Keeping the Balance:
Investigating Covalent Affinity and Passive Permeability of Protease Inhibitors

Dissertation

zur Erlangung des Grades

"Doktor der Naturwissenschaften"

im Promotionsfach Pharmazeutische und Medizinische Chemie

am Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften

der Johannes Gutenberg-Universität Mainz

Collin Zimmer

geb. in Groß-Gerau

Mainz, 2025

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

Dean:

Name of the 1st reviewer:

Name of the 2nd reviewer:

Date of the doctoral examination: 27.02.2025

D77 (Dissertation Universität Mainz)

Declaration of Authorship

I, Collin Zimmer, declare that this thesis titled, “Keeping the Balance: Investigating Covalent Affinity and Passive Permeability of Protease Inhibitors” and the work presented in it are my own.

I confirm that:

• This work was done wholly while in candidature for a research degree at this university.

• Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

• Where I have consulted the published work of others, this is always clearly attributed.

• Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my own work.

• I have acknowledged all main sources of help.

• Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

_____

Usage of artificial intelligence (AI) tools

AI tool	Usage	
Microsoft Word, integrated spelling & grammar examination DeepL Translate	Language support for American English: Correction of single words or sentences Translation of single words or expressions Standard settings	Throughout the entire thesis
Google Scholar Search Algorithm	Literature research Settings: sorted by relevance; no restrictions regarding language and type of work, modular restriction of publication date for specific searches	Throughout the entire thesis
Perplexity Free version	Research for introduction literature to specific topics Knowledge cutoff: Dec-2023	Throughout the introduction part of the thesis

Acknowledgments

Abstract

The era of serendipitously developed covalent drugs has been superseded roughly since the year 2000 by the now rational implementation of covalently reactive warheads. This became especially prominent in the contexts of targeted covalent kinase inhibitors and mechanism-based protease inhibitors. However, there is limited diversity in the approved warheads for this application and many irreversible warheads still face distrust in their clinical applicability—deservedly so if they have high reactivity, or only moderate selectivity towards off-targets. To address this issue and contribute to an improved understanding of covalent protease inhibition, studies on different types of reversible and irreversible warheads attached to a peptide recognition sequence were performed. Rhodesain, the cathepsin L ortholog of *Trypanosoma brucei*, which is a validated drug target of this parasite in the context of Human African Trypanosomiasis, was used to this end as an established model protease. As alternative warheads to irreversible peptide vinylsulfones, reversible α-fluorovinylsulfones and -sulfonates were mechanistically characterized. They elucidated interesting implications of structural changes to residence time with retention of their covalent mode of action. As an extension to the field of dual reactive warheads, keto vinylesters and -sulfones were investigated, which are appreciable for their slowly irreversible reaction towards rhodesain. The last warhead class evaluated are warheads based on nucleophilic aromatic substitutions (S_NAr) which classically react irreversibly with catalytic thiols, but the presented examples expand on reversible options through electrophilicity fine-tuning.

In addition to the investigations on binding mode described above, passive permeability of peptide-based inhibitors was evaluated in the context of a parallel artificial membrane permeation assay (PAMPA). In general, permeability is a critical point of attrition during clinical development, which is the reason why its assessment already in an exploratory phase of drug design merits recognition. Inhibitors targeting different proteases were investigated to obtain information for structurally diverse peptidomimetics. This target-independent characterization yielded valuable information about the negative impact of ionizable groups on passive permeation and clearly identified the necessity for prodrug approaches when carboxylic acid or guanidine groups are present in the peptide backbone. Another study provided a link between in vitro enzyme inhibition and observed effect against isolated *Schistosoma mansoni* lifecycle stages by discussing passive permeability in the context of irreversible inhibitors. For these, it was demonstrated that non-covalent affinity and permeability contribute cooperatively to an effect in a cell experiment, which is valuable information when selecting a chemical probe for cellular target engagement. Lastly, permeability contributions could be quantified for selected reversible warheads as well as non-proteinogenic amino acids. It was shown that nitrile and aldehyde warheads are equivalent in terms of permeability. The same is true for homophenyl alanine (hPhe) compared to *O*-benzyl serine (*O*BnSer). Peptidomimetic options like aryl trifluoroethylamines can effectively contribute to permeability, but at the same time, limitations of lipophilicity to solubility were discussed.

Zusammenfassung

Vor dem Jahr 2000 war ein kovalenter Bindemodus eher ein zufälliges Phänomen der Arzneistoffentwicklung. Seitdem werden kovalent reagierende *Warheads* jedoch rational implementiert. Dies wurde besonders im Zusammenhang von Kinase- und Protease-Inhibitoren prominent, welche jeweils beispielhaft für *targeted-covalent* und *mechanism-based* Inhibitoren sind. Allerdings gibt es nur wenige zugelassene *Warheads* und vor allem solchen mit irreversiblem Mechanismus wird Skepsis bezüglich ihres klinischen Nutzens entgegengebracht, wenn sie eine hohe Reaktivität zeigen oder wenig selektiv gegenüber anderen Zielstrukturen sind. Um dies näher zu beleuchten und dadurch zu einem besseren Verständnis kovalenter Protease-Inhibitoren beizutragen, wurden Studien an peptidischen Erkennungssequenzen mit verschiedenen reversiblen und irreversiblen *Warheads* durchgeführt. Rhodesain, das Cathepsin L-Ortholog in *T. brucei*, welches eine validierte Zielstruktur dieses Parasiten für Arzneistoffe im Kontext der Afrikanischen Trypanosomiasis darstellt, wurde hierbei als Modellprotease genutzt. Als alternative *Warheads* zu irreversiblen Vinylsulfonen wurden reversible α-Fluorvinylsulfone und -sulfonate mechanistisch charakterisiert. Dies verdeutlichte den Einfluss struktureller Änderungen der Inhibitoren auf ihre Verweildauer unter Erhaltung des kovalenten Bindemodus. Als Ergänzung von *Warheads* mit zwei reaktiven Strukturen wurden Ketovinylester und -sulfone mit langsam-irreversiblem Verhalten gegenüber Rhodesain untersucht. Die zuletzt untersuchte *Warhead*-Klasse umfasst solche, die auf einem S_NAr-Mechanismus basieren, welcher klassischerweise in einer irreversiblen Reaktion mit katalytischen Thiolen verläuft. Im vorgestellten Beispiel wurden jedoch durch Modulation der Elektrophilie reversible Optionen entwickelt.

Zusätzlich zu diesen Untersuchungen zum Bindemodus wurde die passive Permeabilität peptidischer Inhibitoren in einem *PAMPA* untersucht. Grundsätzlich ist Permeabilität ein kritischer Parameter der klinischen Entwicklung, weshalb ihre Analyse schon in der explorativen Phase Aufmerksamkeit verdient. Es wurden Inhibitoren untersucht, die verschiedene Proteasen zum Ziel haben, um Informationen strukturell diverser Peptidomimetika zu erhalten. Diese Charakterisierung betonte den negativen Einfluss ionisierbarer Gruppen auf die passive Permeabilität und identifizierte klar die Notwendigkeit von Prodrug-Strategien für Carbonsäuren und Guanidine im Peptidrückgrat. Eine weitere Studie verband in vitro Hemmung einer Protease und beobachteten Effekt gegen kultivierte *S. mansoni* Lebenszyklusformen und diskutierte passive Permeabilität im Kontext irreversibler Inhibitoren. Hierbei wurde eine Kooperativität von nichtkovalenter Affinität und Permeabilität bezüglich eines zellulären Effekts gezeigt. Für die Wahl eines Kandidaten zur Zielstrukturaddressierung im zellulären Kontext stellt dies eine wichtige Information dar. Schlussendlich wurden Beiträge zur Permeabilität verschiedener *Warheads* und nicht-proteinogener Aminosäuren quantifiziert. Es wurde gezeigt, dass Nitril und Aldehyd sowie hPhe und *O*BnSer diesbezüglich jeweils äquivalent sind. Peptidomimetische Alternativen wie Aryltrifluorethylamine können effektiv die Permeabilität steigern; gleichzeitig wurden aber auch ihre Grenzen bezüglich ihres Beitrags zur Löslichkeit diskutiert.

Table of Contents

List of Abbreviations

(e/b)Ro5	(extended / beyond) Rule of 5
(r/g)HAT	(rhodesiense / gambiense) Human African Trypanosomiasis
[I]	Concentration of inhibitor
[S]	Concentration of substrate
5αRed	5α-Reductase
ADMET	Absorption, distribution, metabolism, elimination, toxicity
α-Gluc	α-Glucosidase
AI	Artificial intelligence
AIDS	Acquired immunodeficiency syndrome
Ala	Alanine
AMC	7-Amino-4-methylcoumarin
APOL1	Apolipoprotein L1
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
ATP	Adenosine triphosphate
AU	Arbitrary unit
BCS	Biopharmaceutics Classification System
βR	β-Adrenergic receptors
Caco-2	Cancer coli 2
Cat	Cathepsin
Cbz	Carboxybenzyl
CCL2	C-C motif chemokine ligand 2
CD	Cluster of differentiation
CLIP	Class II-associated invariant chain peptide
CoV	Coronavirus
Cys	Cysteine
DDC	DOPA decarboxylase
dGTP	Desoxyguanosine triphosphate
DNA	Desoxyribonucleic acid
DOPA	L-3,4-Dihydroxyphenyl alanine
DPP	Dipeptidyl peptidase
dUMP	Desoxyuridine monophosphate
EBI	European Bioinformatics Institute
EC	Enzyme commission
ECM	Extracellular matrix
EPR	E prostaglandine receptor
EPR	Enhanced permeability and retention
Eq.	Equation

ER	Estrogen receptor
ER	Endoplasmatic reticulum
ESI	Electrospray ionization
Factor Xa	Coagulation factor Xa
$F_{background}$	Background fluorescence
FPS	Farnesyl pyrophosphate synthase
f_{un}	Fraction unionized
GIT	Gastrointestinal tract
Gln	Glutamine
Glu	Glutamate
GLUT	Glucose transporter
Gly	Glycine
GR	Glucocorticoid receptor
GSH / GS^-	Glutathione (thiol / thiolate)
H2R	Histamine receptor 2
HBVPol	Hepatitis B virus DNA-polymerase
HCVPol	Hepatitis C virus RNA-polymerase
HCVPro	Hepatitis C virus protease
HHVPol	Human herpes virus family DNA polymerase
HIF-1	Hypoxia-induced factor 1
His	Histidine
HIV1Pro	Human immunodeficiency virus 1 protease
HLA	Human leukocyte antigen
HMG-CoA-Red	3-Hydroxy-3-methylglutaryl-coenzyme A reductase
HNE	Human neutrophil elastase
HOMO	Highest occupied molecular orbital
H. sapiens, Hs	*Homo sapiens*
hPhe	Homophenyl alanine
hRFC	Human reduced folate carrier
HSAB	Hard and soft acids and bases
HVPol	Herpes virus DNA-polymerase
IC_{50}	Inhibitor concentration with 50 % reduction in enzyme activity
ICD	Intracellular domain
Ii	Invariant chain
IL	Interleukin
INF	Interferon
ip.	Intraperitoneal
IR	Infrared
IUBMB	International Union of Biochemistry and Molecular Biology
k_{2nd}	Second-order inactivation rate
k_{cat}, k_{chem}, k_{inact}, $k_{1–6}$	Reaction rates

K_i, K_I, K_i*, K_i^{app}	Inhibition constants
K_M	Michaelis-Menten constant
k_{obs}	Rate of change in substrate conversion rate
LAT1	L-type amino acid transporter
LC	Liquid chromatography
Leu	Leucine
LeuT	Leucine transporter
LUMO	Lowest unoccupied molecular orbital
Lys	Lysine
M	Membrane retention
M1, M2	Macrophage polarization states 1 and 2
MALDI	Matrix-assisted laser desorption/ionization
MAOB	Monoamine oxidase B
MCP	Microchannel plate
MCT	Monocarboxylate transporter
MDCK	Madin-Darby kidney cells
Met	Methionine
MHC	Major histocompatibility complex
MMP	Matrix metalloprotease
M^{pro}	Main protease
MR	Mineralocorticoid receptor
MS	Mass spectrometry
MTR	Melatonin receptor
NA	Neuraminidase
NECT	Nifurtimox eflornithin combination therapy
NMR	Nuclear magnetic resonance
*O*BnSer	*O*-Benzyl serine
ODC	Ornithine decarboxylase
OECD	Organisation for Economic Cooperation and Development
PAMPA	Parallel artificial membrane permeation assay
P_{app}, P_a, P_{am}, P_e, P_m, P_o, P_f, P_u, P_{UWL}, P_{ul}	Permeation rates
PDB	Protein Data Bank
PepT	Peptide transporter
Phe	Phenyl alanine
PR	Prostaglandin receptor
PROTAC	Proteolysis-targeting chimeras
Pro	Proline
PSA	Polar surface area
PSA	Prostate-specific membrane antigen
Q	Quadrupole
QM/MM	Quantum mechanics / molecular mechanics

RCSB	Research Collaboratory for Structural Bioinformatics
RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive oxygen species
s	Hill slope
S1PR	Sphingosine-1-phosphate receptor
SARS	Severe acute respiratory syndrome
S. mansoni, Sm	*Schistosoma mansoni*
sc.	Subcutaneous
Ser	Serine
SERD	Selective estrogen receptor degrader
SGLT	Sodium glucose transporter
SLC	Solute carrier
SRA	Serum resistance-associated protein
TCR	T cell receptor
TFA	Trifluoroacetic acid
TGFβ	Transforming growth factor β
TgsGP	*Trypanosoma gambiense*-specific glycoprotein
Thr	Threonine
TMPRSS2	Transmembrane protease serine subtype 2
TNF	Tumor necrosis factor
TOF	Time-of-flight
T_{reg}	Regulatory T cells
T. brucei, Tb	*Trypanosoma brucei*
TS	Thymidylate synthase
Tyr	Tyrosine
UTP	Uridine triphosphate
UV	Ultraviolet
UWL	Unstirred water layer
VAT	Variable antigen type
VEGF	Vascular endothelial growth factor
v_i	Initial substrate conversion rate (velocity)
v_s	Steady-state substrate conversion rate (velocity)
VSG	Variant surface glycoprotein
WHO	World Health Organization
Xaa	Undefined amino acid
XO	Xanthin oxidase

Introduction

Covalent Inhibitors of Proteases Mimicking the Endogenous Substrate Scaffold

Drug Design Strategies: Derivates of Endogenous Ligands

There are several preliminary decisions to make when designing small molecules in the context of disease research. From the start, it needs to be clear in which experimental settings the chemical entity should be employed.

- Is the research focused on mechanistic investigations of an isolated molecular event (ligand-target interaction)?
- Is a multifactorial cellular context supposed to be evaluated (target engagement in cellulo)?
- Is the designed ligand supposed to be evaluated for its effectiveness in the reduction of a real disease burden (clinical study)?

Depending on the answers to these questions, optimizations of different characteristics become relevant. Sufficient affinity is always required, but also selectivity and favorable physicochemical characteristics are important to define molecular interactions. In cellular contexts, permeability and extended chemical stability require assessment and optimization. For drug candidates, aspects of clinical application like metabolic stability and overall bioavailability, but also toxicity are focal points. Since a pharmacodynamic effect is usually the prime criterion, drug discovery campaigns start with the effort to find a high-affinity binder, on which modifications can be implemented. To find this starting molecule, a library of molecules can be screened in silico against a target model (virtual screening), or experimentally against the isolated target (target screening) or the target cells (phenotypic screening). Prior to screening, a library can be filtered for the removal of undesired structures that are, e.g., false-positive hits in many assays (promiscuous), or known causes for toxicity issues. They can also be filtered for desired properties like molecular weight, polar surface area, or hydrogen bond donor and acceptor count (Figure 1).

In this context, virtual screenings usually have the advantage of the highest throughput and access to the most potential ligands, because they are not limited by synthetic constraints. They are especially recognized in early hit identification, lately also supported by artificial intelligence, but also the value of computational support in structure-based drug design to establish and explain structure activity relationships is consensus.[1–3] Experimental target screenings have the advantage of a strong focus due to a low number of interfering variables. They enable a variety of options for manipulation (e.g., target mutation studies to identify crucial moieties, or dependency studies for isolated experimental parameters). Therefore, mechanistic indications can be readily derived, and optimizations can be guided by structure-based decisions. However, target screenings have limited predictability towards behavior in the more complex in cellulo or in vivo environments.[4] In contrast, phenotypic screenings have the advantage of the early inclusion of favorable properties like membrane permeability and information on

an effect in a biologic setting. This comes at the expense of missing information about the multitude of influencing factors in the more complex cellular environment.[5,6] This can obscure reasons for an observed superiority between comparable molecules, but a ligand with favorable characteristics in a cellular environment constitutes the most valuable hit of the three different screening methods.[7] It is obviously reasonable to not only rely on the result of one of these screening methods but to characterize ligands successively in different available assays, and to include specific experiments that assess bottlenecks between them. For example, a promising structure from a virtual screening can exert no experimental effect if it has insufficient aqueous solubility or if its synthesis is impossible to begin with. Similarly, a binder identified in a target-based assay can have poor cellular effect due to limited permeability. These problems can be specifically addressed through structural variations, but only if properly identified.

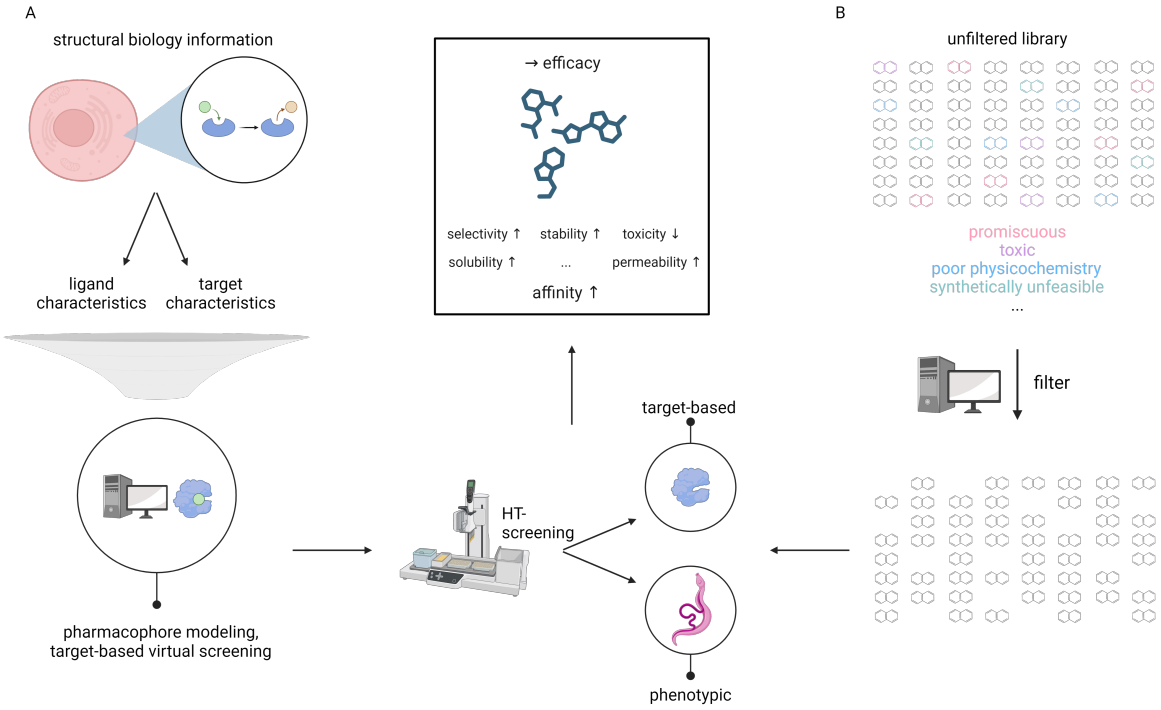


Figure 1: The central goal of drug discovery is to find a chemical entity that has relevant efficacy against a disease, which is strongly driven by affinity for a specific target, but also comprises other parameters (black box). To achieve this, combinations of predictive and experimental methods are used. All these methods result in a small number of ligands with feasible characteristics that are interesting for further drug development. (A) Information on structural biology (crystal structures of target-ligand complexes), from which ligand- and target characteristics can be derived, can be used to optimize predictive methods. Ligand features can be set starting points as pharmacophores, and the target features can be used as an interaction surface for virtual screenings of digital libraries. (B) Libraries of molecules can be filtered for unwanted properties (toxic, promiscuous, poor solubility, etc.). The resulting set of molecules from filtering, together with the selections from virtual screening approaches, has a size that can realistically be handled in experimental settings like screenings on the isolated target or on the target cell type. Figure created with Biorender.com.

A strong similarity between natural and synthetic ligands of a protein target is commonly encountered in approved drugs and in research—an approach that will be rationalized in the following paragraphs. The drugs and experimental probes discussed throughout this thesis conform to this characteristic and can therefore be described as "ligand-based", since they closely resemble the physiological ligand of the target under investigation. It should be considered that this approach is distinct to "ligand-based drug design", which is defined as the design of drug candidates in absence of a crystal structure that visualizes relevant interactions. Instead, experimental affinity characteristics of ligands are used that are transformed into pharmacophore features as scaffolds for further development.[8,9] The difference is that interaction potential in molecules from a ligand-based drug design approach can be exerted by moieties that are structurally different from the natural ligand, and that linking portions between pharmacophores can vastly differ. So, ligand-based drug design yields substances that retain the interaction surface and not necessarily the true structure of the natural ligand. For the reason of avoiding ambiguity, the term "ligand-based" is not used frequently herein, even though it might be a valid descriptor for a set of designed inhibitors.

There are targets that are not sufficiently addressable by any member of a screening library, which can necessitate starting optimization from a known (e.g., the physiologic) ligand. If no library hit is available, this gives a reasonable starting point to structural implementations, which can be expanded upon by a structure-based approach.[10,11] Also, if a ligand characteristic can be identified as essential, it can later be introduced to an existing library hit with already explored binding mode in the effort to optimize it.[12] Starting molecular changes from a physiological ligand with proven affinity and scientifically sound binding mode is a historically successful strategy in drug design. Figure 2 presents a collection of prominent approved drugs that show remarkable congruence with their endogenous counterparts. It does not claim exhaustiveness but highlights targets that are most frequently addressed with molecules designed to mimic their physiological ligand. The large number of structurally and functionally unrelated targets highlights the general applicability of this approach. In this context, a synthetic ligand imitating a substrate, reaction intermediate, or product can be rationally designed from scratch (e.g., protease inhibitors from peptide substrate)[13] or can serendipitously be the optimal scaffold from a screening-based approach (e.g., 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase inhibitors as transition state mimetics)[14].

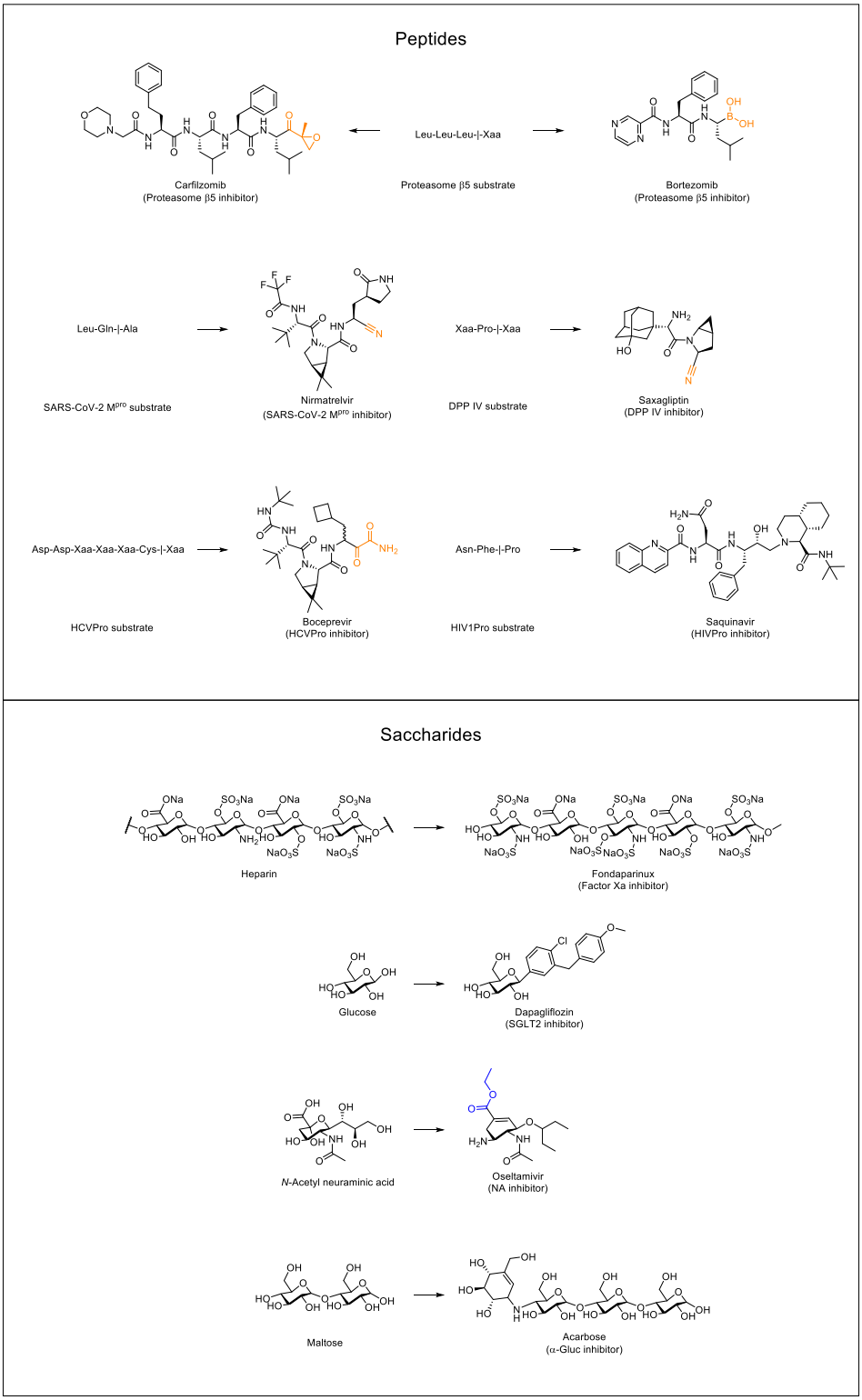


Figure 2A: Collective depiction of several examples of ligand-based drugs (endogenous ligand → drug) grouped by the basic structure they are derived from into peptides, saccharides, steroids, nucleobases, monoamines, and miscellaneous. Structures conveying covalent reactivity are marked orange. Prodrug structures which are either eliminated to liberate the active drug, or drugs that exploit bioconjugation processes are marked blue. Figure created with ChemDraw. -|- = cleavage site, Xaa = undefined amino acid, Leu = leucine, Gln = glutamine, Ala = alanine, Pro = proline, Asp = aspartate, Asn = asparagine, Phe = phenyl alanine, SARS-CoV-2 M^{pro} = severe acute respiratory syndrome coronavirus 2 main protease, DPP IV = dipeptidyl peptidase 4, HCVPro = hepatitis C virus protease, HIV1Pro = human immunodeficiency virus 1 protease, Factor Xa = coagulation factor Xa, SGLT2 = sodium glucose transporter 2, NA = neuraminidase, α-Gluc = α-glucosidase.

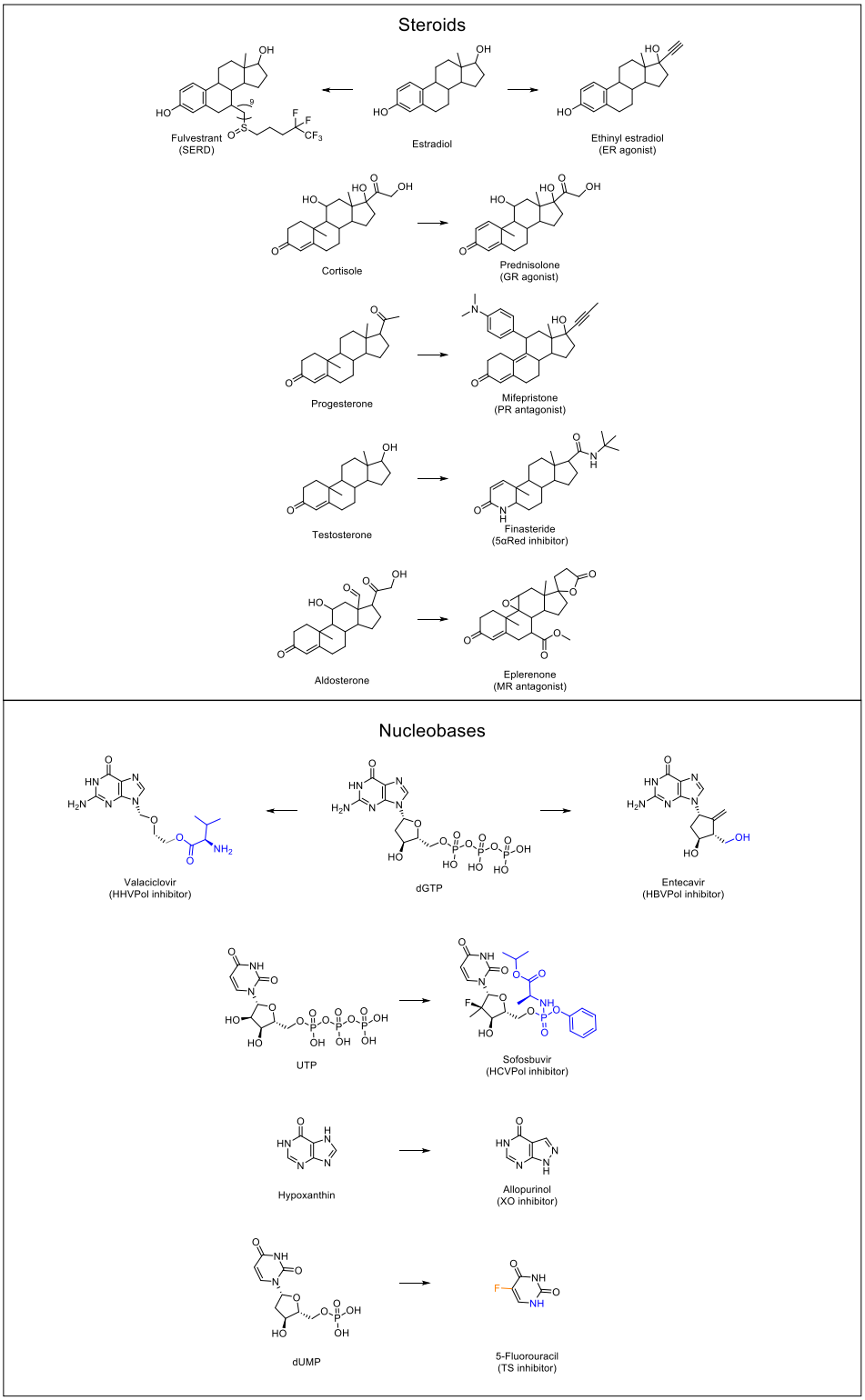


Figure 2B: SERD = selective estrogen receptor degrader, ER = estrogen receptor, GR = glucocorticoid receptor, PR = progesterone receptor, 5αRed = 5α-reductase, MR = mineralocorticoid receptor, HHVPol = human herpes virus family desoxyribonucleic acid (DNA) polymerase, dGTP = desoxyguanosine triphosphate, HBVPol = hepatitis B virus DNA-polymerase, UTP = uridine triphosphate, HCVPol = hepatitis C virus ribonucleic acid (RNA) polymerase, XO = xanthin oxidase, dUMP = desoxyuridine monophosphate, TS = thymidylate synthase.

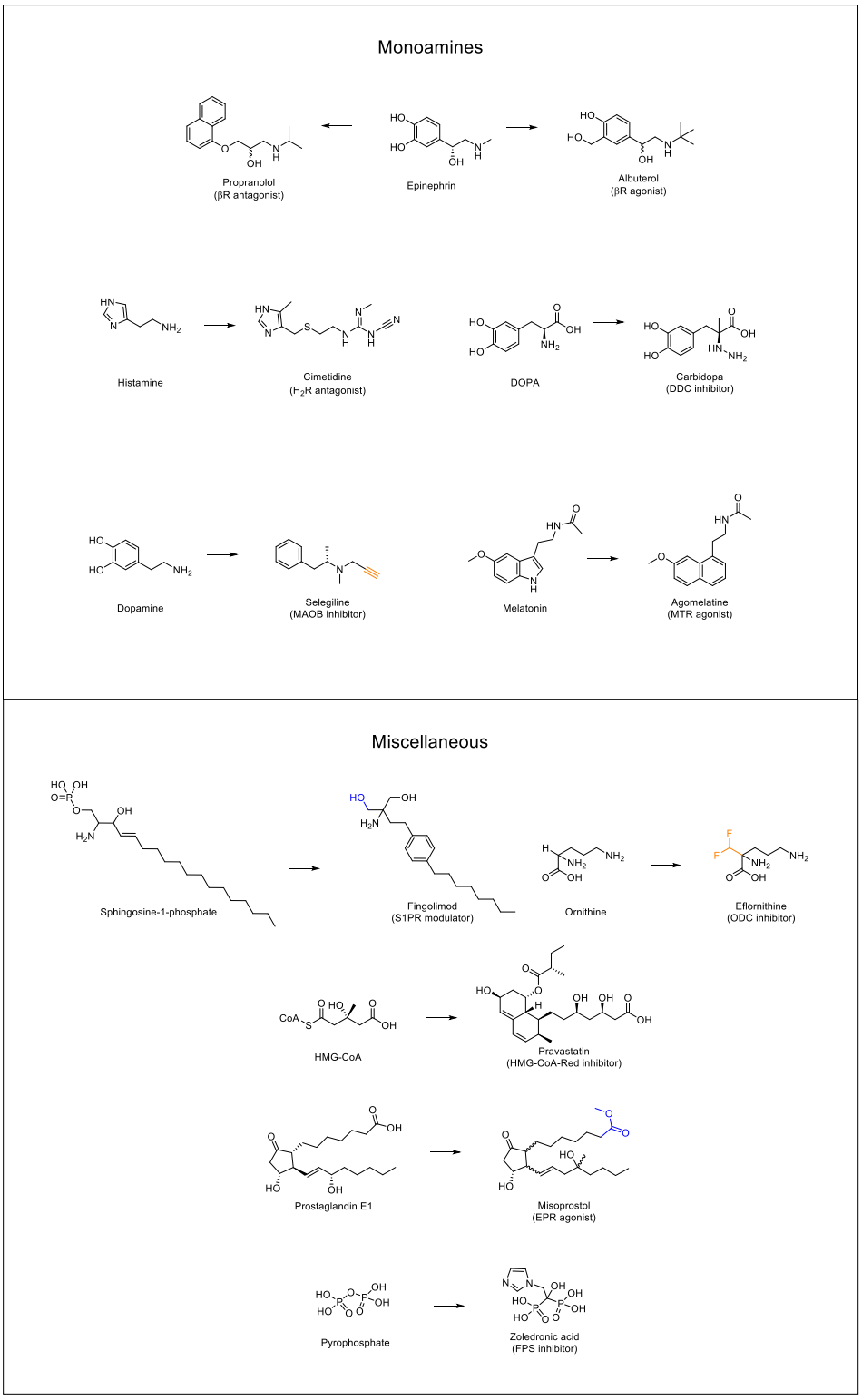


Figure 2C: βR = β-adrenergic receptors, H_2R = histamine receptor 2, DOPA = L-3,4-Dihydroxyphenyl alanine, DDC = DOPA decarboxylase, MAOB = monoamine oxidase B, MTR = melatonin receptor, S1PR = sphingosine-1-phosphate receptor, ODC = ornithine decarboxylase, HMG-CoA-Red = HMG-CoA reductase, EPR = prostaglandin E receptors, FPS = farnesyl pyrophosphate synthase.

A challenge especially for clinical phase I drugs is their pharmacokinetic properties, and membrane permeation as a part of this.[15–17] This is also true for candidates derived from endogenous ligands. Since these ligands are part of intricately regulated physiological functions, they are often not meant to passively distribute throughout the body but are rather confined to a certain environment by physiological membranes. They are transported through specific active mechanisms or never leave the environment in which they are produced. They are often contained in defined intracellular compartments (e.g., secretory vesicles, lysosomes, nucleus, cytosol) or in strictly localized extracellular environments (e.g., synaptic cleft), or are limited to one organ where they exert their function (e.g., vascular system, central nervous system). A property that keeps them in place is their commonly hydrophilic nature, which reduces passive permeation due to the largely lipophilic nature of biological membranes. The general hydrophilicity of endogenous ligands is especially exemplified by the sugar moiety and charge of nucleotides (UTP) or some oligosaccharides (heparin), as well as in the collection of polar amide bonds in oligopeptides, or the zwitterionic nature of amino acids (DOPA, ornithine).

The aim for many novel drugs is to be effective after oral administration, which requires passing several different environments and membranes.[18,19] This makes active and passive permeability, but also chemical and metabolic stability, and plasma protein binding relevant in drug design that require optimization in parallel to an improvement in affinity. To a certain extent, this should also be taken into consideration for research probes for mechanistic investigations to exert an effect in cell-based assays. For hydrophilic ligand-mimetics, this likely entails an initial bias for poor passive membrane permeation. Some workarounds for especially hydrophilic ligand-structures are routinely implemented in drug design, for instance:

- the inclusion of suitable lipophilic groups as an extension of the molecule or a replacement of hydrophilic structures to modulate the overall polarity as in dapagliflozin and nirmatrelvir,
- the chemical, transient masking of hydrophilic structures with cleavable, lipophilic moieties as in the prodrugs oseltamivir and sofosbuvir,
- the exploitation of physiological conjugation mechanisms as in the prodrugs 5-fluorouracil, fingolimod, and entecavir,
- or the reliance on active uptake as in eflornithine and valaciclovir.

Apart from these chemical approaches that change the drug's structure, technological approaches like nanoparticle formulations can be used to override pharmacokinetic properties.[20] Another alternative is to opt for a parenteral application to bypass the gastrointestinal system altogether (e.g., fondaparinux, carfilzomib). All the abovementioned examples are depicted in Figure 2, and those for which the prodrug approach is relevant are highlighted.

It should be noted that there are some endogenous ligands, for which passive transport is possible and relevant to their physiological effect (e.g., steroids). Some of these can therefore be used unchanged as drugs if formulated parenterally to avoid presystemic metabolism (e.g., testosterone, estradiol).[21,22]

Monoamine neurotransmitters like norepinephrine are mainly transported actively but can have relevant passive diffusion over membranes at high concentrations.[23–25] This explains how in some cases permeability does not need to be addressed at all, if sufficiently lipophilic modifications are introduced to optimize binding affinity (e.g., propranolol). An especially interesting field for inhibitors with a scaffold closely resembling a substrate are protease inhibitors. The peptide scaffold enables vast combinatorial options utilizing the 20 proteinogenic amino acids, as well as non-proteinogenic amino acids than can be simple homologs (e.g., phenyl glycine, *tert*-leucine) or remote derivates (e.g., dimethylthioproline, allophenylnorstatine) of the proteinogenic ones.[26] Apart from modifications that are implemented with improved affinity in mind, modifying stereochemistry (D-amino acids), cyclization, or amide *N*-methylation can convey metabolic stability to peptides.[27] A characteristic feature of these protease inhibitors is also the frequent inclusion of electrophilic warheads to react with the nucleophilic active sites of their target enzyme. This derivatization will be discussed in further detail below, and examples are marked in Figure 2A. A more detailed description of some optimization campaigns starting from an endogenous ligand is given below for steroid-, saccharide-, and peptide-derived drugs. A focus is being placed on available information on the optimized parameters (especially affinity, selectivity, permeability, and metabolic stability), which together produce their overall effect.

As exemplified for some steroid drugs in Figure 2B, the derivatization from endogenous ligands has been practiced in ligand optimization efforts for all classes of steroid hormones (glucocorticoids, mineralocorticoids, estrogens, gestagens, androgens). In the case of estrogens, the addition of one ethinyl substituent to estradiol already provided the staple ER agonist in modern clinical practice (i.e., ethinylestradiol in combined oral contraceptives). The introduction of this group was a German invention of the 1930s in an effort to produce orally active estrogens and gestagens.[28] Other approaches to generate orally bioavailable ER agonists were often based on estradiol itself and included a technological optimization (micronization)[29] boosting absorption, or transient chemical derivatization, e.g., to its phenol sulfamate prodrug to adjust metabolism[30]. Leaning away from the steroid skeleton but keeping an estradiol-like conformation led to alkene-rigidized derivates like tamoxifene[31] (or later ospemifene)[32], or heterocyclic drugs like raloxifene[33] (or later bazedoxifene)[34], which are all orally bioavailable selective ER modulators (SERMs) with different intrinsic activities when interacting with ERα or ERβ. They are used for treating dyspareunia[35], breast cancer[36], or osteoporosis[37,38]. Another level of ER interaction-tuning is achieved by targeted degradation with SERDs like the authorized oncologic steroid fulvestrant[39,40] which exerts its effect through competitive antagonism and accelerated receptor degradation. In the same direction go ER proteolysis-targeting chimeras (PROTACs) that link ERs to ubiquitin ligases in the context of proteasomal degradation.[41] All of these examples still closely resemble the estradiol scaffold.

Gliflozins are antidiabetics approved for type II diabetes that exert their effect by inhibiting SGLT2. This transporter is responsible for glucose reabsorption in the kidneys. Hence, the endogenous ligand is glucose but the inhibitors only in half resemble it (Figure 2A). This is because they are technically based

on the chalcone glucoside phlorozin, isolated from apple tree bark, rather than on glucose alone.[42] Initially, this natural product was recognized for its provocation of glucosuria after high oral doses. Later, the SGLTs were identified as targets for its influence on physiological glucose levels.[43,44] Glucose alone has a low affinity to SGLT2, a property greatly enhanced by the aglycone portion.[45] In contrast, however, the glucose portion is also essential for efficient binding, both for the natural product and designed SGLT2 inhibitors like dapagliflozin.[46] The optimization strategies of SGLT2 inhibitors revolved around the relevance of selectivity (mainly towards SGLT1)[47], and in vivo metabolic stability (against glucosidases)[48–50]. The latter weakness is the major reason for a unifying structural element of the SGLT2 inhibitors approved by the European Medicines Agency (EMA) and the US-American Food and Drug Administration (FDA): they are all *C*-glycosides instead of the original *O*-glycosides. This changes the synthetic route quite a bit, but chemically it makes them ethers as opposed to acetals, rendering them stable against glycosidases, which increases their total metabolic stability.[51–53] Some variations of the sugar moiety are allowed. While most are still based on glucose, xylose is possible (Sotagliflozin).[54] Still, most variations are made at the aglycone part of the molecule, with the natural ligand portion (β-D-glucose) mostly unchanged except for the *C*-glycoside modification.

Initial research on the HCV protease was performed as a phenotypic high-throughput screening of different libraries, which did not yield any promising candidate. This was attributed to the active site being “shallow, featureless and highly solvent-exposed”.[13] Therefore, the search was restarted from a known substrate, producing a terminally capped undecapeptide with one central non-proteinogenic amino acid right before the cleavage site.[55–57] It carried a norvaline side-chain and an electrophilic α-keto amide moiety, with which the enzyme was covalent-reversibly inhibited. In this stage, despite its favorable affinity, the detrimental pharmacokinetic properties were recognized (both permeability and metabolic stability), and optimization was necessary. The enzyme favors hydrophobic residues in general, so all hydrophilic amino acids (Asp, serine (Ser), methionine (Met), tyrosine (Tyr)) were either cut off entirely or replaced by more lipophilic structures, resulting in a drastically shortened pentapeptide. The included non-proteinogenic amino acids provided metabolic stability in addition to increasing lipophilicity, which yielded an intermediate inhibitor that already showed good bioavailability from subcutaneous application but was not yet orally bioavailable.[57] The molecule was further reduced in size to a tripeptide, and the remaining non-proteinogenic amino acids were cyclobutyl alanine, a fused Pro/Leu mimetic, and *tert*-leucine. The keto amide was kept constant but was now placed at the peptide *C*-terminus, the *N*-terminus was capped by a *tert*-butyl urea. With these modifications, boceprevir, a high-affinity inhibitor with good oral bioavailability and effect in different animal models could be obtained, that also showed far improved selectivity over the related human neutrophil elastase (HNE).[57]

Protease Classification and Cleavage Mechanism

The main targets of the research described in this thesis are enzymes called “peptidases” or “proteases”; both terms used synonymously in literature.[58,59] “Peptidase” defines their function of cleaving a peptide (amide) bond in analogy to “glycosidases” cleaving glycosidic (acetal) bonds. “Protease” describes the cleaved protein targets in analogy to “nucleases” cleaving nucleic acids. A distinction between “protease” and “peptidase” is generally unnecessary but might be made in the case of an enzyme like oxytocinase.[60] It catalyzes the degradation of small nonapeptide hormones like oxytocin, which are not classified as proteins due to their low molecular weight.[61] Hence, oxytocinase would rather be an oligopeptidase than a protease. Conversely, all proteases are peptidases by function. The most prominent classification systems (described below) employ “peptidase” for optimal coverage of their listed enzymes. While in the medicinal chemistry literature, the term “protease” is encountered more frequently, “peptidase” is used if the function of the protease of interest is important, e.g., when the position of the cleaved bond in a peptide chain is specified (in the case of endo- and exopeptidases, carboxy- and aminopeptidases, or di- and tripeptidyl peptidases).

The realm of proteases comprises a set of diverse enzymes essential to all forms of life. They fulfill tasks ranging from catabolism of nutritional peptides to specific regulatory functions in cellular signaling pathways.[62] Since the 20th century, when methods for protein and gene sequencing became available, their phylogenic relationships have been investigated, and structured systems for classification evolved to handle the ever-increasing number of discovered proteases.[63–65] While analogous in their function to catalyze the hydrolysis of peptide bonds, the mostly separate evolution of the protease types clustered by their catalytic machinery has led to structurally unrelated classes, rarely with readily discernible homologies.[66,67] These protease classes are Asp-, cysteine- (Cys), metallo-, Ser-, and threonine (Thr) proteases, named after the structure (amino acid or metal ion) most crucial to the catalytic mechanism (Figure 3A), with glutamate (Glu) proteases and Asn lyases as marginal additions.[68–71] Some examples with (patho-)physiological relevance are listed in Figure 3B & 3C for each of the major protease classes.

Different organizations aim to compile an up-to-date overview of the proteases available to literature. One example database that curates data on all enzyme classes is the “enzyme commission (EC-) system” maintained by the International Union of Biochemistry and Molecular Biology (IUBMB, London, United Kingdom).[72] Another that only focuses on peptidases and links protease families from different organisms, their substrates, and inhibitors is the “MEROPS database” (European Bioinformatics Institute, EBI, Cambridge/Hinxton, United Kingdom).[73] For structural information on proteins, the “Protein Data Bank” (= PDB, Research Collaboratory for Structural Bioinformatics, RCSB, USA) represents the largest publicly available archive of experimental 3D structures.[74] In this context, impressive advances have been made using artificial intelligence (AI) in recent years (Nobel Prize in Chemistry, 2024) to predict structural information for proteins yet inaccessible to experimental elucidation efforts, with Alphafold (Deepmind, London, United Kingdom) as a prime resource.[75–77]

The absolute total number of proteases in each classification system varies greatly (ca. 400 in the EC-system, to ca. 4000 in the MEROPS database), and the same is true for the distribution per catalytic class. To still give a numerical indication, animals express around 500 different proteases, lower eucaryotic organisms like *Saccharomyces cerevisiae* around 100, and bacteria usually below 100.[59] An indication of class distribution can be inferred from fully described proteases in *Homo sapiens* (close to 600, The Mammalian Degradome Database[78,79], compare Figure 3A), which is in line with other literature.[80] Inside these classes, structural similarity can be inferred, forming the basis for clans, families, and subfamilies in ascending degree of homology (compare Figure 3C). Evolutionary sequence conservation inside a family can clearly be followed for orthologous proteases in different organisms, or for paralogous proteases in one organism (compare Figure 4).[81,82] It indicates that cathepsins from different classes (*Hs*CatA/D/L) are completely different proteins, despite their shared function of lysosomal peptide hydrolysis (as little sequence similarity between each other as to the unrelated GTP hydrolase *Hs*KRAS, similarity <25 %). It also outlines genetic relationships of the human enzymes with parasitic orthologs from *T. brucei* and *S. mansoni* (*Hs*/*Tb*CatL and *Hs*/*Sm*CatB1 similarity ca. 60–75 %) and between human paralogs (*Hs*CatB/L/S, similarity ca. 40–75 %).

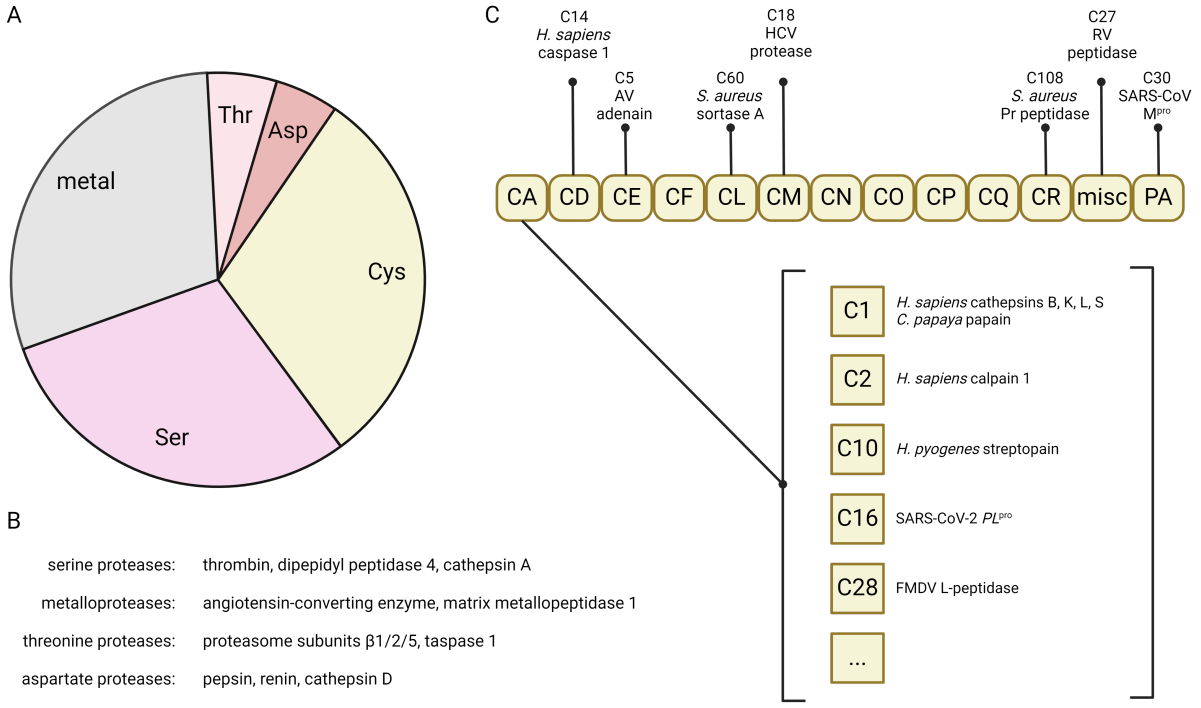


Figure 3: (A) Depiction of protease classes by their proportional distribution within the group of human proteases as compiled in the Mammalian Degradome Database. (B) Examples for human proteases from each main family except cysteine as listed by the MEROPS database. (C) For cysteine proteases, hierarchical organization as clans (rounded boxes) and families (boxes) as listed by the MEROPS database. Apart from the 12 clans consisting only of cysteine proteases, some clans like PA contain proteases from different classes. For some clans, cysteine proteases relevant for human (patho-)physiology are listed. AV = adenovirus, HCV = hepatitis C virus, Pr = phage-related ribosomal, RV = rhinovirus, PL^{pro} = papain-like protease, FMDV = foot-and-mouth disease virus. Figure created with Biorender.com.

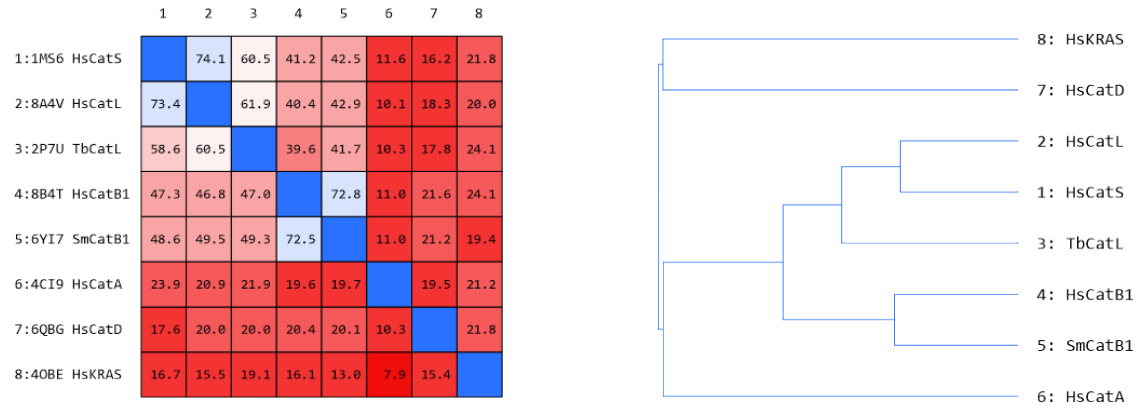


Figure 4: Comparison of overall sequence similarity between cathepsins of different classes related by function (cathepsins A/D/L are lysosomal Ser/Asp/Cys proteases, respectively), and ortho- and paralogs of select Cys proteases (*Hs*/*Tb*CatL, *Hs*/*Sm*CatB1, *Hs*CatB/L/S), and an unrelated human GTP hydrolase (*Hs*KRAS). PDB accession codes are depicted, longest deposited primary sequences of the biological assemblies were compared with the Molecular Operating Environment software (MOE 2022.02, Chemical Computing Group). The value in each square of the similarity matrix (left) is defined as the number of similar residues (BLOSUM62 substitution score >0) between the protein sequences at each intersection divided by the length of the sequence labeled at each column, given in percent. The similarity tree (right) visualizes these numerical implications by connecting closer relationships through shorter paths.

The active site of proteases, to which the hydrolysable peptide substrates bind, can be subdivided into specific functions. Most amino acids forming the active site provide interaction surface for either hydrophilic or hydrophobic interactions, and they sterically restrict the space available for potential ligands. This provides a means of identification for a preferentially cleaved peptide substrate sequence. The formed protease subpockets are usually labeled S, while the substrate amino acids are labeled P, both with numbers corresponding to their relative position to the cleavage site according to the Schechter-Berger nomenclature.[83] Protease active sites are further characterized by a small group of amino acids that contribute “actively” to the catalytic hydrolysis process. In Asp- and metalloproteases, acid/base catalysis is provided for a nucleophilic immobilized water molecule.[84,85] In contrast, the Cys-, Ser-, and Thr proteases convey covalent catalysis through their namesake amino acid residue as a nucleophile.[86,87] In these catalytic assemblies for covalent catalysis (commonly dyads or triads), two roles additional to the nucleophile can be distinguished. There is always a base involved as the second characteristic moiety, and usually a third amino acid that modulates positioning and/or basicity of the others. Together, a microenvironment is established, in which the nucleophilic amino acid that defines the protease class (Cys, Ser, or Thr) gains a strongly polarized side-chain chalcogen-hydrogen bond, which amplifies its nucleophilicity. The catalytic assemblies are highly specific and often characteristic for certain clans of the respective protease classes.[86,87] For example, a Cys→Ser mutation of the catalytic cysteine in bleomycin hydrolase (clan CA, family C1) completely abolished its physiological function.[88] The same is true for Ser→Cys mutations in of the catalytic serine in trypsin (clan PA, family S1) or other classes of hydrolytic enzymes.[66]

Characteristic catalytic mechanisms are depicted for the human proteases CatS, urokinase-type plasminogen activator (uPA), and proteasome β5 subunit in Figure 5. In all three cases, the nucleophile attacks a defined amide group in a peptide to initially form a tetrahedral anionic intermediate (oxyanion). The presence of hydrogen bond donors to form a stabilizing "oxyanion hole" is another modular feature of protease active sites as described below. With the elimination of the amine part of the former amide, an intermediate (thio)ester is formed, in which the remaining part of the substrate is still covalently attached to the enzyme. This (thio)ester is then hydrolyzed through a recruited water molecule, resulting in the restoration of the enzyme after liberation of the carboxylic acid portion of the former amide.

In the case of Cys proteases, Cys acts as the designated nucleophile. Its reactive thiol group is affected by a neighboring histidine (His) side-chain that acts as a proton acceptor for the reaction. The imidazole orientation and basicity, in turn, is regulated by a third hydrogen bonding partner, e.g., an Asn side-chain for clan CA proteases.[86,89] The pK_a values of detached Cys and His in solution are 9 (S-H deprotonation) and 6 (imidazole N^+-H deprotonation).[90] In the active site, however, the pK_a values change to 3 and 8, respectively, indicating the increase in acidity for Cys and in basicity for His.[91] This is the reason why for many cysteine proteases, a Cys^-/His^+ pair is present already prior to catalysis.[86] The intermediate oxyanion can be stabilized by the backbone NH of the catalytic cysteine and a Gln side-chain amide (compare CatS in Figure 5).[86]

A similar general principle is adopted by Ser proteases, in which Ser/His/Asp represents the classical triad. Also here, different architectures for the catalytically involved amino acids are possible for the different Ser proteases.[87,92] In analogy to Asn in the abovementioned Cys proteases, Asp polarizes His, which in turn accepts a proton during catalysis. The major difference is the much lower acidity of Ser (pK_a = 13.6 in solution)[93] in comparison to Cys. This requires a stronger polarization of His than is exerted by the corresponding Asn in Cys proteases, which is provided by the negatively charged Asp.[86] This also explains the higher basicity of the His in the active site (pK_a = 10–12)[94] in comparison to His in a Cys protease triad. However, there is no precatalytical ion pair in Ser proteases. Instead, the serine-β-OH proton is liberated during the nucleophilic attack to the substrate. The intermediate oxyanion can be stabilized by the backbone NHs of the catalytic Ser and a neighboring glycine (Gly) (compare uPA in Figure 5).[87]

For Thr proteases (most prominently the β-subunits 1, 2, and 5 of the proteasome), the principle described above is applied as well but differs in some details. This is mainly because an *N*-terminal Thr side-chain is involved as the nucleophile.[87,95] Of its two possible nucleophilic moieties (α-NH_2 and β-OH), the hydroxyl group in its neutral state is unequivocally the nucleophilic site for catalysis.[96] In its vicinity, there is both the mentioned α-amino group and the side-chain amine of a lysine (Lys) as potential bases for the mechanism. Different mechanisms have been postulated assuming both amines in their deprotonated form[97,98], or one of the two protonated[97,99], discussing whether the Lys or Thr itself functions as the proton acceptor. The most convincing case is made for the Lys already protonated prior to catalysis, as it is expected to be the more basic amine of the two, and the neighboring Asp can also

effectively stabilize the resulting ammonium ion in a salt bridge. This positive charge proximal to the Thr α-amino group can disfavor its spontaneous protonation by water, "preserving" the *N*-terminal basicity for the proton that is eliminated during catalysis.[99] This is especially underlined, because the Thr amino group can partake as a second nucleophile in the reaction for inhibitors with two warheads.[96] For this to be possible, a relevant fraction must be present in its unprotonated state. The oxyanion stabilization is carried out by a single Gly backbone-NH (compare proteasome β5 subunit in Figure 5).[87]

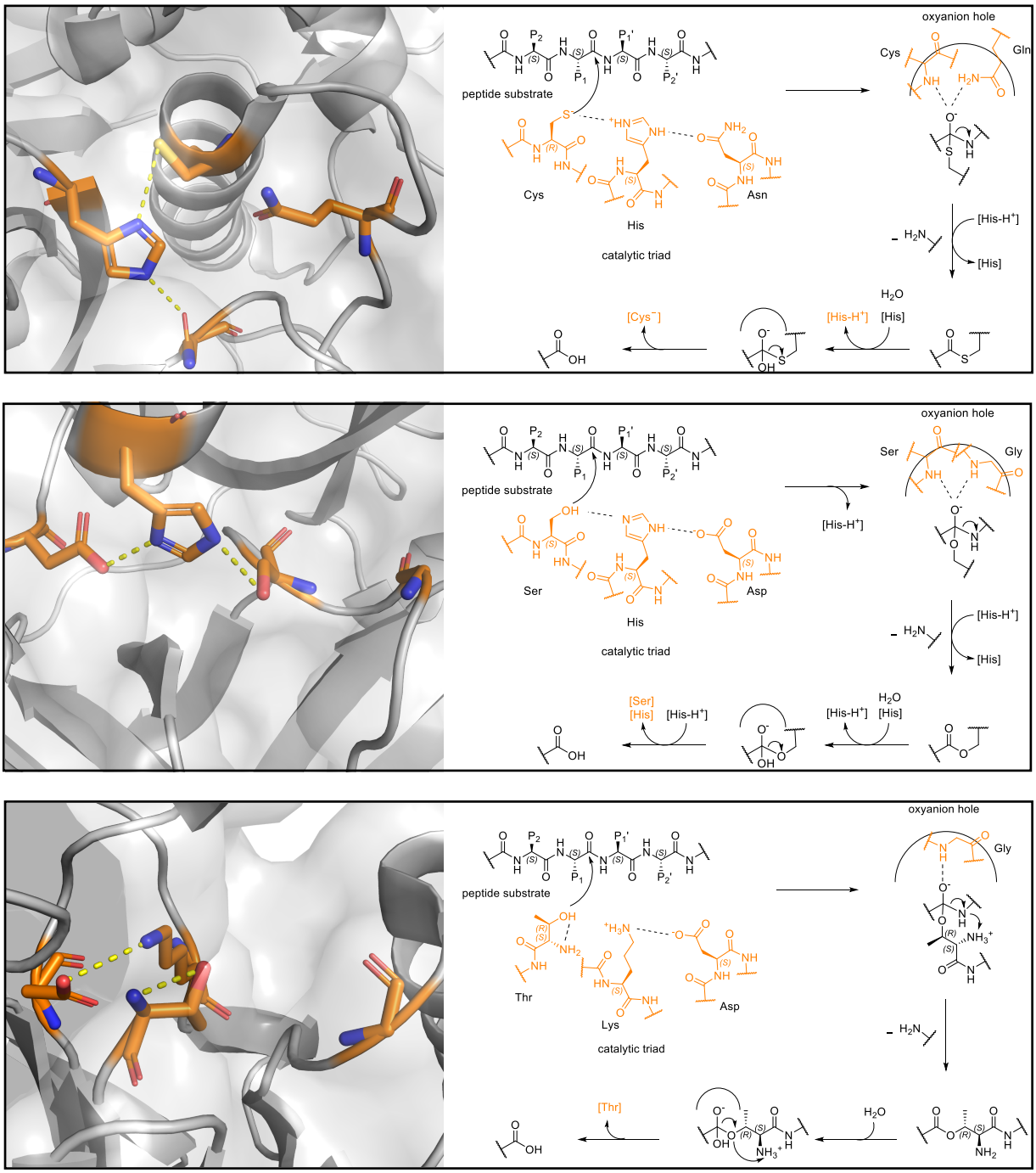


Figure 5: Active site and catalytic mechanisms of the protease classes with covalent catalysis. Example proteases: *Hs*CatS (1ms6, top), *Hs*uPA (1w10[100], center), *Hs*proteasomeβ5 (6rgq[101], bottom). Left panels: surface map and tertiary structure of the active site (grey) with catalytic triad and amino acids involved in oxyanion hole formation (orange). Nucleophilicity-amplifying hydrogen bonding depicted as dashed lines (yellow). Figures created with Pymol (Schrödinger, LLC) and ChemDraw. Right panels: two-dimensional reaction scheme of the catalytic mechanism. Amino acids involved in the catalysis are depicted in their precatalytic protonation state (orange) and relevant hydrogen bonds (black, dashed) are denoted.

The publications included in this thesis have a strong focus on Cys proteases. Investigated targets of this class were *T. brucei* CatL, *H. sapiens* CatB, CatL, and CatS, as well as *S. mansoni* CatB1. Apart from this, human Ser proteases like the transmembrane protease serine subtype 2 (TMPRSS2), uPA, and relevant related proteases (matriptase, tissue plasminogen activator, thrombin, and coagulation factor Xa) were researched. To consciously limit the scope of the introduction to this thesis, information on (patho-)physiological involvement and current research landscape is only detailed for the two main targets, namely trypanosomal CatL and human CatS. For the other investigated proteases, this information is already included in the publications constituting this thesis. Complemented is this, especially regarding own contributions, in each introductory paragraph prior to the respective publication. Examples for this extended information are the implications for drug design drawn from a requirement of an Arg-mimetic structure in TMPRSS2- and uPA inhibitors or from cell membrane properties of *S. mansoni* in the context of the performed experiments.

Electrophilic Warheads on Protease Inhibitors in Clinical and Research Settings

Inhibiting proteases as pathophysiological targets is an important pillar of modern clinical disease management. Inhibitors of renin and angiotensin-converting enzyme (ACE) are options for essential hypertension.[102,103] Anticoagulating drugs against coagulation factors IIa and Xa are in use to prevent cerebral and myocardial infarctions.[104,105] Inhibitors of DPP IV are used for treating type II diabetes.[106,107] Multiple myeloma is a type of bone-marrow cancer treated by inhibition of proteasome subunit β5.[108,109] Prominently, protease inhibitors are in wide clinical use against infectious diseases related to *Hepacivirus hominis*[110] (formerly and in this thesis: HCV), *Lentivirus humimdef1*[111] (formerly and in this thesis: HIV1), and *Betacoronavirus pandemicum*[112] (formerly and in this thesis: SARS-CoV 2).[113–116] This brief list illustrates the utility and value of protease inhibitors in drug design and clinical practice, but it is only the “tip of the iceberg” approved by the drug regulation agencies. Protease inhibitors are also employed in industrial and academic research against a plethora of pathologies like different autoimmune diseases (*Hs*CatS)[117,118], Morbus Alzheimer (*Hs*CatB)[119], osteoporosis (*Hs*CatK)[120], and parasitic diseases like schistosomiasis (*Sm*CatB1)[121] and trypanosomiasis (*Tb*CatL)[122].

As discussed before, the strategy of adding a reactive electrophilic structure to a ligand-inspired scaffold has been a successful strategy for creating effective protease inhibitors, as depicted in Figure 2 for carfilzomib (α-keto epoxide), bortezomib (boronic acid), nirmatrelvir and saxagliptin (nitrile), and boceprevir (α-keto amide). These so-called “warheads” give rise to mechanistic distinctions regarding the kinetics and thermodynamics of target engagement that can have clinical benefits. Still, reactive structures on inhibitors were long frowned upon. They raised concerns of unselective reactivity leading to side-effects and immunological toxicity in the context of haptenization.[123,124] Especially for irreversibly covalent inhibitors and those with high reactivity, these concerns are realistic and need attention.[125] However, since the 1990s, the benefits of covalent interactions are increasingly recognized, which led to a trend towards the rational implementation of covalent drug discovery.[124,126,127] They entail

the option of targeting formerly "undruggable" targets with shallow active sites that are not sufficiently addressable with only non-covalent interactions.[13] The covalent bond formation can result in an increased residence time, which is commonly associated with an improved effect in complex environments, e.g., through kinetic selectivity.[128] Another striking advantage is the decoupling of pharmacokinetics from pharmacodynamics, meaning that an effect is still detectable even if unbound drug is long cleared from circulation, making it possible to reduce dosing strength and frequency, and in turn the metabolic burden.[129]

A ligand's electrophilic warhead reacts with the protease's nucleophilic amino acid. Generally, chemical reactions between electrophiles and nucleophiles are governed by the reactants' chemical potential (μ) and hardness (η), which together determine reactivity. These two values can be derived separately from computations of orbital energies (highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies) and combined to yield a nucleophile's nucleophilic index (ω^-), which defines the numerical propensity for a reaction between a certain nucleophile-electrophile pair.[130–132] To computationally determine electron distribution and reactivity with model or complex nucleophiles, modern techniques use quantum mechanics (QM) or their combination with molecular mechanics (QM/MM), respectively.[133,134] This concept can be applied to biologically relevant nucleophiles like amino acids and relevant electrophile classes.[135,136] However, some contextual differences have to be kept in mind, that render these calculations only estimates. For instance, the polarized serine hydroxyl in a Ser protease's active site is expected to be a transient case between isolated serine-β-OH and serine-β-O^-. Also, the calculation of gas-phase electron distributions of isolated molecules can deviate from the experimental distribution in an enzyme-associated state of the reaction pair.[137]

Classifications for relevant electrophiles and nucleophiles as hard or soft are included in Table 1. High (partial) charge and small distribution volume make up high charge density which characterizes a hard reactant. In contrast, low (partial) charge and large charge distribution volume constitute a soft reactant. Derived from the "hard and soft acids and bases" (HSAB) theory, hard favors reactions with hard and vice versa.[130] Following this classification, electrophiles with localized positive partial charge are considered hard (e.g., aldehydes[138] > nitriles[139]) while those with delocalized partial charge are soft (e.g., Michael acceptors[138] > phosphate esters[140]). In the same spirit, nucleophiles with dense electron distribution are hard (e.g., alcohol(ate)s, amines) and those with diffuse, polarizable electron density are soft (e.g., thiol(ate)s).[135] Hard boron-based warheads follow the HSAB principle and are not efficiently addressed by soft thiols[141], but do react with hard threonine proteases (bortezomib, proteasome)[142] and serine proteases (peptide boronate, dengue virus protease)[143], or other enzymes with hard hydroxyl nucleophiles (taniborbactam, *Klebsiella pneumoniae* carbapenemase 2)[144]. While this classification gives hints to preferred reactions, it does not strictly prohibit cross-reactivity. Aldehydes and nitriles can be very efficient warheads for both serine and cysteine proteases, despite the hard/soft discrepancy. Examples for this are diverse carbonyl-based inhibitors for viral M^{pro} (Cys protease)[145], aldehydes for

human caspase-1 (Cys protease)[146], as well as nitriles for human DPP IV (Ser protease)[147]. Taken together, the HSAB principle is a reason why research on serine proteases often includes carbonyl-based warheads[57,100,148], while variants of Michael acceptors are common in cysteine protease research[149,150].

Table 1: Classification of amino-acid nucleophiles and select electrophiles employed as warheads according to the HSAB principle.

	Soft	**Hard**
Nucleophiles[135]	Cys	Ser, Lys, His
Electrophiles	Michael acceptors[125,132], haloarenes[125,151], nitriles[139]	Ketones[125,132], aldehydes[138], epoxides[132,152], boronic acids[142]

Alternatively to the calculations mentioned before, the propensity of an electrophile to undergo covalent reactions with a specific nucleophile can be estimated experimentally. A simplified approach is to correlate a carbon's signal in a ^{13}C nuclear magnetic resonance (NMR) spectrum with its electrophilicity, as deshielding of an atom's nucleus in its essence is understood as the partial removal of electron density.[153,154] However, to judge the extent of a warhead's reactivity and to assess potential toxicological implications, it is more reliable to use the reaction with small organic nucleophiles. A frequently chosen nucleophile in this context is the tripeptide thiol glutathione (GSH, γGlu-Cys-Gly), because it is an endogenous molecule used for the detoxification of electrophiles.[155] The physiological relevance of kinetic data acquired like this is limited by the fact that in vivo, GSH is only effectively nucleophilic in presence of GSH-*S*-transferase. It is an enzyme that catalyzes the addition of GSH to xenobiotic electrophiles by forming immobilized GS^- to increase its nucleophilicity.[156] In solution also, only the thiolate and not the thiol is relevantly nucleophilic, so the concentration of reactive nucleophile in an assay is defined by the Hendersen-Hasselbalch equation in dependence of the pH value ($pK_{a3}(GSH) = 9$).[157,158] This needs to be considered, when directly comparing reactivity scales from different sources, and is an argument to use this data as indication and ranking rather than as absolute kinetic values of expected reactivity. This pH-dependency of nucleophilicity is also nicely illustrated in the nucleophilicity index ω^- presented before. Depending on the source, Cys-β-S^- is 2–13 times more nucleophilic in comparison to Cys-β-SH, indicating that deprotonation is essential to its reactivity.[135,136] Besides GSH, other possible model thiols include thioglycolic acid and thiophenol[159], cysteine and *N*-acetylcysteine[136], as well as mercaptopropionate[125] and 5-thio-2-nitrobenzoic acid[160]. However, there is a wide variety of model nucleophiles in general.[125] Fortunately, while absolute reactivities are difficult to compare between assays, relative electrophile reactivities are mostly independent of the employed

nucleophile (for the same reaction mechanism), which makes qualitative rankings of warhead reactivity comparable.[125,131,161]

There are many reviews available on the different warheads applied in research not limited to protease targets.[126,162–164] In clinical practice of protease inhibition, only three warheads are currently approved by the EMA[165] and the FDA[166]: nitriles (e.g., nirmatrelvir and saxagliptin), boronic acids (e.g., bortezomib), and epoxyketones (e.g., carfilzomib). Until their discontinuation as corporate decisions due to a change in HCV treatment practice[167,168], e.g., in Europe, USA, Canada[169], and Brazil[170], α-keto amides (boceprevir and telaprevir) were also marketed in the western hemisphere and are still approved, e.g., in Japan[171] (telaprevir). For other diseases, other electrophilic groups can be found in covalent drugs. Some, like the acrylamide in afatinib[172], or the aldehyde in voxelotor[173] were rationally implemented, but there are many with a more serendipitously covalent mechanism like acetylsalicylic acid or penicillin.[174] Example warheads from research campaigns, mostly with a rationally designed covalent mechanism against enzymes with a Ser- or Cys-nucleophile are depicted in Figure 6.

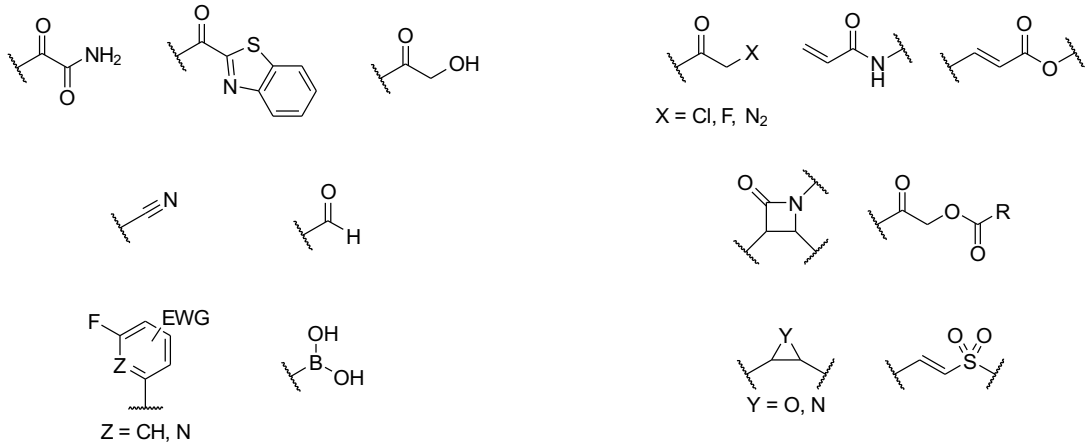


Figure 6: Overview of electrophilic warheads from research on enzymes with nucleophilic serine or cysteine residues. Figure created with ChemDraw. From left to right in rows top to bottom: α-keto amide[145,175,176], α-keto benzothiazole = benzothiazolyl ketone[145,177,178], hydroxymethyl ketone[145], chloro-[179–181]/fluoro-[181,182]/diazomethyl[183,184] ketone, acrylamide[181,185,186], acrylate[145], nitrile[145,187–189], aldehyde[145,173,188], β-lactam[190–192], acyloxymethyl ketone[145,193], S_NAr-warhead = electron-deficient (het)arene[194–196], boronic acid[197,198], epoxide = oxirane[199–201], aziridine[202,203], vinylsulfone[150,204,205]/-sulfonate[150,205]/-sulfonamide[150].

Methodic Background

Characterization of Covalent Inhibition by Mass Spectrometry

Important general methods for the assessment of covalent bond formation during target engagement are nuclear magnetic resonance (NMR) spectroscopy, mass spectrometry (MS), and protein crystallography. All three can be used to assess covalent bond formation between ligand and target but not all are always applied. Bortezomib is shown to form ^{11}B-NMR-detectable adducts with catechols[206], but the covalent interaction with the proteasome is instead characterized by mass spectrometry[207] and crystallography[96]. For the inhibition of cysteine proteases by inhibitors with aldehyde[208], nitrile[209] or α-substituted methyl ketone[210] warheads, also ^{13}C-NMR studies proved the covalency of inhibition. While NMR can be used to observe the native state of the protein-ligand complex in solution, it is a method that can have relatively low sensitivity associated with high sample consumption.[211,212] Crystallography is sought to be applied to all available complexes, since the extensive information gained for the entire interaction profile of the complex speaks for itself. But it can be time-intensive despite automation if conditions require optimization to the protein and/or ligand.[213]

An experimental focus of this work was the use of mass spectrometry in the detection of protease-ligand complexes. As ionization methods in proteomics research, both electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) are established.[214] They are both soft ionization methods that enable the detection of mostly unfragmented analytes, which can therefore be used to monitor the full-length protein. ESI is often used in line after a liquid chromatography (LC) system to separate analytes, and in combination with a quadrupole (Q) detector. On the other hand, MALDI usually starts with a solid-state sample and operates with a time-of-flight (TOF) detector. For investigations involving fragmentation, combined QTOF or alternative orbitrap mass analyzers are in use for proteomics.[215,216]

The detected molecule ions of a defined analyte differ characteristically in charge number and therefore their mass-over-charge ratio m/z when comparing ESI to MALDI, which is due to their different ways of ion formation. In ESI, an electrical potential difference applied to the spray nozzle results in the accumulation of positive charge at the liquid surface, from which positively charged droplets are dispersed (if operated in positive mode) that contain the solvated protein analyte.[217,218] These initial droplets contain an excess of cations from buffer (e.g., NH_4^+) and shrink by solvent evaporation, resulting in charge concentration. This in turn forces coulomb fission into smaller droplets. After desorption of volatile neutral molecules (e.g., NH_3), the combination of these pathways produces isolated multiprotonated analyte ions (e.g., $[M+11H]^{11+}$) that are detected as depicted in Figure 7.

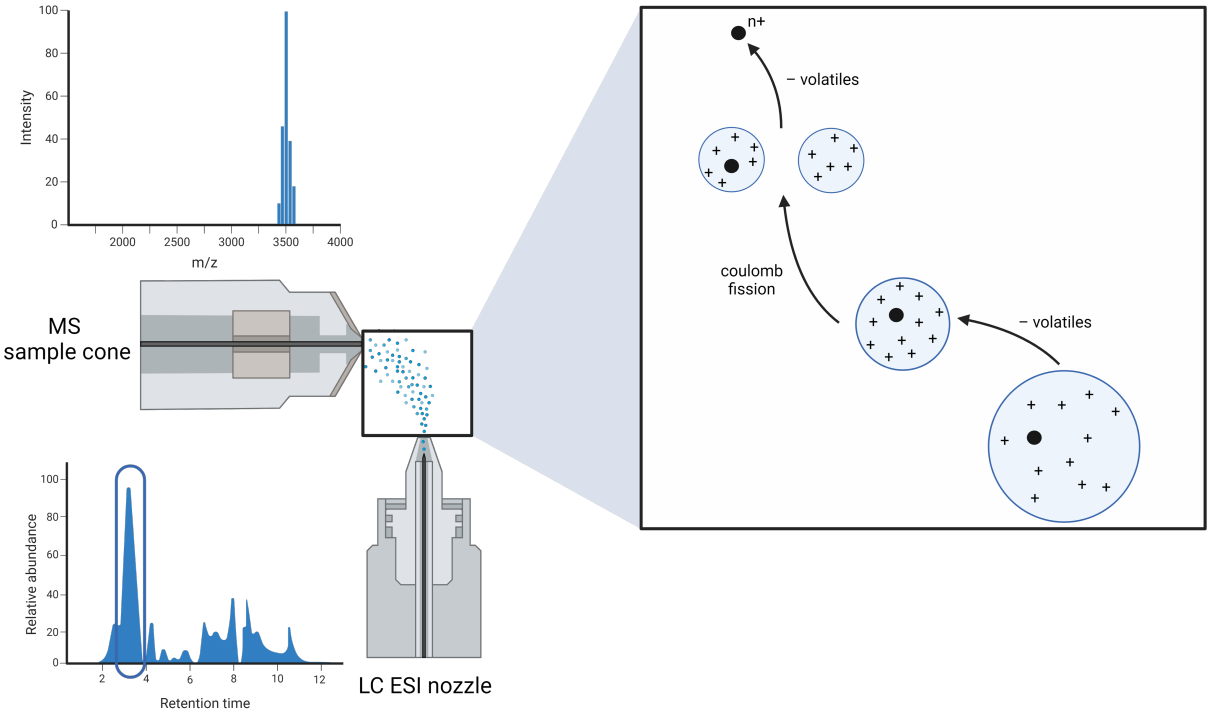


Figure 7: Mass spectrometric information derived from a system based on ESI. Starting from a liquid state, usually in line with a chromatograph (bottom left), the sample is sprayed towards the mass spectrometer. An atmosphere with reduced pressure enables the rapid evaporation of volatiles (i.e., organic solvent, water, uncharged additives), which reduces the droplet volume and induces coulomb fission through electrostatic repulsion. Large molecules therefore carry multiple charges due to the remaining protons from volatile evaporation. Figure created with Biorender.com.

Opposed to that, MALDI does not routinely produce ions from liquid solution, but usually from a solid state, in which the analyte is embedded in matrix (usually an aromatic phenol derived from benzoic or cinnamic acid).[219] This is usually prepared by mixing two separate solutions of matrix and analyte and subsequent drying as a "spot" on a stainless-steel "target". This yields a complex solid dispersion that contains the analyte ions in roughly the protonation state it had in solution, counter ions, and a large surplus of matrix molecules. These spots are then irradiated for ca. 5 ns (ca. 100 nm irradiation depth) by a laser with usually ultraviolet (UV) light of $\lambda = 337$ nm (older N_2 lasers) or 355 nm (modern frequency-tripled Nd:YAG lasers), corresponding to the matrix absorption maximum. Absorption quickly generates large amounts of localized thermal energy, which is converted to kinetic energy due to the low heat dissipation capacity of the matrix ("thermal confinement").[220] This results in thermoelastic ejection processes[221–223] called "ablation" (of clusters of associated molecules) and "desorption" (of single surface molecules, smooth transition to gas state) that take place on a microsecond timescale.[224]

The ionization process during and after desorption from the solid matrix is described by two combined theoretical concepts called "lucky survivor"[225] and "gas phase protonation".[224,226] Initially formed clusters (here for a basic analyte detected in positive mode) can contain different charge states of an analyte, i.e., $\{\text{analyte-H}_n^{n+} + \text{matrix} + m \times \text{counterion}^-\}^{(n-m)+}$. The clusters continuously reduce in size due to matrix evaporation. Evaporation of counterion after protonation (e.g., as uncharged trifluoroacetic

acid, TFA) is a dominant process leading to charge reduction of clusters if the counterion is derived from a sufficiently volatile acid. For the case of an uncharged analyte (n = m = 0), proton transfer from matrix is possible (gas-phase-protonation), gives analyte-H^+, and therefore contributes to the ion yield (signal intensity) for basic analytes.[227] However, it is most probable that proteins are already $n > 0$ times protonated in their preserved solution state. Charged clusters are always formed through "uneven" distribution of counterions in matrix that are separated during the ablation process; the charge of a cluster in deficit of counterion$^-$ corresponds to $(n-m) > 0$.[225] For $(n-m) = 1$, matrix and acid evaporation simply leads to isolated analyte-H^+. Clusters with $(n-m) > 1$ are initially present (or produced by gas-phase-protonation), but charge annihilation takes place, e.g., by reaction with electrons produced through photoionization of the matrix. This charge reduction tendency is quadratically proportional to the charge number, disfavoring multiply charged clusters.[225] The predominant detection of monoprotonated species in MALDI is therefore termed lucky survivor theory, since different ions are formed, but mostly the monoprotonated ones ($[M+H]^+$) survive. These ions are accelerated by an electric field (e.g., U = 20 kV) generated by an open electrode directly above the metal target, after which they fly in a linear motion through the mass analyzer (e.g., s = 1 m). According to the equations below, the ion's kinetic energy is proportional to their charge, resulting in different velocities inversely proportional to their mass. The flight time that named the TOF mass analyzer is therefore proportional to the square root of m/z, with a proportionality factor consisting of physical (elementary charge e, ca. $1.6 * 10^{-19}$ C)[228] and instrument (acceleration voltage U, flight path s) constants (equations (eq.) 1–3 adapted from Lit.[219]).

$$E_{kin} = z \times e \times U = \frac{1}{2} \times m \times v^2 \qquad \text{eq. 1}$$

$$v = \frac{s}{t} \qquad \text{eq. 2}$$

$$t = \frac{s}{\sqrt{2eU}} \times \sqrt{\frac{m}{z}} \qquad \text{eq. 3}$$

The exciting laser can be used to trigger the start of the time measurement, and the arrival of ions at the detector signifies its end. Using the values and equations given above, and for ions with m/z up to 1 MDa this results in flight times below 1 ms. The detector in MALDI-TOF instruments is often a microchannel plate (MCP), to which an electrical field is applied, and therefore functions as an electron multiplier. The impact of ions onto the plate results in the liberation of multiple electrons that, in turn, "bounce" down the microchannels, liberating more electrons that are finally registered. Like this, the signal of a single ion is amplified. An interesting advantage of TOF mass analyzers is their theoretically infinite mass range. However, since MCP detectors have an increasingly poor ion-to-electron conversion rate for heavier (slower) ions, there can be an effective upper limit to the detected masses.[229] But with improved ion conversion dynode detectors, this limit is in the MDa-range, so most (bio-)polymeric analytes are still well accessible with this technology.[230] The MALDI-TOF process is depicted in Figure 8.

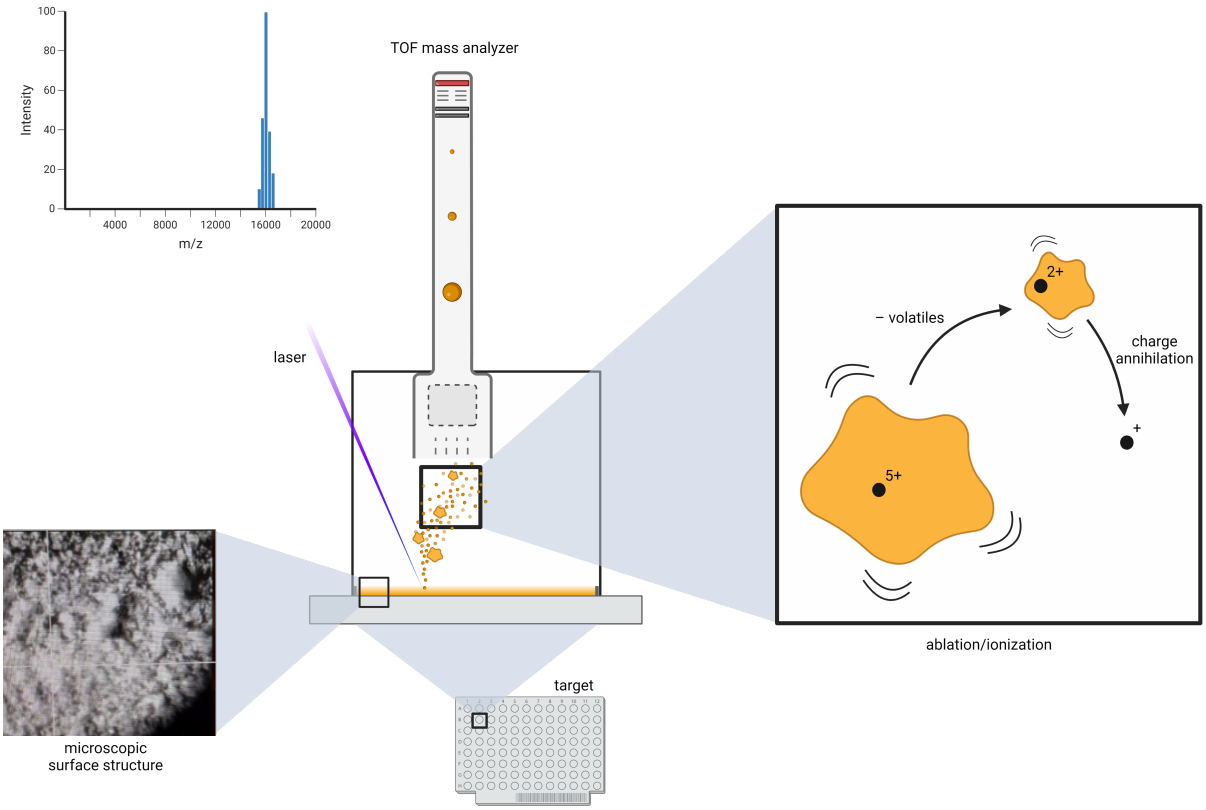


Figure 8.: Mass spectrometric information derived from a system based on MALDI. On a metal plate (target), the dried sample (microscopic surface structure photography captured during measurement) is irradiated by a laser. This desorbs single molecules and ablates excited clusters of molecules. From these, volatiles (e.g., matrix, residual solvent) evaporate, reducing cluster size, and charge annihilation takes place in multi-charged analyte molecules, resulting in ions with low charge numbers. Ion separation is achieved by TOF. Figure created with Biorender.com.

While both methods are suitable for the detection of intact proteins, alternative setups exist, for which a digest of the target protein with low-selectivity proteases can be performed prior to mass spectrometric analysis to generate a characteristic mosaic of peptides.[231] It is often performed in a tandem (MS/MS) approach that entails fragmentation of initially formed ions.[219] This approach requires prior sequence information of the protein for mass allocation, and therefore more computational support than the analysis of the intact protein. In the context of chemically modified proteins (e.g., through inhibitors), intact detection methods give the opportunity to evaluate more fragile complexes. Meanwhile, digestion makes only very stably bound ligands accessible, because the non-covalent binding environment is fully degraded during sample preparation. However, analyzing digestion products gives information about the exact position of a modification, elucidating reactive sites within the target protein. Another advantage of analyzing a digested analyte (or a multiprotonated analyte) is that resolution usually declines with increasing m/z—smaller peptides or higher charged states are detectable with higher mass accuracy and precision.[232] In absolute numbers, a precision of ≤5 Da for 20–70 kDa proteins (≤100 ppm)[233] can be reached with ESI-MS, whereas a precision of <1 Da (<200 ppm)[232] is realistic for peptides <5 kDa. MALDI-TOF can achieve similar precisions for these peptides[232,234,235] and intact proteins[236]. However, ESI is generally considered the more precise method due to the discussed

advantage of detecting lower m/z through the higher charge number. As a confounding factor, proteins can present as a heterogenous population with a group of signals, resulting from inaccuracies in maturation cleavages[236], posttranslational modifications[237], or degradation processes[238,239]. These signals can further be obscured by fragmentation or the presence of matrix[240,241] or metal ion[238] adduct peaks as artifacts from sample preparation. The mass spectrometric signal of a protein sample therefore rarely contains only one peak.

An important question to be asked in the context of characterizing inhibitors and their assumed covalent interaction with the target is whether the employed method can discriminate between non-covalent and covalent complexes. In principle, there are reports for both ESI- and MALDI-based mass spectrometry to detect non-covalent complexes. However, this is far easier with ESI[242] than it is with MALDI[229]. There are case reports with exceptional parameters in MALDI that may enable the detection of non-covalent complexes. Usually, this is described for biologically multimeric structures with large interaction surfaces (e.g., protein oligomers[243,244] or double-strand DNA[245]), or for multiple strong ionic interactions (e.g. polyanion + polycation[246]). To detect non-covalent complexes, the harsh acidic and partially organic standard environment for co-crystallization can be changed to neutral pH and absence of solvent to not interfere with ionization states.[243] Neutral matrices (e.g., 2,6-dihydroxyacetophenone or 6-aza-2-thiothymine)[246] can be used with the same argument. UV irradiation can be changed to softer infrared (IR) irradiation.[244] These example modifications can work in some cases, but are usually considered exceptional. However, the detection of non-covalent complexes in “first shot” spectra is a more frequently described occurrence in MALDI-MS.[247,248] It is assumed that the analyte molecules on the surface of the matrix have the highest probability of retaining non-covalent interactions. Additionally, the “intensity-fading” protocol was described for non-covalent protease inhibitors from biological extracts, that does not monitor the generation of an adduct peak, but the suppression of the ligand peak at low concentrations through the addition of a protein target.[249] Still, the routine conditions in MALDI (acidic crystallization solution in presence of an organic solvent, an acidic matrix, multiple shots per spectrum, high UV laser fluences) are commonly assumed to only enable the detection of covalent complexes, because only covalent interactions tend to survive these conditions.

Modes of Inhibition—Conceptualization and Qualitative Differences

Apart from the classical distinction between competitive/orthosteric, non-competitive/allosteric, and uncompetitive modes of inhibition[250], there are subtleties regarding inhibition kinetics that deserve a closer look. Different structural features on the inhibitor (i.e., reactive warheads) can lead to different intermediate states over the course of inhibition. To start off, some important distinctions for descriptive terms are made here regarding covalency, reversibility, and time-dependence of inhibition.[162]

The theoretical basis and experimental characterization of **covalent** binding was already discussed in previous chapters. For proteases, covalent inhibition is common as they often carry a conserved nucleophilic residue, and their inhibitors are often equipped with a suitable electrophile. Despite this, covalent inhibition is not proven by this alone; positioning is important. A distance criterion from literature to expect productive covalent bond formation between reactants from a non-covalently associated state is <4 Å.[251] Larger distances (e.g., 7 Å) usually do not produce covalent bonds.[252] But purely non-covalent interactions are described crystallographically for a suitable warhead in close proximity (<3 Å) to a Cys protease's cysteine with an electron-density gap.[253]

There are **reversible** and **irreversible** reaction paths. While this seems a strict dichotomy at first, there is a gradient between the two extremes, which is described by the dissociation rate of the formed complex. A truly irreversible reaction has a dissociation rate approaching 0 s^{-1}, which means that the product does not spontaneously dissociate, and physiologically the target is inhibited until it is degraded and resynthesized. Then there are reversible complexes with measurable slow dissociation rate (e.g., 0.16 ms^{-1})[254], meaning that the ligand has a prolonged residence time within the formed complex (e.g., 104 min[254], but the definition "slow" can range from a few minutes to weeks[255,256]). Finally, there are cases with a relatively fast dissociation rate, in which the residence time is small (e.g., 9 min)[254] or even too small for quantitative assessment. This kinetic distinguishment is not reserved for covalent inhibitors, even though it is frequently encountered there in the context of protease inhibitors (e.g., captopril binding to ACE, leupeptin binding to CatG).[255,256] There are well known non-covalent drugs (e.g., methotrexate binding to dihydrofolate reductase, bioactivated acyclovir binding to HHV DNA-polymerase)[255,256] and non-covalent experimental probes (e.g., pepstatin binding to pepsin, compactin binding to HMG-CoA reductase)[255] with slow binding kinetics.

There is **time-independent** and **time-dependent** inhibition. These terms describe whether the equilibration time of the observed reaction happens on a fast (time-independent) or on a slow (time-dependent) scale compared to the timeframe of the experiment. This is tied closely to the reversibility discussed above. In general, inhibitors with low dissociation rates present as time-dependent inhibitors, which means that these inhibitors give an increasingly strong effect within the assay recording period, because the equilibration of a high affinity complex happens on the same timescale as assay progression. Only those with high dissociation rates show time-independent inhibition profiles (Figure 9). As discussed before, this is unrelated to covalency and only in some cases

directly indicative of reversibility, even though it gives hints. For instance, an irreversible mode of inhibition can never be time-independent, but it can be only prominent after hours of recording time.

For further explanations on the modes of inhibition most relevant for research on protease inhibitors, the following setup for protease assays is assumed, from which quantifiable data can be generated (Figure 9). A protease of interest (e.g., *Tb*CatL) is dissolved in a buffer that ensures optimal activity. A fluorogenic peptide is added as the substrate that carries a quenched fluorophore at its *C*-terminus (e.g., carboxybenzyl (Cbz)-Phe-Arg-7-amino-4-methylcoumarin (AMC)), which shows fluorescence only after cleavage.[257] The increase in fluorescence over time is detected. This is performed in absence and presence of inhibitors in specific concentrations to evaluate inhibition potency of the inhibitors in vitro. Depending on the inhibitor, different characteristics regarding time-dependence of inhibition can be observed.

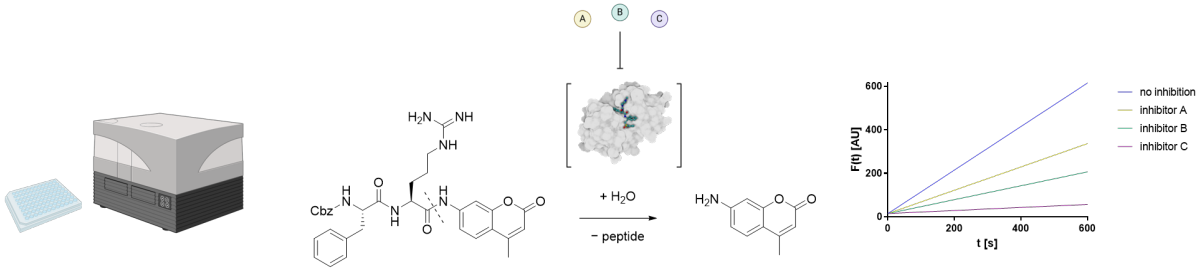


Figure 9: The assay, with which investigations on affinity and reversibility are derived, is performed in a 96-well plate setup with continuous fluorometric detection. The detected fluorophore is generated in a hydrolysis reaction catalyzed by the protease of interest (here *Tb*CatL, pdb: 2p7u) from a suitable substrate (Cbz-Phe-Arg-AMC). In absence of an inhibitor, this results in the constant production of fluorophore per time unit. Its increasing concentration is proportional to the fluorescence detected and the production rate is descriptive of the protease's activity. In presence of an inhibitor, a reduction in this activity is recorded as reduced slopes of the F-t-diagram. This is indicated for three time-independent inhibitors A–C that have different impacts on the protease activity which corresponds to their affinity towards it. Figure created with Biorender.com and ChemDraw.

The most relevant modes of inhibition for proteases in the described assay are depicted in Figure 10. In all cases, a reporter reaction (Figure 10A) indicates substrate conversion by a linear increase of fluorescence over the observation period. For the depicted cases of competitive inhibition, the presence of substrate directly influences the apparent affinity for an inhibitor under investigation. In the presence of substrate, the inhibitor seems weaker, because its binding site in the enzyme can be occupied by the substrate.[256,258] This influences the calculations explained below, but it is necessary for these functional assays, and it is predictive of the situation in vivo, in which multiple potential ligands can compete. If only binding of an inhibitor is investigated (then in absence of a substrate), other methods like isothermal titration calorimetry, differential scanning fluorometry, or microscale thermophoresis can be employed.[259] In general, diffusion is assumed as the fastest process involved in binding, so that all molecular exchange (occupation of unadapted active site by substrate or inhibitor) happens instantaneously.[260] Other processes like conformational change of the active site and association between enzyme and ligand through non-covalent (and covalent) interactions are assumed to happen on a slower timescale which gives the values for affinity and reaction kinetics.

Combinations of covalent and non-covalent states are possible in both reversible and irreversible inhibition, all connected by reaction equilibria. For a purely non-covalent inhibitor, often only one equilibrium is established (Figure 10B), but large conformational changes can introduce a two-step non-covalent reaction.[261] Some covalent inhibitors can bind in a similar "one-phase association" (Figure 10C). These can be irreversible affinity labels, but also designed covalent reversible enzyme inhibitors.[139,256] For these, it is assumed that there is no relevant preliminary non-covalent complex that needs to be formed, so the only kinetic relevance is the one corresponding to warhead reactivity. This one-phase association for covalent inhibitors is usually undesired as it hints towards an unselective process. Instead, for designed inhibitors with a covalent warhead, it is expected that they will initially produce a non-covalent complex with the enzyme. After the necessary conformational adaptations, the covalent bond can form (Figure 10D–E). This entails two sets of on- and off-rates in parallel for the two equilibria; a case commonly described as a "two-phase association". In the case of irreversible inhibition (Figure 10E), the kinetics of these two equilibria are simplified due to the absence of a second off-reaction.

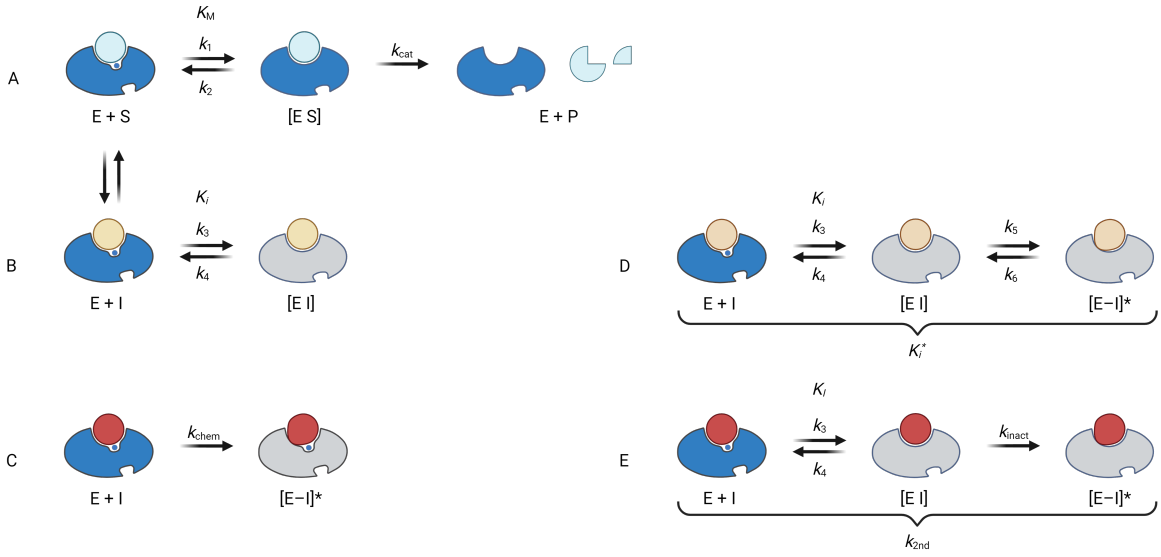


Figure 10: The substrate is depicted in light blue, inhibitors in orange (reversible) or red (irreversible). A distortion of the inhibitor circle represents covalent bond formation. The enzymes are depicted in blue (active) or grey (inactive) conformation, the expulsion of active-site water (small blue circle) represents required conformational adaptation of the enzyme to its ligand and the entropic influence of ligand binding. The presented type of kinetic investigations of protease inhibition is based on a reporter reaction (A) and the changes to the reaction in presence of a competitive inhibitor (putative allosteric pocket shown on the backside of the enzymes). Depicted are four ways in which inhibitors can compete with the substrate in binding to an enzyme. (B) Non-covalent reversible inhibitor, one-phase-association. (C) Covalent irreversible inhibitor, one-phase-association. (D) Covalent-reversible inhibitor, two-phase-association. (E) Covalent-irreversible inhibitor two-phase-association. The commonly reported parameters for kinetics and affinity are already depicted here on the reaction arrows and will be discussed below. It should be noted that a simplifying assumption of $K_M = K_S$ and $K_I = K_i$ is made for reactions A and E, respectively. The distinctions between these terms are discussed in Lit.[256,262] Figure created with Biorender.com.

One major difference between the modes of inhibition depicted in Figure 10 is their reversibility. In Figure 11 the reaction progress for two different two-phase associating covalent inhibitors

(Figure 10D–E) is depicted. In this diagram, the difference in free energy between the reaction intermediates (ΔG) is linked to affinity (K_D), while the activation free energies ($\Delta G^{\ddagger}$) translate to the association and dissociation rates ($k_{on/off}$) of a reaction.[263] In the chosen cases, it is the energetic state of the final product, i.e., the ΔG of the second reaction, that decides if the reaction is reversible. This assumes energetically comparable activation energies for the reaction of two distinct molecules (e.g., warhead-carrying inhibitors that differ in reversibility) with the same enzyme, and therefore similar association rates. This simplification was shown to hold true for similar Cys proteases (rhodesain, cruzain), that are attacked by vinylsulfone, α-fluorovinylsulfone, and nitrile warheads on the same peptide backbone.[264,265] The absolute activation energies (and therefore association rates) predicted by QM/MM or determined experimentally depended on the employed method, but were indeed comparable for the investigated warheads within each method. Hence, the energy liberated during the reaction decided on reversibility (5.9 kcal/mol for a reversible nitrile, 16 kcal/mol for a reversible α-fluorovinylsulfone, and 23–26.7 kcal/mol for an irreversible vinylsulfone) due to the different dissociation rates of the formed products.[264,265]

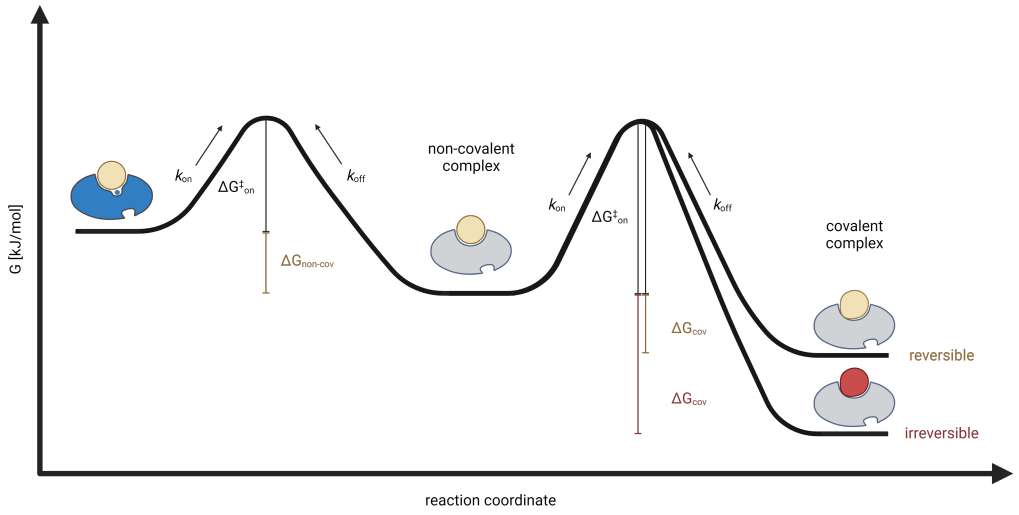


Figure 11: Example reaction path for two inhibitors that react covalently in two-step mechanisms. Inhibitors are depicted in orange (reversible) or red (irreversible) circles (non-covalent) or distorted circles (covalent). Enzyme is depicted in blue (active) or grey (inactive). ΔG between reaction intermediates is proportional to the affinity of the complex. $\Delta G^{\ddagger}_{on}$ denotes the required activation free energy for the forward reaction. $\Delta G^{\ddagger}_{off}$ as the activation free energy of the back reaction is not depicted for the sake of clarity but corresponds to the sum of $\Delta G^{\ddagger}_{on}$ and ΔG. In this constructed case, the $\Delta G^{\ddagger}_{on}$ for the covalent step is larger than for the non-covalent step. The covalent $\Delta G^{\ddagger}_{on}$ is equal to the $\Delta G^{\ddagger}_{off}$ of the non-covalent complex, making the two reactions from the noncovalent complex similarly probable. The main denominator of reversibility is the required $\Delta G^{\ddagger}_{off}$ of the second step, which is represented here by the difference in ΔG due to identical $\Delta G^{\ddagger}_{on}$. In the depicted cases, the $\Delta G^{\ddagger}_{off}$ of the reverse reaction of the red inhibitor is largest and assumed to not be provided by the system, making this reaction the only irreversible one. Figure created with Biorender.com.

Commonly performed experiments to determine reversibility are the (jump) dilution assay (Figure 12) and the dialysis assay (Figure 13).[264,266] For a dilution assay, high concentrations of both enzyme and inhibitor are preincubated for a time that allows equilibration of all reactions included in the inhibition mechanism. Afterwards, this complex is diluted rapidly by the addition of substrate solution to create an

environment of competition between inhibitor and substrate. In the case of an irreversible inhibitor, no substrate conversion is recorded, because competition is impossible if the active site is not vacated by the irreversibly attached inhibitor. For a reversible inhibitor, substrate conversion can be observed and further characterized. If the recovery of enzyme activity is instantaneous, the inhibition equilibria incl. dissociation rates are fast. A slow recovery of enzyme activity is indicative of slow inhibition kinetics.

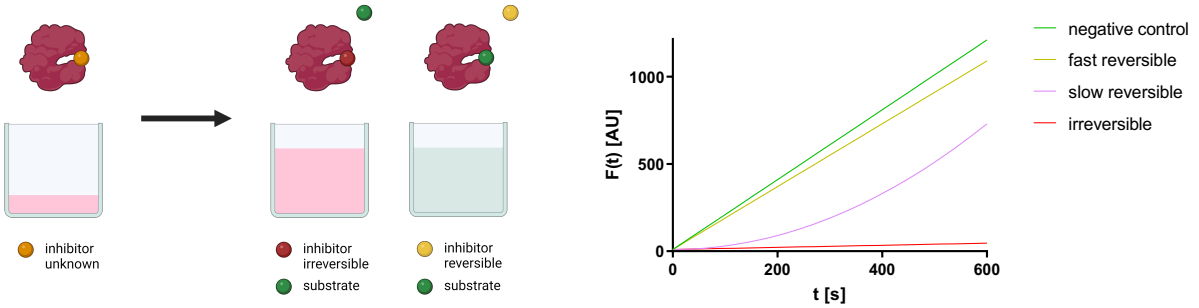


Figure 12: (Left) Schematic depiction of a (jump) dilution assay (dilution is portrayed by the filling level of the wells). Enzyme in red, inhibitor as balls, colored by reversibility. The well contents are colored by enzyme activity in red (inactive) or green (active). After incubation of the enzyme with an inhibitor of unknown reversibility, buffer and substrate are added to investigate how quickly the activity of enzyme recovers. (Right) Progress curves for inhibitors of different reversibility in comparison to a negative control without inhibitor (adapted from Lit.[256]). Figure created with Biorender.com.

A similar principle is followed by a dialysis assay. The enzyme is incubated with excess inhibitor to equilibration (full inhibition). Afterwards, the complex is placed in a compartment connected by a membrane that lets inhibitor but not enzyme pass (hydrodynamic volume cutoff) to a dialyzing flow of buffer. At multiple time points during dialysis, samples are taken, and substrate solution is added. Dialysis is only able to remove unbound inhibitor molecules, so for irreversible inhibitors, no enzymatic activity is recorded at any time. For reversible inhibitors, the inhibitor is still accessible to dialysis at some point, so a recovery of enzymatic activity is recorded with time.

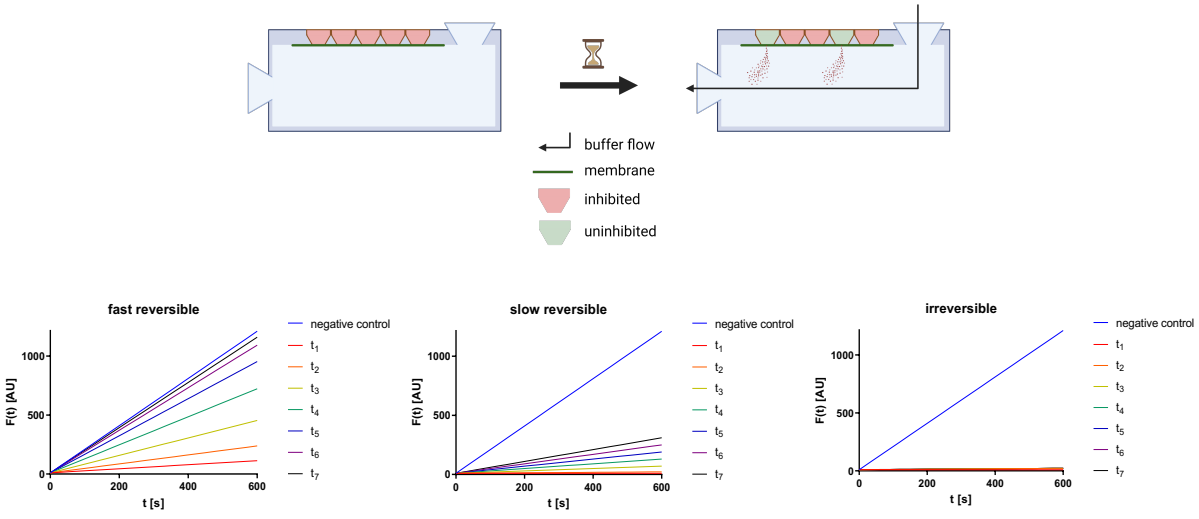


Figure 13: (Top) Schematic depiction of a dialysis assay setup (tubing excluded).[267] Sample wells containing enzyme and inhibitor are incubated in contact with a dialysis chamber. At different times, samples are taken and evaluated for recovered enzyme activity. (Bottom) Progress curves for the samples taken at different times of the dialysis for three inhibitors with different degrees of reversibility. Figure created with Biorender.com.

Modes of Inhibition—Quantitative Descriptors

The quantitative evaluation of data from protease inhibition assays is a field of research that has been discussed in extensive detail in the literature.[256,268,269]

For the discussed modes of inhibition, versions of eq. 4 can describe the fluorescence progression over time.[268] To link the measurement method and the reporter reaction, fluorescence (in arbitrary units AU) is proportional to product formation (Figures 9 and 14). The slope of a section of the F-t-diagram is described by a rate or velocity v [AU/s], further specified as an initial velocity (v_i), or a steady-state velocity (v_s). If v_i and v_s are not identical, so for time-dependent inhibitors, the rate of change from v_i to v_s is described as k_{obs}. The terms v_i, v_s, and k_{obs} are dependent on the inhibitor concentration [I], and the relationship between the three and [I] is characteristic for the investigated inhibitor and its mode of inhibition (described separately for each case below). The background fluorescence ($F_{background}$) is usually constant for the method employed.

$$F(t) = F_{background} + v_s \times t + \frac{v_i - v_s}{k_{obs}} \times (1 - e^{-k_{obs} \times t}) \qquad \text{eq. 4}$$

For time-independent inhibitors (possible for Figure 10B and D), the right summand of eq. 4 is eliminated because $v_i = v_s$, resulting in the simple linear equation 5 as depicted in Figure 14.

$$F(t) = F_{background} + v_s \times t \qquad \text{eq. 5}$$

Eq. 4 can be employed in its entirety if there is a slowly established but reversible equilibrium contributing to the inhibition (possible for Figure 10D). The resulting graph has different v_i and v_s for each [I] and each v_s is approached by the respective v_i with the rate k_{obs} as depicted in Figure 14.

$$F(t) = F_{background} + v_s \times t + \frac{v_i - v_s}{k_{obs}} \times (1 - e^{-k_{obs} \times t}) \qquad \text{eq. 4}$$

Lastly, for irreversible inhibitors $v_s = 0$ (Figure 10E) for all [I], which represents a major difference to the prior example. With time, the enzyme is always fully inactivated. Therefore, the middle summand of eq. 4 is eliminated, and the right summand simplifies to give eq. 6. Graphically this means that for all [I], the F-t-diagram approaches a parallel to the t-axis, and the rate at which it does that (k_{obs}) is dependent on [I], as depicted in Figure 14. Again, v_i is dependent on [I] of the inhibitor of interest.

$$F(t) = F_{background} + \frac{v_i}{k_{obs}} \times (1 - e^{-k_{obs} \times t}) \qquad \text{eq. 6}$$

The dependency of v_i to [I] is the major difference to the case depicted in Figure 10C. For this exceptional one-step irreversible inhibition, v_i is independent of [I] (plotting equation not discussed here).[268]

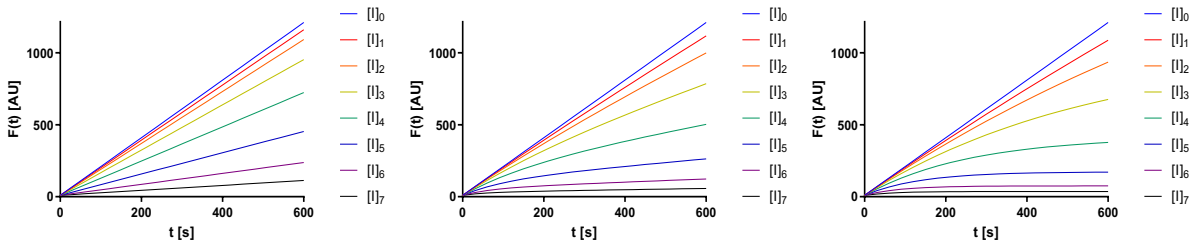


Figure 14: Modelled progress curves with fluorescence signal proportional to product concentration at seven increasing inhibitor concentrations for the time-independent (left, eq. 5), time-dependent & reversible (center, eq. 4), and time-dependent & irreversible (right, eq. 6) characteristics ($[I]_0$ = 0 nM, $[I]_1 \rightarrow [I]_7$ is increasing). Figure created with GraphPad Prism.

Inhibition potency and kinetics can be compared between different inhibitors. To do this, there are defined values associated with each degree of reversibility. A value commonly used indiscriminately is the IC_{50} value, which is related to the observed activity-reduction of an enzyme.[270] In the assay described above, this would be the [I] that results in a reduction of the observed slope to 50 % (e.g., $[I]_5$ in Figure 14 (left)). IC_{50} alone has only limited value because it assumes equilibrium conditions and is therefore easily disturbed by time-dependent assay progression. Also, it is recorded in the presence of substrate, so it is only comparable for identical assay conditions and therefore disturbed, e.g., by modifications to substrate concentration and identity.[271] However, it can be used when the inhibition mechanism is unclear or rough estimates are sufficient for subsequent decisions. Still, the aim is to give experimental results in more comparable metrics that correspond clearly to specific mechanisms. The employed nomenclature to describe inhibition is close to Michaelis-Menten kinetics (K_M and k_{cat}).[272] Equilibrium constants are defined by K that are further distinguished between inhibition modes by sub- or superscript. The terms commonly encountered are K_i, K_i*, and K_I (compare Figure 10, all described in detail below). All equilibrium constants in this notation are already corrected for substrate competition. Without this correction they are further denoted as "apparent", e.g., K_i^{app}. Detailed information concerning reversibility is given by reaction rates k, which are indexed in subscript to allocate meaning. Here, the encountered terms are k_1–k_6, and k_{inact}. While k_{2nd} is denoted similarly, it has a distinct meaning described below.

To further explain the most relevant kinetics, the assay behaviors of three imaginary inhibitors with comparable inhibition parameters were modelled as quantifiable examples using Microsoft Excel and equations 4–6, 8, 11, 14, and 16. The following three cases were constructed:

a) One time-independent (reversible), non-covalent inhibitor with IC_{50} = 15 µM and fast (unresolved) kinetics. (Figure 15)
b) One time-dependent (slowly reversible), covalent inhibitor with K_i^{app} = 15 µM, K_i^{*app} = 1.5 µM, k_5 = 0.018 s^{-1}, and k_6 = 0.002 s^{-1}. (Figures 16 and17)
c) One irreversible, covalent inhibitor with K_I^{app} = 15 µM, and k_{inact} = 0.018 s^{-1}. (Figure 18)

The necessary equations to determine IC_{50} for case (a) are depicted below (eq. 5, 7, and 8) with constant v_s (= v_i) at a certain (variable) inhibitor concentration [I]. Eq. 7 combines absolute v_s at different [I] for plotting, while eq. 8 calculates relative to the enzyme activity in absence of inhibitor (= $[I]_0$). The numerator in eq. 7 gives the slope range (difference between v_s of lowest and highest [I]), which in relative numbers should span 100 %, while the extrapolated background activity should reach 0 %. From experience, deviations from these >10 % can result from a faulty negative control but can also be associated with limited inhibitor solubility.[256] Lastly, s denotes the Hill slope that corresponds to the steepness of this sigmoidal plot (using a log-abscissa), for which large deviation from a value of |s| = 1, e.g., |s| > 2 or |s| < 0.5, can indicate protein denaturation or the presence of two overlapping binding events, respectively.[256] Due to these four parameters that can be modulated to optimize the goodness of fit (range, background, Hill slope, IC_{50}), eq. 7 or 8 is called "four-parameter equation" or "Hill equation".[273] It is a modification of the two-parameter binding isotherm equation that assumes no background and full coverage of the slope range.[256] The graphical depiction of case (a) according to eq. 5 and 8 is shown in Figure 15. The IC_{50} of 15 µM is visualized by the [I] that results in 50 % activity.

$$F(t) = v_s \times t + F_{background} \qquad \text{eq. 5}$$

$$v_s([I]) = \frac{v_{s,range}}{1+(\frac{[I]}{K_i^{app}})^s} + v_{s,background} \qquad \text{eq. 7}$$

$$activity([I])[\%] = \frac{v_s([I])}{v_{s,[I]_0}} = \frac{\frac{v_{s,range}}{v_{s,[I]_0}}}{1+(\frac{[I]}{K_i^{app}})^s} + \frac{v_{s,background}}{v_{s,[I]_0}} \qquad \text{eq. 8}$$

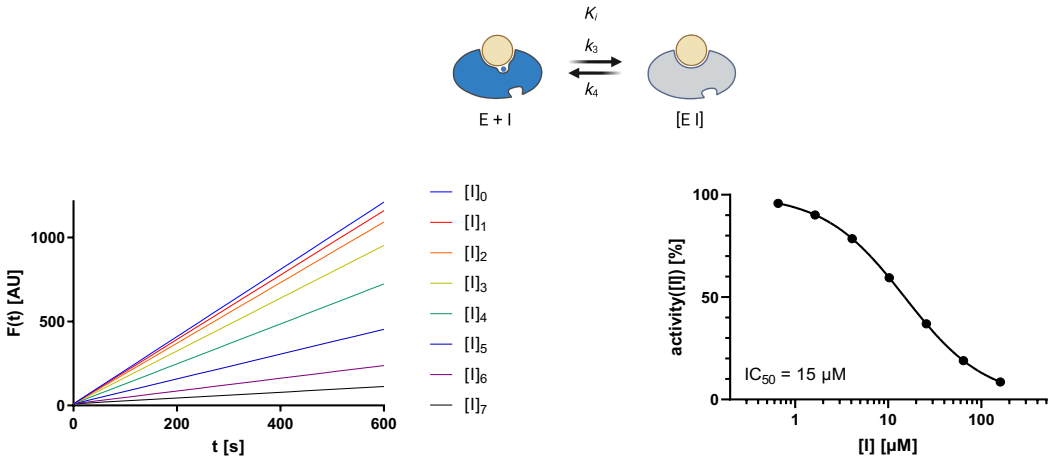


Figure 15: (Top) Repeated illustration of the binding mode of case (a). (Bottom left) Artificial dataset for a time-independent reversible inhibitor using seven different inhibitor concentrations [I] and a negative control without inhibitor $[I]_0$. (Bottom right) Transformation of the artificial dataset for [I] and activity([I]) (black spheres) and fit using eq. 8 (black line) with the following parameters: range = 100 %, background = 0 %, s = 1, IC_{50} = 15 µM. Figures created with Biorender.com and Graphpad Prism.

As a remark, time-independent inhibition can represent a non-covalent or a covalent inhibitor, if it shows fast equilibration kinetics for all contributing equilibria and a reversible reaction in general. Instead of IC_{50}, K_i^{app} is commonly used for the latter covalent reversible case. However, in this covalent case involving two reaction equilibria, K_i^{*app} (see below) should be used instead of K_i^{app}, since only the combined contribution in the steady state (v_s) can be recorded. Still, this is usually not done, and K_i^{app} is used for time-independent covalent inhibitors.

Next are the necessary equations to determine K_i^{app} and K_i^{*app} for case (b). Two possibilities exist to evaluate them. The first one (eq. 9–12, 5, 7, 8, and 13) is essentially the abovementioned case performed twice: once for v_i and once for v_s. It is arguably the more descriptive of the two methods to characterize time-dependent reversible inhibitors. However, it does not provide information on association and dissociation rates, as opposed to the second method described below. For time-dependent inhibitors, $v_i \neq v_s$, which is the distinction from the time-independent case (a). Sometimes, with optimal coverage of the effect range, eq. 11 and 8 can be simplified to a one-parameter equation with only K_i^{app} or K_i^{*app} as the plotting parameter (eq. 12 and 13).[256] An example of how to determine v_i and v_s for each [I], based on the four-parameter equations to subsequently determine K_i^{app} and K_i^{*app}, are shown in Figure 16. The K_i^{app} of 15 µM and K_i^{*app} of 1.5 µM are visualized by the [I] in each of the two graphs that results in 50 % activity.

$$F(t) = v_i \times t + F_{background} \quad \text{eq. 9}$$

$$v_i([I]) = \frac{v_{i,range}}{1+(\frac{[I]}{K_i^{app}})^s} + v_{i,background} \quad \text{eq. 10}$$

$$activity([I])[\%] = \frac{v_i([I])}{v_{i,[I]_0}} = \frac{\frac{v_{i,range}}{v_{i,[I]_0}}}{1+(\frac{[I]}{K_i^{app}})^s} + \frac{v_{i,background}}{v_{i,[I]_0}} \quad \text{eq. 11}$$

$$activity([I])[\%] = \frac{v_i([I])}{v_{i,[I]_0}} = \frac{1}{1+(\frac{[I]}{K_i^{app}})} \quad \text{eq. 12}$$

$$F(t) = v_s \times t + F_{background} \quad \text{eq. 5}$$

$$v_s([I]) = \frac{v_{s,range}}{1+(\frac{[I]}{K_i^{* app}})^s} + v_{s,background} \quad \text{eq. 7}$$

$$activity([I])[\%] = \frac{v_s([I])}{v_{s,[I]_0}} = \frac{\frac{v_{s,range}}{v_{s,[I]_0}}}{1+(\frac{[I]}{K_i^{* app}})^s} + \frac{v_{s,background}}{v_{s,[I]_0}} \quad \text{eq. 8}$$

$$activity([I])[\%] = \frac{v_s([I])}{v_{s,[I]_0}} = \frac{1}{1+(\frac{[I]}{K_i^{* app}})} \quad \text{eq. 13}$$

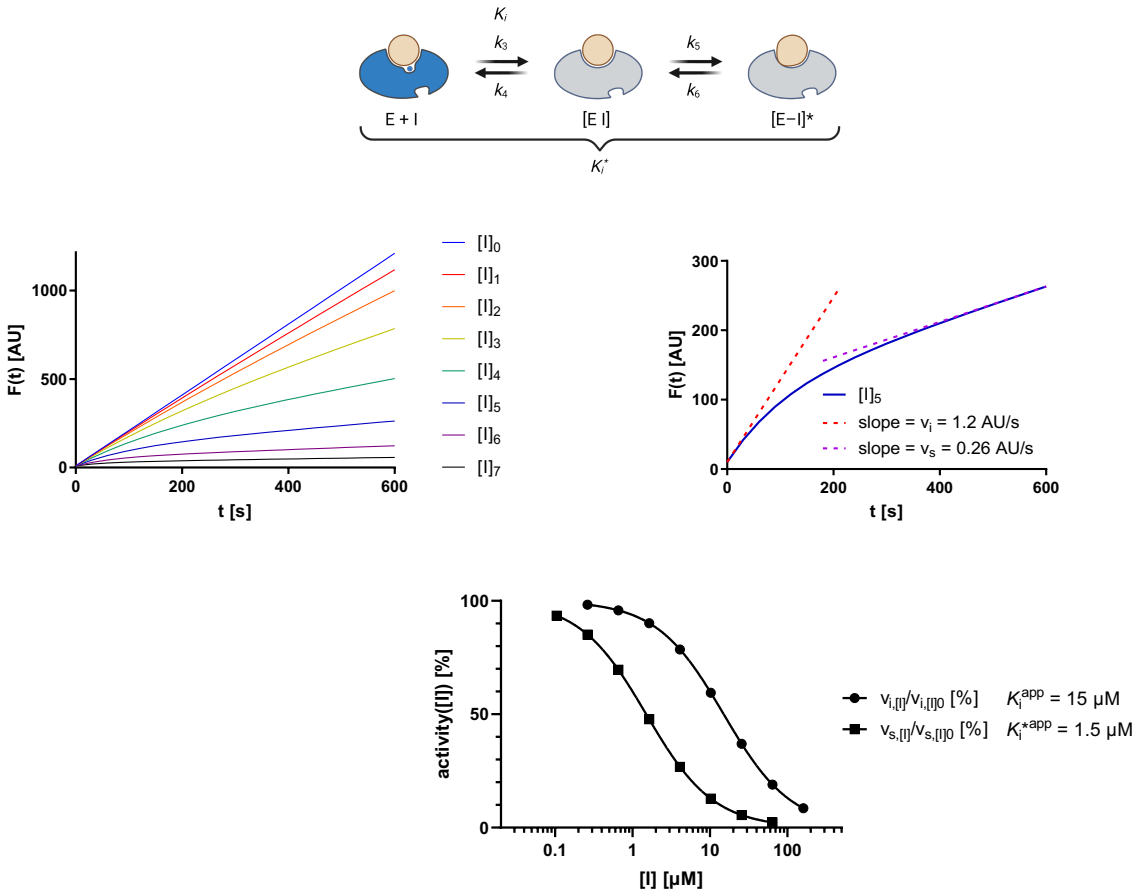


Figure 16: (Top) Repeated illustration of the binding mode of case (b). (Center left) Artificial dataset for a time-dependent reversible inhibitor using seven different inhibitor concentrations [I] and a negative control without inhibitor $[I]_0$. (Center right) Closer investigation of the data for $[I]_5$ to calculate initial and steady-state slopes by linear regression. (Bottom) Transformation of the artificial dataset with [I] and activity([I]) (black spheres using v_i and black squares using v_s) and fit using eq. 11 and 8 (black lines) with the following parameters: range = 100 %, background = 0 %, s = 1, K_i^{app} = 15 µM, K_i^{*app} = 1.5 µM. Figures created with Biorender.com and Graphpad Prism.

The second method does not only use the initial and final state for evaluation but instead focuses on the rate at which the steady state is reached (k_{obs}), as originally described by eq. 4. It uses the plotted k_{obs} of each individual [I] and relates them to each other (eq. 14). From this, the association and dissociation rates (k_5 and k_6 in Figure 17), can be derived separately. In this regard, k_5 is the maximal k_{obs} from v_i to v_s and k_6 is its offset. Therefore, the observed maximal k_{obs} is the sum of both. The non-covalent equilibrium constant is also part of eq. 14 and modulates the "steepness" with which the maximal k_{obs} is approached by the graph; the lower the K_i^{app}, the faster the approach. Using this method, K_i^{app} is derived from Figure 17 as the [I] with 50 % of the maximal k_{obs} (0.020 s^{-1}), which is 15 µM. From the same diagram, k_6 is defined as the intercept of the plot with the k_{obs}-axis (0.002 s^{-1}). A point of concern here: k_6 should not equal the measured k_{obs} of the negative control ([I] = 0 nM). This point should not be included in the plotted data because it should have no k_{obs}, because enzyme activity in absence of inhibitor should be constant ($v_i = v_s$). However, this intercept is very prone to cases where the negative control has an apparent k_{obs}; this at least introduces some error or can make it impossible to accurately determine k_6.[268]

The estimate of K_i*app is then based on a calculation involving K_i^{app} (the non-covalent equilibrium constant) and the rates of the covalent equilibrium (k_5 and k_6) using eq. 15. For the constructed case, $k_5/k_6 = 9$ to connect the K_i^{app} of 15 µM and the K_i*app of 1.5 µM.

$$k_{obs}([I]) = k_6 + \frac{k_5 \times [I]}{K_i^{app} + [I]} \qquad \text{eq. 14}$$

$$K_i^{* app} = \frac{K_i^{app}}{1 + \frac{k_5}{k_6}} \qquad \text{eq. 15}$$

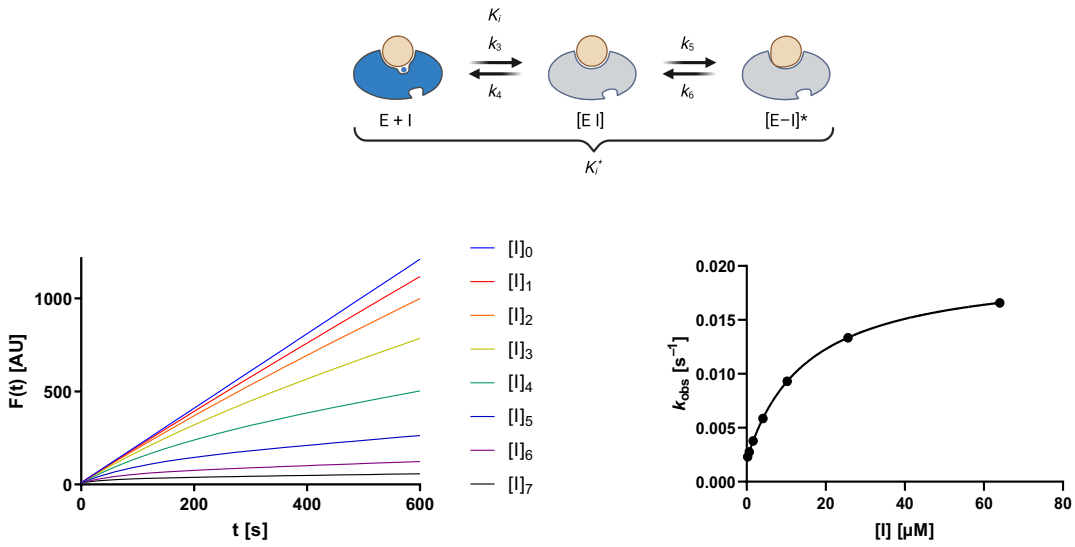


Figure 17: (Top) Repeated illustration of the binding mode of case (b). (Bottom left) Artificial dataset for a time-dependent reversible inhibitor using seven different inhibitor concentrations [I] and a negative control without inhibitor $[I]_0$. (Bottom right) Transformation of the artificial dataset with [I] and activity([I]) (black spheres) and fit using eq. 14 (black line) with the following parameters: $k_5 = 0.018$ s^{-1}, $k_6 = 0.002$ s^{-1}, $K_i^{app} = 15$ µM, K_i*$^{app} = 1.5$ µM. Figures created with Biorender.com and Graphpad Prism.

A point of concern should be mentioned here. Since an "initial" state (v_i) is essentially only recordable at the first moment of the reaction, it is much more prone to error than the steady state. The time between reaction start and data acquisition (usually a few seconds) can be problematic. It distorts the estimates for K_i^{app}, as well as k_5 (and k_6), especially since this is most influential at high [I]. This is true for all cases here, where an initial state is supposed to be characterized. In this context, it has to be kept in mind that the latter of the two methods for case (b) bases the calculation of K_i*^{app} on the result for K_i^{app}, while the first one estimates it separately.

Lastly, the necessary equations to determine K_I^{app}, k_{inact}, and k_{2nd} for case (c) are depicted below (eq. 6, 16, and 17). Most equations relevant here were in principle already described above. To repeat, the difference in eq. 6 to eq. 4 is the simplification due to $v_s = 0$ AU/s. There are two options to gauge the non-covalent affinity. Most commonly, eq. 16 is used to derive K_I^{app} from the observed inactivation kinetics. It assumes that the irreversible reaction is a bimolecular process that is primarily affected by the formation of the non-covalent complex. Therefore, the observed concentration dependency is most descriptive of the formation of the non-covalent complex. Another option is to observe initial velocity

the same way as for case (b) and estimate K_i^{app} through eq. 9 and 10.[268] It should be noted here, that the strict distinction between K_i and K_I for affinity parameters describing a non-covalent complex of an irreversible inhibitor is not always adhered to. This is not problematic if the inhibition mechanism and used calculations are clearly stated. The difference in eq. 16 to describe the kinetics of the off reaction in comparison to eq. 14 is the absence of an offset ($k_6 = 0\ s^{-1}$). So contrary to case (b), this plot crosses the coordinate origin (Figure 18). The maximal k_{obs} here is called k_{inact} instead of k_5, since it represents an irreversible reaction that inactivates the enzyme. For irreversible cases like this one, it is most useful to report both k_{inact} and K_I, since both reactivity and non-covalent binding can be optimized separately. An alternative is combining the two to shift the focus towards an effective inactivation rate of the enzyme. This combined value (eq. 17) of a first-order reaction rate (k_{inact}) with the inhibitor's K_I, is the second order constant k_{2nd}.[266] This value has comparable significance to the catalytic efficiency (k_{cat}/K_M) in Michaelis-Menten kinetics.[262]

$$F(t) = F_{background} + \frac{v_i}{k_{obs}} \times (1 - e^{-k_{obs} \times t}) \qquad \text{eq. 6}$$

$$k_{obs}([I]) = \frac{k_{inact} \times [I]}{K_I^{app} + [I]} \qquad \text{eq. 16}$$

$$k_{2nd} = \frac{k_{inact}}{K_I} \qquad \text{eq. 17}$$

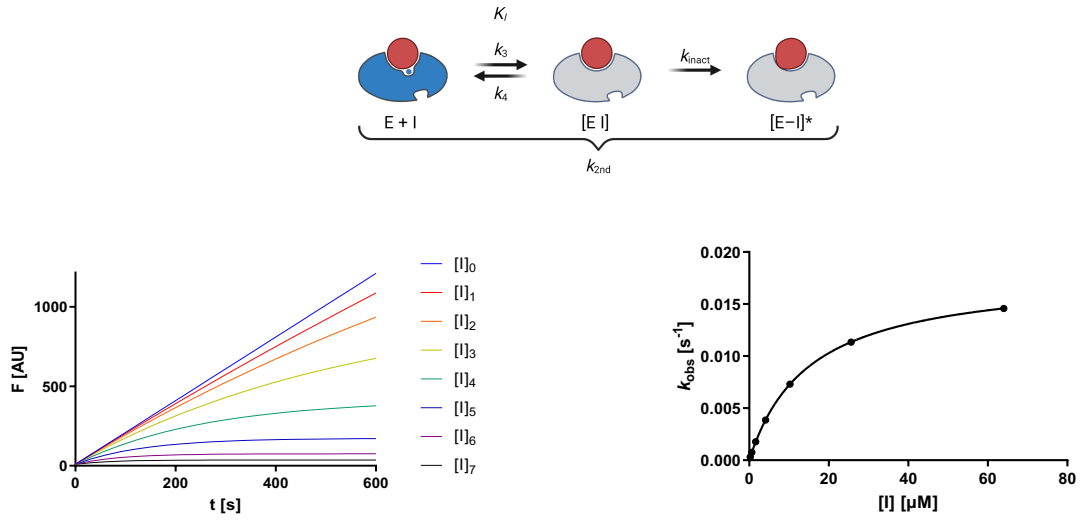


Figure 18: (Top) Repeated illustration of the binding mode of case (c). (Bottom left) Artificial dataset for an irreversible inhibitor using seven different inhibitor concentrations [I] and a negative control without inhibitor $[I]_0$. (Bottom right) Transformation of the artificial dataset with [I] and activity([I]) (black spheres) and fit using eq. 16 (black line) with the following parameters: $k_{inact} = 0.018\ s^{-1}$, $K_i^{app} = 15\ \mu M$. Figures created with Biorender.com and Graphpad Prism.

As shown by eq. 17, it is usually not adequate to use apparent values for (further) characterization. In the depicted case of calculating the k_{2nd}, the K_I representing the non-covalent affinity of the inhibitor in absence of any competing ligand is required. The conversion of apparent values to equilibrium constants can be performed mathematically using the Cheng-Prusoff equation (eq. 18).[271] It describes a dependency that relies on the concentration and K_M of the substrate towards the protease of interest.

$$K_I = \frac{K_I^{app}}{1+\frac{[S]}{K_M}} \qquad \text{eq. 18}$$

The same equation holds true for all apparent values mentioned above.[268] Since the denominator is always ≥1, the corrected values are always lower than the apparent values, because in absence of competition, inhibition is expected to be higher. The deviation is most impactful if the assay is performed at substrate concentrations [S] exceeding K_M. Assuming [S] = 9 µM and K_M = 1 µM for eq. 18 results in a factor 10 difference between K_I and K_I^{app}. The kinetic parameters (e.g., k_{inact}) however are only relevant for an already bound state in which competition is irrelevant and are thus unaffected by the substrate and do not require such correction.[268]

To conclude this chapter, it is important to be fully aware of an inhibitor's mode of inhibition to correctly and comparably describe its affinity and reaction kinetics. Just because an inhibitor carries a certain specific covalent warhead, this does not have to determine the observed results. While an irreversible inhibition always comprises a time-dependent progress, a high residence time is not always associated with a covalent adduct formation but can also be recorded for some purely non-covalent interactions (e.g., cystatin C binding to papain[274]). Also, a covalent mechanism can hide behind a time-independent reaction progress (e.g., nitrile/aldehyde binding to cathepsin K/L/S[139]). The investigation of binding affinity and kinetics is important for all publications 1–6 included in this thesis. Interesting distinctions deviating from the expectations are the difference in reversibility between α-fluorovinylsulfones and -sulfonates (publication 1), the parallel reactions that contribute to inhibition for the dual carbonyl/Michael-acceptor warheads (publication 2), the modifications towards a reversible S_NAr warheads (publication 3), as well as the distinctions between structurally similar α-substituted methyl ketones, aldehyde and the nitrile warheads (publication 6).

Pharmacokinetic Challenges of Peptidomimetic Inhibitors

As alluded to in the previous section on synthetic inhibitors carrying characteristics of their physiologic counterparts, pharmacokinetics can be a challenge. The focus here is placed on peptides, for which this limitation is especially true.[13,275–277] Pharmacokinetics in general are often summarized by ADME parameters (absorption, distribution, metabolism, elimination). As the first major physiological mechanism of a dissolved drug's life cycle during clinical application, absorption is relevant and will be a central element of investigation presented in this thesis. Other mechanisms related to distribution, metabolism, and elimination will not be central elements, even though their importance in drug design is undeniable. ADMET (T for toxicity) has increasingly found recognition since the 1990s as one of the most frequent reasons for failure of promising drug candidates.[278–280]

As a point of terminological concern, "permeability" is a term commonly borrowed to characterize small molecules. However, this is technically a misuse of the word since it is rather conversely a property of a barrier-like structure like a membrane or a matrix: cell membranes are permeable to a drug, a material layer is semi-permeable to molecules with regards to their hydrodynamic radius, Vaseline® is occluding because it is impermeable to water. But since it is established to speak of "permeable drugs", this usage will still be applied herein.

Following oral application, absorption can essentially take place in the entire gastrointestinal tract (GIT). There are vast differences between organ constitutions, cells making up the resorption surface, and surface area available for absorption between mouth, stomach, small and large intestine.[281,282] The organ most targeted for drug resorption is the small intestine, mainly due to its conveniently large internal surface area. Differences in cellular architecture between the three segments of the small intestine, as well as differences in their extracellular milieu (pH, water content, etc.) are simplified to one model herein (Figure 19) that depicts histologic properties of this organ schematically. The high surface area is achieved by macroscopic and microscopic folds of its luminal tissues and cell layers. An intestinal fold is covered in smaller structures called "villi" that each have their own vasculature to continuously provide sink conditions. Additionally, the epithelial cells have electron microscopic microvilli on their apical membrane further increasing the surface area.[283,284] For most drugs, the intestinal epithelium only allows transcellular absorption, meaning a small molecule must pass the membranes and body of an epithelial cell to reach the bloodstream.[285] In principle, there is also the possibility of paracellular absorption. However, this is strongly limited by tight junctions between these cells, making it only accessible to water and small hydrophilic molecules (<200 Da).[286] Still, this route is sometimes discussed as a possibility for drugs.[287] Like all cell membranes, the epithelial membrane is essentially comprised of a phospholipid bilayer, into which functional macromolecules are embedded, and through which transcellular transport needs to progress.[288–290] In principle, passive and active transport can be distinguished as mechanisms for transcellular absorption.[291] Passive transport relies on a molecule's propensity to distribute in and out of the lipid membrane, which is closely related to its lipophilicity. Active transport is a selective process, in which specific transporter proteins tunnel their ligands through

a membrane. These physiological transport mechanisms can be exploited by drugs, which will be discussed in more detail further below. Most orally applied drugs are designed to be sufficiently permeable to be passively absorbed, and do not have to rely on saturable active transport mechanisms.[292]

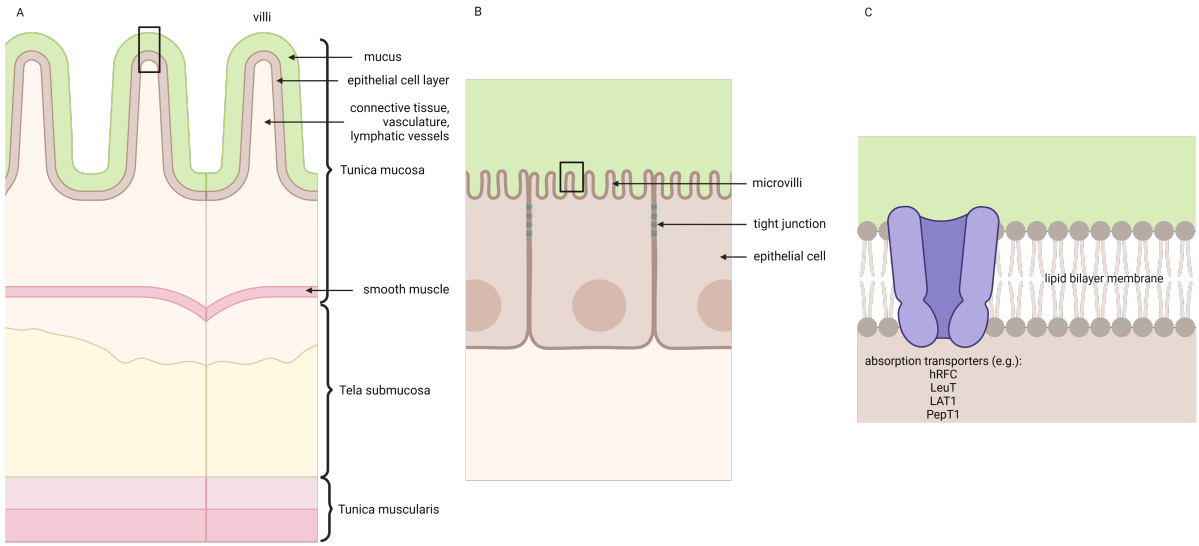


Figure 19: Schematic representation of the microscopic absorption surface of the small intestine. Adapted from Lit.[293,294] (A) The intestinal wall is principally divided between Tunica muscularis (innervated smooth muscles responsible for peristaltic transit consisting of outer longitudinally, and inner circularly contractive fibers), Tela submucosa (loose connective tissue, vasculature, and innervation), and Tunica mucosa (smooth muscles for contractive emptying of the villi's lymphatic system; connective tissue for flexibility which holds vasculature, lymphatic system, innervation, immune cells etc.; epithelial cells specialized for absorption), which is covered with a mucus layer and the intestinal contents. The villi extend ca. 1 mm from crypt to tip. (B) The epithelial cells are further covered with microscopic microvilli (ca. 1 µm), increasing the inner surface area of the small intestine to ca. 30 m^2.[283] Tight-junctions hinder paracellular transport, so most drugs rely on passive or active transcellular diffusion. (C) The epithelial cell membrane (simplified lipid composition of only phospholipids) can enable passive transport or enable active transport through some specialized carriers (e.g., human reduced folate carrier = hRFC, leucine transporter = LeuT, L-type amino acid transporter 1 = LAT1, and peptide transporter 1 = PepT1). Figure created with Biorender.com.

To design permeable molecules, different computed or experimental estimates of critical parameters can be used. The most commonly calculated properties that are related to permeability are logP[295] (distribution coefficient between an aqueous and *n*-octanol phase), $logD_{pH}$[296] (logP with regards to the pH of the aqueous phase and the pK_a value and protonation state of the characterized molecule), polar surface area[297] (PSA), and the number of hydrogen bond donor and acceptor moieties[298]. They can be further extended by flexibility parameters like the number of rotatable bonds or the fraction of sp^3-hybridized carbons.[299,300] These surrogate markers are useful to roughly classify drug candidates, famously established as the "Rule of Five" (Ro5) and its extended versions.[298,301–303] While lipophilicity predictors alone are limited to estimate the contribution of passive transcellular transport to oral bioavailability, they can be predictive to a certain degree.[304,305] In recent years, it is commonly appreciated that the Ro5 in its strict interpretation does not include all drugs suited for approval (including some peptide-based inhibitors).[302] Some have sufficient ADME properties despite violating

the initially postulated rules. To better describe the chemical spaces that also accommodate these structures, versions of the Ro5 were defined, which are compared in Table 2. In the study defining these chemical spaces, peptidic drugs and clinical candidates were evaluated, but the results seem biased by the presence of cyclosporins in this subset. Of protease inhibitors, only HCV and HIV protease inhibitors were used. This should be kept in mind when judging the suitability of peptides.

Table 2: Physicochemical parameters defining chemical spaces according to the Ro5, extended Ro5, and beyond Ro5.

	Ro5[298]	**eRo5**[302,303]	**bRo5**[302,303]
Prediction of oral absorption	likely	possible	n.s.
Molecular weight [Da]	≤500	500–700	700–3000
logP	≤5	0–7.5	<0 or >7.5
Hydrogen bond donor count	≤5	≤5	>5
Hydrogen bond acceptor count	≤10	≤10	>10
Polar surface area [Å^2]	≤140*	≤200	>200
Rotatable bond count	10–20*	≤20	>20

e/bRo5 = extended / beyond "rule of five", n.s. = not specified, * not included in the original Ro5, but extensions that were made prior to the formal definition of the eRo5 space.[302]

An experimental approach aimed to better mimic biological membranes is called parallel artificial membrane permeation assay (PAMPA).[306–309] It reduces the complex membrane constitution to a simpler hydrophobic barrier. Supported by a porous polymer (e.g., polyvinylidene difluoride), a mixture of phospholipids (e.g., phosphatidyl choline) and an apolar solvent (e.g., *n*-dodecane) are immobilized between two aqueous compartments, of which one contains the substance of interest (Figure 20). The phospholipids are assumed to line up with their hydrophilic portion towards the neighboring aqueous phases, and their lipophilic portion immersed in the solvent bulk, separating it completely. This membrane is a lot thicker than cell membranes, and the organic solvent is an unphysiological constituent. But the processes for permeation (namely diffusion and transport by inverse micelles) are assumed to be predictive of the physiological processes (Figure 20). The membrane composition can be varied for this model.[310] This includes using pure *n*-hexadecane without amphiphilic additives[311], an alkane-solution of defined phospholipids (e.g., dioleylphosphatidyl choline)[312], purified lecithins (from egg or soy)[313], or a complex mixture of pure components (e.g., phosphatidylcholine, -ethanol, -serine, -inositol, cholesterol)[314] in an effort to imitate a specific membrane like the epithelium of the blood-brain barrier that carries a pronounced negative surface charge and a characteristic triglyceride-content.[313,315] The used bulk hydrocarbon can also differ.[314]

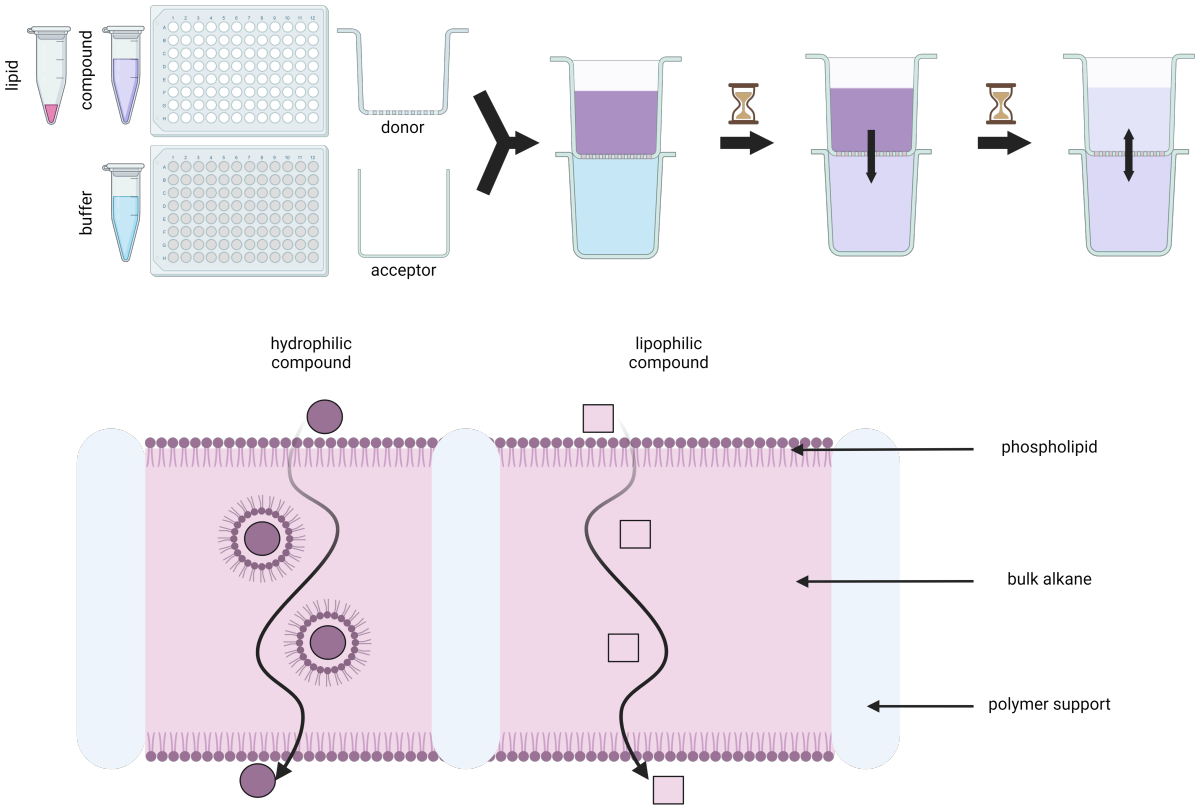


Figure 20: Top: Schematic depiction of the experimental setup required to perform PAMPA. The donor compartment is prepared by adding the lipid mixture onto the polymer support. Then, donor and acceptor compartments are filled with solutions of differing analyte concentration and the setup is assembled to equilibrate. Prior to full equilibration, the assay is read out (e.g. LCMS, spectrophotometry). Bottom: Assumed mass transport mechanisms for hydrophilic and lipophilic molecules over the artificial membrane, adapted from Lit.[316] While lipophilic molecules can permeate unassisted, hydrophilic molecules are assumed to utilize inverse micelles for crossing. Figure created with Biorender.com.

The permeation rate P expresses the rate at which the concentration of a substance of interest equilibrates between two compartments along a concentration gradient. In general, this relationship is derived from the first Fick's law on diffusion stating a proportionality of the equilibration rate to the concentration gradient.[317] In membrane-limited diffusion however, this rate is also dependent on system characteristics like compartment contact area, membrane thickness, and solute lipophilicity.[318,319] Different equations exist in literature to correlate these with permeability.[311,319–321] Of these, eq. 19[311] is arguably the most tangible to assess the apparent permeation rate ("permeability", P_{app}/P_a/P_{am}), even though some simplifying assumptions are made that will be discussed further below. The immediate result for P_{app} is commonly reported in units of [$*10^{-6}$ cm/s] due to the dimensions of the input variables. For comparison in a set of low-permeability molecules, [nm/s] (= [$*10^{-7}$ cm/s]) can be found.[322]

$$P_{app} = -\frac{V_d \times V_a \times \ln\left(1 - \frac{c_{a,t}}{c_{a,\infty}}\right)}{(V_d + V_a) \times A \times t} \qquad \text{eq. 19}$$

In this eq. 19, P_{app} is calculated from an endpoint assay at a specified time (t [s]). The aqueous donor and acceptor volumes (V_d and V_a [cm^3]) that are equilibrated, as well as the contact area between them (A [cm^2]) are fixed parameters. The concentration in the acceptor compartment at the endpoint ($c_{a,t}$) and

after infinite equilibration time ($c_{a,\infty}$) are experimentally determined. Generally, there are multiple ways to measure these concentrations, usually depending on molecule characteristics. If it shows appreciable absorption or fluorescence at a specific wavelength, spectroscopy can be used. If not, mass spectrometric quantification can be employed.

There are further extensions to this equation, or conversely, simplifying assumptions. Apparent permeabilities, in analogy to apparent inhibition constants described in the previous chapter, are not necessarily comparable between methods. They are influenced by the following factors.

a) After incubation, a fraction of the molecule of interest can be inaccessible to the analysis of the aqueous compartments. This can be due to distribution into the lipid layer, but not out of it, which is called mass retention or membrane retention (M). The mass retention for PAMPA experiments is commonly expected to be low[307,322] but can be relevant for very lipophilic molecules, because it is a factor causing an underestimation of permeability.[323] Also it becomes more influential with higher phospholipid content (>10 %) of the employed artificial membrane.[313]
b) Another parameter is the unstirred water layer (UWL) effect, which is dependent on convectional flow within V_d and V_a. It is a confounding factor towards predictability of bioavailability, because this layer is absent in vivo due to sufficient agitation of intestinal contents.[324] It is of special concern to correctly rank high-permeability molecules ($P_{app} > 20 * 10^{-6}$ cm/s), but less important for general classifications or low-permeability molecules.[325]
c) Necessary correction is required of the filter area for its nominal filter porosity ε, because only the porous part is filled with lipid and accessible for mass transport; the rest is solid polymer.[326] The effective area A separating V_d and V_a is therefore smaller than their circular contact area by a factor of ε. It is a property of the employed filter type that should be included in the manufacturer information.[327] With the same argument, the filter already presents a diffusional barrier (P_f) even in absence of artificial membrane, with its porosity and thickness influencing diffusion.[325] It can be assessed by a mock experiment with methanol instead of lipid membrane.[306]
d) It is commonly assumed that only the uncharged state of a (de)protonatable molecule is capable of permeation through strongly lipophilic membranes, so pH is a modifying parameter of the system.[328,329] This is certainly only relevant for molecules that undergo relevant (de)protonation over a physiological pH range.

This results in increasingly complex variations of the experimental corrections required and the underlying calculations. It is therefore a valid approach to include references of known passive permeability into the assessed group instead of employing fully optimized setups. Common high-permeability drugs (P_{app} >10^{-5} cm/s) are verapamil or propranolol.[330] An intermediate class is not necessarily defined but can be quite variable.[331,332] For low-permeability molecules, a distinction between low permeability (e.g., 10^{-7}–10^{-6} cm/s)[332–334] and definitive impermeability is always dependent on the lower quantification limit of the method (e.g., 10^{-9} cm/s)[335]. Drug examples with poor permeability in PAMPA are furosemide[331], methotrexate[330], and theophylline[330]. Here, it is important to note that these references are only for passive permeability as assessed by PAMPA. Active transport that contributes to the permeability in more complex systems as in the case of methotrexate is not assessed.[336]

To sum up these distinctions, a molecule's recorded permeability is always dependent on the permeation experiment. Extents of influencing factors are accessible with the right experimental setups. The required correction experiments characterize the membrane permeability (P_m) being the intrinsic permeability (P_o) corrected for the pH-dependent fraction unionized (f_{un}), the filter permeability (P_f), the permeability of the unstirred water layer (P_u/P_{UWL}/P_{ul}), as well as effects like membrane retention (M). The effective permeability (P_e) can either be defined as the apparent permeability in absence of membrane retention[311], or as the membrane permeability corrected for the UWL effect and protonation state[319]. In commercial applications, P_e can be found equated to P_{app}.[337] This simplification was also made for the publications below, which is reasoned by the factors mentioned above. A distinction would be necessary if values far exceeded 20 *10^{-6} cm/s, as this would introduce problems like membrane retention and UWL influence. This range was not observed for the novel inhibitors tested. Furthermore, protonation could be neglected for most of them as they were either not pH-responsive, or one dominant species was always present in solution. Given a sufficient method description and appropriate references of known permeation behavior, apparent permeabilities are therefore sufficient for comparisons, especially in the context of the inclusion of this method in early drug design. There, it is not crucial to have optimal in vitro-in vivo correlation, but to favor throughput. An automated instrument for high throughput that eliminates UWL effects through efficient stirring inside each well exists, that can reduce assay times to as low as 15 min (e.g., Gutbox®, Pion Inc.).[313,338]

Advancing from the strict focus on passive permeation, cellular model membranes capable of active transport can be employed to apply the same equilibration principles as described above for PAMPA. Example cell types are Cancer coli 2 (Caco-2 cells) and Madin-Darby canine kidney (MDCK cells).[339,340] Both can be cultured to fluence on a porous polymer support[340], and the integrity of the continuous cell-layer can either be checked by electrical resistance measurements[341] prior to transport studies or by integrating an impermeable dye[342] into the measurement set. A predictive advantage of these cell-based models in terms of in vitro-in vivo correlation is that multiple parameters can be assessed coincidentally, namely overall (active and passive) permeability and preliminary metabolic stability.[343] For an investigated probe that does not pass these models, this can complicate the structural optimization

decisions required, because it could be passively impermeable, metabolically unstable, or be a substrate of efflux transporters. But a candidate that passes these models is promising. Overall, cell-based permeability assays are suitable in vitro methods for the assessment of bioavailability. They are accepted for permeability assessment of drug substances for the Biopharmaceutics Classification System (BCS), which is important for regulatory decisions like biowaiver approvals in the context of generic drugs.[342] An optimal assessment of bioavailability is of course only accessible through in vivo studies in humans, but this is inapplicable for early-stage development due to ethical concerns, high cost, low throughput, etc. In contrast, there are no ethical concerns for in vitro methods like PAMPA or the Caco-2 assay. Required pipetting time is in the same order of magnitude for both types of permeability assays, but the latter requires additional days of cell culture that limit flexibility.[339,344] Parallel throughput is classically higher for PAMPA (≥96-well plates) than for Caco-2 (≥24-well plates). In terms of efficiency, PAMPA therefore represents a suitable method for the prediction of passive permeability in an early phase of drug development.

Some specific limitations to permeability for peptide-based inhibitors should be kept in mind. For large molecules (>1000 Da), passive permeation is not an option.[345] This size criterion is usually mended in keeping the peptide recognition sequence of these protease inhibitors short. Despite endogenous inhibitors like cystatins (ca. 120 amino acids long) being highly affine for relevant proteases[346], similar structures are not suitable for drug design, because they do not adhere to ADME criteria. Most approved peptide-based protease inhibitors that are orally bioavailable are tripeptides at most (e.g., saxagliptin, nirmatrelvir, boceprevir, bortezomib). Those that are longer can require a parenteral route of application (e.g., carfilzomib). Adhering to this criterion simultaneously limits the number of polar peptide bonds, which keeps the peptide sufficiently lipophilic.[347] The flexibility as well as effective hydrogen bond donor and acceptor count can be reduced by options such as cyclization exemplified by modern HCV inhibitors or cyclic peptides in general.[348,349]

In general drug design, permanent charge as shown by quaternary amines like hyoscine-*N*-butyl bromide, tiotropium bromide, or methylnaltrexone bromide is commonly regarded as an absolute negative determinant of oral bioavailability. It is therefore expected that transient charge carried by pH-responsive groups also presents a challenge for permeability. But protonation states can be modulated by substructures affecting the pK_a values through chemical modification.[350] In the context of peptides, charged termini and side-chains can be permanently capped or transiently modified by a prodrug approach. Amines can be turned into carbamates, ureas, and amides, which is often applied *N*-terminally as this also reduces backbone hydrophilicity.[351] If an arginine (Arg) mimetic is required, the guanidine can be exchanged for a hydroxy amidine (= amidoxime, e.g., ximelagatran). Free carboxylic acid groups are commonly detrimental to permeability, which is why they are usually capped as esters (e.g., ramipril) or amides (e.g., saquinavir), or exchanged for a more lipophilic group altogether (e.g., nirmatrelvir). It is also uncommon to see polar amino acids such as Lys, Ser, Cys, or Glu. Instead, Phe, Leu, Pro, etc. are used if possible.[347,351] If a polar moiety is necessary, it is often exchanged for a more lipophilic

peptidomimetic option (e.g., the cyclic Gln mimetic in nirmatrelvir), or as mentioned above, modified by the prodrug approach (less basic hydroxyamidine in ximelagatran). There are however famous examples of peptide-derived zwitterions, e.g., in lisinopril or cefadroxil that do not adhere to the abovementioned permeability rules on pH-responsiveness. Despite them carrying multiple charges at intestinal pH, they are bioavailable, which can be understood in the context of active absorption by peptide transporters.[352]

Active transport is a physiological mechanism that contributes to the uptake and distribution of important molecules or the elimination of toxic molecules. Both processes for which it is worth expending energy, either primarily or secondarily driven by an adenosine triphosphate (ATP)-dependent concentration gradient. Important molecules can be nutrients like glucose, vitamins like folic acid, or amino acids. Contrarily, molecules to dispose of are often metabolic products like uric acid or different xenobiotics. In the context of oral bioavailability, only those transporters present on the luminal side of the intestinal epithelium, and only influx transporters are considered here. For drugs already resembling essential α-amino acids, amino acid transporters (e.g., LAT1) are a positive contributor to their permeability; examples are L-DOPA and melphalan.[353] The same is true for methotrexate being structurally close to folate and therefore transported by hRFC.[354] Then, there are prodrugs carrying an amino acid like midodrine, lisdexamfetamine, or valaciclovir and valganciclovir, which can exploit peptide transporters like PepT1 for absorption.[355] This pathway is also available to peptide-derived drugs like β-lactam antibiotics[356], captopril[357], lisinopril[358] and enalapril[357] with free and therefore charged termini.

The peptide transporter most relevant to active drug absorption (PepT1) is classified as part of the major facilitator superfamily of carrier proteins.[352,359] More specifically, it is part of the solute carrier (SLC) family 15 as SLC15A1. It is a H^+-peptide cotransporter that uses an extracellular high proton concentration to transport various di- and tripeptides across cell membranes, albeit with low affinities (K_M in the micromolar to millimolar range).[360,361] This process is secondarily active due to the ATP-dependence of the proton export (e.g., by Na^+-H^+ antiport in combination with Na^+-K^+-ATPase).[362] PepT1 is localized mainly on the apical membrane of intestinal epithelial cells[363], as opposed to its paralog PepT2 which is rather found in kidney epithelia.[364] Both transporters have similar structures and transport mechanisms, but due to their expression patterns, they are involved in different pharmacokinetic processes.[359] While PepT1 is mainly involved in intestinal absorption, PepT2 is relevant for distribution and elimination like clearance from cerebrospinal fluid and renal reabsorption. Their substrate specificity is broad when it comes to side-chain-modified di- and tripeptides, but they do not transport most single amino acids or linear tetrapeptides.[352] Modifications of side-chains or terminal capping in di- and tripeptides, as well as D-configuration also generally reduce uptake.[361,365] Predictions of affinity based on pharmacophore modelling have been established.[352] However, applying these to generate prodrugs that are suitable PepT1 binders does not necessarily result in actual transport and bioavailability.[366,367] The optimal substrates for PepT1 are close mimics of natural oligopeptides,

comprised of amino acids in L-configuration, including many of their charges. One criterion for affinity is *C*- and *N*-terminal charge to occupy the positively and negatively charged subpockets, respectively.[359,368] These criteria contradict the paramount design choices for peptidomimetic drugs, especially in the context of protease inhibition. The backbone of protease inhibitors is often comprised of non-proteinogenic amino acids that convey hydrophobic binding surface and/or metabolic stability. The *C*-terminus is often transformed to an uncharged warhead, and the *N*-terminus is often capped due to synthetic reasons or to aid in target engagement. Also, as described above, charge is detrimental to passive permeability, meaning one has to sacrifice one mode of permeation for the other. So, while PepT1 can be relevant for drugs derived from amino acids, it is not the prime optimization target for peptidomimetic protease inhibitors. It might be an option for peptide ligands for other targets that have an inherently low permeability to begin with.

The investigation of passive permeability was performed for various peptide and peptidomimetic inhibitors with diverse sets of warheads in the context of publications 3–6. The implications on the relevance of prodrug structures not only for oral bioavailability but also for pronounced effects in cellular assays was elucidated in the context of charged moieties in publications 3 and 4. For cellular studies of irreversible inhibitors, it was interesting to observe that low permeability and low non-covalent affinity can be compensated in some cases by an irreversible mode of action as described in publication 5. Lastly, publication 6 provides extensive structure permeability relationships for commonly encountered elements in the design of peptidomimetic protease inhibitors and through that generates value for general drug design.

Disease Contexts for the Main Investigated Targets

Classification of Mammalian Cathepsins and Their Pathophysiological Functions

Cathepsins (Greek *kathepsein* = digest) are specialized eukaryotic proteases first recognized for their ability to degrade exogenic proteins, and for their involvement in the physiological turnover of the cellular proteome. Originally, they were characterized as a group of intracellular proteolytic enzymes active at acidic pH.[369] They perform this task compartmentalized inside the endo-lysosomal system under the reducing[370,371] and acidic[372,373] conditions integral to their activity. All are identified by a seemingly arbitrary one-letter code not coherently derived from order of discovery, function, or relation, the abbreviation being "Cat" plus capital identification letter. Cathepsins A and G are Ser proteases, D and E are Asp proteases, and B, C, F, H, K, L, O, S, V, W, X are Cys proteases. Redundant nomenclature is clarified in Table 3.[374–376] The "missing" letters in this list denote further Cys proteases that share the name but not the role of other cathepsins, like CatT (only one specific substrate)[377,378], CatJ, M, Q, and R (placental proteases in rodents, active at neutral pH)[379], or CatN (limited scientific coverage)[380,381]. Cathepsins are mostly endopeptidases, but CatB and H can catalyze both endo- and exopeptidase reactions. Only CatA, C and X are strict exopeptidases.[369,376] In addition to compartmentalization, the activity of cathepsins is controlled by their expression as inactive zymogens, more specifically as preprocathepsins.[382] After translation of the primary sequence in the endoplasmatic reticulum (ER), the *N*-terminal prepeptide is cleaved in parallel with glycosylation and phosphorylation: a mannose-6-phosphate-containing signal saccharide is unveiled during passage of the Golgi system that generally mediates lysosomal targeting of cellular proteins.[383,384] The *N*-terminal propeptide aids in folding of the enzyme into its functional form and inhibits the catalytic activity while covalently attached. The procathepsin enters the endo-lysosomal system where the propeptide is cleaved off to reveal the functional protease (either inter- or intramolecularly in an autocatalytic fashion[385], or by other already active cathepsins inside the lysosome[386]). If extracellular functions need to be exerted, secretory lysosomes can be triggered to release lysosomal contents into the extracellular space.[383,387–389] In their lysosomal function, cathepsins work in unison and can be redundant in some cases due to overlapping substrate specificity. However, many specific intra- and extracellular tasks have been attributed to cathepsins, making them relevant to a variety of (patho-)physiological contexts.[390] A special impact is given to their functions in modulating immunologic functions[391], and their ability to process extracellular structural proteins.[388] An overview of the involvement of cathepsins in different disease contexts, and the most relevant cathepsin inhibitors in clinical trials are listed in Table 3. It should be emphasized here, that explicitly only human cathepsins are listed. From this, it is evident as well as surprising that despite the obvious involvement of cathepsins in cancer, there have not yet been relevant clinical trials for cathepsin inhibitors in this direction.

Table 3: Collection of cathepsins with identification code, the protease class they belong to, and cleavage site preference as exo- or endopeptidases. (Patho-)Physiological involvement of the different cathepsins is briefly stated, and drugs in clinical trials for these diseases are named incl. the most advanced clinical trial in parentheses.[376,390,392,393]

Cathepsin identification letter (synonyms)	Protease class aspartate (A), cysteine (C), serine (S) (endopeptidase/exopeptidase/mixed)	(Patho-)Physiological involvement Clinical candidates (phase of trial, disease)
A	S (exo)	Vasoregulation, galactosialidosis[394] SAR164653 (I, heart failure)[395]
B (B1)	C (mixed)	Cancer, Morbus Alzheimer, liver, pancreas, and thyroid physiology, Ebola VBY-376 (I, chronic hepatitis C)
C (J, P, DPP I)	C (exo)	Inflammatory diseases, cancer, genetic disorders[396] BI 1291583 (II, bronchiectasis)[397], Brensocatib (III, non-cystic fibrosis bronchiectasis; III Covid-19)[398,399]
D	A (endo)	Non-alcoholic steatohepatitis, cancer, neurodegenerative diseases[400,401]
E	A (endo)	Immunology
F	C (endo)	Neuromuscular development, vascular diseases
G	S (endo)	Immunology, bronchiectasis[397]
H (B3, I)	C (mixed)	Lung development, vascular diseases, immunology
K (O, O2, X)	C (endo)	Osteoporosis, bone-infiltrating cancer[402] Odanacatib (III, Osteoporosis)[403], ONO-5334 (II; Osteoporosis/Osteopenia)[404], MIV-711 (II, Osteoarthritis)[405], Balicatib (II, Osteoporosis/Osteopenia)[406], Relacatib (I, Drug-Drug-Interaction)[407]
L	C (endo)	Cancer, cardiovascular diseases, Covid-19, Ebola
O	C (endo)	Cancer
S	C (endo)	Cancer, Autoimmune diseases[408] Petesicatib (II, Sjogren's Syndrome), VBY-036 (I, neuropathic pain), VBY-891 (I, psoriasis), LY3000328 (I, abdominal aortic aneurysm), SAR114137 (I, pain), RWJ-445380 (II, psoriasis; II, rheumatoid arthritis), CRA-028129 (I, psoriasis)
V (L2, U)	C (endo)	Cancer, vascular diseases[82]
W	C(endo)	Immunology[409,410]
X (B2, P, Y, Z)	C (exo)	Immunology

Covid-19 = coronavirus disease 2019, VBY = Virobay, SAR = Sanofi Aventis, BI = Boehringer Ingelheim, ONO = Ono Pharmaceutical, MIV = Medivir, LY = Eli Lilly, CRA = Celera

The Microenvironment of Solid Tumors

Cancer is recognized as the second most frequent cause of death in countries that are part of the Organisation for Economic Cooperation and Development (OECD), accounting for ca. 20 % of all deaths.[411] However, in terms of treatment, it is often ineffective to subsume the various types of malignant proliferative diseases under the umbrella term "cancer". Since different types of cancer derive from different cell types, their phenotypes, behaviors, and susceptibilities can vastly differ. It is a reason for largely differing symptoms and disease prognosis, meaning their therapy requires adaptation to the specific pathophysiological processes involved. A more detailed analysis shows that solid tumors account for most of the cancer-related deaths.[411] Comparing fractions of cancer-related deaths by organ, lung cancers represent 20 %, and gastrointestinal cancers combined (colorectal, gastric, pancreatic, hepatic) 30 %. Gender-specific tumors located in prostate (10 % of men cancer mortality) and ovaries/cervix (combined 7 % of women cancer mortality), and gender-biased tumors like breast cancer (15 % of women cancer mortality) also account for relevant fractions of cancer-related deaths. In contrast, "non-solid" tumors subsumed as leukemia account for ca. 3 % for men and women.[411] Apart from mortality, survival rates defined as the fraction of patients alive after a defined period of time as a measure of disease severity underline some solid tumors as particularly lethal. Devastatingly low 5-year survival rates of lung (ca. 20 %) and pancreas (ca. 10 %) cancer underline their drastic nature. Of these two, lung cancer mortality is not only driven by lethality but also by incidence. In contrast for hematopoietic tumors, survival rates of 50–90 % are reported for the different types of leukemia, which has drastically improved since the introduction of kinase inhibitors.[412]

At the advent of modern medicine, cancer treatment focused on excessive growth as a prime abnormal characteristic.[413] It was discovered early on that cancerous cells divide with much greater frequency than most of their healthy counterparts.[414] They achieve this with expense of process control, meaning introduced genomic damages are not repaired sufficiently and can result in cell death.[415] Radiotherapy as well as chemotherapeutics like 5-fluorouracil, cisplatin, doxorubicin, bleomycin, etoposide, and cyclophosphamide are therefore based on the principle of damaging DNA, just to name some prime examples of the respective classes.[416] In the same direction go drugs like paclitaxel that directly target mitosis.[417] More specific approaches use targeting strategies aimed at targets characteristic for certain cell types for therapy. Estrogen receptor (ER) in breast cancer and prostate-specific membrane antigen (PSA) in prostate cancer are relevant markers exploited for therapy with SERMs or antibody-conjugated radionuclides, respectively.[418,419]

A unifying feature of solid cancers is their interaction with their surroundings. They are embedded in a tumor microenvironment (TME, Figure 21) with features distinct from healthy tissues. The tumor dogma of unrestricted growth must be sustained by nutrition, which rapidly grows scarce in evolving tumors.[420] Nutrition in this context is often boiled down to oxygen and glucose as the prerequisite of physiological cellular metabolism.[421] In hypoxic tumor tissues, a transcription factor (hypoxia-induced factor I, HIF-1) is active, which promotes angiogenesis by the expression of vascular endothelial growth

factor (VEGF) and glycolysis by the expression of glucose transporters (GLUTs), monocarboxylate transporters (MCTs), and glycolysis-involved enzymes.[422] In this context, boosting glycolysis gives an alternative source of ATP in absence of oxygen and therefore dysfunctional mitochondrial glucose oxidation. The main product of glycolysis is pyruvate, which is converted to lactate that can be H^+-cotransported into the extracellular space by MCTs, resulting in an acidification of the TME.[423]

Another important characteristic of the TME is the presence of infiltrating immune cells (Figure 21), which arrive to have tumor-suppressive functions but can be reprogrammed to have tumor-progressive effects through a cytokine interplay between each other and the tumor cells.[424] The tumor environment is also known for a diverse web of proteases at play which have various functions.[425] This includes remodeling of tissue to accommodate tumor growth and enable metastasis, modulation of angiogenesis to optimize tumor vascularization, and modulation of the cell surface to impact cell communication.[426–428] All of these mechanisms are progressed by different proteases with intra- or extracellular functions. Interestingly, the proteases involved are not necessarily expressed and secreted by the tumor cells, but also by tumor-associated cells like stromal or immune cells.[425,426] The interplay of upregulated and secreted proteases in the context of extracellular matrix (ECM) remodeling and cellular communication comprises different protease classes that process matrix proteins but also regulate each other's activity. Matrix metalloproteases (MMPs) have been clinically investigated in this context already, but the drug candidates showed disappointing results as they lacked either selectivity or oral bioavailability.[429,430] There are also some Ser proteases involved (e.g., uPA), and Cys proteases (mainly cathepsins B, K, L, S, and X) were identified as a relevant group.[425,431]

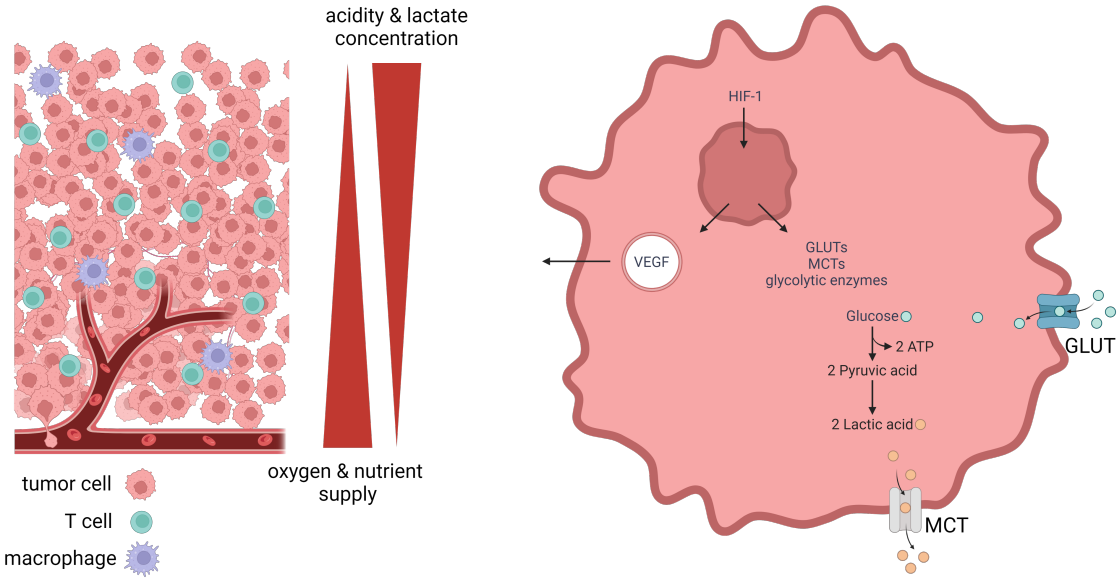


Figure 21: The TME is characterized by an excess of dysregulated cells. Besides the tumor cells, modified professional antigen-presenting cells (macrophages depicted as an example) and antigen-recognizing cells (T cells depicted as an example) contribute to it. Since vasculature is insufficient for the excessively growing tumor, oxygen supply is scarce. By HIF-1 activation, which is dependent on oxygen-depletion, glycolytic flux is increased, which results in an acidification of the TME. By the same mechanism, angiogenesis is induced by the expression and release of growth factors. Figure created with Biorender.com.

Cathepsin S as a Molecular Target for Cancer Therapy

Of the cathepsins involved in the tumor microenvironment, CatS represents a suitable target for therapy, as it is not as widely expressed as CatB or CatL, conveying a layer of selectivity for cell types that goes beyond structure-based drug design.[432] Also, CatS is not as limited to only one type of tissue as, for instance, CatK is in its restriction to bone. The confinement of CatS to antigen-presenting cells (APCs) is rather an advantage, as those cell types with immunologic implications have influence in a variety of solid tumors.

The expression of CatS in general is high in macrophages[433] and tumor-associated macrophages (TAMs) represent the major source of secreted CatS in the tumor microenvironment.[434,435] While CatS is not expressed by physiological epi- and endothelia, expression in those cell types is upregulated during a neoplastic state, which has been shown for solid cancers in several different organs.[433,436–438] These cell populations can also explicitly contribute to secreted CatS.[432,439] On a molecular level, CatS expression in non-macrophage cells is inducible by signal molecules like interferon-γ (INF-γ) or serotonin[440], and cathepsin expression is upregulated in TAMs through interleukin-4 (IL-4).[441,442] The INF-γ-induced expression can also be seen in radiation-induced tumors[443] and tumors under radiotherapy[444], which show radiation-dependent overexpression of CatS. The general involvement of reactive oxygen species (and ROS-generating drugs, i.e., bleomycin) in CatS overexpression is also appreciated[444], while the reports on the effects of taxol treatment is mixed in this regard.[444,445] Since some drugs can increase CatS expression, and CatS has tumor-promoting effects as discussed below, combination therapy approaches of classical chemotherapeutics with CatS inhibitors seems feasible.[446]

The malignant implication of CatS here is closely associated with antigen presentation. Its involvement in immunology has first been acknowledged in several autoimmune diseases including psoriasis, rheumatoid arthritis, Sjogren's syndrome, and celiac disease.[447,448] The principle of rewiring immunologic responses by overexpression or hyperactivity mutation (Y132D) of CatS is found for various cancers.[449–451] As mentioned before, many APCs carry CatS.[452] There, it promotes the presentation of antigens via the major histocompatibility complex II (MHC-II) pathway, and through that the immunologic response evoked by antigen-recognizing T cells.[453–455] The MHCs are an immunologic mechanism of APCs to give indication of intracellular processes on their surface. This mechanism is relevant during viral, bacterial, and parasitic infection, and to distinguish endogenous from "foreign" tissue, e.g., in graft transplants.[456,457] The main task of MHCs is to intracellularly bind a peptide and to extracellularly present it to T cells. Since this mechanism is also active in absence of pathology, it is regulated to avoid this process happening with host-derived peptides and the concurring negative reactions of T cells towards healthy cells. First, binding of host peptides is disfavored at the binding groove of MHCs by requiring a certain peptide length with specific motifs uncommon in host proteins.[457] Second, T cells are selected during maturation to not show excessive reactivity towards MHCs presenting host peptides.[458,459]

In general, two classes of MHC are distinguished (Figure 22): MHC I is responsible for presenting cytosolic peptide fragments to T cells positive for the surface marker cluster of differentiation 8 ($CD8^+$), while MHC II presents peptides from the endo-lysosomal (= endocytic) system to $CD4^+$ T cells.[457,460–462] MHC I is expressed by all cell types. The heterodimer, consisting of a so-called heavy chain anchored to the endoplasmatic reticulum membrane and a β_2 microglobulin, is assembled facing inwards the ER. Oligopeptides generated by the proteasome (and other proteases)[460] are shuttled into the ER lumen and loaded onto MHC-I.[463,464] The heavy chain which binds the peptides is encoded in three polymorphic genes. One of those, the human leukocyte antigen B (HLA-B) gene, is even described as the "most polymorphic gene in the human genome"[465], giving rise to a myriad of possible binding sites for peptide recognition. In its loaded state, the stable MHC-I complex is shuttled to the cell membrane via the endo-lysosomal pathway.[460] There it is recognized by the T cell receptor of $CD8^+$ T cells, which have cytotoxic functions upon activation. On the other hand, MHC II is not as widely distributed, being confined to professional APCs like dendritic cells, macrophages, and B cells in healthy tissue. However, induced by INF-γ, expression is also possible in other cell types in pathophysiological contexts.[460,466] The genetic code for MHC-II is again found in three polymorphic genes. In its initially translated, immature form, MHC-II is assembled in the ER and consists of three α/β heterodimers with their variable ligand binding motif. These three ligand-binding sites are always blocked with an invariant chain (Ii) trimer. The resulting nonamer is therefore not able to bind antigens but requires Ii processing first.[467] The correct function of Ii is crucial for functional MHC quaternary structure, transit to the cell membrane, and presentation of antigens.[457] The entire nonamer is trafficked to the Golgi apparatus through a signal sequence included in Ii. There, it enters the endocytic pathway, where the Ii trimer is cleaved (for example by CatS), which liberates the MHC II dimers that afterwards still carry a short remaining Ii fragment in their antigen binding site (class II-associated invariant chain peptide, CLIP). In a chaperone-mediated mechanism, CLIP is exchanged for an antigen fragment proteolytically generated in the lysosomes. The peptide fragment for loading can be acquired, for instance, by receptor-mediated uptake and degradation of an antibody-bound, foreign antigen.[468] The mature and loaded MHC II is transported to the cell surface for antigen presentation.[467] Via this route, $CD4^+$ T cells are recruited that can have various implications in cancer, depending on their subtype. Regulatory T cells (T_{reg}) are an especially interesting subgroup of $CD4^+$ T cells as they have immunosuppressive functions and therefore provide a favorable environment for tumor growth.[469] Through the promotion of endosomal Ii cleavage, CatS contributes to the complete degradation of Ii, supported by cytosolic proteases like γ-secretase.[470] In its degraded state, the intracellular domain (ICD, meaning the extra-endosomal, cytosolic domain) has an additional function. It is translocated to the nucleus where it activates the transcription of the gene for C-C motif chemokine ligand 2 (CCL2), which is secreted and attracts macrophages. It could be shown that its expression can be upregulated by CatS expression and downregulated by CatS-targeting RNA interference (RNAi) or a peptidomimetic nitrile inhibitor.[470,471] CatS and CCL2 are both negative prognostic markers for different solid cancers and there is an

association between the two that gives a mechanistic suggestion for the observation of an increased density of macrophages in the TME mediated by CatS.[470]

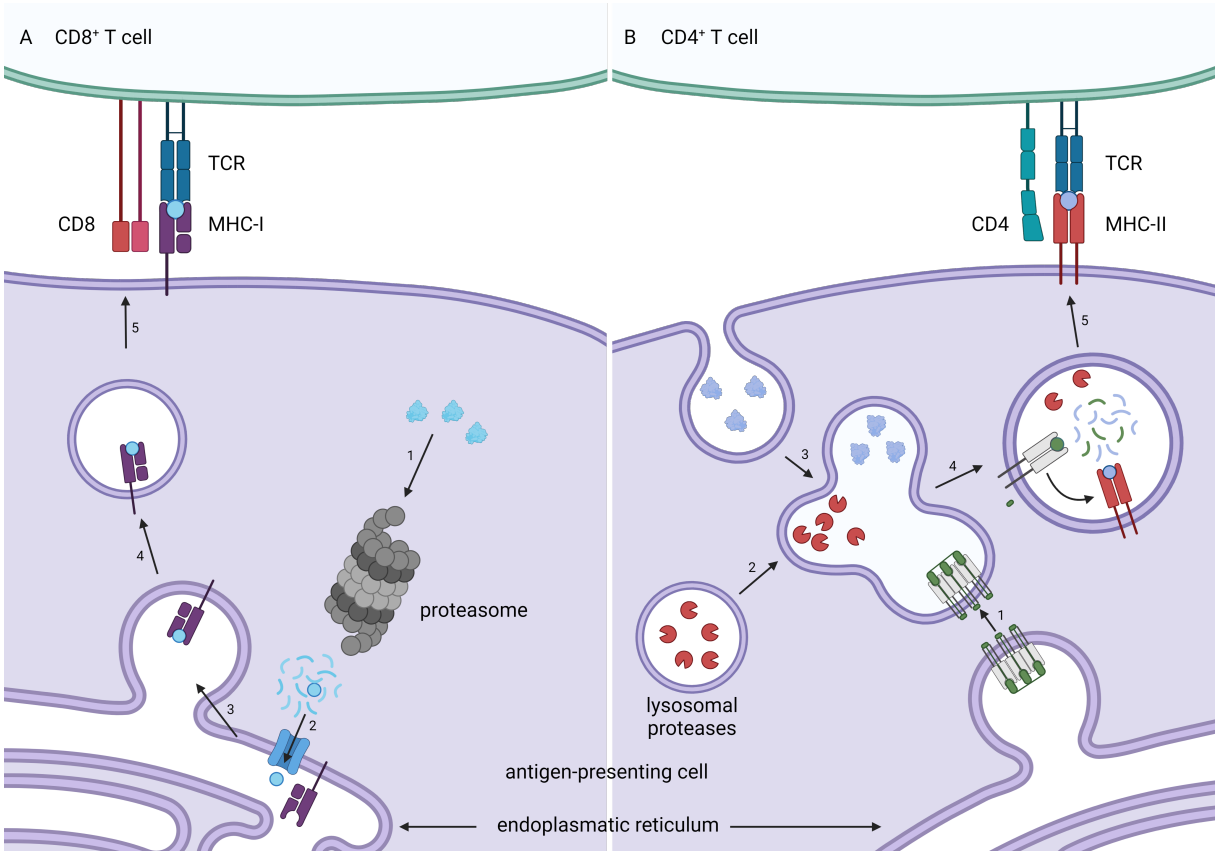


Figure 22: Juxtaposition of MHC pathways I (A) and II (B). (A) 1. Cytosolic (e.g., physiological, viral, or mutated cancerous) proteins are degraded by the proteasome to generate peptide fragments. 2. These fragments are imported into the ER and loaded onto a suitable MHC-I. 3.–5. Loaded MHC-I is transported to the cell membrane where it is recognized by the T cell receptor (TCR) of a $CD8^+$ T cell, which evokes a cytotoxic immunologic response. (B) 1. MHC II is assembled in the ER in an unfunctional (grey) multimeric state associated with Ii (green). For simplification, the short transient presentation of these inactive complexes on the cell surface is omitted.[472] 2. Cleavage of Ii to generate CLIP, the remaining inactivating peptide to MHC-II, changes MHC-II to its functional but still inactivated form. This process is promoted by lysosomal proteases like CatS, which is provided by a fusing lysosome. 3. Extracellular proteins are acquired by endocytosis, and the endosomes fuse to the previously mentioned compartment. 4. The exchange of inhibiting peptide (CLIP, green circle) for a proteolytically generated peptide fragment from the exogenous proteins (purple circle) is mediated in a controlled process. Cytosolic cleavage of Ii by γ-secretase liberates the intracellular domain which boosts transcription of CCL2, a macrophage-attractant. 5. The loaded MHC II complex is shuttled to the cell membrane to be presented to the TCR of $CD4^+$ T cells which promotes regulatory responses. Figure created with Biorender.com, adapted from the template "MHC Class I and II pathways" by Iwasaki and Lee.

There is more than this role of CatS in macrophages by influencing Ii processing during MHC-II maturation. As often in immunologic contexts "macrophage" is an umbrella term for a type of APC that can have several phenotypes, depending on the cytokines they are exposed to. A classically postulated dichotomy is between M1 (= classically activated) and M2 (= alternatively activated) macrophages. However, there are multiple activation paths apart from the INF-γ- and tumor necrosis factor (TNF)-mediated classical path. Stimulation with IL-4 can result in macrophages with tissue repair function, while IL-10 can promote an anti-inflammatory phenotype. The macrophages most common in

the TME of growing tumors best resemble this latter, regulatory type, or a mixed wound-healing/regulatory type. They produce high levels of IL-10 themselves, and thereby promote their own phenotype.[473,474] In general, M2 macrophages are associated with tumor progression by immune-suppression, promoting angiogenesis, and enabling growth around the tumor edges.[475] The TME itself also promotes M2 polarization through a lactate-mediated pathway, which is one of the main metabolic products of tumor cells.[476] A link between CatS and M2 polarization is assumed in its role for autophagy. It is a process of recycling cellular components like mitochondria, which can sustain cell survival in the nutrient-deficient central areas of tumor tissue.[477] CatS-deficient macrophages accumulate unprocessed autophagosomes, as CatS contributes to their fusion with lysosomes and the degradation of their content, which was shown to promote M2 polarization.[478,479] Lastly, CatS is also expressed in T_{reg} cells, another mediator of M2 polarization through its production of IL-10.[449,474] There, it influences the immunosuppressive potential of these cells that reduce a response by $CD8^+$ T cells. The favored development of T_{reg} cells out of other $CD4^+$ T cells is mediated by transforming growth factor β (TGFβ), which can be secreted by tumor tissue.[480] Other $CD4^+$ T cells can contribute to the expression of CatS in TAMs by supplying IL-4 among other interleukins.[481]

Besides these intracellular lysosomal functions, CatS has extracellular roles. This is special because cathepsins are usually assumed to be strongly adapted to the acidic environment of lysosomes, rather than the more neutral extracellular space. The extracellular role of CatS in cancer is enabled by its broad pH-dependent activity.[425,482,483] In comparison, CatB/L[484,485] with primarily lysosomal localization are quickly inactivated at physiological pH, while CatC[486] is more active at more neutral pH. Interestingly, parasitic orthologs of cathepsins also tend to retain their activity at less acidic pH, possibly enabling them to contribute to proteolysis in the cytosol of invaded host cells or the extracellular space.[487] The more acidic pH of the TME (5.5–7.0)[488] in comparison to the generally assumed physiological pH (7.4) contributes to extracellular CatS activity. Regarding TME pH range, a gradient can be assumed as indicated in Figure 21. The directly pericellular space of tumor cells is expected to be the most acidic due to their high glycolysis rate[489] and excretion of lactate (Warburg effect)[490]. The general tumor microenvironment is poorly vascularized, preventing full physiological equilibration, and is therefore still acidic.[491] Lastly, “normal”, unaffected tissue is properly vascularized and shows physiological pH. This adds a spatial layer of complexity to the interplay of different proteases involved in ECM remodeling that is linked to their pH-dependent activity profiles.

In its secreted form, CatS fulfills several functions. In proximity of cells, it acts as a sheddase together with CatL. Sheddases are a group of proteases which cleave off surface proteins from cells. The main specific targets of CatS in this function are receptors involved in kinase signaling and cell adhesion molecules.[492] Loosening of intercellular contacts is also mediated by cleavage of E-cadherin, resulting in enhanced cell mobility and through this, tumor expansion.[493] In a similar manner, CatS can mediate the invasion of brain tissue by degradation of a junctional adhesion molecule (JAM-B), which is a constituent of tight junctions in the blood-brain barrier.[437] CatS also boosts tumor expansion by

modulating vascularization. It degrades anti-angiogenic peptides like canstatin and arresten, and generates pro-apoptotic peptides in the process of degrading laminin-5, a common glycoprotein of the extracellular matrix.[494] However, some anti-angiogenic involvements are discussed, underlining its more balanced role in a healthy environment.[495] The interconnection between the extracellular proteases of the TME has been alluded to above. The role of CatS in this web is as a direct activator of CatC.[425] It has also been shown to be an indirect activator of neutrophil elastase in the context of cystic fibrosis, indicating a possible involvement with other proteases in pathological contexts.[496]

The involvement of cathepsin S has been acknowledged in various cancer types, e.g., in preclinical investigations on pancreas, breast, colorectal and bladder cancer, as well as glioblastoma.[408] In these studies, it could be shown that inhibition of CatS with different small molecule inhibitors can have good prognostic implications. In a cellular model of glioblastoma, covalent peptide inhibitors have been employed, namely a reversible α-keto aldehyde (Cbz-Phe-Leu-COCHO) and an irreversible vinylsulfone (morpholylurea-Leu-hPhe-VS-phenyl).[497] Different modes of action are discussed in this context. CatS inhibition indirectly suppresses kinase-dependent proliferative signals (e.g., activated by TGF-β) and indirectly activates kinase-signaling for autophagy. These effects seem to be promoted by ROS generation resulting from CatS inhibition and cell death is mediated by both autophagy and mitochondrial apoptosis. The TGF-β-related pathway mentioned before for proliferative signaling is also involved in tumor invasiveness by evoking changes in cell phenotype regarding intercellular connectivity (epithelial-to-mesenchymal transition). This TGF-β-mediated reduction in tight-junctions could be effectively counteracted by CatS inhibition.[498] This shows that the effect of CatS inhibition on junctional proteins is more than hampering their degradation in CatS's extracellular functions, but the inhibition also promotes cellular adhesion through indirect modulation of intracellular pathways. The discussed α-keto aldehyde inhibitor was successfully evaluated in a synergistic approach with oxaliplatin in a renal carcinoma model.[499] Investigations on pancreatic cancer with a peptidomimetic α-keto amide inhibitor (VBY-825) targeting CatB, L, and S (all K_i^{app} values in the picomolar range) showed a slight reduction in proliferating cells as well as a slight increase in apoptotic cells, and a slight reduction in invasiveness, together resulting in a reduced tumor growth rate.[500] The weak individual effect size was discussed to be due to drug distribution. It was reasoned that a direct intraperitoneal (ip.) injection that places the drug suspension in direct proximity to the pancreas could result in an increased effect in comparison to the chosen subcutaneous (sc.) route that requires more extensive permeation events to reach the target tissue. Similar investigations have been performed with the epoxide-based inhibitor JPM-OEt applied as an ip. injection, where the invasiveness reduction was also discussed in the context of vascularization.[435] For breast cancer, a correlation between CatS and the formation of brain metastases is published mediated by JAM-B processing as described above. This was successfully reduced using the peptidomimetic nitrile VBY-999.[501] It selectively targets CatS over related cathepsins (selectivity index > 3000) and was applied as a sc. nanoparticulate formulation.[437] From this application in a daily administration regime, effective concentrations could be shown in plasma and brain for this

inhibitor with picomolar K_i^{app}. Success in this field seems to be due to a combination of the need for high concentrations of active CatS required for extravasation, the exceptional affinity of VBY-999 towards CatS, and the sufficient tissue permeation properties of the inhibitor.

Taken together these investigations show that CatS is involved in a plethora of tumor-associated pathologic processes. While its direct effect on tumor cell survival can be limited, its involvement in the immunologic side of the tumor microenvironment, as well as expansiveness and metastatic tendencies of tumors are pronounced. A major advantage is that these features of CatS have been effectively blocked by several peptidomimetic inhibitors. However suitable modifications seem still necessary in terms of bioavailability, which currently is one limitation that hinders the transition of the inhibitors employed in preclinical models towards clinical application.

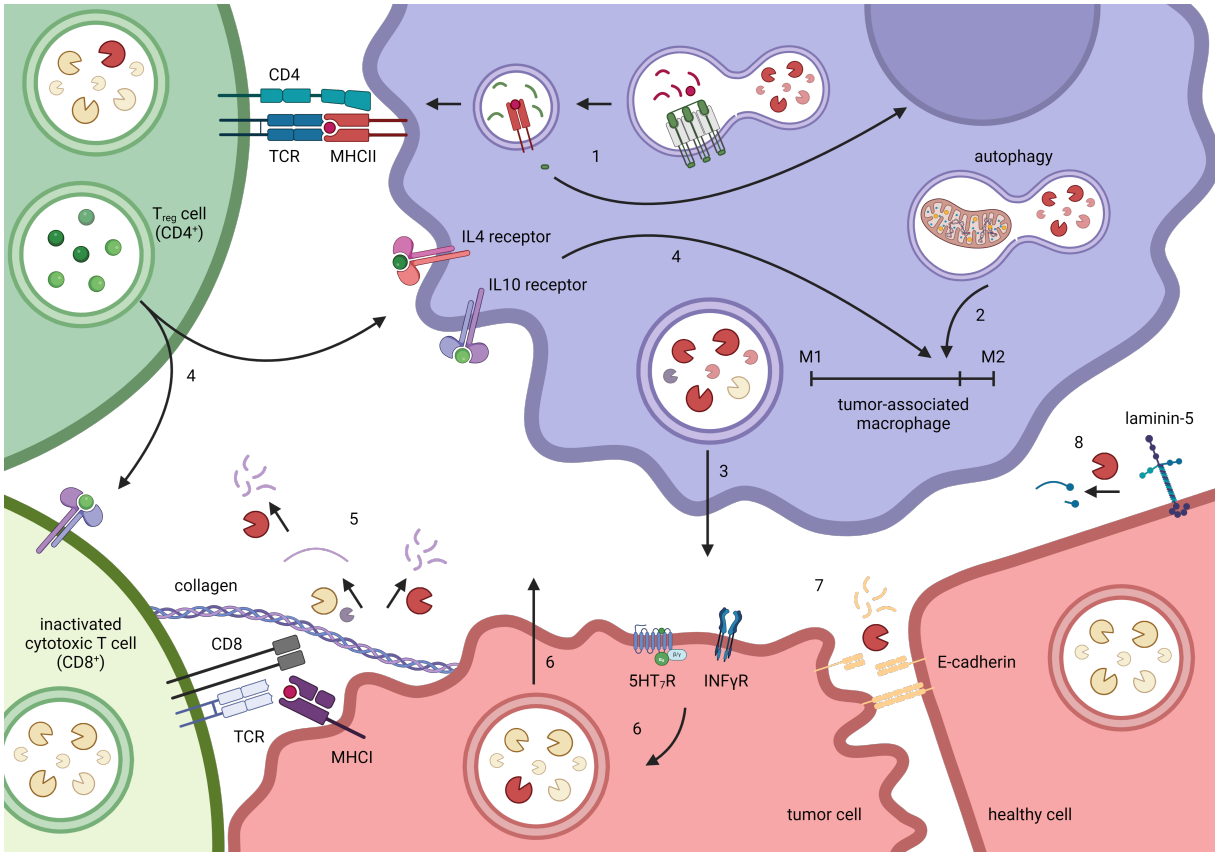


Figure 23: Involvement of CatS in solid tumors. Cells are depicted in green (T cells), purple (macrophage) or red (tissue cells). Cathepsin S is depicted in red, undefined other lysosomal proteases in yellow or purple. 1. Through Ii cleavage, CatS contributes to macrophage infiltration of the TME via induction of CCL2 transcription (in APCs in general, here depicted for a macrophage). It is involved in the maturation of MHC-II, a hallmark of M2 polarized macrophages. 2. Its involvement in autophagic processes promotes the M2 polarization type. 3. M2-type tumor-associated macrophages are the major contributor to extracellular CatS in the TME through secretion. 4. T_{reg} cells express CatS, and their main contribution to the TME is through their suppressive function towards cytotoxic ($CD8^+$) T cells and their M2 macrophage-polarizing function through the secretion of interleukins (IL-4 and -10). 5. Secreted CatS contributes to the degradation of connective tissue and therefore tumor expansion. It also specifically degrades anti-angiogenic cleavage products of collagen produced by other proteases. 6. In tumor cells, CatS expression and secretion is inducible, for instance signaled by serotonin and interferon receptors ($5HT_7R$, INF-γR). 7. CatS can degrade cell adhesion molecules and through that facilitates tumor expansion, angiogenesis, and metastasis. 8. CatS liberates pro-angiogenic peptides from specific surface proteins. Figure created with Biorender.com.

Human African Trypanosomiasis, a Neglected Tropical Disease

As mentioned above (Table 3), some human proteases are involved in virus entry for diseases like Ebola and Covid-19. Besides the mentioned human Cys cathepsins B and L, human Ser proteases like furin or TMPRSS2, and human metalloproteases like ACE can also be involved in this context.[502,503] Apart from this, the involvement of pathogen proteases in infectious diseases is a longstanding field of investigation, which produced effective antiviral drugs for the disease management of acquired immunodeficiency syndrome (AIDS, e.g., darunavir), Hepatitis C (e.g., asunaprevir), and Covid-19 (nirmatrelvir). It is also applied in research for novel antibiotics as well as antiparasitic agents.[113] Human-pathogenic parasites commonly express Cys proteases with essential character to survival or pathogenicity. Examples are the genera *Plasmodium*, *Trypanosoma*, *Leishmania*, and *Schistosoma*.[504] Of the genus *Trypanosoma*, *T. cruzi* and *T. brucei* pair as the two species that are pathogenic to humans.[505] Comparing them and the other major members of this genus[506], there is a plethora of differences concerning their employed insect vector for transmission, their life-cycle stages, mammalian host organisms, etc.[507] Therefore, only a short comparison of the prime differences between *T. brucei* and *T. cruzi* as the major representatives of the genus' subgroups, transmitted via insect saliva or insect feces, respectively, is shown in Table 4.[508,509] They give indication as to why chemotherapies for Human African Trypanosomiasis (HAT) are not necessarily effective in treating American Trypanosomiasis (Chagas disease) despite promising preclinical results.[510,511] To limit the scope here, no further distinctions within this genus or from other pathogenic *Trypanosomatida* (e.g., *Leishmania*) are discussed. The differences between the three *T. brucei* subspecies are presented in more detail below.

Table 4: Direct comparison of key lifestyle aspects of the three subspecies of *T. brucei* in distinction to *T. cruzi*.[508,512–514]

Species	*T. brucei*			*T. cruzi*
Subspecies	*brucei*	*rhodesiense*	*gambiense*	-
Associated disease	Nagana	rHAT	gHAT	Chagas disease
Main mammalian host	non-human	diverse	human	diverse
Localization in infected mammalian host	extracellular skin, bloodstream, brain			intracellular various cell types
Vector genus (common name)	*Glossina* (Tsetse fly)			Different genera in *Reduviidae* family (Kissing bug)
Vector species	n.s.	*G. fuscipes* *G. morsitans* *G. pallidipes*	*G. fuscipes* *G. palpalis*	
Geographic distribution	"*Glossina* belt" (sub-saharan Africa) entire area	(South-) Eastern part	Central / Western part	South America

n.s. = not specified, r/gHAT = rhodesiense / gambiense Human African Trypanosomiasis

HAT is an endemic disease already vaguely characterized in the Middle Ages (ca. 1200 AD), but the link to insect vectors and the formal proof for trypanosomes in blood were made in the latter half of the 19th century.[515,516] Since then, the World Health Organization (WHO) has set subsequent targets for its work against the disease, which is currently targeted for elimination by 2030.[517–519] HAT is characterized by a neurologic symptom to which it owes its common name "sleeping sickness", since it induces severe disruptions of the sleep cycle in 74 % of patients. This and other neurologic symptoms like motoric and sensory malfunctions emerge only in the second (encephalitic) stage of the disease, in which the parasite has infiltrated the brain. During the first (hemolymphatic) stage, when only peripheral tissues are infested, unspecific symptoms common for many infectious diseases like headache, fever, and joint pain dominate.[520–522] There are two clinically and regionally distinct forms called gambiense HAT (gHAT) and rhodesiense HAT (rHAT) named after the parasite subspecies that causes them. The most prominent differences (gHAT vs. rHAT) are the disease progression through the two clinical stages until death (years vs. months), the geographic distribution (West and Central vs. East Africa), the fraction of case numbers (98 % vs. 2 %), and the predominant mammalian reservoirs (anthroponotic vs. zoonotic). Both variants pose a drastic example of the few infectious diseases with almost 100 % lethality when untreated. The small fraction of humans that spontaneously goes into remission without treatment is now recognized, but only described for gHAT cases.[523] In general, there is a specific human immune response towards trypanosomes mediated by human apolipoprotein L1 (APOL1). It is physiologically incorporated into high-density lipoproteins that are taken up by the parasite via receptor-mediated endocytosis, after which the pore-forming ability of APOL1 disrupts the parasite's membrane integrity.[524] This mechanism protects humans from other *Trypanosoma* species (e.g., *T. brucei brucei*, *T. congolense*, *T. vivax*) that infect mammals used as livestock, but it is ineffective against the two HAT trypanosomes, because they each have a characteristic defense mechanism, enabling them to resist APOL1-mediated responses.[525,526] While *T. b. rhodesiense* can inactivate APOL1 through complex formation with serum resistance-associated protein (SRA), *T. b. gambiense* hinders the APOL1 incorporation into endosomal membranes through the membrane-stiffening effect of *T. b. gambiense*-specific glycoprotein (TgsGP). In both cases, proteolytic digestion of inactive APOL1 by cathepsins plays a major role in parasite survival. Subpopulations of humans in West and Central Africa (where gHAT dominates instead of rHAT) have developed APOL1 variants resistant to the defense mechanism that *T. b. rhodesiense* relies on, conferring immunity towards rHAT.[524] In addition to the utilization of different vector species, this might contribute some explanation for the regional distribution of the two HATs. In general, the WHO goal of elimination seems only achievable for gHAT, because *T. b. rhodesiense* has access to other mammal reservoirs, while *T. b. gambiense* is mainly infective to humans.[527,528] In addition, it is assumed that *T. b. rhodesiense* could evolve from *T. b. brucei* in a relatively small timeframe due to their similarity.[529] This could provide explanation for a resurgence of rHAT in areas where it was ruled non-endemic decades ago, but this assumption requires solid proof.[527,530]

For *T. brucei*, cyclical morphological and physiological changes (Figure 24) are always associated with a change in environment.[508] The parasite has to adapt to the conditions of vector gut, vector salivary gland, and mammal bloodstream. The simplified order of progression starts with bloodstream forms, then gut forms, then salivary gland forms, from which the cycle is repeated. For therapeutic purposes, the bloodstream forms inside the human host are relevant. Of these, two phenotypically distinct variants are described as “stumpy” or “slender”. The stumpy form is characterized by an inability to proliferate due to cell cycle arrest and resistance against environmental influences[531], indicating its role as the reservoir for sustained infection and transmission. However, the role of the stumpy form is not only to be transmitted, but to also act as a handle of parasitemia control: if only the dividing slender form existed, then the host might be killed too quickly to enable further transmission events.[532] That is why slender trypanosomes have quorum sensing ability, which is the ability to measure their local population linked to the concentration of oligopeptides liberated during tissue degradation by secreted trypanosomal proteases.[533] They limit their expansive potential by turning into the non-proliferative stumpy form when a high local parasite count results in high tissue damage, which ensures host survival. In contrast, the slender form is best suited for invasion of the mammal host, because of its ability to multiply and to adapt its surface for immunologic evasion. As long as in their slender bloodstream form, they cover themselves homogenously with surface glycoproteins that shield them from recognition by macrophages.[534] It is termed variable surface glycoprotein (VSG) because it can be exchanged by the parasite quite easily.[535,536] This mechanism is a prime element of the parasite’s immune evasion ability. In the host bloodstream, there is always a dominant population of slender trypanosomes with one VSG called the homotype, and minor populations with a different variable antigen type (VAT) called heterotypes.[508] The homotype population gets targeted by the host immune system, that specifically diminishes this population by antibody-mediated pathways.[537] During remission, this leads to the surge of one of the heterotypes as the new homotype with a VAT the immune system has to readapt to. Ca. 2000 VSGs are DNA-encoded, but only one is expressed at a time.[538] There are other immunologic interferences that lead to fatal parasitemia like B cell depletion in the bone marrow to hamper antibody-mediated responses[539], suppressing immunologic memory in general[540], and increasing their membrane fluidity for quicker endocytotic clearance of surface-bound antibodies[541]. Adding another function to the slender bloodstream form and in contrast to the classical belief that only stumpy forms can survive the mammal-to-fly transition and enter the next stages of the parasitic life cycle to procyclic forms, there is experimental proof that this route is available to slender trypanosomes as well.[542]

Much preclinical research towards this disease is performed on the one *T. brucei* subspecies not responsible for HAT (*T. b. brucei*).[543–545] Biohazard concerns, reliance on a non-human mammal model, and established adaptation to laboratory conditions with reliable protocols are all sound reasons for this. However, for more advanced studies *T. b. rhodesiense* and *T. b. gambiense* models are available.[546–548] This comparability relies on the strong genetic similarity between the subspecies, further discussed below. As another level of complexity, investigations can be performed on monomorphic strains

(e.g., MITat 1.4 = Lister 427, slender form that cannot transform to stumpy form)[549,550], or pleomorphic strains (e.g., AnTat 1.1, transformation possible)[549,551]. A homogeneous response from slender forms to select drugs is easier to interpret, and they represent the major pathologic entity as they are the ones that kill their host. But it is also important to characterize pharmacologic effects on the stumpy form that restarts the infectious cycle in terms of disease elimination and therapeutic interruption of transmission.

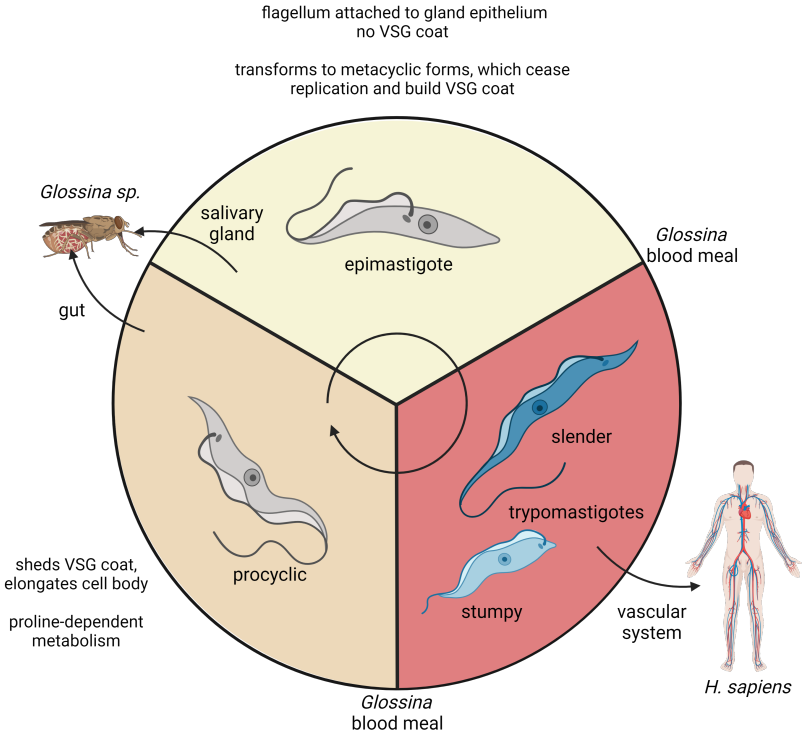


Figure 24: Simplified life cycle of *T. brucei gambiense* with distinctions between host/vector localization (note: human can be exchanged for or supplemented with different other mammals in *T. brucei brucei* or *T. brucei rhodesiense* infections). The characteristics of the trypomastigote (bloodstream) forms are elaborated in detail in the text section above. Different colorization of trypanosomes represents difference in VSG coat. For further differences concerning morphology, metabolism, and replication, as briefly mentioned here, the reader is directed to Lit[508], from which this illustration was adapted. Figure created with Biorender.com.

Pharmacologic Control Efforts Against Human African Trypanosomiasis

The earliest drugs effective against HAT, especially for the hard-to-treat central nervous stage of the disease, were brutal from a modern perspective.[552] Starting with the use of plain sodium arsenite, different organoarsenic drugs were developed and are still in use a century later despite their obvious toxicological risks.[553] These organoarsenic drugs (e.g., (4-aminophenyl)arsenic acid and melarsoprol) were developed in a successful effort to reduce the toxicity of inorganic arsenite.[554] However, major neurological adverse effects like permanent blindness or lethal encephalopathy are still associated with them.[555,556] The mechanism of action for As(III)-derived drugs like melarsoprol is blocking the parasite's glucose metabolism by inhibition of its pyruvate kinase.[554] While melarsoprol was one of the most important drugs for treatment, especially in central-stage cases, since June 2024 its use has been limited to children with rHAT, and in relapsing cases of gHAT.[557]

Structurally inspired by trypanocidal azo dyes like trypane red, suramin was developed as an uncolored derivate suitable for human use.[558] It has a multifactorial mechanism of action that leads to parasite death.[559] From a toxicological point of view it is an improvement from the severe consequences of arsenite therapy, but peripheral polyneuropathy in around 50 % of treated patients, among other relevant organ toxicities, still makes this therapeutic option dangerous.[560] Due to its absent brain-permeation, suramin can only be used in the first stage of the disease and is only in use for children with rHAT.[557]

Symmetric bisamidines represent another group of trypanocidal drugs, with pentamidine being the primary representative.[561] Its mechanism of action is not fully elucidated[562], and it is not without relevant toxicity[563]. The reason this class held great promise was in their investigated potential for oral application as amidoxime prodrugs.[564] All HAT-drugs presented until here need to be applied by an intravenous or intramuscular route due to poor passive permeability. The prodrug melarsoprol, as well as its active form melarsen oxide, are actively transported into trypanosomes by P2-type adenosine transporter.[565] Suramin is actively taken up by endocytosis via invariant surface glycoprotein 75.[566] Lastly, pentamidine is taken up by endocytosis mediated by aquaglyceroporin 2, as well as through the P2-type adenosine and two other transporters.[567,568] To improve passive permeability of bisamidines, an early derivate that was at least similarly effective against HAT compared to pentamidine[563], was chemically derivatized to its bisamidoxime prodrug[569]; a modification that is known to convey oral bioavailability[570]. The effectiveness of this approach was then put to the test for the experimental treatment of trypanosome infections and in clinical trials.[564,571] However, the development was discontinued due to delayed nephrotoxicity during the phase III trial.[572] Despite this, pentamidine is still in use for children with gHAT and optional for adults with gHAT. While evidence in rHAT is limited, it is an option when alternatives are not available.[557]

A dual therapy regime for HAT was introduced by the combination of nifurtimox (oral) and eflornithine (intravenous). The nitrofuran nifurtimox exerts its antiparasitic effect through nitroreductase-generated radicals and subsequent DNA damage.[573] The difluoromethyl-substituted ornithine analog eflornithine is an irreversible inhibitor of the parasite's ornithine decarboxylase, with a weaker effect against *T. b. rhodesiense* due to higher activity and faster turnover of the target enzyme than in the *gambiense* subspecies.[574,575] Eflornithine requires a parenteral route of application and is transported into the parasite cells via amino acid transporter 6.[576] Nifurtimox is orally bioavailable, but is lacking in monotherapy because without a second therapeutic agent, it can require high, toxicity-limited doses for its effect.[577,578] The nifurtimox eflornithine combination therapy (NECT) is an important pillar of treating even severe cases of gHAT.[579] Even though the toxicological profile is still concerning, in light of the severity of central-stage HAT, these adverse effects do usually not result in treatment terminations.[579–581] Therefore, NECT is a first-line therapy for gHAT, especially in the central stage of the disease. It is not used for rHAT.[557]

Structurally linked to nitrofurans is the nitroimidazole fexinidazole, which is the most recently approved oral drug with suggested use against gHAT (since Dec. 2018) and rHAT (since Dec. 2023).[557,582] Both nitrohetaryl classes share their redox-dependent mechanism of action.[583] Fexinidazole is in most cases recommended in favor of both pentamidine and NECT (in gHAT) as well as in favor of suramin or melarsoprol (in rHAT), and generally applied in peripheral and central stages of the disease.[557,582] It requires coadministration with food for 10 consecutive days, and still has some side effects, but this is undeniably a substantial improvement to the comparator therapies.[527]

A novel group of trypanocidal drugs are benzoxaboroles.[584] Acoziborole showed good activity against all subspecies in murine models[585], and was successful in clinical trials against gHAT.[586] The most striking advantage is its oral application as one single dose that is sufficient for therapeutic success. Its molecular target is cleavage- and polyadenylation specificity factor 3, which is a zinc-dependent endonuclease.[544] In its bound state, the boronic acid derivate covalently binds the complexed water molecule within the active site, thereby generating a transition-state mimic of the RNA-phosphate during cleavage.

Most efforts against HAT have been focused on the *gambiense* variant, due to key differentiating factors explained above.[527] An overview is given in Figure 25. The drugs required for treatment are largely supplied by Sanofi (eflornithine, melarsoprol, pentamidine, and fexinidazole) and Bayer (nifurtimox and suramin) via the WHO.[517] Especially with respect to rHAT, which is harder to treat and eliminate than gHAT, alternative treatment options should not be neglected. This is especially true in the context of the evolution of drug resistance, making research on these alternatives valuable from a clinical perspective. The rational design of inhibitors for a yet underexplored target like rhodesain (*Tb*CatL) can potentially fill this gap, which will be discussed next. Beside this immediate benefit, the research on this parasitic cysteine protease is expected to make valuable contributions to the field of protease inhibitors in general. Just like the research on nitroreductases was able to provide valuable drugs for bacterial and parasitic infections by obviously different pathogens, insights gained on inhibitors of rhodesain is expected to be transferrable to the vast field of disease-relevant cysteine proteases described above.

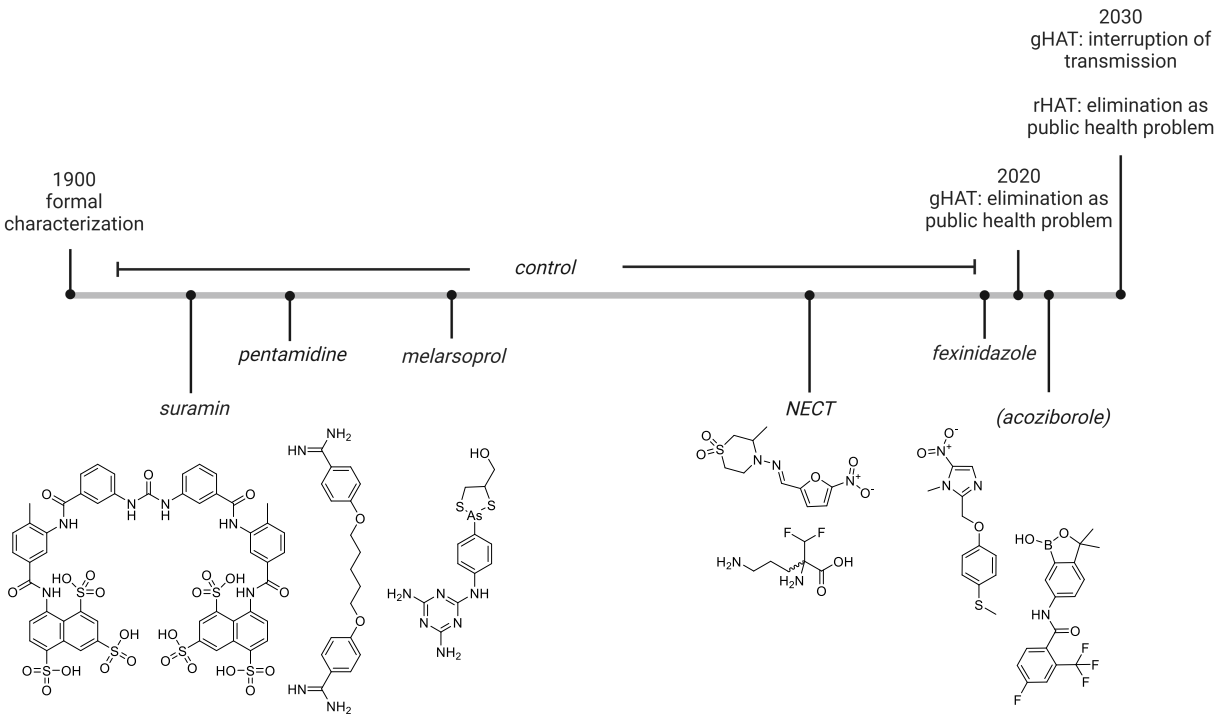


Figure 25: Top information: timeline of Human African Trypanosomiasis from its formal description, followed by the WHO-defined control phase and its elimination as a public health problem, culminating in its targeted elimination. A distinction between g/rHAT is made.[517] Bottom information: timeline for drug advancements against HAT. Dates compiled from the sources mentioned above. Figure created with Biorender.com and ChemDraw.

Rhodesain as a Molecular Target for HAT Therapy

The basis of research interest towards *T. brucei* cathepsin L is the relevance of its activity to disease progression and patient outcome. There is however debate about which name should be used for this enzyme. It is present in ≥98% identity in the three *T. brucei* subspecies *brucei, rhodesiense* and *gambiense*.[587] While distinguished by their infectivity towards humans[525], the phylogenic status of *rhodesiense* as a subspecies besides *brucei* is questioned, further making a separation between rhodesain and brucipain redundant.[513,516] Suggested alternative names indicating the host species, while aiming to omit subspecies ambiguities are *Tbr*CATL and *Tb*CATL.[588,589] For the sake of traceability (i.e., keywords used in published articles that enable search algorithms to find related literature), rhodesain will still appear in this thesis, as well as *Tb*CatL.

The essential status of rhodesain for *T. brucei* was validated in different ways. A study used a murine in vivo model that was infected with genetically modified *T. b. brucei*.[590] The parasite was equipped with a doxycycline-inducible RNAi system targeted to *Tb*CatL. This method for genetic manipulation of trypanosomes can efficiently induce the expression of model proteins (green fluorescent protein, luciferase)[591,592] as well as the induction of RNAi towards essential trypanosomal enzymes[593]. However, the transcription induction of interfering RNA towards *Tb*CatL mRNA resulted only in incomplete silencing in the mentioned study (mRNA reduced by 60 %).[590] In turn, it did not result in curing the infected mice, but still markedly prolonged survival. In the same study, an in vitro model of the

blood-brain barrier could show that this incomplete RNAi towards *Tb*CatL also reduced the parasite's ability to cross a tight membrane cultured from brain microvascular endothelial cells. It was also shown that this enzyme is detectable in higher quantity and with more activity in the human-pathogenic *gambiense* and *rhodesiense* subspecies when compared to *brucei*.[526,594]

In other studies, instead of genetic manipulation of rhodesain levels, a small molecule inhibitor was used to suppress rhodesain activity on a proteomic level to either kill trypanosomes or prevent their passing of the blood-brain barrier.[589,594] To this end, K11777 has become a famous irreversible inhibitor based on a vinylsulfone warhead attached to the relatively lipophilic Phe-hPhe backbone. Vinylsulfones react in a sulfur-analogous Michael-addition with cysteine proteases but not with free thiols, serine or metalloproteases.[204] However, they can bind to threonine proteases.[595] With this backbone, these inhibitors effectively address different cysteine cathepsins with exceptional k_{2nd} towards human CatS, but also efficient inhibition of CatL from human and *T. cruzi*.[204] Selectivity is described over human CatB and CatK, as well as less related cysteine proteases like calpains.[204] The *N*-terminal piperazine capping conveys sufficient solubility through protonation without strong influence on inhibition (k_{2nd}) compared to the original morpholine.[596,597] K11777 is weakly bioavailable, only reaching blood levels sufficient to be effective against parasites without affecting host physiology despite its formal lack of selectivity between *Hs*CatL and *Tb*CatL.[487] While it should be noted that other potential targets of K11777 like the trypanosomal proteasome have been described, it is still assumed that its primary target is *Tb*CatL.[594,596] The treatment of trypanosomes with K11777 was performed in the context of further characterizing the relevance of *Tb*CatL towards blood-brain barrier traversal.[594] It could be shown that the traversal is not propelled by direct digestion of junctional molecules through secreted *Tb*CatL but rather by its modulation of Ca^{2+}-signaling in the endothelial cells. The mechanism of how this is mediated is not fully elucidated, but it is expected to progress in a receptor-mediated way, either through direct interaction of secreted *Tb*CatL with protease-activated receptors or by proteolytic production of agonists at G-protein coupled receptors or tyrosine kinases.

K11777 is a chemical probe that was preclinically investigated in different parasitic disease models (e.g., Chagas disease[598], Schistosomiasis[599], HAT[600], Toxoplasmosis[601]). This diversity for application of cathepsin inhibitors in different parasitologic contexts is an argument making research on the topic of similar structures as inhibitors worthwhile.[602] Especially in the context of covalent peptide-based inhibitors, there are sufficient crystal structures with CatL orthologs from different parasites to enable structure-based drug design on this ligand-inspired scaffold.[597,603] Summarizing, also from the general parasite physiology above, *Tb*CatL in its lysosomal function is involved in parasite nutrition and immune evasion. Peptides are digested for energy or amino acid building blocks, and VSG turnover as well as degradation of inactivated APOL1 is enabled. In its extracellular role, the contribution to quorum sensing by tissue degradation and the receptor-mediated influence on the blood-brain barrier are important. Taken together, lysosomal rhodesain is relevant for parasite survival, while secreted rhodesain is linked to infectivity through the capacity of the parasite to infiltrate the human central

nervous system, which is associated with disease progression and severity. An overview of the function of rhodesain is given in Figure 26.

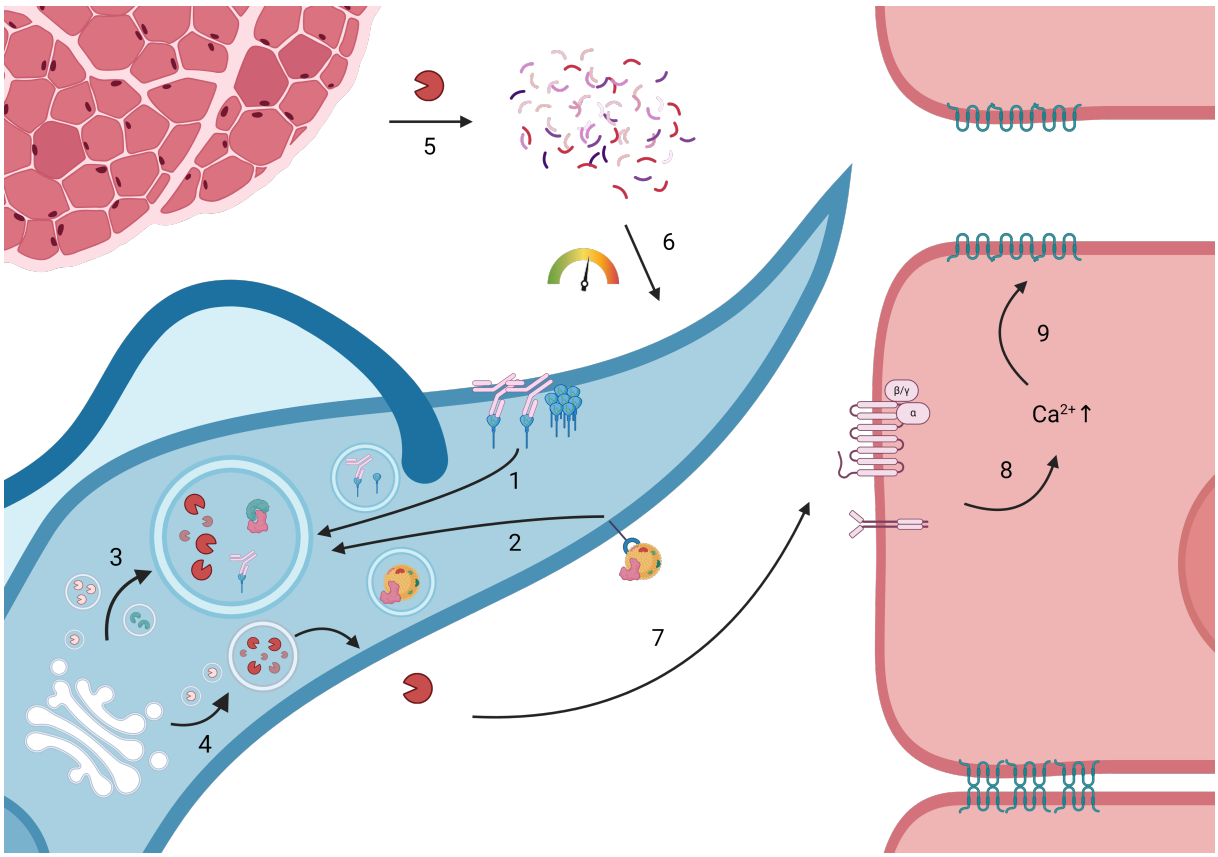


Figure 26: The various roles of *Tb*CatL for parasite survival, immune evasion, and the invasion of the central nervous system. 1. Slender bloodstream forms are covered in a layer of VSG, which is subjected to a constant turnover mediated by lysosomal proteases like rhodesain. 2. One resistance mechanism of humans towards trypanosomes is delivered by high-density lipoproteins (APOL1, pink wedge), which is transported to the lysosome, inactivated, and degraded. 3. Rhodesain is expressed and shuttled to the lysosomal compartment in an inactive form (bright red), where it is proteolytically activated (dark red). SRA or *Tg*sGP (the resistance proteins towards APOL1 toxicity) are shuttled there in the same manner (depicted for SRA, green wedge). 4. Rhodesain in secretory lysosomes is secreted into the pericellular space, where it retains its proteolytic activity. 5. Rhodesain degrades surrounding tissue and the generated peptide fragments contribute to the parasite's quorum sensing ability (6.). 7. Rhodesain promotes the infiltration of the brain through receptor-mediated change in Ca^{2+}-signaling (8.), resulting in loosening of the tight-junctions of the blood-brain barrier (9.). Figure created with Biorender.com.

Project Overview

The publications that are part of this thesis can be divided into two projects. Their connecting quality is the characterization of peptide-based covalent inhibitors of different proteases. Their distinguishing feature is the employed methodology for these characterizations.

In the first project, active site interactions of protein-inhibitor complexes were characterized to elucidate the binding mode of ligand-inspired protease inhibitors with an electrophilic warhead. All included publications investigated rhodesain as a target essential to trypanosomes which are the pathogens of Human African Trypanosomiasis. To this end, a MALDI-TOF-based mass spectrometry method was established with the given complexes to prove covalent target engagement and applied in three publications that distinguish themselves by the employed group of warheads. The first publication focused on α-fluoro-substituted vinylsulfones and -sulfonates that represent a reversibly reacting option to the established irreversible vinylsulfones. The second article focused on keto Michael acceptors that distinguish themselves by the presence of multiple possible electrophilic sites. The third publication focused on electron-deficient arenes as S_NAr warheads which were modulated to generate stable intermediates of this reaction type as the dominant inhibiting complex. For the two latter publications, the target protease was also heterologously expressed in *P. pastoris*, and investigations on binding affinity and kinetics incl. reversibility were performed.

The following publications constitute the results obtained for this project (‡ indicates shared first authorship):

Jung, S.‡; Fuchs, N.‡; Johe, P.; Wagner, A.; Diehl, E.; Yuliani, T.; **Zimmer, C.**; Barthels, F.; Zimmermann, R.A.; Klein, P.; et al. Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure-Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies. *J. Med. Chem.* **2021**, *64*, 12322–12358, doi:10.1021/acs.jmedchem.1c01002.

> Own contributions: Establishing a MALDI-TOF-MS method from different acidic matrices for rhodesain and its inhibitors to distinguish covalent from non-covalent binders, characterizing the covalency of binding for inhibitors **1**, **2a**, **2j** and **2k** in this context, writing parts and editing of the manuscript.
>
> Contribution from others: Inhibitor synthesis and characterization of in vitro affinity, inhibition kinetics, and selectivity, reversibility assessment by dilution assay, non-covalent docking, ESI-MS experiments, QM/MM modeling, antitrypanosomal effect, cytotoxicity, in vitro microsomal metabolism, in vivo distribution study, writing and editing of the manuscript.

Agost-Beltrán, L.; **Zimmer, C.**; Räder, H.J.; Kersten, C.; Schirmeister, T.; Rodríguez, S.; González, F. V. Rhodesain Inhibitors on the Edge of Reversibility-Irreversibility. *Bioorg. Chem.* **2024**, *153*, 107830, doi:10.1016/j.bioorg.2024.107830.

Own contributions: Heterologous expression of rhodesain, preliminary characterization of the in vitro affinity and inhibition kinetics of **11**, **14**, and **15** to validate the approach and verify the results, applying MALDI-TOF-MS to prove the covalency of binding for all final inhibitors, writing parts and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of in vitro affinity, inhibition kinetics, and selectivity, reversibility assessment by dilution and dialysis assay, non-covalent and covalent docking, molecular dynamics simulation, writing and editing of the manuscript.

Zimmer, C.‡; Brauer, J.‡; Ferenc, D.; Meyr, J.; Müller, P.; Räder, H.-J.; Engels, B.; Opatz, T.; Schirmeister, T. Substitution-Induced Mechanistic Switching in S_NAr-Warheads for Cysteine Proteases. *Molecules* **2024**, *29*, 2660, doi:10.3390/molecules29112660.

Own contribution: Heterologous expression of rhodesain, characterization of the in vitro affinity, inhibition kinetics, and selectivity, including inner-filter effect correction, reversibility assessment by dilution assay, characterization of the covalency of inhibitor binding by MALDI-TOF-MS, reactivity assessment towards a model thiol in an aqueous solution, non-covalent docking, assessment of passive permeation by PAMPA, writing and editing of the manuscript.

Contribution from others: Inhibitor synthesis, reactivity study against a model thiolate in an organic solvent, quantum mechanical calculation of the reaction progress, writing and editing of the manuscript.

This last publication of the aforementioned project already introduces the second project, in which protease inhibitors were characterized in a target-independent way to elucidate their potential for passive membrane permeation. To this end, a parallel artificial membrane permeation assay was optimized and applied to the inhibitors in this and three additional publications. They distinguish themselves by their target context and therefore the employed inhibitor structures and warheads, while all inhibitors were built on peptide-based scaffolds. The characterization of esters as prodrug structures in the context of cell-based assays was performed for the S_NAr-based inhibitors mentioned above. In the first new paper of this section, arginine-mimetics were evaluated for their potential influence on the high polarity of the guanidine moiety in the context of permeability. The second paper revolves around inhibitors with the various Michael acceptor warheads alluded to in the first section (both vinylsulfones and keto vinylesters) but now in the context of an effect in a cellular assay. For the last publication, structure-permeability relationships that correlate calculated physicochemical parameters with in vitro permeation results were established on peptidomimetic inhibitors bearing either a nitrile or a carbonyl-based warhead.

The following additional publications constitute the results obtained for this project:

Müller, P.; **Zimmer, C.**; Frey, A.; Holzmann, G.; Weldert, A. C.; Schirmeister, T. Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres. *Int. J. Mol. Sci.* **2024**, *25*, 1375, doi:10.3390/ijms25031375.

Own contribution: Protein similarity calculations for the employed serine proteases, spectroscopic characterization of the α-keto hetaryl warheads, LCMS-based assessment of the inhibitors' stability in an aqueous solution, calculation of physicochemical parameters (pK_a, logP, $logD_{7.4}$) for arginine-mimetic inhibitors, PAMPA assessment of passive permeability, writing parts and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of the in vitro affinity and selectivity, writing and editing of the manuscript.

Fuchs, N.‡; Zimmermann, R.A.‡; Schwickert, M.‡; Gunkel, A.; **Zimmer, C.**; Meta, M.; Schwickert, K.; Keiser, J.; Haeberli, C.; Kiefer, W.; et al. Dual Strategy to Design New Agents Targeting *Schistosoma mansoni*: Advancing Phenotypic and *Sm*CB1 Inhibitors for Improved Efficacy. *ACS Infect. Dis.* **2024**, 10, 1664–1678, doi:10.1021/acsinfecdis.4c00020.

Own contribution: Evaluating the passive permeability of the SJ-series by PAMPA, correlating the passive permeability with inhibition data against isolated target and cellular effect, writing parts and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of the in vitro affinity, inhibition kinetics, and selectivity, antischistosomal effect, correlation of the physicochemical parameters (logP, PSA) with inhibition data against the isolated target and the cellular effect, non-covalent and covalent docking, writing and editing of the manuscript.

Meta, M.‡; **Zimmer, C.**‡; Fuchs, N.; Zecher, M.J.; Lahu, A.; Schirmeister, T. Structural Modifications of Covalent Cathepsin S Inhibitors: Impact on Affinity, Selectivity, and Permeability. *ACS Med. Chem. Lett.* **2024**, *15*, 837–844, doi:10.1021/acsmedchemlett.4c00050.

Own contribution: Evaluating the passive permeability of all final inhibitors by PAMPA, correlating the passive permeability with physicochemical parameters (logP, PSA), LCMS-based aqueous stability analysis and reactivity assessment towards a model thiol for **44**, writing and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of the in vitro affinity, inhibition kinetics, and selectivity, reversibility assessment by dilution assay, non-covalent and covalent docking, hydrazone liberation assay, writing and editing of the manuscript.

Publication record

Publications as part of this dissertation

(‡ indicates shared first authorship)

1) **Investigations of Binding Interactions**

Jung, S.‡; Fuchs, N.‡; Johe, P.; Wagner, A.; Diehl, E.; Yuliani, T.; **Zimmer, C.**; Barthels, F.; Zimmermann, R.A.; Klein, P.; et al. Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure-Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies. *J. Med. Chem.* **2021**, *64*, 12322–12358, doi:10.1021/acs.jmedchem.1c01002.

Agost-Beltrán, L.; **Zimmer, C.**; Räder, H. J.; Kersten, C.; Schirmeister, T.; Rodríguez, S.; González, F. V. Rhodesain Inhibitors on the Edge of Reversibility-Irreversibility. *Bioorg. Chem.* **2024**, *153*, 107830, doi:10.1016/j.bioorg.2024.107830.

Zimmer, C.‡; Brauer, J.‡;_Ferenc, D.; Meyr, J.; Müller, P.; Räder, H.-J.; Engels, B.; Opatz, T.; Schirmeister, T. Substitution-Induced Mechanistic Switching in S_NAr-Warheads for Cysteine Proteases. *Molecules* **2024**, *29*, 2660, doi:10.3390/molecules29112660.

2) **Investigations of Permeability**

Müller, P.; **Zimmer, C.**; Frey, A.; Holzmann, G.; Weldert, A. C.; Schirmeister, T. Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres. *Int. J. Mol. Sci.* **2024**, *25*, 1375, doi:10.3390/ijms25031375.

Fuchs, N.‡; Zimmermann, R. A.‡; Schwickert, M.‡; Gunkel, A.; **Zimmer, C.**; Meta, M.; Schwickert, K.; Keiser, J.; Haeberli, C.; Kiefer, W.; et al. Dual Strategy to Design New Agents Targeting *Schistosoma mansoni*: Advancing Phenotypic and *Sm*CB1 Inhibitors for Improved Efficacy. *ACS Infect. Dis.* **2024**, 10, 1664–1678, doi:10.1021/acsinfecdis.4c00020.

Meta, M.‡; **Zimmer, C.**‡; Fuchs, N.; Zecher, M. J.; Lahu, A.; Schirmeister, T. Structural Modifications of Covalent Cathepsin S Inhibitors: Impact on Affinity, Selectivity, and Permeability. *ACS Med. Chem. Lett.* **2024**, *15*, 837–844, doi:10.1021/acsmedchemlett.4c00050.

Additional publications beyond this dissertation

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.

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.

Di Chio, C.; Previti, S.; De Luca, F.; Bogacz, M.; **Zimmer, C.**; Wagner, A.; Schirmeister, T.; Zappalà, M.; Ettari, R. Drug Combination Studies of the Dipeptide Nitrile CD24 with Curcumin: A New Strategy to Synergistically Inhibit Rhodesain of Trypanosoma Brucei Rhodesiense. *Int. J. Mol. Sci.* **2022**, *23* (22). https://doi.org/10.3390/ijms232214470.

Eurtivong, C.; **Zimmer, C.**; Schirmeister, T.; Butkinaree, C.; Saruengkhanphasit, R.; Niwetmarin, W.; Ruchirawat, S.; Bhambra, A. S. A Structure-Based Virtual High-Throughput Screening, Molecular Docking, Molecular Dynamics and MM/PBSA Study Identified Novel Putative Drug-like Dual Inhibitors of Trypanosomal Cruzain and Rhodesain Cysteine Proteases. *Mol. Divers.* **2023**. https://doi.org/10.1007/s11030-023-10600-2.

Johe, P.; Jung, S.; Endres, E.; Kersten, C.; **Zimmer, C.**; Ye, W.; Sonnichsen, C.; Hellmich, U. A.; Sotriffer, C.; Schirmeister, T.; Neuweiler, H. Warhead Reactivity Limits the Speed of Inhibition of the Cysteine Protease Rhodesain. *ACS Chem. Biol.* **2021**, *16* (4), 661–670. https://doi.org/10.1021/acschembio.0c00911.

Maus, H.‡; Müller, P.‡; Meta, M.; Hoba, S. N.; Hammerschmidt, S. J.; Zimmermann, R. A.; **Zimmer, C.**; Fuchs, N.; Schirmeister, T.; Barthels, F. Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates. *Chem. – A Eur. J.* **2023**, *29* (50). https://doi.org/10.1002/chem.202301855.

Medrano, F. J.; de la Hoz-Rodríguez, S.; Martí, S.; Arafet, K.; Schirmeister, T.; Hammerschmidt, S. J.; Müller, C.; González-Martínez, Á.; Santillana, E.; Ziebuhr, J.; Romero, A.; **Zimmer, C.**; Weldert, A.; Zimmermann, R.; Lodola, A.; Świderek, K.; Moliner, V.; González, F. V. Peptidyl Nitroalkene Inhibitors of Main Protease Rationalized by Computational and Crystallographic Investigations as Antivirals against SARS-CoV-2. *Commun. Chem.* **2024**, *7* (1), 1–16. https://doi.org/10.1038/s42004-024-01104-7.

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.; Rodríguez, S.; Diemer, K.; Neumann, W.; Gonzàlez, 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.

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.

Previti, S.; Ettari, R.; Di Chio, C.; Legac, J.; Bogacz, M.; **Zimmer, C.;** Schirmeister, T.; Rosenthal, P. J.; Zappalà, M. Influence of Amino Acid Size at the P3 Position of N-Cbz-Tripeptide Michael Acceptors Targeting Falcipain-2 and Rhodesain for the Treatment of Malaria and Human African Trypanosomiasis. *Bioorg. Chem.* **2023**, *137* (May), 1–11. https://doi.org/10.1016/j.bioorg.2023.106587.

Welker, A.; Kersten, C.; Müller, C.; Madhugiri, R.; **Zimmer, C.**; Müller, P.; Zimmermann, R.; 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.

Publications

1: Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure-Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies.

The vinylsulfone is a successfully applied irreversible warhead in the development of preclinical inhibitors of cysteine proteases. While its irreversible mode of action sparked doubt about its suitability for clinical applications, K11777 as one of the furthest investigated members of this class has proven biocompatible. It was moderately bioavailable and did not show obvious toxicity, but room for optimization remains. Human African Trypanosomiasis, the disease in which context derivates were evaluated in this study, requires drug distribution into the central nervous system to treat the second stage of the disease, for which only few treatment options are currently available. Even though options have increased in recent years, a lack of alternatives can be concerning considering resistance development. In addition to the need for central nervous bioavailability, increased selectivity is a feature required for optimization of K11777. Apart from modification necessary to the peptide backbone, this ties into the fact that most approved covalent protease inhibitors carry a reversibly covalent warhead, which also requires revisiting the possible modifications to vinylsulfones. Past research had characterized α-fluorovinylsulfones as suitable reversible options for the inhibition of rhodesain.[604–606] Additionally, vinylsulfonates had also been validated, making the combined α-fluorovinylsulfonates potentially viable.[607]

Therefore, α-fluorovinylsulfones and α-fluorovinylsulfonates were evaluated as warhead options, with simultaneous modulation of the dipeptidic recognition sequence and the *N*-terminal capping group. Surprisingly, the two warheads showed distinct behaviors regarding their inhibition kinetics in the standard inhibition assays which necessitated a recharacterization of the entire binding mode for these inhibitors. The mechanistic reason for the reversibility of α-fluorovinylsulfones was probed by a computational QM/MM approach. The initial assumption was that the fluorine in α-position of the double bond electrostatically stabilizes the intermediate carbanion resulting from the nucleophilic attack of rhodesain's active site cysteine thiolate, as described for α-cyanoacrylamides in the context of cysteine-reactive kinase inhibitors.[608] It was shown however by molecular dynamics simulations, that intramolecular torsion is required for the fluorinated double bond that is not necessary for the unfluorinated double bond, because the two bind in different initial non-covalent orientations. The energy required for torsion prior to the covalent reaction is the reason why the exothermicity of the reaction on the fluorinated double bond is reduced, making the entire reaction reversible. The fact that these inhibitors still bind covalently was verified with MALDI-TOF mass spectrometry from acidic matrices that suppress non-covalent interactions. It was shown that structural modifications to the *N*-terminal cap and a change from vinylsulfone (**K11777**) over α-fluorovinylsulfone (**1**) to α-fluorovinylsulfonates with different *N*-terminal capping groups (**2a**, **2j**, **2k**) retained the covalent

binding mode. Reversibility assessment then elucidated a difference in dissociation from the target between the reversible warheads depending on the sulfone or sulfonate moieties. While the α-fluorovinylsulfones showed the fast reversible reaction as observed before, the α-fluorovinylsulfonates showed slower reversibility. Hence, the affinity and kinetics of the two classes were calculated accordingly with time-independent or time-dependent models, respectively. Reasons for selectivity over the related human cathepsins B and L were mainly attributed to specific changes to the P2 and P3 residues. In addition to these mechanistic investigations, the inhibitors were also evaluated in a cell-based assay against *T. brucei* where many showed low micromolar or nanomolar EC_{50} values with low cytotoxicity, especially for the derivates without the *N*-terminal *N*-methyl piperazine. The distribution study in the chosen mouse model showed improved brain permeation and accumulation for an inhibitor, in which the *N*-methyl piperazine was exchanged for pyridine. In vitro metabolism elucidated a possible metabolic hotspot for these inhibitors, which is the *N*-terminal cap; demethylation of *N*-methyl piperazine and *N*-oxidation of *N*-methyl piperazine or pyridine were observed. Taken together this publication is a collaborative investigation of the inhibition mechanism of uncommon reversible variants of the vinylsulfone warhead targeting a cysteine protease involved in a neglected tropical disease. The preclinical results included are promising indicators for future drug design efforts.

Own contributions: Establishing a MALDI-TOF-MS method from different acidic matrices for rhodesain and its inhibitors to distinguish covalent from non-covalent binders, characterizing the covalency of binding for inhibitors **1**, **2a**, **2j** and **2k** in this context, writing parts and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of in vitro affinity, inhibition kinetics, and selectivity, reversibility assessment by dilution assay, non-covalent docking, ESI-MS experiments, QM/MM modeling, antitrypanosomal effect, cytotoxicity, in vitro microsomal metabolism, in vivo distribution study, writing and editing of the manuscript.

This work was published in the Journal of Medicinal Chemistry.

Article reprinted with permission of *J. Med. Chem.* **2021**, 64, 16, 12322–12358, “Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure–Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies”, Copyright © 2021 American Chemical Society. Published by American Chemical Society.

The following publication quoted within “” (pages 73–114) is the same as the manuscript cited here. The appended Supporting Information represents an abridged version. The full version can be accessed online at doi:10.1021/acs.jmedchem.1c01002.

Journal of Medicinal Chemistry

pubs.acs.org/jmc Article

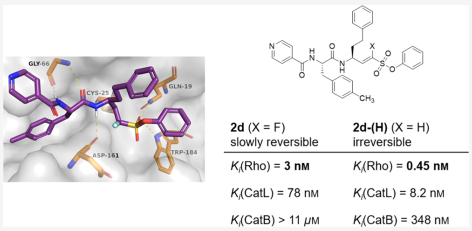
Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure−Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies

Sascha Jung,[‡‡] Natalie Fuchs,[‡‡] Patrick Johe, Annika Wagner, Erika Diehl, Tri Yuliani, Collin Zimmer, Fabian Barthels, Robert A. Zimmermann, Philipp Klein, Waldemar Waigel, Jessica Meyr, Till Opatz, Stefan Tenzer, Ute Distler, Hans-Joachim Räder, Christian Kersten, Bernd Engels, Ute A. Hellmich, Jochen Klein, and Tanja Schirmeister*

Cite This: *J. Med. Chem.* 2021, 64, 12322−12358 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Rhodesain is a major cysteine protease of *Trypanosoma brucei rhodesiense*, a pathogen causing Human African Trypanosomiasis, and a validated drug target. Recently, we reported the development of α-halovinylsulfones as a new class of covalent reversible cysteine protease inhibitors. Here, α-fluorovinylsulfones/-sulfonates were optimized for rhodesain based on molecular modeling approaches. **2d**, the most potent and selective inhibitor in the series, shows a single-digit nanomolar affinity and high selectivity toward mammalian cathepsins B and L. Enzymatic dilution assays and MS experiments indicate that **2d** is a slow-tight binder (K_i = 3 nM). Furthermore, the nonfluorinated **2d-(H)** shows favorable metabolism and biodistribution by accumulation in mice brain tissue after intraperitoneal and oral administration. The highest antitrypanosomal activity was observed for inhibitors with an N-terminal 2,3-dihydrobenzo[*b*][1,4]dioxine group and a 4-Me-Phe residue in P2 (**2e**/**4e**) with nanomolar EC_{50} values (0.14/0.80 μM). The different mechanisms of reversible and irreversible inhibitors were explained using QM/MM calculations and MD simulations.



INTRODUCTION

Human African Trypanosomiasis (HAT, sleeping sickness) is a severe disease classified as a neglected tropical disease (NTD).[1] HAT is caused by the protozoan parasite *Trypanosoma brucei* (*T. brucei*), which is transmitted to humans via the bite of the Tsetse fly.[2] Sleeping sickness is fatal if left untreated. Pentamidine and suramin are used to treat the early, hemolymphatic stage of the disease, while eflornithine and melarsoprol and the combination therapy nifurtimox-eflornithine target the late, neurological stage of the disease.[3] Recently, the nitroimidazole fexinidazole was introduced as the first oral treatment of both, stage-1 and stage-2 *T. b. gambiense* HAT.[4] However, most available drugs show severe toxicity, poor bioavailability, and need long-time administration due to their lack of efficiency.[5] Therefore, there is an urgent need to develop new therapies against this disease and, in addition, *T. brucei* can serve as a valuable model organism for other pathogenic kinetoplastid diseases. The cysteine protease rhodesain (*Tb*CatL) is essential for the development of the parasite and for the progression of the disease.[6,7] Inhibition of the parasitic cysteine protease activity has been validated as a drug target in vitro and in vivo.[8] Consequently, rhodesain represents a promising target for the development of safer drugs against HAT.

Rhodesain belongs to the papain family of cysteine proteases and shares high structural similarity with the human cathepsins, especially cathepsin L (CatL; sequence identity 44.7%, similarity 59.1%, C_α-RMSD 1.35 Å).[9,10] A prominent inhibitor of papain-family cysteine proteases is **K11777**, a peptide-based vinylsulfone that mimics the autoinhibition of prorhodesain and that reacts in a Michael-type addition with the active-site cysteine (Figure 1).[10,11] The arising carbanion is protonated, resulting in the irreversible formation of the covalent enzyme−inhibitor complex. Substitution of the hydrogen at the α-position of the

Received: June 3, 2021
Published: August 11, 2021

ACS Publications © 2021 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Downloaded via UNIV MAINZ on January 4, 2023 at 10:26:29 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

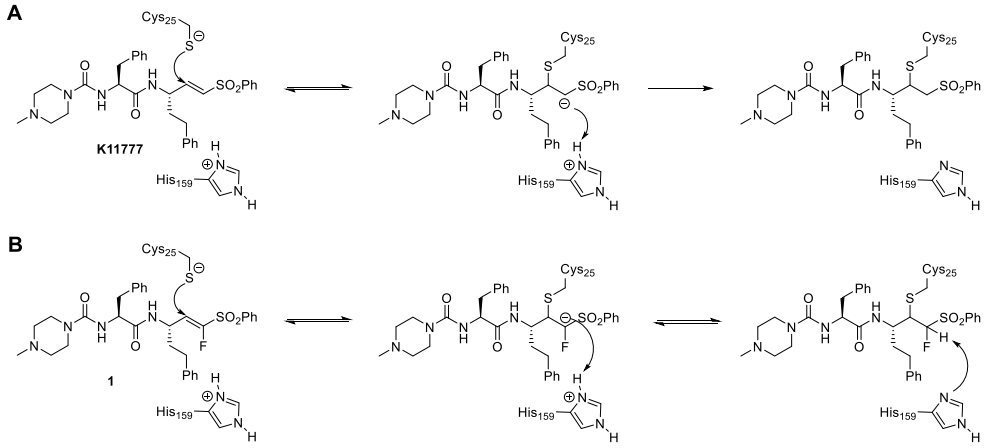
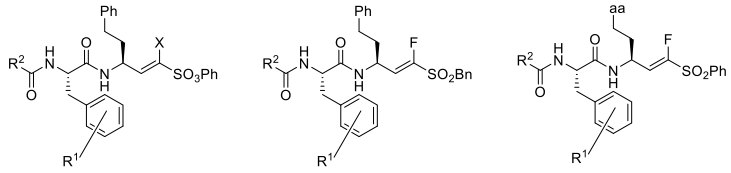


Figure 1. Structures and inhibition mechanisms of irreversible (**K11777**; A) and covalent reversible (**1**; B) peptide-based vinylsulfones.



2a	R^1 = H, R^2 = 4-Pyr, X = F	**3a**	R^1 = H, R^2 = 4-Pyr	**4a**	R^1 = H, R^2 = 4-Pyr, aa = hPhe
2a-(Z)	R^1 = H, R^2 = 4-Pyr, X = F	**3b**	R^1 = H, R^2 = DHBD	**4b**	R^1 = H, R^2 = DHBD, aa = hPhe
2b	R^1 = H, R^2 = DHBD, X = F	**3c**	R^1 = H, R^2 = 3,5-DiFPh	**4c**	R^1 = H, R^2 = 3,5-DiFPh, aa = hPhe
2d	R^1 = 4-Me, R^2 = 4-Pyr, X = F	**3d**	R^1 = 4-Me, R^2 = 4-Pyr	**4d**	R^1 = 4-Me, R^2 = 4-Pyr, aa = hPhe
2d-(H)	R^1 = 4-Me, R^2 = 4-Pyr, X = H	**3e**	R^1 = 4-Me, R^2 = DHBD	**4e**	R^1 = 4-Me, R^2 = DHBD, aa = hPhe
2e	R^1 = 4-Me, R^2 = DHBD, X = F	**3f**	R^1 = 4-Me, R^2 = 3,5-DiFPh	**4f**	R^1 = 4-Me, R^2 = 3,5-DiFPh, aa = hPhe
2f	R^1 = 4-Me, R^2 = 3,5-DiFPh, X = F	**3g**	R^1 = 3-Me, R^2 = 4-Pyr	**4g**	R^1 = 3-Me, R^2 = 4-Pyr, aa = hPhe
2g	R^1 = 3-Me, R^2 = 4-Pyr, X = F	**3h**	R^1 = 3-Me, R^2 = DHBD	**4h**	R^1 = 3-Me, R^2 = DHBD, aa = hPhe
2h	R^1 = 3-Me, R^2 = DHBD, X = F	**3i**	R^1 = 3-Me, R^2 = 3,5-DiFPh	**4i**	R^1 = 3-Me, R^2 = 3,5-DiFPh, aa = hPhe
2i	R^1 = 3-Me, R^2 = 3,5-DiFPh, X = F			**4j**	R^1 = H, R^2 = *N*-MePip, aa = Met
2j	R^1 = H, R^2 = OBn, X = F			**4k**	R^1 = H, R^2 = 4-Pyr, aa = Met
2k	R^1 = H, R^2 = *N*-MePip, X = F			**4l**	R^1 = H, R^2 = *N*-MePip, aa = Cys(Bn)

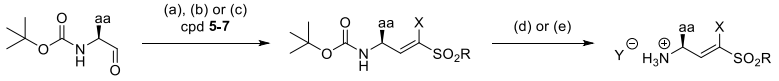
Figure 2. Structures of compounds **2a**−**4l**.

double bond by fluorine (compound **1**) generates an *α*-fluorovinylsulfone, which can undergo a reversible Michael-type addition with thiols (Figure 1).[12]

The development of covalent inhibitors has seen a resurgence in academia as well as in the industry during the past decade.[13] There is much debate on the advantages and disadvantages of covalent inhibition, especially concerning reversible covalent inhibition mechanisms.[14] On the one hand, covalent reversible inhibitors alleviate some of the concerns arising from covalent irreversible protein modifications, such as toxicity emerging from off-target effects,[15] idiosyncratic toxicity,[16] and haptenization,[17] but maintain benefits such as enhanced potency and prolonged residence times.[18,19] In recent studies with reversible fluorinated vinylsulfones and their irreversible counterparts, it was shown that the electrophilic group, the so-called warhead, and not the binding of the peptidic recognition unit limits the kinetics of inhibition of the protease and that the fluorinated vinylsulfone warhead reduces the rate constant of binding.[20,21]

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Scheme 1. Synthesis of Building Blocks 13−17[a]



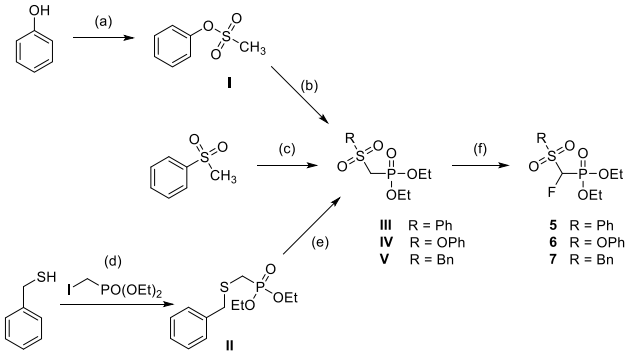
VI aa = hPhe
VII aa = Met
VIII aa = Cys(Bn)

(a) **8** R = Ph, X = F, aa = hPhe
9 R = OPh, X = F, aa = Met
(*Z*)-9 R = OPh, X = F, aa = Cys(Bn)
(b) **10** R = Bn, X = F, aa = hPhe
(c) **9-(H)** R = OPh, X = H, aa = hPhe
11 R = Ph, X = F, aa = Met
12 R = Ph, X = F, aa = Cys(Bn)

(d) **13** Y = Cl, R = Ph, X = F, aa = hPhe
14 Y = Cl, R = OPh, X = F, aa = hPhe
(*Z*)-14 Y = Cl, R = OPh, X = F, aa = hPhe
14-(H) Y = Cl, R = OPh, X = H, aa = hPhe
15 Y = Cl, R = Bn, X = F, aa = hPhe
(e) **16** Y = TFA, R = Ph, X = F, aa = Met
17 Y = TFA, R = Ph, X = F, aa = Cys(Bn)

[a]Reagents and conditions: aa = amino acid side chain; (a) NaH, THF, 0 °C, 1 h, 22−43%; (b) KHMDS, THF, −78 °C for 20 min, 1 h at rt, 47%; (c) LHMDS, THF, −78 °C for 30 min, 12 h at rt, 54−75%; (d) 4 M HCl in dioxane, rt, 30 min, quant.; and (e) TFA, DCM, 0 °C, 1 h, quant.

Scheme 2. Preparation of Phosphonates 5−7[a]



III R = Ph
IV R = OPh
V R = Bn

5 R = Ph
6 R = OPh
7 R = Bn

[a]Reagents and conditions: (a) Ms-Cl, TEA, EtOAc, 0 °C to rt, 30 min, 92%; (b) diethyl chlorophosphite (DECP), KHMDS, THF, −78 to −60 °C, 1 h, 74%; (c) DECP, *n*-BuLi, THF, 0 °C, 1 h, 52%; (d) NaH, THF, 0 °C to rt, 4 h, 84%; (e) *m*CPBA, DCM, 0 °C to rt, 12 h, 99%; and (f) Selectfluor, KHMDS, THF/DMF, −78 °C to rt, 3 h, 48−62%.

Based on the previously reported structure of the covalent reversible inhibitor **1**,[12] in the present study we performed structure−activity relationship (SAR) studies to optimize inhibition potency for rhodesain and selectivity against human cathepsins B and L using molecular docking approaches. Additionally, quantum chemical-based computations were performed to get further information about the differences in the inhibition mechanisms of **K11777** and its fluorinated counterpart **1**. These studies were performed because in our previous study, the computed reaction energies for X = F and X = H in the α-position of the warhead (Figure 1) did not differ significantly. Thus, the transition from irreversible to reversible inhibition by exchange of hydrogen by fluorine could not be explained conclusively. The covalent behavior of the compounds was experimentally evaluated with MS studies, and reversibility was demonstrated with enzymatic dilution and dialysis assays.

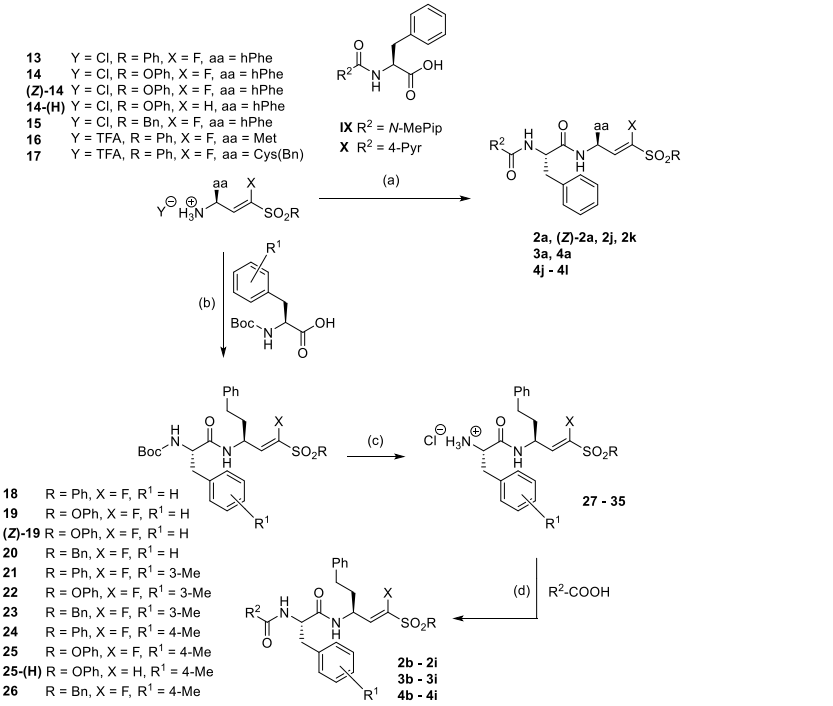
Furthermore, the ADME parameters of compound **1** and the optimized inhibitor **2d-(H)** were investigated via in vitro metabolism studies and in vivo mouse studies in order to determine their organ distribution and their accumulation in brain tissue, thus evaluating their potential as candidates for the treatment of stage-2 HAT.

RESULTS AND DISCUSSION

Design of Inhibitors. Structural variations of the fluorinated vinylsulfone **1** (Figure 1B) were inspired by peptide-based inhibitors of rhodesain, which contain structural elements at the P3, P2, and P1′ positions[21] that are known to either enhance potency against rhodesain and/or increase selectivity against CatL and CatB.[22−25] For the P3 position, introduction of aromatic and heteroaromatic systems was reported to favor inhibition of rhodesain over CatL and CatB.[22−24] Introduction of an additional methyl group to the 3- or 4-position of the phenylalanine aromatic ring at the P2 position can improve potency and selectivity for rhodesain.[22] Additionally, the extension of the phenyl ring at the warhead into the S1′ pocket via linker atoms increases potency for rhodesain.[25] Based on these observations, a virtual library of 511 modified compounds was generated. These compounds were docked at rhodesain (crystal structure of rhodesain bound to **K11777**, protein databank (PDB) 2p7u[11]) using FlexX[26] and DOCKTITE,[27] as reported previously (Table S1).[12] The noncovalent enzyme−inhibitor complex was generated with FlexX, and the scores reflect whether the designed compounds have an improved noncovalent affinity compared to the starting vinylsulfone **1**. Only compounds with a comparable or higher score were selected for synthesis. In addition, the covalent complex was modeled with DOCKTITE.[27] In this case, the scores (Table S1)

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Scheme 3. Synthesis of Compounds 2a−4l from Building Blocks 13−17[a]



[a]Reagents and conditions: (a) TBTU, HOBt, DIEA, DMF, 0 °C to rt, 12 h, 59−81%; (b) TBTU, HOBt, DIEA, DCM, 0 °C to rt, 12 h, 69−83%; (c) 4 M HCl in dioxane, rt, 30 min, quant.; and (d) TBTU, DIEA, DCM, 0 °C to rt, 12 h, 62−84%.

may be interpreted in terms of stability of the protein-bound state.

Based on the results obtained from docking, a series of compounds (**2a−4i**) were selected for synthesis and subsequent evaluation of SAR (Figure 2). For all selected compounds, scores obtained from both docking approaches were generally higher when compared to starting compound **1** (Table S1). Depending on the substitution pattern of the warhead at the P1′ position, compounds can be subdivided into aromatic fluorovinylsulfones (phenyl substituent at P1′, cpds **4a−4i**), aliphatic fluorovinylsulfones (benzyl group at P1′, cpds **3a−3i**), and fluorovinylsulfonates (phenol ester at P1′, cpds **2a−2k**).

Based on noncovalent docking (Table S1), compounds (**4j−4l**) with methionine and *S*-benzylcysteine at P1 were suggested to show a comparable (yet slightly lower) affinity to the respective molecules with a homophenylalanine residue. Therefore, these compounds were included in the present study and investigated for their potency.

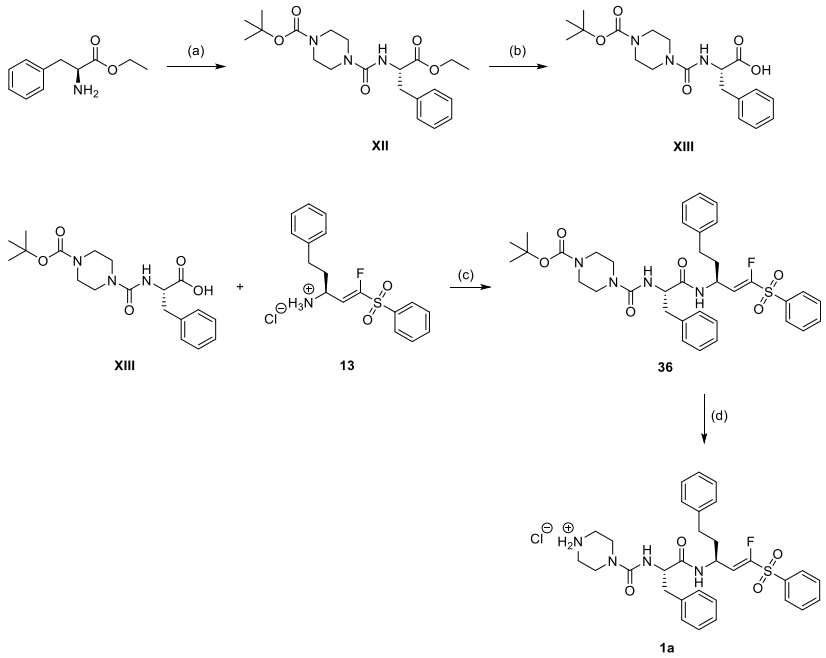
Chemistry. Compounds **2a−4l** (Figure 2) were synthesized using Horner−Wadsworth−Emmons (HWE) chemistry as the key step (Scheme 1). The required boc-protected amino-aldehydes (**VI−VIII**) were prepared using Weinreb chemistry. The appropriate phosphonates (**5−7**) were synthesized in three different ways (Scheme 2). The preparation of phosphonate **5** has been published previously.[12] Phosphonates **6** and **7** were obtained by fluorination of the respective nonhalogenated precursors (**IV**, **V**) with Selectfluor in the presence of KHMDS. The nonhalogenated precursors (**III−V**) were synthesized according to literature procedures.[28]

HWE olefination of the aldehydes (**VI−VII**) with the respective phosphonates (**5−7**) provided the corresponding boc-protected vinylsulfones (**8−12**) as mixtures of (*E*)/(*Z*)-isomers (Scheme 1), whereby the (*E*)-isomer was generally favored under the employed conditions. Overall yields ranged from 59 to 75%. The (*E*)-isomers were isolated by column chromatography in acceptable yields (41−63%) and used for the next steps. In the case of vinylsulfonate **9**, the (*Z*)-isomer was also isolated ((**Z**)-**9**, yield 22%). In the next step, the boc-group was removed using standard protocols, either 4 M HCl in dioxane or TFA in DCM, giving the amine building blocks **13−17** in quantitative yields (Scheme 2). These building blocks were subjected to peptide chemistry based on the boc strategy with TBTU/HOBt as the coupling reagent (Scheme 3). The desired compounds (**2a−4l**) were obtained after one or two coupling and deprotection steps.

The metabolites of compounds **1** and **2d**-(**H**), namely **1a**, **1b**, and **2l**, were synthesized using similar procedures to that described above. For the *N*-demethylated metabolite **1a** (Scheme 4), vinylsulfone **13** was coupled with **XIII**. After removal of the boc group, **1a** was obtained with a yield of 93%.

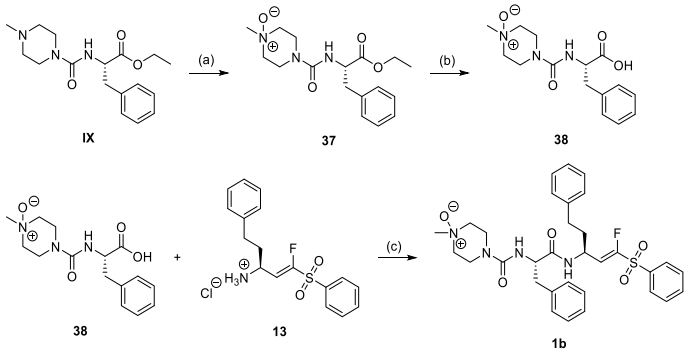
https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Scheme 4. Synthesis of the *N*-Demethylated Metabolite 1a[a]



[a]Reagents and conditions: (a) (1) triphosgene, DCM, 0 °C, 1 h, 85%. (2) 1-Boc piperazine, THF, rt, 18 h, 50%, (b) LiOH, THF, 0 °C to rt, 3 h, 74%, (c) HOBt, TBTU, DIEA, DCM, 0 °C to rt, 12 h, 45%, and (d) 4 M HCl in dioxane, rt, 1 h, 93%.

Scheme 5. Synthesis of the *N*-Oxidized Metabolite 1b[a]



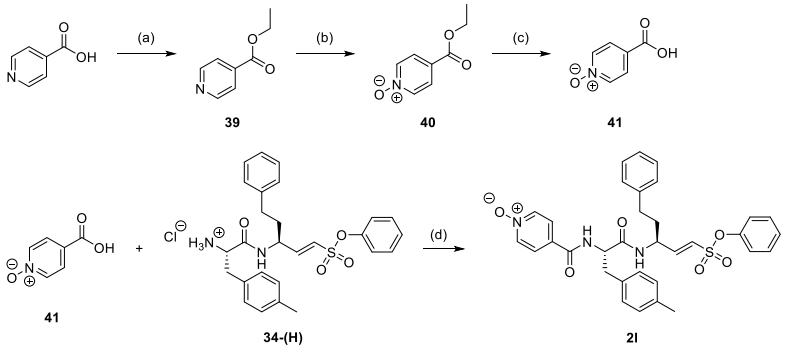
[a]Reagents and conditions: (a) *m*CPBA, DCM, 0 °C, 16 h, 52%; (b) LiOH, THF, 0 °C to rt, 3 h, 53%; and (c) HOBt, TBTU, DIEA, DCM/DMF, 0 °C to rt, 24 h, 23%.

In Scheme 5, the preparation of the metabolite **1b**, the *N*-oxide of compound **1**, is described. The synthesis started with compound **IX**, which was oxidized with *m*CPBA,[29] resulting in ester **37**, which was hydrolyzed with LiOH. Compound **38** was then coupled with vinylsulfone **13**, yielding compound **1b**.

The *N*-oxidized metabolite **2l** of compound **2d-(H)** was prepared as shown in Scheme 6. Isonicotinic acid was esterified and then oxidized with *m*CPBA.[29] The resulting compound **40** was hydrolyzed to **41** and then coupled with **34-(H)**, yielding the *N*-oxide **2l**.

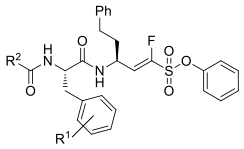
https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Scheme 6. Synthesis of the N-Oxidized Metabolite 2l of Compound 2d-(H)[a]



[a]Reagents and conditions: (a) ethanol, H_2SO_4, reflux, 24 h, 73%; (b) *m*CPBA, DCM, 0 °C, 16 h, 29%; (c) LiOH, THF, 0 °C to rt, 3 h, 90%; and (d) HOBt, TBTU, DIEA, DCM, 0 °C to rt, 24 h, 13%.

Table 1. Inhibition Data for Compounds 2a−2k[a]



	substitution		rhodesain		CatL			CatB
Cpd	R^1	R^2	$K_i/\mu M$[b]	$K_i^*/\mu M$[b]	$K_i/\mu M$[c]	$K_i^*/\mu M$[c]	SI[d]	$K_i/\mu M$
2a	H	4-Pyr	0.098[g]	0.015[g]	0.258[g]	0.060[g]	4	1.7
(*Z*)-**2a**	H	4-Pyr	0.525[e,g]		n.d.		n.d.	n.d.
2b	H	DHBD	0.045[g]	0.009[g]	n.d.	n.d.	n.d.	n.d.
2d	4-Me	4-Pyr	0.024[g]	0.003[g]	0.313[g]	0.078[g]	26	38%[f]
2e	4-Me	DHBD	0.098[g]	0.007[g]	0.348[g]	0.039[g]	6	14%[f]
2f	4-Me	3,5-F_2Ph	0.034[g]	0.005[g]	n.d.	n.d.	n.d.	n.d.
2g	3-Me	4-Pyr	0.094[g]	0.007[g]	0.266[g]	0.030[g]	4	50%[f]
2h	3-Me	DHBD	0.059[g]	0.010[g]	n.d.	n.d.	n.d.	n.d.
2i	3-Me	3,5-F_2Ph	0.152[g]	0.021[g]	n.d.	n.d.	n.d.	n.d.
2j	H	OBn	0.158[e,g]		n.d.		n.d.	n.d.
2k	H	*N*-MePip	0.108[e]		n.d.		n.d.	n.d.

[a]K_i^* denotes the dissociation constant of the high-affinity complex in the case of biphasic, time-dependent inhibition. [b]Calculated with method 1 (see the text). [c]Calculated with method 2 (Dixon equation).[31] [d]K_i^*(CatL)/K_i^*(rhodesain). [e]Calculated from IC_{50} value with the Cheng−Prusoff equation.[32] [f]% inhibition at 11 μM inhibitor concentration (single measurement). [g]Mean value of three independent assays; standard deviations less than 10%.

Enzyme Assays. Inhibition of rhodesain was tested with the fluorogenic substrate Cbz-Phe-Arg-AMC as described previously.[12,30] For the three series of compounds, the results from the fluorometric enzyme assays with rhodesain and the related mammalian enzymes cathepsin L and B are summarized in Tables 1−3.

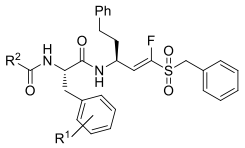
For benzyl and phenyl fluorovinylsulfones (cpds **3a−3i**, **4a−4l**), the progress curves for inhibition of rhodesain were found to be linear in all cases, indicating a fast reversible inhibition (exemplarily shown for compound **3d** in Figure 3A and for **4a** in Figure 3B). In order to confirm the competitive behavior of these compounds, IC_{50} values were measured at seven different substrate concentrations. IC_{50} values were found to increase linearly with ascending substrate concentration, showing competitive inhibition (Figure 3C,D). This was assumed to be the case for all compounds of the series. Consequently, IC_{50} values were converted to K_i values using the Cheng−Prusoff relationship (Tables 2 and 3).[32]

Reversibility was confirmed by dilution assays, that is, the enzyme was incubated with an excess of inhibitor (10-fold the IC_{50} concentration) to ensure full inhibition. Then, the mixture was diluted 100-fold to yield an inhibitor concentration of 0.1-fold the IC_{50} concentration. In the case of reversible inhibition, the enzyme activity should recover, whereas in the case of irreversible inhibition, it should not. For both compounds (**3d**, **4a**), enzyme activity recovers after dilution, whereas for the irreversible inhibitor **K11777** used as a control, enzyme activity does not recover (Figure 4A).

For vinylsulfonates (**2a−2k**), the progress curves for inhibition of rhodesain were not linear for most of the

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

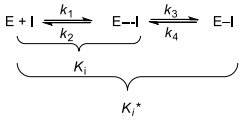
Table 2. Inhibition Data for Compounds 3a−3i



	substitution		rhodesain	CatL		CatB
cpd	R^1	R^2	$K_i/\mu M$	$K_i/\mu M$	SI[a]	$K_i/\mu M$
3a	H	4-Pyr	0.053[d]	0.226[d]	4	44%[b]
3b	H	DHBD	0.046[d]	n.d.	n.d.	n.d.
3c	H	3,5-F_2Ph	0.124[d]	n.d.	n.d.	n.d.
3d	4-Me	4-Pyr	0.015[d]	0.181[d]	12	35%[b]
3e	4-Me	DHBD	0.014[d]	0.076[d]	6	8%[b]
3f	4-Me	3,5-F_2Ph	0.029[d]	n.d.	n.d.	n.d.
3g	3-Me	4-Pyr	0.061[d]	0.122[d]	2	1.7[c]
3h	3-Me	DHBD	0.092[d]	n.d.	n.d.	n.d.
3i	3-Me	3,5-F_2Ph	0.380[d]	n.d.	n.d.	n.d.

[a]K_i(CatL)/K_i(rhodesain). [b]% inhibition at 11 μM (single measurement). [c]Single measurement. [d]Mean value of three independent assays; standard deviations less than 10%.

compounds (**2a−2i**), but showed time-dependency (exemplified for **2d** in Figure 5A). Inhibition by all vinylsulfonates was found to be competitive with respect to the substrate (Figure 5B). Time-dependent inhibition is typical for irreversible inhibitors but may also be observed for covalent reversible inhibition. In the case of irreversible inhibition, the progress curves reach a plateau value with the terminal enzyme activity, that is, the steady-state velocity of substrate turnover in the presence of the inhibitor, $v_s = 0$. For covalent reversible inhibition, time-dependent progress curves reflect a biphasic behavior with the terminal enzyme activity in the presence of the inhibitor $v_s \neq 0$, but with $v_s < v_i$ (v_i = the initial enzyme activity in the presence of the inhibitor). To distinguish between these two scenarios, dilution assays were performed (see above), exemplarily shown for inhibitor **2d** (Figure 4A). The results clearly indicate that vinylsulfonate **2d** is a reversible inhibitor, but dissociates significantly slower compared to compounds **4a** and **3d**, indicating a tight covalent reversible inhibition according to the following inhibition mechanism, with $K_i = k_2/k_1$ as the dissociation constant of the initial noncovalent enzyme−inhibitor (E···I) complex and K_i^* as the dissociation constant of the final covalent, high-affinity complex (E−I)[33]



For these inhibitors, the initial (v_i) and steady-state (v_s) velocities in the presence of the inhibitor as well as the pseudo-first order rate constants k_{obs} were determined for the different inhibitor concentrations by fitting the progress curves (shown for cpd **2d** in Figure 5A) to the slow-binding equation (off = offset)[33]

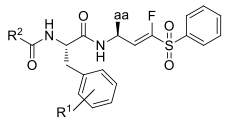
$$[P] = v_s \times t + \frac{v_i - v_s}{k_{obs}}[1 - \exp(-k_{obs} \times t)] + \text{off}$$

The k_{obs} values were replotted against the inhibitor concentrations [I] (Figure 5C) with the equation[33]

$$k_{obs} = k_4 + \left(\frac{k_3 \times [I]}{K_i^{app} + [I]}\right)$$

to yield the apparent dissociation constant K_i^{app} of the initial enzyme−inhibitor complex, as well as the rate constants k_3 and k_4 of the second inhibition step (Method 1). Because the compounds display competitive inhibition with respect to the substrate (Figure 5B), the K_i^{app} value was converted to K_i for the

Table 3. Inhibition Data for Compounds 4a−4l



	substitution			rhodesain	CatL		CatB
cpd	R^1	R^2	aa	$K_i/\mu M$	$K_i/\mu M$	SI[a]	$K_i/\mu M$
1	H	*N*-MePip	hPhe	0.190[c,e]	0.023[e]	0.1	0.47[d]
4a	H	4-Pyr	hPhe	0.032[c,e]	0.110[e]	3	3.13[d]
4b	H	DHBD	hPhe	0.012[e]	0.033[e]	3	4.81[d]
4c	H	3,5-F_2Ph	hPhe	0.035[e]	n.d.	n.d.	n.d.
4d	4-Me	4-Pyr	hPhe	0.008[e]	0.115[e]	14	2.08[d]
4e	4-Me	DHBD	hPhe	0.005[e]	0.023[e]	5	34%[b]
4f	4-Me	3,5-F_2Ph	hPhe	0.010[e]	n.d.	n.d.	n.d.
4g	3-Me	4-Pyr	hPhe	0.025[e]	0.049[e]	2	0.35[d]
4h	3-Me	DHBD	hPhe	0.035[e]	n.d.	n.d.	n.d.
4i	3-Me	3,5-F_2Ph	hPhe	0.329[e]	n.d.	n.d.	n.d.
4j	H	*N*-MePip	Met	0.360[e]	0.173[e]	0.5	1.20[d]
4k	H	4-Pyr	Met	0.056[e]	0.577[e]	10	6.17[d]
4l	H	*N*-MePip	Cys(Bn)	0.630[e]	0.628[e]	1	8.62[d]

[a]K_i(CatL)/K_i(rhodesain). [b]% inhibition at 11 μM. [c]Ref 12 ; n.d.: not determined. [d]Single measurement. [e]Mean value of three independent assays; standard deviations less than 10%.

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

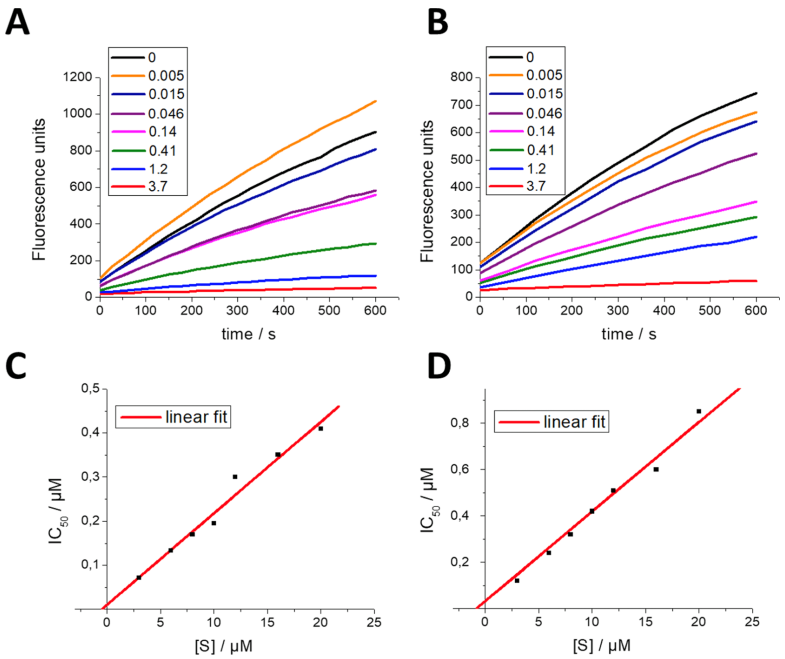


Figure 3. (A) Example for progress curves for inhibition of rhodesain by compound **3d**. Inhibitor concentrations in μM. (B) Example for progress curves for inhibition of rhodesain by compound **4a**. Inhibitor concentrations in μM. (C) Dependence of inhibition potency (IC_{50} values) on substrate concentration for inhibition of rhodesain by compound **3d**. Increasing IC_{50} values at ascending substrate concentrations show competitive inhibition. The K_i value is obtained as the intercept of the regression line with the y-axis (K_i = 25 nM). (D) Dependence of inhibition potency (IC_{50} values) on the substrate concentration for inhibition of rhodesain by compound **4a**. Increasing IC_{50} values at ascending substrate concentrations prove competitive inhibition. The K_i value is obtained as the intercept of the regression line with the y-axis (K_i = 47 nM).

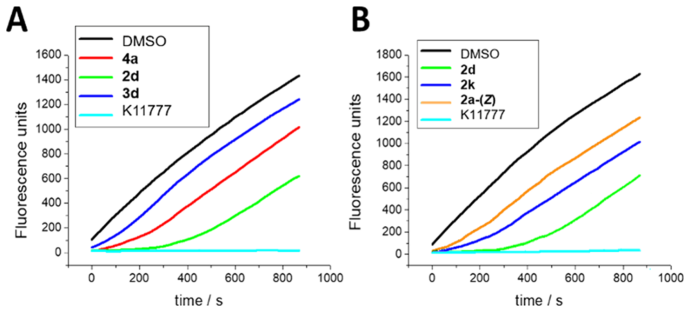


Figure 4. (A) Dilution assays show reversible inhibition of rhodesain (see the text). For all compounds, enzyme activity did recover after dilution, except for the irreversible vinylsulfone **K11777** (control). Compound **3d** showed faster reversibility than compound **4a**. Vinylsulfonate **2d** can be considered a tight-binding, slowly reversible inhibitor. (B) Dilution assays of several compounds of the series of vinylsulfonates. In the case of compound **2d**, which showed two-step inhibition in the enzyme assay, enzymatic activity recovers slower compared to vinylsulfonates **2k** and **2a**-(*Z*), which did not show biphasic behavior. The irreversible inhibitor **K11777** was used as a control.

initial inhibitor complex with the Cheng–Prusoff equation.[32]

$$K_i^* = \frac{K_i}{1 + \left(\frac{k_3}{k_4}\right)}$$

The K_i^* value as the dissociation constant of the final complex was calculated from

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

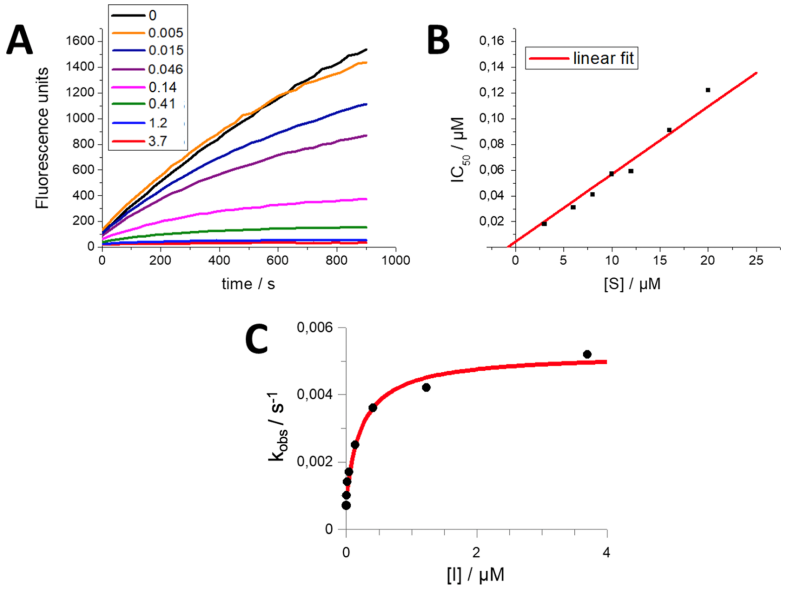


Figure 5. (A) Example for progress curves for inhibition of rhodesain by compound **2d**. Curve shape indicates time-dependent inhibition. Inhibitor concentrations in μM. (B) Dependence of inhibition potency (IC_{50} values) from the substrate concentration for inhibition of rhodesain by compound **2d**. Increasing IC_{50} values at ascending substrate concentrations show competitive inhibition. The K_i value is obtained as the intercept of the regression line with the y-axis (K_i = 19 nM). (C) Plot of the rate constants k_{obs} for the progress curves of compound **2d** from Figure 5A as a function of the inhibitor concentration. k_4 is obtained from the intercept of the regression curve with the y-axis. The maximum value of k_{obs} at infinite inhibitor concentration provides the sum of k_3 and k_4. The concentration of inhibitor yielding a half-maximal value of k_{obs} is equal to K_i^{app}.

Both dissociation constants, K_i and K_i^*, were also calculated by fitting the initial (v_i) and steady-state (v_s) velocities, respectively, against the inhibitor concentrations using the Dixon equation (Method 2)[31]

$$\frac{v_{i,s}}{v_0} = \frac{[I]}{1 + \left(\frac{[I]}{K_i^{(*)app}}\right)}$$

K_i^{app} was obtained from fitting v_i against [I] and $K_i^{(*)app}$ from fitting v_s against [I] (v_0 is the substrate turnover velocity in the absence of the inhibitor), and both constants were converted into K_i and K_i^* with the Cheng–Prusoff equation.[32]

A comparison of the dissociation constants for inhibition of rhodesain by inhibitors **2a–2i** obtained by both methods is shown in Table 4.

The data show that both methods yield similar dissociation constants for the initial low-affinity as well as for the final high-affinity complex proving that both methods are reliable. For all compounds, the rate constant of dissociation of the final complex (k_4) was found to be slower than the rate constant of association (k_3), indicating tight binding of these inhibitors. Interestingly, vinylsulfonates (*Z*)-**2a**, **2k**, and **2j** did not show time dependency of inhibition, but linear progress curves similar to the benzyl and phenyl vinylsulfones. Dilution assays for these three vinylsulfonates [(*Z*)-**2a**, **2k**, and **2j**] indicated a significantly faster recovery of the enzyme activity compared to the time-dependent inhibitor **2d** [shown for compound (*Z*)-**2a** and **2k** in Figure 4B].

Table 4. Inhibition Data and Kinetic Parameters k_3 and k_4 for Time-Dependent Inhibition of Rhodesain by Compounds 2a–2i

	method 1 (slow-binding equation)[a]				method 2 (Dixon equation)[b]	
Cpd	K_i/μM	K_i^*/μM	k_3/s^{-1}	k_4/s^{-1}	K_i/μM	K_i^*/μM
2a	0.098	0.015	0.055	0.010	0.110	0.008
2b	0.045	0.009	0.045	0.011	0.060	0.005
2d	0.024	0.003	0.075	0.010	0.022	0.002
2e	0.098	0.007	0.065	0.005	0.124	0.002
2f	0.034	0.005	0.049	0.009	0.052	0.004
2g	0.094	0.007	0.062	0.005	0.089	0.004
2h	0.059	0.010	0.074	0.015	0.068	0.006
2i	0.152	0.021	0.055	0.010	0.155	0.024

[a]Calculated with slow-binding equation (see the text). [b]Calculated with the Dixon equation (see the text).[31] Mean values of three independent assays, standard deviations less than 10%.

To further quantify the degree of reversibility for the different warheads, compounds **2d**, **3d**, and **4d** with identical recognition units were subjected to a dialysis experiment.[12] Here, rhodesain was incubated with an excess of inhibitor (10-fold the IC_{50} concentration) to ensure full inhibition. Then, the mixture was dialyzed against a continuous flow of enzyme buffer using a 3.5 kDa cut-off dialysis tubing.[34] Samples were taken after 10, 30, 60, and 120 min and analyzed for their residual enzyme activity with the standard fluorometric assay. The results are presented as the

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322–12358

fractional activity of uninhibited rhodesain, which was subjected to dialysis in the same experiment (Figure 6). Compound **3d**

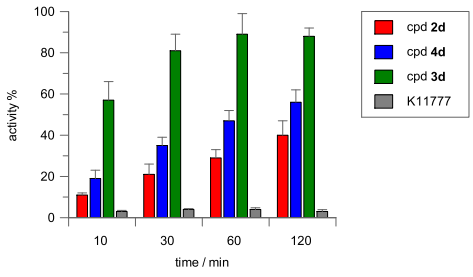


Figure 6. Dialysis experiment for compounds **2d**, **3d**, **4d**, and **K11777**. Rhodesain was incubated with the inhibitors at a concentration of 10-fold the respective IC_{50} value. The mixture was dialyzed against a continuous flow of buffer. Samples were taken after 10, 30, 60, and 120 min and the substrate was added to measure the residual enzyme activity. **K11777** was used as an irreversible control.

showed the fastest reversibility (89% recovery after 1 h), which is in accordance with findings from the dilution assay. Compound **4d** dissociated significantly slower (47% recovery after 1 h). Vinylsulfonate **2d** displayed the slowest reversibility (29% after 1 h), which supports the tight-binding nature of the inhibition with the formation of a very stable high-affinity complex. **K11777** was used as an irreversible control, showing no recovery of the enzymatic activity.

The most potent compounds of each series were also tested against the human enzymes CatL and CatB. For the tested vinylsulfonates (**2a**, **2d**, **2e**, and **2g**), a biphasic behavior was also observed for inhibition of CatL. In these cases, K_i and K_i^* values were calculated with the Dixon equation (Table 2).[31] For inhibition of CatB, no biphasic behavior was observed for these compounds (**2a**, **2d**, **2e**, and **2g**). This can be explained with the low affinity of the compounds for CatB. Notably, compounds **2d**, **2e**, and **2g** showed only very weak inhibition at concentrations of up to 11 μM.

Discussion of SAR. Comparison of the K_i values for rhodesain of the starting compound **1** (K_i = 190 nM) with the P3-modified analogues **4a**–**4c** demonstrates that the exchange of *N*-methyl piperazine against aromatic residues with no or only limited basicity significantly enhances inhibition potency (e.g., cpd **4b** with K_i = 12 nM). This is also reflected by the scores obtained from docking for these compounds, for example, for **4b** noncovalent affinity as well as the stability of the covalent complex are predicted to be higher (Table S1). Additionally, compounds **4a**–**4c** show improved selectivity for rhodesain over the human cathepsins. Lead compound **1** shows a higher affinity for human CatL (K_i = 23 nM) than for rhodesain (K_i = 190 nM), while P3-modified compounds **4a** and **4b** slightly favor inhibition of rhodesain (3-fold). Furthermore, selectivity for rhodesain over CatB is dramatically enhanced by these structural variations. This is a remarkable improvement (2-fold) compared to the weak selectivity of starting compound **1**. A further increase in potency for rhodesain is observed for compounds with an additional 4-methyl substituent at the phenylalanine aromatic ring (**4d**–**4f**), with compound **4e** (K_i = 5 nM) being the most potent inhibitor in the series of phenyl vinylsulfones. As suggested from noncovalent docking, this increase in potency may be directly attributed to additional hydrophobic interactions of the 4-methyl group with lipophilic residues lining the S2 pocket (Ala208, Leu160, Figure 7), which is also reflected in

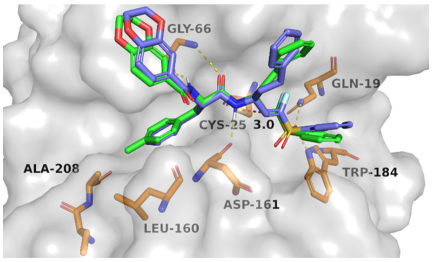


Figure 7. Noncovalent binding mode of compound **4b** (light blue, FlexX score: −31.2, Hyde score: −10.0 kcal mol^{-1}) and **4e** (green, FlexX score: −32.5, Hyde score: −11.7 kcal mol^{-1}) as predicted by FlexX. The electrophilic β-carbon of the fluorovinylsulfone warhead comes in close proximity to the nucleophilic sulfur of Cys25 (3.0 Å). The sulfone group forms hydrogen bonds with Gln19 and Trp184. The recognition unit forms hydrogen bonds to Gly66 and Asp161. The additional 4-methyl group of **4e** at the phenylalanine aromatic ring points into a hydrophobic cavity between Ala208 and Leu160, resulting in additional lipophilic interactions compared to compound **4b**.

higher scores compared to compounds **4a**–**4c** (Table S1). Additionally, introduction of the 4-methyl substituent further increases selectivity against CatL (14-fold for **4d**). In contrast, incorporation of a 3-methyl substituent (**4g**–**4i**) does reduce potency and selectivity compared to compounds **4a**–**4c**. In particular, the combination of the 3-methyl substituent and the 3,5-difluorophenyl moiety (**4i**) is not advantageous: compound **4i** displayed significantly reduced potency (K_i = 329 nM). The considerable difference compared to compound **4e** with a 4-methyl substituent (K_i = 10 nM) cannot be explained exclusively with the slightly reduced docking scores (Table S1).

Within the set of benzyl vinylsulfones (**3a**–**3i**), the SAR relationships are very similar compared to phenyl vinylsulfones (**4a**–**4i**). Consequently, compounds **3d** and **3e** with a 4-methyl substituent and a 4-pyridyl or DHBD moiety, respectively, are the most potent and most selective inhibitors (**3d**: K_i = 15 nM, 12-fold selectivity over CatL). Remarkably, compound **3d** showed even lower activity against CatB (35% inhibition at 11 μM) compared with its counterpart **4d**. Again, compound **3i** with a 3-methyl substituent and 3,5-difluorophenyl residue shows noticeably lower potency (K_i = 380 nM). Comparison of inhibitory potency for benzyl and phenyl vinylsulfones with the same recognition unit (e.g., **3a** vs **4a**) reveals that the benzyl substituent at the warhead has a negative influence on inhibition potency. In general, benzyl vinylsulfones have slightly increased K_i values compared to their phenyl counterparts.

These findings cannot be correlated with scores from docking, which generally predicted a higher affinity for compounds with a benzyl group at P1′. Computational conclusions to explain these differences between theory and experiment would be desirable, but are difficult because the differences in K_i values result from an increase in binding energy of less than 1 kcal mol^{-1} (e.g., **3d** K_i = 0.015 μM, i.e., ΔG = −11.10 kcal mol^{-1} vs **4d** K_i = 0.008 μM, i.e., ΔG = −11.49 kcal mol^{-1}), which is out of the scope of even high-level quantum chemical computations. Possibly, entropic

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

effects due to the higher flexibility of the benzyl group in the noncovalent enzyme inhibitor complex may contribute.

The results from dilution and dialysis assays revealed that benzyl vinylsulfones are faster reversible than their phenyl counterparts, which may be explained with a lower reaction energy for the covalent bond-forming step, resulting in a weaker covalent bond.

Comparison of the K_i values of starting compound **1** (K_i = 190 nM) and the corresponding vinylsulfonate **2k** (K_i = 108 nM) demonstrates a nearly 2-fold improvement in affinity for rhodesain. The increase in potency may be partly attributed to the enlarged substituent at P1′. The results from noncovalent docking suggest that the vinylsulfonate moiety extends further into the S1′ pocket and forms nonpolar interactions with Trp184 (Figure 8), which is also reflected by the slightly

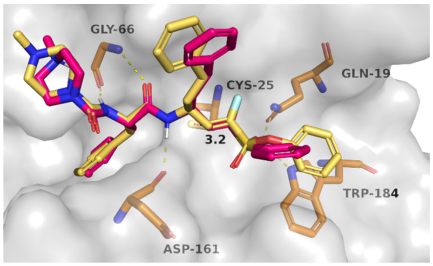


Figure 8. Overlay of noncovalent docking poses of compound **1** (magenta, FlexX score: −27.9, Hyde score: −6.2 kcal mol^{-1}) and compound **2k** (gold, FlexX score: −28.4, Hyde score: −8.8 kcal mol^{-1}). The enlarged vinylsulfonate moiety of compound **2k** extends further into the S1′ pocket and the aromatic ring can form additional hydrophobic interactions with Trp184 (π−π-stacking interactions).

enhanced scores. Again, substitution of the *N*-methyl piperazine against aromatic residues (cpds **2a**−**2i**) further improves affinity for rhodesain. Additionally, and in contrast to compound **2k**, compounds **2a**−**2i** show a biphasic, time-dependent inhibition mechanism.

Dilution and dialysis experiments demonstrated reversibility for compound **2d** (Figures 4 and 6), which is, however, significantly slower compared to vinylsulfonate **2k**, which supports the tight-binding properties of **2d**. The related vinylsufones (**3d** and **4d**) do not show two-step inhibition, indicating that a combination of the properties of the fluorovinylsulfonate warhead with appropriate noncovalent interactions is required for tight binding. This is further supported by vinylsulfonate **2j** with a Cbz group at P3, which also does not show time-dependent inhibition. Among the tight-binding compounds, **2d** shows the highest potency for rhodesain with a dissociation constant of the initial encounter complex K_i = 24 nM, and with a high-affinity complex formed in the second step (K_i^* = 3 nM). Especially, the covalent complex is predicted to be very stable (Figure 9B). This is in excellent accordance with the high scores obtained from both docking approaches for compound **2d** (Table S1). Notably, compound **2d** showed good selectivity over the human cathepsins, with 26-fold selectivity against CatL and virtually no activity against CatB (38% at 11 μM), which is a significant improvement compared to the starting compound **1**.

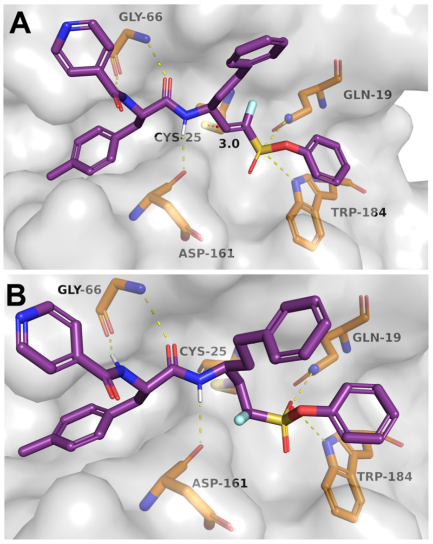


Figure 9. (A) Non-covalent docking pose of compound **2d** predicted with FlexX (FlexX score: −34.3, Hyde score: −38). (B) Geometry of the covalent complex between rhodesain and compound **2d** as predicted with DOCKTITE [score (affinity ΔG, kcal mol^{-1}): −6.2, DSX score: −195.5]. The combination of the vinylsulfonate aromatic ring extending further into the S1′ pocket and forming lipophilic interactions with Trp184, additional hydrophobic interactions of the 4-methyl group at phenylalanine aromatic ring with the S2 pocket and the preferred 4-pyridyl group at P3 results in a high affinity of compound **2d** for the binding site. This is reflected from higher scores obtained from both docking approaches compared to its analogues (Table S1).

Assuming that inhibitor binding to CatL and CatB is similar to rhodesain binding and that the binding mode is not largely altered for the compounds presented herein compared to **K11777**, which is also suggested by the docking (Figures 7−9), the underlying molecular mechanisms for selectivity can be explained by analysis of the known X-ray structures. Rhodesain shares an overall sequence identity of 44.7% and similarity of 59.1% with CatL and 27.9% identity and 47.4% similarity with CatB, as well as a highly similar fold with an C_α-RMSD of 1.35 and 2.15 Å, respectively, known from crystal structures available in the PDB[35] (entries, rhodesain: 2p7u,[11] CatL: 2xu1,[36] and CatB: 3ai8).[37] For residues forming the binding site (defined as all amino acids within 6 Å of reference ligand **K11777**) identity/similarity even increase to 59.1%/70.2% for CatL and 40.4%/49.1% for CatB. Nevertheless, slight structural differences of residues forming the S1−S3 sites can be observed, explaining the selectivity profile of the compounds under investigation. Additionally, CatB, divergent from CatL and rhodesain, contains a so-called occluding loop (residues 104−124), which is crucial for this enzyme's exopeptidase activity.[38] This loop structure closes upon the S′ sites. Neither phenyl- nor benzylfluorovinylsulfone nor fluorovinylsulfonate moieties of the compounds within this study reach far enough toward this sites to form interactions, but—in contrast—may even cause clashes with the residues of CatB (Figure 10A). Further focusing on the S1 site reveals one residue of the CatB occluding loop, which provides

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

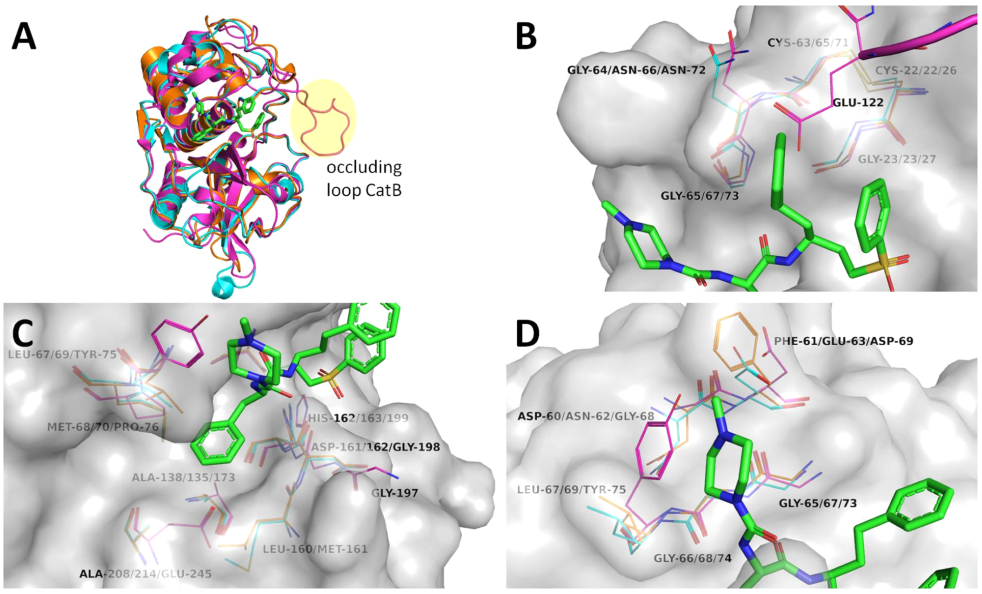


Figure 10. Structure comparison of rhodesain (orange carbon atoms, pdb entry 2p7u), CatL (cyan carbon atoms, pdb entry 2xu1), and CatB (magenta carbon atoms, pdb entry 3ai8) with inhibitor **K11777** shown with green carbon atoms. Residue enumeration is rhodesain/CatL/CatB. For (B–D), the rhodesain surface is shown in gray for orientation. (A) Whole protease structures depicted as cartoons with the CatB occluding loop highlighted. (B) Close view of S1 site residues. CatB-unique occluding loop shown in the upper right with Glu122 as a selectivity determining feature over CatB. (C) Close view of S2 site residues. (D) Close view of S3 site residues reveals acidic Glu63 and Asp69 in cathepsins to form ionic interactions with the basic center of compound **K11777**, which is absent in compounds with more favorable selectivity profiles.

an explanation for selectivity (Figure 10B). Glu122 reaches toward the S1 pocket, not only making this site more polar, but also narrower, clashing with the homophenylalanine moiety of the inhibitors and thereby causing selectivity over CatB. Within the S2 pocket, the largest differences were found for CatB as well (Figure 10C). CatB Gly197 shows a different orientation than the corresponding residues Leu160 in rhodesain and Met161 in CatL. Additionally, Ala208/214 (enumeration is rhodesain/CatL) is exchanged to Glu246, Met69/70 to Pro76, and Leu67/69 to Tyr75. These differences all together result in a more open and polar pocket within CatB, which leads to a higher affinity for rhodesain and CatL for inhibitors carrying (3- or 4-methyl)-phenylalanine moieties. Additionally, the selectivity for rhodesain over CatL introduced by the 4-methyl substitution (compounds **2d**, **3d**, and **4d**) is likely to be caused by non-polar interactions with Leu160 being more favorable compared to the CatL Met161, which is slightly more distant and potentially impaired in its flexibility upon binding. The most significant improvement in selectivity was the displacement of *N*-methyl piperazine by 4-pyridine or—less pronounced—by DHDB (compounds **2d**, **2e**, **3d**, **3e**, **4d**, and **4e**). Within the S3 site, acidic amino acids Glu63 and Asp69 are found in CatL and CatB, respectively, while in rhodesain Phe61 sits at the corresponding position (Figure 1D). By the removal of the positively charged *N*-methyl piperazine, the loss of potential ionic interactions only with the off-target cathepsins is, therefore, likely to improve the selectivity profile of these compounds. Additionally, the Tyr75 residue in CatB compared to Leu in rhodesain and CatL results in a smaller S3 pocket and an enhanced selectivity over CatB, too.

In contrast to all other compounds in this study, the geometry of the double bond of compound (**Z**)-**2a** has a (*Z*)-configuration. Because (*Z*)-isomers are obtained as side products from HWE olefination, it was obvious to explore the influence of the geometry of the double bond on inhibitory potency. Compared to the (*E*)-isomer **2a** (K_i^* = 15 nM), compound (**Z**)-**2a** (K_i = 525 nM) shows markedly reduced potency and forms no high-affinity complex, indicating that the (*Z*)-configuration of the double bond is not favorable.

Compounds with a modified P1 residue (**4j**–**4l**) showed reduced potency for rhodesain compared to the counterparts with homophenylalanine, which was already anticipated from non-covalent docking scores. Nevertheless, compound **4k**, with methionine at P1 and the 4-pyridyl moiety at P3, is still a potent inhibitor of rhodesain (K_i = 56 nM).

MS Analysis. Covalent protein–ligand interactions were verified by ESI MS for compounds **2a**, (**Z**)-**2a**, and **2j** and MALDI-TOF MS for **K11777**, **1**, **2a**, **2j**, and **2k**.

In all three cases (**2a**, (**Z**)-**2a**, and **2j**), the protein–ligand adduct resulting from Michael addition was detected in the ESI mass spectra. The observed mass shift corresponds to the mass of the inhibitor [exemplified for compound **2a** in Figure 11; for compounds (**Z**)-**2a** and **2j**, see Figure S2], which shows covalent bond formation.

To further elucidate possible differences in binding mode related to compound variability, especially to confirm the

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322–12358

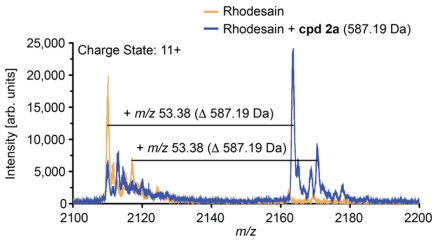


Figure 11. Intact protein−ligand adducts obtained by mass spectrometry for compound **2a**. The observed shift corresponds to the mass of the inhibitor considering the charge state of the protein ($m/z = 11^+$).

covalent inhibition, a representative subset of compounds covering the majority of synthesized modifications was selected and analyzed by MALDI-TOF MS. The aim here was to analyze whether these modifications impede covalent bond formation between the catalytic thiolate and the vinyl moiety. ESI MS is a softer method in terms of transition to the gas phase, where non-covalent adducts can be found as well as covalent adducts. Depending on the matrix, MALDI-TOF MS can discriminate between covalent and non-covalent binding: the formation of non-covalent adducts can be suppressed by using an acidic matrix, which was attributed to the disruption of salt bridges and the subsequent destabilization of the non-covalent protein−ligand complex.[39−41] Therefore, MALDI-TOF MS was applied to further investigate the binding mode.[40,42,43]

In terms of warhead modification, vinyl sulfone **K11777**, α-fluorovinylsulfone **1**, and α-fluorovinylsulfonates **2a**, **2j**, and **2k** were evaluated. Furthermore, the selected subset differed in their substituents in the P3 position, namely isonicotinyl amide (**2a**), benzylcarbamate (**2j**), and 4-methylpiperazine-1-carboxamide (**K11777**, **1**, and **2k**), while sharing the preceding PhehPhe motif in P2 and P1, respectively. **K11777** was used as a known, structurally equivalent control compound for its analogous, though covalent irreversible binding mode. A non-covalent inhibitor of rhodesain (**42**) carrying a fluorodinitrobenzene moiety as an aromatic electrophile was chosen as a control substance for its different inhibition mode (Figure 12) compared to the vinylsulfones. The compound forms a tight π-complex with the catalytic thiolate, the adduct mass of which was detectable in ESI MS experiments with rhodesain.[44]

For all evaluated α-fluorovinyl analogues (**1**, **2a**, **2j**, and **2k**), covalent adducts with rhodesain were observed by MALDI-TOF MS. The resulting spectra consisted of a broad peak corresponding to the protein (ca. 23.3 kDa),[45−47] and a second peak corresponding to the covalent protein−inhibitor adduct (ca. 23.8 kDa) that shows a mass shift corresponding to the mass of the respective compound (Figure 12). The observed multiplicity of the peaks was attributed to additions of matrix molecules [m(sinapinic acid) = 224 Da] to rhodesain, as similar phenomena are described in the literature.[47] The four fluorovinyl derivatives (**1**, **2a**, **2j**, and **2k**) behaved like the irreversible control substance **K11777**, while the non-covalent inhibitor (**42**) did not show any detectable adduct signals under the evaluated conditions. Apart from that, an adduct of **42** was found with ESI MS as published previously.[39,40,44] The identical behavior of **K11777** and the α-fluorovinylsulfones/-sulfonates in the MALDI-TOF MS experiments could be observed using two different matrices, sinapinic acid and a mixture of α-cyano-4-hydroxycinnamic acid and 2,5-dihydroxybenzoic acid,[43] respectively, the latter data are shown in the Figure S1. These findings, combined with the results from ESI MS and the dialysis

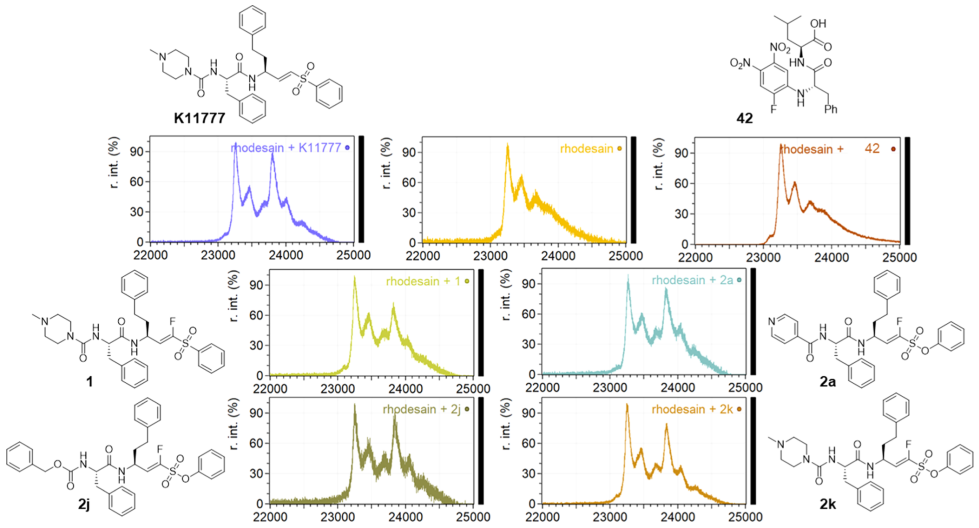


Figure 12. MALDI-TOF mass spectra of rhodesain (ca. 23 kDa) in the presence of different inhibitors (10-fold molar excess of inhibitor to protein). Sinapinic acid was used as the matrix substance. The figures show the relevant range of the spectrum to display $[M + H]^+$, which were baseline corrected.

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

and dilution assays, clearly show that the α-fluorovinylsulfones and -sulfonates are indeed covalent reversible inhibitors of rhodesain.

Comparison of the Inhibition Mechanisms of K11777 and 1 by QM/MM Modeling. In our previous study,[12] the computed reaction energies of **K11777** and **1** differed only marginally so that the computational results could hardly explain the transition from irreversible to reversible inhibition. In that study, both reaction pathways had been calculated starting from the X-ray structure of the covalent enzyme−inhibitor complex of K1177 with rhodesain (PDB: 2p7u) going backwards from the covalent toward the non-covalent complex. To generate the reaction path for **1**, the hydrogen atom had been substituted for fluorine in the X-ray structure. It had been necessary to start the computations from the X-ray structure of the covalent complex because no experimental information on the non-covalent complex was available. Moreover, such procedures had been very successful to explain the stereo- and regioselectivity of various epoxide- and aziridine-based inhibitors and to predict improved inhibitors.[48−50] However, such approaches can be error-prone in cases when the geometries of the covalent and non-covalent enzyme−inhibitor complexes differ largely, especially when substitution leads to large differences in the non-covalent complex.

To obtain more reliable insights into possible differences between the non-covalent enzyme-inhibitor complexes of **K11777** and **1**, at first, non-covalent structures starting from the X-ray structure of **K11777** with rhodesain (PDB: 2p7u) were calculated. For **1**, the hydrogen atom in the α-position was substituted by fluorine. These structures then were used as starting points for MD simulations (10 times 10 ns sequences for each compound). Exemplarily, Figure 13 presents the

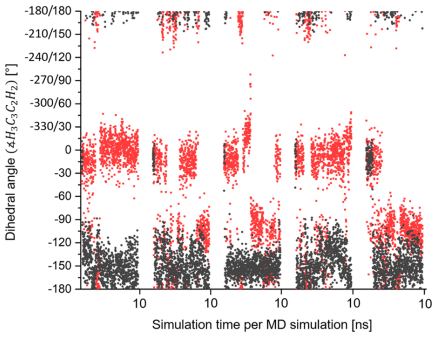


Figure 13. Variation of the dihedral angle $\angle H^3C^3C^2H^2$ along the MD simulations (sequences of 10.0 ns, respectively) for **K11777** with X = H (in red) and the fluorinated vinylsulfone with X = F (in black). Figure 13 defines the dihedral angle and gives the corresponding orientation of the warhead in the active site. Please note that $\angle H^3C^3C^2H^2 = 150°$ is equal to $\angle H^3C^3C^2H^2 = -210°$.

fluctuations in the dihedral angle $\angle H^3C^3C^2H^2$ (see Figure 14 for definition) for five different MD samplings. A variation of $\angle H^3C^3C^2H^2$ from 0 to $\pm 180°$ describes the rotation of the inhibitor around the C^3C^2 single bond, adjacent to the double bond. Due to this rotation, the C^1X bond (X = H1, F) moves from the Gly23-oriented side to the opposite side (Figure 14). For **K11777**, mainly values between −40 and 40° (−320°) were found for the $\angle H^3C^3C^2H^2$ angle. In the following, we denominate alignments with $-40° < \angle H^3C^3C^2H^2 < 40°$ (Figure 14) as "H-orientation". For **1**, $\angle H^3C^3C^2H^2$ adopted values between −120 and −190° (170°), that is, the warhead of the compound oscillates around a position, where the F-atom is oriented toward the backbone NH groups of Cys25, Ser24, and the side-chain NH_2 group of Gln19 (Figure 14b), but mainly adopts positions with $\angle H^3C^3C^2H^2 > -180°$. In the following, geometries with $-190° < \angle H^3C^3C^2H^2 < -120°$ are denominated "F-orientation".

During the MD simulation, **K11777** populated both H- (≈80%) and F-orientation (≈20%) indicating that both alignments are energetically quite similar. In contrast, **1** predominantly adopted the F-orientation due to attractive interactions with the NH backbone group of Cys25, the side chain NH_2 group of Gln19 (known as the oxyanion hole of rhodesain), and the NH backbone group of Ser24. Similar variations in the structures upon fluorination were reported in the literature).[51,52]

While the orientation of the recognition unit (in Figure 14 given in purple) remained fixed in the torsional motion around $\angle H^3C^3C^2H^2$, the relative positions between the attacked double bond of the inhibitor and the involved side chains of Cys25 and His162 changed drastically as shown in Figure 15. In both cases, a syn-addition takes place because the thiolate group of Cys25 and the protonated imidazole ring of His162 are found to be on the same side of the double bond (Figure 15), but for the H-orientation, the CS bond of the Cys25 side chain is orthogonally oriented with respect to the C^1C^2 double bond while it is oriented in parallel for the F-orientation. The position of the protonated imidazole ring of His162 also changed accordingly.

To calculate the influence of the different orientation of the warhead of **K11777** and **1** on the inhibition mechanism (Figure 1), the corresponding reaction paths were computed. However, due to the large number of local minima for the non-covalent complex (the reactant) and the covalent complex (the product) in combination with the roughness of the underlying potential energy surface (PES), the appropriate picture of the inhibition reaction is not that of a single reaction path with one starting point, one transition state, and one product, but that of a very rough PES with various energetically similar pathways starting from and ending at slightly different reactants and products. To achieve representative pathways, we performed geometry optimizations starting from two selected snapshots of the MD simulation for each inhibitor and computed the reaction paths starting from the obtained minima. The geometrical parameters of the respective minima are given in Table S2.

To characterize the reaction mechanism, we first computed two-dimensional scans using the distances $R(S_{cys}-C_2)$ and $R(H_{His}-C_1)$ as main reaction parameters (Figure 14). The resulting transition states served as starting points for subsequent IRC (intrinsic reaction paths) simulations,[53] the course of which generally gives good insights into the reaction mechanism (reaction barrier, reaction energy, and structural changes in the course of the reaction). More details are given in the Experimental Section. In Figure 15, representative IRCs for **K11777** starting from the H-orientation and **1** starting from the F-orientation are compared. Figure 16 also sketches the geometry variations along the IRCs. Further information on the geometries is given in Table S3. Additional paths starting from other reactant minima showing similar shapes can also be found in the Supporting Information (Figures S4 and S7).

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

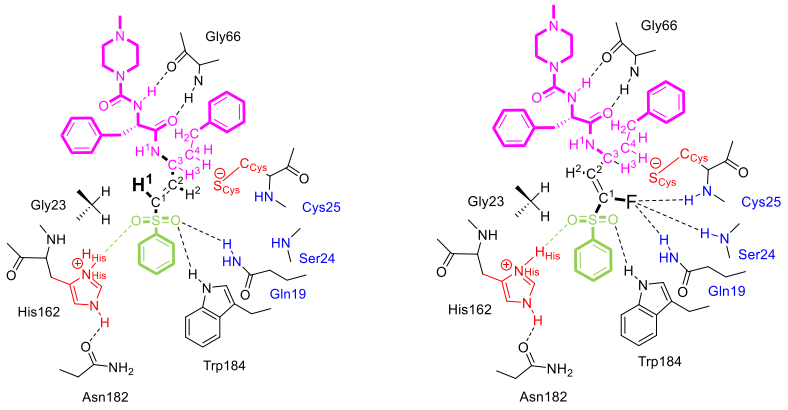


Figure 14. Sketch of the predominant orientation of the warheads (bold) in the active site during the MD simulations given in Figure 13. (a) H-orientation: predominant conformations of **K11777** (X = H) with $-40° \leq \angle H^3C^3C^2H^2 \leq 40°$ (b) F-orientation: predominant conformations for the fluorinated vinylsulfone **1**, with $-190° \leq \angle H^3C^3C^2H^2 \leq -120°$. For more information, see the text.

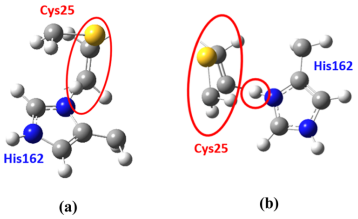


Figure 15. Relative orientation of Cys25 and His162 moieties with respect to the double bond of the vinylsulfone group attacked by Cys25, (a) for **K11777** (X = H) in the H-orientation, and (b) for the fluorinated vinylsulfone (X = F) in the F-orientation.

Figure 16 shows considerable differences in the IRCs of both compounds. While the reaction pathway for the inhibition of rhodesain by **K11777** passes over a barrier of about 12 kcal mol^{-1} and has an exothermic reaction energy of about −20 kcal mol^{-1}, we predict a thermoneutral reaction for the inhibition reaction of **1** with a very high barrier of 25 kcal mol^{-1}. The small reaction barrier and the strong exothermicity nicely explain why **K11777** is an efficient irreversible inhibitor.

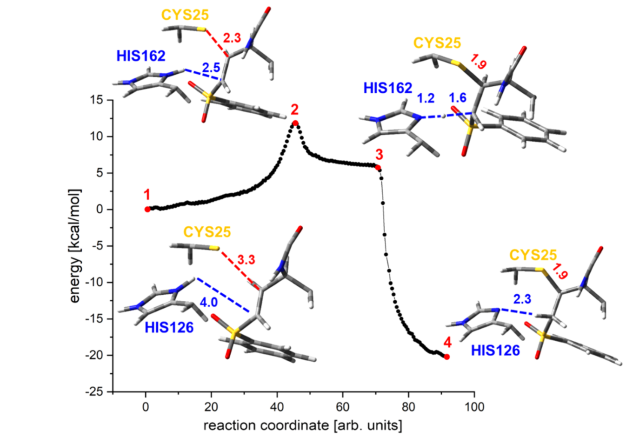
The different shapes of the reaction pathways presented in Figure 16 may result from the different orientation of the olefin in the active site or from changes in the electronic structure due to the fluorine substitution in **1**. To obtain more insights, the reaction of **K11777** from the F-orientation was computed. For this purpose, we first computed the energy profile of the torsional motion of the C^1C^2 double bond around the dihedral angle $\angle H^3C^3C^2H^2$ (Figure S3). As expected from the MD simulations, for **K11777**, the energy difference between the H- and the F-orientation was found to be small (1−2 kcal mol^{-1}) so that the reaction could take place from both orientations. However, the two-dimensional scan starting from the F-orientation predicted barriers of about 33 kcal mol^{-1} and nearly isothermal reactions (ΔE_{reac} = −3 kcal mol^{-1}). This indicates that the different starting orientations are leading forces for the observed differences. Attempts to generate IRCs failed, possibly because the paths energetically lie above the path given in Figure 15a.

This finding could also indicate that a reaction of **1** starting from the H-orientation is more favorable. To answer this question, we calculated a two-step mechanism for inhibitor **1**. In the first step, **1** was found to twist from the F- into the H-orientation (variation of $\angle H^3C^3C^2H^2$). Starting from the obtained local minimum, we then calculated the course of inhibition by compound **1** starting from the H-orientation as the second step of the overall inhibition reaction. The result of this two-step inhibition reaction is shown in Figure 17.

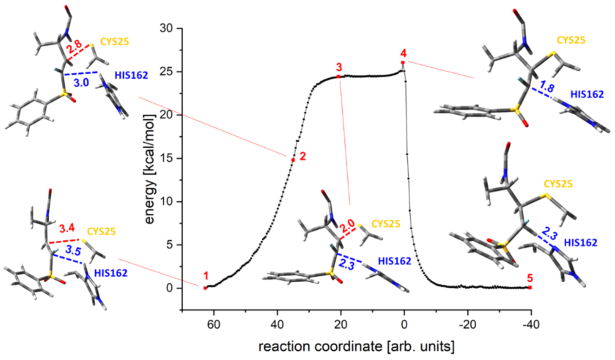
Figure 17 indicates that the shape of the reaction path of **1**, if it starts from the H-orientation, resembles the path found for **K11777** (Figure 16). Starting from the H-orientation (Figure 17, structure 3), the reaction proceeds over a barrier of about 18 kcal mol^{-1} and has a reaction energy of about −12 kcal mol^{-1}. However, the exothermicity of the whole reaction drops to only −4 kcal mol^{-1}, due to the previously necessary transition from the F- to the H-orientation. Other reaction courses given in the Supporting Information provide a similar picture (Figures S8−S10).

Our investigation revealed that the differences between **K11777** and **1** are mainly due to the interaction between the fluorine atom and the oxyanion hole of rhodesain, which induces a flip of the olefin group of the warhead within the active site. This change in the orientation significantly complicates the further course of the inhibition reaction. As a result, for inhibitor **1**, a two-step mechanism becomes more favorable, which contains a torsional movement from the F- to the H-orientation before the actual covalent Michael-type reaction can take place leading to a considerably reduced exothermicity. These differences nicely explain the switch from the irreversible (**K11777**) to the reversible (**1**) inhibition mode and are in line with all experimental data, which indicate a covalent, but reversible bond formation for various fluorinated inhibitors.

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358



(a)



(b)

Figure 16. Representative intrinsic reaction coordinate (IRC) simulation of the inhibition reaction of **K11777** (a) and **1** (b). Selected geometrical parameters of the indicated structures are summarized in Table S3.

Antitrypanosomal Activity and Toxicity. Selected compounds were tested for antitrypanosomal activity and cytotoxicity (Table 5). Trypanocidal activity was measured against *T. brucei brucei* TC211[54,55] or *T. brucei brucei* BS449, as described previously.[56,57] Cytotoxicity was measured for selected compounds in the macrophage cell line J774.1 and in HeLa cells as described before.[12,55] Most inhibitors tested showed significantly improved antitrypanosomal activity compared to starting compound **1**. This correlates very well with the higher inhibitory potency of the compounds for rhodesain. The compounds also displayed an improved safety profile by showing no relevant cytotoxicity in HeLa cells or in the J774.1 macrophage cell line. Most interestingly, from the hPhe series, the compounds with N-terminal DHBD group and 4-Me-Phe residue at the P2 position (**2e**, **4e**) exhibited highest antitrypanosomal potency with the sulfonate **2e** being the most potent compound (EC_{50} = 0.14 μM), similarly active to the irreversible **K11777**. Also, within the 4-Pyr series, the compounds with 4-Me-Phe residue (**2d**, **3d**, and **4d**) are more potent than those with Phe at P2 position (**2a**, **4a**). Probably, the higher lipophilicity and thus better membrane permeability of these compounds contribute to their better antitrypanosomal activity. No differences are observed between the various warheads (phenyl vinylsulfonate **2d**, benzyl vinylsulfone **3d**, and phenyl vinylsulfone **4d**). Within the *N*−Me-Pip series (**1**, **4j**, and **4l**), the compound with Cys(Bn) in the P1 position (**4l**) is most active (3.0 μM) despite being a much less-potent inhibitor of rhodesain. This observation may be explained with the relatively high toxicity of the compound, which seems to be also

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

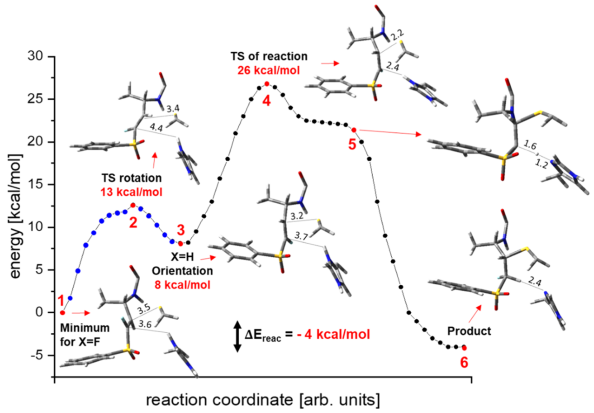
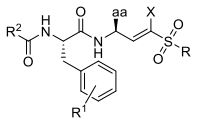


Figure 17. Reaction profile of the two-step inhibition reaction of **1**. The rotation about ∠H^{3}C^{3}C^{2}H^2 is given in blue. Please note that the reaction path is put together from several IRCs. Selected geometrical parameters of the indicated structures are summarized in Table S3. See also Figure S6.

Table 5. Antitrypanosomal Activity and Cytotoxicity of Selected Compounds[a]



	substitution					*T. b. b.* EC_{50}/μM	cytotoxicity CC_{50}/μM	
cpd	R	R^1	R^2	aa	X	48 h	J774.1	HeLa
K11777	Ph	H	*N*-MePip	hPhe	H	0.18 ± 0.03[b,c]	41[b]	>10[d]
1	Ph	H	*N*-MePip	hPhe	F	12.5 ± 0.4[c]	38	10 ± 2
2a	OPh	H	4-Pyr	hPhe	F	4.8 ± 0.9[c]	n.d.	>100
2d	OPh	4-Me	4-Pyr	hPhe	F	1.9 ± 1.8[e]	n.d.	>100
2e	OPh	4-Me	DHBD	hPhe	F	0.14 ± 0.05[e]	n.d.	>100
3d	Bn	4-Me	4-Pyr	hPhe	F	1.4 ± 0.9[e]	n.d.	>100
4a	Ph	H	4-Pyr	hPhe	F	3.0 ± 0.4[c]	>100	>100
4d	Ph	4-Me	4-Pyr	hPhe	F	1.9 ± 1.2[e]	n.d.	>100
4e	Ph	4-Me	DHBD	hPhe	F	0.80 ± 0.5[e]	n.d.	>100
4j	Ph	H	*N*-MePip	Met	F	14.1 ± 0.6[c]	43	77 ± 8
4l	Ph	H	*N*-MePip	Cys(Bn)	F	3.0 ± 0.1[c]	8.7	8 ± 0.4

[a]n.d. not determined. [b]See ref 12. [c]*T. brucei brucei* TC211. [d]See ref 58. [e]*T. brucei brucei* BS449.

connected to the *N*-Me-Pip moiety (compounds **K11777**, **1**, **4j**, and **4l**).

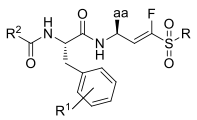
Selectivity over Serine and Threonine Proteases. For selected compounds, inhibitory potency against other proteases (the threonine protease 20S proteasome and the serine protease NS2B/NS3 of the Dengue virus) was tested (Table 6). The activity of the compounds against the different catalytically active subunits of the proteasome (trypsin-like, caspase-like, and α-chymotrypsin-like) was tested separately by the use of specific fluorogenic substrates (see Experimental Section). Most compounds showed no relevant inhibition (less than 50%) at concentrations of 11 μM. Highest percentage inhibition was observed in the case of the caspase-like activity for compounds **2d** (44%) and **4a** (41%). Inhibitor **4l** was the only compound that showed relevant inhibition of the α-chymotrypsin-like activity of the proteasome (71% at 11 μM).

In Vitro Metabolism. The metabolism of compounds **1** and **2d-(H)** was investigated using rat liver microsomes and an NADPH-generating system. The covalent cysteine protease inhibitor **K11777** was used as a control. Previous in vitro metabolism studies by Jacobsen and co-workers revealed three metabolites of **K11777** depicted in Figure 18.[59]

Compound **1** (fluorinated **K11777**) also showed N-demethylation (**1a**) and N-oxidation (**1b**) as shown in Figure S13. The metabolites were analyzed via LC−MS fragmentation. Additionally, the potential metabolites were synthesized as described in the Chemistry section, and their fragments and retention times were compared to those found in the metabolism studies to ensure the identity of the metabolites. In contrast to the reported metabolism of **K11777**, the β-hydroxy homophenylalanine derivative[32] did not occur, neither in the experiments with compound **1** nor in those with **K11777**.

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Table 6. Inhibition Data for 20S Proteasome and Dengue NS2B/NS3 Protease for Selected Compounds



	substitution				human 20S proteasome/%[a]			dengue
cpd	R	R^1	R^2	aa	trypsin	caspase	α-chymotrps.	NS2B/NS3%[a]
1	Ph	H	*N*-MePip	hPhe	n.i.	16	13	n.i.
2a	OPh	H	4-Pyr	hPhe	n.i.	22	11	n.i.
2d	OPh	4-Me	4-Pyr	hPhe	10	44	20	n.i.
2e	OPh	4-Me	DHBD	hPhe	n.i.	31	22	11
3d	Bn	4-Me	4-Pyr	hPhe	n.i.	n.i.	n.i.	n.i.
4a	Ph	H	4-Pyr	hPhe	n.i.	41	10	n.i.
4b	Ph	H	DHBD	hPhe	n.i.	26	n.i.	n.d.
4d	Ph	4-Me	4-Pyr	hPhe	12	28	13	n.d.
4j	Ph	H	*N*-MePip	Met	n.i.	18	n.i.	n.d.
4l	Ph	H	*N*-MePip	Cys(Bn)	14	n.i.	71	13

[a]% inhibition at inhibitor concentrations of 11 μM; n.i. = <10% inhibition at 11 μM; and n.d. = not determined.

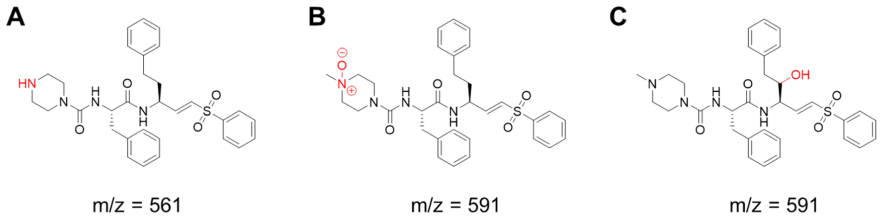
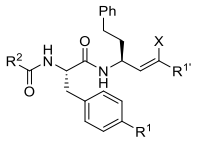


Figure 18. Metabolites of **K11777** as published by Jacobsen and co-workers.[59] N-Demethylation at the *N*-methyl piperazine moiety (A) and N-oxidation (B) are the major metabolic reactions. Additionally, a β-hydroxy homophenylalanine metabolite was found (C).

Table 7. Assay Results of the Metabolites



	substitution				rhodesain		
cpd	R^1	R^2	$R^{1'}$	X	K_i/nm	k_i/s^{-1}	$k_{2nd}/m^{-1}\ s^{-1}$
1	H	*N*-MePip	SO_2Ph	F	190[a,b]	n.a.	n.a.
1a	H	Pip	SO_2Ph	F	192 ± 40[b]	n.a.	n.a.
2d-(H)	Me	4-Pyr	SO_2Ph	H	0.45 ± 0.06	0.020	46×10^6
2l	Me	4-Pyr-*N*-oxide	SO_2Ph	H	1.50 ± 0.46	0.018	12×10^6

[a]See ref 12. [b]K_i calculated from the Cheng−Prusoff equation;[32] n.a. not applicable.

The metabolism studies of **2d-(H)** only resulted in a single metabolite, the N-oxidized derivative **2l** as shown in Figure S14. In order to verify the structure, the metabolite was synthesized as described in the Chemistry section and retention times and fragments were compared as described above.

The synthesized metabolites were also tested in the fluorometric enzyme assay in order to determine their K_i values. The assays were performed as described in the Enzyme assays section and the results are shown in Table 7.

The demethylation of compound **1** to its metabolite **1a** does not decrease its inhibitory activity significantly. Both, compound **2d-(H)** and its metabolite **2l**, show a good inhibition of rhodesain in the low nanomolar range. Therefore, it can be concluded that the metabolites retain most of the inhibitory potency of the parent drug.

In Vivo Distribution. Compound **1** as an example for a covalent reversible and compound **2d-(H)** as an example for a covalent irreversible inhibitor were tested for their biodistribution in vivo in wild type CD1 mice. Their biodistribution after oral (p.o.) or intraperitoneal (i.p.) application was investigated by LC−MS analysis of plasma samples and brain tissues to determine the ability of the compounds to cross the blood−

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

brain barrier. The compounds were chosen for the in vivo assays based on their differences in the mode of inhibition (covalent reversible vs irreversible) and based on their physicochemical properties (Table 8), which are similar for compound **2d-(H)** and its fluorinated counterpart **2d**.

Table 8. Comparison of 1, 2d, and 2d-(H) in Terms of Physicochemical Properties and the Mechanism of Inhibition

cpd	inhibition mode	SlogP (calc.)[a]	TPSA [Å²] (calc.)[a]
1	covalent reversible	3.96	98.82
2d	covalent reversible	5.07	114.46
2d-(H)	covalent irreversible	4.77	114.46

[a]Calculated using MOE.[60]

Compound **1** was found in plasma samples but not in the brain homogenate, whereas compound **2d-(H)** could be found in both, plasma and brain tissue after i.p. and p.o. administration, respectively (Figures 19, S15, and S16). Therefore, it can be concluded that **2d-(H)** is able to cross the blood−brain barrier due to its higher lipophilicity. An accumulation of this compound was also suggested because of a higher AUC after multiple oral administration compared to a single dose i.p. administration, thus making it a possible candidate for the treatment of stage-2 HAT.

Furthermore, the distribution of the compounds in the brain extracellular space was also investigated via microdialysis. Only compound **2d-(H)** could be found in the dialysate with recovery rates from the plasma concentration between 0.8 and 6% of plasma values after i.p. injection. This indicates that the lipophilic compound **2d-(H)** reaches higher concentrations intracellularly than in the extracellular fluid.

Notably, the mice treated with the compounds did not show any signs of toxification.

It can be concluded that the optimization of compound **1** through SAR studies not only enhanced the inhibitory potency against rhodesain and the selectivity over CatB and CatL but also resulted in a potential drug candidate for stage-2 HAT with a higher lipophilicity (calc. *S* log *P* = 4.77 compared to 3.96 for **1**) that is able to cross the blood brain barrier and accumulate in brain tissue.

CONCLUSIONS

In this study, compound **1** (K_i = 190 nM, EC_{50} (*T. brucei*) = 12.5 μM), which was recently identified as a covalent reversible cysteine protease inhibitor by the application of a quantum-chemical-based design protocol,[12] was optimized in terms of inhibitory potency and selectivity for rhodesain. Based on the results obtained from molecular docking and MD simulations, a series of compounds with a modified recognition unit and an altered substitution pattern of the warhead were synthesized. Introduction of aromatic residues at P3 significantly enhanced the potency for rhodesain (**4b**: K_i = 12 nM) and increased the trypanocidal activity against *T. brucei* [**4a**: EC_{50} (*T. brucei*) = 3.0 μM]. Incorporation of a 4-methyl substituent at the phenylalanine aromatic ring additionally improved selectivity against human cathepsins (**4d**: K_i = 8 nM, 12-fold selectivity over CatL, more than 200-fold selectivity over CatB). By alteration of the substitution pattern of the warhead, two new classes of covalent reversible cysteine protease inhibitors with distinct properties were obtained. The compounds from the series of benzyl fluorovinylsulfones (cpds **3a−3i**) showed a similar potency compared to the respective analogues with phenyl substituents (**4a−4i**) but were significantly more rapidly reversible in dilution and dialysis assays. Several compounds in the series of fluorovinylsulfonates (**2a−2i**) showed a biphasic inhibition mechanism, with the formation of a single digit nanomolar, high-affinity complex in the second step (**2d**: K_i^* = 3 nM). This complex was shown to dissociate markedly slower in dialysis experiments for compound **2d**. Therefore, compound **2d** represents a potent, tight binding, and slowly reversible inhibitor of rhodesain. Furthermore, compound **2d** shows selectivity over CatL (26-fold) and is only a weak inhibitor of CatB (38% at 11 μM). In addition, no relevant off-target activity against threonine and serine proteases was observed and the cytotoxicity profile improved considerably. Compounds with a DHBD moiety at the N-terminus and 4-Me-Phe in the P2 position (**2e**, **4e**) are not only nanomolar inhibitors of rhodesain, but are most promising with regard to their the antitrypanosomal activity and cytotoxicity profile, with EC_{50} values comparable to those of the irreversible inhibitor **K11777** (EC_{50} = 0.14−0.80 μM) and no cytotoxic effects against HeLa cells (EC_{50} > 100 μM).

Microsomal stability assays revealed N-oxidation of the 4-Pyr substituent in the P3 position of **2d-(H)**. However, this does not seem to reduce the inhibitory potency. Moreover, the

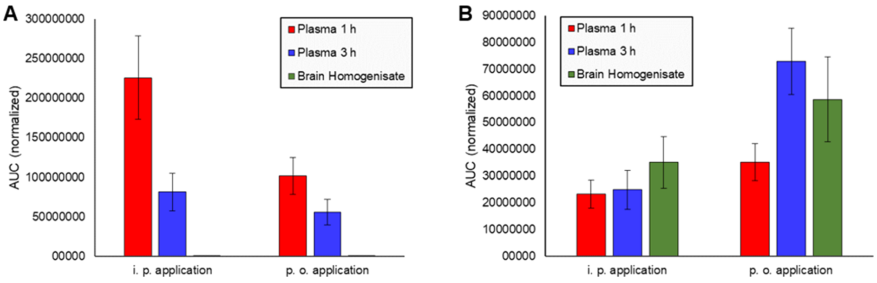


Figure 19. (A) Compound **1** was found in mouse plasma after 1 and 3 h, but not in the brain homogenate after i.p. and p.o. administration, respectively. (B) Compound **2d-(H)** showed plasma levels and distribution in brain tissue after both i.p. and p.o. administration. Both diagrams show the average AUC values of all tested mice (five for each compound and route of administration).

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

accumulation of **2d-(H)** in mouse brain tissue, which did not occur with **1**, indicates a correlation of CNS permeation with the physicochemical properties of the tested inhibitors, such as SlogP values. These findings confirm that the optimized covalent reversible fluorovinylsulfonates should be further investigated as possible candidates for the treatment of both stage-1 and -2 HAT.

Based on these features, the high-affinity, lipophilic, and selective reversible α-fluorovinylsulfones and sulfonates as inhibitors of rhodesain may serve as a basis for the future development of effective and non-toxic drugs for the treatment of HAT.

In addition, we used mass spectrometry to demonstrate the formation of a covalent bond. Using QM/MM and MD computations, we could attribute the differences in inhibition mechanisms between **K11777** (irreversible) and **1** (reversible) to the interaction between the fluorine atom and the oxyanion hole of rhodesain. Due to this interaction, the olefin group of the warhead flips in the active site significantly complicating the further course of the inhibition reaction and leading to a considerably reduced exothermicity and thus a reversible inhibition.

EXPERIMENTAL SECTION

Syntheses. *General.* All reagents and solvents were of analytical grade quality and purchased from Sigma-Aldrich, Alfa Aesar, Acros, or TCI. Chemicals were used without further purification. Solvents were purified by distillation and desiccated by standard methods if necessary. ^{1}H and ^{13}C spectra were recorded on a Bruker Fourier 300 using DMSO-d_6, $CDCl_3$, or CD_2Cl_2 as a solvent. Chemical shifts δ are given in parts per million (ppm) using residual proton peaks of the solvent as the internal standard (^{1}H/^{13}C: DMSO: 2.50/39.52 ppm, $CHCl_3$: 7.26/77.16 ppm, CH_2Cl_2: 5.32/54.00 ppm). The purity of the compounds was determined via HPLC-MS (δ = 254 nm). All compounds are >95% pure according to HPLC analysis. High-resolution mass spectra were obtained on a Waters Q-TOF-Ultima 3-instrument. Alternatively, the mass spectra were obtained from an LC−MS system consisting of a 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 80% acetonitrile, 10% H_2O, and 10% of a 0.1% solution of formic acid in water. Detection wavelength was 254 nm. The molecular mass was detected using an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI) in positive mode. Purification with a preparative HPLC system was performed with a Varian PrepStar system (model 218) with an Agilent Zorbax XDB-C18 21.2 × 150 mm, 5 μm column. Column chromatography was performed with silica gel (0.06−0.02 mm or 0.040−0.063 mm) obtained from Carl Roth. All reactions were monitored by thin-layer chromatography using Macherey-Nagel ALUGRAM Xtra SIL G/UV254 silica gel 60 plates for detection at 254 nm. Melting points were determined in open capillaries using a Stuart SMP10-instrument. Optical rotation $[\alpha]_D^{22}$ was measured on an P3000 polarimeter from Krüss at 22 °C and is reported in cm^3 g^{-1} dm^{-1}.

Synthesis of Starting Materials. Phenyl Methanesulfonate[61] (I). Phenol (2.35 g, 1.0 equiv) was dissolved in EtOAc and cooled to 0 °C. TEA (6.93 mL, 2.0 equiv) and methanesulfonyl chloride (2.52 mL, 1.3 equiv) were added successively and the mixture was allowed to warm to rt. After stirring for 30 min, the solution was washed with water (3×) and brine (1×) and was dried with Na_2SO_4. Evaporation of the solvent yielded phenyl methanesulfonate (**I**) as a colorless powder (3.95 g, 92%). Spectral data matched those reported in the literature.[28]

Diethyl((benzylthio)methyl)phosphonate[28] (II). Benzyl mercaptan (0.84 mL, 1.0 equiv) was dissolved in THF and cooled to 0 °C. NaH (60% in mineral oil, 0.32 g, 1.0 equiv) was added in portions and the mixture was stirred for 15 min. Diethyl iodomethylphosphonate (1.24 mL, 1.0 equiv) was added dropwise and the mixture was stirred for 3.5 h at rt. THF was removed in vacuo and EtOAc was added. The organic phase was washed with water (2×) and brine (1×) and dried with Na_2SO_4. After evaporation of the solvent, the residue was purified by column chromatography (petroleum ether/EtOAc 1:1), yielding the title compound as a colorless liquid. Yield: 1.66 g (84%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.43−7.21 (m, 5H), 4.25−4.08 (m, 4H), 3.90 (s, 2H), 2.54 (d, J_{H-P} = 12.9 Hz, 2H), 1.34 (t, J = 7.1 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 137.3, 129.4, 128.7, 127.4, 62.7 (d, J_{C-P} = 6.7 Hz), 37.0 (J_{C-P} = 2.9 Hz), 23.8 (J_{C-P} = 150 Hz), 16.6 (J_{C-P} = 6.0 Hz). ^{1}H NMR data are consistent with the literature.[12]

Diethyl((phenylsulfonyl)methyl)phosphonate[12] (III). To a solution of methyl phenyl sulfone (3.81 g, 24 mmol) in dry THF, *n*-BuLi (2.5 M in hexane, 21.5 mL, 60 mmol) was added at 0 °C with stirring. After 30 min, diethyl chlorophosphate was added dropwise and the reaction mixture was stirred at 0 °C for 1 h. 25 mL of a saturated solution of ammonium chloride was added and the volatiles were evaporated. The residue was extracted with DCM and the combined extracts were washed with brine, dried with sodium sulfate, concentrated under reduced pressure, and purified by column chromatography (light petroleum/EtOAc 1:5) to afford the title compound as a colorless oil, which solidified upon standing (4.11 g, 52%). NMR data were consistent with the literature.[28,62]

Phenyl(diethoxyphosphoryl)methanesulfonate (IV). Phenyl methanesulfonate (I, 2.5 g, 1.0 equiv) was dissolved in THF and cooled to −78 °C. KHMDS (1.0 M in THF, 15 mL. 1.1 equiv) was added dropwise and the mixture was stirred for an additional 15 min. Diethyl chlorophosphate (1.5 mL, 0.7 equiv) was added slowly and stirred for 1 h at −60 °C. The reaction was quenched by addition of glacial AcOH (1 mL) and then allowed to warm to rt. THF was removed in vacuo and EtOAc was added to the residue. The solution was washed with water (2×) and brine (1×) and then dried with Na_2SO_4. Evaporation of the solvent gave a crude product, which was purified by column chromatography (petroleum ether/EtOAc 1:1 to 1:3), yielding the title compound as a colorless liquid (3.3 g, 74%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.52−7.29 (m, 5H), 4.36−4.20 (m, 4H), 3.81 (d, J_{H-P} = 17 Hz, 2H), 1.44−1.31 (m, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 149.4, 130.2, 127.8, 122.3, 64.2 (d, J_{C-P} = 6.5 Hz), 47.3 (d, J_{C-P} = 139 Hz), 16.4 (d, J_{C-P} = 6.3 Hz).

Diethyl((benzylsulfonyl)methyl)phosphonate[28] (V). Compound **II** (2.55 g, 1.0 equiv) was dissolved in DCM and cooled to 0 °C. Subsequently, *m*CPBA (77%, 5.8 g, 3.0 equiv) was added in portions and the mixture was stirred for 12 h at rt. Then, the solution was filtered and washed with 1 M NaOH (4×), water (1×), and brine (1×) and dried with Na_2SO_4. After evaporation of the solvent, the title compound was obtained as a colorless oil. Yield: 2.82 g (99%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.57−7.47 (m, 2H), 7.44−7.34 (m, 3H), 4.60 (s, 2H), 4.24 (dq, J = 8.1, 7.1 Hz, 4H), 3.36 (d, J = 16 Hz, 2H), 1.37 (dt, J = 7.1, 0.4 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 131.2, 129.3, 129.2, 128.2, 63.9 (d, J_{C-P} = 6.5 Hz), 60.4, 48.0 (d, J_{C-P} = 140 Hz), 16.50 (d, J_{C-P} = 6.4 Hz). Spectral data are consistent with the literature.[12,63]

(S)-tert-Butyl(1-oxo-4-phenylbutan-2-yl)carbamate (VI). L-Homophenylalanine (5.0 g, 28 mmol) was dissolved in THF (15 mL) and a solution of Na_2CO_3 (3 g, 28 mmol) in 100 mL of water was added followed by di-*tert*-butyl dicarbonate (31 mmol) in 75 mL THF. The mixture was stirred overnight, diluted with water (100 mL), and extracted with CH_2Cl_2. The aqueous layer was acidified with $KHSO_4$ to pH 3 and extracted with CH_2Cl_2. The combined organic extracts were dried with sodium sulfate and concentrated under reduced pressure to give crude boc-L-homophenylalanine (7.0 g, 89%). This crude material (4.91 g, 18 mmol) was dissolved in DCM (100 mL) and cooled to 0 °C. EDC−HCl (4.05 g, 21 mmol), HOBt (3.23 g, 21 mmol), and DIEA (13.5 mL, 78 mmol) were added and the mixture was stirred for 15 min before *N*,*O*-dimethylhydroxylamine hydrochloride (2.06 g, 21 mmol) was added. The mixture was allowed to warm to rt and was stirred for 18 h. DCM was removed under reduced pressure, water (60 mL) was added, and the suspension was extracted with EtOAc. The combined organic extracts were washed with saturated aq $NaHCO_3$ (5×) and brine, dried with sodium sulfate, and concentrated under reduced pressure to give *tert*-butyl (*S*)-(1-(methoxy(methyl)amino)-1-oxo-4-phenylbutan-2-yl)carbamate as a crude yellow oil (5.64 g, 99%), which was used in the next step without further purification. ^{1}H NMR (300 MHz, DMSO-d_6): δ 7.46−7.18 (m, 5H), 5.28 (d, J = 9.3 Hz, 1H), 4.73

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

(s, 1H), 3.67 (s, 3H), 3.21 (s, 3H), 2.88−2.61 (m, 2H), 2.18−1.79 (m, 2H), 1.52 (s, 9H). The crude oil (4.05 g, 12.5 mmol) was dissolved in dry diethyl ether (130 mL) and cooled to 0 °C. $LiAlH_4$ (0.59 g, 15.7 mmol) was added portion wise. The mixture was stirred for 30 min and subsequently quenched by addition of aqueous $KHSO_4$ (0.33 M, 65 mL). The organic phase was separated, and the aqueous phase was extracted with diethyl ether. The combined organic extracts were washed with 3 M HCl, saturated aq $NaHCO_3$, and brine, dried with sodium sulfate, concentrated under reduced pressure, and purified by column chromatography (light petroleum/EtOAc 3:1), yielding the title compound as a colorless solid (2.44 g, 74%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 9.55 (s, 1H), 7.40−7.13 (m, 5H), 5.08 (d, J = 4.1 Hz, 1H), 4.26 (q, J = 5.9 Hz, 1H), 2.87−2.63 (m, 2H), 2.35−2.06 (m, 1H), 2.02−1.77 (m, 2H), 1.47 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 199.6, 155.6, 140.7, 128.8, 128.7, 126.5, 80.4, 59.7, 31.6, 31.2, 28.4. NMR data are consistent with the literature.[64]

(S)-tert-Butyl(4-(methylthio)-1-oxobutan-2-yl)carbamate (VII). Boc-L-Met-OH (2.43 g, 1.0 equiv) was dissolved in DCM and cooled to 0 °C. HOBt (1.49 g, 1.0 equiv) and DIEA (5.0 mL, 3.0 equiv) were added successively, and the mixture was stirred until all materials were solubilized. To this mixture, TBTU (3.12 g, 1.0 equiv) was added in one portion and then stirred for 15 min at 0 °C. *N,O*-Dimethylhydroxylamine hydrochloride (1.0 g, 1.1 equiv) was added in one portion and the mixture was stirred for 12 h at rt. DCM was removed in vacuo, and the residue was diluted with EtOAc. The organic layer was washed with water (5×), conc. $NaHCO_3$ (2×), and 1 M aq HCl (2×) and dried with Na_2SO_4. After evaporation of the solvent, the residue was purified by column chromatography (petroleum ether/EtOAc 2:1), yielding *tert*-butyl (*S*)-(1-methoxy(methyl)amino)-4-(methylthio)-1-oxobutan-2-yl)carbamate as a colorless oil (2.42 g, 85%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 5.23 (d, J = 7.7 Hz, 1H), 4.78 (s, 1H), 3.77 (s, 3H), 3.20 (s, 3H), 2.68−2.42 (m, 2H), 2.08 (s, 3H), 2.05−1.92 (m, 1H), 1.89−1.71 (m, 1H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 172.7, 155.7, 79.9, 61.8, 50.0, 32.6, 32.3, 30.3, 28.5, 15.6. This material (2.42 g, 1.0 equiv) was dissolved in THF and cooled to 0 °C. $LiAlH_4$ (0.41 g, 1.3 equiv) was added in portions and stirred for 30 min. The mixture was diluted with 50 mL diethyl ether and then 1 M $KHSO_4$ (50 mL) was added carefully. The layers were separated, and the aqueous phase was extracted twice with diethyl ether. The combined organic layers were washed with 1 M HCl (2×), saturated aq $NaHCO_3$, and brine, and then dried with Na_2SO_4. After evaporation of the solvent, the residue was purified by column chromatography (petroleum ether/EtOAc 2:1) to give the title compound as a colorless solid (1.49 g, 77%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 9.63 (s, 1H), 5.21 (s, 1H), 4.39−4.11 (m, 1H), 2.56 (t, J = 7.2 Hz, 2H), 2.34−2.13 (m, 1H), 2.07 (s, 3H), 1.99−1.83 (m, 1H), 1.44 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 199.2, 155.6, 80.4, 59.1, 29.9, 28.8, 28.4, 15.5. NMR data are consistent with the literature.[65]

(R)-tert-Butyl(1-(benzylthio)-3-oxopropan-2-yl)carbamate (VIII). L-Cysteine hydrochloride (5.20 g, 1.0 equiv) was dissolved in 60 mL 2 M NaOH and 150 mL EtOH. To the mixture, 5.64 g (3.92 mL, 1.0 equiv) benzyl bromide is added dropwise. After stirring for 1 h, the mixture is neutralized by addition of conc. HCl. The precipitate is collected by filtration and washed with water, diethyl ether, and ethanol. *S*-Benzyl L-cysteine is obtained as a colorless solid and directly used for the next step (6.4 g. 91%). ^{1}H NMR (300 MHz, D_2O): δ 7.50−7.24 (m, 5H), 4.17 (q, J = 6.5 Hz, 1H), 3.88−3.71 (s, 2H), 3.09−2.87 (m, 2H). ^{13}C NMR (75 MHz, D_2O): δ 171.7, 138.2, 129.3, 129.1, 127.8, 51.8, 35.5, 30.7. *S*-Benzyl L-cysteine (3.0 g, 1.0 equiv) was dissolved in THF (50 mL), and K_2CO_3 (2.94 g, 1.5 equiv) and water (50 mL) were added. Di-*tert*-butyl-dicarbonate (3.25 g, 1.05 equiv) dissolved in 50 mL THF was added dropwise and stirred for 12 h at rt. THF was removed in vacuo and the residue was extracted twice with DCM. The aqueous layer was acidified to pH 3 by addition of 1 M $KHSO_4$ and extracted with DCM (3×). The combined organic layers were washed with brine (2×), dried with Na_2SO_4, and evaporated under reduced pressure to give crude boc-*S*-benzyl L-cysteine as a colorless oil (3.05 g, 69%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 7.38−7.03 (m, 5H), 4.14−4.02 (m, 1H), 3.71 (s, 2H), 2.67 (m, 2H), 1.35 (s, 9H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 172.6, 155.4, 138.3, 128.9, 128.4, 126.9, 78.3, 53.2, 39.5, 35.2, 32.4, 28.2. This material (3.05 g, 1.0 equiv) was dissolved in DCM and cooled to 0 °C. HOBt (1.49 g, 1.0 equiv) and DIEA (5.0 mL, 3.0 equiv) were added successively, and the mixture was stirred until all materials were solubilized. To this mixture, TBTU (3.40 g, 1.0 equiv) was added in one portion and then stirred for 15 min at 0 °C. *N,O*-Dimethylhydroxylamine hydrochloride (1.0 g, 1.1 equiv) was added in one portion and the mixture was stirred for 12 h at rt. DCM was removed in vacuo, and the residue was diluted with EtOAc. The organic layer was washed with water (5×), conc. $NaHCO_3$ (2×), and 1 M aq HCl (2×) and dried with Na_2SO_4. After evaporation of the solvent, the residue was purified by column chromatography (petroleum ether/EtOAc 2:1), yielding (*R*)-*tert*-butyl (3-(benzylthio)-1-(methoxy-(methyl)amino)-1-oxopropan-2-yl)carbamate as a colorless oil (3.25 g, 95%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.42−7.15 (m, 5H), 5.28 (d, J = 8.6 Hz, 1H), 4.84−4.59 (m, 1H), 3.67 (s, 3H), 3.21 (s, 3H), 2.83−2.54 (m, 2H), 2.15−1.77 (m, 2H), 1.50 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 173.2, 155.7, 141.3, 128.7, 128.5, 126.1, 61.6, 50.2, 38.8, 34.7, 31.8, 28.5. The above compound (3.0 g, 1.0 equiv) was dissolved in diethyl ether and cooled to 0 °C. $LiAlH_4$ (0.42 g, 1.3 equiv) was added in portions and stirred for 30 min. The mixture was diluted with 50 mL diethyl ether and then 1 M $KHSO_4$ (50 mL) was added carefully. The layers were separated, and the aqueous phase was extracted twice with diethyl ether. The combined organic layers were washed with 1 M HCl (2×), saturated aq $NaHCO_3$, and brine, and then dried with Na_2SO_4. After evaporation of the solvent, the residue was purified by column chromatography (petroleum ether/EtOAc 2:1) to give the title compound as a colorless solid (1.70 g, 68%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 9.46 (s, 2H), 7.32−7.21 (m, 13H), 5.29 (d, J = 21.5 Hz, 2H), 4.32−4.16 (m, 2H), 3.67 (s, 5H), 2.86−2.70 (m, 5H), 1.39 (s, 24H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 198.8, 137.7, 129.1, 128.8, 127.3, 80.6, 77.2, 59.3, 37.1, 30.8, 28.4. NMR data are consistent with the literature.[12]

(S)-2-(4-Methylpiperazin-1-ium-1-carboxamido)-3-phenylpropanoate[12] *(IX).* L-Phenylalanine methyl ester hydrochloride (2.0 g, 1.0 equiv) was suspended in CH_2Cl_2 and 25 mL of a saturated aqueous solution of $NaHCO_3$ was added. At 0 °C, triphosgene (0.92 g, 0.3 equiv) was added and the mixture was allowed to stir for 30 min at this temperature. Subsequently, the organic phase was separated, and the aqueous phase was extracted with three portions of DCM. The combined organic extracts were washed with brine, dried with Na_2SO_4, and concentrated under reduced pressure. The residue was dissolved in THF and cooled to 0 °C. *N*-Methyl piperazine (0.93 g, 1.03 mL, 1.0 equiv) was added dropwise. After stirring for 1 h, THF was removed under reduced pressure and the residue was diluted with EtOAc. Aqueous work-up was performed with water (1×), saturated aq $NaHCO_3$ (1×), and brine (1×). The organic layer was dried with Na_2SO_4 and evaporated. The crude product was purified by column chromatography (CH_2Cl_2/MeOH 9:1), giving methyl (4-methylpiperazine-1-carbonyl)-L-phenylalaninate as a colorless oil. Yield: ^{1}H NMR (300 MHz, $CDCl_2$): δ 7.30−7.10 (m, 3H), 7.04 (d, J = 7.3 Hz, 2H), 4.77 (d, J = 7.3 Hz, 1H), 4.61 (m, 1H), 3.61 (s, 3H), 3.29−3.11 (m, 4H), 3.00 (m, 2H), 2.22 (t, J = 5.0 Hz, 4H), 2.16 (s, 3H). This material (3.0 g, 9.4 mmol) was dissolved in THF (20 mL) and LiOH (1.4 g, 33.4 mmol) in 10 mL water was added. The mixture was stirred for 18 h at rt and the volatiles were removed under reduced pressure. Water was added to the residue and the pH was adjusted to 2 with 1 M aq HCl. The title compound was crystallized at 4 °C along with some LiCl and collected by filtration. Recrystallization from methanol gave the pure compound as a colorless powder. ^{1}H NMR (300 MHz, DMSO-d_6): δ 7.33−7.05 (m, 5H), 6.68 (d, J = 7.9 Hz, 1H), 4.28−4.09 (m, 1H), 3.50−3.16 (m, 4H), 2.98 (m, 2H), 2.46−2.30 (m, 4H), 2.25 (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 174.6, 157.1, 139.0, 129.4, 128.0, 126.1, 56.0, 53.1, 44.8, 42.7, 39.5, 36.9.

Isonicotinoyl-L-phenylalanine[66] *(X).* To a solution of L-phenylalanine ethyl ester hydrochloride (2.50 g, 10.88 mmol) and TEA (4.5 mL, 32.65 mmol) in DCM, isonicotinoyl chloride hydrochloride (1.93 g, 10.88 mmol) was added in small portions. The mixture was stirred for 3 h, washed with a saturated solution of $NaHCO_3$, dried with Na_2SO_4, and concentrated under reduced pressure. Ethyl isonicotinoyl-L-phenylalaninate (2.84 g, yield 87%) was obtained as a yellowish oil,

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

which crystallized upon standing and was used in the next step without further purification. ^{1}H NMR (300 MHz, DMSO-d_6): δ 9.18 (d, J = 7.7 Hz, 1H), 8.73 (d, J = 5.2 Hz, 2H), 7.69 (d, J = 5.2 Hz, 2H), 7.37−7.12 (m, 5H), 4.75−4.57 (m, 1H), 4.10 (q, J = 7.1 Hz, 2H), 3.05−3.21 (m, 2H), 1.14 (t, J = 7.1 Hz, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.3, 165.0, 150.3, 140.6, 137.4, 129.1, 128.3, 126.6, 121.3, 60.7, 54.4, 45.7, 39.5, 36.2, 14.0. The above compound (2.84 g, 9.53 mmol) was dissolved in water (30 mL) and THF (30 mL), then LiOH (1.20 g, 28.6 mmol) was added. The mixture was stirred for 2 h before the organic solvent was removed under reduced pressure and the aqueous residue was neutralized with 1 M aq HCl. The title compound was crystallized at 4 °C along with LiCl. Recrystallization from methanol gave the pure product as a colorless powder (1.59 g, 62%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.76−8.57 (m, 2H), 8.22 (d, J = 7.3 Hz, 1H), 7.71−7.57 (m, 2H), 7.25−6.99 (m, 5H), 4.26 (td, J = 7.4, 4.6 Hz, 1H), 3.27−2.95 (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.9, 163.4, 150.2, 142.1, 139.5, 129.4, 127.7, 125.6, 121.0, 56.1, 37.2.

*11-(Tritylthio)undecanoic Acid (**XI**).* Triphenylmethyl chloride (0.63 g, 1.0 equiv) was dissolved in DCM (5 mL). Thereto, a solution of 11-mercaptoundecanoic acid (MUA) (0.5 g, 1.0 equiv) in DCM (15 mL) was added dropwise. The mixture was stirred for 2 h at rt. DCM was removed by distillation, and the residue was purified by column chromatography (petroleum ether/EtOAc 2:1). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.40 (d, J = 7.6 Hz, 6H), 7.31−7.09 (m, 9H), 2.32 (t, J = 7.1 Hz, 2H), 2.11 (t, J = 7.3 Hz, 2H), 1.60 (q, J = 7.4 Hz, 2H), 1.46−0.95 (m, 14H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 179.7, 145.2, 129.8, 127.9, 126.6, 66.5, 34.1, 32.2, 29.5, 29.32, 29.27, 29.16, 29.12, 28.7, 24.8. NMR data are consistent with the literature.[12]

*tert-Butyl (S)-4-((1-Ethoxy-1-oxo-3-phenylpropan-2-yl)-carbamoyl)piperazine-1-carboxylate (**XII**).* To a stirred solution of 0.798 g 1-boc piperazine in THF, a suspension of 0.860 g (*S*)-ethyl-2-isocyanate-3-phenylpropanoate in THF was added dropwise. The mixture was stirred for 18 h and the solvent was removed under reduced pressure. The residue was extracted with EtOAc and the combined extracts were washed with a saturated solution of $NaHCO_3$ and brine, dried with Na_2SO_4, and concentrated under reduced pressure, resulting in a colorless oil (0.81 g, 50%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.49−7.33 (m, 3H), 7.29−7.18 (m, 2H), 5.10−4.83 (m, 1H), 4.28 (dq, J = 15.5, 7.2 Hz, 2H), 3.70−3.57 (m, 1H), 3.57−3.48 (m, 4H), 3.48−3.35 (m, 4H), 3.29−3.22 (m, 2H), 1.60 (s, 9H), 1.39 (td, J = 7.1, 3.3 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 172.6, 156.5, 154.6, 136.3, 129.3, 128.5, 127.0, 80.2, 61.4, 60.4, 54.37, 43.9, 38.4, 28.4, 21.1, 14.2.

*(4-(tert-Butoxycarbonyl)piperazine-1-carbonyl)-L-phenylalanine (**XIII**).* 0.81 g of compound **XII** was dissolved in THF and cooled to 0 °C. A solution of 0.29 g LiOH in water was added dropwise and the mixture was stirred for 3 h at room temperature. The solvent was removed under reduced pressure and the pH of the residue was adjusted to 2 with 1 M HCl. The residue was extracted with ethyl acetate and the combined extracts were washed with brine, dried with Na_2SO_4, and concentrated under reduced pressure. The crude product was dissolved in diethyl ether, *n*-pentane was added, and the product was crystallized at 4 °C, giving a colorless solid (0.56 g, 74%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 9.39 (s, 1H), 7.25−7.14 (m, 3H), 7.14−7.04 (m, 2H), 5.05 (d, J = 7.1 Hz, NH), 4.61 (q, J = 6.2 Hz, 1H), 3.30 (dd, J = 7.9, 4.0 Hz, 4H), 3.24−3.09 (m, 4H), 3.04 (dd, J = 14.0, 6.8 Hz, 2H), 1.39 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 174.7, 157.5, 154.8, 136.4, 129.5, 128.7, 127.2, 80.6, 54.9, 43.7, 37.5, 28.5.

*Preparation of Phosphonates **5**−**7**. Diethyl(fluoro-(phenylsulfonyl)methyl)phosphonate[12,67] (**5**).* To a stirred solution of phosphonate **III** (1.54 g, 1.0 equiv) in THF, KHMDS (1 M in THF, 6.59 mL, 1.25 equiv) was added dropwise at −80 °C. After 30 min at this temperature, Selectfluor (2.80 g, 1.5 equiv) in 10 mL DMF was added and the reaction mixture was stirred for 3 h, slowly warming to 0 °C. 15 mL of a saturated solution of ammonium chloride was added and the volatiles were evaporated. The residue was extracted with DCM and the combined extracts were washed with a saturated solution of $NaHCO_3$ and brine, dried with Na_2SO_4, concentrated under reduced pressure, and purified by column chromatography (light petroleum/EtOAc 1:6) to afford the title compound as a colorless solid. Yield: 1.0 g (62%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.01 (d, J = 7.5 Hz, 2H), 7.74 (t, J = 7.4 Hz, 1H), 7.61 (t, J = 7.7 Hz, 2H), 5.38 (dd, J_{H-F} = 45.5 Hz, J_{H-F} = 6.6 Hz, 1H), 4.40−4.18 (m, 4H), 1.35 (t, J = 7.1 Hz, 6H). NMR data were consistent with the literature.[12]

*Phenyl(diethoxyphosphoryl)fluoromethanesulfonate (**6**).* Phosphonate **IV** (1.73 g, 1.0 equiv) was dissolved in THF (15 mL) and cooled to −78 °C. KHMDS (1 M in THF, 7.0 mL, 1.25 equiv) was added dropwise and the mixture was stirred for an additional 30 min. Selectfluor (3.0 g, 1.5 equiv) was added in one portion and the mixture was stirred for 5 min. Then, DMF (12 mL) was added in one portion and the mixture was allowed to warm to 0 °C. After stirring for 3 h, the reaction was quenched with conc. aq NH_4Cl (5 mL) and THF was removed in vacuo. EtOAc was added and the organic phase was washed with water (3×) and brine (2×) and dried with Na_2SO_4. The solvent was removed under reduced pressure and the residue was purified by column chromatography (petroleum ether/EtOAc 1:1), yielding the title compound as a colorless oil (0.95 g, 52%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.50−7.39 (m, 2H), 7.39−7.28 (m, 3H), 5.63 (dd, J_{H-F} = 45 Hz, J_{H-P} = 7.2 Hz, 1H), 4.45−4.25 (m, 4H), 1.39 (tdd, J = 7.1, 2.0, 0.7 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 149.6, 130.3, 128.0, 122.2, 93.4 (dd, J_{C-F} = 232 Hz, J_{C-P} = 162 Hz), 65.7 (dd, J_{C-P} = 10.7 Hz, J_{C-F} = 6.7 Hz), 16.44 (d, J_{C-P} = 5.8 Hz).

*Diethyl((benzylsulfonyl)fluoromethyl)phosphonate (**7**).* Phosphonate **V** (3.77 g, 1.0 equiv) was dissolved in THF (40 mL) and cooled to −78 °C. KHMDS (1 M in THF, 15.4 mL, 1.25 equiv) was added dropwise and the mixture was stirred for an additional 30 min. Selectfluor (6.55 g, 1.5 equiv) was added in one portion and the mixture was stirred for 5 min. Then, DMF (24 mL) was added in one portion and the mixture was allowed to warm to 0 °C. After stirring for 3 h, the reaction was quenched with conc. NH_4Cl (5 mL) and THF was removed in vacuo. EtOAc was added and the organic phase was washed with water (3×) and brine (2×) and dried with Na_2SO_4. The solvent was removed under reduced pressure and the residue was purified by column chromatography (DCM/EtOAc 9:1), yielding the title compound as a colorless oil (1.92 g, 48%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.51−7.32 (m, 5H), 5.28 (dd, J_{H-F} = 45 Hz, J_{H-P} = 7.1 Hz, 1H), 4.67−4.49 (m, 2H), 4.42−4.21 (m, 4H), 1.39 (td, J = 7.1, 0.7 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 131.4, 129.5, 129.2, 125.8, 95.2 (dd, J_{C-F} = 228 Hz, J_{C-P} = 161 Hz), 65.4 (dd, J_{C-F} = 6.3 Hz, J_{C-P} = 6.2 Hz), 57.5, 16.5 (d, J_{C-P} = 5.8 Hz).

*HWE Olefination (Compounds **8**−**12**). Procedure A.* The specified phosphonate (1.0 equiv) was dissolved in THF and cooled to 0 °C. NaH (60% in mineral oil, 1.1 equiv) was added in portions and the mixture was stirred for 15 min. The corresponding boc-protected aminoaldehyde (1.0 equiv) was added in one portion, and the mixture was allowed to warm to rt. After stirring for 1 h, EtOAc was added and the organic phase was extracted with water, saturated aq $NaHCO_3$, and brine and finally dried with Na_2SO_4. Evaporation of the solvent under reduced pressure provided the boc-protected vinylsulfones as a crude mixture of *E*/*Z* isomers. Purification and separation of the isomers were achieved by column chromatography. The products were obtained as colorless solids.

Procedure B. The specified phosphonate (1.0 equiv) was dissolved in THF and cooled to −78 °C. KHMDS (1 M in THF, 1.0 equiv) or LHMDS (1 M in THF, 1.0 equiv) was added dropwise and it was stirred for additional 20 min. The respective aldehyde (1.0 equiv) was added in one portion and the mixture was allowed to warm to rt over a period of 1 h. After stirring for an additional hour at rt, EtOAc was added and the organic phase was extracted with water (2×) and brine (1×). Evaporation of the solvent under reduced pressure gave the crude product as *E*/*Z* isomers. Purification by column chromatography allowed the isolation of the desired (*E*)-isomers. The products were obtained as colorless solids.

*tert-Butyl (S,E)-(1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)carbamate[12] (**8**).* Prepared following procedure A using 0.35 g of phosphonate **5**, 0.3 g boc-homophenylalaninal or boc-homophenylalanine aldehyde or boc-hPhe-H (**VI**), and 0.05 g NaH in 6 mL THF. Column chromatography: petroleum ether/EtOAc 3:1. (*E*)-isomer: Yield 0.20 g (41%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.95 (d, J = 7.6 Hz, 2H), 7.70 (t, J = 7.4 Hz, 1H), 7.58 (t, J = 7.6 Hz, 2H), 7.36−7.16 (m, 3H), 7.12 (d, J = 7.1 Hz, 2H), 6.19 (dd, J_{H-F} = 32.3 Hz, J_{H-H} = 7.1 Hz,

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

1H), 4.61 (s, 1H), 4.46 (s, 1H), 2.74−2.52 (m, 2H), 2.00−1.77 (m, 2H), 1.39 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 154.9, 154.7 (d, J_{C-F} = 300 Hz), 140.5, 137.2, 134.7, 129.61, 128.8, 128.7, 128.4, 126.4, 118.4 (d, J_{C-F} = 4.6 Hz), 80.3, 46.4 (d, J_{C-F} = 2.7 Hz), 36.2, 32.0, 28.3. (*Z*)-isomer: Yield 0.10 g (21%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.09 (d, J = 4.2 Hz, 2H), 7.69 (t, J = 7.4 Hz, 1H), 7.57 (t, J = 7.7 Hz, 2H), 7.37−7.12 (m, 5H), 5.81 (dd, J_{H-F} = 21 Hz, J_{H-H} = 10 Hz, 1H), 5.35−5.18 (m, 1H), 4.70 (s, 1H), 2.89−2.62 (m, 2H), 2.14−1.86 (m, 2H), 1.46 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 155.0, 152.7 (d, J_{C-F} = 293 Hz), 141.0, 137.7, 134.7, 129.5, 129.1, 128.7, 128.5, 126.3, 121.4 (d, J_{C-F} = 11 Hz), 80.0, 46.6 (d, J_{C-F} = 5 Hz), 37.3, 32.3, 28.5.

*Phenyl (S,E)-3-((tert-Butoxycarbonyl)amino)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**9**).* Prepared following procedure A using 0.66 g of phosphonate **6**, 0.53 g boc-homophenylalaninal or boc-homophenylalanine aldehyde or boc-hPhe-H (**VI**), and 0.09 g NaH in 12 mL THF. Column chromatography: petroleum ether/EtOAc 5:1. (*E*)-isomer: Yield 0.38 g (43%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.46−7.32 (m, 2H), 7.32−7.15 (m, 6H), 7.10 (d, J = 7.0 Hz, 2H), 5.91 (dd, J_{H-F} = 31.3 Hz, J_{H-H} = 8.6 Hz, 1H), 4.68−4.27 (m, 2H), 2.66−2.43 (m, 2H), 1.97−1.66 (m, 2H), 1.43 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 154.8, 150.8, 148.9 (d, J_{C-F} = 296 Hz), 140.2, 130.2, 128.8, 128.4, 128.0, 126.5, 122.4 (d, J_{C-F} = 4.1 Hz), 122.3, 80.4, 46.3 (d, J_{C-F} = 2.1 Hz), 35.9, 31.9, 28.4. (*Z*)-isomer: yield 0.22 g (22%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.44−7.32 (m, 2H), 7.32−7.13 (m, 6H), 7.04 (d, J = 7.2 Hz, 2H), 6.14 (br, 1H), 4.88−4.44 (br, 2H), 2.59−2.46 (m, 1H), 2.46−2.29 (m, 1H), 1.84 (br, 1H), 1.42 (br, 1H), 1.43 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 154.8, 149.3, 147.5 (d, J_{C-F} = 296 Hz), 140.7, 130.1, 128.5, 128.4, 128.1, 126.3, 124.1 (d, J_{C-F} = 9.9 Hz), 122.4, 80.4, 47.1 (d, J_{C-F} = 3.6 Hz), 36.3, 32.0, 28.4.

*Phenyl (S,E)-3-((tert-Butoxycarbonyl)amino)-5-phenylpent-1-ene-1-sulfonate (**9-(H)**).* Prepared following procedure B using 1.45 g of phosphonate **IV**, 1.24 g boc-homophenylalaninal or boc-homophenylalanine aldehyde or boc-hPhe-H (**VI**), and 6.0 mL 1 M LHMDS in 25 mL THF. Column chromatography: cyclohexane/EtOAc 5:1. (*E*)-isomer: Yield 0.80 g (41%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.43−7.35 (m, 2H), 7.35−7.27 (m, 3H), 7.27−7.20 (m, 3H), 7.18−7.10 (m, 2H), 6.76 (dd, J = 15.1, 5.1 Hz, 1H), 6.48 (dd, J = 15.1, 1.5 Hz, 1H), 4.58 (d, J = 8.2 Hz, 1H), 4.33 (s, 1H), 2.77−2.52 (m, 2H), 1.98−1.72 (m, 2H), 1.49 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 154.9, 150.4, 149.6, 140.2, 130.0, 128.8, 128.4, 127.4, 126.6, 124.4, 122.6, 80.5, 51.0, 35.6, 32.0, 28.4.

*tert-Butyl (S,E)-(1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)carbamate (**10**).* Synthesized following procedure B using 0.92 g phosphonate **7**, 0.74 g boc-homophenylalaninal or boc-homophenylalanine aldehyde or boc-hPhe-H (**VI**), and 2.8 mL KHMDS (1 M in THF) in 14 mL THF. Column chromatography: DCM/EtOAc 100:1. (*E*)-isomer: yield: 0.57 g (47%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.39−7.16 (m, 8H), 7.10 (d, J = 7.0 Hz, 2H), 5.76 (dd, J_{H-F} = 32.8 Hz, J_{H-H} = 8.1 Hz, 1H), 4.59−4.38 (m, 2H), 4.35 (s, 2H), 2.62−2.38 (m, 2H), 1.91−1.64 (m, 3H), 1.45 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 154.8, 152.1 (d, J_{C-F} = 300 Hz), 140.5, 131.0, 129.5, 129.2, 128.7, 128.4, 126.8, 126.4, 120.9 (d, J_{C-F} = 3.1 Hz), 80.3, 58.9, 46.1, 36.2, 31.8, 28.5. The (*Z*)-isomer was not isolated.

*tert-Butyl (S,E)-(1-Fluoro-5-(methylthio)-1-(phenylsulfonyl)pent-1-en-3-yl)carbamate (**11**).* Synthesized following procedure B using 0.93 g phosphonate **5**, 0.75 g boc-Met-H, and 3.15 mL LHMDS (1 M in THF) in 20 mL THF. Column chromatography: Petroleum ether/EtOAc 3:1. (*E*)-isomer: Yield 0.88 g (75%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.94 (d, J = 7.8 Hz, 2H), 7.70 (t, J = 7.3 Hz, 1H), 7.58 (t, J = 7.6 Hz, 2H), 6.21 (br d, J_{H-F} = 31.5 Hz, 1H), 4.74 (br s, 1H), 4.55 (br s, 1H), 2.57−2.40 (m, 2H), 2.07 (s, 3H), 2.00−1.78 (m, 2H), 1.37 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 154.9, 154.7 (d, J_{C-F} = 300 Hz), 137.1, 134.7, 129.6, 128.8, 118.0 (d, J_{C-F} = 4.6 Hz), 80.4, 46.0 (d, J_{C-F} = 1.0 Hz), 33.8, 30.2, 28.3, 15.7. The (*Z*)-isomer was not further characterized.

*tert-Butyl (R,E)-(1-(Benzylthio)-4-fluoro-4-(phenylsulfonyl)but-3-en-2-yl)carbamate (**12**).* Synthesized following procedure B using 0.55 g phosphonate **5**, 0.55 g boc-L-Cys(Bn)-H, and 1.85 mL LHMDS (1 M in THF) in 10 mL THF. Column chromatography: Petroleum ether/EtOAc 3:1. (*E*)-isomer: yield 0.44 g (54%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.00−7.92 (m, 3H), 7.75−7.53 (m, 5H), 7.30−7.27 (m, 3H), 6.23 (dd, J_{H-F} = 31.8 Hz, J_{H-H} = 8.5 Hz, 1H), 4.83 (s, 1H), 4.59 (s, 1H), 3.70 (s, 2H), 2.73−2.44 (m, 2H), 1.39 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 156.9, 154.8, 137.5, 137.1, 135.0, 134.7, 129.7, 129.6, 129.2, 129.0, 128.9, 128.8, 127.5, 117.5, 80.5, 46.0, 36.7, 35.4, 28.4.

Peptide Chemistry. *Coupling Protocols. Procedure C.* The specified carboxylic acid (1.0 equiv) and HOBt (1.0 equiv) were dissolved in DCM (8 mL) and cooled to 0 °C. DIEA (3.0 equiv) was added slowly, and the mixture was stirred until all materials were solubilized. TBTU (1.0 equiv) was added in one portion and the mixture was stirred for 15 min at 0 °C. The specific amine (1.0 equiv) was dissolved in DCM and added dropwise to the reaction mixture. After stirring for 12 h at rt, DCM was removed in vacuo and the residue was dissolved in EtOAc. The organic phase was washed with water (5×), conc. $NaHCO_3$ (2×), and 1 M aq HCl (2×) and then dried with Na_2SO_4. Evaporation of the solvent gave the crude product, which was purified by column chromatography.

Procedure D. The specified carboxylic acid (1.0 equiv) and TBTU (1.0 equiv) were dissolved in DCM and cooled to 0 °C. DIEA (3.0 equiv) was added and the mixture was stirred for 15 min. The respective amine (1.0 equiv) was added in one portion. After stirring for 12 h at rt, DCM was removed in vacuo, and the residue was dissolved in EtOAc. The organic phase was washed with water (5×), conc. $NaHCO_3$ (2×), and 1 M aq HCl (2×) and then dried with Na_2SO_4. Evaporation of the solvent gave the crude product, which was purified by column chromatography.

Procedure E. The respective carboxylic acid (**IX** or **X**, 1.0 equiv) and HOBt (1.0 equiv) were dissolved in DMF (5 mL) and cooled to 0 °C. DIEA (3.0 equiv) was added and the mixture was stirred for 15 min. The respective amine (1.0 equiv) was added in one portion and the mixture was stirred for 12 h at rt. EtOAc was added and aqueous work-up was performed with 5% aq LiCl (5×), water (2×), and saturated aq $NaHCO_3$ (2×). After evaporation of the solvent, the residue was purified by column chromatography.

Removal of Boc Groups. Procedure F. Boc-protected compounds (**8**−**10**) were dissolved in 4 M HCl/dioxane solution (5−10 mL). After stirring for 30 min at rt, dioxane was evaporated in vacuo and the residue was washed several times with diethyl ether. The precipitate was collected and dried under reduced pressure. The amines were obtained as hydrochloride salts and were sufficiently pure to be used in the next step without further purification.

Procedure G. Boc-protected compounds (**11**−**12**) were dissolved in DCM (10 mL) and cooled to 0 °C. 1 mL TFA was added dropwise, and the resulting mixture was stirred for 1 h. After evaporation to dryness in high vacuum, the residue was washed with diethyl ether and collected by filtration. The amines were obtained as TFA salt and directly used for the next step.

*(S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-amine Hydrochloride (**13**).* Prepared following procedure F using 0.60 g of compound **8** in 12 mL 4 M HCl/dioxane. Yield: 0.47 g (93%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.65 (s, 3H), 8.03 (d, J = 7.8 Hz, 2H), 7.89 (t, J = 7.3 Hz, 1H), 7.77 (t, J = 7.6 Hz, 2H), 7.32−7.13 (m, 3H), 7.06 (d, J = 7.3 Hz, 2H), 6.54 (dd, J_{H-F} = 33 Hz, J_{H-H} = 9.8 Hz, 1H), 4.09−3.90 (m, 1H), 2.62−2.34 (m, 2H), 2.25−2.07 (m, 1H), 2.07−1.89 (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 155.2 (d, J_{C-F} = 301 Hz), 139.9, 135.9, 135.6, 130.2, 128.54, 128.45, 128.1, 126.2, 114.5 (d, J_{C-F} = 3.7 Hz), 45.0 (d, J_{C-F} = 2.7 Hz), 33.4 (d, J_{C-F} = 1.4 Hz), 30.5.

*Phenyl (S,E)-3-Amino-1-fluoro-5-phenylpent-1-ene-1-sulfonate Hydrochloride (**14**).* Prepared following procedure F using 0.3 g of compound **9** in 7 mL 4 M HCl/dioxane. Yield: 0.23 g (89%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.72 (s, 3H), 7.56−7.47 (m, 2H), 7.47−7.37 (m, 3H), 7.31 (t, J = 7.2 Hz, 2H), 7.26−7.19 (m, 1H), 7.19−7.09 (m, 2H), 6.42 (dd, J_{H-F} = 32 Hz, J_{H-H} = 9.7 Hz, 1H), 4.15 (td, J = 9.2, 5.3 Hz, 1H), 2.48−2.32 (m, 2H), 2.19−2.01 (m, 1H), 2.00−1.81 (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 149.3 (d, J_{C-F} = 300 Hz), 148.6, 140.0, 130.6, 128.5, 128.4, 128.1, 126.3, 122.0, 118.4 (d, J_{C-F} = 3.2 Hz), 45.2 (d, J_{C-F} = 2.2 Hz), 40.4, 33.3, 30.4.

*Phenyl (S,Z)-3-Amino-1-fluoro-5-phenylpent-1-ene-1-sulfonate Hydrochloride (**14-(Z)**).* Prepared following procedure F using 0.2 g of compound **9**-(***Z***) in 7 mL 4 M HCl/dioxane. Yield: 0.16 g (93%). ^{1}H

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

NMR (300 MHz, DMSO-d_6): δ 8.68 (s, 3H), 7.59−7.45 (m, 2H), 7.44−7.13 (m, 6H), 7.08−6.98 (m, 2H), 6.73 (dd, J_{H-F} = 21.2 Hz, J_{H-H} = 10.6 Hz, 1H), 4.27 (q, J = 7.8, 1H), 2.29−2.02 (m, 2H), 1.95−1.80 (m, 1H), 1.68−1.47 (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 148.3, 148.1 (d, J_{C-F} = 288 Hz), 140.1, 130.6, 128.7, 128.4, 128.0, 126.1, 122.1, 120.6 (d, J = 16 Hz), 45.9 (d, J_{C-F} = 6.7 Hz), 34.3, 30.2.

*Phenyl (S,E)-3-Amino-5-phenylpent-1-ene-1-sulfonate Hydrochloride (**14-(H)**).* Prepared following procedure F using 0.78 g of compound **9**-(**H**) in 14 mL 4 M HCl/dioxane. Yield: 0.34 g (51%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.64 (s, 3H), 7.50−7.41 (m, 2H), 7.41−7.34 (m, 2H), 7.34−7.24 (m, 3H), 7.24−7.17 (m, 1H), 7.16−7.10 (m, 2H), 6.81 (dd, J_{H-H} = 15.3, 6.8 Hz, 1H), 4.15−3.95 (m, 1H), 2.61−2.38 (m, 2H), 2.10−1.84 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 149.0, 145.3, 140.3, 130.1, 128.5, 128.2, 127.6, 127.6, 126.2, 122.6, 50.1, 33.3, 30.4.

*(S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-amine Hydrochloride (**15**).* Prepared following procedure F using 0.46 g of compound **10** in 10 mL 4 M HCl/dioxane. Yield: 0.37 g (95%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.66 (s, 3H), 7.48−7.27 (m, 7H), 7.26−7.09 (m, 3H), 6.13 (dd, J_{H-F} = 33 Hz, J_{H-H} = 9.9 Hz, 1H), 4.87 (s, 2H), 4.04 (td, J = 9.3, 5.0 Hz, 1H), 2.39 (t, J = 8.1 Hz, 2H), 2.13−1.97 (m, 1H), 1.91−1.74 (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 154.1 (d, J_{C-F} = 304 Hz), 140.2, 131.2, 129.0, 128.7, 128.5, 128.2, 127.0, 126.2, 115.7 (d, J_{C-F} = 2.9 Hz), 57.6, 45.0 (d, J_{C-F} = 2.2 Hz), 33.5, 30.4.

*(S,E)-1-Fluoro-5-(methylthio)-1-(phenylsulfonyl)pent-1-en-3-aminium 2,2,2-Trifluoroacetate (**16**).* Prepared following procedure G with 0.88 g compound **11**. Yield: 0.60 g (65%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 7.89−7.82 (m, 2H), 7.82−7.63 (m, 1H), 7.57−7.52 (m, 2H), 6.21 (d, J_{H-F} = 32 Hz, 1H), 4.46 (m, 1H), 2.69−2.41 (m, 2H), 2.06 (s, 3H), 2.05−1.75 (m, 2H).

*(R,E)-1-(Benzylthio)-4-fluoro-4-(phenylsulfonyl)but-3-en-2-aminium 2,2,2-Trifluoroacetate (**17**).* Prepared following procedure G with 0.44 g compound **12**. Yield: 0.37 g (81%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.04 (s, 3H), 7.95−7.84 (m, 2H), 7.77−7.55 (m, 4H), 7.25−7.19 (m, 4H), 6.44 (dd, J_{H-F} = 30 Hz, J_{H-H} = 9.1 Hz, 1H), 4.20−4.01 (m, 1H), 3.71 (s, 2H), 2.91−2.62 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 161.4, 160.8, 155.5, 136.6, 135.5, 135.0, 129.9, 129.7, 129.2, 129.1, 128.9, 127.9, 117.3, 113.5, 111.1, 77.2, 46.2, 36.2, 33.3.

*tert-Butyl ((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)-pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (**18**).* Synthesized following procedure C with 0.180 g compound **13**, 0.135 g boc-L-Phe-OH, 0.162 g TBTU, 0.078 g HOBt, and 0.265 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.21 g (73%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.00−7.89 (m, 2H), 7.79−7.68 (m, 1H), 7.67−7.55 (m, 2H), 7.39−7.12 (m, 8H), 7.10−7.00 (m, 2H), 6.00 (dd, J_{H-F} = 32 Hz, J_{H-H} = 8.5 Hz, 1H), 5.82 (d, J = 8.0 Hz, 1H), 4.91 (s, 1H), 4.69 (quint, J = 7.7 Hz, 1H), 4.20 (q, J = 7.5 Hz, 1H), 3.04 (dd, J = 13.7, 6.6 Hz, 1H), 2.95 (dd, J = 13.7, 7.6 Hz, 1H), 2.52 (t, J = 7.2 Hz, 2H), 1.96−1.75 (m, 2H), 1.40 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.8, 155.6, 154.8 (d, J_{C-F} = 301 Hz), 140.4, 137.1, 136.6, 134.8, 129.7, 129.4, 129.0, 128.8, 128.7, 128.4, 127.3, 126.4, 117.3 (d, J_{C-F} = 4.7 Hz), 80.6, 56.2, 45.0, 38.4, 35.7, 31.8, 28.4.

*Phenyl (S,E)-3-((S)-2-((tert-Butoxycarbonyl)amino)-3-phenylpropanamido)-1-fluoro-5-phenylpent-1-Ene-1-sulfonate (**19**).* Synthesized following procedure C with 0.160 g compound **14**, 0.114 g boc-Phe-OH, 0.138 g TBTU, 0.066 g HOBt, and 0.225 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.20 g (78%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.47−7.35 (m, 2H), 7.35−7.10 (m, 9H), 7.08−6.95 (m, 4H), 5.65 (d, J = 7.9 Hz, 1H), 5.60 (dd, J_{H-F} = 32 Hz, J_{H-H} = 8.3 Hz, 1H), 5.02 (d, J = 7.1 Hz, 1H), 4.68 (q, J = 7.4 Hz, 1H), 4.18 (dd, J = 8.6, 5.6 Hz, 1H), 3.04 (dd, J = 13.4, 6.2 Hz, 1H), 2.85 (dd, J = 13.4, 8.4 Hz, 1H), 2.55−2.35 (m, 2H), 1.83−1.60 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.8, 155.8, 149.4, 149.0 (d, J_{C-F} = 297 Hz), 147.0, 140.1, 136.5, 130.2, 129.3, 129.0, 128.8, 128.4, 128.0, 127.4, 126.6, 122.5, 121.3 (d, J_{C-F} = 5.1 Hz), 80.6, 56.2, 44.8, 38.7, 35.3, 31.7, 28.4.

*Phenyl (S,Z)-3-((S)-2-((tert-Butoxycarbonyl)amino)-3-phenylpropanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**19-(Z)**).* Synthesized following procedure C with 0.128 g compound **14**-(**Z**), 0.091 g boc-L-Phe-OH, 0.110 g TBTU, 0.053 g HOBt, and 0.204 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.150 g (81%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.46−7.08 (m, 13H), 6.97 (d, J = 6.8 Hz, 2H), 5.94−5.59 (m, 2H), 4.96 (s, 1H), 4.80 (q, J = 8.2 Hz, 1H), 4.25−4.11 (m, 1H), 3.14−2.87 (m, 2H), 2.44−2.17 (m, 2H), 1.82−1.61 (m, 1H), 1.53−1.42 (m, 1H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 171.0, 155.6, 149.2, 148.0 (d, J_{C-F} = 289 Hz), 140.6, 136.7, 130.2, 129.5, 128.9, 128.6, 128.4, 128.1, 127.2, 126.3, 123.0, 123.0 (d, J_{C-F} = 12 Hz), 80.6, 56.1, 46.1 (d, J_{C-F} = 5.9 Hz), 38.5, 35.7, 31.8, 28.4.

*tert-Butyl ((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (**20**).* Synthesized following procedure C with 0.230 g compound **15**, 0.165 g boc-L-Phe-OH, 0.200 g TBTU, 0.095 g HOBt, and 0.325 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.30 g (85%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.42−7.15 (m, 11H), 7.13−6.97 (m, 4H), 5.68 (d, J = 8.0 Hz, 1H), 5.53 (dd, J_{H-F} = 33 Hz, J_{H-H} = 8.9 Hz, 1H), 4.99 (br s, 1H), 4.68 (quint, J = 7.8 Hz, 1H), 4.35 (s, 2H), 4.18 (q, J = 7.9 Hz, 1H), 3.03 (dd, J = 13.5, 6.4 Hz, 1H), 2.91 (dd, J = 13.5, 8.0 Hz, 1H), 2.49−2.29 (m, 2H), 1.70−1.56 (m, 2H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.6, 155.6, 152.3 (d, J_{C-F} = 301 Hz), 140.3, 136.6, 131.0, 129.5, 129.4, 129.2, 129.0, 128.7, 128.4, 127.3, 126.7, 126.5, 119.7 (d, J_{C-F} = 4.0 Hz), 80.6, 58.8, 56.2, 44.7 (d, J_{C-F} = 1.3 Hz), 38.6, 35.5 (d, J_{C-F} = 1.4 Hz), 31.7, 28.4.

*tert-Butyl ((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)-pent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-yl)carbamate (**21**).* Synthesized following procedure C with 0.220 g compound **13**, 0.173 g boc-L-Phe(3-Me)-OH, 0.198 g TBTU, 0.095 g HOBt, and 0.330 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.28 g (77%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.00−7.88 (m, 2H), 7.76−7.65 (m, 1H), 7.64−7.53 (m, 2H), 7.31−7.12 (m, 4H), 7.11−6.88 (m, 5H), 6.04 (dd, J_{H-F} = 32.1 Hz, J_{H-H} = 8.4 Hz, 1H), 5.93 (d, J = 8.0 Hz, 1H), 4.92 (s, 1H), 4.78−4.59 (m, 1H), 4.21 (q, J = 7.3 Hz, 1H), 3.01 (dd, J = 13.8, 6.6 Hz, 1H), 2.92 (dd, J = 13.8, 7.5 Hz, 1H), 2.58−2.46 (m, 2H), 2.31 (s, 3H), 1.93−1.77 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 155.62, 154.9 (d, J_{C-F} = 303 Hz), 140.4, 138.7, 137.1, 136.5, 134.8, 130.1, 129.7, 128.9, 128.8, 128.7, 128.4, 128.1, 126.4, 126.4, 117.4 (d, J_{C-F} = 3.6 Hz), 80.6, 56.2, 45.2, 38.2, 35.8, 31.8, 28.4, 21.5.

*Phenyl (S,E)-3-((S)-2-((tert-Butoxycarbonyl)amino)-3-(m-tolyl)-propanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**22**).* Synthesized following procedure C with 0.280 g compound **14**, 0.210 g boc-L-Phe(3-Me)-OH, 0.242 g TBTU, 0.115 g HOBt, and 0.40 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.34 g (75%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.40 (t, J = 7.3 Hz, 2H), 7.35−7.08 (m, 7H), 7.08−7.00 (m, 3H), 6.91 (s, 1H), 6.80 (d, J = 7.3 Hz, 1H), 5.78 (d, J = 7.7 Hz, 1H), 5.65 (dd, J_{H-F} = 31.4 Hz, J_{H-H} = 8.3 Hz, 1H), 5.04 (s, 1H), 4.70 (quint, J = 8.2, 7.0 Hz, 1H), 4.18 (quint, J = 8.7, 8.0 Hz, 1H), 3.01 (dd, J = 13.3, 6.1 Hz, 1H), 2.83 (dd, J = 13.3, 8.6 Hz, 1H), 2.55−2.37 (m, 2H), 2.28 (s, 3H), 1.81−1.65 (m, 2H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 155.6, 149.4, 149.0 (d, J_{C-F} = 297 Hz), 140.2, 138.7, 136.5, 130.2, 130.0, 128.9, 128.8, 128.4, 128.1, 128.0, 126.5, 126.4, 122.4, 121.3 (d, J_{C-F} = 4.9 Hz), 80.6, 56.2, 44.9 (d, J_{C-F} = 1.2 Hz), 38.5, 35.4 (d, J_{C-F} = 0.9 Hz), 31.7, 28.4, 21.4.

*tert-Butyl ((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-yl)carbamate (**23**).* Synthesized following procedure C with 0.230 g compound **15**, 0.174 g boc-L-Phe(3-Me)-OH, 0.200 g TBTU, 0.095 g HOBt, and 0.325 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.28 g (69%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.41−7.12 (m, 9H), 7.04 (d, J = 7.3 Hz, 3H), 6.94 (s, 1H), 6.85 (d, J = 7.3 Hz, 1H), 5.73 (d, J = 7.8 Hz, 1H), 5.56 (dd, J_{H-F} = 32.8 Hz, J_{H-H} = 8.8 Hz, 1H), 4.99 (s, 1H), 4.69 (quint, J = 8.3, 7.8 Hz, 1H), 4.34 (s, 2H), 4.18 (q, J = 8.7 Hz, 1H), 3.01 (dd, J = 13.6, 6.5 Hz, 1H), 2.87 (dd, J = 13.6, 7.9 Hz, 1H), 2.49−2.33 (m, 2H), 2.30 (s, 3H), 1.72−1.57 (m, 2H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.7, 155.6, 152.3 (d, J_{C-F} = 301 Hz), 140.4, 138.7, 136.5, 131.1, 130.1, 129.5, 129.2, 128.9, 128.7, 128.4, 128.1, 126.7, 126.5, 119.8 (d, J_{C-F} = 4.2 Hz), 80.6, 58.8, 56.2, 44.8 (d, J_{C-F} = 1.3 Hz), 38.4, 35.6, 31.7, 28.4, 21.5.

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

*tert-Butyl ((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)-pent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)carbamate (**24**).* Synthesized following procedure C with 0.200 g compound **13**, 0.173 g boc-L-Phe(4-Me)-OH, 0.198 g TBTU, 0.095 g HOBt, and 0.30 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.25 g (77%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.94 (d, J = 7.8 Hz, 2H), 7.70 (t, J = 7.2 Hz, 1H), 7.59 (t, J = 7.6 Hz, 2H), 7.35−6.92 (m, 9H), 6.03 (dd, J_{H-F} = 32.0 Hz, J_{H-H} = 8.4 Hz, 1H), 5.89 (d, J = 8.0 Hz, 1H), 4.93 (s, 1H), 4.70 (quint, J = 7.4 Hz, 1H), 4.20 (q, J = 7.2 Hz, 1H), 3.01 (dd, J = 13.6, 6.4 Hz, 1H), 2.91 (dd, J = 13.6, 7.6 Hz, 1H), 2.52 (t, J = 7.5 Hz, 2H), 2.33 (s, 3H), 1.95−1.77 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 155.6, 154.8 (d, J_{C-F} = 301 Hz), 140.4, 137.2, 137.0, 134.7, 133.4, 129.70, 129.65, 129.3, 128.8, 128.7, 128.4, 126.4, 117.4 (d, J_{C-F} = 4.4 Hz), 80.5, 56.2, 45.0, 38.0, 35.8, 31.8, 28.4, 21.2.

*Phenyl (S,E)-3-((S)-2-((tert-Butoxycarbonyl)amino)-3-(p-tolyl)-propanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**25**).* Synthesized following procedure C with 0.292 g compound **14**, 0.220 g boc-L-Phe(4-Me)-OH, 0.252 g TBTU, 0.121 g HOBt, and 0.410 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.39 g (83%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.47−7.35 (m, 2H), 7.35−7.15 (m, 6H), 7.10−6.98 (m, 4H), 6.92 (d, J = 8.0 Hz, 2H), 5.75 (d, J = 7.9 Hz, 1H), 5.62 (dd, J_{H-F} = 31 Hz, J_{H-H} = 9.0 Hz, 1H), 5.05 (d, J = 7.8 Hz, 1H), 4.70 (quint, J = 7.7 Hz, 1H), 4.20−4.08 (m, 1H), 3.00 (dd, J = 13.5, 6.1 Hz, 1H), 2.81 (dd, J = 13.5, 8.3 Hz, 1H), 2.53−2.36 (m, 2H), 2.29 (s, 3H), 1.83−1.62 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 155.6, 149.4, 149.0 (d, J_{C-F} = 297 Hz), 140.2, 137.0, 133.3, 130.2, 129.7, 129.2, 128.8, 128.4, 128.0, 126.5, 122.4, 121.4, 121.3, 80.6, 56.3, 44.8 (d, J_{C-F} = 1.5 Hz), 38.3, 35.4, 31.7, 28.4, 21.1.

*Phenyl (S,E)-3-((S)-2-((tert-Butoxycarbonyl)amino)-3-(p-tolyl)-propanamido)-5-phenylpent-1-ene-1-sulfonate (**25-(H)**).* Synthesized following procedure C with 0.34 g compound **14**-(**H**), 0.28 g boc-L-Phe(4-Me)-OH, 0.32 g TBTU, 0.14 g HOBt, and 0.70 mL DIEA. Purification by column chromatography (cyclohexane/EtOAc 4:1). Yield: 0.24 g (42%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.42−7.32 (m, 2H), 7.31−7.02 (m, 12H), 6.52 (dd, J_{H-H} = 15.2, 4.6 Hz, 1H), 5.92 (dd, J_{H-H} = 37.5, 11.8 Hz, 2H), 5.00 (s, 1H), 4.63−4.47 (m, 1H), 4.31−4.18 (q, 1H), 3.08−2.89 (m, 2H), 2.62−2.42 (m, 2H), 2.31 (s, 3H), 1.88−1.60 (m, 2H), 1.43 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 171.1, 149.5, 149.1, 140.1, 137.3, 133.0, 129.8 (d, J = 3.6 Hz), 129.1 (d, J = 6.7 Hz), 128.6, 128.3, 127.3, 126.4, 124.4, 122.6, 56.5, 49.2, 35.1, 31.7, 28.3, 21.0. mp 93−95 °C.

*tert-Butyl ((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)carbamate (**26**).* Synthesized following procedure C with 0.230 g compound **15**, 0.174 g boc-L-Phe(4-Me)-OH, 0.200 g TBTU, 0.095 g HOBt, and 0.325 mL DIEA. Purification by column chromatography (petroleum ether/EtOAc 3:1). Yield: 0.290 g (79%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.43−7.14 (m, 8H), 7.12−6.99 (m, 4H), 6.96 (d, J = 7.4 Hz, 2H), 5.66 (d, J = 7.8 Hz, 1H), 5.54 (dd, J_{H-F} = 32.9 Hz, J_{H-H} = 8.7 Hz, 1H), 4.98 (s, 1H), 4.69 (quint, J = 8.3 Hz, 1H), 4.34 (s, 2H), 4.15 (q, J = 7.8, 7.4 Hz, 1H), 3.01 (dd, J = 13.6, 5.9 Hz, 1H), 2.85 (dd, J = 13.6, 8.0 Hz, 1H), 2.54−2.33 (m, 2H), 2.30 (s, 3H), 1.70−1.55 (m, 3H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.7, 155.6, 152.3 (d, J_{C-F} = 301 Hz), 140.4, 137.0, 133.4, 131.0, 129.7, 129.5, 129.3, 129.2, 128.7, 128.4, 126.7, 126.5, 119.8 (d, J_{C-F} = 4.0 Hz), 80.5, 58.8, 56.2, 44.7 (d, J_{C-F} = 1.6 Hz), 38.1, 35.6, 31.7, 28.4, 21.2.

*(S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-aminium Chloride (**27**).* Synthesized by following procedure F with 0.21 g of compound **18** in 10 mL 4 M HCl/dioxane. Yield: 0.19 g (99%) The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,E)-1-Fluoro-1-(phenoxysulfonyl)-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-aminium Chloride (**28**).* Synthesized by following procedure F with 0.20 g of compound **19** in 8 mL 4 M HCl/dioxane. Yield: 0.18 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,Z)-1-Fluoro-1-(phenoxysulfonyl)-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-aminium Chloride (**28-(Z)**).* Synthesized by following procedure F with 0.15 g of compound **19**-(**Z**) in 8 mL 4 M HCl/dioxane. Yield: 0.13 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)-amino)-1-oxo-3-phenylpropan-2-aminium Chloride (**29**).* Synthesized by following procedure F with 0.30 g of compound **20** in 10 mL 4 M HCl/dioxane. Yield: 0.26 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-aminium Chloride (**30**).* Synthesized by following procedure F with 0.28 g of compound **21** in 8 mL 4 M HCl/dioxane. Yield: 0.25 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,E)-1-Fluoro-1-(phenoxysulfonyl)-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-aminium Chloride (**31**).* Synthesized by following procedure F with 0.34 g of compound **22** in 10 mL 4 M HCl/dioxane. Yield: 0.30 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)-amino)-1-oxo-3-(m-tolyl)propan-2-aminium Chloride (**32**).* Synthesized by following procedure F with 0.28 g of compound **23** in 8 mL 4 M HCl/dioxane. Yield: 0.25 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-aminium Chloride (**33**).* Synthesized by following procedure F with 0.25 g of compound **24** in 8 mL 4 M HCl/dioxane. Yield: 0.22 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-(((S,E)-1-Fluoro-1-(phenoxysulfonyl)-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-aminium Chloride (**34**).* Synthesized by following procedure F with 0.39 g of compound **25** in 10 mL 4 M HCl/dioxane. Yield: 0.34 g (99%). The crude product was used without further purification and characterization for the next step.

*(S)-1-Oxo-1-(((S,E)-1-(phenoxysulfonyl)-5-phenylpent-1-en-3-yl)-amino)-3-(p-tolyl)propan-2-aminium Chloride (**34-(H)**).* Synthesized by following procedure F with 0.20 g of compound **25**-(**H**) in 4 mL 4 M HCl/dioxane. Yield: 0.15 g (89%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 9.06 (d, J = 8.4 Hz, 1H), 8.57 (s, 3H), 7.48−7.38 (m, 2H), 7.37−7.21 (m, 5H), 7.20−6.97 (m, 7H), 6.58 (dd, J = 15.3, 4.4 Hz, 1H), 6.21 (dt, J = 15.3, 2.1 Hz, 1H), 4.46 (s, 1H), 4.17 (s, 1H), 3.17−2.94 (m, 2H), 2.63−2.50 (m, 2H), 2.23 (d, J = 2.7 Hz, 3H), 1.82−1.62 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 167.8, 150.5, 149.1, 141.1, 136.4, 131.7, 130.0, 129.2 (d, J = 5.5 Hz), 128.3 (d, J = 6.1 Hz), 127.4, 125.9, 123.4, 122.6, 53.5, 49.2, 34.3, 31.1, 20.7.

*(S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)-amino)-1-oxo-3-(p-tolyl)propan-2-aminium Chloride (**35**).* Synthesized by following procedure F with 0.29 g of compound **26** in 10 mL 4 M HCl/dioxane. Yield: 0.26 g (99%). The crude product was used without further purification and characterization for the next step.

*Phenyl (S,E)-1-Fluoro-3-((S)-2-(isonicotinamido)-3-phenylpropanamido)-5-phenylpent-1-ene-1-sulfonate (**2a**).* Synthesized by following procedure E with 0.14 g of compound **14**, 0.10 g compound **X**, 0.13 g TBTU, 0.062 HOBt, and 0.21 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:3 to 0:1. Yield: 0.16 g (73%), colorless solid. mp 71−73 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.70 (d, J = 5.9 Hz, 2H), 7.64 (d, J = 6.0 Hz, 2H), 7.51 (d, J = 7.6 Hz, 1H), 7.45−7.34 (m, 2H), 7.34−7.11 (m, 9H), 7.12−7.02 (m, 2H), 6.97 (d, J = 6.6 Hz, 2H), 6.22 (d, J = 7.6 Hz, 1H), 5.61 (dd, J_{H-F} = 31.0 Hz, J_{H-H} = 9.2 Hz, 1H), 4.80−4.62 (m, 2H), 3.22 (dd, J = 13.3, 6.2 Hz, 1H), 2.96 (dd, J = 13.3, 9.0 Hz, 1H), 2.52−2.27 (m, 2H), 1.82−1.55 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.2, 164.8, 149.3, 149.2 (d, J_{C-F} = 297 Hz), 147.2, 142.1, 139.9, 136.0, 130.3, 129.3, 129.1, 128.8, 128.3, 128.1, 127.7, 126.6, 122.4, 121.8, 120.94 (d, J_{C-F} = 4.8 Hz), 55.7, 44.9 (d, J_{C-F} = 1.9 Hz), 38.9, 35.3, 31.7. mp 71−73 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{32}H_{30}FN_3O_5S$, 588.20; found, 588.3. Purity: 97%. $[\alpha]_D^{22}$ 10° (c 0.5, $CHCl_3$).

*Phenyl (S,Z)-1-Fluoro-3-((S)-2-(isonicotinamido)-3-phenylpropanamido)-5-phenylpent-1-ene-1-sulfonate (**2a-(Z)**).* Synthesized by following procedure E with 0.10 g of compound **14**-(**Z**), 0.073 g compound **X**, 0.086 g TBTU, 0.041 g HOBt, and 0.14 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:3. Yield: 0.12 g

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

(79%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.65 (d, J = 5.5 Hz, 2H), 7.85 (d, J = 7.4 Hz, 1H), 7.60 (d, J = 5.3 Hz, 2H), 7.44−7.07 (m, 14H), 6.98−6.83 (m, 2H), 6.38 (d, J = 7.2 Hz, 1H), 5.58 (dd, J_{H-F} = 20.5 Hz, J_{H-H} = 10.1 Hz, 1H), 4.96 (dt, J = 14.8, 6.6 Hz, 1H), 4.77 (q, J = 7.7 Hz, 1H), 3.24 (dd, J = 13.5, 6.6 Hz, 1H), 3.08 (dd, J = 13.4, 8.7 Hz, 1H), 2.37−2.22 (m, 1H), 2.20−2.08 (m, 1H), 1.69−1.52 (m, 1H), 1.43−1.32 (m, 1H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.7, 164.9, 149.3, 149.1, 148.25 (d, J_{C-F} = 290 Hz), 141.9, 140.3, 136.4, 130.2, 129.5, 129.0, 128.6, 128.7, 128.2, 127.5, 126.4, 122.5 (d, J_{C-F} = 11.4 Hz), 122.3, 121.8, 55.8, 45.9 (d, J_{C-F} = 5.5 Hz), 38.6, 35.8 (d, J_{C-F} = 2.3 Hz), 31.6. mp 173−175 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{32}H_{30}FN_3O_5S$, 588.20; found, 588.3. Purity: 95%. $[\alpha]_D^{22}$ 13° (c 0.5, $CHCl_3$).

*Phenyl (S,E)-3-((S)-2-(2,3-Dihydrobenzo[b][1,4]dioxine-6-carboxamido)-3-phenylpropanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**2b**).* Synthesized by following procedure D with 0.10 g of compound **28**, 0.035 g 1,4-benzodioxane-6-carboxylic acid, 0.062 g TBTU, and 0.10 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1. Yield: 0.088 g (71%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.46−7.35 (m, 2H), 7.35−7.10 (m, 11H), 7.04 (d, J = 6.5 Hz, 2H), 6.96−6.80 (m, 4H), 6.60 (d, J = 7.4 Hz, 1H), 5.60 (dd, J_{H-F} = 31.0 Hz, J_{H-H} = 9.0 Hz, 1H), 4.82 (q, J = 7.4 Hz, 1H), 4.67 (quint, J = 7.8 Hz, 1H), 4.34−4.19 (m, 4H), 3.17 (dd, J = 13.2, 6.2 Hz, 1H), 2.96 (dd, J = 13.2, 8.6 Hz, 1H), 2.45−2.26 (m, 2H), 1.72−1.52 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 167.0, 149.4, 149.0 (d, J_{C-F} = 297 Hz), 147.0, 143.6, 140.1, 136.3, 130.2, 129.4, 129.0, 128.6, 128.3, 128.0, 127.4, 126.4, 122.4, 121.4 (d, J_{C-F} = 4.4 Hz), 120.7, 117.5, 116.8, 64.7, 64.3, 55.2, 44.8 (d, J_{C-F} = 2.0 Hz), 38.9, 35.2 (d, J_{C-F} = 1.5 Hz), 31.6. mp 167−168 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{35}H_{33}FN_2O_7S$, 645.2; found, 645.3. Purity: 98%. $[\alpha]_D^{22}$ 12° (c 1, $CHCl_3$).

*Phenyl (S,E)-1-Fluoro-3-((S)-2-(isonicotinamido)-3-(p-tolyl)-propanamido)-5-phenylpent-1-ene-1-sulfonate (**2d**).* Synthesized by following procedure D with 0.12 g of compound **34**, 0.028 g isonicotinic acid, 0.073 g TBTU, and 0.12 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:3. Yield: 0.10 g (77%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.69 (d, J = 3.5 Hz, 2H), 7.56 (d, J = 3.5 Hz, 2H), 7.44−7.36 (m, 3H), 7.36−7.12 (m, 7H), 7.06 (d, J = 7.7 Hz, 2H), 7.01−6.88 (m, 4H), 6.29 (d, J = 7.7 Hz, 1H), 5.62 (dd, J_{H-F} = 31.0 Hz, J_{H-H} = 9.1 Hz, 1H), 4.81−4.63 (m, 2H), 3.16 (dd, J = 13.4, 5.7 Hz, 1H), 2.92 (dd, J = 12.9, 9.6 Hz, 1H), 2.47−2.30 (m, 2H), 2.27 (s, 3H), 1.80−1.56 (m, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.3, 165.3, 150.4, 149.3, 149.2 (d, J_{C-F} = 297 Hz), 141.0, 139.9, 137.4, 132.8, 130.2, 129.8, 129.2, 128.8, 128.2, 128.1, 126.6, 122.4, 121.2, 120.9 (d, J_{C-F} = 5.2 Hz), 55.6, 44.8 (d, J_{C-F} = 1.6 Hz), 38.6, 35.4, 31.6, 21.1. mp 78−79 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{32}FN_3O_5S$, 602.21; found, 602.3. Purity: 97%. $[\alpha]_D^{22}$ 8° (c 0.5, MeOH).

*Phenyl (S,E)-3-((S)-2-(Isonicotinamido)-3-((p-tolyl)-propanamido)-5-phenylpent-1-ene-1-sulfonate (**2d-(H)**).* Synthesized by following procedure D with 0.21 g of compound **34-(H)**, 0.054 g isonicotinic acid, 0.014 g TBTU, and 0.28 mL DIEA. Column chromatography: cyclohexane/EtOAc 1:3. Yield: 0.080 g (35%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.71 (d, J = 13.0 Hz, 2H), 7.65 (d, J = 21.2 Hz, 2H), 7.52−7.41 (m, 1H), 7.42−7.32 (m, 1H), 7.32−7.23 (m, 3H), 7.22−7.15 (m, 4H), 7.15−7.03 (m, 2H), 7.02−6.90 (m, 2H), 6.71−6.59 (m, 1H), 6.51 (dd, J = 15.2, 5.1 Hz, 1H), 6.32 (d, J = 7.7 Hz, 1H), 5.99 (dd, J = 15.2, 1.5 Hz, 1H), 4.92−4.73 (m, 1H), 4.53 (s, 1H), 3.29−3.13 (m, 1H), 3.07 (dd, J = 13.5, 8.9 Hz, 1H), 2.60−2.40 (m, 2H), 2.32 (s, 3H), 1.83−1.58 (m, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.5, 164.5, 149.4, 148.7, 139.9, 137.6, 132.7, 129.9, 129.1, 128.7, 128.3, 127.4, 126.5, 124.6, 122.5, 126.6, 55.9, 49.6, 35.1, 31.7, 21.1. mp 84−86 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{33}N_3O_5S$, 584.21; found, 584.3. Purity: 98%. $[\alpha]_D^{22}$ 11° (c 0.5, MeOH).

*Phenyl (S,E)-3-((S)-2-(2,3-Dihydrobenzo[b][1,4]dioxine-6-carboxamido)-3-(p-tolyl)propan-amido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**2e**).* Synthesized by following procedure D with 0.100 g of compound **34**, 0.034 g 1,4-benzodioxane-6-carboxylic acid, 0.060 g TBTU, and 0.10 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1. Yield: 0.094 g (78%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.46−7.35 (m, 2H), 7.34−7.23 (m, 5H), 7.23−7.12 (m, 4H), 7.03 (d, J = 7.9 Hz, 2H), 6.97−6.88 (m, 4H), 6.88−6.78 (m, 2H), 6.58 (d, J = 7.7 Hz, 1H), 5.61 (dd, J_{H-F} = 31.2 Hz, J_{H-H} = 9.1 Hz, 1H), 4.85−4.60 (m, 2H), 4.34−4.20 (m, 4H), 3.12 (dd, J = 13.3, 6.1 Hz, 1H), 2.91 (dd, J = 13.3, 8.6 Hz, 1H), 2.44−2.30 (m, 2H), 2.26 (s, 3H), 1.72−1.55 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 166.8, 149.4, 148.9 (d, J_{C-F} = 296 Hz), 147.1, 143.6, 140.1, 137.1, 133.2, 130.2, 129.6, 129.3, 128.6, 128.3, 128.0, 126.4, 122.4, 121.5 (d, J_{C-F} = 4.6 Hz), 120.6, 117.5, 116.8, 64.7, 64.3, 55.3, 44.7 (d, J_{C-F} = 1.7 Hz), 38.5, 35.3, 31.6, 21.1. mp 92−93 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{36}H_{35}FN_2O_7S$, 659.22; found, 659.3. Purity: 95%. $[\alpha]_D^{22}$ 26° (c 1, MeOH).

*Phenyl (S,E)-3-((S)-2-(3,5-Difluorobenzamido)-3-(p-tolyl)-propanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**2f**).* Synthesized by following procedure D with 0.10 g of compound **34**, 0.030 g 3,5-difluorobenzoic acid, 0.060 g TBTU, and 0.10 mL DIEA. Column chromatography: petroleum ether/EtOAc 3:1. Yield: 0.097 g (81%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.46−7.10 (m, 12H), 7.04 (d, J = 7.8 Hz, 2H), 6.99−6.84 (m, 5H), 6.23 (d, J = 7.7 Hz, 1H), 5.58 (dd, J_{H-F} = 31.0 Hz, J_{H-H} = 9.2 Hz, 1H), 4.80−4.60 (m, 2H), 3.15 (dd, J = 13.3, 6.0 Hz, 1H), 2.91 (dd, J = 13.3, 9.1 Hz, 1H), 2.49−2.31 (m, 2H), 2.26 (s, 3H), 1.73−1.56 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.6, 164.9 (t, J_{C-F} = 2.6 Hz), 163.0 (dd, J_{C-F} = 251, 12 Hz), 149.3, 149.2 (d, J_{C-F} = 297 Hz), 140.0, 137.4, 136.8 (t, J_{C-F} = 8.5 Hz), 132.9, 130.2, 129.8, 129.2, 128.7, 128.3, 128.1, 126.6, 122.4, 121.0 (d, J = 5.0 Hz), 110.5 (dd, J_{C-F} = 26.2, 8.8 Hz), 107.4 (t, J_{C-F} = 25.4 Hz), 55.8, 44.9 (d, J_{C-F} = 2.4 Hz), 38.5, 35.4 (d, J_{C-F} = 1.6 Hz), 31.7, 21.1. mp 175−177 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{31}F_3N_2O_5S$, 637.20; found, 637.2. Purity: >99%. $[\alpha]_D^{22}$ 7° (c 0.5, $CHCl_3$).

*Phenyl (S,E)-1-Fluoro-3-((S)-2-(isonicotinamido)-3-(m-tolyl)-propanamido)-5-phenylpent-1-ene-1-sulfonate (**2g**).* Synthesized by following procedure D with 0.11 g of compound **31**, 0.026 g isonicotinic acid, 0.067 g TBTU, and 0.11 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:3. Yield: 0.096 g (77%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.62 (d, J = 4.6 Hz, 2H), 7.49 (d, J = 4.0 Hz, 2H), 7.39−7.03 (m, 11H), 6.97 (d, J = 7.3 Hz, 1H), 6.94−6.81 (m, 3H), 6.74 (d, J = 7.3 Hz, 1H), 6.10 (d, J = 7.5 Hz, 1H), 5.57 (dd, J_{H-F} = 31.0 Hz, J_{H-H} = 9.1 Hz, 1H), 4.74−4.50 (m, 2H), 3.11 (dd, J = 13.1, 5.7 Hz, 1H), 2.88−2.78 (m, 1H), 2.39−2.21 (m, 2H), 2.17 (s, 3H), 1.72−1.49 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.2, 165.3, 150.3, 149.33, 149.25 (d, J_{C-F} = 297 Hz), 147.3, 141.2, 139.9, 138.9, 136.0, 130.2, 130.0, 129.0, 128.8, 128.5, 128.2, 128.1, 126.6, 126.4, 122.4, 121.2, 120.9 (d, J_{C-F} = 5.0 Hz), 55.5, 45.0 (d, J_{C-F} = 2.0 Hz), 38.9, 35.4, 31.6, 21.4. mp 67−68 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{32}FN_3O_5S$, 602.21; found, 602.3. Purity: >99%. $[\alpha]_D^{22}$ 8° (c 0.5, $CHCl_3$).

*Phenyl (S,E)-3-((S)-2-(2,3-Dihydrobenzo[b][1,4]dioxine-6-carboxamido)-3-(m-tolyl)propanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**2h**).* Synthesized by following procedure D with 0.10 g of compound **31**, 0.034 g 1,4-benzodioxane-6-carboxylic acid, 0.060 g TBTU, and 0.10 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1. Yield: 0.088 g (72%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.37−7.27 (m, 2H), 7.27−6.99 (m, 10H), 6.94 (d, J = 7.7 Hz, 1H), 6.91−6.80 (m, 3H), 6.80−6.70 (m, 3H), 6.52 (d, J = 7.7 Hz, 1H), 5.57 (dd, J_{H-F} = 31.1 Hz, J_{H-F} = 9.1 Hz, 1H), 4.73 (q, J = 7.9 Hz, 1H), 4.59 (quint, J = 7.6 Hz, 1H), 4.26−4.09 (m, 4H), 3.07 (dd, J = 13.4, 6.3 Hz, 1H), 2.85 (dd, J = 13.4, 8.5 Hz, 1H), 2.38−2.19 (m, 2H), 2.15 (s, 3H), 1.65−1.46 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 166.9, 149.4, 149.0 (d, J_{C-F} = 297 Hz), 147.1, 143.6, 140.1, 138.7, 136.3, 130.2, 130.1, 128.8, 128.6, 128.3, 128.2, 128.0, 126.8, 126.5, 126.4, 122.4, 121.4 (d, J_{C-F} = 3.6 Hz), 120.6, 117.5, 116.8, 64.7, 64.3, 55.2, 44.9 (d, J_{C-F} = 1.1 Hz), 38.8, 35.3, 31.6, 21.4. mp 126−128 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{36}H_{35}FN_2O_7S$, 659.22; found, 659.3. Purity: 98%. $[\alpha]_D^{22}$ 3° (c 0.5, MeOH).

*Phenyl (S,E)-3-((S)-2-(3,5-Difluorobenzamido)-3-(m-tolyl)-propanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**2i**).* Synthesized by following procedure D with 0.10 g of compound **31**, 0.030 g 3,5-difluorobenzoic acid, 0.060 g TBTU, and 0.10 mL DIEA. Column

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

chromatography: petroleum ether/EtOAc 3:1. Yield: 0.075 g (62%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.36−7.27 (m, 2H), 7.28−7.03 (m, 10H), 7.03−6.94 (m, 2H), 6.94−6.80 (m, 4H), 6.74 (d, J = 7.4 Hz, 1H), 5.82 (d, J = 7.7 Hz, 1H), 5.54 (dd, J_{H-F} = 30.9 Hz, J_{H-H} = 9.1 Hz, 1H), 4.68−4.52 (m, 2H), 3.10 (dd, J = 13.4, 6.0 Hz, 1H), 2.80 (dd, J = 13.3, 9.1 Hz, 1H), 2.41−2.21 (m, J = 7.2 Hz, 2H), 2.19 (s, 3H), 1.68−1.57 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.3, 165.0 (t, J_{C-F} = 2.7 Hz), 163.0 (dd, J_{C-F} = 251, 12 Hz), 149.4, 149.3 (d, J_{C-F} = 298 Hz), 139.9, 139.0, 137.0, 136.9, 136.8, 136.1, 130.2, 130.0, 129.1, 128.8, 128.5, 128.3, 128.1, 126.6, 126.4, 122.4, 120.9 (d, J_{C-F} = 5.0 Hz), 110.5 (dd, J_{C-F} = 26, 8.7 Hz), 107.5 (t, J_{C-F} = 25 Hz), 55.6, 45.0 (d, J_{C-F} = 2.3 Hz), 38.8, 35.4, 31.6, 21.4. mp 129−130 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{31}F_3N_2O_5S$, 637.20; found, 637.3. Purity: 96%. $[\alpha]_D^{22}$ 9° (c 0.5, MeOH).

*Phenyl (S,E)-3-((S)-2-(((Benzyloxy)carbonyl)amino)-3-phenylpropanamido)-1-fluoro-5-phenylpent-1-ene-1-sulfonate (**2j**).* Synthesized by following procedure E with 0.090 g of compound **14**, 0.076 g Cbz-L-Phe-OH, 0.082 g TBTU, 0.039 g HOBt, and 0.13 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1. Yield: 0.12 g (78%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.48−7.14 (m, 17H), 7.06−6.92 (m, 4H), 5.56 (dd, J_{H-F} = 31.1 Hz, J_{H-H} = 9.0 Hz, 2H), 5.32 (d, J = 5.7 Hz, 1H), 5.06 (s, 2H), 4.68 (quint, J = 7.5 Hz, 1H), 4.22 (d, J = 6.6 Hz, 1H), 3.07 (dd, J = 13.4, 6.0 Hz, 1H), 2.83 (dd, J = 13.4, 8.6 Hz, 1H), 2.52−2.30 (m, 2H), 1.80−1.64 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.3, 156.4, 149.4, 149.0 (d, J_{C-F} = 297 Hz), 140.1, 136.2, 136.1, 130.2, 129.3, 129.1, 128.8, 128.7, 128.5, 128.3, 128.2, 128.0, 127.5, 126.6, 122.5, 121.1 (d, J_{C-F} = 4.6 Hz), 67.4, 56.6, 44.9 (d, J_{C-F} = 1.1 Hz), 38.9, 35.2, 31.7. mp 127−128 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{33}FN_2O_6S$, 617.21; found, 617.4. Purity: >99%. $[\alpha]_D^{22}$ 17° (c 0.5, $CHCl_3$).

*Phenyl (S,E)-1-Fluoro-3-((S)-2-(4-methylpiperazine-1-carboxamido)-3-phenylpropanamido)-5-phenylpent-1-ene-1-sulfonate (**2k**).* Synthesized by following procedure E with 0.11 g of compound **14**, 0.086 g compound **IX**, 0.095 g TBTU, and 0.155 mL DIEA. Column chromatography: DCM/MeOH 95:5. Yield: 0.096 g (59%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.45−7.35 (m, 2H), 7.35−7.14 (m, 10H), 7.06−6.93 (m, 4H), 6.62 (d, J = 7.6 Hz, 1H), 5.62 (dd, J_{H-F} = 31.2 Hz, J_{H-H} = 9.1 Hz, 1H), 5.21 (d, J = 7.6 Hz, 1H), 4.66 (quint, J = 7.7 Hz, 1H), 4.45 (q, J = 7.9 Hz, 1H), 3.42−3.24 (m, 4H), 3.05 (dd, J = 13.4, 6.5 Hz, 1H), 2.88 (dd, J = 13.4, 8.4 Hz, 1H), 2.49−2.33 (m, 6H), 2.31 (s, 3H), 1.78−1.52 (m, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 171.8, 156.9, 149.4, 148.9 (d, J_{C-F} = 297 Hz), 140.2, 136.8, 130.2, 129.4, 128.9, 128.7, 128.4, 128.0, 127.3, 126.5, 122.4, 121.5 (d, J_{C-F} = 5.1 Hz), 56.1, 54.4, 45.9, 44.7, 43.6, 39.1, 35.4, 31.6. mp 109−110 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{32}H_{37}FN_4O_5S$, 608.26; found, 609.3. Purity: >99%. $[\alpha]_D^{22}$ 10° (c 0.5, $CHCl_3$).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)isonicotinamide (**3a**).* Synthesized by following procedure E with 0.11 g of compound **15**, 0.080 g compound **X**, 0.095 g TBTU, 0.045 g HOBt, and 0.155 mL DIEA. Column chromatography: pure EtOAc. Yield: 0.14 g (81%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.97 (d, J = 7.9 Hz, 1H), 8.71 (d, J = 4.4 Hz, 2H), 8.35 (d, J = 7.8 Hz, 1H), 7.70 (d, J = 4.4 Hz, 2H), 7.44−7.08 (m, 15H), 5.81 (dd, J_{H-F} = 34.1 Hz, J_{H-H} = 8.8 Hz, 1H), 4.75 (s, 2H), 4.70−4.50 (m, 2H), 3.19−2.87 (m, 2H), 2.49−2.39 (m, 3H), 1.86−1.56 (m, 2H). ^{13}C NMR (75 MHz, DMSO): δ 170.4, 164.8, 151.4 (d, J_{C-F} = 298 Hz), 150.2, 140.9, 140.9, 137.9, 131.0, 129.1, 128.8, 128.6, 128.30, 128.29, 128.1, 127.5, 126.4, 125.9, 121.4, 120.4 (d, J_{C-F} = 3.8 Hz), 57.6, 55.0, 43.7 (d, J_{C-F} = 2.0 Hz), 37.0, 35.0, 30.9. mp 169−171 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{32}FN_3O_4S$, 586.22; found, 586.4. Purity: 96%. $[\alpha]_D^{22}$ 9° (c 0.5, MeOH).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]-dioxine-6-carboxamide (**3b**).* Synthesized by following procedure D with 0.100 g of compound **29**, 0.036 g 1,4-benzodioxane-6-carboxylic acid, 0.064 g TBTU, and 0.104 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1 to pure EtOAc. Yield: 0.096 g (75%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.44 (d, J = 8.1 Hz, 1H), 8.24 (d, J = 8.0 Hz, 1H), 7.45−7.09 (m, 17H), 6.89 (d, J = 8.3 Hz, 1H), 5.82 (dd, J_{H-F} = 34.2 Hz, J_{H-H} = 9.0 Hz, 1H), 4.74 (s, 2H), 4.65−4.47 (m, 2H), 4.34−4.16 (m, 4H), 3.11−2.91 (m, 2H), 2.48−2.35 (m, 2H), 1.87−1.54 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.0, 165.4, 151.3 (d, J_{C-F} = 297 Hz), 146.1, 142.8, 140.9, 138.2, 131.0, 129.1, 128.7, 128.6, 128.3, 128.1, 127.5, 127.0, 126.2, 125.9, 120.9, 120.5 (d, J_{C-F} = 3.9 Hz), 116.6, 116.5, 64.3, 64.0, 57.6, 55.0, 43.6 (d, J_{C-F} = 2.4 Hz), 36.9, 35.0, 30.9. mp 204−206 °C (decomposition). LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{36}H_{35}FN_2O_6S$, 643.23; found, 643.3. Purity: >99%. $[\alpha]_D^{22}$ 7° (c 0.5, $CHCl_3$/MeOH 1:1).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)-3,5-difluorobenzamide (**3c**).* Synthesized by following procedure D with 0.120 g of compound **29**, 0.037 g 3,5-difluorobenzoic acid, 0.075 g TBTU, and 0.12 mL DIEA. Column chromatography: petroleum ether/EtOAc 3:1. Yield: 0.094 g (65%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.86 (d, J = 8.2 Hz, 1H), 8.35 (d, J = 7.9 Hz, 1H), 7.56−7.41 (m, 3H), 7.41−7.10 (m, 16H), 5.81 (dd, J_{H-F} = 34.2 Hz, J_{H-H} = 9.1 Hz, 1H), 4.75 (s, 2H), 4.69−4.51 (m, 2H), 3.09 (dd, J = 13.7, 5.1 Hz, 1H), 2.97 (dd, J = 13.7, 10.1 Hz, 1H), 2.49−2.40 (m, 2H), 1.86−1.57 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 170.48, 163.7 (t, J_{C-F} = 2.8 Hz), 162.2 (dd, J_{C-F} = 251, 12 Hz), 151.4 (d, J_{C-F} = 298 Hz), 140.9, 137.9, 137.4 (t, J_{C-F} = 8.5 Hz), 137.3, 131.0, 129.1, 128.8, 128.6, 128.3, 128.1, 127.5, 126.4, 126.0, 120.4 (d, J_{C-F} = 4.1 Hz), 110.8 (dd, J_{C-F} = 27, 8.7 Hz), 106.9 (t, J_{C-F} = 26 Hz), 57.6, 55.2, 43.7 (d, J_{C-F} = 2.3 Hz), 37.0, 35.0, 30.9. mp 194−196 °C (decomposition). LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{31}F_3N_2O_4S$, 621.2; found, 621.3. Purity: >99%. $[\alpha]_D^{22}$ 6° (c 0.5, $CHCl_3$).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)isonicotinamide (**3d**).* Synthesized by following procedure D with 0.11 g of compound **35**, 0.026 g isonicotinic acid, 0.067 g TBTU, and 0.11 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:3. Yield: 0.091 g (73%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.69 (d, J = 6.0 Hz, 2H), 7.61−7.53 (m, 2H), 7.43−7.28 (m, 6H), 7.25−7.14 (m, 3H), 7.07 (d, J = 7.8 Hz, 2H), 7.01−6.88 (m, 4H), 6.19 (d, J = 7.9 Hz, 1H), 5.55 (dd, J_{H-F} = 32.6 Hz, J_{H-H} = 9.2 Hz, 1H), 4.76−4.59 (m, 2H), 4.36 (s, 2H), 3.14 (dd, J = 13.5, 6.0 Hz, 1H), 2.93 (dd, J = 13.4, 8.6 Hz, 1H), 2.44−2.19 (m, 5H), 1.75−1.50 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.2, 165.1, 152.5 (d, J_{C-F} = 301 Hz), 141.3, 140.1, 137.3, 132.8, 131.0, 129.8, 129.5, 129.3, 129.2, 128.7, 128.3, 126.7, 126.5, 121.3, 119.4 (d, J_{C-F} = 4.3 Hz), 58.7, 55.6, 44.7 (d, J_{C-F} = 2.2 Hz), 38.4, 35.6 (d, J_{C-F} = 1.6 Hz), 31.6, 21.2. mp 181−183 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{34}FN_3O_4S$, 600.24; found, 600.4. Purity: 95%. $[\alpha]_D^{22}$ 13° (c 0.5, MeOH).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)-2,3-dihydrobenzo[b][1,4]-dioxine-6-carboxamide (**3e**).* Synthesized by following procedure D with 0.110 g of compound **35**, 0.037 g 1,4-benzodioxane-6-carboxylic acid, 0.067 g TBTU, and 0.108 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:1 to 0:1. Yield: 0.10 g (76%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.41 (d, J = 8.0 Hz, 1H), 8.23 (d, J = 8.0 Hz, 1H), 7.43−7.23 (m, 9H), 7.23−7.11 (m, 5H), 7.04 (d, J = 7.9 Hz, 2H), 6.89 (d, J = 8.4 Hz, 1H), 5.81 (dd, J_{H-F} = 34.3 Hz, J_{H-H} = 9.1 Hz, 1H), 4.74 (s, 2H), 4.63−4.47 (m, 2H), 4.32−4.18 (m, 4H), 3.05−2.87 (m, 2H), 2.49−2.39 (m, 2H), 2.22 (s, 3H), 1.86−1.57 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.0, 165.4, 151.3 (d, J_{C-F} = 297 Hz), 146.0, 142.8, 140.9, 135.14, 135.10, 131.0, 129.0, 128.8, 128.7, 128.6, 128.3, 127.5, 127.0, 125.9, 121.0, 120.5 (d, J_{C-F} = 3.7 Hz), 116.6, 116.5, 64.3, 64.0, 57.6, 55.1, 43.6 (d, J_{C-F} = 1.6 Hz), 36.6, 35.0, 30.9, 20.6. mp 193−195 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{37}H_{37}FN_2O_6S$, 657.25; found, 657.4. Purity: 97%. $[\alpha]_D^{22}$ 8° (c 0.5, $CHCl_3$/MeOH 1:1).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)-3,5-difluorobenzamide (**3f**).* Synthesized by following procedure D with 0.12 g of compound **35**, 0.036 g 3,5-difluorobenzoic acid, 0.073 g TBTU, and 0.12 mL DIEA. Column chromatography: petroleum ether/EtOAc 3:1 to pure EtOAc. Yield: 0.12 g (81%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.82 (d, J = 8.1 Hz, 1H), 8.32 (d, J = 8.0 Hz, 1H), 7.61−7.40 (m, 3H), 7.40−7.23 (m, 8H), 7.23−7.10 (m, 5H), 7.05 (d, J = 7.9

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Hz, 2H), 5.79 (dd, J_{H-F} = 33.9 Hz, J_{H-H} = 9.3 Hz, 1H), 4.75 (s, 2H), 4.65−4.48 (m, 2H), 3.02 (dd, J = 13.5, 5.1 Hz, 1H), 2.92 (dd, J = 13.6, 10.0 Hz, 1H), 2.48−2.39 (m, 1H), 2.22 (s, 3H), 1.83−1.56 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 170.5, 163.86, 163.7 (t, J_{C-F} = 2.6 Hz), 162.1 (dd, J_{C-F} = 251, 12 Hz), 151.3 (d, J_{C-F} = 298 Hz), 140.9, 137.4 (t, J_{C-F} = 8.4 Hz), 135.3, 134.8, 131.0, 129.0, 128.7, 128.6, 128.3, 127.5, 126.0, 120.4 (d, J_{C-F} = 4.0 Hz), 110.8 (dd, J_{C-F} = 26, 8.4 Hz), 106.9 (t, J_{C-F} = 25 Hz), 57.6, 55.3, 43.7 (d, J_{C-F} = 1.5 Hz), 36.6, 35.0, 30.9, 20.6. mp 196−198 °C (decomposition). LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{35}H_{33}F_3N_2O_4S$, 635.22; found, 635.3. Purity: 95%. $[\alpha]_D^{22}$ 9° (c 0.5, MeOH/$CHCl_3$ 1:1).

*Phenyl (S,E)-1-Fluoro-3-((S)-2-(isonicotinamido)-3-(m-tolyl)-propanamido)-5-phenylpent-1-ene-1-sulfonate (**3g**).* Synthesized by following procedure D with 0.11 g of compound **32**, 0.026 g isonicotinic acid, 0.067 g TBTU, and 0.11 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:3. Yield: 0.086 g (69%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.68 (d, J = 4.7 Hz, 2H), 7.56 (d, J = 5.7 Hz, 2H), 7.40 (d, J = 7.6 Hz, 1H), 7.37−7.28 (m, 5H), 7.25−7.11 (m, 4H), 7.04 (d, J = 7.6 Hz, 1H), 6.99−6.90 (m, 3H), 6.84 (d, J = 7.5 Hz, 1H), 6.24 (d, J = 7.8 Hz, 1H), 5.57 (dd, J_{H-F} = 32.5 Hz, J_{H-H} = 9.1 Hz, 1H), 4.78−4.58 (m, 2H), 4.36 (s, 2H), 3.15 (dd, J = 13.4, 6.2 Hz, 1H), 2.93 (dd, J = 13.4, 8.5 Hz, 1H), 2.43−2.26 (m, 2H), 2.24 (s, 3H), 1.74−1.48 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.2, 165.2, 152.5 (d, J_{C-F} = 301 Hz), 150.3, 141.2, 140.1, 138.8, 136.0, 131.0, 130.0, 129.5, 129.2, 129.0, 128.7, 128.4, 128.3, 126.7, 126.5, 126.5, 121.2, 119.4 (d, J_{C-F} = 4.1 Hz), 58.7, 55.5, 44.8 (d, J_{C-F} = 2.3 Hz), 38.7, 35.6, 35.7, 31.6, 21.4. mp 81−82 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{34}FN_3O_4S$, 600.24; found, 600.4. Purity: 95%. $[\alpha]_D^{22}$ 7° (c 0.5, MeOH).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-yl)-2,3-dihydrobenzo[b][1,4]-dioxine-6-carboxamide (**3h**).* Synthesized by following procedure D with 0.11 g of compound **32**, 0.037 g 1,4-benzodioxane-6-carboxylic acid, 0.067 g TBTU and 0.11 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1 to 0:1. Yield: 0.094 g (69%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.43 (d, J = 8.1 Hz, 1H), 8.25 (d, J = 8.0 Hz, 1H), 7.46−7.23 (m, 9H), 7.23−7.03 (m, 6H), 6.97 (d, J = 7.2 Hz, 1H), 6.90 (d, J = 8.3 Hz, 1H), 5.83 (dd, J_{H-F} = 34.2 Hz, J_{H-H} = 9.0 Hz, 1H), 4.74 (s, 2H), 4.64−4.49 (m, 2H), 4.39−4.17 (m, 4H), 3.08−2.86 (m, 2H), 2.49−2.38 (m, 2H), 2.23 (s, 3H), 1.87−1.55 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.0, 165.5, 151.3 (d, J_{C-F} = 298 Hz), 146.1, 142.8, 140.9, 138.2, 137.0, 131.0, 129.8, 128.7, 128.6, 128.3, 127.9, 127.5, 127.1, 126.9, 126.2, 125.9, 120.9, 120.5 (d, J_{C-F} = 4.2 Hz), 116.6, 116.5, 64.3, 64.0, 57.6, 55.0, 43.6 (d, J_{C-F} = 0.9 Hz), 36.9, 35.1, 30.9, 21.0. mp 188−190 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{37}H_{37}FN_2O_6S$, 657.25; found, 657.4. Purity >99%. $[\alpha]_D^{22}$ 11° (c 0.5, $CHCl_3$).

*N-((S)-1-(((S,E)-1-(Benzylsulfonyl)-1-fluoro-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-yl)-3,5-difluorobenzamide (**3i**).* Synthesized by following procedure D with 0.12 g of compound **32**, 0.036 g 3,5-difluorobenzoic acid, 0.073 g TBTU, and 0.12 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1. Yield: 0.11 g (77%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.83 (d, J = 8.1 Hz, 1H), 8.33 (d, J = 7.8 Hz, 1H), 7.61−7.02 (m, 17H), 6.98 (d, J = 7.1 Hz, 1H), 5.81 (dd, J_{H-F} = 34.1 Hz, J_{H-H} = 9.0 Hz, 1H), 4.74 (s, 2H), 4.67−4.46 (m, 2H), 3.04 (dd, J = 13.5, 4.7 Hz, 1H), 2.93 (dd, J = 13.5, 4.7 Hz, 1H), 2.48−2.39 (m, 2H), 2.23 (s, 3H), 1.83−1.59 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 170.5, 163.8 (t, J_{C-F} = 2.7 Hz), 162.1 (dd, J_{C-F} = 247 Hz, 13 Hz), 151.3 (d, J_{C-F} = 298 Hz), 140.9, 137.9, 137.5 (t, J_{C-F} = 8.5 Hz), 137.1, 131.0, 129.8, 128.8, 128.6, 128.3, 128.0, 127.5, 127.0, 126.2, 126.0, 120.42 (d, J_{C-F} = 3.5 Hz), 110.8 (dd, J_{C-F} = 25 Hz, 8.6 Hz), 106.8 (t, J_{C-F} = 26 Hz), 57.6, 55.2, 43.7 (d, J_{C-F} = 2.0 Hz), 36.9, 35.0, 30.9, 21.0. mp 176−178 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{35}H_{33}F_3N_2O_4S$, 635.22; found, 635.3. Purity >99%. $[\alpha]_D^{22}$ 6° (c 0.5, $CHCl_3$).

*N-((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)isonicotinamide (**4a**).*[11] Synthesized by following procedure E with 0.14 g of compound **13**, 0.107 g compound **X**, 0.13 g TBTU, 0.062 g HOBt, and 0.21 mL DIEA. Column chromatography: DCM/MeOH 19:1. Yield: 0.17 g (75%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.71 (br, 2H), 7.95 (d, J = 7.5 Hz, 2H), 7.76−7.54 (m, 5H), 7.50 (d, J = 7.4 Hz, 1H), 7.40−7.08 (m, 8H), 6.99 (d, J = 6.4 Hz, 2H), 6.40 (d, J = 7.7 Hz, 1H), 6.03 (dd, J_{H-F} = 31.9 Hz, J_{H-H} = 8.8 Hz, 1H), 4.87−4.60 (m, 2H), 3.21 (dd, J = 13.5, 6.3 Hz, 1H), 3.08 (dd, J = 13.5, 8.2 Hz, 1H), 2.49 (t, J = 7.4 Hz, 2H), 1.96−1.68 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.3, 164.9, 155.1 (d, J_{C-F} = 301 Hz), 149.4, 142.1, 140.2, 137.1, 136.2, 134.9, 129.7, 129.4, 129.1, 128.8, 128.7, 128.3, 127.7, 126.50, 121.7, 117.0 (d, J_{C-F} = 5.3 Hz), 55.6, 45.1, 38.8, 35.7 (d, J_{C-F} = 1.9 Hz), 31.8. mp 184−187 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{32}H_{30}FN_3O_4S$, 572.20; found, 572.4. Purity: 95%. $[\alpha]_D^{22}$ 11° (c 0.5, $CHCl_3$).

*N-((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]-dioxine-6-carboxamide (**4b**).* Synthesized by following procedure D with 0.100 g of compound **27**, 0.036 g 1,4-benzodioxane-6-carboxylic acid, 0.064 g TBTU, and 0.104 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:1. Yield: 0.084 g (67%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.43 (d, J = 8.0 Hz, 1H), 8.35 (d, J = 7.8 Hz, 1H), 7.94 (dt, J = 7.3, 1.7 Hz, 2H), 7.85 (tt, J = 7.3, 1.2 Hz, 1H), 7.73 (tt, J = 6.7, 1.1 Hz, 2H), 7.43−7.06 (m, 12H), 6.89 (d, J = 8.3 Hz, 1H), 6.30 (dd, J_{H-F} = 34.0 Hz, J_{H-H} = 8.9 Hz, 1H), 4.68−4.47 (m, 2H), 4.33−4.18 (m, 4H), 3.10−2.90 (m, 2H), 2.59−2.50 (m, 2H), 1.96−1.73 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.0, 165.4, 153.0 (d, J_{C-F} = 296 Hz), 146.1, 142.8, 140.8, 138.2, 136.6, 135.2, 130.1, 129.1, 128.27, 128.26, 128.1, 128.0, 127.0, 126.3, 125.9, 120.9, 119.2 (d, J_{C-F} = 4.8 Hz), 116.6, 116.5, 64.3, 64.0, 54.9, 43.8 (d, J_{C-F} = 1.6 Hz), 36.9, 35.0 (d, J_{C-F} = 2.3 Hz), 31.0. mp 187−189 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{35}H_{33}FN_2O_6S$, 629.2; found, 629.3. Purity >99%. $[\alpha]_D^{22}$ 3° (c 0.5, $CHCl_3$).

*3,5-Difluoro-N-((S)-1-(((S,E)-1-fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)benzamide (**4c**).* Synthesized by following procedure D with 0.124 g of compound **27**, 0.039 g 3,5-difluorobenzoic acid, 0.079 g TBTU, and 0.125 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1. Yield: 0.12 g (81%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.99−7.90 (m, 2H), 7.73−7.63 (m, 1H), 7.57 (t, J = 7.6 Hz, 2H), 7.37−7.09 (m, 12H), 7.00−6.91 (m, 2H), 6.87 (tt, J = 8.5, 2.3 Hz, 1H), 6.37 (d, J = 7.7 Hz, 1H), 5.96 (dd, J_{H-F} = 31.8 Hz, J_{H-H} = 8.9 Hz, 1H), 4.82−4.61 (m, 2H), 3.14 (dd, J = 13.5, 6.4 Hz, 1H), 3.03 (dd, J = 13.5, 8.2 Hz, 1H), 2.45 (t, J = 7.7 Hz, 2H), 1.88−1.72 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.6, 165.0 (t, J_{C-F} = 2.8 Hz), 163.0 (dd, J_{C-F} = 251, 12 Hz), 155.2 (d, J_{C-F} = 301 Hz), 140.2, 137.1, 136.8 (t, J_{C-F} = 8.3 Hz), 136.2, 134.8, 129.7, 129.4, 129.1, 128.8, 128.7, 128.3, 127.6, 126.5, 116.9 (d, J_{C-F} = 5.1 Hz), 110.5 (dd, J_{C-F} = 27, 8.7 Hz), 107.4 (t, J_{C-F} = 25 Hz), 55.7, 45.1 (d, J_{C-F} = 1.5 Hz), 38.7, 35.7, 31.8. mp 90−92 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{29}F_3N_2O_4S$, 607.2; found, 607.3. Purity: 95%. $[\alpha]_D^{22}$ 5° (c 0.5, MeOH).

*N-((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)isonicotinamide (**4d**).* Synthesized by following procedure D with 0.10 g of compound **33**, 0.024 g isonicotinic acid, 0.063 g TBTU, and 0.10 mL DIEA. Column chromatography: pure EtOAc. Yield: 0.094 g (83%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.69 (d, J = 3.8 Hz, 2H), 7.95 (d, J = 7.7 Hz, 2H), 7.70 (t, J = 7.3 Hz, 1H), 7.65−7.46 (m, 4H), 7.30−7.23 (m, 1H), 7.23−7.01 (m, 7H), 6.96 (d, J = 6.8 Hz, 2H), 6.32 (d, J = 7.8 Hz, 1H), 6.01 (dd, J_{H-F} = 31.9 Hz, J_{H-H} = 8.8 Hz, 1H), 4.84−4.60 (m, 2H), 3.15 (dd, J = 13.6, 5.7 Hz, 1H), 3.01 (dd, J = 13.4, 8.2 Hz, 1H), 2.46 (t, J = 7.6 Hz, 2H), 2.32 (s, 3H), 1.90−1.73 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.3, 165.4, 155.12 (d, J_{C-F} = 301 Hz), 150.7, 140.8, 140.1, 137.4, 137.1, 134.8, 132.9, 129.8, 129.7, 129.3, 128.8, 128.7, 128.3, 126.5, 121.0, 116.9 (d, J_{C-F} = 4.8 Hz), 55.4, 45.0 (d, J_{C-F} = 1.7 Hz), 38.4, 35.8, 31.8, 21.2. mp 84−86 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{32}FN_3O_4S$, 586.22; found, 586.4. Purity: 96%. $[\alpha]_D^{22}$ 12° (c 0.5, MeOH).

*N-((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (**4e**).* Synthesized by following procedure D with 0.065 g of compound **33**, 0.023 g 1,4-benzodioxane-6-carboxylic acid, 0.041 g TBTU, and 0.066 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1 to 1:1. Yield: 0.057 g (70%), colorless

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.00−7.90 (m, 2H), 7.75−7.65 (m, 1H), 7.64−7.53 (m, 2H), 7.26 (s, 1H), 7.23−7.03 (m, 8H), 6.93 (dd, J = 7.8, 1.5 Hz, 2H), 6.85 (d, J = 8.4 Hz, 1H), 6.79 (d, J = 7.7 Hz, 1H), 6.70 (d, J = 7.9 Hz, 1H), 6.02 (dd, J_{H-F} = 32.0 Hz, J_{H-H} = 8.7 Hz, 1H), 4.83 (q, J = 7.5 Hz, 1H), 4.68 (quint, J = 7.7 Hz, 1H), 4.34−4.21 (m, 4H), 3.13 (dd, J = 13.6, 6.3 Hz, 1H), 3.02 (dd, J = 13.6, 7.8 Hz, 1H), 2.54−2.38 (m, 2H), 2.31 (s, 3H), 1.87−1.69 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.8, 166.8, 154.8 (d, J_{C-F} = 300 Hz), 147.0, 143.6, 140.3, 137.3, 137.0, 134.7, 133.3, 129.7, 129.6, 129.3, 128.8, 128.6, 128.4, 126.9, 126.3, 120.6, 117.62 (d, J_{C-F} = 5.1 Hz), 117.5, 116.8, 64.7, 64.3, 55.1, 44.9 (d, J_{C-F} = 1.7 Hz), 38.3, 35.7, 31.8, 21.2. mp 95−97 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{36}H_{35}FN_2O_6S$, 643.23; found, 643.3. Purity: 96%. $[\alpha]_D^{22}$ 13° (c 1, $CHCl_3$).

*3,5-Difluoro-N-((S)-1-(((S,E)-1-fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(p-tolyl)propan-2-yl)benzamide (**4f**).* Synthesized by following procedure D with 0.065 g of compound **33**, 0.020 g 3,5-difluorobenzoic acid, 0.041 g TBTU, and 0.066 mL DIEA. Column chromatography: petroleum ether/EtOAc 3:1. Yield: 0.059 g (73%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.00−7.91 (m, 2H), 7.74−7.63 (m, 1H), 7.62−7.53 (m, 2H), 7.31 (d, J = 7.6 Hz, 1H), 7.25−7.09 (m, 7H), 7.05 (d, J = 8.0 Hz, 2H), 6.97−6.82 (m, 3H), 6.39 (d, J = 7.7 Hz, 1H), 5.98 (dd, J_{H-F} = 31.9 Hz, J_{H-H} = 8.8 Hz, 1H), 4.83−4.60 (m, 2H), 3.11 (dd, J = 13.5, 6.3 Hz, 1H), 2.99 (dd, J = 13.5, 8.2 Hz, 1H), 2.52−2.40 (m, 2H), 2.31 (s, 3H), 1.90−1.74 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.7, 164.95 (t, J_{C-F} = 2.9 Hz), 163.02 (dd, J_{C-F} = 251, 12 Hz), 155.12 (d, J_{C-F} = 301 Hz), 140.2, 137.3, 137.2, 136.9 (t, J_{C-F} = 8.3 Hz), 134.8, 133.0, 129.8, 129.7, 129.3, 128.8, 128.7, 128.3, 126.5, 117.05 (d, J_{C-F} = 4.8 Hz), 110.50 (dd, J_{C-F} = 27, 9.4 Hz), 107.37 (t, J_{C-F} = 25 Hz), 55.6, 45.1 (d, J_{C-F} = 1.6 Hz), 38.3, 35.7 (d, J_{C-F} = 1.5 Hz), 31.8, 21.2. mp 174−176 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{31}F_3N_2O_4S$, 621.21; found, 621.3. Purity: 95%. $[\alpha]_D^{22}$ 9° (c 0.5, MeOH).

*N-((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-yl)isonicotinamide (**4g**).* Synthesized by following procedure D with 0.10 g of compound **30**, 0.024 g isonicotinic acid, 0.063 g TBTU, and 0.10 mL DIEA. Column chromatography: pure EtOAc. Yield: 0.071 g (63%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.69 (d, J = 5.2 Hz, 2H), 7.94 (d, J = 7.5 Hz, 2H), 7.74−7.65 (m, 1H), 7.63−7.48 (m, 4H), 7.34−7.11 (m, 6H), 7.11−6.88 (m, 5H), 6.31 (d, J = 7.8 Hz, 1H), 6.01 (dd, J_{H-F} = 31.9 Hz, J_{H-H} = 8.8 Hz, 1H), 4.84−4.59 (m, 2H), 3.16 (dd, J = 13.5, 6.2 Hz, 1H), 3.00 (dd, J = 13.5, 8.4 Hz, 1H), 2.45 (t, J = 7.5 Hz, 2H), 2.29 (s, 3H), 1.89−1.73 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.3, 165.4, 155.10 (d, J_{C-F} = 301 Hz), 150.6, 140.9, 140.1, 138.8, 137.0, 136.0, 134.7, 130.1, 129.7, 129.0, 128.8, 128.7, 128.4, 128.3, 126.5, 126.4, 121.1, 116.9 (d, J_{C-F} = 5.2 Hz), 55.4, 45.1 (d, J_{C-F} = 1.9 Hz), 38.7, 35.8 (d, J_{C-F} = 1.4 Hz), 31.8, 21.5. mp 182−184 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{32}FN_3O_4S$, 586.22; found, 586.4. Purity: 95%. $[\alpha]_D^{22}$ 3° (c 0.5, $CHCl_3$).

*N-((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (**4h**).* Synthesized by following procedure D with 0.10 g of compound **30**, 0.035 g 1,4-benzodioxane-6-carboxylic acid, 0.063 g TBTU, and 0.10 mL DIEA. Column chromatography: petroleum ether/EtOAc 1:1. Yield: 0.10 g (84%), colorless solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.44 (d, J = 8.1 Hz, 1H), 8.38 (d, J = 7.8 Hz, 1H), 7.94 (dt, J = 7.3, 1.5 Hz, 2H), 7.89−7.80 (m, 1H), 7.80−7.68 (m, 2H), 7.42−7.30 (m, 2H), 7.29−7.02 (m, 8H), 6.98 (d, J = 6.9 Hz, 1H), 6.90 (d, J = 8.3 Hz, 1H), 6.31 (dd, J_{H-F} = 34.0 Hz, J_{H-H} = 8.9 Hz, 1H), 4.65−4.46 (m, 2H), 4.27 (d, J = 5.4 Hz, 4H), 3.08−2.84 (m, 2H), 2.61−2.50 (m, 2H), 2.25 (s, 3H), 1.96−1.72 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.1, 165.5, 152.9 (d, J_{C-F} = 295 Hz), 146.1, 142.8, 140.8, 138.1, 137.0, 136.6, 135.2, 130.1, 129.8, 128.3, 128.3, 128.2, 127.9, 127.0, 126.9, 126.2, 125.9, 121.0, 119.3 (d, J_{C-F} = 4.3 Hz), 116.6, 116.5, 64.3, 64.0, 55.0, 43.8 (d, J_{C-F} = 1.5 Hz), 36.8, 35.4, 31.0, 21.0. mp 186−188 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{36}H_{35}FN_2O_6S$, 643.23; found, 643.3. Purity: >99%. $[\alpha]_D^{22}$ 11° (c 1, $CHCl_3$).

*3,5-Difluoro-N-((S)-1-(((S,E)-1-fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-(m-tolyl)propan-2-yl)benzamide (**4i**).* Synthesized by following procedure D with 0.10 g of compound **30**, 0.031 g 3,5-difluorobenzoic acid, 0.063 g TBTU, and 0.10 mL DIEA. Column chromatography: petroleum ether/EtOAc 2:1. Yield: 0.075 g (62%), colorless solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.95 (d, J = 7.5 Hz, 2H), 7.72−7.64 (m, 1H), 7.57 (t, J = 7.6 Hz, 2H), 7.29−7.23 (m, 1H), 7.25−6.81 (m, 13H), 6.38 (d, J = 7.7 Hz, 1H), 6.01 (dd, J_{H-F} = 31.8 Hz, J_{H-H} = 8.8 Hz, 1H), 4.81−4.60 (m, 2H), 3.12 (dd, J = 13.5, 6.5 Hz, 1H), 2.98 (dd, J = 13.5, 8.1 Hz, 1H), 2.45 (t, J = 7.7 Hz, 2H), 2.28 (s, 3H), 1.88−1.74 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.6, 165.02 (t, J_{C-F} = 2.4 Hz), 162.99 (dd, J_{C-F} = 251, 12 Hz), 155.14 (d, J_{C-F} = 301 Hz), 140.2, 138.8, 137.1, 136.8 (t, J_{C-F} = 8.4 Hz), 136.2, 134.8, 130.1, 129.7, 129.0, 128.8, 128.7, 128.4, 128.3, 126.5, 117.0 (d, J_{C-F} = 4.9 Hz), 110.5 (dd, J_{C-F} = 26, 8.7 Hz), 107.4 (t, J_{C-F} = 25 Hz), 55.6, 45.2 (d, J_{C-F} = 1.6 Hz), 38.5, 35.7, 31.8, 21.4. mp 176−178 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{34}H_{31}F3N_2O_4S$, 621.21; found, 621.3. Purity >99%. $[\alpha]_D^{22}$ 5° (c 0.5, MeOH/$CHCl_3$ 1:1).

*N-((S)-1-(((S,E)-1-Fluoro-5-(methylthio)-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)-4-methylpiperazine-1-carboxamide (**4j**).* Synthesized by following procedure E with 0.11 g of compound **16**, 0.089 g compound **IX**, 0.088 g TBTU, 0.042 g HOBt, and 0.14 mL DIEA. Column chromatography: DCM/MeOH 19:1. Yield: 0.11 g (73%), colorless solid. (300 MHz, CD_2Cl_2): δ 7.85 (dd, J = 7.3, 1.7 Hz, 2H), 7.71−7.59 (m, 1H), 7.59−7.47 (m, 2H), 7.25−7.13 (m, 3H), 7.12−7.01 (m, 3H), 5.99 (dd, J_{H-F} = 32.4 Hz, J_{H-H} = 8.8 Hz, 1H), 5.06 (d, J = 7.7 Hz, 1H), 4.76−4.59 (m, 1H), 4.41 (q, J = 7.2 Hz, 1H), 3.21 (q, J = 4.5 Hz, 4H), 3.02−2.79 (m, 2H), 2.31−2.25 (m, 2H), 2.22 (t, J = 5.1 Hz, 4H), 2.16 (s, 3H), 1.90 (s, 3H), 1.82−1.60 (m, 2H). (75 MHz, CD2Cl2): δ 172.3, 157.5, 155.0 (d, J_{C-F} = 298 Hz), 137.6, 137.2, 135.2, 130.1, 129.9, 129.1, 127.4, 118.2 (d, J_{C-F} = 4.9 Hz), 56.3, 55.1, 54.0, 46.4, 44.7 (d, J_{C-F} = 1.8 Hz), 44.3, 39.0, 33.8, 30.4, 15.7. mp 110−112 °C. HR-MS (ESI, m/z): $[M + Na]^+$ calcd for $C_{27}H_{35}FN_4O_4S_2$, 585.1981; found, 585.1973. $[\alpha]_D^{22}$ 12° (c 1, $CHCl_3$).

*N-((S)-1-(((S,E)-1-Fluoro-5-(methylthio)-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)isonicotinamide (**4k**).* Synthesized by following procedure E with 0.11 g of compound **16**, 0.074 g compound **X**, 0.088 g TBTU, 0.042 g HOBt, and 0.14 mL DIEA. Column chromatography: DCM/MeOH 95:5. Yield: 0.098 g (66%), colorless solid. (300 MHz, CD_2Cl_2): δ 8.70 (d, J = 4.5 Hz, 2H), 7.94 (d, J = 7.5 Hz, 2H), 7.76−7.53 (m, 5H), 7.44−7.15 (m, 9H), 7.07 (d, J = 7.3 Hz, 1H), 6.26 (d, J = 7.8 Hz, 1H), 6.04 (dd, J_{H-F} = 32.2 Hz, J_{H-H} = 8.8 Hz, 1H), 4.92−4.55 (m, 2H), 3.23−3.00 (m, 2H), 2.44−2.24 (m, 2H), 2.05−1.91 (m, 3H), 1.88−1.74 (m, 2H). (75 MHz, CD2Cl2): δ 170.0, 165.2, 151.4, 150.6, 140.7, 139.4, 137.0, 136.2, 134.8, 129.6, 129.3, 128.9, 128.6, 127.3, 120.8, 55.1, 45.3, 38.3, 29.8, 15.2. mp 64−66 °C. HR-MS (ESI, m/z): $[M + Na]^+$ calcd for $C_{27}H_{28}FN_3O_4S_2$, 564.1403; found, 564.1411. $[\alpha]_D^{22}$ 21° (c 1, $CHCl_3$).

*N-((S)-1-(((R,E)-1-(Benzylthio)-4-fluoro-4-(phenylsulfonyl)but-3-en-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)-4-methylpiperazine-1-carboxamide (**4l**).* Synthesized by following procedure E with 0.11 g of compound **17**, 0.077 g compound **IX**, 0.076 g TBTU, 0.036 g HOBt, and 0.13 mL DIEA. Column chromatography: DCM/MeOH 9:1. Yield: 0.093 g (63%). (300 MHz, $CDCl_3$): δ 8.02−7.91 (m, 2H), 7.79−7.67 (m, 1H), 7.67−7.54 (m, 2H), 7.43−7.09 (m, 10H), 6.81 (dd, J = 47.2, 7.8 Hz, 1H), 6.16 (dd, J = 43.5, 31.8, 8.6 Hz, 1H), 5.15 (dd, J = 14.7, 7.4 Hz, 1H), 4.93−4.75 (m, 1H), 4.52 (m, 1H), 3.64 (s, 2H), 3.54−3.30 (m, 4H), 3.18−2.99 (m, 2H), 2.67−2.47 (m, 2H), 2.50−2.38 (m, 4H), 2.36 (s, 3H). (75 MHz, $CDCl_3$): δ 171.7, 156.9, 137.4, 137.4, 137.0, 136.93, 136.90, 136.87, 134.8, 129.7, 129.5, 129.4, 129.0, 128.9, 128.85, 128.76, 127.5, 127.3, 127.2, 116.6, 116.57, 116.5, 77.2, 55.9, 5588, 54.4, 54.4, 45.9, 45.8, 44.5, 44.3, 43.55, 38.8, 38.6, 36.5, 35.1, 35.0. mp 84−86 °C. HR-MS (ESI, m/z): $[M + Na]^+$ calcd for $C_{32}H_{37}FN_4O_4S_2$, 647.2138; found, 647.2152. $[\alpha]_D^{22}$ −13° (c 1, $CHCl_3$).

*Synthesis of the Metabolites. tert-Butyl 4-(((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamoyl)piperazine-1-carboxylate (**36**).* Compound **XIII** (0.31 g) was dissolved in DCM and cooled to 0 °C. 0.11 g HOBt, 0.26 g TBTU, and 0.42 mL DIEA were added and the mixture was stirred for 30 min. 0.24 g of compound **13** was added and the mixture was stirred for an additional 24 h at rt. The reaction was stopped with 7 equiv of water and the mixture was washed with a saturated solution of

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

$NaHCO_3$ and brine, dried with Na_2SO_4, and the solvent was removed under reduced pressure. The crude product was purified by column chromatography (cyclohexane/EtOAc 1:1), resulting in a colorless oil (0.21 g, 45%). ^{1}H NMR (300 MHz, CDCl3): δ 7.97−7.88 (m, 2H), 7.74−7.65 (m, 1H), 7.63−7.53 (m, 2H, *H*-27), 7.35−7.11 (m, 8H), 7.02 (ddd, J = 7.8, 3.2, 1.5 Hz, 2H), 6.54 (d, J = 41.8 Hz, 1H), 5.99 (dd, J = 32.0, 8.7 Hz, 1H), 5.11 (d, J = 27.5 Hz, 1H), 4.71−4.56 (m, 1H), 4.47 (d, J_{C-F} = 10.3 Hz), 3.45−3.17 (m, 8H), 3.14−2.91 (m, 2H), 2.56−2.36 (m, 2H), 1.95−1.62 (m, 2H), 1.46 (s, 9H). ^{13}C NMR (75 MHz, CDCl3): δ 171.6, 157.0, 154.7, 140.4, 137.1 (d, J = 7.3 Hz), 136.8, 134.8, 129.7, 129.4, 128.9, 128.8, 128.7, 128.4, 127.3 (d, J = 5.1 Hz), 126.4, 117.4, 80.4, 56.1, 45.1, 43.7, 38.5, 35.8, 31.8, 28.5.

*(S)-4-((1-Ethoxy-1-oxo-3-phenylpropan-2-yl)carbamoyl)-1-methylpiperazine 1-oxide (**28**).* Compound **IX** (1.3 g) was dissolved in DCM and cooled to 0 °C. 0.7 g of 3-chloroperbenzoic acid was added and the mixture was stirred for 16 h. Triphenylphosphine was added to stop the reaction. The solvent was removed under reduced pressure, giving the crude product as a colorless oil (0.82 g, 63%). ^{1}H NMR (300 MHz, methanol-d_4): δ 7.35−7.11 (m, 5H), 4.87 (t, J = 2.0 Hz, 4H), 4.49 (td, J = 5.9, 2.9 Hz, 1H), 4.22−4.03 (m, 2H), 3.47−3.40 (m, 1H), 3.23−2.87 (m, 4H), 2.39 (d, J = 2.9 Hz, 3H), 2.04−1.89 (m, 2H), 1.26−1.10 (m, 3H). ^{13}C NMR (75 MHz, methanol-d_4): δ 174.4, 159.3, 138.8, 130.2, 129.4, 127.8, 62.2, 57.0, 55.2, 45.6, 44.1, 38.5, 14.4.

*(S)-4-((1-Carboxy-2-phenylethyl)carbamoyl)-1-methylpiperazine 1-oxide (**38**).* Compound **37** (0.65 g) was dissolved in THF and cooled to 0 °C, and 0.33 g of LiOH in water was added dropwise. The mixture was stirred for 3 h at room temperature. The product was isolated from the aqueous phase giving a colorless solid (0.31 g, 53%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 11.49 (s, 1H), 7.39−6.82 (m, 5H), 4.15−3.96 (m, 1H), 3.96−3.67 (m, 1H), 3.16−2.68 (m, 8H), 2.50 (s, 2H), 2.33 (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 173.9, 157.1, 138.6, 129.3, 128.2, 126.4, 55.8, 51.9, 42.0, 40.8, 36.4.

*Ethyl Isonicotate (**39**).* Isonicotinic acid (2.46 g) was dissolved in 40 mL ethanol and 1 mL of concentrated H_2SO_4 was added dropwise. The mixture was refluxed for 24 h and the solvent was removed under reduced pressure. The residue was extracted with DCM and the combined extracts were washed with a saturated solution of $NaHCO_3$, resulting in a colorless oil (2.2 g, 73%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.83−8.73 (m, 2H), 7.88−7.73 (m, 2H), 4.34 (q, J = 7.1 Hz, 2H), 1.32 (t, J = 7.1 Hz, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 164.8, 150.8, 137.2, 122.6, 61.7, 14.1.

*4-(Ethoxycarbonyl)pyridine 1-Oxide (**40**).* Compound **39** (1.51 g) was dissolved in DCM and cooled to 0 °C. 1.73 g of 3-chloroperbenzoic acid was added and the mixture was stirred for 16 h. Triphenylphosphine was added to stop the reaction. The solvent was removed under reduced pressure, giving the crude product that was further purified by column chromatography (cyclohexane/EtOAc 1:6), resulting in a colorless oil (0.24 g, 29%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.43−8.19 (m, 2H), 7.97−7.77 (m, 2H), 4.31 (q, J = 7.1 Hz, 2H), 1.31 (t, J = 7.1 Hz, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 163.7, 139.9, 126.85, 125.7, 61.9, 14.5.

*4-Carboxypyridine 1-Oxide (**41**).* Compound **40** (0.20 g) was dissolved in THF and cooled to 0 °C. 0.20 g of LiOH was dissolved in water and added dropwise to the mixture, then stirred for 3 h at rt. The product was isolated from the aqueous phase giving a colorless solid (0.17 g, 100%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.31 (d, J = 6.2 Hz, 2H), 7.82 (d, J = 6.2 Hz, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 165.00, 139.8, 127.5, 127.1.

*4-(((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamoyl)piperazin-1-ium Chloride (**1a**).* A 4 M solution of HCl in dioxane (3 mL) was added to 0.208 g of compound **36**. The mixture was stirred for 12 h and the solvent was removed under reduced pressure. The residue was washed three times with diethyl ether in an ultrasonic bath and then lyophilized, resulting in a colorless solid (0.177 g, 93%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 8.53 (s, 2H), 8.10−7.97 (m, 2H), 7.98−7.66 (m, 3H), 7.40−6.89 (m, 11H), 6.51 (dd, J = 32.7, 9.8 Hz, 1H), 4.58−4.40 (m, 1H), 4.33−3.93 (m, 1H), 3.61−3.44 (m, 8H), 3.09−2.83 (m, 4H), 2.34−1.78 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 163.2, 157.2, 139.3, 136.6, 135.8, 132.6, 130.2, 129.1, 128.1, 126.1, 119.7, 117.8, 115.1, 57.2, 52.4, 46.0, 42.4, 42.4, 36.5, 35.2, 35.0, 33.3. mp 93−94 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{31}H_{35}FN_4O_4S$, 579.24; found, 579.3. Purity: 98%. $[\alpha]_D^{22}$ 10° (c 0.5, MeOH).

*4-(((S)-1-(((S,E)-1-Fluoro-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamoyl)-1-methylpiperazine 1-oxide (**1b**).* Compound **38** (0.20 g) was dissolved in a 1:1 mixture of DCM/DMF and cooled to 0 °C. 0.05 g HOBt, 0.11 g TBTU, and 0.20 mL DIEA were added, and the mixture was stirred for 30 min until all components dissolved. 0.10 g of compound **13** were added, and the mixture was stirred for an additional 48 h at room temperature. The reaction was stopped by adding 7 equiv of water and the residue was extracted with EtOAc and washed with water (2×), $NaHCO_3$ (2×), 1 M HCl (2×), and brine (1×). The solvent was removed under reduced pressure and the crude product was purified by preparative HPLC and eluted with 70% water and 30% ACN, resulting in a colorless oil (0.04 g, 23%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 10.13 (s, 1H), 8.43−7.45 (m, 10H), 7.33−7.03 (m, 5H), 6.33−6.26 (m, 1H), 5.81 (dd, J = 22.1, 10.2 Hz, 1H), 5.50−5.28 (m, 1H), 4.48 (q, J = 8.0 Hz, 1H), 4.36−4.20 (m, 2H), 3.71−3.12 (m, 8H), 2.50 (quint, J = 1.9 Hz, 3H), 2.06 (d, J = 43.8 Hz, 2H), 2.04−1.64 (m, 2H). ^{13}C NMR (300 MHz, DMSO-d_6): δ 172.0, 157.3, 142.4, 139.6, 136.9, 134.0, 129.8, 129.0, 128.9, 128.6, 128.2, 127.7, 126.0, 105.6, 60.2, 57.2, 50.4, 42.5, 36.9, 30.4. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{32}H_{37}FN_4O_5S$, 609.25; found, 609.4. Purity: 99%. $[\alpha]_D^{22}$ 4° (c 0.5, MeOH).

*4-(((S)-1-Oxo-1-(((S,E)-1-(phenoxysulfonyl)-5-phenylpent-1-en-3-yl)amino)-3-(p-tolyl)propan-2-yl)carbamoyl)pyridine 1-Oxide (**2l**).* Compound **41** (0.038 g) was dissolved in DCM and cooled to 0 °C. 0.036 g HOBt, 0.087 g TBTU, and 0.19 mL DIEA were added and the mixture was stirred for 30 min. Then, 0.14 g of compound **34**-(**H**) was added and the mixture was stirred for an additional 24 h at room temperature. The reaction was stopped by adding 7 equiv of water and the residue was extracted with DCM and washed with water (2×), $NaHCO_3$ (2×), 1 M HCl (2×), and brine (1×). The solvent was removed under reduced pressure and the product was further purified by column chromatography (EtOAc/MeOH 9:1), resulting in a colorless solid (0.021 g, 13%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 9.11−8.90 (m, 1H), 8.51−8.39 (m, 1H), 8.36−8.20 (m, 2H), 7.91−7.78 (m, 2H), 7.52−7.38 (m, 2H), 7.37−7.29 (m, 2H), 7.27−7.18 (m, 6H), 7.15−7.00 (m, 4H), 6.68 (dt, J = 15.4, 4.0 Hz, 1H), 6.35 (dd, J = 15.4, 3.0 Hz, 1H), 4.72 (q, J = 10.1, 9.0 Hz, 1H), 4.41 (s, 1H), 3.13−2.92 (m, 2H), 2.66−2.46 (m, 2H), 2.22 (s, 3H), 1.27−1.09 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 171.4, 163.6, 149.6, 141.4, 139.3, 136.0, 135.0, 130.4, 129.5, 129.3, 128.8 (d, J = 3.6 Hz), 127.9, 126.3, 125.5, 123.6, 123.0, 55.9, 49.4, 39.9 (m), 34.5, 31.6, 21.1. mp 112−114 °C. LC−MS (ESI, m/z): $[M + H]^+$ calcd for $C_{33}H_{33}N_3O_6S$, 600.21; found, 600.3. Purity: 97%. $[\alpha]_D^{22}$ 6° (c 0.5, MeOH).

Enzyme Assays. *Fluorometric Assays.* Rhodesain was expressed as published previously.[34] The increase of fluorescence upon cleavage of the fluorogenic substrate Cbz-Phe-Arg-AMC (Bachem) by rhodesain, CatB, or CatL was monitored using a TECAN Infinite F200 Pro fluorimeter (δ excitation: 365 nm, δ emission: 460 nm). The enzymes were diluted from a stock solution (rhodesain: 4 mg/mL in 10 mM sodium citrate buffer, pH 5.5; CatB (human liver, Calbiochem): 0.532 mg/mL; CatL (human liver, Calbiochem): 0.266 mg/mL) with enzyme buffer (rhodesain: 50 mM sodium acetate pH 5.5, 5 mM EDTA, 200 mM NaCl and 2 mM DTT; CatB/CatL: 50 mM Tris−HCl, 5 mM EDTA, 200 mM NaCl, 2 mM DTT, pH 6.5) and were incubated for 1 h at rt. Assays were performed in black, flat-bottom 96-well microtiter plates (Greiner bio-one) with a total volume of 200 μL at 37 °C. Inhibitors and the substrate were prepared as stock solutions in DMSO. Dilution series of inhibitors in DMSO with a minimum of seven different concentrations were prepared in duplicate at least. 180 μL assay buffer (rhodesain: 50 mM sodium acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, and 0.005% Brij35; CatB/CatL: 50 mM Tris−HCl, 5 mM EDTA, 200 mM NaCl, 0.005% Brij35, pH 6.5) was added to the 96-well plates, then 5 μL of the respective enzyme in enzyme buffer, followed by 10 μL DMSO with or without inhibitor and finally 5 μL substrate (final substrate concentrations for rhodesain 10 μM, for CatB 100 μM, and for CatL 6.5 μM). Fluorescence emission was monitored directly after addition of the substrate. The presented data

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

are mean values of three independent measurements. Standard deviations are less than 10% unless otherwise depicted.

Screening of activity against the catalytically active subunits of the proteasome (commercially obtained from Enzo) and Dengue virus NS2/NS3 protease was performed with an inhibitor concentration of 11 μM in DMSO.[68] Human 20S proteasome (from human erythrocytes, 0.5 mg/mL) was diluted with enzyme buffer (for trypsin-like activity: 50 mM Tris−HCl, 50 mM NaCl, 0.5 mM EDTA, pH 7.4; for α-chymotrypsin-/caspase-like activity: 50 mM Tris−HCl, 25 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 0.03% SDS, pH 7.5). Trypsin-like activity was measured with the fluorogenic substrate boc-Leu-Arg-Arg-AMC (Bachem), α-chymotrypsin-like activity with Succ-Leu-Leu-Val-Tyr-AMC (Bachem), and caspase-like activity with Cbz-Leu-Leu-Glu-AMC (Bachem). Measurements were conducted in in black, flat-bottom 96-well microtiter plates (Greiner bio-one) by successive addition of 180 μL enzyme buffer, 5 μL proteasome in enzyme buffer, 10 μL inhibitor (11 μM in DMSO) or 10 μL pure DMSO, and 5 μL of the respective substrate (final substrate concentrations for trypsin-like activity 85 μM, α-chymotrypsin-like activity 70 μM, and caspase-like activity 80 μM). Fluorescence emission was monitored directly after addition of the substrate. Dengue virus NS2B/NS3 protease was diluted from a stock solution (1 mg/mL) with enzyme buffer (50 mM Tris−HCl, 1 mM Chaps, 20% glycerol).[68] Screening was performed with the substrate boc-Gly-Arg-Arg-AMC (Bachem) in black, flat-bottom 96-well microtiter plates (Greiner bio-one) by successive addition of 180 μL enzyme buffer, 5 μL Dengue protease in enzyme buffer, 10 μL inhibitor (11 μM in DMSO) or 10 μL pure DMSO, and 5 μL of the substrate (final substrate concentration 100 μM). Fluorescence emission was monitored directly after addition of the substrate.

The presented inhibition data for all enzymes are mean values of three independent measurements. Standard deviations are less than 10% unless otherwise depicted.

Calculations. The GraFit program (version 5.0.13, 2006, Erithacus Sofware Ltd., UK) was used for data analysis and non-linear regression.

For inhibitors showing time-independent inhibition (**2a-(Z)**, **2k**, **2j**, **3a−3i,** and **4a−4l**), the residual enzyme activity was plotted against the inhibitor concentration. IC_{50} values were obtained by non-linear regression using the equation

$$v_i = \frac{v_0}{1 + \left(\frac{[I]}{IC_{50}}\right)^S}$$

with v_0 = enzyme activity in the absence of inhibitor, v_i = enzyme activity in the presence of inhibitor, [I] = inhibitor concentration, and S = slope factor.

K_i values were obtained by correcting the IC_{50} values to zero substrate concentration using the Cheng−Prusoff relationship[32]

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_M}}$$

Rhodesain: [S] = 10 μM, K_M = 0.8265 μM, CatL: [S] = 6.25 μM, K_M = 6.5 μM, CatB: [S] = 100 μM, and K_M = 150 μM.

Calculation for the inhibitors with slow, tight-binding properties (**2a−2i**) is presented in the Results and Discussion section.

Dilution Assays. Rhodesain (5 μL from 4 mg/mL stock solution) in enzyme buffer (85 μL) was incubated for 30 min with inhibitors (10 μL in DMSO) at concentrations corresponding to 10-fold the IC_{50} value obtained from fluorometric enzyme assay to ensure complete inhibition. These mixtures (2 μL) were diluted 100-fold in assay buffer (198 μL) containing 5 μL substrate (400 μM) to give a final substrate concentration of 10 μM. Recovery of enzyme activity was measured immediately using a fluorescence readout. Rhodesain with DMSO and no inhibitor added was used as a reference while the irreversible inhibitor **K11777** was used as an irreversible control.

Dialysis Experiments. Dialysis experiments for rhodesain were performed in a custom-built dialysis chamber, allowing the parallel examination of five samples.[26] A 3.5 kDa MW cut-off dialysis tubing (Carl Roth, Zellutrans MWCO 3.5 kDa) was cut into half and placed in the instrument, separating a continuous flow of buffer from five cavities at the top of the instrument, where the samples (800 μL) were added. Therefore, 20 μL rhodesain in enzyme buffer was added to 740 μL assay buffer and 40 μL DMSO with or without inhibitor. Inhibitor concentrations of 10-fold the IC_{50} value were chosen in order to guarantee complete inhibition. After incubation for 30 min, the mixtures were transferred to the cavities in the instrument and dialyzed against a continuous flow of assay buffer containing 5% DMSO (300 mL/h). Samples (97.5 μL) were taken at different time points (10, 30, 60, and 120 min) and 2.5 μL of a substrate solution was added to give a final substrate concentration of 10 μM. Enzyme activity was determined by directly measuring the fluorescence emission. The results are given in fractional activity of uninhibited rhodesain used in the same experiment as positive control.

Docking Procedures. *Procedure for Non-covalent Docking with FlexX/LeadIT (vs. 2.1.3).*[11] Non-covalent docking experiments were performed using the crystal structure of rhodesain with covalently bound inhibitor **K11777** (pdb entry 2p7u).[60] The binding site was defined as a 6.5 Å shell around **K11777**. Water molecules present in the crystal structure were omitted except HOH-512, which mediates hydrogen bonding between the inhibitor and the peptide backbone. Generation of 3D-coordinates and energy minimization of the ligands were accomplished with the Molecular Operating Environment (MOE2014.09) using the MMFF94x force field.[26] Docking calculations were executed with LeadIT version 2.1.3.[27] The results presented in Table 1 are those with the best HYDE score[69,70] selected from the 10 highest-priced solutions according to FlexX score (Table 7.1).

Procedure for Covalent Docking with DOCKTITE.[60] Covalent docking was performed with the DOCKTITE software implementation (version 1.2) for Molecular Operating Environment (MOE, 2014.09; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2021**)[11,27] using the crystal structure of rhodesain with covalently bound inhibitor **K11777** (pdb entry 2p7u).[60] Energy minimization of the ligands was performed with MOE using the MMFF94x force field.[71] The different warheads were implemented into the DOCKTITE warhead filter file as described in the DOCKTITE manual. The standard DOCKTITE protocol was followed as described. The main pharmacophore docking step was performed without pharmacophore restriction for the nucleophilic sulfur of Cys25. The Amber12:EHT force field[72] was used for the pharmacophore docking step. Docking solutions were rescored with the MOE implemented scoring functions Affinity dG and additionally with the external empirical scoring function DSX.[73] The results shown in Table S1 are those with the best DSX scores.

Mass Spectrometry (MS). *ESI MS Analysis.* Lyophilized rhodesain was reconstituted at 4 mg/mL in 50 mM NaOAc, 200 mM NaCl, and 5 mM EDTA (pH 5.5). Prior to MS analysis, rhodesain was purified by weak anion exchange (WAX) chromatography using an *ÄKTA* protein purification system (GE Healthcare) equipped with a ProPac WAX-10G 4 × 50 mm guard and a ProPac WAX-10 4 × 250 mm analytical column (Thermo Fisher Scientific). 100 μL of rhodesain stock solution (4 mg/mL) was diluted in 20 mM imidazole, 1 mM DTT in water (pH 6.0) to a final volume of 1 mL, and separated running a gradient from 0 to 40% B in 40 min. Mobile phase A was 20 mM imidazole, 1 mM DTT in water (pH 6) and Solvent B 1 M NaCl, 20 mM imidazole, and 1 mM DTT in water (pH 6). Collected fractions were tested for activity.

For mass spectrometric analysis, the WAX fraction containing active rhodesain was further diluted in 50 mM NaOAc, 200 mM NaCl, 5 mM EDTA, and 5 mM DTT (pH 5.5) and incubated for 1 h at rt. After the addition of the inhibitors at a final concentration of 10 μM, samples were analyzed by LC−MS using a nanoAcquity UPLC system (Waters Corporation) coupled to a nano-ESI-Q-TOF mass spectrometer (Synapt G2-S HDMS, Waters Corporation). Rhodesain without compound served as control. Protein−drug complexes were loaded onto a 200 μm × 5 cm PepSwift Monolithic PS-DVB column from Dionex (Thermo Scientific) using direct injection mode. For LC separation, two mobile phases were used. Mobile phase A contained 0.1% formic acid (FA) and 3% DMSO in ultrapure water, whereas mobile phase B consisted of 0.1% FA and 3% DMSO in ACN. A gradient of 10−90% mobile phase B was run over 7 min at a flow rate of

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

2000 nL/min. Column temperature was set to 45 °C. After separation, the column was rinsed with 90% of mobile phase B and re-equilibrated under initial conditions. All MS analyses were conducted in positive-mode ESI.

MALDI-TOF MS Analysis. For this analysis, a recombinant rhodesain mutant expressed in *Pichia pastoris* was used as described in the enzyme assay section. The lyophilized protein was reconstituted in buffer A (pH 5.5, 50 mM NaOAc, 200 mM NaCl, 5 mM EDTA) at a concentration of 4 mg/mL (= 174 μM). This stock solution was diluted into buffer B (pH 5.5, 50 mM NaOAc, 200 mM NaCl, 5 mM EDTA, 5 mM DTT) to a final protein concentration of 1 or 10 μM and incubated for approximately 1 h, after which the compound of interest (4 mM stock in DMSO) was added at a final inhibitor concentration of 100 μM. Later, the analytes were desalted using Zeba Spin Desalting Columns (7 kDa MWCO, 0.5 mL; Thermo Fisher Scientific) in accordance with the manufacturer's instructions and afterward coprecipitated with a MALDI-matrix, utilizing two separate approaches: For the first method, a thin layer of sinapinic acid (saturated ethanolic solution) was prepared on the target, onto which, after film formation, a volumetric 3:1 mixture of matrix-solution to analyte solution was applied. For the second method, using a mixture of α-cyano-4-hydroxycinnamic acid and 2,5-dihydroxybenzoic acid,[40,44] the preparation of a basal matrix film was omitted and a volumetric 1:1 mixture of matrix-solution to analyte solution was prepared and applied to the target. After evaporation of the solvents (ca. 15 min), measurements were carried out on a rapifleX MALDI-TOF/TOF mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). The instrument is equipped with a scanning smartbeam 10 kHz Nd:YAG laser at a wavelength of 355 nm and a 10 bit 5 GHz digitizer. The acceleration voltage was set to 20 kV and the mass spectra were recorded in positive ion linear mode. Calibration was done with the Bruker protein calibration standard II in a mass range from 10 to 70 kDa. Samples were measured at a laser power of 100% (sinapinic acid) or 70% (matrix mixture), with random walk ionization across the sample spot. As control samples, rhodesain in buffer B with a DMSO concentration matching the samples, rhodesain incubated with a reportedly non-covalent inhibitor,[11] and a known covalent-irreversible inhibitor (**K11777**)[74] were used. Data analysis was performed using the open-source software mMass.[68]

QM/MM Computations. MD simulations were performed with the Amber program package (version 2018) in combination with the FF14SB force field. The unknown parameters for the ligand were calculated using GAFF.[75] The obtained enzymatic system was balanced with sodium ions. We added an octahedral water envelope of 10 Å consisting of TIP3P water molecules Figure S11).[76] All MD simulations were performed under periodic boundary conditions in three consecutive steps. In the first step, the cage of water molecules and then the whole system was minimized. In the next step, the system was heated in a controlled way from 0 to 300 K at 1 bar, using the SHAKE algorithm[77] employing the Langevin or Berendsen thermostat.[77,78] In the last step, the actual MD simulation with a duration of at least 10 ns was performed.

We used the subtractive QM/MM approach employing the electrostatic embedding scheme as implemented in the program package Gaussian (version 2016).[79] The QM part is specified in Figure S12. Please note that larger QM spaces did not lead to considerably changes (Table S4). For the QM part, we employed the ωB97XD functional in combination with the 6-31+G* basis sets.[80−82] The TAO-Toolkit was used.[83] For the intrinsic reaction coordinate (IRC) computations, the transition states were determined using the Berny-Algorithm.[84] All transition states were proved by frequency calculations.

The reaction paths were characterized by a two-step procedure. The covalent step of the inhibition mechanism first consists of the attack of the sulfur center of the thiolate group of Cys25 on the C_2 center of the alkene group. Additionally, a proton transfer from the protonated His152 moiety to the C_1 center occurs (Figure 1). Because it is unclear whether both steps proceed subsequently or simultaneously, we first computed two-dimensional scans selecting the distances $R(S_{Cys}-C_2)$ and $R(H_{His}-C_1)$ as main reaction parameters (Figure 14). By varying the main reaction parameters and optimizing all other internal coordinates for each pair of main coordinates, a minimum energy path (MEP) from the reactant to the product is obtained, which gives the first information about the shape of the reaction path. Starting from the barriers obtained in these scans, we performed intrinsic reaction coordinate (IRC) simulations. Due to the roughness of the computed PES, often more than one transition state was found for the two-dimensional surface. In such cases, we started IRC from each TS to ensure that they lead to similar reactants and products. This was indeed the case. One example is shown in Figures S8 and S10 in which computations started from slightly different TS but yielded comparable shapes for the reaction paths. The same holds for the second step of the reaction path given in Figure 16. For this frame, we found four TS whose IRCs all lead to very similar results.

Antitrypanosomal Activity and Cytotoxicity. Antitrypanosomal activity of **2a**, **2d**, **2e**, **3d**, **4d**, and **4e** against the *T. brucei brucei* BS449 cell line, a descendant of the Lister 427 strain,[85,86] was determined using the ATPlite assay as described previously.[56,57,87] **2a**, **2d**, **2e**, **3d**, **4d**, and **4e** were prepared as 5 mM stock solutions in DMSO and diluted in HMI-9 medium in multiple steps (1:3, then 1:10, and subsequently in ten 1:2 dilution steps using separate microplates). In white 96-well microplates (PerkinElmer), 10 μL of the final 101 dilutions were added to 90 μL of a cell suspension containing 2500 cells/mL, leading to final concentrations of each tested compound from 16.67 μM to 32.55 nM in the microplates. As a negative control, 0.3% DMSO was added to the cell suspension corresponding to the highest DMSO concentration added by compound application. Addition of 10% DMSO served as a positive control because all cells die at this concentration. The plates were prepared as triplicates and incubated for 24 and 48 h at 37 °C and 5% CO_2. After the respective incubation time, 50 μL ATPlite 1 step solution (PerkinElmer) was added to each well. The plate was shaken orbitally for 2 min and luminescence was measured at room temperature with an Infinite M200 PRO plate reader (Tecan Trading AG). The measured luminescence was plotted against the compound concentration to obtain a dose−response curve. The EC_{50} values were determined using GraFit version 5.013 (Erithacus Software Ltd.).

Antitrypanosomal activity of the compounds **1**, **4a**, **4j**, and **4l** was determined as described previously using the Alamar Blue assay.[54,55] Cytotoxicities against the macrophage cell line J774.1 and HeLa cells were investigated according to previously described methods.[12,88]

In Vitro Metabolism Studies. Rat liver microsomes were purchased from Sigma-Aldrich and characterized for cytochrome P450, cytochrome B5, and the activity of CYP1A, CYP3A, CYP2C, and cytochrome *c* reductase. The assay was performed as published previously.[68]

First, NADPH was generated by incubating potassium phosphate buffer (395 μL, 100 mM, pH 7.4), $MgCl_2$ (25 μL, 80 mM), glucose-6-phosphate (25 μL, 100 mM), NADP disodium salt (25 μL, 20 mM), and glucose-6-dehydrogenase (25 μL, 100 IU/mL, Sigma-Aldrich) at 37 °C for 10 min. After addition of the microsomes (25 μL, 20 mg/mL) and incubation (10 min) at 37 °C, the inhibitors (1 μL, 5.21 mM in acetonitrile) were added to the mixture and the incubation continued at 37 °C for 60 min. 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. The samples were centrifuged at 4 °C, 10,000*g* for 10 min. The supernatant was analyzed by LC−MS/MS using an Agilent Poroshell 120 EC-C18 150 × 2.10 mm 4 μm column; mobile phase: compound **1** (35% acetonitrile, 55% H_2O, 10% of a 0.1% solution of formic acid in water) and compound **2d**-(**H**) (55% acetonitrile, 35% 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. The LC−MS chromatograms and their corresponding mass spectra were analyzed using MestReNova (v. 12.0.4).[89]

In Vivo Distribution Studies. *Animals and Treatments.* 60 male and female CD1 mice of 35−45 g body weight (56−62 days of age) were purchased from Charles River Laboratories, Sulzfeld, Germany. Five mice per cage were housed in a controlled room (22 °C, 50−65% humidity; day/night cycle of 12/12 h) with free access to water and the standard laboratory diet (Altromin, Germany). After 7 days of

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

acclimatization, they were randomly divided into eight groups in which each group received four different testing compounds (**1**, **2d-(H)**) at a dose of 25 mg/kg by i.p. injection or oral gavage (p.o.). Compound **1** was diluted in 1% DMSO, while **2d-(H)** was diluted in 2% DMSO. The i.p. injection was given once only while the p.o. administrations were given for 4 days, twice daily (8 doses in total) at 9 and 18 o'clock. The dose of 25 mg/kg was chosen based on preliminary tests. 60 min after the administration of the testing compounds, blood was collected (about 0.2 mL) from the facial vein using a lancet (Goldenrod animal lancet). At the end of the experiment (3 h after the last drug administration), the animals were sacrificed by decapitation under isoflurane anesthesia. Trunk blood was collected, and the brain was removed and homogenized. Experiments were carried out according to the German Law for Animal Protection and were registered with Regierungspräsidium Darmstadt (FR/1021). All efforts were made to minimize animal numbers and animal suffering.

Microdialysis Experiments. Mice were anesthetized with isoflurane (induction dose 5%, maintenance dose 1.5% v/v in synthetic air) (Air Liquide, Düsseldorf, Germany) and placed in a stereotaxic frame (Stoelting, Chicago, IL, USA). A Y-shaped, concentric microdialysis probe with a molecular weight cutoff of 30 kDa and an exchange area of 2 mm was manufactured as described previously.[90] The probe was implanted in the hippocampus with the following coordinates from bregma: AP −2.0 mm; *L* +2.0 mm; DV −2.3 mm.[91] The mice were allowed to recover overnight. On the next day, the probes were perfused (rate: 2 μL/min) with artificial cerebrospinal fluid (aCSF; 147 mM NaCl, 4 mM KCl, 1.2 mM $CaCl_2$, and 1.2 mM $MgCl_2$) for 4 h and dialysates were collected every 20 min. After dialysates were collected for 1 h for equilibration, the test compound was given as described above. A blood sample was collected 1 h after administration of the test compound. At the end of the experiment (3 h after the last drug administration), the animals were sacrificed by decapitation under isoflurane anesthesia. Trunk blood was collected, and the brain was removed and homogenized.

Sample Preparation and LC−MS Analysis. The blood was centrifuged at 4 °C, 1500*g* for 20 min and the plasma supernatant was extracted with acetonitrile and centrifuged again at 4 °C, 10,000*g* for 10 min. The supernatant was analyzed by LC−MS using an Agilent Poroshell 120 EC-C18 150 × 2.10 mm 4 μm column; mobile phase: compound **1** (35% acetonitrile, 55% H_2O, 10% of a 0.1% solution of formic acid in water) and compound **2d-(H)** (55% acetonitrile, 35% 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. The LC−MS chromatograms and their corresponding mass spectra were analyzed using MestReNova (v. 12.0.4).[89]

The brain tissue was homogenized using a mixture of acetonitrile/MilliQ-water (ratio 2:1, 2 mL per 300 mg brain) in a tissue grinder (Potter S, B. Braun, Melsungen, Germany) at 1,500 rpm and 15 strokes. Afterward, the homogenate was centrifuged at 4 °C, 10,000*g* for 10 min and the supernatant was lyophilized. The lyophilizate was extracted with 400 μL acetonitrile and this extract was analyzed via LC−MS as described above. The fractions of the microdialysate were used without further processing. The AUC that was obtained from the extracted ion chromatograms was normalized to a volume of 500 μL for the plasma samples or a weight of 450 mg for the brain homogenates. The average AUC for every sample was calculated from experiments performed in triplicate.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.1c01002.

Scores for non-covalent (FlexX) and covalent docking (DOCKTITE); MALDI-TOF MS spectra of complexes of rhodesain with different inhibitors and different matrix substances; ESI MS spectra of complexes of rhodesain with inhibitors (*Z*)-**2a** and **2j**; additional data for the QM/MM computations; and ESI MS spectra of the in vitro metabolism and in vivo distribution studies of **1** and **2d-(H)** (PDF)

HPLC traces of lead compounds and molecular formula strings (CSV)

■ AUTHOR INFORMATION

Corresponding Author

Tanja Schirmeister − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*; Email: schirmei@uni-mainz.de

Authors

Sascha Jung − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*; Present Address: Faculty of Chemistry and Chemical Biology, TU Dortmund, Otto-Hahn-Str. 6, 44227 Dortmund, Germany

Natalie Fuchs − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*; orcid.org/0000-0001-6404-676X

Patrick Johe − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*

Annika Wagner − *Department of Chemistry, Biochemistry Section, Johannes Gutenberg University, 55128 Mainz, Germany*; Present Address: Institute of Organic Chemistry & Macromolecular Chemistry (IOMC), Friedrich Schiller University, Humboldtstraße 10, 07743 Jena, Germany; Cluster of Excellence "Balance of the Microverse", Friedrich-Schiller-University Jena, Germany.

Erika Diehl − *Department of Chemistry, Biochemistry Section, Johannes Gutenberg University, 55128 Mainz, Germany*

Tri Yuliani − *Institute for Pharmacology and Clinical Pharmacy, Goethe University, 60439 Frankfurt, Germany*; Present Address: Research Center for Chemistry, Indonesian Institute of Sciences (LIPI), Gd. 452 Kawasan PUSPIPTEK, Tangerang Selatan, Indonesia.

Collin Zimmer − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*

Fabian Barthels − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*

Robert A. Zimmermann − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*

Philipp Klein − *Department of Chemistry, Organic Chemistry Section, Johannes Gutenberg University, 55128 Mainz, Germany*

Waldemar Waigel − *Department of Physical and Theoretical Chemistry, Julius-Maximilians-University, 97074 Würzburg, Germany*

Jessica Meyr − *Department of Physical and Theoretical Chemistry, Julius-Maximilians-University, 97074 Würzburg, Germany*

Till Opatz − *Department of Chemistry, Organic Chemistry Section, Johannes Gutenberg University, 55128 Mainz, Germany*; orcid.org/0000-0002-3266-4050

Stefan Tenzer − *Institute for Immunology, University Medical Center, Johannes Gutenberg University, 55131 Mainz, Germany*; orcid.org/0000-0003-3034-0017

Ute Distler − *Institute for Immunology, University Medical Center, Johannes Gutenberg University, 55131 Mainz, Germany;* orcid.org/0000-0002-8031-6384
Hans-Joachim Räder − *Max Planck Institute for Polymer Research, 55128 Mainz, Germany;* orcid.org/0000-0002-7292-4013
Christian Kersten − *Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, 55128 Mainz, Germany*
Bernd Engels − *Department of Physical and Theoretical Chemistry, Julius-Maximilians-University, 97074 Würzburg, Germany;* orcid.org/0000-0003-3057-389X
Ute A. Hellmich − *Department of Chemistry, Biochemistry Section, Johannes Gutenberg University, 55128 Mainz, Germany; Center for Biomolecular Magnetic Resonance (BMRZ), Goethe University, 60438 Frankfurt, Germany;* Present Address: Institute of Organic Chemistry & Macromolecular Chemistry (IOMC), Friedrich Schiller University, Humboldtstraße 10, 07743 Jena, Germany; Cluster of Excellence "Balance of the Microverse", Friedrich-Schiller-University Jena, Germany.
Jochen Klein − *Institute for Pharmacology and Clinical Pharmacy, Goethe University, 60439 Frankfurt, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.1c01002

Author Contributions

‡‡S.J. and N.F. contributed equally.

Funding

Financial support by the Rhine-Main Universities fund (RMU-Initiativfonds Forschung) is gratefully acknowledged. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2051—Project ID390713860.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

E.D. acknowledges a PhD fellowship from the Sibylle Kalkhof-Rose-Stiftung.

■ ABBREVIATIONS

AcOH, acetic acid; aCSF, artificial cerebrospinal fluid; ADME, absorption, distribution, metabolism, excretion; AMC, 7-amido-4-methylcoumaryl; aq, aqueous; Arg, arginine; Asp, aspartate; AUC, area under the curve; Bn, benzyl; boc, *tert*-butyl carbonate; CatB, cathepsin B; CatL, cathepsin L; Cbz, carbobenzoxy; chymotrps., chymotrypsin; cpd, compound; Cys, cysteine; CYP, cytochrome P450; DCM, dichloromethane; DCTB, *trans*-2-[3-(4-*tert*-butylphenyl)-2-methyl-2-propenylidene]malononitrile; DECP, diethyl chlorophosphate; DIEA, *N*,*N*-diisopropylethylamine; DHBD, 2,3-dihydrobenzo-[*b*][1,4]dioxine; DMF, dimethyl formamide; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide; EDTA, ethylenediamine tetraacetic acid; EtOAc, ethyl acetate; ESI, electrospray ionization; Glu, glutamate; Gly, glycine; HAT, Human African *Trypanosomiasis*; HeLa cells, Henrietta Lacks cells; HOBt, hydroxybenzotriazole; HPhe, homophenylalanine; HPLC, high-pressure liquid chromatography; HWE, Horner−Wadsworth−Emmons; i.p., intraperitoneal; KHMDS, potassium bis-(trimethylsilyl)amide; LC, liquid chromatography; Leu, leucin; LHMDS, lithium bis(trimethylsilyl)amide; MALDI, matrix-assisted laser desorption/ionization; *m*CPBA, *meta*-chloroperbenzoic acid; Me, methyl; Met, methionine; MM, molecular mechanics; MOE, molecular operating environment; MS, mass spectrometry; MS-Cl, mesyl chloride; mp, melting point; NADPH, nicotinamide adenine dinucleotide phosphate (reduced form); *n*-BuLi, *n*-butyllithium; n.d., not determined; NHS, *N*-hydroxysuccinimide; NMR, nuclear magnetic resonance; NTD, neglected tropical disease; PDB, Protein Database; Ph, phenyl; Phe, phenylalanine; Pip, piperazine; ppm, parts per million; Pyr, pyridine; QM, quantum mechanics; RMSD, root-mean-square deviation; rt, room temperature; SAR, structure−activity relationship; SPR, surface plasmon resonance; Succ, *N*-succinyl; *T. b. gambiense*, *Trypanosoma brucei gambiense*; *T. b. rhodesiense*, *Trypanosoma brucei rhodesiense*; TBTU, 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate; TEA, triethyl amine; TFA, trifluoroacetic acid; THF, tetrahydrofuran; TLC, thin layer chromatography; TOF, time of flight; Tris, tris(hydroxymethyl)aminomethane; Trp, tryptophan; Tyr, tyrosine; UV, ultraviolet; val, valine; VS, vinyl sulfone; X-ray, energetic high-frequency electromagnetic radiation

■ REFERENCES

(1) Malvy, D.; Chappuis, F. Sleeping Sickness. *Clin. Microbiol. Infect.* **2011**, *17*, 986−995.
(2) Brun, R.; Blum, J.; Chappuis, F.; Burri, C. Human African Trypanosomiasis. *Lancet* **2010**, *375*, 148−159.
(3) Priotto, G.; Kasparian, S.; Mutombo, W.; Ngouama, D.; Ghorashian, S.; Arnold, U.; Ghabri, S.; Baudin, E.; Buard, V.; Kazadi-Kyanza, S.; Ilunga, M.; Mutangala, W.; Pohlig, G.; Schmid, C.; Karunakara, U.; Torreele, E.; Kande, V. Nifurtimox-Eflornithine Combination Therapy for Second-Stage African Trypanosoma Brucei Gambiense Trypanosomiasis: A Multicentre, Randomised, Phase III, Non-Inferiority Trial. *Lancet* **2009**, *374*, 56−64.
(4) Mesu, V. K. B. K.; Kalonji, W. M.; Bardonneau, C.; Mordt, O. V.; Blesson, S.; Simon, F.; Delhomme, S.; Bernhard, S.; Kuziena, W.; Lubaki, J.-P. F.; Vuvu, S. L.; Ngima, P. N.; Mbembo, H. M.; Ilunga, M.; Bonama, A. K.; Heradi, J. A.; Solomo, J. L. L.; Mandula, G.; Badibabi, L. K.; Dama, F. R.; Lukula, P. K.; Tete, D. N.; Lumbala, C.; Scherrer, B.; Strub-Wourgaft, N.; Tarral, A. Oral Fexinidazole for Late-Stage African Trypanosoma Brucei Gambiense Trypanosomiasis: A Pivotal Multicentre, Randomised, Non-Inferiority Trial. *Lancet* **2018**, *391*, 144−154.
(5) Fairlamb, A. H. Chemotherapy of Human African Trypanosomiasis: Current and Future Prospects. *Trends Parasitol.* **2003**, *19*, 488−494.
(6) Steverding, D.; Sexton, D. W.; Wang, X.; Gehrke, S. S.; Wagner, G. K.; Caffrey, C. R. Trypanosoma Brucei: Chemical Evidence That Cathepsin L Is Essential for Survival and a Relevant Drug Target. *Int. J. Parasitol.* **2012**, *42*, 481−488.
(7) Nikolskaia, O. v.; de A Lima, A. P. C.; Kim, Y. v.; Lonsdale-Eccles, J. D.; Fukuma, T.; Scharfstein, J.; Grab, D. J. Blood-Brain Barrier Traversal by African Trypanosomes Requires Calcium Signaling Induced by Parasite Cysteine Protease. *J. Clin. Invest.* **2006**, *116*, 2739−2747.
(8) Scory, S.; Caffrey, C. R.; Stierhof, Y.-D.; Ruppel, A.; Steverding, D. Trypanosoma Brucei: Killing of Bloodstream Formsin Vitroandin Vivo by the Cysteine Proteinase Inhibitor Z-Phe-Ala-CHN2. *Exp. Parasitol.* **1999**, *91*, 327−333.
(9) Kerr, I. D.; Wu, P.; Marion-Tsukamaki, R.; Mackey, Z. B.; Brinen, L. S. Crystal Structures of TbCatB and Rhodesain, Potential Chemotherapeutic Targets and Major Cysteine Proteases of Trypanosoma Brucei. *PLoS Neglected Trop. Dis.* **2010**, *4*, No. e701.
(10) Johé, P.; Jaenicke, E.; Neuweiler, H.; Schirmeister, T.; Kersten, C.; Hellmich, U. A. Structure, Interdomain Dynamics and PH-Dependent Autoactivation of pro-Rhodesain, the Main Lysosomal

Cysteine Protease from African Trypanosomes. *J. Biol. Chem.* **2021**, *296*, 100565.
(11) Kerr, I. D.; Lee, J. H.; Farady, C. J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K. C.; Caffrey, C. R.; Legac, J.; Hansell, E.; Mckerrow, J. H.; Craik, C. S.; Rosenthal, P. J.; Brinen, L. S. Vinyl Sulfones as Antiparasitic Agents and a Structural Basis for Drug Design. *J. Biol. Chem.* **2009**, *284*, 25697−25703.
(12) Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T. H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P. R.; Distler, U.; Tenzer, S.; Johé, P.; Hellmich, U. A.; Engels, B. Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138*, 8332−8335.
(13) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The Resurgence of Covalent Drugs. *Nat. Rev. Drug Discovery* **2011**, *10*, 307−317.
(14) Bauer, R. A. Covalent Inhibitors in Drug Discovery: From Accidental Discoveries to Avoided Liabilities and Designed Therapies. *Drug Discovery Today* **2015**, *20*, 1061−1073.
(15) Kalgutkar, A. S.; Dalvie, D. K. Drug Discovery for a New Generation of Covalent Drugs. *Expert Opin. Drug Discovery* **2012**, *7*, 561−581.
(16) Lammert, C.; Einarsson, S.; Saha, C.; Niklasson, A.; Bjornsson, E.; Chalasani, N. Relationship between Daily Dose of Oral Medications and Idiosyncratic Drug-Induced Liver Injury: Search for Signals. *Hepatology* **2008**, *47*, 2003−2009.
(17) Lee, C.-U.; Grossmann, T. N. Reversible Covalent Inhibition of a Protein Target. *Angew. Chem., Int. Ed.* **2012**, *51*, 8699−8700.
(18) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-Target Residence Time and Its Implications for Lead Optimization. *Nat. Rev. Drug Discovery* **2006**, *5*, 730−739.
(19) Bradshaw, J. M.; McFarland, J. M.; Paavilainen, V. O.; Bisconte, A.; Tam, D.; Phan, V. T.; Romanov, S.; Finkle, D.; Shu, J.; Patel, V.; Ton, T.; Li, X.; Loughhead, D. G.; Nunn, P. A.; Karr, D. E.; Gerritsen, M. E.; Funk, J. O.; Owens, T. D.; Verner, E.; Brameld, K. A.; Hill, R. J.; Goldstein, D. M.; Taunton, J. Prolonged and Tunable Residence Time Using Reversible Covalent Kinase Inhibitors. *Nat. Chem. Biol.* **2015**, *11*, 525−531.
(20) Johe, P.; Jung, S.; Endres, E.; Kersten, C.; Zimmer, C.; Ye, W.; Sönnichsen, C.; Hellmich, U. A.; Sotriffer, C.; Schirmeister, T.; Neuweiler, H. Warhead Reactivity Limits the Speed of Inhibition of the Cysteine Protease Rhodesain. *ACS Chem. Biol.* **2021**, *16*, 661−670.
(21) Schechter, I.; Berger, A. On the Size of the Active Site in Proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157−162.
(22) Jaishankar, P.; Hansell, E.; Zhao, D.-M.; Doyle, P. S.; McKerrow, J. H.; Renslo, A. R. Potency and Selectivity of P2/P3-Modified Inhibitors of Cysteine Proteases from Trypanosomes. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 624−628.
(23) Schirmeister, T.; Schmitz, J.; Jung, S.; Schmenger, T.; Krauth-Siegel, R. L.; Gütschow, M. Evaluation of Dipeptide Nitriles as Inhibitors of Rhodesain, a Major Cysteine Protease of Trypanosoma Brucei. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 45−50.
(24) Yang, P.-Y.; Wang, M.; Li, L.; Wu, H.; He, C. Y.; Yao, S. Q. Design, Synthesis and Biological Evaluation of Potent Azadipeptide Nitrile Inhibitors and Activity-Based Probes as Promising Anti-*Trypanosoma Brucei* Agents. *Chem.—Eur. J.* **2012**, *18*, 6528−6541.
(25) Caffrey, C. R.; Hansell, E.; Lucas, K. D.; Brinen, L. S.; Alvarez Hernandez, A.; Cheng, J.; Gwaltney, S. L.; Roush, W. R.; Stierhof, Y.-D.; Bogyo, M.; Steverding, D.; McKerrow, J. H. Active Site Mapping, Biochemical Properties and Subcellular Localization of Rhodesain, the Major Cysteine Protease of Trypanosoma Brucei Rhodesiense. *Mol. Biochem. Parasitol.* **2001**, *118*, 61−73.
(26) *LeadIT/FlexX*; BioSolveIT GmbH: St. Augustin, Germany 2012.
(27) Scholz, C.; Knorr, S.; Hamacher, K.; Schmidt, B. DOCKTITE-A Highly Versatile Step-by-Step Workflow for Covalent Docking and Virtual Screening in the Molecular Operating Environment. *J. Chem. Inf. Model.* **2015**, *55*, 398−406.
(28) Ng, S. L.; Yang, P.-Y.; Chen, K. Y.-T.; Srinivasan, R.; Yao, S. Q. "Click" Synthesis of Small-Molecule Inhibitors Targeting Caspases. *Org. Biomol. Chem.* **2008**, *6*, 844−847.
(29) Sun, A. C.; McClain, E. J.; Beatty, J. W.; Stephenson, C. R. J. Visible Light-Mediated Decarboxylative Alkylation of Pharmaceutically Relevant Heterocycles. *Org. Lett.* **2018**, *20*, 3487−3490.
(30) Ehmke, V.; Winkler, E.; Banner, D. W.; Haap, W.; Schweizer, W. B.; Rottmann, M.; Kaiser, M.; Freymond, C.; Schirmeister, T.; Diederich, F. Optimization of Triazine Nitriles as Rhodesain Inhibitors: Structure-Activity Relationships, Bioisosteric Imidazopyridine Nitriles, and X-Ray Crystal Structure Analysis with Human Cathepsin L. *ChemMedChem* **2013**, *8*, 967−975.
(31) DIXON, M. The Determination of Enzyme Inhibitor Constants. *Biochem. J.* **1953**, *55*, 170−171.
(32) 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*, 3099−3108.
(33) Purich, D. L. *Enzyme Kinetics, Catalysis and Control, A Reference of Theory and Best-Practice Methods*; Elsevier, 2010.
(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) 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.
(36) Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J. M.; Hartmann, G.; Banner, D. W.; Haap, W.; Diederich, F. Systematic Investigation of Halogen Bonding in Protein-Ligand Interactions. *Angew. Chem., Int. Ed.* **2011**, *50*, 314−318.
(37) Mirković, B.; Renko, M.; Turk, S.; Sosič, I.; Jevnikar, Z.; Obermajer, N.; Turk, D.; Gobec, S.; Kos, J. Novel Mechanism of Cathepsin B Inhibition by Antibiotic Nitroxoline and Related Compounds. *ChemMedChem* **2011**, *6*, 1351−1356.
(38) Illy, C.; Quraishi, O.; Wang, J.; Purisima, E.; Vernet, T.; Mort, J. S. Role of the Occluding Loop in Cathepsin B Activity. *J. Biol. Chem.* **1997**, *272*, 1197−1202.
(39) Zehl, M.; Allmaier, G. n. Investigation of Sample Preparation and Instrumental Parameters in the Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry of Noncovalent Peptide/Peptide Complexes. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1931−1940.
(40) Salih, B.; Zenobi, R. MALDI Mass Spectrometry of Dye - Peptide and Dye - Protein Complexes. *Anal. Chem.* **1998**, *70*, 1536−1543.
(41) Woods, A. S.; Huestis, M. A. A Study of Peptide-Peptide Interaction by Matrix-Assisted Laser Desorption/Ionization. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 88−96.
(42) Heck, A. J. R.; van den Heuvel, R. H. H. Investigation of Intact Protein Complexes by Mass Spectrometry. *Mass Spectrom. Rev.* **2004**, *23*, 368−389.
(43) Laugesen, S.; Roepstorff, P. Combination of Two Matrices Results in Improved Performance of MALDI MS for Peptide Mass Mapping and Protein Analysis. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 992−1002.
(44) Klein, P.; Johe, P.; Wagner, A.; Jung, S.; Kühlborn, J.; Barthels, F.; Tenzer, S.; Distler, U.; Waigel, W.; Engels, B.; Hellmich, U. A.; Opatz, T.; Schirmeister, T. New Cysteine Protease Inhibitors: Electrophilic (Het)Arenes and Unexpected Prodrug Identification for the Trypanosoma Protease Rhodesain. *Molecules* **2020**, *25*, 1451.
(45) Whittal, R. M.; Li, L. High-Resolution Matrix-Assisted Laser Desorption/Ionization in a Linear Time-of-Flight Mass Spectrometer. *Anal. Chem.* **1995**, *67*, 1950−1954.
(46) Bahr, U.; Stahl-Zeng, J.; Gleitsmann, E.; Karas, M. Delayed Extraction Time-of-Flight MALDI Mass Spectrometry of Proteins above 25 000 Da. *J. Mass Spectrom.* **1997**, *32*, 1111−1116.
(47) Kollmeier, A. S.; de la Torre, X.; Müller, C.; Botrè, F.; Parr, M. K. In-Depth Gas Chromatography/Tandem Mass Spectrometry Fragmentation Analysis of Formestane and Evaluation of Mass Spectral Discrimination of Isomeric 3-Keto-4-Ene Hydroxy Steroids. *Rapid Commun. Mass Spectrom.* **2020**, *34*, No. e8937.

(48) Mladenovic, M.; Ansorg, K.; Fink, R. F.; Thiel, W.; Schirmeister, T.; Engels, B. Atomistic Insights into the Inhibition of Cysteine Proteases: First QM/MM Calculations Clarifying the Stereoselectivity of Epoxide-Based Inhibitors. *J. Phys. Chem. B* **2008**, *112*, 11798−11808.
(49) Mladenovic, M.; Junold, K.; Fink, R. F.; Thiel, W.; Schirmeister, T.; Engels, B. Atomistic Insights into the Inhibition of Cysteine Proteases: First QM/MM Calculations Clarifying the Regiospecificity and the Inhibition Potency of Epoxide- and Aziridine-Based Inhibitors. *J. Phys. Chem. B* **2008**, *112*, 5458−5469.
(50) Vicik, R.; Busemann, M.; Gelhaus, C.; Stiefl, N.; Scheiber, J.; Schmitz, W.; Schulz, F.; Mladenovic, M.; Engels, B.; Leippe, M.; Baumann, K.; Schirmeister, T. Aziridide-Based Inhibitors of Cathepsin L: Synthesis, Inhibition Activity, and Docking Studies. *ChemMedChem* **2006**, *1*, 1126−1141.
(51) Müller, K.; Faeh, C.; Diederich, F. Fluorine in Pharmaceuticals: Looking beyond Intuition. *Science* **2007**, *317*, 1881−1886.
(52) Tosstorff, A.; Cole, J. C.; Taylor, R.; Harris, S. F.; Kuhn, B. Identification of Noncompetitive Protein-Ligand Interactions for Structural Optimization. *J. Chem. Inf. Model.* **2020**, *60*, 6595−6611.
(53) Maeda, S.; Harabuchi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. Intrinsic Reaction Coordinate: Calculation, Bifurcation, and Automated Search. *Int. J. Quantum Chem.* **2015**, *115*, 258−269.
(54) Vicik, R.; Hoerr, V.; Glaser, M.; Schultheis, M.; Hansell, E.; McKerrow, J. H.; Holzgrabe, U.; Caffrey, C. R.; Ponte-Sucre, A.; Moll, H.; Stich, A.; Schirmeister, T. Aziridine-2,3-Dicarboxylate Inhibitors Targeting the Major Cysteine Protease of Trypanosoma Brucei as Lead Trypanocidal Agents. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2753−2757.
(55) Ettari, R.; Pinto, A.; Previti, S.; Tamborini, L.; Angelo, I. C.; la Pietra, V.; Marinelli, L.; Novellino, E.; Schirmeister, T.; Zappalà, M.; Grasso, S.; de Micheli, C.; Conti, P. Development of Novel Dipeptide-like Rhodesain Inhibitors Containing the 3-Bromoisoxazoline Warhead in a Constrained Conformation. *Bioorg. Med. Chem.* **2015**, *23*, 7053−7060.
(56) Wagner, A.; Le, T. A.; Brennich, M.; Klein, P.; Bader, N.; Diehl, E.; Paszek, D.; Weickhmann, A. K.; Dirdjaja, N.; Krauth-Siegel, R. L.; Engels, B.; Opatz, T.; Schindelin, H.; Hellmich, U. A. Inhibitor-Induced Dimerization of an Essential Oxidoreductase from African Trypanosomes. *Angew. Chem., Int. Ed.* **2019**, *58*, 3640−3644.
(57) Previti, S.; Ettari, R.; Cosconati, S.; Amendola, G.; Chouchene, K.; Wagner, A.; Hellmich, U. A.; Ulrich, K.; Krauth-Siegel, R. L.; Wich, P. R.; Schmid, I.; Schirmeister, T.; Gut, J.; Rosenthal, P. J.; Grasso, S.; Zappalà, M. Development of Novel Peptide-Based Michael Acceptors Targeting Rhodesain and Falcipain-2 for the Treatment of Neglected Tropical Diseases (NTDs). *J. Med. Chem.* **2017**, *60*, 6911−6923.
(58) Mellott, D. M.; Tseng, C.-T.; Drelich, A.; Fajtová, P.; Chenna, B. C.; Kostomiris, D. H.; Hsu, J.; Zhu, J.; Taylor, Z. W.; Kocurek, K. I.; Tat, V.; Katzfuss, A.; Li, L.; Giardini, M. A.; Skinner, D.; Hirata, K.; Yoon, M. C.; Beck, S.; Carlin, A. F.; Clark, A. E.; Beretta, L.; Maneval, D.; Hook, V.; Frueh, F.; Hurst, B. L.; Wang, H.; Raushel, F. M.; O'Donoghue, A. J.; de Siqueira-Neto, J. L.; Meek, T. D.; McKerrow, J. H. A Clinical-Stage Cysteine Protease Inhibitor Blocks SARS-CoV-2 Infection of Human and Monkey Cells. *ACS Chem. Biol.* **2021**, *16*, 642−650.
(59) Jacobsen, W.; Christians, U.; Benet, L. Z. In Vitro Evaluation of the Disposition of a Novel Cysteine Protease Inhibitor. *Drug Metab. Dispos.* **2000**, *28*, 1343−1351.
(60) *Molecular Operating Environment (MOE)|MOEsaic|PSILO*; Chemical Computing Group Inc.: Montreal, QC, Canada, 2014.
(61) Lei, X.; Jalla, A.; Abou Shama, M.; Stafford, J.; Cao, B. Chromatography-Free and Eco-Friendly Synthesis of Aryl Tosylates and Mesylates. *Synthesis* **2015**, *47*, 2578−2585.
(62) Stokes, B. J.; Liu, S.; Driver, T. G. Rh2(II)-Catalyzed Nitro-Group Migration Reactions: Selective Synthesis of 3-Nitroindoles from *β*-Nitro Styryl Azides. *J. Am. Chem. Soc.* **2011**, *133*, 4702−4705.
(63) Scheidt, K. A.; Roush, W. R.; McKerrow, J. H.; Selzer, P. M.; Hansell, E.; Rosenthal, P. J. Structure-Based Design, Synthesis and Evaluation of Conformationally Constrained Cysteine Protease Inhibitors. *Bioorg. Med. Chem.* **1998**, *6*, 2477−2494.
(64) Ivkovic, J.; Lembacher-Fadum, C.; Breinbauer, R. A Rapid and Efficient One-Pot Method for the Reduction of N-Protected *α*-Amino Acids to Chiral *α*-Amino Aldehydes Using CDI/DIBAL-H. *Org. Biomol. Chem.* **2015**, *13*, 10456−10460.
(65) White, J. D.; Xu, Q.; Lee, C.-S.; Valeriote, F. A. Total Synthesis and Biological Evaluation of (+)-Kalkitoxin, a Cytotoxic Metabolite of the Cyanobacterium Lyngbya Majuscula. *Org. Biomol. Chem.* **2004**, *2*, 2092−2102.
(66) Weissenborn, M. J.; Wehner, J. W.; Gray, C. J.; Šardzík, R.; Eyers, C. E.; Lindhorst, T. K.; Flitsch, S. L. Formation of Carbohydrate-Functionalised Polystyrene and Glass Slides and Their Analysis by MALDI-TOF MS. *Beilstein J. Org. Chem.* **2012**, *8*, 753−762.
(67) Andrei, D.; Wnuk, S. F. Synthesis of the Multisubstituted Halogenated Olefins via Cross-Coupling of Dihaloalkenes with Alkylzinc Bromides. *J. Org. Chem.* **2006**, *71*, 405−408.
(68) 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*, 11359−11382.
(69) Schneider, N.; Lange, G.; Hindle, S.; Klein, R.; Rarey, M. A Consistent Description of HYdrogen Bond and DEhydration Energies in Protein-Ligand Complexes: Methods behind the HYDE Scoring Function. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 15−29.
(70) Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. Towards an Integrated Description of Hydrogen Bonding and Dehydration: Decreasing False Positives in Virtual Screening with the HYDE Scoring Function. *ChemMedChem* **2008**, *3*, 885−897.
(71) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490−519.
(72) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(73) Neudert, G.; Klebe, G. DSX: A Knowledge-Based Scoring Function for the Assessment of Protein-Ligand Complexes. *J. Chem. Inf. Model.* **2011**, *51*, 2731−2745.
(74) Marinas, M.; Sa, E.; Rojas, M. M.; Moalem, M.; Urbano, F. J.; Guillou, C.; Rallo, L. A Nuclear Magnetic Resonance (1 H and 13 C) and Isotope Ratio Mass Spectrometry (d 13 C , d 2 H and d 18 O) Study of Andalusian Olive Oils. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1457−1466.
(75) 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.
(76) 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.
(77) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(78) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(79) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J.

B.; Fox, D. J. *Gaussian 16*, Rev. C.01; Gaussian Inc., Wallingford, CT, 2016.
(80) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973**, *28*, 213−222.
(81) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(82) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724.
(83) Tao, P.; Schlegel, H. B. A Toolkit to Assist ONIOM Calculations. *J. Comput. Chem.* **2010**, *31*, 2363−2369.
(84) Li, X.; Frisch, M. J. Energy-Represented Direct Inversion in the Iterative Subspace within a Hybrid Geometry Optimization Method. *J. Chem. Theory Comput.* **2006**, *2*, 835−839.
(85) Biebinger, S.; Elizabeth Wirtz, L.; Lorenz, P.; Christine Clayton, C. Vectors for Inducible Expression of Toxic Gene Products in Bloodstream and Procyclic Trypanosoma Brucei. *Mol. Biochem. Parasitol.* **1997**, *85*, 99−112.
(86) Cunningham, M. P.; Vickerman, K. Antigenic Analysis in the Trypanosoma Brucei Group, Using the Agglutination Reaction. *Trans. R. Soc. Trop. Med. Hyg.* **1962**, *56*, 48−59.
(87) Klein, P.; Barthels, F.; Johe, P.; Wagner, A.; Tenzer, S.; Distler, U.; Le, T. A.; Schmid, P.; Engel, V.; Engels, B.; Hellmich, U. A.; Opatz, T.; Schirmeister, T. Naphthoquinones as Covalent Reversible Inhibitors of Cysteine Proteases-Studies on Inhibition Mechanism and Kinetics. *Molecules* **2020**, *25*, 2064.
(88) 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.
(89) Mnova 12.0.4—Mestrelab. https://mestrelab.com/download_file/mnova-12-0-4/ (accessed May 3, 2021).
(90) Lietsche, J.; Gorka, J.; Hardt, S.; Karas, M.; Klein, J. Custom-Made Microdialysis Probe Design. *J. Visualized Exp.* **2015**, *2015*, 53048.
(91) Buttner-Ennever, J. *The Rat Brain in Stereotaxic Coordinates*, 3; Paxinos, G.; Watson, C.; Academic Press: San Diego, 1996; Journal of Anatomy 1997, Vol. *191* (2), pp 315−317.

https://doi.org/10.1021/acs.jmedchem.1c01002
J. Med. Chem. 2021, 64, 12322−12358

Supporting Information

Fluorovinylsulfones and -sulfonates as potent covalent reversible inhibitors of the trypanosomal cysteine protease rhodesain: SAR, inhibition mechanism, metabolism and in vivo studies

Sascha Jung,[1,†,‡] Natalie Fuchs,[1,‡] Patrick Johe,[1] Annika Wagner,[2,⸸] Erika Diehl,[2] Tri Yuliani,[3,#] Collin Zimmer,[1] Fabian Barthels,[1] Robert A. Zimmermann,[1] Philipp Klein,[4] Waldemar Waigel,[5] Jessica Meyr,[5] Till Opatz,[4] Stefan Tenzer,[5] Ute Distler,[5] Hans-Joachim Räder,[6] Christian Kersten,[1] Bernd Engels,[5] Ute A. Hellmich,[2,7,⸸] Jochen Klein,[3] and Tanja Schirmeister[1,]*

1 Institute of Pharmaceutical and Biomedical Sciences (IPBS), Johannes Gutenberg University, Staudinger Weg 5, 55128 Mainz, Germany.

2 Department of Chemistry, Biochemistry Section, Johannes Gutenberg University, Hanns-Dieter-Hüsch-Weg 17, 55128 Mainz, Germany.

3 Institute for Pharmacology and Clinical Pharmacy, Goethe University, Max-von-Laue-Str. 9, 60439 Frankfurt, Germany.

4 Department of Chemistry, Organic Chemistry Section, Johannes Gutenberg University, Duesbergweg 10-14, 55128 Mainz, Germany.

5 Department of Physical and Theoretical Chemistry, Julius-Maximilians-University, Emil-Fischer-Str. 42, 97074 Würzburg, Germany.

6 Institute for Immunology, University Medical Center, Johannes Gutenberg University, Langenbeckstraße 1, 55131 Mainz, Germany.

7 Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany.

8 Center for Biomolecular Magnetic Resonance (BMRZ), Goethe University, Max-von-Laue-Str. 9, 60438 Frankfurt, Germany.

* To whom correspondence should be addressed: Prof. Dr. Tanja Schirmeister, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University, Mainz, Email: schirmei@uni-mainz.de.

‡These authors contributed equally.

Table of contents.

Docking:

Mass spectrometry:

QM/MM computations:

- **Table S3:** Variations in the important geometrical parameters along the IRC paths.
- **Figure S3:** Potential energy curves of the torsional motion around $\measuredangle H^3C^3C^2H^2$ obtained for frame 1 of K11777
- **Figure S4:** Intrinsic reaction coordinate (IRC) scan of the inhibition reaction of K11777 starting from the H-orientation
- **Figure S5:** Potential energy curve of the torsion motion around $\measuredangle H^3C^3C^2H^2$ obtained for frame 2 of K11777.
- **Figure S6:** Potential energy curve of the torsional motion around $\measuredangle H^3C^3C^2H^2$ obtained for frame 2 of inhibitor **1**.
- **Figure S7:** Intrinsic reaction coordinate (IRC) scan of the inhibition reaction of inhibitor **1** starting from the F-orientation.
- **Figure S8:** Intrinsic reaction coordinate (IRC) scan of the inhibition reaction of inhibitor **1** starting from the H-orientation of frame 2.
- **Figure S9:** Potential energy curve of the torsional motion around $\measuredangle H^3C^3C^2H^2$ obtained for the frame 2 of inhibitor **1**.
- **Figure S10:** Intrinsic reaction coordinate (IRC) scan of the inhibition reaction of inhibitor **1** starting from the H-orientation of the frame 2 starting from a slightly different TS than used in Figure 8SI.
- **Figure S11:** Composition and division of the QM/MM system.
- **Figure S12:** Description of the standard (black), middle (blue), and large (red) sized QM layer in the QM/MM computations.
- **Table S4:** Computations to test the influence of the size of the QM part within the QM/MM computations for K11777.

Metabolism:

- **Figure S13:** In vitro metabolism of compound **1**.
- **Figure S14:** In vitro metabolism of compound **2d-(H).**

In vivo studies:

- **Figure S15:** Extracted ion chromatograms and mass spectra for the samples of compound **1** obtained in the in vivo studies.
- **Figure S16:** Extracted ion chromatograms and mass spectra for the samples of compound **2d-(H)** obtained in the in vivo studies.

HPLC traces and mass spectra of lead compounds.

MALDI-TOF MS spectra of complexes of rhodesain with different inhibitors and different matrix substances.

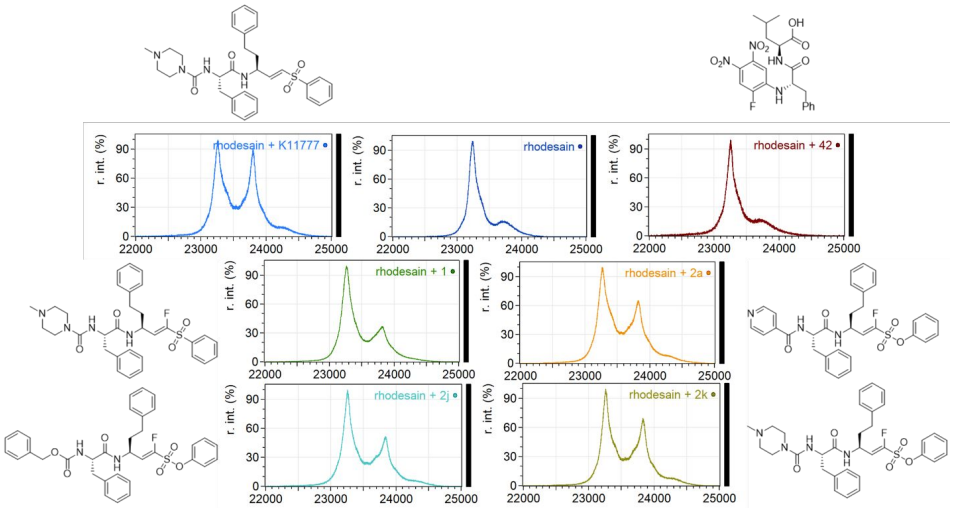


Figure S1. MALDI-TOF spectra of rhodesain (ca. 23 kDa) in the presence of different inhibitors (tenfold molar excess of inhibitor to protein). A matrix mixture of α-cyano-4-hydroxycinnamic acid and 2,5-dihydroybenzoic acid was used as matrix substance. The figures show the relevant range of the spectrum to display $[M+H]^+$ which were baseline corrected.

”

2: Rhodesain Inhibitors on the Edge of Reversibility-Irreversibility.

Commonly, electrophilic warheads each convey one specific reaction mechanism with a catalytic thiol. A nitrile reacts in a reversible addition to a thioimidate, an acrylamide reacts in an irreversible Michael addition. However, there is interest in warheads with multiple possible reaction sites. This was especially researched in the context of the proteasome subunits that have two nucleophilic residues at their *N*-terminal threonine, which can be addressed simultaneously by dual-reactive keto epoxides like carfilzomib.[200] Another example is the halomethyl ketone warhead against serine or cysteine proteases, for which an initially reversible state is rearranged to an irreversible product.[609] Against human cathepsin L, keto vinylester and keto vinylsulfone warheads were investigated in the past, which can have up to four possible positions for a nucleophilic attack.[610] The keto vinylester warhead has been investigated both computationally and experimentally against rhodesain.[611–613] On the topic of the covalency of ketone-based inhibitors of parasitic cysteine proteases, there is one publication that discusses crystallographic data of a putatively non-covalently binding hydroxymethyl ketone with still nanomolar K_i value.[253] This is an uncommon observation since peptidic recognition sequences without a suitable warhead for a covalent reaction tend to show no relevant inhibition.[253,614]

In this study, a side-by-side comparison of dual reactive warheads was performed and complemented by similar, purely irreversible Michael acceptors, all bearing the peptide sequence of K11777. The warheads were acrylate (**11**), hydroxy vinylester (**14**), keto vinylester (**15**), hydroxy vinylsulfone (**18**), and keto vinylsulfone (**19**). Rhodesain was used as a model protease for these mechanistic investigations. As is established for Michael acceptor-based warheads, they were first evaluated assuming irreversible kinetics. But since the two ketone-derived warheads can follow an alternative, reversibly covalent reaction path, and showed slower inactivation rates than the classical Michael acceptors, a time-independent evaluation of the initial state data was also performed to have a metric for the reversible equilibrium. Selectivity against related cathepsins B and L was evaluated. Since the peptidic recognition unit of K11777 was employed, the observed low selectivity was mostly unexpected. To add information to the binding mode of these multifaceted inhibitors, the covalency of binding was verified by MALDI-TOF mass spectrometry. This elucidated that the method could discriminate irreversible binders with low affinity or low reactivity. Reversibility was assessed with dilution as well as dialysis assays, which corroborated the mechanistic assumption that both reactions happen in parallel, generating covalent reversibly and irreversibly inhibited populations of rhodesain at any given time. Naturally, with time, the irreversible mode dominates. A combination of computational methods (namely non-covalent and covalent docking, as well molecular dynamics simulations) showed that both expected reactions can take place with comparable probability. Taken together, this publication gave further mechanistic insight into a class of warheads with a dual mode of action. The observed tendency for the irreversible option proportional to concentration and time can be advantageous with regards to a lysosomotropic effect of these basic molecules.[615] If they accumulate in the lysosome and are trapped there for extended periods, the irreversible reaction with lysosomal proteases becomes favored.

Own contributions: Heterologous expression of rhodesain, preliminary characterization of the in vitro affinity and inhibition kinetics of **11**, **14**, and **15** to validate the approach and verify the results, applying MALDI-TOF-MS to prove the covalency of binding for all final inhibitors, writing parts and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of in vitro affinity, inhibition kinetics, and selectivity, reversibility assessment by dilution and dialysis assay, non-covalent and covalent docking, molecular dynamics simulation, writing and editing of the manuscript.

This work was published in Bioorganic Chemistry.

Article reprinted with permission of *Bioorganic Chemistry* **2024**, 153, 107830, “Rhodesain inhibitors on the edge of reversibility-irreversibility”, Copyright © 2024 Elsevier. Published by Elsevier.

The following publication quoted within “” (pages 117–137) is the same as the manuscript cited here. The appended Supporting Information represents an abridged version. The full version can be accessed online at doi:10.1016/j.bioorg.2024.107830.

“

Bioorganic Chemistry 153 (2024) 107830

Contents lists available at ScienceDirect

Bioorganic Chemistry

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

Rhodesain inhibitors on the edge of reversibility-irreversibility

Laura Agost-Beltrán [a], Collin Zimmer [b], Hans Joachim Räder [c], Christian Kersten [b,d], Tanja Schirmeister [b], Santiago Rodríguez [a], Florenci V. González [a,*]

[a] Departament de Química Inorgànica i Orgànica, Universitat Jaume I, 12071 Castelló de la Plana, Spain
[b] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany
[c] Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany
[d] Institute for Quantitative and Computational Biosciences, Johannes Gutenberg-University, BioZentrum I, Hanns-Dieter-Hüsch.Weg 15, 55128 Mainz, Germany

ARTICLE INFO

Keywords:
Cysteine proteases
Rhodesain
Michael acceptors
Covalent reversible inhibitors

ABSTRACT

A comparative study of Michael acceptor and keto-Michael acceptor inhibitors of the cysteine protease rhodesain has been performed. Five new inhibitors have been prepared bearing the peptide structure of the known cysteine protease inhibitor **K11777** and differing on the warhead. For the preparation of the Michael acceptor warhead, a Horner-Wadsworth-Emmons reaction was used. In the synthetic routes of the keto-Michael acceptor warheads, keto-enoate and keto-vinyl sulfone, a metathesis reaction and a radical sulfonylation were the key steps, respectively. Interestingly, keto-Michael acceptors inhibited rhodesain through a dual mode of action, showing reversibility at low inhibitor concentrations and irreversibility at high inhibitor concentrations.

1. Introduction

Neglected tropical diseases (NTDs) are a diverse group of infectious diseases that mainly affect the impoverished communities without basic sanitation in tropical countries, although some have a much larger geographical distribution. The number of people affected by NTDs is higher than 1 billion and 1.6 billion people require NTD interventions (both preventive and curative) [1].

In particular, Human African Trypanosomiasis (HAT) or sleeping sickness is one of the diseases caused by protozoa of the Trypanosomatid family, together with Chagas disease (Trypanosoma cruzi) and Leishmaniasis. HAT is endemic in Sub-Saharan Africa and it is caused by two subspecies of the protozoan *Trypanosoma brucei*: *T. b. gambiense* and *T. b. rhodesiense* [2]. It is a vector-borne disease transmitted by tse-tse flies. Although many efforts by non-profit agencies and World Health Organization have been made and the number of cases and casualties related to the disease have decreased considerably, millions of people still remain at risk of infection [3]. The disease is fatal when untreated and the current treatments, like nifurtimox-eflornithine combination therapy and fexinidazole [4] cause serious side effects and resistance. Therefore, there is an urgent need to develop new alternatives to mitigate the devastating effects of the disease.

The Target-Based Drug Design is an approach to identify new compounds against NTDs [5]. Among the targets, the cysteine protease rhodesain is a lysosomal protease essential for the development of *T. brucei rhodesiense* and has been identified as a target for the search of new drugs against the disease [6–8]. The mechanism of action of rhodesain resembles that of other cysteine proteases belonging to the papain family with the participation of the catalytic triad Cys25/His162/Asn182.

Vinyl sulfone **K11777** [9] is a well-known irreversible inhibitor of cysteine proteases like rhodesain. The conjugate addition of the thiolate of the cysteine to the double bond results in a carbanion that is subsequently protonated by the histidine (Scheme 1A), originating an irreversible enzyme-inhibitor (E-I) complex.

However, concerns about toxicity of irreversible inhibitors have risen and, consequently, new strategies are being developed to obtain new reversible inhibitors. For example, halogenated vinyl sulfones were reported as an alternative to irreversible vinyl sulfones [10]. Also, previous reports of inhibitors having carbonyl warheads were active against

Abbreviations: HAT, Human African trypanosomiasis; e.g., exempli gratia (for example); CatL, cathepsin L; catB, cathepsin B; NMR, Nuclear Magnetic Resonance; MALDI, Matrix Assisted Laser Desorption/Ionization; TOF, Time of Flight; MS, Mass Spectrometry; R_f, retention factor; IR, Infrared; r.t., room temperature; min, minutes; h, hours; EWG, electron withdrawing group; pdb, protein data bank; Trp, tryptophan; Cys, cysteine; His, histidine; Asn, asparagine; Asp, aspartate; Gly, glycine; Gln, glutamine; Mu, morpholinyl urea; Cha, L-cyclohexylalanine; Ser, serine.

* Corresponding author.
E-mail address: fgonzale@uji.es (F.V. González).

https://doi.org/10.1016/j.bioorg.2024.107830
Received 1 April 2024; Received in revised form 12 September 2024; Accepted 16 September 2024
Available online 19 September 2024
0045-2068/© 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

cysteine proteases [11].

In this work, we have designed and synthesized keto-vinyl sulfones and studied their activity as inhibitors of cysteine proteases. Interestingly, the introduction of a ketone group conjugated with the vinyl sulfone moiety changed the irreversible character of the vinyl sulfones into a dual reversible/irreversible mechanism. The new compounds follow a reversible behavior at low concentrations of the inhibitor, but they act irreversibly at high concentrations. Thus, these results suggest the introduction of a ketone group as an attractive strategy to modify the mode of action of Michael acceptors.

2. Results and discussion

2.1. Design of inhibitors

The chemical structure of a peptidyl protease inhibitor has two parts: the recognition part, consisting of a peptide backbone specifically recognized by the sites of the active center of the protein, and the reactive part, known as warhead, which is an electrophilic group that reacts with the nucleophile of the enzyme, e.g. a thiol in case of cysteine proteases [12]. The warhead controls the inhibition kinetics [13]: in case of covalent reversible inhibitors, it reacts with the nucleophilic thiol and forms a reversible complex with the enzyme that resembles the intermediate thioorthoamide of the native catalytic cycle during substrate cleavage (Scheme 1B). On the contrary, in case of irreversible inhibitors, the warhead reacts with the thiol of the cysteine by forming a stable product, as shown with inhibitor **K11777** in Scheme 1A.

During our studies for the design of new warheads, we became interested in the preparation and study of inhibitors having keto-Michael acceptor warheads as a combination of irreversible Michael acceptors and the reversible attack on the carbonyl of the natural substrate (Fig. 1). In the present work, we performed a comparative study of six inhibitors having the same recognition part, the one from the known inhibitor **K11777**, but differing on the warhead. Two of these inhibitors (unsaturated ester **11** and vinyl sulfone **K11777**) are Michael acceptors. Other two of them, γ-keto-α,β-unsaturated ester **15** and γ-keto-α,β-unsaturated sulfone **19**, are keto-Michael acceptors. Also, γ-hydroxy-α,β-unsaturated ester **14** and γ-hydroxy-α,β-unsaturated sulfone **18** bearing a Michael-acceptor moiety and a secondary alcohol were prepared and tested.

2.2. Synthesis

Inhibitors were synthesized as indicated in Scheme 2.

Carboxylic acid **3** was prepared in a three-step sequence starting from phenylalanine as previously reported [10,14]. Weinreb amide **5**, readily prepared from L-homophenylalanine, was deprotected and coupled with carboxylic acid **3** using standard peptide coupling conditions to afford compound **9**. Then, intermediate **9** was reduced to aldehyde **10**, which was subsequently reacted with triethyl phosphonoacetate via Horner-Wadsworth-Emmons reaction, yielding vinyl ester **11** as an (*E*)-isomer.

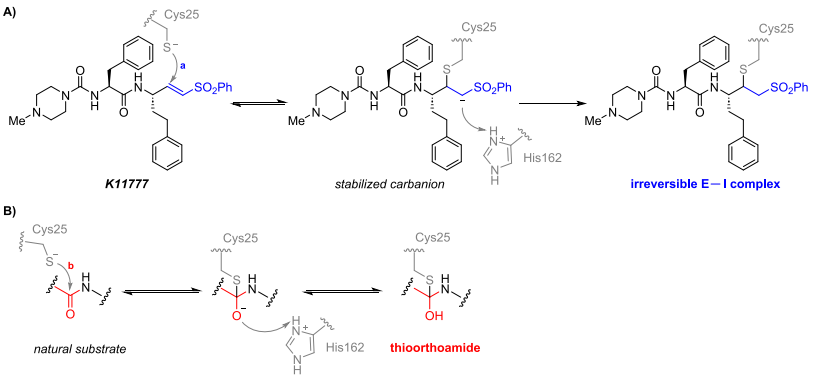
For the preparation of inhibitor **15**, *tert*-butoxy carbonyl homophenylalaninal **6** was reacted with vinylmagnesium bromide to give vinyl alcohol **7** as a 1/0.61 mixture of diastereoisomers, being the *anti* isomer the major one, as predicted by the Felkin-Anh model [15,16]. Then, compound **7** was submitted to metathesis reaction with ethyl acrylate using 2nd generation Hoveyda-Grubb's catalyst. The reaction afforded the γ-hydroxy-α,β-unsaturated ester **12** in good yield (68 %). Then, the metathesis product **12** was deprotected and coupled with building block **3** giving γ-hydroxy-α,β-unsaturated ester **14**, that was oxidized to the keto-vinyl ester **15**.

The synthesis of inhibitor **19** was accomplished by radical alkene sulfenylation [17] of vinyl alcohol **7** using phenyl sulfinate in the presence of iodine and sodium acetate. The resulting compound **16** was then deprotected and coupled with **3**. The obtained alcohol vinyl sulfone **18** was finally oxidized using Dess-Martin periodinane to yield keto-vinyl sulfone **19**.

Inhibitor **K11777** was prepared as previously described [10].

a) Bis(trichloromethyl) carbonate, $NaHCO_3$ (aq), CH_2Cl_2, ice bath, 15 min, 97 %. b) 1-Methylpiperazine, CH_2Cl_2, r.t., 2 h, 94 %. c) Ammonium formate, ethanol, Pd/C, r.t., 2 h, 100 %. d) Di-*tert*-butyl dicarbonate, $NaHCO_3$ (aq), THF, r.t., overnight, 100 %. e) DCC, HOBt·xH_2O, DIPEA, *N,O*-dimethylhydroxylamine hydrochloride, CH_2Cl_2, ice bath to r.t., overnight, 83 %. f) $LiAlH_4$, dry THF, ice bath, 2 h, 87 %. g) Vinylmagnesium bromide (1 M in THF), dry zinc chloride, dry THF, −78 °C to 0 °C, overnight, then Rochelle salt (aq), 52 %. h) TFA, CH_2Cl_2, ice bath, 1 h 5 min. i) DCC, HOBt·xH_2O, DIPEA, **3**, CH_2Cl_2, ice bath to r.t., overnight, 66 % (two steps). j) $LiAlH_4$, dry THF, ice bath, 2 h, 57 %. k) NaH, triethyl phosphonoacetate, dry THF, −78 °C, 10 min, then **10**, −78 °C to 0 °C, overnight, 84 %. l) Ethyl acrylate, 2nd generation Hoveyda-Grubb's catalyst, CH_2Cl_2, reflux at 50 °C, 5 h, 68 %. m) TFA, CH_2Cl_2, ice bath, 1 h 5 min. n) DCC, HOBt·xH_2O, DIPEA, **3**, CH_2Cl_2, ice bath to r.t., overnight, 50 % (two steps). o) DMP, CH_2Cl_2, −5 °C, overnight, 73 %. p) Phenyl sulfinate, iodine, sodium acetate, acetonitrile, reflux, 41 h, 40 %. q) TFA, CH_2Cl_2, ice bath, 1 h 5 min. r) DCC, HOBt·xH_2O, DIPEA, **3**, CH_2Cl_2, ice bath to r.t., overnight, 42 % (two steps). s) DMP, CH_2Cl_2, −5 °C, overnight, 100 %.

Enzyme assays. Compounds **11**, **14**, **15**, **18**, **19** and **K11777** were tested against rhodesain (Table 1) and against human cathepsins B and L



Scheme 1. A) Covalent irreversible inhibition of **K11777**. B) Cleavage of the natural substrate.

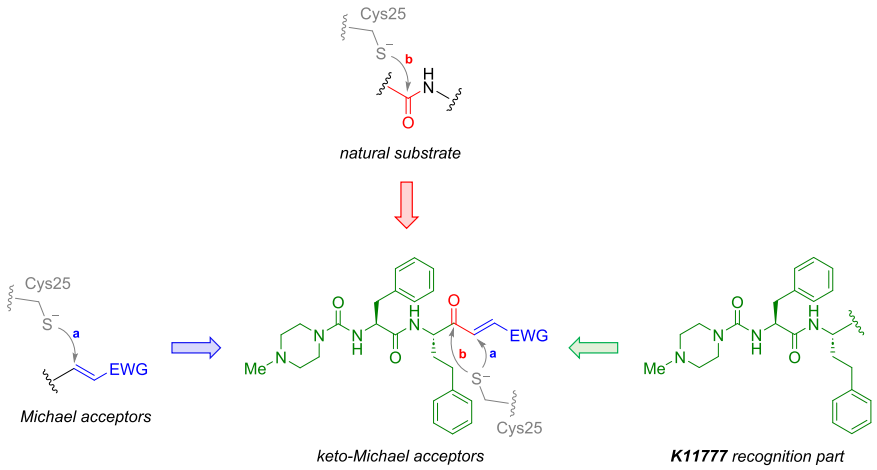
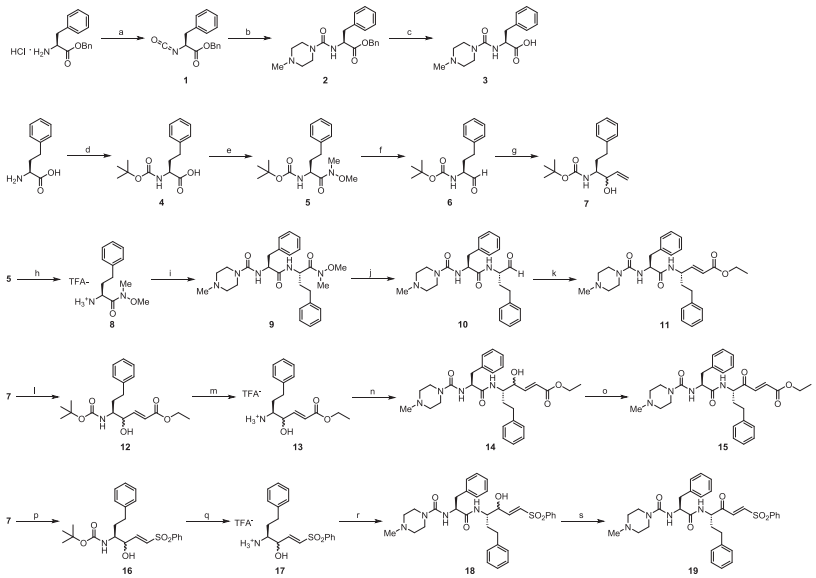


Fig. 1. Rational design of keto-Michael acceptors.



Scheme 2. Synthesis of inhibitors.

for selectivity (Table 2). Michael acceptors **11** and **K11777**, as well as hydroxy Michael acceptor **14**, inhibited rhodesain in an irreversible way. Compound **18** was a poor inhibitor showing less than 50 % inhibition at 20 μM of the inhibitor. Contrastingly, keto Michael acceptors **15** and **19** were potent inhibitors of rhodesain displaying a dual mode of action: they showed reversibility at low concentrations and acted irreversibly at high concentrations (Table 1).

When comparing the same structure with an alcohol group or with a ketone group (**14** vs **15**, **18** vs **19**), the presence of the ketone clearly increases the inhibitory potency, as it can be seen when comparing the k_{2nd} values. The ketone would favor the formation of a thiohemiketal intermediate (mechanism I from Scheme 3 and Fig. 1) similarly to the thioorthoamide of natural substrate (Scheme 1B). Also, the hydroxyl group can possibly force a different conformation of the ligand in the active center placing the electrophilic carbon further away from the thiolate of the cysteine residue. It can form hydrogen bonds with other amino acids that prevent the proper entrance of the inhibitor in the active site. Another explanation for the observed lower activity of the alcohols may be the lower electrophilicity of the β-carbon due to the lack of electron withdrawal as compared to its keto-Michael acceptor counterpart.

When comparing the irreversible kinetic parameters of esters **11** and

Table 1
Inhibition data against rhodesain.

		Rhodesain				
		Irreversible			Reversible	
Kinetic behavior	Compound	k_{inact} (s^{-1})	K_i (μM)	k_{2nd} ($M^{-1}{\cdot}s^{-1}$)	IC_{50} (μM)	K_i (μM)
Irreversible	**11**	0.0094 ± 0.0005	0.1982 ± 0.0391	(4.87 ± 0.87) ·10^4	–	–
Irreversible	**14**	0.0010 ± 0.0001	0.0743 ± 0.0203	(1.35 ± 0.23) ·10^4	–	–
Dual mode	**15**	0.0038 ± 0.0017	0.0485 ± 0.0347	(7.76 ± 2.47) ·10^4	0.391 ± 0.028	0.030 ± 0.002
Irreversible	**K11777** [a]	0.0132	0.020	6.6·10^5	–	–
Irreversible	**18**	n.d.[b]	n.d.[b]	n.d.[b]	–	–
Dual mode	**19**	0.0021 ± 0.0004	0.0123 ± 0.0037	(1.73 ± 0.13) ·10^5	0.509 ± 0.033	0.039 ± 0.003

Values are mean values of at least three measurements.
[a] Previously published data [8].
[b] Less than 50 % inhibition at 20 μM of the inhibitor; n.d. = not determined.

Table 2
Inhibition data against CatL and CatB.

Compound	K_i CatL (μM)	K_i CatB (μM)	SI CatL[a]	SI CatB[b]
11	0.210 ± 0.073 (i)	0.814 ± 0.080 (i)	1.1	4.1
14	1.283 ± 0.007 (i)	0.881 ± 0.114 (i)	17.3	11.9
15	0.179 ± 0.009 (r)	0.134 ± 0.011 (r)	6.0	4.5
K11777 [c]	0.080 (i)	0.390 (i)	4	19.5
18	n.d.	n.d.	–	–
19	0.088 ± 0.003 (r)	0.826 ± 0.062 (r)	2.3	21.2

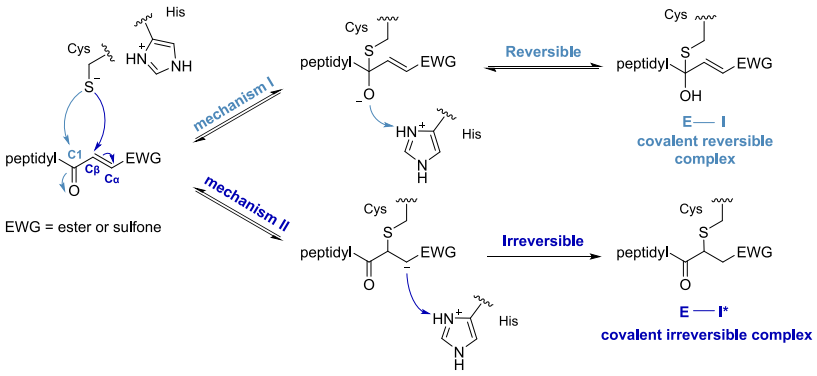
n.d.: not determined; (i) K_i determined using irreversible kinetics equations; (r) K_i determined using reversible kinetics equations.
Values are mean values of at least two independent measurements.
[a] SI CatL=Selectivity Index for CatL = K_i (CatL)/K_i (rhodesain).
[b] SI CatB = Selectivity Index for CatB = K_i (CatB)/K_i (rhodesain).
[c] Previously published data [8].

15 containing the ethyl ester in the P1′ position, the K_i decreases 4-fold when the extra carbonyl is present (compound **15**), which is translated in an increase of the affinity towards the protein. However, compound **15** shows a lower k_{inact} value, meaning that the covalent irreversible reaction is slower. This fact supports the theory that keto-Michael acceptors present a dual mode of action (see Scheme 3), as when only the Michael acceptor group is present, the attack exclusively takes place on the β-position of the double bond giving irreversible inhibition, whereas when the carbonyl is also present, the Michael addition is competing with the reversible addition on the carbonyl, resulting in the deceleration of the irreversible reaction and in a dual mode of action. This decrease in the k_{inact} value can also be related to the fact that the β position to be attacked for the irreversible reaction in compound **15** is one bond away from the position to be attacked in the endogenous substrate, thus the distance between nucleophile and electrophile is higher. These two phenomena (decrease in both K_i and k_{inact}) result in k_{2nd} values of the same order of magnitude for both compounds.

When comparing the irreversible kinetic parameters of compounds **K11777** and **19**, containing the phenyl sulfone in the *P*1′ position, there is a 2-fold decrease on the K_i value when the carbonyl is present (compound **19**), which means that its affinity for the protein is slightly higher. Also, the previously described effect for k_{inact} is observed, so the same explanation could be applied in this case. These two phenomena result again in k_{2nd} values in the same order of magnitude for both compounds.

Finally, when comparing both families of compounds (esters vs. sulfones), it can be stated that the presence of the phenyl sulfone moiety in the *P*1′ site significantly increases the activity of the inhibitors, probably because it extends further into the S1′ pocket, forming interactions that the ester moiety cannot provide (see Docking section).

Compounds **11**, **14**, **15** and **19** gave sub-micromolar activity for cathepsins B and L (Table 2) being keto-vinyl sulfone **19** the most active one with similar values to **K11777**. It is also worth mentioning that the selectivity towards human cathepsin L increases almost 6-fold when comparing compound **15** with compound **11** and is similar in both compounds for human cathepsin B, with the inhibition of rhodesain always being favored (around 4-fold) (Table 2). The selectivity of **14** towards rhodesain is surprisingly good, but its low potency does not make it such an interesting compound. When comparing compound **19** with **K11777**, the affinity towards CatL increases, thus decreasing the selectivity for rhodesain (from 4-fold for **K11777** to 2.3-fold for compound **19**) although there is still a preferential inhibition of the main target, and the affinity towards CatB is still very low in comparison to rhodesain, providing a selective index of about 20-fold.



Scheme 3. Dual mode of action of keto-Michael acceptors.

The progression curves against rhodesain for vinyl ester **11** showed a clear time-dependency while the time-dependency for alcohol vinyl ester **14**, keto-vinyl ester **15** and keto-vinyl sulfone **19** was not so pronounced (Figs. S1 to S4). Thus, for elucidating the mode of action of these inhibitors, dilution and dialysis experiments were performed against rhodesain using previously described methods [18,19].

In the dilution assay, the enzyme was incubated with an excess of inhibitor (10-fold the IC_{50} concentration) to ensure total inhibition. Then, the mixture was diluted 100-fold and the fluorescence emission was measured. DMSO was used as blank control, a previously reported nitrile inhibitor (Mu-Cha-Ser(Bn)-CN) [20] as reversible control and **K11777** as irreversible control. In case of irreversible binding, no recovery should be observed whereas for reversible binding, the enzyme activity should be recovered after dilution. The results are shown in fractional activity of uninhibited rhodesain in Fig. 2.

As it was expected, the irreversible control did not recover any activity and the reversible control recovered almost all of it (96 %). Regarding the compounds of interest, alcohol vinyl ester **14** and alcohol vinyl sulfone **18** recovered to an activity of 50 % and 61 %, respectively, whereas vinyl ester **11**, keto-vinyl ester **15** and keto-vinyl sulfone **19** did not significantly recover any. Even though compounds **14** and **18** behaved as reversible inhibitors, due to their very low reactivity (Table 1), it could happen that they are irreversible inhibitors which slowly react with the enzyme and then, at the point of dilution and also afterwards, there is still rhodesain activity that can produce fluorescence. On the other hand, compounds **11**, **15** and **19** behaved as irreversible inhibitors.

To further study the kinetic behavior of the most active compounds (**15** and **19**), a dialysis assay was performed using a custom-built dialysis chamber. DMSO was used as blank control, nitrile inhibitor [20] as reversible control and compound **K11777** as irreversible control. The samples were dialyzed for 30 min and the fluorescence emission of samples drawn at several time points was recorded after addition of substrate.

For irreversible inhibitors, fluorescence should not be recovered whereas for reversible inhibitors, a recovery of the enzyme activity should be observed due to the diffusion of the inhibitor dissociated from the enzyme across the membrane. The results are shown in fractional activity of uninhibited rhodesain at different times in Fig. 3.

As expected, the irreversible control (**K11777**) did not recover any activity over time and the reversible control recovered to an activity of 63 % after 30 min. Regarding the compounds of interest, both keto-vinyl ester **15** and keto-vinyl sulfone **19** did not significantly recover any activity, which confirms its irreversible behavior. The slight recovery of the activity observed in compounds **15** and **19** could be related to its dual reversible-irreversible character, since in the short period between preparation of solutions and start of the dialysis, a fraction of enzyme is expected to always be reversibly inhibited and thus activity is rescuable by dilution.

Despite these results indicating a covalent irreversible mode of action of keto-Michael acceptors **15** and **19**, the time-dependency on the progression curves was not very pronounced, especially for low concentrations of the inhibitors. Therefore, it can be hypothesized that a dual mode of action takes place, as it was previously demonstrated by computational studies with keto-enoates (Scheme 3) [21].

Due to this dual behavior, the progression curves show the kinetics resulting from two different reactions: the initial reversible attack on the carbonyl, as the original attack position in the endogenous substrate (mechanism I) and the irreversible Michael addition on the double bond conjugated with the EWG, i.e. ester or sulfone (mechanism II). This dual mode mechanism would be at the origin of inhibitors containing Michael acceptors reported by other groups [22].

In reference to the calculations performed for the determination of kinetic parameters, compound **11** showed irreversible behavior in both the progression curves and the dilution assay; thus, the kinetic parameters were calculated using the irreversible kinetics equations (detailed in part **4.3.** of the SI). Compounds **14** and **18** recovered fluorescence in the dilution assay, but this recovery was attributed to their low reactivity, so they were considered to be irreversible and the parameters corresponding to their kinetics were calculated using the irreversible kinetics equations. Compounds **15** and **19** did not show time-dependency in the progression curves, especially at low concentrations, but behaved as irreversible compounds in the dilution and dialysis assays; therefore, in these two cases, the kinetic parameters were

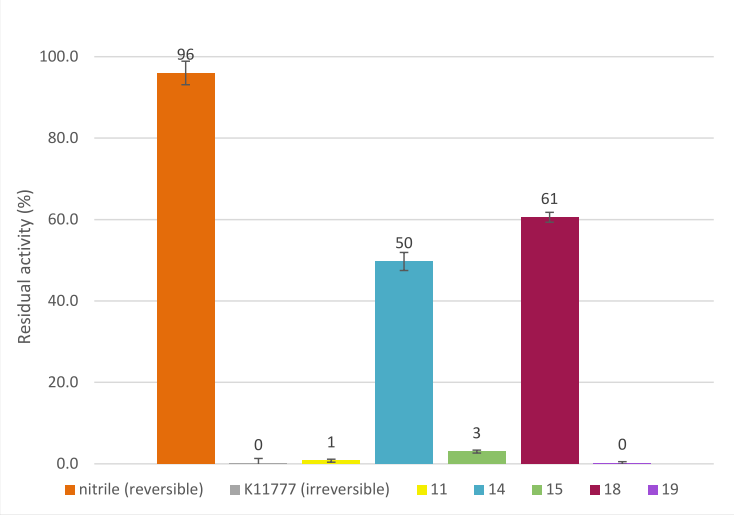


Fig. 2. Dilution assay with rhodesain. Nitrile (orange), **K11777** (grey), **11** (yellow), **14** (light blue), **15** (light green), **18** (dark red) and **19** (purple). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

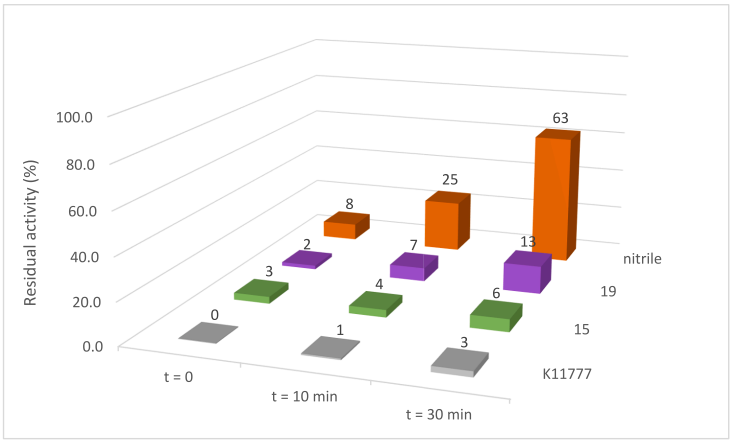


Fig. 3. Dialysis assay with rhodesain. Nitrile (orange), **K11777** (grey), **15** (light green) and **19** (purple). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

obtained using both reversible and irreversible kinetics equations. As it can be seen in Table 1, the K_i values for **15** and **19** calculated with both the reversible and irreversible approaches match quite well. This fact validates the K_i values calculated with the irreversible equations, even though the corresponding errors are higher than 10 % due to the bad fit of k_{obs} vs. inhibitor concentration caused by the lack of curvature of the progression curves.

MS Analysis. To further characterize the interaction between rhodesain and the ligands, MALDI-TOF Mass Spectrometry was performed. In this experiment, **K11777** was used as a covalent control. A covalent protein–ligand adduct was found for **K11777** and compounds **11**, **14**, **15** and **19** (Fig. S19–S27). No adduct was found for alcohol **18**. These findings suggest that the cysteine of the active center covalently reacts with the electrophilic warheads of the cited compounds.

Non-covalent docking. Non-covalent docking of inhibitors **11** (Fig. 4), **15** (Fig. 5) and **19** (Fig. 6) provided rational results, as all the amino acid residues were properly addressed to the binding pockets and the electrophilic carbons were predicted to be near to the catalytic triad (Cys25, His162, and Asn182), especially to the nucleophilic sulfur atom of Cys25, suggesting a high probability of nucleophilic attack and covalent complex formation. A summary of docking parameters is shown in Table S3.

It is worth mentioning that in both inhibitors **15** and **19** the distance to the carbonyl carbon is lower than the distance to the β-carbon (Table S3), which agrees with the addition to the carbonyl group being the first reaction to occur.

The docking poses from **K11777** and **11** were overlapped (Fig. 7) and a high similarity was observed. Even though the atoms forming the warhead are different, they position themselves in a very similar way in which one of the sulfone oxygens from **K11777** overlaps with the ethyl ester oxygen of **11** and both can form a hydrogen bond interaction with Trp184.

The docking poses from **15** and **19** were also overlapped (Fig. 8) and analogous poses were observed. In the first place, both ketone groups are positioned in the same direction. Moreover, one of the sulfone oxygens from **19** overlaps with the oxygen of the ester's carbonyl from compound **15**, allowing the formation of a hydrogen bond with Gln19. Also, the ethyl ester oxygen from **15** is situated towards Trp184, which could also result in a hydrogen bond interaction. Finally, the position of the sulfone's aromatic ring suggests a possible π-π interaction with Trp184. Then, the affinity increase observed in compound **19** with respect to inhibitor **K11777** could be attributed to the fact that the keto-vinyl sulfone moiety can extend further into the S1′ pocket being able to establish an extra interaction with Trp184. This new interaction has been previously observed in fluorovinylsulfonates [23].

Covalent docking. To further support these findings, covalent docking was performed with compounds **K11777** as a positive control, **11**, **15** and **19** (Table S2). The covalent re-docking of **K11777** well reproduced the crystallographic binding mode with an RMSD of 1.4 Å, indicating a suitable docking protocol. For **11**, **15** and **19** (Figs. 9–10), predicted binding modes show similar interactions with the formation of a thiohemiacetal (score: −16.92) being slightly favored over the Michael

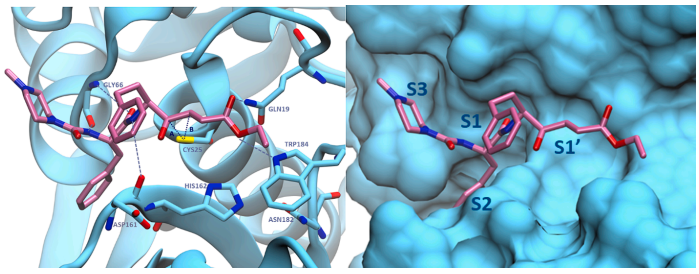


Fig. 4. Non-covalent docking of compound **11**. Distance from S (Cys25) to warhead: A = 4.60 Å.

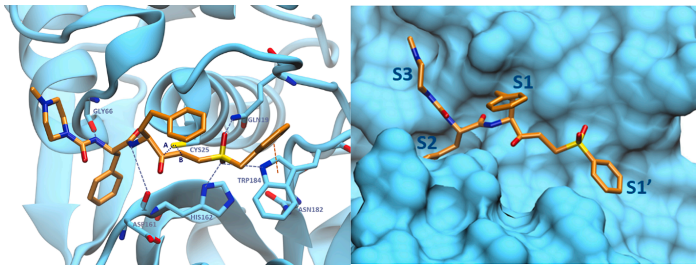


Fig. 5. Non-covalent docking of compound **15**. Distances from S (Cys25) to warhead: A = 3.92 Å; B = 4.31 Å.

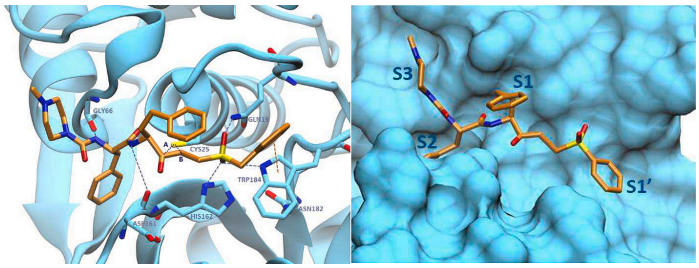


Fig. 6. Non-covalent docking of compound **19**. Distances from S (Cys25) to warhead: A = 4.23 Å; B = 4.97 Å.

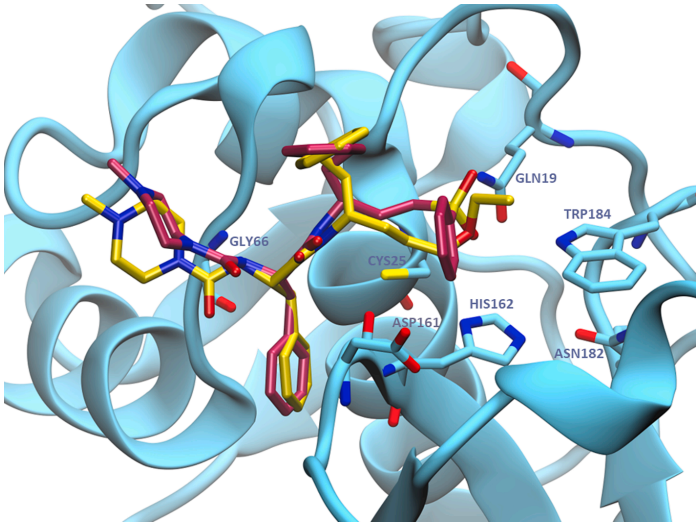


Fig. 7. Overlay of non-covalent docking poses of compounds **K11777** and **11**.

addition reaction (score: −14.06) (Table S2).

Molecular dynamic simulations. While the reaction mechanisms of vinyl sulfones and sulfonates [10,23], and keto-enoates [21] were detailly described previously, keto-vinyl sulfone **19** was followed up by MD simulations. Three independent simulations were performed starting from the non-covalent docking pose (Fig. 6) and untethered structures re-generated from the covalent docking poses for hemithioacetal formation (mechanism I) and Michael addition (mechanism II) thus generating pre-reaction, non-covalent complex structures. All simulations show a very stable rhodesain with backbone RMSD-deviation of around 1 Å compared to the starting structure (Fig. S29). Ligand binding modes quickly equilibrate as well with 2–4 Å deviation to starting structure. Except for the simulation starting from the untethered hemithioacetal, both nucleophilic carbon atoms (C_β of the vinyl sulfone

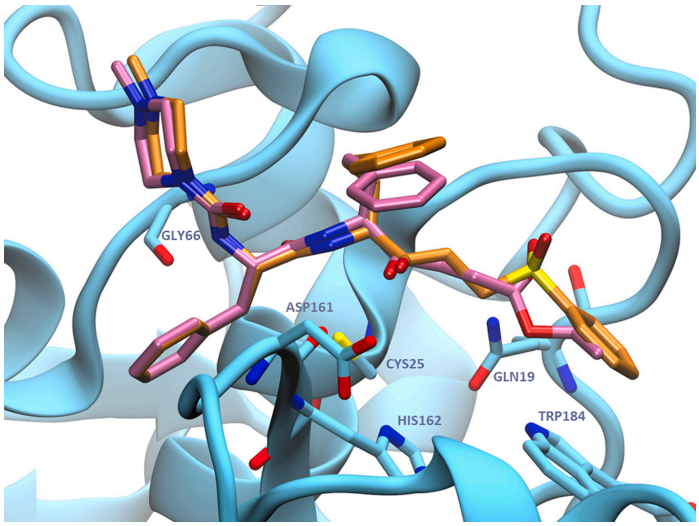


Fig. 8. Overlay of non-covalent docking poses of compounds **15** and **19**.

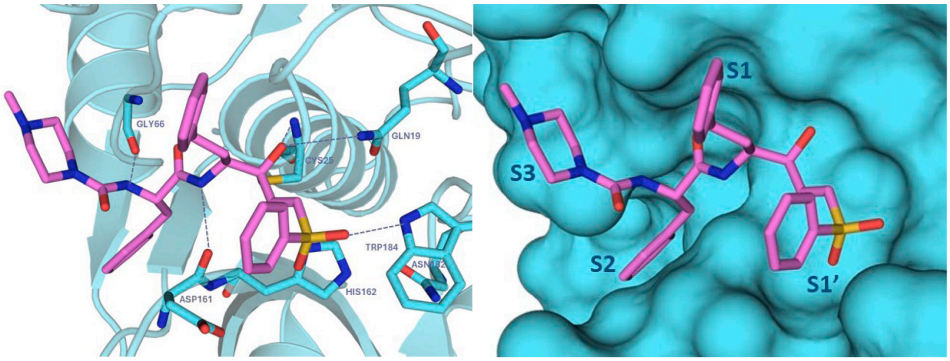


Fig. 9. Covalent docking of compound **19** forming a thiohemiacetal with Cys25 (mechanism I).

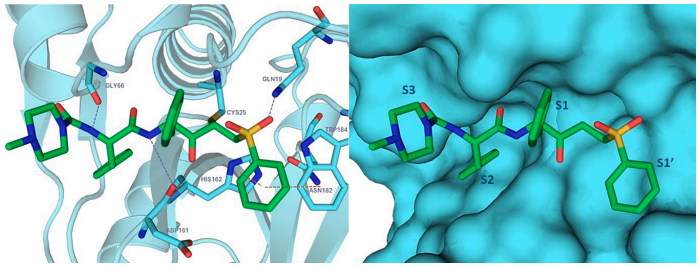


Fig. 10. Covalent docking of compound **19** after Michael addition of Cys25 over the double bond (mechanism II).

moiety and the carbonyl C) stay in proximity to the Cys-25 thiolate with a distance of 3–4 Å (Fig. S29B and D), comparable to C_β of K11777 (Fig. S29 A). Similarly, the thiolate is vastly perpendicular or to the plane of the electrophilic carbon atoms or with angles slightly higher than 90° (Fig. S30). Taken together, both reaction mechanisms (Scheme 3) are possible from geometric analysis due to proximity and angles ≥90° [24], Using Molecular Mechanic Generalized Born Surface Area (MM/GBSA) continuum solvation models, the binding free energies of the different poses were estimated (Table 3). The pose that equilibrates further away from the catalytic Cys-25 (Fig. S29C) has the lowest binding free energy of −37.6 kcal/mol. All other complexes are more favorable and in a similar affinity range between −46.2 and −50.8 kcal/mol. During the simulation starting from the untethered covalent docking pose (mechanism II) shows a conformational change of the ligand after around 10 ns (Fig. S29B) in which the carbonyl oxygen atom orients into the oxyanion hole and after 25 ns adapts a very stable conformation in which both electrophilic carbon atoms are close to Cys-25 thiolate with the carbonyl carbon atom being slightly closer. For this complex (25–50 ns, Table 3), the MM/GBSA predicted binding free energy is −50.8 kcal/mol and hence slightly lower compared to the initial binding mode where the distance is closer to the C_β (−46.2 kcal/mol). For the stable binding mode found in the simulations starting from the non-covalent docking pose, a MM/GBSA score of −50.4 kcal/mol was calculated, but with the thiolate being closer to C_β. These findings indicate that there is not large energetical difference between poses with closer proximity between S^- and either the carbonyl C or C_β. Assuming this allows for reactions with both sites and mechanism I being reversible, over time and/or at high concentrations, the irreversible Michael addition product (mechanism II) might accumulate resulting in the observed time-dependent irreversible binding mode of action.

3. Conclusions

Five new inhibitors have been prepared with the common peptide structure of the known inhibitor **K11777** but differing on the warhead: one of the inhibitors is the unsaturated ester **11** which along with vinyl sulfone **K11777** are Michael acceptors; other two are keto-Michael acceptors: γ-keto-α,β-unsaturated ester **15** and γ-keto-α,β-unsaturated sulfone **19**; and γ-hydroxy-α,β-unsaturated ester **14** and γ-hydroxy-α,β-unsaturated sulfone **18** were also prepared.

A comparative study in terms of *in vitro* activity against the cysteine proteases rhodesain, cathepsin B, and cathepsin L has been carried out. Interestingly, keto-Michael acceptors inhibited rhodesain through a dual mode of action: they are reversible at low inhibitor concentration and irreversible at high inhibitor concentration. In contrast, Michael acceptors without a keto group are irreversible inhibitors. Non-covalent and covalent docking studies explain the dual mode of action of the keto-Michael inhibitors with the ketone and double bond positioned in a close distance to the thiolate of Cys25 and suggest the formation of non-polar interactions between inhibitor **19** and residue Trp184 at S1′ site, unlike ester **15**. Also, the geometric features from the lowest energy poses derived from the molecular dynamic simulations are in agreement to a dual mode of action of the keto-Michael acceptor inhibitors.

This work demonstrates that keto-Michael acceptor follow a reversible/irreversible mechanism as compared to the irreversible mode of Michael acceptor counterparts. Further biological studies with these compounds are underway in our lab and will be reported in the near future.

Table 3
MM/GBSA results for different non-covalent binding modes of compound **19** and reference ligand **K11777**. Calculations were performed for 50 ns production MD omitting initial 1 ns equilibration. Values are average ± standard deviation.

Starting Structure	Time frame	MM/GBSA (kcal/mol)
Untethered rhodesain-K11777 complex (pdb: 2P7U)	0–50 ns	−46.14 ± 5.15
Rhodesain-**19** complex (untethered from covalent docking, mechanism II)	0–50 ns	−48.76 ± 4.91
	0–25 ns (S^- closer to C_β)	−46.20 ± 5.00
	25–50 ns (S^- closer to carbonyl)	−50.82 ± 3.86
Rhodesain-**19** complex (untethered from covalent docking, mechanism I)	0–50 ns (overall large distance between S^- and electrophiles)	−37.64 ± 5.13
Rhodesain-**19** complex (non-covalent docking)	0–50 ns (S^- closer to C_β)	−50.40 ± 3.73

4. Experimental section

4.1. Synthesis

General information. Unless otherwise specified, all reactions were carried out under nitrogen atmosphere with magnetic stirring. All solvents and reagents were obtained from commercial sources and were purified according to standard procedures before use [25]. ^1H and ^{13}C NMR spectra were measured in $CDCl_3$ solution (^1H: singlet at 7.27 ppm; ^{13}C: triplet at 77.0 ppm) or CD_3OD solution (^1H: singlet at 4.87 ppm and pentet at 3.31 ppm; ^{13}C: septet at 49.0 ppm) at 30 °C on a 300 or 400 MHz NMR spectrometer. IR spectra were recorded on an ATR spectrometer. Mass spectra were measured in a QTOF I (quadrupole–hexapole TOF) mass spectrometer with an orthogonal Z-spray-electrospray interface (ESI). Rotary power was measured with a polarimeter. EM Science Silica Gel 60 was used for liquid chromatography, while TLC was performed with precoated plates (Kieselgel 60, F254, 0.25 mm).

Benzyl (S)-2-isocyanato-3-phenylpropanoate (1). Hydrochloride salt (*S*)-phenylalanine benzyl ester (2.180 g, 7.470 mmol, 1 eq) was dissolved in a mixture of saturated $NaHCO_3$ aqueous solution and CH_2Cl_2 (28.5 mL: 28.5 mL). Triphosgene (0.739 g, 2.490 mmol, 0.33 eq) was added under an ice bath. The reaction was stirred for 15 min. Then, organic and aqueous phases were separated, and the aqueous phase was extracted with CH_2Cl_2 (3 × 20 mL). The organic phases were combined, washed with brine (2 × 20 mL), and dried over $MgSO_4$. The solvent was removed under reduced pressure to give benzyl (*S*)-2-isocyanato-3-phenylpropanoate as a white solid (2.030 g, Yield = 97 %).

R_f (Hex:AcOEt, 1:1) = 0.80.

^1H NMR (400 MHz, $CDCl_3$) δ 7.40–7.25 (m, 8H), 7.16–7.09 (m, 2H), 5.26–5.18 (m, 2H), 4.29 (dd, J = 7.5, 4.8 Hz, 1H), 3.15 (dd, J = 13.8, 4.8 Hz, 1H), 3.04 (dd, J = 13.8, 7.5 Hz, 1H) ppm.

^{13}C NMR (75 MHz, $CDCl_3$) δ 170.5, 135.5, 134.9, 129.5, 128.9, 128.9, 128.8, 127.6, 68.3, 58.7, 40.1 ppm.

HRMS (ESI) *m/z* calculated for $C_{17}H_{15}NO_3Na^+$ $[M+Na]^+$: 304.0944, found: 304.0948.

IR (ATR) 3064, 3030, 2974, 2207, 1722, 1566, 1495, 1454, 1379 cm^{-1}.

m.p. 127–129 °C.

Benzyl (4-methylpiperazine-1-carbonyl)-L-phenylalaninate (2). **1** (2.030 g, 7.216 mmol, 1 eq) was dissolved in CH_2Cl_2 (34 mL). 1-Methylpiperazine (831 μL, 7.505 mmol, 1.04 eq) was added and the mixture was stirred overnight. The solvent was removed under reduced pressure and the crude benzyl (4-methylpiperazine-1-carbonyl)-L-phenylalaninate was obtained as a transparent oil (m = 2.69 g, Yield = 94 %) and used in the next step without further purification.

R_f (Hex:AcOEt, 1:1) = 0.83.

^1H NMR (400 MHz, $CDCl_3$) δ 7.30–7.13 (m, 8H), 7.10–7.06 (m, 2H), 5.66 (d, J = 7.8 Hz, 1H), 5.13 (d, J = 12.3 Hz, 1H), 5.03 (d, J = 12.3 Hz, 1H), 4.80 (dd, J = 14.0, 6.9 Hz, 1H), 3.36–3.26 (m, 4H), 3.08 (qd, J = 13.7, 6.6 Hz, 2H), 2.25–2.18 (m, 4H), 2.17 (s, 3H) ppm.

^{13}C NMR (75 MHz, $CDCl_3$) δ 172.5, 156.5, 136.2, 135.3, 129.3, 128.6, 128.4, 128.4, 126.9, 67.1, 54.4, 54.4, 45.8, 43.5, 38.2 ppm.

HRMS (ESI) *m/z* calculated for $C_{22}H_{28}N_3O_3^+$ $[M+H]^+$: 382.2131, found: 382.2133.

(4-Methylpiperazine-1-carbonyl)-L-phenylalanine ***(3)***. **2** (2.690 g, 7.050 mmol) was dissolved in ethanol (150 mL). Ammonium formate (4.580 g, 70.500 mmol, 10 eq) and Pd/C (0.120 g) were added. The reaction was refluxed for 2 h. The resulting solution was filtered through Celite and the solvent was removed under reduced pressure. The crude was recrystallized with MeOH. The pure product was isolated as a white powder (m = 2.048 g, Yield = 100 %).

R_f (Hex:AcOEt, 1:1) = 0.77.

^{1}H NMR (300 MHz, CD_3OD) δ 9.23 (s, 1H), 7.25–7.12 (m, 5H), 5.64 (s, 1H), 4.56 (dd, J = 12.2, 6.1 Hz, 1H), 3.64–3.45 (m, 2H), 3.45–3.28 (m, 2H), 3.20 (dd, J = 13.7, 5.3 Hz, 1H), 3.08 (dd, J = 13.7, 6.3 Hz, 1H), 2.78–2.54 (m, 4H), 2.44 (s, 3H) ppm.

^{13}C NMR (75 MHz, CD_3OD) δ 178.1, 158.8, 139.7, 130.5, 129.2, 127.4, 58.0, 54.6, 44.6, 43.2, 39.2 ppm.

IR (ATR) 3500–2500 (broad), 3024, 3017, 2968, 2945, 1627, 1572, 1539, 1434, 1391, 975 cm^{-1}.

HRMS (ESI) *m/z* calculated for $C_{15}H_{22}N_3O_3^+$ $[M+H]^+$: 292.1656, found: 292.1659.

m.p. 149–151 °C (Lit: [26] 121–123 °C).

(S)-2-((tert-Butoxycarbonyl)amino)-4-phenylbutanoic acid ***(4)***. L-Homophenylalanine (5.120 g, 28.000 mmol, 1 eq) was dissolved in THF (15 mL) and a solution of Na_2CO_3 (2.968 g, 31.000 mmol, 1.1 eq) in distilled water (100 mL) was added followed by addition of di-*tert*-butyl dicarbonate (6.975 g, 31.000 mmol, 1.1 eq) in 75 mL of THF. The mixture was stirred overnight, diluted with distilled water (100 mL), and acidified with $KHSO_4$ to pH=3. The organic product was extracted with CH_2Cl_2 (3 × 75 mL). The combined organic extracts were dried on Na_2SO_4 and concentrated under reduced pressure to give the crude product as a pale yellowish oil (m = 8.144 g, Yield = 100 %), which was used in the next step without further purification.

R_f (Hex:AcOEt, 1:1) = 0.38.

Major rotamer:

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.50 (broad s, 1H), 7.29–7.17 (m, 5H), 5.10 (d, J = 7.2 Hz, 1H), 4.37 (d, J = 4.3 Hz, 1H), 2.72 (t, J = 7.9 Hz, 2H), 2.28–2.13 (m, 1H), 1.99 (td, J = 14.7, 7.4 Hz, 1H), 1.45 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 177.3, 155.8, 140.8, 128.6, 128.6, 126.3, 80.4, 53.4, 34.2, 31.8, 28.4, 27.5 ppm.

IR (ATR) 3317, 3200–2400 (broad), 3015, 2978, 2929, 1707, 1655, 1498, 1460, 1394, 1368, 1260, 1159, 1051, 910, 731, 700 cm^{-1}.

HRMS (ESI) *m/z* calculated for $C_{15}H_{21}NO_4Na^+$ $[M+Na]^+$: 302.1363, found: 302.1366.

Optical rotation: $[\alpha]_D^{25}$ = 0.1412 (c = 0.85, CH_2Cl_2).

tert-Butyl (S)-(1-(methoxy(methyl)amino)-1-oxo-4-phenylbutan-2-yl) carbamate ***(5)***. **4** (6.238 g, 22.332 mmol, 1 eq) was dissolved in THF (125 mL) and the solution was cooled to 0 °C. DCC (5.445 g, 26.128 mmol, 1.17 eq), $HOB_t{\cdot}xH_2O$ (3.640 g, 26.128 mmol, 1.17 eq), and DIPEA (16.7 mL, 96.028 mmol, 4.3 eq) were added and the mixture was stirred for 15 min before N,O-dimethylhydroxylamine hydrochloride (2.600 g, 26.128 mmol, 1.17 eq) was added. Then, the mixture was allowed to warm up to r.t. and stirred overnight. Filtration of side product DCU was carried out several times (immersion of the solution in an ice bath to favor precipitation of DCU is recommended). THF was removed under reduced pressure and the crude was redissolved in CH_2Cl_2 (50 mL). The organic phase was washed with 1 M HCl (3 × 50 mL), $NaHCO_3$ saturated solution (3 × 50 mL), H_2O (50 mL), and brine (50 mL). The organic phase was dried over $MgSO_4$ and concentrated under vacuum. The crude product was purified by Silica Gel Column Chromatography (Hex:AcOEt, gradient from 8:2 to 7:3) to obtain the pure product as a transparent oil (m = 5.952 g, Yield = 83 %).

R_f (Hex:AcOEt, 7:3) = 0.39; R_f (Hex:AcOEt, 1:1) = 0.45.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.26 (s, 2H), 7.21–7.15 (m, 3H), 5.24 (d, J = 8.2 Hz, 1H), 4.68 (s, 1H), 3.62 (s, 3H), 3.16 (s, 3H), 2.79–2.61 (m, 2H), 2.09–1.97 (m, 1H), 1.89–1.77 (m, 1H), 1.45 (s, 9H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2, 155.7, 141.3, 128.6, 128.5, 126.1, 79.7, 61.6, 50.2, 34.7, 32.2, 31.8, 28.5 ppm.

IR (ATR) 3321, 3050, 2974, 2933, 1707, 1655, 1495, 1454, 1445, 1390, 1245, 1163, 992, 701 cm^{-1}.

HRMS (ESI) *m/z* calculated for $C_{17}H_{26}N_2O_4Na^+$ $[M+Na]^+$: 345.1785, found: 345.1786.

Optical rotation: $[\alpha]_D^{25}$ = 0.1371 (c = 1.75, CH_2Cl_2).

tert-Butyl (S)-(1-oxo-4-phenylbutan-2-yl)carbamate ***(6)***. A solution of **5** (3.429 g, 10.643 mmol, 1 eq) in THF (70 mL, 6.5 mL/mmol) was placed in an ice bath. Powdered $LiAlH_4$ (1.249 g, 31.928 mmol, 3 eq) was slowly added. The reaction mixture was stirred until completion (1 h 20 min) and subsequently quenched by addition of aqueous saturated Rochelle salt solution (70 mL). The organic phases were extracted with AcOEt (3 × 70 mL), washed with brine (70 mL), dried over $MgSO_4$, and concentrated over reduced pressure. Purification by Silica Gel Column Chromatography (Hex:AcOEt, gradient from 8:2 to 7:3) was performed and the pure product was obtained as a transparent oil (m = 4.866 g, Yield = 87 %).

R_f (Hex:AcOEt, 8:2) = 0.46, R_f (Hex:AcOEt, 7:3) = 0.51.

^{1}H NMR (300 MHz, $CDCl_3$) δ 9.54 (s, 1H), 7.33–7.16 (m, 5H), 5.18 (d, J = 5.2 Hz, 1H), 4.33–4.15 (m, 1H), 2.71 (t, J = 7.9 Hz, 2H), 2.30–2.14 (m, 1H), 1.88 (td, J = 15.4, 7.8 Hz, 1H), 1.47 (s, 9H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 199.7, 155.6, 140.7, 128.7, 128.5, 126.4, 80.2, 59.6, 31.5, 30.9, 28.4 ppm.

IR (ATR) 3347, 3027, 2978, 2930, 2863, 1685, 1498, 1450, 1380, 1249, 746, 701 cm^{-1}.

HRMS (ESI) *m/z* calculated for $C_{15}H_{21}NO_3Na^+$ $[M+Na]^+$: 286.1414, found: 286.1418.

Optical rotation:$[\alpha]_D^{25}$ = 0.0125 (c = 1.60, CH_2Cl_2).

tert-Butyl ((3S)-4-hydroxy-1-phenylhex-5-en-3-yl)carbamate ***(7)***. To a stirred solution of **6** (0.706 g, 2.841 mmol, 1 eq) in THF (74 mL, 26 mL/mmol), dry $ZnCl_2$ (0.852 g, 6.250 mmol, 2.2 eq) was added. The reaction was stirred for 10 min until $ZnCl_2$ was totally dissolved. Then, it was placed in an acetone-liquid N_2 bath at −78 °C and the vinylmagnesium bromide (12.5 mL 1 M in THF, 12.500 mmol, 4.4 eq) was added dropwise. The mixture was allowed to warm-up to 0 °C and it was followed by TLC until completion (1.5 h). NH_4Cl (75 mL) was added to quench the reaction. Organic and aqueous phases were separated. Then, the aqueous phase was extracted with ethyl acetate (3 × 50 mL), washed with brine, dried over $MgSO_4$, and concentrated under vacuum. The reaction crude was purified by Silica Gel Column Chromatography (Hex:AcOEt, gradient from 9:1 to 7:3) to afford a colorless oil (0.865 g, Yield = 52 %).

R_f (Hex:AcOEt, 7:3) = 0.51.

Proportion of diastereoisomers: A:B = 1:0.61.

Major diastereoisomer **A** (minor R_f):

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38–7.33 (m, 2H), 7.28–7.24 (m, 3H), 6.01–5.85 (m, 1H), 5.37 (dt, J = 17.2, 1.4 Hz, 1H), 5.26 (dt, J = 10.5, 1.3 Hz, 1H), 5.13–4.92 (m, 1H), 4.24–4.15 (m, 1H), 3.77–3.62 (m, 1H), 3.42–3.24 (m, 1H), 2.87–2.65 (m, 2H), 2.06–1.64 (m, 2H), 1.53 (s, 9H) ppm.

^{13}C NMR (75 MHz, $CDCl_3$) δ 156.5, 141.7, 138.2, 128.4, 128.4, 125.8, 116.0, 79.3, 74.5, 54.6, 33.7, 32.6, 28.4 ppm.

Minor diastereoisomer **B** (highest R_f):

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38–7.33 (m, 2H), 7.28–7.24 (m, 3H), 6.01–5.85 (m, 1H), 5.37 (dt, J = 17.2, 1.4 Hz, 1H), 5.26 (dt, J = 10.5, 1.3 Hz, 1H), 5.13–4.92 (m, 1H), 4.34–4.24 (m, 1H), 3.88–3.66 (m, 1H), 3.57–3.42 (m, 1H), 2.87–2.65 (m, 2H), 2.06–1.64 (m, 2H), 1.55 (s, 9H) ppm.

^{13}C NMR (75 MHz, $CDCl_3$) δ 156.7, 141.7, 137.0, 128.4, 128.4, 125.9, 116.5, 79.7, 75.3, 55.0, 33.7, 31.6, 28.4 ppm.

HRMS (ESI) *m/z* calculated for $C_{17}H_{25}NO_3Na^+$ $[M+Na]^+$: 314.1727, found: 314.1732.

IR (ATR) 3362 (broad), 3060, 2978, 2930, 2863, 1684, 1521, 1245, 1167, 1014 cm^{-1}.

Optical rotation: $[\alpha]_D^{25}$ = −0.2087 (c = 1.15, CH_2Cl_2).

(S)-1-(Methoxy(methyl)amino)-1-oxo-4-phenylbutan-2-aminium 2,2,2-

*trifluoroacetate (**8**)*. **5** (2.417 g, 7.502 mmol, 1 eq) was dissolved in CH_2Cl_2 (21.3 mL, 2.84 mL/mmol) and the solution was placed in an ice bath. Then, a 1:1 mixture of TFA and CH_2Cl_2 (0.46 mL TFA: 0.46 mL CH_2Cl_2, 1.5 mL/mmol each) was slowly added and the reaction was stirred until TLC showed completion (1 h and 45 min). Then, the solvent was evaporated and the crude was co-evaporated several times (x3) with CH_2Cl_2 to eliminate the excess of TFA. The product was obtained as a pale brown oil and used in the peptide coupling without further purification.

R_f (Hex:AcOEt, 1:1) = 0.

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.56 (s, 1H), 7.82 (s, 2H), 7.32–7.15 (m, 5H), 4.39–4.31 (m, 1H), 3.55 (s, 3H), 3.17 (s, 3H), 2.83 (ddd, J = 14.2, 8.3, 6.0 Hz, 1H), 2.77–2.69 (m, 1H), 2.31–2.14 (m, 2H) ppm.

HRMS (ESI) *m/z* calculated for $C_{12}H_{19}N_2O_2^+$ $[M]^+$: 223.1441, found: 223.1447.

*N-((S)-1-(((S)-1-(Methoxy(methyl)amino)-1-oxo-4-phenylbutan-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)-4-methylpiperazine-1-carboxamide (**9**)*. Deprotected compound **8** (1.262 g, 3.751 mmol, 1 eq) was dissolved in CH_2Cl_2 (37.5 mL, 10 mL/mmol) and the solution was placed in an ice bath. **3** (1.202 g, 4.126 mmol, 1.1 eq) was added and the mixture was stirred for 10 min. HOBt·xH_2O (0.575 g, 4.126 mmol, 1.1 eq) was added and the reaction was stirred for 10 min. The, DIPEA (2.6 mL, 15.004 mmol, 4 eq) was added and the reaction was stirred for additional 10 min. Finally, EDC·HCl (0.880 g, 4.501 mmol, 1.2 eq) was added and the reaction was stirred overnight at r.t. The reaction was quenched with distilled water (35 mL). Organic and aqueous phases were separated and the aqueous phase was extracted with CH_2Cl_2 (3 × 35 mL). The collected organic phases were washed with brine, dried over $MgSO_4$, and concentrated in vacuum. The crude product was purified by Silica Gel Column Chromatography (CH_2Cl_2:MeOH = 9:1) to obtain the pure product as a white solid (m = 1.222 g, Yield = 66 %).

R_f (CH_2Cl_2:MeOH, 9:1) = 0.44.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.27–7.12 (m, 10*H*), 6.77 (d, J = 8.2 Hz, 1H), 5.02 (d, J = 7.2 Hz, 1H), 4.98–4.86 (m, 1H), 4.64 (q, J = 6.7 Hz, 1H), 3.61 (s, 3H), 3.40–3.29 (m, 4H), 3.16 (s, 3H), 3.09 (d, J = 6.6 Hz, 2H), 2.67–2.52 (m, 2H), 2.38–2.29 (m, 4H), 2.27 (s, 3H), 2.06–1.99 (m, 1H), 1.91–1.81 (m, 1H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.8, 156.9, 141.2, 137.0, 129.6, 128.7, 128.6, 128.5, 127.0, 126.1, 61.6, 55.5, 54.6, 49.1, 46.1, 43.8, 38.7, 34.1, 32.2, 31.6 ppm.

HRMS (ESI) *m/z* calculated for $C_{27}H_{38}N_5O_4^+$ $[M+H]^+$: 496.2918, found: 496.2924.

IR (ATR) 3284, 3060, 3027, 2933, 2851, 1648, 1618, 1532, 1439, 1290, 1264, 1006 cm^{-1}.

m.p. 137–139 °C.

Optical rotation: $[\alpha]_D^{25}$ = −0.2727 (c = 1.10, CH_2Cl_2).

*4-Methyl-N-((S)-1-oxo-1-(((S)-1-oxo-4-phenylbutan-2-yl)amino)-3-phenylpropan-2-yl)piperazine-1-carboxamide (**10**)*. A solution of **9** (1.222 g, 2.466 mmol, 1 eq) in THF (16 mL, 6.5 mL/mmol) was placed in an ice bath. Powdered $LiAlH_4$ (0.236 g, 5.918 mmol, 2.4 eq) was slowly added. The reaction mixture was stirred until completion (3 h) and subsequently quenched by addition of aqueous saturated Rochelle salt solution (30 mL). The organic phases were extracted with AcOEt (3 × 30 mL), washed with brine (80 mL), dried over $MgSO_4$, and concentrated over reduced pressure. Purification by Silica Gel Column Chromatography (CH_2Cl_2:MeOH, gradient from 9:1 to 85:15) was performed and the pure product was obtained as a pale yellow oil (m = 0.532 g, Yield = 57 %).

R_f (CH_2Cl_2:MeOH, 85:15) = 0.43.

Proportion of diastereoisomers: A:B = 1:0.66.

Major diastereoisomer **A** (lowest R_f):

^{1}H NMR (300 MHz, $CDCl_3$) δ 9.39 (s, 1H), 7.32–7.11 (m, 9H), 7.11–7.05 (m, 2H), 5.35–5.23 (m, 1H), 4.78–4.41 (m, 1H), 4.40–4.13 (m, 1H), 3.38–3.29 (m, 4H), 3.09 (d, J = 7.2 Hz, 2H), 2.64–2.43 (m, 2H), 2.39–2.28 (m, 4H), 2.27 (s, 3H), 2.18–2.05 (m, 1H), 1.87–1.72 (m, 1H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) 199.2, 172.9, 157.1, 140.6, 137.0, 129.5, 128.7, 128.7, 128.5, 127.1, 126.4, 58.4, 55.9, 54.5, 46.0, 43.8, 38.7, 31.4, 30.3 ppm.

Minor diastereoisomer **B** (highest R_f):

^{1}H NMR (300 MHz, $CDCl_3$) δ 9.30 (s, 1H), 7.32–7.11 (m, 9H), 7.11–7.05 (m, 2H), 5.35–5.23 (m, 1H), 4.78–4.41 (m, 1H), 4.28 (m, 1H), 3.38–3.29 (m, 4H), 3.09 (d, J = 7.2 Hz, 2H), 2.64–2.43 (m, 2H), 2.39–2.28 (m, 4H), 2.27 (s, 3H), 2.18–2.05 (m, 1H), 1.87–1.72 (m, 1H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) 199.2, 172.9, 157.1, 140.5, 137.1, 129.4, 128.8, 128.7, 128.5, 127.1, 126.4, 58.4, 55.8, 54.5, 46.0, 43.8, 38.7, 31.4, 30.4 ppm.

HRMS (ESI) *m/z* calculated for $C_{25}H_{33}N_4O_3^+$ $[M+H]^+$: 437.2547, found: 437.2553.

IR (ATR) 3280, 3060, 3026, 2930, 2855, 2795, 1656, 1618, 1528, 1439, 1264, 1144 cm^{-1}.

m.p. 79–80 °C.

Optical rotation: $[\alpha]_D^{25}$ = −0.2105 (c = 0.95, CH_2Cl_2).

*Ethyl (S,E)-4-((S)-2-(4-methylpiperazine-1-carboxamido)-3-phenylpropanamido)-6-phenylhex-2-enoate (**11**)*. A solution of triethyl-2-phosphonoacetate (0.06 mL, 0.285 mmol, 1.5 eq) in THF (1 mL, 3.5 mL/mmol) was added dropwise to a −78 °C suspension of 60 % sodium hydride (0.011 g, 0.285 mmol, 1.5 eq) in THF (0.8 mL, 2.7 mL/mmol). The mixture was stirred for 10 min at −78 °C, before a solution of **10** (0.091 g, 0.190 mmol, 1 eq) in THF (0.55 mL, 2.8 mL/mmol) was added dropwise. The mixture was allowed to warm-up to 0 °C and the reaction was stirred overnight. Then, the reaction was quenched with saturated aqueous solution of ammonium chloride (3 mL). The phases were separated and the aqueous phase was extracted with CH_2Cl_2 (3 × 3 mL), dried over Na_2SO_4, and concentrated in vacuum. The crude product was purified by silica gel column chromatography (CH_2Cl_2: MeOH, gradient from 95:5 to 9:1) and the pure product was obtained as a white solid (m = 0.061 g, Yield = 84 %). For major purity, the product was crystallized with acetonitrile.

R_f (CH_2Cl_2:MeOH, 9:1) = 0.5.

Proportion of diastereoisomers: A:B = 1:0.54.

Major diastereoisomer **A**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29–7.14 (m, 8H), 7.10–7.03 (m, 2H), 6.78 (dd, J = 15.7, 5.3 Hz, 1H), 6.59–6.73 (m, 1H), 5.85 (dd, J = 15.7, 1.2 Hz, 1H), 5.28–5.12 (m, 1H), 4.63–4.47 (m, 2H), 4.21–4.10 (m, 2H), 3.38–3.27 (m, 4H), 3.13–3.00 (m, 2H), 2.45 (t, J = 7.9 Hz, 2H), 2.33–2.27 (m, 4H), 2.26 (s, 3H), 1.85–1.61 (m, 2H), 1.24 (t, J = 7.1 Hz, 3H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.8, 166.3, 157.1, 147.3, 140.9, 137.2, 129.5, 128.8, 128.6, 128.4, 127.1, 126.2, 121.4, 60.5, 56.1, 54.6, 49.9, 46.1, 43.8, 38.5, 35.8, 31.9, 14.3 ppm.

Minor diastereoisomer **B**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29–7.14 (m, 8H), 7.10–7.03 (m, 2H), 6.68 (dd, J = 15.7, 5.8 Hz, 1H), 6.59–6.73 (m, 1H), 5.69 (dd, J = 15.7, 1.4 Hz, 1H), 5.28–5.12 (m, 1H), 4.63–4.47 (m, 2H), 4.21–4.10 (m, 2H), 3.38–3.27 (m, 4H), 3.13–3.00 (m, 2H), 2.61–2.50 (m, 2H), 2.33–2.27 (m, 4H), 2.26 (s, 3H), 1.85–1.61 (m, 2H), 1.29 (t, J = 7.2 Hz, 3H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.8, 166.3, 157.1, 147.2, 141.0, 137.0, 129.5, 128.8, 128.6, 128.5, 127.1, 126.2, 121.4, 60.5, 56.1, 54.6, 49.9, 46.1, 43.9, 38.8, 36.0, 31.9, 14.4 ppm.

HRMS (ESI) *m/z* calculated for $C_{29}H_{39}N_4O_4^+$ $[M+H]^+$: 507.2966, found: 507.2971.

IR (ATR) 3276, 3060, 3027, 2933, 2851, 1715, 1655, 1618, 155, 1443, 1290, 1264 cm^{-1}.

m.p. 117–119 °C.

Optical rotation: $[\alpha]_D^{25}$ = −0.0824 (c = 1.70, CH_2Cl_2).

*Ethyl (5S,E)-5-((tert-butoxycarbonyl)amino)-4-hydroxy-7-phenylhept-2-enoate (**12**)*. To a solution of **7** (0.551 g, 1.520 mmol, 1 eq) in CH_2Cl_2 (15 mL, 10 mL/mmol), ethyl acrylate (0.84 mL, 7.599 mmol, 5 eq) and

2nd generation Hoveyda Grubb's catalyst (29.5 mg, 0.046 mmols, 0.03 eq) were added at r.t. and the mixture was refluxed at 50 °C for 5 h. The reaction crude was concentrated under vacuum and purified by Silica Gel Column Chromatography (hexane: ethyl acetate, 7:3) to afford a pale brown oil (m = 0.456 g, Yield = 68 %).

R_f (Hex:AcOEt, 7:3) = 0.31.

Proportion of diastereoisomers: A:B = 1:0.22.

Major diastereoisomer **A**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30–7.23 (m, 2H), 7.20–7.13 (m, 3H), 6.92 (dd, J = 15.6, 4.4 Hz, 1H), 6.08 (dd, J = 15.5 Hz, 1.5 Hz, 1H), 4.85 (d, J = 9.2 Hz, 1H), 4.33 (broad s, 1H), 4.17 (q, J = 7.1 Hz, 2H), 3.70–3.57 (m, 1H), 3.48–3.37 (m, 1H), 2.79–2.69 (m, 1H), 2.69–2.61 (m, 1H), 1.99–1.75 (m, 2H), 1.41 (s, 9H), 1.26 (t, J = 7.1 Hz, 3H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 166.4, 156.5, 147.6, 141.5, 128.6, 128.5, 126.1, 121.8, 79.9, 73.0, 60.5, 54.6, 33.1, 32.6, 28.4, 14.3 ppm.

Minor diastereoisomer **B**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30–7.23 (m, 2H), 7.20–7.13 (m, 3H), 6.88 (dd, J = 15.6, 4.4 Hz, 1H), 6.12 (dd, J = 15.5 Hz, 1.5 Hz, 1H), 4.76 (d, J = 7.9 Hz, 1H), 4.38 (broad s, 1H), 4.18 (q, J = 7.1 Hz, 2H), 3.70–3.57 (m, 1H), 3.48–3.37 (m, 1H), 2.79–2.69 (m, 1H), 2.69–2.61 (m, 1H), 1.99–1.75 (m, 2H), 1.45 (s, 9H), 1.28 (t, J = 7.1 Hz, 3H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 166.4, 157.1, 146.2, 141.2, 128.6, 128.5, 126.2, 122.6, 80.3, 74.4, 60.6, 55.3, 32.6, 32.0, 28.4, 14.3 ppm.

HRMS (ESI) *m/z* calculated for $C_{20}H_{29}NO_5Na^+$ $[M+Na]^+$: 386.1938, found: 386.1943.

IR (ATR) 3358 (broad), 3027, 2978, 2933, 1685, 1498, 1249, 1163 cm^{-1}.

Optical rotation minor diastereoisomer **B**: $[\alpha]_D^{25} = -0.2526$ (c = 1.90, CH_2Cl_2).

*(3S,E)-7-Ethoxy-4-hydroxy-7-oxo-1-phenylhept-5-en-3-aminium 2,2,2-trifluoroacetate **(13)***. **12** (0.221 g, 0.608 mmol, 1 eq) was dissolved in CH_2Cl_2 (1.7 mL, 2.84 mL/mmol) and the solution was placed in an ice bath. Then, a 1:1 mixture of TFA and CH_2Cl_2 (0.9 mL TFA: 0.9 mL CH_2Cl_2, 1.5 mL/mmol each) was slowly added and the reaction was stirred until TLC showed completion (1 h and 35 min). Then, the solvent was evaporated and the crude was co-evaporated several times (x3) with CH_2Cl_2 to eliminate the excess of TFA. The product was obtained as a pale brown oil that became a pale brown solid after long time in the high vacuum pump. The crude was used in the peptide coupling without further purification.

R_f (Hex:AcOEt, 7:3) = 0.

*Ethyl (5S,E)-4-hydroxy-5-((S)-2-(4-methylpiperazine-1-carboxamido)-3-phenylpropanamido)-7-phenylhept-2-enoate **(14)***. **13** (0.608 mmol, 1 eq) was dissolved in CH_2Cl_2 (6 mL, 10 mL/ mmol) and the solution was placed in an ice bath. **3** (0.195 g, 0.669 mmol, 1.1 eq) was added and the mixture was stirred for 5 min. HOBT·xH_2O (0.093 g, 0.608 mmol, 1.1 eq) was added and the reaction was stirred for 10 min. EDC·HCl (0.143 g, 0.730 mmol, 1.2 eq) was added and the reaction was stirred for 5 min. DIPEA (0.42 mL, 2.432 mmol, 4 eq) was finally added and the mixture was stirred for 18 h at r.t. NH_4Cl (6 mL) was added to quench the reaction. The organic phases were separated and the aqueous phase was extracted with CH_2Cl_2 (3 × 6 mL). The collected organic phases were washed with brine (20 mL). The organic layer was dried over $MgSO_4$ and concentrated under vacuum. The reaction crude was purified by Silica Gel Column Chromatography (CH_2Cl_2:MeOH, gradient from 95:5 to 85:15) to afford a white solid (m = 0.180 g, Yield = 50 %).

R_f (CH_2Cl_2:MeOH, 9:1) = 0.37.

Mixture of diastereoisomers: A:B:C:D=1:0.69:0.24:0.17.

Major diastereoisomer **A**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.24–7.04 (m, 11*H*), 6.78 (dd, J = 15.5, 4.3 Hz, 1H), 6.06 (dd, J = 15.6, 1.7 Hz, 1H), 5.25 (d, J = 7.6 Hz, 1H), 4.58–4.50 (m, 1H), 4.40–4.22 (m, 1H), 4.15 (q, J = 7.2 Hz, 2H), 4.04–3.85 (m, 1H), 3.38–3.23 (m, 4H), 3.11–2.97 (m, 2H), 2.64–2.49 (m, 1H), 2.42–2.34 (m, 1H), 2.33–2.25 (m, 4H), 2.22 (s, 3H), 1.97–1.58 (m, 2H), 1.25 (t, J = 7.1 Hz, 3H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 173.1, 166.4, 157.2, 147.4, 141.5, 137.0, 129.4, 128.7, 128.5, 128.5, 127.1, 126.1, 122.0, 72.3, 60.5, 56.1, 54.5, 54.1, 46.1, 43.8, 38.5, 33.1, 32.4, 14.3 ppm.

HRMS (ESI) *m/z* calculated for $C_{30}H_{41}N_4O_5^+$ $[M+H]^+$: 537.3071, found: 537.3080.

IR (ATR) 3280 (broad), 3060, 3027, 2933, 2855, 1715, 1618, 1528, 1454, 1290 cm^{-1}.

m.p. 78–80 °C.

Optical rotation major diastereoisomer **A**: $[\alpha]_D^{25} = -0.1217$ (c = 1.15, CH_2Cl_2).

*Ethyl (S,E)-5-((S)-2-(4-methylpiperazine-1-carboxamido)-3-phenylpropanamido)-4-oxo-7-phenylhept-2-enoate **(15)***. **14** (0.095 g, 0.177 mmol, 1 eq) was dissolved in CH_2Cl_2 (7 mL, 40 mL/mmol) and placed in the cryocool at −5 °C under N_2 atmosphere. Then, DMP (1.139 g, 2.683 mmol, 9 eq) was added in one portion and the reaction was stirred overnight. $NaHCO_3$ and $Na_2S_2O_3$ solution (1.2 g/100 mL each) was added and the reaction was stirred for 1 h. The phases were separated and the product was extracted with CH_2Cl_2 (x3). The collected organic phases were dried over $MgSO_4$ and concentrated over vacuum. The crude product was purified by Silica Gel Column Chromatography (CH_2Cl_2:MeOH, gradient from 9:1 to 88:12) to afford a yellow oil that became a foamy yellow solid in the high vacuum pump (m = 0.117 g, Yield = 73 %).

R_f (CH_2Cl_2:MeOH, 88:12) = 0.59.

Mixture of diastereoisomers: A:B = 1:0.43.

Major diastereoisomer **A**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30–7.12 (m, 9H), 7.13–7.02 (m, 2H), 7.09 (d, J = 15.8 Hz, 1H), 6.70 (d, J = 15.8 Hz, 1H), 5.26–5.10 (m, 1H), 4.72 (dt, J = 11.8, 5.9 Hz, 1H), 4.69–4.60 (m, 1H), 4.25 (q, J = 7.2 Hz, 2H), 3.40–3.31 (m, 4H), 3.11–3.04 (m, 2H), 2.55–2.41 (m, 2H), 2.39–2.32 (m, 4H), 2.28 (s, 3H), 2.17–2.04 (m, 1H), 1.92–1.73 (m, 1H), 1.31 (t, J = 7.1 Hz, 3H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 197.0, 172.3, 165.1, 157.0, 140.4, 137.0, 136.0, 132.6, 129.4, 128.8, 128.6, 128.6, 127.1, 126.5, 61.6, 57.2, 55.8, 54.4, 45.8, 43.6, 38.5, 32.7, 31.4, 14.2 ppm.

Minor diastereoisomer **B**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30–7.12 (m, 9H), 7.13–7.02 (m, 2H), 7.06 (d, J = 15.9 Hz, 1H), 6.65 (d, J = 15.8 Hz, 1H), 5.26–5.10 (m, 1H), 4.72 (dt, J = 11.8, 5.9 Hz, 1H), 4.69–4.60 (m, 1H), 4.22 (d, J = 7.3 Hz, 2H), 3.40–3.31 (m, 4H), 3.11–3.04 (m, 2H), 2.55–2.41 (m, 2H), 2.39–2.32 (m, 4H), 2.29 (s, 3H), 2.17–2.04 (m, 1H), 1.92–1.73 (m, 1H), 1.29 (t, J = 7.1 Hz, 3H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 196.7, 172.3, 165.2, 157.0, 140.5, 136.9, 136.1, 132.6, 129.4, 128.7, 128.6, 128.6, 127.1, 126.4, 61.6, 57.2, 55.8, 54.4, 45.8, 43.6, 38.5, 32.6, 31.4, 14.2 ppm.

HRMS (ESI) *m/z* calculated for $C_{30}H_{39}N_4O_5^+$ $[M+H]^+$: 535.2915, found: 535.2920.

IR (ATR) 3258, 3060, 3027, 2933, 2851, 1703, 1655, 1618, 1536, 1443, 1290, 1260 cm^{-1}.

m.p. 71–73 °C.

Optical rotation of the mixture A:B = 1:0.96: $[\alpha]_D^{25} = -0.2571$ (c = 1.4, CH_2Cl_2).

*tert-Butyl ((3S,E)-4-hydroxy-1-phenyl-6-(phenylsulfonyl)hex-5-en-3-yl) carbamate **(16)***. To a solution of **7** (0.374 g, 1.283 mmol, 1 eq) in acetonitrile (26 mL, 20 mL/mmol) were added dry sodium acetate (0.316 g, 3.850 mmol, 3 eq), benzenesulfinic acid sodium salt (1.290 g, 7.701 mmol, 6 eq), and iodine (0.977 g, 3.850 mmol, 3 eq) and the reaction was stirred under reflux for 41 h. A saturated solution of $Na_2S_2O_3$ (12 mL/mmol) was added to stop the reaction and a saturated solution of $NaHCO_3$ (12 mL/mmol) was added to basify the pH. Organic and aqueous phases were separated and the aqueous phase was extracted with AcOEt (×3). The reaction crude was purified by Silica Gel Column Chromatography (Hex:AcOEt, gradient from 7:3 to 1:1) to afford a white solid (m = 0.221 g, Yield = 40 %).

R_f (Hex:AcOEt, 7:3) = 0.28.

Proportion of diastereoisomers: A:B = 1:0.92.

Major diastereoisomer **A**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90–7.83 (m, 2H), 7.63–7.56 (m, 1H), 7.55–7.48 (m, 2H), 7.31–7.25 (m, 2H), 7.22–7.11 (m, 3H), 6.96 (dd, J = 14.9, 3.2 Hz, 1H), 6.67 (d, J = 14.9 Hz, 1H), 4.83–4.64 (m, 1H), 4.45–4.37 (m, 1H), 3.84–3.69 (m, 1H), 3.65–3.37 (m, 1H), 2.76–2.59 (m, 2H), 2.00–1.71 (m, 2H), 1.40 (s, 9H) ppm.

^{13}C NMR (101 MHz, $CDCl_3$) δ 156.8, 146.1, 141.1, 140.5, 133.5, 131.5, 129.4, 128.7, 128.5, 127.8, 126.3, 80.5, 72.7, 54.8, 32.5, 28.5, 28.4 ppm.

Minor diastereoisomer **B**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90–7.83 (m, 2H), 7.63–7.56 (m, 1H), 7.55–7.48 (m, 2H), 7.31–7.25 (m, 2H), 7.22–7.11 (m, 3H), 6.93 (dd, J = 14.8, 3.1 Hz, 1H), 6.70 (d, J = 14.9 Hz, 1H), 4.83–4.64 (m, 1H), 4.45–4.37 (m, 1H), 3.84–3.69 (m, 1H), 3.65–3.37 (m, 1H), 2.76–2.59 (m, 2H), 2.00–1.71 (m, 2H), 1.43 (s, 9H) ppm.

HRMS (ESI) m/z calculated for $C_{23}H_{29}NO_5SNa^+$ $[M+Na]^+$: 454.1659, found: 454.1664.

IR (ATR) 3381 (broad), 3060, 3027, 2974, 2926, 1659, 1521, 1290, 1249, 1144 cm^{-1}.

Optical rotation: $[\alpha]_D^{25} = -0.6286$ (c = 0.35, CH_2Cl_2).

(3S,E)-4-Hydroxy-1-phenyl-6-(phenylsulfonyl)hex-5-en-3-aminium 2,2,2-trifluoroacetate ***(17)***. **16** (0.243 g, 0.563 mmol, 1 eq) was dissolved in CH_2Cl_2 (1.6 mL, 2.84 mL/mmol) and the solution was placed in an ice bath. Then, a 1:1 mixture of TFA and CH_2Cl_2 (0.84 mL TFA: 0.84 mL CH_2Cl_2, 1.5 mL/mmol each) was slowly added and the reaction was stirred until TLC showed completion (1 h and 45 min). Then, the solvent was evaporated and the crude was co-evaporated several times (x3) with CH_2Cl_2 to eliminate the excess of TFA. The product was obtained as a pale brown oil that became a pale brown solid after long time in the high vacuum pump. The crude was used in the peptide coupling without further purification.

R_f (Hex:AcOEt, 6:4) = 0.

N-((2S)-1-(((3S,E)-4-Hydroxy-1-phenyl-6-(phenylsulfonyl)hex-5-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)-4-methylpiperazine-1-carboxamide ***(18)***. **17** (0.563 mmol, 1 eq) was dissolved in CH_2Cl_2 (6 mL, 10 mL/mmol) and the solution was placed in an ice bath. **3** (0.180 g, 0.619 mmol, 1.1 eq) was added and the mixture was stirred for 5 min. HOBT·xH_2O (0.085 g, 0.619 mmol, 1.1 eq) was added and the reaction was stirred for 10 min. EDC·HCl (0.132 g, 0.676 mmol, 1.2 eq) was added and the reaction was stirred for 5 min. DIPEA (0.39 mL, 2.252 mmol, 4 eq) was finally added and the mixture was stirred for 18 h at r.t. NH_4Cl (6 mL) was added to quench the reaction. The organic phases were separated and the aqueous phase was extracted with CH_2Cl_2 (3 × 6 mL). The collected organic phases were washed with 1 M HCl (3 × 20 mL), $NaHCO_3$ saturated solution (3 × 20 mL), and brine (20 mL). The organic layer was dried over $MgSO_4$ and concentrated under vacuum. The reaction crude was purified by Silica Gel Column Chromatography (CH_2Cl_2: MeOH, gradient from 98:2 to 88:12) to afford a transparent oil that became a white solid after long hours in the high vacuum pump (m = 0.142 g, Yield = 42 %).

R_f (CH_2Cl_2:MeOH, 95:5) = 0.41.

Proportion of diastereoisomers: A:B = 1:0.23.

Major diastereoisomer A (minor R_f):

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.89–7.81 (m, 2H), 7.60–7.54 (m, 1H), 7.53–7.46 (m, 2H), 7.29–7.07 (m, 11*H*), 6.85 (dd, J = 14.9, 3.2 Hz, 1H), 6.65 (dd, J = 14.9, 1.6 Hz, 1H), 5.06 (d, J = 7.3 Hz, 1H), 4.47–4.31 (m, 2H), 3.74–3.58 (m, 1H), 3.47–3.14 (m, 5H), 3.05 (dd, J = 14.1, 6.4 Hz, 1H), 2.92 (dd, J = 14.0, 8.5 Hz, 1H), 2.68–2.52 (m, 2H), 2.33–2.24 (m, 4H), 2.23 (s, 3H), 1.99–1.90 (m, 2H) ppm.

^{13}C NMR (75 MHz, $CDCl_3$) δ 174.0, 157.5, 145.9, 141.1, 140.2, 137.1, 133.6, 131.0, 129.4, 129.4, 128.8, 128.6, 128.5, 127.8, 127.1, 126.2, 72.1, 55.9, 54.8, 54.5, 46.1, 43.8, 37.5, 32.3, 31.7 ppm.

Minor diastereoisomer B (major R_f):

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.87–7.82 (m, 2H), 7.62–7.55 (m, 1H), 7.53–7.47 (m, 2H), 7.35–7.15 (m, 9H), 7.10–7.03 (m, 2H), 6.86 (dd, J = 14.9, 3.2 Hz, 1H), 6.58 (dd, J = 14.9, 2.0 Hz, 1H), 6.48 (d, J = 8.1 Hz, 1H), 4.89 (d, J = 6.1 Hz, 1H), 4.43–4.33 (m, 2H), 4.05–3.95 (m, 1H), 3.40–3.18 (m, 5H), 3.09 (dd, J = 13.7, 7.0 Hz, 1H), 2.99 (dd, J = 13.7, 8.0 Hz, 1H), 2.70–2.43 (m, 2H), 2.33–2.27 (m, 4H), 2.25 (s, 3H), 1.72–1.63 (m, 2H) ppm.

^{13}C NMR (75 MHz, $CDCl_3$) δ 173.3, 157.4, 144.3, 141.0, 140.3, 137.0, 133.6, 132.0, 129.4, 129.3, 129.1, 128.7, 128.5, 127.8, 127.4, 126.3, 72.4, 56.7, 54.5, 54.1, 46.1, 43.9, 38.1, 32.4, 31.1, 30.6 ppm.

HRMS (ESI) m/z calculated for $C_{33}H_{41}N_4O_5S^+$ $[M+H]^+$: 605.2792, found: 605.2798.

IR (ATR) 3280 (broad), 3060, 3030, 2922, 2851, 1617, 1524, 1446, 1290, 1264, 1140, 1085 cm^{-1}.

m.p. 143–145 °C.

Optical rotation major diastereoisomer **A**: $[\alpha]_D^{25} = -0.3478$ (c = 0.35, CH_2Cl_2).

Optical rotation minor diastereoisomer **B**: $[\alpha]_D^{25} = -0.2462$ (c = 0.33, CH_2Cl_2).

4-Methyl-N-((S)-1-oxo-1-(((S,E)-4-oxo-1-phenyl-6-(phenylsulfonyl) hex-5-en-3-yl)amino)-3-phenylpropan-2-yl)piperazine-1-carboxamide ***(19)***. **18** (0.096 g, 0.159 mmol, 1 eq) was dissolved in CH_2Cl_2 (2 mL, 10 mL/mmol) and placed in the cryocool at −5 °C under N_2 atmosphere. Then, DMP (0.208 g, 0.476 mmol, 3 eq) was added in one portion and the reaction was stirred for 18 h. $NaHCO_3$ and $Na_2S_2O_3$ solution (1.2 g/100 mL each) was added, and the reaction was stirred for 1 h. The phases were separated, and the product was extracted with CH_2Cl_2 (x3). The collected organic phases were washed with brine, dried over $MgSO_4$, and concentrated over vacuum. The pure product was obtained as an intense yellow solid (m = 0.096 g, Yield = 100 %).

R_f (CH_2Cl_2:MeOH, 95:5) = 0.60.

Described as a mixture of diastereoisomers A and B:

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90–7.84 (m, 2H), 7.70–7.63 (m, 1H), 7.59–7.53 (m, 2H), 7.31–7.14 (m, 10*H*), 7.11–7.08 (m, 1H), 7.07–7.05 (m, 1H), 7.05–7.03 (m, 1H), 5.14–5.02 (m, 1H), 4.67–4.51 (m, 2H), 3.38–3.28 (m, 4H), 3.10–2.01 (m, 2H), 2.55–2.49 (m, 1H), 2.44 (t, J = 7.6 Hz, 1H), 2.36–2.29 (m, 4H), 2.26 (s, 3H), 2.14–2.01 (m, 1H), 1.79 (dt, J = 15.4, 7.9 Hz, 1H) ppm.

Major diastereoisomer **A**:

^{13}C NMR (75 MHz, $CDCl_3$) δ 195.9, 172.6, 170.9, 157.2, 141.7, 140.2, 138.5, 137.0, 134.5, 132.9, 129.8, 129.4, 128.8, 128.7, 128.6, 128.4, 127.1, 126.5, 57.5, 55.8, 54.4, 46.0, 43.7, 38.1, 32.0, 31.4 ppm.

Minor diastereoisomer **B**:

^{13}C NMR (75 MHz, $CDCl_3$) δ 195.5, 172.5, 170.9, 157.1, 141.7, 140.2, 138.6, 136.9, 134.5, 133.0, 129.8, 129.4, 128.8, 128.7, 128.6, 128.4, 127.1, 126.5, 57.7, 55.7, 54.5, 45.9, 43.7, 38.1, 32.0, 31.4 ppm.

HRMS (ESI) m/z calculated for $C_{33}H_{39}N_4O_5S^+$ $[M+H]^+$: 603.2636, found: 603.2641.

IR (ATR) 3261, 3056, 3027, 2922, 2851, 1618, 1525, 1446, 1290, 1263, 1144, 1081 cm^{-1}.

m.p. 123–124 °C.

Optical rotation mix of **A** and **B**: $[\alpha]_D^{25} = -0.3000$ (c = 0.20, CH_2Cl_2).

Enzyme assays.

a) Inhibition assay. Rhodesain was expressed as previously described [10,27] and human cathepsins B and L were commercial (Sigma-Aldrich, 219364 and 219402, respectively). Proteolytic activities against rhodesain, cathepsin L and cathepsin B were determined using a fluorescence-based assay, as previously described [10,27,28]. Increase of fluorescence upon cleavage of the fluorogenic substrate Cbz-Phe-Arg-AMC (Bachem) was monitored using a TECAN Infinite F200 Pro fluorimeter (λ excitation: 380 nm, λ emission: 460 nm). Assays were performed in white, flat-bottom 96-well microtiter plates (Greiner bio-one) with a total volume of 200 μL consisting of 180 μL assay buffer, 5 μL enzyme in enzyme buffer, 10 μL DMSO with or without inhibitor, and 5 μL substrate. Initial screens were performed at a fixed concentration of 20 μM. Dilution series of active inhibitors with at least seven

different concentrations were prepared in duplicates/triplicates and fluorescence emission was monitored directly after addition of the substrate. For more detailed information, measuring conditions for all the proteases are summarized in Table S2).

b) Dilution assays. Previously activated rhodesain (10 μL from 200 nM solution in enzyme buffer) in assay buffer (8 μL) was incubated for 10 min with inhibitors (2 μL in DMSO) at concentrations corresponding to 10-fold* the IC_{50} value (or K_i^{app} value for irreversible inhibitors) obtained from fluorometric enzyme assay to ensure complete inhibition. These mixtures (2 μL) were diluted 100-fold in assay buffer (193 μL) and substrate (5 μL from 400 μM solution in DMSO). Recovery of enzyme activity was measured immediately using a fluorescence readout. Rhodesain with DMSO (no inhibitor) was used as a blank control, nitrile inhibitor [20] was used as a reversible control, and **K11777** was used as an irreversible control. The results are given in fractional activity of uninhibited rhodesain (only DMSO, no inhibitor) used in the same experiment as blank control.

*For compounds **14** and **18**, 20-fold was used. For the reversible control, 100-fold was used.

c) Dialysis experiments. Dialysis experiments for rhodesain were performed in a custom-built dialysis chamber [18], allowing the parallel examination of five samples. A 13 kDa MW cut-off cellulose tubing was used as the dialysis membrane, separating a chamber with continuous flow of assay buffer from five cavities at the top of the instrument, where the samples were added.

Therefore, rhodesain (25 μL from 80 nM rhodesain in enzyme buffer; 2 nM end concentration of enzyme) was added to 900 μL of assay buffer (containing 0.2 mM of DTT) and 50 μL of DMSO with or without inhibitor. Inhibitor concentrations of 10-fold* the IC_{50} value (or K_i^{app}) during the incubation were chosen to guarantee complete inhibition. After incubation for 10 min, the mixtures (800 μL) were transferred to the cavities in the instrument and dialyzed against a continuous flow of assay buffer containing 0.2 mM of DTT and 5 % DMSO (approximate flow rate of 300 mL/h). Aliquots (58.5 μL) were taken in duplicates at different time points (0, 10, and 30 min) and substrate (1.5 μL from 400 μM solution in DMSO; 10 μM end concentration of substrate) was added. Enzyme activity was immediately determined by measuring the fluorescence emission. The results are given in fractional activity of uninhibited rhodesain (only DMSO, no inhibitor) used in the same experiment as blank control. Also, nitrile [20] was used as a reversible control and **K11777** was used as an irreversible control.

*For the reversible control, 100-fold was used.

Mass spectrometry (MS). Mass spectrometric experiments were performed as described before [23,29]. Lyophilized rhodesain was dissolved to 10 μM in buffer (pH = 5.5, 50 mM NaOAc, 200 mM NaCl, 5 mM EDTA, 5 mM DTT) and incubated at r.t. for 30–60 min to activate the enzyme. Afterwards, each inhibitor (dissolved in DMSO) was added to 100 μL of the solution of activated rhodesain, resulting in 100 μM inhibitor with 2.5 % residual DMSO.

The solutions of protein alone or protein-inhibitor complexes were desalted by using Zeba Spin Desalting Columns (7 kDa MWCO, 0.5 mL; Thermo Fisher Scientific) following the manufacturer's instructions. On the target, desalted sample solutions were mixed 1:1 with a MALDI-matrix, using two different methods: *method a*. Sinapinic acid, saturated solution in ACN/water 1:1 with 0.1 % TFA and/or *method b*. 1 part α-cyano-4-hydroxy-cinnamic acid 20 mg/mL in ACN/5 % formic acid in water (70/30) and 1 part 2,5-dihydroxybenzoic acid 20 mg/mL in ACN/1 % TFA in water (70/30) [30]. Then, the mixtures were left in the fume hood for cocrystallization (ca. 15 min).

The instrument used for the measurements was a rapifleX MALDI-TOF/TOF mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). It is equipped with a scanning smart beam 10 kHz Nd:YAG laser at a wavelength of 355 nm and a 10 bit 5 GHz digitizer. The acceleration voltage was set to 20 kV and the mass spectra were recorded in positive ion linear mode. Calibration was done with the Bruker protein calibration standard II in a mass range from 10 to 70 kDa. Samples were measured at a laser power of 70–100 % with random walk ionization across the sample spot. Rhodesain incubated with compound **K11777**, a known covalent irreversible inhibitor, was used as a control sample.

Data analysis was performed using the open-source software mMass [31]. Since $[M+H]^+$ was evaluated, the spectra were cropped to the range of 22–27 kDa, and the signal intensity was normalized to the highest signal in this range. The baseline was corrected using standard settings. The mass shifts ($\Delta m/z$) are calculated from the rhodesain signal in the negative control subtracted from the adduct peak in the respective sample, if present, and compared to the calculated exact mass of the inhibitor (m_{ex}).

Non-covalent docking. Non-covalent docking using the crystal structure of rhodesain with covalently bound inhibitor **K11777** (pdb: 2p7u) was performed to reveal the affinity and geometry of the pre-reacted enzyme-inhibitor complexes for all the inhibitors. The inhibitors were built based on the **K11777** structure using Discovery Studio software [32]. The binding site was defined as a gridbox (12.75, 19.50 and 14.25 Å) around bound **K11777**. Water molecules present in the crystal structure were omitted. Docking calculations were executed with Autodock Vina 1.2.5 [33,34]. Visualization of selected docking poses was performed with Autodock Tools 1.5.7 [35]. The resulting complexes were analyzed using VMD visualization software 1.9.4a53 [36].

Covalent docking. Covalent docking using the molecular operating environment (MOE) [37] was performed as described previously [20,23]. Briefly, the rhodesain K11777 complex structure (pdb: 2p7u) [38] was protonated in MOE for physiological and assay pH of 5.5. The covalently bound ligand was untethered from Cys-25 and the catalytic diad protonated to form an ion pair. The covalent molecular docking setup was validated by re-docking of K11777 generating 30 poses and using the London dG-scoring function and induced fit refinement to yield 5 final poses (top1 pose: docking score = − 16.49 and RMSD = 1.4 Å) for further visual inspection. Ligands of interest **11**, **15** and **19** were protonated for pH = 5.5 resulting in positively charged piperazine moieties and energetically minimized with the MMFF94x force field [39] using MOE. Docking was performed under identical conditions as K11777 using the Michael addition reaction (**11**, **15**, **19**) and the reversible carbonyl reaction to form a hemithioacetal (**15**, **19**). The resulting complexes were analyzed using PyMOL visualization software 3.0.3. [40]. Best scoring poses of **19** resembling the interactions of K11777 were subjected to MD simulations.

Molecular dynamic simulations. MD simulations were performed as described previously [41,42]. Briefly, non-covalent complex structures of **19** based non-covalent docking, and Michael addition and hemithioacetal-forming covalent docking and untethering the covalent bond. Likewise, for comparison K11777 was also untethered and subjected to MD. Rhodesain was protonated using MOE resulting in standard amino acid protonation states except the catalytic dyad forming the ion pair between His-162 (HIP) and Cys-25 (CYM), protonated Asp-57 (ASH) as described previously [43,44] and protonated His-115 (HIP). Ligands were parameterized using the Generalized Amber Force Field (GAFF2) [45] with AM1-BCC charges [46] using antechamber [47] of the AmberTools24 [48,49]. Complexes including crystallographic water molecules were built with tleap and minimized over 200-time steps with sander prior neutralization with Na+-ions and adding a TIP3P [50] waterbox exceeding the protein–ligand complex by 10.0 Å in every dimension. MD simulations were performed with NAMD2.14 [51] and the AMBER force field (ff19SB) [52]. After 1 ns of equilibration using a 100–300 K heating protocol with gradually decreasing harmonic constraints on protein and ligand atoms, 50 ns of production were performed with an NPT ensemble using periodic boundary conditions and a vdW cut-off at 14.0 Å. Rigid bond length allowed time steps of 2 fs. Simulations were analyzed using VMD-1.9.3.[36] 1-trajectory (1A) MM/GBSA calculations were performed using MMPBSA.py [53] for 50 ns simulations with an interval of 0.1 ns, GB model igb5 [54] from AmberTools and a salt concentration of 0.1 M. If different ligand

conformations were observed during MD, time frames were adapted accordingly (Table 3).

Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

CRediT authorship contribution statement

Laura Agost-Beltrán: Writing – original draft, Methodology, Investigation. **Collin Zimmer:** Investigation, Data curation. **Hans Joachim Räder:** Investigation. **Christian Kersten:** Investigation. **Tanja Schirmeister:** Supervision, Funding acquisition. **Santiago Rodríguez:** Supervision, Funding acquisition. **Florenci V. González:** Writing – original draft, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This research was funded by the Generalitat Valenciana (PROMETEO with ref. CIPROM/2021/079) and Universitat Jaume I (UJI-B2020-03, UJI-B2021-71 and SomUJIcontracovid crowdfunding campaign). L.A. acknowledges a PhD fellowship from Ministerio de Universidades (FPU19/04913).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bioorg.2024.107830.

References

[1] Neglected tropical diseases. Accessed September 5, 2024. https://www.who.int/news-room/questions-and-answers/item/neglected-tropical-diseases.
[2] R. Brun, J. Blum, F. Chappuis, C. Burri, Human African trypanosomiasis, Lancet 375 (2010) 148–159, https://doi.org/10.1016/S0140-6736(09)60829-1.
[3] J.M. Gao, Z.Y. Qian, G. Hide, D.H. Lai, Z.R. Lun, Z.D. Wu, Human African trypanosomiasis: The current situation in endemic regions and the risks for non-endemic regions from imported cases, Parasitology. 147 (9) (2020) 922–931, https://doi.org/10.1017/S0031182020000645.
[4] A. Mullard, FDA approves first all-oral sleeping sickness drug, Nat. Rev. Drug Discov. 20 (9) (2021) 658, https://doi.org/10.1038/d41573-021-00140-5.
[5] Nascimento IJ dos S, Cavalcanti M de AT, de Moura RO. Exploring N-myristoyltransferase as a promising drug target against parasitic neglected tropical diseases. *Eur. J. Med. Chem.* (2023) 258. doi:10.1016/j.ejmech.2023.115550.
[6] R. Ettari, S. Previti, L. Tamborini, G. Cullia, S. Grasso, M. Zappalà, The inhibition of cysteine proteases rhodesain and TbCatB: A valuable approach to treat human African trypanosomiasis, Mini Rev. Med. Chem. 16 (17) (2016) 1374–1391, https://doi.org/10.2174/1389557515666160509125243.
[7] D. Steverding, D.W. Sexton, X. Wang, S.S. Gehrke, G.K. Wagner, C.R. Caffrey, Trypanosoma brucei: chemical evidence that cathepsin L is essential for survival and a relevant drug target, Int. J. Parasitol. 42 (5) (2012) 481–488, https://doi.org/10.1016/j.ijpara.2012.03.009.
[8] A. Latorre, T. Schirmeister, J. Kesselring, et al., Dipeptidyl nitroalkenes as potent reversible inhibitors of cysteine proteases rhodesain and cruzain, ACS Med. Chem. Lett. 7 (12) (2016) 1073–1076, https://doi.org/10.1021/acsmedchemlett.6b00276.
[9] J.T. Palmer, D. Rasnick, J.L. Klaus, D.Y. Bromme, Vinylsulfones as mechanism-based cysteine protease inhibitors, J. Med. Chem. 38 (1995) 3193–3196.
[10] T. Schirmeister, J. Kesselring, S. Jung, et al., Quantum chemical-based protocol for the rational design of covalent inhibitors, J. Am. Chem. Soc. 138 (27) (2016) 8332–8335, https://doi.org/10.1021/jacs.6b03052.
[11] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. L. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors, Science 368 (2020) 409–412, https://doi.org/10.1126/science.abb3405.
[12] Lopes F, Santos MMM, Moreira R. Designing covalent inhibitors: A medicinal chemistry challenge. *Biomed. Chem.*, 2012, Chapter 1.2, 44-59. doi: 10.1515/9783110468755-002.
[13] N.V. Mehta, M.S. Degani, The expanding repertoire of covalent warheads for drug discovery, Drug Discov. Today 28 (12) (2023) 103799–103837, https://doi.org/10.1016/j.drudis.2023.103799.
[14] B.E. Cathers, C. Barrett, J.T. Palmer, R.M. Rydzewski, pH dependence of inhibitors targeting the occluding loop of cathepsin B, Bioorg Chem. 30 (4) (2002) 264–275, https://doi.org/10.1016/S0045-2068(02)00009-3.
[15] Chérest M, Felkin H, and Prudent N. Torsional strain involving partial bonds. The stereochemistry of the lithium aluminium hydride reduction of some simple open-chain ketones. *Tetrahedron Lett.* 1968, 18, 2199–2204. doi:https://doi.org/10.1016/S0040-4039(00)89719-1.
[16] N.T. Anh, Regio- and stereo-selectivities in some nucleophilic reactions, Top. Curr. Chem. 88 (24) (1980) 145–162, https://doi.org/10.1007/BFb0048506.
[17] S. Liang, K. Hofman, M. Friedrich, J. Keller, G. Manolikakes, Recent progress and emerging technologies towards a sustainable synthesis of sulfones, ChemSusChem 14 (22) (2021) 4878–4902, https://doi.org/10.1002/cssc.202101635.
[18] S. Ludewig, M. Kossner, M. Schiller, K. Baumann, T. Schirmeister, Enzyme kinetics and hit validation in fluorimetric protease assays, Curr. Top. Med. Chem. 10 (3) (2010) 368–382, https://doi.org/10.2174/156802610790725498.
[19] P. Klein, F. Barthels, P. Johe, et al., Naphthoquinones as covalent reversible inhibitors of cysteine proteases - studies on inhibition mechanism and kinetics, Molecules 25 (9) (2020) 2064, https://doi.org/10.3390/molecules25092064.
[20] P. Müller, M. Meta, J.L. Meidner, et al., Investigation of the compatibility between warheads and peptidomimetic sequences of protease inhibitors—a comprehensive reactivity and selectivity study, Int. J. Mol. Sci. 24 (8) (2023) 7226, https://doi.org/10.3390/ijms24087226.
[21] K. Arafet, F.V. González, V. Moliner, Elucidating the dual mode of action of dipeptidyl enoates in the inhibition of rhodesain cysteine proteases, Chem. Eur. J. 27 (39) (2021) 10142–10150, https://doi.org/10.1002/chem.202100892.
[22] S. Previti, R. Ettari, E. Calcaterra, et al., Structure-based lead optimization of peptide-based vinyl methyl ketones as SARS-CoV-2 main protease inhibitors, Eur. J. Med. Chem. 247 (2023) 115021, https://doi.org/10.1016/j.ejmech.2022.115021.
[23] S. Jung, N. Fuchs, P. Johe, et al., Fluorovinylsulfones and -sulfonates as potent covalent reversible inhibitors of the trypanosomal cysteine protease rhodesain: structure-activity relationship, inhibition mechanism, metabolism, and in vivo studies, J. Med. Chem. 64 (16) (2021) 12322–12358, https://doi.org/10.1021/acs.jmedchem.1c01002.
[24] H.A. Rodríguez, F.M. Bickelhaupt, I. Fernández, Origin of the Bürgi-Dunitz angle, ChemPhysChem. 24 (17) (2023) 1–6, https://doi.org/10.1002/cphc.202300379.
[25] W.L.F. Armarego, D.D. Perrin, Purification of Laboratory Chemicals, 4th edition, Butterworth-Heinemann, 2000.
[26] N. Fuchs, et al., Subnanomolar cathepsin S Inhibitors with high selectivity: optimizing covalent reversible α-fluorovinylsulfones and α-sulfonates as potential immunomodulators in cancer, ChemMedChem 18 (2023) e202300160.
[27] C.R. Caffrey, E. Hansell, K.D. Lucas, et al., Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of Trypanosoma brucei rhodesiense, Mol. Biochem. Parasitol. 118 (1) (2001) 61–73, https://doi.org/10.1016/S0166-6851(01)00368-1.
[28] F.J. Medrano, S. de la Hoz-Rodríguez, S. Martí, et al., Peptidyl nitroalkene inhibitors of main protease rationalized by computational and crystallographic investigations as antivirals against SARS-CoV-2, Commun. Chem. 7 (1) (2024) 1–16, https://doi.org/10.1038/s42004-024-01104-7.
[29] C. Zimmer, J. Brauer, D. Ferenc, et al., Substitution-induced mechanistic switching in SNAr-warheads for cysteine proteases, Molecules 29 (11) (2024) 1–16, https://doi.org/10.3390/molecules29112660.
[30] S. Laugesen, P. Roepstorff, Combination of two matrices results in improved performance of MALDI MS for peptide mass mapping and protein analysis, J. Am. Soc. Mass Spectrom. 14 (9) (2003) 992–1002, https://doi.org/10.1016/S1044-0305(03)00262-9.
[31] M. Marinas, E. Sa, M.M. Rojas, et al., A nuclear magnetic resonance (1H and 13 C) and isotope ratio mass spectrometry (d 13 C, d 2 H and d 18 O) study of Andalusian olive oils, Rapid Commun. Mass Spectrom. 24 (2010) 1457–1466, https://doi.org/10.1002/rcm.
[32] Discovery Studio: BIOVIA, Dassault Systèmes, BIOVIA Discovery Studio 2021 Client, 21.1.0, San Diego: Dassault Systèmes, 2021.
[33] J. Eberhardt, D. Santos-Martins, A.F. Tillack, S. Forli, AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings, J. Chem. Inf. Model. 61 (8) (2021) 3891–3898, https://doi.org/10.1021/acs.jcim.1c00203.
[34] O.T. Arthur, J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461, https://doi.org/10.1002/jcc.21334.
[35] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, Autodock tools: AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791, https://doi.org/10.1002/jcc.21256.
[36] W. Humphrey, A. Dalke, K.V.M.D. Schulten, Visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38, https://doi.org/10.1016/0263-7855(96)00018-5.
[37] Molecular Operating Environment (MOE) 2022.02. Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7 2022.
[38] I.D. Kerr, J.H. Lee, C.J. Farady, et al., Vinyl sulfones as antiparasitic agents and a structural basis for drug design, J. Biol. Chem. 284 (38) (2009) 25697–25703, https://doi.org/10.1074/jbc.M109.014340.

[39] Halgren, T. A. Merck molecular force field V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comput. Chem.* 1996, 17 (5−6), 616−641. doi:10.1002/(SICI)1096-987X(199604)17:5/6<616::AID-JCC5>3.0.CO;2-X.
[40] PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.
[41] S.J. Hammerschmidt, S. Huber, N.J. Braun, M. Lander, T. Steinmetzer, C. Kersten, Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs, Arch. Pharm. (Weinheim) 356 (4) (2023), https://doi.org/10.1002/ardp.202200518.
[42] C. Kersten, E. Fleischer, J. Kehrein, et al., How to design selective ligands for highly conserved binding sites: A case study using N-myristoyltransferases as a model system, J. Med. Chem. 63 (5) (2020) 2095–2113, https://doi.org/10.1021/acs.jmedchem.9b00586.
[43] P. Johe, S. Jung, E. Endres, et al., Warhead reactivity limits the speed of inhibition of the cysteine protease rhodesain, ACS Chem. Biol. 16 (4) (2021) 661–670, https://doi.org/10.1021/acschembio.0c00911.
[44] P. Johé, E. Jaenicke, H. Neuweiler, T. Schirmeister, C. Kersten, U.A. Hellmich, Structure, interdomain dynamics, and pH-dependent autoactivation of pro-rhodesain, the main lysosomal cysteine protease from African trypanosomes, J. Biol. Chem. 296 (2021) 100565, https://doi.org/10.1016/j.jbc.2021.100565.
[45] D. Vassetti, M. Pagliai, P. Procacci, Assessment of GAFF2 and OPLS-AA general force fields in combination with the water models TIP3P, SPCE, and OPC3 for the solvation free energy of druglike organic molecules, J. Chem. Theory Comput. 15 (3) (2019) 1983–1995, https://doi.org/10.1021/acs.jctc.8b01039.
[46] A. Jakalian, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, J. Comput. Chem. 23 (16) (2002) 1623–1641, https://doi.org/10.1002/jcc.10128.
[47] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, J. Mol. Graph. Model. 25 (2) (2006) 247–260, https://doi.org/10.1016/j.jmgm.2005.12.005.
[48] D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D. S. Cerutti, T.E. Cheatham III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, N. Forouzesh, M. Ghazimirsaeed, G. Giambaşu, T. Giese, M.K. Gilson, H. Gohlke, A. W. Goetz, J. Harris, Z. Huang, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, A. Kovalenko, T. Kurtzman, T.S. Lee, P. Li, Z. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, A. Rahnamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C. L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, J. Wang, X. Wu, Y. Wu, Y. Xiong, Y. Xue, D.M. York, C. Zhao, Q. Zhu, P.A. Kollman, Amber 2024, University of California, San Francisco, 2024.
[49] D.A. Case, H.M. Aktulga, K. Belfon, D.S. Cerutti, G.A. Cisneros, V.W.D. Cruz eiro, N. Forouzesh, T.J. Giese, A.W. Götz, H. Gohlke, S. Izadi, K. Kasavajhala, M.C. Kaymak, E. King, T. Kurtzman, T.-S. Lee, P. Li, J. Liu, T. Luchko, R. Luo, M. Manathunga, M. R. Machado, H.M. Nguyen, K.A. O'Hearn, A.V. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, A. Risheh, S. Schott-Verdugo, A. Shajan, J. Swails, J. Wang, H. Wei, X. Wu, Y. Wu, S. Zhang, S. Zhao, Q. Zhu, T.E. Cheatham III, D.R. Roe, A. Roitberg, C. Simmerling, D.M. York, M.C. Nagan*, and K.M. Merz Jr.* AmberTools. *J. Chem. Inf. Model.63*, 2023, 6183-6191.
[50] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (2) (1983) 926–935, https://doi.org/10.1063/1.445869.
[51] J.C. Phillips, R. Braun, W. Wang, et al., Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802, https://doi.org/10.1002/jcc.20289.
[52] C. Tian, K. Kasavajhala, K.A.A. Belfon, et al., Ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution, J. Chem. Theory Comput. 16 (1) (2020) 528–552, https://doi.org/10.1021/acs.jctc.9b00591.
[53] B.R. Miller, T.D. Mcgee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA. py: An efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (9) (2012) 3314–3321, https://doi.org/10.1021/ct300418h.
[54] R. Anandakrishnan, A. Drozdetski, R.C. Walker, A.V. Onufriev, Speed of conformational change: comparing explicit and implicit solvent molecular dynamics simulations, Biophys J. 108 (5) (2015) 1153–1164, https://doi.org/10.1016/j.bpj.2014.12.047.

Supporting information

Rhodesain inhibitors on the edge of reversibility-irreversibility

Laura Agost-Beltrán [a], *Collin Zimmer* [b], *Hans Joachim Räder* [c], *Christian Kersten* [b,d], *Tanja Schirmeister* [b], *Santiago Rodríguez* [a], *and Florenci V. González* [a,*]

[a] Departament de Química Inorgànica i Orgànica, Universitat Jaume I, 12071 Castelló de la Plana, Spain

[b] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany

[c] Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

[d] Institute for Quantitative and Computational Biosciences, Johannes Gutenberg-University, BioZentrum I, Hanns-Dieter-Hüsch.Weg 15, 55128 Mainz, Germany

5. MALDI-TOF spectra

MALDI is a soft ionization technique that can discriminate between covalent and non-covalent binding, as the formation of non-covalent adducts can be suppressed by using an acidic matrix. This fact is attributed to the disruption of salt bridges and subsequent destabilization of the non-covalent protein-ligand complexes.[2] This technique has been previously used for verifying the covalency of similar covalent-reversible compounds.[3]

In the resulting spectra (**S19** to **S27**), it can be observed a peak corresponding to rhodesain (around 23.3 kDa)[3] or/and, in case of covalent adduct formation, a peak that shows a mass shift corresponding to the mass of the compound bound to rhodesain. In most measurements, rhodesain is detected with one dominant and some minor signals; this multiplicity can be due to adducts with matrix.[3–5] Another possible explanation for the multiplicity would be general heterodispersity in polymeric samples.

5.1. Matrix mixture: 1 part α-cyano-4-hydroxy-cinnamic acid with 1 part 2,5-dihydroxybenzoic acid

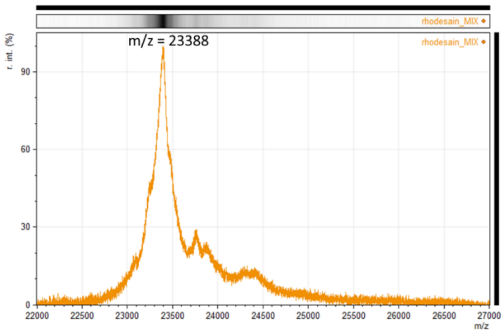


Figure S19. Rhodesain without inhibitor (negative control).

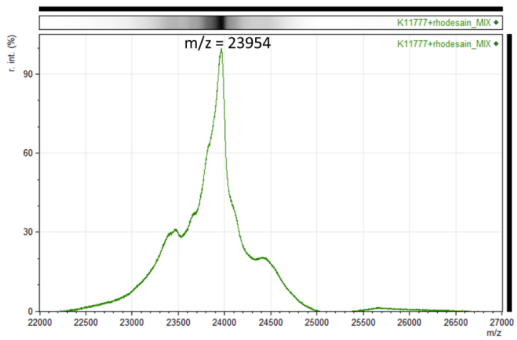


Figure S20. Adduct between rhodesain and **K11777** (Δm/z found = 566 Da (m_{ex} = 574 Da)).

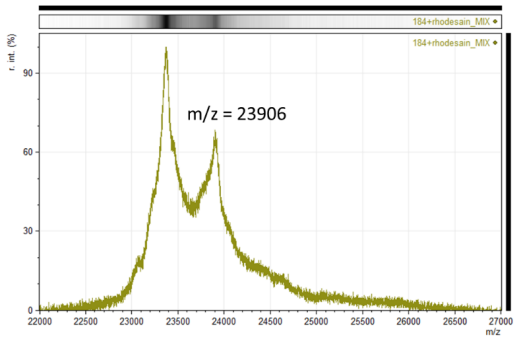


Figure S21. Adduct between rhodesain and **15** (Δm/z found = 518 Da (m_{ex} = 534 Da)).

5.2. Matrix: Sinapinic acid

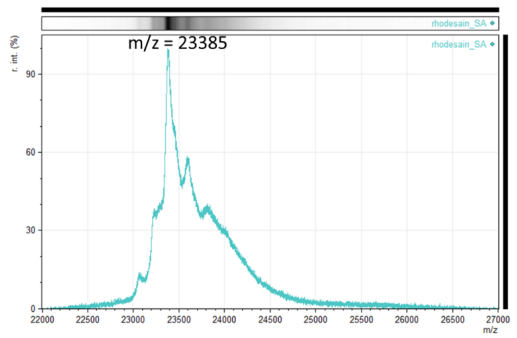


Figure S22. Rhodesain without inhibitor (negative control).

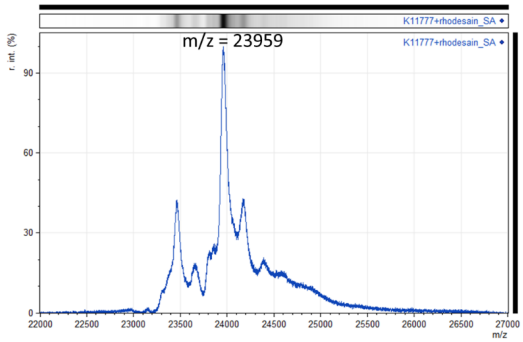


Figure S23. Adduct between rhodesain and **K11777** (Δm/z found = 574 Da (m_{ex} = 574 Da)).

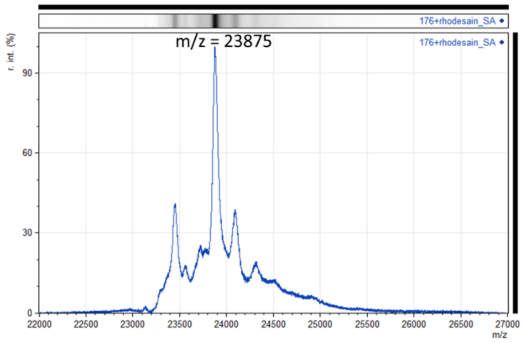


Figure S24. Adduct between rhodesain and **11** (Δm/z found = 490 Da (m_{ex} = 506 Da)).

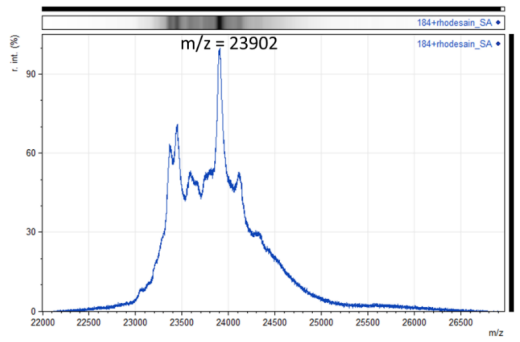


Figure S25. Adduct between rhodesain and **15** (Δm/z found = 517 Da (m_{ex} = 534 Da)).

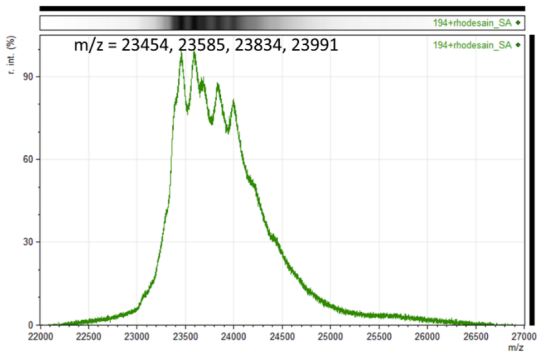


Figure S26. Adduct between rhodesain and **19** (Δm/z found = 606 (m_{ex} = 602 Da)).

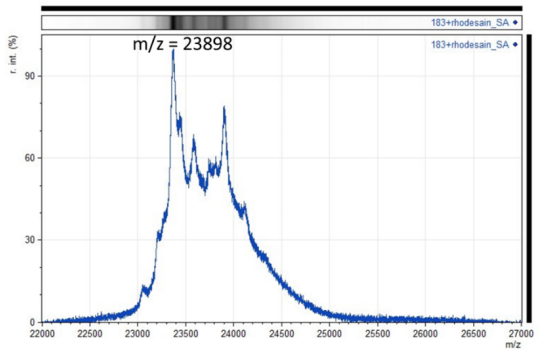


Figure S27. Adduct between rhodesain and **14** (Δm/z found = 513 Da (m_{ex} = 536 Da)).

6. Docking

Docking of alcohols **14** and **18** did not provide any rational pose in which the residues of the ligand were occupying the pockets (S1' to S3) of the enzyme and the electrophilic carbon was near Cys25.

The previously co-crystallized reference ligand **K11777** was redocked with rhodesain (pdb: 2p7u)[6] to prove that the docking procedures could generate reasonable binding modes. As a result, the non-covalent docking provided successful results (**Figure S28**).

"

3: Substitution-Induced Mechanistic Switching in S_NAr-Warheads for Cysteine Proteases.

The principal mode of inhibition by S_NAr warheads culminates in the irreversible substitution of a leaving group, usually a halogen. This has been investigated for inhibitors of a wide variety of enzymes with relevantly reactive cysteines, e.g., proteases and kinases. The underlying reaction path for this nucleophilic substitution crosses a non-covalent π- and a covalent σ-complex prior to the elimination of the halogen. The required number for electron-withdrawing residues additional to the leaving group on these warheads is commonly one or two depending on the targeted nucleophile.[125,194,195] Characteristic is the irreversible nature of these inhibitors. However, in a previous study, an inhibitor with fluorodinitrobenzene warhead positioned at the *N*-terminus of a dipeptide inhibitor with unchanged *C*-terminus was described to arrest the π-complex as a stable reaction intermediate.[251] It was described that it had pronounced effect against rhodesain, and that only its ester prodrug was active in antitrypanosomal cell culture experiments.

To probe the S_NAr reaction path in this study, different electron-deficient arenes were implemented on a peptide backbone with proven affinity towards rhodesain, which was used as the target protease. All potential inhibitors were characterized by their in vitro affinity. This showed that two electron-withdrawing substituents are required for this warhead to show inhibition. It also elucidated the possible level of impact by hetarene analogy, halogen exchange and the exact positioning of the warhead. With one of the warheads that showed inhibition, a retro-inverso approach was designed for the peptide sequence to gain insight on the influence of the backbone, which elucidated a limitation of applicability to the initial peptide sequence. For inhibitors **2**, **3**, and **9** with the most interesting inhibition profiles, the modes of inhibition were elucidated in more detail. The inhibition results were combined with data on covalency by MALDI-TOF mass spectrometry to distinguish the non-covalent π- from the two covalent states. Reversibility was characterized by a dilution assay and apparent kinetics in the inhibition assay to distinguish the full substitution reaction from the reversible states. This was complemented by quantum mechanical computations of the most probable reaction progresses to identify the possibility for stable σ-complexes. Shown by the employed binding mode characterization techniques, the three stable intermediates of the S_NAr reaction could be captured by the inhibitors, namely a non-covalent reversible dinitrophenyl-substituted peptide (π-complex), a covalent reversible fluorodinitrophenyl-substituted peptide (σ-complex), and a covalent irreversible fluorodinitrobenzoyl-substituted peptide (fluoride as the leaving group). For **2**, **3**, and **9**, the selectivity towards the related human cathepsins B and L was described, which was similarly high as for the previously described non-covalent inhibitor that shared the peptide recognition sequence. The abovementioned discrepancy in the past study between inhibition of rhodesain in vitro (free *C*-terminus) and antitrypanosomal activity (esterified *C*-terminus) was experimentally assessed by describing the impact on passive permeability of these modifications on the scaffold under investigation. Indeed, there was a large improvement of permeability through

esterification, as assessed by parallel artificial membrane permeation assay, which rationalizes the published observations experimentally in the context of esters as suitable prodrug modifications for carboxylic acid moieties.[616] Taken together, this study expanded the data available on S_NAr warheads employed against cysteine proteases, by using complementary methods that together can elucidate the crucial factors of covalency and reversibility for electrophilic warheads. It could also shed light on optimization requirements for cell experiments in the context of increasing permeability.

Own contribution: Heterologous expression of rhodesain, characterization of the in vitro affinity, inhibition kinetics, and selectivity, including inner-filter effect correction, reversibility assessment by dilution assay, characterization of the covalency of inhibitor binding by MALDI-TOF-MS, reactivity assessment towards a model thiol in an aqueous solution, non-covalent docking, assessment of passive permeation by PAMPA, writing and editing of the manuscript.

Contribution from others: Inhibitor synthesis, reactivity study against a model thiolate in an organic solvent, quantum mechanical calculation of the reaction progress, writing and editing of the manuscript.

This work was published in Molecules.

Article reprinted with permission of *Molecules* **2024**, *29*(11), 2660, “Substitution-Induced Mechanistic Switching in S_NAr-Warheads for Cysteine Proteases”, Copyright © 2024 The Authors. Published by MDPI.

The following publication quoted within “” (pages 140–173) is the same as the manuscript cited here. The appended Supporting Information represents an abridged version. The full version can be accessed online at doi:10.3390/molecules29112660.

“

molecules

MDPI

Article

Substitution-Induced Mechanistic Switching in S_NAr-Warheads for Cysteine Proteases

Collin Zimmer [1,†], Jan Brauer [2,†], Dorota Ferenc [2], Jessica Meyr [3], Patrick Müller [1], Hans-Joachim Räder [4], Bernd Engels [3], Till Opatz [2,*] and Tanja Schirmeister [1,*]

[1] Institute of Pharmaceutical and Biomedical Sciences, University of Mainz, Staudingerweg 5, 55128 Mainz, Germany; cozimmer@uni-mainz.de (C.Z.); muelpat@uni-mainz.de (P.M.)
[2] Department of Chemistry, University of Mainz, Duesbergweg 10-14, 55128 Mainz, Germany; j.brauer@uni-mainz.de (J.B.); ferenc@uni-mainz.de (D.F.)
[3] Institute of Physical and Theoretical Chemistry, Julius-Maximilians-University Würzburg, Am Hubland, 97074 Würzburg, Germany; jessica.meyr@uni-wuerzburg.de (J.M.); bernd.engels@uni-wuerzburg.de (B.E.)
[4] Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany; raeder@mpip-mainz.mpg.de
* Correspondence: opatz@uni-mainz.de (T.O.); schirmei@uni-mainz.de (T.S.)
† These authors contributed equally to this work.

Abstract: The aim of this study was to investigate the transition from non-covalent reversible over covalent reversible to covalent irreversible inhibition of cysteine proteases by making delicate structural changes to the warhead scaffold. To this end, dipeptidic rhodesain inhibitors with different *N*-terminal electrophilic arenes as warheads relying on the S_NAr mechanism were synthesized and investigated. Strong structure–activity relationships of the inhibition potency, the degree of covalency, and the reversibility of binding on the arene substitution pattern were found. The studies were complemented and substantiated by molecular docking and quantum-mechanical calculations of model systems. Furthermore, the improvement in the membrane permeability of peptide esters in comparison to their corresponding carboxylic acids was exemplified.

Keywords: electrophilic warhead; S_NAr; rhodesain; covalency; reversibility; permeability

check for updates

Citation: Zimmer, C.; Brauer, J.; Ferenc, D.; Meyr, J.; Müller, P.; Räder, H.-J.; Engels, B.; Opatz, T.; Schirmeister, T. Substitution-Induced Mechanistic Switching in S_NAr-Warheads for Cysteine Proteases. *Molecules* **2024**, *29*, 2660. https://doi.org/10.3390/molecules29112660

Academic Editor: Igor Marques

Received: 25 April 2024
Revised: 15 May 2024
Accepted: 20 May 2024
Published: 4 June 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rhodesain (*Tb*CatL), a key cysteine protease of the deadly human parasite *Trypanosoma brucei*, is considered a validated drug target for the treatment of human African trypanosomiasis [1–3]. This infectious disease, caused by two regional subspecies of the pathogen, still poses a burden to countries, mainly in rural central Africa, where 500–1000 people are newly diagnosed per year and three million people are at risk of contracting the disease [4]. In the past decades, the WHO has taken significant efforts to tackle this problem: while chemotherapy has proven effective on the individual level, vector control targeting the transmitting tsetse fly has been effective in controlling the disease spread. This progressed its “elimination as a public health problem” with the goal of “elimination of transmission” by 2030 [5]. The ongoing research and clinical trials on this disease still result in the discovery and even approval of novel drugs, and remarkable advances have been achieved in the last five years in the context of efficacy, application, and side effects (fexinidazole, approved in 2021, and acoziborole, in phase II/III clinical trials) [6,7].

Protease inhibition in general is a relevant contributor to clinical disease management, especially for viral infectious diseases such as hepatitis C, acquired immune deficiency syndrome (AIDS), and COVID-19 [8,9]. In research, the concept is also employed for autoimmune diseases, and different types of cancer [10]. Due to the high similarity of the binding sites of cysteine proteases of the papain family (e.g., cathepsins L and S), rhodesain is an interesting model system for mechanistic investigations on other targets. It can be

produced in large quantities in a straightforward fashion using standard molecular biology methods [11], and in contrast to many commercially available enzymes, it is exceptionally stable and retains catalytic activity for extended periods and under a variety of conditions. On a structural level, it possesses the precatalytically deprotonated active-site cysteine of the CA clan proteases, a highly reactive nucleophile [12].

To address this active-site nucleophile, there are numerous reports on covalent rhodesain inhibitors, but data on non-covalent ones are scarce, especially on a mechanistic level [13–15]. In a previous project on this topic, fluorine- and nitro-substituted arenes as warheads in combination with a dipeptidic binding motif suitable for rhodesain [16,17] were identified as a novel class of inhibitors. The S_NAr-reaction path—the underlying principle of the inhibition mode of these compounds—is characterized by multiple stages, starting with the formation of a non-covalent π-complex between nucleophile and electrophile, followed by the formation of a covalent Meisenheimer-type anion (σ-complex) [18]. The anion can be oxidized, yielding the so-called Zimmermann product; the substitution reaction can take place by elimination of a leaving group, or the addition can revert back to the original reactants via the π-complex. The application of electron-deficient arenes in the context of addressing (non-)catalytic cysteines irreversibly through a substitution reaction has been described on different targets, e.g., peroxisome proliferator-activated receptors (PPARs) [19], the bacterial enzyme sortase A (SrtA) [20], the fibroblast growth factor receptor 4 (FGFR4) [21], S6 kinase β2 (S6K2) [22], and DNA methyltransferase 2 (DNMT2) [23]. While the electrophiles look similar to the ones described for rhodesain, the inhibition mechanism differs substantially. The previously identified arene-based rhodesain inhibitors surprisingly act via a non-covalent mechanism [17,24]. One of these compounds, specifically with a free C-terminus, was identified as a highly affine and selective inhibitor of rhodesain [17]. Initial mechanistic experiments and theoretical investigations pointed to a nucleophilic attack onto the arene that only progresses until the π-, rather than the σ-complex [17]. MALDI-TOF mass spectrometry (MS) later provided further experimental validation that the reaction stops at the non-covalent π-complex [24]. Its effect is supported by anti-trypanosomal data from a cell-based assay and has been explained by the prodrug concept in the context of cell permeability [17].

As a follow-up, we now conducted a systematic structure activity relationship (SAR) study to further elucidate the criteria for affinity and to assess the prerequisite electrophilicity necessary to stop the described reaction on either intermediate on the reaction path to a complete (and irreversible) S_NAr reaction. The inhibition of rhodesain was assessed using classical fluorimetric substrate displacement assays. The covalency of binding was investigated using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) with an acidic matrix to suppress the formation of non-covalent complexes. The (ir)reversibility of inhibition was shown, and the experimental findings on covalency and reversibility were substantiated by quantum-mechanical (QM) calculations and docking to suggest explanations on a molecular level. For the most interesting compounds, selectivity against related human cysteine proteases was investigated. Furthermore, experimental reasoning for the observed anti-trypanosomal activity of ester derivates in a cellular context is given with results from a parallel artificial membrane permeability assay (PAMPA).

2. Results

2.1. Design and Synthesis

To investigate binding orientation and inhibition modes, two strategies were followed:

(A) We used the general structure of Ar-Phe-Leu-OH, relating to the dipeptide phenylalanine-leucine, with "Ar" being an electron-deficient arene attached to the Phe-*N*-terminus, and "-OH" signifying the free carboxylic acid at the *C*-terminal leucine (Figure 1a). With this, we expanded the SAR study for the arene moiety of the not-yet-explored acid counterparts of the previously described esters (compounds **1–12**) [17]. Starting from an arene with only one electron-withdrawing substituent, the influence of substituent

number, identity, and position, and the presence of an adequate leaving group were investigated in an effort to observe differences in interaction with the nucleophilic thiolate in rhodesain.

(B) Employing the concept of retro-inverso peptides [25,26], we combined homophenylalanine (hPhe, a strong interactor with rhodesain's S1 subpocket) and Ala (an expected weak interactor) in an inverted sequence and chirality. With this, we wanted to further probe the directionality of ligand binding (compounds **13–16**). Additionally, we transferred the motif of *N*-terminal, electron-deficient arene and unprotected *C*-terminus onto the dipeptide of the known irreversible pan-cathepsin inhibitor K11777 [27] (Phe-hPhe) and the respective retro-inverso derivate (compounds **17–18**).

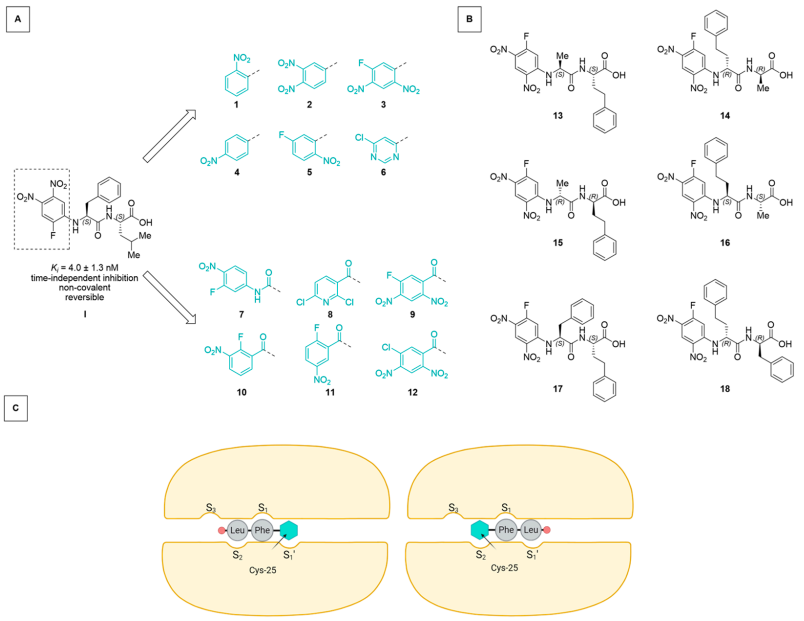
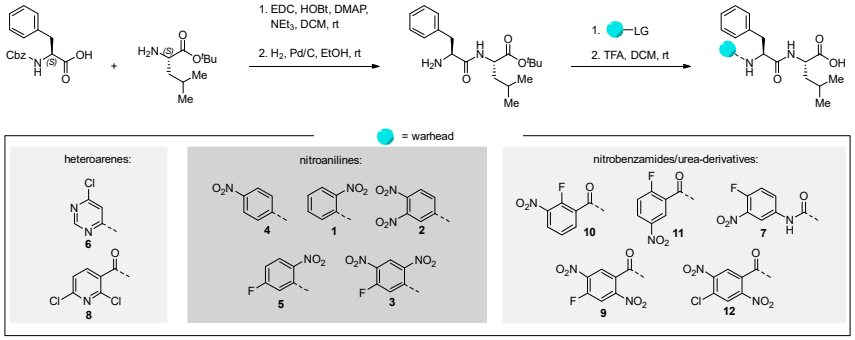


Figure 1. (**A**) Starting from the lead compound **I** [17], an SAR study was performed investigating the effect of the electrophilicity and leaving group properties of the *N*-terminal arene "Ar" of the Ar-Phe-Leu-OH scaffold. (**B**) Dipeptidic motifs (based on Ala-hPhe and Phe-hPhe) with inversed sequence and/or chirality (retro-inverso) were employed to assess the binding directionality. (**C**) Schematic representation of two possible orientations in which the electrophilic arene interacts with the catalytic Cys-25 (inhibitor *C*-terminus as red circle, arene as cyan hexagon). Protein subpocket nomenclature after Schechter and Berger [28] (created using BioRender.com).

All inhibitor candidates were synthesized using known strategies [17]. We started from the commercially available *N*-(carbobenzyloxy)-L-phenylalanine and L-leucine *tert*-butyl ester hydrochloride, which were linked using standard coupling procedures. After hydrogenolytic deprotection of the *N*-terminus, the warheads were attached using different reaction conditions (for further information, we refer to the Supporting Information (SI)). The free carboxylic acid was then obtained via trifluoroacetic acid-promoted ester cleavage (Scheme 1).



Scheme 1. Synthesis of the tested peptides. Yields and detailed reaction conditions can be found in the SI.

We tried to utilize solid-phase synthesis for the preparation of the homophenylalanine-bearing dipeptides. However, this strategy suffered from low yields for two substrates and failed to deliver the envisioned product in one case. We therefore reverted to the previous approach, which worked for all depicted molecules in good yields over four steps. The detailed procedures for the preparation of these compounds can be found in the SI.

2.2. Inhibition Assay

Inhibition of rhodesain was assessed using Z-Phe-Arg-AMC as a fluorogenic substrate [29]. IC_{50} or K_I^{app} were determined depending on the time dependency of the progress curves. Because a competitive mode of inhibition can be assumed [17], K_i (for reversible inhibition) and K_I (for irreversible inhibition) were calculated using the Cheng–Prusoff equation [30]. Selectivity towards human cathepsins L and B (*Hs*CatL, *Hs*CatB) was assessed for the most active compounds. The results can be found in Table 1.

Table 1. Overview of inhibition data (rhodesain) for synthesized compounds. Results are averages including standard deviation of at least two replicates for all compounds.

Compound	K_i or K_I [nM]	Compound	K_i or K_I [nM]
1	$(39 \pm 4) \times 10^3$	**10**	$(14 \pm 2) \times 10^3$
2 #	2.8 ± 0.2	**11**	$(13 \pm 2) \times 10^3$
3 #	1.9 ± 0.1	**12**	$(25 \pm 4) \times 10^3$
4	$(17 \pm 2) \times 10^3$	**13**	$(13 \pm 1) \times 10^3$
5	$(44 \pm 5) \times 10^3$	**14**	$(27 \pm 2) \times 10^3$
6	$(102 \pm 9) \times 10^3$	**15**	$(22 \pm 4) \times 10^3$
7	$(50 \pm 4) \times 10^3$	**16**	$(13 \pm 2) \times 10^3$
8	$(30 \pm 3) \times 10^3$	**17**	$(6 \pm 0) \times 10^3$
9 *	$(53 \pm 7) \times 10^3$	**18**	$(12 \pm 1) \times 10^3$
I #	4.0 ± 1.3 [17]		

= Selectivity indices calculated as [K_i (off-target)/K_i (main target)] towards both *Hs*CatL and *Hs*CatB were >1800 and >2600 for **2** and **3**, respectively, and >375 for **I** (published data [17]); * = irreversible inhibition, k_{inact} = 0.022 ± 0.002 s^{-1}, and *k*2nd = $(4.1 \pm 0.6) \times 10^2$ $M^{-1}s^{-1}$.

An interesting observation is that the number of strongly inhibiting arene motifs is highly limited. The strongest inhibition was displayed by compounds **2** and **3**, showing K_i values in the single-digit nM range. They showed the previously described [17] high degree of selectivity towards the investigated human off-target cathepsins, which seems to

Molecules **2024**, *29*, 2660

be characteristic of the most active inhibitors. All but compound **9** showed linear progress curves, indicating fast-reversible inhibition (Figure 2). The irreversibility of inhibition for compound **9** was proven by dilution assays (SI-Figure S6) as described in the literature [29]. The degree of covalency and therefore a distinction between π- or σ-complex formation in the inhibited state of the complex between rhodesain and **2** or **3** cannot be made with this data alone, but is further assessed below (Sections 2.3 and 2.5). Consequently, the results in Table 1 indicate that from an electrostatic view, aniline derivatives with two nitro groups as strongly electron-withdrawing substituents in the absence or presence of an additional fluorine substituent as a suitable leaving group are necessary for strong reversible inhibition (e.g., cpd. **2** and **3**). From the benzamide derivatives, compound **9** with fluoro-dinitro-substitution shows time-dependent progress curves in the inhibition assay and undergoes an irreversible reaction with rhodesain.

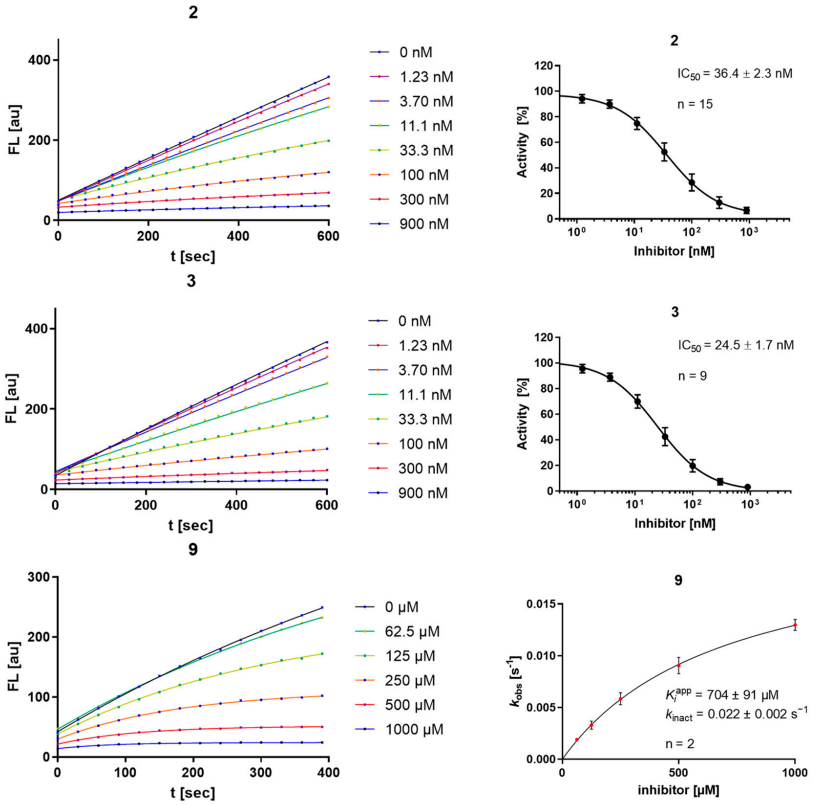


Figure 2. Progress curves (**left**) and IC_{50} determination for the time-independent inhibitors **2** and **3**, and k_{obs}-[I] diagram, K_I^{app}, and k_{inact} for the irreversible inhibitor **9** (**right**); "n" is the number of replicate measurements for each depicted concentration.

The explanation for high-affinity binding with strong inhibition and a distinction between covalent or non-covalent modes of inhibition is naturally based on multiple aspects. Reasoning can, for example, be based on electrostatic properties as prerequisites for the nucleophilic attack, but also on the presence of a suitable leaving group, favorable positioning inside the active site, and optimal bond geometries during transition states. All are important factors in deciding about stable complexes, covalent bond formation,

and substitution. By discussing electrostatic and leaving group properties, the following structure–activity relationships can be inferred. Starting from aniline as the arene portion, substituents with increasing electron-withdrawing properties were employed. First, the mono-nitro-substituted products (*ortho*- (**1**) and *para*- (**4**); *meta*- was not easily accessible) were evaluated; they did not show inhibition. Second, an additional fluorine atom was installed as a −I substituent and a potential leaving group (**5**), which also did not result in inhibition. Following the analogy of replacing nitro groups with ring nitrogen [31], while simultaneously exchanging fluorine with chlorine (**6**), did not produce an active inhibitor either. Only after transitioning to the dinitro-substituted anilines with or without an additional fluorine substituent (**2** and **3**) was strong inhibition observed again with a time-independent character. The K_i values obtained were in the single-digit nM range as for the previously reported fluoro-dinitro-bearing inhibitor **I** (compare Table 1 and Figure 2). Exchanging the fluoro-dinitro-aniline with a benzamide motif (**9**), thereby exchanging the +M-amine nitrogen with a −M-carbonyl substituent, resulted in a drop in affinity, but interestingly in the only compound showing time-dependent inhibition in the series. This is due to irreversible inhibition of rhodesain (compare Figure 2 and SI-Figure S6). A similar substitution of the aniline with a benzamide motif while retaining a monofluoro-mononitro substitution (**10** and **11**) was ineffective. Likewise, limiting the electron-donating effect of the aniline nitrogen by transformation to a urea motif on a monofluoro-mononitro scaffold did not produce an active inhibitor (**7**). A chlorine-substituted hetarene carboxamide (**8**) also did not yield relevant inhibition. Moreover, the exchange of fluorine with chlorine on the fluoro-dinitrobenzamide scaffold (**12**) also abolished the time-dependent mode of inhibition, indicating that a substitution reaction with chloride as leaving group does not take place.

Validating experiments regarding the inner-filter effect and reactivity towards unspecific nucleophiles are depicted in the SI (SI-Figures S1–S5). Most compounds of the aniline series show a relevant inner-filter effect at concentrations exceeding 10 µM, which was considered in the performed calculations. In contrast to compounds **2** and **3**, compound **9** shows relevant reactivity towards unspecific nucleophiles (water and DTT) over the course of minutes to hours, which underlines its high electrophilicity.

2.3. Mass Spectrometric Analysis

Using MALDI-TOF-MS, differentiation between non-covalent and covalent ligand-target complexes is possible if acidic matrices are used that interfere with ionic and dipole interactions (i.e., suppress non-covalent interactions) [32]. Using this method, the non-covalent inhibition mechanism of lead compound **I** was proven, as no adduct could be detected by MALDI-TOF MS, but a respective adduct peak was detected using the native ESI-MS technique, which allows for the detection of non-covalent complexes [17,24].

The MALDI method was applied to the novel compounds (Figure 3). For **2**, no adduct was found using MALDI-TOF-MS, advocating for the dominant formation of the non-covalent π-complex. For **3**, the constitutional isomer of **I**, an adduct corresponding to the mass of the inhibitor was observed, indicating the formation of a covalent σ-complex, but not the elimination of any leaving group. This difference can be explained by the relative substitution pattern with the assumption that the *ipso*-carbon of the fluorine is the most electron-deficient one (as shown by the reaction of **3** and **9** with other nucleophiles; SI-Figures S2–S4). In **I**, the electron density on the fluorine-substituted aryl-carbon is expected to be higher than in the novel compound with a changed substitution pattern. The *o*,*p*-dinitro pattern in **3** can more effectively withdraw electron density than the *m*,*p*-dinitro pattern in **I**, leading to the conclusion that the predominant inhibition product of **3** under assay conditions is the σ-complex depicted below. In addition to differences in electron distribution, unfavorable geometries can affect σ-complex formation or later prohibit the full progression of the substitution. An adduct was also found for **9**, albeit with a shift lower than the mass of the inhibitor, indicating both the formation of the covalent bond and the full substitution of fluoride, which is in line with its irreversible inhibition data (Figure 3),

its reactivity towards other nucleophiles (SI-Figures S2–S4), and the quantum-mechanical calculations explained below (Section 2.5).

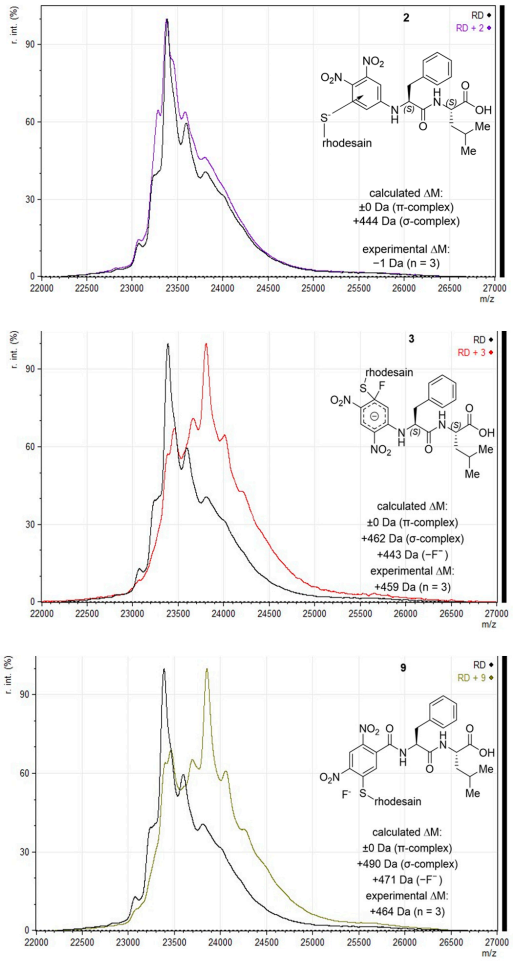


Figure 3. MALDI-TOF-MS (sinapinic acid matrix) with calculated and experimentally determined mass shifts, and depiction of assumed reaction products; "calculated ΔM" states mass shifts for possible S_NAr products. For "experimental ΔM", mean m/z of the most dominant shifted peak for each of the compounds was subtracted from mean m/z of rhodesain $[M+H]^+$; replicate measurements for each compound are indicated as "n = x".

2.4. Docking

All tested compounds were docked into the active site of rhodesain (pdb: 2p7u [33]) using LeadIT. For each compound, the 10 poses ranked highest were manually inspected to identify common binding features. Besides general positioning, the minimal distance between the catalytic thiolate of rhodesain's Cys-25 and any arene carbon (as proxy for π-complex formation) or halogen-bound carbon (expected point of attack for σ-complex formation) was calculated using Pymol.

As expected, two major orientations can be described (substrate or inverse-substrate orientation), with the arene positioned in either the S2 or S1′ pocket, respectively (shown for **I**, **2**, **3**, and **9** in Figure 4). In all benzamide-type and urea-containing inhibitors, the arene was reliably placed in the S2 pocket with substrate orientation. The aniline-derived series can adopt both orientations with different preferences depending on the arene employed. For **2** with an arene warhead in S1′, the Phe side chain is positioned into the S1 pocket, while Leu is placed into the S2 pocket. The smallest predicted distance between Cys-25 and an arene-carbon is 3.0 Å. The same general orientation is adopted by **I**, with the smallest distance of 3.5 Å, and the distance between Cys-25 and the fluorine-bound carbon is 4.8 Å (oriented away). For **3**, the arene is in the S2 pocket with a more substrate-like binding orientation of side chains with a frontal C-terminus. The smallest Cys-25/arene distance here is 3.2 Å and the distance to the fluorine-bound carbon is 4.2 Å, oriented more towards the cysteine compared to **I**. The predicted unfavorable arene position in terms of C-F to Cys-25 distance in **I** compared to **3** can be another argument for the observed difference in the covalency of the inhibition complex (π- vs. σ-complex). Compound **9** is also positioned in the substrate-like orientation with arene in S2, but with larger distances to Cys-25 (3.9 Å as minimal distance to the arene and 4.9 Å to the fluorine-bound carbon), which might be a reason for the observed drop in affinity. The same argument holds true for chlorine analog **12** (SI-Figure S7B).

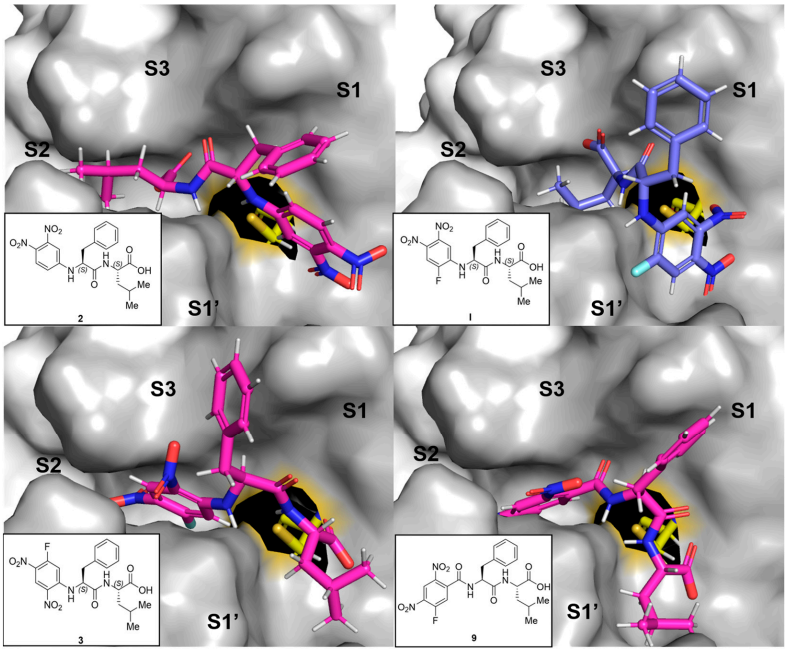


Figure 4. Representative predicted docking poses for compounds **2**, **3**, **9** (in pink), and **I** (in blue) in the active site of rhodesain (pdb 2p7u [27]). Catalytic Cys-25 is depicted as sticks (yellow), and the rest of the binding site as a surface with labeled subpockets (gray).

Inferring from the docking poses, high degrees of inhibition can be achieved by engaging the active site in both orientations, but it seems like covalently addressing the catalytic cysteine is more easily achieved from the S2 pocket in the substrate orientation than from the S1′ site in inverse orientation (with a sufficiently electrophilic arene). This hypothesis requires verification by crystallographic experiments and/or simulations that

take the enzyme flexibility into account, but it indicates a second layer of reason in addition to electron deficiency of the arene as to why some compounds react covalently, while others bind non-covalently. Interestingly, the compounds with *N*-terminal (*R*)-hPhe-based peptidic sequences might be unfavorable in general, as they can force the neighboring arene away from Cys-25 (as depicted for **18** in SI-Figure S7A), indicating why they were inactive in the assays.

2.5. Quantum-Chemical Calculations

To explain the inhibition mechanisms observed using mass spectrometric analysis, we performed quantum-chemical calculations on compounds **I**, **2**, **3**, and **9** reacting with a methyl thiolate nucleophile. As shown in Figure 5, we first calculated the π-complexes of the compounds, which were all very similar in energy (compared to the separated reactants). We chose to compare the electronic energies obtained in the calculation because in an enzymatic environment a non-covalent complex is formed prior to the reaction and the bond formation occurs without the entropic penalty. The corresponding free energy reaction paths are shown in SI-Figure S8. Compounds **I**, **3**, and **9** are able to react in the defined S_NAr reaction, while compound **2**, which does not contain a fluorine substituent, remains in the non-covalent π-complex. The energy barrier associated with the nucleophilic attack (TS1), leading to the formation of the Meisenheimer/σ-complex, is highest for compound **I**, followed by compound **3**. Additionally, the σ-complex of **I** is less stable than the π-complex ($\Delta E = +9$ kJ·mol^{-1}), causing this inhibitor to remain in a non-covalent complex with rhodesain. The arene with the highest number of electron-withdrawing groups, compound **9**, yields the lowest energy barrier for the formation of the σ-complex ($\Delta E = +23$ kJ mol^{-1}). The resulting Meisenheimer complex is exothermic, with $\Delta E = -49$ kJ mol^{-1}, and by crossing a small barrier (TS2, $\Delta E = -24$ kJ mol^{-1}), the strongly exergonic substitution product is obtained. Thus, our calculations closely match the experimental observations of an irreversible reaction of **9** with rhodesain. For compound **3**, the σ-complex is calculated to be slightly exothermic, with $\Delta E = -13$ kJ mol^{-1}, with a barrier of $\Delta E = +6$ kJ mol^{-1} for TS2. Our QM model calculations suggest that both compounds **3** and **9** would react irreversibly. However, the enzymatic environment can strongly influence warhead reactivity patterns, as seen for (fluoro)vinylsulfones [24]. Consequently, for **3**, the substitution might not be able to proceed completely due to enzymatic interactions, as discussed in the mass spectrometry section. In conclusion, although our QM calculations are not able to depict the influence of the enzymatic environment, they can provide insights into the intricate reactivity patterns of electron-deficient arenes with a thiolate.

2.6. Parallel Artificial Membrane Permeability Assay (PAMPA)

For the presented class of inhibitors, cell data are available for carboxylic acid **I** and its ester [17]. While being highly active against rhodesain, the free acid did not show strong effects in cell-based assays, in contrast to its ester, which did not inhibit isolated rhodesain but exhibited notable anti-trypanosomal activity [17]. Due to its higher lipophilicity, the ester was reasoned to be cell-permeable, releasing free acid as a highly potent rhodesain inhibitor after hydrolysis. The carboxylic acid is then trapped inside the cell due to its charge (cpK_A is 4.8 for **2** and 5.2 for **3** [33]). By the same argument, the free acid cannot enter the cells efficiently under cell culture conditions (pH = 7.4). To experimentally substantiate this hypothesis, we compared the passive permeability of the compounds as carboxylic acids (**2** and **3**) and as *tert*-butyl esters (**20** and **21**) using a standard PAMPA setup with propranolol and methotrexate as suitably well- and poorly permeable control substances, respectively, and candesartan and candesartan cilexetil as a comparative model for an acid–ester drug/prodrug pair (Table 2).

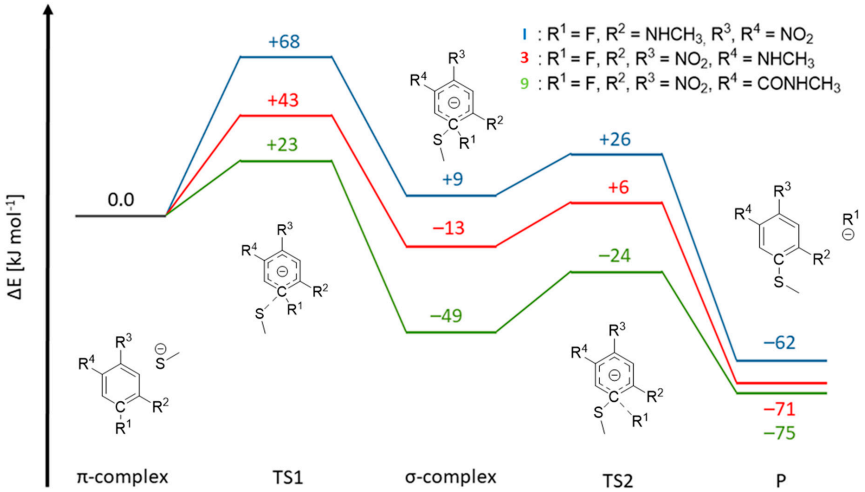
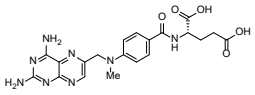
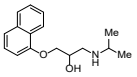
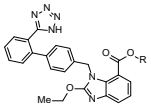
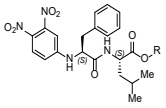
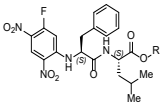


Figure 5. Reaction paths for model compounds **I**, **3**, and **9**. The π-complex of the reactants, the transition state for the formation of the Meisenheimer complex (TS1), the Meisenheimer/σ-complex and the transition state for the elimination of the fluoride (TS2) were calculated, yielding the substitution product (P). The respective C-R1 and S-CR1 distances are given in SI-Table S2.

Table 2. Determined and literature permeabilities for the compounds under investigation. "P_{app} = apparent permeability". Error is only given for compounds that displayed detectable permeation (n for all compounds was ≥2). "-" = No literature value available.

Compound		P_{app} [$\times 10^{-6}$ cm/s]	Lit. P_{app} [$\times 10^{-6}$ cm/s]
	Methotrexate	0.0	0.0 [34]
	Propranolol	9.9 ± 1.0	8–14 [35,36]
	Candesartan, R = H	0.0	0.0 [37]
	Candesartan cilexetil, R =	1.3 ± 0.2	0.6 [37]
	**2**, R = H	0.0	-
	20, R = *t*Bu	9.3 ± 0.7	-
	**3**, R = H	0.0	-
	21, R = *t*Bu	3.4 ± 1.8	-

As expected, all esters (candesartan cilexetil, **20** and **21**) clearly show increased permeability over their acid counterparts. Generally, limited permeability is not uncommon for peptidic compounds, even with blocked termini [38,39]. In this context, the strong increase in permeability from **2** to **20** is remarkable. These findings experimentally clarify the previous observation in cell cultures for the acid–ester pair of **I** (EC_{50} of 20 μM vs. 100 nM) [17].

3. Discussion

To further investigate the mode of inhibition for dipeptidic rhodesain inhibitors carrying an *N*-terminal electrophilic arene as a warhead and an unprotected *C*-terminus, we investigated 18 different compounds following two strategies. By modulating the arene warhead (compounds **1–12**), we identified two highly affine and selective reversible inhibitors with K_i values of 2.8 and 1.9 nM (**2** and **3**) and an irreversible inhibitor with a $k2nd$ value of 4.1×10^2 $M^{-1}s^{-1}$ (**9**). Focusing on a distinction between the S_NAr intermediates of non-covalent π-complex (**2**), covalent reversible σ-complex (**3**), and irreversible substitution (**9**), we identified a representative substitution pattern for the warhead that mediates each of these steps as the predominant binding state, which was experimentally proven by MALDI-TOF mass spectrometry, clarifying binding stability in the context of covalent bond formation. This was underlined by quantum-mechanical calculations, indicating that these are indeed the energetically most favorable states in each case. While they all form an energetically similar π-complex, only **3** and **9** showed an exothermic σ-complex for the attack at the fluorine-bound carbon. The retro-inverso approach (compounds **13–18**) underlined that the peptidic sequence and exact positioning of the arene is also a sensitive deciding factor for affinity. Molecular docking simulations elucidated favorable orientations that also suggested two different engagement possibilities between arene and Cys-25 through either the S2 or S1′ pocket, adding indications for covalent engagement of the catalytic cysteine. Finally, PAMPA was used to prove that the *C*-terminal esters have markedly increased membrane permeability than their acid homologs, providing explanation for the difference in anti-trypanosomal effect described previously.

4. Materials and Methods

See also Supplementary Materials.

4.1. Enzyme Sources

Rhodesain was heterologously expressed from *P. pastoris* as described in the literature [11,29], which is described in detail in the SI. Human cathepsins L and B were commercially available from EMD Millipore (Burlington, MA, USA).

4.2. Enzyme Assays

Fluorimetric enzyme assays were performed as described in the literature [29]. Information regarding buffers, concentrations, the employed substrate, as well as data analysis for linear and non-linear regression is given in the SI. In short, the cleavage of a fluorogenic substrate by the protease of interest is detected in the presence of different concentrations of inhibitors. After regression, apparent values (IC_{50} and K_I^{app}) were obtained, which were then mathematically corrected to comparable affinity values (K_i and K_I). For the irreversible compound with non-linear regression, the kinetic k_{inact} value was also plotted, from which $k2nd$ was calculated.

4.3. MALDI-TOF Mass Spectrometry

Protein mass spectrometry was performed as described before [24], but also described in detail in the SI. In short, rhodesain was incubated with a surplus of inhibitor under conditions similar to those of the enzyme assays described above to enable complex formation. This sample was desalted and then incorporated into a sinapinic acid matrix, which was then subjected to mass spectrometry. Data analysis was performed using mMass [40].

4.4. PAMPA

The principle is described in the literature [41], and the details are described in the SI. An artificial membrane consisting of phosphatidylcholine dissolved in *n*-dodecane was employed to control equilibration between two connected aqueous compartments, of which one contained the compound of interest. This model for compound permeation through a cell membrane was incubated, and the resulting concentration in the acceptor compartment was quantified spectroscopically.

4.5. Docking

The non-covalent docking was performed according to the literature [17] using FlexX, with details described in the SI.

4.6. QM Calculations

All calculations were performed using the ORCA 5.0.4 program package [42]. Geometry optimizations were performed using ωB97X-D3 [43,44] \ma-def2-SVP [45,46] with AutoAux auxiliary basis sets [47]. All stationary points were confirmed by frequency analysis, and implicit solvation in water was included using the CPCM solvation method [48]. Free energies included a concentration correction resulting from the change in standard states going from gas phase to condensed phase [49,50]. The structures depicted in SI-Figure S9 were utilized as model compounds, with methyl thiolate as the model nucleophile.

4.7. Synthetic Procedures

The general procedure for the synthesis of these compounds is shown here. For the detailed procedures as well as the analytical data, we refer to the Supporting Information. Unless stated otherwise, all solvents and reagents were obtained from commercial suppliers and used without prior purification.

4.8. General Procedure for Amide Coupling (GP1)

To a round-bottom flask, benzyloxycarbonyl (Cbz)-protected amino acid (1.01 Eq.), hydroxybenzotriazole monohydrate (HOBt·H_2O 1.01 Eq.), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC·HCl, 1.01 Eq.), 4-dimethylaminopyridine (DMAP, 1.01 Eq.), and the hydrochloride of the corresponding tert-butyl-protected amino acid (1.0 Eq.) were added and dissolved in dry dichloromethane (DCM, ca. 0.1 m). Triethylamine (2.0 Eq.) was added, and the mixture was stirred for 12 h. After the addition of water (20 mL), the phases were separated, and the organic phase was washed with saturated ammonium chloride solution. Afterwards, the organic phase was dried over sodium sulfate (Na_2SO_4), filtered, and the solvent removed under reduced pressure. The product was used without further purification.

4.9. General Procedure (GP2) for the Deprotection of Cbz-Protected Amines

In a round-bottom flask, the Cbz-protected amine was dissolved in ethanol (EtOH) or tetrahydrofurane (THF), and 10 wt% of Pd/C (5 wt%) was added. The flask was evacuated and flushed with hydrogen three times and then stirred under a slight overpressure of hydrogen until TLC showed full conversion of the starting material (roughly one hour in ethanol and 24 h in THF). The reaction mixture was filtered through celite, and the solvent evaporated.

4.10. General Procedure (GP3) for the S_NAr

To a 0.1 m solution of the amine in ethanol, the corresponding aromatic compound (1.0 Eq.) was added. *N*,*N*-Diisopropylethylamine (DIPEA, 2.0 Eq.) was added, and the reaction mixture was stirred until TLC showed full conversion. Twice as much water as ethanol was added, and the resulting mixture was extracted three times with ethylacetate (EtOAc, roughly the same volume as the ethanol used). The combined organic extracts

were dried over Na_2SO_4, and the solvent was removed under reduced pressure. The crude product was purified using flash column chromatography.

4.11. General Procedure (GP4) for the Deprotection of Tert-Butyl-Protected Acids

To an ice-cold solution of the corresponding ester in DCM (1 m), trifluoroacetic acid (TFA, 50 Eq.) was added. The reaction mixture was stirred for 3 h at rt, and the solvent was removed under reduced pressure. Traces of TFA were removed via co-evaporation with toluene. The crude product was purified using either column chromatography or preparative HPLC.

4.12. General Procedure (GP5) for Solid-Phase Synthesis

4.12.1. Loading and Capping

In total, 3.7 g chlorotriethylchloride resin was washed with DCM for 30 min and the DCM was filtered. The fluorenylmethoxycarbonyl (Fmoc)-protected amino acid (0.81 mmol) was dissolved in 100 mL of a 4% collidine/DCM mixture and added to the resin. The mixture was shaken for 12 h, the solvent filtered, and the resin washed three times with 25 mL DCM.

Then, 20 mL of a capping solution (DCM/MeOH/DIPEA, 17:2:1) was added and stirred for 1 h at rt. The solvent was filtered, and the resin was washed four times with 20 mL DCM and dimethylformamide (DMF).

4.12.2. Determination of the Loading

Here, 1.5 mg of the resin was shaken in 1 mL of a 20% piperidine/DMF mixture for 30 min. The solution was filtered and diluted with 5 mL of methanol (MeOH). The mixture was then transferred to a cuvette. The absorption spectrum was measured, and the loading B was calculated using the following equation:

$$B = \frac{A_{289nm} \times V}{\varepsilon_{289nm} \times d \times m} \quad (1)$$

4.12.3. Coupling on the Solid Phase

The resin was washed three times with 20 mL DMF and then shaken for 1 h with 20 mL of a 20% piperidine/DMF solution. The solvent was filtered and a solution consisting of the corresponding amino acid (5.0 Eq.), HATU (4.5 Eq.), and HOAt (4.5 Eq.) in 20 mL of a 20% piperidine/DMF was added, and the mixture was shaken for 12 h.

4.12.4. Capping and Fmoc Deprotection

The resin was filtered from the solvent and washed three times with the same volume of DMF and then three times with 25 mL DCM. Then, 20 mL of the capping solution was added and shaken for 1 h at rt. The solution was filtered, and the resin was washed four times with 20 mL DCM and DMF. In total, 20 mL of a 20% piperidine/DMF mixture was added and shaken for 1 h. The resin was again washed five times with 20 mL DMF and three times with DCM.

4.12.5. Aromatic Substitution and Cleavage from the Resin

1,3-Difluoro-4,6-dinitrobenzene (363 mg, 1.78 mmol, 2.2 Eq.) in 20 mL EtOH was added to the resin. DIPEA (0.28 mL, 1.62 mmol, 2.0 Eq.) was added and shaken overnight at rt. The solution was filtered, and the resin washed three times with 20 mL DCM. The resin was dried for 1 h and shaken with a mixture of 18 mL TFA, 1 mL H_2O, and 1 mL triisopropylsilane for 3 h. The mixture was filtered and removed under reduced pressure. The crude product was purified using preparative HPLC.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/molecules29112660/s1: Inner-filter-effect analysis and spectroscopic characteristics (SI-Figure S1); Stability measurements towards general substitution reactions (SI-Figures S2–S5); Reversibility assessment (SI-Figure S6); Discussion on inactive compounds from strategies A and B with a focus on positional reasoning (SI-Figure S7); Quantum-mechanical free-energy calculations (SI-Figure S8); Methods for assays and QM calculations (SI-Table S1, SI-Figure S9, SI-Table S2); Synthetic procedures and compound characterization (NMR Spectra 1–115), Cartesian coordinates of QM calculation structures. References [40–57] are also cited in Supplementary Materials.

Author Contributions: Conceptualization, T.S., T.O., C.Z. and J.B.; methodology, C.Z., J.B. and J.M.; investigation, C.Z., J.B., D.F., J.M. and P.M.; writing—original draft preparation, C.Z., J.B. and J.M.; writing—review and editing, C.Z., J.B., J.M., T.S., T.O. and B.E.; supervision, T.S., T.O., B.E. and H.-J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article and Supplementary Materials.

Acknowledgments: We thank Johannes C. Liermann (Mainz, Department of Chemistry) for NMR spectroscopy and Christopher Kampf (Mainz, Department of Chemistry) for high-resolution mass spectrometry. Additionally, we thank Jutta Schnee and Stefan Türk (Mainz, Max Planck Institute for Polymer Research) for their aid in conducting the MALDI experiments. The authors gratefully acknowledge the computing time provided to them on the high-performance computer Noctua2 at the NHR Center PC2. This system is funded by the Federal Ministry of Education and Research and the state governments participating on the basis of the resolutions of the GWK for national high-performance computing at universities (www.nhr-verein.de/unsere-partner, accessed on 11 April 2024). The authors gratefully acknowledge the computational and data resources provided by the Leibniz Supercomputing Centre (www.lrz.de, accessed on 11 April 2024). We thank Michael Klein for creating figures with BioRender.com.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Steverding, D.; Sexton, D.W.; Wang, X.; Gehrke, S.S.; Wagner, G.K.; Caffrey, C.R. *Trypanosoma brucei*: Chemical Evidence That Cathepsin L Is Essential for Survival and a Relevant Drug Target. *Int. J. Parasitol.* **2012**, *42*, 481–488. [CrossRef] [PubMed]
2. Nikolskaia, O.V.; Lima, A.P.C.D.A.; Kim, Y.V.; Lonsdale-Eccles, J.D.; Fukuma, T.; Scharfstein, J.; Grab, D.J. Blood-Brain Barrier Traversal by African Trypanosomes Requires Calcium Signaling Induced by Parasite Cysteine Protease. *J. Clin. Investig.* **2006**, *116*, 2739–2747. [CrossRef] [PubMed]
3. Grab, D.J.; Garcia-Garcia, J.C.; Nikolskaia, O.V.; Kim, Y.V.; Brown, A.; Pardo, C.A.; Zhang, Y.; Becker, K.G.; Wilson, B.A.; De, A. Lima, A.P.C.; et al. Protease Activated Receptor Signaling Is Required for African Trypanosome Traversal of Human Brain Microvascular Endothelial Cells. *PLoS Negl. Trop. Dis.* **2009**, *3*, e479. [CrossRef] [PubMed]
4. WHO. Number of Reported Cases of Human African Trypanosomiasis. Available online: https://www.who.int/data/gho/data/themes/topics/human-african-trypanosomiasis (accessed on 10 December 2023).
5. Franco, J.R.; Cecchi, G.; Priotto, G.; Paone, M.; Diarra, A.; Grout, L.; Simarro, P.P.; Zhao, W.; Argaw, D. Monitoring the Elimination of Human African Trypanosomiasis at Continental and Country Level: Update to 2018. *PLoS Negl. Trop. Dis.* **2020**, *14*, e0008261. [CrossRef] [PubMed]
6. Pépin, J. Sleeping Sickness: Time for Dreaming. *Lancet Infect. Dis.* **2023**, *23*, 387–388. [CrossRef] [PubMed]
7. Betu Kumeso, V.K.; Kalonji, W.M.; Rembry, S.; Valverde Mordt, O.; Ngolo Tete, D.; Prêtre, A.; Delhomme, S.; Ilunga Wa Kyhi, M.; Camara, M.; Catusse, J.; et al. Efficacy and Safety of Acoziborole in Patients with Human African Trypanosomiasis Caused by *Trypanosoma brucei* Gambiense: A Multicentre, Open-Label, Single-Arm, Phase 2/3 Trial. *Lancet Infect. Dis.* **2023**, *23*, 463–470. [CrossRef]
8. Agbowuro, A.A.; Huston, W.M.; Gamble, A.B.; Tyndall, J.D.A. Proteases and Protease Inhibitors in Infectious Diseases. *Med. Res. Rev.* **2018**, *38*, 1295–1331. [CrossRef] [PubMed]
9. Owen, D.R.; Allerton, C.M.N.; Anderson, A.S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R.D.; Carlo, A.; Coffman, K.J.; et al. An Oral SARS-CoV-2 M pro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science* **2021**, *374*, 1586–1593. [CrossRef] [PubMed]
10. Biasizzo, M.; Javoršek, U.; Vidak, E.; Zarić, M.; Turk, B. Cysteine Cathepsins: A Long and Winding Road towards Clinics. *Mol. Aspects Med.* **2022**, *88*, 101150. [CrossRef]

11. Caffrey, C.R.; Hansell, E.; Lucas, K.D.; Brinen, L.S.; Alvarez Hernandez, A.; Cheng, J.; Gwaltney, S.L.; Roush, W.R.; Stierhof, Y.D.; Bogyo, M.; et al. Active Site Mapping, Biochemical Properties and Subcellular Localization of Rhodesain, the Major Cysteine Protease of *Trypanosoma brucei rhodesiense*. *Mol. Biochem. Parasitol.* **2001**, *118*, 61–73. [CrossRef]
12. Lewis, S.D.; Johnson, F.A.; Shafer, J.A. Effect of Cysteine-25 on the Ionization of Histidine-159 in Papain as Determined by Proton Nuclear Magnetic Resonance Spectroscopy. Evidence for a Histidine-159-Cysteine-25 Ion Pair and Its Possible Role in Catalysis. *Biochemistry* **1981**, *20*, 48–51. [CrossRef] [PubMed]
13. Jefferson, T.; McShan, D.; Warfield, J.; Ogungbe, I.V. Screening and Identification of Inhibitors of *Trypanosoma brucei* Cathepsin L with Antitrypanosomal Activity. *Chem. Biol. Drug Des.* **2016**, *87*, 154–158. [CrossRef] [PubMed]
14. dos Santos Nascimento, I.J.; de Aquino, T.M.; da Silva-Júnior, E.F. Cruzain and Rhodesain Inhibitors: Last Decade of Advances in Seeking for New Compounds against American and African Trypanosomiases. *Curr. Top. Med. Chem.* **2021**, *21*, 1871–1899. [CrossRef] [PubMed]
15. Schirmeister, T.; Kaeppler, U. Non-Peptidic Inhibitors of Cysteine Proteases. *Mini-Rev. Med. Chem.* **2005**, *3*, 361–373. [CrossRef] [PubMed]
16. Klein, P.; Barthels, F.; Johe, P.; Wagner, A.; Tenzer, S.; Distler, U.; Le, T.A.; Schmid, P.; Engel, V.; Engels, B.; et al. Naphthoquinones as Covalent Reversible Inhibitors of Cysteine Proteases—Studies on Inhibition Mechanism and Kinetics. *Molecules* **2020**, *25*, 2064. [CrossRef] [PubMed]
17. Klein, P.; Johe, P.; Wagner, A.; Jung, S.; Kühlborn, J.; Barthels, F.; Tenzer, S.; Distler, U.; Waigel, W.; Engels, B.; et al. New Cysteine Protease Inhibitors: Electrophilic (Het)Arenes and Unexpected Prodrug Identification for the *Trypanosoma* Protease Rhodesain. *Molecules* **2020**, *25*, 1451. [CrossRef] [PubMed]
18. Meisenheimer, J. Ueber Reactionen Aromatischer Nitrokörper. *Justus Liebigs Ann. Chem.* **1902**, *323*, 205–246. [CrossRef]
19. Leesnitzer, L.M.; Parks, D.J.; Bledsoe, R.K.; Cobb, J.E.; Collins, J.L.; Consler, T.G.; Davis, R.G.; Hull-Ryde, E.A.; Lenhard, J.M.; Patel, L.; et al. Functional Consequences of Cysteine Modification in the Ligand Binding Sites of Peroxisome Proliferator Activated Receptors by GW9662. *Biochemistry* **2002**, *41*, 6640–6650. [CrossRef] [PubMed]
20. 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*, 804970. [CrossRef]
21. Fairhurst, R.A.; Knoepfel, T.; Leblanc, C.; Buschmann, N.; Gaul, C.; Blank, J.; Galuba, I.; Trappe, J.; Zou, C.; Voshol, J.; et al. Approaches to Selective Fibroblast Growth Factor Receptor 4 Inhibition through Targeting the ATP-Pocket Middle-Hinge Region. *MedChemComm* **2017**, *8*, 1604–1613. [CrossRef]
22. Gerstenecker, S.; Haarer, L.; Schröder, M.; Kudolo, M.; Schwalm, M.P.; Wydra, V.; Serafim, R.A.M.; Chaikuad, A.; Knapp, S.; Laufer, S.; et al. Discovery of a Potent and Highly Isoform-Selective Inhibitor of the Neglected Ribosomal Protein S6 Kinase Beta 2 (S6K2). *Cancers* **2021**, *13*, 5133. [CrossRef] [PubMed]
23. Schwickert, M.; Zimmermann, R.A.; Habeck, T.; Hoba, S.N.; Nidoieva, Z.; Fischer, T.R.; Stark, M.M.; Kersten, C.; Lermyte, F.; Helm, M.; et al. Covalent S-Adenosylhomocysteine-Based DNA Methyltransferase 2 Inhibitors with a New Type of Aryl Warhead. *ACS Med. Chem. Lett.* **2023**, *14*, 777–787. [CrossRef]
24. Jung, S.; Fuchs, N.; Johe, P.; Wagner, A.; Diehl, E.; Yuliani, T.; Zimmer, C.; Barthels, F.; Zimmermann, R.A.; Klein, P.; et al. Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure–Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies. *J. Med. Chem.* **2021**, *64*, 12322–12358. [CrossRef]
25. Carriero, M.V.; Bifulco, K.; Ingangi, V.; Costantini, S.; Botti, G.; Ragone, C.; Minopoli, M.; Motti, M.L.; Rea, D.; Scognamiglio, G.; et al. Retro-Inverso Urokinase Receptor Antagonists for the Treatment of Metastatic Sarcomas. *Sci. Rep.* **2017**, *7*, 1312. [CrossRef]
26. Chorev, M.; Goodman, M. Recent Developments in Retro Peptides and Proteins—An Ongoing Topochemical Exploration. *Trends Biotechnol.* **1995**, *13*, 438–445. [CrossRef]
27. Kerr, I.D.; Lee, J.H.; Farady, C.J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K.C.; Caffrey, C.R.; Legac, J.; Hansell, E.; et al. Vinyl Sulfones as Antiparasitic Agents and a Structural Basis for Drug Design. *J. Biol. Chem.* **2009**, *284*, 25697–25703. [CrossRef] [PubMed]
28. Schechter, I.; Berger, A. On the Size of the Active Site in Proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162. [CrossRef] [PubMed]
29. Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T.H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P.R.; Distler, U.; et al. Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138*, 8332–8335. [CrossRef]
30. Cheng, Y.-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*, 3099–3108. [CrossRef]
31. Schlosser, M.; Ruzziconi, R. Nucleophilic Substitutions of Nitroarenes and Pyridines: New Insight and New Applications. *Synthesis* **2010**, *2010*, 2111–2123. [CrossRef]
32. Zehl, M.; Allmaier, G. Investigation of Sample Preparation and Instrumental Parameters in the Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry of Noncovalent Peptide/Peptide Complexes. *Rapid Commun. Mass. Spectrom.* **2003**, *17*, 1931–1940. [CrossRef] [PubMed]
33. ChemAxon Marvin JS 23.11.0. Available online: https://playground.calculators.cxn.io/ (accessed on 19 December 2023).

34. Kerns, E.H.; Di, L.; Petusky, S.; Farris, M.; Ley, R.; Jupp, P. Combined Application of Parallel Artificial Membrane Permeability Assay and Caco-2 Permeability Assays in Drug Discovery. *J. Pharm. Sci.* **2004**, *93*, 1440–1453. [CrossRef] [PubMed]
35. Sugano, K.; Hamada, H.; Machida, M.; Ushio, H. High Throughput Prediction of Oral Absorption: Improvement of the Composition of the Lipid Solution Used in Parallel Artificial Membrane Permeation Assay. *SLAS Discov.* **2001**, *6*, 189–196. [CrossRef] [PubMed]
36. Avdeef, A. *Absorption and Drug Development*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003. [CrossRef]
37. Shimizu, M.; Fukami, T.; Okura, K.; Taniguchi, T.; Nomura, Y.; Nakajima, M. Utility of a Systematic Approach to Selecting Candidate Prodrugs: A Case Study Using Candesartan Ester Analogues. *J. Pharm. Sci.* **2023**, *112*, 1671–1680. [CrossRef] [PubMed]
38. Rafi, S.B.; Hearn, B.R.; Vedantham, P.; Jacobson, M.P.; Renslo, A.R. Predicting and Improving the Membrane Permeability of Peptidic Small Molecules. *J. Med. Chem.* **2012**, *55*, 3163–3169. [CrossRef] [PubMed]
39. Avdeef, A. The Rise of PAMPA. *Expert. Opin. Drug Metab. Toxicol.* **2005**, *1*, 325–342. [CrossRef] [PubMed]
40. Strohalm, M.; Hassman, M.; Košata, B.; Kodíček, M. MMass Data Miner: An Open Source Alternative for Mass Spectrometric Data Analysis. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 905–908. [CrossRef] [PubMed]
41. Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41*, 1007–1010. [CrossRef] [PubMed]
42. Neese, F. The ORCA Program System-Version 5.0. *WIREs Comput. Mol. Sci.* **2022**, *12*, e1606. [CrossRef]
43. Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615. [CrossRef]
44. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [CrossRef] [PubMed]
45. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297. [CrossRef] [PubMed]
46. Zheng, J.; Xu, X.; Truhlar, D.G. Minimally Augmented Karlsruhe Basis Sets. *Theor. Chem. Acc.* **2011**, *128*, 295–305. [CrossRef]
47. Stoychev, G.L.; Auer, A.A.; Neese, F. Automatic Generation of Auxiliary Basis Sets. *J. Chem. Theory Comput.* **2017**, *13*, 554–562. [CrossRef]
48. Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001. [CrossRef]
49. Ben-Naim, A. Standard Thermodynamics of Transfer. Uses and Misuses. *J. Phys. Chem.* **1978**, *82*, 792–803. [CrossRef]
50. Pliego, J.R., Jr.; Riveros, J.M. Gibbs Energy of Solvation of Organic Ions in Aqueous and Dimethyl Sulfoxide Solutions. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622–1627. [CrossRef]
51. 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. [CrossRef] [PubMed]
52. Weitner, T.; Friganović, T.; Šakić, D. Inner Filter Effect Correction for Fluorescence Measurements in Microplates Using Variable Vertical Axis Focus. *Anal. Chem.* **2022**, *94*, 7107–7114. [CrossRef]
53. Halgren, T.A. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 490–519. [CrossRef]
54. Muramatsu, W.; Yamamoto, H. Peptide Bond Formation of Amino Acids by Transient Masking with Silylating Reagents. *J. Am. Chem. Soc.* **2021**, *143*, 6792–6797. [CrossRef] [PubMed]
55. Egawa, H.; Kataoka, M.; Shibamori, K.-I.; Miyamoto, T.; Nakano, J.; Matsumoto, J.-I. A New Synthetic Route to 7-Halo-1-Cyclopropyl-6-Fluoro-1,4-Dihydro-4-Oxoquinoline-3-Carboxylic Acid, an Intermediate for the Synthesis of Quinolone Antibacterial Agents. *J. Heterocycl. Chem.* **1987**, *24*, 181–185. [CrossRef]
56. Draper, R.W.; Hou, D.; Iyer, R.; Lee, G.M.; Liang, J.T.; Mas, J.L.; Vater, E.J. Novel Stereoselective Syntheses of the Fused Benzazepine Dopamine D1 Antagonist (6a*S*,13b*R*)-11-Chloro-6,6a,7,8,9,13b-Hexahydro-7-Methyl-5*H*-Benzo[*d*]Naphth [2,1-*b*]Azepin-12-ol (Sch 39166): 2. l-Homophenylalanine-Based Syntheses. *Org. Process Res. Dev.* **1998**, *2*, 186–193. [CrossRef]
57. Fotakis, C.; Christodouleas, D.; Zoumpoulakis, P.; Kritsi, E.; Benetis, N.-P.; Mavromoustakos, T.; Reis, H.; Gili, A.; Papadopoulos, M.G.; Zervou, M. Comparative Biophysical Studies of Sartan Class Drug Molecules Losartan and Candesartan (CV-11974) with Membrane Bilayers. *J. Phys. Chem. B* **2011**, *115*, 6180–6192. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Supporting Information

Substitution-induced Mechanistic Switching in S_NAr-Warheads for Cysteine Proteases

Collin Zimmer[†,1], Jan Brauer[†,2], Dorota Ferenc[2], Jessica Meyr[3], Patrick Müller[1], Hans-Joachim Räder[4], Bernd Engels[3], Till Opatz[*,2] and Tanja Schirmeister[*,1]

[†] Both authors contributed equally
[1] Institute of Pharmaceutical and Biomedical Sciences, University of Mainz, Staudingerweg 5, 55128, Mainz, Germany
[2] Department of Chemistry, University of Mainz, Duesbergweg 10–14, 55128 Mainz, Germany
[3] Institute of Physical and Theoretical Chemistry, Julius-Maximilians-University, Am Hubland, 97074 Würzburg, Germany
[4]Max Plack Institute for Polymer Research, Ackermannweg 10, 55128 Mainz

*Corresponding authors

TABLE OF CONTENTS

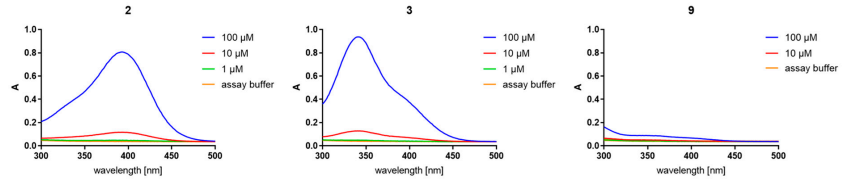
Inner-filter effect and spectroscopic characteristics

Since especially the aniline-derivatives were found to have relatively low-energy absorption maxima, e.g., due to a *p*-nitro-aniline push-pull system (compare **SI-Figure S1A**), the inner filter effect must be ruled out as an influencing factor for the inhibition data.

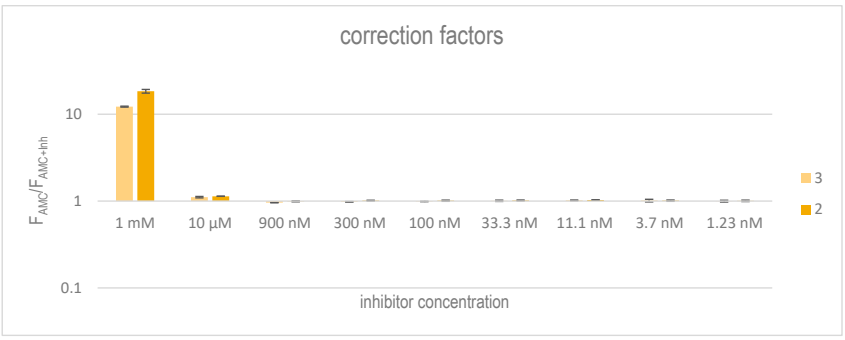
To this end, fluorescence of free AMC (λ_{ex} = 380 nm, λ_{em} = 460 nm; 10 µM in assay buffer) was measured in absence and presence of inhibitor and the resulting reduction in fluorescence was analyzed for the potential to influence kinetic measurements of inhibition. The effect is independent on the concentration of AMC, but dependent on the concentration of the inhibitor, therefore different concentrations of the latter were assessed (compare **SI-Figure S1B**) [1].

It is evident that while the nitro-anilines have the potential for assay interference, this problem does only occur at high concentrations (≥10 µM), which was taken into account for affinity calculations (see the method for the cathepsin assays). For compounds **2** and **3**, the inhibition results were not mathematically corrected since the correction of assay output ([FU/s]) for each inhibitor concentration <10 µM would only have been <5 %. By analysis of the UV-spectra of **2** and **3**, it is evident that their influence at higher concentrations is mainly by primary inner-filter effect (absorption of exciting light), and only to a lesser extent by secondary inner-filter effect (absorption of emitted light) [2]. Even though **9** causes fluorescence interference at highest concentrations (8% at 500 µM and 23% at 1000 µM), this does not need to be corrected for in the utilized time-dependent mode of evaluation (therefore not depicted).

A



B



SI-Figure S1: (**A**) Absorption spectra of **2**, **3** and **9** with different concentrations, visualizing a potential need for mathematical correction at high concentrations for nitroanilines. (**B**) Experimental proof that mathematical correction of attenuation of AMC fluorescence is only necessary for inhibitor concentrations ≥10 µM. For **2** and **3** at 10 µM attenuation of fluorescence signal is 14 and 11 %, respectively. At concentrations relevant for IC_{50} evaluation, no relevant interference is detected.

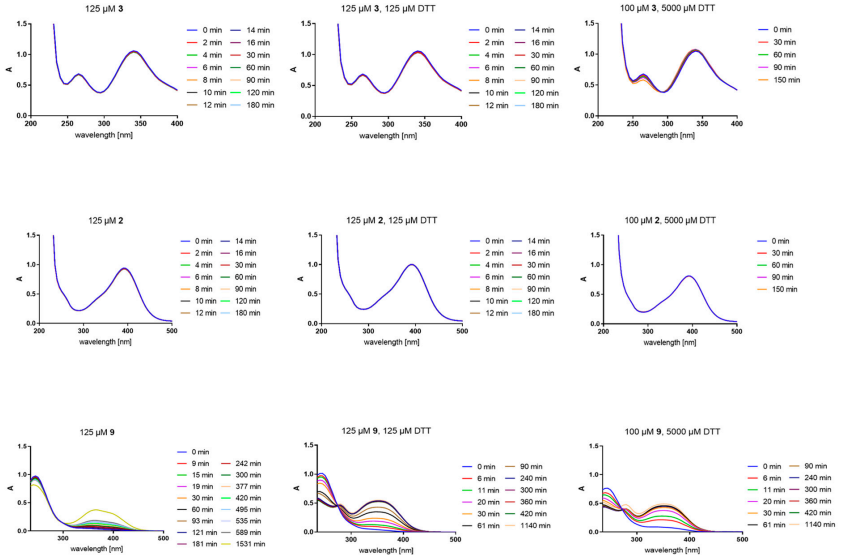
Stability measurements towards general substitution reactions

2, **3** and **9** (from 20 mM stock solutions in DMSO) were diluted into NH_4OAc buffer (pH 5.5) to final concentrations of 100 µM, 125 µM or 1 mM and incubated in buffer alone or in presence of as much DTT as present in the fluorimetric rhodesain assay and the MALDI-TOF experiment, respectively. This was recorded with absorption spectroscopy (results depicted in **SI-Figure S2**) on a Tecan Spark 10M plate reader in UV-transparent 96-well plates (Greiner UV-Star®, 655801, 200 µL volume per sample) or with LCMS (see "measurement setup for LCMS" as described under the method for PAMPA; results depicted in **SI-Figure S3**).

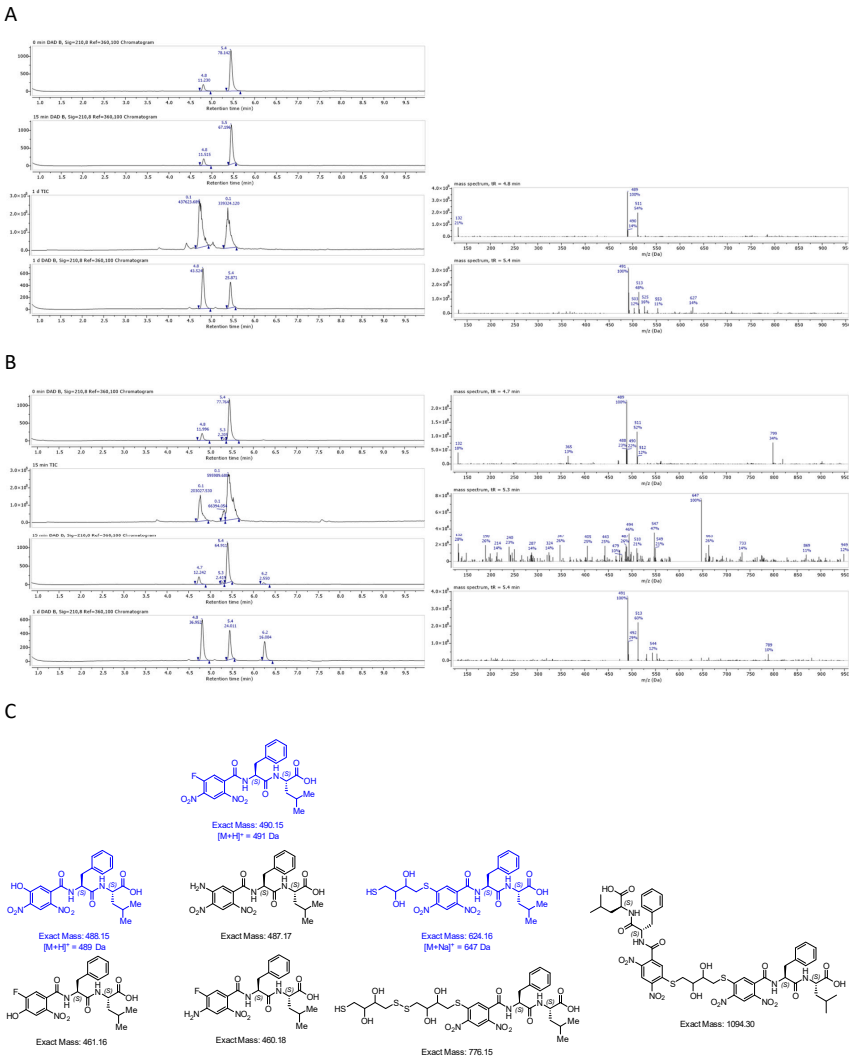
The aniline derivates **2** and **3** show no relevant tendency towards either side reaction in an aqueous system at pH 5.5 over the course of the time necessary for the employed assays. There is a small difference detectable in reactivity of **3** with 5 mM DTT: a slight but time-dependent change in UV-spectrum was found where the maximum at 265 nm decreased in intensity, indicating a small degree of degradation

that is not observed with lower DTT concentration or with **2** at all. The formation takes place on a much larger timescale as the time for detection in the assays (minutes vs. hours) and is therefore irrelevant to the performed evaluations.

However, **9** shows a clear time-dependent change in its UV-spectrum and analysis of the reaction by repeated LCMS analysis indicates the formation of the phenol (substitution of fluoride) over the course of hours. This also seems to happen in the DMSO stock solution to a small degree after some usage (explains the presence of phenol at 0 min). Accordingly, in presence of DTT, both the phenol and the thioether are detected by MS for this compound.

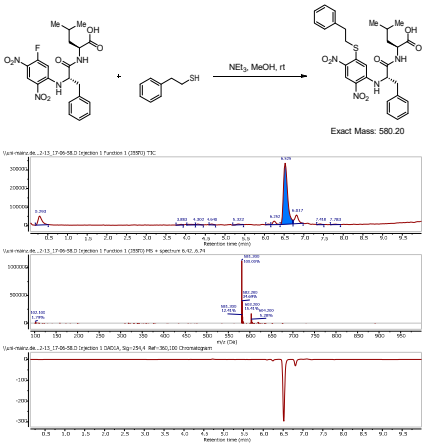


SI-Figure S2: Time-dependent absorption spectra for **2**, **3** and **9** indicating the instability of **9** towards both water and DTT even at slightly acidic pH = 5.5. Spectra for **9** were buffer-corrected prior to display.



SI-Figure S3: Time-dependent degradation of 1 mM **9** to the putative phenol (Δm/z = −2 Da was interpreted as a F→OH exchange) in presence of ammonium acetate buffer pH 5.5 (**A**), and to the putative phenol and a putative DTT adduct in the mentioned buffer + 125 µM DTT (**B**). Some possible reaction products are shown (**C**) with the suggested structures of detected m/z (in either A or B) in blue. Alignment of TIC and DAD chromatograms was performed manually due to a known instrumental error, retention time for column dead volume with DMSO peak was excluded for clarity.

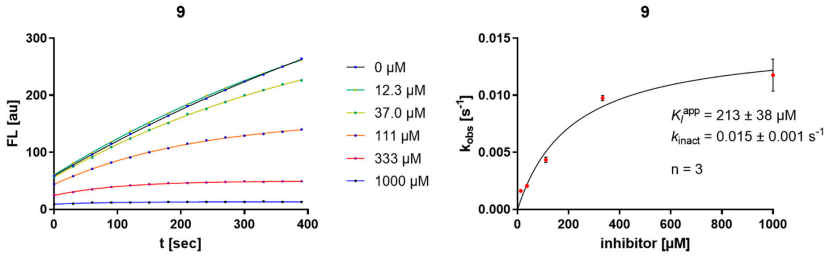
For **3** a substitution experiment in an organic solvent with an *in situ* generated thiolate was also performed: The peptide (46 mg, 0.1 mmol) was dissolved in dry methanol and a mixture of triethylamine (38 µL, 0.3 mmol) and 2-phenylethanethiol (27 µL, 0.2 mmol) was added. The reaction mixture was stirred for one minute and LCMS analysis of the sample was performed. The reaction control only showed one product, resulting from the substitution of fluoride from the aromatic system. No further reaction was observed, even after stirring for prolonged time (up to 3 h).



SI-Figure S4: LCMS analysis of the reaction of **3** with low-MW thiol in the presence of a base indicated almost immediate quantitative conversion to the substitution product.

It is worth noting that in our analysis of the novel compounds, for both the aniline and the benzamide, the substitution of fluoride, which is expected to be the best leaving group of the system, seems to be largely preferred over the substitution of the nitro group reported under forced conditions on **I** in the literature [3]. In this regard, the compounds are therefore expected to react preferentially with the catalytic cysteine in rhodesain under covalent binding (if it occurs) to the fluorine-bearing carbon.

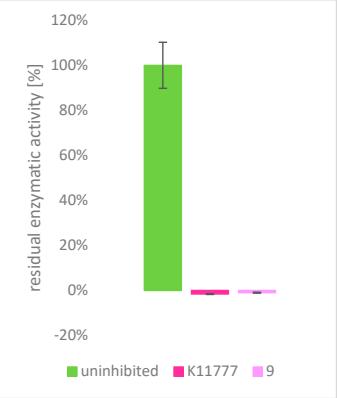
The benzamide is not stable in water in both the presence and absence of DTT. In the presence of DTT, this might already be relevant on a minute-timescale. To investigate this influence, its inhibition of rhodesain was reassessed in DTT-free conditions (tris(2-carboxyethyl)phosphine = TCEP used as a non-nucleophilic reducing agent for enzyme activation). Only slightly improved affinity (K_i^{app} = 0.2 mM instead of 0.7 mM) was recorded as depicted in **SI-Figure S5**. The reactivity towards water should not be seriously relevant on that timescale, even though it is expected to complicate all analysis towards this compound. Reduced affinity due to the elongation of the scaffold by the carbonyl group and the concomitant changes in conformation and binding interactions is expected to be the main reason for the high observed K_i^{app}.



SI-Figure S5: Example progress curves, k_{obs}-[I]-diagram, K_i^{app}, and k_{inact} for the time-dependent inhibitor **9** using TCEP for rhodesain activation to enable inhibitor assessment in absence of low-MW thiol.

Reversibility assessment for time-dependent inhibitor 9

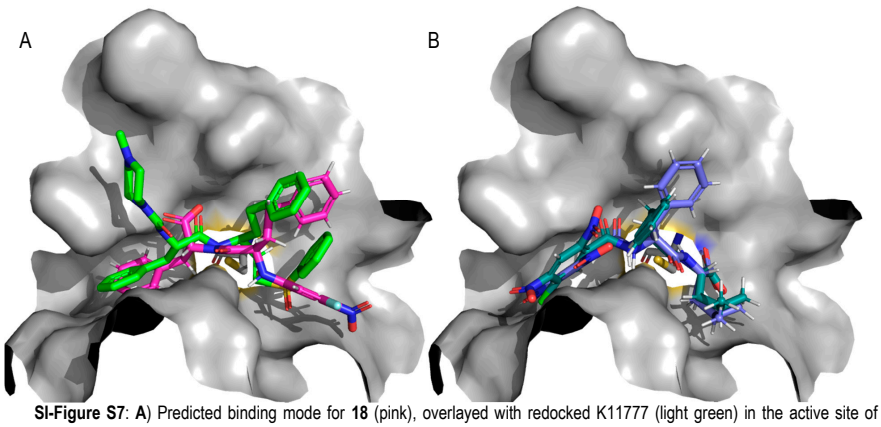
To assess reversibility of the time-dependent inhibitor [4], and to therefore discriminate between slow-reversible and irreversible inhibition, rhodesain was incubated with compound **9** for 1.5 h, at an inhibitor concentration of 400 µM (7.5x K_i) to enable the quantitative progression of a covalent irreversible reaction if one is possible. The same incubation was performed with DMSO alone as a negative control and in presence of 500 nM K11777 as a positive irreversible control. After this incubation period, the samples were diluted 1:100 (to 4 µM of compound **9** (0.075x K_i)) into assay buffer containing substrate to record the degree of recovery of enzymatic activity over the course of 10 min. It is evident that compound **9** could not be displaced by the substrate after incubation and behaved like the irreversible control K11777. The results are depicted in **SI-Figure S6**.



SI-Figure S6: Results of dilution assay to analyze reversibility of inhibition (n = 2). Raw activity of rhodesain after dilution was 0.68 ± 0.05 FU/s which is in the expected range for the concentration and instrument settings. Activity of rhodesain after incubation with either K11777 or **9** and subsequent dilution was −0.01 ±0.00 FU/s in both cases, indicating no recovery of enzymatic activity as expected from irreversible inhibitors.

Discussion on inactive compounds from strategies A and B with a focus on positional reasoning

The compounds containing a hPhe were designed to mimic K11777's binding interactions but overall did not show strong inhibition in the *in vitro* assay. The docking elucidated that this positioning can indeed be unfavorable for the used warhead as it greatly increased its distance from Cys-25. This is also reflected in their assay data: the *N*-terminal (*R*)-hPhe-containing peptides (K_i(**14**) = 27 µM, and K_i(**18**) = 12 µM; docking: arene positioned in S1') are weak inhibitors, probably through forcing the arene away from the catalytic Cys-25 if the (*R*)-hPhe side-chain is positioned optimally (as depicted for **18** in **SI-Figure S7-A**). In contrast, compounds with C-terminal (*S*)-hPhe are predicted to guide the *N*-terminal arene to the S2 pocket, similarly to **3** or **9**, but with different side-chain orientations (K_i(**17**) = 6 µM , and K_i(**13**) = 13 µM). However, since none of the hPhe-containing compounds showed striking inhibition, and the differences between them were marginal, this amino acid is unfavorable when employing *N*-terminal electrophilic arenes as warheads against rhodesain.

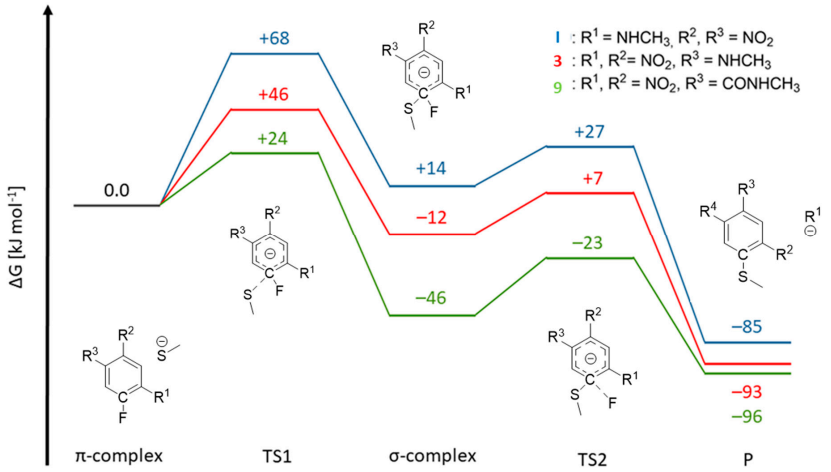


SI-Figure S7: **A**) Predicted binding mode for **18** (pink), overlayed with redocked K11777 (light green) in the active site of rhodesain (pdb: 2p7u). The positioning of Phe and hPhe side chains results in an unfavorable orientation of the electrophilic arene away from Cys-25. **B**) Predicted binding mode for **12** (dark green), and **9** (blue) in the active site of rhodesain (pdb: 2p7u). Both are predicted to be able to assume similar positioning of the arene but with larger distance to Cys-25 for the chlorine-substituted **12** (5.8 Å and 7.0 Å, as smallest distance, and distance to C-Cl, respectively).

Quantum-mechanical free energy calculations

The reaction paths include separately optimized reactants (R), the π-complex of the thiolate and the aromatic compound, transition states (TS), Meisenheimer/σ-complex and the product following the elimination of a fluoride anion (P). The free energies of the reactions are shown in **SI-Figure S8**.

These calculations are in line with the reactivity experiments for **3** and **9** described above and the published reactivity of **I**. Compound **I** slowly reacts with thiolate at room temperature [3], corresponding to the higher energy barrier, and the endergonic σ-complex. Compound **3** performs this reaction more readily, reflecting the lower energy barriers compared to **I**, and the exergonic progression. Finally, compound **9** with the most favorable energy profile for the reaction was shown to undergo substitution with the less nucleophilic thiol already.



SI Figure S8: Free energy reaction paths for model compounds **I**, **3**, and **9.** The π-complex of the reactants, the transition state for the formation of the Meisenheimer complex (TS1), the Meisenheimer/σ-complex and the transition state for the elimination of the fluoride (TS2) were calculated, yielding the substitution product (P). The respective C-R¹ and S-CR¹ distances are given in **SI-Table S2**.

Methods

Rhodesain expression

Rhodesain was recombinantly expressed in *P. pastoris* according to a method adapted from literature [5,6]: The *Pichia pastoris* X-33 mutant, stably transformed with the rhodesain ΔC gene cloned into the pPICZαA vector, was cultured in buffered minimal glycerol supplemented with ampicillin (BMG_{Amp}) at 30 °C to an optical density (λ = 600 nm) of 2–3. To induce AOX1-controlled expression, BMG_{Amp} was exchanged for buffered minimal methanol with ampicillin (BMM_{Amp}). Incubation was continued for 72 h, and 0.5 % (V/V) methanol was added every 12 h. After this period, cells were pelleted by centrifugation and discarded. The culture supernatant containing secreted rhodesain was filtered (0.2 µm filter, cellulose acetate), adjusted to 2 M sodium chloride and loaded overnight onto a Phenylsepharose FF (high sub) column (V = 20 mL), equilibrated to chromatography conditions described below, attached to an ÄKTA start system. Hydrophobic interaction chromatography was performed by gradient elution with falling concentrations of sodium chloride (2 M to 0 M) in 20 mM sodium citrate buffer (pH 5.5). Rhodesain was eluted afterwards with MQ-water. The eluate was concentrated with centrifugal filter units (10 kDa MWCO, regenerated cellulose) to a volume ≤ 5 mL. Subsequently, size exclusion chromatography was performed with 20 mM sodium citrate buffer with 200 mM sodium chloride (pH 5.5) on an ÄKTA start system equipped with a HiLoad 16/600 Superdex 75 pg column. Rhodesain-containing fractions were pooled and then dialyzed against MQ-water for 4 h (dialysis tubing, 6 kDa MWCO, regenerated cellulose). The desalted rhodesain solution was lyophilized overnight and stored at ≤5 °C.

BMG_{Amp} / BMM_{Amp}:
1 % (V/V) glycerol OR 0.5 % (V/V) methanol
100 mM potassium phosphate buffer pH 6
3.4 g/L yeast nitrogen base without amino acids, without ammonium sulphate
10 g/L ammonium sulphate
0.4 mg/L biotin
100 mg/L ampicillin

Rhodesain, CatL, CatB Assays

Continuous fluorometric measurements with a Tecan Spark 10M reader in a 96 well plate format was used to evaluate activity [6]. Rhodesain, *Hs*CatB and *Hs*CatL were diluted in their respective activation buffer prior to use. Rhodesain was incubated in this for 30–60 min prior to use. The enzyme of interest (5 µL) was incubated with the inhibitor in question (10 µL), and its specific substrate (5 µL) in the respective assay buffer. The fluorescence signal was measured every 30 s for 10 min at room temperature with the corresponding excitation/emission wavelengths (λ_{ex} = 380 nm / λ_{em} = 460 nm). The final concentrations and buffer compositions are depicted in **SI-Table S1**. Inner-filter effect corrections were performed as described in literature [1] if deviations >5% were detected.

SI-Table S1: Information on assay conditions. [E] = final enzyme concentration, [S] = final substrate concentration in well.

Enzyme	***Hs*CatB**	***Hs*CatL**	**Rhodesain (*Tb*CatL)**
Enzyme storage buffer	50 mM NaOAc, 1 mM EDTA, pH = 5.0	20 mM malonate, 400 mM NaCl, 1 mM EDTA, pH = 5.5	50 mM NaOAc, 200 mM NaCl, 5 mM EDTA, pH = 5.5
Activation buffer	50 mM TRIS, 200 mM NaCl, 5 mM EDTA, 2 mM DTT, pH = 6.5		50 mM NaOAc, 200 mM NaCl, 5 mM EDTA, 5 mM DTT, pH = 5.5
Assay buffer	50 mM TRIS, 200 mM NaCl, 5 mM EDTA, pH = 6.5		50 mM NaOAc, 200 mM NaCl, 5 mM EDTA, 0.005 % Brij 35, pH = 5.5
Manufacturer	EMD	EMD	In-house expression
[E]	3 nM	5 nM	0.5 nM
Substrate structure	Z-Phe-Arg-AMC		
[S]	100 µM	6.25 µM	10 µM
K_M	150 µM	6.5 µM	0.827 µM

For the evaluation of **9** in absence of DTT (depicted in **SI-Figure S5**): 5 nM final rhodesain concentration that was activated before in presence of TCEP (alternative activation buffer: 5 mM DTT → 5 mM TCEP) was used to generate a similar slope in a control measurement in absence of inhibitor as for the measurements with DTT.

Data evaluation.

Time-independent behavior

For inhibitors that behaved time-independently during the measurements (i.e., F-t-diagrams show constant slope), IC_{50} values using defined dilution series for each compound were calculated with GraphPad PRISM by fitting the remaining enzymatic activity to a four parameter IC_{50} equation with Y [%] as the residual enzyme activity, Y_{max} as the maximum value of the dose response curve at inhibitor concentrations [I] = 0 µM, Y_{min} as the minimum value at high inhibitor concentrations and s as the Hill coefficient.

$$y\ [\%residual\ activity] = \frac{y_{\max} - y_{\min}}{1 + \left(\frac{[\mathrm{I}]}{\mathrm{IC}_{50}}\right)^{s}} + y_{\min} \qquad \textbf{eq. 1}$$

It is assumed that the inhibitors bind competitively in respect to the substrate [7]. Due to the dependence of the IC_{50} value on the substrate affinity and concentration, the K_i values were calculated with the Cheng-Prusoff equation, using the final substrate concentration [S] and the Michaelis-Menten constant K_M to generate comparable data [8].

$$K_{\mathrm{i}} = \frac{\mathrm{IC}_{50}}{1 + \frac{[S]}{K_{\mathrm{M}}}} \qquad \textbf{eq. 2}$$

Time-dependent behavior

For inhibitors that behaved time-dependently during the measurements (i.e., F-t-diagrams show reducing slopes over time), K_i^{app} and k_{inact} values using defined dilution series for each compound were calculated with GraphPad PRISM by fitting the k_{obs} values (derived from plotting the measured F-t-diagram to **eq. 3**) to each concentration. The irreversible reaction necessary for this type of evaluation was shown with rhodesain (**SI-Figure 6**).

$$F_t = F_0 + (F_{max} - F_0) * (1 - e^{-k_{obs} * t}) \qquad \textbf{eq. 3}$$

$$k_{obs} = \frac{k_{inact}[I]}{K_I^{app} + [I]} \qquad \textbf{eq. 4}$$

It is assumed that the inhibitors bind competitively in respect to the substrate [7]. Due to the dependence of the K_i^{app} value on the substrate affinity and concentration, the K_i values were calculated with the Cheng-Prusoff equation, using the final substrate concentration [S] and the Michaelis-Menten constant K_M to generate comparable data.

$$K_{\mathrm{I}} = \frac{K_I^{app}}{1 + \frac{[S]}{K_{\mathrm{M}}}} \qquad \textbf{eq. 5}$$

MALDI-TOF experiment

Mass spectrometric experiments were performed as described before [9]. Lyophilized rhodesain was first reconstituted at 10 µM in buffer containing reducing agent (pH = 5.5, 50 mM NaOAc, 200 mM NaCl, 5 mM EDTA, 5 mM DTT) and incubated at r.t. for 30–60 min to ensure full activation of the enzyme. Afterwards, each inhibitor was added to 100 µL of 10 µM activated rhodesain solution (final concentrations: 100 µM inhibitor, 2.5 % DMSO).

Prior to MS analysis, protein alone or protein-inhibitor complexes were desalted by using Zeba Spin Desalting Columns (7 kDa MWCO, 0.5 mL; Thermo Fisher Scientific) in accordance with the manufacturer's instructions. On the target, desalted sample solutions were mixed 1:1 with a MALDI-matrix: Sinapinic acid, saturated solution in ACN/water 1:1 with 0.1 % TFA. Then, the mixtures were left in the fume hood for cocrystallization until dry.

After evaporation of the solvents (ca. 15 min), measurements were carried out on a rapifleX MALDI-TOF/TOF mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). The instrument is equipped with a scanning smart beam 10 kHz Nd:YAG laser at a wavelength of 355 nm and a 10 bit 5 GHz digitizer. The acceleration voltage was set to 20 kV and the mass spectra were recorded in positive ion linear mode. Calibration was done with the Bruker protein calibration standard II in a mass range from 10 to 70 kDa. Samples were measured at a laser power of 70–100 % with random walk ionization across the sample spot. As control sample, rhodesain incubated with compound K11777, a known covalent irreversible inhibitor, was used which showed the expected mass shift representative of covalent adduct formation (data now shown).

Data analysis was performed using the open-source software mMass [10]. $[M+H]^+$ was evaluated, so cropping from 22–27 kDa was performed. Signal intensity was normalized to highest signal in this range. Baseline was corrected with standard settings. Mass shifts were calculated from rhodesain signal in negative control (m/z was 23383 ± 22 Da (n = 5)) subtracted from adduct peak in the respective sample, if present.

PAMPA

A general literature-known principle was used [11]. Propranolol hydrochloride was sourced from Changzou Yabang Pharmaceutical (Fagron GmbH & Co. KG). Candesartan cilexetil was sourced as a chemical reference standard from the EDQM (catalogue code Y0001388). Candesartan was liberated and purified from this as described at the end of the synthesis section.

Incubation setup consisted of donor (top) plate (Sigma Aldrich, MAIPNTR10), 5 µL artificial membrane (1 % (w/v) L-α-phosphatidyl choline, Sigma Aldrich P3556, in *n*-dodecane, Sigma Aldrich 8205430100), acceptor (bottom) plate (Greiner, 655074). Measurement setup for direct spectroscopy consisted of UV-transparent measurement plate (Greiner UV-Star®, 655801), Tecan Spark 10M® well plate reader, 200 µL sample volume, λ = 200–650 nm. Analysis of absorption spectra was performed with AUC function in GraphPad Prism. Measurement setup for LCMS consisted of an Agilent 1100 series HPLC system coupled to an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI). An Agilent Poroshell 120 EC-C18, 150x2.10 mm, 4 µm column or an Agilent Zorbax SB-Aq 4.6 × 150 mm, 5 µm column was used. A linear gradient was used for elution with a ternary pump using [water/ACN/water + 0.1 % formic acid] that changes ratios from 80/10/10 to 0/90/10 over the course of 6 min, followed by 4 min of isocratic elution (0.7 mL/min). Injection volume was 100 µL. Areas of peaks in the chromatogram (detection-λ = 210 nm) were used to calculate the AUCs. Retention time and mass spectrum recorded in positive ionization mode were used to assign species. Analysis was performed using MestreNova.

Compound solutions were diluted from 20 mM stock solution in DMSO to 100 µM in a buffered (DPBS pH 7.4 or TRIS 50 mM pH 7.4) aqueous solution with 5 % (TRIS) or 50 % (DPBS) final content of DMSO ("donor solution"). Similarly prepared solutions (buffer + solvent) were used as "acceptor solutions". 150 µL of donor solution was applied onto the artificial membrane which had been applied first to the donor plate. This was sealed (Greiner, 676070, Viewseal sealer). 400 µL of acceptor solution was applied to the acceptor plate. Incubation setup was assembled and left for 7 h. After this time, acceptor and reference solutions were analyzed.

The experiment was performed in duplicates on each day. The acceptor solutions of those were measured separately (direct spectroscopic detection) or mixed 1:1 prior to analysis (LCMS detection). Reference solutions were prepared by simply mixing the indicated volumes of donor and acceptor solutions at the start of the incubation period and analyzed later with the acceptor solutions.

Calculations of apparent permeability P_{app} were performed using the following equation with V_D and V_A as volumes of donor and acceptor solutions (0.15 cm^3 and 0.4 cm^3), AUC_A and AUC_{Eq} as the area of the measured and baseline-corrected spectrum of sample and reference solutions, A as the porosity-corrected filter area (0.3019 cm^2*0.7 = 0.2113 cm^2) and t as the incubation time given in seconds.

$$P_{app} = -\frac{V_D * V_A * \ln\left(1 - \frac{AUC_A}{AUC_{Eq}}\right)}{(V_D + V_A) * A * t} \qquad \textbf{eq. 6}$$

All compounds except for esters **20** and **21** were evaluated under standard conditions suitable for both LC-based and direct spectroscopic quantification (5% DMSO, TRIS buffer pH 7.4). Since these esters were found to have poor aqueous solubility to an extent that complicated detection, 50 % DMSO in DPBS

was used for these compounds to ensure dissolution and spectroscopy was used for quantification (**2**, **3**, propranolol and methotrexate were evaluated like this additionally and behaved identically in both conditions).

Docking

Docking was performed as described previously [7]. In short, the crystal structure containing K11777 (pdb: 2p7u) [12] was used for docking without retention of crystallized water. The ligands were energy minimized with MMFF94x forcefield [13] in MOE (v.2022.02, Molecular Operating Environment, Chemical Computing Group Inc., Canada). Docking calculations were performed with FlexX (v.2.3.2, LeadIT/FlexX, BiosolveIT GmbH, Germany) and the top 10 poses were manually inspected for binding interactions in Pymol (v.2.4.0, Schrödinger).

QM calculations

All calculations were performed with the ORCA 5.0.4 program package [14]. Geometry optimizations were performed with ωB97X-D3 [15,16] \ma-def2-SVP [17,18] with the AutoAux auxiliary basis sets [19]. All stationary points were confirmed by frequency analysis and implicit solvation in water was included with the CPCM solvation method [20]. Free energies included a concentration correction resulting from the change in standard states going from gas phase to condensed phase [21,22]. The structures depicted in **SI-Figure S9** were utilized as model compounds with methyl thiolate as the model nucleophile. Starting in the π-complex, the S-CR^1 distances were gradually shortened in a relaxed scan. Subsequently, the highest energy point was subjected to a TS optimization to confirm a true transition state. The same procedure was conducted for TS2, elongating the C-R^1 distances starting from the σ-complex. The C-R^1 and S-CR^1 distances for all calculated structures are given in **SI-Table S2**.

SI Figure S9: Model compounds utilized for QM reaction path calculations.

SI-Table S2: Distances of the leaving group to the attacked aromatic carbon atom (C-R[1]) as well as from the thiolate to the carbon atom (S-CR[1]) for all computed structures of compounds **I**, **3**, and **9**.

[Å]	Cpd **I**		Cpd **3**		Cpd **9**	
	C-R[1]	S-CR[1]	C-R[1]	S-CR[1]	C-R[1]	S-CR[1]
π-complex	1.34	4.26	1.42	4.68	1.32	4.21
TS1	1.36	2.43	1.35	2.47	1.33	2.57
σ-complex	1.43	1.85	1.44	1.86	1.44	1.84
TS2	1.75	1.80	1.81	1.80	1.85	1.77
P	-	1.76	-	1.75	-	1.74

”

4: Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres.

An arginine in P1 position is a common denominator for affinity against a panel of serine proteases with trypsin-like substrate preference. It conveys ionic interactions with an anchoring aspartate moiety in the active site, which is exploited for both peptidomimetic inhibitors (e.g., melagatran)[617], as well as structurally unrelated inhibitors (WX-UK1)[618] which mainly rely on these non-covalent interactions for affinity. Two of these serine proteases are urokinase-type plasminogen activator (uPA), which is involved in the remodeling of tumor microenvironments, and transmembrane protease serine subtype 2 (TMPRSS2), which is involved in cell entry of coronaviruses like SARS-CoV-2. In preliminary studies on uPA and TMPRSS2, inhibitors based on a tri- or tetrapeptide backbone with capped *N*-terminus and an α-keto benzothiazole warhead, which conveys a covalent reversible mode of inhibition, were evaluated.[148,613]

In this study, structural modifications included the substitution of the P1 arginine for different aliphatic or aromatic guanidines, as well as amino acid exchange and shortening of the P2–P4 positions, and variations on the benzothiazole portion of the warhead. These modifications were assessed in the context of their influence on affinity towards the target proteases and structurally related off-targets, as well as in the context of their influence on passive permeability. It is established that most inhibitors bearing arginine-mimetics have poor permeability. This is not always an issue in cellular studies, as the target in these two cases are located on the outer side of the cell membrane, but with regards to in vivo studies, oral bioavailability is desirable. The α-keto benzothiazole warhead is known to positively contribute to lipophilicity and through that, also to permeability[619], so it was evaluated how inhibitors combining the polar arginine-mimetics and the apolar α-keto benzothiazole fair in a parallel artificial membrane permeation assay. While no passive permeation could be observed, the arginine-mimetic moiety seems to at least alleviate solubility problems that can occur for more lipophilic backbones with an α-keto benzothiazole warhead.[620] Both low permeability and good solubility is attributable to their high basicity ($pK_a \geq 10$), which conveys a level of polarity that the lipophilic warhead cannot compensate, underlined by the low $logD_{7.4}$ values. The cyclohexyl guanidine for uPA and the phenyl guanidine for TMPRSS2, respectively, were found to be the most optimal arginine-mimetic in terms of affinity and selectivity. The modulations on the warhead resulted in not more than a moderate increase in affinity, but a reduction in selectivity. Lastly, a special focus can be placed on a shortened dipeptide inhibitor that has exceptional affinity towards TMPRSS2 with good selectivity, and the most promising $logD_{7.4}$ value despite no obvious improvement to passive permeability. Taken together, this study provides a selection of inhibitors with high affinity and good selectivity that are suitable for extended cell-based assays due to their chemical stability and inhibition potency. However, it could also be shown that masking strategies for the arginine-mimetics as prodrugs (e.g., hydroxy guanidine, carbamate)[621,622] or cyclic amidines[623], or alternative approaches to increased permeability through formulation are required for

further drug development efforts. The study also provided inhibitor **33** with improved ligand efficiency and the most drug-like physicochemical parameters suitable for further optimizations.

Own contribution: Protein similarity calculations for the employed serine proteases, spectroscopic characterization of the α-keto hetaryl warheads, LCMS-based assessment of the inhibitors' stability in an aqueous solution, calculation of physicochemical parameters (pK_a, logP, $logD_{7.4}$) for arginine-mimetic inhibitors, PAMPA assessment of passive permeability, writing parts and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of the in vitro affinity and selectivity, writing and editing of the manuscript.

This work was published in the International Journal of Molecular Sciences.

Article reprinted with permission of *Int. J. Mol. Sci.* **2024**, *25*(3), 1375, "Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres", Copyright © 2024 The Authors. Published by MDPI.

The following publication quoted within "" (pages 176–201) is the same as the manuscript cited here. The appended Supporting Information represents an abridged version. The full version can be accessed online at doi:10.3390/ijms25031375.

“

International Journal of
Molecular Sciences

MDPI

Article

Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres

Patrick Müller, Collin Zimmer, Ariane Frey, Gideon Holzmann, Annabelle Carolin Weldert and Tanja Schirmeister *

Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany; muelpat@uni-mainz.de (P.M.); cozimmer@uni-mainz.de (C.Z.); arfrey@uni-mainz.de (A.F.); gholzman@students.uni-mainz.de (G.H.); anwelder@uni-mainz.de (A.C.W.)
* Correspondence: schirmei@uni-mainz.de; Tel.: +49-6131-39-25742

Abstract: Trypsin-like serine proteases are involved in many important physiological processes like blood coagulation and remodeling of the extracellular matrix. On the other hand, they are also associated with pathological conditions. The urokinase-pwlasminogen activator (uPA), which is involved in tissue remodeling, can increase the metastatic behavior of various cancer types when overexpressed and dysregulated. Another member of this protease class that received attention during the SARS-CoV 2 pandemic is TMPRSS2. It is a transmembrane serine protease, which enables cell entry of the coronavirus by processing its spike protein. A variety of different inhibitors have been published against both proteases. However, the selectivity over other trypsin-like serine proteases remains a major challenge. In the current study, we replaced the arginine moiety at the P1 site of peptidomimetic inhibitors with different bioisosteres. Enzyme inhibition studies revealed that the phenylguanidine moiety in the P1 site led to strong affinity for TMPRSS2, whereas the cyclohexylguanidine derivate potently inhibited uPA. Both inhibitors exhibited high selectivity over other structurally similar and physiologically important proteases.

Keywords: trypsin-like serine proteases; covalent reversible inhibitors; enzyme inhibition study; protease inhibitors; peptidomimetic sequence; arginine bioisosteres

check for updates

Citation: Müller, P.; Zimmer, C.; Frey, A.; Holzmann, G.; Weldert, A.C.; Schirmeister, T. Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres. *Int. J. Mol. Sci.* **2024**, *25*, 1375. https://doi.org/10.3390/ijms25031375

Academic Editor: Asim Debnath

Received: 3 January 2024
Revised: 20 January 2024
Accepted: 21 January 2024
Published: 23 January 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With over 600 different proteins, proteases represent an important class of enzymes [1]. Approximately one-third of all known proteolytic enzymes are serine proteases [2]. According to the MEROPS database of peptidases, these enzymes are classified into clans by their catalytic mechanism and into families on the basis of a common ancestry [3]. The largest family of serine proteases are the trypsin-like proteases (TLPs). The catalytic triad of TLPs harbors a nucleophilic serine residue in combination with aspartate and histidine, which increase the nucleophilicity of the serine. The trypsin-like substrate specificity is characterized by the positively charged side chain of arginine or lysine in the P1 position [3,4]. Numerous important physiological processes rely on trypsin-like serine proteases. This includes hemostasis, the immune response system and extracellular matrix remodeling [5–8]. Dysregulation of these enzymes can lead to severe pathological incidents, which range from cardiovascular disorders to cancer progression or neurodegenerative and inflammation processes [8–10]. Moreover, proteases often are virulence factors in infectious diseases. As an example, tropical and subtropical countries are heavily affected by dengue virus infections, where the viral NS2B-NS3 trypsin-like serine protease is essential for the replication process of the virus [11]. Undoubtedly, this class of enzymes includes promising targets in various diseases, and the scientific community still strives to discover more drug candidates [7].

The urokinase-type plasminogen activator (uPA) is one member of the trypsin-like serine proteases. The enzyme is involved in the fibrinolytic system [12]. The binding of

uPA to its specific glycolipid-anchored uPA receptor (uPAR) on cell surfaces enables the conversion of plasminogen to the serine protease plasmin [13]. This mediates extracellular proteolysis and the activation of several further proteases, like activating growth factors and metalloproteases, which catalyze the degradation and remodeling of extracellular matrix components [14,15]. Unfortunately, pathophysiological mechanisms like tumor angiogenesis, tumor progression and metastasis profit from these events, and therefore, inhibition of this protease could be beneficial for the mitigation, or even prevention, of tumor proliferation (Figure 1 left side) [14,15]. Blocking of the catalytic activity was achieved by specific antibodies, overexpression of the endogenous inhibitors PAI-1 and small-molecule inhibitors [16–18]. One of the most promising peptidomimetic inhibitors, mesupron® (upamostat, WX-671, RHB-107, Wilex AG, Heidelberg, Germany), led to reduced metastasis and extended lifespan in clinical trials on pancreatic and breast cancer patients [19]. Hence, uPA can be considered as a promising drug target to block tumor dissemination.

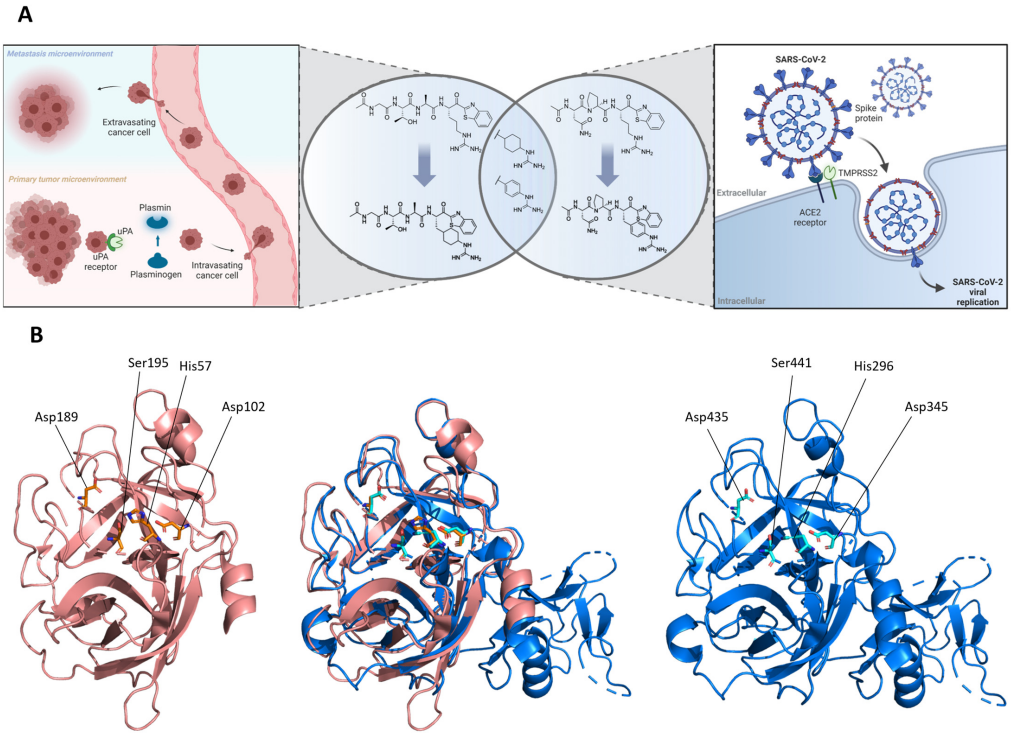


Figure 1. (**A**) Ligand-based design of covalent-reversible uPA (**left**) and TMPRSS2 (**right**) inhibitors with arginine bioisosteres. (**A**) Created with BioRender.com. (**B**) Visualization of the uPA structure (red, PDB: 7VM4) containing the catalytic triad (Asp102, His57, Ser195) and Asp189 in the S1 pocket, TMPRSS2 (blue, PDB: 7MEQ) containing the catalytic triad (Asp345, His296, Ser441) and Asp435 in the S1 pocket and the superimposition of both proteases (red/blue). (**B**) Created with PyMOL (Version 2.4.0, Schrödinger, LLC, New York, NY, USA).

Another proteolytic enzyme that belongs to the trypsin-like serine proteases is the human transmembrane protease serine subtype 2 (TMPRSS2). It has been shown to play an important role for viral host cell entry, and received increased attention during the SARS-CoV-2 pandemic due to its ability to enable cell entry and spread of the coronaviruses SARS-CoV-2, SARS-CoV and MERS-CoV [20–24]. The entry of these viruses is mediated by

the spike protein, which is located at the viral cell surface. TMPRSS2 processes the spike protein after binding of the virus to the angiotensin-converting enzyme 2 receptor (ACE2), initiating the entry into lung cells (Figure 1 right side) [23,25]. Additionally, viral cell entry can occur via the endosomal pathway, whereby the spike protein is processed by cathepsin L [26]. Studies have demonstrated that inhibition of TMPRSS2 blocks the viral host cell entry and replication of SARS-CoV-2 in lung epithelial Calu-3 cells [27,28]. Previous work, in cooperation with Mailänder et al. showed that peptidomimetic inhibitors efficiently reduce TMPRSS2 activity, block SARS-CoV-2 spike-driven entry and prevent SARS-CoV-2 infection in CaCo-2 cells [29]. This highlights the opportunity for an alternative therapeutic strategy, besides targeting of the viral host proteases papain-like protease (PL^{pro}) and the 3C-like- or "main protease" (3CL- or M^{pro}) [30–32].

In the past decades, several uPA inhibitors have been disclosed, most of them with non-covalent reversible or covalent-irreversible inhibition mode [16,33]. On the contrary, only few covalent-reversible inhibitors are found in the literature [34]. Such inhibitors could combine the benefits from both concepts: the high-affinity properties and extended residence time by covalent modification of the catalytic serine residue and the reduced risk for unwanted side effects and toxicity by a reversible binding mechanism [35–37]. Furthermore, in order to minimize the risk for side effects, it is of great importance to inhibit the target protease selectively. This, however, is a major challenge due to the high structural similarity within the trypsin-like serine protease family.

In 2021, the group around Huang et al. created a homology structure model of the TMPRSS2 serine protease domain, and revealed a high similarity between the homology model and the structure of the uPA [38]. This led to the idea to transfer the design of the synthesized uPA inhibitors to the previously published TMPRSS2 inhibitors, to receive an improved set of inhibitors in terms of off-target selectivity (Figure 1) [29].

Herein, we describe the ligand-based development of peptidomimetic inhibitors, which started with Ac-Gly-L-Thr-L-Ala-L-Arg-ketobenzothiazole (kbt) as a covalent-reversible uPA inhibitor discovered in previous work [39]. We substituted the P1-arginine moiety with a variety of bioisosteres, inspired by the serine protease inhibitor camostat, and furthermore modified the benzothiazole structure [40,41]. The cyclohexyl-and phenylguanidine moiety presented the most promising results during the enzyme inhibition studies. Therefore, we translated this structure motif to the suitable peptide sequence Ac-L-Asn-L-Pro-L-Arg-kbt from our previous work towards TMPRSS2 [29]. Within this study, we successfully enhanced the affinity and selectivity for both main-target proteases by systematic variation of different structural elements.

2. Results

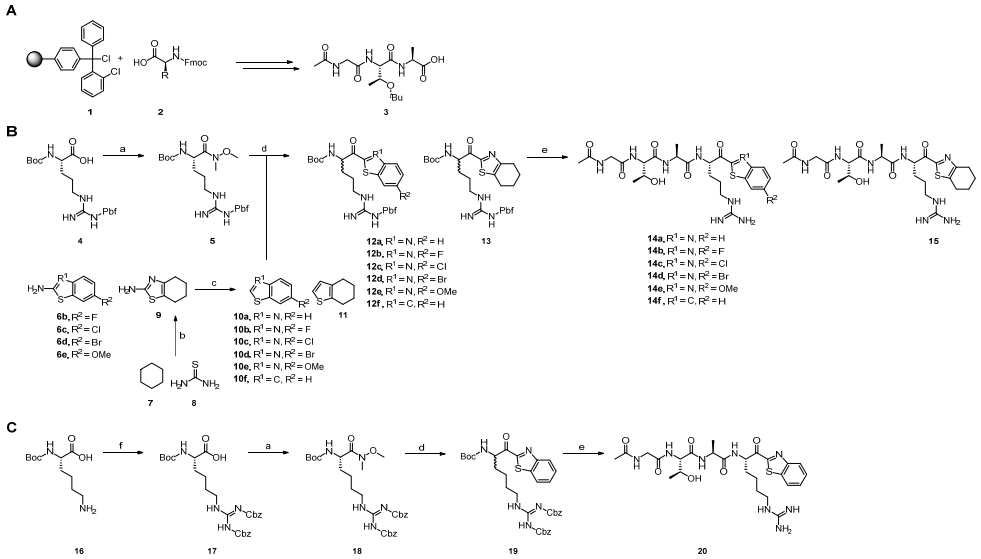
2.1. Chemistry

All tested peptidomimetic inhibitors were synthesized in multistep reactions. First, the peptide sequences (P2–P4) of the inhibitors were prepared via a standard fmoc solid phase peptide synthesis (SPPS) protocol, which is described in detail in the Supporting Information. The P1 derivatives with the ketobenzothiazole moiety as warhead were prepared as described in Schemes 1 and 2.

2.1.1. Synthesis of the (Homo)arginine-Based Inhibitors

Boc-protected N_ω-2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl (pbf)-L-arginine **4**, which was used as the starting material for the arginine-based inhibitors **14a**–**f** and **15**, was modified to the Weinreb amide **5**. The ketobenzothiazole derivatives **12a**–**f** and **13** were obtained by alkylation of **5** with the respective heterocycles **10a**–**f** and **11**. The benzothiazole **10a** and the benzothiophene **10f** were commercially available, whereas the 6-fluoro-, 6-chloro-, 6-bromo-, 6-methoxybenzothiazoles **10b**–**e** and 4,5,6,7-tetrahydrothiazole **11** had to be synthesized by desamination of the commercially available 2-amino precursors **6b**–**e** and **9**. The 2-amino-4,5,6,7-tetrahydrothiazole **9** was prepared from cyclohexane **7** and thiourea **8** with iodine. The preparation of the homoarginine inhibitor **20** started with

the guanylation of boc-protected L-lysine **16** with *N*,*N*′-bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine, yielding compound **17**. Afterwards, **17** was converted to the ketobenzothiazole **19**, in analogy to the arginine derivatives. After boc-deprotection of the amino group, the P1 precursor derivatives were coupled with the Ac-Gly-L-Thr(O^{t}Bu)-L-Ala-OH peptide **3** using 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo [4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU) as the coupling reagent. Final deprotection of the (homo)arginine and threonine side chain under acidic conditions and purification via RP-HPLC yielded the inhibitors **14a–f**, **15** and **20**.



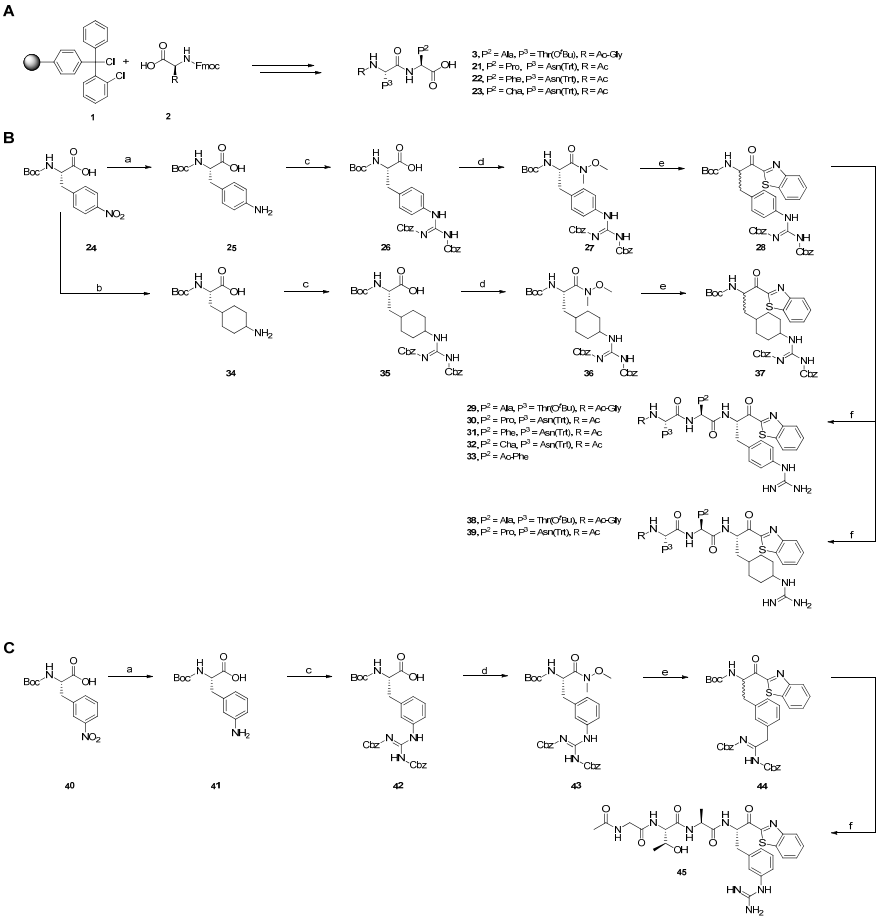
Scheme 1. Synthesis of the (homo)arginine-based inhibitors **14a–f**, **15**, **20**. (**A**) Solid phase peptide synthesis of peptide sequence **3**. (**B**) Preparation of the final arginine compounds **14a–f**, **15**. (**C**) Synthesis of the homoarginine inhibitor **20**. Reaction conditions: (a) *N*,*O*-dimethylhydroxylamine·HCl, TBTU, DIEA, DCM, rt, 12 h, 70–90%; (b) cyclohexane, thiourea, iodine, 130 °C, 12 h, 53%; (c) isopentyl nitrite, THF, reflux, 2–5 h, 46–76%; (d) benzothiazole-derivative, *n*-BuLi, THF, −78 °C, 2 h, 50–88%; (e) 1. TFA, DCM, rt, 0.5 h, 2. HATU, DIEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 4–10%; (f) *N*,*N*′ bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine, Et_3N, DMF, rt, 72 h, quant.

2.1.2. Synthesis of the Phenyl/Cyclohexylguanidine-Based Inhibitors

The preparation of the *p*-phenyl- and *p*-cyclohexylguanidine-based inhibitors both started with boc-protected *p*-nitro-L-phenylalanine **24**. The reduction of the nitro group was carried out with 5% Pd/C in methanol to yield **25**, whereas the hydrogenation of the benzene ring and the nitro group using the Adam's catalyst under acidic conditions yielded the cyclohexane derivative **34**. The amine group of both compounds was guanylated with *N*,*N*′-bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine. The bis-cbz-protected intermediates (**26**, **35**) were converted, in a similar way to the arginine-based inhibitors (Section 2.1.1), to the Weinreb amides **27** and **36** and later to the ketobenzothiazole derivatives **28** and **37.** After removal of the boc-protecting groups, the respective peptide sequences **3**, **21**–**23**, which were synthesized via a standard fmoc solid-phase synthesis (SPPS), were coupled with the *p*-phenyl- and cyclohexylguanidine precursor derivatives. After final deprotection of the side chains in TFA/DCM and purification via RP-HPLC, the inhibitors **29**–**33** and **38**–**39** were obtained. Starting with *m*-nitro-L-phenylalanine **40**, the inhibitor **45** was prepared in analogy to the 5-step synthetic process of the *p*-phenylguanidine derivatives **29**–**33**.

Int. J. Mol. Sci. **2024**, *25*, 1375

The synthesized final compounds **14a–f**, **15**, **20**, **29–33**, **38–39**, **45** showed two peaks with identical m/z ratio and similar retention times in initial chromatographic analyses. This is due to the partial epimerization of the α-carbon in the P1 amino acid portion during the reaction of the Weinreb amide with lithium-benzothiazole solution. Since the faster eluting epimer was always isolated via RP-HPLC in very large excess, while the other diastereomer was obtained only in traces, we supposed the first one to be the L-epimer, and used it for all inhibition studies [42].



Scheme 2. Synthesis of the phenyl/cyclohexylguanidine-based inhibitors **29–33**, **38–41**, **45**. (**A**) Solid phase peptide synthesis of peptide sequences **3**, **21–23**. (**B**) Preparation of the final *p*-phenyl/cyclohexylguanidine compounds **29–33**, **38–41**. (**C**) Synthesis of the *m*-phenylguanidine inhibitor **45**. Reaction conditions: (a) 1. Pd/C 5%, H_2 (3 bar), MeOH, 3 h, (b) PtO_2, H_2 (3 bar), AcOH, MeOH, rt, 24 h, (c) *N*,*N*′-bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine, Et_3N, DMF, rt, 6 h, 90%–quant.; (d) *N*,*O*-dimethylhydroxylamine·HCl, TBTU, DIEA, DCM, rt, 12 h, 51–53%; (e) benzothiazole, *n*-BuLi, THF, −78 °C, 2 h, 64–84%, (f) 1. TFA, DCM, rt, 0.5 h, 2. HATU, DIEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 10–40%.

2.2. Enzyme Inhibition Studies

The inhibitory activity of the synthesized compounds towards the respective main- and off-target proteases was measured via fluorometric and colorimetric assays. Thus, fluorogenic AMC- or colorimetric *p*NA-based substrates with a peptide sequence suitable for the tested protease were utilized (see Supplementary Figure S2). At first, the compounds were screened against five proteases (uPA, TMPRSS2, matriptase, tPA, thrombin, factor Xa) at 20 µM, and a cut-off value of 80% inhibition at this concentration was set, for the differentiation between nonactive (n.a.) and active inhibitors. Due to the reversible inhibition mechanism of the ketobenzothiazole derivates, the IC_{50} values were determined with Graphpad Prism 9, and afterwards converted to the corresponding K_i values for an adequate comparison between the inhibitory activities of the compounds toward all tested proteases. The K_i values were calculated using the Cheng–Prusoff equation [43].

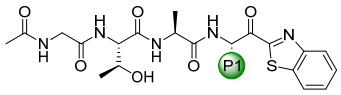
2.2.1. Inhibition Studies with uPA Inhibitors

At first, we investigated the selectivity profile of the starting compound **14a,** which exhibited good inhibition of uPA with a K_i value of 141 nM. **14a** was originally synthesized for the analysis of reactivity and selectivity studies of peptidomimetic covalent inhibitors [40]. The peptide sequence Ac-L-Gly-L-Thr-L-Ala-L-Arg was used, because of its literature-known selectivity for uPA vs. tPA [44]. Due to the similar and important physiological roles of uPA and tPA, a good selectivity is necessary to avoid severe side effects concerning ECM degradation and cell proliferation [5,45,46]. Furthermore, the trypsin-like serine proteases thrombin and factor Xa were chosen because of their important roles in blood coagulation, as well as matriptase as a representative of a transmembrane protease, which is involved in the remodeling of plasma membranes and other lipid matrix formations [47–49]. Due to their structural similarity (calculated sequence similarity is given in Tables 1–3) to uPA and their physiological roles, they resemble important off-targets. As expected, **14a** did not show inhibition of tPA, and only moderate selectivities for uPA towards thrombin (**14a** K_i = 4390 nM) and factor Xa (**14a** K_i = 3360 nM) with inhibition constants in the low micromolar range. In contrast, a lower K_i value was obtained for mapriptase (**14a** K_i = 32 nM). Exchanging the arginine side chain with a *p*-phenyl- or cyclohexylguanidine moiety enhanced the inhibitory properties. Both derivates resulted in more affine inhibitors (**29** K_i = 29 nM, **38** K_i = 39 nM), with a significant improvement in their selectivity profiles. The inhibitors **29**, **38** did not inhibit tPA, thrombin and factor Xa, and the selectivity indices for matriptase (**29** K_i = 132 nM, **38** K_i = 626 nM) were improved. The inhibitors **20** and **45**, which contain the homoarginine and *m*-substituted phenylguanidine moiety, did not show inhibition of all tested proteases at 20 µM, which highlights the importance of the alkyl chain length and the *p*-position of the guanidine element for proper binding into the S1 pocket. Additionally, all compounds were tested against the TMPRSS2 because of the aforementioned structure similarity to uPA [38]. The results indicated a strong affinity to the TMPRSS2 protease with K_i values in the nanomolar range of the arginine, phenyl- and cyclohexyl derivates (**14a** K_i = 5 nM, **29** K_i = 10 nM, **38** K_i = 73 nM). Based on these results, a SAR study with the phenyl- and cyclohexylguanidine moiety as arginine bioisosteres for new TMPRSS2 inhibitors was performed, which is described in Section 2.2.2 [29].

Besides the arginine replacement in the P1 position, we also evaluated the influence of modifications of the benzothiazole moiety (cpds. **14b**–**f**, **15**). The introduction of the electronegative halogen atoms fluorine, chlorine and bromine in position 6 led to an approximately two-fold increase in the affinity for the chloro- and bromo-derivates (**14c** K_i = 82 nM, **14d** K_i = 60 nM), and a three-fold loss of affinity for the fluoro-derivate (**14b** K_i = 388 nM). Other modifications, like the electron-donating methoxy group in position 6 (**14e** K_i = 178 nM), the exchange of the benzene ring system with a cyclohexyl ring (**15** K_i = 435 nM) or the replacement of the benzothiazole with a benzothiophene ring, led to a decrease in or complete loss of the affinity towards the uPA. Selectivity studies were performed with the chloro- and bromo-derivates **14c**–**d**, because they were

the only inhibitors with slightly better affinity than the nonsubstituted ketobenzothiazole inhibitor **14a**. They revealed similar affinity to TMPRSS2 (**14c** K_i = 9 nM, **14d** K_i = 6 nM) and matriptase (**14c** K_i = 59 nM, **14d** K_i = 38 nM), and reduced selectivity vs. thrombin (**14c** K_i = 456 nM, **14d** K_i = 450 nM) and factor Xa (**14c** K_i = 2447 nM, **14d** K_i = 2847 nM) in comparison to **14a**.

Table 1. Inhibition data (K_i values and selectivity indices [SI]) of the synthesized uPA inhibitors **14a**, **20**, **29**, **38**, **45** towards uPA, TMPRSS2, matriptase, tPA, thrombin and factor Xa.



14a, 20, 29, 38, 45

Compound		K_i [nM] [SI]					
P1		**uPA**	**TMPRSS2 55.6%**	**Matriptase 68.9%**	**tPA 66.7**	**Thrombin 60.0%**	**Factor Xa 62.2%**
	**14a**	141 ± 28	5 ± 1 [0.04]	32 ± 14 [0.2]	n.a.	4390 ± 1480 [31]	3360 ± 320 [24]
	**29**	29 ± 2	10 ±4 [0.3]	133 ± 3 [4]	n.a.	n.a.	n.a.
	**38**	39 ± 5	73 ± 16 [1.8]	626 ± 74 [16]	n.a.	n.a.	n.a.
	**20**	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	**45**	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

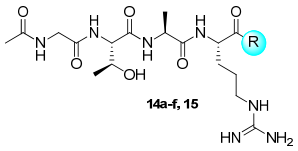
Selectivity indices [SI] represent the quotient of the K_i values for TMPRSS2, matriptase, tPA, thrombin and factor Xa to the K_i value for uPA; n.a. not active, i.e., <80% inhibition at 20 μM; calculated sequence similarity between the binding sites of uPA and the other proteases are given in %.

2.2.2. Inhibition Studies with TMPRSS2 Inhibitors

The selection of the peptide sequence Ac-L-Asn-L-Pro-L-Arg was based on the results of previous work. The published inhibitor Ac-L-Asn-L-Pro-L-Arg-kbt showed K_i values in the single-digit nanomolar range (K_i = 2.5 nM) and a good selectivity vs. thrombin (K_i = 1046 nM) [29]. Unfortunately, only slight selectivity could be observed over factor Xa (K_i = 41.1 nM), and almost no difference in inhibition potency between TMPRSS2 and matriptase (K_i = 5.2 nM). Therefore, we tried to improve the selectivity profile by substituting the arginine side chain in the P1 position with the previously used phenyl- and cyclohexylguanidine moiety. Both derivates **30**, **39** showed an increase in selectivity for TMPRSS2 towards matriptase, with the phenylguanidine-based compound being more affine for TMPRSS2 (**30** K_i = 5 nM) than the cyclohexyl derivate (**39** K_i = 44 nM), but also showing a better inhibition of the matriptase for the phenylguanidine derivate (**30** K_i = 60 nM, **39** K_i = 1198 nM). **30**–**33** and **39** did not inhibit tPA, thrombin and factor Xa. In addition, **30** and **39** showed a moderate selectivity for TMPRSS2 over uPA (**30** K_i = 479 nM, **39** K_i = 936 nM). Based on these results, and due to the overall good affinity and selectivity parameters, we decided to maintain the phenylguanidine moiety in the P1 position and implement P2 modifications with phenyl- and cyclohexylalanine (Phe, Cha) instead of proline. The latter is based on results obtained with hepsin inhibitors from the group of

Kwon et al. [40]. The inhibitor **31** with the P2 phenylalanine residue showed inhibition of TMPRSS2 in the subnanomolar range (**31** K_i = 0.4 nM) and a significant increase in selectivity over matriptase (**31** K_i = 252 nM) and uPA (**31** K_i = 3574 nM). The inhibitor **32** with the cyclohexylalanine residue in P2 position also showed very good selectivity over matriptase (**32** K_i = 3333 nM) and uPA (**32** K_i = 2688 nM), but less affinity to TMPRSS2 (**32** K_i = 34 nM). In an attempt to improve the drug-like properties of the designed inhibitors, we synthesized the shortened compound **33**. This led to a slightly less active TMPRSS2 inhibitor, but still in the low nanomolar range (**33** K_i = 5 nM). The selectivity profile for TMPRSS2 inhibition over matriptase (**33** K_i = 1443 nM) and uPA (**33** K_i = 5264 nM) is still very promising.

Table 2. Inhibition data (K_i values and selectivity indices [SI]) of the synthesized uPA inhibitors **14a**–**f**, **15** towards uPA, TMPRSS2, matriptase, tPA, thrombin and factor Xa.



Compound		K_i [nM] [SI]					
R		uPA	TMPRSS2 55.6%	Matriptase 68.9%	tPA 66.7	Thrombin 60.0%	Factor Xa 62.2%
	**14a**	141 ± 28	5 ± 1 [0.04]	32 ± 14 [0.2]	n.a.	4390 ± 1480 [31]	3360 ± 320 [23]
	**14b**	388 ± 95	-	-	-	-	-
	**14c**	82 ± 6	9 ± 1 [0.1]	59 ± 9 [0.7]	n.a.	456 ± 72 [5]	2447 ± 99 [30]
	**14d**	60 ± 15	6 ± 1 [0.1]	38 ± 2 [0.6]	n.a.	450 ± 69 [8]	2847 ± 844 [47]
	**14e**	178 ± 19	-	-	-	-	-
	**14f**	n.a.	-	-	-	-	-
	**15**	435 ± 33	-	-	-	-	-

Selectivity indices [SI] represent the quotient of the K_i values for TMPRSS2, matriptase, tPA, thrombin and factor Xa to the K_i value for uPA; n.a., not active, i.e., <80% inhibition at 20 µM; "-", not tested; calculated sequence similarity between the binding sites of uPA and the other proteases are given in %.

2.3. Parallel Artificial Membrane Permeation Assay (PAMPA)

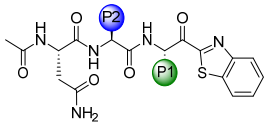
Following the envisioned applications of the presented inhibitors as drug leads for the treatment of cancer or viral infections, cell permeability is an important factor in the characterization process. Since both main targets (uPA and TMPRSS2) are membrane-located, extracellular structures, inhibitors seemingly do not require cell permeation to address their target. However, in the organismic contexts of oral bioavailability (facilitated application) and biodistribution (reaching target tissue), adequate permeation is an important quality. To assess this characteristic, PAMPA was used as a suitable model for passive permeation.

Generally, the inhibitor scaffold combines some favorable features: The Arg-like P1 amino acid (in combination with other hydrophilic amino acids like Asp and Thr)

Int. J. Mol. Sci. **2024**, *25*, 1375

ensures high aqueous solubility, even in the presence of the rather hydrophobic benzoheteroarenes. The latter motif conveys reliable detectability by spectroscopy-based methods (λ_{max} = 305–350 nm, depending on substitution pattern). The inhibitors also were found to be sufficiently stable in the utilized aqueous system (50 mM TRIS, pH = 7.4) over the course of the assay (7 h at room temperature) and under elevated temperature conditions (17 h at 37 °C). Computed physicochemical properties, absorption spectra and stability studies are depicted in Supplementary Figures S4 and S5 and Table S1.

Table 3. Inhibition data (K_i values and selectivity indices [SI]) of the synthesized TMPRSS2 inhibitors **30**–**33**, **39** towards TMPRSS2, uPA, matriptase, tPA, thrombin and factor Xa.



30-33, 39

Compound			K_i [nM] [SI]					
P2	**P1**		**TMPRSS2**	**uPA 53.2%**	**Matriptase 66.7%**	**tPA 62.2%**	**Thrombin 57.8%**	**Factor Xa 66.7%**
Pro	Arg [31]		2.5	-	5.2 [2]	-	1046 [418]	41.1 [16]
Pro		**30**	5 ± 1	479 ± 17 [96]	60 ± 8 [12]	n.a.	n.a.	n.a.
Pro		**39**	44 ± 7	936 ± 100 [21]	1198 ± 181 [27]	n.a.	n.a.	n.a.
Phe		**31**	0.4 ± 0.1	3574 ± 421 [>5000]	252 ± 67 [630]	n.a.	n.a.	n.a.
Cha		**32**	34 ± 4	2688 ± 217 [79]	3333 ± 1141 [98]	n.a.	n.a.	n.a.
		**33**	5 ± 2	5264 ± 922 [1052]	1443 ± 90 [289]	n.a.	n.a.	n.a.

Selectivity indices [SI] represent the quotient of the K_i values for uPA, matriptase, tPA, thrombin and factor Xa to the K_i value for TMPRSS2; n.a., not active, i.e., <80% inhibition at 20 µM; "-", not tested; calculated sequence similarity between the binding sites of TMPRSS2 and the other proteases are given in %.

However, all presented compounds were found to have very low permeabilities ($P_e < 1 \times 10^{-6}$ cm/s) without any indication of improvement between the structural modifications (as exemplified for **38** in Figure 2). This result is not surprising. The pK_a (of the protonated guanidine function) of all compounds is calculated to be ≥10 (Marvin JS 23.11.0), meaning that in assay (or physiologic) conditions, all compounds are expected to be fully (≥99.75%) protonated, and therefore remarkably hydrophilic. Most of the presented compounds have negative $logD_{7.4}$ values, with **33** being the exception ($logD_{7.4}$ = 1.2; compare Supplementary Table S1). This level of lipophilicity, however, was still not enough to exert measurable permeability. For approved drugs with similar structural characteristics (e.g., camostat, melagatran, xylometazoline, metformin), only very limited permeabilities are described as well [50–53]. All this indicates the pronounced hindering effect of the guanidine group for passive permeation.

The discussed properties of the presented compounds can be paralleled to BCS class III compounds, namely their high aqueous solubility and low permeability. For these types of drugs, one major option to improve permeability is to remove charge from the molecule. In amidine-containing drugs, where charge is almost pH-independent due to their immense basicity, this was addressed by conversion to the amidoxime (ximelagatran or mesupron®) or carbamate prodrugs (dabigatran) with lower basicities [19,50,54]. For the guanidine moiety, the conversion to *N*-hydroxyguanidine is possible [55]. In a technological approach

to improved absorption, possible options for oral application are the formulation with permeation enhancing agents, or lipophilic counter ions [56]. For intravenous applications, nanoparticular formulations can be applied (e.g., for doxorubicin or for protease inhibitors) [57,58]. Of course, combinations of both chemical and technological approaches should be employed for optimization.

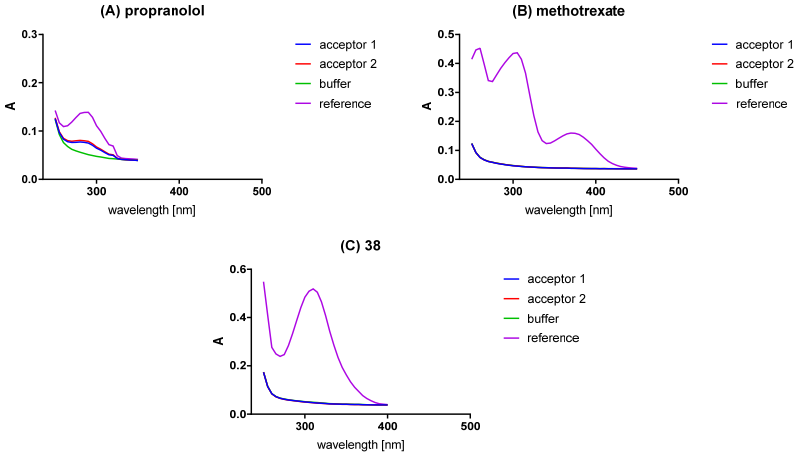


Figure 2. (**A**) Analysis of PAMPA permeable control propranolol HCl. (**B**) Analysis of PAMPA impermeable control methotrexate. (**C**) Analysis of PAMPA for **38**, representative of all synthesized final compounds of this study. All absorption spectra were baseline-corrected with the measured buffer spectrum at $\lambda \geq 450$ nm prior to AUC calculations. The reference spectra represent manually prepared samples, equivalent to a maximally permeated experimental sample. The acceptor spectra result from membrane permeation experiments and their AUC are proportional to the concentration of permeated compound.

3. Discussion

Trypsin-like serine proteases present attractive drug targets for treatment against many diseases, which can be of malignant cellular or viral origin [9,11]. Over the past decades many potent inhibitors were designed with remarkable affinity for the target protease. But most of them lack selectivity because of the highly structural similarity between the proteases. Within our study, we describe a systematic ligand-based approach to enhance affinities and selectivities. Starting from the previous published covalent reversible ketobenzothiazole inhibitor Ac-Gly-L-Thr-L-Ala-L-Arg-kbt **14a**, we modified the P1 arginine side chain with different bioisosteres [39]. The results indicate that the cyclohexylguanidine moiety fits best for uPA inhibition. The inhibitor **38** showed remarkable inhibition with a K_i value of 39 nM and a very good selectivity profile towards the other trypsin-like serine proteases. The modification of the benzothiazole moiety did not improve either the inhibitory properties nor the selectivity profiles, rendering the original kbt warhead the most promising.

The transfer of the P1-arginine replacement with the promising phenyl- and cyclohexylguanidine moieties to the previously published TMPRSS2 inhibitor Ac-L-Asn-L-Pro-L-Arg-kbt was a success, leading to a subnanomolar TMPRSS2 inhibitor **31** (K_i = 0.4 nM), with significantly increased selectivity over other trypsin-like serine proteases [29]. Furthermore, the shortened peptide sequence of the TMPRSS2 inhibitor **31** led to the more drug-like candidate **33,** with still very good inhibitory and selectivity properties. In terms of permeability, the inhibitor scaffold (and especially the shortened compound **33**) leaves the opportunity for improvement in a focused structure-permeability relationship study.

4. Materials and Methods

The materials as well as the methods used for this study are described in the Supporting Information. The authors have cited additional references within the Supporting Information [29,39,40,43,59–72]. Supplementary Figures of the protein similarity calculation (Figure S1), fluorometric inhibition assays (Figures S2 and S3), absorption spectra (Figure S4), stability studies (Figure S5), NMR-spectra and HPLC-chromatograms (Figures S6a–S21c) and Table S1 of the computation of physicochemical parameters can be accessed in the supporting information.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25031375/s1.

Author Contributions: P.M.: design, synthesis and enzyme inhibition studies, writing—original draft; C.Z.: parallel artificial membrane permeation assay (PAMPA), absorption spectra, calculation of physicochemical properties, calculation of protein similarity, stability studies, writing—original draft; A.F.: synthesis and enzyme inhibition studies; G.H.: synthesis and enzyme inhibition studies, A.C.W.: protein preparation and purification; T.S.: validation, review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and Supplementary Materials.

Acknowledgments: We thank Michael Klein for creating the graphical abstract and figures with BioRender.com. The authors wish to thank Torsten Steinmetzer for sharing the pQE-vector containing human matriptase and Simon Huber for sharing his expertise in matriptase 1 preparation.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

Ac	acetyl
Ac-Gly	acetylglycine
AcOH	acetic acid
Ala	alanine
AMC	7-amino-4-methyl coumarine
Asn	asparagine
Boc:	*t*-butyloxy carbonyl
Cbz	benzyloxycarbonyl
Cha	cyclohexylalanine
DCM	dichloromethane
DIEA	*N*,*N*-diisopropylethylamine
DMSO	dimethylsulfoxide
Et_3N	triethylamine
Fmoc	9-Fluorenylmethoxycarbonyl
h	hours
HATU	2-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate
MeOH	methanol
MERS-CoV	Middle East Respiratory Syndrome
n-BuLi	*n*-butyllithium
O^{t}Bu	*O*-*tert*-butyl
Pbf	N_ω-2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl
Phe	phenylalanine
*p*NA	*para*-nitroanilide
Pro	proline
quant	quantitative
RP-HPLC	reversed phase high pressure liquid chromatography

rt	room temperature
SARS-CoV-2	severe acute respiratory syndrome coronavirus type 2
SPPS	solid phase peptide synthesis
TA	thioanisole
TBTU	2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate
TFA	trifluoroacetic acid
THF	tetrahydrofurane
TLP	trypsin-like serine protease
TMPRSS2	transmembrane protease serine subtype 2
tPA	tissue-type plasminogen activator
trt	triphenylmethyl
uPA	urokinase-type plasminogen activator
uPAR	urokinase-type plasminogen activator receptor

References

1. López-Otín, C.; Overall, C.M. Protease degradomics: A new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 509–519. [CrossRef] [PubMed]
2. Di Cera, E. Serine proteases. *IUBMB Life* **2009**, *61*, 510–515. [CrossRef] [PubMed]
3. Rawlings, N.D.; Morton, F.R.; Kok, C.Y.; Kong, J.; Barrett, A.J. MEROPS: The peptidase database. *Nucleic Acids Res.* **2007**, *36*, 320–325. [CrossRef] [PubMed]
4. Ma, W.; Tang, C.; Lai, L. Specificity of Trypsin and Chymotrypsin: Loop-Motion-Controlled Dynamic Correlation as a Determinant. *Biophys. J.* **2005**, *89*, 1183–1193. [CrossRef] [PubMed]
5. Lu, P.; Takai, K.; Weaver, V.M.; Werb, Z. Extracellular Matrix Degradation and Remodeling in Development and Disease. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a005058. [CrossRef]
6. Oncul, S.; Afshar-Kharghan, V. The interaction between the complement system and hemostatic factors. *Curr. Opin. Hematol.* **2020**, *27*, 341–352. [CrossRef]
7. Ferguson, T.E.G.G.; Reihill, J.A.; Martin, S.L.; Walker, B. Novel Inhibitors and Activity-Based Probes Targeting Trypsin-Like Serine Proteases. *Front. Chem.* **2022**, *10*, 782608. [CrossRef]
8. Yaron, J.R.; Zhang, L.; Guo, Q.; Haydel, S.E.; Lucas, A.R. Fibrinolytic Serine Proteases, Therapeutic Serpins and Inflammation: Fire Dancers and Firestorms. *Front. Cardiovasc. Med.* **2021**, *8*, 648947. [CrossRef]
9. Eatemadi, A.; Aiyelabegan, H.T.; Negahdari, B.; Mazlomi, M.A.; Daraee, H.; Daraee, N.; Eatemadi, R.; Sadroddiny, E. Role of protease and protease inhibitors in cancer pathogenesis and treatment. *Biomed. Pharmacother.* **2017**, *86*, 221–231. [CrossRef]
10. Wu, Q.; Kuo, H.-C.; Deng, G.G. Serine proteases and cardiac function. *Biochim. Biophys. Acta-Proteins Proteom.* **2005**, *1751*, 82–94. [CrossRef]
11. Nitsche, C.; Holloway, S.; Schirmeister, T.; Klein, C.D. Biochemistry and Medicinal Chemistry of the Dengue Virus Protease. *Chem. Rev.* **2014**, *114*, 11348–11381. [CrossRef] [PubMed]
12. Dreymann, N.; Wuensche, J.; Sabrowski, W.; Moeller, A.; Czepluch, D.; Van, D.V.; Fuessel, S.; Menger, M.M. Inhibition of Human Urokinase-Type Plasminogen Activator (uPA) Enzyme Activity and Receptor Binding by DNA Aptamers as Potential Therapeutics through Binding to the Different Forms of uPA. *Int. J. Mol. Sci.* **2022**, *23*, 4890–4912. [CrossRef] [PubMed]
13. Mahmood, N.; Mihalcioiu, C.; Rabbani, S.A. Multifaceted Role of the Urokinase-Type Plasminogen Activator (uPA) and Its Receptor (uPAR): Diagnostic, Prognostic, and Therapeutic Applications. *Front. Oncol.* **2018**, *8*, 24. [CrossRef] [PubMed]
14. Mekkawy, A.H.; Pourgholami, M.H.; Morris, D.L. Involvement of Urokinase-Type Plasminogen Activator System in Cancer: An Overview. *Med. Res. Rev.* **2014**, *34*, 918–956. [CrossRef] [PubMed]
15. Masucci, M.T.; Minopoli, M.; Di Carluccio, G.; Motti, M.L.; Carriero, M.V. Therapeutic Strategies Targeting Urokinase and Its Receptor in Cancer. *Cancers* **2022**, *14*, 498. [CrossRef] [PubMed]
16. Buckley, B.J.; Aboelela, A.; Minaei, E.; Jiang, L.X.; Xu, Z.; Ali, U.; Fildes, K.; Cheung, C.-Y.; Cook, S.M.; Johnson, D.C.; et al. 6-Substituted Hexamethylene Amiloride (HMA) Derivatives as Potent and Selective Inhibitors of the Human Urokinase Plasminogen Activator for Use in Cancer. *J. Med. Chem.* **2018**, *61*, 8299–8320. [CrossRef] [PubMed]
17. Ma, D.; Gerard, R.D.; Li, X.-Y.; Alizadeh, H.; Niederkorn, J.Y. Inhibition of Metastasis of Intraocular Melanomas by Adenovirus-Mediated Gene Transfer of Plasminogen Activator Inhibitor Type 1 (PAI-1) in an Athymic Mouse Model. *Blood* **1997**, *90*, 2738–2746. [CrossRef] [PubMed]
18. Ossowski, L.; Russo-Payne, H.; Wilson, L.E. Inhibition of Urokinase-type Plasminogen Activator by Antibodies: The Effect on Dissemination of a Human Tumor in the Nude Mouse. *Cancer Res.* **1991**, *51*, 274–281.
19. Schmitt, M.; Harbeck, N.; Brünner, N.; Jänicke, F.; Meisner, C.; Mühlenweg, B.; Jansen, H.; Dorn, J.; Nitz, U.; Kantelhardt, E.J.; et al. Cancer therapy trials employing level-of-evidence-1 disease forecast cancer biomarkers uPA and its inhibitor PAI-1. *Expert Rev. Mol. Diagn.* **2011**, *11*, 617–634. [CrossRef]
20. Leow, M.K.-S. Correlating Cell Line Studies With Tissue Distribution of DPP4/TMPRSS2 and Human Biological Samples May Better Define the Viral Tropism of MERS-CoV. *J. Infect. Dis.* **2013**, *208*, 1350–1351. [CrossRef]

21. Bertram, S.; Heurich, A.; Lavender, H.; Gierer, S.; Danisch, S.; Perin, P.; Lucas, J.M.; Nelson, P.S.; Pöhlmann, S.; Soilleux, E.J. Influenza and SARS-Coronavirus Activating Proteases TMPRSS2 and HAT Are Expressed at Multiple Sites in Human Respiratory and Gastrointestinal Tracts. *PLoS ONE* **2012**, *7*, e35876. [CrossRef] [PubMed]
22. Simmons, G.; Zmora, P.; Gierer, S.; Heurich, A.; Pöhlmann, S. Proteolytic activation of the SARS-coronavirus spike protein: Cutting enzymes at the cutting edge of antiviral research. *Antivir. Res.* **2013**, *100*, 605–614. [CrossRef] [PubMed]
23. Mahoney, M.; Damalanka, V.C.; Tartell, M.A.; Chung, D.H.; Lourenço, A.L.; Pwee, D.; Mayer Bridwell, A.E.; Hoffmann, M.; Voss, J.; Karmakar, P.; et al. A novel class of TMPRSS2 inhibitors potently block SARS-CoV-2 and MERS-CoV viral entry and protect human epithelial lung cells. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2108728118. [CrossRef] [PubMed]
24. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T.S.; Herrler, G.; Wu, N.H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271–280.e8. [CrossRef]
25. Jackson, C.B.; Farzan, M.; Chen, B.; Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 3–20. [CrossRef]
26. Zhao, M.M.; Yang, W.L.; Yang, F.Y.; Zhang, L.; Huang, W.J.; Hou, W.; Fan, C.F.; Jin, R.H.; Feng, Y.M.; Wang, Y.C.; et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduct. Target. Ther.* **2021**, *6*, 134. [CrossRef]
27. Bestle, D.; Heindl, M.R.; Limburg, H.; Pilgram, O.; Moulton, H.; Stein, D.A.; Hardes, K.; Eickmann, M.; Dolnik, O.; Rohde, C.; et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* **2020**, *3*, e1–e14. [CrossRef]
28. Li, F.; Han, M.; Dai, P.; Xu, W.; He, J.; Tao, X.; Wu, Y.; Tong, X.; Xia, X.; Guo, W.; et al. Distinct mechanisms for TMPRSS2 expression explain organ-specific inhibition of SARS-CoV-2 infection by enzalutamide. *Nat. Commun.* **2021**, *12*, 866. [CrossRef]
29. Wettstein, L.; Knaff, P.M.; Kersten, C.; Müller, P.; Weil, T.; Conzelmann, C.; Müller, J.A.; Brückner, M.; Hoffmann, M.; Pöhlmann, S.; et al. Peptidomimetic inhibitors of TMPRSS2 block SARS-CoV-2 infection in cell culture. *Commun. Biol.* **2022**, *5*, 681. [CrossRef]
30. Sanders, B.C.; Pokhrel, S.; Labbe, A.D.; Mathews, I.I.; Cooper, C.J.; Davidson, R.B.; Phillips, G.; Weiss, K.L.; Zhang, Q.; O'Neill, H.; et al. Potent and selective covalent inhibition of the papain-like protease from SARS-CoV-2. *Nat. Commun.* **2023**, *14*, 1733. [CrossRef]
31. Welker, A.; Kersten, C.; Müller, C.; Madhugiri, R.; Zimmer, C.; Müller, P.; Zimmermann, R.; Hammerschmidt, S.; Maus, H.; Ziebuhr, J.; et al. Structure-Activity Relationships of Benzamides and Isoindolines Designed as SARS-CoV Protease Inhibitors Effective against SARS-CoV-2. *ChemMedChem* **2021**, *16*, 340–354. [CrossRef]
32. Kincaid, J.R.; Caravez, J.C.; Iyer, K.S.; Kavthe, R.D.; Fleck, N.; Aue, D.H.; Lipshutz, B.H. A sustainable synthesis of the SARS-CoV-2 Mpro inhibitor nirmatrelvir, the active ingredient in Paxlovid. *Commun. Chem.* **2022**, *5*, 156. [CrossRef]
33. Joossens, J.; Ali, O.M.; El-Sayed, I.; Surpateanu, G.; Van der Veken, P.; Lambeir, A.-M.; Setyono-Han, B.; Foekens, J.A.; Schneider, A.; Schmalix, W.; et al. Small, Potent, and Selective Diaryl Phosphonate Inhibitors for Urokinase-Type Plasminogen Activator with In Vivo Antimetastatic Properties. *J. Med. Chem.* **2007**, *50*, 6638–6646. [CrossRef]
34. Zeslawska, E.; Jacob, U.; Schweinitz, A.; Coombs, G.; Bode, W.; Madison, E. Crystals of Urokinase Type Plasminogen Activator Complexes Reveal the Binding Mode of Peptidomimetic Inhibitors. *J. Mol. Biol.* **2003**, *328*, 109–118. [CrossRef]
35. Lonsdale, R.; Ward, R.A. Structure-based design of targeted covalent inhibitors. *Chem. Soc. Rev.* **2018**, *47*, 3816–3830. [CrossRef]
36. Baillie, T.A. Targeted Covalent Inhibitors for Drug Design. *Angew. Chem. Int. Ed.* **2016**, *55*, 13408–13421. [CrossRef]
37. De Cesco, S.; Kurian, J.; Dufresne, C.; Mittermaier, A.K.; Moitessier, N. Covalent inhibitors design and discovery. *Eur. J. Med. Chem.* **2017**, *138*, 96–114. [CrossRef]
38. Sun, G.; Sui, Y.; Zhou, Y.; Ya, J.; Yuan, C.; Jiang, L.; Huang, M. Structural Basis of Covalent Inhibitory Mechanism of TMPRSS2-Related Serine Proteases by Camostat. *J. Virol.* **2021**, *95*, 10–1128. [CrossRef]
39. Müller, P.; Meta, M.; Meidner, J.L.; Schwickert, M.; Meyr, J.; Schwickert, K.; Kersten, C.; Zimmer, C.; Hammerschmidt, S.J.; Frey, A.; et al. Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study. *Int. J. Mol. Sci.* **2023**, *24*, 7226. [CrossRef]
40. Kwon, H.; Ha, H.; Jeon, H.; Jang, J.; Son, S.-H.; Lee, K.; Park, S.-K.; Byun, Y. Structure-activity relationship studies of dipeptide-based hepsin inhibitors with Arg bioisosteres. *Bioorg. Chem.* **2021**, *107*, 104521. [CrossRef]
41. Hoffmann, M.; Hofmann-Winkler, H.; Smith, J.C.; Krüger, N.; Arora, P.; Sørensen, L.K.; Søgaard, O.S.; Hasselstrøm, J.B.; Winkler, M.; Hempel, T.; et al. Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *EBioMedicine* **2021**, *65*, 103255. [CrossRef] [PubMed]
42. Han, Z.; Harris, P.K.; Karmakar, P.; Kim, T.; Owusu, B.Y.; Wildman, S.A.; Klampfer, L.; Janetka, J.W. α-Ketobenzothiazole Serine Protease Inhibitors of Aberrant HGF/c-MET and MSP/RON Kinase Pathway Signaling in Cancer. *ChemMedChem* **2016**, *11*, 585–599. [CrossRef] [PubMed]
43. 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. [CrossRef] [PubMed]
44. Li, C.Y.; de Veer, S.J.; Law, R.H.P.; Whisstock, J.C.; Craik, D.J.; Swedberg, J.E. Characterising the Subsite Specificity of Urokinase-Type Plasminogen Activator and Tissue-Type Plasminogen Activator using a Sequence-Defined Peptide Aldehyde Library. *ChemBioChem* **2019**, *20*, 46–50. [CrossRef] [PubMed]

45. Green, K.A.; Lund, L.R. ECM degrading proteases and tissue remodelling in the mammary gland. *BioEssays* **2005**, *27*, 894–903. [CrossRef] [PubMed]
46. Ortiz-Zapater, E.; Peiró, S.; Roda, O.; Corominas, J.M.; Aguilar, S.; Ampurdanés, C.; Real, F.X.; Navarro, P. Tissue Plasminogen Activator Induces Pancreatic Cancer Cell Proliferation by a Non-Catalytic Mechanism That Requires Extracellular Signal-Regulated Kinase 1/2 Activation through Epidermal Growth Factor Receptor and Annexin A2. *Am. J. Pathol.* **2007**, *170*, 1573–1584. [CrossRef] [PubMed]
47. Tanaka, K.A.; Key, N.S.; Levy, J.H. Blood Coagulation: Hemostasis and Thrombin Regulation. *Anesth. Analg.* **2009**, *108*, 1433–1446. [CrossRef]
48. Borensztajn, K.; Spek, C.A. Blood coagulation factor Xa as an emerging drug target. *Expert Opin. Ther. Targets* **2011**, *15*, 341–349. [CrossRef]
49. List, K.; Bugge, T.H.; Szabo, R. Matriptase: Potent Proteolysis on the Cell Surface. *Mol. Med.* **2006**, *12*, 1–7. [CrossRef]
50. Gustafsson, D.; Nyström, J.E.; Carlsson, S.; Bredberg, U.; Eriksson, U.; Gyzander, E.; Elg, M.; Antonsson, T.; Hoffmann, K.J.; Ungell, A.L.; et al. The Direct Thrombin Inhibitor Melagatran and Its Oral Prodrug H 376/95: Intestinal Absorption Properties, Biochemical and Pharmacodynamic Effects. *Thromb. Res.* **2001**, *101*, 171–181. [CrossRef]
51. Sibinovska, N.; Žakelj, S.; Trontelj, J.; Kristan, K. Applicability of RPMI 2650 and Calu-3 Cell Models for Evaluation of Nasal Formulations. *Pharmaceutics* **2022**, *14*, 369. [CrossRef]
52. Elezović, A.; Marić, A.; Biščević, A.; Hadžiabdić, J.; Škrbo, S.; Špirtović-Halilović, S.; Rahić, O.; Vranić, E.; Elezović, A. In vitro pH dependent passive transport of ketoprofen and metformin. *ADMET DMPK* **2020**, *9*, 57–68. [CrossRef]
53. Butnarasu, C.; Caron, G.; Pacheco, D.P.; Petrini, P.; Visentin, S. Cystic Fibrosis Mucus Model to Design More Efficient Drug Therapies. *Mol. Pharm.* **2022**, *19*, 520–531. [CrossRef]
54. Hauel, N.H.; Nar, H.; Priepke, H.; Ries, U.; Stassen, J.-M.; Wienen, W. Structure-Based Design of Novel Potent Nonpeptide Thrombin Inhibitors. *J. Med. Chem.* **2002**, *45*, 1757–1766. [CrossRef]
55. Schade, D.; Kotthaus, J.; Riebling, L.; Kotthaus, J.; Müller-Fielitz, H.; Raasch, W.; Hoffmann, A.; Schmidtke, M.; Clement, B. Zanamivir Amidoxime- and N-Hydroxyguanidine-Based Prodrug Approaches to Tackle Poor Oral Bioavailability. *J. Pharm. Sci.* **2015**, *104*, 3208–3219. [CrossRef]
56. Dave, V.S.; Gupta, D.; Yu, M.; Nguyen, P.; Varghese Gupta, S. Current and evolving approaches for improving the oral permeability of BCS Class III or analogous molecules. *Drug Dev. Ind. Pharm.* **2017**, *43*, 177–189. [CrossRef]
57. Fuchs, N.; Meta, M.; Schuppan, D.; Nuhn, L.; Schirmeister, T. Novel Opportunities for Cathepsin S Inhibitors in Cancer Immunotherapy by Nanocarrier-Mediated Delivery. *Cells* **2020**, *9*, 2021. [CrossRef]
58. Lee, J.; Choi, M.-K.; Song, I.-S. Recent Advances in Doxorubicin Formulation to Enhance Pharmacokinetics and Tumor Targeting. *Pharmaceuticals* **2023**, *16*, 802. [CrossRef]
59. Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41*, 1007–1010. [CrossRef]
60. St-Georges, C.; Désilets, A.; Béliveau, F.; Ghinet, M.; Dion, S.P.; Colombo, É.; Boudreault, P.-L.; Najmanovich, R.J.; Leduc, R.; Marsault, É. Modulating the selectivity of matriptase-2 inhibitors with unnatural amino acids. *Eur. J. Med. Chem.* **2017**, *129*, 110–123. [CrossRef]
61. Costanzo, M.J.; Yabut, S.C.; Almond, H.R.; Andrade-Gordon, P.; Corcoran, T.W.; de Garavilla, L.; Kauffman, J.A.; Abraham, W.M.; Recacha, R.; Chattopadhyay, D.; et al. Potent, Small-Molecule Inhibitors of Human Mast Cell Tryptase. Antiasthmatic Action of a Dipeptide-Based Transition-State Analogue Containing a Benzothiazole Ketone. *J. Med. Chem.* **2003**, *46*, 3865–3876. [CrossRef]
62. Furlan, A.; Colombo, F.; Kover, A.; Issaly, N.; Tintori, C.; Angeli, L.; Leroux, V.; Letard, S.; Amat, M.; Asses, Y.; et al. Identification of new aminoacid amides containing the imidazo[2,1-b]benzothiazol-2-ylphenyl moiety as inhibitors of tumorigenesis by oncogenic Met signaling. *Eur. J. Med. Chem.* **2012**, *47*, 239–254. [CrossRef]
63. Capaldo, L.; Quadri, L.L.; Merli, D.; Ravelli, D. Photoelectrochemical cross-dehydrogenative coupling of benzothiazoles with strong aliphatic C–H bonds. *Chem. Commun.* **2021**, *57*, 4424–4427. [CrossRef]
64. Kerns, E.H.; Di, L.; Petusky, S.; Farris, M.; Ley, R.; Jupp, P. Combined Application of Parallel Artificial Membrane Permeability Assay and Caco-2 Permeability Assays in Drug Discovery. *J. Pharm. Sci.* **2004**, *93*, 1440–1453. [CrossRef]
65. Avdeef, A. *Absorption and Drug Development*; Wiley: Hoboken, NJ, USA, 2003. [CrossRef]
66. Sugano, K.; Hamada, H.; Machida, M.; Ushio, H. High Throughput Prediction of Oral Absorption: Improvement of the Composition of the Lipid Solution Used in Parallel Artificial Membrane Permeation Assay. *SLAS Discov.* **2001**, *6*, 189–196. [CrossRef]
67. Hammerschmidt, S.J.; Maus, H.; Weldert, A.C.; Gütschow, M.; Kersten, C. Improving binding entropy by higher ligand symmetry?—A case study with human matriptase. *RSC Med. Chem.* **2023**, *14*, 969–982. [CrossRef]
68. Ehnebom, J.; Pusa, S.; Björquist, P.; Deinum, J. Comparison of chromogenic substrates for tissue plasminogen activator and the effects on the stability of plasminogen activator inhibitor type-1. *Fibrinolysis Proteolysis* **1997**, *11*, 287–293. [CrossRef]
69. Tapp, H.J.; Grundmann, C.; Kusch, M.; Konig, H. Calibrating Thrombin Generation in Different Samples: Less Effort with a Less Efficient Substrate. *Open Atheroscler. Thromb. J.* **2009**, *2*, 6–11. [CrossRef]
70. Edwards, S.T.; Betz, A.; James, H.L.; Thompson, E.; Yonkovich, S.J.; Sinha, U. Differences between human and rabbit coagulation factor X—Implications for in vivo models of thrombosis. *Thromb. Res.* **2002**, *106*, 71–79. [CrossRef]

71. Wilkinson, D.J.; Habgood, A.; Lamb, H.K.; Thompson, P.; Hawkins, A.R.; Désilets, A.; Leduc, R.; Steinmetzer, T.; Hammami, M.; Lee, M.S.; et al. Matriptase Induction of Metalloproteinase-Dependent Aggrecanolysis In Vitro and In Vivo: Promotion of Osteoarthritic Cartilage Damage by Multiple Mechanisms. *Arthritis Rheumatol.* **2017**, *69*, 1601–1611. [CrossRef]
72. Steinmetzer, T.; Schweinitz, A.; Stürzebecher, A.; Dönnecke, D.; Uhland, K.; Schuster, O.; Steinmetzer, P.; Müller, F.; Friedrich, R.; Than, M.E.; et al. Secondary Amides of Sulfonylated 3-Amidinophenylalanine. New Potent and Selective Inhibitors of Matriptase. *J. Med. Chem.* **2006**, *49*, 4116–4126. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Supporting information

Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres

Patrick Müller, Collin Zimmer, Ariane Frey, Gideon Holzmann, Annabelle Carolin Weldert and Tanja Schirmeister*

Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany; muelpat@uni-mainz.de (P.M.); cozimmer@uni-mainz.de (C.Z.); arfrey@uni-mainz.de (A.F.); gholzman@students.uni-mainz.de (G.H.); anwelder@uni-mainz.de (A.C.W.)

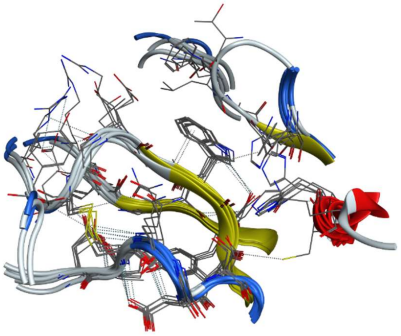
* Correspondence: schirmei@uni-mainz.de; Tel.: +49-6131-39-25742

Content

Protein Similarity Calculations

The similarity of the investigated serine proteases was compared using MOE 2022.02. Their biological assemblies were loaded using their pdb accession codes (uPA: 7VM4, TMPRSS2: 7MEQ, tPA: 1BDA, Factor Xa: 1XKB, Thrombin: 1D6W, Matriptase 1: 2GV6), aligned, and superimposed. For similarity calculations, only the binding sites were selected.

binding site alignment



similarity

	1	2	3	4	5	6
1:uPA		55.6	66.7	62.2	60.0	68.9
2:TMPRSS2	53.2		62.2	66.7	57.8	66.7
3:tPA	63.8	62.2		66.7	57.8	75.6
4:Factor Xa	59.6	66.7	66.7		68.9	68.9
5:Thrombin	57.4	57.8	57.8	68.9		64.4
6:Matriptase 1	66.0	66.7	75.6	68.9	64.4	

Figure S1: (left) visual binding site alignment of all six serine proteases, (right) matrix for calculated sequence similarities of the binding sites.

Fluorometric inhibition assays

The inhibitory activities of the compounds against the proteases were determined by enzyme inhibition assays with fluorogenic or colorimetry-based substrates. The fluorescence, caused by cleaved AMC from the fluorogenic substrates (uPA, matriptase, TMPRSS2, thrombin), was measured in white flat-bottom 96-well plates from Greiner Bio-One using a Tecan Infinite F200 Pro plate reader. The absorption, caused by cleaved *p*NA from colorimetric substrates (factor Xa and tPA), was measured in transparent flat-bottom 96-well plates from Greiner Bio-One using a Tecan Spark 10M plate reader. All measurements with the main-target proteases (uPA and TMRPSS2) were performed as triplicates and with the off-target proteases (matriptase, factor Xa, tPA, thrombin) as duplicates. The substrates and the compounds were prepared as stock solutions in DMSO. Each well contained a total volume of 200 µL, consisting of 185 µL buffer, 5 µL inhibitor in DMSO or pure DMSO as negative control, 5 µL substrate in DMSO and 5 µL enzyme solution in buffer. Dilution series were prepared for the determination of the inhibition constants. The fluorescence signal and absorbance were measured every 30 s for 10 min at 25 or 37 °C with the corresponding excitation/emission (λ_{ex} = 380 nm / λ_{em} = 460 nm) and absorbance (λ_{abs} = 405 nm) wavelengths, depending on the targeted protease. IC_{50} values for the reversible inhibitors were calculated with Graphpad Prism 9 by fitting the remaining enzymatic activity to the four parameter IC_{50} equation with Y [%] as the residual enzyme activity , Y_{max} as the maximum value of the dose response curve at inhibitor concentrations [I] = 0 µM, Y_{min} as the minimum value at high inhibitor concentrations and *s* as the hill coefficient.[1] The fluorescence progress curves and sigmoidal dose-response curve for inhibition of uPa by compound **38** is shown exemplary in Figure **S2**.

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

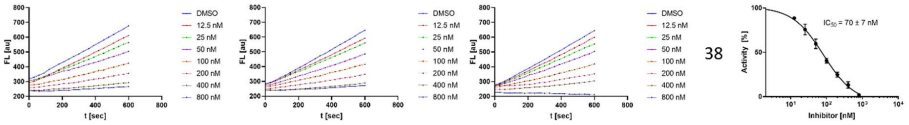


Figure S2: (left) Fluorescence progress curves for inhibition of uPa by cpd. **38**. (right) Plot showing the respective IC_{50} value from sigmoidal fit.

Due to the dependence of the IC_{50} value on the substrate affinity and concentration, the K_i values were calculated with the Cheng-Prusoff equation, using the final substrate concentration [*S*] and the Michaelis-Menten constant K_M, (2) for appropriate comparison of the inhibitory activities to the other enzymes and inhibitors.[1]

$$K_{\mathrm{i}} = \frac{\mathrm{IC}_{50}}{1 + \frac{[S]}{K_{\mathrm{M}}}} \qquad 2$$

K_M values were determined by fitting to the Michaelis-Menten equation using Graphpad Prism 9 with *v* [Δ*F*/min] as the substrate hydrolysis rate, v_{max} as the maximum slope of the dose-response curve, and the substrate [*S*] concentration.[1] The different substrates were serially diluted in DMSO. The measurement was done in analogy to the determination of the inhibitory activity, but with 5 µL pure DMSO instead of inhibitor solution. The fluorescence progress curves and Michaelis-Menten curve of the specific substrate of uPA is shown exemplarily in Figure **S3**.

$$v = \frac{v_{\max} \cdot [\mathrm{S}]}{K_{\mathrm{M}} + [\mathrm{S}]} \qquad 3$$

3

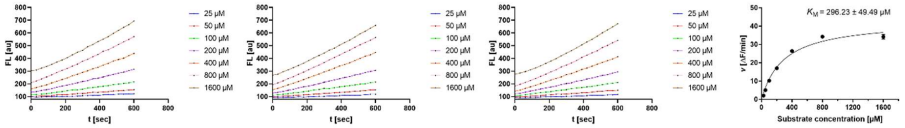


Figure S3: (left) Fluorescence progress curves of cleavage of Z-Gly-Gly-Arg-AMC by uPA. (right) Michaelis-Menten curve showing the respective K_M value.

Buffers and substrates

The following buffers and substrates were used for the respective assays: **uPA** (50 mM Tris HCl pH = 7.9, 150 mM NaCl, 10 mM $CaCl_2$, 0.005% TX-100, 240 µM Z-Gly-Gly-Arg-AMC (K_M: 296.23 ± 49.49 µM), 37 °C, 9 U/mL uPA)[2]; **TMPRSS2** (25 mM Tris HCl pH = 8.0, 150 mM NaCl, 5 mM $CaCl_2$, 0.001% TX-100, 100 µM Boc-Gln-Ala-Arg-AMC (K_M: 68.63 ± 6.64 µM), 25 °C, 3 nM TMPRSS2)[3]; **matriptase** (50 mM Tris HCl pH = 8.0, 150 mM NaCl, 5 mM $CaCl_2$, 0.0001% TX-100, 100 µM Boc-Leu-Arg-Arg-AMC (K_M: 36.13 ± 5.78 µM, 25 °C, 6 nM matriptase)[4]; **tPA** (50 mM Tris HCl pH = 8.3, 250 µM *N*-methylsulfonyl-*D*-Phe-Gly-Arg-*p*NA (K_M: 148.89 ± 29.72 µM), 37 °C, 0.4 ng/µL tPA)[5]; **thrombin** (50 mM Tris HCl pH = 8.0, 100 mM NaCl, 5 mM $CaCl_2$, 0.01% Tween-20, 200 µM Z-Gly-Gly-Arg-AMC (K_M: 49.70 ± 7.15 µM), 25 °C, 10 nM thrombin)[6]; **factor Xa** (50 mM Tris HCl pH = 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0.005% Brij-35, 100 µM Z-*D*-Arg-Gly-Arg-*p*NA (K_M: 249.12 ± 21.59 µM), 37 °C, 0.1 ng/µL factor Xa)[7].

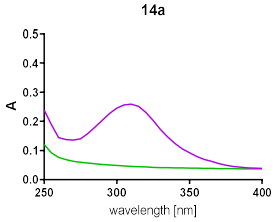
Human uPa and tPA was purchased from Sigma Aldrich, TMPRSS2 from Cusabio, factor Xa and thrombin from Bio-techne. Matriptase was expressed under the conditions described below.

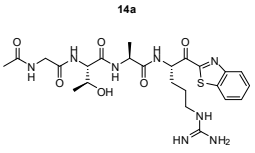
Protein expression and purification of matriptase 1

Recombinant human matriptase 1 was expressed as described previously with slight adaptations.[8,9] In short, the pQE-30 vector containing the zymogen of the catalytic domain of matriptase (uniport: Q9Y5Y6 aa 596–855) with an *N*-terminal hexa-histidine tag, was transformed into competent *Escherichia coli* (*E. coli*) BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA). The transformed bacteria were grown at 37 °C and 160 rpm in LB medium (10 L) containing 100 µg/mL ampicillin. After reaching an optical density (OD_{600}) of ~0.8 overexpression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranosid (IPTG) and the cells were incubated over night at 20 °C. Cells were harvested by centrifugation (9000 rpm at 4 °C for 15 min), resuspended in 250 mL cold lysis buffer (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)) and lysed by sonication (Sonoplus HD 2200; Bandelin, Berlin, Germany). Inclusion bodies were isolated by centrifugation (20000 rpm at 4 °C for 1 h) and resuspended in 500 mL solubilization buffer (50 mM TRIS·HCl pH 8.0, 5% (v/v) glycerol, 6 M urea, 20 mM imidazole). After stirring the suspension overnight at 4 °C the mixture was centrifuged (20000 rpm at 4 °C for 1 h) again to remove insoluble impurities. Matriptase 1 was purified by immobilized metal affinity chromatography (IMAC) using a HisTrap HP 5 ml column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany). The protein was washed with 5 column volumes (CV) wash buffer (50 mM TRIS·HCl pH 8.0, 6 M urea, 20 mM imidazole), and elution was achieved with a linear gradient of elution buffer (50 mM TRIS·HCl pH 8.0, 6 M urea, 200 mM imidazole). Matriptase 1 containing fractions were pooled and refolded by two-step dialysis against 2 L of dialysis buffer 1 (50 mM TRIS·HCl pH 9.0, 1 mM *β*-ME, 3 M urea) and 2 L of dialysis buffer 2 (50 mM TRIS·HCl pH 9.0, 1 mM β-ME) at 4 °C for > 8 h each. For further purification, anion exchange chromatography (AEX) on a HiTrap 5 mL column (Cytiva Europe GmbH) was utilized. Dialysis buffer 2 was used as AEX wash buffer (5 CV) and Matriptase was eluted with a linear gradient of dialysis buffer 2 and AEX elution buffer (50 mM TRIS·HCl pH 9.0, 1 mM β-ME, 1 M NaCl). Fractions containing matriptase, were pooled, flash-frozen in liquid nitrogen and stored at −80°C.

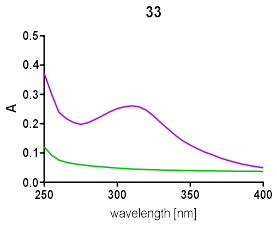
Absorption spectra.

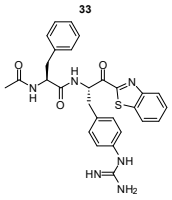
Method: Measurement setup for UV spectroscopy consisted of UV-transparent measurement plate (Greiner UV-Star®, 655801), Tecan Spark 10M® well plate reader, 200 µL sample volume (5% DMSO/Tris buffer pH 7.4), λ = 250–400 nm.



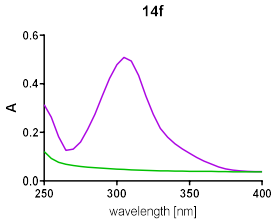


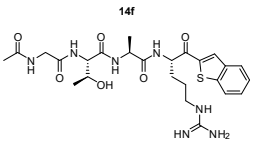
λ_{max} = 310 nm



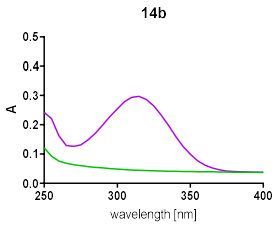


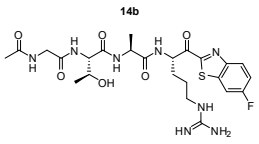
λ_{max} = 310 nm





λ_{max} = 305 nm





λ_{max} = 315 nm

5

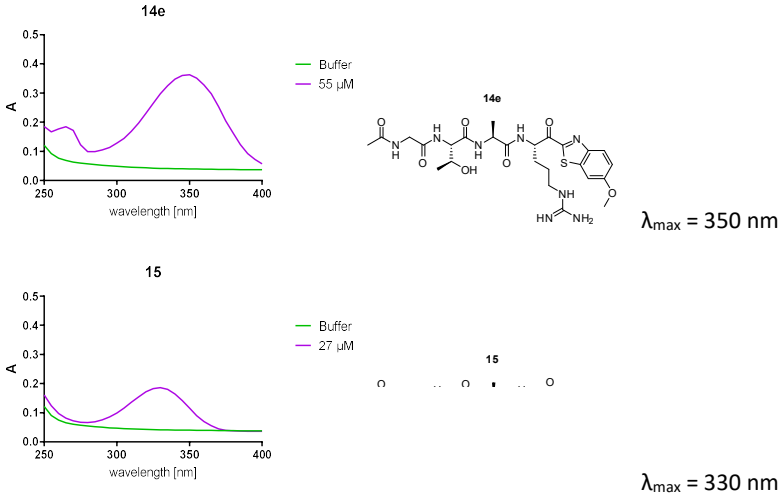


Figure S4: Absorption spectra of a representative set of inhibitors with regards to the warhead λ_{max}. Recorded from 55 µM (27 µM for **15**) solutions in 5% DMSO/Tris buffer pH 7.4.

The absorption spectra show the expected absorption maximum > 250 nm that corresponds to the annealed heteroarenes. The range ≤ 250 nm was not evaluated due to the presence of DMSO, but the absorption maxima of the isolated aromatic systems (Phe, aryl-guanidines) are expected to be detected there.

The spectra change with identity or substitution pattern of the annealed system: benzothiophene has a lower-wavelength λ_{max} = 305 nm compared to benzothiazole (λ_{max} = 310 nm), while cyclohexyl thiazole shows λ_{max} = 330 nm. Halogen substitution of the benzothiazole has a weak bathochromic effect [H (λ_{max} = 310 nm) < F (λ_{max} = 315 nm) < Cl = Br (λ_{max} = 320 nm)], while methoxy substitution has a distinct bathochromic effect (λ_{max} = 350 nm).

Computation of physicochemical parameters.

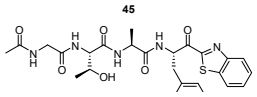
Method: pK_a was calculated using MarvinJS 23.11.0 (https://playground.calculators.cxn.io/). Only the pK_a of the protonated guanidine moiety is reported, since it is the only relevant pH-responsive group in an aqueous medium. logP was calculated using molinspiration v2022.08 (https://www.molinspiration.com/cgi/properties), $logD_{7.4}$ was computed using MarvinJS 23.2.0 (https://plugins.calculators.cxn.io/logd/).

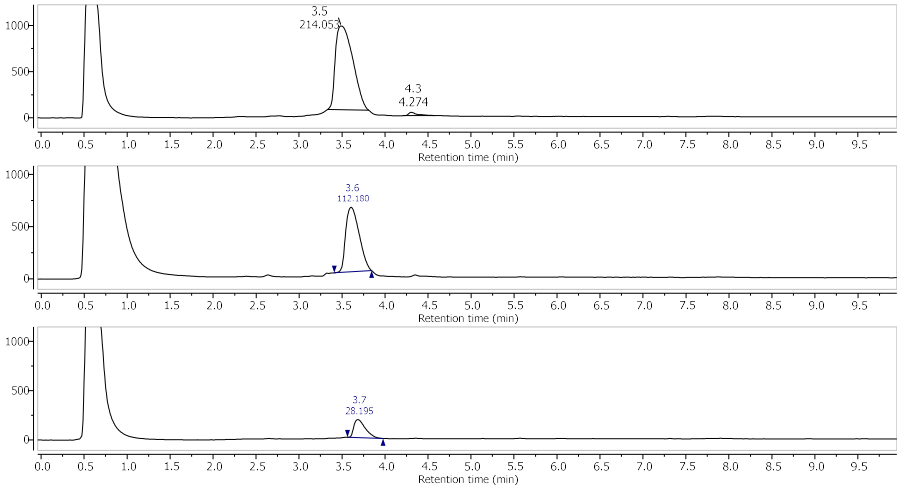
Table S1: List of compounds evaluated for some computed physicochemical properties (pK_A, logP, $logD_{7.4}$) sorted from most hydrophilic to most hydrophobic of the set.

compound	cpK_a (guanidine)	clogP	$clogD_{7.4}$
MTX	-	−2.0	−5.7
15	11.3	−2.2	−4.5
14a	11.7	−2.0	−4.5
14e	11.3	−1.9	−4.5
14b	11.2	−1.8	−4.5
20	11.7	−1.5	−4.2
14f	11.8	−0.9	−4.2
14c	11.2	−1.3	−4.0
14d	11.7	−1.2	−3.8
38	11.8	−0.8	−3.4
45	10.0	−0.7	−3.1
29	10.0	−0.6	−3.1
39	12.1	0.4	−2.9
30	10.0	0.6	−2.3
31	10.0	1.7	−0.9
32	10.0	2.7	−0.2
33	10.0	2.4	1.2
PRO	-	3.0	0.8

Stability studies.

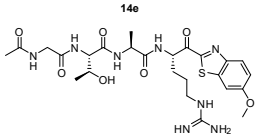
Method: 10–20 mM stock solutions of inhibitors in DMSO were diluted into ACN/water or solutions in ACN/water were prepared directly from solid and then analyzed by LC/MS (“stock”). Alternatively, stock solutions were diluted into DMSO/buffer for PAMPA, incubated either for 7 h at room temperature or 17 h at 37 °C and then analyzed by LC/MS (“room temperature” or “37 °C”). Samples at room temperature were concentrated by lyophilization to ca. 10x the original concentration prior to analysis. The LC/MS method is the same as described in the PAMPA section as “measurement setup for LC/MS”. The chromatograms depicted in the results section are always in the following order: “stock”, then “room temperature”, then “37 °C”.

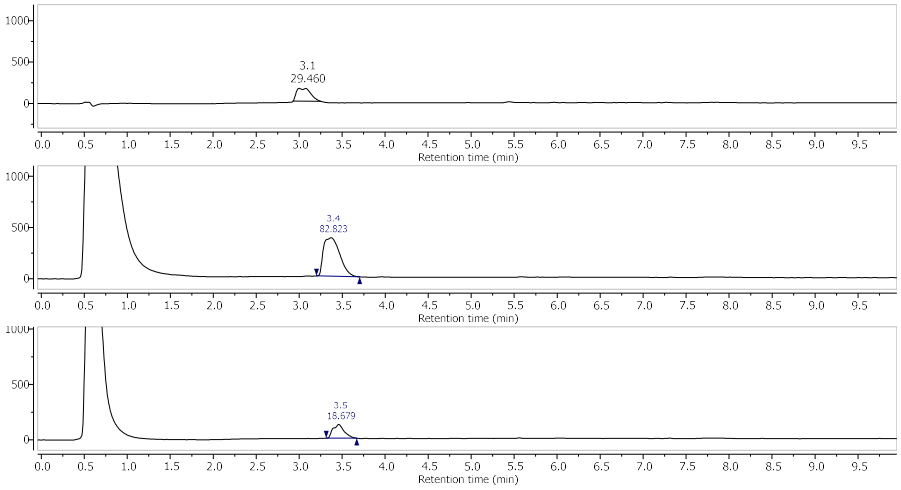




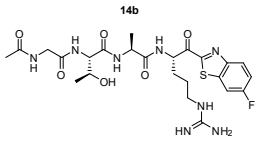
8

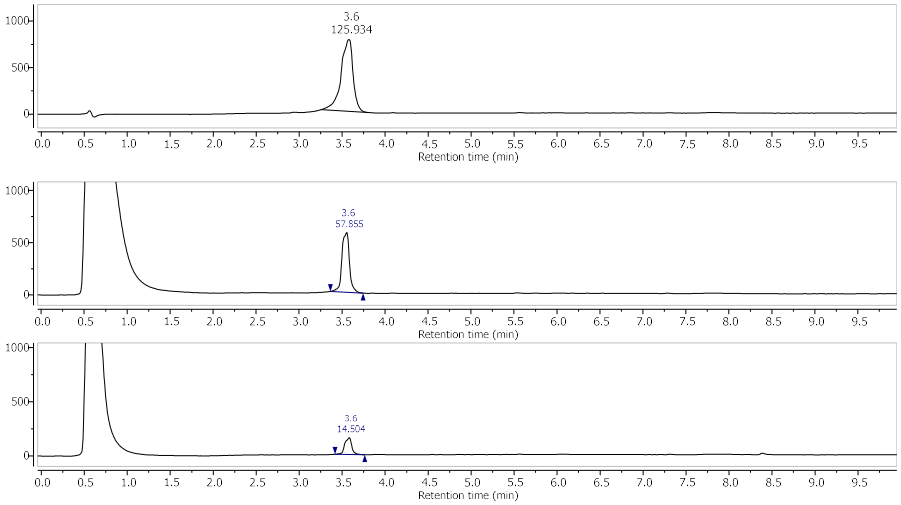
14e



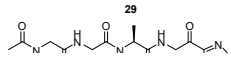

3.1
29.460
3.4
82.823
3.5
18.679
Retention time (min)

14b




3.6
125.934
3.6
57.855
3.6
14.504
Retention time (min)

9



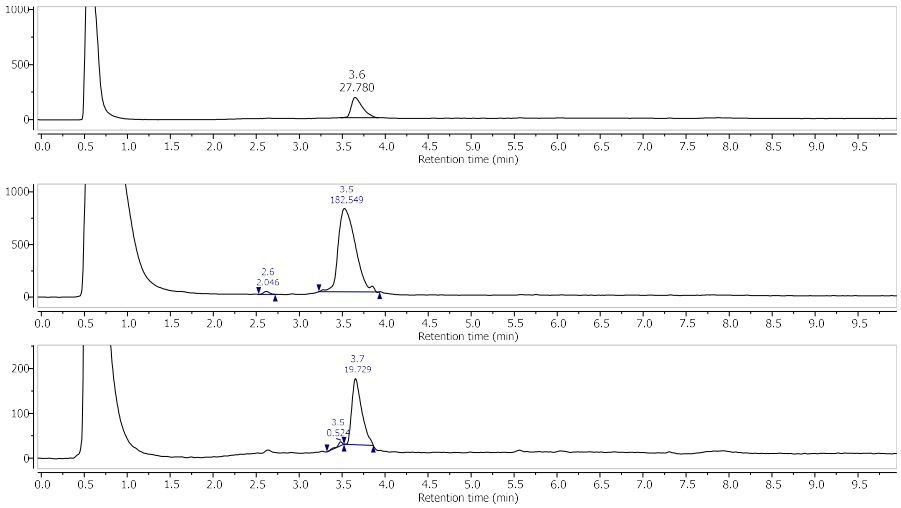


Figure S5: Stability analyses of a representative set of inhibitors. The depictions consist of compound structure and number, then analysis at t_0, then analysis after 7 h at room temperature, then analysis after 12 h at 37 °C.

Parallel Artificial Membrane Permeation Assay

The general principle reported in ref. [10] was used.

Compound solutions were diluted from 10–20 mM stock solutions in DMSO to 100–400 µM in a buffered (TRIS 50 mM pH 7.4) aqueous solution with 5% final content of DMSO ("donor solution"). A similarly prepared solution (buffer + 5% DMSO) was used as "acceptor solutions". Incubation setup consisted of donor (top) plate (Sigma Aldrich, MAIPNTR10), 5 µL artificial membrane (1 % (w/v) L-α-phosphatidylcholine, Sigma Aldrich P3556, in *n*-dodecane, Sigma Aldrich 8205430100), acceptor (bottom) plate (Greiner, 655074). 150 µL of donor solution was applied onto the artificial membrane which was applied first to the donor plate. This upper compartment was sealed (Greiner, 676070, Viewseal sealer). 400 µL of acceptor solution was applied to the acceptor plate. Incubation setup was assembled and left for 7 h at room temperature. Reference solutions were prepared by simply mixing the indicated volumes of donor and acceptor solutions at the start of the incubation period and analyzed later with the acceptor solutions. The experiment was performed in duplicates.

Measurement setup for spectroscopy consisted of UV-transparent measurement plate (Greiner UV-Star®, 655801), Tecan Spark 10M® well plate reader, 200 µL sample volume, λ = 200–650 nm. Analysis of UV data was performed with AUC function in GraphPad Prism using the indicated wavelength range. Acceptor and reference spectra were baseline-corrected to buffer spectrum at a wavelength range without absorption (λ = 450–500 nm).

Measurement setup for LC/MS consisted of Agilent 1100 series HPLC system coupled to an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI). An Agilent Poroshell 120 EC-C18, 150x2.10 mm, 4 µm column was used. A linear gradient was used for elution with a ternary pump using [water/ACN/water + 0.1% formic acid] changing ratios from 80/10/10 to 0/90/10 over the course of 10 min (0.7 mL/min). Injection volume was 100 µL. Areas of peaks in the chromatogram (detection-λ = 210 nm) were used to calculate the AUCs. Retention time and mass spectrum recorded in positive ionization mode were used to assign species. Analysis was performed using Mestrenova 12.0.2.

Calculations of effective permeability P_e were performed using the following equation with V_D and V_A as volumes of donor and acceptor solutions (0.15 cm^3 and 0.4 cm^3, respectively), AUC_{acc} and AUC_{ref} as the area of the measured and baseline-corrected spectrum of acceptor and reference solutions after incubation, A as the porosity-corrected filter area (0.3019 cm^2*0.7 = 0.2113 cm^2) and t as the incubation time given in seconds.

$$P_e = -\frac{V_D * V_A * \ln\left(1-\frac{AUC_{acc}}{AUC_{ref}}\right)}{(V_D+V_A)*A*t} \qquad 4$$

Lit. values for propranolol (Lit-P_e(propranolol) = 8.3–13.9 *10^{-6} cm/s [11,12], measured P_e = 9.0 ± 0.2*10^{-6} cm/s) and methotrexate (Lit-P_e(methotrexate) = 0.0 *10^{-6} cm/s [13], measured P_e = 0*10^{-6} cm/s) were replicated ensuring reliability of data for novel compounds.

Synthesis

General Methods and Materials:

All reagents and solvents were purchased commercially and used as provided by the supplier without further purification. Solvents for synthesis, extraction, and chromatography were of analytical grade. Moisture-sensitive reactions were carried out under argon atmosphere as indicated, and anhydrous solvents were used as provided by the commercial supplier. Reaction progress was monitored by thin-layer chromatography using Alugram Xtra F254 silica plates from Macherey-Nagel and/or LC-MS (Agilent 1100 series HPLC system and an Agilent Poroshell 120 EC-C18, 150 x 2.10 mm, 4 μm column coupled to an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI)). The identities and purities of compounds were determined by the same LC-MS system with a gradient of acetonitrile and water (+0.1% formic acid). Signals were detected at 210/254 nm with quantitation by AUC and masses were determined in positive ionization mode (ESI). HPLC purification was performed with the Agilent 1290 II Infinity Preparative LC System using an InfinityLab Pursuit XRs C18, 30 x 250mm, 5 µm, preparative LC column and a gradient method (10% ACN (0.1% formic acid) to 100% ACN (0.1% formic acid)). Flash chromatography was performed with the Biotage Isolera™ One system using prepacked columns from Biotage. Silica gel (0.040 – 0.063 mm) from Macherey-Nagel was used for column chromatography. Optical rotations $[\alpha]_D^{20}$ were measured on an P3000 polarimeter from Krüss at 20 °C and are reported in ml·dm^{-1}·g^{-1} with the concentration c being g/100 ml. Fourier-transformed ATR-corrected IR spectra were measured on an Avatar 330 single crystal spectrometer from ThermoNicolet. Melting points (uncorrected) were measured with an MPM-H3 using semi-open capillaries. NMR spectra were recorded as stated individually on Bruker Fourier 300 MHz, Bruker Avance DSX 400 MHz and Bruker Avance III 600 MHz. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance ($CDCl_3$ or DMSO-d_6 from Deutero GmbH) as internal standard. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

5: Dual Strategy to Design New Agents Targeting *Schistosoma mansoni*: Advancing Phenotypic and *Sm*CB1 Inhibitors for Improved Efficacy.

Schistosomiasis is a neglected tropical disease currently still defined as a public health problem by the WHO. On the technical side, they judge the available diagnostics to be insufficient and urge the advance of scientific understanding, as well as effective therapeutic interventions.[517] Currently and for the past 50 years, praziquantel is the main drug in use.[624] Its mode of action was only recently elucidated with the description of its binding site at a transient receptor potential melastatin ion channel (TRP_{PZQ}), which activates aberrant Ca^{2+}-influx and subsequent parasite death.[625,626] One obstacle to drug design for this parasite is its membrane constitution. In its adult lifecycle stage, it covers its cell body with at least two bilayer membranes ("tegumental membrane"), hindering cell permeation by drugs.[627–629]

Since translation of in vitro to cell-based assays is complicated by the parasite's tegumental membrane, the study presented here not only included the in vitro characterization of inhibition potency and kinetics of a large set of peptide inhibitors with covalent warheads, but also a phenotypic screening on both the infantile and adult stage of parasite development. The inhibitor sets investigated for this study were initially repurposed from other projects on the human proteasome, *T. brucei* CatL, human CatS, and *Staphylococcus aureus* sortase A. After gaining insight into their phenotypic effects, two strategies were introduced starting from structures with marked phenotypic activity. The first newly synthesized set was based on the keto vinylester-based proteasome inhibitor **KS309** with low affinity for *Sm*CatB1 (= *Sm*CB1). Modifications that increased the lipophilicity, as well as different warheads, and peptidomimetic moieties (azapeptide, *N*-methylation, inversed sequence) were evaluated. A significant correlation between logP and cellular effect could be observed for this set. The second set of novel inhibitors were derived from the two *Tb*CatL inhibitors **SJ605** and **SJ606**. These were optimized for their in vitro inhibition of *Sm*CatB1. A significant correlation with cellular effect was observed for both logP and non-covalent affinity against *Sm*CatB1. Following the suggestions from flexibility correlations, rigid *N*-terminal capping groups were evaluated. Since the recognition sequence was shortened, only low to moderate selectivity against related proteases was achieved. A special focus was put on the effectiveness against adult schistosomes due to the tegumental membrane alluded to before. To rationalize a link between in vitro affinity towards *Sm*CatB1 and this effect, passive permeability was characterized by PAMPA. This elucidated special features for irreversible inhibitors, for which low permeability can also result in a sufficient cellular effect, even with moderate non-covalent affinity. Additionally, the antischistosomal activity of the most effective inhibitor **2h** was corroborated by describing its high permeability in addition to its high affinity. Taken together, this publication presents inhibitors designed by two distinct approaches, for which the predictive value of calculated physicochemical parameters was discussed. These inhibitors were described in the context of their cellular effect against hard-to-target lifecycle stages of *S. mansoni*, and the in vitro data against *Sm*CatB1 was connected to their cellular effect by their passive permeability.

Own contribution: Evaluating the passive permeability of the SJ-series by PAMPA, correlating the passive permeability with inhibition data against isolated target and cellular effect, writing parts and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of the in vitro affinity, inhibition kinetics, and selectivity, antischistosomal effect, correlation of the physicochemical parameters (logP, PSA) with inhibition data against the isolated target and the cellular effect, non-covalent and covalent docking, writing and editing of the manuscript.

This work was published in ACS Infectious Diseases.

Article reprinted with permission from *ACS Infect. Dis.* **2024**, 10, 5, 1664–1678, "Dual Strategy to Design New Agents Targeting Schistosoma mansoni: Advancing Phenotypic and *Sm*CB1 Inhibitors for Improved Efficacy" Copyright © 2024 American Chemical Society. Published by American Chemical Society.

The following publication quoted within "" (pages 204–218) is the same as the manuscript cited here. The Supporting Information can be accessed online at doi:10.1021/acsinfecdis.4c00020.

ACS Infectious Diseases

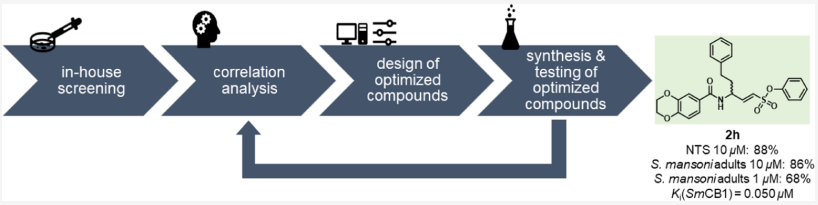
pubs.acs.org/journal/aidcbc Article

Dual Strategy to Design New Agents Targeting *Schistosoma mansoni*: Advancing Phenotypic and *Sm*CB1 Inhibitors for Improved Efficacy

Natalie Fuchs,[§] Robert A. Zimmermann,[§] Marvin Schwickert,[§] Annika Gunkel, Collin Zimmer, Mergim Meta, Kevin Schwickert, Jennifer Keiser, Cécile Haeberli, Werner Kiefer, and Tanja Schirmeister*

Cite This: *ACS Infect. Dis.* 2024, 10, 1664−1678 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: In this study, we have identified and optimized two lead structures from an in-house screening, with promising results against the parasitic flatworm *Schistosoma mansoni* and its target protease *S. mansoni* cathepsin B1 (*Sm*CB1). Our correlation analysis highlighted the significance of physicochemical properties for the compounds' in vitro activities, resulting in a dual approach to optimize the lead structures, regarding both phenotypic effects in *S. mansoni* newly transformed schistosomula (NTS), adult worms, and *Sm*CB1 inhibition. The optimized compounds from both approaches ("phenotypic" vs "*Sm*CB1" approach) demonstrated improved efficacy against *S. mansoni* NTS and adult worms, with **2h** from the "*Sm*CB1" approach emerging as the most potent compound. **2h** displayed nanomolar inhibition of *Sm*CB1 (K_i = 0.050 μM) while maintaining selectivity toward human off-target cathepsins. Additionally, the greatly improved efficacy of compound **2h** toward *S. mansoni* adults (86% dead worms at 10 μM, 68% at 1 μM, 35% at 0.1 μM) demonstrates its potential as a new therapeutic agent for schistosomiasis, underlined by its improved permeability.

KEYWORDS: *schistosomiasis, cathepsin-B-like protease, structure−activity relationship study, covalent inhibition*

Downloaded via UNIV MAINZ on May 16, 2024 at 16:15:08 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Schistosomiasis is an acute or chronic infectious disease that can be caused by different species of the genus *Schistosoma*, a family of human blood flukes. More than 240 million patients are affected by this disease worldwide, most of them in sub-Saharan Africa.[1−3] Of all five *Schistosoma* species that can cause schistosomiasis in humans, *Schistosoma mansoni* is the most widespread, throughout Africa, the Middle East, and even the Americas.[4] Infections can lead to intestinal or hepatic forms of schistosomiasis.[5] During its life cycle, *S. mansoni* depends on two different hosts, freshwater snails of the genus *Biomphalaria* for asexual replication, and a mammalian host for sexual replication.[4] Upon completing asexual replication, the snails release cercariae into the freshwater. These infectious cercariae penetrate human skin exposed to contaminated water and start their sexual replication within the human body. Over the next 5−7 weeks, this larval stage, called schistosomula, develops into matured schistosomes until they begin to mate.[4] Female schistosomes are capable of producing hundreds of eggs daily.[6,7] Egg migration into the intestinal lumen, where they are excreted with feces, completes the sexual reproduction process. Once released into freshwater, the eggs will hatch and release ciliated miracidia, which can now infect the snail hosts to begin asexual replication.[6] Fecal excretion of eggs is not quantitative, meaning that a considerable number of eggs remain in the human body.[7] After 1−2 weeks, the eggs die, no matter if they reach freshwater.[4] Hundreds of trapped, partially dead eggs in the human body can induce severe inflammatory reactions that may cause the full picture of the schistosomiasis disease.[8,9]

Received: January 9, 2024
Revised: April 18, 2024
Accepted: April 18, 2024
Published: April 30, 2024

ACS Publications © 2024 American Chemical Society

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

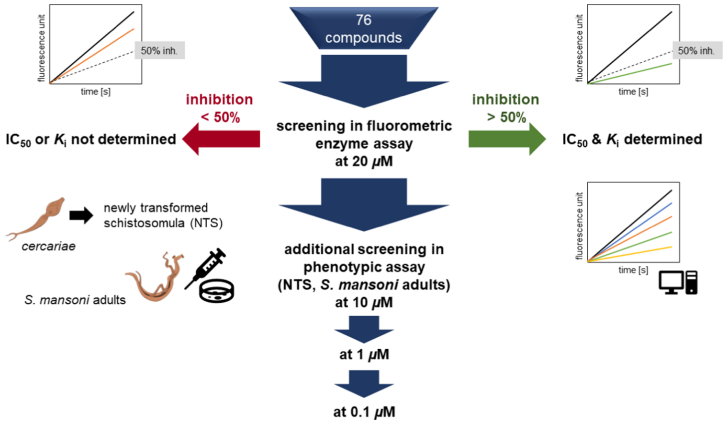
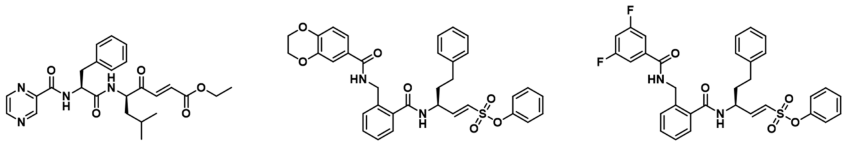


Figure 1. Workflow of the initial screening of in-house compounds. All compounds were screened for *Sm*CB1 inhibition in a fluorometric enzyme assay as well as for their efficacy in a phenotypic assay against NTS and adult worms.



KS309
NTS 10 μM: 100%
S. mansoni adults 10 μM: 76%
K_i(*Sm*CB1) > 100 μM

SJ606
NTS 10 μM: 100%
S. mansoni adults 10 μM: 58%
K_i(*Sm*CB1) = 0.0034 μM

SJ605
NTS 10 μM: 100%
S. mansoni adults 10 μM: 55%
K_i(*Sm*CB1) = 0.0018 μM

Figure 2. Phenotypic screening of in-house library compounds—top 3 compounds with screening results (% dead NTS and *S. mansoni* adult worms at the respective concentration) and K_i values.

Today, the only approved drug for the treatment of schistosomiasis is praziquantel (**PZQ**). Although praziquantel was introduced in the 1970s, very few cases of treatment failure have been reported.[10] With a reported cure rate of ∼90%, **PZQ** remains a powerful tool to fight schistosomiasis. Despite the achieved success, the actual mechanism of action remained elusive until 2021, when Park and co-workers demonstrated that the transient receptor potential melastatin type 8 ion channel (Trpm8) is the target of praziquantel.[11] However, to control or even eradicate this disease, it seems necessary to pursue further drug development to obtain a drug with a broader activity range, also covering the juvenile stage and treat those patients who have shown resistance to **PZQ**. By focusing on additional targets that are vital to the parasite, an improved treatment could be accomplished. Recent publications have shown that the cathepsin-B-like protease of *S. mansoni* (*Sm*CB1) presents a promising target for low-molecular-weight inhibitors, since its inhibition is lethal to *S. mansoni*.[12−14] Previous studies also showed that suppression of *Sm*CB1 activity in early-stage schistosomula had a long-term effect on their growth and development, underlining a certain vulnerability arising from this target.[15]

To date, several *Sm*CB1 inhibitors have been reported. One well-known example is the pan-cathepsin inhibitor **K11777** (IC_{50} = 0.0021 μM, k_{second} = 8.8 × 10^4 M^{-1} s^{-1}), an irreversible vinylsulfone.[13,16] It has proven to be efficient against schistosomes in mice, reducing both worm numbers and egg production.[14] Based on the **K11777** scaffold, Jílková and co-workers have developed the potent irreversible *Sm*CB1 inhibitors **WRR-391** (IC_{50} = 0.2 nM, k_{second} = 2.1 × 10^5 M^{-1} s^{-1}) and **WRR-286** (IC_{50} = 0.6 nM, k_{second} = 2.0 × 10^5 M^{-1} s^{-1}).[13,16] The compounds are effective against newly transformed schistosomula (NTS) at concentrations of 10 μM (**WRR-391** and **WRR-286**) and 1 μM (**WRR-286** only) after 72 h.[13] Additionally, the same authors introduced azapeptide nitriles as covalent-reversible *Sm*CB1 inhibitors with high efficacy against NTS.[12] However, their efficacy against *S. mansoni* adults has not yet been published.

In order to discover new potential antischistosomal compounds, we developed a workflow (see **Graphical Abstract**) starting with a screening that includes 76 compounds, providing us with lead structures (for all structures and data see Supporting Information) for a dual optimization approach. We then performed a correlation analysis of the entire screening data to obtain optimization ideas. From these ideas, we designed new compounds with higher lipophilicity and decreased structural flexibility, also using molecular docking approaches to improve affinity for *Sm*CB1. The new compounds were synthesized and tested on the target enzyme as well as in a phenotypic assay.

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

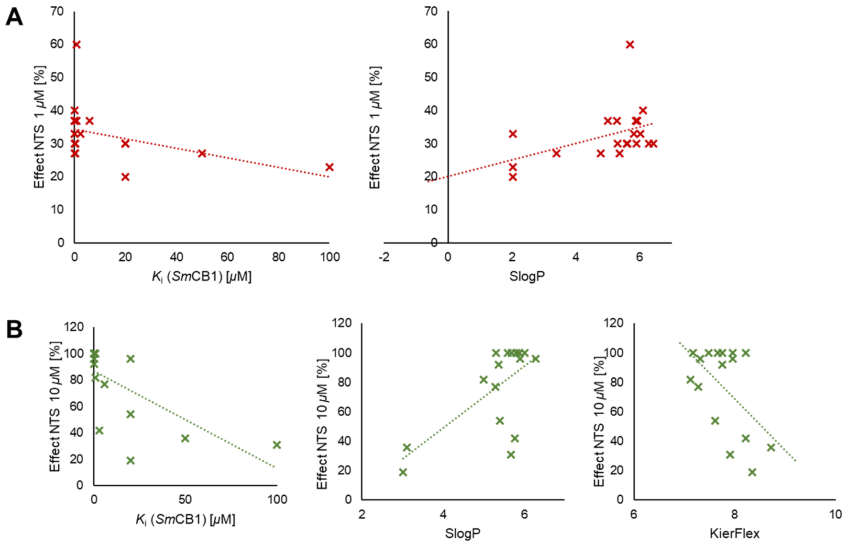


Figure 3. Correlations between K_i values or physicochemical parameters and the effects found in the phenotypic screenings. (A) All compounds (n = 76). Left: effect on NTS at 1 μM vs K_i (*Sm*CB1) in μM (n = 27, p = 0.025). Right: effect on NTS at 1 μM vs $S\log P$ (n = 27, p = 0.021). (B) Compounds of the **SJ600** series (**SJ600**-**SJ620**, n = 16). Left: effect on NTS at 10 μM vs K_i (*Sm*CB1) in μM (n = 16, p = 0.005). Middle: effect on NTS at 10 μM vs $S\log P$ (n = 16, p = 0.005). Right: effect on NTS at 10 μM vs KierFlex (n = 16, p = 0.032).

RESULTS AND DISCUSSION

Screening. We initiated our search for lead structures with antischistosomal activity by screening an in-house library of 76 compounds. The library consisted of several cysteine protease-targeting inhibitors decorated with different covalent warheads, such as fluorovinylsulfon(at)es, nitriles, aldehydes, 4-oxoe-noates, nitroalkenes, and acrylamides.[17−22] The compounds were tested in phenotypic assays on NTS and *S. mansoni* adults.[23] In addition, we evaluated all compounds for their inhibitory effect on *Sm*CB1 using a fluorometric enzyme assay to analyze potential correlations with the results of the phenotypic assay. We determined IC_{50} and K_i values for all compounds exhibiting *Sm*CB1 inhibition >50% at 20 μM. In the phenotypic assay, all substances were initially analyzed at 10 μM on NTS. If an efficacy of >70% dead NTS was achieved, a second screening was performed at 1 μM on NTS and at 10 μM on *S. mansoni* adults. Lower compound concentrations of 1 and 0.1 μM were tested on adult worms when 70% activity (i.e., >70% dead worms) was achieved in the respective previous screening. A table with all data from the screening can be found in the Supporting Information. The procedure of the initial screenings is depicted in Figure 1.

The top three compounds (**KS309, SJ605, SJ606**) from the phenotypic screening against *S. mansoni* adults are shown in Figure 2. While two of them are vinylsulfonate-based cysteine protease inhibitors (**SJ605**, **SJ606**), originally developed for rhodesain,[18,19] the third is a 4-oxoenoate-based bortezomib congener with *S,R*-configuration.[24] All three compounds are highly active against NTS at 10 μM (100% efficacy, i.e., 100% dead NTS). Although **SJ605** and **SJ606** exhibit high *Sm*CB1 inhibition with K_i values in the low nanomolar range, a reduced efficacy against adult worms at 10 μM (55% dead worms for **SJ605**, 58% dead worms for **SJ606**) was achieved. In contrast, **KS309** shows weak inhibition of *Sm*CB1 (K_i > 100 μM) but a stronger effect on adult worms (76% at 10 μM, 29% at 1 μM) in the phenotypic assay, indicating that it addresses another target.

Based on these findings, we pursued a dual strategy to increase the antischistosomal effect of both compound classes on adult *S. mansoni*. In order to maximize their effect, we also focused on improving their physicochemical properties, such as lipophilicity, since the substances have to pass two cell membranes to reach their target site.[25] This resulted in two approaches, the "phenotypic" approach focusing on **KS309** derivatization and the "*Sm*CB1" approach focusing on **SJ605**/**SJ606** derivatizations.[24]

Chemistry. Compounds **1a−m** were prepared by TBTU peptide coupling chemistry. Synthesis of their respective warheads was achieved by Horner−Wadsworth−Emmons (HWE) reactions. All synthetic pathways are described in detail in the Supporting Information.

The fluorovinylsulfone- and -sulfonate-based inhibitors **2a−q** were synthesized as previously published. A schematic overview is shown in Scheme S7 in the Supporting Information.[17,18,25]

Correlations and Hit Optimizations. *Correlation Analysis.* To identify correlations and ideas for optimization, we calculated several important physicochemical parameters ($S\log P$, TPSA, KierFlex) using molecular operating environment (MOE) (see Table S1 in the Supporting Information).[26] With these data, we performed a correlation analysis for each physicochemical parameter as well as for the *Sm*CB1 inhibition with the observed effects in the phenotypic screening, including NTS (10, 1 μM inhibitor concentration) and *S. mansoni* adults (10 μM inhibitor concentration). The resulting correlation factors and their respective p-values are shown in

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

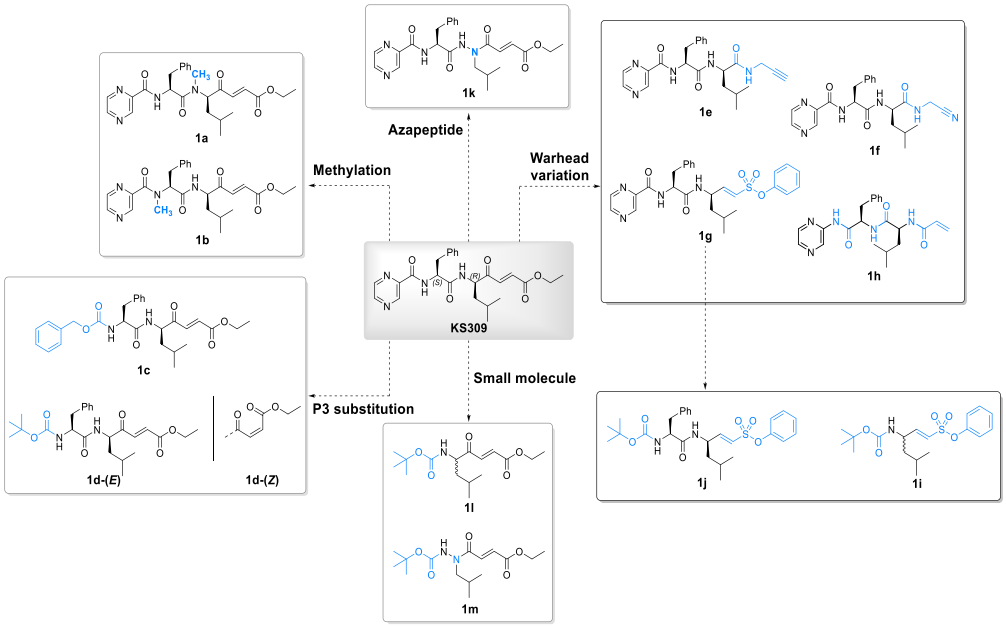


Figure 4. Modifications of **KS309**.

Table S2 in the Supporting Information. Only p-values < 0.05 are considered statistically significant. Respective diagrams are shown in Figure 3.

There is a slight, but statistically significant, correlation between the K_i values and the effects in the phenotypic screening against NTS at 1 μM (p = 0.025). The correlation factor (r) is negative (r = 0.43), implying that lower K_i values lead to higher effects in vitro. For *S. mansoni* adults, this correlation is not significant (p = 0.18). Since the expression levels of proteases fluctuate in different stages of development, adult worms may overcome inhibition of *Sm*CB1 more effectively than NTS. Candidate proteases may be *S. mansoni* cathepsin D or cathepsin L which are also involved in the digestion of blood hemoglobin. Furthermore, an additional second membrane has to be crossed in adult worms which may play a role for the in vitro activity of the compounds.[15,27,28] The positive correlation factor between *S* log *P* and the effect against NTS (r = 0.44 for 1 μM inhibitor concentration) indicates that higher *S* log *P* values are beneficial (p = 0.021). TPSA and molecular flexibility (KierFlex) do not have significant effects.

Since two compounds from the **SJ600** series (i.e., **SJ600**–**SJ620**), namely, **SJ605** and **SJ606**, are among the top three, we decided to analyze the subset containing only these structurally related compounds (n = 16, results in Table S3 in the Supporting Information). The compounds are known to be potent cysteine protease inhibitors with low cytotoxicity as reported by Jung et al.[19]

Again, there is a significant negative correlation (p = 0.005) between the K_i value for *Sm*CB1 and the effects against NTS at 10 μM inhibitor concentration (r = 0.67). The *S* log *P* values correlate with the effect against NTS at 10 μM (r = 0.66, p = 0.005) as well. Within this compound set, an additional correlation between the molecular flexibility (KierFlex) and the effect against NTS at 10 μM (r = 0.54, p = 0.032) can be found, suggesting that lower flexibility is advantageous. KIER's definition of flexibility is structure based upon molecular size, branching, cycles, and number of heteroatoms.[29]

The discussed correlations for the complete compound set and the **SJ600** subgroup are shown in Figure 3. In summary, stronger correlations are found within the **SJ600** subgroup as can also be drawn from the correlation factors (for example NTS 10 μM vs K_i: all compounds r = 0.27 vs **SJ600** subgroup r = 0.67).

Derivatization of KS309. In a first approach, we investigated the structural modifications of **KS309** (Figure 4). To increase lipophilicity, we reduced the number of hydrogen-bond donors by methylating one of each of the amide bond nitrogen atoms (**1a**, **1b**). Furthermore, the effect of replacing the polar pyrazinoyl moiety in P3 with the more lipophilic Cbz (**1c**) and Boc (**1d**) groups was investigated as such modifications have also been shown to be effective in *Sm*CB1 inhibitor design.[12] Since 4-oxoenoates exhibit high reactivity against nucleophiles, we tested alternative warheads such as propargylamide (**1e**), nitrile (**1f**), and vinylsulfonate (**1g**). To analyze the effect of inverted amide bonds, the acrylamide derivative **1h** was prepared. The precursors **1i** and **1j** of vinylsulfonate **1g** were also tested due to their higher lipophilicity. Synthesizing 4-oxoenoates and vinylsulfonates causes epimerization at P1. To avoid the separation of diastereomers at a later stage of the synthetic procedure, we prepared the P1 azapeptide derivative **1k** of **KS309**. In addition, truncated derivatives (**1l**, **1m**) of **KS309** were synthesized and tested.

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

Table 1. Physicochemical Parameters and In Vitro Effects of 1a−m

				NTS effect (%)[b]		*S. mansoni* adults effect (%)[b]	
cpd	*S* log *P*[a]	TPSA (Å^2)	*Sm*CB1 inhibition at 20 μM (%)[b]	at 10 μM	at 1 μM	at 10 μM	at 1 μM
KS309	2.04	127	58 ± 5	100 ± 0	23 ± 12	76 ± 6	29 ± 2
1a	2.38	119	11 ± 6	100 ± 0	48 ± 0	61 ± 4	n. d.
1b	2.38	119	12 ± 9	96 ± 0	44 ± 6	48 ± 0	n. d.
1c	4.01	111	67 ± 0	98 ± 2	15 ± 8	65 ± 2	36 ± 6
1d-(*E*)	3.34	111	42 ± 6	100 ± 0	78 ± 2	61 ± 0	n. d.
1d-(*Z*)	3.34	111	47 ± 1	100 ± 0	40 ± 0	40 ± 2	n. d.
1e	1.10	113	18 ± 7	46 ± 0	n. d.	n. d.	n. d.
1f	0.99	137	10 ± 8	39 ± 0	n. d.	n. d.	n. d.
1g	3.27	127	68 ± 4	100 ± 0	38 ± 4	100 ± 0	43 ± 0
1h	1.86	113	9 ± 4	48 ± 2	n. d.	n. d.	n. d.
1i	3.85	82	9 ± 1	98 ± 2	27 ± 4	93 ± 4	36 ± 2
1j	4.58	111	80 ± 1	100 ± 0	65 ± 0	100 ± 0	45 ± 2
1k	1.45	131	17 ± 6	50 ± 13	33 ± 5	47 ± 10	n. d.
1l	2.61	82	7 ± 8	100 ± 0	14 ± 2	58 ± 2	n. d.
1m	2.03	85	3 ± 8	61 ± 4	32 ± 2	34 ± 0	n. d.

[a]Calculated using *MOE* 2019.01.[26] [b]% Dead after 72 h (compared to DMSO control).

We evaluated the **KS309** derivatives for their *Sm*CB1 inhibition as well as their efficacy on NTS and *S. mansoni* adults. The results are listed in Table 1. Additionally, calculated *S* log *P* values and topological polar surface areas (TPSA) are given.

From this new compound series, **1g**, **1i**, and **1j** showed higher activity on *S. mansoni* adults (93−100% vs 76 ± 6% at 10 μM; 36−45, vs 29 ± 2% at 1 μM) compared to **KS309**. **1a**, **1b**, and **1d**-(*E*) exhibited slightly lower efficacy (61−65% at 10 μM), while all other compounds appeared to be distinctly less efficient (<60% at 10 μM). A correlation between the reduction of the TPSA compared to **KS309** (<127 Å^2) and the increase in efficacy on adult worms could not be observed. Although **1i** and **1j** with TPSAs of 111, and 82 Å^2, respectively, showed higher efficacy, other compounds with TPSAs ranging from 82 to 119 Å^2 (**1a−c**, **1d**-(*E*), **1d**-(*Z*), **1e** and **1h**) were less efficient. On the other hand, a statistically significant correlation between *S* log *P* values and efficacies on NTS at 10 μM could be observed (r = 0.84, p = 0.0001), suggesting that higher *S* log *P* values result in higher efficacy (see Table S4 and Figure S126 in the Supporting Information). Compounds with *S* log *P* values of ≥2.04 (**KS309**, **1a−c**, **1d**-(*E*), **1d**-(*Z*), **1g**, **1i**, **1j**, and **1l**) appeared to have an efficacy of at least 96% at 10 μM. For compounds with *S* log *P* values ranging from 0.99 to 2.03, the activities were only 39−61%. This correlation was also observed for *S. mansoni* adults at 10 μM (r = 0.69, p = 0.015). Overall, **1g** as well as its precursors **1i** and **1j** showed the highest activities within the new compound series on both NTS and *S. mansoni* adults. Interestingly, **1g** and **1j** even exhibited stronger inhibition of *Sm*CB1 compared to **KS309** (68−80, vs 58% at 20 μM), albeit weak compared to the SJ-series (**SJ605**/**SJ606**). Substituting the 4-oxoenoate warhead of **KS309** with a less reactive vinylsulfonate (**1g**), retained the same TPSA (127 Å^2) but increased the *S* log *P* value (3.27 vs 2.04). Replacing the polar pyrazinoyl moiety of **1g** with a more lipophilic Boc group (**1j**) did not change efficacy, although it decreased the TPSA (111 vs 127 Å^2) and increased the *S* log *P* value (4.58 vs 3.27, **1g** and 2.04, **KS309**). In summary, the antischistosomal effect could be increased. It appeared that the vinylsulfonate moiety is the most suitable to achieve higher efficacy, assumably due to its inherent higher lipophilicity. However, the target of the **KS309** derivatives remains unknown. Therefore, rather than focusing on the phenotypic approach by optimizing the series, we focused on optimizing the **SJ605**/**SJ606** inhibitors, since their in vitro effects appeared to correlate with *Sm*CB1 inhibition.

Derivatization of the SJ600 Series Compounds. The correlation analysis suggested that higher rigidity results in higher activity in the phenotypic assay (see Table S3 in the Supporting Information). Nevertheless, the *Sm*CB1 inhibition should not be compromised since it also correlates with the effect in NTS. Figure 5 shows noncovalent docking poses (generated with LEADIT)[30] for lead structures **SJ605** (

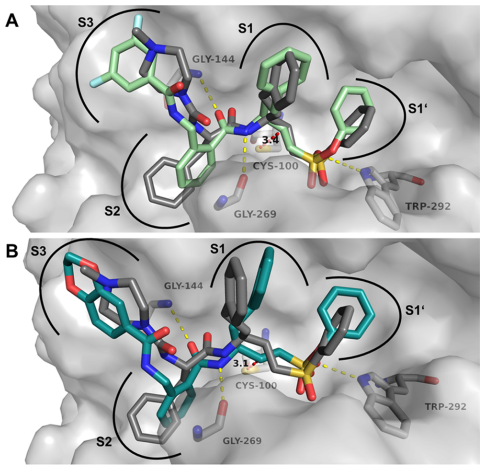


Figure 5. Noncovalent docking poses for lead compounds **SJ605** and **SJ606** obtained with LeadIT.[30] Images were generated using PYMOL.[31] The ligand interaction map was generated with MOE.[26] PDB ID: 3s3r.[16] (A) Docking pose for **SJ605** (pale green). Overlay with **K11777** (gray). Important interactions are shown as yellow dashes. Hyde score: −34 kJ/mol. (B) Docking pose for **SJ606** (deep teal). Overlay with **K11777** (gray). Important interactions are shown as yellow dashes. Hyde score: −25 kJ/mol.

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

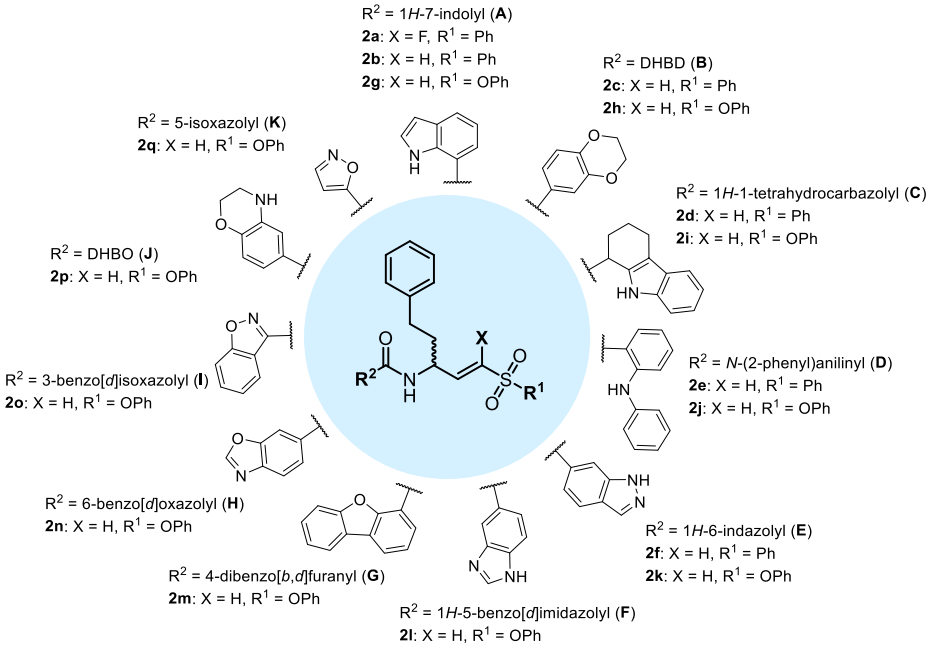
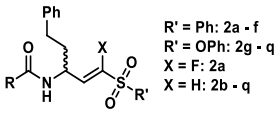


Figure 6. Optimized vinylsulfon(at)e-based compounds **2a**–**q**.

Table 2. Inhibition Data from Fluorometric Enzyme Assay for Fluorovinylsulfon(at)es 2a–q[a,b]



		*Sm*CB1			CatB		CatL	
cpd	R	K_i (μM)	k_{inact} (s^{-1})	k_{second} (M^{-1} s^{-1})	K_i (μM) or %inh @20 μM	SI	K_i (μM) or %inh @ 20 μM	SI
K11777		0.0029 ± 0.0005[b]	4.7×10^{-3}	1.6×10^{6}				
2a	A	5.8 ± 1.5			30%	>3	45%	>3
2b	A	4.7 ± 0.3[b]	1.1×10^{-4}	23	38%	>4	n. i.	>4
2c	B	6.7 ± 0.8[b]	2.3×10^{-4}	34	29%	>3	24%	>3
2d	C	2.2 ± 0.8[b]	1.6×10^{-4}	73	41%	>9	41%	>9
2e	D	>100	n. d.	n. d.	28%	n. d.	30%	n. d.
2f	E	2.5 ± 0.6[b]	2.5×10^{-4}	1.0×10^{2}	22%	>8	32%	>8
2g	A	0.72 ± 0.01[b]	1.8×10^{-3}	2.5×10^{3}	n. i.	>28	43%	>28
2h	B	0.050 ± 0.003[b]	7.0×10^{-4}	1.4×10^{4}	29%	>400	37%	>400
2i	C	1.6 ± 0.4[b]	3.0×10^{-4}	1.9×10^{2}	31%	>13	28%	>13
2j	D	2.3 ± 1.0[b]	5.0×10^{-4}	2.2×10^{2}	20%	>9	26%	>9
2k	E	0.83 ± 0.07[b]	1.3×10^{-3}	1.5×10^{3}	6.5 ± 1.0	7.8	0.54 ± 0.05	7.8
2l	F	0.078 ± 0.007[b]	1.1×10^{-3}	1.4×10^{4}	0.41 ± 0.06	5.3	4.8 ± 0.28	5.3
2m	G	>100	n. d.	n. d.	n. i.	n. d.	20%	n. d.
2n	H	0.030 ± 0.020[b]	2.4×10^{-3}	8.0×10^{4}	1.0 ± 0.2	33	2.2 ± 0.2	73
2o	I	0.050 ± 0.030[b]	1.8×10^{-3}	3.6×10^{4}	48%	>400	0.37 ± 0.03	7.4
2p	J	0.41 ± 0.30[b]	5.4×10^{-3}	1.3×10^{4}	39%	>49	8.2 ± 2.1	20
2q	K	0.15 ± 0.04[b]	2.8×10^{-3}	1.8×10^{4}	0.26 ± 0.01	1.7	2.3 ± 0.2	15

[a]n. d., not determined; n. i., no inhibition at 20 μM inhibitor concentration; %inh at 20 μM, mean inhibition in % from three independent measurements at 20 μM inhibitor concentration with SD < 20%; SI, selectivity index. [b]Time-dependent inhibition.

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

K_i = 1.8 nM) and **SJ606** (K_i = 3.4 nM) revealing that the most important interactions occur with the S1′ and S1 subsites. Gly144 and Gly269 both interact with the amide group between P1 and P2, and Gln94 can form an H-bond with the sulfonyl-oxygen atom as well as with Trp292. The residues in P2/P3 do not form essential interactions in the active site since the residues rather protrude from the binding site.

Based on these interactions, we decided to modify the P2/P3 residues while maintaining the P1′ and P1 residues. To enhance rigidity, we chose various bi- or tricycles in P2, such as indole, DHBD, and several others, resulting in compounds **2a−q** (Figure 6). We prepared irreversible vinylsulfones (**2b−f**) and -sulfonates (**2g−q**) as well as one compound with a reversible α-fluorovinylsulfone warhead (**2a**).

***Sm*CB1 Inhibition and Structure−Activity Relationship.** We tested all new vinylsulfon(at)e-based compounds in fluorometric enzyme assays to evaluate their inhibitory activity for *Sm*CB1 and their selectivity toward off-target cathepsins (assay conditions in the Supporting Information). The results are shown in Table 2.

All new vinylsulfone-based compounds except compound **2e** (K_i > 100 μM) are good *Sm*CB1 inhibitors with K_i values in the low micromolar range. The covalent reaction occurs slowly (k_{inact} around 10^{-4} s^{-1}) and there is moderate selectivity toward cathepsins B and L. **2a**, an α-fluorovinylsulfone, is the only inhibitor in the series that reversibly inhibits the enzyme by a covalent-reversible Michael addition.[17,18] The K_i is in the same range as for the irreversible vinylsulfones with only little selectivity toward CatB and CatL.

Table 2 also shows the inhibition data for vinylsulfonate-based compounds **2g−q**. Notably, their inhibitory potency is much better in comparison to their vinylsulfone counterparts, for example, **2h** (K_i = 0.050 μM, Figure S121 in the Supporting Information) vs **2c** (K_i = 6.7 μM). Moreover, the k_{inact} values of most vinylsulfonates are higher than those of the vinylsulfone counterparts, suggesting a faster covalent bond formation. Since vinylsulfonates proved to be more potent *Sm*CB1 inhibitors, we designed and synthesized additional compounds with improved physicochemical properties, such as *S* log *P* and TPSA (**2l−q**). With these compounds, we obtained K_i values in the nanomolar range. Among these, **2n** (K_i = 0.030 μM) and **2o** (K_i = 0.050 μM) were found to be the top two compounds in terms of inhibitory potency. However, **2h** performs better in terms of selectivity (>400-fold toward CatB and CatL) compared to the moderate selectivity of **2n** (33-fold vs CatB, 73-fold vs CatL) and **2o** (>400-fold vs CatB, 7.4-fold vs CatL).

Docking and SAR Discussion. Docking scores and Hyde scores (Table S6 in the Supporting Information) for the new SJ compounds were generated using LEADIT.[30] Since we maintained the P1′ and P1 residues of **SJ605** and **SJ606** (Figure 7B), we still observed interactions between Gly144 and Gly269 with the amide bond as well as an additional H-bond with Gly143 in the noncovalent docking computations. Gln94 and Trp292 are also involved in H-bond formation with the sulfonyl group. The docking poses for top compound **2h** overlapping with **SJ606** is shown in Figure 7, revealing that the important interactions with Gly144 and Gly269 and the amide bond between P1 and P2 were conserved.

2i surprisingly shows a very weak *SmCB1* inhibition (K_i > 100 μM) which is not reflected by its Hyde score (−33 kJ/mol). Looking at the docking poses of both **2i** and **SJ606**, we noticed that the distance between the β-carbon of the vinylsulfonate double bond and the sulfur atom of the nucleophilic Cys100 that undergoes the covalent reaction is increased from 2.9 to 4.3 Å (Figure S129 in the Supporting Information). Covalent reactions occur more readily at distances up to 3.5 Å, which could explain the poor inhibitory activity.[33] The same may apply for **2m** (K_i > 100 μM) for which a distance of 3.9 Å was found. Additionally, covalent docking for **2i** and **2m** did not yield reasonable docking poses which would resemble the cocrystallized ligand. The covalent docking scores for **2h**, **2i**, **2l**, and **2m** and the distances for compounds **2a−q** are listed in Table S6 in the Supporting Information.

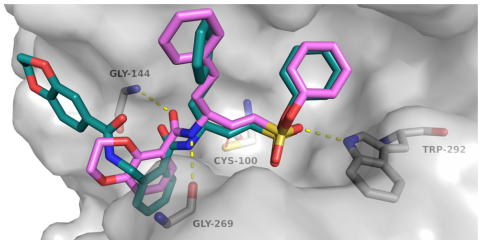


Figure 7. Molecular docking of **2h** and the initial lead **SJ606**. Noncovalent docking with LEADIT, images were generated with PYMOL, ligand interaction map generated with MOE.[26,31] PDB ID: 3s3r.[16] Overlay of **2h** (violet) and **SJ606** (deep teal). Important active site interactions are shown as yellow dashes. Hyde score: −39 kJ/mol.[32]

In Vitro Activity in Phenotypic Assay. The results for the phenotypic assay of **2a−q** compared to leads **SJ605** and **SJ606** as well as the approved drug **PZQ** are shown in Table 3. All compounds were tested against NTS and *S. mansoni* adults.

Compared to starting compounds **SJ605** and **SJ606**, several vinysulfonate-based inhibitors are more active against NTS and *S. mansoni* adults. Especially compounds **2h**, **2i**, **2k**, and **2p** are outstanding since they display a high efficacy (≥75%) against *S. mansoni* adults at 10 μM. **2h** has the highest efficacy in the series with 68% at 1 μM and 35% at 0.1 μM, which is in a similar range as **PZQ** (IC_{50} = 0.16 μM).[35] Notably, **2h** and **2p** (*S. mansoni* adults: 75% dead at 10 μM, 23% dead at 1 μM) share a high structural similarity, also with **SJ606**. This indicates that the DHBD/DHBO moiety is beneficial for phenotypic efficacy, particularly in *S. mansoni* adults. A correlation analysis suggests that lower K_i values result in a higher efficacy in NTS at 10 μM (r = −0.47, p = 0.048, Table S5 in the Supporting Information). Similarly, a slight correlation was found between TPSAs (ranging from 75 up to 110 Å^2) and efficacies (r = 0.55, p = 0.023, Table S5 in the Supporting Information).

Although the initial correlation analysis suggested that a higher lipophilicity could be beneficial, this is not the case for the new vinylsulfon(at)e-based compounds. An example is **2e** with the highest *S* log *P* in the series (6.15) but only limited potency (NTS: 42% at 10 μM, *S. mansoni* adults: n. d.). In contrast, **2k** is one of the top compounds (NTS: 100% at 10 μM, *S. mansoni* adults: 86% at 10 μM) with one of the lowest *S* log *P* values (4.22). Flexibility also seems to have a smaller impact on efficacy than assumed. The top compounds' KierFlex values range from 5.54 to 6.59, whereas compounds with a higher rigidity, such as **2a** (5.09), **2b** (5.14), and **2f** (5.09), did not perform as well in the phenotypic assay.

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

Table 3. Physicochemical Parameters and In Vitro Effects of 2a−q Compared to Praziquantel (PZQ) and SJ605, SJ606

					NTS		*S. mansoni* adults		
cpd	R	*S* log *P*[a]	TPSA (Å^2)[a]	KierFlex[a,b]	effect (%)[c] 10 μM	effect (%)[c] 1 μM	effect (%)[c] 10 μM	effect (%)[c] 1 μM	effect (%)[c] 0.1 μM
SJ605		5.81	101.57	7.96	100 ± 0	33 ± 12	55 ± 4	n. d.	n. d.
SJ606		5.31	120.03	8.22	100 ± 0	30 ± 0	58 ± 5	n. d.	n. d.
2a	A	5.18	79.03	5.30	60 ± 1	38 ± 2	n. d.	n. d.	n. d.
2b	A	4.89	79.03	5.14	85 ± 1	40 ± 3	27 ± 2	n. d.	n. d.
2c	B	4.18	81.70	6.11	92 ± 0	27 ± 1	29 ± 0	n. d.	n. d.
2d	C	5.69	79.03	5.75	100 ± 0	48 ± 1	31 ± 2	n. d.	n. d.
2e	D	6.15	75.27	6.36	42 ± 4	27 ± 1	n. d.	n. d.	n. d.
2f	E	4.28	91.92	5.09	75 ± 2	38 ± 4	47 ± 6	n. d.	n. d.
2g	A	4.82	88.26	5.59	100 ± 0	27 ± 1	63 ± 6	n. d.	n. d.
2h	B	4.11	90.93	6.59	88 ± 0	33 ± 0	86 ± 2	68 ± 0	35 ± 2
2i	C	5.63	88.30	6.19	100 ± 0	42 ± 2	86 ± 2	57 ± 0	n. d.
2j	D	6.08	84.50	6.84	100 ± 0	35 ± 1	31 ± 2	n. d.	n. d.
2k	E	4.22	101.20	5.54	100 ± 0	33 ± 2	86 ± 2	36 ± 0	n. d.
2l	F	4.22	101.20	5.54	100 ± 0	33 ± 2	63 ± 2	n. d.	n. d.
2m	G	6.04	81.70	6.24	100 ± 0	44 ± 4	71 ± 4	29 ± 4	n. d.
2n	H	4.48	98.50	5.74	100 ± 0	48 ± 8	42 ± 4	n. d.	n. d.
2o	I	4.48	98.50	5.74	100 ± 0	40 ± 4	38 ± 8	n. d.	n. d.
2p	J	4.14	93.73	6.37	100 ± 0	58 ± 10	75 ± 0	23 ± 2	n. d.
2q	K	3.33	98.50	5.83	100 ± 0	44 ± 4	40 ± 2	n. d.	n. d.
PZQ[d]		2.63	40.62	3.10	IC_{50}: 1.5 μg/mL[d] = 4.8 μM		IC_{50}: 0.05 μg/mL[d] = 0.16 μM		

[a]Calculated with MOE.[26] [b]KIER molecular flexibility index.[29] [c]% Dead after 72 h (compared to DMSO control). [d]Meister et al.[34]

In summary, several factors combine to influence phenotypic efficacy in schistosomes. Regardless, compound **2h** displays high potency and should be considered for further studies.

Permeability Study. An assessment of permeability by parallel artificial membrane permeability assay (PAMPA) was performed to add another factor for reasonable drug design[36] and to link affinity against *Sm*CB1 with effects in cell-based experiments. In numbers, high permeability was set to $P_{app} \geq 2 \times 10^{-6}$ cm/s[37,38] which was correlated with an effect against adult schistosomes ≥50% at 10 μM (compare Table 3).

Correctly predicted from its favorable physicochemical properties, **PZQ** showed the best permeability, underlining its position as an effective drug. This PAMPA data is in line with reported Caco-2 apparent permeability for **PZQ** ($P_{app,lit.} = 44 \times 10^{-6}$ cm/s vs $P_{app} = 45.4 \pm 18.7 \times 10^{-6}$ cm/s, $n = 6$).[39]

The most promising compound of this study, **2h**, which showed an effect comparable to **PZQ**, also showed high, although slightly lower permeability ($P_{app} = 8.01 \pm 2.83 \times 10^{-6}$ cm/s, $n = 6$) when compared to **PZQ**. This permeability seems to still ensure sufficing concentrations in the targeted cell. In direct contrast to **2h**, compound **2n** has the same affinity to the target enzyme but has a remarkably lower permeability ($P_{app} < 1 \times 10^{-6}$ cm/s, $n = 4$), which underscores the observed inactivity in the cell-based assay.

An interesting case can be described for **2i** as it shows only limited affinity to *Sm*CB1, but both high permeability ($8.32 \pm 2.03 \times 10^{-6}$ cm/s, $n = 6$) and considerable activity (86% at 10 μM) against adult worms. Possible hypotheses are that a compound with sufficient permeability does not require extraordinary affinity to show a desired effect in cellular testing, especially in the context of covalent irreversible inhibitors or that there are relevant other mechanisms contributing to the overall antischistosomal effect. The same phenomenon occurs for **2g** ($P_{app} = 9.70 \pm 2.02 \times 10^{-6}$ cm/s, $n = 6$). On the other hand, **SJ606** protrudes as a compound with low permeability ($P_{app} < 1 \times 10^{-6}$ cm/s, $n = 6$) but remarkably high affinity against *Sm*CB1, and a resulting relevant cellular effect. Similar observations can be made for K11777 which shows low PAMPA permeability, but with high affinity and an irreversible mode of inhibition, antischistosomal effects can be observed over long incubation periods.[13] These cases indicate a plausible cooperative effect between affinity against *Sm*CB1 and permeability, where lacking properties in one aspect can be (at least partially) compensated by the other.

CONCLUSIONS AND PERSPECTIVES

Using a screening with 76 compounds, we have identified two lead structures, **KS309** and the **SJ600**-scaffold (**SJ605**, **SJ606**), which showed promising results against NTS and *S. mansoni* adults and/or the target protease *Sm*CB1. **KS309** was most effective against *S. mansoni* adults at 10 μM but did not show relevant inhibition of the target protease *Sm*CB1 at 100 μM. **SJ605** and **SJ606**, on the other hand, were less effective against *S. mansoni* at 10 μM but displayed strong *Sm*CB1 inhibition in the low nanomolar range. A correlation analysis revealed the importance of several physicochemical properties for the phenotypic efficacy, such as higher *S* log *P* values or, in the case of the **SJ600s**, higher structural rigidity. Therefore, we started a dual approach to optimize the lead structures with regard to physicochemical properties. The "phenotypic approach" using **KS309** resulted in 14 new derivatives with altered physicochemical properties yielding three optimized compounds (**1g**, **1i**, **1j**). **1g** (*Sm*CB1 inhibition: 68% at 20 μM; *S. mansoni* adults: 100% dead at 10 μM, 43% at 1 μM) and **1j** (*Sm*CB1 inhibition: 80% at 20 μM; *S. mansoni* adults: 100% dead at 10 μM, 45% dead at 1 μM) emerged to be the most potent compounds in this series, although their main target remains unknown. The "*Sm*CB1 approach" using **SJ605** and **SJ606** as leads resulted in 17 new derivatives with higher rigidity and lower TPSAs including seven compounds with improved efficacy (**2g**, **2h**, **2j−m**, **2p**). Several compounds displayed strong *Sm*CB1 inhibition in the nanomolar range with improved phenotypic efficacy compared to **SJ605** and **SJ606**. Vinylsulfonates generally proved to be more effective

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

than vinylsulfones. The most potent compound, **2h**, shows nanomolar *Sm*CB1 inhibition while maintaining selectivity toward human off-target cathepsins (K_i [*Sm*CB1] = 50 nM, SI > 400). Underlining its potential for drug development, it shows adequate permeability (P_{app} = 8.01 ± 2.83 × 10^{-6} cm/s) for utilization in living systems. The efficacy of **2h** toward *S. mansoni* adults has improved enormously (86% at 10 μM, 68% at 1 μM and 35% at 0.1 μM), demonstrating efficacy in a similar range as found for the approved drug **PZQ** (*S. mansoni* adults IC_{50} = 0.16 μM).

Further optimization of **2h** has the potential to yield even more potent compounds, offering a promising alternative to the currently only available drug PZQ.[10]

METHODS

General Procedures. *General Procedure A—TBTU Coupling with Amine Hydrochlorides.* To a 0 °C cold solution of the carboxylic acid (1.0 equiv) in DCM and/or DMF, HOBt·H_2O (1.0 equiv) and 2,4,6-collidine (2.0 equiv) or DIPEA (3.5 equiv) were added. After stirring at 0 °C for 30 min, TBTU (1.0 equiv) was added. The solution was stirred for further 30 min at 0 °C, and the amine hydrochloride (1.0 or 1.1 equiv) was added. After stirring at rt overnight, the solvent was removed under reduced pressure. If DMF was used, it was removed by co-distillation with *n*-heptane (3 mL × 150 mL) under reduced pressure. The residue was taken up in ethyl acetate and washed with saturated $NaHCO_3$ solution (3×) and 1 M HCl (3×). The filtrate was concentrated under reduced pressure to yield the desired product. If necessary, further purification by column chromatography was conducted.

General Procedure B—HWE Reaction with Glyoxylic Acid Ethyl Ester. Anhydrous LiCl (1.2 equiv) was placed in a Schlenk flask under an argon line at 0 °C, and a solution of the dimethylphosphonate (1.0 equiv) in dry ACN (10 mL) was added. DIPEA (1.0 equiv) was added, and the mixture was stirred at 0 °C for 15 min. Then, a solution of ethyl glyoxylate (freshly distilled over P_2O_5, 2.0 equiv) in dry ACN (10 mL) was slowly added and stirred for 1 h at 0 °C. The reaction was quenched by the addition of citric acid solution (10% in water, 10 mL), and extracted with ethyl acetate (4 mL × 30 mL). The combined organic extracts were dried over Na_2SO_4, and the solvent was removed by distillation under reduced pressure. The resulting crude product was purified by column chromatography on silica gel.

General Procedure C—HWE Reaction with Other Aldehydes. Anhydrous LiCl (1.2 equiv) was placed in a Schlenk flask under an argon line at 0 °C, and a solution of the diethylphosphonate (1.0 equiv) in dry ACN (10 mL) was added. DBU (1.0 equiv) was added, and the mixture was stirred at 0 °C for 15 min. Then, a solution of the aldehyde (1.0 equiv) in dry ACN (10 mL) was slowly added and stirred for 1 h at 0 °C. The reaction was quenched by the addition of citric acid solution (10% in water, 10 mL), and extracted with ethyl acetate (4 mL× 30 mL). The combined organic extracts were dried over Na_2SO_4, and the solvent was removed by distillation under reduced pressure. The resulting crude product was purified by column chromatography on silica gel.

General Procedure D—HWE Reaction with Boc-Homophenylalaninal.[18] The corresponding phosphonate (1−3, 1 equiv) was dissolved in dry THF and cooled to −80 °C. Then, LHMDS (1 M in THF, 1.4 equiv) was added over 20 min. The aldehyde (5, 0.95 equiv) was dissolved in THF and added to the solution. After stirring for 2 h, the mixture was warmed to rt. THF was removed under reduced pressure and the residue was extracted with EA (3×). The combined organic extracts were washed with water (3×) and brine (1×). The extract was dried over Na_2SO_4, and volatiles were removed under reduced pressure. After column chromatographic workup (CH/EA 6:1), the products were obtained as colorless oils.

General Procedure D—Boc Deprotection. To a solution of the Boc-protected amine in DCM, TFA was added dropwise at 0 °C. The reaction was stirred overnight. Subsequently, all volatiles were removed under reduced pressure. The product was then lyophilized to remove TFA. Alternatively, 4 M HCl in dioxane was used at room temperature.

The fluorovinylsulfone- and -sulfonate-based inhibitors **2a−q** were synthesized as previously published.[17,18,25]

Recombinant Expression and Purification of *Sm*CB1. *Pichia pastoris* yeast cells harboring a pPICZαA vector coding for *Sm*CB1 were kindly donated by the group of Prof. Dr. Michael Mareš (Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czech Republic). Expression was conducted according to the literature.[40] In brief, yeast cells were grown in 100 mL of BMG medium at 30 °C overnight. The medium was then used to inoculate 500 mL of BMG medium, the yeast cells were again grown overnight at 30 °C. Yeast cells were harvested by centrifugation at 20 °C and 4000*g* for 40 min, and then suspended in 2 L of BMM medium to induce overexpression of *Sm*CB1. The medium was split into 250 mL batches and yeast cells were grown for 4 days at 30 °C, pH was adjusted every 12 h with 25% ammonia solution, and overexpression was maintained by adding 2.5 mL methanol every 12 h to 250 mL of medium. After 4 days yeast cells were harvested by centrifugation at 4 °C and 4700*g* for 45 min. The supernatant was then centrifuged again for 15 min and 15,000*g* at 4 °C. Clear supernatant was filtered using a 0.45 μM membrane and DTT was added to a final concentration of 1 mM, pH was adjusted to 6.0 using 25% ammonia solution. To isolate *Sm*CB1 from the medium, hydrophobic interaction chromatography (HIC) was used. Therefore, sodium chloride was added to the medium until a final concentration of 2 M was reached. For isolation and purification, an ÄKTA start device was used. Medium was loaded on a column packed with phenyl sepharose and several column volumes of wash buffer (25 mM sodium citrate, pH 6.0, 2 M sodium chloride, 2.5 mM DTT, 1 mM EDTA) were applied to remove unspecific bound impurities. Finally, *Sm*CB1 was eluted in a gradient from 0 to 100% elution buffer (25 mM sodium citrate, pH 6.0, 2.5 mM DTT, 1 mM EDTA). In a further purification step, *Sm*CB1 was loaded on a Superdex 16/600 75 pg column equilibrated in elution buffer. Until further use, enzyme was stored at −80 °C after flash freezing in liquid nitrogen.

Autoactivation of *Sm*CB1. Inactive zymogen of *Sm*CB1 was activated by autoactivation as described in the literature.[41] Zymogen was diluted to a final concentration of 2 μM in activation buffer (100 mM sodium acetate pH 5.0, 2.5 mM DTT, 1 mM EDTA, 0.1% PEG 6000), dextran sulfate 500 kDa was added to a final concentration of 10 mg/L to induce autocleavage of *Sm*CB1. Cleavage was performed at 37 °C for 6 h, until no inactive zymogen could be detected by SDS-Page.

Fluorometric Enzyme Assays. *SmCB1.* The assay was modified after Jilkova et al.[16] Assay buffer (100 mM sodium citrate, 2.5 mM DTT, 1 mM EDTA, 0.1% PEG 6000, pH 6.0) was incubated with 1 nM activated *Sm*CB1 at 37 °C for 10 min. Afterward, inhibitor in DMSO or DMSO (negative

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

control) was added followed by 20 μM of the substrate *Z*-Phe-Arg-AMC (Bachem, Bubendorf, Switzerland). The mixture was measured at 37 °C for 10 min (screening) or 30 min (IC_{50}, K_i) on a Tecan Spark (δ excitation: 365 nm, δ emission: 460 nm; Tecan Group, Männedorf, Switzerland) plate reader using black *Greiner Bio-One Chimney* 96-well microtiter plates (Greiner Bio-One GmbH, Frickenhausen, Germany). Inhibitor screening concentrations started at 50 μM, followed by 20 μM, 1 μM, 200 nM, and 50 nM. IC_{50} or K_i values were determined for compounds with >50% inhibition at 20 μM.

K_M Determination. The assay was performed as described above using various substrate concentrations (1.25−100 μM). GraFit (version 5.0.13, 2006, Erithracus Software Ltd., U.K.)[42] was used for data analysis and nonlinear regression. The K_M value was calculated as described by Michaelis−Menten

$$v_0 = \frac{v_{max}\cdot[S]}{K_M + [S]}$$

where v_0 is the initial velocity; v_{max} is the maximal velocity; and [S] is the substrate concentration.

IC_{50} and K_i Calculations. We used GraFit (version 5.0.13, 2006, Erithracus Software Ltd., U.K.) for data analysis and nonlinear regression. For compounds without a time-dependent mode of inhibition ((fluorinated) vinylsulfones vs *Sm*CB1, CatB, CatL; vinylsulfonates vs CatB, CatL), we plotted residual enzyme activity in % against the inhibitor concentration in μM, obtaining IC_{50} values by nonlinear regression

$$v_i = \frac{v_o}{1 + \left(\frac{[I]}{IC_{50}}\right)^S}$$

where v_0 is the enzyme activity without inhibitor; v_i is the enzyme activity in the presence of inhibitor; [I] is the inhibitor concentration; and *S* is the slope factor.

K_i values were calculated with the Cheng−Prusoff equation to correct IC_{50} values to zero substrate concentration:[43]

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}}$$

*Sm*CB1: [S] = 20 μM, K_M = 4.3 μM. CatB: [S] = 100 μM, K_M = 150 μM. CatL: [S] = 6.25 μM, K_M = 6.5 μM.

For compounds with a time-dependent mode of inhibition (vinylsulfonates vs *Sm*CB1), the K_I values were calculated as previously published for irreversible binders.[18]

$$E + I \underset{k_2}{\overset{k_1}{\rightleftarrows}} [E\text{---}I] \underset{k_i}{\rightarrow} E - I$$

where E is the enzyme; I is the inhibitor, *k* is the reaction constant; [E---I] is the noncovalent enzyme−inhibitor complex; E—I is the covalent enzyme−inhibitor complex.

Pseudo-first-order rate constants k_{obs} were determined for individual inhibitor concentrations according to a continuous method by Tian and Tsou,[44] monitoring the formation in the presence of inhibitor as a function of time and fitted to the following equation

$$[P]_t = [P]_\infty\cdot(1 - e^{-k_{obs}\cdot t}) + off$$

where $[P]_t$ is the product concentration at time *t*; $[P]_\infty$ is the product concentration at infinite time; and off is the offset value.

Followed by plotting the k_{obs} values against the inhibitor concentrations [I] and fitting to a hyperbolic equation:

$$k_{obs} = \frac{k_i\cdot[I]}{K_I^{app} + [I]}$$

The resulting dissociation constant of the initial enzyme−inhibitor complex K_I^{app} was then corrected to zero substrate concentration with the Cheng−Prusoff relationship, giving the K_I value.[43] The first-order rate constant k_i was further used to calculate second-order rate constants (k_{second}).

$$k_{2nd} = \frac{k_i}{K_I}$$

Selectivity toward Human Cathepsins. Cathepsin B (CatB, Sigma-Aldrich, Darmstadt, Germany) and cathepsin L (CatL, Sigma-Aldrich, Darmstadt, Germany) were incubated in enzyme buffer (50 mM tris-HCl, 5 mM EDTA, 200 mM NaCl, 2 mM DTT, pH 6.5) at room temperature for 20−30 min. Assay buffer (50 mM tris-HCl, 5 mM EDTA, 200 mM NaCl, 0.005% Brij35, pH 6.5) was mixed with cathepsin B or L in enzyme buffer, then inhibitor in DMSO or DMSO (negative control) was added, followed by 100 μM (CatB) or 6.25 μM (CatL) substrate *Z*-Phe-Arg-AMC (Bachem, Bubendorf, Switzerland). Measurements were performed on a Tecan Spark (δ excitation: 365 nm, δ emission: 460 nm; TecanGroup, Männedorf, Switzerland) plate reader on black *Greiner Bio-One Chimney* 96-well microtiter plates (Greiner Bio-One GmbH, Frickenhausen, Germany). Inhibitor screening concentrations started at 20 μM, followed by 1 μM, 200 nM, and 50 nM. K_i values were determined for inhibitors with >50% inhibition at 20 μM.[18]

In Vitro Studies on *S. mansoni*. In vitro and in vivo studies investigating activity against *Schistosoma spp.* were carried out at the Swiss Tropical and Public Health Institute following Swiss national and cantonal regulations on animal welfare (Basel, Switzerland; permission no. 520). Female mice (NMRI strain; age 3 weeks; weight ca. 20−22 g) were purchased from Charles River (Sulzfeld, Germany) and infected by subcutaneously (*s.c.*) injecting approximately 100 *S. mansoni* (Liberian strain, respectively) cercariae. Mice were maintained in polycarbonate cages under environmentally controlled conditions (temperature ∼25 °C; humidity ∼70%; 12 h light and 12 h dark cycle) with free access to water and food.

M199 culture medium and RPMI 1640 culture medium (Gibco, Waltham MA) were used for assays of *S. mansoni* newly transformed schistosomula (NTS) and adult *S. mansoni*, respectively. Both media were supplemented with 1% penicillin (10,000 U/ml) and streptomycin (10 mg/mL) solution (Bioconcept AG, Allschwil, Switzerland) and 5% v/v inactivated fetal calf serum (iFCS, Bioconcept AG, Allschwil, Switzerland).

Cercariae of *S. mansoni* were obtained from infected intermediate host snails (*Biomphalaria glabrata*) and mechanically transformed to NTS as described previously.[23]

Adult *Schistosoma spp.* of both sexes were collected from infected mice, washed with phosphate-buffered saline (PBS, pH 7.4, Sigma-Aldrich, Buchs, Switzerland), and maintained in RPMI 1640 culture medium at 37 °C in an atmosphere of 5% CO_2 until use.

Approximately 50 NTS or at least three adult schistosomes of both sexes, respectively, were incubated in culture medium

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

with the test compounds at 37 °C, 5% CO_2 for 3 days in 96-well plates (cat. no. 83.3924, Sarstedt, Nümbrecht, Germany) Phenotypes were evaluated by microscopic readout (Carl Zeiss, Germany, magnification 10−40×) using a viability scale scoring death, changes in motility, viability, and morphological alterations.[23]

Molecular Docking. *Noncovalent Approach.* The protein crystal structure of *Sm*CB1 with reference ligand **K11777** (3s3r) was loaded into PyMOL and water molecules were removed.[16,31] Molecular modeling was performed using MOE (version 2019.01)[26] with energy minimizing using an MMFF94× force field.[45] Docking calculations were performed using LeadIT (version 2.3.2)[30] from BioSolveIT GmbH and scores were determined using a FlexX and Hyde algorithm.[33] The binding site was defined in a 6.5 Å radius around the reference ligand. Redocking results and score tables are shown below.

Covalent Approach. Covalent docking was performed with MOE. The receptor was prepared using the three-dimensional (3D) protonation tool inside MOE. For the covalent reaction of the warhead, the already existing template reaction was used. Initial 30 poses from the triangle match placement with London ΔG scoring were rescored using the Affinity ΔG scoring function and induced fit refinement implemented in MOE. Ten poses were kept and visually inspected for binding geometry the interactions matching between the docked inhibitor pose and cocrystallized ligand with the enzyme. The poses best matching inspected interaction patterns are further discussed.

Permeability. Permeability was assessed using the parallel artificial membrane permeability assay (PAMPA), originally developed for the purpose of predicting oral absorption of drug candidates.[46] For this, an incubation setup, and a measurement setup both relying on the 96-well format were utilized. The incubation setup consisted of a donor (top) plate (Sigma-Aldrich, MultiScreen IP Filter Plate, 0.45 μm, clear, nonsterile, MAIPNTR10), and an acceptor (bottom) plate (Greiner, white Polystyrene, 655074). The detection setup comprised an ultraviolet (UV)-transparent measurement plate (Greiner UV-Star, 655801) and a Tecan Spark well plate reader. An alternative detection setup comprised a high-performance liquid chromatography−mass spectrometry (HPLC-MS) system as described under the General, Instruments and Software, and Syntheses sections in the SI.

Donor solutions of compounds of interest were made from stock solutions in DMSO or acetonitrile, which were diluted to 100 or 400 μM into Dulbecco's phosphate-buffered saline (DPBS, Sigma D5652, in MQ water; pH = 7.4) or TRIS-HCl buffer (pH = 7.4), with final 5 or 50% (v/v) DMSO or 5% (v/v) acetonitrile, depending on the respective compound's detectability and solubility. For each experiment, 5 μL of artificial membrane (1% (w/v) L-α-phosphatidylcholin, Sigma-Aldrich P3556, in *n*-dodecane, Sigma-Aldrich 8205430100) was first applied to the donor plate membrane, then 150 μL of donor solution was added. This was sealed to avoid evaporation (Greiner, 676070, Viewseal sealer, clear). To the corresponding acceptor well, 400 μL of acceptor solution (DPBS supplemented with the same solvent content as in the respective donor solution) was applied. After assembly of the top and bottom plates, the setup was left to incubate at room temperature without shaking for 5−7 h. To create a reference solution, donor and acceptor solutions were simply mixed in the ratio mentioned above to emulate full equilibration.

After the incubation period, 200 μL of acceptor solution and reference solution were each transferred to the measurement plate and absorption was measured in 5 nm steps over a range of 200−600 nm wavelength. Spectrum-specific wavelength ranges were used for each compound to determine the "area under curve" (AUC) for acceptor and reference solutions. DPBS supplemented with appropriate solvent was used to determine the lower-wavelength limit for the calculations and to baseline-correct the sample spectra. Analysis was performed using GraphPad Prism with the "area under curve" (AUC) function. The baseline for absorption was set to the lowest absorption measured in the range of wavelengths in each experiment. Alternatively, acceptor and reference solutions were analyzed with the same liquid chromatography−mass spectrometry (LC-MS) method and the compound peak areas detected at 210 nm were used for quantification.

Calculation of apparent permeability was performed using the following equation, in which V_D and V_A represent volumes of donor and acceptor solutions (0.15, and 0.4 cm^3), respectively; AUC_A and AUC_{Eq} represent the areas of the measured and baseline-corrected spectrum of sample and reference solution, respectively; A corresponds to the porosity-corrected filter area (0.3019 × 0.7 cm^2 = 0.2113 cm^2); and t indicates the incubation time given in seconds.

$$P_{app} = -\frac{V_D \times V_A \times \ln\left(1 - \frac{AUC_A}{AUC_{Eq}}\right)}{(V_D + V_A) \times A \times t}$$

For evaluation, all compounds with $P_{app} \geq 2.0 \times 10^{-6}$ cm/s were deemed as well permeable and values $<2.0 \times 10^{-6}$ cm/s as poorly permeable.[37,38]

In all performed experiments, propranolol hydrochloride (Fagron, 700556−0004) was used as a sufficiently permeable control substance ($P_{app,Lit}$ = 8.3−13.9 × 10^{-6} cm/s vs P_{app} = 11.0 ± 3.3 × 10^{-6} cm/s, n = 47 in all used PAMPA conditions).[47,48]

PZQ was dissolved in ACN and assayed at 100 or 400 μM donor concentration for 5 or 7 h in DPBS with 5% ACN. **2h** was dissolved in DMSO and assayed at 400 μM donor concentration for 5 h in (DPBS) with 50% DMSO. **2n** was dissolved in DMSO and assayed at 100 μM donor concentration for 7 h in DPBS with 5% DMSO. **2i** was dissolved in DMSO and assayed at 400 μM donor concentration for 5 h in DPBS with 50% DMSO. **2g** was dissolved in DMSO and assayed at 400 μM donor concentration for 5 h in DPBS with 50% DMSO. All aforementioned compounds were quantified using the well plate-based detection method, while K11777 was assayed at 100 μM donor concentration for 7 h in TRIS buffer with 5% DMSO and was quantified using the LC-MS-based method.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsinfecdis.4c00020.

Syntheses, analytical data; NMR spectra and chromatograms; K_M-determination enzyme assay results; in-house screening, correlations; correlation analysis; and molecular docking (PDF)

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

AUTHOR INFORMATION

Corresponding Author

Tanja Schirmeister − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; Email: schirmei@uni-mainz.de

Authors

Natalie Fuchs − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; Present Address: Department of Radiology, Molecular Imaging Innovations Institute (MI3), Weill Cornell Medicine, New York, New York 10065, United States; orcid.org/0000-0001-6404-676X

Robert A. Zimmermann − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; orcid.org/0000-0002-5330-9234

Marvin Schwickert − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; orcid.org/0000-0002-1385-1416

Annika Gunkel − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*

Collin Zimmer − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*

Mergim Meta − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; orcid.org/0000-0003-3601-9523

Kevin Schwickert − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*

Jennifer Keiser − *Swiss Tropical and Public Health Institute, 4123 Allschwil, Switzerland*

Cécile Haeberli − *Swiss Tropical and Public Health Institute, 4123 Allschwil, Switzerland*

Werner Kiefer − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsinfecdis.4c00020

Author Contributions

§N.F., R.A.Z., and M.S. contributed equally to this work. N.F.: design of compounds 2a−q, SmCB1 inhibition testing of compounds 2a−q, correlation analyses, selectivity testing, writing corresponding texts in manuscript and SI; R.A.Z.: SmCB1 expression, purification and autoactivation, SmCB1 inhibition testing of compounds 1a−m, writing corresponding texts in manuscript and SI; M.S.: design and synthesis of compounds 1a−m, writing corresponding texts in manuscript and SI; A.G.: synthesis of compounds 2a−q; C.Z.: PAMPA permeability study; M.M.: covalent docking; K.S.: synthesis of KS-compounds; J.K. and C.H.: *S. mansoni* drug assay; W.K.: conception and consulting; T.S.: conception, editing.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge the support of Sabrina Wagner, Claudia Lenz, and Moritz Weber (all from Johannes Gutenberg University, Mainz) during this project. Furthermore, they thank Dr. Thomas Spangenberg and Dr. Sven Lindemann (EMD, Darmstadt) for the fruitful discussions on this topic.

ABBREVIATIONS

Boc, *tert*-butyl carbamoyl; CatB, human cathepsin B; CatL, human cathepsin L; CH, cyclohexane; cpd, compound; DHBD, 2,3-dihydrobenzo[*b*][1,4]dioxine; EA, ethyl acetate; HIC, hydrophobic interaction chromatography; HWE, Horner−Wadsworth−Emmons; Me, methyl; MOE, molecular operating environment; NTS, newly transformed schistosomula; PAMPA, parallel artificial membrane assay; ppm, parts per million; SI, selectivity index; *Sm*CB1, *Schistosoma mansoni* cathepsin B1; SD, standard deviation; TBTU, 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate; TPSA, topological polar surface area

REFERENCES

(1) James, S. L.; Abate, D.; Abate, K. H.; Abay, S. M.; Abbafati, C.; Abbasi, N.; Abbastabar, H.; Abd-Allah, F.; Abdela, J.; Abdelalim, A.; Abdollahpour, I.; Abdulkader, R. S.; Abebe, Z.; Abera, S. F.; Abil, O. Z.; Abraha, H. N.; Abu-Raddad, L. J.; Abu-Rmeileh, N. M. E.; Accrombessi, M. M. K.; Acharya, D.; Acharya, P.; Ackerman, I. N.; Adamu, A. A.; Adebayo, O. M.; Adekanmbi, V.; Adetokunboh, O. O.; Adib, M. G.; Adsuar, J. C.; Afanvi, K. A.; Afarideh, M.; Afshin, A.; Agarwal, G.; Agesa, K. M.; Aggarwal, R.; Aghayan, S. A.; Agrawal, S.; Ahmadi, A.; Ahmadi, M.; Ahmadieh, H.; Ahmed, M. B.; Aichour, A. N.; Aichour, I.; Aichour, M. T. E.; Akinyemiju, T.; Akseer, N.; Al-Aly, Z.; Al-Eyadhy, A.; Al-Mekhlafi, H. M.; Al-Raddadi, R. M.; Alahdab, F.; Alam, K.; Alam, T.; Alashi, A.; Alavian, S. M.; Alene, K. A.; Alijanzadeh, M.; Alizadeh-Navaei, R.; Aljunid, S. M.; Alkerwi, A.; Alla, F.; Allebeck, P.; Alouani, M. M. L.; Altirkawi, K.; Alvis-Guzman, N.; Amare, A. T.; Aminde, L. N.; Ammar, W.; Amoako, Y. A.; Anber, N. H.; Andrei, C. L.; Androudi, S.; Animut, M. D.; Anjomshoa, M.; Ansha, M. G.; Antonio, C. A. T.; Anwari, P.; Arabloo, J.; Arauz, A.; Aremu, O.; Ariani, F.; Armoon, B.; Ärnlöv, J.; Arora, A.; Artaman, A.; Aryal, K. K.; Asayesh, H.; Asghar, R. J.; Ataro, Z.; Atre, S. R.; Ausloos, M.; Avila-Burgos, L.; Avokpaho, E. F. G. A.; Awasthi, A.; Ayala Quintanilla, B. P.; Ayer, R.; Azzopardi, P. S.; Babazadeh, A.; Badali, H.; Badawi, A.; Bali, A. G.; Ballesteros, K. E.; Ballew, S. H.; Banach, M.; Banoub, J. A. M.; Banstola, A.; Barac, A.; Barboza, M. A.; Barker-Collo, S. L.; Bärnighausen, T. W.; Barrero, L. H.; Baune, B. T.; Bazargan-Hejazi, S.; Bedi, N.; Beghi, E.; Behzadifar, M.; Behzadifar, M.; Béjot, Y.; Belachew, A. B.; Belay, Y. A.; Bell, M. L.; Bello, A. K.; Bensenor, I. M.; Bernabe, E.; Bernstein, R. S.; Beuran, M.; Beyranvand, T.; Bhala, N.; Bhattarai, S.; Bhaumik, S.; Bhutta, Z. A.; Biadgo, B.; Bijani, A.; Bikbov, B.; Bilano, V.; Bililign, N.; Bin Sayeed, M. S.; Bisanzio, D.; Blacker, B. F.; Blyth, F. M.; Bou-Orm, I. R.; Boufous, S.; Bourne, R.; Brady, O. J.; Brainin, M.; Brant, L. C.; Brazinova, A.; Breitborde, N. J. K.; Brenner, H.; Briant, P. S.; Briggs, A. M.; Briko, A. N.; Britton, G.; Brugha, T.; Buchbinder, R.; Busse, R.; Butt, Z. A.; Cahuana-Hurtado, L.; Cano, J.; Cárdenas, R.; Carrero, J. J.; Carter, A.; Carvalho, F.; Castañeda-Orjuela, C. A.; Castillo Rivas, J.; Castro, F.; Catalá-López, F.; Cercy, K. M.; Cerin, E.; Chaiah, Y.; Chang, A. R.; Chang, H.-Y.; Chang, J.-C.; Charlson, F. J.; Chattopadhyay, A.; Chattu, V. K.; Chaturvedi, P.; Chiang, P. P.-C.; Chin, K. L.; Chitheer, A.; Choi, J.-Y. J.; Chowdhury, R.; Christensen, H.; Christopher, D. J.; Cicuttini, F. M.; Ciobanu, L. G.; Cirillo, M.; Claro, R. M.; Collado-Mateo, D.; Cooper, C.; Coresh, J.; Cortesi, P. A.; Cortinovis, M.; Costa, M.; Cousin, E.; Criqui, M. H.; Cromwell, E. A.; Cross, M.; Crump, J. A.; Dadi, A. F.; Dandona, L.; Dandona, R.; Dargan, P. I.; Daryani, A.; Das Gupta, R.; Das Neves, J.; Dasa, T. T.; Davey, G.; Davis, A. C.; Davitoiu, D. V.; De Courten, B.; De La Hoz, F. P.; De Leo, D.; De Neve, J.-W.; Degefa, M. G.; Degenhardt, L.; Deiparine, S.; Dellavalle, R. P.; Demoz, G. T.; Deribe, K.; Dervenis, N.; Des Jarlais, D. C.; Dessie, G. A.; Dey, S.; Dharmaratne, S. D.;

Dinberu, M. T.; Dirac, M. A.; Djalalinia, S.; Doan, L.; Dokova, K.; Doku, D. T.; Dorsey, E. R.; Doyle, K. E.; Driscoll, T. R.; Dubey, M.; Dubljanin, E.; Duken, E. E.; Duncan, B. B.; Duraes, A. R.; Ebrahimi, H.; Ebrahimpour, S.; Echko, M. M.; Edvardsson, D.; Effiong, A.; Ehrlich, J. R.; El Bcheraoui, C.; El Sayed Zaki, M.; El-Khatib, Z.; Elkout, H.; Elyazar, I. R. F.; Enayati, A.; Endries, A. Y.; Er, B.; Erskine, H. E.; Eshrati, B.; Eskandarieh, S.; Esteghamati, A.; Esteghamati, S.; Fakhim, H.; Fallah Omrani, V.; Faramarzi, M.; Fareed, M.; Farhadi, F.; Farid, T. A.; Farinha, C. S. E. sá.; Farioli, A.; Faro, A.; Farvid, M. S.; Farzadfar, F.; Feigin, V. L.; Fentahun, N.; Fereshtehnejad, S.-M.; Fernandes, E.; Fernandes, J. C.; Ferrari, A. J.; Feyissa, G. T.; Filip, I.; Fischer, F.; Fitzmaurice, C.; Foigt, N. A.; Foreman, K. J.; Fox, J.; Frank, T. D.; Fukumoto, T.; Fullman, N.; Fürst, T.; Furtado, J. M.; Futran, N. D.; Gall, S.; Ganji, M.; Gankpe, F. G.; Garcia-Basteiro, A. L.; Gardner, W. M.; Gebre, A. K.; Gebremedhin, A. T.; Gebremichael, T. G.; Gelano, T. F.; Geleijnse, J. M.; Genova-Maleras, R.; Geramo, Y. C. D.; Gething, P. W.; Gezae, K. E.; Ghadiri, K.; Ghasemi Falavarjani, K.; Ghasemi-Kasman, M.; Ghimire, M.; Ghosh, R.; Ghoshal, A. G.; Giampaoli, S.; Gill, P. S.; Gill, T. K.; Ginawi, I. A.; Giussani, G.; Gnedovskaya, E. V.; Goldberg, E. M.; Goli, S.; Gómez-Dantés, H.; Gona, P. N.; Gopalani, S. V.; Gorman, T. M.; Goulart, A. C.; Goulart, B. N. G.; Grada, A.; Grams, M. E.; Grosso, G.; Gugnani, H. C.; Guo, Y.; Gupta, P. C.; Gupta, R.; Gupta, R.; Gupta, T.; Gyawali, B.; Haagsma, J. A.; Hachinski, V.; Hafezi-Nejad, N.; Haghparast Bidgoli, H.; Hagos, T. B.; Hailu, G. B.; Haj-Mirzaian, A.; Haj-Mirzaian, A.; Hamadeh, R. R.; Hamidi, S.; Handal, A. J.; Hankey, G. J.; Hao, Y.; Harb, H. L.; Harikrishnan, S.; Haro, J. M.; Hasan, M.; Hassankhani, H.; Hassen, H. Y.; Havmoeller, R.; Hawley, C. N.; Hay, R. J.; Hay, S. I.; Hedayatizadeh-Omran, A.; Heibati, B.; Hendrie, D.; Henok, A.; Herteliu, C.; Heydarpour, S.; Hibstu, D. T.; Hoang, H. T.; Hoek, H. W.; Hoffman, H. J.; Hole, M. K.; Homaie Rad, E.; Hoogar, P.; Hosgood, H. D.; Hosseini, S. M.; Hosseinzadeh, M.; Hostiuc, M.; Hostiuc, S.; Hotez, P. J.; Hoy, D. G.; Hsairi, M.; Htet, A. S.; Hu, G.; Huang, J. J.; Huynh, C. K.; Iburg, K. M.; Ikeda, C. T.; Ileanu, B.; Ilesanmi, O. S.; Iqbal, U.; Irvani, S. S. N.; Irvine, C. M. S.; Islam, S. M. S.; Islami, F.; Jacobsen, K. H.; Jahangiry, L.; Jahanmehr, N.; Jain, S. K.; Jakovljevic, M.; Javanbakht, M.; Jayatilleke, A. U.; Jeemon, P.; Jha, R. P.; Jha, V.; Ji, J. S.; Johnson, C. O.; Jonas, J. B.; Jozwiak, J. J.; Jungari, S. B.; Jürisson, M.; Kabir, Z.; Kadel, R.; Kahsay, A.; Kalani, R.; Kanchan, T.; Karami, M.; Karami Matin, B.; Karch, A.; Karema, C.; Karimi, N.; Karimi, S. M.; Kasaeian, A.; Kassa, D. H.; Kassa, G. M.; Kassa, T. D.; Kassebaum, N. J.; Katikireddi, S. V.; Kawakami, N.; Karyani, A. K.; Keighobadi, M. M.; Keiyoro, P. N.; Kemmer, L.; Kemp, G. R.; Kengne, A. P.; Keren, A.; Khader, Y. S.; Khafaei, B.; Khafaie, M. A.; Khajavi, A.; Khalil, I. A.; Khan, E. A.; Khan, M. S.; Khan, M. A.; Khang, Y.-H.; Khazaei, M.; Khoja, A. T.; Khosravi, A.; Khosravi, M. H.; Kiadaliri, A. A.; Kiirithio, D. N.; Kim, C.-I.; Kim, D.; Kim, P.; Kim, Y.-E.; Kim, Y. J.; Kimokoti, R. W.; Kinfu, Y.; Kisa, A.; Kissimova-Skarbek, K.; Kivimäki, M.; Knudsen, A. K. S.; Kocarnik, J. M.; Kochhar, S.; Kokubo, Y.; Kolola, T.; Kopec, J. A.; Kosen, S.; Kotsakis, G. A.; Koul, P. A.; Koyanagi, A.; Kravchenko, M. A.; Krishan, K.; Krohn, K. J.; Kuate Defo, B.; Kucuk Bicer, B.; Kumar, G. A.; Kumar, M.; Kyu, H. H.; Lad, D. P.; Lad, S. D.; Lafranconi, A.; Lalloo, R.; Lallukka, T.; Lami, F. H.; Lansingh, V. C.; Latifi, A.; Lau, K. M.-M.; Lazarus, J. V.; Leasher, J. L.; Ledesma, J. R.; Lee, P. H.; Leigh, J.; Leung, J.; Levi, M.; Lewycka, S.; Li, S.; Li, Y.; Liao, Y.; Liben, M. L.; Lim, L.-L.; Lim, S. S.; Liu, S.; Lodha, R.; Looker, K. J.; Lopez, A. D.; Lorkowski, S.; Lotufo, P. A.; Low, N.; Lozano, R.; Lucas, T. C. D.; Lucchesi, L. R.; Lunevicius, R.; Lyons, R. A.; Ma, S.; Macarayan, E. R. K.; Mackay, M. T.; Madotto, F.; Magdy Abd El Razek, H.; Magdy Abd El Razek, M.; Maghavani, D. P.; Mahotra, N. B.; Mai, H. T.; Majdan, M.; Majdzadeh, R.; Majeed, A.; Malekzadeh, R.; Malta, D. C.; Mamun, A. A.; Manda, A.-L.; Manguerra, H.; Manhertz, T.; Mansournia, M. A.; Mantovani, L. G.; Mapoma, C. C.; Maravilla, J. C.; Marcenes, W.; Marks, A.; Martins-Melo, F. R.; Martopullo, I.; März, W.; Marzan, M. B.; Mashamba-Thompson, T. P.; Massenburg, B. B.; Mathur, M. R.; Matsushita, K.; Maulik, P. K.; Mazidi, M.; McAlinden, C.; McGrath, J. J.; McKee, M.; Mehndiratta, M. M.; Mehrotra, R.; Mehta, K. M.; Mehta, V.; Mejia-Rodriguez, F.; Mekonen, T.; Melese, A.; Melku, M.; Meltzer, M.; Memiah, P. T. N.; Memish, Z. A.; Mendoza, W.; Mengistu, D. T.; Mengistu, G.; Mensah, G. A.; Mereta, S. T.; Meretoja, A.; Meretoja, T. J.; Mestrovic, T.; Mezerji, N. M. G.; Miazgowski, B.; Miazgowski, T.; Millear, A. I.; Miller, T. R.; Miltz, B.; Mini, G. K.; Mirarefin, M.; Mirrakhimov, E. M.; Misganaw, A. T.; Mitchell, P. B.; Mitiku, H.; Moazen, B.; Mohajer, B.; Mohammad, K. A.; Mohammadifard, N.; Mohammadnia-Afrouzi, M.; Mohammed, M. A.; Mohammed, S.; Mohebi, F.; Moitra, M.; Mokdad, A. H.; Molokhia, M.; Monasta, L.; Moodley, Y.; Moosazadeh, M.; Moradi, G.; Moradi-Lakeh, M.; Moradinazar, M.; Moraga, P.; Morawska, L.; Moreno Velásquez, I.; Morgado-Da-Costa, J.; Morrison, S. D.; Moschos, M. M.; Mountjoy-Venning, W. C.; Mousavi, S. M.; Mruts, K. B.; Muche, A. A.; Muchie, K. F.; Mueller, U. O.; Muhammed, O. S.; Mukhopadhyay, S.; Muller, K.; Mumford, J. E.; Murhekar, M.; Musa, J.; Musa, K. I.; Mustafa, G.; Nabhan, A. F.; Nagata, C.; Naghavi, M.; Naheed, A.; Nahvijou, A.; Naik, G.; Naik, N.; Najafi, F.; Naldi, L.; Nam, H. S.; Nangia, V.; Nansseu, J. R.; Nascimento, B. R.; Natarajan, G.; Neamati, N.; Negoi, I.; Negoi, R. I.; Neupane, S.; Newton, C. R. J.; Ngunjiri, J. W.; Nguyen, A. Q.; Nguyen, H. T.; Nguyen, H. L. T.; Nguyen, H. T.; Nguyen, L. H.; Nguyen, M.; Nguyen, N. B.; Nguyen, S. H.; Nichols, E.; Ningrum, D. N. A.; Nixon, M. R.; Nolutshungu, N.; Nomura, S.; Norheim, O. F.; Noroozi, M.; Norrving, B.; Noubiap, J. J.; Nouri, H. R.; Nourollahpour Shiadeh, M.; Nowroozi, M. R.; Nsoesie, E. O.; Nyasulu, P. S.; Odell, C. M.; Ofori-Asenso, R.; Ogbo, F. A.; Oh, I.-H.; Oladimeji, O.; Olagunju, A. T.; Olagunju, T. O.; Olivares, P. R.; Olsen, H. E.; Olusanya, B. O.; Ong, K. L.; Ong, S. K.; Oren, E.; Ortiz, A.; Ota, E.; Otstavnov, S. S.; Øverland, S.; Owolabi, M. O.; P A, M.; Pacella, R.; Pakpour, A. H.; Pana, A.; Panda-Jonas, S.; Parisi, A.; Park, E.-K.; Parry, C. D. H.; Patel, S.; Pati, S.; Patil, S. T.; Patle, A.; Patton, G. C.; Paturi, V. R.; Paulson, K. R.; Pearce, N.; Pereira, D. M.; Perico, N.; Pesudovs, K.; Pham, H. Q.; Phillips, M. R.; Pigott, D. M.; Pillay, J. D.; Piradov, M. A.; Pirsaheb, M.; Pishgar, F.; Plana-Ripoll, O.; Plass, D.; Polinder, S.; Popova, S.; Postma, M. J.; Pourshams, A.; Poustchi, H.; Prabhakaran, D.; Prakash, S.; Prakash, V.; Purcell, C. A.; Purwar, M. B.; Qorbani, M.; Quistberg, D. A.; Radfar, A.; Rafay, A.; Rafiei, A.; Rahim, F.; Rahimi, K.; Rahimi-Movaghar, A.; Rahimi-Movaghar, V.; Rahman, M.; Rahman, M. H. ur.; Rahman, M. A.; Rahman, S. U.; Rai, R. K.; Rajati, F.; Ram, U.; Ranjan, P.; Ranta, A.; Rao, P. C.; Rawaf, D. L.; Rawaf, S.; Reddy, K. S.; Reiner, R. C.; Reinig, N.; Reitsma, M. B.; Remuzzi, G.; Renzaho, A. M. N.; Resnikoff, S.; Rezaei, S.; Rezai, M. S.; Ribeiro, A. L. P.; Roberts, N. L. S.; Robinson, S. R.; Roever, L.; Ronfani, L.; Roshandel, G.; Rostami, A.; Roth, G. A.; Roy, A.; Rubagotti, E.; Sachdev, P. S.; Sadat, N.; Saddik, B.; Sadeghi, E.; Saeedi Moghaddam, S.; Safari, H.; Safari, Y.; Safari-Faramani, R.; Safdarian, M.; Safi, S.; Safiri, S.; Sagar, R.; Sahebkar, A.; Sahraian, M. A.; Sajadi, H. S.; Salam, N.; Salama, J. S.; Salamati, P.; Saleem, K.; Saleem, Z.; Salimi, Y.; Salomon, J. A.; Salvi, S. S.; Salz, I.; Samy, A. M.; Sanabria, J.; Sang, Y.; Santomauro, D. F.; Santos, I. S.; Santos, J. V.; Santric Milicevic, M. M.; Sao Jose, B. P.; Sardana, M.; Sarker, A. R.; Sarrafzadegan, N.; Sartorius, B.; Sarvi, S.; Sathian, B.; Satpathy, M.; Sawant, A. R.; Sawhney, M.; Saxena, S.; Saylan, M.; Schaeffner, E.; Schmidt, M. I.; Schneider, I. J. C.; Schöttker, B.; Schwebel, D. C.; Schwendicke, F.; Scott, J. G.; Sekerija, M.; Sepanlou, S. G.; Serván-Mori, E.; Seyedmousavi, S.; Shabaninejad, H.; Shafieesabet, A.; Shahbazi, M.; Shaheen, A. A.; Shaikh, M. A.; Shams-Beyranvand, M.; Shamsi, M.; Shamsizadeh, M.; Sharafi, H.; Sharafi, K.; Sharif, M.; Sharif-Alhoseini, M.; Sharma, M.; Sharma, R.; She, J.; Sheikh, A.; Shi, P.; Shibuya, K.; Shigematsu, M.; Shiri, R.; Shirkoohi, R.; Shishani, K.; Shiue, I.; Shokraneh, F.; Shoman, H.; Shrime, M. G.; Si, S.; Siabani, S.; Siddiqi, T. J.; Sigfusdottir, I. D.; Sigurvinsdottir, R.; Silva, J. P.; Silveira, D. G. A.; Singam, N. S. V.; Singh, J. A.; Singh, N. P.; Singh, V.; Sinha, D. N.; Skiadaresi, E.; Slepak, E. L. N.; Sliwa, K.; Smith, D. L.; Smith, M.; Soares Filho, A. M.; Sobaih, B. H.; Sobhani, S.; Sobngwi, E.; Soneji, S. S.; Soofi, M.; Soosaraei, M.; Sorensen, R. J. D.; Soriano, J. B.; Soyiri, I. N.; Sposato, L. A.; Sreeramareddy, C. T.; Srinivasan, V.; Stanaway, J. D.; Stein, D. J.; Steiner, C.; Steiner, T. J.; Stokes, M. A.; Stovner, L. J.; Subart, M. L.; Sudaryanto, A.; Sufiyan, M. B.; Sunguya, B. F.; Sur, P. J.; Sutradhar, I.; Sykes, B. L.; Sylte, D.

https://doi.org/10.1021/acsinfecdis.4c00020
ACS Infect. Dis. 2024, 10, 1664−1678

O.; Tabarés-Seisdedos, R.; Tadakamadla, S. K.; Tadesse, B. T.; Tandon, N.; Tassew, S. G.; Tavakkoli, M.; Taveira, N.; Taylor, H. R.; Tehrani-Banihashemi, A.; Tekalign, T. G.; Tekelemedhin, S. W.; Tekle, M. G.; Temesgen, H.; Temsah, M.-H.; Temsah, O.; Terkawi, A. S.; Teweldemedhin, M.; Thankappan, K. R.; Thomas, N.; Tilahun, B.; To, Q. G.; Tonelli, M.; Topor-Madry, R.; Topouzis, F.; Torre, A. E.; Tortajada-Girbés, M.; Touvier, M.; Tovani-Palone, M. R.; Towbin, J. A.; Tran, B. X.; Tran, K. B.; Troeger, C. E.; Truelsen, T. C.; Tsilimbaris, M. K.; Tsoi, D.; Tudor Car, L.; Tuzcu, E. M.; Ukwaja, K. N.; Ullah, I.; Undurraga, E. A.; Unutzer, J.; Updike, R. L.; Usman, M. S.; Uthman, O. A.; Vaduganathan, M.; Vaezi, A.; Valdez, P. R.; Varughese, S.; Vasankari, T. J.; Venketasubramanian, N.; Villafaina, S.; Violante, F. S.; Vladimirov, S. K.; Vlassov, V.; Vollset, S. E.; Vosoughi, K.; Vujcic, I. S.; Wagnew, F. S.; Waheed, Y.; Waller, S. G.; Wang, Y.; Wang, Y.-P.; Weiderpass, E.; Weintraub, R. G.; Weiss, D. J.; Weldegebreal, F.; Weldegwergs, K. G.; Werdecker, A.; West, T. E.; Whiteford, H. A.; Widecka, J.; Wijeratne, T.; Wilner, L. B.; Wilson, S.; Winkler, A. S.; Wiyeh, A. B.; Wiysonge, C. S.; Wolfe, C. D. A.; Woolf, A. D.; Wu, S.; Wu, Y.-C.; Wyper, G. M. A.; Xavier, D.; Xu, G.; Yadgir, S.; Yadollahpour, A.; Yahyazadeh Jabbari, S. H.; Yamada, T.; Yan, L. L.; Yano, Y.; Yaseri, M.; Yasin, Y. J.; Yeshaneh, A.; Yimer, E. M.; Yip, P.; Yisma, E.; Yonemoto, N.; Yoon, S.-J.; Yotebieng, M.; Younis, M. Z.; Yousefifard, M.; Yu, C.; Zadnik, V.; Zaidi, Z.; Zaman, S. Bin.; Zamani, M.; Zare, Z.; Zeleke, A. J.; Zenebe, Z. M.; Zhang, K.; Zhao, Z.; Zhou, M.; Zodpey, S.; Zucker, I.; Vos, T.; Murray, C. J. L. Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990−2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* **2018**, *392* (10159), 1789−1858.

(2) Hotez, P. J.; Alvarado, M.; Basáñez, M.-G.; Bolliger, I.; Bourne, R.; Boussinesq, M.; Brooker, S. J.; Brown, A. S.; Buckle, G.; Budke, C. M.; Carabin, H.; Coffeng, L. E.; Fèvre, E. M.; Fürst, T.; Halasa, Y. A.; Jasrasaria, R.; Johns, N. E.; Keiser, J.; King, C. H.; Lozano, R.; Murdoch, M. E.; O'Hanlon, S.; Pion, S. D. S.; Pullan, R. L.; Ramaiah, K. D.; Roberts, T.; Shepard, D. S.; Smith, J. L.; Stolk, W. A.; Undurraga, E. A.; Utzinger, J.; Wang, M.; Murray, C. J. L.; Naghavi, M. The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases. *PLoS Neglected Trop. Dis.* **2014**, *8* (7), No. e2865.

(3) Caldwell, N.; Afshar, R.; Baragaña, B.; Bustinduy, A. L.; Caffrey, C. R.; Collins, J. J.; Fusco, D.; Garba, A.; Gardner, M.; Gomes, M.; Hoffmann, K. F.; Hsieh, M.; Lo, N. C.; McNamara, C. W.; Nono, J. K.; Padalino, G.; Read, K. D.; Roestenberg, M.; Spangenberg, T.; Specht, S.; Gilbert, I. H. Perspective on Schistosomiasis Drug Discovery: Highlights from a Schistosomiasis Drug Discovery Workshop at Wellcome Collection, London, September 2022. *ACS Infect. Dis.* **2023**, *9* (5), 1046−1055.

(4) Colley, D. G.; Bustinduy, A. L.; Secor, W. E.; King, C. H. Human Schistosomiasis. *Lancet* **2014**, *383* (9936), 2253−2264.

(5) Gryseels, B.; Polman, K.; Clerinx, J.; Kestens, L. Human Schistosomiasis. *Lancet* **2006**, *368* (9541), 1106−1118.

(6) Fitzpatrick, J. M.; Johnston, D. A.; Williams, G. W.; Williams, D. J.; Freeman, T. C.; Dunne, D. W.; Hoffmann, K. F. An Oligonucleotide Microarray for Transcriptome Analysis of Schistosoma Mansoni and Its Application/Use to Investigate Gender-Associated Gene Expression. *Mol. Biochem. Parasitol.* **2005**, *141* (1), 1−13.

(7) Sandground, J. H.; Moore, D. V. The Relative Egg Producing Capacity of Schistosoma Mansoni and Schistosoma Japonicum 1,2. *Am. J. Trop. Med. Hyg.* **1956**, *5* (5), 831−840.

(8) Hams, E.; Aviello, G.; Fallon, P. G. The Schistosoma Granuloma: Friend or Foe? *Front. Immunol.* **2013**, *4*, No. 89, DOI: 10.3389/fimmu.2013.00089.

(9) Bueke, M. L.; Jones, M. K.; Gobert, G. N.; Li, Y. S.; Ellis, M. K.; McManus, D. P. Immunopathogenesis of Human Schistosomiasis. *Parasite Immunol.* **2009**, *31* (4), 163−176.

(10) Wang, W.; Wang, L.; Liang, Y.-S. Susceptibility or Resistance of Praziquantel in Human Schistosomiasis: A Review. *Parasitol. Res.* **2012**, *111* (5), 1871−1877.

(11) Park, S.-K.; Friedrich, L.; Yahya, N. A.; Rohr, C. M.; Chulkov, E. G.; Maillard, D.; Rippmann, F.; Spangenberg, T.; Marchant, J. S. Mechanism of Praziquantel Action at a Parasitic Flatworm Ion Channel. *Sci. Transl. Med.* **2021**, *13* (625), No. eabj5832, DOI: 10.1126/scitranslmed.abj5832.

(12) Jílková, A.; Horn, M.; Fanfrlík, J.; Küppers, J.; Pachl, P.; Řezáčová, P.; Lepšík, M.; Fajtová, P.; Rubešová, P.; Chanová, M.; Caffrey, C. R.; Gütschow, M.; Mareš, M. Azanitrile Inhibitors of the SmCB1 Protease Target Are Lethal to *Schistosoma Mansoni*: Structural and Mechanistic Insights into Chemotype Reactivity. *ACS Infect. Dis.* **2021**, *7* (1), 189−201.

(13) Jílková, A.; Rubešová, P.; Fanfrlík, J.; Fajtová, P.; Řezáčová, P.; Brynda, J.; Lepšík, M.; Mertlíková-Kaiserová, H.; Emal, C. D.; Renslo, A. R.; Roush, W. R.; Horn, M.; Caffrey, C. R.; Mareš, M. Druggable Hot Spots in the Schistosomiasis Cathepsin B1 Target Identified by Functional and Binding Mode Analysis of Potent Vinyl Sulfone Inhibitors. *ACS Infect. Dis.* **2021**, *7* (5), 1077−1088.

(14) Abdulla, M.-H.; Lim, K.-C.; Sajid, M.; McKerrow, J. H.; Caffrey, C. R. Schistosomiasis Mansoni: Novel Chemotherapy Using a Cysteine Protease Inhibitor. *PLoS Med.* **2007**, *4* (1), No. e14.

(15) Correnti, J. M.; Brindley, P. J.; Pearce, E. J. Long-Term Suppression of Cathepsin B Levels by RNA Interference Retards Schistosome Growth. *Mol. Biochem. Parasitol.* **2005**, *143* (2), 209−215.

(16) Jílková, A.; Řezáčová, P.; Lepšík, M.; Horn, M.; Váchová, J.; Fanfrlík, J.; Brynda, J.; McKerrow, J. H.; Caffrey, C. R.; Mareš, M. Structural Basis for Inhibition of Cathepsin B Drug Target from the Human Blood Fluke, Schistosoma Mansoni. *J. Biol. Chem.* **2011**, *286* (41), 35770−35781.

(17) Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T. H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P. R.; Distler, U.; Tenzer, S.; Johé, P.; Hellmich, U. A.; Engels, B. Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138* (27), 8332−8335.

(18) Jung, S.; Fuchs, N.; Johe, P.; Wagner, A.; Diehl, E.; Yuliani, T.; Zimmer, C.; Barthels, F.; Zimmermann, R. A.; Klein, P.; Waigel, W.; Meyr, J.; Opatz, T.; Tenzer, S.; Distler, U.; Räder, H.-J.; Kersten, C.; Engels, B.; Hellmich, U. A.; Klein, J.; Schirmeister, T. Fluorovinyl-sulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure−Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies. *J. Med. Chem.* **2021**, *64* (16), 12322−12358.

(19) Jung, S.; Fuchs, N.; Grathwol, C.; Hellmich, U. A.; Wagner, A.; Diehl, E.; Willmes, T.; Sotriffer, C.; Schirmeister, T. New Peptidomimetic Rhodesain Inhibitors with Improved Selectivity towards Human Cathepsins. *Eur. J. Med. Chem.* **2022**, *238*, No. 114460.

(20) Gauthier, J. Y.; Black, W. C.; Courchesne, I.; Cromlish, W.; Desmarais, S.; Houle, R.; Lamontagne, S.; Li, C. S.; Massé, F.; McKay, D. J.; Ouellet, M.; Robichaud, J.; Truchon, J.-F.; Truong, V.-L.; Wang, Q.; Percival, M. D. The Identification of Potent, Selective, and Bioavailable Cathepsin S Inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *17* (17), 4929−4933.

(21) 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.

(22) 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*, No. 804970, DOI: 10.3389/fmolb.2021.804970.

(23) Lombardo, F. C.; Pasche, V.; Panic, G.; Endriss, Y.; Keiser, J. Life Cycle Maintenance and Drug-Sensitivity Assays for Early Drug Discovery in Schistosoma Mansoni. *Nat. Protoc.* **2019**, *14* (2), 461−481.

(24) 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.; Rodríguez, S.; Diemer, K.; Neumann, W.; Gonzàlez, 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), No. 7226, DOI: 10.3390/ijms24087226.
(25) Pearce, E. J.; MacDonald, A. S. The Immunobiology of Schistosomiasis. *Nat. Rev. Immunol.* **2002**, *2* (7), 499−511.
(26) Molecular Operating Environment (MOE). *Chemical Computing Group ULC* Sherbrooke St. W.: Montreal, QC H3A 2R7, Canada; 2023.
(27) Liu, S.; Cai, P.; Piao, X.; Hou, N.; Zhou, X.; Wu, C.; Wang, H.; Chen, Q. Expression Profile of the Schistosoma Japonicum Degradome Reveals Differential Protease Expression Patterns and Potential Anti-Schistosomal Intervention Targets. *PLoS Comput. Biol.* **2014**, *10* (10), No. e1003856.
(28) Furlong, S. T. Unique Roles for Lipids in Schistosoma Mansoni. *Parasitol. Today* **1991**, *7* (2), 59−62.
(29) Kier, L. B. An Index of Molecular Flexibility from Kappa Shape Attributes. *Quant. Struct.-Act. Relat.* **1989**, *8* (3), 221−224.
(30) LeadIT. BioSolveIT GmbH: Sankt Augustin, Germany 2017. Www.Biosolveit.de/LeadIT.
(31) Schrödinger, L.; DeLano, W. *The PyMOL Molecular Graphics System*, Schrödinger, LLC, 2020. Http://Www.Pymol.Org/Pymol.
(32) Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. Towards an Integrated Description of Hydrogen Bonding and Dehydration: Decreasing False Positives in Virtual Screening with the HYDE Scoring Function. *ChemMedChem* **2008**, *3* (6), 885−897.
(33) Gehringer, M.; Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62* (12), 5673−5724.
(34) Meister, I.; Ingram-Sieber, K.; Cowan, N.; Todd, M.; Robertson, M. N.; Meli, C.; Patra, M.; Gasser, G.; Keiser, J. Activity of Praziquantel Enantiomers and Main Metabolites against Schistosoma Mansoni. *Antimicrob. Agents Chemother.* **2014**, *58* (9), 5466−5472.
(35) Meister, I.; Ingram-Sieber, K.; Cowan, N.; Todd, M.; Robertson, M. N.; Meli, C.; Patra, M.; Gasser, G.; Keiser, J. Activity of Praziquantel Enantiomers and Main Metabolites against Schistosoma Mansoni. *Antimicrob. Agents Chemother.* **2014**, *58* (9), 5466−5472.
(36) Di, L.; Artursson, P.; Avdeef, A.; Benet, L. Z.; Houston, J. B.; Kansy, M.; Kerns, E. H.; Lennernäs, H.; Smith, D. A.; Sugano, K. The Critical Role of Passive Permeability in Designing Successful Drugs. *ChemMedChem* **2020**, *15* (20), 1862−1874.
(37) Avdeef, A.; Bendels, S.; Di, Li.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. PAMPA—Critical Factors for Better Predictions of Absorption. *J. Pharm. Sci.* **2007**, *96* (11), 2893−2909.
(38) Cowan, N.; Dätwyler, P.; Ernst, B.; Wang, C.; Vennerstrom, J. L.; Spangenberg, T.; Keiser, J. Activities of *N, N′*-Diarylurea MMV665852 Analogs against Schistosoma Mansoni. *Antimicrob. Agents Chemother.* **2015**, *59* (4), 1935−1941.
(39) Dinora, G.-E.; Julio, R.; Nelly, C.; Lilian, Y.-M.; Cook, H. J. In Vitro Characterization of Some Biopharmaceutical Properties of Praziquantel. *Int. J. Pharm.* **2005**, *295* (1−2), 93−99.
(40) Jílková, A.; Horn, M.; Mareš, M.Structural and Functional Characterization of Schistosoma Mansoni Cathepsin B1. In *Methods in Molecular Biology* 2020; Vol. *2151*, pp 145−158.
(41) Jílková, A.; Horn, M.; Řezáčová, P.; Marešová, L.; Fajtová, P.; Brynda, J.; Vondrášek, J.; McKerrow, J. H.; Caffrey, C. R.; Mareš, M. Activation Route of the Schistosoma Mansoni Cathepsin B1 Drug Target: Structural Map with a Glycosaminoglycan Switch. *Structure* **2014**, *22* (12), 1786−1798.
(42) *GraFit*, Erithracus Software Ltd. 2006.
(43) 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.
(44) Tian, W. X.; Tsou, C. L. Determination of the Rate Constant of Enzyme Modification by Measuring the Substrate Reaction in the Presence of the Modifier. *Biochemistry* **1982**, *21* (5), 1028−1032.
(45) Halgren, T. A. MMFF VI. MMFF94s Option for Energy Minimization Studies. *J. Comput. Chem.* **1999**, *20* (7), 720−729.
(46) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41* (7), 1007−1010.
(47) Sugano, K.; Hamada, H.; Machida, M.; Ushio, H. High Throughput Prediction of Oral Absorption: Improvement of the Composition of the Lipid Solution Used in Parallel Artificial Membrane Permeation Assay. *SLAS Discovery* **2001**, *6* (3), 189−196.
(48) Avdeef, A. *Absorption and Drug Development*; Wiley, 2003.

”

6: Structural Modifications of Covalent Cathepsin S Inhibitors: Impact on Affinity, Selectivity, and Permeability.

Inhibition of CatS is a versatile therapeutic option for treating several solid tumors because the targets are dysregulated tumor-associated immune cells, mainly macrophages. A normalization of the immunological tumor microenvironment can reduce tumor growth and metastasis, and increase susceptibility to cytostatic agents. The progress of CatS inhibitors from preclinical models to clinical applications requires a shift in design strategies surpassing improved affinity and selectivity. Both pharmacokinetic characterization as well as options for functionalization to accommodate potential complex application concepts were identified as requisites.

In general, two sets of inhibitors can be distinguished within this study. The first was focused on modifying the peptidic backbone of inhibitors carrying the nitrile warhead. To this end, recognition motifs of different literature drug candidates were combined or expanded upon. The second set focused on a hydrazone functionalization approach that encompassed different carbonyl-based warheads. For all potential inhibitors investigated, affinity towards cathepsin S was characterized with respect to fast or slowly reversible or irreversible inhibition kinetics. Selectivity was assessed for the related human cathepsins B and L. Since permeability can be a challenge for peptidomimetic inhibitors, it was assessed for all synthesized derivates by PAMPA and correlated with calculated physicochemical parameters (logP and PSA) to derive information that could guide further drug design choices. For the nitrile-bearing inhibitors with variable backbones, a correlation between logP and permeability as well as structure-defined cutoffs for PSA that conveyed passive permeability were identified. Of value for unrelated drug design campaigns was the identification of exchangeable amino acids or capping groups that convey quantifiable influences on permeability. Peptidomimetic aryl trifluoroethylamines were identified to result in higher permeability than a carboxamide-based capping group, and differences between three carboxamide capping groups showed a trend for the explored aliphatic options. For the warhead-exchanged structures, a permeability equivalence could be identified between nitrile and aldehyde warheads. A similar observation was made for hPhe and *O*BnSer as P1 amino acids. Apart from this, the performed functionalization of the carbonyl-based warheads elucidated multiple advantages. The hydrazone group was chosen as a capping group due to its pH-sensitive stability, which was shown to enable liberation of the reactive warhead in the more acidic tumor microenvironment, especially on the promising inhibitor **45**. This masking of the high reversible reactivity of the aldehyde warhead is expected to reduce off-target effects. The employed model hydrazones had poor passive permeability, but even this can be exploited in a parenteral formulation because it could enable favored distribution from the vascular system into the tumor tissue under the effect of the enhanced permeability and retention (EPR) effect.[630] Lastly, choosing different hydrazines or hydrazides for the generation of the hydrazone can in the future enable pH-sensitive reversible linking to surface-modified nanoparticles for a targeted therapy approach. Mechanistically, the analysis of the ketone-based warheads with

different possible leaving groups (sulfonylmethyl ketone **44**, chloromethyl ketone **48**, acylthiomethyl ketone **49**, acyloxymethyl ketone **50**, hydroxymethyl ketone **51**) elucidated cases with a dual-reactive, irreversible warhead that surprisingly favor the reversible reaction of the catalytic cysteine with the carbonyl group. This strongly underlined that the leaving group is a decisive factor towards or against irreversibility. Taken together, the peptidic recognition sequences of established CatS inhibitors were modulated in a combinatorial approach, and the impact of these modifications on inhibition and permeability was investigated. The results on permeability were put into the context of physicochemical parameters and structure-permeability guidelines were discovered for further work on ligand-derived protease inhibitors. In the context of a warhead exchange study, a reversible capping approach was implemented and characterized to enable pH-sensitive linker conjugation, and warhead reactivity was investigated in more detail.

Own contribution: Evaluating the passive permeability of all final inhibitors by PAMPA, correlating the passive permeability with physicochemical parameters (logP, PSA), LCMS-based aqueous stability analysis and reactivity assessment towards a model thiol for **44**, writing and editing of the manuscript.

Contribution from others: Inhibitor synthesis and characterization of the in vitro affinity, inhibition kinetics, and selectivity, reversibility assessment by dilution assay, non-covalent and covalent docking, hydrazone liberation assay, writing and editing of the manuscript.

This work was published in ACS Medicinal Chemistry Letters.

Article reprinted with permission from *ACS Med. Chem. Lett.* **2024**, 15, 6, 837–844, “Structural Modifications of Covalent Cathepsin S Inhibitors: Impact on Affinity, Selectivity, and Permeability” Copyright © 2024 American Chemical Society. Published by American Chemical Society.

The following publication quoted within “” (pages 221–235) is the same as the manuscript cited here. The appended Supporting Information represents an abridged version. The full version can be accessed online at doi:10.1021/acsmedchemlett.4c00050.

“

ACS Medicinal Chemistry Letters

pubs.acs.org/acsmedchemlett Letter

Structural Modifications of Covalent Cathepsin S Inhibitors: Impact on Affinity, Selectivity, and Permeability

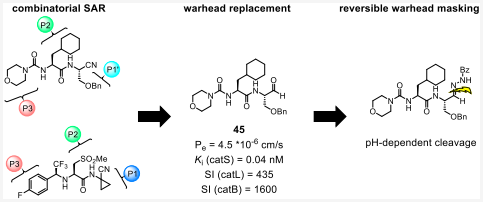
Published as part of ACS Medicinal Chemistry Letters virtual special issue “Exploring Covalent Modulators in Drug Discovery and Chemical Biology”.

Mergim Meta,[‡] Collin Zimmer,[‡] Natalie Fuchs, Maximilian Johannes Zecher, Albin Lahu, and Tanja Schirmeister*

Cite This: ACS Med. Chem. Lett. 2024, 15, 837−844 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Cathepsin S (catS) is a member of the cysteine protease family with limited tissue distribution, which is predominantly found in antigen-presenting cells. Due to overexpression and overactivity of catS in numerous cancers, inhibition of catS is supposed to improve the antitumor response. Here, we explore the potential of small-molecule catS inhibitors emphasizing their in vitro pharmacodynamics and pharmacokinetics. Membrane permeability of selected inhibitors was measured with a Parallel Artificial Membrane Permeation Assay and correlated to calculated physicochemical parameters and inhibition data. The binding kinetics and inhibition types of potent and selective new inhibitors with unexplored warheads were investigated. Our unique approach involves reversible masking of these potent warheads, allowing for further customization without compromising affinity or selectivity. The most promising inhibitors in this study include covalent aldehyde and ketone derivatives reversibly masked as hydrazones as potential candidates for therapeutic interventions targeting catalytic enzymes and modulating the immune response in cancer.



KEYWORDS: *Cathepsin S, covalent inhibition, structure−activity relation, membrane permeability, reversible warheads*

Cysteine cathepsins are essential enzymes of the papain-like family of proteases with several physiological functions, such as extracellular matrix degradation and processing of damaged proteins in the endolysosomal pathway.[1] Cathepsin S (catS) takes a special role, since this protease is stable and catalytically active at neutral pH, and it has a limited tissue distribution (mainly in antigen-presenting cells, e.g. dendritic cells).[2] CatS has emerged as a potential target for several pathological conditions such as Sjögren’s syndrome, psoriasis, and many types of cancer.[3,4] It is overexpressed and overactive in several tumors such as follicular lymphoma, breast cancer, or astrocytoma, and many mechanisms that explain how catS is involved in tumor progression are known.[5,6] CatS is also found extracellularly and is known to degrade extracellular matrix proteins and to drive angiogenesis.[7] It is a major regulator of antigen processing and presentation via the MHC-II pathway in antigen-presenting cells (APCs).[8,9] In a follicular lymphoma model, the overexpression and overactivation of the enzyme led to the shift from the MHC-II to the MHC-I pathway, thus inducing a favored activation of regulatory CD4$^+$ T cells over cytotoxic CD8$^+$ T cells, which ultimately led to a suppression of the antitumor immune response.[5,6,10] One possible strategy to tackle this immune-suppressive milieu of the tumor microenvironment (TME) is the use of small-molecule catS inhibitors, since catS is involved in the polarization of APCs from the M1 phenotype to the M2 phenotype.[3,5,6] The use of inhibitory antibodies, siRNA, and knockout experiments has already shown that the antitumor immune response could be enhanced through the inhibition of catS, by shifting the ratio of CD8$^+$ T cells to CD4$^+$ T cells toward the cytotoxic CD8$^+$ T cells.[5,6] Furthermore, inhibition of several cathepsins with a pan-cathepsin inhibitor led to polarization of tumor-associated macrophages and a shift from M2 macrophages to M1 macrophages, ultimately leading to an increase of pro-inflammatory mediators.[11] To date, several small-molecule inhibitors have been reported, many of them targeting the active site cysteine (Cys-25) in an either reversible or irreversible covalent manner.[12,13] As of today, none of the developed clinical candidates have been approved by the FDA

Received: January 30, 2024
Revised: May 24, 2024
Accepted: May 30, 2024
Published: June 4, 2024

ACS Publications

© 2024 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acsmedchemlett.4c00050
ACS Med. Chem. Lett. 2024, 15, 837−844

Downloaded via UNIV MAINZ on August 30, 2024 at 13:14:33 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

and most of the developed inhibitors were designed for use in autoimmune diseases.[14] There has been an emergency approval for the first nitrile-based cysteine protease-targeting inhibitor nirmatrelvir in combination with the CYP-450 inhibitor ritonavir for the short-term treatment of COVID-19 infections in 2021. This clinical setting is different from the requirements of chronic catS inhibitor administration in autoimmune and oncological diseases. Clinical development of catS inhibitors has not led to approvals so far; nevertheless, the search for covalent modulators of cysteine proteases is on the rise.[15] Figure 1 shows

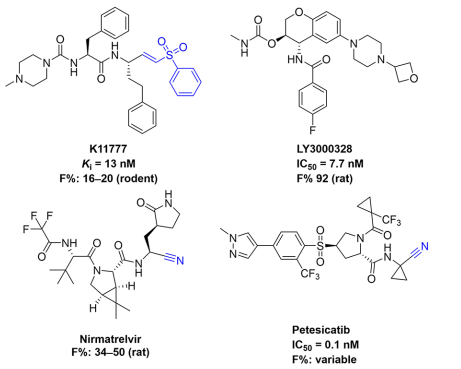


Figure 1. Structures, inhibition data against catS, and bioavailabilities (F%) of developed cathepsin S inhibitors and nirmatrelvir. Warheads are highlighted in blue.[17−19]

the structures of nirmatrelvir and three developed catS inhibitors, consisting of the irreversible pan-cathepsin inhibitor K11777, the noncovalent inhibitor LY3000328, and the covalent-reversible nitrile inhibitor petesicatib.[1,12,16,17]

In this study, we further explored peptidomimetic small-molecule inhibitors starting from nitriles with a focus on their pharmacodynamics (affinity and binding kinetics) and pharmacokinetics (membrane permeability). Additionally, through the exchange of the nitrile warhead with α-substituted ketones or an aldehyde, a set of newly generated inhibitors was synthesized and tested for their affinity and selectivity toward catS (Figure 2). One goal was the development of inhibitors with masked warheads, which can be further decorated with advanced reversible linker systems in the future. So far, in many cases, the developed small-molecule inhibitors have been derivatized with bulky linkers, leading to a dramatic structural change, which can lead to a loss in affinity or selectivity toward the target enzyme.[20] With our approach, highly potent and reactive warheads such as aldehyde or electron-deficient ketones can be masked reversibly as hydrazones, lowering the rate of off-target reactions while retaining the high affinities toward the target protease after controlled release. Two selected inhibitors (an aldehyde and a ketone) were therefore transformed to hydrazones for reversible warhead capping. Subsequently, the ability to release the free inhibitor in an acidic environment was inspected.

We aimed for a combinatorial approach with two catS selective inhibitors as starting points. Through merging and further derivatizing the P1, P2, and P3 substituents, which all have an impact on binding affinity, with aromatic and nonaromatic substituents that have proven to increase affinities toward the target enzyme, we aimed to gain insight into the

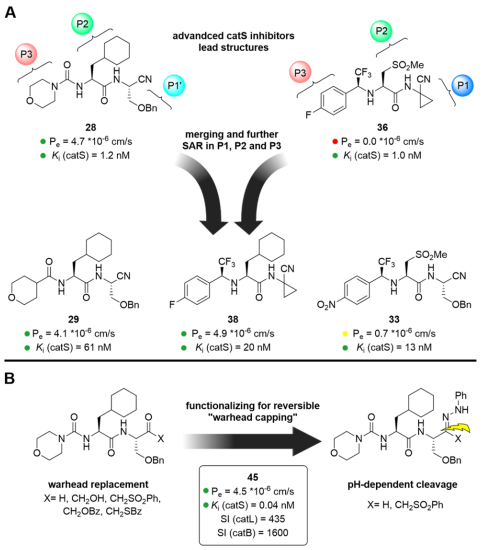


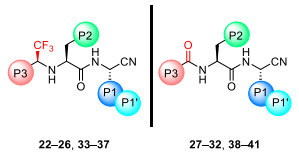
Figure 2. Aim and scope of the present study: (A) combinatorial SAR study through merging of two potent catS inhibitors and further substitutions, and (B) subsequent warhead replacement and reversible capping of the warhead. Created with BioRender.com.

selectivity profiles and membrane permeability of newly generated compounds.[12,13] Therefore, we prepared nitrile-based inhibitors with five different P3 residues frequently encountered in catS inhibitor design: three attached to the peptide as aryl-amides (pyridine-, 2*H*-tetrahydropyran-, and morpholine-4-carboxamide) and two attached as aryl-trifluoroethylamines (TFE) as amide isosters (4-fluoro- and 4-nitrophenyl-TFE). The amide isosters are used with in vivo application in mind, since they offer increased metabolic stability.[12] In P2, we chose L-cyclohexyl alanine (Cha) and a methylated and oxidized cysteine, since these residues in P2 led to the most potent and selective inhibitors in previous studies.[12,13,21] To gain information regarding whether a residue addressing the S1 subpocket or a substituent designed for occupying the S1′ site would be more beneficial, we prepared our inhibitors with three different P1 amino acids containing the nitrile warhead, namely, *O*-benzyl-L-serine (*O*BnSer), L-homophenylalanine (hPhe), and 1-aminocyclopropane-1-carboxylic acid (Acc). Although some of the synthesized inhibitors were already tested against catS in the past, we present them here again with novel data on off-target affinities, as well as membrane permeability.[12,13] Replacement of the nitrile warhead of one of the starting structures with electron-deficient ketones and an aldehyde led to new catS inhibitors from which two could be capped as hydrazones. In the future, this strategy opens the possibility of reversibly attaching the potential immunomodulators on different carrier systems. The pH-dependent release of the free inhibitor from hydrazone was explored for two selected inhibitors via HPLC/MS.

The detailed synthesis procedure of all final inhibitors with Schemes S1−S3 can be accessed in the Supporting Information (SI).

https://doi.org/10.1021/acsmedchemlett.4c00050
ACS Med. Chem. Lett. 2024, 15, 837−844

Table 1A. Inhibition Data for Nitrile-Based Inhibitors 22−41[a]

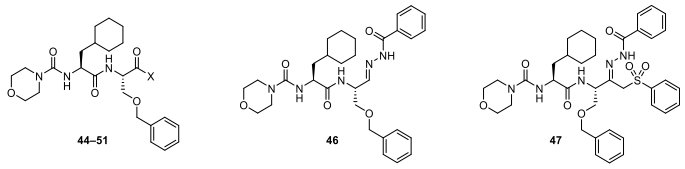


22–26, 33–37 27–32, 38–41

compound	P3	P2	P1/P1′	K_i [nM] catS[b]	catB	catL
	Trifluoroethyl Amines					
34	4-Nitro-phenyl-	SO_2Me	hPhe	3.6 ± 1.8	36 ± 1	>10 000
33	4-Nitro-phenyl-	SO_2Me	*O*BnSer	13 ± 2	82 ± 8[b]	>10 000
37	4-Nitro-phenyl-	SO_2Me	Acc	7.8 ± 0.7	>12 000	>10 000
22	4-Fluoro-phenyl-	SO_2Me	*O*BnSer	2.4 ± 0.2	157[b]	1951[b]
36[c]	4-Fluoro-phenyl-	SO_2Me	Acc	1.0 ± 0.1	1670[b]	7106[b]
25	4-Fluoro-phenyl-	Cha	hPhe	146 ± 15	>12 000	>10 000
26	4-Fluoro-phenyl-	Cha	*O*BnSer	11 ± 3	>12 000	>10 000
38	4-Fluoro-phenyl-	Cha	Acc	20 ± 1	>12 000	1498[b]
	Carboxamides					
27[c]	Morpholine-4-	Cha	hPhe	5.7 ± 0.5	7847[b]	2536[b]
28[c]	Morpholine-4-	Cha	*O*BnSer	1.2 ± 0.1	1054[b]	543[b]
39[c]	Morpholine-4-	Cha	Acc	1.5 ± 0.1	>12 000	385[b]
30	Tetrahydropyran-4-	Cha	hPhe	382 ± 10	>12 000	6335
29	Tetrahydropyran-4-	Cha	*O*BnSer	61 ± 7	>12 000	5036
40	Tetrahydropyran-4-	Cha	Acc	117 ± 8	>12 000	1675
32	Pyridine-4-	Cha	hPhe	13 ± 1	1667[b]	1463[b]
31[c]	Pyridine-4-	Cha	*O*BnSer	11 ± 1	1906[b]	1024[b]
41	Pyridine-4	Cha	Acc	13 ± 1	>12 000	716[b]

[a]Data are shown as mean ± SD of at least duplicate experiments. [b]K_i values were obtained from K_i^{app} values, using the Cheng−Prusoff equation. [c]Inhibition data against catS for these compounds have already been published elsewhere and were reproduced in this study.[12,13]

Table 1B. Inhibition Data for Aldehyde- (45), Ketone-Based Inhibitors (44−51), and Hydrazones (46 and 47)[a]



44–51 46 47

Cpd	X	K_i/K_i^* [nM] catS	k_{inact} (s^{-1})/k_{2nd} (M^{-1} s^{-1}) catS	K_i[b] [nM] catB	catL
45	H	0.039 ± 0.005/0.016 ± 0.001[c]	–	64 ± 6	17 ± 2
44	CH_2SO_2Ph	331 ± 40/43 ± 13[c]	–	>510	>600
46[e]	–	2.2 ± 0.2	–	142 ± 9	21 ± 1
47[e]	–	989 ± 150	–	>12 000	>2500
50	CH_2OBz	25 ± 2[b]	–	>2550	>600
49	CH_2SBz	165 ± 18[b]	–	>10 000	>3000
51	CH_2OH	67 ± 3[b]	–	>2550	>600
48	CH_2Cl	0.5 ± 0.1[b]	0.0023 ± 0.0002/4.3 × 10^6 ± 1.2 × 10^6	n.d.[f]	n.d.[f]
52	vinylsulfone[g]	3.0 ± 0.6[d]	0.020 ± 0.001/7.2 × 10^6 ± 2.4 × 10^5	n.d.[f]	n.d.[f]

[a]Data are shown as mean ± SD of at least duplicate experiments. [b]K_i values were obtained from K_i^{app} values using the Cheng−Prusoff equation. [c]K_i and K_i^* values were obtained from K_i^{app} and $K_i^{*\,app}$. [d]k_{inact} and k_{2nd} were determined from plotting k_{obs} values against inhibitor concentration, as described previously.[22] [e]Inhibitory activity of **46** stems mainly from unmasked **45**, whereas, for hydrazone **47**, the inhibition of liberated **44** is of minor impact, estimated from reported affinities of carbonyl species **44** and **45**, and expected liberated amount by cleavage assays (Figure S7). [f]n.d. = not determined. [g]Data taken from ref 22.

All synthesized inhibitors were tested against catS and off-target cathepsins L and B (catL, catB) in fluorometric enzyme assays and their inhibition constants (K_i for all, K_i^* for slowly reversible, k_{inact} and k_{2nd} for irreversible binders) determined.[21,22] Finally, the membrane permeability of selected inhibitors was measured with a Parallel Artificial Membrane

https://doi.org/10.1021/acsmedchemlett.4c00050
ACS Med. Chem. Lett. 2024, 15, 837−844

Permeation Assay (PAMPA) and correlated with the inhibition data and with calculated physicochemical parameters (log *P*; topological polar surface area, TPSA). For a more-detailed description on the assay procedure, see the methods section in the SI. The results for the nitrile-based inhibitors are summarized in Table 1A; the results for the inhibitors with other warheads are listed in Table 1B.

The exchange of the different side chains between the nitrile-based inhibitors produced eight new inhibitors with affinities in the low double-digit nanomolar range or even lower, similar to the starting compounds **28** and **36**, with K_i values of 1.2 nM and 1.0 nM, respectively (Table 1A).[22,23] It was also observed that the tetrahydropyran derivatives and inhibitor **25** showed significantly higher K_i values, compared to those of **28** and **36**. Overall, the selectivity against the off-targets catB and catL remained in the same range, e.g., compound **37** retained a good selectivity of >1500 over catB, compared to 1670 for starting inhibitor **36**, and even higher selectivity compared to the other starting inhibitor **28** (1054 over catB; see the SI). Compound **34** reached a higher selectivity over catL, compared to the second starting molecule **28** (>2778 over catL vs 453 for compound **28** vs catL, as given in the SI) (see Table 1A).

Exchanging the nitrile warhead of starting compound **28** with an aldehyde (**45**) led to a subnanomolar inhibitor with good selectivities over catB (≥1600; see the SI) and catL (≥435; see the SI), respectively. Further variations of the aldehyde included several α-substituted ketones which are known to react covalent-reversibly (through their ketone moiety) or depending on the quality of the leaving group in the α-position, irreversibly with papain-like cysteine proteases.[24] Visually inspecting the progress curves of the aldehyde (**45**) and the chloromethylketone (**48**) indicated time-dependent inhibition for both inhibitors. Since aldehydes react covalent-reversibly, compound **45** was evaluated as a slow-reversible binder of catS, while the chloromethylketone warhead of **48** is an irreversible inactivator of cysteine-proteases.[25] For the determination of K_i and K_i^* of **45**, we followed a procedure well described for slow-reversible tight binders (see Figure S1).[26,27] The plotting of the initial velocity v_i against the inhibitor concentration leads to K_i^{app} and a similar plotting of the steady-state velocity (v_s) against the inhibitor concentration gives access to $K_i^{*\,app}$, which can then be transformed to the corresponding K_i and $K_i^{*\,app}$ values with the Cheng–Prusoff equation.[26]

For the chloromethylketone (**48**) plotting k_{obs} values against the inhibitor concentration as described previously for irreversible inhibitors, revealed a subnanomolar inhibition constant and a second-order rate of $4.3 \times 10^6 \pm 1.2 \times 10^6$ $M^{-1}\,s^{-1}$, which confirms the high reactivity of this inhibitor class.

To our surprise, **50** did not show time-dependent inhibition in the time scale of the assay. The acyloxymethylketone warhead is known to eventually react irreversibly by substitution of the phenyl carboxylate leaving group.[24] The analogous thio derivative **49** also did not show time-dependent inhibition, leading us to the assumption that, for both inhibitors, a very slow irreversible step follows the reversible enzyme–inhibitor complex formation as described by Brady et al.[29] We were able to determine K_i values for the non-time-dependent first step of the inhibition process (Table 1B). Hydroxymethylketone inhibitor **51** did not exhibit time-dependent inhibition, as expected for this type of warhead; therefore, the K_i value was calculated.[25] Finally, inhibitor **44** with a new type of warhead (α-sulfonylphenyl methylketone) clearly showed time-dependent inhibition. To differ between reversible and irreversible binding, dilution assays (Figure 3) were performed for selected inhibitors with new warheads against catS (**44** and **49**–**52**; see Table 1B).

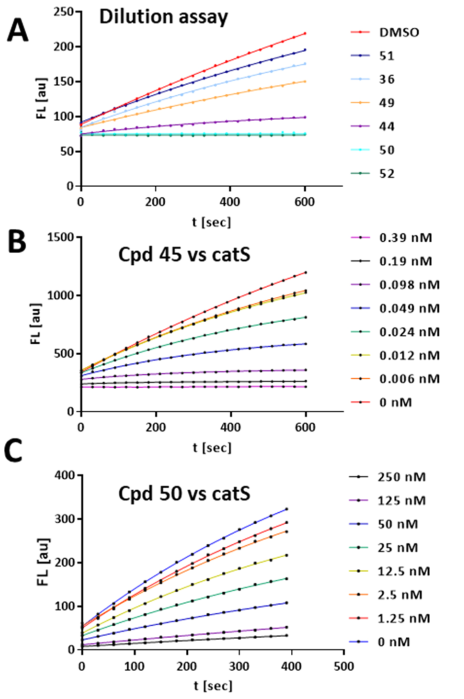


Figure 3. (A) Dilution assay of DMSO control (red), compounds **51** (dark blue), **36** (reversible nitrile, blue), **49** (orange), **44** (purple), **50** (cyan), and irreversible control **52** (vinylsulfone, green). Inhibitors were incubated at 20-fold K_i or $K_i^{*\,app}$ concentrations followed by 100-fold dilution, resulting in enzyme recovery in cases of DMSO and compounds **51**, **36**, **49** (linear progress curves), and **44** (time-dependent progress curve). Progress curves of (B) inhibitor **45**, catS and (C) inhibitor **50**, catS, showing the time-dependent inhibition by cpd **45** and non-time-dependent inhibition by cpd **50**.

K_i and $K_i^{*\,app}$ values for the slow-reversible inhibitor **44** were then calculated as described for **45** (Table 1B).[22]

The results of the dilution assays confirmed the irreversible binding behavior of acyloxymethyl ketone **50** (cyan, Figure 3A) since the enzyme activity did not recover, similar to the irreversible vinylsulfone control inhibitor **52** (green). Comparing the progress curves of the dilution assays of nitrile inhibitor **36** (blue) with the hydroxymethylketone **51** (dark blue) confirmed the reversibility of both warheads, which is well-described in the literature, since the hydroxymethylketone inhibitor **51** has no adequate leaving group.[7,30] Interestingly, inhibitors **49** (orange) and **44** (purple) seem to show a reversible binding behavior, since, for both, a recovery of the enzyme activity was observed. In the case of **44**, the progress curve is time-dependent, which indicates slow, reversible binding. Inspecting the noncovalent docking poses of **44** and **50** inside the active site of catS indicated that the electrophilic carbon atoms of **44** and **50** are in similarly close proximity to the

https://doi.org/10.1021/acsmedchemlett.4c00050
ACS Med. Chem. Lett. 2024, 15, 837–844

nucleophilic thiol of Cys-25 (distance: 2.9 and 3.4 Å for **44**; 2.7 and 3.8 Å for **50**; see Figure 4). Interestingly, inhibitor **50** has

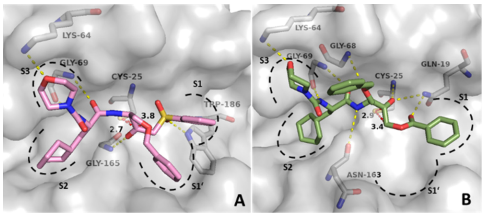


Figure 4. Noncovalent docking poses of **44** ((A) pink C atoms) and (**50** (B), green C atoms) inside the active site of catS (pdb entry: 1MS6). Polar interactions between the inhibitors and active site amino acids are depicted as yellow dashed lines. The distance between the sulfur atom of Cys-25 and the electrophilic C atoms of **44** and **50** is depicted as a red dashed line and the distance is given in Å.

more polar interactions with the binding pocket, compared to **44** (Figure 4), which could help to understand the lower K_i of **50**, in comparison to **44**.

The inhibitory potency of **46** is mainly attributed to traces of cleaved aldehyde **45**, whereas that for **47** (≤5% free **44** under assay conditions) is mainly attributed to the noncleaved hydrazone. The hydrazone derivative **47** showed a 3-fold drop in affinity toward catS, compared to unmasked **44**, and **46** even had a 56-fold drop in affinity, compared to free aldehyde **45**, proving the successful masking of both inhibitors. To investigate the liberation, both hydrazones were incubated in an acidic (pH 4 as artificial lysosomal conditions) and neutral (pH 7.4) environment (see the SI for a detailed description of the methodology). After incubating the respective inhibitor for 24 h at 37 °C, the amount of free aldehyde **45** and ketone **44** was higher at the lower pH value, compared to physiological conditions, indicating a higher cleavage rate under acidic conditions, especially for **46** (80% free inhibitor after 24 h at pH 4). The cleavage of **47** was slower (13% free inhibitor after 24 h at pH 4), indicating a lower hydrolysis rate. These findings highlight the reversible nature of the capped warheads especially under acidic conditions, as they appear in many tumors (Figures S7–S15).[6]

For the clinical candidates presented in Figure 1, oral bioavailability was shown during the drug development process.[31−35] However, some catS inhibitors in the literature suffer from poor experimental permeability (<0.5 × 10^{-6} cm/s).[36] To underline the utility of the newly synthesized compounds in in cellulo and in vivo contexts, their ability to cross membranes was assessed. For this determination, the well-established high-throughput PAMPA approach was used, and detection/quantification was performed by HPLC/MS.[37] By observing a compound's time-dependent permeation of an artificial membrane derived from phospholipids in *n*-dodecane, a good estimation of its passive permeability can be made.[38,39] To rationalize this for future derivatization campaigns, a correlation with the log *P* and TPSA of the compounds is depicted in Figure 5 (all compounds are calculated to be predominantly uncharged in physiologic media).[40] The correlation between physicochemical parameters such as log *P*, PAMPA permeability, and human intestinal absorption can be employed in the context of biopharmaceutics class determination.[41]

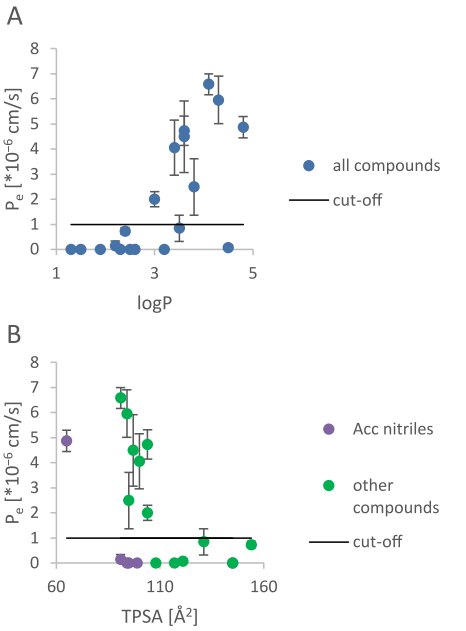


Figure 5. (A) Correlation of log *P* and permeability (P_e) for 19 compounds. (B) Correlation of TPSA and permeability (P_e), divided into two subsets (6 Acc nitriles and 13 others). "Cut-off" depicts the effective permeability value, above which compounds were considered relevantly permeable (P_e > 1 × 10^{-6} cm/s).

For most Acc-containing nitrile inhibitors, very low permeabilities (oftentimes with no permeation detected) were observed. This is also reflected in their low log *P* values of 1.3−2.5. Another highly hydrophilic structure hindering permeation appears to be the methylsulfonyl alanine as the P2 moiety, which can be inferred from the comparison of **36**, **22**, and **38**. While the exchange of P1 to *O*BnSer (**36** → **22**) increases the log *P* value by 1 and does not promote a real improvement in permeability in this case, the exchange of P2 to cyclohexyl alanine (**36** → **38**) increases the log *P* value by ca. 3, resulting in a compound with acceptable permeability. Therefore, **38** is the only compound with Acc in P1 that shows detectable permeation, while no compound with methylsulfonyl alanine as P2 showed appreciable permeation. Note that the combination of the lipophilic fluoro-aryl TFE and Cha moieties in P3 and P2, respectively, can result in unfavorable physicochemical properties due to high lipophilicity if a larger hydrophobic P1 amino acid is used (i.e., **25**, **26**). Excessive lipophilicity is the reason why four inhibitors (**25**, **26**, **49**, and **50**, log *P* ≥ 4.8) could not be assessed with this method, due to a combination of low aqueous solubility and their structurally inherent low detectability (weak ionizability and/or spectrophotometric properties). In contrast, well-balanced lipophilicities can be described for six inhibitors (**27**−**32**) with log *P* values of 3.0−4.3 that, therefore, show appreciable permeabilities ((2−7) × 10^{-6} cm/s). In this subgroup, the hPhe-containing inhibitor always has slightly better permeability than the *O*BnSer-containing analogue. Additionally, a difference in the impact of the P3 residue can

https://doi.org/10.1021/acsmedchemlett.4c00050
ACS Med. Chem. Lett. 2024, 15, 837−844

Table 2. Permeability (P_e) as Determined by PAMPA ($n \geq 4$), and Calculated Physicochemical Parameters log*P* and TPSA for the Listed Compounds Clustered by Structural Features[a]

Cpd	P_e [$\times 10^{-6}$ cm/s]	log *P*	TPSA [Å^2]	Cpd	P_e [$\times 10^{-6}$ cm/s]	log *P*	TPSA [Å^2]
	Nitriles, Acc P1				**Nitriles, Other**		
36	0.0 ± 0.0	1.5	99	**27**	6.0 ± 1.0	4.3	94
37	0.0 ± 0.0	1.3	145	**28**	4.7 ± 0.6	3.6	104
38	4.9 ± 0.4	4.8	65	**29**	4.1 ± 1.1	3.4	100
39	0.0 ± 0.0	2.5	94	**30**	6.6 ± 0.4	4.1	91
40	0.1 ± 0.2	2.2	91	**31**	2.0 ± 0.3	3.0	104
41	0.0 ± 0.0	1.9	95	**32**	2.5 ± 1.1	3.8	95
	Nitriles, Methylsulfonyl Alanine P2				**Aldehyde, Ketones, Hydrazone**		
22	0.0 ± 0.0	2.6	108	**45**	4.5 ± 1.4	3.6	97
33	0.7 ± 0.1	2.4	154	**44**[b]	0.8 ± 0.5	3.5	131
34	0.0 ± 0.0	3.2	145	**51**	0.0 ± 0.0	2.3	117
				46	0.1 ± 0.1	4.5	121

[a]log*P* and TPSA were calculated using the Molinformation molecular properties calculator (https://www.molinspiration.com/).[28] [b]Instability detected (see Figures S3−S6 for stability investigations); only peaks with the expected concomitant *m*/*z* signal were used for quantification.

be observed in which the permeability trends are tetrahydropyran = morpholine > pyridine. Another observation is that the exchange of the warhead between nitrile and aldehyde (**28** → **45**) does not strongly influence the permeability, which is also reflected by identical log *P* values, indicating the possible transfer of the structure-related trends to the respective aldehydes. As depicted in Figure 5A, the trend for the presented compound set (Table 2) is that, in most cases, a log *P* value of >3 is required for good permeation and in the log *P* range of 3−5, a loose proportionality between log *P* and P_e is seen. This underlines the necessity of detailed structure−permeability relationship studies to predict the influence of distinct structural modifications, in addition to the facile calculation of log *P*.

The correlation between the permeability and TPSA also shows interesting relations. The best permeability is observed for compounds with TPSAs of 91−104 Å^2. If the threshold is exceeded (TPSA values of 108−154 Å^2), it results in a drastic decrease in permeation due to increased hydrophilicity. The most impactful moiety in this context seems to be the P3 nitrobenzene (e.g., in **33**, **34**, and **37**) and, after that, the P2 methylsulfonyl alanine. The compounds with TPSA values of >110 Å^2 are assumed to not have the ability to be effectively immersed in an apolar membrane and are therefore not transported through it. However, there are other compounds with seemingly adequate TPSA values (91−99 Å^2) that still have low permeability, namely most Acc nitriles. This can be attributed to the fact that the TPSA is an absolute measure. If the relatively smaller total surface area of the P1 Acc nitriles is considered, it becomes obvious that much of their surface area is polar, in contrast to the other compounds with comparable TPSA values that carry the much larger OBnSer or hPhe in P1. **38** seems to be the exception to the detrimental Acc effect, which indicates that it can be balanced out by the introduction of very lipophilic moieties, as reasoned above. The mentioned distinction of Acc nitriles from other nitriles in the set is also visible in Figure 5B. Each depicted subset has its own cutoff for permeability, which is lower for Acc nitriles (<90 Å^2) than for the other compounds (<110 Å^2).

While **47** could not be assessed for permeability due to solubility limitations, the hydrazone **46** has low permeability ($0.1 \pm 0.1 \times 10^{-6}$ cm/s), probably due to the high TPSA, despite its adequate log *P* value. This is important, because it underlines the utility of the warhead modification. It not only limits its strong electrophilicity but also hinders its cell permeation until it arrives in an environment as acidic as the TME, where the more-permeable and more-reactive warhead is liberated.

In summary, we successfully synthesized 17 nitrile inhibitors as well as five ketone-based inhibitors and one aldehyde inhibitor of catS and tested them on this interesting target and on off-target cathepsins in fluorometric enzyme assays. The exchange of different side chains among nitrile-based inhibitors produced some inhibitors with affinities that were comparable to that of the starting compounds and elucidated some clear affinity limitations. The starting compounds **28** and **36** already exhibited excellent K_i values, laying the background for the subsequent modifications. Two of the newly produced inhibitors even showed improved selectivities against off-target cathepsins B and L over the starting molecules. Introducing an aldehyde warhead in place of the nitrile led to a subnanomolar inhibitor with favorable selectivities.

The aldehyde and ketone derivatives showed time-dependent inhibition and suggested slow-reversible binding for certain compounds. Acyloxymethylketone **50** exhibited irreversible behavior in the dilution assay, while hydroxymethylketone **51** and surprisingly **49** displayed reversible characteristics. The introduction of a new warhead, α-sulfonylphenyl methylketone (**44**), demonstrated time-dependent reversible inhibition. Furthermore, warhead-masking hydrazone derivatives **46** and **47** were synthesized and examined, revealing their susceptibility to cleavage under acidic conditions. The clear difference in liberation kinetics (for **46**: 50% liberated after 24 h at pH 7.4 vs 50% liberated after 2 h at pH 4) are promising for delivery systems with targeted release in the acidic TME (e.g., attachment of the aldehyde to nanoparticles as a hydrazone) where pH values below 7 are commonly encountered.[42] The reversible attachment to nano carriers via the hydrazone moiety is expected to render the low membrane permeability of the hydrazones irrelevant by overriding the pharmacokinetic properties.

The assessment of membrane permeability highlighted some nitriles (esp. **30**) and the aldehyde (**45**) as promising novel candidates for in vivo studies while underlining challenges for inhibitors with strong hydrophilic groups. Correlation analyses between log *P*, TPSA, and permeability provided insights into the structural factors influencing transport across membranes and gave indications for future structural modification campaigns guided by calculated physicochemical properties.

https://doi.org/10.1021/acsmedchemlett.4c00050
ACS Med. Chem. Lett. 2024, 15, 837−844

To conclude, this comprehensive investigation provides valuable insights into the design and behavior of catS inhibitors, paving the way for further optimization and development of potential therapeutic candidates.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsmedchemlett.4c00050.

Full experimental details on synthesis, materials and methods, additional figures on kinetic evaluations, stability analysis, molecular docking score tables, and characterization data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Tanja Schirmeister − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; Email: schirmei@uni-mainz.de

Authors

Mergim Meta − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; orcid.org/0000-0003-3601-9523

Collin Zimmer − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*

Natalie Fuchs − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*; orcid.org/0000-0001-6404-676X

Maximilian Johannes Zecher − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*

Albin Lahu − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsmedchemlett.4c00050

Author Contributions

‡Mergim Meta and Collin Zimmer contributed equally. All authors have given approval to the final version of the manuscript.

Funding

Financial support by the DFG (Deutsche Forschungsgemeinschaft) in the framework of the CRC 1066 (Nanodimensional Polymeric Therapeutics for Tumor Therapy), project Q5 (Targeting and Immunomodulator Structures and their Coupling to Therapeutic Nanosystems for Oncological Application) is gratefully acknowledged.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would kindly like to acknowledge the support of Michael Klein for creating figures with BioRender during this project.

■ ABBREVIATIONS

Acc, 1-aminocyclopropylcarboxylic acid; AMC, 7-amino-4-methylcoumarin; catB, human cathepsin B; catL, human cathepsin L; catS, human cathepsin S; Cha, cyclohexylalanine; cpd, compound; DMP, Dess−Martin periodinane; DTT, dithiothreitol; eq, equivalent; hPhe, homophenylalanine; PAMPA, parallel artificial membrane permeation assay; *O*BnSer, *O*-benzyl-L-serine; TFE, trifluoroethylamine; TME, tumor microenvironment; TPSA, topological polar surface area

■ REFERENCES

(1) Otto, H.-H.; Schirmeister, T. Cysteine Proteases and Their Inhibitors. *Chem. Rev.* **1997**, *97* (1), 133−172.

(2) Shi, G. P.; Webb, A. C.; Foster, K. E.; Knoll, J. H. M.; Lemere, C. A.; Munger, J. S.; Chapman, H. A. Human Cathepsin S: Chromosomal Localization, Gene Structure, and Tissue Distribution. *J. Biol. Chem.* **1994**, *269* (15), 11530−11536.

(3) Fuchs, N.; Meta, M.; Schuppan, D.; Nuhn, L.; Schirmeister, T. Novel Opportunities for Cathepsin S Inhibitors in Cancer Immunotherapy by Nanocarrier-Mediated Delivery. *Cells* **2020**, *9*, 2021 (NLM (Medline), Sept. 2, 2020).

(4) Edman, M. C.; Janga, S. R.; Meng, Z.; Bechtold, M.; Chen, A. F.; Kim, C.; Naman, L.; Sarma, A.; Teekappanavar, N.; Kim, A. Y.; Madrigal, S.; Singh, S.; Ortiz, E.; Christianakis, S.; Arkfeld, D. G.; Mack, W. J.; Heur, M.; Stohl, W.; Hamm-Alvarez, S. F. Increased Cathepsin S Activity Associated with Decreased Protease Inhibitory Capacity Contributes to Altered Tear Proteins in Sjögren's Syndrome Patients. *Sci. Rep.* **2018**, *8* (1), 11044.

(5) Bararia, D.; Hildebrand, J. A.; Stolz, S.; Haebe, S.; Alig, S.; Trevisani, C. P.; Osorio-Barrios, F.; Bartoschek, M. D.; Mentz, M.; Pastore, A.; Gaitzsch, E.; Heide, M.; Jurinovic, V.; Rautter, K.; Gunawardana, J.; Sabdia, M. B.; Szczepanowski, M.; Richter, J.; Klapper, W.; Louissaint, A.; Ludwig, C.; Bultmann, S.; Leonhardt, H.; Eustermann, S.; Hopfner, K. P.; Hiddemann, W.; von Bergwelt-Baildon, M.; Steidl, C.; Kridel, R.; Tobin, J. W. D.; Gandhi, M. K.; Weinstock, D. M.; Schmidt-Supprian, M.; Sárosi, M. B.; Rudelius, M.; Passerini, V.; Mautner, J.; Weigert, O. Cathepsin S Alterations Induce a Tumor-Promoting Immune Microenvironment in Follicular Lymphoma. *Cell Rep.* **2020**, *31* (5), 107522.

(6) Dheilly, E.; Battistello, E.; Katanayeva, N.; Sungalee, S.; Michaux, J.; Duns, G.; Wehrle, S.; Sordet-Dessimoz, J.; Mina, M.; Racle, J.; et al. Cathepsin S Regulates Antigen Processing and T Cell Activity in Non-Hodgkin Lymphoma. *Cancer Cell* **2020**, *37* (5), 674−689.e12.

(7) Wilkinson, R. D. A.; Young, A.; Burden, R. E.; Williams, R.; Scott, C. J. A Bioavailable Cathepsin S Nitrile Inhibitor Abrogates Tumor Development. *Mol. Cancer* **2016**, *15* (1), 29.

(8) Villadangos, J. A.; Riese, R. J.; Peters, C.; Chapman, H. A.; Ploegh, H. L. Degradation of Mouse Invariant Chain: Roles of Cathepsins S and D and the Influence of Major Histocompatibility Complex Polymorphism. *J. Exp. Med.* **1997**, *186* (4), 549−560.

(9) Beers, C.; Burich, A.; Kleijmeer, M. J.; Griffith, J. M.; Wong, P.; Rudensky, A. Y. Cathepsin S Controls MHC Class II-Mediated Antigen Presentation by Epithelial Cells In Vivo. *J. Immunol.* **2005**, *174* (3), 1205−1212.

(10) Jakoš, T.; Pišlar, A.; Jewett, A.; Kos, J. Cysteine Cathepsins in Tumor-Associated Immune Cells. *Front. Immunol.* **2019**, *10*, 2037.

(11) Oelschlaegel, D.; Sadan, T. W.; Salpeter, S.; Krug, S.; Blum, G.; Schmitz, W.; Schulze, A.; Michl, P. Cathepsin Inhibition Modulates Metabolism and Polarization of Tumor-Associated Macrophages. *Cancers (Basel)* **2020**, *12* (9), 2579.

(12) Gauthier, J. Y.; Black, W. C.; Courchesne, I.; Cromlish, W.; Desmarais, S.; Houle, R.; Lamontagne, S.; Li, C. S.; Massé, F.; McKay, D. J.; Ouellet, M.; Robichaud, J.; Truchon, J. F.; Truong, V. L.; Wang, Q.; Percival, M. D. The Identification of Potent, Selective, and Bioavailable Cathepsin S Inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *17* (17), 4929−4933.

(13) Ward, Y. D.; Thomson, D. S.; Frye, L. L.; Cywin, C. L.; Morwick, T.; Emmanuel, M. J.; Zindell, R.; McNeil, D.; Bekkali, Y.; Giradot, M.; et al. Design and Synthesis of Dipeptide Nitriles as Reversible and Potent Cathepsin S Inhibitors. *J. Med. Chem.* **2002**, *45* (25), 5471−5482.

(14) Biasizzo, M.; Javoršek, U.; Vidak, E.; Zarić, M.; Turk, B. Cysteine Cathepsins: A Long and Winding Road towards Clinics. *Mol. Aspects Med.* **2022**, *88*, 101150.
(15) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; et al. An Oral SARS-CoV-2 M pro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science* **2021**, *374* (6575), 1586−1593.
(16) Jadhav, P. K.; Schiffler, M. A.; Gavardinas, K.; Kim, E. J.; Matthews, D. P.; Staszak, M. A.; Coffey, D. S.; Shaw, B. W.; Cassidy, K. C.; Brier, R. A.; Zhang, Y.; Christie, R. M.; Matter, W. F.; Qing, K.; Durbin, J. D.; Wang, Y.; Deng, G. G. Discovery of Cathepsin S Inhibitor LY3000328 for the Treatment of Abdominal Aortic Aneurysm. *ACS Med. Chem. Lett.* **2014**, *5* (10), 1138−1142.
(17) Eng, H.; Dantonio, A. L.; Kadar, E. P.; Obach, R. S.; Di, L.; Lin, J.; Patel, N. C.; Boras, B.; Walker, G. S.; Novak, J. J.; Kimoto, E.; Singh, R. S. P.; Kalgutkar, A. S. Disposition of Nirmatrelvir, an Orally Bioavailable Inhibitor of SARS-CoV-2 3C-Like Protease, across Animals and Humans. *Drug Metab. Dispos.* **2022**, *50* (5), 576−590.
(18) Abdulla, M.-H.; Lim, K.-C.; Sajid, M.; McKerrow, J. H.; Caffrey, C. R. Schistosomiasis Mansoni: Novel Chemotherapy Using a Cysteine Protease Inhibitor. *PLoS Med.* **2007**, *4* (1), e14.
(19) Jadhav, P. K.; Schiffler, M. A.; Gavardinas, K.; Kim, E. J.; Matthews, D. P.; Staszak, M. A.; Coffey, D. S.; Shaw, B. W.; Cassidy, K. C.; Brier, R. A.; Zhang, Y.; Christie, R. M.; Matter, W. F.; Qing, K.; Durbin, J. D.; Wang, Y.; Deng, G. G. Discovery of Cathepsin S Inhibitor LY3000328 for the Treatment of Abdominal Aortic Aneurysm. *ACS Med. Chem. Lett.* **2014**, *5* (10), 1138−1142.
(20) Mikhaylov, G.; Klimpel, D.; Schaschke, N.; Mikac, U.; Vizovisek, M.; Fonovic, M.; Turk, V.; Turk, B.; Vasiljeva, O. Selective Targeting of Tumor and Stromal Cells By a Nanocarrier System Displaying Lipidated Cathepsin B Inhibitor. *Angew. Chem., Int. Ed.* **2014**, *53* (38), 10077−10081.
(21) Fuchs, N.; Meta, M.; Lantzberg, B.; Bros, M.; Ling Kuan, S.; Weil, T.; Schirmeister, T. Subnanomolar Cathepsin S Inhibitors with High Selectivity: Optimizing Covalent Reversible α-Fluorovinylsulfones and α-Sulfonates as Potential Immunomodulators in Cancer. *ChemMedChem.* **2023**, *18* (15), e202300160.
(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.; Rodríguez, S.; Diemer, K.; Neumann, W.; Gonzàlez, 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.
(23) Maus, H.; Müller, P.; Meta, M.; Hoba, S. N.; Hammerschmidt, S. J.; Zimmermann, R. A.; Zimmer, C.; Fuchs, N.; Schirmeister, T.; Barthels, F. Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl amides (NBD Amides) as Class Spanning Protease Substrates. *Chem.—Eur. J.* **2023**, *29* (50), e202301855.
(24) Krantz, A.; Copp, L. J.; Coles, P. J.; Smith, R. A.; Heard, S. B. Peptidyl (Acyloxy)Methyl Ketones and the Quiescent Affinity Label Concept: The Departing Group as a Variable Structural Element in the Design of Inactivators of Cysteine Proteinases. *Biochemistry* **1991**, *30*, 4678−4687.
(25) Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R. J.; Kozminski, K.; Li, L. Y.; Lockner, J. W.; Lou, J.; Marra, M. T.; Mitchell, L. J.; Murray, B. W.; Nieman, J. A.; Noell, S.; Planken, S. P.; Rowe, T.; Ryan, K.; Smith, G. J.; Solowiej, J. E.; Steppan, C. M.; Taggart, B. Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J. Med. Chem.* **2020**, *63* (21), 12725−12747.
(26) Klein, P.; Barthels, F.; Johe, P.; Wagner, A.; Tenzer, S.; Distler, U.; Le, T. A.; Schmid, P.; Engel, V.; Engels, B.; Hellmich, U. A.; Opatz, T.; Schirmeister, T. Naphthoquinones as Covalent Reversible Inhibitors of Cysteine Proteases—Studies on Inhibition Mechanism and Kinetics. *Molecules* **2020**, *25* (9), 2064.
(27) Mons, E.; Roet, S.; Kim, R. Q.; Mulder, M. P. C. A Comprehensive Guide for Assessing Covalent Inhibition in Enzymatic Assays Illustrated with Kinetic Simulations. *Curr. Protoc.* **2022**, *2* (6), e419.
(28) *Molinspiration Cheminformatics Free Web Services*; Molinspiration, Slovensky Grob, Slovak Republic; available via the Internet at: https://www.molinspiration.com.
(29) Brady, K. A Catalytic Mechanism for Caspase-1 and for Bimodal Inhibition of Caspase-1 by Activated Aspartic Ketones. *Bioorg. Med. Chem.* **1999**, *7* (4), 621−631.
(30) Bai, B.; Belovodskiy, A.; Hena, M.; Kandadai, A. S.; Joyce, M. A.; Saffran, H. A.; Shields, J. A.; Khan, M. B.; Arutyunova, E.; Lu, J.; Bajwa, S. K.; Hockman, D.; Fischer, C.; Lamer, T.; Vuong, W.; Van Belkum, M. J.; Gu, Z.; Lin, F.; Du, Y.; Xu, J.; Rahim, M.; Young, H. S.; Vederas, J. C.; Tyrrell, D. L.; Lemieux, M. J.; Nieman, J. A. Peptidomimetic α-Acyloxymethylketone Warheads with Six-Membered Lactam P1 Glutamine Mimic: SARS-CoV-2 3CL Protease Inhibition, Coronavirus Antiviral Activity, and in Vitro Biological Stability. *J. Med. Chem.* **2022**, *65* (4), 2905−2925.
(31) Barr, S. C.; Warner, K. L.; Kornreic, B. G.; Piscitelli, J.; Wolfe, A.; Benet, L.; McKerrow, J. H. A Cysteine Protease Inhibitor Protects Dogs from Cardiac Damage during Infection by Trypanosoma Cruzi. *Antimicrob. Agents Chemother.* **2005**, *49* (12), 5160−5161.
(32) Engel, J. C.; Doyle, P. S.; Hsieh, I.; Mckerrow, J. H. Cysteine Protease Inhibitors Cure an Experimental Trypanosoma cruzi Infection. *J. Exp. Med.* **1998**, *188*, 725−734.
(33) Kratochwil, N. A.; Stillhart, C.; Diack, C.; Nagel, S.; Al Kotbi, N.; Frey, N. Population Pharmacokinetic Analysis of RO5459072, a Low Water-Soluble Drug Exhibiting Complex Food-Drug Interactions. *Br. J. Clin. Pharmacol.* **2021**, *87* (9), 3550−3560.
(34) Payne, C. D.; Deeg, M. A.; Chan, M.; Tan, L. H.; LaBell, E. S.; Shen, T.; DeBrota, D. J. Pharmacokinetics and Pharmacodynamics of the Cathepsin S Inhibitor, LY3000328, in Healthy Subjects. *Br. J. Clin. Pharmacol.* **2014**, *78* (6), 1334−1342.
(35) Olson, J. Antimalarial Effects in Mice of Orally Administered Peptidyl Cysteine Protease Inhibitors. *Bioorg. Med. Chem.* **1999**, *7* (4), 633−638.
(36) Schade, M.; Merla, B.; Lesch, B.; Wagener, M.; Timmermanns, S.; Pletinckx, K.; Hertrampf, T. Highly Selective Sub-Nanomolar Cathepsin S Inhibitors by Merging Fragment Binders with Nitrile Inhibitors. *J. Med. Chem.* **2020**, *63* (20), 11801−11808.
(37) Kansy, M.; Senner, F.; Gubernator, K. Communications to the Editor Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41* (7), 1007−1010.
(38) Zhu, C.; Jiang, L.; Chen, T.-M.; Hwang, K.-K. A Comparative Study of Artificial Membrane Permeability Assay for High Throughput Profiling of Drug Absorption Potential. *Eur. J. Med. Chem.* **2002**, *37*, 399−407.
(39) Kerns, E. H.; Di, L.; Petusky, S.; Farris, M.; Ley, R.; Jupp, P. Combined Application of Parallel Artificial Membrane Permeability Assay and Caco-2 Permeability Assays in Drug Discovery. *J. Pharm. Sci.* **2004**, *93* (6), 1440−1453.
(40) Chemaxon. *Playground v1.6.1*. Available via the Internet at: https://playground.calculators.cxn.io/ (accessed Jan. 28, 2024).
(41) Box, K.; Comer, J. Using Measured pK_a, Log P and Solubility to Investigate Supersaturation and Predict BCS Class. *Curr. Drug Metab.* **2008**, *9* (9), 869−878.
(42) Justus, C. R.; Dong, L.; Yang, L. V. Acidic Tumor Microenvironment and pH-Sensing G Protein-Coupled Receptors. *Front. Physiol.* **2013**, *4*, 354.

https://doi.org/10.1021/acsmedchemlett.4c00050
ACS Med. Chem. Lett. 2024, 15, 837−844

Supporting information

Structural Modifications of Covalent Cathepsin S Inhibitors: Impact on Affinity, Selectivity, and Permeability

Mergim Meta[‡,1], Collin Zimmer[‡,1], Natalie Fuchs[1], Maximilian Johannes Zecher[1], Albin Lahu[1] and Tanja Schirmeister[1]*.

[‡]*Both authors contributed equally*

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

**Corresponding author: Prof. Dr. Tanja Schirmeister, schirmei@uni-mainz.de.*

Content

Permeability (PAMPA)

A General principle after Lit Kansy *et al*. was used.[17] Incubation setup consisted of donor (top) plate (Sigma Aldrich, MAIPNTR10), 5 µL artificial membrane (1 % (w/v) L-α-phosphatidyl-choline, Sigma Aldrich P3556, in *n*-dodecane, Sigma Aldrich 8205430100), acceptor (bottom) plate (Greiner, 655074).

Measurement setup for LC/MS consisted of Agilent 1100 series HPLC system coupled to an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI). An Agilent Poroshell 120 EC-C18, 150 x 2.10 mm, 4 µm column (for all new compounds) or an Agilent Zorbax SB-Aq 4.6 × 150 mm, 5 µm column (for propranolol and methotrexate) was used. A linear gradient was employed for elution with a ternary pump using [water/ACN/water + 0.1% formic acid] that changes the ratio from 80/10/10 to 0/90/10 over the course of 6 min (0.7 mL/min), then stays constant for 4 min. Injection volume was 100 µL. Areas of peaks in the chromatogram (detection-λ = 210 nm or 254 nm) were used to calculate the AUCs. Retention time and mass spectra recorded in positive ionization mode were used to assign species. Analysis was performed using Mestrenova 12.0.2.

Compound solutions were diluted from 5–20 mM stock solution in DMSO to 100 µM in a buffered (TRIS 50 mM pH 7.4) aqueous solution with 5% final content of DMSO ("donor solution"). Similarly prepared solutions (buffer + solvent) were used as "acceptor solutions". 150 µL of donor solution was applied onto the artificial membrane which was applied first to the donor plate. This was sealed (Greiner, 676070, Viewseal sealer). 400 µL of acceptor solution was applied to the acceptor plate. Incubation setup was assembled and left for 7 h. After this time, acceptor and reference solutions were analyzed with LCMS.

On each day, the experiments were performed in duplicates. The acceptor solutions of those were combined prior to analysis. Reference solutions were prepared by simply mixing the indicated volumes of donor solutions and acceptor buffer at the start of the incubation period and analyzed later with the acceptor solutions.

Calculations of effective permeability P_e were performed using the following equation with V_D and V_A as volumes of donor and acceptor solutions (0.15 cm^3 and 0.4 cm^3), AUC_A and AUC_{Eq} as the area of the measured and baseline-corrected spectrum of sample and reference solutions, A as the porosity-corrected filter area (0.3019 cm^2*0.7 = 0.2113 cm^2) and t as the incubation time given in seconds.

$$P_e = -\frac{V_D * V_A * \ln\left(1 - \frac{AUC_A}{AUC_{Eq}}\right)}{(V_D + V_A) * A * t} \quad (6)$$

Propranolol was used as a control with high permeability (P_e(Lit) = 8–14 *10^{-6} cm/s, was determined with n = 18 to 10.4 ± 2.9 *10^{-6} cm/s), metoprolol as a control with poor permeability (P_e(Lit) = 0.7–1.5 *10^{-6} cm/s; with n = 4 experimental P_e = 0.3 ± 0.1 *10^{-6} cm/s), and methotrexate as a control with no permeability (P_e(Lit) = 0 *10^{-6} cm/s, was determined with n = 18 to 0.0 ± 0.0 *10^{-6} cm/s).[18–20]

Molecular Docking

Two different docking approaches were followed, since the inhibitors were designed to react covalently with the active site cysteine-25 of cathepsin S. First, a conventional non-covalent docking was performed, to estimate affinity and geometry of the pre-organized enzyme-inhibitor complex, secondly a covalent docking was used to determine the final covalent enzyme-inhibitor complex.

In both docking setups a crystallographic reference ligand (BLN) was used for validation *via* redocking (**Table S2**).

Molecular docking experiments were performed using the following crystal structures freely available in the protein data bank (PDB):[21] cathepsin S covalently bound to morpholine-4-carboxylic acid [1s-(2-benzyloxy-1r-cyano-ethylcarbamoyl)-3-methyl-butyl]amide (PDB entry 1MS6).[22] All ligands were energetically minimized before docking with Molecular operating environment (MOE 2020.09)[23] using the MMF94x force field.[24]

Docking approach 1: non-covalent docking with LeadIT

The non-covalent docking was performed with LeadIT-2.3.2.[25] The receptor was prepared in MOE with the protonate3D functionality and the covalent bonds between the co-crystallized ligands and the corresponding protease were untethered *via* the Builder tool in MOE. The binding site was defined as a 6.5 Å shell around the bound reference ligand. Water molecules that form at least three hydrogen bonds with the receptor and ligand were kept as part of the binding site. The docking was performed under default settings using the enthalpy-entropy hybrid approach with 2000 solutions *per* iteration and fragmentation. Only the top pose of the initial docking was kept and re-scored using the HYDE scoring function.[26]

Docking approach 2: covalent docking with MOE

Covalent docking was performed with MOE (version 2020.09). The receptors were prepared using the 3D protonation tool inside MOE. For the covalent reaction of the different warheads, the already existing template reactions were used or customized using the combinatorial library tool of MOE. Initial 30 poses from the triangle match placement with London ΔG scoring were re-scored using the Affinity ΔG scoring function and induced fit refinement implemented in MOE. 10 Poses were kept and visually inspected for binding geometry the interactions matching between the docked inhibitor pose and co-crystallized ligand with the enzyme. The poses best matching inspected interaction patterns are further discussed.

Table S1: Redocking of reference ligand BLN.

Enzyme (pdb entry)	Reference ligand ID	Redocking FlexX (RMSD/ Å)	FlexX score (kJ/mol)
Cathepsin S (1MS6)	BLN	0.94	–18.0

Table S2: Results of molecular docking analysis for cathepsin S (pdb-ID: 1MS6).

Cpd	Distance (electrophilic C-Cys25-S) / Å	FlexX score (kJ/mol)	Hyde score (kJ/ mol)	Covalent docking score (Affinity ΔG, MOE/ kcal/mol)
44	2.7, 3.8*	–29.9	–7	–5.8
45	2.8	–26.6	–49	–3.9
49	3.1, 2.8*	–24.9	–20	–4.4
50	2.9, 3.4*	–28.2	–29	–5.9
51	2.6	–26.5	–36	–4.8

*These inhibitors contain two possible electrophilic carbon atoms.

Stability Analyses for Compound 44

Compound **44** carries a α-sulfonylphenyl methylketone warhead that is currently rarely described in literature in the context of protease inhibition.[27] To validate the findings from the enzyme inhibition assays, its stability in aqueous conditions and against thiol nucleophiles in general was analyzed. In **Figure S3**, some hypothetical reactions are depicted (as described for vinyl sulfones, and acyloxy methyl ketones, respectively).[3,28]

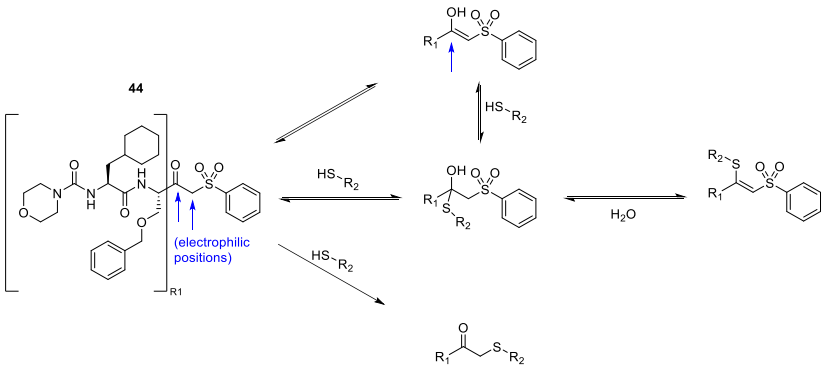


Figure S3: Keto-enol-tautomerism and hypothetical reactivity towards thiol nucleophile in reactions described for similarly structured warheads.

Method: For non-aqueous conditions, DMSO was used as the solvent, and 2-phenylethane-1-thiol (PhEtSH, 2 mM) as nucleophile in presence of triethylamine (TEA, 2 mM) as base to generate $PhEtS^-$ in solution, modified from literature.[7] For aqueous conditions, different 50 mM buffer systems (NH_4HCO_3 for pH 6.5, and TRIS-HCl for pH 7.4, supplemented with DMSO to ensure solubility) with 0 mM or 2.5 mM dithiothreitol (DTT) as a model thiol were utilized, reflecting the employed buffers in enzyme inhibition assays and PAMPA. Incubations were performed in all cases with 200 µM of **44** for the indicated time at 30 °C. Analysis was performed with the Agilent 1100 series HPLC system and software as described earlier. Injection volumes were 100 µL for aqueous conditions and 10 µL for non-aqueous conditions.

First, the results from the reactivity experiment towards $PhEtS^-$ are discussed. It is described in literature that 10 min of incubation time and 2 eq. of nucleophile can be enough to observe relevant formation of covalent adducts.[2] As shown in **Figure S4**, after 1 h of incubation and 10 eq. of nucleophile, no relevant formation of products is observed. While no PhEtSH is consumed, a 10 % reduction in AUC for **44**, and a small additional peak at t_R = 6.4 min was detected that has the same m/z as compound **44** which is therefore attributed to the tautomer depicted in **Figure S4**. With this experiment, we ruled out the irreversible substitution of the sulfone group with elimination of a sulfinic acid leaving group. This is in line with reactivity results described for similarly structured α-sulfonyl amide tool compounds.[29]

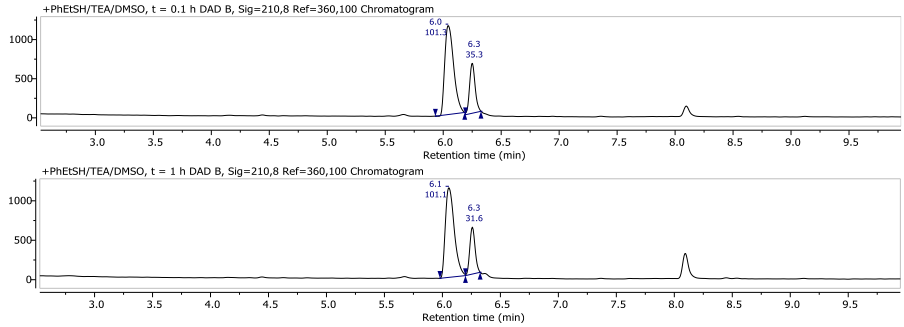


Figure S4: Stability of compound **44** (t_R = 6.3 min, m/z = 600 Da) against thiolate in DMSO. The signal at 6.1 min is attributed to PhEtSH, the signal at 8.1 min is not characterized, but is also formed in background measurements without **44**.

Next, the stability in aqueous systems is described. It is evident from **Figures S5** and **S6** that in aqueous systems, the formation of the putative tautomer (β-hydroxyvinyl sulfone) is the most impactful reaction. Additionally, depending on the presence or absence of DTT, more products are formed indicating insufficient stability over periods >1 . This is expected to have only minor influence over the 10 min period at room temperature that is necessary to record inhibition data. However, over longer periods, as are necessary for dilution assay and PAMPA, this can impact evaluation. Noteworthy are the degradation products with t_R = 5.7 min (m/z = 492 Da) in conditions without DTT and with t_R = 5.3 min (m/z = 646 Da) in conditions with DTT as their difference in m/z is 154 Da which corresponds to the mass of DTT, indicating that the first substance can react with thiols. Their structures were not investigated further, but due to the surplus of DTT present during the dilution assay, it is not expected that this degradation strongly impacts the results from this assay. For PAMPA, however, the formation of degradation products in general (due to unknown formation kinetics and individual permeabilities) introduces a large error in the corresponding value for permeability which is denoted in **Table 2** of the manuscript.

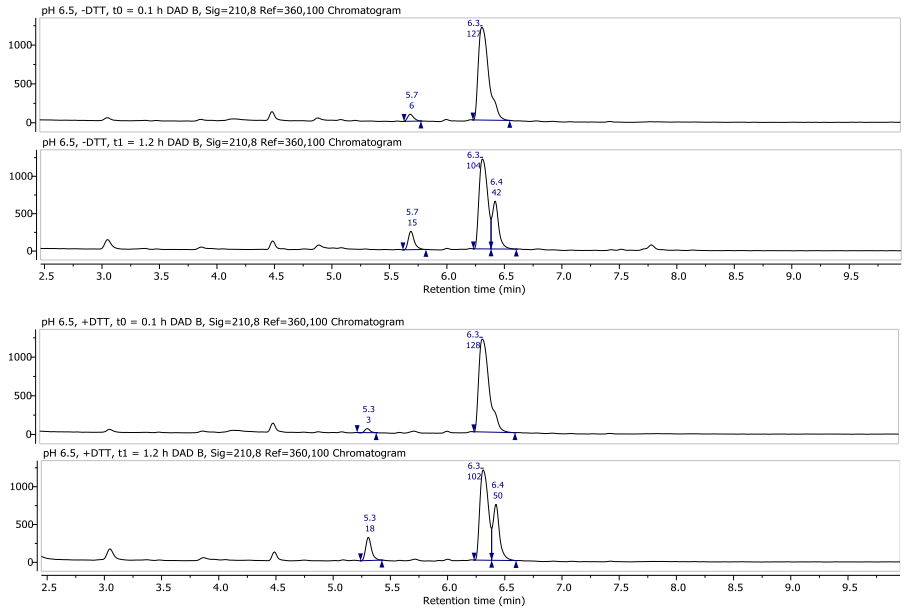


Figure S5: Stability of **44** (t_R = 6.3 min, m/z = 600 Da) in catS-assay-like buffer (top: without DTT. bottom: with DTT) detected at 210 nm. Retention times and AUCs are indicated.

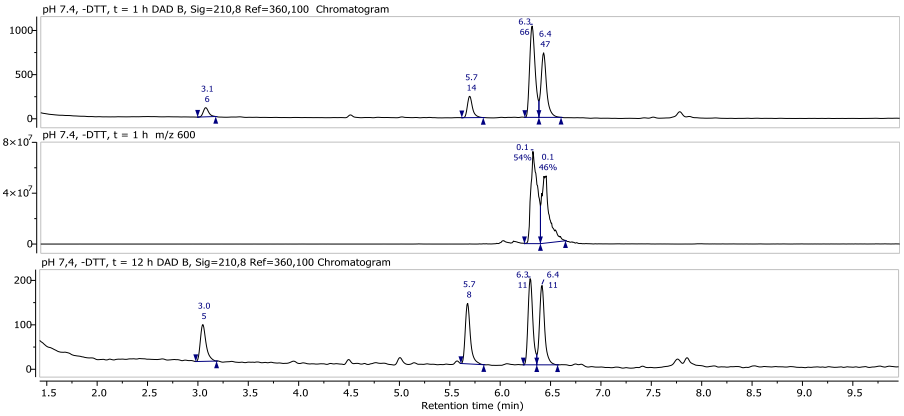


Figure S6: Stability of **44** (t_R = 6.3 min, m/z = 600 Da) in PAMPA buffer (50 mM TRIS pH 7.4) detected at 210 nm (top and bottom) and detected with mass filter of 600 ± 0.25 Da (middle). Spectrophotometrically and mass spectrometrically detected chromatograms were aligned. Retention times and AUCs (absolute or relative) are indicated.

”

Conclusion and Outlook

The published work that is presented in this thesis had the overarching motive to generate data towards the improvement of biological applications of covalent protease inhibitors. To achieve this, two major focuses were placed, namely the in-depth mechanistic investigation of inhibitors harboring different reversibly and irreversibly acting warheads and the possible optimizations towards improved permeability of these inhibitors.

Binding mode investigations were performed that revolved around affinity, selectivity, reversibility and covalency of different types of warheads against the cysteine protease rhodesain that is involved in the pathology of Human African Trypanosomiasis. By establishing a MALDI-TOF-MS-based method for the assessment of covalency for inhibitors, a rational option for clarifying questions on mechanistic implications of structural modifications was realized. It nicely supplements computational methods to model overall reaction progresses like QM/MM calculations, as well as the time-dependent evaluation of the experimental data collected from activity-based assays, and specific reversibility assays. This was applied to inhibitors with α-fluorovinylsulfon(at)e warheads that were found to be distinct from their non-fluorinated progenitors in both binding mode and inhibition kinetics, yet all share the formation of a covalent bond with the target enzyme. Also, inhibitors with dual covalent warheads that can exhibit reversible and irreversible reactions were characterized and the implications of their slowly irreversible mode of action was discussed. The findings from the project on S_NAr-warheads provide a basis for a closer look on sensitive deciding factors for covalency and reversibility, like the electrophilicity and precise substitution pattern on the electron-deficient arene.

Until now, only nitriles have been explored as reversibly covalent warheads in approved cysteine protease inhibitors. Especially the α-fluorovinylsulfon(at)e-based inhibitors have proven to be a promising alternative, based on their efficiency combined with low toxicity and the potential for distribution to the central nervous system when equipped with a suitable recognition sequence. Follow-up experiments require the characterization of these α-fluorovinylsulfon(at)es in animal models of the disease to contextualize their effect in a complex biological system with direct pathophysiological relevance.[631]

Further variations on the S_NAr warheads could include the implementation of alternatives to the nitro groups which are commonly disfavored in terms of toxicity[632] or the investigation of other (het)aryl scaffolds and leaving groups[633]. To fully elucidate the observed bidirectional binding of these inhibitors, longer peptides or extended side-chains could be implemented to address more of rhodesain's open and superficial active site. In a similar way, a decapeptide MMP2 inhibitor with inverse directional binding mode is described.[634]

The second objective of the projects described herein was to give rationalization to ligand-derived protease inhibitors with respect to passive permeability. To this end, a target-independent assessment of

passive permeability was performed for different projects using a parallel artificial membrane permeation assay. This assay has a high throughput and versatile options for detection, enabling the characterization of a broad set of different molecules. The value it brings to drug development is the early inclusion of a pharmacokinetic marker, which can elucidate important factors for structural choices that positively or negatively affect the outcome in preclinical studies. This was exemplified by the influence of charge in the studies on S_NAr warheads against cysteine proteases and α-keto benzothiazole-bearing serine protease inhibitors. In studies involving cellular assays like the one on *S. mansoni*, permeability results can explain observed effects or their absence. In general, knowledge on the equivalence of structural elements can convey security to drug design choices, which was expanded upon in the study on CatS inhibitors.

The relevance of the prodrug approach for peptide-based inhibitors was shown twice. In the study on S_NAr warheads, negative charge in the form of a free carboxylate moiety was identified to be a strong negative determinant for passive permeability. This was exemplified for the dipeptides of the study and for candesartan, all examples for which the ester group represented a prodrug modification suitable to convey permeability. The characterized serine protease inhibitors with an arginine-derived sequence elucidated the same limitation for strong basic moieties. With its almost permanent charge in aqueous environments, a shortened drug-like molecule with strong inhibition data was identified for suitable prodrug modifications. The transformation of the guanidine to a hydroxy guanidine or a carbamate was identified as a required next step for its further development. Another option would be the inclusion of the basic group in an aromatic system, which would reduce affinity but improve permeability.

Vinylsulfone-based *Sm*CatB1 inhibitors were characterized regarding their permeability in the context of a parasitic organism that is especially challenging to penetrate due to its tegumental membrane. In subsequent studies, the irreversible inhibitors of *Sm*CatB1 can be modified to their reversible fluorinated analogs and further modifications to the peptide recognition sequence can be made to improve selectivity. For the most promising inhibitors from the phenotypic approach, cellular target engagement studies would be most logical to rationalize their structural improvements.

For CatS inhibitors as TME modulators, viable structural alternatives were identified in terms of permeability, onto which design strategies can fall back in case one of the preclinical CatS inhibitors encounters development problems related to factors like toxicology or pharmacodynamic effects. Apart from this, the quantified permeability contributions of building blocks are expected to contribute valuably to prospective drug development campaigns against various protease targets. An especially interesting candidate against CatS is the described aldehyde-bearing inhibitor that was identified as permeable and both highly affine and slowly reversible. Additionally, this moiety gives opportunity for reversible conjugation strategies tailored for tumor therapy that can include the introduction of pH-sensitive linkers modulated to respond to the tumor microenvironment, and the conjugation to macromolecular drug delivery systems to influence distribution.

References

(1) Sabe, V. T.; Ntombela, T.; Jhamba, L. A.; Maguire, G. E. M.; Govender, T.; Naicker, T.; Kruger, H. G. Current Trends in Computer Aided Drug Design and a Highlight of Drugs Discovered via Computational Techniques: A Review. *Eur. J. Med. Chem.* **2021**, *224*, 113705. https://doi.org/10.1016/j.ejmech.2021.113705.

(2) Lionta, E.; Spyrou, G.; Vassilatis, D.; Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr. Top. Med. Chem.* **2014**, *14* (16), 1923–1938. https://doi.org/10.2174/1568026614666140929124445.

(3) Moshawih, S.; Bu, Z. H.; Goh, H. P.; Kifli, N.; Lee, L. H.; Goh, K. W.; Ming, L. C. Consensus Holistic Virtual Screening for Drug Discovery: A Novel Machine Learning Model Approach. *J. Cheminform.* **2024**, *16* (1), 62. https://doi.org/10.1186/s13321-024-00855-8.

(4) Croston, G. E. The Utility of Target-Based Discovery. *Expert Opin. Drug Discov.* **2017**, *12* (5), 427–429. https://doi.org/10.1080/17460441.2017.1308351.

(5) Hart, C. P. Finding the Target after Screening the Phenotype. *Drug Discov. Today* **2005**, *10* (7), 513–519. https://doi.org/10.1016/S1359-6446(05)03415-X.

(6) Wang, S.; Wang, Z.; Fang, L.; Lv, Y.; Du, G. Advances of the Target-Based and Phenotypic Screenings and Strategies in Drug Discovery. *Int. J. Drug Discov. Pharmacol.* **2022**, 2. https://doi.org/10.53941/ijddp.v1i1.199.

(7) Swinney, D. C. Phenotypic vs. Target-Based Drug Discovery for First-in-Class Medicines. *Clin. Pharmacol. Ther.* **2013**, *93* (4), 299–301. https://doi.org/10.1038/clpt.2012.236.

(8) Bacilieri, M.; Moro, S. Ligand-Based Drug Design Methodologies in Drug Discovery Process: An Overview. *Curr. Drug Discov. Technol.* **2006**, *3* (3), 155–165. https://doi.org/10.2174/157016306780136781.

(9) Acharya, C.; Coop, A.; E. Polli, J.; D. MacKerell, A. Recent Advances in Ligand-Based Drug Design: Relevance and Utility of the Conformationally Sampled Pharmacophore Approach. *Curr. Comput. Aided-Drug Des.* **2011**, *7* (1), 10–22. https://doi.org/10.2174/157340911793743547.

(10) Blundell, T. L. Structure-Based Drug Design. *Nature* **1996**, *384* (Supp 9), 23–26.

(11) Anderson, A. C. The Process of Structure-Based Drug Design. *Chem. Biol.* **2003**, *10* (9), 787–797. https://doi.org/10.1016/j.chembiol.2003.09.002.

(12) Saar, K. L.; McCorkindale, W.; Fearon, D.; Boby, M.; Barr, H.; Ben-Shmuel, A.; London, N.; von Delft, F.; Chodera, J. D.; Lee, A. A.; Robinson, M. C.; London, N.; Resnick, E.; Zaidmann, D.; Gehrtz, P.; Reddi, R. N.; Gabizon, R.; Barr, H.; Duberstein, S.; Zidane, H.; Shurrush, K.; Cohen, G.; Solmesky, L. J.; Lee, A.; Jajack, A.; Cvitkovic, M.; Pan, J.; Pai, R.; Ripka, E. G.; Nguyen, L.; Shafeev, M.; Matviiuk, T.; Michurin, O.; Chernyshenko, E.; Bilenko, V. A.; Kinakh, S. O.; Logvinenko, I. G.; Melnykov, K. P.; Huliak, V. D.; Tsurupa, I. S.; Gorichko, M.; Shaikh, A.; Pinjari, J.; Swamy, V.; Pingle, M.; BVNBS, S.; Aimon, A.; Delft, F. von; Fearon, D.; Dunnett, L.; Douangamath, A.; Dias, A.; Powell, A.; Neto, J. B.; Skyner, R.; Thompson, W.; Gorrie-Stone, T.; Walsh, M.; Owen, D.; Lukacik, P.; Strain-Damerell, C.; Mikolajek, H.; Horrell, S.; Koekemoer, L.; Krojer, T.; Fairhead, M.; MacLean, E. M.; Thompson, A.; Wild, C. F.; Smilova, M. D.; Wright, N.; Delft, A. von; Gileadi, C.; Rangel, V. L.; Schofield, C.; Salah, E.; Malla, T. R.; Tumber, A.; John, T.; Vakonakis, I.; Kantsadi, A. L.; Zitzmann, N.; Brun, J.; Kiappes, J. L.; Hill, M.; Witt, K. D.; Alonzi, D. S.; Makower, L. L.; Varghese, F. S.; Overheul, G. J.; Miesen, P.; van Rij, R. P.; Jansen, J.; Smeets, B.; Tomésio, S.; Weatherall, C.; Vaschetto, M.; Macdonald, H. B.; Chodera, J. D.; Rufa, D.; Wittmann, M.; Boby, M. L.; Henry, M.; Glass, W. G.; Eastman, P. K.; Coffland, J. E.; Dotson, D. L.; Griffen, E. J.; McCorkindale, W.; Morris, A.; Glen, R.; Cole, J.; Foster, R.; Foster, H.; Calmiano, M.; Tennant, R. E.; Heer, J.; Shi, J.; Jnoff, E.; Hurley, M. F. D.; Lefker, B. A.; Robinson, R. P.; Giroud, C.; Bennett, J.; Fedorov, O.; Reid, S. P.;

Morwitzer, M. J.; Cox, L.; Morris, G. M.; Ferla, M.; Moustakas, D.; Dudgeon, T.; Pšenák, V.; Kovar, B.; Voelz, V.; Carbery, A.; Contini, A.; Clyde, A.; Ben-Shmuel, A.; Sittner, A.; Vitner, B. P. E. B.; Bar-David, E.; Tamir, H.; Achdout, H.; Levy, H.; Glinert, I.; Paran, N.; Erez, N.; Puni, R.; Melamed, S.; Weiss, S.; Israely, T.; Yahalom-Ronen, Y.; Smalley, A.; Oleinikovas, V.; Spencer, J.; Kenny, P. W.; Ward, W.; Cattermole, E.; Ferrins, L.; Eyermann, C. J.; Milne, B. F.; Godoy, A. S.; Noske, G. D.; Oliva, G.; Fernandes, R. S.; Nakamura, A. M.; Gawriljuk, V. O.; White, K. M.; McGovern, B. L.; Rosales, R.; Garcia-Sastre, A.; Carney, D.; Chang, E.; Saikatendu, K. S.; Neyts, L. V. J.; Donckers, K.; Jochmans, D.; Jonghe, S. De; Bowman, G. R.; Borden, B.; Singh, S.; Volkamer, A.; Rodriguez-Guerra, J.; Fate, G.; Hart, S. H.; Bilenko, V. A.; Kinakh, S. O.; Logvinenko, I. G.; Melnykov, K. P.; Huliak, V. D.; Tsurupa, I. S.; Saar, K. L.; Perry, B.; Fraisse, L.; Sjö, P.; Boulet, P.; Hahn, S.; Mowbray, C.; Reid, L.; Rees, P.; Huang, Q. Y. J.; Zvornicanin, S. N.; Shaqra, A. M.; Yilmaz, N. K.; Schiffer, C. A.; Zhang, I.; Pulido, I.; Tomlinson, C.; Taylor, J. C.; Croll, T. I.; Brwewitz, L. Turning High-Throughput Structural Biology into Predictive Inhibitor Design. *Proc. Natl. Acad. Sci.* **2023**, *120* (11). https://doi.org/10.1073/pnas.2214168120.

(13) Venkatraman, S. Discovery of Boceprevir, a Direct-Acting NS3/4A Protease Inhibitor for Treatment of Chronic Hepatitis C Infections. *Trends Pharmacol. Sci.* **2012**, *33* (5), 289–294. https://doi.org/10.1016/j.tips.2012.03.012.

(14) Steinberg, D. Thematic Review Series: The Pathogenesis of Atherosclerosis. An Interpretive History of the Cholesterol Controversy, Part V: The Discovery of the Statins and the End of the Controversy. *J. Lipid Res.* **2006**, *47* (7), 1339–1351. https://doi.org/10.1194/jlr.R600009-JLR200.

(15) Leeson, P. D.; Empfield, J. R. Reducing the Risk of Drug Attrition Associated with Physicochemical Properties; 2010; pp 393–407. https://doi.org/10.1016/S0065-7743(10)45024-1.

(16) Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaudeau, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; Wallace, O.; Weir, A. An Analysis of the Attrition of Drug Candidates from Four Major Pharmaceutical Companies. *Nat. Rev. Drug Discov.* **2015**, *14* (7), 475–486. https://doi.org/10.1038/nrd4609.

(17) van de Waterbeemd, H.; Smith, D. A.; Beaumont, K.; Walker, D. K. Property-Based Design: Optimization of Drug Absorption and Pharmacokinetics. *J. Med. Chem.* **2001**, *44* (9), 1313–1333. https://doi.org/10.1021/jm000407e.

(18) Aungst, B. J. Optimizing Oral Bioavailability in Drug Discovery: An Overview of Design and Testing Strategies and Formulation Options. *J. Pharm. Sci.* **2017**, *106* (4), 921–929. https://doi.org/10.1016/j.xphs.2016.12.002.

(19) Gomez-Orellana, I. Strategies to Improve Oral Drug Bioavailability. *Expert Opin. Drug Deliv.* **2005**, *2* (3), 419–433. https://doi.org/10.1517/17425247.2.3.419.

(20) Glassman, P. M.; Muzykantov, V. R. Pharmacokinetic and Pharmacodynamic Properties of Drug Delivery Systems. *J. Pharmacol. Exp. Ther.* **2019**, *370* (3), 570–580. https://doi.org/10.1124/jpet.119.257113.

(21) Srinivas-Shankar, U.; Wu, F. C. Drug Insight: Testosterone Preparations. *Nat. Clin. Pract. Urol.* **2006**, *3* (12), 653–665. https://doi.org/10.1038/ncpuro0650.

(22) Bertonazzi, A.; Nelson, B.; Salvador, J.; Umland, E. The Smallest Available Estradiol Transdermal Patch: A New Treatment Option for the Prevention of Postmenopausal Osteoporosis. *Women's Heal.* **2015**, *11* (6), 815–824. https://doi.org/10.2217/whe.15.64.

(23) Ingebretsen, O. C.; Flatmark, T. Active and Passive Transport of Dopamine in Chromaffin Granule Ghosts Isolated from Bovine Adrenal Medulla. *J. Biol. Chem.* **1979**, *254* (10), 3833–3839. https://doi.org/10.1016/S0021-9258(18)50662-1.

(24) Scholze, P.; Sitte, H. H.; Singer, E. A. Substantial Loss of Substrate by Diffusion during Uptake in HEK-293 Cells Expressing Neurotransmitter Transporters. *Neurosci. Lett.* **2001**, *309* (3), 173–176. https://doi.org/10.1016/S0304-3940(01)02058-4.

(25) Horackova, H.; Karahoda, R.; Vachalova, V.; Turkova, H.; Abad, C.; Staud, F. Functional Characterization of Dopamine and Norepinephrine Transport across the Apical and Basal Plasma Membranes of the Human Placental Syncytiotrophoblast. *Sci. Rep.* **2022**, *12* (1), 11603. https://doi.org/10.1038/s41598-022-15790-7.

(26) Nguyen, J.-T.; Zhang, M.; Kumada, H.-O.; Itami, A.; Nishiyama, K.; Kimura, T.; Cheng, M.; Hayashi, Y.; Kiso, Y. Truncation and Non-Natural Amino Acid Substitution Studies on HTLV-I Protease Hexapeptidic Inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18* (1), 366–370. https://doi.org/10.1016/j.bmcl.2007.10.066.

(27) Yao, J.-F.; Yang, H.; Zhao, Y.-Z.; Xue, M. Metabolism of Peptide Drugs and Strategies to Improve Their Metabolic Stability. *Curr. Drug Metab.* **2018**, *19* (11), 892–901. https://doi.org/10.2174/1389200219666180628171531.

(28) Inhoffen, H. H.; Hohlweg, W. Neue per Os-Wirksame Weibliche Keimdrüsenhormon-Derivate: 17-Aethinyl-Oestradiol Und Pregnen-in-on-3-Ol-17. *Naturwissenschaften* **1938**, *26* (6), 96–96. https://doi.org/10.1007/BF01681040.

(29) YEN, S. S. C.; MARTIN, P. L.; BURNIER, A. M.; CZEKALA, N. M.; GREANEY, M. O.; CALLANTINE, M. R. Circulating Estradiol, Estrone and Gonadotropin Levels Following the Administration of Orally Active 17β-Estradiol in Postmenopausal Women. *J. Clin. Endocrinol. Metab.* **1975**, *40* (3), 518–521. https://doi.org/10.1210/jcem-40-3-518.

(30) Elger, W.; Barth, A.; Hedden, A.; Reddersen, G.; Ritter, P.; Schneider, B.; Züchner, J.; Krahl, E.; Müller, K.; Oettel, M.; Schwarz, S. Estrogen Sulfamates: A New Approach to Oral Estrogen Therapy. *Reprod. Fertil. Dev.* **2001**, *13* (4), 297. https://doi.org/10.1071/RD01029.

(31) HARPER, M. J. K.; WALPOLE, A. L. Contrasting Endocrine Activities of Cis and Trans Isomers in a Series of Substituted Triphenylethylenes. *Nature* **1966**, *212* (5057), 87–87. https://doi.org/10.1038/212087a0.

(32) Degregorio, M.; Wiebe, V.; Kangas, L.; Härkönen, P.; Väänänen, K.; Laine, A. US 5,912, 273 - Triphenylethylenes Compositions, 1999.

(33) Black, L. J.; Sato, M.; Rowley, E. R.; Magee, D. E.; Bekele, A.; Williams, D. C.; Cullinan, G. J.; Bendele, R.; Kauffman, R. F.; Bensch, W. R. Raloxifene (LY139481 HCI) Prevents Bone Loss and Reduces Serum Cholesterol without Causing Uterine Hypertrophy in Ovariectomized Rats. *J. Clin. Invest.* **1994**, *93* (1), 63–69. https://doi.org/10.1172/JCI116985.

(34) Miller, C. P.; Collini, M. D.; Tran, B. D.; Harris, H. A.; Kharode, Y. P.; Marzolf, J. T.; Moran, R. A.; Henderson, R. A.; Bender, R. H. W.; Unwalla, R. J.; Greenberger, L. M.; Yardley, J. P.; Abou-Gharbia, M. A.; Lyttle, C. R.; Komm, B. S. Design, Synthesis, and Preclinical Characterization of Novel, Highly Selective Indole Estrogens. *J. Med. Chem.* **2001**, *44* (11), 1654–1657. https://doi.org/10.1021/jm010086m.

(35) Bachmann, G. A.; Komi, J. O. Ospemifene Effectively Treats Vulvovaginal Atrophy in Postmenopausal Women. *Menopause* **2010**, *17* (3), 480–486. https://doi.org/10.1097/gme.0b013e3181c1ac01.

(36) Bramwell, V. H. C.; Pritchard, K. I.; Tu, D.; Tonkin, K.; Vachhrajani, H.; Vandenberg, T. A.; Robert, J.; Arnold, A.; O'Reilly, S. E.; Graham, B.; Shepherd, L. A Randomized Placebo-Controlled Study of Tamoxifen after Adjuvant Chemotherapy in Premenopausal Women with Early Breast Cancer (National Cancer Institute of Canada—Clinical Trials Group Trial, MA.12). *Ann. Oncol.* **2010**, *21* (2), 283–290. https://doi.org/10.1093/annonc/mdp326.

(37) Ettinger, B. Reduction of Vertebral Fracture Risk in Postmenopausal Women With Osteoporosis Treated With Raloxifene - Results From a 3-Year Randomized Clinical Trial. *JAMA* **1999**, *282*

(7), 637. https://doi.org/10.1001/jama.282.7.637.

(38) Miller, P. D.; Chines, A. A.; Christiansen, C.; Hoeck, H. C.; Kendler, D. L.; Lewiecki, E. M.; Woodson, G.; Levine, A. B.; Constantine, G.; Delmas, P. D. Effects of Bazedoxifene on BMD and Bone Turnover in Postmenopausal Women: 2-Yr Results of a Randomized, Double-Blind, Placebo-, and Active-Controlled Study. *J. Bone Miner. Res.* **2008**, *23* (4), 525–535. https://doi.org/10.1359/jbmr.071206.

(39) Wakeling, A. E.; Bowler, J. ICI 182,780, a New Antioestrogen with Clinical Potential. *J. Steroid Biochem. Mol. Biol.* **1992**, *43* (1–3), 173–177. https://doi.org/10.1016/0960-0760(92)90204-V.

(40) Wang, G. Fulvestrant as a Reference Antiestrogen and Estrogen Receptor (ER) Degrader in Preclinical Studies: Treatment Dosage, Efficacy, and Implications on Development of New ER-Targeting Agents. *Transl. Cancer Res.* **2020**, *9* (8), 4464–4468. https://doi.org/10.21037/tcr-20-2166.

(41) Lin, X.; Xiang, H.; Luo, G. Targeting Estrogen Receptor α for Degradation with PROTACs: A Promising Approach to Overcome Endocrine Resistance. *Eur. J. Med. Chem.* **2020**, *206*, 112689. https://doi.org/10.1016/j.ejmech.2020.112689.

(42) Ehrenkranz, J. R. L.; Lewis, N. G.; Ronald Kahn, C.; Roth, J. Phlorizin: A Review. *Diabetes. Metab. Res. Rev.* **2005**, *21* (1), 31–38. https://doi.org/10.1002/dmrr.532.

(43) Keller, D. M.; Lotspeich, W. D. Effect of Phlorizin on the Osmotic Behavior of Mitochondria in Isotonic Sucrose. *J. Biol. Chem.* **1959**, *234* (4), 991–994. https://doi.org/10.1016/S0021-9258(18)70217-2.

(44) Alvarado, F.; Crane, R. K. Phlorizin as a Competitive Inhibitor of the Active Transport of Sugars by Hamster Small Intestine, in Vitro. *Biochim. Biophys. Acta* **1962**, *56*, 170–172. https://doi.org/10.1016/0006-3002(62)90543-7.

(45) Kanai, Y.; Lee, W. S.; You, G.; Brown, D.; Hediger, M. A. The Human Kidney Low Affinity Na+/Glucose Cotransporter SGLT2. Delineation of the Major Renal Reabsorptive Mechanism for D-Glucose. *J. Clin. Invest.* **1994**, *93* (1), 397–404. https://doi.org/10.1172/JCI116972.

(46) Bisignano, P.; Ghezzi, C.; Jo, H.; Polizzi, N. F.; Althoff, T.; Kalyanaraman, C.; Friemann, R.; Jacobson, M. P.; Wright, E. M.; Grabe, M. Inhibitor Binding Mode and Allosteric Regulation of Na+-Glucose Symporters. *Nat. Commun.* **2018**, *9* (1), 5245. https://doi.org/10.1038/s41467-018-07700-1.

(47) Turk, E.; Zabel, B.; Mundlos, S.; Dyer, J.; Wright, E. M. Glucose/Galactose Malabsorption Caused by a Defect in the Na+/Glucose Cotransporter. *Nature* **1991**, *350* (6316), 354–356. https://doi.org/10.1038/350354a0.

(48) Washburn, W. N. Development of the Renal Glucose Reabsorption Inhibitors: A New Mechanism for the Pharmacotherapy of Diabetes Mellitus Type 2. *J. Med. Chem.* **2009**, *52* (7), 1785–1794. https://doi.org/10.1021/jm8013019.

(49) Goodwin, N. C.; Mabon, R.; Harrison, B. A.; Shadoan, M. K.; Almstead, Z. Y.; Xie, Y.; Healy, J.; Buhring, L. M.; DaCosta, C. M.; Bardenhagen, J.; Mseeh, F.; Liu, Q.; Nouraldeen, A.; Wilson, A. G. E.; Kimball, S. D.; Powell, D. R.; Rawlins, D. B. Novel L-Xylose Derivatives as Selective Sodium-Dependent Glucose Cotransporter 2 (SGLT2) Inhibitors for the Treatment of Type 2 Diabetes. *J. Med. Chem.* **2009**, *52* (20), 6201–6204. https://doi.org/10.1021/jm900951n.

(50) Cai, W.; Jiang, L.; Xie, Y.; Liu, Y.; Liu, W.; Zhao, G. Design of SGLT2 Inhibitors for the Treatment of Type 2 Diabetes: A History Driven by Biology to Chemistry. *Med. Chem. (Los. Angeles).* **2015**, *11* (4), 317–328. https://doi.org/10.2174/1573406411666150105105529.

(51) TSUJIHARA, K.; HONGU, M.; SAITO, K.; INAMASU, M.; ARAKAWA, K.; OKU, A.; MATSUMOTO, M. Na+-Glucose Cotransporter Inhibitors as Antidiabetics. I. Synthesis and Pharmacological Properties of 4'-Dehydroxyphlorizin Derivatives Based on a New Concept.

Chem. Pharm. Bull. **1996**, *44* (6), 1174–1180. https://doi.org/10.1248/cpb.44.1174.

(52) Sadurní, A.; Kehr, G.; Ahlqvist, M.; Wernevik, J.; Sjögren, H. P.; Kankkonen, C.; Knerr, L.; Gilmour, R. Fluorine-Directed Glycosylation Enables the Stereocontrolled Synthesis of Selective SGLT2 Inhibitors for Type II Diabetes. *Chem. – A Eur. J.* **2018**, *24* (12), 2832–2836. https://doi.org/10.1002/chem.201705373.

(53) Jesus, A. R.; Vila-Viçosa, D.; Machuqueiro, M.; Marques, A. P.; Dore, T. M.; Rauter, A. P. Targeting Type 2 Diabetes with C -Glucosyl Dihydrochalcones as Selective Sodium Glucose Co-Transporter 2 (SGLT2) Inhibitors: Synthesis and Biological Evaluation. *J. Med. Chem.* **2017**, *60* (2), 568–579. https://doi.org/10.1021/acs.jmedchem.6b01134.

(54) Zambrowicz, B.; Freiman, J.; Brown, P. M.; Frazier, K. S.; Turnage, A.; Bronner, J.; Ruff, D.; Shadoan, M.; Banks, P.; Mseeh, F.; Rawlins, D. B.; Goodwin, N. C.; Mabon, R.; Harrison, B. A.; Wilson, A.; Sands, A.; Powell, D. R. LX4211, a Dual SGLT1/SGLT2 Inhibitor, Improved Glycemic Control in Patients With Type 2 Diabetes in a Randomized, Placebo-Controlled Trial. *Clin. Pharmacol. Ther.* **2012**, *92* (2), 158–169. https://doi.org/10.1038/clpt.2012.58.

(55) Bartenschlager, R. The NS3/4A Proteinase of the Hepatitis C Virus: Unravelling Structure and Function of an Unusual Enzyme and a Prime Target for Antiviral Therapy. *J. Viral Hepat.* **1999**, *6* (3), 165–181. https://doi.org/10.1046/j.1365-2893.1999.00152.x.

(56) Arasappan, A.; Njoroge, F. G.; Chan, T.-Y.; Bennett, F.; Bogen, S. L.; Chen, K.; Gu, H.; Hong, L.; Jao, E.; Liu, Y.-T.; Lovey, R. G.; Parekh, T.; Pike, R. E.; Pinto, P.; Santhanam, B.; Venkatraman, S.; Vaccaro, H.; Wang, H.; Yang, X.; Zhu, Z.; Mckittrick, B.; Saksena, A. K.; Girijavallabhan, V.; Pichardo, J.; Butkiewicz, N.; Ingram, R.; Malcolm, B.; Prongay, A.; Yao, N.; Marten, B.; Madison, V.; Kemp, S.; Levy, O.; Lim-Wilby, M.; Tamura, S.; Ganguly, A. K. Hepatitis C Virus NS3-4A Serine Protease Inhibitors: SAR of Moiety with Improved Potency. *Bioorg. Med. Chem. Lett.* **2005**, *15* (19), 4180–4184. https://doi.org/10.1016/j.bmcl.2005.06.091.

(57) Bogen, S. L.; Arasappan, A.; Bennett, F.; Chen, K.; Jao, E.; Liu, Y.-T.; Lovey, R. G.; Venkatraman, S.; Pan, W.; Parekh, T.; Pike, R. E.; Ruan, S.; Liu, R.; Baroudy, B.; Agrawal, S.; Chase, R.; Ingravallo, P.; Pichardo, J.; Prongay, A.; Brisson, J.-M.; Hsieh, T. Y.; Cheng, K.-C.; Kemp, S. J.; Levy, O. E.; Lim-Wilby, M.; Tamura, S. Y.; Saksena, A. K.; Girijavallabhan, V.; Njoroge, F. G. Discovery of SCH446211 (SCH6): A New Ketoamide Inhibitor of the HCV NS3 Serine Protease and HCV Subgenomic RNA Replication. *J. Med. Chem.* **2006**, *49* (9), 2750–2757. https://doi.org/10.1021/jm060077j.

(58) International Union of Biochemistry and Molecular Biology. *EC 3 Introduction, EC 3 Hydrolases, EC 3.4 Acting on Peptide Bonds (Peptidases)*. https://iubmb.qmul.ac.uk/enzyme/EC3/intro.html#1.

(59) Barrett, A. J.; Tolle, D. P.; Rawlings, N. D. Managing Peptidases in the Genomic Era. *Biol. Chem.* **2003**, *384* (6). https://doi.org/10.1515/BC.2003.098.

(60) Nomura, S.; Ito, T.; Yamamoto, E.; Sumigama, S.; Iwase, A.; Okada, M.; Shibata, K.; Ando, H.; Ino, K.; Kikkawa, F.; Mizutani, S. Gene Regulation and Physiological Function of Placental Leucine Aminopeptidase/Oxytocinase during Pregnancy. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1751* (1), 19–25. https://doi.org/10.1016/j.bbapap.2005.04.006.

(61) Moss, G. P.; Smith, P. A. S.; Tavernier, D. Glossary of Class Names of Organic Compounds and Reactivity Intermediates Based on Structure (IUPAC Recommendations 1995). *Pure Appl. Chem.* **1995**, *67* (8–9), 1307–1375. https://doi.org/10.1351/pac199567081307.

(62) López-Otín, C.; Bond, J. S. Proteases: Multifunctional Enzymes in Life and Disease. *J. Biol. Chem.* **2008**, *283* (45), 30433–30437. https://doi.org/10.1074/jbc.R800035200.

(63) Neurath, H. Evolution of Proteolytic Enzymes. *Science (80-.).* **1984**, *224* (4647), 350–357. https://doi.org/10.1126/science.6369538.

(64) Neurath, H.; Walsh, K. A.; Winter, W. P. Evolution of Structure and Function of Proteases.

Science (80-.). **1967**, *158* (3809), 1638–1644. https://doi.org/10.1126/science.158.3809.1638.

(65) Neurath, H.; Bradshaw, R. A. Evolution of Proteolytic Function. *Acc. Chem. Res.* **1970**, *3* (8), 249–257. https://doi.org/10.1021/ar50032a001.

(66) Cen, Y.; Singh, W.; Arkin, M.; Moody, T. S.; Huang, M.; Zhou, J.; Wu, Q.; Reetz, M. T. Artificial Cysteine-Lipases with High Activity and Altered Catalytic Mechanism Created by Laboratory Evolution. *Nat. Commun.* **2019**, *10* (1), 3198. https://doi.org/10.1038/s41467-019-11155-3.

(67) Rawlings, N. D.; Barrett, A. J. Evolutionary Families of Peptidases. *Biochem. J.* **1993**, *290* (1), 205–218. https://doi.org/10.1042/bj2900205.

(68) Rawlings, N. D.; Barrett, A. J. Chapter 1 - Introduction: Aspartic and Glutamic Peptidases and Their Clans. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 3–19. https://doi.org/10.1016/B978-0-12-382219-2.00001-6.

(69) Rawlings, N. D.; Barrett, A. J. Chapter 77 - Introduction: Metallopeptidases and Their Clans. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 325–370. https://doi.org/10.1016/B978-0-12-382219-2.00077-6.

(70) Rawlings, N. D.; Barrett, A. J. Chapter 404 - Introduction: The Clans and Families of Cysteine Peptidases. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 1743–1773. https://doi.org/10.1016/B978-0-12-382219-2.00404-X.

(71) Rawlings, N. D.; Barrett, A. J. Chapter 559 - Introduction: Serine Peptidases and Their Clans. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 2491–2523. https://doi.org/10.1016/B978-0-12-382219-2.00559-7.

(72) McDonald, A. G.; Boyce, S.; Tipton, K. F. ExplorEnz: The Primary Source of the IUBMB Enzyme List. *Nucleic Acids Res.* **2009**, *37* (Database), D593–D597. https://doi.org/10.1093/nar/gkn582.

(73) Rawlings, N. D.; Barrett, A. J.; Thomas, P. D.; Huang, X.; Bateman, A.; Finn, R. D. The MEROPS Database of Proteolytic Enzymes, Their Substrates and Inhibitors in 2017 and a Comparison with Peptidases in the PANTHER Database. *Nucleic Acids Res.* **2018**, *46* (D1), D624–D632. https://doi.org/10.1093/nar/gkx1134.

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

(75) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlin, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630* (8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w.

(76) 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.

(77) 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.

(78) Pérez-Silva, J. G.; Español, Y.; Velasco, G.; Quesada, V. The Degradome Database: Expanding Roles of Mammalian Proteases in Life and Disease. *Nucleic Acids Res.* **2016**, *44* (D1), D351–D355. https://doi.org/10.1093/nar/gkv1201.

(79) The Mammalian Degradome Database. *Human, Mouse and Rat Degradomes in Numbers*. https://degradome.uniovi.es/numbers.html.

(80) López-Otín, C.; Matrisian, L. M. Emerging Roles of Proteases in Tumour Suppression. *Nat. Rev. Cancer* **2007**, *7* (10), 800–808. https://doi.org/10.1038/nrc2228.

(81) Berti, P. J.; Storer, A. C. Alignment/Phylogeny of the Papain Superfamily of Cysteine Proteases. *J. Mol. Biol.* **1995**, *246* (2), 273–283. https://doi.org/10.1006/jmbi.1994.0083.

(82) Lecaille, F.; Chazeirat, T.; Saidi, A.; Lalmanach, G. Cathepsin V: Molecular Characteristics and Significance in Health and Disease. *Mol. Aspects Med.* **2022**, *88*, 101086. https://doi.org/10.1016/j.mam.2022.101086.

(83) Schechter, I.; Berger, A. On the Size of the Active Site in Proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27* (2), 157–162. https://doi.org/10.1016/S0006-291X(67)80055-X.

(84) Wlodawer, A.; Gustchina, A.; James, M. N. G. Chapter 2 - Catalytic Pathways of Aspartic Peptidases. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 19–26. https://doi.org/10.1016/B978-0-12-382219-2.00002-8.

(85) Auld, D. S. Chapter 78 - Catalytic Mechanisms for Metallopeptidases. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 370–396. https://doi.org/10.1016/B978-0-12-382219-2.00078-8.

(86) Polgár, L. Chapter 405 - Catalytic Mechanisms of Cysteine Peptidases. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 1773–1784. https://doi.org/10.1016/B978-0-12-382219-2.00405-1.

(87) Polgár, L. Chapter 560 - Catalytic Mechanisms of Serine and Threonine Peptidases. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 2524–2534. https://doi.org/10.1016/B978-0-12-382219-2.00560-3.

(88) Zheng, Y.; Cui, J.; Wang, Y.; Huang, S.; Lin, E.; Huang, S.; Rudolf, J. D.; Yan, X.; Chang, C. The Structure-Function Relationship of Human Bleomycin Hydrolase: Mutation of a Cysteine Protease into a Serine Protease. *ChemBioChem* **2022**, *23* (12). https://doi.org/10.1002/cbic.202200186.

(89) Czapinska, H.; Bochtler, M. The Nϵ-Rule for Serine, but Not Cysteine Catalytic Triads. *Angew. Chemie* **2022**, *134* (42). https://doi.org/10.1002/ange.202206945.

(90) Harris, T. K.; Turner, G. J. Structural Basis of Perturbed PKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life* **2002**, *53* (2), 85–98. https://doi.org/10.1080/15216540211468.

(91) Pinitglang, S.; Watts, A. B.; Patel, M.; Reid, J. D.; Noble, M. A.; Gul, S.; Bokth, A.; Naeem, A.; Patel, H.; Thomas, E. W.; Sreedharan, S. K.; Verma, C.; Brocklehurst, K. A Classical Enzyme Active Center Motif Lacks Catalytic Competence until Modulated Electrostatically. *Biochemistry* **1997**, *36* (33), 9968–9982. https://doi.org/10.1021/bi9705974.

(92) Dodson, G. Catalytic Triads and Their Relatives. *Trends Biochem. Sci.* **1998**, *23* (9), 347–352. https://doi.org/10.1016/S0968-0004(98)01254-7.

(93) Bruice, T. C.; Fife, T. H.; Bruno, J. J.; Brandon, N. E. Hydroxyl Group Catalysis. II. The

Reactivity of the Hydroxyl Group of Serine. The Nucleophilicity of Alcohols and the Ease of Hydrolysis of Their Acetyl Esters as Related to Their PKa' *. *Biochemistry* **1962**, *1* (1), 7–12. https://doi.org/10.1021/bi00907a002.

(94) Lin, J.; Cassidy, C. S.; Frey, P. A. Correlations of the Basicity of His 57 with Transition State Analogue Binding, Substrate Reactivity, and the Strength of the Low-Barrier Hydrogen Bond in Chymotrypsin. *Biochemistry* **1998**, *37* (34), 11940–11948. https://doi.org/10.1021/bi980278s.

(95) Heinemeyer, W.; Fischer, M.; Krimmer, T.; Stachon, U.; Wolf, D. H. The Active Sites of the Eukaryotic 20 S Proteasome and Their Involvement in Subunit Precursor Processing. *J. Biol. Chem.* **1997**, *272* (40), 25200–25209. https://doi.org/10.1074/jbc.272.40.25200.

(96) Schrader, J.; Henneberg, F.; Mata, R. A.; Tittmann, K.; Schneider, T. R.; Stark, H.; Bourenkov, G.; Chari, A. The Inhibition Mechanism of Human 20 S Proteasomes Enables Next-Generation Inhibitor Design. *Science (80-.).* **2016**, *353* (6299), 594–598. https://doi.org/10.1126/science.aaf8993.

(97) Serrano-Aparicio, N.; Moliner, V.; Świderek, K. Nature of Irreversible Inhibition of Human 20S Proteasome by Salinosporamide A. The Critical Role of Lys–Asp Dyad Revealed from Electrostatic Effects Analysis. *ACS Catal.* **2021**, *11* (6), 3575–3589. https://doi.org/10.1021/acscatal.0c05313.

(98) Saha, A.; Oanca, G.; Mondal, D.; Warshel, A. Exploring the Proteolysis Mechanism of the Proteasomes. *J. Phys. Chem. B* **2020**, *124* (27), 5626–5635. https://doi.org/10.1021/acs.jpcb.0c04435.

(99) Löwe, J.; Stock, D.; Jap, B.; Zwickl, P.; Baumeister, W.; Huber, R. Crystal Structure of the 20 S Proteasome from the Archaeon T. Acidophilum at 3.4 Å Resolution. *Science (80-.).* **1995**, *268* (5210), 533–539. https://doi.org/10.1126/science.7725097.

(100) Zeslawska, E.; Jacob, U.; Schweinitz, A.; Coombs, G.; Bode, W.; Madison, E. Crystals of Urokinase Type Plasminogen Activator Complexes Reveal the Binding Mode of Peptidomimetic Inhibitors. *J. Mol. Biol.* **2003**, *328* (1), 109–118. https://doi.org/10.1016/S0022-2836(03)00267-5.

(101) Toste Rêgo, A.; da Fonseca, P. C. A. Characterization of Fully Recombinant Human 20S and 20S-PA200 Proteasome Complexes. *Mol. Cell* **2019**, *76* (1), 138-147.e5. https://doi.org/10.1016/j.molcel.2019.07.014.

(102) Messerli, F. H.; Williams, B.; Ritz, E. Essential Hypertension. *Lancet* **2007**, *370* (9587), 591–603. https://doi.org/10.1016/S0140-6736(07)61299-9.

(103) Powers, B. J.; Coeytaux, R. R.; Dolor, R. J.; Hasselblad, V.; Patel, U. D.; Yancy, W. S.; Gray, R. N.; Irvine, R. J.; Kendrick, A. S.; Sanders, G. D. Updated Report on Comparative Effectiveness of ACE Inhibitors, ARBs, and Direct Renin Inhibitors for Patients with Essential Hypertension: Much More Data, Little New Information. *J. Gen. Intern. Med.* **2012**, *27* (6), 716–729. https://doi.org/10.1007/s11606-011-1938-8.

(104) Isabel, C.; Calvet, D.; Mas, J.-L. Stroke Prevention. *Presse Med.* **2016**, *45* (12), e457–e471. https://doi.org/10.1016/j.lpm.2016.10.009.

(105) Li, G.; Lip, G. Y. H.; Holbrook, A.; Chang, Y.; Larsen, T. B.; Sun, X.; Tang, J.; Mbuagbaw, L.; Witt, D. M.; Crowther, M.; Thabane, L.; Levine, M. A. H. Direct Comparative Effectiveness and Safety between Non-Vitamin K Antagonist Oral Anticoagulants for Stroke Prevention in Nonvalvular Atrial Fibrillation: A Systematic Review and Meta-Analysis of Observational Studies. *Eur. J. Epidemiol.* **2019**, *34* (2), 173–190. https://doi.org/10.1007/s10654-018-0415-7.

(106) Deacon, C. F.; Holst, J. J. Dipeptidyl Peptidase-4 Inhibitors for the Treatment of Type 2 Diabetes: Comparison, Efficacy and Safety. *Expert Opin. Pharmacother.* **2013**, *14* (15), 2047–2058. https://doi.org/10.1517/14656566.2013.824966.

(107) Craddy, P.; Palin, H.-J.; Johnson, K. I. Comparative Effectiveness of Dipeptidylpeptidase-4 Inhibitors in Type 2 Diabetes: A Systematic Review and Mixed Treatment Comparison. *Diabetes Ther.* **2014**, *5* (1), 1–41. https://doi.org/10.1007/s13300-014-0061-3.

(108) Shah, J. J.; Orlowski, R. Z. Proteasome Inhibitors in the Treatment of Multiple Myeloma. *Leukemia* **2009**, *23* (11), 1964–1979. https://doi.org/10.1038/leu.2009.173.

(109) Manasanch, E. E.; Orlowski, R. Z. Proteasome Inhibitors in Cancer Therapy. *Nat. Rev. Clin. Oncol.* **2017**, *14* (7), 417–433. https://doi.org/10.1038/nrclinonc.2016.206.

(110) International Committee on Taxonomy of Viruses. *History of the Taxon: Hepacivirus hominis*. https://ictv.global/taxonomy/taxondetails?taxnode_id=202303127&taxon_name=Hepacivirus hominis.

(111) International Committee on Taxonomy of Viruses. *History of the Taxon: Lentivirus humimdef1*.

(112) International Committee on Taxonomy of Viruses. *History of the Taxon: Betacoronavirus pandemicum*.

(113) Agbowuro, A. A.; Huston, W. M.; Gamble, A. B.; Tyndall, J. D. A. Proteases and Protease Inhibitors in Infectious Diseases. *Med. Res. Rev.* **2018**, *38* (4), 1295–1331. https://doi.org/10.1002/med.21475.

(114) Chou, R.; Hartung, D.; Rahman, B.; Wasson, N.; Cottrell, E. B.; Fu, R. Comparative Effectiveness of Antiviral Treatment for Hepatitis C Virus Infection in Adults: A Systematic Review. *Ann. Intern. Med.* **2013**, *158* (2), 114. https://doi.org/10.7326/0003-4819-158-2-201301150-00576.

(115) Paton, N. I.; Musaazi, J.; Kityo, C.; Walimbwa, S.; Hoppe, A.; Balyegisawa, A.; Asienzo, J.; Kaimal, A.; Mirembe, G.; Lugemwa, A.; Ategeka, G.; Borok, M.; Mugerwa, H.; Siika, A.; Odongpiny, E. L. A.; Castelnuovo, B.; Kiragga, A.; Kambugu, A.; Kambugu, A.; Kaimal, A.; Castelnuovo, B.; Kiiza, D.; Asienzo, J.; Kisembo, J.; Nsubuga, J.; Okwero, M.; Muyise, R.; Kityo, C.; Nasaazi, C.; Nakiboneka, D. L.; Mugerwa, H.; Namusanje, J.; Najjuuko, T.; Masaba, T.; Serumaga, T.; Alinaitwe, A.; Arinda, A.; Rweyora, A.; Ategeka, G.; Kangah, M. G.; Lugemwa, A.; Kasozi, M.; Tukumushabe, P.; Akunda, R.; Makumbi, S.; Musumba, S.; Myalo, S.; Ahuura, J.; Namusisi, A. M.; Kibirige, D.; Kiweewa, F.; Mirembe, G.; Mabonga, H.; Wandege, J.; Nakakeeto, J.; Namubiru, S.; Nansalire, W.; Siika, A. M.; Kwobah, C. M.; Mboya, C. S.; Mokaya, M. M. B.; Karoney, M. J.; Cheruiyot, P. C.; Cherutich, S.; Njuguna, S. W.; Kirui, V. C.; Borok, M.; Chidziva, E.; Musoro, G.; Hakim, J.; Bhiri, J.; Phiri, M.; Mudzingwa, S.; Manyanga, T.; Kiragga, A.; Banegura, A. M.; Hoppe, A.; Balyegisawa, A.; Agwang, B.; Isaaya, B.; Tumwine, C.; Odongpiny, E. L. A.; Asienzo, J.; Musaazi, J.; Paton, N.; Senkungu, P.; Walimbwa, S.; Kamara, Y.; Amperiize, M.; Allen, E.; Opondo, C.; Mohammed, P.; van Rein-van der Horst, W.; Van Delft, Y.; Boateng, F. A.; Namara, D.; Kaleebu, P.; Ojoo, S.; Bwakura, T.; Katana, M.; Venter, F.; Phiri, S.; Walker, S. Efficacy and Safety of Dolutegravir or Darunavir in Combination with Lamivudine plus Either Zidovudine or Tenofovir for Second-Line Treatment of HIV Infection (NADIA): Week 96 Results from a Prospective, Multicentre, Open-Label, Factorial, Randomised, Non. *Lancet HIV* **2022**, *9* (6), e381–e393. https://doi.org/10.1016/S2352-3018(22)00092-3.

(116) Amani, B.; Amani, B. Efficacy and Safety of Nirmatrelvir/Ritonavir (Paxlovid) for COVID-19: A Rapid Review and Meta-analysis. *J. Med. Virol.* **2023**, *95* (2). https://doi.org/10.1002/jmv.28441.

(117) Kawato, Y.; Fukahori, H.; Nakamura, K.; Kanno, A.; Kubo, K.; Hiramitsu, M.; Matsuda, T.; Hanada, Y.; Furukawa, T.; Nakajima, Y.; Kinugasa, F.; Morokata, T. Potential Benefit of the Cathepsin S Inhibitor, ASP1617, as a Treatment for Systemic Lupus Erythematosus. *Eur. J. Pharmacol.* **2022**, *919*, 174826. https://doi.org/10.1016/j.ejphar.2022.174826.

(118) Bentley, D.; Fisher, B. A.; Barone, F.; Kolb, F. A.; Attley, G. A Randomized, Double-Blind, Placebo-Controlled, Parallel Group Study on the Effects of a Cathepsin S Inhibitor in Primary

Sjögren's Syndrome. *Rheumatology* **2023**, *62* (11), 3644–3653. https://doi.org/10.1093/rheumatology/kead092.

(119) Hook, G.; Hook, V.; Kindy, M. The Cysteine Protease Inhibitor, E64d, Reduces Brain Amyloid-β and Improves Memory Deficits in Alzheimer's Disease Animal Models by Inhibiting Cathepsin B, but Not BACE1, β-Secretase Activity. *J. Alzheimer's Dis.* **2011**, *26* (2), 387–408. https://doi.org/10.3233/JAD-2011-110101.

(120) Bone, H. G.; McClung, M. R.; Roux, C.; Recker, R. R.; Eisman, J. A.; Verbruggen, N.; Hustad, C. M.; DaSilva, C.; Santora, A. C.; Ince, B. A. Odanacatib, a Cathepsin-K Inhibitor for Osteoporosis: A Two-Year Study in Postmenopausal Women with Low Bone Density. *J. Bone Miner. Res.* **2010**, *25* (5), 937–947. https://doi.org/10.1359/jbmr.091035.

(121) dos Santos Nascimento, I. J.; Albino, S. L.; da Silva Menezes, K. J.; de Azevedo Teotônio Cavalcanti, M.; de Oliveira, M. S.; Mali, S. N.; de Moura, R. O. Targeting SmCB1: Perspectives and Insights to Design Antischistosomal Drugs. *Curr. Med. Chem.* **2024**, *31* (16), 2264–2284. https://doi.org/10.2174/0109298673255826231011114249.

(122) Chiba, D. E.; dos Santos Fernandes, G. F.; dos Santos, J. L.; Scarim, C. B. Exploring the Latest Breakthroughs in Rhodesain Inhibitors for African Trypanosomiasis. *Med. Chem. Res.* **2024**, *33* (3), 354–369. https://doi.org/10.1007/s00044-024-03189-0.

(123) Rodriguez-Pena, R.; Antunez, C.; Martin, E.; Blanca-Lopez, N.; Mayorga, C.; Torres, M. J. Allergic Reactions to β-Lactams. *Expert Opin. Drug Saf.* **2006**, *5* (1), 31–48. https://doi.org/10.1517/14740338.5.1.31.

(124) Sutanto, F.; Konstantinidou, M.; Dömling, A. Covalent Inhibitors: A Rational Approach to Drug Discovery. *RSC Med. Chem.* **2020**, *11* (8), 876–884. https://doi.org/10.1039/D0MD00154F.

(125) Schwöbel, J. A. H.; Koleva, Y. K.; Enoch, S. J.; Bajot, F.; Hewitt, M.; Madden, J. C.; Roberts, D. W.; Schultz, T. W.; Cronin, M. T. D. Measurement and Estimation of Electrophilic Reactivity for Predictive Toxicology. *Chem. Rev.* **2011**, *111* (4), 2562–2596. https://doi.org/10.1021/cr100098n.

(126) Boike, L.; Henning, N. J.; Nomura, D. K. Advances in Covalent Drug Discovery. *Nat. Rev. Drug Discov.* **2022**, *21* (12), 881–898. https://doi.org/10.1038/s41573-022-00542-z.

(127) Lonsdale, R.; Ward, R. A. Structure-Based Design of Targeted Covalent Inhibitors. *Chem. Soc. Rev.* **2018**, *47* (11), 3816–3830. https://doi.org/10.1039/C7CS00220C.

(128) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug–Target Residence Time and Its Implications for Lead Optimization. *Nat. Rev. Drug Discov.* **2006**, *5* (9), 730–739. https://doi.org/10.1038/nrd2082.

(129) Kim, H.; Hwang, Y. S.; Kim, M.; Park, S. B. Recent Advances in the Development of Covalent Inhibitors. *RSC Med. Chem.* **2021**, *12* (7), 1037–1045. https://doi.org/10.1039/D1MD00068C.

(130) Pearson, R. G. Hard and Soft Acids and Bases—the Evolution of a Chemical Concept. *Coord. Chem. Rev.* **1990**, *100*, 403–425. https://doi.org/10.1016/0010-8545(90)85016-L.

(131) Jaramillo, P.; Pérez, P.; Contreras, R.; Tiznado, W.; Fuentealba, P. Definition of a Nucleophilicity Scale. *J. Phys. Chem. A* **2006**, *110* (26), 8181–8187. https://doi.org/10.1021/jp057351q.

(132) LoPachin, R. M.; Gavin, T.; DeCaprio, A.; Barber, D. S. Application of the Hard and Soft, Acids and Bases (HSAB) Theory to Toxicant–Target Interactions. *Chem. Res. Toxicol.* **2012**, *25* (2), 239–251. https://doi.org/10.1021/tx2003257.

(133) Ertl, P.; Gerebtzoff, G.; Lewis, R.; Muenkler, H.; Schneider, N.; Sirockin, F.; Stiefl, N.; Tosco, P. Chemical Reactivity Prediction: Current Methods and Different Application Areas. *Mol. Inform.* **2022**, *41* (6). https://doi.org/10.1002/minf.202100277.

(134) Mihalovits, L. M.; Ferenczy, G. G.; Keserű, G. M. The Role of Quantum Chemistry in Covalent Inhibitor Design. *Int. J. Quantum Chem.* **2022**, *122* (8). https://doi.org/10.1002/qua.26768.

(135) LoPachin, R. M.; Geohagen, B. C.; Nordstroem, L. U. Mechanisms of Soft and Hard Electrophile Toxicities. *Toxicology* **2019**, *418*, 62–69. https://doi.org/10.1016/j.tox.2019.02.005.

(136) LoPachin, R. M.; Gavin, T. Reactions of Electrophiles with Nucleophilic Thiolate Sites: Relevance to Pathophysiological Mechanisms and Remediation. *Free Radic. Res.* **2016**, *50* (2), 195–205. https://doi.org/10.3109/10715762.2015.1094184.

(137) Grabowsky, S.; Pfeuffer, T.; Chęcińska, L.; Weber, M.; Morgenroth, W.; Luger, P.; Schirmeister, T. Electron-Density Determination of Electrophilic Building Blocks as Model Compounds for Protease Inhibitors. *European J. Org. Chem.* **2007**, *2007* (17), 2759–2768. https://doi.org/10.1002/ejoc.200601074.

(138) LoPachin, R. M.; Gavin, T. Molecular Mechanisms of Aldehyde Toxicity: A Chemical Perspective. *Chem. Res. Toxicol.* **2014**, *27* (7), 1081–1091. https://doi.org/10.1021/tx5001046.

(139) Löser, R.; Frizler, M.; Schilling, K.; Gütschow, M. Azadipeptide Nitriles: Highly Potent and Proteolytically Stable Inhibitors of Papain-Like Cysteine Proteases. *Angew. Chemie Int. Ed.* **2008**, *47* (23), 4331–4334. https://doi.org/10.1002/anie.200705858.

(140) Ho, T.-L.; Ho, H. C.; Hamilton, L. D. Biochemical Significance of the Hard and Soft Acids and Bases Principle. *Chem. Biol. Interact.* **1978**, *23* (1), 65–84. https://doi.org/10.1016/0009-2797(78)90042-X.

(141) Adams, J.; Behnke, M.; Chen, S.; Cruickshank, A. A.; Dick, L. R.; Grenier, L.; Klunder, J. M.; Ma, Y.-T.; Plamondon, L.; Stein, R. L. Potent and Selective Inhibitors of the Proteasome: Dipeptidyl Boronic Acids. *Bioorg. Med. Chem. Lett.* **1998**, *8* (4), 333–338. https://doi.org/10.1016/S0960-894X(98)00029-8.

(142) Groll, M.; Berkers, C. R.; Ploegh, H. L.; Ovaa, H. Crystal Structure of the Boronic Acid-Based Proteasome Inhibitor Bortezomib in Complex with the Yeast 20S Proteasome. *Structure* **2006**, *14* (3), 451–456. https://doi.org/10.1016/j.str.2005.11.019.

(143) Yin, Z.; Patel, S. J.; Wang, W.-L.; Wang, G.; Chan, W.-L.; Rao, K. R. R.; Alam, J.; Jeyaraj, D. A.; Ngew, X.; Patel, V.; Beer, D.; Lim, S. P.; Vasudevan, S. G.; Keller, T. H. Peptide Inhibitors of Dengue Virus NS3 Protease. Part 1: Warhead. *Bioorg. Med. Chem. Lett.* **2006**, *16* (1), 36–39. https://doi.org/10.1016/j.bmcl.2005.09.062.

(144) Tooke, C. L.; Hinchliffe, P.; Krajnc, A.; Mulholland, A. J.; Brem, J.; Schofield, C. J.; Spencer, J. Cyclic Boronates as Versatile Scaffolds for KPC-2 β-Lactamase Inhibition. *RSC Med. Chem.* **2020**, *11* (4), 491–496. https://doi.org/10.1039/C9MD00557A.

(145) Yang, H.; You, M.; Shu, X.; Zhen, J.; Zhu, M.; Fu, T.; Zhang, Y.; Jiang, X.; Zhang, L.; Xu, Y.; Zhang, Y.; Su, H.; Zhang, Q.; Shen, J. Design, Synthesis and Biological Evaluation of Peptidomimetic Benzothiazolyl Ketones as 3CLpro Inhibitors against SARS-CoV-2. *Eur. J. Med. Chem.* **2023**, *257*, 115512. https://doi.org/10.1016/j.ejmech.2023.115512.

(146) Adriaenssens, Y.; Jiménez Fernández, D.; Vande Walle, L.; Elvas, F.; Joossens, J.; Lambeir, A.; Augustyns, K.; Lamkanfi, M.; Van der Veken, P. Carboxylate Isosteres for Caspase Inhibitors: The Acylsulfonamide Case Revisited. *Org. Biomol. Chem.* **2017**, *15* (35), 7456–7473. https://doi.org/10.1039/C7OB01403A.

(147) Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Burkey, B. F.; Dunning, B. E.; Prasad, K.; Mangold, B. L.; Russell, M. E.; Hughes, T. E. 1-[[(3-Hydroxy-1-Adamantyl)Amino]Acetyl]-2-Cyano-(S)-Pyrrolidine: A Potent, Selective, and Orally Bioavailable Dipeptidyl Peptidase IV Inhibitor with Antihyperglycemic Properties. *J. Med. Chem.* **2003**, *46* (13), 2774–2789. https://doi.org/10.1021/jm0300911.

(148) Wettstein, L.; Knaff, P. M.; Kersten, C.; Müller, P.; Weil, T.; Conzelmann, C.; Müller, J. A.;

Brückner, M.; Hoffmann, M.; Pöhlmann, S.; Schirmeister, T.; Landfester, K.; Münch, J.; Mailänder, V. Peptidomimetic Inhibitors of TMPRSS2 Block SARS-CoV-2 Infection in Cell Culture. *Commun. Biol.* **2022**, *5* (1), 681. https://doi.org/10.1038/s42003-022-03613-4.

(149) Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; Hilgenfeld, R.; Yuen, K. Y.; Wong, L.; Gao, G.; Chen, S.; Chen, Z.; Ma, D.; Bartlam, M.; Rao, Z. Design of Wide-Spectrum Inhibitors Targeting Coronavirus Main Proteases. *PLoS Biol.* **2005**, *3* (10), e324. https://doi.org/10.1371/journal.pbio.0030324.

(150) Jílková, A.; Rubešová, P.; Fanfrlík, J.; Fajtová, P.; Řezáčová, P.; Brynda, J.; Lepšík, M.; Mertlíková-Kaiserová, H.; Emal, C. D.; Renslo, A. R.; Roush, W. R.; Horn, M.; Caffrey, C. R.; Mareš, M. Druggable Hot Spots in the Schistosomiasis Cathepsin B1 Target Identified by Functional and Binding Mode Analysis of Potent Vinyl Sulfone Inhibitors. *ACS Infect. Dis.* **2021**, *7* (5), 1077–1088. https://doi.org/10.1021/acsinfecdis.0c00501.

(151) DiMagno, S.; Sun, H. The Strength of Weak Interactions: Aromatic Fluorine in Drug Design. *Curr. Top. Med. Chem.* **2006**, *6* (14), 1473–1482. https://doi.org/10.2174/156802606777951127.

(152) Singer, R. D.; Oehlschlager, A. C. The Composition and Chemistry of the Mixed Higher-Order Cuprates (PhMe2Si)m(CH3)NCu(CN)Li(M+n). *J. Org. Chem.* **1991**, *56* (11), 3510–3514. https://doi.org/10.1021/jo00011a013.

(153) Lonsdale, R.; Burgess, J.; Colclough, N.; Davies, N. L.; Lenz, E. M.; Orton, A. L.; Ward, R. A. Expanding the Armory: Predicting and Tuning Covalent Warhead Reactivity. *J. Chem. Inf. Model.* **2017**, *57* (12), 3124–3137. https://doi.org/10.1021/acs.jcim.7b00553.

(154) Flanagan, M. E.; Abramite, J. A.; Anderson, D. P.; Aulabaugh, A.; Dahal, U. P.; Gilbert, A. M.; Li, C.; Montgomery, J.; Oppenheimer, S. R.; Ryder, T.; Schuff, B. P.; Uccello, D. P.; Walker, G. S.; Wu, Y.; Brown, M. F.; Chen, J. M.; Hayward, M. M.; Noe, M. C.; Obach, R. S.; Philippe, L.; Shanmugasundaram, V.; Shapiro, M. J.; Starr, J.; Stroh, J.; Che, Y. Chemical and Computational Methods for the Characterization of Covalent Reactive Groups for the Prospective Design of Irreversible Inhibitors. *J. Med. Chem.* **2014**, *57* (23), 10072–10079. https://doi.org/10.1021/jm501412a.

(155) Mitchell, J. R.; Jollow, D. J.; Potter, W. Z.; Gillette, J. R.; Brodie, B. B. Acetaminophen-Induced Hepatic Necrosis. IV. Protective Role of Glutathione. *J. Pharmacol. Exp. Ther.* **1973**, *187* (1), 211–217.

(156) Armstrong, R. N. Structure, Catalytic Mechanism, and Evolution of the Glutathione Transferases. *Chem. Res. Toxicol.* **1997**, *10* (1), 2–18. https://doi.org/10.1021/tx960072x.

(157) Mayer, R. J.; Ofial, A. R. Nucleophilicity of Glutathione: A Link to Michael Acceptor Reactivities. *Angew. Chemie Int. Ed.* **2019**, *58* (49), 17704–17708. https://doi.org/10.1002/anie.201909803.

(158) Bianchi, D. H. A. The Paradoxical Influence of the PKa on the Reactivity of Thiols and Its Biological Relevance. *MaRBLe* **2016**, *2*. https://doi.org/10.26481/marble.2014.v2.316.

(159) Ritchie, C. D.; Sawada, M. Cation-Anion Combination Reactions. 15. Rates of Nucleophilic Aromatic Substitution Reactions in Water and Methanol Solutions. *J. Am. Chem. Soc.* **1977**, *99* (11), 3754–3761.

(160) Resnick, E.; Bradley, A.; Gan, J.; Douangamath, A.; Krojer, T.; Sethi, R.; Geurink, P. P.; Aimon, A.; Amitai, G.; Bellini, D.; Bennett, J.; Fairhead, M.; Fedorov, O.; Gabizon, R.; Gan, J.; Guo, J.; Plotnikov, A.; Reznik, N.; Ruda, G. F.; Díaz-Sáez, L.; Straub, V. M.; Szommer, T.; Velupillai, S.; Zaidman, D.; Zhang, Y.; Coker, A. R.; Dowson, C. G.; Barr, H. M.; Wang, C.; Huber, K. V. M.; Brennan, P. E.; Ovaa, H.; von Delft, F.; London, N. Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *J. Am. Chem. Soc.* **2019**, *141* (22), 8951–8968. https://doi.org/10.1021/jacs.9b02822.

(161) Mayr, H.; Patz, M. Scales of Nucleophilicity and Electrophilicity: A System for Ordering Polar

Organic and Organometallic Reactions. *Angew. Chemie Int. Ed. English* **1994**, *33* (9), 938–957. https://doi.org/10.1002/anie.199409381.

(162) Tuley, A.; Fast, W. The Taxonomy of Covalent Inhibitors. *Biochemistry* **2018**, *57* (24), 3326–3337. https://doi.org/10.1021/acs.biochem.8b00315.

(163) Gehringer, M.; Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62* (12), 5673–5724. https://doi.org/10.1021/acs.jmedchem.8b01153.

(164) Mehta, N. V.; Degani, M. S. The Expanding Repertoire of Covalent Warheads for Drug Discovery. *Drug Discov. Today* **2023**, *28* (12), 103799. https://doi.org/10.1016/j.drudis.2023.103799.

(165) European Medicines Agency (EMA). *EMA Search Engine*. https://www.ema.europa.eu/en/search.

(166) Food and Drug Administration (FDA). *Drugs@FDA: FDA-approved Drugs*. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm.

(167) Bader, P. J.; Merck Sharp & Dohme Corp. *FDA Pharmacy Professional Information Letter: Merck Voluntarily Discontinuing VICTRELIS® (Boceprevir) 200 Mg Capsules*; 2015. https://www.fda.gov/files/drugs/published/Dear-Pharmacy-Professional-Letter.pdf.

(168) European Medicines Agency (EMA). *EMA/599647/2016, Public Statement: Incivo Expiry of the Marketing Authorisation in the European Union*; 2016.

(169) Health Canada. *Health Canada: Drug Product Database online query*. https://health-products.canada.ca/dpd-bdpp/.

(170) Gomes, L. O.; Teixeira, M. R.; Rosa, J. A. da; Feltrin, A. A.; Rodrigues, J. P. V.; Vecchi, M. D.; Carneiro, J. M. M.; Noblat, L. de A. C. B.; Chachá, S. G. F.; Martinelli, A. de L. C.; Pereira, L. R. L.; Silveira, M. P. T.; Blatt, C. R.; Farias, M. R. Hepatitis C in Brazil: Lessons Learned with Boceprevir and Telaprevir. *Rev. Inst. Med. Trop. Sao Paulo* **2018**, *60*. https://doi.org/10.1590/s1678-9946201860029.

(171) Pharmaceutical and Medical Devices Agency. *Pharmaceutical and Medical Devices Agency, Japan: List of Approved Products*. https://www.pmda.go.jp/english/review-services/reviews/approved-information/drugs/0002.html.

(172) Li, D.; Ambrogio, L.; Shimamura, T.; Kubo, S.; Takahashi, M.; Chirieac, L. R.; Padera, R. F.; Shapiro, G. I.; Baum, A.; Himmelsbach, F.; Rettig, W. J.; Meyerson, M.; Solca, F.; Greulich, H.; Wong, K.-K. BIBW2992, an Irreversible EGFR/HER2 Inhibitor Highly Effective in Preclinical Lung Cancer Models. *Oncogene* **2008**, *27* (34), 4702–4711. https://doi.org/10.1038/onc.2008.109.

(173) Blair, H. A. Voxelotor: First Approval. *Drugs* **2020**, *80* (2), 209–215. https://doi.org/10.1007/s40265-020-01262-7.

(174) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The Resurgence of Covalent Drugs. *Nat. Rev. Drug Discov.* **2011**, *10* (4), 307–317. https://doi.org/10.1038/nrd3410.

(175) Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J.; de Wilde, A.; Snijder, E. J.; Liu, H.; Hilgenfeld, R. α-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J. Med. Chem.* **2020**, *63* (9), 4562–4578. https://doi.org/10.1021/acs.jmedchem.9b01828.

(176) Rotella, D. P. The Discovery and Development of Boceprevir. *Expert Opin. Drug Discov.* **2013**, *8* (11), 1439–1447. https://doi.org/10.1517/17460441.2013.843525.

(177) Han, Z.; Harris, P. K. W.; Karmakar, P.; Kim, T.; Owusu, B. Y.; Wildman, S. A.; Klampfer, L.;

Janetka, J. W. A-Ketobenzothiazole Serine Protease Inhibitors of Aberrant HGF/C-MET and MSP/RON Kinase Pathway Signaling in Cancer. *ChemMedChem* **2016**, *11* (6), 585–599. https://doi.org/10.1002/cmdc.201500600.

(178) Costanzo, M. J.; Yabut, S. C.; Almond, H. R.; Andrade-Gordon, P.; Corcoran, T. W.; de Garavilla, L.; Kauffman, J. A.; Abraham, W. M.; Recacha, R.; Chattopadhyay, D.; Maryanoff, B. E. Potent, Small-Molecule Inhibitors of Human Mast Cell Tryptase. Antiasthmatic Action of a Dipeptide-Based Transition-State Analogue Containing a Benzothiazole Ketone. *J. Med. Chem.* **2003**, *46* (18), 3865–3876. https://doi.org/10.1021/jm030050p.

(179) Prorok, M.; Albeck, A.; Foxman, B. M.; Abeles, R. H. Chloroketone Hydrolysis by Chymotrypsin and N-Methylhistidyl-57-Chymotrypsin: Implications for the Mechanism of Chymotrypsin Inactivation by Chloroketones. *Biochemistry* **1994**, *33* (32), 9784–9790. https://doi.org/10.1021/bi00198a050.

(180) Kettner, C.; Shaw, E. Synthesis of Peptides of Arginine Chloromethyl Ketone. Selective Inactivation of Human Plasma Kallikrein. *Biochemistry* **1978**, *17* (22), 4778–4784. https://doi.org/10.1021/bi00615a027.

(181) Chan, A. H.; Lee, W.-G.; Spasov, K. A.; Cisneros, J. A.; Kudalkar, S. N.; Petrova, Z. O.; Buckingham, A. B.; Anderson, K. S.; Jorgensen, W. L. Covalent Inhibitors for Eradication of Drug-Resistant HIV-1 Reverse Transcriptase: From Design to Protein Crystallography. *Proc. Natl. Acad. Sci.* **2017**, *114* (36), 9725–9730. https://doi.org/10.1073/pnas.1711463114.

(182) Imperiali, B.; Abeles, R. H. Inhibition of Serine Proteases by Peptidyl Fluoromethyl Ketones. *Biochemistry* **1986**, *25* (13), 3760–3767. https://doi.org/10.1021/bi00361a005.

(183) Montgomery, J. A.; Thomas, H. J. Diazomethyl Ketone Derivatives of Pyrimidine Nucleosides. *J. Org. Chem.* **1981**, *46* (3), 594–598. https://doi.org/10.1021/jo00316a021.

(184) Watanabe, H.; Green, G. D. J.; Shaw, E. A Comparison of the Behavior of Chymotrypsin and Cathepsin B towards Peptidyl Diazomethyl Ketones. *Biochem. Biophys. Res. Commun.* **1979**, *89* (4), 1354–1360. https://doi.org/10.1016/0006-291X(79)92158-2.

(185) Tan, A. C.; Teh, Y. L.; Lai, G. G. Y.; Tan, D. S. W. Third Generation EGFR TKI Landscape for Metastatic EGFR Mutant Non-Small Cell Lung Cancer (NSCLC). *Expert Rev. Anticancer Ther.* **2019**, *19* (6), 431–435. https://doi.org/10.1080/14737140.2019.1604228.

(186) Ge, R.; Shen, Z.; Yin, J.; Chen, W.; Zhang, Q.; An, Y.; Tang, D.; Satz, A. L.; Su, W.; Kuai, L. Discovery of SARS-CoV-2 Main Protease Covalent Inhibitors from a DNA-Encoded Library Selection. *SLAS Discov.* **2022**, *27* (2), 79–85. https://doi.org/10.1016/j.slasd.2022.01.001.

(187) Augeri, D. J.; Robl, J. A.; Betebenner, D. A.; Magnin, D. R.; Khanna, A.; Robertson, J. G.; Wang, A.; Simpkins, L. M.; Taunk, P.; Huang, Q.; Han, S.-P.; Abboa-Offei, B.; Cap, M.; Xin, L.; Tao, L.; Tozzo, E.; Welzel, G. E.; Egan, D. M.; Marcinkeviciene, J.; Chang, S. Y.; Biller, S. A.; Kirby, M. S.; Parker, R. A.; Hamann, L. G. Discovery and Preclinical Profile of Saxagliptin (BMS-477118): A Highly Potent, Long-Acting, Orally Active Dipeptidyl Peptidase IV Inhibitor for the Treatment of Type 2 Diabetes. *J. Med. Chem.* **2005**, *48* (15), 5025–5037. https://doi.org/10.1021/jm050261p.

(188) Silva, D. G.; Ribeiro, J. F. R.; De Vita, D.; Cianni, L.; Franco, C. H.; Freitas-Junior, L. H.; Moraes, C. B.; Rocha, J. R.; Burtoloso, A. C. B.; Kenny, P. W.; Leitão, A.; Montanari, C. A. A Comparative Study of Warheads for Design of Cysteine Protease Inhibitors. *Bioorg. Med. Chem. Lett.* **2017**, *27* (22), 5031–5035. https://doi.org/10.1016/j.bmcl.2017.10.002.

(189) de Jesus Cortez, F.; Nguyen, P.; Truillet, C.; Tian, B.; Kuchenbecker, K. M.; Evans, M. J.; Webb, P.; Jacobson, M. P.; Fletterick, R. J.; England, P. M. Development of 5N-Bicalutamide, a High-Affinity Reversible Covalent Antiandrogen. *ACS Chem. Biol.* **2017**, *12* (12), 2934–2939. https://doi.org/10.1021/acschembio.7b00702.

(190) Page, M. I. The Mechanisms of Reactions of Beta-Lactam Antibiotics. *Acc. Chem. Res.* **1984**,

17 (4), 144–151. https://doi.org/10.1021/ar00100a005.

(191) Dražić, T.; Kopf, S.; Corridan, J.; Leuthold, M. M.; Bertoša, B.; Klein, C. D. Peptide-β-Lactam Inhibitors of Dengue and West Nile Virus NS2B-NS3 Protease Display Two Distinct Binding Modes. *J. Med. Chem.* **2020**, *63* (1), 140–156. https://doi.org/10.1021/acs.jmedchem.9b00759.

(192) Brandvold, K. R.; Miller, C. J.; Volk, R. F.; Killinger, B. J.; Whidbey, C.; Wright, A. T. Activity-Based Protein Profiling of Bile Salt Hydrolysis in the Human Gut Microbiome with Beta-Lactam or Acrylamide-Based Probes. *ChemBioChem* **2021**, *22* (8), 1448–1455. https://doi.org/10.1002/cbic.202000748.

(193) Thornberry, N. A.; Peterson, E. P.; Zhao, J. J.; Howard, A. D.; Griffin, P. R.; Chapman, K. T. Inactivation of Interleukin-1-Beta Converting Enzyme by Peptide (Acyloxy)Methyl Ketones. *Biochemistry* **1994**, *33* (13), 3934–3940. https://doi.org/10.1021/bi00179a020.

(194) Zhuang, Z.; Byun, W. S.; Kozicka, Z.; Dwyer, B. G.; Donovan, K. A.; Jiang, Z.; Jones, H. M.; Abeja, D. M.; Nix, M. N.; Zhong, J.; Slabicki, M.; Fischer, E. S.; Ebert, B. L.; Gray, N. S. Discovery of Electrophilic Degraders That Exploit SNAr Chemistry. September 27, 2024. https://doi.org/10.1101/2024.09.25.615094.

(195) Chen, K. X.; Lesburg, C. A.; Vibulbhan, B.; Yang, W.; Chan, T.-Y.; Venkatraman, S.; Velazquez, F.; Zeng, Q.; Bennett, F.; Anilkumar, G. N.; Duca, J.; Jiang, Y.; Pinto, P.; Wang, L.; Huang, Y.; Selyutin, O.; Gavalas, S.; Pu, H.; Agrawal, S.; Feld, B.; Huang, H.-C.; Li, C.; Cheng, K.-C.; Shih, N.-Y.; Kozlowski, J. A.; Rosenblum, S. B.; Njoroge, F. G. A Novel Class of Highly Potent Irreversible Hepatitis C Virus NS5B Polymerase Inhibitors. *J. Med. Chem.* **2012**, *55* (5), 2089–2101. https://doi.org/10.1021/jm201322r.

(196) Fairhurst, R. A.; Knoepfel, T.; Leblanc, C.; Buschmann, N.; Gaul, C.; Blank, J.; Galuba, I.; Trappe, J.; Zou, C.; Voshol, J.; Genick, C.; Brunet-Lefeuvre, P.; Bitsch, F.; Graus-Porta, D.; Furet, P. Approaches to Selective Fibroblast Growth Factor Receptor 4 Inhibition through Targeting the ATP-Pocket Middle-Hinge Region. *Medchemcomm* **2017**, *8* (8), 1604–1613. https://doi.org/10.1039/C7MD00213K.

(197) Nelson, D. A.; Nelson, K. T.; Miller, M. W.; Dupe, R.; Chahwala, S. B.; Kennedy, A.; Chander, C.; Fossum, T. W. Pharmacodynamic and Efficacy Profile of TGN 255, a Novel Direct Thrombin Inhibitor, in Canine Cardiopulmonary Bypass and Simulated Mitral Valve Repair. *J. Extracorpor. Technol.* **2008**, *40* (2), 116–122. https://doi.org/10.1051/ject/200840116.

(198) Pattzi, H. M. R.; Pitale, S.; Alpizar, M.; Bennett, C.; O'Farrell, A. M.; Li, J.; Cherrington, J. M.; Guler, H. -P. Dutogliptin, a Selective DPP4 Inhibitor, Improves Glycaemic Control in Patients with Type 2 Diabetes: A 12-week, Double-blind, Randomized, Placebo-controlled, Multicentre Trial. *Diabetes, Obes. Metab.* **2010**, *12* (4), 348–355. https://doi.org/10.1111/j.1463-1326.2010.01195.x.

(199) Marquardt, J. L.; Brown, E. D.; Lane, W. S.; Haley, T. M.; Ichikawa, Y.; Wong, C.-H.; Walsh, C. T. Kinetics, Stoichiometry, and Identification of the Reactive Thiolate in the Inactivation of UDP-GlcNAc Enolpyruvoyl Transferase by the Antibiotic Fosfomycin. *Biochemistry* **1994**, *33* (35), 10646–10651. https://doi.org/10.1021/bi00201a011.

(200) Kuhn, D. J.; Chen, Q.; Voorhees, P. M.; Strader, J. S.; Shenk, K. D.; Sun, C. M.; Demo, S. D.; Bennett, M. K.; van Leeuwen, F. W. B.; Chanan-Khan, A. A.; Orlowski, R. Z. Potent Activity of Carfilzomib, a Novel, Irreversible Inhibitor of the Ubiquitin-Proteasome Pathway, against Preclinical Models of Multiple Myeloma. *Blood* **2007**, *110* (9), 3281–3290. https://doi.org/10.1182/blood-2007-01-065888.

(201) Arafet, K.; Ferrer, S.; Martí, S.; Moliner, V. Quantum Mechanics/Molecular Mechanics Studies of the Mechanism of Falcipain-2 Inhibition by the Epoxysuccinate E64. *Biochemistry* **2014**, *53* (20), 3336–3346. https://doi.org/10.1021/bi500060h.

(202) Vicik, R.; Hoerr, V.; Glaser, M.; Schultheis, M.; Hansell, E.; McKerrow, J. H.; Holzgrabe, U.;

Caffrey, C. R.; Ponte-Sucre, A.; Moll, H.; Stich, A.; Schirmeister, T. Aziridine-2,3-Dicarboxylate Inhibitors Targeting the Major Cysteine Protease of Trypanosoma Brucei as Lead Trypanocidal Agents. *Bioorg. Med. Chem. Lett.* **2006**, *16* (10), 2753–2757. https://doi.org/10.1016/j.bmcl.2006.02.026.

(203) McGregor, N. G. S.; Artola, M.; Nin-Hill, A.; Linzel, D.; Haon, M.; Reijngoud, J.; Ram, A.; Rosso, M.-N.; van der Marel, G. A.; Codée, J. D. C.; van Wezel, G. P.; Berrin, J.-G.; Rovira, C.; Overkleeft, H. S.; Davies, G. J. Rational Design of Mechanism-Based Inhibitors and Activity-Based Probes for the Identification of Retaining α-L-Arabinofuranosidases. *J. Am. Chem. Soc.* **2020**, *142* (10), 4648–4662. https://doi.org/10.1021/jacs.9b11351.

(204) Palmer, J. T.; Rasnick, D.; Klaus, J. L.; Bromme, D. Vinyl Sulfones as Mechanism-Based Cysteine Protease Inhibitors. *J. Med. Chem.* **1995**, *38* (17), 3193–3196. https://doi.org/10.1021/jm00017a002.

(205) Liu, S.; Zhou, B.; Yang, H.; He, Y.; Jiang, Z.-X.; Kumar, S.; Wu, L.; Zhang, Z.-Y. Aryl Vinyl Sulfonates and Sulfones as Active Site-Directed and Mechanism-Based Probes for Protein Tyrosine Phosphatases. *J. Am. Chem. Soc.* **2008**, *130* (26), 8251–8260. https://doi.org/10.1021/ja711125p.

(206) Glynn, S. J.; Gaffney, K. J.; Sainz, M. A.; Louie, S. G.; Petasis, N. A. Molecular Characterization of the Boron Adducts of the Proteasome Inhibitor Bortezomib with Epigallocatechin-3-Gallate and Related Polyphenols. *Org. Biomol. Chem.* **2015**, *13* (13), 3887–3899. https://doi.org/10.1039/C4OB02512A.

(207) Henderson, L. W.; Gautam, A. K. S.; Sharon, E. M.; Johnson, C. R.; Rommel, N. G.; Anthony, A. J.; Russell, D. H.; Jarrold, M. F.; Matouschek, A.; Clemmer, D. E. Bortezomib Inhibits Open Configurations of the 20S Proteasome. *J. Am. Soc. Mass Spectrom.* **2024**, *35* (6), 1063–1068. https://doi.org/10.1021/jasms.4c00080.

(208) Gamcsik, M. P.; Malthouse, J. P. G.; Primrose, W. U.; Mackenzie, N. E.; Boyd, A. S. F.; Russell, R. A.; Scott, A. I. Structure and Stereochemistry of Tetrahedral Inhibitor Complexes of Papain by Direct NMR Observation. *J. Am. Chem. Soc.* **1983**, *105* (20), 6324–6325. https://doi.org/10.1021/ja00358a030.

(209) Moon, J. B.; Coleman, R. S.; Hanzlik, R. P. Reversible Covalent Inhibition of Papain by a Peptide Nitrile. Carbon-13 NMR Evidence for a Thioimidate Ester Adduct. *J. Am. Chem. Soc.* **1986**, *108* (6), 1350–1351. https://doi.org/10.1021/ja00266a066.

(210) Robinson, V. J.; Pauls, H. W.; Coles, P. J.; Smith, R. A.; Krantz, A. Carbon-13 NMR Characterization of the Papain Adduct Formed by Peptidyl Acyloxy-, Aryloxy-, and Chloromethyl Ketone Irreversible Inhibitors. *Bioorg. Chem.* **1992**, *20* (1), 42–54. https://doi.org/10.1016/0045-2068(92)90024-W.

(211) Dalvit, C.; Flocco, M.; Knapp, S.; Mostardini, M.; Perego, R.; Stockman, B. J.; Veronesi, M.; Varasi, M. High-Throughput NMR-Based Screening with Competition Binding Experiments. *J. Am. Chem. Soc.* **2002**, *124* (26), 7702–7709. https://doi.org/10.1021/ja020174b.

(212) Hajduk, P. J.; Gerfin, T.; Boehlen, J.-M.; Häberli, M.; Marek, D.; Fesik, S. W. High-Throughput Nuclear Magnetic Resonance-Based Screening. *J. Med. Chem.* **1999**, *42* (13), 2315–2317. https://doi.org/10.1021/jm9901475.

(213) Maveyraud, L.; Mourey, L. Protein X-Ray Crystallography and Drug Discovery. *Molecules* **2020**, *25* (5), 1030. https://doi.org/10.3390/molecules25051030.

(214) Nadler, W. M.; Waidelich, D.; Kerner, A.; Hanke, S.; Berg, R.; Trumpp, A.; Rösli, C. MALDI versus ESI: The Impact of the Ion Source on Peptide Identification. *J. Proteome Res.* **2017**, *16* (3), 1207–1215. https://doi.org/10.1021/acs.jproteome.6b00805.

(215) El-Aneed, A.; Cohen, A.; Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Appl. Spectrosc. Rev.* **2009**, *44* (3), 210–230.

https://doi.org/10.1080/05704920902717872.

(216) Sinha, A.; Mann, M. A Beginner's Guide to Mass Spectrometry–Based Proteomics. *Biochem. (Lond).* **2020**, *42* (5), 64–69. https://doi.org/10.1042/BIO20200057.

(217) Felitsyn, N.; Peschke, M.; Kebarle, P. Origin and Number of Charges Observed on Multiply-Protonated Native Proteins Produced by ESI. *Int. J. Mass Spectrom.* **2002**, *219* (1), 39–62. https://doi.org/10.1016/S1387-3806(02)00588-2.

(218) Rosell-Llompart, J.; Grifoll, J.; Loscertales, I. G. Electrosprays in the Cone-Jet Mode: From Taylor Cone Formation to Spray Development. *J. Aerosol Sci.* **2018**, *125*, 2–31. https://doi.org/10.1016/j.jaerosci.2018.04.008.

(219) Hillenkamp, F. *MALDI-MS: A Practical Guide to Instrumentation, Methods, and Applications*; 2014.

(220) Dreisewerd, K. The Desorption Process in MALDI. *Chem. Rev.* **2003**, *103* (2), 395–426. https://doi.org/10.1021/cr010375i.

(221) Puretzky, A. A.; Geohegan, D. B.; Hurst, G. B.; Buchanan, M. V.; Luk'yanchuk, B. S. Imaging of Vapor Plumes Produced by Matrix Assisted Laser Desorption: A Plume Sharpening Effect. *Phys. Rev. Lett.* **1999**, *83* (2), 444–447. https://doi.org/10.1103/PhysRevLett.83.444.

(222) Zhigilei, L. V.; Garrison, B. J. Microscopic Mechanisms of Laser Ablation of Organic Solids in the Thermal and Stress Confinement Irradiation Regimes. *J. Appl. Phys.* **2000**, *88* (3), 1281–1298. https://doi.org/10.1063/1.373816.

(223) Knochenmuss, R.; Zhigilei, L. V. Molecular Dynamics Model of Ultraviolet Matrix-Assisted Laser Desorption/Ionization Including Ionization Processes. *J. Phys. Chem. B* **2005**, *109* (48), 22947–22957. https://doi.org/10.1021/jp052945e.

(224) Knochenmuss, R. Ion Formation Mechanisms in UV-MALDI. *Analyst* **2006**, *131* (9), 966. https://doi.org/10.1039/b605646f.

(225) Karas, M.; Glückmann, M.; Schäfer, J. Ionization in Matrix-Assisted Laser Desorption/Ionization: Singly Charged Molecular Ions Are the Lucky Survivors. *J. Mass Spectrom.* **2000**, *35* (1), 1–12. https://doi.org/10.1002/(SICI)1096-9888(200001)35:1<1::AID-JMS904>3.0.CO;2-0.

(226) Jaskolla, T. W.; Karas, M. Compelling Evidence for Lucky Survivor and Gas Phase Protonation: The Unified MALDI Analyte Protonation Mechanism. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 976–988. https://doi.org/10.1007/s13361-011-0093-0.

(227) Kinsel, G. R.; Yao, D.; Yassin, F. H.; Marynick, D. S. Equilibrium Conditions in Laser-Desorbed Plumes: Thermodynamic Properties of α-Cyano-4-Hydroxycinnamic Acid and Protonation of Amino Acids. *Eur. J. Mass Spectrom.* **2006**, *12* (6), 359–367. https://doi.org/10.1255/ejms.821.

(228) Jeckelmann, B.; Piquemal, F. The Elementary Charge for the Definition and Realization of the Ampere. *Ann. Phys.* **2019**, *531* (5). https://doi.org/10.1002/andp.201800389.

(229) Chen, F.; Gülbakan, B.; Weidmann, S.; Fagerer, S. R.; Ibáñez, A. J.; Zenobi, R. Applying Mass Spectrometry to Study Non-covalent Biomolecule Complexes. *Mass Spectrom. Rev.* **2016**, *35* (1), 48–70. https://doi.org/10.1002/mas.21462.

(230) Weidmann, S.; Zenobi, R. High-Mass MALDI-MS Using Ion Conversion Dynode Detectors: Influence of the Conversion Voltage on Sensitivity and Spectral Quality. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (6), 950–954. https://doi.org/10.1007/s13361-014-0867-2.

(231) Giansanti, P.; Tsiatsiani, L.; Low, T. Y.; Heck, A. J. R. Six Alternative Proteases for Mass Spectrometry–Based Proteomics beyond Trypsin. *Nat. Protoc.* **2016**, *11* (5), 993–1006. https://doi.org/10.1038/nprot.2016.057.

(232) Elhanany, E.; Ordentlich, A.; Dgany, O.; Kaplan, D.; Segall, Y.; Barak, R.; Velan, B.; Shafferman, A. Resolving Pathways of Interaction of Covalent Inhibitors with the Active Site of Acetylcholinesterases: MALDI-TOF/MS Analysis of Various Nerve Agent Phosphyl Adducts. *Chem. Res. Toxicol.* **2001**, *14* (7), 912–918. https://doi.org/10.1021/tx0100542.

(233) Barak, D.; Ordentlich, A.; Kaplan, D.; Barak, R.; Mizrahi, D.; Kronman, C.; Segall, Y.; Velan, B.; Shafferman, A. Evidence for P−N Bond Scission in Phosphoroamidate Nerve Agent Adducts of Human Acetylcholinesterase. *Biochemistry* **2000**, *39* (5), 1156–1161. https://doi.org/10.1021/bi992009n.

(234) Doorn, J. A.; Gage, D. A.; Schall, M.; Talley, T. T.; Thompson, C. M.; Richardson, R. J. Inhibition of Acetylcholinesterase by (1 S ,3 S)-Isomalathion Proceeds with Loss of Thiomethyl: Kinetic and Mass Spectral Evidence for an Unexpected Primary Leaving Group. *Chem. Res. Toxicol.* **2000**, *13* (12), 1313–1320. https://doi.org/10.1021/tx000184v.

(235) Gobom, J.; Mueller, M.; Egelhofer, V.; Theiss, D.; Lehrach, H.; Nordhoff, E. A Calibration Method That Simplifies and Improves Accurate Determination of Peptide Molecular Masses by MALDI-TOF MS. *Anal. Chem.* **2002**, *74* (15), 3915–3923. https://doi.org/10.1021/ac011203o.

(236) Falke, S.; Lieske, J.; Herrmann, A.; Loboda, J.; Karničar, K.; Günther, S.; Reinke, P. Y. A.; Ewert, W.; Usenik, A.; Lindič, N.; Sekirnik, A.; Dretnik, K.; Tsuge, H.; Turk, V.; Chapman, H. N.; Hinrichs, W.; Ebert, G.; Turk, D.; Meents, A. Structural Elucidation and Antiviral Activity of Covalent Cathepsin L Inhibitors. *J. Med. Chem.* **2024**, *67* (9), 7048–7067. https://doi.org/10.1021/acs.jmedchem.3c02351.

(237) Zhang, Y.; Go, E. P.; Desaire, H. Maximizing Coverage of Glycosylation Heterogeneity in MALDI-MS Analysis of Glycoproteins with Up to 27 Glycosylation Sites. *Anal. Chem.* **2008**, *80* (9), 3144–3158. https://doi.org/10.1021/ac702081a.

(238) Garden, R. W.; Sweedler, J. V. Heterogeneity within MALDI Samples As Revealed by Mass Spectrometric Imaging. *Anal. Chem.* **2000**, *72* (1), 30–36. https://doi.org/10.1021/ac9908997.

(239) Fedorova, M.; Kuleva, N.; Hoffmann, R. Identification of Cysteine, Methionine and Tryptophan Residues of Actin Oxidized In Vivo during Oxidative Stress. *J. Proteome Res.* **2010**, *9* (3), 1598–1609. https://doi.org/10.1021/pr901099e.

(240) Vestal, M. L.; Juhasz, P.; Martin, S. A. Delayed Extraction Matrix-assisted Laser Desorption Time-of-flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1995**, *9* (11), 1044–1050. https://doi.org/10.1002/rcm.1290091115.

(241) Bahr, U.; Stahl-Zeng, J.; Gleitsmann, E.; Karas, M. Delayed Extraction Time-of-Flight MALDI Mass Spectrometry of Proteins above 25000 Da. *J. Mass Spectrom.* **1997**, *32* (10), 1111–1116. https://doi.org/10.1002/(SICI)1096-9888(199711)32:10<1111::AID-JMS567>3.0.CO;2-Y.

(242) Hofstadler, S. A.; Sannes-Lowery, K. A. Applications of ESI-MS in Drug Discovery: Interrogation of Noncovalent Complexes. *Nat. Rev. Drug Discov.* **2006**, *5* (7), 585–595. https://doi.org/10.1038/nrd2083.

(243) Song, F. A Study of Noncovalent Protein Complexes by Matrix-Assisted Laser Desorption/Ionization. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (7), 1286–1290. https://doi.org/10.1016/j.jasms.2007.03.026.

(244) Rosinke, B.; Strupat, K.; Hillenkamp, F.; Rosenbusch, J.; Dencher, N.; Krüger, U.; Galla, H. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-MS) of Membrane Proteins and Non-covalent Complexes. *J. Mass Spectrom.* **1995**, *30* (10), 1462–1468. https://doi.org/10.1002/jms.1190301012.

(245) Sudha, R.; Zenobi, R. The Detection and Stability of DNA Duplexes Probed by MALDI Mass Spectrometry. *Helv. Chim. Acta* **2002**, *85* (10), 3136–3143. https://doi.org/10.1002/1522-2675(200210)85:10<3136::AID-HLCA3136>3.0.CO;2-F.

(246) Terrier, P.; Tortajada, J.; Zin, G.; Buchmann, W. Noncovalent Complexes between DNA and Basic Polypeptides or Polyamines by MALDI-TOF. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (11), 1977–1989. https://doi.org/10.1016/j.jasms.2007.07.028.

(247) Horneffer, V.; Strupat, K.; Hillenkamp, F. Localization of Noncovalent Complexes in MALDI-Preparations by CLSM. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (11), 1599–1604. https://doi.org/10.1016/j.jasms.2006.06.028.

(248) Wortmann, A.; Pimenova, T.; Alves, S.; Zenobi, R. Investigation of the First Shot Phenomenon in MALDI Mass Spectrometry of Protein Complexes. *Analyst* **2007**, *132* (3), 199. https://doi.org/10.1039/b615411e.

(249) Yanes, O.; Villanueva, J.; Querol, E.; Aviles, F. X. Detection of Non-Covalent Protein Interactions by "intensity Fading" MALDI-TOF Mass Spectrometry: Applications to Proteases and Protease Inhibitors. *Nat. Protoc.* **2007**, *2* (1), 119–130. https://doi.org/10.1038/nprot.2006.487.

(250) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Stryer – Biochemie*; 2014.

(251) Klein, P.; Johe, P.; Wagner, A.; Jung, S.; Kühlborn, J.; Barthels, F.; Tenzer, S.; Distler, U.; Waigel, W.; Engels, B.; Hellmich, U. A.; Opatz, T.; Schirmeister, T. New Cysteine Protease Inhibitors: Electrophilic (Het)Arenes and Unexpected Prodrug Identification for the Trypanosoma Protease Rhodesain. *Molecules* **2020**, *25* (6), 1451. https://doi.org/10.3390/molecules25061451.

(252) Serafimova, I. M.; Pufall, M. A.; Krishnan, S.; Duda, K.; Cohen, M. S.; Maglathlin, R. L.; McFarland, J. M.; Miller, R. M.; Frödin, M.; Taunton, J. Reversible Targeting of Noncatalytic Cysteines with Chemically Tuned Electrophiles. *Nat. Chem. Biol.* **2012**, *8* (5), 471–476. https://doi.org/10.1038/nchembio.925.

(253) Huang, L.; Brinen, L. S.; Ellman, J. A. Crystal Structures of Reversible Ketone-Based Inhibitors of the Cysteine Protease Cruzain. *Bioorg. Med. Chem.* **2003**, *11* (1), 21–29. https://doi.org/10.1016/S0968-0896(02)00427-3.

(254) Chen, X.; Huang, X.; Ma, Q.; Kuzmič, P.; Zhou, B.; Zhang, S.; Chen, J.; Xu, J.; Liu, B.; Jiang, H.; Zhang, W.; Yang, C.; Wu, S.; Huang, J.; Li, H.; Long, C.; Zhao, X.; Xu, H.; Sheng, Y.; Guo, Y.; Niu, C.; Xue, L.; Xu, Y.; Liu, J.; Zhang, T.; Spencer, J.; Zhu, Z.; Deng, W.; Chen, X.; Chen, S.-H.; Zhong, N.; Xiong, X.; Yang, Z. Preclinical Evaluation of the SARS-CoV-2 Mpro Inhibitor RAY1216 Shows Improved Pharmacokinetics Compared with Nirmatrelvir. *Nat. Microbiol.* **2024**, *9* (4), 1075–1088. https://doi.org/10.1038/s41564-024-01618-9.

(255) Meister, A. *Advances in Enzymology and Related Areas of Molecular Biology*; 1988.

(256) Copeland, R. A. *Evaluation of Enzyme Inhibitors in Drug Discovery*; Wiley, 2005. https://doi.org/10.1002/9781118540398.

(257) Tchoupé, J. R.; Moreau, T.; Gauthier, F.; Bieth, J. G. Photometric or Fluorometric Assay of Cathepsin B, L and H and Papain Using Substrates with an Aminotrifluoromethylcoumarin Leaving Group. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1991**, *1076* (1), 149–151. https://doi.org/10.1016/0167-4838(91)90232-O.

(258) 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.

(259) Zhao, Z.; Zhao, L.; Kong, C.; Zhou, J.; Zhou, F. A Review of Biophysical Strategies to Investigate Protein-Ligand Binding: What Have We Employed? *Int. J. Biol. Macromol.* **2024**, *276*, 133973. https://doi.org/10.1016/j.ijbiomac.2024.133973.

(260) Bar-Even, A.; Noor, E.; Savir, Y.; Liebermeister, W.; Davidi, D.; Tawfik, D. S.; Milo, R. The

Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry* **2011**, *50* (21), 4402–4410. https://doi.org/10.1021/bi2002289.

(261) Furfine, E. S.; D'Souza, E.; Ingold, K. J.; Leban, J. J.; Spector, T.; Porter, D. J. T. Two-Step Binding Mechanism for HIV Protease Inhibitors. *Biochemistry* **1992**, *31* (34), 7886–7891. https://doi.org/10.1021/bi00149a020.

(262) Park, C. Visual Interpretation of the Meaning of k Cat / K M in Enzyme Kinetics. *J. Chem. Educ.* **2022**, *99* (7), 2556–2562. https://doi.org/10.1021/acs.jchemed.1c01268.

(263) Chatterjee, P.; Botello-Smith, W. M.; Zhang, H.; Qian, L.; Alsamarah, A.; Kent, D.; Lacroix, J. J.; Baudry, M.; Luo, Y. Can Relative Binding Free Energy Predict Selectivity of Reversible Covalent Inhibitors? *J. Am. Chem. Soc.* **2017**, *139* (49), 17945–17952. https://doi.org/10.1021/jacs.7b08938.

(264) Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T. H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P. R.; Distler, U.; Tenzer, S.; Johé, P.; Hellmich, U. A.; Engels, B. Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138* (27), 8332–8335. https://doi.org/10.1021/jacs.6b03052.

(265) Silva, J. R. A.; Cianni, L.; Araujo, D.; Batista, P. H. J.; de Vita, D.; Rosini, F.; Leitão, A.; Lameira, J.; Montanari, C. A. Assessment of the Cruzain Cysteine Protease Reversible and Irreversible Covalent Inhibition Mechanism. *J. Chem. Inf. Model.* **2020**, *60* (3), 1666–1677. https://doi.org/10.1021/acs.jcim.9b01138.

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

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

(268) Mons, E.; Roet, S.; Kim, R. Q.; Mulder, M. P. C. A Comprehensive Guide for Assessing Covalent Inhibition in Enzymatic Assays Illustrated with Kinetic Simulations. *Curr. Protoc.* **2022**, *2* (6). https://doi.org/10.1002/cpz1.419.

(269) *Enzyme Kinetics - Catalysis and Control*, 1st ed.; Purich, D. L., Ed.; Elsevier, 2010.

(270) Burlingham, B. T.; Widlanski, T. S. An Intuitive Look at the Relationship of Ki and IC50: A More General Use for the Dixon Plot. *J. Chem. Educ.* **2003**, *80* (2), 214. https://doi.org/10.1021/ed080p214.

(271) 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.

(272) Johnson, K. A.; Goody, R. S. The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten Paper. *Biochemistry* **2011**, *50* (39), 8264–8269. https://doi.org/10.1021/bi201284u.

(273) Gubler, H.; Schopfer, U.; Jacoby, E. Theoretical and Experimental Relationships between Percent Inhibition and IC50 Data Observed in High-Throughput Screening. *SLAS Discov.* **2013**, *18* (1), 1–13. https://doi.org/10.1177/1087057112455219.

(274) Olsson, S.-L.; Ek, B.; Björk, I. The Affinity and Kinetics of Inhibition of Cysteine Proteinases by Intact Recombinant Bovine Cystatin C. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1999**, *1432* (1), 73–81. https://doi.org/10.1016/S0167-4838(99)00090-4.

(275) Ghosh, P.; Raj, N.; Verma, H.; Patel, M.; Chakraborti, S.; Khatri, B.; Doreswamy, C. M.;

Anandakumar, S. R.; Seekallu, S.; Dinesh, M. B.; Jadhav, G.; Yadav, P. N.; Chatterjee, J. An Amide to Thioamide Substitution Improves the Permeability and Bioavailability of Macrocyclic Peptides. *Nat. Commun.* **2023**, *14* (1), 6050. https://doi.org/10.1038/s41467-023-41748-y.

(276) Hosono, Y.; Uchida, S.; Shinkai, M.; Townsend, C. E.; Kelly, C. N.; Naylor, M. R.; Lee, H.-W.; Kanamitsu, K.; Ishii, M.; Ueki, R.; Ueda, T.; Takeuchi, K.; Sugita, M.; Akiyama, Y.; Lokey, S. R.; Morimoto, J.; Sando, S. Amide-to-Ester Substitution as a Stable Alternative to N-Methylation for Increasing Membrane Permeability in Cyclic Peptides. *Nat. Commun.* **2023**, *14* (1), 1416. https://doi.org/10.1038/s41467-023-36978-z.

(277) Li, Y.; Li, W.; Xu, Z. Improvement on Permeability of Cyclic Peptide/Peptidomimetic: Backbone N-Methylation as A Useful Tool. *Mar. Drugs* **2021**, *19* (6), 311. https://doi.org/10.3390/md19060311.

(278) Kennedy, T. Managing the Drug Discovery/Development Interface. *Drug Discov. Today* **1997**, *2* (10), 436–444. https://doi.org/10.1016/S1359-6446(97)01099-4.

(279) Tsaioun, K.; Bottlaender, M.; Mabondzo, A. ADDME – Avoiding Drug Development Mistakes Early: Central Nervous System Drug Discovery Perspective. *BMC Neurol.* **2009**, *9* (Suppl 1), S1. https://doi.org/10.1186/1471-2377-9-S1-S1.

(280) Pantaleão, S. Q.; Fernandes, P. O.; Gonçalves, J. E.; Maltarollo, V. G.; Honorio, K. M. Recent Advances in the Prediction of Pharmacokinetics Properties in Drug Design Studies: A Review. *ChemMedChem* **2022**, *17* (1). https://doi.org/10.1002/cmdc.202100542.

(281) Vertzoni, M.; Augustijns, P.; Grimm, M.; Koziolek, M.; Lemmens, G.; Parrott, N.; Pentafragka, C.; Reppas, C.; Rubbens, J.; Van Den Abeele, J.; Vanuytsel, T.; Weitschies, W.; Wilson, C. G. Impact of Regional Differences along the Gastrointestinal Tract of Healthy Adults on Oral Drug Absorption: An UNGAP Review. *Eur. J. Pharm. Sci.* **2019**, *134*, 153–175. https://doi.org/10.1016/j.ejps.2019.04.013.

(282) Vinarov, Z.; Abdallah, M.; Agundez, J. A. G.; Allegaert, K.; Basit, A. W.; Braeckmans, M.; Ceulemans, J.; Corsetti, M.; Griffin, B. T.; Grimm, M.; Keszthelyi, D.; Koziolek, M.; Madla, C. M.; Matthys, C.; McCoubrey, L. E.; Mitra, A.; Reppas, C.; Stappaerts, J.; Steenackers, N.; Trevaskis, N. L.; Vanuytsel, T.; Vertzoni, M.; Weitschies, W.; Wilson, C.; Augustijns, P. Impact of Gastrointestinal Tract Variability on Oral Drug Absorption and Pharmacokinetics: An UNGAP Review. *Eur. J. Pharm. Sci.* **2021**, *162*, 105812. https://doi.org/10.1016/j.ejps.2021.105812.

(283) Helander, H. F.; Fändriks, L. Surface Area of the Digestive Tract – Revisited. *Scand. J. Gastroenterol.* **2014**, *49* (6), 681–689. https://doi.org/10.3109/00365521.2014.898326.

(284) Stenling, R.; Fredrikzon, B.; Nyhlin, H.; Helander, H. F.; Falkmer, S. Surface Infrastructure of the Small Intestine Mucosa in Healthy Children and Adults: A Scanning Electron Microscopic Study with Some Methodological Aspects. *Ultrastruct. Pathol.* **1984**, *6* (2–3), 131–140. https://doi.org/10.3109/01913128409018567.

(285) Lennernäs, H. Intestinal Permeability and Its Relevance for Absorption and Elimination. *Xenobiotica* **2007**, *37* (10–11), 1015–1051. https://doi.org/10.1080/00498250701704819.

(286) *Drug Bioavailability: Estimation of Solubility, Permeability, Absorption and Bioavailability*; van de Waterbeemd, H., Testa, B., Eds.; Methods and Principles in Medicinal Chemistry; Wiley, 2008; Vol. 40. https://doi.org/10.1002/9783527623860.

(287) Brouwers, J.; Mols, R.; Annaert, P.; Augustijns, P. Validation of a Differential in Situ Perfusion Method with Mesenteric Blood Sampling in Rats for Intestinal Drug Interaction Profiling. *Biopharm. Drug Dispos.* **2010**, *31* (5–6), 278–285. https://doi.org/10.1002/bdd.710.

(288) Harayama, T.; Riezman, H. Understanding the Diversity of Membrane Lipid Composition. *Nat. Rev. Mol. Cell Biol.* **2018**, *19* (5), 281–296. https://doi.org/10.1038/nrm.2017.138.

(289) Kawai, K.; Fujita, M.; Nakao, M. Lipid Components of Two Different Regions of an Intestinal Epithelial Cell Membrane of Mouse. *Biochim. Biophys. Acta - Lipids Lipid Metab.* **1974**, *369* (2), 222–233. https://doi.org/10.1016/0005-2760(74)90253-7.

(290) Yang, N. J.; Hinner, M. J. Getting Across the Cell Membrane: An Overview for Small Molecules, Peptides, and Proteins; 2015; pp 29–53. https://doi.org/10.1007/978-1-4939-2272-7_3.

(291) Kokate, A.; Chang, J. H.; Park, M. S. Absorption of Drugs Via Passive Diffusion and Carrier-Mediated Pathways. In *Oral Bioavailability and Drug Delivery*; Wiley, 2023; pp 61–78. https://doi.org/10.1002/9781119660699.ch5.

(292) Sugano, K.; Kansy, M.; Artursson, P.; Avdeef, A.; Bendels, S.; Di, L.; Ecker, G. F.; Faller, B.; Fischer, H.; Gerebtzoff, G.; Lennernaes, H.; Senner, F. Coexistence of Passive and Carrier-Mediated Processes in Drug Transport. *Nat. Rev. Drug Discov.* **2010**, *9* (8), 597–614. https://doi.org/10.1038/nrd3187.

(293) *Anatomie*; Aumüller, G., Aust, G., Doll, A., Engele, J., Kirsch, J., Mense, S., Reißig, D., Salvetter, J., Schmidt, W., Schmitz, F., Schulte, E., Spanel-Borowski, K., Wolff, W., Wurzinger, L. J., Zilch, H.-G., Eds.; Georg Thieme Verlag: Stuttgart, 2010. https://doi.org/10.1055/b-002-46981.

(294) Avdeef, A. *Absorption and Drug Development*; Wiley, 2003. https://doi.org/10.1002/047145026X.

(295) Mannhold, R.; Poda, G. I.; Ostermann, C.; Tetko, I. V. Calculation of Molecular Lipophilicity: State-of-the-Art and Comparison of LogP Methods on More than 96,000 Compounds. *J. Pharm. Sci.* **2009**, *98* (3), 861–893. https://doi.org/10.1002/jps.21494.

(296) Tetko, I. V.; Poda, G. I. Application of ALOGPS 2.1 to Predict Log D Distribution Coefficient for Pfizer Proprietary Compounds. *J. Med. Chem.* **2004**, *47* (23), 5601–5604. https://doi.org/10.1021/jm049509l.

(297) Palm, K.; Stenberg, P.; Luthman, K.; Artursson, P. Polar Molecular Surface Properties Predict the Intestinal Absorption of Drugs in Humans. *Pharm Res* **1997**, *14*, 568–571. https://doi.org/https://doi.org/10.1023/A:1012188625088.

(298) 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.

(299) Lovering, F.; Bikker, J.; Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **2009**, *52* (21), 6752–6756. https://doi.org/10.1021/jm901241e.

(300) Beckers, M.; Sirockin, F.; Fechner, N.; Stiefl, N. Balancing Molecular Size, Activity, Permeability, and Other Properties: Drug Candidates in the Context of Their Chemical Structure Optimization. *J. Chem. Inf. Model.* **2024**, *64* (16), 6636–6647. https://doi.org/10.1021/acs.jcim.4c00898.

(301) Matsson, P.; Doak, B. C.; Over, B.; Kihlberg, J. Cell Permeability beyond the Rule of 5. *Adv. Drug Deliv. Rev.* **2016**, *101*, 42–61. https://doi.org/10.1016/j.addr.2016.03.013.

(302) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* **2014**, *21* (9), 1115–1142. https://doi.org/10.1016/j.chembiol.2014.08.013.

(303) Doak, B. C.; Zheng, J.; Dobritzsch, D.; Kihlberg, J. How Beyond Rule of 5 Drugs and Clinical Candidates Bind to Their Targets. *J. Med. Chem.* **2016**, *59* (6), 2312–2327. https://doi.org/10.1021/acs.jmedchem.5b01286.

(304) Bennion, B. J.; Be, N. A.; McNerney, M. W.; Lao, V.; Carlson, E. M.; Valdez, C. A.; Malfatti,

M. A.; Enright, H. A.; Nguyen, T. H.; Lightstone, F. C.; Carpenter, T. S. Predicting a Drug's Membrane Permeability: A Computational Model Validated With in Vitro Permeability Assay Data. *J. Phys. Chem. B* **2017**, *121* (20), 5228–5237. https://doi.org/10.1021/acs.jpcb.7b02914.

(305) Hou, T. J.; Zhang, W.; Xia, K.; Qiao, X. B.; Xu, X. J. ADME Evaluation in Drug Discovery. 5. Correlation of Caco-2 Permeation with Simple Molecular Properties. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (5), 1585–1600. https://doi.org/10.1021/ci049884m.

(306) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41* (7), 1007–1010. https://doi.org/10.1021/jm970530e.

(307) Avdeef, A. Physicochemical Profiling (Solubility, Permeability and Charge State). *Curr. Top. Med. Chem.* **2001**, *1* (4), 277–351. https://doi.org/10.2174/1568026013395100.

(308) Avdeef, A. The Rise of PAMPA. *Expert Opin. Drug Metab. Toxicol.* **2005**, *1* (2), 325–342. https://doi.org/10.1517/17425255.1.2.325.

(309) Williams, J.; Siramshetty, V.; Nguyễn, Đ.-T.; Padilha, E. C.; Kabir, M.; Yu, K.-R.; Wang, A. Q.; Zhao, T.; Itkin, M.; Shinn, P.; Mathé, E. A.; Xu, X.; Shah, P. Using in Vitro ADME Data for Lead Compound Selection: An Emphasis on PAMPA PH 5 Permeability and Oral Bioavailability. *Bioorg. Med. Chem.* **2022**, *56*, 116588. https://doi.org/10.1016/j.bmc.2021.116588.

(310) Avdeef, A.; Tsinman, O. PAMPA - A Drug Absorption in Vitro Model: 13. Chemical Selectivity Due to Membrane Hydrogen Bonding: In Combo Comparisons of HDM-, DOPC-, and DS-PAMPA Models. *Eur. J. Pharm. Sci.* **2006**, *28* (1–2), 43–50. https://doi.org/10.1016/j.ejps.2005.12.008.

(311) Wohnsland, F.; Faller, B. High-Throughput Permeability PH Profile and High-Throughput Alkane/Water Log P with Artificial Membranes. *J. Med. Chem.* **2001**, *44* (6), 923–930. https://doi.org/10.1021/jm001020e.

(312) Kansy, M.; Fischer, H.; Kratzat, K.; Senner, F.; Wagner, B.; Parrilla, I. High-Throughput Artificial Membrane Permeability Studies in Early Lead Discovery and Development. In *Pharmacokinetic Optimization in Drug Research*; Wiley, 2001; pp 447–464. https://doi.org/10.1002/9783906390437.ch24.

(313) Avdeef, A.; Nielsen, P. E.; Di, C. M. EP 1,521,962 B1: METHOD AND APPARATUS FOR MEASURING MEMBRANE PERMEABILITY, 2009.

(314) Sugano, K.; Hamada, H.; Machida, M.; Ushio, H.; Saitoh, K.; Terada, K. Optimized Conditions of Bio-Mimetic Artificial Membrane Permeation Assay. *Int. J. Pharm.* **2001**, *228* (1–2), 181–188. https://doi.org/10.1016/S0378-5173(01)00845-6.

(315) Di, L.; Kerns, E. H.; Fan, K.; McConnell, O. J.; Carter, G. T. High Throughput Artificial Membrane Permeability Assay for Blood–Brain Barrier. *Eur. J. Med. Chem.* **2003**, *38* (3), 223–232. https://doi.org/10.1016/S0223-5234(03)00012-6.

(316) Assmus, F.; Ross, A.; Fischer, H.; Seelig, J.; Seelig, A. 31P and 1H NMR Studies of the Molecular Organization of Lipids in the Parallel Artificial Membrane Permeability Assay. *Mol. Pharm.* **2017**, *14* (1), 284–295. https://doi.org/10.1021/acs.molpharmaceut.6b00889.

(317) Fick, A. Ueber Diffusion. *Ann. Phys.* **1855**, *170* (1), 59–86. https://doi.org/10.1002/andp.18551700105.

(318) Kansy, M.; Fischer, H.; Kratzat, K.; Senner, F.; Wagner, B.; Parrilla, I. High-Throughput Artificial Membrane Permeability Studies in Early Lead Discovery and Development. *Pharmacokinet. Optim. Drug Res.* **2001**.

(319) Bermejo, M.; Avdeef, A.; Ruiz, A.; Nalda, R.; Ruell, J. A.; Tsinman, O.; González, I.; Fernández, C.; Sánchez, G.; Garrigues, T. M.; Merino, V. PAMPA—a Drug Absorption in Vitro Model: 7.

Comparing Rat in Situ, Caco-2, and PAMPA Permeability of Fluoroquinolones. *Eur. J. Pharm. Sci.* **2004**, *21* (4), 429–441. https://doi.org/10.1016/j.ejps.2003.10.009.

(320) Sawada, G. A.; Barsuhn, C. L.; Lutzke, B. S.; Houghton, M. E.; Padbury, G. E.; Ho, N. F. H.; Raub, T. J. Increased Lipophilicity and Subsequent Cell Partitioning Decrease Passive Transcellular Diffusion of Novel, Highly Lipophilic Antioxidants. *J. Pharmacol. Exp. Ther.* **1999**, *288* (3), 1317–1326.

(321) Palm, K.; Luthmann, K.; Ros, J.; Artursson, P. Effect of Molecular Charge on Intestinal Epithelial Drug Transport: PH-Dependent Transport of Cationic Drugs. *J. Pharmacol. Exp. Ther.* **1999**, *291* (2), 435–443.

(322) Soares, P.; Gadd, M. S.; Frost, J.; Galdeano, C.; Ellis, L.; Epemolu, O.; Rocha, S.; Read, K. D.; Ciulli, A. Group-Based Optimization of Potent and Cell-Active Inhibitors of the von Hippel–Lindau (VHL) E3 Ubiquitin Ligase: Structure–Activity Relationships Leading to the Chemical Probe (2 S ,4 R)-1-((S)-2-(1-Cyanocyclopropanecarboxamido)-3,3-Dimethylbutanoyl)-. *J. Med. Chem.* **2018**, *61* (2), 599–618. https://doi.org/10.1021/acs.jmedchem.7b00675.

(323) Masungi, C.; Mensch, J.; van Dijck, A.; Borremans, C.; Willems, B.; Mackie, C.; Noppe, M.; Brewster, M. E. Parallel Artificial Membrane Permeability Assay (PAMPA) Combined with a 10-Day Multiscreen Caco-2 Cell Culture as a Tool for Assessing New Drug Candidates. *Pharmazie* **2008**, *63*, 194–199.

(324) Lennernaäs, H. Human Intestinal Permeability. *J. Pharm. Sci.* **1998**, *87* (4), 403–410. https://doi.org/10.1021/js970332a.

(325) Korjamo, T.; Heikkinen, A. T.; Mönkkönen, J. Analysis of Unstirred Water Layer in In Vitro Permeability Experiments. *J. Pharm. Sci.* **2009**, *98* (12), 4469–4479. https://doi.org/10.1002/jps.21762.

(326) Nielsen, P. E.; Avdeef, A. PAMPA—a Drug Absorption in Vitro Model 8. Apparent Filter Porosity and the Unstirred Water Layer8. Apparent Filter Porosity and the Unstirred Water Layer. *Eur. J. Pharm. Sci.* **2004**, *22* (1), 33–41. https://doi.org/10.1016/j.ejps.2004.02.003.

(327) Thompson, M.; Lennox, R. B.; McClelland, R. A. Structure and Electrochemical Properties of Microfiltration Filter-Lipid Membrane Systems. *Anal. Chem.* **1982**, *54* (1), 76–81.

(328) Bermejo, M.; Avdeef, A.; Ruiz, A.; Nalda, R.; Ruell, J. A.; Tsinman, O.; González, I.; Fernández, C.; Sánchez, G.; Garrigues, T. M.; Merino, V. PAMPA - A Drug Absorption in Vitro Model: 7. Comparing Rat in Situ, Caco-2, and PAMPA Permeability of Fluoroquinolones. *Eur. J. Pharm. Sci.* **2004**, *21* (4), 429–441. https://doi.org/10.1016/j.ejps.2003.10.009.

(329) Nielsen, P. E.; Avdeef, A. PAMPA - A Drug Absorption in Vitro Model: 8. Apparent Filter Porosity and the Unstirred Water Layer. *Eur. J. Pharm. Sci.* **2004**, *22* (1), 33–41. https://doi.org/10.1016/j.ejps.2004.02.003.

(330) Kerns, E. H.; Di, L.; Petusky, S.; Farris, M.; Ley, R.; Jupp, P. Combined Application of Parallel Artificial Membrane Permeability Assay and Caco-2 Permeability Assays in Drug Discovery. *J. Pharm. Sci.* **2004**, *93* (6), 1440–1453. https://doi.org/10.1002/jps.20075.

(331) Milite, C.; Feoli, A.; Horton, J. R.; Rescigno, D.; Cipriano, A.; Pisapia, V.; Viviano, M.; Pepe, G.; Amendola, G.; Novellino, E.; Cosconati, S.; Cheng, X.; Castellano, S.; Sbardella, G. Discovery of a Novel Chemotype of Histone Lysine Methyltransferase EHMT1/2 (GLP/G9a) Inhibitors: Rational Design, Synthesis, Biological Evaluation, and Co-Crystal Structure. *J. Med. Chem.* **2019**, *62* (5), 2666–2689. https://doi.org/10.1021/acs.jmedchem.8b02008.

(332) Cowan, N.; Dätwyler, P.; Ernst, B.; Wang, C.; Vennerstrom, J. L.; Spangenberg, T.; Keiser, J. Activities of N , N ′-Diarylurea MMV665852 Analogs against Schistosoma Mansoni. *Antimicrob. Agents Chemother.* **2015**, *59* (4), 1935–1941. https://doi.org/10.1128/AAC.04463-14.

(333) Avdeef, A.; Bendels, S.; Di, L. i.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. PAMPA—

Critical Factors for Better Predictions of Absorption. *J. Pharm. Sci.* **2007**, *96* (11), 2893–2909. https://doi.org/10.1002/jps.21068.

(334) Raboisson, P.; de Kock, H.; Rosenquist, Å.; Nilsson, M.; Salvador-Oden, L.; Lin, T.-I.; Roue, N.; Ivanov, V.; Wähling, H.; Wickström, K.; Hamelink, E.; Edlund, M.; Vrang, L.; Vendeville, S.; Van de Vreken, W.; McGowan, D.; Tahri, A.; Hu, L.; Boutton, C.; Lenz, O.; Delouvroy, F.; Pille, G.; Surleraux, D.; Wigerinck, P.; Samuelsson, B.; Simmen, K. Structure–Activity Relationship Study on a Novel Series of Cyclopentane-Containing Macrocyclic Inhibitors of the Hepatitis C Virus NS3/4A Protease Leading to the Discovery of TMC435350. *Bioorg. Med. Chem. Lett.* **2008**, *18* (17), 4853–4858. https://doi.org/10.1016/j.bmcl.2008.07.088.

(335) Klein, V. G.; Townsend, C. E.; Testa, A.; Zengerle, M.; Maniaci, C.; Hughes, S. J.; Chan, K.-H.; Ciulli, A.; Lokey, R. S. Understanding and Improving the Membrane Permeability of VH032-Based PROTACs. *ACS Med. Chem. Lett.* **2020**, *11* (9), 1732–1738. https://doi.org/10.1021/acsmedchemlett.0c00265.

(336) Mikkelsen, T. S.; Thorn, C. F.; Yang, J. J.; Ulrich, C. M.; French, D.; Zaza, G.; Dunnenberger, H. M.; Marsh, S.; McLeod, H. L.; Giacomini, K.; Becker, M. L.; Gaedigk, R.; Leeder, J. S.; Kager, L.; Relling, M. V.; Evans, W.; Klein, T. E.; Altman, R. B. PharmGKB Summary: Methotrexate Pathway. *Pharmacogenet. Genomics* **2011**, *21* (10), 679–686. https://doi.org/10.1097/FPC.0b013e328343dd93.

(337) Schmidt, D.; Lynch, J. *Application Note (Merck): Evaluation of the reproducibility of Parallel Artificial Membrane Permeation Assays (PAMPA)*. https://www.sigmaaldrich.com/DE/de/technical-documents/technical-article/research-and-disease-areas/pharmacology-and-drug-discovery-research/evaluation-of-the-reproducibility-of-pampa.

(338) Avdeef, A.; Nielsen, P. E.; Tsinman, O. PAMPA—a Drug Absorption in Vitro Model: 11. Matching the in Vivo Unstirred Water Layer Thickness by Individual-Well Stirring in Microtitre Plates. *Eur. J. Pharm. Sci.* **2004**, *22* (5), 365–374. https://doi.org/10.1016/j.ejps.2004.04.009.

(339) Volpe, D. A. Drug-Permeability and Transporter Assays in Caco-2 and Mdck Cell Lines. *Future Med. Chem.* **2011**, *3* (16), 2063–2077. https://doi.org/10.4155/fmc.11.149.

(340) Volpe, D. A. Variability in Caco-2 and MDCK Cell-Based Intestinal Permeability Assays. *J. Pharm. Sci.* **2008**, *97* (2), 712–725. https://doi.org/10.1002/jps.21010.

(341) Kus, M.; Ibragimow, I.; Piotrowska-Kempisty, H. Caco-2 Cell Line Standardization with Pharmaceutical Requirements and In Vitro Model Suitability for Permeability Assays. *Pharmaceutics* **2023**, *15* (11), 2523. https://doi.org/10.3390/pharmaceutics15112523.

(342) Food and Drug Administration (FDA). *Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System Guidance for Industry*. https://www.gmp-navigator.com/files/guidemgr/UCM070246.pdf.

(343) Ohta, Y.; Kazuki, K.; Abe, S.; Oshimura, M.; Kobayashi, K.; Kazuki, Y. Development of Caco-2 Cells Expressing Four CYPs via a Mammalian Artificial Chromosome. *BMC Biotechnol.* **2020**, *20* (1), 44. https://doi.org/10.1186/s12896-020-00637-8.

(344) Uchida, M.; Fukazawa, T.; Yamazaki, Y.; Hashimoto, H.; Miyamoto, Y. A Modified Fast (4 Day) 96-Well Plate Caco-2 Permeability Assay. *J. Pharmacol. Toxicol. Methods* **2009**, *59* (1), 39–43. https://doi.org/10.1016/j.vascn.2008.10.006.

(345) Matsson, P.; Kihlberg, J. How Big Is Too Big for Cell Permeability? *J. Med. Chem.* **2017**, *60* (5), 1662–1664. https://doi.org/10.1021/acs.jmedchem.7b00237.

(346) Pemberton, P. A. Proteinase Inhibitors | Cystatins. In *Encyclopedia of Respiratory Medicine*; Elsevier, 2006; pp 511–517. https://doi.org/10.1016/B0-12-370879-6/00329-X.

(347) Vallat, P.; Gaillard, P.; Carrupt, P.; Tsai, R.; Testa, B. Structure-Lipophilicity and Structure-Polarity Relationships of Amino Acids and Peptides. *Helv. Chim. Acta* **1995**, *78* (2), 471–485. https://doi.org/10.1002/hlca.19950780218.

(348) Rezai, T.; Bock, J. E.; Zhou, M. V.; Kalyanaraman, C.; Lokey, R. S.; Jacobson, M. P. Conformational Flexibility, Internal Hydrogen Bonding, and Passive Membrane Permeability: Successful in Silico Prediction of the Relative Permeabilities of Cyclic Peptides. *J. Am. Chem. Soc.* **2006**, *128* (43), 14073–14080. https://doi.org/10.1021/ja063076p.

(349) Bergström, C. A. S.; Bolin, S.; Artursson, P.; Rönn, R.; Sandström, A. Hepatitis C Virus NS3 Protease Inhibitors: Large, Flexible Molecules of Peptide Origin Show Satisfactory Permeability across Caco-2 Cells. *Eur. J. Pharm. Sci.* **2009**, *38* (5), 556–563. https://doi.org/10.1016/j.ejps.2009.10.004.

(350) Morgenthaler, M.; Schweizer, E.; Hoffmann-Röder, A.; Benini, F.; Martin, R. E.; Jaeschke, G.; Wagner, B.; Fischer, H.; Bendels, S.; Zimmerli, D.; Schneider, J.; Diederich, F.; Kansy, M.; Müller, K. Predicting and Tuning Physicochemical Properties in Lead Optimization: Amine Basicities. *ChemMedChem* **2007**, *2* (8), 1100–1115. https://doi.org/10.1002/cmdc.200700059.

(351) Roseman, M. A. Hydrophilicity of Polar Amino Acid Side-Chains Is Markedly Reduced by Flanking Peptide Bonds. *J. Mol. Biol.* **1988**, *200* (3), 513–522. https://doi.org/10.1016/0022-2836(88)90540-2.

(352) Rubio-Aliaga, I.; Daniel, H. Peptide Transporters and Their Roles in Physiological Processes and Drug Disposition. *Xenobiotica* **2008**, *38* (7–8), 1022–1042. https://doi.org/10.1080/00498250701875254.

(353) del Amo, E. M.; Urtti, A.; Yliperttula, M. Pharmacokinetic Role of L-Type Amino Acid Transporters LAT1 and LAT2. *Eur. J. Pharm. Sci.* **2008**, *35* (3), 161–174. https://doi.org/10.1016/j.ejps.2008.06.015.

(354) Wright, N. J.; Fedor, J. G.; Zhang, H.; Jeong, P.; Suo, Y.; Yoo, J.; Hong, J.; Im, W.; Lee, S.-Y. Methotrexate Recognition by the Human Reduced Folate Carrier SLC19A1. *Nature* **2022**, *609* (7929), 1056–1062. https://doi.org/10.1038/s41586-022-05168-0.

(355) Vig, B.; Rautio, J. Amino Acid Prodrugs for Oral Delivery: Challenges and Opportunities. *Ther. Deliv.* **2011**, *2* (8), 959–962. https://doi.org/10.4155/tde.11.75.

(356) Ganapathy, M. E.; Brandsch, M.; Prasad, P. D.; Ganapathy, V.; Leibach, F. H. Differential Recognition of β-Lactam Antibiotics by Intestinal and Renal Peptide Transporters, PEPT 1 and PEPT 2. *J. Biol. Chem.* **1995**, *270* (43), 25672–25677. https://doi.org/10.1074/jbc.270.43.25672.

(357) Zhu, T.; Chen, X.; Steel, A.; Hedinger, M. A.; Smith, D. E. Differential Recognition of ACE Inhibitors in Xenopus Laevis Oocytes Expressing Rat PEPT1 and PEPT2. *Pharm. Res.* **2000**, *17*, 526–532.

(358) Friedman, D. I.; Amidon, G. L. Intestinal Absorption Mechanism of Dipeptide Angiotensin Converting Enzyme Inhibitors of the Lysyl-Proline Type: Lisinopril and SQ 29,852. *J. Pharm. Sci.* **1989**, *78* (12), 995–998. https://doi.org/10.1002/jps.2600781205.

(359) Killer, M.; Wald, J.; Pieprzyk, J.; Marlovits, T. C.; Löw, C. Structural Snapshots of Human PepT1 and PepT2 Reveal Mechanistic Insights into Substrate and Drug Transport across Epithelial Membranes. *Sci. Adv.* **2021**, *7* (45). https://doi.org/10.1126/sciadv.abk3259.

(360) Kottra, G.; Stamfort, A.; Daniel, H. PEPT1 as a Paradigm for Membrane Carriers That Mediate Electrogenic Bidirectional Transport of Anionic, Cationic, and Neutral Substrates. *J. Biol. Chem.* **2002**, *277* (36), 32683–32691. https://doi.org/10.1074/jbc.M204192200.

(361) Knütter, I.; Hartrodt, B.; Theis, S.; Foltz, M.; Rastetter, M.; Daniel, H.; Neubert, K.; Brandsch, M. Analysis of the Transport Properties of Side Chain Modified Dipeptides at the Mammalian Peptide Transporter PEPT1. *Eur. J. Pharm. Sci.* **2004**, *21* (1), 61–67.

https://doi.org/10.1016/S0928-0987(03)00141-6.

(362) Ganapathy, V.; Leibach, F. H. Is Intestinal Peptide Transport Energized by a Proton Gradient? *Am. J. Physiol. Liver Physiol.* **1985**, *249* (2), G153–G160. https://doi.org/10.1152/ajpgi.1985.249.2.G153.

(363) Drozdzik, M.; Gröer, C.; Penski, J.; Lapczuk, J.; Ostrowski, M.; Lai, Y.; Prasad, B.; Unadkat, J. D.; Siegmund, W.; Oswald, S. Protein Abundance of Clinically Relevant Multidrug Transporters along the Entire Length of the Human Intestine. *Mol. Pharm.* **2014**, *11* (10), 3547–3555. https://doi.org/10.1021/mp500330y.

(364) Shen, H.; Smith, D. E.; Yang, T.; Huang, Y. G.; Schnermann, J. B.; Brosius, F. C. Localization of PEPT1 and PEPT2 Proton-Coupled Oligopeptide Transporter MRNA and Protein in Rat Kidney. *Am. J. Physiol. Physiol.* **1999**, *276* (5), F658–F665. https://doi.org/10.1152/ajprenal.1999.276.5.F658.

(365) Daniel, H.; Morse, E. L.; Adibi, S. A. Determinants of Substrate Affinity for the Oligopeptide/H+ Symporter in the Renal Brush Border Membrane. *J. Biol. Chem.* **1992**, *267* (14), 9565–9573. https://doi.org/10.1016/S0021-9258(19)50128-4.

(366) Thomsen, A. E.; Friedrichsen, G. M.; Sørensen, A. H.; Andersen, R.; Nielsen, C. U.; Brodin, B.; Begtrup, M.; Frokjaer, S.; Steffansen, B. Prodrugs of Purine and Pyrimidine Analogues for the Intestinal Di/Tri-Peptide Transporter PepT1: Affinity for HPepT1 in Caco-2 Cells, Drug Release in Aqueous Media and in Vitro Metabolism. *J. Control. Release* **2003**, *86* (2–3), 279–292. https://doi.org/10.1016/S0168-3659(02)00413-3.

(367) Thomsen, A. E.; Christensen, M. S.; Bagger, M. A.; Steffansen, B. Acyclovir Prodrug for the Intestinal Di/Tri-Peptide Transporter PEPT1: Comparison of in Vivo Bioavailability in Rats and Transport in Caco-2 Cells. *Eur. J. Pharm. Sci.* **2004**, *23* (4–5), 319–325. https://doi.org/10.1016/j.ejps.2004.08.005.

(368) Döring, F.; Will, J.; Amasheh, S.; Clauss, W.; Ahlbrecht, H.; Daniel, H. Minimal Molecular Determinants of Substrates for Recognition by the Intestinal Peptide Transporter. *J. Biol. Chem.* **1998**, *273* (36), 23211–23218. https://doi.org/10.1074/jbc.273.36.23211.

(369) Turk, V.; Stoka, V.; Vasiljeva, O.; Renko, M.; Sun, T.; Turk, B.; Turk, D. Cysteine Cathepsins: From Structure, Function and Regulation to New Frontiers. *Biochim. Biophys. Acta - Proteins Proteomics* **2012**, *1824* (1), 68–88. https://doi.org/10.1016/j.bbapap.2011.10.002.

(370) Pisoni, R. L.; Acker, T. L.; Lisowski, K. M.; Lemons, R. M.; Thoene, J. G. A Cysteine-Specific Lysosomal Transport System Provides a Major Route for the Delivery of Thiol to Human Fibroblast Lysosomes: Possible Role in Supporting Lysosomal Proteolysis. *J. Cell Biol.* **1990**, *110* (2), 327–335. https://doi.org/10.1083/jcb.110.2.327.

(371) Adelmann, C. H.; Traunbauer, A. K.; Chen, B.; Condon, K. J.; Chan, S. H.; Kunchok, T.; Lewis, C. A.; Sabatini, D. M. MFSD12 Mediates the Import of Cysteine into Melanosomes and Lysosomes. *Nature* **2020**, *588* (7839), 699–704. https://doi.org/10.1038/s41586-020-2937-x.

(372) Mindell, J. A. Lysosomal Acidification Mechanisms. *Annu. Rev. Physiol.* **2012**, *74* (1), 69–86. https://doi.org/10.1146/annurev-physiol-012110-142317.

(373) Johnson, D. E.; Ostrowski, P.; Jaumouillé, V.; Grinstein, S. The Position of Lysosomes within the Cell Determines Their Luminal PH. *J. Cell Biol.* **2016**, *212* (6), 677–692. https://doi.org/10.1083/jcb.201507112.

(374) Anes, E.; Pires, D.; Mandal, M.; Azevedo-Pereira, J. M. Spatial Localization of Cathepsins: Implications in Immune Activation and Resolution during Infections. *Front. Immunol.* **2022**, *13*. https://doi.org/10.3389/fimmu.2022.955407.

(375) Roberts, R. Lysosomal Cysteine Proteases: Structure, Function and Inhibition of Cathepsins. *Drug News Perspect.* **2005**, *18* (10), 605. https://doi.org/10.1358/dnp.2005.18.10.949485.

(376) *Handbook of Proteolytic Enzymes*, 3rd ed.; Rawlings, N. D., Salvesen, G., Eds.; Elsevier, 2013.

(377) Gohda, E.; Pitot, H. C. A New Thiol Proteinase from Rat Liver. *J. Biol. Chem.* **1981**, *256* (5), 2567–2572. https://doi.org/10.1016/S0021-9258(19)69820-0.

(378) Gohda, E.; Pitot, H. C. Chapter 557 - Cathepsin T. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 2485–2487. https://doi.org/10.1016/B978-0-12-382219-2.00556-1.

(379) Deussing, J. M. Chapter 417 - Placenta-Specific Cathepsins. In *Handbook of Proteolytic Enzymes*; Elsevier, 2013; pp 1845–1851. https://doi.org/10.1016/B978-0-12-382219-2.00416-6.

(380) Maciewicz, R. A.; Etherington, D. J. A Comparison of Four Cathepsins (B, L, N and S) with Collagenolytic Activity from Rabbit Spleen. *Biochem. J.* **1988**, *256* (2), 433–440. https://doi.org/10.1042/bj2560433.

(381) Lutgens, S. P. M.; Cleutjens, K. B. J. M.; Daemen, M. J. A. P.; Heeneman, S. Cathepsin Cysteine Proteases in Cardiovascular Disease. *FASEB J.* **2007**, *21* (12), 3029–3041. https://doi.org/10.1096/fj.06-7924com.

(382) Ishidoh, K.; Kominami, E. Processing and Activation of Lysosomal Proteinases. *Biol. Chem.* **2002**, *383* (12), 1827–1831. https://doi.org/10.1515/BC.2002.206.

(383) Hasilik, A.; Wrocklage, C.; Schroder, B. Intracellular Trafficking of Lysosomal Proteins and Lysosomes. *Int. J. Clin. Pharmacol. Ther.* **2009**, *47* (Suppl. 1), 18–33.

(384) Saftig, P.; Klumperman, J. Lysosome Biogenesis and Lysosomal Membrane Proteins: Trafficking Meets Function. *Nat. Rev. Mol. Cell Biol.* **2009**, *10* (9), 623–635. https://doi.org/10.1038/nrm2745.

(385) Quraishi, O.; Storer, A. C. Identification of Internal Autoproteolytic Cleavage Sites within the Prosegments of Recombinant Procathepsin B and Procathepsin S. *J. Biol. Chem.* **2001**, *276* (11), 8118–8124. https://doi.org/10.1074/jbc.M005851200.

(386) Dahl, S. W.; Halkier, T.; Lauritzen, C.; Dolenc, I.; Pedersen, J.; Turk, V.; Turk, B. Human Recombinant Pro-Dipeptidyl Peptidase I (Cathepsin C) Can Be Activated by Cathepsins L and S but Not by Autocatalytic Processing. *Biochemistry* **2001**, *40* (6), 1671–1678. https://doi.org/10.1021/bi001693z.

(387) Yadati, T.; Houben, T.; Bitorina, A.; Shiri-Sverdlov, R. The Ins and Outs of Cathepsins: Physiological Function and Role in Disease Management. *Cells* **2020**, *9* (7), 1679. https://doi.org/10.3390/cells9071679.

(388) Vidak, E.; Javoršek, U.; Vizovišek, M.; Turk, B. Cysteine Cathepsins and Their Extracellular Roles: Shaping the Microenvironment. *Cells* **2019**, *8* (3), 264. https://doi.org/10.3390/cells8030264.

(389) Yu, J.; Adapala, N. S.; Doherty, L.; Sanjay, A. Cbl-PI3K Interaction Regulates Cathepsin K Secretion in Osteoclasts. *Bone* **2019**, *127*, 376–385. https://doi.org/10.1016/j.bone.2019.07.009.

(390) Frizler, M.; Stirnberg, M.; Sisay, M.; Gutschow, M. Development of Nitrile-Based Peptidic Inhibitors of Cysteine Cathepsins. *Curr. Top. Med. Chem.* **2010**, *10* (3), 294–322. https://doi.org/10.2174/156802610790725452.

(391) Vizovišek, M.; Vidak, E.; Javoršek, U.; Mikhaylov, G.; Bratovš, A.; Turk, B. Cysteine Cathepsins as Therapeutic Targets in Inflammatory Diseases. *Expert Opin. Ther. Targets* **2020**, *24* (6), 573–588. https://doi.org/10.1080/14728222.2020.1746765.

(392) Biasizzo, M.; Javoršek, U.; Vidak, E.; Zarić, M.; Turk, B. Cysteine Cathepsins: A Long and Winding Road towards Clinics. *Mol. Aspects Med.* **2022**, *88*, 101150. https://doi.org/10.1016/j.mam.2022.101150.

(393) Rossi, A.; Deveraux, Q.; Turk, B.; Sali, A. Comprehensive Search for Cysteine Cathepsins in the

Human Genome. *Biol. Chem.* **2004**, *385* (5). https://doi.org/10.1515/BC.2004.040.

(394) Kolli, N.; Garman, S. C. Proteolytic Activation of Human Cathepsin A. *J. Biol. Chem.* **2014**, *289* (17), 11592–11600. https://doi.org/10.1074/jbc.M113.524280.

(395) Tillner, J.; Lehmann, A.; Paehler, T.; Lukacs, Z.; Ruf, S.; Sadowski, T.; Pinquier, J.; Ruetten, H. Tolerability, Safety, and Pharmacokinetics of the Novel Cathepsin A Inhibitor SAR164653 in Healthy Subjects. *Clin. Pharmacol. Drug Dev.* **2016**, *5* (1), 57–68. https://doi.org/10.1002/cpdd.201.

(396) Korkmaz, B.; Caughey, G. H.; Chapple, I.; Gauthier, F.; Hirschfeld, J.; Jenne, D. E.; Kettritz, R.; Lalmanach, G.; Lamort, A.-S.; Lauritzen, C.; Łęgowska, M.; Lesner, A.; Marchand-Adam, S.; McKaig, S. J.; Moss, C.; Pedersen, J.; Roberts, H.; Schreiber, A.; Seren, S.; Thakker, N. S. Therapeutic Targeting of Cathepsin C: From Pathophysiology to Treatment. *Pharmacol. Ther.* **2018**, *190*, 202–236. https://doi.org/10.1016/j.pharmthera.2018.05.011.

(397) Chalmers, J. D.; Gupta, A.; Chotirmall, S. H.; Armstrong, A.; Eickholz, P.; Hasegawa, N.; McShane, P. J.; O'Donnell, A. E.; Shteinberg, M.; Watz, H.; Eleftheraki, A.; Diefenbach, C.; Sauter, W. A Phase 2 Randomised Study to Establish Efficacy, Safety and Dosing of a Novel Oral Cathepsin C Inhibitor, BI 1291583, in Adults with Bronchiectasis: Airleaf. *ERJ Open Res.* **2023**, *9* (3), 00633–02022. https://doi.org/10.1183/23120541.00633-2022.

(398) University of Dundee. *NCT04817332 - STOP-COVID19: Superiority Trial Of Protease Inhibition in COVID-19 (STOP-COVID19)*. https://clinicaltrials.gov/study/NCT04817332?term=Brensocatib &page=2&rank=11.

(399) Insmed Incorporated. *NCT04594369 - A Study to Assess the Efficacy, Safety, and Tolerability of Brensocatib in Participants With Non-Cystic Fibrosis Bronchiectasis (ASPEN)*. https://clinicaltrials.gov/study/NCT04594369.

(400) Houben, T.; Oligschlaeger, Y.; Hendrikx, T.; Bitorina, A. V.; Walenbergh, S. M. A.; van Gorp, P. J.; Gijbels, M. J. J.; Friedrichs, S.; Plat, J.; Schaap, F. G.; Lütjohann, D.; Hofker, M. H.; Shiri-Sverdlov, R. Cathepsin D Regulates Lipid Metabolism in Murine Steatohepatitis. *Sci. Rep.* **2017**, *7* (1), 3494. https://doi.org/10.1038/s41598-017-03796-5.

(401) Mijanovic, O.; Petushkova, A. I.; Brankovic, A.; Turk, B.; Solovieva, A. B.; Nikitkina, A. I.; Bolevich, S.; Timashev, P. S.; Parodi, A.; Zamyatnin, A. A. Cathepsin D—Managing the Delicate Balance. *Pharmaceutics* **2021**, *13* (6), 837. https://doi.org/10.3390/pharmaceutics13060837.

(402) Drake, M. T.; Clarke, B. L.; Oursler, M. J.; Khosla, S. Cathepsin K Inhibitors for Osteoporosis: Biology, Potential Clinical Utility, and Lessons Learned. *Endocr. Rev.* **2017**, *38* (4), 325–350. https://doi.org/10.1210/er.2015-1114.

(403) Brixen, K.; Chapurlat, R.; Cheung, A. M.; Keaveny, T. M.; Fuerst, T.; Engelke, K.; Recker, R.; Dardzinski, B.; Verbruggen, N.; Ather, S.; Rosenberg, E.; de Papp, A. E. Bone Density, Turnover, and Estimated Strength in Postmenopausal Women Treated With Odanacatib: A Randomized Trial. *J. Clin. Endocrinol. Metab.* **2013**, *98* (2), 571–580. https://doi.org/10.1210/jc.2012-2972.

(404) Eastell, R.; Nagase, S.; Ohyama, M.; Small, M.; Sawyer, J.; Boonen, S.; Spector, T.; Kuwayama, T.; Deacon, S. Safety and Efficacy of the Cathepsin K Inhibitor ONO-5334 in Postmenopausal Osteoporosis: The OCEAN Study. *J. Bone Miner. Res.* **2011**, *26* (6), 1303–1312. https://doi.org/10.1002/jbmr.341.

(405) Medivir. *NCT02705625 - A Study to Evaluate the Efficacy, Safety and Tolerability of MIV-711 in Osteoarthritis Patients*. https://clinicaltrials.gov/study/NCT02705625?term=MIV-711&rank=3&tab=results.

(406) Adami, S.; Hala, T.; Brown, J. P.; Garnero, P.; Haemmerle, S.; Ortmann, C. E.; Bouisset, F.; Trechsel, U. 2006 Abstracts: Twenty-Eighth Annual Meeting of the American Society for Bone and Mineral Research: Pennsylvania Convention Convention Center Philadelphia, Pennsylvania,

USA, September 15–19, 2006; 1085. *J. Bone Miner. Res.* **2006**, *21* (S1), S2–S51. https://doi.org/10.1002/jbmr.5650211402.

(407) GlaxoSmithKline. *NCT00411190 - Study To Determine The Effects Of Doses Of Relacatib On The Metabolism Of Acetaminophen, Ibuprofen And Atorvastatin*. https://clinicaltrials.gov/study/NCT00411190?term=SB-462795&rank=1.

(408) McDowell, S. H.; Gallaher, S. A.; Burden, R. E.; Scott, C. J. Leading the Invasion: The Role of Cathepsin S in the Tumour Microenvironment. *Biochim. Biophys. Acta - Mol. Cell Res.* **2020**, *1867* (10), 118781. https://doi.org/10.1016/j.bbamcr.2020.118781.

(409) Linnevers, C.; Smeekens, S. .; Brömme, D. Human Cathepsin W, a Putative Cysteine Protease Predominantly Expressed in CD8 + T-lymphocytes. *FEBS Lett.* **1997**, *405* (3), 253–259. https://doi.org/10.1016/S0014-5793(97)00118-X.

(410) Wex, T.; Bühling, F.; Wex, H.; Günther, D.; Malfertheiner, P.; Weber, E.; Brömme, D. Human Cathepsin W, a Cysteine Protease Predominantly Expressed in NK Cells, Is Mainly Localized in the Endoplasmic Reticulum. *J. Immunol.* **2001**, *167* (4), 2172–2178. https://doi.org/10.4049/jimmunol.167.4.2172.

(411) Organisation for Economic Co-operation and Development (OECD). *Health at a Glance 2023 - OECD Indicators*; Health at a Glance; OECD, 2023. https://doi.org/10.1787/7a7afb35-en.

(412) Siegel, R. L.; Giaquinto, A. N.; Jemal, A. Cancer Statistics, 2024. *CA. Cancer J. Clin.* **2024**, *74* (1), 12–49. https://doi.org/10.3322/caac.21820.

(413) Hajdu, S. I.; Darvishian, F. A Note from History: Landmarks in History of Cancer, Part 5. *Cancer* **2013**, *119* (8), 1450–1466. https://doi.org/10.1002/cncr.27889.

(414) Hajdu, S. I. A Note from History: Landmarks in History of Cancer, Part 3. *Cancer* **2012**, *118* (4), 1155–1168. https://doi.org/10.1002/cncr.26320.

(415) Hosoya, N.; Miyagawa, K. Targeting DNA Damage Response in Cancer Therapy. *Cancer Sci.* **2014**, *105* (4), 370–388. https://doi.org/10.1111/cas.12366.

(416) Woods, D.; Turchi, J. J. Chemotherapy Induced DNA Damage Response. *Cancer Biol. Ther.* **2013**, *14* (5), 379–389. https://doi.org/10.4161/cbt.23761.

(417) Portugal, J.; Mansilla, S.; Bataller, M. Mechanisms of Drug-Induced Mitotic Catastrophe in Cancer Cells. *Curr. Pharm. Des.* **2010**, *16* (1), 69–78. https://doi.org/10.2174/138161210789941801.

(418) Das, S.; Kulkarni, S.; Singh, Y.; Kumar, P.; Thareja, S. Selective Estrogen Receptor Modulators (SERMs) for the Treatment of ER+ Breast Cancer: An Overview. *J. Mol. Struct.* **2022**, *1270*, 133853. https://doi.org/10.1016/j.molstruc.2022.133853.

(419) Meher, N.; VanBrocklin, H. F.; Wilson, D. M.; Flavell, R. R. PSMA-Targeted Nanotheranostics for Imaging and Radiotherapy of Prostate Cancer. *Pharmaceuticals* **2023**, *16* (2), 315. https://doi.org/10.3390/ph16020315.

(420) Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **2011**, *144* (5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013.

(421) Jacquet, P.; Stéphanou, A. A Computational Model of the Spatiotemporal Adaptation of Tumor Cells Metabolism in a Growing Spheroid. September 14, 2023. https://doi.org/10.1101/2023.09.11.557115.

(422) Weidemann, A.; Johnson, R. S. Biology of HIF-1α. *Cell Death Differ.* **2008**, *15* (4), 621–627. https://doi.org/10.1038/cdd.2008.12.

(423) Sonveaux, P.; Copetti, T.; De Saedeleer, C. J.; Végran, F.; Verrax, J.; Kennedy, K. M.; Moon, E. J.; Dhup, S.; Danhier, P.; Frérart, F.; Gallez, B.; Ribeiro, A.; Michiels, C.; Dewhirst, M. W.;

Feron, O. Targeting the Lactate Transporter MCT1 in Endothelial Cells Inhibits Lactate-Induced HIF-1 Activation and Tumor Angiogenesis. *PLoS One* **2012**, *7* (3), e33418. https://doi.org/10.1371/journal.pone.0033418.

(424) Whiteside, T. L. The Tumor Microenvironment and Its Role in Promoting Tumor Growth. *Oncogene* **2008**, *27* (45), 5904–5912. https://doi.org/10.1038/onc.2008.271.

(425) Olson, O. C.; Joyce, J. A. Cysteine Cathepsin Proteases: Regulators of Cancer Progression and Therapeutic Response. *Nat. Rev. Cancer* **2015**, *15* (12), 712–729. https://doi.org/10.1038/nrc4027.

(426) Park, K. C.; Dharmasivam, M.; Richardson, D. R. The Role of Extracellular Proteases in Tumor Progression and the Development of Innovative Metal Ion Chelators That Inhibit Their Activity. *Int. J. Mol. Sci.* **2020**, *21* (18), 6805. https://doi.org/10.3390/ijms21186805.

(427) Kessenbrock, K.; Plaks, V.; Werb, Z. Matrix Metalloproteinases: Regulators of the Tumor Microenvironment. *Cell* **2010**, *141* (1), 52–67. https://doi.org/10.1016/j.cell.2010.03.015.

(428) Miller, M. A.; Sullivan, R. J.; Lauffenburger, D. A. Molecular Pathways: Receptor Ectodomain Shedding in Treatment, Resistance, and Monitoring of Cancer. *Clin. Cancer Res.* **2017**, *23* (3), 623–629. https://doi.org/10.1158/1078-0432.CCR-16-0869.

(429) Overall, C. M.; López-Otín, C. Strategies for MMP Inhibition in Cancer: Innovations for the Post-Trial Era. *Nat. Rev. Cancer* **2002**, *2* (9), 657–672. https://doi.org/10.1038/nrc884.

(430) Almutairi, S.; Kalloush, H. M.; Manoon, N. A.; Bardaweel, S. K. Matrix Metalloproteinases Inhibitors in Cancer Treatment: An Updated Review (2013–2023). *Molecules* **2023**, *28* (14), 5567. https://doi.org/10.3390/molecules28145567.

(431) Prunk, M.; Kos, J. Nanoparticle Based Delivery of Protease Inhibitors to Cancer Cells. *Curr. Med. Chem.* **2018**, *24* (42). https://doi.org/10.2174/0929867323666160922162811.

(432) Burden, R. E.; Gormley, J. A.; Jaquin, T. J.; Small, D. M.; Quinn, D. J.; Hegarty, S. M.; Ward, C.; Walker, B.; Johnston, J. A.; Olwill, S. A.; Scott, C. J. Antibody-Mediated Inhibition of Cathepsin S Blocks Colorectal Tumor Invasion and Angiogenesis. *Clin. Cancer Res.* **2009**, *15* (19), 6042–6051. https://doi.org/10.1158/1078-0432.CCR-09-1262.

(433) Fernández, P. L.; Farré, X.; Nadal, A.; Fernández, E.; Peiró, N.; Sloane, B. F.; Shi, G.-P.; Chapman, H. A.; Campo, E.; Cardesa, A. Expression of Cathepsins B and S in the Progression of Prostate Carcinoma. *Int. J. Cancer* **2001**, *95* (1), 51–55. https://doi.org/10.1002/1097-0215(20010120)95:1<51::AID-IJC1009>3.0.CO;2-J.

(434) Lindahl, C.; Simonsson, M.; Bergh, A.; Thysell, E.; Antti, H.; Sund, M.; Wikström, P. Increased Levels of Macrophage-Secreted Cathepsin S during Prostate Cancer Progression in TRAMP Mice and Patients. *Cancer Genomics Proteomics* **2009**, *6* (3), 149–159.

(435) Joyce, J. A.; Baruch, A.; Chehade, K.; Meyer-Morse, N.; Giraudo, E.; Tsai, F.-Y.; Greenbaum, D. C.; Hager, J. H.; Bogyo, M.; Hanahan, D. Cathepsin Cysteine Proteases Are Effectors of Invasive Growth and Angiogenesis during Multistage Tumorigenesis. *Cancer Cell* **2004**, *5* (5), 443–453. https://doi.org/10.1016/S1535-6108(04)00111-4.

(436) Small, D. M.; Burden, R. E.; Jaworski, J.; Hegarty, S. M.; Spence, S.; Burrows, J. F.; McFarlane, C.; Kissenpfennig, A.; McCarthy, H. O.; Johnston, J. A.; Walker, B.; Scott, C. J. Cathepsin S from Both Tumor and Tumor-associated Cells Promote Cancer Growth and Neovascularization. *Int. J. Cancer* **2013**, *133* (9), 2102–2112. https://doi.org/10.1002/ijc.28238.

(437) Sevenich, L.; Bowman, R. L.; Mason, S. D.; Quail, D. F.; Rapaport, F.; Elie, B. T.; Brogi, E.; Brastianos, P. K.; Hahn, W. C.; Holsinger, L. J.; Massagué, J.; Leslie, C. S.; Joyce, J. A. Analysis of Tumour- and Stroma-Supplied Proteolytic Networks Reveals a Brain-Metastasis-Promoting Role for Cathepsin S. *Nat. Cell Biol.* **2014**, *16* (9), 876–888. https://doi.org/10.1038/ncb3011.

(438) Flannery, T.; McQuaid, S.; McGoohan, C.; McConnell, R. S.; McGregor, G.; Mirakhur, M.; Hamilton, P.; Diamond, J.; Cran, G.; Walker, B.; Scott, C.; Martin, L.; Ellison, D.; Patel, C.; Nicholson, C.; Mendelow, D.; McCormick, D.; Johnston, P. G. Cathepsin S Expression: An Independent Prognostic Factor in Glioblastoma Tumours—a Pilot Study. *Int. J. Cancer* **2006**, *119* (4), 854–860. https://doi.org/10.1002/ijc.21911.

(439) Flannery, T.; Gibson, D.; Mirakhur, M.; McQuaid, S.; Greenan, C.; Trimble, A.; Walker, B.; McCormick, D.; Johnston, P. G. The Clinical Significance of Cathepsin S Expression in Human Astrocytomas. *Am. J. Pathol.* **2003**, *163* (1), 175–182. https://doi.org/10.1016/S0002-9440(10)63641-3.

(440) Gautam, J.; Bae, Y. K.; Kim, J.-A. Up-Regulation of Cathepsin S Expression by HSP90 and 5-HT7 Receptor-Dependent Serotonin Signaling Correlates with Triple Negativity of Human Breast Cancer. *Breast Cancer Res. Treat.* **2017**, *161* (1), 29–40. https://doi.org/10.1007/s10549-016-4027-1.

(441) Storm van's Gravesande, K.; Layne, M. D.; Ye, Q.; Le, L.; Baron, R. M.; Perrella, M. A.; Santambrogio, L.; Silverman, E. S.; Riese, R. J. IFN Regulatory Factor-1 Regulates IFN-γ-Dependent Cathepsin S Expression. *J. Immunol.* **2002**, *168* (9), 4488–4494. https://doi.org/10.4049/jimmunol.168.9.4488.

(442) Gocheva, V.; Wang, H.-W.; Gadea, B. B.; Shree, T.; Hunter, K. E.; Garfall, A. L.; Berman, T.; Joyce, J. A. IL-4 Induces Cathepsin Protease Activity in Tumor-Associated Macrophages to Promote Cancer Growth and Invasion. *Genes Dev.* **2010**, *24* (3), 241–255. https://doi.org/10.1101/gad.1874010.

(443) Lee, H.-J.; Lee, Y.-J.; Kang, C.-M.; Bae, S.; Jeoung, D.; Jang, J.-J.; Lee, S.-S.; Cho, C.-K.; Lee, Y.-S. Differential Gene Signatures in Rat Mammary Tumors Induced by DMBA and Those Induced by Fractionated γ Radiation. *Radiat. Res.* **2008**, *170* (5), 579. https://doi.org/10.1667/RR1106.1.

(444) Seo, H. R.; Bae, S.; Lee, Y. Radiation-induced Cathepsin S Is Involved in Radioresistance. *Int. J. Cancer* **2009**, *124* (8), 1794–1801. https://doi.org/10.1002/ijc.24095.

(445) Shree, T.; Olson, O. C.; Elie, B. T.; Kester, J. C.; Garfall, A. L.; Simpson, K.; Bell-McGuinn, K. M.; Zabor, E. C.; Brogi, E.; Joyce, J. A. Macrophages and Cathepsin Proteases Blunt Chemotherapeutic Response in Breast Cancer. *Genes Dev.* **2011**, *25* (23), 2465–2479. https://doi.org/10.1101/gad.180331.111.

(446) Burden, R. E.; Gormley, J. A.; Kuehn, D.; Ward, C.; Kwok, H. F.; Gazdoiu, M.; McClurg, A.; Jaquin, T. J.; Johnston, J. A.; Scott, C. J.; Olwill, S. A. Inhibition of Cathepsin S by Fsn0503 Enhances the Efficacy of Chemotherapy in Colorectal Carcinomas. *Biochimie* **2012**, *94* (2), 487–493. https://doi.org/10.1016/j.biochi.2011.08.017.

(447) Lee-Dutra, A.; Wiener, D. K.; Sun, S. Cathepsin S Inhibitors: 2004 – 2010. *Expert Opin. Ther. Pat.* **2011**, *21* (3), 311–337. https://doi.org/10.1517/13543776.2011.553800.

(448) Podolin, P. L.; Bolognese, B. J.; Carpenter, D. C.; Davis, T. G.; Johanson, R. A.; Fox, J. H.; Long, E.; Dong, X.; Marquis, R. W.; LoCastro, S. M.; Terfloth, G. J.; Kurali, E.; Peterson, J. J.; Smith, B. R.; McQueney, M. S.; Yamashita, D. S.; Capper-Spudich, E. A. Inhibition of Invariant Chain Processing, Antigen-Induced Proliferative Responses, and the Development of Collagen-Induced Arthritis and Experimental Autoimmune Encephalomyelitis by a Small Molecule Cysteine Protease Inhibitor. *J. Immunol.* **2008**, *180* (12), 7989–8003. https://doi.org/10.4049/jimmunol.180.12.7989.

(449) Yan, X.; Wu, C.; Chen, T.; Santos, M. M.; Liu, C.-L.; Yang, C.; Zhang, L.; Ren, J.; Liao, S.; Guo, H.; Sukhova, G. K.; Shi, G.-P. Cathepsin S Inhibition Changes Regulatory T-Cell Activity in Regulating Bladder Cancer and Immune Cell Proliferation and Apoptosis. *Mol. Immunol.* **2017**, *82*, 66–74. https://doi.org/10.1016/j.molimm.2016.12.018.

(450) Dheilly, E.; Battistello, E.; Katanayeva, N.; Sungalee, S.; Michaux, J.; Duns, G.; Wehrle, S.; Sordet-Dessimoz, J.; Mina, M.; Racle, J.; Farinha, P.; Coukos, G.; Gfeller, D.; Mottok, A.; Kridel, R.; Correia, B. E.; Steidl, C.; Bassani-Sternberg, M.; Ciriello, G.; Zoete, V.; Oricchio, E. Cathepsin S Regulates Antigen Processing and T Cell Activity in Non-Hodgkin Lymphoma. *Cancer Cell* **2020**, *37* (5), 674-689.e12. https://doi.org/10.1016/j.ccell.2020.03.016.

(451) Bararia, D.; Hildebrand, J. A.; Stolz, S.; Haebe, S.; Alig, S.; Trevisani, C. P.; Osorio-Barrios, F.; Bartoschek, M. D.; Mentz, M.; Pastore, A.; Gaitzsch, E.; Heide, M.; Jurinovic, V.; Rautter, K.; Gunawardana, J.; Sabdia, M. B.; Szczepanowski, M.; Richter, J.; Klapper, W.; Louissaint, A.; Ludwig, C.; Bultmann, S.; Leonhardt, H.; Eustermann, S.; Hopfner, K.-P.; Hiddemann, W.; von Bergwelt-Baildon, M.; Steidl, C.; Kridel, R.; Tobin, J. W. D.; Gandhi, M. K.; Weinstock, D. M.; Schmidt-Supprian, M.; Sárosi, M. B.; Rudelius, M.; Passerini, V.; Mautner, J.; Weigert, O. Cathepsin S Alterations Induce a Tumor-Promoting Immune Microenvironment in Follicular Lymphoma. *Cell Rep.* **2020**, *31* (5), 107522. https://doi.org/10.1016/j.celrep.2020.107522.

(452) Gupta, S.; Singh, R. K.; Dastidar, S.; Ray, A. Cysteine Cathepsin S as an Immunomodulatory Target: Present and Future Trends. *Expert Opin. Ther. Targets* **2008**, *12* (3), 291–299. https://doi.org/10.1517/14728222.12.3.291.

(453) Riese, R. J.; Wolf, P. R.; Brömme, D.; Natkin, L. R.; Villadangos, J. A.; Ploegh, H. L.; Chapman, H. A. Essential Role for Cathepsin S in MHC Class II–Associated Invariant Chain Processing and Peptide Loading. *Immunity* **1996**, *4* (4), 357–366. https://doi.org/10.1016/S1074-7613(00)80249-6.

(454) Driessen, C.; Bryant, R. A. R.; Lennon-Duménil, A.-M.; Villadangos, J. A.; Bryant, P. W.; Shi, G.-P.; Chapman, H. A.; Ploegh, H. L. Cathepsin S Controls the Trafficking and Maturation of Mhc Class II Molecules in Dendritic Cells. *J. Cell Biol.* **1999**, *147* (4), 775–790. https://doi.org/10.1083/jcb.147.4.775.

(455) Shi, G.-P.; Villadangos, J. A.; Dranoff, G.; Small, C.; Gu, L.; Haley, K. J.; Riese, R.; Ploegh, H. L.; Chapman, H. A. Cathepsin S Required for Normal MHC Class II Peptide Loading and Germinal Center Development. *Immunity* **1999**, *10* (2), 197–206. https://doi.org/10.1016/S1074-7613(00)80020-5.

(456) Li, X. C.; Raghavan, M. Structure and Function of Major Histocompatibility Complex Class I Antigens. *Curr. Opin. Organ Transplant.* **2010**, *15* (4), 499–504. https://doi.org/10.1097/MOT.0b013e32833bfb33.

(457) Germain, R. N. MHC-Dependent Antigen Processing and Peptide Presentation: Providing Ligands for T Lymphocyte Activation. *Cell* **1994**, *76* (2), 287–299. https://doi.org/10.1016/0092-8674(94)90336-0.

(458) von Boehmer, H.; Kisielow, P. Self-Nonself Discrimination by T Cells. *Science (80-.).* **1990**, *248* (4961), 1369–1373. https://doi.org/10.1126/science.1972594.

(459) Pettmann, J.; Awada, L.; Różycki, B.; Huhn, A.; Faour, S.; Kutuzov, M.; Limozin, L.; Weikl, T. R.; van der Merwe, P. A.; Robert, P.; Dushek, O. Mechanical Forces Impair Antigen Discrimination by Reducing Differences in T-cell Receptor/Peptide–MHC Off-rates. *EMBO J.* **2023**, *42* (7). https://doi.org/10.15252/embj.2022111841.

(460) Neefjes, J.; Jongsma, M. L. M.; Paul, P.; Bakke, O. Towards a Systems Understanding of MHC Class I and MHC Class II Antigen Presentation. *Nat. Rev. Immunol.* **2011**, *11* (12), 823–836. https://doi.org/10.1038/nri3084.

(461) Murphy, K.; Weaver, C. *Janeway Immunologie*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2018. https://doi.org/10.1007/978-3-662-56004-4.

(462) Trombetta, E. S.; Mellman, I. Cell Biology of Antigen Processing in Vitro and in Vivo. *Annu. Rev. Immunol.* **2005**, *23* (1), 975–1028. https://doi.org/10.1146/annurev.immunol.22.012703.104538.

(463) Cresswell, P.; Ackerman, A. L.; Giodini, A.; Peaper, D. R.; Wearsch, P. A. Mechanisms of MHC Class I-restricted Antigen Processing and Cross-presentation. *Immunol. Rev.* **2005**, *207* (1), 145–157. https://doi.org/10.1111/j.0105-2896.2005.00316.x.

(464) Raghavan, M.; Del Cid, N.; Rizvi, S. M.; Peters, L. R. MHC Class I Assembly: Out and About. *Trends Immunol.* **2008**, *29* (9), 436–443. https://doi.org/10.1016/j.it.2008.06.004.

(465) Trowsdale, J. HLA Genomics in the Third Millennium. *Curr. Opin. Immunol.* **2005**, *17* (5), 498–504. https://doi.org/10.1016/j.coi.2005.07.015.

(466) Fehling, H. J.; Viville, S.; Van Ewijk, W.; Benoist, C.; Mathis, D. Fine-Tuning of MHC Class II Gene Expression in Defined Microenvironments. *Trends Genet.* **1989**, *5*, 342–347.

(467) Cresswell, P. Invariant Chain Structure and MHC Class II Function. *Cell* **1996**, *84* (4), 505–507. https://doi.org/10.1016/S0092-8674(00)81025-9.

(468) Roche, P. A.; Furuta, K. The Ins and Outs of MHC Class II-Mediated Antigen Processing and Presentation. *Nat. Rev. Immunol.* **2015**, *15* (4), 203–216. https://doi.org/10.1038/nri3818.

(469) Tay, R. E.; Richardson, E. K.; Toh, H. C. Revisiting the Role of CD4+ T Cells in Cancer Immunotherapy—New Insights into Old Paradigms. *Cancer Gene Ther.* **2021**, *28* (1–2), 5–17. https://doi.org/10.1038/s41417-020-0183-x.

(470) Wilkinson, R. D. A.; Magorrian, S. M.; Williams, R.; Young, A.; Small, D. M.; Scott, C. J.; Burden, R. E. CCL2 Is Transcriptionally Controlled by the Lysosomal Protease Cathepsin S in a CD74-Dependent Manner. *Oncotarget* **2015**, *6* (30), 29725–29739. https://doi.org/10.18632/oncotarget.5065.

(471) Gauthier, J. Y.; Black, W. C.; Courchesne, I.; Cromlish, W.; Desmarais, S.; Houle, R.; Lamontagne, S.; Li, C. S.; Massé, F.; McKay, D. J.; Ouellet, M.; Robichaud, J.; Truchon, J.-F.; Truong, V.-L.; Wang, Q.; Percival, M. D. The Identification of Potent, Selective, and Bioavailable Cathepsin S Inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *17* (17), 4929–4933. https://doi.org/10.1016/j.bmcl.2007.06.023.

(472) Borghese, F.; Clanchy, F. IL. CD74: An Emerging Opportunity as a Therapeutic Target in Cancer and Autoimmune Disease. *Expert Opin. Ther. Targets* **2011**, *15* (3), 237–251. https://doi.org/10.1517/14728222.2011.550879.

(473) Mosser, D. M.; Edwards, J. P. Exploring the Full Spectrum of Macrophage Activation. *Nat. Rev. Immunol.* **2008**, *8* (12), 958–969. https://doi.org/10.1038/nri2448.

(474) Biswas, S. K.; Mantovani, A. Macrophage Plasticity and Interaction with Lymphocyte Subsets: Cancer as a Paradigm. *Nat. Immunol.* **2010**, *11* (10), 889–896. https://doi.org/10.1038/ni.1937.

(475) Quail, D. F.; Joyce, J. A. Microenvironmental Regulation of Tumor Progression and Metastasis. *Nat. Med.* **2013**, *19* (11), 1423–1437. https://doi.org/10.1038/nm.3394.

(476) Mu, X.; Shi, W.; Xu, Y.; Xu, C.; Zhao, T.; Geng, B.; Yang, J.; Pan, J.; Hu, S.; Zhang, C.; Zhang, J.; Wang, C.; Shen, J.; Che, Y.; Liu, Z.; Lv, Y.; Wen, H.; You, Q. Tumor-Derived Lactate Induces M2 Macrophage Polarization via the Activation of the ERK/STAT3 Signaling Pathway in Breast Cancer. *Cell Cycle* **2018**, *17* (4), 428–438. https://doi.org/10.1080/15384101.2018.1444305.

(477) Mathew, R.; Karantza-Wadsworth, V.; White, E. Role of Autophagy in Cancer. *Nat. Rev. Cancer* **2007**, *7* (12), 961–967. https://doi.org/10.1038/nrc2254.

(478) Pan, L.; Li, Y.; Jia, L.; Qin, Y.; Qi, G.; Cheng, J.; Qi, Y.; Li, H.; Du, J. Cathepsin S Deficiency Results in Abnormal Accumulation of Autophagosomes in Macrophages and Enhances Ang II–Induced Cardiac Inflammation. *PLoS One* **2012**, *7* (4), e35315. https://doi.org/10.1371/journal.pone.0035315.

(479) Yang, M.; Liu, J.; Shao, J.; Qin, Y.; Ji, Q.; Zhang, X.; Du, J. Cathepsin S-Mediated Autophagic Flux in Tumor-Associated Macrophages Accelerate Tumor Development by Promoting M2

Polarization. *Mol. Cancer* **2014**, *13* (1), 43. https://doi.org/10.1186/1476-4598-13-43.

(480) Liu, V. C.; Wong, L. Y.; Jang, T.; Shah, A. H.; Park, I.; Yang, X.; Zhang, Q.; Lonning, S.; Teicher, B. A.; Lee, C. Tumor Evasion of the Immune System by Converting CD4+CD25− T Cells into CD4+CD25+ T Regulatory Cells: Role of Tumor-Derived TGF-β. *J. Immunol.* **2007**, *178* (5), 2883–2892. https://doi.org/10.4049/jimmunol.178.5.2883.

(481) Yan, D.; Wang, H.-W.; Bowman, R. L.; Joyce, J. A. STAT3 and STAT6 Signaling Pathways Synergize to Promote Cathepsin Secretion from Macrophages via IRE1α Activation. *Cell Rep.* **2016**, *16* (11), 2914–2927. https://doi.org/10.1016/j.celrep.2016.08.035.

(482) Biniossek, M. L.; Nägler, D. K.; Becker-Pauly, C.; Schilling, O. Proteomic Identification of Protease Cleavage Sites Characterizes Prime and Non-Prime Specificity of Cysteine Cathepsins B, L, and S. *J. Proteome Res.* **2011**, *10* (12), 5363–5373. https://doi.org/10.1021/pr200621z.

(483) Vasiljeva, O.; Dolinar, M.; Pungerčar, J. R.; Turk, V.; Turk, B. Recombinant Human Procathepsin S Is Capable of Autocatalytic Processing at Neutral PH in the Presence of Glycosaminoglycans. *FEBS Lett.* **2005**, *579* (5), 1285–1290. https://doi.org/10.1016/j.febslet.2004.12.093.

(484) Turk, B.; Dolenc, I.; Turk, V.; Bieth, J. G. Kinetics of the PH-Induced Inactivation of Human Cathepsin L. *Biochemistry* **1993**, *32* (1), 375–380.

(485) Xin, X.-Q.; Gunesekera, B.; Mason, R. W. The Specificity and Elastinolytic Activities of Bovine Cathepsins S and H. *Arch. Biochem. Biophys.* **1992**, *299* (2), 334–339. https://doi.org/10.1016/0003-9861(92)90283-3.

(486) Wolters, P. J.; Raymond, W. W.; Blount, J. L.; Caughey, G. H. Regulated Expression, Processing, and Secretion of Dog Mast Cell Dipeptidyl Peptidase I. *J. Biol. Chem.* **1998**, *273* (25), 15514–15520. https://doi.org/10.1074/jbc.273.25.15514.

(487) Sajid, M.; McKerrow, J. H. Cysteine Proteases of Parasitic Organisms. *Mol. Biochem. Parasitol.* **2002**, *120* (1), 1–21. https://doi.org/10.1016/S0166-6851(01)00438-8.

(488) Justus, C. R.; Dong, L.; Yang, L. V. Acidic Tumor Microenvironment and PH-Sensing G Protein-Coupled Receptors. *Front. Physiol.* **2013**, *4*. https://doi.org/10.3389/fphys.2013.00354.

(489) Gatenby, R. A.; Gillies, R. J. Why Do Cancers Have High Aerobic Glycolysis? *Nat. Rev. Cancer* **2004**, *4* (11), 891–899. https://doi.org/10.1038/nrc1478.

(490) Warburg, O.; Wind, F.; Negelein, E. The Metabolism of Tumors in the Body. *J. Gen. Physiol.* **1927**, *8* (6), 519–530. https://doi.org/10.1085/jgp.8.6.519.

(491) de la Cruz-López, K. G.; Castro-Muñoz, L. J.; Reyes-Hernández, D. O.; García-Carrancá, A.; Manzo-Merino, J. Lactate in the Regulation of Tumor Microenvironment and Therapeutic Approaches. *Front. Oncol.* **2019**, *9*. https://doi.org/10.3389/fonc.2019.01143.

(492) Sobotič, B.; Vizovišek, M.; Vidmar, R.; Van Damme, P.; Gocheva, V.; Joyce, J. A.; Gevaert, K.; Turk, V.; Turk, B.; Fonović, M. Proteomic Identification of Cysteine Cathepsin Substrates Shed from the Surface of Cancer Cells. *Mol. Cell. Proteomics* **2015**, *14* (8), 2213–2228. https://doi.org/10.1074/mcp.M114.044628.

(493) Gocheva, V.; Zeng, W.; Ke, D.; Klimstra, D.; Reinheckel, T.; Peters, C.; Hanahan, D.; Joyce, J. A. Distinct Roles for Cysteine Cathepsin Genes in Multistage Tumorigenesis. *Genes Dev.* **2006**, *20* (5), 543–556. https://doi.org/10.1101/gad.1407406.

(494) Wang, B.; Sun, J.; Kitamoto, S.; Yang, M.; Grubb, A.; Chapman, H. A.; Kalluri, R.; Shi, G.-P. Cathepsin S Controls Angiogenesis and Tumor Growth via Matrix-Derived Angiogenic Factors. *J. Biol. Chem.* **2006**, *281* (9), 6020–6029. https://doi.org/10.1074/jbc.M509134200.

(495) Veillard, F.; Saidi, A.; Burden, R. E.; Scott, C. J.; Gillet, L.; Lecaille, F.; Lalmanach, G. Cysteine Cathepsins S and L Modulate Anti-Angiogenic Activities of Human Endostatin. *J. Biol. Chem.*

2011, *286* (43), 37158–37167. https://doi.org/10.1074/jbc.M111.284869.

(496) Taggart, C. C.; Lowe, G. J.; Greene, C. M.; Mulgrew, A. T.; O'Neill, S. J.; Levine, R. L.; McElvaney, N. G. Cathepsin B, L, and S Cleave and Inactivate Secretory Leucoprotease Inhibitor. *J. Biol. Chem.* **2001**, *276* (36), 33345–33352. https://doi.org/10.1074/jbc.M103220200.

(497) Zhang, L.; Wang, H.; Xu, J.; Zhu, J.; Ding, K. Inhibition of Cathepsin S Induces Autophagy and Apoptosis in Human Glioblastoma Cell Lines through ROS-Mediated PI3K/AKT/MTOR/P70S6K and JNK Signaling Pathways. *Toxicol. Lett.* **2014**, *228* (3), 248–259. https://doi.org/10.1016/j.toxlet.2014.05.015.

(498) Wei, L.; Shao, N.; Peng, Y.; Zhou, P. Inhibition of Cathepsin S Restores TGF-β-Induced Epithelial-to-Mesenchymal Transition and Tight Junction Turnover in Glioblastoma Cells. *J. Cancer* **2021**, *12* (6), 1592–1603. https://doi.org/10.7150/jca.50631.

(499) Seo, S. U.; Woo, S. M.; Min, K.; Kwon, T. K. Z-FL-COCHO, a Cathepsin S Inhibitor, Enhances Oxaliplatin-Induced Apoptosis through Upregulation of Bim Expression. *Biochem. Biophys. Res. Commun.* **2018**, *498* (4), 849–854. https://doi.org/10.1016/j.bbrc.2018.03.068.

(500) Elie, B. T.; Gocheva, V.; Shree, T.; Dalrymple, S. A.; Holsinger, L. J.; Joyce, J. A. Identification and Pre-Clinical Testing of a Reversible Cathepsin Protease Inhibitor Reveals Anti-Tumor Efficacy in a Pancreatic Cancer Model. *Biochimie* **2010**, *92* (11), 1618–1624. https://doi.org/10.1016/j.biochi.2010.04.023.

(501) Link, J. O.; Mossman, C. J.; Woo, S. H.; Zipfel, S. M. US 7,547,701 B2 - Haloalkyl Containing Compounds as Cysteine Protease Inhibitors, 2009.

(502) Izaguirre, G. The Proteolytic Regulation of Virus Cell Entry by Furin and Other Proprotein Convertases. *Viruses* **2019**, *11* (9), 837. https://doi.org/10.3390/v11090837.

(503) Robinson, E. L.; Alkass, K.; Bergmann, O.; Maguire, J. J.; Roderick, H. L.; Davenport, A. P. Genes Encoding ACE2, TMPRSS2 and Related Proteins Mediating SARS-CoV-2 Viral Entry Are Upregulated with Age in Human Cardiomyocytes. *J. Mol. Cell. Cardiol.* **2020**, *147*, 88–91. https://doi.org/10.1016/j.yjmcc.2020.08.009.

(504) McKerrow, J. H. Development of Cysteine Protease Inhibitors as Chemotherapy for Parasitic Diseases: Insights on Safety, Target Validation, and Mechanism of Action. *Int. J. Parasitol.* **1999**, *29* (6), 833–837. https://doi.org/10.1016/S0020-7519(99)00044-2.

(505) Stevens, J. R.; Noyes, H. A.; Dover, G. A.; Gibson, W. C. The Ancient and Divergent Origins of the Human Pathogenic Trypanosomes, Trypanosoma Brucei and T. Cruzi. *Parasitology* **1999**, *118* (1), 107–116. https://doi.org/10.1017/S0031182098003473.

(506) VEuPathDB. *TriTryp Database - Trypanosoma*. https://tritrypdb.org/tritrypdb/app/search?q=Trypanosoma&documentType=organism.

(507) Haag, J. The Molecular Phylogeny of Trypanosomes: Evidence for an Early Divergence of the Salivaria. *Mol. Biochem. Parasitol.* **1998**, *91* (1), 37–49. https://doi.org/10.1016/S0166-6851(97)00185-0.

(508) Vickerman, K. Developmental Cycles and Biology of Pathogenic Trypanosomes. *Br. Med. Bull.* **1985**, *41* (2), 105–114. https://doi.org/10.1093/oxfordjournals.bmb.a072036.

(509) Schuster, S.; Lisack, J.; Subota, I.; Zimmermann, H.; Reuter, C.; Mueller, T.; Morriswood, B.; Engstler, M. Unexpected Plasticity in the Life Cycle of Trypanosoma Brucei. *Elife* **2021**, *10*. https://doi.org/10.7554/eLife.66028.

(510) Zuma, A. A.; de Souza, W. Fexinidazole Interferes with the Growth and Structural Organization of Trypanosoma Cruzi. *Sci. Rep.* **2022**, *12* (1), 20388. https://doi.org/10.1038/s41598-022-23941-z.

(511) Pinazo, M.-J.; Forsyth, C.; Losada, I.; Esteban, E. T.; García-Rodríguez, M.; Villegas, M. L.;

Molina, I.; Crespillo-Andújar, C.; Gállego, M.; Ballart, C.; Ramirez, J. C.; Aden, T.; Hoerauf, A.; Pfarr, K.; Vaillant, M.; Marques, T.; Fernandes, J.; Blum, B.; Ribeiro, I.; Sosa-Estani, S.; Barreira, F.; Gascón, J. Efficacy and Safety of Fexinidazole for Treatment of Chronic Indeterminate Chagas Disease (FEXI-12): A Multicentre, Randomised, Double-Blind, Phase 2 Trial. *Lancet Infect. Dis.* **2024**, *24* (4), 395–403. https://doi.org/10.1016/S1473-3099(23)00651-5.

(512) Büscher, P.; Cecchi, G.; Jamonneau, V.; Priotto, G. Human African Trypanosomiasis. *Lancet* **2017**, *390* (10110), 2397–2409. https://doi.org/10.1016/S0140-6736(17)31510-6.

(513) Balmer, O.; Beadell, J. S.; Gibson, W.; Caccone, A. Phylogeography and Taxonomy of Trypanosoma Brucei. *PLoS Negl. Trop. Dis.* **2011**, *5* (2), e961. https://doi.org/10.1371/journal.pntd.0000961.

(514) Noireau, F.; Diosque, P.; Jansen, A. M. Trypanosoma Cruzi : Adaptation to Its Vectors and Its Hosts. *Vet. Res.* **2009**, *40* (2), 26. https://doi.org/10.1051/vetres/2009009.

(515) Steverding, D. The History of African Trypanosomiasis. *Parasit. Vectors* **2008**, *1* (1), 3. https://doi.org/10.1186/1756-3305-1-3.

(516) Baker, J. R. The Subspecific Taxonomy of Trypanosoma Brucei. *Parasite* **1995**, *2* (1), 3–12. https://doi.org/10.1051/parasite/1995021003.

(517) WHO Control of Neglected Tropical Diseases (NTD). *Ending the Neglect to Attain the Sustainable Development Goals: A Road Map for Neglected Tropical Diseases 2021–2030*; Ntuli, M. M., Ed.; 2021.

(518) Franco, J. R.; Priotto, G.; Paone, M.; Cecchi, G.; Ebeja, A. K.; Simarro, P. P.; Sankara, D.; Metwally, S. B. A.; Argaw, D. D. The Elimination of Human African Trypanosomiasis: Monitoring Progress towards the 2021–2030 WHO Road Map Targets. *PLoS Negl. Trop. Dis.* **2024**, *18* (4), e0012111. https://doi.org/10.1371/journal.pntd.0012111.

(519) Sutherland, C. S.; Stone, C. M.; Steinmann, P.; Tanner, M.; Tediosi, F. Seeing beyond 2020: An Economic Evaluation of Contemporary and Emerging Strategies for Elimination of Trypanosoma Brucei Gambiense. *Lancet Glob. Heal.* **2017**, *5* (1), e69–e79. https://doi.org/10.1016/S2214-109X(16)30237-6.

(520) Pays, E.; Radwanska, M.; Magez, S. The Pathogenesis of African Trypanosomiasis. *Annu. Rev. Pathol. Mech. Dis.* **2023**, *18* (1), 19–45. https://doi.org/10.1146/annurev-pathmechdis-031621-025153.

(521) Blum, J.; Schmid, C.; Burri, C. Clinical Aspects of 2541 Patients with Second Stage Human African Trypanosomiasis. *Acta Trop.* **2006**, *97* (1), 55–64. https://doi.org/10.1016/j.actatropica.2005.08.001.

(522) Kennedy, P. G. Clinical Features, Diagnosis, and Treatment of Human African Trypanosomiasis (Sleeping Sickness). *Lancet Neurol.* **2013**, *12* (2), 186–194. https://doi.org/10.1016/S1474-4422(12)70296-X.

(523) Jamonneau, V.; Ilboudo, H.; Kaboré, J.; Kaba, D.; Koffi, M.; Solano, P.; Garcia, A.; Courtin, D.; Laveissière, C.; Lingue, K.; Büscher, P.; Bucheton, B. Untreated Human Infections by Trypanosoma Brucei Gambiense Are Not 100% Fatal. *PLoS Negl. Trop. Dis.* **2012**, *6* (6), e1691. https://doi.org/10.1371/journal.pntd.0001691.

(524) Pays, E.; Vanhollebeke, B.; Uzureau, P.; Lecordier, L.; Pérez-Morga, D. The Molecular Arms Race between African Trypanosomes and Humans. *Nat. Rev. Microbiol.* **2014**, *12* (8), 575–584. https://doi.org/10.1038/nrmicro3298.

(525) Gibson, W. C. The SRA Gene: The Key to Understanding the Nature of Trypanosoma Brucei Rhodesiense. *Parasitology* **2005**, *131* (2), 143–150. https://doi.org/10.1017/S0031182005007560.

(526) Uzureau, P.; Uzureau, S.; Lecordier, L.; Fontaine, F.; Tebabi, P.; Homblé, F.; Grélard, A.; Zhendre, V.; Nolan, D. P.; Lins, L.; Crowet, J.-M.; Pays, A.; Felu, C.; Poelvoorde, P.; Vanhollebeke, B.; Moestrup, S. K.; Lyngsø, J.; Pedersen, J. S.; Mottram, J. C.; Dufourc, E. J.; Pérez-Morga, D.; Pays, E. Mechanism of Trypanosoma Brucei Gambiense Resistance to Human Serum. *Nature* **2013**, *501* (7467), 430–434. https://doi.org/10.1038/nature12516.

(527) Barrett, M. P.; Priotto, G.; Franco, J. R.; Lejon, V.; Lindner, A. K. Elimination of Human African Trypanosomiasis: The Long Last Mile. *PLoS Negl. Trop. Dis.* **2024**, *18* (5), e0012091. https://doi.org/10.1371/journal.pntd.0012091.

(528) Magang, E. M. K.; Kamga, R. M. N.; Telleria, J.; Tichit, M.; Crouzols, A.; Kaboré, J.; Hardy, D.; Bouaka, C. U. T.; Jamonneau, V.; Rotureau, B.; Kuete, V.; Bart, J.-M.; Simo, G. Prevalence of Blood and Skin Trypanosomes in Domestic and Wild Fauna from Two Sleeping Sickness Foci in Southern Cameroon. *PLoS Negl. Trop. Dis.* **2023**, *17* (7), e0011528. https://doi.org/10.1371/journal.pntd.0011528.

(529) Sistrom, M.; Evans, B.; Benoit, J.; Balmer, O.; Aksoy, S.; Caccone, A. De Novo Genome Assembly Shows Genome Wide Similarity between Trypanosoma Brucei Brucei and Trypanosoma Brucei Rhodesiense. *PLoS One* **2016**, *11* (2), e0147660. https://doi.org/10.1371/journal.pone.0147660.

(530) Abera, A.; Mamecha, T.; Abose, E.; Bokicho, B.; Ashole, A.; Bishaw, T.; Mariyo, A.; Bogale, B.; Terefe, H.; Tadese, H.; Belachew, M.; Difabachew, H.; Eukubay, A.; Kinde, S.; Ali, A.; Regasa, F.; Seife, F.; Kebede, Z.; Wossen, M.; Tollera, G.; Hailu, M.; Manaye, N.; van Reet, N.; Priotto, G.; van Griensven, J.; Pareyn, M.; Tasew, G. Reemergence of Human African Trypanosomiasis Caused by Trypanosoma Brucei Rhodesiense , Ethiopia. *Emerg. Infect. Dis.* **2024**, *30* (1). https://doi.org/10.3201/eid3001.231319.

(531) Silvester, E.; McWilliam, K.; Matthews, K. The Cytological Events and Molecular Control of Life Cycle Development of Trypanosoma Brucei in the Mammalian Bloodstream. *Pathogens* **2017**, *6* (3), 29. https://doi.org/10.3390/pathogens6030029.

(532) Vassella, E.; Reuner, B.; Yutzy, B.; Boshart, M. Differentiation of African Trypanosomes Is Controlled by a Density Sensing Mechanism Which Signals Cell Cycle Arrest via the CAMP Pathway. *J. Cell Sci.* **1997**, *110* (21), 2661–2671. https://doi.org/10.1242/jcs.110.21.2661.

(533) Rojas, F.; Silvester, E.; Young, J.; Milne, R.; Tettey, M.; Houston, D. R.; Walkinshaw, M. D.; Pérez-Pi, I.; Auer, M.; Denton, H.; Smith, T. K.; Thompson, J.; Matthews, K. R. Oligopeptide Signaling through TbGPR89 Drives Trypanosome Quorum Sensing. *Cell* **2019**, *176* (1–2), 306-317.e16. https://doi.org/10.1016/j.cell.2018.10.041.

(534) Mosser, D. M.; Roberts, J. F. Trypanosoma Brucei: Recognition in Vitro of Two Developmental Forms by Murine Macrophages. *Exp. Parasitol.* **1982**, *54* (3), 310–316. https://doi.org/10.1016/0014-4894(82)90040-6.

(535) Borst, P.; Cross, G. A. M. Molecular Basis for Trypanosome Antigenic Variation. *Cell* **1982**, *29* (2), 291–303. https://doi.org/10.1016/0092-8674(82)90146-5.

(536) Silva Pereira, S.; Jackson, A. P.; Figueiredo, L. M. Evolution of the Variant Surface Glycoprotein Family in African Trypanosomes. *Trends Parasitol.* **2022**, *38* (1), 23–36. https://doi.org/10.1016/j.pt.2021.07.012.

(537) Seed, J. R. The Role of Immunoglobulins in Immunity to Trypanosoma Brucei Gambiense. *Int. J. Parasitol.* **1977**, *7* (1), 55–60. https://doi.org/10.1016/0020-7519(77)90025-X.

(538) Horn, D. Antigenic Variation in African Trypanosomes. *Mol. Biochem. Parasitol.* **2014**, *195* (2), 123–129. https://doi.org/10.1016/j.molbiopara.2014.05.001.

(539) Bockstal, V.; Guirnalda, P.; Caljon, G.; Goenka, R.; Telfer, J. C.; Frenkel, D.; Radwanska, M.; Magez, S.; Black, S. J. T. Brucei Infection Reduces B Lymphopoiesis in Bone Marrow and Truncates Compensatory Splenic Lymphopoiesis through Transitional B-Cell Apoptosis. *PLoS*

Pathog. **2011**, *7* (6), e1002089. https://doi.org/10.1371/journal.ppat.1002089.

(540) Radwanska, M.; Guirnalda, P.; De Trez, C.; Ryffel, B.; Black, S.; Magez, S. Trypanosomiasis-Induced B Cell Apoptosis Results in Loss of Protective Anti-Parasite Antibody Responses and Abolishment of Vaccine-Induced Memory Responses. *PLoS Pathog.* **2008**, *4* (5), e1000078. https://doi.org/10.1371/journal.ppat.1000078.

(541) Lecordier, L.; Uzureau, S.; Vanwalleghem, G.; Deleu, M.; Crowet, J.-M.; Barry, P.; Moran, B.; Voorheis, P.; Dumitru, A.-C.; Yamaryo-Botté, Y.; Dieu, M.; Tebabi, P.; Vanhollebeke, B.; Lins, L.; Botté, C. Y.; Alsteens, D.; Dufrêne, Y.; Pérez-Morga, D.; Nolan, D. P.; Pays, E. The Trypanosoma Brucei KIFC1 Kinesin Ensures the Fast Antibody Clearance Required for Parasite Infectivity. *iScience* **2020**, *23* (9), 101476. https://doi.org/10.1016/j.isci.2020.101476.

(542) Guegan, F.; Figueiredo, L. A Two-Stage Solution. *Elife* **2021**, *10*. https://doi.org/10.7554/eLife.72980.

(543) Bachmaier, S.; Thanner, T.; Boshart, M. Culturing and Transfection of Pleomorphic Trypanosoma Brucei; 2020; pp 23–38. https://doi.org/10.1007/978-1-0716-0294-2_2.

(544) Wall, R. J.; Rico, E.; Lukac, I.; Zuccotto, F.; Elg, S.; Gilbert, I. H.; Freund, Y.; Alley, M. R. K.; Field, M. C.; Wyllie, S.; Horn, D. Clinical and Veterinary Trypanocidal Benzoxaboroles Target CPSF3. *Proc. Natl. Acad. Sci.* **2018**, *115* (38), 9616–9621. https://doi.org/10.1073/pnas.1807915115.

(545) Jennings, F. W.; Urquhart, G. M. The Use of the 2 Substituted 5-Nitroimidazole, Fexinidazole (Hoe 239) in the Treatment of Chronic T. Brucei Infections in Mice. *Zeitschrift für Parasitenkd. Parasitol. Res.* **1983**, *69* (5), 577–581. https://doi.org/10.1007/BF00926669.

(546) Brun, R.; Jenni, L.; Schönenberger, M.; Schell, K. -F. In Vitro Cultivation of Bloodstream Forms of Trypanosoma Brucei, T. Rhodesiense , and T. Gambiense 1. *J. Protozool.* **1981**, *28* (4), 470–479. https://doi.org/10.1111/j.1550-7408.1981.tb05322.x.

(547) Thuita, J. K.; Kagira, J. M.; Mwangangi, D.; Matovu, E.; Turner, C. M. R.; Masiga, D. Trypanosoma Brucei Rhodesiense Transmitted by a Single Tsetse Fly Bite in Vervet Monkeys as a Model of Human African Trypanosomiasis. *PLoS Negl. Trop. Dis.* **2008**, *2* (5), e238. https://doi.org/10.1371/journal.pntd.0000238.

(548) Torreele, E.; Bourdin Trunz, B.; Tweats, D.; Kaiser, M.; Brun, R.; Mazué, G.; Bray, M. A.; Pécoul, B. Fexinidazole – A New Oral Nitroimidazole Drug Candidate Entering Clinical Development for the Treatment of Sleeping Sickness. *PLoS Negl. Trop. Dis.* **2010**, *4* (12), e923. https://doi.org/10.1371/journal.pntd.0000923.

(549) Ziegelbauer, K.; Quinten, M.; Schwarz, H.; Pearson, T. W.; Overath, P. Synchronous Differentiation of Trypanosoma Brucei from Bloodstream to Procyclic Forms in Vitro. *Eur. J. Biochem.* **1990**, *192* (2), 373–378. https://doi.org/10.1111/j.1432-1033.1990.tb19237.x.

(550) Taleva, G.; Husová, M.; Panicucci, B.; Hierro-Yap, C.; Pineda, E.; Biran, M.; Moos, M.; Šimek, P.; Butter, F.; Bringaud, F.; Zíková, A. Mitochondrion of the Trypanosoma Brucei Long Slender Bloodstream Form Is Capable of ATP Production by Substrate-Level Phosphorylation. *PLOS Pathog.* **2023**, *19* (10), e1011699. https://doi.org/10.1371/journal.ppat.1011699.

(551) MacGregor, P.; Rojas, F.; Dean, S.; Matthews, K. R. Stable Transformation of Pleomorphic Bloodstream Form Trypanosoma Brucei. *Mol. Biochem. Parasitol.* **2013**, *190* (2), 60–62. https://doi.org/10.1016/j.molbiopara.2013.06.007.

(552) Nok, A. J. Arsenicals (Melarsoprol), Pentamidine and Suramin in the Treatment of Human African Trypanosomiasis. *Parasitol. Res.* **2003**, *90* (1), 71–79. https://doi.org/10.1007/s00436-002-0799-9.

(553) Steverding, D. The Development of Drugs for Treatment of Sleeping Sickness: A Historical Review. *Parasit. Vectors* **2010**, *3* (1), 15. https://doi.org/10.1186/1756-3305-3-15.

(554) Flynn, I. W.; Bowman, I. B. R. The Action of Trypanocidal Arsenical Drugs on Trypanosoma Brucei and Trypanosoma Rhodesiense. *Comp. Biochem. Physiol. Part B Comp. Biochem.* **1974**, *48* (2), 261–273. https://doi.org/10.1016/0305-0491(74)90203-X.

(555) McAdams, P. S. The Effect of 606 on the Eye, with the Report of Seven Cases of Serious Eye Complications Following Its Use. *Bost. Med. Surg. J.* **1914**, *170* (9), 308–312. https://doi.org/10.1056/NEJM191402261700903.

(556) Seixas, J.; Atouguia, J.; Josenando, T.; Vatunga, G.; Miaka Mia Bilenge, C.; Lutumba, P.; Burri, C. Clinical Study on the Melarsoprol-Related Encephalopathic Syndrome: Risk Factors and HLA Association. *Trop. Med. Infect. Dis.* **2020**, *5* (1), 5. https://doi.org/10.3390/tropicalmed5010005.

(557) WHO Control of Neglected Tropical Diseases (NTD). *Guidelines for the Treatment of Human African Trypanosomiasis*; 2024.

(558) Dressel, J. (Ralph E. O. The Discovery of Germanin by Oskar Dressel and Richard Kothe. *J. Chem. Educ.* **1961**, *38* (12), 620. https://doi.org/10.1021/ed038p620.

(559) Zoltner, M.; Campagnaro, G. D.; Taleva, G.; Burrell, A.; Cerone, M.; Leung, K.-F.; Achcar, F.; Horn, D.; Vaughan, S.; Gadelha, C.; Zíková, A.; Barrett, M. P.; de Koning, H. P.; Field, M. C. Suramin Exposure Alters Cellular Metabolism and Mitochondrial Energy Production in African Trypanosomes. *J. Biol. Chem.* **2020**, *295* (24), 8331–8347. https://doi.org/10.1074/jbc.RA120.012355.

(560) Von der Ahe, D.; Huehnchen, P.; Balkaya, M.; Peruzzaro, S.; Endres, M.; Boehmerle, W. Suramin-Induced Neurotoxicity: Preclinical Models and Neuroprotective Strategies. *Molecules* **2018**, *23* (2), 346. https://doi.org/10.3390/molecules23020346.

(561) Ashley, J. N.; Barber, H. J.; Ewins, A. J.; Newbery, G.; Self, A. D. H. A Chemotherapeutic Comparison of the Trypanocidal Action of Some Aromatic Diamidines. *J. Chem. Soc.* **1942**, 103. https://doi.org/10.1039/jr9420000103.

(562) Bacchi, C. J. Chemotherapy of Human African Trypanosomiasis. *Interdiscip. Perspect. Infect. Dis.* **2009**, *2009*, 1–5. https://doi.org/10.1155/2009/195040.

(563) Steck, E. A.; Kinnamon, K. E.; Davidson, D. E.; Duxbury, R. E.; Johnson, A. J.; Masters, R. E. Trypanosoma Rhodesiense: Evaluation of the Antitrypanosomal Action of 2,5-Bis(4-Guanylphenyl)Furan Dihydrochloride. *Exp. Parasitol.* **1982**, *53* (1), 133–144. https://doi.org/10.1016/0014-4894(82)90099-6.

(564) Pohlig, G.; Bernhard, S. C.; Blum, J.; Burri, C.; Mpanya, A.; Lubaki, J.-P. F.; Mpoto, A. M.; Mununga, B. F.; N'tombe, P. M.; Deo, G. K. M.; Mutantu, P. N.; Kuikumbi, F. M.; Mintwo, A. F.; Munungi, A. K.; Dala, A.; Macharia, S.; Bilenge, C. M. M.; Mesu, V. K. B. K.; Franco, J. R.; Dituvanga, N. D.; Tidwell, R. R.; Olson, C. A. Efficacy and Safety of Pafuramidine versus Pentamidine Maleate for Treatment of First Stage Sleeping Sickness in a Randomized, Comparator-Controlled, International Phase 3 Clinical Trial. *PLoS Negl. Trop. Dis.* **2016**, *10* (2), e0004363. https://doi.org/10.1371/journal.pntd.0004363.

(565) Carter, N. S.; Fairlamb, A. H. Arsenical-Resistant Trypanosomes Lack an Unusual Adenosine Transporter. *Nature* **1993**, *361* (6408), 173–176. https://doi.org/10.1038/361173a0.

(566) Alsford, S.; Eckert, S.; Baker, N.; Glover, L.; Sanchez-Flores, A.; Leung, K. F.; Turner, D. J.; Field, M. C.; Berriman, M.; Horn, D. High-Throughput Decoding of Antitrypanosomal Drug Efficacy and Resistance. *Nature* **2012**, *482* (7384), 232–236. https://doi.org/10.1038/nature10771.

(567) Song, J.; Baker, N.; Rothert, M.; Henke, B.; Jeacock, L.; Horn, D.; Beitz, E. Pentamidine Is Not a Permeant but a Nanomolar Inhibitor of the Trypanosoma Brucei Aquaglyceroporin-2. *PLOS Pathog.* **2016**, *12* (2), e1005436. https://doi.org/10.1371/journal.ppat.1005436.

(568) De Koning, H. P. Uptake of Pentamidine in Trypanosoma Brucei Brucei Is Mediated by Three

Distinct Transporters: Implications for Cross-Resistance with Arsenicals. *Mol. Pharmacol.* **2001**, *59* (3), 586–592. https://doi.org/10.1124/mol.59.3.586.

(569) Boykin, D. W.; Kumar, A.; Hall, J. E.; Bender, B. C.; Tidwell, R. R. Anti-Pneumocystis Activity of Bis-Amidoximes and Bis-o-Alkylamidoximes Prodrugs. *Bioorg. Med. Chem. Lett.* **1996**, *6* (24), 3017–3020. https://doi.org/10.1016/S0960-894X(96)00557-4.

(570) Weller, T.; Alig, L.; Beresini, M.; Blackburn, B.; Bunting, S.; Hadváry, P.; Hürzeler Müller, M.; Knopp, D.; Levet-Trafit, B.; Lipari, M. T.; Modi, N. B.; Müller, M.; Refino, C. J.; Schmitt, M.; Schönholzer, P.; Weiss, S.; Steiner, B. Orally Active Fibrinogen Receptor Antagonists. 2. Amidoximes as Prodrugs of Amidines. *J. Med. Chem.* **1996**, *39* (16), 3139–3147. https://doi.org/10.1021/jm9509298.

(571) Wenzler, T.; Boykin, D. W.; Ismail, M. A.; Hall, J. E.; Tidwell, R. R.; Brun, R. New Treatment Option for Second-Stage African Sleeping Sickness: In Vitro and In Vivo Efficacy of Aza Analogs of DB289. *Antimicrob. Agents Chemother.* **2009**, *53* (10), 4185–4192. https://doi.org/10.1128/AAC.00225-09.

(572) Harrill, A. H.; DeSmet, K. D.; Wolf, K. K.; Bridges, A. S.; Eaddy, J. S.; Kurtz, C. L.; Hall, J. E.; Paine, M. F.; Tidwell, R. R.; Watkins, P. B. A Mouse Diversity Panel Approach Reveals the Potential for Clinical Kidney Injury Due to DB289 Not Predicted by Classical Rodent Models. *Toxicol. Sci.* **2012**, *130* (2), 416–426. https://doi.org/10.1093/toxsci/kfs238.

(573) Boiani, M.; Piacenza, L.; Hernández, P.; Boiani, L.; Cerecetto, H.; González, M.; Denicola, A. Mode of Action of Nifurtimox and N-Oxide-Containing Heterocycles against Trypanosoma Cruzi: Is Oxidative Stress Involved? *Biochem. Pharmacol.* **2010**, *79* (12), 1736–1745. https://doi.org/10.1016/j.bcp.2010.02.009.

(574) Metcalf, B. W.; Bey, P.; Danzin, C.; Jung, M. J.; Casara, P.; Vevert, J. P. Catalytic Irreversible Inhibition of Mammalian Ornithine Decarboxylase (EC 4.1. 1.17) by Substrate and Product Analogs. *J. Am. Chem. Soc.* **1978**, *100* (8), 2551–2553.

(575) Iten, M.; Mett, H.; Evans, A.; Enyaru, J. C.; Brun, R.; Kaminsky, R. Alterations in Ornithine Decarboxylase Characteristics Account for Tolerance of Trypanosoma Brucei Rhodesiense to D,L-Alpha-Difluoromethylornithine. *Antimicrob. Agents Chemother.* **1997**, *41* (9), 1922–1925. https://doi.org/10.1128/AAC.41.9.1922.

(576) Baker, N.; Alsford, S.; Horn, D. Genome-Wide RNAi Screens in African Trypanosomes Identify the Nifurtimox Activator NTR and the Eflornithine Transporter AAT6. *Mol. Biochem. Parasitol.* **2011**, *176* (1), 55–57. https://doi.org/10.1016/j.molbiopara.2010.11.010.

(577) Pépin, J.; Milord, F.; Meurice, F.; Ethier, L.; Loko, L.; Mpia, B. High-Dose Nifurtimox for Arseno-Resistant Trypanosoma Brucei Gambiense Sleeping Sickness: An Open Trial in Central Zaire. *Trans. R. Soc. Trop. Med. Hyg.* **1992**, *86* (3), 254–256. https://doi.org/10.1016/0035-9203(92)90298-Q.

(578) Pepin, J.; Milord, F.; Mpia, B.; Meurice, F.; Ethier, L.; DeGroof, D.; Bruneel, H. An Open Clinical Trial of Nifurtimox for Arseno-Resistant Trypanosoma Brucei Gambiense Sleeping Sickness in Central Zaire. *Trans. R. Soc. Trop. Med. Hyg.* **1989**, *83* (4), 514–517. https://doi.org/10.1016/0035-9203(89)90270-8.

(579) Priotto, G.; Fogg, C.; Balasegaram, M.; Erphas, O.; Louga, A.; Checchi, F.; Ghabri, S.; Piola, P. Three Drug Combinations for Late-Stage Trypanosoma Brucei Gambiense Sleeping Sickness: A Randomized Clinical Trial in Uganda. *PLoS Clin. Trials* **2006**, *1* (8), e39. https://doi.org/10.1371/journal.pctr.0010039.

(580) Priotto, G.; Kasparian, S.; Ngouama, D.; Ghorashian, S.; Arnold, U.; Ghabri, S.; Karunakara, U. Nifurtimox-Eflornithine Combination Therapy for Second-Stage Trypanosoma Brucei Gambiense Sleeping Sickness: A Randomized Clinical Trial in Congo. *Clin. Infect. Dis.* **2007**, *45* (11), 1435–1442. https://doi.org/10.1086/522982.

(581) Checchi, F.; Piola, P.; Ayikoru, H.; Thomas, F.; Legros, D.; Priotto, G. Nifurtimox plus Eflornithine for Late-Stage Sleeping Sickness in Uganda: A Case Series. *PLoS Negl. Trop. Dis.* **2007**, *1* (2), e64. https://doi.org/10.1371/journal.pntd.0000064.

(582) Lindner, A. K.; Lejon, V.; Chappuis, F.; Seixas, J.; Kazumba, L.; Barrett, M. P.; Mwamba, E.; Erphas, O.; Akl, E. A.; Villanueva, G.; Bergman, H.; Simarro, P.; Kadima Ebeja, A.; Priotto, G.; Franco, J. R. New WHO Guidelines for Treatment of Gambiense Human African Trypanosomiasis Including Fexinidazole: Substantial Changes for Clinical Practice. *Lancet Infect. Dis.* **2020**, *20* (2), e38–e46. https://doi.org/10.1016/S1473-3099(19)30612-7.

(583) Tweats, D.; Bourdin Trunz, B.; Torreele, E. Genotoxicity Profile of Fexinidazole – a Drug Candidate in Clinical Development for Human African Trypanomiasis (Sleeping Sickness). *Mutagenesis* **2012**, *27* (5), 523–532. https://doi.org/10.1093/mutage/ges015.

(584) Ding, D.; Zhao, Y.; Meng, Q.; Xie, D.; Nare, B.; Chen, D.; Bacchi, C. J.; Yarlett, N.; Zhang, Y.-K.; Hernandez, V.; Xia, Y.; Freund, Y.; Abdulla, M.; Ang, K.-H.; Ratnam, J.; McKerrow, J. H.; Jacobs, R. T.; Zhou, H.; Plattner, J. J. Discovery of Novel Benzoxaborole-Based Potent Antitrypanosomal Agents. *ACS Med. Chem. Lett.* **2010**, *1* (4), 165–169. https://doi.org/10.1021/ml100013s.

(585) Nare, B.; Wring, S.; Bacchi, C.; Beaudet, B.; Bowling, T.; Brun, R.; Chen, D.; Ding, C.; Freund, Y.; Gaukel, E.; Hussain, A.; Jarnagin, K.; Jenks, M.; Kaiser, M.; Mercer, L.; Mejia, E.; Noe, A.; Orr, M.; Parham, R.; Plattner, J.; Randolph, R.; Rattendi, D.; Rewerts, C.; Sligar, J.; Yarlett, N.; Don, R.; Jacobs, R. Discovery of Novel Orally Bioavailable Oxaborole 6-Carboxamides That Demonstrate Cure in a Murine Model of Late-Stage Central Nervous System African Trypanosomiasis. *Antimicrob. Agents Chemother.* **2010**, *54* (10), 4379–4388. https://doi.org/10.1128/AAC.00498-10.

(586) Betu Kumeso, V. K.; Kalonji, W. M.; Rembry, S.; Valverde Mordt, O.; Ngolo Tete, D.; Prêtre, A.; Delhomme, S.; Ilunga Wa Kyhi, M.; Camara, M.; Catusse, J.; Schneitter, S.; Nusbaumer, M.; Mwamba Miaka, E.; Mahenzi Mbembo, H.; Makaya Mayawula, J.; Layba Camara, M.; Akwaso Massa, F.; Kaninda Badibabi, L.; Kasongo Bonama, A.; Kavunga Lukula, P.; Mutanda Kalonji, S.; Mariero Philemon, P.; Mokilifi Nganyonyi, R.; Embana Mankiara, H.; Asuka Akongo Nguba, A.; Kobo Muanza, V.; Mulenge Nasandhel, E.; Fifi Nzeza Bambuwu, A.; Scherrer, B.; Strub-Wourgaft, N.; Tarral, A. Efficacy and Safety of Acoziborole in Patients with Human African Trypanosomiasis Caused by Trypanosoma Brucei Gambiense: A Multicentre, Open-Label, Single-Arm, Phase 2/3 Trial. *Lancet Infect. Dis.* **2023**, *23* (4), 463–470. https://doi.org/10.1016/S1473-3099(22)00660-0.

(587) Steverding, D.; Caffrey, C. R. Should the Enzyme Name 'Rhodesain' Be Discontinued? *Mol. Biochem. Parasitol.* **2021**, *245*, 111395. https://doi.org/10.1016/j.molbiopara.2021.111395.

(588) Caffrey, C. R.; Steverding, D. Kinetoplastid Papain-like Cysteine Peptidases. *Mol. Biochem. Parasitol.* **2009**, *167* (1), 12–19. https://doi.org/10.1016/j.molbiopara.2009.04.009.

(589) Steverding, D.; Rushworth, S. A.; Florea, B. I.; Overkleeft, H. S. Trypanosoma Brucei: Inhibition of Cathepsin L Is Sufficient to Kill Bloodstream Forms. *Mol. Biochem. Parasitol.* **2020**, *235*, 111246. https://doi.org/10.1016/j.molbiopara.2019.111246.

(590) Abdulla, M.-H.; O'Brien, T.; Mackey, Z. B.; Sajid, M.; Grab, D. J.; McKerrow, J. H. RNA Interference of Trypanosoma Brucei Cathepsin B and L Affects Disease Progression in a Mouse Model. *PLoS Negl. Trop. Dis.* **2008**, *2* (9), e298. https://doi.org/10.1371/journal.pntd.0000298.

(591) Wirtz, E.; Leal, S.; Ochatt, C.; Cross, G. M. A Tightly Regulated Inducible Expression System for Conditional Gene Knock-Outs and Dominant-Negative Genetics in Trypanosoma Brucei. *Mol. Biochem. Parasitol.* **1999**, *99* (1), 89–101. https://doi.org/10.1016/S0166-6851(99)00002-X.

(592) Niemirowicz, G. T.; Cazzulo, J. J.; Álvarez, V. E.; Bouvier, L. A. Simplified Inducible System for Trypanosoma Brucei. *PLoS One* **2018**, *13* (10), e0205527.

https://doi.org/10.1371/journal.pone.0205527.

(593) Dauchy, F.-A.; Bonhivers, M.; Landrein, N.; Dacheux, D.; Courtois, P.; Lauruol, F.; Daulouède, S.; Vincendeau, P.; Robinson, D. R. Trypanosoma Brucei CYP51: Essentiality and Targeting Therapy in an Experimental Model. *PLoS Negl. Trop. Dis.* **2016**, *10* (11), e0005125. https://doi.org/10.1371/journal.pntd.0005125.

(594) Nikolskaia, O. V. Blood-Brain Barrier Traversal by African Trypanosomes Requires Calcium Signaling Induced by Parasite Cysteine Protease. *J. Clin. Invest.* **2006**, *116* (10), 2739–2747. https://doi.org/10.1172/JCI27798.

(595) Geurink, P. P.; van der Linden, W. A.; Mirabella, A. C.; Gallastegui, N.; de Bruin, G.; Blom, A. E. M.; Voges, M. J.; Mock, E. D.; Florea, B. I.; van der Marel, G. A.; Driessen, C.; van der Stelt, M.; Groll, M.; Overkleeft, H. S.; Kisselev, A. F. Incorporation of Non-Natural Amino Acids Improves Cell Permeability and Potency of Specific Inhibitors of Proteasome Trypsin-like Sites. *J. Med. Chem.* **2013**, *56* (3), 1262–1275. https://doi.org/10.1021/jm3016987.

(596) Yang, P.-Y.; Wang, M.; He, C. Y.; Yao, S. Q. Proteomic Profiling and Potential Cellular Target Identification of K11777, a Clinical Cysteine Protease Inhibitor, in Trypanosoma Brucei. *Chem. Commun.* **2012**, *48* (6), 835–837. https://doi.org/10.1039/C1CC16178D.

(597) Kerr, I. D.; Lee, J. H.; Farady, C. J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K. C.; Caffrey, C. R.; Legac, J.; Hansell, E.; McKerrow, J. H.; Craik, C. S.; Rosenthal, P. J.; Brinen, L. S. Vinyl Sulfones as Antiparasitic Agents and a Structural Basis for Drug Design. *J. Biol. Chem.* **2009**, *284* (38), 25697–25703. https://doi.org/10.1074/jbc.M109.014340.

(598) Barr, S. C.; Warner, K. L.; Kornreic, B. G.; Piscitelli, J.; Wolfe, A.; Benet, L.; McKerrow, J. H. A Cysteine Protease Inhibitor Protects Dogs from Cardiac Damage during Infection by Trypanosoma Cruzi. *Antimicrob. Agents Chemother.* **2005**, *49* (12), 5160–5161. https://doi.org/10.1128/AAC.49.12.5160-5161.2005.

(599) Abdulla, M.-H.; Lim, K.-C.; Sajid, M.; McKerrow, J. H.; Caffrey, C. R. Schistosomiasis Mansoni: Novel Chemotherapy Using a Cysteine Protease Inhibitor. *PLoS Med.* **2007**, *4* (1), e14. https://doi.org/10.1371/journal.pmed.0040014.

(600) Steverding, D. Evaluation of Trypanocidal Activity of Combinations of Anti-Sleeping Sickness Drugs with Cysteine Protease Inhibitors. *Exp. Parasitol.* **2015**, *151–152*, 28–33. https://doi.org/10.1016/j.exppara.2015.01.016.

(601) Chaparro, J. D.; Cheng, T.; Tran, U. P.; Andrade, R. M.; Brenner, S. B. T.; Hwang, G.; Cohn, S.; Hirata, K.; McKerrow, J. H.; Reed, S. L. Two Key Cathepsins, TgCPB and TgCPL, Are Targeted by the Vinyl Sulfone Inhibitor K11777 in in Vitro and in Vivo Models of Toxoplasmosis. *PLoS One* **2018**, *13* (3), e0193982. https://doi.org/10.1371/journal.pone.0193982.

(602) McKerrow, J. H. Update on Drug Development Targeting Parasite Cysteine Proteases. *PLoS Negl. Trop. Dis.* **2018**, *12* (8), e0005850. https://doi.org/10.1371/journal.pntd.0005850.

(603) Kerr, I. D.; Wu, P.; Marion-Tsukamaki, R.; Mackey, Z. B.; Brinen, L. S. Crystal Structures of TbCatB and Rhodesain, Potential Chemotherapeutic Targets and Major Cysteine Proteases of Trypanosoma Brucei. *PLoS Negl. Trop. Dis.* **2010**, *4* (6), e701. https://doi.org/10.1371/journal.pntd.0000701.

(604) Shainyan, B. A. A-Versus B-carbon Nucleophilic Attack in Vinylic Substitution. *J. Phys. Org. Chem.* **1993**, *6* (1), 59–63. https://doi.org/10.1002/poc.610060110.

(605) Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T. H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P. R.; Distler, U.; Tenzer, S.; Johé, P.; Hellmich, U. A.; Engels, B. Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138* (27), 8332–8335. https://doi.org/10.1021/jacs.6b03052.

(606) Johe, P.; Jung, S.; Endres, E.; Kersten, C.; Zimmer, C.; Ye, W.; Sönnichsen, C.; Hellmich, U. A.; Sotriffer, C.; Schirmeister, T.; Neuweiler, H. Warhead Reactivity Limits the Speed of Inhibition of the Cysteine Protease Rhodesain. *ACS Chem. Biol.* **2021**, *16* (4), 661–670. https://doi.org/10.1021/acschembio.0c00911.

(607) Caffrey, C. R.; Hansell, E.; Lucas, K. D.; Brinen, L. S.; Alvarez Hernandez, A.; Cheng, J.; Gwaltney, S. L.; Roush, W. R.; Stierhof, Y.-D.; Bogyo, M.; Steverding, D.; McKerrow, J. H. Active Site Mapping, Biochemical Properties and Subcellular Localization of Rhodesain, the Major Cysteine Protease of Trypanosoma Brucei Rhodesiense. *Mol. Biochem. Parasitol.* **2001**, *118* (1), 61–73. https://doi.org/10.1016/S0166-6851(01)00368-1.

(608) Owens, T. D.; Brameld, K. A.; Verner, E. J.; Ton, T.; Li, X.; Zhu, J.; Masjedizadeh, M. R.; Bradshaw, J. M.; Hill, R. J.; Tam, D.; Bisconte, A.; Kim, E. O.; Francesco, M.; Xing, Y.; Shu, J.; Karr, D.; LaStant, J.; Finkle, D.; Loewenstein, N.; Haberstock-Debic, H.; Taylor, M. J.; Nunn, P.; Langrish, C. L.; Goldstein, D. M. Discovery of Reversible Covalent Bruton's Tyrosine Kinase Inhibitors PRN473 and PRN1008 (Rilzabrutinib). *J. Med. Chem.* **2022**, *65* (7), 5300–5316. https://doi.org/10.1021/acs.jmedchem.1c01170.

(609) Pace, V.; Castoldi, L.; Pregnolato, M. α-Amino-α-Halomethylketones: Synthetic Methodologies and Pharmaceutical Applications as Serine and Cysteine Protease Inhibitors. *Mini Rev. Med. Chem.* **2013**, *13* (7), 988–996.

(610) Fernández-de-la-Pradilla, A.; Royo, S.; Schirmeister, T.; Barthels, F.; Świderek, K.; González, F. V.; Moliner, V. Impact of the Warhead of Dipeptidyl Keto Michael Acceptors on the Inhibition Mechanism of Cysteine Protease Cathepsin L. *ACS Catal.* **2023**, *13* (20), 13354–13368. https://doi.org/10.1021/acscatal.3c02748.

(611) Royo, S.; Rodríguez, S.; Schirmeister, T.; Kesselring, J.; Kaiser, M.; González, F. V. Dipeptidyl Enoates As Potent Rhodesain Inhibitors That Display a Dual Mode of Action. *ChemMedChem* **2015**, *10* (9), 1484–1487. https://doi.org/10.1002/cmdc.201500204.

(612) Arafet, K.; González, F. V.; Moliner, V. Elucidating the Dual Mode of Action of Dipeptidyl Enoates in the Inhibition of Rhodesain Cysteine Proteases. *Chem. – A Eur. J.* **2021**, *27* (39), 10142–10150. https://doi.org/10.1002/chem.202100892.

(613) 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.; Rodríguez, S.; Diemer, K.; Neumann, W.; Gonzàlez, 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.

(614) Westerik, J. O.; Wolfenden, R. Aldehydes as Inhibitors of Papain. *J. Biol. Chem.* **1972**, *247* (24), 8195–8197. https://doi.org/10.1016/S0021-9258(20)81828-6.

(615) Macintyre, A. C.; Cutler, D. J. The Potential Role of Lysosomes in Tissue Distribution of Weak Bases. *Biopharm. Drug Dispos.* **1988**, *9* (6), 513–526. https://doi.org/10.1002/bod.2510090602.

(616) Hamada, Y. Recent Progress in Prodrug Design Strategies Based on Generally Applicable Modifications. *Bioorg. Med. Chem. Lett.* **2017**, *27* (8), 1627–1632. https://doi.org/10.1016/j.bmcl.2017.02.075.

(617) Gustafsson, D.; Nyström, J.-E.; Carlsson, S.; Bredberg, U.; Eriksson, U.; Gyzander, E.; Elg, M.; Antonsson, T.; Hoffmann, K.-J.; Ungell, A.-L.; Sörensen, H.; Någård, S.; Abrahamsson, A.; Bylund, R. The Direct Thrombin Inhibitor Melagatran and Its Oral Prodrug H 376/95: Intestinal Absorption Properties, Biochemical and Pharmacodynamic Effects. *Thromb. Res.* **2001**, *101* (3), 171–181. https://doi.org/10.1016/S0049-3848(00)00399-6.

(618) Stürzebecher, J.; Vieweg, H.; Steinmetzer, T.; Schweinitz, A.; Stubbs, M. T.; Renatus, M.; Wikström, P. 3-Amidinophenylalanine-Based Inhibitors of Urokinase. *Bioorg. Med. Chem. Lett.*

1999, *9* (21), 3147–3152. https://doi.org/10.1016/S0960-894X(99)00541-7.

(619) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S. E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.; Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.; Singh, R. S. P.; Steppan, C. M.; Stewart, A. E.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L.; Yang, Q.; Zhu, Y. An Oral SARS-CoV-2 M pro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science (80-.).* **2021**, *374* (6575), 1586–1593. https://doi.org/10.1126/science.abl4784.

(620) Kneller, D. W.; Li, H.; Phillips, G.; Weiss, K. L.; Zhang, Q.; Arnould, M. A.; Jonsson, C. B.; Surendranathan, S.; Parvathareddy, J.; Blakeley, M. P.; Coates, L.; Louis, J. M.; Bonnesen, P. V.; Kovalevsky, A. Covalent Narlaprevir- and Boceprevir-Derived Hybrid Inhibitors of SARS-CoV-2 Main Protease. *Nat. Commun.* **2022**, *13* (1), 2268. https://doi.org/10.1038/s41467-022-29915-z.

(621) Schade, D.; Kotthaus, J.; Riebling, L.; Kotthaus, J.; Müller-Fielitz, H.; Raasch, W.; Hoffmann, A.; Schmidtke, M.; Clement, B. Zanamivir Amidoxime- and N-Hydroxyguanidine-Based Prodrug Approaches to Tackle Poor Oral Bioavailability. *J. Pharm. Sci.* **2015**, *104* (9), 3208–3219. https://doi.org/10.1002/jps.24508.

(622) Arafa, R. K.; Ismail, M. A.; Munde, M.; Wilson, W. D.; Wenzler, T.; Brun, R.; Boykin, D. W. Novel Linear Triaryl Guanidines, N-Substituted Guanidines and Potential Prodrugs as Antiprotozoal Agents. *Eur. J. Med. Chem.* **2008**, *43* (12), 2901–2908. https://doi.org/10.1016/j.ejmech.2008.02.008.

(623) Lam, P. Y. S.; Clark, C. G.; Li, R.; Pinto, D. J. P.; Orwat, M. J.; Galemmo, R. A.; Fevig, J. M.; Teleha, C. A.; Alexander, R. S.; Smallwood, A. M.; Rossi, K. A.; Wright, M. R.; Bai, S. A.; He, K.; Luettgen, J. M.; Wong, P. C.; Knabb, R. M.; Wexler, R. R. Structure-Based Design of Novel Guanidine/Benzamidine Mimics: Potent and Orally Bioavailable Factor Xa Inhibitors as Novel Anticoagulants. *J. Med. Chem.* **2003**, *46* (21), 4405–4418. https://doi.org/10.1021/jm020578e.

(624) Wang, W.; Wang, L.; Liang, Y.-S. Susceptibility or Resistance of Praziquantel in Human Schistosomiasis: A Review. *Parasitol. Res.* **2012**, *111* (5), 1871–1877. https://doi.org/10.1007/s00436-012-3151-z.

(625) Park, S.-K.; Gunaratne, G. S.; Chulkov, E. G.; Moehring, F.; McCusker, P.; Dosa, P. I.; Chan, J. D.; Stucky, C. L.; Marchant, J. S. The Anthelmintic Drug Praziquantel Activates a Schistosome Transient Receptor Potential Channel. *J. Biol. Chem.* **2019**, *294* (49), 18873–18880. https://doi.org/10.1074/jbc.AC119.011093.

(626) Park, S.-K.; Friedrich, L.; Yahya, N. A.; Rohr, C. M.; Chulkov, E. G.; Maillard, D.; Rippmann, F.; Spangenberg, T.; Marchant, J. S. Mechanism of Praziquantel Action at a Parasitic Flatworm Ion Channel. *Sci. Transl. Med.* **2021**, *13* (625). https://doi.org/10.1126/scitranslmed.abj5832.

(627) Hockley, D. J.; McLaren, D. J. Schistosoma Mansoni: Changes in the Outer Membrane of the Tegument during Development from Cercaria to Adult Worm. *Int. J. Parasitol.* **1973**, *3* (1), 13–20. https://doi.org/10.1016/0020-7519(73)90004-0.

(628) Hockley, D. J.; McLaren, D. J.; Ward, B. J.; Nermut, M. V. A Freeze-Fracture Study of the Tegumental Membrane of Schistosoma Mansoni (Platyhelminthes:Trematoda). *Tissue Cell* **1975**, *7* (3), 485–496. https://doi.org/10.1016/0040-8166(75)90020-8.

(629) McLaren, D. J.; Hockley, D. J. Blood Flukes Have a Double Outer Membrane. *Nature* **1977**, *269* (5624), 147–149. https://doi.org/10.1038/269147a0.

(630) Matsumura, Y.; Maeda, H. A New Concept for Macromolecular Therapeutics in Cancer Chemotherapy: Mechanism of Tumoritropic Accumulation of Proteins and the Antitumor Agent Smancs. *Cancer Res.* **1986**, *46* (12), 6387–6392.

(631) Kennedy, P. G. E. Animal Models of Human African Trypanosomiasis—Very Useful or Too Far Removed? *Trans. R. Soc. Trop. Med. Hyg.* **2007**, *101* (11), 1061–1062. https://doi.org/10.1016/j.trstmh.2007.05.001.

(632) Nepali, K.; Lee, H.-Y.; Liou, J.-P. Nitro-Group-Containing Drugs. *J. Med. Chem.* **2019**, *62* (6), 2851–2893. https://doi.org/10.1021/acs.jmedchem.8b00147.

(633) 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*. https://doi.org/10.3389/fmolb.2021.804970.

(634) Hashimoto, H.; Takeuchi, T.; Komatsu, K.; Miyazaki, K.; Sato, M.; Higashi, S. Structural Basis for Matrix Metalloproteinase-2 (MMP-2)-Selective Inhibitory Action of β-Amyloid Precursor Protein-Derived Inhibitor. *J. Biol. Chem.* **2011**, *286* (38), 33236–33243. https://doi.org/10.1074/jbc.M111.264176.

Curriculum Vitae