JG|U

JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Development of covalent enzyme inhibitors:
RNA methyltransferases and proteases as targets

Dissertation

zur Erlangung des Grades

„DOKTOR DER NATURWISSENSCHAFTEN"

im Promotionsfach Chemie

am Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften der

Johannes Gutenberg-Universität Mainz

Marvin Schwickert

geboren am 12.09.1993 in Bad Kreuznach

Mainz, 2023

D77 (Dissertation Universität Mainz)

Dekanin: ███████████████

1. Gutachterin: ██████████████████

2. Gutachter: ████████████

Die vorliegende Dissertation wurde im Zeitraum von Januar 2019 bis März 2023 im Arbeitskreis von

██████████████████████ am Institut für Pharmazeutische und Biomedizinische Wissenschaften (IPBW) der Johannes Gutenberg-Universität Mainz angefertigt.

Hiermit versichere ich eidesstattlich:

1. Ich habe die vorliegende Dissertation selbstständig angefertigt und alle benutzten Hilfsmittel (Literatur, Apparaturen, Material) in der Arbeit angegeben.

2. Ich habe oder hatte die vorliegende Dissertation nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht.

3. Ich hatte weder die vorliegende Dissertation noch Teile davon bei einer anderen Fakultät bzw. einem anderen Fachbereich als Dissertation eingereicht.

_____ _____

Ort, Datum   Marvin Schwickert

"We must trust to nothing but facts: These are presented to us by nature and cannot deceive. We ought, in every instance, to submit our reasoning to the test of experiment, and never to search for truth but by the natural road of experiment and observation."

– Antoine Lavoisier

Danksagung

Danksagung

VIII

Zusammenfassung

Kovalente Inhibitoren wurden bei der Entwicklung von Arzneistoffen wegen möglichen toxischen Nebenwirkungen durch *Off-Target*-Modifikationen meist vermieden. Mittlerweile werden jedoch sie zunehmend in Betracht gezogen, nicht nur für Arzneistoffe, sondern auch für biochemische *Tools*. Gegenüber nicht-kovalenten Inhibitoren bieten sie nämlich mehrere Vorteile, wie höhere Affinität, effizientere Konkurrenz zu endogenen Substraten, längere Verweildauer und geringere Anfälligkeit für Arzneistoffresistenzen. In dieser Arbeit wurden potenzielle kovalente Inhibitoren für drei verschiedene Enzymklassen entwickelt und analysiert.

Zwar sind die RNA-Methyltransferasen (MTasen) DNMT2, NSUN2 und NSUN6 an verschiedenen physiologischen Prozessen beteiligt und werden u. a. mit Krebs in Verbindung gebracht, ihre biologischen Rollen sind jedoch weitestgehend unbekannt. Aus diesem Grund stellen sie vielversprechende Ziele für die Entwicklung von Medikamenten und *activity-based probes* (ABP) dar, die zu biologischen Untersuchungen dieser Enzyme sowie ihrer RNA-Modifikationen eingesetzt werden können. Ausgehend vom natürlichen Produktinhibitor *S*-Adenosyl-L-homocystein (SAH) wurden zusätzliche Seitenketten (alkylisch, aromatisch, *Warhead*-dekoriert) an das Gerüst angehängt, um das Affinitäts- und Selektivitätsprofil für DNMT2, NSUN2 und NSUN6 zu untersuchen. Die erste Generation von SAH-basierten Inhibitoren lieferte ein Alkin-Derivat als potenten DNMT2-Inhibitor (IC_{50} = 12.9 ± 1.9 μM). Nachfolgend wurde eine zweite Generation, bestehend aus elektronenarmen aromatischen SAH-Derivaten, entwickelt, was nicht nur zum bislang stärksten, sondern auch zum ersten kovalenten DNMT2-Inhibitor (IC_{50} = 1.2 ± 0.1 μM) führte, der selektiv gegenüber NSUN2 und NSUN6 ist. Diese Ergebnisse bilden eine vielversprechende Grundlage für die Entwicklung von ABPs für zukünftige Studien. Des Weiteren wurden selektive NSUN6-Inhibitoren identifiziert, die weiter hinsichtlich ihrer Affinität optimiert werden können.

Schistosoma mansoni ist ein parasitärer Plattwurm, der in 53 Ländern vorkommt und Schistosomiasis verursacht. Zahlreiche *in-house* Inhibitoren wurden auf ihre anti-schistosomale Aktivität und ihre hemmende Wirkung auf *S. mansoni* Cathepsin B1 (*Sm*CB1), die am häufigsten vorkommende Cysteinprotease im Darm des Parasiten, untersucht. Ausgehend von einem peptidomimetischen Inhibitor, der die höchste anti-schistosomale Aktivität aufwies, wurden lipophilere Substrukturen sowie alternative *Warheads* in das Gerüst eingeführt, was zu einer Erhöhung der Zellpermeabilität und Wirksamkeit gegen *S. mansoni* führte. Eine Korrelation mit *Sm*CB1-Inhibition konnte jedoch nicht festgestellt werden, da die im Fokus stehende Strukturklasse nur eine schwache Hemmung dieses Enzyms zeigte.

Sowohl die peptidomimetische Sequenz als auch der *Warhead* wirken sich auf Affinität und Selektivität von Proteaseinhibitoren aus, doch ist der genaue Einfluss jeder Komponente unklar. Daher wurde eine umfassende Studie durchgeführt, in der etablierte peptidomimetische Sequenzen von Inhibitoren für fünf Proteasen mit sieben verschiedenen *Warheads* kombiniert wurden. Die Testungen ergaben, dass die peptidomimetische Sequenz hauptsächlich für die Selektivität verantwortlich, jedoch die Kombination mit einem geeigneten *Warhead* entscheidend ist, um überhaupt eine Aktivität gegenüber einer Protease von Interesse zu erreichen und um eine *Off-Target*-Reaktivität zu vermeiden. Diese Arbeit beschreibt die Entwicklung von Bortezomib-Kongeneren als potenzielle Proteasom-Inhibitoren (einer von fünf Inhibitorensätzen) für diese Studie.

Abstract

Covalent inhibitors were usually avoided in drug development due to concerns of possible toxic side effects caused by off-target modification. Now they are increasingly emerging not only as therapeutic drugs but also as biochemical tools. Compared to non-covalent inhibitors, they have several advantages such as higher affinity, more efficient competition with endogenous substrates, prolonged residence times, and less susceptibility to drug resistance. The aim of this work was the development and analysis of potential covalent inhibitors for three different enzyme classes.

The RNA methyltransferases (MTases) DNMT2, NSUN2, and NSUN6 are involved in various physiological processes and have also been found to be associated to diseases such as cancer, yet their biological roles are largely unknown. Therefore, they represent promising targets for drug discovery and for the development of activity-based probes (ABP) that can be used to improve understanding of these enzymes and their RNA modifications. Based on the natural product inhibitor S-adenosyl-L-homocysteine (SAH), additional side chains (alkyl, aromatic, warhead-decorated) were attached to the scaffold to explore the affinity and selectivity profile for DNMT2, NSUN2, and NSUN6. The first generation of SAH-based inhibitors revealed an alkyne derivative as a potent inhibitor (IC_{50} = 12.9 ± 1.9 μM) of DNMT2. Subsequently, electron-deficient aromatic side chains were analyzed representing the second generation of SAH-based inhibitors, leading to not only the most potent but also the first covalent DNMT2 inhibitor (IC_{50} = 1.2 ± 0.1 μM) that is selective over NSUN2 and NSUN6. These findings represent a promising basis for the development of ABPs for future studies. Additionally, NSUN6 selective inhibitors were identified that can be further optimized to increase affinity.

Schistosoma mansoni is a parasitic flatworm occurring in 53 countries that causes schistosomiasis. An in-house inhibitor library was tested for anti-schistosomal activity and inhibitory effect on *S. mansoni* cathepsin B1 (*Sm*CB1), the most abundant cysteine protease in the parasite's gut. Based on a peptidomimetic inhibitor that exhibited the highest anti-schistosomal activity, more lipophilic substructures as well as alternative warheads were introduced into the scaffold, resulting in an increase in cell permeability and efficacy against *S. mansoni*. However, a correlation to *Sm*CB1 inhibition could not be observed since the structural class in focus showed only weak inhibition of this enzyme.

Both peptidomimetic sequence and warhead affect the affinity and selectivity of protease inhibitors, but the exact impact of each part is unclear. To analyze this, a comprehensive study was conducted combining established peptidomimetic sequences of inhibitors of five proteases with seven different types of warheads. Cross-testing revealed that the peptidomimetic sequence mainly causes selectivity, but the combination with a suitable warhead is crucial to achieve activity against a protease of interest in the first place as well as to avoid off-target reactivity. This work describes the development of bortezomib congeners as potential proteasome inhibitors (one out of five inhibitor sets) for this study.

Content

Abbreviations

$[\alpha]_D^{20}$	Specific rotation at 20 °C
1,2-DCE	1,2-Dichloroethane
2'-*O*-Me	RNA ribose 2'-*O*-methylation
5-FAM	5-Carboxyfluorescein
A	Adenine
ABP	Activity-based probe
ABPP	Activity-based protein profiling
Ad	Adenosine/adenosyl
ADP	Adenosine diphosphate
app	Apparent
ASA	Acetylsalicylic acid
Asn	Asparagine
Asp	Aspartate
ATP	Adenosine triphosphate
ATR	Attenuated total reflection
AUC	Area under the curve
C	Cytosine
CaCo	*Cancer coli* ("colon cancer"), immortalized cell line
CatS	Cathepsin S
CatL	Cathepsin L
CB1	Cathepsin B1
CDI	1,1'-Carbonyldiimidazole
CDK	Cyclin-dependent kinase
CFT	Cysteine-phenylalanine-threonine
COMT	Catechol-*O*-methyltransferase
COMU	[[(*Z*)-(1-cyano-2-ethoxy-2-oxoethylidene)amino]oxymorpholin-4-ylmethylidene]-dimethylazanium hexafluorophosphate
COX	Cyclooxygenase
CRISPR	Clustered regularly interspaced short palindromic repeats
Cryo EM	Cryogenic electron microscopy
Cys	Cysteine
d	Doublet
Dab	(*S*)-2,4-Diaminobutanoic acid
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCA	Dichloroacetic acid
DCC	*N,N'*-Dicyclohexylcarbodiimide
DCM	Dichloromethane
deg	Degree
DIAD	Diisopropyl azodicarboxylate
DIBAL	Diisobutylaluminium hydride
DIPEA	*N,N*-Diisopropylethylamine
DIPN	Diisopropylnaphthaline

Abbreviations

DKFZ	Deutsches Krebsforschungszentrum
DMA	Dimethylamine
DMF	Dimethylformamide
DMP	DESS-MARTIN periodinane
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA *N*-methyltransferase
DOT1L	Disruptor of telomeric silencing 1-like histone lysine methyltransferase
DPPA	Diphenylphosphoryl azide
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
E	Enzyme
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin ligase
Ecm1	*Encephalitozoon cuniculi* mRNA cap (guanine *N*-7) methyl-transferase
ENV	Glutamate-asparagine-valine
Eq.	Equation
equiv.	Equivalents
eRNA	Enhancer RNA
ESI	Electrospray ionization
f^{5}C	5-Formylcytosine
f^{5}Cm	5-Formyl-2'-*O*-methylcytidine
FTAD	5-FAM-triazolyl-adenosyl-Dab
FT-IR	Fourier-transform infrared spectroscopy
G	Guanine
Glu	Glutamate
h	Human
His	Histidine
hm^{5}C	5-Hydroxymethylcytosine
hm^{5}Cm	5-Hydroxymethyl-2'-*O*-methylcytidine
HMT	Histone methyltransferase
HOBt	Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
HTS	High-throughput-screening
HWE	Horner-Wadsworth-Emmons
I	Inhibitor (**Section 1**), Inosine (**Section 2**)
IR	Infrared
ITC	Isothermal titration calorimetry
J	Coupling constant
LC-MS	Liquid chromatography-mass spectrometry
LED	Light-emitting diode

XVI

lncRNA	Long non-coding RNA
Lys	Lysine
m	Multiplet
M.HhaI	*Haemophilus hemolyticus* DNA C5 methyltransferase
m/z	Mass-to-charge ratio
m^{1}A	N^1-Methyladenosine
m^{1}G	1-Methylguanosine
m^{2}G	2'-*O*-Methylguanosine
m^{5}C	5-Methylcytosine
m^{6}A	N^6-Methyladenosine
Me	Methyl
Met	Methionine
METTL	Methyltransferase like
miRNA	MicroRNA
MLL1	Histone methyltransferase mixed lineage leukemia 1
MOE	Molecular Operating Environment
mp	Melting point
M^{pro}	Main protease
mRNA	Messenger RNA
MS	Mass spectrometry
MST	Microscale thermophoresis
MST1	Ste20-like kinase
mt-RNA	Mitochondrial RNA
MTase	Methyltransferase
n-BuLi	*n*-Buthyllithium
NMR	Nuclear magnetic resonance
NSUN	NOL1/Nop2/Sun
NTS	Newly transformed schistosomula
Oxyma	Ethyl cyanohydroxyiminoacetate
p	*para*
PAINS	Pan assay interference compounds
PARP	Poly (ADP-ribose) polymerase
PCQ	Proline-cysteine-glutamine
PDB	Protein Data Bank
PMT	Protein methyltransferase
PZQ	Praziquantel
RlmJ	Ribosomal RNA large subunit methyltransferase J
RMP	RNA-modifying protein
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RNAi	RNA interference
rRNA	Ribosomal RNA
Rsm	16S rRNA C5-methylase
rt	Room temperature

s	Singlet
SAH	*S*-Adenosyl-L-homocysteine
SAM	*S*-Adenosyl-L-methionine
SAR	Structure-activity relationship
SARS-CoV 2	Severe acute respiratory syndrome coronavirus 2
ADME	Absorption, distribution, metabolism, and elimination
Ser	Serine
SFG	Sinefungin
Sm	*Schistosoma mansoni*
*Sm*AE	*Schistosoma mansoni* asparaginyl endopeptidase
sncRNA	Short non-coding RNA
snRNA	Small nuclear RNA
T	Thymine
t	Triplet
Tb	*Trypanosoma brucei*
TBAB	Tetrabutylammonium bromide
TBTU	2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate
TCA	Trichloroacetic acid
TCI	Targeted covalent inhibition
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Thr	Threonine
TPH	Tropical and Public Health Institute
TPSA	Topological polar surface area
TRD	Target recognition domain
TrmD	tRNA (guanine-*N*(1)-)-methyltransferase
tRNA	Transfer RNA
TRP	Transient receptor potential
TRPM	TRP melastatin
TSA	Thermal shift assay
U	Uracil
Ub	Ubiquitin
uPA	Urokinase-type plasminogen activator
UPS	Ubiquitin-proteasome-system
UV	Ultraviolet
vtRNA	Vault tRNA
WHO	World Health Organization
YAP	Yes-associated protein
Ψ	Pseudouridine

XVIII

Annotation

Cooperations

Results in this thesis were achieved in collaboration with the following working groups:

Part I – Joint project with ███ group

➤ ████████████, ██████████, ███████████ (IPBW Johannes Gutenberg University Mainz)

Tritium incorporation assay, LC-MS/MS-based quantification of m^5C.

➤ ████████████, ██████ (DKFZ, Heidelberg)

RNA-Bisulfite-MiSeq

➤ █████████████████, ██████████ (Institute of Pharmaceutical Technology, Goethe University Frankfurt)

CaCo-2 cell permeability assay

➤ ████████████████, ████████ (Technical University of Darmstadt)

Protein mass spectrometry

Part II

➤ ██████████████ (Swiss TPH, Basel)

Phenotypic drug-sensitivity assay in newly transformed schistosomula (NTS) and S. mansoni *adults*

Part III

➤ ███████████████████ (Departament de Química Inorgànica I Orgànica, Universitat Jaume I, Castelló de la Plana, Spain)

Synthesis of rhodesain inhibitors

➤ ███████████████, █████████ (Institute of Physical and Theoretical Chemistry, Julius-Maximilians-University of Würzburg)

QM calculations

Publications and manuscripts of the results obtained in this thesis:

M. Schwickert,[‡] S. N. Hoba,[‡] C. Kersten, A. Möllerke, M. Helm, and T. Schirmeister, *Parallel synthesis of adenosine derivatives for rapid molecule library generation*, **2023**, in preparation.

N. Fuchs,[‡] R. A. Zimmermann,[‡] **M. Schwickert**,[‡] A. Gunkel, C. Zimmer, M. Meta, J. Keiser, W. Kiefer, and T. Schirmeister, *Dual Optimization Strategies for New Agents Targeting Schistosoma mansoni: Advancing Phenotypic and SmCB1 Inhibitors for Improved Efficacy*, **2023**, in preparation.

P. Müller,[‡] M. Meta,[‡] L. Meidner,[‡] **M. Schwickert**, J. Meyer, K. Schwickert, C. Kersten, C. Zimmer, S. Hammerschmidt, A. Frey, A. Lahu, S. de la Hoz-Rodríguez, L. Agost-Beltrán, S. Rodríguez, K. Diemer, W. Neumann, F. V. Gonzàlez, B. Engels, and T. Schirmeister, *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.

M. Schwickert,[‡] R. A. Zimmermann,[‡] T. Habeck, S. N. Hoba, Z. Nidoieva, T. R. Fischer, M. M. Stark, C. Kersten, F. Lermyte, M. Helm and T. Schirmeister, *Covalent S-adenosylhomocysteine-based DNA methyltransferase 2 inhibitors with a new type of aryl warhead*, ACS Med. Chem. Lett. **2023**, in revision.

R. Zimmermann,[‡] **M. Schwickert**,[‡] L. Meidner,[‡] Z. Nidoieva, M. Helm, and T. Schirmeister, *An Optimized Microscale Thermophoresis Method for High-Throughput-Screening of DNA-methyltransferase 2 Ligands*, ACS Pharmacol. Transl. Sci. **2022**, *5*, 1079–1085

M. Schwickert,[‡] T. R. Fischer,[‡] R. A. Zimmermann,[‡] S. N. Hoba, L. Meidner, Ma. Weber, Mo. Weber, M. M. Stark, J. Koch, N. Jung, C. Kersten, M. Windbergs, F. Lyko, M. Helm, and T. Schirmeister, *Discovery of Inhibitors of DNA methyltransferase 2, an Epitranscriptomic Modulator and Potential Target for Cancer Treatment, J. Med. Chem.* **2022***, 65,* 9750–9788.

Additional Publications and manuscripts:

R. A. Zimmermann, T. R. Fischer, **M. Schwickert**, Z. Nidoieva, T. Schirmeister, and C. Kersten, *Chemical space docking against hard-to-drug RNA-methyltransferases DNMT2 and NSUN6*, Int. J. Mol. Sci. **2023**, *24*, 6109.

T. R. Fischer, L. Meidner, **M. Schwickert**, M. Weber, R. A. Zimmermann, C. Kersten, T. Schirmeister, and M. Helm, *Chemical Biology and Medicinal Chemistry of RNA methyltransferases*, Nucleic Acids Res. **2022**, *50*, 4216–4245.

F. Barthels, U. Barthels, **M. Schwickert**, and T. Schirmeister, *FINDUS: An Open-Source 3D Printable Liquid-Handling Workstation for Laboratory Automation in Life Sciences*, SLAS Technol. **2020**, *25*, 190–199.

[‡]These authors contributed equally.

1 Covalent Inhibitors – Introduction

1.1 Timeline of covalent inhibitors

Covalent inhibitors are small organic molecules designed to form a covalent bond with a molecular target of interest, e.g. an enzyme.[1] This process can be either reversible or irreversible, which is determined by the respective warhead.[2] A warhead is referred to as a reactive electrophilic group that allows a covalent reaction with a nucleophilic species of the target, typically an amino acid side chain.[1,3–5] Until recently, covalent inhibitors were usually avoided in drug development, mainly due to concerns of toxic side effects caused by off-target modification.[6] These concerns date back to the discovery of hepatotoxic properties of drugs such as bromobenzene or acetaminophen in the early 1970s. The compounds form highly reactive intermediates during metabolism which can covalently bind to liver proteins.[7] However, after being controversially discussed, covalent inhibitors are increasingly emerging, not only as therapeutic drugs, but also as biochemical tools.[7,8] Although covalent inhibitors have only recently attracted strong interest in medicinal chemistry, their course already started in the late 19th century with the development of acetylsalicylic acid (Aspirin, ASA) by Bayer (**Figure 1**).[1,9,10] ASA inhibits the cyclooxygenase I and II (COX-I and -II) by transferring the acetyl group to a serine residue near the active site.[11] In 1928, penicillin, a β-lactam-based antibiotic was discovered by Alexander Fleming.[12] The drug was later on synthesized by pharma and widely used during WWII in the 1940s. Further covalent drugs approved are 5-fluorouracil (1962), a chemotherapeutic agent,[13] omeprazole (1988), a proton pump inhibitor, fosfomycin (1996), an antibiotic, and the antiplatelet drug clopidogrel (1997).[1,4] Notably, omeprazole and clopidogrel each undergo metabolic modification to generate the active agents that form covalent bonds.[10] While the covalent binding mode was often unintended in early inhibitors, covalent inhibition was increasingly considered in drug design by fine-tuning the reactivity, e.g. by using computer-aided methods.[14–16] Warhead fine-tuning led to bortezomib (2003),[17] a proteasome inhibitor for the treatment of multiple myeloma.[18] Later, carfilzomib (2012) was approved as an alternative proteasome inhibitor for the same indication.[19] In 2009, saxagliptin, an antidiabetic, was introduced to the market,[20] followed by the antiviral drug boceprevir (2011).[21] Further recently approved drugs are dimethyl fumarate (2013) for the treatment of multiple sclerosis, voxelotor (2019) for the treatment of sickle cell disease, and the antiviral nirmatrelvir (2021).[1,4]

Nowadays, the concept of targeted covalent inhibition (TCI) has emerged in drug design. The term 'TCI' is sometimes defined inconsistently in literature, as it refers to either a compound designed to bind covalently to a specific target in general,[22] or to a compound that specifically targets non-catalytic, poorly conserved residues.[7] The latter one originates from the association with kinase inhibitors designed to target non-catalytic cysteine residues by α,β-unsaturated Michael acceptors.[23,24] Prominent examples are

ibrutinib (2013), osimertinib (2015), neratinib (2017/18) that have been approved for cancer treatment in the past decade.[1,4,9,25]
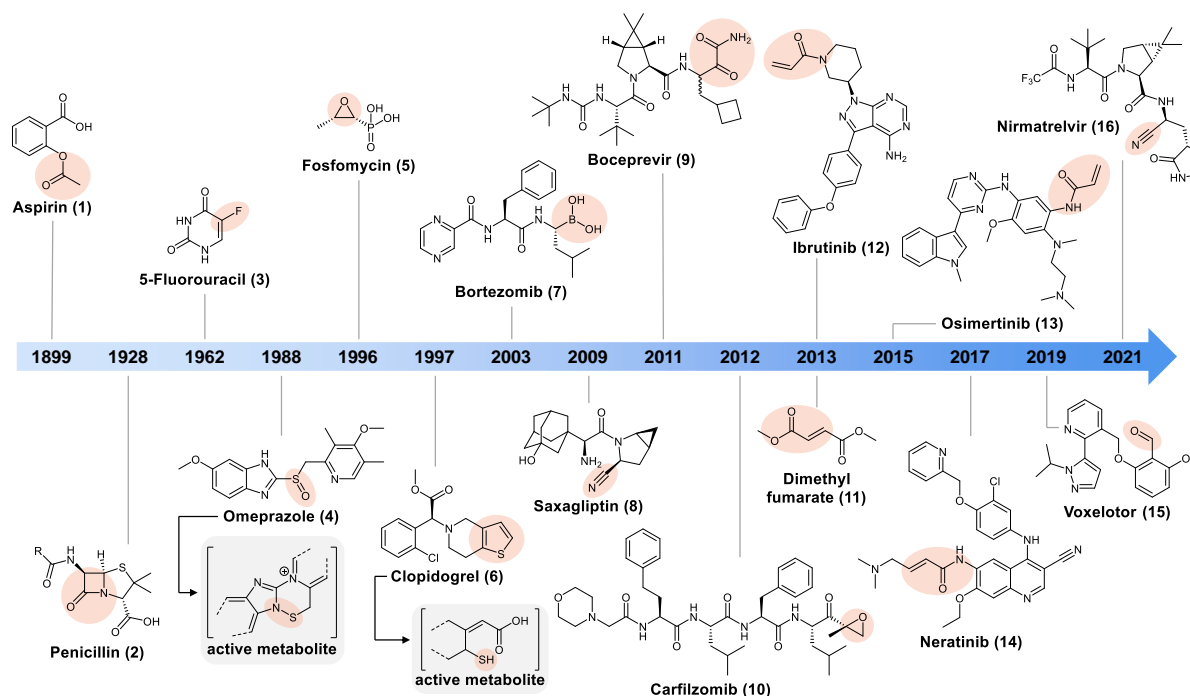
Figure 1: Timeline representation of approved covalent inhibitors in history. The warheads are highlighted in red.

1.2 Design of covalent inhibitors

Covalent drug discovery is a process that includes several steps until a suitable drug candidate is found. At the beginning, a valid biological target needs to be identified. If a hit is found for this target, the binding characteristics of the compound is further investigated and then finally optimized (**Figure 2**).[26] Nevertheless, the design of a covalent drug is a challenging endeavor. It requires detailed knowledge about the structural and mechanistic properties of the target.[26] However, this might not be the most challenging step as advances in cryogenic electron microscopy (cryo EM) and nuclear magnetic resonance (NMR) spectroscopy facilitate the structural elucidation of a biological target.[27,28] Once the information about the target of interest is available, hit finding can be achieved by strategies similar to non-covalent drug discovery: screening of large libraries, such as high-throughput-screening (HTS), virtual screening, and rational design.[26] However, the evaluation of possible hits is crucial, especially if compounds decorated with reactive groups are screened. True positive hits have to be distinguished from false positive Pan Assay Interference Compounds (PAINS), which must be discarded.[26] Yet various PAINS have been published as promising biological active compounds.[29,30] For this reason, substructural features frequently found in PAINS, i.e. "PAIN alerts" are reported in order to quickly identify false binders and discard them in screening libraries.[31] However, PAIN filters are not always reliable, making a small compound hit featuring such a substructure alert not necessarily a PAIN compound. Before a

potential covalent inhibitor would be discarded, additional validation by orthogonal methods should be performed.[32] Moreover, it is important to investigate that the binding does not rely exclusively on a non-specific reaction of the warhead with a nucleophilic residue in the active site.[26] Although designing substructures around a reactive group to build up non-covalent interactions can be a succeeding strategy,[33] the vast majority of designed covalent inhibitors started from a non-covalent hit or lead molecule.[10] By starting from a known non-covalent inhibitor, a derived covalent analog will show high selectivity and less reactive warheads can be used.[26]

Once a covalent inhibitor has been identified, subsequent optimization is usually required, for example, to improve non-covalent interactions and warhead reactivity. Common concepts to improve such structures are structure-activity relationship (SAR),[7] structure-based design,[3] and scaffold hopping,[34] but also bioinformatic approaches such as molecular docking,[35,36] molecular dynamics,[37] and quantum chemical calculations.[14,38] Optimization relates not only to affinity and selectivity, but also to the adjustment of pharmacokinetic parameters, such as absorption, distribution, metabolism, elimination (ADME),[39] which is particularly challenging in drug development.
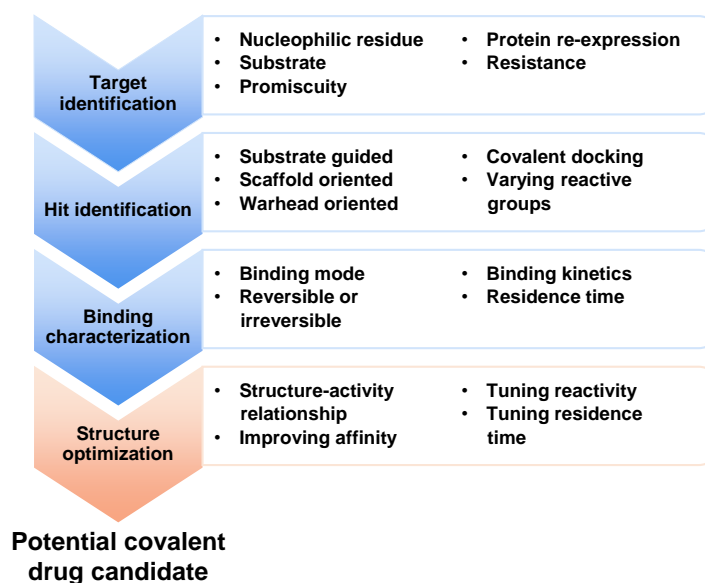
Figure 2: Process of developing a covalent drug.[26]

The warhead is not the only important part for a covalent inhibitor, moreover the combination between non-covalent and covalent interactions characterizes the activity profile. While the warhead ensures reactivity toward the target nucleophile, the non-covalent interactions are crucial for the orientation inside the binding site, bringing the reactive group in the correct position.[40] A reduction in toxicity can therefore be achieved by improving selectivity through optimization of non-covalent interactions and by fine-tuning the electrophilicity of the warhead.[41] Choosing the right warhead depends on several factors, especially the type of nucleophile to be targeted. To form a covalent bond, cysteine, serine, or threonine residues are commonly addressed as nucleophiles,[42] but compounds were also developed targeting

lysine,[43–45] tyrosine,[46] glutamate/aspartate,[23,47–49] and methionine side chains.[50] Among the non-activated residues, cysteine is the most nucleophilic, as shown by the relative nucleophilicity in the neutral states, which differs by several orders of magnitude: cysteine (1) > histidine (10^{-2}) > methionine > (10^{-3}) > lysine, serine (10^{-5}) > threonine and tyrosine (10^{-6}).[51] Given this fact, it is not surprising that most covalent inhibitors are targeting thiol groups.[24,52]

A huge variety of warheads is available to target nucleophilic residues – some of them react preferably with a specific type of nucleophile and some of them can target more than one species (**Figure 3**). While aldehydes and epoxides are unspecific and react with various nucleophiles (Cys, Ser/Thr, Lys; *cis*-configured epoxides also react with Asp[49]), (cyano-)acrylamides, propargyls, alkyl halides, or aziridines react preferably with cysteine.[23,48,53] Serine or threonine can be specifically trapped with β-lactams, boronic acids and fluorophosphonates.[48,53,54] α-Halo ketones, α-ketobenzothiazoles, or keto amides can be used to address cysteine as well as serine/threonine as they react less specifically.[26,53,55,56] These nucleophiles are also targeted by nitriles, which have been receiving much attention in covalent drug design, recently.[57,58] Besides the highly reactive aldehydes and epoxides, there are also a few other warheads that can be used to trap lysine, although most of them are not specific. Cyanamides and vinylsulfones (or -amides) are suitable for lysine but are mainly used to target cysteine.[53] Activated esters are typically used for covalent modification of lysine residues, but these warheads also allow reaction with serine and threonine.[23,53] *N*-Acyl-*N*-alkyl sulfonamides are reported to specifically address lysine, particularly to modify surface-exposed residues of proteins.[23,59] Methionine and glutamate/aspartate are rarely addressed residues. Oxaziridines are used to react highly selectively with methionine via a redox mechanism.[60,61] *N*-Methylisoxazolium structures derived from the "Woodward reagent K" can react with carboxylate groups found in glutamate and aspartate.[23,47,48] Besides the commonly used warheads mentioned in this section, various other reactive groups have been reported.[23,48,53] It is noteworthy that the specificity of all these structures, including those mentioned above, is not always consistently described in the literature.

The choice of the warhead determines not only which nucleophile is to be targeted, but also which binding mode will be performed. For example, the reaction of nitriles, aldehydes, boronic acids, and keto amides are reversible, while epoxides, aziridines, propargyl amides, vinylsulfones, and acryl amides are irreversible covalent modifiers.[23,48,53] β-Lactams are also described as irreversible inhibitors,[54,62] but in some cases a reversible binding mode with the release of an opened β-lactam ring can occur.[63] The binding behavior of some Michael-acceptor warheads can be adjusted by attaching an electron-withdrawing group to the α-position. Such modifications increase the acidity of the α-proton of the covalent adduct, thus facilitating elimination. This allows cyanoacrylamides and fluorovinylsulfones to react reversibly,[64–66] which is a huge advantage: The reactivity with exposed nucleophiles, such as glutathione is less of a

concern as the stability of the covalent complex (residence time) is governed by the stabilization of the non-covalent interactions.[64] Furthermore, the accessibility of the acidic α-proton to water is significantly higher when being solvent exposed instead of being bound in a binding site.[26]
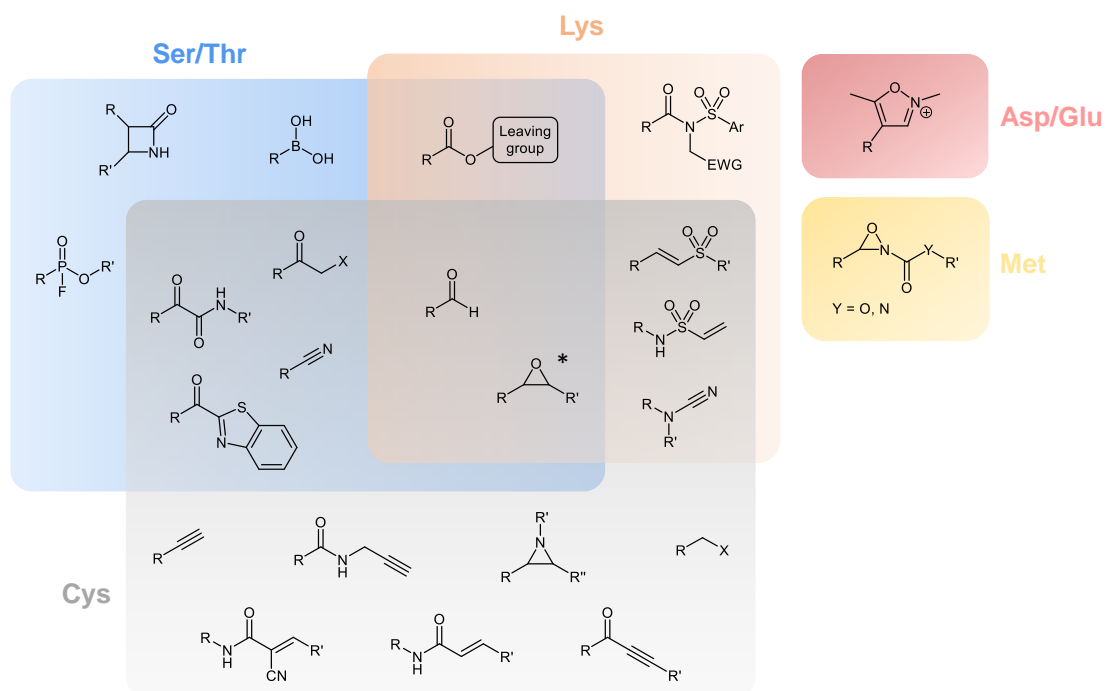
Figure 3: Commonly used warheads for targeting the nucleophiles Cys, Ser/Thr, Lys, Asp/Glu, and Met. *Cis*-configurated epoxides were also found to react with Asp.

1.3 Binding kinetics of covalent inhibitors

In drug development, the characterization of inhibitors is essential. Commonly, inhibitors are characterized by affinity or dissociation constants. Those parameters are usually determined *in vitro* using suitable biophysical assays for evaluation.[67] In the following sections, the different binding kinetics and corresponding parameters are discussed.

1.3.1 Non-covalent inhibitors

When a non-covalent inhibitor (I) binds to an enzyme, it forms an inhibitor-enzyme-complex (E••I), which results from non-covalent interactions between the ligand and the binding site of the target. This process is reversible, resulting in a global dissociation constant K_i defined as the ratio of the microscopic rate constants for the forward (k_1) and reverse (k_{-1}) binding event (**Scheme 1** and **Eq. 1**). Commonly, the on-rate k_{on} is equal to the forward rate constant k_1, while the off-rate k_{off} is equal to the reverse rate constant k_{-1}.[68]
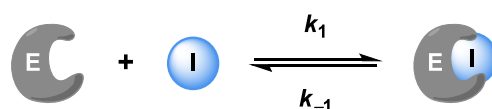
Scheme 1: Kinetic of non-covalent inhibitors.

$$K_i = \frac{k_{-1}}{k_1}$$ **Eq. 1**

1.3.2 Covalent reversible inhibitors

If a second reversible binding step occurs after the formation of a non-covalent complex illustrated in **Scheme 1**, the second equilibrium kinetics k_2/k_{-2} are also taken into account, giving a new global dissociation constant defined as K_i^* (**Scheme 2** and **Eq. 2**).[69] The formation of a covalent bond contributes to the overall affinity of the inhibitor, but subsequent bond cleavage can lower the affinity, which decreases the duration of inhibition.[8]
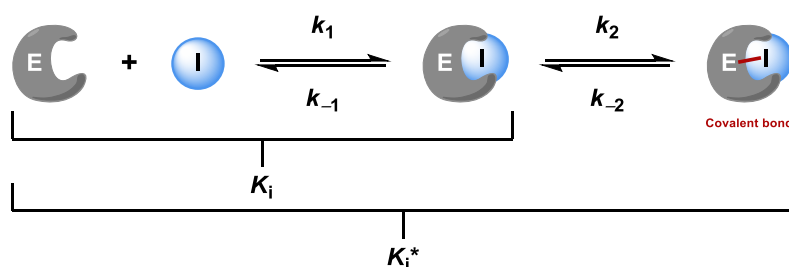
Scheme 2: Kinetic of covalent reversible inhibitors.

$$K_i^* = \frac{K_i}{\left(1 + \frac{k_2}{k_{-2}}\right)}$$ **Eq. 2**

1.3.3 Covalent irreversible inhibitors

If the covalent bond is irreversible, the constant k_{-2} of the reverse binding event is near or equal to zero. In this case, k_2 is referred to as k_{inact} or k_i (**Scheme 3**). To consider the formation and dissociation of the non-covalent complex as well as the total irreversible inhibition step, an additional constant, namely, the second order rate constant of inactivation k_{2nd} can be determined. It is defined as the quotient of k_{inact}/K_i and describes the efficiency of covalent bond formation (**Eq. 3**).[67,70]
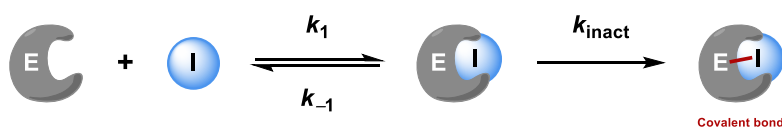
Scheme 3: Kinetic of covalent irreversible inhibitors.

$$k_{2nd} = \frac{k_{inact}}{K_i}$$ **Eq. 3**

1.3.4   Half maximal inhibitory concentration

In medicinal chemistry the biochemical potency of enzyme inhibitors is often expressed in terms of the half maximal inhibitory concentration IC_{50}. The IC_{50} value is related to the K_i value by the CHENG-PRUSOFF equation for competitive inhibitors (**Eq. 4**) where $[S]$ refers to the substrate concentration and K_m is the MICHAELIS constant representing the concentration of half maximal enzyme activity.[71]

$$K_i = \frac{IC_{50}}{\left(1 + \frac{[S]}{K_m}\right)}$$

Eq. 4

Notably, the IC_{50} value must be interpreted carefully regarding irreversible inhibitors. Their IC_{50} value decreases over time, until at asymptotically infinite time it approaches one half of the enzyme concentration. Therefore, the corresponding reaction time is essential if IC_{50} values are reported for a given irreversible inhibitor.[72]

1.3.5   Residence time

The residence time t_r corresponds to the average time a ligand remains bound to the binding site. It is defined as the reciprocal of the dissociation rate (**Eq. 5**).[73–75]

$$t_r = \frac{1}{k_{off}}$$

Eq. 5

The residence time is often considered as the key determinant of *in vivo* pharmacological activity and duration. Not primarily the binding affinity of a drug for its target might be crucial but rather the lifetime of the binary drug-target complex. This follows the assumption that pharmacological activity only persists as long as the drug remains bound.[76]

1.4   Advantages of covalent inhibitors

Although inhibitors with reactive groups were controversially discussed due to concerns of toxic side effects, the advantages of optimized covalent inhibitors now stand out.[6] Covalent inhibitors are highly potent compared to non-covalent analogs. The additional binding energy of a covalent bond allows covalent inhibitors to compete more efficiently with endogenous substrates, given that 80% of targets have an endogenous substrate that competes with the drug.[77,78] Moreover, they can compete with highly concentrated or high affinity natural ligands, which are especially found for kinases (adenosine triphosphate, ATP) or methyltransferases (*S*-adenosyl-L-methionine, SAM).[26,79,80] The same is true for hard-to-drug targets with shallow binding pockets that are difficult to reach through non-covalent interactions.[81,82] Covalent inhibitors provide prolonged target inactivation, and – depending on the warhead – irreversible target inactivation. This allows the administration of smaller drug doses for the patient.[6] With the application of irreversible inhibitors, drug resistance is less of a concern as effective

inhibition will be achieved regardless of any differences in the rate of formation induced by a binding site mutation.[83]

1.5   Development of covalent inhibitors in this work

In this work, three main projects about the development and analysis of potential covalent inhibitors are described:

Part I

S-Adenosyl-L-homocysteine (SAH)-based inhibitors of the DNA *N*-methyltransferase 2 (DNMT2) and representatives of the NOL1/NOP2/sun domain (NSUN) family, namely NSUN2 and NSUN6, all of which are RNA methyltransferases with catalytically active cysteines.

Part II

Covalent peptidomimetics against *Schistosoma mansoni*, a parasitic flatworm responsible for schistosomiasis. Anti-schistosomal compounds were also analyzed for their inhibitory effect on *S. mansoni* cathepsin B1 (*Sm*CB1), a cysteine protease, to identify potential correlations.

Part III

Bortezomib congeners with alternative warheads as potential proteasome inhibitors (threonine protease) for a comprehensive study examining the compatibility between peptidomimetic sequences and warheads and their impacts on affinity as well as selectivity.
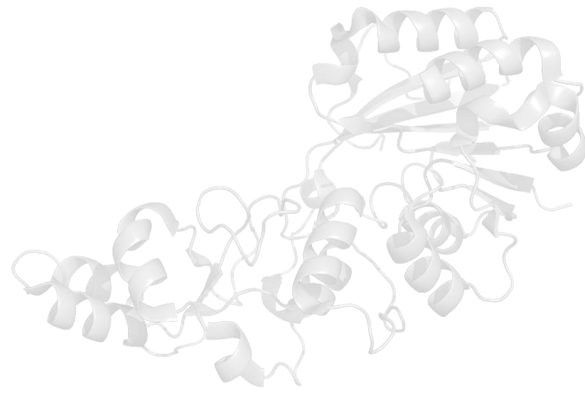
PART I

► **Non-covalent and covalent *S*-adenosyl-ʟ-homocysteine derivatives as RNA methyltransferase inhibitors**

2 Introduction

2.1 Epitranscriptomics

Epigenetics refers to a stably heritable phenotype that results from changes in a chromosome without altering the nucleotide sequence of deoxyribonucleic acid (DNA). Mechanisms that cause such changes include DNA methylation and histone modification.[84] However, not only DNA is affected by modifications. A variety of ribonucleic acid (RNA) modifications were increasingly discovered in the past decades, collectively referred to as the "*epitranscriptome*".[85] While the first RNA modification was discovered in 1957,[86] more than 170 different modifications of coding and non-coding RNA have been found to date, according to the *MODOMICS* database.[87] Chemical modifications occur in diverse RNA species, including messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), long non-coding RNA (lncRNA), and short non-coding RNA (sncRNA).[88–94] In most cases the modification takes place at the nucleobases adenine (A), cytosine (C), guanine (G), and uracil (U). Over 100 types were found, common examples include N^6-methyladenosine (m^{6}A), N^1-methyladenosine (m^{1}A), pseudouridine (Ψ), 5-methylcytosine (m^{5}C), 5-hydroxymethylcytosine (hm^{5}C), 1‑methylguanosine (m^{1}G), and inosine (I) (**Figure 4**).[91,95–97]
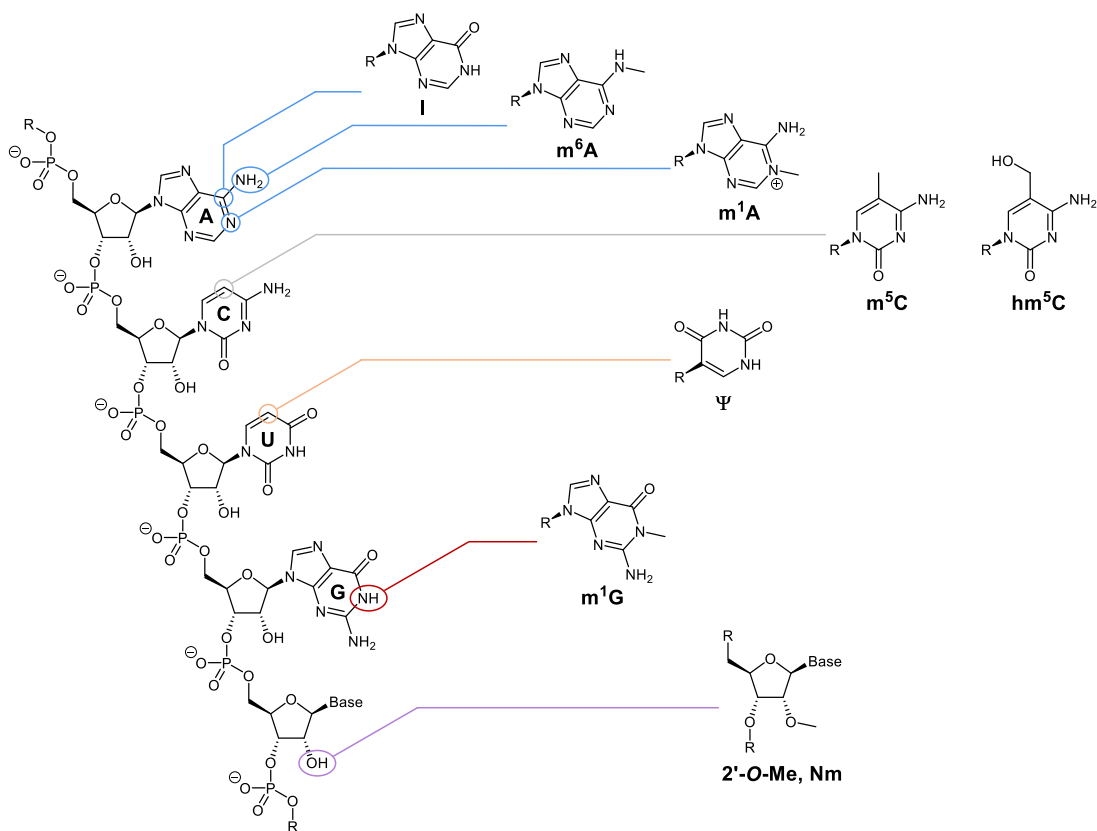
Figure 4: Common modifications found in RNA. A: adenine, C: cytosine, U: uracil, G: guanine, m^{6}A: N^6-methyladenosine, m^{1}A: N^1-methyladenosine, I: inosine, m^{5}C: 5-methylcytosine, hm^{5}C: 5-hydroxymethylcytosine, Ψ: pseudouridine, m^{1}G: 1‑methylguanosine, 2'-*O*-Me, Nm: 2'-*O*-methyl.

However, modifications of RNA molecules are not limited to the nucleobases. It was also found that the ribose moiety is often methylated at the 2'-O-position (2'-O-Me, Nm).[95] Like DNA and histone modifications, RNA modifications are proceeded by RNA-modifying proteins (RMPs) serving as "writers", "erasers", and "readers". Writers deposit RNA chemical marks, erasers remove them. Readers are able to recognize and bind to specific RNA modifications and fulfil different tasks.[98] Although RNA modifications have been found to affect functions in RNA metabolism, including RNA stability, splicing, processing, editing, structure, localization, translation initiation, and gene regulation,[98] the mechanisms and functions of most of them are still unknown.[98,99] Their presence in many species suggests evolutionarily conserved mechanisms which may be involved in the flow of genetic information or the response to environmental challenges.[99,100] Since RNA modifications are widely represented in metabolism, associations with pathological processes have also been found. Alteration of RNA modification patterns caused by dysregulation and mutations in RMPs are linked to various human diseases, including neurological diseases, cancer, genetic birth defects, obesity, and infertility.[101,102] Overall, epitranscriptomics is an emerging field, especially in medicinal chemistry, as it reveals various potential targets for treatment of diseases.[103]

2.2 tRNA cytidine methylation by SAM-dependent methyltransferases

2.2.1 m⁵C modification in tRNA

Among the broad spectrum of RNA modifications, methylation is the most common as it is present in all major RNA species and occurs on diverse positions, including m⁶A, m⁵C, and 2'-O-Me.[87] To date, more than 50 RNA methylation sites have been identified. tRNA molecules exhibit the greatest variety of RNA methylation, some of which are highly conserved in different kingdoms and others specific for a particular branch or group of species.[97]

The methylation of cytosine, in particular at position 5, is a widespread modification not only found in tRNA but also in many other RNA species.[97,104] While m⁵C only occurs in rRNA in bacteria, it is more common in archaea and eukaryotes as it is also present in mRNA and various ncRNAs. The m⁵C modification in DNA has been extensively studied, but in the context of RNA the function is largely unclear.[104] Studies have shown that it is important for the regulation involved in gene expression, including RNA export, ribosome assembly, translation, and RNA stability.[105] The modification affects the tertiary structure of tRNA, but also rRNA, and therefore ensures accurate RNA translation.[104] It could be shown that the small methyl group at position C5 has a huge impact on the conformation of the entire tRNA molecule, despite its size and hydrophobic properties.[106]

2.2.2   tRNA C5 methylation sites of DNMT2, NSUN2, and NSUN6

Methyl groups are transferred to RNA by methyltransferases (MTases).[107] However, these reactions require cofactors, such as N^5-methyltetrahydrofolate or the more commonly used *S*-adenosyl-L-methionine (**17**, SAM).[79,97] SAM is a ubiquitous cofactor in all living organisms, and after adenosine triphosphate (ATP) the second most widely used enzyme substrate.[79,80] In order to methylate cytidine in tRNA, the methyl group of the donor SAM is transferred to the 5-position by the respective methyltransferase, releasing *S*-adenosyl-L-homocysteine (**18**, SAH) as a byproduct (**Scheme 4**).[108]
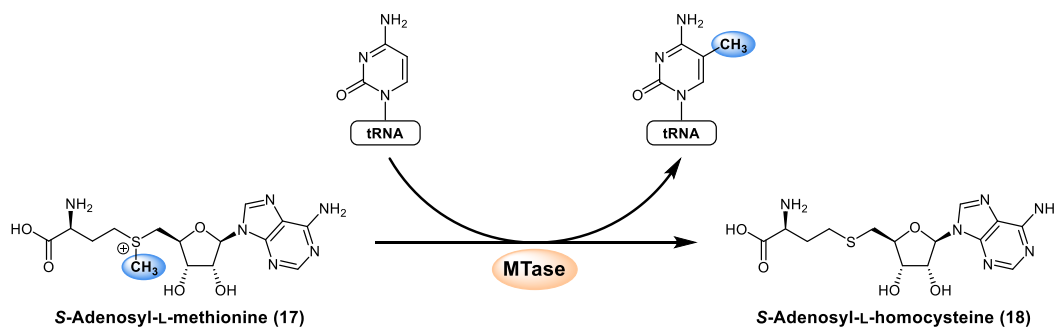
Scheme 4: General representation of the transfer of a methyl group from SAM to tRNA cytidine by MTases such as DNMT2, NSUN2, and NSUN6.

There are eight known RNA methyltransferases that are responsible for m5C modifications in eukaryotes, namely the DNA methyltransferase 2 (DNMT2) and the NOL1/Nop2/Sun family (NSUN1–7).[104] Originally, the eukaryotic DNMT2 was named for its homology to DNA methyltransferases.[109] Although it is a member of the DNA MTase family, its actual substrate is cytosolic tRNA[Asp] at position C38 in the anticodon loop (**Figure 5**).[108,110] Depending on the species, tRNA[Gly], tRNA[Val], and tRNA[Glu] are further substrates of DNMT2.[111] There are also some indications that the enzyme has DNA methylation activity. DNMT2 is able to recognize and modify DNA fragments when they are presented as covalent DNA-RNA hybrids in the structural context of a tRNA.[112] Additionally, it was discussed that DNMT2 can modify DNA at CG residues.[113] However, since it is primarily localized in the cytoplasm instead of the nucleus, it lacks the main properties of a typical DNA methylating enzyme such as its family members DNMT1 and DNMT3.[108]

The members of the NSUN family exhibit different specifications for their RNA substrates. Of the seven members, NSUN2, NSUN3, and NSUN6 are tRNA-methylating enzymes, but differ in terms of their localization. NSUN2 and NSUN6 modify cytoplasmic tRNA.[105] However, due to its predominant localization in the nucleus, NSUN2 appears to play a role in the methylation of tRNA during early-stage biogenesis.[114] NSUN6, on the other hand, is localized in the cytoplasm and, moreover, is enriched in proximity to the Golgi apparatus and pericentriolar matrix,[115] suggesting that the methylation it mediates is a late-stage process.[105] While NSUN6 is limited to m5C72 in the acceptor talk of tRNA[Cys/Thr],[115,116] NSUN2 has a broad substrate spectrum of target tRNAs with multiple methylation sites at C34, C48,

C49, and C50 (**Figure 5**).[114,117–119] Moreover, it also modifies mitochondrial tRNA (mt-tRNA) at positions 48–50[120] and is also able to methylate mRNA.[121]
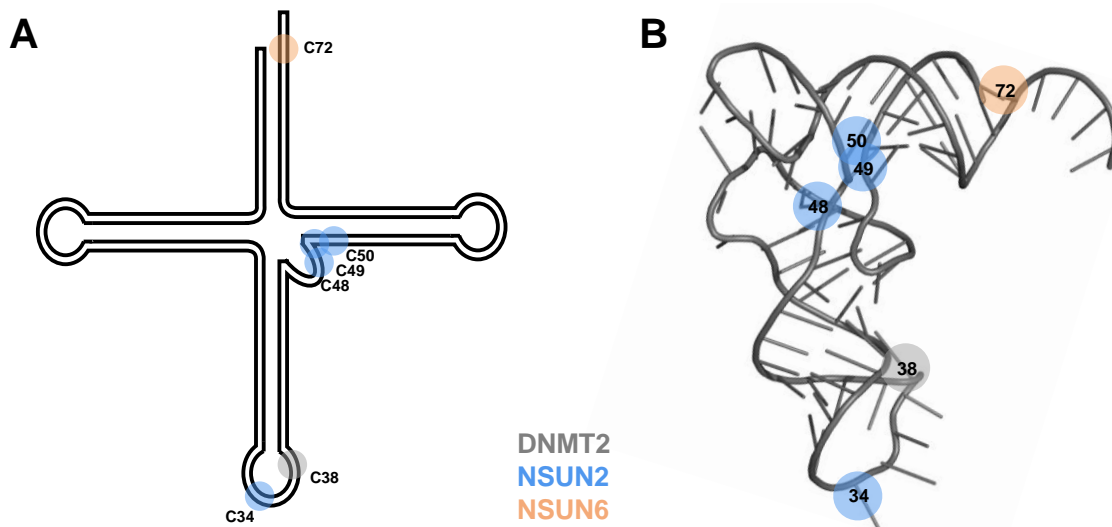
Figure 5: Depiction of the methylation sites of DNMT2 (grey), NSUN2 (blue), and NSUN6 (orange) in cytosolic tRNA. A: Schematic representation. B: Three-dimensional visualization.

Mitochondrial tRNA is mainly modified by another member of the NSUN family. Although NSUN3 and NSUN4 are synthesized on cytoplasmic ribosomes, they are primarily localized in mitochondria where they convert mt-RNA as substrates.[105,122] NSUN3 is required for the m5C34 modification in mt-tRNA^Met,[123–125] its mitochondrial member NSUN4 modifies rRNA.[126] The remaining NSUN family members are not associated with tRNA as substrate. NSUN1 and NSUN5 are specific to cytoplasmic rRNA.[105] The last member NSUN7 is assumed to target different enhancer RNAs (eRNA) at different positions, as NSUN7 depletion resulted in significant decreases in levels of these eRNAs.[127]

2.2.3   Biological functions of cytidine modification by DNMT2, NSUN2, and NSUN6

tRNA modifications are assumed to have different effects depending on the position at which they occur. If they take place within or near the anticodon, codon-anticodon interactions are affected, thus affecting tRNA function in translation. Modifications within the core appear to affect tRNA structure, stability, or both.[105] Methylation of C38 within the anticodon loop by DNMT2 improves the codon recognition and therefore the translation accuracy. It facilitates the discrimination between cognate and near-cognate codons, especially between Asp and Glu codons, which is required for accurate protein synthesis during hematopoiesis. A study demonstrated that lack of tRNA^Asp-m5C38 increases the mis-incorporation rate of amino acids caused by the near cognate codon of tRNA^Glu, which can lead to the production of aberrant proteins.[128] Furthermore, DNMT2-mediated C38 methylation of tRNA^Asp has a regulatory effect on the translation of poly-Asp-containing proteins. The modification increases the preference of the aspartyl-tRNA synthetase to tRNA^Asp-m5C38 four- to fivefold, which results in a higher aminoacylation rate.[129] Besides the improving effects in the translation process, DNMT2 methylation also protects the tRNA

under stress conditions. The enzyme relocalizes to stress granules when cells are exposed to heat shock. Absence of DNMT-mediated m^5C modification resulted in an increased formation of tRNA fragments caused by angiogenin-mediated stress-induced ribonuclease cleavage.[130] This leads to the suggestion that DNMT2 promotes tRNA stability and thus the rate of overall protein synthesis.[118,130] DNMT2 is also involved in other physiological processes, but most of them are not yet fully understood. Loss of DNMT2 in zebrafish resulted in a reduced size of morphants and also affected the development of retina, liver, and brain.[131] In contrast, no morphological effects were observed in flies, mice, and plants.[108] Overexpression in *Drosophila* prolonged life span and increased stress resistance.[132] DNMT2 is also required for the epigenetic transmission of phenotypes associated with the *Kit* and *Sox9* genes in mice, which are responsible for white colored tails and feet (*Kit*) as well as growth (*Sox9*). Neither of these phenotypes occurred in DNMT2 knockout mice.[133]

In the region of the anticodon, another methylation can also take place, namely C34 mediated by NSUN2. In tRNA$^{Leu(CCA)}$ this is not the final modification, but functions as a precursor. For instance, after intron-removal, the methyl group of C34 is oxidized to hm^5C, 5-formylcytosine (f^5C), or both. The RNA is then exported to the cytoplasm where 2'-*O*-ribose methylation occurs to form 5-hydroxymethyl-2'-*O*-methylcytidine (hm^5Cm), 5-formyl-2'-*O*-methylcytidine (f^5Cm), or both.[134,135] In yeast, oxidative stress leads to an increased methylation at C34 of tRNA$^{Leu(CCA)}$, resulting in an enhanced translation of mRNAs required for the biosynthesis of stress response proteins.[136] The other modifications installed by NSUN2 occur outside the anticodon loop, more precisely within the variable loop. C48 and G15 form a reverse Watson-Crick geometry between two antiparallel strands, referred to as the non-canonical "Levitt pair" (**Figure 6**).

Canonical Watson-Crick pair

Reverse Watson-Crick pair (Levitt pair) with m⁵C modification
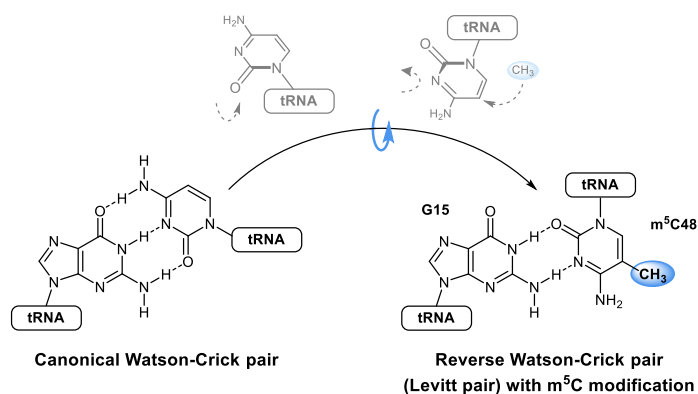
Figure 6: Comparison of the canonical Watson-Crick G-C base pairing and the Levitt pair formed between G15 and C48, which is often methylated (G15–m^5C48).

The formation is crucial for the characteristic L-shaped tertiary structure of most tRNA molecules.[137] NSUN2 methylation at this position increases the hydrophobicity of the nucleobase, which is assumed to enhance base stacking and promote stabilization of the interaction and thus of the tRNA tertiary

structure.[138] According to databases, the methylated Levitt pair G15–m^{5}C48 is found in 26% of tRNA sequences.[139]

The NSUN2 methylation cluster within the variable loop has an overall stabilizing effect as it protects the tRNA against stress-induced angiogenin-mediated endonucleolytic cleavage. Loss of NSUN2 results in fragmentation of tRNAs into short ncRNAs due to lack of m^{5}C48/49/50 modifications, leading to aberrant accumulation of cleaved tRNAs.[117] Overall, NSUN2 plays a significant role in various biological processes. It is involved in translation as it is required to metabolically support high protein synthesis rate.[140] In addition, NSUN2 affects the processing of vault tRNA (vtRNA)[141] and microRNA (miRNA).[142] Studies associate NSUN2 with stem cell function,[143] the differentiation of epidermal,[144] neural,[145] and germ stem cells,[146] as well as cell proliferation.[147] Interestingly, the enzyme appears to promote stress induced premature senescence.[148] NSUN2 was also found to affect mobile genetic element expression and impact genome stability.[149]

Compared to DNMT2 and NSUN2, the function of NSUN6 has been studied much less. The modification occurs at C72 in the acceptor stem close to the 3' CCA terminus of the tRNA molecule. It was shown that NSUN6 recognizes a specific nucleotide sequence as well as a correct folding of the full-length tRNA.[150] The exact biological function underlying this fact remains unknown. In *Pyrococcus horikoshii*, an anaerobic archaeon, the m^{5}C72 modification is assumed to slightly enhance the thermal stability of tRNA.[151] Notably, NSUN6-mediated methylation increases abundance and translation efficiency of mRNA.[152]

2.2.4 Links to pathophysiological processes

DNMT2, NSUN2, and NSUN6 execute various biological functions. Not surprisingly, correlations were also found with pathophysiological processes, e.g. the spread of cancer. DNMT2 was found to be upregulated in various tumor samples.[153] An overexpression of the enzyme was observed in several cancer tissues, such as cervical,[154] bladder tissue,[155] or lymph node metastases.[156] Furthermore, numerous somatic mutations in the enzyme were found in tumors originating from different tissues.[153] Studies suggest that somatic mutations can strongly alter the catalytic activity of DNMT2 and have a functional role in tumorigenesis.[157] Loss of DNMT2 led to a sensitization of cancer cells to radiotherapy and poly (ADP-ribose) polymerase (PARP) inhibitors, suggesting that the enzyme is involved in resistance to these therapeutic measures.[158] There is also evidence that DNMT2 plays a role in other pathophysiological processes besides cancer. Metabolic disorders and their epigenetic inheritance are linked to elevated m^{2}G and m^{5}C levels on sncRNA in mouse models.[159] Deletion of DNMT2, which is responsible for the modifications, abolished sperm sncRNA-mediated transmission of metabolic disorders induced by high-fat-diet.[160]

NSUN2 is a well-studied enzyme, so it is not surprising that much evidence has been found that this enzyme is involved in a variety of cancers.[161,162,171–175,163–170] Those findings are substantiated by the fact that NSUN2 promotes cell growth by regulating the expression of cyclin-dependent kinase 1 (CDK1).[176] Interestingly, NSUN2 can suppress sensitivity of cancer cells to treatment with 5-fluorouracil.[177] The enzyme is also associated to many other diseases, especially with neurological symptoms. Mutations in NSUN2 can cause intellectual disability[178–182] and autism spectrum disorder.[183] Neuronal NSUN2-deficiency results in depression and anxiety.[184] Furthermore, it is linked to the Dubowitz syndrome, an autosomal recessive disorder characterized by a constellation of growth, mild microcephaly, facial gestalt, mental retardation, eczema, and risk of malignancy.[185,186] Notably, accumulation of tRNA fragments due to depletion of NSUN2 has been found to impair brain development in mice, which could be an explanation for the cause of neurodevelopmental disorders in human.[145]

The involvement of NSUN6 in pathophysiological processes is largely unknown. NSUN6 appears to reduce MST1 kinase activity by methylation of Lys59 and activates yes-associated protein (YAP) target genes, which leads to cancer cell-induced osteoclast differentiation and bone metastasis.[187] Another study showed that the expression of NSUN6 is significantly increased in colorectal cancer tissue and classified the enzyme expression as a risk factor for this cancer type.[188]

2.3 Structures and catalytic mechanisms of DNMT2, NSUN2, and NSUN6

2.3.1 Evolution and structures

In 1999, the structural relationship of over 50 proteins, which are likely to be RNA m^{5}C methyltransferases, was identified by comparing genome and proteome sequences.[189] Summarized in a phylogenetic diagram (**Figure 7A**), the NSUN family is strongly represented, as well as DNMT2.[190]
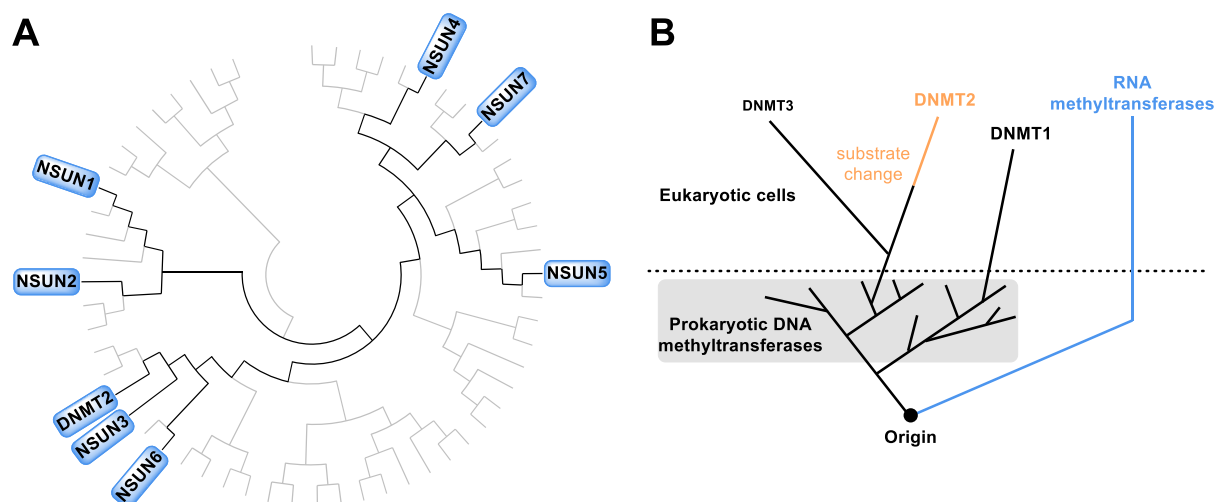

Figure 7: A: Phylogenetic diagram of the RNA methyltransferases DNMT2 and NSUN1–7.[190] B: Consensus model of phylogeny of DNMT2, DNMT1, and DNMT3.[110]

Overall, it is not much known about the origin of those RNA MTases. From an evolutionary point of view DNMT2 seems to be more related to DNMT3 than DNMT1 as several branches of prokaryotic DNA MTases separate their groups (**Figure 7B**).[110] It is possible that DNMT2 and DNMT3 are either derived from one common or two different bacterial precursors, but this cannot be confirmed at present.[110] Furthermore, it is assumed that the eukaryotic DNMT2, DNMT1, and DNMT3 families originated independently from prokaryotic DNA MTase ancestors.[110,111,191] In this context, DNMT2 probably switched its substrate specificity from DNA to tRNA, indicating an evolutionary relationship between cytosine-C5 methylation of tRNA and of DNA.[110,111,118,191–193]

RNA and DNA C5 cytosine methyltransferases all share a common core structure of a mixed seven-stranded β-sheet, the so-called "AdoMet-dependent MTase fold".[194] It consists of up to ten evolutionary conserved motifs (I–X), which can be in different order. Motifs I–III and V are required for the binding of SAM, motifs IV, VI, and VIII are involved in catalysis.[104] In DNMT enzymes a non-conserved region is located between motifs VIII and IX, referred to as the target recognition domain (TRD), which is involved in DNA recognition and specificity.[195,196]

DNMT2 consists of 391 amino acids corresponding to ca. 40 kDa in size, indicating that it is a relatively small protein compared to its family members DNMT1 and DNMT3A/B with 1616, 912, and 853 amino acids, respectively (**Figure 8**).[197–199]
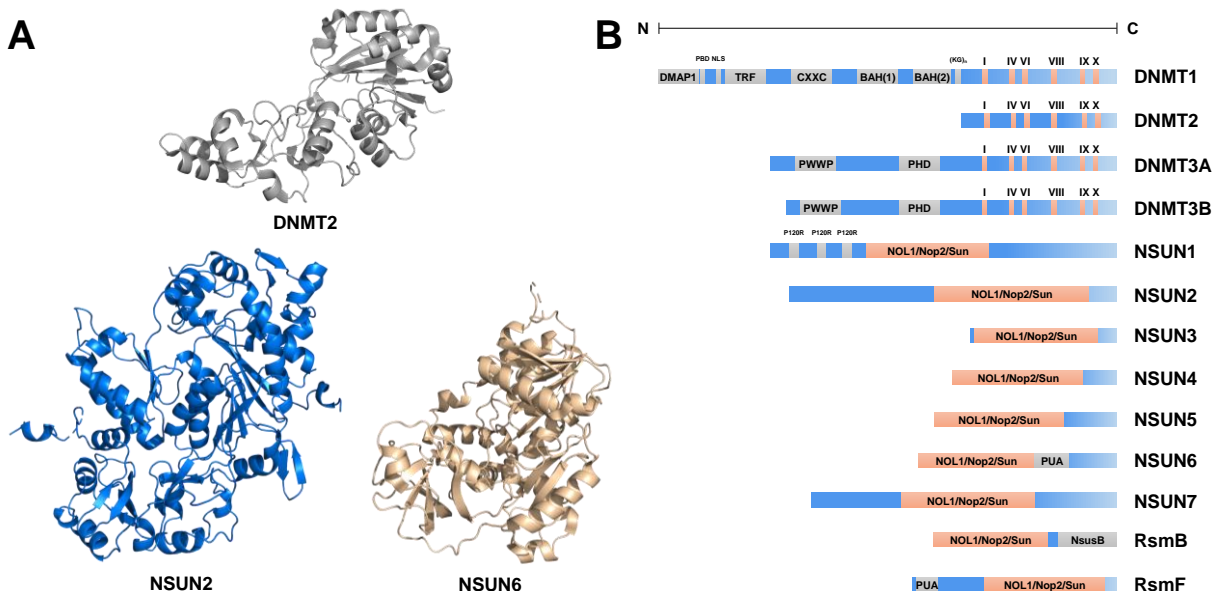
Figure 8: A: Crystal structure of hDNMT2 (grey, PDB-ID: 1G55),[197] AlphaFold model of hNSUN2 (blue),[200,201] and crystal structure of hNSUN6 (orange, PDB-ID: 5WWR).[116] B: Domain organization of animal DNMT family members,[198,199,202] RNA MTases from eukaryotes (NSUN1–7) and bacteria (RsmB and RsmF).[104]

It exclusively consists of the conserved C-terminal catalytic domain but lacks the regulatory N-terminal domain.[109,199] Interestingly, the DNMT2 catalytic domain features a unique cysteine-phenylalanine-threonine (CFT) motif, located between the catalytic motifs VIII and IX, which distinguishes it from other

MTases.[110,197,199] The catalytic motif IV of the DNMT enzymes consists of the common proline-cysteine-glutamine (PCQ) motif, including the catalytically active cysteine. In the crystal structure of human DNMT2, this exact part is disordered.[197] However, in the DNMT2 homologue in *Entamoeba histolytica* a more defined structure of the loop showing an *α*-helical conformation can be found.[203] Another conserved amino acid cluster participating in the catalytic mechanism is motif VI, also known as the ENV motif. It contains a glutamate allowing proton transfer to N3.[199]

NSUN2, at ca. 86 kDa, is the second largest member after NSUN1 of the NSUN family.[204] With ca. 51 kDa in size, NSUN6 is a smaller representative, similar to NSUN4 and NSUN5.[104,116] The NSUN family members all share the conserved catalytic C-terminal domain, which can also be found in the prokaryotic RNA MTases RsmB and RsmF (**Figure 8**). Motifs IV and VI contain both one cysteine, respectively, required for the catalytic mechanism.[104] In addition, motif IV also contains an aspartate, which enables proton transfer to N3 during catalysis.[105]

2.3.2   Catalytic mechanisms of cytidine methylation

Methylation at position 5 of cytosine cannot occur spontaneously even though the cofactor SAM is a very effective methyl group donor. Given the fact that cytosine is an electron-deficient aromatic ring, a nucleophilic attack from position 5 on the methyl group of SAM cannot take place.[196] For this reason, an enzyme-mediated catalysis is required. In the first step, the catalytically active cysteine from motif IV (DNMTs) or motif VI (NSUNs) follows the reaction of a Michael addition (**Scheme 5**) attacking position 6 of cytosine, which results in the formation of a covalent complex between the polynucleotide and the enzyme.[205,206] A transient protonation at position N3 occurs simultaneously, facilitating the nucleophilic attack on position 6.[207]
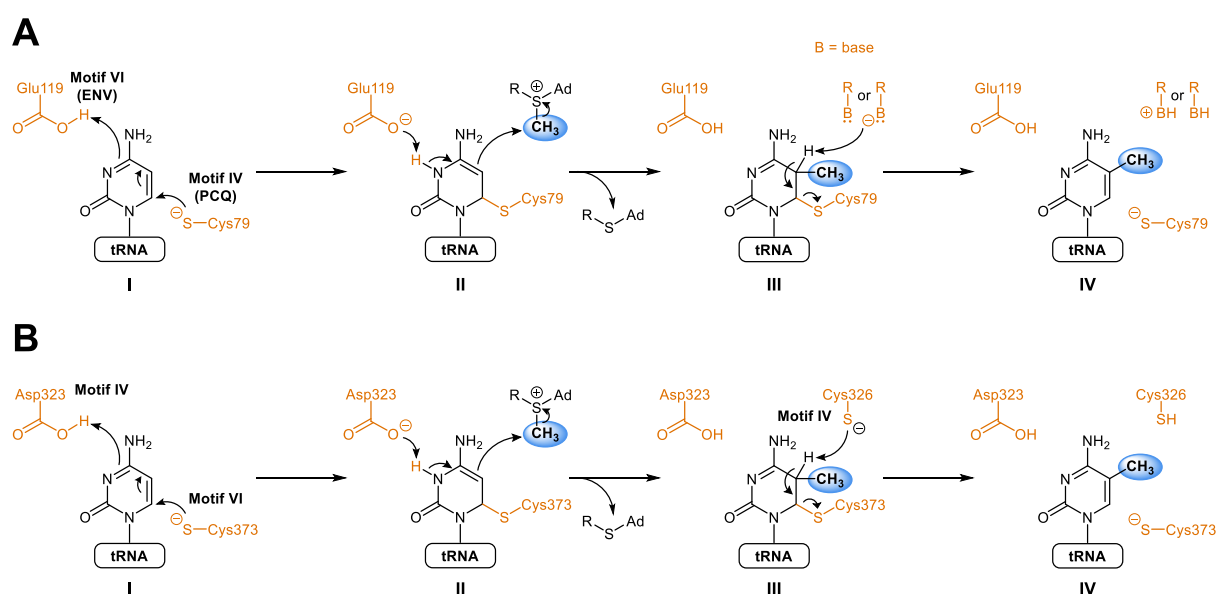
Scheme 5: Catalytic mechanisms of C5 cytosine methylation by A: DNMTs (using the example of DNMT2) and B: NSUNs.

The proton transfer is mediated by either glutamate (DNMTs, motif VI) or aspartate (NSUNs, motif IV).[105,199] The Michael addition leads to a string activation of position 5 enabling the nucleophilic attack on the methyl group of SAM. After the methyl group transfer, base-mediated deprotonation at position 5 results in the elimination of the thiol group of cysteine.[205,206] In NSUN enzymes, this function is carried out by the thiolate group of the second cysteine from motif IV.[105] The covalent bond between polynucleotide and enzyme is resolved and the aromaticity of cytosine restored.

Interestingly, DNMT2 methylates tRNA using a DNA MTase-like mechanism. This fact was discovered as similar residues identified from motifs IV, VI, and VIII, which take part in catalysis, are also present in DNA MTases.[193] It also contains a strongly conserved cysteine (Cys292) in the CFT motif unlike the other DNMT family members. It is assumed that it is involved in catalysis as exchange of Cys292 strongly reduced the catalytic activity of DNMT2.[193]

2.4 Inhibitors of SAM-dependent methyltransferases

Since correlations between RNA methylation and the spread of cancer but also the infectiousness of viruses have been discovered,[208,209] modulation of MTase activity by small molecule inhibitors appears to be a promising strategy as disease treatment. Most currently known RNA MTase inhibitors are derived from natural substrates, such as SAH and SAM.[103] SAH, which results as a reaction product in the methyl group transfer, was found to be a nonselective feedback inhibitor of SAM-dependent MTases.[206,210–214] Similar, sinefungin (**19**, SFG) (**Figure 9**), a natural SAM-related nucleoside that originally was isolated from *Streptomyces griseolus*,[215] showed competitive inhibition of several MTases.[211,214] For DNMT2, SAH and SFG showed inhibition in the low micromolar range with IC_{50} values of 15.8 ± 1.5 μM and 13.2 ± 0.8 μM, respectively.[214] The concept of developing inhibitors of SAM-dependent RNA MTases based on the SAH scaffold has been widely exploited for targets in different organisms: the human mRNA MTase complex METTL3-METTL14,[216–219] the mRNA MTase Ecm1 in *Encephalitozoon cuniculi*,[220,221] the bacterial *E. coli* tRNA MTase TrmD,[222] and rRNA MTase RlmJ,[216] as well as various viral targets.[213,223–228] SAH-based DNMT2 inhibitors, however, were described but showed very low affinity.[229]

Overall, the derivatization of SAH appears to be a successful strategy for the development of inhibitors and tool compounds not only for RNA MTases but also for other SAM-dependent MTases. This concept was carried out for various targets: the catechol-*O*-MTase (COMT)[230–232] and protein MTases (PMTs) including histone MTases (HMTs) such as DOT1L.[233] Notably, pinometostat (**20**) (**Figure 9**), a DOT1L inhibitor is currently being tested in clinical trials.[234] Inhibitors of the DNA MTase family have so far only been developed for DNMT1 and DNMT3B2.[235,236]

In addition to derivatization of SAH, modified nucleoside-based molecules have also been shown to be effective inhibitors. 5-Azacytidine (**21**) and decitabine (**22**) (**Figure 9**), both used as chemotherapeutic agents, inhibit RNA and DNA MTases and induced a change of methylation activity in cancer patients suffering from myelodysplastic syndrome or leukemia.[237–239] Notably, it could be shown that 5-azacytidine treatment reduces activity of DNMT2, which is variably expressed in cancer cell lines.[240] The drugs are randomly incorporated into nascent DNA (decitabine and 5-azacytidine) and RNA (5-azacytidine) by respective polymerases during transcription and replication. Since both structures are substituted with nitrogen in position 5 of the aromatic ring, DNA and RNA MTases remain covalently bound to their substrate, resulting in inhibition of catalytic activity (**Scheme 6**).[241,242]
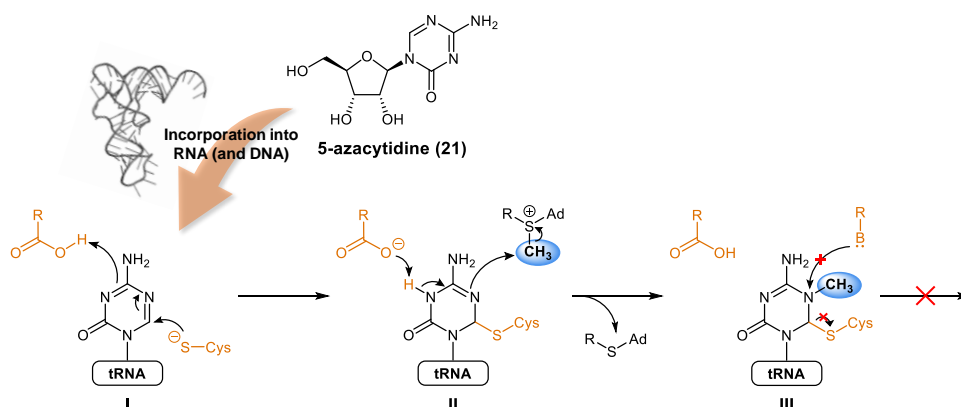
Scheme 6: Mechanism of the inhibition of DNA and RNA MTases by the example of 5-azacytidine.

The pyrimidinone derivative zebularine (**23**) (**Figure 9**) shows a similar mechanism of inhibition. A covalent complex is formed with DNA MTases, such as C5 Mtase from *Haemophilus hemolyticus* (*M.HhaI*) that has close structural resemblance to human DNMT2.[243] Although the inhibition of DNMT2 itself has not been reported yet, an analogous mechanism of inhibition by zebularine is supposed due to its structural similarity to 5-azacytidine. In addition, incorporation of zebularine into tRNA could be confirmed.[244]
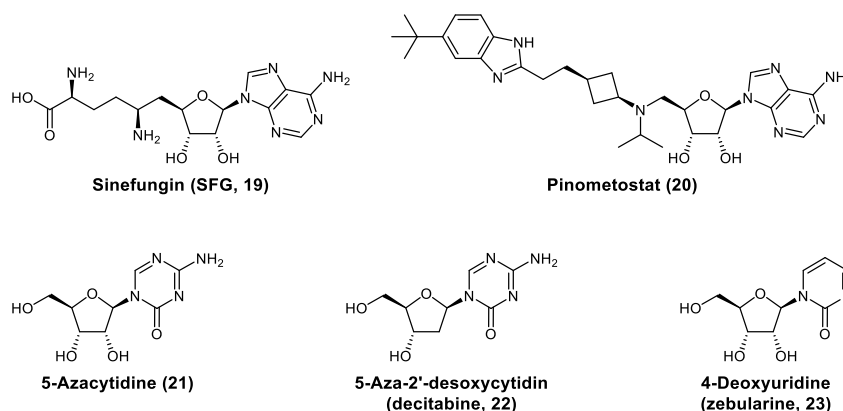

Figure 9: Structures of sinefungin (**19**), pinometostat (**20**), 5-azacytidine (**21**), decitabine (**22**), and zebularine (**23**).

MTase inhibitor research is also focused on developing compounds that strongly differ from the SAH scaffold. Various efficient non-SAH-like inhibitors have been developed for numerous targets of different organisms.[103]

2.5 Biophysical methods for evaluation of MTase inhibitors

2.5.1 Microscale thermophoresis

Microscale thermophoresis (MST) is based on the principle of thermophoresis, which describes the movement of molecules within a temperature gradient.[245] The molecular movement is dependent on various molecular properties, such as size, charge, and solvation entropy.[245,246] Usually, these parameters are slightly altered by ligand binding,[247] making the method highly sensitive to change in molecular properties and thus allowing precise quantification of molecular events.[245] The temperature gradient is generated by an infrared (IR) laser directed at capillaries containing the sample.[245] In most cases, the protein must be fluorescently labeled.[247,248] The labeled sample is excited by an UV LED, and the emitted fluorescence is detected with the same objective through which the laser is focused.[249] When the temperature increases, the local molecule concentration distributes, resulting in a change in fluorescence. During the thermophoresis measurement, different stages are observable (**Figure 10**).
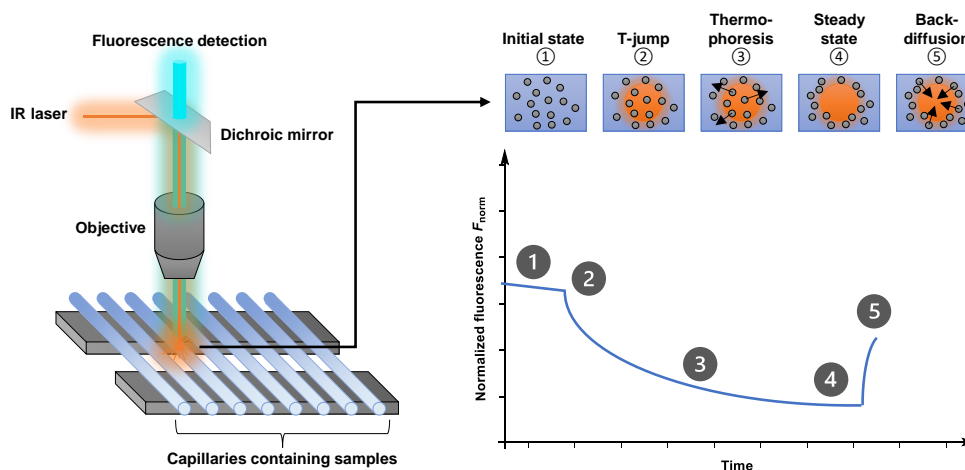
Figure 10: Setup and illustration of an observable MST signal.

The initial stage represents the conditions before the heating, i.e. homogeneous distribution (Stage 1).[247,249] When the laser is activated, an abrupt change in fluorescence intensity is observable, indicated by a "T-jump" (Stage 2).[247] This is affected by the local surroundings of the fluorophore, such as changes in conformation or a binding event near the fluorophore.[247,250,251] Subsequently, slow thermophoretic movement occurs (Stage 3), which creates a concentration gradient as the labeled molecules diffuse out of the heated sample volume.[245,247] This thermophoresis can also be influenced by a binding event.[247] After some measurement time (20–30 s), fluorescence reaches a plateau, which

indicates a steady state (Stage 4). In this stage, the thermodiffusion is counterbalanced by mass diffusion.[247] After deactivation of the IR laser, a "backdiffusion" of molecules can be observed, driven by mass diffusion. This is indicated by an inverse T-jump (Stage 5).[245] For binding analysis, the change in thermophoresis is determined by the change in normalized fluorescence ΔF_{norm}. F_{norm} is defined as the quotient of the measured fluorescence after thermophoresis F_1 and the initial fluorescence (or the fluorescence after the T-jump) F_0 (**Eq. 6**).[245,247]

$$F_{norm} = \frac{F_1}{F_0}$$ **Eq. 6**

In this work, results of tested compounds were determined by different MST methods. At the beginning of this project, a standard MST method was used to identify binders of the target of interest. At a later stage of this project, an optimized MST displacement method using a fluorescent ligand (for details see **Section 3.8**) was applied, which allowed semi-quantitative screening outcomes and a more accurate evaluation of potential inhibitors. If compounds were tested using the optimized MST method, it was indicated.

2.5.2 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a routinely used technique to define the thermodynamics of interactions between biomolecules and ligands.[252] Moreover, the method determines the binding equilibrium between a protein and a ligand by measuring the evolving heat caused by the association.[253] The thermodynamics are characterized by the stoichiometry (n), the association constant (K_a), binding enthalpy (ΔH_b), the free energy (ΔG_b), entropy (ΔS_b), and heat capacity of binding (ΔC_p). Of these parameters, the stoichiometry (n), the association constant (K_a), and the enthalpy (ΔH_b) can be determined in a single experiment, while the free energy, and the entropy result from the value of K_a.[253] Most ITC instruments are based on the dynamic power compensation principle (**Figure 11**).[254]
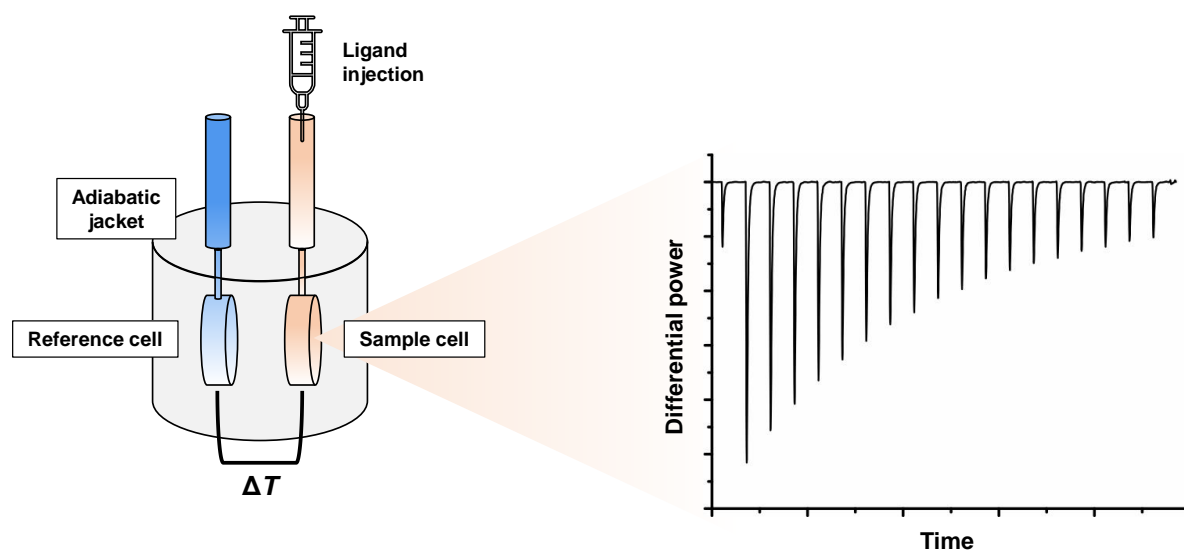

Figure 11: Representation of a power compensation ITC (left) with a resulting thermogram (right).

They consist of two identical coin-shaped cells (sample and reference) enclosed in an adiabatic jacket, a highly efficient thermal conducting material.[253,254] On the sample and reference cell, heaters are located to maintain zero temperature difference between both cells (isothermal).[252] In the sample cell, the protein solution is placed, the reference cell usually contains water or buffer.[253] Then, a ligand is injected into the sample cell. If a binding event takes place, the resulting change of heat is monitored by measuring the differential power applied to the cell heaters.[252] With each injection of the ligand, the binding reaction is triggered, and depending on the concentration of the reactants, a certain amount of complex is formed.[254] During formation, heat is either released (exothermic binding) or absorbed (endothermic binding), resulting in temperature changes in the cells. This directly affects the thermal power applied to compensate the temperature unbalance.[254] The raw signal is the power applied to the control heater as a function of time (μcal/s),[252] which appears as a series of peaks corresponding to multiple injections. By integrating the area under the peak, the amount of heat associated with the injection can be calculated.[252,254]

2.5.3 Differential scanning fluorimetry

Differential scanning fluorimetry (DSF), also known as thermal shift assay (TSA), is a rapid and cost-effective ligand screening method based on the biophysical principle of ligand-induced thermal stabilization of the target protein.[255–258] In this method, the thermal unfolding of the target protein exposed to a temperature gradient in the presence and absence of the ligand is monitored (**Figure 12**).[259]
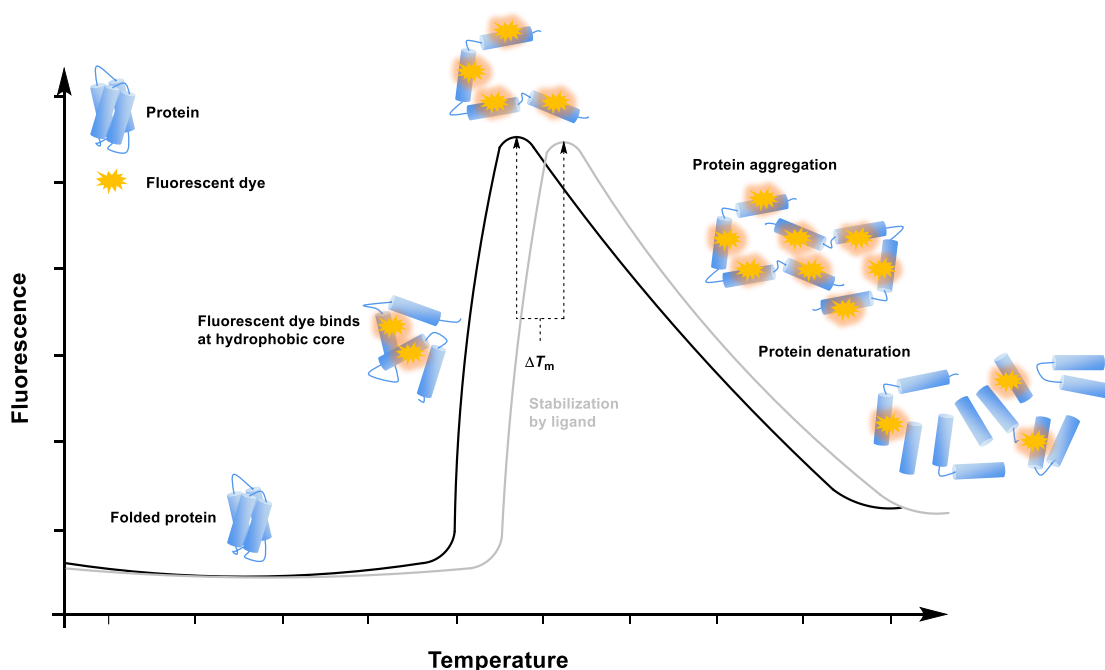

Figure 12: Thermogram of a differential scanning fluorimetry assay with depiction of the different denaturation stages and attachment of fluorescent dye.

A fluorescent dye is often added to the sample solution.[258,260] Being environmentally sensitive, the dye is highly fluorescent in nonpolar environments, whereas it is quenched in aqueous solutions.[259] As the

protein unfolds with rising temperature, its nonpolar amino acids, normally located inside, are exposed to which the dye can bind.[257,259] As a result, the measurable fluorescent increases.[256] In presence of a ligand, the binding event leads to a stabilization of the natively folded protein by reducing the Gibbs free energy of the complex, thus increasing the temperature required for unfolding.[257,259] The observable temperature shift ΔT_{m} caused by the stabilizing effect of the compound is proportional to its concentration, which allows the determination of affinity.[259]

2.5.4 Tritium incorporation assay

A common method to determine MTase activity in presence of inhibitors is the tritium incorporation assay.[112,197,261] In this method, the amount of methylation by utilizing SAM with a tritium-labeled methyl group (**24**, ^{3}H-SAM) is quantified. ^{3}H-SAM, tRNA, MTase, and a ligand of interest are combined (**Scheme 7**). After a certain incubation time, the substrates are precipitated on a filter paper by adding trichloroacetic acid (TCA).[112,261] The excess of unreacted cofactor is removed by washing, and a scintillation cocktail, containing diisopropylnaphthaline isomers (**25**, DIPN) as solvent and 1,4-bis-(2-methylstyryl)-benzene (**26**) and 2,5-diphenyloxazole (**27**) as scintillators is added. The solvent absorbs the energy from the tritium and transfers it to the scintillators, which re-emit it at a higher wavelength (visible light) at around 420 nm.[262] Based on this principle, ^{3}H-labeled tRNA can be quantified by liquid scintillation counting. By referencing the scintillation increase to a positive control (sample minus inhibitor), the inhibition can be determined.[112,261] The lower the signal, the higher the inhibition.
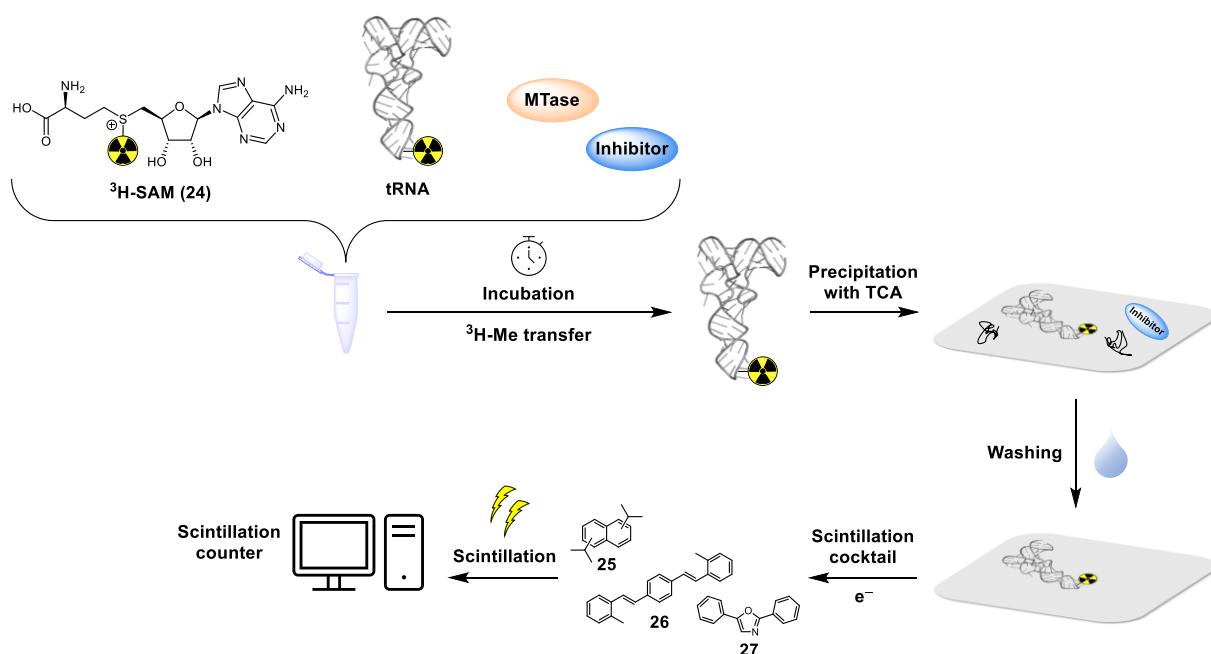
Scheme 7: Workflow of the tritium incorporation assay.

2.6   Drugs vs. activity-based probes

When it comes to the development of inhibitors for a certain target of interest, it is not always the aim to establish a drug. In some cases, an inhibitor is required as a chemical probe – or more precisely, as an activity-based probe (ABP), a tool compound to modulate the function of a biomolecular target, e.g. an enzyme, in cell assays or animal studies by covalent modification of the active site.[263,264] This approach, also referred to as activity-based protein profiling (ABPP), is an often-used concept to help understanding the functional state of enzymes in biological systems.[265,266] It can be used to explore gene function or elucidate the roles of the targeted proteins in healthy and diseased cells and tissues allowing association to a specific disease.[264,267] Since ABPs are used to link a phenotype to a gene, they must meet different quality criteria than drugs (**Table 1**).[264,267] A drug does not have to be highly selective, as long as it is safe and has satisfactory efficacy.[267] ABPs, on the other hand, must exhibit selectivity, high potency, and on-target action, which may refer to a single target or a target family.[263,267] Drug-like properties, such as good pharmacodynamics and oral bioavailability do not have to be met.[267] However, to achieve an observable effect on a cellular basis, the ABP should have sufficient permeability, similar to drugs.

ABPs have become powerful tools in chemical biology and provide a wide range of advantages. Not only are they complementary to genetic approaches, such as CRISPR and RNAi,[267,268] they can rapidly and reversibly inhibit a target in cells or animals and reveal temporal features of target inhibition.[267] As a comparison, using RNAi or gene knockout to achieve transcript and protein depletion requires time (usually of the order of days), which allows cells to compensate for the loss of protein levels or they may not survive due to loss of essential proteins.[269] The application of ABPs is not limited to a specific cell type but can be used in various cell types, even neuronal or primary immune cells that are difficult to manipulate by genetic means.[269] Notably, the span of advantages also extends into the field of drug development, as an ABP approach can be used for translational studies by mimicking the pharmacology of a therapeutic drug, which can be helpful for the development of new medicines.[267]

Table 1: Comparison of drugs and activity-based probes regarding purposes and requirements.

| Drugs	Activity-based probes
➤ Must be safe and effective	➤ Ask a specific biological question
➤ May have undefined mechanism of action	➤ Requires defined mechanism of action
➤ Human bioavailability required	➤ Must have high selectivity
➤ High demands for physiochemical and pharmaceutic properties	➤ Drug-like properties not necessarily required

3 Results and discussion

Note: Unless stated otherwise, inhibitor assays were performed using human MTases.
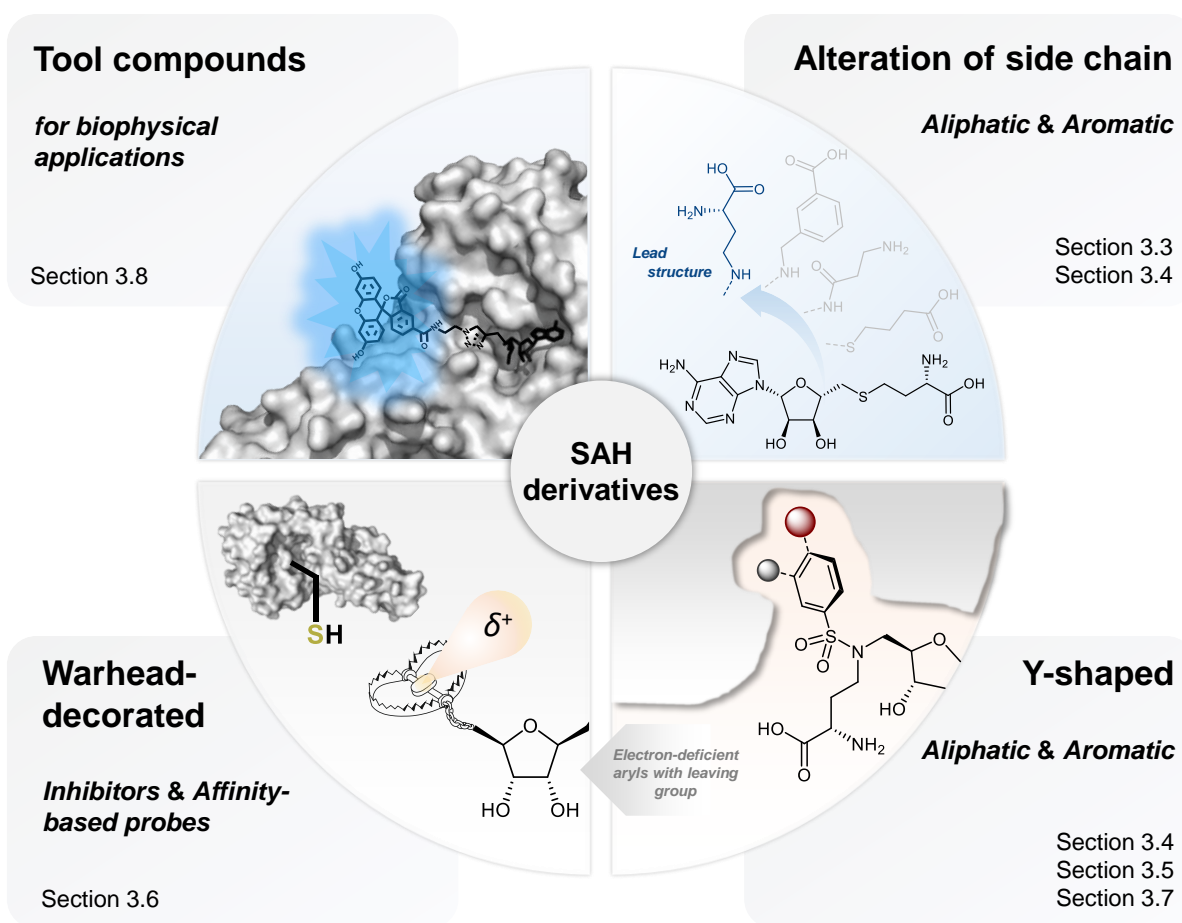
3.1 Project overview

This project was a collaboration between the ████████ group and the ████ group (both Johannes Gutenberg University Mainz) and aimed at the discovery and optimization of potential lead structures that selectively inhibit the MTases DNMT2, NSUN2, and NSUN6, with a primary focus on DNMT2 (**Scheme 8**). If a lead structure has been identified, it should be optimized in terms of affinity and selectivity. Once this has been accomplished, the structure should be modified with reactive groups to enable covalent binding with catalytically active cysteine residues inside the binding sites. Potent inhibitors were tested in a cellular context to analyze their membrane permeability as well as their effects in the biological system. By attaching fluorescent dyes to a potent inhibiting structure, an ABP should be created that can be applied to biological assays. This approach helps to understand the functional states of the MTases and to investigate the biological impact of their RNA modifications, but also facilitates the development of new inhibitors as potential drugs.
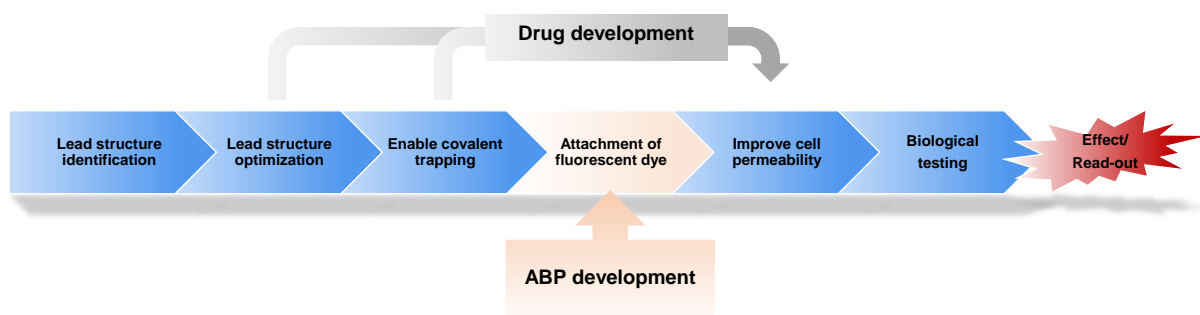
Scheme 8: Process flowchart of the collaborative project between the ████████ group and the ████ group.

The development of efficient inhibitors and ABPs is a challenging endeavor, especially for methyltransferases such as DNMT2, NSUN2, or NSUN6. So far, only very few inhibitors are known for these targets – mainly natural ligands like SAH or SFG that function as pan-inhibitors with low selectivity. Notably, both structures highly resemble each other, which provides little information about potential structure-activity relationships. Therefore, different approaches were pursued to generate a variety of structures based on the SAH scaffold, which serve different applications (**Figure 13**). First, the amino acid side chain of SAH was altered to investigate the pharmacophoric significance of functional groups and to reveal alternative structural elements (**Sections 3.3** and **3.4**). Furthermore, possible replacements, particularly nitrogen-containing groups for the sulfur unit were analyzed that might be used for additional structural attachments targeting subpockets such as the cytidine site. After a suitable side chain had been found that provides such attachment points, various side chains based on aliphatic or aromatic substructures were examined to generate Y-shaped inhibitors with improved affinity and selec-

tivity (**Sections 3.4**, **3.5**, and **3.7**). Moreover, electron-deficient aromatic moieties decorated with leaving groups were developed to enable a covalent reaction with catalytically active cysteine residues of the targets (**Section 3.5**). To extend the pool of potential warheads, SAH derivatives modified with different electrophilic groups were prepared and analyzed for their inhibitory activity (**Section 3.6**). Covalent SAH-derived inhibitors provide a suitable basis for the development of fluorescently labeled ABPs. Such tool compounds are useful for future studies to understand RNA methyltransferases and the biological impact of their RNA modifications. Another part of this work was about the synthesis of a fluorescent tool compound that can be used for the establishment of new biophysical methods to analyze the binding behavior of MTase inhibitors (**Section 3.8**).

Figure 13: Overview of projects that include the development of SAH congeners for different applications.

3.2 Method establishment for the synthesis of SAH derivatives

To enable a synthetic procedure that allows fast generation of a large molecule library, a feasible reaction type with suitable building blocks must be found. According to literature, nucleophilic substitution is one of the most used procedures in medicinal chemistry to generate amines.[270] By considering a retrosynthesis of amino-based SAH derivatives (**Scheme 9**), the structure can be divided into two main building blocks allowing nucleophilic substitutions to be applied. The adenosine component **31** bearing a leaving group at the 5'-position enables the conversion with various amines, e.g. (*S*)-2,4-diaminobutanoic acid (Dab)-based building blocks (**30**). For the amino acid substructure, *tert*-butyl ester and a Boc group were chosen as protecting groups while an isopropylidene group was chosen for the ribosyl structure because all these groups are expected to be cleaved in a single deprotection step.
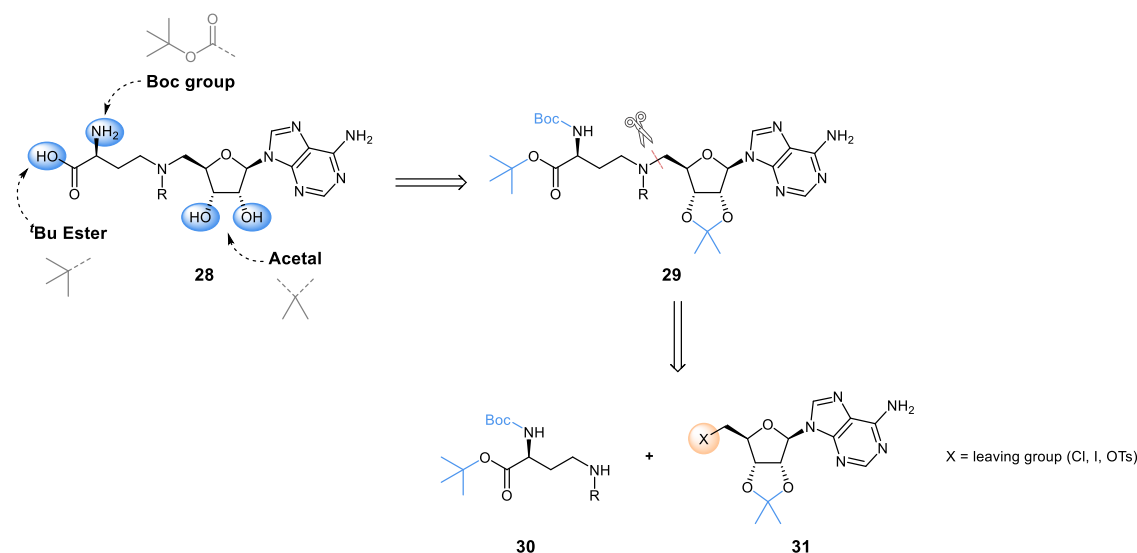
Scheme 9: Retrosynthesis of amine-based SAH derivatives to educts allowing nucleophilic substitutions to be applied.

In order to synthesize the leaving group bearing building blocks, adenosine (**32**) was first protected using acetone and catalytic amounts of *p*-toluenesulfonic acid at room temperature (**Scheme 10**). The resulting protected adenosine **33** was first treated with iodine, triphenylphosphine, and pyridine using an adapted procedure by MARTÍNEZ-MONTERO *et al.* to give the 5'-iodo derivative **34**.[271] Unfortunately, the desired product could not be isolated as purification approaches led to decomposition (determined by thin-layer chromatography). In an alternative approach, **33** was modified with tosyl chloride in pyridine according to THOMPSON *et al.* to yield the 5'-*O*-tosylated adenosine **35**.[272] Attempts were made to react this building block with different amines, such as *n*-propylamine (**36**), benzylamine (**37**), and 4-aminobutanoic acid (**38**) in DMF or pyridine at room temperature and under heating (70–80 °C). None of these reactions showed any conversion into the desired products **39**, **40**, and **41**. Another building block tested was 5'-chloro-5'-deoxyadenosine (**42**), which was obtained by treating adenosine first with thionyl chloride, followed by ammonia.[273] A test experiment using *n*-propylamine in DMF under heating did not yield any product.
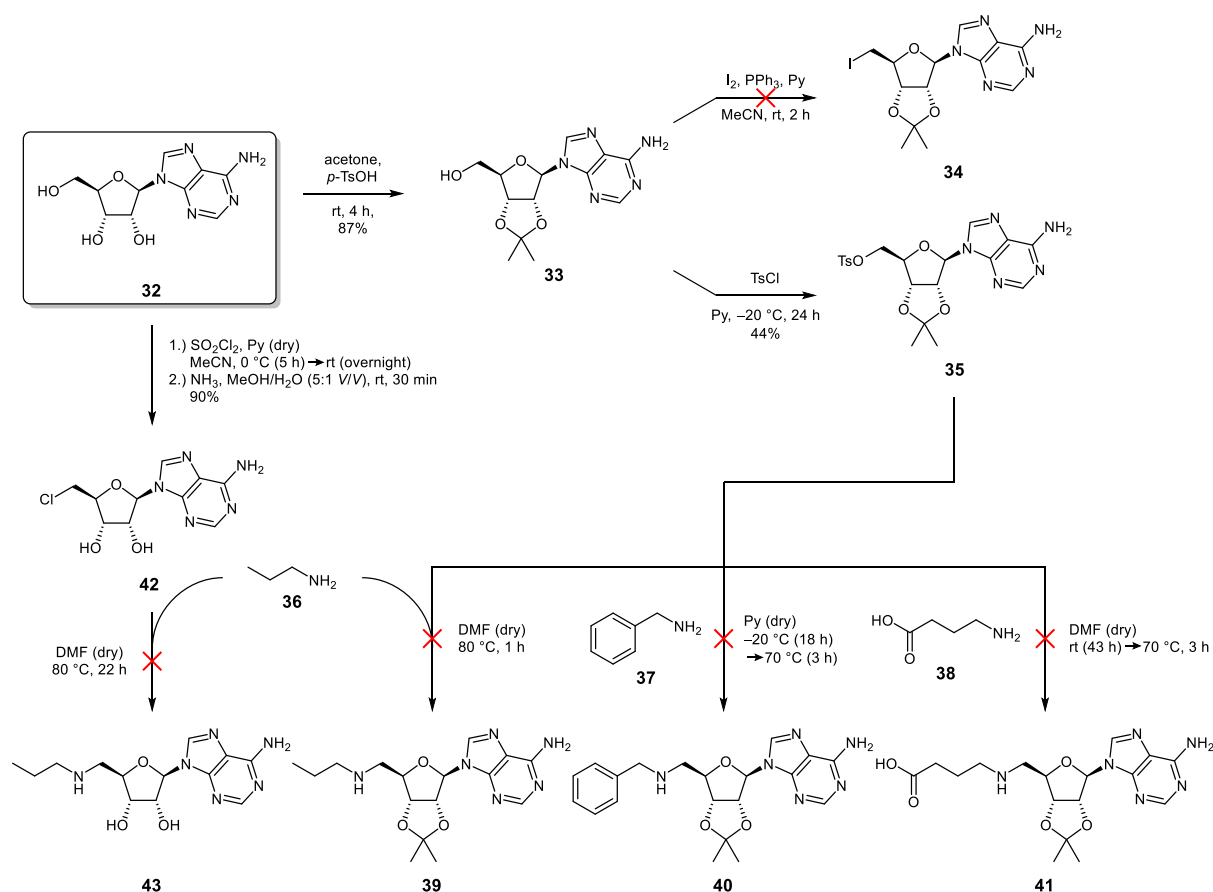
Scheme 10: Approaches to synthesize adenosine building blocks with leaving groups at the 5'-position, followed by nucleophilic substitution with different amines.

Since the nucleophilic substitution did not result in a suitable method for the synthesis of SAH derivatives, an alternative procedure was investigated. Among the most applied reactions in medicinal chemistry, the reductive amination is a common reaction type for amine synthesis.[274,275] Retrosynthesis shows the breakdown into an amine (**30**) and an aldehyde (**44**) building block for the application of reductive amination (**Scheme 11**).
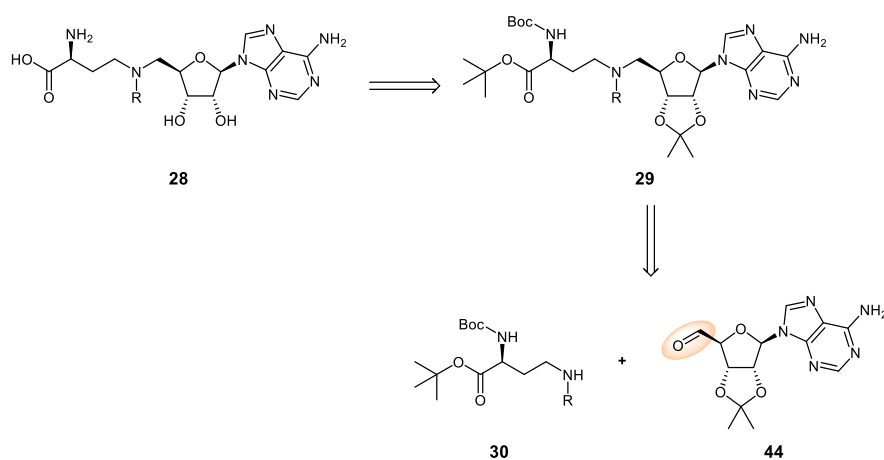
Scheme 11: Retrosynthesis of amine-based SAH derivatives to educts allowing reductive aminations to be applied.

According to BARTON *et al*., the protected adenosine **33** was treated with *N,N'*-dicyclohexylcarbodiimide (DCC), DMSO and dichloroacetic acid (DCA) in a PFITZNER-MOFFATT oxidation (**Scheme 12**).[276] Unfortunately, the desired product could not be detected by LC-MS. Testing the reaction by replacing DCA with trifluoroacetic acid (TFA) and pyridine,[277] the same result was observed. In a third approach, DESS-MARTIN periodinane (DMP) in DCM was used, but the oxidation also failed.
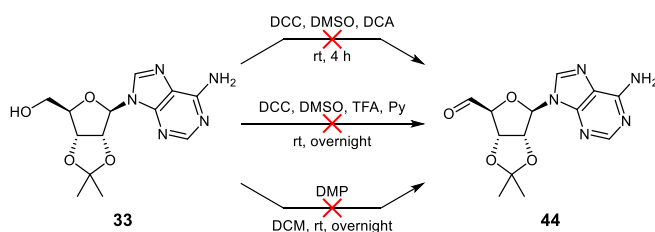
Scheme 12: Approaches to oxidize 2',3'-*O*-isopropylideneadenosine **33** at the 5'-position to the respective aldehyde **44**.

The synthesis experiments showed that the adenosine-like building block was not suitable to bear the aldehyde function. For this reason, the strategy of swapping the functions of the building blocks was pursued. In this context, the 5'-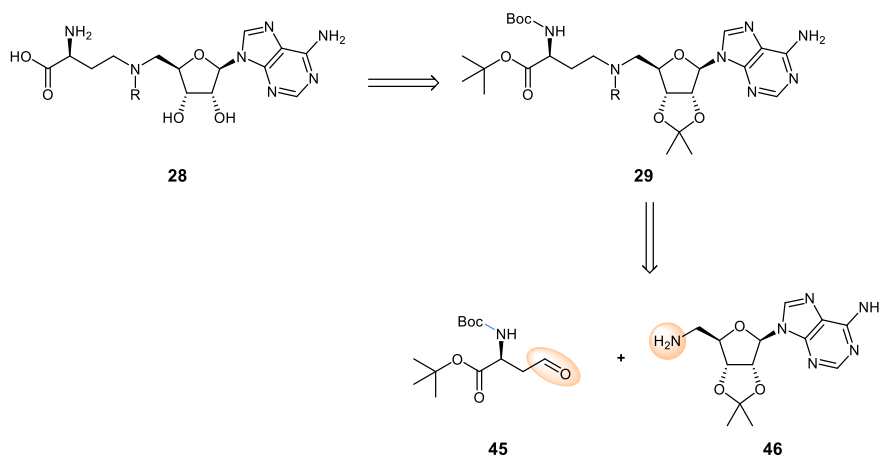amino-5'-deoxy derivative **46** was selected as the main building block for conversion with different aldehydes (**45**) to create a molecular library (**Scheme 13**).

Scheme 13: Retrosynthesis of amine-based SAH derivatives into alternative building blocks bearing amine and aldehyde function.

The synthesis was carried out via a GABRIEL synthesis according to LIU *et al*., starting from the protected adenosine **33** (**Scheme 14**).[278] In the first step, it was modified with phthalimide via a MITSUNOBU reaction using diisopropyl azodicarboxylate (DIAD) to yield compound **47**. In the following reaction, the phthalimide **47** was cleaved with hydrazine to yield 5'-amino-5'-deoxy-2',3'-*O*-isopropylidene adenosine (**46**).
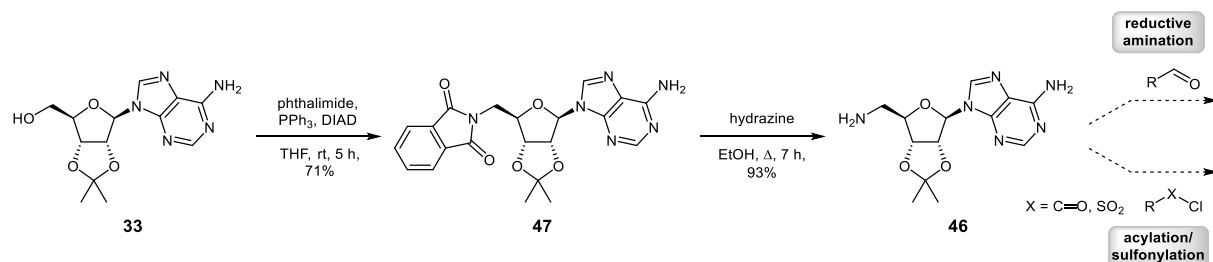
Scheme 14: Synthesis of 5'-amino-5'-deoxy-2',3'-*O*-isopropylidene adenosine (**46**).

To establish the complete synthetic pathway to a final compound that includes a valid deprotection procedure, *N*-adenosyl-(*S*)-2,4-diaminobutanoic acid (**51**, adenosyl-Dab) was intended to be synthesized according to ZHANG *et al.* (**Scheme 15**).[279] First, Boc and *tert*-butyl protected aspartate (**48**) was converted into the respective Weinreb amide **49** using 1,1'-carbonyldiimidazole (CDI). In the following step, **49** was reduced utilizing diisobutylaluminium hydride (DIBAL) at −78 °C. The resulting aldehyde **45** was reacted with the amine building block **46** by reductive amination to give the protected adenosyl-Dab **50**.
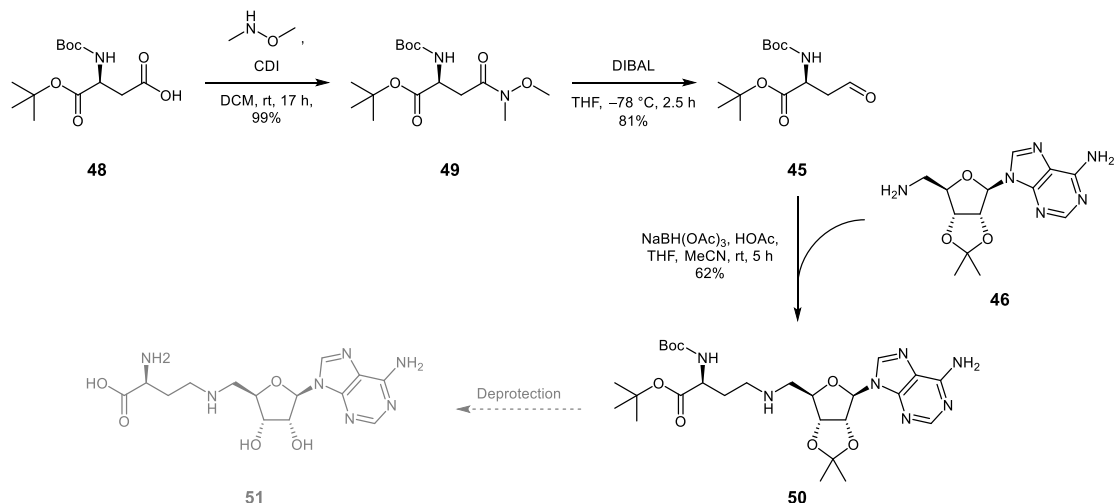
Scheme 15: Synthesis of the protected adenosyl-Dab **50**.

The reaction was carried out using sodium triacetoxyborohydride and acetic acid in THF at room temperature. Besides the formation of the desired product, LC-MS analysis revealed the presence of two byproducts showing a mass-to-charge ratio of $m/z = 490$ ([M+H]$^+$) and $m/z = 821$ ([M+H]$^+$), respectively. The molecular weight of 820 Da could be assigned to the double alkylated product **52** formed by two successive reductive amination reactions with the aldehyde **45** (**Scheme 16A**). A compound with 489 Da suggests a cyclic analog (**57**) that forms in the presence of protons as proposed in **Scheme 16B**. The proton-mediated activation of the ester function allows an intramolecular nucleophilic attack of the secondary amine, which results in a tetrahedral intermediate **53**. After proton transfer, *tert*-butanol is eliminated, yielding the protonated species **56**. In the final step, the loss of a proton results in the final cyclic byproduct **57**.

A

B
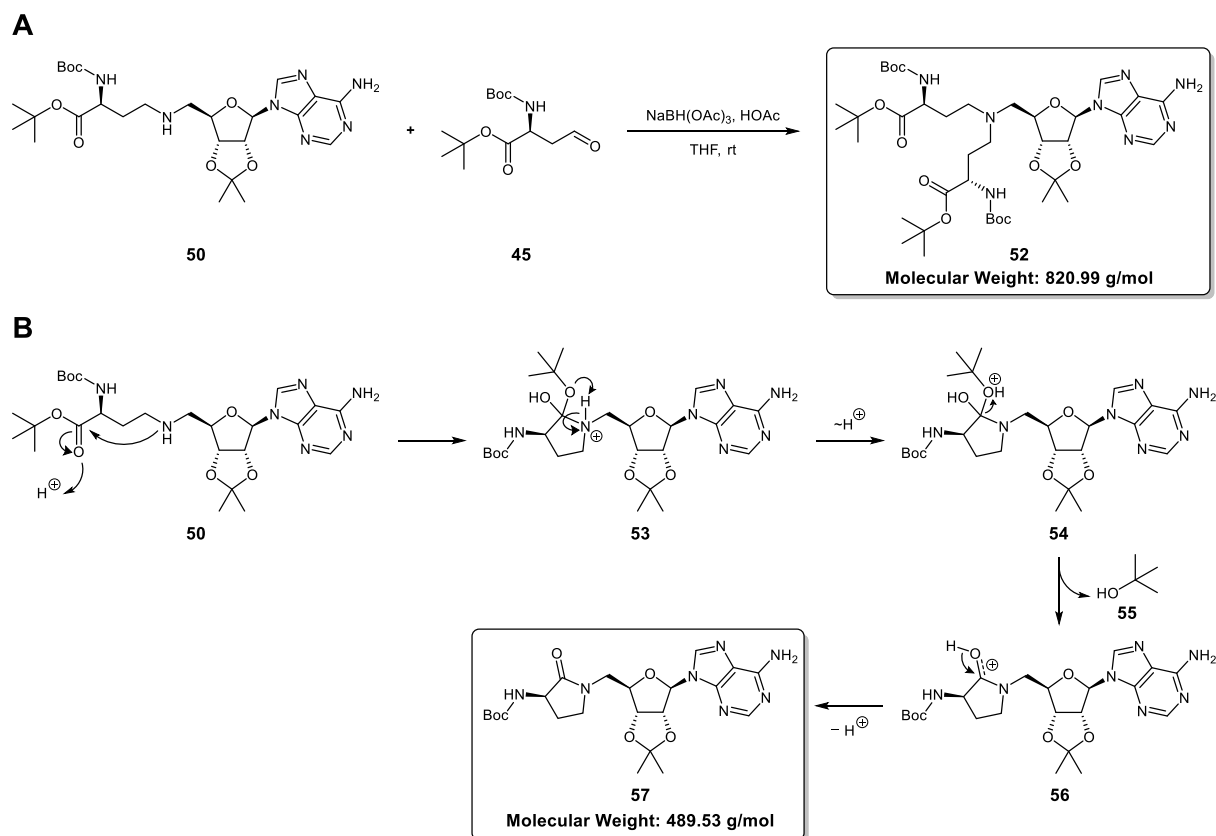
Scheme 16: A: formation of the double alkylated byproduct **52**. B: postulated mechanism for the formation of the cyclic byproduct **57** in the reductive amination procedure.

By conducting the reductive amination at 0–5 °C in a mixture of THF and MeCN, the formation of the cyclic byproduct could be largely suppressed (▮▮▮▮▮▮▮▮▮, master thesis, under supervision of MARVIN SCHWICKERT, ▮▮▮▮▮▮▮ group).

To cleave alle protecting groups of the precursor **50**, a slightly modified deprotection method by ZHANG *et al.* was tested.[279] For this purpose, the compound was treated with TFA/H$_2$O (95:5, *V*/*V*) at room temperature (**Scheme 17**). Although the desired product **51** could be detected, the procedure largely led to decomposition of the compound. A method described by DOWDEN *et al.* using a lower amount of water (2.5 vol%) yielded a similar result.[280]

Since the deprotection methods described in the literature showed unsatisfying results, an alternative procedure under mild conditions was designed. Instead of cleaving all three protecting groups in one step, a two-step one-pot procedure was utilized. The Boc and *tert*-butyl groups were first removed by treating the compound with 50 vol% TFA in DCM at 5 °C. After this, TFA was removed by codestillation with DCM. Finally, 14 vol% TFA in water was used at 5 °C to cleave the acetal protecting group. Lyophilization yielded the final product in high purity (>95%).
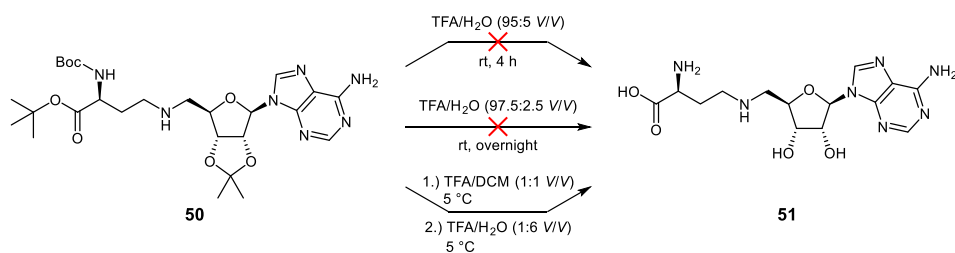
Scheme 17: Attempts to cleave the protecting groups of the precursor **50**.

3.3 Variation of side chain by parallel synthesis

To quickly generate a SAH-like inhibitor library with varying side chains, a parallel synthesis was established using a *Synthesis 1 parallel synthesizer* from *Heidolph Instruments GmbH & CO. KG*. The experiments were carried out by ████████████ (master thesis, under supervision of MARVIN SCHWICKERT, ████████ group). The 5'-amino building block **46** was treated with different in-house acyl and sulfonyl chlorides in DCM using triethylamine (**Scheme 18**).
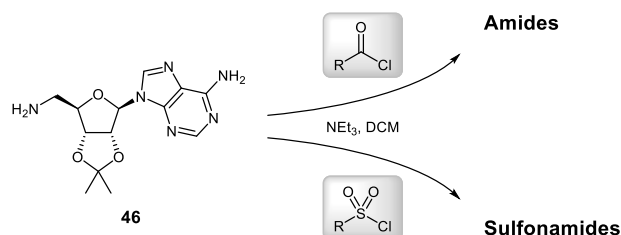
Scheme 18: General reactions used for parallel synthesis to rapidly generate a molecule library of adenosine derivatives.

A detailed depiction of the workflow is shown in **Scheme 19**. The reactants were mixed for 16 h at room temperature. To allow complete conversion of the acyl/sulfonyl chloride, the amine was used in excess (1.11 equiv.). The organic phase was then washed five times with a saturated $NaHCO_3$ solution and two times with water, which completely removes the residual amine **46**. In the following step, the organic solvent was removed by distillation. The residue was treated with either an aqueous HCl solution (33 vol%) or an aqueous TFA solution (14 vol%) for 16 h at room temperature. After lyophilization, the final product was obtained as the respective hydrochloride or trifluoroacetate salt.
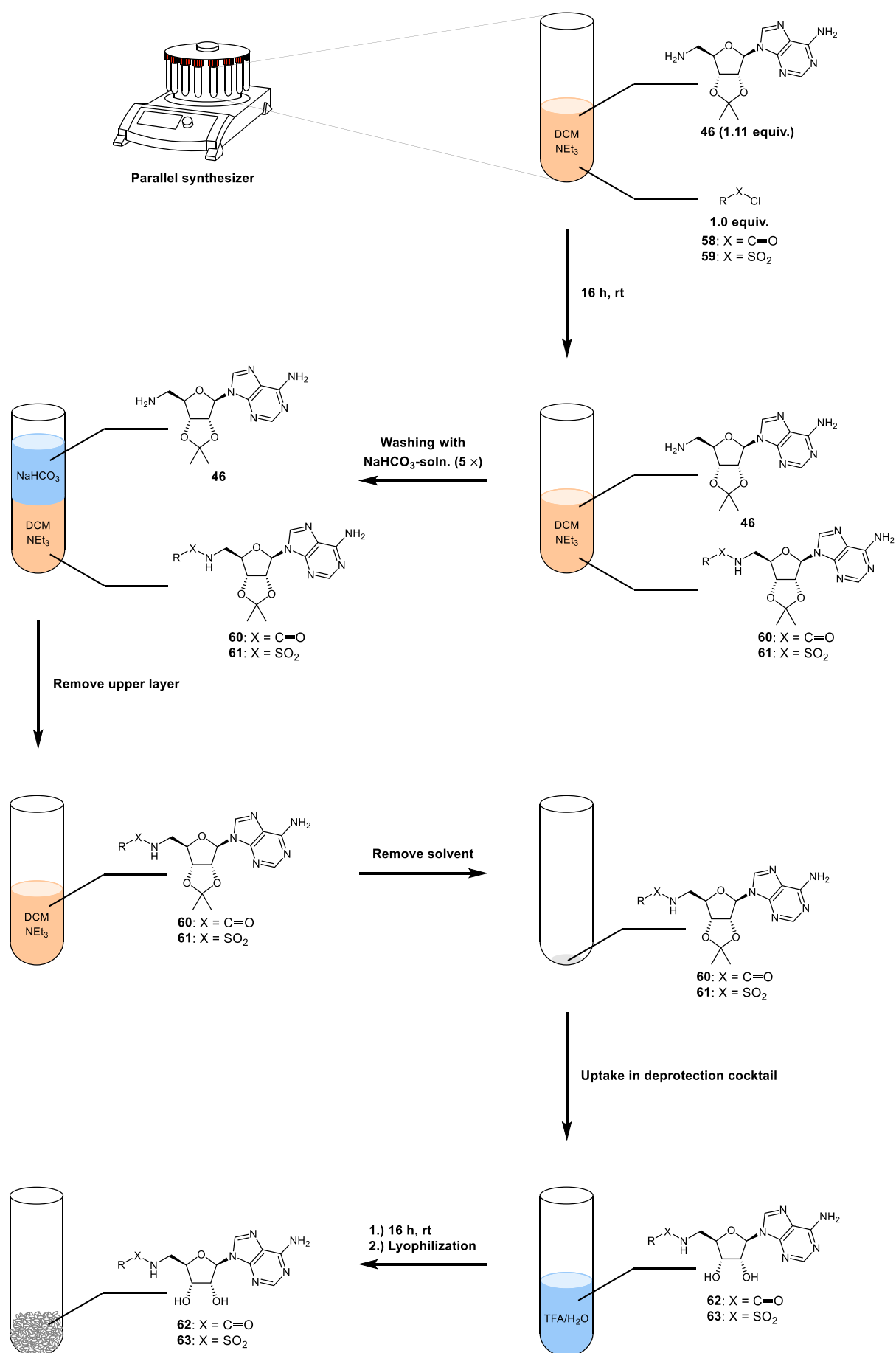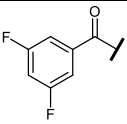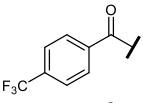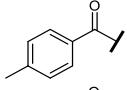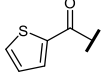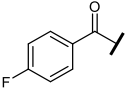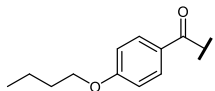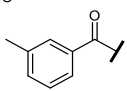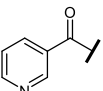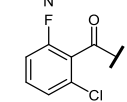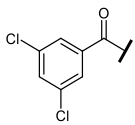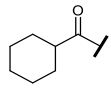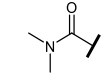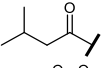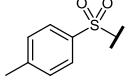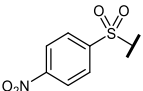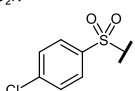
Scheme 19: Workflow of the parallel synthesis.

The structures of the compounds obtained with the parallel synthesis method are shown in **Table 2**. Yields and purities are given.

Table 2: Compounds obtained via parallel synthesis with yields and purities given.
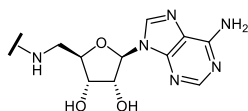
62a–m, 63a–c

Compound		Deprotection method	Yield	Purity
62a		HCl/water, rt	99%	98%
62b		HCl/water, rt	99%*	72%
62c		HCl/water, rt	99%	99%
62d		HCl/water, rt	99%	>99%
62e		HCl/water, rt	88%	95%
62f		TFA/water, 5 °C	48%	97%
62g		TFA/water, 5 °C	99%	96%
62h		TFA/water, 5 °C	77%	96%
62i		TFA/water, 5 °C	99%	96%
62j		TFA/water, 5 °C	99%	96%
62k		TFA/water, 5 °C	99%	95%
62l		TFA/water, 5 °C	67%	95%
62m		TFA/water, 5 °C	85%	98%
63a		TFA/water, 5 °C	99%	98%
63b		TFA/water, 5 °C	95%	98%
63c		TFA/water, 5 °C	98%	97%

*In this case, the value refers to conversion and not yield.

Table 2 shows that the parallel synthesis procedure yielded the compounds in moderate to high yields. 15 out of 16 compounds showed high purity of ≥95% without further purification steps. Although 99% conversion was achieved for compound **62b**, a second peak was observed in the LC-MS purity measurement, which could be assigned to adenine cleaved using the deprotection procedure consisting of HCl/water. Areas under the curves (AUC) revealed a ratio of 76:23 (**62b**/adenine).

The compounds were tested using MST, DSF, and a tritium incorporation assay. MST measurements were conducted by ███████████ (under supervision of ███████████████, ███████████ group). DSF measurements were also conducted by ███████████ (under supervision of MARVIN SCHWICKERT, ███████████ group). The measurements of the tritium incorporation assay were carried out by ███████████ and ███████████ (███ group). All results are listed in **Table 3**.

The results obtained with MST and DSF (columns 2 and 3 in **Table 3**) are fairly consistent, with the MST showing higher sensitivity in identifying binders, which might be due to the higher concentrations (250 μM vs. 100 μM) of the compounds used in this assay. Two exceptions, however, could be seen in the case of the 4-methylbenzoyl derivative **62c** and the nicotinoyl derivative **62h**, which were classified as non-binders by MST but as binders by DSF. Interestingly, the 4-methylbenzoyl derivative **62c** induced a melting point decrease, which might be due to the stabilization of a conformation with lower melting point or the disruption of enzyme stabilizing interactions. Although 9 out of 16 compounds have been identified as binders (MST), no inhibition could be detected in the tritium incorporation assay. For this reason, none of these structures were considered for further optimization.

Nevertheless, this parallel synthesis method represents a suitable basis to rapidly generate linear SAH or adenosine derivatives. Even though the compound classes provided with this method did not seem to be suitable for the inhibition of DNMT2, such scaffolds represent promising inhibitors as they have been used extensively for other targets,[230,281,282] most recently for inhibition of the SARS-CoV-2 nsp14 *N*7-methyltransferase.[283]

For a more detailed discussion, as well as the presentation of the experimental procedures for the synthesis of compounds **62a–m** and **63a–c**, the reader is referred to the corresponding master thesis:

███████████, *DNMT2-Inhibitoren als Tools für biophysikalische Anwendungen*, **2020**, Johannes Gutenberg University Mainz.

Own contribution: Design of the parallel synthesis method and inhibitors, supervising the master thesis about the conduction of parallel synthesis of compounds **62a–m** and **63a–c** as well as the testing using DSF.

A manuscript about the procedure is in preparation and will be submitted to *ACS Journal of Combinatorial Sciences*.

Table 3: Results of compounds **62a**–**m** and **63a**–**c**.

62a–m, 63a–c
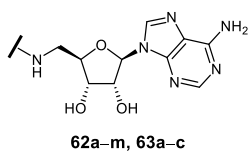
Compound		MST Binding at 250 μM	DSF Binding at 100 μM	^{3}H-Assay % inhibition at 100 μM
62a		✓	✗	n. i.
62b		✓	✗	n. i.
62c		✗	✓*	n. i.
62d		✗	✗	n. i.
62e		✗	✗	n. i.
62f		✗	✗	n. i.
62g		✓	✓	14 ± 11 (n. i.)
62h		✗	✓	n. i.
62i		✓	✗	n. i.
62j		✗	✗	n. i.
62k		✓	✓	n. i.
62l		✓	✗	n. i.
62m		✗	✗	11 ± 7.1 (n. i.)
63a		✓	✗	n. i.
63b		✓	✓	n. i.
63c		✓	✓	n. i.

*Instead of an increase, a decrease in melting point was observed; n. i. = no inhibition.

3.4 Variation of side chain by rational design and development of Y-shaped inhibitors

Note: The substance numbers from the corresponding publication are listed *in italics* behind the substance numbers in this thesis.

3.4.1 Summary and own contribution

As an alternative approach to the parallel synthesis, rational design was used to vary the amino acid side chain of SAH. Based on this, three different compound series were designed and prepared to analyze their effects in a structure-activity relationship study (**Figure 14**):
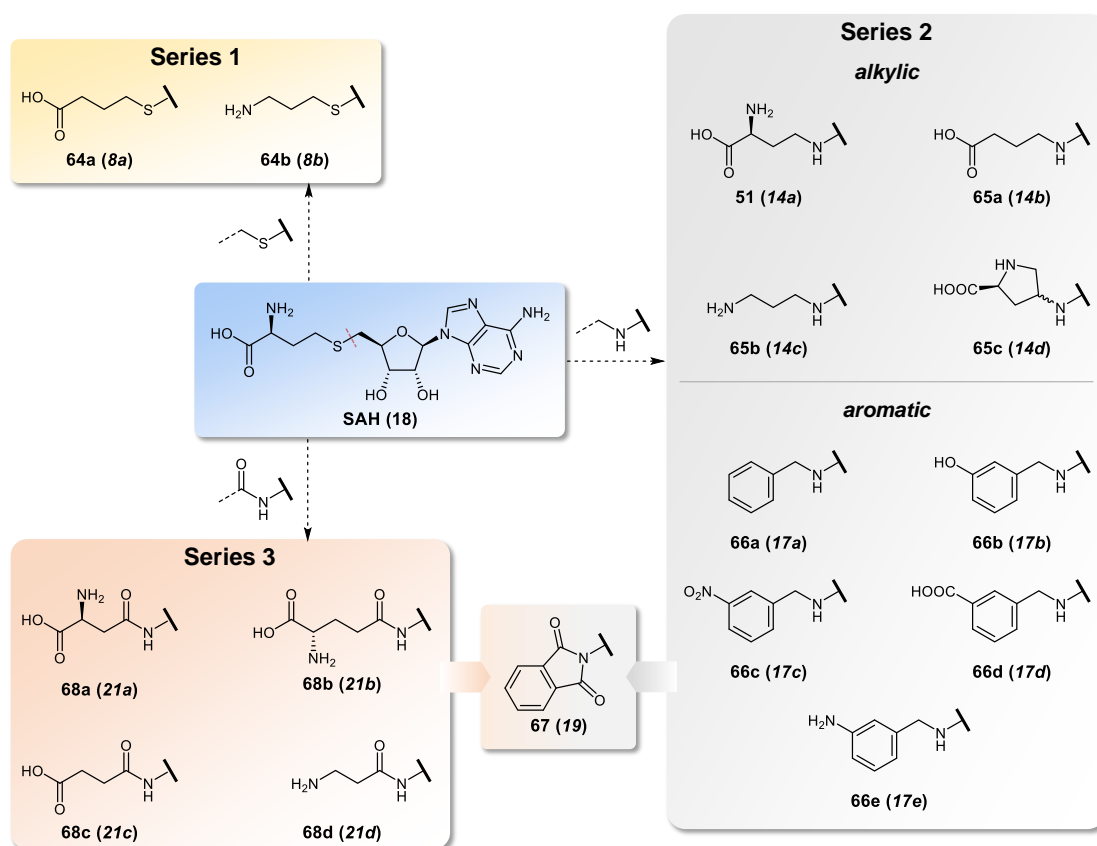
Figure 14: Variations of the SAH amino acid side chain.

1. To analyze the pharmacophoric significance of the amino and carboxylic acid groups, SAH analogs lacking either of these moieties were synthesized (████████████, master thesis, under supervision of ████████, ████ group and MARVIN SCHWICKERT, ████████ group).

2. The homocysteine's sulfur atom was exchanged with nitrogen allowing an additional side chain to be attached. Additionally, it introduces a basic functionality that can be protonated depending on the pH value. The resulting positive charge may imitate the sulfonium function found in SAM. Analogously to the compounds of series 1, two compounds were synthesized, each lacking one of the side chain functional groups. To constrain the alkyl side chains flexibility, a five-membered ring was

implemented. As an alternative rigidizing element, derivatives with benzene-based side chains were developed with and without polar functions.

3. An amide function was investigated as a replacement of the thioether group. It represents an alternative nitrogenous but non-basic and rigid function allowing side chain attachment. Structures lacking either the amine or carboxylic acid group were synthesized. Additionally, a derivative with an extended aliphatic side chain was generated since the amide functions yield a shorter bond length compared to thioether- or amine-based linkers.

The compounds of series 1–3, SAH, and SFG were tested for binding to human DNMT2 in a microscale thermophoresis assay established by ▆▆▆▆▆▆▆▆▆▆▆ (▆▆▆▆▆▆▆ group). Inhibition was measured by ▆▆▆▆▆▆ (▆▆▆ group) using a tritium incorporation assay. Besides SAH and SFG, N-adenosyl-(S)-2,4-diaminobutanoic acid (adenosyl-Dab) **51** was identified as a binder of DNMT2, indicating that both functionalities of the amino acid side chain are crucial for binding. However, inhibition was only found for SAH (IC$_{50}$ = 15.8 ± 1.5 μM) and SFG (IC$_{50}$ = 13.2 ± 0.8 μM). Although adenosyl-Dab did not exhibit inhibition, the structure was considered for further optimizations as the replacement of sulfur with nitrogen enables side chain attachment to target the cytidine site of DNMT2. A variety of N-alkyl substituents were introduced to generate Y-shaped inhibitors (**Figure 15**), starting from small modifications such as methyl or ethyl, to branched alkyl chains (e.g. isobutyl, neopentyl). Furthermore, unsaturated systems such as benzyl, phenethyl, and various alkyne-based moieties were considered as side chain modifications.
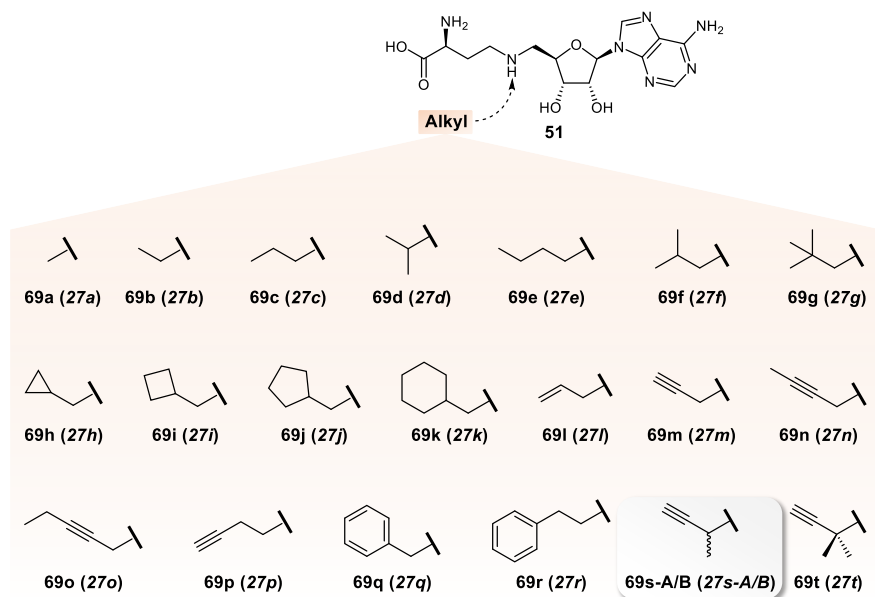
Figure 15: Different N-alkyl substituents attached to the adenosyl-Dab scaffold (**51**).

Out of the series of N-alkylated SAH derivatives, the testing revealed a compound with a (SR)-N-but-3-yn-2-yl group (**69s-A/B** (**27s-A/B**)) as the most active (IC$_{50}$ = 12.9 ± 1.9 μM), exhibiting similar potency as the natural ligands SAH and SFG. Both the S- and R-epimers were found to be equipotent.

Furthermore, **69s-A/B** (*27s-A/B*) was selective toward NSUN2 and NSUN6 as it did not show inhibition in the tritium incorporation assay at 100 μM. Based on this compound, an ethyl ester prodrug (*28-A/B*) was prepared to test it in a cellular context. The results showed very low but significant inhibition of tRNA methylation in cells by the ester prodrug but not by the acid. Moreover, CaCo-2 assays revealed that the prodrug exhibited poor membrane permeability due to rapid hydrolysis, which was further confirmed by stability testing using LC-MS.

For a more detailed discussion, as well as the presentation of the experimental procedures, the reader is referred to the corresponding publication (**Section 3.4.2**) and Supporting Information (**Appendix**) (**DOI: 10.1021/acs.jmedchem.2c00388**).

Own contribution: Design and synthesis of inhibitors, supervising a master thesis (█████████████████) about the synthesis of **64-a** (*8a*) and **64-b** (*8b*), supervising a master thesis (████████████) about the synthesis of Y-shaped inhibitors, synthesis and separation of diastereomers **69s-A** (*27s-A*) and **69s-B** (*27s-B*), synthesis of prodrug (*28-A/B*) and stability testing by LC-MS, analysis of CaCo-2 assays by LC-MS, writing the corresponding parts in the manuscript and Supporting Information, writing the introduction of the manuscript, creation of graphical abstract, **Figure 1**, **Figure 2**, **Figure 6**, **Schemes 1–7**.

3.4.2 Publication

Discovery of Inhibitors of DNA Methyltransferase 2, an Epitranscriptomic Modulator and Potential Target for Cancer Treatment

Published as part of the Journal of Medicinal Chemistry special issue "Epigenetics 2022".

Marvin Schwickert,[‡] Tim R. Fischer,[‡] Robert A. Zimmermann,[‡] Sabrina N. Hoba, J. Laurenz Meidner, Marlies Weber, Moritz Weber, Martin M. Stark, Jonas Koch, Nathalie Jung, Christian Kersten, Maike Windbergs, Frank Lyko, Mark Helm,* and Tanja Schirmeister*

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Selective manipulation of the epitranscriptome could be beneficial for the treatment of cancer and also broaden the understanding of epigenetic inheritance. Inhibitors of the tRNA methyltransferase DNMT2, the enzyme catalyzing the *S*-adenosyl-methionine-dependent methylation of cytidine 38 to 5-methylcyti-dine, were designed, synthesized, and analyzed for their enzyme-binding and -inhibiting properties. For rapid screening of potential DNMT2 binders, a microscale thermophoresis assay was established. Besides the natural inhibitors *S*-adenosyl-L-homocysteine (SAH) and sinefungin (SFG), we identified new synthetic inhibitors based on the structure of *N*-adenosyl-2,4-diaminobutyric acid (Dab). Structure−activity relationship studies revealed the amino acid side chain and a Y-shaped substitution pattern at the 4-position of Dab as crucial for DNMT2 inhibition. The most potent inhibitors are alkyne-substituted derivatives, exhibiting similar binding and inhibitory potencies as the natural compounds SAH and SFG. CaCo-2 assays revealed that poor membrane permeabilities of the acids and rapid hydrolysis of an ethylester prodrug might be the reasons for the insufficient activity in cellulo.
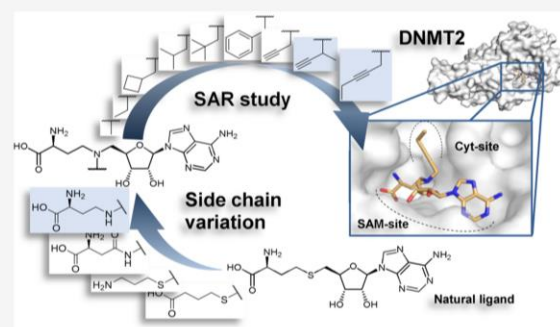
■ INTRODUCTION

Epigenetic inheritance is not only mediated by modifications of DNA and histones but also driven by RNA and RNA modifications.[1−6] Various data suggest that RNA species are involved in the heredity of specific phenotypes such as mental[7−10] or metabolic disorders.[4,11−15] As an example, metabolic disorders and their epigenetic transmission were found to be linked to elevated levels of m[2]G and m[5]C modifications on small noncoding RNAs (sncRNAs) in mouse models,[4] and it could be shown that deletion of the enzyme responsible for these modifications, namely, the DNA methyltransferase 2 (DNMT2), abolished sperm sncRNA-mediated transmission of high-fat-diet-induced metabolic disorders to descendants.[11] The enzyme DNMT2 is also essential for the epigenetic transmission of phenotypes linked to the *Kit* and *Sox9* genes between subsequent generations of mice. These genes encode for a tyrosine kinase and a transcription factor, variants of which lead to white colored tails and feet (*Kit*) or to an enhanced growth (*Sox9*) of *Dnmt2*[+/+] mice but not *Dnmt2*[−/−] mice.[16]

According to its sequence and structure, human DNMT2 is a member of the DNA methyltransferase (MTase) family.[17,18]

Also, the catalytic motif is strongly conserved and highly characteristic for DNA methyltransferases. In contrast to DNMT1 and DNMT3, however, DNMT2 is primarily localized in the cytoplasm instead of the nucleus and therefore lacks the properties of a typical DNA methylating enzyme.[19] Although it was discussed that DNMT2 also modifies DNA at CG residues,[20] its main function is the methylation of tRNA[Asp] at position C38 in the anticodon loop.[19] Further substrates are tRNA[Gly], tRNA[Val], and tRNA[Glu], depending on the species.[21]

Besides its role in epigenetics, DNMT2 is involved in other physiological processes, some of which are not quite understood so far. While induced loss of DNMT2 reduced the size of zebrafish morphants and affected retina, liver, and brain development,[22] no morphological effects could be observed in

Figure 1. Modification of tRNA by DNMT2 with SAM as a cosubstrate. 5-Azacytidine and sinefungin are inhibitors of DNMT2. Zebularine is a known DNA MTase inhibitor and is supposed to exhibit an analogous mechanism as 5-azacytidine as it has been shown to be incorporated into tRNA.[39] Its inhibitory potential against DNMT2 has not been reported yet. However, the M.*Hha*I C5 Mtase, which shows close similarity to DNMT2, is a known target of zebularine.[38]

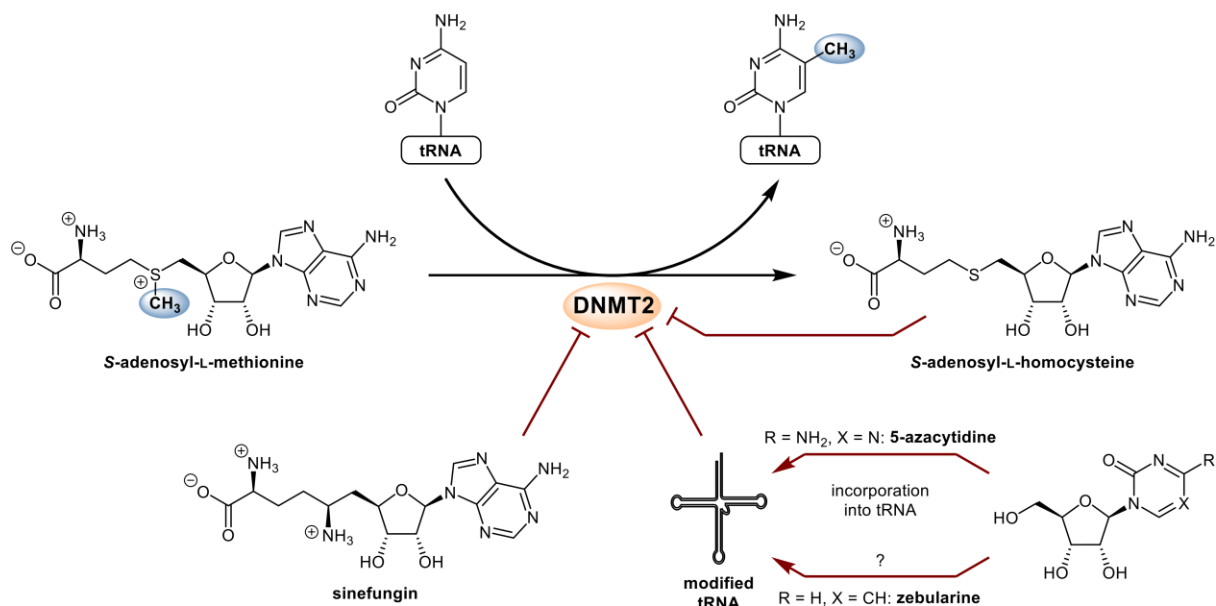flies, mice, and plants.[19] An overexpression of DNMT2 in *Drosophila* flies led to life span prolongation and an increased stress resistance.[23] This is substantiated by the fact that *Drosophila* loss-of-function DNMT2 mutants showed reduced viability under stress conditions.

Methylation of tRNA by DNMT2 prevents ribonuclease-mediated cleavage of tRNA,[24] thus regulating RNA stability and correspondingly leading to higher rates of overall protein synthesis.[25] Moreover, the methylation of C38-tRNA^{Asp} by DNMT2 is required to ensure accurate polypeptide synthesis by improving codon recognition. Methylation increases the capacity of tRNA to discriminate between Asp and Glu codons improving translational accuracy, which prevents the production of aberrant proteins.[26]

There are also indications that DNMT2 plays a role in malignancies as it is overexpressed in several cancer tissues, such as cervical[27] or bladder tissue.[28] According to the COSMIC database, various tumor samples showed upregulation of DNMT2 and numerous somatic mutations in the enzyme were found in tumors originating from different tissues.[29]

Taken together, DNMT2 represents a potential target in cancer treatment and epigenetic drug discovery.

In 2001, Dong et al. revealed the crystal structure of human DNMT2 (Protein Data Bank (PDB) entry 1G55) in 1.8 Å resolution. Consisting of 391 amino acids corresponding to a size of ca. 40 kDa, the enzyme is a relatively small protein.[30] A comparison to DNMT1 and DNMT3 illustrates that in contrast to other eukaryotic MTases DNMT2 lacks the large N-terminal domain.[17] The conserved C-terminal catalytic domain of DNMT2 features a unique cysteine-phenylalanine-threonine (CFT) motif between the catalytic motifs VIII and IX, which is not found in other MTases.[30,31] In the catalytic motif IV, the protein contains the common proline-cysteine-glutamine (PCQ) motif including the catalytically active cysteine. However, in the crystal structure, this exact part of the catalytic

center is disordered.[30] A more defined structure of the loop can be found in the DNMT2 homologue of *Entamoeba histolytica*, which shows an α-helical conformation.[32]

DNMT2 requires *S*-adenosyl-L-methionine (SAM) as cofactor, transferring SAM's methyl group to the 5-position of the tRNA substrate's cytidine, yielding m5C38-tRNA^{Asp} and *S*-adenosyl-L-homocysteine (SAH) (Figure 1).[19]

Currently, only a few DNMT2 inhibitors have been reported. The anticancer drugs 5-azacytidine (5-azaC) and 5-aza-2′-deoxycytidine (decitabine) inhibit RNA and DNA MTases (Figure 1) leading to a change of the methylation activity in cancer patients.[33−35] During transcription and replication, the drugs are randomly incorporated into nascent DNA (decitabine and 5-azaC) and RNA (5-azaC) by respective polymerases. Due to the substitution of cytidine's carbon atom in position 5 with nitrogen in decitabine and 5-azaC, DNA and RNA MTases remain covalently bound to the target nucleic acids, thus inhibiting the catalytic activity of the enzymes.[36,37] A similar mechanism of inhibition is exhibited by the pyrimidinone derivative zebularine. It forms a covalent complex with DNA MTases, such as C5 Mtase from *Haemophilus hemolyticus* (M.*Hha*I), which shows close similarity to human DNMT2.[38] However, inhibition of DNMT2 itself has not been reported yet. Due to the structural similarity to 5-azaC, an analogous mechanism of inhibition by zebularine is supposed as incorporation into tRNA has been found.[39]

The natural SAM-related nucleoside sinefungin (adenosyl-ornithine, SFG),[40] originally isolated from *Streptomyces griseolus*, exhibits competitive inhibition of several SAM-dependent MTases.[41] However, the inhibitory potential toward DNMT2 has not been quantified yet. Similarly, the inhibition of DNMT2 by SAH, which has been reported to be a nonselective feedback inhibitor of several SAM-dependent MTases, has not been investigated in detail.[42]
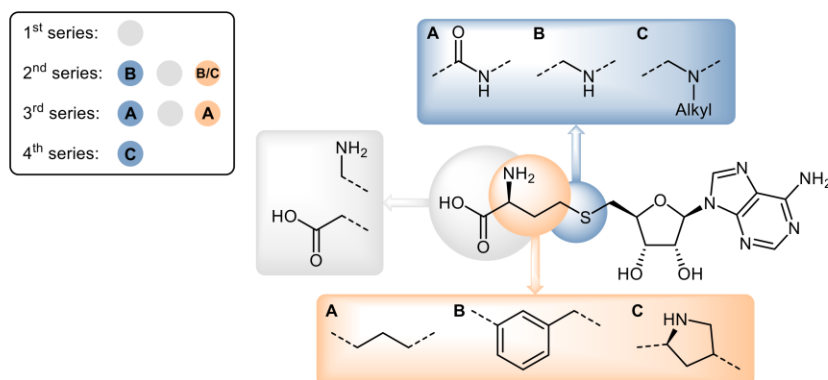
Figure 2. Inhibitor design based on the SAH scaffold.

Scheme 1. Syntheses of Thioethers 8a, 8b: (A) Synthesis of Alkyl Bromides 2 and 4; (B) Synthesis of Acetyl-Protected Thiol Building Block 6 Followed by Nucleophilic Substitution and Deprotection[a]
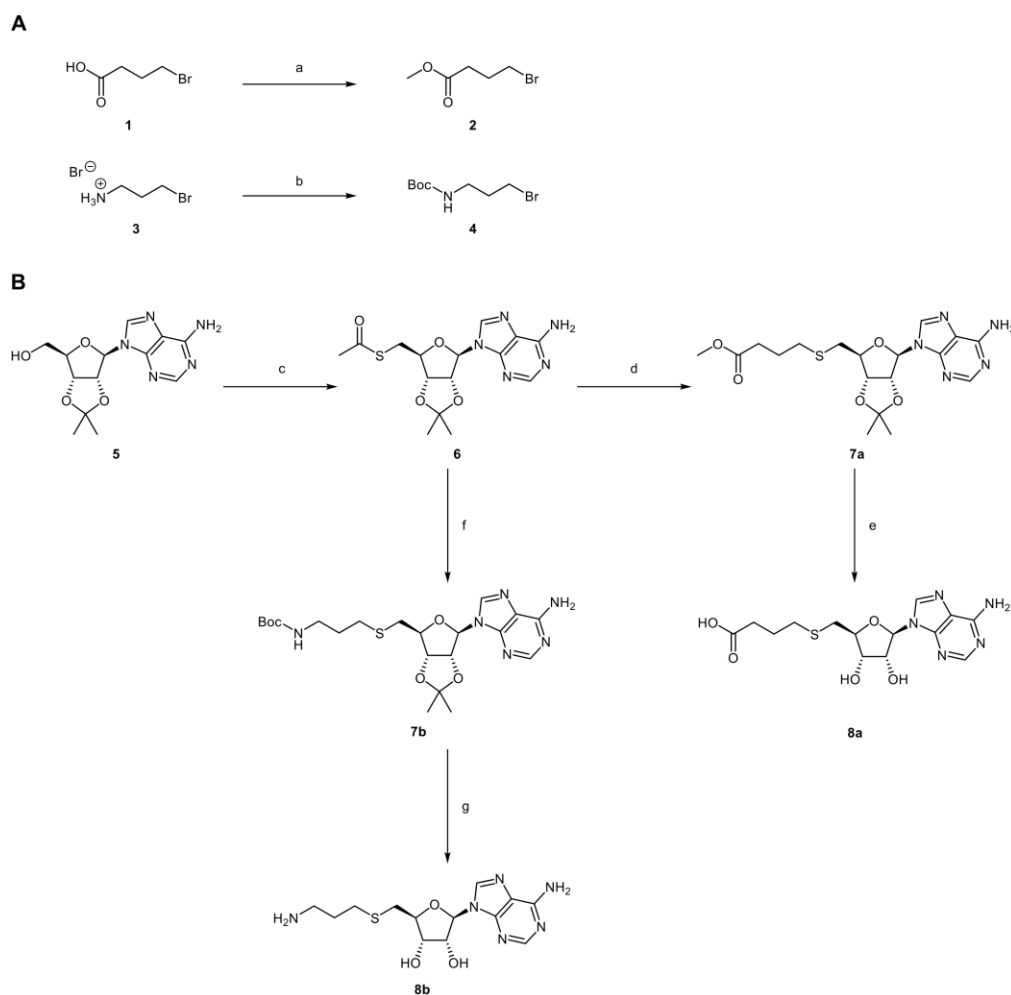
[a]Reagents and conditions: (a) MeOH, SOCl$_2$, rt, 16 h, 79%; (b) Boc$_2$O, THF, NEt$_3$, rt, 16 h, 70%; (c) PPh$_3$, DIAD, AcSH, THF, 0 °C, 18 h, 95%; (d) **2**, NaOMe, MeOH, rt, 16 h, 75%; (e) (1) LiOH, THF/H$_2$O, rt, 1 h; (2) DCM/TFA (1:1), H$_2$O, 5 °C, 1 h, 96%; (f) **4**, NaOMe, MeOH, rt, 16 h, 84%; (g) HCOOH, H$_2$O, 0 °C, 2 d, 96%.

Article

Scheme 2. Syntheses of Aliphatic Amine Derivatives: (A) Synthesis of Aldehyde Building Blocks 11a−c and Structure of the Commercially Available Ketone 11d; (B) Reductive Amination and Deprotection to Yield the Amines 14a−d[a]

[a]Reagents and conditions: (a) *N,O*-dimethylhydroxylamine, CDI, DCM, rt, 18−25 h, 81−89%; (b) DIBAL, THF, −78 °C, 2−2.5 h, 34−81%; (c) **11a−c**, NaBH(OAc)$_3$, HOAc, THF, rt, 5−72 h, 57−76%; (d) (1) TFA/DCM 1:1, −20 °C; (2) TFA/H$_2$O 1:6, 8 °C, 99%; (e) **11d**, NaBH(OAc)$_3$, HOAc, THF, rt, overnight, 68%; (f) (1) TFA/DCM 1:1, −20 °C; (2) TFA/H$_2$O 1:6, 8 °C, 99%.
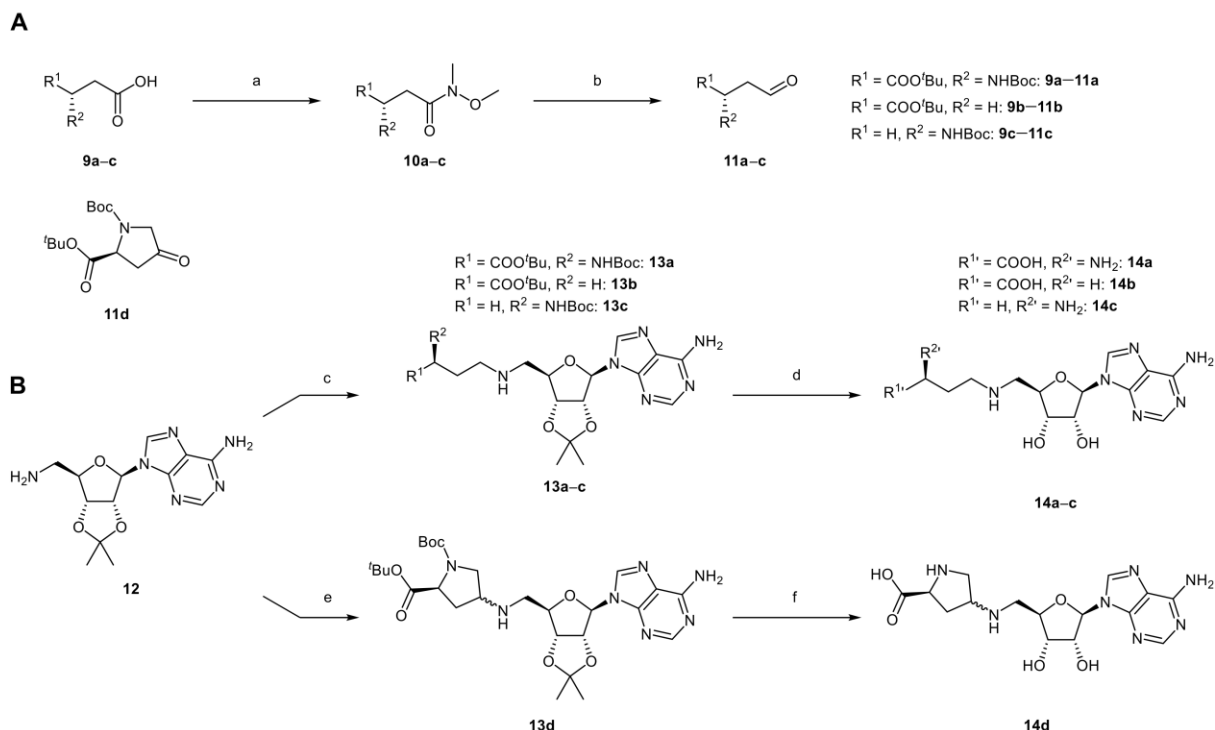
As a successful strategy in the search for inhibitors or tools for study of SAM-dependent MTases, the derivatization of the natural binder SAH has been exploited for various targets such as the catechol-*O*-MTase (COMT),[43−45] protein MTases (PMTs) including histone MTases (HMTs) such as DOT1L,[46] RNA MTases such as METTL3,[47] and the SARS-CoV-2 mRNA cap MTase[48,49] or the previously mentioned DNA MTases DNMT1 and DNMT3B2.[50,51] The DOT1L inhibitor pinometostat is currently being tested in clinical trials.[52] For DNMT2, synthetic inhibitors based on SAH were previously described, albeit only with very low affinity.[53] However, a systematic approach to identify and improve such inhibitors or tools has not yet been followed so far.

In this work, we present the discovery of potent DNMT2 inhibitors based on SAH's adenosyl scaffold. While this nucleoside substructure was retained, we performed structure−activity relationship studies by investigating different side-chain types (Figure 2). We prepared four compound series:

(1) First, SAH analogs lacking either the amino or the acid group (**8a**, **8b**) were tested to identify the pharmacophoric significance of these moieties.

(2) In order to extend the possibilities for substitution, we exchanged the homocysteine's sulfur atom with nitrogen yielding *N*-adenosyl-2,4-diaminobutyric acid (adenosyl-Dab) derivatives (**14a−d**, **17a−e**). This additionally introduces a basic functionality, which—depending on the pH value—can be protonated but does not provide a permanently charged group as found in SAM. Analo-

gously to the sulfide derivatives, we generated two compounds, each of which lacks one of the functional groups in the side chain (**14b**, **14c**). To constrain the flexibility of the aliphatic side chain, we implemented a five-membered ring (**14d**). In addition, we developed derivatives with benzene as a rigidizing element, substituted with and without polar functions (**17a−e**).

(3) We also investigated the effect of an amide function as a replacement of the thioether group (**21a−d**). To enable a comparison of the sulfide and amine derivatives, we also created structures lacking either the amine or the carboxylic acid group of the amino acid side chain (**21c**, **21d**). Since the amide function yields a shorter bond length compared to thioether- or amine-based linkers, we generated a structure with an extended aliphatic side chain (**21b**).

(4) In order to target both the SAM- and cytidine-binding site, we designed Y-shaped structures (**27a−27t**).

The compounds were tested for binding to DNMT2 using microscale thermophoresis (MST). Both DNMT2 binders and nonbinders were evaluated with regard to their inhibitory effect on DNMT2-catalyzed methylation of substrate tRNA^{Asp} in a tritium incorporation assay. This procedure was also followed in order to investigate if DNMT2 binding directly resulted in DNMT2 inhibition and if the MST binding assay can be used as a prescreening method prior to the laborious, time-consuming, and expensive radioactive assay.

Scheme 3. Syntheses of Aromatic Amine Derivatives 17a–e[a]

R = H: **15a**
R = OH: **15b**
R = NO$_2$: **15c**
R = COO^{t}Bu: **15d**
R = NHBoc: **15e**

R = H: **17a**
R = OH: **17b**
R = NO$_2$: **17c**

R = H: **16a**   R = COO^{t}Bu: **16d**
R = OH: **16b**   R = NHBoc: **16e**
R = NO$_2$: **16c**

R = COOH: **17d**
R = NH$_2$: **17e**
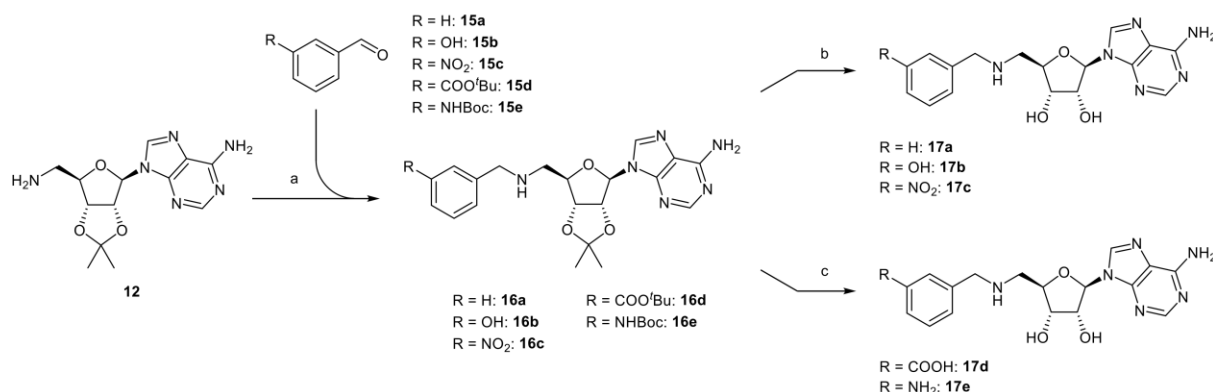
[a]Reagents and conditions: (a) **15a–e**, NaBH(OAc)$_3$, HOAc, THF or 1,2-DCE/MeCN, rt, 3.5–24 h, 55–92%; (b) TFA/H$_2$O 1:6, 5 °C, 91–99%; (c) (1) TFA/DCM 1:1, −20 °C; (2) TFA/H$_2$O 1:6, 5 °C, 97–99%.

Scheme 4. Syntheses of Amide Derivatives 21a–d and Phthalimide 19[a]

R^1 = COO^{t}Bu, R^2 = NHBoc, n = 1: **20a**
R^1 = COO^{t}Bu, R^2 = NHBoc, n = 2: **20b**
R^1 = COO^{t}Bu, R^2 = H, n = 1: **20c**
R^1 = H, R^2 = NHBoc, n = 1: **20d**

R^1 = COOH, R^2 = NH$_2$, n = 1: **21a**
R^1 = COOH, R^2 = NH$_2$, n = 2: **21b**
R^1 = COOH, R^2 = H, n = 1: **21c**
R^1 = H, R^2 = NH$_2$, n = 1: **21d**
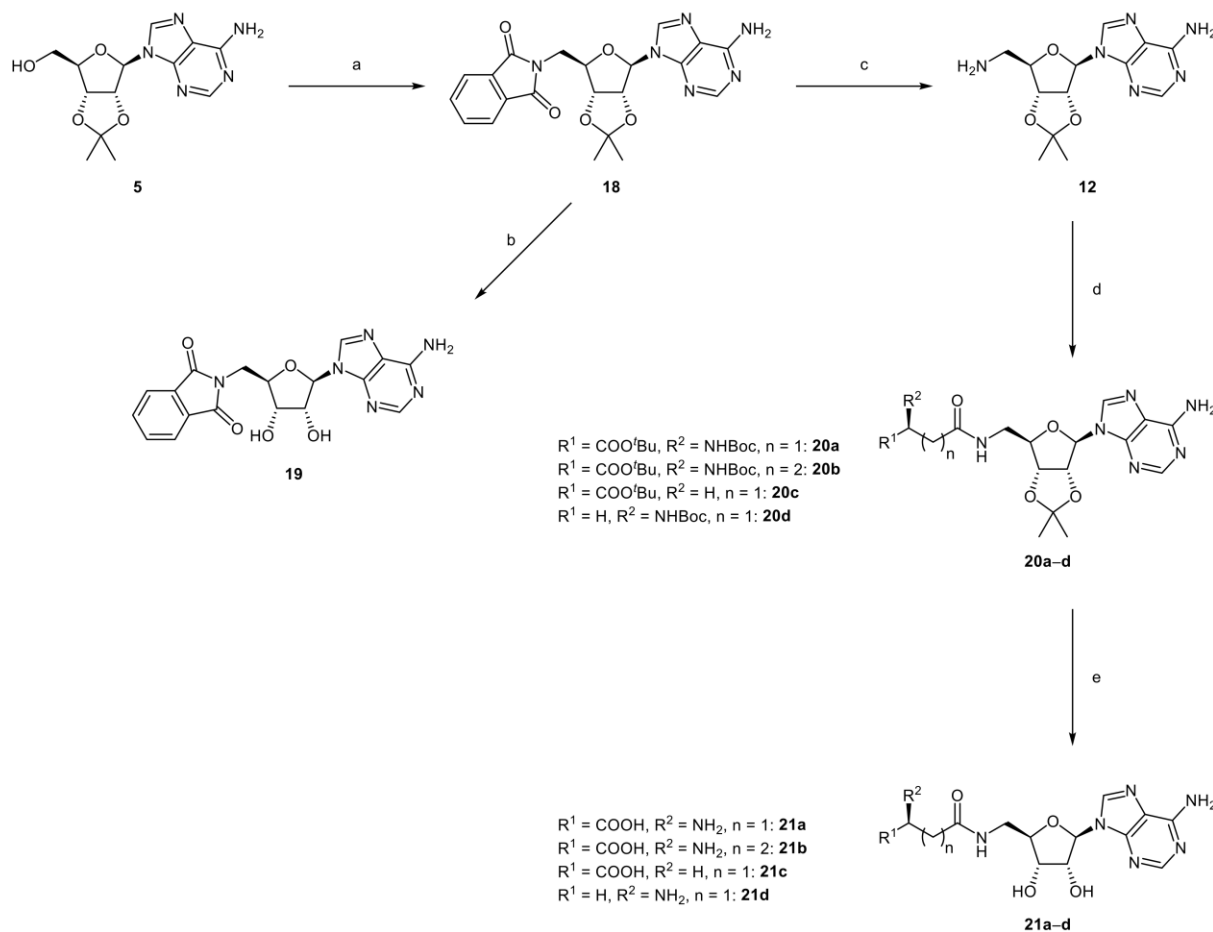
[a]Reagents and conditions: (a) phthalimide, PPh$_3$, DIAD, THF, rt, 5 h, 71%; (b) TFA/H$_2$O 1:6, 8 °C, 2 h, 96%; (c) hydrazine, EtOH, Δ, 7 h, 93%; (d) R^1−CHR^2−(CH$_2$)$_n$−COOH, TBTU, DIPEA, DMF, rt, 2.5–30 h, 76–97%; (e) (1) TFA/DCM 1:1, −20 °C; (2) TFA/H$_2$O 1:6, 8 °C, 99%.

For the most potent methylation inhibitors, IC$_{50}$ values were determined in the tritium incorporation assay, and K_D values were determined using isothermal titration calorimetry (ITC).

In our studies, we identified novel SAH-based DNMT2 binders and inhibitors with low micromolar affinity.

Scheme 5. Synthesis of Y-Shaped Methyl Derivative 27a[a]
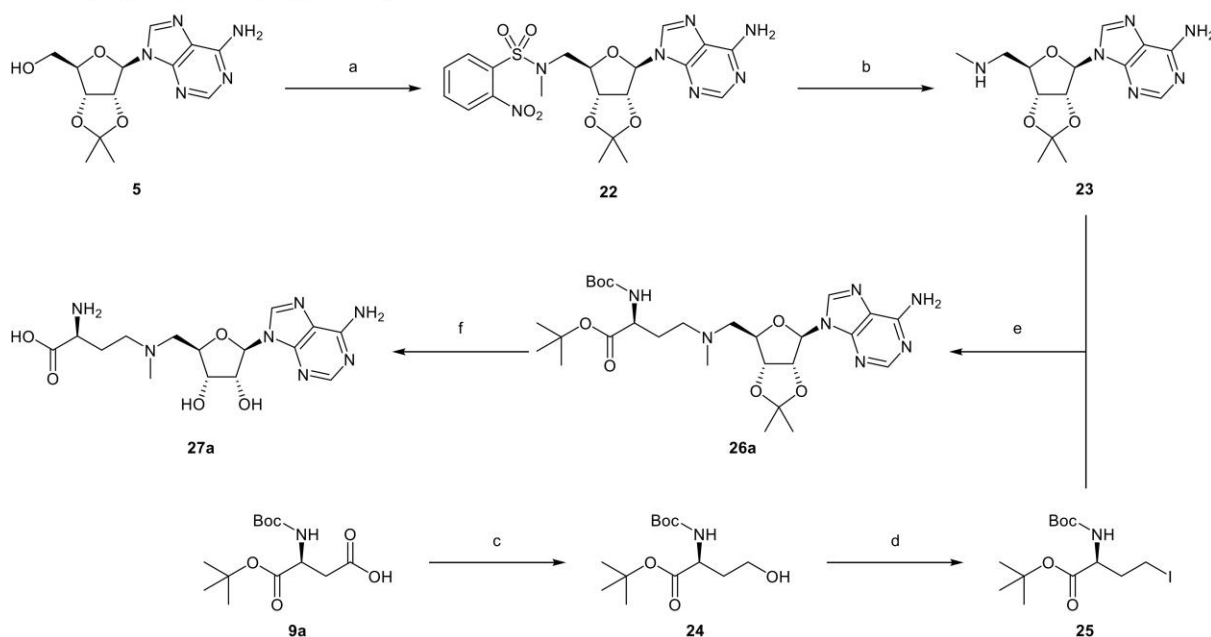
[a]Reagents and conditions: (a) 2-nitro-N-methylbenzenesulfonamide, PPh$_3$, DIAD, THF, rt, 24 h, 23%; (b) Cs$_2$CO$_3$, thiophenol, MeCN, rt, 72 h, 65%; (c) (1) ethyl chloroformate, N-methyl morpholine, THF, −10 °C to −5 °C, 15 min; (2) NaBH$_4$, THF/H$_2$O, 5 °C to rt, 18 h, 73%; (d) (1) tosyl chloride, DMAP, NEt$_3$, 1,2-DCE, 0 °C to rt, 4 h; (2) NaI, acetone, rt, 42 h, 66%; (e) DIPEA, MeCN (dry), 18.5 h at rt, 6 h at 55 °C, 28%; (f) TFA/DCM 1:1, H$_2$O, 5 °C, 99%.

■ RESULTS AND DISCUSSION

Chemistry. First, SAH analogs lacking either the amino or the acid group (**8a**, **8b**, series 1, Figure 2) were synthesized to identify the pharmacophoric significance of these moieties (Scheme 1). The acetyl-protected thiol building block **6**, used as precursor for both compounds, was synthesized from 2′,3′-O-isopropylideneadenosine **5** in a Mitsunobu reaction using thioacetic acid and diisopropyl azodicarboxylate (DIAD). The methyl ester **7a** was synthesized by nucleophilic substitution reaction of **6** with methyl-4-bromobutyrate (**2**), which was obtained from 4-bromobutyric acid (**1**) by esterification. The Boc-protected amine **7b** was obtained analogously by reaction of **6** with 3-(Boc-amino)propyl bromide (**4**), which was obtained by reaction of 3-aminopropyl bromide hydrobromide (**3**) with Boc anhydride. The obtained sulfide **7a** was deprotected using lithium hydroxide in THF/water followed by treatment with trifluoroacetic acid (TFA) and water in DCM at 5 °C to yield **8a**. Compound **7b** was deprotected by treatment with formic acid in water at 0 °C yielding **8b**.

In series 2 compounds, the homocysteine's sulfur atom was exchanged with nitrogen yielding N-adenosyl-2,4-diaminobutyric acid (adenosyl-Dab) derivatives (**14a–d**, **17a–e**). Also in this series, we generated two compounds that lack either the amino or the carboxylic acid function in the side chain (**14b**, **14c**). In addition, a five-membered ring (**14d**) was introduced in order to constrain the flexibility of the side chain. The necessary precursor aldehydes **11a–c** for reductive amination of 5′-amino-5′-deoxy-2′,3′-O-isopropylidene adenosine **12** (for the synthesis of **12**, see Scheme 3) were prepared in two-step syntheses starting from the protected carboxylic acids **9a–c** (Scheme 2). These were converted to the respective Weinreb amides **10a–c** with N,O-dimethylhydroxylamine and 1,1′-carbonyl-diimida-

zole (CDI) followed by reduction with diisobutylaluminum hydride (DIBAL) at −78 °C. The secondary amines **13a–d** were prepared by reductive amination of the aldehydes **11a–c** and the commercially available ketone **11d**, respectively, with amine **12** using sodium triacetoxyborohydride and acetic acid. Deprotection of **13a–d** to yield **14a–d** was achieved by treatment with 50% (v/v) TFA in dichloromethane at −20 °C followed by 14% (v/v) TFA in water at 8 °C.

The amine building block **12** was also modified with the commercially available aldehydes **15a–e** using sodium triacetoxyborohydride and acetic acid to form the respective benzylic amines **16a–e** (Scheme 3). These were treated with 14% (v/v) TFA in water at 8 °C to yield the target compounds **17a–c**. To accomplish the deprotection of **16d** and **16e** to **17d** and **17e**, respectively, a two-step procedure was required, using 50% (v/v) TFA in dichloromethane at −20 °C first, followed by treatment with 14% (v/v) TFA in water at 8 °C.

For the syntheses of the amides **21a–d** (series 3), 2′,3′-O-isopropylideneadenosine **5** was modified with phthalimide via Mitsunobu reaction with diisopropyl azodicarboxylate (DIAD) (Scheme 4) to yield compound **18**. In the following Gabriel synthesis, the phthalimide **18** was cleaved with hydrazine to yield 5′-amino-5′-deoxy-2′,3′-O-isopropylidene adenosine **12**. This intermediate was coupled with different Boc- and tert-butyl ester-protected carboxylic acids in the presence of 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate (TBTU) to obtain the corresponding amides **20a–d**. Concomitant deprotection of all protecting groups (Boc, tert-butyl ester, isopropylidene group) to yield compounds **21a–d** was achieved by performing a two-step procedure, using 50% (v/v) TFA in dichloromethane at −20 °C first, followed by treatment with 14% (v/v) TFA in water at 8 °C. To obtain the

47

Scheme 6. Synthesis of Y-Shaped Amine Derivatives 27b−t (Series 4)[a]
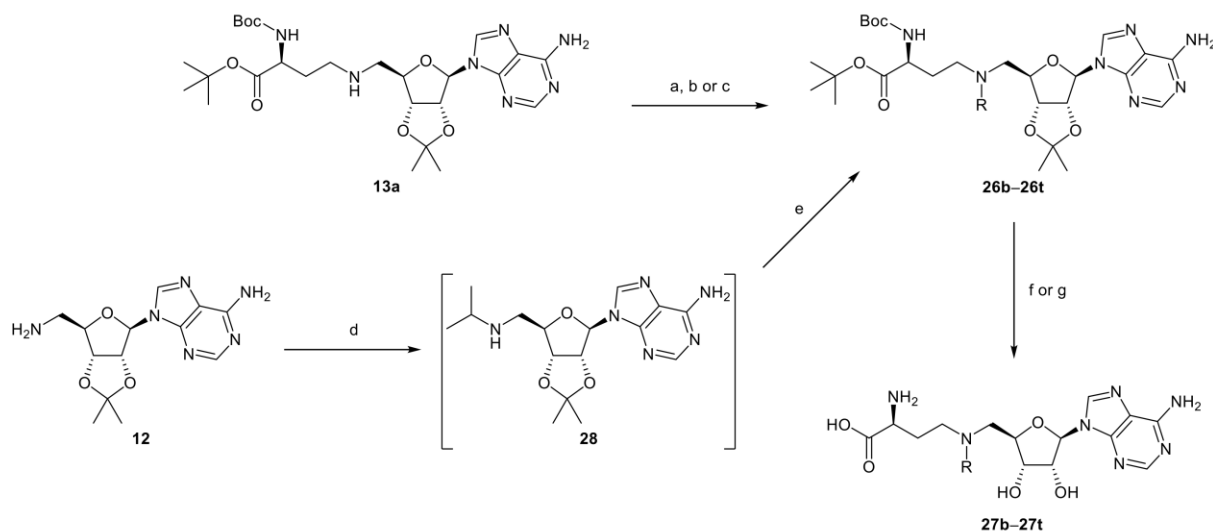
[a]Reagents and conditions: (a) RCHO, NaBH(OAc)$_3$, HOAc, 1,2-DCE, 0 °C → rt, overnight, 14−89%; (b) RBr, DIPEA, DMF, rt, overnight, 55−74%; (c) RBr or RCl, DIPEA, CuBr, DMF, rt, 16−48 h, 41−49%; (d) acetone, NaBH$_3$CN, HOAc, MeOH, 0 °C, 0.5 h; (e) 11a, NaBH(OAc)$_3$, HOAc, 1,2-DCE, 0 °C → rt, overnight, 66%; (f) TFA/DCM 1:1, H$_2$O, 5 °C, 99%; (g) (1) TFA/DCM 1:1, 5 °C; (2) TFA/H$_2$O 1:6, 5 °C, 99%. The structures of compounds 27b−27t are presented in Table 1.

deprotected phthalimide derivative 19, the isopropylidene-protected phthalimide 18 was treated analogously using 14% (v/v) TFA in water.

For the synthesis of the Y-shaped methyl derivate 27a of series 4 (Scheme 5), 2′,3′-O-isopropylideneadenosine 5 was functionalized with 2-nitro-N-methylbenzenesulfonamide under Mitsunobu conditions. The resulting sulfonamide 22 was hydrolyzed using cesium carbonate and thiophenol. In the following step, the methylated amine 23 was reacted with the alkyl iodide 25 to form the protected Y-shaped methyl derivative 26a, which was deprotected using TFA and water at 5 °C to give 27a. The alkyl iodide 25 was obtained by reducing the protected aspartate 9a with ethyl chloroformate and sodium borohydride followed by treatment first with tosyl chloride, then with sodium iodide.

The other Y-shaped compounds of series 4, 27b−t (Scheme 6, Table 1), were synthesized starting from amine 13a. This intermediate was either reacted with various aldehydes via reductive amination or different alkyl bromides and chlorides via nucleophilic substitution to obtain the tertiary amines 26b−t. To generate the isopropyl derivative 27d, the primary amine 12 was first alkylated with acetone and then with aldehyde 11a, each in a reductive amination reaction. In the final step, all protective groups were removed by treatment with TFA and water at 5 °C yielding compounds 27b−t. Compound 27s was obtained as a 50:50 mixture of two epimers (27s-A/B), but separation by HPLC (hydrophilic C$_{18}$ column MZ-Aqua Perfect) failed. Therefore, the single epimers 27s-A and 27s-B were obtained by separation of the precursor epimers 26s-A and 26s-B by HPLC (C$_{18}$) followed by deprotection of these single epimers by treatment with TFA and water at 5 °C.
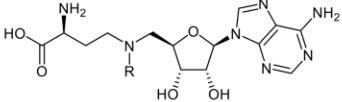
To evaluate the effect of compound 27s-A/B in cells, an ethyl ester prodrug was prepared. Therefore, 27s-A/B was treated with thionyl chloride in ethanol at 60 °C to give the corresponding ethyl ester 28-A/B as a 50:50-mixture of two epimers (Scheme 7). The separation of the epimers was achieved by HPLC yielding 28-A and 28-B. The assignment of
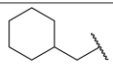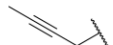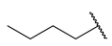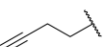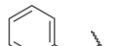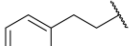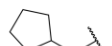
the absolute configuration of the chiral center within the alkyl side chain was not possible, neither for 27s nor for 28. However, since both epimers turned out to be equipotent (see below), we refrained from further efforts to assign the configuration.

Biological Evaluation: DNMT2 Binding and Inhibition. As a preliminary screening, all compounds including SFG and SAH were tested for binding to human DNMT2 at concentrations of 100 μM via MST to discriminate between binders and nonbinders (Figure 3). The readout of the assay was based on the normalized fluorescence signal. If a shift in the normalized fluorescence signal was detected, this was attributed to the formation of a protein−ligand complex. To discriminate binders and nonbinders, the 99% confidence intervals of the respective fluorescence signals were calculated. If the difference between the confidence interval of the measurement and that of the control measurement was more than 1‰ relative to the normalized fluorescence, the compound was defined as a binder. To evaluate the inhibitory potential of the compounds, tritium incorporation assays using DNMT2, substrate tRNA^{Asp}, and ^{3}H-SAM as cosubstrate were performed with compound concentrations of 100 μM. The results of these assays are presented in Table 2.

To evaluate the quality of an assay, especially of high-throughput screenings (HTS) Zhang et al. developed a tool called Z-factor.[54] This value can be used to compare assays, or to optimize their quality. A Z-factor of 1 represents an ideal assay, whereas Z-factors > 0.5 classify excellent assays. The same interpretation can be applied to the Z′-factor, which describes the quality of an assay itself. The only difference is that Z′ instead compares between negative and positive controls. For the prescreening assay via MST, SFG was chosen as positive control, since it was described as a potent pan-methyltransferase inhibitor in the literature and also showed high inhibition against DNMT2 within this study. Furthermore, 27s-A/B was chosen as another positive control compound because it was the most potent inhibitor within this study. Since both epimers

Table 1. Structures of N-Substituted Adenosyl-Dab Derivatives 27a−t (Series 4)[a]

R	Compound 27	R	Compound 27
	a		k
	b		l
	c		m
	d		n
	e		o
	f		p
	g		q
	h		r
	i		s-A/B*
	j		t

[a]* indicates that **27s-A/B** is a 1:1 mixture of two epimers with different configuration (*R* or *S*) at the stereocenter within the alkyl side chain. The single epimers with unknown configuration at this position are **27s-A** and **27s-B**.

Scheme 7. Synthesis of the Ethyl Esters 28-A and 28-B of Compound 27s-A/B[a]
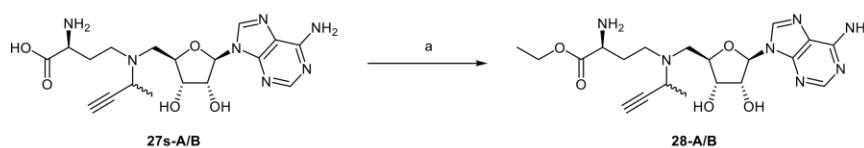
[a]Reagents and conditions: (a) SOCl$_2$, EtOH, 0 °C → 60 °C, 9 h, 27%.

showed comparable inhibition of DNMT2, the mixture was used. As negative control DMSO was chosen. A quartet of runs with technical septets were performed, each by three different persons, and revealed Z′-factors of 0.72 (SFG) and 0.76 (**27s-A/B**) for the prescreening assay via MST. It was stated by Zhang et al. that Z′-factors are always higher than their corresponding Z-factors, but they also stated that a Z-factor > 0 is sufficient for a "yes/no" type assay. Therefore, Z′-factors of 0.72 and 0.76 can be considered sufficient for the desired purposes within this study.

For selected compounds exhibiting both binding to DNMT2 and at least 60% inhibition of the enzymatic methylation activity at a concentration of 100 μM (**27m, 27n, 27o, 27s-A/B**), K_D and IC$_{50}$ values for DNMT2 binding and DNMT2 inhibition, respectively, were determined using ITC and the tritium

incorporation assay. The results are summarized in Table 3. Examples of dose−response curves for determination of the IC$_{50}$ values and examples of ITC thermograms are displayed in Figures 4 and 5.

In general, a very good correlation between DNMT2 binding and DNMT2 inhibition was found (Table 2). Compounds that did not emerge as DNMT2 binders also did not show significant DNMT2 inhibition at 100 μM. Only two compounds (**14a** and **27b**) that were identified as binders in the MST assay were found to be inactive in the tritium incorporation assay. In summary, the MST assay as implemented for this study exhibited good predictive value: 88.9% of all detected binders could be verified as inhibitors in the enzymatic assay. Furthermore, all nonbinders were found to be inactive in the enzyme inhibition assays. Therefore, the MST binding assay is
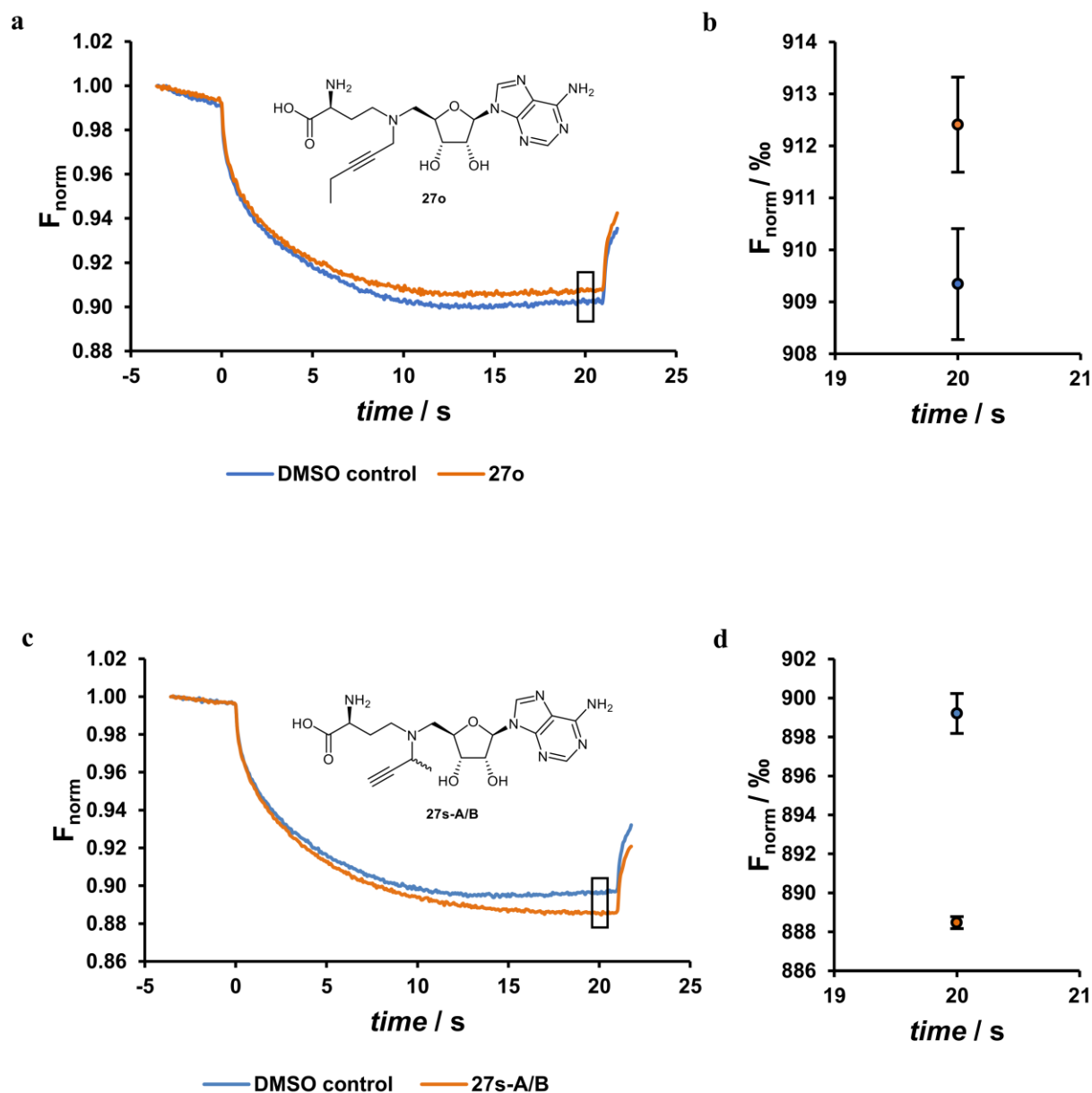
Figure 3. (a, c) MST traces of DNMT2 in absence (control, blue) and presence of 100 μM **27o** or **27s-A/B** (orange), shown as average of $n = 4$. F_{norm} indicates the fluorescence signal normalized to the initial fluorescence. (b, d) Fluorescence signal at 20 s MST-on time for control and for **27o** and **27s-A/B**; 99% confident intervals are given.

validated as a suitable method for early discrimination of promising compounds from those with no or low inhibitory activity.

In line with these general findings, both SFG and SAH could be classified as binders in the MST assay and also showed inhibition of the enzymatic activity (ca. 85% at 100 μM). IC_{50} values for SFG and SAH were determined to be 13.2 $\pm$ 0.8 μM and 15.8 $\pm$ 1.5 μM, respectively, while the K_D values, determined by ITC, were 7.5 $\pm$ 3.5 μM and 13.6 $\pm$ 4.4 μM.

The truncated SAH analogues **8a** and **8b** (series 1) exhibited neither DNMT2 binding nor DNMT2 inhibition indicating that both functionalities, that is, the amine and the carboxylic acid,

are crucial for an interaction with the enzyme. Within the secondary amine series with aliphatic chains (**14a–d**, series 2), only the amino acid derivative **14a** showed binding to DNMT2, but obviously without affecting the enzymatic activity. Compound **14a** was thus investigated at a higher concentration of 1 mM, at which it indeed inhibited the enzyme by 44%, confirming the low affinity toward DNMT2. Also, all the compounds of series 2 with aromatic moieties within the side chain exhibited neither binding to nor inhibition of DNMT2 (**17a–e**).

Comparably to the amines **14a–c**, the amides **21a–d** (series 3) contain either an amino acid chain or only the amine or the

50

Table 2. Binding of Compounds to DNMT2 as Determined by MST and Inhibition of DNMT2 as Determined in the Tritium Incorporation Assays
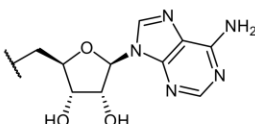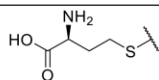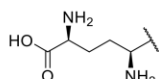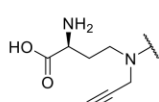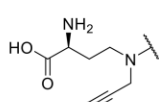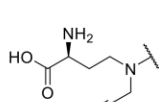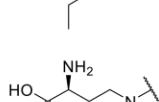
SAH, SFG, 17a–e, 19, 8a–b, 14a–d, 21a–e   27a–t

Compound	Binding at 100 μM (MST)	Inhibition at 100 μM [%][a]	Compound	Binding at 100 μM (MST)	Inhibition at 100 μM [%][a]
SAH	✓	85.7 ± 1.7*	27a	✓	30.3 ± 1.9**
SFG	✓	83.5 ± 1.1*	27b	✓	n.i.
17a	✗	n.i.	27c	✗	n.i.
17b	✗	n.i.	27d	✓	44.0 ± 2.7**
17c	✗	n.i.	27e	✗	n.i.
17d	✗	n.i.	27f	✗	n.i.
17e	✗	n.i.	27g	✗	n.i.
19	✗	n.i.	27h	✓	27.5 ± 0.4*
8a	✗	n.i.	27i	✓	32.0 ± 2.0*
8b	✗	n.i.	27j	✓	24.5 ± 8.5*
14a	✓	n.i.	27k	✗	n.i.
14b	✗	n.i.	27l	✓	53.9 ± 6.4**
14c	✗	n.i.	27m	✓	72.2 ± 1.2**
14d	✗	n.i.	27n	✓	61.3 ± 2.1*
21a	✗	n.i.	27o	✓	62.8 ± 0.1*
21b	✗	n.i.	27p	✓	30.2 ± 0.8*
21c	✗	n.i.	27q	✓	56.8 ± 6.8**
21d	✗	n.i.	27r	✓	23.2 ± 8.2*
			27s-A/B	✓	81.6 ± 2.1****
			27t	✓	17.3 ± 1.4***

[a]Mean values $\pm$ standard deviations of three independent measurements. n.i. = no inhibition. *$p < 0.05$. **$p < 0.01$. ***$p < 0.005$. ****$p < 0.001$.

Table 3. K_D Values, as Determined by ITC, and IC$_{50}$ Values of the Most Potent DNMT2 Inhibitors[d]

Compound		K_D [μM][a,b]	N	ΔG [kJ mol^{-1}]	ΔH [kJ mol^{-1}]	$-T\Delta S$ [kJ mol^{-1}]	IC$_{50}$ [μM][b]
SAH		13.6 ± 4.4[c]	1.1 ± 0.1	-27.8 ± 0.8	-55.3 ± 5.8	27.5 ± 6.6	15.8 ± 1.5
SFG		7.5 ± 3.5	1.3 ± 0.2	-29.3 ± 1.1	-24.5 ± 4.5	-4.8 ± 5.6	13.2 ± 0.8
27m		11.4 ± 2.4	0.93 ± 0.03	-28.2 ± 0.5	-37.8 ± 3.6	9.6 ± 4.1	77.1 ± 5.3
27n		10.4 ± 2.2	0.89 ± 0.03	-28.5 ± 0.5	-27.0 ± 1.2	-1.4 ± 3.1	39.7 ± 9.2
27o		10.5 ± 3.3	1.1 ± 0.1	-28.4 ± 0.8	-27.0 ± 3.6	-1.4 ± 4.4	32.2 ± 4.3
27s-A/B		8.1 ± 1.4	1.04 ± 0.04	-29.1 ± 0.4	-58.7 ± 4.1	29.6 ± 4.6	12.9 ± 1.9

[a]As determined by ITC. [b]Mean values ± standard deviations of three independent measurements. [c]Mean values ± standard deviations of six independent measurements. [d]N denominates the binding stoichiometry as determined by ITC.

carboxylic acid function with **21b** bearing a glutamate instead of an aspartate residue in the side chain. This should counter the shorter chain length generally found in amides compared to amines or thioethers. All compounds of this series showed neither binding to DNMT2 nor inhibition of the enzymatic activity. The same holds true for phthalimide **19**.

Next, Y-shaped tertiary amines (series 4, Figure 2), which structurally are more related to SFG, were investigated. The "inverse" SFG derivative with a methyl residue at the nitrogen atom (**27a**) displayed binding at 100 μM and also inhibited the DNMT2 activity by 30%. For the ethyl analogue **27b**, binding could be detected but not DNMT2 inhibition. The *n*-propyl and *n*-butyl derivatives **27c** and **27e** displayed neither binding nor inhibition, suggesting that incorporating linear aliphatic residues at this position is less suitable for the design of DNMT2 inhibitors.

Similar behavior could be observed for the branched derivatives **27d**, **27f**, and **27g**. The smallest derivative, **27d** (*i*-Pr), exhibited binding to the enzyme at 100 μM in the MST assay and inhibited the enzymatic activity by 44%. The larger

branched derivatives **27f** (*i*-Bu) and **27g** (neopentyl) showed neither binding nor inhibition at 100 μM.

The preference for smaller residues further became evident when investigating cyclic aliphatic residues. For compound **27h**, containing a cyclopropane residue, binding and inhibition (28%) were detected. Analogue **27i** with a four-membered ring was also determined as a binder and a weak inhibitor (32%). Compound **27j**, derivatized with a cyclopentane moiety, displayed binding and weak inhibition (25%). The compound with the largest investigated cyclic aliphatic residue, namely, a cyclohexane ring (**27k**), showed neither inhibition nor binding at 100 μM.

The best inhibitors of the methylation activity of DNMT2 were found within the series of the Y-shaped amines that contain an unsaturated residue. The allylic derivative **27l** was determined as a binder by MST and also inhibited the DNMT2 activity by 54%. Similar behavior was observed for the aromatic compounds **27q** and **27r**, with benzyl or a phenethyl groups, respectively. Both compounds bound to the enzyme at a concentration of 100 μM. Moreover, both compounds reduced the enzymatic activity at 100 μM by ca.
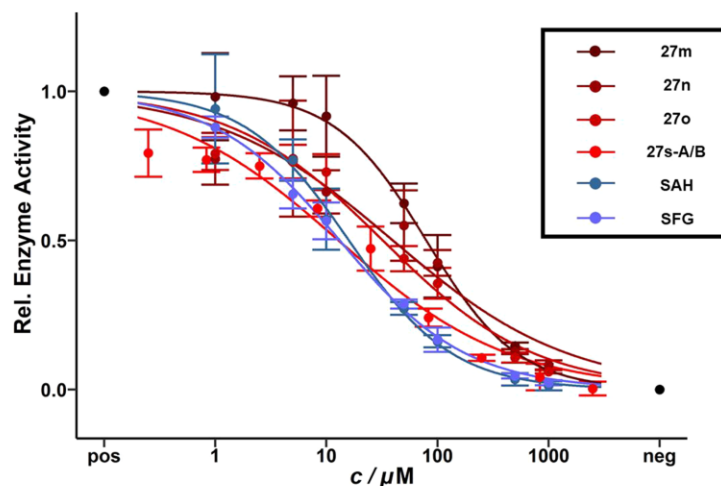
Figure 4. Dose−response curves for determination of the IC_{50} values for DNMT2 inhibition by the investigated compounds as determined by the tritium incorporation assay. Error bars refer to biological triplicates; "pos" describes the relative enzyme activity without compound and "neg" the relative enzyme activity without addition of substrate tRNA^{Asp}.

57% and 23%, respectively. The propargylic derivative **27m** also was found to bind to the enzyme at 100 μM and inhibited it by 72%.

For this compound, we determined the IC_{50} (Figure 4) and K_D values (Figure 5), using the tritium incorporation assay and ITC. While the K_D value of 11.4 ± 2.4 μM was very similar to that of the natural products SAH and SFG, the IC_{50} value of 77.1 ± 5.5 μM was higher. In order to identify the possible binding site of **27m**, a replacement titration via ITC was performed. For this purpose, DNMT2 was preincubated with **27m** and titrated against SAH. No binding enthalpy was detected suggesting, that **27m** and SAH address the same binding site.

In the next step, **27m** was further derivatized to attain the first structure−activity relationship of this compound class. The propargyl residue was exchanged with a 3-butinyl moiety (**27p**). This compound was found to bind to DNMT2, but only led to weak inhibition of the enzyme (30%) at 100 μM. Also, a methyl or ethyl group was attached to the triple bond. The corresponding compounds **27n** and **27o** were both identified as binders by MST. They also inhibited DNMT2 by ca. 62%. The determined IC_{50} values of 39.7 ± 9.2 μM and 32.2 ± 4.3 μM (Figure 4) were in the same range as those found for SAH and SFG and slightly better than that of **27m**. With 10.4 ± 2.2 μM for **27n** and 10.5 ± 3.3 μM for **27o**, the K_D values, as determined by ITC, were also in the same range.

Since both **27d** and **27m** exhibited substantial inhibition at 100 μM, both structures were merged, leading to **27s-A/B** and **27t**. These compounds were identified as binders in the MST assay. While **27t** reduced the enzymatic activity only by 17%, **27s-A/B** inhibited DNMT2 around 82% at 100 μM. The IC_{50} and K_D values (Figure 5) of 12.9 ± 1.9 μM and 8.1 ± 1.4 μM were comparable to those of SAH and SFG, and the IC_{50} value was found to be nearly 1 order of magnitude lower than that of the unsubstituted propargyl derivative **27m**.

Compound **27s-A/B** is a mixture of two epimers, with either (*S*) or (*R*) configuration of the stereocenter within the alkyl side chain. To assess if the single isomers exhibit different biological activities, the separated epimers **27s-A** and **27s-B** were analyzed regarding the inhibition of DNMT2. At 86% and 82%, the inhibition of both epimers was comparable and well in the range

of the 1:1 mixture **27s-A/B** (82% inhibition at 100 μM) indicating no different activities. An exact assignment of the separated epimers to the *R* or *S* derivatives was not possible by NMR, but since no preference for one epimer was observed, this was not investigated further.

With the exception of SAH and **27s-A/B**, the K_D values for the tested compounds (SAH, SFG, **27m**, **27n**, **27o**, **27s-A/B**) are generally lower than the IC_{50} values with ratios of $K_D/IC_{50} \approx 2$ (SFG), 7 (**27m**), 4 (**27n**), 3 (**27o**); that is, binding to DNMT2 does not lead to a "productive" inhibition to the same degree. This may be due to the different conditions used in both assays leading to slightly different hydration shells or surface charges, which might result in different binding behaviors. Furthermore, IC_{50} determination is strongly dependent on the experimental conditions, potentially adding to this bias.[55] The most profound difference between the assays is the lack of the substrate tRNA^{Asp} in ITC measurements. Binding of such a large and strongly charged molecule might induce conformational changes in the enzyme and is probably altering its surface charge significantly. Such differences would explain distinct binding modes of compounds in the presence or absence of RNA.

Our data clearly demonstrated that the amino acid residue is crucial for the interaction with DNMT2. This could be shown by investigating derivatives with aromatic linkers and with analogues lacking either the amino or carboxylic acid function. Exchange of the sulfur atom of SAH with secondary amines or amides abolished the activity of the compounds. With tertiary Y-shaped amines (series 4 compounds) however, DNMT2 activity could be successfully inhibited. In general, smaller residues were superior. The strongest inhibition could be achieved by unsaturated residues, with the propargylic derivatives **27m**, **27n**, **27o**, and **27s-A/B** being the most promising compounds. Remarkably, **27s-A/B**, which resulted from merging the most potent aliphatic compound **27d** with the propargylic derivative **27m**, was the most potent inhibitor of DNMT2, with IC_{50} and K_D values being similar to those of SAH and SFG. The thermodynamic data collected with the ITC measurements indicated that **27s-A/B** is the most enthalpically driven inhibitor (−58.7 kJ mol^{-1}) within this study, even more enthalpically driven than SAH (−55.3 kJ mol^{-1}). On the other hand, SFG

53
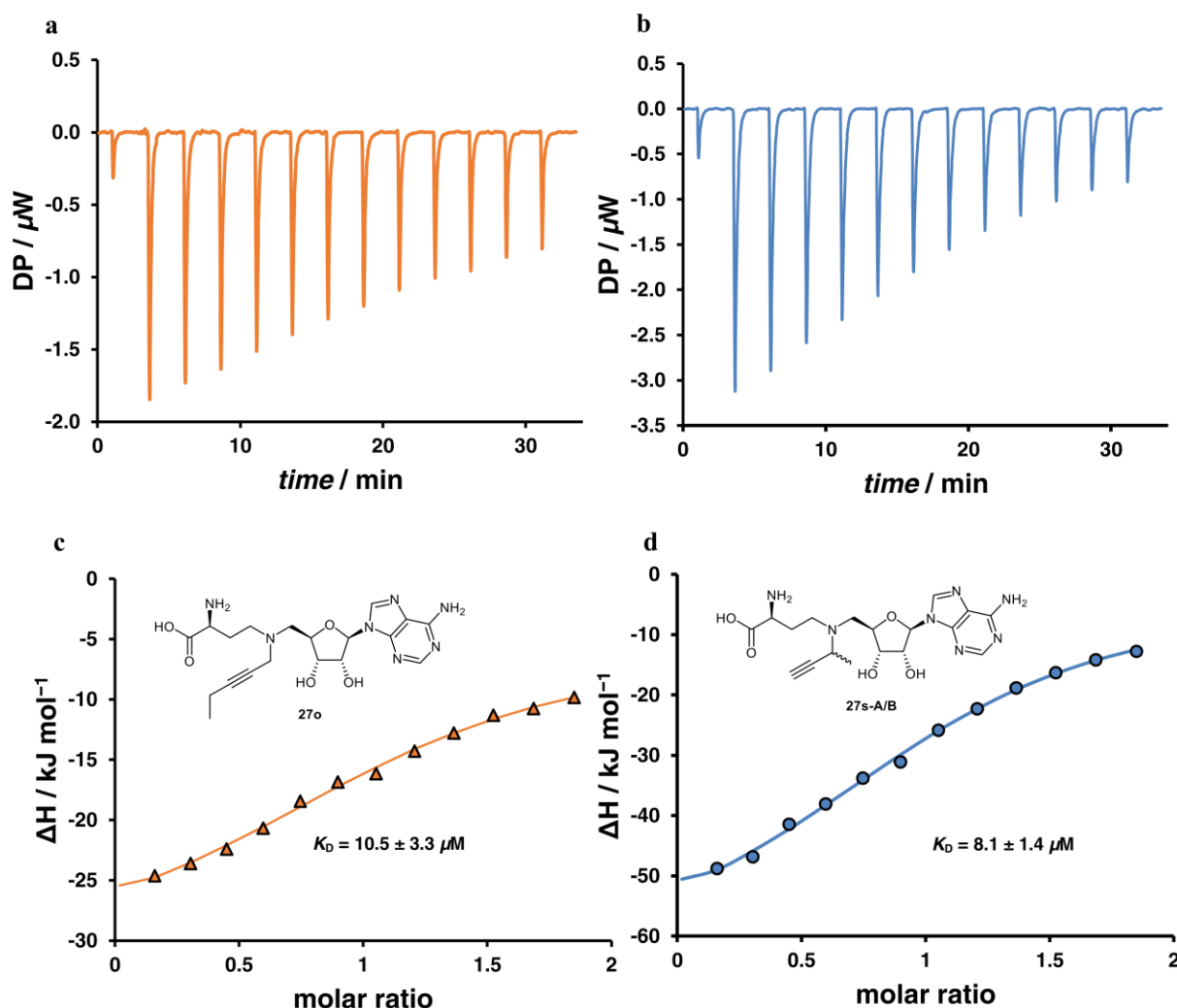
Figure 5. ITC thermograms (a, b) and titration curve fittings (c, d) of **27o** (a, c) and **27s-A/B** (b, d); K_D values, thermodynamic data, and binding stoichiometry are listed in Table 3.

Table 4. Selectivity Profiles of SAH, SFG, and Several DNMT2 Inhibitors against a Selection of SAM-Dependent MTases[a]

compound	DNMT3A-3L[b]	NSUN2[b]	NSUN6[b]	G9a[b]	DNMT2[b]
SAH	96.7 ± 2.8	97.3 ± 0.4	99.6 ± 4.6	94.3 ± 1.2	85.7 ± 1.7
SFG	98.0 ± 3.2	77.1 ± 2.2	70.2 ± 1.8	43.0 ± 9.8	83.5 ± 1.1
27m	97.6 ± 5.6	n.i.	25.9 ± 3.5	22.8 ± 0.3[c]	72.2 ± 1.2
27n	93.9 ± 2.1	n.i.	12.9 ± 7.6	37.6 ± 5.9	61.3 ± 2.1
27o	95.3 ± 1.6	n.i.	22.7 ± 4.9	32.0 ± 14.5	62.8 ± 0.1
27s-A/B	84.2 ± 0.7	n.i.	n.i.	71.0 ± 5.6	81.6 ± 2.1

[a]The inhibition of the enzymes was investigated at 100 μM inhibitor concentration in tritium incorporation assays. [b]Mean values ± standard deviations of three independent measurements. [c]Mean value ± standard deviation of two independent measurements. n.i. = no inhibition.

benefits the most from an entropic gain (-4.8 kJ mol^{-1}). Although **27m** showed a substantial gain in enthalpy (-37.8 kJ mol^{-1}), this effect is partially compensated by a loss of entropy (9.6 kJ mol^{-1}). The same applies for **27s-A/B** with an enthalpic gain of -58.7 kJ mol^{-1} but also a considerable loss of entropy (29.6 kJ mol^{-1}). For the elongated propargylic derivatives **27n** and **27o**, a remarkable gain in enthalpy (-27.0 kJ mol^{-1}) and a slight gain in entropy (-1.4 kJ mol^{-1}) could be detected.

Selectivity toward Different SAM-Dependent MTases.
Due to the close structural resemblance of the synthesized inhibitors to the ubiquitous MTase cosubstrate SAM, selectivity toward other SAM-dependent MTases is of major concern. The compounds listed in Table 4, including the natural products SAH and SFG, were subjected to tritium incorporation enzyme assays at a concentration of 100 μM with several different MTases to assess their selectivity profiles. The chosen MTases

54

included DNMT3A-3L[18] due to its close structural resemblance to DNMT2, the functionally related tRNA m5C MTases NSUN2 and NSUN6,[56−58] as well as the functionally unrelated but SAM-dependent histone MTase G9a.[59]

All investigated compounds showed better inhibition of the DNMT3A-3L construct compared to DNMT2. SAH and SFG both displayed around 100% inhibition at 100 μM, while **27m**, **27n**, and **27o** reduced the enzyme's activity by 98%, 94%, and 95%, respectively. At 84%, the inhibition by **27s-A/B** was the weakest one in this comparison exhibiting similar inhibition of DNMT2 and DNMT3A-3L.

SAH inhibited both NSUN2 and NSUN6 nearly to 100%, while SFG displayed weaker inhibition of 77% and 70%, respectively. All investigated synthetic compounds did not significantly inhibit NSUN2. NSUN6 activity was only weakly reduced by **27m**, **27n**, and **27o**, while **27s-A/B** inhibited neither NSUN2 nor NSUN6.

The G9a activity was reduced by all investigated inhibitors. SAH and SFG inhibited G9a by 94% and 43%, respectively. While **27m**, **27n**, and **27o** inhibited the histone MTase only weakly (23−38%), **27s-A/B** displayed stronger inhibition (71%).

The profound inhibition of DNMT3A-3L by all synthetic compounds clearly demonstrates that the compounds are not selective in this regard. Due to the high structural similarity between DNMT2 and DNMT3A this is not surprising, especially if one considers the structural similarity of the synthetic inhibitors to the ubiquitous cofactor SAM.

Inhibition of DNMT3A by SAH and SFG was reported with IC50 values in the low- to sub-micromolar range,[60,61] supporting the obtained results. High inhibition by SAH was also determined for the closely related DNMT3B (IC50 < 500 nM).[50,51]

In the comparison with the functionally related enzymes NSUN2 and NSUN6, however, the investigated compounds revealed remarkable selectivity. The most active synthetic DNMT2 inhibitor, **27s-A/B**, was not active at 100 μM against NSUN2 and NSUN6. This might be attributed to the Y-shaped structure, which is also suggested by the slight preference of SAH over SFG by both enzymes.

The inhibition of G9a by SFG was weak compared to other MTases with only around 40% inhibition at 100 μM. However, reported inhibition of G9a by SFG varies strongly, with IC50 values ranging from the low micromolar range to over 500 μM, and therefore this observation is not surprising.[62−64] Furthermore, the reported trend of inhibition by SAH and SFG, with more profound inhibition by SAH (94% at 100 μM) was reproduced.[63]

Compounds **27m**, **27n**, and **27o** exhibited only weak inhibition of G9a, which is also lower compared to that of DNMT2. On the other hand, **27s-A/B** reduced G9a activity strongly and comparably to DNMT2.

As expected, the designed DNMT2 inhibitors do not display very pronounced selectivity, most likely due to their close resemblance to the natural cofactor SAM. Especially for DNMT3A-3L, this problem becomes very evident since both proteins share conserved structural motives. Since DNMT1 shares a comparable high structural conservation, a similarly unselective behavior of the compounds can be anticipated.[18] This holds true especially when considering reported inhibition for SFG, SAH, and their most potent synthetic analogues toward DNMT1, which exhibit IC50 values in the low- to sub-micromolar range.[50,51,65,66] Also, for G9a, no profound

preference of the compounds toward DNMT2 was observed, particularly not for **27s-A/B**, the best synthetic DNMT2 inhibitor so far. Surprisingly, in the comparison with NSUN2 and NSUN6, no or only weak inhibition by the synthetic inhibitors was observed, proving that at least partial selectivity was achieved.

Inhibition of tRNA Methylation In Cellulo. To first determine the hydrolysis of the ethyl ester in aqueous solution at physiological pH value, compound **28-A** (100 μM) was dissolved in 100 mM TRIS buffer, pH 7.4, and incubated at 37 °C. The amount of hydrolysis product was determined by LC-MS after 3 h, followed by 6 h intervals (Figure 6).
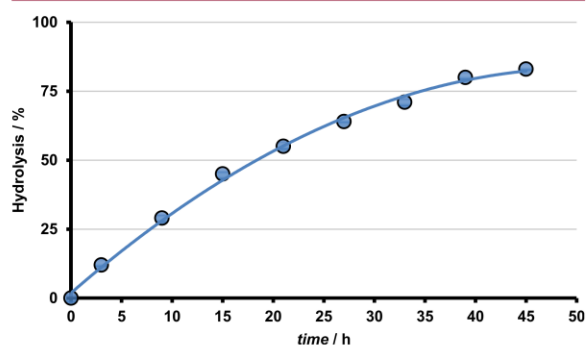
Figure 6. Hydrolysis of prodrug **28-A** (100 μM) in 100 mM TRIS buffer at pH 7.4 and 37 °C as determined by LC-MS.

Figure 6 shows that the ethyl ester is hydrolyzed to ca. 60% during the first 24 h of incubation.

The most promising compound **27s-A/B**, its ester derivative **28-A/B**, SFG, and a negative control (**27c**) were tested in a cellular context. Since no preference for one compound was identified (*vide supra*), a mixture of the epimers was used both in case of **27s** and **28**.

First, the cell toxicities were investigated in cell viability assays. Therefore, **27s-A/B**, SFG, **28-A/B**, and **27c** were subjected to MTT assays. All compounds showed negligible toxicity against HEK-293 cells up to the highest concentration in the assay, revealing 50% cytotoxic concentration (CC50) values >500 μM. According to these results, a concentration of 100 μM was chosen for cell-based experiments. Next, the inhibitors' potencies for modulating DNMT2 activity was elucidated in HEK-293 cells. SFG and derivative **27s-A/B** were chosen as representative molecules. The structurally related compound **27c** containing a propyl side chain was chosen as a negative control since it did not show inhibition of DNMT2 in vitro. For estimating inhibition in vivo, HEK-293 cells were incubated with 100 μM compound dissolved in DMSO or with DMSO only as a control for 24 h at 37 °C. Total RNA was extracted from treated cells, and subsequently total tRNA was isolated by gel electrophoresis. To determine m5C levels, LC-MS/MS and bisulfite sequencing were conducted. LC-MS/MS analysis was performed using total tRNA. This method allows highly accurate quantification, but information on sequence is lost as the RNA is digested down to nucleosides.[67] Therefore, bisulfite sequencing was chosen as an orthogonal method being capable of analyzing RNA on nucleotide level while the information about its sequence context is maintained.[68] With this approach, changes in m5C levels at position C38 of tRNA^Asp, which are attributed to DNMT2 activity, can be mapped.
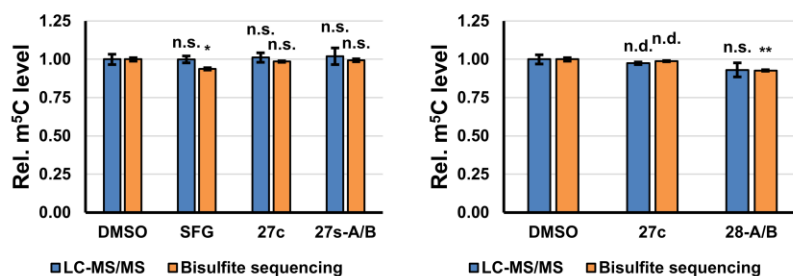
Figure 7. (a) Relative m^{5}C levels in HEK-293 cells after treatment with 100 μM SFG, **27c**, or **27s-A/B** for 24 h at 37 °C. Displayed are relative m^{5}C levels normalized to a DMSO control as elucidated by LC-MS/MS (blue) and bisulfite sequencing (orange). Error bars refer to three independent biological replicates, *p < 0.05, n.s. = not significant. (b) Relative m^{5}C levels in HEK-293 cells after treatment with 100 μM **27c** or **28-A/B** for 24 h at 37 °C. Displayed are relative m^{5}C levels normalized to a DMSO control as elucidated by LC-MS/MS (blue) and bisulfite sequencing (orange). Error bars refer to four (DMSO, **28-A/B**) or two (**27c**) independent biological replicates, **p < 0.01, n.s. = not significant, n.d. = significance was not determined due to low number of replicates.

Bisulfite sequencing and LC-MS/MS analysis did not reveal any significant decrease in m^{5}C levels relative to the DMSO control for the synthetic compounds **27c** and **27s-A/B** as shown in Figure 7a. Treatment with SFG on the other hand slightly reduced the m^{5}C level down to 94% according to bisulfite sequencing, whereas no significant change could be observed in LC-MS/MS measurements. According to in vitro enzyme assays, **27s-A/B** is able to bind and inhibit DNMT2 with comparable affinity as SFG. Changing the in vitro system with its limited components to a more complex cellular environment comes along with an increased number of possible target MTases, as well as the process of cell penetration. Consequently, the overall inhibitor concentration that is accessible for DNMT2 is significantly decreased. Thus, a prodrug approach was followed by converting the acidic moiety of **27s-A/B** into the ethyl ester **28-A/B**. The enhanced lipophilicity is expected to increase cell permeability, while the ester group can be cleaved by esterases in cellulo to yield the active compound **27s-A/B**.[69]

In vitro activity of ester **28-A/B** was elucidated using the tritium incorporation assay revealing no significant inhibition at a concentration of 100 μM. To assess the prodrug's activity in cellulo, HEK-293 cells were treated with the compound as described above.

As depicted in Figure 7b, the usage of ester **28-A/B** led to a slight reduction in m^{5}C levels in LC-MS/MS and bisulfite sequencing experiments, which was not significant for LC-MS/MS experiments but was significant for the latter ones.

Overall, it can be observed that the investigated compounds are not able to inhibit DNMT2 significantly in a cellular environment. But there is a good perspective in optimizing their physicochemical properties by applying a diverse range of prodrug approaches. Compared to the free acid **27s-A/B**, its ester derivative **28-A/B** seems to slightly reduce the m^{5}C levels indicating an improvement in its cellular concentration. However, this reduction was only perceived when applying bisulfite sequencing. Accordingly, the well-known pan MTase inhibitor SFG also showed a minor decrease in the modification level, only observable when evaluated by bisulfite sequencing.

Among others, possible explanations could be that studied compounds compromise multiple charged functional groups hindering permeation of the molecule through the cell membrane in a quantitative manner. Moreover, differences are present between bisulfite sequencing and LC-MS/MS analysis, which are possibly due to their principle of analyzing RNA: bisulfite sequencing is capable of identifying changes in the m^{5}C

level specific to position 38, which is targeted by DNMT2, whereas in LC-MS/MS measurements the determined m^{5}C level is restricted to the input RNA and loses its sequence information. Consequently, m^{5}C levels in LC-MS/MS data include the total amount of methylated cytidines independent from their position and target enzyme. In combination with the weak inhibitory effect in cellulo, no significant reduction in m^{5}C levels could be detected by LC-MS/MS whereas this was possible with bisulfite sequencing.

Cell Permeabilities. To determine the cell permeabilities of acid **27s** and its supposed ester prodrug **28**, CaCo-2 permeability assays were performed. A CaCo-2 monolayer separating a donor and acceptor compartment was treated with either **27s-A/B** or **28-B** at 37 °C for 24 h, and the solutions from each compartment were analyzed by LC-MS. Analyses of the solutions from the assays with **28-B** revealed 20% of ester **28-B** and 78% of acid in the donor compartment, whereas the acceptor compartment consisted of 2% acid. In the assays with the acid **27s-A/B**, only traces of the compound (0.6%) could be detected in the acceptor compartment. In conclusion, neither **27s-A/B** nor **28-B** was found to pass the cell barrier in significant amounts. In the case of the ester **28-B**, this might be due to its rapid hydrolysis rate observed under the tested conditions. This is in accordance with the results of the assays to detect inhibition of tRNA methylation in cellulo (see above): these assays showed very low, but significant inhibition of tRNA methylation in cells by the ester prodrug but not by the acid.

For stronger cellular effects, future investigations to obtain more potent inhibitors and to optimize the compounds' membrane permeabilities, for example, by strategies yielding prodrugs less susceptible to hydrolysis, are needed.

Structure−Activity Relationship (SAR) Elucidation by Molecular Docking and Solvent Analysis. To explain the SAR observed in the MST and tritium incorporation assays, molecular docking studies were performed. The chosen docking setup was able to reproduce the crystallographic binding mode of SAH (PDB-ID 1G55,[30] redocking RMSD = 0.77 Å, FlexX-score = −42.0 kJ mol^{-1}) indicating its ability to predict realistic binding modes. Due to the common SAH-related substructure of the molecules reported herein, binding poses showed strong overlap with the reference ligand as well as an overall narrow score distribution complicating discrimination of binders and nonbinders simply by docking score (Table S1). However, some general trends could be observed. For Y-shaped ligands (**27a−t**, series 4), orientation of the aliphatic side chains toward the

56

a   **b**
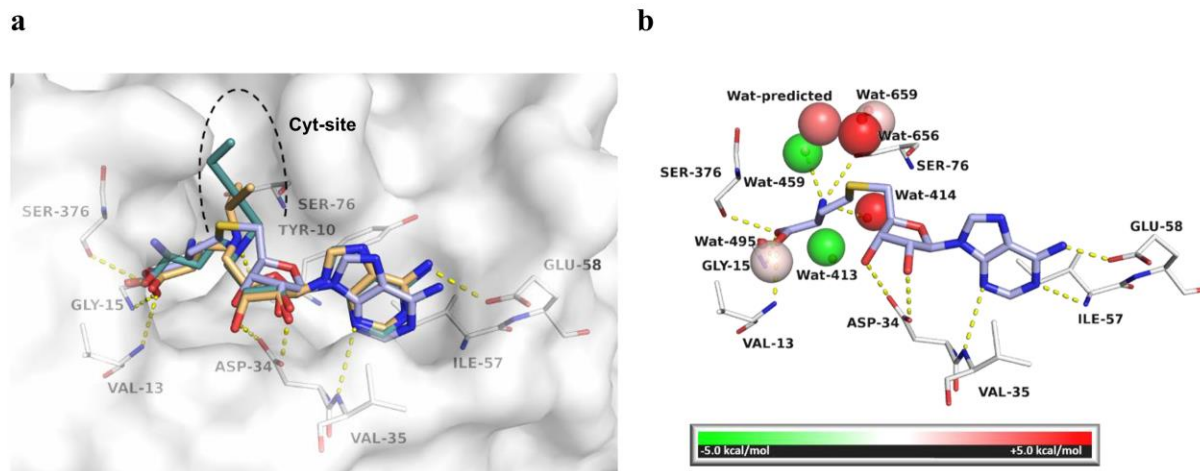
Figure 8. (a) Predicted binding modes of **27s-A/B** ((*R*)-epimer, light orange carbon atoms) and **27o** (light teal carbon atoms) in complex with hDNMT2 (white carbon atoms and surface) in superposition with crystallographic ligand SAH (light blue carbon atoms). (b) 3D-RISM predicted hydration sites and corresponding water molecules within 4 Å of the SAH homocysteine substructure colored by estimated interaction energy. Identical point of view for panels a and b. For clear view, only residues forming polar interactions (yellow dashes) are depicted as lines and labeled. PDB-ID 1G55. Figures were made with PyMOL.[71]

tRNA cytidine binding site was observed (Figure 8a) according to the design hypothesis and described by a favorable score for binders (in order of ascending FlexX-score −44.2 to −39.2 kJ mol⁻¹, Table S1) **27m, 27a, 27r, 27p, 27o, 27q, 27l, 27b, 27h,** and **27n** and a slightly lower score for bulky, potentially with the pocket clashing, nonbinders **27f, 27k,** and **27g** (FlexX score −28.5 to −38.5 kJ mol⁻¹, Table S1). This nicely agrees with the experimental results (Table 2, Table 3) but underestimates the affinity of **27s-A/B** (−37.0 kJ mol⁻¹ for the *R*- and −34.5 kJ mol⁻¹ for the *S*-epimer) and the weak binder **27t** (−24.9 kJ mol⁻¹). Also, in agreement with the experiments, removal of the acidic moiety resulted in a reduced score (**8b** FlexX-score −30.4 kJ mol⁻¹, **14c** −36.4 kJ mol⁻¹, **21d** −35.7 kJ mol⁻¹). Differently from the MST and tritium assay results, removal of the primary amine of the homocysteine moiety was tolerated in the docking studies (**8a** FlexX-score −42.1 kJ mol⁻¹, **17d** −40.2 kJ mol⁻¹, **21c** −49.0 kJ mol⁻¹), as was an amide linker (**21a** −41.4 kJ mol⁻¹, **21b** −51.1 kJ mol⁻¹). However, it seems reasonable that basic amines are more likely to mimic the charged sulfur atom of the native substrate SAM and bind with higher affinity compared to uncharged linkers like amides. To further elucidate the discrepancies between docking and experiment, solvent analysis was performed using the three-dimensional reference inter-action site model (3D-RISM, Table S2) approach. The detailed analysis of hydration sites within 4 Å of the derivatized homocysteine substructure of SAH demonstrated that crystallo-graphic water molecule positions could be reproduced with high accuracy (Figure 8b). This also includes water molecules being part of the docking setup (Wat-413, Wat-414, Wat-459, and Wat-495) forming three polar interactions with the enzyme or ligand with RMSD between crystallographic positions and 3D-RISM-predicted hydration sites within 0.16−0.93 Å. In close proximity to the primary amine, two hydration sites correspond-ing to Wat-459 and Wat-413 were further predicted to be highly stable (predicted ΔG of −7.2 and −10.8 kcal mol⁻¹, respectively). This might hold the explanation for the importance of the primary amine group as removal would disrupt the H-bond and water network potentially destabilizing the two hydration sites. Additionally to Wat-659 and Wat-656,

one more hydration site ("Wat-predicted", Figure 8b) not corresponding to a crystallographic water molecule was identified in the region toward the cytidine binding site. All three hydration sites are rather unfavorable (predicted ΔG of +0.6, +3.3 and +5.1 kcal mol⁻¹, respectively), and removal of the corresponding water molecules by aliphatic groups like in the Y-shaped molecules of series 4 (**27a−t**), except those with too bulky groups like in **27f, 27g,** and **27k**, should result in a beneficial desolvation and increased potency. This might also hold an explanation for the more favorable temperature-dependent entropy contribution to binding of **27m, 27n,** and **27o** (−*T*Δ*S* of +9.6, −1.4 and −1.4 kJ mol⁻¹, respectively, Table 3) compared to SAH (+27.5 kJ mol⁻¹). While desolvation of a hydrophobic sub-pocket is usually accompanied by a gain (at least a partial one) of entropy,[70] SAH lacks a substituent to displace these water molecules. This might also account for **27s-A/B** whose substituent might not reach deep enough into that sub-pocket (Figure 8a).

Crystal structure analysis indicated that the higher potency of **27m−o** against DNMT2 over NSUN6 might be attributed to a close shape-complementarity without larger gaps in the cytidine sub-pocket, which was found to be larger and open in NSUN6 (Figures S239 and S240).[72] While for NSUN2 no structure is available in the PDB, the most sequentially similar proteins are bacterial methyltransferases from *Pyrococcus horikoshii, Enter-ococcus faecium* or *Methanocaldococcus jannaschii* sharing only 29−32% sequence identity. Hence an Alphafold2-derived homology model[73,74] of NSUN2 was analyzed. It revealed both structural and sequential differences between NSUN2 and DNMT2, especially within the binding site, where only 17% identical and 30% similar residues were found. Additionally, an α helix forming one site of the cytidine binding site in DNMT2 (residues V366−L377) is not present in NSUN2 and NSUN6 highlighting a variety of potential selectivity determining features (Figure S241). Differently DNMT2 and DNMT3A share a highly similar fold (Figure S242) and slightly higher sequence similarity of 39%. However, within the binding site, the interaction profile is nearly identical (Figure S243) as is that of the cytidine-site residues. Nevertheless, some of the differences

57

like L377/W893, N375/R891 or S76/G707 might hold opportunities for selective ligand design.

CONCLUSION

In this study, we introduced new inhibitors of DNMT2 based on the adenosyl scaffold as novel examples for non-natural m^{5}C RNA MTase inhibitors, outclassing previously described synthetic specimens but not the natural compounds SAH and SFG.[53] Amino and amide analogues of SAH were designed, synthesized, and screened using MST and an enzymatic tritium incorporation assay. A clear correlation between the binding assay (MST) and the inhibition assay (^{3}H incorporation assays) with a high predictivity of the MST was found. Thus, in future studies, the laborious, time-consuming, and expensive tritium-based activity assay can be omitted for initial screenings and can be applied for hit validation. The most promising hits discovered in the present study were further investigated using ITC and detailed enzymatic assays. These revealed the compounds possessing similar binding and inhibition constants as SAH and SFG. As expected, these compounds compete with SAH for the same binding pocket. Our findings clearly indicated the need for a complete amino acid side chain for potent inhibition. Furthermore, a Y-shaped structure with a third substituent at the nitrogen atom was shown to be essential. This structure–activity relationship (SAR) is nicely backed by a computational analysis, which predicted the replacement of unfavorable water molecules by this third residue. As substituents at this position, small aliphatic moieties are favored, with alkyne groups being most potent. In summary, our findings will contribute to the further development of DNMT2 inhibitors and tool compounds, which might help to shed light on the underlying functions of this enzyme in the development of cancer or in epigenetic inheritance.

EXPERIMENTAL SECTION

Syntheses. *General Information.* All reagents and solvents were commercial grade and used without further purification. Reaction progress was monitored by thin-layer chromatography using Alugram Xtra F254 silica plates from Machery-Nagel. Column chromatography was performed with silica gel (40−63 μm) from Machery-Nagel. NMR spectra were recorded on Bruker Fourier 300 at 300 MHz. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance (CDCl$_3$, DMSO-d_6 or CD$_3$OD from Deutero GmbH) as internal standard. The identities and purities of final compounds were determined by combined HPLC/ESI-MS analysis using an Agilent 1100 series HPLC system with an Agilent Zorbax SB-Aq (4.6 × 150 mm^2; mobile phase MeCN/H$_2$O = 20:80 + 0.1% HCOOH; flow rate 0.7 mL/min) column. Samples were applied using 5 μL injection with quantitation by AUC at 254 nm. 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 semiopen capillaries. Specific rotations $[\alpha]_D^{20}$ were determined with a Krüss P3000 polarimeter and are given in deg cm^3 g^{-1} dm^{-1}. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

Methyl 4-Bromobutanoate (2). 4-Bromobutyric acid (2.06 g, 11.98 mmol, 1.0 equiv) was dissolved in methanol (20 mL); then thionyl chloride was added dropwise while stirring the solution. The mixture was kept stirring overnight at rt. The solvent was evaporated under reduced pressure at 40 °C, and the residue was dissolved in ethyl acetate (20 mL). After being washed with saturated NaHCO$_3$ solution (3 × 10 mL) and with brine (10 mL), the organic phase was dried over anhydrous Na$_2$SO$_4$. Evaporation of the solvent under reduced pressure at 40 °C afforded the desired product as an orange oil (1.72 g, 9.50 mmol, 79%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 3.68 (s, 3H), 3.45 (t, J = 6.4 Hz, 2H), 2.50 (t, J = 7.1 Hz, 1H), 2.23−2.09 (m, 2H). ^{13}C

NMR (75 MHz, CDCl$_3$): δ/ppm = 173.0, 51.7, 32.7, 32.2, 27.8. FT-IR: ν/cm^{-1} = 2952, 2028, 1733, 1436, 1367, 1313, 1250, 1206, 1172,1131, 1060, 1026, 994, 874, 779.

tert-Butyl (3-Bromopropyl)carbamate (3). 3-Bromopropylamine hydrobromide (1.01 g, 4.57 mmol, 1.0 equiv), di-*tert*-butyl dicarbonate (1.20 g, 5.48 mmol, 1.2 equiv), and triethylamine (1.4 mL, 10.05 mmol, 2.2 equiv) were suspended in ethyl acetate (20 mL), and the mixture was stirred overnight at rt. The mixture was then washed with saturated NaHCO$_3$ solution (3 × 10 mL), 1 M HCl (3 × 10 mL), and brine (10 mL), and the organic phase was dried over anhydrous Na$_2$SO$_4$. Evaporation of the solvent under reduced pressure at 40 °C afforded the desired product as a colorless solid (707 mg, 3.22 mmol, 70%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 3.43 (t, J = 6.5 Hz, 2H), 3.26 (t, J = 6.6 Hz, 2H), 2.09−1.98 (m, 2H), 1.43 (s, 9H,). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 156.0, 79.5, 39.1, 32.7, 30.8, 28.4. FT-IR: ν/cm^{-1} = 3375, 2980, 2950, 1682, 1520, 1436, 1387, 1363, 1340,1295, 1278, 1247, 1221, 1161, 1134, 1086, 1040, 1030, 991, 916, 873. Mp: 39−42 °C.

S-(((3aS,4S,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl) Ethane-thioate (6). According to the method of Pignot et al.,[75] a solution of triphenylphosphine (1.69 g, 6.44 mmol, 2.2 equiv) in dry THF (10 mL) was cooled to 0 °C in an ice bath. Then DIAD (1.26 mL, 6.44 mmol, 2.2 equiv) was added dropwise while stirring, over a period of 10 min. The resulting suspension was stirred for an additional 30 min while keeping it at 0 °C. Compound 5 (902 mg, 2.93 mmol, 1.0 equiv) was added, and the mixture was stirred and cooled for 15 min. Thioacetic acid (0.46 mL, 6.44 mmol, 2.2 equiv) was diluted with dry THF (2 mL) and added dropwise to the suspension, while stirring. The resulting mixture was stirred at 0 °C for an additional hour and subsequently overnight at rt. The solvent was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography (DCM/MeOH = 97:3) to afford the desired product as a pale-yellow solid (1.02 g, 2.79 mmol, 95%). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = δ 8.25 (s, 1H), 8.24 (s, 1zH), 6.18 (d, J = 2.3 Hz, 1H), 5.53 (dd, J = 6.4, 2.3 Hz, 1H), 4.99 (dd, J = 6.4, 3.1 Hz, 1H), 4.26 (td, J = 6.9, 3.1 Hz, 1H), 3.29−3.12 (m, 2H), 2.32 (s, 3H), 1.57 (s, 3H), 1.37 (s, 3H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 194.2, 155.3, 151.9, 148.1, 139.7, 118.5, 113.4, 89.5, 85.0, 83.2, 82.8, 30.0, 28.3, 25.3, 23.4. FT-IR: ν/cm^{-1} = 3332, 3183, 2987, 2936, 1690, 1640, 1597, 1578, 1475, 1423, 1373, 1329, 1296, 1237, 1206, 1155, 1134, 1076, 1002, 968. Mp: 54−57 °C. $[\alpha]_D^{20}$ = −32 (10 mg/mL; MeOH).

Methyl 4-((((3aS,4S,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)thio)-butanoate (7a). According to the method of Pignot et al.,[75] to a solution of compound 6 (150 mg, 0.41 mmol, 1.0 equiv) and 2 (111 mg, 0,62 mmol, 1.5 equiv) in dry methanol (10 mL) was added sodium methoxide (49 mg, 0.90 mmol, 2.2 equiv) while stirring under an argon atmosphere. After stirring overnight at rt, the solvent was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography (DCM/MeOH = 97:3) to afford the desired product as a colorless solid (125 mg, 0.30 mmol, 72%). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.29 (s, 1H), 8.24 (s, 1H), 6.19 (d, J = 2.4 Hz, 1H), 5.55 (dd, J = 6.4, 2.4 Hz, 1H), 5.08 (dd, J = 6.4, 3.0 Hz, 1H), 4.35 (td, J = 6.9, 3.0 Hz, 1H), 3.65 (s, 3H), 2.81 (d, J = 6.9 Hz, 2H), 2.53 (t, J = 7.2 Hz, 2H), 2.36 (t, J = 7.2 Hz, 2H), 1.78 (p, J = 7.3 Hz, 2H), 1.60 (s, 3H), 1.40 (s, 3H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 173.1, 155.3, 151.9, 139.8, 113.3, 89.6, 86.2, 83.1, 83.1, 49.9, 32.8, 31.2, 30.3, 25.2, 23.6, 23.4. FT-IR: ν/cm^{-1} = 3297, 3135, 2990, 2935, 2026, 1732, 1732, 1670, 1600, 1561, 1473, 1365, 1328, 1304, 1267, 1200, 1175, 1092, 1057, 971. Mp: 99−102 °C. $[\alpha]_D^{20}$ = −14 (10 mg/mL; MeOH).

tert-Butyl (3-((((3aS,4S,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)thio)propyl)-carbamate (7b). To a solution of compound 6 (150 mg, 0.41 mmol, 1.0 equiv) and 3 (146 mg, 0,62 mmol, 1.5 equiv) in dry methanol (10 mL) was added sodium methoxide (49 mg, 0.90 mmol, 2.2 equiv) while stirring under an argon atmosphere. After stirring overnight at rt, the solvent was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography (DCM/MeOH = 97:3) to afford the desired product as a colorless solid (165 mg, 0.34 mmol,

84%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.33 (s, 1H), 8.17 (s, 1H), 7.34 (s, 2H), 6.16 (d, J = 2.4 Hz, 1H), 5.50 (dd, J = 6.3, 2.4 Hz, 1H), 4.99 (dd, J = 6.3, 2.7 Hz, 1H), 4.26–4.21 (m, J = 8.0, 6.4, 2.7 Hz, 1H), 2.95–2.89 (m, J = 6.4 Hz, 2H), 2.81–2.66 (m, J = 13.5, 7.2 Hz, 2H), 2.44 (t, J = 7.3 Hz, 2H), 1.53 (s, 3H), 1.36 (s, 9H), 1.33 (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 156.1, 155.5, 152.7, 148.7, 140.0, 120.1, 113.2, 89.2, 85.7, 83.34, 83.1, 77.4, 33.4, 29.5, 28.7, 28.2, 26.8, 25.1. FT-IR: ν/cm^{-1} = 3323, 3181, 2978, 2930, 1695, 1638, 1596, 1508, 1475, 1458, 1420, 1366, 1329, 1248, 1207, 1161, 1075, 982, 868, 798, 782, 723. Mp: 74–77 °C. $[\alpha]_D^{20}$ = −9 (10 mg/mL; MeOH).

4-((((2S,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetra-hydrofuran-2-yl)methyl)thio)butanoic Acid (8a). To a solution of compound **7a** (175 mg, 0.41 mmol, 1.0 equiv) in THF (4 mL) and water (3 mL) was added lithium hydroxide monohydrate (52 mg, 1.24 mml, 3.0 equiv). After stirring at rt for 1 h, the mixture was acidified with 1 M HCl and extracted with CHCl$_3$/^{i}PrOH (3:1). The combined organic extracts were dried over Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C. The residue was taken up in DCM (1.5 mL), and trifluoroacetic acid (1.5 mL), followed by water (200 μL), was added dropwise at 5 °C. After stirring at 5 °C for 1 h, the solvent was removed under reduced pressure and co-distilled with DCM (3 × 40 mL) at 40 °C. Water was added, and the mixture was lyophilized to give the desired product as a colorless resin (146 mg, 0.39 mmol, 96%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.45 (s, 1H), 8.35 (s, 1H), 6.00 (d, J = 4.9 Hz, 1H), 4.69 (t, J = 4.9 Hz, 1H), 4.27 (t, J = 4.9 Hz, 1H), 4.21–4.12 (m, 1H), 2.97–2.80 (m, 2H), 2.60–2.51 (m, 2H), 2.34–2.26 (m, 2H), 1.85–1.72 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 176.9, 152.4, 150.1, 145.9, 144.0, 120.6, 90.5, 86.0, 75.3, 73.9, 35.1, 33.5, 33.1, 25.9. FT-IR: ν/cm^{-1} = 3099, 2934, 1693, 1508, 1420, 1324, 1197, 1136, 1096, 1054, 836, 798, 785, 723, 681. $[\alpha]_D^{20}$ = −2 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 370.10 (100%), 371.04 (21.1%), 372.09 (6.4%). Calculated: 370.12 (100%), 371.12 (18.2%), 372.12 (1.4%). Purity: 95% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.17 min.

(2R,3R,4S,5S)-2-(6-Amino-9H-purin-9-yl)-5-(((3-aminopropyl)-thio)methyl)tetrahydrofuran-3,4-diol (8b). To a suspension of compound **7b** (50 mg, 0.10 mmol) in water (1 mL), formic acid (1 mL) was added at 0 °C. The mixture was stirred for 2 days at rt. The solvent was removed by lyophilization, and the desired product was obtained as a pale-yellow solid (34 mg, 0.99 mmol, 96%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.17 (s, 1H), 7.34 (s, 2H), 5.89 (d, J = 5.6 Hz, 1H), 4.75 (t, J = 5.6 Hz, 1H), 4.14 (t, J = 4.4 Hz, 1H), 4.00 (d, J = 4.4 Hz, 1H), 3.02–2.71 (m, 2H), 2.10 (t, J = 7.5 Hz, 1H), 1.72 (q, J = 7.5 Hz, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 174.2, 156.4, 152.9, 149.8, 140.3, 119.5, 87.9, 84.4, 72.89, 34.2, 33.0, 31.8, 25.6. FT-IR: ν/cm^{-1} = 3332, 3183, 2927, 2360, 2341, 1643, 1574, 1477, 1421, 1374, 1333, 1301, 1248, 1209, 1174, 1127, 1088, 1045, 1003, 826, 796, 719. Mp: 66–69 °C. $[\alpha]_D^{20}$ = −22 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 341.13 (100%), 342.11 (15.9%), 343.12 (6.1%). Calculated: 341.14 (100%), 342.14 (17.2%), 343.14 (5.6%). Purity: 97% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 4.00 min.

tert-Butyl N^2-(tert-Butoxycarbonyl)-N^4-methoxy-N^4-methyl-L-as-paraginate (10a). According to the method of Zhang et al.,[76] to a solution of Boc-Asp-O^{t}Bu (19.4 g, 67.0 mmol, 1.0 equiv) in DCM (130 mL) was added 1,1-carbonyldiimidazole (12.0 g, 73.7 mmol, 1.1 equiv), and the reaction was stirred at rt for 1 h. Then, N,O-dimethylhydroxyl-amine hydrochloride (7.19 g, 73.7 mmol, 1.1 equiv) was added, and the reaction mixture was stirred at rt for a further 18 h. The mixture was diluted with EtOAc (100 mL), and the residue was filtered off. The filtrate was washed several times with 1 M HCl, saturated NaHCO$_3$ solution, and brine until no side products were detected by TLC. The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C to give **10a** (18.7 g, 56.3 mmol, 84%) as a colorless oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 5.65 (d, J = 8.9 Hz, 1H), 4.49–4.38 (m, 1H), 3.67 (s, 3H), 3.15 (s, 3H), 3.13–2.80 (m, 2H), 1.44 (s, 9H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 171.9, 170.7, 155.9, 81.9, 79.6, 61.3, 50.5, 34.8, 32.1, 28.5, 28.0. FT-IR: ν/cm^{-1} = 3252, 2984, 2934, 1729, 1703, 1630, 1520, 1250, 1141, 1000, 960, 841. $[\alpha]_D^{20}$ = +15 (10 mg/mL; CHCl$_3$).

tert-Butyl 4-(Methoxy(methyl)amino)-4-oxobutanoate (10b). 4-(tert-Butoxy)-4-oxobutanoic acid (500 mg, 2.87 mmol, 1 equiv) and 1,1'-carbonyldiimidazole (512 mg, 3.16 mmol, 1.1 equiv) were dissolved in DCM (10 mL) and stirred at room temperature for 1 h. N,O-Dimethylhydroxylamine hydrochloride (308 mg, 3.16 mmol, 1.1 equiv) was added, and the reaction mixture was allowed to stir at room temperature for 19 h. After the solid matter was filtered off, the filtrate was diluted with DCM (10 mL) and washed with 1 M HCl (15 mL), saturated NaHCO$_3$ solution (15 mL), and brine (15 mL). The organic layer was dried over anhydrous Na$_2$SO$_4$ and concentrated under reduced pressure at 40 °C to give the desired product as a colorless oil (552 mg, 2.54 mmol, 89%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 3.68 (s, 3H), 3.15 (s, 3H), 2.67 (t, J = 7.3 Hz, 2H), 2.52 (t, J = 7.3 Hz, 2H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 173.1, 172.3, 80.5, 61.3, 32.4, 29.9, 28.2, 27.1. FT-IR: ν/cm^{-1} = 2984, 2934, 1720, 1663, 1360, 1248, 1159, 1105, 987, 888, 847, 751.

tert-Butyl (3-(Methoxy(methyl)amino)-3-oxopropyl)carbamate (10c). Boc-β-Ala-OH (500 mg, 2.64 mmol, 1 equiv) and 1,1'-carbonyldiimidazole (471 mg, 2.91 mmol, 1.1 equiv) were dissolved in DCM (10 mL) and stirred at room temperature for 1 h. N,O-Dimethylhydroxylamine hydrochloride (284 mg, 2.91 mmol, 1.1 equiv) was added, and the reaction mixture was allowed to stir at room temperature for 17 h. After the solid matter was filtered off, the filtrate was diluted with DCM (10 mL) and washed with 1 M HCl (15 mL), saturated NaHCO$_3$ solution (15 mL), and brine (15 mL). The organic layer was dried over anhydrous Na$_2$SO$_4$ and concentrated under reduced pressure at 40 °C to give the desired product as a colorless oil (537 mg, 2.31 mmol, 88%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 6.72 (t, J = 5.7 Hz, 1H), 3.65 (s, 3H), 3.14 (q, J = 7.0 Hz, 2H), 3.08 (s, 3H), 2.52 (t, J = 7.0 Hz, 2H), 1.37 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 171.6, 155.5, 77.5, 60.9, 35.9, 31.7, 31.6, 28.2. FT-IR: ν/cm^{-1} = 3346, 2975, 2937, 1698, 1651, 1505, 1449, 1390, 1365, 1249, 1166, 995, 971, 782.

tert-Butyl (S)-2-((tert-Butoxycarbonyl)amino)-4-oxobutanoate (11a). According to the method of Zhang et al.,[76] a solution of diisobutylaluminum hydride (DIBAL) in hexane (1 M, 4.06 mmol, 4.06 mL, 1.5 equiv) was added dropwise to a solution of compound **10a** (900 mg, 2.71 mmol, 1 equiv) in anhydrous THF (25 mL) at −78 °C. The reaction mixture was stirred at −78 °C for 2.5 h. KHSO$_4$ (0.35 M, 20 mL) was added, and the mixture was extracted with diethyl ether (3 × 80 mL). The combined organic extracts were washed with 1 M HCl (3 × 50 mL), saturated NaHCO$_3$ solution (3 × 50 mL), and brine (3 × 50 mL). After drying with anhydrous Na$_2$SO$_4$, the solution was concentrated under reduced pressure at 40 °C to obtain the desired product as a colorless oil (600 mg, 2.20 mmol, 81%) that crystallized after storage at −20 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 9.72 (s, 1H), 5.36 (d, J = 8.1 Hz, 1H), 4.54–4.39 (m, 1H), 3.07–2.87 (m, 2H), 1.44 (s, 9H), 1.43 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 199.5, 170.0, 155.4, 82.7, 80.0, 49.4, 46.4, 28.4, 28.0. FT-IR: ν/cm^{-1} = 3458, 2977, 2932, 1732, 1711, 1494, 1266, 1333, 1223, 1055, 846. Mp: 67–70 °C. $[\alpha]_D^{20}$ = +16 (10 mg/mL; CHCl$_3$).

tert-Butyl 4-Oxobutanoate (11b). A solution of diisobutylalumi-num hydride (DIBAL) in hexane (1 M, 3.05 mmol, 3.05 mL, 1.5 equiv) was added dropwise to a solution of compound **10b** (442 mg, 2.03 mmol, 1 equiv) in anhydrous THF (12 mL) at −78 °C. The reaction mixture was stirred at −78 °C for 2 h. KHSO$_4$ solution (0.35 M, 10 mL) was added, and the mixture was extracted with ethyl acetate (3 × 25 mL). The combined organic extracts were washed with 1 M HCl (3 × 15 mL), saturated NaHCO$_3$ solution (3 × 15 mL), and brine (3 × 15 mL). After drying with anhydrous Na$_2$SO$_4$, the solution was concentrated under reduced pressure at 40 °C to obtain the desired product as a colorless oil (109 mg, 0.69 mmol, 34%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 10.09 (s, 1H), 3.05–2.98 (m, 2H), 2.87–2.79 (m, 2H), 1.73 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 200.5, 171.6, 81.0, 38.8, 30.4, 28.1. FT-IR: ν/cm^{-1} = 2978, 2932, 1724, 1456, 1392, 1366, 1246, 1151, 1073, 949, 875, 845, 754.

tert-Butyl (3-Oxopropyl)carbamate (11c). A solution of diisobuty-laluminum hydride (DIBAL) in hexane (1 M, 2.83 mmol, 2.83 mL, 1.5 equiv) was added dropwise to a solution of compound **10c** (438 mg, 1.89 mmol, 1 equiv) in anhydrous THF (12 mL) at −78 °C. The

reaction mixture was stirred at −78 °C for 2 h. KHSO$_4$ solution (0.35 M, 6 mL) was added, and the mixture was extracted with diethyl ether (3 × 25 mL). The combined organic extracts were washed with 1 M HCl (3 × 15 mL), saturated NaHCO$_3$ solution (3 × 15 mL), and brine (3 × 15 mL). After drying with anhydrous Na$_2$SO$_4$, the solution was concentrated under reduced pressure at 40 °C to obtain the desired product as a colorless oil (110 mg, 0.64 mmol, 34%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 10.09 (s, 1H), 5.32 (d, J = 7.5 Hz, 1H), 3.71 (q, J = 6.0 Hz, 2H), 2.99 (t, J = 6.0 Hz, 2H), 1.71 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 201.5, 155.9, 79.4, 44.4, 34.1, 28.4. FT-IR: ν/cm^{-1} = 3367, 2978, 2932, 1684, 1510, 1366, 1274, 1249, 1163, 1040, 1003, 751, 666.

9-((3aR,4R,6R,6aR)-6-(Aminomethyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine (12). According to the method of Liu et al.,[77] to a suspension of compound **18** (3.63 g, 8.32 mmol, 1 equiv) in ethanol (100 mL) was added hydrazine monohydrate (6.47 mL, 133.08 mmol, 16 equiv), and the mixture was refluxed for 1 h. After cooling, the precipitate was filtered off, and the solvent was removed under reduced pressure at 40 °C. The residues were dissolved in ethanol (60 mL) and stirred for 15 min at room temperature, followed by filtration. The filtrate was concentrated under reduced pressure at 40 °C, and the residues were purified by column chromatography (DCM/MeOH + 0.1% NEt$_3$) to afford the desired product as a white powder (2.32 g, 7.57 mmol, 91%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.37 (s, 1H), 8.15 (s, 1H), 7.34 (s, 2H), 6.08 (d, J = 3.2 Hz, 1H), 5.45 (dd, J = 6.3, 3.2 Hz, 1H), 4.98 (dd, J = 6.3, 2.8 Hz, 1H), 4.09 (td, J = 5.8, 2.8 Hz, 1H), 2.70 (dd, J = 5.8, 2.8 Hz, 2H), 1.53 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 156.2, 152.8, 149.0, 140.0, 119.2, 113.2, 89.1, 87.1, 82.7, 81.6, 43.7, 27.1, 25.2. FT-IR: ν/cm^{-1} = 3538, 3331, 3176, 1746, 1701, 1660, 1601, 1477, 1381, 1330, 1303, 1270, 1218, 1091, 1016, 902, 872, 798. Mp: 212−215 °C. [α]$_D^{20}$ = −32 (10 mg/mL; MeOH).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (13a). To a 0 °C cooled solution of **12** (7.18 g, 28.9 mmol, 1.1 equiv) in THF/MeCN (1:3 = 160 mL) under argon atmosphere were added a suspension of **11a** (7.18 g, 26.3 mmol, 1.0 equiv) in THF/MeCN (1:3 = 12 mL) and HOAc (1.50 mL, 26.3 mmol, 1.0 equiv). After stirring for 30 min, NaBH(OAc)$_3$ (8.36 g, 39.4 mmol, 1.5 equiv) was added, and the reaction was stirred for 18 h at 5 °C. The suspension was diluted with saturated NaHCO$_3$ solution (250 mL), and the aqueous layer was extracted with DCM (5 × 50 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1 to 10:1 + 0.5% NEt$_3$) to give **13a** (9.11 g, 16.2 mmol, 62%) as a colorless solid. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.26 (s, 1H), 7.88 (s, 1H), 6.40 (s, 2H), 6.05 (d, J = 8.2 Hz, 1H), 5.93 (d, J = 3.4 Hz, 1H), 5.45 (dd, J = 6.5, 3.4 Hz, 1H), 5.03 (dd, J = 6.5, 3.4 Hz, 1H), 4.37−4.30 (m, 1H), 4.29−4.19 (m, 1H), 2.96−2.66 (m, 3H), 2.65−2.53 (m, 1H), 1.98−1.71 (m, 2H), 1.56 (s, 3H), 1.43−1.29 (m, 21H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 171.9, 156.0, 155.7, 153.1, 149.4, 139.9, 120.4, 114.6, 91.1, 85.0, 83.1, 82.2, 81.7, 79.4, 53.0, 51.4, 46.3, 32.3, 28.4, 28.1, 27.4, 25.5. FT-IR: ν/cm^{-1} = 2980, 1703, 1640, 1597, 1504, 1476, 1425, 1367, 1330, 1297, 1213, 1152, 1075, 857, 798, 750, 665. Mp: 96−100 °C. [α]$_D^{20}$ = −33 (10 mg/mL; CHCl$_3$).

tert-Butyl 4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-butanoate (13b). To a solution of compound **12** (93 mg, 0.30 mmol, 1 equiv) and compound **11b** (48 mg, 0.30 mmol, 1 equiv) in THF (7 mL) were added acetic acid (17 µL, 0.30 mmol, 1 equiv) and NaBH(OAc)$_3$ (97 mg, 0.46 mmol, 1.5 equiv). After the mixture was stirred at room temperature overnight, the reaction was quenched by addition of saturated NaHCO$_3$ solution (15 mL), followed by extraction with DCM (4 × 15 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 20:1 + 0.1% NEt$_3$) afforded the desired product as a colorless oil (101 mg, 0.23 mmol, 75%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm =

8.30 (s, 1H), 7.88 (s, 1H), 6.31 (s, 2H), 5.98 (d, J = 3.2 Hz, 1H), 5.43 (dd, J = 6.5, 3.2 Hz, 1H), 4.99 (dd, J = 6.5, 3.2 Hz, 1H), 4.37−4.28 (m, 1H), 2.94−2.78 (m, 2H), 2.67−2.50 (m, 2H), 2.22 (t, J = 7.4 Hz, 2H), 1.77−1.65 (m, 2H), 1.57 (s, 3H), 1.38 (s, 9H), 1.35 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 172.9, 156.0, 153.2, 149.5, 139.9, 120.5, 114.6, 90.9, 85.6, 83.5, 82.4, 80.2, 51.3, 49.1, 33.4, 28.2, 27.4, 25.5, 25.4. FT-IR: ν/cm^{-1} = 2985, 2938, 2852, 1669, 1507, 1426, 1370, 1197, 1178, 1132, 1078, 868, 831, 798, 720. [α]$_D^{20}$ = −22 (10 mg/mL; CHCl$_3$).

tert-Butyl (3-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-propyl)carbamate (13c). To a solution of compound **12** (131 mg, 0.43 mmol, 1 equiv) and compound **11c** (74 mg, 0.43 mmol, 1 equiv) in THF (7 mL) was added acetic acid (24 µL, 0.43 mmol, 1 equiv) and NaBH(OAc)$_3$ (136 mg, 0.64 mmol, 1.5 equiv). After the mixture was stirred at room temperature overnight, the reaction was quenched by addition of saturated NaHCO$_3$ solution (15 mL), followed by extraction with DCM (4 × 15 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 30:1 + 0.1% NEt$_3$) afforded the desired product as a colorless powder (150 mg, 0.32 mmol, 76%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.27 (s, 1H), 7.94 (s, 1H), 6.16 (s, 2H), 6.00 (d, J = 3.2 Hz, 1H), 5.44 (dd, J = 6.4, 3.2 Hz, 1H), 5.33 (t, J = 6.4 Hz, 1H), 5.06 (dd, J = 6.4, 3.2 Hz, 1H), 4.48−4.38 (m, 1H), 3.29−3.11 (m, 2H), 3.02−2.89 (m, 2H), 2.85−2.60 (m, 2H), 1.76−1.63 (m, 2H), 1.59 (s, 3H), 1.41−1.34 (m, 12H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 156.3, 155.8, 153.2, 149.4, 140.1, 120.4, 114.8, 91.1, 85.0, 83.5, 82.4, 79.2, 51.2, 47.7, 39.1, 29.5, 28.5, 27.4, 25.5. FT-IR: ν/cm^{-1} = 3326, 3182, 2980, 2937, 1642, 1598, 1475, 1366, 1250, 1210, 1159, 1076, 867. Mp: 46−49 °C. [α]$_D^{20}$ = −22 (10 mg/mL; CHCl$_3$).

Di-tert-butyl (2S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-amino)pyrrolidine-1,2-dicarboxylate (13d). To a solution of compound **12** (103 mg, 0.34 mmol, 1 equiv) and compound **11d** (96 mg, 0.34 mmol, 1 equiv) in THF (5 mL) was added acetic acid (19 µL, 0.34 mmol, 1 equiv) and NaBH(OAc)$_3$ (107 mg, 0.50 mmol, 1.5 equiv). After the mixture was stirred at room temperature overnight, the reaction was quenched by addition of saturated NaHCO$_3$ solution (15 mL), followed by extraction with DCM (4 × 15 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 30:1 + 0.1% NEt$_3$) afforded the desired product as a colorless oil (130 mg, 0.23 mmol, 68%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.31 (s, 1H), 7.90 (s, 1H), 6.02−5.92 (m, 3H), 5.42−5.35 (m, 1H), 5.08−5.01 (m, 1H), 4.41−4.30 (m, 1H), 4.18−4.04 (m, 1H), 3.78−3.67 (m, 1H), 3.34−3.23 (m, 2H), 2.98−2.88 (m, 2H), 2.49−2.29 (m, 1H), 1.91−1.77 (m, 1H), 1.59 (s, 3H), 1.45−1.35 (m, 21H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 172.4, 155.8, 153.9, 149.4, 140.2, 140.1, 120.5, 114.8, 91.0, 85.2, 83.6, 82.2, 81.5, 80.2, 58.7, 56.5, 51.9, 49.636 36.3, 28.5, 28.1, 27.4, 25.5. FT-IR: ν/cm^{-1} = 3331, 3181, 2978, 2934, 1694, 1643, 1597, 1476, 1366, 1212, 1152, 1075, 855, 769. [α]$_D^{20}$ = −29 (10 mg/mL; CHCl$_3$).

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)amino)butanoic Acid Trifluor-oacetate Salt (14a). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **13a** (72 mg, 0.13 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of Boc and ester groups. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 µL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized. The desired product was obtained as the trifluoroacetate salt (colorless solid, 78 mg, 0.13 mmol, 99%). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.38 (s, 1H), 8.26 (s, 1H), 6.02 (d, J = 5.4 Hz, 1H), 4.74−4.67 (m, 1H), 4.36−4.27 (m, 2H), 3.95 (dd, J = 8.6, 4.8 Hz, 1H), 3.58−3.47 (m, 1H), 3.45−3.36 (m, 1H), 3.34−3.26 (m, 2H), 2.34−2.21 (m, 1H), 2.20−2.07 (m, 1H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/

ppm = 172.7, 153.5, 149.8, 147.6, 144.1, 120.7, 91.1, 81.8, 75.1, 73.2, 53.1, 50.7, 46.5, 28.1. FT-IR: ν/cm^{-1} = 3308, 3166, 3072, 1672, 1508, 1424, 1324, 1198, 1132, 1074, 837, 799, 723, 667. Mp: 58−61 °C. $[\alpha]_D^{20}$ = +2 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 368.15 (100%), 369.14 (17.0%), 370.15 (2.3%). Calculated: 368.17 (100%), 369.17 (18.2%), 370.17 (2.5%). Purity: > 98% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.48 min.

4-((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetra-hydrofuran-2-yl)methyl)amino)butanoic Acid (14b). Cold trifluoro-acetic acid (1.5 mL) was added to a solution of compound **13b** (26 mg, 0.06 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of the ester group. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized to give the desired product as a slightly yellow resin (20 mg, 0.06 mmol, quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.43−8.35 (m, 2H), 6.05 (d, J = 5.0 Hz, 1H), 4.78 (t, J = 4.9 Hz, 1H), 4.39−4.26 (m, 2H), 3.61−3.48 (m, 1H), 3.45−3.35 (m, 1H), 3.09 (t, J = 7.8 Hz, 2H), 2.39 (t, J = 6.9 Hz, 2H), 1.98−1.81 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 176.1, 153.0, 149.8, 146.9, 144.5, 121.2, 91.9, 81.6, 74.8, 73.3, 50.6, 48.6, 31.5, 22.2. FT-IR: ν/cm^{-1} = 3088, 2846, 1667, 1508, 1423, 1323, 1182, 1129, 1047, 919, 834, 798, 721. $[\alpha]_D^{20}$ = +2 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 353.13 (100%), 354.11 (17.5%), 355.13 (2.5%). Calculated: 353.16 (100%), 354.16 (15.6%), 355.16 (2.5%). Purity: > 98% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.85 min.

(2R,3R,4S,5R)-2-(6-Amino-9H-purin-9-yl)-5-(((3-aminopropyl)-amino)methyl)tetrahydrofuran-3,4-diol Trifluoroacetate Salt (14c). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **13c** (46 mg, 0.10 mmol) in DCM (1.5 mL) at 5 °C. The solution was stirred at 5 °C until LC-MS analysis indicated complete removal of the Boc group. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized. The desired product was obtained as the trifluoroacetate salt (colorless solid, 59 mg, 0.10 mmol, quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.37 (s, 1H), 8.35 (s, 1H), 6.05 (d, J = 4.9 Hz, 1H), 4.78 (t, J = 4.9 Hz, 1H), 4.41−4.31 (m, 2H), 3.64−3.53 (m, 1H), 3.48−3.40 (m, 1H), 3.19−3.13 (m, 2H), 3.04−2.96 (m, 2H), 2.11−1.98 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 153.7, 149.9, 148.0, 144.1, 121.1, 91.8, 81.4, 74.8, 73.2, 50.7, 46.1, 37.8, 25.2. FT-IR: ν/cm^{-1} = 3065, 2847, 1666, 1427, 1180, 1123, 1065, 1044, 834, 797, 759, 721, 667. Mp: 66−68 °C. $[\alpha]_D^{20}$ = +4 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 324.17 (100%), 325.14 (16.3%), 326.11 (1.7%). Calculated: m/z: 324.18 (100%), 325.18 (17.0%). Purity: 98% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 1.93 min.

(2S)-4-((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl)methyl)amino)pyrrolidine-2-carboxylic Acid (14d). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **13d** (50 mg, 0.09 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of Boc and ester groups. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized to give the desired product as a yellow resin (34 mg, 0.09 mmol, quant.). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.51 (s, 1H), 8.29 (s, 1H), 8.09 (s, 2H), 6.00 (d, J = 5.8 Hz, 1H), 4.79−4.72 (m, 1H), 4.47−4.37 (m, 1H), 4.26−4.17 (m, 2H), 4.09−3.97 (m, 1H), 3.61−3.49 (m, 2H), 3.48−3.37 (m, 2H), 2.86−2.71 (m, 1H), 2.18−2.01 (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 169.1, 154.3, 150.2, 148.9, 141.4, 119.4, 88.4, 80.1, 72.6, 71.6, 58.0, 54.8, 48.4, 46.4, 30.5. FT-IR: ν/cm^{-1} =

3011, 1669, 1506, 1422, 1180, 1128, 1047, 1024, 997, 827, 798, 763, 720. ESI-MS: m/z [M + H]$^+$ = 380.15 (100%), 381.12 (20.3%), 382.14 (2.6%). Calculated: 380.17 (100%), 381.17 (19.3%), 382.17 (2.7%). Purity: 98% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.75 min.

9-((3aR,4R,6R,6aR)-6-((Benzylamino)methyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine (16a). To a solution of compound **12** (140 mg, 0.46 mmol, 1 equiv) and benzaldehyde (**15a**; 46 μg, 0.46 mmol, 1 equiv) in 1,2-DCE/MeCN (2:1, 9 mL) were added acetic acid (26 μL, 0.46 mmol, 1 equiv) and NaBH(OAc)$_3$ (145 mg, 0.69 mmol, 1.5 equiv). After the mixture was stirred at room temperature overnight, the reaction was quenched by addition of MeOH (2 mL), and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 30:1 + 0.1% NEt$_3$) afforded the desired product as a colorless powder (116 mg, 0.40 mmol, 87%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.07 (s, 1H), 7.85 (s, 1H), 7.30−7.17 (m, 5H), 6.25 (s, 2H), 5.96 (d, J = 3.2 Hz, 1H), 5.44 (dd, J = 6.4, 3.2 Hz, 1H), 5.05 (dd, J = 6.4, 3.2 Hz, 1H), 4.42−4.35 (m, 1H), 3.77 (s, 2H), 2.96−2.81 (m, 2H), 1.58 (s, 3H), 1.35 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 155.9, 153.1, 149.4, 140.0, 139.8, 128.5, 128.2, 127.1, 120.5, 114.6, 91.2, 85.5, 83.4, 82.4, 53.8, 50.7, 27.4, 25.5. FT-IR: ν/cm^{-1} = 3316, 3159, 2987, 2935, 1643, 1596, 1474, 1373, 1329, 1207, 1073, 863, 732, 698. Mp: 57−60 °C. $[\alpha]_D^{20}$ = −24 (10 mg/mL; CHCl$_3$).

3-(((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-methyl)phenol (16b). To a solution of compound **12** (157 mg, 0.51 mmol, 1 equiv) and 3-hydroxybenzaldehyde (**15b**; 63 mg, 0.51 mmol, 1 equiv) in THF (2 mL) were added acetic acid (29 μL, 0.51 mmol, 1 equiv) and NaBH(OAc)$_3$ (163 mg, 0.77 mmol, 1.5 equiv). After the mixture was stirred at room temperature for 7 h, the reaction was quenched by addition of MeOH (1 mL), and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 10:1 + 0.1% NEt$_3$) afforded the desired product as a colorless oil (193 mg, 0.47 mmol, 92%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.22 (s, 1H), 8.05 (s, 1H), 7.04 (t, J = 7.8 Hz, 1H), 6.69−6.59 (m, 3H), 6.12 (d, J = 2.8 Hz, 1H), 5.45 (dd, J = 6.3, 2.8 Hz, 1H), 5.00 (dd, J = 6.3, 3.2 Hz, 1H), 4.35 (td, J = 6.0, 3.2 Hz, 1H), 3.62 (s, 2H), 2.83 (d, J = 6.0 Hz, 2H), 1.57 (s, 3H), 1.35 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 158.6, 157.3, 153.9, 150.2, 142.0, 141.6, 130.4, 120.7, 120.4, 116.2, 115.5, 115.2, 91.8, 86.9, 84.8, 83.9, 54.2, 51.5, 27.5, 25.6. FT-IR: ν/cm^{-1} = 3543, 3339, 2982, 1645, 1600, 1480, 1457, 1424, 1376, 1331, 1050, 969, 869, 788, 696. Mp: 247 °C (decomposition). $[\alpha]_D^{20}$ = −13 (10 mg/mL; CHCl$_3$).

9-((3aR,4R,6R,6aR)-2,2-Dimethyl-6-(((3-nitrobenzyl)amino)-methyl)tetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine (16c). To a solution of compound **12** (402 mg, 1.31 mmol, 1 equiv) and 3-nitrobenzaldehyde (**15c**; 198 mg, 1.31 mmol, 1 equiv) in THF (10 mL) were added acetic acid (75 μL, 1.31 mmol, 1 equiv) and NaBH(OAc)$_3$ (417 mg, 1.97 mmol, 1.5 equiv). After the mixture was stirred at room temperature overnight, the reaction was quenched by addition of MeOH (15 mL), and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 40:1 + 0.1% NEt$_3$) afforded the desired product as a colorless oil (358 mg, 0.81 mmol, 62%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.19−8.13 (m, 1H), 8.11−8.03 (m, 2H), 7.86 (s, 1H), 7.60 (d, J = 7.8 Hz, 1H), 7.43 (t, J = 7.8 Hz, 1H), 6.26 (s, 2H), 5.96 (d, J = 3.2 Hz, 1H), 5.48 (dd, J = 6.4, 3.2 Hz, 1H), 5.08 (dd, J = 6.4, 3.2 Hz, 1H), 4.45−4.34 (m, 1H), 3.90−3.84 (m, 2H), 2.99−2.81 (m, 2H), 1.60 (s, 3H), 1.38 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 155.9, 152.9, 149.3, 148.5, 142.5, 140.2, 134.1, 129.3, 122.9, 122.2, 120.4, 114.7, 91.3, 85.6, 83.3, 82.3, 53.0, 50.7, 27.4, 25.5. FT-IR: ν/cm^{-1} = 3323, 3159, 2982, 2935, 1643, 1597, 1525, 1475, 1348, 1207, 1074, 867, 798, 751. $[\alpha]_D^{20}$ = −16 (10 mg/mL; CHCl$_3$).

tert-Butyl 3-(((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-methyl)benzoate (16d). To a solution of compound **12** (200 mg, 0.65 mmol, 1 equiv) and *tert*-butyl 3-formylbenzoate (**15d**; 135 mg, 0.65 mmol, 1 equiv) in THF (5 mL) was added acetic acid (37 μL, 0.65 mmol, 1 equiv) and NaBH(OAc)$_3$ (208 mg, 0.98 mmol, 1.5 equiv). After the mixture was stirred at room temperature overnight, the

reaction was quenched by addition of saturated NaHCO$_3$ solution (25 mL), followed by extraction with DCM (4 × 20 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 30:1 + 0.1% NEt$_3$) afforded the desired product as a yellowish oil (177 mg, 0.36 mmol, 55%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.12 (s, 1H), 7.92−7.89 (m, 1H), 7.88−7.83 (m, 2H), 7.49−7.43 (m, 1H), 7.33 (t, J = 7.6 Hz, 1H), 6.06−5.93 (m, 3H), 5.46 (dd, J = 6.4, 3.3 Hz, 1H), 5.07 (dd, J = 6.4, 3.3 Hz, 1H), 4.43−4.36 (m, 1H), 3.83 (s, 2H), 2.99−2.82 (m, 2H), 1.60 (s, 3H), 1.57 (s, 9H), 1.37 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 165.9, 155.8, 153.2, 149.5, 140.1, 132.3, 132.2, 129.2, 128.4, 128.3, 120.6, 114.7, 91.2, 85.5, 83.4, 82.4, 81.1, 53.5, 50.7, 28.3, 27.5, 25.6. FT-IR: ν/cm^{-1} = 3322, 3159, 2980, 2933, 1706, 1643, 1475, 1368, 1294, 1157, 1076, 909, 850, 729. $[\alpha]_D^{20}$ = −18 (10 mg/mL; CHCl$_3$).

tert-Butyl (3-(((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-methyl)phenyl)carbamate (16e). To a solution of compound **12** (151 mg, 0.49 mmol, 1 equiv) and *tert*-butyl (3-formylphenyl)carbamate (**15e**) (109 mg, 0.49 mmol, 1 equiv) in 1,2-DCE (5 mL) were added acetic acid (28 μL, 0.49 mmol, 1 equiv) and NaBH(OAc)$_3$ (157 mg, 0.74 mmol, 1.5 equiv). After the mixture was stirred at room temperature overnight, the reaction was quenched by addition of MeOH (2 mL), and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography (DCM/MeOH = 30:1 + 0.1% NEt$_3$) afforded the desired product as a colorless powder (214 mg, 0.42 mmol, 85%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.15 (s, 1H), 7.95 (s, 1H), 7.33 (s, 1H), 7.25−7.15 (m, 2H), 6.95−6.89 (m, 1H), 6.01 (d, J = 3.1 Hz, 1H), 5.94 (s, 2H), 5.44 (dd, J = 6.4, 3.1 Hz, 1H), 5.05 (dd, J = 6.4, 3.1 Hz, 1H), 4.45−4.38 (m, 1H), 3.85−3.71 (m, 2H), 2.95−2.80 (m, 2H), 1.60 (s, 3H), 1.50 (s, 9H), 1.38 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 155.8, 153.3, 153.0, 149.5, 140.6, 140.1, 138.8, 129.1, 122.8, 120.5, 118.4, 117.6, 114.6, 91.3, 85.8, 83.7, 82.5, 80.6, 53.8, 50.7, 28.5, 27.4, 25.5. FT-IR: ν/cm^{-1} = 3319, 3178, 2977, 2977, 2934, 1716, 1643, 1594, 1367, 1239, 1156, 1072, 867, 733. Mp: 68−71 °C. $[\alpha]_D^{20}$ = −12 (10 mg/mL; CHCl$_3$).

(2R,3R,4S,5R)-2-(6-Amino-9H-purin-9-yl)-5-((benzylamino)-methyl)tetrahydrofuran-3,4-diol Trifluoroacetate Salt (17a). Cold trifluoroacetic acid (250 μL) was added to a suspension of compound **16a** (29 mg, 0.07 mmol) in water (1.5 mL) at 5 °C. The mixture was stirred at 5 °C for 48 h. Water (5 mL) was added, and the mixture was lyophilized to give the desired product as the trifluoroacetate salt (colorless solid, 34 mg, 0.07 mmol, quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.36 (s, 1H), 8.16 (s, 1H), 7.42 (s, 5H), 6.06 (d, J = 4.7 Hz, 1H), 4.76 (t, J = 4.7 Hz, 1H), 4.43−4.35 (m, 2H), 4.26 (d, J = 2.8 Hz, 2H), 3.56 (dd, J = 13.0, 9.2 Hz, 1H), 3.43 (dd, J = 13.0, 2.7 Hz, 1H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 154.1, 149.8, 148.5, 143.9, 132.1, 131.2, 130.8, 130.3, 121.2, 92.0, 81.5, 74.8, 73.3, 52.4, 49.9. FT-IR: ν/cm^{-1} = 3326, 3186, 2800, 1644, 1574, 1459, 1423, 1327, 1206, 1067, 825, 797, 699. Mp: 61−63 °C. $[\alpha]_D^{20}$ = +6 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 357.16 (100%), 358.12 (20.8%), 359.14 (2.0%). Calculated: 357.17 (100%), 358.17 (18.7%), 359.17 (2.6%). Purity: 97% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.44 min.

(2R,3R,4S,5R)-2-(6-Amino-9H-purin-9-yl)-5-(((3-hydroxybenzyl)-amino)methyl)tetrahydrofuran-3,4-diol Trifluoroacetate Salt (17b). Cold trifluoroacetic acid (250 μL) was added to a suspension of compound **16b** (37 mg, 0.09 mmol) in water (1.5 mL) at 5 °C. The mixture was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized to give the desired product as the trifluoroacetate salt (yellowish solid, 43 mg, 0.09 mmol, 99%). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.35 (s, 1H), 8.18 (s, 1H), 7.16 (t, J = 8.0 Hz, 1H), 6.85−6.74 (m, 3H), 6.02 (d, J = 4.6 Hz, 1H), 4.70 (t, J = 4.6 Hz, 1H), 4.37−4.26 (m, 2H), 4.11 (s, 2H), 3.52−3.31 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 159.3, 153.0, 149.7, 146.8, 144.4, 133.3, 131.4, 121.8, 121.2, 117.9, 117.6, 92.1, 81.5, 74.8, 73.3, 52.3, 49.8. FT-IR: ν/cm^{-1} = 3087, 2840, 1669, 1592, 1427, 1195, 1129, 1040, 834, 797, 721, 697. Mp: 61 °C (decomposition). $[\alpha]_D^{20}$ = +4 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 373.17 (100%), 374.14 (20.1%), 375.17 (2.6%). Calculated: 373.16 (100%), 374.17 (18.8%),

375.17 (2.5%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 4.06 min.

(2R,3R,4S,5R)-2-(6-Amino-9H-purin-9-yl)-5-(((3-nitrobenzyl)-amino)methyl)tetrahydrofuran-3,4-diol Trifluoroacetate Salt (17c). Cold trifluoroacetic acid (500 μL) was added to a suspension of compound **16c** (96 mg, 0.22 mmol) in water (3 mL) at 5 °C. The mixture was stirred at 5 °C overnight. Undissolved residues were filtered off, and water (5 mL) was added to the filtrate. The mixture was lyophilized to give the desired product as a colorless oil (103 mg, 0.20 mmol, 91%). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.37 (s, 1H), 8.32−8.29 (m, 1H), 8.27 (s, 1H), 8.23−8.17 (m, 1H), 7.85−7.77 (m, 1H), 7.61 (t, J = 8.0 Hz, 1H), 6.04 (d, J = 4.5 Hz, 1H), 4.70 (t, J = 4.5 Hz, 1H), 4.44−4.28 (m, 4H), 3.61−3.41 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 153.9, 149.8 (2x C), 148.2, 143.9, 137.4, 134.3, 131.5, 126.03, 125.4, 121.1, 99.0, 81.3, 74.8, 73.3, 51.4, 50.3. FT-IR: ν/cm^{-1} = 3326, 3201, 3090, 1672, 1531, 1509, 1482, 1427, 1354, 1199, 1134, 834, 799, 722. $[\alpha]_D^{20}$ = +14 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 402.14 (100%), 403.1 (21.2%), 404.17 (2.7%). Calculated: 402.15 (100%), 403.16 (18.8%), 404.16 (2.7%). Purity: 99% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.37 min.

3-(((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxyte-trahydrofuran-2-yl)methyl)amino)methyl)benzoic Acid Trifluoroa-cetate Salt (17d). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **16d** (79 mg, 0.16 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of the ester group. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized. The desired product was obtained as the trifluoroacetate salt (colorless solid, 82 mg, 0.16 mmol, quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.34 (s, 1H), 8.21 (s, 1H), 8.04 (d, J = 1.6 Hz, 1H), 7.98 (dt, J = 7.7, 1.6 Hz, 1H), 7.61 (dt, J = 7.7, 1.6 Hz, 1H), 7.46 (t, J = 7.7 Hz, 1H), 6.02 (d, J = 4.5 Hz, 1H), 4.68 (t, J = 4.6 Hz, 1H), 4.37−4.24 (m, 4H), 3.55−3.36 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 168.8, 153.0, 149.7, 146.7, 144.4, 135.6, 133.0, 132.7, 132.3, 131.8, 130.5, 121.2, 92.0, 81.4, 74.9, 73.3, 52.0, 50.1. FT-IR: ν/cm^{-1} = 3076, 3076, 2847, 1667, 1509, 1427, 1280, 1182, 1129, 1042, 826, 797, 749, 721. mp: 65−67 °C. $[\alpha]_D^{20}$ = +10 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 401.14 (100%), 402.12 (22.2%), 403.14 (2.5%). Calculated: 401.16 (100%), 402.16 (19.9%), 403.16 (3.3%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.91 min.

(2R,3R,4S,5R)-2-(6-Amino-9H-purin-9-yl)-5-(((3-aminobenzyl)-amino)methyl)tetrahydrofuran-3,4-diol Trifluoroacetate Salt (17e). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **16e** (108 mg, 0.21 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of the Boc group. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized. The desired product was obtained as the trifluoroacetate salt (1:1.3; ocherous solid, 111 mg, 0.21 mmol, quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.33 (s, 1H), 8.18 (s, 1H), 7.39−7.32 (m, 1H), 7.25−7.20 (m, 2H), 7.19−7.14 (m, 1H), 6.01 (d, J = 4.7 Hz, 1H), 4.72−4.67 (m, 1H), 4.36−4.28 (m, 2H), 4.25−4.14 (m, 2H), 3.56−3.45 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 153.4, 149.7, 147.4, 144.2, 140.0, 133.9, 131.6, 126.8, 122.4, 122.0, 121.1, 92.0, 81.4, 74.8, 73.3, 52.0, 50.1. FT-IR: ν/cm^{-1} = 2979, 2637, 1666, 1507, 1465, 1427, 1188, 1128, 1041, 836, 797, 721, 695. Mp: 69−72 °C. $[\alpha]_D^{20}$ = +6 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 372.16 (100%), 373.13 (20.9%), 374.15 (2.2%). Calculated: 372.18 (100%), 373.18 (21.3%), 374.19 (1.7%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.88 min.

2-(((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)isoindoline-1,3-dione (18). According to the method of Liu et al.,[77] to a suspension

of compound **5** (1000 mg, 3.25 mmol, 1 equiv) in THF (40 mL) were added phthalimide (479 mg, 3.25 mmol, 1 equiv) and triphenylphosphine (853 mg, 3.25, 1 equiv). DIAD (639 μL, 3.25 mmol, 1 equiv) was added dropwise, and the mixture was stirred at room temperature for 5 h. The precipitate was filtered off and washed with cyclohexane to give the desired product as a colorless solid. More product could be obtained by adding cyclohexane to the filtrate and stirring the mixture overnight. The precipitate was filtered off and washed with cyclohexane. Yield: 1.01 g, 2.31 mmol, 71%. ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.28 (s, 1H), 7.86 (s, 1H), 7.81 (s, 4H), 7.31 (s, 2H), 6.18 (d, J = 2.0 Hz, 1H), 5.44 (dd, J = 6.3, 2.0 Hz, 1H), 5.18 (dd, J = 6.3, 3.7 Hz, 1H), 4.43−4.33 (m, 1H), 3.97−3.83 (m, 2H), 1.50 (s, 3H), 1.30 (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 167.8, 156.1, 152.5, 148.6, 140.3, 134.5, 131.4, 123.1, 119.2, 113.5, 88.7, 83.8, 83.4, 81.8, 27.0, 25.3. FT-IR: ν/cm^{-1} = 3326, 3161, 2979, 1709, 1661, 1597, 1397, 1327, 1208, 1074, 869, 798, 725. Mp: 147−150 °C. $[\alpha]_D^{20}$ = +18 (10 mg/mL; DMSO).

*2-(((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)isoindoline-1,3-dione (**19**).* Cold trifluoroacetic acid (250 μL) was added to a suspension of compound **18** (40 mg, 0.09 mmol) in water (1.5 mL) at 5 °C. The mixture was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized to give the desired product as a colorless solid (35 mg, 0.09 mmol, 96%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.67 (s, 1H), 8.27 (s, 1H), 7.90−7.79 (m, 4H), 5.92 (d, J = 5.3 Hz, 1H), 4.74 (t, J = 5.3 Hz, 1H), 4.24 (t, J = 4.4 Hz, 1H), 4.02−3.85 (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 167.9, 158.5, 151.9, 148.6, 147.2, 142.0, 134.6, 131.5, 123.2, 113.0, 87.9, 81.8, 73.1, 71.5. FT-IR: ν/cm^{-1} = 3318, 3120, 2944, 1679, 1395, 1188, 1127, 1029, 1000, 972, 874, 797, 738. mp: 143−147 °C. $[\alpha]_D^{20}$ = +32 (10 mg/mL; DMSO). ESI-MS: m/z [M + H]$^+$ = 397.12 (100%), 398.06 (19.7%), 399.11 (2.5%). Calculated: 397.13 (100%), 398.13 (19.9%), 399.13 (3.3%). Purity: 97% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 4.20 min.

*tert-Butyl N^4-(((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-N^2-(tert-butoxycarbonyl)-L-asparaginate (**20a**).* To a solution of compound **12** (200 mg, 0.65 mmol, 1 equiv) and DIPEA (111 μL, 0.65 mmol, 1 equiv) in DMF (2 mL) was added HBTU (248 mg, 0.65 mmol, 1 equiv), and the mixture was stirred at room temperature for 15 min. Boc-Asp-O^{t}Bu (189 mg, 0.65 mmol, 1 equiv) was added, and the mixture was stirred for 1 h. The solution was diluted with saturated NaHCO$_3$ solution and extracted with ethyl acetate (3 × 20 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C to give the desired product as a yellowish solid (352 mg, 0.61 mmol, 93%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.39 (s, 1H), 7.86 (s, 1H), 6.27 (s, 2H), 5.80 (d, J = 4.8 Hz, 1H), 5.73 (d, J = 8.0 Hz, 1H), 5.25 (t, J = 5.7 Hz, 1H), 4.80 (dd, J = 6.3, 2.1 Hz, 1H), 4.50−4.41 (m, 2H), 4.22−4.08 (m, 1H), 3.27−3.17 (m, 1H), 2.96−2.89 (m, 2H), 1.60 (s, 3H), 1.45 (s, 9H), 1.37−1.30 (m, 12H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 170.7, 170.6, 155.9, 155.8, 152.7, 148.9, 140.7, 120.1, 114.9, 92.8, 83.5, 82.3, 82.2, 81.6, 79.8, 51.3, 41.1, 38.4, 28.3, 28.1, 27.6, 25.4. FT-IR: ν/cm^{-1} = 3323, 2980, 2932, 1709, 1643, 1480, 1367, 1296, 1249, 1153, 1096, 1077, 852, 798. Mp: 85 °C (decomposition). $[\alpha]_D^{20}$ = −136 (10 mg/mL; CHCl$_3$).

*tert-Butyl N^5-(((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-N^2-(tert-butoxycarbonyl)-L-glutaminate (**20b**).* To a solution of compound **12** (200 mg, 0.65 mmol, 1 equiv) and DIPEA (111 μL, 0.65 mmol, 1 equiv) in DMF (2 mL) was added TBTU (210 mg, 0.65 mmol, 1 equiv), and the mixture was stirred at room temperature for 40 min. Boc-Glu-OtBu (198 mg, 0.65 mmol, 1 equiv) was added, and the mixture was stirred for 2.5 h. The solution was diluted with saturated NaHCO$_3$ solution and extracted with ethyl acetate (3 × 20 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C. The residues were dissolved in DCM (2 mL), and *n*-heptane (20 mL) was added. The resulting precipitate was filtered off and washed with *n*-heptane to give the desired product as a colorless solid (322 mg, 0.54 mmol, 84%). ^{1}H

NMR (300 MHz, CDCl$_3$): δ/ppm = 8.33 (s, 1H), 8.18 (s, 1H), 8.13 (t, J = 5.8 Hz, 1H), 7.36 (s, 2H), 7.10 (d, J = 7.7 Hz, 1H), 6.11 (d, J = 3.1 Hz, 1H), 5.41 (dd, J = 6.4, 3.1 Hz, 1H), 4.88 (dd, J = 6.4, 3.1 Hz, 1H), 4.22−4.14 (m, 1H), 3.83−3.72 (m, 1H), 3.37−3.32 (m, 2H), 2.25−2.13 (m, 2H), 1.97−1.81 (m, 1H), 1.80−1.64 (m, 1H), 1.53 (s, 3H), 1.40−1.34 (m, 18H), 1.31 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 171.7, 171.6, 156.2, 155.5, 152.7, 148.7, 140.1, 119.3, 113.5, 89.2, 84.1, 82.7, 81.7, 80.2, 78.0, 53.9, 40.8, 31.5, 28.2, 27.6, 27.0, 26.5, 25.2. FT-IR: ν/cm^{-1} = 3323, 3212, 2980, 2932, 1709, 1643, 1480, 1367, 1213, 1153, 1096, 852, 798. Mp: 83−86 °C. $[\alpha]_D^{20}$ = −91 (10 mg/mL; CHCl$_3$).

*tert-Butyl 4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-4-oxobutanoate (**20c**).* To a solution of compound **12** (204 mg, 0.67 mmol, 1 equiv) and DIPEA (113 μL, 0.67 mmol, 1 equiv) in DMF (2 mL) was added TBTU (214 mg, 0.67 mmol, 1 equiv), and the mixture was stirred at room temperature for 40 min. 4-(tert-Butoxy)-4-oxobutanoic acid (116 mg, 0.67 mmol, 1 equiv) was added, and the mixture was stirred for 2.5 h. The solution was diluted with saturated NaHCO$_3$ solution and extracted with ethyl acetate (3 × 20 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C to give the desired product as a yellowish solid (300 mg, 0.65 mmol, 97%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.49 (d, J = 8.9 Hz, 1H), 8.35 (s, 1H), 7.85 (s, 1H), 6.17 (s, 2H), 5.80 (d, J = 5.0 Hz, 1H), 5.38−5.32 (m, 1H), 4.84 (dd, J = 6.2, 2.0 Hz, 1H), 4.52−4.44 (m, 1H), 4.26−4.13 (m, 1H), 3.27−3.14 (m, 1H), 2.82−2.61 (m, 2H), 2.59−2.44 (m, 2H), 1.60 (s, 3H), 1.40 (s, 9H), 1.34 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 172.4, 172.3, 156.2, 152.9, 149.0, 140.6, 121.2, 114.6, 92.9, 83.7, 82.3, 81.7, 80.6, 41.1, 30.9, 30.5, 28.2, 27.7, 25.3. FT-IR: ν/cm^{-1} = 3326, 3186, 2980, 2926, 1722, 1642, 1598, 1367, 1332, 1151, 1095, 852, 798. Mp: 62−65 °C. $[\alpha]_D^{20}$ = −164 (10 mg/mL; CHCl$_3$).

*tert-Butyl (3-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-3-oxopropyl)carbamate (**20d**).* To a solution of compound **12** (200 mg, 0.65 mmol, 1 equiv) and DIPEA (111 μL, 0.65 mmol, 1 equiv) in DMF (2 mL) was added TBTU (210 mg, 0.65 mmol, 1 equiv), and the mixture was stirred at room temperature for 30 min. Boc-β-Ala-OH (124 mg, 0.65 mmol, 1 equiv) was added, and the mixture was stirred overnight. The solution was diluted with saturated NaHCO$_3$ solution and extracted with ethyl acetate (3 × 20 mL). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C to give the desired product as a colorless solid (298 mg, 0.62 mmol, 96%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.36 (s, 1H), 8.22 (d, J = 8.2 Hz, 1H), 7.87 (s, 1H), 6.42 (s, 2H), 5.83 (d, J = 4.5 Hz, 1H), 5.46−5.36 (m, 1H), 5.30−5.23 (m, 1H), 4.80 (dd, J = 6.3, 2.4 Hz, 1H), 4.49−4.42 (m, 1H), 4.17−4.03 (m, 1H), 3.52−3.40 (m, 2H), 3.34−3.23 (m, 1H), 2.67−2.46 (m, 2H), 1.60 (s, 3H), 1.35 (s, 9H), 1.32 (s, 3H). ^{13}C NMR (75 MHz, CDCl$_3$): δ/ppm = 172.4, 156.2, 156.0, 152.8, 148.9, 140.6, 121.1, 114.9, 92.5, 83.6, 82.5, 81.5, 79.3, 41.1, 36.9, 36.3, 28.5, 27.6, 25.4. FT-IR: ν/cm^{-1} = 3323, 3207, 2980, 2935, 1643, 1600, 1507, 1366, 1247, 1210, 1164, 1077, 854, 750. Mp: 98−101 °C. $[\alpha]_D^{20}$ = −132 (10 mg/mL; CHCl$_3$).

*N^4-(((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-L-asparagine Trifluoroacetate Salt (**21a**).* Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **20a** (106 mg, 0.18 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of Boc and ester groups. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized. The desired product was obtained as the trifluoroacetate salt (colorless solid, 89 mg, 0.18 mmol, quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.35 (s, 1H), 8.29 (s, 1H), 5.91 (d, J = 5.3 Hz, 1H), 4.64 (t, J = 5.3 Hz, 1H), 4.22−4.12 (m, 2H), 4.08−4.00 (m, 1H), 3.61−3.40 (m, 2H), 2.94−2.76 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 171.4, 171.3, 153.2, 149.9, 147.1, 144.1, 120.8, 90.8, 85.0, 75.1, 72.8, 51.0, 42.3, 35.4. FT-IR:

ν/cm^{-1} = 3087, 2981, 1664, 1507, 1423, 1323, 1186, 1129, 1069, 899, 836, 798, 721. Mp: 72 °C (decomposition). $[\alpha]_D^{20}$ = −20 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 382.13 (100%), 383.12 (17.2%), 384.13 (2.1%). Calculated: 382.15 (100%), 383.15 (15.6%), 384.15 (2.7%). Purity: 95% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.68 min.

N^5-(((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxyte-trahydrofuran-2-yl)methyl)-ʟ-glutamine Trifluoroacetate Salt (21b). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **20b** (101 mg, 0.17 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of Boc and ester groups. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized. The desired product was obtained as the trifluoroacetate salt (yellowish solid, 87 mg, 0.17 mmol quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.36 (s, 1H), 8.30 (s, 1H), 5.91 (d, J = 5.4 Hz, 1H), 4.66−4.61 (m, 1H), 4.16−4.10 (m, 1H), 4.08−4.00 (m, 1H), 3.95−3.88 (m, 1H), 3.59−3.39 (m, 2H), 2.46−2.38 (m, 2H), 2.19−1.99 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 174.5, 171.6, 152.8, 150.0, 146.5, 144.2, 120.9, 90.9, 85.1, 75.2, 72.9, 53.6, 42.5, 32.4, 27.2. FT-IR: ν/cm^{-1} = 3082, 2945, 1667, 1556, 1507, 1424, 1323, 1184, 1129, 898, 836, 797, 721. Mp: 77−80 °C. $[\alpha]_D^{20}$ = −17 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 396.15 (100%), 397.14 (18%), 398.15 (2.6%). Calculated: 396.16 (100%), 397.17 (16.7%), 398.17 (2.5%). Purity: 96% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.22 min.

4-((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetra-hydrofuran-2-yl)methyl)amino)-4-oxobutanoic Acid (21c). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **20b** (50 mg, 0.11 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of the ester group. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized to give the desired product as a colorless solid (40 mg, 0.11 mmol, quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.40−8.28 (m, 2H), 5.94 (d, J = 5.9 Hz, 1H), 4.77−4.71 (m, 1H), 4.21−4.10 (m, 2H), 3.77−3.66 (m, 1H), 3.41 (dd, J = 14.3, 4.1 Hz, 1H), 2.63−2.55 (m, 2H), 2.55−2.47 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 175.0, 174.9, 154.6, 150.1, 149.5, 143.5, 120.9, 90.8, 85.6, 74.9, 72.7, 42.1, 31.6, 30.0. FT-IR: ν/cm^{-1} = 3305, 3102, 2935, 1646, 1553, 1420, 1195, 1129, 1059, 834, 798, 721. Mp: 64−67 °C. $[\alpha]_D^{20}$ = −64 (5 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 367.12 (100%), 368.09 (16.8%), 369.10 (2.3%). Calculated: 367.14 (100%), 368.14 (15.6%), 369.14 (2.7%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 4.19 min.

3-Amino-N-(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihy-droxytetrahydrofuran-2-yl)methyl)propanamide Trifluoroacetate Salt (21d). Cold trifluoroacetic acid (1.5 mL) was added to a solution of compound **20d** (83 mg, 0.17 mmol) in DCM (1.5 mL) at −20 °C. The solution was stirred at −20 °C until LC-MS analysis indicated complete removal of Boc and ester groups. DCM (40 mL) was added, and the solvent was removed under reduced pressure at 40 °C. After co-distillation with DCM (3 × 40 mL), the residues were dissolved in water (1.5 mL) and cooled down to 5 °C. Cold trifluoroacetic acid (250 μL) was added, and the solution was stirred at 5 °C overnight. Water (5 mL) was added, and the mixture was lyophilized. The desired product was obtained as the trifluoroacetate salt (yellow resin, 78 mg, 0.17 mmol quant.). ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.44 (s, 1H), 8.38 (s, 1H), 5.99 (d, J = 5.2 Hz, 1H), 4.75−4.68 (m, 1H), 4.25−4.18 (m, 1H), 4.16−4.08 (m, 1H), 3.65−3.51 (m, 2H), 3.20−3.14 (m, 2H), 2.66−2.58 (m, 2H). ^{13}C NMR (75 MHz, CD$_3$OD): δ/ppm = 171.1, 152.2, 148.7, 146.4, 142.4, 119.5, 89.5, 83.6, 73.7, 71.5, 41.0, 35.7, 31.4. FT-IR: ν/cm^{-1} = 3271, 3088, 1666, 1557, 1507, 1423, 1325, 1183, 1126, 1074, 835, 798, 721. $[\alpha]_D^{20}$ = −23 (10 mg/mL; MeOH). ESI-MS: m/z [M +

H]$^+$ = 338.13 (100%), 339.12 (15.5%), 340.11 (2.0%). Calculated: 338.16 (100%), 339.16 (14.4%), 340.16 (2.1%). Purity: 95% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 4.02 min.

N-(((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-N-methyl-2-nitrobenzenesulfonamide (22). To a solution of 2′,3′-O-isopropyli-denadenosine (**5**) (0.5 g, 1.6 mmol, 1.0 equiv) in THF (12 mL) at 0 °C were added 2-nitro-N-methylbenzenesulfonamide (387 mg, 1.8 mmol, 1.1 equiv), PPh$_3$ (726 mg, 2.8 mmol, 1.7 equiv), and DIAD (0.54 mL, 560 mg, 2.8 mmol, 1.7 equiv), and the mixture was stirred at rt for 24 h. The reaction was concentrated, and the residue was recrystallized in hot MeOH (25 mL) and further purified by column chromatography (DCM/MeOH = 25:1) to give **22** (192 mg, 0.38 mmol, 23%) as a yellow solid. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.30 (d, J = 9.0 Hz, 2H), 8.17 (s, 1H), 7.92 (dd, J = 8.3, 1.4 Hz, 1H), 7.85−7.78 (m, 2H), 7.76−7.66 (m, 1H), 7.36 (s, 2H), 6.21 (d, J = 2.3 Hz, 1H), 5.45 (dd, J = 6.4, 2.3 Hz, 1H), 5.08 (dd, J = 6.4, 3.4 Hz, 1H), 4.40−4.27 (m, 1H), 3.63 (dd, J = 14.5, 5.3 Hz, 1H), 3.51−3.40 (m, 1H), 2.77 (s, 3H), 1.53 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 156.1, 152.7, 148.7, 147.8, 140.3, 134.5, 132.3, 130.4, 129.8, 124.3, 119.2, 113.6, 89.0, 84.4, 83.2, 81.8, 51.4, 35.7, 27.0, 25.2. FT-IR: ν/cm^{-1} = 3114, 1543, 1372, 1347, 1217, 1205, 1170, 1103, 1072, 1057, 972, 877, 854, 777, 767, 693. Mp: 230−233 °C. $[\alpha]_D^{20}$ = −10 (10 mg/mL; MeOH). R_f = 0.25 (DCM/MeOH = 20:1).

9-((3aR,4R,6R,6aR)-2,2-Dimethyl-6-((methylamino)methyl)-tetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine (23). To a solution of **22** (172 mg, 0.34 mmol, 1.0 equiv) in MeCN (3 mL) were added Cs$_2$CO$_3$ (332 mg, 1.02 mmol, 3.0 equiv) and thiophenol (0.14 mL, 1.36 mmol, 4.0 equiv), and the suspension was stirred at rt for 72 h. By the addition of 1 M NaOH (20 mL), the reaction was quenched, and the aqueous layer was extracted with DCM (3 × 20 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 90:10 to 70:30) to give **23** (71 mg, 0.22 mmol, 65%) as a yellow oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.26 (s, 1H), 7.87 (s, 1H), 6.56 (s, 2H), 5.97 (d, J = 3.3 Hz, 1H), 5.42 (dd, J = 6.6, 3.2 Hz, 1H), 5.01−4.94 (m, 1H), 4.37−4.30 (m, 1H), 2.83 (d, J = 4.9 Hz, 2H), 2.37 (s, 3H), 1.55 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 156.0, 153.1, 149.3, 139.8, 120.3, 114.6, 90.9, 85.2, 83.5, 82.4, 53.5, 36.4, 27.3, 25.4. FT-IR: ν/cm^{-1} = 3163, 2986, 1643, 1596, 1474, 1374, 1328, 1296, 1265, 1208, 1155, 1074, 866, 798, 732, 702. $[\alpha]_D^{20}$ = −17 (10 mg/mL; MeOH). R_f = 0.11 (DCM/MeOH = 10:1).

tert-Butyl (tert-Butoxycarbonyl)-ʟ-homoserinate (24). To a solution of Boc-Asp-O^{t}Bu (2.5 g, 8.7 mmol, 1.0 equiv) in dry THF (100 mL) under argon atmosphere at −10 °C was added N-methyl morpholine (0.95 mL, 875 mg, 8.7 mmol, 1.0 equiv). After 1 min, ethyl chloroformate (0.82 mL, 938 mg, 8.7 mmol, 1.0 equiv) was added, and the mixture was stirred for 15 min at −5 °C. Then, the precipitates were filtered off, and the filtrate was dropwise added to a 5 °C chilled solution of NaBH$_4$ (0.67 mL, 736 mg, 19.5 mmol, 2.3 equiv) in water. The reaction was slowly warmed up to rt and was stirred at rt for 18 h. The mixture was cooled to 5 °C and acidified to pH 2 with 3 M HCl. Then, the aqueous phase was extracted with EtOAc (2 × 25 mL), and the combined organic layers were washed with brine (3 × 25 mL), dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (cyclohexane/EtOAc = 1:1) to give **24** (1.73 g, 6.3 mmol, 73%) as a colorless oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 5.48−5.23 (m, 1H), 4.45−4.22 (m, 1H), 3.76−3.55 (m, 2H), 3.33 (s, 1H), 2.20−2.04 (m, 1H), 1.44 (d, J = 6.4 Hz, 18H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 172.1, 156.8, 82.4, 80.5, 58.4, 51.0, 36.7, 31.3, 28.4, 28.4, 28.1, 28.0. FT-IR: ν/cm^{-1} = 3377, 2979, 1715, 1505, 1456, 1392, 1367, 1251, 1154, 1054, 846. $[\alpha]_D^{20}$ = −39 (10 mg/mL; MeOH). R_f = 0.40 (cyclohexane/EtOAc = 2:1).

tert-Butyl 2-((tert-Butoxycarbonyl)amino)-4-iodobutanoate (25). To a solution of **24** (0.5 g, 1.8 mmol, 1.0 equiv) in 1,2-dce at 0 °C were added DMAP (22 mg, 0.18 mmol, 0.1 equiv), tosyl chloride (693 mg, 3.6 mmol, 2.0 equiv), and NEt$_3$ (1.26 mL, 9.1 mmol, 5.0 equiv). After the mixture was stirred at rt for 4 h, the solvent was removed under

reduced pressure at 40 °C. The obtained residue was dissolved in EtOAc (20 mL) and washed with 1 M HCl (3 × 20 mL), saturated NaHCO$_3$ solution (3 × 20 mL), and brine (3 × 20 mL). The organic layer was dried over MgSO$_4$, filtered, and concentrated under reduced pressure at 40 °C to give a yellow oil. To a solution of the oil in acetone (5 mL) was added NaI (4.1 g, 27.3 mmol, 3.0 equiv), and the solution was stirred at rt for 42 h under light exclusion. The reaction was quenched by the addition of 10 mL of ice water and then extracted with EtOAc (3 × 15 mL). The organic layer was washed with brine (3 × 15 mL), dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The obtained oil was purified via column chromatography to give **25** (462 mg, 1.2 mmol, 66%) as a brown oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 5.15−5.00 (m, 1H), 4.18 (d, J = 5.1 Hz, 1H), 3.25−3.07 (m, 2H), 2.44−2.28 (m, 1H), 2.18−2.07 (m, 1H), 1.46 (d, J = 8.0 Hz, 18H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 170.7, 155.4, 82.7, 80.2, 55.1, 37.8, 28.5, 28.1, −0.4. FT-IR: ν/cm^{-1} = 2978, 2934, 1714, 1504, 1455, 1392, 1367, 1251, 1153, 1046, 1022, 846. $[\alpha]_D^{20}$ = −13 (10 mg/mL; MeOH). R_f = 0.61 (cyclohexane/EtOAc = 2:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-(methyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (26a). To a solution of **23** (62 mg, 0.19 mmol, 1.0 equiv) in dry MeCN (1 mL) was added dropwise a solution of **25** (90 mg, 0.23 mmol, 1.2 equiv) in dry MeCN (1 mL). After stirring for 30 min, DIPEA (40 µL, 0.23 mmol, 1.2 equiv) was added, and the mixture was stirred at rt for futher 18 h. After the reaction was heated to 55 °C for 6 h, the reaction was quenched by the addition of brine (10 mL), and the aqueous layer was extracted with EtOAc (3 × 15 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26a** (30 mg, 0.05 mmol, 28%) as a colorless oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.33 (s, 1H), 7.94 (s, 1H), 6.05 (d, J = 2.2 Hz, 1H), 5.95 (s, 2H), 5.59 (d, J = 8.3 Hz, 1H), 5.48−5.40 (m, 1H), 5.02−4.86 (m, 1H), 4.47−4.33 (m, 1H), 4.23−4.10 (m, 1H), 2.85−2.35 (m, 4H), 2.27 (s, 3H), 2.06−1.70 (m, 3H), 1.60 (s, 3H), 1.46−1.32 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.6, 155.8, 155.6, 153.2, 149.4, 140.2, 120.5, 114.7, 90.9, 84.9, 84.0, 83.3, 81.9, 79.7, 59.4, 54.3, 52.9, 42.7, 29.3, 28.5, 28.1, 27.3, 25.5. FT-IR: ν/cm^{-1} = 2978, 1704, 1642, 1597, 1475, 1366, 1330, 1296, 1249, 1211, 1152, 1076, 909, 870, 729. $[\alpha]_D^{20}$ = −12 (10 mg/mL; MeOH). R_f = 0.41 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)(ethyl)-amino)-2-((tert-butoxycarbonyl)amino)butanoate (26b). To a solution of **13a** (204 mg, 0.36 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added acetaldehyde (48 mg, 1.09 mmol, 61 µL, 3.0 equiv) and HOAc (31 µL, 0.54 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (130 mg, 0.62 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26b** as a colorless oil (71 mg, 0.12 mmol, 34%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.30 (s, 1H), 7.92 (s, 1H), 6.23 (s, 2H), 6.03 (d, J = 2.1 Hz, 1H), 5.73 (d, J = 8.1 Hz, 1H), 5.52−5.41 (m, 1H), 4.95 (dd, J = 6.5, 3.4 Hz, 1H), 4.36−4.27 (m, 1H), 4.21−4.08 (m, 1H), 2.79−2.69 (m, 1H), 2.63−2.37 (m, 5H), 1.97−1.81 (m, 1H), 1.78−1.62 (m, 1H), 1.58 (s, 3H), 1.44−1.34 (m, 21H), 1.32−1.26 (m, 2H), 0.92 (t, J = 6.7 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.9, 155.9, 155.6, 153.1, 149.3, 140.2, 120.4, 114.4, 91.0, 85.6, 83.9, 83.4, 81.7, 79.4, 55.6, 53.1, 50.1, 48.0, 29.8, 29.2, 28.4, 28.1, 27.2, 25.2, 11.3. FT-IR: ν/cm^{-1} = 2977, 1704, 1643, 1597, 1366, 1297, 1248, 1210, 1152, 1061, 909, 871, 799, 729. $[\alpha]_D^{20}$ = −9 (10 mg/mL; MeOH). R_f = 0.51 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)(propyl)-amino)-2-((tert-butoxycarbonyl)amino)butanoate (26c). To a sol-

ution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added propionaldehyde (51 µL, 0.71 mmol, 2.0 equiv) and HOAc (30 µL, 0.53 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26c** as a colorless oil (129 mg, 0.21 mmol, 60%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.27 (s, 1H), 7.89 (s, 1H), 6.47 (s, 2H), 6.01 (d, J = 2.1 Hz, 1H), 5.84 (d, J = 8.1 Hz, 1H), 5.52−5.43 (m, 1H), 4.94 (dd, J = 6.4, 3.3 Hz, 1H), 4.32−4.25 (m, 1H), 4.17−4.07 (m, 1H), 2.79−2.65 (m, 1H), 2.60−2.20 (m, 5H), 1.94−1.80 (m, 1H), 1.71−1.60 (m, 1H), 1.55 (s, 3H), 1.45−1.30 (m, 21H), 0.76 (t, J = 6.8 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 172.0, 156.0, 155.6, 153.0, 149.2, 140.1, 120.3, 114.2, 91.0, 85.6, 83.8, 83.5, 81.5, 79.3, 56.6, 56.1, 53.1, 50.8, 29.1, 28.4, 28.0, 27.2, 25.5, 19.8, 11.8. FT-IR: ν/cm^{-1} = 2978, 1704, 1642, 1597, 1475, 1367, 1297, 1248, 1210, 1153, 1076, 907, 871, 799, 728. $[\alpha]_D^{20}$ = −22 (10 mg/mL; MeOH). R_f = 0.57 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-(isopropyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (26d). To a 0 °C cooled solution of **12** (150 mg, 0.49 mmol, 1.0 equiv) were added acetone (72 µL, 0.98 mmol, 2.0 equiv) and HOAc (56 µL, 0.98 mmol, 2.0 equiv). After stirring for 30 min at rt, NaBH$_3$CN (154 mg, 2.45 mmol, 5.0 equiv) was added, and the reaction was stirred at rt for 18 h. The solvent was removed under reduced pressure at 40 °C, and the residue was dissolved in water (15 mL). The aqueous layer was then extracted with EtOAc (3 × 20 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was utilized without further purification. The residue (64 mg, 0.21 mmol, 1.0 equiv) was dissolved in 1,2-dce (4 mL) and was cooled to 0 °C. Compound **11a** (114 mg, 0.42 mmol, 2.0 equiv) and HOAc (18 µL, 0.31 mmol, 1.5 equiv) were added, and the mixture was stirred at 0 °C for 30 min. After the addition of NaBH(OAc)$_3$ (71 mg, 0.33 mmol, 1.6 equiv), the reaction was stirred at rt for 18 h. By the addition of saturated NaHCO$_3$ solution (20 mL), the reaction was quenched. The aqueous layer was extracted with EtOAc (3 × 20 mL), and the combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26d** (84 mg, 0.14 mmol, 66%) as a colorless oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.28 (s, 1H), 7.90 (s, 1H), 6.38 (s, 2H), 6.02 (d, J = 2.1 Hz, 1H), 5.84 (d, J = 8.1 Hz, 1H), 5.53−5.44 (m, 1H), 4.98 (dd, J = 6.5, 3.2 Hz, 1H), 4.33−4.21 (m, 1H), 4.18−4.03 (m, 1H), 3.00−2.83 (m, 1H), 2.77−2.63 (m, 1H), 2.58−2.33 (m, 3H), 1.99−1.85 (m, 1H), 1.78−1.63 (m, 1H), 1.57 (s, 3H), 1.46−1.32 (m, 21H), 0.98 (d, J = 6.5 Hz, 3H), 0.77 (d, J = 6.5 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 172.0, 156.0, 155.6, 153.0, 149.2, 140.3, 120.3, 114.3, 91.1, 86.4, 83.8, 83.4, 81.5, 79.3, 53.1, 52.1, 50.9, 47.1, 30.4, 29.7, 28.4, 28.0, 27.2, 25.5, 19.4, 16.2. FT-IR: ν/cm^{-1} = 2976, 1704, 1643, 1597, 1366, 1298, 1248, 1210, 1153, 1075, 908, 870, 799, 728. $[\alpha]_D^{20}$ = −20 (10 mg/mL; MeOH). R_f = 0.45 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)(butyl)-amino)-2-((tert-butoxycarbonyl)amino)butanoate (26e). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added butanal (51 mg, 0.71 mmol, 64 µL, 2.0 equiv) and HOAc (30 µL, 0.71 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26e** as a colorless oil (155 mg, 0.25 mmol, 71%).

65

^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.29 (s, 1H), 7.90 (s, 1H), 6.32 (s, 2H), 6.02 (s, 1H), 5.79 (d, J = 8.0 Hz, 1H), 5.57−5.41 (m, 1H), 4.95 (dd, J = 6.5, 3.3 Hz, 1H), 4.37−4.24 (m, 1H), 4.19−4.08 (m, 1H), 2.79−2.65 (m, 1H), 2.61−2.21 (m, 5H), 1.98−1.79 (m, 1H), 1.77−1.62 (m, 1H), 1.57 (s, 3H), 1.43−1.34 (m, 21H), 1.32−1.08 (m, 5H), 0.81 (t, J = 7.1 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.9, 155.9, 155.6, 153.1, 149.2, 140.2, 120.4, 114.3, 91.0, 85.6, 83.9, 83.5, 81.6, 79.3, 56.1, 54.3, 53.1, 50.8, 29.1, 28.7, 28.4, 28.0, 27.2, 25.5, 20.5, 14.0. FT-IR: ν/cm^{-1} = 2978, 1704, 1642, 1597, 1475, 1367, 1330, 1297, 1248, 1210, 1153, 1076, 908, 871, 799, 728. $[\alpha]_D^{20}$ = −23 (10 mg/mL; MeOH). R_f = 0.48 (DCM/MeOH = 20:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-(isobutyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (**26f**). To a solution of **13a** (250 mg, 0.44 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added isobutanal (49 μL, 0.53 mmol, 1.2 equiv) and HOAc (38 μL, 0.67 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26f** as a colorless oil (162 mg, 0.26 mmol, 59%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.24 (s, 1H), 7.86 (s, 1H), 6.59 (s, 2H), 6.00 (d, J = 2.1 Hz, 1H), 5.91 (d, J = 8.0 Hz, 1H), 5.53−5.41 (m, 1H), 4.94 (dd, J = 6.5, 3.1 Hz, 1H), 4.31−4.22 (m, 1H), 4.17−3.96 (m, 1H), 2.70 (dd, J = 13.5, 7.2 Hz, 1H), 2.58−2.25 (m, 3H), 2.12−1.99 (m, 2H), 1.96−1.78 (m, 1H), 1.54 (s, 5H), 1.42−1.26 (m, 21H), 0.82 (d, J = 6.6 Hz, 3H), 0.72 (d, J = 6.6 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 172.1, 156.0, 155.6, 153.0, 149.1, 140.1, 120.2, 114.1, 91.1, 85.6, 83.7, 83.6, 81.4, 79.2, 63.7, 56.6, 53.1, 51.4, 28.8, 28.3, 28.0, 27.1, 26.3, 25.4, 20.9, 20.8. FT-IR: ν/cm^{-1} = 2978, 1705, 1640, 1597, 1475, 1367, 1248, 1210, 1153, 1091, 907, 872, 727. $[\alpha]_D^{20}$ = −34 (10 mg/mL; MeOH). R_f = 0.52 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-(neopentyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (**26g**). To a solution of **13a** (250 mg, 0.44 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added pivalaldehyde (97 μL, 0.89 mmol, 2.0 equiv) and HOAc (38 μL, 0.67 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26g** as a colorless oil (97 mg, 0.15 mmol, 35%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.29 (s, 1H), 7.88 (s, 1H), 6.27 (s, 2H), 6.02 (d, J = 2.2 Hz, 1H), 5.57 (d, J = 7.1 Hz, 1H), 5.47 (d, J = 6.3 Hz, 1H), 5.04−4.95 (m, 1H), 4.39−4.26 (m, 1H), 4.16−4.02 (m, 1H), 2.85−2.72 (m, 1H), 2.70−2.53 (m, 2H), 2.51−2.40 (m, 1H), 2.20−2.11 (m, 2H), 1.95−1.80 (m, 1H), 1.57 (s, 4H), 1.48−1.31 (m, 21H), 0.80 (s, 9H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.9, 155.9, 155.5, 153.1, 149.3, 140.1, 120.4, 114.3, 91.0, 85.9, 83.9, 83.5, 81.7, 79.4, 68.4, 58.4, 53.2, 53.1, 29.5, 28.4, 28.3, 28.0, 27.3, 25.5. FT-IR: ν/cm^{-1} = 2977, 1705, 1643, 1598, 1478, 1366, 1248, 1209, 1152, 1075, 909, 871. $[\alpha]_D^{20}$ = −16 (10 mg/mL; MeOH). R_f = 0.55 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-(cyclopropylmethyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (**26h**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added cyclopropanecarbaldehyde (75 μL, 0.71 mmol, 2.0 equiv) and HOAc (30 μL, 0.53 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred

overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26h** (123 mg, 0.20 mmol, 57%) as a colorless oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.25 (s, 1H), 7.89 (s, 1H), 6.53 (s, 2H), 6.01 (d, J = 2.1 Hz, 1H), 5.94 (d, J = 8.0 Hz, 1H), 5.53−5.40 (m, 1H), 4.95 (dd, J = 6.4, 3.3 Hz, 1H), 4.37−4.24 (m, 1H), 4.13 (dd, J = 7.6, 3.0 Hz, 1H), 2.81 (dd, J = 13.5, 6.7 Hz, 1H), 2.68−2.41 (m, 3H), 2.27 (d, J = 6.5 Hz, 2H), 1.96−1.78 (m, 1H), 1.78−1.59 (m, 1H), 1.54 (s, 3H), 1.39−1.28 (m, 21H), 0.83−0.66 (m, 1H), 0.37 (dd, J = 8.0, 3.6 Hz, 2H), 0.06−0.10 (m, 2H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.9, 156.0, 155.6, 153.0, 149.1, 140.0, 120.2, 114.2, 90.9, 85.4, 83.8, 83.4, 81.5, 79.2, 59.3, 56.0, 53.1, 50.6, 29.0, 28.4, 28.2, 28.0, 27.1, 25.5, 8.3, 4.2, 3.7. FT-IR: ν/cm^{-1} = 2980, 1704, 1640, 1367, 1248, 1211, 1153, 1091, 1077, 907, 871, 727. $[\alpha]_D^{20}$ = −17 (10 mg/mL; MeOH). R_f = 0.50 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydro-furo[3,4-d][1,3]dioxol-4-yl)-methyl)-(cyclobutylmethyl)amino)-2-((tert-butoxycarbonyl)amino)-butanoate (**26i**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added cyclobutancarbaldehyde (64 μL, 0.71 mmol, 2.0 equiv) and HOAc (30 μL, 0.53 mmol, 1.5 equiv) The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26i** (97 mg, 0.15 mmol, 43%) as a colorless oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.29 (s, 1H), 7.89 (s, 1H), 6.38 (s, 2H), 6.02 (d, J = 2.1 Hz, 1H), 5.90 (d, J = 8.0 Hz, 1H), 5.56−5.43 (m, 1H), 4.93 (dd, J = 6.4, 3.2 Hz, 1H), 4.29 (td, J = 6.9, 3.2 Hz, 1H), 4.17−4.06 (m, 1H), 2.70 (dd, J = 13.5, 7.0 Hz, 1H), 2.59−2.26 (m, 6H), 2.07−1.60 (m, 7H), 1.57 (s, 3H), 1.53−1.46 (m, 1H), 1.45−1.29 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 172.0, 156.0, 155.6, 153.1, 149.2, 140.2, 120.3, 114.2, 91.1, 85.6, 83.8, 83.5, 81.5, 79.2, 61.0, 56.2, 53.2, 51.1, 33.4, 28.6, 28.4, 28.0, 27.4, 27.2, 27.1, 25.5, 18.7. FT-IR: ν/cm^{-1} = 2975, 1705, 1644, 1598, 1366, 1297, 1248, 1209, 1152, 1075, 1029, 871, 799. $[\alpha]_D^{20}$ = −28 (10 mg/mL; MeOH). R_f = 0.50 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-(cyclopentylmethyl)amino)-2-((tert-butoxycarbonyl)amino)-butanoate (**26j**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added cyclopentanecarbaldehyde (76 μL, 0.71 mmol, 2.0 equiv) and HOAc (30 μL, 0.53 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min and then NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26j** as a colorless oil (67 mg, 0.10 mmol, 29%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.25 (s, 1H), 7.85 (s, 1H), 6.13 (s, 2H), 5.99 (d, J = 2.1 Hz, 1H), 5.88 (d, J = 8.0 Hz, 1H), 5.53−5.39 (m, 1H), 4.94 (dd, J = 6.4, 3.2 Hz, 1H), 4.36−4.22 (m, 1H), 4.14−4.03 (m, 1H), 2.81−2.63 (m, 1H), 2.62−2.12 (m, 5H), 1.94−1.62 (m, 5H), 1.57 (s, 3H), 1.49−1.40 (m, 4H), 1.39−1.32 (m, 21H), 1.16−0.91 (m, 2H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 172.0, 155.9, 155.7, 153.1, 149.2, 140.3, 120.4, 114.3, 91.2, 85.7, 83.9, 83.7, 81.5, 79.3, 60.8, 56.3, 53.4, 51.4, 37.7, 31.3, 31.2, 28.7, 28.5, 28.1, 27.2, 25.5, 25.3, 25.2. FT-IR: ν/cm^{-1} = 2950, 1705, 1643, 1598, 1366, 1297, 1248, 1209, 1153, 1075, 909, 871, 799, 730, 682. $[\alpha]_D^{20}$ = −30 (10 mg/mL; MeOH). R_f = 0.53 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-(cyclohexylmethyl)amino)-2-((tert-butoxycarbonyl)-amino)-butanoate (**26k**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added cyclohexanecarbaldehyde (51 μL, 0.43 mmol, 1.2 equiv) and HOAc (30 μL, 0.53 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26k** (207 mg, 0.31 mmol, 89%) as a colorless oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.26 (s, 1H), 7.87 (s, 1H), 6.55 (s, 2H), 6.00 (d, J = 1.4 Hz, 1H), 5.91 (d, J = 8.0 Hz, 1H), 5.46 (d, J = 6.4 Hz, 1H), 4.94 (d, J = 3.3 Hz, 1H), 4.34−4.22 (m, 1H), 4.16−4.04 (m, 1H), 2.76−2.64 (m, 1H), 2.60−2.40 (m, 2H), 2.40−2.24 (m, 1H), 2.20−1.98 (m, 2H), 1.93−1.70 (m, 2H), 1.69−1.48 (m, 7H), 1.42−1.30 (m, 21H), 1.30−1.12 (m, 3H), 1.10−0.99 (m, 2H), 0.80−0.62 (m, 2H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 172.0, 156.0, 155.6, 153.0, 149.1, 140.1, 120.3, 114.1, 91.1, 85.6, 83.8, 83.6, 81.4, 79.2, 62.3, 56.7, 53.2, 51.6, 35.7, 31.7, 31.6, 28.9, 28.4, 28.0, 27.1, 26.7, 26.1, 26.0, 25.4. FT-IR: ν/cm^{-1} = 2926, 1705, 1640, 1597, 1476, 1367, 1330, 1211, 1153, 1076, 907, 871, 799, 727. $[\alpha]_D^{20}$ = −22 (10 mg/mL; MeOH). R_f = 0.59 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-(Allyl(((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydro-furo[3,4-d][1,3]dioxol-4-yl)methyl)-amino)-2-((tert-butoxycarbonyl)-amino)butanoate (**26l**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in DMF (4 mL) were added allyl bromide (30 μL, 0.39 mmol, 1.1 equiv) and DIPEA (66 μL, 0.39 mmol, 1.1 equiv), and the mixture was stirred at rt overnight. The reaction was quenched by the addition of brine (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26l** as a colorless oil (117 mg, 0.19 mmol, 55%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.29 (s, 1H), 7.90 (s, 1H), 6.44−6.18 (m, 2H), 6.03 (d, J = 2.2 Hz, 1H), 5.81−5.60 (m, 2H), 5.48−5.42 (m, 1H), 5.09−4.98 (m, 2H), 4.97−4.90 (m, 1H), 4.38−4.27 (m, 1H), 4.19−4.09 (m, 1H), 3.18−3.08 (m, 1H), 3.05−2.95 (m, 1H), 2.80−2.67 (m, 1H), 2.64−2.38 (m, 3H), 1.98−1.82 (m, 1H), 1.79−1.62 (m, 1H), 1.58 (s, 3H), 1.46−1.32 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.8, 155.9, 155.6, 153.1, 149.3, 140.1, 134.8, 120.4, 118.2, 114.4, 90.8, 85.5, 83.9, 83.4, 81.7, 79.4, 57.6, 55.7, 52.9, 50.4, 29.3, 28.4, 28.1, 27.2, 25.5. FT-IR: ν/cm^{-1} = 2979, 1705, 1642, 1597, 1476, 1366, 1330, 1297, 1248, 1210, 1152, 1075, 911, 871, 799, 730. $[\alpha]_D^{20}$ = −15 (10 mg/mL; MeOH). R_f = 0.56 (DCMMeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)(prop-2-yn-1-yl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (**26m**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in DMF (4 mL) were added propargyl bromide (58 mg, 0.49 mmol, 37 μL, 1.1 equiv) and DIPEA (66 μL, 0.39 mmol, 1.1 equiv), and the mixture was stirred at rt overnight. The reaction was quenched by the addition of brine (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26m** as a colorless oil (155 mg, 0.26 mmol, 74%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.27 (s, 1H), 7.92 (s, 1H), 6.48 (s, 2H), 6.03 (d, J = 2.2 Hz, 1H), 5.68 (d, J = 8.2 Hz, 1H), 5.48−5.40 (m, 1H), 4.99−4.85 (m, 1H), 4.36−4.25 (m, 1H), 4.22−4.13 (m, 1H), 3.36 (t, J = 2.4 Hz, 2H), 2.81−2.60 (m, 2H), 2.52 (t, J = 6.8 Hz, 2H), 2.07 (s, 1H), 1.95−1.80 (m, 1H), 1.80−1.64 (m, 1H), 1.55 (s, 3H), 1.42−1.30 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.8, 155.9, 155.5, 153.0, 149.2, 140.0, 120.3, 114.4, 90.8, 85.5, 83.9, 83.1, 81.7, 79.5, 77.9, 73.5, 55.3, 52.7, 50.2, 42.6, 29.8, 28.4, 28.0, 27.1, 25.4. FT-IR: ν/cm^{-1} =

2980, 1704, 1640, 1597, 1367, 1329, 1297, 1248, 1211, 1153, 1075, 908, 871, 728. $[\alpha]_D^{20}$ = −13 (10 mg/mL; MeOH). R_f = 0.68 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)(but-2-yn-1-yl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (**26n**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in DMF (4 mL) were added 1-bromo-2-butyne (43 μL, 0.49 mmol, 1.4 equiv) and DIPEA (66 μL, 0.49), and the mixture was stirred at rt overnight. The reaction was quenched by the addition of brine (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26n** as a colorless solid (128 mg, 0.21 mmol, 60%). ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.31 (s, 1H), 7.96 (s, 1H), 6.32 (s, 2H), 6.05 (d, J = 2.2 Hz, 1H), 5.63 (d, J = 8.2 Hz, 1H), 5.45 (d, J = 6.4 Hz, 1H), 4.99−4.86 (m, 1H), 4.43−4.30 (m, 1H), 4.23−4.12 (m, 1H), 3.38−3.30 (m, 2H), 2.90−2.62 (m, 2H), 2.60−2.49 (m, 2H), 2.03−1.75 (m, 2H), 1.58 (s, 3H), 1.45−1.35 (m, 21H), 1.22 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.8, 155.8, 155.6, 152.9, 149.3, 140.1, 120.3, 114.5, 90.9, 85.4, 84.0, 83.2, 81.7, 79.5, 77.4, 72.8, 55.5, 52.8, 50.3, 42.9, 29.7, 28.0, 27.2, 25.5, 3.4. FT-IR: ν/cm^{-1} = 2929, 1705, 1642, 1597, 1367, 1329, 1296, 1248, 1210, 1153, 1077, 908, 871, 799, 729. Mp: 83−86 °C. $[\alpha]_D^{20}$ = −10 (10 mg/mL; MeOH). R_f = 0.49 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)(pent-2-yn-1-yl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (**26o**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in DMF (4 mL) were added 1-bromo-2-pentyne (50 μL, 0.49 mmol, 1.4 equiv) and DIPEA (66 μL, 0.39 mmol, 1.1 equiv), and the mixture was stirred at rt overnight. The reaction was quenched by the addition of brine (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26o** (115 mg, 0.18 mmol, 51%) as a colorless solid. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.33 (s, 1H), 7.97 (s, 1H), 6.10 (s, 2H), 6.06 (d, J = 2.2 Hz, 1H), 5.58 (d, J = 8.2 Hz, 1H), 5.49−5.40 (m, 1H), 5.01−4.82 (m, 1H), 4.41−4.31 (m, 1H), 4.25−4.13 (m, 1H), 3.44−3.31 (m, 2H), 2.88−2.63 (m, 2H), 2.59−2.51 (m, 2H), 2.11 (qt, J = 7.5, 2.1 Hz, 2H), 2.01−1.71 (m, 2H), 1.59 (s, 3H), 1.47−1.29 (m, 21H), 1.06 (t, J = 7.5 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.8, 155.7, 155.6, 153.1, 149.4, 140.1, 120.4, 114.5, 90.9, 87.5, 85.4, 84.0, 83.2, 81.8, 79.6, 72.9, 55.5, 52.9, 42.9, 29.7, 28.5, 28.1, 27.3, 25.5, 14.3, 12.4. FT-IR: ν/cm^{-1} = 2926, 1705, 1642, 1366, 1247, 1210, 1153, 1080, 1028, 908, 871, 848, 730. Mp: 51−54 °C. $[\alpha]_D^{20}$ = −11 (10 mg/mL; MeOH). R_f = 0.46 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[-3,4-d][1,3]dioxol-4-yl)methyl)(but-3-yn-1-yl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (**26p**). To a solution of **13a** (200 mg, 0.35 mmol, 1.0 equiv) in DMF (4 mL) were added 4-bromo-1-butyne (46 μL, 0.49 mmol, 1.4 equiv) and DIPEA (66 μL, 0.39 mmol, 1.1 equiv), and the mixture was stirred at rt overnight. After the reaction was heated to 95 °C for 3 days, the reaction was quenched by the addition of brine (20 mL). The aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give **26p** (30 mg, 0.05 mmol, 14%) as a colorless solid. ^{1}H NMR (300 MHz, CDCl$_3$): δ/ppm = 8.35 (s, 1H), 7.93 (s, 1H), 6.05 (d, J = 2.2 Hz, 1H), 5.84 (s, 2H), 5.57−5.45 (m, 2H), 5.04−4.98 (m, 1H), 4.40−4.32 (m, 1H), 4.22−4.13 (m, 1H), 2.89−2.47 (m, 6H), 2.29 (td, J = 7.3, 2.7 Hz, 2H), 1.94 (s, 1H), 1.85−1.67 (m, 2H), 1.61 (s, 3H), 1.48−1.37 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ/ppm = 171.8, 155.6, 153.1, 149.3, 140.5, 124.9, 120.5, 114.6, 91.1, 85.5, 83.9, 83.4, 82.5, 81.9, 79.7, 69.6, 55.8, 53.1, 52.8, 50.7, 29.8, 28.5, 28.1, 27.3, 25.5, 16.7. FT-IR: ν/cm^{-1} = 2976, 2928, 1709, 1643, 1598, 1476, 1366, 1330, 1297, 1248, 1210, 1153, 1079, 871. Mp: 64−67

°C. $[\alpha]_D^{20}$ = −21 (10 mg/mL; MeOH). R_f = 0.55 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(benzyl)-amino)-2-((tert-butoxycarbonyl)amino)butanoate (26q). To a solution of 13a (200 mg, 0.35 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added benzaldehyde (72 μL, 0.71 mmol, 2.0 equiv) and HOAc (30 μL, 0.53 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)₃ (128 mg, 0.60 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO₃ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na₂SO₄, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give 26q as a colorless oil (142 mg, 0.23 mmol, 63%). ¹H NMR (300 MHz, CDCl₃): δ/ppm = 8.19 (s, 1H), 7.82 (s, 1H), 7.24−7.15 (m, 5H), 6.24 (s, 2H), 6.01 (d, J = 2.1 Hz, 1H), 5.67 (d, J = 8.1 Hz, 1H), 5.33 (d, J = 5.9 Hz, 1H), 4.84 (dd, J = 6.4, 3.4 Hz, 1H), 4.41−4.29 (m, 1H), 4.22−4.09 (m, 1H), 3.69 (d, J = 13.5 Hz, 1H), 3.44 (d, J = 13.5 Hz, 1H), 2.82−2.70 (m, 1H), 2.66−2.42 (m, 3H), 2.03−1.89 (m, 1H), 1.84−1.69 (m, 1H), 1.57 (s, 3H), 1.45−1.31 (m, 21H). ¹³C NMR (75.5 MHz, CDCl₃): δ/ppm = 171.8, 155.8, 155.5, 153.1, 149.2, 139.9, 138.3, 129.1, 128.3, 127.3, 120.3, 114.4, 90.8, 85.3, 83.9, 83.5, 81.7, 79.4, 59.0, 55.7, 53.0, 50.8, 29.3, 28.5, 28.0, 27.2, 25.5. FT-IR: ν/cm⁻¹ = 2979, 1704, 1642, 1597, 1367, 1247, 1209, 1152, 1074, 908, 870, 728, 699. $[\alpha]_D^{20}$ = −14 (10 mg/mL; MeOH). R_f = 0.66 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-(phenethyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (26r). To a solution of 13a (250 mg, 0.44 mmol, 1.0 equiv) in 1,2-dce (4 mL) at 0 °C under argon atmosphere were added phenylacetaldehyde (103 μL, 0.89 mmol, 2.0 equiv) and HOAc (38 μL, 0.67 mmol, 1.5 equiv). The mixture was stirred at 0 °C for 30 min, and then NaBH(OAc)₃ (160 mg, 0.75 mmol, 1.7 equiv) was added. The reaction was slowly warmed up to rt and was then stirred overnight. The reaction was quenched by the addition of saturated NaHCO₃ solution (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over Na₂SO₄, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give 26r as a colorless oil (140 mg, 0.21 mmol, 48%). ¹H NMR (300 MHz, CDCl₃): δ/ppm = 8.30 (s, 1H), 7.87 (s, 1H), 7.24−7.01 (m, 5H), 6.52 (s, 2H), 6.03 (d, J = 2.1 Hz, 1H), 5.74 (d, J = 8.1 Hz, 1H), 5.52−5.42 (m, 1H), 4.95 (dd, J = 6.5, 3.4 Hz, 1H), 4.39−4.28 (m, 1H), 4.16−4.06 (m, 1H), 2.94−2.78 (m, 1H), 2.77−2.47 (m, 7H), 2.04−1.86 (m, 1H), 1.83−1.66 (m, 1H), 1.58 (s, 3H), 1.46−1.32 (m, 21H). ¹³C NMR (75.5 MHz, CDCl₃): δ/ppm = 171.8, 155.9, 155.5, 152.9, 149.1, 140.1, 139.9, 128.6, 128.4, 126.0, 120.2, 114.3, 90.9, 85.5, 83.7, 83.3, 81.6, 79.4, 56.2, 55.8, 52.8, 50.6, 33.0, 29.2, 28.0, 27.9, 27.1, 25.4. FT-IR: ν/cm⁻¹ = 2978, 1704, 1642, 1598, 1367, 1248, 1209, 1152, 1075, 909, 870, 729, 700. $[\alpha]_D^{20}$ = −14 (10 mg/mL; MeOH). R_f = 0.50 (DCM/MeOH = 10:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)((R)/(S)-but-3-yn-2-yl)amino)-2-((tert-butoxycarbonyl)-amino)butanoate (26s). To a solution of 13a (200 mg, 0.35 mmol, 1.1 equiv) in DMF (4 mL) were added 3-bromo-1-butyne (32 μL, 0.35 mmol, 1.0 equiv), DIPEA (66 μL, 0.39 mmol, 1.1 equiv), and CuBr (7 mg, 0.035 mmol, 10 mol %), and the mixture was stirred at rt overnight. The reaction was quenched by the addition of brine (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give 26s (107 mg, 0.17 mmol, 49%) as a colorless solid. The diastereomers were separated by semipreparative HPLC (reversed phase C₁₈, MeCN/H₂O = 25:75 + 0.1% TFA, t_R(26s-B) = 24 min, t_R(26s-A) = 26 min) and directly used in the next step without further characterization. ¹H NMR (300 MHz, CDCl₃): δ/ppm

= 8.29 (s, 2H), 7.68 (s, 2H), 6.15 (dd, J = 4.8, 2.2 Hz, 1H), 5.44 (s, 1H), 5.25 (td, J = 6.4, 2.1 Hz, 1H), 4.97 (q, J = 4.9 Hz, 1H), 4.67−4.50 (m, 1H), 4.24 (d, J = 21.4 Hz, 1H), 4.12−4.00 (m, 1H), 3.55−2.79 (m, 4H), 2.63−2.41 (m, 1H), 2.25−2.04 (m, 1H), 1.99−1.79 (m, 1H), 1.60 (s, 3H), 1.48 (dd, J = 20.0, 6.7 Hz, 3H), 1.44−1.31 (m, 21H). ¹³C NMR (75.5 MHz, CDCl₃): δ/ppm = 170.6, 155.7, 155.2, 148.1, 145.0, 142.7, 120.1, 115.9, 90.7, 84.3, 83.9, 82.6, 82.4, 82.3, 80.2, 78.2, 53.6, 52.4, 50.8, 48.9, 28.3, 28.2, 27.9, 27.1, 25.4, 18.0. FT-IR: ν/cm⁻¹ = 2982, 1695, 1505, 1456, 1427, 1369, 1181, 1140, 1082, 911, 832, 799, 722. Mp: 80−83 °C. $[\alpha]_D^{20}$ = +3 (10 mg/mL; CDCl₃). R_f = 0.69 (DCM/MeOH = 9:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-Amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-methylbut-3-yn-2-yl)amino)-2-((tert-butoxycarbonyl) amino)butanoate (26t). To a solution of 13a (150 mg, 0.27 mmol, 1.0 equiv) in DMF (4 mL) were added 3-chlor-3-methyl-1-butyne (33 μL, 0.29 mmol, 1.1 equiv), DIPEA (50 μL, 0.29 mmol, 1.1 equiv), and CuBr (38 mg, 0.27 mmol, 1.0 equiv), and the mixture was stirred at rt for 48 h. The reaction was quenched by the addition of brine (20 mL), and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography (DCM/MeOH = 30:1) to give 26t (69 mg, 0.11 mmol, 41%) as a colorless oil. ¹H NMR (300 MHz, CDCl₃): δ/ppm = 8.30 (s, 1H), 7.91 (s, 1H), 6.27 (s, 2H), 6.08−5.98 (m, 1H), 5.52−5.41 (m, 1H), 5.36 (d, J = 8.0 Hz, 1H), 5.00 (dd, J = 6.6, 3.4 Hz, 1H), 4.31 (td, J = 6.6, 3.4 Hz, 1H), 4.14−3.97 (m, 1H), 2.97 (dd, J = 14.1, 6.7 Hz, 1H), 2.70 (ddd, J = 16.7, 11.0, 6.7 Hz, 3H), 2.13 (s, 1H), 2.06−1.87 (m, 1H), 1.82−1.69 (m, 1H), 1.57 (s, 3H), 1.38 (d, J = 11.3 Hz, 24H), 1.25 (s, 3H). ¹³C NMR (75.5 MHz, CDCl₃): δ/ppm = 171.9, 155.8, 155.5, 152.9, 149.3, 140.4, 120.5, 114.4, 91.0, 87.2, 86.9, 84.0, 83.2, 81.8, 79.6, 70.8, 55.0, 53.7, 52.9, 48.3, 32.4, 29.1, 28.8, 28.4, 28.0, 27.3, 25.6. FT-IR: ν/cm⁻¹ = 3307, 2980, 2935, 1705, 1644, 1598, 1477, 1367, 1249, 1211, 1155, 1094, 871, 755. $[\alpha]_D^{20}$ = +6 (10 mg/mL; CHCl₃). R_f = 0.50 (DCM/MeOH = 9:1).

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(methyl)amino)butanoic Acid Trifluoroacetate Salt (27a). To a solution of 26a (21 mg, 0.04 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H₂O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H₂O (7 mL) and was dried by lyophilization to give 27a (20 mg, 0.04 mmol, 99%, 1.0 equiv TFA) as a colorless trifluoroacetate salt. ¹H NMR (300 MHz, CD₃OD): δ/ppm = 8.50 (s, 1H), 8.26 (s, 1H), 6.11 (d, J = 5.4 Hz, 1H), 4.74 (t, J = 5.4 Hz, 1H), 4.53 (d, J = 10.9 Hz, 1H), 4.31 (t, J = 4.4 Hz, 1H), 4.04 (dd, J = 9.2, 3.6 Hz, 1H), 3.77 (t, J = 12.1 Hz, 1H), 3.62−3.50 (m, 2H), 3.49−3.38 (m, 1H), 2.93 (s, 3H), 2.48−2.30 (m, 1H), 2.24−2.08 (m, 1H). ¹³C NMR (75.5 MHz, CD₃OD): δ/ppm = 171.3, 153.3, 149.8, 147.3, 144.0, 120.4, 90.5, 80.3, 74.9, 73.4, 59.4, 56.1, 53.3, 39.7, 26.1. FT-IR: ν/cm⁻¹ = 3093, 1668, 1507, 1425, 1325, 1196, 1129, 835, 799, 722. Mp: 86−89 °C. $[\alpha]_D^{20}$ = +14 (6 mg/mL; MeOH). ESI-MS: m/z [M + H]⁺ = 382.2 (100%), 383.1 (18.7%), 384.1 (2.5%). Calculated: 382.2 (100%), 383.2 (16.2%), 384.2 (1.2%). Purity: 99% (HPLC, MeCN/H₂O = 20:80 + 0.1% HCOOH); t_R = 1.89 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(ethyl)amino)butanoic Acid Trifluoroacetate Salt (27b). To a solution of 26b (66 mg, 0.11 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H₂O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H₂O (7 mL) and was dried by lyophilization to give 27b (69 mg, 0.11 mmol, 99%, 2.0 equiv TFA) as a colorless trifluoroacetate salt. ¹H NMR (300 MHz, CD₃OD): δ/ppm = 8.35 (s,

1H), 8.25 (s, 1H), 6.01 (d, J = 4.2 Hz, 1H), 4.61 (t, J = 4.7 Hz, 1H), 4.43−4.24 (m, 2H), 3.94 (dd, J = 8.2, 4.5 Hz, 1H), 3.70−3.46 (m, 2H), 3.46−3.10 (m, 4H), 2.39−2.19 (m, 1H), 2.19−2.03 (m, 1H), 1.20 (t, J = 7.2 Hz, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 170.7, 151.5, 148.3, 145.2, 142.9, 119.4, 90.2, 78.8, 73.4, 72.0, 54.5, 51.2, 50.2, 24.6, 7.2. FT-IR: ν/cm^{-1} = 2471, 2074, 1663, 1506, 1424, 1183, 1128, 973, 833, 799, 721. Mp: 80−83 °C. $[\alpha]_D^{20}$ = +18 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 396.2 (100%), 397.2 (20.1%), 398.2 (2.5%). Calculated: 396.2 (100%), 397.2 (17.3%), 398.2 (1.4%). Purity: 99% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.82 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(propyl)amino)butanoic Acid Trifluoroacetate Salt (27c). To a solution of **26c** (121 mg, 0.20 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27c** (106 mg, 0.20 mmol, 99%, 1.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.19 (s, 1H), 8.13 (s, 1H), 5.85 (d, J = 4.0 Hz, 1H), 4.45 (t, J = 4.5 Hz, 1H), 4.26−4.12 (m, 2H), 3.78 (dd, J = 7.8, 5.0 Hz, 1H), 3.54−3.11 (m, 4H), 3.05−2.84 (m, 2H), 2.25−2.07 (m, 1H), 2.05−1.91 (m, 1H), 1.47 (dt, J = 9.3, 7.0 Hz, 2H), 0.68 (t, J = 7.3 Hz, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 171.6, 152.5, 149.6, 146.0, 144.4, 120.9, 91.9, 80.0, 74.7, 73.4, 56.4, 56.2, 52.1, 51.8, 25.9, 17.9, 11.0. FT-IR: ν/cm^{-1} = 2979, 2473, 2074, 1662, 1506, 1425, 1183, 1130, 973, 834, 798, 721. Mp: 64−67 °C. $[\alpha]_D^{20}$ = +21 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 410.2 (100%), 411.2 (24.5%), 412.2 (3.6%). Calculated: 410.2 (100%), 411.2 (18.4%), 412.2 (1.6%). Purity 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.61 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(isopropyl)amino)butanoic Acid Trifluoroacetate Salt (27d). To a solution of **26d** (65 mg, 0,11 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27d** (77 mg, 0.11 mmol, 99%, 2.5 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.34 (s, 1H), 8.29 (s, 1H), 6.01 (d, J = 3.6 Hz, 1H), 4.58 (s, 1H), 4.37−4.29 (m, 2H), 3.88 (dd, J = 8.4, 4.4 Hz, 1H), 3.78−3.63 (m, 1H), 3.58−3.49 (m, 2H), 3.46−3.23 (m, 2H), 3.19 (d, J = 3.0 Hz, 1H), 2.41−2.22 (m, 1H), 2.13 (d, J = 12.2 Hz, 1H), 1.28 (d, J = 6.5 Hz, 3H), 1.19 (d, J = 5.7 Hz, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.0, 152.8, 149.7, 146.6, 144.2, 121.0, 92.0, 81.0, 74.7, 73.5, 57.9, 53.8, 52.7, 49.9, 27.0, 17.0, 16.0. FT-IR: ν/cm^{-1} = 2992, 2488, 1663, 1506, 1425, 1182, 1128, 975, 833, 799, 721. Mp: 88−91 °C. $[\alpha]_D^{20}$ = +19 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 410.2 (100%), 411.2 (18.4%), 412.2 (1.6%). Calculated: 410.2 (100%), 411.2 (20.2%), 412.2 (3.0%). Purity: 98% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.81 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(butyl)amino)butanoic Acid Trifluoroacetate Salt (27e). To a solution of **26e** (150 mg, 0.24 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27e** (123 mg, 0.24 mmol, 99%, 1.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.19 (s, 1H), 8.13 (s, 1H), 5.85 (d, J = 3.9 Hz, 1H), 4.45 (t, J = 4.4 Hz, 1H), 4.24−4.13 (m, 2H), 3.78 (dd, J = 7.7, 5.0 Hz, 1H), 3.57−3.11 (m, 4H),

2.97 (d, J = 5.0 Hz, 2H), 2.21−2.07 (m, 1H), 2.01 (dd, J = 9.6, 5.0 Hz, 1H), 1.40 (dd, J = 11.9, 6.0 Hz, 2H), 1.07 (dd, J = 7.5, 2.3 Hz, 2H), 0.63 (t, J = 6.7, 6.0 Hz, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 171.5, 152.5, 149.7, 146.0, 144.5, 120.9, 91.9, 80.0, 74.7, 73.4, 56.4, 54.6, 52.1, 51.7, 26.3, 25.9, 20.7, 13.8. FT-IR: ν/cm^{-1} = 2969, 1663, 1506, 1424, 1182, 1129, 976, 834, 799, 721. Mp: 74−77 °C. $[\alpha]_D^{20}$ = +17 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 424.2 (100%), 425.2 (19.8%), 426.2 (3.3%). Calculated: 424.2 (100%), 425.2 (19.5%), 426.2 (1.8%). Purity: 99% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.63 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(isobutyl)amino)butanoic Acid Trifluoroacetate Salt (27f). To a solution of **26f** (160 mg, 0.26 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27f** (200 mg, 0.26 mmol, 99%, 3.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.14 (d, J = 12.3 Hz, 2H), 5.84 (d, J = 3.2 Hz, 1H), 4.36 (t, J = 3.9 Hz, 1H), 4.23−4.13 (m, 2H), 3.75 (dd, J = 7.8, 4.8 Hz, 1H), 3.53−3.26 (m, 3H), 3.24−3.11 (m, 1H), 2.94−2.71 (m, 2H), 2.24−1.92 (m, 2H), 1.90−1.72 (m, 1H), 0.70 (dd, J = 14.9, 6.5 Hz, 6H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 171.6, 152.6, 149.6, 146.0, 144.4, 120.9, 92.2, 79.6, 74.7, 73.4, 63.0, 56.7, 53.0, 52.3, 25.7, 25.5, 20.5, 20.4. FT-IR: ν/cm^{-1} = 2476, 2241, 2072, 1666, 1426, 1186, 1136, 973, 834, 799, 722. Mp: 67−70 °C. $[\alpha]_D^{20}$ = +14 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 424.2 (100%), 425.2 (21.5%), 426.2 (2.9%). Calculated: 424.2 (100%), 425.2 (19.5%), 426.2 (1.8%). Purity: 98% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.88 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(neopentyl)amino)butanoic Acid Trifluoroacetate Salt (27g). To a solution of **26g** (67 mg, 0.11 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27g** (73 mg, 0.11 mmol, 99%, 2.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.35 (d, J = 9.2 Hz, 2H), 6.04 (d, J = 2.9 Hz, 1H), 4.51 (dd, J = 4.8, 2.9 Hz, 1H), 4.45−4.35 (m, 2H), 3.90 (dd, J = 8.8, 4.2 Hz, 1H), 3.77−3.34 (m, 4H), 3.17−3.01 (m, 2H), 2.41−2.23 (m, 1H), 2.23−2.08 (m, 1H), 1.02 (s, 9H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.0, 152.8, 149.6, 146.5, 144.1, 121.0, 92.4, 79.9, 74.6, 73.6, 68.4, 59.4, 56.4, 53.0, 32.6, 27.9, 26.2. FT-IR: ν/cm^{-1} = 2969, 1667, 1507, 1485, 1423, 1185, 1132, 976, 833, 799, 721. Mp: 86−89 °C. $[\alpha]_D^{20}$ = −13 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 438.2 (100%), 439.2 (24.6%), 440.3 (3.9%). Calculated: 438.3 (100%), 439.3 (20.5%), 440.3 (2.0%). Purity:100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.82 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(cyclopropylmethyl)amino)-butanoic Acid Trifluoroacetate Salt (27h). To a solution of **26h** (114 mg, 0.18 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27h** (124 mg, 0.18 mmol, 99%, 2.4 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.48 (s, 1H), 8.39 (s, 1H), 6.14 (d, J = 3.9 Hz, 1H), 4.71 (t, J = 4.5 Hz, 1H), 4.56−4.42 (m, 2H), 4.06 (dd, J = 8.3, 4.5 Hz, 1H), 3.86−3.74 (m, 2H), 3.71−3.51 (m, 2H), 3.26−3.18 (m, 2H), 2.54−2.37 (m,

1H), 2.35−2.19 (m, 1H), 1.19−1.05 (m, 1H), 0.82−0.67 (m, 2H), 0.45 (q, J = 3.9, 2.7 Hz, 2H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.0, 152.8, 149.7, 146.5, 144.2, 120.8, 91.8, 80.0, 74.8, 73.4, 59.5, 56.3, 52.6, 52.2, 26.1, 6.2, 5.1. FT-IR: ν/cm^{-1} = 2973, 1667, 1507, 1467, 1424, 1322, 1185, 1129, 1064, 949, 898, 833, 799, 721. mp: 70−73 °C. $[\alpha]_D^{20}$ = +17 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 422.2 (100%), 423.2 (17.8%), 424.2 (2.1%). Calculated: 422.2 (100%), 423.2 (19.5%), 424.2 (1.8%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.13 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(cyclobutylmethyl)amino)-butanoic Acid Trifluoroacetate Salt (27i). To a solution of **26i** (114 mg, 0.18 mmol) in DCM (1.5 mL) at 5 °C was added TFA (1.5 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid was detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL), and the residue was dissolved in water (1.8 mL) and TFA (0.3 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was dried by lyophilization to give **27i** (91 mg, 0.13 mmol, 99%, 2.3 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.49 (s, 1H), 8.38 (s, 1H), 6.14 (d, J = 3.9 Hz, 1H), 4.72 (t, J = 4.5 Hz, 1H), 4.55−4.37 (m, 2H), 4.02 (dd, J = 8.8, 4.1 Hz, 1H), 3.79−3.24 (m, 6H), 2.89−2.70 (m, 1H), 2.50−2.31 (m, 1H), 2.29−2.05 (m, 3H), 2.01−1.76 (m, 4H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.5, 153.1, 149.7, 147.0, 144.1, 117.9, 114.1, 91.6, 80.1, 74.7, 73.5, 59.6, 56.6, 53.0, 31.5, 28.2, 28.0, 26.1, 19.4. FT-IR: ν/cm^{-1} = 2977, 2461, 2073, 1664, 1506, 1423, 1182, 1128, 975, 833, 799, 721. Mp: 63−66 °C. $[\alpha]_D^{20}$ = +19 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 436.2 (100%), 437.2 (23.4%), 438.2 (3.3%). Calculated: 436.2 (100%), 437.2 (20.5%), 438.2 (1.0%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.89 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(cyclopentylmethyl)amino)-butanoic Acid Trifluoroacetate Salt (27j). To a solution of **26j** (55 mg, 0.09 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL), and was dried by lyophilization to give **27j** (51 mg, 0.09 mmol, 99%, 1.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.36 (d, J = 21.6 Hz, 2H), 6.11 (d, J = 3.5 Hz, 1H), 4.74 (s, 1H), 4.50 (d, J = 7.2 Hz, 2H), 4.05−3.84 (m, 1H), 3.80−3.40 (m, 4H), 3.30−3.11 (m, 2H), 2.56−2.04 (m, 3H), 1.72−1.43 (m, 4H), 1.38−1.11 (m, 2H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 173.0, 154.8, 149.9, 149.7, 143.3, 120.9, 91.9, 80.3, 74.4, 73.7, 60.7, 56.5, 54.4, 54.3, 36.5, 32.1, 32.0, 26.3, 25.9, 25.8. FT-IR: ν/cm^{-1} = 2961, 1669, 1421, 1199, 1131, 978, 833, 799, 721. Mp: 121−124 °C. $[\alpha]_D^{20}$ = +15 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 450.2 (100%), 451.2 (25.2%), 452.2 (3.9%). Calculated: 450.3 (100%), 451.3 (21.6%), 452.3 (2.2%). Purity: 95% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.99 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(cyclohexylmethyl)amino)-butanoic Acid Trifluoroacetate Salt (27k). To a solution of **26k** (150 mg, 0.23 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27k** (122 mg, 0.23 mmol, 99%, 0.6 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.28 (s, 1H), 8.23 (s, 1H), 5.96 (d, J = 3.3 Hz, 1H), 4.50 (t, J = 3.9 Hz, 1H), 4.36−4.24 (m, 2H), 3.84 (dd, J = 8.6, 4.3 Hz, 1H), 3.65−3.22 (m, 4H), 3.02−2.82 (m, 2H), 2.32−2.14 (m, 1H), 2.14−1.99 (m, 1H), 1.69−1.37 (m, 6H), 1.16−0.66 (m, 6H). ^{13}C NMR (75.5 MHz,

CD$_3$OD): δ/ppm = 172.2, 153.0, 149.7, 146.8, 144.2, 120.9, 92.1, 79.9, 74.7, 73.5, 62.0, 56.8, 53.8, 52.8, 34.5, 31.7, 31.6, 26.8, 26.4, 25.9. FT-IR: ν/cm^{-1} = 2932, 2857, 1666, 1506, 1423, 1182, 1130, 976, 833, 799, 721. Mp: 126−129 °C. $[\alpha]_D^{20}$ = +16 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 464.2 (100%), 464.6 (10.2%), 465.5 (4.6%). Calculated: 464.3 (100%), 465.3 (22.7%), 466.3 (2.5%). Purity: 96% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.77 min.

(S)-4-(Allyl(((2R,3S,4R,5R)-5-(6-Amino-9H-purin-9-yl)-3,4-dihy-droxytetrahydrofuran-2-yl)methyl)amino)-2-aminobutanoic Acid Trifluoroacetate Salt (27l). To a solution of **26l** (100 mg, 0.17 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27l** (119 mg, 0.17 mmol, 99%, 2.6 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.19 (s, 1H), 8.13 (s, 1H), 5.85 (d, J = 4.0 Hz, 1H), 5.76−5.59 (m, 1H), 5.40−5.26 (m, 2H), 4.45 (t, J = 4.7 Hz, 1H), 4.30−4.10 (m, 2H), 3.78 (dd, J = 7.9, 5.0 Hz, 1H), 3.73−3.57 (m, 2H), 3.54−3.10 (m, 4H), 2.25−2.08 (m, 1H), 2.07−1.92 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 171.7, 152.5, 149.7, 146.0, 144.5, 127.6, 126.8, 120.8, 91.7, 80.0, 74.7, 73.4, 56.7, 56.3, 52.2, 51.5, 25.9. FT-IR: ν/cm^{-1} = 2981, 1672, 1509, 1427, 1196, 1136, 836, 799, 723. Mp: 78−81 °C. $[\alpha]_D^{20}$ = +14 (10 mg/ mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 408.2 (100%), 409.2 (20.8%), 410.2 (3.4%). Calculated: 408.2 (100%), 409.2 (18.4%), 410.2 (1.6%). Purity:100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.71 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(prop-2-yn-1-yl)amino)-butanoic Acid Trifluoroacetate Salt (27m). To a solution of **26m** (134 mg, 0.22 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27m** (157 mg, 0.22 mmol, 99%, 2.7 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.23 (s, 1H), 8.06 (s, 1H), 5.83 (d, J = 4.9 Hz, 1H), 4.45 (t, J = 4.9 Hz, 1H), 4.20 (d, J = 7.0 Hz, 1H), 4.09 (t, J = 4.9 Hz, 1H), 3.88 (t, J = 2.8 Hz, 2H), 3.85−3.76 (m, 1H), 3.51−3.28 (m, 2H), 3.24 (t, J = 7.1 Hz, 2H), 2.96 (s, 1H), 2.22−2.04 (m, 1H), 1.99−1.84 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.2, 152.3, 149.7, 145.8, 144.5, 120.6, 91.1, 81.1, 80.7, 74.9, 73.4, 72.9, 56.8, 52.8, 52.5, 42.9, 26.3. FT-IR: ν/cm^{-1} = 3097, 1671, 1509, 1428, 1323, 1195, 1137, 837, 799, 723. Mp: 76−79 °C. $[\alpha]_D^{20}$ = +12 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 406.2 (100%), 407.2 (21.0%), 408.1 (2.5%). Calculated: 406.2 (100%), 407.2 (18.4%), 408.2 (1.6%). Purity: 97% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.70 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(but-2-yn-1-yl)amino)-butanoic Acid Trifluoroacetate Salt (27n). To a solution of **26n** (119 mg, 0.19 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27n** (131 mg, 0.19 mmol, 99%, 2.4 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.61 (s, 1H), 8.31 (s, 1H), 6.17 (d, J = 5.4 Hz, 1H), 4.77 (t, J = 5.3 Hz, 1H), 4.62−4.50 (m, 1H), 4.38 (t, J = 4.4 Hz, 1H), 4.28−4.09 (m, 3H), 3.85−3.50 (m, 4H), 2.55−2.36 (m, 1H), 2.33−2.13 (m, 1H), 1.92 (s, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 173.0, 152.4, 149.7, 146.1, 144.3, 120.2, 90.4, 89.3, 80.6, 75.1, 73.3, 67.4, 56.6, 53.2,

52.8, 42.7, 26.1, 3.2. FT-IR: ν/cm^{-1} = 2925, 1666, 1507, 1423, 1184, 1131, 976, 897, 835, 799, 722. Mp: 121−124 °C. $[\alpha]_D^{20}$ = +13 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 420.2 (100%), 421.2 (20.6%), 422.2 (3.5%). Calculated: 420.2 (100%), 421.2 (19.5%), 422.2 (1.8%). Purity: 99% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 1.95 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(pent-2-yn-1-yl)amino)-butanoic Acid Trifluoroacetate Salt (27o). To a solution of **26o** (98 mg, 0.16 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27o** (93 mg, 0.16 mmol, 99%, 1.3 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.61 (s, 1H), 8.29 (s, 1H), 6.17 (d, J = 5.5 Hz, 1H), 4.75 (t, J = 5.3 Hz, 1H), 4.59−4.49 (m, 1H), 4.38 (t, J = 4.4 Hz, 1H), 4.30−4.10 (m, 3H), 3.83−3.47 (m, 4H), 2.46 (dt, J = 15.6, 7.9 Hz, 1H), 2.36−2.26 (m, 2H), 2.25−2.14 (m, 1H), 1.17 (t, J = 7.5 Hz, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 173.3, 152.7, 149.7, 146.5, 144.2, 120.1, 94.6, 90.3, 80.7, 75.1, 73.3, 67.9, 56.6, 53.5, 53.1, 42.4, 26.2, 13.7, 13.0. FT-IR: ν/cm^{-1} = 2981, 1669, 1507, 1422, 1321, 1198, 1130, 1061, 900, 833, 799, 721. Mp: 110−113 °C. $[\alpha]_D^{20}$ = +14 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 434.2 (100%), 435.4 (9.4%), 436.2 (2.7%). Calculated: 434.2 (100%), 435.2 (20.5%), 436.2 (1.0%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.78 min.

(S)-2-Amino-4-(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(but-3-yn-1-yl)amino)-butanoic Acid Trifluoroacetate Salt (27p). To a solution of **26p** (25 mg, 0.04 mmol) in DCM (1.5 mL) at 5 °C was added TFA (1.5 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid was detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL), and the residue was dissolved in water (1.8 mL) and TFA (0.3 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was dried by lyophilization to give **27p** (28 mg, 0.04 mmol, 99%, 2.5 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.63−8.24 (m, 2H), 6.11 (s, 1H), 4.74 (s, 1H), 4.55−4.35 (m, 2H), 4.01 (dd, J = 13.9, 7.3 Hz, 1H), 3.74−3.33 (m, 6H), 2.75−2.64 (m, 2H), 2.47 (d, J = 2.3 Hz, 1H), 2.43−2.26 (m, 1H), 2.19−1.99 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.6, 154.1, 148.7, 148.7, 143.2, 125.5, 91.6, 80.9, 80.2, 74.6, 73.5, 73.1, 56.7, 53.8, 53.6, 52.8, 26.5, 15.3. FT-IR: ν/cm^{-1} = 2926, 1673, 1506, 1423, 1200, 1132, 834, 800, 722. Mp: 163−166 °C. $[\alpha]_D^{20}$ = +28 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 420.2 (100%), 421.2 (21.8%), 422.2 (2.9%). Calculated: 420.2 (100%), 421.2 (19.5%), 422.2 (1.8%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.94 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(benzyl)amino)butanoic Acid Trifluoroacetate Salt (27q). To a solution of **26q** (115 mg, 0.18 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27q** (104 mg, 0.18 mmol, 99%, 1.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.19 (s, 1H), 8.04 (s, 1H), 7.27−7.19 (m, 2H), 7.17−7.08 (m, 3H), 5.90 (d, J = 3.3 Hz, 1H), 4.39 (dd, J = 5.2, 3.3 Hz, 1H), 4.36−4.11 (m, 4H), 3.76−3.67 (m, 1H), 3.52−3.20 (m, 4H), 2.29−2.10 (m, 1H), 2.10−1.96 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.5, 152.8, 149.6, 146.5, 144.2, 132.3, 131.1, 130.6, 130.3, 120.8, 92.0, 80.1, 74.8, 73.5, 58.9, 56.1, 53.1, 52.9, 26.2. FT-IR: ν/cm^{-1} = 3088, 2475, 1664, 1504, 1423, 1182, 1127, 975, 833, 799, 721, 702. Mp: 122−125 °C. $[\alpha]_D^{20}$ =

+19 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 458.2 (100%), 459.2 (24.4%), 460.2 (3.9%). Calculated: 458.2 (100%), 459.2 (22.7%), 460.2 (2.5%). Purity:100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.90 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(phenethyl)amino)butanoic Acid Trifluoroacetate Salt (27r). To a solution of **26r** (113 mg, 0.17 mmol) in DCM (1.5 mL) at 5 °C were added TFA (1.5 mL) and H$_2$O (200 μL). The solution was kept at 5 °C for 3 d. After the reaction was diluted and co-distilled with DCM (3 × 20 mL), the residue was once more dissolved in DCM (1.5 mL) and TFA (1.5 mL). The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was diluted and co-distilled with DCM (5 × 20 mL). The residue was dissolved in H$_2$O (7 mL) and was dried by lyophilization to give **27r** (103 mg, 0.17 mmol, 99%, 1.5 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.31 (s, 1H), 8.19 (s, 1H), 7.17−6.93 (m, 5H), 5.98 (d, J = 4.2 Hz, 1H), 4.61 (t, J = 4.7 Hz, 1H), 4.45−4.36 (m, 1H), 4.34−4.28 (m, 1H), 3.91 (dd, J = 8.2, 4.7 Hz, 1H), 3.72−3.28 (m, 6H), 2.94−2.82 (m, 2H), 2.43−2.23 (m, 1H), 2.21−2.07 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 171.8, 152.9, 149.7, 146.7, 144.3, 137.2, 129.9, 129.7, 128.2, 121.0, 91.9, 80.3, 74.5, 73.5, 56.3, 55.7, 52.5, 52.3, 30.6, 26.0. FT-IR: ν/cm^{-1} = 3066, 1666, 1505, 1425, 1187, 1131, 976, 835, 799, 722, 702. Mp: 79−82 °C. $[\alpha]_D^{20}$ = +18 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 472.2 (100%), 473.3 (26.7%), 474.2 (3.8%). Calculated: 472.2 (100%), 473.2 (23.8%), 474.2 (2.7%). Purity: 99% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.87 min.

(2S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(but-3-yn-2-yl)amino)-butanoic Acid Trifluoroacetate Salt (27s-A/B). To a solution of **26s** (31 mg, 0.05 mmol) in DCM (1.5 mL) at 5 °C was added TFA (1.5 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid was detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL), and the residue was dissolved in water (1.8 mL) and TFA (0.3 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was dried by lyophilization to give **27s** (39 mg, 0.05 mmol, 99%, 3.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.43−8.28 (m, 2H), 6.02 (t, J = 3.8 Hz, 1H), 4.68−4.56 (m, 1H), 4.38−4.12 (m, 3H), 4.00−3.83 (m, 1H), 3.53−3.25 (m, 4H), 3.04 (dd, J = 3.6, 2.2 Hz, 1H), 2.32−2.14 (m, 1H), 2.13−1.98 (m, 1H), 1.42 (dd, J = 29.3, 6.9 Hz, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 172.1, 171.7, 152.9, 149.8, 146.7, 144.1, 121.0, 91.8, 91.5, 82.6, 81.6, 79.6, 79.1, 78.4, 78.1, 74.8, 73.5, 73.3, 55.0, 54.6, 53.7, 52.3, 51.4, 51.2, 50.7, 27.4, 27.3, 18.2, 18.2. FT-IR: ν/cm^{-1} = 3091, 1670, 1508, 1427, 1323, 1196, 1134, 835, 799, 722. Mp: 60−63 °C. $[\alpha]_D^{20}$ = +12 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 420.2 (100%), 421.1 (21.7%), 422.2 (3.2%). Calculated: 420.2 (100%), 421.2 (19.5%), 422.2 (1.8%). Purity: 100% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.95 min.

(2S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(but-3-yn-2-yl)amino)-butanoic Acid Trifluoroacetate Salt (27s-A). To a solution of **26s-A** (36 mg, 0.06 mmol) in DCM (1.5 mL) at 5 °C was added TFA (1.5 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid was detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL), and the residue was dissolved in water (3 mL) and TFA (0.5 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was dried by lyophilization to give **27s-A** (39 mg, 0.06 mmol, 99%, 2.1 equiv TFA) as a colorless trifluoroacetate salt (resin). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.53 (s, 1H), 8.36 (s, 1H), 5.94 (d, J = 4.9 Hz, 1H), 4.63 (t, J = 4.9 Hz, 1H), 4.20−4.06 (m, 2H), 3.99−3.85 (m, 2H), 3.33 (d, J = 2.0 Hz, 1H), 3.16−3.04 (m, 1H), 2.9−2.81 (m, 2H), 2.80−2.63 (m, 1H), 2.04−1.92 (m, 2H), 1.32 (d, J = 6.9 Hz, 3H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ/ppm = 170.8, 153.4, 148.9, 148.8, 141.4, 119.2, 88.3, 82.5, 81.5, 76.3, 72.8, 71.9, 53.6, 50.8, 49.2, 47.9, 27.7, 19.1. FT-IR: ν/cm^{-1} = 2952, 2256, 2126, 1670, 1505, 1419, 1324, 1197, 1127, 1024, 997, 826, 798, 719. $[\alpha]_D^{20}$ = −3 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 420.2 (100%), 421.1 (21.3%),

422.2 (3.5%). Calculated: 420.2 (100%), 421.2 (19.5%), 422.2 (1.8%). Purity: 96% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.31 min.

(2S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(but-3-yn-2-yl)amino)-butanoic Acid Trifluoroacetate Salt (27s-B). To a solution of **26s-B** (74 mg, 0.12 mmol) in DCM (1.5 mL) at 5 °C was added TFA (1.5 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid was detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL), and the residue was dissolved in water (3 mL) and TFA (0.5 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was dried by lyophilization to give **27s-B** (58 mg, 0.12 mmol, 99%, 0.8 equiv TFA) as a colorless trifluoroacetate salt (resin). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.57 (s, 1H), 8.43 (s, 1H), 5.97 (d, J = 4.7 Hz, 1H), 4.62 (t, J = 4.7 Hz, 1H), 4.25−4.13 (m, 2H), 4.07−3.90 (m, 2H), 3.51−3.42 (m, 1H), 3.23−3.09 (m, 1H), 3.07−2.81 (m, 3H), 2.15−2.02 (m, 1H), 2.02−1.87 (m, 1H), 1.27 (d, J = 6.7 Hz, 3H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ/ppm = 171.0, 153.1, 149.1, 148.3, 142.2, 119.5, 88.8, 81.3 (2× C), 77.5, 73.6, 72.1, 53.7, 51.5, 49.4, 48.8, 27.3, 18.7. FT-IR: ν/cm^{-1} = 2941, 2255, 2127, 1674, 1505, 1418, 1322, 1197, 1128, 1024, 1001, 824, 798, 719. $[\alpha]_D^{20}$ = +20 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 420.2 (100%), 421.1 (21.9%), 422.2 (3.2%). Calculated: 420.2 (100%), 421.2 (19.5%), 422.2 (1.8%). Purity: 98% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.39 min.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-di-hydroxytetrahydrofuran-2-yl)methyl)(2-methylbut-3-yn-2-yl)-amino)butanoic Acid Trifluoroacetate Salt (27t). To a solution of **26t** (25 mg, 0.04 mmol) in DCM (1.5 mL) at 5 °C was added TFA (1.5 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid was detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL), and the residue was dissolved in water (1.8 mL) and TFA (0.3 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the reaction was dried by lyophilization to give **27t** (73 mg, 0.11 mmol, 99%, 2.0 equiv TFA) as a colorless trifluoroacetate salt. ^{1}H NMR (300 MHz, CD$_3$OD): δ/ppm = 8.24 (s, 2H), 5.95 (d, J = 3.5 Hz, 1H), 4.51 (t, J = 3.5 Hz, 1H), 4.38−4.24 (m, 2H), 3.85 (t, J = 6.4 Hz, 1H), 3.69−3.35 (m, 4H), 3.17 (s, 1H), 2.37−2.20 (m, 1H), 2.16−1.97 (m, 1H), 1.55 (d, J = 4.7 Hz, 6H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ/ppm = 170.8, 151.8, 145.8, 145.7, 116.5, 112.8, 90.5, 81.5, 80.6, 77.0, 73.4, 72.1, 61.1, 54.3, 49.9, 27.4, 25.1, 24.9. FT-IR: ν/cm^{-1} = 3095, 1663, 1506, 1427, 1186, 1133, 834, 799, 722. Mp: 74−77 °C. $[\alpha]_D^{20}$ = +21 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 434.2 (100%), 435.2 (22.6%), 436.2 (2.9%). Calculated: 434.2 (100%), 435.2 (20.5%), 436.2 (2.0%). Purity: 96% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 2.18 min.

Ethyl (2S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(but-3-yn-2-yl)amino)-butanoate (28-A/B). To a solution of **27-A/B** (99 mg, 0.13 mmol) in absolute ethanol (6 mL) was added thionyl chloride (77 µL, 1.07 mmol, 8.0 equiv) at 0 °C. The solution was stirred at 0 °C for 5 min and then heated for 9 h at 50 °C. After the solvent was evaporated under reduced pressure at 40 °C, the residue was purified by column chromatography (DCM/MeOH = 10:1 + 1% NEt$_3$) to give **28-A/B** (15 mg, 0.04 mmol, 27%) as a colorless powder. The diastereomers **28-A** and **28-B** were separated using a hydrophilic column (MZ Aqua Perfect), MeCN/H$_2$O = 9:91 + 0.1% TFA. After lyophilization **28-A** (10 mg, 0.01 mmol, 2.7 equiv TFA) and **28-B** (6 mg, 0.01 mmol, 2.0 equiv TFA) were obtained as colorless trifluoroacetate salts (resins). **28-A**: ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.52 (s, 1H), 8.35 (s, 1H), 5.93 (d, J = 5.0 Hz, 1H), 4.61 (t, J = 5.0 Hz, 1H), 4.25−4.12 (m, 3H), 4.11−4.01 (m, 2H), 3.91−3.80 (m, 1H), 3.35−3.27 (m, 1H), 3.12−2.99 (m, 1H), 2.94−2.72 (m, 1H), 2.68−2.56 (m, 1H), 2.04−1.90 (m, 2H), 1.31 (d, J = 6.9 Hz, 3H), 1.20 (t, J = 7.1 Hz, 3H). $[\alpha]_D^{20}$ = −1 (10 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 448.2 (100%), 449.2 (24.1%), 450.2 (3.2%). Calculated: 448.2 (100%), 449.2 (24.2%), 450.2 (3.4%). Purity: 99% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.78 min. **28-B**: ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.46 (s, 1H), 8.30 (s, 1H),

5.92 (d, J = 4.9 Hz, 1H), 4.61 (t, J = 4.9 Hz, 1H), 4.23−4.03 (m, 5H), 3.87−3.76 (m, 1H), 3.38−3.28 (m, 1H), 3.04−2.90 (m, 1H), 2.89−2.75 (m, 2H), 2.74−2.65 (m, 1H), 2.03−1.84 (m, 2H), 1.25−1.17 (m, 6H). $[\alpha]_D^{20}$ = +33 (3 mg/mL; MeOH). ESI-MS: m/z [M + H]$^+$ = 448.2 (100%), 449.2 (26.0%), 450.1 (3.6%). Calculated: 448.2 (100%), 449.2 (24.2%), 450.2 (3.4%). Purity: 99% (HPLC, MeCN/H$_2$O = 20:80 + 0.1% HCOOH); t_R = 3.68 min.

Molecular Docking. Molecular docking was conducted using the FlexX algorithm within the LeadIT-2.3.2 software under default parameters.[78,79] The protein structure of the human DNMT2−SAH complex (PDB-ID 1G55)[30] was obtained from the Protein Data Bank (PDB).[80] The binding site was defined to include all amino acids within a 6.5 Å radius around the crystallographic ligand SAH and water molecules 413, 414, 459, and 495 in a predefined orientation forming at least three interactions with SAH and the binding site. The receptor was protonated using the Protoss module[81] within LeadIT. For molecular docking, all ligands were energy-minimized using the Merck molecular force field (MMFF94x)[82] within MOE 2020.09.[83] The docking setup was validated by redocking of SAH (Table S1 Supporting Information).

Solvent analysis was performed using the 3D-RISM approach[84] within MOE.[83] The DNMT2−SAH complex structure PDB-ID 1G55 was prepared using the QuickPrep module to add missing atoms and for assignment of protonation states without further energy minimization to preserve crystallographic water positions (structure resolution 1.80 Å). Receptor and ligand atoms 10 Å around SAH were considered in solvent analysis using a grid spacing of 0.25 Å.

Expression and Purification of Human Full-Length DNMT2. The plasmid coding for the full-length protein of human DNMT2 as described by Dong et al.[30] was kindly provided by Albert Jeltsch (University of Stuttgart, Germany). The protein was expressed in *E. coli* Rosetta 2 (DE3) pLysS cells (Merck KGaA, Darmstadt, Germany), using a pET-28a vector to obtain an N-terminal polyhistidine tagged protein. The sequence was then verified by Eurofins Genomics Europe (Ebersberg, Germany). Cells were grown in TB medium at 37 °C until they reached an optical density at 600 nm of 0.8. Overexpression was induced by adding IPTG up to a final concentration of 500 µM and was maintained overnight at 20 °C. Cells were harvested by centrifugation at 4000g for 30 min at 4 °C; afterward the cell pellets were resuspended in lysis buffer (50 mM sodium phosphate, pH 8.0, 150 mM NaCl, 25 mM imidazole, 0.1% polysorbate-20). After incubation on ice for 40 min with lysozyme (Carl Roth, Karlsruhe, Germany), cells were disrupted by sonication on ice in 15 s intervals. Cell debris was removed by centrifugation at 17 500g for 60 min at 4 °C. The clear supernatant was filtered using Chromafil Xtra RC 25 mm 0.45 µM syringe filters (Macherey-Nagel, Düren, Germany). For purification, an ÄKTA start system (Cytiva, Marlborough, Chicago, USA) was used. Therefore, the supernatant was loaded on a HisTrap HP 5 mL column (Cytiva, Marlborough, Chicago, USA), a wash-out with several column volumes lysis buffer was conducted, and finally DNMT2 was eluted using a gradient up to 100% elution buffer (50 mM sodium phosphate, pH 8.0, 150 mM NaCl, 800 mM imidazole, 0.1% polysorbate-20). To remove DNA and RNA fragments, eluted DNMT2 was diluted 1:10 with IEX buffer A (10 mM sodium phosphate, pH 8.0, 0.1% polysorbate-20) and loaded on a HiTrap Q HP 5 mL column (Cytiva, Marlborough, USA). Several column volumes of IEX buffer A were used for wash-out of unspecific bound anions, followed by a gradient up to 100% IEX buffer B (10 mM sodium phosphate, 1 M NaCl, 0.1% polysorbate-20) to elute DNMT2. DNA and RNA fragments eluted at higher concentrations of NaCl compared to DNMT2. For further purification, size exclusion chromatography using a Superdex 16/600 75 pg column (Cytiva, Marlborough, USA) equilibrated in SEC buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1 mM EDTA, 2 mM DTT, 0.1% polysorbate-20) was performed. DNMT2 was then concentrated using Amicon Ultra 4 10K centrifugal filters (Merck KGaA, Darmstadt, Germany) and diluted 1:4 with storage buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1 mM EDTA, 2 mM DTT, 0.1% polysorbate-20, 60% glycerol). Protein was stored in liquid state at −20 °C for further use.

Expression and Purification of Murine DNMT3A-3L Single Chain Fusion Protein. The plasmid coding for the murine

DNMT3A-3L single chain fusion protein (pET-Dnmt3A3L-sc27) was kindly gifted from Albert Jeltsch (Addgene plasmid no. 71827; http://n2t.net/addgene:71827; RRID:Addgene_71827).[85] The sequence was verified by Eurofins genomics (Ebersberg, Germany). For protein expression, *E. coli* Rosetta 2 (DE3) pLysS cells (Merck KGaA, Darmstadt, Germany) were transformed and grown in TB medium to an optical density of 0.5. Overexpression was induced by addition of IPTG to a final concentration of 500 μM, and cells were maintained overnight at 16 °C. Cells were harvested by centrifugation at 4500*g* at 4 °C for 30 min; afterward cells were washed in lysis buffer (30 mM sodium phosphate, pH 7.3, 500 mM NaCl, 10 mM imidazole, 0.2 mM DTT, 0.1% polysorbate-20, 10% glycerol) and centrifuged again at 10 000*g* at 4 °C for 10 min. Clear supernatant was removed, and cells were flash frozen in liquid nitrogen to be stored at −80 °C until further use.

Frozen cells were resuspended in lysis buffer and incubated with lysozyme on ice for 30 min. Afterward, cells were lysed by sonication in 45 s intervals on ice, and cell debris was removed by centrifugation at 17 000*g* at 4 °C for 90 min. Clear supernatant was filtered using Chromafil Xtra RC 25 mm 0.45 μM syringe filters (Macherey-Nagel, Düren, Germany) and loaded on a HisTrap HP 5 mL column (Cytiva, Marlborough, Chicago, USA). A wash-out step with several column volumes of lysis buffer was conducted, and finally the DNMT3A-3L single chain fusion protein was eluted using a gradient up to 100% elution buffer (30 mM sodium phosphate, pH 7.3, 500 mM NaCl, 500 mM imidazole, 0.2 mM DTT, 0.1% polysorbate-20, 10% glycerol). A further purification step was conducted using a Superdex 16/600 75 pg column (Cytiva, Marlborough, USA) equilibrated in SEC buffer (20 mM HEPES, pH 7.3, 200 mM sodium chloride, 1 mM DTT, 0.1% polysorbate-20, 10% glycerol). Eluted protein was then concentrated using Amicon Ultra 4 10K centrifugal filters (Merck KGaA, Darmstadt, Germany) and adjusted to a final concentration of 40% glycerol. Protein was stored at −20 °C until further use.

Expression and Purification of Human NSUN2. The gene coding for the full-length human NSUN2 (Uniprot ID Q08J23) was synthesized and inserted into a pET28a(+) vector between *Eco*RI and *Xho*I sites by Genscript (Piscataway, New Jersey, USA) as previously described.[86] The sequence was verified by Eurofins Genomics Europe (Ebersberg, Germany). The plasmid used for this study will be available at Addgene (Addgene ID: 188059). For protein expression, *E. coli* Rosetta 2 (DE3) pLysS cells (Merck KGaA, Darmstadt, Germany) were transformed and grown as described for the murine DNMT3A-3L single chain fusion protein. Purification was conducted quite similarly to that described for DNT3A-3L with some adjustments to the buffers used. Cells were lysed and washed with lysis buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1% polysorbate-20) and then eluted with elution buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1% polysorbate-20). For size-exclusion chromatography, a Superdex 16/600 75 pg column (Cytiva, Marlborough, USA) equilibrated in SEC buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1 mM EDTA, 2 mM DTT, 0.1% polysorbate-20) was used. The eluted protein was then concentrated using Amicon Ultra 4 10K centrifugal filters (Merck KGaA, Darmstadt, Germany) and adjusted to a final concentration of 40% glycerol. The protein was stored at −20 °C until further use.

Expression and Purification of Human NSUN6. The gene coding for the full-length human NSUN6 (Uniprot ID Q8TEA1) was synthesized and inserted into a pET22b(+) plasmid between *Nde*I and *Xho*I sites by Genscript (Piscataway, New Jersey, USA), as described earlier in literature.[87] The sequence was verified by Eurofins Genomics Europe (Ebersberg, Germany). The plasmid used for this study will be available at Addgene (Addgene ID: 188060). For protein expression, *E. coli* Rosetta 2 (DE3) pLysS cells (Merck KGaA, Darmstadt, Germany) were transformed and grown as described for the murine DNMT3A-3L single chain fusion protein. Purification was conducted quite similarly to that described for DNMT3A-3L with some adjustments to the buffers used. Cells were lysed and washed with lysis buffer (50 mM sodium phosphate, pH 7.5, 150 mM NaCl, 10 mM imidazole, 0.1% polysorbate-20) and then eluted with elution buffer (50 mM sodium phosphate, 7.5, 150 mM NaCl, 500 mM imidazole, 0.1% polysorbate-

20). For size-exclusion chromatography, a Superdex 16/600 75 pg column (Cytiva, Marlborough, USA) equilibrated in SEC buffer (25 mM HEPES, pH 7.5, 300 mM NaCl, 1 mM DTT, 0.1% polysorbate-20, 10% glycerol) was used. The protein was concentrated using Amicon Ultra 4 10K centrifugal filters (Merck KGaA, Darmstadt, Germany) and diluted 1:4 with storage buffer (25 mM HEPES, pH 7.5, 300 mM NaCl, 1 mM DTT, 0.1% polysorbate-20, 60% glycerol). Protein was stored at −20 °C until further use.

Expression and Purification of the Methyltransferase Domain of Human EHMT2 (G9a). The plasmid coding for the methyltransferase domain of human EHMT2 (G9a) was a kind gift from Cheryl Arrowsmith (Addgene plasmid no. 25503; http://n2t.net/addgene:25503; RRID:Addgene_25503). The sequence was verified by Eurofins Genomics Europe (Ebersberg, Germany). For protein expression, *E. coli* Rosetta 2 (DE3) pLysS cells (Merck KGaA, Darmstadt, Germany) were transformed and grown as described for the murine DNMT3A-3L single chain fusion protein. The purification was conducted quite similarly to that described for DNMT3A-3L with some adjustments to the buffers used. Cells were lysed in lysis buffer (100 mM sodium phosphate, pH 7.4, 250 mM NaCl, 0.1% 2-mercaptoethanol, 0.1% polysorbate-20, 5% glycerol), washed with wash buffer (100 mM sodium phosphate, pH 7.4, 250 mM NaCl, 50 mM imidazole, 0.1% polysorbate-20, 5% glycerol), and then eluted with elution buffer (100 mM sodium phosphate, pH 7.4, 250 mM NaCl, 250 mM imidazole, 0.1% polysorbate-20, 5% glycerol). For size-exclusion chromatography, a Superdex 16/600 75 pg column (Cytiva, Marlborough, USA) equilibrated in SEC buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% polysorbate-20) was used. The protein was flash frozen in liquid nitrogen and stored at −80 °C until further use.

Expression and Purification of T7 Polymerase. For the protein expression, competent BL21 (DE3) pLysS cells (Merck KGaA, Darmstadt, Germany) were transformed with plasmid obtained by Addgene (pQE 80L Kan T7WT was a gift from Andrew Ellington (Addgene plasmid no. 102790; http://n2t.net/addgene:102790; RRID:Addgene_102790). Cells were grown in LB medium at 37 °C until an optical density of 0.8 at 600 nm was reached. After induction of the expression by 500 μM IPTG, the expression was maintained overnight at 37 °C. Disruption of cells was performed as described above using lysis buffer$_{T7}$ (25 mM sodium phosphate, pH 7.4, 500 mM NaCl, 25 mM imidazole). For purification, the clear supernatant was loaded onto a HisTrap HP 5 mL column, and protein was eluted performing a gradient up to 100% elution buffer$_{T7}$ (25 mM sodium phosphate, pH 7.4, 500 mM NaCl, 500 mM imidazole). For storage, the eluted protein was rebuffered into 2× storage buffer$_{T7}$ (100 mM Tris-HCl, pH 7.8, 200 mM NaCl, 0.2 mM EDTA, 2 mM DTT, 0.2% Triton X-100) using Amicon Ultra 4 10K centrifugal filters (Merck KGaA, Darmstadt, Germany) and was diluted 1:1 very gently with glycerol. Protein was stored until further use at −20 °C.

Polymerase Chain Reaction (PCR). For the preparation of IVT templates for tRNA^{Asp}, PCR reactions were carried out in final volumes of 200 μL. Taq DNA Polymerase (0.05 U μL^{-1}; NEB; units defined by the supplier) was added to the reaction buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl$_2$) containing 3 mM MgCl$_2$, 400 μM dNTP-Mix, 2 μM of each primer (forward, 5′-CGC GCG AAG CTT AAT ACG ACT CAC TAT A-3′; reverse, 5′-TGG CGG GCC GTC G-3′) and 10 nM template (5′-TGG CGG GCC GTC GGG GAA TCG AAC CCG CGT CTC CCG CGT GAC AGG CGG GGA TAC TCA CCA CTA TAC TAA CGA CCC TAT AGT GAG TCG TAT T-3′). After an initial denaturation step (2 min at 90 °C), 35 PCR cycles were performed with annealing (30 s at 59 °C), elongation (45 s at 72 °C), and denaturation (30 s at 90 °C) and a final elongation step of 5 min.

The template of tRNA^{Thr} was prepared accordingly with primers 5′-CGC GCG AAG CTT AAT ACG ACT CAC TAT A-3′ (forward) and 5′-TGG AGG CCC CGC TGG GAG TCG AA-3′ (reverse) and the following template: 5′-TAG TCG TAA GCT GAT ATG GCT GAT TAG TCG GAA GCA TCG AAC GCT GAT TGG AGG CCC CGC TGG GAG TCG AAC CCA GGA TCT CCT GTT TAC TAG ACA GGC GCT TTA ACC AAC TAA GCT ACG GAG CCT ATA GTG AGT CGT ATT A-3′.

In Vitro Transcription (IVT) and Purification of tRNA^Asp and tRNA^Thr.

A portion (400 μL) of the PCR reaction product without further purification was added to reaction mix to reach a final volume of 1 mL in transcription buffer (40 mM Tris-HCl, pH 8.8, 1 mM spermidine, 5 mM DTT, 0.01% Triton X-100) containing 30 mM MgCl$_2$, 5 mM DTT, 5 mM of each NTP, 2.5 μg mL^{-1} BSA, and 50 μg mL^{-1} of T7 polymerase. Reactions were carried out at 37 °C for 4 h.

The DNA template was digested by addition of 1 U RNase-free DNase I (Thermo Scientific; units defined by the supplier) and 100 μL DNase I buffer (100 mM Tris-HCl, pH 7.5, 25 mM MgCl$_2$, 1 mM CaCl$_2$, Thermo Scientific) to the transcription reaction and incubation for 1 h at 37 °C.

In vitro transcribed RNA was purified by 10% denaturing polyacrylamide gel electrophoresis (PAGE) for 90 min at 15 W. RNA was visualized by UV shadowing and the corresponding gel area was excised from the gel. After overnight elution in 0.5 M NH$_4$OAc and Nanosep filtering (0.45 μm, VWR), the RNA was precipitated with EtOH.

Purified RNA was dissolved in Milli-Q water. The quality was checked on 10% denaturing PAGE gels after staining with GelRed (Biotrend Chemikalien GmbH, Köln, Germany) and visualizing RNA on a Typhoon 9400 at an excitation wavelength of 525 nm. The concentration was determined on a Nanodrop 2000 Spectrometer (Thermo Scientific, Waltham, USA) by measuring the absorption at 260 and 280 nm.

Microscale Thermophoresis (MST).

A screening of all compounds for binding was performed on a Monolith NT.115 Pico instrument (NanoTemper Technologies, Muenchen, Germany). His6-tagged DNMT2 was labeled using the Monolith His-Tag Labeling Kit RED-Tris-NTA second generation, according to the manufacturer's instructions. The labeling strategy was chosen due to the existence of a His6 tag within the recombinant protein, which should allow specific labeling without interfering with the binding site of the protein. Labeled protein was diluted to 10 nM into MST buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM MgCl$_2$, 1 mM DTT, 0.05% polysorbate-20, 0.1% PEG-8000), and the compound was added in a final concentration of 100 μM. HEPES was recommended by the manufacturer as a reasonable buffer; moreover, addition of 0.05% polysorbate-20 and 0.1% PEG-8000 stabilized the protein and prevented aggregation during the measurements. Measurements were performed in quadruplets in MST buffer at 25 °C and at medium MST power in Monolith NT.115 Capillaries (NanoTemper Technologies, Muenchen, Germany). Medium MST power was sufficient to induce a shift in thermophoresis; therefore the usage of high MST power was waived to avoid stress for the protein. Furthermore, special care was taken to always use fresh protein and buffers to reduce adsorption and aggregation and also to maintain reducing conditions for the protein, which was crucial for the interactions. The protein was incubated with compound for 5 min at room temperature prior to measurements. The raw data were analyzed using the MO.Affinity Analysis software (NanoTemper Technologies, Muenchen, Germany). For binder/nonbinder discrimination, a 99% confidence interval of normalized fluorescence was calculated. If there was at least a 1‰ fluorescence shift, based on normalized fluorescence between the confidence intervals of testing compound and control, the compound was defined as a binder.

Isothermal Titration Calorimetry (ITC).

Calorimetric experiments for determination of binding affinity of detected inhibitors were conducted with a MicroCal PEAQ-ITC Automated (Malvern Panalytical, Malvern, UK). DNMT2 was concentrated and buffer was exchanged into ITC buffer (50 mM sodium phosphate, pH 7.5, 300 mM NaCl, 0.5 mM EDTA, 2 mM β-mercaptoethanol, 0.05% polysorbate-20) using Amicon Ultra 4 10K centrifugal filters (Merck KGaA, Darmstadt, Germany) to a final concentration of 30 μM. Stocks of compounds were provided as 25 mM stocks in DMSO and were diluted to 300 μM with ITC buffer. Measurements were performed in 13 injections at 25 °C in triplicate. Data were analyzed and fitted using MicroCal PEAQ-ITC Analysis Software 1.21. (Malvern Panalytical, Malvern, UK). Values were not buffer-corrected, considering the low ionization heat of phosphate buffers.[88]

For analysis of the binding site, DNMT2 was spiked with 300 μM **27m** instead of DMSO and the titration was performed with 300 μM SAH, which served as a known binder. Parameters and analysis were as described in the preceding.

DNMT2 Activity Assay.

The enzymatic assay was carried out in 100 mM Tris-HCl, pH 8, 100 mM NH$_4$OAc, 0.1 mM EDTA, 10 mM MgCl$_2$, and 10 mM DTT. SAM was added as a mixture of ^{3}H-SAM (Hartmann Analytics) and cold SAM (NEB) to final concentrations of 0.025 μCi μL^{-1} and 0.9 μM. tRNA^Asp was heated to 75 °C for 5 min and cooled down to room temperature before it was added to the reaction mixture to a final concentration of 5 μM. The amount of DMSO was adjusted to 5%. The reactions were started by adding 250 nM DNMT2 to a final volume of 20 μL and carried out at 37 °C.

At 0 and 20 min, aliquots of 8 μL were taken from the reaction mixture and spotted on Whatman glass microfiber filters (GF/C, 25 mm). The RNA was precipitated on the filters in 5% ice cold TCA for at least 15 min. The filters were washed twice with 5% TCA at room temperature for 20 and 10 min and once in EtOH for 10 min. After drying, the filters were transferred into scintillation vials and 3 mL of Gold MV liquid scintillation cocktail (PerkinElmer, Waltham, USA) was added. Scintillation was measured on a scintillation counter (TriCarb Liquid Scintillation Analyzer 4810TR) with a measurement time of 1 min.

For compound screening, compounds were added to final concentrations of 100 μM. Percentage of inhibition at this concentration was obtained by referencing scintillation increase to a positive control without compound. Errors refer to the standard deviation of three independent measurements.

For IC$_{50}$ value determination, compounds were analyzed at a minimum of seven different concentrations in experimental triplicates. IC$_{50}$ values were calculated by exponential fitting of the relative enzymatic activities against the inhibitor concentrations using the LL.2 function of the drc package, version 3.0-1,[89] in RStudio.[90] IC$_{50}$ errors are given as standard errors.

DNMT3A-3L Activity Assay.

The enzymatic assay was carried out in 50 mM KPi, pH 7.8, 1.0 mM EDTA, 20 mM NaCl, 0.2 mg mL^{-1} BSA, and 1.0 mM DTT. ^{3}H-SAM (Hartmann Analytics) was added to a final concentration of 0.63 μM and 0.05 μCi μL^{-1}.

Substrate DNA was prepared by annealing 7.5 μM of oligos 5'-GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC-3' and 5'-GAC GAC GAC GAC GAC GAC GAC GAC GAC GAC GAC GAC-3' (Biomers) in 100 mM HEPES, pH 7.5, and 15 mM KOAc by heating to 95 °C and slowly cooling down to room temperature.

Duplex DNA was added to the reaction mixture to a final concentration of 750 nM. The amount of DMSO was adjusted to 5%. The reactions were started by adding 1 μM DNMT3A-3L to a final volume of 20 μL and carried out at 37 °C.

At 0 and 150 min, aliquots of 8 μL were taken from the reaction mixture and spotted on Whatman glass microfiber filters (GF/C, 25 mm). The DNA was precipitated on the filters in 5% ice cold TCA for at least 15 min. The filters were washed twice with 5% TCA at room temperature for 10 and 20 min and once in EtOH for 10 min, respectively. After drying, the filters were transferred into scintillation vials, and 3 mL of Gold MV liquid scintillation cocktail (PerkinElmer, Waltham, USA) was added. Scintillation was measured on a scintillation counter (TriCarb Liquid Scintillation Analyzer 4810TR) with a measurement time of 1 min.

Compounds were added to final concentrations of 100 μM. Percentage of inhibition at this concentration was obtained by referencing scintillation increase to a positive control without compound. Errors refer to the standard deviation of three independent measurements.

NSUN2 Activity Assay.

The enzymatic assay was carried out in 50 mM Tris-HCl, pH 7.5, 5.0 mM EDTA, 5.0 mM MgCl$_2$, 10% glycerol, and 1.5 mM DTT. SAM was added as a mixture of ^{3}H-SAM (Hartmann Analytics) and cold SAM (NEB) to final concentrations of 0.025 μCi μL^{-1} and 0.9 μM. Total tRNA of *E. coli* (Roche Diagnostics GmbH) was heated to 75 °C for 5 min and cooled down to room temperature before it was added to the reaction mixture to a final concentration of 123 ng μL^{-1}. The amount of DMSO was adjusted to 5%. The reactions

were started by adding 250 nM NSUN2 to a final volume of 20 μL and carried out at 37 °C.

At 0 and 20 min, aliquots of 8 μL were taken from the reaction mixture and spotted on Whatman glass microfiber filters (GF/C, 25 mm). The RNA was precipitated on the filters in 5% ice cold TCA for at least 15 min. The filters were washed twice with 5% TCA at room temperature for 10 and 20 min and once in EtOH for 10 min, respectively. After drying, the filters were transferred into scintillation vials and 3 mL of Gold MV liquid scintillation cocktail (PerkinElmer, Waltham, USA) was added. Scintillation was measured on a scintillation counter (TriCarb Liquid Scintillation Analyzer 4810TR) with a measurement time of 1 min.

Compounds were added to final concentrations of 100 μM. Percentage of inhibition at this concentration was obtained by referencing scintillation increase to a positive control without compound. Errors refer to the standard deviation of three independent measurements.

NSUN6 Activity Assay. The enzymatic assay was carried out in 50 mM Tris-HCl, pH 7.0, 50 mM NaCl, 5.0 mM MgCl$_2$, and 1.0 mM DTT. SAM was added as a mixture of ^{3}H-SAM (Hartmann Analytics) and cold SAM (NEB) to final concentrations of 1.2 μM and 0.038 μCi μL^{-1}. tRNA^{Thr} was heated to 75 °C for 5 min and cooled down to room temperature before it was added to the reaction mixture to a final concentration of 1 μM. The amount of DMSO was adjusted to 5%. The reactions were started by adding 30 nM NSUN6 to a final volume of 20 μL and carried out at 37 °C.

At 0 and 20 min, aliquots of 8 μL were taken from the reaction mixture and spotted on Whatman glass microfiber filters (GF/C, 25 mm). The RNA was precipitated on the filters in 5% ice cold TCA for at least 15 min. The filters were washed twice with 5% TCA at room temperature for 20 and 10 min and once in EtOH for 10 min. After drying, the filters were transferred into scintillation vials, and 3 mL of Gold MV liquid scintillation cocktail (PerkinElmer, Waltham, USA) was added. Scintillation was measured on a scintillation counter (TriCarb Liquid Scintillation Analyzer 4810TR) with a measurement time of 1 min.

Compounds were added to final concentrations of 100 μM. Percentage of inhibition at this concentration was obtained by referencing scintillation increase to a positive control without compound. Errors refer to the standard deviation of three independent measurements.

G9a Activity Assay. The enzymatic assay was carried out in 20 mM KPi, pH 8.0, 2.0 mM MgCl$_2$, 0.01% Tween-20. SAM was added as a mixture of ^{3}H-SAM (Hartmann Analytics) and cold SAM (NEB) to final concentrations of 0.025 μCi μL^{-1} and 0.9 μM. Histone H3 (AAs 1−25, Eurogentec) was added to the reaction mixture to a final concentration of 1.6 μM. The amount of DMSO was adjusted to 5%. The reactions were started by adding 30 nM G9a to a final volume of 20 μL and carried out at 37 °C.

At 0 and 20 min, aliquots of 8 μL were taken from the reaction mixture and spotted on Whatman glass microfiber filters (GF/C, 25 mm). The peptide was precipitated on the filters in 0.2 M ice cold NH$_4$OAc for at least 15 min. The filters were washed twice with 0.2 M NH$_4$OAc at room temperature for 10 and 20 min and once in acetone for 10 min. After drying, the filters were transferred into scintillation vials, and 3 mL of Gold MV liquid scintillation cocktail (PerkinElmer, Waltham, USA) was added. Scintillation was measured on a scintillation counter (TriCarb Liquid Scintillation Analyzer 4810TR) with a measurement time of 1 min.

Compounds were added to final concentrations of 100 μM. Percentage of inhibition at this concentration was obtained by referencing scintillation increase to a positive control without compound. Errors refer to the standard deviation of three independent measurements.

Cell Culture. HEK-293 cells were cultured in DMEM high glucose medium (Thermo Scientific) supplied with 10% (v/v) fetal bovine serum (FBS, Thermo Scientific) and 1% (v/v) Penicillin-Streptomycin (10000 U mL^{-1}, Thermo Scientific) in 175 cm^2 cell culture flasks at 37 °C and 5% CO$_2$. At a confluency of 80−90%, passaging was performed by washing once with 5 mL of Dulbecco's phosphate-buffered saline

(DPBS, Thermo Scientific) followed by adding 1 mL of trypsin-EDTA (Thermo Scientific) before splitting in a ratio of 1:20. Cells were passaged every 2−3 days.

Cell Viability Assays. Cell viability assays were performed as described in the literature[91] with minor adaptions. Briefly, HEK-293 cells were cultured in a humidified incubator at 37 °C with 5% CO$_2$ atmosphere as described in the previous section. Cell viability assay were performed in Dulbecco's modified Eagle medium (DMEM, no glucose, no glutamine, no phenol red, Thermo Scientific) supplemented with 10% (v/v) fetal bovine serum, 1% pyruvate (Thermo Scientific), and 1% Penicillin-Streptomycin (10000 U mL^{-1}). Cells (10000 per well) were seeded into 96-well microplates in a volume of 200 μL of supplemented DMEM. Compounds, provided as 50 mM stocks in DMSO, were added in a range of 500 μM to 800 nM (1% to 0.0016% DMSO, respectively) in duplicates to the wells. Negative controls were performed with DMSO treatment in the same concentrations as the compound solutions. After incubation for 24 h at 37 °C and 5% CO$_2$, 40 μL of a 3 mg mL^{-1} solution of 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT, Sigma-Aldrich) was added to each well. After a further incubation at 37 °C and 5% CO$_2$ for 20 min, the medium was removed and 225 μL of an 8:1 mixture of DMSO/(0.1 M glycine pH 10.5 (Merck KGaA), 0.1 M NaCl) was added to each well. Plates were shaken for 20 min, and the absorbance was determined at 595 and 670 nm using an Infinite M200 Pro plate reader (Tecan, Männedorf, Switzerland). Background absorbance at 670 nm was subtracted from absorbance at 595 nm. The cell viability was calculated as the ratio of corrected absorbance in compound treated cells to DMSO treated cells with corresponding DMSO concentrations.

In Vivo Methylation Assay. 750 000 cells per well were seeded in 6-well plates, and 2.5 mL of DMEM supplemented with 10% FBS and 1% Penicillin-Streptomycin was added. After 24 h, the medium was removed, and the experiment was initiated by the addition of 2.5 mL of supplemented DMEM including a final concentration of 100 μM compound (5 μL of a 50 mM stock in DMSO) or 5 μL of DMSO. Incubation was carried out for 24 h at 37 °C and 5% CO$_2$. To terminate the experiment, the medium was removed, and cells were washed with DPBS and directly harvested using 0.5 mL of Tri Reagent (Sigma-Aldrich).

CaCo-2 Cell Permeability Assay. Cell permeabilities of **27s-A/B** and **28-B** were assessed in transport studies across monolayers of a human enterocyte cell line. CaCo-2 cells (clone C2Bbe1, passage 49, LGC Standards, Wesel, Germany) were seeded at a density of 5 × 10^4 cells/cm^2 in 24-well Transwell inserts with a pore diameter of 3.0 μm (Corning Life Science, Corning, USA) in DMEM cell culture medium (Gibco, Thermo Fisher Scientific, Waltham, USA) supplemented with 10% fetal calf serum (Sigma-Aldrich, St. Louis, USA) and 1% nonessential amino acids (Gibco, Thermo Fisher Scientific, Waltham, USA). Cells were cultured for 14 days at 37 °C and 5% CO$_2$, and medium was exchanged every 2−3 days. Before and after the transport study, barrier integrity was evaluated via the transepithelial electrical resistance using chop stick electrodes (Millicell ERS-2, Merck Millipore, Burlington, USA). Compounds **27s-A/B** and **28-B** were dissolved in DMSO and applied to the apical side of CaCo-2 cell layers at a final concentration of 200 μM in cell culture medium (final concentration of 1% DMSO v/v) in four separate experiments. Compound-free DMSO solutions were used as controls. After 24 h, cell culture media from the apical and basolateral compartments were collected and lyophilized. The residues were each taken up in 200 μL of MeOH, centrifuged, and analyzed by LC-MS using an Agilent Zorbax SB-Aq column (4.6 × 150 mm^2) with the mobile phases MeCN/H$_2$O = 10:90 + 0.1% HCOOH or MeCN/H$_2$O = 5:95 + 0.1% HCOOH. The flow rate was 0.7 mL/min. Samples were applied using 5 μL injection, and the quantitation of the compounds in the apical and basolateral compartment was conducted using the AUCs of the respective peaks at 254 nm. The retention times using the two different mobile phases were as follows: t_R (**27s-A/B**) = 3.21 min (MeCN/H$_2$O = 10:90 + 0.1% HCOOH); t_R (**28-B**) = 9.25 min (MeCN/H$_2$O = 10:90 + 0.1% HCOOH); 4.96 min (**27s-A**, MeCN/H$_2$O = 5:95 + 0.1% HCOOH), 5.68 min (**27s-B**, MeCN/H$_2$O = 5:95 + 0.1% HCOOH). Analyses of

the solutions from the assays with **28-B** revealed 20% of ester **28-B** and 78% of acid in the donor compartment, whereas the acceptor compartment contained 2% acid. In the assays with the acid **27s-A/B**, only traces of the compound (ca. 0.6%) could be detected in the acceptor compartment. All values are mean values from the four independent experiments.

Extraction of Total RNA. Total RNA was extracted in Tri Reagent according to the manufacturer's protocol with chloroform (Honeywell Riedel-de Haën). Chloroform (200 μL per milliliter of Tri Reagent) was added, and the samples were vortexed and incubated at 25 °C for 2 min followed by centrifugation for 15 min at 12 000g at 4 °C. Total RNA contained in the upper phase was transferred into a new tube, and the procedure was repeated. Afterward, RNA was precipitated by addition of 500 μL of isopropanol (Honeywell Riedel-de Haën), incubation for 10 min, and centrifugation at 13 000g at 4 °C for 30−45 min. The supernatant was discarded, and pellets were washed with 300 μL of 75% ethanol (Carl Roth) and centrifuged again. Following this procedure, the pellet was dried and RNA was dissolved in 10 μL of RNase-free water. Concentrations were determined on a Nanodrop 2000 Spectrometer (Thermo Scientific, Waltham, USA) by measuring the absorption at 260 and 280 nm.

Isolation of Total tRNA. Total tRNA was isolated from total RNA by 10% denaturing polyacrylamide gel elution as described in previous sections with minor adaptions: the gel bands containing small RNA species (75 to 120 nucleotides) were excised after running the gel for 90 min at 15 W, staining, and visualization as described previously. RNA was eluted from the gel by adding 300 μL of NH$_4$OAc (0.5 M), shaking overnight at 850 rpm and 20 °C, and filtering through a Nanosep device (0.45 μm, VWR). Finally, RNA was EtOH precipitated, and the concentration was determined on a Nanodrop 2000 Spectrometer (Thermo Scientific, Waltham, USA) by measuring the absorption at 260 and 280 nm.

LC-MS/MS-Based Quantification of m^{5}C. Before measuring m^{5}C levels via LC-MS/MS, samples were digested down to nucleosides. Therefore, RNA was incubated for 2 h at 37 °C in 5 mM Tris-HCl (pH 8), 1 mM MgCl$_2$, 0.3 U of nuclease P1 from *Penicillium citrinum* (Merck KGaA), 0.1 U of phosphodiesterase I (venom exonuclease, Worthington Biochemical Corporation), 1 U of alkaline phosphatase from bovine intestinal mucosa (Merck KGaA), and 10 U of Benzonase Nuclease (Merck KGaA).

For relative quantification of m^{5}C, deuterated m^{5}C (D$_3$-m^{5}C) was added to digested samples as an internal standard. Samples were measured using an Agilent 1260 HPLC Infinity system equipped with a diode array detector (DAD) and a Synergy Fusion RP column (4 μm particle size, 80 Å pore size, 250 mm length, 2 mm inner diameter, Phenomenex). The HPLC was coupled to a triple quadrupole mass spectrometer (Agilent 6470) equipped with an electrospray ion source (ESI, Agilent Jet Stream). Column temperature was set to 35 °C and elution was performed at a flow rate of 0.350 mL min^{-1} using a gradient over 30 min with the following mobile phases: A (5 mM NH$_4$OAc adjusted to pH 5.3 using acetic acid) and B (100% acetonitrile, CHROMASOLV, Honeywell Riedel-de Haën). The gradient included a linear increase from 0−8% B (0−10 min) followed by 8−40% B (10−20 min), 40−0% B (20−23 min), and a final constant composition of 0% B (23−30 min). HPLC separation was followed by a photometrical measurement of the main nucleosides using a DAD at 254 nm prior to entering the triple quadrupole mass spectrometer, which was run at a gas temperature of 300 °C, gas flow of 7 L min^{-1}, nebulizer pressure of 60 psi, sheath gas temperature of 400 °C, sheath gas flow of 12 L min^{-1}, capillary voltage of 3000 V, and nozzle voltage of 0 V. Measurement was performed in the positive ion mode using the Agilent MassHunter software, and modified nucleosides were monitored in the dynamic multiple reaction monitoring (dynamic MRM) mode. The following nucleoside-to-base mass transitions were used for detection of m^{5}C and D$_3$-m^{5}C: 258 → 126 (m^{5}C), 261 → 129 (D$_3$-m^{5}C) with a retention time (m^{5}C and D$_3$-m^{5}C) of 7.6 min. Analysis was performed using the Agilent MassHunter Qualitative Analysis Software V10.0. For relative quantification, the ratio of unlabeled ^{12}C-m^{5}C and D$_3$-m^{5}C was calculated and normalized to the UV peak area of adenosine. The final

m^{5}C levels were related to the negative control with DMSO, which was set to 1.

tRNA-Bisulfite-MiSeq. For tRNA-bisulfite-MiSeq, primers were designed to in silico converted (C > U), unmethylated regions of tRNA^{Asp} according to Bormann et al.[92] Within the amplified region, target cytosine residue 38 was contained. For compatibility with Illumina index and sequencing adapters, 5′ and 3′ Illumina overhang adapters were appended to the tRNA^{Asp}-specific primer sequences (*vide infra*). For sequencing library preparation, 1 μg of total RNA isolated from compound-treated HEK-293 cells was bisulfite-converted using the EZ RNA Methylation Kit (Zymo Research). Target amplicons were generated by RT-PCR. cDNA first-strand synthesis was performed using the SuperScript III First-Strand Synthesis kit (Invitrogen). Briefly, 200 ng of bisulfite-converted total RNA was reverse-transcribed using the tRNA^{Asp}-specific reverse primer listed below and subsequently amplified via PCR. PCR reactions were separated on agarose gels and purified using the QIAquick Gel Extraction Kit (Qiagen). Dual indices and sequencing adapters were attached to the amplicons by an indexing PCR using the Illumina Nextera XT Index Kit v2 Set B (Illumina) and the 2× KAPA HiFi Hot Start Ready Mix (KAPA Biosystems). Amplicons were purified using AMPure XP beads (Beckman Coulter Genomics) according to the 16S Metagenomic Sequencing Library Preparation Protocol (Illumina). The final library was quantified via the Qubit dsDNA BR Assay Kit (Life Technologies), pooled at equimolar ratios, and submitted for MiSeq at the Genomic and Proteomic Core Facility (German Cancer Research Center). Sequencing data processing and methylation heatmap generation were performed using the BisAMP pipeline.[92] The forward primer used for library preparation was 5′-tcg tcg gca gcg tca gat gtg tat aag aga cag TGT TAG TAT AGT GGT GAG TAT-3′ and the reverse primer 5′-gtc tcg tgg gct cgg aga tgt gta taa gag aca gCT CCC CAT CAA AAA ATT A-3′. Capital letters correspond to target-specific sequences that were designed to in silico converted (C > U), unmethylated regions of tRNA^{Asp}. Small letters represent Illumina overhang adapter sequences that are compatible with Illumina index and sequencing adapters.

Statistics. P-values were determined with the Rstatix package[93] in RStudio[90] using unpaired two-tailed Welsh t tests and adjusted to reduce the number of false positive hits appearing due to performing multiple comparisons. Adjustment of p-values was conducted using the Benjamini−Hochberg method, a method based on the false discovery rate (FDR), reducing false positive hits while minimizing adjustment-based false negative results.[94,95]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.2c00388.

> 3D-RISM results for hydration sites, synthesis protocols and analytical data for all compounds, MST traces, ITC curves, NMR spectra, mass spectra, and LC-MS chromatograms of all tested compounds (PDF)
>
> Molecular docking results and predicted binding modes (ZIP)
>
> Molecular formula strings (CSV)

■ AUTHOR INFORMATION

Corresponding Authors

Mark Helm − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ⊙ orcid.org/0000-0002-0154-0928; Phone: ++49 (0) 6131 39 25731; Email: mhelm@uni-mainz.de

Tanja Schirmeister − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* Phone: ++49 (0) 6131 39 25742; Email: schirmei@uni-mainz.de

Authors

Marvin Schwickert − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ● orcid.org/0000-0002-1385-1416

Tim R. Fischer − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany*

Robert A. Zimmermann − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany*

Sabrina N. Hoba − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany*

J. Laurenz Meidner − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany*

Marlies Weber − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany*

Moritz Weber − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany*

Martin M. Stark − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany*

Jonas Koch − *Division of Epigenetics, DKFZ-ZMBH Alliance, German Cancer Research Center, D-69120 Heidelberg, Germany;* ● orcid.org/0000-0002-2080-6293

Nathalie Jung − *Institute of Pharmaceutical Technology, Goethe University Frankfurt, D-60438 Frankfurt am Main, Germany*

Christian Kersten − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ● orcid.org/0000-0001-9976-7639

Maike Windbergs − *Institute of Pharmaceutical Technology, Goethe University Frankfurt, D-60438 Frankfurt am Main, Germany*

Frank Lyko − *Division of Epigenetics, DKFZ-ZMBH Alliance, German Cancer Research Center, D-69120 Heidelberg, Germany*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jmedchem.2c00388

■ ABBREVIATIONS USED

5-azaC, 5-azacytidine; C, cytidine; CC_{50}, half maximal cytotoxic concentration; CDI, 1,1′-carbonyldiimidazole; COMT, catechol-*O*-MTase; Dab, 2,4-diaminobutyric acid; 1,2-DCE, 1,2-dichloroethane; DCM, dichloromethane; DIAD, diisopropyl azodicarboxylate; DIBAL, diisobutylaluminum hydride; DIPEA, *N*,*N*-diisopropylethylamine; DMAP, 4-dimethylaminopyridine; DMEM, Dulbecco's modified Eagle medium; DMF, *N*,*N*-dimethylformamide; DMSO, dimethyl sulfoxide; DNMT, DNA methyltransferase; DOT1L, histone−lysine *N*-methyltransferase, H3 lysine-79 specific; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HBTU, hexafluorophosphate benzotriazole tetramethyl uronium; HMT, histone methyltransferase; IEX, ion exchange chromatography; ITC, isothermal titration calorimetry; m^5C, 5-methylcytidine; m^2G, 2-methylguanosine; METTL, N^6-adenosine MTase; MMFF94x, Merck molecular force field; MTase, methyltransferase; MST, microscale thermophoresis; NTMT, N-terminal methyltransferase; PCR, polymerase chain reaction; PDB, Protein Data Bank; PMT, protein methyltransferase; 3D-RISM, three-dimensional reference interaction site model; RMSD, root-mean-square deviation of atomic positions; SAH, *S*-adenosyl-ʟ-homocysteine; SAM, *S*-adenosyl-ʟ-methionine; SAR, structure−activity relationship; SARS-CoV-2, severe acute respiratory syndrome coronavirus type 2; SFG, sinefungin; sncRNA, small noncoding RNA; TBTU, (1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate; TCA, trichloroacetic acid; THF, tetrahydrofuran; Tris, tris(hydroxymethyl)aminomethane

■ REFERENCES

(1) Ganesan, A.; Arimondo, P. B.; Rots, M. G.; Jeronimo, C.; Berdasco, M. The Timeline of Epigenetic Drug Discovery: From Reality to Dreams. *Clin. Epigenetics* **2019**, *11* (1), 174.

(2) Jonkhout, N.; Tran, J.; Smith, M. A.; Schonrock, N.; Mattick, J. S.; Novoa, E. M. The RNA Modification Landscape in Human Disease. *RNA* **2017**, *23* (12), 1754−1769.

(3) Chen, Q.; Yan, W.; Duan, E. Epigenetic Inheritance of Acquired Traits through Sperm RNAs and Sperm RNA Modifications. *Nat. Rev. Genet.* **2016**, *17* (12), 733−743.

(4) Chen, Q.; Yan, M.; Cao, Z.; Li, X.; Zhang, Y.; Shi, J.; Feng, G.; Peng, H.; Zhang, X.; Zhang, Y.; Qian, J.; Duan, E.; Zhai, Q.; Zhou, Q. Sperm TsRNAs Contribute to Intergenerational Inheritance of an Acquired Metabolic Disorder. *Science* **2016**, *351* (6271), 397−400.

(5) Rassoulzadegan, M.; Grandjean, V.; Gounon, P.; Vincent, S.; Gillot, I.; Cuzin, F. RNA-Mediated Non-Mendelian Inheritance of an Epigenetic Change in the Mouse. *Nature* **2006**, *441* (7092), 469−474.

(6) Liebers, R.; Rassoulzadegan, M.; Lyko, F. Epigenetic Regulation by Heritable RNA. *PLoS Genet.* **2014**, *10* (4), No. e1004296.

(7) Wang, Y.; Chen, Z.-P.; Hu, H.; Lei, J.; Zhou, Z.; Yao, B.; Chen, L.; Liang, G.; Zhan, S.; Zhu, X.; Jin, F.; Ma, R.; Zhang, J.; Liang, H.; Xing, M.; Chen, X.-R.; Zhang, C.-Y.; Zhu, J.-N.; Chen, X. Sperm MicroRNAs Confer Depression Susceptibility to Offspring. *Sci. Adv.* **2021**, *7* (7), eabd7605.

(8) Gapp, K.; Jawaid, A.; Sarkies, P.; Bohacek, J.; Pelczar, P.; Prados, J.; Farinelli, L.; Miska, E.; Mansuy, I. M. Implication of Sperm RNAs in Transgenerational Inheritance of the Effects of Early Trauma in Mice. *Nat. Neurosci.* **2014**, *17* (5), 667−669.

(9) Rodgers, A. B.; Morgan, C. P.; Leu, N. A.; Bale, T. L. Transgenerational Epigenetic Programming via Sperm MicroRNA Recapitulates Effects of Paternal Stress. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (44), 13699−13704.

(10) Gapp, K.; van Steenwyk, G.; Germain, P. L.; Matsushima, W.; Rudolph, K. L. M.; Manuella, F.; Roszkowski, M.; Vernaz, G.; Ghosh, T.; Pelczar, P.; Mansuy, I. M.; Miska, E. A. Alterations in Sperm Long RNA Contribute to the Epigenetic Inheritance of the Effects of Postnatal Trauma. *Mol. Psychiatry* **2020**, *25* (9), 2162−2174.

(11) Zhang, Y.; Zhang, X.; Shi, J.; Tuorto, F.; Li, X.; Liu, Y.; Liebers, R.; Zhang, L.; Qu, Y.; Qian, J.; Pahima, M.; Liu, Y.; Yan, M.; Cao, Z.; Lei, X.; Cao, Y.; Peng, H.; Liu, S.; Wang, Y.; Zheng, H.; Woolsey, R.; Quilici, D.; Zhai, Q.; Li, L.; Zhou, T.; Yan, W.; Lyko, F.; Zhang, Y.; Zhou, Q.; Duan, E.; Chen, Q. Dnmt2 Mediates Intergenerational Transmission of Paternally Acquired Metabolic Disorders through Sperm Small Non-Coding RNAs. *Nat. Cell Biol.* **2018**, *20* (5), 535−540.

(12) Sharma, U.; Conine, C. C.; Shea, J. M.; Boskovic, A.; Derr, A. G.; Bing, X. Y.; Belleannee, C.; Kucukural, A.; Serra, R. W.; Sun, F.; Song, L.; Carone, B. R.; Ricci, E. P.; Li, X. Z.; Fauquier, L.; Moore, M. J.; Sullivan, R.; Mello, C. C.; Garber, M.; Rando, O. J. Biogenesis and Function of TRNA Fragments during Sperm Maturation and Fertilization in Mammals. *Science* **2016**, *351* (6271), 391−396.

(13) Sarker, G.; Sun, W.; Rosenkranz, D.; Pelczar, P.; Opitz, L.; Efthymiou, V.; Wolfrum, C.; Peleg-Raibstein, D. Maternal Overnutrition Programs Hedonic and Metabolic Phenotypes across Generations through Sperm TsRNAs. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (21), 10547−10556.

(14) Swanson, G. M.; Estill, M.; Diamond, M. P.; Legro, R. S.; Coutifaris, C.; Barnhart, K. T.; Huang, H.; Hansen, K. R.; Trussell, J. C.; Coward, R. M.; Zhang, H.; Goodrich, R.; Krawetz, S. A. Human Chromatin Remodeler Cofactor, RNA Interactor, Eraser and Writer Sperm RNAs Responding to Obesity. *Epigenetics* **2020**, *15* (1−2), 32−46.

(15) Zhang, Y.; Shi, J.; Rassoulzadegan, M.; Tuorto, F.; Chen, Q. Sperm RNA Code Programmes the Metabolic Health of Offspring. *Nat. Rev. Endocrinol.* **2019**, *15* (8), 489−498.

(16) Kiani, J.; Grandjean, V.; Liebers, R.; Tuorto, F.; Ghanbarian, H.; Lyko, F.; Cuzin, F.; Rassoulzadegan, M. RNA-Mediated Epigenetic Heredity Requires the Cytosine Methyltransferase Dnmt2. *PLoS Genet.* **2013**, *9* (5), No. e1003498.

(17) Yoder, J. A.; Bestor, T. H. A Candidate Mammalian DNA Methyltransferase Related to Pmt1p of Fission Yeast. *Hum. Mol. Genet.* **1998**, *7* (2), 279−284.

(18) Lyko, F. The DNA Methyltransferase Family: A Versatile Toolkit for Epigenetic Regulation. *Nat. Rev. Genet.* **2018**, *19* (2), 81−92.

(19) Goll, M. G.; Kirpekar, F.; Maggert, K. A.; Yoder, J. A.; Hsieh, C. L.; Zhang, X.; Golic, K. G.; Jacobsen, S. E.; Bestor, T. H. Methylation of TRNAAsp by the DNA Methyltransferase Homolog Dnmt2. *Science (80-.).* **2006**, *311* (5759), 395−398.

(20) Hermann, A.; Schmitt, S.; Jeltsch, A. The Human Dnmt2 Has Residual DNA-(Cytosine-C5) Methyltransferase Activity. *J. Biol. Chem.* **2003**, *278* (34), 31717−31721.

(21) Jeltsch, A.; Ehrenhofer-Murray, A.; Jurkowski, T. P.; Lyko, F.; Reuter, G.; Ankri, S.; Nellen, W.; Schaefer, M.; Helm, M. Mechanism and Biological Role of Dnmt2 in Nucleic Acid Methylation. *RNA Biol.* **2017**, *14* (9), 1108−1123.

(22) Rai, K.; Chidester, S.; Zavala, C. V.; Manos, E. J.; James, S. R.; Karpf, A. R.; Jones, D. A.; Cairns, B. R. Dnmt2 Functions in the Cytoplasm to Promote Liver, Brain, and Retina Development in Zebrafish. *Genes Dev.* **2007**, *21*, 261−266.

(23) Lin, M. J.; Tang, L. Y.; Reddy, M. N.; Shen, C. K. J. DNA Methyltransferase Gene DDnmt2 and Longevity of Drosophila. *J. Biol. Chem.* **2005**, *280* (2), 861−864.

(24) Schaefer, M.; Pollex, T.; Hanna, K.; Tuorto, F.; Meusburger, M.; Helm, M.; Lyko, F. RNA Methylation by Dnmt2 Protects Transfer RNAs against Stress-Induced Cleavage. *Genes Dev.* **2010**, *24* (15), 1590−1595.

(25) Tuorto, F.; Liebers, R.; Musch, T.; Schaefer, M.; Hofmann, S.; Kellner, S.; Frye, M.; Helm, M.; Stoecklin, G.; Lyko, F. RNA Cytosine Methylation by Dnmt2 and NSun2 Promotes TRNA Stability and Protein Synthesis. *Nat. Struct. Mol. Biol.* **2012**, *19*, 900.

(26) Tuorto, F.; Herbst, F.; Alerasool, N.; Bender, S.; Popp, O.; Federico, G.; Reitter, S.; Liebers, R.; Stoecklin, G.; Gröne, H.; Dittmar, G.; Glimm, H.; Lyko, F. The TRNA Methyltransferase Dnmt2 Is Required for Accurate Polypeptide Synthesis during Haematopoiesis. *EMBO J.* **2015**, *34* (18), 2350−2362.

(27) Li, L.; Wang, S. DNA Methyltransferase (DNMTs) Expression in Cervical Cancer Tissues and Its Relationship with HPV Infection and Tumor Malignancy. *J. Hainan Med. Univ.* **2017**, *23*, 136−139.

(28) Towns, W. L.; Begley, T. J. Transfer RNA Methyltransferases and Their Corresponding Modifications in Budding Yeast and Humans: Activities, Predications, and Potential Roles in Human Health. *DNA and Cell Biology.* **2012**, *31*, 434−454.

(29) Forbes, S. A.; Beare, D.; Gunasekaran, P.; Leung, K.; Bindal, N.; Boutselakis, H.; Ding, M.; Bamford, S.; Cole, C.; Ward, S.; Kok, C. Y.; Jia, M.; De, T.; Teague, J. W.; Stratton, M. R.; McDermott, U.; Campbell, P. J. COSMIC: Exploring the World's Knowledge of Somatic Mutations in Human Cancer. *Nucleic Acids Res.* **2015**, *43* (D1), D805−D811.

(30) Dong, A. Structure of Human DNMT2, an Enigmatic DNA Methyltransferase Homolog That Displays Denaturant-Resistant Binding to DNA. *Nucleic Acids Res.* **2001**, *29* (2), 439−448.

(31) Jurkowski, T. P.; Jeltsch, A. On the Evolutionary Origin of Eukaryotic DNA Methyltransferases and Dnmt2. *PLoS One* **2011**, *6* (11), e28104.

(32) Schulz, E. C.; Roth, H. M.; Ankri, S.; Ficner, R. Structure Analysis of *Entamoeba histolytica* DNMT2 (EhMeth). *PLoS One* **2012**, *7* (6), e38728.

(33) Tran, H. T. T.; Kim, H. N.; Lee, I.-K.; Kim, Y.-K.; Ahn, J.-S.; Yang, D.-H.; Lee, J.-J.; Kim, H.-J. DNA Methylation Changes Following 5-Azacitidine Treatment in Patients with Myelodysplastic Syndrome. *J. Korean Med. Sci.* **2011**, *26* (2), 207−213.

(34) Mund, C.; Hackanson, B.; Stresemann, C.; Lübbert, M.; Lyko, F. Characterization of DNA Demethylation Effects Induced by 5-Aza-2'-Deoxycytidine in Patients with Myelodysplastic Syndrome. *Cancer Res.* **2005**, *65* (16), 7086−7090.

(35) Yang, A. S.; Doshi, K. D.; Choi, S.-W.; Mason, J. B.; Mannari, R. K.; Gharybian, V.; Luna, R.; Rashid, A.; Shen, L.; Estecio, M. R. H.; Kantarjian, H. M.; Garcia-Manero, G.; Issa, J.-P. J. DNA Methylation Changes after 5-Aza-2'-Deoxycytidine Therapy in Patients with Leukemia. *Cancer Res.* **2006**, *66* (10), 5495−5503.

(36) Santi, D. V.; Norment, A.; Garrett, C. E. Covalent Bond Formation between a DNA-Cytosine Methyltransferase and DNA Containing 5-Azacytosine. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81* (22), 6993−6997.

(37) Khoddami, V.; Cairns, B. R. Identification of Direct Targets and Modified Bases of RNA Cytosine Methyltransferases. *Nat. Biotechnol.* **2013**, *31* (5), 458−464.

(38) Zhou, L.; Cheng, X.; Connolly, B. A.; Dickman, M. J.; Hurd, P. J.; Hornby, D. P. Zebularine: A Novel DNA Methylation Inhibitor That Forms a Covalent Complex with DNA Methyltransferases. *J. Mol. Biol.* **2002**, *321* (4), 591−599.

(39) Cho, H. D.; Oyelere, A. K.; Strobel, S. A.; Weiner, A. M. Use of Nucleotide Analogs by Class I and Class II CCA-Adding Enzymes (TRNA Nucleotidyltransferase): Deciphering the Basis for Nucleotide Selection. *RNA* **2003**, *9* (8), 970−981.

(40) Hamill, R. L.; Hoehn, M. M. A9145, a New Adenine-Containing Antifungal Antibiotic: I. Discovery and Isolation. *J. Antibiot. (Tokyo).* **1973**, *26* (8), 463−465.

(41) Borchardt, R. T.; Eiden, L. E.; Wu, B. S.; Rutledge, C. O. Sinefungin, a Potent Inhibitor of S-Adenosylmethionine: Protein O-Methyltransferase. *Biochem. Biophys. Res. Commun.* **1979**, *89* (3), 919−924.

(42) Richon, V. M.; Johnston, D.; Sneeringer, C. J.; Jin, L.; Majer, C. R.; Elliston, K.; Jerva, L. F.; Scott, M. P.; Copeland, R. A. Chemogenetic Analysis of Human Protein Methyltransferases. *Chem. Biol. Drug Des.* **2011**, *78* (2), 199−210.

(43) Lerner, C.; Masjost, B.; Ruf, A.; Gramlich, V.; Jakob-Roetne, R.; Zürcher, G.; Borroni, E.; Diederich, F. Bisubstrate Inhibitors for the Enzyme Catechol-O-Methyltransferase (COMT): Influence of Inhibitor Preorganisation and Linker Length between the Two Substrate Moieties on Binding Affinity. *Org. Biomol. Chem.* **2003**, *1* (1), 42−49.

(44) Paulini, R.; Lerner, C.; Jakob-Roetne, R.; Zürcher, G.; Borroni, E.; Diederich, F. Bisubstrate Inhibitors of the Enzyme Catechol O-Methyltransferase (COMT): Efficient Inhibition despite the Lack of a Nitro Group. *Chembiochem* **2004**, *5* (9), 1270−1274.

(45) Ellermann, M.; Paulini, R.; Jakob-Roetne, R.; Lerner, C.; Borroni, E.; Roth, D.; Ehler, A.; Schweizer, W. B.; Schlatter, D.; Rudolph, M. G.; Diederich, F. Molecular Recognition at the Active Site of Catechol-O-Methyltransferase (COMT): Adenine Replacements in Bisubstrate Inhibitors. *Chemistry* **2011**, *17* (23), 6369−6381.

(46) Anglin, J. L.; Deng, L.; Yao, Y.; Cai, G.; Liu, Z.; Jiang, H.; Cheng, G.; Chen, P.; Dong, S.; Song, Y. Synthesis and Structure-Activity Relationship Investigation of Adenosine-Containing Inhibitors of Histone Methyltransferase DOT1L. *J. Med. Chem.* **2012**, *55* (18), 8066−8074.

(47) Bedi, R. K.; Huang, D.; Eberle, S. A.; Wiedmer, L.; Śledź, P.; Caflisch, A. Small-Molecule Inhibitors of METTL3, the Major Human Epitranscriptomic Writer. *ChemMedChem* **2020**, *15* (9), 744−748.

(48) Ferron, F.; Decroly, E.; Selisko, B.; Canard, B. The Viral RNA Capping Machinery as a Target for Antiviral Drugs. *Antiviral Res.* **2012**, *96* (1), 21−31.

(49) Ahmed-Belkacem, R.; Sutto-Ortiz, P.; Guiraud, M.; Canard, B.; Vasseur, J.-J.; Decroly, E.; Debart, F. Synthesis of Adenine Dinucleosides SAM Analogs as Specific Inhibitors of SARS-CoV Nsp14 RNA Cap Guanine-N7-Methyltransferase. *Eur. J. Med. Chem.* **2020**, *201*, 112557.

(50) Saavedra, O. M.; Isakovic, L.; Llewellyn, D. B.; Zhan, L.; Bernstein, N.; Claridge, S.; Raeppel, F.; Vaisburg, A.; Elowe, N.; Petschner, A. J.; Rahil, J.; Beaulieu, N.; MacLeod, A. R.; Delorme, D.; Besterman, J. M.; Wahhab, A. SAR around (l)-S-Adenosyl-l-Homocysteine, an Inhibitor of Human DNA Methyltransferase (DNMT) Enzymes. *Bioorg. Med. Chem. Lett.* **2009**, *19* (10), 2747−2751.

(51) Isakovic, L.; Saavedra, O. M.; Llewellyn, D. B.; Claridge, S.; Zhan, L.; Bernstein, N.; Vaisburg, A.; Elowe, N.; Petschner, A. J.; Rahil, J.; Beaulieu, N.; Gauthier, F.; MacLeod, A. R.; Delorme, D.; Besterman, J. M.; Wahhab, A. Constrained (l-)-S-Adenosyl-l-Homocysteine (SAH) Analogues as DNA Methyltransferase Inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19* (10), 2742−2746.

(52) Stein, E. M.; Garcia-Manero, G.; Rizzieri, D. A.; Tibes, R.; Berdeja, J. G.; Savona, M. R.; Jongen-Lavrenic, M.; Altman, J. K.; Thomson, B.; Blakemore, S. J.; Daigle, S. R.; Waters, N. J.; Suttle, A. B.; Clawson, A.; Pollock, R.; Krivtsov, A.; Armstrong, S. A.; DiMartino, J.; Hedrick, E.; Löwenberg, B.; Tallman, M. S. The DOT1L Inhibitor Pinometostat Reduces H3K79 Methylation and Has Modest Clinical Activity in Adult Acute Leukemia. *Blood* **2018**, *131* (24), 2661−2669.

(53) Halby, L.; Marechal, N.; Pechalrieu, D.; Cura, V.; Franchini, D.-M.; Faux, C.; Alby, F.; Troffer-Charlier, N.; Kudithipudi, S.; Jeltsch, A.; Aouadi, W.; Decroly, E.; Guillemot, J.-C.; Page, P.; Ferroud, C.; Bonnefond, L.; Guianvarc'h, D.; Cavarelli, J.; Arimondo, P. B. Hijacking DNA Methyltransferase Transition State Analogues to Produce Chemical Scaffolds for PRMT Inhibitors. *Philos. Trans. R. Soc. London. Ser. B, Biol. Sci.* **2018**, *373* (1748), 20170072.

(54) Zhang, J. H.; Chung, T. D.; Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **1999**, *4* (2), 67−73.

(55) Kalliokoski, T.; Kramer, C.; Vulpetti, A.; Gedeck, P. Comparability of Mixed IC$_{50}$ Data - a Statistical Analysis. *PLoS One* **2013**, *8* (4), No. e61007.

(56) Chellamuthu, A.; Gray, S. G. The RNA Methyltransferase NSUN2 and Its Potential Roles in Cancer. *Cells* **2020**, *9* (8), 1758.

(57) Trixl, L.; Lusser, A. The Dynamic RNA Modification 5-Methylcytosine and Its Emerging Role as an Epitranscriptomic Mark. *Wiley Interdiscip. Rev. RNA* **2019**, *10* (1), No. e1510.

(58) Xue, C.; Zhao, Y.; Li, L. Advances in RNA Cytosine-5 Methylation: Detection, Regulatory Mechanisms, Biological Functions and Links to Cancer. *Biomark. Res.* **2020**, *8*, 43.

(59) Shankar, S. R.; Bahirvani, A. G.; Rao, V. K.; Bharathy, N.; Ow, J. R.; Taneja, R. G9a, a Multipotent Regulator of Gene Expression. *Epigenetics* **2013**, *8* (1), 16−22.

(60) Yokochi, T.; Robertson, K. D. Dimethyl Sulfoxide Stimulates the Catalytic Activity of de Novo DNA Methyltransferase 3a (Dnmt3a) in Vitro. *Bioorg. Chem.* **2004**, *32* (4), 234−243.

(61) Gros, C.; Chauvigné, L.; Poulet, A.; Menon, Y.; Ausseil, F.; Dufau, I.; Arimondo, P. B. Development of a Universal Radioactive DNA Methyltransferase Inhibition Test for High-Throughput Screening and Mechanistic Studies. *Nucleic Acids Res.* **2013**, *41* (19), No. e185.

(62) Liu, Q.; Cai, X.; Yang, D.; Chen, Y.; Wang, Y.; Shao, L.; Wang, M. W. Cycloalkane Analogues of Sinefungin as EHMT1/2 Inhibitors. *Bioorg. Med. Chem.* **2017**, *25* (17), 4579−4594.

(63) Rotili, D.; Tarantino, D.; Marrocco, B.; Gros, C.; Masson, V.; Poughon, V.; Ausseil, F.; Chang, Y.; Labella, D.; Cosconati, S.; Di Maro, S.; Novellino, E.; Schnekenburger, M.; Grandjenette, C.; Bouvy, C.; Diederich, M.; Cheng, X.; Arimondo, P. B.; Mai, A. Properly Substituted Analogues of BIX-01294 Lose Inhibition of G9a Histone Methyltransferase and Gain Selective Anti-DNA Methyltransferase 3A Activity. *PLoS One* **2014**, *9* (5), No. e96941.

(64) Devkota, K.; Lohse, B.; Liu, Q.; Wang, M.-W.; Stærk, D.; Berthelsen, J.; Clausen, R. P. Analogues of the Natural Product Sinefungin as Inhibitors of EHMT1 and EHMT2. *ACS Med. Chem. Lett.* **2014**, *5* (4), 293−297.

(65) Zheng, W.; Ibáñez, G.; Wu, H.; Blum, G.; Zeng, H.; Dong, A.; Li, F.; Hajian, T.; Allali-Hassani, A.; Amaya, M. F.; Siarheyeva, A.; Yu, W.; Brown, P. J.; Schapira, M.; Vedadi, M.; Min, J.; Luo, M. Sinefungin Derivatives as Inhibitors and Structure Probes of Protein Lysine Methyltransferase SETD2. *J. Am. Chem. Soc.* **2012**, *134* (43), 18004−18014.

(66) Kilgore, J. A.; Du, X.; Melito, L.; Wei, S.; Wang, C.; Chin, H. G.; Posner, B.; Pradhan, S.; Ready, J. M.; Williams, N. S. Identification of DNMT1 Selective Antagonists Using a Novel Scintillation Proximity Assay. *J. Biol. Chem.* **2013**, *288* (27), 19673−19684.

(67) Thüring, K.; Schmid, K.; Keller, P.; Helm, M. Analysis of RNA Modifications by Liquid Chromatography-Tandem Mass Spectrometry. *Methods* **2016**, *107*, 48−56.

(68) Schaefer, M.; Pollex, T.; Hanna, K.; Lyko, F. RNA Cytosine Methylation Analysis by Bisulfite Sequencing. *Nucleic Acids Res.* **2008**, *37* (2), No. e12.

(69) Rautio, J.; Meanwell, N. A.; Di, L.; Hageman, M. J. The Expanding Role of Prodrugs in Contemporary Drug Design and Development. *Nat. Rev. Drug Discovery* **2018**, *17* (8), 559−587.

(70) Biela, A.; Nasief, N. N.; Betz, M.; Heine, A.; Hangauer, D.; Klebe, G. Dissecting the Hydrophobic Effect on the Molecular Level: The Role of Water, Enthalpy, and Entropy in Ligand Binding to Thermolysin. *Angew. Chem., Int. Ed. Engl.* **2013**, *52* (6), 1822−1828.

(71) *The PyMOL Molecular Graphics System*; Schrödinger, LLC.

(72) Liu, R.-J.; Long, T.; Li, J.; Li, H.; Wang, E.-D. Structural Basis for Substrate Binding and Catalytic Mechanism of a Human RNA:M5C Methyltransferase NSun6. *Nucleic Acids Res.* **2017**, *45* (11), 6684−6697.

(73) 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.

(74) 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.

(75) Pignot, M.; Pljevaljcic, G.; Weinhold, E. Efficient Synthesis of S-Adenosyl-L-Homocysteine Natural Product Analogues and Their Use to Elucidate the Structural Determinant for Cofactor Binding of the

DNA Methyltransferase M·HhaI. *Eur. J. Org. Chem.* **2000**, *2000* (3), 549−555.

(76) Zhang, G.; Richardson, S. L.; Mao, Y.; Huang, R. Design, Synthesis, and Kinetic Analysis of Potent Protein N-Terminal Methyltransferase 1 Inhibitors. *Org. Biomol. Chem.* **2015**, *13* (14), 4149−4154.

(77) Liu, Q.; Cai, X.; Yang, D.; Chen, Y.; Wang, Y.; Shao, L.; Wang, M.-W. Cycloalkane Analogues of Sinefungin as EHMT1/2 Inhibitors. *Bioorg. Med. Chem.* **2017**, *25* (17), 4579−4594.

(78) Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the FlexX Incremental Construction Algorithm for Protein- Ligand Docking. *Proteins Struct. Funct. Genet.* **1999**, *37* (2), 228−241.

(79) *LeadIT*; BioSolveIT GmbH: Sankt Augustin, Germany, 2017.

(80) 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* (1), 235−242.

(81) Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M. Protoss: A Holistic Approach to Predict Tautomers and Protonation States in Protein-Ligand Complexes. *J. Cheminform.* **2014**, *6* (1), 12.

(82) 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.

(83) *Molecular Operating Environment (MOE)*; Chemical Computing Group, 2019.

(84) Beglov, D.; Roux, B. An Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water. *J. Phys. Chem. B* **1997**, *101* (39), 7821−7826.

(85) Siddique, A. N.; Nunna, S.; Rajavelu, A.; Zhang, Y.; Jurkowska, R. Z.; Reinhardt, R.; Rots, M. G.; Ragozin, S.; Jurkowski, T. P.; Jeltsch, A. Targeted Methylation and Gene Silencing of VEGF-A in Human Cells by Using a Designed Dnmt3a-Dnmt3L Single-Chain Fusion Protein with Increased DNA Methylation Activity. *J. Mol. Biol.* **2013**, *425* (3), 479−491.

(86) Zhang, X.; Liu, Z.; Yi, J.; Tang, H.; Xing, J.; Yu, M.; Tong, T.; Shang, Y.; Gorospe, M.; Wang, W. The TRNA Methyltransferase NSun2 Stabilizes P16INK^4 MRNA by Methylating the 3′-Untranslated Region of P16. *Nat. Commun.* **2012**, *3*, 712.

(87) Long, T.; Li, J.; Li, H.; Zhou, M.; Zhou, X.-L.; Liu, R.-J.; Wang, E.-D. Sequence-Specific and Shape-Selective RNA Recognition by the Human RNA 5-Methylcytosine Methyltransferase NSun6. *J. Biol. Chem.* **2016**, *291* (46), 24293−24303.

(88) Goldberg, R. N.; Kishore, N.; Lennen, R. M. Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data* **2002**, *31* (2), 231−370.

(89) Ritz, C.; Baty, F.; Streibig, J. C.; Gerhard, D. Dose-Response Analysis Using R. *PLoS One* **2015**, *10* (12), No. e0146021.

(90) *RStudio*; RStudio Team: Boston, MA, 2020.

(91) 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.

(92) Bormann, F.; Tuorto, F.; Cirzi, C.; Lyko, F.; Legrand, C. BisAMP: A Web-Based Pipeline for Targeted RNA Cytosine-5 Methylation Analysis. *Methods* **2019**, *156*, 121−127.

(93) Kassambara, A. *Rstatix: Pipe-Friendly Framework for Basic Statistical Tests*; 2021.

(94) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, *57* (1), 289−300.

(95) Jafari, M.; Ansari-Pour, N. Why, When and How to Adjust Your P Values? *Cell J.* **2019**, *20* (4), 604−607.

3.5 Electron-deficient aromatic SAH derivatives as potential covalent DNMT2 inhibitors

Note: The substance numbers from the corresponding manuscript are listed *in italics* behind the substance numbers in this thesis.

3.5.1 Summary and own contribution

Apart from effective alkyne-base inhibitors, the study described in **Section 3.4** yielded the *N*-benzyl derivative **69q** with moderate inhibitory activity against DNMT2 ($56.8 \pm 6.8\%$ at 100 μM). Compared to alkynes, the phenyl moiety allows various possible modifications to create a suitable substitution pattern for a cytidine-site targeting substructure. Based on the *N*-benzyl derivative **69q**, a structure-activity relationship study was performed using the TOPLISS scheme (**70a–n** (*32–45*)). This is an operational scheme in drug design for the derivatization of aromatic compounds considering hydrophobic (π), electronic (σ), and steric (E_s) values of different substituents.[284] Depending on the effect on activity – more active, equipotent, or less active – different modifications are proposed. The principle of the scheme was followed, i.e. if a modification increased the affinity, the proposed path of modification was followed and vice versa. Selected derivatives of non-proposed paths were also tested for comparison. The compounds were evaluated for their binding affinity to DNMT2 via MST by ██████████ (██████████ group). Their inhibitory activity was determined using a tritium incorporation assay conducted by ████████ (████ group) and ██████████ (██████████ group). By evaluating the inhibition data with substituent effects taken into account, a clear correlation with the TOPLISS scheme suggestions became apparent. Moreover, it could be observed that introducing electron-withdrawing groups increased the inhibition of DNMT2, as demonstrated by the following branch (**Figure 16**): 4-Cl (**70a** (*32*)) $\rightarrow$ 3,4-diCl (**70g** (*40*)) $\rightarrow$ 4-Cl-3-CF$_3$ (**70m** (*41*)) $\rightarrow$ 4-NO$_2$-3-CF$_3$ (**70n** (*45*)) with increasing inhibition of 72% $\rightarrow$ 74% $\rightarrow$ 86% $\rightarrow$ 94% at 100 μM, respectively. Out of this compound series, the 4-nitro-3-trifluoromethyl derivative **70n** (*45*) showed the highest inhibition with an IC$_{50}$ value of 2.5 ± 0.2 μM.
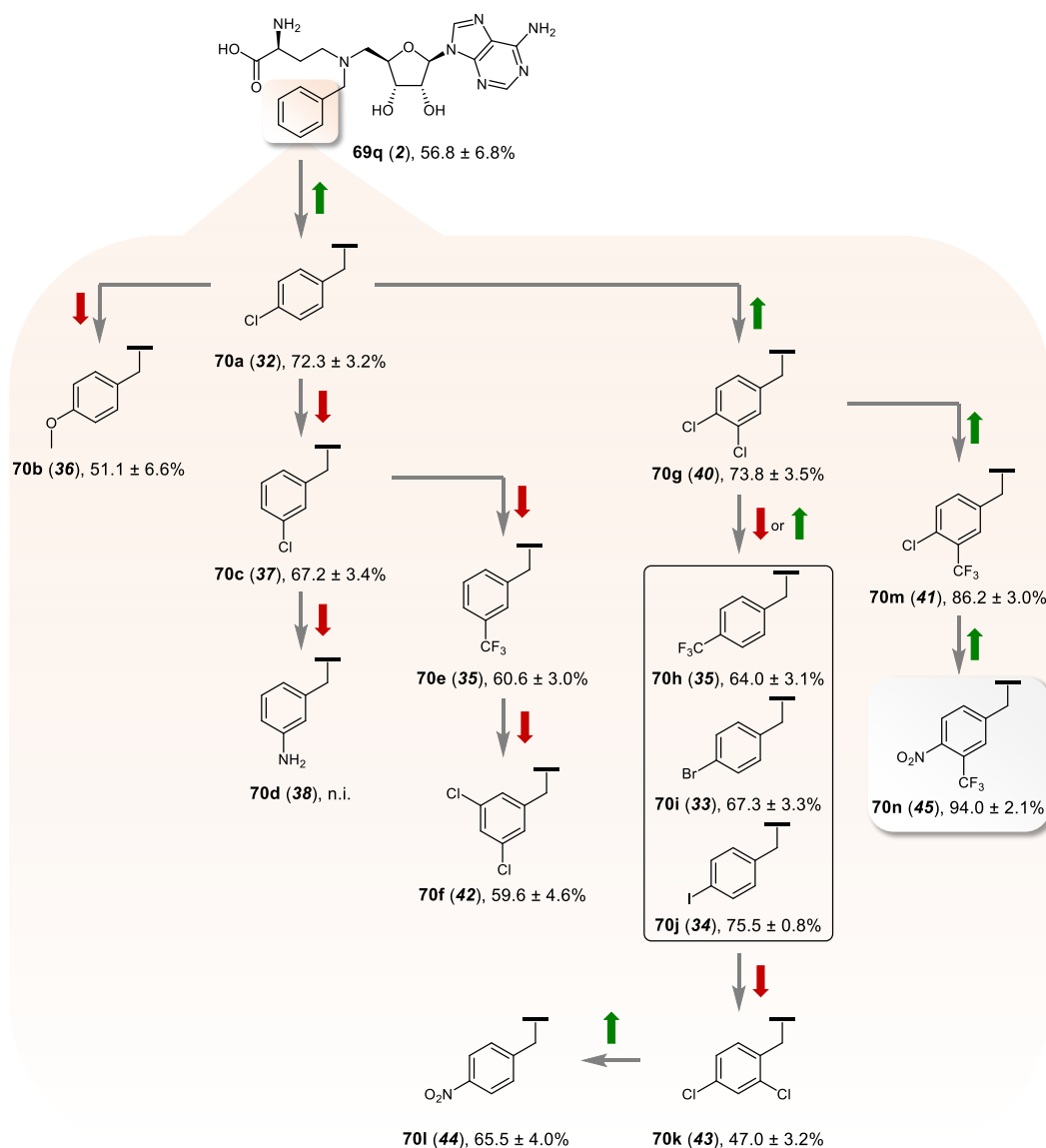
Figure 16: Synthesized compounds and their corresponding results with depiction of their positions in the TOPLISS scheme. Arrows indicate either increase (green) or decrease (red) in inhibition of DNMT2.

To further increase the electrophilicity of the aromatic ring by introducing additional −I and −M effects, the methylenamine substructure was replaced with a sulfonamide. Based on this modification, a subsequent structure-activity relationship study was carried out (**71a–o** (*72–84*, *90*, and *91*)). The aryl sulfonyl building blocks were decorated with strong electron-withdrawing groups such as NO$_2$ and CF$_3$ as well as intracyclic nitrogen atoms (**Figure 17**). Furthermore, halogen leaving groups were attached to the aromatic ring to enable a possible covalent reaction with the catalytically active cysteine in the cytidine site.
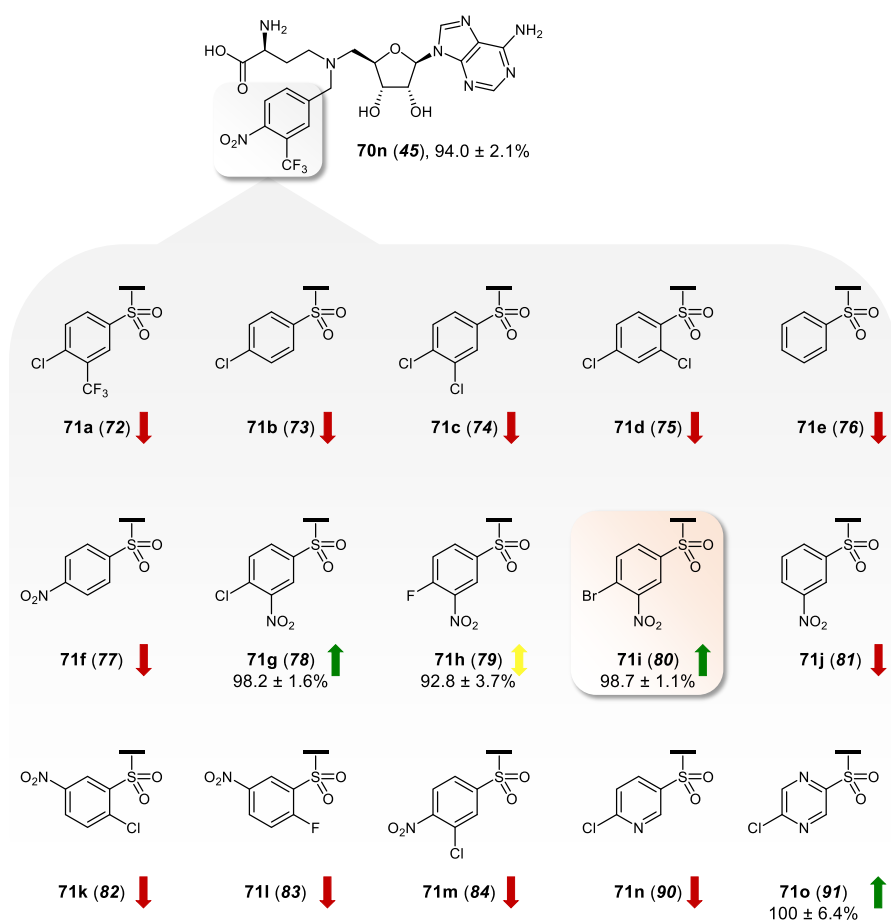
Figure 17: Synthesized sulfonamides and their corresponding results. Arrows indicate either increase (green), no change (yellow), or decrease (red) in inhibition of DNMT2.

The sulfonamide derivatives showed significantly lower inhibition compared to their benzyl analogs, e.g. **70a** (*32*) vs. **71b** (*73*) (4-Cl), **70g** (*40*) vs. **71c** (*74*) (3,4-diCl), and **70m** (*41*) vs. **71a** (*72*) (4-Cl-3-CF$_3$). Interestingly, the 4-Cl/Br-3-NO$_2$ substituted sulfonamide-based derivatives **71g** (*78*) and **71i** (*80*) as well as the 4-chloropyrazinyl derivative **71o** (*91*) exhibited the highest inhibitions (IC$_{50}$ < 2.5 μM). Protein mass spectrometry conducted by ▮▮▮▮▮▮▮ (▮▮▮▮▮ group, TU Darmstadt) revealed that the 4-halogen-3-NO$_2$-decorated phenylsulfonamide derivatives **71g–i** (*78–80*) and the chloropyrazinyl structure **71o** (*91*) reacted covalently with the catalytically active Cys79 in the cytidine site of DNM2 (**Scheme 20**).

For compounds **71h** (*79*) and **71o** (*91*) a second binding site was identified as they also bound to Cys287 on the protein surface. This observation was due to their high reactivity, which could further be proven with the fluoro-analog **71h** (*79*) that showed rapid conversion with the reducing agent dithiothreitol (DTT) in a reactivity test.
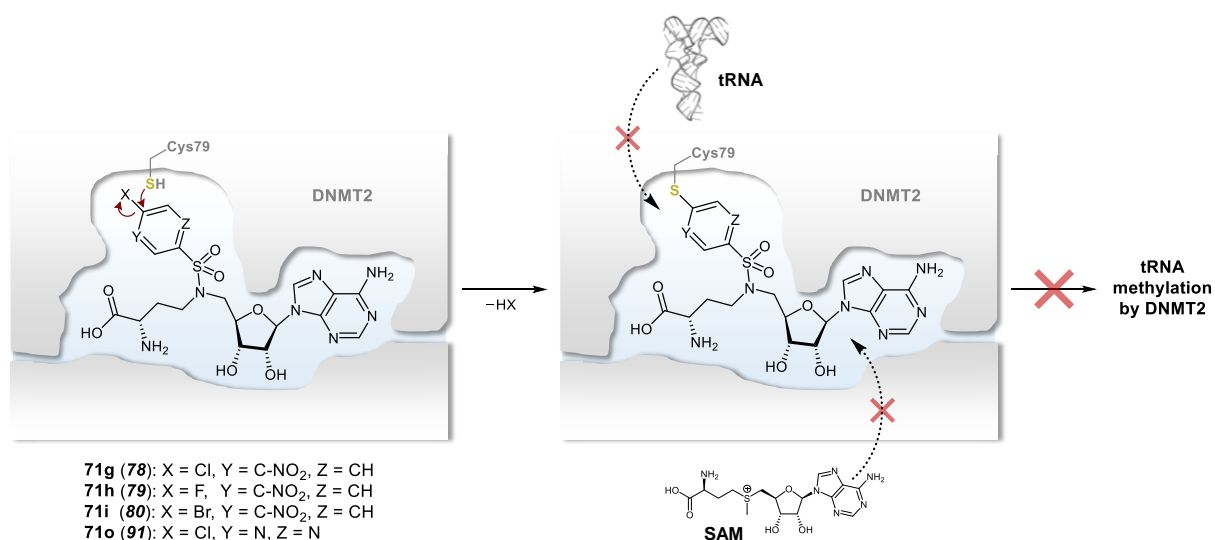
Scheme 20: Covalent reaction between inhibitors **71g–i** (*78–80*) and **71o** (*91*) with DNMT2 resulting in irreversible inhibition.

The most promising inhibitor out of this study turned out to be the 4-Br-3-NO$_2$-phenylsulfonamide derivative **71i** (*80*). With an IC$_{50}$ value of 1.2 ± 0.1 μM, compound **71i** (*80*) represented an improvement of one order of magnitude compared to the (*SR*)-*N*-but-3-yn-2-yl-based SAH derivative **69s-A/B** (**Section 3.4**) and outclassed the natural ligands SAH and SFG. Furthermore, the inhibitor featured high selectivity for DNMT2 compared with the RNA MTases NSUN2 and NSUN6.

For a more detailed discussion, as well as the presentation of the experimental procedures, the reader is referred to the corresponding manuscript (**Section 3.5.2**) and Supporting Information (**Appendix**). Experimental procedures conducted or supervised by MARVIN SCHWICKERT (████████ group) are described in **Section 3.5.3**.

Own contribution: Design and synthesis of inhibitors, supervising a master thesis (████████) about the synthesis of benzyl-based inhibitors **70a–k** (*32–43*) and **70m** (*41*), reactivity testing with DTT by LC-MS, writing the manuscript except for the section about protein mass spectrometry, creation of graphical abstract, figures, and schemes except for **Figures 4–6**.

3.5.2 Manuscript

Covalent *S*-Adenosylhomocysteine-based DNA Methyltransferase 2 Inhibitors with a New Type of Aryl Warhead

Marvin Schwickert,[‡,#] Robert A. Zimmermann,[‡,#] Tanja Habeck,[†] Sabrina N. Hoba,[#] Zarina Nidoieva,[#] Tim R. Fischer,[#] Martin M. Stark,[#] Christian Kersten,[#] Frederik Lermyte,[†] Mark Helm,[#*] and Tanja Schirmeister[#*]

[#]Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany

[†]Technical University of Darmstadt, Alarich-Weiss-Str. 4, D-64287 Darmstadt, Germany

*Corresponding authors:

Prof. Dr. Tanja Schirmeister, Phone: +49 6131 39-25742, E-Mail: schirmei@uni-mainz.de.

Prof. Dr. Mark Helm, Phone: +49 6131 39-25731, E-Mail: mhelm@uni-mainz.de

[‡]These authors contributed equally.

ABSTRACT: The DNA methyltransferase 2 (DNMT2) is an RNA modifying enzyme associated with pathophysiological processes, such as mental and metabolic disorders, or cancer. Although the development of methyltransferase inhibitors remains challenging, DNMT2 is not only a promising target for drug discovery, but also for the development of activity-based probes. Here, we present covalent SAH-based DNMT2 inhibitors decorated with a new type of aryl warhead. Based on a non-covalent DNMT2 inhibitor with *N*-benzyl substituent, the *Topliss* scheme was followed for optimization. The results showed that electron-deficient benzyl moieties highly increased affinity. By decorating the structures with strong electron-withdrawing moieties and leaving groups, we adjusted the electrophilicity to create covalent DNMT2 inhibitors. A 4-bromo-3-nitrophenylsulfonamide (80) decorated SAH derivative turned out to be the most potent ($IC_{50} = 1.2 \pm 0.1$ μM) and selective inhibitor. Protein mass spectrometry confirmed the covalent reaction with the catalytically active cysteine-79.
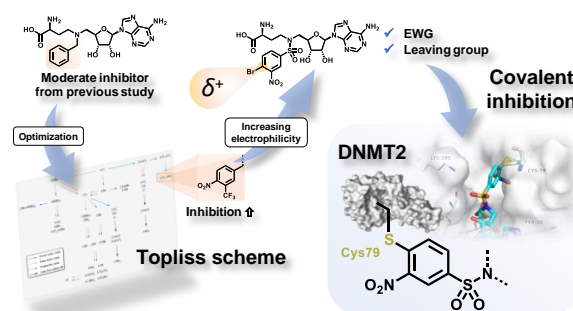
KEYWORDS: DNMT2, NSUN, covalent SAH-based inhibitors, aryl warhead, Topliss scheme, protein mass spectrometry

INTRODUCTION

RNA and its modifications play a significant role in epigenetic inheritance.[1,2] Studies revealed that some RNA modifications are linked to mental[3] and metabolic disorders.[4] Inheritance of metabolic disorders has been found to be caused by increased levels of m^2G and m^5C modifications.[5] The human DNA methyltransferase 2 (hDNMT2), which amongst others, is responsible for m^5C modifications, is involved in this process.[6] Due to its similar sequence and structure hDNMT2 is part of the DNA methyltransferase family, but its main substrate is tRNA at position C38.[7,8] hDNMT2 plays a role in different physiological processes, but is also linked to cancer.[8]

To transfer a methyl group to tRNA, hDNMT2 requires S-adenosyl-L-methionine (SAM) as a cofactor, releasing m^5C38-tRNA^{Asp} and S-adenosyl-L-homocysteine (SAH) as a byproduct.[7] Besides the natural inhibitors SAH and sinefungin,[9] the chemotherapeutic agents 5-azacytidine, decitabine, and zebularine are known to inactivate DNMTs.[10]

Methyltransferases are hard-to-drug targets, especially due to the high cellular concentration of SAM[11] that competes with a potential inhibitor. Furthermore, most of the over 200 different methyltransferases (MTases)[12] in humans are SAM-dependent, which increases the challenge to find selective SAH-based inhibitors. Since covalent inhibitors are capable to compete with high natural ligand concentrations,[13] these issues can be overcome with a warhead-decorated SAH derivative. Moreover, high selectivity can be achieved by targeting the catalytically active cysteines found in the hDNMT[14] or NOL1/NOP2/sun domain (NSUN)[15] families as various other MTases follow different mechanisms.[16] Such covalent modifiers can also be a suitable basis for the development of fluorescently labeled activity-based probes (ABPs),[17] which can be used to improve understanding of RNA methyltransferases and their RNA modifications.

Here, we present covalent SAH-based hDNMT2 inhibitors with 4-halo-3-nitrophenylsulfonamide warheads. This warhead class represents a well-fitting substructure for the cytidine binding site of the enzyme as it not only mimics the cytidine residue of tRNA, but also provides proper orientation and length to reach the catalytically active cysteine-79. Recently, we published hDNMT2 inhibitors based on the SAH scaffold.[9] We replaced the sulfur atom with various N-alkylated substructures, and the N-but-3-yn-2-yl derivative **1** turned out to be the most potent inhibitor of hDNMT2 in a tritium incorporation assay (^{3}H-assay) ($IC_{50} = 12.9 \pm 1.9 \ \mu M$). We also tested the N-benzylated derivative **2,** which showed moderate affinity (57% at 100 μM). Since the phenyl moiety allows a huge space for modifications, we considered the scaffold as a basis for optimization.

Starting from this structure, we developed N-benzyl containing compounds according to the *Topliss* scheme (**Figure 1**). This is an operational scheme in drug design for the derivatization of aromatic compounds considering hydrophobic (π), electronic (σ), and steric (E_s) values of different substituents.[18]

Starting from an unsubstituted phenyl moiety, the 4-chloro analog is initially proposed as it increases the π value. Subsequently, we followed the scheme suggestions depending on the inhibitory effects. Selected derivatives of non-proposed paths were also tested for comparison.

We performed an additional structure activity-relationship study based on the scaffold that exhibited the highest activity. The methylenamine substructure of the *N*-benzyl moiety was replaced with a sulfonamide to test the effect of an additional −I and −M effect. To allow a covalent reaction with the catalytically active cysteine in the cytidine site, electron-deficient aryls with halogen leaving groups were designed to increase electrophilicity. To confirm the covalent reaction between ligand and hDNMT2, protein mass spectrometric experiments were conducted.
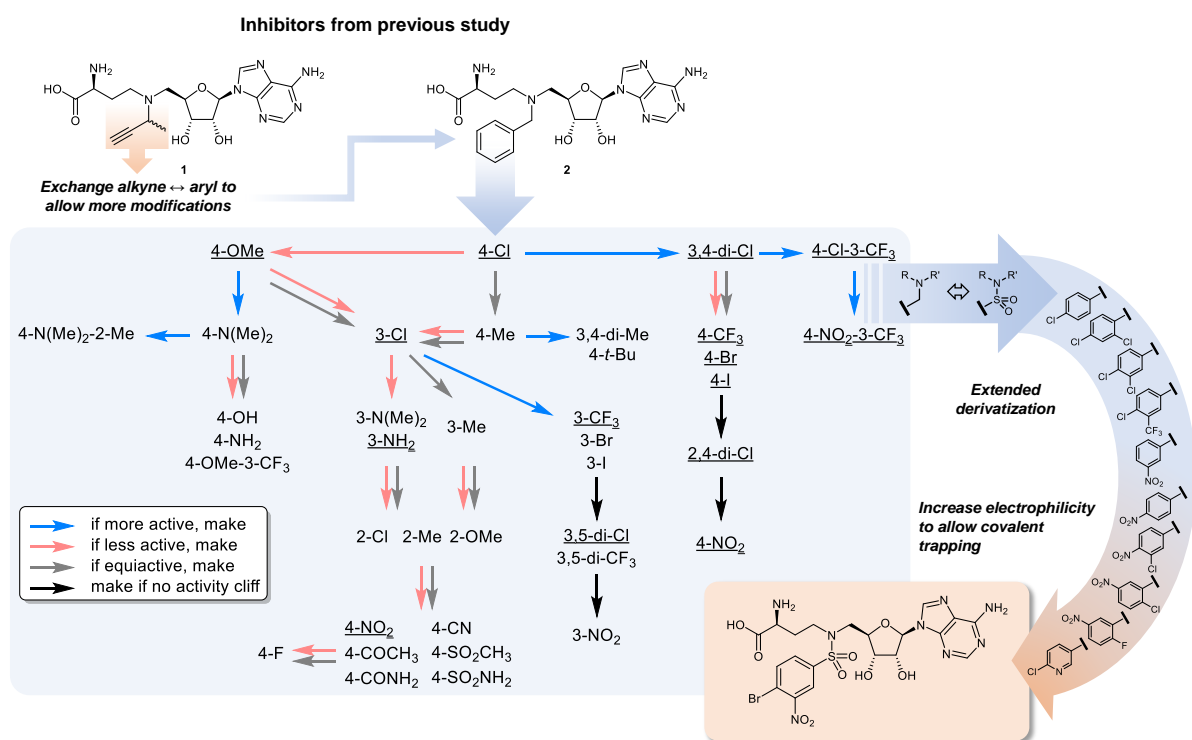
Figure 1: Derivatization of the benzyl amine derivative **2** according to the *Topliss* scheme. Based on the 4-chloro-3-trifluoromethyl and the 4-nitro-3-trifluoromethyl derivatives, an additional SAR study was conducted. First, the aminomethyl substructure was exchanged with a sulfonamide unit to increase the electrophilicity. The aryl moiety was decorated with electron-withdrawing groups and leaving groups to allow a potential covalent reaction with the catalytically active cysteine of hDNMT2.

RESULTS AND DISCUSSION

Chemistry

The benzyl derivatives were synthesized by reductive amination using protected adenosyl-(*S*)-2,4-diaminobutanoic acid (adenosyl-Dab)[9] **3** and substituted benzaldehydes **4**–**17** (**Scheme 1**). For this, sodium triacetoxyborohydride and acetic acid were used as reagents in 1,2-dichloroethane. The building block **3** was prepared according to a previously described procedure.[9] In the final step, the protecting

groups were removed by treatment with 50% (*V/V*) TFA in dichloromethane at 5 °C, followed by treatment with 14% (*V/V*) TFA in water at 5 °C.

Scheme 1: Synthesis of benzyl amine derivatives **32–45**.[a]
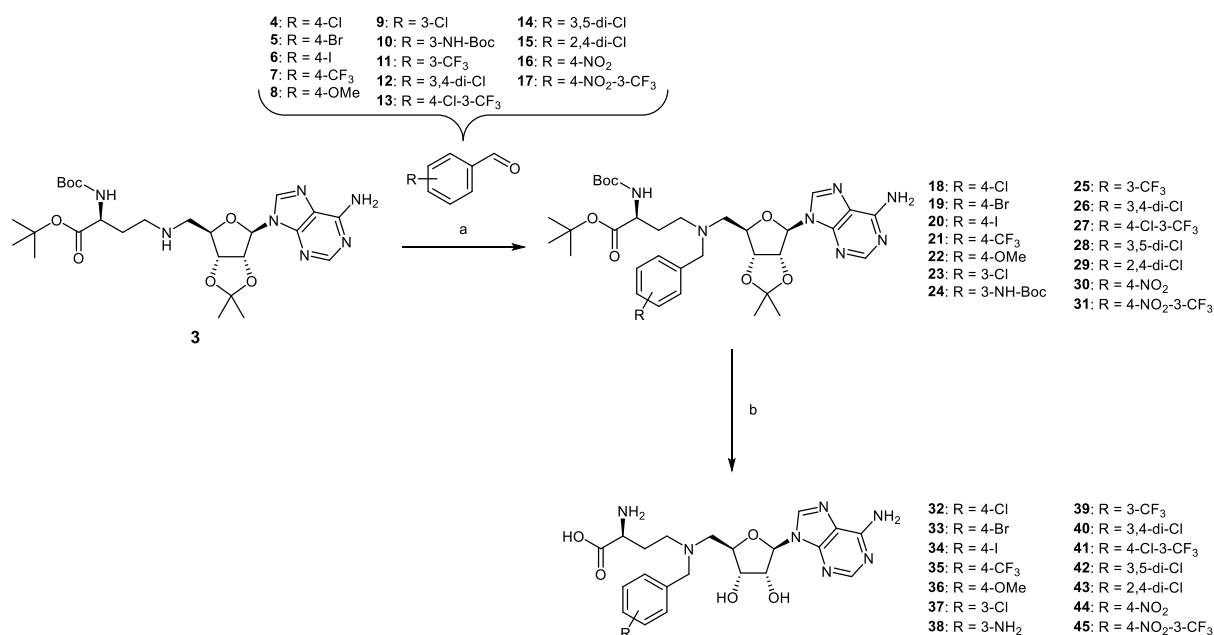
[a]Reagents and conditions: (a) NaBH(OAc)$_3$, HOAc, 1,2-DCE, 0 °C to rt, overnight, 37–96%; (b) (1) TFA/DCM (1:1 *V/V*), 5 °C; (2) TFA/H$_2$O (1:6 *V/V*), 5 °C, 99%.

To further increase the electrophilicity of the aromatic ring, the methylenamine substructure of the benzyl moiety was exchanged by a sulfonamide group. With its −I and −M effects, this compound class shows interesting properties as electron-withdrawing substituents seemed to increase the affinity for hDNMT2. Electron-withdrawing groups that can potentially act as leaving groups, such as halides, were chosen as substituents to enable a possible covalent reaction with the catalytically active cysteine of hDNMT2. To obtain the sulfonamide-based inhibitors, building block **3** was brought to reaction with substituted phenylsulfonyl chlorides **46–58** either in presence of triethylamine in DCM under reflux or using a two-phase system consisting of DCM and saturated NaHCO$_3$ solution at room temperature (**Scheme 2**). In the final step, the resulting precursors **59–71** were deprotected using 50% (*V/V*) TFA in dichloromethane at 5 °C, followed by treatment with 14% (*V/V*) TFA in water at 5 °C to give the inhibitors **72–84**.

Scheme 2: Synthesis of sulfonamide derivatives **72–84**.[a]

46: R = 4-Cl-3-CF₃   53: R = 4-F-3-NO₂
47: R = 4-Cl   54: R = 4-Br-3-NO₂
48: R = 3,4-di-Cl   55: R = 3-NO₂
49: R = 2,4-di-Cl   56: R = 2-Cl-5-NO₂
50: R = H   57: R = 2-F-5-NO₂
51: R = 4-NO₂   58: R = 3-Cl-4-NO₂
52: R = 4-Cl-3-NO₂

59: R = 4-Cl-3-CF₃   66: R = 4-F-3-NO₂
60: R = 4-Cl   67: R = 4-Br-3-NO₂
61: R = 3,4-di-Cl   68: R = 3-NO₂
62: R = 2,4-di-Cl   69: R = 2-Cl-5-NO₂
63: R = H   70: R = 2-F-5-NO₂
64: R = 4-NO₂   71: R = 3-Cl-4-NO₂
65: R = 4-Cl-3-NO₂

72: R = 4-Cl-3-CF₃   79: R = 4-F-3-NO₂
73: R = 4-Cl   80: R = 4-Br-3-NO₂
74: R = 3,4-di-Cl   81: R = 3-NO₂
75: R = 2,4-di-Cl   82: R = 2-Cl-5-NO₂
76: R = H   83: R = 2-F-5-NO₂
77: R = 4-NO₂   84: R = 3-Cl-4-NO₂
78: R = 4-Cl-3-NO₂

[a]Reagents and conditions: (a) NEt₃, DCM, Δ, 2 h, 52–73%; (b) NaHCO₃, H₂O, DCM, rt, overnight, 43–82%; (c) (1) TFA/DCM (1:1 V/V), 5 °C; (2) TFA/H₂O (1:6 V/V), 5 °C, 99%.

To investigate the effect of intracyclic nitrogens as replacements for the nitro groups, pyridine and pyrazine derivatives decorated with chlorine were synthesized (**Scheme 3**). The pyridine derivative **90** was prepared by combining the building block **3** and 6-chloropyridine-3-sulfonyl chloride (**85**) with triethylamine in DCM under reflux to yield the protected precursor **88**. For the synthesis of the pyrazine derivative **91**, 2,5-dichloropyrazine (**86**) was substituted with benzyl mercaptan in presence of sodium hydride in THF to give **87**. In the next step, **87** was treated with sulfuryl chloride,[19] and was reacted with the building block **3** using a two-phase system consisting of DCM and saturated NaHCO₃ solution at room temperature to yield the protected pyrazinyl derivative **89**. The precursors **88** and **89** were finally deprotected using 50% (V/V) TFA in dichloromethane at 5 °C, followed by treatment with 14% (V/V) TFA in water at 5 °C to yield the inhibitors **90** and **91**. A 2-chloropyrimidine derivative was also to be synthesized, but due to its high reactivity already with weak nucleophiles such as methanol and water, which led to substitution of chlorine, it was not further pursued.

Scheme 3: Synthesis of the heterocycles **90** and **91**.[a]
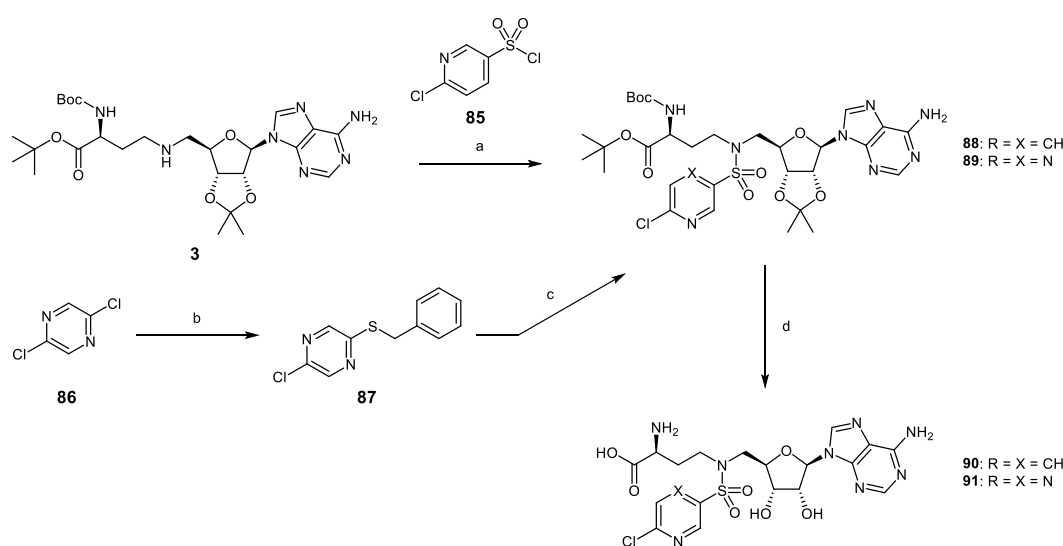
[a]Reagents and conditions: (a) NEt$_3$, DCM, Δ, 2 h, 62%; (b) benzyl mercaptan, NaH, THF, 1 h at 0 °C, 16 h at rt, 91%, (c) (1) sulfuryl chloride, −5 °C, 2 h; (2) **3**, NaHCO$_3$, H$_2$O, DCM, rt, overnight, 34%; (d) (1) TFA/DCM (1:1 *V*/*V*), 5 °C; (2) TFA/H$_2$O (1:6 *V*/*V*), 5 °C, 99%.

Biological Evaluation: hDNMT2 Binding and Inhibition
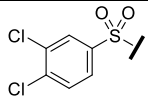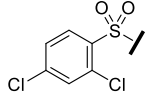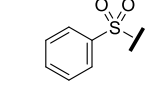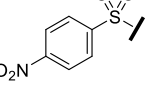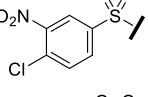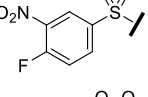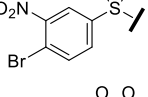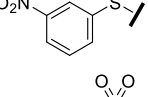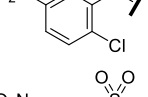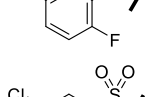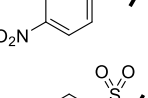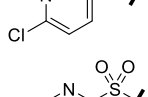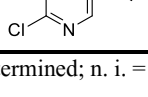
To determine the binding affinity of the compounds towards hDNMT2, a screening was performed with full-length hDNMT2 (if not described otherwise, full-length protein was used) based on a microscale thermophoresis (MST) displacement method using the fluorescent ligand FTAD.[20] Potent binders were defined as compounds that were able to displace FTAD at least in the same extent as sinefungin (MST shift ≥13. For all potent binders an apparent K_D value (K_D^{app}) was determined using MST. All other compounds were not further evaluated. Ligands with the highest affinity towards hDNMT2 (**41**, **45**, **78**–**80**, and **91**; **32** for comparison), were subjected to isothermal titration calorimetry (ITC), which served as an orthogonal method to quantify the binding affinity of those ligands. The results from ITC measurements confirmed the data of the MST measurements as the results of both methods showed a high consistency. Finally, to investigate their actual inhibitory effect, the most potent binders and some selected weak binders were evaluated using a ^{3}H-assay. The assay was performed with tRNA^{Asp} as substrate and tritium-labeled SAM (^{3}H-SAM) as cosubstrate at a compound concentration of 100 μM. For the most potent inhibitors IC$_{50}$ values were determined. All results are summarized in **Table 1**.

Table 1: Binding of compounds to hDNMT2 as determined by MST and inhibition of hDNMT2 as determined in the ^{3}H-assays.
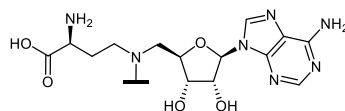
32–45, 72–84, 90, 91

Compound	MST		ITC	^{3}H-assay	
	Shift at 20 μM / ‰	K_D^{app} / μM	K_D / μM	% Inhibition at 100 μM	IC_{50} / μM
SFG	13.3	20.9 ± 0.7[20]	7.5 ± 3.5[9]	83.5 ± 1.1[9]	13.2 ± 0.8[9]
32 (4-Cl)	10.7	28.2 ± 1.4	32.1 ± 8.4	72.3 ± 3.1	62.8 ± 3.0
33 (4-Br)	8.5	n. d.	n. d.	67.3 ± 3.4	n. d.
34 (4-I)	10.8	18.1 ± 1.0	n. d.	75.5 ± 0.8	n. d.
35 ($4\text{-}CF_3$)	11.3	23.7 ± 1.0	n. d.	51.1 ± 6.6	n. d.
36 (4-OMe)	6.2	55.8 ± 4.9	n. d.	64.0 ± 3.1	n. d.
37 (3-Cl)	9.8	n. d.	n. d.	67.2 ± 3.4	n. d.
38 ($3\text{-}NH_2$)	6.3	n. d.	n. d.	n. i.	n. d.
39 ($3\text{-}CF_3$)	8.9	45.1 ± 2.3	n. d.	60.6 ± 3.0	n. d.
40 (3,4-Cl)	11.3	17.9 ± 0.9	n. d.	73.8 ± 3.5	n. d.
41 ($3\text{-}CF_3$, 4-Cl)	14.2	10.4 ± 0.4	6.0 ± 0.7	86.2 ± 3.0	18.3 ± 4.0
42 (3,5-Cl)	8.5	n. d.	n. d.	59.6 ± 4.6	n. d.
43 (2,4-Cl)	6.9	n. d.	n. d.	47.0 ± 3.2	n. d.
44 ($4\text{-}NO_2$)	11.8	14.8 ± 0.6	n. d.	65.5 ± 4.0	n. d.
45 ($3\text{-}CF_3$, $4\text{-}NO_2$)	17.1	4.2 ± 0.3	0.94 ± 0.05	94.0 ± 2.1	2.5 ± 0.2
72 ($3\text{-}CF_3$, 4-Cl, sulfonyl)	11.7	19.5 ± 1.4	n. d.	54.3 ± 3.7	n. d.
73 (4-Cl, sulfonyl)	2.9	n. d.	n. d.	n. i.	n. d.

Table 1 continued.

Compound		MST		ITC	³H-assay	
		Shift at 20 μM / ‰	K_D^{app} / μM	K_D / μM	% Inhibition at 100 μM	IC$_{50}$ / μM
74		7.9	n. d.	n. d.	27.8 ± 17.0	n. d.
75		0.8	n. d.	n. d.	n. d.	n. d.
76		1.2	n. d.	n. d.	n. d.	n. d.
77		2.5	n. d.	n. d.	n. d.	n. d.
78		18.0	2.5 ± 0.1	4.3 ± 0.2	98.2 ± 1.6	2.3 ± 0.5
79		16.3	11.2 ± 0.6	17.3 ± 3.6	92.8 ± 3.7	8.5 ± 1.3
80		18.3	2.3 ± 0.1	4.9 ± 0.4	98.7 ± 1.1	1.2 ± 0.1
81		1.4	n. d.	n. d.	n. i.	n. d.
82		0.2	n. d.	n. d.	n. d.	n. d.
83		2.0	n. d.	n. d.	n. d.	n. d.
84		10.9	22.7 ± 0.9	n. d.	67.0 ± 1.6	n. d.
90		4.1	n. d.	n. d.	33.2 ± 1.9	n. d.
91		17.2	2.6 ± 0.2	1.9 ± 0.3	100.0 ± 6.4	1.1 ± 0.2

n. d. = not determined; n. i. = no inhibition.

Within the benzyl series, 8 out of 14 compounds (**32**, **34**, **35**, **40**, **44**, and **45**) could be classified as promising binders as they caused an MST shift ≥10. Starting with the 4-chloro substituent (**32**) in the *Topliss* scheme, an increase in inhibition from 57% to 72% was achieved. According to the scheme, the proposed modification in this case is the 3,4-dichloro substitution (**40**). With an inhibition of 74% it did

not show a significant increase in potency, resulting in the proposed modifications 4-CF$_3$, 4-Br and 4-I. While the 4-bromo (**33**) and 4-trifluoro (**35**) derivatives showed reduced inhibition (67% and 51% respectively), the 4-iodo modification (**34**) is equipotent (75%). Based on alternative modifications in this branch, we tested the 2,4-dichloro derivative **43**, which further reduced inhibition to 47%. Since the proposed modifications of this branch did not significantly improved inhibition, we followed the branch proposed if the 3,4-dichloro substitution (**40**) would be classified as more potent (73.8 $\pm$ 3.5% vs. 72.3 $\pm$ 3.1%) than the 4-chloro modification (**32**) after all. As a result, the proposed modification 4-chloro-3-trifluoromethyl (**41**) was tested and exhibited a significant increase in inhibition of up to 86%. Based on this result, the scheme proposes the 4-nitro-3-trifluoromethyl derivative **45** as a final modification. Notably, an even higher inhibition of 94% (IC$_{50}$ = 2.5 $\pm$ 0.2 μM) was achieved. To identify potential inhibitors that were likely omitted due to incorrect prediction of the scheme we selected different structures of the remaining branches: 4-OMe (**36**), 3-Cl (**37**), 3-NH$_2$ (**38**), 3-CF$_3$ (**39**), 3,5-diCl (**42**), and 4-NO$_2$ (**44**). The tests revealed that the derivatives **36**, **37**, **39**, and **42** showed moderate inhibition of 59–67%, whereas the 3-amino structure **38** was inactive. Interestingly, the nitro derivative **44** showed a significant increase in inhibition compared to its direct branch precursor **43** (66% vs. 47%). However, its inhibition was still lower than that of compounds **33**–**35** of the same branch. Evaluation of the inhibition of hDNMT2 in correlation with the substituent effects revealed that the results were in high accordance with the *Topliss* scheme suggestions (**Figure 2**). It could be observed that the more electron-withdrawing groups were introduced, the stronger was the inhibition of hDNMT2, which was demonstrated by the following branch: 4-Cl (**32**) $\longrightarrow$ 3,4-diCl (**40**) $\longrightarrow$ 4-Cl-3-CF$_3$ (**41**) $\longrightarrow$ 4-NO$_2$-3-CF$_3$ (**45**) with increasing inhibition of 72% $\longrightarrow$ 74% $\longrightarrow$ 86% $\longrightarrow$ 94% at 100 μM, respectively. While substituents in position 2 significantly reduced inhibition, probably due to steric hindrance within the binding pocket, introduction of several groups at position 3 was tolerated. However, the potency compared to the phenyl derivative **2** was only slightly increased from 56% to 67% by introducing the 3-Cl substituent (**37**). Position 4 also allowed several substituents, e.g. 4-Cl (**32**), 4-Br (**33**), 4-I (**34**), 4-CF$_3$ (**35**), 4-OMe (**36**), and 4-NO$_2$ (**44**) all of which increased the potency compared to the unsubstituted inhibitor **2**. A significant increase in inhibition was achieved with the 4-Cl (72%) and the 4-I (75%) substituents. Interestingly, they differ in electronegativity and volume, yet they showed equipotent inhibition of hDNMT2. So far, a disubstituted, electron-deficient benzyl group appeared to be most effective at enhancing inhibition.
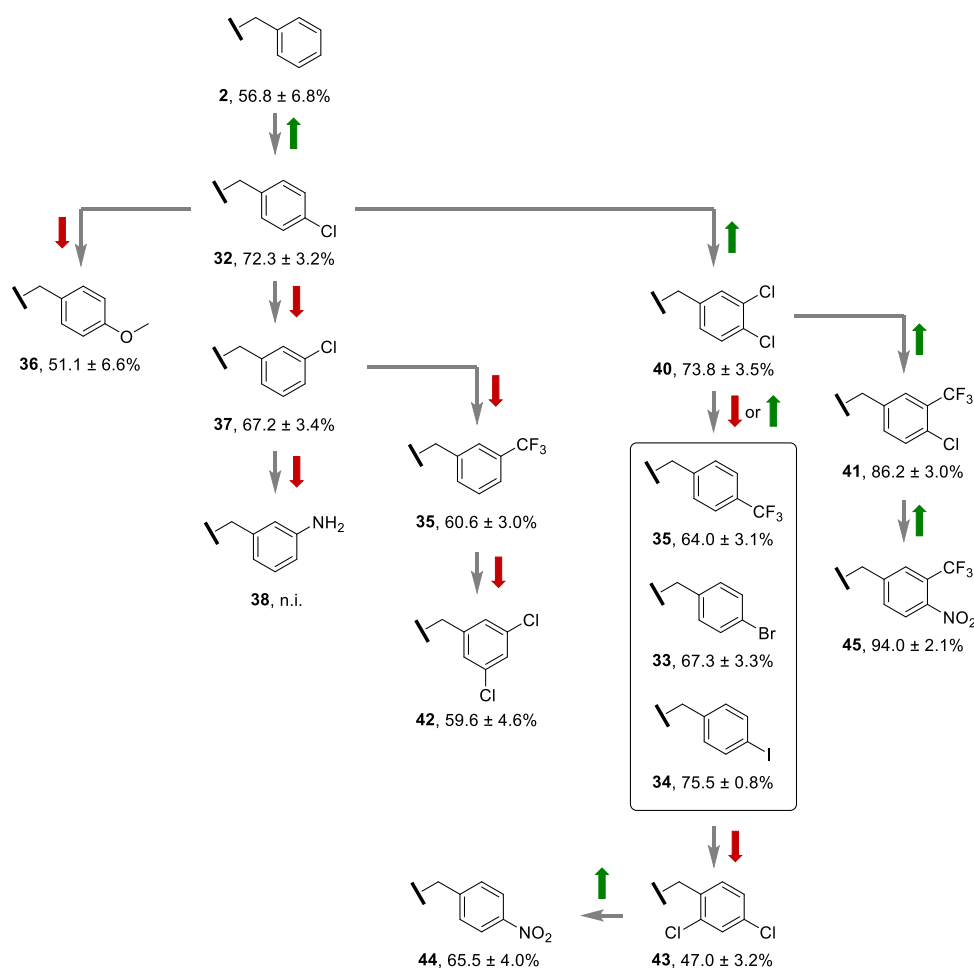
Figure 2: Synthesized compounds and corresponding results with depiction of positions in the *Topliss* scheme. Arrows indicate either increase (green) or decrease (red) in inhibition.

To further increase electrophilicity of the aromatic ring, we replaced the methylenamine substructure of the *N*-benzyl moiety with a sulfonamide that introduces additional −I and −M effects. Based on this modification, a subsequent structure-activity relationship (SAR) study was conducted using different aryl sulfonyl moieties with strong electron-withdrawing groups such as NO_2 and CF_3 as well as intracyclic nitrogen atoms. To enable a possible covalent reaction with the catalytically active cysteine, the aromatic rings were decorated with halogen leaving groups. First, we started with the 4-chloro-3-trifluoromethyl analog **72** for comparison with one of the most active compounds (**41**). With an inhibition of 54% it showed a significantly lower inhibition than **41** with 86%, which indicated a negative effect of the sulfonamide group. The same holds true for **73** (4-Cl, no inhibition), **74** (3,4-diCl, 28% inhibition), **75** (2,4-diCl, no inhibition), **76** (unsubstituted, no inhibition), and **77** (4-NO_2, no inhibition), which caused significantly lower inhibition compared to their corresponding benzyl derivatives. Interestingly, the 4-chloro-3-nitro derivative **78** had the highest inhibitory activity of 98% ($IC_{50} = 2.3 \pm 0.5$ μM). Given **78** as the new lead structure, we analyzed changes in the substitution pattern. For the 4-F-3-NO_2 modification (**79**) a slightly lower inhibition of 93% ($IC_{50} = 8.5 \pm 1.3$ μM) was observed while the 4-Br-3-NO_2 derivative **80** was found to be equipotent with an inhibition of 99% ($IC_{50} = 1.2 \pm 0.1$ μM). Other structural

changes such as 3-NO$_2$ (**81**), 2-Cl-5-NO$_2$ (**82**), and 2-F-5-NO$_2$ (**83**) did not lead to inhibition of hDNMT2. The 3-Cl-4-NO$_2$ derivative **84** showed only moderate inhibition of 67%. Replacing the 3-nitro group of **78** with an intracyclic nitrogen (**90**) atom resulted in a strong reduction of inhibition to 33%. However, a second intracyclic nitrogen located in the opposite position, as found in pyrazinyl moieties, increased the inhibition to 100% (**91**, IC$_{50}$ = 1.1 ± 0.2 μM). Comparing the benzyl derivatives with the sulfonamide-based compounds, an obvious trend is observable. The sulfonamide derivatives showed significantly lower inhibition compared to their benzyl analogs, e. g. **32** vs. **73** (4-Cl), **40** vs. **74** (3,4-diCl), and **41** vs. **72** (4-Cl-3-CF$_3$). Yet the 4-Cl/Br-3-NO$_2$ substituted sulfonamide-based derivatives **78** and **80** as well as the 4-chloropyrazinyl derivative **91** showed the highest inhibition (IC$_{50}$ < 2.5 μM). The results also highlight the quality of our FTAD-based MST screening method[20] since the screening results showed high consistency with the measured inhibition in the ^{3}H-assay.

Warhead stability in presence of DTT

We found that the inhibition of hDNMT2 by compound **79** was highly dependent on the presence of the reducing agent used in the assays. While inhibition of 93% (IC$_{50}$ = 8.5 ± 1.3 μM) was measured in the presence of TCEP, no inhibition could be observed in the presence of DTT. These findings strongly suggest a covalent reaction of the thiol-based agent with the electrophilic inhibitor. For this reason, we conducted stability tests with DTT and the inhibitors **78**–**80** in TRIS buffer pH = 8.0 at room temperature under assay-like conditions. The inactivation was determined by LC-MS after 2 min followed by 15 min intervals (**Figure 3**).
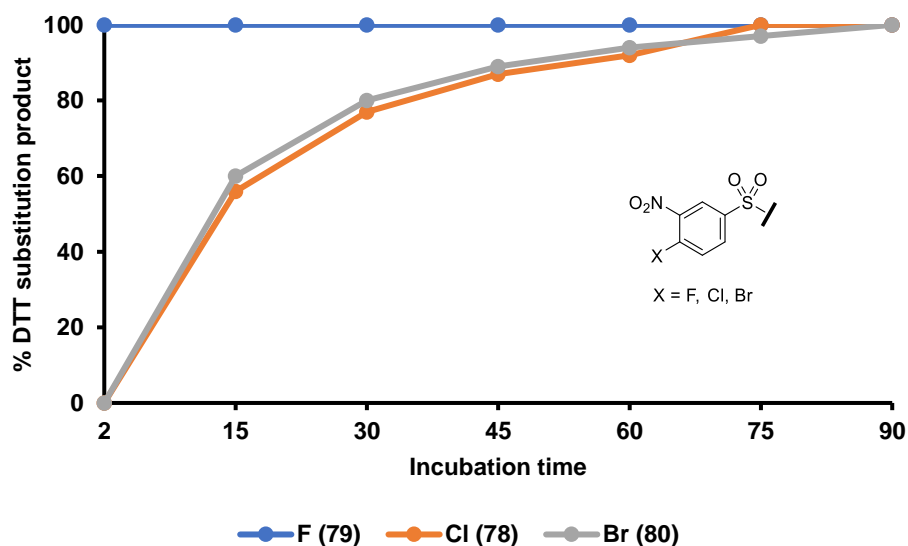
Figure 3: Stability test of inhibitors **78**–**80** in presence of DTT under assay-like conditions. Percent inactivation was determined initially after 2 min and then at 15 min intervals.

LC-MS measurements confirmed the formation of a DTT substitution product in all three cases. **Figure 3** shows that **79** already completely reacted with DTT after 2 min. **78** reacts with a rate comparable to **80**,

but much slowlier than **79**. In the first 15 min both react very quickly until a conversion of ca. 60% was reached. A full conversion was observed after 90 min, respectively.

Protein mass spectrometry

To investigate the binding of the ligands **45**, **78–80**, and **91** to hDNMT2, we used intact-protein LC-MS under denaturing conditions, which only preserve covalently bound ligands. Additionally, direct infusion nano-electrospray ionization experiments were performed under non-denaturing conditions, which allow preservation of non-covalent binding. hDNMT2 harbors a flexible loop (residues 191–237) that decreases protein stability and prevents crystallization. However, a construct without this loop was co-crystallized with SAH (PDB-ID: 1G55). Based on this construct a deletion mutant without residues 191–237 (hDNMTΔ) was designed which was better suited for mass spectrometry than hDNMT2. hDNMT2Δ showed decreased but measurable activity in the ^{3}H-assay, which supports the hypothesis that hDNMT2Δ binds SAH. For all experiments, the known non-covalent ligand SFG was used as a control. All LC-MS spectra exhibited the broad charge state distribution and high charge states typical for denatured proteins. In **Figure 4A** the results of these measurements, zoomed in on the region around charge state 32+ (*m/z* 1293), are shown. For **78–80**, the same mass shift of 551 Da, corresponding to addition of the structure highlighted in orange was observed. This indicates loss of the halogen substituent and the formation of a covalent bond between the protein and ligand. A second addition of ligand **79** was detected, which suggests the existence of two potential binding sites of this ligand and therefore a less specific binding of the desired target cysteine. A comparable result was observed with ligand **91** where two mass shifts of 508 Da, corresponding to the blue-highlighted structure, were detected. Here, a very low amount of remaining free protein signal was observed, which indicates a fast and favorable reaction of this ligand with hDNMT2Δ. For **45** and the known non-covalent ligand SFG, no mass shift and thus no covalent irreversible binding was demonstrated. A table with observed and calculated masses, mass errors, and intensities can be found in the **Supporting Information**. Additional measurements under near-native conditions were performed with SFG and **45** (**Figure 4B**). As is typical in native MS, these spectra show a narrow charge state distribution with low charge states. For both ligands, a shift by the corresponding mass was detected (570 Da for **45**; 381 Da for SFG). Furthermore, a second binding event of ligand **45** was observed. This demonstrates the non-covalent binding of **45** and SFG to hDNMT2Δ. This second, but weaker binding event observed for **45** and **74** could also explain the increased binding stoichiometry determined in the ITC experiments (1.37 ± 0.01 and 1.34 ± 0.04 respectively). On the other hand, for **91** a decreased binding stoichiometry (0.84 ± 0.01) was measured leading to the assumption that this compound is too reactive and therefore could cause partial protein aggregation. All other compounds investigated with ITC showed binding stoichiometries in the range from 0.9 to 1.1 towards hDNMT2Δ, which correlates nicely with data obtained from protein MS. We determined the binding sites of the covalent ligands with a tryptic digest followed by bottom-up LC-MS/MS. A high sequence coverage of
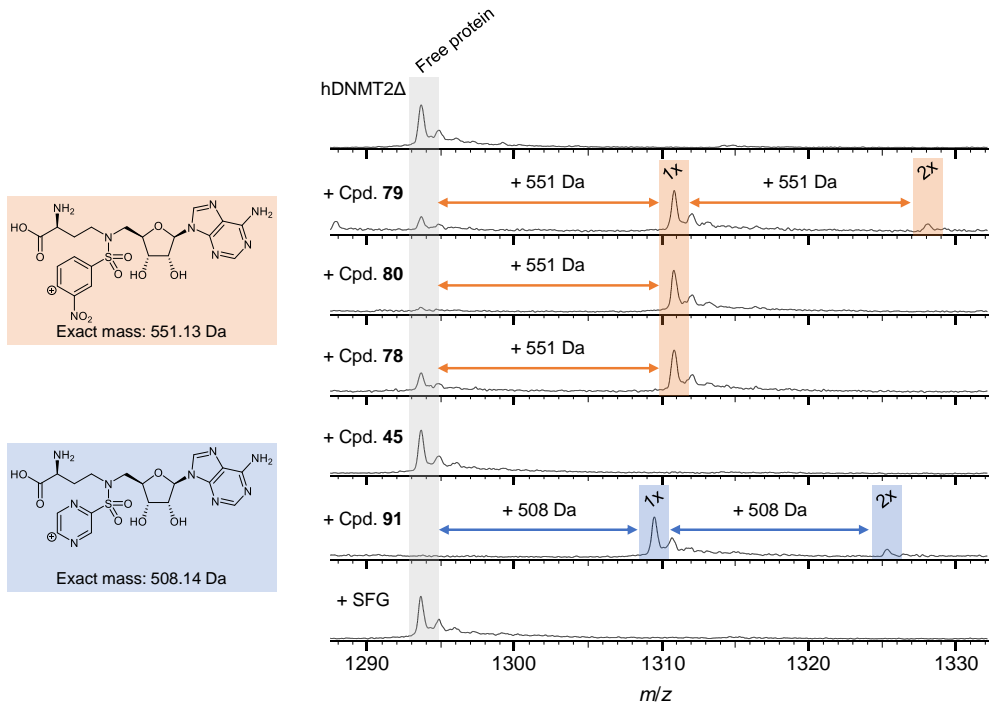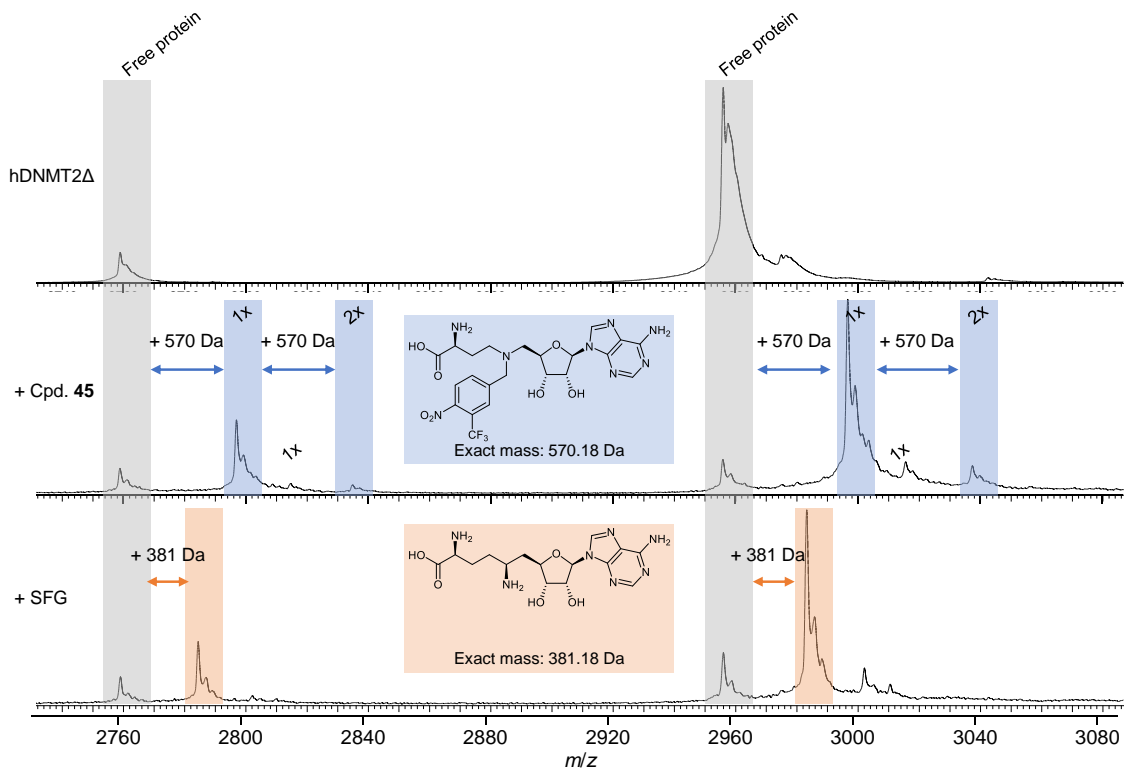
Figure 4: (A) LC-MS of intact, denatured hDNMT2Δ with and without ligands. Corresponding peak for free protein signal is highlighted in grey (32+ charge state). For **45** and SFG, no binding was detected under denaturing conditions. Binding to **78**–**80** all led to addition of the same 551 Da moiety (mass shift and corresponding structure shown in orange). Binding to **91** led to addition of a 508 Da moiety (structure and mass shift shown in blue). (B) Native MS of hDNMT2Δ in complex with **45** and SFG. Peaks corresponding to free protein signal (charge states 14+ and 15+) are highlighted in grey. Binding to **45** (two binding events) and SFG (one event) are highlighted in blue and orange, respectively.

over 80% was achieved for all samples. For **78**–**80**, binding to Cys79, which is known to be the catalytically active cysteine, was detected. In agreement with the denatured intact mass measurements, which indicated a second binding site for **79**, an additional modified peptide containing the cysteine residue corresponding to Cys287 in the wild-type protein was detected. Similarly, binding of **91** was also observed at Cys79 and Cys287. As expected, no modified peptides were detected for **45** and SFG, which is consistent with a non-covalent binding mode. These findings indicate two covalent binding sites for **79** and **91**, and the binding to only the desired target Cys79 for **78** and **80**. A detailed overview of the results is shown in the **Supporting Information**.

Selectivity

Selectivity of the most promising inhibitors (**45**, **78**–**80**, and **91**) towards other tRNA modifying m^{5}C MTases (NSUN2 and NSUN6) was measured in a ^{3}H-assay at 100 μM (**Figure 5**). All selected inhibitors showed higher selectivity compared to SFG[9] with compound **80** being the most selective one as it did not inhibit NSUN2 and NSUN6. Compounds **78** and **91** appear to be highly selective as they showed only slight inhibition of NSUN2 (12 ± 6.8% and 13 ± 9.2%) and NSUN6 (n. i. and 7 ± 4.9%). Lower selectivity was observed for **45** (12 ± 1.3% for NSUN2; 30±0.1% for NSUN6) and **79** (23 ± 4.9% for NSUN2; 69 ± 4.8% for NSUN6). Given that **79** is highly reactive, a lack of selectivity was expected. Furthermore, detection of beyond active site binding to hDNMT2Δ for **45** and **79** via LC-MS already indicated a promiscuous binding behavior of those compounds.
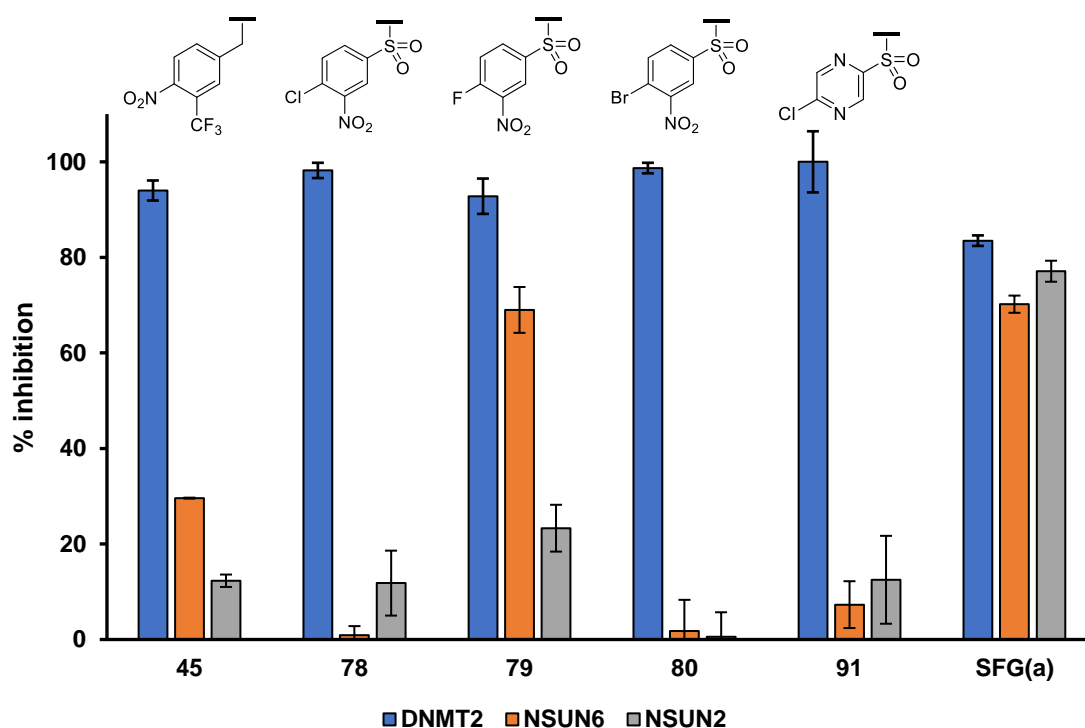
Figure 5: Inhibition of different RNA methyltransferases by most potent DNMT2 inhibitors at 100 μM. Triplicates ± SD are given. (a) Data from previous study.

Docking studies confirm proper orientation allowing covalent reaction

To investigate the binding mode of compounds **45**, **78**–**80**, and **91** and confirm a proper orientation that enables a covalent reaction, docking studies using FlexX[21] and MOE (*Molecular Operating Environment*, 2022.02 Chemical Computing Group ULC, Montreal, Canada, 2023) were performed (**Supporting Information**). Since the catalytic loop (residues 79–96) is not resolved in the crystal structure of hDNMT2 (PDB-ID: 1G55), it was introduced using the 'Loop Modeler' functionality within MOE.

Based on the docking-predicted binding modes, the benzyl derivative **45** as well as the aromatic sulfonamides **78**–**80** and **91** expand from the SAM-site towards the cytidine binding site. **Figure 6** shows the binding poses of compound **80** as an example. The proximity of the electrophilic carbon atoms of to the nucleophilic sulfur atom of the catalytic Cys79 of 3.9–4.2 Å (**Table S3**) indicated high likeliness of a covalent reaction.
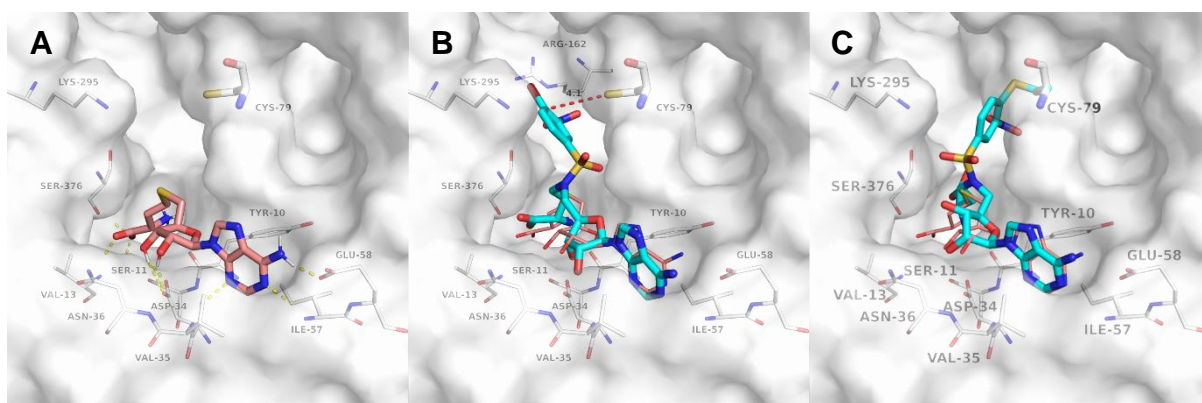
Figure 6. Non-covalent and covalent docking of compound **80** (cyan) with SAH as reference (salmon). A: Crystal structure of SAH (PDB-ID: 1G55). B: Non-covalent docking of compound **80**. The distance of the electrophilic carbon to Cys79 is indicated with a red dotted line. C: Covalent docking of compound **80**.

CONCLUSION

Here, we presented covalent SAH-based hDNMT2 inhibitors with a new type of aryl warhead. By successfully applying the *Topliss* scheme to the moderate inhibitor *N*-benzyl-adenosyl-Dab **2**[9], electron-deficient benzyl derivatives (4-Cl-3-CF$_3$ **41**, 4-NO$_2$-3-CF$_3$ **45**) exhibiting stronger hDNMT2 inhibition were identified. Based on these findings, the electrophilicity was further increased by replacing the methylenamine substructure with a sulfonamide function. We performed a subsequent SAR study using different aryl sulfonyl building blocks with strong electron-withdrawing groups such as NO$_2$ and CF$_3$ as well as intracyclic nitrogen atoms. By attaching halogen leaving groups the aromatic rings were adjusted to enable a covalent reaction with the catalytically active cysteine of hDNMT2. Protein mass spectrometry revealed that the 4-halogen-3-NO$_2$-decorated phenylsulfonamide derivatives **78**–**80** and the chloropyrazinyl structure **91** reacted covalently with the catalytically active Cys79 in the cytidine site of hDNMT2. However, compounds **79** and **91** exhibited too high reactivity as they also bound to Cys287 on the protein surface. Furthermore, a stability test showed that **79** was quickly inactivated by reaction

with DTT (**Figure 3**). Our most promising covalent inhibitor turned out to be the 4-Br-3-NO_2-phenylsulfonamide derivative **80**, with an IC_{50} value of 1.2 ± 0.1 μM resulting in an improvement of one order of magnitude compared to our previously published inhibitors.[9] Therefore, it outclasses the natural ligands SAH and SFG. Moreover, compound **80** showed high selectivity towards NSUN2 and NSUN6 (**Figure 5**). With the discovery of covalent hDNMT2 inhibitors, this study provides a suitable basis for the development of fluorescently ABPs. Such tool compounds can be used for future studies to improve understanding of RNA methyltransferases and the biological impact of their RNA modifications.

ASSOCIATED CONTENT

Supporting Information

Synthesis protocols and analytical data of all compounds, NMR spectra, LC-MS chromatograms of all tested compounds, MST traces, ITC curves, molecular docking results (PDF).

AUTHOR CONTRIBUTION

‡M.S. and R.A.Z. contributed equally.

FUNDING

NOTES

Mark Helm is a consultant for Moderna Inc.

ABBREVIATIONS

ABP, activity-based probe; 5-FAM, 5-carboxyfluorescein; Dab, (*S*)-2,4-diaminobutanoic acid; DCE, dichloroethane; DCM, dichloromethane; DNMT, DNA *N*-methyltransferase; DTT, dithiothreitol; FTAD, 5-FAM-triazolyl-adenosyl-Dab; ITC, isothermal titration calorimetry; MST, microscale thermophoresis; MTase, methyltransferase; NSUN, NOL1/NOP2/sun domain; RMSD, root-mean-square deviation; SAH, *S*-adenosylhomocysteine; SAM, *S*-adenosylmethionine; SFG, sinefungin; TCEP, tris(2-carboxyethyl)phosphine; TFA, trifluoroacetic acid; THF, tetrahydrofuran; TRIS, tris(hydroxymethyl)aminomethane.

REFERENCES

1   Jonkhout, N.; Tran, J.; Smith, M. A.; Schonrock, N.; Mattick, J. S.; Novoa, E. M. The RNA Modification Landscape in Human Disease. *RNA* **2017**, *23* (12), 1754–1769. https://doi.org/10.1261/rna.063503.117.

2   Liebers, R.; Rassoulzadegan, M.; Lyko, F. Epigenetic Regulation by Heritable RNA. *PLoS Genet.* **2014**, *10* (4). https://doi.org/10.1371/journal.pgen.1004296.

3   Wang, Y.; Chen, Z.-P.; Hu, H.; Lei, J.; Zhou, Z.; Yao, B.; Chen, L.; Liang, G.; Zhan, S.; Zhu, X.; Jin, F.; Ma, R.; Zhang, J.; Liang, H.; Xing, M.; Chen, X.-R.; Zhang, C.-Y.; Zhu, J.-N.; Chen, X. Sperm MicroRNAs Confer Depression Susceptibility to Offspring. *Sci. Adv.* **2021**, *7* (7). https://doi.org/10.1126/sciadv.abd7605.

4   Zhang, Y.; Shi, J.; Rassoulzadegan, M.; Tuorto, F.; Chen, Q. Sperm RNA Code Programmes the Metabolic Health of Offspring. *Nat. Rev. Endocrinol.* **2019**, *15* (8), 489–498. https://doi.org/10.1038/s41574-019-0226-2.

5   Chen, Q.; Yan, M.; Cao, Z.; Li, X.; Zhang, Y.; Shi, J.; Feng, G.; Peng, H.; Zhang, X.; Zhang, Y.; Qian, J.; Duan, E.; Zhai, Q.; Zhou, Q. Sperm TsRNAs Contribute to Intergenerational Inheritance of an Acquired Metabolic Disorder. *Science* **2016**, *351* (6271), 397–400. https://doi.org/10.1126/science.aad7977.

6   Zhang, Y.; Zhang, X.; Shi, J.; Tuorto, F.; Li, X.; Liu, Y.; Liebers, R.; Zhang, L.; Qu, Y.; Qian, J.; Pahima, M.; Liu, Y.; Yan, M.; Cao, Z.; Lei, X.; Cao, Y.; Peng, H.; Liu, S.; Wang, Y.; Zheng, H.; Woolsey, R.; Quilici, D.; Zhai, Q.; Li, L.; Zhou, T.; Yan, W.; Lyko, F.; Zhang, Y.; Zhou, Q.; Duan, E.; Chen, Q. Dnmt2 Mediates Intergenerational Transmission of Paternally Acquired Metabolic Disorders through Sperm Small Non-Coding RNAs. *Nat. Cell Biol.* **2018**, *20* (5), 535–540. https://doi.org/10.1038/s41556-018-0087-2.

7   Goll, M. G.; Kirpekar, F.; Maggert, K. A.; Yoder, J. A.; Hsieh, C. L.; Zhang, X.; Golic, K. G.; Jacobsen, S. E.; Bestor, T. H. Methylation of TRNAAsp by the DNA Methyltransferase Homolog Dnmt2. *Science* **2006**, *311* (5759), 395–398. https://doi.org/10.1126/science.1120976.

8   Jeltsch, A.; Ehrenhofer-Murray, A.; Jurkowski, T. P.; Lyko, F.; Reuter, G.; Ankri, S.; Nellen, W.; Schaefer, M.; Helm, M. Mechanism and Biological Role of Dnmt2 in Nucleic Acid Methylation. *RNA Biol.* **2017**, *14* (9), 1108–1123. https://doi.org/10.1080/15476286.2016.1191737.

9   Schwickert, M.; Fischer, T. R.; Zimmermann, R. A.; Hoba, S. N.; Meidner, J. L.; Weber, M.; Weber, M.; Stark, M. M.; Koch, J.; Jung, N.; Kersten, C.; Windbergs, M.; Lyko, F.; Helm, M.; Schirmeister, T. Discovery of Inhibitors of DNA Methyltransferase 2, an Epitranscriptomic Modulator and Potential Target for Cancer Treatment. *J. Med. Chem.* **2022**, *65* (14), 9750–9788. https://doi.org/10.1021/acs.jmedchem.2c00388.

10 Ganesan, A.; Arimondo, P. B.; Rots, M. G.; Jeronimo, C.; Berdasco, M. The Timeline of Epigenetic Drug Discovery: From Reality to Dreams. *Clin. Epigenetics* **2019**, *11* (1), 174. https://doi.org/10.1186/s13148-019-0776-0.

11 Fontecave, M.; Atta, M.; Mulliez, E. S-Adenosylmethionine: Nothing Goes to Waste. *Trends Biochem. Sci.* **2004**, *29* (5), 243–249. https://doi.org/10.1016/j.tibs.2004.03.007.

12 Petrossian, T. C.; Clarke, S. G. Uncovering the Human Methyltransferasome. *Mol. Cell. Proteomics* **2011**, *10* (1), M110.000976. https://doi.org/10.1074/mcp.M110.000976.

13 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. https://doi.org/10.1016/j.ejmech.2017.06.019.

14   Lyko, F. The DNA Methyltransferase Family: A Versatile Toolkit for Epigenetic Regulation. *Nat. Rev. Genet.* **2018**, *19* (2), 81–92. https://doi.org/10.1038/nrg.2017.80.

15   Bohnsack, K. E.; Höbartner, C.; Bohnsack, M. T. Eukaryotic 5-Methylcytosine (m^{5}C) RNA Methyltransferases: Mechanisms, Cellular Functions, and Links to Disease. *Genes* **2019**, *10* (2). https://doi.org/10.3390/genes10020102.

16   Sun, Q.; Huang, M.; Wei, Y. Diversity of the Reaction Mechanisms of SAM-Dependent Enzymes. *Acta Pharm. Sin. B* **2021**, *11* (3), 632–650. https://doi.org/10.1016/j.apsb.2020.08.011.

17   Lee, J.; Schapira, M. The Promise and Peril of Chemical Probe Negative Controls. *ACS Chem. Biol.* **2021**, *16*, 579–585. https://doi.org/10.1021/acschembio.1c00036.

18   Topliss, J. G. Utilization of Operational Schemes for Analog Synthesis in Drug Design. *J. Med. Chem.* **1972**, *15* (10), 1006–1011. https://doi.org/10.1021/jm00280a002.

19   Cooper, M.; Miller, D.; Macleod, A.; Van Wiltenburg, J.; Thom, S.; St-Gallay, Stephen Shannon, J.; Alanine, T.; Onions, S.; Strutt, I. Novel Sulfonamide Carboxamide Compounds. WO2019/8025 A1, 2019.

20   Zimmermann, R. A.; Schwickert, M.; Meidner, J. L.; Nidoieva, Z.; Helm, M.; Schirmeister, T. An Optimized Microscale Thermophoresis Method for High-Throughput Screening of DNA Methyltransferase 2 Ligands. *ACS Pharmacol. Transl. Sci.* **2022**, *5* (11), 1079–1085. https://doi.org/10.1021/acsptsci.2c00175.

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

3.5.3 Experimental section of the manuscript

Note: Syntheses conducted or supervised by MARVIN SCHWICKERT (████████ group) are described in this section. Experiments performed by others can be found in the complete Supporting Information file in the **Appendix**. Compound numbers in this section refer to the numbering in the corresponding manuscript.

General information

All reagents and solvents were commercial grade and used without further purification. Reaction progress was monitored by thin-layer chromatography using ALUGRAM® Xtra SIL G UV$_{254}$ silica plates from *Machery-Nagel*. Column chromatography was performed with silica (40–63 μm) from *Machery-Nagel*. Flash chromatography was performed with a *Biotage Isolera^{TM} One* system using prepacked C$_{18}$ columns from *Biotage*. NMR spectra were recorded on a *Bruker Fourier 300* at 300 MHz. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance (CDCl$_3$, DMSO-d_6, or CD$_3$OD from *Deutero GmbH*) as internal standard. The identities and purities of final compounds were determined by combined HPLC/ESI-MS analysis using an *Agilent 1100 series* HPLC system with an *Agilent Zorbax SB-Aq* (4.6 × 150 mm, 5 μm; mobile phase: MeCN/MilliQ®-H$_2$O + 0.1% HCOOH = 20:80 or 30:70; flow rate: 0.7 mL/min) column at 40 °C oven temperature. Samples were applied using 5 μL injection with quantitation by AUC at 254 nm. Electro spray ionization (ESI-) mass spectra were recorded on an *Agilent 1100 series LC/MSD Ion trap* spectrometer in the positive ion mode. 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. Specific rotations $[\alpha]_D^{20}$ were determined with a *Krüss P3000* polarimeter and are given in deg cm^3 g^{-1} dm^{-1}. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

General Procedure A – Reductive Amination

To a solution of **3** (1.0 equiv.) in 1,2-DCE (5 mL) at 0 °C under argon atmosphere were added an aldehyde (2.0 equiv.) and HOAc (1.5 equiv.). The mixture was stirred at 0 °C for 30 min and then NaBH(OAc)$_3$ (1.7 equiv.) was added. The reaction was slowly warmed up to room temperature and stirred overnight. The mixture was quenched by the addition of saturated NaHCO$_3$ solution (20 mL), and the aqueous layer was extracted with ethyl acetate (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (DCM/MeOH = 30:1).

General Procedure B – Reductive Amination

To a solution of **3** (1.0 equiv.) in 1,2-DCE (5 mL) at 0 °C under argon atmosphere were added an aldehyde (2.0 equiv.) and HOAc (1.5 equiv.). The mixture was stirred at 0 °C for 30 min and then NaBH(OAc)$_3$ (1.7 equiv.) was added. The reaction was slowly warmed up to room temperature and stirred overnight. The mixture was quenched by the addition of MeOH (5 mL) and the solvent was removed under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (DCM/MeOH = 30:1).

General Procedure C – Deprotection

To a solution of a protected compound in DCM (2 mL) at 5 °C was added TFA (2 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid was detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL) at 40 °C, and the residue was dissolved in water (3 mL) and TFA (0.5 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then, the mixture was diluted with water (15 mL) and dried by lyophilization to obtain the final compound.

General Procedure D – Sulfonamide formation 1

To a solution of **3** (1.0 equiv.) and NEt$_3$ (1.0 equiv.) in DCM (6 mL) was added the sulfonyl chloride (1.0 equiv.). The solution was heated to reflux for 2 h and the solvent was removed under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (DCM/MeOH = 30:1).

General Procedure E – Sulfonamide formation 2

To a mixture of **3** (1.0 equiv.) in DCM (5 mL) and saturated NaHCO$_3$ solution (5 mL) was added the sulfonyl chloride (1.5 equiv.). After stirring overnight, the organic phase was separated, and the aqueous phase was extracted with DCM (2 × 10 mL). The combined organic extracts were dried over Na$_2$SO$_4$, filtered, and the solvent was removed under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (DCM/MeOH = 30:1) and, if necessary, on C$_{18}$ (MeCN/H$_2$O = 10:90 → 100:0).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d]-[1,3]dioxol-4-yl)methyl)(4-chlorobenzyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (18)
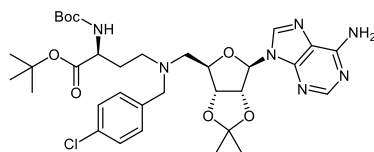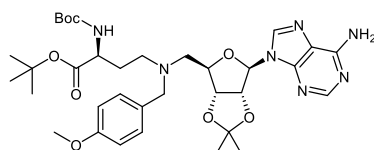

The compound was prepared from **3** (200 mg, 0.35 mmol, 1.0 equiv.), 4-chlorobenzaldehyde (100 mg, 0.71 mmol, 2.0 equiv.), HOAc (30 μL, 0.53 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless solid (137 mg, 0.20 mmol, 57%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.16 (s, 1H), 7.82 (s, 1H), 7.16–7.08 (m, 4H), 6.49 (s, 2H), 6.00 (d, J = 2.0 Hz, 1H), 5.66 (d, J = 8.1 Hz), 5.42–5.31 (m, 1H), 4.86 (dd, J = 6.5, 3.5 Hz), 4.30 (pseudo-td, J ≈ 6.5, 3.5 Hz), 4.19–4.09 (m, 1H), 3.59 (d, J = 13.7 Hz, 1H), 3.38 (d, J = 13.7 Hz, 1H), 2.74 (dd, J = 13.8, 6.3 Hz, 1H), 2.59 (dd, J = 13.3, 6.8 Hz, 2H), 2.48–2.38 (m, 1H), 2.04–1.83 (m, 1H), 1.78–1.63 (m, 1H), 1.55 (s, 3H), 1.41–1.30 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.9, 155.5, 152.9, 149.0, 139.9, 137.1, 132.8, 130.2, 128.3, 120.2, 114.3, 90.7, 85.4, 83.8, 83.5, 81.7, 79.4, 58.3, 55.7, 52.8, 50.6, 29.4, 28.4, 27.9, 27.2, 25.5. FT-IR: ν / cm^{-1} = 2980, 1704, 1642, 1597, 1490, 1367, 1210, 1153, 1088, 907, 870, 799, 728, 665. $[\alpha]_D^{20}$ = −23 (10 mg/mL; MeOH). mp: 84–87 °C. ESI-MS: m/z calculated for [C$_{33}$H$_{46}$ClN$_7$O$_7$+H]$^+$ ([M+H]$^+$) = 688.3, found: 688.3.

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d]-[1,3]dioxol-4-yl)methyl)(4-bromobenzyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (19)
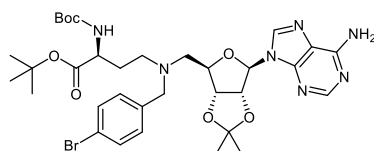
The compound was prepared from **3** (150 mg, 0.27 mmol, 1.0 equiv.), 4-bromobenzaldehyde (98 mg, 0.53 mmol, 2.0 equiv.), HOAc (23 μL, 0.40 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (96 mg, 0.45 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless oil (192 mg, 0.26 mmol, 96%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.16 (s, 1H), 7.83 (s, 1H), 7.29 (d, J = 8.1 Hz, 2H), 7.06 (d, J = 8.1 Hz, 2H), 6.50 (s, 2H), 6.00 (d, J = 1.9 Hz, 1H), 5.64 (d, J = 8.1 Hz, 1H), 5.40–5.29 (m, 1H), 4.86 (dd, J = 6.4, 3.5 Hz, 1H), 4.29 (pseudo-td, J ≈ 6.6, 3.4 Hz, 1H), 4.21–4.05 (m, 1H), 3.57 (d, J = 13.7 Hz, 1H), 3.37 (d, J = 13.8 Hz, 1H), 2.74 (dd, J = 13.8, 6.3 Hz, 1H), 2.58 (dd, J = 13.4, 6.8 Hz, 2H), 2.50–2.36 (m, 1H), 2.00–1.80 (m, 1H), 1.76–1.61 (m, 1H), 1.55 (s, 3H), 1.42–1.31 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.9, 155.5, 153.0, 149.1, 140.0, 137.7, 131.3, 130.5, 120.9, 120.3, 114.3, 90.7, 85.4, 83.8, 83.5, 81.7, 79.5, 58.4, 55.7, 52.8, 50.6, 29.4, 28.4), 27.9, 27.2, 25.4. FT-IR: ν / cm^{-1} = 2980, 1704, 1641, 1596, 1486, 1367, 1329, 1214, 1152, 1070, 1011, 870, 798, 749,

666. $[\alpha]_D^{20} = -6$ (10 mg/mL; CHCl$_3$). ESI-MS: *m/z* calculated for $[C_{33}H_{46}BrN_7O_7+H]^+$ ($[M+H]^+$) = 732.3, found: 732.2.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(4-iodobenzyl)amino)-2-((_tert_-butoxycarbonyl)amino)butanoate (20)
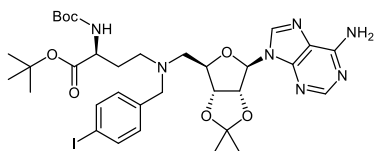
The compound was prepared from **3** (150 mg, 0.27 mmol, 1.0 equiv.), 4-iodobenzaldehyde (123 mg, 0.53 mmol, 2.0 equiv.), HOAc (23 μL, 0.40 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (96 mg, 0.45 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless solid (156 mg, 0.20 mmol, 74%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.19 (s, 1H), 7.84 (s, 1H), 7.52 (d, *J* = 7.9 Hz, 2H), 6.96 (d, *J* = 7.9 Hz, 2H), 6.29 (s, 2H), 6.02–6.00 (m, 1H), 5.57 (d, *J* = 8.0 Hz, 1H), 5.41–5.32 (m, 1H), 4.87 (dd, *J* = 6.5, 3.3 Hz, 1H), 4.36–4.27 (m, 1H), 4.16 (d, *J* = 7.3 Hz, 1H), 3.59 (d, *J* = 13.7 Hz, 1H), 3.39 (d, *J* = 13.7 Hz, 1H), 2.83–2.68 (m, 1H), 2.65–2.54 (m, 2H), 2.46 (d, *J* = 7.0 Hz, 1H), 2.03–1.85 (m, 1H), 1.81–1.66 (m, 1H), 1.57 (s, 3H), 1.43–1.30 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.9, 155.5, 153.0, 149.1, 140.0, 138.3, 137.3, 130.9, 120.4, 114.4, 92.6, 90.8, 85.5, 83.9, 83.5, 81.8, 79.5, 58.5, 55.8, 52.8, 50.6, 29.5, 28.5, 28.0, 27.2, 25.5. FT-IR: v / cm^{-1} = 2979, 1704, 1640, 1482, 1367, 1248, 1210, 1153, 1075, 1007, 907, 870, 798, 727. $[\alpha]_D^{20} = -26$ (10 mg/mL; MeOH). mp: 76–79 °C. ESI-MS: *m/z* calculated for $[C_{33}H_{46}IN_7O_7+H]^+$ ($[M+H]^+$) = 780.3, found: 780.3.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(4-(trifluoromethyl)benzyl)amino)-2-((_tert_-butoxycarbonyl)amino)-butanoate (21)
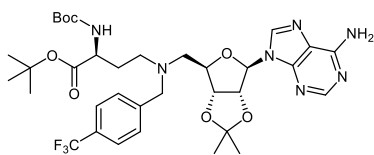
The compound was prepared from **3** (150 mg, 0.27 mmol, 1.0 equiv.), 4-(trifluoromethyl)benzaldehyde (73 μL, 0.53 mmol, 2.0 equiv.), HOAc (23 μL, 0.40 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (96 mg, 0.45 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless oil (119 mg, 0.16 mmol, 59%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.18 (s, 1H), 7.85 (s, 1H), 7.47 (d, *J* = 7.9 Hz, 2H), 7.35 (d, *J* = 7.9 Hz, 2H, H-26), 6.32 (s, 2H), 6.02 (s, 1H, H-9), 5.52 (d, *J* = 7.9 Hz, 1H), 5.43–5.34 (m, 1H), 4.91 (dd, *J* = 6.4, 3.5 Hz, 1H), 4.36–4.28 (m, 1H), 4.21–4.10 (m, 1H), 3.70 (d, *J* = 14.0 Hz, 1H), 3.52 (d, *J* = 14.0 Hz, 1H), 2.82 (dd, *J* = 14.0, 6.0 Hz, 1H), 2.72–2.51 (m, 2H), 2.50–2.37 (m, 1H), 2.02–1.84 (m, 1H), 1.83–1.64 (m, 1H), 1.57 (s, 3H, H-16), 1.42–1.31 (m, 21H).

^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.9, 155.5, 152.9, 149.1, 143.1, 140.1, 129.4 (q, J = 32.4 Hz), 129.1, 125.2, 124.3 (q, J = 271.7 Hz), 120.4, 114.5, 90.8, 85.5, 83.9, 83.5, 81.8, 79.6, 58.7, 55.9, 52.8, 50.6, 29.6, 28.4, 28.0, 27.2, 25.5. FT-IR: ν / cm^{-1} = 2980, 1705, 1642, 1367, 1324, 1248, 1211, 1154, 1125, 1065, 1018, 908, 847, 799, 728. $[\alpha]_D^{20}$ = −18 (10 mg/mL; MeOH). ESI-MS: m/z calculated for $[C_{34}H_{46}F_3N_7O_7+H]^+$ ($[M+H]^+$) = 722.4, found: 722.3.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(4-methoxybenzyl)amino)-2-((_tert_-butoxycarbonyl)amino)butanoate (22)
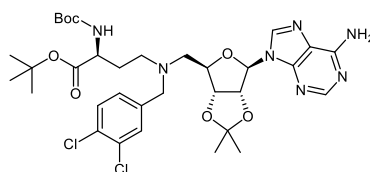

The compound was prepared from **3** (200 mg, 0.35 mmol, 1.0 equiv.), 4-methoxybenzaldehyde (86 μL, 0.71 mmol, 2.0 equiv.), HOAc (30 μL, 0.53 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv.) according to the general procedure A to afford the desired product as a colorless oil (105 mg, 0.15 mmol, 43%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.21 (s, 1H), 7.82 (s, 1H), 7.12 (d, J = 8.5 Hz, 2H), 6.74 (d, J = 8.5 Hz, 2H), 6.39–6.28 (m, 2H), 6.01 (d, J = 2.2 Hz, 1H), 5.71 (d, J = 7.8 Hz, 1H), 5.35–5.28 (m, 1H), 4.82 (dd, J = 6.4, 3.4 Hz, 1H), 4.36–4.27 (m, 1H), 4.18–4.11 (m, 1H), 3.74 (s, 3H), 3.61 (d, J = 13.3 Hz, 1H), 3.37 (d, J = 13.3 Hz, 1H), 2.73 (dd, J = 13.6, 6.7 Hz, 1H), 2.62–2.42 (m, 3H), 2.03–1.86 (m, 1H), 1.83–1.69 (m, 1H), 1.57 (s, 3H), 1.44–1.32 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 158.8, 155.9, 155.5, 153.1, 149.2, 139.8, 130.3, 120.3, 114.3, 113.6, 90.8, 85.3, 83.9, 83.5, 81.6, 79.4, 58.3, 55.5, 55.2, 53.0, 50.6, 29.2, 28.4, 28.0, 27.2, 25.5. FT-IR: ν / cm^{-1} = 3326, 2978, 1705, 1643, 1598, 1511, 1366, 1299, 1245, 1210, 1152, 1075, 909, 848, 799, 730. $[\alpha]_D^{20}$ = −20 (10 mg/mL; MeOH). ESI-MS: m/z calculated for $[C_{34}H_{49}N_7O_8+H]^+$ ($[M+H]^+$) = 684.4, found: 684.4.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(3-chlorobenzyl)amino)-2-((_tert_-butoxycarbonyl)amino)butanoate (23)
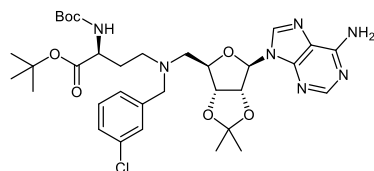
The compound was prepared from **3** (170 mg, 0.30 mmol, 1.0 equiv.), 3-chlorobenzaldehyde (68 μL, 0.60 mmol, 2.0 equiv.), HOAc (26 μL, 0.45 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (109 mg, 0.51 mmol, 1.7 equiv.) according to the general procedure A to afford the desired product as a colorless solid (139 mg, 0.20 mmol, 67%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.17 (s, 1H), 7.83 (s,), 7.22 (d, J = 1.7 Hz, 1H), 7.16–7.03 (m, 3H), 6.50 (s, 2H), 6.00 (d, J = 2.1 Hz, 1H), 5.60 (d, J = 8.1 Hz, 1H), 5.43–5.30 (m, 1H),

4.87 (dd, J = 6.4, 3.5 Hz, 1H4.35–4.26 (m, 1H), 4.20–4.07 (m, 1H), 3.60 (d, J = 13.7 Hz, 1H), 3.41 (d, J = 13.7 Hz, 1H), 2.76 (dd, J = 13.7, 6.3 Hz, 1H), 2.61 (dd, J = 13.1, 6.7 Hz, 2H), 2.53–2.39 (m, 1H), 2.01–1.83 (m, 1H), 1.81–1.65 (m, 1H), 1.56 (s, 3H), 1.44–1.29 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.9, 155.5, 152.9, 149.1, 134.1, 129.4, 128.9, 127.3, 127.0, 120.3, 114.4, 90.8, 85.4, 83.8, 83.4, 81.7, 79.5, 58.5, 55.8, 52.8, 50.7, 29.6, 28.4, 27.9, 27.1, 25.4. FT-IR: ν / cm^{-1} = 2980, 1704, 1641, 1597, 1475, 1426, 1367, 1330, 1210, 1153, 1075, 907, 870, 781, 727, 683. $[\alpha]_D^{20}$ = −15 (10 mg/mL; MeOH). mp: 77–80 °C. ESI-MS: m/z calculated for $[C_{33}H_{46}ClN_7O_7+H]^+$ ($[M+H]^+$) = 688.3, found: 688.3.

***tert*-Butyl (*S*)-4-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*]-[1,3]dioxol-4-yl)methyl)(3-((*tert*-butoxycarbonyl)amino)benzyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate (24)**
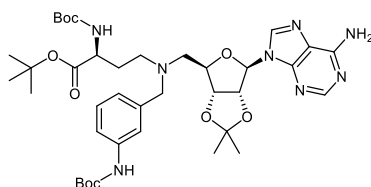
The compound was prepared from **3** (150 mg, 0.27 mmol, 1.0 equiv.), *tert*-butyl(3-formyl phenyl)carbamate (118 mg, 0.53 mmol, 2.0 equiv.), HOAc (23 μL, 0.40 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (96 mg, 0.45 mmol, 1.7 equiv.) according to the general procedure A to afford the desired product as a colorless oil (78 mg, 0.10 mmol, 37%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.23 (s, 1H), 7.88 (s, 1H), 7.34 (d, J = 8.2 Hz, 1H), 7.23 (s, 1H), 7.12 (*pseudo*-t, J ≈ 7.8 Hz, 1H), 7.02 (s, 1H), 6.86 (d, J = 7.5 Hz, 1H), 6.19 (s, 2H), 6.03 (d, J = 2.1 Hz, 1H), 5.57 (d, J = 8.2 Hz, 1H), 5.34 (d, J = 5.9 Hz, 1H), 4.88 (d, J = 5.9 Hz, 1H), 4.41–4.31 (m, 1H), 4.24–4.12 (m, 1H), 3.63 (d, J = 13.7 Hz, 1H), 3.49 (d, J = 13.7 Hz, 1H), 2.81–2.72 (m, 1H), 2.69–2.47 (m, 3H), 2.02–1.84 (m, 1H), 1.83–1.73 (m, 1H), 1.58 (s, 3H), 1.48 (s, 9H), 1.43–1.31 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.7, 155.6, 153.0, 149.2, 140.0, 138.7, 128.9, 123.5, 120.3, 119.3, 117.7, 114.5, 90.8, 85.2, 84.0, 83.4, 81.8, 80.4, 79.6, 59.1, 55.9, 52.9, 50.7, 29.4, 28.5, 28.0, 27.2, 25.5. FT-IR: ν / cm^{-1} = 2978, 1702, 1642, 1595, 1477, 1366, 1300, 1238, 1153, 1067, 909, 870, 782, 729. $[\alpha]_D^{20}$ = −2 (10 mg/mL; CHCl$_3$). ESI-MS: m/z calculated for $[C_{38}H_{56}N_8O_9+H]^+$ ($[M+H]^+$) = 769.4, found: 769.4.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(3-(trifluoromethyl)benzyl)amino)-2-((_tert_-butoxycarbonyl)amino)-butanoate (25)
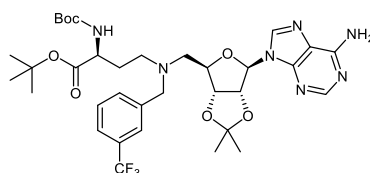
The compound was prepared from **3** (206 mg, 0.37 mmol, 1.0 equiv.), 3-(trifluoromethyl) benzaldehyde (98 μL, 0.73 mmol, 2.0 equiv.), HOAc (31 μL, 0.55 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (132 mg, 0.62 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless oil (190 mg, 0.26 mmol, 70%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.14 (s, 1H), 7.81 (s, 1H), 7.52 (s, 1H), 7.44–7.38 (m, 2H), 7.30 (d, _J_ = 7.6 Hz, 1H), 6.55 (s, 2H), 6.00 (d, _J_ = 2.1 Hz, 1H), 5.57 (d, _J_ = 8.1 Hz, 1H), 5.40–5.32 (m, 1H), 4.89 (dd, _J_ = 6.5, 3.6 Hz, 1H), 4.30 (_pseudo_-td, _J_ ≈ 6.6, 3.5 Hz, 1H), 4.18–4.09 (m, 1H), 3.68 (d, _J_ = 14.0 Hz, 1H), 3.49 (d, _J_ = 14.0 Hz, 1H), 2.78 (dd, _J_ = 13.5, 6.6 Hz, 1H), 2.67–2.42 (m, 3H), 1.99–1.85 (m, 1H), 1.81–1.65 (m, 1H), 1.55 (s, 3H), 1.39–1.27 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 156.0, 155.5, 152.9, 149.0, 140.1, 139.9, 132.1, 130.5 (q, _J_ = 32.0 Hz), 128.6, 125.5 (d, _J_ = 4.0 Hz), 124.2 (q, _J_ = 272.3 Hz), 123.91 (d, _J_ = 4.0 Hz), 120.3, 114.4, 90.6, 85.4, 83.8, 83.4, 81.7, 79.5, 58.6, 55.8, 52.7, 50.7, 29.6, 28.3, 27.9, 27.1, 25.3. FT-IR: ν / cm^{-1} = 2980, 1705, 1367, 1328, 1248, 1200, 1154, 1126, 1093, 1072, 907, 870, 798, 728, 703. $[\alpha]_D^{20}$ = −11 (10 mg/mL; MeOH). ESI-MS: _m/z_ calculated for [C$_{34}$H$_{46}$F$_3$N$_7$O$_7$+H]$^+$ ([M+H]$^+$) = 722.4, found: 722.3.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(3,4-dichlorobenzyl)amino)-2-((_tert_-butoxycarbonyl)amino)butanoate (26)

The compound was prepared from **3** (200 mg, 0.35 mmol, 1.0 equiv.), 3,4-dichlorobenzaldehyde (124 mg, 0.71 mmol, 2.0 equiv.), HOAc (30 μL, 0.53 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv.) according to general procedure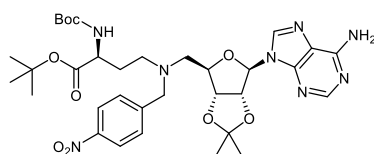 A to afford the desired product as a colorless oil (187 mg, 0.26 mmol, 74%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.16 (s, 1H), 7.82 (s, 1H), 7.31 (d, _J_ = 2.0 Hz, 1H), 7.22 (d, _J_ = 8.2 Hz, 1H), 7.02 (dd, _J_ = 8.2, 2.0 Hz, 1H), 6.42 (s, 2H), 5.99 (d, _J_ = 2.0 Hz, 1H), 5.53 (d, _J_ = 8.1 Hz, 1H), 5.42–5.32 (m, 1H), 4.89 (dd, _J_ = 6.5, 3.7 Hz, 1H), 4.29 (_pseudo_-td, _J_ ≈ 6.5, 3.5 Hz, 1H), 4.21–4.07 (m, 1H), 3.55 (d, _J_ = 14.0 Hz, 1H), 3.39 (d, _J_ = 14.1 Hz, 1H), 2.76 (dd, _J_ = 13.7, 6.1 Hz, 1H), 2.69–2.40 (m, 3H), 2.00–1.82 (m, 1H), 1.82–1.63 (m, 1H), 1.56 (s, 3H), 1.42–1.32 (m, 21H).

109

^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.9, 155.5, 152.9, 149.0, 140.0, 139.4, 132.2, 130.9, 130.6, 130.1, 128.1, 120.3, 114.4, 90.7, 85.5, 83.8, 83.4, 81.8, 79.6, 58.0, 55.8, 52.7, 50.6, 29.7, 28.4, 28.0, 27.2, 25.4. FT-IR: ν / cm^{-1} = 2980, 1705, 1640, 1597, 1472, 1367, 1329, 1248, 1209, 1153, 1075, 1030, 907, 870, 799, 727. $[\alpha]_D^{20}$ = −12 (10 mg/mL; MeOH). ESI-MS: m/z calculated for [C$_{33}$H$_{45}$Cl$_2$N$_7$O$_7$+H]$^+$ ([M+H]$^+$) = 722.3, found: 722.3.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(4-chloro-3-(trifluoromethyl)benzyl)amino)-2-((_tert_-butoxycarbonyl)amino)butanoate (27)
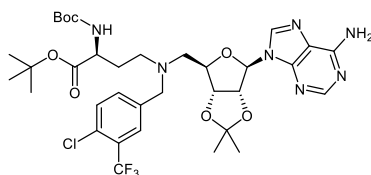
The compound was prepared from **3** (150 mg, 0.27 mmol, 1.0 equiv.), 4-chloro-3(trifluoromethyl) benzaldehyde (77 μL, 0.53 mmol, 2.0 equiv.), HOAc (23 μL, 0.40 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (96 mg, 0.45 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless oil (166 mg, 0.22 mmol, 81%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.16 (s, 1H), 7.84 (s, 1H), 7.59 (d, _J_ = 1.9 Hz, 1H), 7.39–7.27 (m, 2H), 6.48 (s, 2H), 6.00 (d, _J_ = 2.0 Hz, 1H), 5.50–5.42 (m, 1H), 5.39–5.32 (m, 1H), 4.91 (dd, _J_ = 6.6, 3.7 Hz, 1H), 4.29 (_pseudo_-td, _J_ ≈ 6.6, 3.7 Hz, 1H), 4.19–4.07 (m, 1H), 3.63 (d, _J_ = 14.2 Hz, 1H), 3.48 (d, _J_ = 14.2 Hz, 1H), 2.80 (dd, _J_ = 13.7, 6.3 Hz, 1H), 2.70–2.40 (m, 3H), 1.97–1.86 (m, 1H), 1.80–1.65 (m, 1H), 1.56 (s, 3H), 1.46–1.34 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.7, 155.9, 155.5, 152.8, 149.1, 140.1, 138.5, 133.0, 131.3, 130.7, 128.3, 127.8 (q, _J_ = 7.2 Hz), 122.9 (q, _J_ = 273.3 Hz), 120.3, 114.6, 90.7, 85.5, 83.9, 83.5, 81.9, 79.7, 58.2, 55.9, 52.7, 50.7, 29.8, 28.4, 27.9, 27.2, 25.4. FT-IR: ν / cm^{-1} = 2980, 1705, 1367, 1316, 1258, 1205, 1143, 1110, 1094, 1035, 907, 869, 728, 662. $[\alpha]_D^{20}$ = −14 (10 mg/mL; MeOH). ESI-MS: m/z calculated for [C$_{34}$H$_{45}$ClF$_3$N$_7$O$_7$+H]$^+$ ([M+H]$^+$) = 756.3, found: 756.3.

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(3,5-dichlorobenzyl)amino)-2-((_tert_-butoxycarbonyl)amino)butanoate (28)
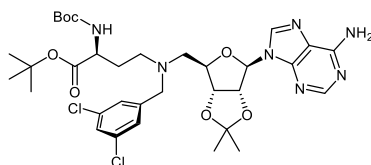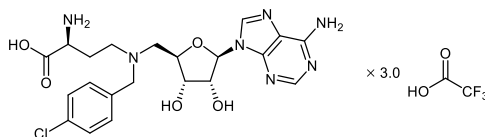

The compound was prepared from **3** (175 mg, 0.31 mmol, 1.0 equiv.), 3,5-dichlorobenzaldehyde (109 mg, 0.62 mmol, 2.0 equiv.), HOAc (27 μL, 0.47 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (112 mg, 0.53 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless oil (148 mg,

0.21 mmol, 68%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.19 (s, 1H), 7.83 (s, 1H), 7.17–7.09 (m, 3H), 6.31 (s, 2H), 6.00 (d, J = 2.1 Hz, 1H), 5.51–5.40 (m, 1H), 5.39–5.32 (m, 1H), 4.90 (dd, J = 6.5, 3.7 Hz, 1H), 4.29 (*pseudo*-td, $J \approx$ 6.6, 3.7 Hz, 1H), 4.19–4.09 (m, 1H), 3.56 (d, J = 14.2 Hz, 1H), 3.43 (d, J = 14.2 Hz), 2.84–2.42 (m, 4H), 2.02–1.85 (m, 1H), 1.81–1.66 (m, 1H), 1.58 (s, 3H), 1.44–1.32 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.7, 155.8, 155.5, 152.9, 149.1, 142.8, 140.1, 134.8, 127.3, 127.2, 120.4, 114.6, 90.7, 85.4, 83.8, 83.4, 81.9, 79.7, 58.2, 55.9, 52.7, 50.8, 29.8, 28.4, 28.0, 27.2, 25.4. FT-IR: v / cm^{-1} = 2979, 1704, 1643, 1595, 1367, 1248, 1209, 1152, 1075, 908, 850, 796, 729. $[\alpha]_D^{20}$ = −6 (10 mg/mL; MeOH). ESI-MS: m/z calculated for $[C_{33}H_{45}Cl_2N_7O_7+H]^+$ ($[M+H]^+$) = 722.3, found: 722.3.

tert-Butyl (*S*)-4-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*]-[1,3]dioxol-4-yl)methyl)(2,4-dichlorobenzyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate (29)
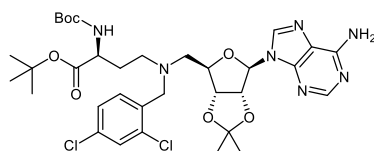
The compound was prepared from **3** (150 mg, 0.27 mmol, 1.0 equiv.), 2,4-dichlorobenzaldehyde (93 mg, 0.53 mmol, 2.0 equiv.), HOAc (23 μL, 0.40 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (96 mg, 0.45 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless oil (135 mg, 0.19 mmol, 70%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.14 (s, 1H), 7.80 (s, 1H), 7.29 (d, J = 8.3 Hz, 1H), 7.19 (d, J = 2.1 Hz, 1H), 6.99 (dd, J = 8.3, 2.1 Hz, 1H), 6.38 (s, 2H), 5.96 (d, J = 1.9 Hz, 1H), 5.46 (d, J = 8.1 Hz, 1H), 5.36 (d, J = 6.4 Hz, 1H, H-8), 4.85 (dd, J = 6.4, 3.4 Hz, 1H), 4.34–4.27 (m, 1H), 4.18–4.03 (m, 1H), 3.66–3.46 (m, 2H), 2.82–2.41 (m, 4H), 1.96–1.81 (m, 1H), 1.74–1.60 (m, 1H), 1.52 (s, 3H), 1.39–1.28 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.9, 155.5, 152.9, 149.0, 140.1, 135.1, 134.4, 133.1, 131.6, 129.0, 126.9, 120.3, 114.3, 90.9, 85.5, 83.8, 83.5, 81.8, 79.5, 56.1, 55.3, 52.7, 50.7, 29.6, 28.4, 28.0, 27.2, 25.4. FT-IR: v / cm^{-1} = 2979, 1704, 1642, 1474, 1367, 1248, 1209, 1152, 1094, 908, 865, 728. $[\alpha]_D^{20}$ = +4 (10 mg/mL; CHCl$_3$). ESI-MS: m/z calculated for $[C_{33}H_{45}Cl_2N_7O_7+H]^+$ ($[M+H]^+$) = 722.3, found: 722.3.

tert-Butyl (*S*)-4-(((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*]-[1,3]dioxol-4-yl)methyl)(4-nitrobenzyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate (30)

The compound was prepared from **3** (200 mg, 0.35 mmol, 1.0 equiv.), 4-nitrobenzaldehyde (107 mg, 0.71 mmol, 2.0 equiv.), HOAc (30 μL, 0.53 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (128 mg, 0.60 mmol, 1.7 equiv.) according to general procedure B to afford the desired product as a slightly yellowish solid

(220 mg, 0.31 mmol, 89%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.09 (s, 1H), 8.05–7.97 (m, 2H), 7.85 (s, 1H), 7.40–7.32 (m, 2H), 6.57 (s, 2H), 6.06–5.98 (m, 1H), 5.42–5.32 (m, 2H), 4.92 (dd, J = 6.5, 3.6 Hz, 1H), 4.41–4.32 (m, 1H), 4.25–4.12 (m, 1H), 3.74–3.55 (m, 2H), 2.85–2.69 (m, 2H), 2.68–2.47 (m, 2H), 2.05–1.87 (m, 1H), 1.81–1.64 (m, 1H), 1.58 (s, 3H), 1.45–1.32 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.7, 155.8, 155.5, 152.5, 148.8, 147.1, 146.8, 140.2, 129.3, 123.4, 119.9, 114.6, 90.9, 85.8, 84.0, 83.5, 82.0, 79.7, 58.5, 56.2, 52.7, 50.8, 29.9, 28.5, 28.1, 27.2, 25.5. FT-IR: ν / cm^{-1} = 3335, 3185, 2978, 2934, 1705, 1643, 1599, 1519, 1366, 1344, 1209, 1152, 1074, 847, 737. $[\alpha]_D^{20}$ = +18 (10 mg/mL; CHCl$_3$). mp: 81–83 °C. ESI-MS: m/z calculated for [C$_{33}$H$_{46}$N$_8$O$_9$+H]$^+$ ([M+H]$^+$) = 699.4, found: 699.3.

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d]-[1,3]dioxol-4-yl)methyl)(4-nitro-3-(trifluoromethyl)benzyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (31)
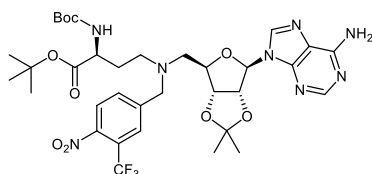
The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.), 4-nitro-3-(trifluoromethyl)benzaldehyde (134 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 μL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure B to afford the desired product as a slightly yellowish solid (242 mg, 0.32 mmol, 72%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.14 (s, 1H), 7.87 (s, 1H), 7.78–7.74 (m, 1H), 7.74–7.68 (m, 1H), 7.60–7.53 (m, 1H), 6.63 (s, 2H), 6.06–5.98 (m, 1H), 5.40–5.22 (m, 2H), 4.94 (dd, J = 6.5, 3.9 Hz, 1H), 4.37–4.27 (m, 1H), 4.23–4.10 (m, 1H), 3.77–3.58 (m, 2H), 2.92–2.68 (m, 2H), 2.68–2.47 (m, 2H), 2.02–1.86 (m, 1H), 1.82–1.65 (m, 1H), 1.57 (s, 3H), 1.40–1.33 (m, 21H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.6, 155.6, 155.5, 152.2, 148.9, 146.9, 145.7, 140.3, 132.6, 128.0 (q, J = 5.5 Hz), 125.2, 123.6 (q, J = 33.9 Hz), 121.8 (q, J = 273.6 Hz), 120.0, 114.8, 90.8, 85.7, 84.0, 83.4, 82.1, 79.9, 58.2, 56.2, 52.5, 51.0, 30.2, 28.4, 28.0, 27.2, 25.4. FT-IR: ν / cm^{-1} = 3338, 3186, 2979, 2936, 1708, 1644, 1598, 1539, 1366, 1310, 1250, 1151, 1095, 846, 738. $[\alpha]_D^{20}$ = +16 (10 mg/mL; CHCl$_3$). mp: 77–79 °C. ESI-MS: m/z calculated for [C$_{34}$H$_{45}$F$_3$N$_8$O$_9$+H]$^+$ ([M+H]$^+$) = 767.3, found: 767.2.

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-chlorobenzyl)amino)butanoic acid, trifluoroacetate salt (32)
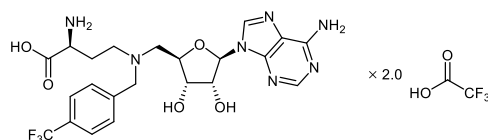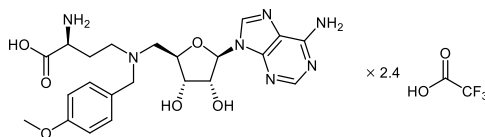

The compound was prepared from **18** (132 mg, 0.19 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (158 mg, 0.19 mmol, 99%, 3.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.22 (s, 1H), 8.07 (s, 1H), 7.23 (d, J = 8.3 Hz, 2H), 7.09 (d, J = 8.3 Hz, 2H), 5.92 (d, J = 3.4 Hz, 1H), 4.41 (dd, J = 5.2, 3.4 Hz, 1H), 4.35–4.20 (m, 3H), 4.13 (d, J = 13.5 Hz, 1H), 3.71 (dd, J = 8.8, 4.1 Hz, 1H), 3.49–3.17 (m, 4H), 2.29–2.13 (m, 1H), 2.11–1.94 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.7, 152.9, 149.5, 146.7, 144.2, 136.9, 133.8, 130.2, 117.9, 114.1, 92.1, 80.2, 74.8, 73.5, 58.1, 56.0, 53.4, 53.2, 26.4. FT-IR: ν / cm^{-1} = 2476, 2240, 2072, 1666, 1496, 1416, 1200, 1120, 1093, 1017, 973, 834, 799, 722. $[\alpha]_D^{20}$ = +16 (10 mg/mL; MeOH). mp: 74–77 °C. ESI-MS: m/z calculated for $[C_{21}H_{26}ClN_7O_5+H]^+$ ($[M+H]^+$) = 492.2, found: 492.2. Purity: 95% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.44 min).

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-bromobenzyl)amino)butanoic acid, trifluoroacetate salt (33)
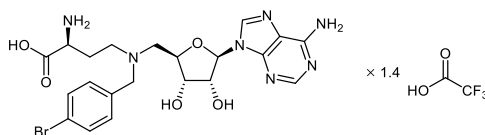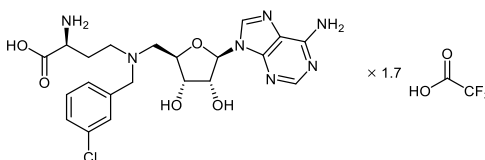

The compound was prepared from **19** (126 mg, 0.17 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (119 mg, 0.17 mmol, 99%, 1.4 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.25 (s, 1H), 8.13 (s, 1H), 7.31 (d, J = 8.2 Hz, 2H), 7.21 (d, J = 8.2 Hz, 2H), 5.98 (d, J = 3.2 Hz, 1H), 4.49 (*pseudo*-t, $J \approx$ 4.0 Hz, 1H), 4.42–4.27 (m, 3H), 4.14 (d, J = 13.4 Hz, 1H), 3.71 (dd, J = 8.9, 4.0 Hz, 1H), 3.55–3.27 (m, 4H), 2.36–2.14 (m, 1H), 2.13–1.94 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 173.0, 153.1, 149.5, 147.0, 144.1, 133.9, 133.2, 130.5, 125.0, 120.9, 92.2, 80.2, 74.8, 73.5, 58.3, 55.8, 53.3, 26.4. FT-IR: ν / cm^{-1} = 3105, 1666, 1434, 1189, 1131, 1074, 1014, 841, 799, 723. $[\alpha]_D^{20}$ = +8 (10 mg/mL; MeOH). mp: 100–103 °C. ESI-MS: m/z calculated for $[C_{21}H_{26}BrN_7O_5+H]^+$ ($[M+H]^+$) = 536.1, found: 536.1. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 30:70 + 0.1% HCOOH, t_R = 1.99 min).

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-iodobenzyl)amino)butanoic acid, trifluoroacetate salt (34)
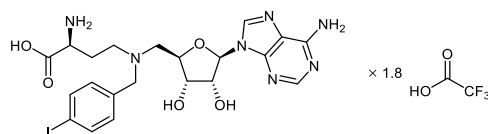
The compound was prepared from **20** (156 mg, 0.20 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (158 mg, 0.20 mmol, 99%, 1.8 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.06 (s, 1H), 7.95 (s, 1H), 7.29 (d, J = 8.2 Hz, 2H), 6.85 (d, J = 8.2 Hz, 2H), 5.76 (d, J = 3.0 Hz, 1H), 4.23 (dd, J = 4.9, 3.0 Hz, 1H), 4.18–4.06 (m, 3H), 3.98 (d, J = 13.4 Hz, 1H), 3.62 (dd, J = 8.1, 4.8 Hz, 1H), 3.38–3.05 (m, 4H), 2.18–2.02 (m, 1H), 2.00–1.89 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.8, 152.3, 149.4, 145.7, 144.4, 139.4, 134.0, 130.3, 120.9, 96.9, 92.3, 79.8, 74.8, 73.4, 58.5, 55.8, 52.5, 52.4, 26.1. FT-IR: ν / cm^{-1} = 3095, 2484, 2075, 1662, 1506, 1426, 1181, 1132, 1062, 1009, 972, 835, 798, 722. $[\alpha]_D^{20}$ = +11 (10 mg/mL; MeOH). mp: 72–75 °C. ESI-MS: m/z calculated for $[C_{21}H_{26}IN_7O_5+H]^+$ ($[M+H]^+$) = 584.1, found: 584.1. Purity: >99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.76 min).

(S)-2-Amino-4-(((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-(trifluoromethyl)benzyl)amino)butanoic acid, trifluoroacetate salt (35)

The compound was prepared from **21** (94 mg, 0.18 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (98 mg, 0.13 mmol, 99%, 2.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.26 (s, 1H), 8.16 (s, 1H), 7.57–7.44 (m, 4H), 5.97 (d, J = 3.2 Hz, 1H), 4.52–4.19 (m, 5H), 3.79 (dd, J = 8.2, 4.6 Hz, 1H), 3.57–3.21 (m, 4H), 2.35–2.18 (m, 1H), 2.18–2.01 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.1, 152.7, 149.6, 146.3, 144.3, 136.0, 132.8, 132.4, 126.9 (q, J = 3.9 Hz), 125.2 (q, J = 270.8 Hz), 120.9, 92.2, 80.2, 74.8, 73.5, 58.3, 56.2, 53.0, 52.8, 26.4. FT-IR: ν / cm^{-1} = 3082, 1668, 1506, 1423, 1325, 1183, 1125, 1068, 1020, 834, 800, 722. $[\alpha]_D^{20}$ = +12 (10 mg/mL; MeOH). mp: 55–58 °C. ESI-MS: m/z calculated for $[C_{22}H_{26}F_3N_7O_5+H]^+$ ($[M+H]^+$) = 526.2, found: 526.2. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 4.36 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-methoxybenzyl)amino)butanoic acid, trifluoroacetate salt (36)
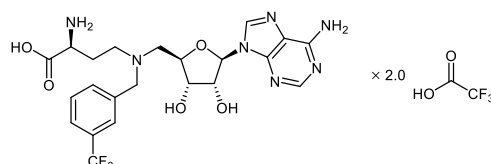

The compound was prepared from **22** (70 mg, 0.10 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (76 mg, 0.10 mmol, 99%, 2.4 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.26 (s, 1H), 8.10 (s, 1H), 7.18 (d, J = 8.4 Hz, 2H), 6.68 (d, J = 8.3 Hz, 2H), 5.95 (d, J = 3.3 Hz, 1H), 4.45 (dd, J = 5.1, 3.4 Hz, 1H), 4.40–4.07 (m, 4H), 3.77 (dd, J = 8.8, 4.2 Hz, 1H), 3.58 (s, 3H), 3.55–3.22 (m, 4H), 2.24 (dd, J = 14.8, 7.8 Hz, 1H), 2.09 (s, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.3, 162.3, 152.8, 149.6, 146.4, 144.3, 133.8, 122.1, 115.5, 92.0, 80.0, 74.9, 73.5, 58.5, 52.8, 52.6, 26.2. FT-IR: ν / cm^{-1} = 2972, 2469, 2073, 1663, 1613, 1516, 1426, 1308, 1255, 1181, 1130, 1029, 973, 834, 799, 721. $[\alpha]_D^{20}$ = +15 (10 mg/mL; MeOH). mp: 79–82 °C. ESI-MS: m/z calculated for [C$_{22}$H$_{29}$N$_7$O$_6$+H]$^+$ ([M+H]$^+$) = 488.2, found: 488.2. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.05 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(3-chlorobenzyl)amino)butanoic acid, trifluoroacetate salt (37)

The compound was prepared from **23** (91 mg, 0.13 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (89 mg, 0.13 mmol, 99%, 1.7 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.20 (s, 1H), 8.06 (s, 1H), 7.25 (d, J = 2.1 Hz, 1H), 7.21–7.05 (m, 3H), 5.92 (d, J = 3.1 Hz, 1H), 4.40 (dd, J = 4.9, 3.1 Hz, 1H), 4.35–4.21 (m, 3H), 4.14 (d, J = 13.4 Hz, 1H), 3.73 (dd, J = 8.6, 4.3 Hz, 1H), 3.47–3.21 (m, 4H), 2.29–2.14 (m, 1H), 2.09–1.95 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.7, 153.0, 149.5, 146.8, 144.1, 135.9, 133.6, 132.0, 131.7, 130.9, 130.5, 120.9, 92.2, 80.2, 74.8, 73.5, 58.3, 56.0, 53.3, 26.4. FT-IR: ν / cm^{-1} = 2478, 2073, 1662, 1506, 1427, 1185, 1132, 972, 834, 799, 785, 722, 684. $[\alpha]_D^{20}$ = +17 (10 mg/mL; MeOH). mp: 133–136 °C. ESI-MS: m/z calculated for [C$_{21}$H$_{26}$ClN$_7$O$_5$+H]$^+$ ([M+H]$^+$) = 492.2, found: 492.2. Purity: 96% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.04 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(3-aminobenzyl)amino)butanoic acid, trifluoroacetate salt (38)
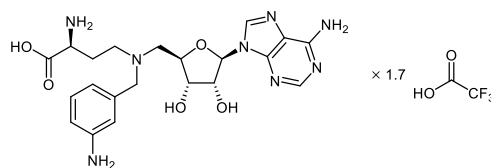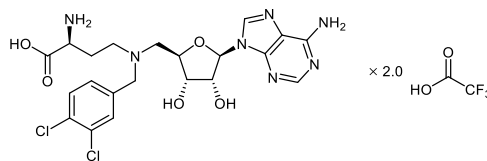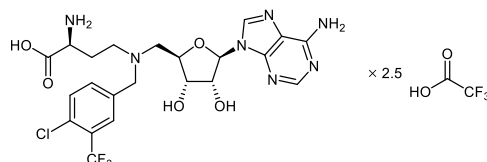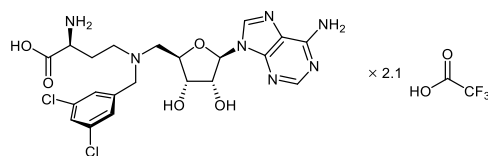

The compound was prepared from **24** (53 mg, 0.07 mmol) according to the general procedure C to afford the final product as a colorless trifluoroacetate salt (46 mg, 0.07 mmol, 99%, 1.7 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.31 (s, 1H), 8.23 (s, 1H), 7.28 (*pseudo*-t, $J \approx$ 7.8 Hz, 1H), 7.19 (d, J = 1.9 Hz, 1H), 7.16–7.04 (m, 2H), 6.03 (d, J = 3.4 Hz, 1H), 4.54 (dd, J = 5.2, 3.4 Hz, 1H), 4.45–4.30 (m, 3H), 4.21 (d, J = 13.4 Hz, 1H), 3.82 (dd, J = 8.6, 4.3 Hz, 1H), 3.59–3.31 (m, 4H), 2.42–2.22 (m, 1H), 2.19–2.05 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.2, 153.0, 149.7, 146.8, 144.1, 140.0, 133.4, 131.6, 127.9, 123.2, 122.2, 121.0, 92.2, 80.4, 74.8, 73.5, 58.8, 56.2, 53.1, 26.5. FT-IR: ν / cm^{-1} = 3014, 1668, 1508, 1428, 1195, 1132, 838, 799, 723. $[\alpha]_D^{20}$ = +11 (10 mg/mL; MeOH). mp: 68–71 °C. ESI-MS: *m/z* calculated for [C$_{21}$H$_{28}$N$_8$O$_5$+H]$^+$ ([M+H]$^+$) = 473.2, found: 473.2. Purity: 95% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.20 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(3-(trifluoromethyl)benzyl)amino)butanoic acid, trifluoroacetate salt (39)

The compound was prepared from **25** (173 mg, 0.24 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (185 mg, 0.24 mmol, 99%, 2.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.03 (s, 1H), 7.91 (s, 1H), 7.42 (s, 1H), 7.36 (d, J = 7.8 Hz, 1H), 7.27 (d, J = 7.8 Hz, 1H), 7.14 (*pseudo*-t, $J \approx$ 7.8 Hz, 1H), 5.75 (d, J = 3.0 Hz, 1H), 4.26–4.01 (m, 5H), 3.58 (dd, J = 8.2, 4.7 Hz, 1H), 3.33–3.01 (m, 4H), 2.15–1.98 (m, 1H), 1.99–1.82 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.6, 152.8, 149.5, 146.4, 144.2, 135.9, 133.0, 132.31 (q, J = 32.6 Hz), 131.0, 128.9 (q, J = 3.8 Hz), 127.5 (q, J = 3.4 Hz), 120.8, 118.0 (q, J = 292.0 Hz), 92.2, 80.1, 74.8, 73.4, 58.4, 56.1, 53.2, 53.0, 26.4. FT-IR: ν / cm^{-1} = 2462, 2075, 1663, 1506, 1427, 1330, 1182, 1125, 1076, 973, 834, 798, 722, 704, 665. $[\alpha]_D^{20}$ = +15 (10 mg/mL; MeOH). mp: 73–76 °C. ESI-MS: *m/z* calculated for [C$_{22}$H$_{26}$F$_3$N$_7$O$_5$+H]$^+$ ([M+H]$^+$) = 526.2, found: 526.2. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.23 min).

(*S*)-2-Amino-4-(((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(3,4-dichlorobenzyl)amino)butanoic acid, trifluoroacetate salt (40)

The compound was prepared from **26** (124 mg, 0.17 mmol) according to the general procedure C to afford the final product as a colorless trifluoroacetate salt (127 mg, 0.17 mmol, 99%, 2.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.35 (s, 1H), 8.23 (s, 1H), 7.55 (d, *J* = 1.9 Hz, 1H), 7.44–7.29 (m, 2H), 6.05 (d, *J* = 2.8 Hz, 1H), 4.50 (d, *J* = 3.1 Hz, 1H), 4.45–4.39 (m, 2H), 4.33 (d, *J* = 11.4 Hz, 2H), 3.95 (dd, *J* = 7.9, 5.0 Hz, 1H), 3.60–3.34 (m, 4H), 2.44–2.31 (m, 1H), 2.30–2.18 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.8, 152.5, 149.4, 145.9, 144.4, 134.8, 134.0, 133.9, 132.3, 132.1, 131.8, 121.1, 92.5, 80.0, 74.9, 73.4, 57.8, 55.8, 52.8, 52.7, 26.4. FT-IR: ν / cm^{-1} = 3094, 1673, 1507, 1474, 1427, 1200, 1134, 1036, 831, 799, 722. $[\alpha]_D^{20}$ = +17 (10 mg/mL; MeOH). mp: 87–90 °C. ESI-MS: *m/z* calculated for [C$_{21}$H$_{25}$Cl$_2$N$_7$O$_5$+H]$^+$ ([M+H]$^+$) = 526.1, found: 526.1. Purity: 95% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.84 min).

(*S*)-2-Amino-4-(((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-chloro-3-(trifluoromethyl)benzyl)amino)butanoic acid, trifluoroacetate salt (41)
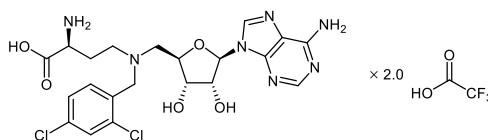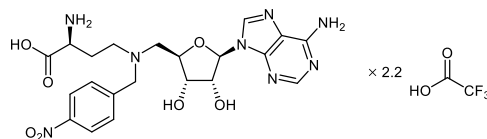

The compound was prepared from **27** (89 mg, 0.12 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (101 mg, 0.12 mmol, 99%, 2.5 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.20 (s, 1H), 8.10 (s, 1H), 7.67 (d, *J* = 2.1 Hz, 1H), 7.48 (dd, *J* = 8.2, 2.1 Hz, 1H), 7.32 (d, *J* = 8.2 Hz, 1H), 5.91 (d, *J* = 3.0 Hz, 1H), 4.37 (dd, *J* = 4.9, 3.0 Hz, 1H), 4.32–4.12 (m, 4H), 3.79 (dd, *J* = 7.9, 4.9 Hz, 1H), 3.43–3.18 (m, 4H), 2.28–2.12 (m, 1H), 2.11–1.95 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.9, 152.7, 149.5, 146.2, 144.3, 137.0, 134.3, 133.2, 132.1, 131.1, 129.6 (q, *J* = 31.6 Hz), 123.89 (q, *J* = 272.7 Hz), 121.0, 92.3, 80.2, 74.9, 73.4, 57.9, 56.0, 52.9, 52.7, 26.6. FT-IR: ν / cm^{-1} = 3080, 1663, 1506, 1484, 1427, 1322, 1267, 1179, 1129, 1039, 975, 834, 799, 722, 664. $[\alpha]_D^{20}$ = +15 (10 mg/mL; MeOH). mp: 79–82 °C. ESI-MS: *m/z* calculated for [C$_{22}$H$_{25}$ClF$_3$N$_7$O$_5$+H]$^+$ ([M+H]$^+$) = 560.1, found: 560.2. Purity: 96% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.42 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(3,5-dichlorobenzyl)amino)butanoic acid, trifluoroacetate salt (42)

The compound was prepared from **28** (119 mg, 0.16 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (123 mg, 0.16 mmol, 99%, 2.1 equiv. TFA). ^{1}H NMR (300 MHz, DMSO-d_6): δ / ppm = 8.48 (s, 1H), 8.33 (s, 1H), 7.47 (s, 1H), 7.41 (d, J = 1.9 Hz, 2H), 5.98 (d, J = 4.4 Hz, 1H), 4.54 (*pseudo*-t, $J \approx$ 4.4 Hz, 1H), 4.32–4.18 (m, 2H), 4.08–3.90 (m, 3H), 3.26–2.85 (m, 4H), 2.25–1.96 (m, 2H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ / ppm = 170.6, 153.1, 148.6, 148.2, 141.8, 138.9, 134.1, 128.7, 127.8, 119.3, 89.1, 80.0, 73.1, 72.0, 56.1, 55.1, 50.5, 49.9, 26.0. FT-IR: ν / cm^{-1} = 3090, 1674, 1572, 1508, 1431, 1200, 1136, 834, 800, 723, 684. $[\alpha]_D^{20}$ = +14 (10 mg/mL; MeOH). mp: 69–72 °C. ESI-MS: *m/z* calculated for $[C_{21}H_{25}Cl_2N_7O_5+H]^+$ ($[M+H]^+$) = 526.1, found: 526.1. Purity: 97% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.81 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(2,4-dichlorobenzyl)amino)butanoic acid, trifluoroacetate salt (43)

The compound was prepared from **29** (117 mg, 0.16 mmol) according to the general procedure C to afford the final product as a colorless resin (123 mg, 0.16 mmol, 99%, 2.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.24 (s, 1H), 8.17 (s, 1H), 7.40 (d, J = 8.3 Hz, 1H), 7.29 (d, J = 2.1 Hz, 1H), 7.10 (dd, J = 8.3, 2.1 Hz, 1H), 5.92 (d, J = 3.2 Hz, 1H), 4.44 (dd, J = 5.0, 3.3 Hz, 1H), 4.33–4.19 (m, 4H), 3.77 (dd, J = 7.8, 5.2 Hz, 1H), 3.48–3.21 (m, 4H), 2.28–2.14 (m, 1H), 2.14–2.01 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.8, 152.5, 149.5, 146.0, 144.4, 137.3, 137.2, 135.1, 130.9, 128.9, 121.0, 92.2, 80.7, 74.8, 73.5, 56.9, 56.0, 52.9, 26.7. FT-IR: ν / cm^{-1} = 3088, 1667, 1506, 1478, 1423, 1390, 1197, 1131, 1057, 833, 799, 722. $[\alpha]_D^{20}$ = +16 (10 mg/mL; MeOH). ESI-MS: *m/z* calculated for $[C_{21}H_{25}Cl_2N_7O_5+H]^+$ ($[M+H]^+$) = 526.1, found: 526.1. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.99 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-nitrobenzyl)amino)butanoic acid, trifluoroacetate salt (44)

The compound was prepared from **30** (135 mg, 0.19 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (145 mg, 0.19 mmol, 99%, 2.2 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.21 (s, 1H), 8.13 (s, 1H), 7.97–7.91 (m, 2H), 7.51 (d, J = 8.7 Hz, 2H), 5.93 (d, J = 3.0 Hz, 1H), 4.44–4.38 (m, 1H), 4.37–4.19 (m, 4H), 3.87–3.78 (m, 1H), 3.44–3.20 (m, 4H), 2.32–2.16 (m, 1H), 2.15–2.01 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.0, 152.6, 149.6, 149.5, 146.2, 144.3, 140.0, 133.0, 124.8, 120.9, 92.2, 80.5, 74.9, 73.5, 58.3, 56.4, 53.0, 52.9, 26.7. FT-IR: ν / cm^{-1} = 3085, 1666, 1609, 1524, 1421, 1348, 1323, 1182, 1125, 858, 830, 798, 741, 720, 699. $[\alpha]_D^{20}$ = +18 (10 mg/mL; MeOH). mp: 92 °C (decomposition). ESI-MS: m/z calculated for $[C_{21}H_{26}N_8O_7+H]^+$ ($[M+H]^+$) = 503.2, found: 503.0. Purity: >99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.34 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(4-nitro-3-(trifluoromethyl)benzyl)amino)butanoic acid, trifluoroacetate salt (45)
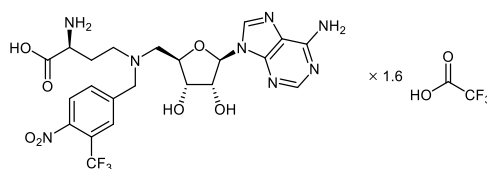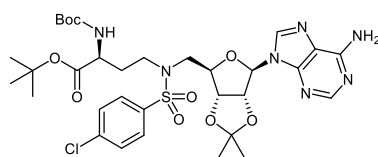

The compound was prepared from **31** (183 mg, 0.24 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (180 mg, 0.24 mmol, 99%, 1.6 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.22 (s, 1H), 8.14 (s, 1H), 7.84 (s, 1H), 7.78–7.69 (m, 2H), 5.93 (d, J = 3.1 Hz, 1H), 4.44–4.38 (m, 1H), 4.32–4.12 (m, 4H), 3.95–3.86 (m, 1H), 3.24–3.10 (m, 4H), 2.27–2.14 (m, 1H), 2.13–2.01 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.8, 152.6, 149.52, 148.9, 145.9, 144.2, 141.8–141.5 (m), 136.4, 130.6 (q, J = 6.4 Hz), 126.5, 124.2 (q, J = 34.1 Hz), 123.3 (q, J = 272.5 Hz), 120.9, 115.9, 92.1, 81.3, 74.9, 73.4, 58.1, 56.6, 52.7, 52.5, 27.5. FT-IR: ν / cm^{-1} = 3077, 1667, 1542, 1507, 1426, 1359, 1320, 1281, 1181, 1126, 1051, 901, 832, 798, 721. $[\alpha]_D^{20}$ = +13 (10 mg/mL; MeOH). mp: 71–73 °C. ESI-MS: m/z calculated for $[C_{22}H_{25}F_3N_8O_7+H]^+$ ($[M+H]^+$) = 571.2, found: 571.0. Purity: 96% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 4.10 min).

tert-Butyl (_S_)-4-((_N_-(((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-_d_][1,3]dioxol-4-yl)methyl)-4-chloro-3-(trifluoromethyl)phenyl)sulfonamido)-2-((_tert_-butoxy-carbonyl)amino)butanoate (59)
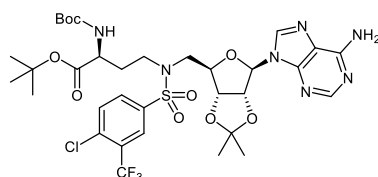
The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.), 4-chloro-3-(trifluoromethyl)benzenesulfonyl chloride (124 mg, 0.44 mmol, 1.0 equiv.), and NEt$_3$ (61 μL, 0.44 mmol, 1.0 equiv.) according to general procedure D to afford the desired product as a colorless solid (261 mg, 0.32 mmol, 73%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.32 (s, 1H), 8.02–7.96 (m, 1H), 7.83 (s, 1H), 7.74–7.66 (m, 1H), 7.52–7.41 (m, 1H), 6.26 (s, 1H), 5.99 (d, _J_ = 1.8 Hz, 1H), 5.40 (dd, _J_ = 6.4, 1.8 Hz, 1H), 5.27–5.16 (m, 1H), 5.09 (dd, _J_ = 6.4, 3.5 Hz, 1H), 4.38–4.28 (m, 1H), 4.09–3.96 (m, 1H), 3.77–3.63 (m, 1H), 3.49–3.35 (m, 1H), 3.33–3.06 (m, 2H), 2.14–1.98 (m, 1H), 1.90–1.74 (m, 1H), 1.56 (s, 3H), 1.44–1.38 (m, 18H), 1.35 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 155.8, 155.6, 152.9, 149.0, 140.3, 138.9, 137.1, 132.3, 131.4, 129.4 (q, _J_ = 32.6 Hz), 126.5 (q, _J_ = 5.2 Hz), 122.1 (q, _J_ = 273.0 Hz), 120.5, 114.8, 90.9, 85.675, 84.2, 82.7, 82.6, 80.1, 52.1, 50.1, 45.9, 32.1, 28.4, 28.0, 27.2, 25.4. FT-IR: ν / cm^{-1} = 3420, 3334, 3223, 2982, 2940, 1706, 1650, 1598, 1357, 1309, 1153, 1078, 871, 836, 749. $[\alpha]_D^{20}$ = +33 (10 mg/mL; CHCl$_3$). mp: 169 °C (decomposition). ESI-MS: _m/z_ calculated for [C$_{33}$H$_{43}$ClF$_3$N$_7$O$_9$S+H]$^+$ ([M+H]$^+$) = 806.3, found: 806.2.

tert-Butyl (_S_)-4-((_N_-(((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-_d_][1,3]dioxol-4-yl)methyl)-4-chlorophenyl)sulfonamido)-2-((_tert_-butoxycarbonyl)amino)-butanoate (60)

The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.), 4-chlorobenzenesulfonyl chloride (94 mg, 0.44 mmol, 1.0 equiv.), and NEt$_3$ (61 μL, 0.44 mmol, 1.0 equiv.) according to general procedure D to afford the desired product as a colorless solid (234 mg, 0.32 mmol, 72%). ^{1}H NMR (300 MHz, CDCl$_3$/CD$_3$OD = 1:1): δ / ppm = 8.15 (s, 1H), 7.84 (s, 1H), 7.49–7.43 (m, 2H), 7.30–7.20 (m, 2H), 5.91 (d, _J_ = 2.1 Hz, 1H), 5.26 (dd, _J_ = 6.4, 2.1 Hz, 1H), 4.96 (dd, _J_ = 6.5, 3.7 Hz, 1H), 4.28–4.13 (m, 1H), 3.83–3.65 (m, 1H), 3.60–3.44 (m, 1H), 3.25–2.92 (m, 3H), 2.00–1.81 (m, 1H), 1.78–1.60 (m, 1H), 1.45 (s, 3H), 1.32–1.26 (m, 18H), 1.24 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$/CD$_3$OD = 1:1): δ / ppm = 171.0, 155.7, 155.5, 152.6, 148.6, 140.1, 139.2, 137.4, 129.2, 128.4, 119.6, 114.7, 90.4, 85.2, 83.8, 82.4, 82.3,

79.9, 52.0, 50.2, 45.9, 31.4, 28.0, 27.6, 26.8, 25.0. FT-IR: ν / cm^{-1} = 3413, 3339, 3226, 2979, 2936, 1727, 1697, 1655, 1603, 1320, 1152, 1076, 874, 829, 773. $[\alpha]_D^{20}$ = +21 (10 mg/mL; CHCl$_3$/MeOH = 1:1). mp: 232–234 °C. ESI-MS: m/z calculated for [C$_{32}$H$_{44}$ClN$_7$O$_9$S+H]$^+$ ([M+H]$^+$) = 738.3, found: 738.3.

tert-Butyl (S)-4-((N-(((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-3,4-dichlorophenyl)sulfonamido)-2-((tert-butoxycarbonyl)amino)-butanoate (61)
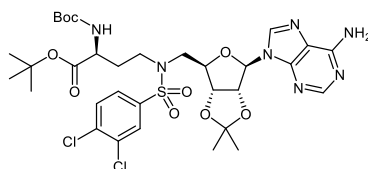
The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.), 3,4-dichlorobenzenesulfonyl chloride (124 mg, 0.44 mmol, 1.0 equiv.), and NEt$_3$ (61 μL, 0.44 mmol, 1.0 equiv.) according to general procedure D to afford the desired product as a colorless solid (177 mg, 0.23 mmol, 52%). ^{1}H NMR (300 MHz, CDCl$_3$/CD$_3$OD = 1:1): δ / ppm = 8.21 (s, 1H), 7.85 (s, 1H), 7.71–7.66 (m, 1H), 7.42–7.34 (m, 2H), 5.95 (d, J = 2.0 Hz, 1H), 5.30 (dd, J = 6.5, 2.1 Hz, 1H), 5.00 (dd, J = 6.4, 3.7 Hz, 1H), 4.30–4.19 (m, 1H), 3.98–3.83 (m, 1H), 3.67–3.52 (m, 1H), 3.36–2.98 (m, 3H), 2.06–1.88 (m, 1H), 1.84–1.66 (m, 1H), 1.50 (s, 3H), 1.36–1.32 (m, 18H), 1.28 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$/CD$_3$OD = 1:1): δ / ppm = 170.9, 155.7, 155.5, 152.7, 148.7, 140.2, 139.0, 137.5, 133.6, 131.0, 128.9, 126.1, 119.7, 114.8, 90.6, 85.2, 83.9, 82.4, 80.0, 52.0, 50.1, 45.9, 31.6, 28.2, 27.8, 26.9, 25.1. FT-IR: ν / cm^{-1} = 3340, 3225, 2979, 2936, 2571, 2500, 1702, 1617, 1323, 1218, 1155, 1074, 873, 824, 744. $[\alpha]_D^{20}$ = +32 (10 mg/mL; CHCl$_3$). mp: 203 °C (decomposition). ESI-MS: m/z calculated for [C$_{32}$H$_{43}$Cl$_2$N$_7$O$_9$S+H]$^+$ ([M+H]$^+$) = 772.2, found: 772.2.

tert-Butyl (S)-4-((N-(((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-2,4-dichlorophenyl)sulfonamido)-2-((tert-butoxycarbonyl)amino)-butanoate (62)
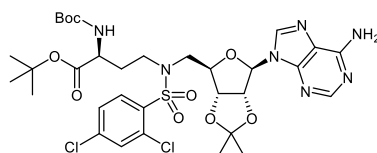
The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and 2,4-dichloro-nitrobenzenesulfonyl chloride (163 mg, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless solid (270 mg, 0.35 mmol, 80%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.30 (s, 1H), 7.81 (s, 1H), 7.73 (d, J = 8.5 Hz, 1H), 7.35 (d, J = 2.0 Hz, 1H), 7.07 (d, J = 8.5 Hz, 1H), 6.34 (s, 2H), 6.01–5.91 (m, 1H), 5.31 (dd, J = 6.4, 1.9 Hz, 1H), 5.17 (d, J = 8.1 Hz, 1H), 5.01 (dd, J

= 6.4, 3.7 Hz, 1H), 4.35–4.25 (m, 1H), 4.02 (s, 1H), 3.77–3.65 (m, 1H), 3.65–3.40 (m, 2H), 3.34–3.18 (m, 1H), 2.11–1.95 (m, 1H), 1.84–1.68 (m, 1H), 1.55 (s, 3H), 1.40 (d, J = 1.0 Hz, 18H), 1.33 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 155.6, 155.5, 152.5, 148.9, 140.3, 139.3, 136.3, 133.1, 132.6, 131.4, 127.0, 120.4, 114.7, 90.7, 85.5, 84.3, 82.7, 82.5, 79.9, 52.1, 49.6, 45.0, 31.5, 28.4, 28.0, 27.2, 25.4. FT-IR: ν / cm^{-1} = 2979, 2932, 1705, 1644, 1598, 1575, 1456, 1368, 1331, 1249, 1153, 1099, 869, 821, 752. $[\alpha]_D^{20}$ = +27 (10 mg/mL; CHCl$_3$). mp: 82–84 °C. ESI-MS: m/z calculated for $[C_{32}H_{43}Cl_2N_7O_9S+H]^+$ ($[M+H]^+$) = 772.2, found: 772.1.

tert-Butyl (_S_)-4-(_N_-(((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-_d_][1,3]dioxol-4-yl)methyl)phenylsulfonamido)-2-((_tert_-butoxycarbonyl)amino)butanoate (63)
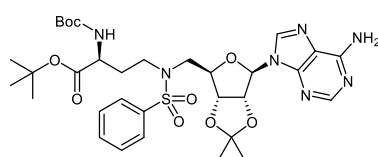
The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and benzenesulfonyl chloride (85 μL, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless solid (252 mg, 0.36 mmol, 81%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.34 (s, 1H), 7.88 (s, 1H), 7.69–7.61 (m, 2H), 7.52–7.44 (m, 1H), 7.42–7.34 (m, 2H), 6.32 (s, 2H), 6.01 (d, J = 2.0 Hz, 1H), 5.44 (dd, J = 6.4, 2.0 Hz, 1H), 5.25 (d, J = 8.0 Hz, 1H), 5.14 (dd, J = 6.4, 3.4 Hz, 1H), 4.39–4.29 (m, 1H), 4.07–3.95 (m, 1H), 3.70 (dd, J = 14.4, 6.6 Hz, 1H), 3.32–2.98 (m, 3H), 2.10–1.94 (m, 1H), 1.90–1.74 (m, 1H), 1.56 (s, 3H), 1.42–1.38 (m, 18H), 1.35 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.1, 155.7, 155.5, 152.7, 149.1, 140.4, 139.1, 132.8, 129.1, 127.2, 120.4, 114.7, 90.9, 85.7, 84.0, 82.8, 82.3, 79.9, 52.2, 50.5, 46.2, 32.2, 28.4, 28.0, 27.2, 25.4. FT-IR: ν / cm^{-1} = 3342, 3177, 2979, 2936, 1708, 1639, 1598, 1477, 1367, 1211, 1152, 1088, 869, 745, 690. $[\alpha]_D^{20}$ = +29 (10 mg/mL; CHCl$_3$). mp: 84–86 °C. ESI-MS: m/z calculated for $[C_{32}H_{45}N_7O_9S+H]^+$ ($[M+H]^+$) = 704.3, found: 704.2.

tert-Butyl (_S_)-4-((_N_-(((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-_d_][1,3]dioxol-4-yl)methyl)-4-nitrophenyl)sulfonamido)-2-((_tert_-butoxycarbonyl)amino)-butanoate (64)
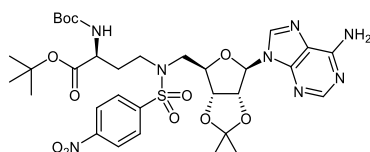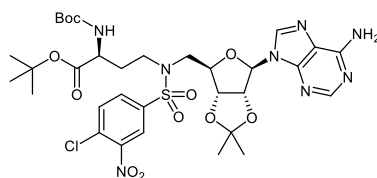

The compound was prepared from **3** (135 mg, 0.24 mmol, 1.0 equiv.) and 4-nitrobenzenesulfonyl chloride (80 mg, 0.36 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless solid (134 mg, 0.18 mmol, 75%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.32 (s, 1H), 8.17–8.06 (m, 2H), 7.89–7.79 (m, 1H), 7.77–7.69 (m, 2H), 6.38 (s, 2H), 5.96 (d, J = 1.9 Hz, 1H), 5.39 (dd, J =

6.4, 1.9 Hz, 1H), 5.28 (d, J = 7.9 Hz, 1H), 5.09 (dd, J = 6.4, 3.6 Hz, 1H), 4.37–4.27 (m, 1H), 4.11–3.97 (m, 1H), 3.74–3.60 (m, 1H), 3.54–3.09 (m, 3H), 2.15–2.00 (m, 1H), 1.91–1.77 (m, 1H), 1.56 (s, 3H), 1.43–1.40 (m, 18H), 1.34 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 155.7, 155.6, 152.7, 149.8, 148.9, 145.3, 140.3, 128.3, 124.1, 120.4, 114.8, 90.7, 85.3, 84.2, 82.7, 82.5, 80.0, 52.1, 49.9, 45.5, 32.0, 28.4, 28.0, 27.1, 25.4. FT-IR: ν / cm^{-1} = 3414, 3337, 3227, 2979, 2937, 1726, 1697, 1650, 1603, 1535, 1348, 1154, 1063, 871, 741. $[\alpha]_D^{20}$ = +31 (10 mg/mL; CHCl$_3$). mp: 209 °C (decomposition). ESI-MS: m/z calculated for $[C_{32}H_{44}N_8O_{11}S+H]^+$ ($[M+H]^+$) = 749.3, found: 749.2.

tert-Butyl (*S*)-4-((*N*-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-*d*][1,3]dioxol-4-yl)methyl)-4-chloro-3-nitrophenyl)sulfonamido)-2-((*tert*-butoxycarbonyl)-amino)butanoate (65)
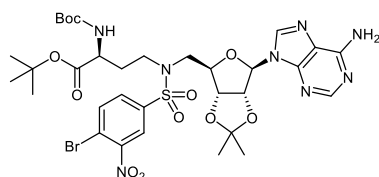

The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and 4-chloro-3-nitrobenzenesulfonyl chloride (170 mg, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless oil (149 mg, 0.19 mmol, 43%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.29 (s, 1H), 8.12 (d, J = 2.1 Hz, 1H), 7.82 (s, 1H), 7.66 (dd, J = 8.4, 2.2 Hz, 1H), 7.44 (d, J = 8.4 Hz, 1H), 6.31 (s, 2H), 5.99 (d, J = 1.8 Hz, 1H), 5.42–5.28 (m, 2H), 5.08 (dd, J = 6.4, 3.6 Hz, 1H), 4.40–4.27 (m, 1H), 4.12–3.91 (m, 1H), 3.74–3.47 (m, 2H), 3.40–3.11 (m, 2H), 2.20–2.00 (m, 1H), 1.93–1.75 (m, 1H), 1.55 (s, 3H), 1.45–1.36 (m, 18H), 1.34 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 156.0, 155.6, 153.2, 148.9, 147.6, 140.2, 140.0, 132.6, 131.4, 131.1, 124.4, 120.4, 114.7, 90.8, 85.3, 84.3, 82.7, 82.7, 80.1, 52.1, 49.8, 45.4, 31.9, 28.4, 28.0, 27.1, 25.3. FT-IR: ν / cm^{-1} = 3328, 2980, 2933, 1707, 1644, 1597, 1540, 1506, 1476, 1425, 1366, 1293, 1251, 1212, 1154, 1096, 1050, 887, 871, 754, 665. $[\alpha]_D^{20}$ = +50 (10 mg/mL; DCM). ESI-MS: m/z calculated for $[C_{32}H_{43}ClN_8O_{11}S+H]^+$ ($[M+H]^+$) = 783.3, found: 783.1.

***tert*-Butyl (*S*)-4-((*N*-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-*d*][1,3]dioxol-4-yl)methyl)-4-fluoro-3-nitrophenyl)sulfonamido)-2-((*tert*-butoxycarbonyl)-amino)butanoate (66)**
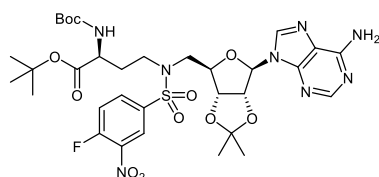
The compound was prepared from **3** (258 mg, 0.46 mmol, 1.0 equiv.), 4-fluoro-3-nitrobenzenesulfonyl chloride (110 mg, 0.46 mmol, 1.0 equiv.), and NEt$_3$ (63 μL, 0.46 mmol, 1.0 equiv.) according to general procedure D to afford the desired product as a yellowish oil (242 mg, 0.32 mmol, 70%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.36–8.23 (m, 2H), 7.89–7.75 (m, 2H), 7.17 (*pseudo*-t, $J \approx$ 9.4 Hz, 1H), 6.51 (s, 2H), 5.99 (d, J = 1.7 Hz, 1H), 5.37 (dd, J = 6.4, 1.8 Hz, 1H), 5.08 (dd, J = 6.4, 3.5 Hz, 1H), 4.38–4.26 (m, 1H), 4.10–3.98 (m, 1H), 3.74–3.46 (m, 2H), 3.38–3.06 (m, 2H), 2.15–2.02 (m, 1H), 1.93–1.75 (m, 1H), 1.54 (s, 3H), 1.42–1.37 (m, 19H), 1.33 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 157.3 (d, J = 268.0 Hz), 155.8, 155.5, 152.7, 148.8, 140.4, 137.0 (d, J = 7.5 Hz), 134.1 (d, J = 9.9 Hz), 125.7, 120.3, 119.3 (d, J = 22.0 Hz), 114.7, 90.7, 85.2, 85.2, 84.2, 82.6, 80.0, 52.0, 49.7, 45.3, 31.8, 28.3, 28.0, 27.1, 25.3. FT-IR: ν / cm^{-1} = 3340, 3195, 2980, 2936, 1708, 1642, 1605, 1540, 1478, 1350, 1239, 1155, 1072, 871, 755. $[\alpha]_D^{20}$ = +40 (10 mg/mL; CHCl$_3$). ESI-MS: *m/z* calculated for [C$_{32}$H$_{43}$FN$_8$O$_{11}$S+H]$^+$ ([M+H]$^+$) = 767.3, found: 767.3.

***tert*-Butyl (*S*)-4-((*N*-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-*d*][1,3]dioxol-4-yl)methyl)-4-bromo-3-nitrophenyl)sulfonamido)-2-((*tert*-butoxycarbonyl)-amino)butanoate (67)**

The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and 4-bromo-3-nitrobenzenesulfonyl chloride (200 mg, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless solid (252 mg, 0.30 mmol, 69%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.28 (s, 1H), 8.06 (d, J = 2.0 Hz, 1H), 7.84 (s, 1H), 7.66–7.49 (m, 2H), 6.55 (s, 2H), 5.99 (d, J = 1.8 Hz, 1H), 5.35 (dd, J = 6.4, 1.8 Hz, 1H), 5.06 (dd, J = 6.4, 3.5 Hz, 1H), 4.35–4.25 (m, 1H), 4.11–3.97 (m, 1H), 3.70–3.47 (m, 2H), 3.36–3.09 (m, 2H), 2.15–2.00 (m, 1H), 1.90–1.74 (m, 1H), 1.53 (s, 3H), 1.41–1.36 (m, 18H), 1.32 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 155.9, 155.5, 152.9, 149.5, 148.8, 140.6, 140.3, 135.8, 130.9, 124.2, 120.3, 119.1, 114.7, 90.7, 85.2, 84.2, 82.6, 82.6, 80.0,

52.0, 49.8, 45.4, 31.8, 28.3, 27.9, 27.1, 25.3. FT-IR: ν / cm^{-1} = 3342, 2977, 2934, 1705, 1637, 1590, 1540, 1353, 1250, 1151, 1093, 1032, 869, 775, 661. $[\alpha]_D^{20}$ = +40 (10 mg/mL; CHCl$_3$). mp: 84–86 °C. ESI-MS: m/z calculated for [C$_{32}$H$_{43}$BrN$_8$O$_{11}$S+H]$^+$ ([M+H]$^+$) = 827.2, found: 827.2.

tert-Butyl (S)-4-((N-(((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-3-nitrophenyl)sulfonamido)-2-((tert-butoxycarbonyl)amino)-butanoate (68)
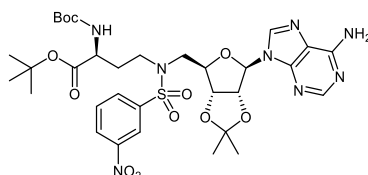
The compound was prepared from **3** (251 mg, 0.45 mmol, 1.0 equiv.), 3-nitrobenzenesulfonyl chloride (99 mg, 0.45 mmol, 1.0 equiv.), and NEt$_3$ (62 μL, 0.45 mmol, 1.0 equiv.) according to general procedure D to afford the desired product as a colorless solid (220 mg, 0.29 mmol, 64%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.46 (*pseudo*-t, $J \approx$ 2.0 Hz, 1H), 8.32–8.28 (m, 1H), 8.27–8.20 (m, 1H), 7.90–7.85 (m, 1H), 7.85–7.79 (m, 1H), 7.47 (*pseudo*-t, $J \approx$ 8.0 Hz, 1H), 6.42 (s, 2H), 5.98 (d, J = 1.8 Hz, 1H), 5.42–5.29 (m, 2H), 5.09 (dd, J = 6.4, 3.5 Hz, 1H), 4.39–4.27 (m, 1H), 4.13–3.91 (m, 1H), 3.75–3.61 (m, 1H), 3.56–3.42 (m, 1H), 3.40–3.03 (m, 2H), 2.17–2.01 (m, 1H), 1.94–1.78 (m, 1H), 1.54 (s, 3H), 1.43–1.37 (m, 18H), 1.33 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 155.9, 155.6, 153.0, 148.9, 148.1, 141.7, 140.2, 132.6, 130.2, 126.9, 122.3, 120.3, 114.7, 90.7, 85.2, 84.2, 82.6, 82.6, 80.0, 52.1, 49.9, 45.4, 31.9, 28.3, 28.0, 27.1, 25.3. FT-IR: ν / cm^{-1} = 3331, 3179, 2981, 2936, 1705, 1640, 1598, 1533, 1351, 1211, 1153, 1070, 908, 871, 727. $[\alpha]_D^{20}$ = +40 (10 mg/mL; CHCl$_3$). mp: 91–93 °C. ESI-MS: m/z calculated for [C$_{32}$H$_{44}$N$_8$O$_{11}$S+H]$^+$ ([M+H]$^+$) = 749.3, found: 749.3.

tert-Butyl (S)-4-((N-(((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-2-chloro-5-nitrophenyl)sulfonamido)-2-((tert-butoxycarbonyl)-amino)butanoate (69)
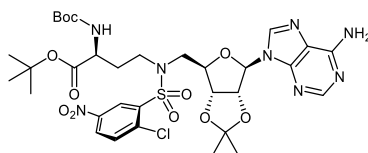
The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and 2-chloro-5-nitrobenzenesulfonyl chloride (170 mg, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a yellowish solid (175 mg, 0.22 mmol, 51%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.65–8.56 (m, 1H), 8.26 (s, 1H), 8.10–8.00 (m, 1H), 7.79–7.68 (m, 1H), 7.49–7.40 (m, 1H), 6.36 (s, 2H), 5.95–5.86 (m, 1H), 5.27 (dd, J = 6.4, 1.8 Hz, 1H), 5.21 (d, J = 7.7 Hz, 1H), 4.99 (dd, J =

6.4, 3.9 Hz, 1H), 4.36–4.26 (m, 1H), 4.06 (s, 1H), 3.84–3.53 (m, 4H), 3.36–3.22 (m, 1H), 2.21–2.05 (m, 1H), 1.91–1.74 (m, 1H), 1.55 (s, 3H), 1.45–1.37 (m, 18H), 1.32 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.8, 155.5, 155.5, 152.6, 148.7, 145.6, 140.2, 139.7, 139.0, 132.5, 127.3, 126.3, 120.2, 115.0, 90.4, 84.8, 84.2, 82.7, 82.5, 80.1, 52.2, 49.3, 44.8, 31.6, 28.4, 28.0, 27.2, 25.3. FT-IR: ν / cm^{-1} = 3344, 3104, 2980, 2936, 1706, 1639, 1600, 1528, 1347, 1242, 1153, 1075, 1040, 870, 739. $[\alpha]_D^{20}$ = +27 (10 mg/mL; CHCl$_3$). mp: 98–100 °C. ESI-MS: m/z calculated for [C$_{32}$H$_{43}$ClN$_8$O$_{11}$S+H]$^+$ ([M+H]$^+$) = 783.3, found: 783.2.
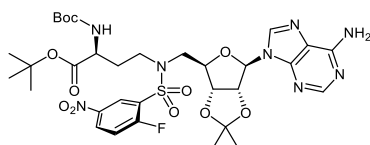
***tert*-Butyl (*S*)-4-((*N*-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-*d*][1,3]dioxol-4-yl)methyl)-2-fluoro-5-nitrophenyl)sulfonamido)-2-((*tert*-butoxycarbonyl)-amino)butanoate (70)**

The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and 2-fluoro-5-nitrobenzenesulfonyl chloride (159 mg, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless solid (277 mg, 0.36 mmol, 82%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.47–8.37 (m, 1H), 8.27 (s, 1H), 8.21–8.11 (m, 1H), 7.72 (s, 1H), 7.12 (*pseudo*-t, $J \approx$ 8.9 Hz, 1H), 6.39 (s, 2H), 5.88 (d, J = 2.0 Hz, 1H), 5.28 (dd, J = 6.4, 2.0 Hz, 1H), 5.01 (dd, J = 6.4, 3.9 Hz, 1H), 4.35–4.27 (m, 1H), 4.15–4.02 (m, 1H), 3.71 (d, J = 6.3 Hz, 2H), 3.58 (s, 1H), 3.30–3.16 (m, 1H), 2.19–2.05 (m, 1H), 1.97–1.81 (m, 1H), 1.54 (s, 3H), 1.44–1.38 (m, 18H), 1.31 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 162.0 (d, J = 265.2 Hz), 163.7, 160.2, 155.7, 155.6, 152.8, 148.7, 143.3, 140.0, 130.1 (d, J = 17.1 Hz), 129.5 (d, J = 10.6 Hz), 126.2, 120.1, 117.9 (d, J = 24.1 Hz), 114.9, 90.3, 84.6, 84.2, 82.5, 82.4, 80.0, 52.1, 49.4, 44.7, 31.8, 28.4, 28.0, 27.2, 25.3. FT-IR: ν / cm^{-1} = 3342, 2980, 2936, 1705, 1640, 1586, 1534, 1474, 1350, 1254, 1153, 1058, 870, 844, 744. $[\alpha]_D^{20}$ = +19 (10 mg/mL; CHCl$_3$). mp: 84 °C (decomposition). ESI-MS: m/z calculated for [C$_{32}$H$_{43}$FN$_8$O$_{11}$S+H]$^+$ ([M+H]$^+$) = 767.3, found: 767.2.

tert-Butyl (*S*)-4-((*N*-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-*d*][1,3]dioxol-4-yl)methyl)-3-chloro-4-nitrophenyl)sulfonamido)-2-((*tert*-butoxycarbonyl)-amino)butanoate (71)
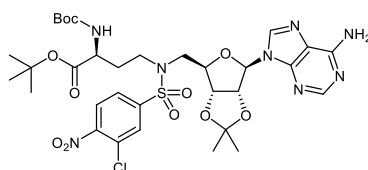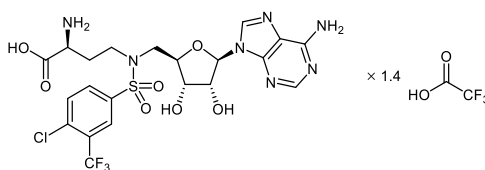

The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and 3-chloro-4-nitrobenzenesulfonyl chloride (170 mg, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless solid (249 mg, 0.32 mmol, 72%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.32 (s, 1H), 7.86–7.78 (m, 2H), 7.76–7.69 (m, 1H), 7.59–7.53 (m, 1H), 6.33 (s, 2H), 5.99 (d, J = 1.8 Hz, 1H), 5.39 (dd, J = 6.4, 1.8 Hz, 1H), 5.28 (d, J = 8.0 Hz, 1H), 5.09 (dd, J = 6.4, 3.5 Hz, 1H), 4.37–4.29 (m, 1H), 4.12–3.98 (m, 1H), 3.73–3.52 (m, 2H), 3.39–3.13 (m, 2H), 2.17–2.02 (m, 1H), 1.92–1.75 (m, 1H), 1.56 (s, 3H), 1.44–1.40 (m, 18H), 1.35 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.8, 155.7, 155.6, 152.8, 149.8, 148.9, 144.3, 140.4, 130.5, 128.0, 126.2, 126.0, 120.4, 114.8, 90.9, 85.4, 84.2, 82.7, 80.1, 52.1, 50.0, 45.7, 32.1, 28.4, 28.0, 27.2, 25.3. FT-IR: ν / cm^{-1} = 3331, 2980, 2934, 1705, 1639, 1594, 1536, 1366, 1211, 1153, 1096, 869, 797, 749, 672. $[\alpha]_D^{20}$ = +32 (10 mg/mL; CHCl$_3$). mp: 97–99 °C. ESI-MS: *m/z* calculated for $[C_{32}H_{43}ClN_8O_{11}S+H]^+$ ($[M+H]^+$) = 783.3, found: 783.2.

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-4-chloro-3-(trifluoromethyl)phenyl)sulfonamido)butanoic acid, trifluoroacetate salt (72)

The compound was prepared from **59** (92 mg, 0.11 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (88 mg, 0.11 mmol, 99%, 1.4 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.35 (s, 1H), 8.29 (s, 1H), 8.06 (d, J = 2.2 Hz, 1H), 7.98 (dd, J = 8.4, 2.2 Hz, 1H), 7.72–7.64 (m, 1H), 5.92 (d, J = 4.2 Hz, 1H), 4.58–4.52 (m, 1H), 4.25–4.18 (m, 1H), 4.17–4.09 (m, 1H), 3.94–3.85 (m, 1H), 3.72–3.52 (m, 2H), 3.50–3.29 (m, 2H), 2.34–2.19 (m, 1H), 2.12–1.97 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.4, 152.8, 149.8, 146.5, 143.9, 140.2, 138.0, 134.0, 133.3, 130.05 (q, J = 32.2 Hz), 127.6 (q, J = 5.4 Hz), 123.5 (q, J = 273.3 Hz), 120.8, 91.1, 83.6, 75.1, 73.0, 52.4, 51.5, 46.6, 31.2. FT-IR: ν / cm^{-1} = 3103, 2969, 1671, 1509, 1426, 1311, 1259, 1184, 1130, 1036, 968, 837, 798, 722, 660. $[\alpha]_D^{20}$ = +25 (10 mg/mL; MeOH). mp: 61 °C (decomposition). ESI-MS:

m/z calculated for $[C_{21}H_{23}ClF_3N_7O_7S+H]^+$ ($[M+H]^+$) = 610.1, found: 610.0. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 30:70 + 0.1% HCOOH, t_R = 3.02 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-4-chlorophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (73)
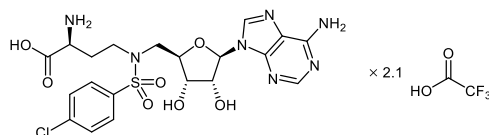
The compound was prepared from **60** (104 mg, 0.14 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (110 mg, 0.14 mmol, 99%, 2.1 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.31 (s, 1H), 8.24 (s, 1H), 7.71–7.64 (m, 2H), 7.41–7.34 (m, 2H), 5.87 (d, J = 4.3 Hz, 1H), 4.55–4.48 (m, 1H), 4.19–4.13 (m, 1H), 4.12–4.04 (m, 1H), 3.86–3.78 (m, 1H), 3.63–3.53 (m, 1H), 3.47–3.18 (m, 3H), 2.25–2.10 (m, 1H), 2.03–1.88 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.4, 152.5, 149.8, 146.0, 144.1, 140.4, 139.0, 130.6, 130.0, 120.8, 91.1, 84.0, 75.1, 73.0, 52.7, 51.6, 46.8, 31.2. FT-IR: ν / cm^{-1} = 3097, 2939, 1673, 1476, 1412, 1334, 1198, 1131, 1091, 1013, 972, 828, 798, 764, 721. $[\alpha]_D^{20}$ = +18 (10 mg/mL; MeOH). mp: 82–84 °C. ESI-MS: m/z calculated for $[C_{20}H_{24}ClN_7O_7S+H]^+$ ($[M+H]^+$) = 542.1, found: 542.0. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.70 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-3,4-dichlorophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (74)
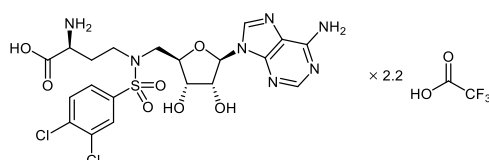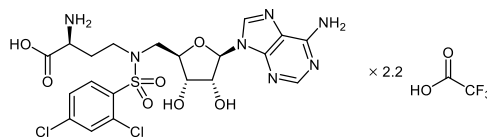

The compound was prepared from **61** (76 mg, 0.10 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (81 mg, 0.10 mmol, 99%, 2.2 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.35 (s, 1H), 8.29 (s, 1H), 7.89 (d, J = 2.1 Hz, 1H), 7.69–7.64 (m, 1H), 7.61–7.56 (m, 1H), 5.92 (d, J = 4.3 Hz, 1H), 4.62–4.54 (m, 1H), 4.26–4.20 (m, 1H), 4.18–4.09 (m, 1H), 3.93–3.86 (m, 1H), 3.70–3.52 (m, 2H), 3.50–3.28 (m, 2H), 2.33–2.19 (m, 1H), 2.12–1.97 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.5, 152.9, 149.8, 146.6, 143.9, 140.6, 138.5, 134.4, 132.6, 130.2, 128.0, 120.8, 91.1, 83.7, 75.0, 73.0, 52.4, 51.6, 46.6, 31.2. FT-IR: ν / cm^{-1} = 3096, 1670, 1508, 1457, 1426, 1372, 1337, 1188, 1132, 1033, 967, 825, 798, 722, 676. $[\alpha]_D^{20}$ = +18 (10 mg/mL; MeOH). mp: 60 °C (decomposition). ESI-MS: m/z calculated for $[C_{20}H_{23}Cl_2N_7O_7S+H]^+$ ($[M+H]^+$) = 576.1, found: 576.0. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 4.87 min).

(S)-2-Amino-4-((N-(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-2,4-dichlorophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (75)

The compound was prepared from **62** (159 mg, 0.21 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (169 mg, 0.21 mmol, 99%, 2.2 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.26–8.21 (m, 2H), 7.80–7.74 (m, 1H), 7.38 (d, J = 2.1 Hz, 1H), 7.13 (dd, J = 8.6, 2.1 Hz, 1H), 5.80 (d, J = 4.0 Hz, 1H), 4.46–4.40 (m, 1H), 4.15–4.09 (m, 1H), 4.09–4.00 (m, 1H), 3.83–3.74 (m, 1H), 3.71–3.49 (m, 3H), 3.46–3.33 (m, 1H), 2.26–2.10 (m, 1H), 2.02–1.86 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.3, 152.7, 149.7, 146.3, 144.0, 140.5, 137.5, 134.3, 133.9, 132.6, 128.4, 120.7, 91.1, 83.3, 75.1, 73.0, 51.6, 51.4, 45.9, 30.5. FT-IR: ν / cm^{-1} = 3087, 1668, 1574, 1506, 1427, 1373, 1328, 1185, 1130, 1038, 966, 821, 798, 721, 676. $[\alpha]_D^{20}$ = +30 (10 mg/mL; MeOH). mp: 83–85 °C. ESI-MS: m/z calculated for [C$_{20}$H$_{23}$Cl$_2$N$_7$O$_7$S+H]$^+$ ([M+H]$^+$) = 576.1, found: 576.0. Purity: >99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.75 min).

(S)-2-Amino-4-(N-(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)phenylsulfonamido)butanoic acid, trifluoroacetate salt (76)
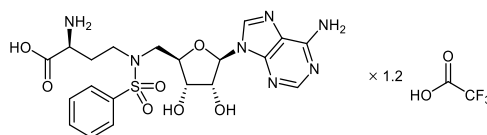
The compound was prepared from **63** (102 mg, 0.15 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (94 mg, 0.15 mmol, 99%, 1.2 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.33 (s, 1H), 8.24 (s, 1H), 7.76–7.70 (m, 2H), 7.55–7.47 (m, 1H), 7.46–7.38 (m, 2H), 5.90 (d, J = 4.4 Hz, 1H), 4.58–4.52 (m, 1H), 4.21–4.15 (m, 1H), 4.15–4.06 (m, 1H), 3.88–3.81 (m, 1H), 3.64–3.55 (m, 1H), 3.44–3.20 (m, 3H), 2.26–2.11 (m, 1H), 2.05–1.90 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.5, 152.5, 149.8, 146.1, 144.1, 140.0, 134.2, 130.5, 128.3, 120.8, 91.2, 84.2, 75.1, 73.1, 53.0, 51.6, 47.0, 31.4. FT-IR: ν / cm^{-1} = 3098, 1669, 1507, 1447, 1424, 1326, 1189, 1130, 1088, 966, 835, 798, 745, 721, 689. $[\alpha]_D^{20}$ = +24 (10 mg/mL; MeOH). mp: 78–80 °C. ESI-MS: m/z calculated for [C$_{20}$H$_{25}$N$_7$O$_7$S+H]$^+$ ([M+H]$^+$) = 508.2, found: 508.0. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.73 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-4-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (77)
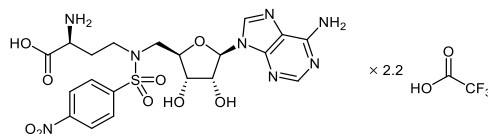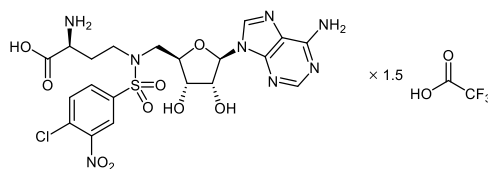

The compound was prepared from **64** (86 mg, 0.12 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (92 mg, 0.12 mmol, 99%, 2.2 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.28 (s, 1H), 8.23 (s, 1H), 8.17–8.11 (m, 2H), 7.93–7.87 (m, 2H), 5.82 (d, J = 4.2 Hz, 1H), 4.52–4.46 (m, 1H), 4.18–4.12 (m, 1H), 4.10–4.01 (m, 1H), 3.86–3.79 (m, 1H), 3.68–3.59 (m, 1H), 3.55–3.23 (m, 3H), 2.27–2.11 (m, 1H), 2.06–1.91 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.4, 152.5, 151.5, 149.8, 146.1, 144.0, 129.8, 125.4, 120.7, 91.0, 83.6, 75.1, 72.9, 52.4, 51.5, 46.6, 31.1. FT-IR: ν / cm^{-1} = 3106, 1666, 1530, 1424, 1350, 1313, 1131, 967, 855, 835, 798, 758, 742, 721, 684. $[\alpha]_D^{20}$ = +16 (10 mg/mL; MeOH). mp: 84 °C (decomposition). ESI-MS: m/z calculated for [C$_{20}$H$_{24}$N$_8$O$_9$S+H]$^+$ ([M+H]$^+$) = 553.2, found: 553.0. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.78 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-4-chloro-3-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (78)
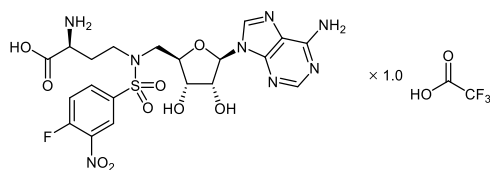

The compound was prepared from **65** (140 mg, 0.18 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (136 mg, 0.18 mmol, 99%, 1.5 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.32 (s, 1H), 8.28 (s, 1H), 8.23 (d, J = 2.1 Hz, 1H), 7.94 (dd, J = 8.5, 2.2 Hz, 1H), 7.68 (d, J = 8.5 Hz, 1H), 5.89 (d, J = 4.2 Hz, 1H), 4.56–4.50 (m, 1H), 4.23–4.17 (m, 1H), 4.15–4.07 (m, 1H), 3.92–3.85 (m, 1H), 3.71–3.53 (m, 2H), 3.52–3.29 (m, 2H), 2.33–2.18 (m, 1H), 2.12–1.96 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.3, 152.5, 149.8, 149.2, 146.1, 144.0, 141.0, 134.1, 132.7, 132.0, 125.5, 120.8, 91.1, 83.4, 75.1, 72.9, 52.2, 51.5, 46.4, 31.1. FT-IR: ν / cm^{-1} = 3093, 2944, 1673, 1538, 1426, 1347, 1170, 1133, 1049, 968, 888, 834, 797, 721, 664. $[\alpha]_D^{20}$ = +26 (10 mg/mL; MeOH). mp: 138 °C (decomposition). ESI-MS: m/z calculated for [C$_{20}$H$_{23}$ClN$_8$O$_9$S+H]$^+$ ([M+H]$^+$) = 587.1, found: 587.0. Purity: 96% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 4.76 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-4-fluoro-3-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (79)
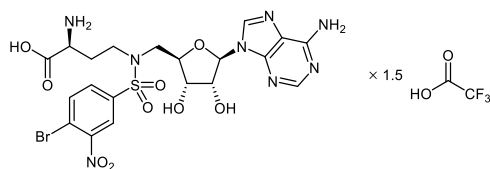

The compound was prepared from **66** (88 mg, 0.12 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (83 mg, 0.12 mmol, 99%, 1.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.32 (dd, J = 6.8, 2.3 Hz, 1H), 8.29–8.19 (m, 2H), 8.07–7.98 (m, 1H), 7.45–7.36 (m, 1H), 5.82 (d, J = 4.1 Hz, 1H), 4.50–4.43 (m, 1H), 4.17–4.10 (m, 1H), 4.08–4.00 (m, 1H), 3.86–3.78 (m, 1H), 3.66–3.45 (m, 2H), 3.43–3.21 (m, 2H), 2.27–2.11 (m, 1H), 2.04–1.91 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.4, 158.8 (d, J = 270.2 Hz), 152.5, 149.8, 146.1, 144.0, 138.5 (d, J = 8.6 Hz), 137.81 (d, J = 4.1 Hz), 135.7 (d, J = 10.3 Hz), 126.8, 121.0 (d, J = 22.4 Hz), 120.8, 91.1, 83.4, 75.1, 72.9, 52.2, 51.5, 46.4, 31.1. FT-IR: ν / cm^{-1} = 3106, 1673, 1606, 1537, 1416, 1349, 1270, 1167, 1132, 1073, 972, 898, 818, 799, 721. $[\alpha]_D^{20}$ = +26 (10 mg/mL; MeOH). mp: 66–68 °C. ESI-MS: m/z calculated for [C$_{20}$H$_{23}$FN$_8$O$_9$S+H]$^+$ ([M+H]$^+$) = 571.1, found: 571.2. Purity: 95% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.76 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-4-bromo-3-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (80)

The compound was prepared from **67** (198 mg, 0.24 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (192 mg, 0.24 mmol, 99%, 1.5 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.29 (s, 1H), 8.25 (s, 1H), 8.16–8.13 (m, 1H), 7.85–7.78 (m, 2H), 5.85 (d, J = 4.2 Hz, 1H), 4.52–4.46 (m, 1H), 4.19–4.14 (m, 1H), 4.12–4.04 (m, 1H), 3.88–3.81 (m, 1H), 3.67–3.50 (m, 2H), 3.49–3.24 (m, 2H), 2.30–2.15 (m, 1H), 2.08–1.93 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.4, 152.5, 151.3, 149.8, 146.1, 144.0, 141.6, 137.4, 132.4, 125.3, 120.8, 119.9, 91.1, 83.4, 75.1, 72.9, 52.2, 51.5, 46.5, 31.1. FT-IR: ν / cm^{-1} = 3077, 1667, 1538, 1426, 1351, 1169, 1130, 1033, 968, 886, 833, 798, 775, 721, 661. $[\alpha]_D^{20}$ = +24 (10 mg/mL; MeOH). mp: 81 °C (decomposition). ESI-MS: m/z calculated for [C$_{20}$H$_{23}$BrN$_8$O$_9$S+H]$^+$ ([M+H]$^+$) = 631.1, found: 630.9. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 4.46 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-3-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (81)
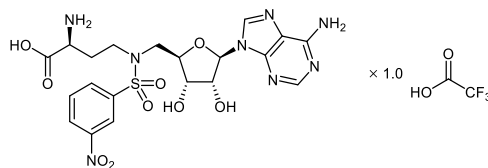
The compound was prepared from **68** (103 mg, 0.14 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (92 mg, 0.14 mmol, 99%, 1.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.41–8.34 (m, 1H), 8.27–8.16 (m, 3H), 8.06–7.99 (m, 1H), 7.65–7.56 (m, 1H), 5.79 (d, J = 4.1 Hz, 1H), 4.51–4.43 (m, 1H), 4.18–4.10 (m, 1H), 4.07–3.98 (m, 1H), 3.85–3.76 (m, 1H), 3.67–3.45 (m, 2H), 3.44–3.21 (m, 1H), 2.28–2.09 (m, 1H), 2.07–1.89 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.5, 152.6, 149.8, 149.5, 146.2, 144.0, 142.5, 134.0, 132.1, 128.4, 123.2, 120.7, 91.1, 83.5, 75.0, 72.9, 52.3, 51.6, 46.5, 31.1. FT-IR: ν / cm^{-1} = 3090, 1673, 1530, 1418, 1352, 1165, 1124, 967, 879, 834, 798, 777, 721, 671, 659. $[\alpha]_D^{20}$ = +25 (10 mg/mL; MeOH). mp: 73–75 °C. ESI-MS: *m/z* calculated for [C$_{20}$H$_{24}$N$_8$O$_9$S+H]$^+$ ([M+H]$^+$) = 553.2, found: 553.0. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.16 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-2-chloro-5-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (82)
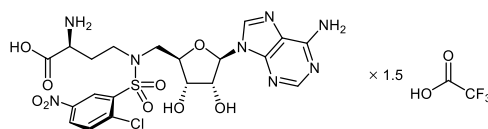

The compound was prepared from **69** (91 mg, 0.12 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (88 mg, 0.12 mmol, 99%, 1.5 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.50 (d, J = 2.7 Hz, 1H), 8.25 (s, 1H), 8.17 (s, 1H), 8.00 (dd, J = 8.7, 2.7 Hz, 1H), 7.60 (d, J = 8.7 Hz, 1H), 5.72 (d, J = 3.6 Hz, 1H), 4.44–4.37 (m, 1H), 4.23–4.15 (m, 1H), 4.12–4.02 (m, 1H), 3.90–3.74 (m, 1H), 3.73–3.66 (m, 2H), 3.56–3.42 (m, 1H), 2.36–2.19 (m, 1H), 2.13–1.98 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.5, 153.0, 149.5, 147.0, 146.9, 143.8, 140.7, 140.1, 134.2, 128.9, 127.2, 120.7, 91.3, 82.5, 75.1, 72.8, 51.6, 50.8, 45.8, 30.5. FT-IR: ν / cm^{-1} = 3103, 1670, 1601, 1525, 1423, 1348, 1195, 1127, 1039, 966, 886, 836, 797, 739, 721. $[\alpha]_D^{20}$ = +24 (10 mg/mL; MeOH). mp: 80–82 °C. ESI-MS: *m/z* calculated for [C$_{20}$H$_{23}$ClN$_8$O$_9$S+H]$^+$ ([M+H]$^+$) = 587.1, found: 587.0. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.78 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-2-fluoro-5-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (83)
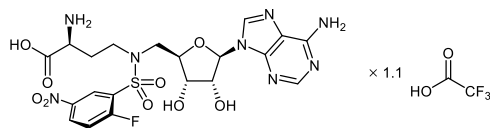

The compound was prepared from **70** (160 mg, 0.21 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (145 mg, 0.21 mmol, 99%, 1.1 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.33 (dd, J = 5.9, 2.9 Hz, 1H), 8.24 (s, 1H), 8.17–8.11 (m, 2H), 7.36–7.28 (m, 1H), 5.71 (d, J = 3.7 Hz, 1H), 4.44–4.39 (m, 1H), 4.22–4.15 (m, 1H), 4.11–4.02 (m, 1H), 3.90–3.83 (m, 1H), 3.77–3.62 (m, 3H), 3.49–3.36 (m, 1H), 2.33–2.18 (m, 1H), 2.13–1.98 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.4, 163.5 (d, J = 263.5 Hz), 152.8, 149.5, 146.7, 144.6, 143.9, 131.2 (d, J = 11.0 Hz), 130.8 (d, J = 17.7 Hz), 127.1 (d, J = 3.0 Hz), 120.7, 119.6 (d, J = 24.9 Hz), 91.2, 82.5, 75.1, 72.8, 51.6, 51.2, 45.5, 30.8. FT-IR: ν / cm^{-1} = 3110, 1668, 1533, 1474, 1424, 1351, 1258, 1129, 1058, 970, 896, 837, 798, 744, 721. $[\alpha]_D^{20}$ = +11 (10 mg/mL; MeOH). mp: 80–82 °C. ESI-MS: m/z calculated for $[C_{20}H_{23}FN_8O_9S+H]^+$ ($[M+H]^+$) = 571.1, found: 571.0. Purity: >99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.81 min).
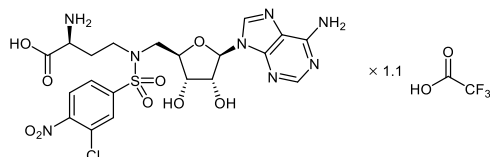
(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-3-chloro-4-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (84)

The compound was prepared from **71** (179 mg, 0.23 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (162 mg, 0.23 mmol, 99%, 1.1 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.32 (s, 1H), 8.26 (s, 1H), 7.97 (d, J = 1.8 Hz, 1H), 7.89 (d, J = 8.4 Hz, 1H), 7.83 (dd, J = 8.4, 1.8 Hz, 1H), 5.88 (d, J = 4.3 Hz, 1H), 4.55–4.49 (m, 1H), 4.22–4.16 (m, 1H), 4.15–4.06 (m, 1H), 3.90–3.83 (m, 1H), 3.71–3.52 (m, 2H), 3.51–3.27 (m, 2H), 2.32–2.17 (m, 1H), 2.10–1.95 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.3, 152.5, 151.6, 149.8, 146.1, 145.1, 144.0, 131.5, 128.2, 128.2, 127.5, 120.8, 91.1, 83.6, 75.1, 72.9, 52.4, 51.5, 46.6, 31.2. FT-IR: ν / cm^{-1} = 3101, 1669, 1534, 1507, 1424, 1346, 1297, 1186, 1164, 1130, 968, 835, 797, 721, 671. $[\alpha]_D^{20}$ = +20 (10 mg/mL; MeOH). mp: 73–75 °C. ESI-MS: m/z calculated for $[C_{20}H_{23}ClN_8O_9S+H]^+$ ($[M+H]^+$) = 587.1, found: 586.9. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 4.53 min).

tert-Butyl (_S_)-4-((_N_-(((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-_d_][1,3]dioxol-4-yl)methyl)-6-chloropyridine)-3-sulfonamido)-2-((_tert_-butoxycarbonyl)amino)-butanoate (88)
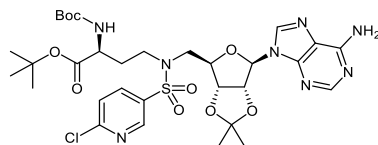
The compound was prepared from **3** (251 mg, 0.45 mmol, 1.0 equiv.), 6-chloropyridine-3-sulfonyl chloride (94 mg, 0.45 mmol, 1.0 equiv.), and NEt$_3$ (62 μL, 0.45 mmol, 1.0 equiv.) according to general procedure D to afford the desired product as a slightly yellowish solid (207 mg, 0.28 mmol, 62%). ^{1}H NMR (300 MHz, DMSO-d_6): δ / ppm = 8.67 (d, J = 2.5 Hz, 1H), 8.32 (s, 1H), 8.19 (s, 1H), 8.04 (dd, J = 8.3, 2.5 Hz, 1H), 7.60 (d, J = 8.3 Hz, 1H), 7.37 (s, 2H), 7.19 (d, J = 7.8 Hz, 1H), 6.18 (d, J = 2.3 Hz, 1H), 5.41 (dd, J = 6.3, 2.3 Hz, 1H), 5.03 (dd, J = 6.3, 3.2 Hz, 1H), 4.37–4.26 (m, 1H), 3.84–3.72 (m, 1H), 3.64–3.44 (m, 2H), 3.34–3.01 (m, 2H), 1.94–1.67 (m, 2H), 1.52 (s, 3H), 1.39–1.33 (m, 18H), 1.30 (s, 3H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ / ppm = 171.1, 156.1, 155.5, 154.0, 152.6, 148.6, 147.9, 140.2, 138.1, 135.0, 125.0, 119.3, 113.5, 88.9, 83.8, 83.3, 82.0, 80.5, 78.2, 52.1, 49.8, 45.6, 29.8, 28.2, 27.6, 26.9, 25.1. FT-IR: ν / cm^{-1} = 3415, 3337, 3227, 2980, 2936, 1727, 1702, 1651, 1602, 1357, 1157, 1075, 872, 784, 742. $[\alpha]_D^{20}$ = +37 (10 mg/mL; CHCl$_3$). mp: 222 °C (decomposition). ESI-MS: m/z calculated for [C$_{31}$H$_{43}$ClN$_8$O$_9$S+H]$^+$ ([M+H]$^+$) = 739.3, found: 739.3.

2-(Benzylthio)-5-chloropyrazine (87)
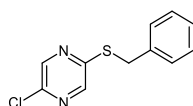
Using a modified procedure from C OOPER _et al._, _Novel Sulfonamide Carboxamide Compounds._ WO2019/8025 A1, **2019**.

To a solution of benzyl mercaptan (0.79 g, 0.95 mmol, 0.95 equiv.) in THF (30 mL) was added NaH (60% dispersion in mineral oil, 0.23 g, 9.46 mmol, 1.41 equiv.) at 0 °C. After the suspension was stirred at 0 °C for 10 min, a solution of 2,5-dichloropyrazine (1.00 g, 6.71 mmol, 1.0 equiv.) in THF (5 mL) was added dropwise, and the mixture was stirred for an additional hour at 0 °C. The mixture was warmed up to room temperature and stirred for 16 h. Methanol (1 mL) was added at 0 °C, and stirring was continued for 5 min at 0 °C. The mixture was extracted with DCM (1 × 75 mL) and the separated organic layer was dried over Na$_2$SO$_4$. After filtration, the solvent was removed under reduced pressure at 40 °C to yield the desired product as an orange oil (1.37 g, 5.78 mmol, 91%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.42 (d, J = 1.5 Hz, 1H), 8.22 (d, J = 1.5 Hz, 1H), 7.43–7.24 (m, 5H), 4.43–4.40 (m, 2H). ^{13}C NMR (75.5 MHz,

CDCl$_3$): δ / ppm = 154.8, 144.9, 143.6, 142.5, 136.9, 129.1, 128.8, 127.7, 34.6. FT-IR: ν / cm^{-1} = 3062, 3029, 2923, 2852, 1495, 1439, 1418, 1335, 1297, 1281, 1145, 1012, 894, 766, 697.

tert-Butyl (_S_)-4-((_N_-(((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-_d_][1,3]dioxol-4-yl)methyl)-5-chloropyrazine)-2-sulfonamido)-2-((_tert_-butoxycarbonyl)amino)-butanoate (89)

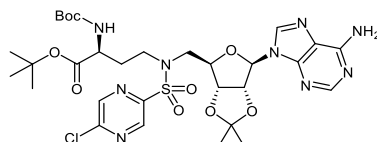1.) Using a modified procedure from COOPER _et al._, _Novel Sulfonamide Carboxamide Compounds._ WO2019/8025 A1, **2019**.

To a solution of compound **87** (0.43 g, 1.82 mmol, 1.0 equiv.) in DCM (9 mL) was added water (0.9 mL). The mixture was cooled to −5 °C, and sulfuryl chloride (1.66 g, 12.30 mmol, 6.77 equiv.) was added. After the mixture was stirred at −5 °C for 2 h, ice (ca. 8 g) was added, and the organic phase was separated. The aqueous phase was extracted with DCM (2 × 15 mL) and dried over MgSO$_4$. After filtration, the solvent was removed under reduced pressure at 40 °C to yield 5-chloropyrazine-2-sulfonyl chloride as a yellowish oil, which was directly used in the next step.

2.) The compound was prepared from **3** (250 mg, 0.44 mmol, 1.0 equiv.) and 5-chloropyrazine-2-sulfonyl chloride (step 1, 142 mg, 0.67 mmol, 1.5 equiv.) according to general procedure E to afford the desired product as a colorless solid (112 mg, 0.15 mmol, 34%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.60 (s, 1H), 8.48 (d, J = 1.3 Hz, 1H), 8.31 (s, 1H), 7.81 (s, 1H), 6.20 (s, 2H), 5.92 (d, J = 1.9 Hz, 1H), 5.39 (dd, J = 6.4, 1.9 Hz, 1H), 5.25 (d, J = 8.2 Hz, 1H), 5.06 (dd, J = 6.4, 3.4 Hz, 1H), 4.38–4.28 (m, 1H), 4.11–3.95 (m, 1H), 3.85–3.73 (m, 1H), 3.65–3.43 (m, 2H), 3.37–3.21 (m, 1H), 2.13–2.00 (m, 1H), 1.92–1.77 (m, 1H), 1.54 (s, 3H), 1.42–1.39 (m, 18H), 1.33 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.0, 155.9, 155.6, 153.1, 152.1, 151.8, 149.0, 144.2, 142.4, 140.2, 120.4, 114.7, 90.7, 85.5, 84.2, 82.7, 82.5, 80.0, 52.1, 50.6, 46.4, 32.1, 28.4, 28.0, 27.1, 25.3. FT-IR: ν / cm^{-1} = 3411, 3340, 3230, 2980, 2937, 1698, 1653, 1602, 1332, 1135, 1075, 994, 872, 804, 751. $[\alpha]_D^{20}$ = +23 (10 mg/mL; CHCl$_3$). mp: 216 °C (decomposition). ESI-MS: m/z calculated for [C$_{30}$H$_{42}$ClN$_9$O$_9$S+H]$^+$ ([M+H]$^+$) = 740.3, found: 740.2.

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-6-chloropyridine)-3-sulfonamido)butanoic acid, trifluoroacetate salt (90)
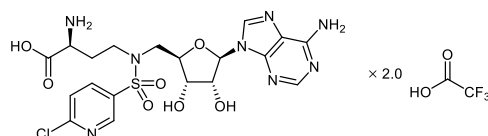
The compound was prepared from **88** (110 mg, 0.15 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (114 mg, 0.15 mmol, 99%, 2.0 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.46 (d, J = 2.5 Hz, 1H), 8.20 (s, 1H), 8.16 (s, 1H), 7.92 (dd, J = 8.5, 2.5 Hz, 1H), 7.29 (d, J = 8.5 Hz, 1H), 5.73 (d, J = 4.2 Hz, 1H), 4.44–4.38 (m, 1H), 4.10–4.03 (m, 1H), 4.02–3.93 (m, 1H), 3.78–3.70 (m, 1H), 3.55–3.37 (m, 2H), 3.36–3.11 (m, 2H), 2.19–2.03 (m, 1H), 1.99–1.82 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.3, 156.3, 152.3, 149.8, 149.4, 145.7, 144.1, 139.2, 136.5, 126.1, 120.8, 91.1, 83.4, 75.0, 72.9, 52.2, 51.5, 46.4, 31.1. FT-IR: ν / cm^{-1} = 3497, 3359, 3286, 3041, 2891, 2444, 2197, 1633, 1597, 1348, 1167, 1113, 1040, 783, 747. $[\alpha]_D^{20}$ = +22 (10 mg/mL; MeOH). mp: 71–73 °C. ESI-MS: m/z calculated for [C$_{19}$H$_{23}$ClN$_8$O$_7$S+H]$^+$ ([M+H]$^+$) = 543.1, found: 543.2. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.70 min).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-5-chloropyrazine)-2-sulfonamido)butanoic acid, trifluoroacetate salt (91)
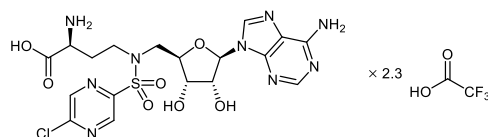
The compound was prepared from **89** (86 mg, 0.12 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (94 mg, 0.12 mmol, 99%, 2.3 equiv. TFA). ^{1}H NMR (300 MHz, DMSO-d_6): δ / ppm = 8.65 (d, J = 1.4 Hz, 1H), 8.44 (d, J = 1.4 Hz, 1H), 8.36 (s, 1H), 8.24 (s, 1H), 5.57 (d, J = 5.5 Hz, 1H), 4.39–4.31 (m, 1H), 3.91–3.80 (m, 2H), 3.63 (s, 1H), 3.53–3.30 (m, 3H), 3.27–3.11 (m, 1H), 1.98–1.83 (m, 1H), 1.82–1.66 (m, 1H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ / ppm = 170.5, 151.8, 151.4, 151.0, 148.6, 146.8, 144.9, 142.2, 142.0, 119.0, 87.7, 81.8, 73.1, 71.2, 51.0, 49.9, 45.4, 29.8. FT-IR: ν / cm^{-1} = 3086, 1671, 1509, 1431, 1348, 1300, 1188, 1131, 1018, 971, 924, 827, 798, 764, 722. $[\alpha]_D^{20}$ = +27 (10 mg/mL; MeOH). mp: 76–87 °C. ESI-MS: m/z calculated for [C$_{18}$H$_{22}$ClN$_9$O$_7$S+H]$^+$ ([M+H]$^+$) = 544.1, found: 544.0. Purity: 95% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.53 min).

3.6   Warhead-decorated adenosine and SAH derivatives

To identify other potential warheads suitable for inhibition of DNMT2, a study was conducted in which various electrophilic groups were attached to the SAH mimic adenosyl-Dab **51** to yield reactive Y-shaped inhibitors. In addition, warheads were attached to the 5'-position of adenosine using different functional groups for connection of the substructures. This strategy was used to investigate if the amino acid side chain is crucial for proper warhead orientation and if a covalent reaction can still occur inside the binding site. Structural simplification of biologically active substances can be beneficial, for example to improve their pharmacokinetic profiles or to reduce synthetic effort.[285] For this project, warheads targeting cysteine residues were selected, including nitriles, propargyl amides, acrylamides, disulfides, maleimides, alkyl bromide, vinylsulfonamides, and electron-deficient aromats (**Figure 18**).
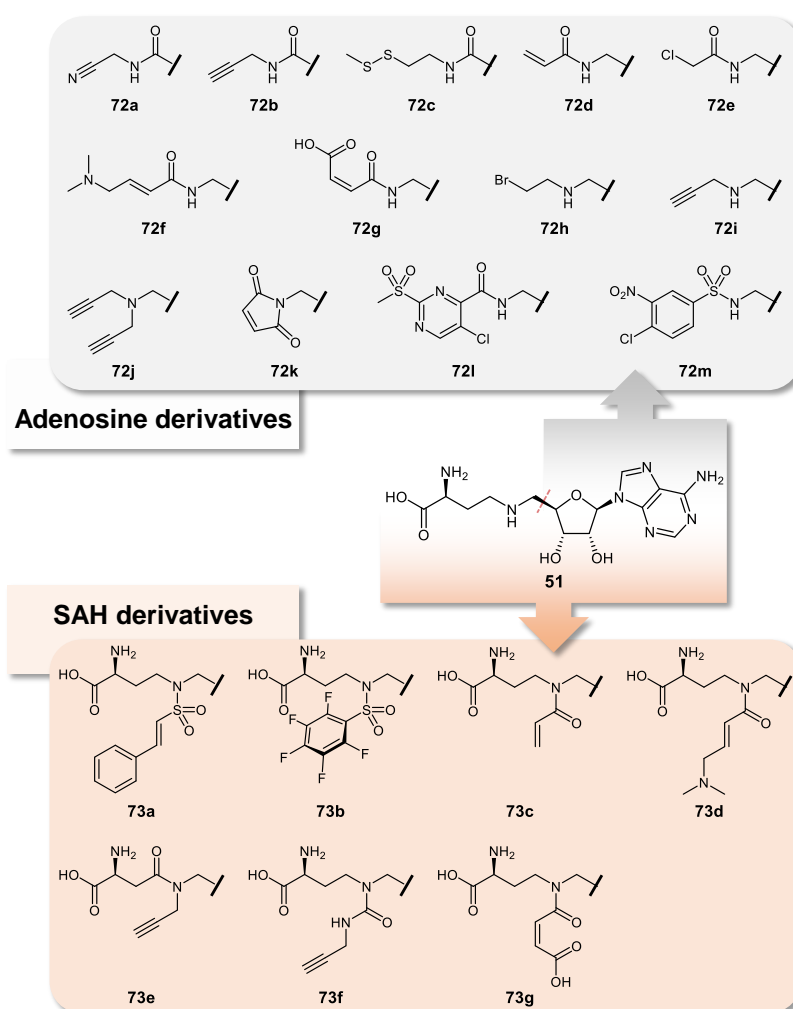
Figure 18: Different types of warheads attached to the adenosine as well as the SAH scaffold.

Compounds **72a–m** and **73c–g** were synthesized by ▉▉▉▉▉▉▉ (master thesis, under supervision of MARVIN SCHWICKERT, ▉▉▉▉▉▉ group). The vinylsulfonamide **73a** and the pentafluorophenyl derivative **73b** were synthesized by MARVIN SCHWICKERT. Protected adenosyl-Dab **50** was brought to reaction with the respective sulfonyl chlorides **74a** and **74b** using a two-phase system consisting of DCM

and saturated NaHCO$_3$ solution at room temperature (**Scheme 21**). The resulting sulfonamides **75a** and **75b** were treated with 50 vol% TFA in dichloromethane at 5 °C, followed by treatment with 14 vol% TFA in water at 5 °C to give the final inhibitors **73a** and **73b**.
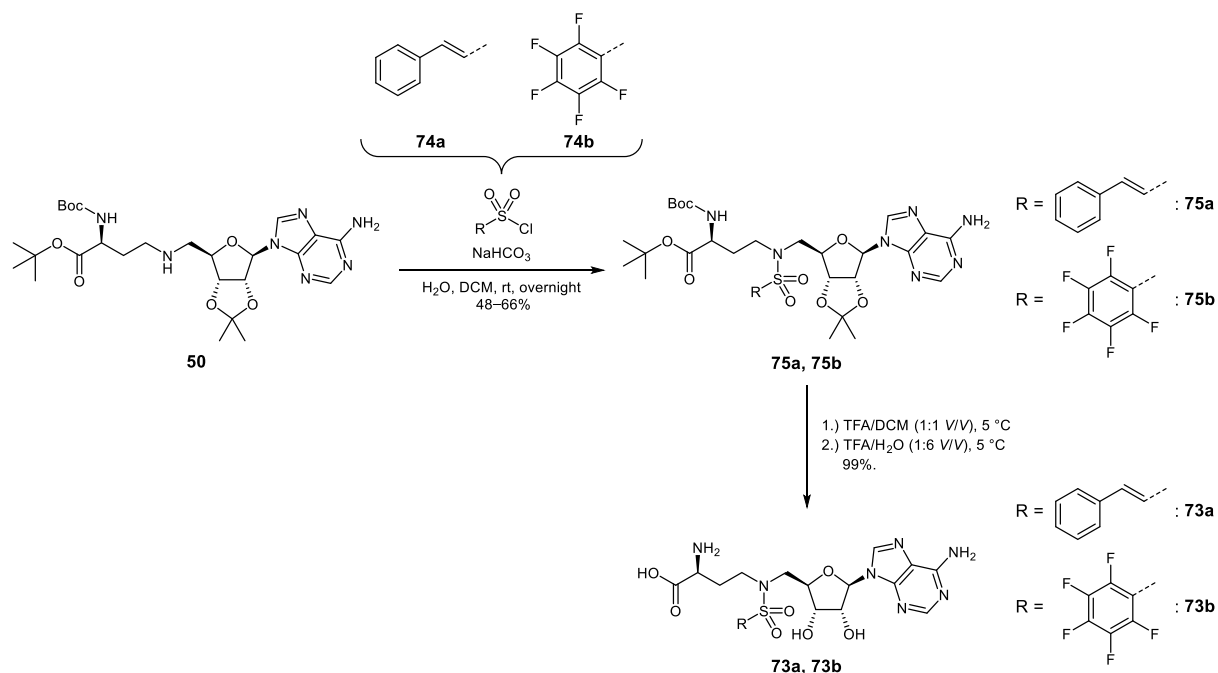
Scheme 21: Synthesis of compounds **73a** and **73b**.
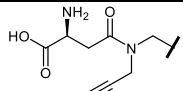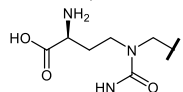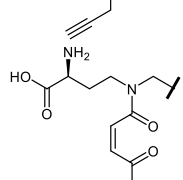
Binding analysis of **72a**–**m** and **73c**–**g** was conducted by ▮▮▮▮▮▮▮▮▮ using MST (under supervision of ▮▮▮▮▮▮▮▮▮▮▮▮▮ (▮▮▮▮▮▮▮ group) and DSF (under supervision of MARVIN SCHWICKERT). Compounds **73a** and **73b** were tested at a later stage of this project using an optimized MST displacement method by ▮▮▮▮▮▮▮▮▮▮ (for details see **Section 3.8**). To evaluate their inhibitory potential, the compounds were subjected to a tritium incorporation assay measured by ▮▮ ▮▮▮ and ▮▮▮▮▮▮▮ (▮▮▮ group). All results are listed in **Table 4**.

Table 4: Results of compounds **72a–m** and **73a–g**.
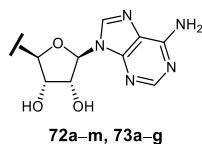
72a–m, 73a–g

Compound		MST Binding at 250 μM	DSF Binding at 100 μM	^{3}H-Assay % inhibition at 100 μM
72a		✓	✗	n. i.
72b		✓	✗	n. i.
72c		✓	✗	n. i.
72d		✓	✗	n. i.
72e		✓	✗	n. i.
72f		✗	✗	n. i.
72g		✓	✗	n. i.
72h		✗	✗	n. i.
72i		✓	✗	n. i.
72j		✗	✗	n. i.
72k		✗	✗	n. i.
72l		✗	✗	n. i.
72m		✓	✗	n. i.
73a		✗*	n. d.	n. d.
73b		✗*	n. d.	n. d.
73c		✗	✗	n. i.
73d		✗	✓	n. i.

139

Table 4 continued.

Compound	MST Binding at 250 μM	DSF Binding at 100 μM	^{3}H-Assay % inhibition at 100 μM
73e	✗	✗	n. i.
73f	✓	✗	n. i.
73g	✓	✗	14 ± 4

*Determined by MST-based displacement method at 20 μM using a fluorescent dye (for details see **Section 3.8**); n. d. = not determined; n. i. = no inhibition.

The testing revealed that of the 20 compounds, ten (**72a–e**, **72g**, **72i**, **72m**, **73f**, and **73g**) were identified as potential binders of DNMT2 by MST. Using DSF, only compound **73d** was identified as a binder, which was in contrast to its corresponding MST result. Notably, the MST and DSF results are barely consistent, which could be explained by the different compound concentrations (250 μM vs. 100 μM) used in both methods. Regarding the inhibition, only the maleic acid-derived compound **73g** showed weak activity (14 ± 4% inhibition at 100 μM) in the tritium incorporation assay. Due to the low inhibition, compound **73g** was not further investigated for potential covalent binding.

For a more detailed discussion, as well as the presentation of the experimental procedures for the synthesis of compounds **72a–m** and **73c–g**, the reader is referred to the corresponding master thesis:

██████████, *Entwicklung von Adenosin-Derivaten mit elektrophilen Warheads als potenzielle kovalente Dnmt2-Inhibitoren*, **2020**, Johannes Gutenberg University Mainz.

Own contribution: Design of inhibitors, synthesis of **73a** and **73b**, supervising the master thesis (██████████) about the synthesis of warhead-decorated inhibitors **72a–m** and **73c–g** as well as the testing using DSF.

3.7 Development of NSUN6-selective inhibitors based on the SAH scaffold

Since SAM-dependent MTases like DNMT2, NSUN2, and NSUN6 share similar binding pockets for their cofactor SAM, the development of selective SAH-derived inhibitors is highly challenging. These enzymes address different positions of their substrate tRNA (**Section 2.2.2**), consequently, they structurally differ in their cytidine sites. Therefore, selectivity could rather be achieved by attaching suitable cytidine-site targeting side chains to the SAH scaffold than by direct changes of the SAH scaffold itself. This theory was substantiated by results of first studies described in **Sections 3.4** and **3.5** as (*SR*)-*N*-but-3-yn-2-yl-adenosyl-Dab **69s-A/B** as well as a the 4-bromo-3-nitrophenylsulfonamide-based SAH derivative **71i** showed high selectivity toward NSUN2 and NSUN6. Apart from the results presented in the first published study (**Section 3.4**),[214] several *N*-alkyl derivatives (**69f**, **69g**, **69k**, and **69l**) were tested for NUSN6 inhibition by ▮▮▮▮ (▮▮▮ group) in a tritium incorporation assay (**Table 5**). In addition, K_D values were obtained by ▮▮▮▮▮ (▮▮▮▮ group) using ITC.

Table 5: Inhibition of DNMT2 and NSUN6 by compounds **69f**, **69g**, **69k**, **69l**, SAH, and SFG.

69f, 69g, 69k, 69l
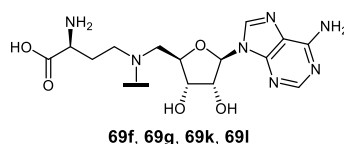
	Compound	DNMT2	NSUN6	
		³H-Assay	³H-Assay	ITC
		% inhibition at 100 μM	% inhibition at 100 μM	K_D / μM
18	**SAH**	85.7 ± 1.7	100	6.8 ± 0.6
19	**SFG**	83.5 ± 1.1	70.2 ± 1.8	5.5 ± 0.9
69f		n. i.	24.9 ± 1.4	32.6 ± 16.1
69g		n. i.	10.1 ± 2.0	n. d.
69k		n. i.	45.8 ± 6.1	10.9 ± 1.9
69l		22.4 ± 2.6	22.4 ± 1.8	n. d.

n. d. = not determined; n. i. = no inhibition.

The testing revealed that NSUN6 preferred bulkier side chains (inhibition of 10–46% at 100 μM) while these modifications resulted in no inhibitory activity against DNMT2, as demonstrated with compounds **69f**, **69g**, and **69k**. The cyclohexyl derivative **69k** showed the strongest discrepancy with 45.8 ± 6.1% inhibition of NSUN6 compared to no inhibition of DNMT2. Notably, with a K_D value of 10.9 ± 1.9 μM, compound **69k** showed a binding affinity to NSUN6 that approximated the binding affinity of the natural ligands SAH and SFG, which exhibited high inhibition of 100% and 70.2 ± 1.8% at 100 μM, respectively. Since the cyclohexyl moiety presumably targets the NSUN6 cytidine site, unlike SAH and SFG, competition with the tRNA might have a stronger impact on inhibition of **69k**. Interestingly, the discrepancy between K_D values and observed inhibitions of NSUN6 was lower compared to Y-shaped

DNMT2 inhibitors presented in **Section 3.4**, which might be due to higher affinity of the tRNA for NSUN6. Small unsaturated systems such as the allyl (**69l**) derivative did not affect selectivity. The findings could be explained by comparing the cytidine sites of DNMT2, NSUN2, and NSUN6 (**Figure 19**).
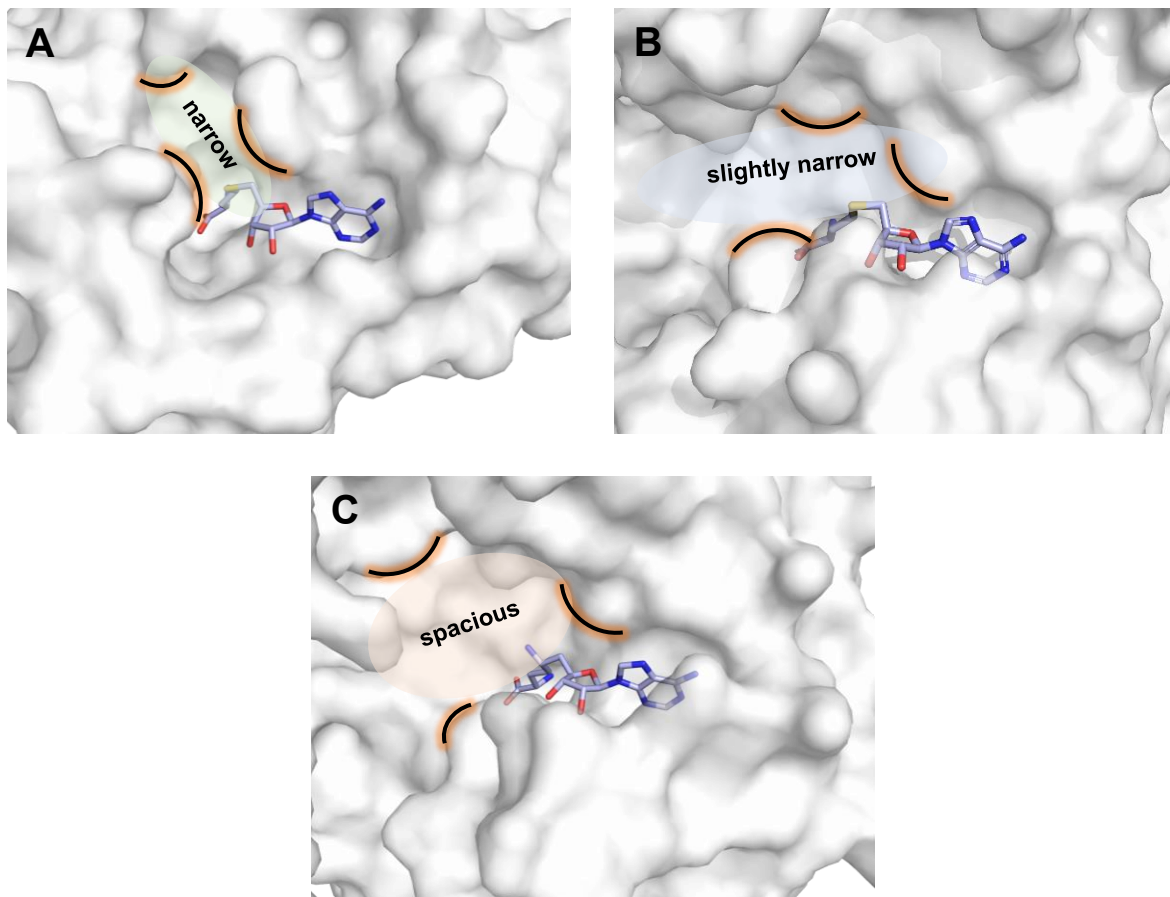
Figure 19: A: Crystal structure of SAH in complex with hDNMT2 (PDB-ID:1G55).[197] B: AlphaFold homology model of hNSUN2[200,201] aligned with SAH of the hDNMT2 crystal structure (PDB-ID:1G55). C: Crystal structure of SFG in complex with hNSUN6 (PDB-ID: 5WWR).[116] Glowing lines indicate potential clashing points of an inhibitor side chain.

The cytidine site of DNMT2 appears to be narrow, presumably allowing only less rigid, small planar structures to be positioned, which is consistent with findings described in **Section 3.4**. In the structure of NSUN2 a slightly wider cytidine site can be found compared to DNMT2. NSUN6, on the other hand, exhibits a large cavity that should allow binding of sterically more complex substructures. Consequently, developing SAH derivatives with large and/or bulky side chains might allow binding to NSUN6, whereas such structures should lead to steric clash with the other two MTases. Based on this idea, sterically complex compounds were designed (**Figure 20**) to force steric clash toward DNMT2 and NSUN2 and enable selectivity. The modifications included variations the cyclohexyl moiety (**76a–i**) as the derivative **69k** showed moderate inhibition of NSUN6 and selectivity toward DNMT2, as well as large aromatic systems (**77a–k**). The aim of this SAR was to establish a "clash profile" that provides information about compatibility of cytidine sites and sterically complex side chain substructures.

142
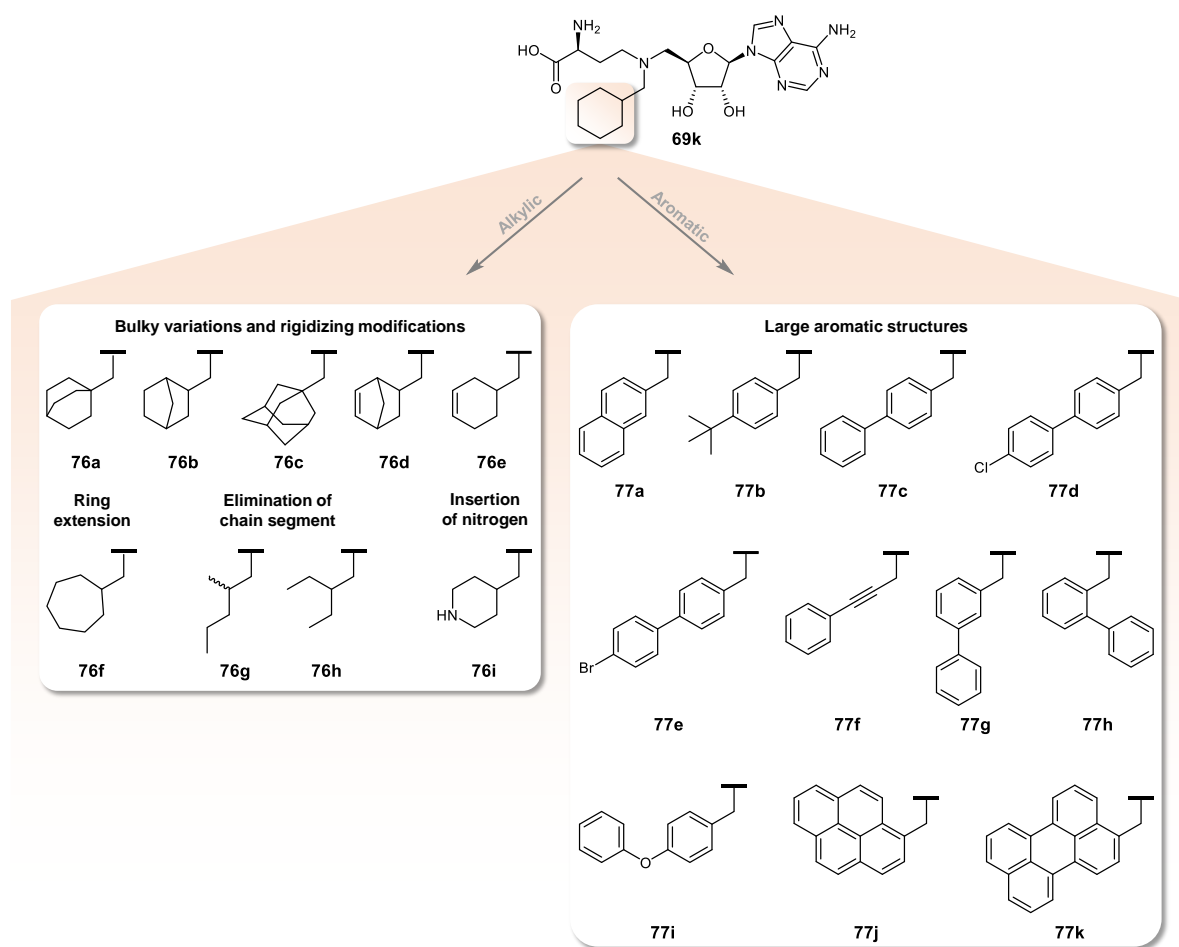
Figure 20: Design of SAH derivatives for testing on NSUN6. Some of the structures were selected together with ▮▮▮▮▮ (▮▮▮▮▮ group).

In this work, a small set of the designed compounds (**76a**, **76c**, **76f**, **77c**, and **77i–k**, **Figure 20**) was generated to obtain first data about a potential selectivity profile. They were synthesized starting from the secondary amine **50** which was reacted with the aldehydes **79a–g** in a reductive amination using sodium triacetoxyborohydride yielding the precursors **81a–c** and **82a–d** (**Scheme 22**). Bicyclo[2.2.2]octane-1-carbaldehyde (**79a**) was prepared by oxidization of the corresponding alcohol **78** using pyridinium chlorochromate. In the final step, **81a–c** and **82a–d** were treated with 50 vol% TFA in dichloromethane at 5 °C, followed by treatment with 14 vol% TFA in water at 5 °C to give the inhibitors **76a**, **76c**, **76f**, **77c**, and **77i–k**.

A

B
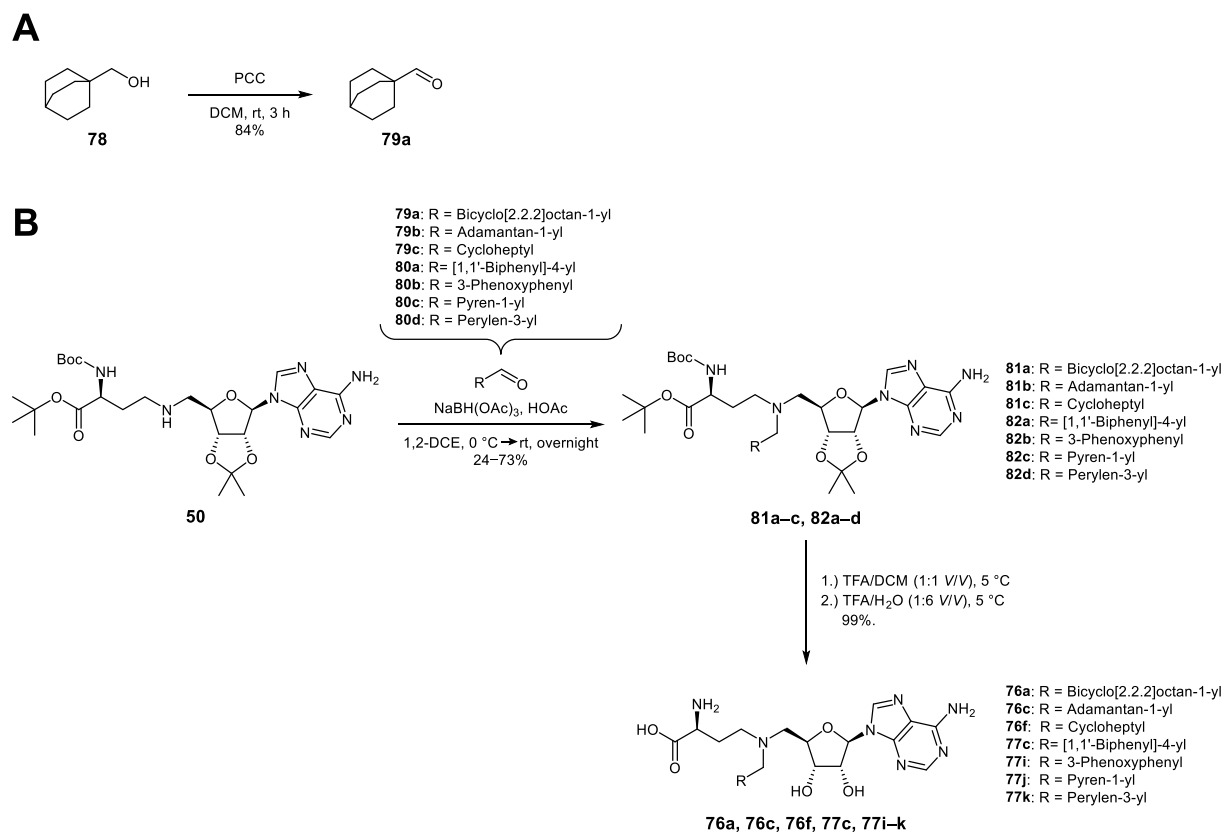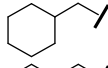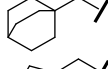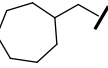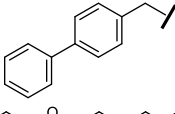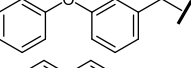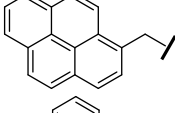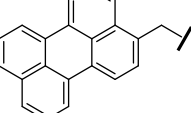
Scheme 22: A: Synthesis of bicyclo[2.2.2]octane-1-carbaldehyde (**79a**). B: Synthesis of **76a**, **76c**, **76f**, **77c**, and **77i–k**.

Compounds **76a**, **76c**, **76f**, **77c**, and **77i–k** were evaluated for their binding affinities to DNMT2 by ▮▮▮▮▮▮▮▮▮▮ (▮▮▮▮▮▮ group) using an MST displacement method. The inhibitory potential to DNMT2, NSUN2, and NSUN6 was determined in a tritium incorporation assay by ▮▮▮▮ ▮▮▮▮ (▮▮▮▮▮▮ group). All results are given in **Table 6**.

Bulkier modifications of the cyclohexyl moiety (**76a** and **76c**) or ring extension (**76f**) resulted in a slight decrease in inhibition of NSUN6 (21–35% inhibition at 100 μM) suggesting adverse effects. However, these structures exhibited selectivity toward DNMT2 as they were not considered as binders in the MST displacement assay (shift = 1.0–2.1‰). Selectivity toward NSUN2 could also be observed as **76a** and **76f** showed no inhibition, whereas the adamantyl derivative **76c** was slightly active with $11.7 \pm 0.8\%$. The phenoxyphenyl derivative **77i** is equipotent to **69k** ($42.7 \pm 3.9\%$ vs. $45.8 \pm 6.1\%$) but shows weak binding to DNMT2 (MST shift = 8.1‰) and NSUN2 ($10.6 \pm 2.9\%$ inhibition). Notably, the biphenyl derivative **77c** increased the inhibition of NSUN6 ($73.2 \pm 4.7\%$) but also showed moderate activity against DNMT2 ($49.8 \pm 4.2\%$) and NSUN2 (52.7 ± 8.3). Large aromatic systems such as pyrene (**77j**) or perylene (**77k**) were either not evaluable or caused inaccurate data due to aggregation effects.

A dataset generated by seven compounds is not yet sufficient to establish a proper MTase clash profile. To increase the size of the molecule library by preparing, among others, compounds **76b**, **76d**, **76e**, **76g**, **76h**, **76i**, **77a**, **77b**, and **77d–h**, the project will be continued by ▮▮▮▮▮▮ (▮▮▮▮▮▮ group).

Table 6: Inhibition of DNMT2, NSUN2, and NSUN6 by compounds **76a**, **76c**, **76f**, **77c**, and **77i–k**.

76a, 76c, 76f, 77c, 77i–k
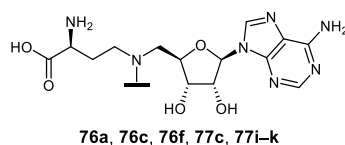
Compound		DNMT2		NSUN2	NSUN6
		MST* shift at 20 μM / ‰	^{3}H-Assay % inhibition at 100 μM	^{3}H-Assay % inhibition at 100 μM	^{3}H-Assay % inhibition at 100 μM
69k		n. d.	n. i.	t. b. d.	45.8 ± 6.1
76a		1.0	n. d.	n. i.	20.8 ± 5.7
76c		1.2	n. d.	11.7 ± 0.8	35.9 ± 4.8
76f		2.1	n. d.	n. i.	34.6 ± 8.7
77c		10.4	49.8 ± 4.2	52.7 ± 8.3	73.2 ± 4.7
77i		8.1	n. d.	10.6 ± 2.9	42.7 ± 3.9
77j		13.4	67.6 ± 10.9**	n. i.	25.4 ± 7.1**
77k		n. e.	n. d.	27.1 ± 9.5**	39.1 ± 3.4**

*Determined by MST-based displacement method using a fluorescent dye (for details see **Section 3.8**); **not accurate due to bad solubility or induced aggregation caused by compound; n. d. = not determined; n. e. = not evaluable; n. i. = no inhibition; t. b. d. = to be determined.

Nevertheless, the dataset provides preliminary evidence for a selectivity profile as bulky moieties derived from the cyclohexyl substructure appeared to generate selectivity for NSUN6. The combination of a selectivity generating moiety (**69k**) with an affinity generating substructure (**77c**) could lead to a suitable activity and selectivity profile. An exemplary structure is shown in **Figure 21**. However, this hypothesis needs to be validated by additional data.
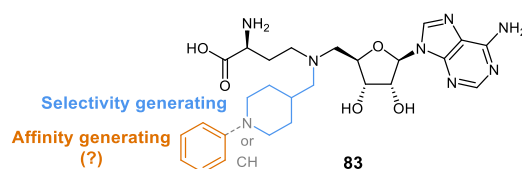
Figure 21: Possible inhibitor that combines the selectivity generating substructure of **69k** with the affinity generating moiety of **77c**.

3.8 Development of tool compounds for biophysical applications to assay methyltransferase inhibitors

Note: The substance numbers from the corresponding publication are listed *in italics* behind the substance numbers in this thesis.

3.8.1 Summary and own contribution

The development of MTase inhibitors is challenging since reliable methods to identify promising binders are required. Compared to well established fluorometric assays used for the evaluation of protease inhibitors, the methods that can be used for measurement of MTase ligands have several disadvantages. Tritium incorporation assays using ^{3}H-SAM are currently the most reliable and therefore often used methods to assay MTases.[193,286] However, these assays are very time-consuming and cost-intensive. Other methods, such as LC-MS, also require a lot of working steps and time. Therefore, fast, reliable, and cost-efficient screening methods are needed to pre-evaluate compounds and reduce their number subjected to tritium incorporation assays or LC-MS-based methods. Based on a literature-known fluorescent SAH derivative, 5-carboxyfluorescein (5-FAM)-triazolyl-adenosyl-Dab (FTAD, **84** (**6**)), originally developed for a fluorescence polarization assay to investigate the histone methyltransferase mixed-lineage leukemia 1 (MLL1),[287] an optimized MST screening method was established for DNMT2. In this assay, FTAD forms a complex with DNMT2, which allows fluorescence tracing. Adding a ligand of interest displaces FTAD, resulting in an altered fluorescence signal due to dissociation of the FTAD-DNMT2 complex.
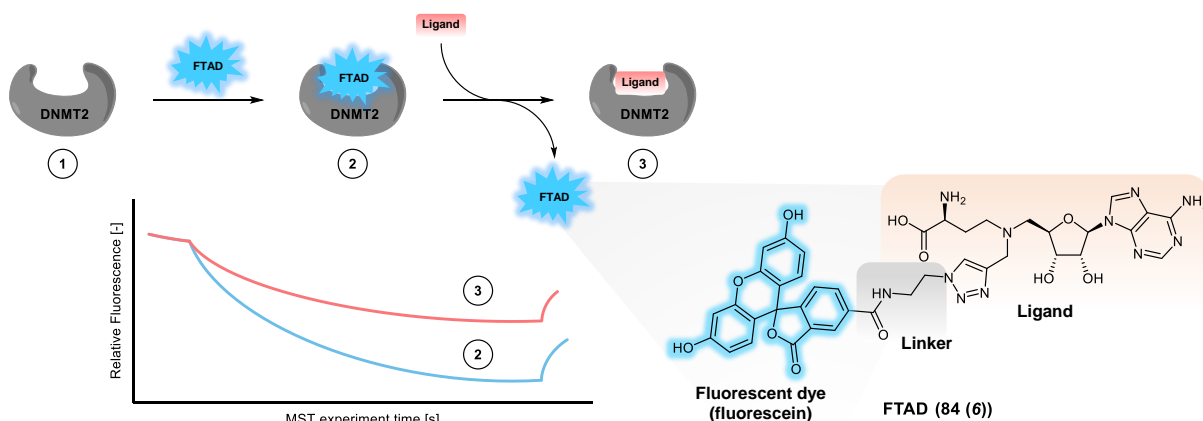
Figure 22: Left: principle of the MST displacement method. Right: Structure of FTAD (**84** (**6**)).

The method was evaluated using SAH, SFG, and five SAH derivatives (four alkyne-based inhibitors (**69s-A/B** (**9**), **69m**–**o** (**11**–**13**)) and one non-binder (**69c** (**10**))) introduced in a previous study (**Section 3.4**).[214] It was found that the thermophoresis shifts in this displacement assay correlated with the measured inhibition from the tritium incorporation assay. In addition, binding affinity could be determined for all compounds. Overall, this optimized FTAD displacement assay circumvents various problems inherent to standard MST methods by providing the following improvements:

➢ reduced aggregation due to lower concentrations of ligand

➢ half-quantitative screenings outcomes: compounds can be pre-evaluated for potential inhibition of DNMT2

➢ determination of active site interactions

➢ no enzyme labeling required

For a more detailed discussion, as well as the presentation of the experimental procedures, the reader is referred to the corresponding publication (**Section 3.8.2**) and Supporting Information (**Appendix**) (**DOI: 10.1021/acsptsci.2c00175**).

Own contribution: synthesis of FTAD (preparation of precursor *4*, conducting step iv), writing the corresponding part in the manuscript and Supporting Information. Creation of graphical abstract, **Scheme 1**, and **Scheme 2**.

3.8.2   Publication

ACS Pharmacology & Translational Science

pubs.acs.org/ptsci

Letter

An Optimized Microscale Thermophoresis Method for High-Throughput Screening of DNA Methyltransferase 2 Ligands

Robert Alexander Zimmermann,[‡] Marvin Schwickert,[‡] J. Laurenz Meidner,[‡] Zarina Nidoieva, Mark Helm, and Tanja Schirmeister*

Cite This: *ACS Pharmacol. Transl. Sci.* 2022, 5, 1079–1085

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Developing methyltransferase inhibitors is challenging, since most of the currently used assays are time-consuming and cost-intensive. Therefore, efficient, fast, and reliable methods for screenings and affinity determinations are of utmost importance. Starting from a literature-known fluorescent *S*-adenosylhomocysteine derivative, 5-FAM-triazolyl-adenosyl-Dab, developed for a fluorescence polarization assay to investigate the histone methyltransferase mixed-lineage leukemia 1, we herein describe the applicability of this compound as a fluorescent tracer for the investigation of DNA-methyltransferase 2 (DNMT2), a human RNA methyltransferase. Based on these findings, we established a microscale thermophoresis (MST) assay for DNMT2. This displacement assay can circumvent various problems inherent to this method. Furthermore, we optimized a screening method via MST which even indicates if the detected binding is competitive and gives the opportunity to estimate the potency of a ligand, both of which are not possible with a direct binding assay.

KEYWORDS: Drug discovery, High-throughput-screening, Microscale thermophoresis, RNA methyltransferase DNMT2, Fluorescein-labeling
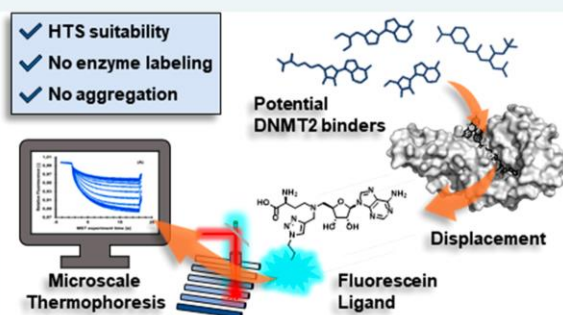
O riginally the human DNA methyltransferase 2 (DNMT2) was considered to be a DNA-modifying enzyme, as it shares most of its sequence with other members of the DNMT family (DNMT1, DNMT3A, DNMT3B).[1,2] However, in 2006, Goll et al. discovered that the main substrate of DNMT2 is tRNA. DNMT2 catalyzes the methylation of a cytosine at position 38 in the anticodon loop of tRNA^Asp. During this modification, the cofactor *S*-adenosylmethionine (SAM) is converted to *S*-adenosylhomocysteine (SAH), a known product inhibitor of DNMT2.[3] Subsequently, tRNA^Val and tRNA^Gly were also found to be substrates of DNMT2, underlining its role as an RNA methyltransferase, especially in humans.[4] Since then, a lot of possible physiological roles of DNMT2 were discussed. It was proven that DNMT2 together with other RNA-methyltransferases like NSUN2 has a critical influence on tRNA stability and protein translation in cells.[5–7] During oxidative stress, an up-regulation of DNMT2 could be detected, leading to the assumption that DNMT2 also helps cells coping with exogenous stress factors.[8–10] DNMT2 was as well found to be overexpressed in cancer cells, where mutations on DNMT2 are assumed to play a functional role in tumorigenesis.[11,12] Furthermore, DNMT2 is also involved in the epigenetic inheritance process.[13] Taken all this into account, DNMT2 appears to be a promising drug target for medicinal chemistry. Therefore, reliable methods to identify potent binders are required.

Tritium incorporation assays using ³H-SAM are still the most reliable and therefore often used methods to assay methyltransferases like DNMT2, despite all the disadvantages, such as time consumption and high costs.[14,15] Other methods, like LC-MS, still require a lot of working steps and time.[16] Therefore, fast, reliable, and cost-efficient screening methods are needed to reduce the number of compounds subjected to tritium incorporation or LC-MS assays to a minimum. At present, various biophysical methods are available for affinity-based screening assays such as isothermal titration calorimetry (ITC),[17] differential scanning fluorimetry (DSF),[18] surface plasmon resonance (SPR),[19] or microscale thermophoresis (MST).[20] All have their advantages but also their disadvantages. ITC measurements not only determine ligand affinities,

by measuring the binding enthalpy, but also reveal thermodynamic data of the binding as well as binding stoichiometry. Although this is a very robust method, the enormous amounts of protein and ligand needed can be a major obstacle.[21−23] DSF is a fast-screening method based on the measurement of protein melting curves in the presence of a fluorescent dye.[18,24,25] A disadvantage is the fact that some ligands induce only small thermal shifts, leading to a false negative result in a screening campaign.[26] Today, SPR is one key technology applied in pharmaceutical research.[27] Although it offers several advantages, such as fast, reliable, and label-free measurements, one cannot ignore that the necessity of surface immobilization remains a major challenge. Besides SPR, the significance of MST also emerges.[28−31] In contrast to SPR or ITC, thermodynamic properties cannot be analyzed using MST with a single measurement but require several measurements at different temperatures.[32] However, especially in early drug discovery, this sample-saving method can be of great benefit. It is based on the principle that molecules migrate in a temperature gradient. This so-called thermophoresis is reproducible at given conditions, but even small changes such as ligand binding can alter that behavior.[32] Furthermore, the change in the thermophoresis behavior correlates with the extent of this interaction. This allows an affinity determination of a protein−ligand interaction by observing fluorescence changes.[20,32,33] Besides thermophoresis, other effects can cause these changes, e.g., temperature differences and changes in the local environment of the fluorophore.[34,35] The advantages of this method are certainly its sensitivity and scalability, but also the fact that screenings can be performed directly with cell lysates or blood sera, which makes it a powerful method for drug discovery.[20,26,32,36−40] Determining binding affinity by MST, however, comes with some inherent problems. In most cases, the protein of interest must be fluorescently labeled, either in a covalent manner, by thiol or amide coupling, or in a non-covalent manner, by His-tag labeling with a fluorescent dye.[32,33] Direct measuring of protein−ligand interactions by MST only shows binding events but cannot reveal if the detected binding is competitive with respect to the active site of the protein, or if the ligand binding occurs on an allosteric site. It can also be very challenging to reach saturation conditions for the bound protein−ligand complex, especially if ligands only have binding affinities in the low single-digit micromolar range like the currently known inhibitors of DNMT2. Achieving saturation conditions often needs high protein and ligand concentrations resulting in protein and/or ligand aggregation and/or unspecific binding. Furthermore, the maximal effect of a ligand on the thermophoresis behavior of the protein−ligand complex is not predictable. Therefore, it is not uncommon, that low-affinity ligands have a stronger influence on the thermophoresis behavior than high-affinity ones.

The main cause of these problems is notably due to the necessity of fluorescent-labeling of the protein. If the fluorescent dye is not attached to the enzyme but rather a known ligand (fluorescence tracer)—especially one with a strong thermophoresis behavior—a displacement assay could be established to circumvent a lot of this problems. In this paper, we introduce an optimized MST method with excellent suitability for high-throughput-screening of DNMT2 ligands. We developed a displacement assay, using a literature-known fluorescein-based tool compound 5-FAM-triazolyl-adenosyl-Dab (**6**, FTAD, see Scheme 1), that originally was applied by

Luan et al. in a fluorescence polarization (FP) assay for the histone methyltransferase mixed-lineage leukemia 1 (MLL1).[41] Our method circumvents the inherent problems previously described, and provides the following improvements:

- reduced aggregation due to lower concentrations of ligand
- half-quantitative screenings outcomes
- determination of active site interactions
- no enzyme labeling required

■ RESULTS AND DISCUSSION

Chemistry. For the preparation of the tool compound, we followed a synthetic procedure similar to the one used by Luan et al. (Scheme 1).[41] First the azide **2** was prepared in a

Scheme 1. Synthesis of FTAD (6)[a]
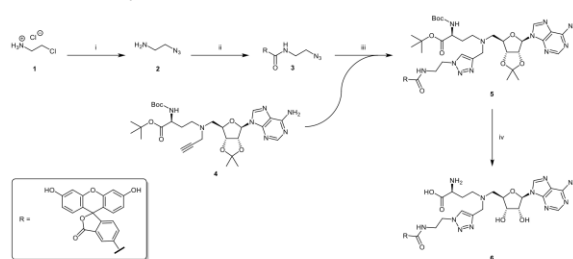
[a]Reagents and conditions: (i) NaN$_3$, H$_2$O, 85 °C, 24 h, 67%; (ii) 5-FAM, TPTU, DIPEA, DMF, 0−20 °C, 48 h, 96%; (iii) CuSO$_4$, sodium ascorbate, MeOH/H$_2$O, 16 h, 61%; (iv) 1. TFA, DCM, 5 °C; 2. TFA, H$_2$O, 5 °C, 99%.

nucleophilic substitution reaction of 2-chloroethylamine hydrochloride **1** and sodium azide. This product was then coupled to 5-carboxyfluorescein (5-FAM) using O-(2-oxo-1(2H)pyridyl)-N,N,N′,N′-tetramethyluronium tetrafluoroborate (TPTU) and N-ethyldiisopropylamine (DIPEA), yielding compound **3**. In the subsequent copper(I)-catalyzed azide−alkyne cycloaddition (CuAAC), the alkyne **4** was connected to the azide **3**, forming the triazole product **5**. The synthesis of building block **4** was carried out according to literature.[42] In the final step, all protecting groups were cleaved using a two-step procedure. First, 50% (v/v) trifluoroacetic acid (TFA in DCM was used at 5 °C, then 14% TFA in water at 5 °C to finally yield FTAD (**6**) as its trifluoroacetate salt.

Establishment of DNMT2 Fluorescence Polarization Assay. Binding affinity of FTAD (**6**) to DNMT2 was investigated using the fluorescence polarization assay protocol described for the histone methyltransferase MLL1.[41] Therefore, a saturation curve was measured, which resulted in a K_D value of 2.4 μM. These results indicated that this fluorescent probe is not only suitable to assay MLL1 but also DNMT2. To check the quality of this assay, the Z-factor was evaluated. The obtained value of 0.92 for the Z-factor is very good and sufficient to establish an assay for the determination of binding affinities.[43] Next, the literature known binding affinities of S-adenosylhomocysteine (**7**, SAH) and sinefungin (**8**, SFG) could be confirmed. K_D values of 12.2 μM and 6.5 μM for SAH and SFG were measured, which are quite similar to reported values from literature (K_D = 13.6 μM and K_D = 7.5 μM).[42] Furthermore, the binding affinity of a recently published DNMT2 inhibitor (**9**) was determined, revealing a K_D of 7.8

μM which was also found to fit to the published value of 8.1 μM.[42]

To investigate if this tool compound may be suitable for pan-methyltransferase assays, binding toward the human tRNA methyltransferases NSUN2 and NSUN6 as well as the histone methyltransferase EHMT2 of the KMT family was tested. Unfortunately, FTAD showed only weak affinities to those methyltransferases with approximate K_D values in the higher two-digit or the three-digit micromolar range (data not shown).

Nevertheless, these findings proved that this tool compound which originally was designed to assay MLL1 via FP was also suitable for DNMT2.

Establishment of DNMT2 Microscale Thermophoresis Assay. In a next step, FTAD (**6**) was used as a fluorescence tracer for a microscale thermophoresis (MST) displacement assay. The assay is based on the formation of a FTAD-DNMT2 complex, which allows fluorescence tracing. By adding a ligand of interest FTAD is displaced, which results in an altered fluorescence signal, since FTAD is no longer bound to DNMT2 (Scheme 2). Excitation and emission

Scheme 2. Establishment of DNMT2 Microscale Thermophoresis Assay[a]
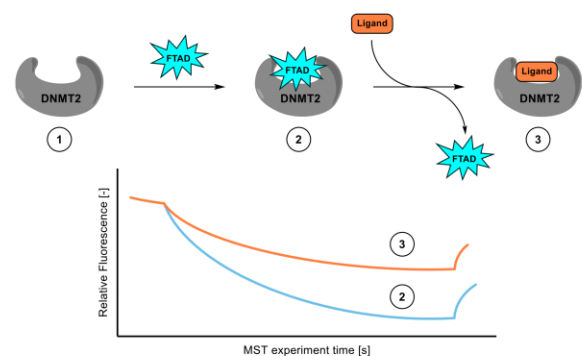
[a]Steps: (1) Unlabeled DNMT2 cannot be observed by MST. (2) DNMT2 is pre-incubated with FTAD; the FTAD-DNMT2 complex can be observed by MST. (3) Displacement of FTAD with ligand of interest; free FTAD can be observed by MST.
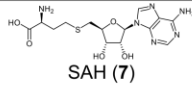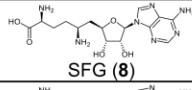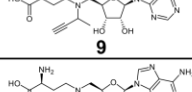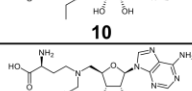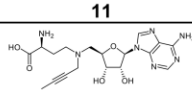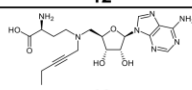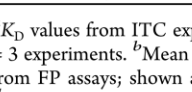
wavelengths of FTAD were sufficient for MST measurements with blue light settings (excitation: 465−490 nm and emission: 500−550 nm). At first, an assay concentration for FTAD was evaluated, starting with 50 nM, as described in the FP protocol. Since the fluorescence signal was not optimal, the concentration was increased. Finally, a concentration of 100 nM was chosen, which allowed measurements with sufficient fluorescence signal, at the lowest possible concentration. Furthermore, no surface absorption or bleaching was detected in standard capillaries. In a dilution series of DNMT2 at constant FTAD concentrations, binding affinity of FTAD against DNMT2 was investigated using MST, revealing a K_D value of 1.8 μM, which is in very good accordance with the binding affinity determined using FP (Figure 1A,B). Furthermore, a Z-factor of 0.90 (Figure 1C) for this assay indicated that this assay had a very good quality. In the next step, the binding affinities of SAH and SFG were verified by MST. Since displacement assays only provide EC$_{50}$ values, those results were corrected for the binding competition with the fluorescence tracer according to the instructions of
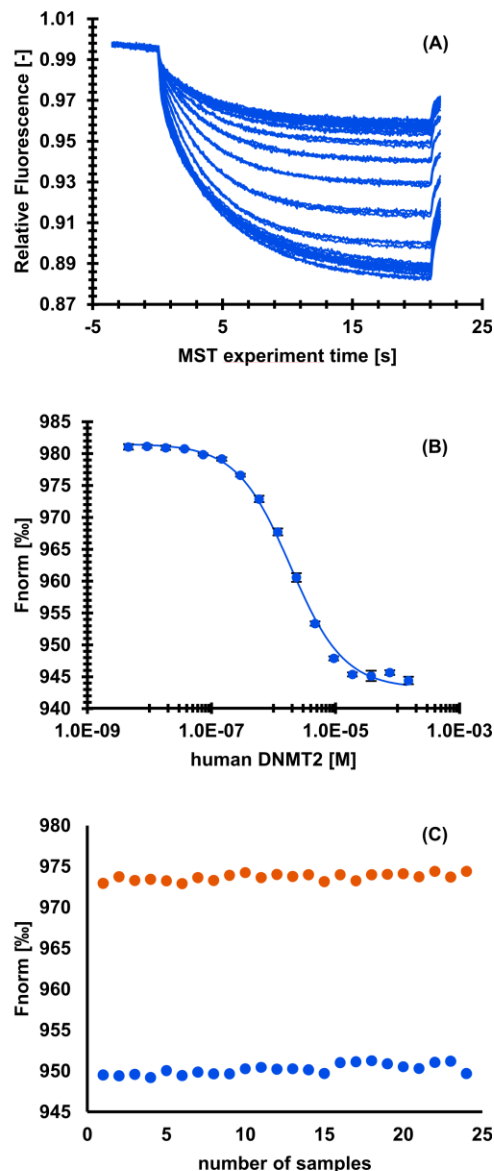
Figure 1. Microscale thermophoresis assay for FTAD against a dilution series of human DNMT2. (A) Raw traces of runs. (B) Normalized fluorescence plotted against concentration of human DNMT2. Data is given as mean ± SD of triplicates. K_D of FTAD against DNMT2 was found to be 1.8 μM ± 0.1 μM. (C) Z-factor determination of microscale thermophoresis assay. In orange, free FTAD without DNMT2; in blue, FTAD in the presence of 2 μM human DNMT2.

Nanotemper Technologies to receive K_D values. For SAH, K_D = 6.5 μM was determined, while for SFG, K_D = 9.2 μM was found. Both values were in high accordance with literature, where K_D values of 13.6 μM and 7.5 μM, respectively, were described (Table 1). For compound **9**, a K_D value of 5.2 μM was found, which correlates quite well with the literature value (K_D = 8.1 μM[42]). Interestingly, also for compound **10**, a literature known non-inhibitor of DNMT2,[42] a very slight protein−ligand interaction could be detected in this assay with

Table 1. Binding Affinities of Different Ligands toward DNMT2 Measured by Orthogonal Biophysical Methods

Compound	ITC[a] / μM	FP[c] / μM	MST[d] / μM
SAH (7)	13.6 ± 4.4[b]	12.2 ± 0.9	6.5 ± 1.4
SFG (8)	7.5 ± 3.5	6.5 ± 0.5	9.2 ± 0.5
9	8.1 ± 1.4	7.8 ± 1.1	5.2 ± 0.6
10	n.d.*	n.d.	> 70
11	11.4 ± 2.4	n.d.	15.3 ± 3.1
12	10.2 ± 2.2	n.d.	11.3 ± 0.8
13	10.5 ± 3.3	n.d.	11.0 ± 1.1

[a]K_D values from ITC experiments; shown are mean values ± SD of n = 3 experiments. [b]Mean values ± SD of n = 6 experiments. [c]K_D values from FP assays; shown are mean values ± SD of n = 3 experiments. [d]K_D values from MST displacement assays; shown are mean values ± SD of n = 4 experiments. *n.d. = could not be determined due to the high concentrations of enzyme and ligand required to quantify low-affinity binding.

an EC_{50} of >100 μM, this indicates that the method is even capable to identify very low affinity binders if needed. To further confirm the potential of this assay for affinity determination, three additional SAH-derived DNMT2 inhibitors, 11−13,[42] were selected and the K^D values were determined by the MST displacement assay and by ITC as an orthogonal method.

The data are in very good agreement, proving the reliability of the assay. Moreover, in comparison to the ITC method, the amounts of samples, especially of protein, are much lower.

One of the most considerable findings was the fact that with this displacement method the plateaus for total unbound and bound protein could be resolved very well (Figure 2). Since the maxima of the shifts can easily be measured by control runs, the fits derived from those measurements are far more reliable than those of direct binding assays based on MST. Aggregation, up to concentrations needed to reach a plateau, was not observed in any run. Furthermore, the signal-to-noise ratio was highly increased compared to direct measurements of labeled DNMT2. This may derive from the largely increased shift, which is probably due to the fact, that the large DNMT2 molecule influences the thermophoresis behavior of the small fluorescence tracer much more, than vice versa a small molecular ligand a >40 kDa protein like DNMT2.

MST can be used as a fast and efficient screening method.[42] Although common MST assays using labeled protein are highly

sensitive screening methods that can detect binders with almost negligible affinity, there remain some problems. One is for sure the comparability of the detected shifts. The extent of the included shifts in thermophoresis does not correlate very well with the affinity of different ligands; i.e., a less affine ligand can induce a larger shift than a ligand with higher affinity. In our recently published structure−activity relationship (SAR) study for SAH-derived DNMT2 inhibitors, MST was used as an initial screening method. Some ligands were found to induce noteworthy shifts in the thermophoretic behavior, but a tritium incorporation assay revealed that a lot of those ligands showed only poor inhibitory properties.[42] By plotting the induced shifts against the found inhibition, no correlation could be detected (Figure 3A).

Therefore, a screening with our novel displacement method was performed for a set of binders and non-binders of this SAR study. A ligand concentration of 20 μM was found to be suitable for this screening. Since for a displacement assay the detectable shift is known from the beginning and depends only on the fraction of bound and unbound fluorescence tracer, this seemed to open the possibility to achieve a correlation between the thermophoresis shifts and the actual inhibition of DNMT2. Indeed, a correlation between the thermophoresis shifts in this displacement assay and the measured inhibition from the tritium incorporation assay could be found (Figure 3B). This led to the conclusion that ligands inducing a thermophoretic shift <5‰ should not be considered for further measurements, while ligands inducing shifts ~10‰ and higher are likely to be very promising DNMT2 inhibitors. The displacement assay not only allows researchers to discriminate good and poor binders. Given the fact that FTAD binds to the active site of DNMT2, this assay also indicates if the ligand is active-site directed, which can be very helpful for the screening of DNMT2 inhibitors that structurally do not resemble the natural ligand. It is noteworthy that the displacement assay can be performed at decreased ligand concentrations in a range of 20 μM instead of 100 μM, which can prevent unspecific binding to the protein and therefore misleading hits from the beginning.

■ CONCLUSION

Within this study we present a novel method for a fast and reliable ligand screening for the human RNA methyltransferase DNMT2 using microscale thermophoresis. The assay can easily be extended to a high-throughput screening. This displacement assay allows screenings at low ligand, dye, and protein concentrations. Together with the small volumes required for microscale thermophoresis, this results in a cost-effective but still robust method. Under the conditions described, no adsorption, bleaching, or aggregation was detected, which usually can cause severe interferences in direct microscale thermophoresis assays. Furthermore, this assay reveals accurate screening results, which facilitates early-stage drug discovery in the field of DNMT2. Due to the fact that the fluorescence tracer is bound to the active site of DNMT2, thermophoresis shifts detected by this method always indicate active-site targeting, making further binding-site verification unnecessary. As a result of the increased thermophoresis shifts compared to direct microscale thermophoresis assays with labeled protein, this method can be used for reliable binding affinity determination up to the low micromolar range. In summary, we are highly convinced that this method can be a
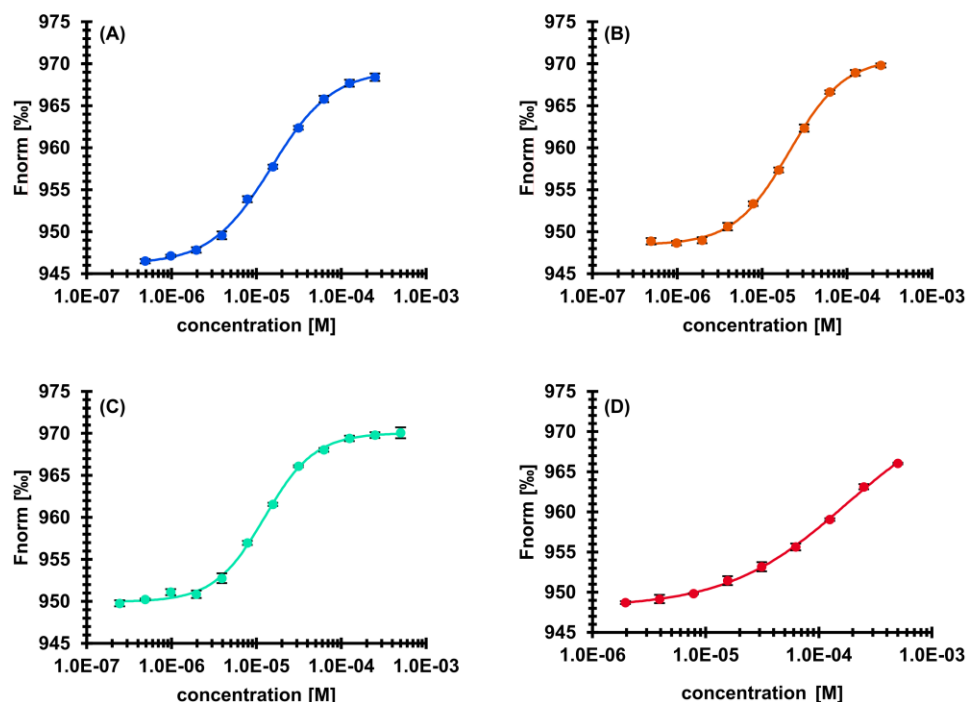
Figure 2. Displacement of FTAD by literature known inhibitors in a microscale thermophoresis assay. Data is given as mean ± SD of quadruplicates. (A) SAH (**7**), EC$_{50}$ = 15.0 ± 0.7 μM; (B) SFG (**8**), EC$_{50}$ = 20.9 ± 0.7 μM; (C) compound **9**, EC$_{50}$ = 12.4 ± 0.5 μM; and (D) compound **10**, EC$_{50}$ = > 100 μM.
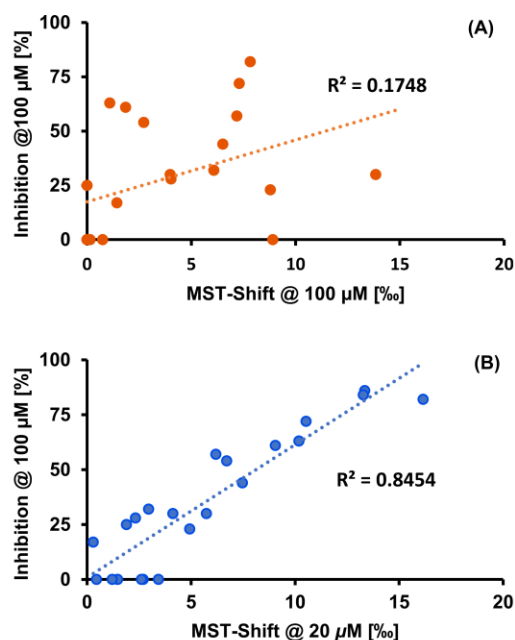
Figure 3. Correlation of MST shifts at 20 μM compound concentrations and inhibition at 100 μM in a tritium incorporation assay for literature-known compounds: (A) for direct MST assay and (B) for MST displacement assay.

powerful tool accelerating the search for potent DNMT2 inhibitors.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsptsci.2c00175.

Microscale thermophoresis raw traces, NMR spectra, and chromatograms (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

Authors

Robert Alexander Zimmermann − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ⓞ orcid.org/0000-0002-5330-9234

Marvin Schwickert − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ⓞ orcid.org/0000-0002-1385-1416

J. Laurenz Meidner − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ⓞ orcid.org/0000-0001-5488-0646

Zarina Nidoieva − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ⓞ orcid.org/0000-0001-5364-7527

Mark Helm − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany;* ⊙ orcid.org/0000-0002-0154-0928

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsptsci.2c00175

Author Contributions

‡R.A.Z., M.S., and J.L.M. contributed equally to the project.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

CCR2, CC chemokine receptor 2; CCL2, CC chemokine ligand 2; CCR5, CC chemokine receptor 5; Dab, 2,4-diaminobutyric acid; 5-FAM, 5-carboxyfluorescein; FTAD, 5-FAM-triazolyl-adenosyl-Dab; TLC, thin-layer chromatography

■ REFERENCES

(1) Van den Wyngaert, I.; Sprengel, J.; Kass, S. U.; Luyten, W. H. Cloning and Analysis of a Novel Human Putative DNA Methyltransferase. *FEBS Lett.* **1998**, *426* (2), 283−289.

(2) Dong, A. Structure of Human DNMT2, an Enigmatic DNA Methyltransferase Homolog That Displays Denaturant-Resistant Binding to DNA. *Nucleic Acids Res.* **2001**, *29* (2), 439−448.

(3) Goll, M. G.; Kirpekar, F.; Maggert, K. A.; Yoder, J. A.; Hsieh, C. L.; Zhang, X.; Golic, K. G.; Jacobsen, S. E.; Bestor, T. H. Methylation of TRNAAsp by the DNA Methyltransferase Homolog Dnmt2. *Science* **2006**, *311* (5759), 395−398.

(4) Schaefer, M.; Pollex, T.; Hanna, K.; Tuorto, F.; Meusburger, M.; Helm, M.; Lyko, F. RNA Methylation by Dnmt2 Protects Transfer RNAs against Stress-Induced Cleavage. *Genes Dev.* **2010**, *24* (15), 1590−1595.

(5) Tuorto, F.; Liebers, R.; Musch, T.; Schaefer, M.; Hofmann, S.; Kellner, S.; Frye, M.; Helm, M.; Stoecklin, G.; Lyko, F. RNA Cytosine Methylation by Dnmt2 and NSun2 Promotes TRNA Stability and Protein Synthesis. *Nat. Struct. Mol. Biol.* **2012**, *19* (9), 900−905.

(6) Shanmugam, R.; Fierer, J.; Kaiser, S.; Helm, M.; Jurkowski, T. P.; Jeltsch, A. Cytosine Methylation of TRNA-Asp by DNMT2 Has a Role in Translation of Proteins Containing Poly-Asp Sequences. *Cell Discovery* **2015**, *1*, 15010.

(7) Tuorto, F.; Herbst, F.; Alerasool, N.; Bender, S.; Popp, O.; Federico, G.; Reitter, S.; Liebers, R.; Stoecklin, G.; Gröne, H.; Dittmar, G.; Glimm, H.; Lyko, F. The TRNA Methyltransferase Dnmt2 Is Required for Accurate Polypeptide Synthesis during Haematopoiesis. *EMBO J.* **2015**, *34* (18), 2350−2362.

(8) Becker, M.; Müller, S.; Nellen, W.; Jurkowski, T. P.; Jeltsch, A.; Ehrenhofer-Murray, A. E. Pmt1, a Dnmt2 Homolog in Schizosaccharomyces Pombe, Mediates TRNA Methylation in Response to Nutrient Signaling. *Nucleic Acids Res.* **2012**, *40* (22), 11648−11658.

(9) Lin, M.-J.; Tang, L.-Y.; Reddy, M. N.; Shen, C.-K. J. DNA Methyltransferase Gene DDnmt2 and Longevity of Drosophila. *J. Biol. Chem.* **2005**, *280* (2), 861−864.

(10) Kaul, G.; Thippeswamy, H. Role of Heat Shock Proteins in Diseases and Their Therapeutic Potential. *Indian J. Microbiol.* **2011**, *51* (2), 124−131.

(11) Forbes, S. A.; Beare, D.; Gunasekaran, P.; Leung, K.; Bindal, N.; Boutselakis, H.; Ding, M.; Bamford, S.; Cole, C.; Ward, S.; Kok, C. Y.; Jia, M.; De, T.; Teague, J. W.; Stratton, M. R.; McDermott, U.; Campbell, P. J. COSMIC: Exploring the World's Knowledge of Somatic Mutations in Human Cancer. *Nucleic Acids Res.* **2015**, *43*, D805−11.

(12) Elhardt, W.; Shanmugam, R.; Jurkowski, T. P.; Jeltsch, A. Somatic Cancer Mutations in the DNMT2 TRNA Methyltransferase Alter Its Catalytic Properties. *Biochimie* **2015**, *112*, 66−72.

(13) Zhang, Y.; Zhang, X.; Shi, J.; Tuorto, F.; Li, X.; Liu, Y.; Liebers, R.; Zhang, L.; Qu, Y.; Qian, J.; Pahima, M.; Liu, Y.; Yan, M.; Cao, Z.; Lei, X.; Cao, Y.; Peng, H.; Liu, S.; Wang, Y.; Zheng, H.; Woolsey, R.; Quilici, D.; Zhai, Q.; Li, L.; Zhou, T.; Yan, W.; Lyko, F.; Zhang, Y.; Zhou, Q.; Duan, E.; Chen, Q. Dnmt2Mediates Intergenerational Transmission of Paternally Acquired Metabolic Disorders through Sperm Small Non-Coding RNAs. *Nat. Cell Biol.* **2018**, *20* (5), 535−540.

(14) Jurkowski, T. P.; Meusburger, M.; Phalke, S.; Helm, M.; Nellen, W.; Reuter, G.; Jeltsch, A. Human DNMT2Methylates TRNA(Asp) Molecules Using a DNA Methyltransferase-like Catalytic Mechanism. *RNA* **2008**, *14* (8), 1663−1670.

(15) Müller, S.; Windhof, I. M.; Maximov, V.; Jurkowski, T.; Jeltsch, A.; Förstner, K. U.; Sharma, C. M.; Gräf, R.; Nellen, W. Target Recognition, RNA Methylation Activity and Transcriptional Regulation of the Dictyostelium Discoideum Dnmt2-Homologue (DnmA). *Nucleic Acids Res.* **2013**, *41* (18), 8615−8627.

(16) Salyan, M. E. K.; Pedicord, D. L.; Bergeron, L.; Mintier, G. A.; Hunihan, L.; Kuit, K.; Balanda, L. A.; Robertson, B. J.; Feder, J. N.; Westphal, R.; Shipkova, P. A.; Blat, Y. A General Liquid Chromatography/Mass Spectroscopy-Based Assay for Detection and Quantitation of Methyltransferase Activity. *Anal. Biochem.* **2006**, *349* (1), 112−117.

(17) Szlag, V. M.; Jung, S.; Rodriguez, R. S.; Bourgeois, M.; Bryson, S.; Schatz, G. C.; Reineke, T. M.; Haynes, C. L. Isothermal Titration Calorimetry for the Screening of Aflatoxin B1 Surface-Enhanced Raman Scattering Sensor Affinity Agents. *Anal. Chem.* **2018**, *90* (22), 13409−13418.

(18) Gao, K.; Oerlemans, R.; Groves, M. R. Theory and Applications of Differential Scanning Fluorimetry in Early-Stage Drug Discovery. *Biophys. Rev.* **2020**, *12* (1), 85−104.

(19) Olaru, A.; Bala, C.; Jaffrezic-Renault, N.; Aboul-Enein, H. Y. Surface Plasmon Resonance (SPR) Biosensors in Pharmaceutical Analysis. *Crit. Rev. Anal. Chem.* **2015**, *45* (2), 97−105.

(20) Jerabek-Willemsen, M.; Wienken, C. J.; Braun, D.; Baaske, P.; Duhr, S. Molecular Interaction Studies Using Microscale Thermophoresis. *Assay Drug Dev. Technol.* **2011**, *9* (4), 342−353.

(21) Bastos, M.; Velazquez-Campoy, A. Isothermal Titration Calorimetry (ITC): A Standard Operating Procedure (SOP). *Eur. Biophys. J.* **2021**, *50* (3−4), 363−371.

(22) Baranauskiene, L.; Kuo, T.-C.; Chen, W.-Y.; Matulis, D. Isothermal Titration Calorimetry for Characterization of Recombinant Proteins. *Curr. Opin. Biotechnol.* **2019**, *55*, 9−15.

(23) Kabiri, M.; Unsworth, L. D. Application of Isothermal Titration Calorimetry for Characterizing Thermodynamic Parameters of Biomolecular Interactions: Peptide Self-Assembly and Protein Adsorption Case Studies. *Biomacromolecules* **2014**, *15* (10), 3463−3473.

(24) Niesen, F. H.; Berglund, H.; Vedadi, M. The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability. *Nat. Protoc.* **2007**, *2* (9), 2212−2221.

(25) Zhang, R.; Monsma, F. Fluorescence-Based Thermal Shift Assays. *Curr. Opin. Drug Discovery Devel.* **2010**, *13* (4), 389−402.

(26) Linke, P.; Amaning, K.; Maschberger, M.; Vallee, F.; Steier, V.; Baaske, P.; Duhr, S.; Breitsprecher, D.; Rak, A. An Automated Microscale Thermophoresis Screening Approach for Fragment-Based Lead Discovery. *J. Biomol. Screen.* **2016**, *21* (4), 414−421.

(27) Hinman, S. S.; McKeating, K. S.; Cheng, Q. Surface Plasmon Resonance: Material and Interface Design for Universal Accessibility. *Anal. Chem.* **2018**, *90* (1), 19−39.

(28) Rainard, J. M.; Pandarakalam, G. C.; McElroy, S. P. Using Microscale Thermophoresis to Characterize Hits from High-Throughput Screening: A European Lead Factory Perspective. *SLAS Discovery Adv. life Sci. R D* **2018**, *23* (3), 225−241.

(29) 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.

(30) Kozielski, F.; Sele, C.; Talibov, V. O.; Lou, J.; Dong, D.; Wang, Q.; Shi, X.; Nyblom, M.; Rogstam, A.; Krojer, T.; Fisher, Z.; Knecht, W. Identification of Fragments Binding to SARS-CoV-2 Nsp10 Reveals Ligand-Binding Sites in Conserved Interfaces between Nsp10 and Nsp14/Nsp16. *RSC Chem. Biol.* **2022**, *3* (1), 44−55.

(31) Feoli, A.; Pisapia, V.; Viviano, M.; Castellano, S.; Bartoschik, T.; Sbardella, G. Development of a Microscale Thermophoresis-Based Method for Screening and Characterizing Inhibitors of the Methyl-Lysine Reader Protein MRG15. *SLAS Discovery Adv. life Sci. R D* **2021**, *26* (1), 77−87.

(32) Seidel, S. A. I.; Dijkman, P. M.; Lea, W. A.; van den Bogaart, G.; Jerabek-Willemsen, M.; Lazic, A.; Joseph, J. S.; Srinivasan, P.; Baaske, P.; Simeonov, A.; Katritch, I.; Melo, F. A.; Ladbury, J. E.; Schreiber, G.; Watts, A.; Braun, D.; Duhr, S. Microscale Thermophoresis Quantifies Biomolecular Interactions under Previously Challenging Conditions. *Methods* **2013**, *59* (3), 301−315.

(33) Bartoschik, T.; Gupta, A.; Kern, B.; Hitchcock, A.; Adams, N. B. P.; Tschammer, N. Quantifying the Interaction of Phosphite with ABC Transporters: MicroScale Thermophoresis and a Novel His-Tag Labeling Approach. *Methods Mol. Biol.* **2020**, *2168*, 51−62.

(34) López-Méndez, B.; Uebel, S.; Lundgren, L. P.; Sedivy, A. Microscale Thermophoresis and Additional Effects Measured in NanoTemper Monolith Instruments. *Eur. Biophys. J.* **2021**, *50* (3−4), 653−660.

(35) Royer, C. A. Probing Protein Folding and Conformational Transitions with Fluorescence. *Chem. Rev.* **2006**, *106* (5), 1769−1784.

(36) Wienken, C. J.; Baaske, P.; Rothbauer, U.; Braun, D.; Duhr, S. Protein-Binding Assays in Biological Liquids Using Microscale Thermophoresis. *Nat. Commun.* **2010**, *1*, 100.

(37) Seidel, S. A. I.; Wienken, C. J.; Geissler, S.; Jerabek-Willemsen, M.; Duhr, S.; Reiter, A.; Trauner, D.; Braun, D.; Baaske, P. Label-Free Microscale Thermophoresis Discriminates Sites and Affinity of Protein-Ligand Binding. *Angew. Chem., Int. Ed. Engl.* **2012**, *51* (42), 10656−10659.

(38) Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. MicroScale Thermophoresis: Interaction Analysis and Beyond. *J. Mol. Struct.* **2014**, *1077*, 101−113.

(39) Bartoschik, T.; Galinec, S.; Kleusch, C.; Walkiewicz, K.; Breitsprecher, D.; Weigert, S.; Muller, Y. A.; You, C.; Piehler, J.; Vercruysse, T.; Daelemans, D.; Tschammer, N. Near-Native, Site-Specific and Purification-Free Protein Labeling for Quantitative Protein Interaction Analysis by MicroScale Thermophoresis. *Sci. Rep.* **2018**, *8* (1), 4977.

(40) Magnez, R.; Thiroux, B.; Taront, S.; Segaoula, Z.; Quesnel, B.; Thuru, X. PD-1/PD-L1 Binding Studies Using Microscale Thermophoresis. *Sci. Rep.* **2017**, *7* (1), 17623.

(41) Luan, Y.; Blazer, L. L.; Hu, H.; Hajian, T.; Zhang, J.; Wu, H.; Houliston, S.; Arrowsmith, C. H.; Vedadi, M.; Zheng, Y. G. Design of a Fluorescent Ligand Targeting the S-Adenosylmethionine Binding Site of the Histone Methyltransferase MLL1. *Org. Biomol. Chem.* **2016**, *14* (2), 631−638.

(42) Schwickert, M.; Fischer, T. R.; Zimmermann, R. A.; Hoba, S. N.; Meidner, J. L.; Weber, M.; Stark, M. M.; Koch, J.; Jung, N.; Kersten, C.; Windbergs, M.; Lyko, F.; Helm, M.; Schirmeister, T. Discovery of Inhibitors of DNA Methyltransferase 2, an Epitranscriptomic Modulator and Potential Target for Cancer Treatment. *J. Med. Chem.* **2022**, *65* (14), 9750−9788.

(43) Zhang, J. H.; Chung, T. D.; Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **1999**, *4* (2), 67−73.

3.8.3   Experimental section of the publication

Note: Experimental procedures conducted by Marvin Schwickert (⬛⬛⬛⬛⬛⬛⬛⬛ group) are described in this section. Experiments performed by others can be found in the complete Supporting Information file in the **Appendix**. Compound numbers in this section refer to the numbering in the corresponding manuscript. The synthesis of compound **4** is described **Section 3.4.2**.

General information

Reagents and solvents were of commercial quality and were used without further purification. Reaction progress was monitored by thin layer chromatography using ALUGRAM® Xtra SIL G UV$_{254}$ silica plates from *Machery-Nagel*. For compound purification by column chromatography silica (40–63 μm) from *Machery-Nagel* was used. NMR spectra were recorded at 300 MHz on a *Bruker Fourier 300* and at 600 MHz on a *Bruker Avance III 600*. Chemical shifts are indicated in parts per million (ppm), using the solvent resonance (CDCl$_3$, DMSO-d_6, or CD$_3$OD from *Deutero GmbH*) as internal standard. The identity and purity of the final compound was determined by combined HPLC/ESI-MS analysis using an *Agilent 1100 series* HPLC system with an *Agilent Zorbax SB-Aq* (4.6 × 150 mm, 5 μm; mobile phase: MeCN/H$_2$O = 30:70 + 0.1% HCOOH; flow rate: 0.7 mL/min) column. The sample was applied using 5 μL injection with quantitation by AUC at 210 nm, 254 nm, and 280 nm. 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. Specific rotations were determined with a *Krüss P3000* polarimeter and are given in deg cm^3 g^{-1} dm^{-1}. The purity of all compounds tested in the described biological assays was ≥95% as determined by LC-MS.

(S)-2-Amino-4-(((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)((1-(2-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthene]-5-carboxamido)-ethyl)-1H-1,2,3-triazol-4-yl)methyl)amino)butanoic acid, trifluoroacetate salt (6, FTAD)
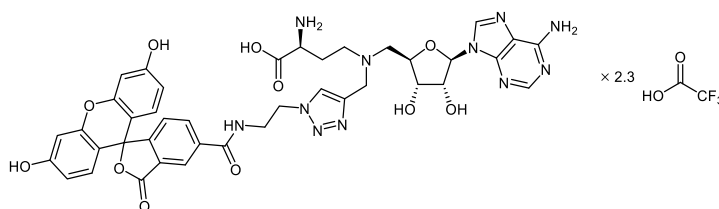
To a solution of compound **5** (182 mg, 0.17 mmol) in DCM (4 mL) at 5 °C was added TFA (4 mL). The solution was kept at 5 °C for 24 h. After the mixture was diluted and co-distilled with DCM (3 × 40 mL), the residue was dissolved in H$_2$O (3 mL) and cooled to 5 °C. TFA (0.5 mL) was added, and the solution was kept at 5 °C for 5 d. The mixture was diluted with water (12 mL) and dried by lyophilization to give FTAD (**6**, 194 mg, 0.17 mmol, 99%, 2.3 equiv. TFA) as an orange trifluoroacetate salt. ^{1}H NMR (600 MHz, CD$_3$OD): δ / ppm = 8.31–8.19 (m, 3H), 8.16–8.11 (m, 1H), 8.08–8.03 (m, 1H), 7.23 (d, J =

8.0 Hz, 1H), 6.65–6.62 (m, 2H), 6.53–6.44 (m, 4H), 6.05–5.98 (m, 1H), 4.69–4.58 (m, 3H), 4.51–4.43 (m, 1H), 4.38–4.25 (m, 3H), 3.89–3.79 (m, 3H), 3.48–3.40 (m, 1H), 3.36–3.27 (m, 2H), 2.36–2.25 (m, 1H), 2.13–2.01 (m, 1H), 1.33 − 1.29 (m, 1H). ^{13}C NMR (150 MHz, CD$_3$OD): δ / ppm = 168.4, 166.5, 159.3, 154.5, 151.8, 147.8, 140.9, 135.3, 133.3, 128.0, 126.4, 125.5, 123.6, 122.8, 111.6, 108.5, 101.4, 89.2, 79.1, 72.5, 71.3, 55.3, 54.4, 53.6, 52.1, 50.9, 49.1, 45.7, 39.1, 24.5, 16.5, 15.1. FTIR: ν / cm^{-1} = 3090, 2609, 1742, 1670, 1506, 1452, 1317, 1248, 1178, 1128, 994, 836, 798, 761, 720. mp: 67 °C (decomposition). $[\alpha]_D^{20}$ = +17 (10 mg/mL; MeOH). ESI-MS: m/z calculated for [C$_{40}$H$_{39}$N$_{11}$O$_{11}$+2H]$^{2+}$ ([M+2H]$^{2+}$) = 425.7; found: 425.6. Purity: 95% (HPLC, MeCN/H$_2$O = 30:70 + 0.1% HCOOH); t_R = 2.38 min.

3.9 Additional results

3.9.1 Modification of 4-bromo-3-nitrophenylsulfonyl-adenosyl-Dab to address additional subpockets

Based on the most promising inhibitor, compound **71i** (**Section 3.5**), further modifications were considered to potentially enhance inhibition of DNMT2. Regarding the binding pose of **71i** (**Figure 23**), a subpocket above the adenine site is visible that is not targeted by the inhibitor. By positioning a phenyl moiety into the subpocket, hydrophobic and aromatic interactions with Pro78 and Tyr10 as well as induced dipole interactions with Ser98 might be achieved and could contribute to the overall affinity and selectivity profile. A small methylene unit was chosen as a linker, yielding the corresponding N^6-benzyl derivative **85** of compound **71i**.
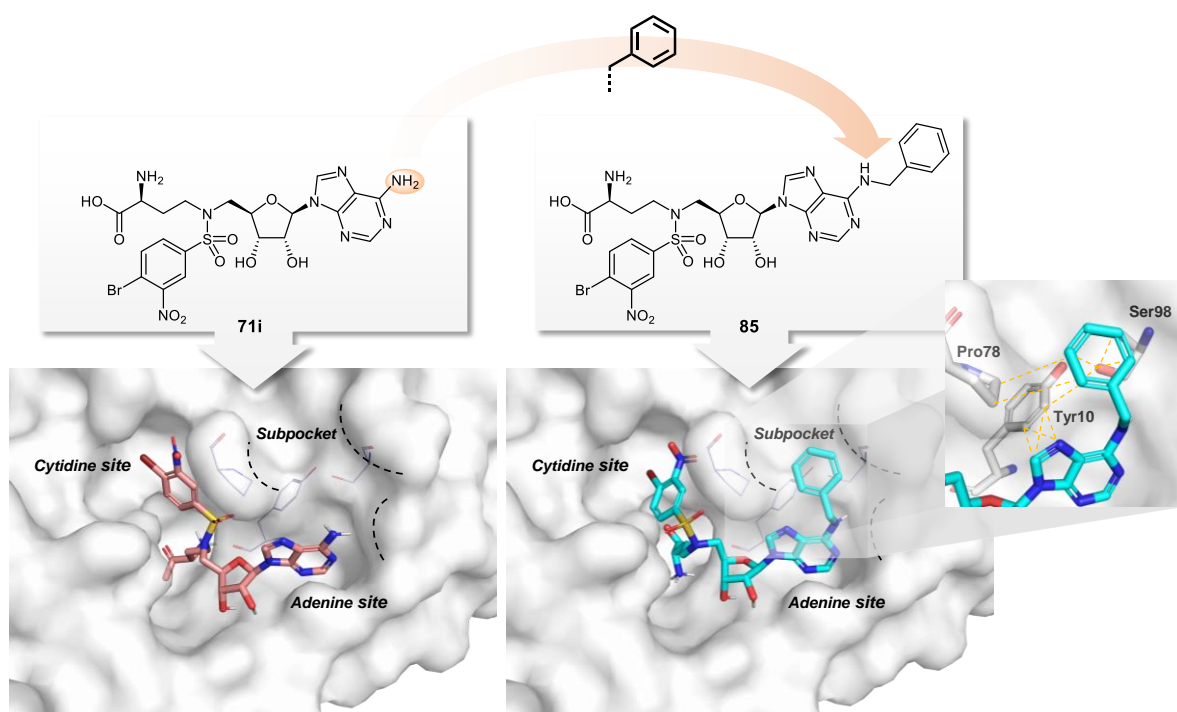
Figure 23: Docking poses of compound **71i** (salmon, left) and compound **85** (cyan, right) in DNMT2 (PDB-ID: 1G55)[197] with illustration of potential interactions of the benzyl group inside the subpocket.

The synthesis of compound **85** was carried out starting from commercially available N^6-benzyladenosine (**86**), which was protected using acetone in presence of *p*-toluenesulfonic acid yielding **87** (**Scheme 23**). Interestingly, **87** could also be prepared according to a procedure from SARFATI *et al.* by alkylating 2',3'-*O*-isopropylideneadenosine (**33**) with benzyl bromide at position $N1$.[288] The resulting intermediate **88** was treated with dimethylamine (DMA) in methanol to generate **87** in a DIMROTH rearrangement. **87** was brought to reaction with diphenylphosphoryl azide (DPPA) under basic conditions, followed by sodium azide in presence of 15-crown-5 to give the azide **89**. Hydrogenation of **89** yielded the primary amine **90**, which was reacted with the aldehyde **45** in a reductive amination using sodium triacetoxyboro-

hydride. The secondary amine **91** was treated with 4-bromo-3-nitrobenzenesulfonyl chloride (**92**) in a two-phase system consisting of DCM and saturated NaHCO₃ solution. In the last step, the resulting sulfonamide **93** was treated with 50 vol% TFA in dichloromethane at 5 °C, followed by treatment with 14 vol% TFA in water at 5 °C to give the final inhibitor **85**.
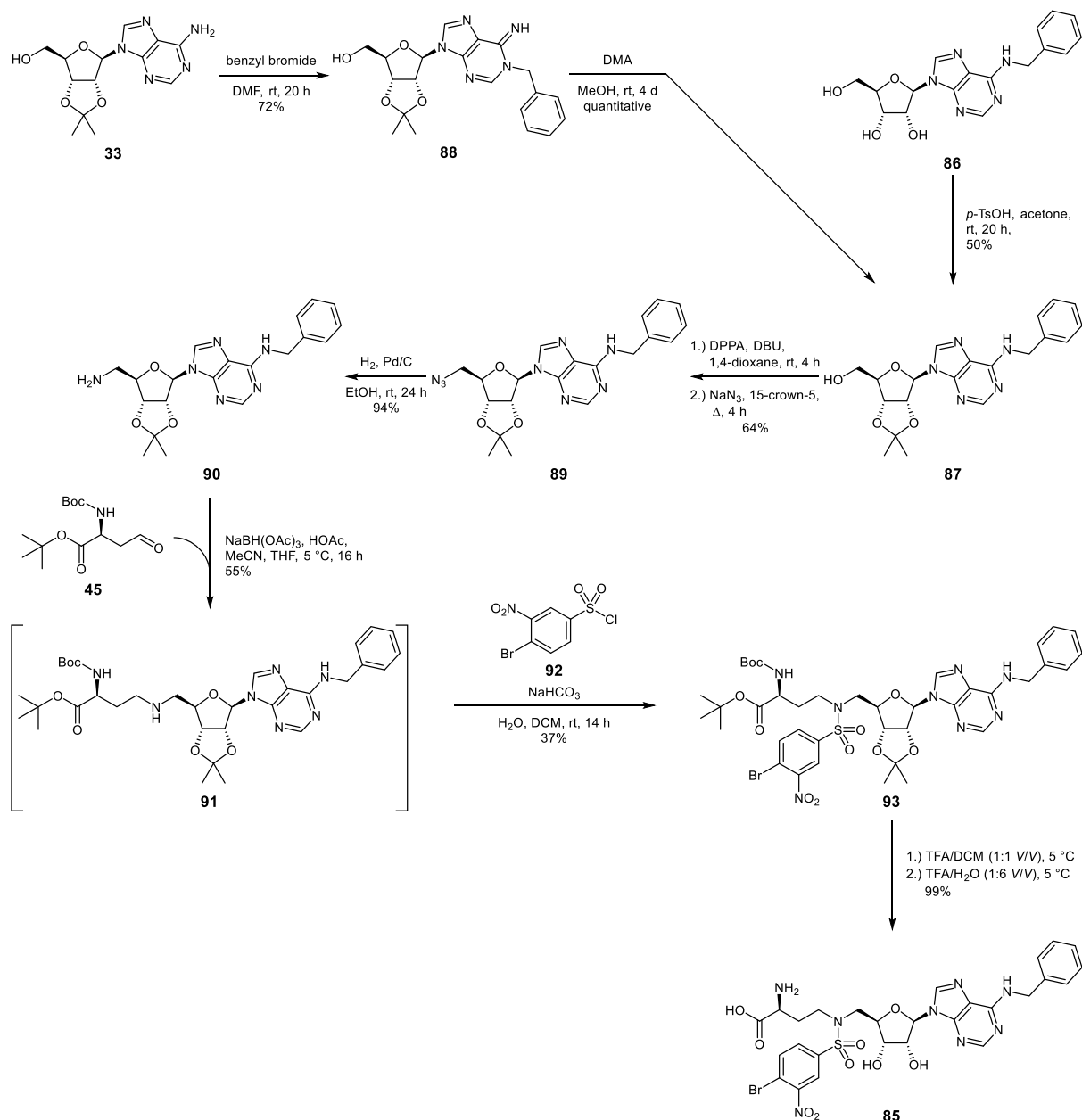
Scheme 23: Synthesis of compound **85**.

Compound **85** was evaluated for its binding affinity to DNMT2 by ▓▓▓▓▓▓▓▓▓▓ (▓▓▓▓▓▓▓ group) using an MST displacement method (**Section 3.8**). The inhibitory potential was determined in a tritium incorporation assay by ▓▓▓▓▓▓▓▓ (▓▓▓▓▓▓ group). Results are given in **Table 7**.

Table 7: Binding of compounds **71i** and **85** to DNMT2 as determined by MST and inhibition of DNMT2 as determined in the tritium incorporation assay.

71i, 85
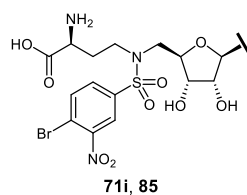
Compound		MST*	³H-Assay	
		K_D^{app} / μM	% inhibition at 100 μM	IC_{50} / μM
71i		2.3 ± 0.1	98.7 ± 1.1	1.6 ± 0.2
85		11.2 ± 1.1	91.0 ± 1.9	30.6 ± 2.2

*Determined by MST-based displacement method using a fluorescent dye (for details see **Section 3.8**).

The attachment of a benzyl group to *N*6 of compound **71i** resulted in a decrease in inhibition (IC_{50} = 30.6 ± 2.2 μM vs. 1.6 ± 0.2 μM) and slightly lower binding affinity (K_D^{app} = 11.2 ± 1.1 μM vs. 2.3 ± 0.1 μM). While conducting the assays, an unsatisfactory water solubility was observed. The deterioration of the obtained results might be due to solubility rather than to adverse interactions inside the binding site caused by the benzyl moiety. However, this project was not further pursued.

3.9.2 Method establishment for synthesis of potential prodrugs of SAH-based inhibitors

SAH-based inhibitors exhibit poor pharmacokinetic properties since they bear several polar functional groups, such as the vicinal hydroxy groups at the ribose scaffold, the amine function of adenine, as well as the amine and carboxyl groups of the amino acid side chain. The latter two make cell permeability an even bigger challenge as both functions – an acidic and a basic function – dissociate at physiological pH of 7.4 resulting in a zwitterionic substructure. In order to reach sufficient cell permeability, at least one of these dissociating functions should be capped by a suitable prodrug group to avoid the zwitterionic state before cell entry. In the first study (**Section 3.4**) an approach was pursued to cap the carboxyl group of **69s-A/B** as an ethyl ester since it is a commonly used concept for drug delivery.[289] Unfortunately, it rapidly hydrolyzed in aqueous environment leading to insufficient cell permeability.

As a more promising strategy, a prodrug form should be considered that is less prone to spontaneous hydrolysis and more likely to be cleaved by an enzyme (e.g. an esterase) inside the cell. The prodrug of the angiotensin receptor blocker (AT$_1$ antagonist) candesartan, namely candesartan cilexetil (**94**) provides an acylal-based group capping the carboxyl function (**Figure 24**).[290]
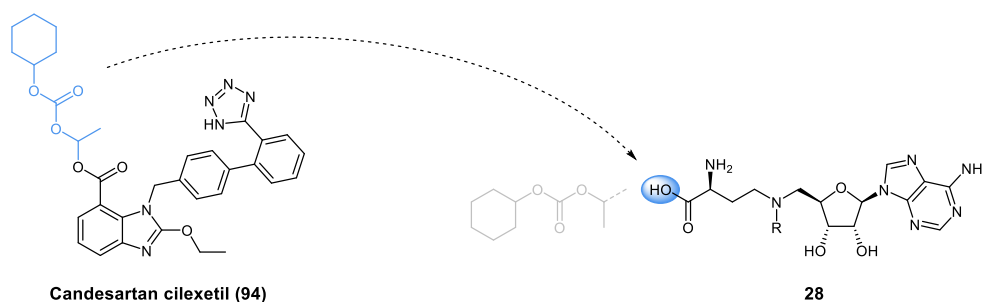
Figure 24: Left: Structure of candesartan cilexetil. The acylal prodrug substructure is highlighted in blue. Right: Attachment point for the acylal group to generate SAH-based prodrugs.

In a test approach *N-n*-butyl-adenosyl-Dab **69e** was treated with 1-chloroethyl cyclohexyl carbonate (**95**) using either potassium carbonate or *N,N*-diisopropylethylamine (DIPEA) in DMF at different temperatures to analyze the chemoselectivity of this reaction (**Scheme 24**). Conducting the reactions with potassium carbonate at room temperature or 50 °C did not lead to a conversion. The same was true for using DIPEA at room temperature. Interestingly, when the mixtures were heated at 70–90 °C, a product was formed exhibiting a mass-to-charge ratio of *m/z* = 550, which could be assigned to the cyclohexyl carbamate **97** or a respective constitutional isomer.
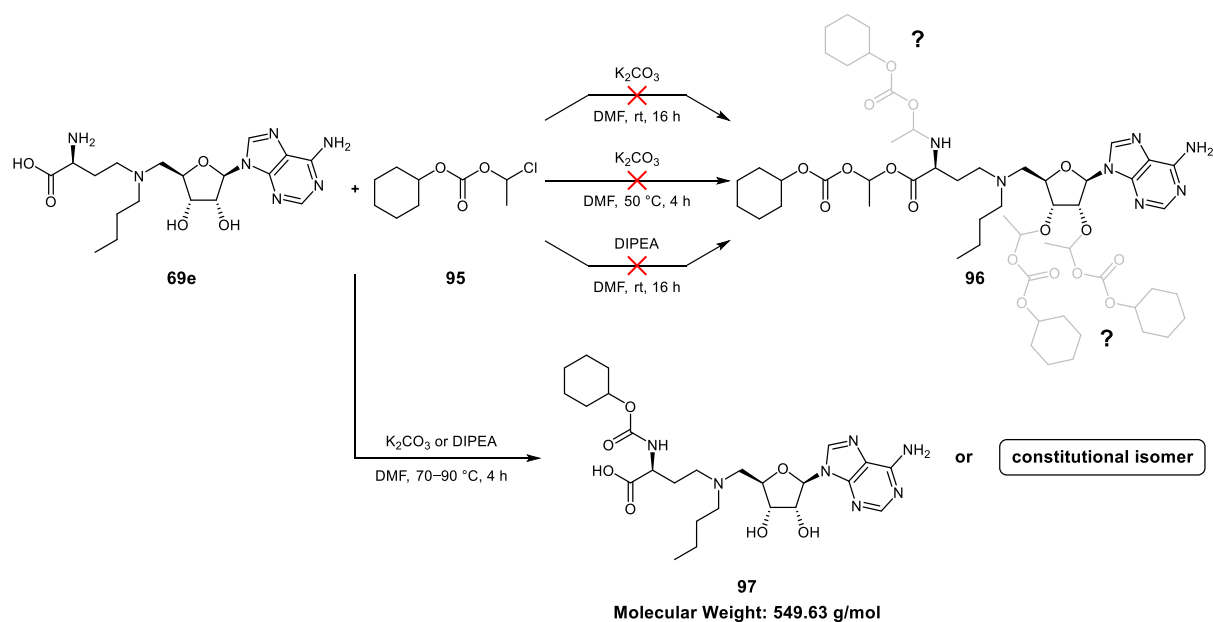
Scheme 24: Test approaches to introduce an acylal prodrug group to *N-n*-butyl-adenosyl-Dab (**69e**).

Since several approaches to introduce the acylal prodrug group to an SAH-like scaffold failed, another strategy was pursued. Instead of capping the carboxyl function, modification of the more nucleophilic amino function should be synthetically more feasible. This concept is used for 5-fluorouracil **3**, which is applied as the prodrug capecitabine (**98**) with the amino function capped as pentyl carbamate (**Figure 25**). In the organism the carbamate is cleaved by carboxylesterases releasing the free amine.[291,292]
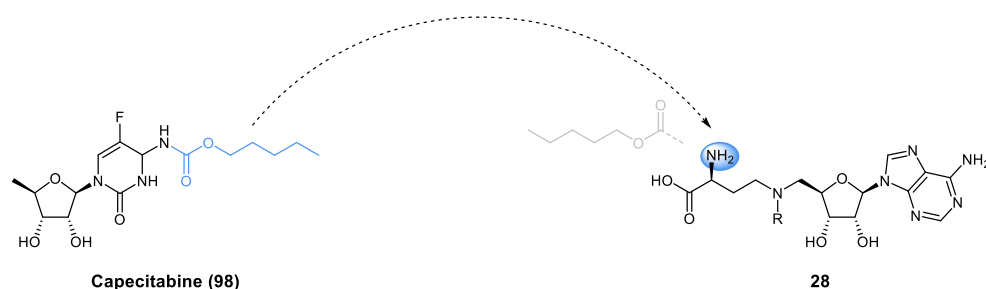
Figure 25: Left: Structure of capecitabine with the pentyl carbamate group highlighted in blue. Right: Attachment point for the pentyl carbamate group to generate SAH-based prodrugs.

For the design of SAH-based inhibitors, the hexyl carbamate group was considered as a suitable prodrug substructure as it is more lipophilic than pentyl carbamate and should therefore compensate the high hydrophilicity of the SAH scaffold. A test approach was conducted treating 4-bromo-3-nitrophenylsulfonyl-adenosyl-Dab (**71i**) with *n*-hexyl chloroformate (**99**) in presence of sodium hydrogencarbonate and tetrabutylammonium bromide (TBAB) in THF and water at room temperature (**Scheme 25**).
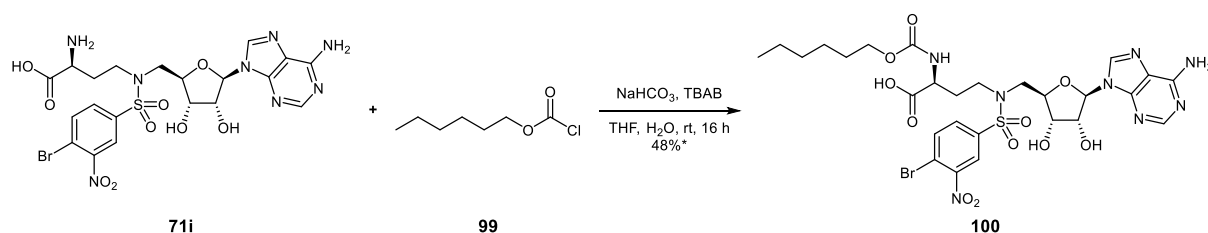
Scheme 25: Test approach to synthesize the hexyl carbamate prodrug **100** (procedure adapted from PONZANO *et al.*).[293] *Determined via LC-MS (AUC).

The chromatogram of the reaction mixture showed a huge extension of retention time (5.2 vs. 2.9 min) of the prodrug compared to the free inhibitor on a C_{18} column (**Figure 26**), suggesting a significant increase in lipophilicity. In addition, it appeared to be more stable as hydrolysis in aqueous environment could not be observed compared to the ester prodrug *28-A/B* presented in **Section 3.4**. The suitability of the prodrug needs to be further validated by analyzing cell permeability and performing kinetic studies of enzymatic cleavage to determine the required incubation time for cell testing.
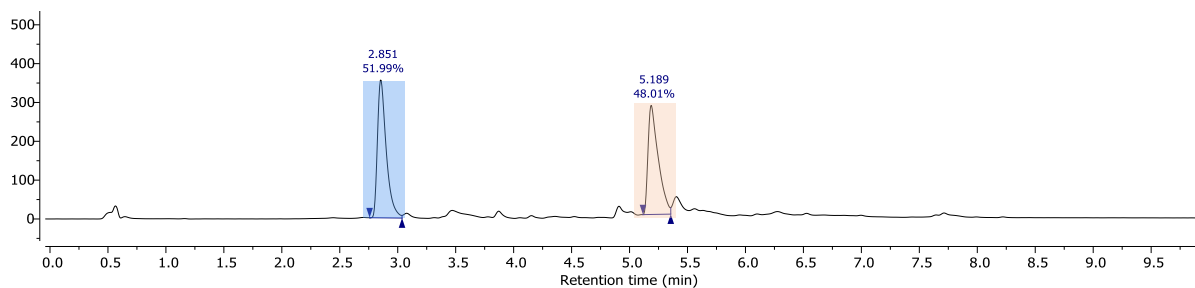
Figure 26: Chromatogram showing reaction mixture of hexyl carbamate prodrug (**100**) synthesis after 16 h. Educt peak (**71i**) is highlighted in blue, product peak (**100**) is highlighted in orange. C_{18} column, MeCN/H$_2$O gradient method (10:90 → 90:10).

4 Conclusion and outlook

The RNA methyltransferases DNMT2, NSUN2, and NSUN6 do not only play significant roles in various physiological processes but are also associated to diseases such as cancer. Although efforts are made by researchers to improve understanding of the biological functions of these enzymes, particularly of their RNA modifications, they are not perfectly understood yet. Therefore, these RNA methyltransferases represent promising targets for drug discovery as well as for the development of activity-based probes, which can be used to further study the biological impacts of the enzymes. However, the development of MTase inhibitors is challenging, especially since suitable lead structures are not known for each target. Literature often only provides knowledge about the inhibitory effects of the natural ligands SAH and SFG – structures that highly resemble each other. Therefore, the structural possibilities of the SAH scaffold within the chemical space was explored in this thesis to find potent inhibitors of DNMT2, NSUN2, and NSUN6 with a primary focus on DNMT2 (**Scheme 26**).

In a first study, the amino acid side chain of SAH was varied, which led to the conclusion that retaining the carboxyl and amino functions is essential for affinity to DNMT2. Exchanging the sulfur with nitrogen allowed the attachment of additional side chains to target the cytidine site. Various alkyl-based moieties were analyzed for their structure-activity relationship. A (*SR*)-*N*-but-3-yn-2-yl-decorated derivative (**69s-A/B**) was identified as the most potent ($IC_{50} = 12.9 \pm 1.9$ μM) representative of the first generation of SAH-based inhibitors. Although it was only slightly more potent than SAH ($IC_{50} = 15.8 \pm 1.5$ μM) and SFG ($IC_{50} = 13.2 \pm 0.8$ μM), it exhibited an improved selectivity profile as it did not inhibit NSUN2 and NSUN6. To investigate the effect on a cellular level, an ethyl ester prodrug of **69s-A/B** was prepared and tested, resulting in only a slight but significant reduction of the m^5C tRNA level. CaCo-2 assays revealed poor cell permeability due to rapid hydrolysis of the ester prodrug in aqueous environment, leading to reconsideration of the concept for future inhibitor and prodrug design.

To allow more space for side chain modification, the *N*-benzyl derivative **69q** ($56.8 \pm 6.8\%$ inhibition at 100 μM) from the first generation was taken for optimization. According to the TOPLISS scheme suggestions, the phenyl moiety was decorated with different substituents yielding electron-deficient aromatic structures with high inhibition of DNMT2 (e.g. **70n**, $IC_{50} = 2.5 \pm 0.2$ μM) as the second generation of SAH-based inhibitors. By attaching strong electron-withdrawing groups combined with leaving groups, the electrophilicity could be adjusted providing a new type of aryl warhead (4-bromo-3-nitrophenylsulfonyl) that covalently reacted with the catalytically active Cys79 of DNMT2. This warhead not only provides optimal length and positioning of the electrophilic center for Cys79, but it also mimics the natural substrate, namely cytidine of tRNA. With an IC_{50} value of 1.2 ± 0.1 μM, compound **71i** showed an improved inhibition of one order of magnitude compared to **69s-A/B**, hence outclassing the natural ligands SAH and SFG. Furthermore, it was selective toward NSUN2 and NSUN6. The concept of a cova-

lent aromatic SAH derivative may be highly beneficial for the overall selectivity profile as various other MTases do not exhibit a catalytically active cysteine.
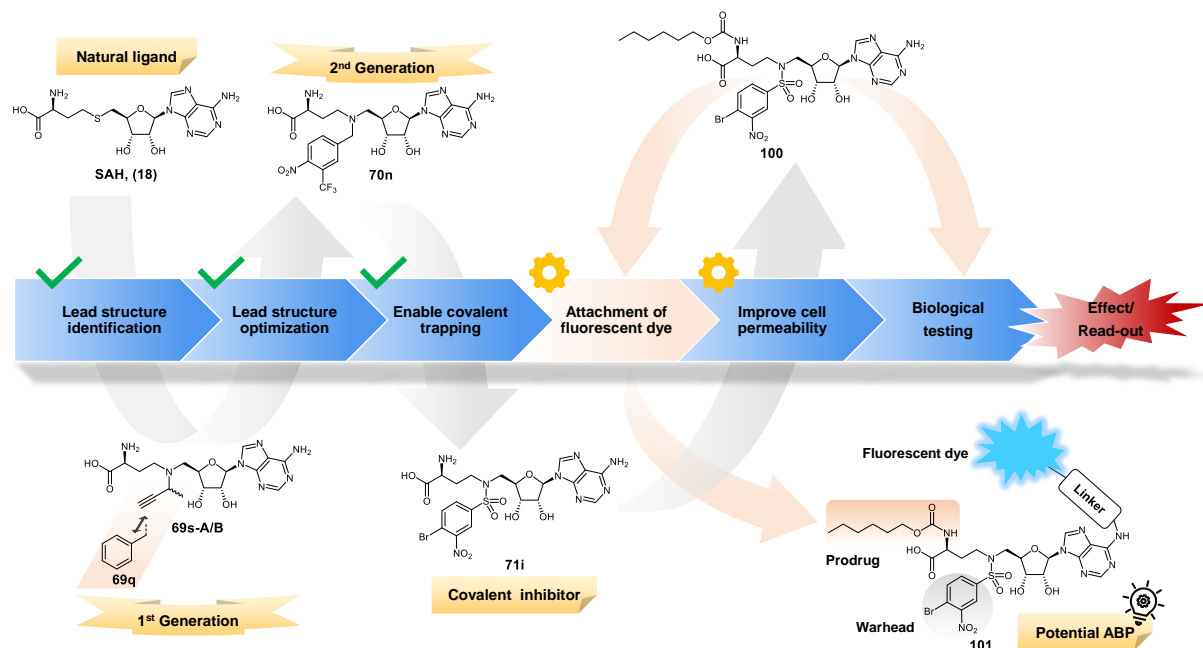
Scheme 26: Process flowchart of the DNMT2 project showing the milestones.

Evaluating alternative prodrug groups led to a hexyl carbamate **100** that showed promising properties as it significantly increased lipophilicity of **71i**. In addition, spontaneous hydrolysis could not be observed compared to the ethyl ester prodrug *28-A/B* of the first study (**Section 3.4**). Nevertheless, the prodrug needs to be further validated by an enzymatic cleavage study, determination of cell permeability as well as cell testing. If the prodrug group was found to be suitable for *in cellulo* testing, it can be used as a basis for ABP development by attaching a fluorescent dye (via a linker) to *N*6 of the adenine moiety. An exemplary structure (**101**) is illustrated in **Scheme 26**. Notably, another formulation strategy for SAH derivatives could be a liposome-based carrier system that allows the delivery of highly polar, charged molecules into cells. A method is currently being established by ▉▉▉▉▉▉▉ (▉▉▉▉ group).

Creating selective inhibitors for other MTases may be achieved by exploring the structural properties of cytidine-site targeting side chains since the enzymes address different locations of their target tRNAs and, therefore, differ in composition, particularly size of their corresponding subpockets. A cyclohexyl derivative (**69k**) inhibiting NSUN6 (45.8 ± 6.1% at 100 μM) was found to be selective toward DNMT2 (n. i.) suggesting that bulky moieties result in binding to NSUN6 (large cytidine site), but steric clash toward DNMT2 (narrow cytidine site). Based on the concept of enforcing selectivity through steric clash, various SAH-based compounds with bulky cyclohexyl derived moieties and aromatic systems of different sizes were designed for testing on DNMT2, NSUN2, and NSUN6. The results should provide information on the substructures that fit into each cytidine site of the three enzymes to create a 'clash profile' that can be used for inhibitor design. A small set of the design compounds for the study was prepared giving

163

evidence about the preference of NSUN6 for bulky moieties. To make precise conclusions, especially for the creation of a clash profile, the dataset needs to be completed.

Within the MTase project, it was possible to establish an optimized MST method designed by ▮▮▮▮▮▮ ▮▮▮▮▮▮▮▮ (▮▮▮▮▮▮▮ group) that is based on the displacement of the fluorescent ligand FTAD (**84**) used as an ABP. The method allows fast screening of active-site inhibitors of DNMT2 with half-quantitative screening outcomes. The measurements are highly beneficial for pre-evaluation of potential inhibitors since it reduces the number of compounds subjected to the time- and cost-intensive tritium incorporation assay.

As a future perspective, disposal of SAH-like characteristics could be beneficial, especially in terms of selectivity as well as pharmacokinetic properties. By creating non-SAH-like inhibitors (**Figure 27**), high structural similarity to the natural ligand would be eliminated and open space for new modifications. Since the amino acid side chain functions appeared to be essential for affinity, this part of the molecule should either be retained or substituted by suitable bioisosteres. The same is true for established cytidine-site targeting substructures as they improved the affinity and selectivity profile. To avoid potential mutagenic effects caused by the nitro group of **71i**,[294] it may be replaced by a cyano group. However, for SAH-like derivatives, this requires an alternative synthetic approach as the currently used deprotection method using TFA/water will probably lead to hydration of this group. Overall, replacement of the adenosine substructure may yield potential MTase inhibitors. Adenine may be replaced using scaffold hopping that yields an aromatic structure exhibiting a similar hydrogen bond donor/acceptor profile. A potential replacement for the ribose was identified by ▮▮▮▮▮▮▮▮▮▮▮▮▮ (▮▮▮▮▮▮ group) using computer-aided methods. This structure mimics the vicinal diol of the pentose, which interacts with Asp34 of DNMT2.
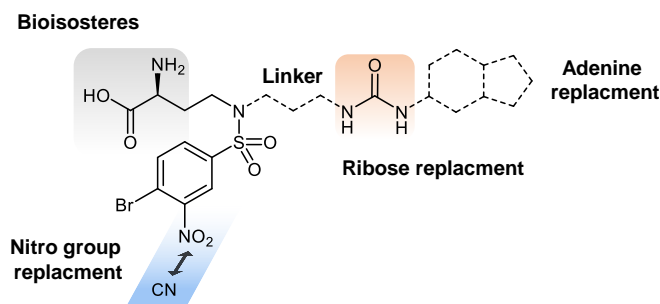
Figure 27: Possible modifications/approaches to create non-SAH-like derivatives retaining established substructures of SAH-based inhibitors.

5 Experimental section

5.1 General information

All reagents and solvents were commercial grade and used without further purification. Reaction progress was monitored by thin-layer chromatography using ALUGRAM® Xtra SIL G UV$_{254}$ silica plates from *Machery-Nagel*. Column chromatography was performed with silica (40–63 μm) from *Machery-Nagel*. NMR spectra were recorded on *Bruker Fourier 300* at 300 MHz. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance (CDCl$_3$, DMSO-d_6, or CD$_3$OD from *Deutero GmbH*) as internal standard. Flash chromatography was performed with a *Biotage Isolera^{TM} One* system using prepacked C$_{18}$ columns from *Biotage*. The identities and purities of final compounds were determined by combined HPLC/ESI-MS analysis using an *Agilent 1100 series* HPLC system with an *Agilent Zorbax SB-Aq* (4.6 × 150 mm, 5 μm) column and an *Agilent Poroshell 120 EC-C$_{18}$* (150 mm × 2.10 mm, 4 μm) column at 40 °C oven temperature, respectively. As mobile phase, mixtures of MeCN and MilliQ®-H$_2$O + 0.1% HCOOH were used at a flow rate of 0.7 mL/min. Samples were applied using 5 μL injection with quantitation by AUC at 254 nm. Electro spray ionization (ESI) mass spectra were recorded on an *Agilent 1100 series LC/MSD Ion trap* spectrometer in the positive ion mode. 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. Specific rotations $[\alpha]_D^{20}$ were determined with a *Krüss P3000* polarimeter and are given in deg cm^3 g^{-1} dm^{-1}. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

5.2 Synthetic protocols

5.2.1 General procedure A – reductive Amination

To a solution of **50** (1.0 equiv.) in 1,2-DCE (5 mL) at 0 °C under argon atmosphere were added an aldehyde (2.0 equiv.) and HOAc (1.5 equiv.). The mixture was stirred at 0 °C for 30 min and then NaBH(OAc)$_3$ (1.7 equiv.) was added. The reaction was slowly warmed up to room temperature and was stirred overnight. The mixture was quenched by the addition of MeOH (5 mL) and the solvent was removed under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (DCM/MeOH = 30:1) and, if necessary, on C$_{18}$ (MeCN/H$_2$O = 10:90 $\rightarrow$ 100:0).

5.2.2 General procedure B – sulfonamide formation

To a mixture of **50** (1.0 equiv.) in DCM (5 mL) and saturated NaHCO$_3$ solution (5 mL) was added the sulfonyl chloride (1.5 equiv.). After vigorous stirring overnight, the organic phase was separated, and the aqueous phase was extracted with DCM (2 × 10 mL). The combined organic extracts were dried over Na$_2$SO$_4$, filtered, and the solvent was removed under reduced pressure at 40 °C. The residue was purified

by column chromatography on silica (DCM/MeOH = 30:1) and, if necessary, on C_{18} (MeCN/H$_2$O = 10:90 → 100:0).

5.2.3 General procedure C – deprotection

To a solution of a protected compound in DCM (2 mL) at 5 °C was added TFA (2 mL). The solution was kept at 5 °C until complete deprotection of the amine and carboxylic acid were detected by LC-MS. The reaction was diluted and co-distilled with DCM (3 × 20 mL) at 40 °C, and the residue was dissolved in water (3 mL) and TFA (0.5 mL) at 5 °C. The solution was kept again at 5 °C until full conversion was detected by LC-MS. Then the mixture was diluted with water (15 mL) and dried by lyophilization to obtain the final compound.

5.2.4 Compounds and analytical data

***tert*-Butyl (*S*)-4-(((*E*)-*N*-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydro-furo[3,4-*d*][1,3]dioxol-4-yl)methyl)-2-phenylvinyl)sulfonamido)-2-((*tert*-butoxycarbonyl)amino)-butanoate (75a)**
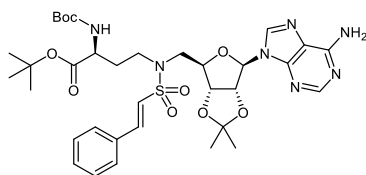
The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.) and (*E*)-2-phenylethene-1-sulfonyl chloride (135 mg, 0.67 mmol, 1.5 equiv.) according to general procedure B to afford the desired product as a colorless solid (211 mg, 0.29 mmol, 66%). FT-IR: ν / cm^{-1} = 3428, 3342, 3232, 2978, 2941, 1727, 1697, 1650, 1332, 1139, 1074, 870, 825, 750, 692. mp: 230 °C (decomposition). R$_f$: 0.25 (DCM/MeOH = 30:1).

Note: the product was only soluble in organic solvents under heating. NMR spectra and specific rotation could not be obtained.

tert-Butyl (S)-4-((N-(((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-d][1,3]dioxol-4-yl)methyl)-2,3,4,5,6-pentafluorophenyl)sulfonamido)-2-((tert-butoxy-carbonyl)amino)butanoate (75b)
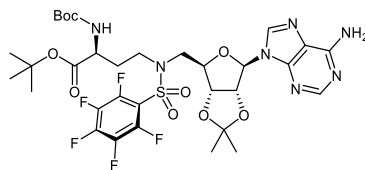
The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.) and pentafluorobenzenesulfonyl chloride (99 μL, 0.67 mmol, 1.5 equiv.) according to general procedure B to afford the desired product as a colorless oil (167 mg, 0.21 mmol, 48%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.31 (s, 1H), 7.79 (s, 1H), 6.30 (s, 2H), 5.96 (d, J = 1.8 Hz, 1H), 5.35 (dd, J = 6.4, 1.9 Hz, 1H), 5.25 (d, J = 7.5 Hz, 1H), 5.06 (dd, J = 6.4, 3.8 Hz, 1H), 4.39–4.30 (m, 1H), 4.14–3.99 (m, 1H), 3.83–3.48 (m, 3H), 3.31–3.16 (m, 1H), 2.20–2.06 (m, 1H), 1.95–1.78 (m, 1H), 1.57 (s, 3H), 1.43 (s, 9H), 1.41 (s, 9H), 1.34 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.8, 155.7, 155.6, 152.8, 148.8, 140.1, 120.3, 114.9, 90.6, 85.0, 84.3, 82.7, 82.6, 80.1, 52.1, 49.9, 45.4, 31.9, 28.4, 28.0, 27.2, 25.3. FT-IR: ν / cm^{-1} = 3327, 3182, 2980, 2937, 1708, 1643, 1520, 1495, 1367, 1251, 1153, 1097, 990, 870, 750. $[\alpha]_D^{20}$ = +23 (10 mg/mL; CHCl$_3$). R$_f$: 0.18 (DCM/MeOH = 30:1).

Note: Some of the peaks of the quaternary carbon atoms coupling with fluorine were weakly resolved and had low intensity in the ^{13}C NMR.

(S)-2-amino-4-(((E)-N-(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-2-phenylvinyl)sulfonamido)butanoic acid, trifluoroacetate salt (73a)
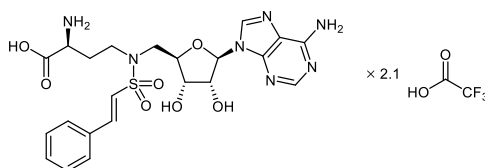
The compound was prepared from **75a** (128 mg, 0.18 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (136 mg, 0.18 mmol, 99%, 2.1 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.29 (s, 1H), 8.16 (s, 1H), 7.32–7.13 (m, 6H), 6.81 (d, J = 15.4 Hz, 1H), 5.90 (d, J = 4.5 Hz, 1H), 4.51 (pseudo-t, J ≈ 4.5 Hz, 1H), 4.21–4.10 (m, 2H), 3.84 (pseudo-t, J ≈ 6.4 Hz, 1H), 3.61–3.42 (m, 2H), 3.40–3.20 (m, 2H), 2.27–2.11 (m, 1H), 2.06–1.90 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.8, 152.9, 149.9, 146.8, 143.7, 143.3, 134.0, 131.9, 130.1, 129.3, 124.7, 120.7, 91.0, 83.8, 75.1, 73.0, 52.2, 51.8, 46.1, 31.1. FT-IR: ν / cm^{-1} = 3065, 1668, 1614, 1506, 1423, 1323, 1186, 1128, 970, 861, 821, 798, 748, 721, 689. $[\alpha]_D^{20}$ = +31 (10 mg/mL; MeOH). mp: 85–87 °C. ESI-MS: m/z

calculated for $[C_{22}H_{27}N_7O_7S+H]^+$ ($[M+H]^+$) = 534.2, found: 534.9. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 3.52 min).

(S)-2-amino-4-((N-(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-2,3,4,5,6-pentafluorophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (73b)
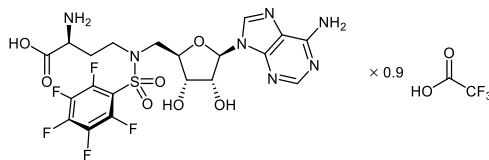
The compound was prepared from **75b** (110 mg, 0.14 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (97 mg, 0.14 mmol, 99%, 0.9 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.27 (s, 2H), 5.82 (d, J = 4.1 Hz, 1H), 4.53 (*pseudo*-t, $J \approx$ 5.1 Hz, 1H), 4.19 (*pseudo*-t, $J \approx$ 5.1 Hz, 1H), 4.16–4.07 (m, 1H), 3.82 (*pseudo*-t, $J \approx$ 6.4 Hz, 1H), 3.76–3.59 (m, 3H), 3.50–3.33 (m, 1H), 2.34–2.17 (m, 1H), 2.13–1.97 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.2, 152.6, 149.7, 146.1, 144.1, 120.8, 91.2, 83.2, 75.0, 73.0, 52.0, 51.5, 46.3, 31.0. FT-IR: ν / cm^{-1} = 3104, 1670, 1521, 1495, 1425, 1358, 1299, 1168, 1131, 1097, 990, 896, 834, 798, 721. $[\alpha]_D^{20}$ = +15 (10 mg/mL; MeOH). mp: 71–73 °C. ESI-MS: m/z calculated for $[C_{20}H_{20}F_5N_7O_7S+H]^+$ ($[M+H]^+$) = 598.1, found: 598.2. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.79 min).

Note: Some of the peaks of the quaternary carbon atoms coupling with fluorine were weakly resolved and had low intensity in the ^{13}C NMR.

Bicyclo[2.2.2]octane-1-carbaldehyde (79a)

Using a modified procedure from TAKEUCHI *et al.*, *Synthesis* **1987**, 612–615.

To a solution of bicyclo[2.2.2]octan-1-ylmethanol (**78**, 300 mg, 2.14 mmol, 1.0 equiv.) in DCM (5 mL) was added pyridinium chlorochromate (692 mg, 3.21 mmol, 1.5 equiv.). After stirring at room temperature for 3 h, the mixture was filtered over silica to yield the desired product as a colorless solid (248 mg, 1.79 mmol, 84%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.39 (s, 1H), 1.70–1.65 (m, 1H), 1.61–1.58 (m, 12H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 206.7, 43.1, 25.5, 25.0, 24.6. FT-IR: ν / cm^{-1} = 2937, 2916, 2861, 2695, 1720, 1694, 1455, 1408, 1267, 1139, 1073, 912, 893, 822, 727. mp: 57–59 °C. R$_f$: 0.67 (cyclohexane/ethyl acetate = 10:1).

tert-Butyl (_S_)-4-((((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_]-[1,3]dioxol-4-yl)methyl)(bicyclo[2.2.2]octan-1-ylmethyl)amino)-2-((_tert_-butoxycarbonyl)amino)-butanoate (81a)
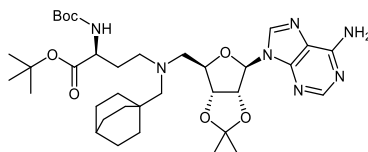
The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.), bicyclo[2.2.2]octane-1-carbaldehyde (**79a**, 123 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 μL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless solid (71 mg, 0.10 mmol, 24%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.31 (s, 1H), 7.89 (s, 1H), 6.13–5.99 (m, 3H), 5.59–5.43 (m, 2H), 4.96 (dd, _J_ = 6.4, 3.4 Hz, 1H), 4.35–4.23 (m, 1H), 4.14–3.88 (m, 1H), 2.76–2.64 (m, 1H), 2.64–2.45 (m, 2H), 2.45–2.33 (m, 1H), 2.08–1.96 (m, 2H), 1.93–1.76 (m, 1H), 1.59 (s, 4H), 1.49–1.36 (m, 28H), 1.28–1.18 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.9, 155.8, 155.5, 153.2, 149.3, 140.2, 120.5, 114.4, 91.0, 86.0, 83.9, 83.5, 81.6, 79.4, 67.3, 58.7, 53.2, 32.7, 29.9, 29.6, 28.5, 28.1, 27.3, 26.2, 25.6, 24.7. FT-IR: ν / cm^{-1} = 3333, 3171, 2977, 2932, 2859, 1706, 1638, 1597, 1475, 1366, 1209, 1152, 1075, 870, 753. $[\alpha]_D^{20}$ = +5 (10 mg/mL; CHCl$_3$). mp: 73–75 °C. R$_f$: 0.17 (DCM/MeOH = 30:1).

tert-Butyl (_S_)-4-((((3_R_,5_R_,7_R_)-adamantan-1-yl)methyl)(((3a_R_,4_R_,6_R_,6a_R_)-6-(6-amino-9_H_-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-_d_][1,3]dioxol-4-yl)methyl)amino)-2-((_tert_-butoxycarbonyl)-amino)butanoate (81b)
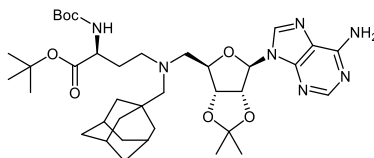
The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.), 1-adamantanecarboxaldehyde (**79b**, 146 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 μL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless solid (81 mg, 0.11 mmol, 26%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.25 (s, 1H), 7.91 (s, 1H), 6.60 (s, 2H), 6.02 (d, _J_ = 2.1 Hz, 1H), 5.59 (d, _J_ = 7.6 Hz, 1H), 5.51–5.42 (m, 1H), 4.96 (dd, _J_ = 6.4, 3.5 Hz, 1H), 4.37–4.25 (m, 1H), 4.14–3.86 (m, 1H), 2.77–2.48 (m, 3H), 2.47–2.36 (m, 1H), 1.99 (s, 2H), 1.90–1.80 (m, 4H), 1.68–1.50 (m, 11H), 1.43–1.38 (m, 19H), 1.38–1.34 (m, 7H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.9, 155.9, 155.5, 152.7, 149.1, 140.1, 120.0, 114.4, 91.0, 86.0, 83.9, 83.5, 81.6, 79.4, 69.4, 58.8, 53.4, 53.1, 41.3, 37.2, 34.8, 29.6, 28.5, 28.4, 28.0, 27.3, 25.6. FT-IR: ν / cm^{-1} = 3328, 3181, 2978,

2902, 2846, 1705, 1643, 1598, 1476, 1366, 1247, 1152, 1075, 871, 753. $[\alpha]_D^{20}$ = +2 (10 mg/mL; CHCl$_3$). mp: 75–77 °C. R$_f$: 0.18 (DCM/MeOH = 30:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(cycloheptylmethyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (81c)
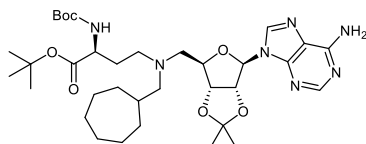
The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.), cycloheptanecarbaldehyde (**79c**, 112 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 µL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless solid (95 mg, 0.14 mmol, 32%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.30 (s, 1H), 7.88 (s, 1H), 6.22 (s, 2H), 6.02 (d, J = 2.1 Hz, 1H), 5.71 (d, J = 7.9 Hz, 1H), 5.57–5.45 (m, 1H), 4.95 (dd, J = 6.4, 3.2 Hz, 1H), 4.36–4.24 (m, 1H), 4.19–3.94 (m, 1H), 2.79–2.67 (m, 1H), 2.61–2.44 (m, 2H), 2.42–2.30 (m, 1H), 2.20–1.99 (m, 2H), 1.96–1.81 (m, 1H), 1.79–1.64 (m, 2H), 1.63–1.42 (m, 11H), 1.42–1.39 (m, 18H), 1.36 (s, 3H), 1.33–1.17 (m, 2H), 1.15–0.90 (m, 2H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 172.0, 155.9, 155.6, 153.1, 149.2, 140.2, 120.4, 114.3, 91.2, 85.7, 83.9, 83.7, 81.6, 79.4, 62.4, 56.7, 53.2, 51.6, 37.2, 32.8, 32.2, 29.1, 28.8, 28.6, 28.1, 27.2, 26.4, 26.4, 25.5. FT-IR: ν / cm^{-1} = 3333, 3178, 2977, 2924, 2851, 1709, 1642, 1597, 1476, 1366, 1208, 1152, 1073, 870, 754. $[\alpha]_D^{20}$ = −12 (10 mg/mL; CHCl$_3$). mp: 76–78 °C. R$_f$: 0.50 (DCM/MeOH = 20:1).

tert-Butyl (S)-4-(([1,1'-biphenyl]-4-ylmethyl)(((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-2-((tert-butoxycarbonyl)amino)-butanoate (82a)
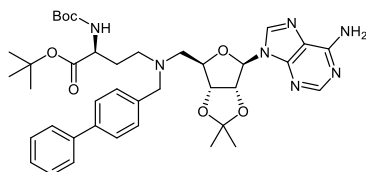
The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.), biphenyl-4-carboxaldehyde (**80a**, 162 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 µL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless solid (231 mg, 0.32 mmol, 72%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.21 (s, 1H), 7.85 (s, 1H), 7.59–7.53 (m, 2H), 7.50–7.45 (m, 2H), 7.44–7.38 (m, 2H), 7.35–7.27 (m, 3H), 6.09 (s, 2H), 6.03 (d, J = 2.1 Hz, 1H), 5.86–5.61 (m, 1H), 5.48–5.33 (m, 1H), 4.90 (dd, J = 6.5, 3.4 Hz, 1H), 4.43–4.34 (m, 1H), 4.26–4.14 (m, 1H), 3.78–3.68 (m, 1H), 3.58–3.45 (m, 1H), 2.86–2.75 (m, 1H), 2.73–2.59 (m, 2H), 2.58–2.48 (m,

170

1H), 2.08–1.93 (m, 1H), 1.88–1.72 (m, 1H), 1.59 (s, 3H), 1.42 (s, 9H), 1.39 (s, 9H), 1.37 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.8, 155.6, 153.1, 149.2, 141.0, 140.1, 140.0, 137.7, 129.5, 128.8, 127.3, 127.1, 127.0, 120.4, 114.4, 90.9, 85.5, 83.9, 83.5, 81.7, 79.5, 58.7, 55.8, 53.0, 50.8, 29.5, 28.5, 28.0, 27.2, 25.5. FT-IR: ν / cm^{-1} = 3332, 3161, 2978, 2934, 2824, 1709, 1640, 1598, 1487, 1366, 1152, 1075, 848, 760, 698. $[\alpha]_D^{20}$ = −20 (10 mg/mL; CHCl$_3$). mp: 79–81 °C. R$_f$: 0.38 (DCM/MeOH = 20:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-phenoxybenzyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (82b)

The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.), 3-phenoxybenzaldehyde (**80b**, 176 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 μL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a colorless solid (181 mg, 0.24 mmol, 55%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.22 (s, 1H), 7.83 (s, 1H), 7.33–7.26 (m, 2H), 7.23–7.16 (m, 1H), 7.09–7.02 (m, 1H), 7.01–6.94 (m, 4H), 6.87–6.81 (m, 1H), 6.17 (s, 2H), 6.01 (d, J = 2.2 Hz, 1H), 5.73–5.52 (m, 1H), 5.47–5.31 (m, 1H), 4.87 (dd, J = 6.5, 3.5 Hz, 1H), 4.36–4.27 (m, 1H), 4.22–3.92 (m, 1H), 3.72–3.61 (m, 1H), 3.51–3.38 (m, 1H), 2.85–2.71 (m, 1H), 2.69–2.54 (m, 2H), 2.53–2.41 (m, 1H), 2.03–1.87 (m, 1H), 1.83–1.67 (m, 1H), 1.57 (s, 3H), 1.43–1.37 (m, 18H), 1.33 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 157.4, 157.3, 155.8, 155.5, 153.1, 149.3, 141.0, 139.9, 129.8, 129.6, 123.9, 123.2, 120.4, 119.6, 118.8, 117.7, 114.5, 90.8, 85.4, 83.9, 83.4, 81.7, 79.5, 58.9, 55.8, 52.9, 50.8, 29.5, 28.5, 28.0, 27.2, 25.5. FT-IR: ν / cm^{-1} = 3332, 3176, 2978, 2935, 2823, 1705, 1583, 1484, 1366, 1248, 1152, 1073, 870, 753, 692. $[\alpha]_D^{20}$ = −7 (10 mg/mL; CHCl$_3$). mp: 67–69 °C. R$_f$: 0.26 (DCM/MeOH = 30:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(pyren-1-ylmethyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (82c)
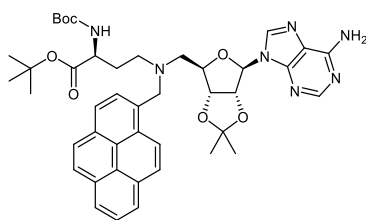
The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.), 1-pyrenecarboxaldehyde (**80c**, 204 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 μL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as a slightly yellowish solid (248 mg, 0.32 mmol, 73%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.43 (d, J = 9.3 Hz, 1H), 8.15 (d, J = 7.8 Hz, 2H), 8.08–7.94 (m, 6H), 7.88 (d, J = 7.8 Hz, 1H), 7.73 (s, 1H), 6.10–5.90 (m, 3H), 5.66–5.32 (m, 1H), 5.19–5.04 (m, 1H), 4.63 (dd, J = 6.5, 3.0 Hz, 1H), 4.44–4.31 (m, 2H), 4.25–4.05 (m, 2H), 2.89–2.58 (m, 4H), 2.09 (s, 1H), 1.90 (s, 1H), 1.49 (s, 3H), 1.38 (s, 9H), 1.31 (s, 9H), 1.12 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.6, 155.5, 153.0, 149.1, 139.7, 132.1, 131.4, 130.9, 129.8, 128.3, 127.5, 127.4, 127.3, 126.0, 125.1, 125.0, 124.8, 124.5, 124.0, 120.2, 114.2, 91.1, 85.3, 83.6, 83.4, 81.8, 79.5, 57.7, 55.7, 52.9, 51.2, 29.4, 28.4, 28.0, 27.1, 25.1. FT-IR: ν / cm^{-1} = 3332, 3167, 2978, 2933, 2360, 1706, 1636, 1474, 1366, 1245, 1151, 1075, 845, 754, 710. $[\alpha]_D^{20}$ = −28 (10 mg/mL; CHCl$_3$). mp: 103–105 °C. R$_f$: 0.31 (DCM/MeOH = 20:1).

tert-Butyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d]-[1,3]dioxol-4-yl)methyl)(perylen-3-ylmethyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate (82d)

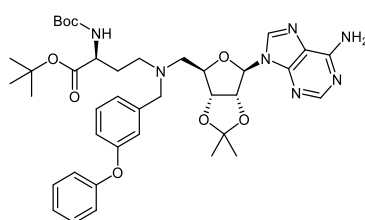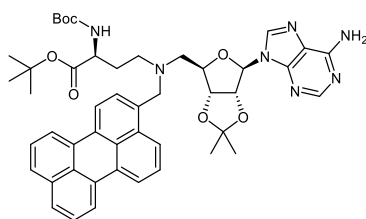The compound was prepared from **50** (250 mg, 0.44 mmol, 1.0 equiv.), 3-perylenecarboxaldehyde (**80d**, 249 mg, 0.89 mmol, 2.0 equiv.), HOAc (38 μL, 0.67 mmol, 1.5 equiv.), and NaBH(OAc)$_3$ (160 mg, 0.75 mmol, 1.7 equiv.) according to general procedure A to afford the desired product as an orange-yellow solid (215 mg, 0.26 mmol, 59%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.17–8.05 (m, 4H), 8.00 (d, J = 8.4 Hz, 1H), 7.93 (d, J = 7.7 Hz, 1H), 7.76 (s, 1H), 7.63 (d, J = 8.1 Hz, 2H), 7.49–7.39 (m, 3H), 7.30 (d, J = 7.7 Hz, 1H), 6.19 (s, 2H), 5.95 (d, J = 2.2 Hz, 1H), 5.45–5.30 (m, 1H), 5.25–5.13 (m, 1H), 4.69 (dd, J = 6.4, 3.0 Hz, 1H), 4.47–4.35 (m, 1H), 4.19 (s, 1H), 4.07–3.93 (m, 1H), 3.85–3.73 (m, 1H),

2.88–2.55 (m, 4H), 2.08 (s, 1H), 1.93–1.79 (m, 1H), 1.53 (s, 3H), 1.40 (s, 9H), 1.37 (s, 9H), 1.23 (s, 3H).
^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.8, 155.6, 155.5, 152.7, 149.0, 139.7, 134.7, 133.9, 133.5, 131.5, 131.4, 131.2, 131.0, 128.0, 129.53, 128.5, 127.8, 126.7, 126.6, 126.4, 124.5, 120.3, 120.2, 120.1, 119.5, 114.2, 91.2, 85.4, 83.7, 83.5, 81.9, 79.6, 57.8, 55.6, 52.9, 51.2, 29.4, 28.4, 28.0, 27.1, 25.2. FT-IR: ν / cm^{-1} = 3332, 3175, 3051, 2977, 2934, 1705, 1643, 1593, 1476, 1366, 1248, 1151, 1091, 814, 767. $[\alpha]_D^{20}$ = −21 (10 mg/mL; CHCl$_3$). mp: 86 °C (decomposition). R$_f$: 0.38 (DCM/MeOH = 20:1).

(S)-2-Amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(bicyclo[2.2.2]octan-1-ylmethyl)amino)butanoic acid, trifluoroacetate salt (76a)
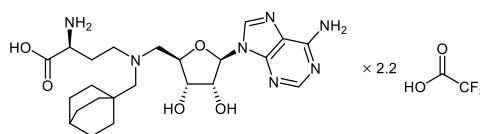
The compound was prepared from **81a** (52 mg, 0.08 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (56 mg, 0.08 mmol, 99%, 2.2 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.38 (s, 1H), 8.35 (s, 1H), 6.06 (d, J = 2.9 Hz, 1H), 4.58–4.51 (m, 1H), 4.47–4.36 (m, 2H), 3.92 (dd, J = 8.7, 4.3 Hz, 1H), 3.73–3.63 (m, 1H), 3.62–3.46 (m, 2H), 3.45–3.32 (m, 1H), 3.11–2.91 (m, 2H), 2.40–2.25 (m, 1H), 2.23–2.09 (m, 1H), 1.61–1.42 (m, 13H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.1, 153.1, 149.7, 147.1, 144.0, 121.0, 92.4, 80.0, 74.5, 73.6, 67.5, 59.3, 56.5, 53.2, 32.5, 29.7, 26.3, 25.1. FT-IR: ν / cm^{-1} = 2946, 2867, 1670, 1559, 1508, 1458, 1423, 1322, 1185, 1131, 987, 904, 829, 798, 721. $[\alpha]_D^{20}$ = +18 (10 mg/mL; MeOH). mp: 74 °C (decomposition). ESI-MS: m/z calculated for [C$_{23}$H$_{35}$N$_7$O$_5$+H]$^+$ ([M+H]$^+$) = 490.3, found: 490.1. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.79 min).

(S)-4-(((((3R,5R,7R)-adamantan-1-yl)methyl)(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)-2-aminobutanoic acid, trifluoroacetate salt (76c)
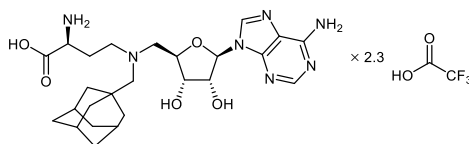
The compound was prepared from **81b** (63 mg, 0.09 mmol) according to general procedure C to afford the final product as a slightly yellowish trifluoroacetate salt (69 mg, 0.09 mmol, 99%, 2.3 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.37 (s, 1H), 8.34 (s, 1H), 6.05 (d, J = 2.9 Hz, 1H), 4.56–4.50 (m, 1H), 4.47–4.36 (m, 2H), 3.93 (dd, J = 8.6, 4.4 Hz, 1H), 3.75–3.64 (m, 1H), 3.63–3.46 (m, 2H), 3.45–3.33 (m, 1H), 3.07–2.87 (m, 2H), 2.42–2.26 (m, 1H), 2.24–2.09 (m, 1H), 1.95–1.87 (m, 3H), 1.73–1.50 (m, 12H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.8, 152.7, 149.6, 146.5, 144.2, 121.0, 92.4, 79.9, 74.6, 73.60, 69.0, 59.9, 56.4, 52.8, 40.7, 37.2, 34.5, 29.3, 26.2. FT-IR: ν / cm^{-1} = 3083, 2910, 2853, 1668,

1558, 1507, 1423, 1319, 1183, 1131, 980, 907, 832, 798, 721. $[\alpha]_D^{20}$ = +15 (10 mg/mL; MeOH). mp: 77–80 °C. ESI-MS: m/z calculated for $[C_{25}H_{37}N_7O_5+H]^+$ ([M+H]$^+$) = 516.3, found: 516.2. Purity: 96% (HPLC, 254 nm, MeCN/H$_2$O = 30:70 + 0.1% HCOOH, t_R = 2.21 min).

(S)-2-amino-4-(((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(cycloheptylmethyl)amino)butanoic acid, trifluoroacetate salt (76f)
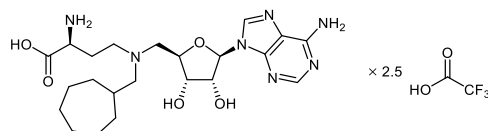
The compound was prepared from **81c** (71 mg, 0.11 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (80 mg, 0.11 mmol, 99%, 2.5 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.34 (s, 1H), 8.31 (s, 1H), 6.01 (d, J = 3.2 Hz, 1H), 4.57–4.51 (m, 1H), 4.43–4.30 (m, 2H), 3.97–3.87 (m, 1H), 3.68–3.42 (m, 3H), 3.41–3.28 (m, 1H), 3.08–2.90 (m, 2H), 2.37–2.23 (m, 1H), 2.22–2.10 (m, 1H), 1.92–1.78 (m, 1H), 1.71–1.61 (m, 1H), 1.59–1.48 (m, 2H), 1.47–1.28 (m, 6H), 1.25–1.02 (m, 3H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.5, 152.8, 149.6, 146.5, 144.3, 121.0, 92.3, 79.7, 74.7, 73.5, 62.4, 56.8, 53.4, 52.4, 36.0, 32.9, 29.2, 29.2, 26.7, 25.8. FT-IR: ν / cm^{-1} = 3075, 2929, 2860, 1667, 1507, 1464, 1424, 1355, 1322, 1182, 1129, 898, 831, 798, 721. $[\alpha]_D^{20}$ = +13 (10 mg/mL; MeOH). mp: 61–63 °C. ESI-MS: m/z calculated for $[C_{22}H_{35}N_7O_5+H]^+$ ([M+H]$^+$) = 478.3, found: 478.1. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 20:80 + 0.1% HCOOH, t_R = 2.74 min).

(S)-4-(([1,1'-Biphenyl]-4-ylmethyl)(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetra-hydrofuran-2-yl)methyl)amino)-2-aminobutanoic acid, trifluoroacetate salt (77c)
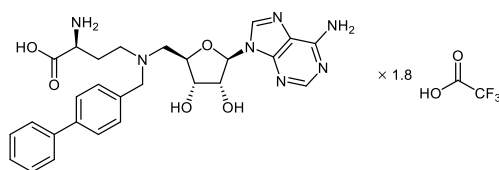
The compound was prepared from **82a** (162 mg, 0.22 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (164 mg, 0.22 mmol, 99%, 1.8 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.29 (s, 1H), 8.14 (s, 1H), 7.48–7.36 (m, 6H), 7.32–7.24 (m, 2H), 7.23–7.16 (m, 1H), 5.99 (d, J = 3.2 Hz, 1H), 4.49–4.38 (m, 3H), 4.36–4.24 (m, 2H), 3.83 (dd, J = 8.6, 4.3 Hz, 1H), 3.64–3.34 (m, 4H), 2.40–2.24 (m, 1H), 2.22–2.07 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.4, 152.8, 149.5, 146.6, 144.1, 143.9, 140.8, 132.8, 130.0, 129.6, 129.0, 128.5, 127.9, 120.8, 92.1, 80.1, 74.8, 73.5, 58.7, 56.0, 53.1, 52.9, 26.3. FT-IR: ν / cm^{-1} = 3064, 1669, 1506, 1488, 1418, 1322, 1182, 1126, 1009, 897, 831, 798, 763, 720, 698. $[\alpha]_D^{20}$ = +8 (10 mg/mL; MeOH). mp: 103 °C (decomposition). ESI-MS: m/z calculated for $[C_{27}H_{31}N_7O_5+H]^+$ ([M+H]$^+$) = 534.3, found: 534.1. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 30:70 + 0.1% HCOOH, t_R = 2.34 min).

174

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(3-phenoxybenzyl)amino)butanoic acid, trifluoroacetate salt (77i)

The compound was prepared from **82b** (112 mg, 0.15 mmol) according to general procedure C to afford the final product as a colorless trifluoroacetate salt (115 mg, 0.15 mmol, 99%, 1.9 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.32 (s, 1H), 8.12 (s, 1H), 7.28–7.07 (m, 4H), 7.01–6.97 (m, 1H), 6.97–6.89 (m, 1H), 6.88–6.82 (m, 1H), 6.82–6.75 (m, 2H), 6.01 (d, J = 3.5 Hz, 1H), 4.52–4.46 (m, 1H), 4.45–4.23 (m, 4H), 3.84 (dd, J = 8.8, 4.2 Hz, 1H), 3.62–3.32 (m, 4H), 2.31–2.16 (m, 1H), 2.15–2.03 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.5, 159.4, 157.8, 152.8, 149.5, 146.5, 144.1, 132.9, 131.8, 131.0, 126.9, 124.9, 122.2, 120.9, 120.76, 120.1, 91.8, 80.2, 74.9, 73.5, 58.5, 56.4, 53.4, 53.2, 26.4. FT-IR: ν / cm^{-1} = 2980, 1669, 1585, 1506, 1488, 1423, 1320, 1257, 1196, 1127, 894, 831, 798, 720, 692. $[\alpha]_D^{20}$ = +15 (10 mg/mL; MeOH). mp: 90–92 °C. ESI-MS: m/z calculated for [C$_{27}$H$_{31}$N$_7$O$_6$+H]$^+$ ([M+H]$^+$) = 550.2, found: 550.1. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 30:70 + 0.1% HCOOH, t_R = 2.37 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(pyren-1-ylmethyl)amino)butanoic acid, trifluoroacetate salt (77j)

The compound was prepared from **82c** (190 mg, 0.24 mmol) according to general procedure C to afford the final product as a slightly yellowish trifluoroacetate salt (177 mg, 0.24 mmol, 99%, 1.3 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.03 (s, 1H), 8.00–7.95 (m, 1H), 7.94–7.87 (m, 3H), 7.87–7.77 (m, 3H), 7.75–7.66 (m, 2H), 7.36 (s, 1H), 5.85 (d, J = 2.4 Hz, 1H), 4.98–4.90 (m, 2H), 4.48–4.37 (m, 1H), 4.26–4.15 (m, 2H), 3.89–3.78 (m, 1H), 3.73–3.51 (m, 4H), 2.50–2.34 (m, 1H), 2.32–2.16 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 172.8, 151.9, 148.4, 145.4, 143.7, 133.1, 132.1, 131.3, 131.0, 129.6, 129.4, 129.3, 128.0, 127.6, 127.2, 126.8, 125.8, 125.3, 124.8, 124.6, 123.0, 120.2, 92.4, 79.8, 74.9, 73.5, 56.2, 55.8, 54.2, 53.7, 26.6. FT-IR: ν / cm^{-1} = 3053, 1670, 1505, 1472, 1418, 1320, 1197, 1180, 1125, 832, 798, 757, 720, 710, 681. $[\alpha]_D^{20}$ = +52 (10 mg/mL; MeOH). mp: 135 °C (decomposition). ESI-MS: m/z calculated for [C$_{31}$H$_{31}$N$_7$O$_5$+H]$^+$ ([M+H]$^+$) = 582.3, found: 582.0. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 30:70 + 0.1% HCOOH, t_R = 2.28 min).

(*S*)-2-Amino-4-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)-methyl)(perylen-3-ylmethyl)amino)butanoic acid, trifluoroacetate salt (77k)
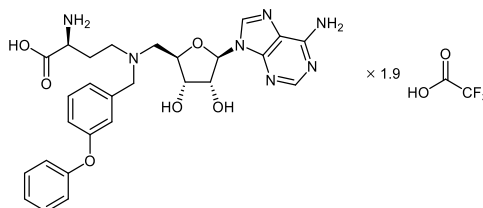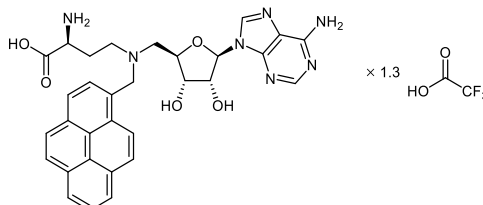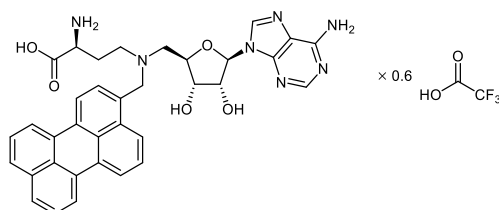

The compound was prepared from **82d** (139 mg, 0.17 mmol) according to general procedure C to afford the final product as a yellow trifluoroacetate salt (118 mg, 0.17 mmol, 99%, 0.6 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.05 (s, 1H), 7.74 (s, 1H), 7.64–7.49 (m, 3H), 7.46–7.38 (m, 1H), 7.37–7.24 (m, 3H), 7.13–7.01 (m, 3H), 6.99–6.89 (m, 1H), 5.84 (s, 1H), 4.43–4.19 (m, 5H), 3.82–3.71 (m, 1H), 3.63–3.39 (m, 4H), 2.39–2.23 (m, 1H), 2.21–2.07 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 173.2, 152.8, 149.0, 147.0, 143.5, 135.5, 133.9, 133.8, 132.4, 131.5, 131.2, 130.8, 129.5, 129.4, 129.1, 128.8, 128.1, 127.6, 127.5, 126.2, 123.4, 121.9, 121.7, 121.4, 120.4, 120.1, 116.2, 92.0, 80.2, 74.8, 73.5, 56.6, 56.2, 54.3, 53.9, 26.5. FT-IR: ν / cm^{-1} = 3051, 1670, 1602, 1503, 1473, 1420, 1388, 1328, 1197, 1127, 895, 812, 799, 767, 720. $[\alpha]_D^{20}$: not possible to determine due to strong absorption. mp: 92 °C (decomposition). ESI-MS: *m/z* calculated for [C$_{35}$H$_{33}$N$_7$O$_5$+H]$^+$ ([M+H]$^+$) = 632.3, found: 632.0. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O = 30:70 + 0.1% HCOOH, t_R = 3.52 min).

((3a*R*,4*R*,6*R*,6a*R*)-6-(1-Benzyl-6-imino-1,6-dihydro-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro-[3,4-*d*][1,3]dioxol-4-yl)methanol (88)
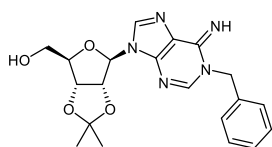
Using a procedure from SARFATI and KANSAL, *Tetrahedron* **1988**, *44*, 6367–6372.

To a solution of 2',3'-*O*-isopropylideneadenosine (**33**, 1.00 g, 3.25 mmol, 1.0 equiv.) in dry DMF (10 mL) was added benzyl bromide (1.26 mL, 10.60 mmol, 3.26 equiv.) under argon atmosphere. After stirring overnight at room temperature, the solvent was removed under reduced pressure at 40 °C. The residue was taken up in DCM, and the resulting precipitate was filtered off and washed two times with DCM to yield the desired product as a colorless solid (933 mg, 2.35 mmol, 72%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 10.50 (s, 1H), 10.12 (s, 1H), 8.95 (s, 1H), 8.29 (s, 1H), 7.59–7.50 (m, 2H), 7.43–7.32 (m, 3H), 6.34–6.24 (m, 2H), 5.90 (d, *J* = 15.3 Hz, 1H), 4.91 (dd, *J* = 5.8, 1.5 Hz, 1H), 4.61–4.55 (m, 1H), 4.51–4.46 (m, 1H), 4.15–4.05 (m, 1H), 3.83–3.72 (m, 1H), 1.65 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 150.3, 147.9, 146.4, 142.9, 132.2, 129.6, 129.5, 118.2, 114.3, 92.4, 86.6, 86.4, 82.0, 62.0, 56.0, 27.6, 25.5. FT-IR: ν / cm^{-1} = 3288, 3055, 2988, 2938, 2359, 1684, 1574, 1506, 1422, 1374,

1210, 1154, 1080, 850, 705. $[\alpha]_D^{20}$ = +60 (10 mg/mL; CHCl$_3$). mp: 186 °C (decomposition). R$_f$: 0.31 (DCM/MeOH = 30:1).

((3a*R*,4*R*,6*R*,6a*R*)-6-(6-(Benzylamino)-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]-dioxol-4-yl)methanol (87)
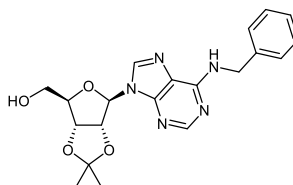
Method A: Using a modified procedure from SARFATI and KANSAL, *Tetrahedron* **1988**, *44*, 6367–6372.

To compound **88** (973 mg, 2.45 mmol) was added DMA (4 M in MeOH, 20 mL) and the resulting solution was stirred at room temperature for 24 h. Additional 20 mL of DMA solution were added and stirring was continued for 3 d at room temperature. The solvents were removed under reduced pressure at 40 °C yielding the desired product as a colorless solid (973 mg, 2.45 mmol, quantitative).

Method B: Using a modified procedure from FELLS *et al*., *Bioorg*. *Med*. *Chem*. **2008**, *16*, 6207–6217.

To a solution of *N*[6]-benzyladenosine (**86**, 1.08 g, 3.01 mmol, 1.0 equiv.) in acetone (300 mL) was added *p*-toluenesulfonic acid monohydrate (5.72 g, 30.08 mmol, 10.0 equiv.). After stirring overnight at room temperature, the precipitate was filtered off and NaHCO$_3$ solution (1 M, 50 mL) was added to the filtrate. The emulsion was stirred for 10 min and the solvents were removed under reduced pressure at 40 °C. Subsequent purification of the residue by column chromatography on silica (DCM/MeOH = 20:1) and on C$_{18}$ (MeCN/H$_2$O = 10:90 → 100:0) yielded the desired product as a colorless solid (596 mg, 1.50 mmol, 50%).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.34 (s, 1H), 7.60 (s, 1H), 7.40–7.24 (m, 5H), 5.80 (d, *J* = 4.9 Hz, 1H), 5.22–5.16 (m, 1H), 5.13–5.07 (m, 1H), 4.85 (s, 2H), 4.55–4.49 (m, 1H), 4.01–3.92 (m, 1H), 3.82–3.73 (m, 1H), 1.64 (s, 2H), 1.37 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 155.2, 152.9, 147.6, 139.8, 138.3, 128.9, 127.9, 127.7, 121.4, 114.1, 94.5, 86.2, 83.2, 81.9, 63.5, 44.6, 27.8, 25.4. FT-IR: ν / cm^{-1} = 3209, 3031, 2982, 2935, 2867, 1615, 1480, 1333, 1213, 1154, 1074, 849, 794, 744, 698. $[\alpha]_D^{20}$ = −92 (10 mg/mL; MeOH). mp: 56–58 °C. R$_f$: 0.31 (DCM/MeOH = 30:1).

177

9-((3a*R*,4*R*,6*R*,6a*R*)-6-(Azidomethyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-*N*-benzyl-9*H*-purin-6-amine (89)
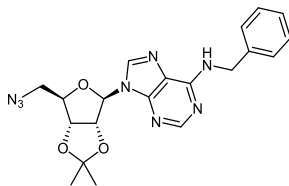
Using a modified procedure from IKEUCHI *et al.*, *Bioorg. Med. Chem.* **2009**, *17*, 6641–6650.

To a solution of compound **87** (556 mg, 1.40 mmol, 1.0 equiv.) in dry 1,4-dioxane (10 mL) were added diphenylphosphoryl azide (602 μL, 2.80 mmol, 2.0 equiv.) and DBU (627 μL, 4.20 mmol, 3.0 equiv.) at argon atmosphere. After stirring for 4 h at room temperature, 15-crown-5 (28 μL, 0.14 mmol, 0.1 equiv.) and sodium azide (455 mg, 7.00 mmol, 5.0 equiv.) were added. The mixture was refluxed for 4 h and then stirred at room temperature overnight. After filtration over Celite®, the solvent of the filtrate was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography on silica (DCM/MeOH = 50:1) and on C_{18} (MeCN/H$_2$O = 10:90 $\rightarrow$ 100:0) to yield the desired product as a colorless solid (380 mg, 0.90 mmol, 64%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.39 (s, 1H), 7.74 (s, 1H), 7.41–7.26 (m, 5H), 6.07 (d, *J* = 2.3 Hz, 1H), 5.45 (dd, *J* = 6.3, 2.3 Hz, 1H), 5.06 (dd, *J* = 6.3, 3.4 Hz, 1H), 4.98–4.79 (m, 2H), 4.41–4.33 (m, 1H), 3.64–3.49 (m, 2H), 1.61 (s, 3H), 1.39 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 154.9, 153.3, 148.5, 139.4, 138.5, 128.8, 127.8, 127.6, 120.5, 114.8, 90.8, 85.8, 84.2, 82.2, 52.4, 44.6, 27.2, 25.5. FT-IR: ν / cm^{-1} = 3346, 3092, 2979, 2931, 1732, 1595, 1553, 1513, 1369, 1291, 1227, 1141, 1030, 872, 749. $[\alpha]_D^{20}$ = +7 (10 mg/mL; MeOH). mp: 54–56 °C. R$_f$: 0.32 (DCM/MeOH = 50:1).

9-((3a*R*,4*R*,6*R*,6a*R*)-6-(Aminomethyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-*N*-benzyl-9*H*-purin-6-amine (90)
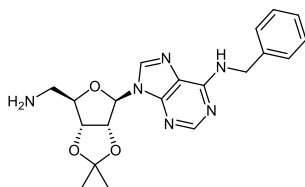
Using a modified procedure from ZHANG *et al.*, *Org. Biomol. Chem.* **2015**, *13*, 4149–4154.

To a solution of compound **89** (463 mg, 1.10 mmol) in EtOH (20 mL) was added palladium (10% on activated charcoal; 46 mg, 10 wt%) in an autoclave. Air inside the autoclave was removed under reduced pressure and hydrogen was added until a pressure of 3.5 bar was reached. The suspension was stirred at room temperature overnight und then filtered over Celite®. The solvent of the filtrate was removed under reduced pressure at 40 °C to yield the desired product as a colorless solid (410 mg, 1.03 mmol, 94%).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.35 (s, 1H), 7.65 (s, 1H), 7.39–7.21 (m, 5H), 6.57 (t, J = 6.0 Hz, 1H), 5.86 (d, J = 3.3 Hz, 1H), 5.74–5.51 (m, 2H), 5.31–5.20 (m, 2H), 4.92–4.73 (m, 2H), 4.48–4.41 (m, 1H), 3.37–3.16 (m, 2H), 1.59 (s, 3H), 1.36 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 154.9, 153.3, 147.9, 139.4, 138.6, 128.8, 128.0, 127.6, 120.5, 114.9, 91.8, 84.2, 83.4, 81.6, 44.6, 42.6, 27.3, 25.4. FT-IR: ν / cm^{-1} = 3271, 3029, 2983, 2934, 1614, 1479, 1375, 1330, 1212, 1154, 1074, 868, 794, 747, 699. $[\alpha]_D^{20}$ = −48 (10 mg/mL; CHCl$_3$). mp: 85–87 °C. R$_f$: 0.36 (DCM/MeOH = 10:1).

tert-Butyl (*S*)-4-((*N*-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-(benzylamino)-9*H*-purin-9-yl)-2,2-dimethyltetra-hydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)-4-bromo-3-nitrophenyl)sulfonamido)-2-((*tert*-butoxy-carbonyl)amino)butanoate (93)
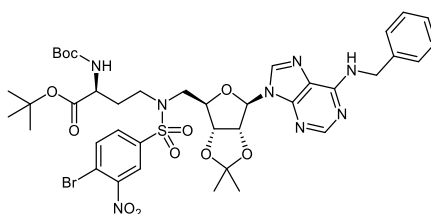
To a solution of compound **90** (410 mg, 1.03 mmol, 1.0 equiv.) in MeCN (6 mL) and THF (3 mL) was added **45** (257 mg, 0.94 mmol, 0.9 equiv.) and HOAc (54 μL, 0.94 mmol, 0.9 equiv.) at 5 °C. After the mixture was stirred at 5 °C for 30 min, sodium triacetoxyborohydride (299 mg, 1.41 mmol, 1.4 equiv.) was added. Stirring was continued at 5 °C overnight and then quenched by addition of MeOH (5 mL). The solvents were removed by distillation under reduced pressure at 40 °C, and the residue was purified by column chromatography on silica (DCM/MeOH = 30:1 + 1% NEt$_3$) yielding the secondary amine **91** (340 mg, 0.52 mmol, 55%), which was directly used in the next step. **91** (340 mg, 0.52 mmol, 1.0 equiv.) was reacted with 4-bromo-3nitrobenzenesulfonyl chloride (**92**, 197 mg, 0.66 mmol, 1.3 equiv.) according to general procedure B to afford the desired product as a colorless solid (170 mg, 0.19 mmol, 37%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.36 (s, 1H), 8.13–8.07 (m, 1H), 7.68 (s, 1H), 7.61–7.52 (m, 2H), 7.42–7.36 (m, 2H), 7.35–7.23 (m, 3H), 5.96 (d, J = 1.8 Hz, 1H), 5.37 (dd, J = 6.4, 1.8 Hz, 1H), 5.21–5.06 (m, 2H), 4.99–4.74 (m, 2H), 4.38–4.28 (m, 1H), 4.12–3.99 (m, 1H), 3.73–3.50 (m, 2H), 3.40–3.12 (m, 2H), 2.17–2.02 (m, 1H), 1.92–1.76 (m, 1H), 1.57 (s, 3H), 1.43 (s, 9H), 1.41 (s, 9H), 1.35 (s, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.8, 155.6, 154.8, 153.3, 149.6, 148.1, 140.7, 139.7, 138.4, 135.8, 130.9, 128.8, 127.9, 127.7, 124.3, 120.6, 119.2, 114.7, 90.8, 85.4, 84.3, 82.8, 82.7, 80.1, 52.1, 49.9, 45.5, 44.6, 31.9, 28.4, 28.0, 27.1, 25.3. FT-IR: ν / cm^{-1} = 3365, 3277, 3088, 2980, 2935, 1707, 1616, 1539, 1351, 1213, 1150, 1092, 869, 750, 662. $[\alpha]_D^{20}$ = +40 (10 mg/mL; CHCl$_3$). mp: 66 °C (decomposition). R$_f$: 0.53 (DCM/MeOH = 30:1).

(*S*)-2-Amino-4-((*N*-(((2*R*,3*S*,4*R*,5*R*)-5-(6-(benzylamino)-9*H*-purin-9-yl)-3,4-dihydroxytetrahydro-furan-2-yl)methyl)-4-bromo-3-nitrophenyl)sulfonamido)butanoic acid, trifluoroacetate salt (85)
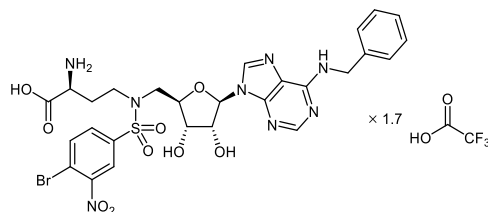
The compound was prepared from **93** (111 mg, 0.13 mmol) according to general procedure C to afford the final product as a yellowish trifluoroacetate salt (111 mg, 0.13 mmol, 99%, 1.7 equiv. TFA). ^{1}H NMR (300 MHz, CD$_3$OD): δ / ppm = 8.22 (s, 2H), 8.16–8.10 (m, 1H), 7.83–7.71 (m, 2H), 7.32–7.24 (m, 2H), 7.24–7.08 (m, 3H), 5.84 (d, *J* = 4.2 Hz, 1H), 4.83–4.64 (m, 2H), 4.57–4.44 (m, 1H), 4.21–4.13 (m, 1H), 4.12–4.03 (m, 1H), 3.86 (*pseudo*-t, *J* ≈ 6.4 Hz, 1H), 3.70–3.53 (m, 2H), 3.51–3.26 (m, 2H), 2.32–2.15 (m, 1H), 2.10–1.95 (m, 1H). ^{13}C NMR (75.5 MHz, CD$_3$OD): δ / ppm = 171.3, 151.2, 149.3, 142.7, 141.6, 138.0, 137.3, 132.4, 129.8, 128.9, 128.7, 125.2, 119.8, 119.7, 115.8, 90.9, 83.2, 75.0, 72.9, 52.1, 51.5, 46.3, 46.0, 31.1. FT-IR: ν / cm^{-1} = 3095, 2911, 1668, 1538, 1428, 1352, 1169, 1135, 1033, 969, 886, 834, 798, 721, 661. $[\alpha]_D^{20}$ = +26 (10 mg/mL; MeOH). mp: 73 °C (decomposition). ESI-MS: *m/z* calculated for $[C_{27}H_{29}BrN_8O_9S+H]^+$ ($[M+H]^+$) = 721.1, found: 721.0. Purity: 95% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 → 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 6.37 min).

5.3 Non-covalent docking and structural alignment

For non-covalent docking, crystal structure of DNMT2 was extracted from the Protein Data Bank (PDB-ID:1G55)[197,295] to *FlexX/LeadIT* (Version 2.3.2).[296,297] The binding site for docking was defined as a 6.5 Å sphere around the crystallographic reference ligand SAH. Other parameters remained set to the default settings of the software. 2D structures were transferred into energy minimized 3D structures using the *MMFF94x* force field implemented in *MOE* (*Molecular Operating Environment*, 2019.01).[298] Docking was performed using protonation states in aqueous solution as recommended by *FlexX/LeadIT*. The docking setup was validated by re-docking of SAH (RMSD: 0.80 Å). Top 10 poses were visually inspected for interaction profiles with the SAM- and cytidine binding sites. Structural alignments of crystal structures or AlphaFold models were conducted using *PyMOL* (Version 2.3.0).[299]

Part II

► **Development of covalent peptidomimetics against *Schistosoma mansoni* by ligand-based drug design**

6 Covalent peptidomimetics against *Schistosoma mansoni*

6.1 Introduction

6.1.1 Global distribution of schistosomiasis

Schistosomiasis, also known as bilharzia, is an infectious tropical disease caused by blood-dwelling fluke worms of the genus *Schistosoma*.[300,301] It is one of the most significant parasitic diseases affecting more than 230 million people worldwide according to estimates.[302–304] For humans, six *Schistosoma* species are relevant that cause severe diseases.[305] These species differ biologically and in their geographical distribution (**Figure 28**). Notably, they cause different types of symptoms.[302] Intestinal schistosomiasis is caused by *Schistosoma mansoni* (occuring in 53 countries in Africa, the Middle East, the Caribbean and South America), *Schistosoma japonicum* (occurring in China, Indonesia, and the Philippines), *Schistosoma mekongi* (occurring in Cambodia and Laos), and *Schistosoma guineensis* as well as its related species *Schistosoma intercalatum* (both occurring in rain forest areas of central Africa). Another form of schistosomiasis, urogenital schistosomiasis, is caused by *Schistosoma haematobium,* which occurs in Africa, the Middle East, and Corsica.[305]
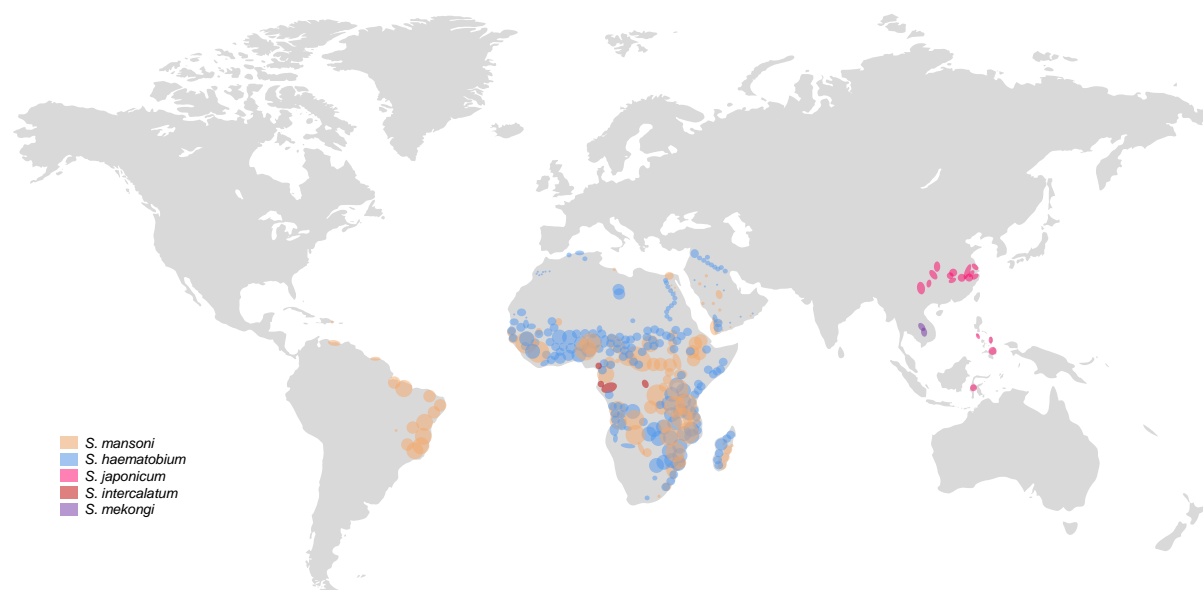
Figure 28: Global distribution of Schistosoma species.[301]

6.1.2 Symptoms and pathology

The pathogenesis of schistosomiasis can manifest itself in an acute as well as in a chronic form. Acute schistosomiasis, often referred to as the KATAYAMA syndrome, appears between 14–84 days after first exposal of non-immune individuals to first infection or heavy reinfection with schistosome cercariae.[306,307] The disease is a hypersensitivity reaction against the migration of schistosomula and egg deposition, which leads to the release of egg antigen and the host's granulomatous and immune complex response.

Individuals often suffer from nocturnal fever, cough, malaise, myalgia, headache, eosinophilia, fatigue, and abdominal pain.[300,306]

The morbidity of chronic schistosomiasis is not caused by adult worms, but by the host's immune response to the schistosomes eggs.[308–311] Many eggs are not excreted and are trapped in the intestines or liver (*S. mansoni, S. japonicum, and S. mekongi*) or in the bladder and urogenital system (*S. haematobium*). There, the eggs secrete proteolytic enzymes inducing eosinophilic inflammatory and granulomatous reactions, which result in fibrotic deposits in host tissues.[301,308–311] As a consequence, this causes an intestinal or urogenital schistosomiasis. Individuals with intestinal schistosomiasis suffer from abdominal pain, diarrhea, and blood in stool. In advanced cases, enlargement of the liver or spleen is common. The most commonly observed symptom of urogenital schistosomiasis is blood in urine (haematuria). In advanced cases, kidney damage as well as fibrosis of the bladder and ureter can occur, followed by bladder cancer in later stages.[305]

6.1.3 Characteristics of *S. mansoni*

Adult schistosomes are worms with a white-greyish, cylindrical body that measure 7–20 mm in length.[301] They have two terminal suckers, a complex tegument, a blind digestive tract, and organs for reproduction. Schistosomes have separate sexes, unlike other trematodes. Male schistosomes exhibit a gynaecophoric canal that holds the longer and thinner female.[301] Schistosomes of the genus *mansoni* live an estimated average of 5–10 years in their human host, based on a statistical analysis that revealed a significantly lower egg output after more than 5 years.[312] Interestingly, in some cases this lifespan can be exceeded. In 1985, CHABASSE *et al.* reported the case of a 56-year-old man who was still infected after 37 years.[313] The worms digest blood cells and globulins, the debris are regurgitated in the human bloodstream. They have an anaerobic metabolism, which is mainly used for the movements of the male schistosomes and the egg production of the females.[314]

6.1.4 Schistosome life cycle

The schistosome life cycle (**Scheme 27**) is described in detail in several reviews.[300–302,314,315] Infection with *Schistosoma* occurs through contact with infested surface water containing free-living larval forms of the parasite (cercariae).[302,315] They penetrate the human skin, shed their bifurcated tails, and undergo several other changes transforming them into a new form, the schistosomula.[302,315] The schistosomula enter capillaries and lymphatic vessels, migrate with the bloodstream to the heart, through the lungs and into the liver and the portal vein.[301,314,315] There, they grow and reach sexual maturity.[302,315] The male and female worms pair and then migrate to the superior mesenteric veins (*S. mansoni*), the inferior mesenteric and superior hemorrhoidal veins (*S. japonicum*), or the vesical plexus and veins draining the ureters (*S. haematobium*) where egg production occurs.[302,315] The production usually begins four to six weeks after infection and continues for the life of the worm.[315] In this period, a female worm can produce

hundreds (*S. mansoni* and *S. haematobium*) to thousands (*S. japonicum*) of eggs per day.[301,314] Each egg contains a ciliated larva (miracidium) that secretes proteolytic enzymes allowing the eggs to migrate from the blood vessels through the adjacent tissues into the lumen of the intestine (*S. mansoni* and *S. japonicum*) or the bladder (*S. haematobium*).[301,314] Most eggs are trapped in the host tissues where they induce inflammation and die.[300,314] However, up to one third are excreted into the environment through feces (*S. mansoni* and *S. japonicum*) or urine (*S. haematobium*) and can stay viable for up to seven days.[301,314,315] If the eggs reach freshwater, free-living ciliated miracidia are released and swim around to search suitable intermediate hosts, usually freshwater snails.[300,301,314] The miracidium penetrates the snail where it undergoes asexual replication into multicellular sporocysts over a period of four to six weeks, shedding tens of thousands of infectious cercariae into the water.[300,301,314] At this point, the life cycle is completed. The cercariae leave the snail under the stimulation of light and whirl around the water for up to 72 hours. It mainly occurs during daytime when the sun is high and human water contact most frequent.[301,314] Notably, one snail infected by one miracidium can release thousands of infectious cercariae into the water daily for months.[301]
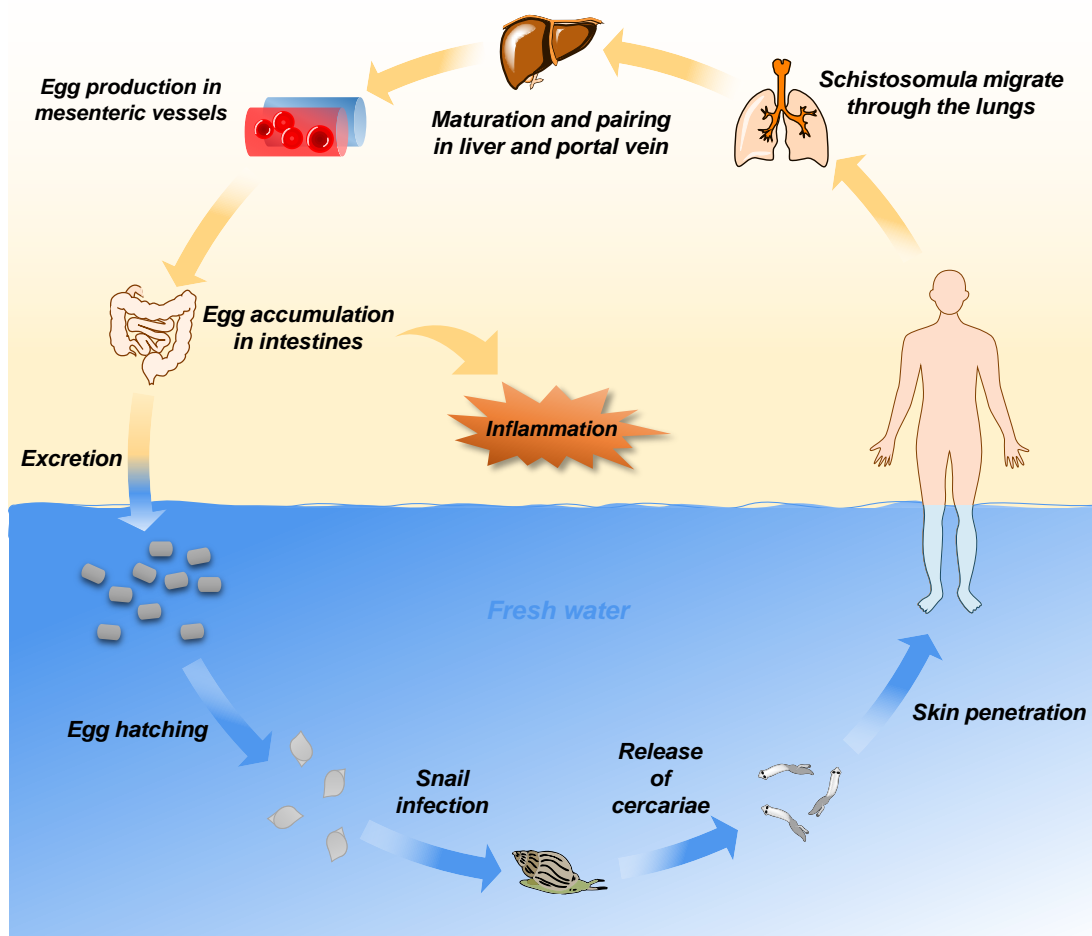
Scheme 27: Life cycle of *S. mansoni*.

6.1.5   Treatment of schistosomiasis

For the treatment of schistosomiasis, Praziquantel (**102**, PZQ, **Figure 29**) is the drug of choice. Originally, a library of compounds developed as potential antipsychotics by Merck (formerly E. Merck) was provided to Bayer for testing against worm parasites. In 1972, this led to the discovery of EMBAY-8440, which was further developed by both companies to yield PZQ. Finally, in 1980, the drug was approved as an anthelmintic under the brandname Biltricide[®].[316,317] PZQ is effective against all *Schistosoma* species, however, it acts against adult worms and poorly against immature larvae.[300] Although the drug is applied as a racemic mixture, the *R*-enantiomer is primarily responsible for the anthelmintic activity causing a rapid contraction of worms and damage to the worm surface.[318–321] The mechanism of action has not been understood for a long time. In 2019, a transient receptor potential (TRP) channel of the TRP melastatin (TRPM) subfamily activated by PZQ was identified in *S. mansoni*.[322] Two years later, a hydrophobic ligand binding pocket for PZQ within the voltage sensor-like domain of the channel could be found. This lead to the assumption that PZQ directly binds to *Sm.*TRPM causing calcium influx and worm paralysis.[323] For the treatment of most *Schistosoma* species, such as *S. mansoni* and *S. haematobium*, a standard dose of 40 mg/kg of body weight is recommended. Due to the low toxicity and high tolerance of PZQ,[318,324] even higher doses of 60 mg/kg can be applied, which is often required for *S. japonicum* and *S. mekongi* or intense diseases.[300,306,315] However, studies have shown that the dosage of 40 mg/kg and 60 mg/kg led to equivalent outcomes for infections either with *S. mansoni* or *S. japonicum*.[325] According to the WHO guidelines in 2022, the dose administered is now determined using the PZQ dose pole considering body measurements.[326] PZQ achieves moderate to high cure rates of 59–90% with field studies showing dose dependence in some cases.[318,327] Notably, an effective host antibody response is crucial for full treatment efficacy.[328,329] Other drugs used for the treatment of schistosomiasis are oxamniquine (**103**) and metrifonate (**104**) (**Figure 29**), which are only applied against specific species, *S. mansoni* and *S. haematobium*, respectively.[330,331] Unfortunately, these drugs are no longer readily available.[300,315,332]

Praziquantel (PZQ, 102)   **Oxamniquine (103)**   **Metrifonate (104)**
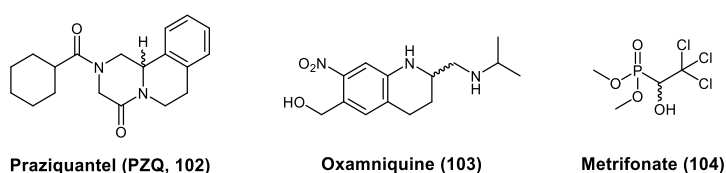
Figure 29: Structure of praziquantel (PZQ), oxamniquine, and metrifonate. All drugs are applied as racemic mixtures.

For 40 years, treatment of schistosomiasis has relied primarily on PZQ, which significantly increases concerns about drug resistance.[333] There are already published field studies showing unexpected low cure rates after treatment of *S. mansoni* with PZQ.[334,335] Furthermore, reduced susceptibility of *S. mansoni* and *S. hematobium* isolates to the drug has already been observed.[336] To date, no other therapeutic options are available, leading the WHO to call for the development of new interventions, including alternatives to praziquantel.[337]

6.1.6 *S. mansoni* cathepsin B1 – a possible target

S. mansoni cathepsin B1 (*Sm*CB1) is the most abundant papain-like cysteine protease in the lumen of the schistosomes gut and has also been found in the gastrodermal cells.[338,339] Like other cathepsin B enzymes, it has exopeptidase as well as endopeptidase functionality,[340] and is required for the digestion of host hemoglobin,[341] and necessary for normal schistosome growth.[342] The enzyme is synthesized as an inactive 41 kDa glycosylated zymogen that needs to be activated by proteolysis of the pro-peptide to yield the mature 31 kDa glycosylated protein. The pro-peptide cleavage can be catalyzed by another gut-associated cysteine peptidase, the *S. mansoni* asparaginyl endopeptidase (*Sm*AE, also known as legumain).[339] Although this step has been found to be catalyzed by *Sm*AE, the protease is not essential for the activation *in vivo*.[343] *Sm*CB1 consists of a single peptide chain with a papain-like fold divided into L and R domains. At the interface between those domains, the active site is located with the catalytic triad consisting of Cys100, His270, and Asn290.[344,345] Like other cathepsin B peptidases, *Sm*CB1 has an "occluding loop" formed by residues 175–194, a well-known substructure element that restricts access to the primed region of the active site. Two histidines (His180 and His181) located in the center of the occluding loop are required for important interactions with the C-terminal end of the substrate. Depending on the pH value, the protonation of these two histidines affects the endo- or exopeptidase activity of cathepsin B.[346,347] **Scheme 28** illustrates the cleavage mechanism proceeded by *Sm*CB1.[348]
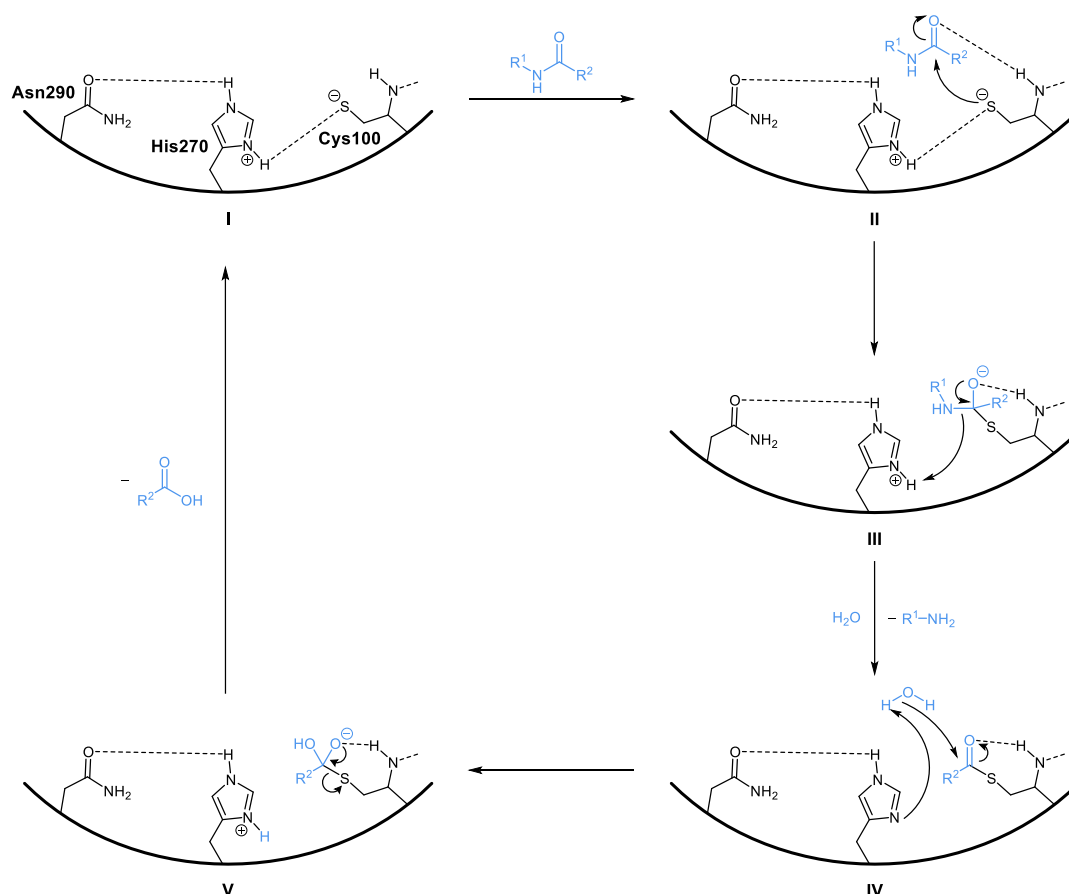
Scheme 28: Peptide cleavage mechanism proceeded by *Sm*CB1.

It starts with the nucleophilic attack of Cys100 at the peptide bond, resulting in the formation of a tetrahedral intermediate with the peptide substrate. Then, cleavage of the peptide bond occurs, releasing the N-terminus of the C-terminal cleavage product. The remaining acyl intermediate is hydrolyzed by nucleophilic attack of a water molecule to finally yield the C-terminus. During the mechanism, Asn290 ensures the correct positioning of His270 that is required for proton transfer.[348]

Meanwhile, *Sm*CB1 has been validated as a chemotherapeutic target for the cure of schistosomiasis.[344,349] In 2011, *Sm*CB1 was co-crystallized in complex with **K11777** (**Figure 30**),[344] a literature-known, covalent inhibitor of cruzain, a cysteine protease of the protozoan parasite *Trypanosoma cruzi*, and cathepsins B and L.[350–353] **K11777** was found to inhibit *Sm*CB1 by treatment of infected mice.[349] This structure features a peptidomimetic sequence connected to a vinylsulfone warhead that traps the catalytically active nucleophile, Cys100, inside the active site of the enzyme. The reaction follows an irreversible Michael addition of the thiol group at the β-position of the vinylsulfone system,[354] that is located in the S1' subsite (according to the nomenclature of SCHECHTER and BERGER).[355] The homophenylalanine moiety is placed in the S1 subpocket, the phenylalanine in S2, and the terminal *N*-methylpiperidine group in S3.[344]
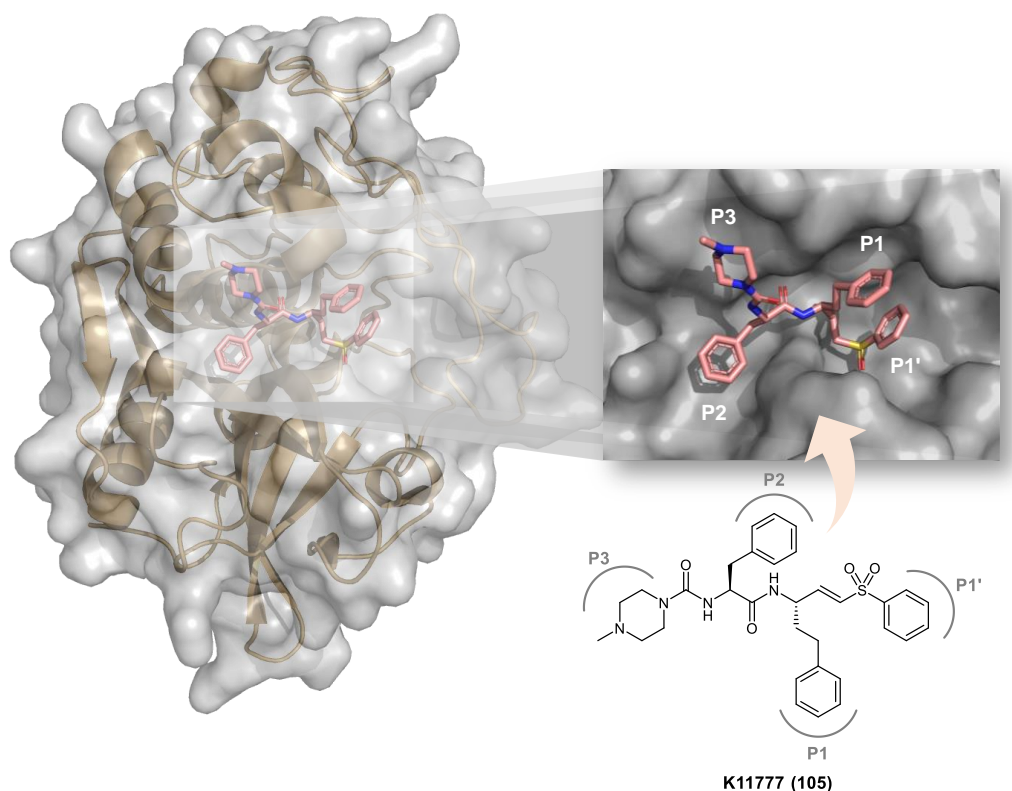
Figure 30: Left: *Sm*CB1 in complex with **K11777** (cyan), PDB-ID:3S3R.[344] Right: 2D Structure of **K11777** with subpocket-targeting units labeled according to the SCHECHTER and BERGER nomenclature.[355]

The epoxide-based covalent inhibitor **CA074** (**Figure 31**) was also found to inhibit *Sm*CB1. Interestingly, this structure was designed to address the S1' and S2' sites of cathepsin B, especially allowing interactions with His180 and His181 of the occluding loop with a carboxyl function.[344,356,357] Recently, further peptidomimetic *Sm*CB1 inhibitors based on azanitrile warheads were developed that showed lethal activity against *S. mansoni* newly transformed schistosomula (NTS).[358]
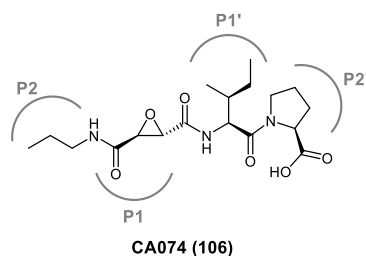
CA074 (106)

Figure 31: Structure of **CA074** (**106**) with subpocket-targeting units labeled according to the SCHECHTER and BERGER nomenclature.[355]

6.2   Summary and own contribution

Note: The substance numbers from the corresponding manuscript are listed *in italics* behind the substance numbers in this thesis.

To find new anti-schistosomal lead structures, an in-house library consisting of 76 cysteine-targeting inhibitors was screened in a phenotypic assay on newly transformed schistosomula (NTS) and *S. mansoni* adults as well as for the inhibitory effect on *Sm*CB1 in fluorometric enzyme assay. The phenotypic screening revealed three hits, two vinylsulfonate-based structures (**SJ605** and **SJ606**) and the 4-oxoenoate-based diastereomeric bortezomib congener **KS309** (**Table 8**).

Table 8: Structures of **KS309**, **SJ605**, and **SJ606** with indication of anti-schistosomal activity and inhibition of *Sm*CB1 in the fluorometric enzyme assay.
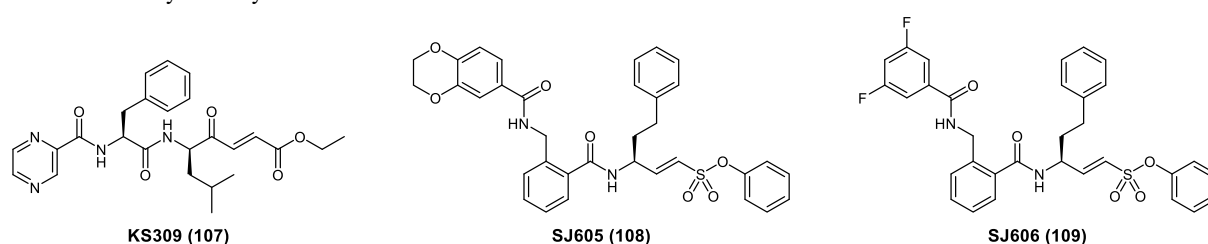
KS309 (107)   SJ605 (108)   SJ606 (109)

Compound	*Sm*CB1 inhibition* K_i / μM	NTS % effect** at 10 μM	*S. mansoni* adults % effect** at 10 μM	at 1 μM
KS309	>100	100	76	29
SJ605	0.0018	100	55	n. d.
SJ606	0.0034	100	58	n. d.

*Measured by ▉▉▉▉▉▉ and ▉▉▉▉▉ (▉▉▉▉▉ group); **dead after 72 h; n. d. = not determined.

Although **KS309** showed weak inhibition of *Sm*CB1 (K_i >100 μM), it was taken as a lead structure for further optimization as it had the highest effect against adult worms (76% at 10 μM, 29% at 1 μM). To maximize the effect, lipophilicity was increased by methylation of amides (*1a* and *1b*) as well as P3 substitution (*1c* and *1d*) to improve cell permeability and reach higher compound concentrations in the worm organism for a stronger effect (**Figure 32**). Furthermore, alternative warheads (*1e–h*), an azapeptide (*1k*), and small molecule derivatives were analyzed (*1l* and *1m*).

Based on this strategy, three optimized compounds were identified (vinylsulfonate **110** (*1g*) as well as its precursors **111** (*1j*) and **112** (*1i*)) exhibiting higher efficacy against *S. mansoni* adults (93–100% at 10 μM, 36–45% at 1 μM) compared to **KS309**. However, their inhibitions of *Sm*CB1 were low (9–80% at 20 μM). To facilitate potential optimizations of these compounds in future studies, structural knowledge about the target would be beneficial. For this reason, optimization of **SJ605** and **SJ606** may be a more promising strategy as their effects appeared to correlate with *Sm*CB1 inhibition. As part of this project, derivatives of **SJ605** and **SJ606** were developed by ▉▉▉▉▉▉ (▉▉▉▉▉ group).
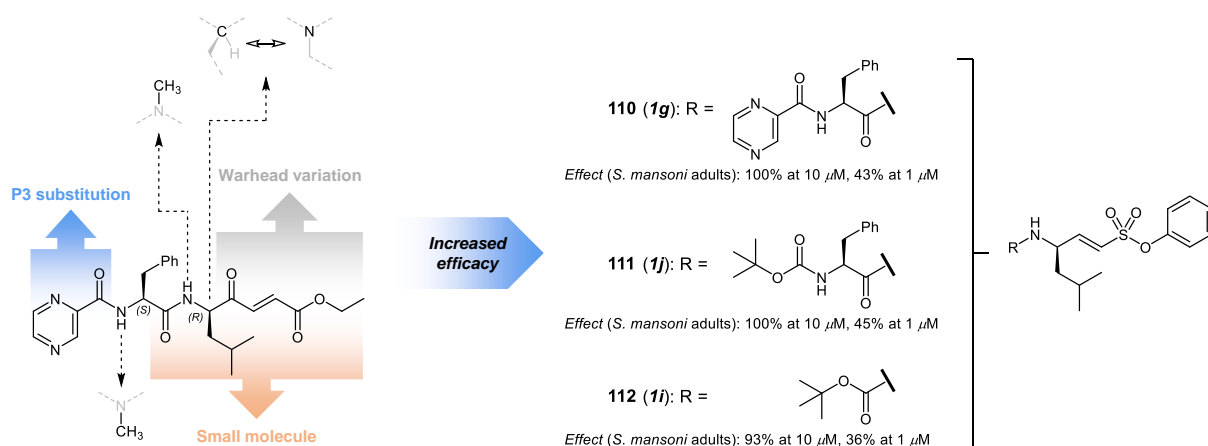
Figure 32: Left: Variation of **KS309**. Right: Structures with increased anti-schistosomal effect.

For a more detailed discussion, as well as the presentation of all experimental procedures, the reader is referred to the corresponding manuscript (**Section 6.3**) and Supporting Information (**Appendix**) Synthetic procedures conducted by MARVIN SCHWICKERT (　　　　　group) are described in **Section 6.4**. The corresponding experimental section can be found in **Section 6.5**.

Own contribution: Design and synthesis of inhibitors *1a–m*, writing the corresponding parts in the manuscript and Supporting Information, creation of **Schemes 1**–**6** in the Supporting Information.

6.3   Manuscript

Dual Strategy to Design New Agents Targeting *Schistosoma mansoni*: Advancing Phenotypic and *Sm*CB1 Inhibitors for Improved Efficacy

Natalie Fuchs,[1*] Robert A. Zimmermann,[1*] Marvin Schwickert,[1*] Annika Gunkel,[1] Collin Zimmer,[1] Mergim Meta,[1] Kevin Schwickert,[1] Jennifer Keiser,[2] Werner Kiefer,[1] and Tanja Schirmeister[1*]

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

[2]Swiss Tropical and Public Health Institute, Kreuzstrasse 2, 4123 Allschwil, Switzerland

*All correspondence to Prof. Dr. Tanja Schirmeister, schirmei@uni-mainz.de

[*]N. F., R. A. Z., and M. S. contributed equally.

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/or its target protease *S. mansoni* cathepsin B1 (*Sm*CB1). Our correlation analysis highlighted the significance of physicochemical properties for the compounds' efficacy, resulting in a dual approach to optimize the lead structures regarding both phenotypic effects and *Sm*CB1 inhibition. The optimized compounds from both approaches ("phenotypic" vs. "*Sm*CB1" approach) demonstrated improved efficacy against *S. mansoni* adult worms, with **2h** from the "*Sm*CB1" approach emerging as the most potent compound. **2h** displayed nanomolar inhibition against *Sm*CB1 ($K_i = 0.050$ μM) while maintaining selectivity towards human off-target cathepsins, and greatly improved efficacy towards *S. mansoni* adults (86% at 10 μM, 68% at 1 μM), demonstrating potential as a new therapeutic agent for schistosomiasis.

2h
NTS 10 μM: 88%
S. mansoni adults 10 μM: 86%
S. mansoni adults 1 μM: 68%
K_i(*Sm*CB1) = 0.050 μM
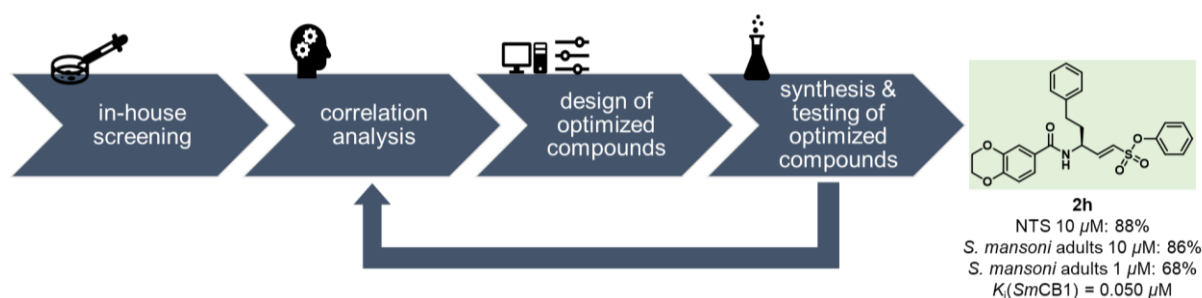
INTRODUCTION

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 140 million patients are affected by this disease worldwide, most of them in sub-Saharan Africa.[1,2] 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.[3] Infections can lead to intestinal or hepatic forms of schistosomiasis.[4] 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.[3] 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.[3] Female schistosomes are capable to produce hundreds of eggs daily.[5,6] Egg migration into the intestinal lumen where they are excreted with faeces 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.[5] Fecal excretion of eggs is not quantitative, meaning that a considerable number of eggs remain in the human body. An early study with hamsters as mammalian hosts found that only ~20% eggs were found in faeces, while ~80% of the eggs remained in the body.[6] After 1–2 weeks, the eggs will die, independent if they reach freshwater.[3] The 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.[7,8]

Today, the only approved drug for the treatment of schistosomiasis is praziquantel (PZQ). Although praziquantel was introduced in the 1970s, very few cases of resistance have been reported.[9] Until recently, even attempts to induce a resistance against PZQ in the laboratory has failed.[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 ion channel is the target of praziquantel.[11] However, to control or even eradicate this disease, it seems necessary to pursue further drug development to also 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 small molecular 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 that arises from this target.[15]

To date, several *Sm*CB1 inhibitors have been reported. One well-known example is the pan-cathepsin inhibitor **K11777** ($IC_{50} = 0.0021\ \mu M$, $k_{2nd} = 8.8 \cdot 10^4\ M^{-1}s^{-1}$), an irreversible vinylsulfone.[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, $k_{2nd} = 2.1 \cdot 10^5\ M^{-1}s^{-1}$) and **WRR-286** ($IC_{50} = 0.6$ nM, $k_{2nd} = 2.0 \cdot 10^5\ M^{-1}s^{-1}$).[13,16] The compounds are effective against newly transformed schistosomula (NTS) at concentrations of 10 μM for both and 1 μM for **WRR-286** after 72 hours.[13] Additionally, they 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 anti-schistosomal compounds, we developed a workflow (see **Graphical Abstract**) starting with an in-house screening including 76 compounds that provided 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 using molecular docking approaches and input from literature. We synthesized several optimized structures and tested them on the target enzyme as well as in the phenotypic assay.

RESULTS AND DISCUSSION

In-house screening

We initiated our search for lead structures with anti-schistosomal activity by screening our in-house library including 76 compounds in a phenotypic assay on newly transformed schistosomula (NTS) and *S. mansoni* adults.[17] This library consisted of several cysteine-targeting inhibitors decorated with different covalent warheads, such as (F-)vinylsulfon(at)es, nitriles, aldehydes, 4-oxoenoates, nitroalkenes, and acrylamides.[18–23] 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 with *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 was achieved, a second screening was performed at 1 μM on NTS and at 10 μM on *S. mansoni* adults. Lower concentrations of 1 μM and 0.1 μM were tested on adult worms when 70% efficacy was achieved in the respective previous screening. A table with all data from the in-house screening can be found in the **Supporting Information**. The procedure of the initial screenings is described in **Figure 1**.
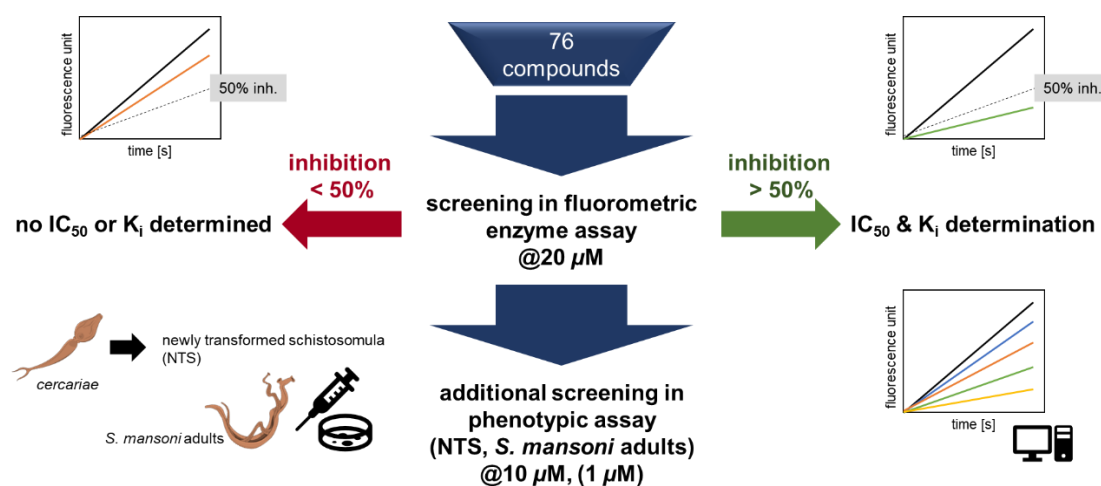
Figure 1: Workflow of the initial in-house screening. All compounds were screened for their *Sm*CB1 inhibition in a fluorometric enzyme assay as well as their efficacy in a phenotypic assay against NTS and adult worms.

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 inhibitors (**SJ605**, **SJ606**), originally developed for rhodesain,[19,20] the third is a 4-oxoenoate-based diastereomeric bortezomib congener (unpublished results). All three compounds were highly efficient against NTS at 10 μM (100% efficacy). Although **SJ605** and **SJ606** exhibited high *Sm*CB1 inhibition with K_i values in the low nanomolar range, only moderate efficacy against adult worms at 10 μM (55% for **SJ605**, 58% for **SJ606**) was achieved. In contrast, **KS309** showed weak inhibition of *Sm*CB1 (K_i >100 μM) but a strong effect on adult worms (76% at 10 μM, 29% at 1 μM) in the phenotypic assay, indicating that it addresses another target.

KS309
NTS (10 μM): 100%
S. mansoni adults (10 μM): 76%
K_i (*Sm*CB1) = >100 μM

SJ605
NTS (10 μM): 100%
S. mansoni adults (10 μM): 58%
K_i (*Sm*CB1) = 0.0034 μM

SJ606
NTS (10 μM): 100%
S. mansoni adults (10 μM): 55%
K_i (*Sm*CB1) = 0.0018 μM
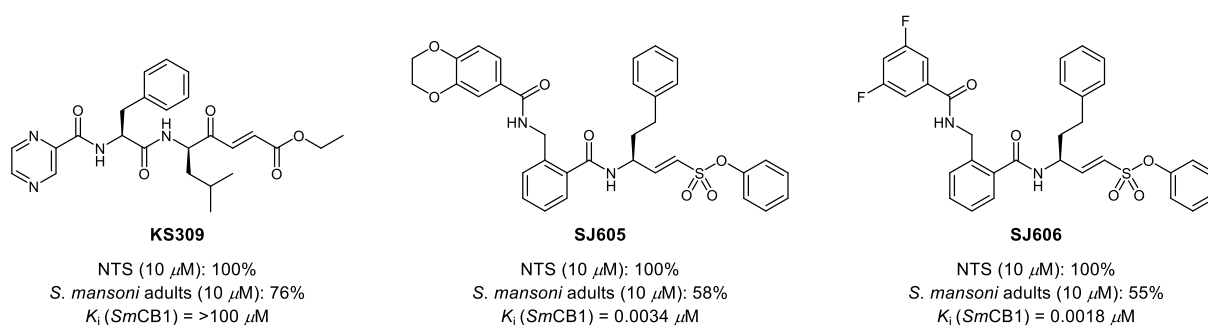
Figure 2: Phenotypic screening of in-house library – top three compounds with screening results and K_i values.

Based on these findings, we pursued a dual strategy to increase the anti-schistosomal effect of both compound classes on adult *S. mansoni*. In order to maximize their effect, we focused primarily on improving their physicochemical properties, such as lipophilicity, since the substances have to pass two cell membranes to reach their target site.[24] This resulted in two approaches, the "phenotypic" approach focusing on **KS309** derivatization and the "*Sm*CB1" approach including **SJ605**/**SJ606** derivatizations. The syntheses are described in the **Supporting Information**.

Derivatization of KS309

In our first approach, we investigated various changes in the structure of **KS309** (**Figure 3**). To increase lipophilicity, we reduced the number of hydrogen bond donors by methylating one of each of the amide bonds (**1a**, **1b**). Furthermore, the effect of replacing the polar pyrazinoyl moiety in P3 with 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, 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, small molecule derivatives (**1l**, **1m**) of **KS309** and **1k** were tested.
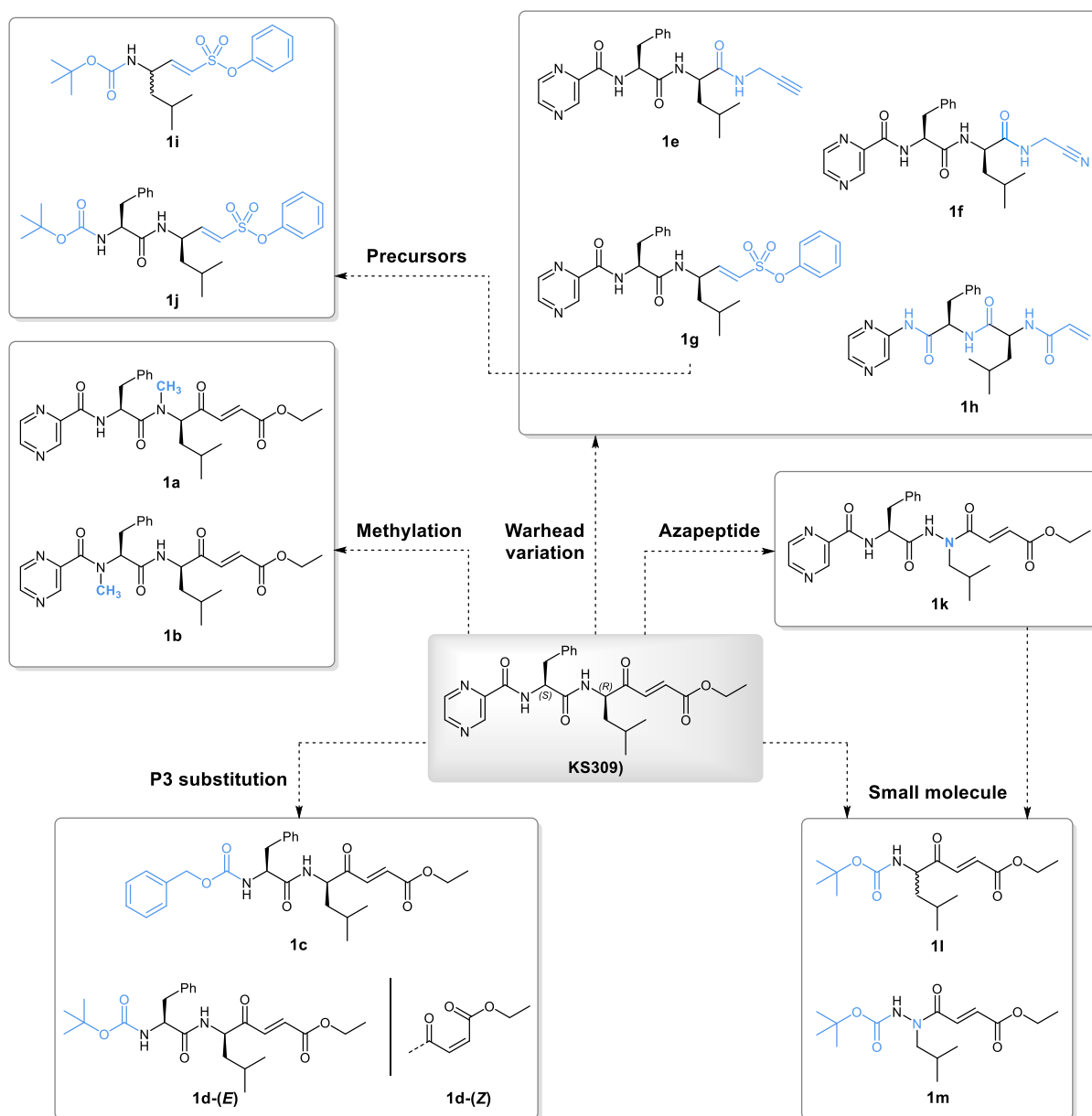
Figure 3: Derivatization of **KS309**.

The **KS309** derivatives were evaluated 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 SlogP values and topological polar surface areas (TPSA) are given.

Table 1: Physicochemical parameters and *in vitro* effects of **1a**–**m**.

Cpd.	SlogP[a]	TPSA / Å²	*Sm*CB1 % inhibition at 20 μM	NTS % effect[b] at 10 μM	NTS % effect[b] at 1 μM	*S. mansoni* adults % effect[b] at 10 μM	*S. mansoni* adults % effect[b] at 1 μM
KS309	2.04	127	58	100 ± 0	23 ± 12	76 ± 6	29 ± 2
1a	2.38	119	11	100 ± 0	48 ± 0	61 ± 4	n. d.
1b	2.38	119	12	96 ± 0	44 ± 6	48 ± 0	n. d.
1c	4.01	111	67	98 ± 2	15 ± 8	65 ± 2	36 ± 6
1d-(*E*)	3.34	111	42	100 ± 0	78 ± 2	61 ± 0	n. d.
1d-(*Z*)	3.34	111	47	100 ± 0	40 ± 0	40 ± 2	n. d.
1e	1.10	113	18	46 ± 0	n. d.	n. d.	n. d.
1f	0.99	137	10	39 ± 0	n. d.	n. d.	n. d.
1g	3.27	127	68	100 ± 0	38 ± 4	100 ± 0	43 ± 0
1h	1.86	113	9	48 ± 2	n. d.	n. d.	n. d.
1i	3.85	82	9	98 ± 2	27 ± 4	93 ± 4	36 ± 2
1j	4.58	111	80	100 ± 0	65 ± 0	100 ± 0	45 ± 2
1k	1.45	131	17	50 ± 13	33 ± 5	47 ± 10	n. d.
1l	2.61	82	7	100 ± 0	14 ± 2	58 ± 2	n. d.
1m	2.03	85	3	61 ± 4	32 ± 2	34 ± 0	n. d.

a: calculated using *MOE* 2019.01;[25] b: % dead after 72 h.

Regarding the results in **Table 1**, **1g**, **1i**, and **1j** showed higher efficacy on *S. mansoni* adults compared to **KS309** (93–100% vs. 76% at 10 μM; 36–45% vs 29% at 1 μM). **1a**, **1b**, and **1d-(*E*)** caused 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 in TPSA compared to **KS309** (<127 Å²) and the increase in efficacy on adult worms could not be observed. Although **1i** and **1j** showed higher efficacy with TPSAs of 111 Å² and 82 Å², respectively, several compounds with TPSAs ranging from 82 Å² to 119 Å² (**1a**–**c**, **1d-(*E*)**, **1d-(*Z*)**, **1e**, and **1h**) were less efficient. On the other hand, a statistically significant correlation (described as correlation factor *r*) between SlogP value and efficacy against NTS at 10 μM could be observed (*r* = 0.84, *p* = 0.0001), suggesting that higher SlogP values resulted in higher efficacy (see **Table 4**, **Figure S2** in the **Supporting Information**). Compounds with SlogP 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%. For SlogP values ranging from 0.99 to 2.03, efficacy was 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 efficacy in the compound series against 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 vinylsulfonate (**1g**) retained the same TPSA (127 Å²) but increased the SlogP value (3.27 vs. 2.04), which might have improved cell permeability. Furthermore, the vinylsulfonate might be less prone to off-target reactions compared to the highly reactive 4-oxoenoate substructure. Replacing the polar pyrazinoyl

moiety of **1g** in P3 with a more lipophilic Boc group (**1j**) did not change efficacy, although it decreased the TPSA (111 Å^2 vs. 127 Å^2) and increased the SlogP value (4.58 vs. 3.27, **1g** and 2.04, **KS309**). In summary, the anti-schistosomal effect could be increased, however, the target of **KS309** derivatives remains unknown. Therefore, we did not further focus on the phenotypic approach by optimizing this series, but rather on the optimization of the **SJ605/SJ606** inhibitors as their effects appeared to correlate with *Sm*CB1 inhibition.

Correlations and hit optimizations

Correlation analysis

For correlations and optimization ideas, we started by calculating several important physicochemical parameters (logP, TPSA etc.) using *MOE* (see **Table S1** in **Supporting Information**).[25] We then performed a correlation analysis for each physicochemical parameter or the K_i values with the observed effects from phenotypic screening, including NTS (10 μM, 1 μM inhibitor) and *S. mansoni* adults (10 μM inhibitor). The resulting correlations are shown in **Table S2** in the **Supporting Information** with their respective *p*-values. Only *p*-values <0.05 were considered statistically significant.

There was a slight, but statistically significant, correlation between K_i values and the effect in the phenotypic screening against NTS at 1 μM ($p = 0.025$). The correlation factor (r) was negative ($r = -0.43$), implying that lower K_i values lead to a higher effect. The negative correlation for K_i and the effect on *S. mansoni* adults was not significant ($p = 0.18$). The adult worms seemed to compensate the inhibition more effectively than the NTS. It was reported earlier that several proteases are involved in the digestion of blood haemoglobin. One explanation for this observation could be that inhibition of *Sm*CB1 alone can be compensated by other proteases such as *Sm* cathepsin D or *Sm* cathepsin L. Since the expression levels of proteases fluctuate in different stages of development, adult worms may overcome inhibition of *Sm*CB1 more effectively than NTS.[15,26] Furthermore an additional second membrane has to be crossed in adult worms.[27] The positive correlation factor between SlogP and the effect against NTS ($r = 0.44$ at 1 μM) indicated that higher logP values are beneficial ($p = 0.048, 0.021$). TPSA and molecular flexibility (KierFlex) did not have a significant effect on the screening results.

Since two compounds from the **SJ600s** series (**SJ605**, **SJ606**) were among the top three, we decided to investigate a sub-set containing only these 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*. (2022).[20]

Again, there was a significant correlation ($p = 0.005$) between the K_i value for *Sm*CB1 and the effect against NTS at 10 μM inhibitor concentration ($r = -0.67$) like the correlation described above. The SlogP correlates with the effect against NTS at 10 μM ($r = 0.66, p = 0.005$) as well. Here, we had an additional

correlation between the molecular flexibility (KierFlex) and the effect against NTS at 10 μM ($r = -0.54$, $p = 0.032$), suggesting that lower flexibility is advantageous. KIER'S definition of flexibility is entirely structure-based upon molecular size, branching, cycles and number of heteroatoms.[28]

The significant correlations for the complete compound set and the **SJ600** subgroup are shown in **Figure 4**, suggesting stronger correlations within the **SJ600** subgroup as can also be drawn from the correlation factors (e.g. NTS 10 μM vs. K_i: all compounds $r = -0.27$ vs. **SJ600s** $r = -0.67$).
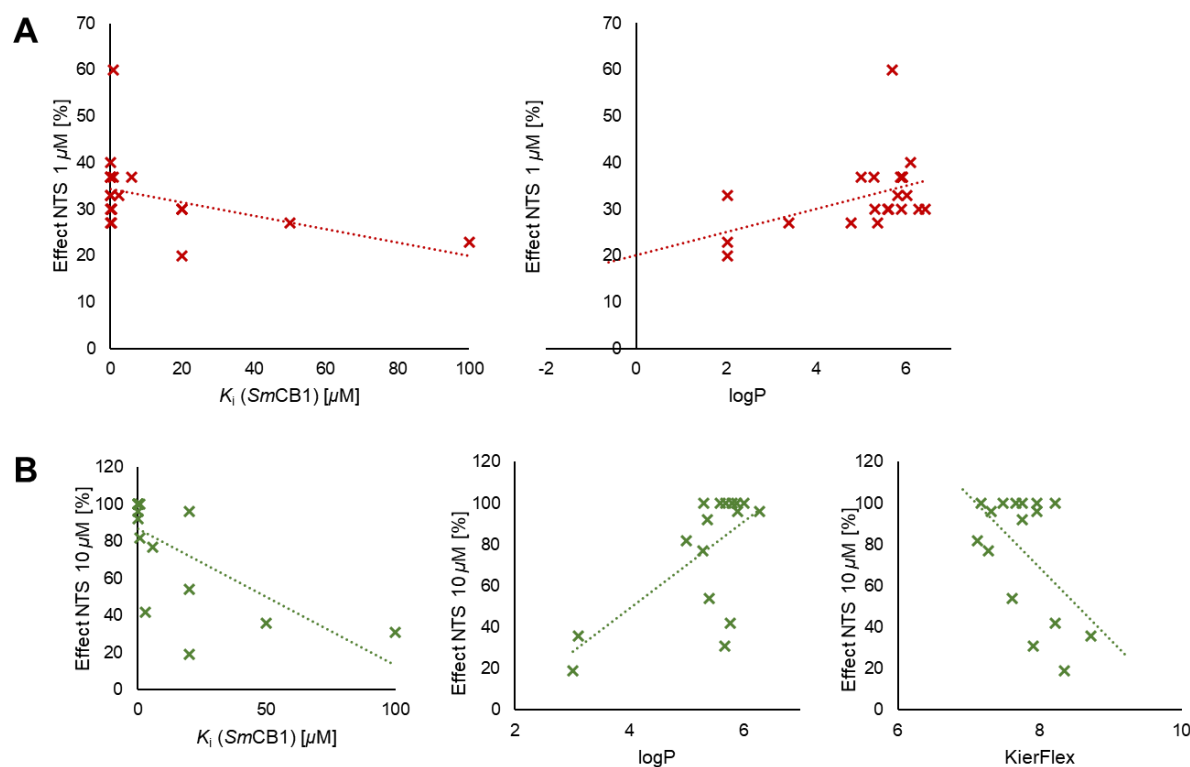
Figure 4: Correlations between K_i values or physicochemical parameters and the effect in the phenotypic screening. (A) All compounds (n = 76). Left: effect on NTS at 1 μM vs. K_i (*Sm*CB1) in μM in red (n = 27). Right: effect on NTS at 1 μM vs. logP in red (n = 27). (B) **SJ600s** (n = 16). Left: effect on NTS at 10 μM vs. K_i (*Sm*CB1) in μM in green (n = 16). Middle: effect on NTS at 10 μM vs. logP in green (n = 16). Right: effect on NTS at 10 μM vs. KierFlex in green (n = 16).

Inhibitor design

The correlation analysis suggested that a higher rigidity resulted in a higher activity in the phenotypic assay (**Figure 4**, **Table S3** in the **Supporting Information**). Nevertheless, the *Sm*CB1 inhibition should not be compromised since it also correlateed with the effect in NTS. **Figure 5** shows non-covalent docking poses (generated with *LeadIT*)[29] for lead structures **SJ605** ($K_i = 1.8$ nM) and **SJ606** ($K_i = 3.4$ nM), revealing that the most important interactions occur in the S1' and S1 subsites. Gly144 and Gly269 both interact with the amide group between P1 and P2, and Gln94 can form a H-bond with the sulfonyl-oxygen atom as well as Trp292. The residues in P2/P3 do not form essential interactions in the active site since the residues rather protrude from the binding site.
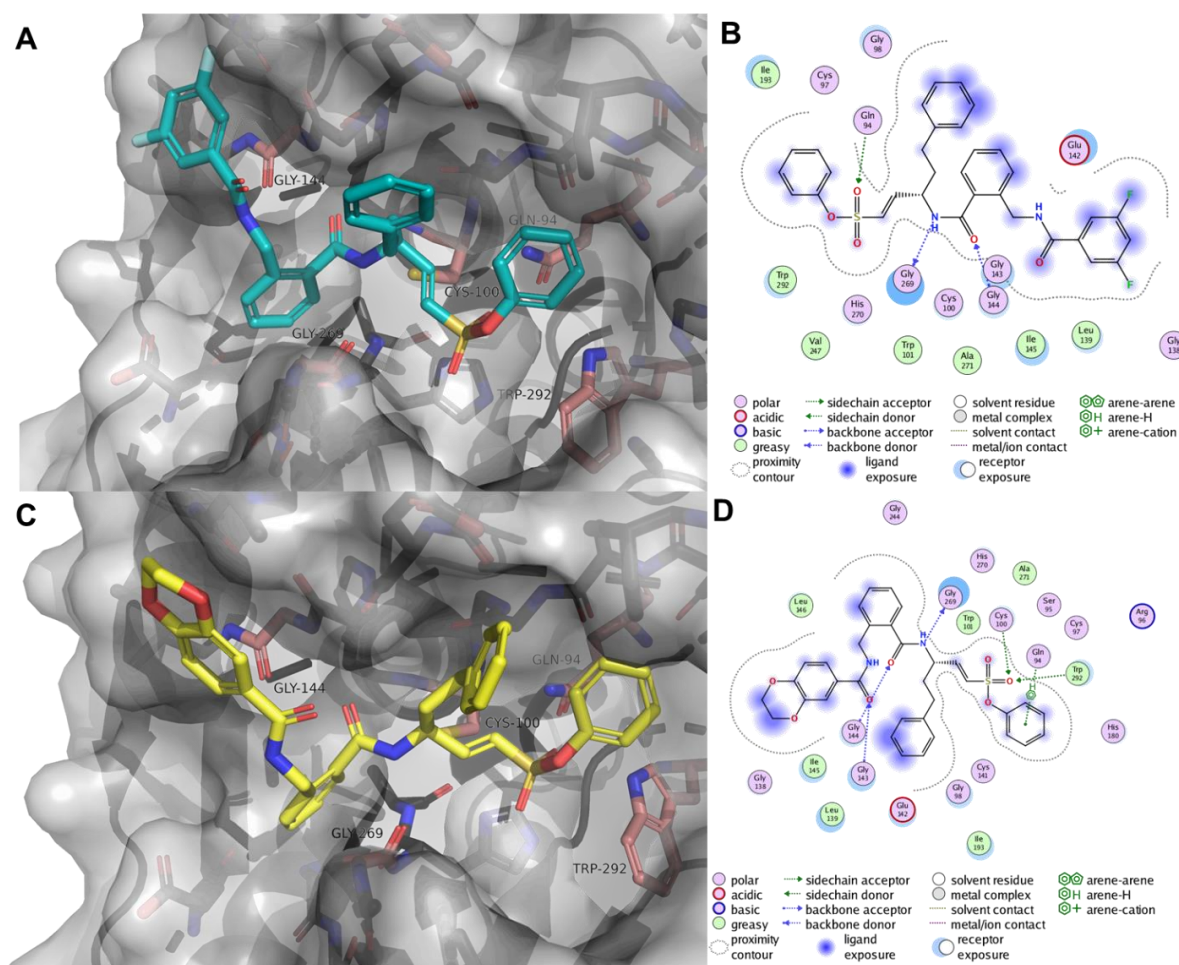
Figure 5: Non-covalent docking for lead compounds **SJ605** and **SJ606** with *LeadIT*.[29] Images generated using *PyMOL*.[30] Ligand interaction map generated with *MOE*.[25] PDB-ID: 3S3R.[16] (A) Docking pose for **SJ605** (teal). Important interaction partners shown in salmon. Hyde score: −34 kJ/mol. (B) Ligand interactions of **SJ605** with the *Sm*CB1 active site. (C) Docking pose for **SJ606** (yellow). Important interaction partners shown in salmon. Hyde score: −25 kJ/mol. (D) Ligand interactions of **SJ606** with the *Sm*CB1 active site.

Based on these interactions, we decided to modify P2/P3 while maintaining the P1' and P1 residues. To enhance rigidity, we chose various bi- or tricycles in P2, such as indole, 2,3-dihydrobenzo[*b*][1,4]dioxine (DHBD), and several others, resulting in compounds **2a**–**q** (**Figure 6**). We prepared corresponding irreversible vinylsulfones (**2b**–**f**) and -sulfonates (**2g**–**q**) as well as one compound with a reversibly reacting *α*-fluorovinylsulfone warhead (**2a**).

R² = 1*H*-7-indolyl (**A**)
2a: X = F, R¹ = Ph
2b: X = H, R¹ = Ph
2g: X = H, R¹ = OPh

R² = DHBD (**B**)
2c: X = H, R¹ = Ph
2h: X = H, R¹ = OPh

R² = 5-isoxazolyl (**K**)
2q: X = H, R¹ = OPh

R² = 1*H*-1-tetrahydrocarbazolyl (**C**)
2d: X = H, R¹ = Ph
2i: X = H, R¹ = OPh

R² = DHBO (**J**)
2p: X = H, R¹ = OPh

R² = 3-benzo[*d*]isoxazolyl (**I**)
2o: X = H, R¹ = OPh

R² = *N*-(2-phenyl)anilinyl (**D**)
2e: X = H, R¹ = Ph
2j: X = H, R¹ = OPh

R² = 6-benzo[*d*]oxazolyl (**H**)
2n: X = H, R¹ = OPh

R² = 4-dibenzo[*b,d*]furanyl (**G**)
2m: X = H, R¹ = OPh

R² = 1*H*-6-indazolyl (**E**)
2f: X = H, R¹ = Ph
2k: X = H, R¹ = OPh

R² = 1*H*-5-benzo[*d*]imidazolyl (**F**)
2l: X = H, R¹ = OPh

Figure 6: Designed (F-)vinylsulfon(at)e-based compounds **2a–q**.

*Sm*CB1 inhibition and structure-activity relationship

Fluorometric enzyme assay

We tested all compounds in a fluorometric enzyme assay to evaluate their inhibitory activity for *Sm*CB1 and their selectivity towards off-target cathepsins (procedures in **Supporting Information**). The results are shown in **Table 2**.

The vinylsulfone-based compounds were moderate *Sm*CB1 inhibitors with K_i values in the low micromolar range except for **2e** (K_i >100 μM). The covalent reaction occured slowly (k_{inact} around 10^{-4} s^{-1}) and there was moderate selectivity toward cathepsins B and L. **2a**, an α-fluorovinylsulfone, was the only inhibitor in the series that reversibly inhibited the enzyme by a covalent-reversible Michael addition.[18,19] The K_i was 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 had improved in all cases compared to the vinylsulfone counterparts, for example **2h** (K_i = 0.050 μM) vs. **2c** (K_i = 6.7 μM). Since vinylsulfonates proved to be more potent *Sm*CB1 inhibitors, we designed and synthesized additional compounds (**2l–q**) and with these obtained K_i values mostly in the nanomolar range. The optimizations for **2l–q** were performed based on similar or even improved physicochemical properties compared to **2a–k** and their feasibility was tested in molecular docking studies.

Overall, **2h** ($K_i = 0.050$ μM, **Figure S6** in the **Supporting Information**), **2n** ($K_i = 0.030$ μM), and **2o** ($K_i = 0.050$ μM) were the top three compounds in terms of inhibitory potency. **2h** also had the highest selectivity in the series (>400-fold toward CatB and CatL) compared to the moderate selectivity of **2n** and **2o**. Notably, the k_{inact} values for most vinylsulfonates were higher than for the vinylsulfones, suggesting a faster covalent bond formation.

Table 2: Inhibition data from fluorometric enzyme assay for (F-)vinylsulfon(at)es **2a–q**.

2a–f: R' = Ph
2g–q: R' = OPh
2a: X = F
2b–q: X = H
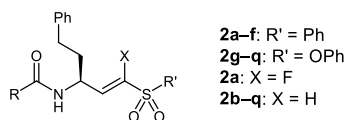
| Cpd. | R | SmCB1 | | | CatB | | CatL | |
		K_i / μM	k_{inact} / s^{-1}	k_{2nd} / M^{-1}s^{-1}	K_i / μM or % inh. at 20 μM	SI	K_i / μM or % inh. at 20 μM	SI
2a	A	5.8 ± 1.5	-	-	30%	> 3	45%	> 3
2b	A	4.7 ± 0.3	$1.1 \cdot 10^{-4}$	23	38%	> 4	n. i.	> 4
2c	B	6.7 ± 0.8	$2.3 \cdot 10^{-4}$	34	29%	> 3	24%	> 3
2d	C	2.2 ± 0.8	$1.6 \cdot 10^{-4}$	7.3	41%	> 9	41%	> 9
2e	D	> 100	n. d.	n. d.	28%	n. d.	30%	n. d.
2f	E	2.5 ± 0.6	$2.5 \cdot 10^{-4}$	10	22%	> 8	32%	> 8
2g	A	0.72 ± 0.005*	$1.8 \cdot 10^{-3}$	$2.5 \cdot 10^{3}$	n. i.	>28	43%	>28
2h	B	0.050 ± 0.003*	$7.0 \cdot 10^{-4}$	$1.4 \cdot 10^{4}$	29%	>400	37%	>400
2i	C	1.6 ± 0.36*	$3.0 \cdot 10^{-4}$	$1.9 \cdot 10^{2}$	31%	>13	28%	>13
2j	D	2.3 ± 1.0*	$5.0 \cdot 10^{-4}$	$2.2 \cdot 10^{2}$	20%	> 9	26%	>9
2k	E	0.83 ± 0.065*	$1.3 \cdot 10^{-3}$	$1.5 \cdot 10^{3}$	6.5 ± 1.0	7.8	0.54 ± 0.05	7.8
2l	F	0.078 ± 0.007*	$1.1 \cdot 10^{-3}$	$1.4 \cdot 10^{4}$	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*	$2.4 \cdot 10^{-3}$	$8.0 \cdot 10^{4}$	1.0 ± 0.16	33	2.2 ± 0.19	73
2o	I	0.050 ± 0.030*	$1.8 \cdot 10^{-3}$	$3.6 \cdot 10^{4}$	48%	>400	0.37 ± 0.03	7.4
2p	J	0.41 ± 0.30*	$5.4 \cdot 10^{-3}$	$1.3 \cdot 10^{4}$	39%	>49	8.2 ± 2.1	20
2q	K	0.15 ± 0.040*	$2.8 \cdot 10^{-3}$	$1.8 \cdot 10^{4}$	0.26 ± 0.01	1.7	2.3 ± 0.23	15

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. *time-dependent inhibition.

SAR discussion

Docking scores and Hyde scores were generated using *LeadIT* and are shown in **Table S6** in the **Supporting Information**.[29] Since we maintained P1' and P1 compared to **SJ605** and **SJ606** (**Figure7B**), we still observed interactions between Gly144 and Gly269 with the amide bond as well as an additional H-bond with Gly143 during non-covalent docking. Gln94 and Trp292 are also involved in H-bond

formation with the sulfonyl group. The docking poses for top compound **2h** overlapping with **SJ606** or **2p** are shown in **Figure 7**. The poses for **2h** and **2p** are very similar (**Figure 7A**) as expected from Hyde scores (**2h**: −39 kJ/mol, **2p**: −37 kJ/mol). Nonetheless, there was a difference in *Sm*CB1 inhibition. **2h** was one of the most potent inhibitors in the series ($K_i = 0.050$ μM), whereas **2p** was only a moderate inhibitor ($K_i = 0.41$ μM). Their k_{2nd} values, on the other hand, were in the same range ($10^4\,\text{M}^{-1}\text{s}^{-1}$) due to the faster irreversible reaction of **2p** compared to **2h** (**Table 2**).

2i surprisingly showed very weak inhibition of *Sm*CB1 (K_i >100 μM), which is not reflected in 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 that undergoes the covalent reaction with the nucleophilic Cys100 was increased from 2.9 Å to 4.3 Å (**Figure S5** in the **Supporting Information**). Covalent reactions occur more readily at distances up to 3.5 Å, which could explain the poor inhibitory activity.[31] The same might apply for **2m** (K_i >100 μM) for which a distance of 3.9 Å was found. The distances for compounds **2a**–**q** are listed in **Table S6** in the **Supporting Information**.
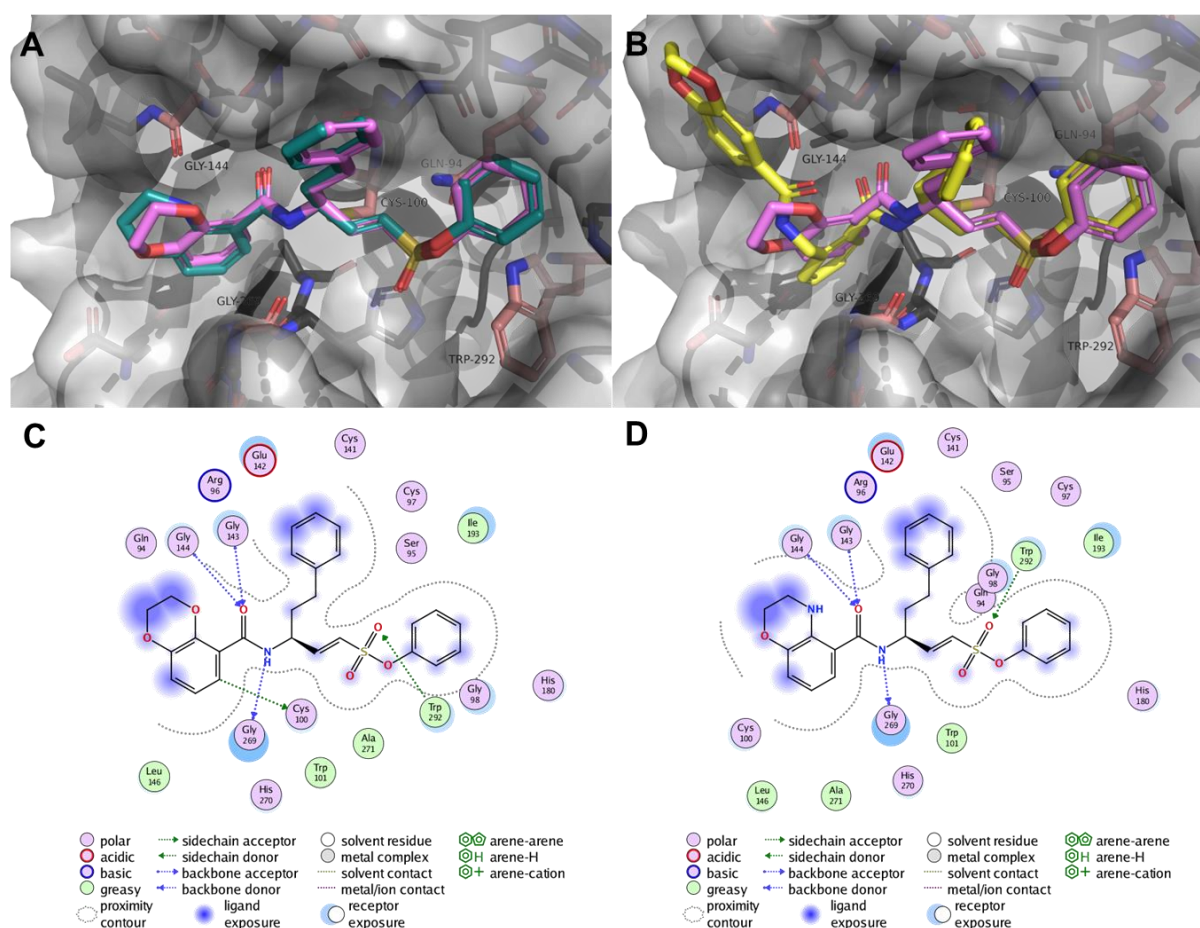
Figure 7: Molecular docking of **2h**, **2p**, and initial lead structure **SJ606**. Non-covalent docking with *LeadIt*, images were generated with *PyMOL*, ligand interaction map generated with *MOE*.[25,30] PDB-ID: 3S3R.[16] (A) Overlay of **2h** (violet) and **2p** (deepteal). Important actives site residues are shown in salmon. Both poses are very similar, which is reflected by their Hyde scores (**2h**: −39 kJ/mol, **2p**: −37 kJ/mol).[32] (B) Overlay of **2h** (violet) and **SJ606** (yellow). Important residues in the binding site are shown in salmon. The binding poses are similar except for additional interactions of the DHBD moiety of **SJ606** with Ile145 and Gly118. (C) Ligand interactions for **2h**. (D) Ligand interactions for **2p**.

In vitro activity in phenotypic assay

Table 3: Physicochemical parameters and *in vitro* effects of **2a–q** compared to praziquantel (**PZQ**), **SJ605**, and **SJ606**.

Cpd.	R	SlogP[a]	TPSA[a] / Å²	Kier-Flex[a,b]	NTS % effect[c] at 10 μM	at 1 μM	*S. mansoni* adults % effect[c] at 10 μM	at 1 μM	at 0.1 μM
SJ605	-	5.81	102	7.96	100 ± 0	33 ± 12	55 ± 4	n. d.	n. d.
SJ606	-	5.31	120	8.22	100 ± 0	30 ± 0	58 ± 5	n. d.	n. d.
2a	A	5.18	79	5.30	60 ± 1	38 ± 2	n. d.	n. d.	n. d.
2b	A	4.89	79	5.14	85 ± 1	40 ± 3	27 ± 2	n. d.	n. d.
2c	B	4.18	82	6.11	92 ± 0	27 ± 1	29 ± 0	n. d.	n. d.
2d	C	5.69	79	5.75	100 ± 0	48 ± 1	31 ± 2	n. d.	n. d.
2e	D	6.15	75	6.36	42 ± 4	27 ± 1	n. d.	n. d.	n. d.
2f	E	4.28	92	5.09	75 ± 2	38 ± 4	47 ± 6	n. d.	n. d.
2g	A	4.82	88	5.59	100 ± 0	27 ± 1	63 ± 6	n. d.	n. d.
2h	B	4.11	91	6.59	88 ± 0	33 ± 0	86 ± 2	68 ± 0	35 ± 2
2i	C	5.63	88	6.19	100 ± 0	42 ± 2	86 ± 2	57 ± 0	n. d.
2j	D	6.08	85	6.84	100 ± 0	35 ± 1	31 ± 2	n. d.	n. d.
2k	E	4.22	101	5.54	100 ± 0	33 ± 2	86 ± 2	36 ± 0	n. d.
2l	F	4.22	101	5.54	100 ± 0	33 ± 2	63 ± 2	n. d.	n. d.
2m	G	6.04	82	6.24	100 ± 0	44 ± 4	71 ± 4	29 ± 4	n. d.
2n	H	4.48	99	5.74	100 ± 0	48 ± 8	42 ± 4	n. d.	n. d.
2o	I	4.48	99	5.74	100 ± 0	40 ± 4	38 ± 8	n. d.	n. d.
2p	J	4.14	94	6.37	100 ± 0	58 ± 10	75 ± 0	23 ± 2	n. d.
2q	K	3.33	99	5.83	100 ± 0	44 ± 4	40 ± 2	n. d.	n. d.
PZQ[d]	-	2.63	41	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;[25] b: KIER molecular flexibility index;[28] c: % dead after 72 hours; d: MEISTER *et al.*[33]

Compared to starting compounds **SJ605** and **SJ606**, several vinysulfonate-based inhibitors were more efficient against NTS and *S. mansoni* adults. Especially compounds **2h**, **2i**, **2k**, and **2p** were outstanding as they displayed high efficacy (≥75%) against *S. mansoni* adults at 10 μM. **2h** had 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).[33] Notably, **2h** and **2p** (*S. mansoni* adults: 75% at 10 μM, 23% 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 suggested that lower K_i values resulted in higher efficacy against NTS at 10 μM (r = –0.47, p = 0.048, **Table S5** in **Supporting Information**) as well as a slight correlation between higher TPSA (ranging from 75 Å² up to 110 Å²) and higher efficacy (r = 0.55, p = 0.023, **Table S5** in **Supporting Information**).

Although the initial correlation analysis suggested that a higher lipophilicity could be beneficial, this was not the case for all optimized compounds. An example was **2e** with the highest SlogP in the series (6.15) but only limited potency (NTS: 42% at 10 μM, *S. mansoni* adults: n. d.). In contrast, **2k** was one of the top compounds (NTS: 100% at 10 μM, *S. mansoni* adults: 86% at 10 μM) with one of the lowest SlogP values (4.22). Flexibility also seemed to have a smaller impact on efficacy than assumed. The top compounds' KierFlex values ranged 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.

In summary, several factors combine to influence phenotypic efficacy in schistosomes. Regardless, compound **2h** displayed high potency and should be considered for further studies.

CONCLUSION

Using an in-house screening with 76 compounds, we have identified two lead structures, **KS309** and the **SJ600**-scaffold (**SJ605**, **SJ606**), which showed promising results against *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 only moderately 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 logP values or, in case of the **SJ600s**, more 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 derivatives with altered physicochemical properties yielding three optimized compounds: **1g** (*Sm*CB1 inhibition: 68% at 20 μM; *S. mansoni* adults: 100% at 10 μM, 43% at 1 μM), **1i** (*Sm*CB1 inhibition: 9% at 20 μM; *S. mansoni* adults: 93% at 10 μM, 36% at 1 μM), and **1j** (*Sm*CB1 inhibition: 80% at 20 μM; *S. mansoni* adults: 100% at 10 μM, 45% at 1 μM). However, their main target remains unknown. The "*Sm*CB1 approach" using **SJ605** and **SJ606** as leads resulted in 17 derivatives with higher rigidity and lower TPSA including seven optimized compounds in terms of 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 than vinylsulfones. The most potent compound, **2h**, showed nanomolar inhibition of *Sm*CB1 while maintaining selectivity towards human off-target cathepsins (K_i = 50 nM, SI >400). The efficacy of **2h** toward *S. mansoni* adults 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 praziquantel (*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 available drug PZQ, which is facing rising resistance.[9] For further validation, metabolism studies and plasma stability tests should be considered to prevent problems that may arise in

future *in vivo* studies. Finally, preclinical *in vivo* studies could give valuable insights into biodistribution, followed up by *in vivo* efficacy, e.g. with infected mice.

ABBREVIATIONS

AMC, 7-amino-4-methylcoumarin; aq., aqueous; Boc, tert-butyl carbamoyl; CatB, human cathepsin B; CatL, human cathepsin L; CH, cyclohexane; cpd, compound; DCM, dichloromethane; DECP, diethyl chlorophosphate; DHBD, 2,3-dihydrobenzo[*b*][1,4]dioxine; DIPEA, *N,N*-diisopropylethylamine; DMF, dimethylformamide; DTT, dithiothreitol; EA, ethyl acetate; EDTA, ethylenediaminetetraacetic acid; eq., equivalent; ESI, electrospray ionization; HIC, hydrophobic interaction chromatography; HOBt, 1-hydroxybenzotriazole; hPhe, homophenylalanine; HWE, Horner-Wadsworth-Emmons; KHMDS, potassium bis(trimethylsilyl)amide; LHMDS, lithium bis(trimethylsilyl)amide; Me, methyl; MOE, Molecular Operating Environment; NTS, newly transformed schistosomula; 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; TEA, triethyl amine; THF, tetrahydrofuran; TPSA, topological polar surface area.

ACKNOWLEDGEMENTS

REFERENCES

1   James, S. L.; Abate, D.; Abate, K. H.; Abay, S. M.; Abbafati, C.; Abbasi, N.; Abbastabar, H.; Vos, T.; Murray, C. J. L.; Al., E. 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. https://doi.org/10.1016/S0140-6736(18)32279-7.

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 Negl. Trop. Dis.* **2014**, *8* (7). https://doi.org/10.1371/JOURNAL.PNTD.0002865.

3   Colley, D. G.; Bustinduy, A. L.; Secor, W. E.; King, C. H. Human Schistosomiasis. *Lancet (London, England)* **2014**, *383* (9936), 2253–2264. https://doi.org/10.1016/S0140-6736(13)61949-2.

4   Gryseels, B.; Polman, K.; Clerinx, J.; Kestens, L. Human Schistosomiasis. *Lancet (London, England)* **2006**, *368* (9541), 1106–1118. https://doi.org/10.1016/S0140-6736(06)69440-3.

5   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. https://doi.org/10.1016/J.MOLBIOPARA.2005.01.007.

6   MOORE, D. V.; SANDGROUND, J. H. The Relative Egg Producing Capacity of Schistosoma Mansoni and Schistosoma Japonicum. *Am. J. Trop. Med. Hyg.* **1956**, *5* (5), 831–840. https://doi.org/10.4269/AJTMH.1956.5.831.

7   Hams, E.; Aviello, G.; Fallon, P. G. The Schistosoma Granuloma: Friend or Foe? *Front. Immunol.* **2013**, *4* (APR), 89. https://doi.org/10.3389/FIMMU.2013.00089/BIBTEX.

8   Burke, 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. https://doi.org/10.1111/J.1365-3024.2009.01098.X.

9   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/TABLES/1.

10   de Souza Dias, L. C.; Olivier, C. E. Failure at Inducing Resistance to Schistosomicidal Drugs in a Brazilian Human Strain of Schistosoma Mansoni. *Rev. do Inst. Med. Trop. São Paulo* **1986**, *28* (5), 352–357.

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). https://doi.org/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. https://doi.org/10.1021/ACSINFECDIS.0C00644/ASSET/IMAGES/LARGE/ID0C00644_0006.JPEG.

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. https://doi.org/10.1021/ACSINFECDIS.0C00501/ASSET/IMAGES/LARGE/ID0C00501_0004.JPEG.

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), 0130–0138. https://doi.org/10.1371/journal.pmed.0040014.

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. https://doi.org/10.1016/J.MOLBIOPARA.2005.06.007.

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. https://doi.org/10.1074/JBC.M111.271304.

17   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 142* **2019**, *14* (2), 461–481. https://doi.org/10.1038/s41596-018-0101-y.

18   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.

19   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. 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* (16), 12322–12358. https://doi.org/10.1021/ACS.JMEDCHEM.1C01002/SUPPL_FILE/JM1C01002_SI_002.CSV.

20   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*. https://doi.org/10.1016/J.EJMECH.2022.114460.

21   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. *Bioorganic Med. Chem. Lett.* **2007**, *17* (17), 4929–4933. https://doi.org/10.1016/j.bmcl.2007.06.023.

22   Barthels, F.; Marincola, G.; Marciniak, T.; Konhäuser, M.; Hammerschmidt, S.; Bierlmeier, J.; Distler, U.; Wich, P. R.; Tenzer, S.; Schwarzer, D.; Ziebuhr, W.; Schirmeister, T. Asymmetric Disulfanylbenzamides as Irreversible and Selective Inhibitors of Staphylococcus Aureus Sortase A. *ChemMedChem* **2020**, *15* (10), 839–850. https://doi.org/10.1002/CMDC.201900687.

23   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*, 1284. https://doi.org/10.3389/FMOLB.2021.804970/BIBTEX.

24   Pearce, E. J.; MacDonald, A. S. The Immunobiology of Schistosomiasis. *Nat. Rev. Immunol. 2002 27* **2002**, *2* (7), 499–511. https://doi.org/10.1038/nri843.

25   Molecular Operating Environment (MOE). Chemical Computing Group ULC: 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7 2020.

26   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), e1003856. https://doi.org/10.1371/JOURNAL.PCBI.1003856.

27   Furlong, S. T. Unique Roles for Lipids in Schistosoma Mansoni. *Parasitol. Today* **1991**, *7* (2), 59–62. https://doi.org/10.1016/0169-4758(91)90192-Q.

28   Kier, L. B. An Index of Molecular Flexibility from Kappa Shape Attributes. *Quant. Struct. Relationships* **1989**, *8* (3), 221–224. https://doi.org/10.1002/QSAR.19890080307.

29   LeadIT. BioSolveIT GmbH: Sankt Augustin, Germany 2017. www.biosolveit.de/LeadIT.

30   Schrödinger, L.; DeLano, W. The PyMOL Molecular Graphics System. Schrödinger, LLC. 2020. http://www.pymol.org/pymol.

31   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/ASSET/IMAGES/LARGE/JM-2018-011539_0038.JPEG.

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. https://doi.org/10.1002/cmdc.200700319.

33   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. https://doi.org/10.1128/AAC.02741-14/ASSET/5080265F-AEA4-4E8C-8A03-FD570B397423/ASSETS/GRAPHIC/ZAC0091432400003.JPEG.

6.4   Syntheses

Note: Syntheses conducted by Marvin Schwickert (⬛⬛⬛⬛⬛⬛ group) are described in this section. Synthetic procedures performed by others can be found in the complete Supporting Information file in the **Appendix**. Compound numbers in this section refer to the numbering in the corresponding manuscript and Supporting Information.

The synthesis of compounds **1b–d**, and **1g** started from Boc-D-leucine (**4**) (**Scheme 29**). First, the Weinreb amide **5** was prepared by TBTU coupling. Treatment with *n*-butyllithium and dimethyl methylphosphonate at −78 °C yielded the phosphonate **6** as an enantiomeric mixture (ratio not determined). To generate the *N*-methylated oxoenoate **1b** (P3-*N*-Me-P2), **6** was deprotected by treatment with 4 M HCl, and TBTU-coupling with Boc-*N*-methyl-L-phenylalanine was performed yielding **7** as a diastereomeric mixture. After Boc cleavage and TBTU-coupling with pyrazinoic acid, followed by diastereomer separation, the phosphonate **8** was obtained. In the final step, a Horner-Wadsworth-Emmons (HWE) reaction using ethyl glyoxylate, lithium chloride, and DIPEA was conducted to form the *N*-methylated oxoenoate **1b** (P3-*N*-Me-P2). To generate the P3 congeners **1c** (P3-Cbz) and **1d** (P3-Boc) of **KS309**, the phosphonate **6** was deprotected with 4 M HCl in dioxane in each case, followed by TBTU coupling with either Boc-L-phenylalanine or Cbz-L-phenylalanine. The resulting phosphonates **9** and **10** were finally treated with ethyl glyoxylate, lithium chloride and DIPEA in a HWE reaction to give **1c** and **1d**. For the synthesis of the vinylsulfonate **1g**, the Weinreb amide **5** was brought to reaction with lithium aluminium hydride at 0 °C to generate the aldehyde **11**. Then, a HWE reaction was conducted using the phosphonate **12** (synthesis was previously described),[359] lithium chloride and DBU yielding the intermediate **1i**. After Boc cleavage and TBTU coupling with Boc-L-phenylalanine, **1j** was obtained. In the final step, Boc cleavage and TBTU coupling with pyrazinoic acid was conducted to give the final vinylsulfonate **1g**.
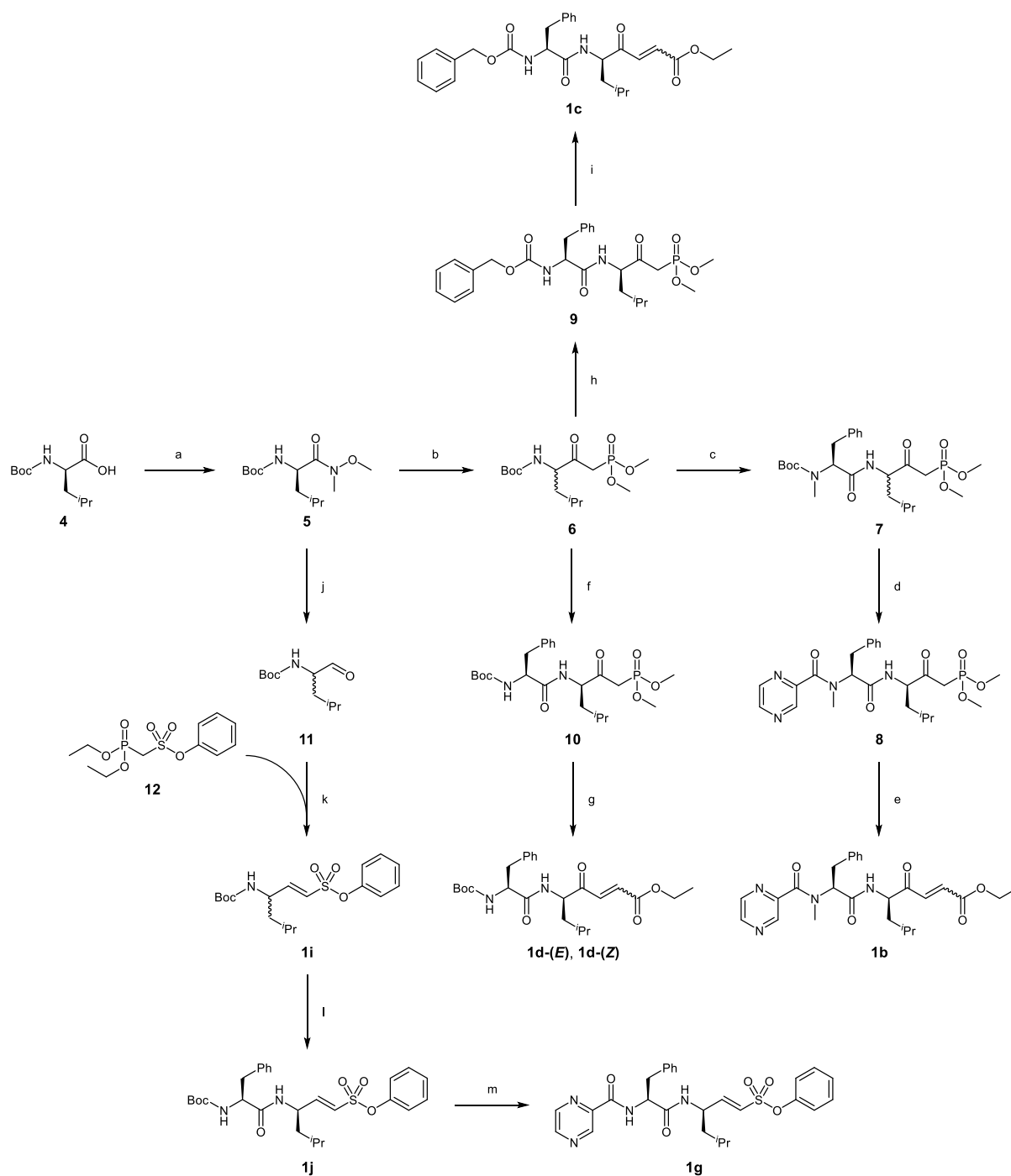
Scheme 29: Synthesis of compounds **1b–d**, and **1g**. a: *N,O*-dimethylhydroxylamine • HCl, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 94%; b: dimethyl methyl phosphonate, *n*-BuLi, THF (dry), −78 °C, 1.5 h, 93%; c: 1.) 4 M HCl in dioxane, rt, 1.5 h, quantitative; 2.) Boc-*N*-Me-ʟ-Phe-OH, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 63%; d: 1.) 4 M HCl in dioxane, rt, 1.5 h, quantitative; 2.) pyrazinoic acid, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, DMF, 0 °C to rt, overnight, 41%; e: ethyl glyoxylate, DIPEA, LiCl, MeCN (dry), 0 °C, 75 min, 61%; f: 1.) 4 M HCl in dioxane, rt, 1.5 h, quantitative; 2.) Boc-ʟ-Phe-OH, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 50%; g: ethyl glyoxylate, DIPEA, LiCl, MeCN (dry), 0 °C, 75 min, 53%; h: 1.) 4 M HCl in dioxane, rt, 1.5 h, quantitative; 2.) Cbz-ʟ-Phe-OH, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 61%; i: ethyl glyoxylate, DIPEA, LiCl, MeCN (dry), 0 °C, 75 min, 45%; j: LiAlH₄, THF (dry), 0 °C, 30 min, 98%; k: DBU, LiCl, MeCN (dry), 0 °C, 75 min, 74%; l: 1.) 4 M HCl in dioxane, rt, 1.5 h, quantitative; 2.) Boc-ʟ-Phe-OH, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 89%; m: 1.) 4 M HCl in dioxane, rt, 1 h, quantitative; 2.) pyrazinoic acid, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 60%.

The other *N*-methylated oxoenoate **1a** (P2-*N*-Me-P1) was prepared analogously to **1b** in **Scheme 29**. Instead of Boc-D-leucine, Boc-*N*-methyl-D-leucine (**13**) was used in the first step, and Boc-L-phenylalanine was used instead of Boc-*N*-methyl-L-phenylalanine in the third step (**Scheme 30**).
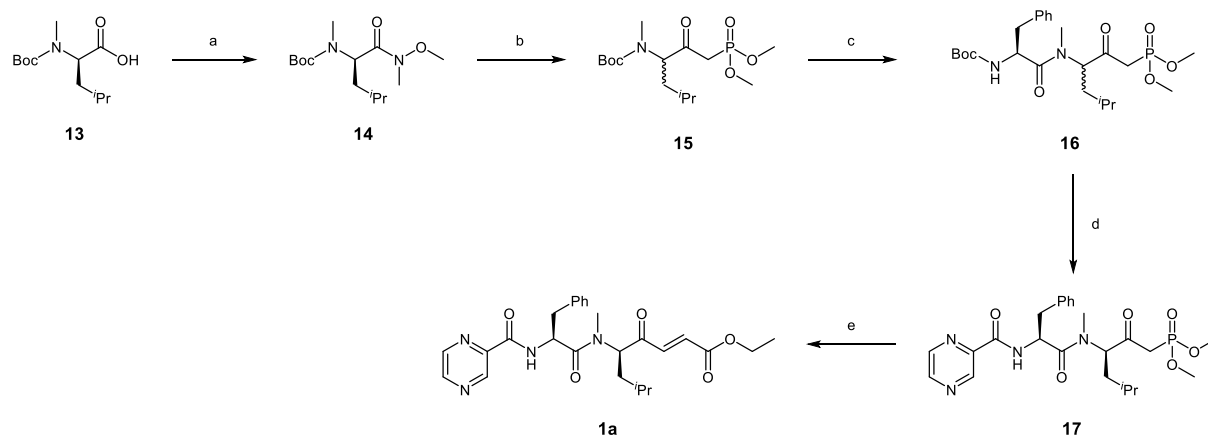

Scheme 30: Synthesis of compound **1a**. a: *N,O*-dimethylhydroxylamine • HCl, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 86%; b: dimethyl methyl phosphonate, *n*-BuLi, THF (dry), −78 °C, 1.5 h, 96%; c: 1.) 4 M HCl in dioxane, rt, 1 h, quantitative; 2.) Boc-L-Phe-OH, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 41%; d: 1.) 4 M HCl in dioxane, rt, 1.5 h, quantitative; 2.) pyrazinoic acid, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, DMF, 0 °C to rt, overnight, 32%; e: ethyl glyoxylate, DIPEA, LiCl, MeCN (dry), 0 °C, 75 min, 44%.

The azapeptide **1k** was prepared starting from Cbz-L-phenylalanine (**18**), which was brought to reaction with *tert*-butyl carbazate using TBTU coupling (**Scheme 31**). Hydrogenolytic cleavage of the Cbz group of **19** led to the free amine **20**, which was used in a TBTU coupling with pyrazinoic acid to give **21**. Then, the Boc group was cleaved using 4 M HCl in dioxane. The resulting hydrazide **22** was alkylated by a two-step reductive amination with isobutyraldehyde and sodium borohydride to yield **23**. In the last step, reaction with ethyl fumaroyl chloride generated the final azapeptide **1k**.
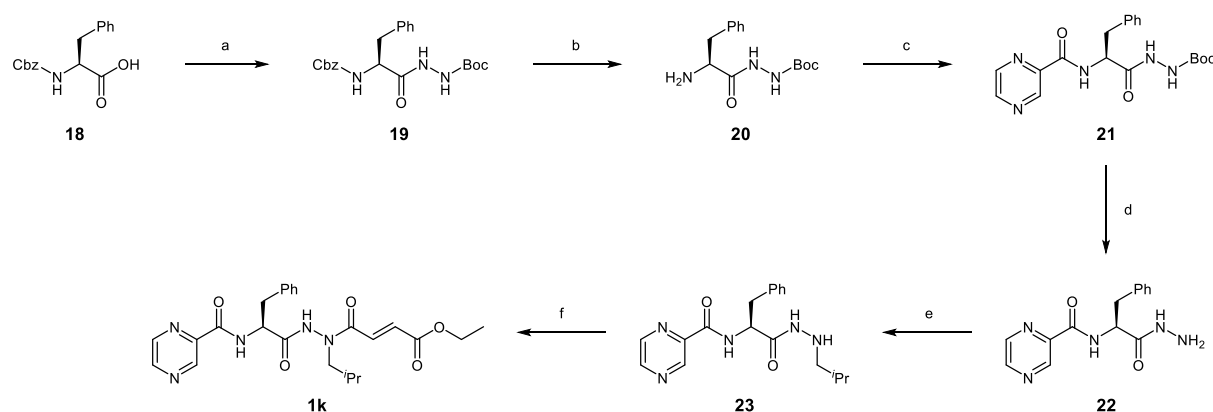
Scheme 31: Synthesis of compound **120**. a: *tert*-Butyl carbazate, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 88%; b: H₂, Pd/C, MeOH, rt, 21 h, 97%; c: pyrazinoic acid, TBTU, HOBt • H₂O, 2,4,6-collidine, DCM, DMF, 0 °C to rt, overnight, 89%; d: 4 M HCl in dioxane, rt, 1 h, 95%; e: isobutyraldehyde, NaBH₄, THF, rt, 7.5 h, 36%; f: ethyl fumaroyl chloride, NEt₃, DCM, rt, 22 h, 55%.

Propargyl and nitrile compounds **1e** and **1f** were prepared by the same synthetic pathway (**Scheme 32**). First, L-leucine methyl ester hydrochloride (**24**) was converted with Boc-L-phenylalanine using TBTU coupling, which yielded the dipeptide **25**. After Boc cleavage and TBTU coupling with pyrazinoc acid, **26** was obtained, which was hydrolyzed using basic conditions. The resulting free acid **27** was reacted with either aminoacetonitrile hydrochloride or propargylamine using TBTU coupling to yield the final compounds **1e** and **1f**.
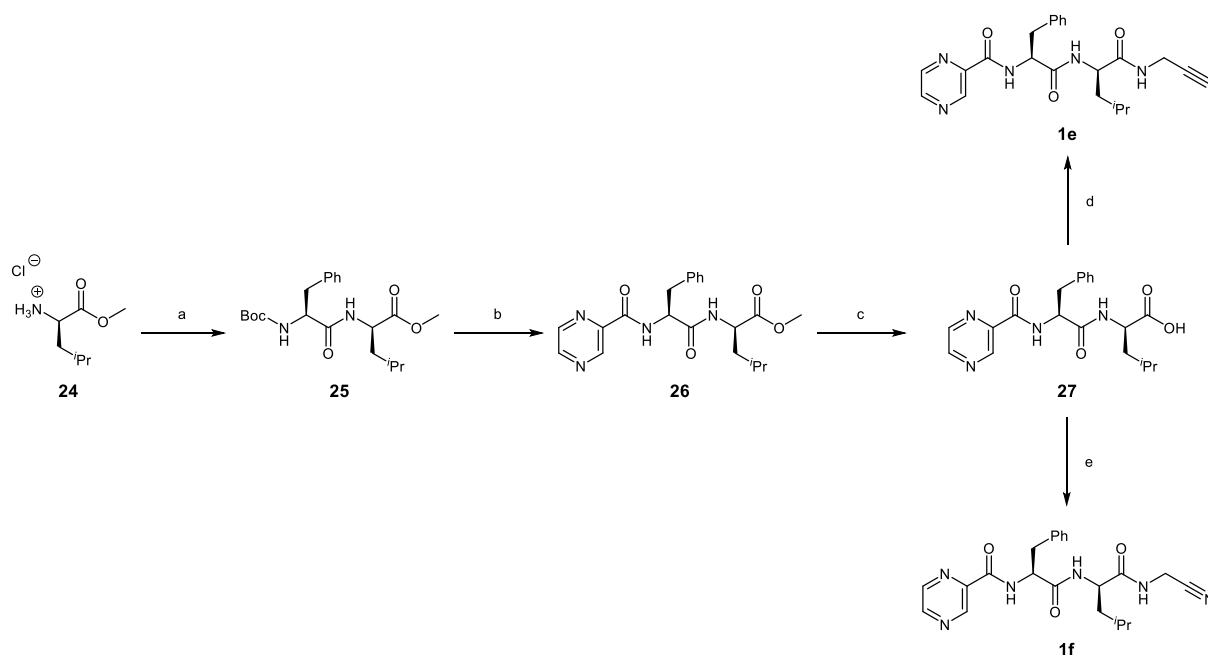
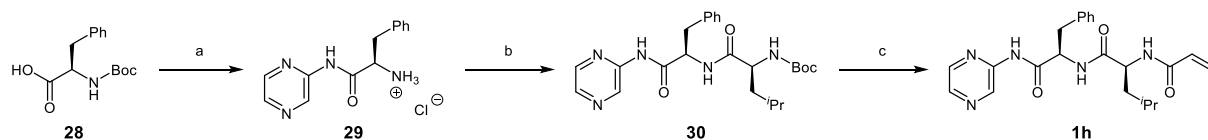**Scheme 32:** Synthesis of compounds **1e** and **1f**. a: Boc-L-Phe-OH, TBTU, HOBt • H$_2$O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 94%; b: 1.) 4 M HCl in dioxane, rt, 1 h, quantitative; 2.) pyrazinoic acid, TBTU, HOBt • H$_2$O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 88%; c: LiOH • H$_2$O, THF, H$_2$O, rt, 6 h, quantitative; d: aminoacetonitrile hydrochloride, TBTU, HOBt • H$_2$O, 2,4,6-collidine, DCM, 0 °C to rt, 26 h, 92%; e: propargylamine, TBTU, HOBt • H$_2$O, 2,4,6-collidine, DCM, 0 °C, 1 h, 53%.

The synthesis of the acrylamide decorated compound **1h** with inverted amide bonds was conducted starting from Boc-D-phenylalanine (**28**) (**Scheme 33**). First, coupling with 2-aminopyrazine using COMU, followed by deprotection with 4 M HCl in dioxane to yield the hydrochloride salt **29**. In the next step, **29** was reacted with Boc-L-leucine by TBTU coupling to give **30**. Finally, the Boc group of **30** was removed by treatment with 4 M HCl in dioxane, and the acrylamide function was introduced by reaction with acryloyl chloride yielding the final product **1h**.

Scheme 33: Synthesis of compound **117**. a: 1.) 2-aminopyrazine, COMU, oxyma, 2,4,6-collidine, 0 °C to rt, overnight; 2.) 4 M HCl in dioxane, rt, 1 h, 24%; b: Boc-L-Leu-OH, TBTU, HOBt • H$_2$O, 2,4,6-collidine, DCM, 0 °C to rt, overnight, 65%; c: 1.) 4 M HCl in dioxane, rt, 1 h; 2.) acryloyl chloride, NEt$_3$, DCM, 0 °C, 3 h, 51%.

The small molecule derivatives **1l** and **1m** were each prepared in a two-step synthesis (**Scheme 34**). Reaction of Weinreb amide **5** with vinylmagnesium bromide at −78 °C yielded the acryl derivative **31**, which was treated with ethyl acrylate using Hoveyda Grubbs II catalyst to give the final compound **1l**. The azapeptide analog **1m** was prepared starting from *tert*-butyl carbazate (**32**). In the first step, **32** was alkylated by a two-step reductive amination using isobutyraldehyde followed by treatment with Pd/C under hydrogen atmosphere. The resulting intermediate **33** was then reacted with ethyl fumaroyl chloride to yield the azapeptide analog **1m**.

A

B
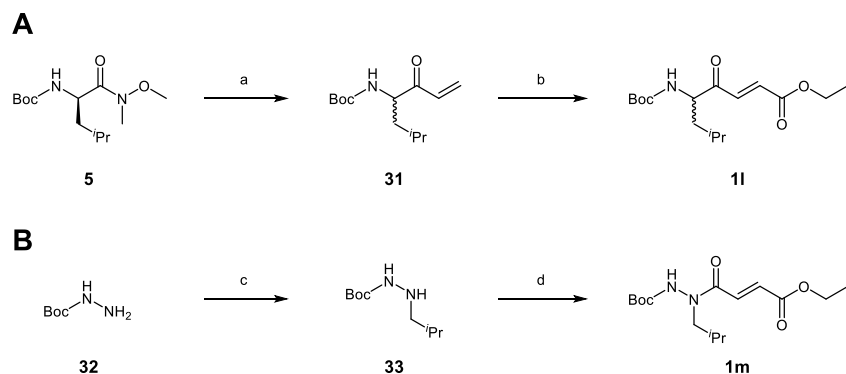
Scheme 34: A: synthesis of compound **1l**. B: synthesis of compound **1m**. a: vinylmagnesium bromide, THF (dry), −78 °C, 30 min, 50%; b: ethyl acrylate, Hoveyda Grubbs II catalyst, DCM (dry), Δ, 4 h, 37%; c: 1.) isobutyraldehyde, toluene, 50 °C, 1 h; 2.) H$_2$ (3 bar), Pd/C, rt, 20 h, 37%; d: ethyl fumaroyl chloride, NEt$_3$, DCM, rt, 1 h, 68%.

6.5 Experimental section of the manuscript

Note: Experimental procedures conducted or supervised by MARVIN SCHWICKERT (███████ group) are described in this section. Experiments performed by others can be found in the complete Supporting Information file in the **Appendix**. Compound numbers in this section refer to the numbering in the corresponding manuscript.

General information

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, and anhydrous solvents were used as provided by the commercial supplier. Reaction progress was monitored by thin-layer chromatography using ALUGRAM® Xtra SIL G UV$_{254}$ silica plates from *Macherey-Nagel* and/or LC-MS. For LC-MS analysis, an *Agilent 1100 series* HPLC system and an *Agilent Poroshell 120 EC-C$_{18}$*, 150 × 2.10 mm, 4 μm column coupled to an *Agilent 1100 series* LC/MSD Trap with electron spray ionization (ESI) was used. The identities and purities of compounds were determined by the same LC-MS system with a gradient of MeCN and MilliQ®-H$_2$O + 0.1% HCOOH (flow rate: 0.7 mL/min). 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 a *Varian PrepStar* system using an *Agilent Zorbax PrepHT XDB C$_{18}$* (150 mm × 21.2 mm, 5 μm) column. Flash chromatography was performed with a *Biotage Isolera^{TM} One* system using prepacked columns from *Biotage*. Silica (40–63 μm) from *Macherey-Nagel* was used for column chromatography. Optical rotations $[\alpha]_D^{20}$ were measured on a *P3000* polarimeter from *Krüss* and are reported in deg cm^3 g^{-1} dm^{-1}. 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 on a *Bruker Fourier 300*. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance (CDCl$_3$, DMSO-d_6, or CD$_3$OD from *Deutero GmbH*) as internal standard. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

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 were added HOBt • H$_2$O (1.0 equiv.) and 2,4,6-collidine (2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and the amine hydrochloride (1.0 or 1.1 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C. If DMF was used, it was removed by co-distillation with *n*-heptane (3 × 150 mL) under reduced

pressure at 40 °C. The residue was taken up in ethyl acetate and washed with saturated $NaHCO_3$ solution (3 ×) and 1 M HCl (3 ×). The combined organic extracts were dried over Na_2SO_4, and the solvent was removed by distillation under reduced pressure at 40 °C 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 MeCN (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 MeCN (10 mL) was slowly added and the mixture was stirred for 1 h at 0 °C. The reaction was quenched by addition of citric acid solution (10% in water, 10 mL) and extracted with ethyl acetate (4 × 30 mL). The combined organic extracts were dried over Na_2SO_4, and the solvent was removed by distillation under reduced pressure at 40 °C. The resulting crude product was purified by column chromatography on silica.

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 MeCN (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 MeCN (10 mL) was slowly added and the mixture was stirred for 1 h at 0 °C. The reaction was quenched by addition of citric acid solution (10% in water, 10 mL) and extracted with ethyl acetate (4 × 30 mL). The combined organic extracts were dried over Na_2SO_4, and the solvent was removed by distillation under reduced pressure at 40 °C. The resulting crude product was purified by column chromatography on silica.

***tert*-Butyl (*R*)-(1-(methoxy(methyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (5)**
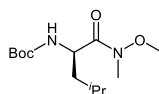
The desired compound was prepared from Boc-D-Leu-OH (1.00 g, 4.32 mmol, 1.0 equiv.), HOBt • H_2O (0.66 g, 4.32 mmol, 1.0 equiv.), 2,4,6-collidine (1.15 mL, 8.65 mmol, 2.0 equiv.), TBTU (1.39 g, 4.32 mmol, 1.0 equiv.), and *N,O*-dimethylhydroxylamine hydrochloride (0.42 g, 4.32 mmol, 1.0 equiv.) in DCM (30 mL) according to general procedure A to afford the desired product as a colorless oil (1.12 g, 4.07 mmol, 94%). ¹H NMR (300 MHz, CDCl₃): δ / ppm = 5.06 (d, *J* = 9.6 Hz, 1H), 4.80–4.61 (m, 1H), 3.77 (s, 3H), 3.18 (s, 3H), 1.79–1.62 (m, 1H), 1.49–1.33 (m, 11H), 0.99–0.88 (m, 6H). ¹³C NMR (75.5 MHz, CDCl₃): δ / ppm = 174.0, 155.8, 79.6, 61.7, 49.1, 42.2, 32.3, 28.5, 24.8, 23.5, 21.7. FT-IR: ν / cm⁻¹ = 3325, 2958, 2937, 2870, 1709, 1659, 1501, 1389, 1366, 1250, 1165, 1045, 1016, 989, 876. $[\alpha]_D^{20}$ = +23 (10 mg/mL; MeOH). R_f: 0.47 (cyclohexane/ethyl acetate = 2:1).

tert-Butyl (1-(dimethoxyphosphoryl)-5-methyl-2-oxohexan-3-yl)carbamate (6)
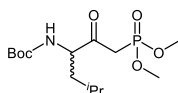
Using a modified procedure by Ho *et al., Eur. J. Org. Chem*. **2005**, *2005*, 4829–4834.

A solution of *n*-BuLi (2.5 M in hexane, 16.20 mL, 40.51 mmol, 5.3 equiv.) was slowly added to a solution of dimethyl methylphosphonate (4.91 mL, 45.86 mmol, 6.0 equiv.) in dry THF (25 mL) at −78 °C under argon atmosphere. Stirring was continued for 1 h at −78 °C, and then a solution of **5** (2.10 g, 7.64 mmol, 1.0 equiv.) in dry THF (8 mL) was added dropwise, and the mixture was stirred for an additional 30 min at −78 °C. The reaction mixture was quenched by addition of saturated NH$_4$Cl solution (80 mL) and extracted with ethyl acetate (3 × 70 mL). The combined organic phases were washed with water (4 × 100 mL), dried over Na$_2$SO$_4$, and the solvent was removed by distillation under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (cyclohexane/ethyl acetate = 1:3) to yield the title compound as a slightly yellowish oil (2.41 g, 7.14 mmol, 93%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 5.21 (d, J = 8.6 Hz, 1H), 4.39–4.25 (m, 1H), 3.80–3.78 (m, 3H), 3.77–3.74 (m, 3H), 3.41–3.24 (m, 1H), 3.17–3.01 (m, 1H), 1.74–1.58 (m, 2H), 1.45–1.38 (m, 10H), 0.96–0.90 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 202.5 (d, J = 6.2 Hz), 155.7, 80.1, 59.0, 53.2 (d, J = 6.6 Hz), 40.0, 38.1 (d, J = 130.2 Hz), 28.4, 24.9, 23.4, 21.6. FT-IR: ν / cm^{-1} = 3273, 2957, 2904, 2871, 1705, 1540, 1366, 1300, 1237, 1163, 1036, 870, 853, 824, 804. $[\alpha]_D^{20}$ = +44 (10 mg/mL; MeOH). R$_f$: 0.38 (cyclohexane/ethyl acetate = 1:3).

Dimethyl (5-methyl-3-((*S*)-2-(*N*-methylpyrazine-2-carboxamido)-3-phenylpropanamido)-2-oxo-hexyl)phosphonate (7)
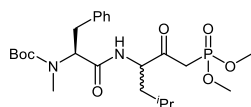
1.) Compound **6** (1.95 g, 5.78 mmol) was treated with 4 M HCl in dioxane (10 mL) at room temperature for 1.5 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless foam (1.58 g, 5.78 mmol, quantitative), which was directly used in the next step.
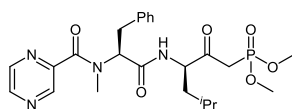
2.) The desired compound was prepared from Boc-*N*-Me-L-Phe-OH (1.61 g, 5.77 mmol, 1.0 equiv.), HOBt · H$_2$O (0.88 g, 5.77 mmol, 1.0 equiv.), 2,4,6-collidine (1.53 mL, 11.55 mmol, 2.0 equiv.), TBTU (1.85 g, 5.77 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (1.58 g, 5.77 mmol, 1.0 equiv.) in DCM (50 mL) according to general procedure A, followed by purification by column chromatography on silica (DCM/MeOH = 50:1) to afford the desired product as a colorless oil, which was present as a mixture of two diastereomers ($\sum$ = 1.80 g, 3.61 mmol, 63%, ratio not determined).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.33–7.11 (m, 5H), 6.89–6.58 (m, 1H), 4.92–4.75 (m, 1H), 4.73–4.50 (m, 1H), 3.84–3.69 (m, 6H), 3.45–2.86 (m, 4H), 2.85–2.70 (m, 3H), 1.77–1.48 (m, 2H), 1.44–1.25 (m, 10H), 0.97–0.80 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 201.6–201.2 (m, diastereomer or rotamer A and B), 171.1 (diastereomer or rotamer A), 170.6 (diastereomer or rotamer B), 156.7 (diastereomer or rotamer A), 155.2 (diastereomer or rotamer B), 138.0 (diastereomer or rotamer B), 137.4 (diastereomer or rotamer A), 129.1, 128.5, 126.8 (diastereomer or rotamer B), 126.7 (diastereomer or rotamer A), 80.7, 60.0, 57.7 (d, J = 6.1 Hz), 53.3, 53.2, 39.5, 38.2 (d, J = 131.2 Hz), 34.2, 31.2, 28.3, 24.8, 23.3, 21.5. FT-IR: ν / cm^{-1} = 3282, 2957, 2932, 2870, 1674, 1454, 1389, 1366, 1323, 1248, 1144, 1029, 815, 748, 699. $[\alpha]_D^{20}$ = −39 (10 mg/mL; CHCl$_3$). R$_f$: 0.32 (cyclohexane/ethyl acetate = 1:2).

Dimethyl ((*R*)-5-methyl-3-((*S*)-2-(*N*-methylpyrazine-2-carboxamido)-3-phenylpropanamido)-2-oxohexyl)phosphonate (8)

1.) Compound **7** (1.80 g, 3.61 mmol) was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1.5 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless foam (1.57 g, 3.61 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from pyrazinoic acid (0.41 g, 3.28 mmol, 1.0 equiv.), HOBt • H$_2$O (0.50 g, 3.28 mmol, 1.0 equiv.), 2,4,6-collidine (0.87 mL, 6.56 mmol, 2.0 equiv.), TBTU (1.05 g, 3.28 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (1.57 g, 3.61 mmol, 1.1 equiv.) in DCM (10 mL) and DMF (10 mL) according to general procedure A, followed by purification by column chromatography on silica (DCM/MeOH = 40:1) to afford the desired product as a yellowish oil, which was present as a mixture of two diastereomers ($\sum$ = 0.68 g, 1.35 mmol, 41%, ratio not determined). Separation of the diastereomers was achieved by HPLC (C$_{18}$, MeCN/H$_2$O = 25:75).

(*S*,*R*)-Diasteromer: ^{1}H NMR (300 MHz, DMSO-d_6, 60 °C): δ / ppm = 8.76–8.67 (m, 1H), 8.66–8.55 (m, 1H), 8.34–8.13 (m, 2H), 7.38–7.15 (m, 4H), 7.06 (s, 1H), 5.45–4.72 (m, 1H), 4.51–4.35 (m, 1H), 3.69 (s, 3H), 3.66 (s, 3H), 3.39–3.33 (m, 1H), 3.32–3.26 (m, 1H), 3.22–3.13 (m, 2H), 3.05–3.00 (m, 2H), 2.98–2.93 (m, 1H), 1.64–1.40 (m, 3H), 0.95–0.80 (m, 6H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ / ppm = 201.7 (d, J = 6.4 Hz) (rotamer B), 201.8 (d, J = 6.4 Hz) (rotamer A), 169.6 (rotamer B), 169.4 (rotamer A), 166.8 (rotamer A), 166.7 (rotamer B), 149.5 (rotamer B), 148.9 (rotamer A), 145.6 (rotamer A), 145.4 (rotamer B), 144.6 (rotamer A), 144.2 (rotamer B), 143.2 (rotamer B), 142.7 (rotamer A), 137.4 (rotamer B), 137.2 (rotamer A), 128.9 (rotamer B), 128.8 (rotamer A), 128.3, 126.6 (rotamer A), 126.4 (rotamer B), 62.1 (rotamer A), 58.1 (rotamer B), 57.3 (d, J = 3.6 Hz), 52.6, 52.5, 37.6 (rotamer A), 37.5

(rotamer B), 36.1 (d, J = 130.5 Hz), 34.6, 34.1, 33.9 (rotamer B), 29.7 (rotamer A), 24.1 (rotamer A), 24.1 (rotamer B), 23.1, 21.1. FT-IR: ν / cm^{-1} = 3307, 2956, 2870, 1720, 1676, 1635, 1524, 1453, 1389, 1249, 1019, 856, 812, 750, 700. $[\alpha]_D^{20}$ = −47 (10 mg/mL; CHCl$_3$). R$_f$: 0.23 (DCM/MeOH = 40:1).

Ethyl (*R,E*)-7-methyl-5-(((*S*)-2-(*N*-methylpyrazine-2-carboxamido)-3-phenylpropanamido)-4-oxo-oct-2-enoate (1b)
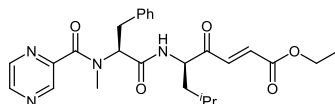
The desired compound was prepared from **8** (284 mg, 0.56 mmol, 1.0 equiv.), LiCl (29 mg, 0.68 mmol, 1.2 equiv.), DIPEA (96 μL, 0.56 mmol, 1.0 equiv.), and ethyl glyoxylate (112 μL, 1.13 mmol, 2.0 equiv.) in dry MeCN (in total 20 mL) according to general procedure B, followed by purification by column chromatography on silica (cyclohexane/ethyl acetate = 5:1) to afford the *E*-isomer as a slightly yellowish oil (164 mg, 0.34 mmol, 61%) and the *Z*-isomer as a colorless oil (25 mg, 0.05 mmol, 9%) giving a total yield of $\sum$ = 189 mg, 0.39 mmol, 70%. *E*-Isomer: ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.73 (d, J = 1.5 Hz, 1H, rotamer B), 8.55 (d, J = 2.5 Hz, 1H, rotamer A + B), 8.50–8.45 (m, 1H, rotamer B), 8.32–8.26 (m, 1H, rotamer A), 8.16 (d, J = 1.5 Hz, 1H, rotamer A), 7.75 (d, J = 7.2 Hz, 1H, rotamer A), 7.26–7.05 (m, 7H, rotamer A + B), 6.83–6.73 (m, 6H, rotamer A + B), 6.70 (d, J = 15.8 Hz, 1H, rotamer B), 5.42 (t, J = 7.9 Hz, 1H, rotamer B), 4.83–4.71 (m, 1H, rotamer A + B), 4.64 (dd, J = 10.8, 4.2 Hz, 1H, rotamer A), 4.26–4.17 (m, 2H, rotamer A), 4.17–4.10 (m, 2H, rotamer B), 3.38 (dd, J = 14.2, 8.0 Hz, 1H, rotamer B), 3.10–2.99 (m, 5H, rotamer A), 2.98–2.87 (m, 5H, rotamer B), 1.79–1.42 (m, 3H, rotamer A + B), 1.29–1.24 (m, 3H, rotamer A), 1.23–1.16 (m, 3H, rotamer B), 1.00–0.91 (m, 6H, rotamer A), 0.89–0.80 (m, 6H, rotamer B). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 197.8 (rotamer B), 197.3 (rotamer A), 169.7 (rotamer B), 169.4 (rotamer A), 168.4 (rotamer B), 167.2 (rotamer A), 165.2 (rotamer A + B), 149.1 (rotamer A), 148.6 (rotamer B), 146.3 (rotamer A), 145.8 (rotamer A), 145.6 (rotamer B), 145.5 (rotamer B), 142.9 (rotamer B), 141.3 (rotamer A), 136.9 (rotamer A), 136.8 (rotamer B), 136.3 (rotamer B), 136.2 (rotamer A), 132.6 (rotamer A), 132.4 (rotamer B), 129.2 (rotamer B), 128.9 (rotamer A), 128.8 (rotamer A), 128.7 (rotamer B), 127.2 (rotamer A), 126.9 (rotamer B), 63.2 (rotamer A), 61.6 (rotamer A), 61.6 (rotamer B), 58.0 (rotamer B), 56.7 (rotamer A), 56.4 (rotamer B), 39.6 (rotamer A), 39.6 (rotamer B), 33.6 (rotamer A), 33.4 (rotamer B), 33.3 (rotamer B), 29.7 (rotamer A), 25.3 (rotamer A), 25.0 (rotamer B), 23.4 (rotamer A), 23.3 (rotamer B), 21.8 (rotamer A), 21.6 (rotamer B), 14.2 (rotamer A), 14.2 (rotamer B). FT-IR: ν / cm^{-1} = 3326, 3060, 3028, 2957, 2871, 1635, 1523, 1389, 1288, 1186, 1081, 1019, 858, 751, 700. $[\alpha]_D^{20}$ = −93 (10 mg/mL; CHCl$_3$). ESI-MS: *m/z* calculated for $[C_{26}H_{32}N_4O_5+H]^+$ ($[M+H]^+$) = 503.2, found: 503.2. Purity: >99% HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 → 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 5.90 min). R$_f$: 0.38 (cyclohexane/ethyl acetate = 1:1).

Benzyl ((*S*)-1-(((*R*)-1-(dimethoxyphosphoryl)-5-methyl-2-oxohexan-3-yl)amino)-1-oxo-3-phenyl-propan-2-yl)carbamate (9)
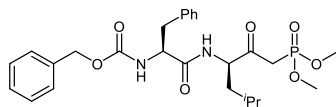

1.) Compound **6** (700 mg, 2.08 mmol) was treated with 4 M HCl in dioxane (5 mL) at room temperature for 1.5 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless foam (568 mg, 2.08 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from Cbz-L-Phe-OH (621 mg, 2.08 mmol, 1.0 equiv.), HOBt • H$_2$O (318 mg, 2.08 mmol, 1.0 equiv.), 2,4,6-collidine (550 μL, 4.15 mmol, 2.0 equiv.), TBTU (666 mg, 2.08 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (568 mg, 2.08 mmol, 1.0 equiv.) in DCM (15 mL) according to general procedure A, followed by purification by column chromatography on silica (DCM/MeOH = 40:1) to afford the desired product (*S,R*-diastereomer) as a yellowish oil (653 mg, 1.26 mmol, 61%).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.37–7.16 (m, 10H), 7.11 (d, *J* = 8.7 Hz, 1H), 5.66 (d, *J* = 7.9 Hz, 1H), 5.11–4.98 (m, 2H), 4.66–4.43 (m, 2H), 3.74 (s, 3H), 3.70 (s, 3H), 3.50–3.29 (m, 1H), 3.17–2.84 (m, 3H), 1.65–1.50 (m, 1H), 1.39–1.20 (m, 2H), 0.91–0.78 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 201.8 (d, *J* = 5.8 Hz), 171.4, 156.1, 136.4, 136.3, 129.4, 128.7, 128.6, 128.2, 128.0, 127.1, 67.0, 57.5 (d, *J* = 1.9 Hz), 56.6, 53.4, 53.3, 39.5, 38.5, 38.4 (d, *J* = 128.8 Hz), 24.5, 23.3, 21.4. FT-IR: ν / cm^{-1} = 3275, 3062, 3031, 2955, 2869, 1715, 1660, 1530, 1454, 1386, 1239, 1026, 816, 740, 697. $[\alpha]_D^{20}$ = +27 (10 mg/mL; CHCl$_3$). R$_f$: 0.39 (DCM/MeOH = 40:1).

Ethyl (*R*,*E*)-5-((*S*)-2-((((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-7-methyl-4-oxooct-2-enoate (1c)
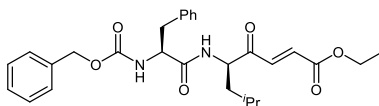
The desired compound was prepared from **9** (322 mg, 0.62 mmol, 1.0 equiv.), LiCl (32 mg, 0.75 mmol, 1.2 equiv.), DIPEA (106 μL, 0.62 mmol, 1.0 equiv.), and ethyl glyoxylate (123 μL, 1.24 mmol, 2.0 equiv.) in dry MeCN (in total 20 mL) according to general procedure B, followed by purification by column chromatography on silica (cyclohexane/ethyl acetate = 5:1) to afford the *E*-isomer (138 mg, 0.28 mmol, 45%) and the *Z*-isomer (33 mg, 0.07 mmol, 11%) each as a colorless oil. *E*-Isomer: ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.35–7.10 (m, 11H), 6.78 (d, *J* = 15.8 Hz, 1H), 6.44 (d, *J* = 7.8 Hz, 1H), 5.43 (d, *J* = 8.0 Hz, 1H), 5.07 (s, 2H), 4.85–4.74 (m, 1H), 4.57–4.43 (m, 1H), 4.25 (q, *J* = 7.1 Hz, 2H), 3.07 (d, *J* = 7.2 Hz, 2H), 1.52–1.42 (m, 1H), 1.41–1.34 (m, 1H), 1.31 (t, *J* = 7.1 Hz, 3H), 1.27–1.18 (m, 1H), 0.93–

0.78 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 197.6, 171.1, 165.2, 156.1, 136.4, 136.3, 136.2, 132.7, 129.4, 128.8, 128.6, 128.3, 128.1, 127.2, 67.2, 61.6, 56.4, 56.0, 40.2, 38.7, 24.8, 23.2, 21.8, 14.2. FT-IR: ν / cm^{-1} = 3293, 3063, 3034, 2951, 2927, 2868, 1688, 1654, 1528, 1285, 1240, 1186, 1040, 736, 696. $[\alpha]_D^{20}$ = −22 (10 mg/mL; CHCl$_3$). ESI-MS: m/z calculated for [C$_{28}$H$_{34}$N$_2$O$_6$+H]$^+$ ([M+H]$^+$) = 495.3, found: 495.2. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O = 55:45 + 0.1% HCOOH, t_R = 4.65 min). R$_f$: 0.27 (cyclohexane/ethyl acetate = 4:1).

tert-Butyl ((_S_)-1-(((_R_)-1-(dimethoxyphosphoryl)-5-methyl-2-oxohexan-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (10)
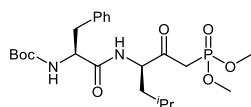
1.) Compound **6** (4.21 g, 12.47 mmol) was treated with 4 M HCl in dioxane (5 mL) at room temperature for 1.5 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless foam (3.41 g, 12.47 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from Boc-L-Phe-OH (3.31 g, 12.47 mmol, 1.0 equiv.), HOBt • H$_2$O (1.91 g, 12.47 mmol, 1.0 equiv.), 2,4,6-collidine (3.31 mL, 24.94 mmol, 2.0 equiv.), TBTU (4.01 g, 12.47 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (3.41 g, 12.47 mmol, 1.0 equiv.) in DCM (15 mL) according to general procedure A, followed by purification by column chromatography on silica (DCM/MeOH = 30:1 → 25:1) to afford the desired product as a colorless solid, which was mainly present as a mixture of two diastereomers ($\sum$ = 3.01 g, 6.21 mmol, 50%). Fractions with pure (_S,R_)-diastereomer could also be obtained and were used for the next reaction step.

(_S,R_)-Diasteromer: ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.32–7.15 (m, 5H), 7.01–6.87 (m, 1H), 5.26–5.09 (m, 1H), 4.63–4.48 (m, 1H), 4.37 (_pseudo_-q, $J \approx$ 7.4 Hz, 1H), 3.82–3.69 (m, 6H), 3.54–3.32 (m, 1H), 3.14–2.88 (m, 3H), 1.63–1.50 (m, 1H), 1.45–1.17 (m, 11H), 0.91–0.76 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 202.0 (d, J = 7.2 Hz), 171.8, 155.5, 136.6, 129.4, 128.8, 127.0, 80.2, 57.5 (d, J = 1.8 Hz), 56.3, 53.3, 53.2, 39.6, 38.4 (d, J = 128.2 Hz), 38.3, 28.3, 24.5, 23.3, 21.4. FT-IR: ν / cm^{-1} = 3359, 3255, 3053, 2959, 2920, 2869, 1696, 1676, 1538, 1229, 1172, 1015, 816, 744, 698. $[\alpha]_D^{20}$ = +32 (10 mg/mL; CHCl$_3$). mp: 109–111 °C. R$_f$: 0.23 (DCM/MeOH = 50:1).

Ethyl (*R,E*)-5-((*S*)-2-((*tert*-butoxycarbonyl)amino)-3-phenylpropanamido)-7-methyl-4-oxooct-2-enoate (1d-(*E*))
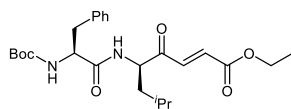
The desired compound was prepared from **10** (301 mg, 0.62 mmol, 1.0 equiv.), LiCl (32 mg, 0.75 mmol, 1.2 equiv.), DIPEA (106 μL, 0.62 mmol, 1.0 equiv.), and ethyl glyoxylate (123 μL, 1.24 mmol, 2.0 equiv.) in dry MeCN (in total 20 mL) according to general procedure B, followed by purification by column chromatography on silica (cyclohexane/ethyl acetate = 5:1) to afford the *E*-isomer as a slightly yellowish oil (151 mg, 0.33 mmol, 53%) and the *Z*-isomer as a colorless oil (53 mg, 0.12 mmol, 18%). *E*-Isomer: ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.33–7.06 (m, 6H), 6.77 (d, *J* = 15.8 Hz, 1H), 6.57 (d, *J* = 7.8 Hz, 1H), 5.12 (d, *J* = 7.9 Hz, 1H), 4.86–4.71 (m, 1H), 4.47–4.33 (m, 1H), 4.24 (q, *J* = 7.1 Hz, 2H), 3.04 (d, *J* = 7.0 Hz, 2H), 1.54–1.21 (m, 15H), 0.99–0.77 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 197.7, 171.5, 165.2, 155.5, 136.7, 136.4, 132.5, 129.3, 128.7, 127.0, 80.4, 61.6, 55.8, 40.1, 38.4, 28.3, 24.8, 23.2, 21.7, 14.2. FT-IR: ν / cm^{-1} = 3301, 3064, 3030, 2960, 2932, 2871, 1702, 1653, 1497, 1366, 1286, 1165, 1027, 979, 699. $[\alpha]_D^{20}$ = −29 (10 mg/mL; CHCl$_3$). ESI-MS: *m/z* calculated for [C$_{25}$H$_{36}$N$_2$O$_6$+Na]$^+$ ([M+Na]$^+$) = 483.3, found: 483.2. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O gradient + 0.1% HCOOH, t_R = 6.57 min). R$_f$: 0.34 (cyclohexane/ethyl acetate = 4:1).

tert-Butyl (4-methyl-1-oxopentan-2-yl)carbamate (11)
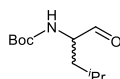
To a solution of compound **5** (1.72 g, 4.86 mmol, 1.0 equiv.) in dry THF (20 mL) was added LiAlH$_4$ (0.24 g, 6.32 mmol, 1.3 equiv.) portionwise at 0 °C under argon atmosphere. After stirring at 0 °C for 30 min, diethyl ether (20 mL) and KHSO$_4$ solution (10 mL, 0.33 M) were added. The suspension was filtered, and the filtrate was extracted with diethyl ether (2 × 50 mL). The combined organic extracts were washed with 1 M HCl (2 × 40 mL) and saturated NaHCO$_3$ solution (2 × 40 mL) and dried over Na$_2$SO$_4$. After removing the solvent under reduced pressure at 40 °C, the desired product was obtained as a colorless oil (1.03 g, 4.78 mmol, 98%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.55 (s, 1H), 5.07–4.91 (m, 1H), 4.28–4.13 (m, 1H), 1.85–1.69 (m, 1H), 1.69–1.52 (m, 1H), 1.42 (s, 9H), 1.40–1.31 (m, 1H), 0.97–0.91 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 200.6, 155.7, 80.1, 58.5, 38.2, 28.4, 24.8, 23.2, 22.0. FT-IR: ν / cm^{-1} = 3342, 2957, 2933, 2871, 1688, 1506, 1455, 1391, 1366, 1250, 1164, 1044, 1011, 873, 779. $[\alpha]_D^{20}$ = +43 (10 mg/mL; MeOH). R$_f$: 0.41 (cyclohexane/ethyl acetate = 5:1).

Phenyl (*E*)-3-((*tert*-butoxycarbonyl)amino)-5-methylhex-1-ene-1-sulfonate (1i)
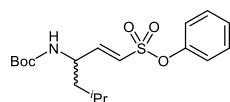
The desired compound was prepared from **12** (1.39 g, 4.50 mmol, 1.0 equiv.), LiCl (0.23 g, 5.40 mmol, 1.2 equiv.), DBU (0.67 mL, 4.50 mmol, 1.0 equiv.), and **11** (0.97 g, 4.50 mmol, 1.0 equiv.) in dry MeCN (in total 20 mL) according to general procedure C, followed by purification by column chromatography on silica (cyclohexane/ethyl acetate = 10:1) to afford the *E*-isomer (1.23 g, 3.32 mmol, 74%) as a slightly yellowish resin. ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.24–7.01 (m, 5H), 6.52 (dd, J = 15.1, 5.2 Hz, 1H), 6.26 (dd, J = 15.1, 1.5 Hz, 1H), 4.33 (d, J = 8.1 Hz, 1H), 4.15 (s, 1H), 1.51–1.35 (m, 1H), 1.27 (s, 10H), 1.21–1.11 (m, 2H), 0.84–0.68 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 155.0, 151.1, 149.6, 129.9, 127.4, 123.9, 122.7, 80.3, 49.7, 43.1, 28.4, 24.8, 22.7, 22.1. FT-IR: ν / cm^{-1} = 3320, 2960, 2871, 1678, 1532, 1487, 1372, 1277, 1164, 1143, 1052, 909, 860, 772, 687. $[\alpha]_D^{20}$ = +14 (10 mg/mL; CHCl$_3$). ESI-MS: *m/z* calculated for [C$_{18}$H$_{27}$NO$_5$S+Na]$^+$ ([M+Na]$^+$) = 392.2, found: 392.1. Purity: >99% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 $\rightarrow$ 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 6.96 min). R$_f$: 0.29 (cyclohexane/ethyl acetate = 10:1).

Phenyl (*R*,*E*)-3-((*S*)-2-((*tert*-butoxycarbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-1-sulfonate (1j)
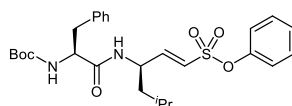
1.) Compound **1i** (1.19 g, 3.22 mmol) was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless solid (0.98 g, 3.22 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from Boc-L-Phe-OH (770 mg, 2.90 mmol, 1.0 equiv.), HOBt • H$_2$O (444 mg, 2.90 mmol, 1.0 equiv.), 2,4,6-collidine (770 μL, 5.81 mmol, 2.0 equiv.), TBTU (932 mg, 2.90 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (888 mg, 2.90 mmol, 1.0 equiv.) in DCM (20 mL) according to general procedure A, followed by purification by column chromatography on silica (DCM/MeOH = 10:1) to afford the desired product as a colorless solid (1336 mg, 2.59 mmol, 89%).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.42–7.34 (m, 2H), 7.33–7.15 (m, 8H), 6.59 (dd, J = 15.1, 4.7 Hz, 1H), 6.45 (d, J = 15.2 Hz, 1H), 5.94 (d, J = 8.0 Hz, 1H), 5.06 (d, J = 7.6 Hz, 1H), 4.59–4.45 (m, 1H), 4.33–4.22 (m, 1H), 3.15–2.97 (m, 2H), 1.42 (s, 9H), 1.24–1.12 (m, 3H), 0.87–0.68 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.1, 155.7, 150.0, 149.6, 136.4, 129.9, 129.3, 129.0, 127.4, 127.3, 124.2,

122.8, 80.9, 56.5, 48.3, 42.5, 37.9, 28.4, 24.5, 22.8, 21.9. FT-IR: ν / cm^{-1} = 3307, 3030, 2960, 2931, 2870, 1662, 1487, 1366, 1246, 1143, 1022, 864, 779, 727, 691. $[\alpha]_D^{20}$ = −5 (10 mg/mL; CHCl$_3$). mp: 46–48 °C. ESI-MS: *m/z* calculated for $[C_{27}H_{36}N_2O_6S+Na]^+$ ($[M+Na]^+$) = 539.2, found: 539.1. Purity: 97% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 → 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 7.11 min). R$_f$: 0.25 (DCM/MeOH = 50:1).

Phenyl (*R,E*)-5-methyl-3-(((*S*)-3-phenyl-2-(pyrazine-2-carboxamido)propanamido)hex-1-ene-1-sulfonate (1g)
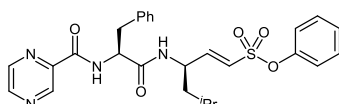
1.) Compound **1j** (300 mg, 0.58 mmol) was treated with 4 M HCl in dioxane (3 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless solid (263 mg, 0.58 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from pyrazinoic acid (72 mg, 0.58 mmol, 1.0 equiv.), HOBt • H$_2$O (89 mg, 0.58 mmol, 1.0 equiv.), 2,4,6-collidine (154 μL, 1.16 mmol, 2.0 equiv.), TBTU (186 mg, 0.58 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (263 mg, 0.58 mmol, 1.0 equiv.) in DCM (20 mL) according to general procedure A, followed by purification by column chromatography on silica (ethyl acetate) to afford the desired product as a colorless solid (181 mg, 0.35 mmol, 60%).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.28 (d, *J* = 1.5 Hz, 1H), 8.76 (d, *J* = 2.4 Hz, 1H), 8.59–8.52 (m, 1H), 8.38 (d, *J* = 7.7 Hz, 1H), 7.36–7.12 (m, 10H), 6.62 (dd, *J* = 15.1, 4.9 Hz, 1H), 6.51 (dd, *J* = 15.2, 1.2 Hz, 1H), 6.22 (d, *J* = 8.0 Hz, 1H), 4.86–4.75 (m, 1H), 4.61–4.48 (m, 1H), 3.33–3.12 (m, 2H), 1.25–1.11 (m, 3H), 0.84–0.73 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.1, 163.3, 149.9, 149.5, 147.7, 144.3, 143.9, 143.0, 136.2, 129.8, 129.3, 129.0, 127.4, 127.3, 124.3, 122.7, 55.4, 48.4, 42.4, 38.4, 24.4, 22.7, 22.0. FT-IR: ν / cm^{-1} = 3369, 3053, 3029, 2960, 2918, 2866, 1652, 1504, 1370, 1141, 1022, 865, 773, 721, 690. $[\alpha]_D^{20}$ = −4 (10 mg/mL; CHCl$_3$). mp: 119–121 °C. ESI-MS: *m/z* calculated for $[C_{27}H_{30}N_4O_5S+H]^+$ ($[M+H]^+$) = 523.2, found: 523.3. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 → 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 6.44 min). R$_f$: 0.39 (cyclohexane/ethyl acetate = 1:1).

tert-Butyl (*R*)-(1-(methoxy(methyl)amino)-4-methyl-1-oxopentan-2-yl)(methyl)carbamate (14)
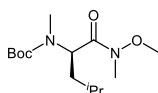
The desired compound was prepared from Boc-*N*-Me-D-Leu-OH (2.12 g, 8.63 mmol, 1.0 equiv.), HOBt · H$_2$O (1.32 g, 8.63 mmol, 1.0 equiv.), 2,4,6-collidine (2.29 mL, 17.26 mmol, 2.0 equiv.), TBTU (2.77 g, 8.63 mmol, 1.0 equiv.), and *N*,*O*-dimethylhydroxylamine hydrochloride (0.84 g, 8.63 mmol, 1.0 equiv.) in DCM (30 mL) according to general procedure A to afford the desired product as a colorless oil (2.13 g, 7.38 mmol, 86%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 5.36–4.95 (m, 1H), 3.75–3.63 (m, 3H), 3.15 (s, 3H), 2.81 (s, 3H), 1.69–1.56 (m, 1H), 1.54–1.32 (m, 11H), 0.95–0.88 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 156.2, 155.5, 80.1 (rotamer B), 79.6 (rotamer A), 61.6 (rotamer B), 61.5 (rotamer A), 53.1 (rotamer B), 51.9 (rotamer A), 37.8 (rotamer B), 37.6 (rotamer A), 32.4 (rotamer A), 32.2 (rotamer B), 30.0 (rotamer B), 29.5 (rotamer A), 28.5 (rotamer B), 28.5 (rotamer A), 25.0 (rotamer A), 24.6 (rotamer B), 23.3 (rotamer A), 21.9 (rotamer B). FT-IR: ν / cm^{-1} = 3317, 2956, 2934, 2870, 1669, 1453, 1366, 1323, 1260, 1149, 1127, 995, 749, 700, 661. $[\alpha]_D^{20}$ = +60 (10 mg/mL; CHCl$_3$). R$_f$: 0.60 (cyclohexane/ethyl acetate = 2:1).

tert-Butyl (1-(dimethoxyphosphoryl)-5-methyl-2-oxohexan-3-yl)(methyl)carbamate (15)
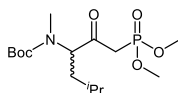
Using a modified procedure by Ho *et al., Eur. J. Org. Chem.* **2005**, *2005*, 4829–4834.

A solution of *n*-BuLi (2.5 M in hexane, 13.48 mL, 33.69 mmol, 5.3 equiv.) was slowly added to a solution of dimethyl methylphosphonate (4.08 mL, 38.14 mmol, 6.0 equiv.) in dry THF (25 mL) at −78 °C under argon atmosphere. Stirring was continued for 1 h at −78 °C, and then a solution of **14** (1.83 g, 6.36 mmol, 1.0 equiv.) in dry THF (8 mL) was added dropwise, and the mixture was stirred for an additional 30 min at −78 °C. The reaction mixture was quenched by addition of saturated NH$_4$Cl solution (100 mL) and extracted with ethyl acetate (3 × 70 mL). The combined organic phases were washed with water (5 × 80 mL), dried over Na$_2$SO$_4$, and the solvent was removed by distillation under reduced pressure at 40 °C to yield the title compound as a colorless oil (2.15 g, 6.12 mmol, 96%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 4.73–4.37 (m, 1H), 3.78–3.70 (m, 6H), 3.38–3.15 (m, 1H), 3.06–2.82 (m, 1H), 2.77–2.64 (m, 3H), 1.68–1.56 (m, 1H), 1.55–1.36 (m, 11H), 0.93–0.85 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 200.3 (d, *J* = 6.7 Hz) (rotamer B), 200.1 (d, *J* = 6.2 Hz) (rotamer A), 156.1 (rotamer B), 155.1 (rotamer A), 81.1 (rotamer A), 80.5 (rotamer B), 64.1 (d, *J* = 2.0 Hz) (rotamer B), 64.7 (d, *J* = 2.0) (rotamer A), 53.2 (rotamer B), 53.1 (rotamer B), 53.0 (rotamer A), 52.9 (rotamer A), 37.5 (d, *J* = 131.2 Hz), 36.2 (rotamer B), 35.4 (rotamer A), 31.7 (rotamer B), 31.2 (rotamer A), 28.3, 24.9 (rotamer B), 24.7

(rotamer A), 23.3 (rotamer A), 23.2 (rotamer B), 21.9 (rotamer B), 21.8 (rotamer A). FT-IR: ν / cm^{-1} = 2957, 2934, 2871, 1688, 1454, 1389, 1367, 1310, 1254, 1152, 1027, 909, 869, 811, 772. $[\alpha]_D^{20}$ = +170 (10 mg/mL; CHCl$_3$). R$_f$: 0.42 (cyclohexane/ethyl acetate = 1:3).

***tert*-Butyl ((2*S*)-1-((1-(dimethoxyphosphoryl)-5-methyl-2-oxohexan-3-yl)(methyl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (16)**
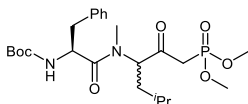
1.) Compound **15** (2.18 g, 6.20 mmol) was treated with 4 M HCl in dioxane (10 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless foam (1.78 g, 6.20 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from Boc-L-Phe-OH (1.65 g, 6.20 mmol, 1.0 equiv.), HOBt • H$_2$O (0.95 g, 6.20 mmol, 1.0 equiv.), 2,4,6-collidine (1.64 mL, 12.40 mmol, 2.0 equiv.), TBTU (1.99 g, 6.20 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (1.78 g, 6.20 mmol, 1.0 equiv.) in DCM (40 mL) according to general procedure A, followed by purification by column chromatography on silica (DCM/MeOH = 30:1) to afford the desired product as a colorless solid, which was present as a mixture of two diastereomers ($\sum$ = 0.88 g, 1.35 mmol, 41%, ratio not determined).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.32–7.13 (m, 5H), 5.31–5.20 (m, 1H), 5.17–5.06 (m, 1H), 4.89–4.75 (m, 1H), 3.78–3.68 (m, 6H), 3.20–2.88 (m, 4H), 2.65 (s, 3H), 1.61–1.47 (m, 1H), 1.38 (s, 9H), 1.31–1.16 (m, 1H), 1.11–0.96 (m, 1H), 0.84–0.77 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 199.4 (d, J = 6.7 Hz), 173.2, 155.4, 136.1, 129.4, 128.6, 127.1, 80.1, 61.1 (d, J = 2.9 Hz), 53.2 (diastereomer or rotamer A), 53.1 (diastereomer or rotamer A), 52.9 (diastereomer or rotamer B), 52.8 (diastereomer or rotamer B), 51.9, 39.2, 37.8 (d, J = 131.7 Hz), 35.1, 31.4, 28.3, 24.4, 23.1, 22.0. FT-IR: ν / cm^{-1} = 3244, 3030, 2956, 2934, 2868, 1688, 1648, 1537, 1365, 1239, 1169, 1035, 860, 739, 699. $[\alpha]_D^{20}$ = +174 (10 mg/mL; CHCl$_3$). mp: 114–117 °C. R$_f$: 0.30 (DCM/MeOH = 40:1).

Dimethyl ((*R*)-5-methyl-3-((*S*)-*N*-methyl-3-phenyl-2-(pyrazine-2-carboxamido)propanamido)-2-oxohexyl)phosphonate (17)
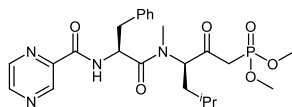

1.) Compound **16** (1.06 g, 2.12 mmol) was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1.5 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless foam (0.92 g, 2.12 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from pyrazinoic acid (0.24 g, 1.93 mmol, 1.0 equiv.), HOBt • H$_2$O (0.30 g, 1.93 mmol, 1.0 equiv.), 2,4,6-collidine (0.51 mL, 3.85 mmol, 2.0 equiv.), TBTU (0.62 g, 1.93 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (0.92 g, 2.12 mmol, 1.1 equiv.) in DCM (10 mL) and DMF (10 mL) according to general procedure A, followed by purification by column chromatography on silica (DCM/MeOH = 30:1) to afford the (*S*,*R*)-diastereomer) as a yellowish oil (0.31 g, 0.61 mmol, 32%). Fractions consisting of a mixture of the two diastereomers were also collected (0.11 g, 0.21 mmol, 11%) to give a total yield of $\sum$ = 0.42 g, 0.82 mmol, 43%.

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.34 (d, J = 1.5 Hz, 1H), 8.76 (d, J = 2.5 Hz, 1H), 8.56 (dd, J = 2.5, 1.5 Hz, 1H), 8.44 (d, J = 8.3 Hz, 1H), 7.34–7.20 (m, 5H), 5.45–5.34 (m, 1H), 5.21–5.13 (m, 1H), 3.78–3.68 (m, 6H), 3.25–2.93 (m, 4H), 2.77 (s, 3H), 1.67–1.56 (m, 1H), 1.38–1.25 (m, 1H), 1.20–1.04 (m, 1H), 0.89–0.83 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 199.2 (d, J = 6.8 Hz), 172.3, 162.7, 147.6, 144.4, 144.2, 142.9, 135.8, 129.5, 128.8, 127.4, 61.42 (d, J = 2.8 Hz), 53.3 (rotamer A), 53.2 (rotamer A), 53.0 (rotamer B), 52.9 (rotamer B), 50.9, 39.3, 38.1 (d, J = 130.9 Hz), 35.2, 31.6, 24.6, 23.1, 22.0. FT-IR: ν / cm^{-1} = 3390, 3292, 2956, 2869, 1717, 1639, 1515, 1398, 1255, 1180, 1019, 868, 815, 746, 701. $[\alpha]_D^{20}$ = +160 (5 mg/mL; CHCl$_3$). R$_f$: 0.23 (DCM/MeOH = 30:1).

Ethyl (*R*,*E*)-7-methyl-5-((*S*)-*N*-methyl-3-phenyl-2-(pyrazine-2-carboxamido)propanamido)-4-oxo-oct-2-enoate (1a)
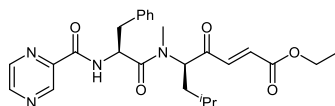
The desired compound was prepared from **17** (277 mg, 0.55 mmol, 1.0 equiv.), LiCl (28 mg, 0.66 mmol, 1.2 equiv.), DIPEA (93 μL, 0.55 mmol, 1.0 equiv.), and ethyl glyoxylate (112 μL, 1.10 mmol, 2.0 equiv.) in dry MeCN (in total 20 mL) according to general procedure B, followed by purification by column chromatography on silica (cyclohexane/ethyl acetate = 2:1) to afford the *E*-isomer (117 mg, 0.24 mmol, 44%) as a yellowish oil. ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.36–9.27 (m, 1H), 8.78–8.70 (m, 1H),

8.56–8.51 (m, 1H), 8.47 (d, J = 8.6 Hz, 1H), 7.32–7.16 (m, 5H), 7.03 (d, J = 15.8 Hz, 1H), 6.71 (d, J = 15.8 Hz, 1H), 5.46–5.32 (m, 2H), 4.16–4.02 (m, 2H), 3.27–3.08 (m, 2H), 2.73 (s, 3H), 1.64–1.50 (m, 1H), 1.44–1.23 (m, 1H), 1.17 (t, J = 7.1 Hz, 3H), 1.13–1.00 (m, 1H), 0.89–0.79 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 197.0, 172.2, 165.1, 162.5, 147.4, 144.4, 144.2, 142.9, 136.8, 135.8, 132.2, 129.4, 128.7, 127.3, 61.3, 59.6, 50.7, 39.4, 35.4, 31.1, 24.5, 23.2, 21.9, 14.0. FT-IR: ν / cm^{-1} = 3382, 2956, 2870, 1723, 1640, 1514, 1397, 1284, 1184, 1081, 1020, 979, 864, 753, 700. $[\alpha]_D^{20}$ = +134 (10 mg/mL; CHCl$_3$). ESI-MS: m/z calculated for [C$_{26}$H$_{32}$N$_4$O$_5$+Na]$^+$ ([M+Na]$^+$) = 503.2, found: 503.1. Purity: 96% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 $\rightarrow$ 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 6.49 min). R$_f$: 0.40 (cyclohexane/ethyl acetate = 1:1).

tert-Butyl 2-(((benzyloxy)carbonyl)-L-phenylalanyl)hydrazine-1-carboxylate (19)
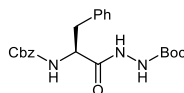
The desired compound was prepared from Cbz-L-Leu-OH (2.00 g, 6.68 mmol, 1.0 equiv.), HOBt • H$_2$O (1.02 g, 6.68 mmol, 1.0 equiv.), 2,4,6-collidine (0.89 mL, 6.68 mmol, 1.0 equiv.), TBTU (2.15 g, 6.68 mmol, 1.0 equiv.), and *tert*-butyl carbazate (0.88 mg, 6.68 mmol, 1.0 equiv.) in DCM (30 mL) according to general procedure A to afford the desired product as a colorless solid (2.44 g, 5.90 mmol, 88%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.67–8.38 (m, 1H), 7.35–7.15 (m, 10H), 6.86–6.71 (m, 1H), 5.78–5.59 (m, 1H), 5.08–4.92 (m, 2H), 4.67–4.51 (m, 1H), 3.25–2.92 (m, 2H), 1.45 (s, 9H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.0, 156.4, 155.4, 136.3, 136.1, 129.5, 128.7, 128.6, 128.2, 128.1, 127.1, 82.0, 67.3, 54.7, 38.3, 28.3. FT-IR: ν / cm^{-1} = 3282, 3062, 3032, 2977, 2934, 1677, 1497, 1454, 1367, 1237, 1156, 1046, 851, 738, 696. $[\alpha]_D^{20}$ = −7 (10 mg/mL; CHCl$_3$). mp: 56–58 °C. R$_f$: 0.64 (cyclohexane/ethyl acetate = 1:1).

tert-Butyl 2-(L-phenylalanyl)hydrazine-1-carboxylate (20)
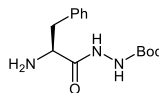
To a solution of compound **19** (2.36 g, 5.71 mmol, 1.0 equiv.) in methanol (30 mL) was added palladium (10% on activated charcoal; 0.24 g, 10 wt%). The suspension was stirred under hydrogen atmosphere (760 Torr) for 21 h at room temperature and filtered over Celite®. Distillation of the solvent under reduced pressure at 40 °C yielded the desired product as a colorless solid (1.56 g, 5.57 mmol, 97%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.44–7.19 (m, 5H), 3.87–3.74 (m, 1H), 3.36–3.23 (m, 1H), 2.88–2.71 (m, 1H), 1.47 (s, 9H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 173.1, 155.5, 137.3, 129.5, 128.9, 127.1, 81.8, 55.6, 40.7, 28.3. FT-IR: ν / cm^{-1} = 3355, 3293, 3228, 2984, 2930, 1739, 1682, 1493, 1366, 1240, 1152,

913, 859, 754, 706. $[\alpha]_D^{20}$ = −54 (10 mg/mL; CHCl$_3$). mp: 105–107 °C. R$_f$: 0.28 (cyclohexane/ethyl acetate = 1:4).

tert-Butyl 2-((pyrazine-2-carbonyl)-L-phenylalanyl)hydrazine-1-carboxylate (21)
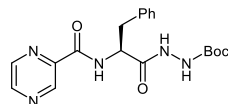
The desired compound was prepared from pyrazinoic acid (0.61 g, 4.90 mmol, 1.0 equiv.), HOBt • H$_2$O (0.75 g, 4.90 mmol, 1.0 equiv.), 2,4,6-collidine (1.30 mL, 9.80 mmol, 2.0 equiv.), TBTU (1.57 g, 4.90 mmol, 1.0 equiv.), and **20** (1.51 g, 5.39 mmol, 1.1 equiv.) in DCM (10 mL) and DMF (10 mL) according to general procedure A to afford the desired product as a colorless solid (1.68 g, 4.35 mmol, 89%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.30–9.22 (m, 1H), 8.73–8.66 (m, 1H), 8.51–8.43 (m, 1H), 8.41–8.31 (m, 1H), 7.29–7.12 (m, 5H), 6.88–6.70 (m, 1H), 5.09–4.96 (m, 1H), 3.37–3.14 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.4, 163.4, 155.4, 147.6, 144.4, 143.8, 142.9, 136.2, 129.4, 128.7, 127.2, 81.9, 53.0, 38.1, 28.2. FT-IR: ν / cm^{-1} = 3271, 2978, 2928, 1662, 1518, 1455, 1393, 1367, 1239, 1154, 1047, 1020, 867, 750, 699. $[\alpha]_D^{20}$ = −42 (10 mg/mL; CHCl$_3$). mp: 70–72 °C. R$_f$: 0.19 (cyclohexane/ethyl acetate = 1:1).

(_S_)-_N_-(1-hydrazinyl-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (22)
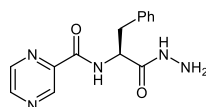
Compound **21** (1.95 g, 5.78 mmol) was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1 h. After the solvent was removed under reduced pressure at 40 °C, the residue was taken up in DCM (100 mL) and washed with saturated NaHCO$_3$ solution (80 mL). The aqueous phase was extracted with DCM (2 × 60 mL), and the combined organic extracts were dried over Na$_2$SO$_4$. Removal of the solvent by distillation under reduced pressure at 40 °C yielded the desired product as a colorless foam (1.12 g, 3.91 mmol, 95%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.30 (d, J = 1.5 Hz, 1H), 8.73 (d, J = 2.5 Hz, 1H), 8.50 (dd, J = 2.5, 1.5 Hz, 1H), 8.40 (d, J = 8.5 Hz, 1H), 4.94–4.83 (m, 1H), 3.29–3.13 (m, 2H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.1, 163.2, 147.7, 144.5, 143.9, 142.9, 136.3, 129.3, 128.8, 127.3, 53.5, 38.4. FT-IR: ν / cm^{-1} = 3371, 3258, 3027, 2920, 2851, 1654, 1511, 1389, 1260, 1148, 1023, 989, 872, 739, 696. $[\alpha]_D^{20}$ = −28 (10 mg/mL; CHCl$_3$). mp: 89–91 °C. R$_f$: 0.12 (cyclohexane/ethyl acetate = 1:1 + 1% NEt$_3$).

(*S*)-*N*-(1-(2-isobutylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (23)
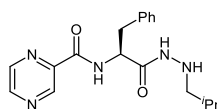
To a solution of compound **22** (300 mg, 1.05 mmol, 1.0 equiv.) in THF (5 mL) was added isobutyraldehyde (106 μL, 1.16 mmol, 1.1 equiv.) dropwise. After stirring at room temperature for 6 h, NaBH$_4$ (40 mg, 1.05 mmol, 1.0 equiv.) was added. Stirring was continued for an additional 1.5 h and the suspension was then quenched by the addition of saturated NaHCO$_3$ solution (20 mL). The mixture was extracted with DCM (2 × 25 mL), and the combined organic extracts were dried over Na$_2$SO$_4$. The solvent was removed by distillation under reduced pressure at 40 °C, and the residue was purified by column chromatography on silica (cyclohexane/ethyl acetate = 1:1.3) to afford the desired product as a pale beige solid (130 mg, 0.38 mmol, 36% over two steps). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.28 (d, J = 1.5 Hz, 1H), 8.72 (d, J = 2.5 Hz, 1H), 8.50 (dd, J = 2.5, 1.5 Hz, 1H), 8.42 (d, J = 8.6 Hz, 1H), 7.25–7.15 (m, 5H), 4.90–4.79 (m, 1H), 3.18 (d, J = 7.4 Hz, 2H), 2.54–2.40 (m, 2H), 1.59–1.46 (m, 1H), 0.83 (d, J = 6.7 Hz, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 169.6, 163.0, 147.6, 144.4, 144.0, 142.9, 136.3, 129.4, 128.7, 127.2, 59.8, 53.5, 38.7, 26.8, 20.5, 20.4. FT-IR: ν / cm^{-1} = 3310, 3063, 3023, 2953, 2925, 2868, 1650, 1516, 1387, 1243, 1152, 1022, 867, 749, 698. $[\alpha]_D^{20}$ = −21 (10 mg/mL; CHCl$_3$). mp: 120–122 °C. R$_f$: 0.35 (cyclohexane/ethyl acetate = 1:3).

Ethyl (*E*)-4-(1-isobutyl-2-((pyrazine-2-carbonyl)-L-phenylalanyl)hydrazinyl)-4-oxobut-2-enoate (1k)
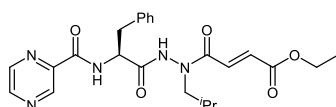
To a solution of compound **23** (107 mg, 0.31 mmol, 1.0 equiv.) in DCM (4 mL) was added NEt$_3$ (43 μL, 0.31 mmol, 1.0 equiv.) and ethyl fumaroyl chloride (42 μL, 0.31 mmol, 1.0 equiv.). After stirring at room temperature for 22 h, the solvent was removed under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (cyclohexane/ethyl acetate = 1:1 $\rightarrow$ 1:1.8) to afford the desired product as a colorless solid (80 mg, 0.17 mmol, 55%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.33–9.08 (m, 2H), 8.83–8.68 (m, 1H), 8.59–8.47 (m, 1H), 8.37 (d, J = 8.1 Hz, 1H), 7.29–7.14 (m, 6H), 6.71 (d, J = 15.3 Hz, 1H), 5.14–4.94 (m, 1H), 4.30–3.86 (m, 2H), 3.65–2.94 (m, 4H), 1.62–1.45 (m, 1H), 1.13 (t, J = 7.1 Hz, 3H), 0.87–0.69 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 169.9, 166.5, 165.4, 163.8, 147.8, 144.4, 143.6, 143.0, 135.7, 132.6, 132.4, 129.4, 128.9, 127.4, 61.0, 55.4, 53.2, 37.7, 26.4, 20.1, 19.9, 14.1. FT-IR: ν / cm^{-1} = 3265, 3029, 2960, 2929, 2872, 1659, 1516, 1403, 1288, 1166, 1128, 1020, 974, 750, 700. $[\alpha]_D^{20}$ = −42 (10 mg/mL; CHCl$_3$). mp: 57–59 °C. ESI-MS: *m/z* calculated for [C$_{24}$H$_{29}$N$_5$O$_5$+H]$^+$ ([M+H]$^+$) = 468.2, found: 468.3. Purity: 98% (HPLC, 254 nm, MeCN/H$_2$O + 0.1%

HCOOH = 10:90 → 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 5.40 min). R_f: 0.29 (cyclohexane/ethyl acetate = 1:1).

Methyl (*tert*-butoxycarbonyl)-L-phenylalanyl-D-leucinate (25)
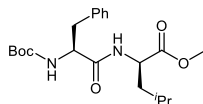
The desired compound was prepared from Boc-L-Phe-OH (1.20 mg, 3.77 mmol, 1.0 equiv.), HOBt • H_2O (0.58 g, 3.77 mmol, 1.0 equiv.), 2,4,6-collidine (1.00 mL, 7.54 mmol, 2.0 equiv.), TBTU (1.21 g, 3.77 mmol, 1.0 equiv.), and H-D-Leu-OMe • HCl (0.69 g, 3.77 mmol, 1.0 equiv.) in DCM (20 mL) according to general procedure A to afford the desired product as a colorless solid (1.39 g, 3.53 mmol, 94%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.34–7.16 (m, 5H), 6.34 (d, J = 8.3 Hz, 1H), 5.09 (s, 1H), 4.60–4.47 (m, 1H), 4.40 (d, J = 8.7 Hz, 1H), 3.70 (s, 3H), 3.07 (d, J = 7.0 Hz, 2H), 1.40 (s, 12H), 0.92–0.79 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 173.2, 171.2, 155.4, 136.8, 129.4, 128.7, 127.0, 80.3, 55.9, 52.3, 50.7, 41.4, 38.6, 28.3, 24.6, 22.8, 21.9. FT-IR: ν / cm^{-1} = 3333, 3069, 2954, 2931, 2869, 1742, 1688, 1631, 1515, 1366, 1271, 1153, 1019, 747, 704. $[\alpha]_D^{20}$ = −1 (10 mg/mL; CHCl$_3$). mp: 138–140 °C. R_f: 0.60 (cyclohexane/ethyl acetate = 2:1).

Methyl (pyrazine-2-carbonyl)-L-phenylalanyl-D-leucinate (26)
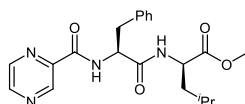
1.) Compound **25** (1.17 g, 2.98 mmol) was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless foam (0.98 g, 2.98 mmol, quantitative), which was directly used in the next step.

2.) The desired compound was prepared from pyrazinoic acid (0.37 g, 2.98 mmol, 1.0 equiv.), HOBt • H_2O (0.46 g, 2.98 mmol, 1.0 equiv.), 2,4,6-collidine (0.79 mL, 5.96 mmol, 2.0 equiv.), TBTU (0.96 mg, 2.98 mmol, 1.0 equiv.), and the amine hydrochloride from step 1 (0.98 mg, 2.98 mmol, 1.0 equiv.) in DCM (10 mL) according to general procedure A to afford the product as a colorless resin (1.05 g, 2.62 mmol, 88%).

^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.30 (d, J = 1.5 Hz, 1H), 8.72 (d, J = 2.4 Hz, 1H), 8.51 (dd, J = 2.5, 1.5 Hz, 1H), 8.42 (d, J = 8.3 Hz, 1H), 7.27–7.16 (m, 5H), 6.51 (d, J = 8.2 Hz, 1H), 5.03–4.90 (m, 1H), 4.57–4.45 (m, 1H), 3.62 (s, 3H), 3.30–3.11 (m, 2H), 1.51–1.39 (m, 1H), 1.38–1.23 (m, 2H), 0.86–0.78 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 173.1, 170.2, 162.9, 147.5, 144.3, 144.1, 142.9, 136.5, 129.4, 128.7, 127.1, 54.7, 52.3, 50.8, 41.3, 38.9, 24.6, 22.8, 21.8. FT-IR: ν / cm^{-1} = 3338, 3029,

2951, 2869, 1746, 1646, 1584, 1511, 1379, 1148, 1025, 875, 779, 750, 698. $[\alpha]_D^{20} = -15$ (10 mg/mL; CHCl$_3$). R$_f$: 0.54 (cyclohexane/ethyl acetate = 1:3).

(Pyrazine-2-carbonyl)-L-phenylalanyl-D-leucine (27)
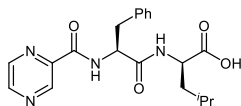
To a solution of compound **26** (945 mg, 2.37 mmol, 1.0 equiv.) in THF (25 mL) and water (25 mL) was added LiOH • H$_2$O (298 mg, 7.11 mmol, 3.0 equiv.). After stirring at room temperature for 6 h, THF was removed by distillation under reduced pressure at 40 °C, and the aqueous solution was acidified to pH 1 with 1 M HCl. The resulting suspension was extracted with ethyl acetate and the combined organic extracts were dried over Na$_2$SO$_4$. After removing the solvent by distillation under reduced pressure at 40 °C, the title compound was obtained as a colorless foam (1.05 g, 2.72 mmol, quantitative). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.90 (s, 1H), 9.19 (d, J = 1.4 Hz, 1H), 8.68 (d, J = 2.5 Hz, 1H), 8.58 (d, J = 8.7 Hz, 1H), 8.50 (dd, J = 2.5, 1.5 Hz, 1H), 7.25–7.16 (m, 5H), 7.07 (d, J = 8.1 Hz, 1H), 5.25–5.12 (m, 1H), 4.56–4.45 (m, 1H), 3.18 (d, J = 7.5 Hz, 2H), 1.56–1.44 (m, 1H), 1.41–1.29 (m, 2H), 0.83 (d, J = 6.1 Hz, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 175.4, 170.8, 163.1, 147.3, 144.2, 144.0, 143.2, 136.3, 129.5, 128.7, 127.1, 54.7, 50.8, 41.2, 39.2, 24.6, 22.9, 21.9. FT-IR: ν / cm^{-1} = 3297, 3062, 2957, 2870, 2582, 1725, 1651, 1521, 1467, 1403, 1151, 1020, 867, 746, 698. $[\alpha]_D^{20} = +4$ (10 mg/mL; CHCl$_3$). mp: 65–67 °C. R$_f$: 0.26 (cyclohexane/ethyl acetate = 1:3 + 1% HOAc).

N-((S)-1-(((R)-4-methyl-1-oxo-1-(prop-2-yn-1-ylamino)pentan-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (1e)
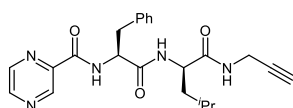
To a solution of compound **27** (200 mg, 0.52 mmol, 1.0 equiv.), HOBt • H$_2$O (80 mg, 0.52 mmol, 1.0 equiv.), and TBTU (167 mg, 0.52 mmol, 1.0 equiv.) in DCM (5 mL) was added propargylamine (33 μL, 0.52 mmol, 1.0 equiv.) and 2,4,6-collidine (69 μL, 0.52 mmol, 1.0 equiv.) at 0 °C. After the solution was stirred at 0 °C for 1 h, the solvent was removed under reduced pressure at 40 °C. The residue was taken up in ethyl acetate (25 mL) and washed with saturated NaHCO$_3$ solution (2 × 25 mL) and 1 M HCl (2 × 25 mL). The combined organic extracts were dried over Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C. Purification of the residue by column chromatography on silica (DCM/MeOH = 30:1) yielded the desired product as a colorless solid (115 mg, 0.27 mmol, 53%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.35–9.27 (m, 1H), 8.73 (d, J = 2.4 Hz, 1H), 8.55–8.47 (m, 1H), 8.40 (d, J = 7.3 Hz, 1H), 7.28–7.15 (m, 5H), 6.84 (d, J = 8.5 Hz, 1H), 4.88–4.77 (m, 1H), 4.47–4.36 (m,

1H), 3.95 (dd, J = 5.3, 2.6 Hz, 2H), 3.20 (d, J = 7.5 Hz, 2H), 2.11 (t, J = 2.5 Hz, 1H), 1.65–1.53 (m, 1H), 1.39–1.27 (m, 1H), 1.24–1.09 (m, 1H), 0.81–0.74 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 171.6, 170.8, 163.4, 147.7, 144.5, 143.9, 142.9, 136.1, 129.3, 128.8, 127.2, 79.6, 71.5, 55.6, 51.8, 40.6, 38.4, 29.3, 24.4, 23.1, 21.8. FT-IR: ν / cm^{-1} = 3291, 3063, 2958, 2926, 2870, 1636, 1522, 1370, 1223, 1194, 1156, 1020, 867, 748, 698. $[\alpha]_D^{20}$ = +41 (10 mg/mL; CHCl$_3$). mp: 170–172 °C. ESI-MS: m/z calculated for $[C_{23}H_{27}N_5O_3+H]^+$ ($[M+H]^+$) = 444.2, found: 444.1. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 $\rightarrow$ 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 4.94 min). R$_f$: 0.28 (DCM/MeOH = 30:1).

N-((S)-1-(((R)-1-((cyanomethyl)amino)-4-methyl-1-oxopentan-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (1f)
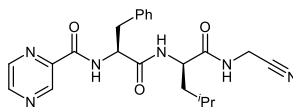
To a solution of compound **27** (100 mg, 0.26 mmol, 1.0 equiv.), aminoacetonitrile hydrochloride (24 mg, 0.26 mmol, 1.0 equiv.), HOBt • H$_2$O (40 mg, 0.26 mmol, 1.0 equiv.), and TBTU (84 mg, 0.26 mmol, 1.0 equiv.) in DCM (5 mL) was added 2,4,6-collidine (69 μL, 0.52 mmol, 2.0 equiv.) at 0 °C. The solution was stirred for 2 h at 0 °C, then 24 h at room temperature. After removal of the solvent under reduced pressure at 40 °C, the residue was taken up in ethyl acetate (25 mL) and washed with saturated NaHCO$_3$ solution (2 × 25 mL) and 1 M HCl (2 × 25 mL). The combined organic extracts were dried over Na$_2$SO$_4$, and the solvent was removed under reduced pressure at 40 °C to yield the desired product as a colorless solid (101 mg, 0.24 mmol, 92%). ^{1}H NMR (300 MHz, DMSO-d_6): δ / ppm = 9.16 (d, J = 1.5 Hz, 1H), 8.88 (d, J = 2.5 Hz, 1H), 8.78–8.66 (m, 3H), 8.56 (d, J = 8.1 Hz, 1H), 7.27–7.13 (m, 5H), 4.94–4.80 (m, 1H), 4.34–4.20 (m, 1H), 4.15 (d, J = 5.5 Hz, 2H), 3.19–3.02 (m, 2H), 1.52–1.32 (m, 3H), 0.88–0.76 (m, 6H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ / ppm = 172.6, 170.4, 162.4, 147.9, 144.0, 143.5, 143.4, 137.0, 129.3, 128.1, 126.4, 117.5, 54.1, 50.8, 40.4, 37.9, 27.1, 24.0, 23.0, 21.2. FT-IR: ν / cm^{-1} = 3285, 3061, 2952, 2926, 2868, 1647, 1522, 1257, 1230, 1161, 1049, 1020, 876, 748, 699. $[\alpha]_D^{20}$ = +68 (10 mg/mL; MeOH). mp: 175–177 °C. ESI-MS: m/z calculated for $[C_{22}H_{26}N_6O_3+Na]^+$ ($[M+Na]^+$) = 445.2, found: 445.1. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 $\rightarrow$ 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 4.84 min). R$_f$: 0.32 (cyclohexane/ethyl acetate = 1:5).

(*R*)-2-Amino-3-phenyl-*N*-(pyrazin-2-yl)propenamide hydrochloride (29)
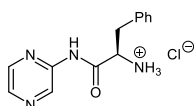
To a solution of Boc-D-Phe-OH (500 mg, 1.88 mmol, 1.0 equiv.) was added Oxyma (268 mg, 1.88 mmol, 1.0 equiv.), COMU (807 mg, 1.88 mmol, 1.0 equiv.), and 2,4,6-collidine (250 μL, 1.88 mmol, 1.0 equiv.) at 0 °C. After stirring at 0 °C for 15 min, 2-aminopyrazine (179 mg, 1.88 mmol, 1.0 equiv.) was added and stirring at 0 °C was continued for an additional hour. The solution was allowed to warm up to room temperature and stirred overnight. Then the mixture was washed with saturated $NaHCO_3$ solution (2 × 25 mL) and 1 M HCl (2 × 25 mL), and the combined organic extracts were dried over Na_2SO_4. After the solvent was removed under reduced pressure at 40 °C, the residue was purified by column chromatography on silica (cyclohexane/ethyl acetate = 2:1) to yield the Boc-protected intermediate as a colorless solid. The intermediate was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1 h. Then, the solvent was removed under reduced pressure at 40 °C to yield the desired product (**29**) as a colorless solid (129 mg, 0.46 mmol, 24%). ^{1}H NMR (300 MHz, DMSO-d_6): δ / ppm = 11.44 (s, 1H), 9.31–9.15 (m, 1H), 8.85–8.60 (m, 3H), 8.47–8.37 (m, 2H), 7.35–7.17 (m, 5H), 4.50–4.37 (m, 1H), 3.31–3.12 (m, 2H). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ / ppm = 168.1, 147.9, 143.0, 140.5, 136.3, 134.7, 129.6, 128.6, 127.3, 53.8, 36.8. FT-IR: ν / cm^{-1} = 3415, 3300, 3192, 2895, 2652, 2518, 2096, 1707, 1534, 1490, 1418, 1213, 1055, 834, 704. $[\alpha]_D^{20}$ = −77 (10 mg/mL; MeOH). mp: 252 °C (decomposition).

tert-Butyl ((*S*)-4-methyl-1-oxo-1-(((*R*)-1-oxo-3-phenyl-1-(pyrazin-2-ylamino)propan-2-yl)amino)-pentan-2-yl)carbamate (30)
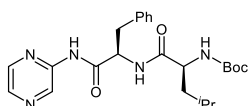
The desired compound was prepared from Boc-L-Leu-OH (85 mg, 0.37 mmol, 1.0 equiv.), HOBt • H_2O (56 mg, 0.37 mmol, 1.0 equiv.), 2,4,6-collidine (97 μL, 0.37 mmol, 1.0 equiv.), TBTU (118 mg, 0.37 mmol, 1.0 equiv.), and **29** (102 mg, 0.37 mmol, 1.0 equiv.) in DCM (8 mL) according to general procedure A, followed by purification by column chromatography on silica (cyclohexane/ethyl acetate = 2:1) to afford the desired product as a colorless foam (110 mg, 0.24 mmol, 65%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 10.16–9.99 (m, 1H), 9.50 (d, *J* = 1.4 Hz, 1H), 8.34–8.23 (m, 2H), 7.24–7.13 (m, 5H), 6.25–6.04 (m, 1H), 5.54–5.39 (m, 1H), 4.34–4.14 (m, 1H), 3.32–3.06 (m, 2H), 1.59–1.32 (m, 12H), 0.95–0.82 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 173.6, 170.4, 155.9, 148.2, 142.2, 140.1, 137.3, 136.0, 129.5, 128.7, 127.2, 80.2, 54.4, 53.5, 41.9, 39.0, 28.4, 24.8, 22.9, 22.5. FT-IR: ν / cm^{-1} = 3267, 3060, 2957, 2870, 1647, 1541, 1412, 1366, 1295, 1165, 1013, 910, 843, 732, 699. $[\alpha]_D^{20}$ = −26 (10 mg/mL; CHCl$_3$). mp: 81–83 °C. R$_f$: 0.19 (cyclohexane/ethyl acetate = 2:1).

(*S*)-2-Acrylamido-4-methyl-*N*-((*R*)-1-oxo-3-phenyl-1-(pyrazin-2-ylamino)propan-2-yl)pentanamide (1h)
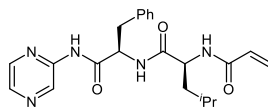
Compound **30** (110 mg, 0.24 mmol) was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C, the residue was taken up in DCM (4 mL), and NEt$_3$ (66 μL, 0.48 mmol, 2.0 equiv.) was added. After the mixture was cooled down to 0 °C, acryloyl chloride (20 μL, 0.24 mmol, 1.0 equiv.) was added dropwise over a period of 1 h. The mixture was allowed to stir for an additional 2 h, and the solvent was removed under reduced pressure at 40 °C. Purification by column chromatography on silica (DCM/MeOH = 30:1) yielded the desired product as a colorless solid (50 mg, 0.12 mmol, 51%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.83 (s, 1H), 9.42 (s, 1H), 8.34–8.23 (m, 1H), 8.14 (s, 1H), 7.70 (d, J = 8.6 Hz, 1H), 7.25–7.11 (m, 6H), 6.25 (d, J = 16.9 Hz, 1H), 6.02 (dd, J = 16.9, 10.2 Hz, 1H), 5.54 (d, J = 10.2 Hz, 1H), 5.37–5.23 (m, 1H), 4.75–4.60 (m, 1H), 3.35–3.23 (m, 1H), 3.17–3.03 (m, 1H), 1.55–1.38 (m, 3H), 0.90–0.78 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 173.2, 170.3, 165.9, 148.2, 142.0, 140.2, 137.4, 136.2, 130.3, 129.4, 128.7, 127.7, 127.2, 54.9, 52.2, 41.4, 38.5, 24.9, 22.8, 22.5. FT-IR: ν / cm^{-1} = 3262, 3056, 2957, 2930, 2870, 1650, 1537, 1409, 1295, 1062, 1011, 982, 842, 746, 698. $[\alpha]_D^{20}$ = −16 (10 mg/mL; CHCl$_3$). mp: 88–90 °C. ESI-MS: *m/z* calculated for [C$_{22}$H$_{27}$N$_5$O$_3$+H]$^+$ ([M+H]$^+$) = 410.2, found: 410.2. Purity: 99% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 $\rightarrow$ 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 4.05 min). R$_f$: 0.23 (DCM/MeOH = 30:1).

tert-Butyl (6-methyl-3-oxohept-1-en-4-yl)carbamate (31)
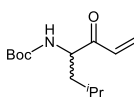
To a solution of compound **5** (265 mg, 1.15 mmol, 1.0 equiv.) in dry THF (7 mL) was added vinylmagnesium bromide (0.7 M in THF, 4.91 mL, 3.44 mmol, 3.0 equiv.) dropwise at −78 °C under argon atmosphere. After stirring at −78 °C for 30 min, the solution was allowed to warm up to 0 °C, and 1 M HCl was slowly added. The mixture was extracted with ethyl acetate (3 × 50 mL), and the combined organic extracts were filtrated over silica. The solvent was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography on silica (cyclohexane/acetone = 20:1) to yield the desired product as a colorless oil (137 mg, 0.57 mmol, 50%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 6.54–6.29 (m, 2H), 5.85 (dd, J = 9.8, 2.0 Hz, 1H), 5.19–5.03 (m, 1H), 4.70–4.54 (m, 1H), 1.79–1.66 (m, 1H), 1.58–1.46 (m, 1H), 1.41 (s, 9H), 1.39–1.28 (m, 1H), 1.01–0.87 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 199.4, 155.7, 133.5, 129.9, 79.8, 55.8, 41.6, 28.4, 25.0, 23.4, 22.0. FT-IR: ν / cm^{-1} =

3333, 2959, 2933, 2872, 1693, 1505, 1455, 1392, 1367, 1250, 1161, 1045, 1022, 848, 777. $[\alpha]_D^{20} = +13$ (10 mg/mL; MeOH). R_f: 0.33 (cyclohexane/acetone = 20:1).

Ethyl (*E*)-5-((*tert*-butoxycarbonyl)amino)-7-methyl-4-oxooct-2-enoate (1l)
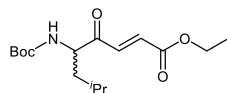
A solution of compound **31** (49 mg, 0.20 mmol, 1.0 equiv.), ethyl acrylate (221 μL, 2.03 mmol, 1.0 equiv.), and Hoveyda Grubbs II catalyst (6 mg, 0.01 mmol, 0.05 equiv.) in dry DCM (7 mL) was heated under reflux for 4 h under argon atmosphere. The solvent was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography on silica (cyclohexane/acetone = 20:1) to yield the desired product as a colorless oil (23 mg, 0.07 mmol, 37%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.20 (d, J = 15.8 Hz, 1H), 6.80 (d, J = 15.8 Hz, 1H), 5.17–4.94 (m, 1H), 4.71–4.50 (m, 1H), 4.27 (q, J = 7.1 Hz, 2H), 1.81–1.67 (m, 1H), 1.43 (s, 9H), 1.32 (t, J = 7.1 Hz, 3H), 1.02–0.91 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 198.8, 165.4, 155.7, 136.5, 132.4, 80.2, 61.6, 57.4, 28.4, 25.1, 23.4, 21.9, 14.3. FT-IR: ν / cm^{-1} = 3365, 2958, 2928, 2872, 1704, 1507, 1392, 1367, 1302, 1168, 1025, 981, 873, 784, 668. $[\alpha]_D^{20} = -12$ (10 mg/mL; CHCl$_3$). ESI-MS: *m/z* calculated for $[C_{16}H_{27}NO_5+Na]^+$ ([M+Na]$^+$) = 336.2, found: 336.2. Purity: 95% (HPLC, 254 nm, MeCN/H$_2$O + 0.1% HCOOH = 10:90 $\rightarrow$ 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 5.47 min). R_f: 0.44 (cyclohexane/ethyl acetate = 5:1).

tert-Butyl 2-isobutylhydrazine-1-carboxylate (33)
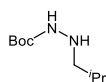
To a solution of *tert*-butyl carbazate (500 mg, 3.78 mmol, 1.0 equiv.) in toluene (8 mL) was added isobutyraldehyde (380 μL, 4.16 mmol, 1.1 equiv.) dropwise. After heating at 50 °C for 1 h, the solution was allowed to cool down to room temperature and palladium (10% on activated charcoal; 50 mg, 10 wt%) was added. The suspension was stirred under hydrogen atmosphere in an autoclave (3.0 bar, air inside the autoclave was removed under reduced pressure) for 20 h at room temperature and filtered over Celite®. The solvent was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography on silica (cyclohexane/ethyl acetate = 9:1) to afford the desired product as a colorless oil (264 mg, 1.40 mmol, 37% over two steps). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 2.62 (d, J = 6.8 Hz, 2H), 1.77–1.62 (m, 1H), 1.42 (s, 9H), 0.88 (d, J = 6.8 Hz, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 156.9, 80.4, 60.0, 28.4, 26.9, 20.6. FT-IR: ν / cm^{-1} = 3253, 2979, 2956, 2869, 2829, 1699, 1536, 1470, 1364, 1268, 1172, 1145, 1121, 1071, 855. R_f: 0.26 (cyclohexane/ethyl acetate = 5:1).

tert-Butyl (E)-2-(4-ethoxy-4-oxobut-2-enoyl)-2-isobutylhydrazine-1-carboxylate (1m)
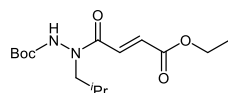
To a solution of compound **33** (225 mg, 1.20 mmol, 1.0 equiv.) in DCM (5 mL) was added NEt_3 (166 μL, 1.20 mmol, 1.0 equiv.) and ethyl fumaroyl chloride (159 μL, 1.20 mmol, 1.0 equiv.). After stirring at room temperature for 1 h, the solvent was removed under reduced pressure at 40 °C. The residue was purified by column chromatography on silica (cyclohexane/ethyl acetate = 5:1) to afford the desired product as a colorless solid (258 mg, 0.82 mmol, 68%). ^{1}H NMR (300 MHz, $CDCl_3$): δ / ppm = 7.41 (d, J = 15.5 Hz, 1H), 6.92–6.65 (m, 2H), 4.23 (q, J = 7.1 Hz, 2H), 3.95–2.94 (m, 2H), 2.02–1.85 (m, 1H), 1.46 (s, 9H), 1.29 (t, J = 7.1 Hz, 3H), 0.93 (d, J = 6.7 Hz, 6H). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ / ppm = 167.3, 165.7, 154.1, 132.7, 132.4, 82.7, 61.2, 55.6, 28.3, 26.4, 20.2, 14.3. FT-IR: ν / cm^{-1} = 3207, 2977, 2935, 2873, 1721, 1626, 1521, 1421, 1366, 1267, 1158, 1139, 1041, 981, 754. mp: 89–91 °C. ESI-MS: m/z calculated for $[C_{15}H_{26}N_2O_5+Na]^+$ ($[M+Na]^+$) = 337.2, found: 337.0. Purity: 96% (HPLC, 254 nm, $MeCN/H_2O$ + 0.1% HCOOH = 10:90 → 90:10 over 6 min, isocratic 90:10 for 2 min, t_R = 5.79 min). R_f: 0.43 (cyclohexane/ethyl acetate = 4:1).
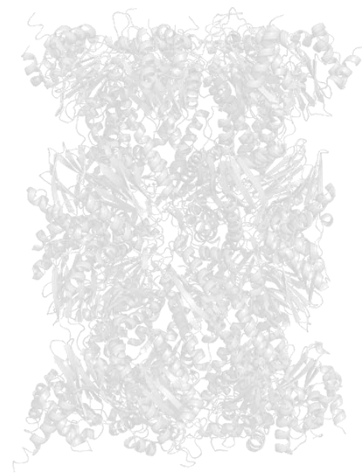
Part III

► **Investigation of the compatibility between warheads and peptidomimetic sequences of protease inhibitors: Bortezomib congeners with diverse warheads as proteasome inhibitors**

7 Bortezomib congeners with diverse warheads as proteasome inhibitors

7.1 Introduction

7.1.1 The ubiquitin-proteasome-system (UPS)

The ubiquitin-proteasome-system (UPS) represents a key element in protein homeostasis as it is responsible for approximately 80% of the protein degradation in human cells.[360–362] Therefore, it affects the regulation of most cellular processes, such as apoptosis, cell cycle progression, DNA repair, and ensures rapid degradation of misfolded and regulatory proteins.[363–365] The degradation process is initiated by covalent tagging of proteins with ubiquitin (Ub) at lysine residues, involving three enzymes. In the initial step, requiring ATP, the ubiquitin-activating enzyme (E1) binds to ubiquitin,[366] which is subsequently transferred to the ubiquitin-conjugating enzyme (E2).[367] Finally, specific ubiquitin ligases (E3) catalyze the transfer of ubiquitin to the substrate with the formation of isopeptide bonds at lysine residues.[368,369] In further progress, more ubiquitin units are attached by the three enzymes to form a polyubiquitin chain.[369] It is suggested that at least four ubiquitin units are required to be recognized by the proteasome.[370] However, the proteasomal proteolytic signal is far more complex as different chain variations can occur. It has also been found that the proteasome can recognize a single ubiquitin moiety.[371]
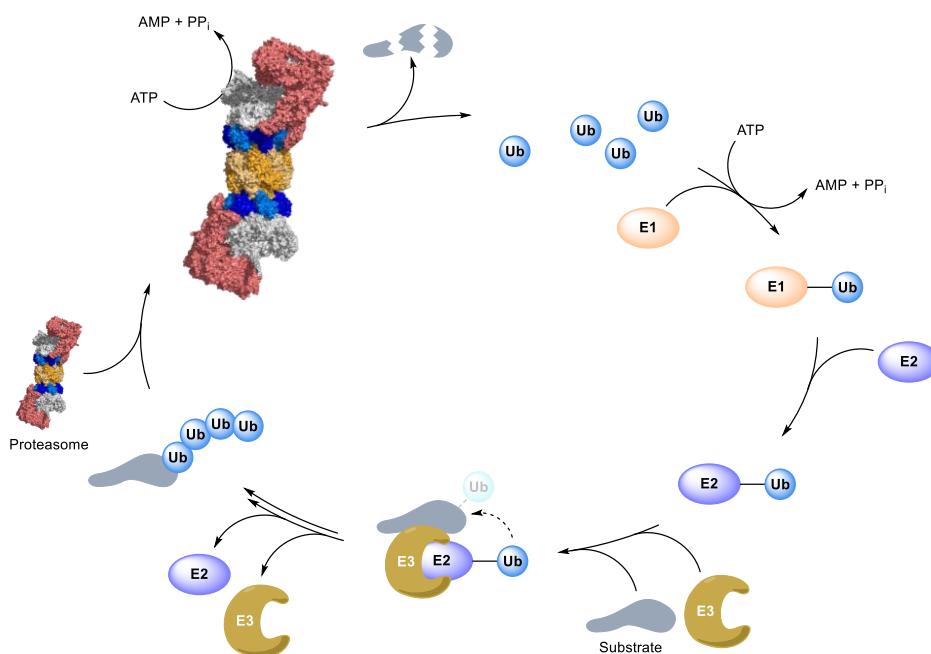
Scheme 35: Process of protein degradation by the ubiquitin-proteasome-system (UPS).

7.1.2 Structure and assembly

In mammals, different forms of the proteasome are known. It can exist as the free core particle (20S proteasome, ca. 718 kDa)[372] or be combined with two 19S regulatory caps (ca. 925 kDa) at both ends (26S proteasome, ca. 2568 kDa) (**Figure 33**).[373–375] The central core consists of four stacked rings arranged as complexes of seven subunits each (α_{1-7}, β_{1-7}, β_{1-7}, α_{1-7}), of which three subunits of each β-ring (β_1, β_2, and β_5) contain the active sites.[373–375] Depending on their substrate specificity, they are categorized as caspase-like (β_1, preference for acidic residues),[376] trypsin-like (β_2, preference for basic residues), or chymotrypsin-like (β_5, preference for hydrophobic residues).[377] However, this categorization may not reflect the true nature of proteasome activity as it appears to be more complex.[378]
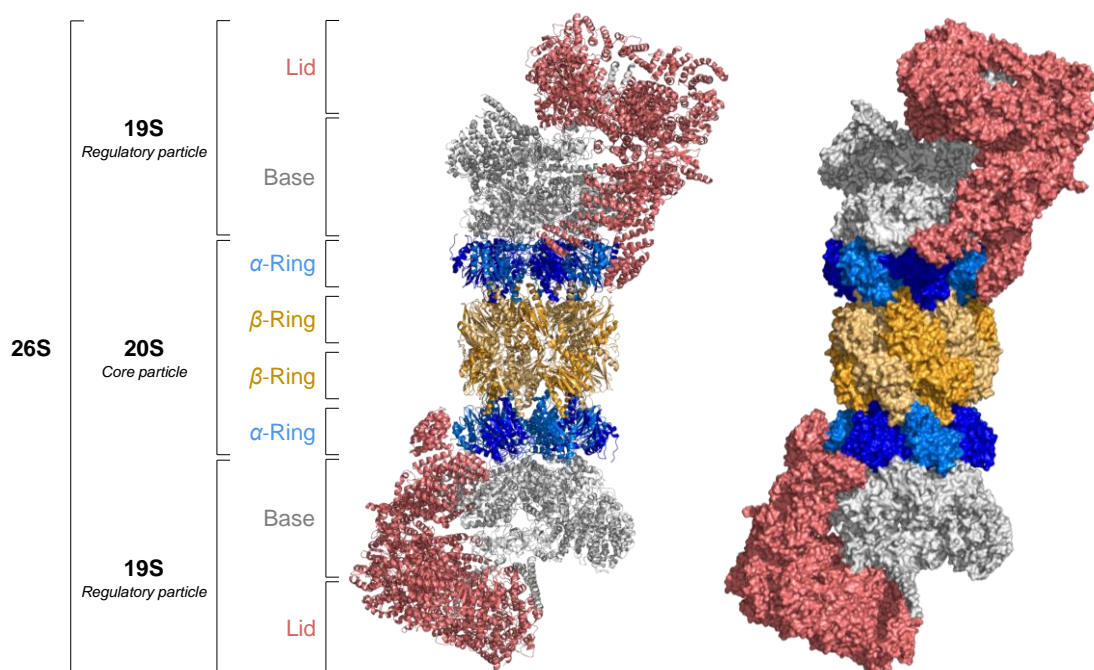
Figure 33: Assembly of the human 26S proteasome (PDB-ID: 5GJR).[374] Left: Cartoon model. Right: surface model.

7.1.3 Catalytic mechanism

The N-terminal Thr1 represents the catalytically active nucleophile that is suggested to be part of a catalytic triad including Lys33 and Asp17.[373,379] Regardless of their substrate specificity, all subunits cleave peptides by the same mechanism shown in **Scheme 36**.[379] The catalysis is initiated by deprotonation of Thr1OH via Lys33NH$_2$, followed by nucleophilic attack of Thr1O$^\gamma$ at the carbonyl function of the peptide bond. In this process, the positively charged Thr1NH$_3^+$ facilitates the orientation of the peptide bond and donates a proton to the emerging N-terminus of the C-terminal cleavage product. After the nucleophilic attack, a tetrahedral intermediate is formed, which subsequently collapses releasing the amine fragment. The deprotonated Thr1NH$_2$ activates a water molecule resulting in a nucleophilic attack at the ester bond, followed by release of the C-terminus, which restores the initial state of the catalytic triad.[379]
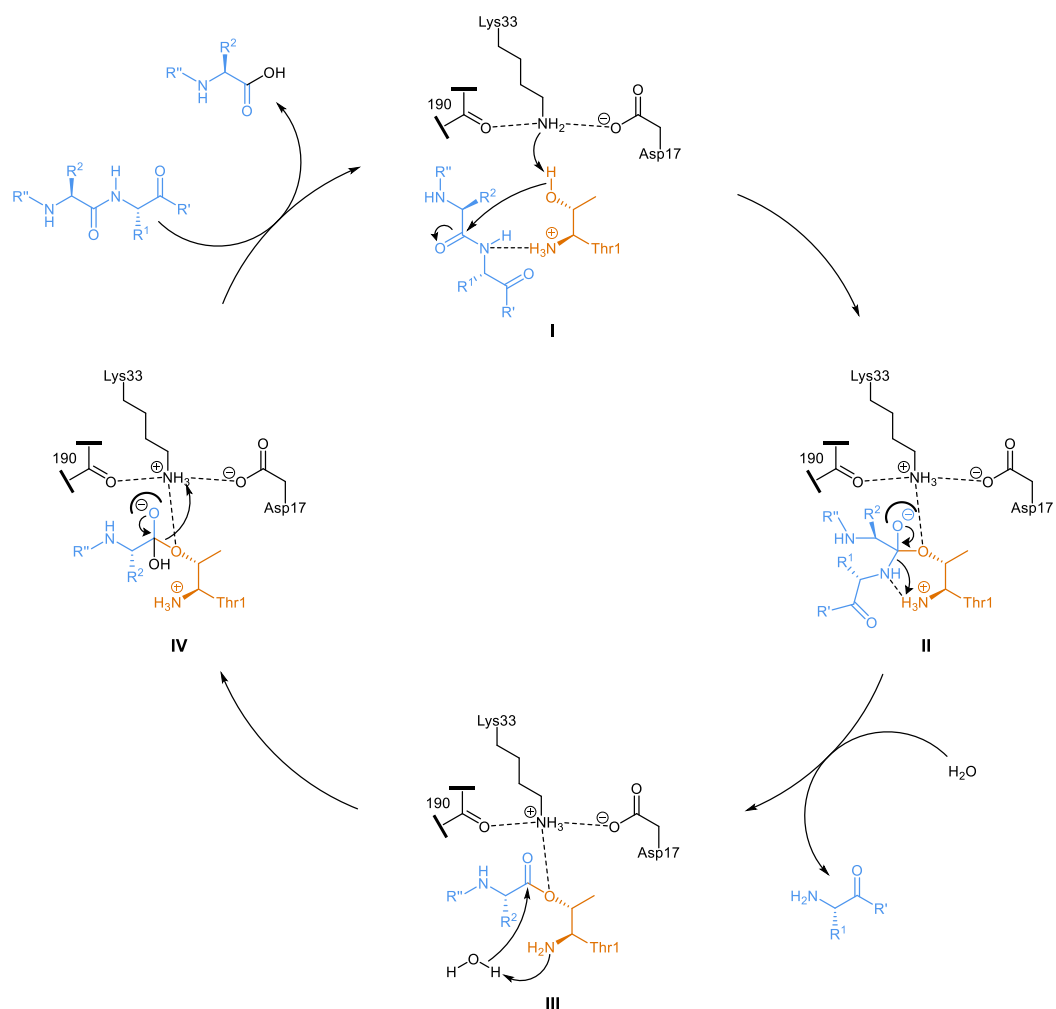
Scheme 36: Proposed mechanism of peptide hydrolysis proceeded by the catalytic triad Thr1-Lys33-Asp17 of the proteasome.[379]

7.1.4 Pathophysiological role

Besides its tasks in cellular processes mentioned before, the UPS is linked to malignancies as either deregulation or overexpression has been found in different cancer types.[380] In past decades, several anti-cancer drugs were developed,[381] such as bortezomib **7** and carfilzomib **10** for treatment of multiple myeloma.[18,19] By covalently trapping Thr1, proteasome inhibitors induce a programmed cell death due to accumulation misfolded proteins and proapoptotic factors.[382] Over the course of time, various new proteasome inhibitors with different peptidomimetic sequences and diverse types of warheads were developed exhibiting different selectivity profiles to the β-subunits.[383]

7.2 Summary and own contribution

Note: This section is about a joint project of ████████████, ██████████, ████████████, and MARVIN SCHWICKERT (████████████ group) in cooperation with ████████████ (██████ group, University of Würzburg) and the ████████ group (Universitat Jaume I, Spain). The substance numbers from the corresponding manuscript are listed *in italics*.

To achieve effective inactivation of a target of interest by covalent inhibitors, the combination of the non-covalent as well as the covalent part is important, especially for peptidomimetics. While the peptidomimetic sequence resembles the natural substrate and may ensure selectivity, the warhead is necessary for reactivity with the catalytically active nucleophile. However, it is not always clear which component has how much impact on affinity and selectivity. To analyze this, a comprehensive study was conducted to investigate the compatibility between warheads and peptidomimetic sequences of different protease inhibitors. Five targets were selected for this study: the urokinase-type plasminogen activator (uPA) as a serine,[384] the proteasome β_5-subunit as a threonine,[373] and human cathepsin S (CatS),[385] *T. brucei* rhodesain (*Tb*CatL)[351] as well as the SARS-CoV-2 main protease (M[pro])[386] as representatives of cysteine proteases. For the inhibitor design, peptidomimetic sequences were chosen according to established structures from literature that showed high affinity and selectivity or are even part of approved drugs (**Figure 34**).[18,55,385,387,388] Vinylsulfones, fluorovinylsulfones, nitroalkenes, α-ketobenzothiazoles, 4-oxoenoates, nitriles, and β-lactams were selected as warheads. Notably, some combinations could not be synthesized due to intramolecular reactivity, which was particularly true for the arginine in the uPA inhibitor sequence.

The inhibitors were tested on each target using fluorometric enzyme assays to determine affinity and selectivity. Covalent and non-covalent docking studies were conducted to explain the obtained results. In addition, a reactivity study was carried out using model compounds containing the seven different warhead types (Boc-L-Leu-warhead), which were reacted with hydroxyl and thiol model nucleophiles representing Ser, Thr, and Cys proteases to evaluate the chemoselectivity. Quantum mechanics simulations were used to analyze and confirm the experimentally determined findings.

The *in vitro* results showed that the sequence mainly determined selectivity as inhibition was achieved when using the sequence designed for the respective target. Nevertheless, the combination with a suitable warhead for the specific type of protease nucleophile was crucial. It ensured high affinity to the target or even activity in the first place as demonstrated with the bortezomib congeners: with a boronic acid (bortezomib), inhibition of the proteasome β_5-subunit was achieved, while the other tested warheads failed. An unsuitable warhead can also lead to off-target inhibition, which could be observed with bortezomib congeners with Michael acceptor- and nitrile-based warheads inhibiting CatS and rhodesain at nanomolar concentrations. The papain-like proteases CatS and rhodesain were found to be prone for

inhibitor cross reactivity, despite the design of well-defined peptidomimetic sequences, which might be due to their structural similarity. For this reason, the combination of a highly reactive warhead designed for the target nucleophile with a suitable peptidomimetic sequence can lead to potent and selective inhibitors as demonstrated with the following examples:

➢ **uPA**: uPA-inhibitor sequence + α-ketobenzothiazole (***103***)

➢ **CatS**: CatS-inhibitor sequence + nitrile (***30***)

➢ **Rhodesain**: rhodesain-inhibitor sequence + nitroalkene (***13***)

➢ **M^pro**: M^pro-inhibitor sequence + α-ketobenzothiazole (***90***)

➢ **Proteasome**: Proteasome-inhibitor sequence + boronic acid (bortezomib)
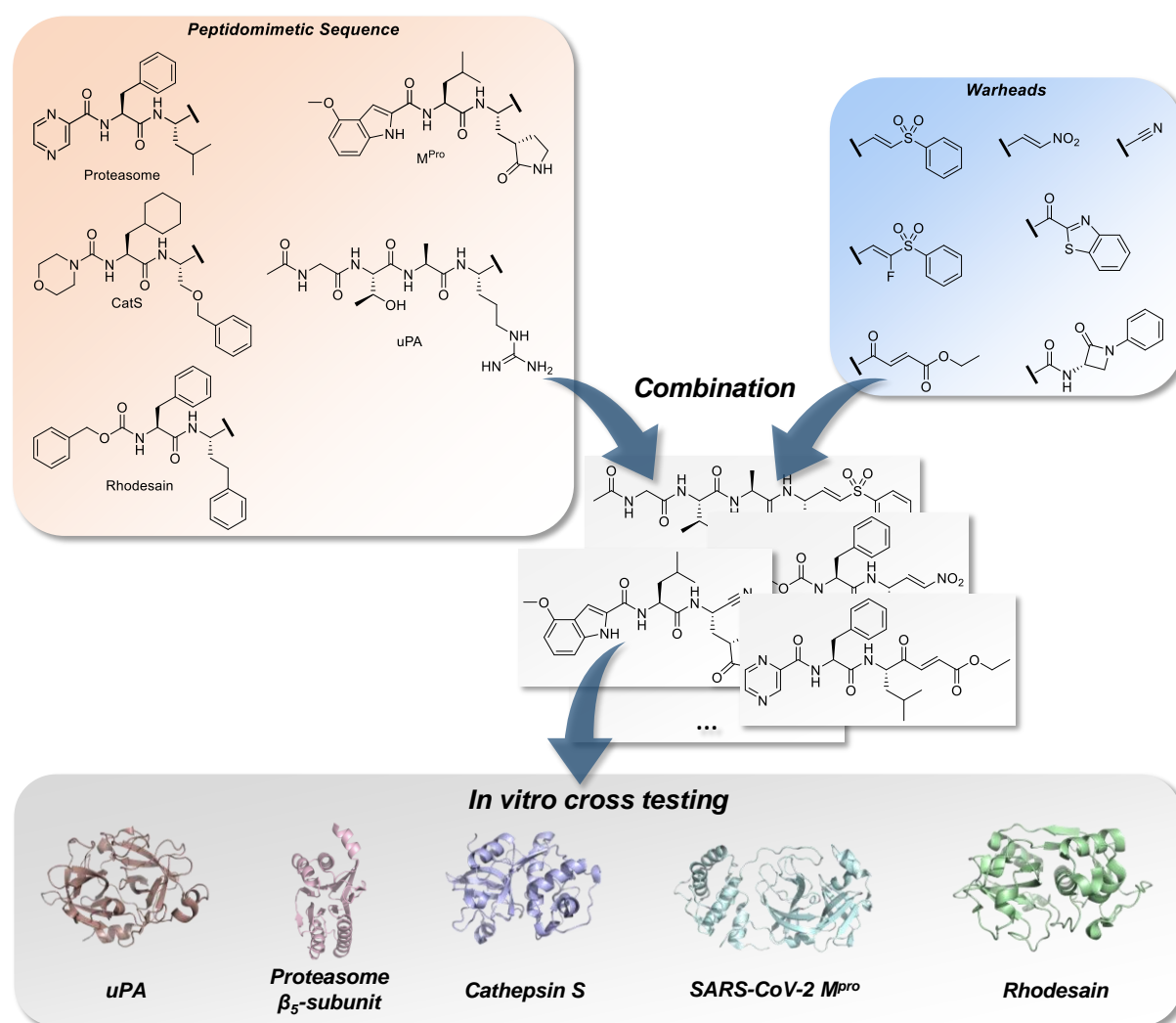
Figure 34: Combination of characteristic peptidomimetic inhibitor sequences for the targets urokinase-type plasminogen activator (uPA, PDB-ID: 1W10),[384] proteasome β_5-subunit (PDB-ID: 5LF3),[389] cathepsin S (PDB-ID: 1MS6),[385] SARS-CoV-2 main protease (M^pro, PDB-ID: 6XR3),[386] and rhodesain (PDB-ID: 2P7U)[351] with selected warheads (vinylsulfone, fluorovinylsulfone, nitroalkene, α-ketobenzothiazole, 4-oxoenoate, nitrile, and β-lactam). The resulting compounds were tested on each target to determine affinity and selectivity.

245

Non-covalent docking confirmed reasonable binding modes for all compounds, resembling interactions of the crystallographic reference ligands and peptide recognition sequences in their expected subpockets. In addition, the warheads were found in close proximity to the nucleophilic catalytic amino acids except for the *β*-lactams, which tended to be positioned in the S1' subpockets.

The reactivity tests confirmed high reactivity of the 4-oxoenoate, the fluorovinylsulfones and the nitroalkene warheads toward the deprotonated thiol nucleophile/cysteine model. Analogously to the *in vitro* studies, the *α*-ketobenzothiazole warhead was found to be a potent electrophilic trap for both cysteine and serine nucleophiles. Nevertheless, differences in reactivity of some warheads could be observed, which might be explained by different conditions used in the chemical test system and the biochemical *in vitro* studies.

Overall, this study highlighted the importance of carefully designing covalent peptidomimetic inhibitors. By using a suitable peptidomimetic sequence, high selectivity can be achieved, but it can only be ensured by combining it with a proper warhead that is not prone for off-target reactivity. Yet the warhead must have sufficient reactivity toward the target nucleophile to enable inhibition.

For a more detailed discussion, as well as the presentation of all experimental procedures, the reader is referred to the corresponding manuscript (**Section 7.3**) and Supporting Information (**Appendix**). Experimental procedures conducted or supervised by MARVIN SCHWICKERT (████████ group) are described in **Section 7.4**.

Own contribution: Synthesis of bortezomib congeners (vinylsulfone, fluorovinylsulfone, nitrile, *α*-ketobenzothiazole, *β*-lactam), writing the corresponding parts in the manuscript and Supporting Information, designing parts of the reactivity study, among others, NMR-based reactivity study of ^{13}C-labeled test compounds, synthesis of ^{13}C-labeled compound *115*, creation of **Figure 1**.

7.3   Manuscript

Article

Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study

Patrick Müller [1,†], Mergim Meta [1,†], Jan Laurenz Meidner [1,†], Marvin Schwickert [1], Jessica Meyr [2],
Kevin Schwickert [1], Christian Kersten [1], Collin Zimmer [1], Stefan Josef Hammerschmidt [1], Ariane Frey [1],
Albin Lahu [1], Sergio de la Hoz-Rodríguez [3], Laura Agost-Beltrán [3], Santiago Rodríguez [3], Kira Diemer [2],
Wilhelm Neumann [2], Florenci V. Gonzàlez [3], Bernd Engels [2] and Tanja Schirmeister [1,*]

[1]   Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz,
      Staudinger Weg 5, D-55128 Mainz, Germany
[2]   Institute of Physical and Theoretical Chemistry, Julius-Maximilians-University of Wuerzburg,
      Emil-Fischer-Straße 42 Süd, D-97074 Wuerzburg, Germany
[3]   Departament de Química Inorgànica i Orgànica, Universitat Jaume I, 12080 Castelló de la Pana, Spain
*   Correspondence: schirmei@uni-mainz.de; Tel.: +49-6131-39-25742
†   These authors contributed equally to this work.

Abstract: Covalent peptidomimetic protease inhibitors have gained a lot of attention in drug development in recent years. They are designed to covalently bind the catalytically active amino acids through electrophilic groups called warheads. Covalent inhibition has an advantage in terms of pharmacodynamic properties but can also bear toxicity risks due to non-selective off-target protein binding. Therefore, the right combination of a reactive warhead with a well-suited peptidomimetic sequence is of great importance. Herein, the selectivities of well-known warheads combined with peptidomimetic sequences suited for five different proteases were investigated, highlighting the impact of both structure parts (warhead and peptidomimetic sequence) for affinity and selectivity. Molecular docking gave insights into the predicted binding modes of the inhibitors inside the binding pockets of the different enzymes. Moreover, the warheads were investigated by NMR and LC-MS reactivity assays against serine/threonine and cysteine nucleophile models, as well as by quantum mechanics simulations.

Keywords: covalent inhibitors; in vitro study; protease inhibitors; peptidomimetic sequence; warhead; reactivity and selectivity study

check for
updates

1. Introduction

The human organism expresses about 600 different proteases falling into five different catalytic classes: aspartic, cysteine, metallo, serine and threonine proteases [1,2]. With their ability to catalyze irreversible protein hydrolysis, these members of the degradome manage the functions of many proteins through various mechanisms, such as activating or inactivating, e.g., growth factors, cytokines and other enzymes. As a result, they play an important role in physiological and developmental processes. These include DNA replication, cell proliferation and differentiation, but also tissue remodeling and neuronal outgrowth [3,4]. Due to their essential roles in such vital processes, dysregulation of these proteins causes severe pathologic conditions, such as cancer and neurodegenerative or cardiovascular disorders [5,6]. Furthermore, proteases play a key role in infectious diseases of, for example, parasitic or viral origin. African trypanosomiasis, also called sleeping sickness, and Chagas disease are caused by parasites and are classified as neglected tropical diseases and constitute important health issues in Latin American and Sub-Saharan African countries. For both diseases, proteases have been identified, which are essential for the

development of the parasites and the progression of the disease [7,8]. The 2019–2020 coronavirus (SARS-CoV-2) outbreak is the most recent example of a viral disease with global impact and burden. The viral replication and spreading is associated with proteases playing crucial roles in the viral life cycle, turning them into valid targets for the design of new anti-infectives [9,10].

Over the course of time, various protease inhibitors have been discovered either by targeted design or serendipity. Depending on the target binding site and inhibition mechanism, the molecular structures vary significantly. These range from small molecules to macrocyclic drugs and from non-covalent to covalent inhibition types [11–13]. Until recently, covalent modifiers which consist of an electrophilic trap (warhead) were controversially discussed as therapeutics due to the possibility of unselective reactions with off-target proteins and associated immunogenicity and toxicity. These compounds are emerging as potential drugs due to various inherent advantages, such as longer residence times and an accompanying lower drug dosage necessary for effective therapy [14]. There are many covalent drugs that have been approved, including some protease inhibitors, such as the proteasome inhibitors bortezomib or carfilzomib, for treatment of multiple myeloma, which inhibit the proteasome's β5-subunit in an irreversible manner, due to the permanent covalent bond to the catalytically active Thr-1. On the other hand, the nitriles saxagliptin and vildagliptin for treatment of type 2 diabetes and the recent first-approved cysteine protease inhibitor nirmatrelvir for treatment of COVID-19 bind covalent-reversibly to their target proteases, due to the decomposition of the (thio)-imidate adduct formed between the inhibitor and the amino acid of the protease, which is preferable since covalent-reversible inhibition leads to a lower risk of haptenization and binding to off-targets [15–17].

The binding of such covalent protease inhibitors proceeds in two stages. A peptidic or peptidomimetic recognition sequence is mainly responsible for the non-covalent interactions (first step) with the substrate binding pockets. It mainly determines the selectivity profile of the inhibitor towards the protease of interest, due to polar and non-polar interactions between the peptidic residues and the enzyme sub pockets. In the second step, the reaction between the warhead and an active site amino acid residue leads to the formation of a covalent bond, either reversibly or irreversibly, between the drug and the enzyme. This step mainly determines the affinity of the inhibitor to the target protease [14,18]. However, the warhead must be suitable for the respective nucleophilic amino acid residue in the active site. Depending on the type of nucleophile, different warheads can be used to target thiol or hydroxy groups of amino acid residues. Functional groups, such as β-lactams, but also boronic acids, which are all considered hard electrophiles with regard to the HSAB theory, are warheads targeting mainly serine and threonine-based proteases. Unsaturated, vinylogous Michael-acceptor-like structures, which are considered soft electrophiles, preferably react with cysteine proteases [18–21]. There are also warheads, e.g., ketones, aldehydes and nitriles, that are similarly suitable for serine-, threonine- and cysteine-based proteases [22–25]. Thus, exchanging the warhead can lead to different reactivity and affinity profiles, and alterations to the peptidomimetic/peptidic sequence may affect the selectivity of an inhibitor.

Within this extensive systematic study, we selected peptidomimetic sequences specifically to ensure a high affinity towards the protease of interest, which will be discussed below. We collected information about different kinds of warheads regarding their electrophilic properties and inhibition mechanisms to obtain a well-balanced assortment to potentially target cysteine and serine-/threonine proteases and combined them with the sequences (Figure 1) [18,21,22,24]. In vitro testing of all inhibitors on every target, first with the suited peptidomimetic sequence with differing warheads for their on-target and afterwards towards the off-target proteases, revealed the impact of the peptidomimetic sequences and the warheads on affinity and selectivity. The results indicate that, depending on the protease, every tested warhead behaved differently. The experimental results were compared with molecular docking results, visualizing putative binding modes in order to achieve a better understanding of the characteristics of the tested compounds.

Additionally, a reactivity study was carried out using model compounds containing the seven different warhead types, which were reacted with hydroxy and thiol model nucleophiles representing serine, threonine and cysteine proteases. Quantum mechanical computations of the reactions between the warheads and model nucleophiles were used to explain the experimental reactivity test data. These data highlight the preference of the warheads for specific active site residues.

To our knowledge, this is the first systematic study of this extent to evaluate the inhibition properties of peptidomimetic inhibitors with different warheads described in the literature, including in vitro testing towards a series of selected proteases, reactivity tests of the warheads in solution with model nucleophiles and in silico studies (docking and quantum mechanics and kinetic simulations) to explain the experimentally obtained data.

For our studies, the urokinase-type plasminogen activator (uPA) was chosen as a serine, the $\beta5$-subunit of the proteasome as a threonine and human cathepsin S (CatS), SARS-CoV-2 main protease (M^{pro}) and *T. brucei* rhodesain (TbCatL) as representatives of cysteine proteases.

The uPA belongs to the trypsin-like serine protease superfamily and contains a catalytic triad consisting of Ser195, His57 and Asp102 [26]. The enzyme is involved in several physiological functions, such as the degradation of the extracellular matrix (ECM), cell migration and thrombolysis [27,28]. Dysregulation of the uPA is involved in the metastasis of several cancer species [29]. We chose Ac–(L)Gly–(L)Thr–(L)Ala–(L)Arg–(warhead) as the specific peptidomimetic sequence for the uPA-inhibitors because of its high selectivity, which has been reported in the literature [30].

The 20S proteasome is responsible for most of the protein degradation in cells but can also lead to cancer by dysfunction [31]. It consists of three β-subunits ($\beta1$, $\beta2$ and $\beta5$), each containing a catalytic threonine. Here, we focus on the $\beta5$-subunit with the catalytic triad Thr1, Lys33 and Asp17, as it has the greatest impact on the proteolytic activity of the 20S proteasome. We selected the peptidomimetic sequence of bortezomib Pyz–(L)Phe–(L)Leu–(warhead) because of its clinically proven properties as a potent drug [32].

As cysteine proteases, we chose CatS, M^{pro} and rhodesain. Since CatS and rhodesain are both members of the papain family, they would allow a closer examination of the selectivity of the tested inhibitors towards related proteases [33]. CatS contains a catalytic dyad consisting of Cys25 and His164 [34]. It is partly tethered at the cell surface and involved in tissue remodeling, which can lead to cancer cell growth and spreading [35]. We utilized the peptidomimetic sequence morpholine–(L)cyAla–(L)Ser(OBn)–(warhead) which has been reported in the literature because of its described affinity and selectivity properties [36].

In contrast to the aforementioned proteases, rhodesain and M^{pro} do not originate from the human organism but play significant roles in the progression of infectious diseases. Rhodesain is essential for the development of the parasite *Trypanosoma brucei rhodesiense*, which is responsible for the sleeping sickness "Human African Trypanosomiasis". Analogously to CatS, it contains a catalytic dyad consisting of Cys25 and His159 [37]. There are various peptidomimetic sequences that have been published for rhodesain inhibitors. We decided to utilize Cbz–(L)Phe–(L)*h*Phe–(warhead), as it is a commonly used sequence with great affinity and selectivity [38]. M^{pro} originates from SARS-CoV-2 and plays a key role in the virus replication. The active site contains Cys145 and His164 as a catalytic dyad [39]. Similar to the newly published M^{pro} inhibitors, we chose 4-(OMe)-1*H*-indole–(L)Leu–3-[(3*S*)-2-oxopyrrolidin-3-yl]-(L)Ala–(warhead) as the general structure [40]. All peptidomimetic sequences and warheads are illustrated in Figure 1.
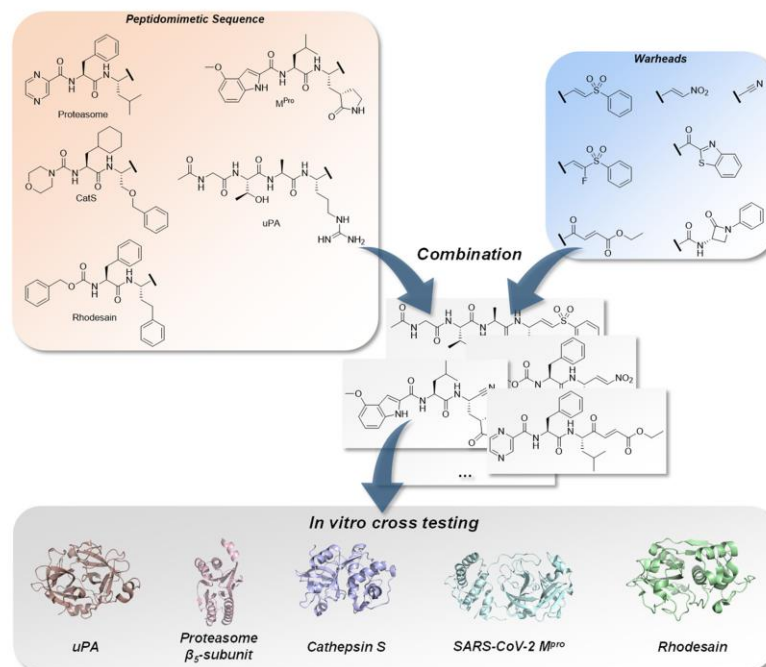
Figure 1. Combination of characteristic peptidomimetic inhibitor sequences for the targets: urokinase-type plasminogen activator (uPA), PDB-ID: 1W10 [41], proteasome $\beta 5$-subunit, PDB-ID: 5LF3 [42], cathepsin S, PDB-ID: 1MS6 [43], SARS-CoV-2 main protease (M^{pro}), PDB-ID: 6XR3 [44] and rhodesain, PDB-ID: 2P7U [45], with selected warheads (vinyl sulfone, F-vinyl sulfone, nitroalkene, α-ketobenzothiazole, 4-oxoenoate, nitrile and β-lactam). The resulting compounds were tested on each target to determine affinity and selectivity.

2. Results

2.1. Chemistry

2.1.1. Synthesis of Precursors

All tested substances were synthesized in multi-step reactions [19,21]. Regarding the synthesis of the (F-)vinyl sulfone and β-lactam compounds, the same precursor molecules were used repeatedly. The preparation of these precursors is shown in Scheme 1.
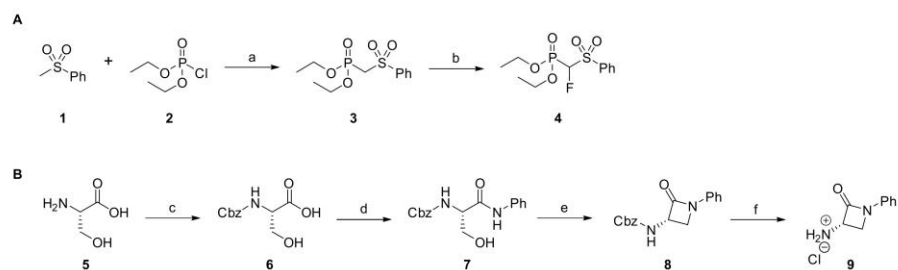
Scheme 1. Synthesis of precursor molecules. (**A**) Synthesis of phosphonate building block **4**. (**B**) Synthesis of β-Lactam building block **9**. Reaction conditions: (a) *n*-BuLi, DECP, THF, −78 °C; (b) **3**, LHMDS, Selectfluor®, THF, DMF, −78 °C, 3 h, 49%; (c) Cbz-Cl, NaHCO$_3$, H$_2$O, 12 h, rt, 90%; (d) aniline, TBTU, HOBt · H$_2$O, EtOAc, 12 h, rt, 74%; (e) ImSO$_2$, NaH, DMF, F20 °C, 1.5 h, 77%; (f) Pd/C, H$_2$, THF, 88%.

In a substitution reaction on diethyl chlorophosphate (DECP) **2** using methyl phenyl sulfone **1** and *n*-butyllithium (*n*-BuLi), the phosphonate **3** was prepared. Subsequent fluorination of **3** with Selectfluor® led to phosphonate **4**. These precursors were used for the synthesis of vinyl sulfone warheads.

The synthesis of the *β*-lactam precursor **9** was conducted from L-serine. Benzyloxy-carbonyl (Cbz) protection followed by amide coupling of the free carboxylic acid moiety with aniline led to the intermediate **7**. The following cyclisation was performed using 1,1'-sulfonyldiimidazol (ImSO$_2$) and sodium hydride (NaH). Cbz deprotection with hydrogen and palladium on carbon (Pd/C) yielded precursor **9**.

2.1.2. Rhodesain Inhibitors

The synthesis of substances with the targeting structure designed for rhodesain was conducted according to Scheme 2.
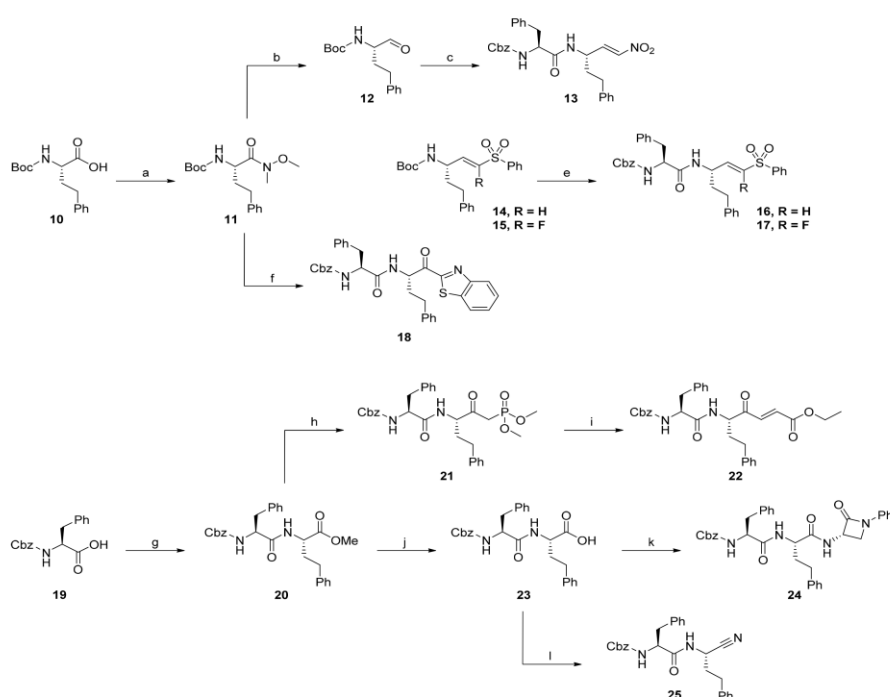
Scheme 2. Synthesis of rhodesain compounds. Reaction conditions: (a) *N,O*-dimethylhydroxylamine · HCl, DCC, HOBt · H$_2$O, DIPEA, THF, rt, 12 h, 46%; (b) LAH, THF, 0 °C, 2 h, 67%; (c) 1. MeNO$_2$, Et$_3$N, DCM, rt, 8 h, 2. TFA, DCM, rt, 0.5 h, 3. Cbz-(L)Phe-OH, EDC · HCl, TEA, DCM, rt, 12 h, 4. MsCl, DIPEA, DCM, rt, 2 h, 75%; (d) **3/4**, LiHMDS, THF, −80 °C, 12 h, 59% (**14**), 57% (**15**); (e) TFA, DCM, Cbz-(L)Phe-OH, T3P, DIPEA, DMF, rt, 12 h, 65% (**16**), 48% (**17**); (f) 1. benzothiazole, *n*-BuLi, THF, −78 °C, 3 h, 2. TFA, DCM, rt, 2 h, 3. Cbz-(L)Phe-OH, EDC · HCl, HOBt · H$_2$O, Et$_3$N, DCM, rt, 8 h, 56%; (g) H-(L)*h*Phe-OMe, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, quant.; (h) DMMP, *n*-BuLi, THF, –70 °C, 2 h, 98%; (i) ethyl glyoxylate, K$_2$CO$_3$, EtOH, rt, 2 h, 76%; (j) LiOH, THF/H$_2$O, rt, 16 h, quant.; (k) **9**, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 67%; (l) 1. EDC · HCl, HOBt · H$_2$O, NH$_4$OH, DMF, rt, 12 h, 2. TFAA, pyridine, DMF, 0 °C, 0.1 h, 40%.

The first step of the synthesis of rhodesain inhibitors was the conversion of Boc-(L)*h*Phe-OH **10** into Weinreb amide **11**. From this intermediate, the nitroalkene inhibitor **13** was accessible by reduction to aldehyde **12** and subsequent Henry reaction followed by standard deprotection and amide coupling to connect the P2-P3 residues. In a similar way, the vinyl sulfone **16** and F-vinyl sulfone **17** were obtained, whereby the aldehyde **12**

was used in a Horner–Wadsworth–Emmons (HWE) reaction with the precursors **3** and **4** followed by the attachment of the P2-P3 residues. The α-ketobenzothiazole inhibitor **18** was prepared by alkylation of the Weinreb amide **11** with benzothiazole and subsequent attachment of the P2-P3 residues. Starting from Boc-(L)*h*Phe-OH, the methyl ester **20** was prepared by amide coupling. A following alkylation with dimethyl methylphosphonate (DMMP) and HWE reaction with ethyl glyoxylate led to the 4-oxoenoate **22**. For the synthesis of β-lactam **24**, hydrolysis of methyl ester **20** and amide coupling with precursor **9** yielded the desired product. Nitrile **25** was prepared from carboxylic acid **23** via amide coupling with ammonia followed by dehydration.

2.1.3. Cathepsin S Inhibitors

Compounds designed for the inhibition of cathepsin S were synthesized according to Scheme 3.
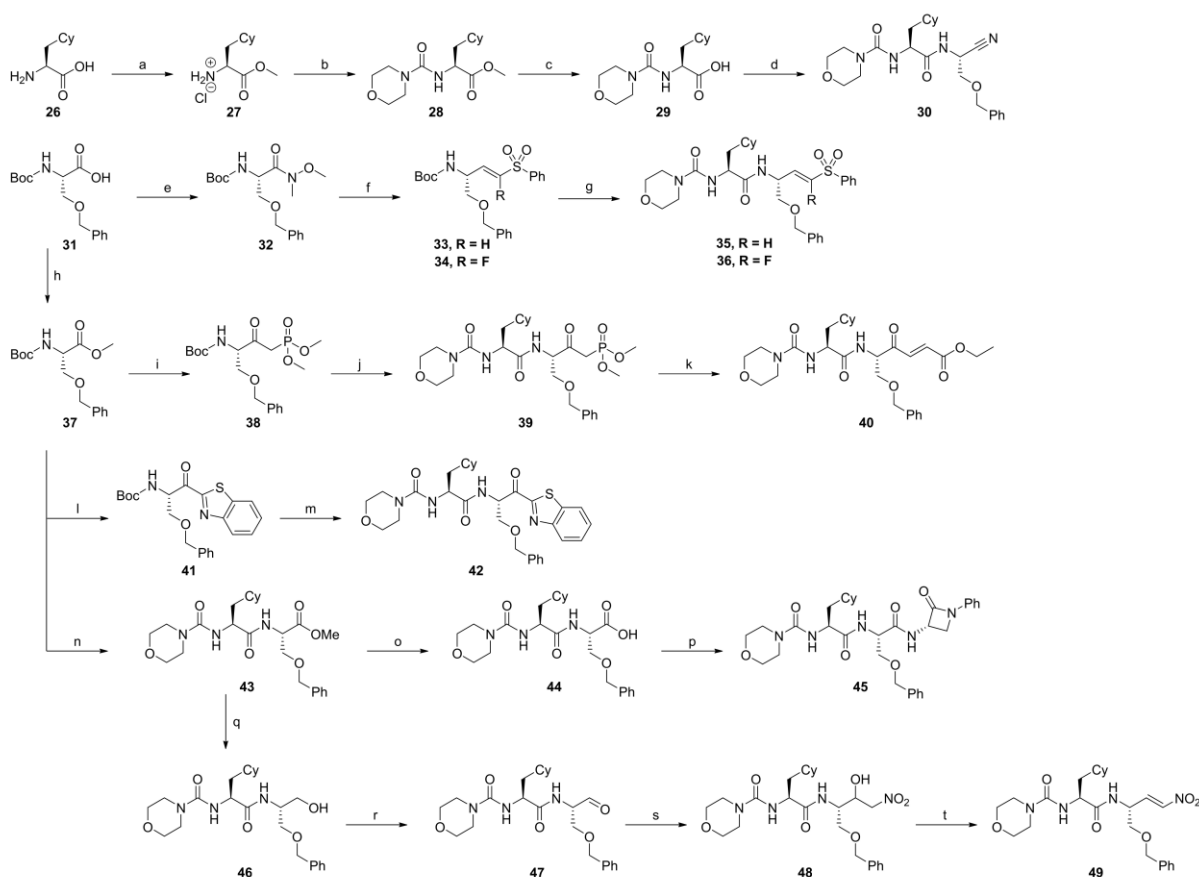
Scheme 3. Synthesis of cathepsin S compounds. Reaction conditions: (a) SOCl₂, MeOH, −10 °C, 16 h, 91%; (b) morpholine, triphosgene, NaHCO₃, CHCl₂, 0 °C; 16 h, 98%; (c) LiOH, THF/H₂O, 3 h, 97%; (d) 1. NaCN, NH₄Cl, NH₃, 2-(benzyloxy)acetaldehyde, Et₂O, 2. HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 53%; (e) *N,O*-dimethylhydroxylamine · HCl, DCC, HOBt · H₂O, DIPEA, THF, −15–0°C, 16 h, 80%; (f) 1. LAH, Et₂O, 0 °C, 2 h, 2. **3**/**4**, KHMDS/ LHMDS, THF, −78 °C, 3 h; 75% (**33**), 44% (**34**); (g) 1. 4 N HCl in 1,4-dioxane, 2. **29**, HATU, collidine, DCM/DMF, rt, 16 h, 63% (**35**), 60% (**36**); (h) MeI, DMF,

0 °C, 16 h, 97%; (i) *n*-BuLi, DMMP, THF, −78 °C, 3 h, 79%; (j) 1. 4 N HCl in 1,4-dioxane, 2. **29**, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 52%; (k) LiCl, ethyl glyoxylate, DIPEA, MeCN, 0 °C, 2 h, 39%; (l) benzothiazole, *n*-BuLi, THF, −78 °C, 3 h, 38%; (m) 1. 4 N HCl in 1,4-dioxane, 2. **29** HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 43%; (n) 1.4 N HCl in 1,4-dioxane, 2. **29**, HATU, collidine, DCM/DMF, rt, 16 h, 90%; (o) LiOH, THF/H$_2$O, rt, 3 h, 99%; (p) **9**, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 81%; (q) NaBH$_4$, MeOH, THF, 0 °C, 16 h, 91%; (r) Dess–Martin–Periodinan, DCM, rt, 16 h, 70%; (s) NaH, MeNO$_2$, THF, 0 °C, 1 h, 58%; (t) MsCl, Et$_3$N, DCM, 0 °C, 3 h, 45%.

For the synthesis of cathepsin S inhibitors, the P2–P3 intermediate **29** was used repeatedly. It was prepared by attaching a morpholino-urea residue to cyclohexyl alanine **26** followed by hydrolysis of the methyl ester. In a direct conversion from **29**, the nitrile inhibitor **30** was prepared by amide coupling with ammonia and dehydration. From Boc-(L)Ser(OBn)-OH **31**, the vinyl sulfone **35** and F-vinyl sulfone **36** were obtained by conversion into Weinreb amide **32** followed by reduction, HWE reaction with the precursors **3** and **4** and subsequent standard deprotection and amide coupling with intermediate **29**. Boc-(L)Ser(OBn)-OH **31** was also converted into the methyl ester **37**, which was used for the synthesis of the 4-oxoenoate **40**. Therefore, an alkylation with DMMP and subsequent introduction of the P2 and P3 residues by deprotection and amide coupling led to the phosphonate intermediate **39**, which was converted into the desired product by HWE reaction with ethyl glyoxylate. The α-ketobenzothiazole **42** was prepared from methyl ester **37** in an alkylation reaction with benzothiazole and attachment of the P2-P3 residues by deprotection and amide coupling with intermediate **29**. Starting from methyl ester **37**, deprotection and amide coupling with intermediate **29** led to the methyl ester intermediate **43**, which was converted into the β-lactam **45** by hydrolysis and amide coupling with precursor **9**. The nitroalkene **49** also was prepared from methyl ester **43** by firstly converting it to the alcohol **46** and then to aldehyde **47**, which was used in a Henry reaction with nitromethane and subsequent dehydration.

2.1.4. Proteasome β5-Subunit Inhibitors

Compounds designed for the inhibition of the proteasome β5-subunit were synthesized according to Scheme 4.

The synthesis of proteasome β5-subunit targeting compounds started from Boc-(L)Leu-OH **50**, which was converted into the Weinreb amide **51**. From this, the F-vinyl sulfone **55** was prepared by reduction to aldehyde **52** and subsequent HWE reaction followed by a standard deprotection and amide coupling procedure connecting the P2 and P3 residues. For the vinyl sulfone inhibitor **58**, a different route was taken. First, the Weinreb amide intermediate **56** containing the P2 and P3 residues was prepared by standard amide coupling. Subsequent reduction and HWE reaction led to the desired inhibitor. The Weinreb amide **57** was also the intermediate for nitroalkene **59**, which was prepared by reduction and Henry reaction with subsequent dehydration. From Weinreb amide **51**, the α-ketobenzothiazole moiety was introduced by alkylation. The attachment of the P2 and P3 residues by standard deprotection and amide coupling yielded the α-ketobenzothiazole **62**.

The 4-oxoenoate inhibitor **68** was prepared by HWE reaction of ethyl glyoxylate with the phosphonate intermediate **67**. The latter was synthesized by starting with the Boc protection of H-(L)Leu-OMe · HCl **63**, followed by alkylation of the methyl ester with DMMP and successive deprotection/amide coupling to introduce the P2 and P3 residues. In the same way, the introduction of the P2 and P3 residues to H-(L)Leu-OMe · HCl **63** led to the methyl ester intermediate **70**, from which the β-lactam **72** was prepared by hydrolysis and subsequent amide coupling with precursor **9**. Methyl ester **70** was also converted into the nitrile **74** by ammonolysis and dehydration.
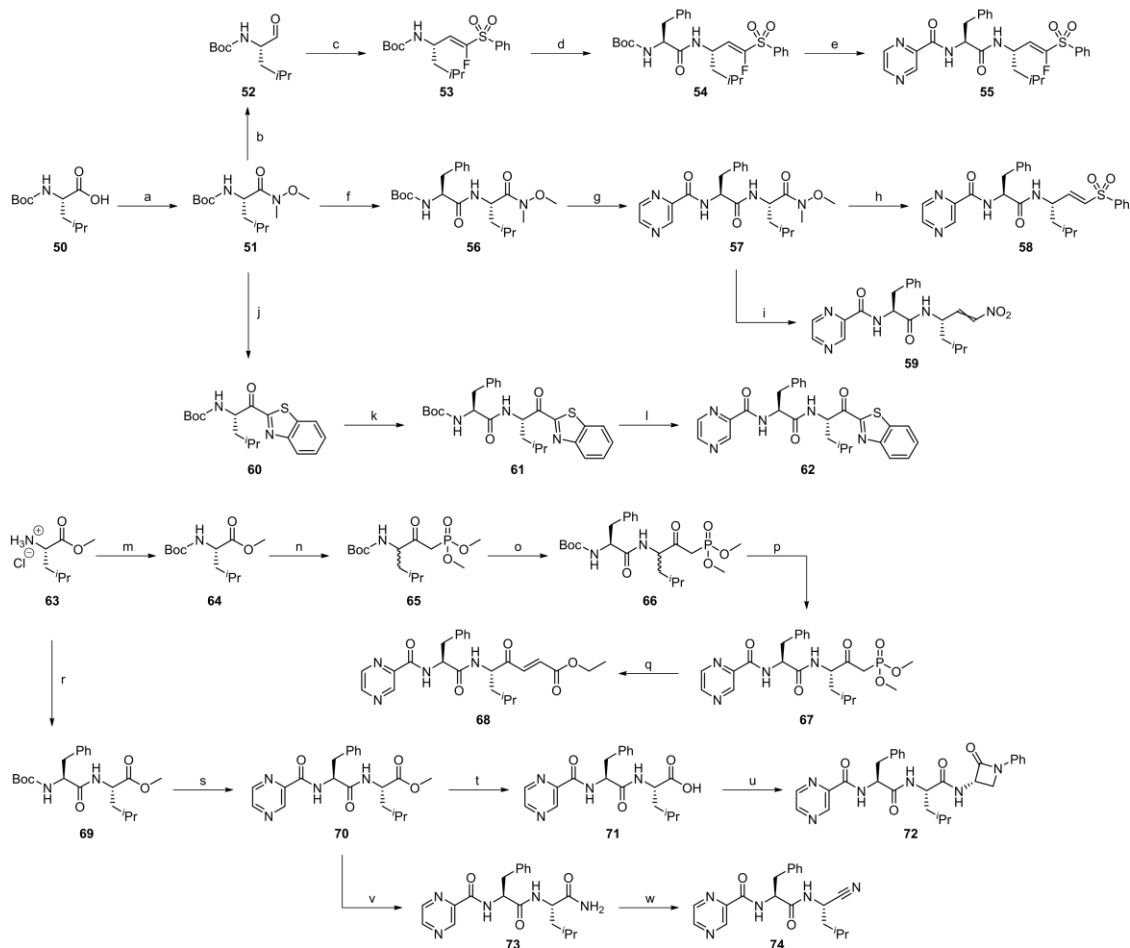
Scheme 4. Synthesis of proteasome *β*5-subunit compounds. Reaction conditions: (a) *N,O*-dimethylhydroxylamine · HCl, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 86%; (b) LAH, THF, 0 °C, 30 min, 29%; (c) **4**, LiCl, DBU, MeCN, 0 °C, 1 h, 80%; (d) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-Phe-OH, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0 °C, 16 h, 97%; (e) 1. 4 N HCl in 1,4-dioxane, rt, 1.5 h, 2. pyrazinoic acid, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0 °C, 16 h, 67%; (f) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-(L)Phe-OH, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 97%; (g) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 88%; (h) 1. LAH, THF, 0 °C, 1 h, 2. **3**, LiCl, DBU, MeCN, 0 °C, 1.5 h, 11%; (i) 1. LAH, THF, 0 °C, 1 h, 2. MeNO$_2$ Et$_3$N, DCM, 0–20 °C, 16h, 3. MsCl, DIPEA, DCM, rt, 3 h, 19%; (j) benzothiazole, *n*-BuLi, THF, –78 °C, 6 h, 65%; (k) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-(L)Phe-OH, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 68%; (l) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 53%; (m) Boc$_2$O, NaHCO$_3$, water, 1,4-dioxane, 3 h, rt, 99%; (n) *n*-BuLi, THF, –78 °C, 6 h, 88%; (o) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-(L)Phe-OH, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, DMF, 0–20 °C, 16 h, 72%; (p) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, DMF, 0–20 °C, 16 h, 66%; (q) ethyl glyoxylate, LiCl, DIPEA, MeCN, 1 h, 0 °C, 79%; (r) Boc-(L)Phe-OH, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 82%; (s) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 82%; (t) LiOH, water, THF, rt, 17 h, quant.; (u) **9**, TBTU, HOBt · H$_2$O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 65%; (v) 7 N NH$_3$ in MeOH, rt, 48 h, 89%; (w) cyanuric chloride, DMF, 0 °C, 48 h, 47%.

2.1.5. SARS-CoV-2 Mᵖʳᵒ Inhibitors

Compounds designed for the inhibition of SARS-CoV-2 Mᵖʳᵒ were synthesized according to Scheme 5.
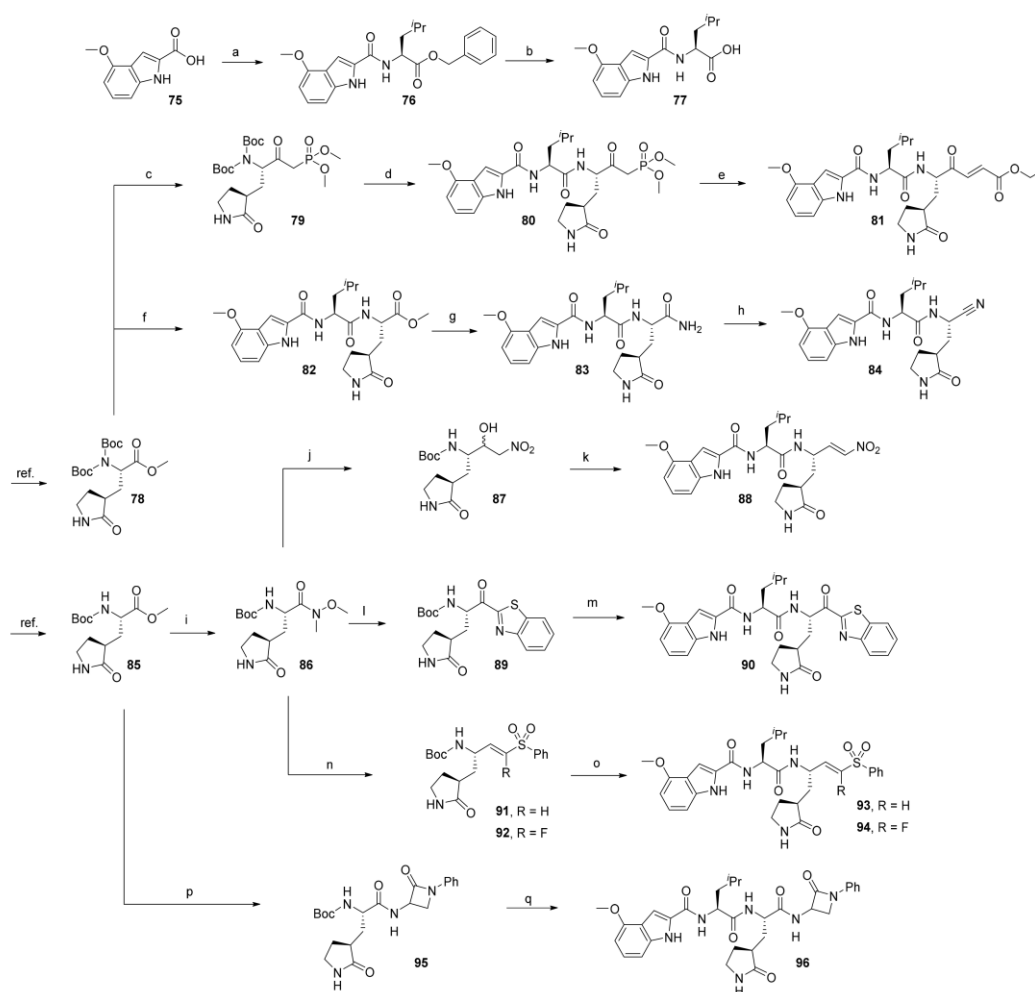
Scheme 5. Synthesis of SARS-CoV-2 Mᵖʳᵒ compounds. Reaction conditions: (a) H-(L)Leu-OBn · *p*TsOH, TBTU, DIPEA, DCM, 0–20 °C, 16 h, 93%; (b) H₂, Pd/C (10%), EtOH, quant.; (c) DMMP, *n*-BuLi, THF, −78 °C, 5 h, 24%; (d) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, −20 °C, 2 h, 40%; (e) ethyl glyoxylate, LiCl, DIPEA, MeCN, 0 °C, 2 h, 47%; (f) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, HATU, 2,4,6-collidine, DMF, 0–20 °C, 16 h, 83%; (g) 1. LiOH, THF/H₂O, 0–4 °C, 16 h, 2. NH₄OH, HATU, OxymaPure®, 2,4,6-colllidine, DMF, 0–20 °C, 16 h, 65%; (h) burgess reagent, DCM, rt, 2 h, 67%; (i) 1. LiOH, THF/MeOH/H₂O, 0–4 °C, 16 h, 2. *N,O*-dimethylhydroxylamine · HCl, HATU, 2,4,6-colllidine, DMF, 0–20 °C, 16 h, 76%; (j) 1. LAH, THF, −20 °C, 2 h, 2. MeNO₂, EtN₃, DCM, rt, 15 h, 71%; (k) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EDC · HCl, HOBt · H₂O, DIPEA, DCM, 0–20 °C, 16 h, 3. MsCl, DIPEA, DCM, 0–20 °C, 16 h, 50%; (l) benzothiazole, *n*-BuLi, THF, −78 °C, 5 h, 59%; (m) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, −20 °C, 2 h, 54%; (n) 1. LAH, THF, −20 °C, 2 h, 2. **3/4**, LiCl, DBU, MeCN, 0 °C, 2 h, 65% (**89**), 35% (**90**); (o) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, −20 °C, 2 h, 18% (**91**), 15% (**92**); (p) 1. LiOH, water, THF, 16 h, 2. **9**, TBTU, HOBt · H₂O, DIPEA, DCM, 0–20 °C, 48 h, 39%; (q) TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, −20 °C, 2 h, 80%.

Potential SARS-CoV-2 M^{pro} inhibitors were synthesized, starting from the rigidized glutamine analogs **78** and **85**, which had been prepared according to methods reported in the literature [46,47]. The P2–P3 residues fragment of the potential inhibitors was prepared by standard amide coupling with **75** and subsequent deprotection, yielding the intermediate **77**. From glutamine analog **78**, the 4-oxoenoate **81** was prepared by alkylation with DMMP and subsequent deprotection and amide coupling with **77** followed by HWE reaction with ethyl glyoxylate. Also starting from **78**, deprotection and amide coupling with **77** followed by hydrolysis and coupling with ammonia and subsequent dehydration yielded the nitrile inhibitor **84**. Starting with the preparation of Weinreb amide **86** from glutamine analog **85**, the nitroalkene **88** was accessible through reduction, a subsequent Henry reaction with nitromethane followed by dehydration and final deprotection and amide coupling with **77**. Introduction of the α-ketobenzothiazole moiety to **86** and connection of the P2–P3 residues by deprotection and coupling with **77** led to α-ketobenzothiazole **90**. Similarly, the reduction of **86** and HWE reaction with the precursors **3** and **4** and subsequent attachment of the P2–P3 residues yielded the vinyl sulfone **93** and F-vinyl sulfone **94**. For the β-lactam **96**, hydrolysis of **85** and amide coupling with precursor **9** followed by attachment of the P2–P3 residues yielded the desired product.

2.1.6. uPA Inhibitors

Compounds designed for the inhibition of the uPA were synthesized according to Scheme 6.
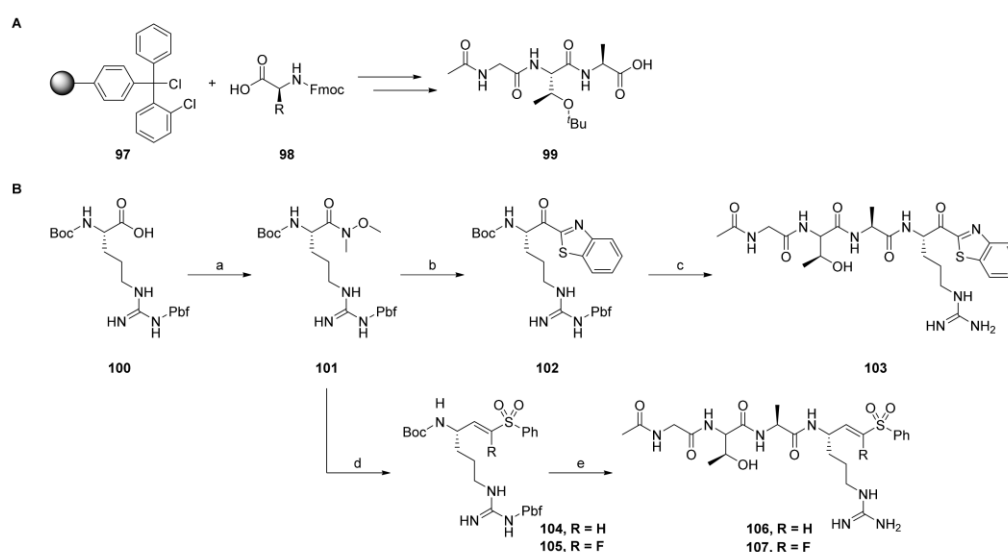
Scheme 6. Synthesis of the uPA compounds. (**A**) solid phase peptide synthesis of building block **99**. (**B**) combined synthesis of the final uPA compounds. Reaction conditions: (a) *N,O*-dimethylhydroxylamine· HCl, TBTU, DIPEA, DCM, rt, 12 h, 95%; (b) benzothiazole, *n*-BuLi, THF, −78 °C, 2 h, 76%; (c) 1. TFA, DCM, rt, 0.5 h, 2. HATU, DIPEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 10%; (d) 1. LAH, THF, 0 °C, 2. **3/4**, LiCl, DBU, MeCN, 0 °C, 1 h, 72% (**104**), 31% (**105**); (e) 1. TFA, DCM, rt, 0.5 h, 2. **99**, HATU, DIPEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 16% (**106**), 11% (**107**).

The potential uPa inhibitors are based on a peptide sequence which was synthesized via a standard Fmoc solid-phase peptide synthesis (SPPS) protocol. The obtained peptide **99** was coupled to the α-ketobenzothiazole intermediate **102**, which had been prepared from Boc- (L)Arg(Pbf)-OH **100** by alkylation of its Weinreb amide with benzothiazole to yield the α-ketobenzothiazole **103**, after deprotection. The vinyl sulfone **106** and F-vinyl sulfone **107** were prepared by reduction of Weinreb amide **101**, followed by a subsequent

HWE reaction with the precursors **104** and **105**, which were then coupled with **99** and finally deprotected. The inhibitors with the β-lactam, nitrile and 4-oxoeonoate moiety were not synthetically accessible due to the acidic conditions for the Pbf-deprotection to obtain the final inhibitors.

2.1.7. Synthesis of Reactivity Probes

Substances designed for reactivity assay were synthesized according to Scheme 7.
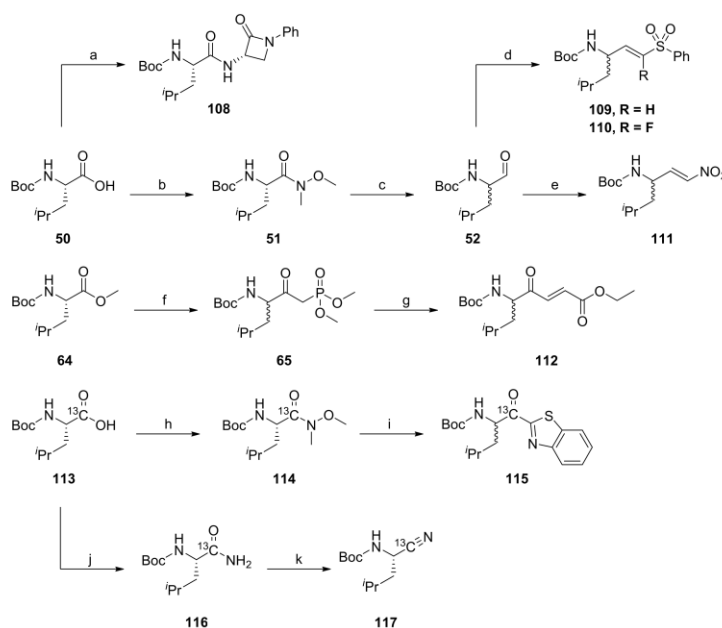
Scheme 7. Synthesis of the reactivity probes. Reaction conditions: (a) **9**, HATU, collidine, DCM, DMF, rt, 16 h, 97%; (b) TBTU, HOBt · H$_2$O, *N,O*-dimethylhydroxylamine · HCl, 2,4,6-collidine, rt, 12 h, quant.; (c) 1. LAH, Et$_2$O, 0 °C, 1 h, 89%; (d) **3/4**, LiCl, DBU, MeCN, 0 °C, 1 h, 50% (**109**), 30% (**110**); (e) 1. NaH, MeNO$_2$, THF, 0 °C, 1 h, 2. MsCl, Et$_3$N, DCM, 0 °C, 1 h, 15%; (f) DMMP, *n*-BuLi, THF, –78 °C, 1.5 h, 70%; (g) ethyl glyoxylate, LiCl, DBU, MeCN, 0 °C, 1 h, 45%; (h) TBTU, HOBt · H$_2$O, *N,O*-dimethylhydroxylamine · HCl, 2,4,6-collidine, 0–20 °C, 16 h, 95%; (i) benzothiazole, *n*-BuLi, THF, –78 °C, 3.5 h, 56%; (j) EDC · HCl, HOBt · H$_2$O, NH$_4$OH, DMF, rt, 16 h, 14%; (k) TFAA, pyridine, THF, –10 °C, 2 h, 74%.

For the synthesis of the reactivity probes, leucine was chosen as the model amino acid due to availability and to avoid side-chain reactivity. The different warheads were synthesized in the same way as described above for the full peptidic/peptidomimetic inhibitors. The β-lactam **108**, (F-)vinyl sulfone **109**, **110** and nitroalkene **111** reactivity probes were synthesized starting from Boc-(L)Leu-OH **50**, whereas the 4-oxoenoate **112** was prepared from Boc-(L)Leu-OMe **64**. Boc-(L)Leu-1-^{13}C-OH **113** was the starting material for the ^{13}C-labelled α-ketobenzothiazole **115** and nitrile **117** reactivity probes.

2.2. Reactivity Tests

To investigate the reactivity between the different warheads towards the three classes of proteases (serine, threonine and cysteine proteases), their behavior in model systems under the same reaction conditions (solvent, nucleophile and base) using either NMR or LC-MS analysis was investigated. We used reactivity probes with a Boc-L-Leu-(warhead) sequence. Leucin was chosen as a P1 amino acid to minimize influences of the side chain and due to synthetic accessibility. 2-Phenylethanethiol was used as a model nucleophile to mimic the thiol moiety of cysteine proteases, and sodium ethoxide was used as a ser-

ine/threonine replacement. DMSO-d_6 was used as solvent. Under these conditions, the nucleophile is deprotonated, simulating the activated serine or threonine in the catalytic triad of serine and threonine proteases, while ethanol as protonated alcohol species turned out to be unreactive in preliminary test reactions. The reactivity tests using 2-phenylethanthiol were carried out in the presence and absence of triethylamine as a base. This allowed for a reactivity comparison of the warheads towards protonated and deprotonated nucleophilic thiol species. Generating a deprotonated thiol species in the presence of triethylamine simulates the deprotonated cysteine in the catalytic dyad of cysteine proteases. Scheme 8 illustrates the reaction of the reactivity assay with both model nucleophiles and the vinyl sulfone moiety **109** as an example.
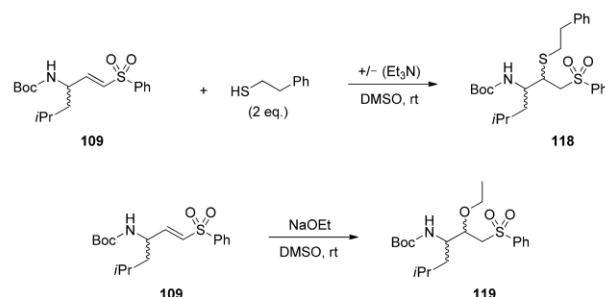
Scheme 8. Reaction scheme of the reactivity assay with both model nucleophiles under equal reaction conditions and the vinyl sulfone moiety **109**.

The reactivity tests of all Michael acceptors, **109**, **110**, **111** and **112**, the α-ketobenzothiazole **115** and nitrile **117** were investigated using an NMR-based analysis method, while the β-lactam **108** reactivity was investigated via LC-MS, due to its lack of proton signals, which could be used for evaluation of the reactivity in the ^{1}H-NMR studies, and the irreversible reaction mechanism, which allowed the LC-MS analysis. Additionally, LC-MS analyses of all reactions were performed in order to prove the formation of the expected reaction products. Formation of the expected adducts with the nitrile **117** (PhEtSH/PhEtS$^-$/EtONa), the α-ketobenzothiazole **115** (PhEtSH/PhEtS$^-$) and the nitroalkene **111** (EtONa) could not be observed. This may have been due to the covalent reversible reaction mechanism of the nitrile and α-ketobenzothiazole and the overall difficult ionization of the specific compounds by an electron spray ionization mass spectrometer.

Method A (NMR): ^{1}H-NMR spectra were recorded for the respective warhead and nucleophile mixture, before the addition of the nucleophile (0 min) and after 5, 30, 60, 120 and 240 min reaction time. For quantification, the double bond-signals (doublet/doublet of doublets, around 7.4–6.7 ppm) of the Michael acceptors were integrated relative to 1,3-dioxolane as an internal standard. The acetal CH$_2$ signal of the internal standard at 5.3 ppm was used as a reference.

The α-ketobenzothiazole **115** and nitrile **117** were similarly analyzed by ^{13}C-NMR. Therefore, the corresponding ^{13}C-leucin derivates were synthesized (Scheme 7). Quantifications of the reactions were carried out by using the integral of the carbonyl carbon atom signal at 195 ppm for the α-ketobenzothiazole and 120 ppm for the nitrile moiety. The reference signal of DMSO-d_6 was set to 39.52 ppm.

In Figure 2, the ^{1}H-NMR spectra of the test reaction of the 4-oxoenoate **112** with 2-phenylethanthiol **118** are shown exemplarily. After four hours, 92% conversion of the inhibitor to the product **120** was observed. LC-MS analysis confirmed the diastereomeric formation of the expected product **120**.
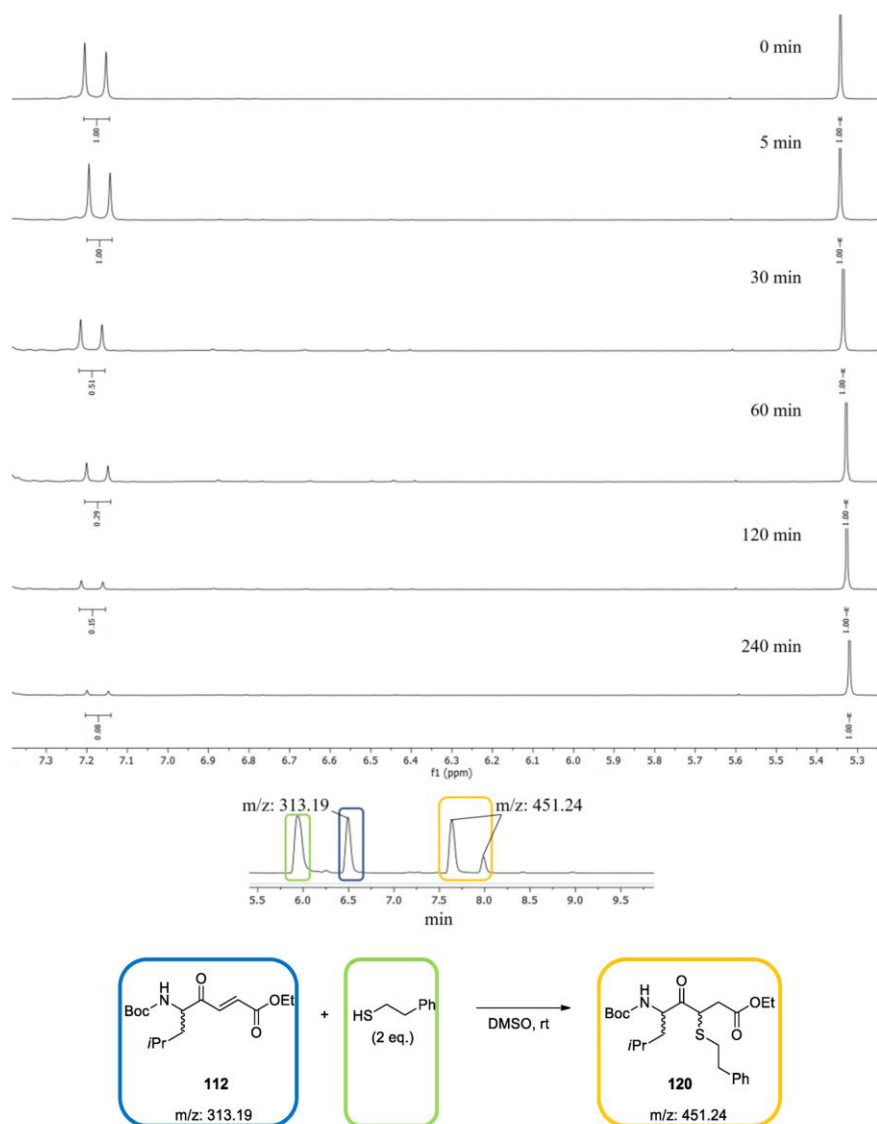
Figure 2. ^{1}H-NMR spectra of the 4-oxoenoate **112** before the addition (0 min) and after 5, 30, 60, 120 and 240 min reaction time with 2-phenylethanthiol in the absence of triethylamine. The integrals of the β-proton of the double bond at 7.17 ppm in relation to the 2-CH$_2$ signal of the internal standard 1,3-dioxolane at 5.3 ppm are given. LC-MS analysis of the same reaction at 30 min.

Method B (LC-MS): The reactivity of the β-lactam test compound **108** with the nucleophiles was investigated using an LC-MS-based method. To quantify the conversion, the AUCs were determined at 254 nm. In Figure 3, the UV spectra of the test reaction of the β-lactam **108** with EtONa are shown exemplarily. After one hour, complete conversion of the inhibitor to the adduct **122** was observed.
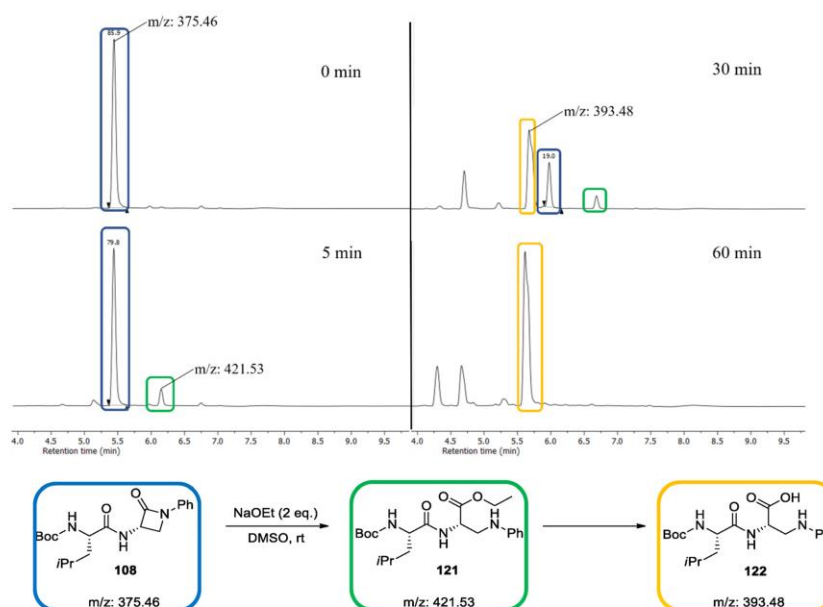
Figure 3. LC-MS spectra of the *β*-lactam **108** before (0 min) the addition and after 5, 30 and 60 min reaction time with sodium ethoxide.

All ^{1}H-NMR/^{13}C-NMR spectra and chromatograms of the reactivity tests are presented in the Supporting Information (Figures S1–S18). The reactivity test results of all warhead compounds with PhEtSH, PhEtSH + Et$_3$N and EtONa are shown in Figure 4.
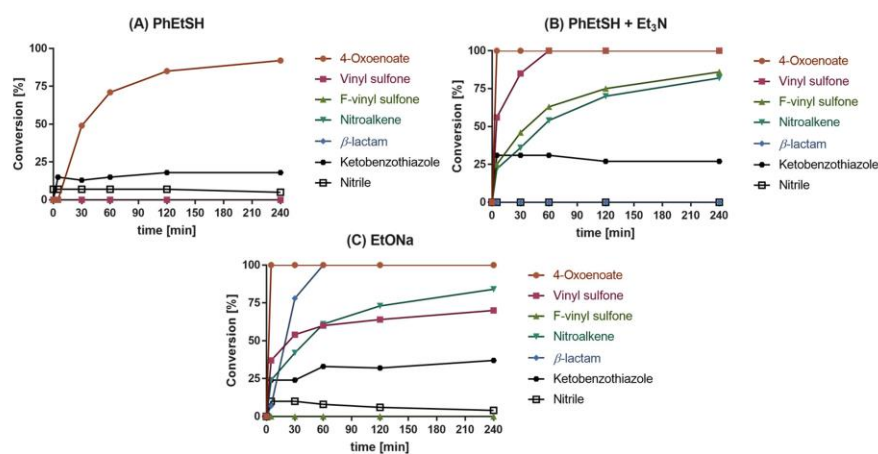
Figure 4. Progress curves of the reactions of the different warhead compounds with the model nucleophiles PhEtSH, PhEtSH + Et$_3$N and EtONa as measured by NMR and LC-MS analysis.

As depicted in Figure 4A, the 4-oxoenoate **112**, the *α*-ketobenzothiazole **115** and the nitrile **117** moiety did indeed react with PhEtSH under non-basic conditions. In contrast, conversion was not observed with the vinyl sulfone **109**, F-vinyl sulfone **110**, *β*-lactam **108** and nitroalkene **111** warheads. After 240 min, the 4-oxoenoate **112** had nearly completely (92%) reacted with PhEtSH, while only 18% conversion of the *α*-ketobenzothiazole **115** was observed. The equilibrium of the *α*-ketobenzothiazole **116** was reached after 5 min. Similarly, with the nitrile moiety **117**, only 7% conversion was detected, indicating that the

formed thioimidate adduct is relatively unstable (Figure 4A). With the addition of Et$_3$N (Figure 4B) the overall reactivity increased. Every warhead except the β-lactam **108** and the nitrile **117** reacted with the deprotonated thiol species. Full conversion of the 4-oxoenoate **112** could be observed after 5 min, followed by the vinyl sulfone **109**, which took 60 min for complete reaction. The F-vinyl sulfone **110** and nitroalkene **111** both showed similar reactivity with the thiolate species, with a maximum conversion of 86% and 82% after 240 min, respectively. The α-ketobenzothiazole **115** also showed an increased reactivity, with around 30% conversion. The reactivity tests with EtONa as nucleophile revealed the 4-oxoenoate **112** moiety as the most reactive warhead, which was completely consumed after 5 min (Figure 4C). However, LC-MS analysis did not prove the formation of the expected product but rather unspecific conversion of **112** (see Supplementary Materials, Figure S8). The nitroalkene **111** showed a similar behavior in comparison to the reactivity test with the deprotonated thiolate species, with a conversion of 84% after 240 min. The β-lactam **108** compound showed full conversion after 60 min. In contrast to the deprotonated PhEtSH species, the results indicated a much slower reactivity of the vinyl sulfone **109** with a conversion of 70% after 240 min. No conversion with EtONa was observed for the F-vinyl sulfone **110**. The α-ketobenzothiazole **115** showed a higher conversion in the presence of EtONa (37%) than with PhEtSH, but reached this maximum only after 60 min, showing a slower reaction rate compared to the deprotonated thiol species at 5 min. The equilibrium between the α-ketobenzothiazole **115** and hemiacetal shifted to 37% conversion and was higher compared to the reactivity test with the thiol nucleophiles. The nitrile **117** showed a similar conversion at 5 min with EtONa compared to the protonated thiol species, with 10% conversion, but again decreased after a period of time, which again indicates the instability of the imidate adduct under basic conditions.

The high reactivity of the 4-oxoenoate **112** warhead with the thiolate is in accordance with the high inhibitory potency of dipeptidyl 4-oxoenoate-based compounds against cysteine protease [48]. The missing reactivity of both vinyl sulfones **109** and **110** toward protonated thiol species and the high reactivity with deprotonated thiols are also in agreement with the high activity of vinyl sulfone inhibitors against cysteine proteases with a thiolate residue in the catalytic center, as reported in the literature.

Nitroalkenes are classified as cysteine targeting warheads, which is also confirmed by the observed reactivity with the model thiolate nucleophile [49].

β-lactams are commonly known as warheads in antibacterial agents with transpeptidase-inhibiting properties but have also been used in the development of serine protease inhibitors [21,50,51]. The reactivity tests demonstrate the preference for alcoholate-based nucleophiles, since they only reacted with EtONa and not with PhEtSH/PhEtS$^-$.

α-Ketobenzothiazole derivatives are used as potent serine and cysteine protease inhibitors [52,53]. Therefore, the reactivity of the α-ketobenzothiazole **115** moiety towards all three model nucleophiles was expected. In accordance with the HSAB concept, the stability of the tetrahedral (thio)hemiacetal decreased from the hard sodium ethoxide to the soft thiol/thiolate nucleophiles (EtONa>PhEtS$^-$PhEtSH) after 240 min.

The observed reaction of the nitrile **117** with both nucleophiles (PhEtSH/EtONa) is in accordance with the well-known reactivity of nitrile-based drugs. The observed instability of the (thio)-imidate adduct might have been due to the neutral or basic reaction conditions in solution [54,55]. In contrast, the (thio)imidate adduct is stabilized by interaction with amino acid residues of the enzyme pocket [56].

2.3. Quantum Mechanics Simulations

As model nucleophiles for the QM simulations, methanethiol/ate and methanolate were used. While the formed products were identical, the warheads vinyl sulfone, F-vinyl sulfone and nitroalkene exhibited varying reactivities for PhEtSH and PhEtSH in the presence of triethylamine. Only for 4-oxoenoate and α-ketobenzothiazole was significant reactivity towards PhEtSH observed, whereas for PhEtSH + Et$_3$N, all warheads except the nitrile and the β-lactam showed reactivity (Figure 4). Since most of the reaction energies

with both MeSH and MeSH + Et$_3$N were computed to be exergonic (Figure 5A), this cannot be explained merely by thermodynamics. For instance, the experimental results do not show reactivity of the warheads vinyl sulfone, F-vinyl sulfone and nitroalkene with PhEtSH, despite a computed negative free energy of reaction. Thus, to determine whether a reaction can be expected to take place, it is important to consider the whole reaction path, including the activation barriers, which determine the kinetics. Previous calculations have revealed that MeSH is often insufficiently nucleophilic to allow a reaction to occur at room temperature [57]. A base, such as triethylamine, serves as interim storage for the thiol proton before it is transferred to the warhead. By deprotonating the thiol prior to the nucleophilic attack, the nucleophilicity of MeSH is strongly increased, thereby decreasing the associated activation barriers considerably (Figures S22C, S23B and S24A–C). Following the addition of the nucleophile, the proton is transferred back from the base to the anionic intermediate. Unlike the 4-oxoenoate, vinyl sulfone, F-vinyl sulfone and nitroalkene warheads, the β-lactam warhead does not show any reactivity with PhEtSH + Et$_3$N. The computed reaction mechanism revealed three consecutive steps to obtain the product (Figure 5C). First, the nucleophile attacks the amide carbonyl group (TS1), resulting in a tetrahedral anionic intermediate (Int1). The rate-determining step is the opening of the lactam ring in the second step (TS2). This was computed to be about 33 kcal mol^{-1} for MeSH + NEt$_3$, which is in excellent agreement with the experimental data. In the last step, the former amide nitrogen is protonated by the base to yield the final product (TS3). For the nitrile warhead, a weak reaction with PhEtSH but none with PhEtSH + triethylamine was observed experimentally, which cannot be explained by reference to the computational data. As described in the reactivity tests, this might have been due to the instability of thioimidates in basic solution.

As a result of our calculations, the difference in reactivity between PhEtSH and PhEtSH + Et$_3$N for the vinyl sulfone, F-vinyl sulfone and nitroalkene warheads was attributed to a significant reduction in activation barriers caused by proton transfer from the nucleophile to the base prior to the nucleophilic attack. We therefore investigated the reason for the reactivity of 4-oxoenoate and α-ketobenzothiazole warheads with PhEtSH in DMSO in the absence of a base. For 4-oxoenoate, a conversion of 92% was observed experimentally, which corresponds to a computed free energy of reaction of about –9 kcal mol^{-1} for the nucleophilic attack at C$_\alpha$ and about –11 kcal mol^{-1} for the addition at C$_\beta$. The reaction of PhEtSH with α-ketobenzothiazole, however, showed only about 18% conversion, and the corresponding product was computed to be 8 kcal mol^{-1} (Figure 5B) and 9 kcal mol^{-1} for the thermodynamic calculation with a bigger basis set (Figure 5A). The solvent used in the experiments was not completely free of water, and, as a result, water molecules were able to catalyze the nucleophilic attack for α-ketobenzothiazole and 4-oxoenoate, as well as the keto-enol tautomerization for the latter (Figure 5B and Figure S22A,B) [58]. Our calculations demonstrate that traces of water in the solvent can function as a base to catalyze the reaction of MeSH with the warhead. The activation barrier for α-ketobenzothiazole is reduced from more than 40 kcal mol^{-1} to roughly 25 kcal mol^{-1}, and the product energy is lowered to 1 kcal mol^{-1} (Figure 5B). Similarly, water catalyzes both the nucleophilic attack of MeSH at the 4-oxoenoate warhead and the subsequent keto-enol tautomerization, leading to a decreased activation barrier of 26 kcal mol^{-1} for the first step (TS1) and one of 20–25 kcal mol^{-1} for the second step (TS2). Additionally, the product energy is even more exergonic at –17-(–18) kcal mol^{-1} (Figure S22B). Contrary to the reaction without water, the proton does not have to be transferred directly from the thiol to the atom to be protonated. Instead, it is shuffled along a chain of water molecules. The keto-enol tautomerization is favored for the C$_\beta$-addition, but the barrier associated with the rate-determining nucleophilic attack is nearly identical (Figure S22B). Thus, it is expected that both reactions should occur in solution. For the reaction with an enzyme, the conformation of the binding pocket will likely determine at which carbon atom the nucleophilic attack will occur.
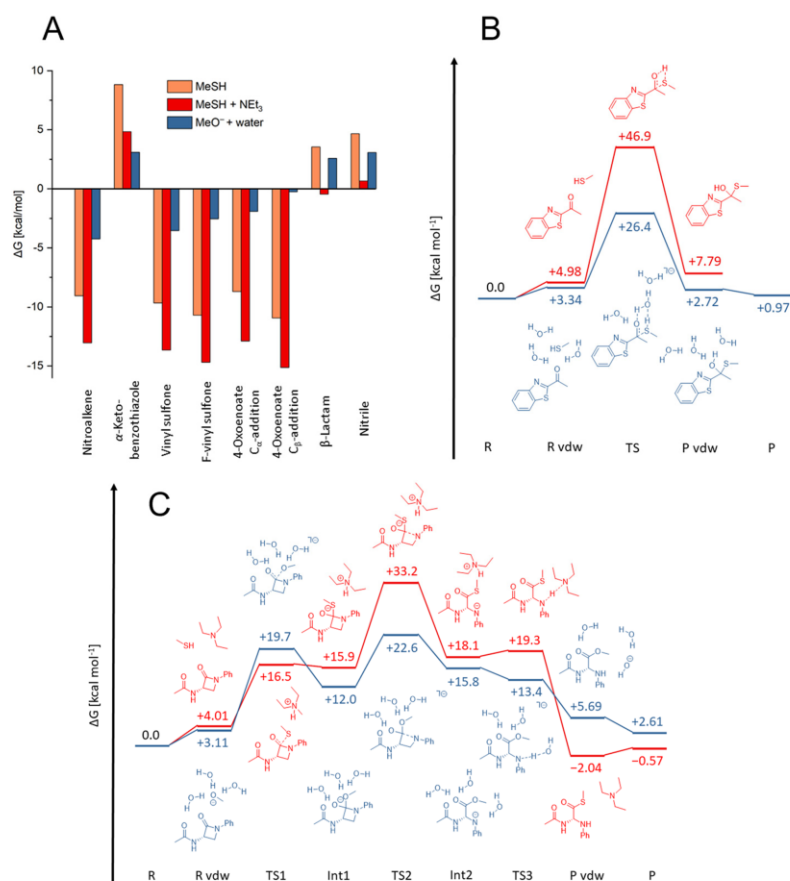
Figure 5. (**A**) Free energies of the reactions for all inhibitor warheads with MeSH, MeSH + Et$_3$N and MeO$^-$ + 3H$_2$O, computed as described in the Supplementary Materials section and Figures S24 by ωB97X−D/6−311++G**//ωB97X−D/6-31+G* calculations. (**B**) Free energy reaction paths of the α-ketobenzothiazole warhead with MeSH (red) and with MeSH in the presence of three water molecules (blue). For MeSH (red), the van der Waals complex (P vdw) and separated product molecules (P) are identical since the reaction yields only a single product molecule. For MeSH + water (blue), the energies are referenced on MeSH + 2H$_2$O and the α-ketobenzothiazole warhead + H$_2$O (R). (**C**) Free energy reaction paths of the β-lactam warhead with MeSH + NEt$_3$ (red) and with MeO$^-$ in the presence of three water molecules (blue). The reaction proceeds in three consecutive steps: first, nucleophilic attack at the amide carbonyl group (TS1); second, the opening of the lactam ring (TS2); and third, the proton transfer from the base (NEt$_3$ or H$_2$O) to the former amide nitrogen (TS3).

To mimic the reaction of the warheads with NaOEt, we calculated the reaction path with MeO$^-$ and included three water molecules to allow for protonation of the intermediates to obtain the final products and stabilize the reactive anionic species (Figures 5C, S22D, S23D and S24A–C). The reaction can either terminate at the anionic intermediate or proceed to the neutral adduct by transferring one proton from a water molecule, depending on the basicity of the intermediate, i.e., the intermediate carbanion is poorly stabilized for the vinyl sulfone, hence the reaction progresses to form the neutral addition product (Figure S24A). The nitroalkene carbanion, however, is strongly stabilized, and our calculations suggest that the reaction might stop at the intermediate (Figure S24C). Analogously, the α-ketobenzothiazole forms a deprotonated hemiacetal (Figure S23B). Experimentally, no reactivity of the F-vinyl sulfone warhead with NaOEt was observed, which was not supported by our calculations

and is contradictory to chemical intuition (Figure S24B). As previously stated, the barrier for the β-lactam ring opening in reaction with MeSH + Et$_3$N was computed to be over 30 kcal mol^{-1}, explaining the lack of reactivity in the experiments. Since the anionic species and ring opening are better stabilized in the reaction with MeO$^-$ + 3H$_2$O, only 23 kcal mol^{-1} is required in this step, which is consistent with the experimental data (Figure 5C).

2.4. In Vitro Evaluation of the Synthesized Compounds

Inhibition of the target enzymes was tested via fluorometric assays. Therefore, fluorogenic AMC- or FRET-based substrates with appropriate peptide sequences for the different proteases were used (see Supplementary Materials, Figures S19–S20).

The potential inhibitors were initially screened against all five target enzymes at 20 μM. A cut-off value of 80% inhibition at this concentration was set to differentiate the non-active (n.a.) compounds from active ones.

For the reversible inhibitors (α-ketobenzothiazole, nitroalkene, F-vinyl sulfone and nitrile), the IC$_{50}$ values were determined and converted to corresponding K_i values using the Cheng–Prusoff equation [56]. Regarding the irreversible inhibitors (vinyl sulfone, 4-oxoenoate and β-lactam) the K_i, k_{inact} and k_{2nd} values were determined (see Table S1) [56]. For a better overview, the pK_i values were calculated and are presented in Figure 6.
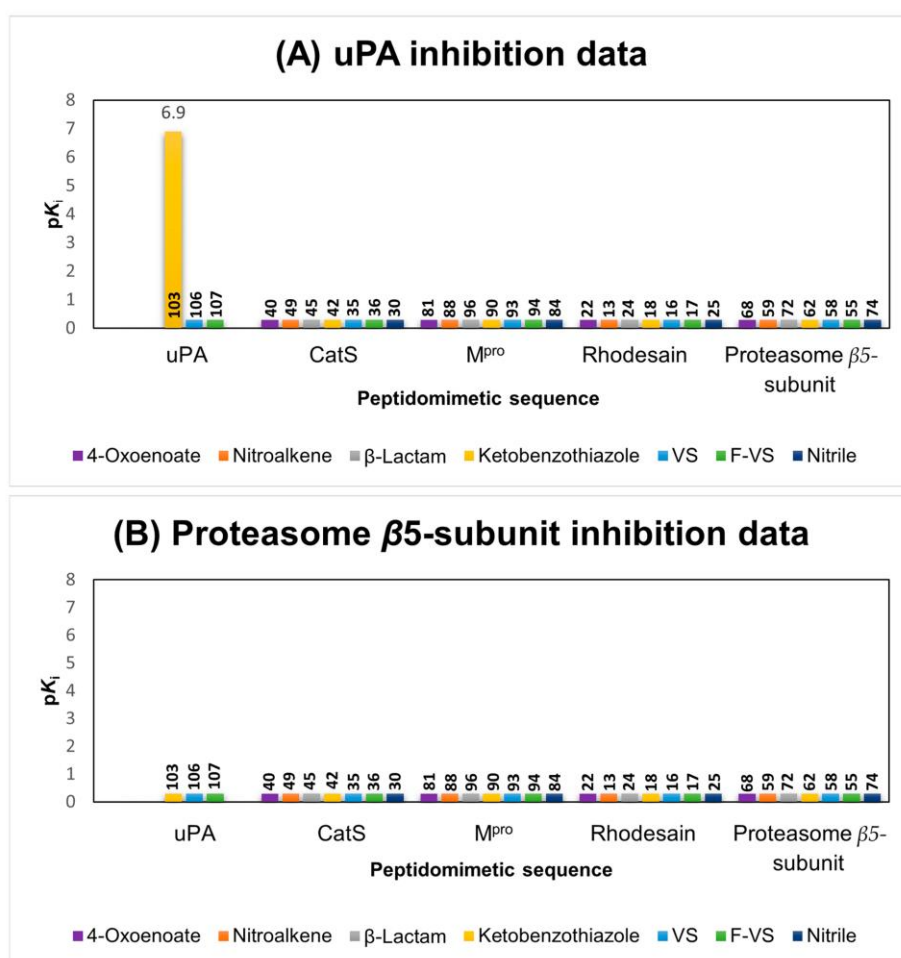
Figure 6. *Cont.*
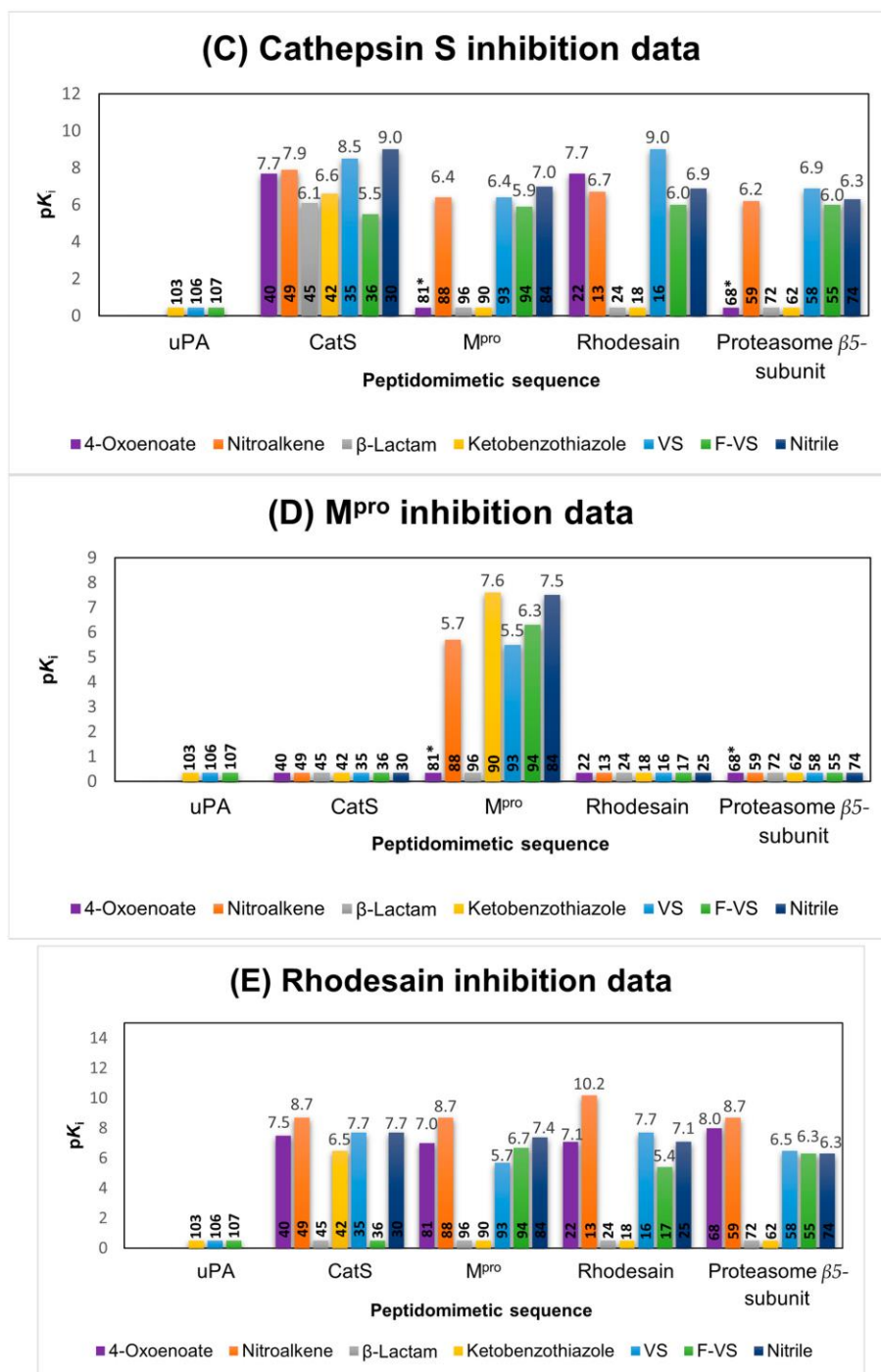
Figure 6. Inhibition data for the assays with uPA (**A**), proteasome *β5*-subunit (**B**), cathepsine S (**C**), SARS-CoV-2 M^pro (**D**) and rhodesain (**E**). pK_i values were calculated from the K_i values ($-\log_{10}(\underline{K_i}/M)$) [59]. The height of a bar indicates the inhibitory potency of an inhibitor towards the target enzyme, and the color of a bar indicates the warhead of the inhibitor; the peptidomimetic sequence is indicated by the enzyme name under the bars, e.g., the purple bar with the height value

of 7.5 for rhodesain inhibition (Figure 7E) by inhibitor **40** with the oxoenoate warhead and the CatS sequence. * Respective compounds were inactive in the in vitro assay due to instability towards DTT in the buffer.
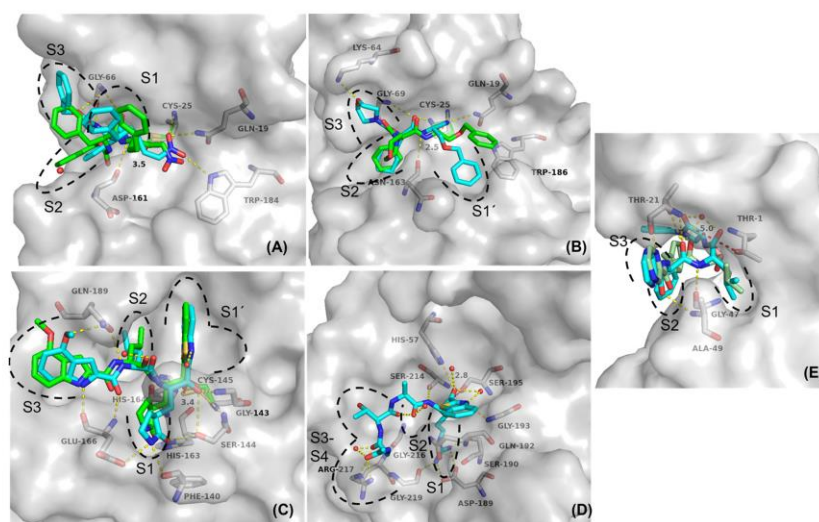
Figure 7. Predicted binding modes and polar interactions (yellow dashed lines) of different inhibitor classes with different enzymes (white carbon atoms and surface). For non-covalent docking poses, the distance between electrophilic carbon and nucleophilic sulfur or oxygen is shown as a red dashed line, with the distance measured in Å. For a clear view, only amino acids that form polar interactions with the ligands are shown as sticks and labelled. Black dashed lines indicate subpocket locations. Non-covalent docking poses are shown with cyan C-atoms and the covalent docking poses with green C-atoms. (**a**) Superposition of the non-covalent and the covalent docking pose of **13** with rhodesain, PDB-ID (2P7U). (**b**) Superposition of the non-covalent and the covalent docking pose of **30** with CatS, PDB-ID (1MS6). (**c**) Superposition of the non-covalent and the covalent docking pose of **90** with SARS-CoV-2-Mᵖʳᵒ, PDB-ID (6XR3). (**d**) Predicted binding mode of non-covalently docked **103** with uPA, PDB-ID (1W10). (**e**) Superposition of the non-covalent docking pose of **72** and bortezomib (palegreen C-atoms) with the *β*-5 subunit of human 20S-proteasome, PDB-ID (5LF3).

In the following, the inhibition data will be analyzed for each enzyme, first with their suited peptidomimetic sequences (Figure 6, parts A, B, C, D and E), followed by cross testing against the other enzymes.

uPA. Only the *α*-ketobenzothiazole inhibitor **103** was found to be active. The combination of the appropriate sequence for uPA with the *α*-ketobenzothiazole warhead resulted in a potent inhibitor with a pK_i value of 6.9. Other enzymes were not inhibited (Figure 6A).

Proteasome *β*5-subunit. None of the compounds with the Pyz–(L)Phe–(L)Leu sequence (**55**, **85**, **59**, **62**, **68**, **72** and **74**), which is well-known from the potent boronic acid-based inhibitor bortezomib, showed inhibition of the proteasome at 20 µM, independently of the warhead used (Figure 6B). Moreover, none of the other compounds with any of the other peptidomimetic sequences showed any inhibition. This highlights the general difficulty of addressing this protease with peptidomimetic inhibitors [56]. An alternative warhead which reacts preferably with Ser or Thr proteases is the epoxide functionality, which is also present in the approved proteasome inhibitor carfilzomib. Although very potent, due to its unpredictable reaction mechanism, this warhead was not included in this study [60].

CatS. Regarding the in vitro testing of the cysteine protease CatS, a total of 20 hits were detected (Figure 6C). The most potent inhibitors with the fitting CatS sequence were the nitrile **30** ($pK_i = 9$) and the vinyl sulfone **35** ($pK_i = 8.5$). The nitroalkene **49** ($pK_i = 7.9$) also showed high affinity towards CatS, followed by the 4-oxoenoate **40** ($pK_i = 7.7$), the α-ketobenzothiazole **42** ($pK_i = 6.5$), β-lactam **45** ($pK_i = 6.1$) and F-vinyl sulfone **36** ($pK_i = 5.5$). Since CatS and rhodesain are both papain-like cysteine proteases with similar active sites, cross reactivity between these two series was expected and has been well described in the literature [61]. The vinyl sulfone with the rhodesain-targeting sequence **16** ($pK_i = 9$) showed the same inhibition constant as the corresponding inhibitor with the CatS sequence. The vinyl sulfones with the proteasome and the M^{Pro} sequences inhibited CatS to lower degrees ($pK_i = 6.9$ and 5.5). A comparison of the 4-oxoenoates of the CatS and rhodesain series yielded the same results, since both exhibited the same pK_i value of 7.7 for inhibition of CatS. The 4-oxoenoates designed for targeting the proteasome **68** and the M^{Pro} **81** were essentially inactive against CatS (no inhibition in the initial screening at 20 μM). This can be explained by the instability of these compounds in the CatS assay buffer containing dithiothreitol (DTT).

The F-vinyl sulfones, which are reversibly reacting counterparts of the vinyl sulfones, inhibited CatS to a lower degree, and exchange of the peptidomimetic sequence (**36**, $pK_i = 5.5$ vs. **17**, $pK_i = 6$ vs. **55**, $pK_i = 6$ vs. **94**, $pK_i = 5.9$) had little to no effect, except for the compounds with the uPA sequence (**103**, **106** and **107**), which was not active at 20 μM against CatS.

The nitroalkene inhibitor which contains the CatS sequence showed a high on-target affinity but changing the sequence to any of the other targeting sequences led to less potent inhibitors (**13**, $pK_i = 6.7$ vs. **88**, $pK_i = 6.4$ μM vs. **59**, $pK_i = 6.2$). Interestingly, the α-ketobenzothiazole- (**42**, $pK_i = 6.6$) and β-lactam- (**45**, $pK_i = 6.1$) based inhibitors showed only significant inhibition of CatS if connected to the respective CatS sequence, indicating the strong dependency of a suitable peptidomimetic sequence combined with one of these warheads.

M^{Pro}. In comparison to the M^{Pro} inhibitors α-ketobenzothiazole **90** ($pK_i = 7.6$) and nitrile **84** ($K_i = 7.5$) described in the literature, vinyl sulfone **93** ($pK_i = 5.5$), F-vinyl sulfone **94** ($pK_i = 6.3$) and nitroalkene **88** ($pK_i = 5.7$), all of which contain the appropriate M^{Pro} peptidic sequence, showed weaker inhibition (Figure 6D) [40,62]. A clear preference of the protease for specific warheads could be observed. The vinylogous warheads (vinyl sulfone, F-vinyl sulfone and nitroalkene) showed significantly weaker inhibition than the α-ketobenzothiazole- and nitrile-based compounds. As also observed in the model reactivity studies, the in vitro studies with CatS and both 4-oxoenoate inhibitors with the M^{Pro} **81** and the proteasome β5-subunit **68** sequences revealed instability in the buffer with DTT. The β-lactam **96** as well as all compounds containing a targeting structure designed for other proteases were inactive at 20 μM. This indicates a high specificity of the M^{Pro} towards its peptidomimetic sequence.

Rhodesain. The results showed similar trends to those found for CatS, with 20 compounds active in the assays (Figure 6E). The most potent was the nitroalkene **13** which contains the corresponding rhodesain peptidic sequence ($pK_i = 10.2$), followed by the vinyl sulfone **16**, 4-oxoenoate **22** and nitrile **25**, which showed similar inhibition constants ($pK_i = 7.1–7.7$). The F-vinyl sulfone **17** showed moderate inhibition ($pK_i = 5.4$), and the β-lactam **24** and α-ketobenzothiazole **18** were inactive at 20 μM, indicating a preference of rhodesain for vinylogous warheads. Comparable to the CatS study, inhibitors lacked selectivity between rhodesain and CatS due to the structural similarity of the proteases. This is evident through the high pK_i values of the synthesized CatS inhibitors with 4-oxoenoate- **40** ($pK_i = 7.5$), nitroalkene- **49** ($pK_i = 8.7$), vinyl sulfone- **35** ($pK_i = 7.7$) and nitrile- **30** ($pK_i = 7.7$) moieties. Surprisingly, the α-ketobenzothiazole **42** designed for targeting CatS showed significant inhibition ($pK_i = 6.5$), whereas the α-ketobenzothiazole with the rhodesain peptidic sequence **18** was inactive. Among the compounds designed for M^{Pro} and the proteasome, the nitroalkene derivatives **88** and **59** showed the same potency as the CatS analogue **49**, both

with a pK_i value of 8.7. Interestingly, the vinyl sulfones with the M^{pro} **93** (pK_i = 5.7) and the proteasome sequence **58** (pK_i = 6.5) showed significantly lower affinity compared to the vinyl sulfone designed for rhodesain **16** (pK_i = 4.7). Differently, the F-vinyl sulfones **55** (pK_i = 6.3) and **94** (pK_i = 7.7) showed higher affinities than the analogue with the rhodesain sequence **17** (pK_i = 5.4). The 4-oxoenoate inhibitor **68** (pK_i-= 8.0) designed for the proteasome and the one designed for the M^{pro} **81** (pK_i = 7.0) also showed strong inhibition. All other inhibitors with the β-lactam and α-ketobenzothiazole moiety were inactive, as well as the compounds containing the uPA sequence (**103**, **106** and **107**).

2.5. Molecular Docking

To further elucidate the impact of the different warhead types on the binding modes of the inhibitors, protease–inhibitor complexes were investigated with non-covalent and covalent docking [63,64]. For the non-covalent docking, special emphasis was laid on the distances between the reactive nucleophilic carbon atoms of the corresponding warheads to the thiol(ate) or hydroxyl(ate) side chains of the cysteine/serine(threonine) active site amino acids, respectively, as estimates for covalent-bond-formation likeliness. Additionally, the impact of the different warhead moieties on the binding conformation of the inhibitors with otherwise identical peptidomimetic recognition sequences was analyzed. The covalent docking setup was used to investigate whether realistic poses for the covalent complexes could be generated and whether larger conformational rearrangements of the ligand may occur after the covalent reaction.

Conventional non-covalent docking yielded generally reasonable binding modes for all complexes resembling interactions of the crystallographic reference ligands and peptidomimetic recognition sequences in their expected subpockets. Additionally, electrophilic warheads were regularly found in close proximity to the nucleophilic catalytic amino acids (Figure 7, Tables S2 and S3).

The docking with rhodesain (pdb entry 2P7U) indicated that the introduction of a Michael-acceptor system as the warhead led to binding poses similar to the co-crystallized reference ligand, with all the essential interactions between inhibitor and enzyme being nearly identical, as exemplified for the docking poses of the nitroalkene inhibitor **13** (Figure 7a). The poses of the covalent and the non-covalent docking showed that the overall orientation of the inhibitor inside the active site should not change much after the covalent reaction, since the final covalent enzyme–inhibitor complex is very similar to the non-covalent complex (Figure 7a). The corresponding electrophilic C-atoms of all warheads were predicted to be in close proximity to the sulfur atom of Cys25 (2.54–3.50 Å), suggesting a high probability for a nucleophilic attack. High docking scores were also found for the nitroalkene inhibitor **13** (FlexX score: –24.03 kJ/mol; MOE score: –2.66), indicating that it should form very favorable non-covalent interactions while correctly placing the electrophilic warhead (distance to Cys25 sulfur: 3.50 Å). This is consistent with the in vitro data, showing that the nitroalkene moiety represents the most potent inhibitor class for rhodesain. Since the α-ketobenzothiazole designed for CatS (**42**) surprisingly inhibited rhodesain with a submicromolar affinity, we compared the non-covalent docking poses between **42** and the ketobenzothiazole with the rhodesain sequence (**18**) (Figure S21). Superposition of the non-covalent docking poses showed that both inhibitors have almost the same positioning with the warhead close to Cys-25 (2.5 Å) inside the active site of rhodesain, indicating that both compounds should have similar affinities towards rhodesain. This makes it hard to explain why inhibitor **42** had a significantly higher affinity for rhodesain in the in vitro testing. Since molecular docking is an inaccurate method, flawed docking poses are no rarity. The non-covalent docking method used in this case might not be suited to explaining this in vitro result. The results of the docking with CatS (1MS6) showed similar trends, since the distances between the electrophilic C-atoms of the warheads and the sulfur atom of Cys25 were again in close proximity in all cases (2.69–3.37 Å). The vinyl sulfone **35** and the nitrile **25** had high scores (FlexX score: –27.35/–26.22 kJ/mol; MOE score: –5.32/–3.00 kcal/mol) combined with similar binding geometries for the covalent

and the non-covalent docking poses (shown for nitrile inhibitor **25**, Figure 7b). These data are in accordance with the in vitro data showing that the nitrile warhead was the most potent one, but other warheads also led to productive enzyme inhibition.

For SARS-CoV-2 M^{pro} (6XR3), the distances between the electrophilic C-atoms and the Cys145 sulfur atom were overall slightly higher (2.90–4.91 Å) compared to the papain-like cysteine proteases. The α-ketobenzothiazole warhead seems to have a very favorable positioning in the binding pocket, as illustrated by the close proximity (3.41 Å) of the electrophilic C-atom to the thiol of the enzyme Cys145 (Figure 7c). Superposition of the covalent and non-covalent docking poses of **90** showed almost identical positioning of the inhibitor inside the enzyme, with most of the polar interactions retained.

Out of all the investigated warheads in this series, only the nitrile, the α-ketobenzothiazole and the β-lactam warheads are known to react with oxygen containing amino acid residues in serine (uPA) or threonine (proteasome) proteases.

For the uPA, which was the only target with only one hit in the enzymatic assay, non-covalent docking revealed a large distance between the electrophilic C-atom and the hydroxy-group in the active site for the β-lactam (5.07Å) as a known serine warhead. Only the α-ketobenzothiazole inhibitor **103**, which had one of the highest scores out of all the inhibitors (FlexX score: –51.59 kJ/mol), was in close proximity to the oxygen of Ser195 (2.84 Å distance to the electrophilic C-atom). This inhibitor also showed a high potency in the in vitro study (Figure 7d). Finally, docking of the β-lactam containing inhibitor **72** designed for the proteasome revealed that the warhead position was, again, too far away from the threonine oxygen (4.95 Å), possibly preventing a covalent reaction (Figure 7e). This could be explained by the shifted positioning of the lactam moiety compared to the other warheads. Although the docking of the nitrile and α-ketobenzothiazole inhibitors **74** and **62** might suggest that these compounds should inhibit their target sufficiently since the warheads are positioned correctly and in close proximity (2.27 Å/3.16 Å) to the Thr-1 oxygen atom, there was still no inhibition with these warheads in the in vitro study. This might have been due to wrongly generated binding poses, since docking approaches are not always reliable and cannot be considered flawless in all cases. A possible explanation why none of the compounds designed to address the β5-subunit of the proteasome showed any inhibition might be the catalytic dyad in the active site consisting of Lys33 and Thr1 compared to the catalytic dyads or even triads in the other enzymes, where the deprotonation of the active site residue is assisted by histidine and/or asparagine. The lysine residue might not always be able to deprotonate the threonine in the active site, depending on the inhibitor, and thus facilitate the covalent reaction step with a warhead [65].

2.6. Comparison of the Reactivity Assay Results with the In Vitro Study

Based on the reactivity assay, all Michael acceptors (4-oxoenoate **112**, (F-) vinyl sulfone **109**/**110** and nitroalkene **111**) showed high reactivity toward the deprotonated cysteine model nucleophile, which is congruent with the observed behavior of the synthesized compounds designed for CatS and rhodesain inhibition in the in vitro studies. Furthermore, the α-ketobenzothiazole warhead **115** showed a strong reactivity for both model nucleophiles (PhEtS$^-$/EtONa), which is consistent with the correspondent uPA and M^{pro} inhibitors **103** and **90** in the protease assays. However, the nitrile **117** showed no reaction with the deprotonated cysteine but with the serine model nucleophile, which contradicts the high inhibitory activity against the cysteine proteases and the missing inhibition by the proteasome β5-subunit inhibitor **74**. This might have been due to the aforementioned instability of the thioimidate adduct in basic conditions compared to the stabilized adduct in the enzyme pocket and the overall difficulty of addressing the proteasome β5-subunit. The β-lactam **108** showed only a strong reactivity towards the serine model nucleophile, but the corresponding bortezomib derivative **72** did not inhibit the proteasome β5-subunit, which might have been due to the shift of the electrophilic center of the β-lactam moiety into the S1' pocket and the resulting increase in distance. The 4-oxoenoate moiety **112**

was the only warhead that showed high reactivity toward the protonated cysteine model nucleophile, which might hint at non-selective reactivity behavior toward thiol species under physiological conditions. This could also be observed in the in vitro studies. The 4-oxoenoate compounds designed for the M^{pro} **81** and proteasome-β5-subunit **68** both reacted quickly with DTT in the respective buffer solutions and appeared to be inactive.

3. Discussion

Covalent targeting has become a popular and powerful concept in drug discovery, and great efforts have been devoted to developing and repurposing different warheads [66]. In this first extensive systematic study, we aimed to achieve a deeper insight into the reactivities and selectivities of a selection of electrophilic traps combined with established peptidomimetic sequences for the uPA, CatS, β5-subunit of the proteasome, SARS-CoV-2 M^{pro} and rhodesain, which represent cysteine, serine and threonine proteases. Based on these peptidomimetic sequences, we synthesized compounds decorated with warheads of different specificities. We chose the Michael acceptors ((F-)vinyl sulfone, nitroalkene and 4-oxoenoate) as cysteine-targeting and β-lactam as serine/threonine-targeting representatives. Furthermore, nitriles and α-ketobenzothiazoles were used, as they are applicable for both hydroxy- and thiol-containing nucleophiles. The compounds were tested on each target to analyze their affinities as well as their selectivity profiles.

Based on the in vitro studies, it is evident that the peptidomimetic sequences of the synthesized compounds play a crucial role in the selectivity towards the tested on-target and off-target proteases. This could be observed by the selectivity profile towards the cysteine protease M^{pro} and serine protease uPA. Only the inhibitors with the suited peptidomimetic sequence for M^{pro} (**84**, **88**, **90**, **93** and **94**) and for uPA (**103**) displayed inhibitory activity towards their targeted protease. Furthermore, the selection of a suitable warhead for the specific type of protease nucleophile ensures high affinity to the target or even activity in the first place, as demonstrated with the bortezomib congeners and the α-ketobenzothiazole inhibitor **103** as the only affine compound towards the uPA. The structurally similar papain-like proteases CatS and rhodesain showed that cross reactivity can occur, despite the design of well-defined peptidomimetic sequences. Therefore, the combination of both a highly reactive warhead towards the target protease, for example, the nitrile **30** group for CatS or the nitroalkene **13** for rhodesain, with a suitable peptidomimetic sequence can lead to potent inhibitors with promising pharmacodynamic properties.

Non-covalent docking yielded reasonable binding modes for all compounds resembling interactions of the crystallographic reference ligands and peptide recognition sequences in their expected subpockets. Additionally, electrophilic warheads were regularly found in close proximity to the nucleophilic catalytic amino acids, except for the β-lactams.

A reactivity test system with tool compounds of the used warheads and model nucleophiles was established to evaluate chemoselectivity. The findings confirmed the high reactivity of the 4-oxoenoate, the (F-)vinyl sulfones and the nitroalkene moieties towards the deprotonated thiol nucleophile/cysteine model, and high affinity of the Michael acceptor inhibitors towards the cysteine proteases was observed. Analogously to the in vitro studies of the uPA and M^{pro} target, the α-ketobenzothiazole warhead was found to be a potent electrophilic trap for both cysteine and serine proteases. Nevertheless, some major differences in reactivity could be observed, which might have been due to different conditions used in the chemical test system and the biochemical in vitro studies. To the best of our knowledge, this is the first extensive study in which different warhead types were combined with different peptidic recognition units and in which the resulting compounds were cross tested against different protease types. Similar published studies limited their focus to testing different warheads on one target or exchanging the peptidic backbone while retaining the same warhead [37,59,67,68].

4. Material and Methods

The material 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 [21,42,43,46,47,63,64,69–89] (Supplementary Figures of the reactivity study (Figures S1–S18), of the fluorometric inhibition assays (Figures S19 and S20), of molecular docking (Figure S21), of quantum mechanics simulation (Figures S22–S25) and of the NMR-spectra and HPLC-chromatograms of the final inhibitors (Figures S26–S137) can be accessed in the supporting information).

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. Grozdanić, M.; Vidmar, R.; Vizovišek, M.; Fonović, M. Degradomics in Biomarker Discovery. *Proteom. Clin. Appl.* **2019**, *13*, 1800138. [CrossRef] [PubMed]
3. Ruggiano, A.; Ramadan, K. DNA–protein crosslink proteases in genome stability. *Commun. Biol.* **2021**, *4*, 11. [CrossRef]
4. Lee, C.W.; Stankowski, J.N.; Chew, J.; Cook, C.N.; Lam, Y.W.; Almeida, S.; Carlomagno, Y.; Lau, K.F.; Prudencio, M.; Gao, F.B.; et al. The lysosomal protein cathepsin L is a progranulin protease. *Mol. Neurodegener.* **2017**, *12*, 55. [CrossRef] [PubMed]
5. 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] [PubMed]
6. Liu, C.L.; Guo, J.; Zhang, X.; Sukhova, G.K.; Libby, P.; Shi, G.P. Cysteine protease cathepsins in cardiovascular disease: From basic research to clinical trials. *Nat. Rev. Cardiol.* **2018**, *15*, 351–370. [CrossRef]

7.   Previti, S.; Ettari, R.; Calcaterra, E.; Di Chio, C.; Ravichandran, R.; Zimmer, C.; Hammerschmidt, S.; Wagner, A.; Bogacz, M.; Cosconati, S.; et al. Development of Urea-Bond-Containing Michael Acceptors as Antitrypanosomal Agents Targeting Rhodesain. *ACS Med. Chem. Lett.* **2022**, *13*, 1083–1090. [CrossRef]

8.   Rocha, D.A.; Silva, E.B.; Fortes, I.S.; Lopes, M.S.; Ferreira, R.S.; Andrade, S.F. Synthesis and structure-activity relationship studies of cruzain and rhodesain inhibitors. *Eur. J. Med. Chem.* **2018**, *157*, 1426–1459. [CrossRef]

9.   Kincaid, J.R.A.; 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]

10.   Müller, P.; Maus, H.; Hammerschmidt, S.J.; Knaff, P.M.; Mailänder, V.; Schirmeister, T.; Kersten, C. Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy. *Curr. Med. Chem.* **2022**, *29*, 635–665. [CrossRef]

11.   Knaff, P.M.; Müller, P.; Kersten, C.; Wettstein, L.; Münch, J.; Landfester, K.; Mailänder, V. Structure-Based Design of High-Affinity and Selective Peptidomimetic Hepsin Inhibitors. *Biomacromolecules* **2022**, *23*, 2236–2242. [CrossRef] [PubMed]

12.   Tsantrizos, Y.S.; Bolger, G.; Bonneau, P.; Cameron, D.R.; Goudreau, N.; Kukolj, G.; LaPlante, S.R.; Llinàs-Brunet, M.; Nar, H.; Lamarre, D. Macrocyclic inhibitors of the NS3 protease as potential therapeutic agents of hepatitis C virus infection. *Angew. Chemie Int. Ed.* **2003**, *42*, 1356–1360. [CrossRef] [PubMed]

13.   Maus, H.; Barthels, F.; Hammerschmidt, S.J.; Kopp, K.; Millies, B.; Gellert, A.; Ruggieri, A.; Schirmeister, T. SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases. *Bioorg. Med. Chem.* **2021**, *47*, 116392. [CrossRef] [PubMed]

14.   Martin, J.S.; MacKenzie, C.J.; Fletcher, D.; Gilbert, I.H. Characterising covalent warhead reactivity. *Bioorg. Med. Chem.* **2019**, *27*, 2066–2074. [CrossRef] [PubMed]

15.   Wang, Y.-H.; Zhang, F.; Diao, H.; Wu, R. Covalent Inhibition Mechanism of Antidiabetic Drugs—Vildagliptin vs Saxagliptin. *ACS Catal.* **2019**, *9*, 2292–2302. [CrossRef]

16.   Lamb, Y.N. Nirmatrelvir Plus Ritonavir: First Approval. *Drugs* **2022**, *82*, 585–591. [CrossRef]

17.   Robak, P.; Robak, T. Bortezomib for the Treatment of Hematologic Malignancies: 15 Years Later. *Drugs R D* **2019**, *19*, 73–92. [CrossRef]

18.   Johe, P.; Jung, S.; Endres, E.; Kersten, C.; Zimmer, C.; Ye, W.; Sönnichsen, C.; Hellmich, U.A.; Sotriffer, C.; Schirmeister, T.; et al. Warhead Reactivity Limits the Speed of Inhibition of the Cysteine Protease Rhodesain. *ACS Chem. Biol.* **2021**, *16*, 661–670. [CrossRef]

19.   Santos, M.; Moreira, R. Mini-Reviews. *Med. Chem.* **2007**, *7*, 1040–1050.

20.   Adams, J.; Kauffman, M. Development of the Proteasome Inhibitor Velcade^TM (Bortezomib). *Cancer Investig.* **2004**, *22*, 304–311. [CrossRef]

21.   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*, 140–156. [CrossRef] [PubMed]

22.   Fleming, F.F.; Yao, L.; Ravikumar, P.C.; Funk, L.; Shook, B.C. Nitrile-Containing Pharmaceuticals: Efficacious Roles of the Nitrile Pharmacophore. *J. Med. Chem.* **2010**, *53*, 7902–7917. [CrossRef] [PubMed]

23.   Bullock, T.L.; Breddam, K.; Remington, J.S. Peptide Aldehyde Complexes with Wheat Serine Carboxypeptidase II: Implications for the Catalytic Mechanism and Substrate Specificity. *J. Mol. Biol.* **1996**, *255*, 714–725. [CrossRef] [PubMed]

24.   Hu, X.; Lin, C.; Xu, Q.; Zhou, X.; Zeng, P.; McCormick, P.J.; Jiang, H.; Li, J.; Zhang, J. Structural Basis for the Inhibition of Coronaviral Main Proteases by a Benzothiazole-Based Inhibitor. *Viruses* **2022**, *14*, 2075. [CrossRef] [PubMed]

25.   Akiyama, Y.; Tsutsumi, S.; Hatsushiba, E.; Ohuchi, S.; Okonogi, T. Peptidyl α-keto thiazole as potent thrombin inhibitors. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 533–538. [CrossRef]

26.   Vincenza Carriero, M.; Patrizia Stoppelli, M. The Urokinase-type Plasminogen Activator and the Generation of Inhibitors of Urokinase Activity and Signaling. *Curr. Pharm. Des.* **2011**, *17*, 1944–1961. [CrossRef] [PubMed]

27.   Ismail, A.A.; Shaker, B.T.; Bajou, K. The plasminogen–activator plasmin system in physiological and pathophysiological angiogenesis. *Int. J. Mol. Sci.* **2021**, *23*, 337. [CrossRef] [PubMed]

28.   Smith, H.W.; Marshall, C.J. Regulation of cell signalling by uPAR. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 23–36. [CrossRef]

29.   Kumar, A.A.; Buckley, B.J.; Ranson, M. The Urokinase Plasminogen Activation System in Pancreatic Cancer: Prospective Diagnostic and Therapeutic Targets. *Biomolecules* **2022**, *12*, 152. [CrossRef]

30.   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]

31.   Collins, G.A.; Goldberg, A.L. The Logic of the 26S Proteasome. *Cell* **2017**, *169*, 792–806. [CrossRef] [PubMed]

32.   Kisselev, A.F.; van der Linden, W.A.; Overkleeft, H.S. Proteasome Inhibitors: An Expanding Army Attacking a Unique Target. *Chem. Biol.* **2012**, *19*, 99–115. [CrossRef] [PubMed]

33.   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*, e701. [CrossRef] [PubMed]

34.   Pauly, T.A.; Sulea, T.; Ammirati, M.; Sivaraman, J.; Danley, D.E.; Griffor, M.C.; Kamath, A.V.; Wang, I.K.; Laird, E.R.; Seddon, A.P.; et al. Specificity determinants of human cathepsin S revealed by crystal structures of complexes. *Biochemistry* **2003**, *42*, 3203–3213. [CrossRef] [PubMed]

35. Wilkinson, R.D.A.; Williams, R.; Scott, C.J.; Burden, R.E. Cathepsin S: Therapeutic, diagnostic, and prognostic potential. *Biol. Chem.* **2015**, *396*, 867–882. [CrossRef] [PubMed]

36. Chen, J.-C.; Uang, B.-J.; Lyu, P.-C.; Chang, J.-Y.; Liu, K.-J.; Kuo, C.-C.; Hsieh, H.-P.; Wang, H.-C.; Cheng, C.-S.; Chang, Y.-H.; et al. Design and Synthesis of α-Ketoamides as Cathepsin S Inhibitors with Potential Applications against Tumor Invasion and Angiogenesis. *J. Med. Chem.* **2010**, *53*, 4545–4549. [CrossRef] [PubMed]

37. 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]

38. Previti, S.; Ettari, R.; Cosconati, S.; Schirmeister, G.; Chouchene, K.; Wagner, A.; Hellmich, U.A.; Ulrich, K.; Krauth-Siegel, R.L.; Wich, P.R.; et al. 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. [CrossRef]

39. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289–293. [CrossRef]

40. 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]

41. 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] [PubMed]

42. Schrader, J.; Henneberg, F.; Mata, R.A.; Tittmann, K.; Schneider, T.R.; Stark, H.; Bourenkov, G.; Chari, A. The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. *Science* **2016**, *353*, 594–598. [CrossRef] [PubMed]

43. Ward, Y.D.; Thomson, D.S.; Frye, L.L.; Cywin, C.L.; Morwick, T.; Emmanuel, M.J.; Zindell, R.; McNeil, D.; Bekkali, Y.; Marc Girardot, M.; et al. Design and synthesis of dipeptide nitriles as reversible and potent Cathepsin S inhibitors. *J. Med. Chem.* **2002**, *45*, 5471–5482. [CrossRef] [PubMed]

44. Hattori, S.-I.; Higashi-Kuwata, N.; Hayashi, H.; Allu, S.R.; Raghavaiah, J.; Bulut, H.; Das, D.; Anson, B.J.; Lendy, E.K.; Takamatsu, Y.; et al. A small molecule compound with an indole moiety inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat. Commun.* **2021**, *12*, 668. [CrossRef] [PubMed]

45. 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]

46. Vuong, W.; Vederas, J.C. Improved Synthesis of a Cyclic Glutamine Analogue Used in Antiviral Agents Targeting 3C and 3CL Proteases Including SARS-CoV-2 M pro. *J. Org. Chem.* **2021**, *86*, 13104–13110. [CrossRef]

47. Tian, Q.; Nayyar, N.K.; Babu, S.; Chen, L.; Tao, J.; Lee, S.; Tibbetts, A.; Moran, T.; Liou, J.; Guo, M.; et al. An efficient synthesis of a key intermediate for the preparation of the rhinovirus protease inhibitor AG7088 via asymmetric dianionic cyanomethylation of N-Boc-l-(+)-glutamic acid dimethyl ester. *Tetrahedron Lett.* **2001**, *42*, 6807–6809. [CrossRef]

48. 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*, 1484–1487. [CrossRef]

49. Latorre, A.; Schirmeister, T.; Kesselring, J.; Jung, S.; Johé, P.; Hellmich, U.A.; Heilos, A.; Engels, B.; Krauth-Siegel, R.L.; Dirdjaja, N.; et al. Dipeptidyl Nitroalkenes as Potent Reversible Inhibitors of Cysteine Proteases Rhodesain and Cruzain. *ACS Med. Chem. Lett.* **2016**, *7*, 1073–1076. [CrossRef]

50. Shah, S.K.; Finke, P.E.; Brause, K.A.; Chandler, G.O.; Ashe, B.M.; Weston, H.; Maycock, A.L.; Mumford, R.A.; Doherty, J.B. Mono-cyclic β-lactam inhibitors of human leukocyte elastase. Stereospecific synthesis and activity of 3,4-disubstituted-2-azetidinones. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2295–2298. [CrossRef]

51. Han, W.T.; Trehan, A.K.; Kim Wright, J.J.; Federici, M.E.; Seiler, S.M.; Meanwell, N.A. Azetidin-2-one derivatives as inhibitors of thrombin. *Bioorg. Med. Chem.* **1995**, *3*, 1123–1143. [CrossRef] [PubMed]

52. Steert, K.; Berg, M.; Mottram, J.C.; Westrop, G.D.; Coombs, G.H.; Cos, P.; Maes, L.; Joossens, J.; Van der Veken, P.; Haemers, A.; et al. α-Ketoheterocycles as Inhibitors of Leishmania mexicana Cysteine Protease CPB. *ChemMedChem* **2010**, *5*, 1734–1748. [CrossRef] [PubMed]

53. Costanzo, M.J.; Almond, H.R.; Hecker, L.R.; Schott, M.R.; Yabut, S.C.; Zhang, H.-C.; Andrade-Gordon, P.; Corcoran, T.W.; Giardino, E.C.; Kauffman, J.A.; et al. In-Depth Study of Tripeptide-Based α-Ketoheterocycles as Inhibitors of Thrombin. Effective Utilization of the S1 'Subsite and Its Implications to Structure-Based Drug Design. *J. Med. Chem.* **2005**, *48*, 1984–2008. [CrossRef] [PubMed]

54. DiNinno, F.; Ernest, V.L. Facile Synthesis of β-Thioxo Esters from β-Enamino Esters. *J. Org. Chem.* **1979**, *44*, 3271–3273. [CrossRef]

55. Delprino, L.; Giacomotti, M.; Dosio, F.; Brusa, P.; Ceruti, M.; Grosa, G.; Cattel, L. Toxin-Targeted Design for Anticancer Therapy. I: Synthesis and Biological Evaluation of New Thioimidate Heterobifunctional Reagents. *J. Pharm. Sci.* **1993**, *82*, 506–512. [CrossRef] [PubMed]

56. Brogi, S.; Ibba, R.; Rossi, S.; Butini, S.; Calderone, V.; Gemma, S.; Campiani, G. Covalent Reversible Inhibitors of Cysteine Proteases Containing the Nitrile Warhead: Recent Advancement in the Field of Viral and Parasitic Diseases. *Molecules* **2022**, *27*, 2561. [CrossRef]

57.   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]

58.   Paasche, A.; Schiller, M.; Schirmeister, T.; Engels, B. Mechanistic Study of the Reaction of Thiol-Containing Enzymes with *α*,*β*-Unsaturated Carbonyl Substrates by Computation and Chemoassays. *ChemMedChem* **2010**, *5*, 869–880. [CrossRef]

59.   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.; et al. A comparative study of warheads for design of cysteine protease inhibitors. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 5031–5035. [CrossRef]

60.   Kim, K.B.; Crews, C.M. From epoxomicin to carfilzomib: Chemistry, biology, and medical outcomes. *Nat. Prod. Rep.* **2013**, *30*, 600. [CrossRef]

61.   Cianni, L.; Feldmann, C.W.; Gilberg, E.; Gütschow, M.; Juliano, L.; Leitão, A.; Bajorath, J.; Montanari, C.A. Can Cysteine Protease Cross-Class Inhibitors Achieve Selectivity? *J. Med. Chem.* **2019**, *62*, 10497–10525. [CrossRef] [PubMed]

62.   Konno, S.; Kobayashi, K.; Senda, M.; Funai, Y.; Seki, Y.; Tamai, I.; Schäkel, L.; Sakata, K.; Pillaiyar, T.; Taguchi, A.; et al. 3CL Protease Inhibitors with an Electrophilic Arylketone Moiety as Anti-SARS-CoV-2 Agents. *J. Med. Chem.* **2022**, *65*, 2926–2939. [CrossRef] [PubMed]

63.   Chemical Computing Group ULC. *Molecular Operating Environment (MOE)*; Chemical Computing Group ULC: Montreal, QC, Canada, 2020; Available online: https://www.chemcomp.com/. (accessed on 23 March 2022).

64.   *LeadIT/FlexX*, Version 2.3.2; GmbH, BioSolveIT: Sankt Augustin, Germany, 2017.

65.   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*, 3575–3589. [CrossRef]

66.   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*, 5673–5724. [CrossRef]

67.   Chenna, B.C.; Li, L.; Mellott, D.M.; Zhai, X.; Siqueira-Neto, J.L.; Calvet Alvarez, C.; Bernatchez, J.A.; Desormeaux, E.; Alvarez Hernandez, E.; Gomez, J.; et al. Peptidomimetic Vinyl Heterocyclic Inhibitors of Cruzain Effect Antitrypanosomal Activity. *J. Med. Chem.* **2020**, *63*, 3298–3316. [CrossRef]

68.   Vankadara, S.; Dawson, M.D.; Fong, J.Y.; Oh, Q.Y.; Ang, Q.A.; Liu, B.; Chang, H.Y.; Koh, J.; Koh, X.; Tan, Q.W.; et al. A Warhead Substitution Study on the Coronavirus Main Protease Inhibitor Nirmatrelvir. *ACS Med. Chem. Lett.* **2022**, *13*, 1345–1350. [CrossRef] [PubMed]

69.   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]

70.   Barthels, F.; Marincola, G.; Marciniak, T.; Konhäuser, M.; Hammerschmidt, S.; Bierlmeier, J.; Distler, U.; Wich, P.R.; Tenzer, S.; Schwarzer, D.; et al. Irreversible and Selective Inhibitors of Staphylococcus aureus Sortase A. *ChemMedChem* **2020**, *15*, 839–850. [CrossRef]

71.   Amendola, G.; Ettari, R.; Previti, S.; Di Chio, C.; Messere, A.; Di Maro, S.S.; Hammerschmidt, J.; Zimmer, C.; Zimmermann, R.A.; Schirmeister, T.; et al. Lead Discovery of SARS-CoV-2 Main Protease Inhibitors through Covalent Docking-Based Virtual Screening. *J. Chem. Inf. Model.* **2021**, *61*, 2062–2073. [CrossRef]

72.   Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T.H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P.R.; Distler, U.; et al. Engels, Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138*, 8332–8335. [CrossRef]

73.   Caffrey, C.R.; Hansell, E.; Lucas, K.D.; Brinen, L.S.; Hernandez, A.A.; Cheng, J.; Roush, W.R.; Stierhof, Y.-D.; Bogyo, M.; Steverding, D.; 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]

74.   Berman, H.M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef]

75.   Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* **2003**, *10*, 980. [CrossRef] [PubMed]

76.   Halgren, T.A. MMFF94s option for energy minimization studies. *J. Comput. Chem.* **1999**, *20*, 720–729. [CrossRef]

77.   *The PyMOL Molecular Graphics System*, version 2.5.2; Schrödinger, LLC: New York, NY, USA, 2021.

78.   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. [CrossRef] [PubMed]

79.   Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Peterson, G.A.; Nakatsuji, H.; et al. *Gaussian 16 (Revision A.03)*; Gaussian Inc.: Wallingford, CT, USA, 2016.

80.   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]

81.   Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [CrossRef]

82.   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]

83. Ben-Naim, A. Standard Thermodynamics of Transfer. Uses and Misuses. *J. Phys. Chem.* **1978**, *82*, 792–803. [CrossRef]
84. Spina, R.; Colacino, E.; Martinez, J.; Lamaty, F. Poly(ethylene glycol) as a Reaction Matrix in Platinum- or Gold-Catalyzed Cycloisomerization: A Mechanistic Investigation. *Chem.-A Eur. J.* **2013**, *19*, 3817–3821. [CrossRef]
85. Ho, A.; Cyrus, K.; Kim, K.-B. Towards Immunoproteasome-Specific Inhibitors: An Improved Synthesis of Dihydroeponemycin. *Eur. J. Org. Chem.* **2005**, *2005*, 4829–4834. [CrossRef]
86. 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] [PubMed]
87. 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]
88. Engel-Andreasen, J.; Wellhöfer, I.; Wich, K.; Olsen, C.A. Backbone-Fluorinated 1,2,3-Triazole-Containing Dipeptide Surrogates. *J. Org. Chem.* **2017**, *82*, 11613–11619. [CrossRef] [PubMed]
89. Dutton, F.E.; Lee, B.H.; Johnson, S.S.; Coscarelli, E.M.; Lee, P.H. Restricted Conformation Analogues of an Anthelmintic Cyclodepsipeptide. *J. Med. Chem.* **2003**, *46*, 2057–2073. [CrossRef]

7.4   Experimental section of the manuscript

Note: Experimental procedures conducted or supervised by MARVIN SCHWICKERT (█████████ group) are described in this section. Experiments performed by others can be found in the complete Supporting Information file in the **Appendix**. Compound numbers in this section refer to the numbering in the corresponding manuscript.

General information

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, and anhydrous solvents were used as provided by the commercial supplier. Reaction progress was monitored by thin-layer chromatography using ALUGRAM® Xtra SIL G UV$_{254}$ silica plates from *Macherey-Nagel* and/or LC-MS. For LC-MS analysis, an *Agilent 1100 series* HPLC system and an *Agilent Poroshell 120 EC-C$_{18}$*, 150 × 2.10 mm, 4 μm column coupled to an *Agilent 1100 series* LC/MSD Trap with electron spray ionization (ESI) was used. The identities and purities of compounds were determined by the same LC-MS system with a gradient of MeCN and MilliQ®-H$_2$O + 0.1% HCOOH (flow rate: 0.7 mL/min). 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 a *Varian PrepStar* system using an *Agilent Zorbax PrepHT XDB C$_{18}$* (150 mm × 21.2 mm, 5 μm) column or an *Agilent 1290 II Infinity* preparative LC system using an *InfinityLab Pursuit XRs* C$_{18}$, 30 × 250 mm, 5 μm, preparative LC column. Flash chromatography was performed with a *Biotage Isolera^{TM} One* system using prepacked columns from *Biotage*. Silica (40–63 μm) from *Macherey-Nagel* was used for column chromatography. Optical rotations $[\alpha]_D^{20}$ were measured on a *P3000* polarimeter from *Krüss* and are reported in deg cm^3 g^{-1} dm^{-1}. 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 a *Bruker Fourier 300 MHz*, a *Bruker Avance DSX 400 MHz*, and a *Bruker Avance III 600 MHz*. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance (CDCl$_3$, DMSO-d_6 or CD$_3$OD from *Deutero GmbH*) as internal standard. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

tert-Butyl (S)-(1-(methoxy(methyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (51)
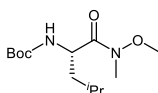
To a 0 °C cold solution of Boc-L-Leu-OH (**50**, 2.04 g, 8.81 mmol, 1.0 equiv.) in DCM (80 mL) were added HOBt • H$_2$O (1.35 g, 8.81 mmol, 1.0 equiv.) and 2,4,6-collidine (2.34 mL, 17.62 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (2.83 g, 8.81 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and N,O-dimethylhydroxylamine • HCl (0.89 g, 8.81 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (100 mL). The mixture was washed with saturated NaHCO$_3$ solution (3 × 80 mL) and 1 M HCl (3 × 80 mL) and filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C to yield the desired product as a colorless oil (2.24 g, 8.18 mmol, 93%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 5.05 (d, J = 9.5 Hz, 1H), 4.70 (s, 1H), 3.77 (s, 3H), 3.18 (s, 3H), 1.78–1.62 (m, 1H), 1.47–1.38 (m, 11H), 0.99–0.88 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 174.0, 155.8, 79.6, 61.7, 49.1, 42.2, 32.3, 28.5, 24.8, 23.5, 21.7. $[\alpha]_D^{20}$ = −9 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3324, 2958, 2936, 2870, 1709, 1658, 1501, 1389, 1366, 1250, 1165, 1045, 1016, 989, 876. ESI-MS: m/z calculated for [C$_{13}$H$_{26}$N$_2$O$_4$+Na]$^+$ ([M+Na]$^+$): 297.2, found: 297.1.

tert-Butyl (4-methyl-1-oxopentan-2-yl)carbamate (52)
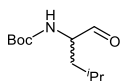
To a solution of **51** (1.51 g, 5.50 mmol, 1.0 equiv.) in dry THF (20 mL) was added LiAlH$_4$ (271 mg, 7.15 mmol, 1.3 equiv.) portionwise at 0 °C under argon atmosphere. After stirring at 0 °C for 30 min, diethyl ether (50 mL) and KHSO$_4$ solution (0.33 M, 80 mL,) were added. The suspension was filtered, and the filtrate was extracted with diethyl ether (2 × 50 mL). The combined organic extracts were washed with 1 M HCl (2 × 40 mL) and saturated NaHCO$_3$ solution (2 × 40 mL) and dried over anhydrous Na$_2$SO$_4$. After removing the solvent under reduced pressure at 40 °C, the desired product was obtained as a colorless oil (345 mg, 1.60 mmol, 29%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.57 (s, 1H), 4.95 (d, J = 7.3 Hz, 1H), 4.23 (s, 1H), 1.81–1.70 (m, 1H), 1.67–1.58 (m, 1H), 1.44 (s, 9H), 1.39–1.32 (m, 1H), 0.98–0.94 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 200.5, 155.8, 80.2, 58.5, 38.3, 28.4, 24.8, 23.2. $[\alpha]_D^{20}$ = −9 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3352, 2958, 2932, 2871, 1689, 1507, 1455, 1391, 1366, 1249, 1164, 1045, 1010, 873, 779.

tert-Butyl (*E*)-(1-fluoro-5-methyl-1-(phenylsulfonyl)hex-1-en-3-yl)carbamate (53)
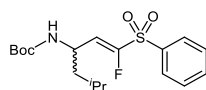
To a solution of **4** (451 mg, 1.45 mmol, 1.0 equiv.) and DBU (217 μL, 1.45 mmol, 1.0 equiv.) in dry MeCN (20 mL) was added anhydrous LiCl (74 mg, 1.74 mmol, 1.2 equiv.) at 0 °C under argon atmosphere. The solution was stirred for 20 min, and a solution of **52** (313 mg, 1.45 mmol, 1.0 equiv.) in dry MeCN (6 mL) was added. After stirring for 1 h at 0 °C under an argon atmosphere, citric acid solution (10% in water, 60 mL) was added. The mixture was extracted with ethyl acetate (3 × 30 mL), and the combined organic layers were filtered over a silica. The filtrate was concentrated under reduced pressure at 40 °C, and the residue was purified by column chromatography (cyclohexane/ethyl acetate = 10:1) to yield the desired product as a colorless oil (249 mg *E*-isomer + 183 mg *E/Z* mixture, $\sum$ = 432 mg, 1.16 mmol, 80%). *E*-isomer: ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.97–7.90 (m, 2H), 7.73–7.63 (m, 1H), 7.62–7.51 (m, 2H), 6.26–5.96 (m, 1H), 4.64–4.35 (m, 2H), 1.66–1.25 (m, 12H), 0.92–0.85 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 155.0, 154.5 (d, J = 298.7 Hz), 137.3, 134.6, 129.6, 128.7, 119.2 (d, J = 5.3 Hz), 80.1, 45.1, 43.5, 28.3, 24.8, 22.5, 22.4. $[\alpha]_D^{20}$ = +21 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3390, 2959, 2933, 2871, 1699, 1699, 1448, 1332, 1247, 1158, 1093, 1013, 754, 720, 686. ESI-MS: *m/z* calculated for [C$_{18}$H$_{26}$FNO$_4$S+Na]$^+$ ([M+Na]$^+$): 394.2, found: 394.1.

tert-Butyl ((2*S*)-1-(((*E*)-1-fluoro-5-methyl-1-(phenylsulfonyl)hex-1-en-3-yl)amino)-1-oxo-3-phenyl-propan-2-yl)carbamate (54)
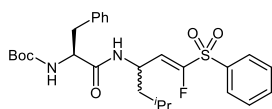
(1) **53** (223 mg, 0.60 mmol) was treated with 4 M HCl in dioxane (3 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C to yield the deprotected amine hydrochloride as a colorless solid (184 mg, 0.60 mmol, quantitative), which was directly used in the next step.

(2) To a 0 °C cold solution of Boc-L-Phe-OH (160 mg, 0.60 mmol, 1.0 equiv.) in DCM (5 mL) were added HOBt • H$_2$O (92 mg, 0.60 mmol, 1.0 equiv.) and 2,4,6-collidine (160 μL, 1.20 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (194 g, 0.60 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and the amine hydrochloride from step 1 (184 mg, 0.60 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (30 mL). The mixture was washed with saturated NaHCO$_3$ solution (3 × 25 mL) and 1 M HCl (3 × 25 mL) and filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C to yield the desired product as a colorless foam (302 mg, 0.58 mmol, 97%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 7.96–7.89 (m, 2H), 7.73–7.65 (m, 1H), 7.62–7.54 (m, 2H), 7.33–7.22 (m, 3H), 7.17–7.12 (m, 2H), 6.16–5.86 (m, 2H), 5.17–5.01 (m, 1H), 4.81–4.67 (m, 1H), 4.30–4.18 (m, 1H), 3.08–2.89 (m, 2H), 1.52–

1.44 (m, 1H), 1.41–1.38 (m, 9H), 1.34–1.22 (m, 1H), 0.85–0.77 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.9, 155.6, 154.9 (d, J = 299.7 Hz), 137.2, 136.6, 134.7, 129.6, 129.4, 128.9, 128.7, 127.2, 118.1 (d, J = 5.1 Hz), 80.5, 56.1, 43.5, 43.2, 38.5, 28.4, 24.7, 22.6, 22.1. mp: 47–49 °C. $[\alpha]_D^{20}$ = +11 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3302, 3064, 3031, 2959, 2931, 2871, 1655, 1521, 1366, 1333, 1165, 1092, 753, 720, 686. ESI-MS: m/z calculated for [C$_{27}$H$_{35}$FN$_2$O$_5$S+Na]$^+$ ([M+Na]$^+$): 541.2, found: 541.2.

N-((S)-1-(((S,E)-1-fluoro-5-methyl-1-(phenylsulfonyl)hex-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (55)
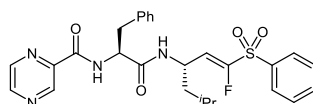

(1) **54 (**254 mg, 0.49 mmol) was treated with 4 M HCl in dioxane (3 mL) at room temperature for 1.5 h. The solvent was removed under reduced pressure at 40 °C to yield the amine hydrochloride as a colorless solid (224 mg, 0.49 mmol, quantitative), which was directly used in the next step.

(2) To a 0 °C cold solution of pyrazinecarboxylic acid (61 mg, 0.49 mmol, 1.0 equiv.) in DCM (6 mL) were added HOBt • H$_2$O (75 mg, 0.49 mmol, 1.0 equiv.) and 2,4,6-collidine (130 μL, 0.98 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (158 g, 0.49 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and the amine hydrochloride (224 mg, 0.49 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (30 mL). The mixture was washed with saturated NaHCO$_3$ solution (3 × 25 mL) and 1 M HCl (3 × 25 mL) and dried over anhydrous Na$_2$SO$_4$. After the solvent was removed under reduced pressure at 40 °C, the residue was purified by column chromatography (cyclohexane/ethyl acetate = 1:1) to yield the product as a mixture of two diastereomers (colorless resin, 174 mg, 0.33 mmol, 67%). 40 mg of diastereomeric mixture were separated by preparative HPLC (MeCN/H$_2$O = 25:75) to give 20 mg pure (S,S)- and 10 mg pure (S,R)-diastereomer. (S,S)-diastereomer: ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.34–9.25 (m, 1H), 8.75 (d, J = 2.5 Hz, 1H), 8.57–8.49 (m, 1H), 8.42 (d, J = 8.2 Hz, 1H), 7.98–7.91 (m, 2H), 7.73–7.66 (m, 1H), 7.63–7.55 (m, 2H), 7.32–7.16 (m, 5H), 6.19–6.14 (m, 1H), 5.93 (dd, J = 31.9, 8.9 Hz, 1H), 4.83–4.71 (m, 2H), 3.23–3.05 (m, 2H), 1.44–1.36 (m, 1H), 1.34–1.21 (m, 2H), 0.79–0.74 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 169.8, 163.0, 154.6 (d, J = 300.1 Hz), 147.7, 144.2, 144.0, 143.0, 137.2, 136.3, 134.7, 129.6, 129.4, 128.0, 129.7, 127.4, 117.9 (d, J = 5.3 Hz), 55.0, 43.6, 43.1, 38.8, 24.7, 22.4, 22.3. mp: 68–70 °C. $[\alpha]_D^{20}$ = +1 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3368, 3302, 3063, 3030, 2957, 2927, 2870, 1652, 1519, 1332, 1167, 1019, 752, 720, 685. ESI-MS: m/z calculated for [C$_{27}$H$_{29}$FN$_4$O$_4$S+H]$^+$ ([M+H]$^+$): 525.2, found: 525.1. Purity: 99%.

N-((S)-1-(((S,E)-5-Methyl-1-(phenylsulfonyl)hex-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)-pyrazine-2-carboxamide (58)
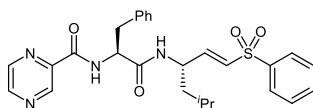
(1) Under argon atmosphere, LiAlH$_4$ (89 mg, 2.34 mmol, 2.0 equiv.) was added portionwise to a solution of **57** (500 mg, 1.17 mmol, 1.0 equiv.) in dry THF (20 mL) at 0 °C. After stirring for 1 h at 0 °C, the reaction mixture was quenched by addition of KHSO$_4$ solution (0.3 M, 50 mL), filtered, and extracted with ethyl acetate (3 × 50 mL). The combined organic extracts were washed with saturated NaHCO$_3$ solution (2 × 100 mL), dried over anhydrous Na$_2$SO$_4$, and the solvent was removed by distillation under reduced pressure at 40 °C. The peptidylaldehyde (ca. 200 mg) obtained in the form of an orange oil was directly used in the next step.

(2) To a solution of **3** (155 mg, 0.53 mmol, 1.0 equiv.) and DBU (79 μL, 0.53 mmol, 1.0 equiv.) in dry MeCN (20 mL) was added dried LiCl (27 mg, 0.64 mmol, 1.2 equiv.) at 0 °C under argon atmosphere. The solution was stirred for 20 min, and a solution of the peptidylaldehyde in dry MeCN (6 mL) was added. After stirring for 1 h at 0 °C and argon atmosphere, citric acid solution (10% in water, 60 mL) was added. The mixture was extracted with ethyl acetate (3 × 30 mL), and the combined organic extracts were filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C, and the residue was purified by preparative HPLC to yield the desired product ((S,S)-diastereomer) as a colorless solid (29 mg, 0.06 mmol, 11%). ^{1}H NMR (600 MHz, CDCl$_3$): δ / ppm = 9.33 (s, 1H), 8.77 (s, 1H), 8.54 (s, 1H), 8.32 (d, J = 8.0 Hz, 1H), 7.87–7.84 (m, 2H), 7.67–7.63 (m, 1H), 7.59–7.55 (m, 2H), 7.25–7.18 (m, 5H), 6.74 (dd, J = 15.1, 5.0 Hz, 1H), 6.07 (dd, J = 15.1, 1.6 Hz, 1H), 5.93–5.86 (m, 1H), 4.77–4.71 (m, 1H), 4.70–4.63 (m, 1H), 3.20–3.10 (m, 2H), 1.53–1.46 (m, 1H), 1.35–1.31 (m, 2H), 0.83–0.80 (m, 6H). ^{13}C NMR (150 MHz, CDCl$_3$): δ / ppm = 169.9, 163.3, 147.8, 145.8, 144.3, 143.9, 143.0, 140.2, 136.1, 133.7, 130.3, 129.5, 129.3, 129.1, 127.8, 127.6, 55.2, 48.2, 43.1, 38.3, 24.7, 22.8, 22.0. mp: 78–81 °C. $[\alpha]_D^{20}$ = −22 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3299, 3062, 2957, 2929, 2870, 1656, 1517, 1446, 1306, 1144, 1085, 1020, 751, 719, 687. ESI-MS: m/z calculated for [C$_{27}$H$_{30}$N$_4$O$_4$S+H]$^+$ ([M+H]$^+$): 507.2, found: 507.1. Purity: 99%.

tert-Butyl (1-(benzo[d]thiazol-2-yl)-4-methyl-1-oxopentan-2-yl)carbamate (60)
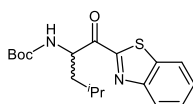
To a solution of benzothiazole (2.66 g, 19.68 mmol, 10.0 equiv.) in dry THF (50 mL) at −75 °C was added n-BuLi (2.5 M in n-hexane, 5.5 mL, 13.78 mmol, 7.0 equiv.) dropwise over 15 min. The mixture was stirred at −75 °C for 1 h, and a solution of **51** (540 mg, 1.97 mmol, 1.0 equiv.) in dry THF (10 mL) was added. After stirring was continued at −75 °C for 5 h, saturated NH$_4$Cl solution (80 mL) was added. The mixture was extracted with ethyl acetate (3 × 80 mL), and the combined organic extracts were dried

over anhydrous Na$_2$SO$_4$. The solvent was removed by distillation under reduced pressure at 40 °C, and the residue was purified by column chromatography (cyclohexane/ethyl acetate = 20:1) to yield the desired product as a yellowish solid (448 mg, 1.29 mmol, 65%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.22–8.17 (m, 1H), 8.01–7.95 (m, 1H), 7.62–7.49 (m, 2H), 5.71–5.54 (m, 1H), 5.33–5.18 (m, 1H), 1.90–1.77 (m, 2H), 1.60–1.50 (m, 1H), 1.43 (s, 9H), 1.09 (d, J = 6.0 Hz, 3H), 0.96 (d, J = 6.3 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 194.9, 164.2, 155.6, 153.7, 137.4, 128.0, 127.1, 125.95, 122.5, 80.0, 55.4, 42.2, 28.4, 25.4, 23.4, 21.9. mp: 101–103 °C. $[\alpha]_D^{20}$ = +33 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3362, 2965, 2930, 2871, 1702, 1679, 1518, 1481, 1366, 1230, 1162, 878, 828, 759, 731. ESI-MS: m/z calculated for [C$_{18}$H$_{24}$N$_2$O$_3$S+Na]$^+$ ([M+Na]$^+$): 371.1, found: 371.1.

tert-Butyl ((2_S_)-1-((1-(benzo[d]thiazol-2-yl)-4-methyl-1-oxopentan-2-yl)amino)-1-oxo-3-phenyl-propan-2-yl)carbamate (61)
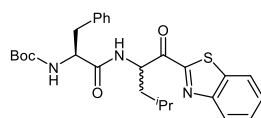
(1) **60** (274 mg, 0.79 mmol) was treated with 4 M HCl in dioxane (4 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C to yield the deprotected amine hydrochloride as a colorless solid (223 mg, 0.79 mmol, quantitative), which was directly used in the next step.

(2) To a 0 °C cold solution of Boc-L-Phe-OH (208 mg, 0.78 mmol, 1.0 equiv.) in DCM (6 mL) were added HOBt • H$_2$O (120 mg, 0.78 mmol, 1.0 equiv.) and 2,4,6-collidine (208 μL, 1.57 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (252 g, 0.78 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and the deprotected amine hydrochloride from step 1 (223 mg, 0.78 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (30 mL). The mixture was washed with saturated NaHCO$_3$ solution (3 × 25 mL) as well as 1 M HCl (3 × 25 mL) and filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C to yield the desired product as a yellowish solid (264 mg, 0.53 mmol, 68%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.20–8.13 (m, 1H), 8.00–7.94 (m, 1H), 7.62–7.50 (m, 2H), 7.24–6.95 (m, 5H), 6.79 (d, J = 8.5 Hz, 1H), 5.79–5.65 (m, 1H), 5.23–5.02 (m, 1H), 4.48–4.31 (m, 1H), 3.14–2.95 (m, 2H), 1.84–1.51 (m, 3H), 1.4 –1.39 (m, 9H), 1.03 (d, J = 6.2 Hz, 3H), 0.90 (d, J = 6.3 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 193.4 (diastereomer B), 192.9 (diastereomer A), 171.1 (diastereomer B), 171.0 (diastereomer A), 164.0, 155.5, 153.6 (diastereomer B), 153.5 (diastereomer A), 137.3 (diastereomer B), 137.2 (diastereomer A), 136.7, 129.3, 128.7 (diastereomer B), 128.6 (diastereomer A), 128.1, 127.2, 127.0 (diastereomer B), 126.8 (diastereomer A), 125.9, 122.4, 80.3, 56.3, 54.6 (diastereomer A), 54.3 (diastereomer B), 42.2 (diastereomer A), 41.8 (diastereomer B), 38.7 (diastereomer A), 38.6 (diastereomer B), 28.4, 25.2, 23.3 (diastereomer B), 23.2 (diastereomer

A), 21.9 (diastereomer A), 21.7 (diastereomer B). mp: 133–135 °C. $[\alpha]_D^{20} = -13$ (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3275, 3065, 2962, 2930, 2870, 1686, 1639, 1518, 1366, 1248, 1169, 885, 758, 726, 699. ESI-MS: *m/z* calculated for [C$_{27}$H$_{33}$N$_3$O$_4$S+Na]$^+$ ([M+Na]$^+$): 518.2, found: 518.0.

N-((*S*)-1-(((*S*)-1-(Benzo[*d*]thiazol-2-yl)-4-methyl-1-oxopentan-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (62)
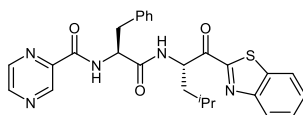
(1) **61** (223 mg, 0.45 mmol) was treated with 4 M HCl in dioxane (4 mL) at room temperature for 1 h. The solvent was removed under reduced pressure at 40 °C to yield the deprotected amine hydrochloride as a colorless solid (194 mg, 0.45 mmol, quantitative), which was directly used in the next step.

(2) To a 0 °C cold solution of pyrazinecarboxylic acid (56 mg, 0.45 mmol, 1.0 equiv.) in DCM (6 mL) were added HOBt • H$_2$O (69 mg, 0.45 mmol, 1.0 equiv.) and 2,4,6-collidine (120 μL, 0.90 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (145 g, 0.45 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and the deprotected amine hydrochloride from step 1 (194 mg, 0.45 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (30 mL). The mixture was washed with saturated NaHCO$_3$ solution (3 × 25 mL) as well as 1 M HCl (3 × 25 mL) and filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C, and the residue was purified by column chromatography (cyclohexane/ethyl acetate = 2:1) to yield the desired product as a colorless solid (mixture of two diastereomers, $\sum$ = 119 mg, 0.24 mmol, 53%). Separation of the diastereomers was achieved by preparative HPLC. (*S,S*)-diastereomer: ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.35 (d, *J* = 1.5 Hz, 1H), 8.75 (d, *J* = 2.5 Hz, 1H), 8.53 (dd, *J* = 2.5, 1.5 Hz, 1H), 8.44 (d, *J* = 8.2 Hz, 1H), 8.13–8.08 (m, 1H), 8.01–7.96 (m, 1H), 7.61–7.50 (m, 2H), 7.25–7.19 (m, 2H), 7.14–7.07 (m, 2H), 7.03–6.96 (m, 1H), 6.77 (d, *J* = 8.3 Hz, 1H), 5.72–5.63 (m, 1H), 4.96–4.87 (m, 1H), 3.28–3.08 (m, 2H), 1.82–1.75 (m, 1H), 1.67–1.53 (m, 2H), 0.97 (d, *J* = 6.3 Hz, 3H), 0.87 (d, *J* = 6.3 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 192.7, 170.1, 164.0, 163.0, 153.5, 147.6, 144.4, 144.2, 142.9, 137.2, 136.4, 129.3, 128.7, 128.1, 127.2, 127.0, 125.9, 122.5, 55.1, 54.9, 42.1, 38.9, 25.3, 23.1, 22.0. mp: 65–67 °C. $[\alpha]_D^{20} = -12$ (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3307, 3061, 2956, 2925, 2869, 1651, 1519, 1370, 1207,1155, 1019, 882, 760, 730, 699. ESI-MS: *m/z* calculated for [C$_{27}$H$_{27}$N$_5$O$_3$S+H]$^+$ ([M+H]$^+$): 502.2, found: 502.1. Purity: 99%.

Methyl (*tert*-butoxycarbonyl)-L-phenylalanyl-L-leucinate (69)
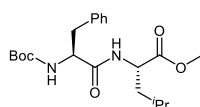
To a 0 °C cold solution of Boc-L-Phe-OH (1.34 mg, 5.04 mmol, 1.0 equiv.) in DCM (20 mL) were added HOBt • H_2O (0.77 g, 5.04 mmol, 1.0 equiv.) and 2,4,6-collidine (1.34 mL, 1.20 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (1.62 g, 5.04 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and H_2N-L-Leu-OMe • HCl (**63**, 0.92 g, 5.04 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (100 mL). The mixture was washed with saturated $NaHCO_3$ solution (3 × 80 mL) as well as 1 M HCl (3 × 80 mL) and filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C to yield the desired product as a colorless solid (1.63 mg, 4.15 mmol, 82%).

^{1}H NMR (300 MHz $CDCl_3$): δ / ppm = 7.32–7.14 (m, 5H), 6.33 (d, J = 8.2 Hz, 1H), 5.12–4.95 (m, 1H), 4.62–4.48 (m, 1H), 4.42–4.27 (m, 1H), 3.68 (s, 3H), 3.06 (d, J = 6.8 Hz, 2H), 1.64–1.43 (m, 3H), 1.40 (s, 9H), 0.89 (t, J = 5.6 Hz, 6H). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ / ppm = 173.0, 171.1, 155.5, 136.7, 129.5, 128.8, 127.1, 80.4, 55.9, 52.4, 50.9, 41.7, 38.2, 28.4, 24.8, 22.9, 22.0. mp: 98–100 °C. $[\alpha]_D^{20}$ = −14 (10 mg/mL; $CHCl_3$). FT-IR: ν / cm^{-1} = 3289, 3067, 2956, 2871, 1747, 1682, 1647, 1532, 1440, 1366, 1251, 1169, 1048, 1027, 702. ESI-MS: m/z calculated for $[C_{21}H_{32}N_2O_5+Na]^+$ ([M+Na]$^+$): 415.2, found: 415.1.

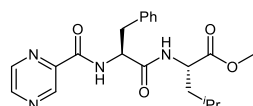**Methyl (pyrazine-2-carbonyl)-L-phenylalanyl-L-leucinate (70)**

(1) **69** (1.44 g, 3.67 mmol) was treated with 4 M HCl in dioxane (8 mL) at room temperature for 1.5 h. The solvent was removed under reduced pressure at 40 °C to yield the deprotected amine hydrochloride as a colorless foam (1.21 g, 3.67 mmol, quantitative), which was directly used in the next step.

(2) To a 0 °C cold solution of pyrazinecarboxylic acid (0.44 g, 3.55 mmol, 1.0 equiv.) in DCM (20 mL) were added HOBt • H_2O (0.54 g, 3.55 mmol, 1.0 equiv.) and 2,4,6-collidine (940 μL, 7.09 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (1.14 g, 3.55 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and the deprotected amine hydrochloride from step 1 (1.17 g, 3.55 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (70 mL). The mixture was washed with saturated $NaHCO_3$ solution (3 × 50 mL) as well as 1 M HCl (3 × 50 mL) and filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C to yield the desired product as a colorless solid (1.15 mg, 2.89 mmol, 82%). ^{1}H NMR (300 MHz, $CDCl_3$): δ / ppm = 9.35 (d, J = 1.5 Hz, 1H), 8.75 (d, J = 2.5 Hz, 1H), 8.55–8.51 (m, 1H), 8.39 (d, J = 8.3 Hz,

1H), 7.28–7.19 (m, 5H), 6.44 (d, $J = 8.0$ Hz, 1H), 4.98–4.88 (m, 1H), 4.61–4.50 (m, 1H), 3.71 (s, 3H), 3.21 (d, $J = 7.0$ Hz, 2H), 1.62–1.40 (m, 3H), 0.88–0.82 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 172.8, 170.2, 163.0, 147.5, 144.3, 144.1, 143.0, 136.4, 129.5, 128.8, 127.2, 54.6, 52.4, 51.1, 41.6, 38.5, 24.9, 22.7, 22.1. mp: 50–53 °C. $[\alpha]_D^{20} = -29$ (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3294, 3063, 3030, 2955, 2870, 1742, 1651, 1519, 1438, 1370, 1200, 1152, 1019, 746, 699. ESI-MS: m/z calculated for [C$_{21}$H$_{26}$N$_4$O$_4$+Na]$^+$ ([M+Na]$^+$): 421.2, found: 421.1.

(P$_{YRAZINE}$-2-carbonyl)-L-phenylalanyl-L-leucine (71)
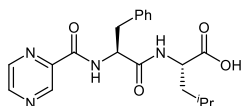
To a solution of **70** (1.09 g, 2.72 mmol, 1.0 equiv.) in THF (15 mL) and water (15 mL) was added LiOH • H$_2$O (0.34 g, 8.17 mmol, 3.0 equiv.). After stirring at room temperature for 17 h, THF was removed by distillation under reduced pressure at 40 °C, and the aqueous solution was acidified to pH = 1 with 1 M HCl. The resulting suspension was extracted with CHCl$_3$, and the combined organic extracts were dried over anhydrous Na$_2$SO$_4$. After removing the solvent by distillation under reduced pressure at 40 °C, the title compound was obtained as a colorless foam (1.05 g, 2.72 mmol, quantitative). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.31 (d, $J = 1.4$ Hz, 1H), 8.74 (d, $J = 2.5$ Hz, 1H), 8.56–8.54 (m, 1H), 8.51 (d, $J = 8.5$ Hz, 1H), 7.25–7.07 (m, 6H), 6.98 (d, $J = 8.0$ Hz, 1H), 5.10–5.00 (m, 1H), 4.64–4.53 (m, 1H), 3.29–3.13 (m, 2H), 1.71–1.47 (m, 3H), 0.88–0.82 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 175.5, 170.8, 163.1, 147.2, 144.3, 144.0, 143.3, 136.3, 129.5, 128.7, 127.2, 54.6, 51.2, 41.3, 38.6, 24.9, 22.8, 22.0. mp: 62–64 °C. $[\alpha]_D^{20} = -3$ (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3297, 3064, 3031, 2956, 2870, 1724, 1648, 1522, 1467, 1199, 1151, 1020, 866, 746, 698. ESI-MS: m/z calculated for [C$_{20}$H$_{24}$N$_4$O$_4$+H]$^+$ ([M+H]$^+$): 385.2, found: 385.1.

N-((S)-1-(((S)-4-Methyl-1-oxo-1-(((S)-2-oxo-1-phenylazetidin-3-yl)amino)pentan-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (72)
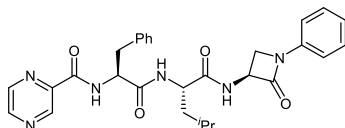
To a 0 °C cold solution of **72** (131 mg, 0.34 mmol, 1.0 equiv.), HOBt • H$_2$O (52 mg, 0.34 mmol, 1.0 equiv.), and TBTU (109 mg, 0.34 mmol, 1.0 equiv.) in DCM (20 mL) were added **9** (68 mg, 0.34 mmol, 1.0 equiv.) and 2,4,6-collidine (90 μL, 0.68 mmol, 2.0 equiv.), and the mixture was stirred at room temperature overnight. The solvent was removed under reduced pressure at 40 °C, and the residue was purified by column chromatography (DCM/MeOH 30:1) to yield the desired product as a colorless solid (117 mg, 0.22 mmol, 65%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.27 (d, $J = 1.4$ Hz, 1H), 8.69 (d, $J = 2.5$ Hz, 1H), 8.46–8.40 (m, 1H), 8.35 (d, $J = 7.5$ Hz, 1H), 7.51 (d, $J = 7.5$ Hz, 1H), 7.33–7.27 (m,

4H), 7.25–7.13 (m, 5H), 7.12–7.05 (m, 1H), 7.00 (d, J = 8.1 Hz, 1H), 4.97–4.85 (m, 2H), 4.55–4.45 (m, 1H), 3.89–3.82 (m, 1H), 3.59 (dd, J = 5.8, 2.7 Hz, 1H), 3.24–3.18 (m, 2H), 1.73–1.43 (m, 3H), 0.84–0.76 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 172.4, 170.9, 163.8, 163.6, 147.7, 144.3, 143.8, 142.9, 137.8, 136.2, 129.4, 129.3, 128.9, 127.3, 124.5, 116.8, 55.62, 55.0, 52.0, 47.1, 40.5, 38.0, 24.8, 23.0, 21.9. mp: 187–189 °C. $[\alpha]_D^{20}$ = −15 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3390, 3280, 3066, 2957, 2870, 1757, 1653, 1598, 1498, 1464, 1387, 1147, 1020, 750, 693. ESI-MS: m/z calculated for [C$_{29}$H$_{32}$N$_6$O$_4$+H]$^+$ ([M+H]$^+$): 529.3, found: 529.2. Purity: 95%.

N-((*S*)-1-(((*S*)-1-Amino-4-methyl-1-oxopentan-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carboxamide (73)
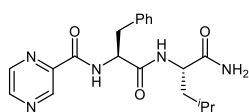
To a 0 °C cold solution of **70** (365 mg, 0.92 mmol) in MeOH (5 mL) was added NH$_3$ solution (7 N in MeOH, 12 mL). The solution was stirred at 0 °C for 15 min and at room temperature for 24 h. Again, NH$_3$ solution (7 N in MeOH, 12 mL) was added, and the mixture was stirred at room temperature for an additional 24 h. The solvent was removed by distillation to yield the title compound as a colorless solid (315 mg, 0.82 mmol, 89%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.16–9.07 (m, 1H), 8.87 (d, J = 2.6 Hz, 1H), 8.77–8.62 (m, 2H), 8.27 (d, J = 8.3 Hz, 1H), 7.37–7.29 (m, 1H), 7.28–7.09 (m, 5H), 7.07–6.93 (m, 1H), 4.89–4.76 (m, 1H), 4.36–4.24 (m, 1H), 3.23–3.02 (m, 2H), 1.67–1.40 (m, 3H), 0.92–0.78 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 173.9, 170.0, 162.3, 147.8, 144.1, 143.4, 137.3, 129.3, 128.1, 126.4, 53.9, 50.9, 41.1, 37.5, 24.3, 23.0, 21.7. mp: 201–203 °C. $[\alpha]_D^{20}$ = −4 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3368, 3308, 3195, 3054, 2959, 1686, 1645, 1530, 1449, 1398, 1227, 1020, 774, 746, 693. ESI-MS: m/z calculated for [C$_{20}$H$_{25}$N$_5$O$_3$+Na]$^+$ ([M+Na]$^+$): 406.2, found: 406.1.

N-((*S*)-1-(((*S*)-1-Cyano-3-methylbutyl)amino)-1-oxo-3-phenylpropan-2-yl)pyrazine-2-carbox-amide (74)
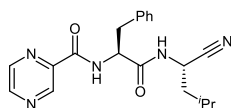
To a 0 °C cold solution of **73** (211 mg, 0.55 mmol, 1.0 equiv.) in dry DMF (4 mL) was added cyanuric chloride (112 mg, 0.61 mmol, 1.1 equiv.) portionwise. The solution was stirred at 0 °C for 1 h and at room temperature overnight. An additional quantity of cyanuric chloride (112 mg, 0.61 mmol, 1.1 equiv.) was added, and the mixture was stirred at room temperature for an additional 24 h. The solvent was removed by distillation under reduced pressure at 40 °C, and the residue was purified by reversed phase flash chromatography (MeCN/H$_2$O = 10:90 → 100:0) to yield the desired product as a yellowish solid

(95 mg, 0.26 mmol, 47%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 9.33 (d, J = 1.5 Hz, 1H), 8.78 (d, J = 2.5 Hz, 1H), 8.55 (dd, J = 2.5, 1.5 Hz, 1H), 8.38 (d, J = 8.1 Hz, 1H), 7.34–7.19 (m, 5H), 6.54 (d, J = 8.2 Hz, 1H), 4.86–4.74 (m, 2H), 3.22 (d, J = 7. 1 Hz, 2H), 1.77–1.45 (m, 3H), 0.93–0.85 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 170.1, 163.4, 147.9, 144.4, 143.7, 143.0, 135.9, 129.4, 129.1, 127.5, 118.2, 54.9, 41.7, 39.2, 38.4, 24.9, 22.3, 21.8. mp: 46 °C (decomposition). $[\alpha]_D^{20}$ = −32 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3288, 3061, 2959, 2872, 2247, 1657, 1519, 1467, 1402, 1236, 1154, 1020, 910, 729, 700. ESI-MS: m/z calculated for [C$_{20}$H$_{23}$N$_5$O$_2$+H]$^+$ ([M+H]$^+$): 366.2, found: 366.0. Purity: 95%.

(*tert*-Butoxycarbonyl)-L-leucine-1-^{13}C (113)
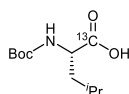
Adapted from DUTTON *et al.*, *J. Med. Chem.* **2003**, *46*, 2057–2073.

To a 0 °C cold solution of H-Leu-L-^{13}C-OH (0.98 g, 7.42 mmol, 1.0 equiv.) in 1,4-dioxane (5 mL) and water (3 mL) were added NaOH solution (1 M, 15 mL) and Boc$_2$O (3.66 g, 16.77 mmol, 2.26 equiv.). The solution was stirred at room temperature overnight, and water (30 mL) was added. The mixture was washed with *n*-pentane (4 × 80 mL) and acidified to pH = 1 with 1 M HCl. The resulting suspension was extracted with ethyl acetate (4 × 50 mL), and the combined organic extract were dried over anhydrous Na$_2$SO$_4$. After removing the solvent by distillation under reduced pressure at 40 °C, the desired product was obtained as a colorless oil (1.72 g, 7.42 mmol, quantitative). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 4.89 (d, J = 8.5 Hz, 1H), 4.41–4.05 (m, 1H), 1.82–1.50 (m, 3H), 1.45 (s, 9H), 0.95 (d, J = 6.2 Hz, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 178.3, 155.9, 80.4, 52.16 (d, J = 58.6 Hz), 41.6, 28.4, 24.9, 23.0, 21.9. $[\alpha]_D^{20}$ = −5 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3446, 3333, 2959, 2940, 2870, 2561, 1684, 1533, 1392, 1366, 1237, 1161, 1048, 872, 757. ESI-MS: m/z calculated for [C$_{10}$^{13}CH$_{21}$NO$_4$+Na]$^+$ ([M+Na]$^+$): 255.1, found: 255.0.

tert-Butyl (*S*)-(1-(methoxy(methyl)amino)-4-methyl-1-oxopentan-2-yl-1-^{13}C)carbamate (114)
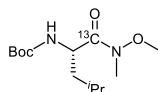
To a 0 °C cold solution of Boc-Leu-L-^{13}C-OH (**113**, 1.70 g, 7.32 mmol, 1.0 equiv.) in DCM (80 mL) were added HOBt • H$_2$O (1.12 g, 7.32 mmol, 1.0 equiv.) and 2,4,6-collidine (1.94 mL, 14.64 mmol, 2.0 equiv.). After stirring at 0 °C for 30 min, TBTU (2.35 g, 7.32 mmol, 1.0 equiv.) was added. The solution was stirred for another 30 min at 0 °C, and *N,O*-dimethylhydroxylamine • HCl (0.71 g, 7.32 mmol, 1.0 equiv.) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure at 40 °C, and the residue was taken up in ethyl acetate (100 mL). The mixture

was washed with saturated NaHCO$_3$ solution (3 × 80 mL) as well as 1 M HCl (3 × 80 mL) and filtered over silica. The filtrate was concentrated under reduced pressure at 40 °C to yield the desired product as a colorless oil (1.91 g, 6.92 mmol, 95%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 5.14–4.95 (m, 1H), 4.79–4.60 (m, 1H), 3.77 (s, 3H), 3.22–3.15 (m, 3H), 1.78–1.61 (m, 1H), 1.49–1.37 (m, 11H), 0.97–0.88 (m, 6H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm = 174.0, 155.8, 79.6, 61.7, 49.10 (d, J = 54.6 Hz), 42.2, 32.3, 28.5, 24.9, 23.5, 21.0. $[\alpha]_D^{20}$ = −10 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3321, 2957, 2935, 2870, 1708, 1619, 1499, 1365, 1250, 1165, 1045, 1017, 986, 875, 758. ESI-MS: m/z calculated for [C$_{12}$^{13}CH$_{26}$N$_2$O$_4$+Na]$^+$ ([M+Na]$^+$): 298.2, found: 298.1.

tert-Butyl (1-(benzo[d]thiazol-2-yl)-4-methyl-1-oxopentan-2-yl-1-^{13}C)carbamate (115)
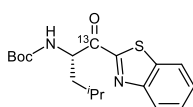
To a solution of benzothiazole (2.65 g, 19.61 mmol, 10.0 equiv.) in dry THF (50 mL) at −75 °C was added n-BuLi (2.5 M in hexanes, 5.5 mL, 13.73 mmol, 7.0 equiv.) dropwise over 15 min. The mixture was stirred at −75 °C for 1 h and a solution of 114 (540 mg, 1.96 mmol, 1 equiv.) in dry THF (10 mL) was added. After stirring was continued at −75 °C for 2.5 h, saturated NH$_4$Cl solution (30 mL) was added. The mixture was extracted with ethyl acetate (3 × 80 mL), and the combined organic extracts were filtered over a small silica column. The filtrate was concentrated by distillation under reduced pressure at 40 °C, and the residue was purified by column chromatography (cyclohexane/ethyl acetate = 20:1) to yield the desired product as a yellowish solid (380 mg, 1.09 mmol, 56%). ^{1}H NMR (300 MHz, CDCl$_3$): δ / ppm = 8.22–8.16 (m, 1H), 8.00–7.94 (m, 1H), 7.61–7.49 (m, 2H), 5.76–5.52 (m, 1H), 5.40–5.18 (m, 1H), 1.88–1.77 (m, 2H), 1.60–1.50 (m, 1H), 1.43 (s, 9H), 1.09 (d, J = 6.0 Hz, 3H), 0.96 (d, J = 6.2 Hz, 3H). ^{13}C NMR (75.5 MHz, CDCl$_3$): δ / ppm =194.9, 164.2 (d, J = 64.8 Hz), 155.6, 153.7 (d, J = 7.8 Hz), 137.4, 128.0, 127.1, 125.9, 122.5, 80.0, 55.4 (d, J = 46.0 Hz), 42.2, 28.4, 25.4, 23.4, 21.8. mp: 108–110 °C. $[\alpha]_D^{20}$ = +40 (10 mg/mL; CHCl$_3$). FT-IR: ν / cm^{-1} = 3364, 3066, 2965, 2929, 2871, 1682, 1659, 1518, 1480, 1366, 1161, 876, 824, 760, 731. ESI-MS: m/z calculated for [C$_{17}$^{13}CH$_{24}$N$_2$O$_3$S+Na]$^+$ ([M+Na]$^+$): 372.1, found: 372.0. Purity: 98%.

Part IV

▶ **References and appendix**

8 References

(1)   Sutanto, F. et al. *RSC Med. Chem.* **2020**, *11* (8), 876–884.

(2)   Abdeldayem, A. et al. *Chem. Soc. Rev.* **2020**, *49* (9), 2617–2687.

(3)   Lonsdale, R. et al. *Chem. Soc. Rev.* **2018**, *47* (11), 3816–3830.

(4)   Boike, L. et al. *Nat. Rev. Drug Discov.* **2022**, *21* (12), 881–898.

(5)   Adeniyi, A. A. et al. *Expert Opin. Drug Discov.* **2016**, *11* (1), 79–90.

(6)   Martin, J. S. et al. *Bioorg. Med. Chem.* **2019**, *27* (10), 2066–2074.

(7)   Singh, J. et al. *Nat. Rev. Drug Discov.* **2011**, *10* (4), 307–317.

(8)   Tuley, A. et al. *Biochemistry* **2018**, *57* (24), 3326–3337.

(9)   Aljoundi, A. et al. *Protein J.* **2020**, *39* (2), 97–105.

(10)   Bauer, R. A. *Drug Discov. Today* **2015**, *20* (9), 1061–1073.

(11)   Shimokawa, T. et al. *J. Biol. Chem.* **1992**, *267* (17), 12387–12392.

(12)   Fleming, A. *Br. J. Exp. Pathol.* **1929**, *10* (3), 226–236.

(13)   Thomas, D. M. et al. *Clin. Exp. Pharmacol. Physiol.* **1998**, *25* (11), 887–895.

(14)   Flanagan, M. E. et al. *J. Med. Chem.* **2014**, *57* (23), 10072–10079.

(15)   Oballa, R. M. et al. *Bioorg. Med. Chem. Lett.* **2007**, *17* (4), 998–1002.

(16)   Smith, J. M. et al. *J. Comput. Aided. Mol. Des.* **2015**, *29* (8), 725–735.

(17)   Adams, J. et al. *Bioorg. Med. Chem. Lett.* **1998**, *8* (4), 333–338.

(18)   Kisselev, A. F. et al. *Chem. Biol.* **2012**, *19* (1), 99–115.

(19)   Siegel, D. S. et al. *Blood* **2012**, *120* (14), 2817–2825.

(20)   Tahrani, A. A. et al. *Adv. Ther.* **2009**, *26* (3), 249–262.

(21)   Foote, B. S. et al. *Ann. Pharmacother.* **2011**, *45* (9), 1085–1093.

(22)   Baillie, T. A. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (43), 13408–13421.

(23)   Gehringer, M. et al. *J. Med. Chem.* **2019**, *62* (12), 5673–5724.

(24)   Leproult, E. et al. *J. Med. Chem.* **2011**, *54* (5), 1347–1355.

(25)   Ghosh, A. K. et al. *ChemMedChem* **2019**, *14* (9), 889–906.

(26)   De Cesco, S. et al. *Eur. J. Med. Chem.* **2017**, *138*, 96–114.

(27)   Renaud, J.-P. et al. *Nat. Rev. Drug Discov.* **2018**, *17* (7), 471–492.

(28)   Softley, C. A. et al. *J. Biomol. NMR* **2020**, *74* (6–7), 287–309.

(29)   Baell, J. et al. *Nature* **2014**, *513* (7519), 481–483.

(30)   Dahlin, J. L. et al. *Assay Drug Dev. Technol.* **2016**, *14* (3), 168–174.

(31)   Baell, J. B. et al. *J. Med. Chem.* **2010**, *53* (7), 2719–2740.

(32)   Capuzzi, S. J. et al. *J. Chem. Inf. Model.* **2017**, *57* (3), 417–427.

(33)   Hartenfeller, M. et al. *Methods Mol. Biol.* **2011**, *672*, 299–323.

(34)   Callis, T. B. et al. *J. Med. Chem.* **2022**, *65* (20), 13483–13504.

(35)   Scarpino, A. et al. *Curr. Pharm. Des.* **2020**, *26* (44), 5684–5699.

(36)   Kumalo, H. M. et al. *Molecules* **2015**, *20* (2), 1984–2000.

(37)   Zhang, S. et al. *J. Mol. Model.* **2009**, *15* (12), 1481–1490.

(38)   Schmidt, T. C. et al. *Chemphyschem* **2014**, *15* (15), 3226–3235.

(39)   Fan, J. et al. *Biochem. Pharmacol.* **2014**, *87* (1), 93–120.

(40)   Zhu, K. et al. *J. Chem. Inf. Model.* **2014**, *54* (7), 1932–1940.

(41)   Lonsdale, R. et al. *J. Chem. Inf. Model.* **2017**, *57* (12), 3124–3137.

(42)   Powers, J. C. et al. *Chem. Rev.* **2002**, *102* (12), 4639–4750.

(43)   Pettinger, J. et al. *Angew. Chem. Int. Ed. Engl.* **2017**, *56* (48), 15200–15209.

(44)   Tizón, L. et al. *Org. Biomol. Chem.* **2015**, *13* (3), 706–716.

(45)   Cross, B. C. S. et al. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (15), E869-78.

(46)   Hett, E. C. et al. *ACS Chem. Biol.* **2015**, *10* (4), 1094–1098.

(47)   Woodward, R. B. et al. *J. Am. Chem. Soc.* **1961**, *83* (4), 1010–1012.

(48)   Dalton, S. E. et al. *Chembiochem* **2020**, *21* (8), 1080–1100.

(49)   Degel, B. et al. *ChemMedChem* **2008**, *3* (2), 302–315.

(50)   Kharenko, O. A. et al. *J. Med. Chem.* **2018**, *61* (18), 8202–8211.

(51)   Way, J. C. *Curr. Opin. Chem. Biol.* **2000**, *4* (1), 40–46.

(52)   Gan, J. et al. *Chem. Res. Toxicol.* **2009**, *22* (4), 690–698.

(53)   McAulay, K. et al. *Pharmaceuticals* **2022**, *15* (11).

(54)   Wilmouth, R. C. et al. *Biochemistry* **1999**, *38* (25), 7989–7998.

(55)   Owen, D. R. et al. *Science* **2021**, *374* (6575), 1586–1593.

(56)   Han, Z. et al. *ChemMedChem* **2016**, *11* (6), 585–599.

(57)   Brogi, S. et al. *Molecules* **2022**, *27* (8).

(58)   Fleming, F. F. et al. *J. Med. Chem.* **2010**, *53* (22), 7902–7917.

292

(59)   Dalton, S. E. et al. *J. Am. Chem. Soc.* **2018**, *140* (3), 932–939.

(60)   Lin, S. et al. *Science* **2017**, *355* (6325), 597–602.

(61)   Bizet, V. et al. *Chem. Soc. Rev.* **2015**, *44* (11), 3378–3390.

(62)   Brahms, A. et al. *Water*. 2021.

(63)   Dražić, T. et al. *J. Med. Chem.* **2020**, *63* (1), 140–156.

(64)   Serafimova, I. M. et al. *Nat. Chem. Biol.* **2012**, *8* (5), 471–476.

(65)   Schirmeister, T. et al. *J. Am. Chem. Soc.* **2016**, *138* (27), 8332–8335.

(66)   Johe, P. et al. *ACS Chem. Biol.* **2021**, *16* (4), 661–670.

(67)   Ludewig, S. et al. *Curr. Top. Med. Chem.* **2010**, *10* (3), 368–382.

(68)   McWhirter, C. In *The Design of Covalent-Based Inhibitors*; Ward, R. A., Grimster, N. P. B. T.-A. R. in M. C., Eds.; Academic Press, 2021; Vol. 56, pp 1–31.

(69)   Morrison, J. F. et al. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1988**, *61*, 201–301.

(70)   Strelow, J. M. *SLAS Discov. Adv. life Sci. R D* **2017**, *22* (1), 3–20.

(71)   Cheng, Y. et al. *Biochem. Pharmacol.* **1973**, *22* (23), 3099–3108.

(72)   Kuzmič, P. *bioRxiv* **2022**, 2020.06.25.171207.

(73)   Copeland, R. A. *Methods Biochem. Anal.* **2005**, *46*, 1–265.

(74)   Copeland, R. A. et al. *Nat. Rev. Drug Discov.* **2006**, *5* (9), 730–739.

(75)   Copeland, R. A. *Expert Opin. Drug Discov.* **2021**, *16* (12), 1441–1451.

(76)   Copeland, R. A. *Nat. Rev. Drug Discov.* **2016**, *15* (2), 87–95.

(77)   Smith, A. J. T. et al. *J. Med. Chem.* **2009**, *52* (2), 225–233.

(78)   Swinney, D. C. *Curr. Top. Med. Chem.* **2006**, *6* (5), 461–478.

(79)   Cantoni, G. L. *Annu. Rev. Biochem.* **1975**, *44*, 435–451.

(80)   Fontecave, M. et al. *Trends Biochem. Sci.* **2004**, *29* (5), 243–249.

(81)   Lin, C. et al. *Infect. Disord. Drug Targets* **2006**, *6* (1), 3–16.

(82)   Njoroge, F. G. et al. *Acc. Chem. Res.* **2008**, *41* (1), 50–59.

(83)   Kwak, E. L. et al. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (21), 7665–7670.

(84)   Berger, S. L. et al. *Genes Dev.* **2009**, *23* (7), 781–783.

(85)   Saletore, Y. et al. *Genome Biol.* **2012**, *13* (10), 175.

(86)   DAVIS, F. F. et al. *J. Biol. Chem.* **1957**, *227* (2), 907–915.

(87)   Boccaletto, P. et al. *Nucleic Acids Res.* **2022**, *50* (D1), D231–D235.

(88)   Li, S. et al. *Annu. Rev. Genomics Hum. Genet.* **2014**, *15*, 127–150.

(89)   Meyer, K. D. et al. *Nat. Rev. Mol. Cell Biol.* **2014**, *15* (5), 313–326.

(90)   Kirchner, S. et al. *Nat. Rev. Genet.* **2015**, *16* (2), 98–112.

(91)   Roundtree, I. A. et al. *Cell* **2017**, *169* (7), 1187–1200.

(92)   Gilbert, W. V et al. *Science* **2016**, *352* (6292), 1408–1412.

(93)   Jaffrey, S. R. *Nat. Struct. Mol. Biol.* **2014**, *21* (11), 945–946.

(94)   Song, C.-X. et al. *Nat. Biotechnol.* **2012**, *30* (11), 1107–1116.

(95)   Xuan, J.-J. et al. *Nucleic Acids Res.* **2018**, *46* (D1), D327–D334.

(96)   Helm, M. et al. *Nat. Rev. Genet.* **2017**, *18* (5), 275–291.

(97)   Motorin, Y. et al. *Wiley Interdiscip. Rev. RNA* **2022**, *13* (1), e1691.

(98)   Esteve-Puig, R. et al. *Cancer Lett.* **2020**, *474*, 127–137.

(99)   Schaefer, M. et al. *Open Biol.* **2017**, *7* (5).

(100)   Cayir, A. et al. *Environ. Res.* **2020**, *189*, 109885.

(101)   Jonkhout, N. et al. *RNA* **2017**, *23* (12), 1754–1769.

(102)   Barbieri, I. et al. *Nat. Rev. Cancer* **2020**, *20* (6), 303–322.

(103)   Fischer, T. R. et al. *Nucleic Acids Res.* **2022**, *50* (8), 4216–4245.

(104)   Kuznetsova, S. A. et al. *Biochemistry. (Mosc).* **2019**, *84* (8), 851–869.

(105)   Bohnsack, K. E. et al. *Genes (Basel).* **2019**, *10* (2).

(106)   Chen, Y. et al. *Biochemistry* **1993**, *32* (38), 10249–10253.

(107)   Schubert, H. L. et al. *Trends Biochem. Sci.* **2003**, *28* (6), 329–335.

(108)   Goll, M. G. et al. *Science* **2006**, *311* (5759), 395–398.

(109)   Yoder, J. A. et al. *Hum. Mol. Genet.* **1998**, *7* (2), 279–284.

(110)   Jurkowski, T. P. et al. *PLoS One* **2011**, *6* (11), e28104.

(111)   Jeltsch, A. et al. *RNA Biol.* **2017**, *14* (9), 1108–1123.

(112)   Kaiser, S. et al. *RNA Biol.* **2017**, *14* (9), 1241–1251.

(113)   Hermann, A. et al. *J. Biol. Chem.* **2003**, *278* (34), 31717–31721.

(114)   Brzezicha, B. et al. *Nucleic Acids Res.* **2006**, *34* (20), 6034–6043.

(115)   Haag, S. et al. *RNA* **2015**, *21* (9), 1532–1543.

(116)   Liu, R.-J. et al. *Nucleic Acids Res.* **2017**, *45* (11), 6684–6697.

(117)   Blanco, S. et al. *EMBO J.* **2014**, *33* (18), 2020–2039.

(118) Tuorto, F. et al. *Nat. Struct. Mol. Biol.* **2012**.

(119) Auxilien, S. et al. *RNA Biol.* **2012**, *9* (11), 1331–1338.

(120) Shinoda, S. et al. *Nucleic Acids Res.* **2019**, *47* (16), 8734–8745.

(121) Yang, X. et al. *Cell Res.* **2017**, *27* (5), 606–625.

(122) Cámara, Y. et al. *Cell Metab.* **2011**, *13* (5), 527–539.

(123) Van Haute, L. et al. *Nat. Commun.* **2016**, *7*, 1–10.

(124) Haag, S. et al. *EMBO J.* **2016**, *35* (19), 2104–2119.

(125) Nakano, S. et al. *Nat. Chem. Biol.* **2016**, *12* (7), 546–551.

(126) Metodiev, M. D. et al. *PLoS Genet.* **2014**, *10* (2), e1004110.

(127) Aguilo, F. et al. *Cell Rep.* **2016**, *14* (3), 479–492.

(128) Tuorto, F. et al. *EMBO J.* **2015**, *34* (18), 2350–2362.

(129) Shanmugam, R. et al. *Cell Discov.* **2015**, *1*, 15010.

(130) Schaefer, M. et al. *Genes Dev.* **2010**, *24* (15), 1590–1595.

(131) Rai, K. et al. *Genes Dev.* **2007**, *21* (3), 261–266.

(132) Lin, M.-J. et al. *J. Biol. Chem.* **2005**, *280* (2), 861–864.

(133) Kiani, J. et al. *PLoS Genet.* **2013**, *9* (5), e1003498.

(134) Païs de Barros, J. P. et al. *Nucleic Acids Res.* **1996**, *24* (8), 1489–1496.

(135) Kawarada, L. et al. *Nucleic Acids Res.* **2017**, *45* (12), 7401–7415.

(136) Chan, C. T. Y. et al. *Nat. Commun.* **2012**.

(137) Levitt, M. *Nature* **1969**, *224* (5221), 759–763.

(138) Väre, V. Y. P. et al. *Biomolecules* **2017**, *7* (1).

(139) Oliva, R. et al. *RNA* **2007**, *13* (9), 1427–1436.

(140) Gkatza, N. A. et al. *PLoS Biol.* **2019**, *17* (6), e3000297.

(141) Hussain, S. et al. **2013**, 255–261.

(142) Yuan, S. et al. *Mol. Cell. Biol.* **2014**, *34* (19), 3630–3641.

(143) Blanco, S. et al. *Nature* **2016**, *534* (7607), 335–340.

(144) Blanco, S. et al. *PLoS Genet.* **2011**, *7* (12), e1002403.

(145) Flores, J. V et al. *Stem cell reports* **2017**, *8* (1), 112–124.

(146) Hussain, S. et al. *Mol. Cell. Biol.* **2013**, *33* (8), 1561–1570.

(147) Wang, W. *Wiley Interdiscip. Rev. RNA* **2016**, *7* (6), 838–842.

(148)   Cai, X. et al. *Oncotarget* **2016**, *7* (15), 19099–19110.

(149)   Genenncher, B. et al. *Cell Rep.* **2018**, *22* (7), 1861–1874.

(150)   Long, T. et al. *J. Biol. Chem.* **2016**, *291* (46), 24293–24303.

(151)   Li, J. et al. *Nucleic Acids Res.* **2019**, *47* (4), 2041–2055.

(152)   Selmi, T. et al. *Nucleic Acids Res.* **2021**, *49* (2), 1006–1022.

(153)   Forbes, S. A. et al. *Nucleic Acids Res.* **2015**, *43*, D805-11.

(154)   Li, L. et al. *J. Hainan Med. Univ.* **2017**.

(155)   Towns, W. L. et al. *DNA Cell Biol.* **2012**, *31* (4), 434–454.

(156)   Tzelepi, V. et al. *Pathology* **2020**, *52* (2), 218–227.

(157)   Elhardt, W. et al. *Biochimie* **2015**, *112*, 66–72.

(158)   Chen, H. et al. *Nat. Commun.* **2020**, *11* (1), 2834.

(159)   Chen, Q. et al. *Science* **2016**, *351* (6271), 397–400.

(160)   Zhang, Y. et al. *Nat. Cell Biol.* **2018**, *20* (5), 535–540.

(161)   Wu, J. et al. *Int. J. Genomics* **2021**, *2021*, 3803724.

(162)   Li, X. et al. *J. Cell. Mol. Med.* **2021**, *25* (3), 1383–1393.

(163)   Zhai, C.-T. et al. *Kaohsiung J. Med. Sci.* **2021**, *37* (11), 991–999.

(164)   Huang, Z. et al. *Front. cell Dev. Biol.* **2021**, *9*, 657547.

(165)   Hu, Y. et al. *Cell Death Dis.* **2021**, *12* (9), 842.

(166)   Xiang, S. et al. *Front. Mol. Biosci.* **2020**, *7*, 599340.

(167)   Manning, M. et al. *RNA Biol.* **2020**, *17* (4), 474–486.

(168)   Lu, L. et al. *Oral Oncol.* **2020**, *101*, 104554.

(169)   Bhawe, K. et al. *Mol. Neurobiol.* **2020**, *57* (9), 3827–3845.

(170)   Zhu, Y. et al. *Cancer Manag. Res.* **2019**, *11*, 1383–1390.

(171)   Lu, L. et al. *Cancer Invest.* **2018**, *36* (4), 246–253.

(172)   Yi, J. et al. *Oncotarget* **2017**, *8* (13), 20751–20765.

(173)   Yang, J.-C. et al. *Future Oncol.* **2017**, *13* (22), 1981–1990.

(174)   Okamoto, M. et al. *DNA Cell Biol.* **2012**, *31* (5), 660–671.

(175)   Mei, L. et al. *Cell Death Dis.* **2020**, *11* (4), 270.

(176)   Xing, J. et al. *Mol. Cell. Biol.* **2015**, *35* (23), 4043–4052.

(177)   Okamoto, M. et al. *PLoS Genet.* **2014**, *10* (9), e1004639.

(178) Kato, K. et al. *Am. J. Med. Genet. A* **2021**, *185* (1), 282–285.

(179) Sun, S. et al. *Mol. Genet. genomic Med.* **2020**, *8* (12), e1518.

(180) Komara, M. et al. *J. Mol. Neurosci.* **2015**, *57* (3), 393–399.

(181) Abbasi-Moheb, L. et al. *Am. J. Hum. Genet.* **2012**, *90* (5), 847–855.

(182) Khan, M. A. et al. *Am. J. Hum. Genet.* **2012**, *90* (5), 856–863.

(183) Doan, R. N. et al. *Nat. Genet.* **2019**, *51* (7), 1092–1098.

(184) Blaze, J. et al. *Nat. Commun.* **2021**, *12* (1), 4913.

(185) Innes, A. M. et al. *Am. J. Med. Genet. C. Semin. Med. Genet.* **2018**, *178* (4), 387–397.

(186) Martinez, F. J. et al. *J. Med. Genet.* **2012**, *49* (6), 380–385.

(187) Li, C. et al. *Nat. Cell Biol.* **2017**, *19* (2), 106–119.

(188) Fang, X. et al. *J. Clin. Lab. Anal.* **2022**, *36* (4), e24303.

(189) Reid, R. et al. *Nucleic Acids Res.* **1999**, *27* (15), 3138–3145.

(190) Chellamuthu, A. et al. *Cells* **2020**, *9* (8).

(191) Iyer, L. M. et al. *Prog. Mol. Biol. Transl. Sci.* **2011**, *101*, 25–104.

(192) Sunita, S. et al. *J. Mol. Biol.* **2008**, *383* (3), 652–666.

(193) Jurkowski, T. P. et al. *RNA* **2008**, *14* (8), 1663–1670.

(194) Cheng, X. et al. *Nucleic Acids Res.* **2001**, *29* (18), 3784–3795.

(195) Cheng, X. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 293–318.

(196) Jeltsch, A. *Chembiochem* **2002**, *3* (4), 274–293.

(197) Dong, A. et al. *Nucleic Acids Res.* **2001**, *29* (2), 439–448.

(198) Cheng, X. et al. *Structure* **2008**, *16* (3), 341–350.

(199) Lyko, F. *Nat. Rev. Genet.* **2018**, *19* (2), 81–92.

(200) Jumper, J. et al. *Nature* **2021**, *596* (7873), 583–589.

(201) Varadi, M. et al. *Nucleic Acids Res.* **2022**, *50* (D1), D439–D444.

(202) Gros, C. et al. *Biochimie* **2012**, *94* (11), 2280–2296.

(203) Schulz, E. C. et al. *PLoS One* **2012**, *7* (6), e38728.

(204) *Nucleic Acids Res.* **2021**, *49* (D1), D480–D489.

(205) Santi, D. V et al. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81* (22), 6993–6997.

(206) Wu, J. C. et al. *J. Biol. Chem.* **1987**, *262* (10), 4778–4786.

(207) Chen, L. et al. *J. Am. Chem. Soc.* **1993**, *115* (12), 5318–5319.

(208) Cheng, C. C. *J. Pharm. Sci.* **1972**, *61* (4), 645–649.

(209) Both, G. W. et al. *Cell* **1975**, *6* (2), 185–195.

(210) Richon, V. M. et al. *Chem. Biol. Drug Des.* **2011**, *78* (2), 199–210.

(211) Borchardt, R. T. et al. *Biochem. Biophys. Res. Commun.* **1979**, *89* (3), 919–924.

(212) Jamaluddin, M. et al. *Biochemistry* **1975**, *14* (4), 694–698.

(213) Chen, H. et al. *PLoS One* **2013**, *8* (10), e76900.

(214) Schwickert, M. et al. *J. Med. Chem.* **2022**, *65*, 9750−9788.

(215) Hamil, R. L. et al. *J. Antibiot. (Tokyo).* **1973**, *26* (8), 463–465.

(216) Oerum, S. et al. *RNA Biol.* **2019**, *16* (6), 798–808.

(217) Bedi, R. K. et al. *ChemMedChem* **2020**, *15* (9), 744–748.

(218) Moroz-Omori, E. V et al. *ChemMedChem* **2021**, *16* (19), 3035–3043.

(219) Buker, S. M. et al. *SLAS Discov. Adv. life Sci. R D* **2020**, *25* (4), 361–371.

(220) Hausmann, S. et al. *J. Biol. Chem.* **2005**, *280* (21), 20404–20412.

(221) Zheng, S. et al. *J. Biol. Chem.* **2006**, *281* (47), 35904–35913.

(222) Lahoud, G. et al. *RNA* **2011**, *17* (7), 1236–1246.

(223) Lim, S. P. et al. *J. Biol. Chem.* **2011**, *286* (8), 6233–6240.

(224) Jain, R. et al. *Sci. Rep.* **2017**, *7* (1), 1632.

(225) Ahmed-Belkacem, R. et al. *Eur. J. Med. Chem.* **2020**, *201*, 112557.

(226) Otava, T. et al. *ACS Infect. Dis.* **2021**, *7* (8), 2214–2220.

(227) Devkota, K. et al. *SLAS Discov. Adv. life Sci. R D* **2021**, *26* (9), 1200–1211.

(228) Bobiļeva, O. et al. *ACS Med. Chem. Lett.* **2021**, *12* (7), 1102–1107.

(229) Halby, L. et al. *Philos. Trans. R. Soc. London. Ser. B, Biol. Sci.* **2018**, *373* (1748).

(230) Lerner, C. et al. *Org. Biomol. Chem.* **2003**, *1* (1), 42–49.

(231) Paulini, R. et al. *Chembiochem* **2004**, *5* (9), 1270–1274.

(232) Ellermann, M. et al. *Chemistry* **2011**, *17* (23), 6369–6381.

(233) Anglin, J. L. et al. *J. Med. Chem.* **2012**, *55* (18), 8066–8074.

(234) Stein, E. M. et al. *Blood* **2018**, *131* (24), 2661–2669.

(235) Saavedra, O. M. et al. *Bioorg. Med. Chem. Lett.* **2009**, *19* (10), 2747–2751.

(236) Isakovic, L. et al. *Bioorg. Med. Chem. Lett.* **2009**, *19* (10), 2742–2746.

(237) Tran, H. T. T. et al. *J. Korean Med. Sci.* **2011**, *26* (2), 207–213.

(238) Mund, C. et al. *Cancer Res.* **2005**, *65* (16), 7086–7090.

(239) Yang, A. S. et al. *Cancer Res.* **2006**, *66* (10), 5495–5503.

(240) Schaefer, M. et al. *Cancer Res.* **2009**, *69* (20), 8127–8132.

(241) Santi, D. V. et al. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81* (22 I), 6993–6997.

(242) Khoddami, V. et al. *Nat. Biotechnol.* **2013**, *31* (5), 458–464.

(243) Zhou, L. et al. *J. Mol. Biol.* **2002**, *321* (4), 591–599.

(244) Cho, H. D. et al. *RNA* **2003**, *9* (8), 970–981.

(245) Jerabek-Willemsen, M. et al. *J. Mol. Struct.* **2014**, *1077*, 101–113.

(246) Duhr, S. et al. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (52), 19678–19682.

(247) Seidel, S. A. I. et al. *Methods* **2013**, *59* (3), 301–315.

(248) Bartoschik, T. et al. *Methods Mol. Biol.* **2020**, *2168*, 51–62.

(249) Wienken, C. J. et al. *Nat. Commun.* **2010**, *1*, 100.

(250) Royer, C. A. *Chem. Rev.* **2006**, *106* (5), 1769–1784.

(251) Liu, Y. et al. *Analyst* **2015**, *140* (8), 2762–2770.

(252) Freyer, M. W. et al. *Methods Cell Biol.* **2008**, *84*, 79–113.

(253) Pierce, M. M. et al. *Methods* **1999**, *19* (2), 213–221.

(254) Velázquez-Campoy, A. et al. *Curr. Protoc. cell Biol.* **2004**, *Chapter 17*, Unit 17.8.

(255) Martinez Molina, D. et al. *Science* **2013**, *341* (6141), 84–87.

(256) Lo, M.-C. et al. *Anal. Biochem.* **2004**, *332* (1), 153–159.

(257) Gao, K. et al. *Biophys. Rev.* **2020**, *12* (1), 85–104.

(258) Pantoliano, M. W. et al. *J. Biomol. Screen.* **2001**, *6* (6), 429–440.

(259) Niesen, F. H. et al. *Nat. Protoc.* **2007**, *2* (9), 2212–2221.

(260) Poklar, N. et al. *Biochemistry* **1997**, *36* (47), 14345–14352.

(261) Tovy, A. et al. *J. Vis. Exp.* **2010**, No. 44.

(262) PerkinElmer. *Liquid Scintillation Cocktails*. https://www.perkinelmer.com/de/lab-products-and-services/application-support-knowledgebase/radiometric/liquid-scintillation-cocktails.html#Liquidscintillationcocktails-Roleofscintillator.

(263) Fang, H. et al. *Chem. Sci.* **2021**, *12* (24), 8288–8310.

(264) Lee, J. et al. *ACS chemical biology*. United States April 2021, pp 579–585.

(265) Sanman, L. E. et al. *Annu. Rev. Biochem.* **2014**, *83*, 249–273.

(266) Niphakis, M. J. et al. *Annu. Rev. Biochem.* **2014**, *83*, 341–377.

(267) Arrowsmith, C. H. et al. *Nat. Chem. Biol.* **2015**, *11* (8), 536–541.

(268) Weiss, W. A. et al. *Nat. Chem. Biol.* **2007**, *3* (12), 739–744.

(269) Wijdeven, R. H. et al. *Trends Cell Biol.* **2014**, *24* (12), 751–760.

(270) Boström, J. et al. *Nat. Rev. Drug Discov.* **2018**, *17* (10), 709–727.

(271) Martínez-Montero, S. et al. *J. Org. Chem.* **2015**, *80* (6), 3083–3091.

(272) Thompson, M. J. et al. *J. Org. Chem.* **1999**, *64* (20), 7467–7473.

(273) Robins, M. J. et al. *Can. J. Chem.* **1991**, *69* (9), 1468–1474.

(274) Hartenfeller, M. et al. *J. Chem. Inf. Model.* **2011**, *51* (12), 3093–3098.

(275) Brown, D. G. et al. *J. Med. Chem.* **2016**, *59* (10), 4443–4458.

(276) Barton, D. H. R. et al. *J. Chem. Soc. Perkin Trans. 1* **1991**, No. 5, 981–985.

(277) Hollmann, J. et al. *Liebigs Ann. der Chemie* **1984**, *1984* (1), 98–107.

(278) Liu, Q. et al. *Bioorganic Med. Chem.* **2017**, *25* (17), 4579–4594.

(279) Zhang, G. et al. *Org. Biomol. Chem.* **2015**, *13* (14), 4149–4154.

(280) Dowden, J. et al. *Org. Biomol. Chem.* **2011**, *9* (22), 7814–7821.

(281) Aronov, A. M. et al. *Bioorg. Med. Chem. Lett.* **1998**, *8* (24), 3505–3510.

(282) Van Calenbergh, S. et al. *J. Med. Chem.* **1995**, *38* (19), 3838–3849.

(283) Ahmed-Belkacem, R. et al. *J. Med. Chem.* **2022**, *65* (8), 6231–6249.

(284) Topliss, J. G. *J. Med. Chem.* **1972**, *15* (10), 1006–1011.

(285) Wang, S. et al. *Acta Pharm. Sin. B* **2019**, *9* (5), 880–901.

(286) Müller, S. et al. *Nucleic Acids Res.* **2013**, *41* (18), 8615–8627.

(287) Luan, Y. et al. *Org. Biomol. Chem.* **2016**, *14* (2), 631–638.

(288) Sarfati, S. R. et al. *Tetrahedron* **1988**, *44* (20), 6367–6372.

(289) Beaumont, K. et al. *Curr. Drug Metab.* **2003**, *4* (6), 461–485.

(290) McClellan, K. J. et al. *Drugs* **1998**, *56* (5), 847–869.

(291) Shimma, N. et al. *Bioorg. Med. Chem.* **2000**, *8* (7), 1697–1706.

(292) Ishikawa, T. et al. *Biochem. Pharmacol.* **1998**, *55* (7), 1091–1097.

(293) Ponzano, S. et al. *J. Med. Chem.* **2013**, *56* (17), 6917–6934.

(294) Nepali, K. et al. *J. Med. Chem.* **2019**, *62* (6), 2851–2893.

(295) Berman, H. M. et al. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.

300

(296) Rarey, M. et al. *J. Mol. Biol.* **1996**, *261* (3), 470–489.

(297) BioSolveIT GmbH: St. Augustin 2017.

(298) Chemical Computing Group ULC: Montreal 2019. https://www.chemcomp.com/Research-Citing_MOE.htm.

(299) Schrödinger, LLC.

(300) Colley, D. G. et al. *Lancet* **2014**, *383* (9936), 2253–2264.

(301) Gryseels, B. et al. *Lancet (London, England)* **2006**, *368* (9541), 1106–1118.

(302) Waine, G. J. et al. *Bioessays* **1997**, *19* (5), 435–443.

(303) Vos, T. et al. *Lancet (London, England)* **2012**, *380* (9859), 2163–2196.

(304) Chitsulo, L. et al. *Nat. Rev. Microbiol.* **2004**, *2* (1), 12–13.

(305) *World Health Organisation (2022). Schistosomiasis [Fact Sheet].* https://www.who.int/news-room/fact-sheets/detail/schistosomiasis.

(306) Ross, A. G. et al. *Lancet. Infect. Dis.* **2007**, *7* (3), 218–224.

(307) Bottieau, E. et al. *J. Infect.* **2006**, *52* (5), 339–345.

(308) Pearce, E. J. et al. *Nat. Rev. Immunol.* **2002**, *2* (7), 499–511.

(309) Cheever, A. W. et al. *Immunology today*. England September 2000, pp 465–466.

(310) Fairfax, K. et al. *Semin. Immunopathol.* **2012**, *34* (6), 863–871.

(311) Boros, D. L. et al. *J. Exp. Med.* **1970**, *132* (3), 488–507.

(312) Warren, K. S. et al. *Am. J. Trop. Med. Hyg.* **1974**, *23* (5), 902–909.

(313) Chabasse, D. et al. *Bull. Soc. Pathol. Exot. Filiales* **1985**, *78* (5), 643–647.

(314) Gryseels, B. *Infect. Dis. Clin. North Am.* **2012**, *26* (2), 383–397.

(315) Ross, A. G. P. et al. *N. Engl. J. Med.* **2002**, *346* (16), 1212–1220.

(316) Spangenberg, T. *ACS Infect. Dis.* **2021**, *7* (5), 939–942.

(317) Fenwick, A. *Future medicinal chemistry*. England 2015, pp 677–680.

(318) Andrews, P. et al. *Med. Res. Rev.* **1983**, *3* (2), 147–200.

(319) Meister, I. et al. *Antimicrob. Agents Chemother.* **2014**, *58* (9), 5466–5472.

(320) Pax, R. et al. *Naunyn. Schmiedebergs. Arch. Pharmacol.* **1978**, *304* (3), 309–315.

(321) Staudt, U. et al. *Parasitol. Res.* **1992**, *78* (5), 392–397.

(322) Park, S.-K. et al. *J. Biol. Chem.* **2019**, *294* (49), 18873–18880.

(323) Park, S.-K. et al. *Sci. Transl. Med.* **2021**, *13* (625), eabj5832.

(324)   Dayan, A. D. *Acta Trop.* **2003**, *86* (2–3), 141–159.

(325)   Olliaro, P. L. et al. *PLoS Negl. Trop. Dis.* **2011**, *5* (6), e1165.

(326)   *WHO GUIDELINE on Control and Elimination of Human Schistosomiasis*; Lo, N. C., Ed.; World Health Organization, 2022.

(327)   Groll, E. *Adv. Pharmacol. Chemother.* **1984**, *20*, 219–238.

(328)   Brindley, P. J. et al. *J. Immunol.* **1987**, *139* (1), 215–220.

(329)   Doenhoff, M. J. et al. *Trans. R. Soc. Trop. Med. Hyg.* **1987**, *81* (6), 947–951.

(330)   Foster, R. *Trans. R. Soc. Trop. Med. Hyg.* **1987**, *81* (1), 55–59.

(331)   Feldmeier, H. et al. *Acta Trop.* **1987**, *44* (3), 357–368.

(332)   Katz, N. et al. *Am. J. Trop. Med. Hyg.* **1991**, *44* (5), 509–512.

(333)   Olveda, D. U. et al. *Curr. Opin. Infect. Dis.* **2016**, *29* (6), 595–608.

(334)   Stelma, F. F. et al. *Am. J. Trop. Med. Hyg.* **1995**, *53* (2), 167–170.

(335)   Guisse, F. et al. *Am. J. Trop. Med. Hyg.* **1997**, *56* (5), 511–514.

(336)   Wang, W. et al. *Parasitol. Res.* **2012**, *111* (5), 1871–1877.

(337)   *Ending the Neglect to Attain the Sustainable Development Goals: A Road Map for Neglected Tropical Diseases 2021–2030*; Malecela Mwelecele Ntuli, Ed.; World Health Organization, 2022.

(338)   Caffrey, C. R. et al. *Parasitol. Res.* **1997**, *83* (6), 632–635.

(339)   Sajid, M. et al. *Mol. Biochem. Parasitol.* **2003**, *131* (1), 65–75.

(340)   Polgár, L. et al. *J. Biol. Chem.* **1987**, *262* (30), 14448–14453.

(341)   Chappell, C. L. et al. *Exp. Parasitol.* **1986**, *61* (2), 160–167.

(342)   Correnti, J. M. et al. *Mol. Biochem. Parasitol.* **2005**, *143* (2), 209–215.

(343)   Krautz-Peterson, G. et al. *Mol. Biochem. Parasitol.* **2008**, *159* (1), 54–58.

(344)   Jílková, A. et al. *J. Biol. Chem.* **2011**, *286* (41), 35770–35781.

(345)   Musil, D. et al. *EMBO J.* **1991**, *10* (9), 2321–2330.

(346)   Illy, C. et al. *J. Biol. Chem.* **1997**, *272* (2), 1197–1202.

(347)   Novinec, M. et al. *Biomol. Concepts* **2013**, *4* (3), 287–308.

(348)   Otto, H.-H. et al. *Chem. Rev.* **1997**, *97* (1), 133–172.

(349)   Abdulla, M.-H. et al. *PLoS Med.* **2007**, *4* (1), e14.

(350)   Engel, J. C. et al. *J. Exp. Med.* **1998**, *188* (4), 725–734.

(351)   Kerr, I. D. et al. *J. Biol. Chem.* **2009**, *284* (38), 25697–25703.

(352) Yang, P.-Y. et al. *Chem. Commun. (Camb).* **2012**, *48* (6), 835–837.

(353) Jacobsen, W. et al. *Drug Metab. Dispos.* **2000**, *28* (11), 1343–1351.

(354) Palmer, J. T. et al. *J. Med. Chem.* **1995**, *38* (17), 3193–3196.

(355) Schechter, I. et al. *Biochem. Biophys. Res. Commun.* **1967**, *27* (2), 157–162.

(356) Murata, M. et al. *FEBS Lett.* **1991**, *280* (2), 307–310.

(357) Towatari, T. et al. *FEBS Lett.* **1991**, *280* (2), 311–315.

(358) Jílková, A. et al. *ACS Infect. Dis.* **2021**, *7* (1), 189–201.

(359) Jung, S. et al. *J. Med. Chem.* **2021**, *64* (16), 12322–12358.

(360) Finley, D. *Annu. Rev. Biochem.* **2009**, *78*, 477–513.

(361) Zhao, J. et al. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (52), 15790–15797.

(362) Cromm, P. M. et al. *ACS Cent. Sci.* **2017**, *3* (8), 830–838.

(363) Collins, G. A. et al. *Cell* **2017**, *169* (5), 792–806.

(364) Tyedmers, J. et al. *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (11), 777–788.

(365) Fredrickson, E. K. et al. *Semin. Cell Dev. Biol.* **2012**, *23* (5), 530–537.

(366) Haas, A. L. et al. *J. Biol. Chem.* **1982**, *257* (5), 2543–2548.

(367) Yang, Y. et al. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2003**, *17* (8), 790–799.

(368) Risseeuw, E. P. et al. *Plant J.* **2003**, *34* (6), 753–767.

(369) Glickman, M. H. et al. *Physiol. Rev.* **2002**, *82* (2), 373–428.

(370) Thrower, J. S. et al. *EMBO J.* **2000**, *19* (1), 94–102.

(371) Ciechanover, A. et al. *Biochim. Biophys. Acta* **2014**, *1843* (1), 86–96.

(372) Sahu, I. et al. *Nat. Commun.* **2021**, *12* (1), 6173.

(373) Groll, M. et al. *Nature* **1997**, *386* (6624), 463–471.

(374) Huang, X. et al. *Nat. Struct. Mol. Biol.* **2016**, *23* (9), 778–785.

(375) Budenholzer, L. et al. *J. Mol. Biol.* **2017**, *429* (22), 3500–3524.

(376) Kisselev, A. F. et al. *J. Biol. Chem.* **1999**, *274* (6), 3363–3371.

(377) Orlowski, M. *Biochemistry* **1990**, *29* (45), 10289–10297.

(378) Groll, M. et al. *Int. J. Biochem. Cell Biol.* **2003**, *35* (5), 606–616.

(379) Huber, E. M. et al. *Nat. Commun.* **2016**, *7*, 10900.

(380) Ding, F. et al. *Front. Biosci. (Landmark Ed.* **2014**, *19* (6), 886–895.

(381) Goldberg, A. L. *J. Cell Biol.* **2012**, *199* (4), 583–588.

(382)   Shen, M. et al. *Expert Opin. Ther. Targets* **2013**, *17* (9), 1091–1108.

(383)   Verdoes, M. et al. *European J. Org. Chem.* **2009**, No. 20, 3301–3313.

(384)   Zeslawska, E. et al. *J. Mol. Biol.* **2003**, *328* (1), 109–118.

(385)   Ward, Y. D. et al. *J. Med. Chem.* **2002**, *45* (25), 5471–5482.

(386)   Hattori, S.-I. et al. *Nat. Commun.* **2021**, *12* (1), 668.

(387)   Li, C. Y. et al. *Chembiochem* **2019**, *20* (1), 46–50.

(388)   Previti, S. et al. *J. Med. Chem.* **2017**, *60* (16), 6911–6923.

(389)   Schrader, J. et al. *Science* **2016**, *353* (6299), 594–598.

9 Appendix

CD containing NMR spectra, chromatograms, and Supporting Information files.

Curriculum Vitae

██████   ████████████████████████████████
                    ██████████████████████████ █
████████   ████████████████████████████
██████   ████████████████████████
██

█████████ █████████████

███████████████   ███████ █████ █ ████ ████ ███ █████████ █████ █
                       ███
                       ████████████ ████

                       ██
                       ██
                       ██
                       ██████████████████████████████

████████████   ██
                       ██
                       █████████████ ████

███████████████   ███████ ████████ █ █ ██ ██ █████████ █████ █ █
                       ██
                       ████████████ ████

                       ██

█████████████   ██
                       █████████████████████████████████

                       ████████████████████████████████████

█████████████   ██
                       ████████████████████████████

                       ████████████████████████████████████

███████   ██████ ████████ ██████ █████████ ███████ ████ ██
                    ████████████████████████
██

████

████████   ██
                     ██

██████████   ██████████████████████████████████████
██

██████████████

██████████████   █████████████████████████████████████