Unraveling the Secrets of Pathogenic RNA: Biophysical Approaches to Study RNA–Small Molecule Interactions

Doctoral thesis

submitted in the fulfillment of the regulations for attaining the degree of
“Doktor der Naturwissenschaften (Dr. rer. nat.)“
on the subject of Pharmaceutical and Medicinal Chemistry

at the faculty of Chemistry, Pharmaceutical Sciences, Geography and Geosciences
at the Johannes Gutenberg University of Mainz

Elisabeth Kallert
Born in Kaufbeuren

Mainz, 2024

Submitted at the Faculty of Chemistry, Pharmaceutical Sciences, Geography and Geosciences:

Dean: [redacted]

Name of 1st reviewer: [redacted]
Name of 2nd reviewer: [redacted]

Date of doctoral examination: November 12th, 2024

D77 (Dissertation Universität Mainz)

Declaration of Authorship

I, Elisabeth Kallert, declare that this thesis titled, “Unraveling the Secrets of pathogenic RNA: Biophysical Approaches to Study RNA–Small Molecule Interactions” and the work presented in it are my own. I confirm that:

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

Signed:

Date:

"*The secret of success is making your vocation your vacation*."

Mark Twain

Abstract

The targeting of RNA molecules represents a frontier in drug discovery to be poised to revolutionize the treatment of bacterial and viral infections. While traditional drug development has predominantly focused on protein targets, the distinctive structures and functions of bacterial and viral RNA offer hitherto unexploited prospects for therapeutic intervention. However, the transition from concept to clinical application is beset with challenges. RNA's inherent instability, diverse conformational dynamics and intricate cellular interactions present significant challenges in the identification and optimization of effective small molecules binders. This dissertation explores the high stakes of RNA-targeted drug discovery, highlighting the transformative potential of such approaches in combating infectious diseases.

A critical element of this undertaking is the utilization of sophisticated biophysical techniques to elucidate the interactions between small molecules and RNA targets. Microscale thermophoresis (MST), isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), and Förster resonance energy transfer (FRET) assays emerge as indispensable techniques for this endeavor. MST impresses with a rapid and sample-saving affinity determination, ITC provides quantitative insights into binding thermodynamics, SPR offers real-time interactions kinetics, and FRET enables the visualization of conformational changes within RNA complexes.

In this dissertation, significant strides in this field are presented. The $preQ_1$ riboswitch has been a focal point, where an MST assay was employed to validate protein-based docking programs for RNA docking, thereby identifying small molecule binders for this RNA target. Moreover, an MST assay utilizing non-covalent dyes for studying RNA–ligand interactions was devised to augment the sensitivity and adaptability of this methodology. The investigation into HCV-IRES binding events using FRET yielded significant insights into the conformational changes that occur upon small molecule binding to this target. Furthermore, the relationship between charge and binding affinity and kinetics was investigated for the $preQ_1$ and SAM-VI riboswitch through a combination of MST, (kin)ITC, and SPR, revealing intricate interaction mechanisms. A structure-activity relationship (SAR) study on SAM-I and SAM-VI riboswitches with SAM analogues provided further insights into the subtle interplay between ligand structure and RNA binding efficacy.

Zusammenfassung

Das gezielte Adressieren von RNA-Molekülen birgt ein großes Potenzial in der Arzneimittelforschung, bereit die Behandlung bakterieller und viraler Infektionen weiter voranzutreiben. Während sich die traditionelle Arzneimittelentwicklung vorwiegend auf Protein-Zielstrukturen konzentriert hat, eröffnen die einzigartigen Strukturen und Funktionen bakterieller und viraler RNA bislang ungenutzte Möglichkeiten für Therapieansätze. Der Übergang von Konzept zu klinischer Anwendung ist allerdings mit erheblichen Herausforderungen verbunden. Die intrinsische Instabilität von RNA-Molekülen, ihre vielfältigen dynamischen Konformationen sowie komplexe Interaktionen auf zellulärer Ebene erschweren die Identifikation und Optimierung von RNA-bindenden *small Molecules*. Diese Dissertation widmet sich den großen Herausforderungen der auf RNA als Zielstruktur ausgelegten Wirkstoffforschung und hebt das transformative Potenzial solcher Ansätze im Kontext der Bekämpfung von Infektionskrankheiten hervor.

Ein wesentlicher Aspekt dieses Ansatzes ist die Nutzung biophysikalischer Techniken, um die Interaktionen zwischen *small Molecules* und RNA-Zielstrukturen zu erforschen. *Microscale Thermophoresis* (MST), Isothermale Titrationskalorimetrie (*isothermal titration calorimetry*, ITC), Oberflächenplasmonenresonanz (*surface plasmon resonance*, SPR) und Förster-Resonanz-Energie-Transfer (FRET) erweisen sich dabei als unabdingbare Methoden. MST ermöglicht eine schnelle und materialsparende Methode zur Affinitätsbestimmung, ITC liefert quantitative Einblicke in die Bindungsthermodynamik, SPR bietet Echtzeit-Kinetiken von Interaktionen und FRET ermöglicht die Verfolgung von Konformationsänderungen von RNA-Komplexen. In dieser Dissertation werden bedeutende Fortschritte auf diesem Gebiet präsentiert.

Ein Schwerpunkt war der $preQ_1$ *Riboswitch*, bei dem ein MST-*Assay* zur Validierung proteinbasierter Docking-Programme eingesetzt wurde, wodurch neue *small Molecules*, die an diese RNA binden, identifiziert wurden. Darüber hinaus wurde ein auf nicht-kovalenten Farbstoffen basierender MST-*Assay* entwickelt, um die Sensibilität und Anwendbarkeit dieser Methodik zu erhöhen. Die Untersuchung von Hepatitis-C-Virus *internal ribosome entry site* (HCV IRES)-Bindungsereignissen mit Hilfe von FRET hat bedeutende Einblicke in die Konformationsänderungen der RNA, die durch die Bindung von *small Molecules* auftreten, geliefert. Des Weiteren wurde die Beziehung zwischen Ladung und Bindungsaffinität und -kinetik für den $preQ_1$ *Riboswitch* und den SAM-VI *Riboswitch* untersucht. Dafür wurden die Methoden MST, (kin)ITC und SPR angewandt und thermodynamische und kinetische Bindungsprofile charakterisiert. Zusätzlich wurde die Struktur-Wirkungsbeziehung für die SAM-I und den SAM-VI *Riboswitches* genauer untersucht. Als potenzielle Liganden wurden SAM-Analoga verwendet, um tiefere Einblicke in das subtile Zusammenspiel zwischen Ligandenstruktur und RNA-Bindungseffizienz zu gewinnen.

Acknowledgments

I would like to thank numerous people who have accompanied and supported me over the last few years during my doctorate and who have contributed to the success of this doctoral thesis.

First of all, I want to heartfully thank [redacted] for the opportunity to write this thesis in her working group. I also would like to thank [redacted] for the interesting and challenging projects. Your scientific expertise, exchange and vigorous support within all stages of each project was incredibly valuable to find the right conclusions and move one step further. Additionally, I am grateful that you read this thesis and provided suggestions for improvement of the first draft versions.

Special thanks also go to [redacted] for carefully reviewing this thesis and for the permission to use several devices and laboratories of his working group. Furthermore, I would like to thank [redacted] and [redacted] for the many SPR measurements and their support during the evaluation.

I would also like to thank my colleagues [redacted], [redacted], [redacted], [redacted], [redacted] and [redacted] for the great collaboration within several projects and supervision of the students' practical course. Furthermore, I would like to thank [redacted] for the kind introduction in working with RNA.

Furthermore, I want to thank [redacted], [redacted] and [redacted] for always having an open ear for all kinds of concerns occurring during the last three years.

During my time in this working group, I was allowed to supervise several students: [redacted], [redacted] and [redacted] provided great contributions to several projects within their Master/Bachelor Thesis. [redacted], [redacted], [redacted] and [redacted] worked on various small projects during their internship. Your scientific curiosity and diligence were an honor to work with and I am very proud of your achievements.

Last but not least, I want to thank my family and friends, without your unconditionally support, this journey would not have been possible.

[redacted] and [redacted] accompanied me during all stages of my doctorate and listened to all of my bigger and smaller problems. Whether drinking coffee, cooking, theater

visitations, pub quiz evenings or long walks along the rhine, spending time with you always balanced me out.

I also want to thank [redacted] from the bottom of my heart for always being there for me in every situation, motivated and grounded me with an incredible calm and patience.

Furthermore, I would like to say a special thank you to [redacted] and [redacted] for always “adopting” me, game nights, excursions, and of course especially the great food from the master chef himself.

My deepest gratitude goes to my parents, who enabled me everything – always and unconditionally. Without your support, I would not be at this point today. I am blessed to have such great people around me.

Table of Contents

Declaration of Authorship ..... **i**

Abstract ..... **iii**

Zusammenfassung ..... **iv**

Acknowledgments ..... **v**

List of Abbreviations ..... **ix**

1. Introduction **1**

1.1. Targeting RNA with Small Molecules – General Process and Challenges *1*

1.1.1. The Relevance of Disease-Related RNAs 2

1.1.2. Target Identification and Validation 6

1.1.3. Identification of Small Molecules – Common Strategies 7

1.1.4. Determination of RNA Structures 12

1.1.5. Structure-Based Small Molecule Design and Optimization 15

1.2. Antibiotic Drug Discovery *18*

1.2.1. Antibiotics and Multi-Drug Resistance 18

1.2.2. Perspectives for Novel Antibiotic Drugs 20

1.3. Riboswitches *21*

1.3.1. Gene Control Mechanisms of Riboswitches 21

1.3.2. Riboswitch Classification System 24

1.3.3. Riboswitches as Antibiotic Drug Targets 29

1.4. $PreQ_1$ Riboswitches *31*

1.4.1. Structural Features and Ligand Recognition 31

1.4.2. Regulation Mechanisms 33

1.4.3. $PreQ_1$ Riboswitches as Potential Drug Targets 37

1.4.4. $PreQ_1$ Riboswitch as Model System 38

1.5. SAM Riboswitch Family *39*

1.5.1. Biological Functions and Relevance 39

1.5.2. SAM-I Riboswitches – Structure, Ligand Recognition and Distribution 40

1.5.3. SAM-I Riboswitch as Potential Drug Target 42

1.5.4. SAM-VI Riboswitch – Structure, Ligand Recognition and Distribution 43

1.5.5. SAM-VI Riboswitches as Model System for SAM-III Riboswitches 45

1.6. Antiviral Drug Discovery *48*

1.6.1. Antiviral Drugs and Their Limitation 48

1.6.2. Viral RNAs as Drug Targets 51

1.7. Hepatitis C Virus *51*

1.7.1. Prevalence and Risk Factors 51
1.7.2. Organization of the HCV Genome 52
1.7.3. HCV IRES 53
1.7.4. HCV IRES as Potential Drug Target 55
1.7.5. HCV IRES as a Model System 55

1.8. Unveiling RNA–Small Molecule Interactions Using Biophysical Methods 56
1.8.1. Challenges in Assay Development 56
1.8.2. Microscale Thermophoresis 61
1.8.3. Isothermal Titration Calorimetry 64
1.8.4. Förster Resonance Energy Transfer (FRET) 69
1.8.5. Surface Plasmon Resonance (SPR) 71

2. List of Publications and Manuscripts 76
2.1. Publications and Manuscripts as Part of this Doctoral Thesis 76
2.2. Publication Beyond This Doctoral Thesis 79

3. Discovery of Novel Binders for the $PreQ_1$ Riboswitch by MST 80
3.1. Context, Project Summary, and Own Contribution 80
3.2. Publication 82

4. Non-Covalent MST as a New Method for RNA-Ligand Interaction Studies 132
4.1. Context, Project Summary, and Own Contribution 132
4.2. Publication 134

5. Biophysical Investigation of HCV IRES-Ligand Interactions 154
5.1. Context, Project Summary, and Own Contribution 154
5.2. Publication 156

6. Electrostatic Anchoring – Exploring the Impact of Charge to RNA–Small Molecule Binding Affinity, Kinetics, Thermodynamics and Selectivity 240
6.1. Context, Project Summary, and Own Contribution 240
6.2. Draft Manuscript 242

7. SAR Study on the SAM-I and SAM-VI Riboswitches – Preliminary Data 335
7.1. Introduction 335
7.2. Results and Discussion 341
7.3. Experimental Section 350

8. Conclusion and Outlook 354

9. Bibliography 365

List of Abbreviations

ALS	Amyotrophic lateral sclerosis
2-AP	2-aminopurine
5'-ADA	5'-amino-5'-desoxyadenosine
A	Adenine
ABSSSI	Acute bacterial skin and skin structure infections
ADE	Antibody-dependent enhancement
AdoCbl	Adenosyl cobalamin
AIDS	Acquired immunodeficiency syndrome
ALIS	Automated ligand identification system
AS-MS	Affinity selection-mass spectrometry
ASO	Antisense oligonucleotide
Asp	Aspartate
ATP	Adenosine triphosphate
Ba	*Bifidobacterium angulatum*
BACE1	β-site amyloid precursor protein cleaving enzyme 1
Bs	*Bacillus subtilis*
C	Cytosine
CABP	Community-acquired bacterial pneumonia
CADD	Computer-aided drug-design
c-di-GMP	Cyclic-di-guanosine monophosphate
Chem-CLIP	Chemical cross-linking and isolation by pull-down
COVID-19	Coronavirus disease 2019
CRISPR	Clustered regularly interspaced short palindromic repeats
Cryo-EM	Cryogenic electron microscopy
Cy3	Cyanine-3
Cy5	Cyanine-5
DAA	Direct acting antiviral
DABA	2,4-Diaminobutyric acid
DEL	DNA encoded library
DENV	Dengue virus
DM1	Myotonic dystrophy type 1
DMPK	Dystrophia myotonica protein kinase
DMS	Dimethyl sulfate
DNA	Desoxyribonucleic acid
DNMT2	Desoxyribonucleic acid methyltransferase 2
ESCMID	European Society of Clinical Microbiology and Infectious Diseases
ESKAPE	*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter pneumoniae, Pseudomonas aeruginosa, Enterobacter species*
ETC	Equilibration time curve
FBLD	Fragment-based ligand design
FDA	U.S. Food and Drug Administration
FID	Fluorescence indicator displacement
FMN	Flavin mononucleotide
FP	Fluorescence polarization
FQ	Fit quality

FRET	Förster resonance energy transfer
G	Guanine
GlmS	Glucosamine-6-phosphate
HBV	Hepatitis B virus
HCMV	Human cytomegalovirus
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HSV	Herpes simplex virus
HTS	High-throughput screening
IC_{50}	Half-maximal inhibitory concentration
IRES	Internal ribosome entry site
ITC	Isothermal titration calorimetry
K_A	Association constant
K_D	Dissociation constant
k_{on}	Association rate
k_{off}	Dissociation rate
kinITC	Kinetic isothermal titration calorimetry
LC-MS	Liquid chromatography-mass spectrometry
LE	Ligand efficiency
LLE	Ligand-lipophilicity efficiency
lncRNA	Long non-coding ribonucleic acid
logP	Logarithm of the octanol-water partition coefficient
MALAT1	Metastasis-associated lung adenocarcinoma transcript-1
MERS	Middle east respiratory syndrome
METTL	Methyltransferase-like
MoCo	Molybdenum cofactor
M^{Pro}	Main protease
mRNA	messenger ribonucleic acid
miRNA	micro ribonucleic acid
MRSA	Methicillin-resistant *Staphylococcus aureus*
MS	Mass spectrometry
MST	Microscale thermophoresis
ncRNA	non-coding ribonucleic acid
NMR	Nuclear magnetic resonance
NS	Non-structural
ORF	Open reading frame
oQ	Epoxyqueuosine
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pIC_{50}	negative logarithm of the IC_{50} value
$PreQ_0$	Pre-queuosine-0, 7-cyano-7-deazaguanine
$PreQ_1$	Pre-queuosine-1; 7-aminomethyl-7-deaza-guanine
Q	Queuosine
qPCR	Quantitative polymerase chain reaction
RBS	Ribosome-binding site
RiboTAC	Ribonuclease targeting chimera
RNA	Ribonucleic acid
RSV	Respiratory syncytial virus
RU	Response unit
SAB	*Staphylococcus aureus* bacteremia
SAH	*S*-Adenosylhomocysteine

SAM	*S*-Adenosylmethionine
SAR	Structure-activity relationship
SARS	Severe acute respiratory syndrome
SARS-CoV-2	Severe acute respiratory syndrome coronavirus type 2
SBD	Structure-based design
SDS	Shine Dalgarno sequence
SEC	Size exclusion chromatography
SFG	Sinefungin
SHAPE	Selective 2'-hydroxyl acylation analyzed by primer extension
siRNA	small interfering ribonucleic acid
SMA	Spinal muscular atrophy
smFRET	single molecule Förster resonance energy transfer
SMM	Small molecule microarray
SMN	Human survival of motor neuron
SPR	Surface plasmon resonance
STR	Short tandem repeats
TAR	Transactivating response region
TGT	transfer ribonucleic acid-guanine transglycosylase
THF	Tetrahydrofolate
TPP	Thiamine pyrophosphate
tPSA	topological polar surface area
tRNA	transfer ribonucleic acid
Tte	*Thermoanaerobacter tengcongensis*
U	Uracil
UTR	Untranslated region
VZV	Varicella-zoster virus
WHO	World Health Organization

1. Introduction

The objective of this thesis was the biophysical characterization of various RNA–small molecule interactions regarding fundamental research in structure-based drug design. This thesis is structured into 4 topics:

(A) The general process of targeting RNA, including different approaches and methodologies.
(B) Antibiotic drug discovery with focus on riboswitches as bacterial RNA targets.
(C) Anti-viral drug discovery with focus on the hepatitis C virus internal ribosome entry site (HCV IRES) as model system for viral RNA targeting.
(D) Biophysical assay methods to investigate and characterize those RNA-ligand interactions: Microscale thermophoresis (MST), isothermal titration calorimetry (ITC), Förster resonance energy transfer (FRET), and surface plasmon resonance (SPR).

1.1. Targeting RNA with Small Molecules – General Process and Challenges

Targeting RNA has emerged as a promising strategy for therapeutic intervention, offering precise modulation of RNA function, stability, and gene expression with potential applications across various human diseases, including bacterial or viral infections, cancer, and neurological disorders. The discovery of RNA's diverse roles beyond its canonical function as a messenger between DNA and protein has spurred significant interest in targeting RNA for therapeutic purposes.[1,2] The establishment of small molecules as RNA modulators would result in a significant increase in the number of druggable macromolecules, as many of the approximately 3,000 disease-related human proteins are currently considered undruggable (Figure 1). This is since they lack distinct binding sites for high affinity and specificity small molecule binding.[3–5] Consequently, by targeting mRNA, these undruggable protein gene products can be influenced by modulating the translation efficiency.[5] However, recent approaches also target non-coding (nc) RNA if those structures are found to be relevant for the survival of the pathogenic organisms or processes.[1,6]

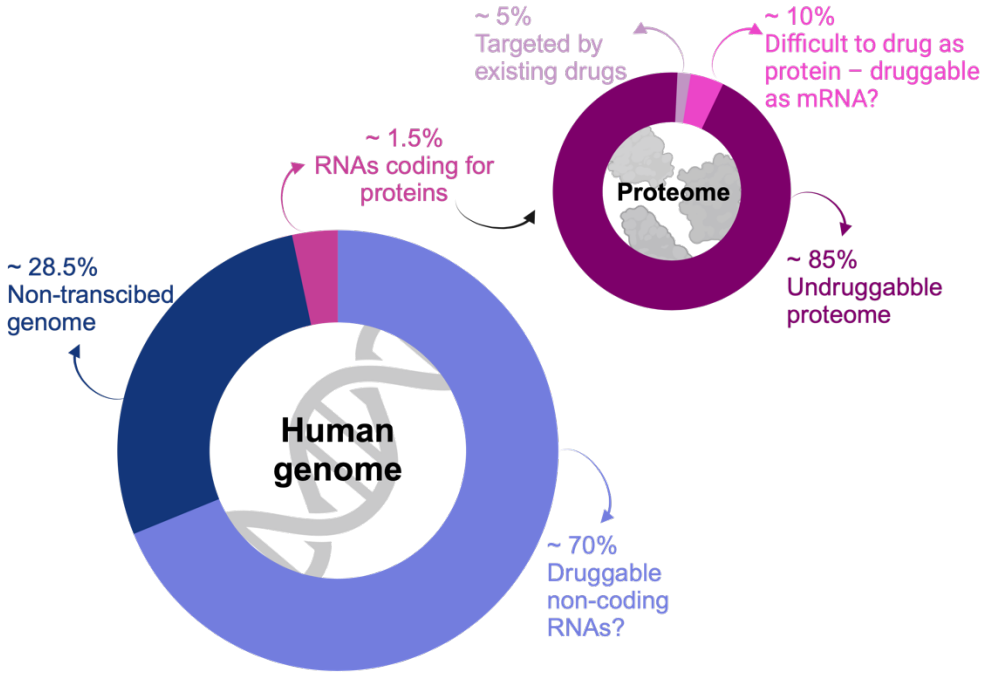


Figure 1: RNA-targeted parts of the human genome relevant for potential drug development. ~ 70% of the human genome is transcribed into ncRNAs, including microRNAs or long ncRNAs as relevant drug targets. Only ~ 1.5% of the human genome is transcribed into protein coding RNAs with only 0.2% of the genome refers to disease-related proteins (~ 15% of the proteome). Currently, ~ 5% of the proteome is targeted by existing drugs and ~ 10% represent relevant proteins that might be druggable as mRNA. ~ 85% of the proteome is undruggable. This figure was created with BioRender.com according to reference [5].

1.1.1. The Relevance of Disease-Related RNAs

The etiology of human diseases can be attributed to a multitude of factors, including the involvement of RNA molecules. External sources may include bacteria or viruses, which are involved in infection incidents. It is often the case that cancer, metabolic or neurogenetic disorders originate from human RNAs. The typical classification of RNA molecules is according to their capacity to code for proteins. Messenger RNAs (mRNAs) undergo translation into proteins, while transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) are non-coding RNAs (ncRNAs). However, mRNAs also contain untranslated regions (UTRs) and are derived from pre-mRNAs through intron splicing. A significant number of ncRNAs are involved in pivotal cellular processes, including amino acid transport for protein translation, peptide polymerization catalysis, transcription regulation, and signaling pathways.[1] Some functional RNAs are found almost exclusively in a particular group of organisms. A prominent example of these RNAs are riboswitches, which are found in the 5'-UTR of bacterial mRNAs and act as

control elements for translation and transcription.[7–9] The designation of riboswitches as targets, as exemplified by pre-queuosine-1 ($preQ_1$) riboswitches, provides a preliminary framework for the rational design of ligands for these structures.[9,10] Despite the plethora of antibiotic drugs currently available on the market, the potential for toxic side effects (e.g., aminoglycosides[11,12], fluoroquinolones[13]) and the emergence of antibiotic resistance necessitates continued vigilance with regard to alternative antibiotic targets.[9,14,15]

In contrast to riboswitches, relevant viral RNAs lack natural ligands. Their functions rely for example on replication initiation by interacting with specific proteins.[10,16] However, the highly conserved tertiary structure of viral ncRNA might possess pockets for small molecule ligand binding, which renders them an attractive target for further investigation.[1,10,16] The most advanced studies in this field are on the hepatitis C virus (HCV)[17–21] and human immunodeficiency viruses (HIV)[16,22,23]. Promising antiviral therapeutic strategies include the use of antisense oligonucleotides (ASOs) and clustered regularly interspaced short palindromic repeats (CRISPR) gene editing as future perspectives fighting viral infections.[24,25] These approaches are additional to long-time established nucleoside and non-nucleoside structured drugs, which affect viral proteins such as RNA polymerase.

Targeting human RNAs for cancer or neurogenic disease treatment (e.g., spinal muscular atrophy (SMA)) is a more recent approach. The relevant targets include micro RNAs (miRNAs), long non-coding RNAs (lncRNAs), pre-mRNAs, as well as distinct sequence motifs such as short tandem repeats or repeat expansions.[1] As with the targeting of viral RNA, the two common approaches are the use of RNA-based drugs (e.g., ASOs, small interfering RNAs, and CRISPR technologies)[26–28] and non-RNA-based drugs (e.g., small molecules).[29–31] While ASOs, siRNAs and CRISPR editing have greatly advanced chemical biology, their clinical usefulness has been hindered by delivery challenges and notable adverse effects. In contrast, small molecules present a promising alternative, offering potential advantages such as oral bioavailability, the ability to penetrate the blood-brain-barrier and the independency from primary RNA sequences. Leveraging insights from medicinal chemistry, these molecules can be systematically optimized to enhance affinity, pharmacokinetic properties, and selectivity, underscoring their significance in therapeutic development.[2,32] In 2020, these approaches led to the approval of risdiplam (Figure 2 A) for the treatment of SMA, the first small molecule drug targeting other RNA than

the ribosome.[33] SMA is characterized by a deficiency in the human survival of motor neuron (SMN) protein, which is caused by a mutation or deletion of the related SMN1 gene (Figure 2 B). The SMN2 gene also produces SMN, although at levels that are insufficient for the maintenance of normal function. This is due to a point mutation that results in the exclusion of exon 7 in the majority of mature SMN2 transcripts, thereby leading to the production of unstable SMN proteins (Figure 2 C). Risdiplam acts as a splicing modulator of the SMN2 pre-mRNA. By binding to two sites within the SMN2 exon 7 the inclusion of this exon in the mature transcript is promoted resulting in the production of sufficient functional SMN proteins to compensate for the loss of SMN1 gene (Figure 2 D).[29,34]

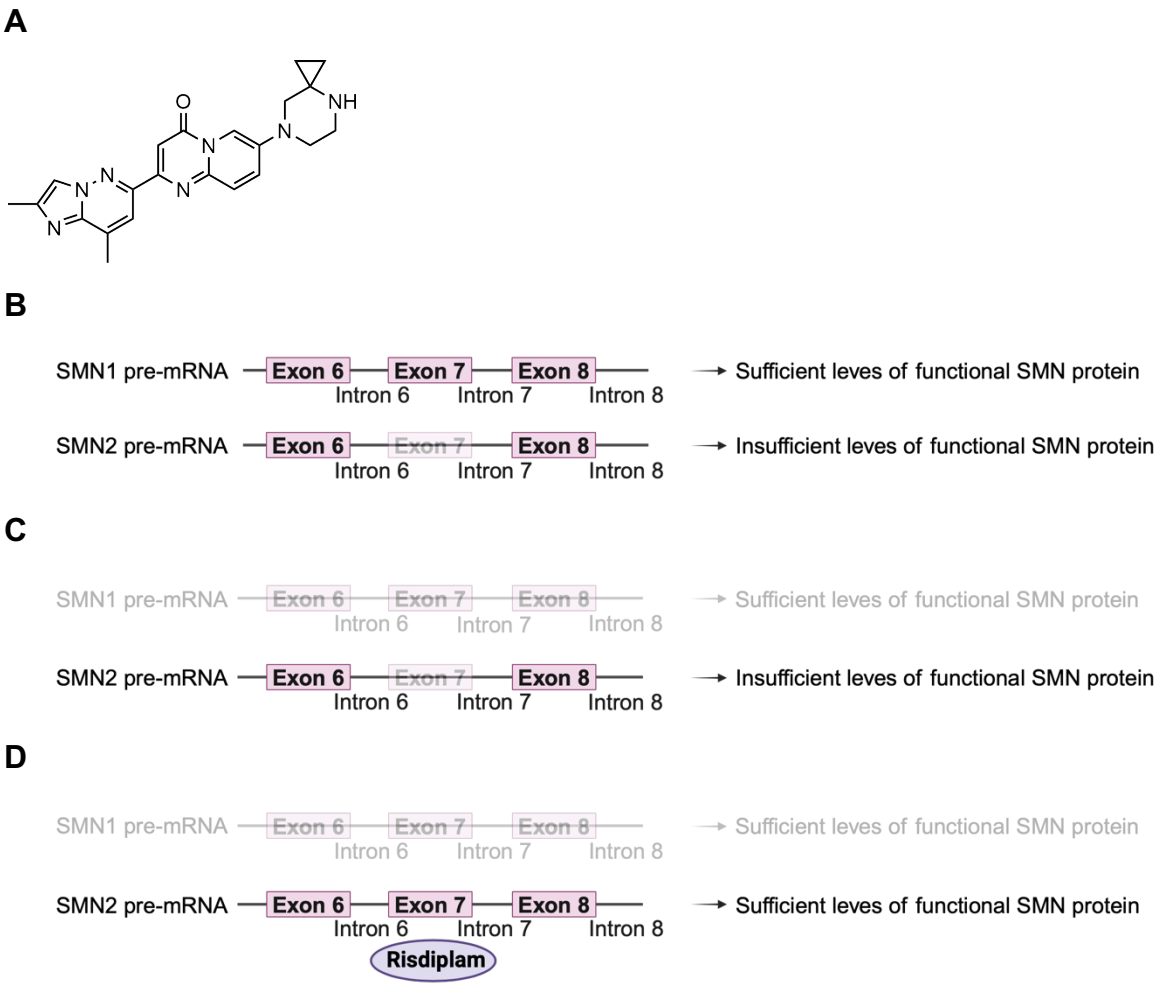


Figure 2: Structure and mechanism of action of the splicing modulator risdiplam. **A**, Structure of risdiplam. **B**, Sufficient functional SMN protein production in healthy humans by SMN1 pre-mRNA splicing. **C**, Deletion, or mutation of the SMN1 gene in SMA patients leads to insufficient functional SMN protein production by SMN2 pre-mRNA due to exclusion of exon 7 during splicing. **D**, Treatment with risdiplam leads to inclusion of the exon 7 in SMN2 pre-mRNA splicing process and production of sufficient levels of functional SMN protein. This figure was compiled according to references [29,34] . Panel **A** was created with ChemDraw 20, panels **B**–**D** were created with BioRender.com.

Currently, druggable and potential druggable RNA targets include bacterial ribosomes[35–39], viral RNA motifs[19–21,40], bacterial riboswitches[41–44], precursor mRNA[29– 31], and miRNAs[44–48] (Table 1). These RNAs exhibit highly structured and functional sites for small molecule binding besides the abundance in the target pathogens and cells.[32] Nevertheless, identification of RNAs and targeting those structures selectively remains a challenge due to their high chemical similarity as well as the limited chemical functionality and highly dynamic structure.[49,50]

Table 1: List of potential/approved drugs and their target RNAs, type, (potential) application and current development status. This table was compiled according to reference [32].

Target RNA	Compound (Class)	Type	(Potential) Application	Current Status
Ribosomal RNA	• Aminoglycosides • Macrolides • Oxazolidinones • Tetracyclines	Small molecules	Antibiotic	Approved
SMN2 pre-mRNA	• Risdiplam	Small molecule	SMA-treatment	Approved
HCV IRES	• Benzimidazoles	Small molecules	HCV infections	Preclinical development
FMN Riboswitch	• Ribocils	Small molecules	Antibiotic	Preclinical development
Pri-miRNA	• Targaprimir-96	ASO	Anticancer	Preclinical development
Pre-miRNA	• Targapremir-210	ASO	Anticancer	Preclinical development

1.1.2. Target Identification and Validation

The identification of disease related RNAs can be challenging, particularly when there is no prior knowledge of their structure. Currently, there are several experimental and computational strategies for identifying these RNAs. These approaches are more concerned with the RNA's effect in cells than with its physical structure so that there is no structure model needed initially. The process of finding relevant RNAs involve transcriptomic analyses, functional genomics studies, or bioinformatics to pinpoint relevant RNA targets.[51] Advances in sequencing technologies, such as RNA sequencing (RNA-seq), have enabled the discovery of novel RNA species and their regulatory roles. The sequenced RNA populations in a diseased tissue sample can be compared to RNAs in healthy tissues. Thus, differentially expressed RNAs potentially associated with the disease can be identified. While less comprehensive than RNA-seq, microarrays can be used to profile the expression of thousands of known RNA sequences simultaneously, identifying those differentially expressed in disease conditions. In this field, computational methods employ known genomic data to predict RNA sequences associated with disease processes. Subsequently, the identified RNA must be validated. This entails confirming the involvement of the identified RNA in the disease process and its functional role, which could be targeted for therapeutic intervention. For this, the expression of the RNA is analyzed. The most common methods for this purpose are quantitative polymerase chain reaction (qPCR), northern blotting, and *in situ* hybridization. Functional studies, including knockdown or overexpression experiments, can be employed to observe phenotypic changes that are directly linked to specific RNAs. Furthermore, functional assays such as reporter assays or rescue experiments, can be carried out. High-throughput screenings (HTS) are frequently employed for target validation. HTS involves testing large libraries of molecules or genetic constructs (e.g., siRNAs) to identify those that produce a desired effect in a biological assay. HTS screens often utilize RNA interference (RNAi) or clustered regularly interspaced short palindromic repeats (CRISPR) systems, which target specific RNA sequences rather than structures. The efficacy of these models is contingent upon the complementarity of the target RNA sequence. Given that the primary objective of HTS in this instance is to observe phenotypic changes, the structure of the RNA is not a relevant factor. Recent studies employed phenotypic screenings of large compound libraries to identify potential small molecule drugs.

Subsequent analysis revealed that these drugs target specific RNAs.[30,31,41] Phenotypic screening represents a target agnostic approach used in drug discovery to identify potential therapeutic compounds based on their effects on observable phenotypes or characteristics of cells, tissues, or whole organisms. Unlike target-based screening, where compounds are tested for their ability to interact with a specific molecular target, believed to be involved in a disease process, phenotypic screening focuses on the overall functional outcome of drug treatment. Relevant cellular models, such as cell lines, or organoids, are selected to reflect the physiological context of the target disease or pathway. Assays are designed to measure specific phenotypic endpoints, often utilizing imaging, flow cytometry or fluorometric biological assays.[2,52] Currently, there is one RNA-targeting drug on the market that was developed using phenotypic screening: the splicing modulator risdiplam for the treatment of spinal muscular atrophy (SMA).[29–31] The non-approved ribocil, which targets the FMN riboswitch, was also discovered by phenotypic screening of the bacterial riboflavin pathway.[41] Another approach for target validation in cells is chemical cross-linking and isolation by pull-down (Chem-CLIP). A small molecule of interest is modified to include a chemical reactive group, the cross-linker, that forms a covalent bond with its target RNA upon interaction. This step captures the small molecule in direct contact with its target RNA. Additionally, the small molecule is tagged with a label (e.g., hemagglutinin or biotin) that can be recognized by specific antibodies or streptavidin. After cross-linking, the complex is isolated from the cell using immunoprecipitation, which involves pulling-down the labelled complex via an antibody (e.g., anti-hemagglutinin antibody) or affinity tag (e.g., streptavidin-biotin). After complex isolation, the bound RNA can be identified by techniques like RNA-seq for exact determination.[45,53–56]

1.1.3. Identification of Small Molecules – Common Strategies

Traditionally, RNA-binding small molecules are identified using HTS of large compound libraries against the RNA target. Several well-described in vitro methods, including fluorescence-based and label-free methods, as well as DNA-encoded library (DEL) screenings, and computational methods can be used to identify potential small molecule RNA binders.[2,6] The interest in the development of small molecules that target RNA for the treatment of various human diseases has grown considerably in recent years.[2,32,57] Given that there are RNA-targeting drugs on the market that do not

adhere to Lipinski's rule of five, which were subsequently found to have severe off-target effects (e.g., aminoglycosides), the latest approaches in this field have focused on following these conventional rules.[2,32,58] An example of successful adaptation of the protein drug discovery concept is linezolid (target-RNA: 50S ribosomal subunit. Linezolid exhibits selective binding to a distinct binding pocket in the target RNA, with molecular weight below 500 Da and low tPSA values, which are indicative of good membrane permeability.[5,35] Linezolid was identified using traditional drug discovery methods, particularly screening large compound libraries and medicinal chemistry for improvement of potency, selectivity, and other drug-like properties.[35,59] Ribocil (target-RNA: FMN riboswitch)[41,60] and risdiplam (target-RNA: SMN2-pre-mRNA)[30,31], which were discovered more than one decade later, were identified by phenotypic screening and subsequent optimization strategies. These examples illustrate that RNA-binding small molecules are predominantly identified through HTS screens, whereas rational drug design represents an exception.[1]

The most performed assay types to identify small molecule binding to RNA are fluorescence-based assays adapted from protein interaction studies.[6] Commonly, the RNA is labelled at its 3'- or 5'-terminus with a fluorescent probe, such as cyanine (Cy3, Cy5) or Alexa Fluor dyes or alternatively with modified nucleotides. Upon ligand binding, physical properties of the target RNA change and subsequently can be monitored by the fluorescence signal. These fluorescence-based assays offer high sensitivity, specificity, and scalability, making them valuable tool for identifying small molecules that bind RNA (Figure 3). As each assay has advantages and limitations, the choice depends on specific requirements of the screening experiment.[2]

Fluorescence indicator displacement (FID) assays (Figure 3 A) utilize the displacement of a fluorescent intercalator dye, such as TO-PRO-1 or Sybr dyes, by a small molecule. When a small molecule binds to the RNA, it displaces the intercalator dye, leading to changes in fluorescence signal.[61] Another method is the measurement of changes in fluorescence using the 2-aminopurine (2-AP) assay (Figure 3 B). In this assay, internal adenine residues of the target RNA are replaced with 2-AP, a fluorescent analogue of adenine, whose fluorescence is highly sensitive to its microenvironment. This sensitivity is exploited to detect binding events. Both assays require sample amounts typically in the nM to µM range and provide a fluorescence intensity read-out.

FRET assays (Figure 3 D) utilize the transfer of energy between two fluorescent molecules, a donor, and an acceptor, when they are in close proximity. In RNA-binding

assays, the RNA is typically labelled with a donor fluorophore (e.g., Cy3) at one end and an acceptor fluorophore (e.g., Cy5) at the other end. Due to conformational changes in the RNA structure upon ligand binding, changes in the FRET signal are observed.[1,2,6] As the fluorescent read-out depends on the distance of the reporter dyes, this method is more restricted than other fluorescence-based methods.

Fluorescence polarization (FP) assays (Figure 3 E) measure changes in the rotational speed of fluorescently labelled RNA upon binding to a small molecule. When a small molecule binds to the RNA, it alters the rotational speed, resulting in changes in polarization. While having a low sample requirement (nM to µM), FP is restricted by the size of the interacting molecules.

MST assays (Figure 3 F) are based on the altered molecular movement in a microscopic temperature gradient. The thermophoresis depends strongly on properties such as size, charge, hydration shell and conformation, which are influenced by ligand binding and is tracked via fluorescence over time.[6,62] MST has the great advantage of requiring minimal sample amount (pM to nM). The observed thermophoretic movement strongly depends on the interaction, making MST a screening method that might not be suitable in general.

An additional approach in this field is the use of small molecule microarrays (SMMs) with labelled RNA (Figure 3 C). A microarray is a high-throughput experimental tool for the simultaneous analysis of the expression levels or presence of thousands of genes or other biomolecules in a sample. In a small molecule microarray (SMM), the small molecules are immobilized onto a solid support, such as a glass slide, each spot on this microarray represents a different compound. The sample containing the target RNA is applied to the array and subsequently incubated under conditions that allow binding interactions between the RNA and immobilized small molecules. Following the removal of unbound components, the array is subjected to analysis in order to identify and quantify binding events. Detection methods vary depending on the target and experimental goals, in most cases the RNA is fluorescently labelled. Recently, the SMM technology was successfully employed to identify binders for the $preQ_1$ riboswitch[63] and the human immunodeficiency virus (HIV) transactivating response (TAR) region[22].

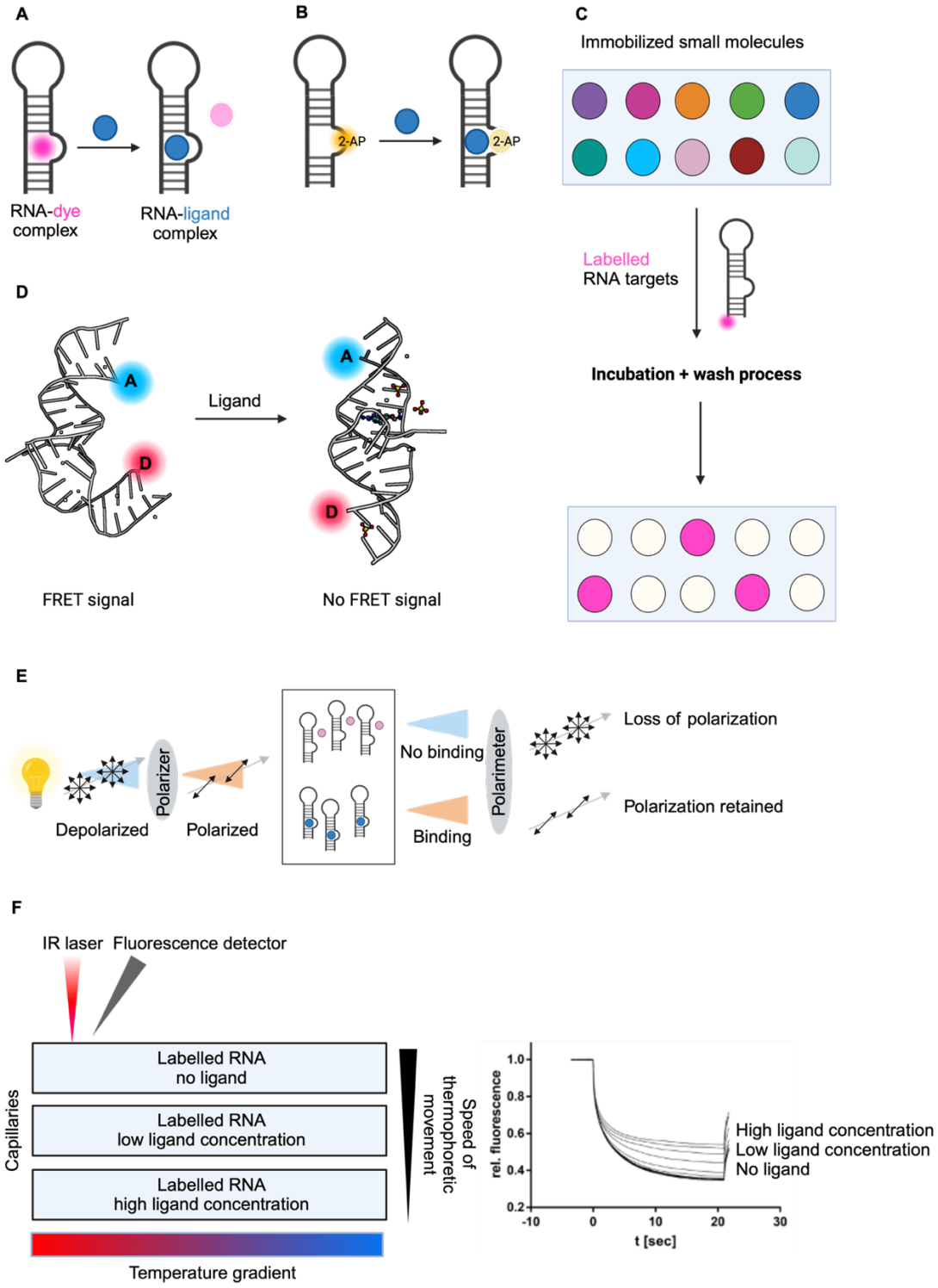


Figure 3: Fluorescence-based assay types for identification of small molecule RNA binders. **A**, FID, a non-selective RNA-binding dye is displaced by a ligand resulting in loss of fluorescence signal. **B**, 2-AP assay, ligand binding changes the microenvironment of the modified nucleotide 2-AP resulting in loss of fluorescence signal. **C**, SMM screening, a panel of small molecules is immobilized on a surface and incubated with labelled RNAs. After a washing process, the bound molecules can be identified. **D**, HCV IRES FRET assay, the RNA is labelled with donor and acceptor dye. Due to ligand binding a conformational change induced a change in FRET signal. **E**, FP, plane polarized light is used to measure the molecular rotation of a labelled sample. **F**, MST, measurement of the thermophoretic movement of a labelled sample in a temperature gradient which depends on molecular properties. This figure was compiled according to references [2,6]. This figure was created with BioRender.com.

Alternatively, label-free screening methods include mass spectrometry (MS) and DEL, as well as SPR, which is discussed in more detail below.

MS is a technique employed to ascertain the mass-to-charge ratio of ions, thereby furnishing data regarding the molecular composition and structure of analytes present in a sample.[2] Affinity selection mass spectrometry (AS-MS) provides a label-free variant of MS utilized for the direct identification of target-ligand complexes through mass spectrometry following the separation of unbound ligands using size-exclusion chromatography (SEC). While extensively applied in protein studies, its utilization in RNA target identification is relatively recent. This method was established by Rizvi *et al*. who demonstrated the feasibility of this application by screening a large compound library against the FMN riboswitch and identified a few specific hits.[60] A modified version of AS-MS, the automated ligand identification system (ALIS), indirectly detects target-ligand interactions by dissociating the formed complex prior to liquid chromatography-mass spectrometry (LC-MS) to identify the bound ligand.[64] The extremely low sample requirement (fM to nM) is in stark contrast to the more complex preparation.

DNA-encoded libraries (DEL) are large collections of small molecule compounds, each of which is covalently linked to a unique DNA-sequence, first introduced in 1992.[65] DEL technology offers a powerful and efficient approach for screening large chemical spaces to identify potential RNA binding small molecules. The DNA-tagged small molecule library is incubated with the RNA of interest. The binding of small molecules to RNA results in the formation of complexes, in contrast to the unbound molecules. This process of separation due to complex formation is known as affinity selection. Subsequently, the mixture is partitioned to separate bound complexes from unbound molecules. Bound complexes retain the DNA-“barcodes”, while unbound molecules are removed. The DNA-“barcodes” of bound molecules are then amplified using polymerase-chain reaction (PCR) to increase their concentration for downstream analysis. The identified hit molecules are then synthesized “off-DNA” and validated using secondary assays like fluorescence-based assays and can further undergo optimization processes. The identification of RNA-binding small molecules remains a challenging task due to the high risk of false positive DNA-RNA binding events. This issue can be addressed by statistical differentiation of DNA-RNA and small molecule-RNA interactions, which allows for the identification of true positive signals during DEL

selection.[2,66] Two recent studies have demonstrated the capacity of DELs to identify small molecules that bind to their target RNA with high affinity and selectivity.[67,68] Furthermore, SPR can be utilized as a screening method to identify small molecules that bind to a specific RNA.[69,70] Surface Plasmon Resonance (SPR) occurs when polarized light interacts with a metal surface situated at the junction between material with distinct refractive indices. When the angle of incidence is precise, electrons positioned on the metal surface become excited and oscillate, absorbing a portion of the light while reflecting the remainder. Altering the composition of one of the surfaces affects its refractive index, consequently modifying the resonance angle at which light is absorbed.[6,70]

1.1.4. Determination of RNA Structures

Structure-based design requires knowledge of the architecture and 3D structure of RNA molecules to design small molecules that can interact with specific binding sites or functional regions. This understanding can be used to design small molecules that complement or disrupt these structures, modulating their function.

RNA architecture can be divided into primary, secondary, and tertiary structure. The primary structure is represented by the nucleotide sequence consisting of the four nucleobases adenine (A), cytosine (C), guanine (G) and uracil (U). These aromatic nitrogen-containing bases are linked to ribose groups by phosphodiester bonds, the so called negatively charged “backbone” (Figure 4 A).[1] RNA can adopt a variety of secondary structures, including bulges, helices, loops, and stems (Figure 4 E). This huge structural variability is referred to different interactions such as Watson-Crick-Franklin (WCF) hydrogen bonding, stacking between planar nucleobase surfaces or Hoogsteen-sugar edge conformations (Figure 4 B). Additionally, the glycosidic bond can be orientated *cis* or *trans* (Figure 4 C). Spatial interactions of secondary structure elements enable the formation of tertiary structures, such as multiway-junctions and pseudoknots (Figure 4 D). In addition, quaternary structures, which are known as RNA complexes, can be formed.[1,71,72] In particular, larger, macromolecular RNA can form pockets for high affinity small molecule binding. Given that RNA is composed of only four nucleotide building blocks (A, U, C and G), it appears to be less variable than protein structures which consist of 20 proteinogenic amino acids. However, numerous

additional interactions are feasible via the nucleobase, ribose, and phosphate backbone than initially apparent.[49]

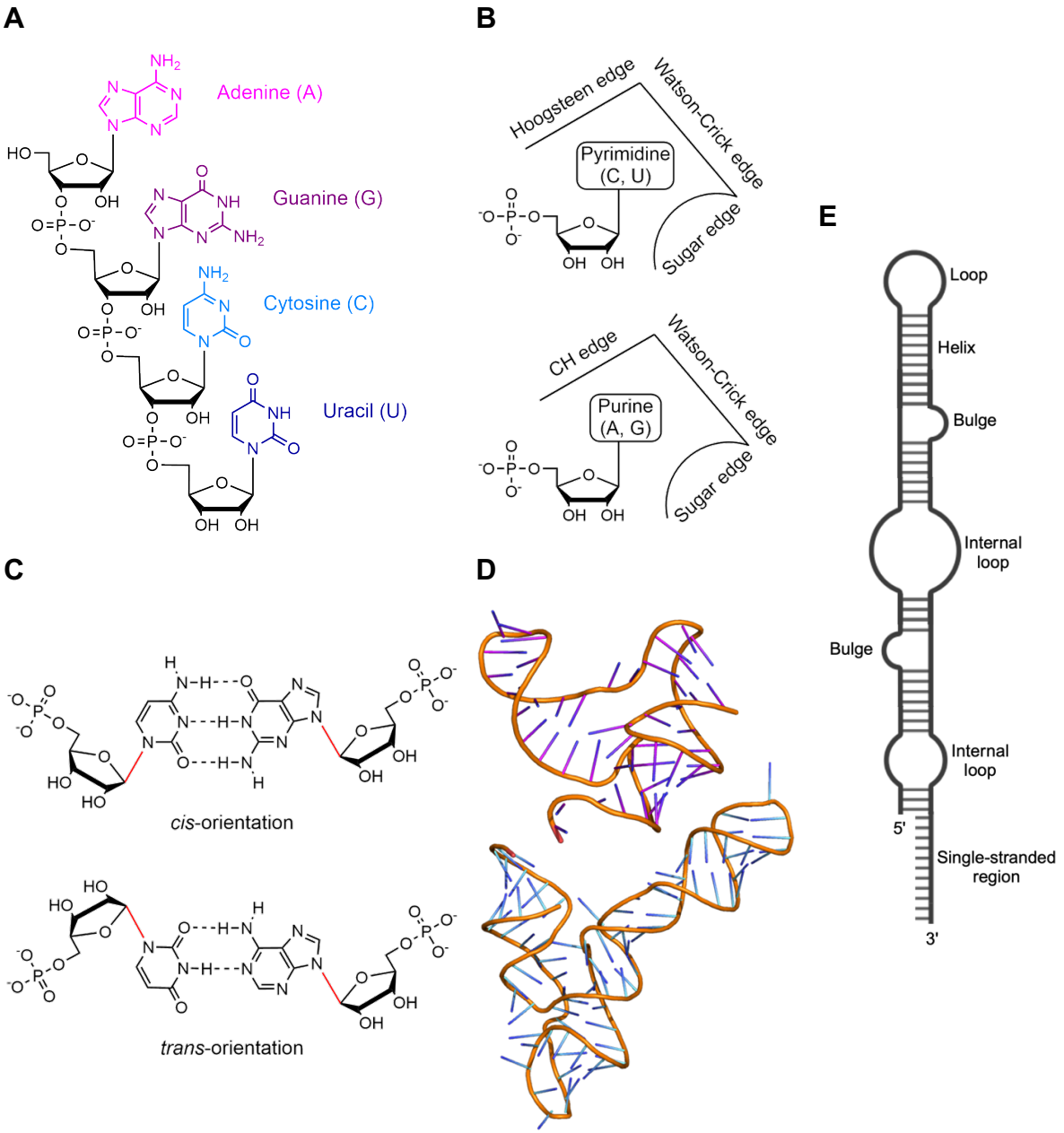


Figure 4: RNA primary, secondary and tertiary structures. **A**, Nucleotide building blocks of RNA primary structure. Each consists of a ribose sugar, phosphodiester groups, which connect the 5'- and 3'-hydroxyl groups of the ribose sugars and adenine (A, pink), guanine (G, purple), cytosine (C, light blue), or uracil (U, dark blue) nucleobases. **B**, Interaction sides of nucleotides. Purine nucleotides can interact via Hoogsteen, Watson-Crick or sugar edges (top); pyrimidine nucleotides can interact via CH, Watson-Crick or sugar edge (bottom). **C**, *Cis*- (top) and *trans*- (bottom) orientation of the glycosidic bond (red). **D**, Example of canonical secondary structure elements, which can be observed in various RNAs. Solid lines represent A–U and G–C Watson-Crick-Franklin base pairs, bulges represent wobble base pairs. **E**, Pseudoknot, and multiway junction as examples of tertiary structures. Pseudoknot structure in the fluoride riboswitch (PDB: 3VRS, left) and three-way junction in the glycine riboswitch (PDB: 3OX0, right). Panels **A**–**C** were created with Chemdraw 20, panel **D** was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC, and panel **E** was created with BioRender.com.

Additionally, eukaryotic RNA can adopt post-transcriptional modifications, such as N^6- methyladenosine (m^6A), pseudouridine (Ψ) and 5-methylcytidine (m^5C).[73,74] Currently, there are more than 170 RNA modifications reported.[75] These small chemical changes strongly influence RNA folding and function. Thus, a dysregulation in those modifications is involved into the pathogenesis of various human diseases, including cancer, neurological and metabolic disorders.[73,74]

The high dynamic variability of RNAs presents a significant challenge to their exact structural determination, which may appear to be a relatively straightforward process at first glance.[2,76]

Techniques such as nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography, and cryo-electron microscopy (cryo-EM) are extensively used to determine RNA structures. While cryo-EM traditionally studies large molecules[2,76], efforts are ongoing to develop approaches for smaller RNA structures[77].

RNA structures can also be predicted using computational methods such as RNAfold (secondary structure) and FARFAR2 (tertiary structure).[2] Typically, the RNA structures are investigated by NMR spectroscopy or X-ray crystallography. Subsequently, computational docking studies and bioassays are used to identify small molecule binders.[2,78] Furthermore, computational tools can be used to simulate their molecular dynamics from NMR or crystal structure models.[2]

Computational methods utilize algorithms such as free energy minimization and phylogenetic comparison to predict RNA structures from sequence data. Free energy minimization calculates the minimum free energy structure, along with suboptimal structures, based on experimentally obtained thermodynamic parameters. Programs like RNAfold can predict base pairs with an accuracy of approximately 70%.[2,5] Experimental techniques, including chemical modifications of unpaired bases with dimethyl sulfate (DMS) and Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE) probing, enhance prediction reliability.[79–81] These techniques provide insights into RNA structural profiles and conformational changes under different biological conditions.

Genetic differences that preserve RNA secondary structure, known as covariations, offer insights into RNA structure. Covariation-based prediction methods, although challenging due to dataset size and complexity, enhance accuracy when coupled with other prediction methods like homology modelling and free energy minimization. Tertiary RNA structures can be predicted using programs like FARFAR2.[2,82] These

methods, though less mature compared to protein structure prediction, show promise in predicting 3D RNA structures from sequence data.[2]

Subsequently, these structure models can be used for structure-based virtual screening and docking. This approach is also adapted from protein-based drug discovery and is still less common than the traditional strategies for identification of small molecules. A reason might be that there are less 3D structures available for RNAs than for proteins.[76] Additionally, there are several difficulties to be considered similar to protein studies, including the tautomeric[42] and protonation[83] states of the ligands, the RNA's dynamics[84], and solvation[85].

The collected data in this process are essential for the identification of high-specificity and high-affinity druggable RNA structures. Once target RNA structures are identified, various RNA-targeting molecules can be designed and optimized to selectively interact with them.[2,5,32,86]

1.1.5. Structure-Based Small Molecule Design and Optimization

In comparison to proteins, RNA is relatively hydrophilic due to its phosphate backbone. This negative charged phosphate groups generates a substantial electrostatic potential, which exerts a considerable influence on its surrounding ionic atmosphere and necessitates the presence of cations such as Mg^{2+} for folding.[87–89] As described in several electrostatic theory models, polyelectrolyte molecules such as RNA favor the surrounding of ions for neutralization.[87,89–91] Therefore, positively charged (small) molecules can interact with the negatively charged RNA backbone with high affinity but low selectivity which can result in toxic off-target effects and low cellular uptake due to the positive charge.[5] The antibiotic class of aminoglycosides represents a category of basic RNA-binding molecules. The first member of this class, streptomycin (Figure 5), was introduced in 1944 as an anti-tuberculosis agent. Aminoglycosides consist of a core amino sugar moiety connected by glycosidic linkers to a dibasic aminocyclitol moiety, which binds with high affinity to the A-site of 16S ribosomal RNA of 30S subunits. Subsequently, the protein synthesis is altered resulting in the assembly of polypeptides that damage the bacterial cells.[11] With molecular weights greater than 500 Da and their basic properties, aminoglycosides violate the Lipinski rule of five and lack selectivity, which results in suboptimal pharmacological properties.[32,58,92] Further serious side effects of aminoglycosides are nephrotoxicity

after glomerular filtration in the proximal convoluted tubule and ototoxicity by damaging the sensory hair cells.[12]

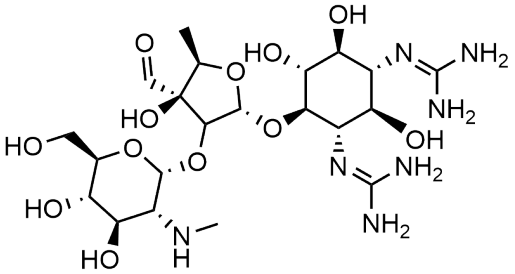


Figure 5: Structure of the aminoglycoside antibiotic streptomycin. This figure was created with Chemdraw 20.

Nevertheless, the design of tight and selective RNA-binding small molecules that exploit the negatively charged phosphodiester backbone in combination with the binding site, is still a promising strategy.[93] Typically, the association rates of small molecules binding to RNA are slow compared to their diffusion, or in a broader context, compared to protein-ligand binding.[94] This strategy is known as electrostatic anchoring and has been well-studied for protein interactions with charged molecules.[95–98] By attaching a positively charged moiety to a known selective RNA binder, this may enhance the association speed.[10,93,99,100] The positive charge simultaneously anchors the molecule strongly to the RNA surface.[93,99] However, a remaining challenge for this concept is to find a balance between charge and selectivity. The small molecule must fit well in the binding pocket after electrostatic attraction and subsequently be anchored. It can be postulated that otherwise the charge will dominate the interaction resulting in a loss of selectivity and a fast on/fast off kinetic behavior. However, this assumption still needs to be corroborated.

Furthermore, fragment-based ligand discovery (FBLD) can be used within a rational and systematic structure-based approach to identify binders. This strategy involves screening libraries of small, low molecular weight (≤ 300 g/mol) compounds, known as fragments, to identify those that bind to specific regions or motifs within the RNA molecule. Instead of screening large, diverse compound libraries, FBLD screens a smaller library of fragments. These fragments are designed to cover a broad range of chemical space while maintaining drug-like properties. Given that they are fragments, their binding affinity is often relatively low, requiring sensitive biophysical methods such as NMR, SPR, and X-ray crystallography for detection. Fragments, that bind to the target RNA are identified as hits and subsequently characterized based on their

binding affinities, binding modes, and structural features. As a second step, chemical modifications involving addition of functional groups, known as fragment growing or linking fragments together are performed. The synthetical extension of fragments can enhance their fit into distinct binding sites. If two fragments bind into the same pocket at different positions, they can be linked synthetically using flexible or rigid chemical linkers for optimal positioning into the binding site. This process leads to optimization of affinity, pharmacokinetics, and selectivity, maintaining favorable drug-like properties.[2,101,102]

These properties include the key metrics ligand efficiency (LE), ligand-lipophilicity efficiency (LLE), and fit quality (FQ).[103,104] LE measures the binding affinity of a ligand relative to its molecular size, whereas affinity solely quantifies the strength of an interaction without considering the size of the ligand. Smaller fragments typically exhibit reduced binding affinities compared to larger molecules due to the limited number of interactions with the target. LE addresses this discrepancy by evaluating the contribution of each heavy atom to the overall affinity, providing a more accurate representation of the fragment's utilization of its heavy atoms. LE is expressed as the binding energy (ΔG) per heavy atom (HA) in units of kcal/mol per non-hydrogen atom (equation (1)). This metric helps identifying fragments that make efficient use of their size to achieve strong interactions with the target molecule.[103–105] The recommended range for fragment hits during optimization is an LE value ≥ 0.3 kcal/mol.[103]

$$LE = -\frac{\Delta G}{HA} = -\frac{RT \cdot \ln(K_D)}{HA} \quad (1)$$

LLE is a measure that balances binding affinity with lipophilicity, promoting the selection of compounds that achieve high affinity without excessive lipophilicity, which can lead to poor solubility and undesirable pharmacokinetic properties. This is an important parameter during optimization campaigns, as the affinity often is enhanced through the introduction of lipophilic moieties. LLE is expressed as the difference between the pIC_{50} and the logP value (equation (2)).[103,104,106] The recommended range for fragment hits during optimization is an LLE value ≥ 3.[103]

$$LLE = pIC_{50} - logP \quad (2)$$

FQ evaluates how well the observed binding affinity scales with its molecular size, helping to prioritize compounds that show efficient use of molecular weight in enhancing affinity.[103,104,107] It compares the actual LE to the expected efficiency based on its molecular weight. To obtain FQ, the LE is divided by the calculated maximal achievable LE (LE_Scale; equation (3)).[103,104,107] LE_Scale can be calculated in different ways[108–110]. The recommended range for fragment hits during optimization is an FQ value of 0.8–1. A high FQ indicates that a compound uses its size effectively to achieve binding, making it a valuable metric in FBLD where efficient growth from fragments to drug-like molecules is critical.[103]

$$FQ = \frac{LE}{LE_Scale} \tag{3}$$

Throughout FBLD, as fragments are optimized, the aim is to improve affinity while maintaining or improving LE, LLE, and FQ. This ensures that the final compounds are both potent and developable, minimizing undesired properties like high lipophilicity or molecular weight.[103,104] In recent times, there has been a notable increase in the number of RNA-binding small molecules, which have been identified by FBLD as a valuable addition to the toolbox of fundamental methods in this research field aimed at combating human diseases.[55,111–116]

1.2. Antibiotic Drug Discovery

1.2.1. Antibiotics and Multi-Drug Resistance

Nature has served as an enduring wellspring of medicinal resources throughout history, yielding numerous beneficial bioactive compounds derived from plant origins. However, the beginning of the 20th century marked a pivotal juncture with the emerge of penicillins, catalyzing a paradigm shift in pharmaceutical exploration towards microbial reservoirs. A significant proportion of pharmaceutical agents have since evolved from bacterial secondary metabolites. These drugs, have found diverse applications, encompassing diagnostic, palliative, therapeutic and prophylactic interventions against various diseases.[117] The zenit of microbial product screening, spanning the golden era from the 1930s to 1970s, witnessed exhaustive efforts involving the screening of millions of microorganisms.[118] This leaded to a plethora of

microbial metabolites, constituting a vast reservoir of therapeutic entities fighting bacterial infections, notably including widely prescribed antibacterial therapeutics such as described in table 2.[118–120]

Table 2: Overview of the most important currently approved antibiotic drug classes, their main targets and effect on the microorganisms. The table was compiled according to classification from Lewis *et al*.[121]

Bacterial Target	Antibiotic Class	Effect
Folic acid metabolism	• Sulphonamides • Trimethoprim	Bacteriostatic
Protein synthesis (30S ribosomal subunit)	• Spectinomycin • Tetracyclines	
Protein synthesis (50S ribosomal subunit)	• Chloroamphenicol • Macrolides • Oxazolidinones	
Cell wall synthesis	• Carbapenems • Cephalosporins • Glycopeptides • Monobactams • Penicillins	Bactericidal
Cell membrane	• Polymyxins	
DNA gyrase	• Quinolones	
Protein synthesis (30S ribosomal subunit)	• Aminoglycosides	
RNA polymerase	• Ansamycines	

However, the success of this golden age has gradually faded as the effectiveness of traditional antibiotics has been surpassed by the persistent development of antibiotic resistance among pathogenic microorganisms.[122,123] There are many mechanisms of bacterial drug resistance, including enzymatic degradation (e.g., β-lactamases) and target alteration (e.g., topoisomerases) as most prominent examples.[124] An increasing number of pathogens are multidrug resistant (MDR), withstanding potentially lethal

doses of at least three antibiotic classes in vitro. This group of bacteria is called ESAKPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae, Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter species*) pathogens and responsible for most hospital-acquired infections.[15,123]

The transition of the industry from modifying existing antibiotics to pursuing novel therapeutic agents capable of tackling resistant pathogens has been filled with persistent setbacks and disillusionment. The escalating crisis of antibiotic resistance poses an imminent threat to global public health, transcending socio-economic boundaries and geographical divides.[122,123] Clinical isolates of numerous pathogenic bacterial species, including *Mycobacterium tuberculosis*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus*, exhibited alarming levels of resistance to conventional antibiotics, rendering once effective treatments obsolete.[125] In recent years, there has been a significant increase in global initiatives to combat antibiotic resistance. These efforts are led by prominent health organizations such as the World Health Organization (WHO) or the European Society for Clinical Microbiology and Infectious Diseases (ESCMID).[123] The WHO publishes a list of priority pathogens in order to guide and promote research and development of new antibiotic agents, including pathogenic bacteria from the ESKAPE group.[126] In contrast, pharmaceutical companies increasingly reduce their effort developing novel antibiotic drugs. This circumstance can be explained by several economic considerations.[121,127] Novel antibacterial drugs yield a comparatively modest return on investment in contrast to other drug classes like antihypertensive or cholesterol-lowering medicals. They are predominantly for the management of severely ill patients (last-resort antibiotics), they are no permanent medications leading to brief treatment duration, and their clinical efficacy is frequently compromised by the swift emerge of resistance, curtailing their utility over time.[124,128]

1.2.2. Perspectives for Novel Antibiotic Drugs

In April 2024 after 17 years of development, the FDA approved ceftobiprole (Zevtera®) for the treatment of *Staphylococcus aureus* bloodstream infections (SAB), acute bacterial skin and skin structure infections (ABSSSI) and community-acquired bacterial pneumonia (CABP).[129] This fifth-generation cephalosporin antibiotic is effective against Gram-positive pathogens, including methicillin-resistant *Staphylococcus aureus*

(MRSA), and Gram-negative pathogens, such as *Pseudomonas aeruginosa* and *Haemophilus influenzae*, but only those without extended-spectrum β-lactamase.[130] Over the past 25 years, two novel classes of antibiotic drugs have been identified and approved: oxazolidinones (including linezolid) and cyclic lipopeptides (including daptomycin). However, these drugs are only active against Gram-positive bacteria. Other newly approved antibacterial drugs are new generation derivatives of the major antibiotic classes, such as ceftobiprole (a cephalosporin) and eravacycline (a tetracycline), which delay the development of resistance rather than eliminating it.[131] Developing new classes of antibiotics with a lower risk of resistance development remains a challenging task. One approach is to design small molecule ligands that target RNA structures in bacteria. Therefore, structures, that only occur in bacteria, such as riboswitches, are promising targets for fighting bacterial infections because they are sensitive to small molecules and are responsible for the synthesis or transport of metabolites essential for bacterial survival. Several compounds have been found to exhibit antibacterial activity through interaction with riboswitch structures, demonstrating this principle of action.[7,14,132,133]

1.3. Riboswitches

1.3.1. Gene Control Mechanisms of Riboswitches

Riboswitches are RNA-based regulatory elements sensing small molecules or elemental ions and undergo structural changes modulating the expression of the relevant genes. Bacterial riboswitches are typically located in the 5‘-untranslated region (5‘-UTR) of certain mRNAs and consist of at least one conserved aptamer domain and a flexible downstream overlapping expression platform. The aptamer region binds selective to the target ligand and interfaces to the expression platform to affect gene expression. Thus, the aptamer region is highly conserved among different species, while the non-conserved expression platform undergoes a conformational change upon ligand binding. Subsequently, this formed structure directly affects the transcription termination or the translation initiation (Figure 6). Therefore, the term “riboswitch” is associated with a ribonucleic acid structure that can switch on or off transcription or translation in response to a ligand-induced conformational change. Many bacteria obtain essential metabolites by transport from their environment or *de*

novo synthesis from precursor molecules, such as $preQ_1$. Distinct proteins are necessary for these processes. Riboswitches usually modulate the expression of genes coding for these proteins involved in transport or synthesis of metabolites, which are usually necessary for bacterial survival. Therefore, their specific ligand often is a precursor or product of the pathway they control. OFF-switches are often part of a feedback loop where the concentration of the ligand controls its own production or uptake: When the intracellular ligand concentration exceeds a specified threshold concentration, it binds to its riboswitch, initiating the control mechanism. Conversely, when the concentration is below the threshold, the gene expression remains unaffected.[8,134–137]

The most common mechanisms of gene expression modulation are transcriptional control and translational control. However, mRNA degradation and alternative splicing mechanisms have also been reported.[136,138–140]

The designation of OFF- or ON-switch is dependent on the relative promotion or prevention of transcription or translation. A transcriptional OFF-riboswitch, such as the coenzyme B_{12} or $preQ_1$ riboswitch, functions by forming an intrinsic terminator stem in the ligand-bound state. An intrinsic terminator represents an extension of the stem loop typically comprised by six or more U residues. Consequently, transcription is terminated prior to the production of the coding region of mRNA, thereby preventing gene expression (Figure 6 A). Conversely, when ligand concentrations are below a certain threshold, transcription of the complete mRNA occurs because the expression platform includes an anti-terminator stem. Transcriptional ON-switches, such as the adenine-specific riboswitch, operate in reverse by forming an anti-terminator stem in the ligand-bound state to promote transcription (Figure 6 A). However, this type of riboswitch is not as prevalent as transcriptional OFF-switches because there is a higher priority preventing cells from accumulating excess amounts of metabolites. The control mechanism at the translational level functions by affecting the mRNA structure. These conformational changes can either lock or open the access to the ribosome-binding site (RBS), thereby preventing or promoting the initiation of protein translation (Figure 6 B).[134,136]

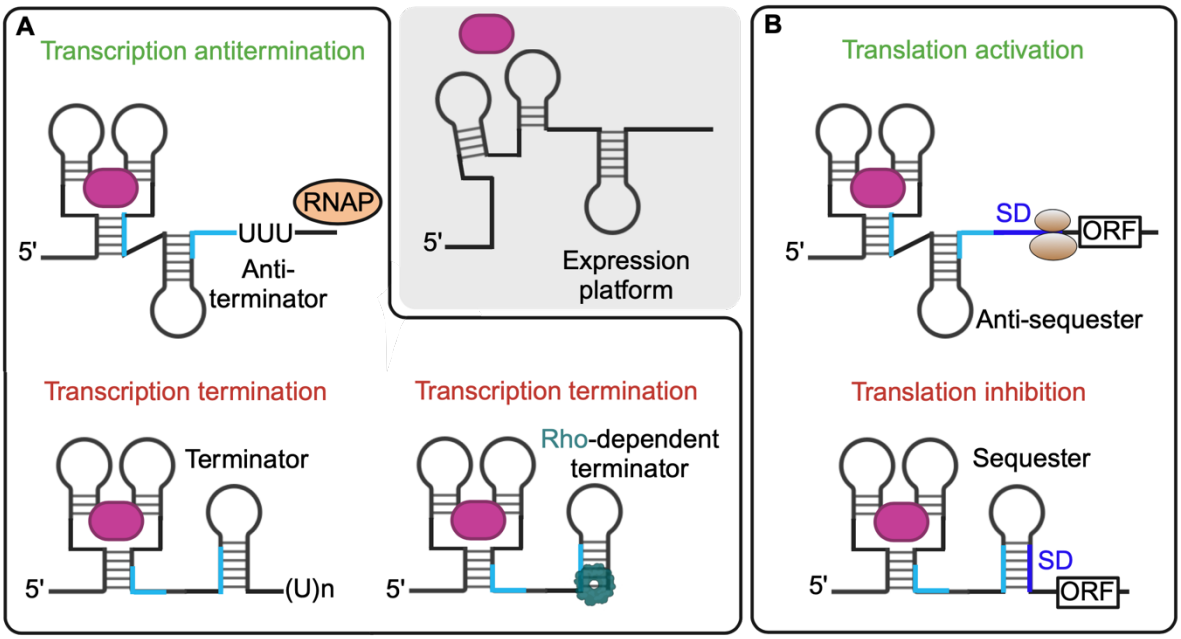


Figure 6: Common gene control mechanisms of riboswitches in bacteria. **A**, Transcriptional control mechanisms: Ligand (pink) binding stabilizes the antiterminator hairpin to allow the RNA polymerase (RNAP) the termination of transcription (top). Ligand binding promotes the formation of a terminator hairpin (Rho-independent: bottom left, Rho-dependent: bottom right) to stop transcription at a premature level. Complementary regions are shown in light blue. **B**, Translational control mechanisms: Ligand binding promotes the formation of a SD-antisequester hairpin for ribosome (brown) binding and translation initiation (top). Ligand binding causes the sequestration of the SD sequence (blue) and the repression of ORF expression, leading to translation inhibition (bottom). Complementary regions are shown in light blue. This figure was compiled with BioRender.com according to reference [136].

Recent studies have demonstrated that the function of transcriptional riboswitches underlies also significant kinetic control. For transcription control, the ligand must bind and affect the expression platform structure before the RNA polymerase proceeds through it. This results in 5'-UTR transcriptional pausing for ligands with slow association rates, which permits ligand-dependent structural rearrangements.[141–143] These mechanisms have been elucidated with several biophysical techniques, which have revealed the kinetic impact on folding and base-pairing, as influenced by the ligand.[144–146]

1.3.2. Riboswitch Classification System

Since the first riboswitches were experimentally validated in 2002[147–150], the collection of different riboswitch structures has been steadily growing. Riboswitches can be organized into distinct classes based on their ligand and folding of the aptamer region. The structure of these highly conserved regions can be studied by X-ray diffraction.[136,137,151,152] Currently, there are approximately 400 PDB (Protein Data Bank) entries for riboswitch structures.[153] Riboswitches are mainly sorted into two groups (types) depending on the architecture of their conserved aptamer domains, while the majority of them possess multihelical junctions and pseudoknot structures (Figure 7). Junctional riboswitches include riboswitches folded into multihelical junctions, such as TPP riboswitches.[154] This type is divided into two subtypes Ia (Figure 7 A) and Ib (Figure 7 B) depending on the position of their key junction in relation to the regulatory helix P1. In type Ia, the junction is centrally positioned, joining the helix, while in type Ib the junction is located in the periphery, far from the helix.[136,137]

Type II includes riboswitches based on pseudoknot structures (Figure 7 C), such as $preQ_1$[155–157] and fluoride[158] riboswitches.[136,137] Pseudoknots are formed by Watson-Crick base-pairing between a stem-loop region and an external region resulting in a knot-shaped conformation. Pseudoknots always consist of at least two stems.[159]

The architecture of the riboswitches is not correlated with their ligand. This is evident in the SAM sensing riboswitches which contain different junctional and pseudoknot structures.[135,160–162]

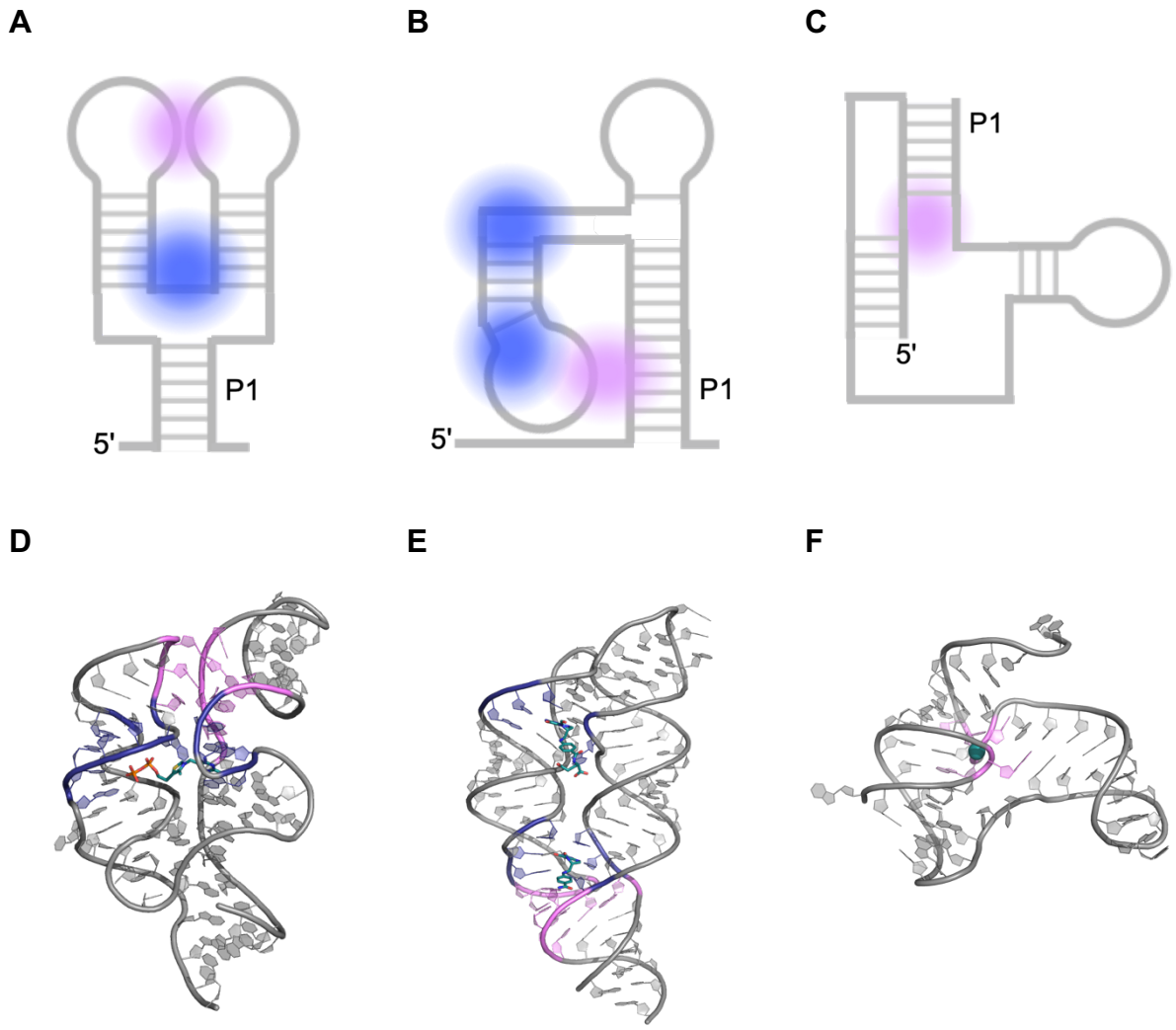


Figure 7: Riboswitch classification system. **A**, Schematic presentation of a type Ia junctional riboswitch with a centrally "straight" to P1 positioned key junction. **B**, Schematic presentation of a type Ib junctional riboswitch with a preipheric "inverse" to P1 positioned key junction. **C**, Schematic presentation of a type II pseudoknot riboswitch. **D**, The TPP riboswitch (shown as cartoon, grey) bound to TPP (shown as sticks, green) as a member of type Ia junctional riboswitches; PDB: 2GDI. **E**, The THF riboswitch (shown as cartoon, grey) bound to THF (shown as sticks, green) as a member of type Ib junctional riboswitches; PDB 4LVV. **F**, The fluoride riboswitch (shown as cartoon, grey) bound to fluoride (shown as sticks, green) as a member of type II pseudoknot riboswitches; PDB: 4ENC. Ligand binding sites are colored in blue and long-distance tertiary interactions are colored in violet. This figure was compiled according to reference [136]. Panels **A**–**C** were created with BioRender.com, and panels **D**–**F** were created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

Riboswitches sense a wide range of different ligands, including ions such as fluoride, amino acids, sugars, enzymatic cofactors (e.g., SAM), and RNA-associated molecules (e.g., $preQ_1$). The sensed molecules have in common that they are involved into essential biological processes and functions that ensure bacterial survival. Currently, there are 56 different classes of riboswitches found based on their aptamer architecture and natural ligands, while the biggest class are the cofactor-binding riboswitches, including AdoCbl-, SAM- and TPP- riboswitches (Table 3).[163]

Table 3: List of different well-studied riboswitch classes, containing at least one example per ligand type, number of representatives for each class (count), control mechanism, fold type, and proven prevalence in relevant human pathogenic bacteria. Notes: (i) SAM-I and -IV are listed in one line due to their highly similar architecture. (ii) *B. angulatum* (*Ba*, SAM-VI) is not a human pathogen, as surrogate structure of SAM-III, SAM-VI serves as a model system. This table was compiled according to Breaker[163] and the specified references.

Class	Count	Common Control Mechanisms	Fold Type	Ligand Function	Prevalence
TPP	16701	Transcription-OFF, translation-OFF[164]	Junction[137]	Cofactor	*P. aeruginosa, S. aureus*, *M. tuberculosis*[9]
AdoCbl	14646	Transcription-OFF, translation-OFF[164]	Junction, pseudo-knot[165]	Cofactor	*S. aureus*, *E. faecium*, *P. aeruginosa*[166]
SAM-I, -IV	10012	Transcription-OFF[160]	Junction, pseudo-knot[160]	Cofactor	*S. aureus*, *P. aeruginosa*[166]
c-di-GMP-I	6971	Alternative self-splicing, translation-ON[136]	Junction[167]	Signal molecule	*V. cholerae*[168–170]
Glycine	6815	Transcription-ON[134]	Junction[137]	Amino Acid	*H. influenzae*, *S. pneumonia*, *S. aureus*[9]

Table 3: List of different well-studied riboswitch classes, containing at least one example per ligand type, number of representatives for each class (count), control mechanism, fold type, and proven prevalence in relevant human pathogenic bacteria. Notes: (i) SAM-I and -IV are listed in one line due to their highly similar architecture. (ii) *B. angulatum* (*Ba*, SAM-VI) is not a human pathogen, as surrogate structure of SAM-III, SAM-VI serves as a model system. This table was compiled according to Breaker[163] and the specified references.

Class	Count	Common Control Mechanisms	Fold Type	Ligand Function	Prevalence
FMN	5615	Transcription-OFF, translation-OFF[164]	Junction[137]	Cofactor	*A. baumannii*, *P. aeruginosa*, *S. aureus*[9]
Fluoride	2583	Transcription-OFF, translation-OFF[171]	Pseudoknot[165]	Detoxification	*A. baumannii*, *P. aeruginosa*, *E. faecium*[9]
Guanine	2334	Transcription-OFF[172]	Junction, pseudoknot[165]	RNA precursor	*S. aureus*[9]
$PreQ_1$-I	2387	Transcription-OFF, translation-OFF[173]	Pseudoknot[165]	tRNA modification	*E. faecium*, *N. gonorrhoeae*[9]

Table 3: List of different well-studied riboswitch classes, containing at least one example per ligand type, number of representatives for each class (count), control mechanism, fold type, and proven prevalence in relevant human pathogenic bacteria. Notes: (i) SAM-I and -IV are listed in one line due to their highly similar architecture. (ii) *B. angulatum* (*Ba*, SAM-VI) is not a human pathogen, as surrogate structure of SAM-III, SAM-VI serves as a model system. This table was compiled according to Breaker[163] and the specified references.

Class	Count	Common Control Mechanisms	Fold Type	Ligand Function	Prevalence
GlmS	2132	mRNA degradation after ribozyme cleavage[136]	Junction, pseudoknot[165]	Sugar	*S. aureus*, *E. faecium*, *S. typhi*[166]
Guanidine-IV	1275	Transcription-OFF[174]	Pseudoknot[174]	Metabolic intermediate	*S. aureus*[174]
MoCo	1023	Transcription-OFF[175]	Junction[175]	Cofactor	*K. pneumoniae*, *S. typhi*[170]
$PreQ_1$-II	263	Translation-OFF[173]	Pseudoknot[165]	tRNA modification	*S. pneumoniae*[173,176]
SAM-III	220	Translation-OFF[160]	Junction[160]	Cofactor	*E. faecalis*[177,178]
SAM-VI	17	Translation-OFF[162]	Junction[165]	Cofactor	*B. angulatum*[161,162]

1.3.3. Riboswitches as Antibiotic Drug Targets

Riboswitches are found ubiquitously in bacteria due to their role as control elements for essential metabolic pathways. The fact that they often show high-affinity binding to specific ligands makes them a promising target for designing novel antibacterial drugs.[9] Moreover, riboswitches are found almost exclusively in bacteria. The TPP riboswitch is the only class known for being present in eukaryotes.[9] There is an increasing number of riboswitches with different ligands, but they do not occur simultaneously in all bacteria. Therefore, a selective selection of potential targets can be made based on existing riboswitch structures in relevant human pathogens. The most widely spread riboswitches in human pathogens are TPP- and FMN-riboswitches, found in almost all relevant pathogenic bacteria, including WHO's priority pathogens *P. aeruginosa*, *S. aureus*, and *E. faecium*.[166]

Crucial detailed structural and biochemical knowledge has advanced over time, with to date 399 crystal structures now available in the PDB[153], revealing diverse architectural characteristics. Riboswitches have evolved distinct strategies for achieving high specificity, such as shape complementarity and hydrogen bonding with ligands. These strategies allow for precise and efficient regulation of gene expression. Interactions with nucleobases, ribose, and phosphate groups, as well as aromatic rings, contribute to the diversity of ligand binding. Therefore, small molecules similar to the natural ligand often show binding to the target riboswitch, but with decreased affinity.[7,14,132,133] However, several challenges remain in harnessing riboswitches as drug targets. A critical requirement for antibiotic drugs targeting riboswitches is that they selectively inhibit bacterial riboswitches while not affecting human cellular functions. This necessitates identifying riboswitches that control the expression of genes crucial for bacterial survival but are absent of functionally distinct in human cells. In contrast, some human enzymes utilize the same metabolites as bacterial riboswitches for their cofactors.[14] Two prominent examples are the binding of *S*-adenosylmethionine (SAM) by both bacterial SAM riboswitches and human methyltransferases as well as thiamine pyrophosphate (TPP) riboswitches and TPP-binding proteins. Fortunately, the recognition of the metabolite differs between riboswitches and proteins. This allows for the rational design of riboswitch-targeting compounds that selectively affect their intended target.[154,179] Chemical dissimilarity to the metabolite can enhance this selectivity and reduce off-target effects.

When selecting the target riboswitch, it is also crucial to consider that the metabolite, whose synthesis, or transport is regulated, cannot be obtained from the host tissue cells. Therefore, the Flavin mononucleotide (FMN) riboswitch is a promising target in Gram-negative pathogens since the FMN riboswitch regulates the enzymes for FMN de novo synthesis by co-transcriptional repression of the *rib*-genes as well as the expression of riboflavin-transporter for external uptake.[14,180] By screening a small molecule library containing 57000 compounds, ribocil was identified binding to the FMN riboswitch.[41] This hit compound was further improved leading to ribocil-c (Figure 8 B) with a K_D-value in low nanomolar range against the *E. coli* FMN riboswitch.[42] This small molecule compound bears no chemical similarity to the natural ligand FMN (Figure 8 A), but shares critical interactions with the binding site.[41]

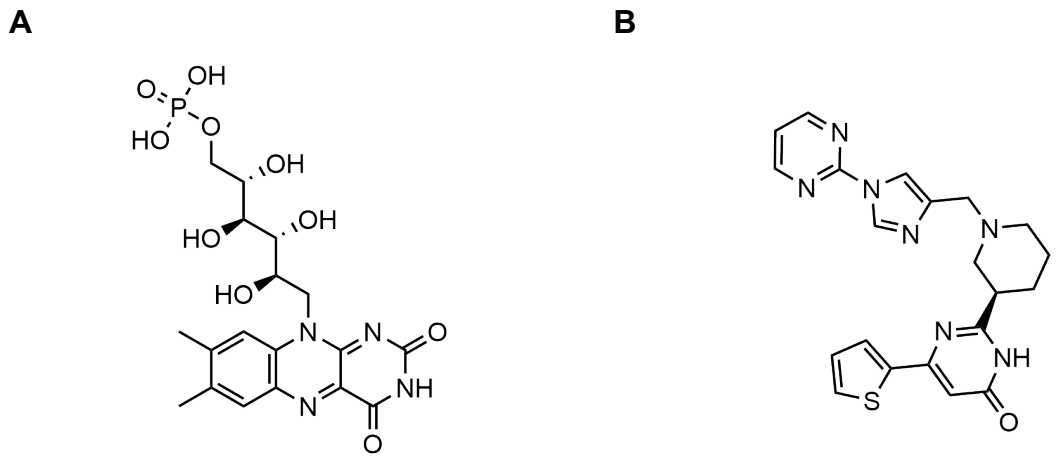


Figure 8: Structures of the natural FMN riboswitch ligand FMN (**A**) and its synthetic ligand ribocil-c (**B**). This figure was created with ChemDraw 20.

It is important to note that not all riboswitches are equally promising for drug discovery, as their presence in pathogenic bacteria and suitability as drug targets may vary. While some riboswitches have potential for broad-spectrum antibiotics due to their presence in multiple pathogenic bacteria, others may be better suited for narrow-spectrum antibiotics to minimize harm to the microbiome and reduce the development of resistance. Riboswitches must have a binding site that can accommodate drug-like ligands with high affinity, making riboswitches with binding pockets that are too small, too polar or too solvent-exposed difficult to target.[9]

1.4. PreQ$_1$ Riboswitches

1.4.1. Structural Features and Ligand Recognition

The discovery of the preQ$_1$ riboswitch structure in 2007 by Roth *et al*.[181] represents a significant milestone in understanding RNA-mediated gene regulation. This exploration culminated the discovery of three distinct structural classes of preQ$_1$ riboswitches (I–III) and their subtypes. Each class has unique structural features and regulation mechanisms. However, they all share a common feature: the use of a conserved cytidine forming Watson-Crick interactions with the guanine-like face of preQ$_1$ for tight binding.[157,176,181–184] These findings underscore the intricate sophistication of RNA-based regulatory systems and their pivotal role in cellular processes. Class I preQ$_1$ riboswitches are among the smallest known riboswitches with an aptamer region of a minimum of 34 nucleotides.[181–183] They are also the most widely distributed preQ$_1$ riboswitch class.[183] To date, three subtypes with different length and nucleotide composition in the loop region of the hairpin have been identified. Although the nucleotides in this loop vary, the binding of preQ$_1$ occurs in all three subtypes through a *cis*-Watson-Crick interaction between preQ$_1$ and a conserved cytidine at the end of the loop (Figure 9 A, B).[155,156,181–183] In contrast, class II preQ$_1$ riboswitches are distinguished by a larger and more complex structure than preQ$_1$-I riboswitches. The aptamer region of class II riboswitches is on average 58 nucleotides in length and forms up to five base-paired pseudoknot substructures. The recognition of preQ$_1$ is also driven by a conserved loop cytidine, similar to that observed in class I riboswitches. However, in class II, a *trans*-Watson-Crick interaction is used for ligand binding (Figure 9 C, D).[157,176] The most recent class III represents the largest aptamer region among preQ$_1$ riboswitches forming multiple substructures to create a sophisticated binding pocket. PreQ$_1$ binds tightly with low nanomolar affinity in a three-way helical junction forming *trans*-Watson-Crick interaction with a conserved cytidine (Figure 9 E, F).[183,184]

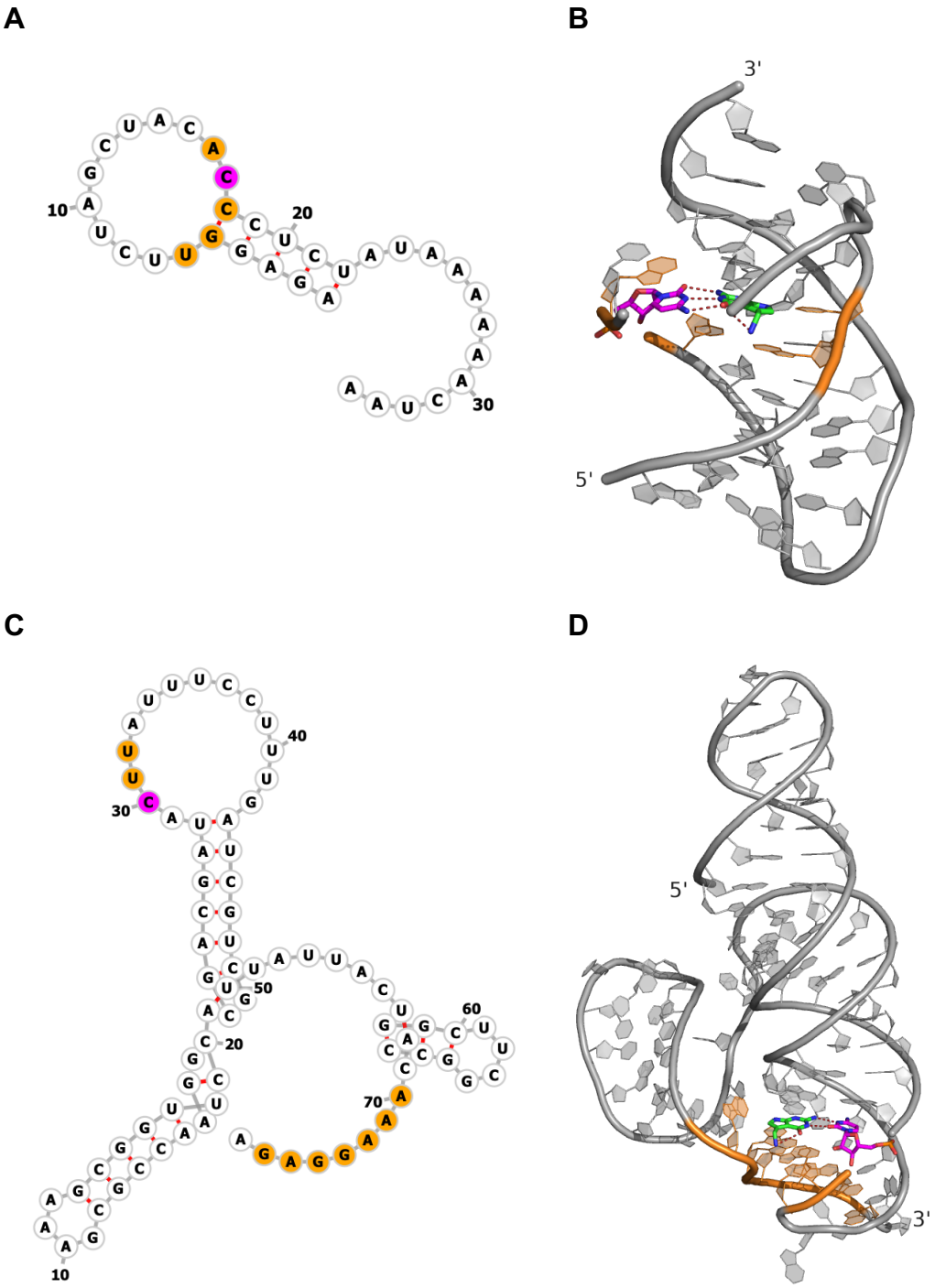


Figure 9: Structural features of the three preQ$_1$ riboswitch classes I–III. **A**, Secondary structure of a class I (type 2) preQ$_1$ riboswitch (PDB: 3FU2) with the conserved nucleotides colored in orange and the *cis*-acting cytosine colored in magenta. **B**, Tertiary structure of a class I (type 2) preQ$_1$ riboswitch (PDB: 3FU2). The conserved nucleotides are colored in orange (shown as cartoon), the *cis*-acting cytosine is colored in magenta (shown as sticks) and the ligand preQ$_1$ is colored in green (shown as sticks). **C**, Secondary structure of a class II preQ$_1$ riboswitch (PDB: 4JF2) with conserved nucleotides colored in orange and the *trans*-acting cytosine colored in magenta. **D**, Tertiary structure of a class II preQ$_1$ riboswitch (PDB: 4JF2). The conserved nucleotides are colored in orange (shown as cartoon), the *trans*-acting cytosine is colored in magenta (shown as sticks) and the ligand preQ$_1$ is colored in green (shown as sticks).

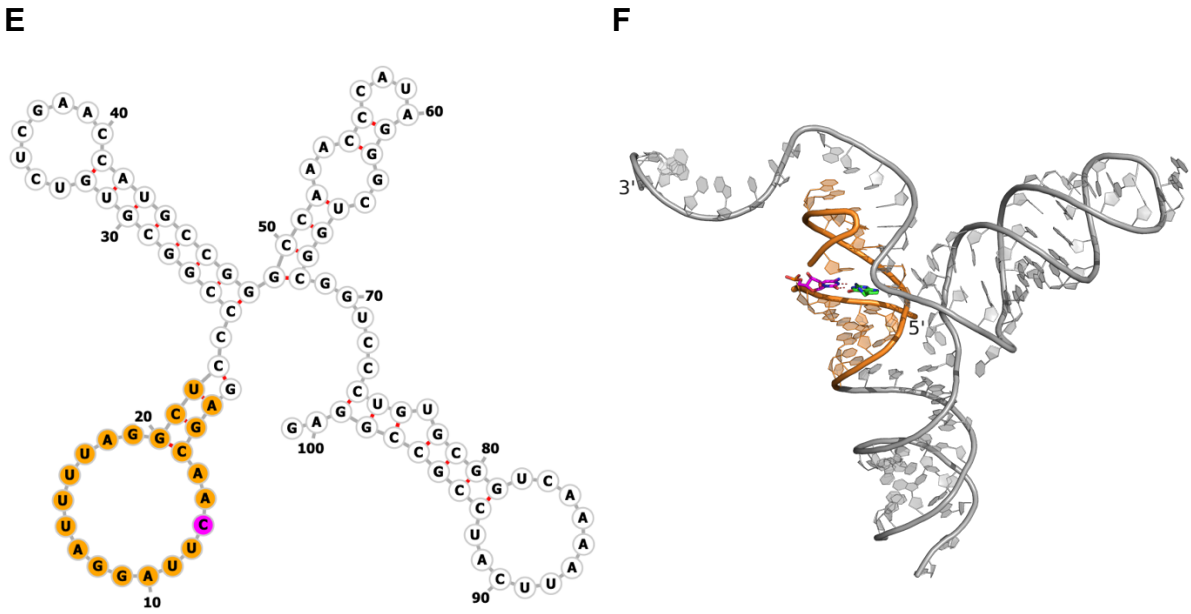


Figure 9: Structural features of the three preQ$_1$ riboswitch classes I–III. **E**, Secondary structure of a class III preQ$_1$ riboswitch (PDB: 4RZD) with conserved nucleotides colored in orange and the *trans*-acting cytosine colored in magenta. **F**, Tertiary structure of a class III preQ$_1$ riboswitch (PDB: 4RZD). The conserved nucleotides are colored in orange (shown as cartoon), the *trans*-acting cytosine is colored in magenta (shown as sticks) and the ligand preQ$_1$ is colored in green (shown as sticks). This figure was compiled according to references [183,184]. Panels **A**, **C**, and **E** was created with the forna web tool.[185] Panels **B**, **D**, and **F** was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

1.4.2. Regulation Mechanisms

As a member of riboswitches involved in the regulation of the purine metabolism, preQ$_1$ riboswitches control the biosynthesis and transport of preQ$_1$, a precursor of the hypermodified guanine derivate queuosine (Q) both at transcriptional and translational level.[156,157,176,181,183,184,186,187] PreQ$_1$ is only synthesized *de novo* from GTP in a multistep enzymatic pathway in procaryotes, while eucaryotes receive queuosine directly from scavenge pathways. PreQ$_1$ is incorporated into tRNAs at the wobble position of anticodons containing the GUN sequence. This posttranscriptional exchange of guanine with preQ$_1$ is catalyzed by tRNA-guanine transglycosylase (TGT). Once preQ$_1$ is attached to tRNAs, it serves as a precursor for Q biosynthesis, including the modification forming epoxyqueuosine (oQ), which is further transformed into Q (Figure 10). Queuosine-modified tRNAs play important roles in translation as Q enhances the accuracy and efficiency of protein synthesis by promoting correct codon-anticodon interactions and improving the decoding of degenerate codons. This

modification is particularly relevant for organisms, that lack certain aminoacyl-tRNA synthetases or have specialized translational needs. Queuosine-modified tRNAs also are implicated in various cellular processes, including the regulation of gene expression, stress response, and adaptation to environmental changes. The absence of queuosine biosynthesis or salvage enzymes can lead to deleterious phenotypes in organisms, highlighting the importance of $preQ_1$ metabolism in cellular physiology.[181,187–190]

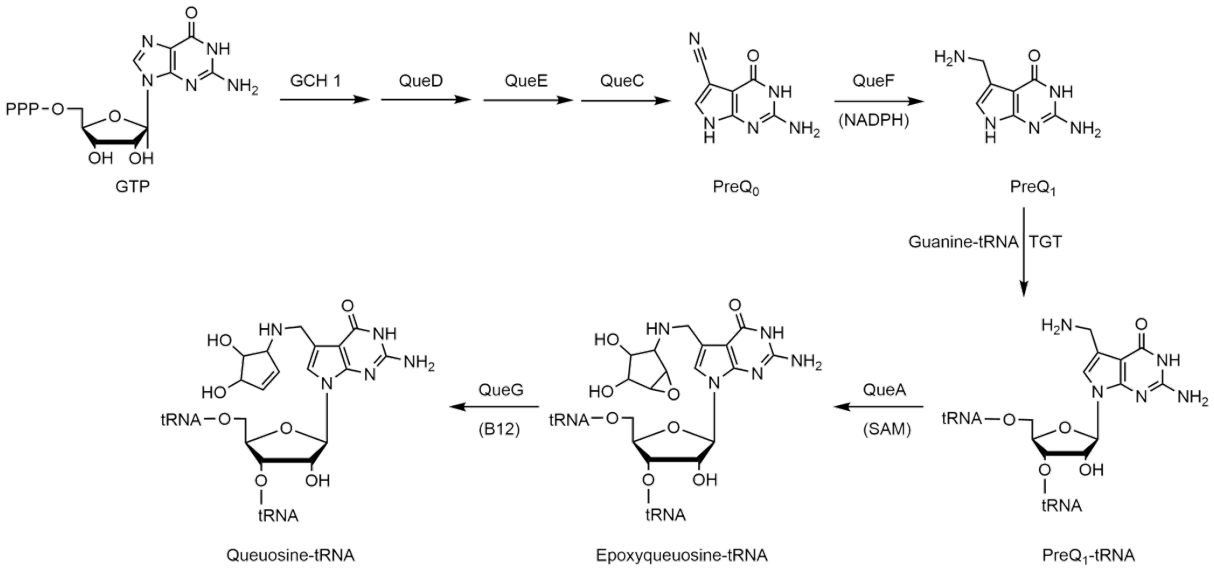


Figure 10: Synthesis of a queuosine-modified tRNA starting from GTP with the intermediates $preQ_0$ and $preQ_1$. This figure was created with ChemDraw 20.

The control mechanism of this pathway is created by a negative feedback loop that depends on intracellular $preQ_1$ concentrations.[155–157,176,181–184,186,187] In several members of class I, such as *Bacillus subtilis* (*Bs*), this occurs at the level of transcription. Binding of $preQ_1$ to the aptamer rearranges the downstream expression platform to form a terminator stem that aborts the transcription. Consequently, the region of the mRNA that encodes for enzymes required for $preQ_1$ synthesis, is not made.[63,155,156,181–183,187,191] However, there are also members of class I that have also been reported (e.g., *Thermoanaerobacter tengcongensis*, *Tte*[63,186,191] and *Escherichia coli*, *E. coli*[192]) to affect translation initiation in a manner similar to members of class II and III, which exclusively switch off translation. In this case, the ribosome-binding site (RBS) or Shine Dalgarno sequence (SDS) at the end of the expression platform is exposed for translation in the unbound state. Binding of $preQ_1$ results in conformational changes that modulate the interactions of the nucleobases, leading to the sequestration of the SDS. Consequently, the initiation of translation of proteins

required for $preQ_1$ synthesis is inhibited.[157,176,184,187] In *Tte* $preQ_1$-I riboswitches, the L2 loop and S2 stem of the pseudoknot exhibit reduced base pairing and less same strand stacking interactions in the ligand-free state. This leads to the exposure of the SDS and promotion of protein translation. Upon binding of $preQ_1$, the nucleobases of L2 and S2 are allowed for base pairing and stacking interactions with L2 in a spine-like conformation. This rearrangement of the expression platform results in the formation of an anti-SDS strand, which prevents translation initiation.[186,191]

Despite their different regulation mechanisms, the *Bs* and *Tte* $preQ_1$ riboswitch ligand recognition pattern is very similar: Both riboswitches form distinct binding pockets from bases of the same loop and stem regions to bind $preQ_1$ (Figure 11).[187]

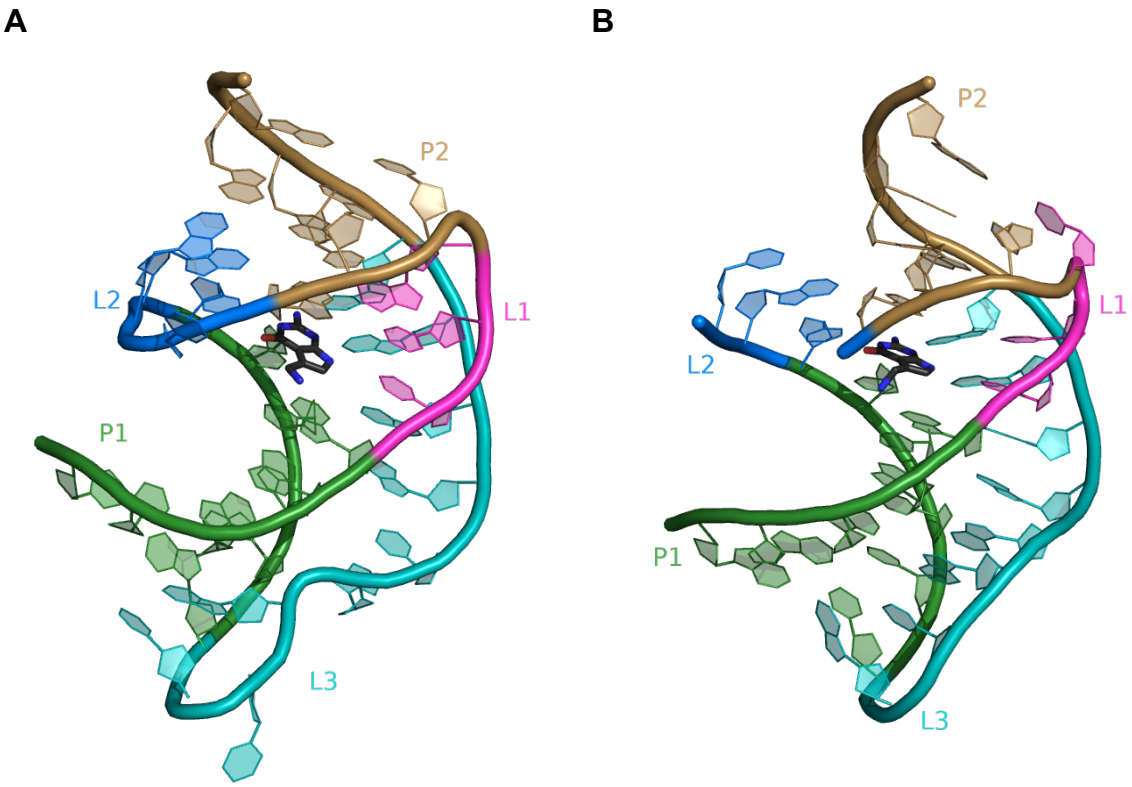


Figure 11: Tertiary structure, binding pocket composition and ligand recognition of the *Tte* (PDB: 3Q50) and *Bs* (PDB: 3FU2) $preQ_1$ riboswitch aptamers. **A**, Tertiary structure of the *Tte* $preQ_1$ riboswitch aptamer (shown as cartoon) in complex with $preQ_1$ (shown as sticks). **B**, Tertiary structure of the *Bs* $preQ_1$ riboswitch aptamer (shown as cartoon) in complex with $preQ_1$ (shown as sticks). This figure was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

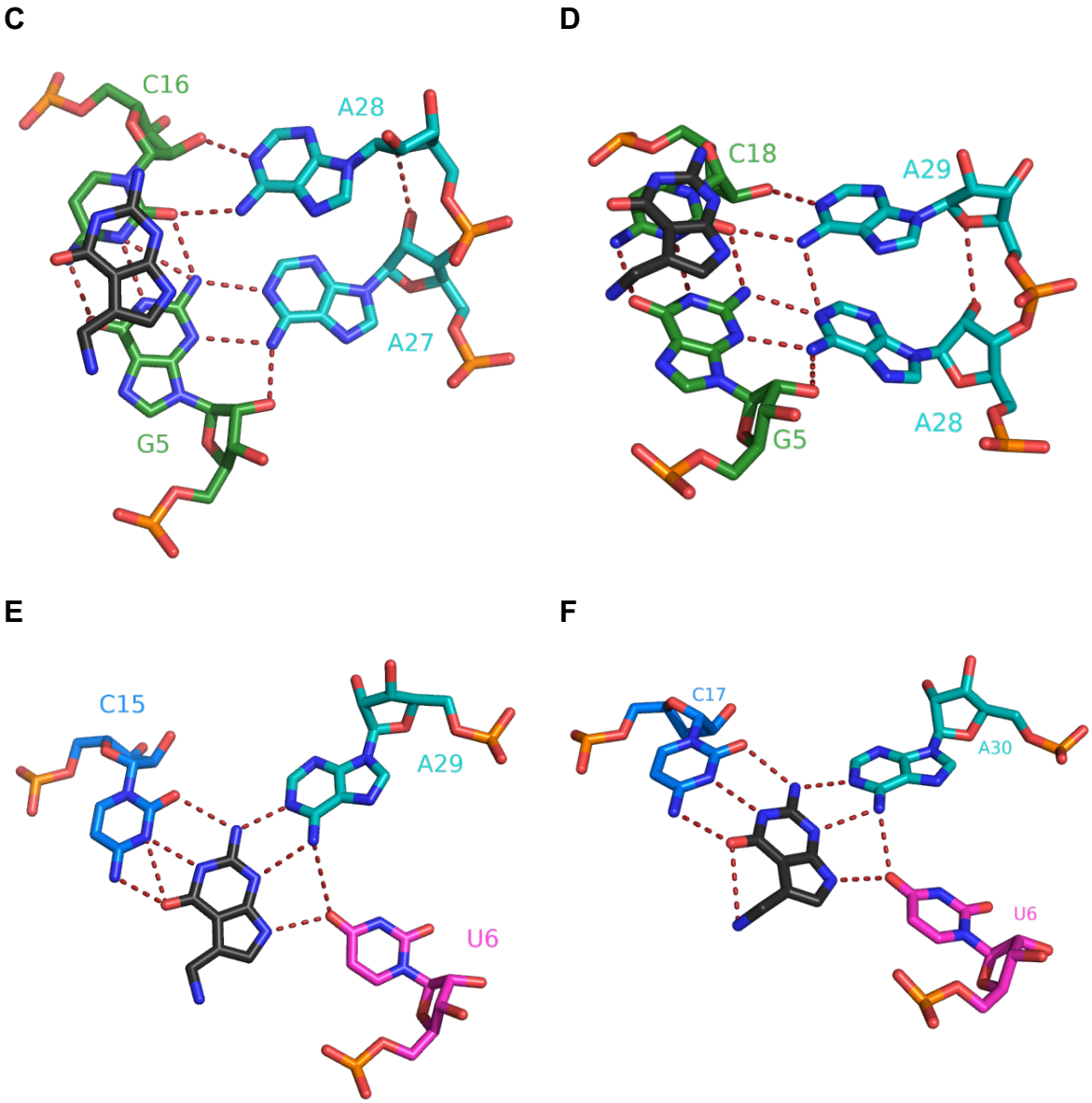


Figure 11: Tertiary structure, binding pocket composition and ligand recognition of the *Tte* (PDB: 3Q50) and *Bs* (PDB: 3FU2) preQ$_1$ riboswitch aptamers. **C**, *Tte* binding pocket "floor": P1–L3 A28–C16–G5–A27 base quartet. **D**, *Bs* binding pocket "floor": P1–L3 A29–C18–G5–A28 base quartet. **E**, *Tte* ligand recognition: preQ$_1$ forms hydrogen bonding interactions with U6 of L1, C15 of L2, and A29 of L3. **F**, *Bs* ligand recognition: preQ$_1$ forms hydrogen bonding interactions with U6 of L1, C17 of L2, and A30 of L3. This figure was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

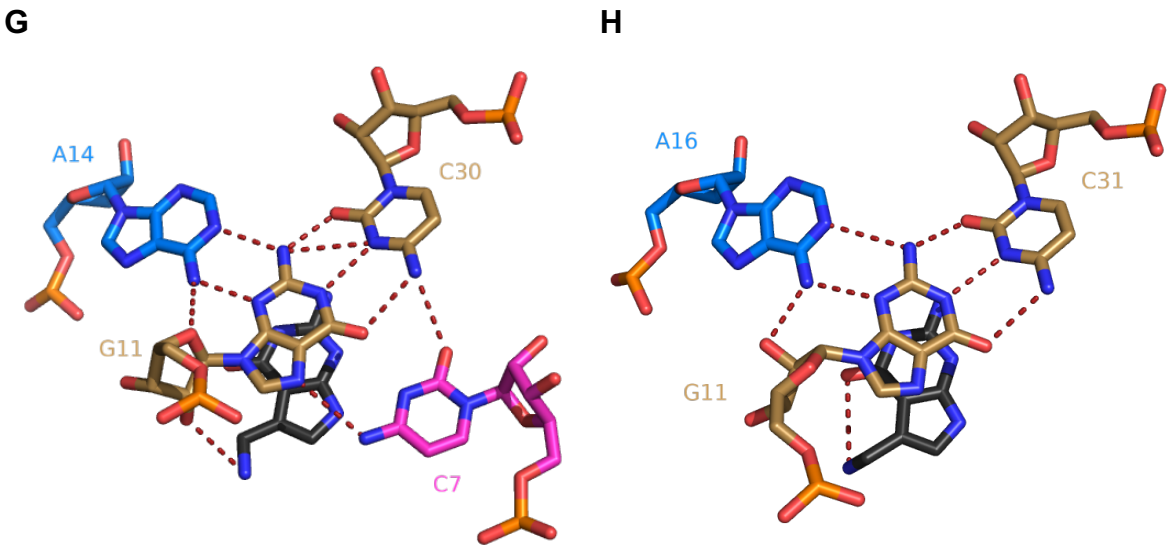


Figure 11: Tertiary structure, binding pocket composition and ligand recognition of the *Tte* (PDB: 3Q50) and *Bs* (PDB: 3FU2) preQ$_1$ riboswitch aptamers. **G**, *Tte* binding pocket "ceiling": L2–P2–L1 A14–G11–C30–C7 base quartet. **H**, *Bs* binding pocket "ceiling": L2–P2 A15–G11–C31 base triple. This figure was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

1.4.3. PreQ$_1$ Riboswitches as Potential Drug Targets

After their discovery in 2007[181], several classes of preQ$_1$ riboswitches were unearthed across various bacterial species, showcasing their widespread occurrence and functional diversity.[157,176,183,184,191] PreQ$_1$ riboswitches control pathways associated with the synthesis of the hypermodified guanine base queuosine, which is known for its important role in bacterial survival strategies.[155,181,187,188,192] PreQ$_1$ riboswitch classes are widely spread among different bacteria, including relevant human pathogens carrying multidrug resistance such as *Enterococcus faecalis*, *Haemophilus influenzae*, *Neisseria gonorrhoeae*[9] (class I), and *Streptococcus pneumoniae*[176,187] (class II).[183] The preQ$_1$ pathway represents an attractive point of attack since preQ$_1$ is exclusively synthesized in bacteria[187] to further generate queuosine, and this pathway is controlled by riboswitches, which are not found in eukaryotic cells resp. humans.[8] Therefore, preQ$_1$ riboswitches fulfill important criteria of potential drug targets: The metabolite, whose pathway is controlled, cannot be obtained from the environment, human queuosine pathways are not affected since they work in a different way, and molecular recognition patterns are well studied by now allowing the design of binding small molecules.[9]

1.4.4. PreQ$_1$ Riboswitch as Model System

The class I preQ$_1$ riboswitches represent an excellent model system to study RNA-small molecule interactions due to its predictable highly conserved aptamer structure which is responsible for ligand binding.[63,155,156,181–183,187] These structures have been extensively studied by X-ray crystallography and NMR. To date, there is a huge number of PDB entries that provide deep insights in the unique structural features of several preQ$_1$ riboswitches.[155,156,182,192] Another important point making this riboswitch a good model system, is its accessibility for experimental manipulation such as side directed mutagenesis and chemical modifications.[63,156,182] The preQ$_1$ riboswitch shows also a good assayability and reproducibility, proven by the employment of a large set of diverse biophysical methods.[63,181,193–195] These methods include equilibrium dialysis, fluorescence intensity assays, in-line probing, ITC and, MST.

The class I *Bs* preQ$_1$ riboswitch aptamer is one of the most studied preQ$_1$ riboswitch structures with available crystal structures in the unbound state and bound state with different ligands allowing insights in the folding and ligand binding of this riboswitch.[63,155,156,181–183,187,191]

In this context, the aim of this dissertation is the elucidation of challenges in targeting RNA with small molecules at the level of fundamental research using the well-studied *bs*-preQ$_1$ riboswitch as a model system. The following publication is part of the results obtained in this preQ$_1$ riboswitch project:

Kallert, E.; Fischer, T. R.; Schneider, S.; Grimm, M.; Helm, M.; Kersten, C. Protein-Based Virtual Screening Tools Applied for RNA–Ligand Docking Identify New Binders of the PreQ$_1$-Riboswitch. *J. Chem. Inf. Model.* **2022**, 62(17), 4134–4148.

1.5. SAM Riboswitch Family

1.5.1. Biological Functions and Relevance

S-adenosylmethionine (SAM) is the key metabolite in sulfur metabolism and an essential compound involved in numerous biological processes.[196,197] SAM is synthesized from adenosine triphosphate (ATP) and methionine catalyzed by SAM synthetase.[198] Primarily, SAM is used as a cofactor in biochemical methylation reactions, such as tRNA methylation, which are catalyzed by methyltransferases (Figure 12). The transfer of the methyl group from SAM to a substrate results in the formation of *S*-adenosylhomocysteine (SAH). Consequently, the positively charged sulfonium is turned into a neutral thioether.[199] Due to the importance of SAM for cellular processes, the associated riboswitch classes are the largest group of riboswitches sensing a single ligand in bacteria. The SAM riboswitch family is involved in the regulation of the SAM biosynthesis and consequently in the maintenance of its cellular levels. Hence, the discrimination against SAH assumes pivotal importance in bacterial viability, as failure to do so may lead to the downregulation of SAM synthetase expression in cells. This is attributable to the functional mechanism of most SAM riboswitches acting as translational or transcriptional OFF switches.[200,201] To date, seven types of SAM-sensing riboswitches are known[161,202–208], which can be divided into four classes based on their binding pocket similarity[160,161] (Table 4). Those riboswitches exhibit a pronounced discrimination between SAM and SAH.[160,209] Based on their structural similarity, the positive charge of SAM is favored for ligand binding.[160,210,211] Furthermore, a so called charge-assisted chalcogen bond, in which oxygen atoms of RNA bases interact with the σ-hole of the SAM sulfur atom is also discussed in the literature for the SAM-I, SAM-I/IV, SAM-II, SAM-III and SAM-V riboswitches.[212]

Additionally, one SAH-sensing riboswitch and one riboswitch that binds SAM and SAH with comparative affinity, are reported (Table 4).

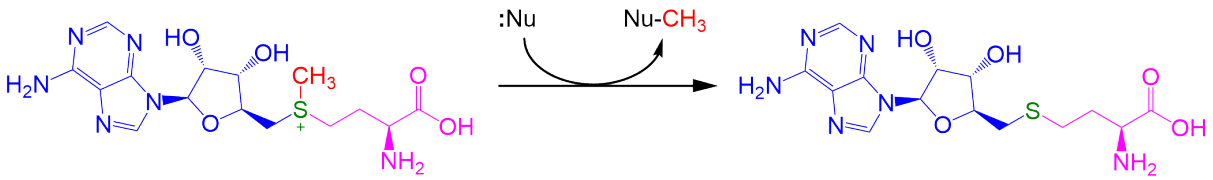


Figure 12: SAM-dependent enzymatic methylation reaction. Blue: adenosyl-moiety, green: positively charged sulfonium moiety (SAM) and neutral thioether moiety (SAH), magenta: methionine-tail, red: transferred methyl group. This figure was created with Chemdraw 20.

Table 4: Organization of SAM/SAH-binding riboswitch classes and types. Members of each class appear to form nearly the same binding pocket for ligand recognition. This table was compiled according to reference [161].

Class	Types	Favored Ligand
SAM-I	SAM-I, SAM-IV, SAM-I/IV	SAM
SAM-II	SAM-II, SAM-V	SAM
SAM-III	SAM-III	SAM
SAH	SAH	SAH
SAM-SAH	SAM-SAH	SAM, SAH
SAM-VI	SAM-VI	SAM

1.5.2. SAM-I Riboswitches – Structure, Ligand Recognition and Distribution

The SAM-I riboswitch was first described as “S-box regulon” in 1998 by Grundy and Henkin.[213] Five years later, Winkler *et al.* identified this RNA element as a SAM-responsive riboswitch.[202] Montange and Batey resolved the tertiary structure of the *Tte* SAM-I riboswitch by X-ray crystallography in 2006.[179] The SAM-I riboswitch belongs to the class of the same name, which also includes the SAM-IV and SAM-I/IV riboswitches. The three members of this class exhibit a high degree of structural similarity and SAM-binding interactions.[160,161] However, the SAM-I riboswitches are the most prevalent members of the SAM-I class and are typically found in Gram-positive bacteria, where they function as control elements for transcription termination.[160,202]
The tertiary structure of the *Tte* SAM-I riboswitch mainly consists of four stems (P1–P4), which form two coaxially stacked helices (P1/4 and P2/3). These are connected to one central four-way junctional internal loop L1 (Figure 13 A, B). An additional pseudoknot-stem is formed by long-range interactions of the kink-turn of P2 and C65–G68 within J3/4. This pseudoknot between P2 and P4 stabilizes the tertiary structure. Furthermore, the canonical (A85–U64) and non-canonical (A85·A24) also contribute

to the structure of the SAM-I riboswitch. The SAM-binding pocket is located between the minor grooves of P1 and P3. The pocket is capped by the junctional region J1/2. The parallel-folded methionine- and adenine-moieties are anchored in the binding pocket by distinct interactions (Figure 13 C, D). The methionine-tail interacts with the triple base pair (G58–C44)·G11 from P3 and J1/2 (Figure 13 D). The adenine forms hydrogen bonding interactions with A45 and U57, as well as stacking interactions with C47–G56 from P3 (Figure 13 C). Additionally, the ribose moiety of SAM forms a hydrogen bond with the ribose of C47 by adopting a C3'-*endo* sugar pucker conformation (Figure 13 C). The positively charged sulfonium group is recognized by U7 and U88 of P1 (Figure 13 E). This electrostatic interaction is the basis for discrimination between SAM and SAH.[179,211]

At low intracellular SAM levels, the formation of a transcription antiterminator stem-loop is favored. In contrast, at high SAM levels, an alternative terminator structure is formed. In this structure, one of the antiterminator stems is not part of P1, which leads to transcription termination.[202]

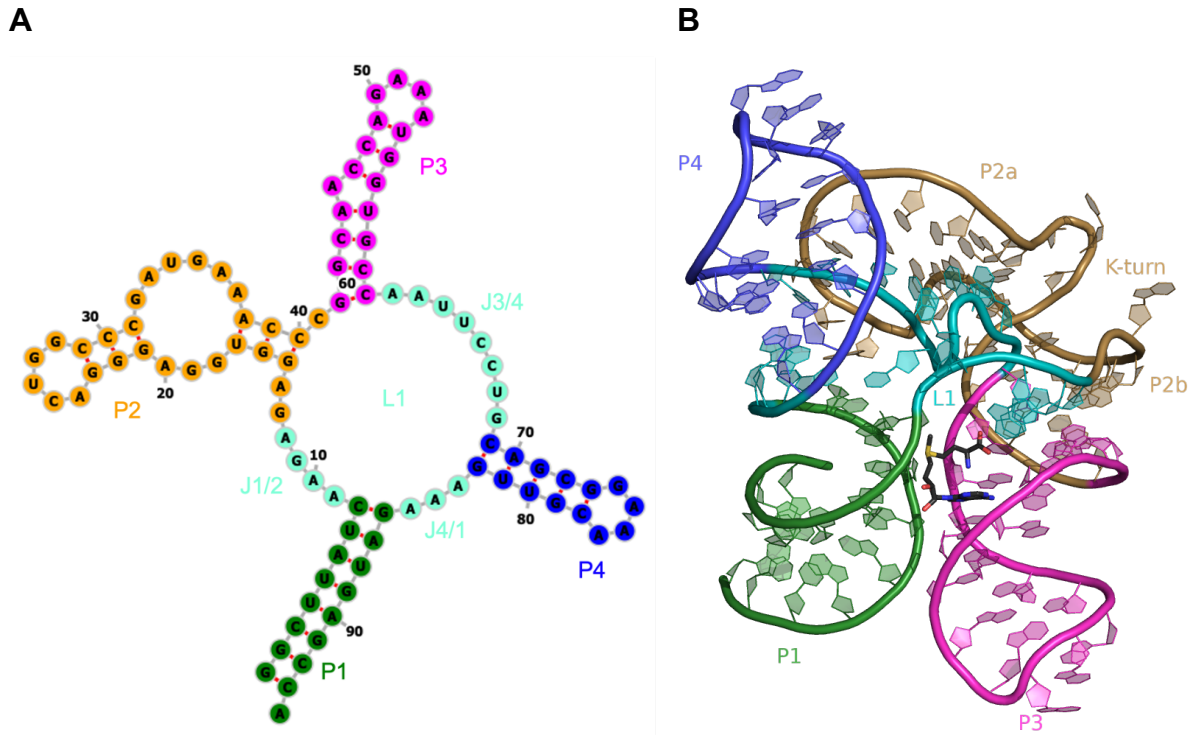


Figure 13: Structure and ligand recognition of the *Tte* SAM-I riboswitch (PDB: 2GIS). **A**, Sequence and secondary structure of SAM-I riboswitch. **B**, The tertiary structure of SAM-I riboswitch (shown as cartoon) complexed with SAM (shown as sticks). Panel **A** was created with the forna web tool[185], panel **B** was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

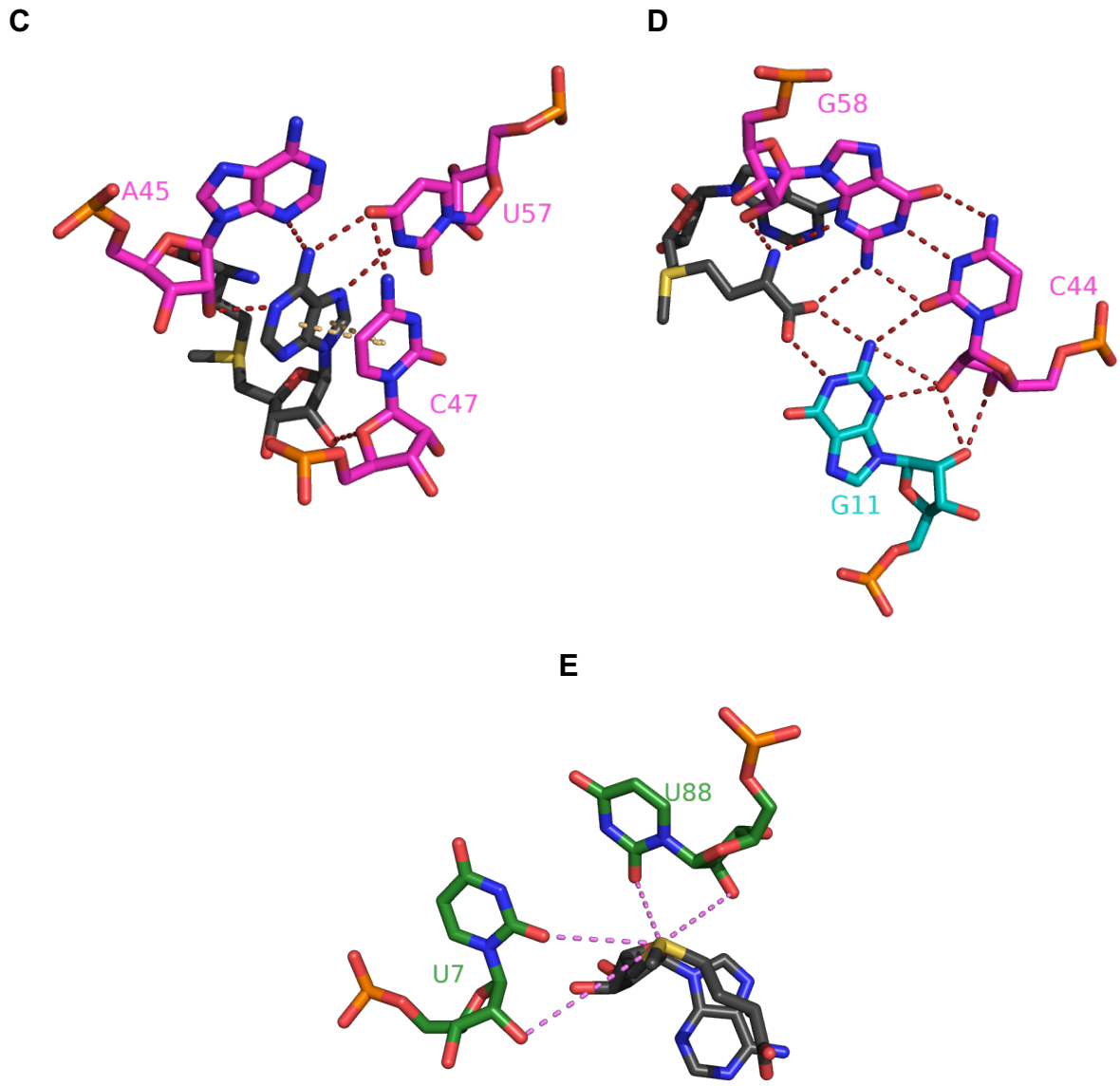


Figure 13: Structure and ligand recognition of the *Tte* SAM-I riboswitch (PDB: 2GIS). **C**, The adenine moiety of SAM forms hydrogen bond interactions with A45 and U57 from P2 and stacking interactions with C47. The ribose moiety forms one hydrogen bond with the ribose of C47. **D**, The methionine moiety of SAM is recognized by the base triple (G58–C44)·G11. **E**, Electrostatic interactions between the positively charged sulfonium moiety of SAM and O2 of U7 and U88. Panels **C**–**E** were created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

1.5.3. SAM-I Riboswitch as Potential Drug Target

The SAM-I riboswitch represents a promising potential drug target for the treatment of bacterial infections. The *Tte* SAM-I riboswitch, which has been extensively studied[179,202,211], exhibits a high degree of sequence similarity found in the *S. aureus* genome.[214] This bacterium belongs to the WHO's priority group of human pathogens.[126] Given the sequence similarity, it can be assumed that the tertiary structures of the two riboswitches are also similar, as are their ligand recognition

capabilities. The deep structural knowledge may facilitate the design of small molecules that bind with high affinity and specificity to the SAM-I riboswitch. In this regard, this thesis includes a small SAR study to elucidate the interaction between the SAM-I riboswitch and various SAM analogue compounds as well as several amino acids. These preliminary data provide a starting point for further investigations about the binding behavior of the SAM-I riboswitch. This includes a virtual screening of small molecules distinct from the SAM/SAH structural motif based on the results of the SAR study. Subsequently, identified hits should be synthesized and derivatized in order to obtain potent small molecule binders of the SAM-I riboswitch.

1.5.4. SAM-VI Riboswitch – Structure, Ligand Recognition and Distribution

The SAM-VI riboswitch represents the sixth SAM riboswitch class. It was found in 2018 by comparative sequence analysis of the Bifido-*metK* gene. This gene encodes for the SAM synthetase which catalyzes SAM synthesis from methionine and ATP and is predominantly found in *Bifidobacteriae*.[161] The tertiary structure of the *Ba* SAM-VI riboswitch consists of three stems (P1, P2 and P3) which are connected by a central three-way junction (Figure 14 A). P1 and P2 are arranged coaxial and are fixed by stacking interactions between one junctional Watson-Crick base pair (G7– G33) and the adenine moiety of bound SAM which is recognized by U8 from stem P2 (Figure 14 D). P3 is arranged almost orthogonally to the coaxial P1–P2 unit and is connected to P1 via the SDS. The positively charged sulfonium moiety is recognized by oxygens of U6 and U8 (Figure 14 B).[161,162] Due to the spatial proximity of the junction to a chain of the P2 double helix, H-bonds are formed between junctional bases A37–U39 and G7–C10 of P2, including the ribose 2'-OH groups (Figure 14 C). This distinct motif is known as the ribose zipper.[162,215]

The SAM binding site is located at the junction between P1 and P2 in the large groove that forms the coaxial unit from these two stems. This binding pocket consists of the Watson-Crick base pair G9–C32 from P2, the non-canonical *trans* Watson-Crick base pair G7–G33, four continuously stacked bases (A34, A36, A37 and G38), and the junctional U8 interacting with the SAM adenine moiety. Additionally, the adenine forms π–π stacking with G9–C32 and G33 of the riboswitch, while the methionine is overhanging the binding site.[162]

At low intracellular SAM levels, the SAM-VI riboswitch establishes a co-transcriptional P0 stem with the 5'-end of the mRNA not originating from the riboswitch. Following the conformational selection, SAM does not bind to P0. SAM binding only occurs when the transcript reaches a length at which the P1 stem can be fully formed. Subsequently, by refolding the stem into P1, the Shine-Dalgarno sequence is blocked and thus translation is prevented.[162]

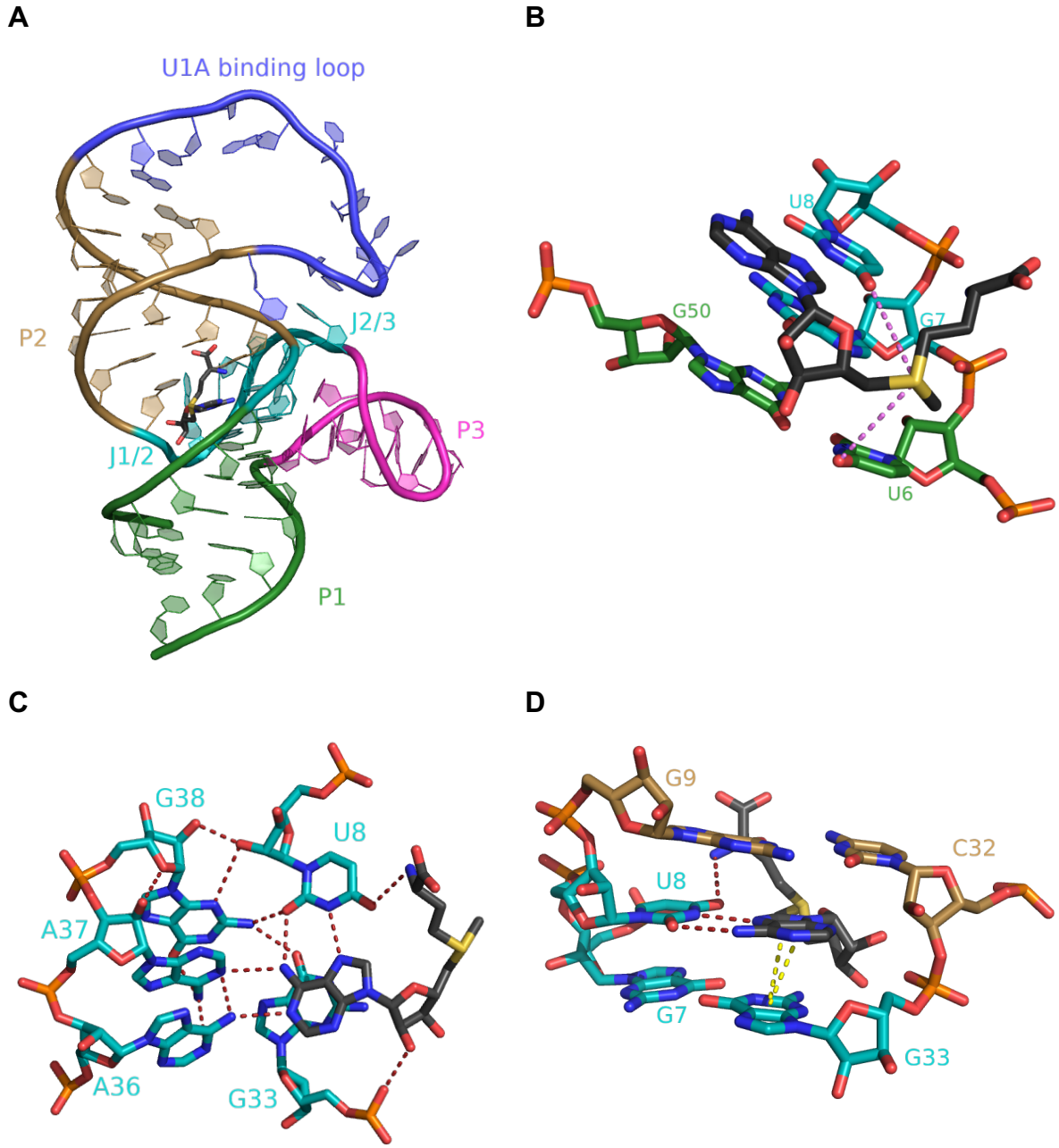


Figure 14: Structure and molecular recognition of the SAM-VI riboswitch (PDB: 6LAS). **A**, Tertiary structure of SAM-VI riboswitch (shown as cartoon) in complex with SAM (shown as sticks). **B**, Electrostatic interaction between the positively charged sulfonium moiety of SAM and O4 of U6 and U8 in the binding pocket. **C**, Hydrogen bonding interactions between SAM and U8, G33, A36 and A37 with G38 interacting with U8 in the binding pocket. **D**, The adenine moiety of SAM pairs with U8 and becomes stacked between G9–C32 and G7·G33 in the binding pocket. This figure was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

1.5.5. SAM-VI Riboswitches as Model System for SAM-III Riboswitches

When the SAM-VI riboswitch aptamer structure was discovered, it seemed to show some similarity to the previously reported SAM-III riboswitch aptamer.[152,177] Both secondary structure models consist of three-way junctions (P1, P2 and P3) with a short P3 stem which is linked to P1 by the SDS (Figure 15). In addition, five nucleotides contacting SAM in the binding pocket are identical and in equivalent positions, assuming that SAM binding operates under similar interactions in both riboswitch models and discriminates strongly between SAM and SAH (Figure 15 A, B). However, the binding pockets of the two riboswitches in the center of the three-stem junction differ in the number of nucleotides. While SAM-VI has six highly conserved nucleotides in J2–3, SAM-III has only three. Additionally, the SAM-VI secondary structure does not exhibit a bulge in P2, which is known to be essential for the formation of the binding site in SAM-III. This indicates that the architecture of the binding site, may be distinct, despite the similarities in the interactions utilized for SAM binding. In comparison to SAM-III, SAM-VI incorporates the start codon AUG directly into the terminal region of P1 and comprises a greater number of conserved nucleotides.[161,162]

In the SAM III riboswitch binding pocket, the adenine moiety of SAM forms hydrogen bonding interactions with the G26 sugar edge and A73 (Figure 15 E). In the SAM VI riboswitch binding pocket, the equivalent base is G7 which forms hydrogen bonding with G33 and stacking with the adenine moiety. A36 interacts with the N1 of adenine as an equivalent of A73 in the SAM III pocket (Figure 15 F). While the ribose moiety of SAM is recognized by a single hydrogen bond with G33 phosphate in the SAM VI pocket, the SAM III riboswitch forms two hydrogen bonds between the ribose and the phosphate and *N*7 of G89 (Figure 15 G, H). The sulfonium moiety forms electrostatic interactions with both riboswitches: In the SAM III binding pocket, G71 and U72 are involved, U6 and U8 interact in the SAM VI riboswitch (Figure 15 I, J).[162]

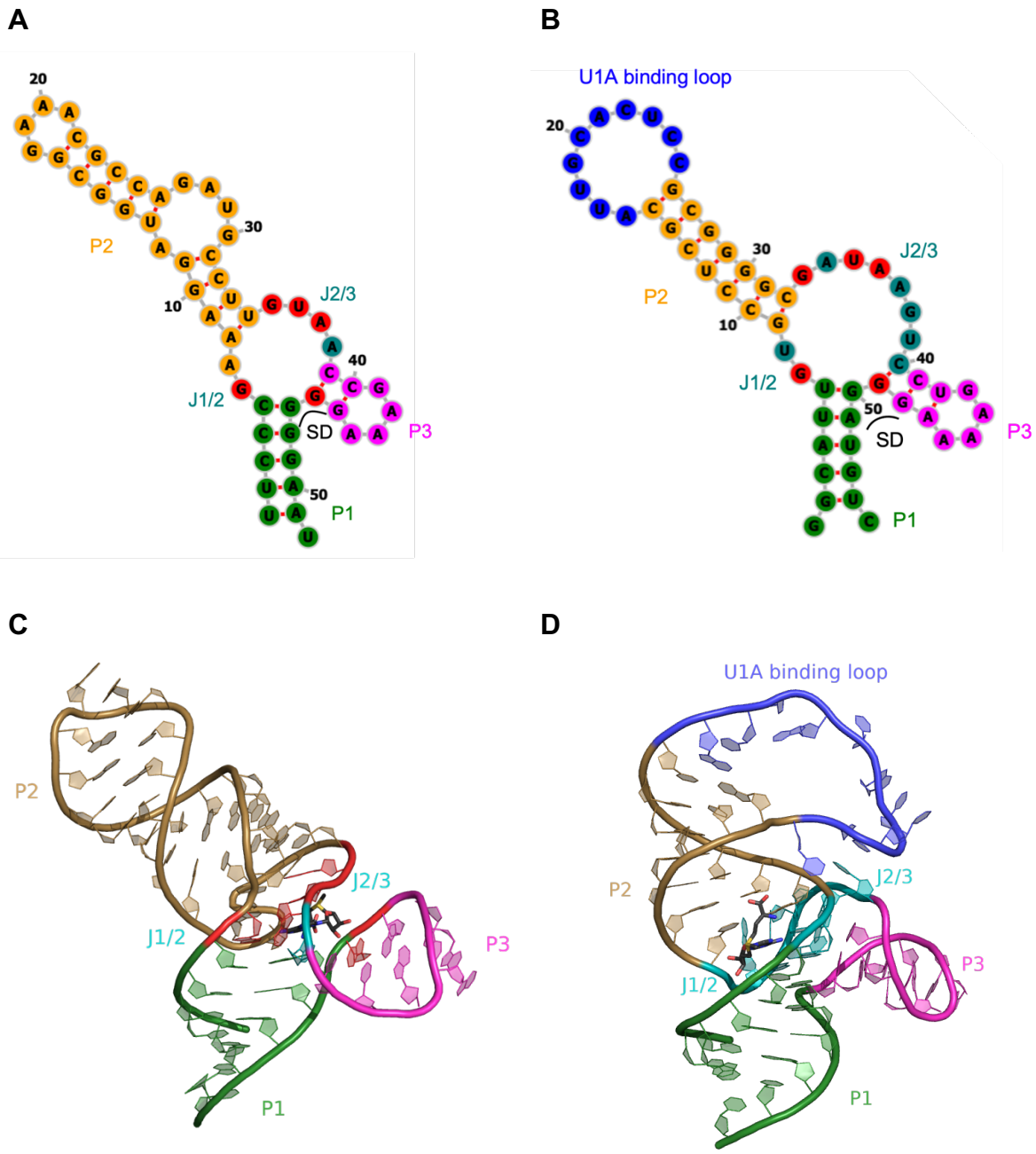


Figure 15: Sequence, tertiary structure and binding pocket comparison of the SAM-III (PDB:3E5C) and the SAM-VI (PDB: 6LAS) riboswitch. Both riboswitches exhibit a high sequence similarity in the binding pocket, with five conserved nucleotides each (red) which are identical in both riboswitches and form interactions with SAM. **A**, Sequence of the SAM-III riboswitch; green: stem P1, orange: stem P2, magenta: stem P3, teal: junctional region, red: conserved nucleotides forming interactions with SAM. **B**, Sequence of the SAM-VI riboswitch; green: stem P1, orange: stem P2, magenta: stem P3, blue: U1A binding loop, teal: junctional region, red: conserved nucleotides forming interactions with SAM. **C**, Tertiary structure of the SAM III riboswitch (shown as cartoon) complexed with SAM (shown as sticks). **D**, Tertiary structure of the SAM VI riboswitch (shown as cartoon) complexed with SAM (shown as sticks). Panels **A** and **B** were created with the forna web tool.[185] Panels **C** and **D** were created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

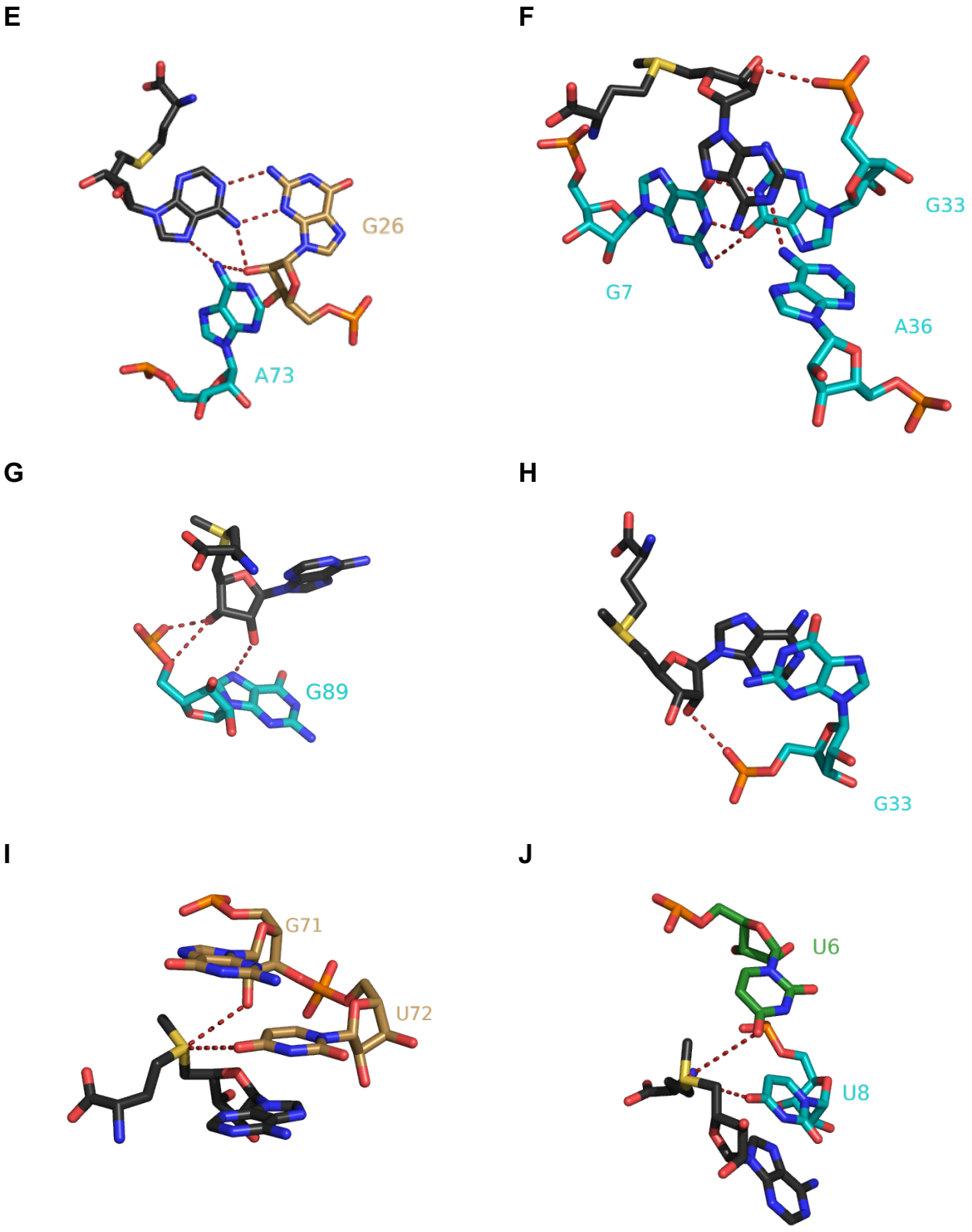


Figure 15: Sequence, tertiary structure and binding pocket comparison of the SAM-III (PDB:3E5C) and the SAM-VI (PDB: 6LAS) riboswitch. Both riboswitches exhibit a high sequence similarity in the binding pocket, with five conserved nucleotides each (red) which are identical in both riboswitches and form interactions with SAM. **E**, Recognition of the adenine moiety of SAM by G26 and A73 in the SAM III binding pocket. **F**, Recognition of the adenine moiety of SAM by G7, G33 and A36 in the SAM VI binding pocket. **G**, Hydrogen bonding between G89 of SAM III and the ribose moiety of SAM. **H**, Hydrogen bonding between G33 of SAM VI and the ribose moiety of SAM VI. **I**, Electrostatic interactions between G71 and U72 oxygens of SAM III and the charged sulfonium moiety of SAM. **J**, Electrostatic interactions between U6 and U8 oxygens of SAM VI and the charged sulfonium moiety of SAM. Panels **E**–**J** were created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

Due to the similarities in their structural and functional characteristics, the SAM-VI riboswitch can serve as a model system with enhanced assayability for the SAM-III riboswitch in fundamental research. This may facilitate a more profound comprehension of the folding dynamics and ligand recognition processes pertinent to SAM-III, which could inform the further development of novel antibiotic active substances. This riboswitch is present in human pathogens such as *E. faecalis*[177,178], and the binding pocket is highly conserved, suggesting its potential as a target for the development of new antibiotics.

1.6. Antiviral Drug Discovery

1.6.1. Antiviral Drugs and Their Limitation

Antiviral drug discovery has become increasingly important in recent decades as viruses, notably RNA viruses, are predominant among the WHO's current list of ten global health threats. The emergence of novel viral pathogens, such as human immunodeficiency virus (HIV), severe acute respiratory syndrome coronavirus (SARS-CoV-1), Middle East respiratory syndrome coronavirus (MERS-CoV), dengue virus (DENV) and more recently, SARS-CoV-2 causing the 2019 pandemic, has highlighted the urgent need for effective antiviral therapies. These viruses can cause devastating epidemics and pandemics, underscoring the critical importance of developing drugs combating them. Particularly, globalization and international travel facilitate the rapid spread of viral infections.[216] Although significant progress has been made in exploring potential targets and discovering antiviral agents, many viral diseases with high mortality rates still exist.[216,217]

The two strategies combating viral diseases primarily involve prevention through vaccines and treatment using antiviral drugs, typically small molecules, and antibodies. Despite more than 200 viruses known to infect humans, with 80% naturally persistent in animals, as of 2019, there were only 87 approved vaccines targeting less than 20 human viruses, mainly influenza. The economic risk associated with developing vaccines for viral diseases is high due to the lengthy development process and relatively low success rates, as seen with vaccines like Dengvaxia®.[218,219] The coronavirus disease 2019 (COVID-19) pandemic has spurred a global effort to develop a vaccine, although challenges remain due to the rapid emergence of SARS-CoV-2

and the logistical hurdles of mass production. In absence of a vaccine, effective treatments are crucial during viral outbreaks. Antibody and antiviral drug therapies are currently explored, although challenges such as resistance and antibody-dependent enhancement (ADE) need to be addressed.[216]

There are two most common approaches in antiviral drug discovery: targeting viral proteins ensuring viral survival as direct inhibition and targeting host factors which are essential for incidence of infection leading to an indirect inhibition mechanism. Most of these drugs addresses viral targets, which can be divided into five classes, including viral polymerases, proteases, integrases, structural and accessory proteins, summarized in Table 5.[217,220] In 2021, there are 118 antiviral therapies approved by the FDA, with 72% being monotherapies and the remaining being combination therapies of 2–4 drugs. More than 80% of these therapy options are directed against viral components.[24]

Table 5: Overview of human viruses which currently can be treated with antiviral drugs. 1–2 examples of approved small molecule drugs for each target are listed. The table was compiled according to references [24,217,220,221].

Human Virus	Viral Target	Approved Drug(s)
Hepatitis B virus (HBV)	• DNA polymerase	• Entecavir, tenofovir
Hepatitis C virus (HCV)	• NS3/4A protease • NS5A phosphoprotein • NS5B polymerase	• Grazoprevir, simeprevir • Daclatasvir, elbasvir • Dasabuvir, sofosbuvir
Herpes simplex virus (HSV)	• DNA polymerase	• Brivudine, acyclovir
Human cytomegalovirus (HCMV)	• DNA polymerase • Terminase	• Foscarnet, ganiciclovir • Letermovir
Human immunodeficiency virus (HIV)	• Protease • Reverse trancriptase • Integrase	• Fosamprenavir, ritonavir • NRTIs: Emtricitabine, tenofovir • NNRTIs: Efavirenz, nevirapine • Elvitegravir, raltegravir

Table 5: Overview of human viruses which currently can be treated with antiviral drugs. 1–2 examples of approved small molecule drugs for each target are listed. The table was compiled according to references [24,217,220,221].

Human Virus	Viral Target	Approved Drug(s)
Human influenza virus	• RNA polymerase • Neuraminidase • Matrix protein 2	• Baloxavir marboxil • Oseltamivir, zanamivir • Rimantadine
Respiratory syncytial virus (RSV)	• RNA polymerase • Fusion glycoprotein	• Ribavirin • Palivizumab
Severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2)	• Main protease M^{pro}	• Nirmatrelvir[221]
Varicella-zoster virus (VZV)	• DNA polymerase	• Brivudine, acyclovir, valaciclovir

Like bacteria, viruses can develop resistance to antiviral medications over time.[220] The widespread use of antiviral drugs, particularly in the context of chronic infections such as HCV and HIV, has led to the emerge of drug-resistant strains. The common resistance mechanisms are mutations in the viral genome that cause the target structure to escape the drug. For instance, resistance of HCV against NS5A inhibitor daclatasvir rises from mutations in the NS5A region of the viral genome.[222]

Certain viral infections such as HIV, hepatitis B and C, and herpesviruses, can establish long-term, chronic infections in individuals. While existing antiviral therapies can effectively suppress viral replication and improve clinical outcomes, they are often unable to eliminate the virus from the host's body. Most viral infections can also coexist with other infectious diseases or underlying health conditions, leading to complex clinical scenarios (e.g., acquired immunodeficiency syndrome (AIDS)).[220] Effective antiviral drugs are essential for managing these co-infections and reducing the risk of complications in vulnerable populations.

In summary, the increasing prevalence of viral threats, coupled with challenges such as drug resistance and chronic infections, undergoes the critical importance of antiviral drug discovery in safeguarding public health and combating infectious diseases.

1.6.2. Viral RNAs as Drug Targets

Currently, all approved antiviral drugs target viral proteins, including proteases, integrases, neuraminidases, and RNA polymerases (Table 5). Thus, the resistance rate increases since there are mutations in the viral genome coding for the target proteins.[222] A novel approach in antiviral drug development is directly targeting viral ncRNA with small molecules, particularly highly conserved regions to reduce the degree of freedom for mutations that are compatible with virus functions. Small molecules binding to ncRNA might modulate stability, interactions, and structural features like conformation in order to inhibit essential viral processes. As of the current time, there are multiple studies demonstrating this concept for various viruses, including HIV and HCV. However, target selectivity remains still a huge challenge in this field. Therefore, the fundamental research on the mechanisms of action of these viral ncRNAs is crucial for the development of novel antiviral therapeutics against clinically relevant viruses like bloodborne viruses (HIV, HBV, and HCV), and insect-borne viruses (dengue and zika virus) as well as filoviruses like Ebola and Marburg.[16]

1.7. Hepatitis C Virus

1.7.1. Prevalence and Risk Factors

Hepatitis C is a liver inflammation caused by infection with the hepatitis C virus (HCV). The disease can be acute or become a chronic manifestation leading to serious lifelong illness including liver cirrhosis and cancer. HCV is primarily transmitted through exposure to infected blood, such as through unsafe injection practices, healthcare procedures, blood transfusions, and sexual practices. The WHO reports that 50 million people worldwide are affected by chronic HCV infection, resulting in over 200,000 deaths due to liver cirrhosis and hepatocellular carcinoma in 2022. While effective medications such as sofosbuvir exist for HCV infection, there is currently no vaccine for prevention. Unfortunately, access to diagnosis and treatment remains limited, with only 36% of those affected knowing their diagnosis and even fever (20%) having received treatment with direct acting antivirals (DAAs). This highlights the necessity for additional strategies to address the lack of access to healthcare, particularly in countries with a high prevalence of HCV.[223]

1.7.2. Organization of the HCV Genome

HCV represents an enveloped positive single-stranded RNA virus belonging to the genus *Hepacivirus* of the *Flaviviridae* family, with a genome spanning approximately 9600 nucleotides organized in a linear, uncapped structure. Central to its genetic makeup is a lengthy open reading frame (ORF), flanked by short highly structured UTR at both its 5'- and 3'-ends. The 5'-UTR, comprising roughly 340 nucleotides, encompasses the pivotal internal ribosome entry site (IRES), crucial for initiating the translation of the viral genome. Conversely, the 3'-UTR consists of three distinct segments: a variable region spanning approximately 40 nucleotides downstream of the HCV coding sequence, a poly(U/UC) tract of variable length, and a notably conserved sequence of 98 nucleotides recognized as the X-tail. Both the poly(U/UC) tract and the X-tail play indispensable roles in HCV replication and infectiveness. The expression of viral proteins is orchestrated through a polyprotein precursor, subsequently undergoing proteolytic cleavage to yield structural proteins (core, envelope proteins E1 and E2) and non-structural (NS) proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B), alongside the hydrophobic peptide p7. Notably, a *cis*-acting RNA element (CRE) is situated within the 3'-terminal coding region of NS5B, facilitating long distance RNA-RNA interactions with the X-tail, a necessity for efficient viral replication.[224]

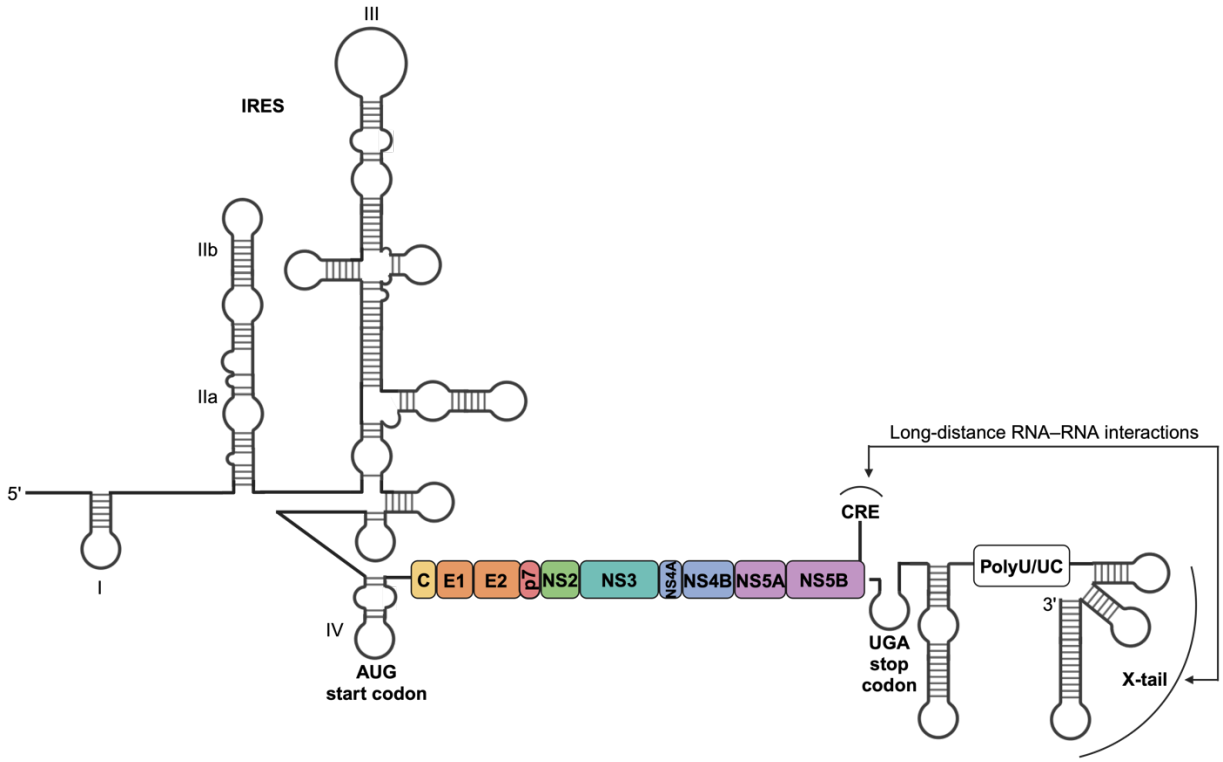


Figure 16: Schematic of the HCV genome. The IRES is located within the 5'-UTR (left) and carries the four for replication essential subdomains II–IV. Domain IV carries the start codon AUG followed by the coding region (colored). The 3'-UTR carries the UGA stop codon, followed by a poly(U/UC) tract and a conserved X-tail. This X-tail forms long-distance RNA–RNA interactions with the CRE, which is located at the 3'-end of the NS5B coding region immediately upstream of the stop codon. This figure was created with BioRender.com according to references[224,225].

1.7.3. HCV IRES

The translation of the HCV genome is regulated by an IRES element located in the 5'- UTR. The IRES binds directly with high affinity to the 40S ribosomal subunit and requires only the two eukaryotic initiation factors eIF2[20] and eIF3 for assembling into translation-competent ribosomes directly located at the start-codon.[226] This leads to a 5'-cap-independent translation mechanism. The 5'-UTR of the HCV can be divided into four distinct domains (I-IV) connected by flexible linkers. Domain I contains the stem loop, which is not related to IRES activity. Domain II mediates the release of eIF2 and participates in the formation of translation-competent ribosomes.[226] For this process, the L-shape of subdomain II is crucial since it steers the apical hairpin loop IIb in the direction of the ribosomal E site near the active site. Once the ribosome is associated, a conformational change closes the mRNA binding cleft.[21,40,227] This conformational change (Figure 17 B, C) has been exploited in various studies in order to develop small

molecule inhibitors (e. g., the benzimidazole compound[40] shown in figure 17 A) that binds selectively to the subdomain IIa.[17–21,40,228–231]

Domain III is capable of binding to eIF3 and ribosomal subunits[226], whereas domain IV contains the start codon, which is exposed during initiation assembly.[17] Domains II-IV consist of independently folding structures essential for virus activity and survival.[17]

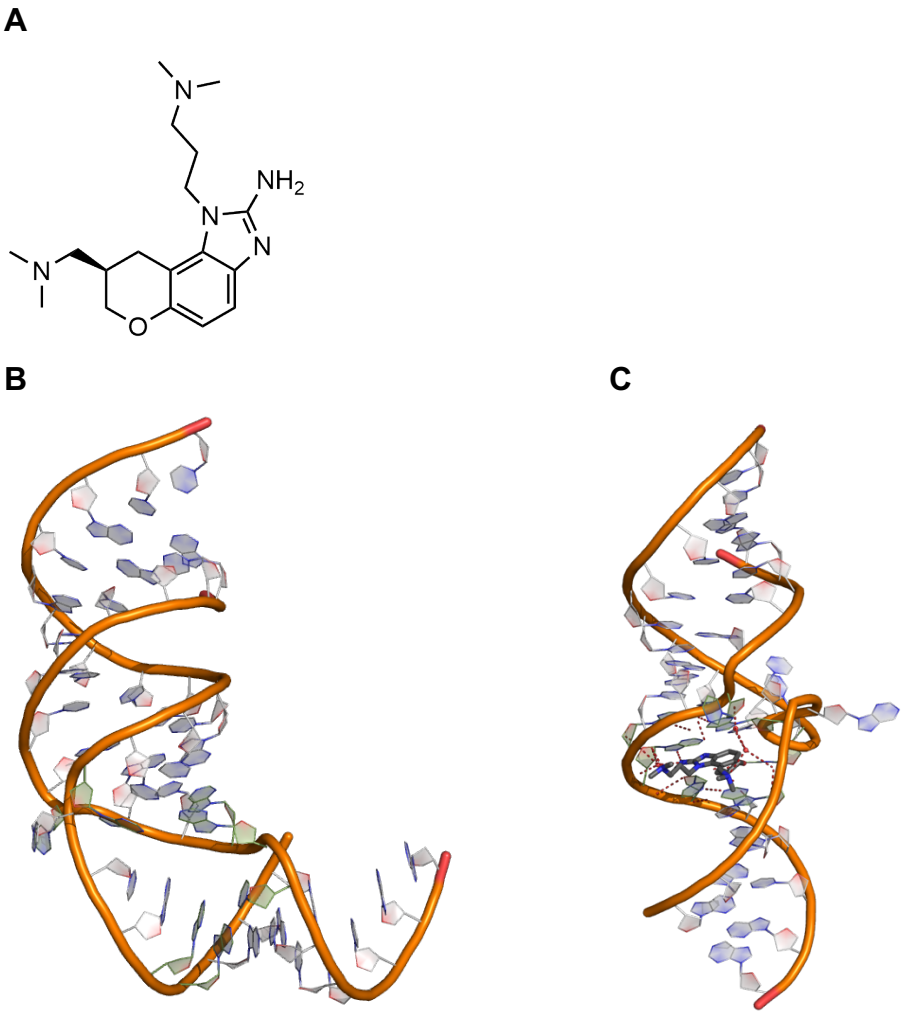


Figure 17: Conformational change of the HCV IRES tertiary structure upon ligand binding. **A**, Structure of the literature-known[40] HCV IRES subdomain IIa binder. **B**, "Kinked" tertiary structure (PDB: 2NOK). **C**, "Straight" tertiary structure (shown as cartoon) complexed with the benzimidazole inhibitor (shown as sticks; PDB: 3TZR). Panel **A** was created with ChemDraw 20, panels **B** and **C** were created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

1.7.4. HCV IRES as Potential Drug Target

As the IRES subdomain IIa is the most conserved sequence of the HCV genome, it became a promising drug target. This conservation implies that any drug targeting this region has the potential to be effective against a broad spectrum of HCV genotypes, as well as limited potential for resistance. Any mutations, that may confer resistance to drugs targeting IIa are likely to come with a significant fitness cost to the virus.[232] The subdomain IIa plays a crucial role in the initiation of translation during the viral replication process. It functions by interacting with the host ribosome and other cellular factors to facilitate the translation of viral RNA into proteins essential for viral replication.[19–21] Although IIa interacts with host cellular components, it possesses distinct structural and functional features compared to host RNA elements involved in translation initiation.[225] This distinction allows for the development of specific inhibitors that selectively target viral RNA without showing significant cytotoxicity.[17,229,230] Several studies have demonstrated the feasibility of targeting subdomain IIa with small molecule inhibitors *in vitro* and in animal models, showing promising antiviral activity.[17– 21,40] These findings provide strong foundation for further development and optimization of drugs targeting this region for clinical use.

Overall, the combination of its essential role in viral replication, high conservation, distinctiveness from host cellular machinery, limited potential for resistance, accessibility to small molecule inhibitors, and demonstrated druggability show that the HCV-IRES subdomain IIa is an attractive and promising drug target for the treatment of hepatitis C infections.

1.7.5. HCV IRES as a Model System

Although IRES elements have similar functions for viral replication, they lack of universal conserved features. In contrast to bacterial riboswitches, viral RNAs do not have native small molecule ligands. Viral RNAs mainly interact with proteins to maintain their function, and host factors for translation.[233] Consequently, there is not *per se* a defined binding site, but highly conserved structural features such as IRES elements that possess distinct “pockets” due to their folding.[16] Due to a drastic conformational change of IIa upon ligand binding resulting in translation inhibition, this RNA is an attractive target for rational drug design.[17–21]

The HCV IRES can be used as model system due to its extensively investigation since discovery, resulting in a wealth of structural and functional data that serve as a solid foundation for further studies. The HCV IRES structure is relatively compact compared to other viral IRES elements, making it more amenable to structural studies for developing small molecules.[234] Currently, there are several small molecule inhibitors against the HCV IRES subdomain IIa reported, providing a good starting point for further improvements and new investigations.[17,21,228–230]
In this regard, the aim of this dissertation is the proof-of-concept for structure-based ligand design for RNA targets using computer-aided drug design (CADD) and fluorescence-based assays for subsequent *in vitro* binding studies.
The following publication is part of the results obtained in this project:

Kallert, E.; Almena-Rodriguez, L.; Husmann, J.- Å.; Blatt, K.; Kersten, C. Structure-based virtual screening of unbiased and RNA-focused libraries to identify new ligands for the HCV IRES model system. *RSC Med. Chem.* **2024**, 15(5), 1527–1538.

1.8. Unveiling RNA–Small Molecule Interactions Using Biophysical Methods

1.8.1. Challenges in Assay Development

Understanding the intricacies of molecular interactions lies at the heart of numerous scientific endeavors, spanning fields from biochemistry to pharmacology. Consequently, the utilization of biophysical methods to characterize molecular interactions is of critical significance in order to achieve a more profound understanding. The established methods for studying RNA-small molecule interactions include microscale thermophoresis (MST)[194], isothermal titration calorimetry (ITC)[235,236], surface plasmon resonance (SPR)[237,238], and FRET applications[19,23], in addition to in-line probing[148,239], chemical probing[22,240], circular dichroism[241], FID[61], NMR[112,242], electrophoretic mobility shift[243] and size exclusion filtration[177]. However, the latter are not discussed in detail in this thesis.
The primary objective of assays studying those interactions is the measurement of binding events. For further investigations and validation of detected binding molecules, the knowledge of the RNA's function (e.g., influence of transcription or translation) is

crucial. In comparison to functional assays, such as those applied for enzymatic reactions, binding assays are more challenging. For enzymatic assays, there is always a substrate used which serves as a positive control. It is also the case that several RNAs (e.g., riboswitches) also bind to natural ligands, which can serve as positive control for assay validation. However, for RNAs that lack natural ligands (e.g., viral RNAs like HCV IRES), the development of accurate binding assays is quite challenging, especially in the absence of a reported binder that can be used as a positive control. Any signals that can be interpreted as binding in comparison to negative control signals, must be validated. For this purpose, assays with appropriate read-outs are necessary. Binding events that affect the transcription of DNA into RNA, are typically validated using gel electrophoresis[63,148,150] or promoter reporter assays.[144,244] Those assays rely on DNA templates with a reporter gene (e.g., *β*-galactosidase, luciferase) downstream of a promoter. The reporter gene product is then quantified, providing a measure of transcriptional activity.[144,244] To confirm effects at translational level, protein synthesis using recombinant elements (PURE)[245] or fluorescent reporter assays[246–250] can be applied. PURE systems allow the quantification of synthesized proteins using radioactive or fluorescent labels.[245] Reporter assays are based on mRNA constructs encoding a reporter protein (e.g., luciferase[18,247,248], green fluorescent protein (GFP)[246,247]), whose fluorescence can serve as a measure of translation efficiency. But also, dual read-out assays are possible: Bains *et al.* reported a system using fluorogenic Mango-(IV) RNA that shows fluorescence upon binding to thiazole orange (TO) and shifted green fluorescent protein (sGFP). The function of transcriptional and translational riboswitches was confirmed using this assay.[246]

Additionally, the use of RNA-containing samples in binding assays presents several challenges. RNA molecules in aqueous solutions are inherently labile and prone to degradation in several manners.[251] In comparison to DNA, RNA is more reactive due to its 2'-hydroxyl group, which can be deprotonated. This so-generated nucleophile can readily attack the 5'-phosphodiester group resulting in backbone cleavage.[252–255] This hydrolysis process is catalyzed by water, RNases, divalent cations and Brønsted acids and bases, which highlights the problem of stability in solution.[254,255] Besides hydrolytic processes, RNA degradation can also occur through oxidation with reactive oxygen species.[256] Additionally, the single stranded conformation is less stable than the double stranded DNA.[257] With regard to assay performance, buffering species, salt, pH and

temperature are also critical factors for physical RNA stability.[251,253,258] It has been reported that Tris-HCl, Tris-borate and sodium borate are buffer species that significantly enhance RNA degradation in the presence of Mg^{2+} at temperatures above 37 °C.[253] Additionally, high Mg^{2+} concentrations result in degradation by forming insoluble aggregates with RNA.[251,253,258] This evidence demonstrates the significance of monitoring the Mg^{2+} concentration, as this cation is essential for RNA folding[87–89] and therefore indispensable in assay buffers. Since it is well established that acids and bases promote RNA hydrolysis, it is important to choose an appropriate pH value.[251,253,255,257] Consequently, a meticulous selection of buffer composition and pH value is imperative.

Typically, DMSO is employed as co-solvent for the solubilization of hydrophobic compounds due to its miscibility with water and lack of reactivity.[259] It is well-established that high DMSO concentrations (> 75%) exert a significant influence on RNA structure and folding energetics.[260,261] For binding assays, a commonly utilized maximum DMSO content is 10%. However, recent studies have demonstrated that DMSO concentrations above 5% destabilize noncanonical residues within bulges and loops resulting in a significant reduction in binding affinities.[262,263] This directly affects considerations in drug discovery with regard to hydrophobic compounds. It is essential to strike a balance between lipophilicity and solubility in appropriate solvents.

The characterization of RNA conformational dynamics and an understanding of their impact on small molecule interactions is of paramount importance for the accurate interpretation of binding assay results. The selection of an appropriate assay method for the study of RNA-small molecule interactions is contingent upon factors such as the binding affinity range, sample requirements, and the desired output parameters.

Each assay method possesses its own set of advantages and limitations, and the selection of the most suitable method requires careful consideration of the specific characteristics of the RNA sample and research objectives.

In the event of a limited sample quantity, it is advisable to employ spectroscopic assays such as MST, SPR or FRET-based plate reader assays in preference over ITC. However, MST and FRET assays are fluorescence-based methods, and it is therefore crucial to select an appropriate labeling method to avoid any risk of altering the binding event or read-out interference. Traditionally, RNA molecules are covalently labeled with a fluorescent probe at either their 3’- or 5’end. These dyes may alter the physical

properties of the RNA, potentially influencing the interaction with the ligand, particularly when this binding occurs near the labelled position.

In the case of SPR, one interaction partner (e.g., RNA) must be immobilized on a metal surface. It is important to select an immobilization method that avoids affecting RNA activity or accessibility to small molecules. The selection of the buffer needs to be adjusted to the nature of the ligand and the RNA sequence. Increasing the ionic strength may result in a decrease of binding affinity of positively charged ligands to the negatively charged RNA due to shielding effects. Additionally, the regeneration conditions must be harsh enough to remove formed complexes, but mild enough to keep the RNA unharmed. As for many molecules DMSO is added as a co-solvent, a buffer correction has to be applied to compensate the change in the refractive index.[264] Nevertheless, SPR is one of the most frequently employed spectroscopic techniques for the study of the kinetics of biomolecular interactions.[265]

MST has the significant advantage of requiring an extremely low sample volume (≤ 10 µL) and being an immobilization-free method. However, it should be noted that careful considerations must be made when the expected affinity is in the high micromolar range or kinetic properties are of interest. The accurate determination of affinities necessitates the presence of sufficient ligand concentrations above their K_D values to achieve saturation binding. For low-affinity ligands, there is the possibility of solubility issues at those high concentrations, which may result in aggregation and undetectable signals. Furthermore, MST does not provide kinetic data, as it is an equilibrium-based technique.[62,265,266]

FRET assays have been employed in numerous studies investigating RNA interactions.[2,18,19,267–270] For instance, the binding of aminoglycosides to the bacterial ribosome has been examined using constructs comprising an isosteric nucleoside analogue as donor and tagging the antibiotics with an acceptor.[267] It is also possible to exploit conformational changes of RNA molecules upon ligand binding. For this purpose, the RNA is labelled with donor and acceptor fluorophore and due to conformational change, the FRET signal might increase or decrease.[18,19,269] Designing an appropriate FRET system is a complex process that requires careful consideration of numerous factors. This includes the development of a structural model (typically databanks such as the PDB are helpful for this purpose), selecting the FRET pair, and implementing a labelling strategy that yields the desired positioning. Furthermore, the estimated FRET efficiency and mechanism must be confirmed.[271] For instance, for

FRET assays that rely on the conformational change of the RNA target, it is essential to design appropriate constructs to accurately report the interaction of interest. The RNA must be adjusted in a length to allow the FRET pair to move in or out of its Förster radius. This implies that the FRET pair must be chosen carefully, depending on the required radius. Additionally, it is important to ensure that the labelling does not negatively affect the folding of the RNA or its interaction with other molecules.[19] Consequently, this method may not be suitable in general.

The ITC represents the only label-free method of these four assays, which are discussed below. This method provides thermodynamic and, in some cases, kinetic parameters with high reliability.[272] In contrast to other methods, it requires a high sample volume, especially for weaker binding events, might show lower sensitivity, and is time-consuming.[236] However, it is essential to consider a number of variables that may impact the outcome of ITC experiments and subsequent interpretation. To prevent erroneous calculations, it is imperative to ascertain the precise concentration of the RNA within the cell and the ligand in the syringe as well as the correct folding of the RNA.[235,273] Furthermore, an appropriate concentration ratio between RNA and ligand must be selected to achieve saturation within a reasonable number of injections. When preparing samples using organic co-solvents such as DMSO, it is of the utmost importance to ensure that all buffer components are perfectly matched. In the event of a discrepancy in the concentration between the cell and the syringe, the result is likely to be a significant heat effect.[236,273] Additionally, it is crucial to be aware that certain interactions may result in a protonation change when selecting the buffer medium. Typically, the protons are released or absorbed by the buffering agent. In this instance, a buffer ionization correction must be conducted in order to obtain the true binding enthalpy of an interaction.[236,274] Therefore, it is advisable to perform ITC experiments after narrowing the number of small molecules of interest. The wealth of information that can be obtained in a single experiment renders this method the current predominant approach in RNA-small molecule binding studies.[265]

To address these challenges, it is necessary to employ rigorous sample preparation techniques, meticulous experimental design, and strict validation of assay conditions to ensure the reliability and reproducibility of binding assay results involving RNA-containing samples.

1.8.2. Microscale Thermophoresis

Microscale Thermophoresis (MST) stands as a testament to the ingenuity of modern biophysical techniques, offering a sophisticated means to probe molecular interactions in solution. Its emergence has reshaped the landscape of biochemical and biophysical research, providing scientists with a highly effective instrument for studying various biomolecular interactions in solution, with minimal sample volume required.[62,266] The essence of MST lies in its exploitation of the thermodynamic phenomenon known as thermophoresis, whereby molecules migrate in response to a temperature gradient, reflecting their intrinsic properties and interactions.[62,266,275] This technique has garnered significant attention due to its inherent advantages, including high sensitivity, minimal sample consumption, and versatility across a spectrum of biomolecular systems.[62,266] Recently, MST has found wide-ranging applications in various fields, including biochemistry, biophysics, and drug discovery.[194,276–279] Moreover, this method has been employed in fragment-based drug discovery, screening small molecule libraries for potential drug candidates, and optimizing lead compounds for binding affinity and specificity.[280,281]

The thermophoresis is described as the mobility of the molecules in response to a temperature gradient (usually 2–5 °C) and was first described in the 19th century by Ludwig.[282] A few years later, the Swiss scientist Soret described the phenomenon of a salt solution in a tube with different temperatures at each end, which does not contain an overall equivalent concentration. The solution was found to be more concentrated in the region of the tube with lower temperature, which suggests a temperature-dependent movement of the salt molecules.[283,284] Thermophoresis can be mathematically described as a Soret coefficient S_T of the free molecules in solution:

$$S_T = \frac{A}{kT} \cdot (-\Delta S_{hyd}(T) + \frac{\beta \sigma_{eff}^2}{4\varepsilon\varepsilon_0 T} \cdot \lambda_{DH} \quad (4)$$

where the S_T correlates linearly with the surface area of the molecules A and the Debye radius λ_{DH}, a measure of a charged molecule's net electrostatic effect in solutions[285]. σ_{eff} is the effective charge, ΔS_{hyd} is the hydration shell effect, ε is the dielectric constant, and β is temperature derivative of ε.[275] Consequently, MST strongly depends on inherent molecular properties such as charge, shape and size, but also on external

factors like salt concentration, which are affected by binding events.[62,265,266,275] To visualize the thermophoretic movement in the heated spot, a molecule must be labelled with a fluorescent reporter. The sample containing the molecules of interest are placed in a microliter-scale capillary and then a temperature gradient is applied by heating with an IR-laser (Figure 18 A).[62,266,286] The MST signal can be separated into two different effects by time: fast T-Jump (< 1 s and slower thermophoresis (> 1 s). The T-Jump is characterized by its fast and strong decrease in fluorescence based on the temperature-related intensity change (TRIC) of the fluorescent dye, whereas the thermophoresis is more related to the binding event. When the binding event leads to a change of the internal state of the labelled molecule, the T-Jump signal might be affected and is used to investigate the interaction (Figure 18 B).[62,266]

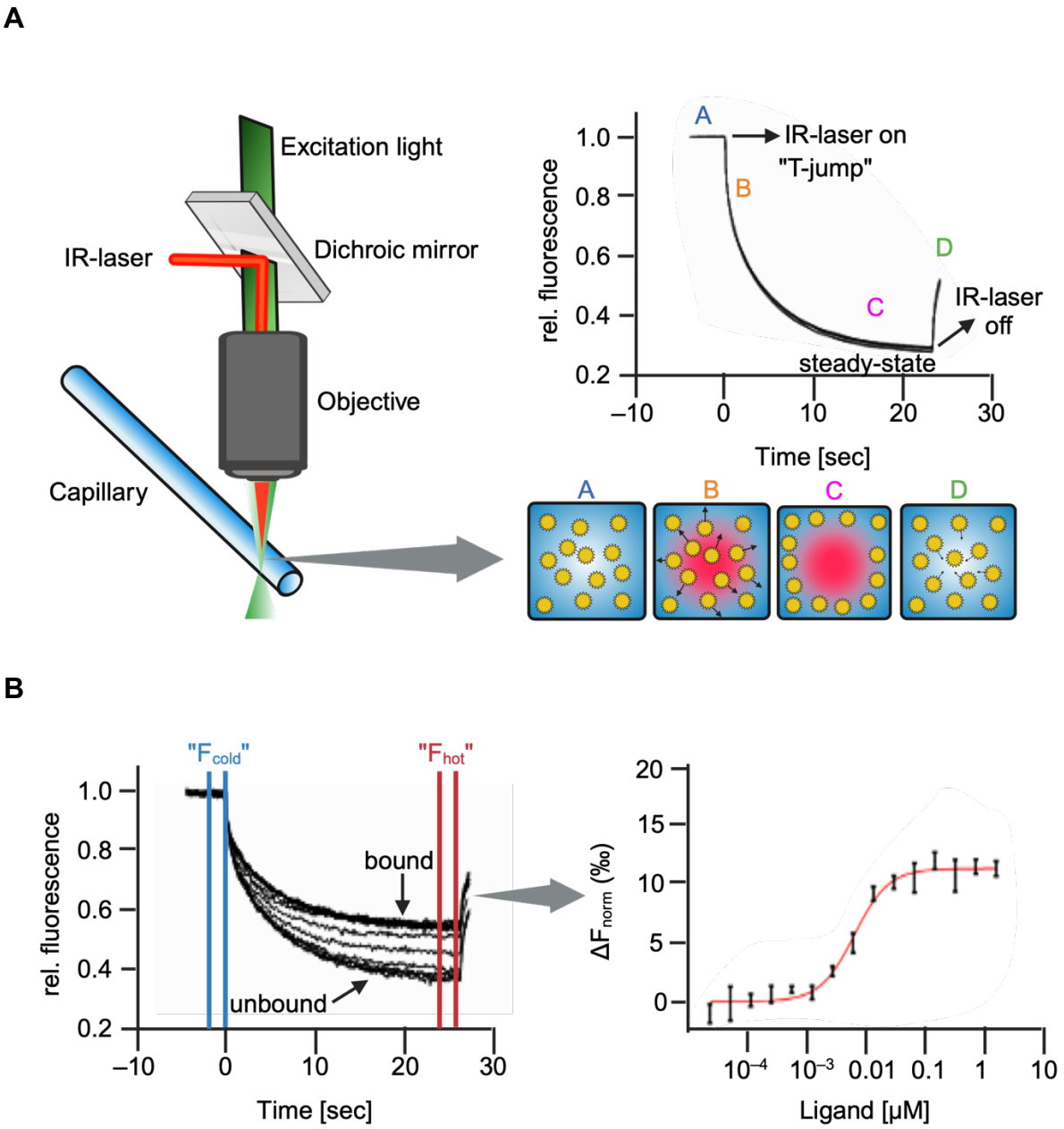


Figure 18: MST experimental setup and results explanation. **A**, IR-laser induced temperature gradient leads to thermophoretic movement of molecules in the capillary. Homogenous distribution of the molecules in solution (A), fast movement of the molecules along the temperature gradient (B, "T-jump"), all molecules are moved out of the heated region (C, steady-state), "backdiffusion" of molecules to homogenous distribution when the laser is switched off (D). **B**, Unbound molecules move faster than bound molecules along the temperature gradient (left). Plotting of the ligand concentration against the change in normalized fluorescence at a specific time during thermophoresis leads to affinity determination (right). This figure was created with BioRender.com according to reference [286]

Traditionally, the target RNA is covalently labelled with reporter dyes such as cyanine or fluorescein derivatives attached to the 3' or 5'-end or to modified nucleosides.[56,194,287,288] This method prevents steric interactions of the fluorophore with the investigated interface but might lead to a decrease of the T-Jump effects.[286,289]

Alternatively, RNA can be labelled with non-covalent fluorescent dyes such as from the Sybr or Syto series to study interactions. Due to the statistical binding of these dyes to nucleic acids[290], the investigated interface is not affected.
In this regard, the aim of this thesis is to introduce a novel non-covalent labelling technique for MST demonstrating the usefulness of this method studying small molecule RNA interactions. This includes a broad spectrum of RNA structures with different size and folding as well as ligands with different molecular properties.
The following publication is part of this RNA-MST project:

Kallert, E.; Behrendt, M., Frey, A.; Kersten, C.; Barthels, F. Non-covalent dyes in microscale thermophoresis for studying RNA ligand interactions and modifications. *Chem. Sci.* **2023**, 14(36), 9827–9837.

Additionally, MST experiments were employed within the $preQ_1$, HCV IRES and electrostatic anchoring projects as part of this thesis.
Furthermore, MST experiment were conducted to obtain binding constants for the SAM-I and SAM-VI riboswitch within a structure/activity relationship study using SAH-analogue compounds. This project will be continued beyond the scope of this thesis and is not yet ready for publication.

1.8.3. Isothermal Titration Calorimetry

Among the techniques available for probing molecular interactions, ITC stands out as a powerful and versatile tool. ITC offers unique insights into the thermodynamic and – in some cases, kinetic parameters[272,291,292] governing binding events between biomolecules, small molecules, or macromolecules.[236,274,293] While the first measurements of reaction enthalpies go back to the 18th century[294], the origins of the ITC can be traced back to the 1960s.[295–298] These developments were crucial in demonstrating the potential of calorimetry to measure heat changes associated with molecular interactions. Throughout the 1980s and 1990s, significant advancements were made in ITC technology, including improvements in sensitivity, data analysis software, and automation.[293] Hence, the application of ITC has surged, driven by its ability to provide direct, label-free measurements with high sensitivity and precision. Fundamentally, ITC operates on the principle of quantifying heat changes associated

with a chemical process occurring at constant temperature. When a ligand is titrated into a solution containing its target, heat is either released or absorbed depending on the nature of the interaction. By monitoring the heat flow accompanying the titration of a ligand into a solution containing its target, ITC enables the determination of thermodynamic parameters such as binding affinity (K_D), stoichiometry (n), enthalpy (ΔH), and entropy (ΔS) of the interaction. The binding isotherm can be obtained by plotting the integrated peak intensities against the molar ratio. The inflection point provides the stoichiometry (n) and the slope of the inflection point is equal to the K_D value. The Gibb's free energy (ΔG) is correlated to the affinity via the following equation, with R as the universal gas constant and T as the absolute temperature in Kelvin.[235,236,274,293]

$$\Delta G = R \cdot T \cdot \ln (K_D) \tag{5}$$

The Gibb's free energy can also be described as the sum of the changes in enthalpy (ΔH) and the negative product of the change in the system's entropy and temperature (– TΔS). The binding enthalpy (ΔH) is obtained from the heat changes of the first injections before saturation. From this relationship, the entropy (ΔS) can be calculated.[235,236,274,293]

$$\Delta G = \Delta H - T\Delta S \tag{6}$$

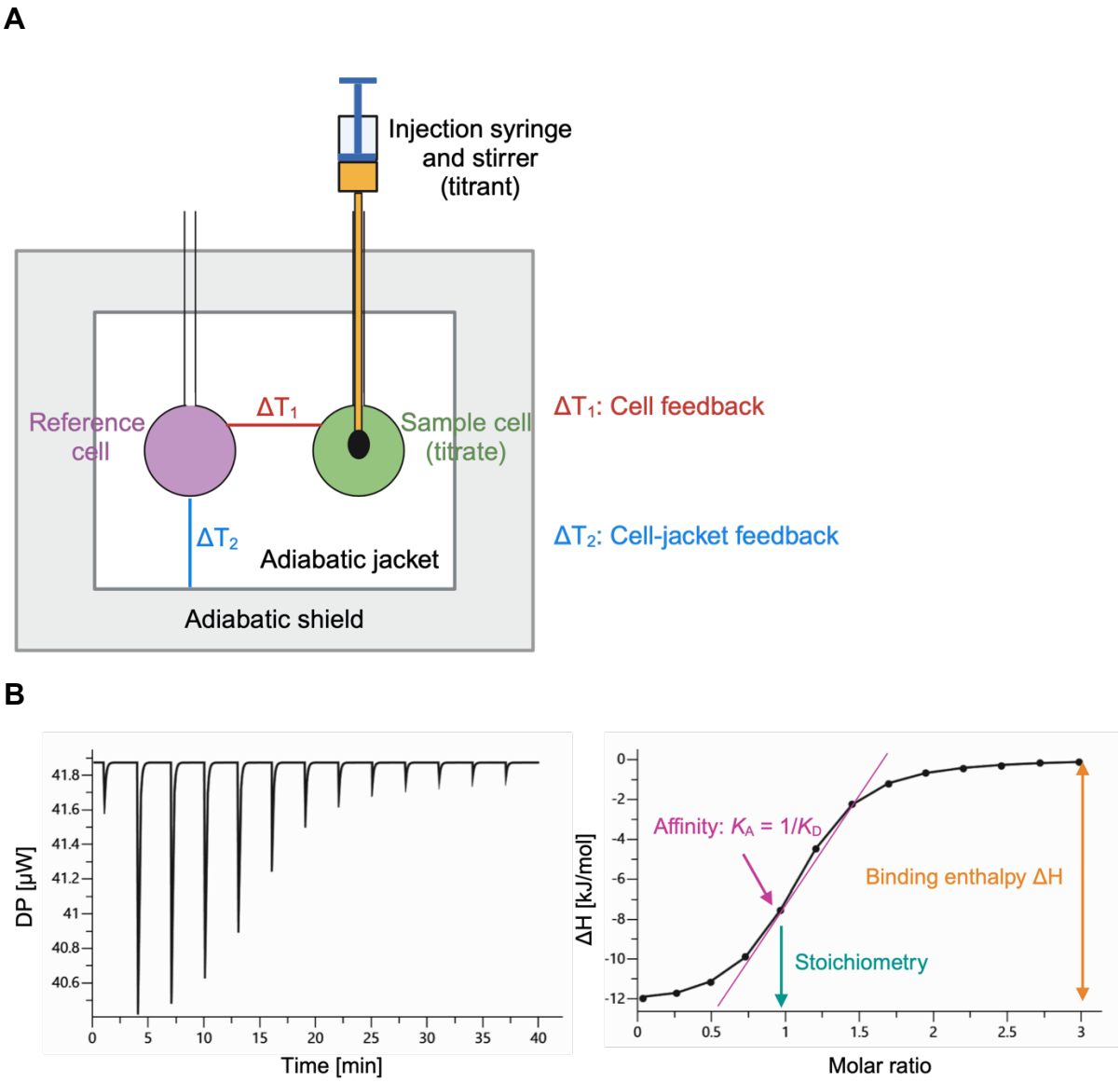


Figure 19: ITC experimental setup and analysis. **A**, Schematic diagram of a typical isothermal titration calorimeter with sample and reference cell in an adiabatic jacket. **B**, Example ITC data. The raw data (left) can be expressed as thermogram where the integrated peak intensities are plotted against the molar ratio of titrant and titrate (right). From the thermogram the binding enthalpy ΔH (orange), affinity (magenta) and stoichiometry (teal) can be obtained directly. This figure was created with BioRender.com.

A more recent approach is the kinetic ITC (kinITC), which allows the simultaneous measurement of both the thermodynamic and kinetic profile of a reaction. kinITC extends the method's capability by analyzing the time course of the heat flow to extract kinetic information about the binding interaction, specifically the association (k_{on}) and dissociation (k_{off}) rate constants. A simple one-step kinetic reaction is represented by A + B ⇌ C with an equilibrium dissociation constant K_D linking thermodynamics and kinetics (equation (5) and (7)).[272,291,292,299]

$$K_D = \frac{k_{off}}{k_{on}} \tag{7}$$

The heat evolved during the formation of one mole product C at a constant pressure is equivalent to the molar variation of enthalpy of the reaction, designated as ΔH. A significant aspect of ITC is that the instrument does not directly measure the total heat (Q) evolved during the reaction; rather, it measures the rate of production or absorption of heat during the reaction. Consequently, the primary signal is the heat power (P_s). By integration, Q can be obtained[272,291]:

$$P_s(t) = \frac{dQ}{dt} \tag{8}$$

As ΔH can be described as the heat evolved during product formation, a direct link between the heat power and production rate of C can be made[272,291]:

$$P_s(t) = V_{cell} \cdot \Delta H \cdot \frac{d[C]}{dt} \tag{9}$$

Where V_{cell} is the volume of the reaction cell. Thus, as $P_s(t)$, ΔH and V_{cell} are known, the rate of production can be obtained. The kinetics resp. the production rate of a one-step reaction is characterized by k_{on} and k_{off}[272,291]:

$$\frac{d[C]}{dt} = k_{on}[A][B] - k_{off}[C] \tag{10}$$

with A being the compound in the cell and B the compound in the syringe.
Equation (10) can be transformed into to permit the manipulation of molar fractions with a clear meaning, rather than absolute concentrations.[272,291]

$$\frac{d[C]}{dt} = k_{on}[A]_0 \cdot AB - k_{off}[C] \tag{11}$$

where $[A]_0$ is the total concentration of the titrand in the cell, $A = [A]/[A]_0$, $B = [B]/[A]_0$, and $C = [C]/[A]_0$.[272,291]
The determination of on- and off-rates is based on the equilibration time curve (ETC). The equilibration time of each injection is plotted against the stoichiometric ratio of the

titrand and titrator. Typically, the equilibration time slows down close to mid-titration of a reaction described by equation (11) when [A] = [B], because the concentrations of both reactants are low. This was also proven by Burnouf *et al.*[272], who demonstrated that the characteristic time of return to baseline τ for any injection can be described as

$$\tau = \frac{1/s}{k_{off}\left[\left(\frac{1}{s}+B_0+1\right)^2 - 4B_0\right]^{1/2}} \tag{12}$$

where $s = [A]_0/K_D$ is the Wiseman parameter and $B_0 = [B]_0/[A]_0$ is the stoichiometric ratio. While the concentration of B (titrator, syringe) and the stoichiometric ratio B_0 increase from injection to injection, the concentration of A (titrand, cell) only decreases slightly by dilution. Assuming that $[A]_0$ is constant, equation (12) shows that τ reaches its maximum at $B_0 = 1 - 1/s$ which is related to mid-titration's equal concentrations of A and B. As the K_D value is known from the thermodynamic ITC results, it can serve as a constrain determining k_{on} from equation (12).[272,291]

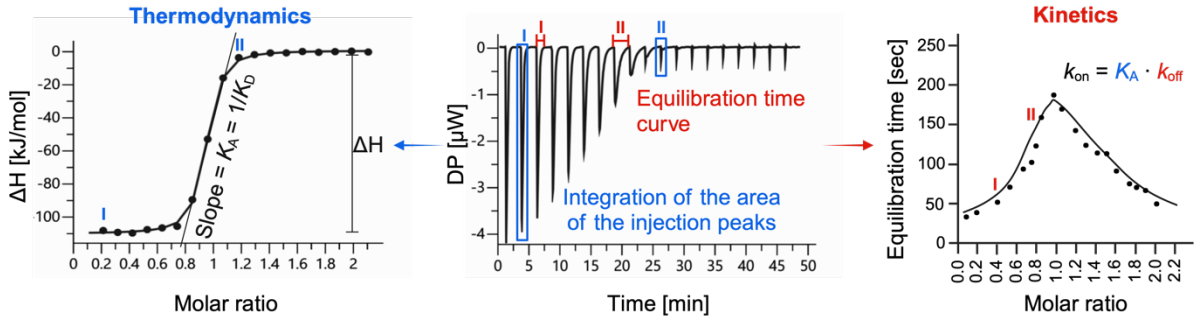


Figure 20: Determination of the kinetic parameters k_{on} and k_{off} from an ITC experiment. Middle: Raw data of an ITC experiment with 19 injections that exhibit longer equilibration times around the mid-titration. Left: K_D determination of the interaction from the slope. The thermogram is obtained by integration of the area of the injection peaks and plotted against the molar ratio. Right: ETC obtained by plotting the equilibration time of each injection against the molar ratio to calculated k_{off}. The association rate k_{on} can be determined by the dependency of k_{off} and K_A: $k_{on} = K_A \cdot k_{off}$. This figure was created with BioRender.com.

This holistic characterization offers profound insights into the driving forces and kinetic properties governing molecular interactions. The versatility of ITC extends across a spectrum of applications, including studies of RNA-small molecule interactions up to macromolecular interactions involving enzymes.[162,193,211,235,236,300–302] RNA-small molecule interactions frequently accompany conformational changes with significant enthalpies, rendering them especially well-suited for ITC analysis.[235] Moreover, its capacity to discern subtle changes in binding affinity due to modifications in molecular

structure or environmental condition renders ITC invaluable in drug discovery, rational design of therapeutics, and elucidation of biological pathways.[300,303]
In this thesis, ITC experiments were employed to investigate the thermodynamic and kinetic binding properties of charged molecules to the *Tte* preQ$_1$ riboswitch and the *Ba* SAM-VI riboswitch within the electrostatic anchoring project.
ITC was also used within the SAR study for the SAM-I and SAM-VI riboswitches. The data for the SAM-VI riboswitch were obtained previously from the electrostatic anchoring project. For the SAM-I riboswitch, first ITC experiments were established using the positive controls SAM and SAH.

1.8.4. Förster Resonance Energy Transfer (FRET)

Over the past two decades, a large number of FRET applications have been established due to the technique's capacity to provide information about intra- and intermolecular changes that occur upon binding events by measuring distances in the nanometer range.[271] FRET relies on the non-radiative transfer of energy between two chromophores, typically fluorophores like Cy3 and Cy5. The excitation energy of one fluorophore (donor) is transferred to another nearby fluorophore (acceptor) through dipole-dipole coupling. The efficiency of FRET depends on several factors. There must be a spectral overlap between the donor emission and the acceptor absorption to provide the required energetic resonance. This spectral overlap (J) can be described by the integral

$$J = \int \bar{I_D} \cdot \varepsilon_A \cdot \lambda^4 \cdot d\lambda \qquad (13)$$

where I_D is the normalized donor emission intensity, ε_A is the molar extinction coefficient of the acceptor and λ the wavelength. In addition to a good spectral overlap, the acceptor molecule must have a large extinction coefficient. Another requirement for FRET is the distance between the two fluorophores. This so called Förster radius R_0 corresponds to the donor-acceptor distance resulting in 50% energy transfer efficiency. R_0 depends on the spectral overlap, the orientation of the transition dipole moments (κ^2), donor quantum yield (Φ_D), and the refractive index of the used medium (n).[271,304]

This dependency can be described as

$$R_0^6 = 0.021 \cdot J \cdot \kappa^2 \cdot \Phi_D \cdot n^{-4} \quad (14)$$

At distances typically within 0.1–10 nm, FRET efficiency (E_{FRET}) follows an inverse sixth power relationship with distance (equation (15)), making it highly sensitive to molecular interactions and conformational changes.[271,304]

The FRET efficiency can be calculated from the single donor-acceptor distance R, the Förster Radius R_0, and the number of acceptors a[271,304]:

$$E_{FRET} = \frac{aR_0^6}{aR_0^6 \cdot R^6} = 1 - \frac{1}{R^6} \quad (15)$$

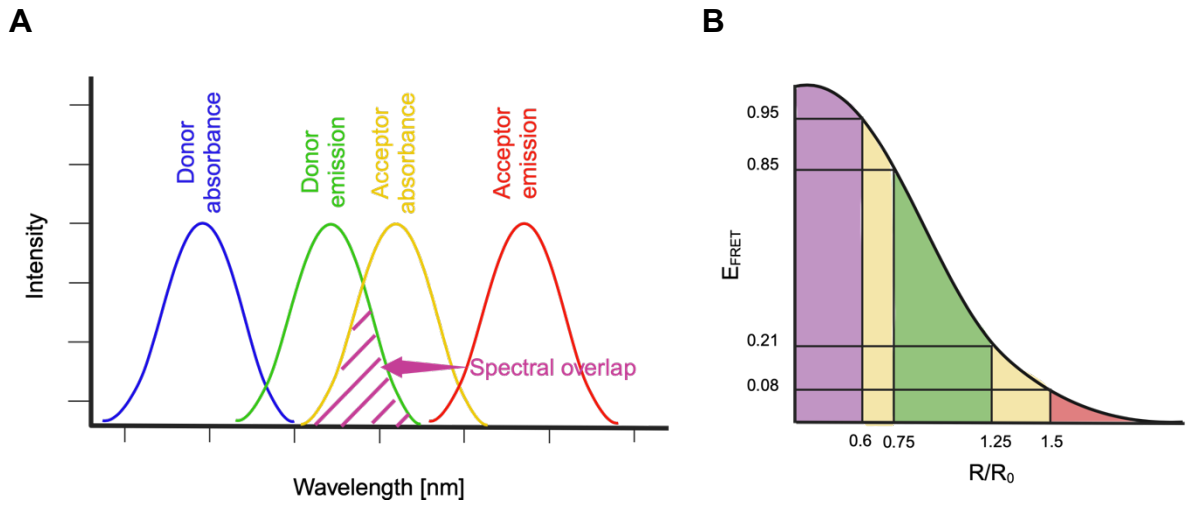


Figure 21: **A**, Overlap (magenta) of donor emission (green) and acceptor absorption (yellow) spectra. **B**, FRET efficiency over distance with a decrease of the dynamic range from the optimal green zone to the yellow to the red and violet zones. This figure was created with BioRender.com according to *reference* [271].

However, to investigate small molecule-RNA interactions, there are two common approaches of using FRET based assays: labelling both the small molecule and RNA or labelling only the RNA with donor and acceptor dye at specific sites.[50] When two labelled molecules are used, the distance between donor and acceptor decreases upon ligand binding. This intermolecular event results in an increase of FRET efficiency. This assay type was used for the screening against the bacterial rRNA A-site.[267] FRET experiments that involve only one labelled molecule (e.g., RNA) are commonly employed to monitor intramolecular conformational changes that occur when a ligand binds. By meticulously affixing donor and acceptor fluorophores at

distinct positions, open and close states can be ascertained by changing FRET efficiency.[18,19,145,184] Therefore, FRET can facilitate the screening of small molecule libraries to identify compounds that selectively bind to RNA targets of interest. By monitoring the changes in FRET signals upon small molecule binding, high-throughput assays can be developed for drug discovery and lead optimization campaigns targeting RNA-mediated diseases such as viral infections, cancer, and neurodegenerative disorders.[50]

In this thesis, FRET was employed within the HCV IRES project to identify small molecules that bind to the HCV IRES. By labelling the RNA with the donor/acceptor pair Cy3/Cy5, the conformational change upon ligand binding was investigated by a decreasing FRET signal.

1.8.5. Surface Plasmon Resonance (SPR)

First introduced in the 1980s for biosensing purposes[305], surface plasmon resonance (SPR) has emerged as a pivotal technique for probing molecular interactions, particularly in the realm of RNA-small molecule interactions.[264,306,307] SPR offers an optical, label-free technique for investigating various biomolecular interactions in real-time, providing insights into affinity, kinetic, and specificity properties.[69,70,264,308] Recently, the interest in developing SPR assays to investigate RNA-small molecule interactions has increased.[264] Understanding the dynamics of these interactions is crucial for elucidating RNA function, designing RNA-targeted therapeutics, and unraveling the intricate mechanisms underlying gene expression regulation.[1] SPR enables the study of various molecular interactions, thin film characterization, and label-free detection with remarkable sensitivity and precision.[69,70] From studying RNA-small molecule interactions[156,306] over monitoring DNA hybridization events[309] to protein-small molecule[310,311], and protein-protein interactions[312,313], SPR offers unparalleled sensitivity and specificity, making it an indispensable tool in drug discovery, clinical diagnostics, and fundamental biological research.[69,70] Moreover, the development of functionalized surfaces and biosensor arrays has extended the applicability of SPR to multiplexed analysis and high-throughput screening.[1,69,308]

At the heart of SPR lies the interaction between electromagnetic waves and free electrons at the interface of a metal and a dielectric medium. When light moves from a medium with a higher density to one with a lower density, the incident angle changes.

At a specific angle, the light undergoes total internal reflection. In SPR, total reflected polarized light hits a metal surface (typically a gold surface) that is placed between two different media (air and sensor medium). At specific conditions, some part of the polarized light is absorbed by electrons in the metal surface. As a result of this energy absorption, the electrons begin to oscillate, forming the surface plasmon. The SPR angle is dependent on the refractive indices of the sensor medium and metal surface. Due to a binding event, the refractive index of the sensor medium changes, which affects the SPR angle and the production of plasmon.[69,70,308]

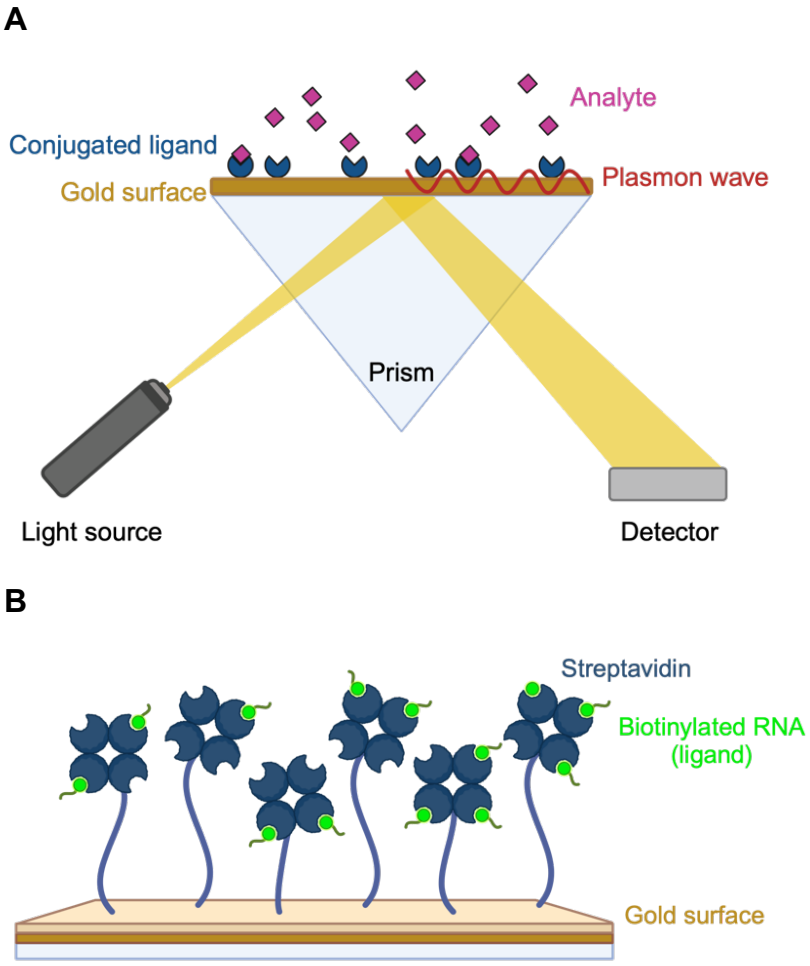


Figure 22: Concept of a SPR experiment for studying RNA-small molecule interactions. **A**, Typical SPR setup with the immobilized ligand on a gold surface, flow cell with analyte, light source, prism and detector. **B**, Streptavidin-functionalized sensor chip for immobilization of biotinylated ligands, e.g., biotinylated RNA. This figure was created with BioRender.com according to references [69,306].

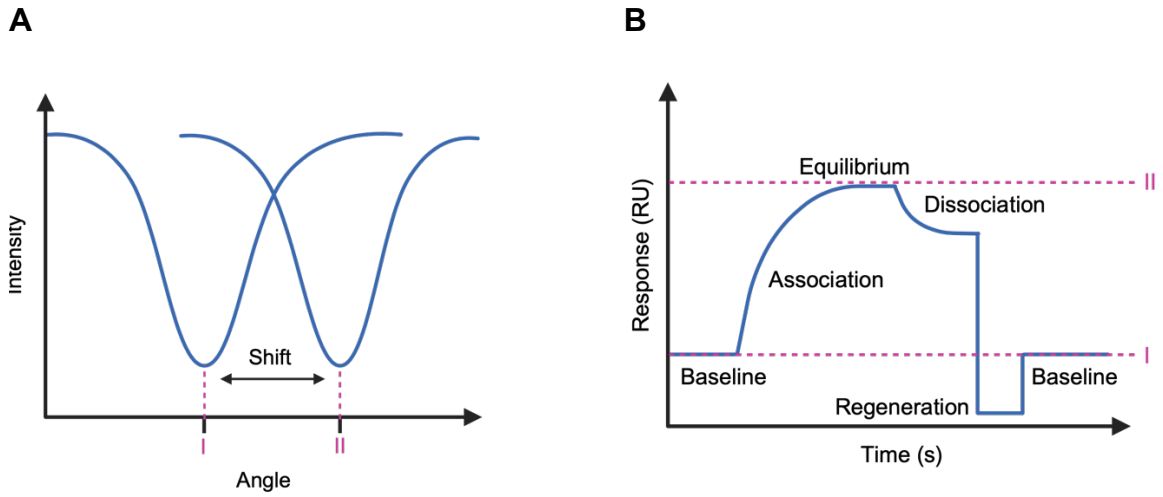


Figure 23: Output of a SPR experiment. **A**, Spectrum of reflected light before (I) and after (II) refractive index change. **B**, Typical sensorgram with refractive index changes during association, equilibrium, and dissociation over time. This figure was created with BioRender.com according to references [69,306].

Experimental setups for SPR typically involve a prism-coupled configuration or a grating-coupled configuration, where light is incident at an angle conductive SPR excitation.[308,314] First, (bio)molecules of interest (e.g., RNA, protein) are immobilized on the metal surface, either by covalent attachment (e.g, by amide-coupling) or capturing (e.g., biotin/streptavidin sensor, protein L/antibody sensor). Then, a solution of a (potential) interaction partner is placed into a flow channel allowing the contact with the coated surface. Any interaction between analyte and ligand results in a change in the refractive index and affects the production of the evanescent wave of the plasmon.[264] The interaction between the evanescent wave and the sample on the metal surface induces changes in the reflectance or transmission spectra, providing valuable information about the sample properties.[314,315] The read-out of SPR experiments are sensorgrams with response units (RU) plotted against the time. 1 RU is equivalent to a critical angle shift (10^{-4} °) of the incident light. At the beginning of any experiment, the starting critical angle resp. RU value is determined.[69] By fitting the sensorgrams in an appropriate binding model, the association and dissociation rates and subsequent, the K_D (equation 7) value can be calculated. k_{on} can be obtained from the increase in the RU signal during the association phase. k_{off} is determined from the rate of decrease in the RU signal during the dissociation phase.[69,264]

For a reaction type with k_{on}- and k_{off}-rates and a resulting K_D (equation 7), the equilibrium binding affinity K_A is defined as[264]:

$$K_A = \frac{[C]}{[A][B]} = \frac{k_{on}}{k_{off}} = \frac{1}{K_D} \quad (16)$$

Where [A] is the concentration of the injected ligand, [B] is the concentration of the unbound immobilized molecule, and [C] is the concentration of the bound complex. The association is described as the increase of [C] as a function of time[264]:

$$\frac{d[C]}{dt} = k_{on} \cdot [A][B] + k_{off}[C] \quad (17)$$

The dissociation is described as the decrease of [C] as a function of time[264]:

$$-\frac{d[C]}{dt} = k_{off} \cdot [C] \quad (18)$$

If a saturation is reached, the K_D can be calculated from the SPR response in the steady-state plateau (assuming a 1:1 binding model)[264]:

$$\frac{K_A \cdot A_{free}}{1 + K_A \cdot A_{free}} = \frac{RU_{obs}}{RU_{max}} \quad (19)$$

Where A_{free} is the free ligand concentration in equilibrium with the complex, RU_{obs} is the observed RU in the plateau region, and RU_{max} is the predicted maximum RU for a 1:1 binding. A_{free} is constant as the free ligands are in flow.
The RU_{max} can be calculated from the molecular weights of the ligand molecule (MW_{lig}) and the immobilized molecule (MW_{imm}) and the RU value of the immobilized molecule (RU_{imm})[264]:

$$\frac{RU_{max}}{RU_{imm}} = \frac{MW_{lig}}{MW_{imm}} \quad (20)$$

In addition to the theoretical calculation, RU_{max} can be determined experimentally. In cases where the experimental RU_{max} is greater than the theoretical RU_{max}, it can be

assumed that there is more than one binding site, as equation (17) applies for a 1:1 binding model.[264]

A typical sensorgram shows a rapid increase in RU during the analyte injection (association phase), followed by a plateau if equilibrium is reached, and then a gradual decrease in RU once the injection stops (dissociation phase). Sensorgrams showing both fast association and dissociation phases, suggest a non-specific binding. This phenomenon can be observed by electrostatic interactions between charged molecules, e. g., aminoglycosides and ribosomal RNA.[238]

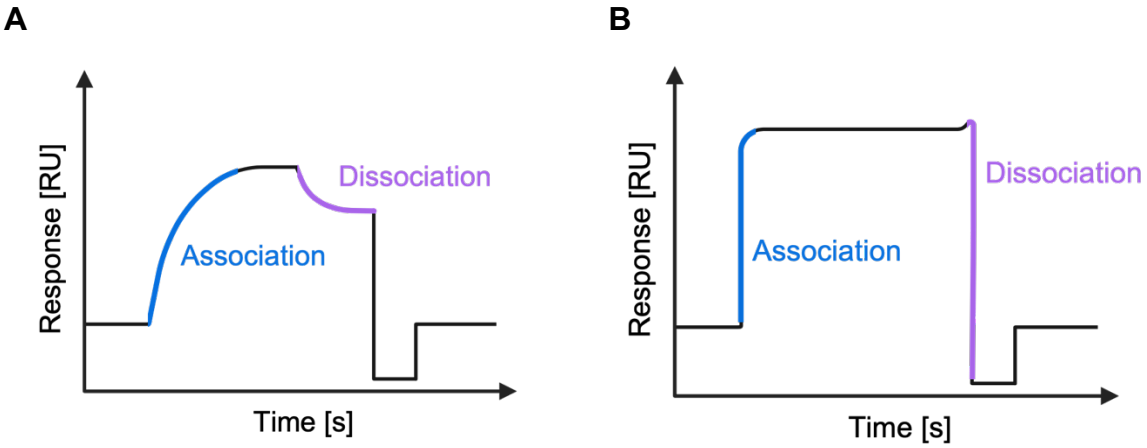


Figure 24: Sensorgrams of specific and non-specific binding between two interaction partners. **A**, Specific binding sensorgram with moderate association (blue) and dissociation (violet) rates. **B**, Un-specific binding sensorgram with typical fast association (blue) and dissociation (violet) rates. This figure was created with BioRender.com.

To investigate RNA–small molecule interactions, the RNA molecules are typically immobilized onto a metal surface by streptavidin-biotin capturing. The small molecules are then allowed to flow over the sensor surface.[306] The resulting binding events induce alterations in the SPR signal, manifested as shifts in resonance angle or changes in reflectance intensity.[70,306]

In this thesis, SPR experiments were thankfully conducted and carefully evaluated by Kirsten Schaubruch and Prof. Dr. Ralf Heermann within the electrostatic anchoring project. The aim was to obtain thermodynamic and kinetic binding properties of differently charged small molecules to the *Tte* preQ$_1$ riboswitch and the *Ba* SAM-VI riboswitch.

2. List of Publications and Manuscripts

2.1. Publications and Manuscripts as Part of this Doctoral Thesis

(1) **Kallert, E.**; Fischer, T. R.; Schneider, S.; Grimm, M.; Helm, M.; Kersten, C. Protein-Based Virtual Screening Tools Applied for RNA–Ligand Docking Identify New Binders of the $PreQ_1$-Riboswitch. *J. Chem. Inf. Model.* **2022**, 62(17), 4134–4148.

Own contributions as part of this doctoral thesis: Within this publication, covalent MST assay was established and evaluated to discover novel binders for the $preQ_1$ riboswitch. MST measurements were carried out using a Monolith NT.115 instrument (NanoTemper Technologies, Munich, Germany) using standard coated capillaries, the results were analyzed using the MO Affinity software (version 2.3).

Other contributions: Within this publication, molecular and structure-based virtual screening were conducted using LeadIT, FlexX, DOCK, FRED, HYBRID and rDOCK to identify potential binders the $preQ_1$ riboswitch.

(2) **Kallert, E.**; Behrendt, M., Frey, A.; Kersten, C.; Barthels, F. Non-Covalent Dyes in Microscale Thermophoresis For Studying RNA Ligand Interactions and Modifications. *Chem. Sci.* **2023**, 14(36), 9827-9837.

Own contributions as part of this doctoral thesis: Within this publication, different RNA oligos were divided into distinct classes using the RNAfold webserver (ViennaRNA web services). RNA- and dye-dependent effects were investigated by non-covalent MST with the different oligo classes. Riboswitch aptamers were prepared and labelled with non-covalent dyes for studying ligand binding in MST. Modified RNAs were labelled with non-covalent dyes for investigation of modifications. Additionally, antibody-assisted detection of the labelled modified RNAs was conducted. MST measurements were carried out using a Monolith NT.115 instrument (NanoTemper Technologies, Munich, Germany) using standard coated capillaries, the results were analyzed using GraphPad Prism 7.01 (GraphPad Software, Boston, MA, USA).

Other contributions: Within this publication, fluorescence measurements were carried out on a Tecan Spark 10 M plate reader (Tecan Group AG, Männedorf, Switzerland). All oligo MST data were analyzed, and a METTL3/14 MST enzyme assay was established. Additionally, a ^{3}H-based enzyme assay was conducted as control experiment. Furthermore, a covalent MST assay for the SAM-VI and preQ$_1$ riboswitch was conducted.

(3) **Kallert, E.**; Almena-Rodriguez, L.; Husmann, J.- Å.; Blatt, K.; Kersten, C. Structure-Based Virtual Screening of Unbiased and RNA-Focused Libraries to Identify New Ligands For the HCV IRES Model System. *RSC Med. Chem.* **2024**, 15(5), 1527–1538.

Own contributions as part of this doctoral thesis: Within this publication, FRET and covalent MST assays were established in order to investigate the interaction between HCV IRES and potential ligands as well as their selective binding to this RNA motif. MST measurements were carried out using a Monolith NT.115 instrument (NanoTemper Technologies, Munich, Germany) using standard coated capillaries. FRET measurements were carried out using a Tecan Spark 10 M plate reader (Tecan Group AG, Männedorf, Switzerland). The results were analyzed using GraphPad Prism 8.0 (GraphPad Software, Boston, MA, USA).

Other contributions: Within this publication, virtual screening of unbiased and RNA-focused libraries was conducted. MST and FRET assay screening and affinity measurements were provided as well as the MST selectivity assay. MST measurements were carried out using a Monolith NT.115 instrument (NanoTemper Technologies, Munich, Germany) using standard coated capillaries. FRET measurements were carried out using a Tecan Spark 10 M plate reader (Tecan Group AG, Männedorf, Switzerland). The results were analyzed using GraphPad Prism 8.0 (GraphPad Software, Boston, MA, USA).

Elisabeth Kallert and **Laura Almena-Rodriguez** contributed equally to this manuscript.

(4) Almena-Rodriguez, L.; **Kallert, E.**; Husmann, J.-Å.; Schaubruch, K.; Schwickert, M.; Hoba, S. N.; Heermann, R.; Kersten, C. Electrostatic Anchoring in RNA-Ligand Design – Dissecting the Effects of Positive charges on Affinity, Selectivity, Binding Kinetics and Thermodynamics.

Manuscript (4) is a draft version in which parts especially concerning SPR measurements still need to be added by other authors.

Own contributions as part of this doctoral thesis: Within this publication, MST and ITC experiments were established to investigate the influence of charge-carrying ligands to binding affinities (MST) thermodynamics and kinetics (ITC). MST measurements were carried out on a Monolith NT.115 instrument (NanoTemper Technologies, Munich, Germany) using standard coated capillaries and the results were analyzed using GraphPad Prism 8.0 (GraphPad Software, Boston, MA, USA). ITC measurements were carried out on a MicroCal PEAQ-ITC automated system (Malvern Panalytical, Malvern, UK) and the results were analyzed using the MicroCal PEAQ-ITC Analysis software and the AFFINIMETER Web tool (Software 4 Science Developments, Santiago de Compostela, Spain).

Other contributions: Within this manuscript, the SAM analogues, dibenzofuran and carbazole derivatives were synthesized. The purity of each compound was confirmed by LC-MS and NMR spectroscopy. SPR measurements were established using the Biacore T200 (Cytiva, Marlborough, MA, USA) and analyzed with the software (version 3.2.1) provided by the manufacturer. Additionally, SPR measurements were analyzed by interaction map analysis using the Trace Drawer software 1.10.1 (Ridgeview AB, Uppsala, Sweden).

Laura Almena-Rodriguez and **Elisabeth Kallert** contributed equally to this manuscript.

2.2. Publication Beyond This Doctoral Thesis

(5) Barthels, F.; Hammerschmidt, SJ.; Fischer, TR.; Zimmer, C.; **Kallert, E.**; Helm, M.; Kersten, C.; Schirmeister, T. A low-cost 3D-printable Differential Scanning Fluorometer for Protein and RNA Melting Experiments. *HardwareX*, **2022**, 11, e00256.

3. Discovery of Novel Binders for the $PreQ_1$ Riboswitch by MST

3.1. Context, Project Summary, and Own Contribution

Recently, the increasing number of multi-drug resistant bacteria including the ESKAPE pathogens has been underlined the relevance of the discovery of novel antibacterial therapeutics.[15] Therefore, bacterial RNA is a promising target for the development of small molecules modulating their function to fight bacterial infections. Currently, many of the approved drugs have several problems, e. g., toxic side effects of aminoglycosides, and systemic drug delivery issues of siRNAs or ASOs.[2,32] Consequently, selective small molecules represent a promising concept to overcome these issues.[2] For this purpose, various strategies can be adopted from the protein-targeting field, including structure-based small molecule design using computational methods alongside biophysical assays.[2] We evaluated the capabilities of protein-based docking programs in RNA-ligand interactions using the well-studied *Bs* $preQ_1$ riboswitch[63,155,156,181–183,191] as a model system. Our results demonstrated that these programs could reproduce native binding modes of small molecule-RNA ligands and distinguish between known binders and decoys. However, challenges such as protomer and tautomer selections, target dynamics, and explicit solvent presence are similar to those in protein-ligand docking. A prospective virtual screening was performed using the tools FRED, HYBRID and FlexX on the $preQ_1$ riboswitch aptamer domain. 23 hit compounds were selected based on their docking scores and for mimicking key interactions of the native ligand. By employing MST, we identified six binders with potencies ranging from 29.5 nM to 11.0 µM out of the 23 hits. This highlights the significant role of biophysical assays, particularly MST, in validating and characterizing the binding of RNA-targeting ligands identified through VS. This validation step is crucial for verifying the efficacy of the computationally identified hits and ensuring they interact with the RNA target as predicted by the docking studies. With this study, we contributed to the research field in structure based-small molecule design for targeting RNA and demonstrated the feasibility of structure-based VSs against RNA targets.

Own contributions: MST assay establishment, MST screening and affinity measurements, MST data evaluation, HPLC-ESI/MS purity determinations, writing of the original draft & editing of the manuscript.

Contributions from other authors: Molecular docking and structure-based virtual screening, writing of the original draft & editing of the manuscript.

This work has been published in: Journal of Chemical Information and Modeling (impact factor: 5.6).

Article reprinted with permission from *J. Chem. Inf. Model*. **2022**, 62(17), 4134–4148. 'Protein-Based Virtual Screening Tools Applied for RNA−Ligand Docking Identify New Binders of the $preQ_1$-Riboswitch.' ©2022 American Chemical Society.

3.2. Publication

The following publication quoted (within “”) from page 82 to page 131 is exactly the same as the manuscript cited on page 76: “

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

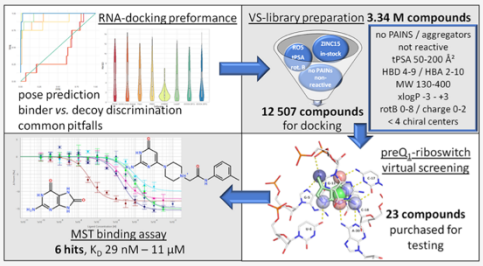
Protein-Based Virtual Screening Tools Applied for RNA−Ligand Docking Identify New Binders of the preQ$_1$-Riboswitch

Elisabeth Kallert, Tim R. Fischer, Simon Schneider, Maike Grimm, Mark Helm, and Christian Kersten*

Cite This: J. Chem. Inf. Model. 2022, 62, 4134−4148 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Targeting RNA with small molecules is an emerging field. While several ligands for different RNA targets are reported, structure-based virtual screenings (VSs) against RNAs are still rare. Here, we elucidated the general capabilities of protein-based docking programs to reproduce native binding modes of small-molecule RNA ligands and to discriminate known binders from decoys by the scoring function. The programs were found to perform similar compared to the RNA-based docking tool rDOCK, and the challenges faced during docking, namely, protomer and tautomer selection, target dynamics, and explicit solvent, do not largely differ from challenges in conventional protein-ligand docking. A prospective VS with the *Bacillus subtilis* preQ$_1$-riboswitch aptamer domain performed with FRED, HYBRID, and FlexX followed by microscale thermophoresis assays identified six active compounds out of 23 tested VS hits with potencies between 29.5 nM and 11.0 μM. The hits were selected not solely based on their docking score but for resembling key interactions of the native ligand. Therefore, this study demonstrates the general feasibility to perform structure-based VSs against RNA targets, while at the same time it highlights pitfalls and their potential solutions when executing RNA−ligand docking.



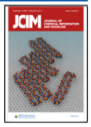
Downloaded via UNIV MAINZ on March 21, 2024 at 08:28:41 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

Because of its fundamental role in many biological processes, RNA has become of increasing interest as a potential drug target over the last years.[1−9] Several antisense oligonucleotides and small interfering RNAs (siRNAs) were approved by the Food and Drug Administration (FDA) and European Medicines Agency or are in clinical development.[10] Targeting RNA with small molecules, however, is still a novel concept, and only one molecule except ribosome-targeting antibiotics is approved by the FDA. With Risdiplam,[11,12] the first approved small-molecule splicing modifier is available for the treatment of spinal muscular atrophy, and further splicing modifiers, micro-RNA (miRNA) ligands, and binders of extended trinucleotide repeats are under elucidation for many indications including cancer[8,13−16] and neurological and genetic disorders.[2,17−25] While only around 1.5% of the human genome is transcribed into proteins, the common drug-targets, about 70−90% is transcribed into RNA.[7,26,27] These mainly noncoding RNAs (ncRNAs) represent promising targets for novel therapeutics. Furthermore, viral RNAs[28] like the SARS-CoV-2 frameshift pseudoknot,[29,30] enterovirus internal ribosomal entry site,[31] or the HIV *trans*-activation response element (TAR)[28,32] were successfully targeted with small molecules. To fight emerging multidrug resistant bacteria, riboswitches (RSs) are attractive targets for the development of new antibiotics.[33−35]

While structure-based computational methods like molecular docking are regularly used for protein targets, most identified RNA-binding small molecules, however, emerged from high-throughput or fragment-screening campaigns and derivatization of a known ligand or serendipity rather than structure-based design by molecular docking.[29,36−45] One challenge in this regard is the low number of available 3D RNA structures compared to the number of proteins deposited in the protein data bank (PDB),[46] often additionally struggling with low resolution. However, these structures already show clearly that the limited chemical diversity of just four RNA bases compared to 20 canonical amino acids in proteins can still result in many different folds, for example, hairpins, G-quadruplexes, multiway junctions, (pseudo)knots, L-shaped tRNAs, triple helices, ribozymes, and the ribosome complex.[7] While docking tools are occasionally used to generate an RNA−ligand binding mode hypothesis,[13,31] prospective virtual screenings (VSs)[47−49] and extensive validation of the used docking programs for RNA targets are rare.[50−53] However,

Received: June 14, 2022
Published: August 22, 2022



ACS Publications © 2022 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

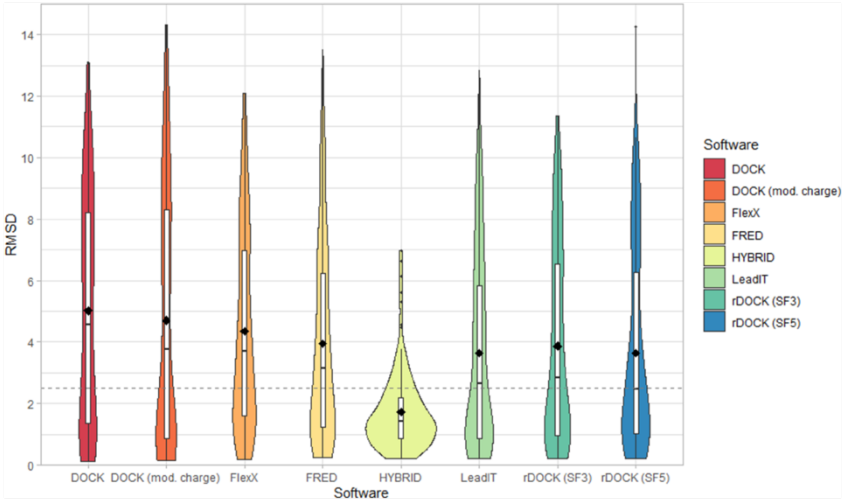


Figure 1. Violin plots of re-docking performance in an automated docking workflow against 150 RNA targets for DOCK 6.9 (red), DOCK 6.9 with modified charges (orange, adding +1 partial charge on phosphorus atoms), FlexX-3 (light orange), FRED 3.3.0.3 (beige), HYBRID 3.3.0.3 (yellow), LeadIT-2.3.2 (light green), and rDOCK (SF3 green, SF5 blue). Bar plots around the median are depicted with 25−75% range in white and mean value as a black dot. The dashed line at RMSD = 2.5 Å resembles the cutoff for a re-docking to be defined as "successful".

validation of conventional protein-based docking tools for its applicability against RNA targets is crucial because parameters cannot easily be transferred and nucleic acids hold their own unique challenges different from proteins for docking programs, like high intrinsic flexibility, a negatively charged phosphate backbone which is eventually masked by cations, or a limited chemical diversity.[54−57] Consequently, over time, some specialized RNA−ligand docking algorithms and scoring functions were developed.[58−67]

Once validated against a given data set to demonstrate the ability to correctly predict crystallographic binding modes (posing) and accurately discriminate binders from nonbinders or decoys by scoring, the true challenge for a docking program is the identification of novel scaffolds for a target by VS. While for proteins prospective docking community challenges like the drug design data resource (D3R) grand challenges[68−72] are performed from time to time, challenges for RNA structure prediction[73−76] but not for molecular docking have been reported so far. For our prospective case study, we selected the *Bacillus subtilis* (*Bs*) prequeosine 1 (7-aminomethyl-7-deaza-guanine, preQ$_1$, **1**) RS aptamer domain as a model system. RSs are noncoding, *cis*-acting RNA elements in the 5′ untranslated region of mRNA regulating gene expression. RSs consist of a ligand binding domain and an expression platform and are usually found in prokaryotes. The preQ$_1$-RS aptamer was selected because of its reported assayability,[77−79] available crystal structures with different ligands,[80,81] and proved druggability as small drug-like ligands structurally distinct from the native ligand preQ$_1$ were reported.[81] PreQ$_1$-RSs are also present in *Enterococcus faecium*, *Neisseria gonorrhoeae*, *Streptococcus pneumoniae*, and *Haemophilus influenzae*, pathogens from the World Health Organization priority list,[34] making the preQ$_1$-RS an attractive target for the development of novel antibiotics against drug-resistant bacteria.[82]

POSE PREDICTION

Prior prospective VSs, the tools under elucidation were evaluated in terms of their ability to correctly predict native binding modes of X-ray and nuclear magnetic resonance (NMR) complex structures and to discriminate known ligands from decoys with similar physicochemical properties.[83,84] The protein-based docking programs tested were FlexX[85,86] and its derivative LeadIT[87] and FRED[88] with its template-docking modification HYBRID.[89] DOCK6[90,91] was included as it was validated besides AutoDock for RNA−ligand docking previously[51,52] and demonstrated to be successful in a prospective VS against a purine RS as well.[47,92] rDOCK,[59] a successor of RiboDOCK,[58] represents a tool, developed explicitly for RNA−ligand docking. It is noteworthy that all dockings were performed in a highly automated process with minimal user intervention and hence are very likely to hold optimization potential.

To elucidate the precision in pose prediction, a panel of 150 RNA−ligand complexes was compiled from previous studies[51,52,59] including ribosome structures, RSs, viral RNA, and artificial RNA aptamers (Table S1). Re-docking performance of the complexes was elucidated by the ligand heavy atom root-mean-square deviation (RMSD) between the docking pose and experimentally determined binding mode. A cutoff for a successful re-docking was defined at 2.5 Å as done previously for RNA−ligand docking.[50,51,67] This slightly higher cutoff compared to the more common 2.0 Å is justified by the challenges arising from the transfer of protein-ligand docking to RNA targets. Without experienced user intervention during receptor and ligand preparation prior docking, the success rates for re-docking were rather poor compared to what might be expected by expert-defined docking protocols. Success rates ranged from 35−50% with median RMSD values of 2.5−4.7 Å (Figure 1, Table 1). However, these success rates lie within the

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

Table 1. Re-Docking Result Summary of 150 RNA−Ligand Complexes

software	mean RMSD (Å)	median RMSD (Å)	success rate (%)[a]
DOCK 6.9	5.2	4.7	35
DOCK 6.9 (mod. charge)	4.7	3.8	45
FlexX-3	4.3	3.7	40
FRED 3.0.3	4.0	3.2	43
HYBRID 3.0.3	1.7	1.4	80
LeadIT-2.3.2	3.6	2.6	49
rDOCK 2013.1-1 (SF3)	3.9	2.9	46
rDOCK 2013.1-1 (SF5)	3.6	2.5	50

[a]The success rate was defined as percent of re-docking poses within an RMSD of 2.5 Å compared to the reference ligand.

expected range of automated, nonoptimized docking screens.[72,93] The exception is HYBRID with a success rate of 80%, but it is to be noted that HYBRID while being based on FRED additionally to the ChemGauss scoring function uses an additional scoring step called chemical gaussian overlay which considers the reference ligand during docking.[89] Therefore, a better performance for re-docking using this template-based approach is not surprising. rDOCK, which was originally developed for RNA−ligand docking, performed only slightly better than protein-based docking programs. As recommended from the original authors, from the two scoring functions SF3 and SF5 within rDOCK, SF5 performed better, likely because of its improved desolvation terms.[59] DOCK, which was improved and validated for RNA−ligand docking previously[52] and was already successfully applied in a prospective VS against the adenine riboswitch,[47] was outperformed with a success rate of only 35% and a median RMSD of 4.7 Å when using a common protein-based protocol. However, it is reported that force-field-based scoring functions tend to overestimate ionic interactions.[54] While the backbone of RNA is highly charged, these charges are neutralized or at least masked by cations in solution. For compensation, the overall backbone charges are usually modified for docking either by adding counter ions[52] as also done in molecular dynamic (MD) simulations[94−96] or by adaptation of the partial charges either of the backbone oxygen[47] or phosphorous atoms.[51] By increasing the partial charge of P by +1 to neutralize the backbone, the success rate of DOCK ("mod. charge", Table 1) was improved by 10%.

BINDER-DECOY DISCRIMINATION

For a prospective VS, docking programs are required not only to identify correct binding modes of known ligands but also to be able to discriminate binders from nonbinders by their docking score.[84] While nonbinders are often difficult to receive from the literature, a common method is the usage of decoys that have similar physicochemical properties like the ligands but are structurally dissimilar to be classified as nonbinders. For the ribosomal 16S aminoglycoside binding site (A-site)[97] and seven RSs (adenine,[47,98] thiamine pyrophosphate (TPP),[99,100] *S*-adenosyl methionine (SAM),[101] flavin mononucleotide (FMN),[40,102] preQ$_1$,[78] guanine,[47] and tetrahydrofolate (THF)[103]), ligands were collected from the literature and corresponding decoys were generated from the "database of useful decoys: enhanced" (DUD-E).[83] After docking of ligands and decoys against the selected target structures, the discrimination was evaluated by receiver operating characteristic (ROC) curves and their area under the curve (AUC) for overall performance as well as the enrichment[104] of top 2, 5, and 10% of the screened data sets (Table 2, Figure S2).

Overall, a good discrimination between reported binders and decoys was observed with ROC AUCs over 0.6 (moderate enrichment) for all except five RS cases of DOCK either with or without charge modification, which might require some further refinement like the inclusion of explicit solvent molecules (as done for the TPP- and adenine-RS) or modification of the binding site size (as for the guanine-RS). In particular, the adenine-RS is an exemplary demonstration of the importance of user intervention prior docking and screening library preparation, as the selected ligands of this RS originate from a VS with DOCK in which an explicit solvent molecule was included in the receptor preparation.[47] For the rRNA A-site, DOCK performed better than other conventional protein-based tools with high ROC AUCs of 0.93 and 0.91 with or without charge modification, respectively.

Similarly, the RNA-specialized rDOCK (SF5) reached a ROC AUC 0.93 for the rRNA A-site. Furthermore, different from DOCK, rDOCK also showed an overall very good performance for all of the RSs as well, with ROC AUCs between 0.60 and 1.00.

The protein-based tools FlexX, FRED, HYBRID, and LeadIT also reached high ROC AUCs and good enrichments, especially for RSs. HYBRID, due to its template-based approach with ligands often being limited in diversity and structurally similar to the crystallographic reference ligand, demonstrated its suitability for RNA−ligand docking with ROC AUCs between 0.71 and 1.00, recovering between 11

Table 2. Binder vs Decoy Discrimination Performance for the Ribosomal A-Site and Seven RSs

				ROC AUC (% of ligands recovered in top 2% of screened database)						
PDB-ID	target	binders	decoys	DOCK	DOCK (mod. charge)	FlexX	FRED	HYBRID	LeadIT	rDOCK (SF5)
1LC4	rRNA A-site	18	1050	0.93 (22)	0.91 (17)	0.72 (6)	0.73 (0)	0.75 (11)	0.68 (0)	0.93 (44)
1Y26	adenine-RS	13	900	0.66[a] (0)	0.48 (0)	0.97 (54)	1.00 (100)	1.00 (100)	0.99 (92)	0.99 (83)
2GDI	TPP-RS	9	550	0.81[b] (38)	0.57 (33)	0.67 (22)	0.79 (33)	0.86 (44)	0.58 (33)	0.86 (44)
2GIS	SAM-RS	3	150	0.77 (0)	0.76 (0)	0.88 (0)	0.99 (100)	1.00 (100)	0.82 (33)	0.92 (67)
3F2Q	FMN-RS	26	600	0.83 (12)	0.87 (46)	0.64 (12)	0.88 (50)	0.91 (58)	0.70 (31)	0.99 (67)
3FU2	preQ$_1$-RS	8	350	0.44 (0)	0.44 (0)	0.94 (75)	0.88 (63)	0.99 (100)	0.97 (75)	1.00 (87.5)
4FE5	guanine-RS	19	1150	0.86[c] (3)	0.58 (16)	0.74 (8)	0.65 (0)	0.71 (2.5)	0.81 (34)	0.83 (31.5)
4LVW	THF-RS	10	550	0.40 (5)	0.46 (5)	0.82 (5)	0.71 (25)	0.73 (30)	0.83 (10)	0.60 (10)

[a]Explicit water molecule ID 290 was manually included in binding site definition. [b]Several water molecules around reference ligand TPP were included in binding site definition. [c]The binding site size was reduced to 4 Å around the reference ligand.

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

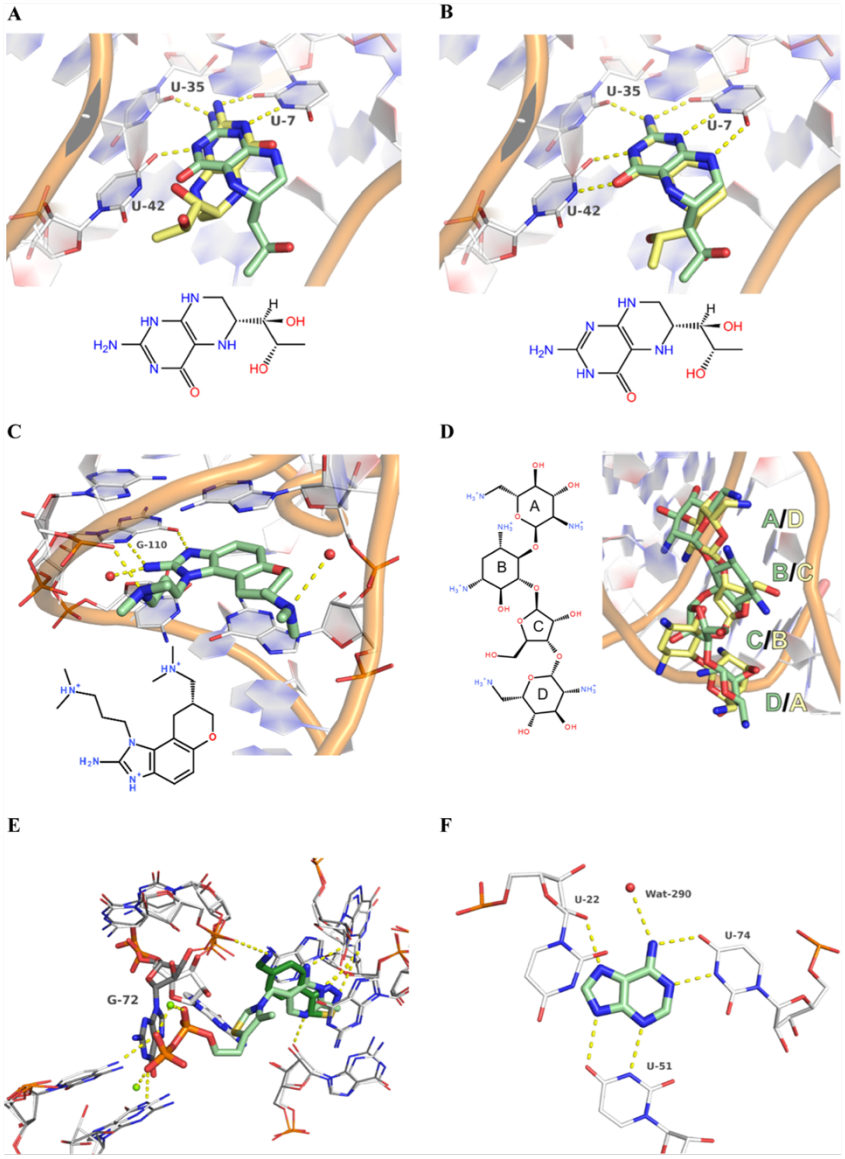


Figure 2. Identified challenges of RNA−ligand docking. RNA is depicted with white carbon atoms, crystallographic ligands with pale green carbon atoms, docking poses with light yellow carbon atoms, and polar interactions as yellow dashed lines. (A) Predicted binding mode of the tetrahydrobiopterin 1H-tautomer in the THF-RS (re-docking RMSD 4.0 Å, FlexX-score −28.3 kJ/mol, PDB-ID 4LVX). (B) Predicted binding mode of the tetrahydrobiopterin 3H-tautomer in the THF-RS (re-docking RMSD 1.6 Å, FlexX-score −39.4 kJ/mol, PDB-ID 4LVX). (C) Crystallographic binding mode of 2-aminobenzimidazole in its threefold protonated state forming two H-bonds as a donor to G-110 (PDB-ID 3TZR). (D) Predicted binding mode of neomycin in complex with HIV TAR (PDB-ID 1QD3) reveals an inverted binding mode of the docking pose compared to the NMR structure, however displaying highly similar interactions. (E) Superposition of the TPP-RS in complex with TPP (white RNA carbon atoms and pale green ligand carbon atoms, PDB-ID 2GDI) and a fragment ligand (gray RNA carbon atoms and dark green ligand carbon atoms, PDB-ID 4NYB). G-72 adopts an altered conformation depending on which ligand is bound. (F) Adenine binding to its RS (PDB-ID 1Y26). Inclusion of Wat-290 is crucial to improve binder-decoy discrimination with DOCK.

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

and 100% of the ligands within the top 2% by score of the docked data sets.

The results obtained with the nontemplate-based tools FRED, FlexX, and LeadIT were similar with only slightly reduced ROC AUCs and often reaching very good enrichments. Especially for the adenine-RS, SAM-RS, preQ$_1$-RS, and THF-RS, all tools could reach ROC AUCs over 0.7. Hence, it can be concluded that protein-based docking tools can generally be used against RNA targets.

■ COMMON PITFALLS AND SOLUTIONS

In favor of comparability between the different tools under elucidation, automated workflows for docking were established; however, a more refined and user-controlled procedure of receptor and ligand preparation and docking parameters should be considered as best practice.[105−109] While doing so for all 150 targets with all docking programs under elucidation is beyond the scope of this manuscript, some selected cases will be discussed to highlight common pitfalls and possible solutions when performing RNA−ligand docking. While the best scoring pose did not always reproduce the native binding mode, lower scoring poses with RMSD values <2.5 Å were regularly found for nearly all programs and targets, indicating that correct scoring of generated poses is more challenging than coverage of generated binding complex geometries. However, in some cases docking failure can be assigned to specific features of the docking setup.

Tautomers. For re-docking of tetrahydrobiopterin to the THF-RS (PDB-ID 4LVX),[103] the RMSD was 4.0 Å with a score of −28.3 kJ/mol using the automatic generated 1H tautomer with LeadIT (Figure 2A). Manually changing to the 3H tautomer (Figure 2B) improved the re-docking result of the top pose to an RMSD of 1.6 Å with H-bond complementarity to residues U-7, U-35, and U-42 as found in the crystal structure. The improved score of −39.4 kJ/mol further shows the ability of the scoring function to select the right tautomer over the wrong one. Similarly, the importance of tautomeric states was reported previously for the FMN-RS ligand ribocil (PDB-ID 5C45) for which the oxygen of the pyrimidinone tautomer is essential as an H-bond acceptor excluding the pyrimidinol tautomer.[36]

Protomers. For re-docking of a synthetic ligand to the hepatitis C virus internal ribosomal entry site (HCV IRES, PDB-ID: 3TZR),[110] re-docking with FRED resulted in an RMSD of 5.5 Å (score −11.2 kcal/mol). In this case, the ligand was protonated by Protoss[111] prior docking, resulting in a twofold protonated molecule with an uncharged 2-aminobenzimidazole moiety. However, two crucial H-bonds with G-110 of HCV IRES require the protonated state of this substructure (Figure 2C). With a predicted pK_a of 7.6 (predicted with MOE 2019)/8.4 (MarvinSketch-18.13.0, ChemAxon Ltd., http://www.chemaxon.com), the relative occurrence of the +3 protomer is about 61%/90%, respectively, at physiological pH. Docking of this protomer resulted in a correct pose with RMSD = 1.5 Å and a slightly better score (−12.0 kcal/mol). Interestingly, docking of the deprotonated ligand resulted in RMSD values between 0.92 and 2.14 Å for rDOCK, LeadIT, and FlexX (Table S1), indicating that the wrong protonation state can be compensated by other interactions and only FRED being sensitive to the correct protonation state. Also, different protonation states of multibasic ligands like aminoglycosides should be considered. While aliphatic amines are clearly basic and protonated under physiologic conditions if observed alone, the relative occurrence of protomers might be of importance in RNA−ligand recognition.[112] Exemplarily, the sixfold protonated state of neomycin is dominant at pH 7.4, but with only 22% (calculated with MOE)/43% (calculated with MarvinSketch, Figure 2D). Several minor species with lower formal charges are also possible under these conditions, and potential binding protomers of the target RNA or changes of protonation states might occur upon binding.

Beyond "Rule of Five" Ligands. Aminoglycosides hold further challenges by being large and polar, exceeding the (extended) Lipinski's rule of five (RO5) for drug-like, orally bioavailable molecules.[113] Additionally, aminoglycosides can have multiple binding modes to one RNA-binding site and further hold the potential of unspecific RNA binding.[114,115] Keeping in mind that molecular docking is often applied in the VS to identify lead structures, molecular docking programs are often trained with drug-like molecules. As within the re-docking set of 150 RNA−ligand complexes only 46 ligands do not violate the RO5, the programs are not only challenged by the target RNA, but also by many of the ligands. Especially for aminoglycosides like neomycin with a molecular weight of 621 g/mol, 13 H-bond donor and 13 H-bond acceptors, a tPSA of 363 Å^2, and 16 rotatable bonds, docking predictions are challenging. Considering the linear construction of four ring systems however, even high RMSD re-docking poses might be considered as no absolute failure. The roughly symmetric neomycin is inverted within the binding site, but aliphatic rings are fairly aligned and polar interactions reasonably saturated as observed for the LeadIT-predicted pose in complex with the human immunodeficiency virus *trans*-activation response element (HIV TAR) (Figure 2D, PDB-ID 1QD3).[116] While the RMSD for the re-docking pose is 9.5 Å, pseudo similarity (a metric within LeadIT to describe nearest atoms of same type instead of the identical atom ID) is 1.9 Å. While this predicted binding mode is still incorrect and no useful starting point for, for example, ligand optimization or derivatization, it can at least hint for interaction hot spots within the binding site to be addressed by polar, charged, or aliphatic moieties. Especially for HIV TAR a variety of different ligands with differing binding modes was reported further highlighting the combined challenge of complex ligands and flexible targets.[32,117,118]

Target Dynamics. Many RNA molecules are of highly dynamic nature.[54−56] Hence, docking against a single conformation might not be sufficient to represent RNA−ligand binding behavior. While re-docking was always performed against frame 1 of the NMR structures, this might not necessarily be the best frame selection method even though frame number 1 usually represents the lowest energy conformation.[51] Other strategies might be the selection of the frame closest to average coordinates or docking against an ensemble of frames either derived from NMR structures or MD simulations. When docking against all frames of 11 NMR complex structures for which multiple frames were deposited in the PDB, the mean re-docking RMSD was reduced from 6.6 to 3.7 Å (Table S2). It is noteworthy that for some complexes the ligand RMSD compared to average coordinates reached values up to 5.2 Å over the different NMR frames, highlighting the high flexibility even within the bound state of the complex. Additionally, it is to mention that docking against NMR structures is usually referred to be less precise and accurate compared to docking against X-ray structures.[51] In addition to

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

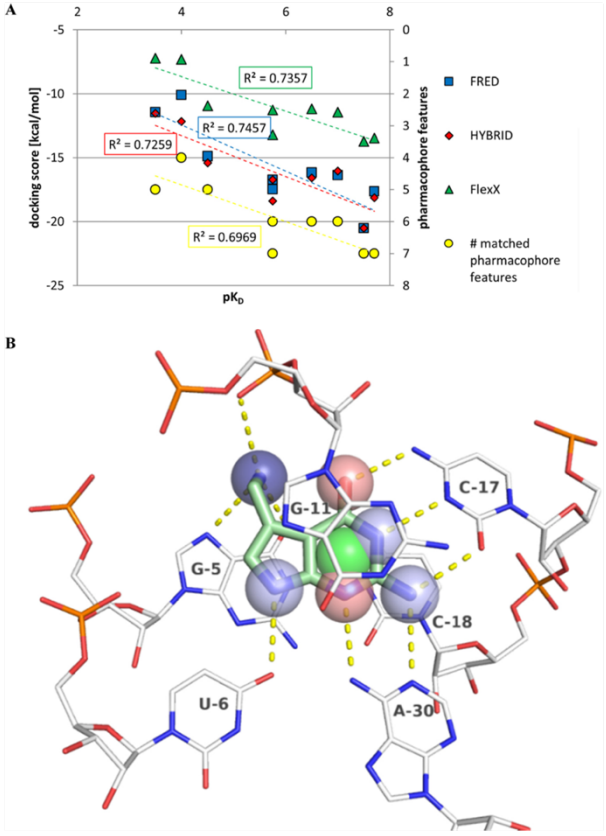


Figure 3. (A) Correlation plots between the docking score (left *y*-axis) and reported pK_D values for protein-based docking tools FRED (blue), HYBRID (red), and FlexX (green) against the preQ$_1$-RS and between the number of matched pharmacophoric features (right *y*-axis, yellow) and pK_D. (B) PreQ$_1$ (pale green carbon atoms) in complex with the preQ1-RS (white carbon atoms) including polar interactions (yellow dashed lines) and pharmacophore features described in the text: dark blue = H-bond donor or cation, light blue = H-bond donor, salmon red = H-bond acceptor, and green = aromatic center, PDB-ID 3FU2.

intrinsic dynamics, distinct conformational changes of residues, as reported for the TPP-RS, might need consideration. In the complex of several fragments with the TPP-RS (e.g., PDB-ID 4NYB), G-72 adopts a different orientation compared to the TPP- and analogue-bound state (PDB-ID 2GDI, Figure 2E).[100,119] To consider this conformational change, target flexibility should be taken into account during docking either by a protocol allowing side-chain flexibility[67] or a strategy with an ensemble of multiple RNA structures.[120,121]

Solvation. Finally, solvation and desolvation effects upon ligand binding and water-mediated interactions should be elucidated. While explicit solvation sites within protein binding pockets can have a huge impact on affinity and selectivity,[122,123] it is more than likely that similar effects account for RNA-binding sites as well. However, within the RNA−ligand docking test set presented, only 18 structures with a resolution below 2.0 Å allow sufficiently accurate determination of explicit water molecules. For the adenine-RS, a structurally important water molecule (PDB-ID 1Y26)[124] was reported previously (Figure 2F).[47] Including this water molecule in the docking setup of the binder-decoy discrimination of DOCK improved the ROC AUC from 0.47 to 0.66 (Table 2). Similarly, the enrichment for the TPP-RS was also improved by inclusion of crystallographic solvent molecules.

Further Considerations. While we identified the importance of the correct selection of ligand protomers and tautomers (Figure 2A−C), similarly rare protomers and tautomers of RNA might also play a role for ligand recognition in specific cases.[125,126] Additionally, it is to be taken into account that besides the four nucleobases, over 170 RNA modifications are reported,[127−129] which occur in biological systems and require individual tautomer and protomer considerations and potentially parameterization.[130]

Another aspect worth consideration is the functional role of ions. While a diffuse ion atmosphere masks the negative charge of the RNA backbone,[131] especially free and metabolite-bound Mg^{2+} ions play an important role in RNA folding, stabilization, and protection from degradation.[132,133] However, the identi-

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

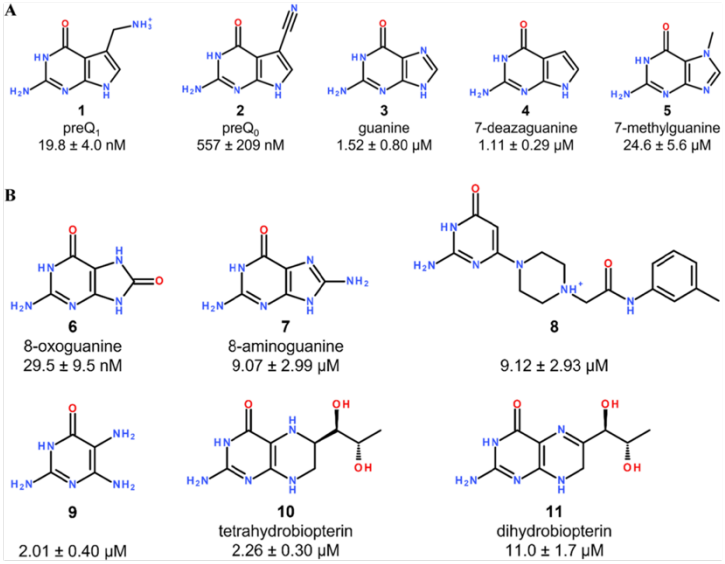


Figure 4. MST-derived K_D values for known preQ$_1$-RS ligands (A) and new ligands identified via VSs (B).

fication of explicit (Mg^{2+}) ion binding sites in RNA is not trivial as only high-resolution X-ray structures allow for accurate structural determination. Furthermore, because of the high salt concentrations in crystallography, ion types and binding sites might differ from native conditions. MD simulations can help to confirm cation binding sites, if structural information is available,[95] but cation binding site prediction for structures lacking this information (e.g., NMR structures or those where ions of crystallization buffers largely differ from physiological conditions) can be challenging as simulations might need rather long convergence times for divalent cations.[96] Even with the structural information of ions in the binding site available, binding strength between a certain cation and RNA might differ and a potential ligand can be required to either replace an ion or indirectly interact with the RNA via the ion.[134]

VS AGAINST THE PREQ$_1$-RIBOSWITCH

Because of their overall good performance, the fast protein-based docking programs FlexX, FRED, and HYBRID were selected for a prospective VS against the *Bacillus subtilis* preQ$_1$-RS. All three programs showed high accuracy in binding mode prediction with re-docking RMSD values of 0.8, 0.4, and 0.5 Å and binder-decoy discrimination with ROC AUCs of 0.94, 0.88, and 0.99, respectively (Table S1, Table 2). Furthermore, not only the discrimination of binders from nonbinders, but also the ranking within a series of nine reported ligands[78] spanning four orders of magnitude in potency, showed reasonable correlation between docking scores and reported pK_D values (Figure 3A, Table S3). Likewise, a correlation between the number of matched interactions resembling those of the native ligand preQ$_1$ with its RS (Figure 3B) and the pK_D was observed for the nine molecules as well. These interactions include three Watson−Crick-like H-bonds with C-17, additional H-bonds with U-6, and A-30 and possible H-bonds between the basic amine and G-5 and the phosphate backbone. Furthermore, preQ$_1$ stacks between G-11 and G-5 with face-to-face $\pi-\pi$-interactions. From this interaction profile, a pharmacophore model containing these seven features nicely correlating with affinity (Figure 3A,B) was derived.[80] For VSs, our in-house VS-library of 3.34 million compounds was filtered for physicochemical criteria (see Materials and Methods), and the remaining 12,507 compounds were docked without any constraints. Molecules for purchase were selected based on their docking scores and matching of the pharmacophoric features for at least one of the three programs, including molecules with (1) high scores and many met pharmacophore features, (2) high scores and little reflection of native ligands' interactions valuing the docking score over "MedChem intuition", and (3) low scoring molecules that resemble many pharmacophore features valuing human decision making over the scoring function. In this regard, special emphasis was placed on the identification of molecules which could mimic a Watson−Crick-like base pairing with C-17. The corresponding H-bond donor and acceptor functionalities were usually embedded in an heteroaromatic system, additionally allowing stacking between G-11 and G-5. Subsequently, 23 compounds were purchased for testing (Table S4). Because of previous filtering, the selected molecules were not flagged as *pan*-assay interference compounds (PAINS).[135] However, as a note of caution PAINS definition for RNA binding might differ from conventional protein-PAINS definition. For example, multi-basic *pan*-RNA binders spermine and spermidine[136] are also not flagged as PAINS and the concept of promiscuous binding might need further elucidation in the context of RNA ligands.

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

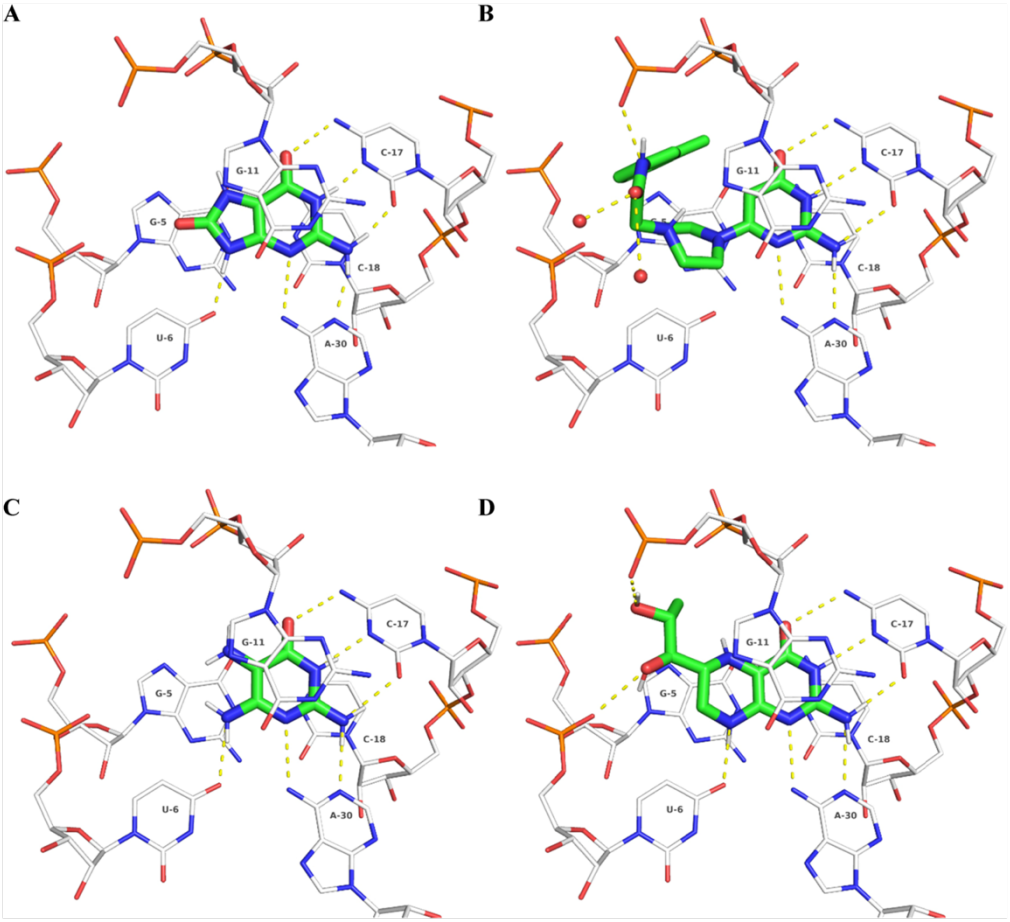


Figure 5. Predicted binding modes of **6** (A), **8** (B), **9** (C), and **10** (**D**) in complex with the preQ$_1$-RS aptamer (PDB-ID 3FU2). For clear view, only the predicted binding modes of the FlexX docking are shown. Ligands are depicted with green carbon atoms including polar hydrogens, RNA with white carbon atoms. H-bonds are shown as yellow dashed lines. H-bonds known from preQ$_1$ (Figure 3B) with C-17 and A-30 as well as aromatic stacking with G-11 are conserved among all ligands and one H-bond with U-6 for all except **8**.

■ EXPERIMENTAL RESULTS

PreQ$_1$-RS aptamer binding of the purchased compounds was elucidated by microscale thermophoresis (MST) using a Cy5-labeled RNA. The assay was validated by K_D-value determination of the reported[77,78] preQ$_1$-RS ligands preQ$_1$, preQ$_0$, guanine, 7-deazaguanine, and 7-methylguanine (**1**−**5**, Figure 4A) with K_D values similar to reported results (Table S3). For VS hits, a binding prescreening using three different concentrations of 100, 31.6, and 10 μM was performed. For 6 out of 23 tested compounds, **6**−**11**, which showed a concentration-dependent shift in thermophoresis, K_D values were subsequently determined (Figures 4B and S3).

Compounds **6**, **7**, **8**, **10**, and **11** matched 5−6 pharmacophore features of preQ$_1$ (Figure 3B) and additionally showed high docking scores in at least one of the three docking setups (FlexX, FRED, HYBRID, Table S4), while **9** matched 6 of 7 pharmacophore features with a comparatively low score. The conserved interactions according to the generated docking poses among all newly identified binders include the Watson−Crick like H-bonds with C-17, two H-bonds with A-30, and aromatic stacking with G-11 for all novel ligands (Figure 5A−D). Additionally, all ligands except compound **8** cover an H-bond donor to interact with U-6. Finally, none of the molecules with high docking scores and a low number of matched pharmacophore features turned out to bind to the preQ$_1$-RS (Table S4). Hence, the importance of proper pose inspection after docking is highlighted[137] even though the structural diversity and novelty are thus limited. This especially accounts for **6** (the strongest binder 8-oxoguanine) and **7** (8-aminoguanine) being highly similar to previously reported ligands containing the guanine substructure. The exceptional

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

high potency of **6** with a K_D of 29.5 nM compared to **2**−**4** and **7**, however, cannot be easily assigned to additional direct H-bonds to the RNA based on the predicted binding mode (Figure 5A). Potentially water-mediated interactions which require an H-bond acceptor at the 8-position are formed. This could also explain the much lower potency of the close homologue **7** with a K_D of 9.07 μM, which carries an H-bond donor at the corresponding position but obtains very similar docking scores. Structurally most distinct from preQ$_1$ and other guanine-derivatives, **8** still matches aromatic stacking interactions and the H-bond patterns with C-17 and A-30 but lacks the H-bond donor to U-6 according to docking poses predicted by FlexX and HYBRID. This compound however is positively charged which might compensate for the missing H-bond with U-6, and it might also form additional polar or hydrophobic/aromatic contacts (Figure 5B). **9** is the smallest binder identified in this VS and showed the lowest docking scores being ranked 1326th, 5948th, and 398th out of 12,507 for FlexX, FRED, and HYBRID, respectively. However, based on the predicted binding modes, this small molecule still matches 6 out of 7 pharmacophore features of preQ$_1$ (Figure 5C) in all docking setups, only lacking one H-bond donor/cationic moiety, resulting in a K_D of 2.01 μM. Finally, for tetra- and 7,8-dihydrobiopterin (**10** and **11**) binding with K_D values of 2.26 and 11.0 μM, respectively, a crucial role of the H-bond donor to interact with U-6 (Figure 5D) was indicated. For the aromatic pterins biopterin and neopterin (**12** and **13**, Table S4) which lack the hydrogen bond donor functionality at the corresponding position being replaced by a hydrogen bond acceptor, no binding was detected in the MST assay. This highlights the importance of this matched interaction, as the predicted binding modes of **12** and **13** had high scores and still matched 5 of the remaining pharmacophore features including the interactions with C-17, A-30, and aromatic stacking. Similarly, this was observed for 7-methylguanine (**5**) previously, which lacks this hydrogen bond donor and binds less strongly to the preQ$_1$-RS compared to **3** and **4**, which are capable of the H-bond interaction with U-6 (Figure 4A).[78]

■ CONCLUSIONS

Targeting RNAs with small molecules is an emerging field, but structure-based approaches are limited to very few cases because of a lower number of known RNA−ligand complexes compared to proteins. Hence, the knowledge basis for molecular docking is also restricted, and RNA holds its own challenges for docking-based VS. The basic physical principles of molecular recognition, however, do not differ between RNA and proteins, and molecular docking is applicable. Using a compilation of 150 RNA−ligand complexes (Figure 1, Table 1), the capability to reproduce native binding modes for protein-based docking tools without major modifications was similar to or even better than DOCK which was improved[52] and successfully used[47] for RNA−ligand docking previously and competitive to the RNA-specific rDOCK.[59] Additionally, the ability to discriminate known ligands from decoys for the rRNA A-site and several RSs was demonstrated (Table 2). The challenges faced for RNA−ligand docking include the selection of correct protomers and tautomers, target flexibility, and (de-) solvation effects, all of which are known for proteins as well, and are not RNA-specific (Figure 2). Subsequently, in a prospective VS with the preQ$_1$-RS aptamer, six new ligands with K_D values between 29.5 nM (**6**) and 11.0 μM (**11**) were discovered (Figure 4B). Based on the hit-selection criteria, it can be concluded that a good docking score only in combination with binding mode interpretation based on a pharmacophore hypothesis was superior to the docking score alone. As already identified in the initial validation set of ligands, it turned out that matching preQ$_1$'s pharmacophoric interaction features correlate well with affinity (Figure 3A,B, Table S3). This especially accounts for the Watson−Crick-like pairing with C-17 and hydrogen bonds to A-30 incorporated into a heteroaromatic framework. While the lack of a hydrogen bond donor to interact with U-6 in some cases like **8** can be tolerated, an acceptor-acceptor mismatch like in **12** and **13** (Table S4) abolishes binding or at least reduces affinity (**5**). Newly identified ligands **6**−**8**, **10**, and **11** nicely match the pharmacophore pattern and additionally show good docking scores. In contrast, no molecule with a high score and a low number of matched pharmacophore features was identified as a preQ$_1$-RS binder (Table S4). In lieu thereof, ligand **9** (K_D = 2.01 μM) matched 6 of 7 pharmacophore features with a relatively low score further highlighting the importance of proper pose inspection. Consequently, we conclude that RNA−ligand docking with conventional software is generally feasible but requires a certain extent of human intervention in both docking setup validation and hit selection. Hence, the strategy used herein may be used as a guideline to identify and circumvent potential pitfalls and support structure-based VSs against RNA targets in the future.

■ MATERIALS AND METHODS

Molecular Docking and Structure-Based VS. RNA−ligand complex structures were obtained from the PDB.[46] For pose prediction, PDB entries from previous publications[51,52,59] were combined and visually inspected using PyMOL.[138] If multiple chains of the target structure were present in the asymmetric unit of X-ray structures, chain A was selected. In case of NMR structures, frame 1 usually representing the lowest energy conformation[51] was used. If ions were present within the binding site, these ions were kept during docking. Except for LeadIT, which by default includes water molecules forming at least three interactions with the receptor and ligand, water molecules were removed from the complexes. All docking programs under elucidation treated the ligand to be flexible and the receptor to be rigid. If not stated differently, only the best scoring pose was further evaluated. Decoys to be discriminated from reported ligands (Table 2, Figures S5−S12) were generated using the DUD-E web service.[83]

For LeadIT-2.3.2 docking,[87] the graphical user interface was used for receptor preparation using standard settings and a binding site definition including residues within 6.5 Å around the reference ligand, which was protonated within LeadIT using the Protoss module.[111] The docking was performed under default settings using the enthalpy/entropy hybrid approach with 200 iterations per placement and fragmentation.

For FlexX[85,86] docking, the receptor and ligand were exported from LeadIT, and docking was performed without any further modifications from command-line.

For FRED/HYBRID docking,[88,89,139,140] the receptor was prepared using the Make_Receptor functionality of the OpenEye tools under default conditions. Ligand conformers for docking were generated using OMEGA[141,142] based on the Merck Molecular Force Filed (MMFF94).[143] For macrocyclic ligands, OMEGA with the keyword "macrocycle" was used. Docking with FRED and HYBRID was subsequently performed without deviations from default settings.

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

Receptor preparation for rDOCK-2013.1-1[59] was performed using the molecular operating environment (MOE 2015).[144] For rDOCK, a 6.0 Å binding site was defined around the reference ligand, and 50 poses were generated and subsequently scored using the SF3 and SF5 scoring functions.

For DOCK 6.9,[90,91] a protocol based on the tutorial of the developers was followed with only minor modifications. The receptor and ligand were protonated, and AM1-BCC ligand charges were applied within MOE 2015. To define the ligand binding site, a grid exceeding the ligand by 6.0 Å with a grid spacing of 0.3 Å was generated and receptor parameters were obtained from the AMBER99 force field. The extra margin of the receptor box was 5.0 Å. A second docking protocol was performed adding an additional charge to the phosphate backbone of the RNA to obtain a formal charge of zero per nucleotide to mimic counter ions masking the formal charge (phosphorous atom partial charge was changed from 1.1662 to 2.1662 as described previously).[51] Because of a poor performance under these settings in some cases of the binder-decoy discrimination set, explicit solvent molecules were included for docking against the adenine- and TPP-RS, and the binding site was reduced to 4.0 Å around the reference ligand for the guanine-RS.

For the prospective VS against the preQ$_1$-RS, PDB-ID 3FU2,[80] chain A was used. Docking with FlexX-4.1,[145] FRED/HYBRID-3.3.0.3 was performed as described above. Our in-house virtual molecule library of trusted vendors derived from ZINC15[146,147] contained 3.34 million molecules with at most 1 violation of Lipinski's RO5,[113] no reactive and no PAINS[135] moieties, a tPSA below 200 Å^2, and less than 11 rotatable bonds. This library was filtered after protonation with MOE 2019 using OpenEye FILTER to include only molecules without known aggregator functionalities, 4−9 H-bond donors, 2−10 H-bond acceptors, a molecular weight between 130 and 400 g/mol, xlogP between −3 and +3, tPSA over 50 Å^2, 0−8 rotatable bonds, a formal charge of 0−2, and less than 4 chiral centers to reduce chemical complexity. The remaining 12,507 molecules were subsequently energetically minimized using OMEGA-classic for FlexX docking, or conformers were generated by OMEGA-pose for FRED/HYBRID docking. After docking without constraints, poses were sorted by the docking score and elucidated by visual inspection and their accordance to the preQ$_1$ pharmacophore features (Figure 3B) within MOE 2019 prior selection for purchase (Table S4). Some structural analogues of promising scaffolds, including **7** (analogue of **6**), **10**, and **11** (analogues of **12**, **13**, and **29**, Table S4) not being part of the initial screening library, were docked and added for testing separately after the initial VS.

Microscale Thermophoresis. The HPLC-purified, Cy5-labeled, *Bs*preQ$_1$-RS aptamer, sequence 5′-Cy5-AGAGGUUCUAGCUACACCCUCUAUAAAAAACUAA-3′ was purchased from Eurofins genomics, Germany GmbH. MST experiments were performed on a Monolith Pico (NanoTemper Technologies, Munich, Germany) using standard uncoated capillaries. The RNA was diluted to 20 nM in 50 mM Tris−HCl buffer (pH 7.5) containing 100 mM KCl and 25 mM $MgCl_2$ similar to a protocol described previously.[77] The preQ$_1$-RS was heated to 75 °C for 5 min and cooled down to room temperature over 60 min. Ligands were added to final concentrations of 100, 31.6, and 10 μM in the initial screening. If a concentration-dependent shift of thermophoresis was observed, for K_D-determination concentrations ranging from 1.0 mM (if permitted by solubility, otherwise starting from 316 or 100 μM) to 0.1 nM in a 3.16-fold (half-logarithmic) dilution series and one DMSO negative control without ligand were prepared. Final DMSO concentration was always 2%. All measurements were performed at least as triplicates. Results were analyzed from signals after 1.5 s laser on-time with the MO.Affinity Analysis software, version 2.3 using the K_D-fit model. Identities and purities >90% for all compounds (Table S4) were stated in the vendor's certificates of analysis (CoA). For all newly discovered preQ$_1$-RS binders **6**−**11**, identities and purities >97% were confirmed by HPLC-ESI/MS (Figure S4). The mass spectra were obtained from a 1100 series HPLC system from Agilent with a Poroshell 120 EC-C18 150 × 2.10 mm, 4 μm column. The mobile phase consisted of 80% acetonitrile, 10% H_2O, and 10% of a 0.1% solution of formic acid in water. The UV detection wavelength was 254 nm. The molecular mass was detected using an Agilent 1100 series mass-selective detector trap with positive mode electron spray ionization. The LC−MS chromatograms and their corresponding mass spectra were analyzed using MestReNova (v.12.0.4) from Mestrelab Research.

Data and Software Availability. Molecular dockings were performed with LeadIT, FlexX, DOCK, FRED, HYBRID, and rDOCK using the cited versions. Additionally, FILTER, Make_receptor and OMEGA from the OE-toolkit, and MOE were used for the described steps in receptor and ligand preparation. Used PDB entries for re-docking, ligands, and decoys (derived from the DUD-E webservice) for discrimination, and SMILES of newly discovered preQ$_1$-RS ligands with their predicted binding modes are provided in the Supporting Information.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.2c00751.

Detailed results of re-docking and binder-decoy discrimination; SMILES, docking scores, MST-results, and LC−MS data of newly discovered ligands (XLSX)

SMILES of compounds **1**-**11** (XLSX)

Molecular structuresof binders and decoys (ZIP)

Docking poses of compounds **6**-**12** (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Christian Kersten − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz 55128, Germany*; orcid.org/0000-0001-9976-7639; Email: kerstec@uni-mainz.de

Authors

Elisabeth Kallert − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz 55128, Germany*

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

Simon Schneider − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz 55128, Germany*

Maike Grimm − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz 55128, Germany*

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

Mark Helm − *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz 55128, Germany;* orcid.org/0000-0002-0154-0928

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.2c00751

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Tanja Schirmeister for scientific discussion and proof-reading of the manuscript and Jessica Emsermann for support during elaboration of the MST assay. The MST instrument was provided by the Bundesministerium für Bildung und Forschung (BMBF) to M.H. (BMBF/01ED1804). The project was funded by the Internal University Research Funding (Stufe-I) of the Johannes Gutenberg-University. We further thank OpenEye Scientific and the Kuntz Lab (UCSF) for free academic licenses.

■ ABBREVIATIONS

A-site, aminoglycoside binding site; ASO, antisense oligonucleotide; AUC, area under the curve; *Bs*, *Bacillus subtilis*; CGO, chemical gaussian overlay; DUD-E, database of useful decoys enhanced; EMA, European Medicines Agency; ESI, electron spray ionization; FDA, Food and Drug Administration; FMN, flavin mononucleotide; GUI, graphical user interface; HCV, hepatitis C virus; IRES, internal ribosomal entry site; MD, molecular dynamic; miRNA, micro-RNA; MSD, mass-selective detector; ncRNA, noncoding RNA; PAINS, *pan*-assay interference compounds; PDB, protein data bank; preQ$_{1/0}$, prequeosine-1/0; RMSD, root-mean-square deviation; RO5, rule of five; ROC, receiver operating characteristic; RNA, ribosomal RNA; RS, riboswitch; SAM, *S*-adenosyl methionine; siRNA, small interfering RNA; TAR, trans-activation response element; THF, tetrahydrofolate; TPP, thiamine pyrophosphate; UTR, untranslated region; VS, virtual screening; WHO, World Health Organization

■ REFERENCES

(1) Rizvi, N. F.; Santa Maria, J. P.; Nahvi, A.; Klappenbach, J.; Klein, D. J.; Curran, P. J.; Richards, M. P.; Chamberlin, C.; Saradjian, P.; Burchard, J.; Aguilar, R.; Lee, J. T.; Dandliker, P. J.; Smith, G. F.; Kutchukian, P.; Nickbarg, E. B. Targeting RNA with Small Molecules: Identification of Selective, RNA-Binding Small Molecules Occupying Drug-Like Chemical Space. *SLAS Discovery* **2020**, *25*, 384−396.

(2) Rizvi, N. F.; Smith, G. F. RNA as a Small Molecule Druggable Target. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 5083−5088.

(3) Morgan, B. S.; Forte, J. E.; Culver, R. N.; Zhang, Y.; Hargrove, A. E. Discovery of Key Physicochemical, Structural, and Spatial Properties of RNA-Targeted Bioactive Ligands. *Angew. Chem., Int. Ed.* **2017**, *56*, 13498−13502.

(4) Hargrove, A. E. Small Molecule-RNA Targeting: Starting with the Fundamentals. *Chem. Commun.* **2020**, *56*, 14744−14756.

(5) Zafferani, M.; Hargrove, A. E. Small Molecule Targeting of Biologically Relevant RNA Tertiary and Quaternary Structures. *Cell Chem. Biol.* **2021**, *28*, 594−609.

(6) Sheridan, C. First Small-Molecule Drug Targeting RNA Gains Momentum. *Nat. Biotechnol.* **2021**, *39*, 6−8.

(7) Falese, J. P.; Donlic, A.; Hargrove, A. E. Targeting RNA with Small Molecules: From Fundamental Principles towards the Clinic. *Chem. Soc. Rev.* **2021**, *50*, 2224−2243.

(8) Di Giorgio, A.; Duca, M. Synthetic Small-Molecule RNA Ligands: Future Prospects as Therapeutic Agents. *MedChemComm* **2019**, *10*, 1242−1255.

(9) Connelly, C. M.; Moon, M. H.; Schneekloth, J. S. The Emerging Role of RNA as a Therapeutic Target for Small Molecules. *Cell Chem. Biol.* **2016**, *23*, 1077−1090.

(10) Winkle, M.; El-Daly, S. M.; Fabbri, M.; Calin, G. A. Noncoding RNA Therapeutics — Challenges and Potential Solutions. *Nat. Rev. Drug Discovery* **2021**, *20*, 629−651.

(11) Ratni, H.; Ebeling, M.; Baird, J.; Bendels, S.; Bylund, J.; Chen, K. S.; Denk, N.; Feng, Z.; Green, L.; Guerard, M.; Jablonski, P.; Jacobsen, B.; Khwaja, O.; Kletzl, H.; Ko, C. P.; Kustermann, S.; Marquet, A.; Metzger, F.; Mueller, B.; Naryshkin, N. A.; Paushkin, S. V.; Pinard, E.; Poirier, A.; Reutlinger, M.; Weetall, M.; Zeller, A.; Zhao, X.; Mueller, L. Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (SMN2) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA). *J. Med. Chem.* **2018**, *61*, 6501−6517.

(12) Ratni, H.; Scalco, R. S.; Stephan, A. H. Risdiplam, the First Approved Small Molecule Splicing Modifier Drug as a Blueprint for Future Transformative Medicines. *ACS Med. Chem. Lett.* **2021**, *12*, 874−877.

(13) Vo, D. D.; Tran, T. P. A.; Staedel, C.; Benhida, R.; Darfeuille, F.; Digiorgio, A.; Duca, M. Oncogenic MicroRNAs Biogenesis as a Drug Target: Structure-Activity Relationship Studies on New Aminoglycoside Conjugates. *Chem. − Eur. J.* **2016**, *22*, 5350−5362.

(14) Maucort, C.; Vo, D. D.; Aouad, S.; Charrat, C.; Azoulay, S.; Di Giorgio, A.; Duca, M. Design and Implementation of Synthetic RNA Binders for the Inhibition of MiR-21 Biogenesis. *ACS Med. Chem. Lett.* **2021**, *21*, 899.

(15) Vo, D. D.; Staedel, C.; Zehnacker, L.; Benhida, R.; Darfeuille, F.; Duca, M. Targeting the Production of Oncogenic MicroRNAs with Multimodal Synthetic Small Molecules. *ACS Chem. Biol.* **2014**, *9*, 711−721.

(16) Zhao, R.; Fu, J.; Zhu, L.; Chen, Y.; Liu, B. Designing Strategies of Small - Molecule Compounds for Modulating Non - Coding RNAs in Cancer Therapy. *J. Hematol. Oncol.* **2022**, *15*, 1−19.

(17) Khan, E.; Mishra, S. K.; Mishra, R.; Mishra, A.; Kumar, A. Discovery of a Potent Small Molecule Inhibiting Huntington's Disease (HD) Pathogenesis via Targeting CAG Repeats RNA and Poly Q Protein. *Sci. Rep.* **2019**, *9*, 1−15.

(18) Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. Dynamic Combinatorial Selection of Molecules Capable of Inhibiting the (CUG) Repeat RNA-MBNL1 Interaction in Vitro: Discovery of Lead Compounds Targeting Myotonic Dystrophy (DM1). *J. Am. Chem. Soc.* **2008**, *130*, 16254−16261.

(19) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. A Simple Ligand That Selectively Targets CUG Trinucleotide Repeats and Inhibits MBNL Protein Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 16068−16073.

(20) Parkesh, R.; Childs-Disney, J. L.; Nakamori, M.; Kumar, A.; Wang, E.; Wang, T.; Hoskins, J.; Tran, T.; Housman, D.; Thornton, C. A.; Disney, M. D. Design of a Bioactive Small Molecule That Targets the Myotonic Dystrophy Type 1 RNA via an RNA Motif-Ligand Database and Chemical Similarity Searching. *J. Am. Chem. Soc.* **2012**, *134*, 4731−4742.

(21) Rzuczek, S. G.; Southern, M. R.; Disney, M. D. Studying a Drug-like, RNA-Focused Small Molecule Library Identifies Compounds That Inhibit RNA Toxicity in Myotonic Dystrophy. *ACS Chem. Biol.* **2015**, *10*, 2706−2715.

(22) Angelbello, A.; Benhamou, R.; Rzuczek, S.; Choudhary, S.; Tang, Z.; Chen, J.; Roy, M.; Wang, K. W.; Yildirim, I.; Jun, A.; Thornton, C.; Disney, M. A Small Molecule That Binds an RNA Repeat Expansion Stimulates Its Decay via the Exosome Complex. *Cell Chem. Biol.* **2020**, 1−12.

(23) Wagner-Griffin, S.; Abe, M.; Benhamou, R. I.; Angelbello, A. J.; Vishnu, K.; Chen, J. L.; Childs-disney, J. L.; Disney, M. D. A Druglike Small Molecule That Targets r(CCUG) Repeats in Myotonic

Dystrophy Type 2 Facilitates Degradation by RNA Quality Control Pathways. *J. Med. Chem.* **2021**, *64*, 8474−8485.
(24) Donlic, A.; Swanson, E. G.; Chiu, L.-Y.; Wicks, S. L.; Juru, A. U.; Cai, Z.; Kassam, K.; Laudeman, C.; Sanaba, B. G.; Sugarman, A.; Han, E.; Tolbert, B. S.; Hargrove, A. E. R-BIND 2.0: An Updated Database of Bioactive RNA-Targeting Small Molecules and Associated RNA Secondary Structures. *ACS Chem. Biol.* **2022**, *17*, 1556−1566.
(25) Sabnis, R. W. Novel Substituted Thiadiazoles as Splicing Modulators for Modulating Splicing of MRNA and Uses Thereof for Treating Diseases. *ACS Med. Chem. Lett.* **2021**, *12*, 864−865.
(26) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like Small Molecules. *Nat. Rev. Drug Discovery* **2018**, *17*, 547−558.
(27) Birney, E.; Stamatoyannopoulos, J. A.; Dutta, A.; Guigó, R.; Gingeras, T. R.; Margulies, E. H.; Weng, Z.; Snyder, M.; Dermitzakis, E. T.; Thurman, R. E.; Kuehn, M. S.; Taylor, C. M.; Neph, S.; Koch, C. M.; Asthana, S.; Malhotra, A.; Adzhubei, I.; Greenbaum, J. A.; Andrews, R. M.; Flicek, P.; Boyle, P. J.; Cao, H.; Carter, N. P.; Clelland, G. K.; Davis, S.; Day, N.; Dhami, P.; Dillon, S. C.; Dorschner, M. O.; Fiegler, H.; Giresi, P. G.; Goldy, J.; Hawrylycz, M.; Haydock, A.; Humbert, R.; James, K. D.; Johnson, B. E.; Johnson, E. M.; Frum, T. T.; Rosenzweig, E. R.; Karnani, N.; Lee, K.; Lefebvre, G. C.; Navas, P. A.; Neri, F.; Parker, S. C. J.; Sabo, P. J.; Sandstrom, R.; Shafer, A.; Vetrie, D.; Weaver, M.; Wilcox, S.; Yu, M.; Collins, F. S.; Dekker, J.; Lieb, J. D.; Tullius, T. D.; Crawford, G. E.; Sunyaev, S.; Noble, W. S.; Dunham, I.; Denoeud, F.; Reymond, A.; Kapranov, P.; Rozowsky, J.; Zheng, D.; Castelo, R.; Frankish, A.; Harrow, J.; Ghosh, S.; Sandelin, A.; Hofacker, I. L.; Baertsch, R.; Keefe, D.; Dike, S.; Cheng, J.; Hirsch, H. A.; Sekinger, E. A.; Lagarde, J.; Abril, J. F.; Shahab, A.; Flamm, C.; Fried, C.; Hackermüller, J.; Hertel, J.; Lindemeyer, M.; Missal, K.; Tanzer, A.; Washietl, S.; Korbel, J.; Emanuelsson, O.; Pedersen, J. S.; Holroyd, N.; Taylor, R.; Swarbreck, D.; Matthews, N.; Dickson, M. C.; Thomas, D. J.; Weirauch, M. T.; Gilbert, J.; Drenkow, J.; Bell, I.; Zhao, X.; Srinivasan, K. G.; Sung, W. K.; Ooi, H. S.; Chiu, K. P.; Foissac, S.; Alioto, T.; Brent, M.; Pachter, L.; Tress, M. L.; Valencia, A.; Choo, S. W.; Choo, C. Y.; Ucla, C.; Manzano, C.; Wyss, C.; Cheung, E.; Clark, T. G.; Brown, J. B.; Ganesh, M.; Patel, S.; Tammana, H.; Chrast, J.; Henrichsen, C. N.; Kai, C.; Kawai, J.; Nagalakshmi, U.; Wu, J.; Lian, Z.; Lian, J.; Newburger, P.; Zhang, X.; Bickel, P.; Mattick, J. S.; Carninci, P.; Hayashizaki, Y.; Weissman, S.; Hubbard, T.; Myers, R. M.; Rogers, J.; Stadler, P. F.; Lowe, T. M.; Wei, C. L.; Ruan, Y.; Struhl, K.; Gerstein, M.; Antonarakis, S. E.; Fu, Y.; Green, E. D.; Karaöz, U.; Siepel, A.; Taylor, J.; Liefer, L. A.; Wetterstrand, K. A.; Good, P. J.; Feingold, E. A.; Guyer, M. S.; Cooper, G. M.; Asimenos, G.; Dewey, C. N.; Hou, M.; Nikolaev, S.; Montoya-Burgos, J. I.; Löytynoja, A.; Whelan, S.; Pardi, F.; Massingham, T.; Huang, H.; Zhang, N. R.; Holmes, I.; Mullikin, J. C.; Ureta-Vidal, A.; Paten, B.; Seringhaus, M.; Church, D.; Rosenbloom, K.; Kent, W. J.; Stone, E. A.; Batzoglou, S.; Goldman, N.; Hardison, R. C.; Haussler, D.; Miller, W.; Sidow, A.; Trinklein, N. D.; Zhang, Z. D.; Barrera, L.; Stuart, R.; King, D. C.; Ameur, A.; Enroth, S.; Bieda, M. C.; Kim, J.; Bhinge, A. A.; Jiang, N.; Liu, J.; Yao, F.; Vega, V. B.; Lee, C. W. H.; Ng, P.; Yang, A.; Moqtaderi, Z.; Zhu, Z.; Xu, X.; Squazzo, S.; Oberley, M. J.; Inman, D.; Singer, M. A.; Richmond, T. A.; Munn, K. J.; Rada-Iglesias, A.; Wallerman, O.; Komorowski, J.; Fowler, J. C.; Couttet, P.; Bruce, A. W.; Dovey, O. M.; Ellis, P. D.; Langford, C. F.; Nix, D. A.; Euskirchen, G.; Hartman, S.; Urban, A. E.; Kraus, P.; Van Calcar, S.; Heintzman, N.; Hoon Kim, T.; Wang, K.; Qu, C.; Hon, G.; Luna, R.; Glass, C. K.; Rosenfeld, M. G.; Aldred, S. F.; Cooper, S. J.; Halees, A.; Lin, J. M.; Shulha, H. P.; Zhang, X.; Xu, M.; Haidar, J. N. S.; Yu, Y.; Iyer, V. R.; Green, R. D.; Wadelius, C.; Farnham, P. J.; Ren, B.; Harte, R. A.; Hinrichs, A. S.; Trumbower, H.; Clawson, H.; Hillman-Jackson, J.; Zweig, A. S.; Smith, K.; Thakkapallayil, A.; Barber, G.; Kuhn, R. M.; Karolchik, D.; Armengol, L.; Bird, C. P.; De Bakker, P. I. W.; Kern, A. D.; Lopez-Bigas, N.; Martin, J. D.; Stranger, B. E.; Woodroffe, A.; Davydov, E.; Dimas, A.; Eyras, E.; Hallgrímsdóttir, I. B.; Huppert, J.; Zody, M. C.; Abecasis, G. R.; Estivill, X.; Bouffard, G. G.; Guan, X.; Hansen, N. F.; Idol, J. R.; Maduro, V. V. B.; Maskeri, B.; McDowell, J. C.; Park, M.; Thomas, P. J.; Young, A. C.; Blakesley, R. W.; Muzny, D. M.; Sodergren, E.; Wheeler, D. A.; Worley, K. C.; Jiang, H.; Weinstock, G. M.; Gibbs, R. A.; Graves, T.; Fulton, R.; Mardis, E. R.; Wilson, R. K.; Clamp, M.; Cuff, J.; Gnerre, S.; Jaffe, D. B.; Chang, J. L.; Lindblad-Toh, K.; Lander, E. S.; Koriabine, M.; Nefedov, M.; Osoegawa, K.; Yoshinaga, Y.; Zhu, B.; De Jong, P. J. Identification and Analysis of Functional Elements in 1% of the Human Genome by the ENCODE Pilot Project. *Nature* **2007**, *447*, 799−816.
(28) Hermann, T. Small Molecules Targeting Viral RNA. *Wiley Interdiscip. Rev.: RNA* **2016**, *7*, 726−743.
(29) Haniff, H. S.; Tong, Y.; Liu, X.; Chen, J. L.; Suresh, B. M.; Andrews, R. J.; Peterson, J. M.; O'Leary, C. A.; Benhamou, R. I.; Moss, W. N.; Disney, M. D. Targeting the SARS-COV-2 RNA Genome with Small Molecule Binders and Ribonuclease Targeting Chimera (RiboTAC) Degraders. *ACS Cent. Sci.* **2020**, *6*, 1713−1721.
(30) Di Giorgio, A.; Duca, M. New Chemical Modalities Enabling Specific RNA Targeting and Degradation: Application to SARS-CoV-2 RNA. *ACS Cent. Sci.* **2020**, *6*, 1647−1650.
(31) Davila-Calderon, J.; Patwardhan, N. N.; Chiu, L. Y.; Sugarman, A.; Cai, Z.; Penutmutchu, S. R.; Li, M. L.; Brewer, G.; Hargrove, A. E.; Tolbert, B. S. IRES-Targeting Small Molecule Inhibits Enterovirus 71 Replication via Allosteric Stabilization of a Ternary Complex. *Nat. Commun.* **2020**, *11*, 4775.
(32) Davis, B.; Afshar, M.; Varani, G.; Murchie, A. I. H.; Karn, J.; Lentzen, G.; Drysdale, M.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T.; Aboul-Ela, F. Rational Design of Inhibitors of HIV-1 TAR RNA through the Stabilisation of Electrostatic "Hot Spots.". *J. Mol. Biol.* **2004**, *336*, 343−356.
(33) Rekand, I. H.; Brenk, R. Ligand Design for Riboswitches, an Emerging Target Class for Novel Antibiotics. *Future Med. Chem.* **2017**, *9*, 1649−1663.
(34) Panchal, V.; Brenk, R. Riboswitches as Drug Targets for Antibiotics. *Antibiotics* **2021**, *10*, 45.
(35) Deigan, K. E.; Ferré-D'Amaré, A. R. Riboswitches: Discovery of Drugs That Target Bacterial Gene-Regulatory RNAs. *Acc. Chem. Res.* **2011**, *44*, 1329−1338.
(36) Howe, J. A.; Wang, H.; Fischmann, T. O.; Balibar, C. J.; Xiao, L.; Galgoci, A. M.; Malinverni, J. C.; Mayhood, T.; Villafania, A.; Nahvi, A.; Murgolo, N.; Barbieri, C. M.; Mann, P. A.; Carr, D.; Xia, E.; Zuck, P.; Riley, D.; Painter, R. E.; Walker, S. S.; Sherborne, B.; De Jesus, R.; Pan, W.; Plotkin, M. A.; Wu, J.; Rindgen, D.; Cummings, J.; Garlisi, C. G.; Zhang, R.; Sheth, P. R.; Gill, C. J.; Tang, H.; Roemer, T. Selective Small-Molecule Inhibition of an RNA Structural Element. *Nature* **2015**, *526*, 672−677.
(37) Zeller, M. J.; Nuthanakanti, A.; Li, K.; Aubé, J.; Serganov, A.; Weeks, K. M. Subsite Ligand Recognition and Cooperativity in the TPP Riboswitch: Implications for Fragment-Linking in RNA Ligand Discovery. *ACS Chem. Biol.* **2022**, *17*, 438−448.
(38) Sreeramulu, S.; Richter, C.; Berg, H.; Wirtz Martin, M. A.; Ceylan, B.; Matzel, T.; Adam, J.; Altincekic, N.; Azzaoui, K.; Bains, J. K.; Blommers, M. J. J.; Ferner, J.; Fürtig, B.; Göbel, M.; Grün, J. T.; Hengesbach, M.; Hohmann, K. F.; Hymon, D.; Knezic, B.; Martins, J.; Mertinkus, K. R.; Niesteruk, A.; Peter, S. A.; Pyper, D. J.; Qureshi, N. S.; Scheffer, U.; Schlundt, A.; Schnieders, R.; Stirnal, E.; Sudakov, A.; Tröster, A.; Vögele, J.; Wacker, A.; Weigand, J. E.; Wirmer-Bartoschek, J.; Wöhnert, J.; Schwalbe, H. Exploring the Druggability of Conserved RNA Regulatory Elements in the SARS-CoV-2 Genome. *Angew. Chem., Int. Ed.* **2021**, *60*, 19191−19200.
(39) Martin, W. J.; Grandi, P.; Marcia, M. Screening Strategies for Identifying RNA- and Ribonucleoprotein-Targeted Compounds. *Trends Pharmacol. Sci.* **2021**, *42*, 758−771.
(40) Vicens, Q.; Mondragón, E.; Reyes, F. E.; Coish, P.; Aristoff, P.; Berman, J.; Kaur, H.; Kells, K. W.; Wickens, P.; Wilson, J.; Gadwood, R. C.; Schostarez, H. J.; Suto, R. K.; Blount, K. F.; Batey, R. T. Structure-Activity Relationship of Flavin Analogues That Target the Flavin Mononucleotide Riboswitch. *ACS Chem. Biol.* **2018**, *13*, 2908−2919.

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

(41) Lim, J.; Winkler, W. C.; Nakamura, S.; Scott, V.; Breaker, R. R. Molecular-Recognition Characteristics of SAM-Binding Riboswitches. *Angew. Chem., Int. Ed.* **2006**, *45*, 964−968.
(42) Sztuba-Solinska, J.; Shenoy, S. R.; Gareiss, P.; Krumpe, L. R. H.; Le Grice, S. F. J.; O'Keefe, B. R.; Schneekloth, J. S. Identification of Biologically Active, HIV TAR RNA-Binding Small Molecules Using Small Molecule Microarrays. *J. Am. Chem. Soc.* **2014**, *136*, 8402−8410.
(43) Jordan, D.; Yang, M.; Schneekloth, J. S. Three-Color Imaging Enables Simultaneous Screening of Multiple RNA Targets on Small Molecule Microarrays. *Curr. Protoc. Chem. Biol.* **2020**, *12*, 1−11.
(44) Lundquist, K. P.; Panchal, V.; Gotfredsen, C. H.; Brenk, R.; Clausen, M. H. Fragment-based Drug Discovery for RNA Targets. *ChemMedChem* **2021**, *16*, 2588.
(45) Zeller, M. J.; Favorov, O.; Li, K.; Nuthanakanti, A.; Hussein, D.; Michaud, A.; Lafontaine, D. A.; Busan, S.; Serganov, A.; Aubé, J.; Weeks, K. M. SHAPE-Enabled Fragment-Based Ligand Discovery for RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, 1−11.
(46) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(47) Daldrop, P.; Reyes, F. E.; Robinson, D. A.; Hammond, C. M.; Lilley, D. M.; Batey, R. T.; Brenk, R. Novel Ligands for a Purine Riboswitch Discovered by RNA-Ligand Docking. *Chem. Biol.* **2011**, *18*, 324−335.
(48) Filikov, A. V.; Mohan, V.; Vickers, T. A.; Griffey, R. H.; Cook, P. D.; Abagyan, R. A.; James, T. L. Identification of Ligands for RNA Targets via Structure-Based Virtual Screening: HIV-1 TAR. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 593−610.
(49) Miranda, V. G.; Mollasalehi, N.; Gokhale, V.; Khanna, M. In Silico Targeting of the Long Noncoding RNA MALAT1. *ACS Med. Chem. Lett.* **2021**, *12*, 915−921.
(50) Li, Y.; Shen, J.; Sun, X.; Li, W.; Liu, G.; Tang, Y. Accuracy Assessment of Protein-Based Docking Programs against RNA Targets. *J. Chem. Inf. Model.* **2010**, *50*, 1134−1146.
(51) Detering, C.; Varani, G. Validation of Automated Docking Programs for Docking and Database Screening against RNA Drug Targets. *J. Med. Chem.* **2004**, *47*, 4188−4201.
(52) Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. DOCK 6: Combining Techniques to Model RNA-Small Molecule Complexes. *RNA* **2009**, *15*, 1219−1230.
(53) Chen, L.; Calin, G. A.; Zhang, S. Novel Insights of Structure-Based Modeling for RNA-Targeted Drug Discovery. *J. Chem. Inf. Model.* **2012**, *52*, 2741−2753.
(54) Luo, J.; Wei, W.; Waldispühl, J.; Moitessier, N. Challenges and Current Status of Computational Methods for Docking Small Molecules to Nucleic Acids. *Eur. J. Med. Chem.* **2019**, *168*, 414−425.
(55) Feng, Y.; Yan, Y.; He, J.; Tao, H.; Wu, Q.; Huang, S. Y. Docking and Scoring for Nucleic Acid−Ligand Interactions: Principles and Current Status. *Drug Discovery Today* **2022**, *27*, 838−847.
(56) Tessaro, F.; Tessaro, F.; Scapozza, L.; Scapozza, L. How "Protein-Docking" Translates into the New Emerging Field of Docking Small Molecules to Nucleic Acids? *Molecules* **2020**, *25*, 2749.
(57) Manigrasso, J.; Marcia, M.; De Vivo, M. Computer-Aided Design of RNA-Targeted Small Molecules: A Growing Need in Drug Discovery. *Chem* **2021**, *7*, 2965−2988.
(58) Morley, S. D.; Afshar, M. Validation of an Empirical RNA-Ligand Scoring Function for Fast Flexible Docking Using RiboDock®. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 189−208.
(59) Ruiz-Carmona, S.; Alvarez-Garcia, D.; Foloppe, N.; Garmendia-Doval, A. B.; Juhos, S.; Schmidtke, P.; Barril, X.; Hubbard, R. E.; Morley, S. D. RDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLoS Comput. Biol.* **2014**, *10*, 1−7.
(60) Pfeffer, P.; Gohlke, H. DrugScoreRNA - Knowledge-Based Scoring Function to Predict RNA - Ligand Interactions. *J. Chem. Inf. Model.* **2007**, *47*, 1868−1876.
(61) Krüger, D. M.; Bergs, J.; Kazemi, S.; Gohlke, H. Target Flexibility in RNA-Ligand Docking Modeled by Elastic Potential Grids. *ACS Med. Chem. Lett.* **2011**, *2*, 489−493.
(62) Philips, A.; Milanowska, K.; Ach, G. L.; Bujnicki, J. M. LigandRNA: Computational Predictor of RNA − Ligand Interactions. *RNA* **2013**, *19*, 1605−1616.
(63) Yan, Z.; Wang, J. SPA-LN: A Scoring Function of Ligand-Nucleic Acid Interactions via Optimizing Both Specificity and Affinity. *Nucleic Acids Res.* **2017**, *45*, 1−8.
(64) Sun, L. Z.; Jiang, Y.; Zhou, Y.; Chen, S. J. RLDOCK: A New Method for Predicting RNA-Ligand Interactions. *J. Chem. Theory Comput.* **2020**, *16*, 7173−7183.
(65) Feng, Y.; Huang, S. Y. ITScore-NL: An Iterative Knowledge-Based Scoring Function for Nucleic Acid-Ligand Interactions. *J. Chem. Inf. Model.* **2020**, *60*, 6698−6708.
(66) Feng, Y.; Zhang, K.; Wu, Q.; Huang, S. Y. NLDock: A Fast Nucleic Acid-Ligand Docking Algorithm for Modeling RNA/DNA-Ligand Complexes. *J. Chem. Inf. Model.* **2021**, *61*, 4771.
(67) Guilbert, C.; James, T. L. Docking to RNA via Root-Mean-Square-Deviation-Driven Energy Minimization with Flexible Ligands and Flexible Targets. *J. Chem. Inf. Model.* **2008**, *48*, 1257−1268.
(68) Gathiaka, S.; Liu, S.; Chiu, M.; Yang, H.; Stuckey, J. A.; Kang, Y. N.; Delproposto, J.; Kubish, G.; Dunbar, J. B.; Carlson, H. A.; Burley, S. K.; Walters, W. P.; Amaro, R. E.; Feher, V. A.; Gilson, M. K. D3R Grand Challenge 2015: Evaluation of Protein−Ligand Pose and Affinity Predictions. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 651−668.
(69) Gaieb, Z.; Liu, S.; Gathiaka, S.; Chiu, M.; Yang, H.; Shao, C.; Feher, V. A.; Walters, W. P.; Kuhn, B.; Rudolph, M. G.; Burley, S. K.; Gilson, M. K.; Amaro, R. E. D3R Grand Challenge 2: Blind Prediction of Protein−Ligand Poses, Affinity Rankings, and Relative Binding Free Energies. *J. Comput.-Aided Mol. Des.* **2018**, *32*, 1−20.
(70) Gaieb, Z.; Parks, C. D.; Chiu, M.; Yang, H.; Shao, C.; Walters, W. P.; Lambert, M. H.; Nevins, N.; Bembenek, S. D.; Ameriks, M. K.; Mirzadegan, T.; Burley, S. K.; Amaro, R. E.; Gilson, M. K. D3R Grand Challenge 3: Blind Prediction of Protein−Ligand Poses and Affinity Rankings. *J. Comput.-Aided Mol. Des.* **2019**, *33*, 1−18.
(71) Parks, C. D.; Gaieb, Z.; Chiu, M.; Yang, H.; Shao, C.; Walters, W. P.; Jansen, J. M.; McGaughey, G.; Lewis, R. A.; Bembenek, S. D.; Ameriks, M. K.; Mirzadegan, T.; Burley, S. K.; Amaro, R. E.; Gilson, M. K. D3R Grand Challenge 4: Blind Prediction of Protein−Ligand Poses, Affinity Rankings, and Relative Binding Free Energies. *J. Comput.-Aided Mol. Des.* **2020**, *34*, 99−119.
(72) Wagner, J. R.; Churas, C. P.; Liu, S.; Swift, R. V.; Chiu, M.; Shao, C.; Feher, V. A.; Burley, S. K.; Gilson, M. K.; Amaro, R. E. Continuous Evaluation of Ligand Protein Predictions: A Weekly Community Challenge for Drug Docking. *Structure* **2019**, *27*, 1326−1335.e4.
(73) Cruz, J. A.; Blanchet, M. F.; Boniecki, M.; Bujnicki, J. M.; Chen, S. J.; Cao, S.; Das, R.; Ding, F.; Dokholyan, N. V.; Flores, S. C.; Huang, L.; Lavender, C. A.; Lisi, V.; Major, F.; Mikolajczak, K.; Patel, D. J.; Philips, A.; Puton, T.; Santalucia, J.; Sijenyi, F.; Hermann, T.; Rother, K.; Rother, M.; Serganov, A.; Skorupski, M.; Soltysinski, T.; Sripakdeevong, P.; Tuszynska, I.; Weeks, K. M.; Waldsich, C.; Wildauer, M.; Leontis, N. B.; Westhof, E. RNA-Puzzles: A CASP-like Evaluation of RNA Three-Dimensional Structure Prediction. *RNA* **2012**, *18*, 610−625.
(74) Miao, Z.; Adamiak, R. W.; Blanchet, M. F.; Boniecki, M.; Bujnicki, J. M.; Chen, S. J.; Cheng, C.; Chojnowski, G.; Chou, F. C.; Cordero, P.; Cruz, J. A.; Ferré-D'Amaré, A. R.; Das, R.; Ding, F.; Dokholyan, N. V.; Dunin-Horkawicz, S.; Kladwang, W.; Krokhotin, A.; Lach, G.; Magnus, M.; Major, F.; Mann, T. H.; Masquida, B.; Matelska, D.; Meyer, M.; Peselis, A.; Popenda, M.; Purzycka, K. J.; Serganov, A.; Stasiewicz, J.; Szachniuk, M.; Tandon, A.; Tian, S.; Wang, J.; Xiao, Y.; Xu, X.; Zhang, J.; Zhao, P.; Zok, T.; Westhof, E. RNA-Puzzles Round II: Assessment of RNA Structure Prediction Programs Applied to Three Large RNA Structures. *RNA* **2015**, *21*, 1066−1084.
(75) Miao, Z.; Adamiak, R. W.; Antczak, M.; Batey, R. T.; Becka, A. J.; Biesiada, M.; Boniecki, M. J.; Bujnicki, J. M.; Chen, S.; Cheng, C. Y.

RNA-Puzzles Round III: 3D RNA Structure Prediction of Five Riboswitches and One Ribozyme. *RNA* **2017**, *23*, 655−672.
(76) Miao, Z.; Adamiak, R. W.; Antczak, M.; Boniecki, M. Ł. J.; Bujnicki, J.; Chen, S. J.; Cheng, C. Y. U.; Cheng, Y. I.; Chou, F. C.; Das, R.; Dokholyan, N. V.; Ding, F.; Geniesse, C.; Jiang, Y.; Joshi, A.; Krokhotin, A.; Magnus, M.; Mailhot, O.; Major, F.; Mann, T. H.; Piątkowski, P.; Pluta, R.; Popenda, M.; Sarzynska, J.; Sun, L.; Szachniuk, M.; Tian, S.; Wang, J. U. N. J.; Wang, J. U. N. J.; Watkins, A. M.; Wiedemann, J.; Xiao, Y. I.; Xu, X.; Yesselman, J. D.; Zhang, D.; Zhang, Y. I.; Zhang, Z.; Zhao, C.; Zhao, P.; Zhou, Y.; Zok, T.; Zyla, A.; Ren, A.; Batey, R. T.; Golden, B. L.; Huang, L.; Lilley, D. M.; Liu, Y.; Patel, D. J.; Westhof, E.; Tkowski, P. Ł. P. I. A.; Pluta, R.; Popenda, M.; Sarzynska, J.; Sun, L.; Szachniuk, M.; Tian, S.; Wang, J. U. N. J.; Wang, J. U. N. J.; Watkins, A. M.; Wiedemann, J.; Xiao, Y. I.; Xu, X.; Yesselman, J. D.; Zhang, D.; Zhang, Y. I.; Zhang, Z.; Zhao, C.; Zhao, P.; Zhou, Y. RNA-Puzzles Round IV: 3D Structure Predictions of Four Ribozymes and Two Aptamers. *RNA* **2020**, *26*, 982−995.
(77) Moon, M. H.; Hilimire, T. A.; Sanders, A. M.; Schneekloth, J. S. Measuring RNA-Ligand Interactions with Microscale Thermophoresis. *Biochemistry* **2018**, *57*, 4638−4643.
(78) Roth, A.; Winkler, W. C.; Regulski, E. E.; Lee, B. W. K.; Lim, J.; Jona, I.; Barrick, J. E.; Ritwik, A.; Kim, J. N.; Welz, R.; Iwata-Reuyl, D.; Breaker, R. R. A Riboswitch Selective for the Queuosine Precursor PreQ1 Contains an Unusually Small Aptamer Domain. *Nat. Struct. Mol. Biol.* **2007**, *14*, 308−317.
(79) Jenkins, J. L.; Krucinska, J.; McCarty, R. M.; Bandarian, V.; Wedekind, J. E. Comparison of a PreQ1 Riboswitch Aptamer in Metabolite-Bound and Free States with Implications for Gene Regulation. *J. Biol. Chem.* **2011**, *286*, 24626−24637.
(80) Klein, D. J.; Edwards, T. E.; Ferré-D'Amaré, A. R. Cocrystal Structure of a Class I PreQ1 Riboswitch Reveals a Pseudoknot Recognizing an Essential Hypermodified Nucleobase. *Nat. Struct. Mol. Biol.* **2009**, *16*, 343−344.
(81) Connelly, C. M.; Numata, T.; Boer, R. E.; Moon, M. H.; Sinniah, R. S.; Barchi, J. J.; Ferré-D'Amaré, A. R.; Schneekloth, J. S. Synthetic Ligands for PreQ 1 Riboswitches Provide Structural and Mechanistic Insights into Targeting RNA Tertiary Structure. *Nat. Commun.* **2019**, *10*, 1−12.
(82) Eichhorn, C. D.; Kang, M.; Feigon, J. Structure and Function of PreQ1 Riboswitches. *Biochim. Biophys. Acta, Gene Regul. Mech.* **2014**, *1839*, 939−950.
(83) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582−6594.
(84) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49*, 6789−6801.
(85) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.
(86) BioSolveIT. *FlexX. version 3.0*; BioSolveIT GmbH: St. Augustin, Germany, 2019.
(87) BioSolveIT. *LeadIT-2.3.2*; BioSolveIT GmbH: St. Augustin, Gernamy, 2017.
(88) McGann, M. FRED Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* **2011**, *51*, 578−596.
(89) McGann, M. FRED and HYBRID Docking Performance on Standardized Datasets. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 897−906.
(90) DesJarlais, R. L.; Seibel, G. L.; Kuntz, I. D.; Sheridan, R. P.; Venkataraghavan, R.; Scott Dixon, J. Using Shape Complementarity as an Initial Screen in Designing Ligands for a Receptor Binding Site of Known Three-Dimensional Structure. *J. Med. Chem.* **1988**, *31*, 722−729.
(91) Allen, W. J.; Balius, T. E.; Mukherjee, S.; Brozell, S. R.; Moustakas, D. T.; Lang, P. T.; Case, D. A.; Kuntz, I. D.; Rizzo, R. C. DOCK 6: Impact of New Features and Current Docking Performance. *J. Comput. Chem.* **2015**, *36*, 1132−1156.
(92) Daldrop, P.; Brenk, R. Structure-Based Virtual Screening for the Identification of RNA-Binding Ligands. *Methods Mol. Biol.* **2014**, *1103*, 127−139.
(93) Irwin, J. J.; Shoichet, B. K.; Mysinger, M. M.; Huang, N.; Colizzi, F.; Wassam, P.; Cao, Y. Automated Docking Screens: A Feasibility Study. *J. Med. Chem.* **2009**, *52*, 5712−5720.
(94) Hashem, Y.; Auffinger, P. A Short Guide for Molecular Dynamics Simulations of RNA Systems. *Methods* **2009**, *47*, 187−197.
(95) Fischer, N. M.; Polěto, M. D.; Steuer, J.; Van Der Spoel, D. Influence of Na + and Mg 2+ Ions on RNA Structures Studied with Molecular Dynamics Simulations. *Nucleic Acids Res.* **2018**, *46*, 4872−4882.
(96) Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurečka, P.; Walter, N. G.; Otyepka, M. RNA Structural Dynamics as Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* **2018**, *118*, 4177−4338.
(97) Thomas, J. R.; Hergenrother, P. J. Targeting RNA with Small Molecules. *Chem. Rev.* **2008**, *108*, 1171−1224.
(98) Mandal, M.; Breaker, R. R. Adenine Riboswitches and Gene Activation by Disruption of a Transcription Terminator. *Nat. Struct. Mol. Biol.* **2004**, *11*, 29−35.
(99) Edwards, T. E.; Ferré-D'Amaré, A. R. Crystal Structures of the Thi-Box Riboswitch Bound to Thiamine Pyrophosphate Analogs Reveal Adaptive RNA-Small Molecule Recognition. *Structure* **2006**, *14*, 1459−1468.
(100) Warner, K. D.; Homan, P.; Weeks, K. M.; Smith, A. G.; Abell, C.; Ferré-D'Amaré, A. R. Validating Fragment-Based Drug Discovery for Biological RNAs: Lead Fragments Bind and Remodel the TPP Riboswitch Specifically. *Chem. Biol.* **2014**, *21*, 591−595.
(101) Montange, R. K.; Mondragón, E.; van Tyne, D.; Garst, A. D.; Ceres, P.; Batey, R. T. Discrimination between Closely Related Cellular Metabolites by the SAM-I Riboswitch. *J. Mol. Biol.* **2010**, *396*, 761−772.
(102) Lee, E. R.; Blount, K. F.; Breaker, R. R. Roseoflavin Is a Natural Antibacterial Compound That Binds to FMN Riboswitches and Regulates Gene Expression. *RNA Biol.* **2009**, *6*, 187.
(103) Trausch, J. J.; Batey, R. T. A Disconnect between High-Affinity Binding and Efficient Regulation by Antifolates and Purines in the Tetrahydrofolate Riboswitch. *Chem. Biol.* **2014**, *21*, 205−216.
(104) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(105) Ban, F.; Dalal, K.; Li, H.; LeBlanc, E.; Rennie, P. S.; Cherkasov, A. Best Practices of Computer-Aided Drug Discovery (CADD): Lessons Learned from the Development of a Preclinical Candidate for Prostate Cancer with a New Mechanism of Action. *J. Chem. Inf. Model.* **2017**, *57*, 1018.
(106) Damm-Ganamet, K. L.; Arora, N.; Becart, S.; Edwards, J. P.; Lebsack, A. D.; McAllister, H. M.; Nelen, M. I.; Rao, N. L.; Westover, L.; Wiener, J. J. M.; Mirzadegan, T. Accelerating Lead Identification by High Throughput Virtual Screening: Prospective Case Studies from the Pharmaceutical Industry. *J. Chem. Inf. Model.* **2019**, *59*, 2046−2062.
(107) Bender, B. J.; Gahbauer, S.; Luttens, A.; Lyu, J.; Webb, C. M.; Stein, R. M.; Fink, E. A.; Balius, T. E.; Carlsson, J.; Irwin, J. J.; Shoichet, B. K. A Practical Guide to Large-Scale Docking. *Nat. Protoc.* **2021**, *16*, 4799−4832.
(108) Irwin, J. J. Docking Screens for Novel Ligands Conferring New Biology. *J. Med. Chem.* **2015**, 2, 4103−4120.
(109) Kolb, P.; Irwin, J. J. Docking Screens: Right for the Right Reasons? *Curr. Top. Med. Chem.* **2009**, *9*, 755−770.
(110) Dibrov, S. M.; Ding, K.; Brunn, N. D.; Parker, M. A.; Bergdahl, B. M.; Wyles, D. L.; Hermann, T. Structure of a Hepatitis C Virus RNA Domain in Complex with a Translation Inhibitor Reveals a Binding Mode Reminiscent of Riboswitches. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5223−5228.

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

(111) Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M. Protoss: A Holistic Approach to Predict Tautomers and Protonation States in Protein-Ligand Complexes. *Aust. J. Chem.* **2014**, *6*, 1−12.
(112) Freire, F.; Cuesta, I.; Corzana, F.; Revuelta, J.; González, C.; Hricovini, M.; Bastida, A.; Jiménez-Barbero, J.; Asensio, J. L. A Simple NMR Analysis of the Protonation Equilibrium That Accompanies Aminoglycoside Recognition: Dramatic Alterations in the Neomycin-B Protonation State upon Binding to a 23-Mer RNA Aptamer. *Chem. Commun.* **2007**, 174−176.
(113) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.
(114) François, B.; Russell, R. J. M.; Murray, J. B.; Aboul-ela, F.; Masquida, B.; Vicens, Q.; Westhof, E. Crystal Structures of Complexes between Aminoglycosides and Decoding A Site Oligonucleotides: Role of the Number of Rings and Positive Charges in the Specific Binding Leading to Miscoding. *Nucleic Acids Res.* **2005**, *33*, 5677−5690.
(115) Kulik, M.; Goral, A. M.; Jasiński, M.; Dominiak, P. M.; Trylska, J. Electrostatic Interactions in Aminoglycoside-RNA Complexes. *Biophys. J.* **2015**, *108*, 655−665.
(116) Faber, C.; Sticht, H.; Schweimer, K.; Rösch, P. Structural Rearrangements of HIV-1 Tat-Responsive RNA upon Binding of Neomycin B. *J. Biol. Chem.* **2000**, *275*, 20660−20666.
(117) Murchie, A. I. H.; Davis, B.; Isel, C.; Afshar, M.; Drysdale, M. J.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T. M.; Mirza, S.; Prescott, C. D.; Vaglio, P.; Aboul-ela, F.; Karn, J. Structure-Based Drug Design Targeting an Inactive RNA Conformation: Exploiting the Flexibility of HIV-1 TAR RNA. *J. Mol. Biol.* **2004**, *336*, 625−638.
(118) Du, Z.; Lind, K. E.; James, T. L. Structure of TAR RNA Complexed with a Tat-TAR Interaction Nanomolar Inhibitor That Was Identified by Computational Screening. *Chem. Biol.* **2002**, *9*, 707−712.
(119) Serganov, A.; Polonskaia, A.; Phan, A. T.; Breaker, R. R.; Patel, D. J. Structural Basis for Gene Regulation by a Thiamine Pyrophosphate-Sensing Riboswitch. *Nature* **2006**, *441*, 1167−1171.
(120) Campbell, A. J.; Lamb, M. L.; Joseph-McCarthy, D. Ensemble-Based Docking Using Biased Molecular Dynamics. *J. Chem. Inf. Model.* **2014**, *54*, 2127−2138.
(121) Ganser, L. R.; Lee, J.; Rangadurai, A.; Merriman, D. K.; Kelly, M. L.; Kansal, A. D.; Sathyamoorthy, B.; Al-Hashimi, H. M. High-Performance Virtual Screening by Targeting a High-Resolution RNA Dynamic Ensemble. *Nat. Struct. Mol. Biol.* **2018**, *25*, 425−434.
(122) Spyrakis, F.; Ahmed, M. H.; Bayden, A. S.; Cozzini, P.; Mozzarelli, A.; Kellogg, G. E. The Roles of Water in the Protein Matrix: A Largely Untapped Resource for Drug Discovery. *J. Med. Chem.* **2017**, *60*, 6781−6827.
(123) Kersten, C.; Fleischer, E.; Kehrein, J.; Borek, C.; Jaenicke, E.; Sotriffer, C.; Brenk, R. How to Design Selective Ligands for Highly Conserved Binding Sites: A Case Study Using N-Myristoyltransferases as a Model System. *J. Med. Chem.* **2020**, *63*, 2095−2113.
(124) Serganov, A.; Yuan, Y. R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A. T.; Hobartner, C.; Micura, R.; Breaker, R. R.; Patel, D. J. Structural Basis for Discriminative Regulation of Gene Expression by Adenine- and Guanine-Sensing MRNAs. *Chem. Biol.* **2004**, *11*, 1729−1741.
(125) Westhof, E. Isostericity and Tautomerism of Base Pairs in Nucleic Acids. *FEBS Lett.* **2014**, *588*, 2464−2469.
(126) Singh, V.; Fedeles, B. I.; Essigmann, J. M. Role of Tautomerism in RNA Biochemistry. *RNA* **2015**, *21*, 1−13.
(127) Dunin-Horkawicz, S.; Czerwoniec, A.; Gajda, M. J.; Feder, M.; Grosjean, H.; Bujnicki, J. M. MODOMICS: A Database of RNA Modification Pathways. *Nucleic Acids Res.* **2006**, *34*, 145−149.
(128) Boccaletto, P.; MacHnicka, M. A.; Purta, E.; Pitkowski, P.; Baginski, B.; Wirecki, T. K.; De Crécy-Lagard, V.; Ross, R.; Limbach, P. A.; Kotter, A.; Helm, M.; Bujnicki, J. M. MODOMICS: A Database of RNA Modification Pathways. 2017 Update. *Nucleic Acids Res.* **2018**, *46*, D303−D307.
(129) Boccaletto, P.; Stefaniak, F.; Ray, A.; Cappannini, A.; Mukherjee, S.; Purta, E.; Kurkowska, M.; Shirvanizadeh, N.; Destefanis, E.; Groza, P.; Avşar, G.; Romitelli, A.; Pir, P.; Dassi, E.; Conticello, S. G.; Aguilo, F.; Bujnicki, J. M. MODOMICS: A Database of RNA Modification Pathways. 2021 Update. *Nucleic Acids Res.* **2022**, *50*, D231−D235.
(130) Jones, E. L.; Mlotkowski, A. J.; Hebert, S. P.; Schlegel, H. B.; Chow, C. S. Calculations of PKa Values for a Series of Naturally Occurring Modified Nucleobases. *J. Phys. Chem. A* **2022**, *126*, 1518−1529.
(131) Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. Understanding Nucleic Acid-Ion Interactions. *Annu. Rev. Biochem.* **2014**, *83*, 813−841.
(132) Yamagami, R.; Sieg, J. P.; Bevilacqua, P. C. Functional Roles of Chelated Magnesium Ions in RNA Folding and Function. *Biochemistry* **2021**, *60*, 2374−2386.
(133) Stefan, L. R.; Zhang, R.; Levitan, A. G.; Hendrix, D. K.; Brenner, S. E.; Holbrook, S. R. MeRNA: A Database of Metal Ion Binding Sites in RNA Structures. *Nucleic Acids Res.* **2006**, *34*, 131−134.
(134) Jenner, L.; Starosta, A. L.; Terry, D. S.; Mikolajka, A.; Filonava, L.; Yusupov, M.; Blanchard, S. C.; Wilson, D. N.; Yusupova, G. Structural Basis for Potent Inhibitory Activity of the Antibiotic Tigecycline during Protein Synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 3812−3816.
(135) Baell, J.; Walters, M. A. Chemical Con Artists Foil Drug Discovery. *Nature* **2014**, *513*, 481−483.
(136) Lightfoot, H. L.; Hall, J. Endogenous Polyamine Function - The RNA Perspective. *Nucleic Acids Res.* **2014**, *42*, 11275−11290.
(137) Fischer, A.; Smieško, M.; Sellner, M.; Lill, M. A. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. *J. Med. Chem.* **2021**, *64*, 2489−2500.
(138) *The PyMOL Molecular Graphics System, Version 2.0*; Schrödinger, LLC.
(139) FRED 3.3.0.3. *OpenEye Sci. Software, St. Fe, NM.*
(140) HYBRID 3.3.0.3. *OpenEye Sci. Software, St. Fe, NM.*
(141) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50*, 572−584.
(142) OMEGA 3.1.0.3. *OpenEye Sci. Software, St. Fe, NM.*
(143) Halgren, T. A. Merck Molecular Force Field. *J. Comput. Chem.* **1996**, *17*, 490−519.
(144) MOE: Molecular Operating Environment (MOE). *Chem. Comput. Gr. ULC, 1010 Sherbooke St. West, Suite #910, Montr. QC, Canada, H3A 2R7, 2022* 2015.
(145) BiosolveIT. *FlexX. version 4.1*; BioSolveIT GmbH: St. Augustin, Germany, 2019.
(146) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757−1768.
(147) Sterling, T.; Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55*, 2324−2337.

https://doi.org/10.1021/acs.jcim.2c00751
J. Chem. Inf. Model. 2022, 62, 4134−4148

Protein-based Virtual Screening Tools applied for RNA-Ligand Docking identify new Binders of the $preQ_1$-Riboswitch

Elisabeth Kallert[1], Tim R. Fischer[1], Simon Schneider[1], Maike Grimm[1], Mark Helm[1], Christian Kersten[1,*]

[1] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Staudingerweg 5, 55128 Mainz, Germany.
*corresponding author: kerstec@uni-mainz.de

Content

Tables S1-S4
Figures S1-S12
References

Additional SI files:

SMILES.xlsx
docking-poses_6-12.zip (poses as SD files, receptor as PDB file)
binders+decoys.zip (SD files of ligands and decoys used for binder vs decoy discrimination)

Table S1. Summary of re-docking results of 150 RNA targets with LeadIT, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, FCharge: formal charge of dominant protomer.

Target				Ligand							
PDB-ID	source	resolution / no. of conformers submitted (calculated)	type	Weight	lip_don	lip_acc	SlogP	lip_violation	TPSA	opr_nrot	FCharge
1AKX	NMR	1 (20)	HIV-2 TRANS ACTIVATING REGION RNA COMPLEX WITH ARGININAMIDE, NMR, MINIMIZED AVERAGE STRUCTURE	160,2	8	5	-3,8	1	108	6	2
1AM0	NMR	8 (8)	AMP RNA APTAMER COMPLEX, NMR, 8 STRUCTURES	345,2	4	12	-4,1	1	180	5	-2
1BYJ	NMR	38 (38)	GENTAMICIN C1A A-SITE COMPLEX	454,6	17	12	-7,9	2	225	12	5
1EHT	NMR	10 (50)	THEOPHYLLINE-BINDING RNA IN COMPLEX WITH THEOPHYLLINE, NMR, 10 STRUCTURES	180,2	1	6	0,1	0	69,3	0	0
1EI2	NMR	17 (50)	STRUCTURAL BASIS FOR RECOGNITION OF THE RNA MAJOR GROOVE IN THE TAU EXON 10 SPLICING REGULATOR	620,7	25	19	-13,2	3	363	16	6
1F1T	X-Ray	2,8	CRYSTAL STRUCTURE OF THE MALACHITE GREEN APTAMER COMPLEXED WITH TETRAMETHYL-ROSAMINE	346,5	0	3	5,2	1	15,7	3	0
1F27	X-Ray	1,3	CRYSTAL STRUCTURE OF A BIOTIN-BINDING RNA PSEUDOKNOT	243,3	2	5	-0,5	0	81,3	5	-1
1FMN	NMR	5 (5)	SOLUTION STRUCTURE OF FMN-RNA APTAMER COMPLEX, NMR, 5 STRUCTURES	453,3	3	13	-2,5	1	183	7	-3
1FYP	NMR	29(100)	EUKARYOTIC DECODING REGION A-SITE RNA-PAROMOMYCIN COMPLEX	620,7	23	19	-12,4	3	355	16	5
1HNW	X-Ray	3,4	STRUCTURE OF THE THERMUS THERMOPHILUS 30S RIBOSOMAL SUBUNIT IN COMPLEX WITH TETRACYCLINE	443,4	5	10	-0,7	0	181	3	-1
1J7T	X-Ray	2,5	Complex between Paromomycin and the 16S-rRNA A-site at 2.5 A resolution	620,7	23	19	-12,4	3	355	16	5
1JZX	X-Ray	3,1	Structural Basis for the Interaction of Antibiotics with the Peptidyl Transferase Center in Eubacteria	426,0	5	7	-1,0	0	103	10	1
1K73	X-Ray	3,01	Co-crystal Structure of Anisomycin Bound to the 50S Ribosomal Subunit	267,3	2	5	0,4	0	67,8	5	0
1KC8	X-Ray	3,01	Co-crystal Structure of Blasticidin S Bound to the 50S Ribosomal Subunit	425,5	10	13	-5,6	2	220	11	1
1KOC	NMR	1 (1)	RNA APTAMER COMPLEXED WITH ARGININE, NMR	175,2	8	6	-5,2	1	131	6	1
1KOD	NMR	1 (1)	RNA APTAMER COMPLEXED WITH CITRULLINE, NMR	175,2	6	6	-3,2	1	123	5	0
1LC4	**X-Ray**	**2,54**	**Crystal Structure of Tobramycin Bound to the Eubacterial 16S rRNA A Site**	**472,6**	**20**	**14**	**-9,9**	**2**	**276**	**12**	**5**
1LVJ	NMR	12 (12)	STRUCTURE OF TAR RNA COMPLEXED WITH A TAT-TAR INTERACTION NANOMOLAR INHIBITOR THAT WAS IDENTI	327,5	1	3	3,0	0	24,8	5	1
1MWL	X-Ray	2,4	Crystal structure of geneticin bound to the eubacterial 16S rRNA A site	500,6	17	14	-8,8	3	258	12	4
1NBK	NMR	1 (1)	The structure of RNA aptamer for HIV Tat complexed with two argininamide molecules	175,2	8	6	-5,2	1	131	6	1
1NEM	NMR	9 (100)	Saccharide-RNA recognition in the neomycin B / RNA aptamer complex	620,7	25	19	-13,2	3	363	16	6
1NJI	X-Ray	3	Structure of chloramphenicol bound to the 50S ribosomal subunit	323,1	3	7	1,0	0	115	6	0
1NJM	X-Ray	3,6	The crystal structure of the 50S Large ribosomal subunit from Deinococcus radiodurans complexed with a tRNA acc	349,5	5	7	0,2	0	111	10	0
1NTA	X-Ray	2,9	2.9 A crystal structure of Streptomycin RNA-aptamer	584,6	19	19	-12,4	3	339	16	3
1NTB	X-Ray	2,9	2.9 A crystal structure of Streptomycin RNA-aptamer complex	586,6	20	19	-12,6	3	343	16	3
1O15	NMR	1 (250)	THEOPHYLLINE-BINDING RNA IN COMPLEX WITH THEOPHYLLINE, NMR, REGULARIZED MEAN STRUCTURE, REFINE	180,2	1	6	0,1	0	69,3	0	0
1P9X	X-Ray	3,4	THE CRYSTAL STRUCTURE OF THE 50S LARGE RIBOSOMAL SUBUNIT FROM DEINOCOCCUS RADIODURANS COMPLE	813,0	2	15	3,8	2	173	25	1
1PBR	NMR	1 (1)	STRUCTURE OF 16S RIBOSOMAL RNA, NMR, MINIMIZED AVERAGE STRUCTURE	620,7	23	19	-12,4	3	355	16	5
1Q8N	NMR	25 (100)	Solution Structure of the Malachite Green RNA Binding Aptamer	330,5	0	2	5,0	0	6,48	5	0
1QD3	NMR	17 (100)	HIV-1 TAR RNA/NEOMYCIN B COMPLEX	620,7	25	19	-13,2	3	363	16	6
1RAW	NMR	10 (100)	ATP BINDING RNA APTAMER IN COMPLEX WITH AMP, NMR, 10 STRUCTURES	345,2	4	12	-4,1	1	180	5	-2
1TOB	NMR	7 (50)	SACCHARIDE-RNA RECOGNITION IN AN AMINOGLYCOSIDE ANTIBIOTIC-RNA APTAMER COMPLEX, NMR, 7 STRUCTU	472,6	20	14	-9,9	2	276	12	5
1UTS	NMR	16 (50)	Designed HIV-1 TAR Binding Ligand	410,6	8	6	0,1	1	89,1	12	3
1UUD	NMR	1 (50)	NMR structure of a synthetic small molecule, rbt203, bound to HIV-1 TAR RNA	354,5	12	9	-4,5	1	162	14	3
1UUI	NMR	1 (50)	NMR structure of a synthetic small molecule, rbt158, bound to HIV-1 TAR RNA	323,4	7	7	-0,9	1	104	9	2
1VQ8	X-Ray	2,2	The structure of CCDA-PHE-CAP-BIO and the antibiotic sparsomycin bound to the large ribosomal subunit of haloarc	361,4	4	8	-0,8	0	144	8	0
1XBP	X-Ray	3,5	Inhibition of peptide bond formation by pleuromutilins: The structure of the 50S ribosomal subunit from Deinococcu	494,8	2	5	3,6	0	68	11	1
1XPF	X-Ray	2,3	HIV-1 subtype A genomic RNA Dimerization Initiation Site	206,4	10	4	-3,8	1	88,5	11	4
1Y26	**X-Ray**	**2,1**	**A-riboswitch-adenine complex**	**135,1**	**3**	**5**	**-0,1**	**0**	**80,5**	**0**	**0**
1YI2	X-Ray	2,62	Crystal Structure Of Erythromycin Bound To The G2099A Mutant 50S Ribosomal Subunit Of Haloarcula Marismortui	734,9	6	14	0,4	3	195	24	1
1YRJ	X-Ray	2,7	Crystal Structure of Apramycin bound to a Ribosomal RNA A site oligonucleotide	544,6	20	16	-10,8	3	295	12	5

Table S1. Summary of re-docking results of 150 RNA targets with LeadIT, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, FCharge: formal charge of dominant protomer.

	Re-docking RMSD [Å]								
PDB-ID	**LeadIT-2.3.2**	**FlexX-3**	**HYBRID 3.3.0.3**	**FRED 3.3.0.3**	**rDOCK (SF3)**	**rDOCK (SF5)**	**DOCK6**	**DOCK6 (mod. charge)**	**ligand smiles**
1AKX	3,57	4,35	2,03	4,43	4,79	3,39	4,64	3,69	O=CC([NH3+])CCCNC(=[NH2+])N
1AM0	3,86	4,06	1,43	1,37	1,58	1,85	1,32	1,49	P(=O)(OCC1C(O)C(O)C(O)C(n2c3ncnc(N)c3nc2)O1)([O-])[O-]
1BYJ	5,36	1,6	1,55	4,8	4,75	1,15	7,39	9,58	O(C1C(O)C(OC2C(O)C([NH2+]C)C(O)(C)CO2)C([NH3+])CC1[NH3+])C1C([NH3+])CCC(C[NH3+])O1
1EHT	3,69	3,68	3,62	4,09	6,58	6,57	8,62	8,91	O=C1N(C)C(=O)c2nc[nH]c2N1C
1EI2	8,1	7,31	6,96	6,74	5,97	6,88	3,61	8,65	O(C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
1F1T	0,82	1,26	6,15	6,15	0,7	0,63	0,98	0,29	N(C)(C)c1cc2Oc3c(C(C4C=CCC=C4)c2cc1)ccc(N(C)C)c3
1F27	1,3	5,22	1,92	6,6	8,99	9,19	10,16	10,80	O=C([O-])CCCCC1SCC2NC(=O)NC12
1FMN	2,1	3,79	2,09	3,18	2,63	2,6	1,95	2,15	P(=O)(OCC(O)C(O)C(O)CN1c2c(N=C3C(=O)[N-]C(=O)N=C13)cc(C)c(C)c2)([O-])[O-]
1FYP	3,71	8,29	2,2	3,65	3,06	10,64	5,77	3,83	O(C1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
1HNW	6,47	6,1	2,31	4,92	7,06	7,08	8,2	8,27	O=C(N)C1C(=O)C2(O)C([O-])=C3C(=O)c4c(O)cccc4C(O)(C)C3CC2C(N(C)C)C1=O
1J7T	1,29	4,81	1,03	10,4	2,37	1,17	9,97	9,16	O(C1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
1JZX	9,19	2,06	1,56	7,88	8,23	2,33	8,21	7,70	ClC(C(NC(=O)C1[NH+](C)CC(CCC)C1)C1C(O)C(O)C(O)C(SC)O1)C
1K73	6,52	8,82	1,41	1,27	1,52	3,85	1,52	1,62	O=C(OC1C(CC2C=CC(OC)C=C2)NCC1O)C
1KC8	5,49	3,67	1,83	1,86	2,41	1,76	8,39	11,64	O=C(NC1C(C(=O)[O-])OC(N2C(=O)N=C(N)C=C2)CC1)CC([NH3+])CCN(C(=[NH2+])N)C
1KOC	8,2		2,86	3,71	2,36	2,7	2,4	2,05	O=C([O-])C([NH3+])CCCNC(=[NH2+])N
1KOD	4,15	9,51	2,05	4,68	2,22	2,95	4,39	5,12	O=C(NCCCC([NH3+])C(=O)[O-])N
1LC4	**1,15**	**5,01**	**0,94**	**1,26**	**9,54**	**8,67**	**7,5**	**12,63**	**O(C1C(O)C(OC2C([NH3+])CC(O)C(C[NH3+])O2)C([NH3+])CC1[NH3+])C1C(O)C([NH3+])C(O)C(CO)O1**
1LVJ	9,17	7,53	5,31	7,59	6,88	4,31	6,08	5,75	O=C(C)c1cc2N(CCC[NH+](C)C)c3c(Sc2cc1)cccc3
1MWL	1,15	2,26	1,7	6,72	6,3	5,86	20,7	6,89	O(C1C(O)C(OC2C(O)C([NH2+]C)C(O)(C)CO2)C([NH3+])CC1[NH3+])C1C([NH3+])C(O)C(O)C(C(O)C)O1
1NBK	7,11	3,71	1,33	4,25	3,89	1,95	7,02	4,95	O=C([O-])C([NH3+])CCCNC(=[NH2+])N
1NEM	6,79	8,11	0,9	1,2	0,96	8,54	0,7	2,36	O(C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
1NJI	6,93	6,48	1,62	6,49	6,32	6,8		6,79	ClC(Cl)C(=O)NC(C(O)c1ccc([N+](=O)[O-])cc1)CO
1NJM	8,62	4,37	2,69	9,73	10,1	10,29	10,9	10,72	S(CSC)CC(NC(O)CCC=1C(=O)NC(=O)NC=1C)CO
1NTA	4,47		2,93	3,91	2,71	2,71	6,94	2,39	O=CC1(O)C(OC2C([NH2+]C)C(O)C(O)C(CO)O2)C(OC2C(O)C(O)C(NC(=[NH2+])N)C(O)C2NC(=[NH2+])N)OC1C
1NTB	5,34	1,83	1,69	1,74	1,2	2,26	7,2	0,86	O(C1C(O)(CO)C(C)OC1OC1C(O)C(O)C(NC(=[NH2+])N)C(O)C1NC(=[NH2+])N)C1C([NH2+]C)C(O)C(O)C(CO)O1
1O15	3,69	3,59	3,63	3,52	3,37	3,35	1,38	0,78	O=C1N(C)C(=O)c2nc[nH]c2N1C
1P9X	4,71	3,75	2,89	13,52	3,58	14,24	15,33	13,82	O(C)C1(C)C(OC2C(O)C([NH+](C)C)CC(C)O2)C(C)C(=O)C(C)C(=O)OC(CC)C2(C)OC(=O)N(CCCCn3cc(-c4cnccc4)nc3)C2C(C)C(=O)C(C)C1
1PBR	9,19	2,89	1,38	1,38	4,8	1,54	0,5	0,50	O(C1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
1Q8N	3,06	3,99	1,82	6,77	6,77	6,76	9,32	7,09	N(C)(C)c1ccc(C(c2ccc(N(C)C)cc2)c2ccccc2)cc1
1QD3	9,54	9,65	3,27	4,08	9,53	8,94	8,63	8,71	O(C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
1RAW	1,78	6,36	1,77	6,29	4,47	2,48	3,81	3,82	P(=O)(OCC1C(O)C(O)C(n2c3ncnc(N)c3nc2)O1)([O-])[O-]
1TOB	3,96	8,8	0,77	5,85	9,53	9,24	3,32	3,07	O(C1C(O)C(OC2C([NH3+])CC(O)C(C[NH3+])O2)C([NH3+])CC1[NH3+])C1C(O)C([NH3+])C(O)C(CO)O1
1UTS	10,19	7,47	1,33	10,13	9,66	10,31	10,15	10,45	O(CCC[NH3+])c1c(C[NH2+]CCN2CC[NH2+]CC2)cc(-c2cc3c([nH]cc3)cc2)cc1
1UUD	5,15	5,13	1,92	7,16	5,67	4,82	8,24	9,94	O(CCNC(=[NH2+])N)c1c(C[NH2+]CCCCNC(=[NH2+])N)cc(OC)cc1
1UUI	7,38	7,52	5,6	7,07	4,89	6,61	9,26	9,86	O(CCC[NH3+])c1c(CN2CC[N+](=C(N)N)CC2)cc(OC)cc1
1VQ8	1,1	1,21	1,26	1,29	8,39	4,03	9,7	7,73	[SH0](=O)(CSC)CC(NC(=O)C=CC=1C(=O)NC(=O)NC=1C)CO
1XBP	9,93	10,31	1,42	1,47	9,58	9,6	12,75	14,33	S(CC(=O)OC1C2(C)C(C)CCC3(C(C)C(O)C(C=C)(C)C1)C2C(=O)CC3)CC[NH+](CC)CC
1XPF	5,66	7,28	1,33	6,67	4,09	4,16	6,86	6,27	[NH3+]CCC[NH2+]CCCC[NH2+]CCC[NH3+]
1Y26	**0,56**	**0,58**	**0,28**	**0,24**	**0,59**	**0,61**	**0,33**	**0,43**	**Nc1ncnc2[nH]cnc12**
1YI2	7,14	7,69	2,55	10,85	2,51	1,39	10,83	9,96	O(C)C1(C)C(O)C(C)OC(OC2C(C)C(=O)OC(CC)C(O)(C)C(O)C(C)C(=O)C(C)CC(O)(C)C(OC3C(O)C([NH+](C)C)CC(C)O3)C2C)C1
1YRJ	12,83	10,7	0,83	2,93	1,7	2,19	9,59	9,67	O(C1C(O)C(O)C([NH3+])CC1[NH3+])C1C([NH3+])CC2OC(OC3C(O)C(O)C([NH3+])C(CO)O3)C([NH2+]C)C(O)C2O1

Table S1. Summary of re-docking results of 150 RNA targets with LeadIT, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, FCharge: formal charge of dominant protomer.

Target				Ligand							
PDB-ID	source	resolution / no. of conformers submitted (calculated)	type	Weight	lip_don	lip_acc	SlogP	lip_violation	TPSA	opr_nrot	FCharge
2AU4	NMR	1 (1)	Class I GTP aptamer	519,1	5	19	-8,1	2	282	9	-4
2B57	X-Ray	2,15	Guanine Riboswitch C74U mutant bound to 2,6-diaminopurine	150,1	5	6	-0,5	0	107	0	0
2BE0	X-Ray	2,63	Complex Between Paromomycin Derivative JS5-39 and the 16S-Rrna A-Site.	757,9	25	21	-12,5	3	373	22	6
2BEE	X-Ray	2,6	Complex Between Paromomycin derivative JS4 and the 16S-Rrna A Site	719,8	26	21	-13,4	3	387	22	6
2EES	X-Ray	1,75	Guanine riboswitch A21U, U75A mutant bound to hypoxanthine	136,1	2	5	-0,2	0	70,1	0	0
2EET	X-Ray	1,95	Guanine Riboswitch A21G, U75C mutant bound to hypoxanthine	136,1	2	5	-0,2	0	70,1	0	0
2ESJ	X-Ray	2,2	Complex between Lividomycin A and the 16S-Rrna A Site	766,8	25	23	-13,6	3	414	21	5
2ET3	X-Ray	2,8	Complex Between Gentamicin C1A and the 16S-RRNA A-Site	454,6	17	12	-7,9	2	225	12	5
2ET4	X-Ray	2,4	Complex Between Neomycin B and the 16S-RRNA A-Site	620,7	25	19	-13,2	3	363	16	6
2ET5	X-Ray	2,2	Complex Between Ribostamycin and the 16S-RRNA A-Site	458,5	18	14	-9,5	2	269	11	4
2ET8	X-Ray	2,5	Complex Between Neamine and the 16S-RRNA A-Site	326,4	16	10	-8,0	1	210	7	4
2F4S	X-Ray	2,8	A-site RNA in complex with neamine	326,4	16	10	-8,0	1	210	7	4
2F4T	X-Ray	3	Asite RNA + designer antibiotic	514,6	21	15	-10,3	3	288	17	5
2F4U	X-Ray	2,6	Asite RNA + designer antibiotic	529,7	20	15	-10,1	3	285	18	5
2FCX	X-Ray	2	HIV-1 DIS kissing-loop in complex with neamine	326,4	16	10	-8,0	1	210	7	4
2FCY	X-Ray	2,2	HIV-1 DIS kissing-loop in complex with Neomycin	620,7	25	19	-13,2	3	363	16	6
2FCZ	X-Ray	2,01	HIV-1 DIS kissing-loop in complex with ribostamycin	458,5	18	14	-9,5	2	269	11	4
2FD0	X-Ray	1,8	HIV-1 DIS kissing-loop in complex with lividomycin	766,8	25	23	-13,6	3	414	21	5
2GDI	**X-Ray**	**2,05**	**Crystal structure of thiamine pyrophosphate-specific riboswitch in complex with thiamine pyrophosphate**	**424,3**	**2**	**11**	**-2,7**	**1**	**159**	**8**	**-2**
2GIS	**X-Ray**	**2,9**	**Structure of the S-adenosylmethionine riboswitch mRNA regulatory element**	**399,5**	**7**	**11**	**-3,9**	**2**	**187**	**8**	**1**
2KMJ	NMR	10 (200)	High resolution NMR solution structure of a complex of HIV-2 TAR RNA and a synthetic tripeptide in a 1:2 stoichiom	512,7	19	15	-9,3	3	282	19	4
2L1V	NMR	20 (200)	Solution structure of a preQ1 riboswitch (Class I) aptamer bound to preQ1	180,2	7	6	-1,3	1	111	1	1
2L8H	NMR	10 (100)	Chemical probe bound to HIV TAR RNA	333,4	10	7	-1,4	1	133	9	2
2MIY	NMR	18 (200)	Solution NMR structure of a preQ1 Class II riboswitch from Streptococcus pneumoniae	180,2	7	6	-1,3	1	111	1	1
2O3V	X-Ray	2,8	Crystal Structure of the Homo sapiens Cytoplasmic Ribosomal Decoding Site complexed with paromamine derivativ	518,6	19	16	-10,2	3	298	14	4
2O3W	X-Ray	2,8	Crystal Structure of the Homo sapiens Cytoplasmic Ribosomal Decoding Site in presence of paromomycin	620,7	23	19	-12,4	3	355	16	5
2O3X	X-Ray	2,9	Crystal Structure of the Prokaryotic Ribosomal Decoding Site Complexed with Paromamine Derivative NB30	458,5	18	14	-9,5	2	269	11	4
2OE5	X-Ray	1,51	1.5 A X-ray crystal structure of Apramycin complex with RNA fragment GGCGUCGCUAGUACCG/GGUACUAAAAGUC	544,6	20	16	-10,8	3	295	12	5
2OE8	X-Ray	1,8	1.8 A X-ray crystal structure of Apramycin complex with RNA fragment GGGCGUCGCUAGUACC/CGGUACUAAAAGU	544,6	20	16	-10,8	3	295	12	5
2PWT	X-Ray	1,8	Crystal structure of the bacterial ribosomal decoding site complexed with aminoglycoside containing the L-HABA gr	868,0	25	23	-12,5	3	407	27	6
2QWY	X-Ray	2,8	SAM-II riboswitch bound to S-adenosylmethionine	399,5	7	11	-3,9	2	187	8	1
2TOB	NMR	13 (40)	SOLUTION STRUCTURE OF THE TOBRAMYCIN-RNA APTAMER COMPLEX, NMR, 13 STRUCTURES	472,6	20	14	-9,9	2	276	12	5
2Z75	X-Ray	1,7	T. tengcongensis glmS ribozyme bound to glucosamine-6-phosphate	258,1	6	9	-6,2	1	158	5	-1
3C44	X-Ray	2	Crystal structure of HIV-1 subtype F DIS extended duplex RNA bound to paromomycin	620,7	23	19	-12,4	3	355	16	5
3CPW	X-Ray	2,7	The structure of the antibiotic LINEZOLID bound to the large ribosomal subunit of HALOARCULA MARISMORTUI	340,4	1	7	0,4	0	70,9	6	1
3D0U	X-Ray	2,8	Crystal Structure of Lysine Riboswitch Bound to Lysine	147,2	6	4	-3,2	1	95,4	5	1
3D2X	X-Ray	2,5	Structure of the thiamine pyrophosphate-specific riboswitch bound to oxythiamine pyrophosphate	423,3	1	11	-3,2	1	149	8	-2
3DIG	X-Ray	2,8	CRYSTAL STRUCTURE OF THE THERMOTOGA MARITIMA LYSINE RIBOSWITCH BOUND TO S-(2-aminoethyl)-L-cystei	165,2	6	4	-3,7	1	95,4	5	1
3DIL	X-Ray	1,9	Crystal structure of the Thermotoga maritima lysine riboswitch bound to lysine	147,2	6	4	-3,2	1	95,4	5	1
3DIM	X-Ray	2,9	Crystallization of the Thermotoga maritima lysine riboswitch bound to lysine, Cs+ Soak	147,2	6	4	-3,2	1	95,4	5	1
3DIO	X-Ray	2,4	Crystallization of the Thermotoga maritima lysine riboswitch bound to lysine, IRIDIUM HEXAMINE SOAK	147,2	6	4	-3,2	1	95,4	5	1
3DIQ	X-Ray	2,7	Crystallization of the Thermotoga maritima lysine riboswitch bound to homoarginine	189,2	8	6	-4,8	1	131	7	1
3DIR	X-Ray	2,9	Crystallization of the Thermotoga maritima lysine riboswitch bound to N6-1-iminoethyl-L-Lysine	188,3	6	5	-3,7	1	105	7	1
3DIY	X-Ray	2,71	Crystallization of the Thermotoga maritima lysine riboswitch bound to lysine, Mn2+ soak	147,2	6	4	-3,2	1	95,4	5	1
3DIZ	X-Ray	2,85	Crystallization of the Thermotoga maritima lysine riboswitch bound to lysine in the Absence of Mg2+	147,2	6	4	-3,2	1	95,4	5	1

Table S1. Summary of re-docking results of 150 RNA targets with LeadIT, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, FCharge: formal charge of dominant protomer.

PDB-ID	LeadIT-2.3.2	FlexX-3	HYBRID 3.3.0.3	FRED 3.3.0.3	rDOCK (SF3)	rDOCK (SF5)	DOCK6	DOCK6 (mod. charge)	ligand smiles
2AU4	2,46	2,38	1,44	3,1	2,87	3,77	5,77	5,88	P(=O)(OP(=O)(OCC1C(O)C(O)C(n2c3N=C(N)NC(=O)c3nc2)O1)[O-])(OP(=O)([O-])[O-])[O-]
2B57	0,35	0,27	0,23	0,37	0,21	0,22	0,11	0,16	Nc1nc(N)c2nc[nH]c2n1
2BE0	5,72	6,94	1,34	11,04	9,13	6,04	10,97	10,94	O(CC[NH2+]CC1=CNC=CC1)C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)OC1OC1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O
2BEE	5,56	7,2	2,58	9,83	6,26	8,74	8,37	8,26	O(CC[NH2+]CC=CN)C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)OC1OC1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O
2EES	0,31	3,43	0,44	0,39	2,46	0,27	3,45	0,20	O=C1NC=Nc2[nH]cnc12
2EET	0,32	2,43	0,27	0,55	0,31	0,26	3,41	0,30	O=C1NC=Nc2[nH]cnc12
2ESJ	4,74	8,45	1,45	10,21	7,65	7,26	9,49	8,41	O(C1C(OC2C([NH3+])CC(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(OC3C(O)C(O)C(O)C(CO)O3)C([NH3+])O2)C(CO)O1
2ET3	1,02	4,08	1,29	1,65	7,26	9,3	10,03	11,71	O(C1C(O)C(OC2C(O)C([NH2+]C)C(O)C)CO2)C([NH3+])CC1[NH3+])C1C([NH3+])CCC(C[NH3+])O1
2ET4	1,44	1,88	0,95	3,34	1,3	1,6	4,52	10,00	O(C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
2ET5	2,38	8,67	0,81	0,75	8,02	0,62		7,52	O(C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(O)C(CO)O1
2ET8	0,68	1,54	0,33	0,58	6,89	6,69	5,41	0,81	O(C1C(O)C(O)C([NH3+])CC1[NH3+])C1C([NH3+])C(O)C(O)C(C[NH3+])O1
2F4S	1,01	2,61	0,63	0,72	0,87	1,26	0,9	0,88	O(C1C(O)C(O)C([NH3+])CC1[NH3+])C1C([NH3+])C(O)C(O)C(C[NH3+])O1
2F4T	1,73		0,99	1,97	9,66	9,03	10,77	0,92	O=C(NC1C(OCCNCC[NH3+])C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])C1)C(O)CC[NH3+]
2F4U	2,05	1,66	1,2	2,02	9,01	8,6	10,98	1,28	O=C(NC1C(OCC[NH2+]CCC[NH3+])C(O)C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])C1)C(O)CC[NH3+]
2FCX	8,18	7,36	1,35	8,15	6,54	6,17	7,98	7,10	O(C1C(O)C(O)C([NH3+])CC1[NH3+])C1C([NH3+])C(O)C(O)C(C[NH3+])O1
2FCY	2,74	4,94	1,57	3,08	5,14	9,35	9,13	9,53	O(C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
2FCZ	3,27	6,79	1,23	6,23	0,54	6,26	9,46	0,86	O(C1C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(O)C(CO)O1
2FD0	6,5	8,59	2,45	6,96	2,82	0,86	12	12,23	O(C1C(OC2C([NH3+])CC(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(OC3C(O)C(O)C(O)C(CO)O3)C(C[NH3+])O2)C(CO)O1
2GDI	**1,48**	**0,43**	**1,85**	**1,79**	**1,51**	**1,91**	**6,57**	**9,75**	**P(=O)(OP(=O)([O-])[O-])(OCCC1C(C)[N+](Cc2c(N)nc(C)nc2)=CS1)[O-]**
2GIS	**0,83**	**1,67**	**1,28**	**4,22**	**4,3**	**2,76**	**5,32**	**5,28**	**[S+](CCC([NH3+])C(=O)[O-])(CC1C(O)C(O)C(n2c3ncnc(N)c3nc2)O1)C**
2KMJ	9,4	8,53	2,03	8,5	9,3	7,33	13,12	12,38	O=C(NC(C(=O)NC(C(=O)N)CCCNC(=[NH2+])N)CCCC1=[NH+]CC=CN1)C([NH3+])CCCNC(=[NH2+])N
2L1V	0,43	0,53	0,49	5,17	0,51	0,5	4,72	4,64	O=C1NC(N)=Nc2[nH]cc(C[NH3+])c12
2L8H		10,23	3,74	5,6	5,25	4,84	7,5	4,67	O(C)c1c2c(cc(NC(O)C([NH3+])CCCNC(=[NH2+])N)c1)cccc2
2MIY	0,81	0,83	0,97	3,61	6,1	4,61	6,4	7,01	O=C1NC(N)=Nc2[nH]cc(C[NH3+])c12
2O3V	11,14	10,62	2,31	8,4	5,52	7,27		9,66	O(COC1C(O)C(CO)OC(O)C1[NH3+])C1C(O)C(CO)OC(OC2C(O)C(O)C([NH3+])CC2[NH3+])C1[NH3+]
2O3W	9,32	2,22	1,21	6,52	8,72	9,88	9,01	8,38	O(C1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
2O3X	0,87	2,22	1,21	6,52	8,68	1,1	8,82	11,98	O(C1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(O)C(C[NH3+])O1
2OE5	3,7	6,24	0,71	5,23	1,28	3,05	2,54	0,40	O(C1C(O)C(O)C([NH3+])CC1[NH3+])C1C([NH3+])CC2OC(OC3C(O)C(O)C([NH3+])C(CO)O3)C([NH2+]C)C(O)C2O1
2OE8	0,63	9,81	0,74	2,84	1,85	2,36	2,4	0,62	O(C1C(O)C(O)C([NH3+])CC1[NH3+])C1C([NH3+])CC2OC(OC3C(O)C(O)C([NH3+])C(CO)O3)C([NH2+]C)C(O)C2O1
2PWT	6,17	4,15	2,21	3,39	11,14	9,88	10,7	11,91	O=C(NC1C(O)C(OC2C(OCC[NH2+]CCc3ccccc3)C(OC3C([NH3+])C(O)C(O)C(C[NH3+])O3)C(CO)O2)C(OC2C([NH3+])C(O)C(O)C(C=O)O2)C([NH3+])C1)C(O)CC[NH3+]
2QWY	1,3	3,13	4,55	5,84	4,13	1,26	0,9	0,63	[S+](CCC([NH3+])C(=O)[O-])(CC1C(O)C(O)C(n2c3ncnc(N)c3nc2)O1)C
2TOB	10,81	10,92	0,87	2,15	1,52	1,38	2,22	1,80	O(C1C(O)C(OC2C([NH3+])CC(O)C(C[NH3+])O2)C([NH3+])CC1[NH3+])C1C(O)C([NH3+])C(O)C(CO)O1
2Z75	1,31	0,85	2,72	1,7	7,3	7,21	6,4	7,16	P(=O)(OCC1C(O)C(O)C([NH3+])C(O)O1)([O-])[O-]
3C44	1,46	4,89	0,67	3,38	0,73	1,17	7,86	8,48	O(C1C(OC2C([NH3+])C(O)C(O)C(CO)O2)C([NH3+])CC([NH3+])C1O)C1C(O)C(OC2C([NH3+])C(O)C(O)C(C[NH3+])O2)C(CO)O1
3CPW	0,69	0,69	1,51	1,5	3,1	8,64	1,6	1,10	FC1C(=[N+]2CCOCC2)CCC(N2C(=O)OC(CNC(=O)C)C2)=C1
3D0U	0,4	0,62	0,52	0,52	0,58	0,96	0,87	1,17	O=C([O-])C([NH3+])CCCC[NH3+]
3D2X	1,76	1,24	2,66	1,25	1,25	2,51	7,03	1,59	P(=O)(OP(=O)([O-])[O-])(OCCc1c(C)[n+](CC=2C(=O)N=C(C)NC=2)cs1)[O-]
3DIG	1,43	12,09	0,71	1,07	0,84	1,33	3,93	1,78	S(CC([NH3+])C(=O)[O-])CC[NH3+]
3DIL	0,45	1,82	0,4	0,71	0,8	1,01	0,31	0,39	O=C([O-])C([NH3+])CCCC[NH3+]
3DIM	5,35	4,58	0,76	0,51	0,96	1,03	0,64	1,07	O=C([O-])C([NH3+])CCCC[NH3+]
3DIO	0,54	0,75	0,62	0,63	0,88	1,45	0,4	1,10	O=C([O-])C([NH3+])CCCC[NH3+]
3DIQ	1,26	10,92	1,05	1,07	0,64	0,74	0,85	1,06	O=C([O-])C([NH3+])CCCCNC(=[NH2+])N
3DIR	1,39	1,6	1,22	1,33	0,75	1	1,16	1,01	O=C([O-])C([NH3+])CCCCNC(=[NH2+])C
3DIY	0,4	1,39	0,59	0,62	1,23	1,04	0,38	1,00	O=C([O-])C([NH3+])CCCC[NH3+]
3DIZ	0,95	11,54	0,61	0,51	0,6	0,37	0,56	1,24	O=C([O-])C([NH3+])CCCC[NH3+]

Table S1. Summary of re-docking results of 150 RNA targets with LeadIT, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, FCharge: formal charge of dominant protomer.

Target				Ligand							
PDB-ID	source	resolution / no. of conformers submitted (calculated)	type	Weight	lip_don	lip_acc	SlogP	lip_violation	TPSA	opr_nrot	FCharge
3DJ0	X-Ray	2,5	Crystallization of the Thermotoga maritima lysine riboswitch bound to L-4-oxalysine	149,2	6	5	-4,4	1	105	5	1
3DJ2	X-Ray	2,5	Crystallization of the Thermotoga maritima lysine riboswitch bound to lysine, Tl+ Soak	147,2	6	4	-3,2	1	95,4	5	1
3DLL	X-Ray	3,5	The oxazolidinone antibiotics perturb the ribosomal peptidyl-transferase center and effect tRNA positioning	337,4	1	7	1,1	0	71,1	6	0
3EGZ	X-Ray	2,2	Crystal structure of an in vitro evolved tetracycline aptamer and artificial riboswitch	482,9	9	10	-1,4	1	192	2	0
3F2Q	**X-Ray**	**2,95**	**Crystal structure of the FMN riboswitch bound to FMN**	**453,3**	**3**	**13**	**-2,5**	**1**	**183**	**7**	**-3**
3F4E	X-Ray	3,05	Crystal structure of the FMN riboswitch bound to FMN, split RNA.	453,3	3	13	-2,5	1	183	7	-3
3F4G	X-Ray	3,01	Crystal structure of the FMN riboswitch bound to riboflavin.	375,4	4	10	-0,3	0	143	5	-1
3FU2	**X-Ray**	**2,85**	**Cocrystal structure of a class-I preQ1 riboswitch**	**180,2**	**7**	**6**	**-1,3**	**1**	**111**	**1**	**1**
3G4S	X-Ray	3,2	Co-crystal structure of Tiamulin bound to the large ribosomal subunit	494,8	2	5	3,6	0	68	11	1
3GER	X-Ray	1,7	Guanine riboswitch bound to 6-chloroguanine	169,6	3	5	0,6	0	80,5	0	0
3GX2	X-Ray	2,9	TteSAM-I riboswitch variant A94GU34C bound to sinefungin	384,4	11	12	-5,2	2	213	8	1
3IRW	X-Ray	2,7	Structure of a c-di-GMP riboswitch from V. cholerae	687,4	7	24	-7,6	3	333	12	-1
3IWN	X-Ray	3,2	Co-crystal structure of a bacterial c-di-GMP riboswitch	687,4	7	24	-7,6	3	333	12	-1
3K1V	X-Ray	2,2	Cocrystal structure of a mutant class-I preQ1 riboswitch	180,2	7	6	-1,3	1	111	1	1
3OWI	X-Ray	2,85	Crystal structure of the glycine riboswitch bound to glycine	75,1	3	3	-3,0	0	67,8	1	0
3Q50	X-Ray	2,75	Structural analysis of a class I PreQ1 riboswitch aptamer in the metabolite-bound state	180,2	7	6	-1,3	1	111	1	1
3SUH	X-Ray	2,65	Crystal structure of THF riboswitch, bound with 5-formyl-THF	471,4	6	14	-4,5	2	221	10	-2
3SUX	X-Ray	2,9	Crystal structure of THF riboswitch, bound with THF	181,2	5	6	-1,8	0	91,5	1	0
3TZR	X-Ray	2,21	Structure of a Riboswitch-like RNA-ligand complex from the Hepatitis C Virus Internal Ribosome Entry Site	334,5	5	6	-1,5	0	63,2	6	3
4FE5	**X-Ray**	**1,32**	**Crystal structure of the xpt-pbuX guanine riboswitch aptamer domain in complex with hypoxanthine**	**136,1**	**2**	**5**	**-0,2**	**0**	**70,1**	**0**	**0**
4JF2	X-Ray	2,5	Structure of a class II preQ1 riboswitch reveals ligand recognition by a new fold	180,2	7	6	-1,3	1	111	1	1
4L81	X-Ray	2,95	Structure of the SAM-I/IV riboswitch (env87(deltaU92, deltaG93))	399,5	7	11	-3,9	2	187	8	1
4LVV	X-Ray	2,1	Structure of the THF riboswitch	343,3	7	10	-1,8	1	155	5	0
4LVW	**X-Ray**	**1,77**	**Structure of the THF riboswitch bound to 7-deazaguanine**	**150,1**	**4**	**5**	**-0,3**	**0**	**83,3**	**0**	**0**
4LVX	X-Ray	1,9	Structure of the THF riboswitch bound to tetrahydrobiopterin	241,3	7	8	-3,1	1	132	3	0
4LVY	X-Ray	2	Structure of the THF riboswitch bound to pemetrexed	297,3	6	7	0,6	1	126	4	0
4LVZ	X-Ray	1,77	Structure of the THF riboswitch bound to 2,6-diaminopurine	150,1	5	6	-0,5	0	107	0	0
4LW0	X-Ray	1,89	Structure of the THF riboswitch bound to adenine	135,1	3	5	-0,1	0	80,5	0	0
4NYA	X-Ray	2,65	Crystal structure of the E. coli thiM riboswitch in complex with 5-(azidomethyl)-2-methylpyrimidin-4-amine	164,2	2	6	1,3	0	89,2	3	0
4NYG	X-Ray	3,05	Crystal structure of the E. coli thiM riboswitch in complex with thiamine	265,4	3	5	0,9	0	75,9	4	1
4P20	X-Ray	2,7	Crystal structures of the bacterial ribosomal decoding site complexed with amikacin	589,6	21	18	-11,3	3	338	16	4
4Q9R	X-Ray	3,12	Crystal structure of an RNA aptamer bound to trifluoroethyl-ligand analog in complex with Fab	320,2	0	4	2,4	0	49,7	2	0
4RZD	X-Ray	2,75	Crystal Structure of a PreQ1 Riboswitch	180,2	7	6	-1,3	1	111	1	1
4TS2	X-Ray	2,88	Crystal structure of the Spinach RNA aptamer in complex with DFHBI, magnesium ions	698,9	4	13	2,2	2	183	24	0
4U24	X-Ray	2,9	Crystal structure of the E. coli ribosome bound to dalfopristin.	252,2	0	4	1,7	0	55,5	2	0
4U3M	X-Ray	3	Crystal structure of Anisomycin bound to the yeast 80S ribosome	265,3	2	5	0,5	0	67,8	5	0
4V56	X-Ray	3,93	Crystal structure of the bacterial ribosome from Escherichia coli in complex with spectinomycin.	334,4	7	9	-5,0	1	139	4	2
4V7S	X-Ray	3,25	Crystal structure of the E. coli ribosome bound to telithromycin.	815,0	3	15	2,9	2	176	25	1
4V7T	X-Ray	3,19	Crystal structure of the E. coli ribosome bound to chloramphenicol.	323,1	3	7	1,0	0	115	6	0
4V7U	X-Ray	3,1	Crystal structure of the E. coli ribosome bound to erythromycin.	734,9	6	14	0,4	3	195	24	1
4V7V	X-Ray	3,29	Crystal structure of the E. coli ribosome bound to clindamycin.	426,0	5	7	-1,0	0	103	10	1
4V7W	X-Ray	3	Structure of the Thermus thermophilus ribosome complexed with chloramphenicol.	323,1	3	7	1,0	0	115	6	0
4V7X	X-Ray	3	Structure of the Thermus thermophilus ribosome complexed with erythromycin.	734,9	6	14	0,4	3	195	24	1
4V7Z	X-Ray	3,1	Structure of the Thermus thermophilus 70S ribosome complexed with telithromycin.	813,0	2	15	3,8	2	173	25	1
4V8A	X-Ray	3,2	The structure of thermorubin in complex with the 70S ribosome from Thermus thermophilus.	598,5	1	12	3,3	2	189	9	-2

Table S1. Summary of re-docking results of 150 RNA targets with LeadIT, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, FCharge: formal charge of dominant protomer.

	Re-docking RMSD [Å]								
PDB-ID	LeadIT-2.3.2	FlexX-3	HYBRID 3.3.0.3	FRED 3.3.0.3	rDOCK (SF3)	rDOCK (SF5)	DOCK6	DOCK6 (mod. charge)	ligand smiles
3DJ0	0,82	11,43	1,24	1,17	1,31	1,06	3,77	0,49	O=C([O-])C([NH3+])COCC[NH3+]
3DJ2	0,8	1,64	0,51	0,56	0,8	0,84	0,41	1,03	O=C([O-])C([NH3+])CCCC[NH3+]
3DLL	1,37	1,81	1,54	2,46	1,74	1,67	3,85	4,60	Fc1c(N2CCOCC2)ccc(N2C(=O)OC(CNC(=O)C)C2)c1
3EGZ	7,01	5,58	1,66	1,78	2,1	2,1	1,83	1,57	Clc1c2C(O)(C)C3C(=C(O)c2c(O)cc1)C(O)C1(O)C([O-])=C(C(=O)N)C(O)C([NH+](C)C)C1C3
3F2Q	**0,59**	**1,48**	**0,74**	**0,8**	**0,73**	**1,26**	**1,51**	**1,06**	**P(=O)(OCC(O)C(O)C(O)CN1c2c(N=C3C(=O)[N-]C(=O)N=C13)cc(C)c(C)c2)([O-])[O-]**
3F4E	0,68	1,53	0,72	0,76	0,49	0,41	1,13	1,74	P(=O)(OCC(O)C(O)C(O)CN1c2c(N=C3C(=O)[N-]C(=O)N=C13)cc(C)c(C)c2)([O-])[O-]
3F4G	3,32	0,96	1,25	1,24	7,33	1,32	0,34	0,25	O=C1N=C2N(CC(O)C(O)C(O)CO)c3c(N=C2C(=O)[N-]1)cc(C)c(C)c3
3FU2	**1,01**	**0,83**	**0,5**	**0,41**	**0,74**	**0,62**	**0,2**	**0,27**	**O=C1NC(N)=Nc2[nH]cc(C[NH3+])c12**
3G4S	1,65	7,01	1,06	1,89	3,33	2,27	7,85	7,86	S(CC(=O)OC1C2(C)C(C)CCC3(C(C)C(O)C(C=C)(C)C1)C2C(=O)CC3)CC[NH+](CC)CC
3GER	0,47	2,88	0,37	0,47	2,85	2,88	0,28	0,27	Clc1nc(N)nc2[nH]cnc12
3GX2	1,59	6,07	1,51	2,12	7,04	6,94	6,01	5,80	O=C([O-])C([NH3+])CCC([NH3+])CC1C(O)C(O)C(n2c3NCN=C(N)c3nc2)O1
3IRW	0,66	6,02	2,58	2,45	0,5	0,36	0,55	0,25	O=P1([O-])OC2C(O)C([N+]=3C4=NC(N)=NC(=O)C4=NC=3)OC2COP(=O)([O-])OC2C(O)C(n3c4N=C(N)NC(=O)c4nc3)OC2CO1
3IWN	1,31	6,13	2,66	2,64	0,49	0,51	7,1	0,50	O=P1([O-])OC2C(O)C([N+]=3C4=NC(N)=NC(=O)C4=NC=3)OC2COP(=O)([O-])OC2C(O)C(n3c4N=C(N)NC(=O)c4nc3)OC2CO1
3K1V	0,33	0,31	0,85	0,86	0,33	0,3	0,42	0,31	O=C1NC(N)=Nc2[nH]cc(C[NH3+])c12
3OWI	0,85	1,23	2,3	3,28	1,15	1,13	6,4	6,46	O=C([O-])C[NH3+]
3Q50	0,58	0,77	0,54	0,82	0,29	0,31	0,49	0,43	O=C1NC(N)=Nc2[nH]cc(C[NH3+])c12
3SUH	6,26	7,87	2,89	7,69	3,85	3,59	9,57	8,84	O=C([O-])C(NC(=O)c1ccc(NCC2N(C=O)C=3C(=O)NC(N)=NC=3NC2)cc1)CCC(=O)[O-]
3SUX	3,07	3,15	2,91	3,05	2,92	2,92	2,55	0,44	O=C1NC(N)=NC=2NCC(C)NC1=2
3TZR	1,17	2,14	1,61	5,5	1,44	0,92	5,85	7,00	[NH+](CCCn1c(N)[nH+]c2c1c1c(OCC(C[NH+](C)C)C1)cc2)(C)C
4FE5	**0,41**	**3,39**	**0,35**	**0,38**	**0,38**	**0,52**	**3,41**	**0,23**	**O=C1NC=Nc2[nH]cnc12**
4JF2	0,35	0,18	0,37	0,56	0,52	0,45	0,42	0,25	O=C1NC(N)=Nc2[nH]cc(C[NH3+])c12
4L81	1,37	1,18	1,12	2,7	5,62	2,34			[S+](CCC([NH3+])C(=O)[O-])(CC1C(O)C(O)C(n2c3ncnc(N)c3nc2)O1)C
4LVV	3,19	3,71	3,03	3,98	5,99	3,34	6,23	3,96	O=C(N)c1ccc(NCC2N(C=O)C=3C(=O)N=C(N)NC=3NC2)cc1
4LVW	**0,42**	**0,55**	**0,3**	**0,3**	**0,48**	**0,38**	**0,4**	**0,40**	**O=C1NC(N)=Nc2[nH]ccc12**
4LVX	3,96	3,77	3,42	3,7	3,87	3,97	0,7	0,62	O=C1NC(N)=NC=2NCC(C(O)C(O)C)NC1=2
4LVY	4,69	5,03	3,32	5,02	0,77	5,7	4,42	4,37	O=C(N)c1ccc(CCc2c3C(=O)NC(N)=Nc3[nH]c2)cc1
4LVZ	0,37	0,49	3,77	3,81	0,32	0,32	0,33	0,27	Nc1nc(N)c2nc[nH]c2n1
4LW0	0,23	0,57	0,29	3,13	0,61	0,49	0,33	0,31	Nc1ncnc2[nH]cnc12
4NYA	1,38	1,27	1,8	1,4	4,95	0,99	2,77	2,76	[N+](#N)[N-]Cc1c(N)nc(C)nc1
4NYG	4,37	4,42	2,33	2,9	7,48	4,73	4,38	4,47	OCCc1c(C)[n+](Cc2c(N)nc(C)nc2)cs1
4P20	6,2	6,97	1,3	1,5	1,54	8,31	1,5	1,49	O=C(NC1C(OC2C(O)C([NH3+])C(O)C(CO)O2)C(O)C(OC2C(O)C(O)C(O)C([NH3+])O2)C([NH3+])C1)C(O)CC[NH3+]
4Q9R	1,74	1,77	1,33	6,23	2,59	2,63	6,72	0,68	FC(F)(F)CN1C(=O)C(C=C2C=C(F)C(=O)C(F)=C2)N=C1C
4RZD	0,5	0,37	0,38	0,47	0,49	0,58	0,27	0,39	O=C1NC(N)=Nc2[nH]cc(C[NH3+])c12
4TS2	0,53	2,04	2,5	8,08	6,75	0,82	2,13	2,14	Fc1c([O-])c(F)cc(CC=2C(=O)[N+](C)=C(C)N=2)c1
4U24	4,62	4,48			1	3,52			S(=O)(=O)(CCN(CC)CC)C1C2C(O)OC(C(C)C)C(C)C=CC(=O)NCCCC(C)CC(O)CC(O)Cc3occ(C(=O)N2CC1)n3
4U3M	4,9	4,49	1,28	5,35	4,48	6,8	5,94	5,74	O=C(OC1C(Cc2ccc(OC)cc2)NCC1O)C
4V56	1,18	1,07	0,61	0,63	0,9	0,32	5,83	5,76	O=C1C2(O)OC3C([NH2+]C)C(O)C([NH2+]C)C(O)C3OC2OC(C)C1
4V7S			3,07	11,75	11,36	2,75	10,53	13,98	O(C)C1(C)C(OC2C(O)C([NH+](C)C)CC(C)O2)C(C)C(=O)C(C)C(=O)OC(CC)C2(C)OC(O)N(CCCCn3cc(-c4cnccc4)nc3)C2C(C)C(=O)C(C)C1
4V7T	4,44	1,56	1,67	1,99	1,71	0,92	1,21	0,43	ClC(Cl)C(=O)NC(C(O)c1ccc([N+](=O)[O-])cc1)CO
4V7U	8,08	9,11	2,14	10,08	0,65	8,81	9,62	9,30	O(C)C1(C)C(O)C(C)OC(OC2C(C)C(=O)OC(CC)C(O)C)C(O)C(C)C(=O)C(C)CC(O)C)C(OC3C(O)C([NH+](C)C)CC(C)O3)C2C)C1
4V7V	7,62	1,45	0,78	1,22	1,11	1,03	0,4	0,39	ClC(C(NC(=O)C1[NH+](C)CC(CCC)C1)C1C(O)C(O)C(O)C(SC)O1)C
4V7W	7,22	7,91	1,22	1,17	1,59	1,79	11,44		ClC(Cl)C(=O)NC(C(O)c1ccc([N+](=O)[O-])cc1)CO
4V7X	7,37	8,44	1,86	7,54	8,52	8,62	10,93	11,41	O(C)C1(C)C(O)C(C)OC(OC2C(C)C(=O)OC(CC)C(O)C)C(O)C(C)C(=O)C(C)CC(O)C)C(OC3C(O)C([NH+](C)C)CC(C)O3)C2C)C1
4V7Z	6,25	6,61	2,23	11,83	5,23	0,97			O(C)C1(C)C(OC2C(O)C([NH+](C)C)CC(C)O2)C(C)C(=O)C(C)C(=O)OC(CC)C2(C)OC(=O)N(CCCCn3cc(-c4cnccc4)nc3)C2C(C)C(=O)C(C)C1
4V8A	4,2								O=C([O-])Cc1c(C(=O)CC(=O)c2c(O)cccc2)c(OC)c2c(OC)c3c([O-])c4C(=O)OC(C(=O)OC)=Cc4cc3cc2c1

Table S1. Summary of re-docking results of 150 RNA targets with LeadFr, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, fCharge: formal charge of dominant protomer.

Target				Ligand								
PDB-ID	**source**	**resolution / no. of conformers submitted (calculated)**	**type**	**Weight**	**lip_don**	**lip_acc**	**SlogP**	**lip_violation**	**TPSA**	**opr_nrot**	**FCharge**
4V9A	X-Ray	3,3	Crystal Structure of the 70S ribosome with tetracycline.	443,4	5	10	-0,7	0	181	3	-1
4V9B	X-Ray	3,1	Crystal Structure of the 70S ribosome with tigecycline.	585,7	7	13	-1,3	3	210	8	0
4V9Q	X-Ray	3,4	Crystal Structure of Blasticidin S Bound to Thermus Thermophilus 70S Ribosome.	425,5	10	13	-5,6	2	220	11	1
4WF9	X-Ray	3,43	The crystal structure of the large ribosomal subunit of Staphylococcus aureus in complex with telithromycin	813,0	2	15	3,8	2	173	25	1
4WFA	X-Ray	3,39	The crystal structure of the large ribosomal subunit of Staphylococcus aureus in complex with linezolid	337,4	1	7	1,1	0	71,1	6	0
4YHH	X-Ray	3,42	Crystal structure of the 30S ribosomal subunit from Thermus thermophilus in complex with tigecycline	585,7	7	13	-1,3	3	210	8	0
5BJO	X-Ray	2,35	Crystal structure of the Corn RNA aptamer in complex with DFHO, site-specific 5-iodo-U	280,2	0	6	2,1	0	85,2	3	-1
5BTP	X-Ray	2,82	Fusobacterium ulcerans ZTP riboswitch bound to ZMP	336,2	6	12	-4,9	2	197	6	-2
5DSL	X-Ray	2,5	PreQ1-II riboswitch with an engineered G-U wobble pair bound to Cs+	180,2	7	6	-1,3	1	111	1	1
5HKV	X-Ray	3,66	The crystal structure of the large ribosomal subunit of Staphylococcus aureus in complex with lincomycin	407,6	6	8	-2,3	1	124	10	1
5J5B	X-Ray	2,8	Structure of the WT E coli ribosome bound to tetracycline	443,4	5	10	-0,7	0	181	3	-1
5J91	X-Ray	2,96	Structure of the Wild-type 70S E coli ribosome bound to Tigecycline	585,7	7	13	-1,3	3	210	8	0
5KPY	X-Ray	2	Structure of a 5-hydroxytryptophan aptamer	220,2	5	5	-1,2	0	104	3	0
5OB3	X-Ray	2	iSpinach aptamer	253,2	1	4	2,7	0	65,2	2	-1
5V7Q	X-Ray	3,7	Cryo-EM structure of the large ribosomal subunit from Mycobacterium tuberculosis bound with a potent linezolid ar	317,4	1	6	2,3	0	71,5	4	0
6E1S	X-Ray	1,8	Crystal structure of a class I PreQ1 riboswitch complexed with a synthetic compound 1: 2-[(dibenzo[b,d]furan-2-yl)o	228,3	3	6	2,2	0	50	3	1
6E1U	X-Ray	1,94	Crystal structure of a class I PreQ1 riboswitch complexed with a synthetic compound 2: 2-[(dibenzo[b,d]furan-2-yl)o	256,3	1	3	2,1	0	26,8	4	1
6E1V	X-Ray	2,56	Crystal structure of a class I PreQ1 riboswitch complexed with a synthetic compound 3: 2-[(9H-carbazol-3-yl)oxy]-N,	255,3	2	3	1,8	0	29,5	4	1
6E1W	X-Ray	1,69	Crystal structure of a class I PreQ1 riboswitch complexed with PreQ1	180,2	7	6	-1,3	1	111	1	1
			Average	389,3	8,8	10,1	-3,6	1,3	172	8,6	1,4
			Median	348,0	6	9	-2,5	1,0	143	7	1
	count		**Min**	75,1	0	2	-13,6	0,0	6,48	0	-4
NMR	27		**Max**	868,0	26	24	5,2	3,0	414	27	6
X-Ray	123		**StdDev**	192,7	7,6	5,5	5,0	1,1	101	6,7	2,3
X-Ray < 2 Å	18										
				Lip. Viol.	count						
				0	46						
				1	47						
				2	23						
				3	34						

Table S1. Summary of re-docking results of 150 RNA targets with LeadFr, FlexX, FRED, HYBRID and rDOCK. Entries in bold were additionally subject to a binder-decoy discrimination. Lip_don: H-bond donors, lip_acc: H-bond acceptors, lip_violation: number of violations of Lipinski rule of five, opr_nrot: number of rotatable bonds, FCharge: formal charge of dominant protomer.

| | | Re-docking RMSD [Å] | | |

Table S2. Comparison of LeadIT-2.3.2 re-docking results for 11 NMR-structures using only the first or all frames including dynamic information for ligand over submitted frames. [a]Calculated with VMD-1.9.3.[1]

PDB	no. of conformers submitted to PDB (no. of conformers calculated)	re-docking RMSD in Å (frame 1)	re-docking RMSD in Å ("best frame")	lowest RMSD frame no.	NMR-derived RMSD of all frames compared to average ligand coordinates average ± standard deviation (min - max) in Å [a]
1AM0	8 (8)	3,74	1,38	5	1.7 ±0.2 (1.3 - 2.2)
1EHT	10 (50)	3,69	3,69	1	1.7 ±1.1 (0.5 - 3.9)
1EI2	17 (50)	8,10	6,96	6	3.4 ±1.1 (1.7 - 5.2)
1FYP	29(100)	3,71	3,00	22	1.4 ±0.4 (0.8 - 2.7)
1LVJ	12 (12)	9,17	3,17	4	1.2 ±0.4 (0.8 - 1.9)
1NEM	9 (100)	6,79	1,74	2	1.3 ±0.4 (0.8 - 2.0)
1Q8N	25 (100)	3,06	2,88	25	0.3 ±0.1 (0.1 -0.6)
1QD3	17 (100)	9,54	3,74	5	2.3 ±0.5 (1.5 - 3.2)
1TOB	7 (50)	3,96	3,23	2	1.5 ±0.5 (1.0 - 2.3)
1UTS	16 (50)	10,19	7,48	14	2.2 ±0.8 (1.2 - 4.8)
2TOB	13 (40)	10,81	0,83	11	0.5 ±0.3 (0.3 - 1.6)

Table S3. Correlation between (p)K_D-values, docking scores and number of matched pharmacophore features of previously reported $preQ_1$-RS ligands. [a] K_D-values are estimated from Roth *et al.* [2]

Ligand	SMILES	K_D [a]	pK_D	# matched pharmacophore freatures	FlexX-score (kJ/mol)	FRED-score (kcal/mol)	HYBRID-score (kcal/mol)
$preQ_1$ (**1**)	NCC1C=NC2=C1C(=O)NC(=N2)N	2,00E-08	7,7	7	-53,9	-17,6	-18,1
$preQ_0$ (**2**)	NC1=NC2=C(C(C=N2)C#N)C(=O)N1	1,00E-07	7,0	6	-45,7	-16,4	-16,0
guanine (**3**)	NC1=Nc2[nH]cnc2C(=O)N1	3,1623E-07	6,5	6	-44,7	-16,1	-16,6
7-Deazaguanin (**4**)	NC1=Nc2[nH]ccc2C(=O)N1	1,7783E-06	5,8	6	-45,1	-16,7	-16,7
7-methylguanin (**5**)	Cn1cnc2N=C(N)NC(=O)c12	3,1623E-05	4,5	5	-43,7	-14,9	-15,4
7-carboxamide-7-deazaguanine	NC(=O)c1c[nH]c2N=C(N)NC(=O)c12	3,1623E-08	7,5	7	-54,9	-20,5	-20,5
7-*NN*-dimethylaminomethyl-7-deazaguanine	CN(C)Cc1c[nH]c2N=C(N)NC(=O)c12	1,7783E-06	5,8	7	-52,8	-17,5	-18,4
2-aminopurine	Nc1ncc2nc[nH]c2n1	0,0001	4,0	4	-29,3	-10,1	-12,2
hypoxanthine	c1[nH]c2c(n1)[nH]cnc2=O	0,00031623	3,5	5	-28,9	-11,4	-11,5

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested). **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine). n.d.: not docked by HYBRID. [a] **12** was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	structure	SMILES	FlexX-score (kJ/mol)	FlexX-rank	FlexX-pose # matched pharmacophore features	FRED Chemgauss4 score (kcal/mol)	FRED-rank	FRED-pose # of matched pharmacophore features	HYBRID Chemgauss4 score (kcal/mol)	HYBRID-rank	HYBRID pose # of matched pharmacophore features
6		O=C1NC=2C(=O)NC(N)=NC=2N1	-48	76	6	-17	95	6	-16	14	6
7		C12=C(N=C(N1)N)N=C(NC2=O)N	-45	X	6	-18	X	6	X	X	X
8		O=C(Nc1cc(C)ccc1)CN1CCN(C=2N=C(N)NC(=O)C=2)CC1	-50	20	5	-10	9021	0	-11	871	5
9		O=C1C(N)=C(N)N=C(N)N1	-41	1326	6	-12	5948	6	-12	398	6
10		CC(C(C1CNC2=C(N1)C(=O)N=C(N2)N)O)O	-48	X	6	-18	X	6	X	X	X
11		CC(C(C1=NC2=C(NC1)N=C(NC2=O)N)O)O	-51	X	6	-19	X	6	X	X	X

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested). **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine). n.d.: not docked by HYBRID. [a] **12** was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	structure	SMILES	FlexX-score (kJ/mol)	FlexX-rank	FlexX-pose # matched pharmacophore features	FRED Chemgauss4 score (kcal/mol)	FRED-rank	FRED-pose # of matched pharmacophore features	HYBRID Chemgauss4 score (kcal/mol)	HYBRID-rank	HYBRID pose # of matched pharmacophore features
12		O=C2N=C(/N=c1\nc(nc12)[C@@H](O)[C@@H](O)CN	-46 (-51)[a]	166	5 (5)	-17	46	5	-17	6	5
13		O=C1N=C(N)Nc2nc([C@H](O)[C@H](O)CO)nc12	-47	124	5	-17	58	5	-18	3	5
14		S(CC(=O)Nc1scc(-c2ccccc2)n1)c1[nH]c2N=C(N)NC(=O)c2n1	-54	1	6	-13	2180	0	n.d.	n.d.	n.d.
15		O=C(NCC1(O)C[NH+]2[CCC1)c1cc(-c2cc(C(=O)N)cnc2)ccc1	-53	2	3	-15	594	0	-10	986	2
16		O=C(NCC1(O)C[NH+]2[CCC1)C=1C=C(C)NC(=O)C=1	-51	12	3	-16	285	1	-12	462	3
17		[N+H3]C1CC(c2nc(N)nc(NCC(C)(C)N3CC[N+H](CC)CC3)c2)C1	-49	48	4	-10	9434	0	-12	329	4

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested). **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine). n.d.: not docked by HYBRID. *__12__ was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	structure	SMILES	FlexX-score (kJ/mol)	FlexX-rank	FlexX-pose # matched pharmacophore features	FRED Chemgauss4 score (kcal/mol)	FRED-rank	FRED-pose # of matched pharmacophore features	HYBRID Chemgauss4 score (kcal/mol)	HYBRID-rank	HYBRID pose # of matched pharmacophore features
1		O=C(N)CN1CCC(C(=O)N2CCN(C=3N=C(N)NC(=O)C=3)CC2)CCC1	-49	32	5	-16	234	1	-11	575	0
1		O=C(Nc1cc(C(=O)NCc2c(Cl)3CCOCC3)ccc2)ccc1)N	-49	30	3	-15	445	3	n.d.	n.d.	0
2		O=C(NCC1(O)C(N+H2)CC1)C=1C=C(Cl)NC(=O)C=1	-49	34	3	-15	532	2	-12	464	2
2		O=C(NCCc1[nH]c(N)nn1)C12C(C[N+H2]C1)CNC2	-49	25	4	-15	460	0	-15	34	4
2		O(C)c1nc(N)nc(NCC[N+H]2CC(CO)CCC2)c1	-48	67	4	-14	1250	1	-11	653	4
2		O=C(NCc1[nH]c(N)nn1)c1cc(OCCC[N+H3])c(Cl)cc1	-45	284	4	-16	200	3	-16	24	4

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested). **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine), n.d.: not docked by HYBRID. *__12__ was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	structure	SMILES	FlexX-score (kJ/mol)	FlexX-rank	FlexX-pose # matched pharmacophore features	FRED Chemgauss4 score (kcal/mol)	FRED-rank	FRED-pose # of matched pharmacophore features	HYBRID Chemgauss4 score (kcal/mol)	HYBRID-rank	HYBRID pose # of matched pharmacophore features
24		O=C(NCc1n[nH]c2c1CCC2)c1c(C)[nH]c(N)n1	-44	450	4	-17	45	4	-16	21	5
25		O(C(CO)CO)Cn1c2N=C(N)NC(=O)c2nc1	-39	2158	5	-16	216	5	-17	4	5
26		NC1=NC=2C(=N1)C=C1N=C(N)N=C1C=2	-35	4952	2	-16	208	1	-17	13	3
27		N(=C(N)N)c1[nH]c2c([nH]1)cccc2	-31	8599	0	-19	11	3	-18	1	3
28		O(C)c1ccc(CCNC2=NC(N)=NC(=O)C2)cc1	-28	10528	1	-17	49	4	-17	9	4
29		O=C([C@@H](O)C)C1=NC=2C(=O)NC(N)=NC=2NC=1	-51	10	6	-18	27	6	-18	2	6

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested), **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine). n.d.: not docked by HYBRID. [a] **12** was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	Vendor	Purity (vendor certificate of analysis CoA)	Purity & mass found (m/z, in-house HPLC-ESI/MS, see Figure S4)
6	Cayman Chemical *via* Biomol	>90% (vendor CoA, IR spectrum conforms structure, mass found 167.1 g/mol)	> 99%, 167.0 g/mol
7	Sigma Aldrich (Merck)	>98% (vendor CoA, HPLC)	> 99%, 167.0 g/mol
8	Hit2Lead (ChemBridge Corporation)	97% (vendor CoA, mass found 343.0 g/mol)	> 97%, 343.1 g/mol
9	Alfa Aesar	>97% (vendor CoA, aqueous acid-base titration and HPLC)	> 99%, 142.0 g/mol
10	Sigma Aldrich (Merck)	>96% (vendor CoA, HPLC, IR spectrum conforms to structure)	> 99%, 242.0 g/mol
11	Sigma Aldrich (Merck)	>98.4% (vendor CoA, HPLC, IR spectrum conforms to structure)	> 99%, 240.0 g/mol

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested), **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine). n.d.: not docked by HYBRID. [a] **12** was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	Vendor	Purity (vendor certificate of analysis CoA)	Purity & mass found (m/z, in-house HPLC-ESI/MS, see Figure S4)
12	Sigma Aldrich (Merck)	>98% (vendor CoA, HPLC, IR spectrum conforms to structure)	
13	Sigma Aldrich (Merck)	>97.5% (vendor CoA, HPLC, specific rotation and IR spectrom conform to structure)	
14	Hit2Lead (ChemBridge Corporation)	purity undetermined, identity confirmed via 1H-NMR (vendor)	
15	Hit2Lead (ChemBridge Corporation)	100% (vendor CoA, mass found 355.2 g/mol)	
16	Hit2Lead (ChemBridge Corporation)	100% (vendor CoA, mass found 266.0 g/mol)	
17	Hit2Lead (ChemBridge Corporation)	100% (vendor CoA, mass found 348.0 g/mol)	

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested), **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine). n.d.: not docked by HYBRID. [a] **12** was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	Vendor	Purity (vendor certificate of analysis CoA)	Purity & mass found (m/z, in-house HPLC-ESI/MS, see Figure S4)
18	Hit2Lead (ChemBridge Corporation)	100% (vendor CoA, mass found 364.0 g/mol)	
19	Enamine Ltd.	90% (vendor)	
20	Hit2Lead (ChemBridge Corporation)	97% (vendor CoA, mass found 252.0 g/mol)	
21	Hit2Lead (ChemBridge Corporation)	94% (vendor CoA, mass found 266.0 g/mol)	
22	Hit2Lead (ChemBridge Corporation)	99% (vendor, CoA, mass found 282.2 g/mol)	
23	Hit2Lead (ChemBridge Corporation)	100% (vendor CoA, mass found 305.0 g/mol)	

Table S4. Docking results of virtual screening hits selected for purchase and testing. Ranks out of 12,507 molecules subjected to docking. X: Compounds were not part of initial VS library. **10** (tetrahydrobiopterin) and **11** (dihydrobiopterin) were selected based on structural similarity to **12** (biopterin), **13** (neopterin) and the commercially inavailable **29** (L-sepiapterin, not purchased and tested), **7** (8-aminoguanine) on similarity to **6** (8-oxoguanine). n.d.: not docked by HYBRID. [a] **12** was initially docked as the 3H protomer and not the depicted 3H protomer (FlexX score of 3H protomer in parenthesis).

ligand	Vendor	Purity (vendor certificate of analysis CoA)	Purity & mass found (m/z, in-house HPLC-ESI/MS, see Figure S4)
24	Hit2Lead (ChemBridge Corporation)	96% (vendor CoA, mass found 260.9 g/mol)	
25	Enamine Ltd.	95% (vendor)	
26	Hit2Lead (ChemBridge Corporation)	^{1}H-NMR (vendor)	
27	Sigma Aldrich (Merck)	>98.8% (vendor CoA, titration $HClO_4$ and HPLC, IR spectrum conforms to structure)	
28	Hit2Lead (ChemBridge Corporation)	^{1}H-NMR (vendor)	
29	commercially not available	commercially not available	

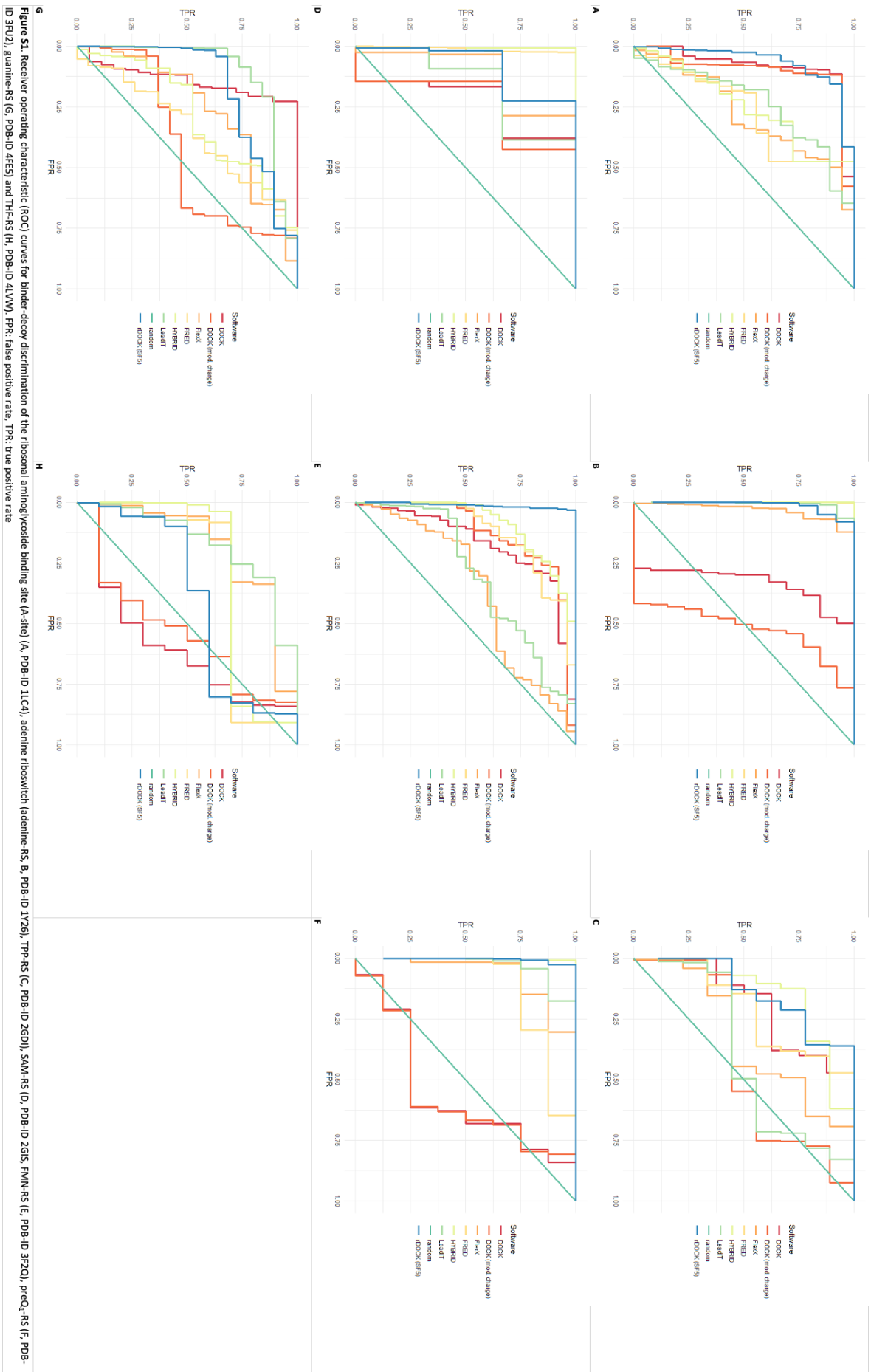


Figure S1. Receiver operating characteristic (ROC) curves for binder-decoy discrimination of the ribosonal aminoglycoside binding site (A-site) (A, PDB-ID 1LC4), adenine riboswitch (adenine-RS, B, PDB-ID 1Y26), TPP-RS (C, PDB-ID 2GDI), SAM-RS (D, PDB-ID 2GIS, FMN-RS (E, PDB-ID 3F2Q), preQ$_1$-RS (F, PDB-ID 3FU2), guanine-RS (G, PDB-ID 4FE5) and THF-RS (H, PDB-ID 4LVW). FPR: false positive rate, TPR: true positive rate

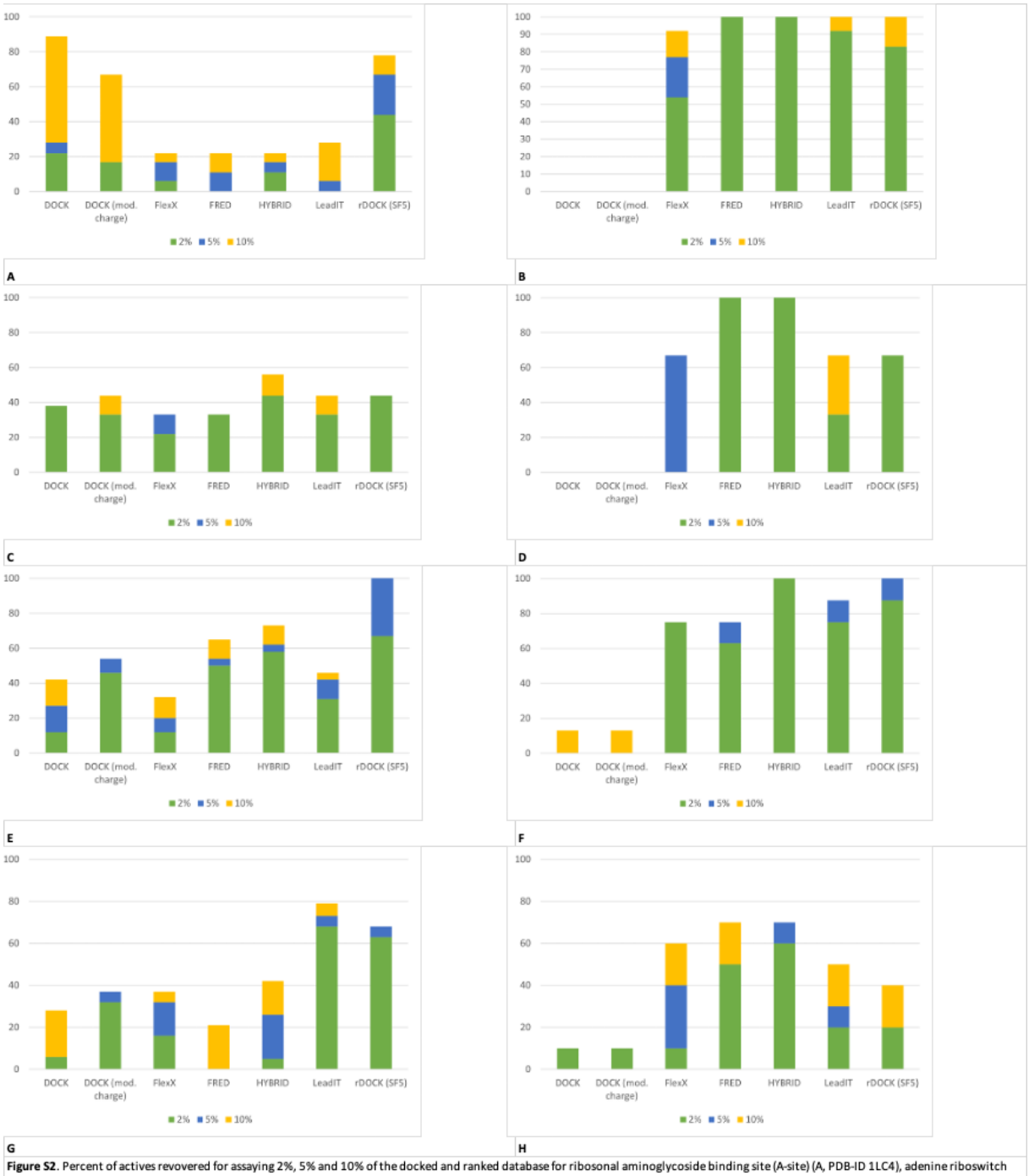


Figure S2. Percent of actives revovered for assaying 2%, 5% and 10% of the docked and ranked database for ribosonal aminoglycoside binding site (A-site) (A, PDB-ID 1LC4), adenine riboswitch (adenine-RS, B, PDB-ID 1Y26), TPP-RS (C, PDB-ID 2GDI), SAM-RS (D, PDB-ID 2GIS, FMN-RS (E, PDB-ID 3F2Q), preQ1-RS (F, PDB-ID 3FU2), guanine-RS (G, PDB-ID 4FE5) and THF-RS (H, PDB-ID 4LVW).

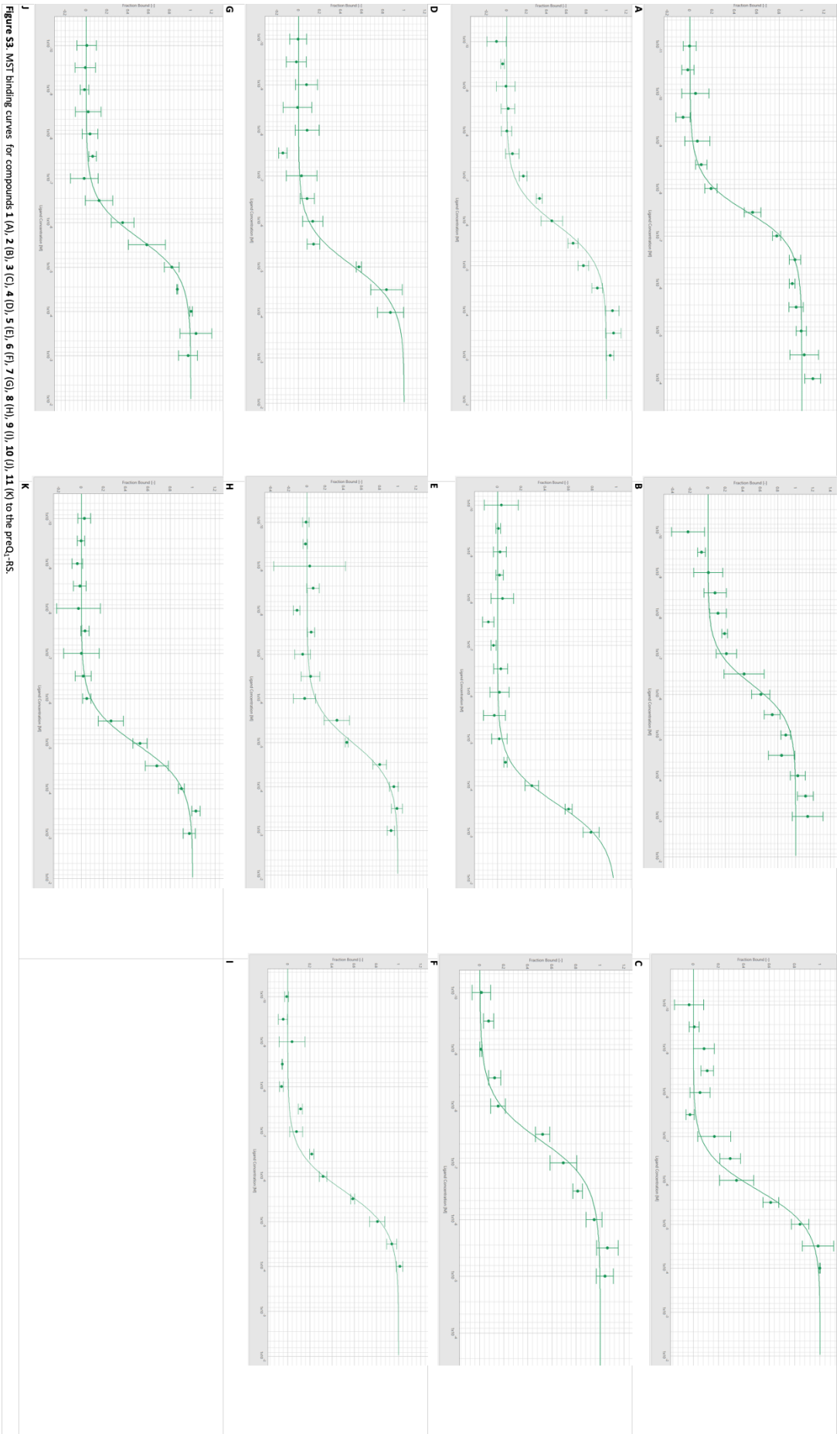
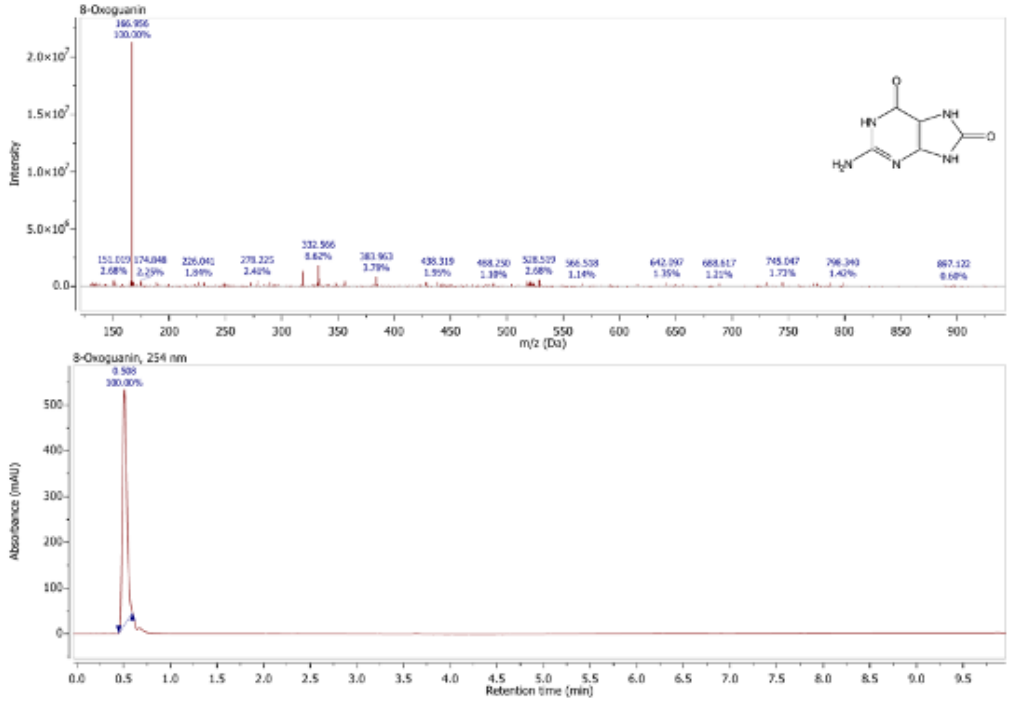
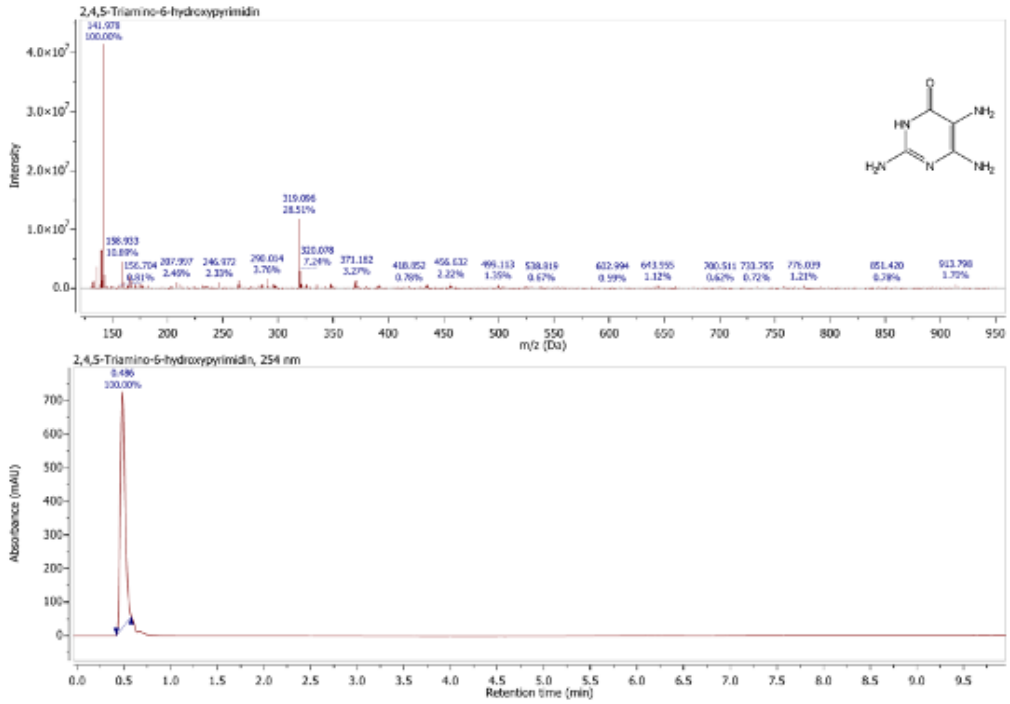


Figure S3. MST binding curves for compounds **1** (A), **2** (B), **3** (C), **4** (D), **5** (E), **6** (F), **7** (G), **8** (H), **9** (I), **10** (J), **11** (K) to the preQ$_1$-RS.



A



E

Figure S4. HPLC-ESI/MS results for **6** (A), **7** (B), **8** (C), **9** (D), **10** (E) and **11** (F).

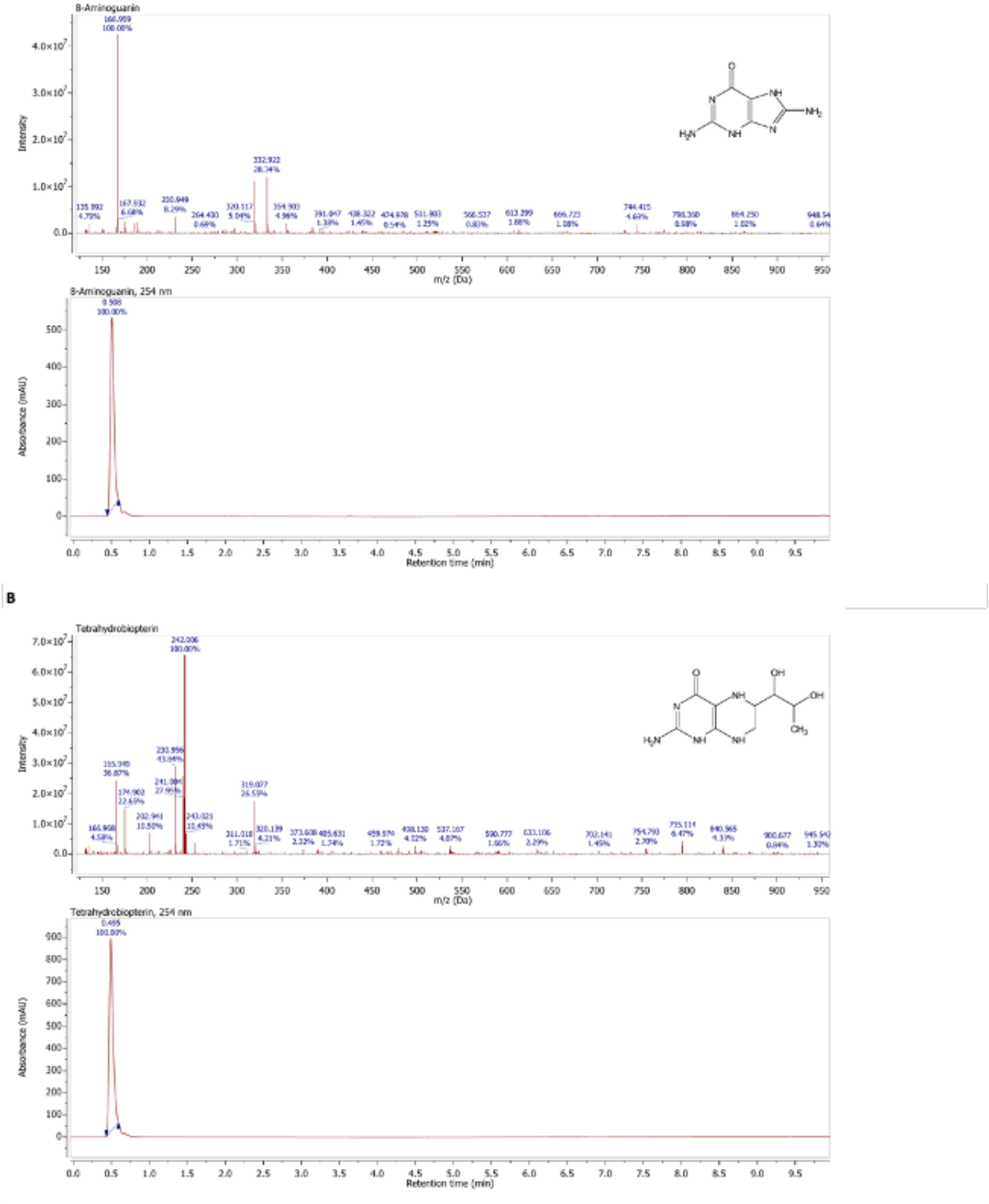


Figure S4. HPLC-ESI/MS results for **6** (A), **7** (B), **8** (C), **9** (D), **10** (E) and **11** (F).

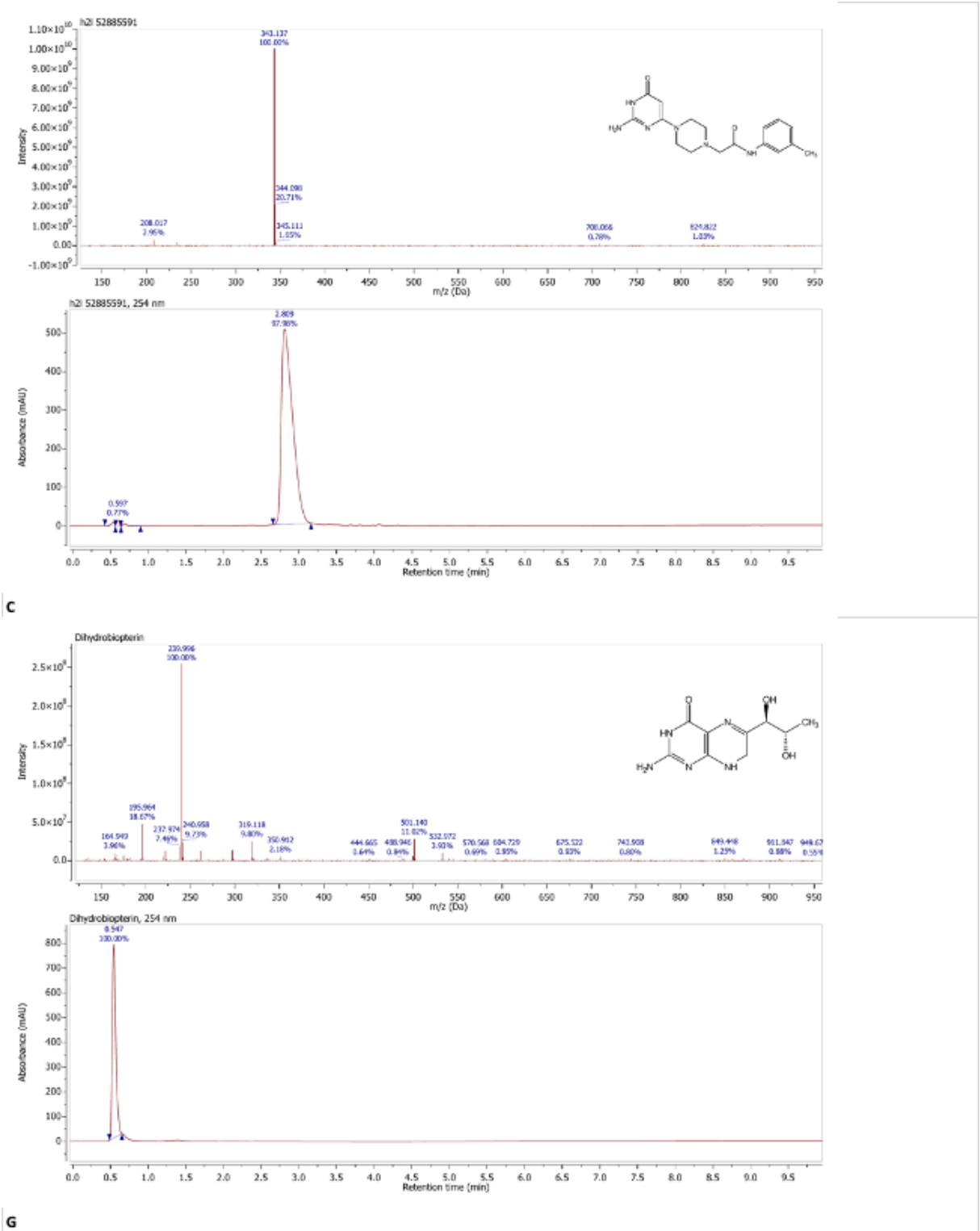


Figure S4. HPLC-ESI/MS results for **6** (A), **7** (B), **8** (C), **9** (D), **10** (E) and **11** (F).

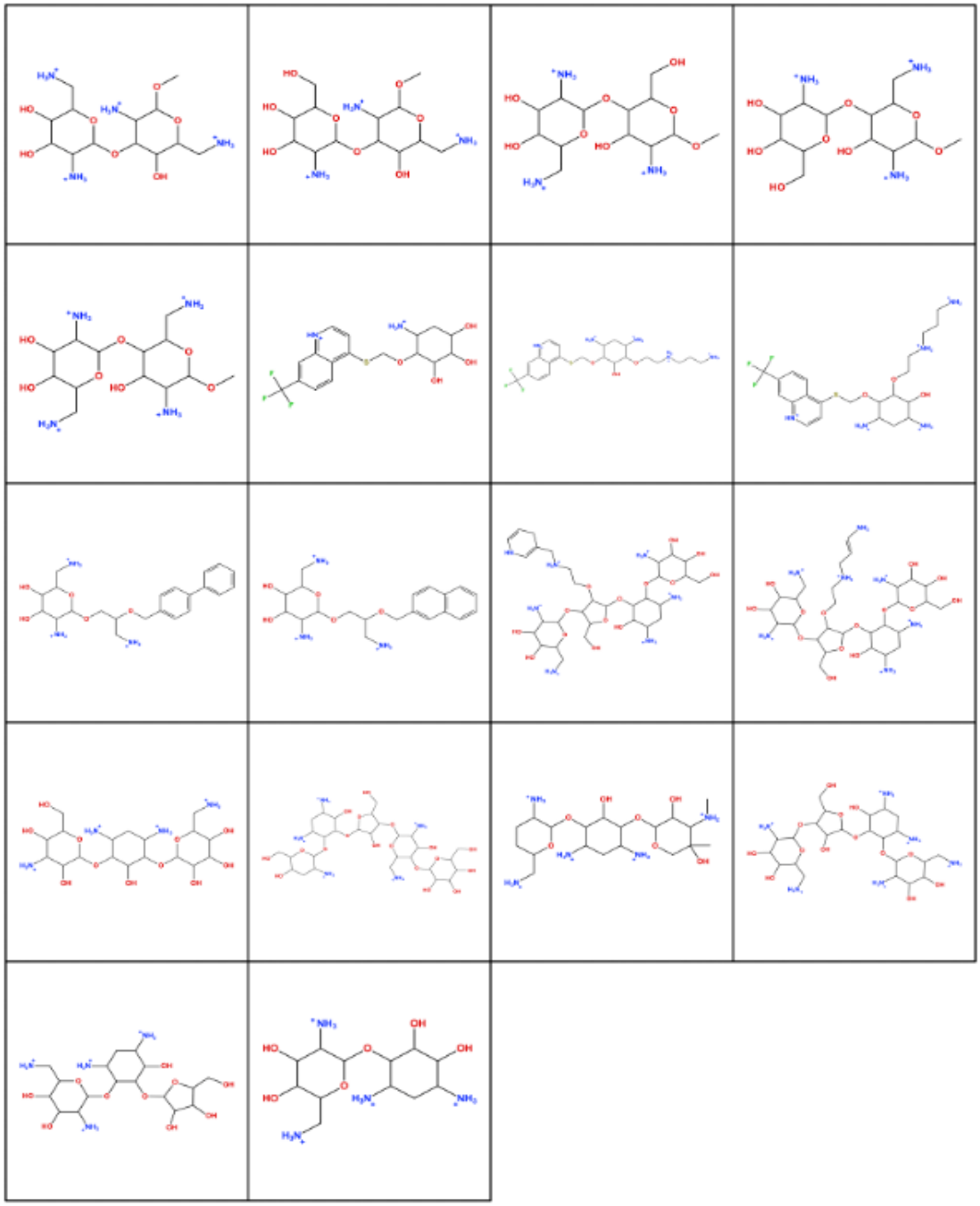


Figure S5. Ribosomal A-site ligands of binder-decoy discrimination set. Stereochemistry not shown. For stereo information please view corresponding SD file.

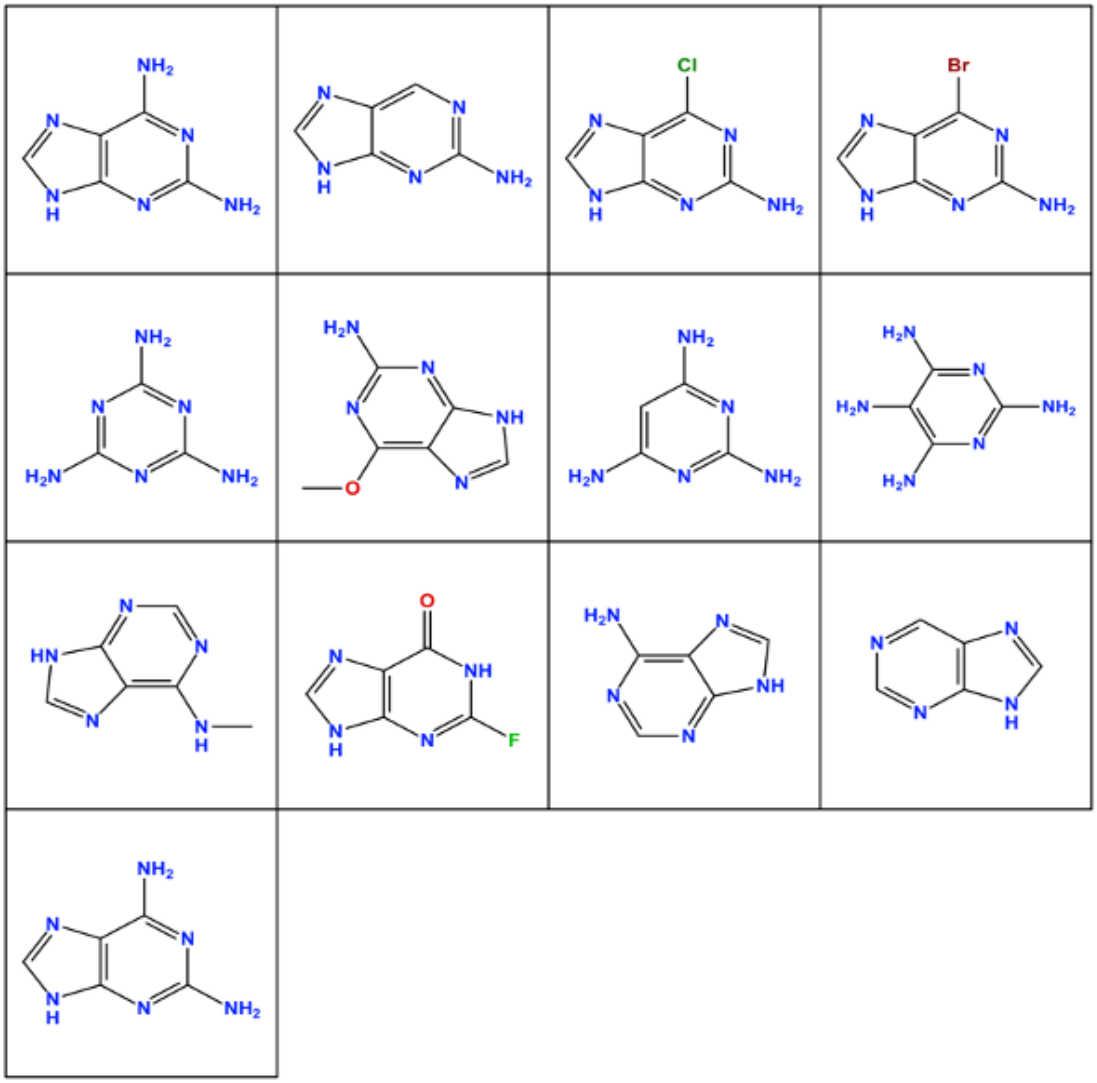


Figure S6. Adenine-RS ligands of binder-decoy discrimination set.

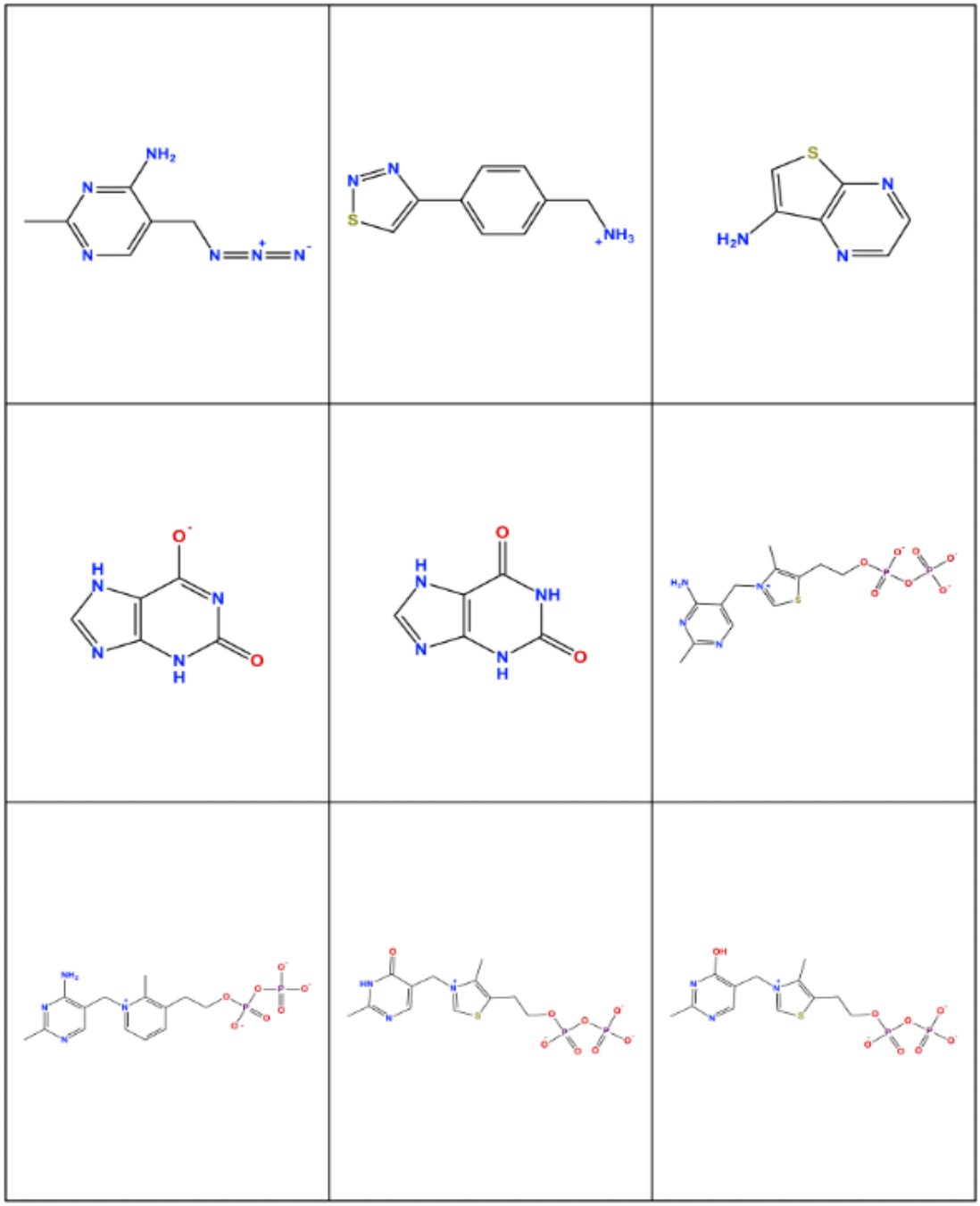


Figure S7. TPP-RS ligands of binder-decoy discrimination set.

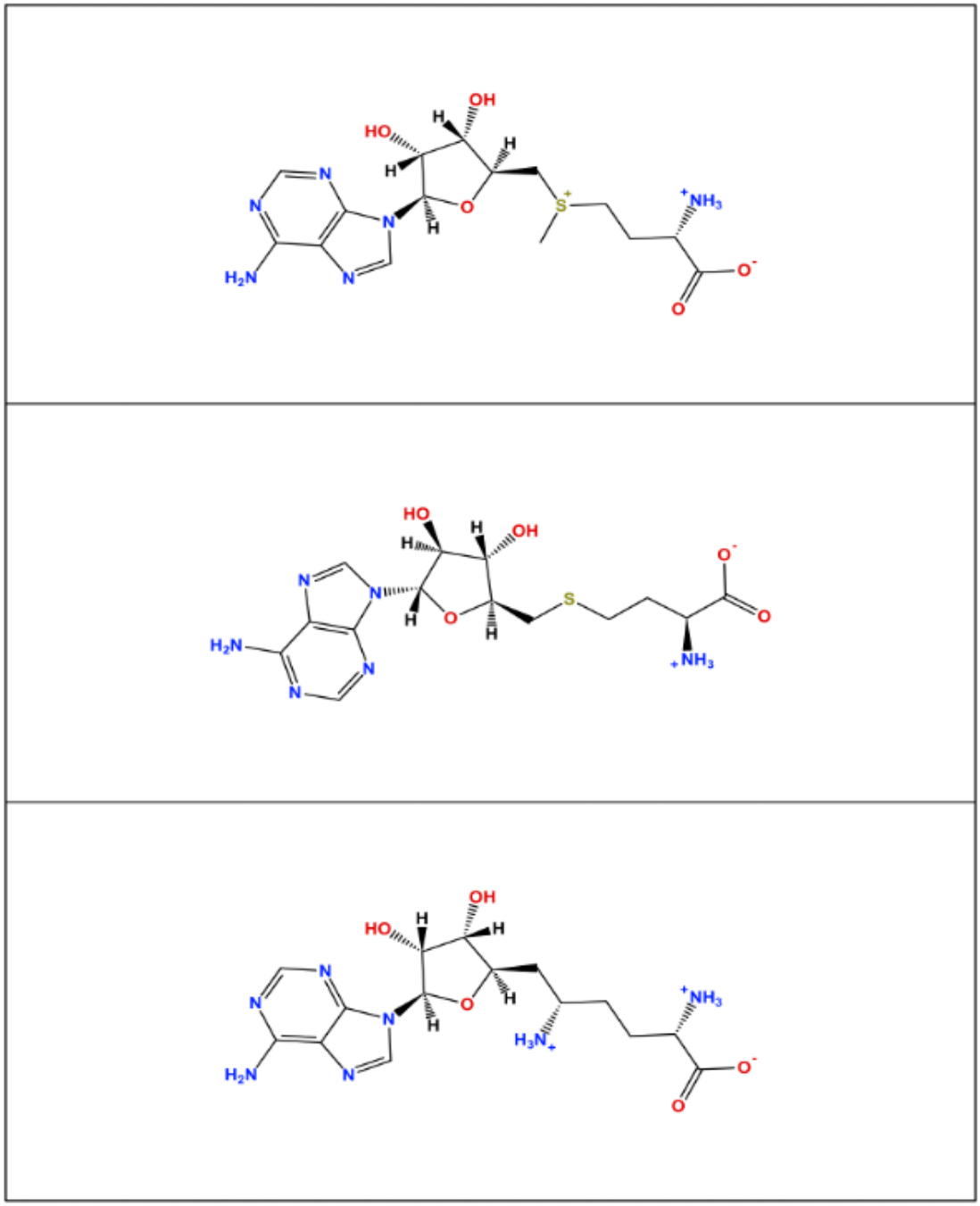


Figure S8. SAM-RS ligands of binder-decoy discrimination set.

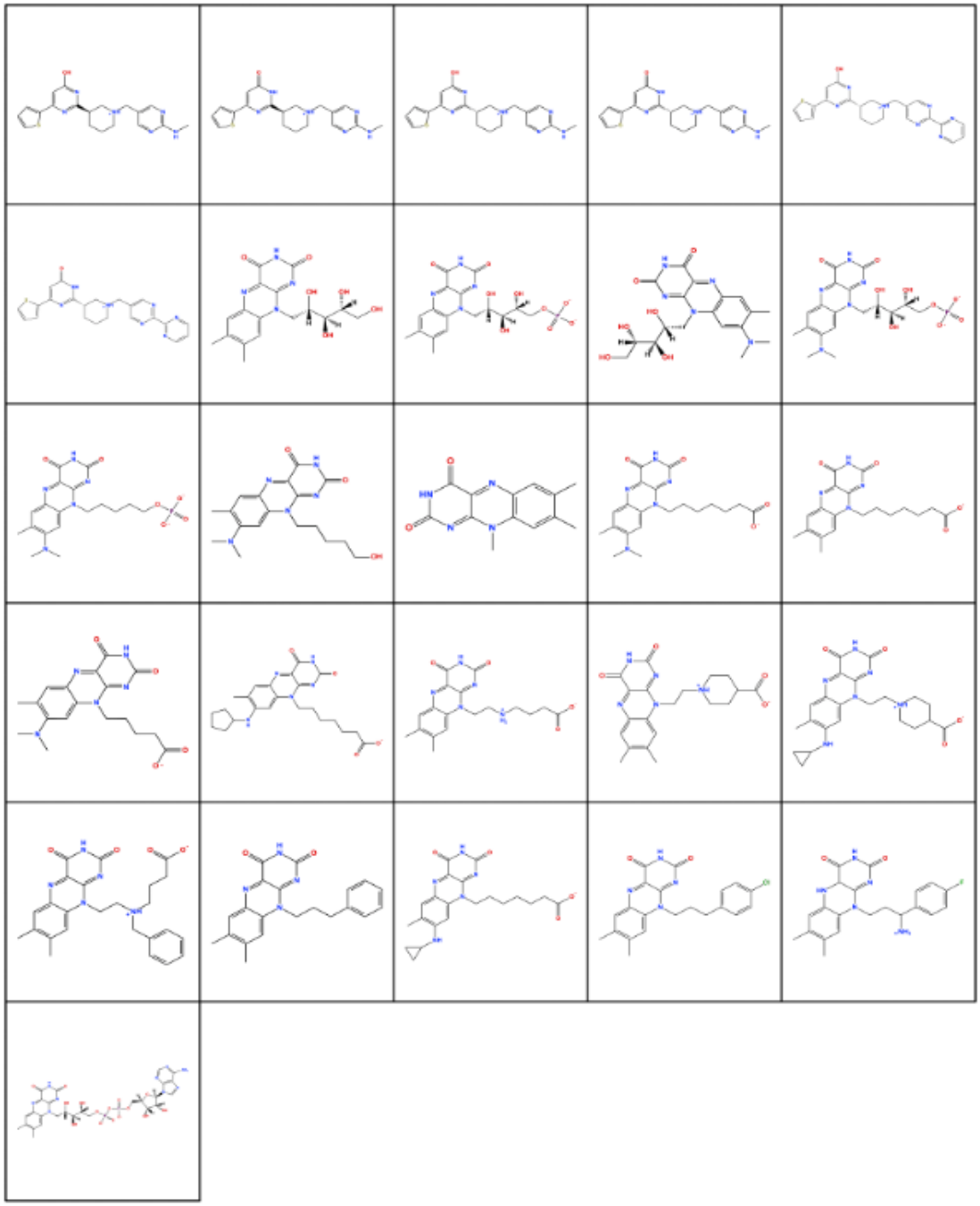


Figure S9. FMN-RS ligands of binder-decoy discrimination set.

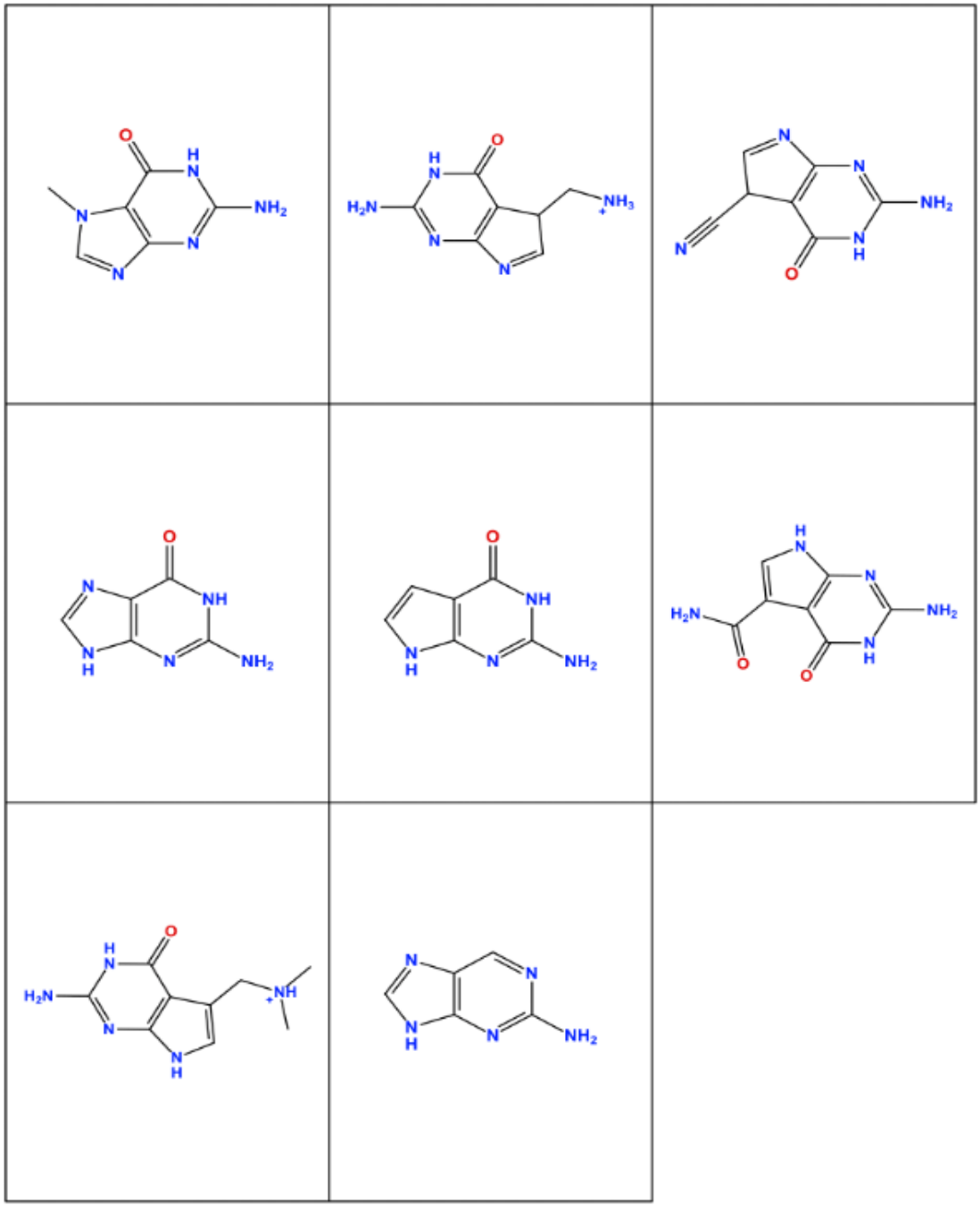


Figure S10. $PreQ_1$-RS ligands of binder-decoy discrimination set.

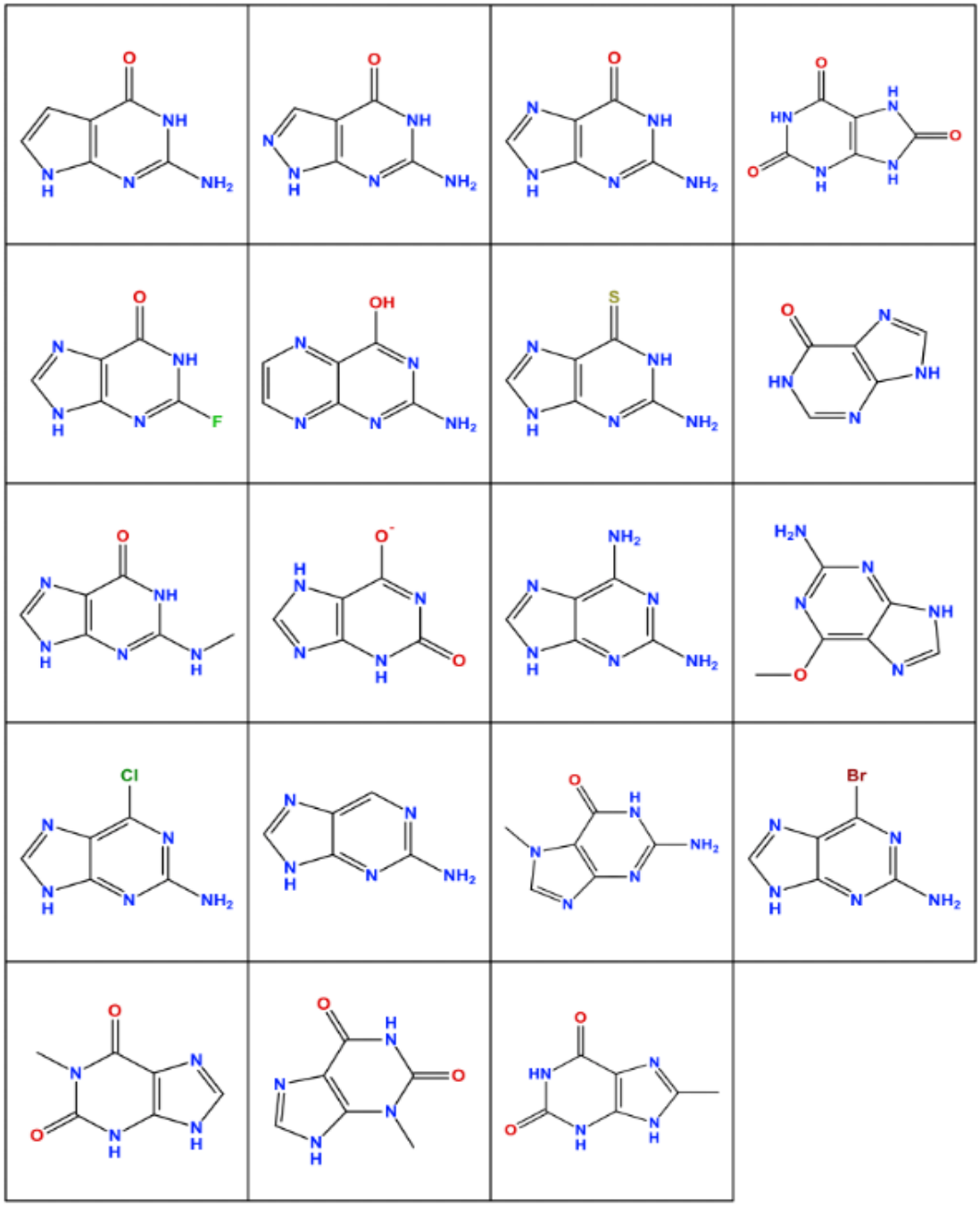


Figure S11. Guanine-RS ligands of binder-decoy discrimination set.

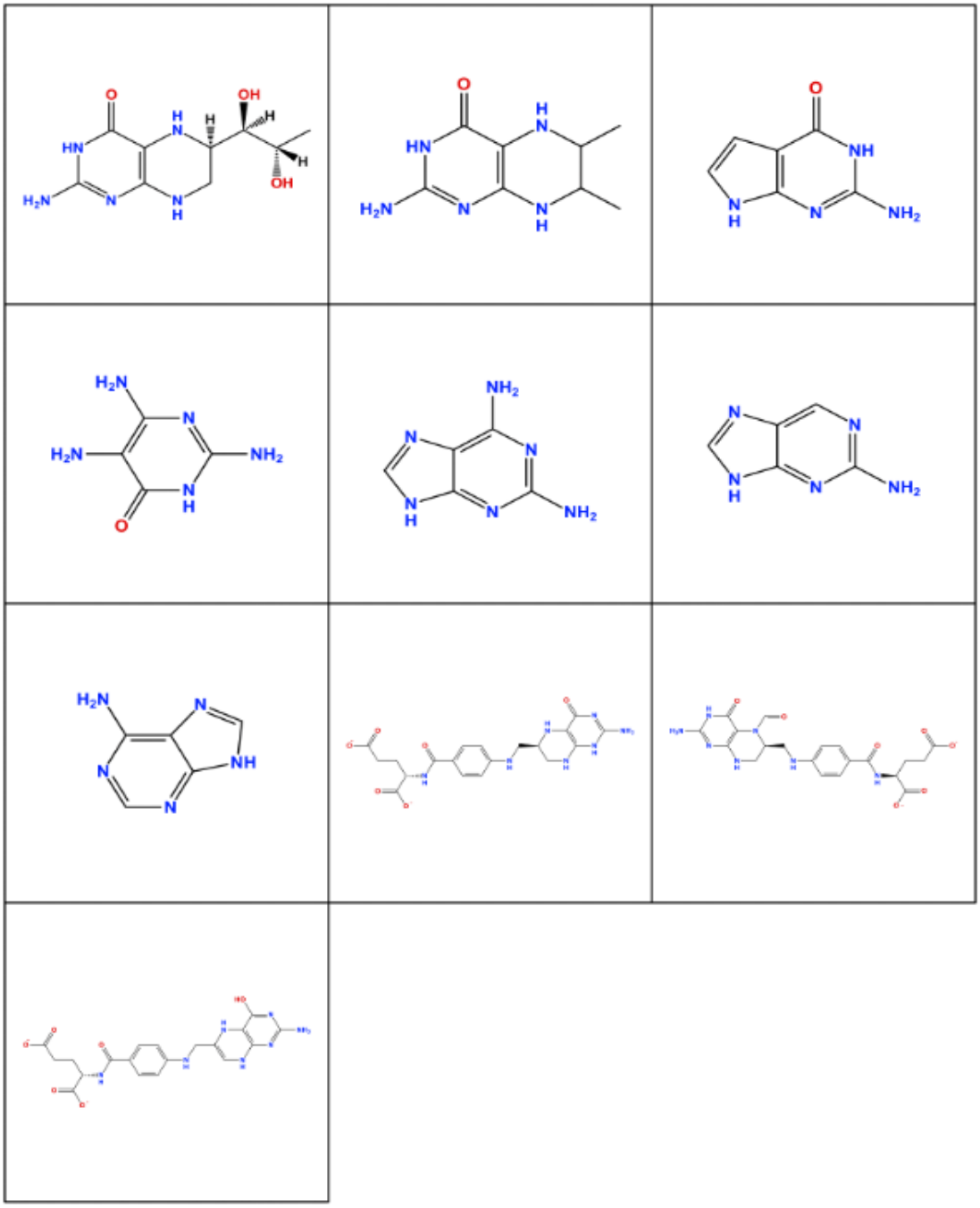


Figure S12. THF-RS ligands of binder-decoy discrimination set.

References	
1	Humphrey, W.; Dalke, A.; Schulten, K. Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
2	Roth, A.; Winkler, W. C.; Regulski, E. E.; Lee, B. W. K.; Lim, J.; Jona, I.; Barrick, J. E.; Ritwik, A.; Kim, J. N.; Welz, R.; Iwata-Reuyl, D.; Breaker, R. R. A Riboswitch Selective for the Queuosine Precursor PreQ1 Contains an Unusually Small Aptamer Domain. *Nat. Struct. Mol. Biol.* **2007**, *14* (4), 308–317. https://doi.org/10.1038/nsmb1224.

Compound	SMILES	K_D *Bacillus subtilis* preQ$_1$-riboswitch aptamer (MST, nM)
1	NCC1C=NC2=C1C(=O)NC(=N2)N	19,8
2	NC1=NC2=C(C(C=N2)C#N)C(=O)N1	557
3	NC1=Nc2[nH]cnc2C(=O)N1	1,52E+03
4	NC1=Nc2[nH]ccc2C(=O)N1	1,11E+03
5	Cn1cnc2N=C(N)NC(=O)c12	2,46E+04
6	O=C1NC=2C(=O)NC(N)=NC=2N1	2,95E+01
7	C12=C(N=C(N1)N)N=C(NC2=O)N	9,07E+03
8	O=C(Nc1cc(C)ccc1)CN1CCN(C=2N=C(N)NC(=O)C=2)CC1	9,15E+03
9	O=C1C(N)=C(N)N=C(N)N1	2,01E+03
10	CC(C(C1CNC2=C(N1)C(=O)N=C(N2)N)O)O	2,26E+03
11	CC(C(C1=NC2=C(NC1)N=C(NC2=O)N)O)O	1,10E+04
SMILES of compounds **1**–**11**		

”

4. Non-Covalent MST as a New Method for RNA-Ligand Interaction Studies

4.1. Context, Project Summary, and Own Contribution

The interest in RNA as a therapeutic target has been on the rise recently, driven by its central role in various disease-relevant biological processes.[2,100] For the accurate characterization of RNA-ligand interactions a wide range of biophysical methods can be employed. Especially for fundamental research, fast, simple and low sample consuming methods are favored.[2] MST has become one of the most attractive novel methods for studying those interactions in recent years.[280] Based on the thermophoretic behavior of molecules along a temperature gradient, this method is extremely sensitive to small changes in physicochemical properties that occur upon binding events. As the thermophoresis is detected by fluorescence, the RNA of interest must be labelled with a fluorescent probe.[62,286] Traditionally, RNA labelling is done by covalent attachment of fluorescent dyes such as Cy5 at the 3'- or 5'-end of the RNA construct.[56,194,287,288] With this, several issues come along the way, including the steric prevention of interactions.[286,289] To overcome those issues, we developed and validated an alternative MST assay protocol using non-covalent RNA-sensitive dyes. Mixing those dyes with RNA immediately leads to the statistical binding of one fluorophore per RNA molecule.[290] First, we investigated RNA- and dye-dependent effects on thermophoresis by using RNA oligos of different length and structure complexity as well as a set of seven dyes from the Syto and Sybr series. From these experiments, we concluded that there is no prediction of MST parameters based on the RNA sequence. Additionally, we found the most suitable dyes for MST experiments by evaluating their photobleaching rate, tendency forming aggregates and dissociation from the RNA complex during thermophoresis. Subsequently, we validated the method using literature known RNA-binding small molecules. We investigated interactions with different RNA-ligand size ratios to confirm the broad range application of the method. For instance, we used aptamers of the $preQ_1$ riboswitch, the fluoride riboswitch, and the adenosyl cobalamin riboswitch as well as $tRNA^{Asp}$ binding to the DNMT2 methyltransferase. These experiments prove that the assay can be easily applied to screen (small) molecule libraries against various RNA targets due to its fast and simple

“mix and measure” protocol. Therefore, it provides a robust platform for understanding RNA-small molecule interactions crucial for the design of RNA-targeted drugs.

Own contribution: Oligo structure analysis and categorization in distinct classes, preparation of riboswitch aptamers; MST measurements with seven different non-covalent dyes, oligo/functional RNA/RNA modification MST measurements, MST data evaluation, writing of the original draft & editing of the manuscript.

Contributions from other authors: MST-based METTL3/14 enzyme assay, ^{3}H-based enzyme assay, evaluation of oligo MST data, covalent MST assay, writing of the original draft & editing of the manuscript.

This work has been published in Chemical Science (impact factor: 7.6).

Article reprinted with permission from *Chem. Sci*. **2023**, 14(36), 9827–9837. ‘Non-covalent dyes in microscale thermophoresis for studying RNA ligand interactions and modifications.’ ©2022 The Royal Society of Chemistry.

4.2. Publication

The following publication quoted (within “”) from page 134 to page 153 is exactly the same as the manuscript cited on page 76: “

Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2023, **14**, 9827

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 12th June 2023
Accepted 27th August 2023

DOI: 10.1039/d3sc02993j

rsc.li/chemical-science

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Non-covalent dyes in microscale thermophoresis for studying RNA ligand interactions and modifications†

Elisabeth Kallert, Malte Behrendt, Ariane Frey, Christian Kersten and Fabian Barthels *

Microscale Thermophoresis (MST) is a powerful biophysical technique that measures the mobility of biomolecules in response to a temperature gradient, making it useful for investigating the interactions between biological molecules. This study presents a novel methodology for studying RNA-containing samples using non-covalent nucleic acid-sensitive dyes in MST. This “mix-and-measure” protocol uses non-covalent dyes, such as those from the Syto or Sybr series, which lead to the statistical binding of one fluorophore per RNA oligo showing key advantages over traditional covalent labelling approaches. This new approach has been successfully used to study the binding of ligands to RNA molecules (*e.g.*, SAM- and $PreQ_1$ riboswitches) and the identification of modifications (*e.g.*, m^6A) in short RNA oligos which can be written by the RNA methyltransferase METTL3/14.

Introduction

Understanding the molecular foundations of RNA biomolecular interactions is crucial for the study of their cellular involvement and the development of RNA-targeting therapeutics.[1,2] Microscale thermophoresis (MST) is a powerful biophysical technique for the study of the interactions between molecules such as proteins or nucleic acids with high sensitivity and specificity.[3,4] MST measures their mobility in response to a temperature gradient (usually 2–5 °C), and thus, the measured thermophoresis is dependent on the inherent properties of the biomolecule, such as its size, shape, and charge, while the origin is governed by the Soret coefficient.[5] In recent years, MST has been utilized in the study of RNA molecules, providing valuable insights into the dynamics and interactions of these intensively researched biomolecules.[6–8] The target RNA is commonly labelled with a covalent fluorescent reporter; subsequently, the sample is placed in a microliter-scale capillary, and a temperature gradient is applied by heating with an IR laser. Traditional MST methods for RNA studies rely on reporter dyes such as Cy5 or fluorescein, which are synthetically attached to the 3′- resp. 5′-ends, or to modified nucleosides.[6,8–11] While this is effective to prevent steric interaction of the fluorophore with the investigated biomolecular interface, it might also reduce the effect of local structural changes (T-jump effects) on the dyes' fluorescence.[12,13]

In this study, we present a novel MST-based methodology for the investigation of RNA-containing samples using nucleic acid-sensitive dyes, such as those from the Syto or Sybr series, which statistically bind to nucleic acids through non-covalent interaction.[14] It is important to note that most of these dyes do not significantly alter overall RNA stability which is a prerequisite for the study of biomolecular interactions.[15] To our knowledge, there is only one published study which has used non-covalent SybrGold labelling of DNA oligos using a non-microscale thermophoresis setup (microscope detection) to show that Sybr-DNA duplexes do not dissociate under common thermophoretic conditions.[16] In this previous study, however, no relation was made to applications in the recently popular MST experimental set-up nor to method development for nucleic acid analysis.

This led us to hypothesize that such nucleic acid dyes are also suitable for RNA labelling in MST experiments, and hence, we elaborated on the experimental scope, structure–activity relationships and practical applications for RNA–ligand and RNA-modification elucidation. Here, we report a “mix-and-measure” protocol that shows key advantages for relevant RNA-MST applications compared to covalent labelling approaches which are technically more challenging, time-consuming, and expensive (Fig. 1A). This approach thus pays homage to the dye-nitrilotriacetic acid (NTA) protocol developed for protein MST, which allows non-covalent labelling of proteins and has subsequently become more widely used than the original covalent maleimide- and NHS-labelling protocols.[17]

Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University of Mainz, Staudingerweg 5, 55128 Mainz, Germany. E-mail: barthels@uni-mainz.de

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3sc02993j

© 2023 The Author(s). Published by the Royal Society of Chemistry

View Article Online
Edge Article Chemical Science

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

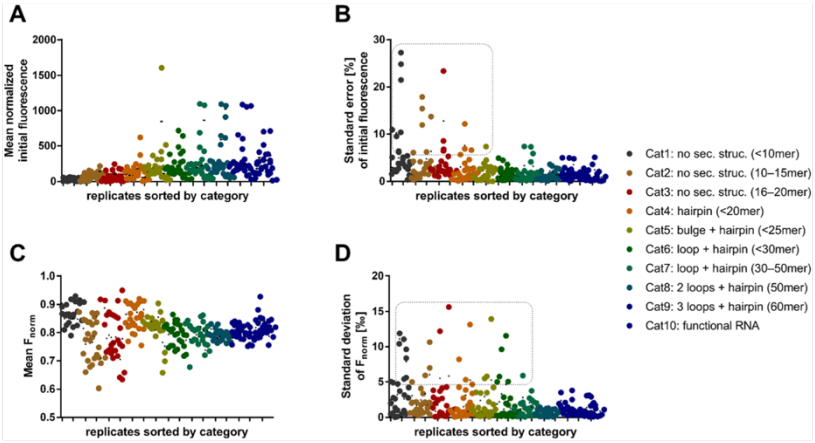


Fig. 2 Initial fluorescence and F_{norm} as RNA-dependent MST characteristics. (A) Mean normalized initial fluorescence of all dye combinations for each RNA oligo. (B) Relative standard error of the mean normalized initial fluorescence. For Syto 62 combinations with short RNAs (category 1–5) the error is >5% (dashed box). (C) The key MST descriptor F_{norm} (t = 2 s) displays mostly values between 0.75 and 0.90 across all categories and dyes. (D) Standard deviation of the F_{norm} values (t = 2 s).

loop/bulge content *vs.* F_{norm} yielded a significant correlation between the given quantities (Fig. S3†).

Another aspect of structure–activity relationships was the analysis of dye-dependent parameters on characteristics of thermophoresis curves described by the following descriptors (Fig. 3F): (1) slope of the pre-MST fluorescence traces (measure of dye photobleaching), (2) steady-state (t = 15–20 s) slope of thermophoresis traces (measure of the RNA–dye complex dissociation in the temperature gradient), (3) goodness (R^2) of the hyperbolic thermophoresis fit, (4) sum of the least squares error of a fluorescence–time curve triplicate (reproducibility of the whole MST experiment), (5) normalized initial fluorescence (measure for the brightness of the intrinsic dye fluorescence).

These results showed that Syto 60, 62, 63, and 64 exhibited increased susceptibility to photobleaching, with Syto 62 showing a particularly strong effect (Fig. 3A). In contrast, Syto 17, 59, and 61 as well as SybrGold and RiboGreen were mostly resistant to photobleaching. Furthermore, the dye–RNA complex of Syto 60, 62, and 63 seem to be prone to premature dissociation during thermophoresis, as a positive steady-state slope was observed for these dye combinations (Fig. 3B). We also found varying degrees of susceptibility to aggregation for Syto 17, 60, 62, and 63 during thermophoresis, while the other dyes were stable across all types of RNA samples. Regarding the reproducibility of an MST experiment triplicate, Syto 62 as well as Syto 60 and 63 led to significantly lower reproducibility compared to the other dyes (Fig. 3D). Lastly, Syto 59 is by far the brightest in terms of its intrinsic fluorescence compared to the other red dyes (Fig. 3E).

To summarize, we found that Syto 59 (red, pico channel) and SybrGold (blue, nano channel) are the most suitable (of the investigated) dyes for RNA-MST experiments, as they showed

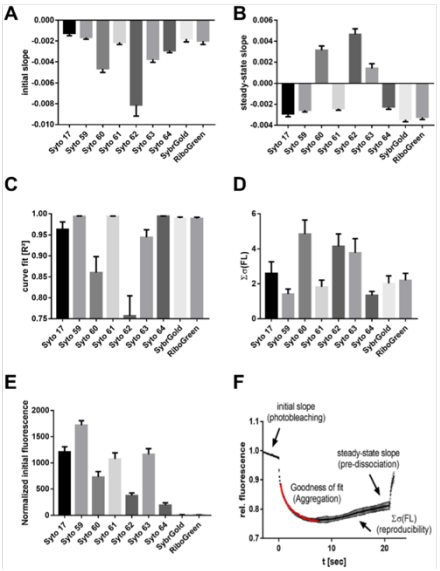


Fig. 3 Dye-dependent effects on the MST characteristics. (A) Initial slope indicating the photobleaching rate. (B) Steady-state (t = 15–20 s) slope showing the stability of dye–RNA complexes. (C) Goodness of the curve fit (R^2). (D) Dye-dependent reproducibility. (E) Normalized initial fluorescence. (F) Depiction of the MST curve descriptors analyzed in subfigures (A)–(D).

© 2023 The Author(s). Published by the Royal Society of Chemistry

View Article Online
Chemical Science
Edge Article

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

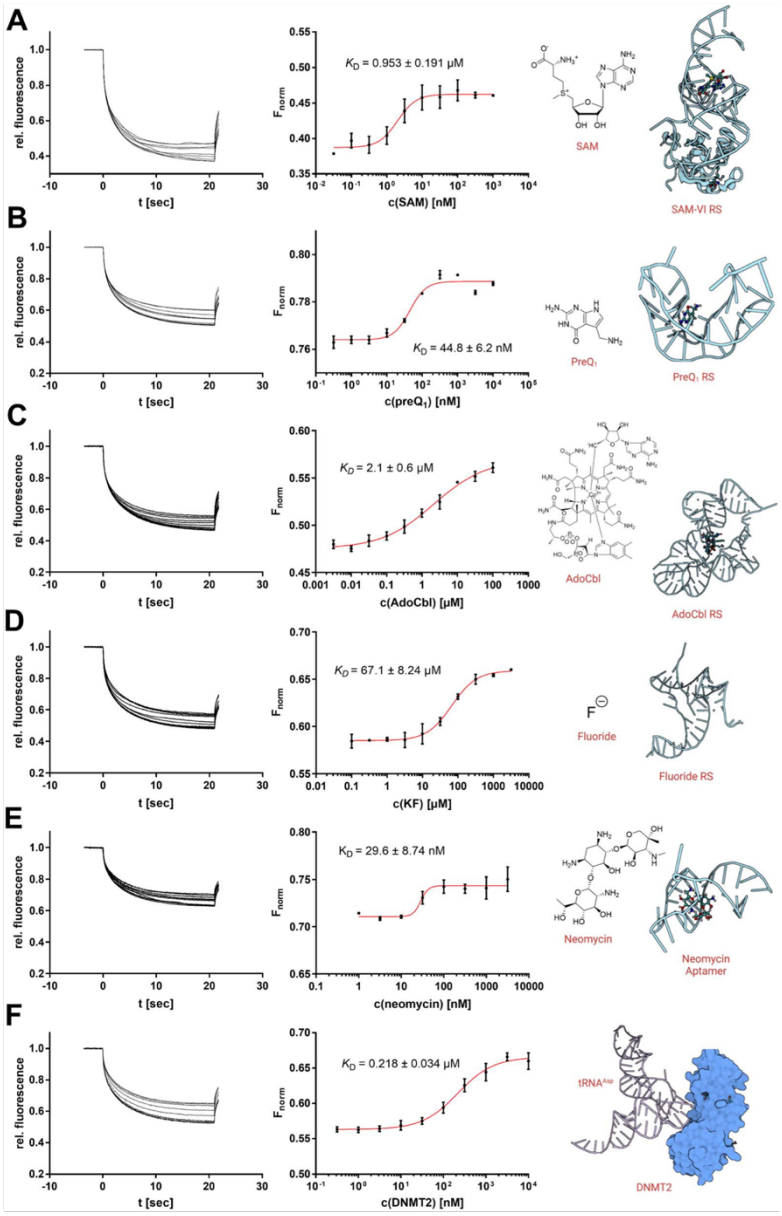


Fig. 4 Analysis of RNA–ligand binding by *in situ* labelling with SybrGold or Syto 59 (1×). Dissociation constants (K_D) were determined using a Boltzmann fit. (A) Interaction between SAM-VI RS and SAM (F_{norm} at 20 s, SybrGold). (B) PreQ$_1$ RS and the PreQ$_1$ ligand (F_{norm} at 1 s, SybrGold). (C) AdoCbl RS and the AdoCbl ligand (F_{norm} at 15 s, Syto 59). (D) Fluoride RS and the potassium fluoride ligand (F_{norm} at 5 s, Syto 59). (E) Neomycin aptamer and the neomycin ligand (F_{norm} at 5 s, Syto 59). (F) tRNA^{Asp} and DNMT2 (F_{norm} at 10 s, SybrGold).

© 2023 The Author(s). Published by the Royal Society of Chemistry

View Article Online
Chemical Science
Edge Article

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

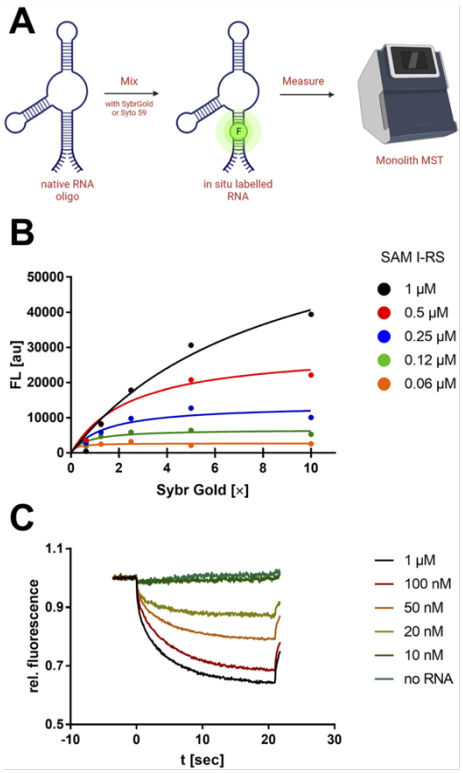


Fig. 1 Non-covalent nucleic acid dyes to study RNA by MST. (A) "Mix-and-measure" protocol for the thermophoretic analysis of native RNA oligos. (B) Fluorometric titration of SAM-I riboswitch with SybrGold. Samples were measured by a Tecan Spark 10 M plate reader ($\lambda_{ex/em}$: 490/550 nm). The apparent binding constant (Langmuir isotherm) was determined to be K_D = 85.3 nM. (C) RNA concentration-dependent MST curves with SybrGold (1×) and SAM-I riboswitch (variable concentration, blue laser, 8% excitation power). The apparent K_D from the F_{norm} (2 s) was 61.4 nM.

Results and discussion

Establishment of the non-covalent dye MST protocol

First, a qualitative working range of useful SybrGold and RNA concentrations was determined by fluorometry. With a sample matrix of RNA (62.5 nM–1 μM) and SybrGold (1×–10×) concentrations, the stoichiometry and apparent binding constant of the dye to the *S*-adenosylmethionine (SAM)-I riboswitch was determined (Fig. 1B).

It was found that 1× SybrGold produced sufficient fluorescence for MST (Monolith NT.115), and at such SybrGold concentrations, one fluorophore molecule is likely to bind statistically to the RNA of interest (K_D = 85.3 nM). Statistical 1 : 1 binding stoichiometry was concluded as a one-step binding event was observed to display the same apparent binding affinity independent of the RNA concentration. Hence, it was found that the minimum RNA concentration during MST was primarily limited by the binding affinity of SybrGold to the respective RNA sequence. For the SAM-I riboswitch, a significant thermophoretic response was saturated at an RNA concentration of approx. 100 nM, which was strongly distinguished from the unbound dye. F_{norm} (2 s): 0.75 *vs.* 0.96 (Fig. 1C).

As a next step, systematic investigations of structure–activity-relationships between fluorescent dyes and RNA in MST were performed. First, we investigated the effect of RNA length and its secondary structures on the MST characteristics, such as the initial fluorescence level and F_{norm}, using nine different dyes (Syto 17, Syto 59–64, SybrGold, and RiboGreen). For this purpose, we defined ten RNA categories containing each three RNA oligos which comprise different lengths and structural features (Table S1†). Across all dye and RNA combinations, MST mixtures were prepared as triplicates with 100 nM RNA and 1× dye. The mean initial fluorescence of different dyes was normalized by a concentration-dependent regression (Fig. S1†). In general, more complex RNA structures and longer oligonucleotide sequences resulted in higher absolute initial fluorescence (Fig. 2A). The central MST descriptor F_{norm} (2 s) was mostly between 0.75 and 0.90 across all combinations (Fig. 2C). Lower F_{norm} values (0.60–0.75) were eventually found for shorter RNA, however, a significant dependency on the dye or the category was not ascertained.

We found that shorter RNAs (<20 nt) exhibit more sensitive fluctuations in F_{norm}, suggesting that these are more susceptible to biological and chemical modifications (Fig. 2C). This can possibly be exploited to detect small changes in sequence or conditions for biological assay design. Depending on the dye, however, the standard error of the absolute fluorescence and F_{norm} was coincidently higher for these short RNAs. The standard deviation of the F_{norm} values (t = 2 s) was mostly <5‰ across all categories. An exception was the use of the dyes Syto 62 and RiboGreen. The data points in the dashed box of Fig. 2D are exclusively populated by RNA–dye combinations with Syto 62 or RiboGreen. For Syto 62, this may be due to the highly variable absolute fluorescence of the RNA–dye conjugates (Fig. 2B). The use of these two dyes might be not recommended for RNA-MST experiments.

After the broad analysis regarding the influence of RNA-dependent parameters on the initial fluorescence and F_{norm}, more detailed correlation studies with RNA descriptors (length, secondary structure, and base composition) were performed. The secondary structures of RNA oligos were predicted by the RNAfold web server.[18] By analyzing the relationship between F_{norm} and RNA length, we confirmed the trend shown in Fig. 2, wherein shorter RNAs (<20 nt) lead to a greater diversity of F_{norm} (Fig. S2A†). Notably, the F_{norm} parameters of the two blue dyes (SybrGold and RiboGreen) were generally lower than that of the red Syto dyes. High standard deviations (>5‰) were found only for Syto 62 and RiboGreen. As further correlation parameters, the proportion of hybridized bases (Fig. S2C†) and the secondary structural content were investigated (Fig. S2E†). Neither the analysis of the paired bases nor that of the

© 2023 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

the best fluorescence benchmarks and highest reproducibility rates.

Investigation of RNA ligand binding

Targeting physiologically relevant RNA with small molecule drugs has become of increasing interest as a potential drug target over the last years.[1,19,20] Currently, the standard repertoire of bioanalytical techniques includes techniques such as nuclear magnetic resonance (NMR), surface plasmon resonance (SPR), and isothermal titration calorimetry (ITC).[21,22] However, in recent years, MST has probably become one of the most popular techniques for the characterization of RNA-small molecule interactions using covalently dye-labelled RNA targets.[6,8,10] Hence, we aimed to demonstrate four examples where the use of SybrGold and Syto 59 (Fig. 4 and S4†) showed comparable results to RNA ligand binding profiles as determined with orthogonal techniques.

Thus, we investigated three types of nucleotide-derived ligand riboswitch (RS) aptamer domains: the *T. tengcongensis* SAM-I RS, the *B. angulatum* SAM-VI RS, and the *T. tengcongensis* pre-queuosine 1 ($preQ_1$) RS, with their natural ligands SAM, *S*-adenosylhomocysteine (SAH), and $preQ_1$, respectively.[23–25] For the SAM-I RS, our MST data yielded a $K_D = 157$ nM for SAM and $K_D = 137$ μM for SAH agreeing with previously reported values obtained by ITC measurements.[26] For the SAM-VI RS, our MST results for SAM ($K_D = 953$ nM) and SAH ($K_D = 10$ μM) were also consistent with our covalently Cy5-labelled MST experiments (Fig. S5†) and literature reports.[27,28] Similarly, our results showed that $preQ_1$ binds the $preQ_1$ RS with an apparent $K_{D,app}$ of 44.8 nM.

While this value is in the same order of magnitude as our previously reported Cy5-labelled RNA affinity (Fig. S5†), it should be noted that high-affinity ligands (<50 nM) will result in tight-binding to the RNA receptor ([RNA] = 100 nM) and thus imply some scope limitation of this presented protocol.[6,8] For the existing protocols using a covalently labelled RNA target, however, this mostly represents the same technical limitation.

As a next step, we investigated three additional RS aptamer domains against chemically more diverse chemotypes of ligands: a marine metagenome AdoCbl RS, the *T. petrophila* fluoride anion RS, and a synthetic neomycin aptamer.[29–31] The AdoCbl RS naturally binds different variants of the cobalamin co-factor family, which are exceptionally large ligands with a molecular weight of >1000 Da. Our MST experiments using an env8 mutant AdoCbl RS with Syto 59 labelling determined a K_D value of 2.1 μM with adenosylcobalamin (AdoCbl) as a ligand, which is in agreement with the literature.[29] Thus, the non-covalent RNA-MST method is also suitable for non-typically large ligands and their riboswitches.

Next, we studied the opposite case; a riboswitch that is able to bind very small ligands. Aptamers against the smallest known ligands are fluoride anion aptamers. The study of the *T. petrophila* fluoride anion RS showed that this protocol is also useful for such applications if ligand binding results in a substantial conformational change of the aptamer. Hence, non-covalent RNA-MST experiments with Syto 59 ($K_D = 67.1$ μM) are in the same order as dissociation constants determined by ITC experiments in the literature.[30]

As a final example, we examined a neomycin-binding aptamer of synthetic origin. The apparent binding affinity of neomycin to the 27-mer RNA oligo could be determined by $K_{D,app}$ of 29.6 nM.[31]

Of note, the titration of *in situ* labelled RNA samples with their natural ligands mostly does not lead to a change in the fluorescence intensity by displacement of the dye with the ligand (see MST capillary scans in Fig. S8†). In this regard, an RNA target usually has just one ligand binding resp. interaction site but possibly several dye binding sites, some of which overlap with the ligand binding site and some of which do not. If we choose the concentrations of RNA (100 nM) and dyes (1×) such that statistically at most 1 dye molecule is bound per RNA (Fig. 1B), there will be functional, fluorescent RNA molecules in solution that we can (a) address with the ligand resp. interaction and (b) measure their altered thermophoresis behaviour. For the very short $preQ_1$ riboswitch or the neomycin aptamer (both 27 nt), however, we quantified a slight effect on the initial fluorescence at higher ligand concentrations which supports our model of statistical labelling, since for such a short RNA receptor there are fewer opportunities for dye binding that do not overlap with the ligand binding site (Fig. S8A&G†).

Furthermore, we also investigated a macromolecular interaction of RNA with an endogenous protein–ligand. The $tRNA^{Asp}$ represents the physiological substrate of the RNA methyltransferase DNMT2.[32] Using SybrGold, we were able to determine a dissociation constant of this macromolecular interaction ($K_D = 0.218$ μM) similar to a value determined by fluorescence polarization.[33]

Investigation of RNA nucleobase modifications

The development of methods for the detection and quantification of RNA modifications such as nucleobase methylation is a recent topic of research due to their physiological relevance and disease involvement.[34] Currently, the gold standard repertoire of bioanalytical techniques includes mass spectrometric, radiometric, and sequencing methods.[35] Due to the simple and fast preparation of an MST experiment, the development of an MST protocol for the quantification of RNA modifications might offer practical advantages. Thus, the two most promising fluorescent dyes (SybrGold & Syto 59) were harnessed to test for the detection of specific RNA nucleoside modifications.

First, as a proof of concept, two variously modified model RNA oligos AUUAACCUUUUAA (Fig. 5) and CCA-CAACCAUGGUGAGCAA (Fig. 6 and S6†) were investigated for changes in their F_{norm} parameters in dependence on specific RNA modifications. In summary, these model experiments showed that larger or charge-altering RNA modifications (*e.g.*, 5-taurinomethyl-2-thiouridylation) have a more pronounced effect on F_{norm} than single nucleoside methylations (*e.g.*, 6-methyladenine, or 2′-*O*-methylation), which often resulted in scarcely detectable F_{norm} shifts (Fig. 5, 6, and S6†).

In the first step, the influence of different physiologically relevant modifications such as pseudouridylation (ψ), 5-uridine

© 2023 The Author(s). Published by the Royal Society of Chemistry

View Article Online
Chemical Science
Edge Article

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

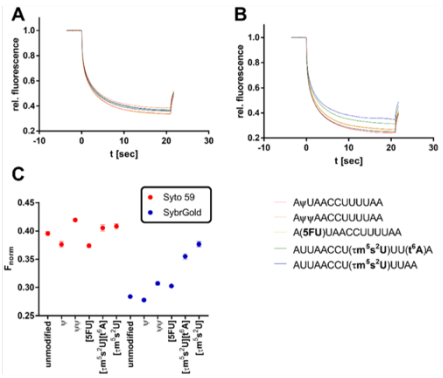


Fig. 5 Analysis of various modifications of the model oligo AUUAACCUUUAA (100 nM) by MST. (A) Thermophoresis curves for the usage of Syto 59 as a reporter dye. (B) Thermophoresis curves for the usage of SybrGold as a reporter dye. (C) MST response (F_{norm} at 10 s) of the modified oligos in comparison to their unmodified analogues.

fluorination (5FU), 5-taurinomethyl-2-thiouridiylation (τm^5s^2U), and threonylcarbamoyladenosylation (t^6A) on the short non-structured model oligonucleotide AUUAACCUUAA as a fragment of the hmt-tRNA^{Lys} was investigated.[36]

Significant divergences were found between the two dyes SybrGold and Syto 59 in the differential analysis of the F_{norm} values. While Syto 59 showed only slight changes in F_{norm} values depending on the modifications (Fig. 5A), significant increases in F_{norm} values were measurable when SybrGold was used, especially for the two oligos with the 5-taurinomethyl-2-thiouridiylation (Fig. 5B). A more systematic investigation of a second model oligonucleotide (CCACAACCAUGGUGAGCAA) was conducted. This is a generically designed oligo that has a hairpin with three regions (stem, loop, tail). The hypothesis was to show at which position RNA methylations have a particular effect on the MST shift For this purpose, the adenine positions of the unmodified parent oligo were systematically methylated either at the N^1 or N^6 position of an adenine base (Fig. 6), or the 2′-OH group of ribose was methylated (Fig. S6†).

In contrast to the results shown in Fig. 5, the detection of adenine methylation with Syto 59 led to more sensitive changes in F_{norm} values than with SybrGold (Fig. 6C). In particular, m^1A

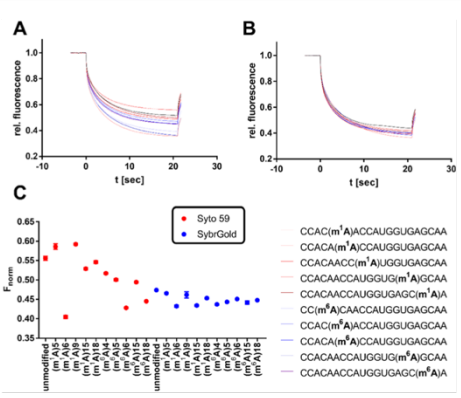


Fig. 6 Analysis of various modifications of the model oligo CCACAACCAUGGUGAGCAA (100 nM) by MST. (A) Thermophoresis curves for the usage of Syto 59 as a reporter dye. (B) Thermophoresis curves for the usage of SybrGold as a reporter dye. (C) MST response (F_{norm} at 10 s) of the modified oligos in comparison to their unmodified analogues.

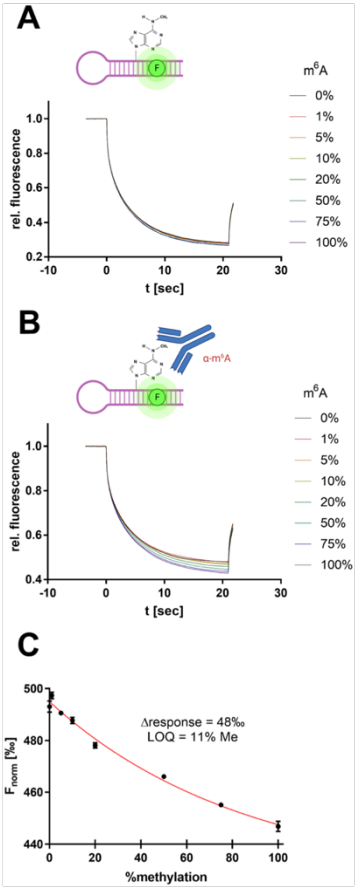


Fig. 7 Antibody-assisted detection of RNA m^6A modification by SybrGold MST. (A) MST traces for METTL3/14 substrate methylation levels (total RNA: 100 nM). (B) Same reaction mixture as in (A) but supplemented with 600 nM anti-m^6A antibody clone 9B7.[40] (C) MST response (F_{norm} at 10 s) as a function of substrate methylation derived from (B). The limit of quantitation (LOQ) was defined by 3σ around $F_{norm} = 0\%$.

© 2023 The Author(s). Published by the Royal Society of Chemistry

View Article Online
Edge Article Chemical Science

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

or m^6A methylation of the 6th adenine base resulted in significantly altered F_{norm} values compared to the unmodified oligo indicating that local influences of RNA modification either affect the binding of the fluorophore or these modifications modulate the global topology resp. folding of the RNA and thus lead to an altered thermophoresis behaviour. The detection of ribose methylation followed a similar trend, however, with this form of modification leading to generally lower changes in F_{norm} values (Fig. S6C†).

Development of an MST-based METTL3/14 enzyme assay

To derive a relevant biotechnological application from this model concept, an MST-based enzyme assay was developed for the RNA methyltransferase METTL3/14 allowing the tracking of the enzymatic insertion of a single methyl group to an RNA substrate. METTL3/14, which has recently become popular as a potential drug target, is an RNA methyltransferase that transfers a methyl group from SAM to the N^6 atom of an adenine base found in a GGACU motif.[37] Thus, we focused on the development of an enzymatic assay that can follow the 6-methyladenine (m^6A) methylation of the METTL3/14 substrate oligo AACUUAAUGUUUGCAUUGG(m^6A)CUUGAGUUA.

MST offers the potential to evaluate turnover rates of enzymatic reactions, especially in cases in which no direct detection *via* fluorescence reporter exists.[38,39] Therefore, we sought to determine the intrinsic difference of the substrate's m^6A modification in thermophoresis behaviour. For this purpose, we simulated different m^6A methylation levels from 0–100% by

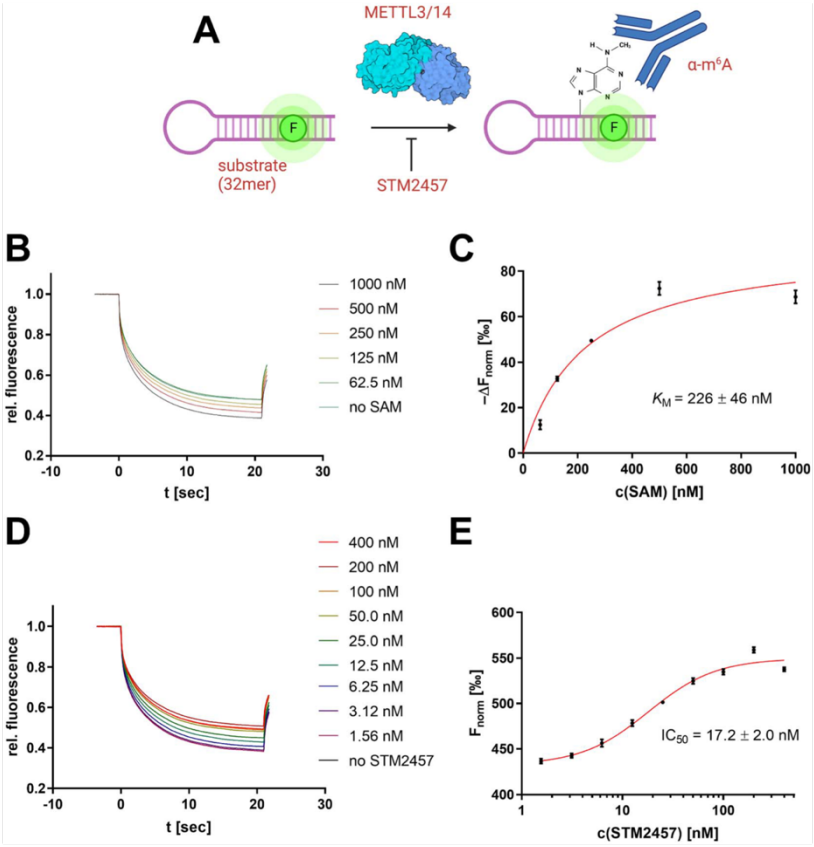


Fig. 8 Development of an MST-based METTL3/14 assay. (A) Illustration of the assay concept: a 32mer METTL3/14 substrate oligo is converted enzymatically which can be detected using an anti-m^6A antibody and SybrGold MST. (B) Enzymatic substrate methylation with varying concentrations of SAM as cofactor. After 2 h of incubation time, the crude reaction mixture was supplemented with the anti-m^6A antibody and analyzed by SybrGold MST. (C) Michaelis–Menten plot of the MST response (F_{norm} at 10 s) *vs.* the SAM concentration. A K_M value (226 nM) was calculated by non-linear regression. (D) MST curves of the enzymatic substrate methylation in the presence of STM2457 as METTL3/14 inhibitor. (E) Determination of the STM2457 IC_{50} value by dose–response fit of F_{norm} *vs.* [STM2457].

© 2023 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

mixing the synthetic substrate oligo with its methylated analogue at a total RNA concentration of 100 nM (Fig. 7A).

However, the use of 1× SybrGold and Syto 59 as a reporter dye resulted in very little intrinsic difference between the unmethylated and methylated oligos' F_{norm} values ($\Delta = 5‰$), so discrimination was not sufficient for the evaluation of enzymatic turnover. Therefore, we harnessed an antibody-coupled approach to amplify the F_{norm} discrimination signal. The addition of an m^6A-specific antibody (9B7 clone, 600 nM) resulted in a 10-fold dispersion ($\Delta = 48‰$, LOQ = 11% methylation) of the RNA oligos' F_{norm} (Fig. 7B and C).[40] This concept thus resembles an immuno-MST assay which has been previously described for non-RNA substrates but in this form claims novelty as an RNA methyltransferase assay.[41,42]

In this regard, the RNA substrate (120 nM), METTL3/14 enzyme (10 nM), and a variable concentration of the cofactor SAM (0–1000 nM) were incubated for 2 h at room temperature. Subsequently, the reaction was added with the anti-m^6A antibody (600 nM) and mounted in MST capillaries as a crude reaction mixture without further purification of the RNA substrate (Fig. 8A). The resulting thermophoresis curves showed a dependence of the F_{norm} response on the concentration of the SAM cofactor (Fig. 8B), which can be attributed to a different methylation content of the RNA substrate. The differential analysis of this F_{norm} shift at different SAM concentrations agreed with the calibration of the m^6A content (Fig. 7C) and indicated near quantitative methylation at the highest SAM concentration of 1 μM (Fig. 8C). The initial addition of SAH (10 μM) as a known inhibitor of methylation reaction yielded no shifts of the F_{norm} values (Fig. S7A†), confirming the effective inhibition of the enzyme and the overall assay concept. Michaelis–Menten regression of this differential methylation content (ΔF_{norm}) allowed the determination of an apparent K_M value for SAM of 226 nM which is in agreement with a ^{3}H-based enzyme assay (Fig. S7B†) and literature reports.[43]

To demonstrate the usefulness of the developed METTL3/14 assay for drug discovery applications, we also performed an inhibition experiment with the known literature inhibitor STM2457 (Fig. 8C), which revealed an apparent inhibition constant IC_{50} of 17.2 nM in agreement with the literature-determined value.[44]

Conclusions

In conclusion, we could show that the use of non-covalent nucleic acid dyes is suitable to determine some RNA modifications by a simple "mix-and-measure" protocol. Comparing *in situ* labelled RNA samples with orthogonal ligand binding assays, we found that the usage of non-covalent dyes mostly does not alter ligand binding of the RNA molecules, and hence, this new protocol might result in a wide range of opportunities for research applications on the function of RNA molecules. Meanwhile, small modifications that do not significantly affect F_{norm} could be enhanced by the use of a modification-specific antibody, as shown for the METTL3/14 assay, allowing the detection of minimal chemical modifications such as single-site methylation of larger RNA molecules. We hope to have laid the first step for the development of further RNA-immuno-MST assays and for numerous studies in this field.

Experimental section

DNA/RNA oligos

DNA/RNA oligos except for the ones prepared by *in vitro* transcription (IVT, see below) were purchased in HPLC-purified quality (for oligo sequences and manufacturers see Tables S1 and S2†). All experiments were performed under RNase-avoiding conditions. The concentrations of RNA, DNA, and enzymes were determined by a ThermoScientific NanoDrop2000.

RNA preparation

PCR. The DNA templates for the IVT were amplified using PCR. The PCR reaction mixture contained 10 nM of template DNA, 2 μM of each forward and reverse primer, 400 μM dNTP Mix (NEB), 3 mM $MgCl_2$, 0.05 U μL^{-1} Q5 High-Fidelity DNA Polymerase (NEB) and 1× Q5 reaction buffer (NEB) in a final volume of 200 μL. The reaction was performed using a ^{3}Prime thermal cycler with the following cycling conditions: initial denaturation (95 °C, 30 s), followed by 30 cycles of denaturation (95 °C, 30 s), annealing (temperature depends on the primer sequence and was determined by NEB T_M calculator, 30 s) and extension (72 °C, 45 s), and a final extension (72 °C, 2 min). The PCR product homogeneity was checked on a 0.4% agarose gel to confirm the size and purity.

***In vitro* transcription (IVT).** The crude PCR products were used as templates for *in vitro* transcription using the T7 polymerase system. The IVT reaction mixture contained 400 μL PCR product (without further purification), 5 mM of each NTP, 5 mM DTT, 2.5 μg mL^{-1} BSA, $MgCl_2$ (concentration depends on RNA type and was determined before using micro reactions with different $MgCl_2$ concentrations), 50 U mL^{-1} T7 RNA polymerase (ThermoFisher) and 1× reaction buffer (ThermoFisher) in a final volume of 1 mL. The reaction was incubated at 37 °C for 2 hours, and then, another aliquot of T7 RNA polymerase was added and incubated for 4 hours at 37 °C.

RNA extraction and purification. The formed pyrophosphate was removed by centrifugation at 14 000 g for 5 min directly after transcription and the cleared supernatant was transferred to a micro reaction tube for further sample processing. To remove the PCR- and IVT enzymes, a phenol/chloroform extraction was performed twice using ROTI® phenol/chloroform/isoamyl alcohol (Carl Roth) and the standard protocol from the supplier. For a better gel resolution, the extracted samples were desalted using Zeba spin desalting columns (ThermoFisher) following the standard protocol from the supplier. Furthermore, ethanol precipitation was applied to remove chloroform residues and concentrate the RNA. For this, the samples were mixed with 2.5× volume ethanol, 0.1× volume 3 M NaOAc (pH 4.7), and 1 μL glycogen (20 g L^{-1}, ThermoFisher), cooled at −80 °C for 2 hours, then centrifuged for 45 min at 4 °C and 16 000 g. The pellet was washed twice with

 © 2023 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

75% ice-cold ethanol and centrifuged again for 10 min at 4 °C and 16 000 g. The air-dried pellet was redissolved in a small volume of RNase-free water and an equal volume of 2× gel loading buffer (2× TBE, 67% formamide) was added.

For the purification of RNA, a native polyacrylamide gel electrophoresis was utilized. A 10% acrylamide gel solution was prepared using the ROTIPHORESE® DNA sequencing system and 10× TBE buffer. After sample loading, the electrophoresis separation was performed at 15 W until the dye markers of the ladder showed sufficient resolution in the separation range of interest. The RNA was visualized by UV shadowing (254 nm), extracted by cutting out the bands of interest, and the excised parts were aliquoted into small pieces and frozen at −20 °C for 30 min. The frozen sample bands were extracted overnight at 15 °C in a 0.3 M NaOAc (pH 4.7) solution and the samples were centrifuged for 10 min at room temperature and 10 000 g, the supernatant and the gel residues were added into 0.45 μm Nanosep (Pall corporation) tubes and centrifuged again for 10 min and 10 000 g. The flow-through was precipitated with ethanol as described above. The pellet was redissolved in RNase-free water and stored until usage at −20 °C. The RNA product homogeneity was checked on a native 10% polyacrylamide gel to confirm the size and purity.

Setup of a standard mix-and-measure RNA-MST experiment (non-covalent dyes)

The reaction mixtures used for a typical MST experiment contained 100 nM of the investigated RNA and 1× of the respective dye both dissolved in MST buffer (50 mM Tris–HCl pH 7.5, 100 mM KCl, and 25 mM $MgCl_2$) and supplemented with ligands or additives if needed. Previous to the measurement, the RNA stock solution was heated for 5 min at 75 °C and then cooled down to room temperature in the heating block.

MST experiments were carried out using a Monolith NT.115 instrument (NanoTemper Technologies) using Monolith standard capillaries. The instrument was calibrated according to the manufacturer's instructions. The laser type was chosen dye-dependently (NanoBlue laser for SybrGold and RiboGreen and PicoRed laser for Syto dyes). General settings were applied for all MST experiments as follows: manual temperature control: 25 °C, fluorescence measurement before MST: 4 s, MST (IR laser) on: 20 s, fluorescence after MST: 2 s, delay: 25 s. The laser power was adjusted to optimize the signal-to-noise ratio and the fluorescence signal. LED and MST power settings were chosen individually for each sample by adjusting the LED power to yield fluorescence signals of at least 300 units (blue laser) or 3000 units (red laser) and the MST power to achieve an appropriate thermophoretic response (standard setting: medium MST power). MST measurements were analyzed using the NT analysis software (version 1.5.41) and exported for statistical analysis and plotting in GraphPad Prism 7.01.

Advice for the establishment of new RNA–ligand assay systems

We recommend the following procedure for assaying new RNA targets: initially, the new experimental setup should be tested with 100 nM RNA and 1× SybrGold (for NanoBlue devices) or 1× Syto 59 (for PicoRed devices). Initial assays include four controls consisting of 100 nM RNA, 1× dye, and a known reference ligand ($c = 10\times K_D$) in a buffer that has been shown to be appropriate for the biological function of the RNA: (1) +RNA/+dye/+ligand, (2) +RNA/+dye/−ligand, (3) −RNA/+dye/−ligand, (4) −RNA/+dye/+ligand. Samples 3 and 4 should not show a significant thermophoresis curve due to the absence of RNA. Sample 1 should optimally result in an MST shift compared to Sample 2 as a function of reference ligand concentration.

If a ligand-dependent MST shift is not achieved with this initial setup, we recommend the following steps with the same assay setup: (1) variation of RNA dyes as described above, (2) variation of RNA concentration, (3) variation of thermophoresis settings (laser power, temperature, *etc.*).

MST experiments with 5′-Cy5-labelled RNA

MST experiments were performed on a Monolith NT.115 using standard uncoated capillaries. The 5′-Cy5-labelled RNA was heated to 75 °C for 5 min and cooled to room temperature for 60 min. The Cy5-labelled $PreQ_1$ RS and SAM-VI RS were diluted to 20 nM in MST buffer (50 mM Tris–HCl pH 7.5, 100 mM KCl, and 25 mM $MgCl_2$) and supplemented with ligands in appropriate concentrations. Data acquisition was performed identically to the setup of the non-covalently labelled MST experiments (*vide supra*). MST measurements were analyzed using the NT analysis software (version 1.5.41) and exported for statistical analysis and plotting in GraphPad Prism 7.01.

METTL3/14 MST assay (K_M & IC_{50}) determination

MST-based METTL3/14 assays were carried out by mixing the RNA substrate AACUUAAUGUUGCAUUGGACUUGAGUUA (Iba, final: 120 nM), the METTL3/14 enzyme (Biomol, final: 10 nM) and a variable concentration of the cofactor *S*-adenosylmethionine (Sigma Aldrich, final: 0–1000 nM) in 50 μL of assay buffer (20 mM Tris pH 7.5, 1 mM DTT, 0.01% Triton X-100, 50 mM KCl and 0.25 mM $MgCl_2$). After an incubation period of 2 h at room temperature, the reaction was quenched by the addition of SAH (Sigma Aldrich, final: 10 μM, K_I = 260 nM), added with the anti-m^6A antibody (clone 9B7, final: 600 nM), SybrGold (final: 1×), and mounted in MST capillaries in triplicates. The laser power was adjusted to 20% for the LED laser and "high" for MST laser. MST measurements were analyzed using the NT analysis software (version 1.5.41) and exported for statistical analysis and plotting in GraphPad Prism 7.01. The decrease in F_{norm} *vs.* [SAM] is a measure of the activity of the enzyme, and thus, the MST traces follow the Michaelis–Menten equation which was used for enzyme kinetic analysis: $v_0 = (v_{max} \times [S]/K_M + [S])$, where v_0 is the initial reaction rate, at a given substrate concentration [S]. v_{max} is the maximum reaction rate, while K_M indicates [S] at which the conversion rate is half-maximal.

To determine the IC_{50} value of the reference inhibitor STM2457, the assay was performed as described above with the following adaptations. 500 nM SAM was used as substrate. STM2457 was added as DMSO stock (final DMSO conc.: 3%).

© 2023 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

The enzyme reactions were incubated for 60 min at room temperature and quenched with SAH (10 μM). F_{norm} values *vs.* [STM2457] were evaluated using a 4-parameter Hill equation to determine an IC_{50} value: $y(F_{norm}) = \text{bottom} + ([\text{STM2457}]^{slope}) \times (\text{top} - \text{bottom})/([\text{STM2457}]^{slope} + IC_{50}^{slope})$.

METTL3/14 tritium-based assay

The radiometric METTL3/14 assay was carried out as described previously by mixing the RNA substrate AACUUAAUGUUGCAUUGGACUUGAGUUA (Iba, final: 200 nM), the METTL3/14 enzyme (Biomol, final: 20 nM) and a variable concentration of the cofactor ^{3}H-SAM (Hartmann Analytics 15 Ci mmol^{-1}, final: 0–1000 nM) in 10 μL of assay buffer (20 mM HEPES pH 7.5, 50 mM KCl, 250 μM $MgCl_2$, 1 mM DTT, 0.01% Tween-20).[43] The reactions were started by the addition of the cofactor ^{3}H-SAM and were carried out at room temperature for 30 min. Aliquots of 8 μL were taken from the reaction mixture and spotted on Whatman glass microfiber filters (GF/C, 25 mm, Cytiva). The RNA was precipitated on the filters in 5% ice-cold TCA (Sigma Aldrich) for 35 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) was added. Scintillation was measured on a scintillation counter (TriCarb Liquid Scintillation Analyzer 4810TR) with a measurement time of 1 min. Statistical analysis and plotting were performed in GraphPad Prism 7.01 as described above.

Data availability

The datasets supporting this article have been uploaded as part of the ESI.†

Author contributions

Elisabeth Kallert: formal analysis, methodology, validation, writing—original draft. Malte Behrendt: methodology, validation. Ariane Frey: methodology, formal analysis. Christian Kersten: conceptualization, writing—review & editing. Fabian Barthels: conceptualization, visualization, writing—review & editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Financial support by the German Research Foundation (DFG) in the framework of the Transregio Collaborative Research Center RMAP (TRR319 RMaP) is gratefully acknowledged. We thank Prof. Mark Helm (University of Mainz) for providing various RNA oligos and Prof. Gunter Meister (University of Regensburg) for providing the anti-m^{6}A antibody. The figures in this article were created with **https://www.biorender.com/**.

Notes and references

1 J. L. Childs-Disney, X. Yang, Q. M. R. Gibaut, Y. Tong, R. T. Batey and M. D. Disney, *Nat. Rev. Drug Discovery*, 2022, **21**, 736–762.
2 A. L. Garner, *ACS Med. Chem. Lett.*, 2023, **14**, 251–259.
3 M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske and S. Duhr, *Assay Drug Dev. Technol.*, 2011, **9**, 342–353.
4 S. E. Deeb, A. Al-Harrasi, A. Khan, M. Al-Broumi, G. Al-Thani, M. Alomairi, P. Elumalai, R. A. Sayed and A. E. Ibrahim, *Methods Appl. Fluoresc.*, 2022, **10**, 042001.
5 S. Duhr and D. Braun, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 19678–19682.
6 M. H. Moon, T. A. Hilimire, A. M. Sanders and J. S. Schneekloth, *Biochemistry*, 2018, **57**, 4638–4643.
7 D. Jacob, K. Thüring, A. Galliot, V. Marchand, A. Galvanin, A. Ciftci, K. Scharmann, M. Stock, J.-Y. Roignant, S. A. Leidel, Y. Motorin, R. Schaffrath, R. Klassen and M. Helm, *Angew. Chem., Int. Ed.*, 2019, **58**, 9565–9569.
8 E. Kallert, T. R. Fischer, S. Schneider, M. Grimm, M. Helm and C. Kersten, *J. Chem. Inf. Model.*, 2022, **62**, 4134–4148.
9 S. Balaratnam, C. Rhodes, D. D. Bume, C. Connelly, C. C. Lai, J. A. Kelley, K. Yazdani, P. J. Homan, D. Incarnato, T. Numata and J. S. Schneekloth Jr, *Nat. Commun.*, 2021, **12**, 5856.
10 C. Entzian and T. Schubert, *Methods*, 2016, **97**, 27–34.
11 A. Chramiec-Głąbik, M. Rawski, S. Glatt and T.-Y. Lin, in *RNA–Protein Complexes and Interactions: Methods and Protocols*, ed. R.-J. Lin, Springer US, New York, NY, 2023, pp. 29–53.
12 M. Jerabek-Willemsen, T. André, R. Wanner, H. M. Roth, S. Duhr, P. Baaske and D. Breitsprecher, *J. Mol. Struct.*, 2014, **1077**, 101–113.
13 S. A. I. Seidel, P. M. Dijkman, W. A. Lea, G. van den Bogaart, M. Jerabek-Willemsen, A. Lazic, J. S. Joseph, P. Srinivasan, P. Baaske, A. Simeonov, I. Katritch, F. A. Melo, J. E. Ladbury, G. Schreiber, A. Watts, D. Braun and S. Duhr, *Methods*, 2013, **59**, 301–315.
14 P. J. Kolbeck, W. Vanderlinden, G. Gemmecker, C. Gebhardt, M. Lehmann, A. Lak, T. Nicolaus, T. Cordes and J. Lipfert, *Nucleic Acids Res.*, 2021, **49**, 5143–5158.
15 R. Silvers, H. Keller, H. Schwalbe and M. Hengesbach, *ChemBioChem*, 2015, **16**, 1109–1114.
16 Y. T. Maeda, T. Tlusty and A. Libchaber, *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, 17972–17977.
17 T. Bartoschik, S. Galinec, C. Kleusch, K. Walkiewicz, D. Breitsprecher, S. Weigert, Y. A. Muller, C. You, J. Piehler, T. Vercruysse, D. Daelemans and N. Tschammer, *Sci. Rep.*, 2018, **8**, 4977.
18 A. R. Gruber, R. Lorenz, S. H. Bernhart, R. Neuböck and I. L. Hofacker, *Nucleic Acids Res.*, 2008, **36**, W70–W74.
19 C. M. Connelly, M. H. Moon and J. S. Schneekloth, *Cell Chem. Biol.*, 2016, **23**, 1077–1090.
20 K. E. Deigan and A. R. Ferré-D'Amaré, *Acc. Chem. Res.*, 2011, **44**, 1329–1338.
21 B. D. Blakeley, S. M. Deporter, U. Mohan, R. Burai, B. S. Tolbert and B. R. Mc Naughton, *Tetrahedron*, 2012, **68**, 8837–8855.

© 2023 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 29 August 2023. Downloaded on 3/21/2024 8:27:55 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

22 C. P. Jones, G. Piszczek and A. R. Ferré-D'Amaré, *Methods Mol. Biol.*, 2019, **1964**, 75–87.
23 R. T. Batey, *Wiley Interdiscip. Rev.: RNA*, 2011, **2**, 299–311.
24 A. Serganov, *Curr. Opin. Struct. Biol.*, 2009, **19**, 251–259.
25 R. R. Breaker, *Biochemistry*, 2022, **61**, 137–149.
26 R. K. Montange, E. Mondragón, D. van Tyne, A. D. Garst, P. Ceres and R. T. Batey, *J. Mol. Biol.*, 2010, **396**, 761–772.
27 A. Sun, C. Gasser, F. Li, H. Chen, S. Mair, O. Krasheninina, R. Micura and A. Ren, *Nat. Commun.*, 2019, **10**, 5728.
28 C. M. Connelly, T. Numata, R. E. Boer, M. H. Moon, R. S. Sinniah, J. J. Barchi, A. R. Ferré-D'Amaré and J. S. Schneekloth, *Nat. Commun.*, 2019, **10**, 1501.
29 J. T. Polaski, S. M. Webster, J. E. Johnson and R. T. Batey, *J. Biol. Chem.*, 2017, **292**, 11650–11658.
30 A. Ren, K. R. Rajashankar and D. J. Patel, *Nature*, 2012, **486**, 85–89.
31 H. Gustmann, A.-L. J. Segler, D. B. Gophane, A. J. Reuss, C. Grünewald, M. Braun, J. E. Weigand, S. T. Sigurdsson and J. Wachtveitl, *Nucleic Acids Res.*, 2019, **47**, 15–28.
32 A. Jeltsch, A. Ehrenhofer-Murray, T. P. Jurkowski, F. Lyko, G. Reuter, S. Ankri, W. Nellen, M. Schaefer and M. Helm, *RNA Biol.*, 2017, **14**, 1108–1123.
33 S. Johannsson, P. Neumann, A. Wulf, L. M. Welp, H.-D. Gerber, M. Krull, U. Diederichsen, H. Urlaub and R. Ficner, *Sci. Rep.*, 2018, **8**, 8880.
34 N. Jonkhout, J. Tran, M. A. Smith, N. Schonrock, J. S. Mattick and E. M. Novoa, *RNA*, 2017, **23**, 1754–1769.
35 M. Helm, M. C. Schmidt-Dengler, M. Weber and Y. Motorin, *Adv. Biol.*, 2021, **5**, 2100866.
36 Y. Motorin and M. Helm, *Biochemistry*, 2010, **49**, 4934–4944.
37 F. Fiorentino, M. Menna, D. Rotili, S. Valente and A. Mai, *J. Med. Chem.*, 2023, **66**, 1654–1677.
38 H. Yu and Q. Zhao, *Anal. Chem.*, 2023, **95**, 2152–2156.
39 A. Topf, P. Franz and G. Tsiavaliaris, *BioTechniques*, 2017, **63**, 187–190.
40 F. Weichmann, R. Hett, A. Schepers, T. Ito-Kureha, A. Flatley, K. Slama, F. D. Hastert, N. B. Angstman, M. C. Cardoso, J. König, S. Hüttelmaier, C. Dieterich, S. Canzar, M. Helm, V. Heissmeyer, R. Feederle and G. Meister, *RNA*, 2020, **26**, 1489–1506.
41 C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun and S. Duhr, *Nat. Commun.*, 2010, **1**, 100.
42 S. Lippok, S. A. I. Seidel, S. Duhr, K. Uhland, H.-P. Holthoff, D. Jenne and D. Braun, *Anal. Chem.*, 2012, **84**, 3523–3530.
43 S. M. Buker, Z. A. Gurard-Levin, B. D. Wheeler, M. D. Scholle, A. W. Case, J. L. Hirsch, S. Ribich, R. A. Copeland and P. A. Boriack-Sjodin, *SLAS Discovery*, 2020, **25**, 361–371.
44 E. Yankova, W. Blackaby, M. Albertella, J. Rak, E. De Braekeleer, G. Tsagkogeorga, E. S. Pilka, D. Aspris, D. Leggate, A. G. Hendrick, N. A. Webster, B. Andrews, R. Fosbeary, P. Guest, N. Irigoyen, M. Eleftheriou, M. Gozdecka, J. M. L. Dias, A. J. Bannister, B. Vick, I. Jeremias, G. S. Vassiliou, O. Rausch, K. Tzelepis and T. Kouzarides, *Nature*, 2021, **593**, 597–601.

© 2023 The Author(s). Published by the Royal Society of Chemistry

Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2023

Non-covalent Dyes in Microscale Thermophoresis for Studying RNA Ligand Interactions and Modifications

Elisabeth Kallert,[a] Malte Behrendt,[a] Ariane Frey,[a] Christian Kersten,[a] Fabian Barthels*[a]

[a] E. Kallert, M. Behrendt, A. Frey, Dr. C. Kersten, Dr. F. Barthels
Institute of Pharmaceutical and Biomedical Sciences
Johannes Gutenberg-University of Mainz
Staudingerweg 5, 55128 Mainz, Germany
E-mail: barthels@uni-mainz.de

RNA and DNA oligo sequences

Table S1. Oligos used for RNA-MST experiments. **Category 1**: no structure, <10mer; **category 2**: no structure, 10–15mer; **category 3**: no structure, 16–20mer; **category 4**: hairpin, 16–20mer; **category 5**: bulge/internal loop + hairpin, 21–25mer; **category 6**: internal loop + hairpin, 26–30mer; **category 7**: bulge + internal loop + hairpin(s), 31–50mer; **category 8**: 2 internal loops + 1 hairpin, 50mer; **category 9**: 3 internal loops + 1 hairpin, 60mer; **category 10**: various functional RNA.

OligoID	Cat.	RNA sequence	Manufacturer
MH993	1	GUUAUUC	Iba
MH996	1	UUUUAUUUU	Biomers
MH474	1	GGUACCCAA	Iba
MH122	2	ACAGCUAUCCAUUG	Iba
MH847	2	UCUGAGGGUCCAGGG	Biomers
MH268	2	CCCGGUAAUCGCAUA	Salifu
MH266	3	AAACUUAAAACUUUACUGU	Salifu
MH907	3	ACAAGAAGGGUUUGAUCAUC	Iba
MH272	3	AGAAAUAUGUCUGAUAAA	Salifu
MH355	4	CGCGGGAGACCGGGGUUCGA	Dharmacon
MH261	4	GCUUGAAACCAGCUUUGGG	Salifu
MH1428	4	GGAUCCUCUAGAGUCGACCU	Idt
MH515	5	GGUGGGGUUCCCGAGCGGCCAAAG	Iba
MH499	5	GUUCGAUCCACAGAAUUCGCACCA	Biomers
MH649	5	GGUUCGAUUCCCCGACGGGGAGCCA	Iba
MH857	6	UCUAGAGGAUCCCGGGUACCGAGCUC	Biomers
MH876	6	UGCCGCCCGUCUUGAAACACGGACCAAGGA	Iba
MH1145	6	AACUUAAUGUUUUUGCAUUGGACUUUGAGUUA	Iba
MH761	7	GGCGUGUAGGAUAUUAAAUCAUACCAGAUCGCGG	Iba
MH646	7	UCCUCGUUAGUAUAGUGGUGAGUAUCCCCGCCUAUAUC	Iba
MH566	7	GUCACGCGGGAGACCGGGGUUCGAUUCCCCGACGGGGAGCCA	Iba
MH723	8	CCGCGAUCUGGAGAGGUGAAGAAUACGACCACCUACAAGAAGGACACGCC	Iba
MH721	8	CCGCGAUCUGGAGAGGUGAAGAAUACGACCACCUAAGAAGGACACGCC	Iba
MH762	8	CCGCGAUCUGGAGAGGUGAAGAAUUCGACCACCUACCAGAAGGACACGCC	Iba
MH950	9	CGGAGCCAGCGAGUCUAACCUUGGCCGAGAGGUCUUGGUAAUCUUGUGAAACUC CGUCGU	Iba

1

MH952	9	GGCCUCAGGAUCUGCUUAGAGAAGGGGGCAACUCCAUCUCAGAGCGGAGAAUUUGGACAA	Iba
MH953	9	UCUUGGAUUUAUUGAAGACUAACUACUGCGAAAGCAUUUGCCAAGGACGUUUUCAUUAAU	Iba
tRNA^{Asp}	10	UCCUCGUUAGUAUAGUGGUGAGUAUCCCCGCCUGUCACGCGGGAGACCGGGGUUCGAUUCCCCGACGGGGAG	In house IVT
PreQ$_1$ RS (*Thermoanaerobacter tengcongensis*)	10	GGAGAGGUUCUAGCUACACCUCUAUAAAAAACUAA	biomers
SAM-I RS (*Thermoanaerobacter tengcongensis*)	10	GGCUUAUCAAGAGAGGGGGAGCGACUGGCGCGAAGAACCCCGGCAACCAGAAAUGGUGCCAAUUCCUGCAGCGGAAACGUUGAAAGAUGAGCCG	In house IVT
SAM-VI RS (*Bifidobacterium angulatum*)	–	GGCAUUGUGCCUCGCAUUGCACUCCGCGGGGCGAUAAGUCCUGAAAAGGGAUGUC	In house IVT
Cy5-PreQ$_1$ RS (*Bacillus subtilis*)	–	Cy5-AGAGGUUCUAGCUACACCCUCUAUAAAAAACUAA	Eurofins Genomics
Cy5-SAM-VI RS (*Bifidobacterium angulatum*)	–	Cy5-GGCAUUGUGCCUCGCAUUGCACUCCGCGGGGCGAUAAGUCCUGAAAAGGGAUGUC	Eurofins Genomics
AdoCbl riboswitch (Marine metagenome *env8* 4WJvar)	–	GGCCUGAAAGCGUGGUGGGGAAGUGACGUGAAAUUCGUCCAGAUUACUUGAUACGGUUAUACUCCGAGUGCCACCCAGGCCAUACAACGAGCAAGGAGACUC	GenScript
Fluoride riboswitch (*Thermotoga petrophila*)	–	GGGCGAUGAGGCCCGCCCAAACUGCCCUGAAAAGGGCUGAUGGCCUCUACUG	GenScript
Neomycin aptamer (synthetic)	–	GGCUGCUUGUCCUUUAAUGGUCCAGUC	GenScript
METTL3/14 substrate	–	AACUUAAUGUUGCAUUGGACUUGAGUUA	Iba

2

Table S2. DNA templates and primers used for PCR and IVT.

OligoID	description	DNA sequence	Manufacturer
MH53	Universal forward primer	CGCGCGAAGCTTAATACGACTCACTATA	IDT
EK0004	Reverse primer SAM-I RS	CGGCTCATCTTTCAACGTTTCCG	IDT
EK0010	Reverse primer SAM-VI RS	GACATCCCTTTTCAGGACTTATCGCCCCG	IDT
MH1102	Reverse primer $tRNA^{Asp}$	TGGCGGGCCGTCG	Iba
EK0003	Template SAM-I RS	CGGCTCATCTTTCAACGTTTCCGCTGCAGGAATTGGCACCATTTCTGGTTGCCGGGGTTCTTC GCGCCAGTCGCTCCCCCTCTCTTGATAAGCCTATAGTGAGTCGTATTA	IDT
EK0009	Template SAM-VI RS	GACATCCCTTTTCAGGACTTATCGCCCCGCGGAGTGCAATGCGAGGCACAATGCCTATAGT GAGTCGTATTA	IDT
MH1101	Template $tRNA^{Asp}$	TGGCGGGCCGTCGGGGAATCGAACCCCGGTCTCCCGCGTGACAGGCGGGGATACTCACCA CTATACTAACGACCCTATAGTGAGTCGTATT	Iba

Supplementary Figures

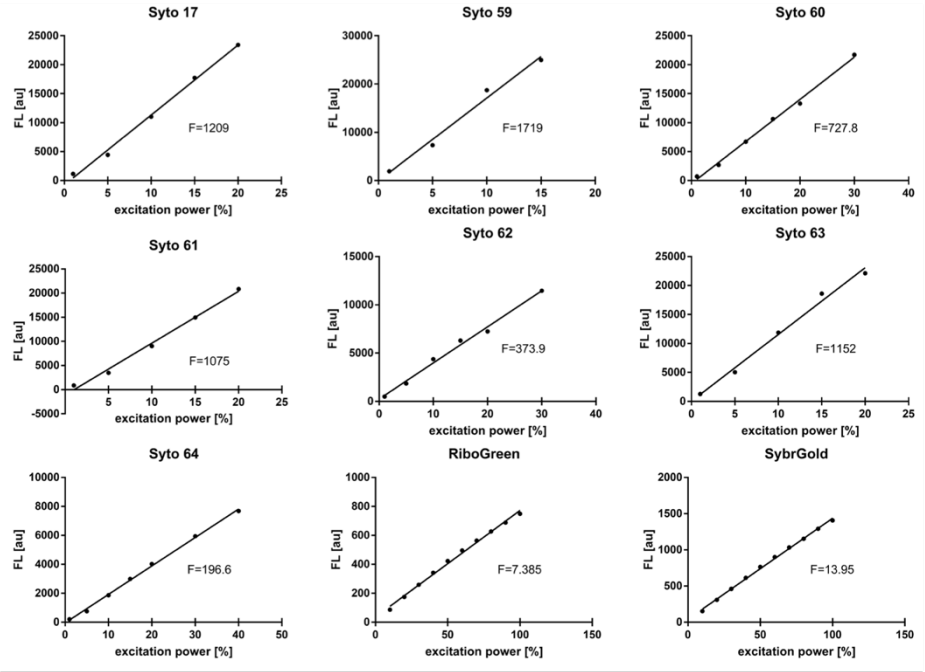


Figure S1. Regression plots for dye fluorescence normalization. Each dye (1×) was mixed with 100 nM of MH646 as RNA model oligonucleotide (UCCUCGUUAGUAUAGUGGUGAGUAUCCCCGCCUAUAUC). The initial fluorescence intensity as a function of the excitation power was determined pre-MST with either the NanoBlue laser for SybrGold and RiboGreen or the PicoRed laser for the Syto dyes. Linear regression yielded a slope-derived sensitivity factor (F) for the normalization of fluorescence intensity during structure-activity-relationship experiments.

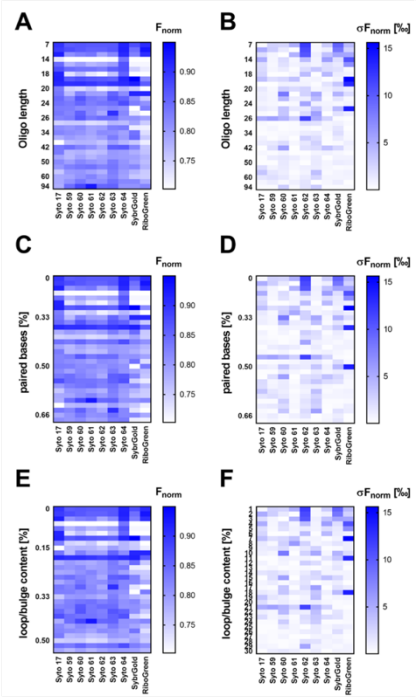


Figure S2. Structural RNA-dependent effects on F_{norm}. (**A**) Correlation between oligo length and F_{norm}. (**B**) Correlation between oligo length and σF_{norm}. (**C**) Correlation between hybridized base content and F_{norm}. (**D**) Correlation between hybridized base content and σF_{norm}. (**E**) Correlation between loop/bulge content and F_{norm}. (**F**) Correlation between loop/bulge content and σF_{norm}.

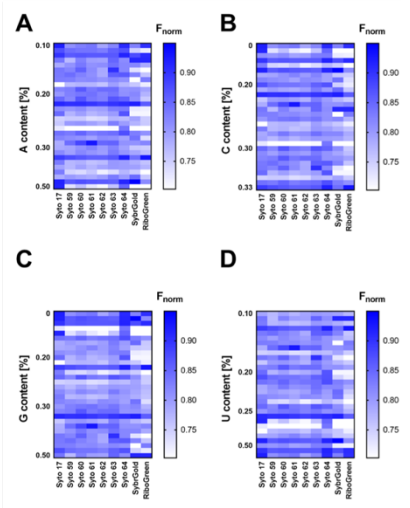


Figure S3. RNA sequence-dependent effects on F_{norm}. (**A**) Correlation between adenine content and F_{norm}. (**B**) Correlation between cytosine content and F_{norm}. (**C**) Correlation between guanine content and F_{norm}. (**D**) Correlation between uridine content and F_{norm}.

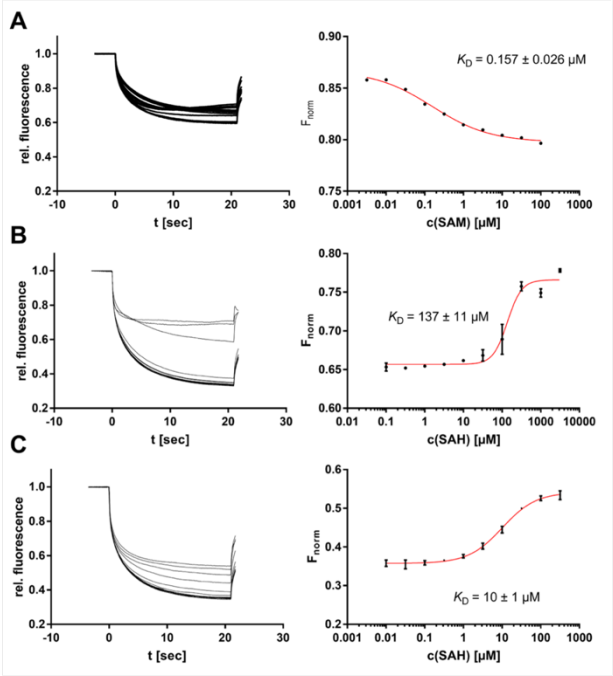


Figure S4. Analysis of RNA-ligand interaction by in situ labelling with SybrGold (1×). (**A**) Interaction between SAM-I RS and SAM (F_{norm} at 1 s). (**B**) SAM-I RS and SAH (F_{norm} at 1 s). (**C**) SAM-VI RS and SAH (F_{norm} at 15 s).

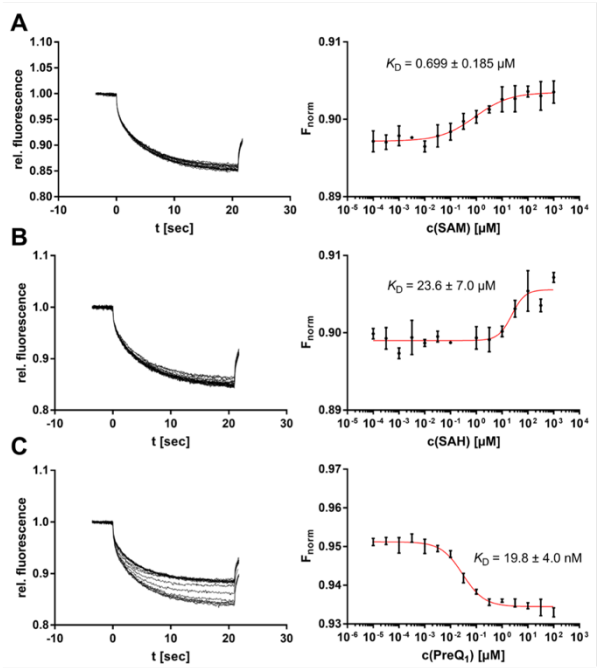


Figure S5. Analysis of RNA-ligand interaction by Cy5-labelled RNA-MST (**A**) Interaction between SAM-VI RS and SAM (F_{norm} at 5 s). (**B**) SAM-VI RS and SAH (F_{norm} at 5 s). (**C**) $PreQ_1$ RS and $PreQ_1$ (F_{norm} at 1 s).

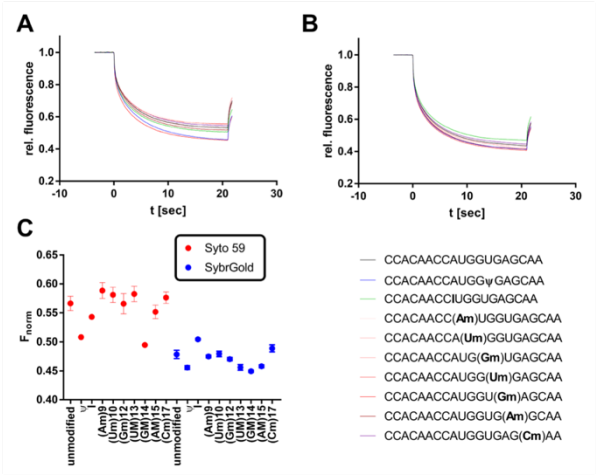


Figure S6. Analysis of various modifications of the model oligo CCACAACCAUGGUGAGCAA (100 nM) by MST. (**A**) Thermophoresis curves for the usage of Syto 59 as a reporter dye. (**B**) Thermophoresis curves for the usage of SybrGold as a reporter dye. (**C**) MST response (F_{norm} at 10 s) of the modified oligos in comparison to their unmodified analogues.

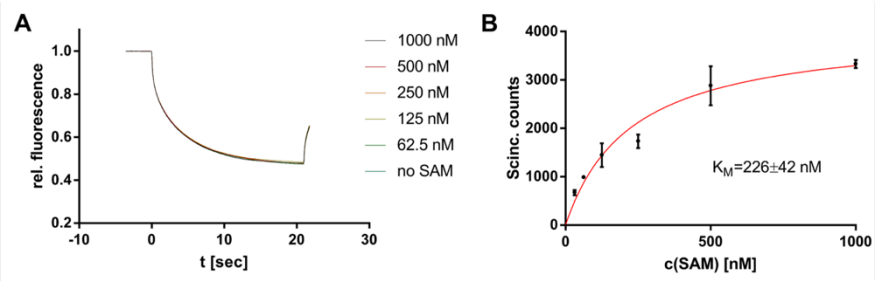


Figure S7. Supplementary data for the establishment of the METTL3/14 assay. (**A**) Thermophoresis curves for the METTL3/14 reaction inhibited by the initial addition of SAH (10 µM). Enzymatic substrate methylation with varying concentrations of SAM as the cofactor. After 2 h of incubation time, the crude reaction mixture was supplemented with the anti-m^{6}A antibody and analyzed by SybrGold MST. (**B**) Tritium-based METTL3/14 methyltransferase assay. The Michaelis-Menten plot was derived from radiometric analysis of ^{3}H incorporation into RNA substrate AACUUAAUGUUUGCAUUGGACUUGAGUUA. A K_M value (226 nM) was calculated by non-linear regression (Michaelis Menten).

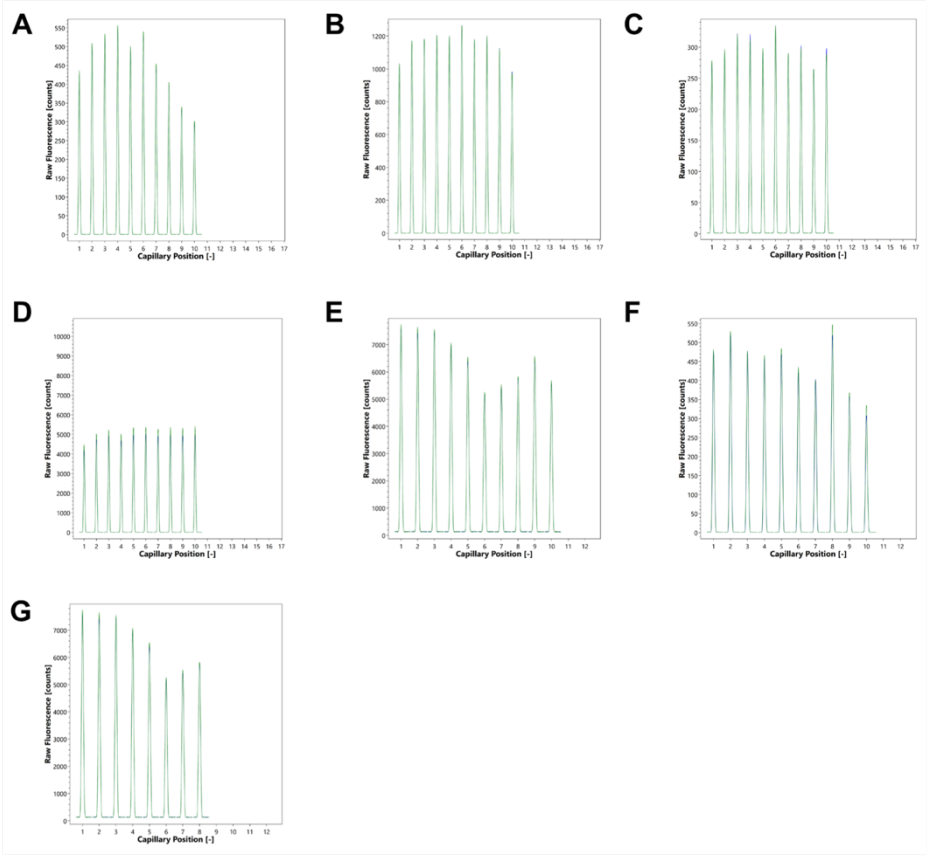


Figure S8. Capillary scans of RNA-ligand interaction samples for the assessment of dye displacement by ligand binding. Capillary scans are presented with increasing concentration corresponding to the MST curves in Figure 4 and Figure S4. (**A**) $PreQ_1$ RS vs. $PreQ_1$. (**B**) SAM-I RS vs. SAM. (**C**) SAM-VI RS vs. SAM. (**D**) Fluoride RS vs. potassium fluoride. (**E**) AdoCbl RS vs. adenosyl cobalamin. (**F**) DNMT2 vs. tRNA. (**G**) Neomycin aptamer vs. neomycin.

”

5. Biophysical Investigation of HCV IRES-Ligand Interactions

5.1. Context, Project Summary, and Own Contribution

Given the continued prevalence of severe viral infections such as chronic hepatitis C, which is exacerbated by the emergence of drug resistance, there is an urgent need for novel strategies in antiviral drug discovery.[216,217] Similarly, as in bacteria, viral RNA also contains highly conserved regions which can serve as potential targets for small molecules.[16] We previously demonstrated that the integration of computational and biophysical experimental techniques provides a robust pipeline for RNA-targeted drug discovery (Chapter 3). Nevertheless, this approach is still in its nascent stages. Consequently, we employed this methodology to identify and characterize small molecules targeting the HCV IRES structure. This RNA motif represents an optimal model system as it is highly conserved, essential for the viral replication system, and has a solved crystal structure. We validated the top hits from the VS of RNA-focused and non-biased libraries by FRET and MST assay, ensuring their efficacy in binding to the HCV IRES. The distinctive ligand-induced conformational change of this RNA motif was exploited to study its interactions with small molecules by FRET assay. Furthermore, we performed an MST assay to corroborate the findings of the FRET assay. We identified ten selective hit molecules with potencies in the high to medium micromolar range. Among those compounds, we identified a 2-guanidino-quinazoline moiety as a suitable motif for selective binding to the HCV IRES. Compared to previously reported binders, those molecules provide moderate lipophilicity and reduced basicity making them a potential starting point for further optimization campaigns. In this regard, we were able to contribute to the field of structure-based small molecule design against highly relevant viral RNA targets.

Own contribution: MST and FRET assay establishment, screening and affinity measurements, MST selectivity assay, FRET competition selectivity assay, data evaluation, HPLC-ESI/MS purity determinations, writing of the original draft & editing of the manuscript.

Contributions of other authors: Virtual screening, MST and FRET assay screening and affinity measurements, MST selectivity assay, data evaluation, writing of the original draft & editing of the manuscript.

Elisabeth Kallert and **Laura Almena-Rodriguez** contributed equally to this manuscript.

This work has been published in RSC Medicinal Chemistry (impact factor: 4.1).

Article reprinted with permission from *RSC Med. Chem.* **2024**. ‘Structure-based virtual screening of unbiased and RNA focused libraries to identify new ligands for the HCV IRES model system.’ ©2024 The Royal Society of Chemistry.

5.2. Publication

The following publication quoted (within “”) from page 156 to page 239 is exactly the same as the manuscript cited on page 77: “

RSC Medicinal Chemistry

ROYAL SOCIETY OF CHEMISTRY

RESEARCH ARTICLE

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/d3md00696d

Structure-based virtual screening of unbiased and RNA-focused libraries to identify new ligands for the HCV IRES model system†

Elisabeth Kallert, ‡[a] Laura Almena Rodriguez, ‡[a] Jan-Åke Husmann, [a] Kathrin Blatt [a] and Christian Kersten *[ab]

Targeting RNA including viral RNAs with small molecules is an emerging field. The hepatitis C virus internal ribosome entry site (HCV IRES) is a potential target for translation inhibitor development to raise drug resistance mutation preparedness. Using RNA-focused and unbiased molecule libraries, a structure-based virtual screening (VS) by molecular docking and pharmacophore analysis was performed against the HCV IRES subdomain IIa. VS hits were validated by a microscale thermophoresis (MST) binding assay and a Förster resonance energy transfer (FRET) assay elucidating ligand-induced conformational changes. Ten hit molecules were identified with potencies in the high to medium micromolar range proving the suitability of structure-based virtual screenings against RNA-targets. Hit compounds from a 2-guanidino-quinazoline series, like the strongest binder, compound **8b** with an EC_{50} of 61 μM, show low molecular weight, moderate lipophilicity and reduced basicity compared to previously reported IRES ligands. Therefore, it can be considered as a potential starting point for further optimization by chemical derivatization.

Received 8th December 2023,
Accepted 16th March 2024

DOI: 10.1039/d3md00696d

rsc.li/medchem

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Introduction

Hepatitis C virus infections

Hepatitis C virus (HCV) infection can cause acute and chronic inflammation of the liver leading to hepatic fibrosis and cirrhosis in most of the chronic cases. The risk for development of hepatocellular carcinoma (HCC) is significantly increased by chronic HCV infection.[1] HCC is the most frequent liver cancer, accounting for 70–80% of all cases, and is the fourth leading cause of cancer mortality.[2] Approximately 290 000 people died from the corresponding disease, hepatitis C (and mostly due to cirrhosis and HCC), in 2019 following WHO estimations.[3] Currently, around 58 million people are affected by chronic HCV infection worldwide.[3] Central Asia and central Africa have one of the highest HCV prevalence rates followed by other regions in Asia, Africa, Eastern Europe, Australia, Latin America, and the Middle East identified in a worldwide study.[4] HCV is a blood-borne virus, it bears high genetic diversity and produces diverse mutant clouds in infected people. Thus, development of an effective vaccine against hepatitis C was not successful so far and still is challenging. Current therapies target three significant non-structural (NS) protein components of the virus. The so-called class of direct acting antivirals (DAAs) belong either to protease or polymerase inhibitors and were subsequently approved for treating HCV infections starting in 2011. The combination of DAAs achieved up to 90% sustained virologic response, which is more than twice as good compared to previous treatments. However, the efficacy of DAAs is perturbed by the emerging resistance-associated amino acid substitution at drug target sites. Drug resistance can lead to treatment discontinuation or failure and is one of the major challenges in DAA treatment.[5]

HCV IRES

Discovering new treatment options and addressing HCV on the RNA-level could circumvent this emerging resistance sustainably. HCV is an enveloped positive single-stranded RNA virus of the *Flaviviridae* family discovered in 1989.[2] The translation of its RNA genome is regulated by an internal ribosome entry site (IRES), which is located in the 5′-untranslated region (UTR). IRES represents a virus' “trick” to accelerate the translation of its genome by avoiding most host cell initiation factors. It directly binds the 40S ribosomal subunit and requires only the eukaryotic initiation factors (eIF) 2 and eIF3 to assemble into translation-competent

[a] *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany.*
E-mail: kerstec@uni-mainz.de
[b] *Institute for Quantitative and Computational Biosciences, Johannes Gutenberg-University, BioZentrum I, Hanns-Dieter-Hüsch-Weg 15, 55128 Mainz, Germany*
† Electronic supplementary information (ESI) available. See DOI: **https://doi.org/10.1039/d3md00696d**
‡ EK and LAR contributed equally to this manuscript.

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

ribosomes directly at the start codon.[6] Thus, translation is initiated through a 5′-cap-independent mechanism. The 5′ UTR of the HCV can be divided into four distinct domains (I–IV), which are connected by flexible linkers. The stem loop in domain I is not essential for IRES activity and domain IV contains the start codon, which becomes exposed during initiation assembly.[7] Domains II and III are crucial for IRES activity, in which domain III represents the core domain binding to eIF3 and ribosomal subunits.[6] The release of eIF2 is mediated by domain II which also participates in formation of translation-competent ribosomes.[8] Structures within domain II to IV can independently fold and the resulting three-dimensional organization of the IRES is crucial for its activity.[7] The HCV IRES is highly structured, essential for virus survival and is the most sequence-conserved region of the HCV genome, thus being a promising drug target. The development of selective translation inhibitors offers new opportunities for less resistance prone treatment options.[9]

Small molecule HCV translation inhibitors

Inhibitors targeting different domains of the HCV IRES were identified since its discovery ranging from oligonucleotide-based inhibitors[10] and peptidic inhibitors[11] to small molecules.[7] The field of small molecules covers different substance classes like phenazines[12] (**1**), biarylguanidines[13] (**2**), diamino-piperidines[14] (DAP, **3**) and 2-aminobenzimidazoles[15] (**4**, **5**) (Fig. 1). The initial 2-aminobenzimidazole hit (**4**) showed a K_D-value of 100 μM (determined by a mass spectrometry (MS)-based assay) and subsequent structure–activity-relationship (SAR) analysis and derivatization achieved benzimidazole derivatives with sub-micromolar affinities. One optimized inhibitor is **5** with a K_D-value of 0.86 μM which showed only minimal toxicity and antiviral activity in a cellular replicon assay.[15] Thus, **5** is the most potent inhibitor reported so far. Further SAR exploration around this scaffold resulted in equally or less potent aminobenzoxazoles[16] (**6**) and quinazolines[17] (**7**). The classes of 2-aminobenzimidazole, quinazolines and DAP-conjugates were identified to bind to the IIa subdomain of the HCV IRES. DAP conjugates show effects on the IRES conformation upon binding different from 2-aminobenzimidazoles and quinazolines.[14] X-ray crystallography studies[18,19] of the native IRES subdomain IIa showed that two closely spaced divalent metal ions stabilize a right-angled bend at the intersection of two helices resulting in an L-shape of the RNA (Fig. 2A). This conformation is essential for IRES-driven translation to position the IIb hairpin at the 40S ribosomal subunit.[18–20] Nuclear magnetic resonance (NMR) structural analysis of subdomain IIa in complex with 2-aminobenzimidazole **5** provided insights into an induced conformational switch. The inhibitor changes the approximately 90° bend to a more stretched shape with 23° widening of the interhelical angle. This provides a potential reason for translation inhibition (Fig. 2A).[8] In addition, X-ray analysis and comparison of subdomain IIa with and without

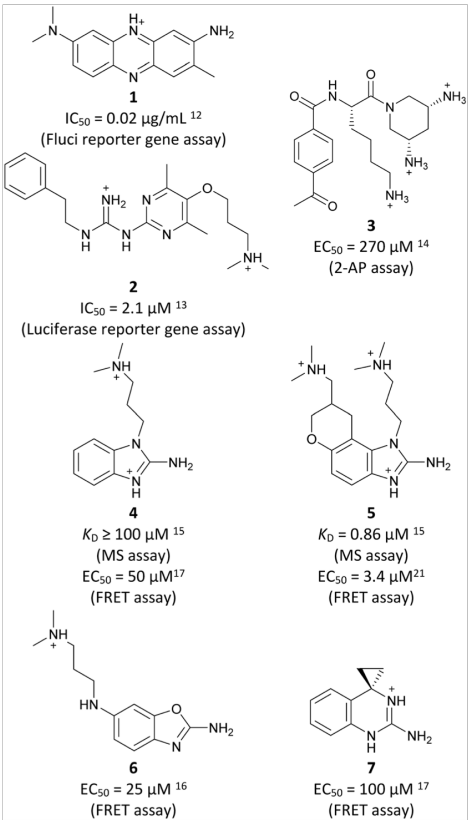


Fig. 1 Molecular structures, binding and inhibition results of previously reported HCV IRES ligands. Molecules are depicted as the dominant protomer under physiological conditions. FRET: Förster resonance energy transfer, 2-AP: 2-aminopurine.

5 bound revealed that three magnesium ions are present in both structures. They are regarded as intrinsic structural components which stabilize both, the free and the ligand-bound, states. The ions are not displaced upon ligand binding and are necessary for high-affinity interactions. Moreover, the ions undergo adaptive reorganization during the binding process.[21] In contrast, DAP conjugates bind in competition with structural divalent metal ions and immobilize the L-shape of the RNA, which might be an alternative mechanism of inhibition.[14] These conformational changes were monitored and confirmed *via* two different fluorescence assays.[14,22]

Binding of 2-aminobenzimidazole **5** to the highly conserved subdomain IIa induces a conformational change which forms a deep and well-defined binding pocket. An X-ray structure of subdomain IIa in complex with **5** at 2.2 Å resolution shows how the pocket encapsulates the ligand

 This journal is © The Royal Society of Chemistry 2024

View Article Online
RSC Medicinal Chemistry Research Article

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

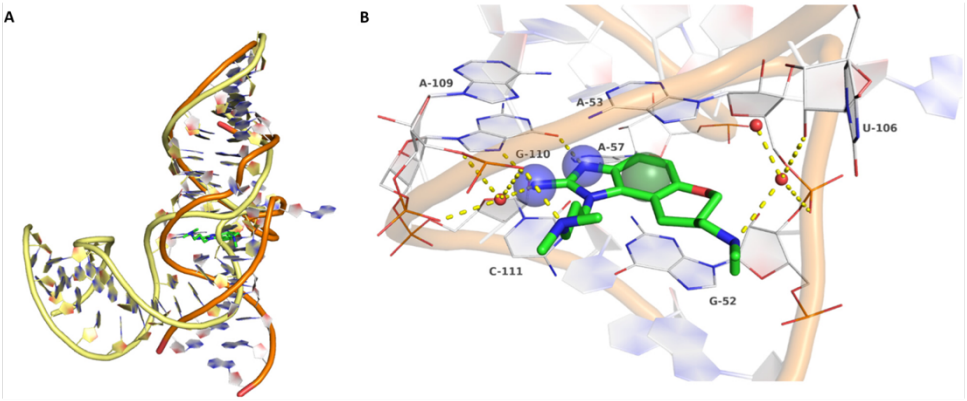


Fig. 2 A) Superposition of ligand-free, kinked (yellow backbone, PDB-ID: 1P5M) and compound **5**-bound, stretched (orange backbone, PDB-ID: 3TZR) HCV IRES subdomain IIa. B) Crystallographic binding mode of **5** (green carbon atoms) in complex with the HCV IRES subdomain IIa (white carbon atoms, PDB-ID: 3TZR). Direct and water-mediated (red spheres) polar interactions are depicted as yellow dashed lines. Key pharmacophore features are depicted as blue transparent spheres for hydrogen bond donors and a green sphere for an aromatic centre (pharmacophore feature radius 1.2 Å).

(Fig. 2B, PDB-ID 3TZR).[21] The roof and the floor of the binding pocket are formed by stacking interactions between A53 and G52 with the benzimidazole scaffold. Further important RNA–ligand interactions comprise hydrogen bonding of the 2-amino-imidazole group to the guanine Hoogsteen edge of G110 as well as hydrogen bonding of the protonated dimethylamino-propyl side chain to the phosphate group of A109. If these key interactions with G110 and A109 are missing, no or weak binding to the HCV IRES subdomain IIa was observed.[17] Thus, the 2-aminobenzimidazole and the dimethylaminopropyl functionality are crucial for this type of inhibitors. In addition, a basic substituent at the tetrahydropyran site is considered as beneficial and further substituents attached to the aromatic ring are not tolerated due to the tight fit into the pocket.[17]

The binding motif of 2-aminobenzimidazole translation inhibitors with the Hoogsteen edge of G110 and the intercalation between G52 and A53, may mimic interactions of arginine or guanosine with RNA nucleobases. A statistical evaluation of the amino acid–nucleotide interaction (AANT)[23] database showed that the most likely interaction between an amino acid and any RNA nucleotide includes arginine or lysine residues. In the case of arginine, interactions with cytidine or guanosine were occurring more often compared to the other nucleotides.[23] In addition, Hoogsteen pseudo pairs were found for G and arginine forming two hydrogen bonds. Arginine has three donor nitrogen atoms and is able to form four distinct pseudo pair geometries for both types of interactions. Interestingly, the Hoogsteen edge of A could be addressed by various amino acids, but the Hoogsteen edge of G was only recognized by arginine.[24] In a more concrete example, the human immunodeficiency virus Tat protein can bind specifically to an RNA stem-loop structure (TAR) mediated by a single arginine residue. Arginine-rich sequences can also inhibit translation. Evidence can be found in a similar case, in which an arginine attenuator peptide can regulate its own translation by interacting with ribosomes.[25] Arginine and G have been found to interact similarly with G:C base pairs.[26] Hydrogen bond networks between G:C pairs with arginine or C:G–G triplets are similar to the observed Hoogsteen edge interactions between C58: G110 and 2-aminobenzimidazole **5** (Fig. S1A–C†).[21,27,28] In a previous study, arginine, G and their derivatives were tested for HCV IRES subdomain IIa binding. While for arginine no binding could be observed, G and its derivatives showed some weak binding behavior (G, EC_{50} = 1.05 mM).[29] Like 2-aminobenzimidazoles, G and its derivatives induced a conformational change by stretching the L-shape of the RNA and inhibited IRES-driven translation. Other nucleobases were inactive. It was hypothesized that a guanosine of the same RNA strand located in domain IV autoregulates HCV IRES activity. At the beginning, trigger guanosine is sequestered in a hairpin loop in domain IV to allow ribosomal assembly. Positioning of the start codon at the decoding site induces movement in domain IV which unmask the trigger guanosine. Guanosine induces a conformational change by stretching subdomain IIa, which facilitates IRES release from the ribosome.[29]

In this work, a virtual screening was performed to identify new scaffolds with reduced basicity (compared to **5** and derivatives), binding to the HCV IRES subdomain IIa. This will not only obtain alternative molecules as starting points for optimization in terms of HCV DAA resistance-mutation preparedness, but also serve as a proof-of concept model system for structure-based RNA–ligand design for RNA

View Article Online
Research Article RSC Medicinal Chemistry

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

targets which do not have a native small molecule ligand like riboswitches.[30–34] The crystal structure of the HCV IRES subdomain IIa in complex with **5** and previous SAR studies were used as a starting point for the structure-based virtual screening.

Results

Structure-based virtual screening

Despite the increasing interest[35–38] in RNA-targeting small molecules, structure-based virtual screenings for RNA-targets are still in its infancy.[39–41] One challenge is the high flexibility of RNA, which is simplified in this study by using the RNA–ligand complex structure for docking, assuming similar conformational changes upon ligand binding as observed for compound **5**. Likewise, while unbiased screening libraries against RNAs showed hit-rates comparable to screenings against proteins,[33,34,42,43] several studies identified privileged scaffolds and physicochemical properties for RNA-targeted chemical spaces.[44–49] To account for this recent development, additionally to our unbiased in-house virtual library, also RNA-focused libraries of different chemical suppliers were separately considered in the virtual screening against the HCV IRES subdomain IIa. Except for Otava and Reaxense, the RNA-focused libraries showed only little overlap (Fig. S2A†). Physicochemical properties were comparable to the unbiased VS library and showed only slight differences in moderately increased numbers of aromatic and basic (N) atoms in line with previous observations from R-bind (Table S1†).[45,50] Higher formal charges were found for the Asinex and Enamine libraries, while reduced log*P*-values or higher numbers of hydrogen bond donors as described previously[50] for RNA–ligands, were not found. After further filtering for target-specific physicochemical properties (Table S2†), molecules were subjected to a two-step docking process (Fig. S2B†). The molecular docking workflow was validated by re-docking and binder *vs.* non-binder and decoy discrimination (Fig. S3A–I†). During prospective docking, not only docking scores, but also generated binding poses were evaluated. Especially matching of the crystallographic reference ligand's key interactions (Fig. 2B) was heavily weighted in this process. As indicated in previous SAR studies, aromatic stacking between G52 and A53, as well as two hydrogen bond donor functionalities to interact with the Hoogsteen edge of G110 are common themes in many binders. Exemplarily, the exchange of the 2-aminobenzimidazole (**5**) moiety to 2-aminobenzoxazole (**6**) resulted in a loss of potency[16] (Fig. 1) as the hydrogen bond donor functionality of the protonated 2-aminobenzimidazole is lacking in the less basic and non-protonated 2-aminobenzoxazole. 18 molecules (**8–25**, Table S3†) with high docking scores and best mimicking of the interaction profile of **5** were selected for *in vitro* testing. Notably, even though the fraction of molecules passing each triage step was higher for RNA-focused libraries compared to the unbiased library, in the end only molecules from the unbiased library were selected for testing. Some derivatives of obtained compounds were also present in the Enamine RNA-focused library, as well as molecules from this supplier in the unbiased library that were not part of the RNA-focused library. For the given example case of the HCV IRES subdomain IIa, this finding indicates that RNA-focusing can increase virtual hit rates, but current commercial libraries are too small to result in overall high hit numbers (Fig. S2B†).

Binding studies and SAR analysis

None of the molecules selected for *in vitro* testing were flagged as pan-assay interference compounds (PAINs).[51,52] However, so far there is no unified "RNA-PAINs" or promiscuous binder concept available. Therefore, HCV IRES subdomain IIa interactions were assessed with two independent assays. These assays were validated with reported HCV IRES ligands **4** and **5**. In a FRET-assay, the conformational change of subdomain IIa upon ligand binding (Fig. 2A) was monitored as described previously (Fig. S4, S6 and S7†).[22,53,54] In an MST-assay, binding affinity was measured (Fig. S5 and S8†). To exclude interference with the Cy3- and Cy5-labels, excitation and emission spectra of all compounds were recorded showing no interference. Identity and purity >95% of most compounds was verified by LC-MS and NMR analytics (Table S3, ESI:† 4. Analytical data of virtual screening hits). However, from the molecules identified as hits in the assays, **9**, **10** and **12** showed purity below 95%. After an initial screening at four different ligand concentrations (Fig. S4 and S5†), dose–response curves were recorded for compounds that showed binding in the initial screening (Table 1, Fig. S6 and S8†). In some cases, EC_{50}- and K_D-values could not unambiguously be determined due to weak binding affinity or limited solubility over 1 mM concentrations (indicated as a value lower limit *via* ≥).

To exclude promiscuous binding behavior, two selectivity assays were performed. First, interactions with an unrelated and assayable RNA, the *Thermoanaerobacter tengcongensis* (*Tte*) preQ$_1$-riboswitch aptamer domain, was evaluated for all compounds that showed binding to the HCV IRES *via* MST (Fig. S9A–O†). For compounds **5** and **9**, which showed binding to the preQ$_1$-riboswitch at very high concentrations, an additional selectivity MST assay with the *B. angulatum* SAM VI-riboswitch aptamer domain was conducted (Fig. S9P and R–T†). There, no MST shift and hence, no general RNA-binding promiscuity for these compounds was found. Second, the FRET assay was repeated for hit compounds in presence of a 5-fold (m/m) excess of competing total RNA (Fig. S10†). Six VS hits (**8–13**, Table 1, Fig. 3) were identified as specific binders of the HCV IRES subdomain IIa.

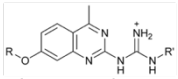
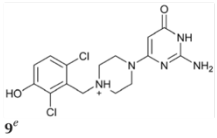
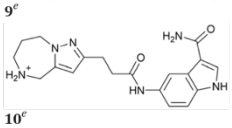
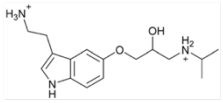
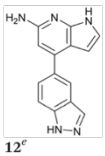
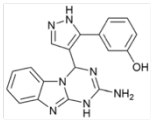
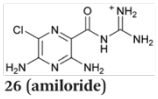
Compound **8** with EC_{50} = 413 μM and K_D ≥ 440 μM contains a 2-guanidino-pyrimidine moiety, a feature already reported in HCV IRES ligand **2** (Fig. 1) which is reminiscent of the hypothesized G110–arginine interaction (Fig. 3A and S1B†).[7] Consequently, further derivatives of compound **8** were

 This journal is © The Royal Society of Chemistry 2024

View Article Online
RSC Medicinal Chemistry Research Article

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 1 FRET and MST results for HCV IRES reference ligands **4** and **5**, virtual screening hits **8**–**13**, **8**-derivatives **8a**–**8c** and amiloride (**26**). Pre-screening, MST time-traces, dose–response curves, and selectivity assays are summarized in Fig. S4–S10.† Molecules are depicted in their dominant protomeric/tautomeric state under physiological conditions. For **12**, 1*H*-pyrimidine protomers/tautomers are also reasonable. N.b. no binding observed. Me: methyl, Et: ethyl

Compound	FRET EC_{50} [μM]	MST K_D[a] [μM]	Selectivity[b]
4	78 ± 35	x	✓[c]
5	4.9 ± 2.6	5.2 ± 0.8	✓[c]
			
8 (R: Et, R′: Me)	413 ± 97	≥440	✓
8a (R: Et, R′: H)	88 ± 12	189 ± 20	✓
8b (R: Me, R′: Me)	61 ± 19	88 ± 2	✓
8c (R: Me, R′: H)	120 ± 24	134 ± 43	✓
 **9**[e]	217 ± 41	n.b.	x[c] (K_D = 7.8 ± 3.3 μM $preQ_1$ riboswitch)
 **10**[e]	≥500[d]	n.b.	✓
 **11**	≥500	≥100	✓
 **12**[e]	≥300	≥50	✓
 **13**	≥1000	≥100	✓
 **26 (amiloride)**	≥500	≥100	✓

[a] MST binders are defined as molecules that clearly show a concentration-dependent MST-shift, but do not reach saturation due to solubility limitations and therefore no upper plateau in the dose–response curve. This is indicated by the presentation of K_D-values as a lower limit (≥). [b] Molecules are defined to be selective if no MST-shift is observed for the unrelated $preQ_1$-riboswitch (Fig. S9†) and the FRET signal in presence of competing total RNA is vastly unaffected (Fig. S10†). [c] **5** and **9** showed MST-shifts in the $preQ_1$-riboswitch MST assays at high concentrations, but not for SAM-VI-riboswitch binding. EC_{50}-values of **4** and **9** were reduced in presence of total RNA and did not reach saturation. [d] Based on decreased Cy5-emission. Cy3-emission did not increase for cpd. **10**. [e] Purity of **9**, **10** and **12** was below 95% determined by LC-MS and NMR analytics.

obtained. Derivatives **8a**–**8c** showed binding to the HCV IRES subdomain IIa in the MST assay, induction of the conformational change in the FRET assay, and selectivity both in the total RNA competition assay and $preQ_1$ binding assay (Fig. S6–S10†). From these molecules, **8b** showed the highest potency with EC_{50} = 61 μM and K_D = 88 μM. A small,

View Article Online

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

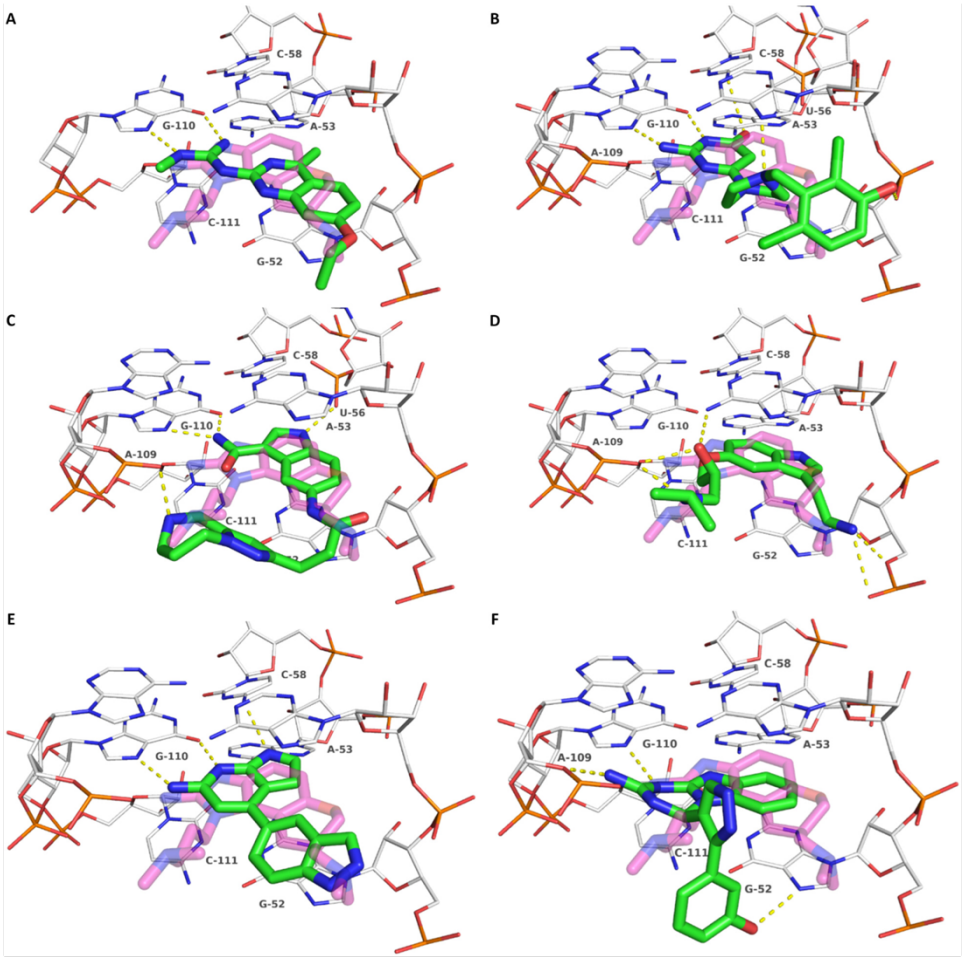


Fig. 3 Docking-predicted binding modes of **8** (A), **9** (B), **10** (C) and **11** (D), **12** (E) and **13** (F). Ligands are depicted with green carbon atoms, HCV IRES with white carbon atoms (PDB-ID: 3TZR). The crystallographic reference ligand (**5**) is shown with transparent, magenta-colored carbon atoms for orientation. Polar interactions are depicted as yellow dashed lines.

fragment-like size (18 heavy atoms, ligand efficiency LE 0.31 kcal mol^{-1}), moderate lipophilicity (log*P* 1.1,[55] lipophilic ligand efficiency LLE 2.9)[56] and reduced basicity (one basic center compared to three in compound **5** and 2-aminobenzimidazole derivatives) makes this molecule a potential starting point for further optimization. Guanidinium-containing molecules like in several amiloride derivatives, were repeatedly reported as RNA–ligands.[57–63] Adding amiloride (**26**) to the testing panel showed some weak activity with $EC_{50} \geq 500$ μM and $K_D \geq 100$ μM (Fig. S6L and S8L†) as well as selectivity over the preQ$_1$-riboswitch (Fig. S9O†). Camostat (**27**, Fig. S11A†), a benzylguanidine, revealed no binding to subdomain IIa (Fig. S4Y and S5Y†). This indicates that, despite their reoccurrence as RNA–ligands, guanidine moieties attached to aromatic systems might be considered as privileged scaffolds rather than promiscuous RNA-binders. Except compounds **8**, **8a–8c**, hits from the virtual screening show only low affinity in the FRET and MST assays. While still being selective over unrelated RNAs (Fig. S9 and S10†), only limited insights in interaction patterns with the HCV IRES subdomain IIa are possible from these compounds **9–13**. Compound **9** with EC_{50} = 217 μM, but no significant MST shift, contains a Watson–Crick-like edge as found in G within its 2-amino-1*H*-pyrimidin-6-one

 This journal is © The Royal Society of Chemistry 2024

View Article Online
RSC Medicinal Chemistry Research Article

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

substructure. Based on the predicted binding mode (Fig. 3B), this moiety mimics interactions as found for G–G:C interactions (Fig. S1C†). Further molecules containing this substructure (**14**, **15**, Table S3†) showed no binding. As the essential structural difference between **9** and **14**/**15**, the basic moiety of **9** in the tertiary aliphatic amine of the piperazine ring interacting with A53 is assumed. **14** and **15** contain no basic nitrogen, but an amide. Notably, differently from the other hits, **9** also binds to the preQ$_1$-riboswitch with a K_D-value of 7.8 μM (Fig. S9G and P†). This lack of selectivity can be explained by the preference of that riboswitch to G-like compounds.[31,64] In a recent study,[31] compound **28** (Fig. S11B†) with high similarity to **9** was identified as a preQ$_1$-riboswitch ligand. Cross-testing of **28** against the HCV IRES subdomain IIa, however, showed no binding (Fig. S5Z†). Additional selectivity tests of **9** and **28** against the unrelated SAM-VI-riboswitch also showed no binding (Fig. S9S and T†). Thus, **9** seems to be not generally promiscuous, but eventually preferred in RNA recognizing G-like Watson–Crick edges. This is also reflected in the FRET competition assay in presence of total RNA, where the EC_{50}-value of **9** is reduced, but the reached plateau for Cy5-fluorescence shows higher values, indicative for off-target RNA-binding. Compounds **10** and **11** share an indole core with two substituents at the 3- and 5-position connecting to one (**10**) or two (**11**) basic centres resulting in EC_{50}-values of ≥ 500 μM, respectively. Molecular docking poses show different indole orientations (Fig. 3C and D) and cannot satisfy all pharmacophore features. Predicted binding modes still resemble an overall similar shape to 2-aminobenzimidazole **5**. While **10**'s and **11**'s basic centers are close to the phosphate backbone of A109, the docking pose of **11** additionally shows an ionic interaction with the backbone of G52. Only **11**, but not **10** also showed binding in the MST-assay with a K_D-value of ≥ 100 μM. Compound **12** with an EC_{50} ≥ 300 μM and K_D-value ≥ 50 μM, contains a slightly basic 4-substitued pyrrolo[2,3-*b*] pyridine-6-imine moiety well meeting the desired pharmacophore features (Fig. 3E). 3,4-Dihydrobenzo[4,5] imidazo[1,2-*a*][1,3,5]triazin-2-amines (**13**, **16**–**22**, Table S3†), as well as benzamidines (**23** and **24**) are structurally related and reoccurring scaffolds among the best scoring molecules from the virtual screening. However, these showed no to very weak HCV IRES subdomain IIa binding. Only compound **13** from the 3,4-dihydrobenzo[4,5]imidazo[1,2-*a*][1,3,5]triazin-2-amine series (Fig. 3F, S6K and S8H†) showed very weak binding in both FRET (EC_{50} ≥ 1 mM) and MST (K_D ≥ 100 μM) assays. **18**, **21** and **22** showed only slight, concentration-dependent MST-shifts (Fig. S5P, S and T†), but no large changes in FRET-signals (Fig. S4†) were observed. Overall, limited ligand potency hinders strong conclusions except for compounds **8** and **8a**–**8c**. Additionally, identified hits do not represent the molecules with the highest docking scores (Table S3†), but those well representing interactions observed in the HCV IRES IIa – compound **5** crystal structure (Fig. 2B) which is indicative for the requirement of some (human introduced) ligand-bias during hit selection for RNA-targeting molecular docking screens.

Discussion and conclusions

A virtual screening of several RNA-focused and an unbiased library of drug-like small molecules against the HCV IRES subdomain IIa was performed. VS hit triage revealed that a larger fraction of molecules from the RNA-focused libraries is passing the respective filter steps compared to the unbiased library (Fig. S2B†). However, the overall numbers of commercially available compounds for focused libraries were too small for final hit selection. Thus, all molecules selected for testing originate from the unbiased library. We hypothesize that focusing of larger VS libraries or even make-on-demand chemical spaces[65–68] has the potential to result in higher hit rates and numbers. Eventually, the general concept of RNA-focusing might be insufficient. A more refined approach for the specific target or target class of interest should be used like for proteins rather than thinking of "RNA-focused". Instead of "protein-focused" libraries, kinase-, GPCR- or serine protease-focused libraries are used there. While this study is only a one-target test case, further studies are needed to refine our understanding of RNA-targeting (or riboswitch-, three-way junction-, internal loop-, triple helix-targeting, …) chemical space. Likewise, the knowledge of key-pharmacophore features of previously reported HCV IRES subdomain IIa ligands (Fig. 2B) was exploited during pose inspection to select VS hits for testing. Using two independent assays, an MST-binding assay and a FRET-based assay to monitor conformational changes of the target RNA, several VS hits could be confirmed as binders of the HCV IRES subdomain IIa with low to moderate potency (Table 1, Fig. 3). From six FRET- and seven MST-hits, four compounds (**8**, **11**–**13**) were identified in both assays. For the 2-guanidino-quinazoline compound **8**, further derivatives (**8a**–**8c**) were obtained for testing, representing the strongest binders in this study. The most potent ligand, compound **8b** (EC_{50} = 61 μM, K_D = 88 μM), shows structural similarities with the previously reported ligands **2** and **7** (Fig. 1), reminiscent of the interaction profile of arginine with G Hoogsteen edges (Fig. 3A and S1B†). Testing of amiloride (**26**) due to structural analogy, a privileged RNA-binding scaffold,[57–63] showed some weak binding as well. From the less potent single-assay hits, on the one hand, **9** and **10** showed signals in the FRET assay, but not in MST. Notably, MST-shifts are highly ligand-dependent and even though being a very sensitive[69] method, MST can therefore miss hits (false negatives). Hence, **9** might still be considered a weak IRES binder that induces a conformational change, especially when considering the prevalence of G-like moieties for G Hoogsteen edge binding in G–G:C motifs (Fig. S1C†).[29] This is also reflected in the observed preQ$_1$-riboswitch binding and altered FRET-behaviour in presence of competing RNA (Fig. S10G†). Compound **10** only showed a decrease in the Cy5-signal, but no increase of Cy3-emission (Fig. S7H†) which

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

might be an indicator for unspecific binding behavior. On the other hand, compounds **18**, **21** and **22** induced a slight MST-shift, but no changes in the FRET-assay, hinting for no to very weak or even unspecific binding that does not effectively induce conformational changes of the IRES subdomain IIa. A selectivity counter-screen with the structurally unrelated preQ$_1$-riboswitch and a competition assay in presence of a five-fold (m/m) excess of total RNA, showed selectivity of all VS hits except **9**. However, G-derivatives[31,64] are known binders of the preQ$_1$-riboswitch recognizing the 1*H*-pyrimidin-6-one moiety of compound **9**. This site might be recognized by other RNAs present in the used total RNA as well. Additional testing with the SAM-VI-riboswitch showed no general RNA-promiscuity of compound **9**. Finally, from the virtual screening, most hits showed only very low potency (**9**–**13**). Only compound **8** and derivatives **8a**–**8c** reached saturation in dose–response curves for both FRET and MST assays while being selective in the competition assay and the preQ$_1$-riboswitch counter screen. Therefore, only these compounds can unambiguously be described as specific HCV IRES subdomain IIa ligands. Identified 2-guanidino-quinazoline hits like compound **8b** (LE: 0.3 kcal mol^{-1}, LLE: 2.9) from this study can be considered as a potential starting point for derivatization and optimization. None of the hit compounds reaches the potency of previously reported 2-aminobenzimidazole ligands.[15,17,21,22,53,70–72] However, the reduced basicity of hit compounds compared to three basic nitrogen atoms in compound **5**, might result in beneficial physico-chemical properties. Additionally, this study can be considered as a proof-of-concept for the so far very rare field of structure-based virtual screenings against RNA-targets. Using protein-based tools and additional care during hit selection by pharmacophore elucidation, small molecular weight binders of the HCV IRES subdomain IIa were identified. Consequently, findings from this study can help to improve the field of RNA-targeting molecular docking screens hinting towards the implementation of bias from known ligands.

Material and methods

Virtual screening

For molecular docking, the HCV IRES subdomain IIa in complex with the 2-aminobenzimidazole ligand **5** was derived from the protein data bank (PDB,[73] PDB-ID: 3TZR[21]). FRED[74,75] (FRED 4.0.0.0 OpenEye Scientific Software, Santa Fe, NM. **http://www.eyesopen.com**), LeadIT (LeadIT-2.3.2, BioSolveIT GmbH, St. Augustin, Germany, 2017, **https://www.biosolveit.de/**) and FlexX[76] (FlexX Version 4.5.1, BioSolveIT GmbH, St. Augustin, Germany, 2020, **https://www.biosolveit.de/FlexX**) were demonstrated to be suitable for RNA–ligand docking without further modifications and treating the target RNA as rigid previously.[31] Docking setups were validated by re-docking and binder-decoy discrimination using 40 reported binders and 15 non-binders[15–17] as well as 1200 decoys generated using the database of useful decoys enhanced (DUD-E).[77] For LeadIT and FlexX docking, molecules were protonated with MOE (Molecular Operating Environment (MOE); 2020.09; Chemical Computing Group ULC: 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2020. **https://www.chemcomp.com/index.htm**) and energetically minimized using OMEGA[78] (OMEGA 4.1.0.0: OpenEye Scientific Software, Santa Fe, NM, USA. **http://www.eyesopen.com**, 2019). For FRED docking, conformers were generated with OMEGA using the "pose" keyword. For FRED and FlexX docking, water molecules and magnesium ions were removed from the binding site. Two magnesium ions (MG68 and MG70) close to the reference ligand were only indirectly considered by this removal to avoid overestimation of ion-complexing ligand moieties. However, ligands expanding deeply into the MG68 sub-pocket were also removed from hit-selection as binding strength/impact of displacement of Mg^{2+} to RNA is difficult to estimate. For LeadIT docking, water molecules that form at least three interactions with the ligand or target, namely water molecules HOH-6, HOH-71 and HOH-92 from PDB-ID 3TZR (Fig. 2B), were kept, but allowed to be displaceable or re-oriented during docking to allow poses with both direct or indirect, water-mediated interactions. Re-docking root-mean-square-deviation (RMSD) values of 1.22 Å, 1.23 Å and 2.08 Å (Fig. S3A–C†), receiver operating characteristics (ROC) area under the curve (AUC) of 0.88, 0.97 and 0.70 (Fig. S3D–F†) and adjusted ROC–log AUC$_{0.1–100\%}$[79–81] of 0.44, 0.35 and 0.09 (Fig. S3G–I†) for FRED, LeadIT and FlexX, respectively were obtained. Due to the weak performance of FlexX, likely caused by automatic deprotonation of 2-aminobenzimidazole-moieties during docking (Fig. S3C, F and I†), the prospective virtual screening was continued with FRED as a first step, followed by LeadIT, where pre-defined protonation states from MOE were used instead of using the Protoss[82] routines of LeadIT/FlexX. RNA-focused libraries were obtained from the respective supplier homepages of Asinex (combined focused libraries: "RNA-targeting", "dinucleoside mimetics", "fragments for RNA", "macrocycles for RNA", **https://www.asinex.com/screening-libraries**, accessed: 2021/08/13), Enamine (**https://enamine.net/compound-libraries/targeted-libraries**, accessed: 2021/08/13), Life Chemicals (**https://lifechemicals.com/screening-libraries/pre-plated-focused-libraries**, accessed: 2021/08/13), Otava (**https://www.otavachemicals.com/products/targeted-libraries-and-focused-libraries/other-focused-libraries**, accessed: 2021/08/13) and Reaxense (**https://www.reaxense.com/products/focused-libraries/rna-targeted-focused-library**, accessed: 2021/08/13). Molecules were "washed" to remove ions and protonated with MOE. PAINs[51,52] removal and physicochemical property filters (Table S2†) were applied using FILTER (OMEGA 4.1.0.0: OpenEye Scientific Software, Santa Fe, NM, USA. **http://www.eyesopen.com**, 2019). Molecules were docked with FRED as described for the model validation. Molecules with a ChemGauss score of ≤−13.0 kcal mol^{-1} or matching at least two pharmacophore features were

 This journal is © The Royal Society of Chemistry 2024

View Article Online
RSC Medicinal Chemistry Research Article

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

processed and subsequently docked with LeadIT prior visual pose inspection[83] and selection for testing.

Chemical analytics

Reference compound **4** was purchased from Chembridge/Hit2Lead. Compound **5** was synthesized according to published procedures[70,72] (see extended Material and methods section of the ESI†). Virtual screening hits were purchased from Chembridge/Hit2Lead (**9–12**, **14–16**, **25**), Enamine Ltd. (**17**, **18**, **20–23**) and Vitas-M (**8**, **13**, **19**). All purchased compounds were provided with confirmed identity and purity by LC-MS or ^{1}H-NMR analytics. Identity and purity of all compounds were re-determined in-house by HPLC-ESI/MS (Table S3†) using a 1100 series HPLC system from Agilent with a Poroshell 120 EC-C18 150 × 2.10 mm, 4 μm column if not stated differently and ^{1}H- and ^{13}C-NMR analytics (ESI,† 4. Analytical data of virtual screening hits) on a Bruker Avance Neo 400 MHz NMR spectrometer. The LC-MS measurements were conducted with a gradient of acetonitrile and water (+0.1% formic acid) ranging from 10% to 90% acetonitrile over 10 min with a flow rate of 0.7 mL min^{-1}. Signals were detected at 254 nm with quantification by area under curve (AUC). The molecular mass was detected using an Agilent 1100 series mass-selective detector trap with positive mode electron spray ionization. The LC-MS chromatograms and their corresponding mass spectra were analyzed using MestReNova (version 12.0.4, Mestrelab Research S.L. Feliciano Barrera 9B, Bajo, 15706 Santiago de Compostela, Spain. **http://www.mestrelab.com**). To exclude interference with the Cy3- and Cy5-labels, excitation and emission spectra of all compounds were recorded. None of the molecules showed excitation or emission in the range of Cy3 (excitation: 480 nm, emission: 535 nm) and Cy5 (excitation: 633 nm, emission: 670 nm).

FRET assay

For the FRET assay, a two stranded HCV IRES subdomain IIa carrying a 5′-Cy3-label on one and 5′-Cy5-label on the other strand (Cy3-UCGGAGGAACUACUGUCUUCACGCC, Cy5-UGCGUGUCGUGCAGCCUCCGG, Biomers.net GmbH) was used as described previously.[22,53] All samples containing labelled RNA were treated under light protection. FRET experiments were performed on a Tecan Spark 10 M plate reader (excitation/emission: 480/670 nm) using black half area 96 well microplates (Greiner Bio-One), based on a protocol by Zhou *et al.*[53] The 5′-Cy3- and 5′-Cy5-labelled single strands were annealed by heating to 75 °C in FRET buffer (10 mM HEPES pH 7.0, 2 mM $MgCl_2$) and cooled down to room temperature for at least 20 min. The 5′-Cy3/Cy5-labelled RNA was diluted to 100 nM in FRET buffer and supplemented with compounds **8–28** in appropriate concentrations. Compounds **4** and **5** were used as positive controls. For the initial screening, compounds were added in final concentrations of 1000, 100, 10 and 0.001 μM (Fig. S4†). If a concentration dependent reduction of the FRET signal (Cy5-emission) was observed, EC_{50}-values were determined using ligand concentrations from 1 mM to 100 nM in 3.16-fold (semi-logarithmic) dilution series. For those compounds, the Cy3 emission was measured at 535 nm to confirm a conformational change dependent reduction in the Cy5 emission signal *via* Cy3 emission increase (Fig. S7†). For each experiment, one DMSO negative control was carried out. The final DMSO concentration was always 2%. Excitation and emission spectra were recorded for the Cy3/Cy5-labelled RNA (100 nM) in FRET buffer to confirm excitation (Cy3) and emission (Cy5) wavelength as well as for each compound (1 μM) in FRET buffer to exclude any interference with the cyanine dyes. All measurements were performed at least as triplicates for screening and EC_{50}-determination.

For the HCV IRES-binding compounds (**4**, **5**, **8–13**, **8a–8c**), a selectivity FRET assay in presence of total RNA from yeast (Roche) was performed. The ratio of the FRET HCV IRES subdomain IIa construct and the total RNA was set to 1 : 5 m/m. The preparation of the labelled RNA and the compounds was the same as in the binding assay. The yeast RNA was diluted to 7.88 mg L^{-1} and supplemented with 100 nM HCV IRES and appropriate concentrations of the compounds (3.16-fold dilution series) in FRET buffer (Fig. S10†). For each compound, triplicate measurements were performed. For each experiment, one DMSO negative control was carried out. The final DMSO concentration was always 2%.

FRET experiments were analyzed using GraphPad Prism (version 8.0.1., GraphPad Software, Boston, Massachusetts USA, **https://www.graphpad.com**).

MST assay

For the MST assay, the same sequence of the two stranded HCV IRES subdomain IIa as for the FRET assay was used. One strand was 5′-Cy5-labelled and the other strand was unmodified (UCGGAGGAACUACUGUCUUCACGCC, Cy5-UGCGUGUCGUGCAGCCUCCGG, Biomers.net GmbH). All samples containing labelled RNA were treated under light protection. MST experiments were performed on a Monolith NT.115 from NanoTemper Technologies using standard uncoated capillaries. The strands were annealed by heating to 75 °C in MST buffer (10 mM HEPES pH 7.0, 2 mM $MgCl_2$) and cooled down to room temperature for at least 20 min. The 5′-Cy5-labelled HCV IRES domain IIa RNA was diluted to 30 nM in MST buffer and supplemented with compounds **8–25** in appropriate concentrations. Compounds **4** and **5** were used as positive controls. For the initial screening, compounds were added to final concentrations of 1000 (or 316 if limited by solubility), 100, 10 and 0.001 μM (Fig. S5†). If a concentration-dependent shift of thermophoresis was observed, K_D-values were determined using concentrations ranging from 1 mM (if permitted by solubility, otherwise starting from 316 or 100 μM) to 0.01 μM in a 3.16-fold (semi-logarithmic) dilution series. For each experiment, one DMSO negative control was carried out. Final DMSO concentration was always 2%.

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Off-target counter screens by MST were performed as described previously.[31,84] Briefly, the 5′-Cy5-labelled *Tte* preQ$_1$-Riboswitch aptamer domain (Biomers.net GmbH) was heated to 75 °C for 5 min, cooled down to room temperature for at least 20 min and then diluted to 20 nM in MST buffer (50 mM Tris-HCl pH 7.5, 100 mM KCl, 25 mM $MgCl_2$) and supplemented with compounds **4**, **5**, **8**–**13**, **18**, **21**, **22** and **26** to final concentrations of 1000, 100, 10 and 0.001 μM. PreQ$_1$ was used as a positive control. The 5′-Cy5-labelled *B. angulatum* SAM VI-Riboswitch aptamer domain (Eurofins Genomics GmbH) was heated to 75 °C for 5 min, cooled down to room temperature for at least 20 min and then diluted to 20 nM in MST buffer (50 mM Tris-HCl pH 7.5, 100 mM KCl, 25 mM $MgCl_2$) and supplemented with compounds **5**, **9** and **28** to final concentrations of 1000, 100, 10 and 0.001 μM. SAM was used as a positive control. For each experiment, one DMSO negative control was carried out. The final DMSO concentration was 2%. All measurements were performed at least as duplicates for pre-screening and triplicates for K_D-determination. MST measurements were analyzed from signals after 18 s laser on-time using the MO.Affinity Analysis software (version 2.3) and exported for statistical analysis and plotting in GraphPad Prism (version 8.0.1., GraphPad Software, Boston, Massachusetts USA, **https://www.graphpad.com**).

Author contributions

EK and LAR: investigation, methodology, formal analysis, visualization, validation, writing – original draft, writing – review & editing; JÅH: investigation, software; KB: investigation; CK: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, visualization, writing – original draft, writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Parts of this research were conducted using the supercomputer MOGON 2 and/or advisory services offered by Johannes Gutenberg University Mainz (**https://hpc.uni-mainz.de**), which is a member of the AHRP (Alliance for High Performance Computing in Rhineland Palatinate, **https://www.ahrp.info**) and the Gauss Alliance e.V. The authors gratefully acknowledge the computing time granted on the supercomputer MOGON 2 at Johannes Gutenberg University Mainz (hpc.uni-mainz.de). We further thank OpenEye Scientific (**https://www.eyesopen.com**) for academic licenses. The project was financially supported by the Johannes Gutenberg-University Internal Research Funding (Stufe-I). We thank Tanja Schirmeister (Johannes Gutenberg-University, Mainz) for scientific discussions and research facility access and Bernt Mikael Bergdahl (San Diego State University) for valuable hints on compound **5** synthesis. We also thank Mark Helm (Johannes Gutenberg-University, Mainz) for access to the MST instrument granted by the Bundesministerium für Bildung und Forschung (BMBF/01ED1804). Funding of the Bruker Avance Neo 400 MHz NMR spectrometer was supported by the German Research Council (Deutsche Forschungsgemeinschaft, DFG) through grant "400 MHz NMR-Spektrometerkonsole" (INST 247/1108-1 FUGG).

Notes and references

1 P. Axley, Z. Ahmed, S. Ravi and A. K. Singal, *J. Clin. Transl. Hepatol.*, 2018, **6**, 1–6.
2 K. Goto, A. A. Roca Suarez, F. Wrensch, T. F. Baumert and J. Lupberger, *Int. J. Mol. Sci.*, 2020, **21**, 3057.
3 WHO, Hepatitis C - Key facts, **https://www.who.int/news-room/fact-sheets/detail/hepatitis-c**, (accessed 1st December 2023).
4 A. Petruzziello, S. Marigliano, G. Loquercio, A. Cozzolino and C. Cacciapuoti, *World J. Gastroenterol.*, 2016, **22**, 7824.
5 M. A. Izhari, *Diagnostics*, 2023, **13**, 3102.
6 H. Ji, C. S. Fraser, Y. Yu, J. Leary and J. A. Doudna, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 16990–16995.
7 S. M. Dibrov, J. Parsons, M. Carnevali, S. Zhou, K. D. Rynearson, K. Ding, E. Garcia Sega, N. D. Brunn, M. A. Boerneke, M. P. Castaldi and T. Hermann, *J. Med. Chem.*, 2014, **57**, 1694–1707.
8 R. B. Paulsen, P. P. Seth, E. E. Swayze, R. H. Griffey, J. J. Skalicky, T. E. Cheatham and D. R. Davis, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 7263–7268.
9 D. R. Davis and P. P. Seth, *Antiviral Chem. Chemother.*, 2011, **21**, 117–128.
10 R. Hanecak, V. Brown-Driver, M. C. Fox, R. F. Azad, S. Furusako, C. Nozaki, C. Ford, H. Sasmor and K. P. Anderson, *J. Virol.*, 1996, **70**, 5203–5212.
11 R. Pudi, S. S. Ramamurthy and S. Das, *J. Virol.*, 2005, **79**, 9842–9853.
12 W. Wang, P. Préville, N. Morin, S. Mounir, W. Cai and M. A. Siddiqui, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 1151–1154.
13 E. A. Jefferson, P. P. Seth, D. E. Robinson, D. K. Winter, A. Miyaji, S. A. Osgood, E. E. Swayze and L. M. Risen, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 5139–5143.
14 M. Carnevali, J. Parsons, D. L. Wyles and T. Hermann, *ChemBioChem*, 2010, **11**, 1364–1367.
15 P. P. Seth, A. Miyaji, E. A. Jefferson, K. A. Sannes-Lowery, S. A. Osgood, S. S. Propp, R. Ranken, C. Massire, R. Sampath, D. J. Ecker, E. E. Swayze and R. H. Griffey, *J. Med. Chem.*, 2005, **48**, 7099–7102.
16 K. D. Rynearson, B. Charrette, C. Gabriel, J. Moreno, M. A. Boerneke, S. M. Dibrov and T. Hermann, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 3521–3525.
17 B. P. Charrette, M. A. Boerneke and T. Hermann, *ACS Chem. Biol.*, 2016, **11**, 3263–3267.
18 Q. Zhao, Q. Han, C. R. Kissinger, T. Hermann and P. A. Thompson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2008, **64**, 436–443.

 This journal is © The Royal Society of Chemistry 2024

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

19 S. M. Dibrov, H. Johnston-Cox, Y. H. Weng and T. Hermann, *Angew. Chem., Int. Ed.*, 2007, **46**, 226–229.
20 C. M. T. Spahn, J. S. Kieft, R. A. Grassucci, P. A. Penczek, K. Zhou, J. A. Doudna and J. Frank, *Science*, 2001, **291**, 1959–1962.
21 S. M. Dibrov, K. Ding, N. D. Brunn, M. A. Parker, B. M. Bergdahl, D. L. Wyles and T. Hermann, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 5223–5228.
22 J. Parsons, M. P. Castaldi, S. Dutta, S. M. Dibrov, D. L. Wyles and T. Hermann, *Nat. Chem. Biol.*, 2009, **5**, 823–825.
23 M. M. Hoffman, M. A. Khrapov, J. C. Cox, J. Yao, L. Tong and A. D. Ellington, *Nucleic Acids Res.*, 2004, **32**, 174–181.
24 J. Kondo and E. Westhof, *Nucleic Acids Res.*, 2011, **39**, 8628–8637.
25 Z. Wang, P. Fang and M. S. Sachs, *Mol. Cell. Biol.*, 1998, **18**, 7528–7536.
26 A. S. Abu Almakarem, A. I. Petrov, J. Stombaugh, C. L. Zirbel and N. B. Leontis, *Nucleic Acids Res.*, 2012, **40**, 1407–1423.
27 M. D. Shortridge, P. T. Wille, A. N. Jones, A. Davidson, J. Bogdanovic, E. Arts, J. Karn, J. A. Robinson and G. Varani, *Nucleic Acids Res.*, 2019, **47**, 1523–1531.
28 P. Bénas, G. Bec, G. Keith, R. Marquet, C. Ehresmann, B. Ehresmann and P. Dumas, *RNA*, 2000, **6**, 1347–1355.
29 M. A. Boerneke, S. M. Dibrov, J. Gu, D. L. Wyles and T. Hermann, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15952–15957.
30 P. Daldrop, F. E. Reyes, D. A. Robinson, C. M. Hammond, D. M. Lilley, R. T. Batey and R. Brenk, *Chem. Biol.*, 2011, **18**, 324–335.
31 E. Kallert, T. R. Fischer, S. Schneider, M. Grimm, M. Helm and C. Kersten, *J. Chem. Inf. Model.*, 2022, **62**, 4134–4148.
32 Q. Vicens, E. Mondragón, F. E. Reyes, P. Coish, P. Aristoff, J. Berman, H. Kaur, K. W. Kells, P. Wickens, J. Wilson, R. C. Gadwood, H. J. Schostarez, R. K. Suto, K. F. Blount and R. T. Batey, *ACS Chem. Biol.*, 2018, **13**, 2908–2919.
33 C. M. Connelly, T. Numata, R. E. Boer, M. H. Moon, R. S. Sinniah, J. J. Barchi, A. R. Ferré-D'Amaré and J. S. Schneekloth, *Nat. Commun.*, 2019, **10**, 1501.
34 M. J. Zeller, O. Favorov, K. Li, A. Nuthanakanti, D. Hussein, A. Michaud, D. A. Lafontaine, S. Busan, A. Serganov, J. Aubé and K. M. Weeks, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2122660119.
35 J. L. Childs-Disney, X. Yang, Q. M. R. Gibaut, Y. Tong, R. T. Batey and M. D. Disney, *Nat. Rev. Drug Discovery*, 2022, **21**, 736–762.
36 J. P. Falese, A. Donlic and A. E. Hargrove, *Chem. Soc. Rev.*, 2021, **50**, 2224–2243.
37 C. Sheridan, *Nat. Biotechnol.*, 2021, **39**, 6–8.
38 H. Ratni, R. S. Scalco and A. H. Stephan, *ACS Med. Chem. Lett.*, 2021, **12**, 874–877.
39 Y. Feng and S.-Y. Huang, *J. Chem. Inf. Model.*, 2020, **60**, 6698–6708.
40 Y. Zhou, Y. Jiang and S. Chen, *WIREs Comput. Mol. Sci.*, 2022, **12**, e1571.
41 J. Manigrasso, M. Marcia and M. De Vivo, *Chem*, 2021, **7**, 2965–2988.
42 O. Binas, V. de Jesus, T. Landgraf, A. E. Völklein, J. Martins, D. Hymon, J. Kaur Bains, H. Berg, T. Biedenbänder, B. Fürtig, S. Lakshmi Gande, A. Niesteruk, A. Oxenfarth, N. Shahin Qureshi, T. Schamber, R. Schnieders, A. Tröster, A. Wacker, J. Wirmer-Bartoschek, M. A. Wirtz Martin, E. Stirnal, K. Azzaoui, C. Richter, S. Sreeramulu, M. J. José Blommers and H. Schwalbe, *ChemBioChem*, 2021, **22**, 423–433.
43 S. Sreeramulu, C. Richter, H. Berg, M. A. Wirtz Martin, B. Ceylan, T. Matzel, J. Adam, N. Altincekic, K. Azzaoui, J. K. Bains, M. J. J. Blommers, J. Ferner, B. Fürtig, M. Göbel, J. T. Grün, M. Hengesbach, K. F. Hohmann, D. Hymon, B. Knezic, J. N. Martins, K. R. Mertinkus, A. Niesteruk, S. A. Peter, D. J. Pyper, N. S. Qureshi, U. Scheffer, A. Schlundt, R. Schnieders, E. Stirnal, A. Sudakov, A. Tröster, J. Vögele, A. Wacker, J. E. Weigand, J. Wirmer-Bartoschek, J. Wöhnert and H. Schwalbe, *Angew. Chem., Int. Ed.*, 2021, **60**, 19191–19200.
44 A. Umuhire Juru and A. E. Hargrove, *J. Biol. Chem.*, 2021, **296**, 100191.
45 K. Wang, Y. Jian, H. Wang, C. Zeng and Y. Zhao, *Bioinformatics*, 2018, **34**, 3131–3136.
46 T. Tran and M. D. Disney, *Nat. Commun.*, 2012, **3**, 1125.
47 J. L. Childs-Disney, T. Tran, B. R. Vummidi, S. P. Velagapudi, H. S. Haniff, Y. Matsumoto, G. Crynen, M. R. Southern, A. Biswas, Z.-F. Wang, T. L. Tellinghuisen and M. D. Disney, *Chem*, 2018, **4**, 2384–2404.
48 K. P. Lundquist, V. Panchal, C. H. Gotfredsen, R. Brenk and M. H. Clausen, *ChemMedChem*, 2021, **16**, 2588–2603.
49 M. D. Disney, A. M. Winkelsas, S. P. Velagapudi, M. Southern, M. Fallahi and J. L. Childs-Disney, *ACS Chem. Biol.*, 2016, **11**, 1720–1728.
50 A. Donlic, E. G. Swanson, L.-Y. Chiu, S. L. Wicks, A. U. Juru, Z. Cai, K. Kassam, C. Laudeman, B. G. Sanaba, A. Sugarman, E. Han, B. S. Tolbert and A. E. Hargrove, *ACS Chem. Biol.*, 2022, **17**, 1556–1566.
51 J. Baell and M. A. Walters, *Nature*, 2014, **513**, 481–483.
52 J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719–2740.
53 S. Zhou, K. D. Rynearson, K. Ding, N. D. Brunn and T. Hermann, *Bioorg. Med. Chem.*, 2013, **21**, 6139–6144.
54 M. A. Boerneke and T. Hermann, *Methods*, 2015, **91**, 35–39.
55 S. A. Wildman and G. M. Crippen, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 868–873.
56 A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees and C. H. Reynolds, *Nat. Rev. Drug Discovery*, 2014, **13**, 105–121.
57 N. N. Patwardhan, L. R. Ganser, G. J. Kapral, C. S. Eubanks, J. Lee, B. Sathyamoorthy, H. M. Al-Hashimi and A. E. Hargrove, *MedChemComm*, 2017, **8**, 1022–1036.
58 N. N. Patwardhan, Z. Cai, A. Umuhire Juru and A. E. Hargrove, *Org. Biomol. Chem.*, 2019, **17**, 9313–9320.
59 J. Davila-Calderon, N. N. Patwardhan, L.-Y. Chiu, A. Sugarman, Z. Cai, S. R. Penutmutchu, M.-L. Li, G. Brewer, A. E. Hargrove and B. S. Tolbert, *Nat. Commun.*, 2020, **11**, 4775.

Open Access Article. Published on 18 March 2024. Downloaded on 4/15/2024 11:10:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

60 M. Zafferani, C. Haddad, L. Luo, J. Davila-Calderon, L.-Y. Chiu, C. S. Mugisha, A. G. Monaghan, A. A. Kennedy, J. D. Yesselman, R. J. Gifford, A. W. Tai, S. B. Kutluay, M.-L. Li, G. Brewer, B. S. Tolbert and A. E. Hargrove, *Sci. Adv.*, 2023, **7**, eabl6096.
61 J. Vögele, D. Hymon, J. Martins, J. Ferner, H. R. A. Jonker, A. E. Hargrove, J. E. Weigand, A. Wacker, H. Schwalbe, J. Wöhnert and E. Duchardt-Ferner, *Nucleic Acids Res.*, 2023, gkad762.
62 A. C. Stelzer, A. T. Frank, J. D. Kratz, M. D. Swanson, M. J. Gonzalez-Hernandez, J. Lee, I. Andricioaei, D. M. Markovitz and H. M. Al-Hashimi, *Nat. Chem. Biol.*, 2011, **7**, 553–559.
63 D. Hymon, J. Martins, C. Richter, S. Sreeramulu, A. Wacker, J. Ferner, N. N. Patwardhan, A. E. Hargrove and H. Schwalbe, *RSC Med. Chem.*, 2024, **15**, 165–177.
64 A. Roth, W. C. Winkler, E. E. Regulski, B. W. K. Lee, J. Lim, I. Jona, J. E. Barrick, A. Ritwik, J. N. Kim, R. Welz, D. Iwata-Reuyl and R. R. Breaker, *Nat. Struct. Mol. Biol.*, 2007, **14**, 308–317.
65 R. A. Zimmermann, T. R. Fischer, M. Schwickert, Z. Nidoieva, T. Schirmeister and C. Kersten, *Int. J. Mol. Sci.*, 2023, **24**, 6109.
66 W. A. Warr, M. C. Nicklaus, C. A. Nicolaou and M. Rarey, *J. Chem. Inf. Model.*, 2022, **62**, 2021–2034.
67 A. Neumann, L. Marrison and R. Klein, *ACS Med. Chem. Lett.*, 2023, **14**, 466–472.
68 T. Hoffmann and M. Gastreich, *Drug Discovery Today*, 2019, **24**, 1148–1156.
69 P. Linke, K. Amaning, M. Maschberger, F. Vallee, V. Steier, P. Baaske, S. Duhr, D. Breitsprecher and A. Rak, *SLAS Discovery*, 2016, **21**, 414–421.
70 M. A. Parker, E. Satkiewicz, T. Hermann and B. M. Bergdahl, *Molecules*, 2011, **16**, 281–290.
71 K. Ding, A. Wang, M. A. Boerneke, S. M. Dibrov and T. Hermann, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 3113–3117.
72 D. Schmit, U. Milewicz, M. A. Boerneke, S. Burley, K. Walsworth, J. Um, D. Hecht, T. Hermann and B. M. Bergdahl, *Aust. J. Chem.*, 2020, **73**, 212–221.
73 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
74 M. McGann, *J. Chem. Inf. Model.*, 2011, **51**, 578–596.
75 M. McGann, *J. Comput.-Aided Mol. Des.*, 2012, **26**, 897–906.
76 M. Rarey, B. Kramer, T. Lengauer and G. Klebe, *J. Mol. Biol.*, 1996, **261**, 470–489.
77 M. M. Mysinger, M. Carchia, J. J. Irwin and B. K. Shoichet, *J. Med. Chem.*, 2012, **55**, 6582–6594.
78 P. C. D. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson and M. T. Stahl, *J. Chem. Inf. Model.*, 2010, **50**, 572–584.
79 M. M. Mysinger and B. K. Shoichet, *J. Chem. Inf. Model.*, 2010, **50**, 1561–1573.
80 B. J. Bender, S. Gahbauer, A. Luttens, J. Lyu, C. M. Webb, R. M. Stein, E. A. Fink, T. E. Balius, J. Carlsson, J. J. Irwin and B. K. Shoichet, *Nat. Protoc.*, 2021, **16**, 4799–4832.
81 C. Kersten, S. Clower and F. Barthels, *J. Chem. Inf. Model.*, 2023, **63**, 2218–2225.
82 S. Bietz, S. Urbaczek, B. Schulz and M. Rarey, *Aust. J. Chem.*, 2014, **6**, 12.
83 A. Fischer, M. Smieško, M. Sellner and M. A. Lill, *J. Med. Chem.*, 2021, **64**, 2489–2500.
84 E. Kallert, M. Behrendt, A. Frey, C. Kersten and F. Barthels, *Chem. Sci.*, 2023, **14**, 9827–9837.

 This journal is © The Royal Society of Chemistry 2024

Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.
This journal is © The Royal Society of Chemistry 2024

Supporting Information:

Structure-based virtual screening of unbiased and RNA-focused libraries to identify new ligands for the HCV IRES model system

Elisabeth Kallert,[a,†] Laura Almena Rodriguez,[a,†] Jan-Åke Husmann,[a] Kathrin Blatt,[a] and Christian Kersten[a,b,*]

[a]Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany.
[b]Institute for Quantitative and Computational Biosciences, Johannes Gutenberg-University, BioZentrum I, Hanns-Dieter-Hüsch.Weg 15, 55128 Mainz, Germany
† EK and LAR contributed equally to this manuscript.
*correspondence CK: kerstec@uni-mainz.de

Table of Contents

1. Figures S1-11 2
2. Tables S1-3 15
3. Extended material and methods: synthesis of compound 5 19
4. Analytical data of virtual screening hits 28
5. References 72

1. Figures S1-11

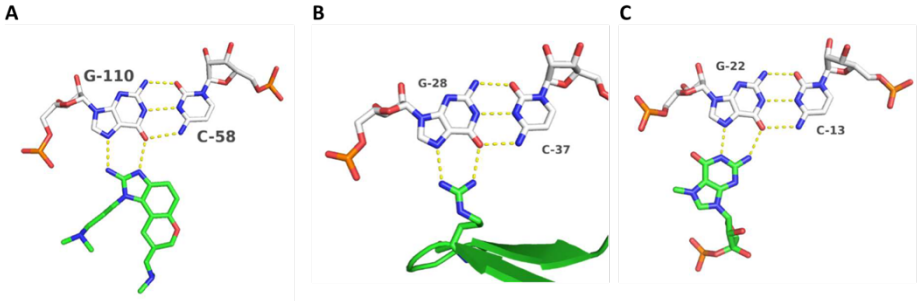


Figure S 1: Hoogsteen edge interactions of G:C pairs (white carbon atoms) with **A)** 2-aminobenzimidazole ligands (HCV IRES, PDB-ID: 3TZR), **B)** arginine (HIV trans-activation response element - peptidic macrocycle complex, PDB-ID 6D2U) and **C)** 7-methylguanosine (HIV reverse-transcription primer tRNA, PDB-ID: 1FIR).

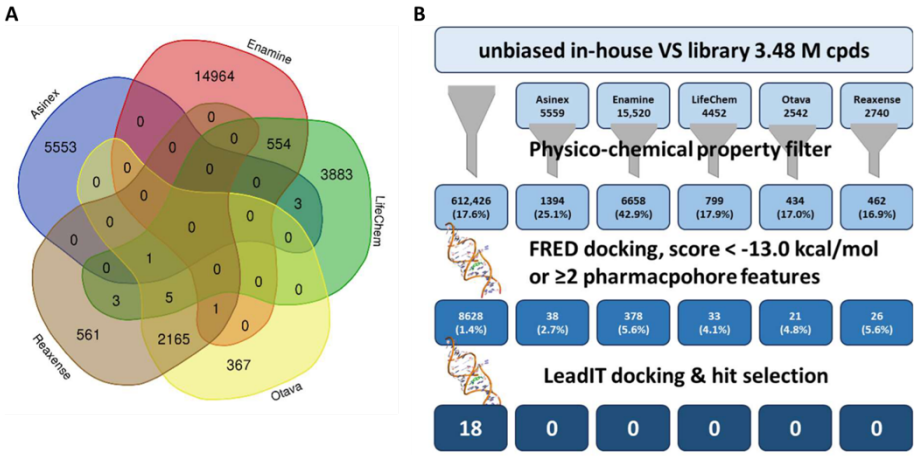


Figure S 2: A) Venn diagram showing little overlap of RNA-focused libraries from different suppliers (except Otava/Reaxense). Figure was made with the Bioinformatics & Evolutionary Genomics web tool (https://bioinformatics.psb.ugent.be/webtools/Venn/). **B)** Schematic representation of virtual screening triage workflow covering physicochemical property filtering, a two-step docking process and human pose inspection for hit selection. Values in parentheses indicate faction of molecules passing the previous step in %.

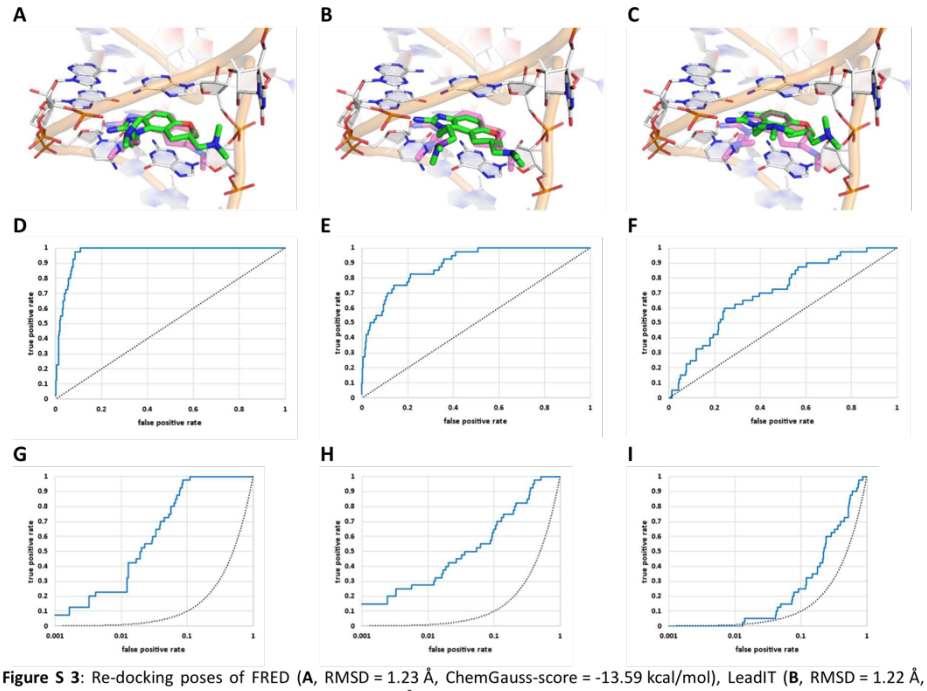
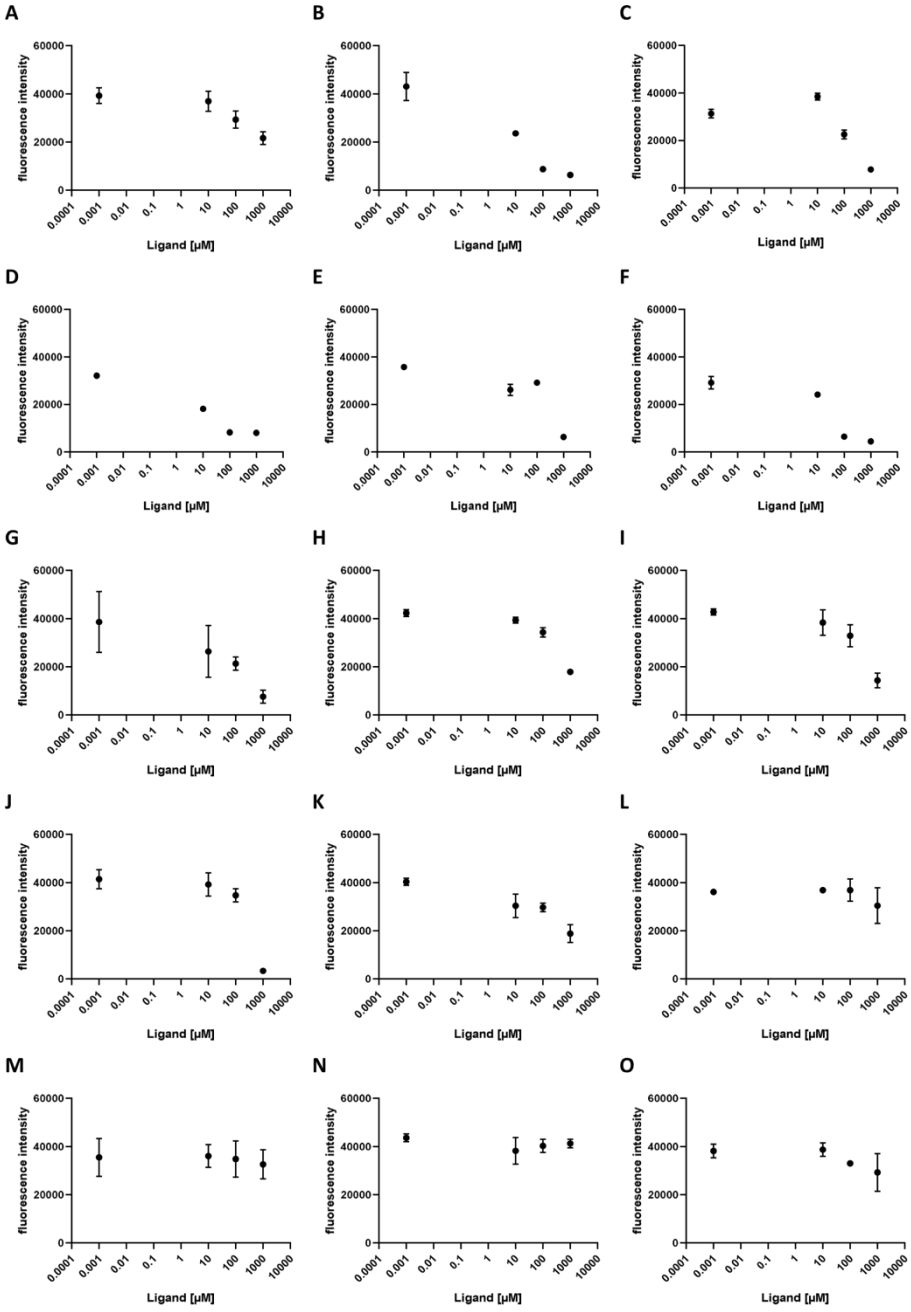


Figure S 3: Re-docking poses of FRED (**A**, RMSD = 1.23 Å, ChemGauss-score = -13.59 kcal/mol), LeadIT (**B**, RMSD = 1.22 Å, FlexX-score = -37.3 kJ/mol) and FlexX (**C**, RMSD = 2.08 Å, FlexX-score = -27.7 kJ/mol). HCV IRES is depicted with white carbon atoms and transparent cartoon (PDB-ID: 3TZR), docking poses with green carbon atoms and the crystallographic reference ligand with transparent, magenta-colored carbon atoms. Receiver operating characteristics curves for FRED (**D**, ROC-AUC = 0.97), LeadIT (**E**, ROC-AUC = 0.88) and FlexX (**F**, ROC-AUC = 0.70). Semilogarithmic ROC curves between 0.1 % and 100 % for FRED (**G**, ROC-logAUC = 0.44), LeadIT (**H**, ROC-logAUC = 0.36) and FlexX (**I**, ROC-logAUC = 0.09). ROC-logAUC values are adjusted to obtain values of 0 for random distribution to 0.855 for prefect discrimination. FlexX was not followed up for the prospective virtual screening.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
fluorescence intensity
60000
40000
20000
0
0.0001
0.001
0.01
0.1
1
10
100
1000
10000
Ligand [µM]

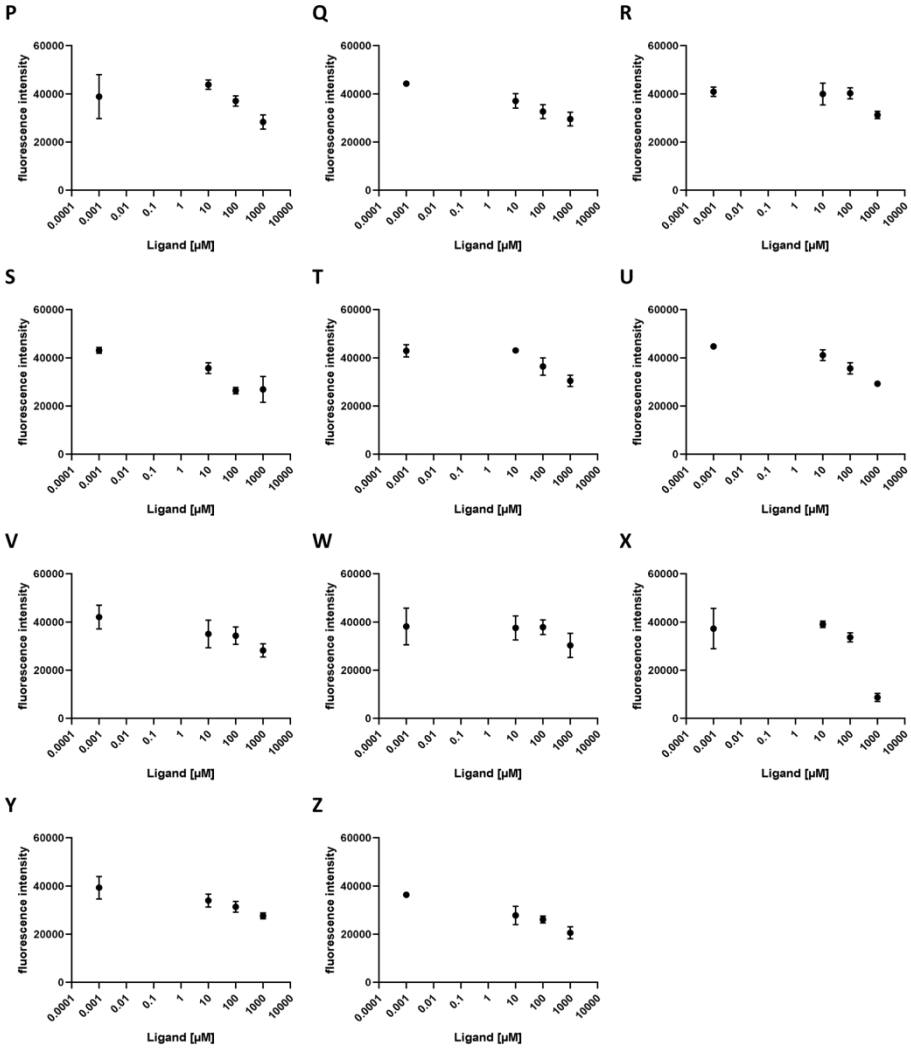
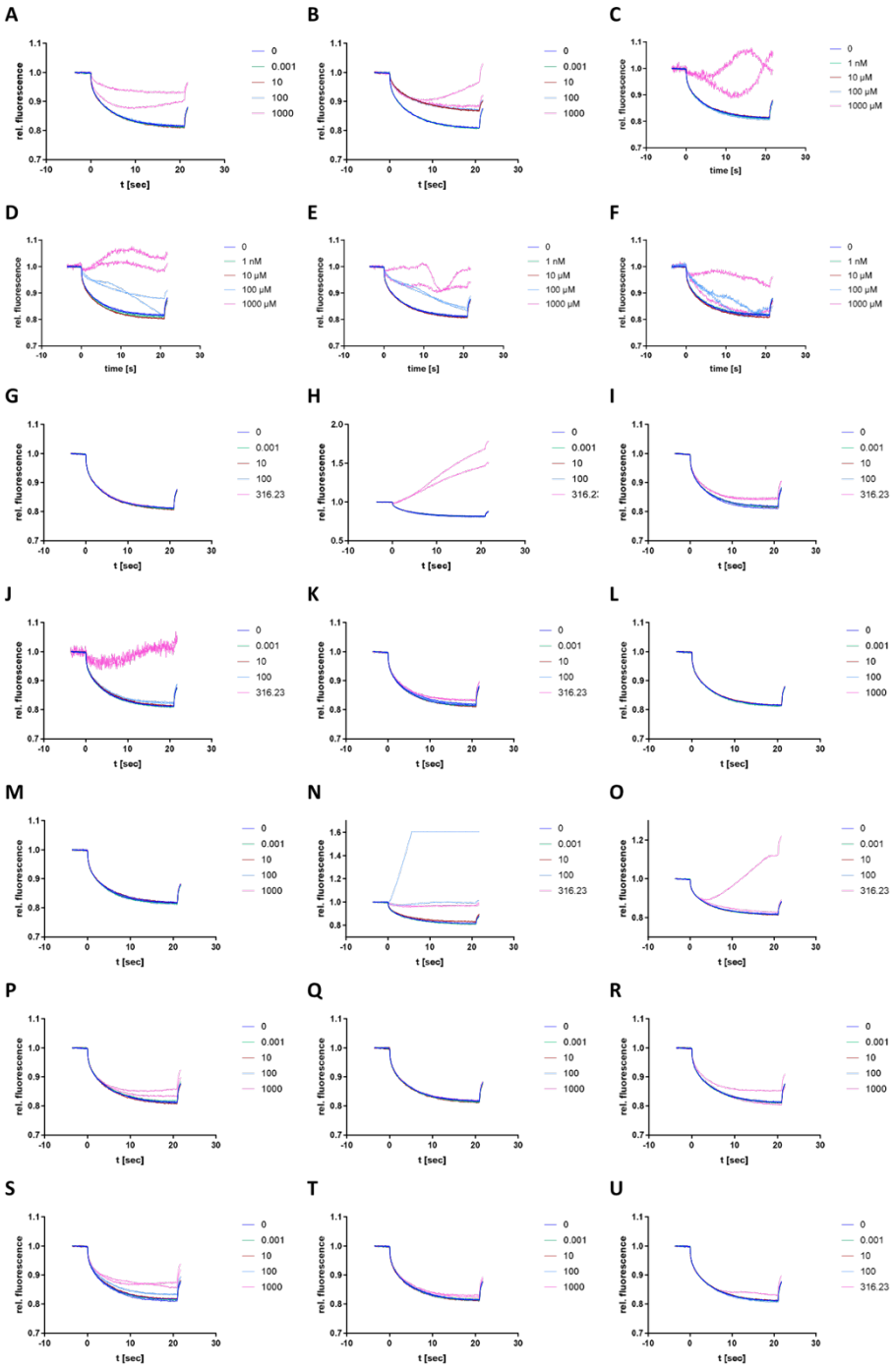


Figure S 4: FRET pre-screening results (Cy5 emission at 670 nm) at concentrations of 1000, 100, 10 and 0.001 µM for compounds **4** (A), **5** (B), **8** (C), **8a** – **c** (D-F), **9** – **28** (G-Z).


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
rel. fluorescence
t [sec]
time [s]
0
0.001
1 nM
10
10 µM
100
100 µM
1000
1000 µM
316.23

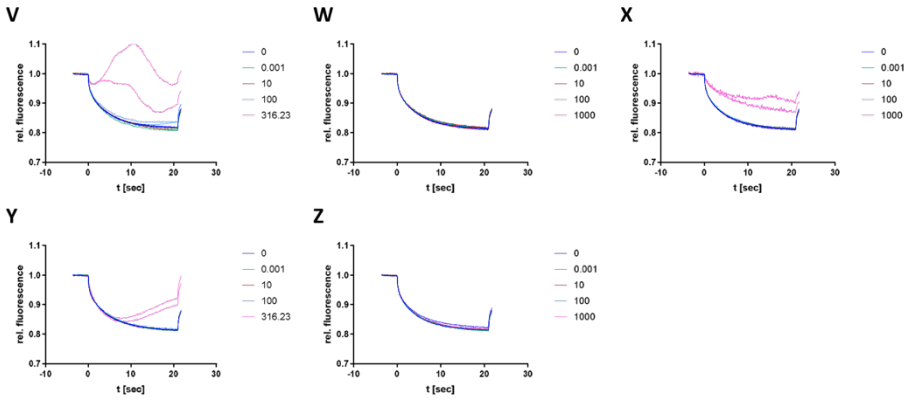


Figure S 5: MST pre-screening results at concentrations of 1000 (316.2 in case of poor solubility), 100, 10 and 0.001 µM for compounds **4** (A), **5** (B), **8** (C), **8a** (D), **8b** (E), **8c** (F), **9** – **28** (G-Z). Aggregation at high concentrations was observed for compounds **5** (B), **8** (C), **8a** (D) **8b** (E), **8c** (F), **10** (H), **12** (J), **16** (N), **17** (O), **24** (V), **26** (X), **27** (Y).

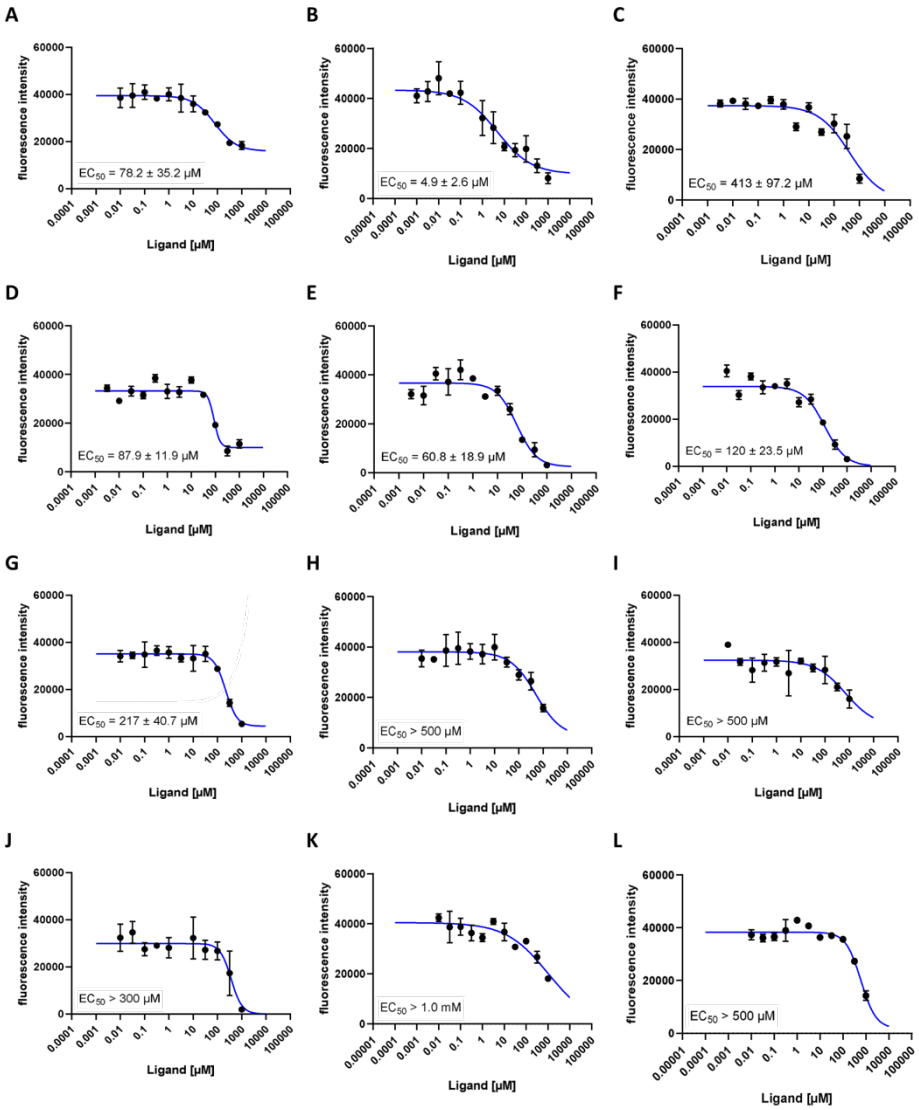


Figure S 6: FRET dose-response curves (Cy5 emission at 670 nm) for compounds **4** (A), **5** (B), **8** (C), **8a** (D), **8b** (E), **8c** (F), **9** (G), **10** (H), **11** (I), **12** (J), **13** (K) and amiloride (**26**, L).

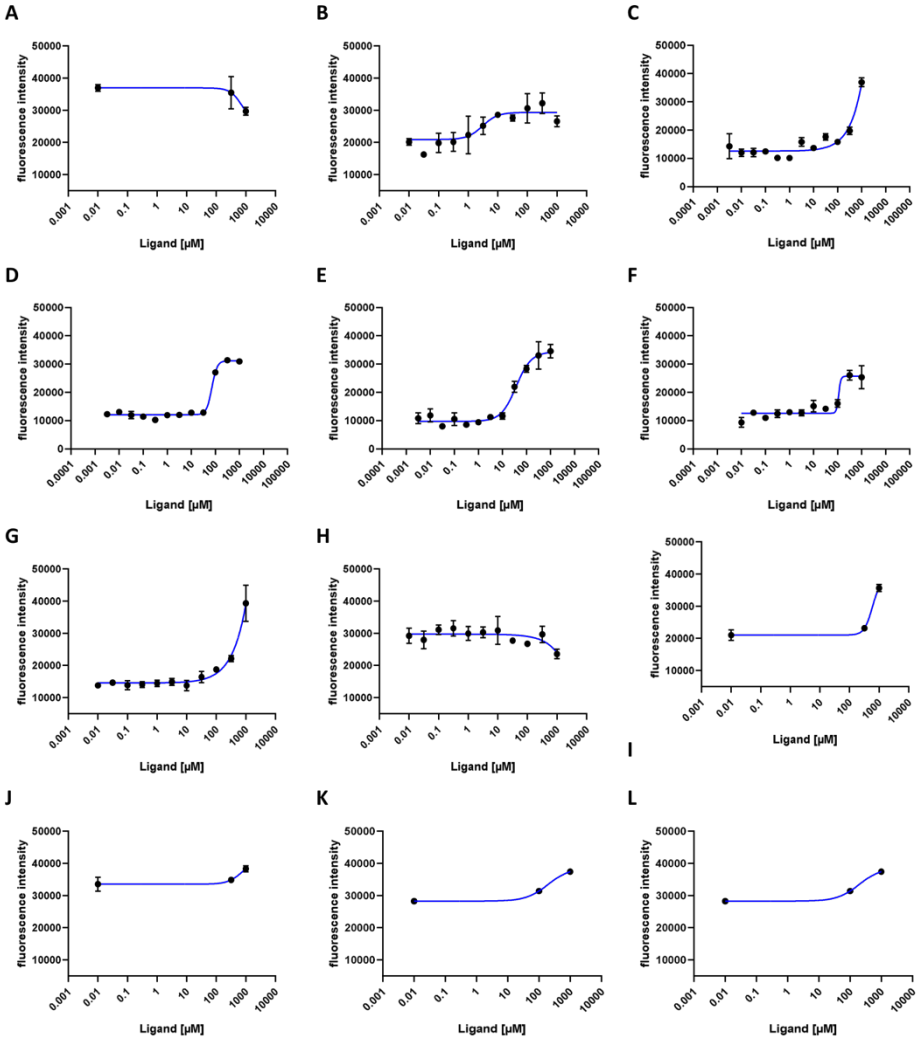


Figure S 7: Cy3 emission fluorescence intensity (535 nm) dose-response curves for compounds **4** (A), **5** (B), **8** (C), **8a** (D), **8b** (E), **8c** (F), **9** (G), **10** (H), **11** (I), **12** (J), **13** (K) and amiloride (**26**, L).

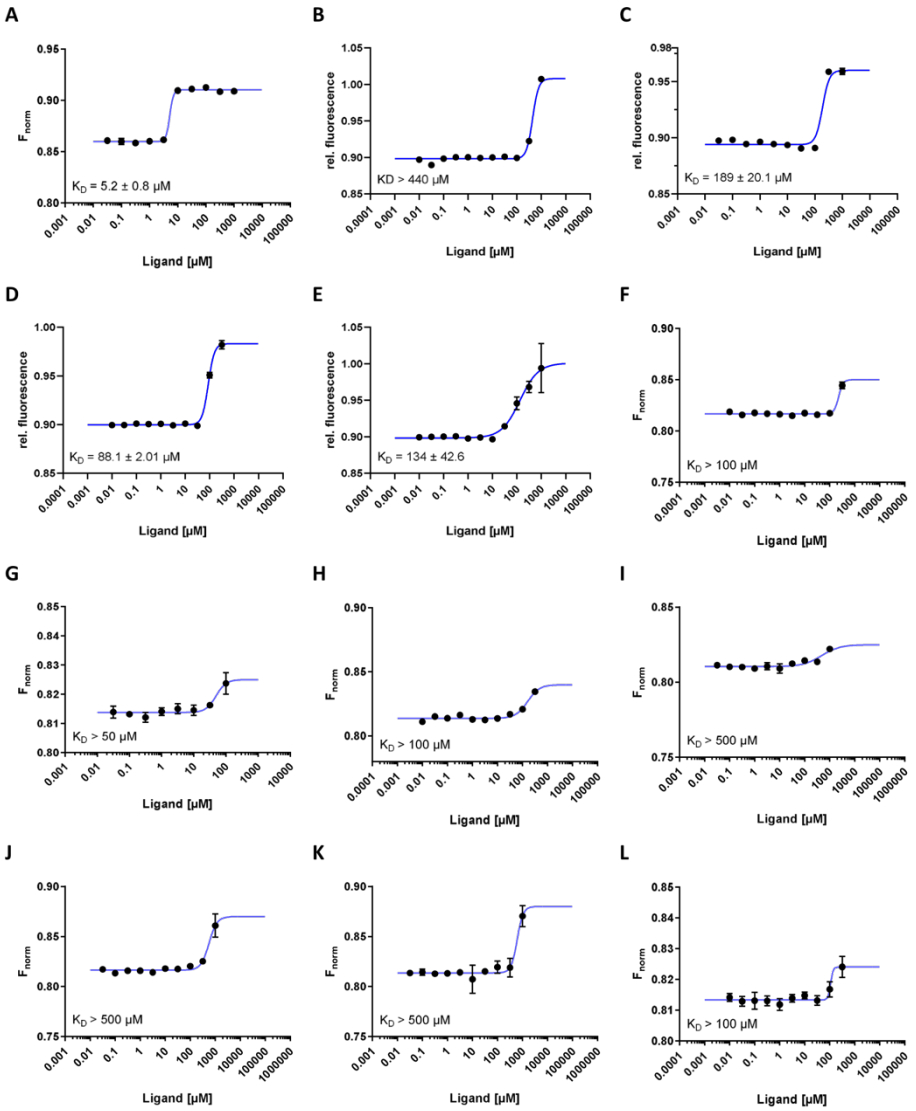
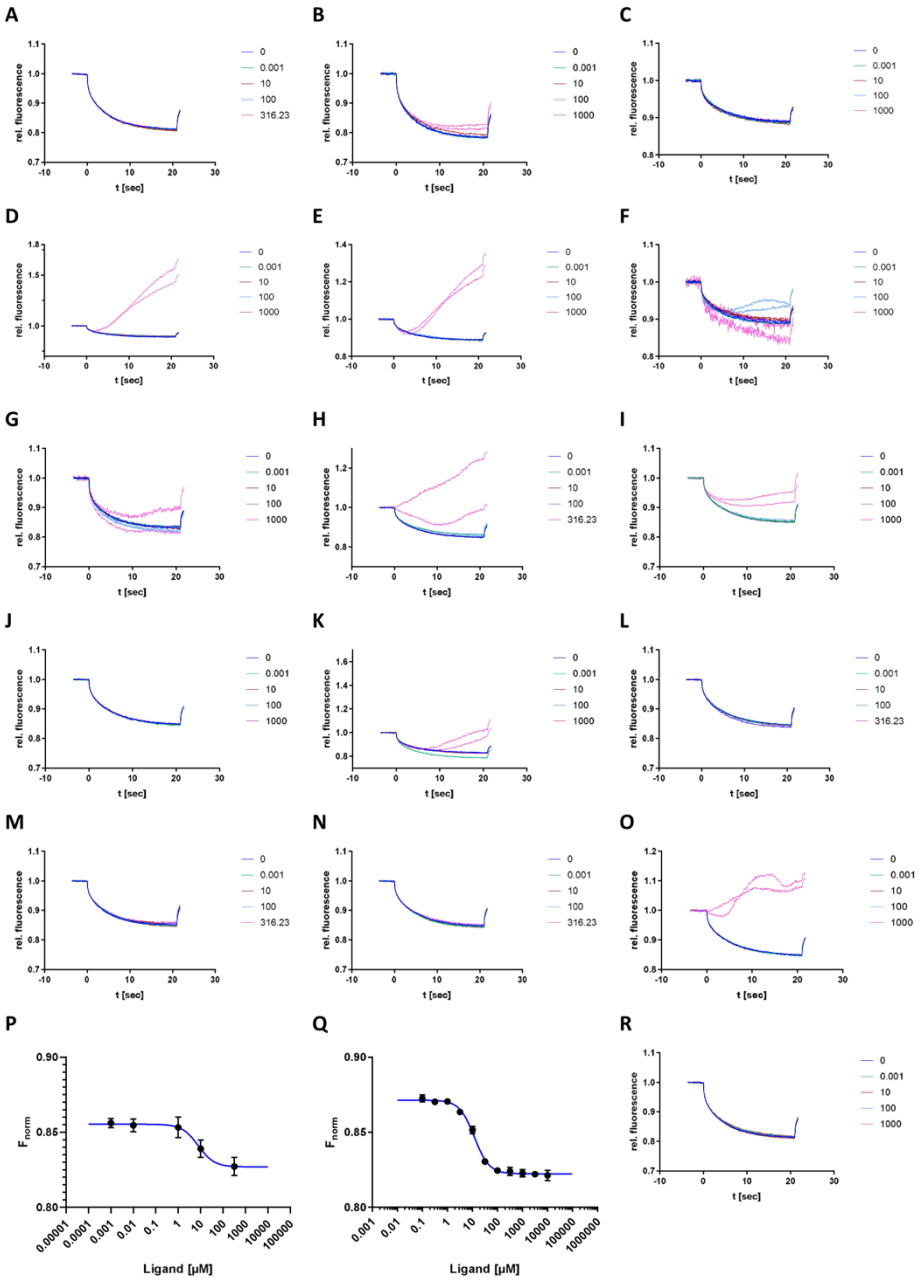


Figure S 8: MST dose-response curves for compounds **5** (A), **8** (B), **8a** (C), **8b** (D), **8c** (E). **11** (F), **12** (G), **13** (H), **18** (I) **21** (J), **22** (K) and amiloride (**26**, L).


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
rel. fluorescence
t [sec]
Fnorm
Ligand [µM]
0
0.001
10
100
316.23
1000

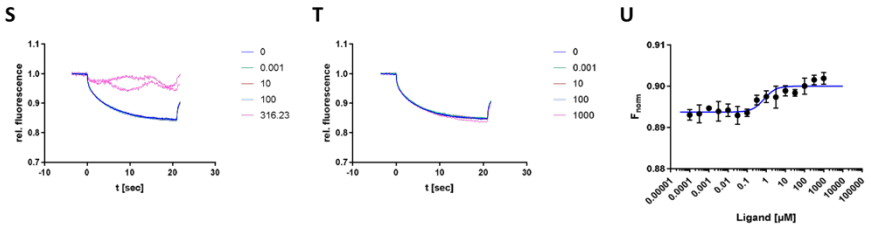


Figure S 9: MST-traces from the selectivity counter-screening against the preQ$_1$-riboswitch. Compound **4** (A), **5** (B), **8** (C), **8a** (D), **8b** (E), **8c** (F), **9** (G), **10** (H), **11** (I), **12** (J), **13** (K), **18** (L), **21** (M), **22** (N) and amiloride (**26**, O). Aggregation was observed for high concentrations of compounds **8a** (D), **8b** (E), **8c** (F), **9** (G), **10** (H), **11** (I), **13** (K) and **26** (O). Weak preQ$_1$-riboswitch binding was observed for compound **5** (B) and **9** (G). P) Dose-response curve of compound **9** binding to the preQ$_1$-riboswitch (K_D = 7.8 µM). Q) Dose-response curve of preQ$_1$ binding to its riboswitch (positive control, K_D = 12 nM). MST-traces of SAM-VI-riboswitch in presence of compound **5** (R), **9** (S) and **28** (T) show no binding. Aggregation was observed for high concentrations of compound **9** (S). U) Dose-response curve of SAM binding to the SAM-VI-riboswitch (positive control, K_D = 772 nM).

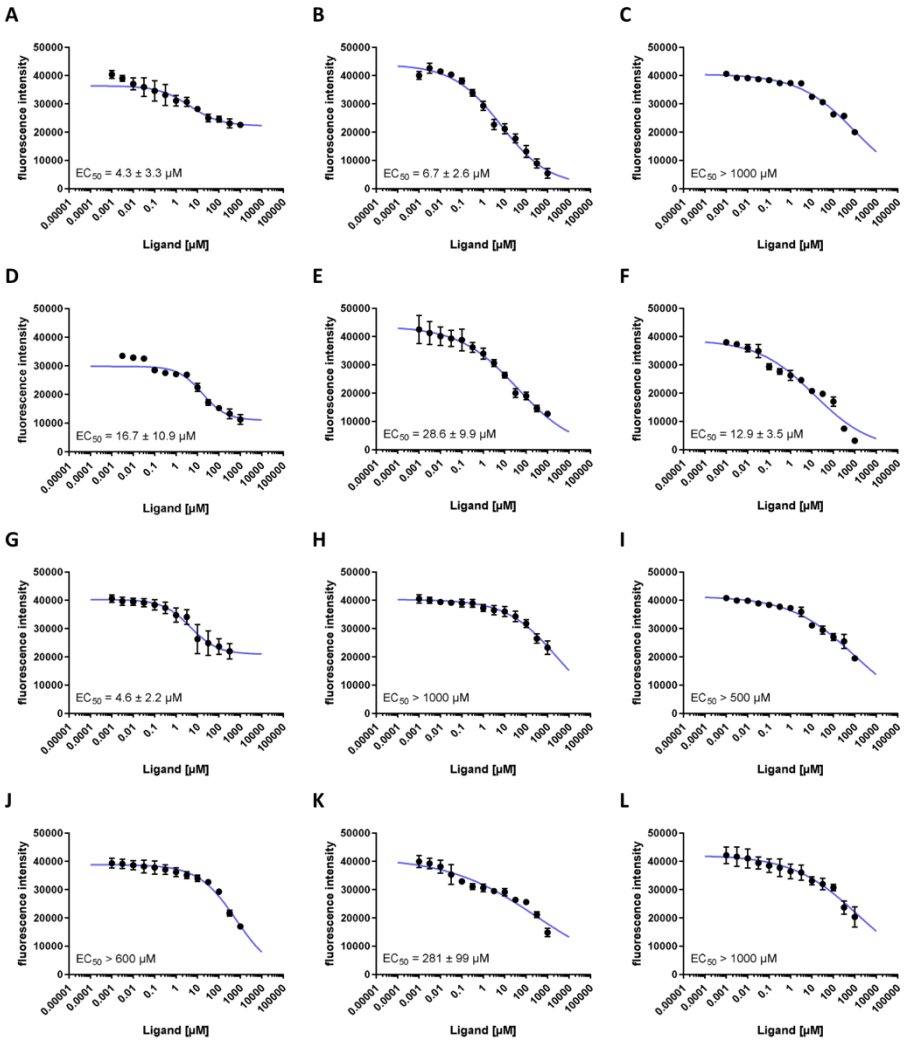


Figure S 10: FRET dose-response curves from the selectivity screening against HCV IRES in presence of total RNA from yeast (1:5 HCV-IRES: total RNA ratio, m/m). Compound **4** (A), **5** (B), **8** (C), **8a** (D), **8b** (E), **8c** (F), **9** (G), **10** (H), **11** (I), **12** (J), **13** (K) and amiloride (**26**, L). For **4** and **9** lower EC_{50}-values are observed in presence of total RNA compared to HCV IRES alone (Figure S6 A and G, respectively), but the remaining fluorescence intensity is higher indicating impairment of the conformational change.

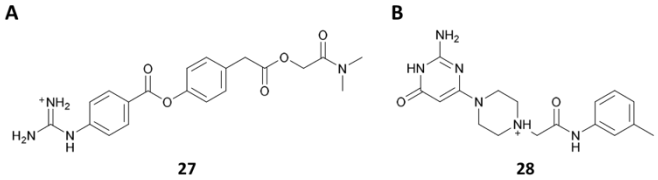


Figure S 11: Molecular structures of camostat mesylate **27** (A) and the previously reported preQ$_1$-riboswitch ligand **28** (B).

2. Tables S1-3

Table S 1: Physico-chemical parameter comparison between the unbiased in-house VS-library, RNA-focussed libraries from different suppliers and R-bind 2.0 (small molecules). Except of the total library size, numbers are mean values. RO5: Lipinski rule-of-five for oral bioavailability, MW: molecular weight (g/mol), HBA: hydrogen bond acceptors, HBD: hydrogen bond donors, tPSA: topological polar surface area (Å²).

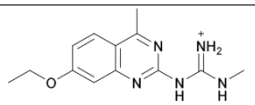
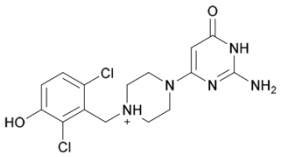
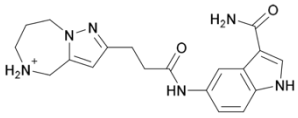
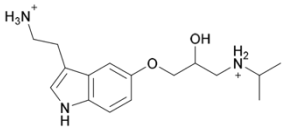
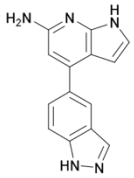
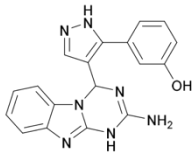
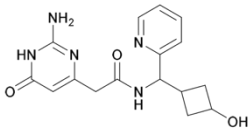
parameter	library						
	unbiased	Asinex	Enamine	LifeChem	Otava	Reaxense	R-bind
library size	4,1654169[b]	5559	15,520	4452	2542	2740	67
RO5 violations	0.1	0.3	0.1	0.1	0.1	0.1	0.3
MW	340.4	390.4	344.9	365.7	335.5	338.2	350.1
HBA	5.5	6.8	6.3	6.4	5.2	5.3	5.9
HBD	1.4	2.0	1.9	1.0	1.8	1.8	3.5
logP	2.1	0.8	1.7	2.7	2.4	2.3	0.2
$logD_{7.4}$	2.5	1.7	2.0	2.9	2.7	2.7	1.1
rotatable bonds	5.4	9.1	5.3	5.1	4.9	5.0	4.9
tPSA	68.5	79.2	77.0	68.1	69.0	69.6	79.0
heavy atoms	23.9	28.2	24.6	26.1	23.9	24.0	25.6
acidic centers	0.22	0.03	0.15	0.52	0.16	0.16	0.16
basic centers	0.25	0.64	0.43	0.48	0.32	0.32	1.99
aromatic atoms	11.0	12.0	13.3	12.4	12.9	13.0	13.9
aromatic/heavy	0.45	0.42	0.54	0.47	0.55	0.55	0.54
nitrogens	3.0	4.5	4.4	4.0	2.8	2.9	4.3
oxygens	2.5	2.2	1.8	2.4	2.4	2.4	1.6
stereo centers	0.6	0.8	0.6	0.3	0.3	0.3	0.3
formal charge	0.12	0.55	0.26	0.13	0.20	0.20	1.06
number of rings	3.0	3.7	3.4	3.6	3.3	3.3	3.7

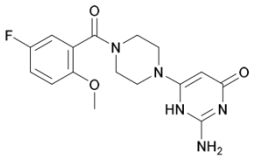
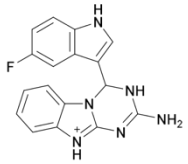
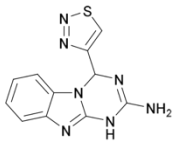
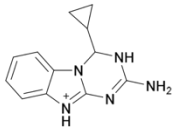
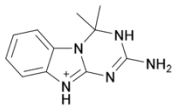
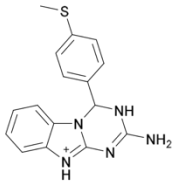
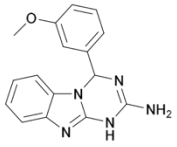
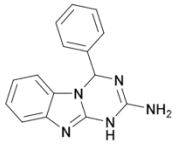
[b]For molecular docking, molecules with less than 2 HBD and without aromatic atoms were removed according to the HCV IRES pharmacophore hypothesis resulting in 3,481,646 molecules.

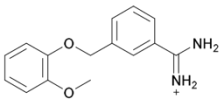
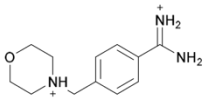
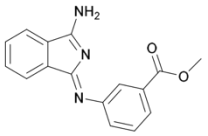
Table S 2: Applied physico-chemical property filters using FILTER (OMEGA 4.1.0.0: OpenEye Scientific Software, Santa Fe, NM, USA. http://www.eyesopen.com, 2019). PAINs were removed from all libraries in a separate, preceding step.

AGGREGATORS true "Eliminate known aggregators"
MAX_LIPINSKI_ACCEPTORS 10 "Maximum number of oxygen & nitrogen atoms"
MIN_LIPINSKI_DONORS 2 "Minimum number O & N atoms with hydrogens"
MAX_LIPINSKI_DONORS 5 "Maximum number O & N atoms with hydrogens"
MIN_MOLWT 130 "Minimum molecular weight"
MAX_MOLWT 400 "Maximum molecular weight"
MIN_XLOGP -3.0 "Minimum XLogP"
MAX_XLOGP 3 "Maximum XLogP"
MIN_2D_PSA 50.0 "Minimum 2-Dimensional (SMILES) Polar Surface Area"
MAX_2D_PSA 200.0 "Maximum 2-Dimensional (SMILES) Polar Surface Area"
MIN_ROT_BONDS 0 "Minimum number of rotatable bonds"
MAX_ROT_BONDS 8 "Maximum number of rotatable bonds"
MIN_COUNT_FORMAL_CRG 0 "Minimum number formal charges"
MAX_COUNT_FORMAL_CRG 3 "Maximum number of formal charges"
MIN_SUM_FORMAL_CRG 0 "Minimum sum of formal charges"
MAX_SUM_FORMAL_CRG 3 "Maximum sum of formal charges"
MIN_CHIRAL_CENTERS 0 "Minimum chiral centers"
MAX_CHIRAL_CENTERS 3 "Maximum chiral centers"
ELIMINATE_METALS Sc,Ti,V,Cr,Mn,Fe,Co,Ni,Cu,Zn,Y,Zr,Nb,Mo,Tc,Ru,Rh,Pd,Ag,Cd
ALLOWED_ELEMENTS H,C,N,O,F,P,S,Cl,Br,I

Table S 3: Molecular docking results of virtual screening hits selected for testing including identity and purity confirmation by LC-MS analytics (for details see supporting information 4. Analytical data of virtual screening hits). Molecules are depicted as their likely protomeric/tautomeric states under physiological conditions. For **15** – **22** different protomers/tautomers are reasonable. CoA: certificate of analysis.

compound	SMILES	FRED docking		LeadIT docking		Analytics		
		score kcal/mol	ph4-features	score kJ/mol	ph4-features	mass calc. / mass found	purity [%]	Supplier CoA
 **8**	O(CC)c1cc2nc(NC(=[N+])NC)nc(C)c2cc1	-14.20	2	-28.73	2	260.2 / 260.0 [M + H⁺]	98.8	identity: ^{1}H NMR purity state-ment: >90 %
 **9**	Clc1c(C[N+]2CCN(C=3N=C(N)NC(=O)C=3)CC2)c(Cl)ccc1O	-13.88	2	-31.82	2	370.1 / 370.0 [M + H⁺]	93.7	LC-MS (mass: 370.0, purity: 96.6%)
 **10**	O=C(Nc1cc2c(C(=O)N)c[nH]c2cc1)CCc1nn2c(c1)C[N+]CCC2	-13.91	1	-39.90	1	367.2 / 367.1 [M + H⁺]	90.5	LC-MS (mass: 367.1, purity: 95.7%)
 **11**	O(CC(O)C[N+]C(C)C)c1cc2c(CC[N+])c[nH]c2cc1	-15.86	1	-28.57	1	292.2 / 292.1 [M + H⁺]	99.1	identity: ^{1}H NMR purity state-ment: >85%
 **12**	Nc1nc2[nH]ccc2c(-c2cc3c([nH]nc3)cc2)c1	-15.77	1	-40.28	2	250.1 / 250.0 [M + H⁺]	93.8	LC-MS (mass: 250.1, purity: 95.4%)
 **13**	Oc1cc(-c2c(C3n4c5c(nc4NC(N)=N3)cccc5)cn[nH]2)ccc1	-16.02	1	-34.17	2	346.1 / 346.0 [M + H⁺]	99.2	identity: ^{1}H NMR purity state-ment: >90%
 **14**	O=C(NC(c1nccccc1)C1CC(O)C1)CC=1N=C(N)NC(=O)C=1	-14.87	1	-33.61	2	330.2 / 330.0 [M + H⁺]	98.9	LC-MS (mass: 330.2, purity: 100%)

 **15**	Fc1cc(C(=O)N2CCN(C=3NC(N)=NC(=O)C=3)CC2)c(OC)cc1	-13.05	1	-32.14	2	348.1 / 348.0 [M + H$^+$]	98.2	LC-MS (mass: 348.0, purity: 100%)
 **16**	Fc1cc2c(C3n4c5c([n+]c4N=C(N)N3)cccc5)c[nH]c2cc1	-14.37	3	-30.37	3	321.1 / 321.0 [M + H$^+$]	98.3	identity: ^{1}H NMR purity state-ment: >85%
 **17**	NC1=NC(c2nnsc2)n2c3c(nc2N1)cccc3	-13.47	3	-31.02	2	272.3 [M + H$^+$] / 310.0 [M + K$^+$]	97.9	LC-MS (mass: 272.2, purity: 100%)
 **18**	NC1=Nc2[n+]c3c(n2C(C2CC2)N1)cccc3	-12.17	3	-29.74	2	228.1 / 228.0 [M + H$^+$]	98.6	LC-MS (mass: 228.2, purity: 99.1%)
 **19**	NC1=Nc2[n+]c3c(n2C(C)(C)N1)cccc3	-13.47	1	-29.81	3	216.1 / 216.0 [M + H$^+$]	100.0	identity: ^{1}H NMR purity state-ment: >90 %
 **20**	S(C)c1ccc(C2n3c4c([n+]c3N=C(N)N2)cccc4)cc1	-11.54	1	n.d.	n.d.	310.4 / 310.0 [M + H$^+$]	97.8	LC-MS (mass: 310.2, purity: 100%)
 **21**	O(C)c1cc(C2n3c4c(nc3NC(N)=N2)cccc4)ccc1	-11.27	1	n.d.	n.d.	294.1 / 294.0 [M + H$^+$]	99.7	LC-MS (mass: 294.1 , purity: 100%)
 **22**	NC1=NC(c2ccccc2)n2c3c(nc2N1)cccc3	-11.26	3	-29.04	0	264.1 / 264.0 [M + H$^+$]	97.9	LC-MS (mass: 264.0, purity: 97.7%)

 **23**	O(Cc1cc(C(=[N+])N)ccc1)c1c(OC)cccc1	-13.64	1	-27.57	2	257.1 / 257.0 [M + H$^+$]	99.8	LC-MS (mass: 257.1, purity: 100%)
 **24**	[N+]=C(N)c1ccc(C[N+]2CCOCC2)cc1	-13.64	2	-28.71	2	220.2 / 220.0 [M + H$^+$]	97.6	LC-MS (mass: 220.0, purity statement: >90 %)
 **25**	O=C(OC)c1cc(N=C2N=C(N)c3c2cccc3)ccc1	-14.53	3	-29.00	1	280.1 / 280.0 [M + H$^+$]	92.6	identity: ^{1}H NMR purity statement: >85%

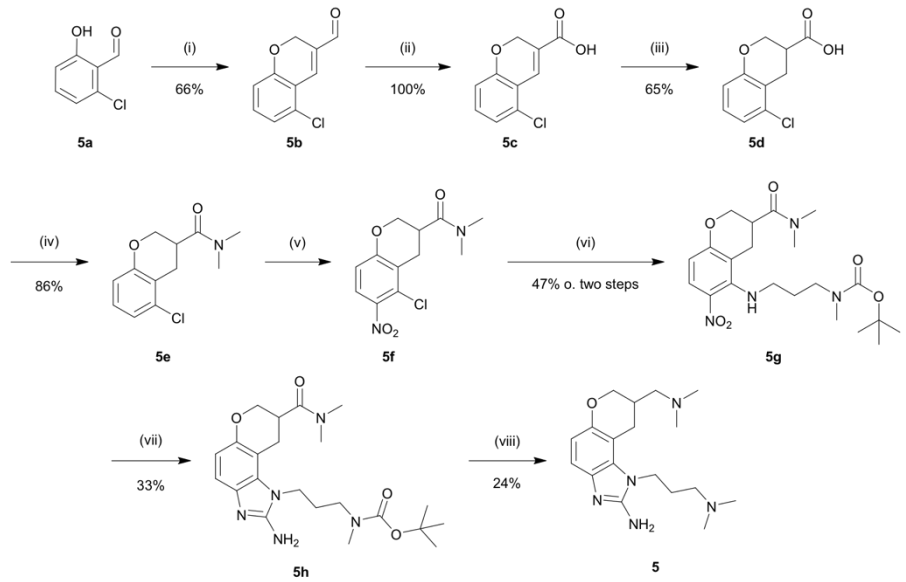
3. Extended material and methods: synthesis of compound 5

General experimental methods

All reagents and solvents were commercial grade and used without further purification. Reaction progress was monitored by thin layer chromatography (**TLC**) using Alugram Xtra F254 silica plates from Machery-Nagel or neutral aluminium oxide 60 F254 plates from Merck KGaA. In addition, high-performance liquid chromatography/electron spray ionization mass spectrometry (**HPLC/ESI-MS**) was used to control reaction's conversion and to determine the identity as well as the purity of compounds tested in the assay. An Agilent 1100 series HPLC system and either an Agilent Zorbax SB-Aq (4.6 mm x 150 mm, 5 µm) or Poroshell 120 EC-C18 (2.10 mm x 150 mm, 4 µM) column coupled to an Agilent 1100 series LC/MSD Trap with ESI, was used. The measurements were conducted with a gradient of acetonitrile and water (+0.1% formic acid) with 10%–90% acetonitrile over 10 min with a flow rate of 0.7 mL/min. Signals were detected at 254 nm with quantification by area under curve (AUC) and masses were determined in positive ionization mode (ESI) unless otherwise stated. **Column chromatography** was performed with silica gel 60 (40–63 µm) from Machery-Nagel. **Flash column chromatography** was performed with the Biotage Isolera™ One system. For reversed-phase flash chromatography prepacked columns of the type Biotage®Sfär C18 Duo from Biotage® were used. In the case of normal-phase flash chromatography, columns were self-packed with silica gel 60 (15–40 µm) from Merck KGaA. **Preparative HPLC** purification was performed with an Agilent 1290 II Infinity Preparative LC System using a MZ-Aqua Perfect C18 20 x 250 mm, 7 µm preparative LC column and acetonitrile/water +0.1% formic acid as mobile phase. **Melting points** (uncorrected) were measured with an MPM-H3 using semi-open capillaries. Nuclear magnetic resonance (**NMR**) spectra were recorded as stated individually on Bruker Fourier 300 MHz or Bruker Avance III 600 MHz. The chemical shift was abbreviated to δ and has the unit ppm. Tetramethyl silane (TMS) (δ = 0 ppm) was used as the reference substance. The chemical shifts were referenced to the solvent peaks in ^{1}H: δ = 7.26 (Chloroform-*d*), 2.50 (DMSO-*d*6) ppm and in ^{13}C: δ = 77.16 (Chloroform-*d*), 39.52 (DMSO-d6) ppm purchased from Deutero GmbH. In addition, the following abbreviations for the multiplicities of the peaks were defined: s (singlet), d (doublet), dd (doublet of doublet), t (triplet), td (triplet of doublet), q (quartet), qd (quartet of doublet), h (sextet), hept (septet) and m (multiplet). In the evaluation of the 1H-NMR spectrum, the coupling constant *J* was given in Hz and denoted together with the number of signaling hydrogen atoms. The MestReNova 12.0.4-22023 NMR spectrum processing program from *Mestrelab Research* was used to evaluate the NMR spectra and to determine the purity of the compound through LC-MS.

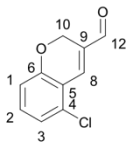
Overview

The following synthetic route is based on previous reported procedures[1,2], which were slightly modified at some points.



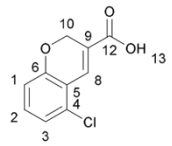
Scheme 1: Reagents and conditions: (i) Acrolein, DABCO, dioxane, 95 °C; (ii) silver nitrate, NaOH, EtOH/H_2O, 85° C; (iii) Na/Hg, NaOH/H_2O, room temperature (rt); (iv) EDC HCl, HOBt, dimethylamine hydrochloride, *N*-Methyl morpholine, dichloromethane, rt; (v) sodium nitrate, trifluoro acetic acid, dichloroacetic acid, 0 °C-rt; (vi) *N*-(3-aminopropyl)-*N*-methylcarbamate, 1-methyl-2-pyrrolidinone, 75 °C; (vii) H_2 Pd/C, EtOH, then cyanogen bromide/EtOH, rt; (viii) LAH, THF, 60 °C.

5-Chloro-2*H*-chromene-3-carbaldehyde (5b)



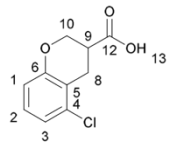
A mixture of 6-chlorosalicylaldehyde (**5a)** (4.84 g, 30.9 mmol, 1.00 equiv), acrolein (2.89 g, 3.44 mL, 46.4 mmol, 1.50 equiv) and 1,4-diazabicyclo[2.2.2]octane (1.73 g, 15.5 mmol, 0.500 equiv) in dioxane (16.9 mL) was heated under inert gas atmosphere at 95 °C for 2 h in a sealed flask. The reaction mixture was cooled to room temperature and diluted with dichloromethane (75.0 ml). The organic phase was washed with 3M HCl (3 x 30.0 mL). The aqueous layer was extracted with dichloromethane (3 x 30.0 mL). Combined organic layers were washed with aqueous saturated sodium chloride solution, dried over sodium sulfate, filtered, and concentrated under reduced pressure. Crude product was purified *via* column chromatography (isocratic, cyclohexane (Cy): ethyl acetate (EA); 50:1). Target compound **5b** (3.95 g, 20.3 mmol, **66% yield**) was isolated as yellow solid. T_m = 60–62 °C. **^{1}H-NMR, COSY** (300 MHz, Chloroform-*d*, TMS) δ_H 9.63 (1H, s, 12-H), 7.59 (1H, q, *J* = 1.3 Hz, 8-H), 7.19 (1H, t, *J* = 8.1 Hz, 2-H), 6.99 (1H, dd, *J* = 8.1, 1.0 Hz, 1-H), 6.78 (1H, dt, *J* = 8.1, 1.0 Hz, 3-H), 4.99 (2H, d, *J* = 1.3 Hz, 10-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, Chloroform-*d*, TMS) δ_C 189.83 (C-12), 157.21 (C-6), 137.18 (C-8), 133.95 (C-4), 133.14 (C-2), 132.42 (C-9), 122.74 (C-1), 119.37 (C-5), 115.45 (C-3), 63.13 (C-10). **MS (ESI)** found: m/z = 194.9 [M+H$^+$], calculated: m/z = 195.0 [M+H$^+$].

5-Chloro-2*H*-chromene-3-carboxylic acid (5c)



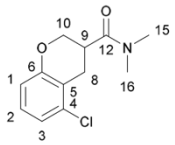
A solution of sodium hydroxide (3.25 g, 81.2 mmol, 4.00 equiv) in water (25.7 mL) was added to absolute ethanol (52.1 mL). A solution of silver nitrate (7.24 g, 42.6 mmol, 2.10 equiv) in water (25.7 mL) was added portionwise. Then **5b** (3.95 g, 20.3 mmol, 1.00 equiv) was added and the mixture was heated at 85°C for 75 min. The supernatant was decanted after cooling down at room temperature and the solid was washed with a 1:1 ethanol/water solution (4 × 20.0 mL). The washings were combined with the decanted supernatant and were diluted with an excess of 1M aqueous HCl (150 mL) resulting in a voluminous light-yellow precipitate. The aqueous suspension was extracted with methylene chloride (3 x 75.0 mL). Combined organic layers were dried over sodium sulfate, filtered, and concentrated under reduced pressure. Target compound **5c** (4.26 g, 20.2 mmol, **100% yield**) was isolated as light-yellow solid. T_m = 190–192 °C. **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 7.53 (1H, q, *J* = 1.5 Hz, 8-H), 7.27 (1H, t, *J* = 8.1 Hz, 2-H), 7.08 (1H, dd, *J* = 8.1, 1.1 Hz, 1-H), 6.86 (1H, dt, *J* = 8.1, 1.1 Hz, 3-H), 4.92 (2H, d, *J* = 1.5 Hz, 10-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 165.09 (C-12), 155.70 (C-6), 132.36 (C-2), 132.08 (C-4), 127.34 (C-8), 125.43 (C-9), 122.39 (C-1), 118.99 (C-5), 115.12 (C-3), 64.09 (C-10). **MS (ESI)** found: m/z = 210.9 [M+H$^+$], calculated: m/z = 211.0 [M+H$^+$].

5-Chlorochromane-3-carboxylic acid (5d)



To a solution of **5c** (4.25 g, 20.2 mmol, 1.00 equiv) in 10% aqueous sodium hydroxide solution (100 mL) was added portionwise 20% sodium amalgam (8.80 g) over a period of 30 min (caution: gas and heat evolution). The grey suspension was stirred for 3 h at room temperature. Afterwards, the progress of the reaction was controlled via TLC. If starting material is left, an appropriate amount of 20% sodium amalgam was added portionwise (additional total amount of 20% sodium amalgam of 6.20 g was added in four portions). This process was repeated till reaction's completion, which resulted in a white suspension. The supernatant was decanted from the liquid mercury, and the mercury was washed with 10% aqueous sodium hydroxide solution (3 x 20.0 mL). The combined supernatants were acidified to a pH of 2 with concentrated hydrochloric acid, and extracted with dichloromethane (3 x 50.0 mL). Combined organic layers were dried over sodium sulfate, filtered, and concentrated under reduced pressure. Target compound **5d** (2.80 g, 13.2 mmol, **65% yield**) was isolated as colorless solid. $\boldsymbol{T_m}$ = 190–192 °C. **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 12.69 (1H, br s, 13-H), 7.11 (1H, t, *J* = 8.1 Hz, 2-H), 7.00 (1H, dd, *J* = 8.1, 1.3 Hz, 1-H), 6.77 (1H, dd, *J* = 8.1, 1.3 Hz, 3-H), 4.33 – 4.12 (2H, m, 10-H), 3.07 (1H, tdd, *J* = 7.1, 6.7, 3.4 Hz, 9-H), 2.91 (2H, d, *J* = 6.7 Hz, 8-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 173.04 (C-12), 155.35 (C-6), 133.68 (C-4), 127.79 (C-2), 120.97 (C-1), 119.78 (C-5), 115.32 (C-3), 66.07 (C-10), 37.20 (C-9), 24.89 (C-8). **MS (ESI)** (in negative mode) found: m/z = 210.7 [M-H$^-$], calculated: m/z = 211.0 [M-H$^-$].

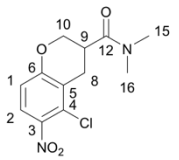
5-Chloro-*N,N*-dimethylchromane-3-carboxamide (5e)



To a solution of **5d** (1.40 g, 6.58 mmol, 1.00 equiv), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (1.89 g, 9.88 mmol, 1.50 equiv), dimethylamine hydrochloride (1.34 g, 16.5 mmol, 2.50 equiv) and *N*-hydroxy benzotriazole (1.51 g, 9.88 mmol, 1.50 equiv) in dichloromethane (60.0 mL) was added 4-methylmorpholine (3.33 g, 3.66 mL, 32.9 mmol, 5.00 equiv) under inert gas atmosphere and the mixture stirred for 66 h at room temperature. The reaction mixture was diluted with additional dichloromethane, and an equal volume of saturated aqueous sodium bicarbonate solution was added. The organic phase was separated and the aqueous phase was washed with dichloromethane (3x 25.0 mL). The combined organic phases were washed with aqueous saturated sodium chloride solution (1 x 30.0 mL), dried over sodium sulfate, filtered, and concentrated under reduced pressure. The crude product was purified *via* flash chromatography (gradient 0–25% EA in Cy; product eluates at 15% EA) and lyophilized afterwards. Target compound **5e** (1.36 g, 5.66 mmol, **86% yield**) was isolated as colorless oil. **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 7.12 (1H, t, *J* = 8.0 Hz, 2-H), 7.00 (1H, d, *J* = 8.0 Hz, 1-H), 6.79 (1H d, *J* = 8.0 Hz, 3-H), 4.31 (1H, d, *J* = 10.3 Hz, 10-H), 3.84 (1H, t, *J* = 10.3 Hz, 10-H), 3.40 – 3.25 (1H, m, 9-H), 3.10 (3H, s, 15/16-H), 2.95 – 2.67 (5H, m, 15/16-H, 8-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 170.83 (C-12), 155.25 (C-6), 133.66 (C-4), 127.70 (C-2), 120.80 (C-

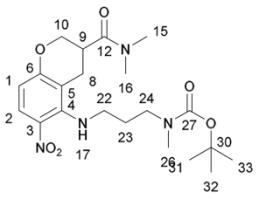
1), 120.19, (C-5) 115.19 (C-3), 66.45 (C-10), 36.74 (C-15/16), 35.07 (C-15/16), 34.15 (C-9), 26.18 (C-8). **MS (ESI)** found: m/z = 240.0 [M+H$^+$], calculated: m/z = 240.1 [M+H$^+$].

5-Chloro-*N,N*-dimethyl-6-nitrochromane-3-carboxamide (5f)



Pestled sodium nitrate (2.89 g, 34.0 mmol, 3.00 equiv) and **5e** (2.72 g, 11.3 mmol, 1.00 equiv) were mixed in dichloroacetic acid (140 mL) under ice cooling and inert gas atmosphere. Trifluoroacetic acid (465 mg, 312 µL, 4.08 mmol, 0.360 equiv) was added and the mixture stirred for 28 h covering the ice bath and allowing it to warm up overnight. The resulting dark crude reaction mixture was then slowly poured into a rapidly stirred solution of sodium carbonate (10%, 1.20 L) (Caution CO2 evolution). After rapid stirring for 10 minutes, the aqueous phase was extracted with dichloromethane (3 x 75.0 mL). Combined organic layers were dried over sodium sulfate, filtered, and concentrated under reduced pressure. Crude product of **5f** (red resin-like oil, 1.57 g) was used without further purification. For analytics a small portion was purified *via* column chromatography (isocratic, Cy:EA; 3:2) to separate the regioisomers. Para-substituted isomer eluates first, then product. Major isomer: **5-chloro-*N,N*-dimethyl-6-nitrochromane-3-carboxamide (5f)** light yellow solid. T_m = 126–128 °C. **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 7.87 (1H, d, *J* = 9.0 Hz, 2-H), 6.98 (1H, d, *J* = 9.0 Hz, 1-H), 4.47 – 4.34 (1H, m, 10-H), 4.08 – 3.95 (1H, m, 10-H), 3.47 – 3.36 (1H, m, 9-H), 3.11 (3H, s, 15/16-H), 3.04 – 2.77 (5H, m, 15/16/H, 8-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 170.38 (C-12), 158.18 (C-6), 141.14 (C-3), 126.93 (C-4), 124.59 (C-2), 122.79 (C-5), 115.47 (C-1), 66.94 (C-10), 36.77 (C-15/16), 35.12 (C-15/16), 33.49 (C-9), 26.58 (C-8). **MS (ESI)** found: m/z= = 285.0 [M+H$^+$], calculated: m/z = 285.1 [M+H$^+$]. Undesired isomer: 5-chloro-*N,N*-dimethyl-8-nitrochromane-3-carboxamide: **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 7.79 (1H, d, *J* = 8.8 Hz), 7.21 (1H, d, *J* = 8.8 Hz), 4.60 – 4.35 (1H, m), 4.17 – 3.94 (1H, m), 3.54 – 3.39 (1H, m), 3.11 (3H, s), 3.01 – 2.71 (5H, m). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 170.23, 148.72, 138.42, 137.37, 123.79, 123.49, 120.41, 67.40, 36.76, 35.13, 33.19, 26.54.

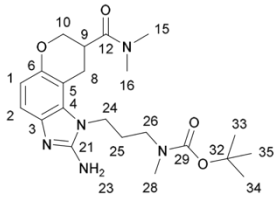
***Tert*-butyl(3-((3-(dimethylcarbamoyl)-6-nitrochroman-5-yl)amino)propyl)(methyl) carbamate (5g)**



Crude product of **5f** (600 mg, 2.11 mmol, 1.00 equiv) and *N*-(3-aminopropyl)-*N*-methylcarbamate (1.98 g, 10.5 mmol, 5.00 equiv) in 1-methyl-2-pyrrolidinone (2.00 mL) were mixed in a sealed tube under inert gas atmosphere. The mixture was heated at 75 °C for 24 h. After cooling down, the reaction mixture was diluted with diethyl ether (100 mL) and washed with water (4 × 100 mL) and 5% aqueous potassium biphosphate solution (80 mL). The organic phase was dried with sodium sulfate, filtered and concentrated under reduced pressure. Crude product was purified *via* column chromatography (first

isocratic Cy:EA; 1:1, when product starts to eluate percentage of EA can be increased). Target compound **5g** (430 mg, 985 µmol, **47% yield** over two steps) was isolated as bright yellow oil. **^{1}H-NMR, COSY** (300 MHz, Chloroform-*d*) δ_H 7.88 (1H, d, *J* = 9.4 Hz, 2-H), 6.28 (1H, d, *J* = 9.4 Hz, 1-H), 4.40 – 4.28 (1H, m, 10-H), 4.11 – 3.97 (1H, m, 10-H), 3.45 – 3.27 (1H, m, 24-H), 3.27 – 3.12 (2H, m, 22-H), 3.11 – 3.01 (2H, m, 9-H, 24-H), 3.01 – 2.84 (7H, m, 15-H, 16-H, 8-H), 2.83 – 2.60 (4H, m, 8-H, 26-H), 1.73 (2H, p, *J* = 6.9 Hz, 23-H), 1.36 (9H, s, 31-33-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, Chloroform-*d*) δ_C 171.30 (C-12), 160.03 (C-6), 155.63 (C-27), 148.70 (C-4), 131.40 (C-3), 126.31 (C-2), 110.94 (C-5), 109.24 (C-1), 79.44 (C-30), 67.28 (C-10), 46.17 (C-22), 45.69 (C-24), 37.24 (C-9), 35.62 (C-15/16), 35.15 (C-15/16), 34.29 (C-26), 28.36 (C-23), 27.77 (C-31-33). **MS (ESI)** found: m/z = 437.1 [M+H$^+$], calculated: m/z = 437.2 [M+H$^+$].

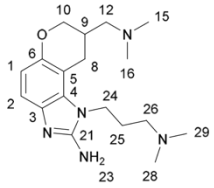
***Tert*-butyl (3-(2-amino-8-(dimethylcarbamoyl)-8,9-dihydrochromeno [5,6-*d*] imidazol-1(7*H*)-yl)propyl)(methyl)carbamate (5h)**



A solution of **5g** (130 mg, 298 µmol, 1.00 equiv) in absolute ethanol (13.5 mL) was mixed with 10% palladium on carbon (135 mg, 127 µmol, 0.425 equiv) under inert gas atmosphere. After addition, the atmosphere was replaced by a hydrogen atmosphere (1 atm) and the mixture stirred for 2 h at room temperature. The mixture was purged with argon upon hydrogen atmosphere removal. Afterwards, the mixture was quickly filtered through Celite® and washed with 10.0 mL ethanol in 3 portions directly in the next reaction flask. To this light yellow solution, cyanogen bromide (34.7 mg, 328 µmol, 1.10 equiv) in 1.0 mL ethanol was added under inert gas atmosphere. The mixture was purged again and covered with aluminum foil and stirred for 21 h. The solvent was removed under reduced pressure and the crude residue was dissolved in a mixture of dichloromethane (20.0 mL) and a saturated sodium bicarbonate solution (20.0 mL) and purged with argon. The mixture has turned from purple to light brown. After the mixture was stirred for 20 min, it was transferred to a separation funnel and the aqueous phase was extracted twice with dichloromethane (2 x 20.0 mL). The combined organic phases were dried over anhydrous sodium sulfate, filtered and concentrated under reduced pressure. The crude product was purified via column chromatography with basic aluminum oxide 60 (isocratic, 3% methanol in diethyl ether, when side-product eluted completely the percentage of methanol was increased to 10%). The product was further purified via reversed-phase flash (gradient 10–100% acetonitrile (ACN) in water+0.1% formic acid, product eluates at 40% ACN). Target compound **5h** (43.0 mg, 99.5 µmol, **33% yield**) was isolated as light purple solid (caution: air sensitive). $\boldsymbol{T_m}$ = 130–132 °C. **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 6.87 (1H, d, *J* = 8.4 Hz, 2-H), 6.43 (1H, d, *J* = 8.4 Hz, 1-H), 6.15 (2H, br, s, 23-H), 4.23 (1H, dd, *J* = 10.3, 3.1 Hz, 10-H), 4.10 – 3.97 (2H, br, m, 24-H), 3.78 (1H, t, *J* = 10.3 Hz, 10-H), 3.33 – 3.02 (8H, m, 8-H, 9-H, 15/16-H, 26-H), 2.88 (3H, s, 15/16-H), 2.76 (3H, s, 28-H), 1.82 (2H, m, 25-H), 1.48 – 1.15 (9H, br, m, 33-35-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 171.46 (C-12), 154.91 (C-29), 154.45 (C-21), 147.64 (C-6), 136.60 (C-3), 131.50 (C-4), 113.67 (C-2), 109.53 (C-1), 104.08 (C-5), 78.49 (C-32), 66.13 (C-10), 45.61 (C-26), 41.12 (C-24), 36.79 (C-15/16), 35.07 (C-15/16), 34.72 (C-9), 33.70 (C-28), 28.88 (C-25), 28.00 (C-33-35), 23.87 (C-8). **MS (ESI)** found: m/z = 432.2 [M+H$^+$], calculated: m/z = 432.3 [M+H$^+$]. Formation of *tert*-butyl (3-(8-

(dimethylcarbamoyl)-2-methyl-8,9-dihydrochromeno[5,6-*d*]imidazol-1(7*H*)-yl)propyl)(methyl)carbamate as side-product was observed: **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 7.21 (1H, d, *J* = 8.6 Hz), 6.62 (1H, d, *J* = 8.6 Hz), 4.30 (1H, dd, *J* = 10.1, 3.1 Hz), 4.24 – 4.06 (2H, br, m), 3.84 (1H, t, *J* = 10.1 Hz), 3.32 – 3.17 (5H, m), 3.13 (3H, s), 2.88 (3H, s), 2.79 (3H, s), 2.46 (3H, s), 1.98 – 1.77 (2H, br, m), 1.56 – 1.13 (9H, br, m). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 171.29, 154.60, 150.72, 149.71, 136.89, 133.00, 116.92, 111.58, 104.76, 78.62, 66.20, 45.69, 42.41, 36.81, 35.11, 34.45, 33.80, 29.91, 27.98, 23.78, 13.58. **MS (ESI)** found: m/z = 431.2 [M+H$^+$], calculated: m/z = 431.3 [M+H$^+$].

8-((Dimethylamino)methyl)-1-(3-(dimethylamino)propyl)-1,7,8,9-tetrahydrochromeno[5,6-*d*]imidazol-2-amine (5)



Lithium aluminum hydride (LAH) (9.48 mg, 250 µmol, 3.76 equiv) was added portionwise to a solution of **5h** (28.7 mg, 66.4 µmol, 1.00 equiv) in dry tetrahydrofuran (THF) (3.50 mL) under inert gas atmosphere. The mixture was heated at 60°C for 2 h in a sealed tube, then the same amount of LAH (9.48 mg, 250 µmol, 3.76 equiv) was added and heated for 5 h. Reaction's progress was monitored by LC-MS. After cooling down, the excess hydride was quenched with water (80 µL) followed by 15% sodium hydroxide solution (80 µL). Then additional water was added (240 µL) and a brown solid precipitated. The mixture was filtered over Celite® and the filter pad washed with THF. The filtrate was dried over sodium sulfate, filtered, and concentrated under reduced pressure. The crude product was taken in DMSO and purified *via* reversed phase flash chromatography (isocratic 0% ACN in water+0.1% formic acid; product eluates with DMSO at 0% ACN). The crude product was further purified *via* preparative HPLC (gradient over 10 min 0–5% ACN+0.1% formic acid in water+0.1% formic acid). Target compound **5** (5.23 mg, 15.8 µmol, **24% yield, 99% purity**) was isolated as colorless solid which gets brown upon air contact (keep under inert gas atmosphere). This reaction is recommended to do in the indicated scale. ***T*$_m$** = 107–109 °C. **^{1}H-NMR, COSY** (300 MHz, DMSO-d_6, TMS) δ_H 6.88 (1H, d, *J* = 8.1 Hz, 2-H), 6.45 (1H, d, *J* = 8.1 Hz, 1-H), 6.21 (br, s, 23-H), 4.15 (1H, dd, *J* = 10.6, 2.7 Hz, 10-H), 4.10 – 3.98 (2H, br, m, 24-H), 3.74 (1H, dd, *J* = 10.5, 7.6 Hz, 10-H), 3.17 (1H, dd, *J* = 16.0, 5.5 Hz, 8-H), 2.78 (1H, dd, *J* = 16.0, 7.6 Hz, 8-H), 2.43 – 2.27 (4H, m, 26-H, 12-H), 2.26 (6H, s, 28-29-H), 2.23 (7H, s, 15-16-H, 9-H), 1.94 – 1.74 (2H, br, m, 25-H). **^{13}C-NMR, HSQC, HMBC** (75 MHz, DMSO-d_6, TMS) δ_C 154.22 (C-21), 148.36 (C-6), 134.54 (C-3), 131.22 (C-4), 112.95 (C-2), 109.99 (C-1), 104.53 (C-5), 67.61 (C-10), 60.61 (C-12), 54.76 (C-26), 45.37 (C-15-16), 44.30 (C-28-29), 40.77 (C-24), 29.57 (C-9), 27.61 (C-25), 24.58 (C-8). **MS (ESI)** found: m/z = 332.2 [M+H$^+$], calculated: m/z = 332.2 [M+H$^+$].

Analytics compound 5

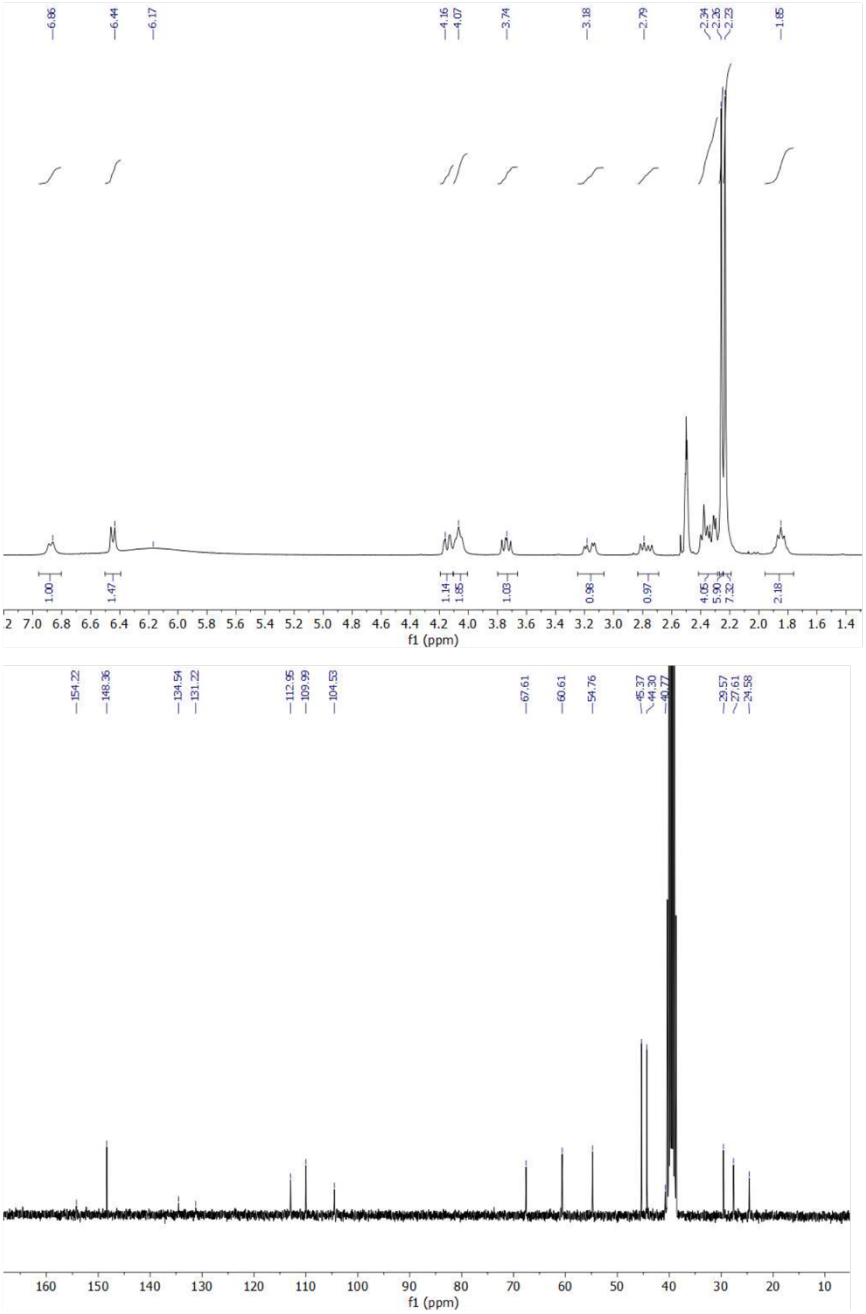


Figure S 7: NMR-spectra of final compound **5**: ^{1}H-NMR spectrum measured in DMSO-d6 (top), ^{13}C-NMR spectrum measured in DMSO-d6 (bottom).

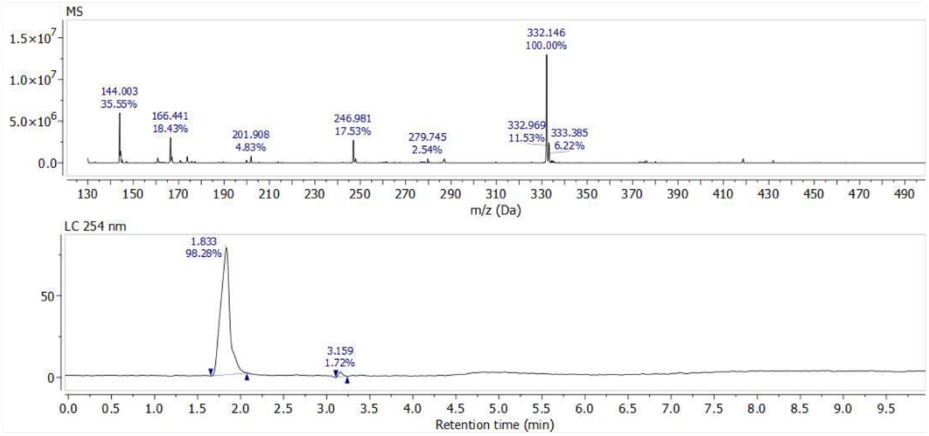


Figure S 8: LC-ESI-MS (measured with Zorbax column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min) traces of compound **5**, top: mass spectrum, bottom: LC chromatogram.

4. Analytical data of virtual screening hits

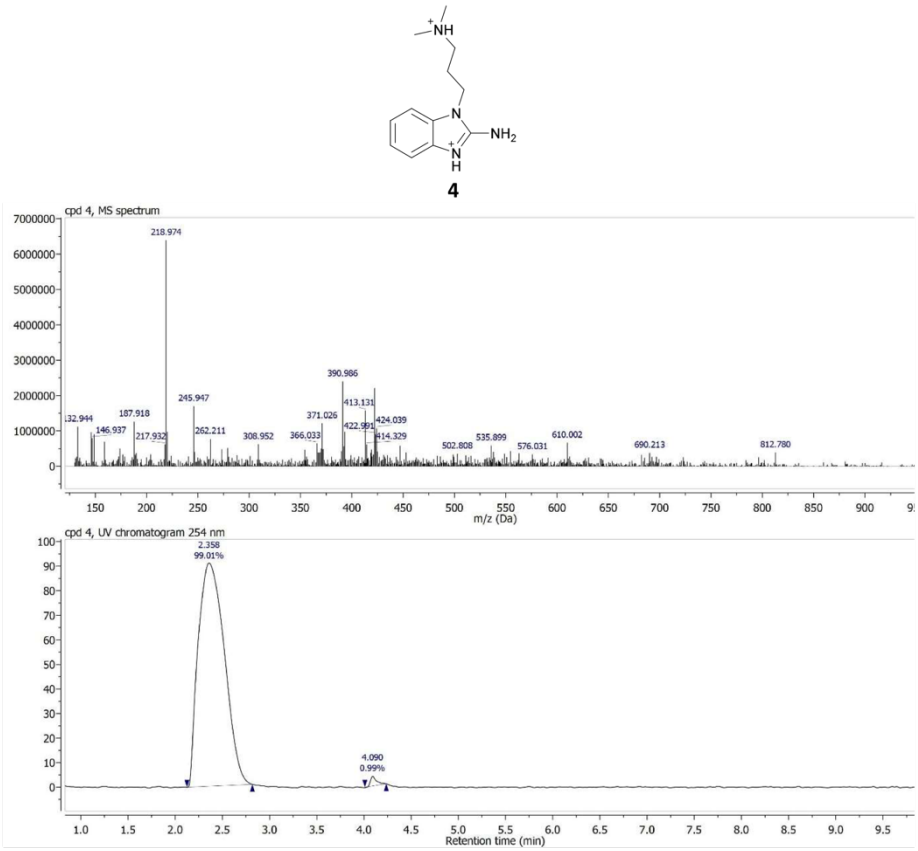
General experimental methods

High-performance liquid chromatography/electron spray ionization mass spectrometry (**HPLC/ESI-MS**) was used to determine the identity as well as the purity of compounds tested in the assay. An Agilent 1100 series HPLC system and either an Agilent Zorbax SB-Aq (4.6 mm x 150 mm, 5 µm) or Poroshell 120 EC-C18 (2.10 mm x 150 mm, 4 µM) column coupled to an Agilent 1100 series LC/MSD Trap with ESI, was used. The measurements were conducted with a gradient of acetonitrile and water (+0.1% formic acid) with 10%–90% acetonitrile over 10 min with a flow rate of 0.7 mL/min. Signals were detected at 254 nm with quantification by area under curve (AUC) and masses were determined in positive ionization mode (ESI) unless otherwise stated. Nuclear magnetic resonance (**NMR**) spectra were recorded on a Bruker Avance Neo 400 MHz NMR spectrometer. The chemical shift was abbreviated to δ and has the unit ppm. Tetramethyl silane (TMS) (δ = 0 ppm) was used as the reference substance. The chemical shifts were referenced to the solvent peaks in ^{1}H: δ = 3.31 (Methanol-d_4), 2.50 (DMSO-*d*6), 4.79 (Deuterium oxide) ppm and in ^{13}C: δ = 49.00 (Methanol-d_4), 39.52 (DMSO-d6) ppm purchased from Deutero GmbH. TMS was used as reference substance for ^{13}C spectra measured in deuterium oxide. In addition, the following abbreviations for the multiplicities of the peaks were defined: s (singlet), d (doublet), dd (doublet of doublet), t (triplet), dt (doublet of triplet), q (quartet), qd (quartet of doublet) and m (multiplet). In the evaluation of the 1H-NMR spectrum, the coupling constant *J* was given in Hz and denoted together with the number of signalling hydrogen atoms. The MestReNova 12.0.4-22023 NMR spectrum processing program from *Mestrelab Research* was used to evaluate the NMR spectra and to determine the purity of the compounds through LC-MS.

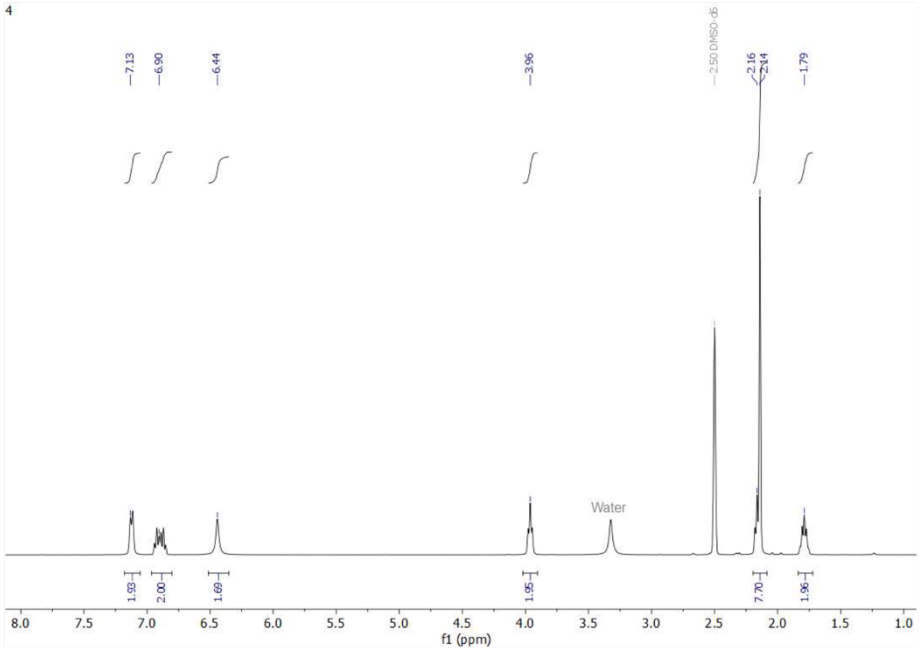
For compounds **8** and **8b**, additional purification by **preparative HPLC** was performed with an Agilent 1290 II Infinity Preparative LC System using a MZ-Aqua Perfect C18 20 x 250 mm, 7 µm preparative LC column and acetonitrile/water +0.1% formic acid as the mobile phase. The impurity of compound **8b** was isolated and resulted in compound **8c**.

Spectra of VS hits

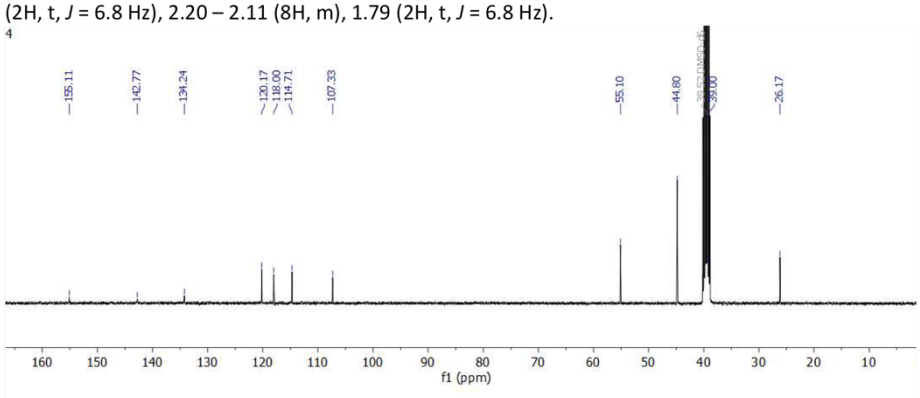
Analytical data of the purchased VS hits and compound 4 including LC-MS analytics (top: mass spectrum, bottom: LC chromatogram at 254 nm, column used is indicated) and NMR spectroscopy (^{1}H-NMR and ^{13}C-NMR, if possible). Due to low amounts of purchased compounds, determination of ^{1}H- and ^{13}C-NMR spectra was not always feasible. In these cases, ^{1}H-NMR spectra of the vendor are depicted, if provided. Therefore, ^{13}C-NMR spectra are not available for compounds **9**, **12**, **15**, **17**, **20**, **23** and **24**. For compound **9**, also no ^{1}H-NMR spectrum is available.



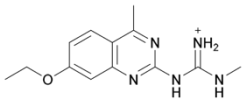
LC-ESI-MS (measured with Zorbax column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 219.3 [M + H^+], found: m/z = 219.0 [M + H^+], purity 99.0%.



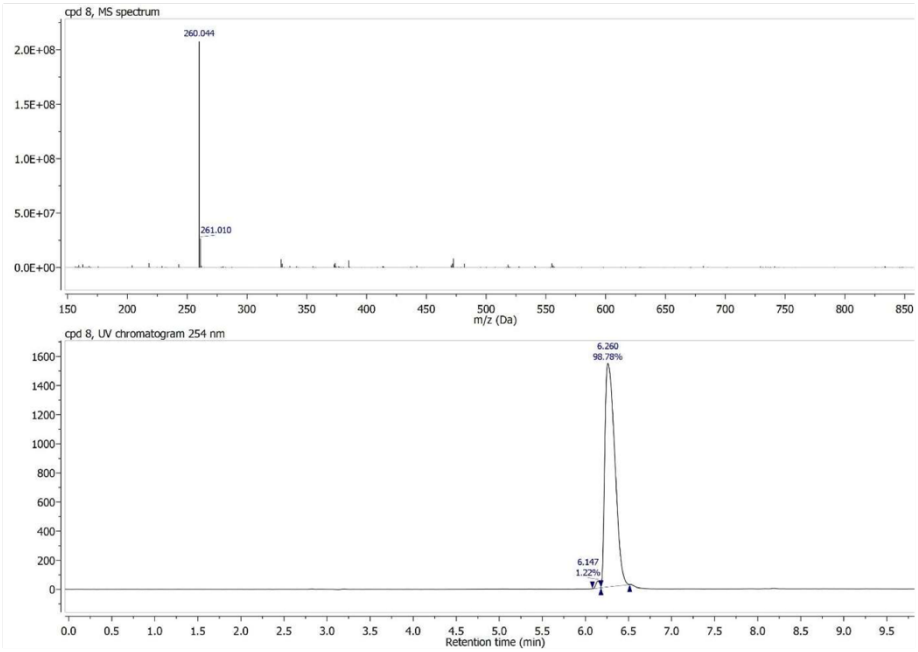
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 7.12 (2H, d, J = 7.5 Hz), 6.98 – 6.77 (2H, m), 6.44 (2H, s), 3.96 (2H, t, J = 6.8 Hz), 2.20 – 2.11 (8H, m), 1.79 (2H, t, J = 6.8 Hz).



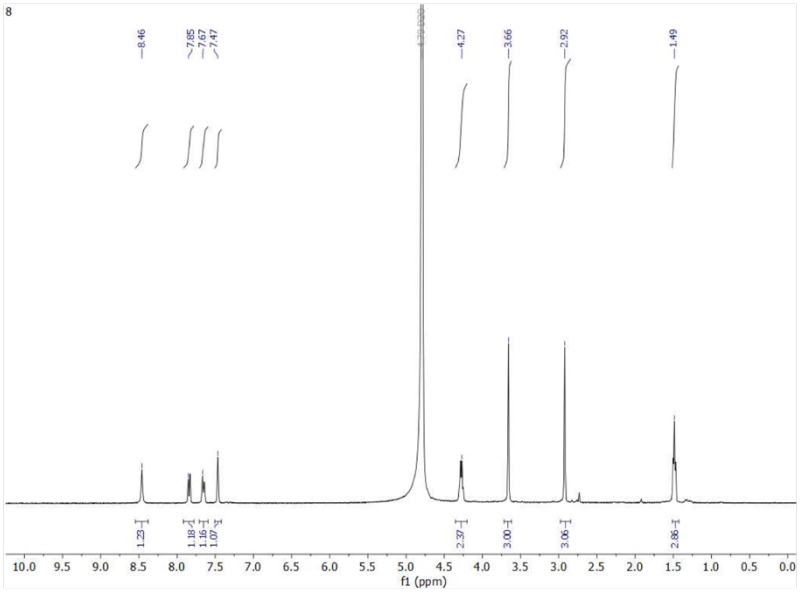
^{13}C NMR (101 MHz, DMSO-d_6, TMS) δ_C 155.11, 142.77, 134.24, 120.17, 118.00, 114.71, 107.33, 55.10, 44.80, 39.00, 26.17.



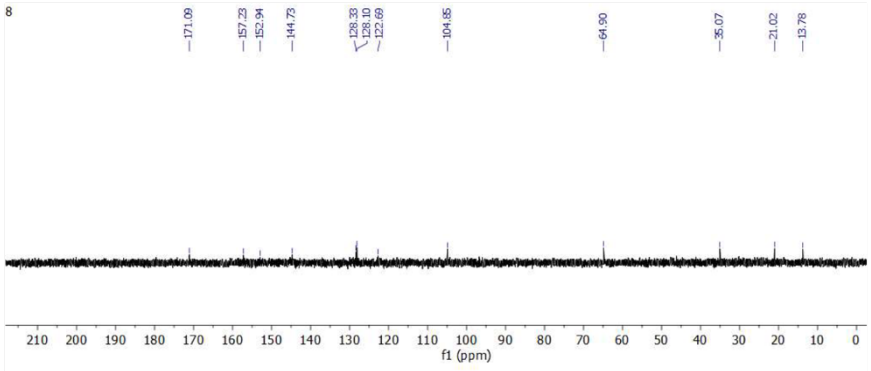
8



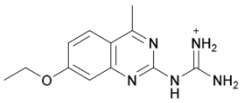
LC-ESI-MS (measured with Zorbax column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 260.2 [M + H^+], found: m/z = 260.0 [M + H^+], purity 98.8%.



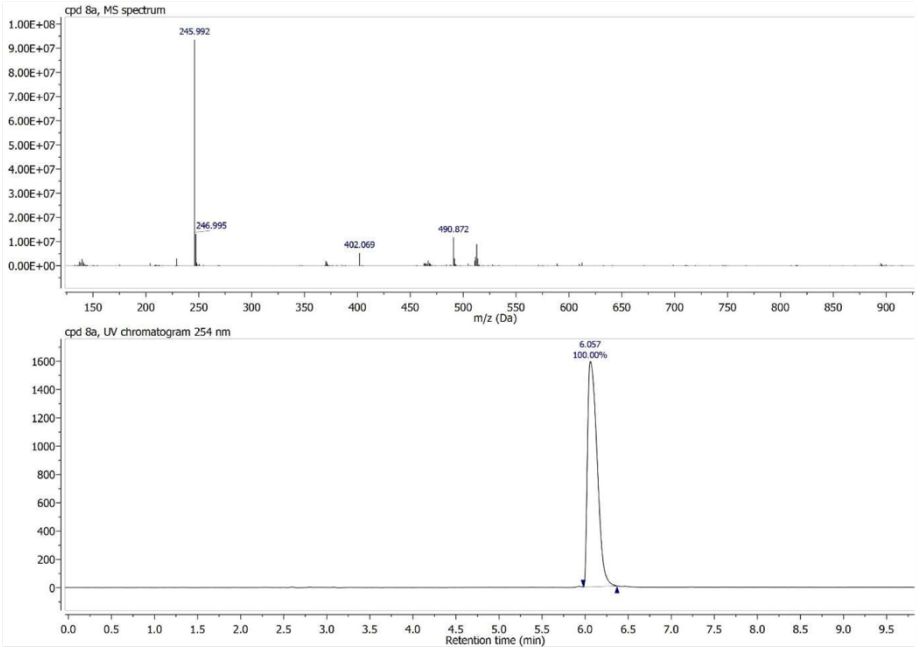
^{1}H NMR (400 MHz, Deuterium Oxide, TMS) δ_H 8.46 (1H, s), 7.93 – 7.78 (1H, m), 7.72 – 7.58 (1H, m), 7.52 – 7.41 (1H, m), 4.28 (2H, q, *J* = 7.5, 7.1 Hz), 3.66 (3H, s), 2.92 (3H, s), 1.49 (3H, t, *J* = 6.8, 3.6 Hz).



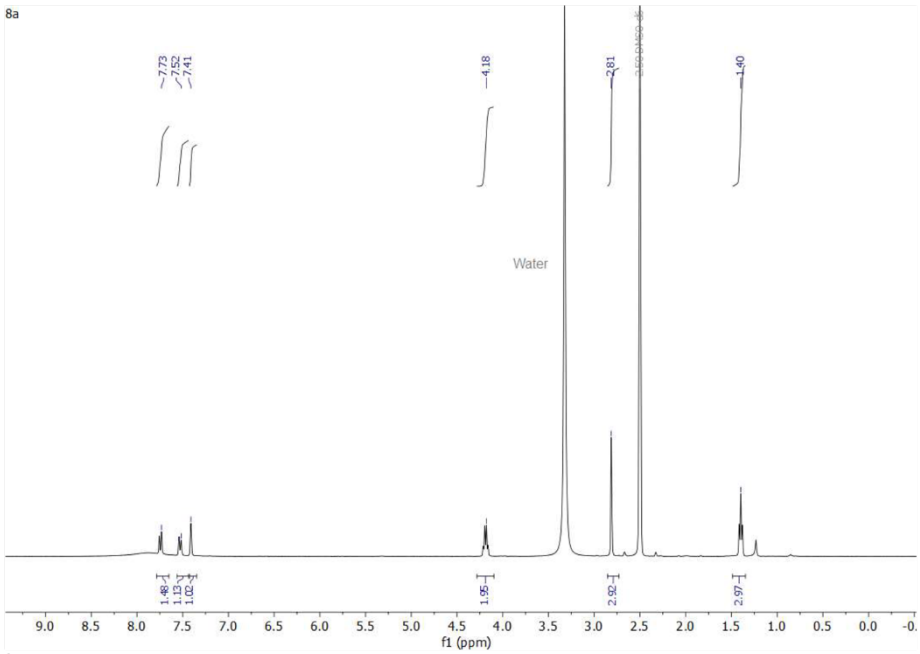
^{13}C NMR (101 MHz, Deuterium Oxide, TMS) δ_C 171.09, 157.23, 152.94, 144.73, 128.33, 128.10, 122.69, 104.85, 64.90, 35.07, 21.02, 13.78.



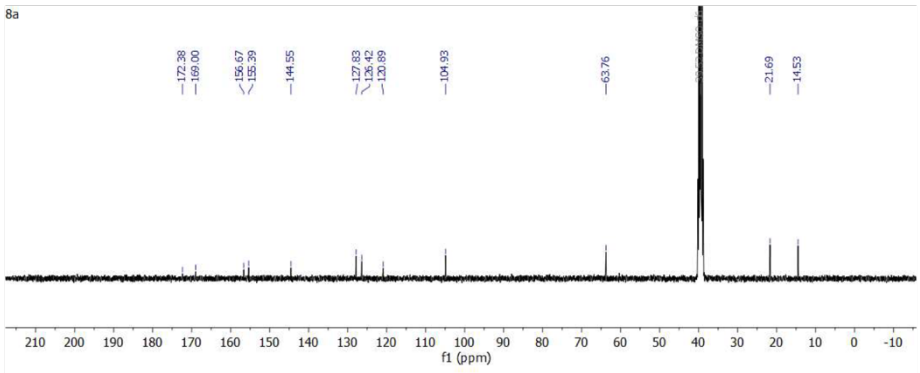
8a



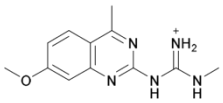
LC-ESI-MS (measured with Zorbax column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 246.1 [M + H^+], found: m/z = 246.0 [M + H^+], purity 100.0%.



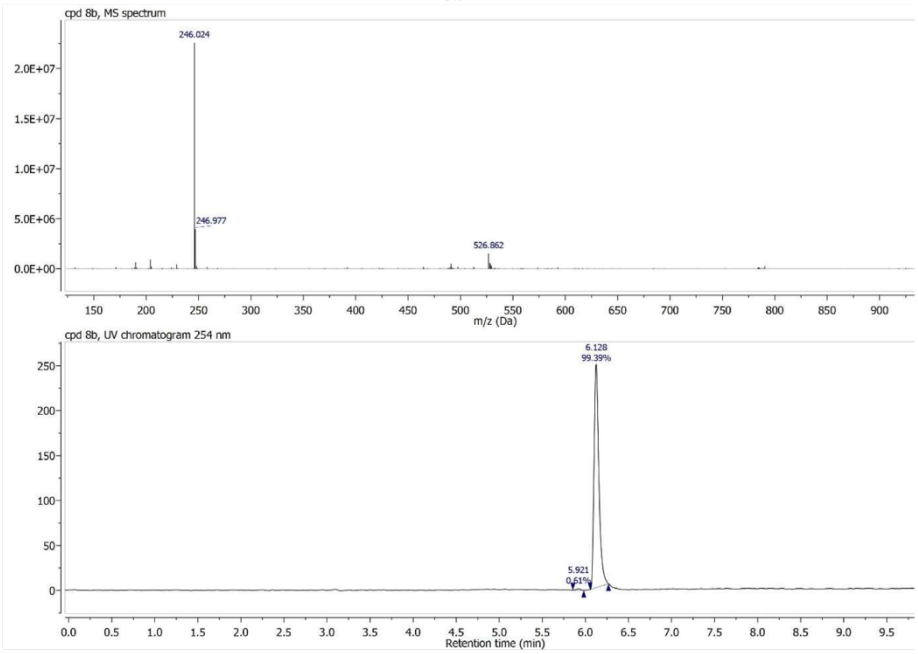
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 7.74 (1H, dd, *J* = 9.1, 2.0 Hz), 7.53 (1H, dt, *J* = 9.1, 3.8, 2.4 Hz), 7.41 (1H, s), 4.19 (2H, q, *J* = 6.4 Hz), 2.81 (3H, s), 1.40 (3H, t, *J* = 6.1 Hz).



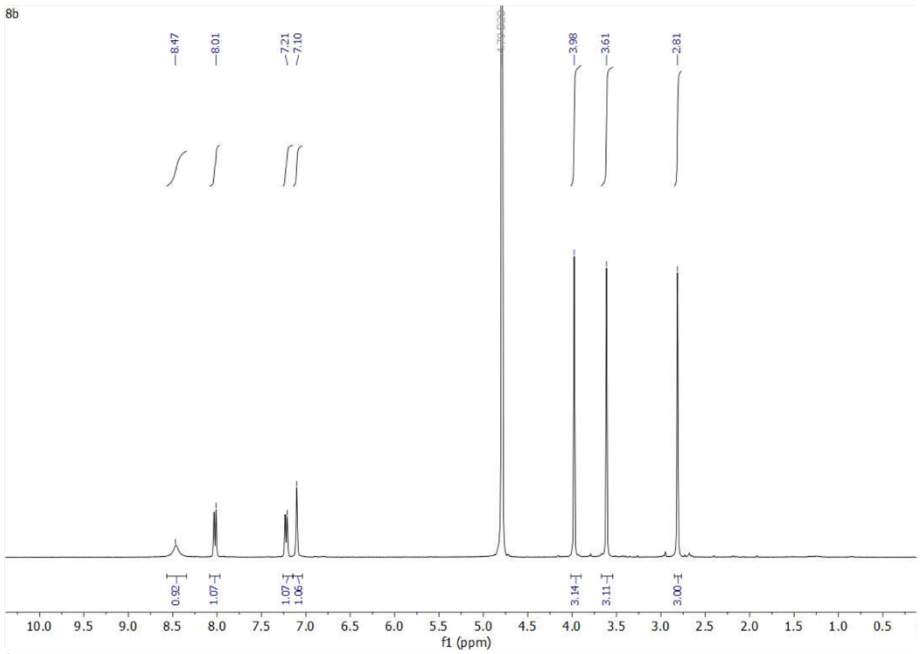
^{13}C NMR (101 MHz, DMSO-d_6, TMS) δ_C 172.38, 169.00, 156.67, 155.39, 144.55, 127.83, 126.42, 120.89, 104.93, 63.76, 21.69, 14.53.



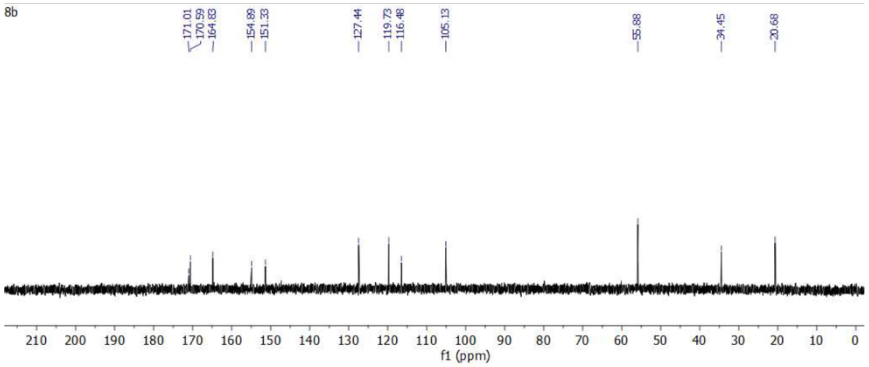
8b



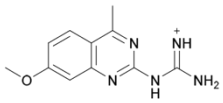
LC-ESI-MS (measured with Zorbax column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 246.1 [M + H^+], found: m/z= 246.0 [M + H^+], purity 99.4%.



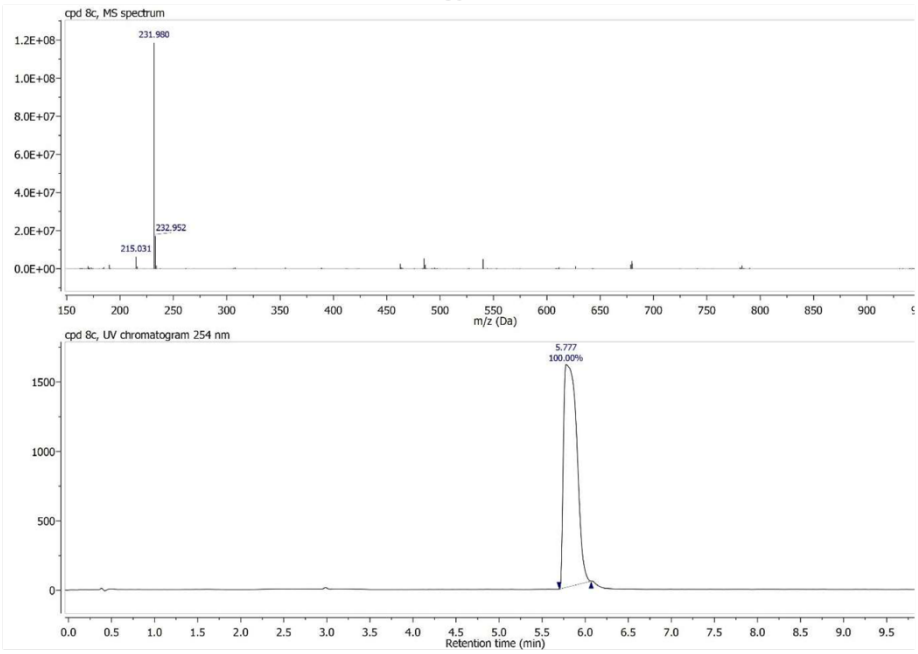
^{1}H NMR (400 MHz, Deuterium Oxide, TMS) δ_H 8.46 (1H, s), 8.02 (1H, d, *J* = 9.1 Hz), 7.22 (1H, d, *J* = 9.1 Hz), 7.10 (1H, s), 3.98 (3H, s), 3.61 (3H, s), 2.81 (3H, s).



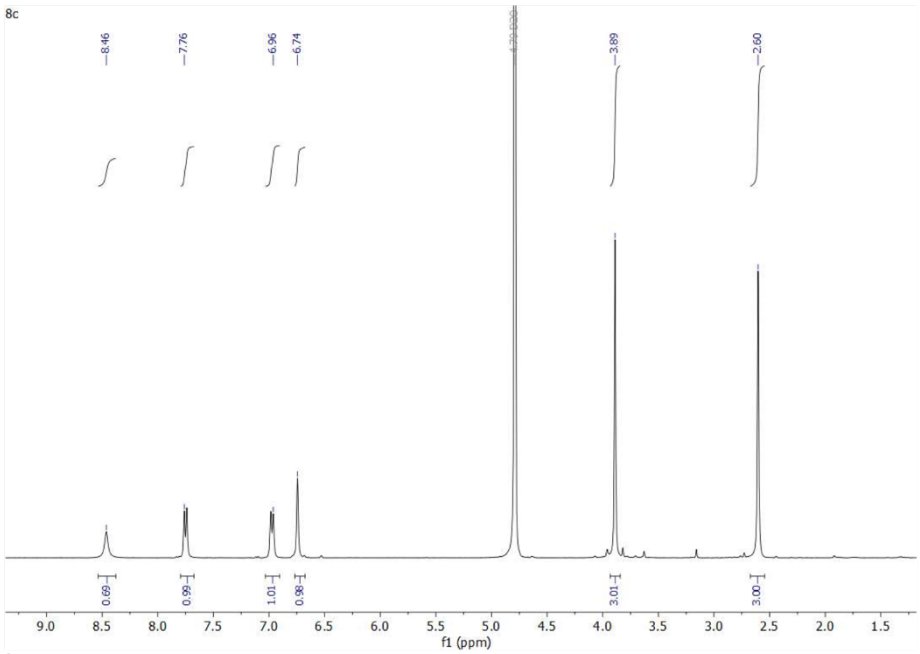
^{13}C NMR (101 MHz, Deuterium Oxide, TMS) δ_C 171.01, 170.59, 164.83, 154.89, 151.33, 127.44, 119.73, 116.48, 105.13, 55.88, 34.45, 20.68.



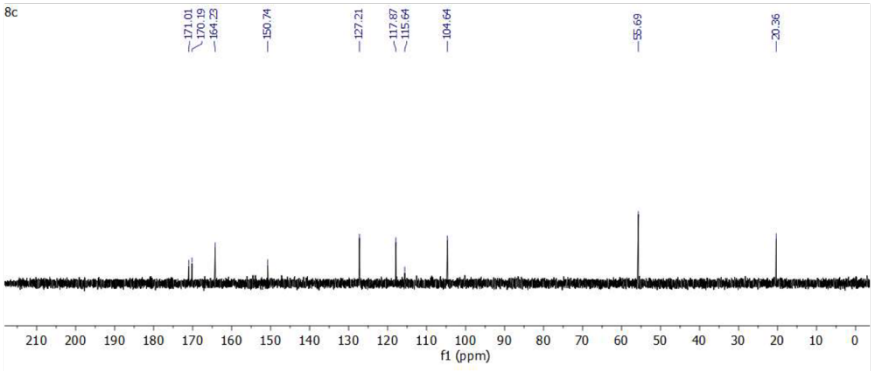
8c



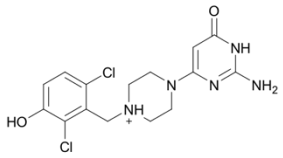
LC-ESI-MS (measured with Zorbax column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 232.1 [M + H^+], found: m/z = 232.0 [M + H^+], purity 100.0%.



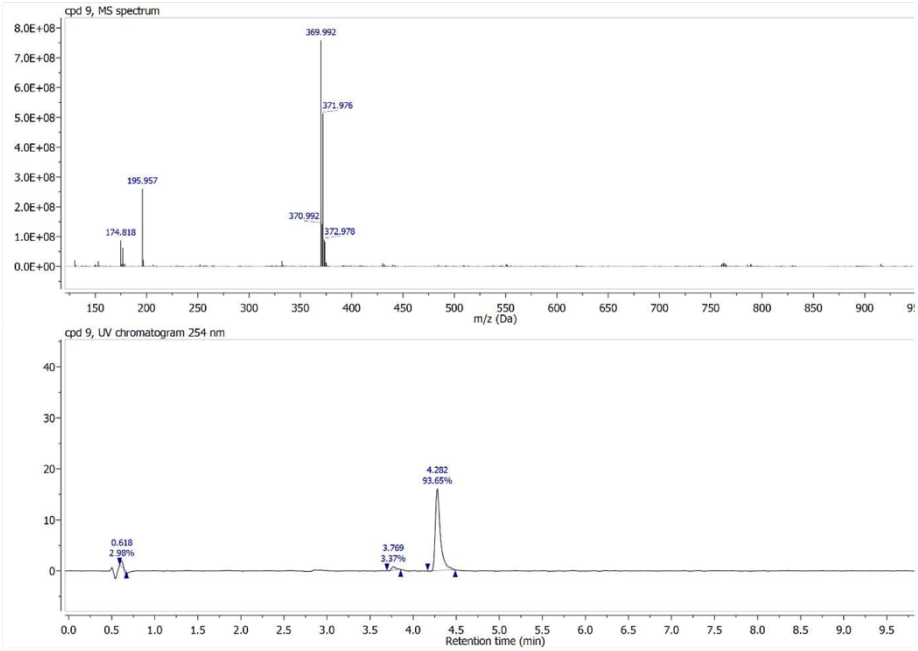
^{1}H NMR (400 MHz, Deuterium Oxide, TMS) δ_H 8.46 (1H, s), 7.75 (1H, d, *J* = 9.1 Hz), 6.97 (1H, d, *J* = 9.0 Hz), 6.74 (1H, s), 3.89 (3H, s), 2.60 (3H, s).



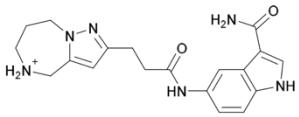
^{13}C NMR (101 MHz, Deuterium Oxide, TMS) δ_C 171.01, 170.19, 164.23, 150.74, 127.21, 117.87, 115.64, 104.64, 55.69, 20.36.



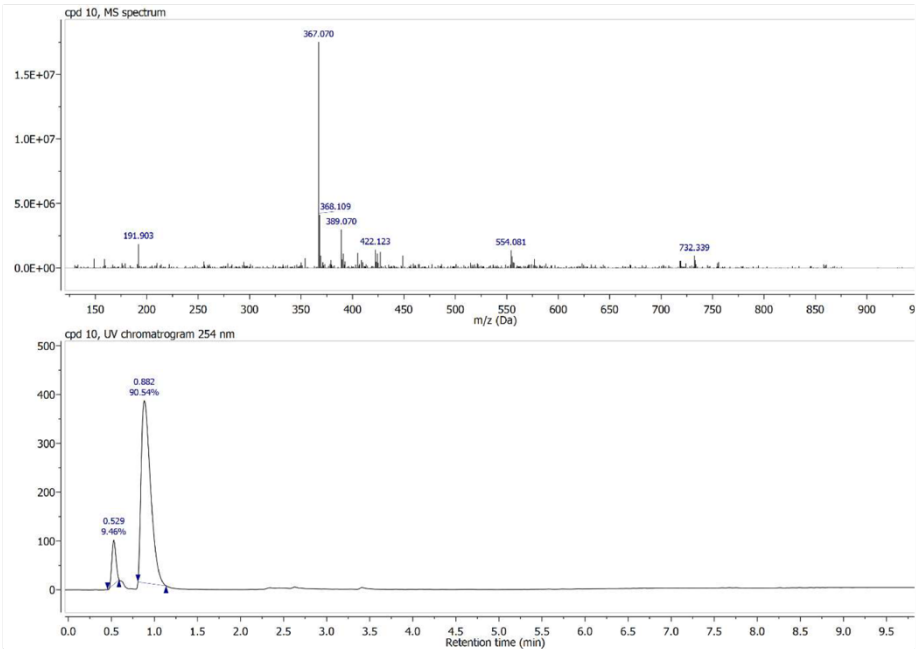
9



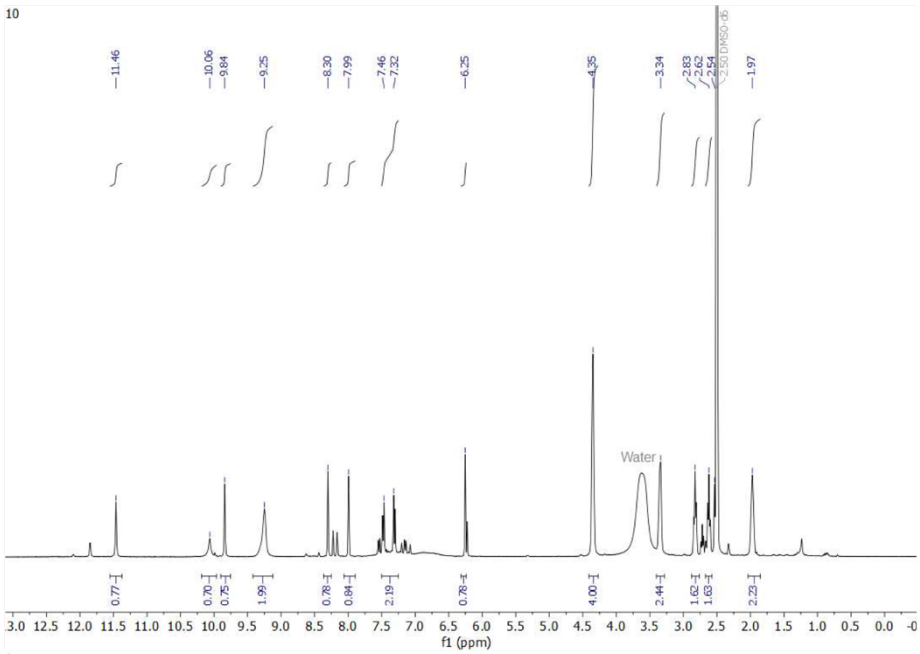
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 370.1 [M + H^+], found: m/z = 370.0 [M + H^+], purity 93.7%.



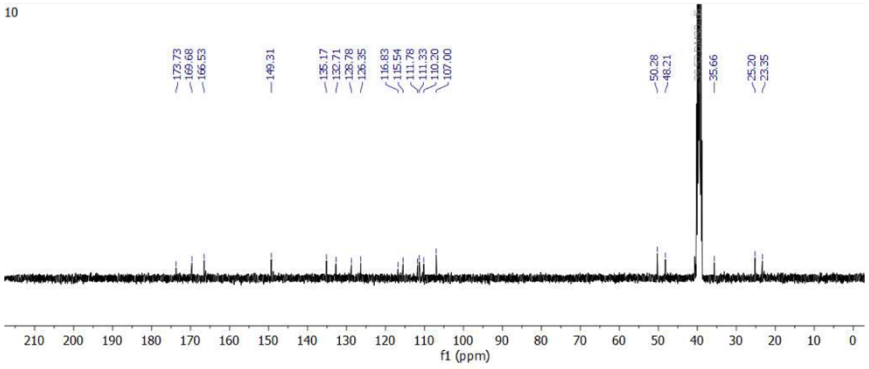
10



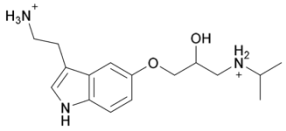
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 367.2 [M + H^+], found: m/z = 367.1 [M + H^+], purity 90.5%.



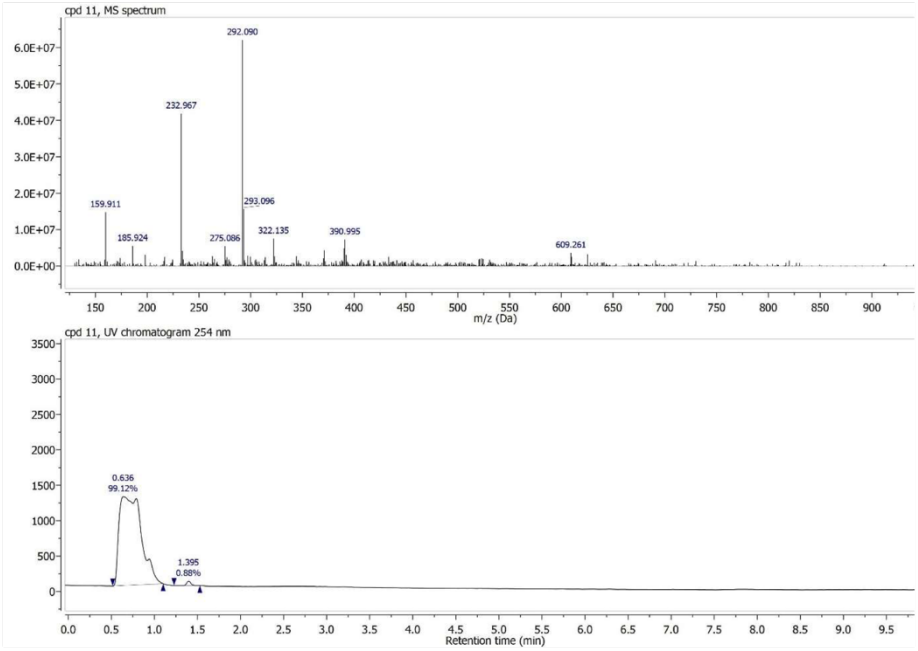
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 11.46 (1H, s), 10.06 (1H, s), 9.84 (1H, s), 9.25 (2H, s), 8.30 (1H, s), 8.00 (1H, d, J = 2.5 Hz), 7.48 – 7.30 (2H, m), 7.31 (1H, d, J = 8.9 Hz), 6.25 (1H, s), 4.44 – 4.23 (4H, m), 3.43 – 3.25 (2H, m), 2.83 (2H, t, J = 7.7 Hz), 2.62 (2H, t, J = 7.8 Hz), 2.08 – 1.86 (2H, m).



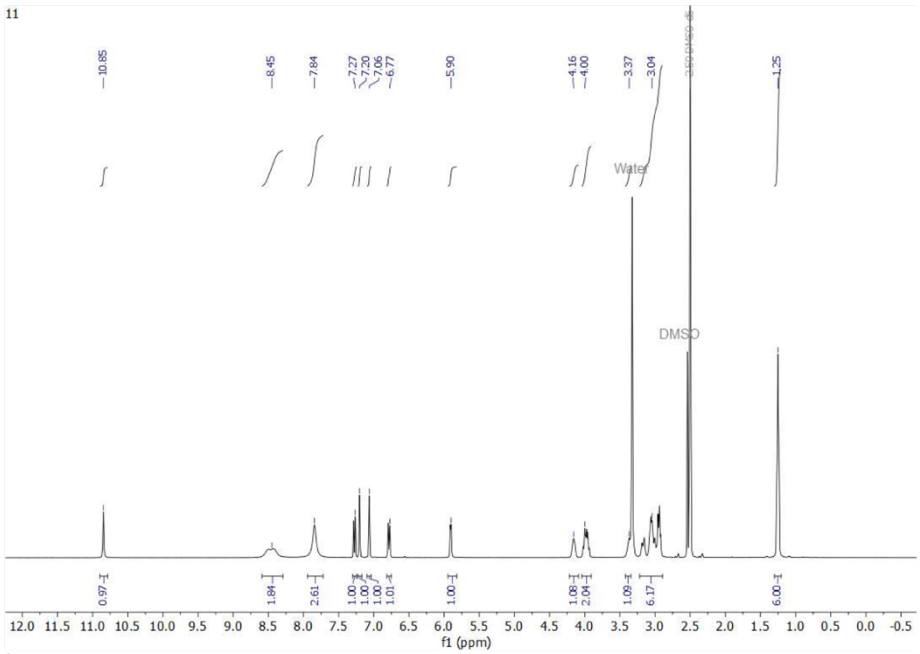
^{13}C NMR (101 MHz, DMSO-d_6, TMS) δ_C 173.73, 169.68, 166.53, 149.31, 135.17, 132.71, 128.78, 126.35, 116.83, 115.54, 111.78, 111.33, 110.20, 107.00, 50.28, 48.21, 35.66, 25.20, 23.35.



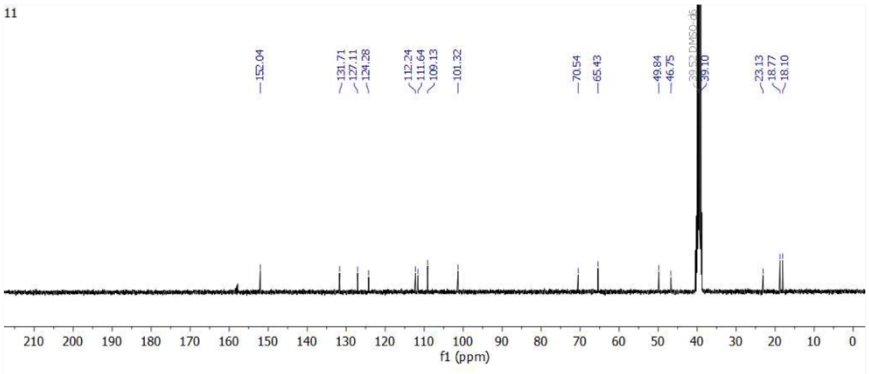
11



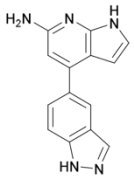
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 292.2 [M + H^+], found: m/z = 292.1 [M + H^+], purity 99.1%.



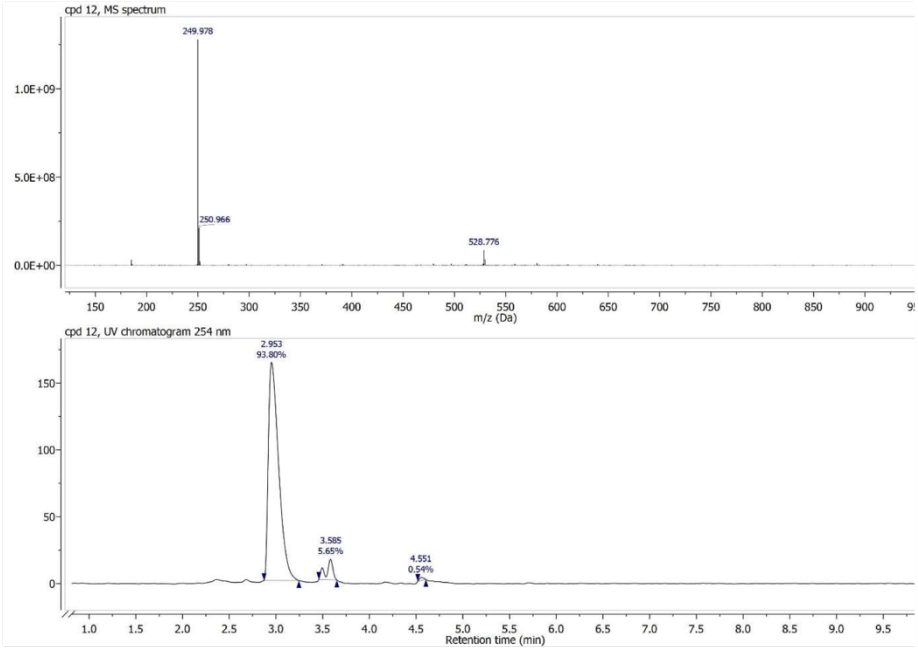
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 10.85 (1H, s), 8.66 – 8.28 (2H, m), 7.84 (3H, s), 7.28 (1H, dd, *J* = 8.8, 1.9 Hz), 7.20 (1H, d, *J* = 2.8 Hz), 7.06 (1H, d, *J* = 2.6 Hz), 6.79 (1H, dd, *J* = 8.7, 2.2 Hz), 5.91 (1H, d, *J* = 5.0 Hz), 4.15 (1H, dd, *J* = 9.9, 5.4 Hz), 3.98 (2H, qd, *J* = 9.8, 4.9 Hz), 3.42 – 3.35 (1H, m), 3.23 – 2.86 (6H, m), 1.25 (6H, t, *J* = 5.9 Hz).



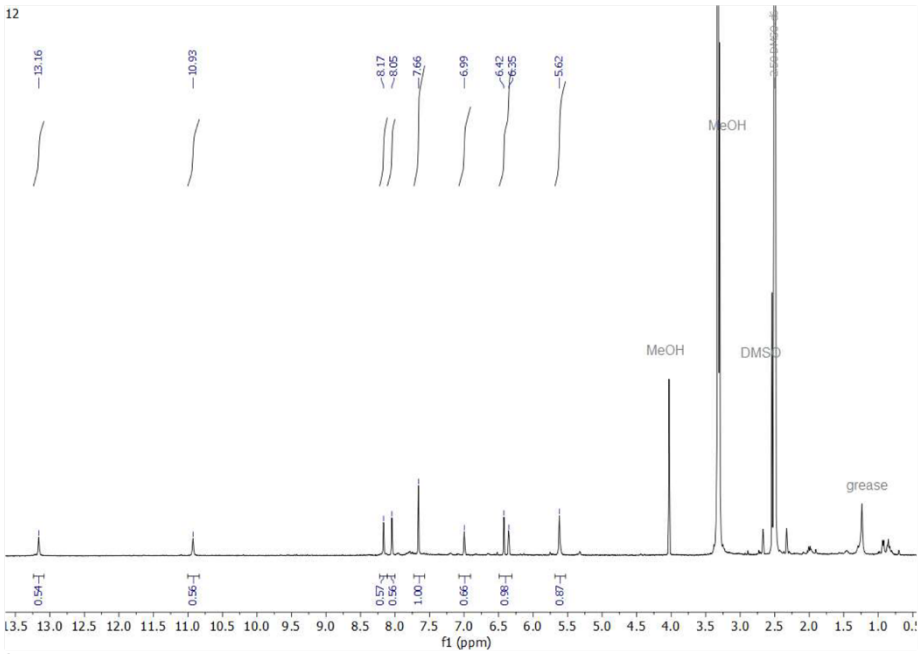
^{13}C NMR (101 MHz, DMSO-d_6, TMS) δ_C 152.04, 131.71, 127.11, 124.28, 112.24, 111.64, 109.13, 101.32, 70.54, 65.43, 49.84, 46.75, 39.10, 23.13, 18.77, 18.10.



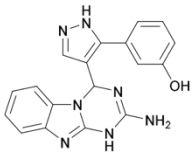
12



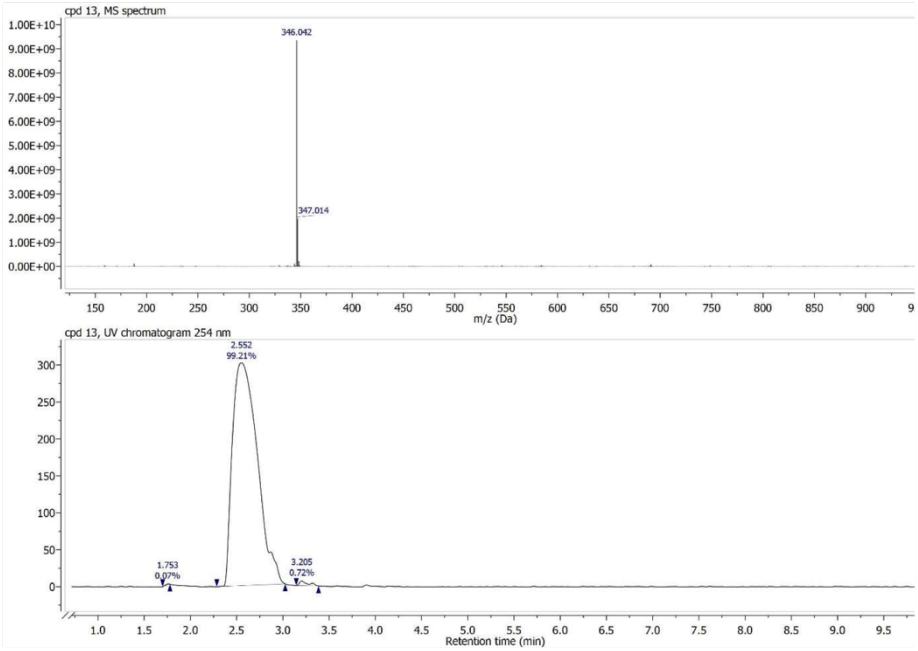
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 250.1[M + H^+], found: m/z = 250.0[M + H^+], purity 93.8%.



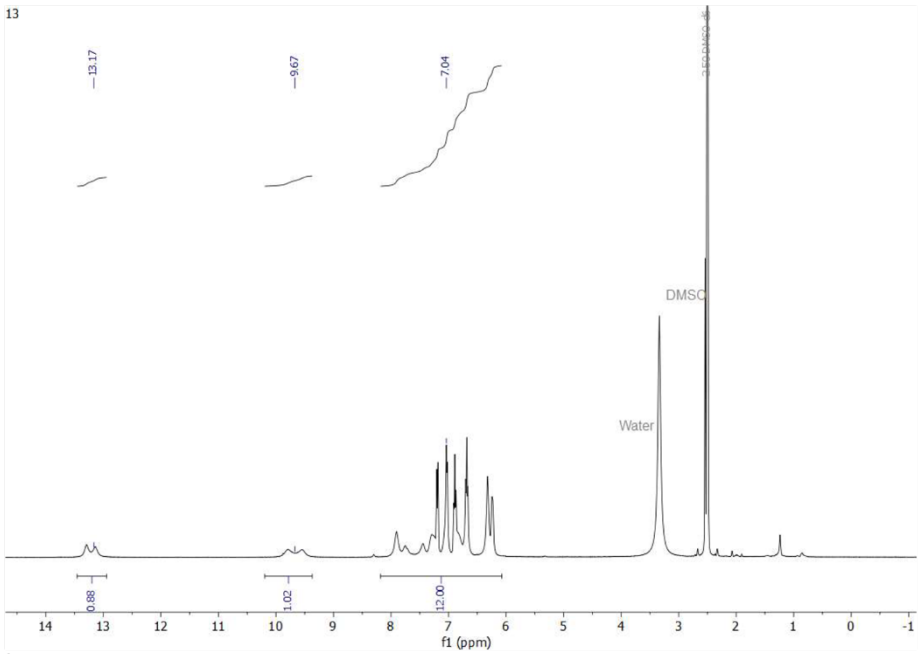
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 13.16 (1H, s), 10.93 (1H, s), 8.17 (1H, s), 8.05 (1H, s), 7.66 (1H, s), 7.13 – 6.95 (1H, m), 6.44 – 6.33 (1H, m), 5.62 (1H, s).



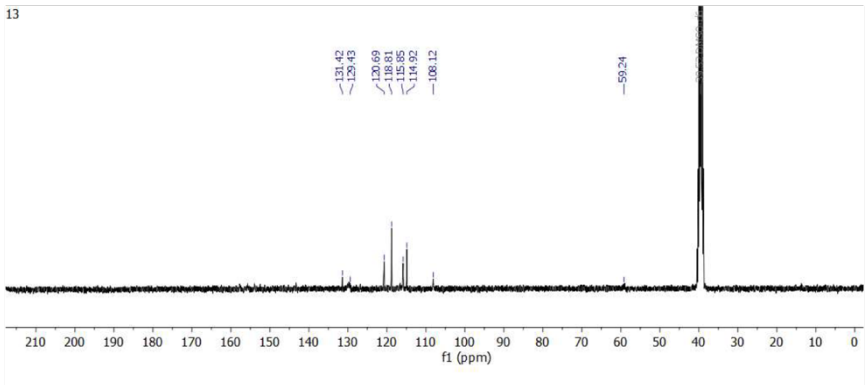
13



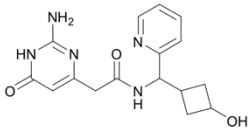
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 346.1 [M + H^+], found: m/z = 346.0[M + H^+], purity 99.2%.



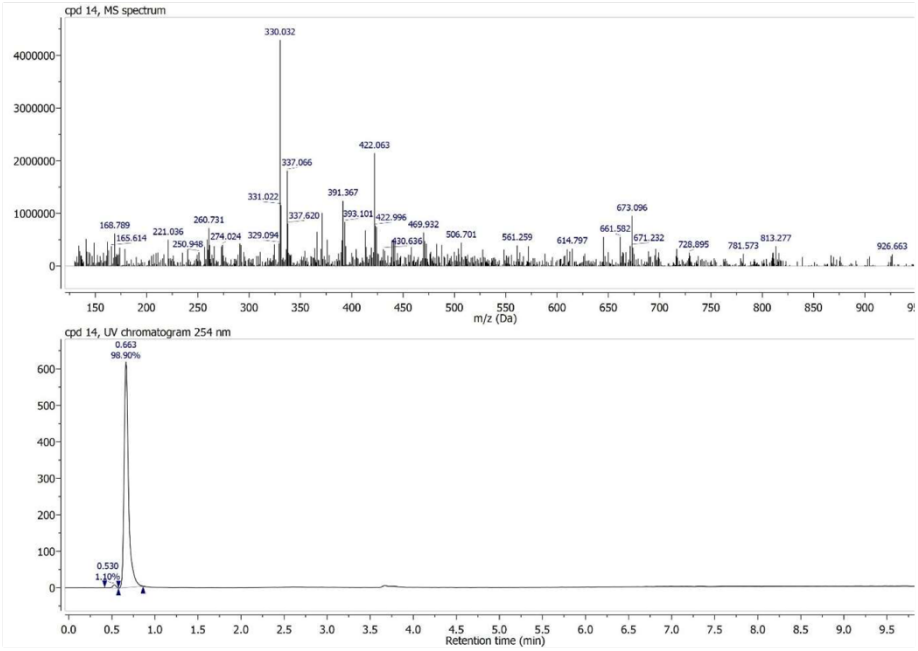
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 13.61 – 12.91 (1H, m), 10.08 – 9.28 (1H, m), 8.25 – 5.99 (12H, m).



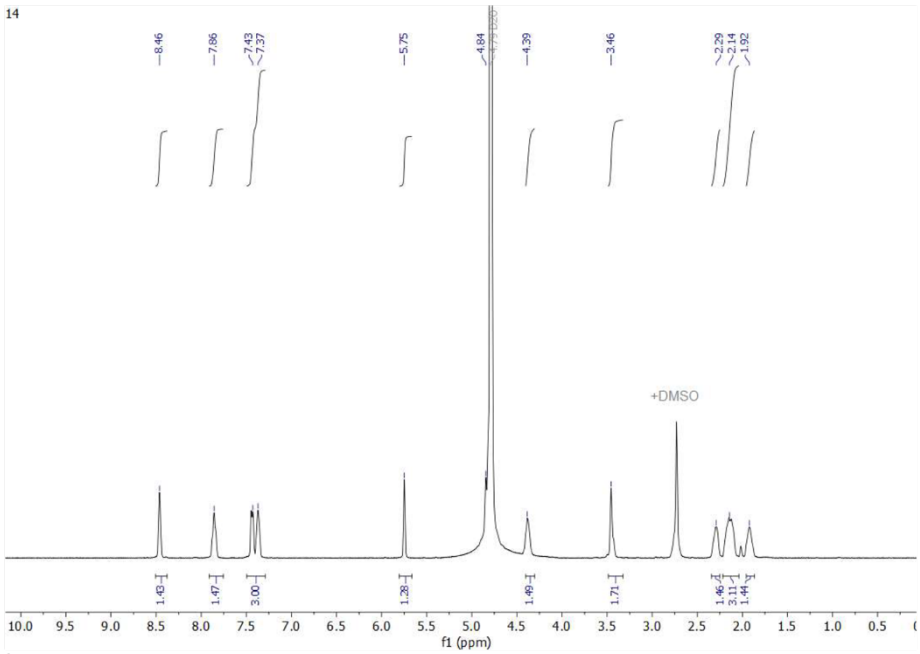
^{13}C NMR (101 MHz, DMSO-d_6, TMS) δ_C 131.42, 129.43, 120.69, 118.81, 115.85, 114.92, 108.12, 59.24.



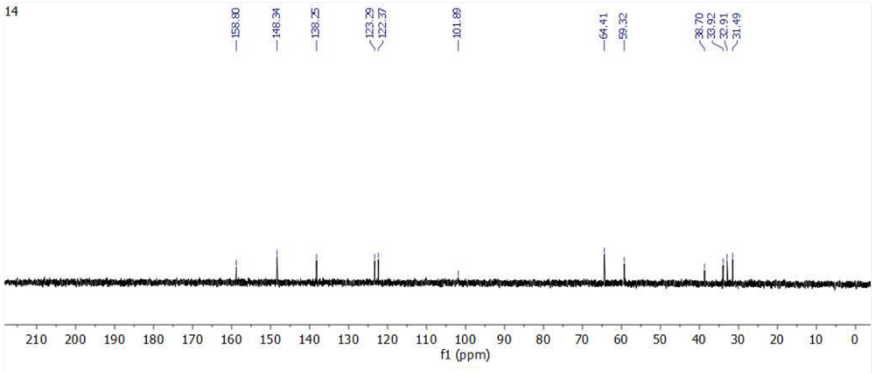
14



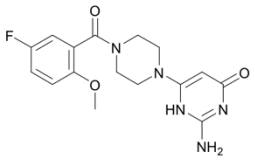
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 330.2 [M + H^+], found: m/z = 330.0 [M + H^+], purity 98.9%.



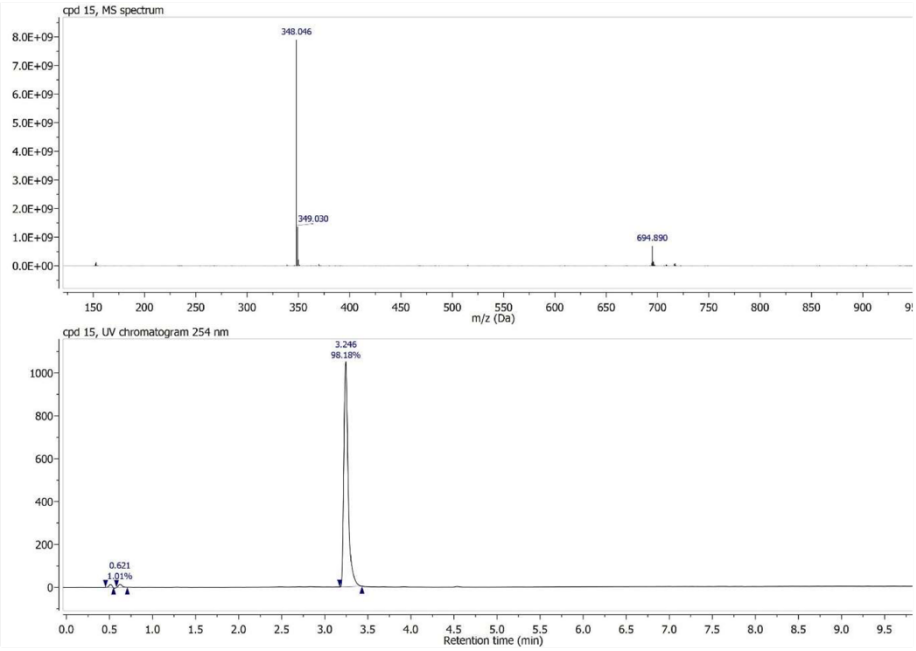
^{1}H NMR (400 MHz, Deuterium Oxide, TMS) δ_H 8.54 – 8.36 (1H, m), 7.95 – 7.78 (1H, m), 7.54 – 7.30 (3H, m), 5.75 (1H, s), 4.92 – 4.82 (m, 6H), 4.49 – 4.27 (1H, m), 3.52 – 3.35 (2H, m), 2.41 – 2.22 (1H, m), 2.22 – 2.03 (3H, m), 1.97 – 1.81 (1H, m).



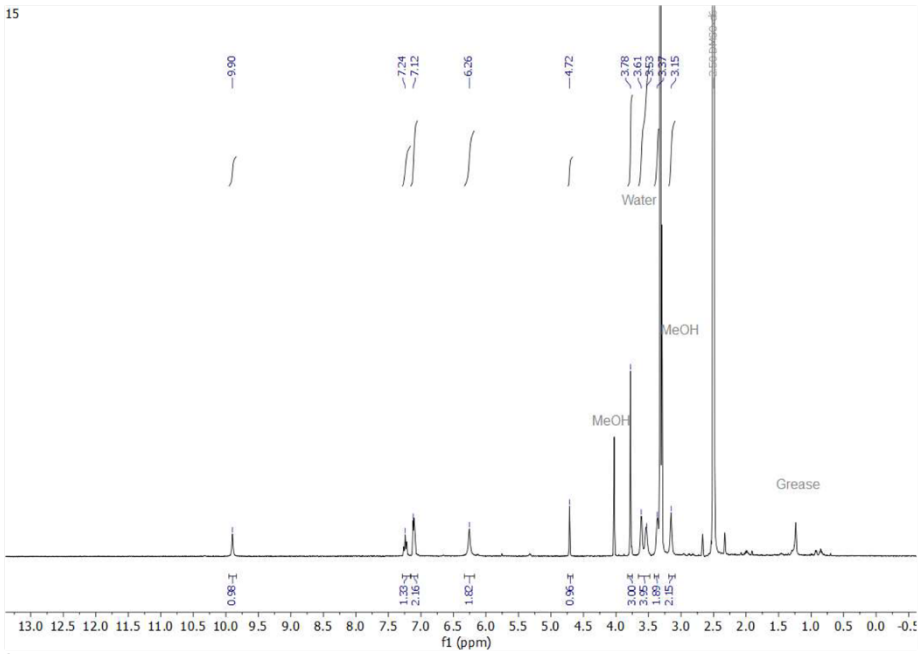
^{13}C NMR (101 MHz, Deuterium Oxide, TMS) δ_C 158.80, 148.34, 138.25, 123.29, 122.37, 101.89, 64.41, 59.32, 38.70, 33.92, 32.91, 31.49.



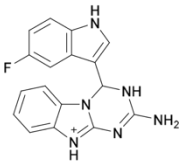
15



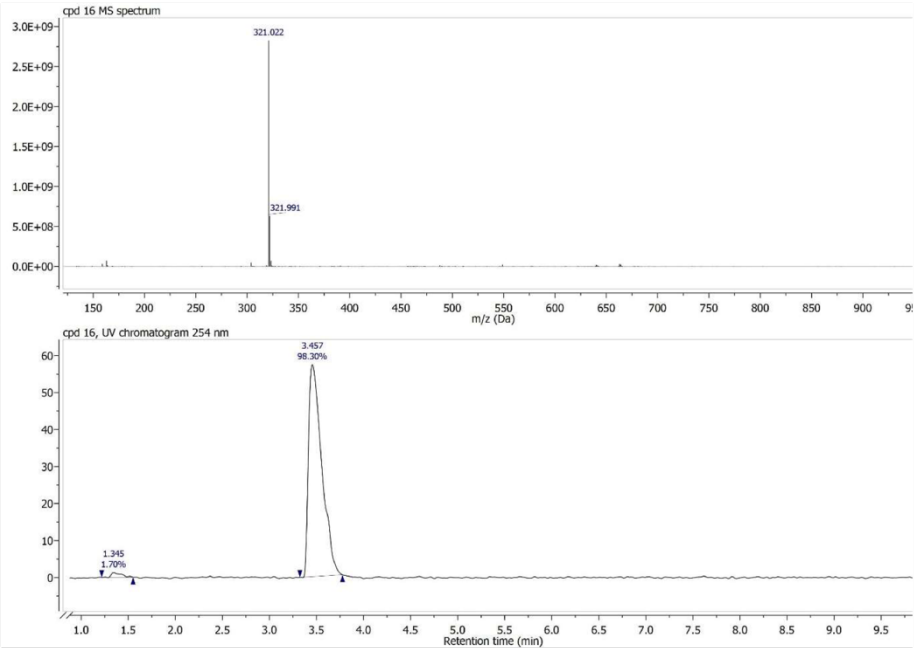
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 348.1 [M + H^+], found: m/z = 348.0 [M + H^+], purity 98.2%.



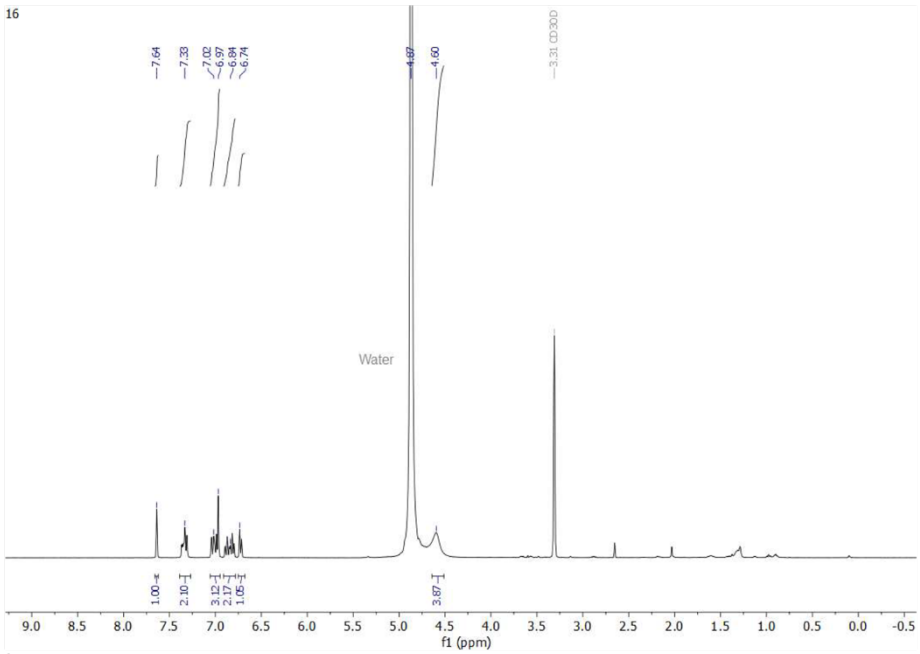
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 9.90 (1H, s), 7.34 – 7.16 (1H, m), 7.17 – 7.02 (2H, m), 6.26 (2H, s), 4.72 (1H, s), 3.82 – 3.72 (3H, m), 3.63 – 3.49 (4H, m), 3.38 – 3.33 (2H, m), 3.22 – 3.08 (2H, m).



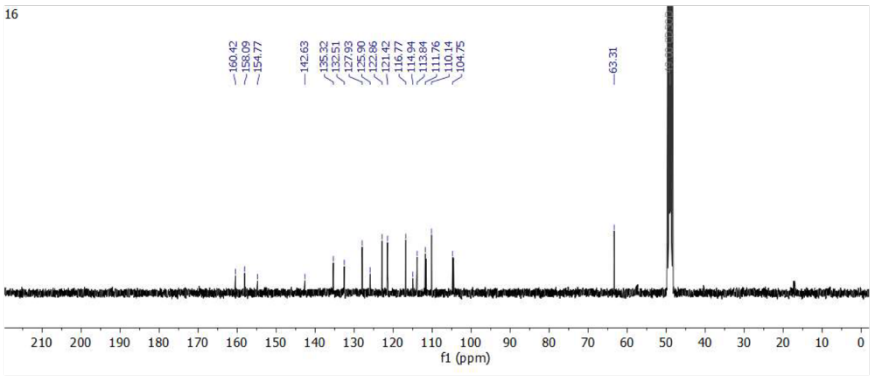
16



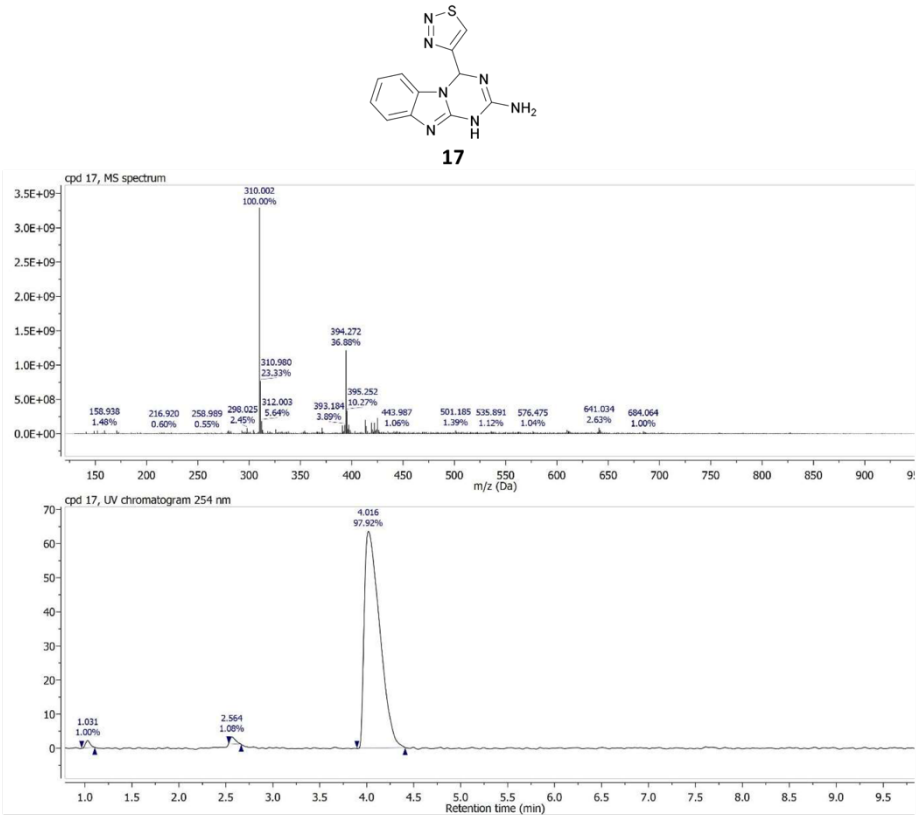
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 321.1 [M + H^+], found: m/z = 321.0 [M + H^+], purity 98.3%.



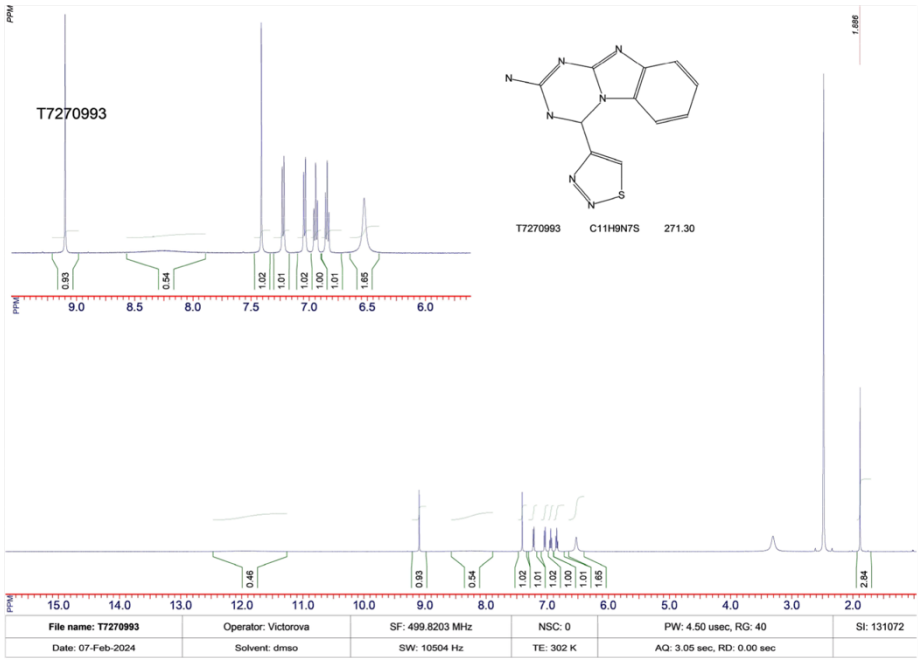
^{1}H NMR (400 MHz, Methanol-d_4, TMS) δ_H 7.64 (1H, s), 7.43 – 7.25 (2H, m), 7.11 – 6.93 (3H, m), 6.84 (2H, dt, J = 20.3, 8.3 Hz), 6.73 (1H, d, J = 8.0 Hz), 4.60 (4H, s).



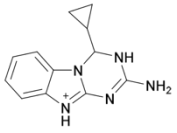
^{13}C NMR (101 MHz, Methanol-d_4, TMS) δ_C 160.42, 158.09, 154.77, 142.63, 135.32, 132.51, 127.93, 125.90, 122.86, 121.42, 116.77, 114.94, 113.84, 111.76, 110.14, 104.75, 63.31.



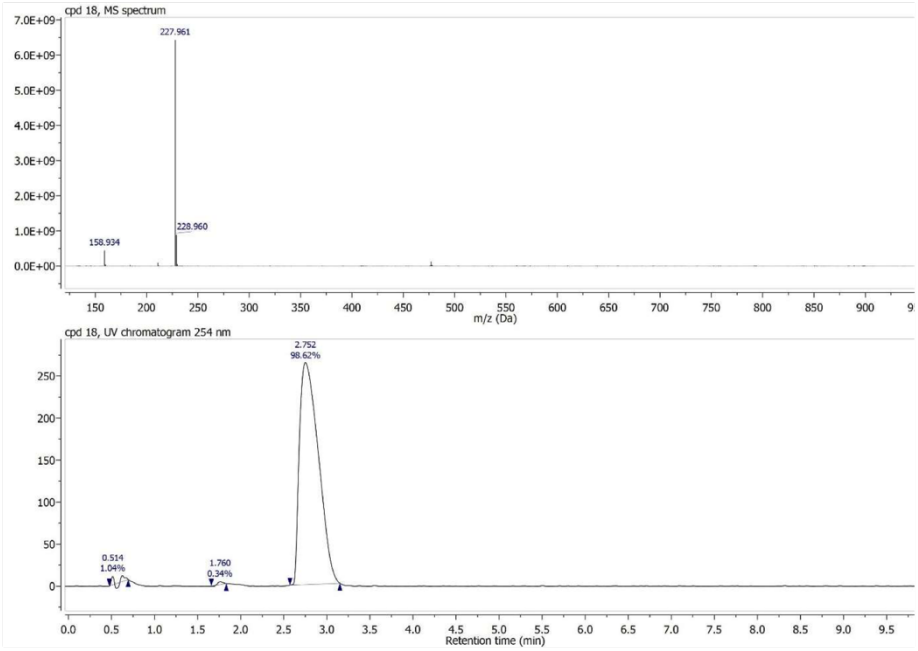
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 272.3 [M + H^+], found: m/z = 310.0 [M + K^+],, purity 97.9%.



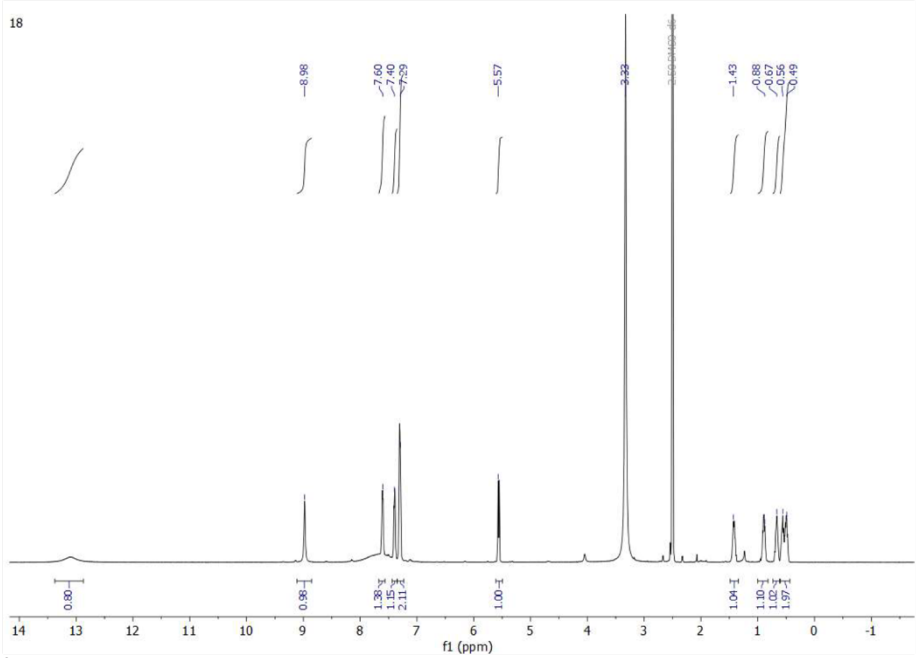
^{1}H NMR provided by Enamine Ltd.



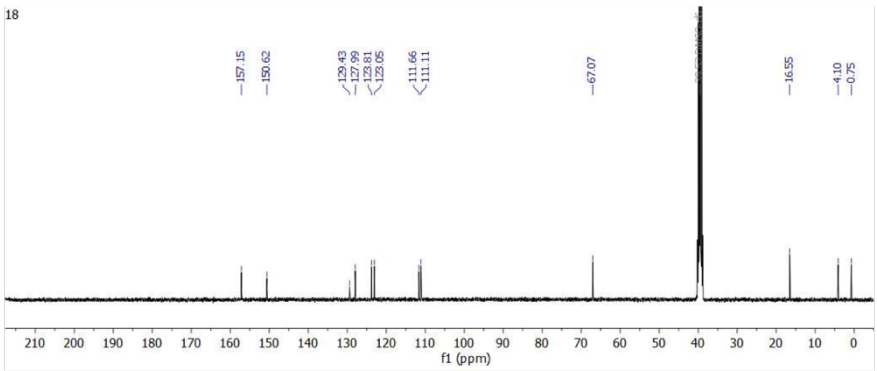
18



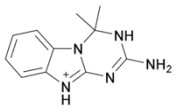
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 228.1 [M + H^+], found: m/z = 228.0 [M + H^+], purity 98.6%.



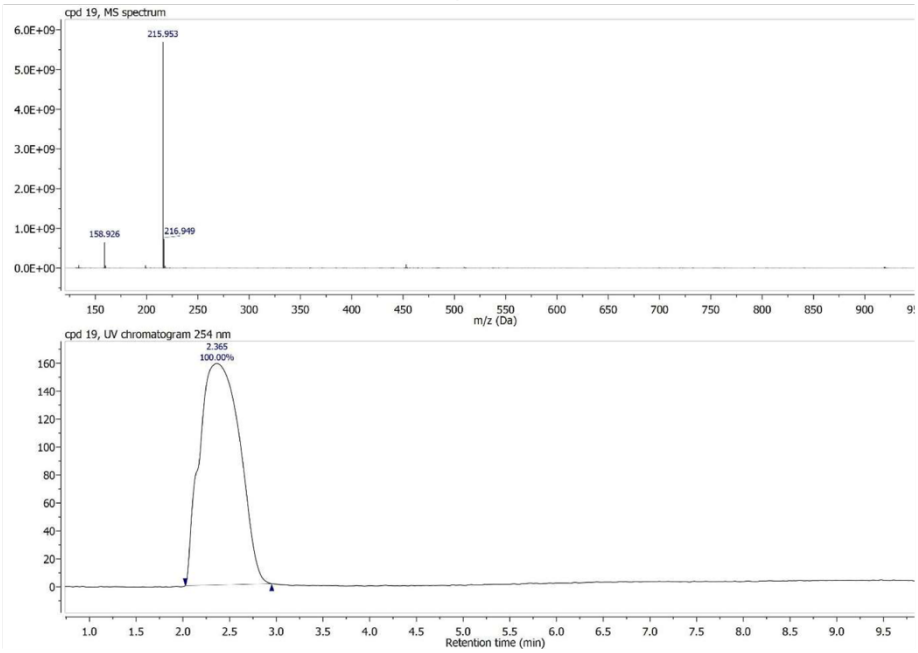
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 13.10 (1H, s), 8.98 (1H, s), 7.73 – 7.53 (1H, m), 7.47 – 7.33 (1H, m), 7.34 – 7.18 (2H, m), 5.56 (1H, d, J = 7.6 Hz), 1.54 – 1.34 (1H, m), 1.02 – 0.83 (1H, m), 0.72 – 0.61 (1H, m), 0.61 – 0.43 (2H, m).



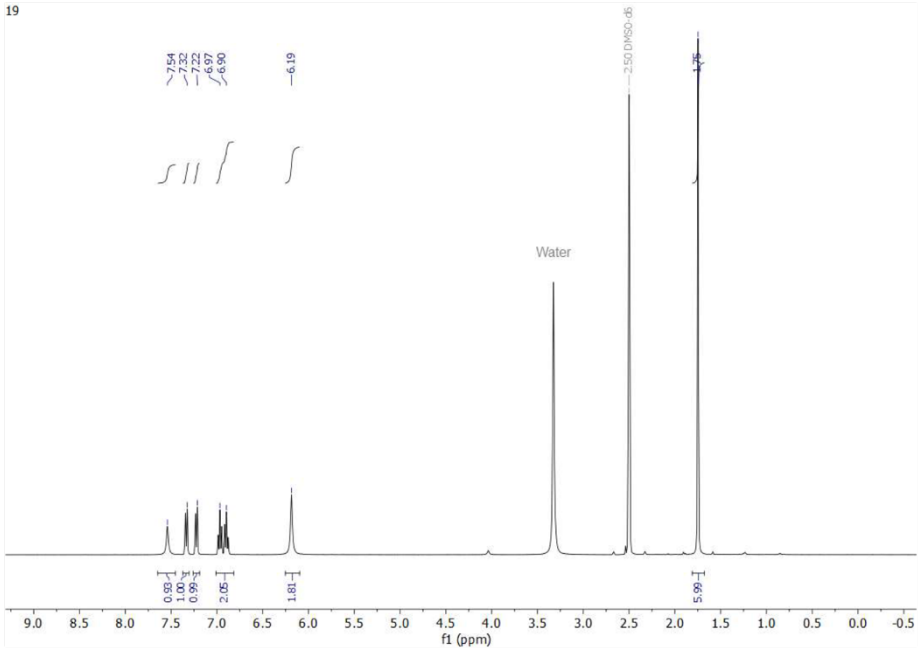
^{13}C NMR (101 MHz, DMSO-d_6, TMS) δ_C 157.15, 150.62, 129.43, 127.99, 123.81, 123.05, 111.66, 111.11, 67.07, 16.55, 4.10, 0.75.



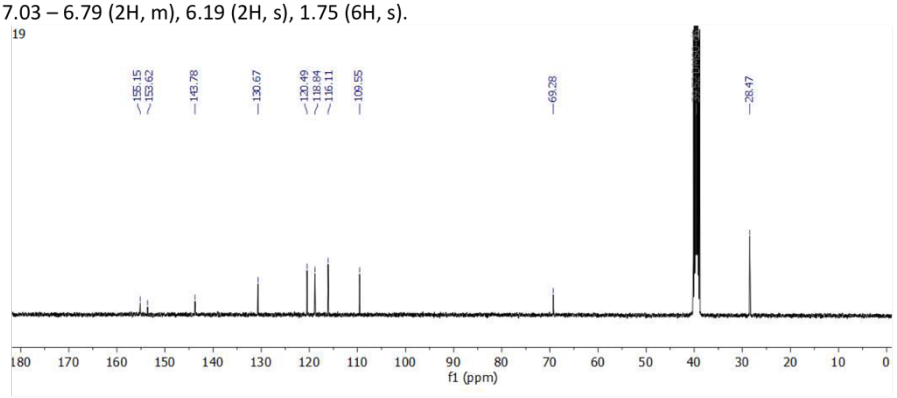
19



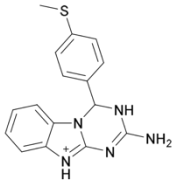
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 216.1 [M + H^+], found: m/z = 216.0 [M + H^+], purity 100.0%.



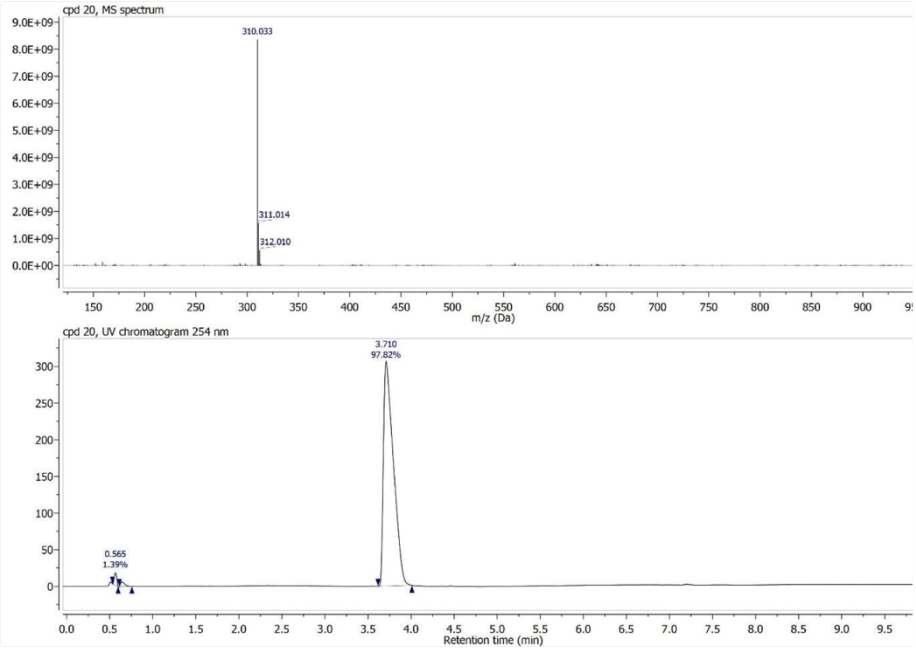
^{1}H NMR (400 MHz, DMSO-d_6, TMS) δ_H 7.54 (1H, s), 7.33 (1H, d, J = 7.8 Hz), 7.22 (1H, d, J = 7.7 Hz), 7.03 – 6.79 (2H, m), 6.19 (2H, s), 1.75 (6H, s).



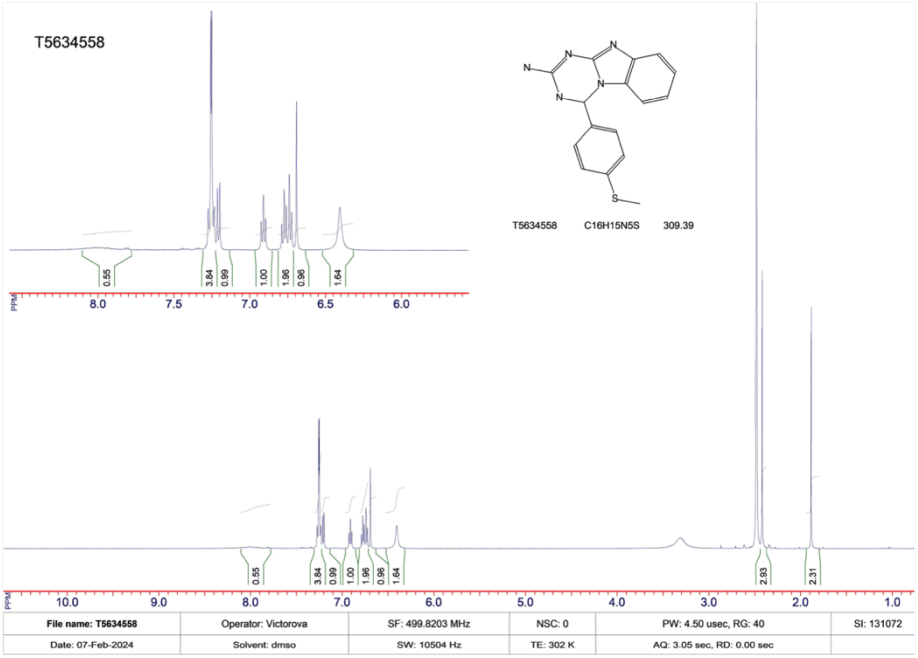
^{13}C NMR (101 MHz, DMSO-d_6, TMS) δ_C 155.15, 153.62, 143.78, 130.67, 120.49, 118.84, 116.11, 109.55, 69.28, 28.47.



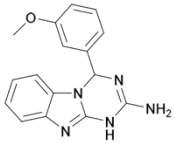
20



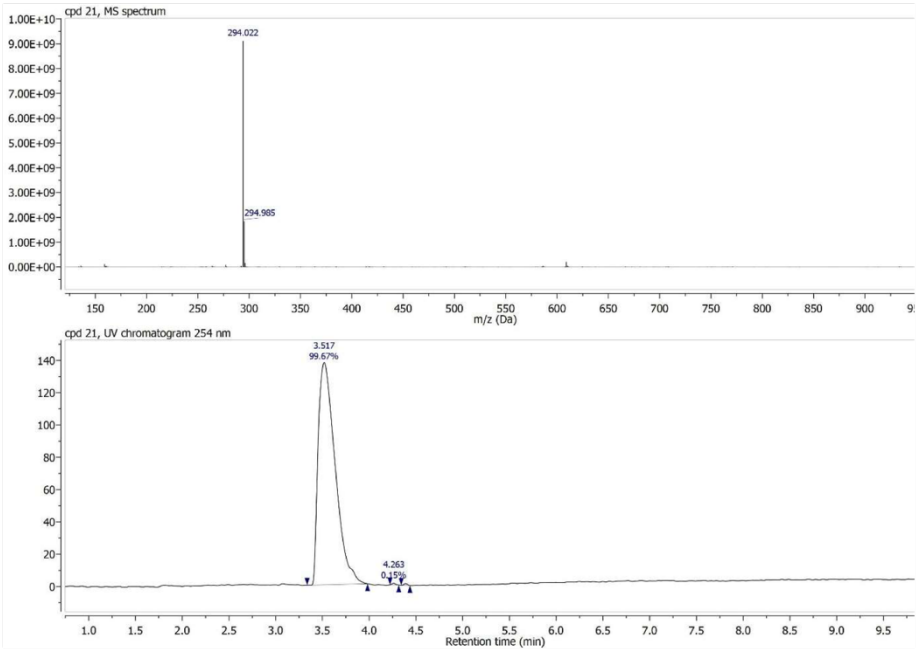
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 310.4 [M + H^+], found: m/z = 310.0 [M + H^+], purity 97.8%.



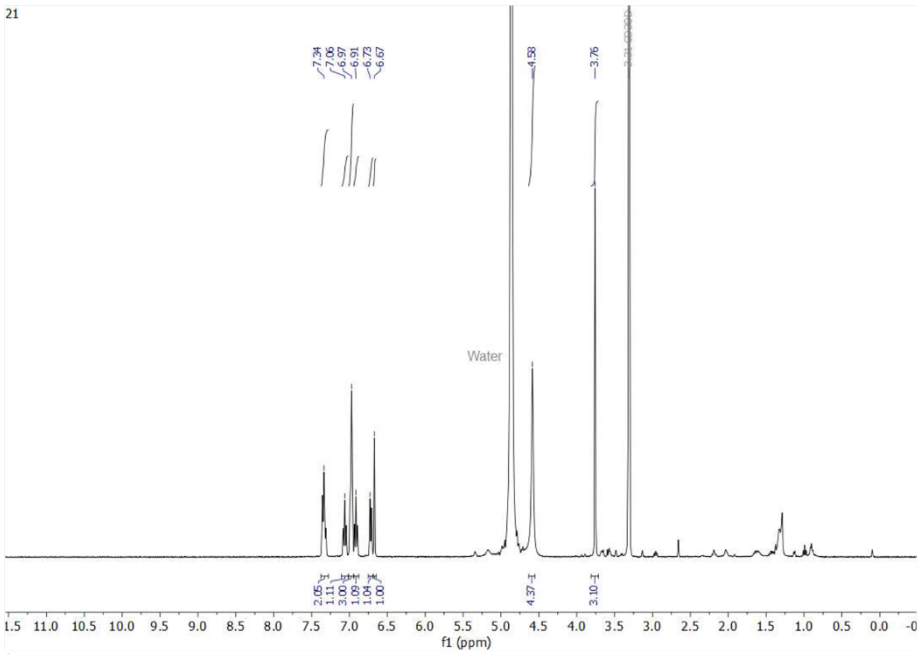
^{1}H NMR provided by Enamine Ltd.



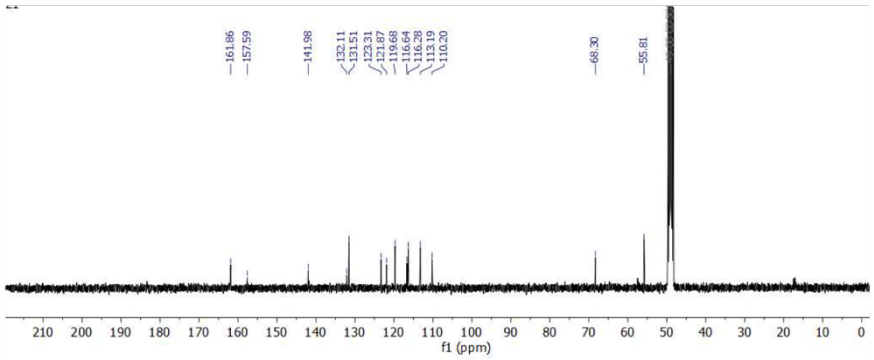
21



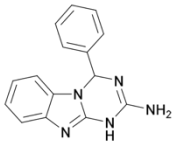
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 294.1 [M + H^+], found: m/z = 294.0 [M + H^+], purity 99.7%.



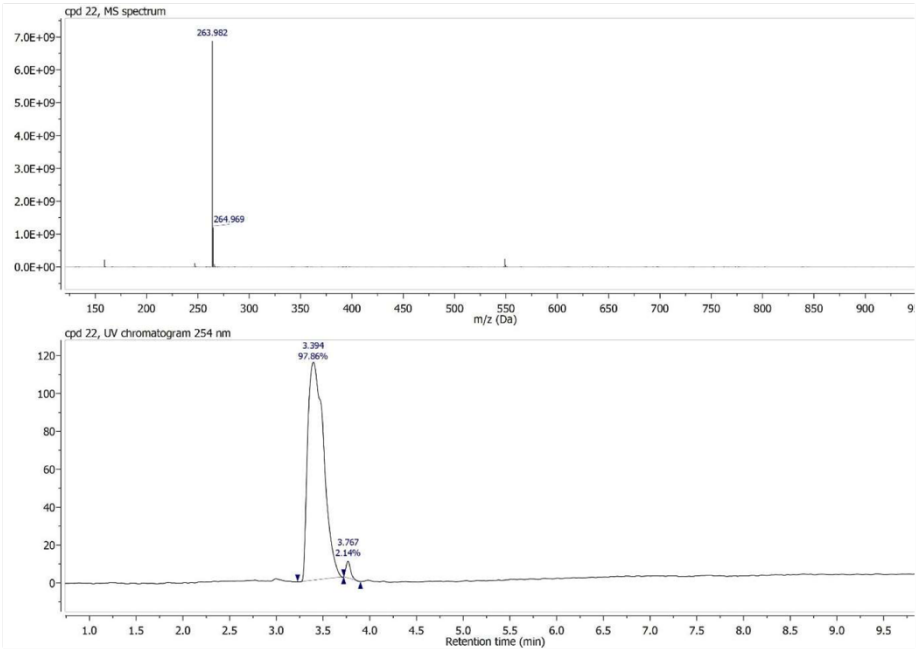
^{1}H NMR (400 MHz, Methanol-d_4, TMS) δ_H 7.38 – 7.29 (2H, m), 7.06 (1H, t, J = 7.7 Hz), 7.01 – 6.94 (3H, m), 6.91 (1H, t, J = 7.8 Hz), 6.72 (1H, d, J = 8.0 Hz), 6.67 (1H, d, J = 1.9 Hz), 4.58 (4H, s), 3.76 (3H, s).



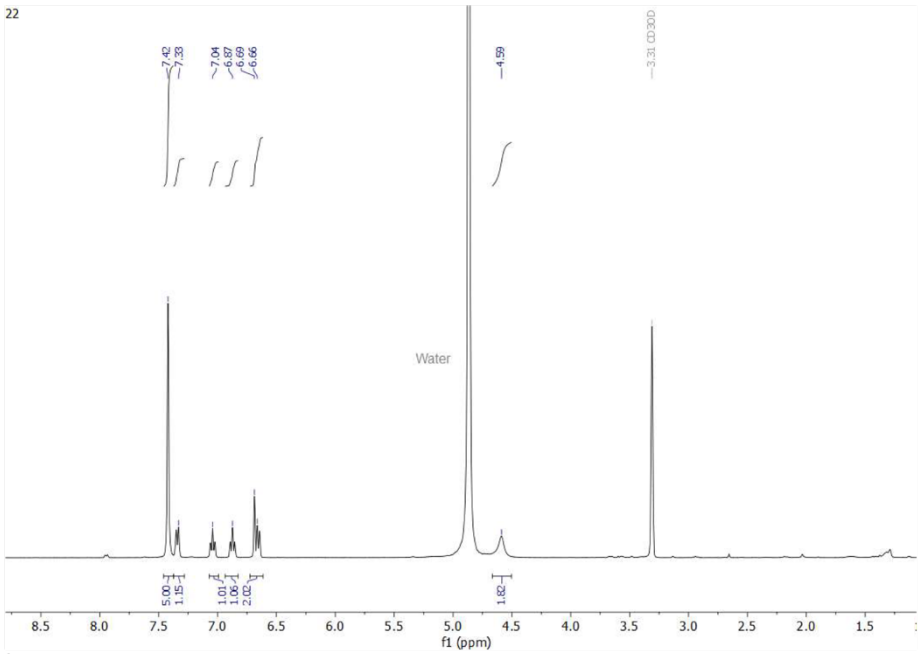
^{13}C NMR (101 MHz, Methanol-d_4, TMS) δ_C 161.86, 157.59, 141.98, 132.11, 131.51, 123.31, 121.87, 119.68, 116.64, 116.28, 113.19, 110.20, 68.30, 55.81.



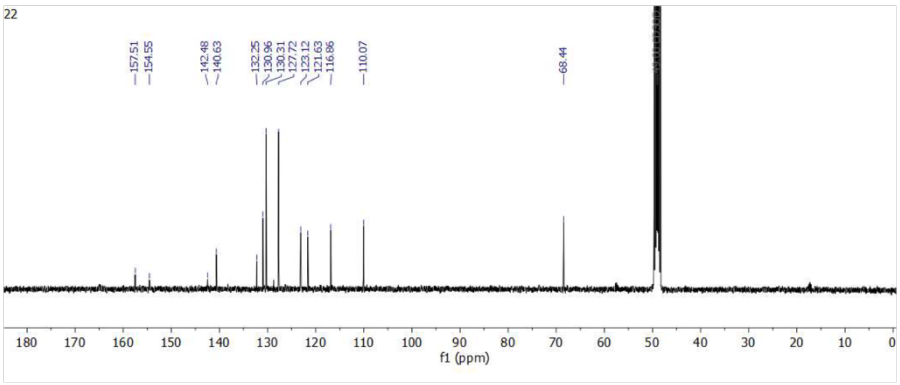
22



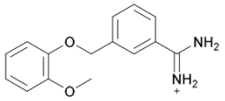
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 264.1 [M + H^+], found: m/z = 264.0 [M + H^+], purity 97.9%.



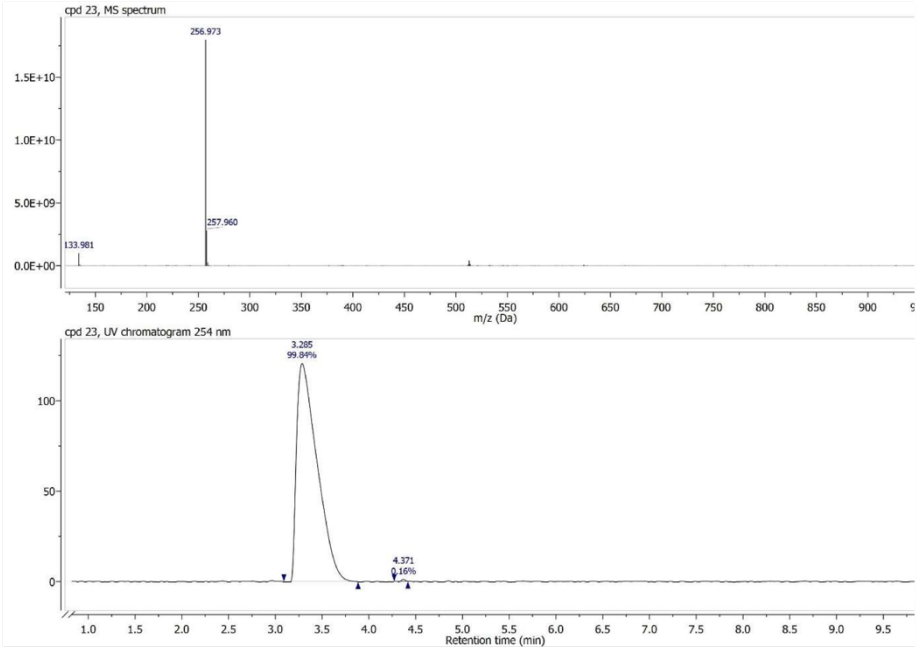
^{1}H NMR (400 MHz, Methanol-d_4, TMS) δ_H 7.48 – 7.37 (5H, m), 7.34 (1H, d, J = 8.0 Hz), 7.13 – 6.96 (1H, m), 6.94 – 6.78 (1H, m), 6.74 – 6.58 (2H, m), 4.59 (2H, s).



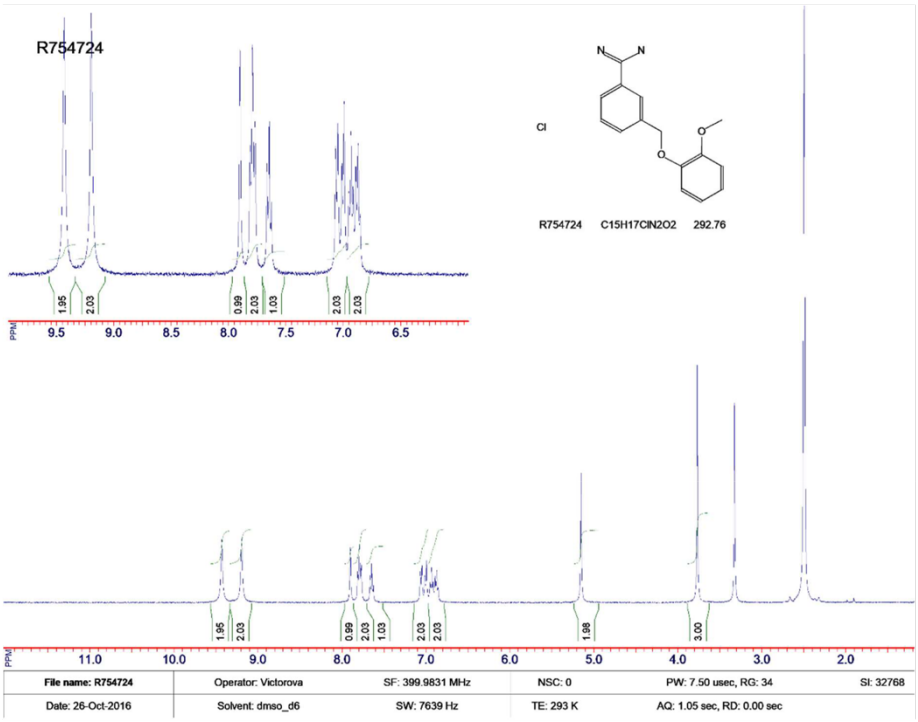
^{13}C NMR (101 MHz, Methanol-d_4, TMS) δ_C 157.51, 154.55, 142.48, 140.63, 132.25, 130.96, 130.31, 127.72, 123.12, 121.63, 116.86, 110.07, 68.44.



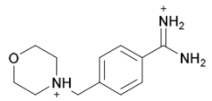
23



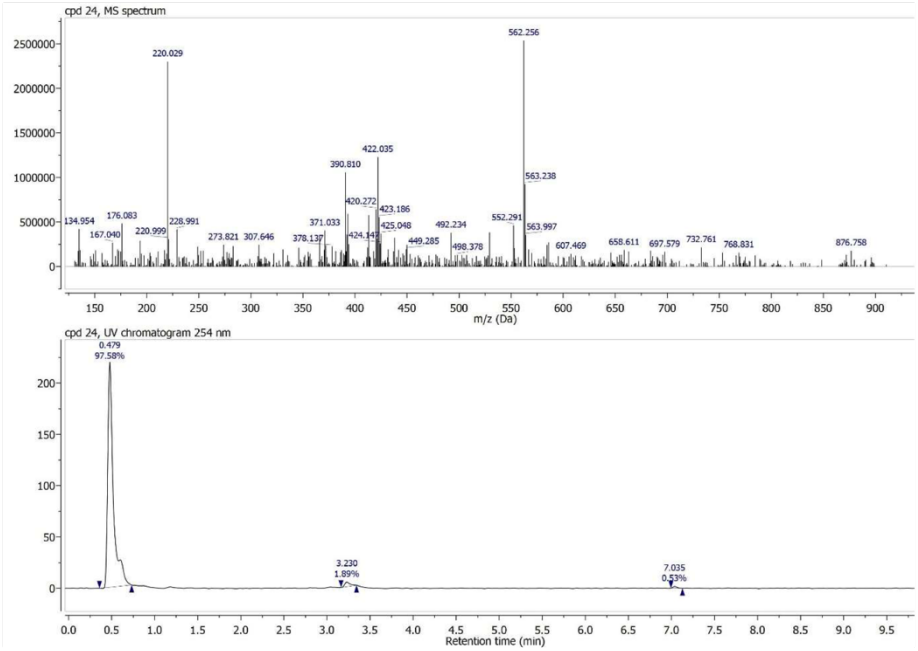
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 257.1 [M + H^+], found: m/z = 257.0 [M + H^+], purity 99.8%.



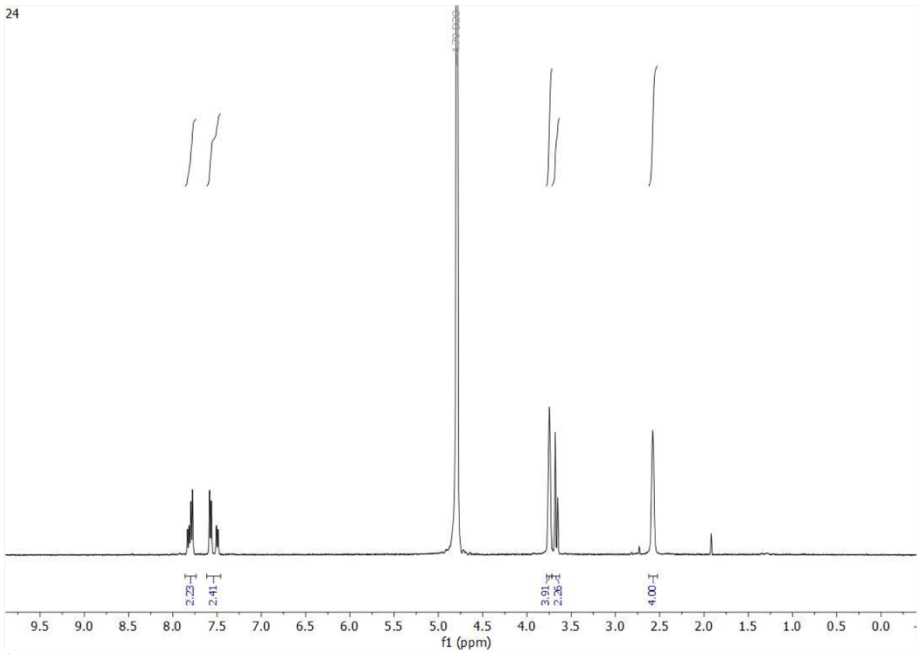
^{1}H NMR provided by Enamine Ltd.



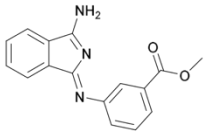
24



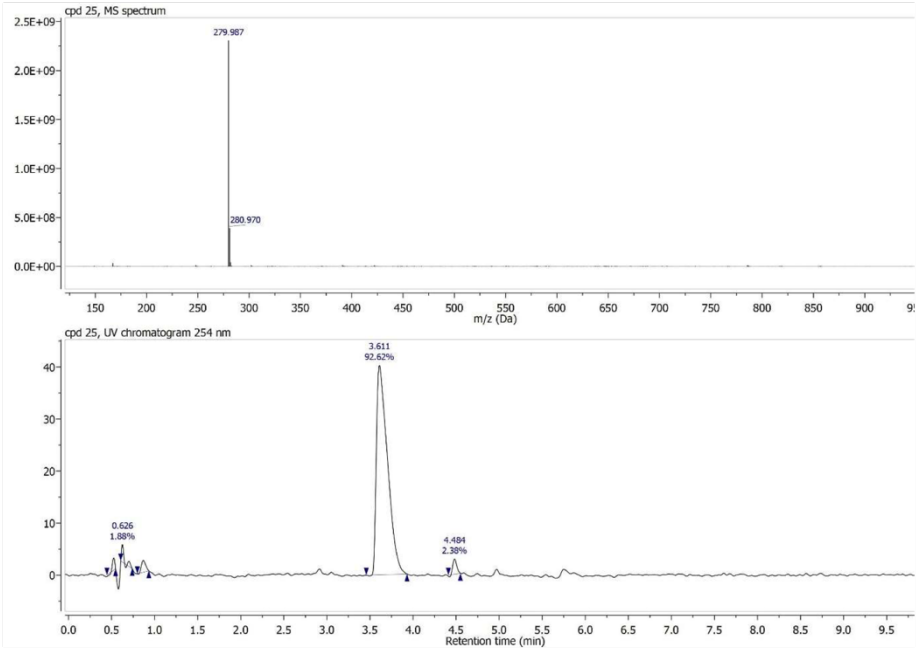
LC-ESI-MS (measured with Poroshell column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 220.2 [M + H^+], found: m/z = 220.0 [M + H^+], purity 97.6%.



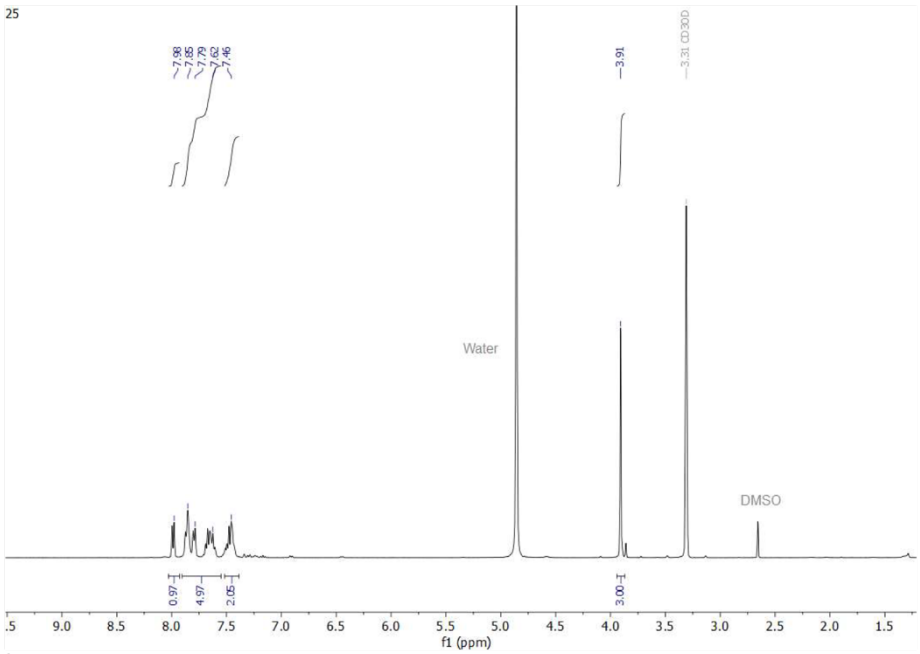
^{1}H NMR (400 MHz, Deuterium Oxide, TMS) δ_H 7.86 – 7.75 (2H, m), 7.60 – 7.46 (2H, m), 3.79 – 3.72 (4H, m), 3.70 – 3.63 (2H, m), 2.65 – 2.51 (4H, m).



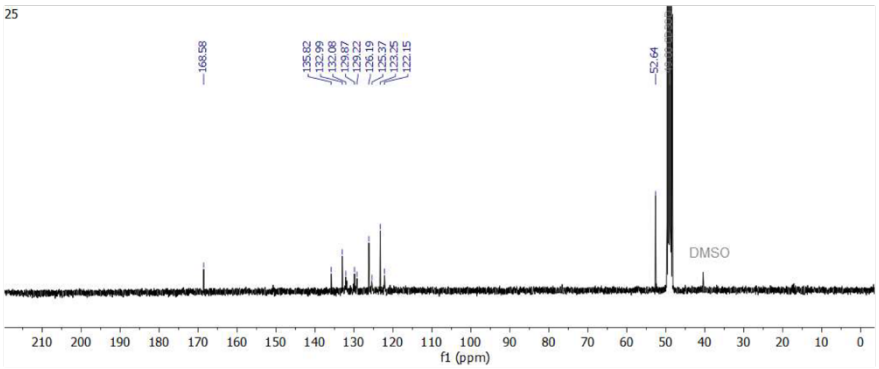
25



LC-ESI-MS (measured with Zorbax column, gradient 10–90% ACN in water + 0.1% formic acid over 10 min), top: mass spectrum, bottom: LC chromatogram at 254 nm. Calculated: m/z = 280.1 [M + H^+], found: m/z = 280.0 [M + H^+], purity 92.6%.



^{1}H NMR (400 MHz, Methanol-d_4, TMS) δ_H 7.99 (1H, d, J = 7.2 Hz), 7.91 – 7.55 (5H, m), 7.56 – 7.37 (2H, m), 3.91 (3H, s).



^{13}C NMR (101 MHz, Methanol-d_4, TMS) δ_C 168.58, 135.82, 132.99, 132.08, 129.87, 129.22, 126.19, 125.37, 123.25, 122.15, 52.64.

5. References

1 D. Schmit, U. Milewicz, M. A. Boerneke, S. Burley, K. Walsworth, J. Um, D. Hecht, T. Hermann and B. M. Bergdahl, *Aust J Chem*, 2020, **73**, 212–221.

2 M. A. Parker, E. Satkiewicz, T. Hermann and B. M. Bergdahl, *Molecules*, 2011, **16**, 281–290.

”

6. Electrostatic Anchoring – Exploring the Impact of Charge to RNA–Small Molecule Binding Affinity, Kinetics, Thermodynamics and Selectivity

6.1. Context, Project Summary, and Own Contribution

In the context of RNA's growing potential as a promising drug target for a range of diseases, the principle of electrostatic anchoring has emerged as a rational approach, given the charged nature of RNA molecules.[100] The concept is not novel in the context of the antibiotic aminoglycoside group, but further development is required with a strong focus on pharmacological properties, such as absorption and toxicity.[3,5,10] The incorporation of positively charged groups has been demonstrated to significantly enhance the affinity via a fast association with RNA molecules.[93,99] However, this approach can also lead to a loss in selectivity, which may consequently result in toxic side effects. To address these challenges, RNA-binding small molecules can be derivatized with charged moieties to facilitate anchoring in the binding site, enhance the association rate and consequently their affinity.

We investigated the impact of charge on RNA–small molecule binding kinetics and affinity on two different RNA targets: The *Tte* preQ$_1$ riboswitch and the *Ba* SAM-VI riboswitch. Both riboswitches are capable of discriminating between charged (preQ$_1$, SAM) and uncharged (preQ$_0$, SAH) natural ligands, thereby serving as suitable model systems for the experiments described herein. We initially determined the kinetic and thermodynamic binding profiles of the natural ligands using ITC, SPR and MST assays. For both riboswitches, we observed enhanced association rates of the charged ligands, while the dissociation rates remained in the same order of magnitude. Consequently, the enhanced association is the primary factor contributing to the enhanced affinities of those charged ligands. Furthermore, we sought to determine if there exists a limit in the proportionality between k_{on} and net charge. To this end, we investigated the binding profiles of small molecules with increasing net charges (0, +1, +2, and +3).

The results of our experiments demonstrated a correlation between charge, association speed, and affinity. The results also demonstrated a discernible proclivity for selectivity to diminish with rising net charge, in conjunction with a constraint on the value of k_{on}. This suggests that the principle of electrostatic anchoring may be an effective strategy to improve the pharmacokinetic profiles of actual known RNA-binding small molecules when a positive charge is introduced. Therefore, this concept may not be universally applicable, as we observed non-specific and non-selective binding for small molecules with increased positive charge. In conclusion, the consistent data from three different biophysical assays (ITC, SPR, and MST), demonstrated that the development of small molecules for RNA targeting requires a delicate balance between affinity and selectivity. Our study contributed to the field of rational RNA-targeted drug design and highlighted the impact of charge on selectivity and affinity for those binding events.

Own contribution: Preparation of the SAM-VI riboswitch aptamer, MST assay establishment and affinity measurements, ITC assay establishment and measurements, MST and ITC data evaluation (affinity, thermodynamic and kinetic properties), writing of the original draft & editing of the manuscript.

Contributions of other authors: Synthesis and analytics of all compounds, SPR assay establishment and measurements, SPR data evaluation, interaction map analysis, writing of the original draft & editing of the manuscript.

Laura Almena-Rodriguez, L. and **Elisabeth Kallert** contributed equally to this manuscript.

The following manuscript is a draft version to which sections concerning SPR methodology and analysis still need to be added. The illustrations of the sensorgrams in the manuscript and supplementary information are also preliminary. Prof. Dr. Ralf Hermann and Kirsten Schaubruch are currently working on this part of the manuscript.

6.2. Draft Manuscript

Electrostatic Anchoring in RNA-Ligand Design – Dissecting the Effects of Positive Charges on Affinity, Selectivity, Binding Kinetics and Thermodynamics

Laura Almena Rodriguez[1,‡], Elisabeth Kallert[1,‡], Jan-Åke Husmann[1], Kirsten Schaubruch[2] , Marvin Schwickert[1], Sabrina N. Hoba[1], Ralf Heermann[2], Christian Kersten[1,3,*]

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

[2] Institute of Molecular Physiology, Microbiology and Biotechnology, Johannes Gutenberg-University, Hanns-DieterHüsch-Weg 17, 55128 Mainz, Germany

[3] Institute for Quantitative and Computational Biosciences, Johannes Gutenberg-University, BioZentrum I, Hanns-Dieter-Hüsch Weg 15, 55128 Mainz, Germany

‡ LAR & EK contributed equally to this manuscript

*correspondence: kerstec@uni-mainz.de

KEYWORDS: RNA-binding small molecules, binding thermodynamics, binding kinetics, selectivity, rational ligand design, riboswitches, electrostatic interactions, RNA

INTRODUCTION

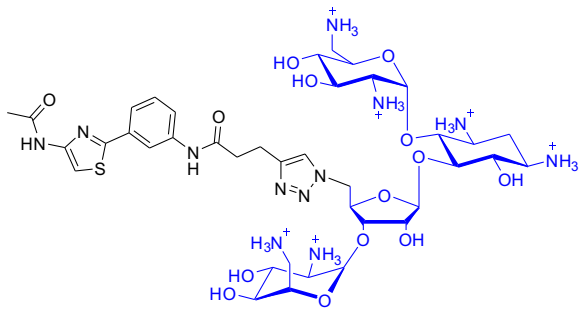
In recent years, interest in the use of ribonucleic acids (RNAs) as a therapeutic agent, as well as their potential as a drug target has been growing. This has been reflected in the development of new vaccines, antisense oligonucleotides (ASOs), and small interfering RNAs (siRNAs). Fruther, targeting bacterial riboswitches, viral RNA, or human dysregulated RNA elements holds the potential for novel antibiotics, antivirals, and treatments for previously untreatable diseases.[1–3] However, oligonucleotide-based therapeutic agents encounter several challenges, including delivery issues and the occurrence of adverse effects.[4,5] In contrast, small molecules offer superior pharmacokinetic properties and the potential for chemical optimization campaigns.[1,5] Further, the development of RNA-binding small molecules would significantly expand

the number of potential drug targets. At present, approximately 3,000 disease-related human proteins are known, many of them considered undruggable. However, targeting those proteins at the messenger RNA (mRNA) level with small molecules represents a promising approach.[6,7] In 2020, risdiplam, the first small molecule to target RNA other than the bacterial ribosome, received FDA approval. Risdiplam binds the survival motor neuron 2 (SMN2) pre-mRNA as a splicing modulator for the therapy of spinal muscular atrophy (SMA) and was identified in a phenotypic screening.[8–10] While several examples of RNA-binding small molecules were reported, optimization often proofed to be difficult.[7] One general challenge in RNA hit optimization and affinity enhancement is the requirement of a high information content at the target site arising from unpaired bases and a subsequent complex tertiary structure.[7,11] Based on this theory, high information content RNAs tend to be more likely to accommodate more potent ligands. Likewise, these regions are rare which also should result in selectivity, particularly in the cellular context where numerous similar low information content RNA motifs are present, but a structurally similar high information content RNA off-target is unlikely. By acquiring knowledge about the structural properties and potential unique binding sites of a target RNA, ligands can be rationally optimized.[7,11] However, there is currently no guideline on how to design selective, drug-like[12,13] RNA-targeting small molecules, especially for lower information content targets. Affinity optimization was addressed through a variety of different strategies in these cases, but no universal approach has been identified.[1–3]

Different strategies are under elucidation to overcome limited affinity. Ribonucleic targeting chimeras (RiboTACs)[5,14–19] and proximity-induced nucleic acid degraders (PINADs)[20] aim for the degradation of the target RNA requiring only sub-stoichiometric amounts of ligands like it is established for proteins in form of proteolysis targeting chimeras (PROTACs).[5,14–22] An intuitive concept to increase RNA-ligand affinity, is the conjugation of RNA-binding motifs with positively charged groups to interact with the negatively charged RNA backbone. Plain attractive forces of ionic interactions are masked by a diffuse ion atmosphere around nucleic acids.[23–25] Still, statistical evaluation of RNA-binding small molecules indicated beneficial effects arising from positively charged groups.[26,27] This concept was further explored by Maria Duca and coworkers using aminoglycoside-conjugates.[28–31] Aminoglycosides are known

antibiotics targeting the bacterial ribosome. Polar interactions between their amino and hydroxyl groups with the RNA backbone or nucleobases can likely occur[28] and play a leading role in short-range interactions,[32–34] but also in the diffusion towards the ribosomal A-site.[35,36] A correlation between higher charge and improved affinity or activity to its target was shown using computational[33,37] and experimental[34,38] methods. The polycationic character of aminoglycosides enhances binding but reduces their specificity causing severe side-effects like nephrotoxicity or ototoxicity.[39–41] This is also reflected in the re-occurrence of aminoglycosides as ligands for multiple unrelated RNA targets.[29,31,42–46] In the attempts to target other RNAs and to exploit their beneficial electrostatic interactions, aminoglycosides have been conjugated with various specific RNA-ligands to increase their selectivity and affinity.[28] Exemplarily, neomycin was linked to an aminophenyl-thiazole-based artificial nucleobase resulting in an improved oncogenic pre-microRNA (miRNA) binder (Figure 1A).[31] This artificial nucleobase was designed to specifically interact with T-A pairs.[47,48] Aminoglycoside conjugates were also found to target viral[30,43,49] and oncogenic[29,31,50–52] RNAs and to be sufficiently selective in cellular assays.[29,31,50,51] The implementation of basic amino acids showed similar beneficial effects.[29,30] For example, basic amino acids were linked to an artificial nucleobase to bind to the human immunodeficiency virus trans-activation response elements (HIV TAR, (Figure 1B).[30] Likewise, functionalized polyamines were identified to inhibit biogenesis of oncogenic pre-miRNAs in a screening. The highly active polyamine (Figure 1C) affected only a small set of oncogenic miRNAs and showed specific inhibition of cancer cell proliferation.[53] Implementing cationic moieties, like spermidines,[53,54] within RNA-targeting ligands can thus be used to increase affinity while not necessarily resulting in unselective, promiscuous binders.[3]

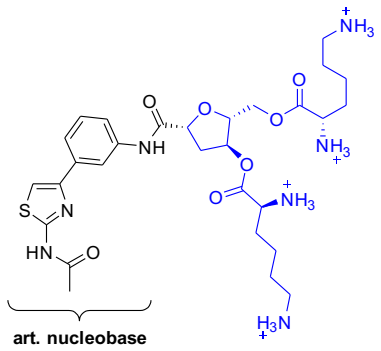
A



neomycin: K_D (pre-miRNA-372)= 7400 ± 120 nM[31]

conjugate: K_D (pre-miRNA-372)= 16 ± 8.8 nM[31]

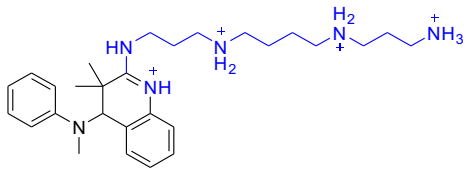
B



art. nucleobase: K_D (HIV-1 TAR) = nb[30]

conjugate: K_D (HIV-1 TAR) = 8 ± 1 µM[30]

C



K_D (pre-miRNA-372)= 0.15 ± 0.01 µM[53]

Figure 1. Molecular structures and binding constants K_D of **A**, Neomycin-nucleobase conjugate binding to pre-miRNA-372,[31] **B**, Artificial (art.) nucleobase basic amino acid conjugate binding to HIV-1 TAR RNA,[30] and **C**, A polyamine binding to pre-miRNA-372.[53] Polycationic moieties are highlighted in blue.

RNA-small molecule interactions are commonly described by binding affinities or biological activity. However, the additional determination of kinetic and thermodynamic parameters provides insights into fundamental principles of molecular recognition between RNA and small molecules.[55] Binding of some polycationic molecules, like the aminoglycoside neomycin[56] or polyamide amino acids[57], to their target RNA were

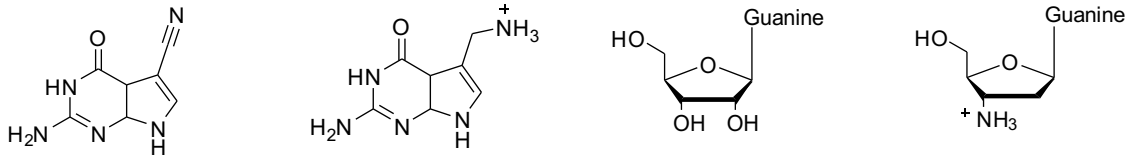
found to be enthalpy-driven. It was suggested that their polycationic moieties are involved in specific noncovalent interactions.[55,57] Polyamines were found to modulate RNA functions via specific interactions and to play a structural role distinct from divalent cations.[58] Comparing the thermodynamic binding profiles of similar compounds can give insight into the energetic contributions of small structural ligand changes, which can guide rational design.[59] The binding profiles of RNA ligands for on- and off-targets can help to explain selectivity patterns.[60] In addition, for bacterial riboswitches that are under kinetic control, a fast-associating ligand will likely be more efficient, since it has to compete with the natural ligand.[61]

The influence of cationic moieties within non-aminoglycoside RNA-binding small molecules on their kinetic profiles was evaluated in a few studies.[62–64] Wedekind and coworkers compared the natural ligands pre-queuosine-1 ($preQ_1$) and $preQ_0$ (Figure 1 A) binding to the *thermoanaerobacter tengcongensis* (*Tte*) $preQ_1$ riboswitch aptamer. $PreQ_1$ bears a methylamine group, which is replaced by a nitrile group in the case of $preQ_0$. Thus, $preQ_1$ is positively charged under physiological pH in contrast to $preQ_0$ which is neutral. The association rate of $preQ_1$ was found to be 12-fold faster than $preQ_0$ while dissociation rates were similar. This results in the higher affinity observed for $preQ_1$.[62] Micura and coworkers synthesized amino- and azido-functionalized derivatives of $preQ_1$ and tested them in *in vitro* and *in vivo* assays. Amino-functionalized derivatives showed enhanced k_{on}-values and affinities in agreement with the expectation of adding electrostatic interactions.[63] For the *tetrahymena* group I ribozyme, one hydroxy group of the cofactor guanosine was exchanged with an amino moiety (Figure 1B). It was hypothesized that the introduced basic group replaces a magnesium ion from the binding site. This led to a faster association rate and longer residence times of an early binding complex, which were necessary to enhance productive binding. Herschlag and coworkers termed this cationic group a “molecular anchor”. The strategy was proposed to be a universal principle to enhance association rates for well-placed anchors.[64] In a more poetical way, this effect of electrostatic anchoring can be regarded as a type of “Circe effect”.[65] This effect describes attractive long-range interactions to conduct a substrate into the active site of an enzyme, where it will be converted. It was defined by William P. Jencks and is named after Circe, a goddess in Greek mythology, luring men and turning them

into pigs. In case of RNA-binding small molecules this effect applies to the association towards the binding site without the following conversion.

Consequently, introduction of positive charges into RNA-ligands might be a promising design strategy to improve affinity by enhanced association rates. However, conjugation of RNA-ligands to poly-cationic moieties like aminoglycosides may not hold up as a universal RNA-ligand design concept due to toxicity, selectivity or pharmacokinetic issues. A deeper understanding of how positively charged moieties impact RNA-ligand binding affinity, selectivity, kinetics and thermodynamics is required to improve current design and development concepts for RNA-targeting small molecules. Therefore, we used the $preQ_1$ riboswitch and SAM-VI riboswitch aptamers as model systems to investigate the impact of ligand charges on affinity, binding kinetics, thermodynamics and selectivity. Surface plasmon resonance (SPR) spectroscopy, isothermal titration calorimetry (ITC) and microscale thermophoresis (MST) were used to examine the effects of electrostatic interactions on binding behavior. We aim to answer the questions: (1) Are positively charged groups generally leading to more potent RNA-ligands, (2) do these additional charges impair selectivity, (3) (how much) do positive charges enhance association rates, (4) what are the resulting thermodynamic binding profiles and (5) can general design principles around charged moieties be deduced for the development of RNA-binding small molecules?

A **B**



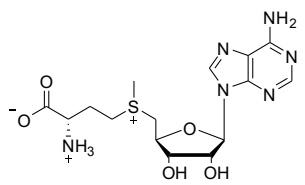
$PreQ_0$[62]	**$PreQ_1$**[62]	**Guanosine (G)**[64]	**G (2'H, 3'NH_3^+)**[64]
K_D = 35.10 ± 6.07 nM	K_D = 2.05 ± 0.29 nM	K_D = 110 µM	K_D = 0.094 µM
k_{on} = (0.65 ± 0.13) $\cdot 10^4$ $M^{-1}s^{-1}$	k_{on} = (7.77 ± 3.45) $\cdot 10^4$ $M^{-1}s^{-1}$	k_{on} = 4.0$\cdot 10^5$ $M^{-1}min^{-1}$	k_{on} > 2.1$\cdot 10^7$ $M^{-1}min^{-1}$
k_{off} = (2.22 ± 0.14) $\cdot 10^{-4}$ s^{-1}	k_{off} = (1.53 ± 0.47) $\cdot 10^{-4}$ s^{-1}	k_{off} = 44 min^{-1}	k_{off} > 2 min^{-1}

Figure 2. Molecular structures, binding constants K_D and kinetic parameters k_{on}/k_{off} of **A**, $PreQ_1$ compared to $preQ_0$ binding to the *Tte* $preQ_1$ riboswitch[62] and **B**, Guanosine and its amine derivative binding to *tetrahymena* group I ribozyme.[64]

RESULTS AND DISCUSSION

Rational design of charged ligands. Several RNA motifs are capable to discriminate between charged and uncharged natural ligands, including preQ$_1$ and SAM riboswitches, which strongly discriminate between preQ$_1$ and preQ$_0$ (Figure 2A),[62] and *S*-adenosyl methionine (SAM) and *S*-adenosyl homocysteine (SAH), respectively (Figure 3A)[66,67] Therefore, these RNAs can serve as ideal model systems for the investigation of charge effects on binding behavior. Particularly, the enhancement of association rates as a direct consequence of increased positive ligand charges was investigated. In this study, we selected two well-studied riboswitches for analysis: the *Tte* preQ$_1$ and the *Bifidobacterium angulatum* (*Ba*) SAM-VI riboswitch aptamer domains. We modified reported ligands of the respective riboswitches (Figure 3) to elucidate the impact of charges on binding affinity, selectivity, kinetics and thermodynamics. Charged moieties were added at ligand moieties oriented out of the binding site to minimize effects via direct ionic interactions with binding site residues and to better describe long-range attraction and solvent-exposed interactions.

A

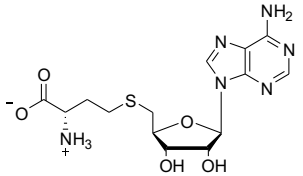


+1

SAM

K_D = 0.33 ± 0.06 µM[66]

k_{on} = (2.46 ± 0.07) · 10^4 $M^{-1}s^{-1}$ [66]

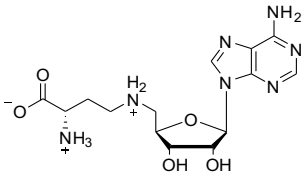


0

SAH

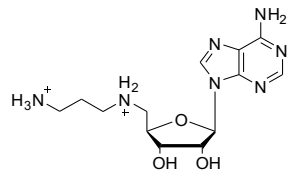
K_D = 10.9 ± 1.3 µM[66]

k_{on} = (5.50 ± 0.27) · 10^3 $M^{-1}s^{-1}$ [66]



+1

1



+2

2

Figure 3. Polycationic small molecules under elucidation. **A**, Ligands of the *Ba* SAM-VI riboswitch. Formal charges are shown in blue.

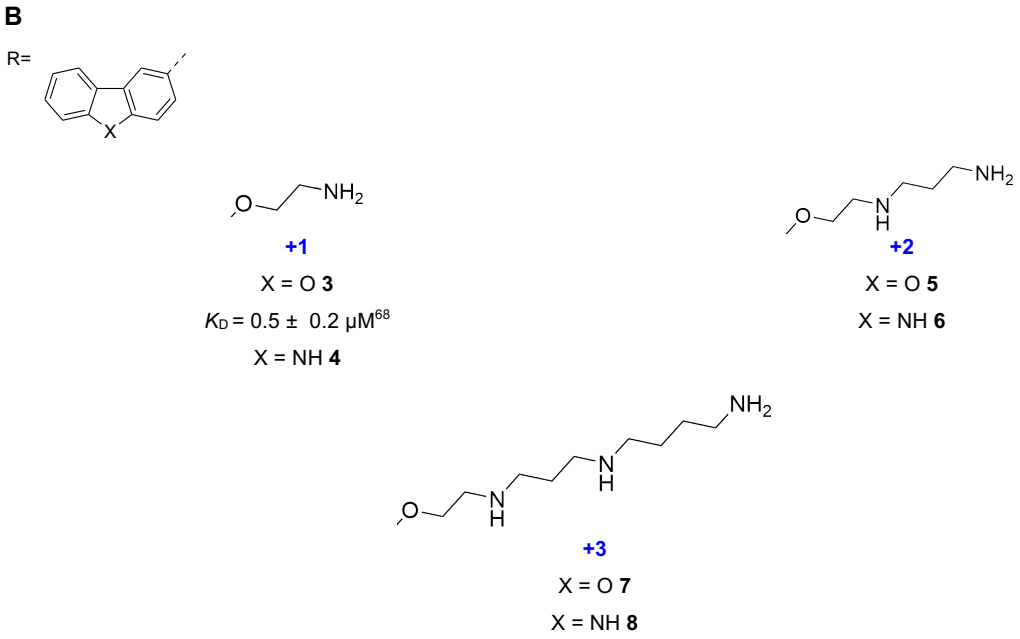


Figure 3. Polycationic small molecules under elucidation. **B**, Synthetic ligands of the *Tte* preQ$_1$ riboswitch. Formal charges are shown in blue.

The SAM-VI riboswitch belongs to the widespread family of SAM-sensing riboswitches, which are involved in the regulation of SAM metabolism.[66,69] SAM is a ubiquitous methylation cofactor, and these riboswitches are found in the majority of known bacterial species.[70,71] SAM riboswitches strongly discriminate between the positively charged SAM and its neutral methylation side product SAH. For example, the *Ba* SAM-VI riboswitch has a 33-fold higher affinity to SAM compared to SAH.[66,67] Previously, it was demonstrated that the positively charged sulfonium moiety of SAM is the primary factor responsible for its selective binding to the associated riboswitches.[66,72–74] The SAM-binding site of the SAM-VI riboswitch lies at the junction between stems P1 and P2. The binding pocket consists of the canonical Watson-Crick base pair G9–C32 and the non-canonical *trans* Watson-Crick base pair G7–G33. Additionally, the four continuously stacked bases A34, A36, A37, and G38 and the junctional U8 form a unique binding pocket. SAM adopts an extend conformation with the methionine tail exceeding the binding site. The adenine moiety forms hydrogen bonds with the Watson-Crick edge of U8, A36 and A37 as well as stacking interactions with G9–C32 and G33.[66] Notably, the positively charged sulfonium moiety does not directly interact with binding site residues (Figure 4A).[66] The closest potential interacting atoms are oxygens (O4) of U6 and U8 which might occasionally interact by charge-assisted

chalcogen bonds as described for SAM-I riboswitches.[75] To investigate the influence of charges on binding behavior in detail, the overall neutral, zwitterionic SAH, SAM and its aza-analogue compound **1** carrying one positive net charge, and **2** which carries two positive net charges due to removal of the carboxylic acid were tested (Figure 3A). Even though the SAM-VI riboswitch strongly discriminates between SAM and SAH, none of the modified groups directly interacts with the riboswitch's buried binding site (Figure 4A).

The second model system is the *Tte* preQ$_1$ riboswitch aptamer. Additional to its discrimination of the natural ligands preQ$_1$ (one positive net charge) and preQ$_0$ (neutral),[62] Connelly *et al.* reported the selective binding of dibenzofuran- and carbazole-based ligands with sub-micromolar potency (compound **3**, Figure 3B).[68] PreQ$_1$ and preQ$_0$ were used for assay establishment. The high affinity and association-rate of preQ$_1$ (K_D = 2.1 nM, k_{on} = 7.8·10^4 1/Ms)[62] limits further elucidation of additional charges as higher affinities and association rates are at the resolution limit of the assays used.[76,77] Therefore, polycationic derivatives of the lower-affinity, synthetic ligands were obtained and tested (**3**–**8**, Figure 3B). The *Tte* preQ$_1$ riboswitch binding pocket consists of the base quartets A28–C16–G5–A27 forming the "floor" and A14–G11–C30–C7 forming the "ceiling". PreQ$_1$ stacks in between these residues and interacts via additional hydrogen bonds to U6, C15 and A29 (Figure 4B). PreQ$_0$ differs from preQ$_1$ with a nitrile moiety instead of an amino group. This amino group contributes to the stability of the RNA-preQ$_1$ complex by hydrogen bonds with G11. The loss of this charged group results in a reduced affinity and association rate.[62,68] The dibenzofuran ligand **3** is positioned similar to preQ$_1$ in the binding pocket (Figure 4C). The dibenzofuran ring is stacked between G11 "ceiling" and G5–C16 "floor". The furan oxygen atom forms a hydrogen bond to A29.[68] Differently from the native ligand's binding mode, C15 is rotated to form a face-to-edge π-stacking interaction with compound **3** instead of the Watson-Crick-like pair found in the preQ$_1$-complex. Like for the SAM-VI riboswitch, the sidechain of ligand **3** is oriented out of the binding site (Figure 4C) making it an ideal attachment point for additional basic centers (ligands **5**–**8**).

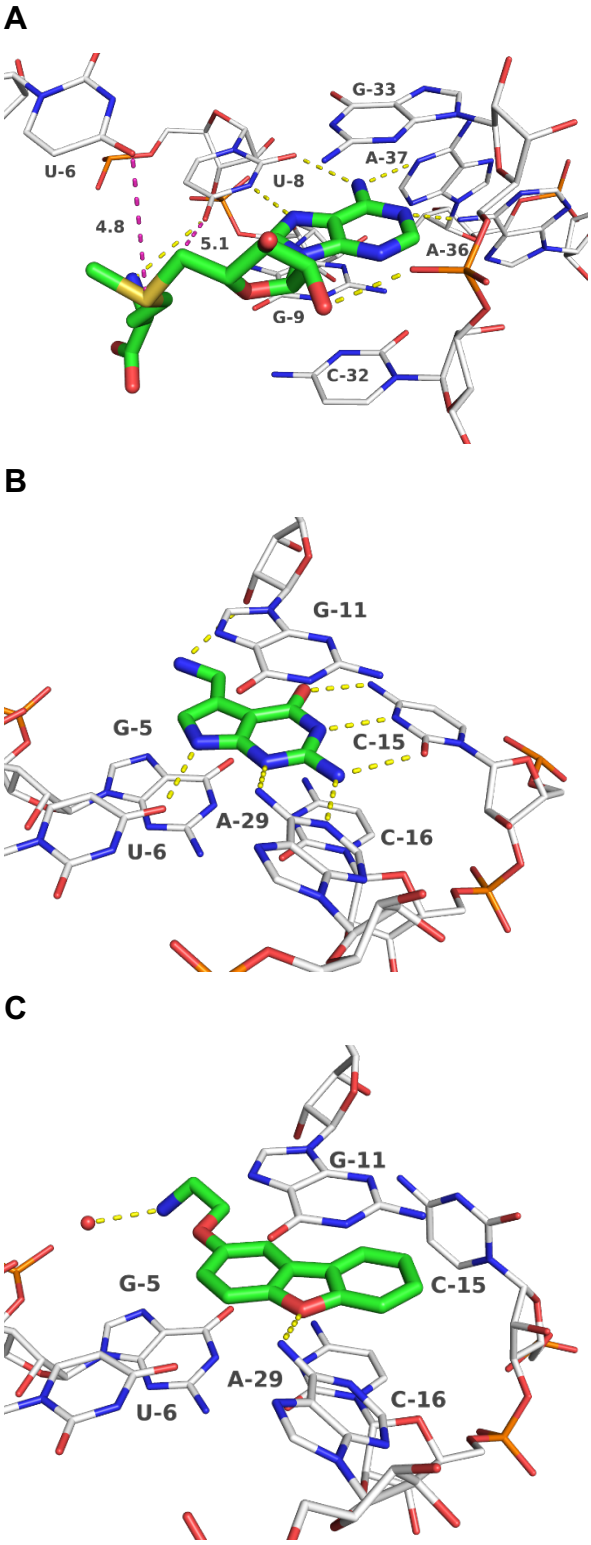


Figure 4. **A**, *Ba* SAM-VI riboswitch in complex with SAM (PDB-ID: 6LAS). Distances between sulfonium ion and oxygen atoms (O4) of U6 and U8, respectively, are depicted as magenta dashed lines with distances in Å.316 **B**, *Tte* preQ$_1$ riboswitch in complex with preQ$_1$ (PDB-ID: 3Q50).317 **C**, PreQ$_1$ riboswitch in complex with synthetic ligand **3** (PDB-ID: 6E1S).317 Ligands are depicted with green, RNA with white carbon atoms. Polar interactions are shown as yellow dotted lines.

SAM-VI riboswitch binding studies. SPR, MST and ITC experiments of SAM, SAH, compounds **1** and **2** with the *Ba* SAM-VI riboswitch aptamer domain were performed (Table 1). Results for the reference ligands SAM and SAH are in good agreement with literature reported values (Figure 3A) and variations between SPR, MST and ITC are within the usual range when comparing different assay methods.[78–80] In contrast to the zwitterionic, net neutral SAH, SAM carries an additional methyl group at the sulfur atom and a formal charge of +1. Strong discrimination between SAM and SAH was observed across the different assays. Discrimination factors for the binding affinities showed that SAH binds between 21-fold and 49-fold weaker than SAM. These results are in good agreement with previously reported ITC data showing a discrimination factor of 33[66] in K_D-values. Either the additional positive charge or the methyl group or a combination thereof is responsible for the observed discrimination between SAM and SAH. Lim *et al.* reported previously, that the positively charged sulfonium moiety is the key factor of SAM selectivity in SAM-I riboswitches as binding modes of SAM and SAH are highly similar.[72] Likewise, similar ligand conformations for SAM and SAH are also found for the SAM-VI riboswitch (Figure 3A, Figure S23). The aza-derivative of SAH, **1** carries one positive formal charge like SAM, but lacks the sulfur methyl group. The binding affinity of compound **1** with a K_D-value of 55.5 µM was comparable to SAH (34.4 µM) in MST. For ITC a higher (3.86 µM) and in SPR a slightly higher affinity (12.3 µM) than SAH (K_D = 15 µM in ITC and 150 µM in SPR, respectively) was observed. The exchange of sulfur for nitrogen and with it introducing an additional positive charge compared to SAH, did still result in lower binding affinities compared to SAM (K_D-values between 0.7 µM in MST and ITC, and 3.7 µM in SPR). A charge-assisted chalcogen bond is discussed as the driving factor of the tight binding of SAM to various SAM riboswitches. The Coulombic attraction between the positive charge of the sulfonium ion and an electron donor is enhanced by the interaction between the sulfur σ-hole and the electron donor. In the case of the *Ba* SAM-VI riboswitch, oxygen atoms (O4) of U6 or U8 could potentially act as the electron donor atoms (Figure 4A).[75] This might explain the higher affinity of SAM compared to its aza-analogue **1** despite the same formal charge of +1. However, different size between sulfur and nitrogen and the presence of a methyl group may also contribute to affinity. Further the sulfonium ion is permanently charged while ammonium could be partially de-protonated (pK_a-values of

7.6 and 7.8 for **1** and **2**, respectively, calculated using MOE2022.02). Compound **2** lacks the carboxylate functionality of the other SAM-derivatives resulting in two positive formal charges. The binding affinities determined with all three methods (SPR, MST, ITC) were comparable with values in the one-digit micromolar range. Compound **2** showed enhanced affinity compared to the net uncharged SAH and **1** (formal charge +1). Thus, the series SAH, compound **1** and **2** revealed a trend in terms of increased affinity by higher formal charges with SAM being superior to non-sulfonium-containing molecules.

Comparison of the kinetic parameters determined by SPR and kinITC (Table 1, Figures S1, S2, S15, and S16) showed similar trends in binding kinetics. Especially in SPR, the gold standard method for determination of binding kinetics, the sum of formal charges correlated with faster association rates. SAH showed a k_{on}-value of $3.9 \cdot 10^2$ 1/Ms, which was increased up to $1.3 \cdot 10^4$ 1/Ms for compound **2**. Additional positively charged moieties seem to be responsible for this effect. Notably, increasing formal charges from +1 in compound **1** to +2 in compound **2** had only minor impact (k_{on}-factor 4.4 in SPR). In kinITC, the association rates showed overall similar trends, with SAH being the slowest one binding, even though absolute k_{on}-values differ from SPR (k_{on} = $3.9 \cdot 10^2$ 1/Ms in SPR, $3.4 \cdot 10^3$ 1/Ms in kinITC). Here, the effect of additional charges was also less pronounced for the exchange from +1 to +2 formal charge resulting in overall similar k_{on}-values for SAM, **1** and **2**. The dissociation rates were broadly similar across both methods and all compounds showed k_{off}-values between $0.98 \cdot 10^{-2}$ 1/s and $5.57 \cdot 10^{-2}$ 1/s. Thus, different binding affinities can be primarily attributed to differences in increased association rates as observed in SPR.

The binding of all SAM-VI riboswitch ligands was highly enthalpy-driven with ΔH ranging from – 67.4 kJ/mol (SAM) to –108 kJ/mol (**1**). Highest enthalpic contributions were observed for compounds **1** and **2**. Overall, binding of compounds **1** and **2** results in an only moderate increase of ΔG compared to SAH due to a higher entropic penalty (– TΔS = +77.2 kJ/mol for **1** and +54.5 kJ/mol for **2**, respectively) indicative for a strong enthalpy-entropy compensation event.[81,82] Similar enthalpy-entropy compensation effects were described for other polycationic compounds binding to RNA.[57,83,84] The replacement of sulfur (SAM, SAH) by nitrogen in compounds **1** and **2** resulted in an increased enthalpic contribution with comparable binding free energies and

dissociation constants to SAM. This more exothermic process indicates strong RNA-ligand interactions at the cost of entropy, suggesting a higher intrinsic stability of the formed complex.

Pan-RNA promiscuity and selectivity was investigated by testing SAM, SAH, **1** and **2** against the structurally unrelated *Tte* preQ$_1$ riboswitch by SPR and MST (Table 1). In MST, all SAM-VI riboswitch ligands were selective with respect to their target. In SPR, binding of compound **2** to the preQ$_1$ riboswitch was observed. However, differently from the on-target binding sensorgrams (Figure 5A), the off-target sensorgrams showed fast-on/fast-off characteristics which are indicative for loose, unspecific binding events (Figure 5B).[85] In contrast, the sensorgrams of compound **2** binding to the SAM-VI riboswitch showed the typical shape for specific binding interactions with higher response units (Figure 5A). The effect of fast-on/fast-off rates was only observed for the highest charged compound **2** (Figure 5B). Thus, a correlation between unspecific binding and the sum of positive formal charges of the molecule can be assumed. In agreement with MST, steady-state evaluation of the SPR data revealed saturation of compound **2** binding to the SAM-VI riboswitch (Figure 5C). For this ligand **2**, no saturation of the off-target could be reached (Figure 5D). This suggests accumulation of the compound on the off-target's surface, eventually by loose, rather unspecific ionic interactions.

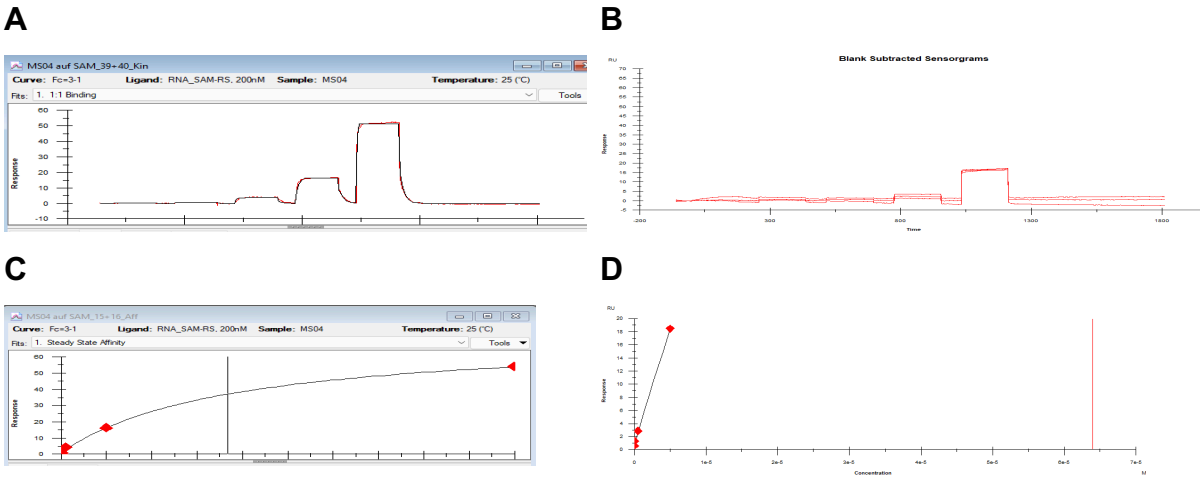


Figure 5. **A**, Representative SPR-sensorgram of **2** binding to the SAM-VI riboswitch (K_D = 4.17 µM). **B**, Representative SPR-sensorgram of **2** showing loose, unspecific preQ$_1$ riboswitch binding. **C**, SPR steady-state response of **2** binding to the SAM-VI riboswitch reaching saturation (K_D = 8.64 µM). **D**, SPR steady-state response of **2** binding to the preQ$_1$ riboswitch showing no saturation within tested concentrations (K_D = not determinable).

PreQ$_1$ riboswitch binding studies. For the *Tte* preQ$_1$ riboswitch aptamer binding of preQ$_1$, preQ$_0$ and synthetic ligands **3**–**8** was investigated by SPR, MST and ITC (Table 2). We observed an enhanced affinity of the positively charged preQ$_1$ over the neutral preQ$_0$ with high consistency in all experiments as described in literature.[62] PreQ$_1$ showed a 3-fold higher affinity compared to preQ$_0$ in SPR analysis, with being 6-fold in MST and 4-fold in ITC.

Further, the kinetic binding parameters showed association rate constants of the positively charged preQ$_1$ to be 2-fold higher in SPR and in kinITC compared to the uncharged preQ$_0$. As described previously, k_{off}-values for preQ$_1$ and preQ$_0$ were similar and the higher potency of preQ$_1$ can be annotated to faster association. Again, absolute values differ between SPR and kinITC which might be due to fundamental differences in the assays including immobilization for SPR or high concentrations in ITC experiments. Still, relative values between preQ$_1$ and preQ$_0$ agree with literature.[62] The enhanced on-rate while maintaining similar off-rates resulted in the observed higher binding affinity of preQ$_1$ over preQ$_0$ which was attributed to the different ligand charge.

Sensorgrams of preQ$_1$ binding to its riboswitch showed better fits when using a heterogeny target model compared to the standard 1:1 binding model. Additionally, the theoretically calculated maximal response units (47 RU, Figure S1 F) were exceeded (70 RU, Figure S1 F), indicating binding of more than one molecule per riboswitch. This suggests that the *Tte* preQ$_1$ riboswitch may have at least two binding sites. The recently published crystal structure of the structurally similar *Carnobacterium antarcticum* preQ$_1$ riboswitch showed two bound preQ$_1$ molecules (Figure S24, PDB 8FB3).[86] Probably, a similar second binding event is also possible in the *Tte* preQ$_1$ riboswitch.

The 1:1 binding model underlies the steady-state analysis, thus not representing the best fitting model for preQ$_1$. This can be the reason, why our binding affinities determined by steady-state analysis of the SPR data were similar and did not show discrimination between both natural ligands.

The natural ligands preQ$_1$ and preQ$_0$ showed thermodynamic binding profiles with large enthalpic contributions of –150 kJ/mol and –129 kJ/mol, respectively (Table 2). As described for the SAM-VI riboswitch, the highly exothermic binding is compensated

by an entropic penalty of –TΔS = +106 kJ/mol for $preQ_1$ and –TΔS = +88.5 kJ/mol for $preQ_0$. The substitution of the nitrile group in $preQ_0$ by the more flexible methylamino group in $preQ_1$ resulted in slightly favorable enthalpic and unfavorable entropic contributions. The positively charged amine group is involved in specific polar interactions (Figure 4B), but rigidification upon binding. This is similar to the results obtained for the SAM-VI riboswitch where additional charges also led to more favorable binding enthalpy partially compensated by unfavorable entropy.

The synthetic dibenzofuran ligand **3** with one positive charge was reported to bind the *Tte* $preQ_1$ riboswitch with K_D = 0.5 ± 0.2 µM.[68] Several dibenzofuran and carbazole derivatives with one positive charge showed similar affinities.[68,87,88] We selected dibenzofuran **3** and the carbazole analog **4**, both carrying one positive charge, as starting points for further polycationic derivatives **5**–**8** (Figure 3B). In SPR, sensorgrams showed fast-on-fast-off shapes, while steady-state K_D-evaluation showed no saturation in dose-response curves (Figure S1 G, H). For MST poor signal to noise-ratios were observed prohibiting accurate determination of K_D-values (Figure S3 C–H, Figure S5 C–H). Likewise, ITC showed no heat signals beyond ligand dilution during titrations (Figure S11 G–L). The reason why no binding was detected for **3**–**8** in these methods (MST, ITC and steady-state SPR signals) might lie in a loose binding behavior like described for unspecific binding of **2** to the $preQ_1$ riboswitch (Figure 5B, D). Probably, the positive charge and intercalation properties of the three-membered ring systems are responsible for unspecific interactions. These results do not necessarily contradict the results published previously.[68,87] Possibly, the unspecific contributions overlap the specific ones in a manner that only the unspecific binding can be observed in SPR where a high negative charge density on the sensor surface is facing a (poly)cationic ligand. The synthesized multibasic dibenzofuran and carbazole derivatives **5** and **6** with two positive charges and **7** and **8** with three positive charges were designed following the findings in spermidine-like RNA binders.[53,54] No binding was detected for these polycationic compounds in MST and ITC. The SPR sensorgrams and steady-state evaluation were similar to the one-basic compounds **3** and **4**, showing unspecific loose binding without reaching saturation (Figure S1 I–L). Especially for the three-basic compound **8**, the theoretically calculated maximal response units (234 RU, Figure S1 L) were exceeded (Figure S1 L), indicating binding

of more than one molecule. This suggests accumulation of the compound on the target's surface by electrostatic interactions and/or intercalation.

For selectivity studies, binding of *Tte* $preQ_1$ riboswitch ligands to the SAM-VI riboswitch were investigated using SPR and MST (Table 2). In MST, $preQ_1$ and $preQ_0$ were selective for the *Tte* $preQ_1$ riboswitch showing no binding to the *Ba* SAM-VI riboswitch. Further, no binding was detected in steady-state evaluation of SPR. The sensorgrams of $preQ_1$ showed partially fast-on/fast-off shapes and low response units (5 RU vs. maximal 40 RU calculated, Figure S1F) similar to the ones of compound **2** against its off-target (Figure 5B) in contrast to the sensorgrams involving its target ($preQ_1$ riboswitch, Figure S1F). The single positive charge of $preQ_1$ probably increases the possibility of weak loose binding events in SPR. Dibenzofuran and carbazole derivatives **3**–**8** did not show binding in the MST assay but showed sensorgram shapes of unspecific loose binding to the SAM-VI riboswitch similar to their proposed target, the $preQ_1$ riboswitch. In case of the highest-charged compounds, the theoretically calculated maximal response units (**7**: 200 RU, **8**: 190 RU, Figure S1K,L) were exceeded (**7**: 458RU, **8**: ; Figure S1K,L) in case of the highest-charged compounds.

Table 1. SPR, MST, ITC and selectivity over *Tte* preQ$_1$ riboswitch results for SAM-VI riboswitch ligands. Kinetic parameters for ITC experiments were determined using kinITC.[89–91]

Cpd	Sum of formal charges	SPR				MST	ITC						Selectivity	
		K_D [µM]	k_{on} [1/Ms]	k_{off} [1/s]	$K_{D,ss}$ [µM]	K_D [µM]	K_D [µM]	ΔG [kJ/mol]	ΔH [kJ/mol]	–TΔS [kJ/mol]	k_{on} [1/Ms]	k_{off} [1/s]	SPR	MST
SAM	+1	3.72 ± 0.41	(2.98 ± 0.69) ·10^3	(9.80 ± 3.06) ·10^{-3}	4.53 ± 0.64	0.699 ± 0.185	0.712 ± 0.022	–35.1 ± 0.1	–67.4 ± 3.7	32.1 ± 3.4	(2.02 ± 0.51) ·10^4	(1.44 ± 0.33) ·10^{-2}	✓	✓
SAH	±0	150 ± 39	(3.88 ± 1.18) ·10^2	(5.40 ± 0.67) ·10^{-2}	121 ± 32	34.4 ± 12.0	15.0 ± 3.4	–27.6 ± 0.5	–35.5 ± 1.6	7.87 ± 2.09	(3.43 ± 1.44) ·10^3	(4.81 ± 1.30) ·10^{-2}	✓	✓
1	+1	12.3 ± 1.3	(2.94 ± 0.15) ·10^3	(3.61 ± 0.19) ·10^{-2}	20.7 ± 3.3	55.5 ± 13.3	3.86 ± 0.41	–30.9 ± 0.3	–108 ± 13	77.2 ± 13.2	(1.40 ± 0.46) ·10^4	(5.57 ± 1.60) ·10^{-2}	✓	✓
2	+2	4.17 ± 1.56	(1.28 ± 0.34) ·10^4	(4.81 ± 0.27) ·10^{-2}	15.9 ± 13.8	8.64 ± 3.30	1.73 ± 0.51	–33.0 ± 0.8	–87.5 ± 1.8	54.5 ± 1.2	(1.36 ± 0.10) ·10^4	(2.38 ± 0.84) ·10^{-2}	✓[a]	✓

[a]SPR-sensorgrams show fast-on and fast off-rates indicating a loose, unspecific binding event (see Figure 5B, D)

Table 2. SPR, MST, ITC and selectivity over SAM-VI riboswitch results for *Tte* preQ$_1$ riboswitch ligands. Kinetic parameters for ITC experiments were determined using kinITC.[89–91]

Cpd	Sum of formal charges	SPR				MST	ITC						Selectivity	
		K_D [nM]	k_{on} [1/Ms]	k_{off} [1/s]	$K_{D,SS}$ [nM]	K_D [nM]	K_D [nM]	ΔG [kJ/mol]	ΔH [kJ/mol]	–TΔS [kJ/mol]	k_{on} [1/Ms]	k_{off} [1/s]	SPR	MST
PreQ$_0$	±0	74.0 ± 29.4	(5.72 ± 4.49) ·10^3	(3.07 ± 0.48) ·10^{-4}	687 ± 197	70.4 ± 6.8	85.4 ± 3.3	–40.4 ± 0.1	–129 ± 16	88.5 ± 15.9	(3.28 ± 0.59) ·10^4	(2.81 ± 0.58) ·10^{-3}	✓	✓
PreQ$_1$	+1	24.1 ± 9.9	(1.10 ± 0.11) ·10^4	(2.60 ± 0.96) ·10^{-4}	503 ± 184	12.0 ± 0.8	23.1 ± 7.4	–43.7 ± 0.9	–150 ± 7	106 ± 7	(7.34 ± 2.88) ·10^4	(1.62 ± 0.69) ·10^{-3}	✓[a]	✓
het. Fit 1 (51%)[b]		14.4	2.77·10^4	3.99·10^{-4}	-									
het. Fit 2 (30%)[b]		950	5.10·10^2	4.84·10^{-4}	-									

[a]SPR-sensorgrams show partially fast-on and fast off-rates indicating a loose, unspecific binding event (Figure S1F)

[b]IM calculation based on heterogeny ligand fit for a representative SPR sensorgram of preQ$_1$. Percentage of binding mode occurrence (Figure S2).

CONCLUSIONS

We investigated the addition of positive ligand charges as a design strategy in RNA-binding small molecule drug discovery. In this regard, the impact of additional ligand charges on binding affinity, kinetics, thermodynamics and selectivity was examined using the *Tte* preQ$_1$ and *Ba* SAM-VI riboswitch aptamer domains as model systems. Combined methods of SPR, ITC and MST were used to examine the effects of protonated amines as “electrostatic anchors” on binding characteristics. For the SAM-VI riboswitch, the natural ligands SAM and SAH as well as synthetic positively charged SAM-analogs **1** and **2** were tested (Figure 3A). For the *Tte* preQ$_1$ riboswitch, preQ$_1$ and preQ$_0$ as natural binders (Figure 2A), and synthetic dibenzofuran and carbazole derivatives (Figure 3B) were analyzed. Our results for reported ligands agree with previous findings[62,66,69], which validated the methods under investigation. Overall, the addition of positively charged moieties had only minor effects on the binding affinities and association rates (Tables 1 and 2). For both compound series, SAM, SAH, compounds **1** and **2,** and preQ$_0$, preQ$_1$, only a slight trend that higher positive charges can lead to enhanced association rates, and thus affinity to the RNA targets, was observed. These changes were only moderate and cannot be considered as a universal design principle to overcome affinity cliffs.

Thermodynamic binding profiles of all ligands were dominated by enthalpy, partially compensated by unfavorable entropy terms. This observation was more pronounced for ligands carrying multiple positive charges. Like for protein-ligand interactions, this suggests that strong interactions like salt bridges reduce flexibility of interacting moieties for both ligand and target resulting in an entropic penalty.[55,84,92] As a consequence, this enthalpy-entropy compensation leads to smaller changes in ΔG and affinity than additional interaction sites might suggest.

In the selectivity study, all ligands of a particular target, SAM, SAH, **1** and **2** for the SAM-VI riboswitch, and preQ$_1$ and preQ$_0$ for the preQ$_1$ riboswitch, discriminated their target from the respective off-target. Compounds **3**–**8** from the dibenzofuran and carbazole series showed no binding in ITC and MST. For these synthetic, putative preQ$_1$ riboswitch ligands, however, fast association and dissociation rates were observed in SPR experiments for both riboswitches. The shapes of SPR sensorgrams

and the steady-state analysis of SPR data showed clear differences for specific ligands compared to such a loose binding behavior. Compounds, which did not bind in MST and are positively charged, had a higher probability to be identified as fast-on/fast-off loose binders[85] in SPR. Therefore, a *caveat* arises for SPR-screenings against RNA-targets: Sensorgrams of loose and likely unspecific binders showed fast-off/fast-on rates, and no saturation of the target for steady-state analysis was reached. This was also the case for the SAM-VI riboswitch ligand **2** binding to the preQ$_1$ riboswitch off-target. We hypothesize this loose, unspecific binding is caused by the high ligand charges leading to surface interactions between protonated amine moieties of the ligand with the phosphate-backbone of immobilized RNA. This behavior further suggests accumulation of the compounds on the target's surface probably by electrostatic interactions like the "ion atmosphere" found around nucleic acids.[23–25] For **3**–**8** this effect might be further enhanced by intercalation of the three-ring core scaffold overshadowing the previously reported specific interactions[68,87] with the preQ$_1$ riboswitch. Similarly, increased potency by acridine-conjugation was described previously for multiple unrelated RNA-targets.[93,94] However, such loose binding events to the off-targets probably might also be an artifact of immobilization-based techniques like SPR, as no binding was observed via MST and ITC for these compounds. Hence, it may not have biological relevance through RNA interaction competition. Poly-cationic compounds still can be selective for their target.[29,31,50,51,53] Consequently, ligands carrying one or probably more positive charges are no exclusion criterium *per se* when it comes to binding affinity and selectivity. Still, negative pharmacokinetic effects on permeability and cellular uptake[63] must be considered in RNA-targeting drug design. Subsequently, our findings demonstrate that introducing a positive ligand charge for specific interactions does not need to be generally avoided. The designed ligands of this study introduced positive charges mostly in molecular regions oriented away from the binding site. This aimed to minimize specific interactions with binding site residues and to elucidate the impact of long-range ionic interactions in sense of electrostatic anchoring or enhanced association in a Circe-like effect. Overall, charge influence on affinity and association rates was low, but larger affinity gains can be obtained when placing direct ionic interactions or charge-assisted hydrogen bonds. An example can be found in the amine-substituted guanosine-derivative (Figure 2B) binding to the

tetrahymena group I ribozyme with a 1000-fold affinity increase.[64] Another example is the here and previously described discrimination of $preQ_1$ riboswitches between $preQ_1$ and $preQ_0$.[62,95] Within the series of SAM-derivatives, the biggest affinity increase was reached from SAH when exchanging neutral sulfur to a sulfonium ion in SAM and partly when changed to a protonatable amine (compound **1**). Even though no direct interaction is indicated by the SAM-VI riboswitch – SAM complex structure, one could speculate if the flexibility of the amino acid substructure and the intrinsic flexibility of RNA are sufficient to occasionally form specific interactions between RNA and the sulfonium or ammonium groups. A likely interaction partner for a charge-assisted chalcogen bond with SAM as described previously for the SAM-I riboswitch[75], or ion-dipole interaction with compounds **1** and **2** would be oxygen (O4) of the U6 or U8 residues (Figure 4A). Further examples, where charged ligand moieties led to potent ligands are described for pre-miRNA-372[31] (Figure 1A), the thiamine pyrophosphate (TPP) riboswitch[96], the hepatitis C virus internal ribosome entry site (HCV IRES)[97,98] and myotonic dystrophy type 2 (DM2)-causing CCUG-repeats.[45]

In summary, our kinetic and thermodynamic *in vitro* studies provide valuable insights into molecular recognition between RNA and small molecules to answer the initially raised questions: (1 - affinity) The implementation of solvent exposed, positively charged moieties leads only to minor increases in potency insufficient to overcome affinity cliffs. Differently, purposefully placed moieties forming ion-ion or ion-dipole interactions can have larger impact.[62,64,66,69] (2 - kinetics) Small to moderate improvements in affinity from ligand charges seem to arise from faster association rates agreeing with electrostatic anchoring and a Circe-like effect. (3 - selectivity) No general loss of selectivity and off-target (*pan*-RNA) binding was observed with up to +3 formal charges (compounds **7**, **8**) in MST for our model systems. SPR sensorgrams showed fast-on/fast-off kinetics and loose unspecific binding of multi-basic compounds. This behavior can be identified from the sensorgrams (fast-on/fast-off) while saturation in dose-response curves is not reached. Awareness and testing with orthogonal methods can protect against following-up false positive multi-charged hits in SPR screenings against RNA. *Vice versa*, the implementation of a well-placed charged moiety does not impair selectivity *per se*. (4 - thermodynamics) All thermodynamic binding profiles were dominated by high enthalpic contributions and

unfavorable entropy. This enthalpy-entropy compensation is also well described for proteins where enthalpically strong ion-ion and ion-dipole interactions result in an entropic penalty.[55,84,92] (5 - design strategy) Taken together, implementation of positively charged groups cannot serve as a universal strategy for RNA-ligand design. The low contribution to affinity is insufficient to overcome affinity cliffs (1) even though association rates are slightly enhanced (2). However, a well-placed charged moiety could enhance potency without necessarily having negative impact on selectivity (3). Still, physicochemical properties and their effect on pharmacokinetics must be considered in hit to lead optimization. These findings contribute to a better understanding of RNA-small molecule recognition and provide a piece to the puzzle of how to design and optimize selective drug-like RNA-ligands.

EXPERIMENTAL SECTION

General synthetic methods

All reagents and solvents were commercial grade and used without further purification. Reaction progress was monitored by thin layer chromatography (**TLC**) using Alugram Xtra F254 silica plates from Machery-Nagel. In addition, high-performance liquid chromatography/electron spray ionization mass spectrometry (**HPLC/ESI-MS**) was used to control reaction's conversion and to determine the identity as well as the purity of compounds tested in the assay. An Agilent 1100 series HPLC system and either an Agilent Zorbax SB-Aq (4.6 mm x 150 mm, 5 µm) or Poroshell 120 EC-C18 (2.10 mm x 150 mm, 4 µM) column coupled to an Agilent 1100 series LC/MSD Trap with ESI, was used. The measurements were conducted with a gradient of acetonitrile and water (+0.1% formic acid) with 10%–90% acetonitrile over 10 min with a flow rate of 0.7 mL/min unless otherwise stated. Signals were detected at 254 nm with quantification by area under curve (AUC) and masses were determined in positive ionization mode (ESI). For the determination of purity and identity for the most polar compounds **7** and **8**, a Waters Alliance e2695 HPLC system and a MZ-Aqua Perfect C18 (4.6 x 250 mm, 10 µm) analytical HPLC column coupled to a Waters ACQUITY QDa single quadrupole detector with ESI was used. **Column chromatography** was

performed with silica gel 60 (40–63 µm) from Machery-Nagel. **Flash column chromatography** was performed with the Biotage Isolera™ One system. For reversed-phase flash chromatography prepacked columns of the type Biotage®Sfär C18 Duo from Biotage® were used. In the case of normal-phase flash chromatography, columns were self-packed with silica gel 60 (15–40 µm) from Merck KGaA. **Preparative HPLC** purification was performed with a Varian PrepStar system (model 218) with a MZ-Aqua Perfect C18 20 x 250 mm, 7 µm preparative LC column and acetonitrile/water +0.1% trifluoroacetic acid as mobile phase. **Melting points (mp)** (uncorrected) were measured with an MPM-H3 using semi-open capillaries. Proton (^{1}H) and carbon (^{13}C) nuclear magnetic resonance (**NMR**) spectra were recorded on Bruker Fourier 300 MHz (300 MHz for ^{1}H and 75 MHz for ^{13}C). The chemical shift was abbreviated to δ and has the unit ppm. The chemical shifts were referenced to the solvent peaks in ^{1}H: δ = 7.26 ($CDCl_3$), 2.50 (DMSO-d_6), 3.31 (CD_3OD) ppm and in ^{13}C: δ = 77.16 (Chloroform-*d*), 39.52 (DMSO-d_6), 49.00 (CD_3OD) ppm purchased from Deutero GmbH. In addition, the following abbreviations for the multiplicities of the peaks were defined: s (singlet), d (doublet), dd (doublet of doublet), t (triplet), td (triplet of doublet), q (quartet), p (pentet), and m (multiplet). Coupling constant *J* was given in Hz and denoted together with the number of signaling hydrogen atoms for ^{1}H-NMR spectra. MestReNova 12.0.4-22023 NMR spectrum processing program from *Mestrelab Research* was used to evaluate the NMR spectra and to determine the purity of the compounds through LC-MS. The purity of all compounds tested in biochemical assays was ≥95% as determined by LC-MS. NMR spectra and LC-MS traces for the tested compounds were included in the SI (Figures S17–22).

General method A: Molecular weight determination for trifluoroacetate salts. The molecular weights of trifluoroacetate salt compounds **5**–**8** (two benzofurans and two carbazoles) were determined using the corresponding LC system. Compounds **3d** and **4a** were used as reference for the respective compound class. First, the reference compound and the undetermined compound were mixed in a 1:1 ratio in methanol assuming the molecular weight of the free base. The percentual ratio of both compounds in the sample were evaluated by calculating the AUCs at 254 nm wavelength. The molecular weight of the tested compound was recalculated with the

determined ratio. For confirmation, a new mixed 1:1 sample was prepared including the determined molecular weight. The ratio of the new determined AUCs at 254 nm had to be between 0.8 and 1.2. Different absorption properties of the reference compounds and the tested compounds at 254 nm were neglected. The aliphatic amine tail was assumed to not influence absorption at 254 nm wavelength. The hydrochloride salts of the final compounds **3** and **4** were evaluated, too, for further validation of this method and could confirm their molecular weights.

General method B: Amide couplings. The desired Boc-protected amino acid (1.20 equiv) and benzotriazol-1-yloxy(dimethylamino)methylidene]-dimethylazanium tetrafluoroborate (TBTU) (3.00 equiv) were suspended in DCM (0.6 M) and cooled to 0 °C under inert gas atmosphere. *N,N*-Di*iso*propylethylamine (DIPEA) (4.00 equiv) was added dropwise and the mixture was stirred for 1 h. The carbazole or benzofuran amine (1.00 equiv) was added portionwise and the mixture was allowed to reach room temperature and was stirred for 23 h. The reaction was quenched with saturated sodium bicarbonate solution. The aqueous layer was extracted with DCM unless otherwise stated. Combined organic layers were washed with aqueous saturated sodium chloride solution, dried over sodium sulfate, filtered, and concentrated under reduced pressure. Crude product was purified *via* column chromatography with either a cyclohexane-ethyl acetate or DCM-methanol system.

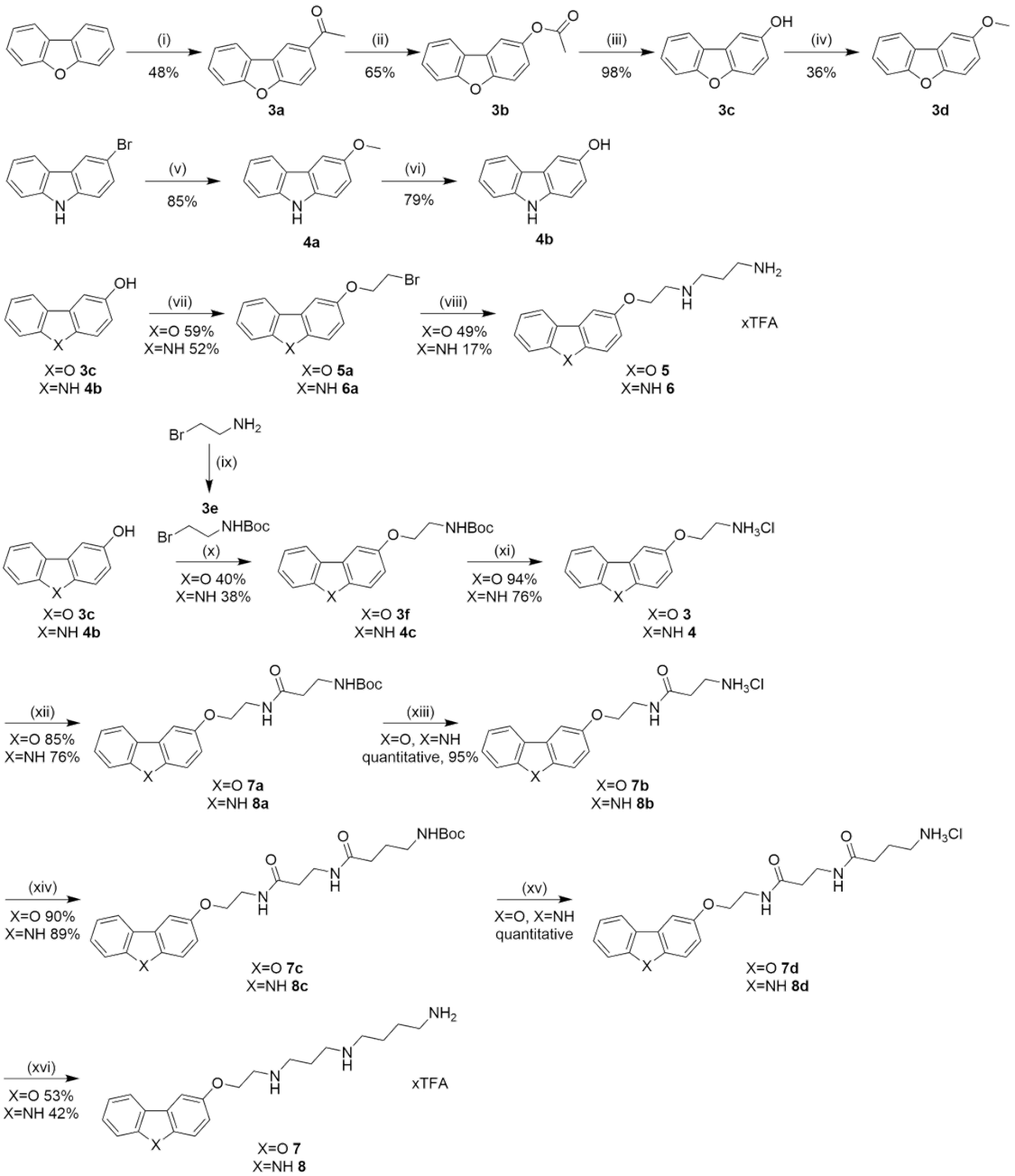
General method C: Boc deprotection. Deprotection was performed according to Han *et al.* reaction conditions.[318] The Boc protected compound (1.00 equiv) was dissolved in 1,4-dioxane, treated with 4 M HCl in 1,4-dioxane (20.0 eq) and stirred overnight under inert gas atmosphere. Ammonium salt precipitation was completed with cold petroleum ether and product isolation was done via vacuum filtration unless otherwise stated.

Synthesis SAH analogs 1 and 2

(*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)amino)butanoic acid trifluoroacetate salt (1). Complete synthesis and analytical data of SAH analog 1 was reported by Schwickert *et al*.[100]. Same batch was used. 99% purity.

(2*R*,3*R*,4*S*,5*R*)-2-(6-Amino-9*H*-purin-9-yl)-5-(((3-aminopropyl)-amino)methyl)tetrahydrofuran-3,4-diol trifluoroacetate salt (2). Complete synthesis and analytical data of SAH analog **2** was reported by Schwickert *et al*.[100]. Same batch was used. 99% purity.

Scheme 1. Synthesis of compounds **3**–**8**.



Reagents and conditions: (i) acetyl chloride, $AlCl_3$, chloroform, rt, 4 h; (ii) 2,2,2-trifluoroacetic acid, 3-chloroperoxybenzoic acid, DCM, 0 °C→rt, 3 d; (iii) sodium methanolate, methanol, 0 °C→rt, 1 h; (iv) NaH, THF, 0 °C, 10 min, then iodomethane, rt, 1 h; (v) sodium methanolate, CuI, DMF, rt→120 °C; (vi) 47% HBr, acetic acid glacial, reflux, 6 h; (vii) 1,2-dibromoethane, K_2CO_3, acetonitrile, 80 °C, 22–24 h; (viii) propane-1,3-diamine, acetonitrile, 80 °C, 3 h; (ix) Boc_2O, triethylamine, DCM, 0 °C→rt, 2 h; (x) K_2CO_3, acetonitrile, 80 °C, 18–20 h; (xi) 4 M HCl in 1,4-dioxane, rt, 3–20 h; (xii) 3-(*tert*-butoxycarbonylamino)propionic acid, TBTU, DIPEA, DCM, 0 °C→rt, 23 h; (xiii) 4 M HCl in 1,4-dioxane, rt, 1.5–20 h; (xiv) 4-(*tert*-butoxycarbonylamino)butyric acid, TBTU, DIPEA, DCM, 0 °C→rt, 23 h; (xv) 4 M HCl in 1,4-dioxane, rt, 1–20 h; (xvi) 1 M THF-BH_3, THF, 0 °C→60 °C, 3 d.

Synthesis preQ$_1$-related compounds 3–8

1-Dibenzofuran-2-ylethanone (3a). Synthesis was performed according to Chiranjeevi *et al.*[101] instructions. A solution of acetyl chloride (1.12 g, 1.02 mL, 14.3 mmol, 1.20 equiv) and anhydrous trichloroalumane (1.90 g, 14.2 mmol, 1.20 equiv) in chloroform (20.0 mL) was added dropwise to a stirred solution of dibenzofuran (2.00 g, 11.9 mmol, 1.00 equiv) in chloroform (20.0 mL) under inert gas atmosphere. The reaction mixture was stirred at room temperature for 4 h. The reaction mixture was poured over ice and the precipitated aluminum hydroxide was dissolved with concentrated hydrochloric acid. The aqueous layer was extracted with chloroform (2x 20.0 mL). The combined organic layers were dried with sodium sulfate, filtered and concentrated under reduced pressure. The crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 5:1). Target compound **3a** (1.19 g, 5.66 mmol, 48% yield) was isolated as a colorless solid. Mp: 77–79 °C. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 8.61 – 8.54 (m, 1H), 8.09 (dd, *J* = 8.6, 1.8 Hz, 1H), 8.03 – 7.94 (m, 1H), 7.62 – 7.54 (m, 2H), 7.54 – 7.45 (m, 1H), 7.42 – 7.33 (m, 1H), 2.70 (s, 3H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 197.4, 159.0, 157.0, 132.6, 128.1, 124.7, 123.9, 123.5, 121.7, 121.0, 112.1, 111.7, 26.9. **MS (ESI)**: found: m/z = 211.0 [M+H$^+$], calculated: m/z = 211.1 [M+H$^+$].

Dibenzofuran-2-yl acetate (3b). Synthesis was performed according to Chiranjeevi *et al.*[101] instructions. A solution of 1-dibenzofuran-2-ylethanone (**3a**) (1.60 g, 7.61 mmol, 1.00 equiv) in dry DCM (25.0 mL) was cooled to 0 °C under inert gas atmosphere. 2,2,2-trifluoroacetic acid (3.55 g, 2.40 mL, 31.2 mmol, 4.09 equiv) was added dropwise, followed by a solution of 3-chloroperoxybenzoic acid (3.41 g, 15.2 mmol, 2.00 equiv) in DCM (13.0 mL), which was added carefully. After addition, the mixture was allowed to reach room temperature and was stirred for 3 d. The reaction was quenched with aqueous $FeSO_4$ heptahydrate (around 2.5 g) solution and basified with saturated sodium bicarbonate solution under gas evolution. Emerging solid was filtered over Celite® and the filtrate was washed with saturated sodium bicarbonate solution. The aqueous layer was extracted with DCM (2x 25.0 mL) and the combined organic layers were washed with saturated sodium chloride solution, dried with sodium sulfate,

filtered and concentrated under reduced pressure. The dark brown crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 7:1). The content of isomeric 1-(dibenzo[b,d]furan-3-yl)ethan-1-one was reduced by recrystallization from methanol (15 mL) from 10% to 3% isomeric 1-(dibenzo[b,d]furan-3-yl)ethan-1-one. The target compound **3b** (1.16 g, 4.95 mmol, 65% yield) was isolated as beige solid. Mp: 113–115 °C. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 7.95 – 7.87 (m, 1H), 7.72 – 7.67 (m, 1H), 7.61 – 7.53 (m, 2H), 7.52 – 7.43 (m, 1H), 7.39 – 7.30 (m, 1H), 7.17 (dd, *J* = 8.8, 2.4 Hz, 1H), 2.36 (s, 3H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 170.1, 157.1, 153.8, 146.3, 127.8, 125.1, 124.1, 122.9, 121.0, 120.8, 113.7, 112.2, 112.0, 21.3. **MS (ESI)**: found: m/z = 184.9 [M-Acetyl+H^+], calculated: m/z = 227.1 [M+H^+], m/z = 185.1 [M-Acetyl+H^+].

2-Dibenzofuranol (3c). Synthesis was performed according to Chiranjeevi *et al*.[101] instructions. A suspension of dibenzofuran-2-yl acetate (**3b**) (1.10 g, 4.86 mmol, 1.00 equiv) in methanol (19.0 mL) was added dropwise to a suspension of freshly prepared sodium methanolate in methanol (559 mg elementary sodium in 12.0 mL methanol) under ice cooling under inert gas atmosphere. The mixture was allowed to reach room temperature. After 1 h, the reaction was quenched with 2 M HCl solution (8.00 mL). Methanol was evaporated under reduced pressure and the remaining residue was taken in water. The aqueous layer was acidified with concentrated HCl to protonate the product. Then it was extracted with chloroform (3x) and the organic layer was washed with saturated sodium chloride solution, dried with sodium sulfate, filtered and concentrated under reduced pressure. The target compound **3c** (876 mg, 4.75 mmol, 98% yield) was isolated as pale-yellow solid. Mp: 130–132 °C. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 7.91 – 7.84 (m, 1H), 7.55 (dt, *J* = 8.3, 0.9 Hz, 1H), 7.49 – 7.40 (m, 2H), 7.40 – 7.35 (m, 1H), 7.35 – 7.28 (m, 1H), 6.97 (dd, *J* = 8.8, 2.6 Hz, 1H), 4.95 (s, 1H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 127.4, 125.2, 124.3, 122.6, 120.8, 115.4, 112.3, 111.9, 106.4. **MS (ESI):** found: m/z = 184.9 [M+H^+], calculated: m/z = 185.1 [M+H^+].

2-Methoxydibenzofuran (3d). Synthesis was performed according to Yempala *et al*.[102] instructions. Sodium hydride (15.6 mg, 391 µmol, 1.20 equiv) was added to a solution of 2-dibenzofuranol (**3c**) (60.0 mg, 326 µmol, 1.00 equiv) in dry THF (1 mL)

under ice cooling and inert gas atmosphere and stirred for 10 min. After, iodomethane (55.5 mg, 24.3 µL, 391 µmol, 1.20 equiv) was added and was allowed to reach room temperature and was stirred for 1 h. Further 0.60 equiv iodomethane were added and stirred for 1 h. The reaction was quenched with 5 mL water and the aqueous layer was extracted with DCM (3x 15.0 mL). Combined organic layers were washed with aqueous saturated sodium chloride solution, dried over sodium sulfate, filtered, and concentrated under reduced pressure. The crude product was purified *via* flash chromatography (cyclohexane:ethyl acetate; gradient 0% to 100% ethyl acetate) and was applied as dry load. The target compound **3d** (23.7 mg, 118 µmol, 36% yield) was isolated as pink solid. Mp: 47–49 °C. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 7.96 – 7.88 (m, 1H), 7.56 (dt, *J* = 8.3, 1.0 Hz, 1H), 7.51 – 7.39 (m, 3H), 7.33 (td, *J* = 7.5, 1.1 Hz, 1H), 7.06 (dd, *J* = 8.9, 2.6 Hz, 1H), 3.92 (s, 3H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 157.1, 156.0, 151.0, 127.2, 124.8, 124.6, 122.5, 120.7, 115.3, 112.2, 111.9, 103.9, 56.2. **MS (ESI)**: found: m/z = 199.0 [M+H$^+$], calculated: m/z = 199.1 [M+H$^+$].

3-Methoxy-9*H*-carbazole (4a). Synthesis was performed according to Ku *et al.*[103] instructions. A solution of 3-bromo-9*H*-carbazole (2.00 g, 8.13 mmol, 1.00 equiv) in dry DMF (11.0 mL) was added to a suspension of freshly prepared sodium methanolate (3.10 g elementary sodium in 20.0 mL methanol; excess of methanol was removed under reduced pressure) in dry DMF (20.0 mL). Copper(I)iodide (3.10 g, 16.3 mmol, 2.00 equiv) was added and the mixture was stirred for 15 min at room temperature while several color changes were observed. Afterwards, the dark green mixture was heated to 120 °C and refluxed for 22 h under inert gas atmosphere. The reaction mixture was filtered over silica gel, which was rinsed several times with ethyl acetate (in total 150 mL). The organic phase was washed with water and the resulting precipitate was filtered off over silica gel. The filtrate was washed with saturated sodium chloride solution (3x 30.0 mL, 5x 10.0 mL) and the organic layer was dried with sodium sulfate, filtered and concentrated under reduced pressure. The brown crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 5:1). The target compound **4a** (1.36 g, 6.88 mmol, 85% yield) was isolated as beige solid. Mp: 145–147 °C. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 8.07 (dt, *J* = 7.8, 1.0 Hz, 1H), 7.88 (s, 1H), 7.60 (d, *J* = 2.5 Hz, 1H), 7.49 – 7.36 (m, 2H), 7.32 (dd, *J* = 8.8, 0.6 Hz, 1H), 7.29 – 7.21

(m, 1H), 7.10 (dd, *J* = 8.8, 2.5 Hz, 1H), 3.96 (s, 3H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 154.0, 140.4, 134.5, 125.9, 123.9, 123.5, 120.4, 119.2, 115.2, 111.4, 110.9, 103.3, 56.2. **MS (ESI):** found: m/z = 198.0 [M+H$^+$], calculated: m/z = 198.1 [M+H$^+$].

9*H*-Carbazol-3-ol (4b). Synthesis was performed according to Milne *et al*[104]*.* instructions. A mixture of 3-methoxy-9*H*-carbazole (**4a**) (1.30 g, 6.59 mmol, 1.00 equiv), 47% hydrobromic acid (2.90 mL) and acetic acid glacial (24.0 mL) was heated to reflux (at 135 °C) under inert gas atmosphere. The progress of the reaction was monitored by TLC every hour. After 6 h, the reaction mixture was cooled to room temperature and acetone (20.0 mL) and water (10.0 mL) were added. The aqueous layer was extracted with DCM (1x 20.0 mL, 4x 10.0 mL) and concentrated under reduced pressure. The residue was taken in a water/ acetonitrile mixture and lyophilized overnight. The crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 5:1 to 2:1). Traces of starting material were removed by washing the solid with chloroform (3x 3.00 mL). The target compound **4b** (956 mg, 5.22 mmol, 79% yield) was isolated as pale green solid. Mp: > 255 °C under decomposition. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 10.87 (s, 1H), 8.92 (s, 1H), 7.98 (d, *J* = 7.8 Hz, 1H), 7.47 – 7.36 (m, 2H), 7.36 – 7.26 (m, 2H), 7.15 – 7.00 (m, 1H), 6.91 (dd, *J* = 8.6, 2.4 Hz, 1H). **^{13}C-NMR** (75 MHz, DMSO-d_6): δ = 150.4, 140.4, 133.8, 125.2, 123.1, 122.3, 120.1, 117.7, 115.0, 111.3, 110.8, 104.9. **MS (ESI):** found: m/z = 183.8 [M+H$^+$], calculated: m/z = 184.1 [M+H$^+$].

Synthesis of **5a** and **6a** was performed according to Pájaro *et al*.[105] reaction conditions. 2-Dibenzofuranol (**3c**) or 9*H*-carbazol-3-ol (**4b**) (100 mg, 1.00 equiv) was added portionwise to a suspension of 1,2-dibromoethane (5.00 equiv) and anhydrous potassium carbonate (3.00 equiv) in dry acetonitrile (2 mL) heated at 80 °C under inert gas atmosphere. After addition, the mixture was heated to 80 °C and stirred for 22–24 h. The mixture was filtered and the filtrate was concentrated under reduced pressure. Crude product was purified *via* column chromatography (cyclohexane:ethyl acetate).

2-(2-Bromoethoxy)dibenzofuran (5a). Crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 20:1) yielding **5a** (96.4 mg, 318 µmol, 59% yield) as colorless solid. Mp: 69–71 °C. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 7.94 – 7.87 (m, 1H), 7.56 (dt, *J* = 8.3, 0.9 Hz, 1H), 7.50 – 7.42 (m, 3H), 7.37 – 7.30

(m, 1H), 7.07 (dd, *J* = 8.9, 2.7 Hz, 1H), 4.38 (t, *J* = 6.3 Hz, 2H), 3.69 (t, *J* = 6.3 Hz, 2H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 157.1, 154.4, 151.5, 127.4, 124.9, 124.4, 122.7, 120.7, 116.0, 112.4, 111.9, 105.7, 69.2, 29.4. **MS (ESI):** found: not ionizable, calculated: m/z = 291.0 [M+H$^+$].

3-(2-Bromoethoxy)-9*H*-carbazole (6a). Crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 4:1) yielding **6a** (81.9 mg, 282 µmol, 52% yield) as colorless solid. Mp: 134–136 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 11.08 (s, 1H), 8.15 – 8.05 (m, 1H), 7.74 (d, *J* = 2.5 Hz, 1H), 7.53 – 7.31 (m, 3H), 7.17 – 7.04 (m, 2H), 4.40 (t, *J* = 5.5 Hz, 2H), 3.84 (t, *J* = 5.5 Hz, 2H). **^{13}C-NMR** (75 MHz, DMSO- d_6): δ = 151.5, 140.5, 134.9, 125.5, 122.8, 122.4, 120.3, 118.1, 115.3, 111.6, 111.0, 104.8, 68.7, 31.8. **MS (ESI):** found: m/z = 289.8/291.8 [M+H$^+$], calculated: m/z = 290.0/292.0 [M+H$^+$].

Synthesis of **5** and **6** was performed according to Zhang *et al*.[106] reaction conditions. A solution of propane-1,3-diamine (10.0 equiv) in acetonitrile (3 M) was prepared and heated to 80 °C. Then a solution of 2-(2-bromoethoxy)dibenzofuran (**5a**) or 3-(2-bromoethoxy)-9H-carbazole (**6a**) (1.00 equiv) in acetonitrile (1 M) was added dropwise under inert gas atmosphere. After addition, the mixture was stirred at that temperature for 3 h. The reaction mixture was concentrated under reduced pressure. The crude product was purified *via* preparative HPLC (gradient 10–100% acetonitrile in water+0.1% TFA, eluates at 40-50%). The molecular weight was determined according to general method A.

***N*^{1}-(2-(Dibenzo[b,d]furan-2-yloxy)ethyl)propane-1,3-diamine trifluoroacetate salt (5). 5a** (35.0 mg, 120 µmol, 1.00 equiv) yielded the target compound **5** as colorless solid (600.4 g/mol (~3xTFA), 40.4 mg, 0.067 mmol, 49.0% yield). Mp: 179–181 °C. **^{1}H- NMR** (300 MHz, DMSO-d_6): δ = 8.10 (d, *J* = 7.7 Hz, 1H), 7.76 (s, 1H), 7.64 (t, *J* = 7.9 Hz, 2H), 7.50 (t, *J* = 7.8 Hz, 1H), 7.38 (t, *J* = 7.6 Hz, 1H), 7.17 (d, *J* = 8.9 Hz, 1H), 4.43 – 4.28 (m, 2H), 3.43 – 3.39 (m, 2H), 3.23 – 3.07 (m, 2H), 3.02 – 2.85 (m, 2H), 2.12 – 1.90 (m, 2H). **^{13}C-NMR** (75 MHz, DMSO-d_6): δ = 156.4, 154.3, 150.6, 127.9, 124.4, 123.9, 123.1, 121.3, 116.3, 112.5, 111.9, 105.7, 64.4, 46.4, 44.6, 36.4, 23.9. **MS (ESI):** found: m/z = 285.1 [M+H$^+$], calculated: m/z = 285.2 [M+H$^+$]. 99% purity.

***N*^{1}-(2-((9*H*-Carbazol-3-yl)oxy)ethyl)propane-1,3-diamine trifluoroacetate salt (6).** **6a** (76.0 mg, 262 µmol, 1.00 equiv) yielded the target compound **6** as colorless solid (855.8 g/mol (~5x TFA), 38.5 mg, 0.045 mmol, 17% yield). Mp: >194 °C under decomposition**.** **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 11.20 (s, 1H), 8.08 (d, *J* = 7.7 Hz, 1H), 7.75 (d, *J* = 2.4 Hz, 1H), 7.45 (t, *J* = 8.5 Hz, 2H), 7.36 (t, *J* = 7.5 Hz, 1H), 7.17 – 7.05 (m, 2H), 4.33 (t, *J* = 4.9 Hz, 2H), 3.59 – 3.26 (m, 2H), 3.15 (t, *J* = 7.6 Hz, 2H), 2.94 (t, *J* = 7.5 Hz, 2H), 2.01 (p, *J* = 7.5 Hz, 2H). **^{13}C- NMR** (75 MHz, DMSO-d_6): δ = 151.8, 140.9, 135.5, 126.0, 123.2, 122.8, 120.6, 118.5, 115.8, 112.1, 111.5, 105.1, 64.8, 46.7, 44.8, 36.7, 24.2. **MS (ESI):** found: m/z = 284.1 [M+H$^+$], calculated: m/z = 284.2 [M+H$^+$]. 100% purity.

***tert*-Butyl *N*-(2-bromoethyl) carbamate (3e).** A solution of di-*tert*-butyldicarbonate (3.20 g, 14.6 mmol, 1.00 equiv) in DCM (10.0 mL) was added dropwise to a solution of 2-bromoethylamine hydrobromide (3.00 g, 14.6 mmol, 1.00 equiv) triethylamine (1.48 g, 2.03 mL, 14.6 mmol, 1.00 equiv) in DCM (30.0 mL) under ice cooling and inert gas atmosphere. After addition, the mixture was allowed to reach room temperature and was stirred for 2 h. The organic phase was washed with 1 M HCl (3x 20.0 mL), saturated sodium bicarbonate solution (3x 20.0 mL), aqueous saturated sodium chloride solution (1x 20 mL), dried over sodium sulfate, filtered, and concentrated under reduced pressure. The crude product of **3e** (2.967 g) showed around 20% di-*tert*-butyldicarbonate contamination and was used without further purification.**^{1}H- NMR** (300 MHz, $CDCl_3$): δ = 4.99 (s, 1H), 3.57 – 3.46 (m, 2H), 3.46 – 3.38 (m, 2H), 1.43 (s, 9H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 155.7, 79.9, 42.5, 32.8, 28.4. **MS (ESI):** found: m/z = 167.7/169.8 [M-*tert*-butyl+H$^+$], calculated: m/z = 224.0/226.0 [M+H$^+$], m/z = 168.0/170.0 [M-*tert*-butyl+H$^+$].

Synthesis of **3f** and **4c** was performed according to Pájaro *et al*.[105] reaction conditions. 2-Dibenzofuranol (**3c**) or 9*H*-carbazol-3-ol (**4b**) (1.00 equiv), *tert*- butyl- *N*- (2- bromoethyl)carbamate (**3e**) (3.00 equiv, 80% purity) and anhydrous potassium carbonate (5.00 equiv) were mixed in dry acetonitrile (**3c/4b** 0.3 M). The mixture was heated at 80 °C and stirred for 18–20 h. The mixture was concentrated under reduced pressure. The residue was taken in water and was extracted with DCM. Combined organic layers were washed with water and aqueous saturated sodium chloride solution, dried over sodium sulfate, filtered, and concentrated under reduced

pressure. The crude product was purified *via* flash chromatography or column chromatography.

***tert*-Butyl (2-(dibenzo[b,d]furan-2-yloxy)ethyl)carbamate (3f). 3c** (750 mg, 3.26 mmol, 1.00 equiv) was used. Crude product was purified *via* flash chromatography (cyclohexane:ethyl acetate; gradient 5% to 12% ethyl acetate). Mixed product fractions with contamination of *tert*-butyl *N*-(2-bromoethyl)carbamate were purified *via* column chromatography (cyclohexane:ethyl acetate; 7:1) yielding target compound **3f** as pale-yellow oil (425 mg, 1.30 mmol, 40% yield). **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 7.79 – 7.72 (m, 1H), 7.45 – 7.27 (m, 3H), 7.24 (s, 1H), 7.23 – 7.12 (m, 1H), 6.89 (dd, *J* = 8.9, 2.6 Hz, 1H), 5.03 (s, 1H), 3.97 (t, *J* = 5.2 Hz, 2H), 3.47 (d, *J* = 5.7 Hz, 2H), 1.37 (s, 9H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 157.0, 156.0, 154.8, 151.1, 127.2, 124.8, 124.4, 122.5, 120.6, 115.6, 112.2, 111.8, 104.8, 79.6, 68.2, 40.4, 28.5. **MS (ESI)**: found: m/z = 350.1 [M+Na$^+$], calculated: m/z = 328.2 [M+H$^+$], m/z = 350.1 [M+Na$^+$].

***tert*-Butyl (2-((9*H*-carbazol-3-yl)oxy)ethyl)carbamate (4c). 4b** (642 mg, 2.80 mmol, 1.00 equiv) was used. The crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 3:1). The target compound **4c** (350 mg, 1.07 mmol, 38% yield) was isolated as yellow solid. Mp: 154–158 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 11.01 (s, 1H), 8.13 – 8.03 (m, 1H), 7.69 (d, *J* = 2.5 Hz, 1H), 7.47 – 7.30 (m, 3H), 7.14 – 7.07 (m, 1H), 7.03 (dd, *J* = 8.7, 2.5 Hz, 1H), 4.05 (t, *J* = 5.9 Hz, 2H), 3.40 – 3.31 (m, 2H), 1.40 (s, 9H). **^{13}C-NMR** (75 MHz, DMSO-d_6): δ = 155.7, 152.0, 140.4, 134.7, 125.4, 122.8, 122.4, 120.2, 118.0, 115.3, 111.5, 110.9, 104.2, 77.7, 67.3, 39.7, 28.2. **MS (ESI)**: found: m/z = 227.0 [M-Boc+H$^+$], calculated: m/z = 327.2 [M+H$^+$], m/z = 227.1 [M-Boc+H$^+$].

2-Dibenzofuran-2-yloxyethylamine hydrochloride (3). Synthesis was performed according to general method C. **3f** (1.73 g, 5.29 mmol, 1.00 equiv) was used. The target compound **3** (1.31 g, 4.97 mmol, 94% yield) was isolated as colorless solid. Mp: > 260 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 8.45 (s, 3H), 8.15 (d, *J* = 7.6 Hz, 1H), 7.81 (d, *J* = 2.6 Hz, 1H), 7.65 (t, *J* = 8.5 Hz, 2H), 7.51 (t, *J* = 7.7 Hz, 1H), 7.38 (t, *J* = 7.4 Hz, 1H), 7.17 (dd, *J* = 9.0, 2.7 Hz, 1H), 4.32 (t, *J* = 5.2 Hz, 2H), 3.25 (s, 2H). **^{13}C-NMR** (75 MHz, DMSO-d_6): δ = 156.2, 154.3, 150.3, 127.6, 124.2, 123.8, 122.9,

121.3, 116.3, 112.3, 111.7, 105.5, 65.2, 38.3. **MS (ESI)**: found: m/z = 228.0 [M+H$^+$], calculated: m/z = 228.1 [M+H$^+$]. 100% purity.

2-((9*H*-Carbazol-3-yl)oxy)ethan-1-amine hydrochloride (4). Synthesis was performed according to general method C. **4c** (1.02 g, 3.13 mmol, 1.00 equiv) was used. The target compound **4** (629 mg, 2.39 mmol, 76% yield) was isolated as colorless solid. Mp: > 260 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 11.21 (s, 1H), 8.42 (s, 3H), 8.09 (d, *J* = 7.8 Hz, 1H), 7.75 (d, *J* = 2.4 Hz, 1H), 7.50–7.38 (m, 2H), 7.41 – 7.29 (m, 1H), 7.17–7.05 (m, 2H), 4.29 (t, *J* = 5.2 Hz, 2H), 3.24 (q, *J* = 5.3 Hz, 2H). **^{13}C- NMR** (75 MHz, DMSO-d_6): δ = 151.4, 140.4, 135.0, 125.5, 122.7, 122.4, 120.3, 118.1, 115.5, 111.7, 111.1, 104.6, 65.2, 38.5. **MS (ESI)**: found: m/z = 227.0 [M+H$^+$], calculated: m/z = 227.1 [M+H$^+$]. 100% purity.

***tert*-butyl (3-((2-(dibenzo[b,d]furan-2-yloxy)ethyl)amino)-3-oxopropyl)carbamate (7a).** Synthesis was performed according to general method B. **3** (377 mg, 1.43 mmol, 1.00 equiv) was used. Crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 1:2). The target compound **7a** (485 mg, 1.22 mmol, 85% yield) was isolated as colorless solid. Mp: 117–119 °C. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ = 7.92 – 7.85 (m, 1H), 7.57 – 7.38 (m, 4H), 7.31 (td, *J* = 7.5, 1.1 Hz, 1H), 7.01 (dd, *J* = 8.9, 2.6 Hz, 1H), 6.33 (s, 1H), 5.18 (s, 1H), 4.12 (t, *J* = 5.1 Hz, 2H), 3.70 (q, *J* = 5.3 Hz, 2H), 3.42 (t, *J* = 6.1 Hz, 2H), 2.52 – 2.38 (m, 2H), 1.41 (s, 9H). **^{13}C- NMR** (75 MHz, $CDCl_3$): δ = 171.8, 157.0, 156.2, 154.7, 151.2, 127.4, 124.9, 124.4, 122.6, 120.7, 115.6, 112.3, 111.9, 105.0, 79.5, 67.7, 39.2, 36.9, 36.4, 28.5. **MS (ESI)**: found: m/z = 299.1 [M-Boc+H$^+$], calculated: m/z = 399.2 [M+H$^+$], m/z = 299.1 [M- Boc+H$^+$].

***tert*-Butyl (3-((2-((9*H*-carbazol-3-yl)oxy)ethyl)amino)-3-oxopropyl)carbamate (8a).** Synthesis was performed according to general method B. **6** (319 mg, 1.21 mmol, 1.00 equiv) was used. Crude product was purified *via* column chromatography (cyclohexane:ethyl acetate; 1:2). The target compound **8a** (430 mg, 920 µmol, 76% yield) was isolated as colorless solid. Mp: 149–152 °C.**^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 11.02 (s, 1H), 8.15 (t, *J* = 5.5 Hz, 1H), 8.08 (d, *J* = 7.8 Hz, 1H), 7.70 (d, *J* = 2.5 Hz, 1H), 7.48 – 7.29 (m, 3H), 7.16 – 6.99 (m, 2H), 6.74 (t, *J* = 5.8 Hz, 1H), 4.07 (t, *J* = 5.8 Hz, 2H), 3.48 (q, *J* = 5.7 Hz, 2H), 3.24 – 3.10 (m, 2H), 2.30 (t, *J* = 7.3 Hz, 2H), 1.37 (s, 9H). **^{13}C-NMR** (75 MHz, DMSO-d_6): δ = 170.7, 155.5, 152.0, 140.4, 134.7,

125.4, 122.8, 122.4, 120.3, 118.0, 115.3, 111.6, 111.0, 104.3, 77.6, 67.2, 38.4, 36.8, 35.8, 28.2. **MS (ESI)**: found: m/z = 298.1 [M-Boc+H$^+$], calculated: m/z = 398.2 [M+H$^+$], m/z = 298.2 [M-Boc+H$^+$].

3-Amino-*N*-(2-(dibenzo[b,d]furan-2-yloxy)ethyl)propenamide hydrochloride (7b). Synthesis was performed according to general method C. **7a** (453 mg, 1.14 mmol, 1.00 equiv) was used and the reaction was stirred for 1.5 h. The solvent was removed under reduced pressure and the residue was lyophilized in water. The target compound **7b** (379 mg, 1.13 mmol, quantitative) was isolated as colorless solid. Mp: 171–173 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 8.54 (t, *J* = 5.5 Hz, 1H), 8.22 – 8.07 (m, 4H), 7.75 (d, *J* = 2.7 Hz, 1H), 7.66 – 7.55 (m, 2H), 7.53 – 7.44 (m, 1H), 7.36 (td, *J* = 7.5, 1.0 Hz, 1H), 7.10 (dd, *J* = 8.9, 2.7 Hz, 1H), 4.11 (t, *J* = 5.7 Hz, 2H), 3.51 (q, *J* = 5.7 Hz, 2H), 3.00 (t, *J* = 7.1 Hz, 2H), 2.59 (t, *J* = 7.1 Hz, 2H). **^{13}C- NMR** (75 MHz, DMSO-d_6): δ = 169.8, 156.2, 154.8, 150.1, 127.5, 124.2, 123.9, 122.8, 121.3, 116.0, 112.2, 111.7, 105.3, 67.1, 39.0, 38.4, 35.3, 32.1. **MS (ESI)**: found: m/z = 299.0 [M+H$^+$], calculated: m/z = 298.7 [M+H$^+$].

***N*-(2-((9*H*-Carbazol-3-yl)oxy)ethyl)-3-aminopropanamide hydrochloride (8b).** Synthesis was performed according to general method C. **8a** (400 mg, 1.01 mmol, 1.00 equiv) was used. The solvent was removed under reduced pressure and the residue was lyophilized in water. The target compound **8b** (318 mg, 953 µmol, 95% yield) was isolated as colorless solid. Mp: 175–177 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 11.16 (s, 1H), 8.51 (t, *J* = 5.6 Hz, 1H), 8.20 – 8.02 (m, 4H), 7.70 (d, *J* = 2.4 Hz, 1H), 7.49 – 7.29 (m, 3H), 7.14 – 6.99 (m, 2H), 4.09 (t, *J* = 5.7 Hz, 2H), 3.51 (q, *J* = 5.7 Hz, 2H), 3.01 (q, *J* = 6.3 Hz, 2H), 2.59 (t, *J* = 7.1 Hz, 2H). **^{13}C- NMR** (75 MHz, DMSO-d_6): δ = 169.7, 151.9, 140.4, 134.7, 125.4, 122.8, 122.4, 120.3, 118.0, 115.3, 111.6, 111.0, 104.2, 67.1, 38.5, 35.3, 32.1. **MS (ESI)**: found: m/z = 298.1 [M+H$^+$], calculated: m/z = 298.2 [M+H$^+$].

***tert*-Butyl (4-((3-((2-(dibenzo[b,d]furan-2-yloxy)ethyl)amino)-3-oxopropyl)amino)-4-oxobutyl)carbamate (7c).** Synthesis was performed according to general method B. **7b** (330 mg, 986 µmol, 1.00 equiv) was used. Crude product was purified *via* column chromatography (DCM:methanol; 20:1). The target compound **7c** (428 mg, 885 µmol, 90% yield) was isolated as colorless solid. Mp: 119–121 °C. **^{1}H- NMR** (300 MHz, $CDCl_3$): δ = 7.90 – 7.84 (m, 1H), 7.54 – 7.36 (m, 4H), 7.33 – 7.25

(m, 1H), 6.99 (dd, *J* = 8.9, 2.6 Hz, 1H), 6.95 – 6.72 (m, 2H), 4.91 (s, 1H), 4.10 (t, *J* = 5.2 Hz, 2H), 3.67 (q, *J* = 5.4 Hz, 2H), 3.52 (q, *J* = 5.9 Hz, 2H), 3.07 (t, *J* = 6.7 Hz, 2H), 2.47 (t, *J* = 6.1 Hz, 2H), 2.16 (t, *J* = 7.1 Hz, 2H), 1.74 (t, *J* = 6.9 Hz, 2H), 1.40 (s, 9H). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ = 173.1, 172.0, 157.0, 156.4, 154.7, 151.2, 127.4, 124.9, 124.3, 122.6, 120.6, 115.6, 112.3, 111.8, 105.0, 79.4, 67.6, 39.2, 35.8, 33.7, 28.5, 26.2. **MS (ESI)**: found: m/z = 484.1 [M+H$^+$], calculated: m/z = 484.2 [M+H$^+$].

***tert*-Butyl (4-((3-((2-((9*H*-carbazol-3-yl)oxy)ethyl)amino)-3-oxopropyl)amino)-4-oxobutyl)carbamate (8c).** Synthesis was performed according to general method B. **8b** (279 mg, 836 µmol, 1.00 equiv) was used. The aqueous layer was extracted with DCM (1x 20.0 mL). Next, a mixture of chloroform and *iso*propanol (2:1) was used to extract the aqueous phase (3x 20.0 mL). Combined organic layers were washed with aqueous saturated sodium chloride solution, dried over sodium sulfate, filtered, and concentrated under reduced pressure. Crude product was purified *via* column chromatography (DCM:methanol; 15:1). The target compound **8c** (359 mg, 744 µmol, 89% yield) was isolated as colorless/pale yellow solid. Mp: 160–163 °C. **^{1}H- NMR** (300 MHz, DMSO-d_6): δ = 11.02 (s, 1H), 8.16 (t, *J* = 5.6 Hz, 1H), 8.08 (d, *J* = 7.8 Hz, 1H), 7.84 (t, *J* = 5.7 Hz, 1H), 7.70 (d, *J* = 2.4 Hz, 1H), 7.48 – 7.29 (m, 3H), 7.15 – 7.00 (m, 2H), 6.77 (t, *J* = 5.7 Hz, 1H), 4.08 (t, *J* = 5.8 Hz, 2H), 3.49 (q, *J* = 5.7 Hz, 2H), 3.34 – 3.21 (m, 2H), 2.89 (q, *J* = 6.6 Hz, 2H), 2.31 (t, *J* = 7.0 Hz, 2H), 2.04 (t, *J* = 7.5 Hz, 2H), 1.58 (p, *J* = 7.3 Hz, 2H), 1.36 (s, 9H). **^{13}C- NMR** (75 MHz, DMSO-d_6): δ = 171.9, 170.8, 155.6, 152.0, 140.4, 134.7, 125.4, 122.8, 122.5, 120.3, 118.0, 115.3, 111.6, 111.0, 104.3, 77.4, 67.2, 39.5, 38.5, 35.4, 35.3, 32.9, 28.3, 25.9. **MS (ESI)**: found: m/z = 483.1 [M+H$^+$], calculated: m/z = 483.3 [M+H$^+$].

4-Amino-*N*-(3-((2-(dibenzo[b,d]furan-2-yloxy)ethyl)amino)-3-oxopropyl)butanamide hydrochloride (7d). Synthesis was performed according to general method C. **7c** (100 mg, 207 µmol, 1.00 equiv) was used and the reaction was stirred for 1 h. Solvent was concentrated under reduced pressure and the residue was lyophilized in water. The target compound **7d** (86.0 mg, 205 µmol, quantitative) was isolated as colorless solid. Mp: 206–209 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 8.28 (t, *J* = 5.6 Hz, 1H), 8.19 – 8.04 (m, 5H), 7.75 (d, *J* = 2.6 Hz, 1H), 7.61

(dd, *J* = 16.3, 8.6 Hz, 2H), 7.54 – 7.43 (m, 1H), 7.36 (t, *J* = 7.5 Hz, 1H), 7.10 (dd, *J* = 9.0, 2.7 Hz, 1H), 4.10 (t, *J* = 5.8 Hz, 2H), 3.48 (q, *J* = 5.7 Hz, 2H), 3.28 (q, *J* = 6.7 Hz, 2H), 2.75 (t, *J* = 7.5 Hz, 2H), 2.32 (t, *J* = 7.1 Hz, 2H), 2.18 (t, *J* = 7.2 Hz, 2H), 1.78 (p, *J* = 7.3 Hz, 2H). **^{13}C-NMR** (75 MHz, DMSO-d_6): δ = 171.3, 170.8, 156.1, 154.8, 150.1, 127.5, 124.2, 123.9, 122.8, 121.2, 116.0, 112.1, 111.6, 105.3, 67.2, 38.4, 38.3, 35.4, 35.3, 32.1, 23.2. **MS (ESI)**: found: m/z = 384.1 [M+H$^+$], calculated: m/z = 384.2 [M+H$^+$].

***N*-(3-((2-((9*H*-Carbazol-3-yl)oxy)ethyl)amino)-3-oxopropyl)-4-aminobutanamide hydrochloride (8d).** Synthesis was performed according to general method C. **8c** (320 mg, 663 µmol, 1.00 equiv) was used. Solvent was concentrated under reduced pressure and the residue was lyophilized in water. The target compound **8d** (275 mg, 656 µmol, quantitative) was isolated as colorless solid. Mp: 240–243 °C. **^{1}H- NMR** (300 MHz, DMSO-d_6): δ = 11.13 (s, 1H), 8.26 (s, 1H), 8.20 – 7.97 (m, 5H), 7.70 (d, *J* = 2.4 Hz, 1H), 7.49 – 7.26 (m, 3H), 7.17 – 6.97 (m, 2H), 4.07 (t, *J* = 5.8 Hz, 2H), 3.48 (q, *J* = 6.0 Hz, 2H), 3.28 (q, *J* = 6.7 Hz, 2H), 2.85 – 2.64 (m, 2H), 2.32 (t, *J* = 7.1 Hz, 2H), 2.18 (t, *J* = 7.2 Hz, 2H), 1.79 (p, *J* = 7.4 Hz, 2H). **^{13}C- NMR** (75 MHz, DMSO-d_6): δ = 171.3, 170.8, 152.0, 140.4, 134.7, 125.4, 122.8, 122.4, 120.3, 118.0, 115.3, 111.6, 111.0, 104.3, 67.2, 38.4, 35.4, 35.4, 32.1, 23.2. **MS (ESI)**: found: m/z = 383.1 [M+H$^+$], calculated: m/z = 383.2 [M+H$^+$].

Synthesis of **7** and **8** was performed according to Pirali *et al.*[107] reaction conditions. Tetrahydrofuran-BH_3 complex (1.00 M in THF, 20.0 equiv) was added dropwise to a solution of amide **7d** or **8d** (50.0 mg, 1.00 equiv) in dry THF (0.5 mL) under ice cooling and inert gas atmosphere. After addition, the mixture was let to reach room temperature and was stirred for 3 d at 60 °C. The reaction mixture was quenched with methanol (gas evolution) and the solvent was evaporated. The resulting borane-amine complex was dissolved differently.

***N*^{1}-(3-((2-(dibenzo[b,d]furan-2-yloxy)ethyl)amino)propyl)butane-1,4-diamine trifluoroacetate salt (7).** The mixture was taken in 2.4 mL water. 0.6 mL TFA was added (in total 20 vol%) and the mixture stirred for 30 min to destroy the borane-amine complex. The solution was lyophilized and the crude product was purified *via* reversed phase flash chromatography (acetonitrile:water 0.1% TFA; gradient 0%–25% acetonitrile). The product was further purified *via* preparative HPLC (gradient 10–100%

acetonitrile in water+0.1% TFA, eluates at 30%). The molecular weight was determined according to the described general method A. The target compound **7** (874.6 g/mol (~5 xTFA), 54.8 mg, 0.063 mmol, 53%) was isolated as colorless solid. Mp: 233–236 °C. **^{1}H-NMR** (300 MHz, CD_3OD): δ = 8.01 (dd, *J* = 7.7, 1.3 Hz, 1H), 7.67 (d, *J* = 2.6 Hz, 1H), 7.59 – 7.44 (m, 3H), 7.36 (td, *J* = 7.4, 1.2 Hz, 1H), 7.19 (dd, *J* = 8.9, 2.6 Hz, 1H), 4.42 (t, *J* = 5.1 Hz, 2H), 3.56 (t, *J* = 4.9 Hz, 2H), 3.36 – 3.25 (m, 2H), 3.25 – 3.15 (m, 2H), 3.11 (t, *J* = 7.4 Hz, 2H), 3.00 (t, *J* = 7.1 Hz, 2H), 2.33 – 2.15 (m, 2H), 1.91 – 1.70 (m, 4H). **^{13}C-NMR** (75 MHz, CD_3OD): δ = 158.4, 155.6, 152.8, 128.6, 126.1, 125.5, 123.8, 121.8, 117.0, 113.2, 112.6, 106.3, 65.3, 48.4, 48.3, 46.0, 45.8, 40.0, 25.6, 24.2, 24.0. **MS (ESI)**: found: m/z = 356.0 [M+H$^+$], calculated: m/z = 356.2 [M+H$^+$]. 98% purity.

N^1-(3-((2-((9*H*-Carbazol-3-yl)oxy)ethyl)amino)propyl)butane-1,4-diamine trifluoroacetate salt (8). The mixture was taken in 2.4 mL water. 0.6 mL TFA were added (in total 20 vol%) and the mixture was stirred in total for 6 h to destroy the amine-borane complexes, which was not completed. The mixture was lyophilized and taken in 2.2 mL water and 0.4 mL 10% sodium hydroxide solution. Stirring for 30 min resulted in gas evolution and completely dissolved amine-borane complexes. The mixture was lyophilized and the crude product was purified *via* flash chromatography reversed phase (acetonitrile:water 0.1% TFA; gradient 0%–25% acetonitrile). Crude product was further purified *via* HPLC (gradient 5–100% acetonitrile in water+0.1% TFA, eluates at 10–20%). The molecular weight was determined according to the described general method A. The target compound **8** (827.2 g/mol (ca. 5x TFA), 41.0 mg, 0.050 mmol, 42% yield) was isolated as colorless solid (pale brown in solution). Mp: 211–213 °C. **^{1}H-NMR** (300 MHz, DMSO-d_6): δ = 11.18 (s, 1H), 8.07 (d, *J* = 7.8 Hz, 1H), 7.74 (d, *J* = 2.5 Hz, 1H), 7.52 – 7.29 (m, 3H), 7.18 – 7.05 (m, 2H), 4.34 (t, *J* = 4.9 Hz, 2H), 3.60 – 3.30 (m, 2H), 3.15 (t, *J* = 7.7 Hz, 2H), 3.04 (t, *J* = 7.7 Hz, 2H), 2.94 (t, *J* = 7.2 Hz, 2H), 2.82 (t, *J* = 7.0 Hz, 2H), 2.15 – 1.96 (m, 2H), 1.63 (m, 4H). **^{13}C- NMR** (75 MHz, DMSO-d_6): δ = 151.4, 140.5, 135.1, 125.6, 122.8, 122.4, 120.2, 118.1, 115.4, 111.7, 111.1, 104.7, 64.4, 46.3, 46.1, 44.4, 43.9, 38.2, 24.1, 22.6, 22.3. **MS (ESI):** found: m/z = 355.0 [M+H$^+$], calculated: m/z = 355.3 [M+H$^+$]. 97% purity.

RNA oligonucleotides. RNA oligonucleotides except for the label-free *Ba* SAM-VI riboswitch which was prepared by polymerase chain reaction (PCR, see below) and *in vitro* transcription (IVT, see below), were purchased from Biomers.net GmbH (Ulm, Germany) (5'-biotin-*Ba* SAM-VI riboswitch, 5'-biotin-*Tte* preQ$_1$ riboswitch, *Tte* preQ$_1$ riboswitch) and Integrated DNA Technologies (Coralville, IA, USA) (5'-Cy5-*Tte* preQ$_1$ riboswitch, 5'-Cy5-*Ba* SAM-VI riboswitch) in HPLC-purified quality (for oligo sequences see Tables S6 and S7). The concentrations of RNA were determined using a NanoDrop2000 (ThermoFisher Scientific, Waltham, MA, USA). All experiments were performed under RNase-avoiding conditions. All experiments including Cy5- and biotin-labelled RNA were performed under light protection.

Polymerase chain reaction. The DNA template for the IVT was amplified using PCR. The PCR reaction mixture contained 10 nM of template DNA, 2 µM of each forward and reverse primer, 400 µM dNTP Mix (New England Biolabs GmbH, Frankfurt, Germany)), 3 mM $MgCl_2$, 0.05 U mL^{-1} *Taq* DNA Polymerase (New England Biolabs GmbH, Frankfurt, Germany) and 1× standard *Taq* reaction buffer (New England Biolabs GmbH, Frankfurt, Germany) in a final volume of 200 µL. The reaction was performed using a ^{3}Prime thermal cycler (Cole-Parmer Ltd, Stone, St, UK) with the following cycle parameters: initial denaturation (95 °C, 30 s), followed by 30 cycles of denaturation (95 °C, 30 s), annealing (temperature depends on the primer sequence and was determined by NEB Tm calculator, 30 s) and extension (72 °C, 45 s), and a final extension (72 °C, 4 min). The PCR product homogeneity was determined on a 10% native PAGE gel to confirm the size and purity.

***In vitro* transcription.** The crude PCR products were used as templates for IVT using the T7 polymerase system. The IVT reaction mixture contained 400 µL PCR product, (without further purification), 7.5 mM of each NTP (Jena Bioscience, Jena, Germany), 10 mM DTT, 30 mM $MgCl_2$, 5 µL HighYield T7 RNA polymerase mix (Jena Bioscience, Jena, Germany) and 1× HighYield T7 reaction buffer (Jena Bioscience) in a final volume of 1 mL. The reaction was incubated at 37 °C for 4 hours. The formed pyrophosphate was removed by centrifugation at 14,000 g for 5 min directly after transcription and the cleared supernatant was transferred to a micro reaction tube for further sample processing. To remove the PCR- and IVT-enzymes, phenol/chloroform extraction was performed twice using ROTI® phenol/chloroform/isoamyl alcohol (Carl

Roth, Karlsruhe, Germany) and the standard protocol from the supplier. For further purification and concentration of the RNA the RNA Clean & Concentrator-5 kit (Zymo Research Europe GmbH, Freiburg, Germany) including DNase I was used. The RNA product was determined on a 10% denaturing PAGE gel to confirm size and purity.

Isothermal titration calorimetry. ITC experiments were performed on an MicroCal PEAQ-ITC Automated calorimeter (Malvern Panalytical, Malvern, UK) at 20 °C. Prior to titration, the RNA was desalted using Zebra spin desalting columns (ThermoFisher Scientific, Waltham, MA, USA) following the standard protocol from the supplier, and lyophilized. RNA was resolved in ITC buffer (50 mM HEPES pH 7.0, 100 mM NaCl, 6 mM $MgCl_2$ for *Tte* $preQ_1$ riboswitch and 40 mM HEPES pH 7.0, 50 mM KCl, 10 mM $MgCl_2$ for *Ba* SAM-VI riboswitch). For refolding, the RNA was heated to 75 °C for 5 min and cooled down to room temperature for 20 min. For the *Tte* $preQ_1$ riboswitch experiments, the RNA was diluted to 20 µM in ITC buffer, $preQ_1$ and $preQ_0$ were diluted to 200 µM in ITC buffer. For each experiment, 19 ligand injections in a volume of 2.0 µL, with a 0.5 µL s^{-1} rate, 150 s intervals between injections and a reference power of 41.9 µW were performed. For the *Ba* SAM-VI riboswitch experiments, the RNA was diluted to 15 µM (SAM, MS02, MS04 measurements) and 30 µM (SAH measurements) in ITC buffer. SAM, **1** and **2** were diluted to 150 µM in ITC buffer, SAH was diluted to 300 µM in ITC buffer. For each experiment, 13 ligand injections in a volume of 3.0 µL, with a 0.5 µL s^{-1} rate, 150 s intervals between injections and a reference power of 41.9 µW were performed. All experiments contained a final DMSO concentration of 0.2%. Due to high sample consumption, thermodynamic profiles are not corrected for buffer ionization effects.[78,108] Integrated heat data were analyzed using a one-site binding model via the MicroCal PEAQ-ITC Analysis Software (version 1.21), provided by the manufacturer. Kinetic data were analyzed using the AFFINImeter web tool (Software 4 Science Developments). All ITC experiments were independently repeated three times.

Microscale Thermophoresis. MST experiments were performed on a Monolith NT.115 (NanoTemper Technologies, Munich, Germany) using standard uncoated capillaries as described previously.[109–111] Briefly, for refolding, the 5'-Cy5-labelled RNA was heated to 75 °C for 5 min in MST buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl, 25 mM $MgCl_2$) and cooled down to room temperature for 20 min. The RNA was diluted

to 20 nM in MST buffer and supplemented with ligands in appropriate concentrations ranging from 1 mM to 0.1 nM in 3.16-fold (semi-logarithmic) dilution series. Cross-selectivity experiments were performed under the same conditions, but with ligand concentrations of 1 mM, 100 µM, 1 µM, 100 nM and 1 nM. For each experiment, one DMSO negative control was carried out. Final DMSO concentration was always 2%. MST measurements were analyzed using the NT analysis software (version 1.5.41) and exported for statistical analysis and plotting in GraphPad Prism (version 8.0.1, GraphPad Software, Boston, MA, USA, https://www.graphpad.com).

ASSOCIATED CONTENT

Supporting Information. Data of SPR measurements and representative sensorgrams; data of MST measurements, MST raw traces and dose-response plots; data of (kin)ITC experiments, ITC injection diagrams, isotherms, signature plots, and ETCs; NMR spectra, mass spectra, and LC-MS chromatograms of all compounds tested in assays.
Molecular formula strings
This material is available free of charge via the Internet.

AUTHOR INFORMATION

Corresponding Author

* correspondence: CK, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany, e-mail: kerstec@uni-mainz.de

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript. / ‡These authors (LAR, EK) contributed equally.

Funding Sources

ACKNOWLEDGMENT

ABBREVIATIONS

AUC, area under curve; ASO, antisense oligonucleotide; *Ba*, *Bifidobacterium angulatum*; Boc, *tert*-butoxycarbonyl; Boc_2O, di-*tert*-butyl decarbonate; Cy5, Cyanine5; DCM, methylene chloride; DIPEA, *N,N*-di*iso*propylethylamine; DM2, myotonic dystrophy type 2; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; DNA,

deoxyribonucleic acid; DNAse, deoxyribonuclease; dNTP, deoxynucleotide triphosphate; ETC, equilibration time curve; ESI, electron spray ionization; HCV IRES, hepatitis C virus internal ribosome entry site; HEPES, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); HPLC, high-performance liquid chromatography; mp, melting point; IVT, *in vitro* transcription; ITC, isothermal titration calorimetry; MS, mass spectrometry; MST, microscale thermophoresis; mRNA, messenger RNA; NMR, nuclear magnetic resonance; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; PINAD, proximity-induced nucleic acid degrader; $preQ_1$, pre-queuosine-1; $preQ_0$, pre-queuosine-0; PROTAC, proteolysis targeting chimera; RiboTAC, ribonuclease targeting chimera; RNase, ribonuclease; RU, response units; SAH, *S*-adenosyl homocysteine; SAM, *S*-adenosyl methionine; siRNA, small interfering RNA; SMA, spinal muscular atrophy; SMN2, survival motor neuron 2; SPR, surface plasmon resonance; *Taq*, *Thermus aquaticus*; TAR, trans-activation response element; TBTU, benzotriazol-1-yloxy(dimethylamino)methylidene]-dimethylazanium tetrafluoroborate; TFA, trifluoroacetic acid; THF, tetrahydrofuran; TLC, thin layer chromatography; Tm, melting temperature; TPP, thiamine pyrophosphate; Tris-HCl, Tris(hydroxymethyl)aminomethane hydrochloride; *Tte*, *Thermoanaerobacter tengcongensis*.

REFERENCES

(1) Falese, J. P.; Donlic, A.; Hargrove, A. E. Targeting RNA with Small Molecules: From Fundamental Principles towards the Clinic. *Chem Soc Rev* **2021**, *50* (4), 2224–2243. https://doi.org/10.1039/D0CS01261K.

(2) Hargrove, A. E. Small Molecule–RNA Targeting: Starting with the Fundamentals. *Chemical Communications* **2020**, *56* (94), 14744–14756. https://doi.org/10.1039/D0CC06796B.

(3) Kovachka, S.; Panosetti, M.; Grimaldi, B.; Azoulay, S.; Di Giorgio, A.; Duca, M. Small Molecule Approaches to Targeting RNA. *Nat Rev Chem* **2024**, *8* (2), 120–135. https://doi.org/10.1038/s41570-023-00569-9.

(4) Yu, A.-M.; Choi, Y. H.; Tu, M.-J. RNA Drugs and RNA Targets for Small Molecules: Principles, Progress, and Challenges. *Pharmacol Rev* **2020**, *72* (4), 862–898. https://doi.org/10.1124/pr.120.019554.

(5) Childs-Disney, J. L.; Yang, X.; Gibaut, Q. M. R.; Tong, Y.; Batey, R. T.; Disney, M. D. Targeting RNA Structures with Small Molecules. *Nat Rev Drug Discov* **2022**, *21* (10), 736–762. https://doi.org/10.1038/s41573-022-00521-4.

(6) Costales, M. G.; Childs-Disney, J. L.; Haniff, H. S.; Disney, M. D. How We Think about Targeting RNA with Small Molecules. *J Med Chem* **2020**, *63* (17), 8880–8900. https://doi.org/10.1021/acs.jmedchem.9b01927.

(7) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like Small Molecules. *Nat Rev Drug Discov* **2018**, *17* (8), 547–558. https://doi.org/10.1038/nrd.2018.93.

(8) Paik, J. Risdiplam: A Review in Spinal Muscular Atrophy. *CNS Drugs* **2022**, *36* (4), 401–410. https://doi.org/10.1007/s40263-022-00910-8.

(9) Ratni, H.; Ebeling, M.; Baird, J.; Bendels, S.; Bylund, J.; Chen, K. S.; Denk, N.; Feng, Z.; Green, L.; Guerard, M.; Jablonski, P.; Jacobsen, B.; Khwaja, O.; Kletzl, H.; Ko, C.-P.; Kustermann, S.; Marquet, A.; Metzger, F.; Mueller, B.; Naryshkin, N. A.; Paushkin, S. V.; Pinard, E.; Poirier, A.; Reutlinger, M.; Weetall, M.; Zeller, A.; Zhao, X.; Mueller, L. Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (*SMN2*) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA). *J Med Chem* **2018**, *61* (15), 6501–6517. https://doi.org/10.1021/acs.jmedchem.8b00741.

(10) Ratni, H.; Scalco, R. S.; Stephan, A. H. Risdiplam, the First Approved Small Molecule Splicing Modifier Drug as a Blueprint for Future Transformative Medicines. *ACS Med Chem Lett* **2021**, *12* (6), 874–877. https://doi.org/10.1021/acsmedchemlett.0c00659.

(11) Carothers, J. M.; Oestreich, S. C.; Davis, J. H.; Szostak, J. W. Informational Complexity and Functional Activity of RNA Structures. *J Am Chem Soc* **2004**, *126* (16), 5130–5137. https://doi.org/10.1021/ja031504a.

(12) Ritchie, T. J.; Macdonald, S. J. F. How Drug-like Are 'Ugly' Drugs: Do Drug-Likeness Metrics Predict ADME Behaviour in Humans? *Drug Discov Today* **2014**, *19* (4), 489–495. https://doi.org/10.1016/j.drudis.2014.01.007.

(13) Lipinski, C. A. Lead- and Drug-like Compounds: The Rule-of-Five Revolution. *Drug Discov Today Technol* **2004**, *1* (4), 337–341. https://doi.org/10.1016/j.ddtec.2004.11.007.

(14) Haniff, H. S.; Tong, Y.; Liu, X.; Chen, J. L.; Suresh, B. M.; Andrews, R. J.; Peterson, J. M.; O'Leary, C. A.; Benhamou, R. I.; Moss, W. N.; Disney, M. D. Targeting the SARS-CoV-2 RNA Genome with Small Molecule Binders and Ribonuclease Targeting Chimera (RIBOTAC) Degraders. *ACS Cent Sci* **2020**, *6* (10), 1713–1721. https://doi.org/10.1021/acscentsci.0c00984.

(15) Meyer, S. M.; Tanaka, T.; Zanon, P. R. A.; Baisden, J. T.; Abegg, D.; Yang, X.; Akahori, Y.; Alshakarchi, Z.; Cameron, M. D.; Adibekian, A.; Disney, M. D. DNA-Encoded Library Screening To Inform Design of a Ribonuclease Targeting Chimera (RiboTAC). *J Am Chem Soc* **2022**, *144* (46), 21096–21102. https://doi.org/10.1021/jacs.2c07217.

(16) Tong, Y.; Lee, Y.; Liu, X.; Childs-Disney, J. L.; Suresh, B. M.; Benhamou, R. I.; Yang, C.; Li, W.; Costales, M. G.; Haniff, H. S.; Sievers, S.; Abegg, D.; Wegner, T.; Paulisch, T. O.; Lekah, E.; Grefe, M.; Crynen, G.; Van Meter, M.; Wang, T.; Gibaut, Q. M. R.; Cleveland, J. L.; Adibekian, A.; Glorius, F.; Waldmann, H.; Disney, M. D. Programming Inactive RNA-Binding Small Molecules into Bioactive Degraders. *Nature* **2023**, *618* (7963), 169–179. https://doi.org/10.1038/s41586-023-06091-8.

(17) Costales, M. G.; Aikawa, H.; Li, Y.; Childs-Disney, J. L.; Abegg, D.; Hoch, D. G.; Pradeep Velagapudi, S.; Nakai, Y.; Khan, T.; Wang, K. W.; Yildirim, I.; Adibekian, A.; Wang, E. T.; Disney, M. D. Small-Molecule Targeted Recruitment of a Nuclease to

Cleave an Oncogenic RNA in a Mouse Model of Metastatic Cancer. *Proceedings of the National Academy of Sciences* **2020**, *117* (5), 2406–2411. https://doi.org/10.1073/pnas.1914286117.

(18) Tong, Y.; Gibaut, Q. M. R.; Rouse, W.; Childs-Disney, J. L.; Suresh, B. M.; Abegg, D.; Choudhary, S.; Akahori, Y.; Adibekian, A.; Moss, W. N.; Disney, M. D. Transcriptome-Wide Mapping of Small-Molecule RNA-Binding Sites in Cells Informs an Isoform-Specific Degrader of QSOX1 MRNA. *J Am Chem Soc* **2022**, *144* (26), 11620–11625. https://doi.org/10.1021/jacs.2c01929.

(19) Tong, Y.; Zhang, P.; Yang, X.; Liu, X.; Zhang, J.; Grudniewska, M.; Jung, I.; Abegg, D.; Liu, J.; Childs-Disney, J. L.; Gibaut, Q. M. R.; Haniff, H. S.; Adibekian, A.; Mouradian, M. M.; Disney, M. D. Decreasing the Intrinsically Disordered Protein α-Synuclein Levels by Targeting Its Structured MRNA with a Ribonuclease-Targeting Chimera. *Proceedings of the National Academy of Sciences* **2024**, *121* (2). https://doi.org/10.1073/pnas.2306682120.

(20) Mikutis, S.; Rebelo, M.; Yankova, E.; Gu, M.; Tang, C.; Coelho, A. R.; Yang, M.; Hazemi, M. E.; Pires de Miranda, M.; Eleftheriou, M.; Robertson, M.; Vassiliou, G. S.; Adams, D. J.; Simas, J. P.; Corzana, F.; Schneekloth, J. S.; Tzelepis, K.; Bernardes, G. J. L. Proximity-Induced Nucleic Acid Degrader (PINAD) Approach to Targeted RNA Degradation Using Small Molecules. *ACS Cent Sci* **2023**, *9* (5), 892–904. https://doi.org/10.1021/acscentsci.3c00015.

(21) Sun, X.; Gao, H.; Yang, Y.; He, M.; Wu, Y.; Song, Y.; Tong, Y.; Rao, Y. PROTACs: Great Opportunities for Academia and Industry. *Signal Transduct Target Ther* **2019**, *4* (1), 64. https://doi.org/10.1038/s41392-019-0101-6.

(22) He, M.; Cao, C.; Ni, Z.; Liu, Y.; Song, P.; Hao, S.; He, Y.; Sun, X.; Rao, Y. PROTACs: Great Opportunities for Academia and Industry (an Update from 2020 to 2021). *Signal Transduct Target Ther* **2022**, *7* (1), 181. https://doi.org/10.1038/s41392-022-00999-9.

(23) Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. Understanding Nucleic Acid–Ion Interactions. *Annu Rev Biochem* **2014**, *83* (1), 813–841. https://doi.org/10.1146/annurev-biochem-060409-092720.

(24) Chu, V. B.; Bai, Y.; Lipfert, J.; Herschlag, D.; Doniach, S. A Repulsive Field: Advances in the Electrostatics of the Ion Atmosphere. *Curr Opin Chem Biol* **2008**, *12* (6), 619–625. https://doi.org/10.1016/j.cbpa.2008.10.010.

(25) Draper, D. E.; Grilley, D.; Soto, A. M. Ions and RNA Folding. *Annu Rev Biophys Biomol Struct* **2005**, *34* (1), 221–243. https://doi.org/10.1146/annurev.biophys.34.040204.144511.

(26) Donlic, A.; Swanson, E. G.; Chiu, L.-Y.; Wicks, S. L.; Juru, A. U.; Cai, Z.; Kassam, K.; Laudeman, C.; Sanaba, B. G.; Sugarman, A.; Han, E.; Tolbert, B. S.; Hargrove, A. E. R-BIND 2.0: An Updated Database of Bioactive RNA-Targeting Small Molecules and Associated RNA Secondary Structures. *ACS Chem Biol* **2022**, *17* (6), 1556–1566. https://doi.org/10.1021/acschembio.2c00224.

(27) Allen, T. E. H.; McDonagh, J. L.; Broncel, M.; Bryant, C. J.; Incarnato, D.; Vasudevan, A.; Khan, R. T. Physicochemical Principles Driving Small Molecule Binding to RNA. *bioRxiv* **2024**, 2024.01.31.578268. https://doi.org/10.1101/2024.01.31.578268.

(28) Aradi, K.; Di Giorgio, A.; Duca, M. Aminoglycoside Conjugation for RNA Targeting: Antimicrobials and Beyond. *Chemistry – A European Journal* **2020**, *26* (54), 12273–12309. https://doi.org/10.1002/chem.202002258.

(29) Maucort, C.; Vo, D. D.; Aouad, S.; Charrat, C.; Azoulay, S.; Di Giorgio, A.; Duca, M. Design and Implementation of Synthetic RNA Binders for the Inhibition of MiR-21 Biogenesis. *ACS Med Chem Lett* **2021**, *12* (6), 899–906. https://doi.org/10.1021/acsmedchemlett.0c00682.

(30) Duca, M.; Malnuit, V.; Barbault, F.; Benhida, R. Design of Novel RNA Ligands That Bind Stem–Bulge HIV-1 TAR RNA. *Chemical Communications* **2010**, *46* (33), 6162. https://doi.org/10.1039/c0cc00645a.

(31) Vo, D. D.; Staedel, C.; Zehnacker, L.; Benhida, R.; Darfeuille, F.; Duca, M. Targeting the Production of Oncogenic MicroRNAs with Multimodal Synthetic Small Molecules. *ACS Chem Biol* **2014**, *9* (3), 711–721. https://doi.org/10.1021/cb400668h.

(32) Ma, C.; Baker, N. A.; Joseph, S.; McCammon, J. A. Binding of Aminoglycoside Antibiotics to the Small Ribosomal Subunit: A Continuum Electrostatics Investigation. *J Am Chem Soc* **2002**, *124* (7), 1438–1442. https://doi.org/10.1021/ja016830+.

(33) Yang, G.; Trylska, J.; Tor, Y.; McCammon, J. A. Binding of Aminoglycosidic Antibiotics to the Oligonucleotide A-Site Model and 30S Ribosomal Subunit: Poisson−Boltzmann Model, Thermal Denaturation, and Fluorescence Studies. *J Med Chem* **2006**, *49* (18), 5478–5490. https://doi.org/10.1021/jm060288o.

(34) Wang, H.; Tor, Y. Electrostatic Interactions in RNA Aminoglycosides Binding. *J Am Chem Soc* **1997**, *119* (37), 8734–8735. https://doi.org/10.1021/ja9715105.

(35) Długosz, M.; Antosiewicz, J. M.; Trylska, J. Association of Aminoglycosidic Antibiotics with the Ribosomal A-Site Studied with Brownian Dynamics. *J Chem Theory Comput* **2008**, *4* (4), 549–559. https://doi.org/10.1021/ct700210n.

(36) Długosz, M.; Trylska, J. Aminoglycoside Association Pathways with the 30S Ribosomal Subunit. *J Phys Chem B* **2009**, *113* (20), 7322–7330. https://doi.org/10.1021/jp8112914.

(37) Kulik, M.; Goral, A. M.; Jasiński, M.; Dominiak, P. M.; Trylska, J. Electrostatic Interactions in Aminoglycoside-RNA Complexes. *Biophys J* **2015**, *108* (3), 655–665. https://doi.org/10.1016/j.bpj.2014.12.020.

(38) Wang, H.; Tor, Y. RNA–Aminoglycoside Interactions: Design, Synthesis, and Binding of “Amino-Aminoglycosides” to RNA. *Angewandte Chemie International Edition* **1998**, *37* (1–2), 109–111. https://doi.org/10.1002/(SICI)1521-3773(19980202)37:1/2<109::AID-ANIE109>3.0.CO;2-0.

(39) Hobbie, S. N.; Akshay, S.; Kalapala, S. K.; Bruell, C. M.; Shcherbakov, D.; Böttger, E. C. Genetic Analysis of Interactions with Eukaryotic RRNA Identify the Mitoribosome as Target in Aminoglycoside Ototoxicity. *Proceedings of the National Academy of Sciences* **2008**, *105* (52), 20888–20893. https://doi.org/10.1073/pnas.0811258106.

(40) Lopez-Novoa, J. M.; Quiros, Y.; Vicente, L.; Morales, A. I.; Lopez-Hernandez, F. J. New Insights into the Mechanism of Aminoglycoside Nephrotoxicity: An Integrative Point of View. *Kidney Int* **2011**, *79* (1), 33–45. https://doi.org/10.1038/ki.2010.337.

(41) Mingeot-Leclercq, M.-P.; Tulkens, P. M. Aminoglycosides: Nephrotoxicity. *Antimicrob Agents Chemother* **1999**, *43* (5), 1003–1012. https://doi.org/10.1128/AAC.43.5.1003.

(42) Disney, M. D.; Winkelsas, A. M.; Velagapudi, S. P.; Southern, M.; Fallahi, M.; Childs-Disney, J. L. Inforna 2.0: A Platform for the Sequence-Based Design of Small Molecules Targeting Structured RNAs. *ACS Chem Biol* **2016**, *11* (6), 1720–1728. https://doi.org/10.1021/acschembio.6b00001.

(43) Luedtke, N. W.; Baker, T. J.; Goodman, M.; Tor, Y. Guanidinoglycosides: A Novel Family of RNA Ligands. *J Am Chem Soc* **2000**, *122* (48), 12035–12036. https://doi.org/10.1021/ja002421m.

(44) Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Lee, J.; Andricioaei, I.; Markovitz, D. M.; Al-Hashimi, H. M. Discovery of Selective Bioactive Small Molecules by Targeting an RNA Dynamic Ensemble. *Nat Chem Biol* **2011**, *7* (8), 553–559. https://doi.org/10.1038/nchembio.596.

(45) Lee, M. M.; Pushechnikov, A.; Disney, M. D. Rational and Modular Design of Potent Ligands Targeting the RNA That Causes Myotonic Dystrophy 2. *ACS Chem Biol* **2009**, *4* (5), 345–355. https://doi.org/10.1021/cb900025w.

(46) Thomas, J. R.; Hergenrother, P. J. Targeting RNA with Small Molecules. *Chem Rev* **2008**, *108* (4), 1171–1224. https://doi.org/10.1021/cr0681546.

(47) Guianvarc'h, D.; Benhida, R.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S. Incorporation of a Novel Nucleobase Allows Stable Oligonucleotide-Directed Triple Helix Formation at the Target Sequence Containing a Purine·pyrimidine Interruption. *Chemical Communications* **2001**, *1* (18), 1814–1815. https://doi.org/10.1039/b103743a.

(48) Guianvarc'h, D.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S.; Benhida, R. Design of Artificial Nucleobases for the Recognition of the AT Inversion by Triple-Helix Forming Oligonucleotides: A Structure–Stability Relationship Study and Neighbour Bases Effect. *Bioorg Med Chem* **2003**, *11* (13), 2751–2759. https://doi.org/10.1016/S0968-0896(03)00229-3.

(49) Cabrera, C.; Gutiérrez, A.; Barretina, J.; Blanco, J.; Litovchick, A.; Lapidot, A.; Clotet, B.; Esté, J. A. Anti-HIV Activity of a Novel Aminoglycoside-Arginine Conjugate. *Antiviral Res* **2002**, *53* (1), 1–8. https://doi.org/10.1016/S0166-3542(01)00188-7.

(50) Vo, D. D.; Tran, T. P. A.; Staedel, C.; Benhida, R.; Darfeuille, F.; Di Giorgio, A.; Duca, M. Oncogenic MicroRNAs Biogenesis as a Drug Target: Structure–Activity

Relationship Studies on New Aminoglycoside Conjugates. *Chemistry – A European Journal* **2016**, *22* (15), 5350–5362. https://doi.org/10.1002/chem.201505094.

(51) Vo, D. D.; Becquart, C.; Tran, T. P. A.; Di Giorgio, A.; Darfeuille, F.; Staedel, C.; Duca, M. Building of Neomycin–Nucleobase–Amino Acid Conjugates for the Inhibition of Oncogenic MiRNAs Biogenesis. *Org Biomol Chem* **2018**, *16* (34), 6262–6274. https://doi.org/10.1039/C8OB01858H.

(52) Sucheck, S. J.; Greenberg, W. A.; Tolbert, T. J.; Wong, C.-H. Design of Small Molecules That Recognize RNA: Development of Aminoglycosides as Potential Antitumor Agents That Target Oncogenic RNA Sequences. *Angewandte Chemie* **2000**, *112* (6), 1122–1126. https://doi.org/10.1002/(SICI)1521-3757(20000317)112:6<1122::AID-ANGE1122>3.0.CO;2-V.

(53) Staedel, C.; Tran, T. P. A.; Giraud, J.; Darfeuille, F.; Di Giorgio, A.; Tourasse, N. J.; Salin, F.; Uriac, P.; Duca, M. Modulation of Oncogenic MiRNA Biogenesis Using Functionalized Polyamines. *Sci Rep* **2018**, *8* (1), 1667. https://doi.org/10.1038/s41598-018-20053-5.

(54) Maucort, C.; Bonnet, M.; Ortuno, J.-C.; Tucker, G.; Quissac, E.; Verreault, M.; Azoulay, S.; Di Giorgio, C.; Di Giorgio, A.; Duca, M. Synthesis of Bleomycin-Inspired RNA Ligands Targeting the Biogenesis of Oncogenic MiRNAs. *J Med Chem* **2023**, *66* (15), 10639–10657. https://doi.org/10.1021/acs.jmedchem.3c00797.

(55) Umuhire Juru, A.; Patwardhan, N. N.; Hargrove, A. E. Understanding the Contributions of Conformational Changes, Thermodynamics, and Kinetics of RNA–Small Molecule Interactions. *ACS Chem Biol* **2019**, *14* (5), 824–838. https://doi.org/10.1021/acschembio.8b00945.

(56) Xi, H.; Gray, D.; Kumar, S.; Arya, D. P. Molecular Recognition of Single-stranded RNA: Neomycin Binding to Poly(A). *FEBS Lett* **2009**, *583* (13), 2269–2275. https://doi.org/10.1016/j.febslet.2009.06.007.

(57) Pascale, L.; Azoulay, S.; Di Giorgio, A.; Zenacker, L.; Gaysinski, M.; Clayette, P.; Patino, N. Thermodynamic Studies of a Series of Homologous HIV-1 TAR RNA Ligands Reveal That Loose Binders Are Stronger Tat Competitors than Tight Ones. *Nucleic Acids Res* **2013**, *41* (11), 5851–5863. https://doi.org/10.1093/nar/gkt237.

(58) Lightfoot, H. L.; Hall, J. Endogenous Polyamine Function—the RNA Perspective. *Nucleic Acids Res* **2014**, *42* (18), 11275–11290. https://doi.org/10.1093/nar/gku837.

(59) Pilch, D. S.; Kaul, M.; Barbieri, C. M.; Kerrigan, J. E. Thermodynamics of Aminoglycoside–RRNA Recognition. *Biopolymers* **2003**, *70* (1), 58–79. https://doi.org/10.1002/bip.10411.

(60) Frolov, L.; Dix, A.; Tor, Y.; Tesler, A. B.; Chaikin, Y.; Vaskevich, A.; Rubinstein, I. Direct Observation of Aminoglycoside–RNA Binding by Localized Surface Plasmon Resonance Spectroscopy. *Anal Chem* **2013**, *85* (4), 2200–2207. https://doi.org/10.1021/ac3029079.

(61) Wickiser, J. K.; Cheah, M. T.; Breaker, R. R.; Crothers, D. M. The Kinetics of Ligand Binding by an Adenine-Sensing Riboswitch. *Biochemistry* **2005**, *44* (40), 13404–13414. https://doi.org/10.1021/bi051008u.

(62) Jenkins, J. L.; Krucinska, J.; McCarty, R. M.; Bandarian, V.; Wedekind, J. E. Comparison of a PreQ1 Riboswitch Aptamer in Metabolite-Bound and Free States with Implications for Gene Regulation. *Journal of Biological Chemistry* **2011**, *286* (28), 24626–24637. https://doi.org/10.1074/jbc.M111.230375.

(63) Neuner, E.; Frener, M.; Lusser, A.; Micura, R. Superior Cellular Activities of Azido- over Amino-Functionalized Ligands for Engineered PreQ1riboswitches in E.Coli. *RNA Biol* **2018**, *15* (10), 1376–1383. https://doi.org/10.1080/15476286.2018.1534526.

(64) Sengupta, R. N.; Herschlag, D. Enhancement of RNA/Ligand Association Kinetics via an Electrostatic Anchor. *Biochemistry* **2019**, *58* (24), 2760–2768. https://doi.org/10.1021/acs.biochem.9b00231.

(65) Jencks, W. P. Binding Energy, Specificity, and Enzymic Catalysis: The Circe Effect; 1975; pp 219–410. https://doi.org/10.1002/9780470122884.ch4.

(66) Sun, A.; Gasser, C.; Li, F.; Chen, H.; Mair, S.; Krasheninina, O.; Micura, R.; Ren, A. SAM-VI Riboswitch Structure and Signature for Ligand Discrimination. *Nat Commun* **2019**, *10* (1), 5728. https://doi.org/10.1038/s41467-019-13600-9.

(67) Price, I. R.; Grigg, J. C.; Ke, A. Common Themes and Differences in SAM Recognition among SAM Riboswitches. *Biochimica et Biophysica Acta (BBA) - Gene*

Regulatory Mechanisms **2014**, *1839* (10), 931–938. https://doi.org/10.1016/j.bbagrm.2014.05.013.

(68) Connelly, C. M.; Numata, T.; Boer, R. E.; Moon, M. H.; Sinniah, R. S.; Barchi, J. J.; Ferré-D'Amaré, A. R.; Schneekloth, J. S. Synthetic Ligands for PreQ1 Riboswitches Provide Structural and Mechanistic Insights into Targeting RNA Tertiary Structure. *Nat Commun* **2019**, *10* (1), 1501. https://doi.org/10.1038/s41467-019-09493-3.

(69) Mirihana Arachchilage, G.; Sherlock, M. E.; Weinberg, Z.; Breaker, R. R. SAM-VI RNAs Selectively Bind S-Adenosylmethionine and Exhibit Similarities to SAM-III Riboswitches. *RNA Biol* **2018**, *15* (3), 371–378. https://doi.org/10.1080/15476286.2017.1399232.

(70) Wang, J. X.; Breaker, R. R. Riboswitches That Sense S -Adenosylmethionine and S -AdenosylhomocysteineThis Paper Is One of a Selection of Papers Published in This Special Issue, Entitled CSBMCB — Systems and Chemical Biology, and Has Undergone the Journal's Usual Peer Review Process. *Biochemistry and Cell Biology* **2008**, *86* (2), 157–168. https://doi.org/10.1139/O08-008.

(71) Batey, R. T. Recognition of S-adenosylmethionine by Riboswitches. *WIREs RNA* **2011**, *2* (2), 299–311. https://doi.org/10.1002/wrna.63.

(72) Lim, J.; Winkler, W. C.; Nakamura, S.; Scott, V.; Breaker, R. R. Molecular-Recognition Characteristics of SAM-Binding Riboswitches. *Angewandte Chemie* **2006**, *118* (6), 978–982. https://doi.org/10.1002/ange.200503198.

(73) Lu, C.; Smith, A. M.; Fuchs, R. T.; Ding, F.; Rajashankar, K.; Henkin, T. M.; Ke, A. Crystal Structures of the SAM-III/SMK Riboswitch Reveal the SAM-Dependent Translation Inhibition Mechanism. *Nat Struct Mol Biol* **2008**, *15* (10), 1076–1083. https://doi.org/10.1038/nsmb.1494.

(74) Priyakumar, U. D. Atomistic Details of the Ligand Discrimination Mechanism of S MK /SAM-III Riboswitch. *J Phys Chem B* **2010**, *114* (30), 9920–9925. https://doi.org/10.1021/jp1042427.

(75) Piña, M. de las N.; Frontera, A.; Bauza, A. Charge Assisted S/Se Chalcogen Bonds in SAM Riboswitches: A Combined PDB and *Ab Initio* Study. *ACS Chem Biol* **2021**, *16* (9), 1701–1708. https://doi.org/10.1021/acschembio.1c00417.

(76) Feig, A. L. Applications of Isothermal Titration Calorimetry in RNA Biochemistry and Biophysics. *Biopolymers* **2007**, *87* (5–6), 293–301. https://doi.org/10.1002/bip.20816.

(77) El Deeb, S.; Al-Harrasi, A.; Khan, A.; Al-Broumi, M.; Al-Thani, G.; Alomairi, M.; Elumalai, P.; Sayed, R. A.; Ibrahim, A. E. Microscale Thermophoresis as a Powerful Growing Analytical Technique for the Investigation of Biomolecular Interaction and the Determination of Binding Parameters. *Methods Appl Fluoresc* **2022**, *10* (4), 042001. https://doi.org/10.1088/2050-6120/ac82a6.

(78) Hammerschmidt, S. J.; Barthels, F.; Weldert, A. C.; Kersten, C. Advanced Isothermal Titration Calorimetry for Medicinal Chemists with ITCcalc. *J Chem Educ* **2024**, *101* (3), 1086–1095. https://doi.org/10.1021/acs.jchemed.3c01133.

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

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

(81) Starikov, E. B.; Nordén, B. Entropy–Enthalpy Compensation as a Fundamental Concept and Analysis Tool for Systematical Experimental Data. *Chem Phys Lett* **2012**, *538*, 118–120. https://doi.org/10.1016/j.cplett.2012.04.028.

(82) Ryde, U. A Fundamental View of Enthalpy–Entropy Compensation. *Med. Chem. Commun.* **2014**, *5* (9), 1324–1336. https://doi.org/10.1039/C4MD00057A.

(83) Marky, L. A.; Kupke, D. W. Enthalpy-Entropy Compensations in Nucleic Acids: Contribution of Electrostriction and Structural Hydration. In *Methods in Enzymology*; Academic Press, 2000; Vol. 323, pp 419–441. https://doi.org/https://doi.org/10.1016/S0076-6879(00)23376-4.

(84) Kumar, G. S.; Basu, A. The Use of Calorimetry in the Biophysical Characterization of Small Molecule Alkaloids Binding to RNA Structures. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2016**, *1860* (5), 930–944. https://doi.org/10.1016/j.bbagen.2015.10.026.

(85) Hendrix, M.; Priestley, E. S.; Joyce, G. F.; Wong, C.-H. Direct Observation of Aminoglycoside−RNA Interactions by Surface Plasmon Resonance. *J Am Chem Soc* **1997**, *119* (16), 3641–3648. https://doi.org/10.1021/ja964290o.

(86) Schroeder, G. M.; Akinyemi, O.; Malik, J.; Focht, C. M.; Pritchett, E. M.; Baker, C. D.; McSally, J. P.; Jenkins, J. L.; Mathews, D. H.; Wedekind, J. E. A Riboswitch Separated from Its Ribosome-Binding Site Still Regulates Translation. *Nucleic Acids Res* **2023**, *51* (5), 2464–2484. https://doi.org/10.1093/nar/gkad056.

(87) Wang, Y.; Parmar, S.; Schneekloth, J. S.; Tiwary, P. Interrogating RNA–Small Molecule Interactions with Structure Probing and Artificial Intelligence-Augmented Molecular Simulations. *ACS Cent Sci* **2022**, *8* (6), 741–748. https://doi.org/10.1021/acscentsci.2c00149.

(88) Parmar, S.; Bume, D. D.; Conelly, C.; Boer, R.; Prestwood, P. R.; Wang, Z.; Labuhn, H.; Sinnadurai, K.; Feri, A.; Ouellet, J.; Homan, P.; Numata, T.; Schneekloth, J. S. Mechanistic Analysis of Riboswitch Ligand Interactions Provides Insights into Pharmacological Control over Gene Expression. *bioRxiv* **2024**, 2024.02.23.581746. https://doi.org/10.1101/2024.02.23.581746.

(89) Muñoz, E.; Piñeiro, A. AFFINImeter Software: From Its Beginnings to Future Trends- A Literature Review. *J Appl Bioanal* **2018**, *4* (4), 124–139. https://doi.org/10.17145/jab.18.017.

(90) Dumas, P.; Ennifar, E.; Da Veiga, C.; Bec, G.; Palau, W.; Di Primo, C.; Piñeiro, A.; Sabin, J.; Muñoz, E.; Rial, J. Extending ITC to Kinetics with KinITC; 2016; pp 157–180. https://doi.org/10.1016/bs.mie.2015.08.026.

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

(92) Salim, N. N.; Feig, A. L. Isothermal Titration Calorimetry of RNA. *Methods* **2009**, *47* (3), 198–205. https://doi.org/10.1016/j.ymeth.2008.09.003.

(93) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. A Simple Ligand That Selectively Targets CUG Trinucleotide Repeats and Inhibits MBNL Protein Binding. *Proceedings of the National Academy of Sciences* **2009**, *106* (38), 16068–16073. https://doi.org/10.1073/pnas.0901824106.

(94) Binas, O.; de Jesus, V.; Landgraf, T.; Völklein, A. E.; Martins, J.; Hymon, D.; Kaur Bains, J.; Berg, H.; Biedenbänder, T.; Fürtig, B.; Lakshmi Gande, S.; Niesteruk, A.; Oxenfarth, A.; Shahin Qureshi, N.; Schamber, T.; Schnieders, R.; Tröster, A.; Wacker, A.; Wirmer-Bartoschek, J.; Wirtz Martin, M. A.; Stirnal, E.; Azzaoui, K.; Richter, C.; Sreeramulu, S.; José Blommers, M. J.; Schwalbe, H. 19 F NMR-Based Fragment Screening for 14 Different Biologically Active RNAs and 10 DNA and Protein Counter-Screens. *ChemBioChem* **2021**, *22* (2), 423–433. https://doi.org/10.1002/cbic.202000476.

(95) Klein, D. J.; Edwards, T. E.; Ferré-D'Amaré, A. R. Cocrystal Structure of a Class I PreQ1 Riboswitch Reveals a Pseudoknot Recognizing an Essential Hypermodified Nucleobase. *Nat Struct Mol Biol* **2009**, *16* (3), 343–344. https://doi.org/10.1038/nsmb.1563.

(96) Hucke, O.; Gottschling, D.; Schnapp, G.; Ries, B.; Kreuz, S.; Klugmann, M.; Hedwig, V.; Spöring, M.; Scharsich, S.; Hartig, J.; Aspnes, G. Designing Drug-like High Affinity Ligands for a TPP- Orthogonal Riboswitch Derived from the ThiM-Aptamer. *ACS Fall* **2023**.

(97) Schmit, D.; Milewicz, U.; Boerneke, M. A.; Burley, S.; Walsworth, K.; Um, J.; Hecht, D.; Hermann, T.; Bergdahl, B. M. Syntheses and Binding Testing of N 1-Alkylamino-Substituted 2-Aminobenzimidazole Analogues Targeting the Hepatitis C Virus Internal Ribosome Entry Site. *Aust J Chem* **2020**, *73* (3), 212–221. https://doi.org/10.1071/CH19526.

(98) Seth, P. P.; Miyaji, A.; Jefferson, E. A.; Sannes-Lowery, K. A.; Osgood, S. A.; Propp, S. S.; Ranken, R.; Massire, C.; Sampath, R.; Ecker, D. J.; Swayze, E. E.; Griffey, R. H. SAR by MS: Discovery of a New Class of RNA-Binding Small Molecules for the Hepatitis C Virus: Internal Ribosome Entry Site IIA Subdomain. *J Med Chem* **2005**, *48* (23), 7099–7102. https://doi.org/10.1021/jm050815o.

(99) Han, G.; Tamaki, M.; Hruby, V. J. Fast, Efficient and Selective Deprotection of the Tert-butoxycarbonyl (Boc) Group Using HCl/Dioxane (4 M). *The Journal of Peptide Research* **2001**, *58* (4), 338–341. https://doi.org/10.1034/j.1399-3011.2001.00935.x.

(100) 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.

(101) Chiranjeevi, B.; Koyyada, G.; Prabusreenivasan, S.; Kumar, V.; Sujitha, P.; Kumar, C. G.; Sridhar, B.; Shaik, S.; Chandrasekharam, M. Iron-Catalyzed Aryl-Aryl Cross Coupling Route for the Synthesis of 1-(2-Amino)-Phenyl)Dibenzo[b,d]Furan-2-Ol Derivatives and Their Biological Evaluation. *RSC Adv* **2013**, *3* (37), 16475. https://doi.org/10.1039/c3ra43345e.

(102) Yempala, T.; Cassels, B. K. Synthesis, Scope, 1H and 13C Spectral Assignments of Isomeric Dibenzofuran Carboxaldehydes. *Research on Chemical Intermediates* **2017**, *43* (3), 1291–1299. https://doi.org/10.1007/s11164-016-2698-1.

(103) Ku, S.-K.; Lee, J.-H.; O, Y.; Lee, W.; Song, G.-Y.; Bae, J.-S. Vascular Barrier Protective Effects of 3-N- or 3-O-Cinnamoyl Carbazole Derivatives. *Bioorg Med Chem Lett* **2015**, *25* (19), 4304–4307. https://doi.org/10.1016/j.bmcl.2015.07.079.

(104) Milne, A. H.; Tomlinson, M. L. 526. 1 : 2 : 3 : 4-Tetrahydro-6-Hydroxycarbazole and 3-Hydroxycarbazole. *Journal of the Chemical Society (Resumed)* **1952**, 2789. https://doi.org/10.1039/jr9520002789.

(105) Pájaro, Y.; Sathicq, Á.; Puello-Polo, E.; Pérez, A.; Romanelli, G.; Trilleras, J. An Efficient K $_2$ CO $_3$ -Promoted Synthesis of 1-Bromo-2-Aryloxyethane Derivatives and Evaluation of Larval Mortality against *Aedes Aegypti*. *J Chem* **2017**, *2017*, 1–7. https://doi.org/10.1155/2017/6175315.

(106) Zhang, Y.; Zheng, W.; Luo, Q.; Zhao, Y.; Zhang, E.; Liu, S.; Wang, F. Dual-Targeting Organometallic Ruthenium(<scp>ii</Scp>) Anticancer Complexes Bearing EGFR-Inhibiting 4-Anilinoquinazoline Ligands. *Dalton Transactions* **2015**, *44* (29), 13100–13111. https://doi.org/10.1039/C5DT01430A.

(107) Pirali, T.; Callipari, G.; Ercolano, E.; Genazzani, A. A.; Giovenzana, G. B.; Tron, G. C. A Concise Entry into Nonsymmetrical Alkyl Polyamines. *Org Lett* **2008**, *10* (19), 4199–4202. https://doi.org/10.1021/ol801612r.

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

(109) Kallert, E.; Behrendt, M.; Frey, A.; Kersten, C.; Barthels, F. Non-Covalent Dyes in Microscale Thermophoresis for Studying RNA Ligand Interactions and Modifications. *Chem Sci* **2023**, *14* (36), 9827–9837. https://doi.org/10.1039/D3SC02993J.

(110) Kallert, E.; Almena Rodriguez, L.; Husmann, J.-Å.; Blatt, K.; Kersten, C. Structure-Based Virtual Screening of Unbiased and RNA-Focused Libraries to Identify New Ligands for the HCV IRES Model System. *RSC Med Chem* **2024**, *15* (5), 1527–1538. https://doi.org/10.1039/D3MD00696D.

(111) Kallert, E.; Fischer, T. R.; Schneider, S.; Grimm, M.; Helm, M.; Kersten, C. Protein-Based Virtual Screening Tools Applied for RNA–Ligand Docking Identify New Binders of the PreQ1 -Riboswitch. *J Chem Inf Model* **2022**, *62* (17), 4134–4148. https://doi.org/10.1021/acs.jcim.2c00751.

Supporting Information

Electrostatic Anchoring in RNA-Ligand Design – Dissecting the Effects of Positive Charges on Affinity, Selectivity, Binding Kinetics and Thermodynamics

Laura Almena Rodriguez[1,‡], Elisabeth Kallert[1,‡], Jan-Åke Husmann[1], Kirsten Schaubruch[2] , Marvin Schwickert[1], Sabrina N. Hoba[1], Ralf Heermann[2], Christian Kersten[1,3,*]

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

[2] Institute of Molecular Physiology, Microbiology and Biotechnology, Johannes Gutenberg-University, Hanns-DieterHüsch-Weg 17, 55128 Mainz, Germany

[3] Institute for Quantitative and Computational Biosciences, Johannes Gutenberg-University, BioZentrum I, Hanns-Dieter-Hüsch Weg 15, 55128 Mainz, Germany

‡ LAR & EK contributed equally to this manuscript

*correspondence: kerstec@uni-mainz.de

Table of Contents

1. *SPR Experiments* 300
2. *MST Experiments* 308
3. *ITC Experiments* 311
4. *Analytical data compounds 3–8* 321
5. *Miscellaneous* 333

1. SPR Experiments

Table S1. Binding affinities and kinetic parameters from SPR measurements for SAM, SAH, **1**, and **2** against the *Ba* SAM-VI riboswitch, 1:1 binding fit model.

SAM-VI	**SPR**			
Compound	**K_D [µM]**	**k_{on} [1/Ms]**	**k_{off} [1/s]**	**$K_{D,SS}$ [µM]**
SAM	3.02	$2.51 \cdot 10^3$	$7.59 \cdot 10^{-3}$	5.27
	3.74	$4.76 \cdot 10^3$	$1.78 \cdot 10^{-2}$	4.83
	3.51	$2.52 \cdot 10^3$	$8.85 \cdot 10^{-3}$	5.17
	3.63	$2.38 \cdot 10^3$	$8.64 \cdot 10^{-3}$	4.60
	3.47	$2.44 \cdot 10^3$	$8.46 \cdot 10^{-3}$	4.68
	3.78	$3.72 \cdot 10^3$	$1.40 \cdot 10^{-2}$	5.15
	3.36	$2.61 \cdot 10^3$	$8.78 \cdot 10^{-3}$	4.67
	3.21	$2.60 \cdot 10^3$	$8.34 \cdot 10^{-3}$	4.85
	2.39	$2.90 \cdot 10^3$	$6.93 \cdot 10^{-3}$	3.40
	2.96	$3.16 \cdot 10^3$	$9.36 \cdot 10^{-3}$	3.54
	2.86	$3.14 \cdot 10^3$	$8.98 \cdot 10^{-3}$	3.69
SAM Average	**3.72±0.41**	**$(2.98 \pm 0.68) \cdot 10^3$**	**$(9.80 \pm 3.06) \cdot 10^{-3}$**	**4.53±0.64**
SAH	64.9	$6.95 \cdot 10^2$	$4.51 \cdot 10^{-2}$	58.3
	172	$2.86 \cdot 10^2$	$4.92 \cdot 10^{-2}$	102
	164	$2.93 \cdot 10^2$	$4.82 \cdot 10^{-2}$	105
	167	$2.90 \cdot 10^2$	$4.85 \cdot 10^{-2}$	127
	79.0	$5.38 \cdot 10^2$	$4.25 \cdot 10^{-2}$	80.0
	128	$4.74 \cdot 10^2$	$6.04 \cdot 10^{-2}$	89.2
	192	$3.29 \cdot 10^2$	$6.31 \cdot 10^{-2}$	141
	189	$3.16 \cdot 10^2$	$5.97 \cdot 10^{-2}$	157
	171	$3.35 \cdot 10^2$	$5.73 \cdot 10^{-2}$	160
	168	$3.70 \cdot 10^2$	$6.21 \cdot 10^{-2}$	140
	149	$3.75 \cdot 10^2$	$5.60 \cdot 10^{-2}$	137
	156	$3.61 \cdot 10^2$	$5.61 \cdot 10^{-2}$	154
SAH Average	**150±39**	**$(3.88 \pm 1.18) \cdot 10^2$**	**$(5.40 \pm 0.67) \cdot 10^{-2}$**	**121±32**

1	14.2	$2.72 \cdot 10^3$	$3.88 \cdot 10^{-2}$	24.4
	14.0	$2.73 \cdot 10^3$	$3.83 \cdot 10^{-2}$	25.3
	14.2	$2.77 \cdot 10^3$	$3.93 \cdot 10^{-2}$	26.0
	11.8	$2.95 \cdot 10^3$	$3.48 \cdot 10^{-2}$	17.6
	11.8	$3.00 \cdot 10^3$	$3.54 \cdot 10^{-2}$	18.5
	11.6	$3.05 \cdot 10^3$	$3.53 \cdot 10^{-2}$	20.0
	11.3	$3.05 \cdot 10^3$	$3.44 \cdot 10^{-2}$	18.4
	11.0	$3.12 \cdot 10^3$	$3.41 \cdot 10^{-2}$	18.2
	11.2	$3.11 \cdot 10^3$	$3.47 \cdot 10^{-2}$	18.3
1 Average	**12.3±1.3**	**$(2.94\pm0.15) \cdot 10^3$**	**$(3.61\pm0.19) \cdot 10^{-2}$**	**20.7±3.3**
2	6.88	$7.16 \cdot 10^3$	$4.93 \cdot 10^{-2}$	37.8
	5.70	$8.64 \cdot 10^3$	$4.93 \cdot 10^{-2}$	33.8
	6.40	$8.46 \cdot 10^3$	$5.41 \cdot 10^{-2}$	34.7
	3.09	$1.49 \cdot 10^4$	$4.59 \cdot 10^{-2}$	6.02
	3.10	$1.48 \cdot 10^4$	$4.59 \cdot 10^{-2}$	6.05
	3.28	$1.48 \cdot 10^4$	$4.85 \cdot 10^{-2}$	5.99
	3.00	$1.50 \cdot 10^4$	$4.49 \cdot 10^{-2}$	6.15
	2.96	$1.55 \cdot 10^4$	$4.60 \cdot 10^{-2}$	6.16
	3.08	$1.58 \cdot 10^4$	$4.87 \cdot 10^{-2}$	6.29
2 Average	**4.17±1.56**	**$(1.28\pm0.34) \cdot 10^4$**	**$(4.81\pm0.27) \cdot 10^{-2}$**	**15.9±13.8**

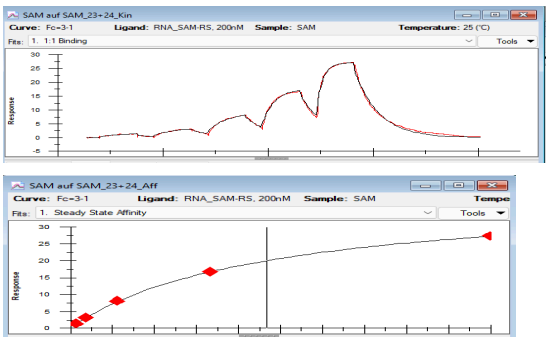
Table S2. Binding affinities and kinetic parameters from SPR measurements for $preQ_0$ and $preQ_1$ against the *Tte* $preQ_1$ riboswitch, 1:1 binding fit model.

***Tte* $preQ_1$**	**SPR**			
Compound	**K_D [nM]**	**k_{on} [1/Ms]**	**k_{off} [1/s]**	**$K_{D,SS}$ [nM]**
$preQ_0$	64.3	$3.18 \cdot 10^3$	$2.05 \cdot 10^{-4}$	633
	107	$3.13 \cdot 10^3$	$3.35 \cdot 10^{-4}$	769
	107	$3.11 \cdot 10^3$	$3.32 \cdot 10^{-4}$	761
	88.3	$3.17 \cdot 10^3$	$2.80 \cdot 10^{-4}$	939
	95.2	$3.81 \cdot 10^3$	$3.63 \cdot 10^{-4}$	754

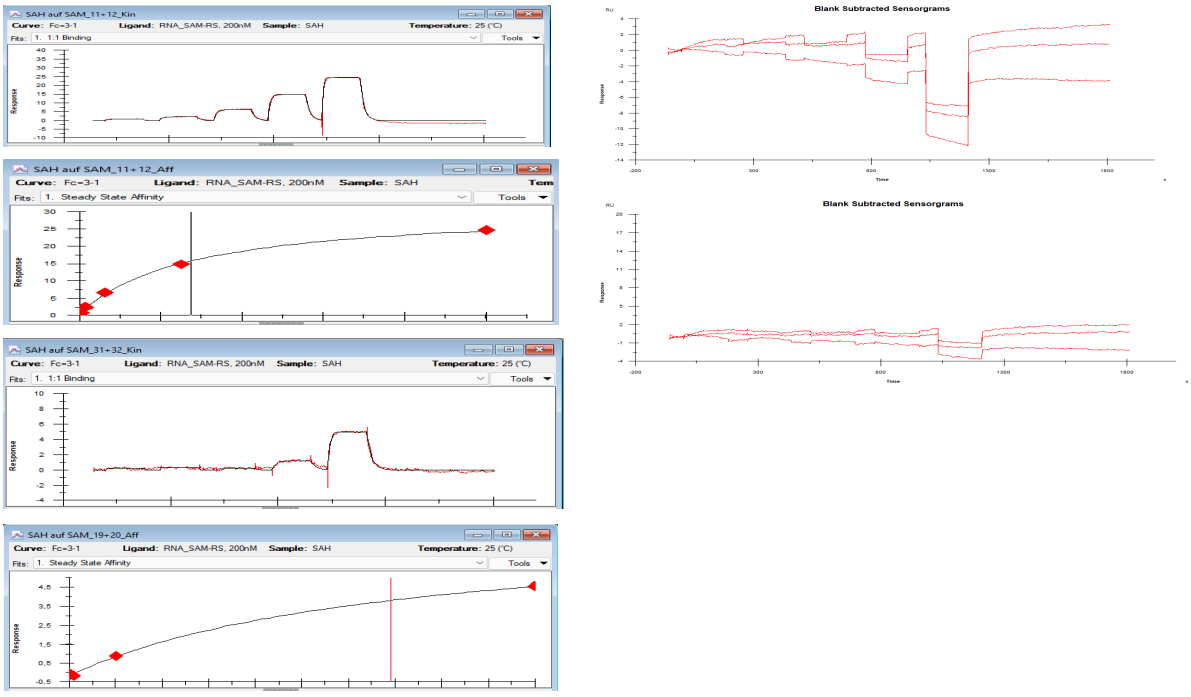
	89.8	$3.80 \cdot 10^{3}$	$3.41 \cdot 10^{-4}$	783
	53.2	$4.26 \cdot 10^{3}$	$2.26 \cdot 10^{-4}$	641
	85.5	$3.99 \cdot 10^{3}$	$3.41 \cdot 10^{-4}$	829
	83.1	$4.06 \cdot 10^{3}$	$3.38 \cdot 10^{-4}$	839
	21.7	$1.45 \cdot 10^{4}$	$3.14 \cdot 10^{-4}$	313
	19.4	$1.59 \cdot 10^{4}$	$3.08 \cdot 10^{-4}$	300
$preQ_0$ Average	**74.0±29.4**	**$(5.72±4.49) \cdot 10^{3}$**	**$(3.07±0.48) \cdot 10^{-4}$**	**687±197**
$preQ_1$	26.6	$1.04 \cdot 10^{4}$	$2.77 \cdot 10^{-4}$	389
	19.1	$1.05 \cdot 10^{4}$	$2.00 \cdot 10^{-4}$	410
	18.5	$1.03 \cdot 10^{4}$	$1.90 \cdot 10^{-4}$	431
	26.3	$9.60 \cdot 10^{3}$	$2.52 \cdot 10^{-4}$	474
	21.6	$1.00 \cdot 10^{4}$	$2.17 \cdot 10^{-4}$	446
	23.0	$9.65 \cdot 10^{3}$	$2.22 \cdot 10^{-4}$	528
	44.0	$9.81 \cdot 10^{3}$	$4.32 \cdot 10^{-4}$	396
	47.4	$9.99 \cdot 10^{3}$	$4.74 \cdot 10^{-4}$	455
	40.1	$1.00 \cdot 10^{4}$	$4.03 \cdot 10^{-4}$	482
	35.8	$1.15 \cdot 10^{4}$	$4.10 \cdot 10^{-4}$	309
	32.8	$1.20 \cdot 10^{4}$	$3.94 \cdot 10^{-4}$	297
	34.9	$1.17 \cdot 10^{4}$	$4.07 \cdot 10^{-4}$	329
	26.1	$1.12 \cdot 10^{4}$	$2.91 \cdot 10^{-4}$	329
	19.5	$1.19 \cdot 10^{4}$	$2.32 \cdot 10^{-4}$	300
	19.1	$1.18 \cdot 10^{4}$	$2.27 \cdot 10^{-4}$	327
	21.4	$1.02 \cdot 10^{4}$	$2.18 \cdot 10^{-4}$	736
	14.1	$1.19 \cdot 10^{4}$	$1.67 \cdot 10^{-4}$	712
	14.1	$1.19 \cdot 10^{4}$	$1.68 \cdot 10^{-4}$	768
	23.5	$9.61 \cdot 10^{3}$	$2.25 \cdot 10^{-4}$	838
	19.2	$1.14 \cdot 10^{4}$	$2.19 \cdot 10^{-4}$	777
	24.7	$1.01 \cdot 10^{4}$	$2.49 \cdot 10^{-4}$	914
	28.4	$1.11 \cdot 10^{4}$	$3.14 \cdot 10^{-4}$	265
	15.7	$1.23 \cdot 10^{4}$	$1.94 \cdot 10^{-4}$	254
	15.3	$1.20 \cdot 10^{4}$	$1.83 \cdot 10^{-4}$	284

	42.5	$9.29 \cdot 10^{3}$	$3.94 \cdot 10^{-4}$	593
	10.6	$1.12 \cdot 10^{4}$	$1.19 \cdot 10^{-4}$	579
	11.0	$1.09 \cdot 10^{4}$	$1.20 \cdot 10^{-4}$	668
	19.5	$1.05 \cdot 10^{4}$	$2.04 \cdot 10^{-4}$	607
	14.4	$1.32 \cdot 10^{4}$	$1.90 \cdot 10^{-4}$	572
	15.3	$1.31 \cdot 10^{4}$	$2.00 \cdot 10^{-4}$	619
preQ$_1$ Average	**24.1±9.9**	**$(1.10\pm0.11) \cdot 10^{4}$**	**$(2.60\pm0.96) \cdot 10^{-4}$**	**503±184**
preQ1 het?				

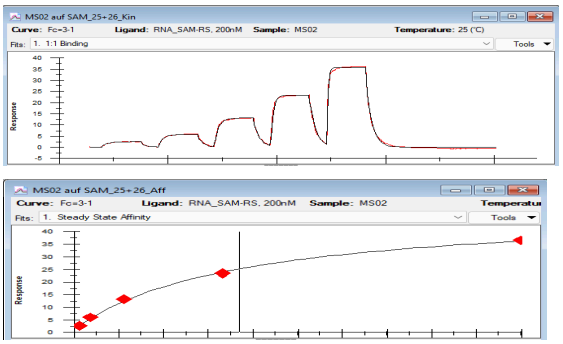
A

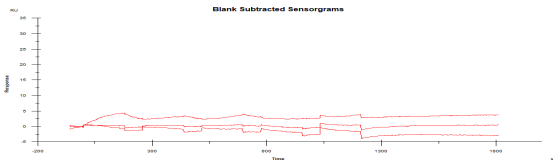


B

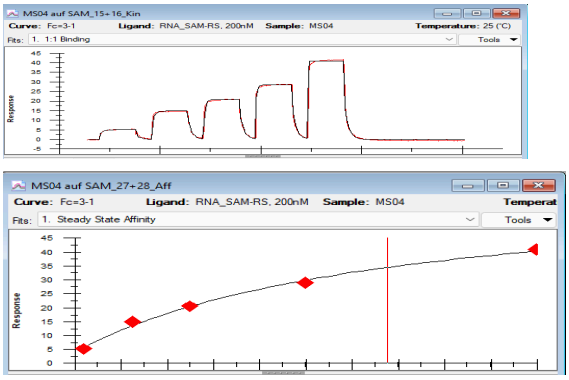


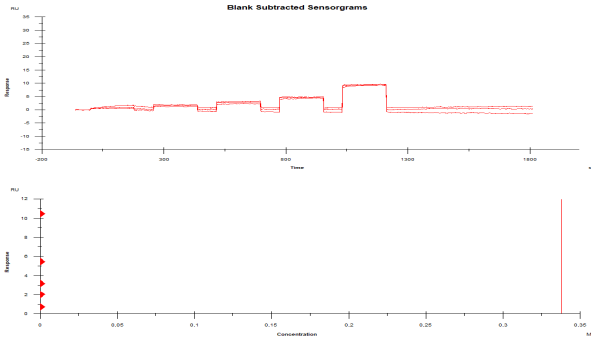
C


MS02 auf SAM_25+26_Kin
Curve: Fc=3-1 Ligand: RNA_SAM-RS, 200nM Sample: MS02 Temperature: 25 (°C)
Fits: 1. 1:1 Binding
Tools
Response
MS02 auf SAM_25+26_Aff
Curve: Fc=3-1 Ligand: RNA_SAM-RS, 200nM Sample: MS02
Fits: 1. Steady State Affinity
Tools
Response

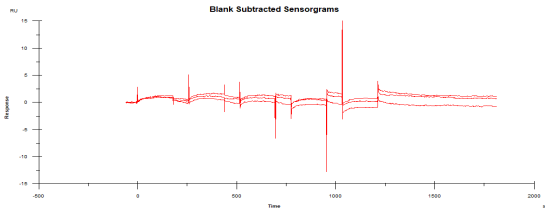

Blank Subtracted Sensorgrams

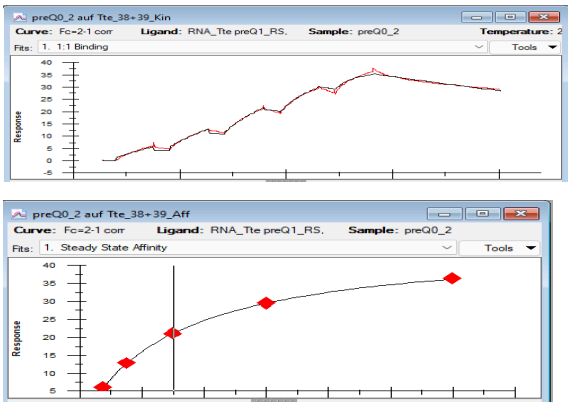
D


MS04 auf SAM_15+16_Kin
Curve: Fc=3-1 Ligand: RNA_SAM-RS, 200nM Sample: MS04 Temperature: 25 (°C)
Fits: 1. 1:1 Binding
Tools
Response
MS04 auf SAM_27+28_Aff
Curve: Fc=3-1 Ligand: RNA_SAM-RS, 200nM Sample: MS04
Fits: 1. Steady State Affinity
Tools
Response

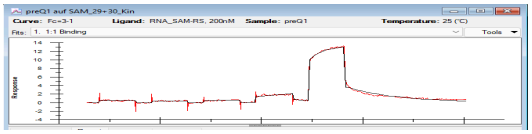

Blank Subtracted Sensorgrams
Concentration

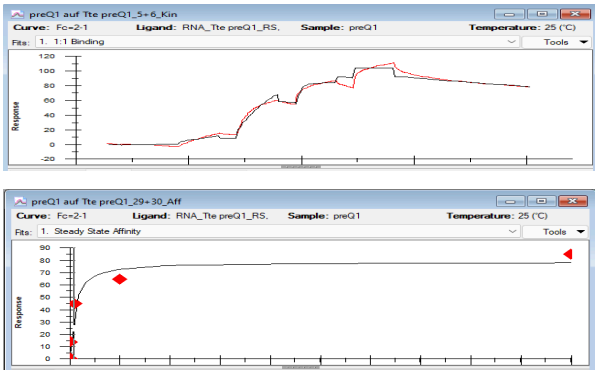
E


Blank Subtracted Sensorgrams

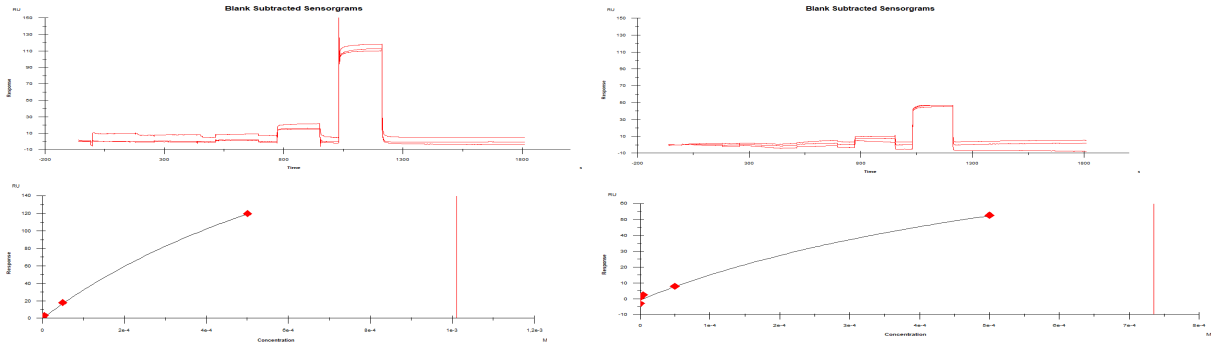

preQ0_2 auf Tte_38+39_Kin
Curve: Fc=2-1 corr Ligand: RNA_Tte preQ1_RS, Sample: preQ0_2 Temperature: 2
Fits: 1. 1:1 Binding
Tools
Response
preQ0_2 auf Tte_38+39_Aff
Curve: Fc=2-1 corr Ligand: RNA_Tte preQ1_RS, Sample: preQ0_2
Fits: 1. Steady State Affinity
Tools
Response

F

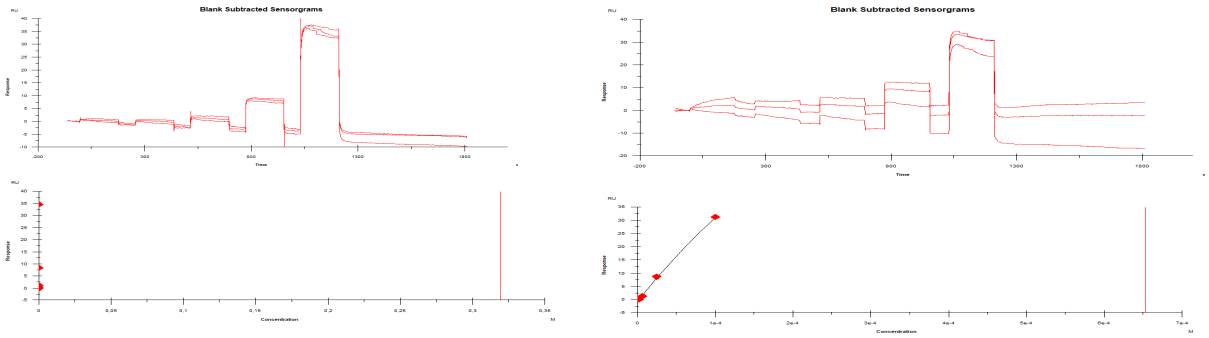

preQ1 auf SAM_29+30_Kin
Curve: Fc=3-1 Ligand: RNA_SAM-RS, 200nM Sample: preQ1 Temperature: 25 (°C)
Fits: 1. 1:1 Binding
Tools
Response


preQ1 auf Tte preQ1_5+6_Kin
Curve: Fc=2-1 Ligand: RNA_Tte preQ1_RS, Sample: preQ1 Temperature: 25 (°C)
Fits: 1. 1:1 Binding
Tools
Response
preQ1 auf Tte preQ1_29+30_Aff
Curve: Fc=2-1 Ligand: RNA_Tte preQ1_RS, Sample: preQ1 Temperature: 25 (°C)
Fits: 1. Steady State Affinity
Tools
Response

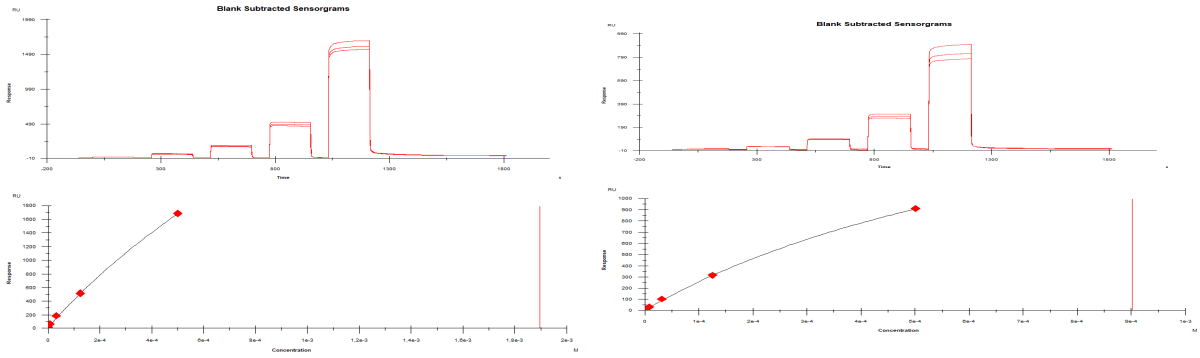
G


Blank Subtracted Sensorgrams
Blank Subtracted Sensorgrams
Concentration
Concentration

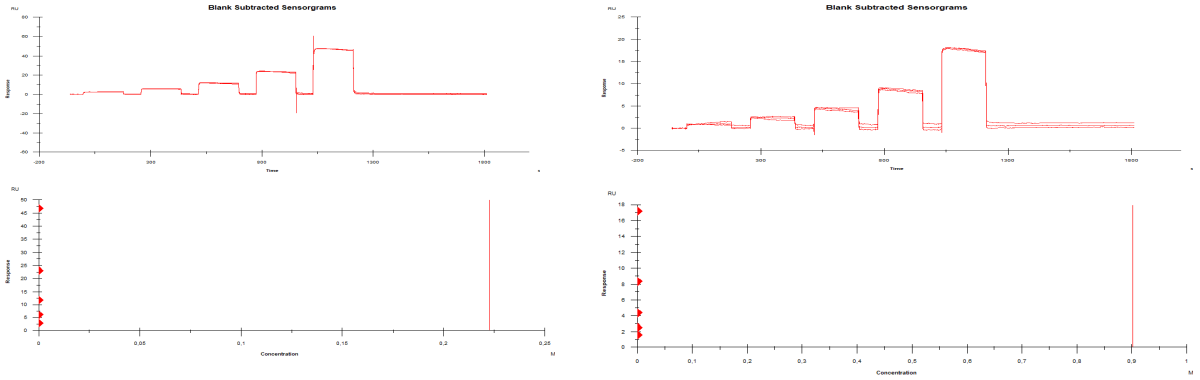
H


Blank Subtracted Sensorgrams
Blank Subtracted Sensorgrams
Concentration
Concentration

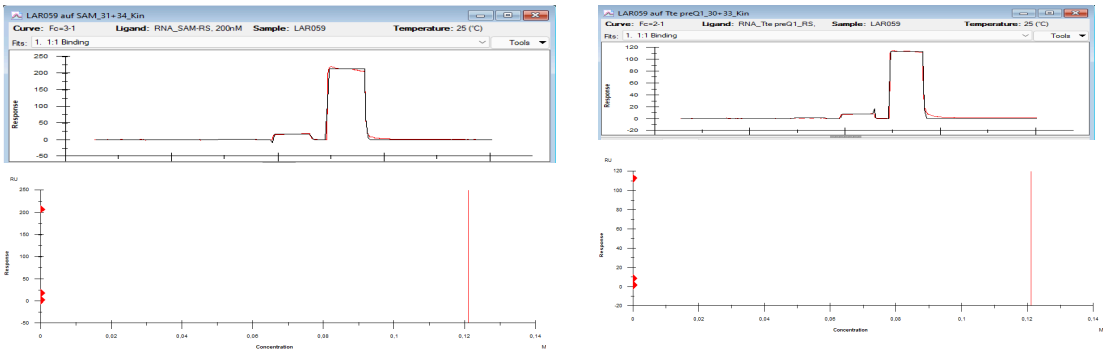
I


Blank Subtracted Sensorgrams
Blank Subtracted Sensorgrams
Concentration
Concentration

J


Blank Subtracted Sensorgrams
Blank Subtracted Sensorgrams
Concentration
Concentration

K



L

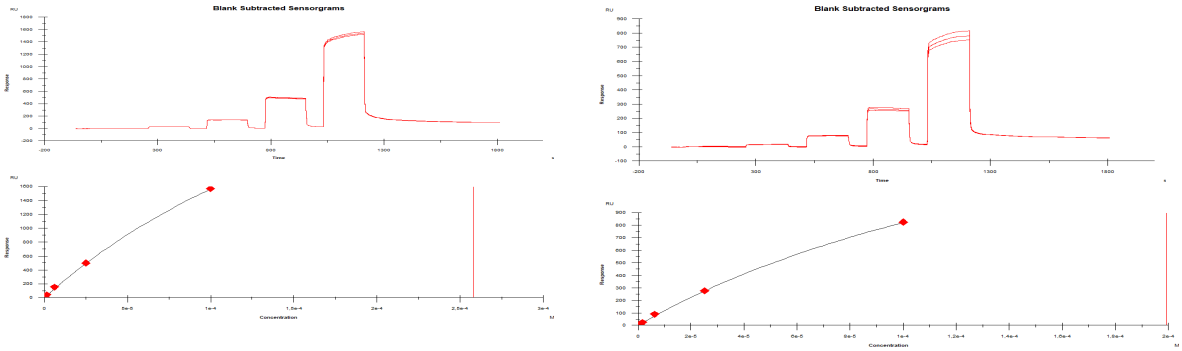
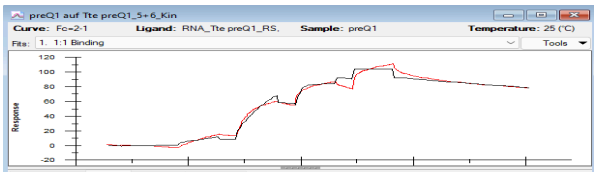
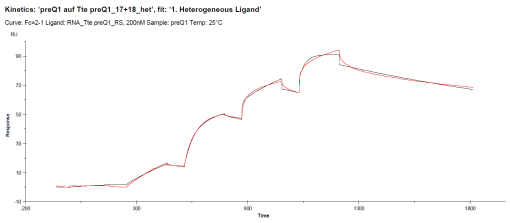


Figure S1. Representative SPR sensorgrams for SAM (**A**), SAH (**B**), **1** (**C**), **2** (**D**), preQ$_0$ (**E**), preQ$_1$ (**F**), **3** (**G**), **4** (**H**), **5** (**I**), **6** (**J**), **7** (**K**) and **8** (**L**) binding to the *Ba* SAM-VI riboswitch (left panel) and binding to the *Tte* preQ$_1$ riboswitch (right panel). black; red 1:1 binding mode fit.

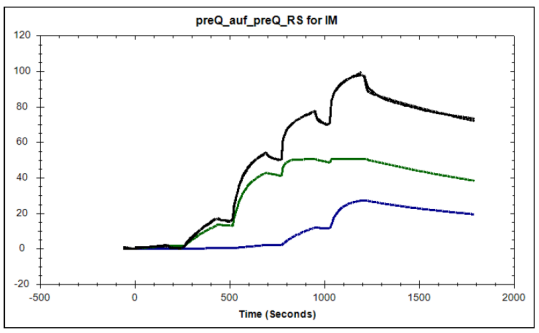
A



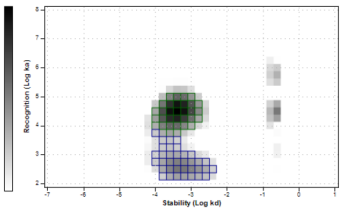
B



C



Description:
Crv loaded Scaled all curves with factor 1,062 Data point reduction 1,510% 2024-08-14 16:33 File IMData_8103.txt
Regularization factor = 4
tolerance 1,0E-4 reached, exiting.



Peak color	Id	Pos. ka	Pos. kd	KD	Weight (%)	Width ka	Width kd
DarkGreen	0	2,77e4	3,99e-4	1,44e-8	50,73	1,26e4 - 8,91e4	1,66e-4 - 1,00e-3
DarkBlue	1	5,10e2	4,84e-4	9,50e-7	29,57	1,50e2 - 1,50e3	1,91e-4 - 1,82e-3

Figure S2. Representative SPR sensorgram for preQ$_1$ binding to the *Tte* preQ$_1$ riboswitch analyzed with a 1:1 binding mode fit (**A**) or a heterogeny target model fit (**B**). IM analysis of the same dataset (**C**).

2. MST Experiments

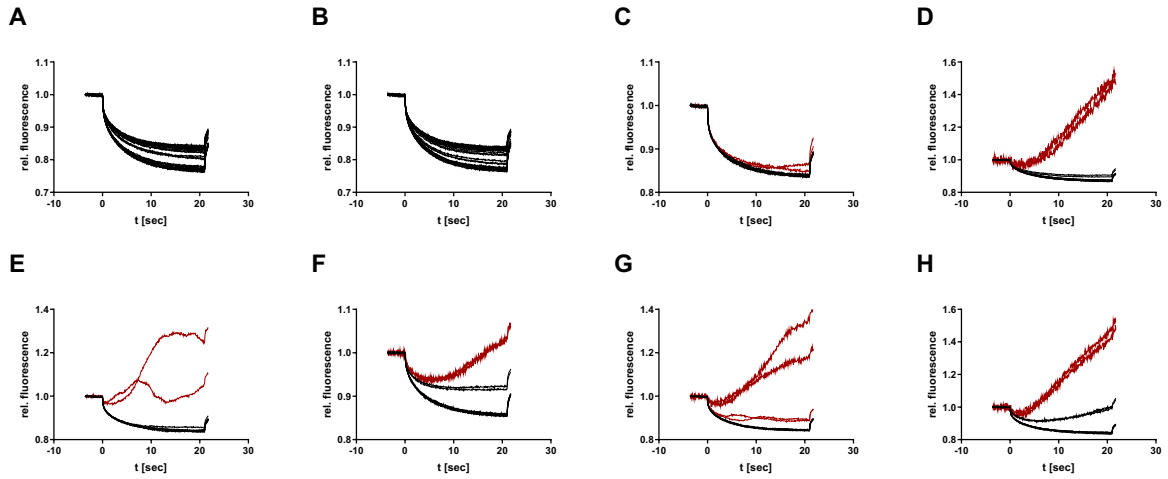


Figure S3. MST results for preQ$_1$ (**A**), preQ$_0$ (**B**), **3** (**C**), **4** (**D**), **5** (**E**), **6** (**F**), **7** (**G**), and **8** (**H**) binding to the *Tte* preQ$_1$ riboswitch. Aggregation at high concentrations is colored in red.

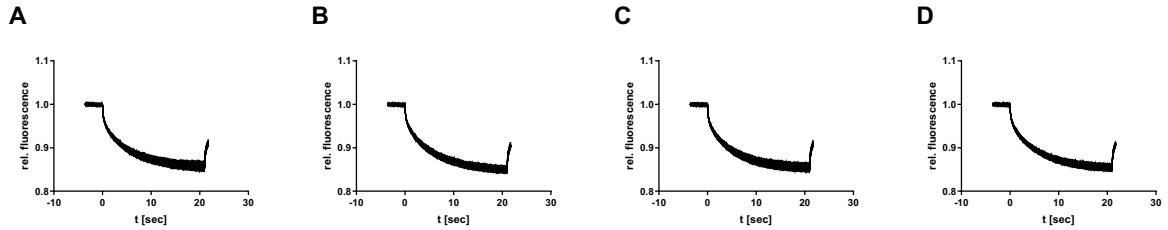


Figure S4. MST results for SAM (**A**), SAH (**B**), **1** (**C**), and **2** (**D**) binding to the *Ba* SAM-VI riboswitch.

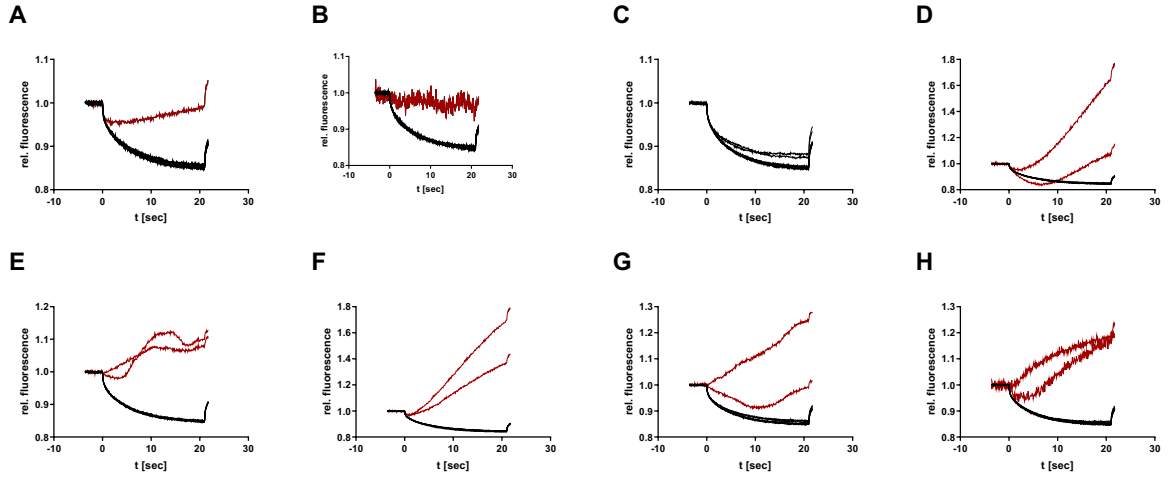


Figure S5. MST selectivity results at concentrations of 1 nM–1 mM for preQ$_1$ (**A**), preQ$_0$ (**B**), **3** (**C**), **4** (**D**), **5** (**E**), **6** (**F**), **7** (**G**), and **8** (**H**) binding to the *Ba* SAM-VI riboswitch. Aggregation at high concentrations is colored in red.

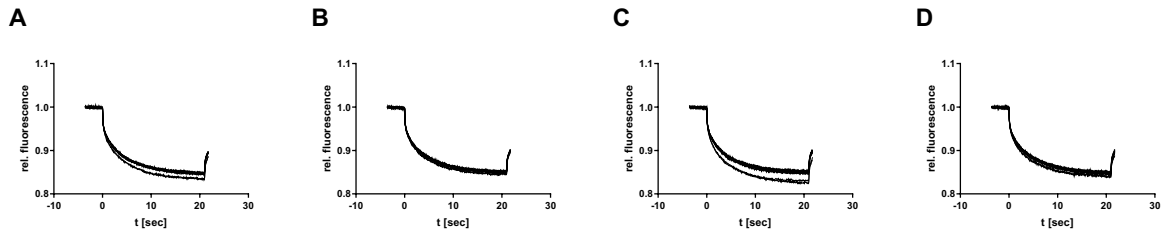


Figure S6. MST selectivity results at concentrations of 1 nM–1 mM for SAM (**A**), SAH (**B**), **1** (**C**), and **2** (**D**) binding to the *Tte* preQ$_1$ riboswitch.

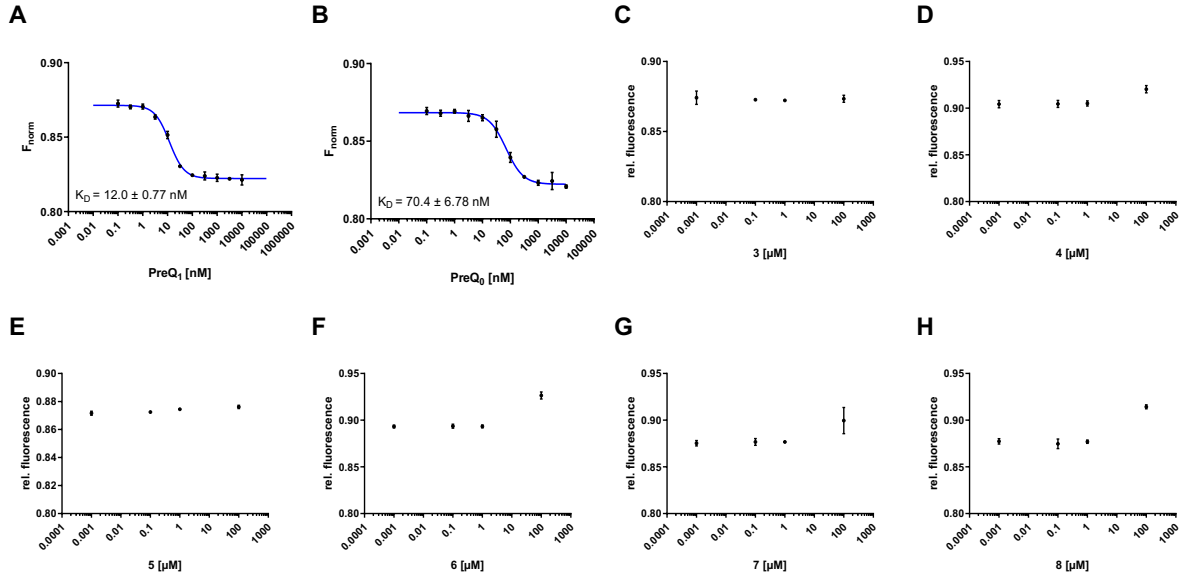


Figure S7. MST dose-response results for preQ$_1$ (**A**), preQ$_0$ (**B**), **3** (**C**), **4** (**D**), **5** (**E**), **6** (**F**), **7** (**G**), and **8** (**H**) binding to the *Tte* preQ$_1$ riboswitch.

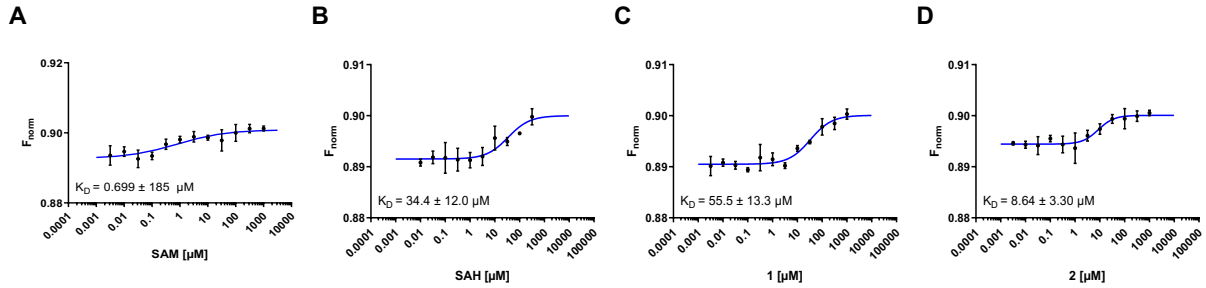


Figure S8. MST dose-response curves for SAM (**A**), SAH (**B**), **1** (**C**), and **2** (**D**) binding to the *Ba* SAM-VI riboswitch.

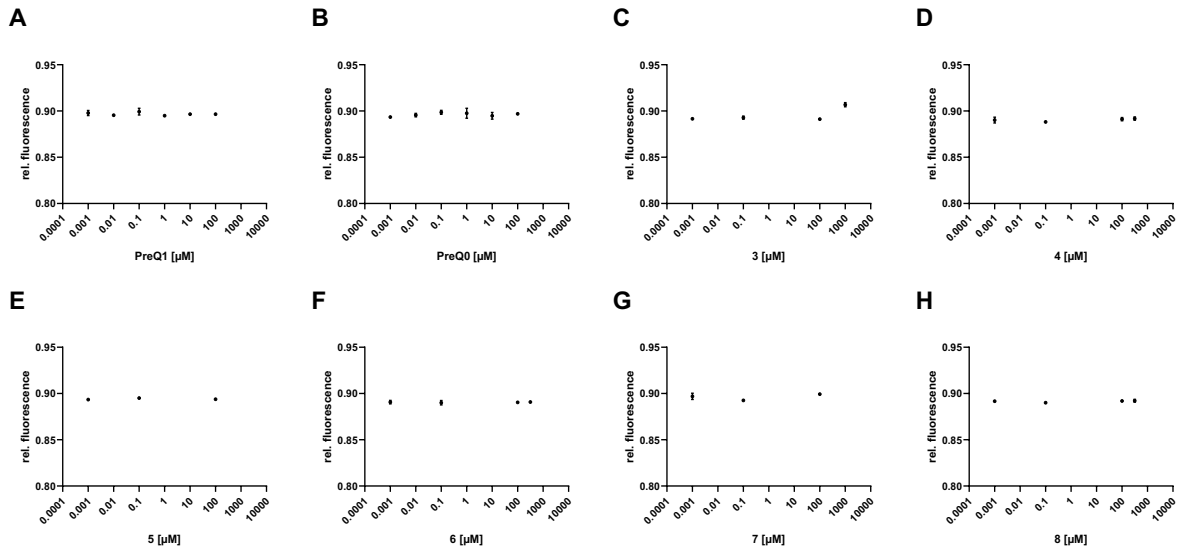


Figure S9. MST selectivity dose-response results for $preQ_1$ (**A**), $preQ_0$ (**B**), **3** (**C**), **4** (**D**), **5** (**E**), **6** (**F**), **7** (**G**), and **8** (**H**) not binding to the *Ba* SAM-VI riboswitch.

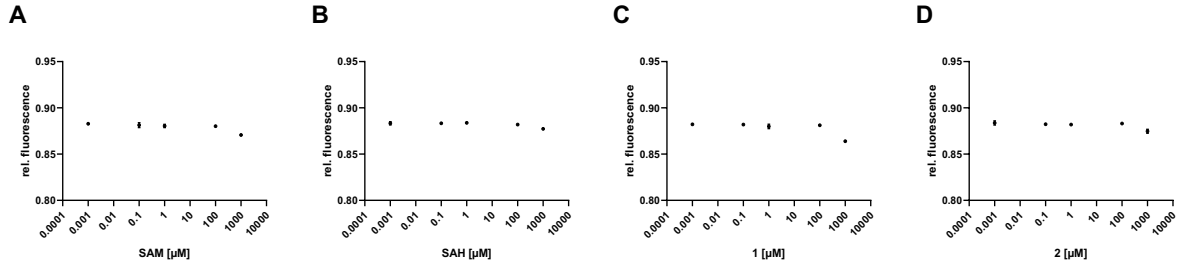


Figure S10. MST selectivity dose-response results for SAM (**A**), SAH (**B**), **1** (**C**), and **2** (**D**) not binding to the *Tte* $preQ_1$ riboswitch.

3. ITC Experiments

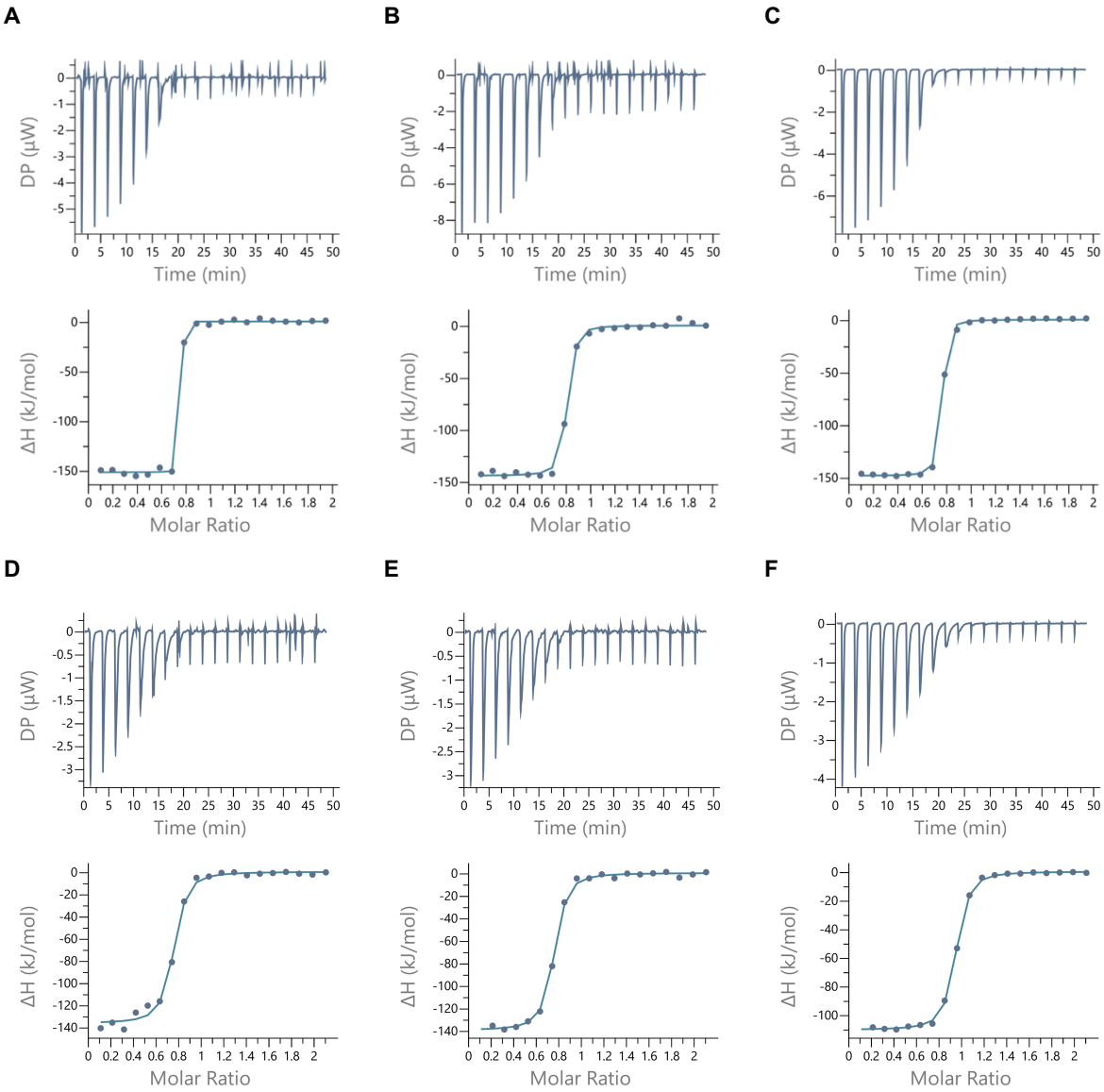


Figure S11. Thermograms (top and isotherms) from ITC experiments for preQ$_1$ (**A–C**), preQ$_0$ (**D–F**), **3** (**G**), **4** (**H**), **5** (**I**), **6** (**J**), **7** (**K**), and **8** (**L**) binding to the *Tte* preQ$_1$ riboswitch. For compounds **3**–**8**, no binding signals during titration were observed.

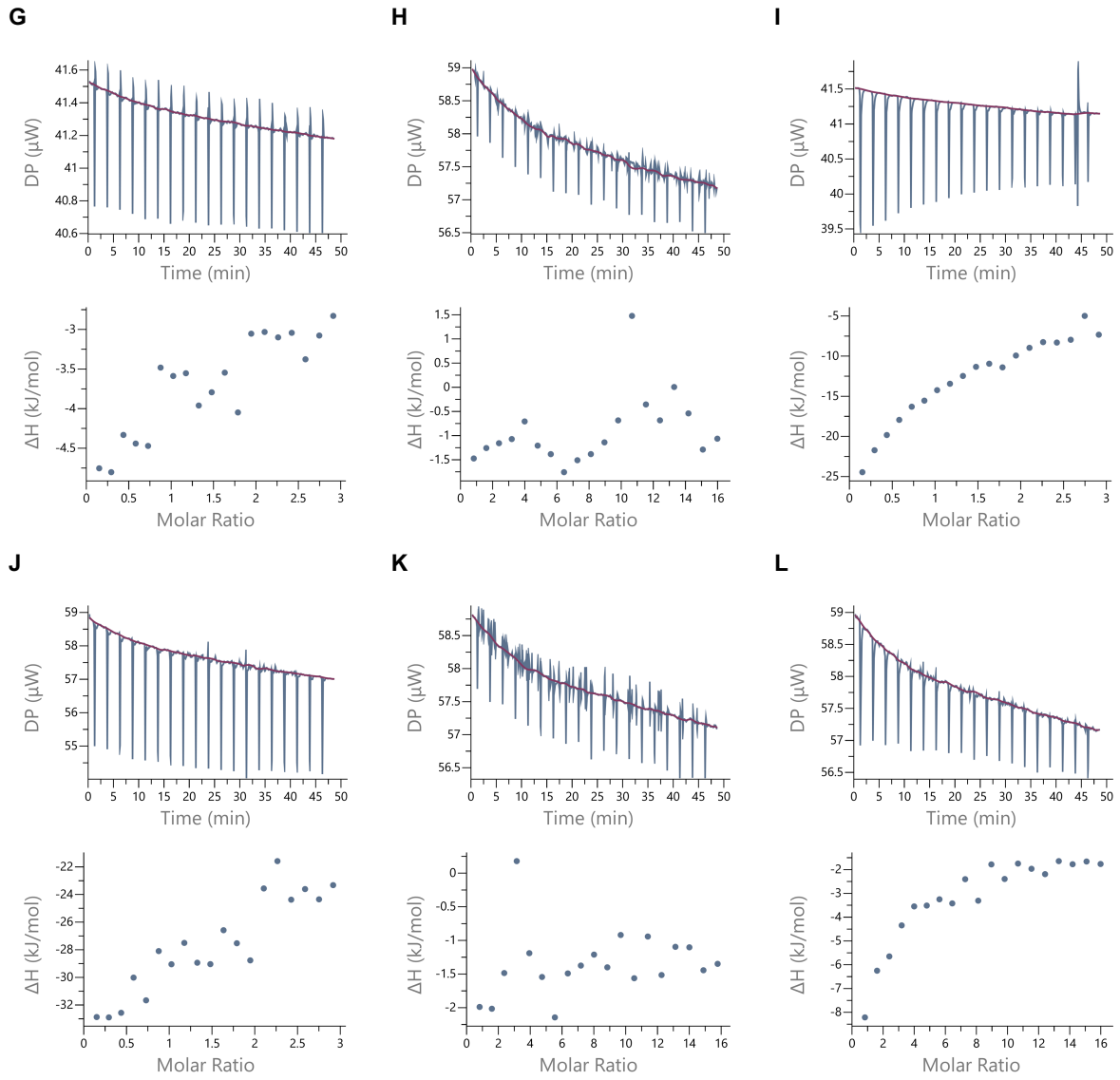


Figure S12. Thermograms (top and isotherms) from ITC experiments for preQ$_1$ (**A–C**), preQ$_0$ (**D–F**), **3** (**G**), **4** (**H**), **5** (**I**), **6** (**J**), **7** (**K**), and **8** (**L**) binding to the *Tte* preQ$_1$ riboswitch. For compounds **3**–**8**, no binding signals during titration were observed.

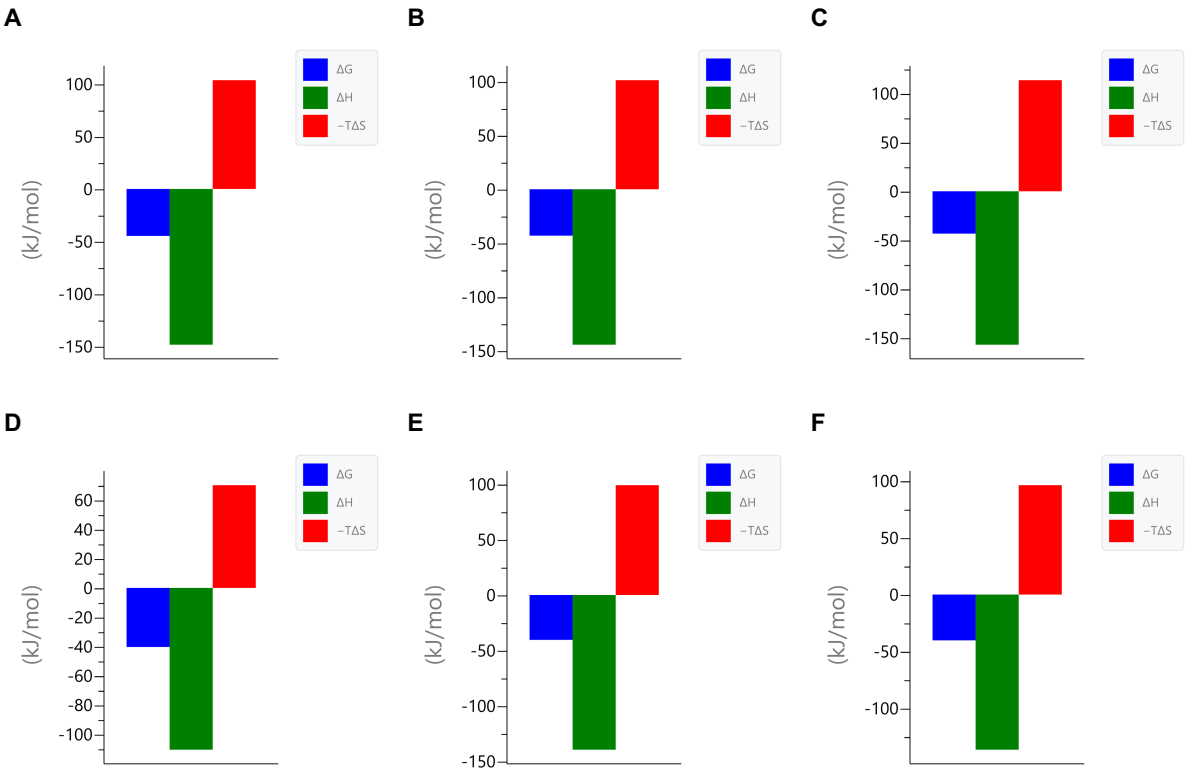
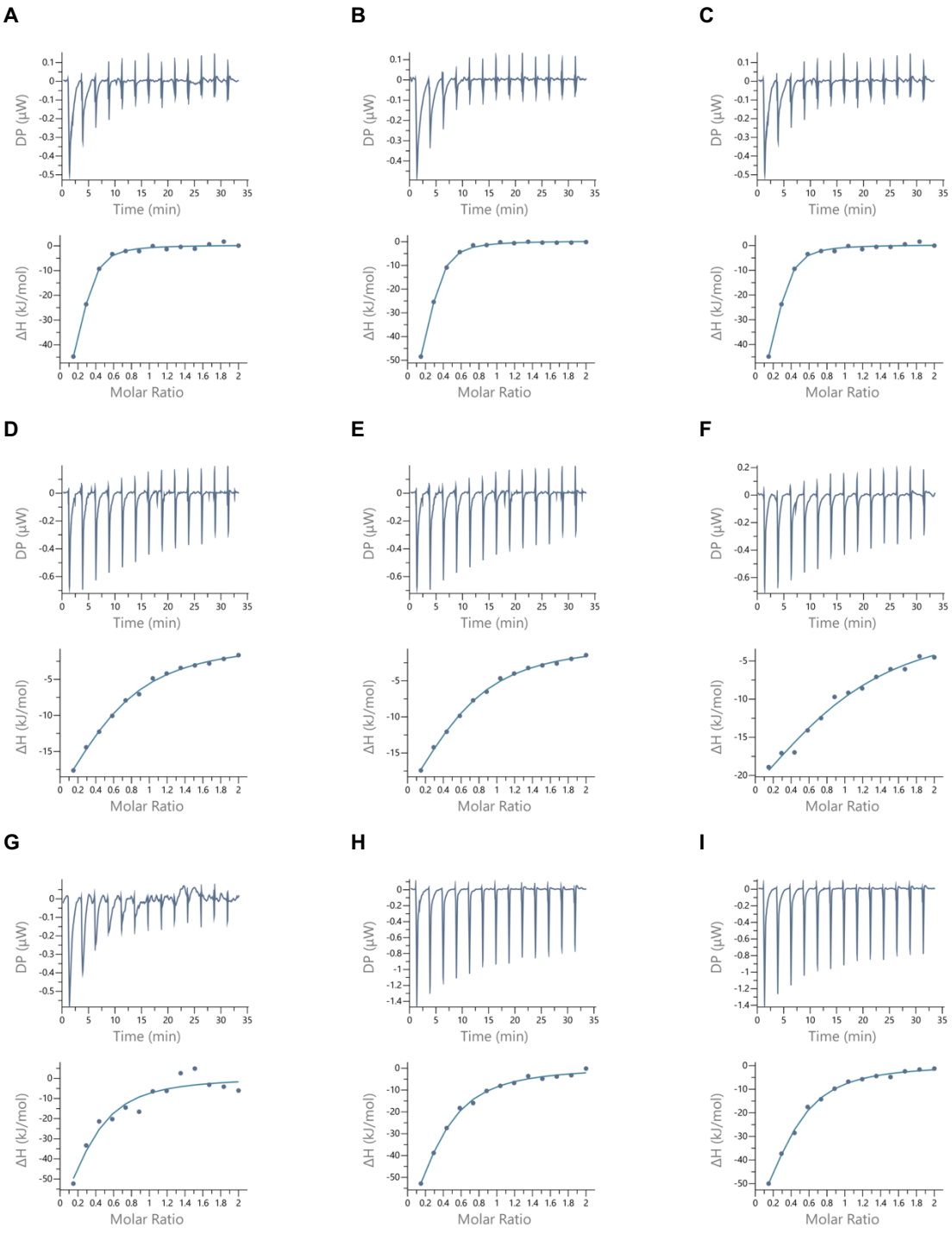


Figure S13. Thermodynamic signature plots of preQ$_1$ (**A**) and preQ$_0$ (**B**) binding to the *Tte* preQ$_1$ riboswitch.


A
B
C
D
E
F
G
H
I
DP (µW)
Time (min)
ΔH (kJ/mol)
Molar Ratio

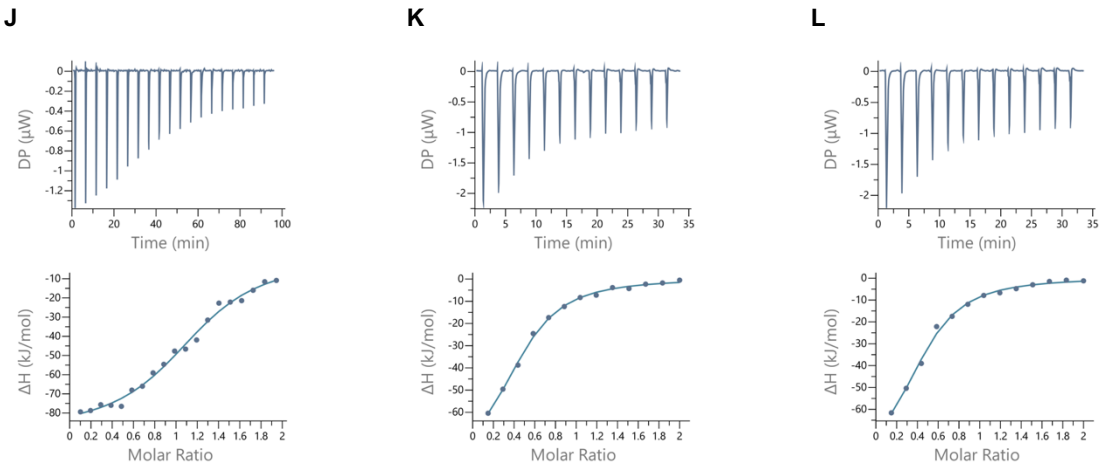


Figure S14. Thermograms (top and isotherms) from ITC experiments for SAM (**A–C**), SAH (**D–F**), **1** (**G–I**), and **2** (**J–L**) binding to the *Ba* SAM-VI riboswitch.

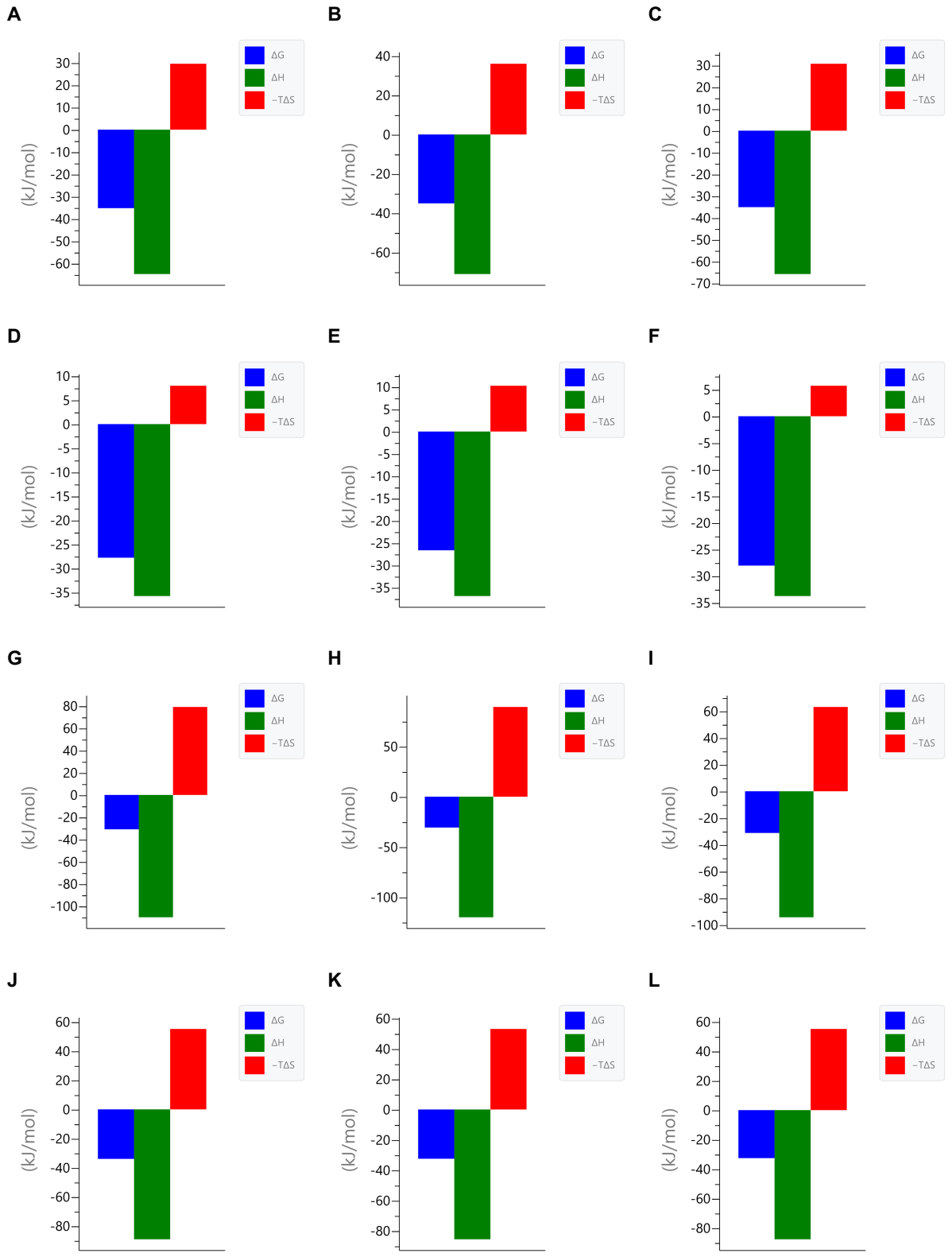


Figure S15. Thermodynamic signature plots of AM (**A**), SAH (**B**), **1** (**C**), and **2** (**D**) binding to the *Ba* SAM-VI riboswitch.

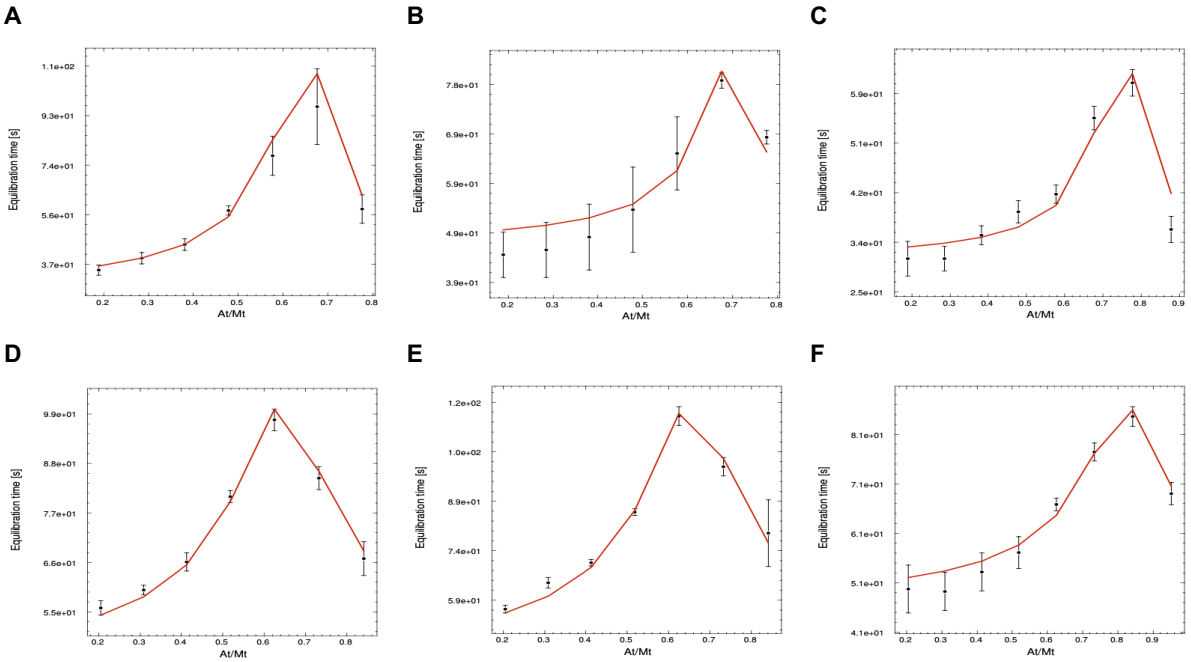


Figure S16. Equilibration-time curves (ETC) from kinITC analysis for preQ$_1$ (**A**–**B:** injections 2–8, **C**: injections 2–9) and preQ$_0$ (**D**–**E**: injections 2-8, **F**: injections 2–9) binding to the *Tte* preQ$_1$ riboswitch.

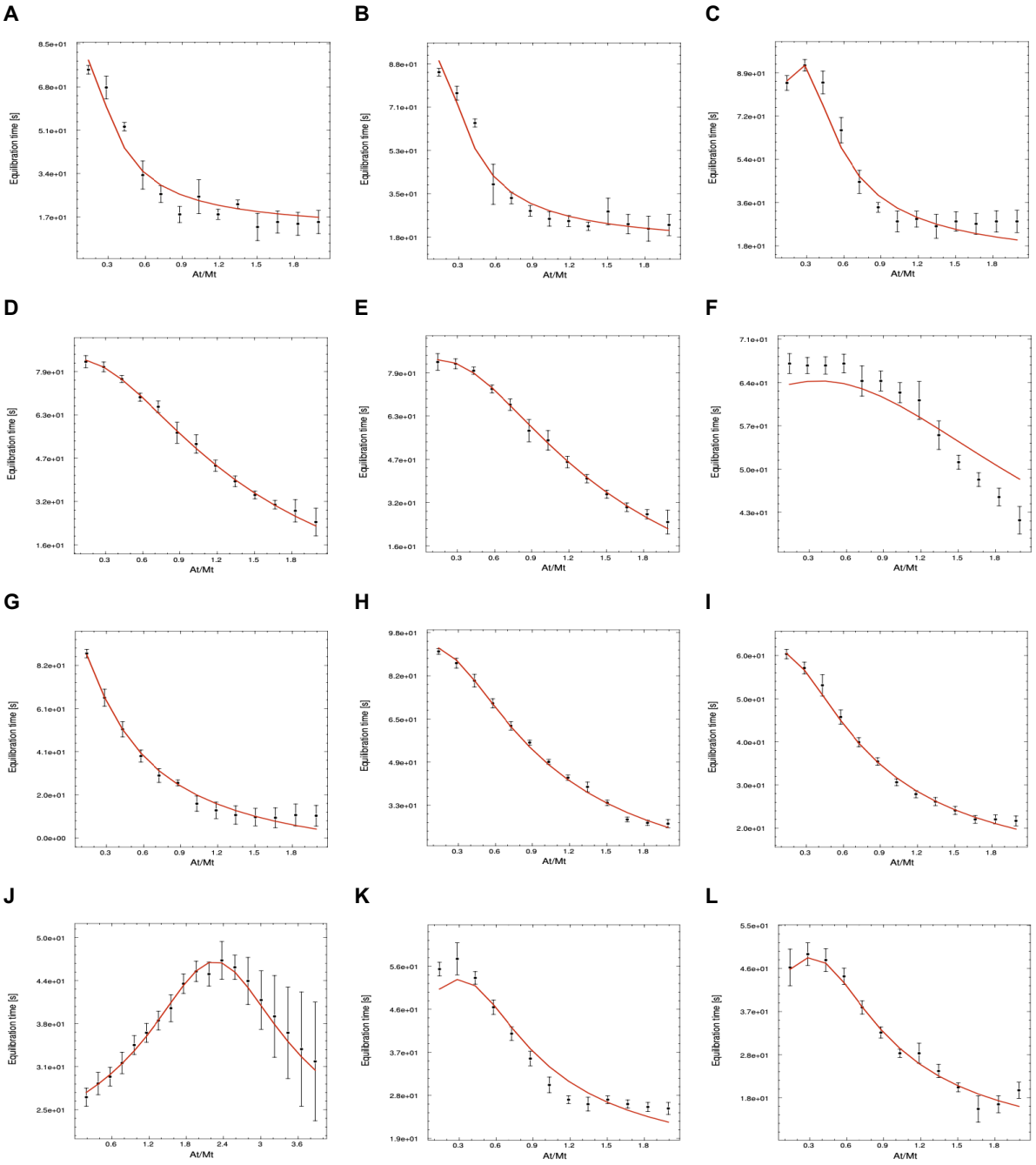


Figure S17. ETCs from kinITC analysis for SAM (**A**–**C**), SAH (**D**–**F**), **1** (**G**–**I**), and **2** (**J**–**L**) binding to the *Ba* SAM-VI riboswitch.

Table S3. Thermodynamic analysis results from ITC measurements for SAM, SAH, **1**, and **2** against the *Ba* SAM-VI riboswitch.

Ligand	K_D [µM]	ΔG [kJ/mol]	ΔH [kJ/mol]	–TΔS [kJ/mol]
SAM	0.688	–35.2	–64.8	29.6
	0.729	–35.1	–71.1	36.0
	0.720	–35.1	–65.8	30.7
SAM Average	**0.712 ± 0.022**	**–35.1 ± 0.1**	**–67.4 ± 3.7**	**32.1 ± 3.4**
SAH	13.7	–27.8	–35.8	7.98
	12.4	–28.0	–33.7	5.72
	18.8	–27.0	–36.9	9.90
SAH Average	**15.0 ± 3.4**	**–27.6 ± 0.5**	**–35.5 ± 1.6**	**7.87 ± 2.09**
1	3.90	–30.9	–110	79.1
	4.25	–30.7	–120	89.3
	3.44	–31.2	–94.3	63.1
1 Average	**3.86 ± 0.41**	**–30.9 ± 0.3**	**–108 ± 13**	**77.2 ± 13.2**
2	1.14	–34.0	–89.1	55.2
	2.07	–32.5	–85.6	53.1
	1.97	–32.6	–87.7	55.1
2 Average	**1.73 ± 0.51**	**–33.0 ± 0.8**	**–87.5 ± 1.8**	**54.5 ± 1.2**

Table S4. Kinetic analysis results from ITC measurements for SAM, SAH, **1**, and **2** against the *Ba* SAM-VI riboswitch.

Ligand	k_{on} [1/Ms]	k_{off} [1/s]	Chi^2
SAM	$(2.50 \pm 0.08) \cdot 10^4$	$(1.72 \pm 0.06) \cdot 10^{-2}$	3.60
	$(2.08 \pm 0.06) \cdot 10^4$	$(1.51 \pm 0.04) \cdot 10^{-2}$	4.37
	$(1.49 \pm 0.05) \cdot 10^4$	$(1.07 \pm 0.04) \cdot 10^{-2}$	1.57
SAM Average	$\mathbf{(2.02 \pm 0.51) \cdot 10^4}$	$\mathbf{(1.44 \pm 0.33) \cdot 10^{-2}}$	
SAH	$(4.19 \pm 0.04) \cdot 10^3$	$(5.75 \pm 0.06) \cdot 10^{-2}$	0.19
	$(1.77 \pm 0.02) \cdot 10^3$	$(3.33 \pm 0.03) \cdot 10^{-2}$	4.98
	$(4.32 \pm 0.03) \cdot 10^3$	$(5.36 \pm 0.04) \cdot 10^{-2}$	0.18
SAH Average	$\mathbf{(3.43 \pm 1.44) \cdot 10^3}$	$\mathbf{(4.81 \pm 1.30) \cdot 10^{-2}}$	
1	$(1.64 \pm 0.03) \cdot 10^4$	$(6.39 \pm 0.12) \cdot 10^{-2}$	0.46
	$(0.88 \pm 0.02) \cdot 10^4$	$(3.73 \pm 0.10) \cdot 10^{-2}$	1.44
	$(1.04 \pm 0.03) \cdot 10^4$	$(3.57 \pm 0.09) \cdot 10^{-2}$	0.60
1 Average	$\mathbf{(1.40 \pm 0.46) \cdot 10^4}$	$\mathbf{(5.57 \pm 1.60) \cdot 10^{-2}}$	
2	$(1.25 \pm 0.08) \cdot 10^4$	$(1.42 \pm 0.09) \cdot 10^{-2}$	0.13
	$(1.42 \pm 0.05) \cdot 10^4$	$(2.95 \pm 0.10) \cdot 10^{-2}$	4.44
	$(1.41 \pm 0.07) \cdot 10^4$	$(2.78 \pm 0.13) \cdot 10^{-2}$	0.74
2 Average	$\mathbf{(1.36 \pm 0.10) \cdot 10^4}$	$\mathbf{(2.38 \pm 0.84) \cdot 10^{-2}}$	

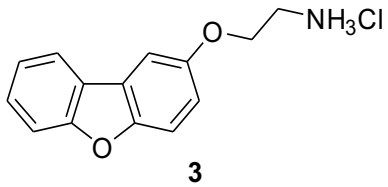
Table S5. Thermodynamic analysis results from ITC measurements for $preQ_1$ and $preQ_0$ against the *Tte* $preQ_1$ riboswitch.

Ligand	K_D [nM]	ΔG [kJ/mol]	ΔH [kJ/mol]	–TΔS [kJ/mol]
$PreQ_1$	26.0	–43.3	–157	114
	28.7	–43.1	–144	101
	14.7	–44.8	–148	104
$PreQ_1$ Average	**23.1 ± 7.4**	**–43.7 ± 0.9**	**–150 ± 7**	**106 ± 7**
$PreQ_0$	88.5	–40.3	–137	96.2
	81.9	–40.5	–140	99.1
	85.9	–40.4	–111	70.2
$PreQ_0$ Average	**85.4 ± 3.3**	**–40.4 ± 0.1**	**–129 ± 16**	**88.5 ± 15.9**

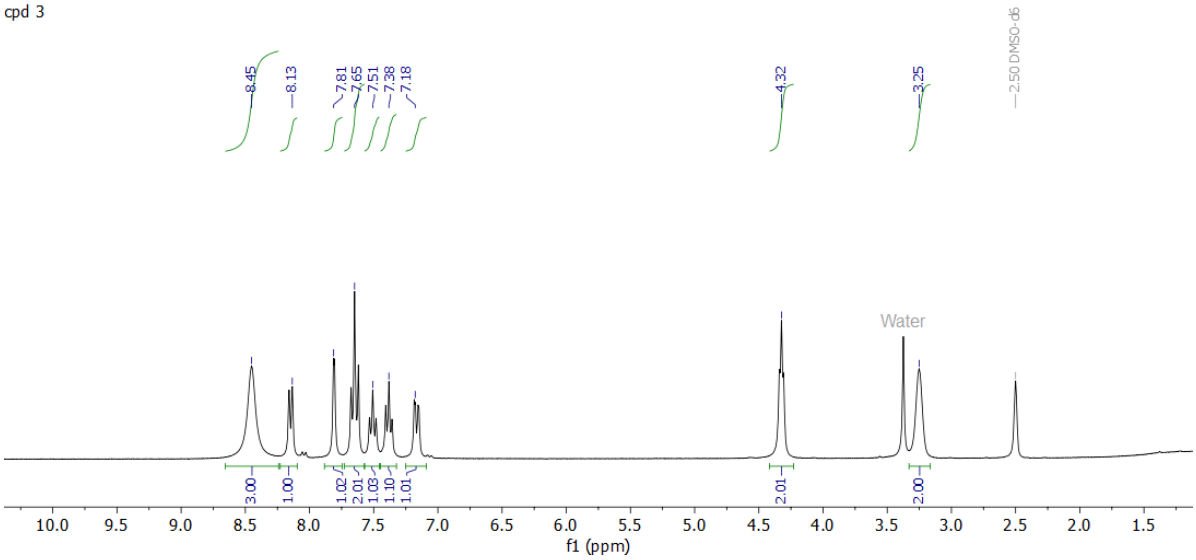
Table S6. Kinetic analysis results from ITC measurements for $preQ_1$ and $preQ_0$ against the *Tte* $preQ_1$ riboswitch.

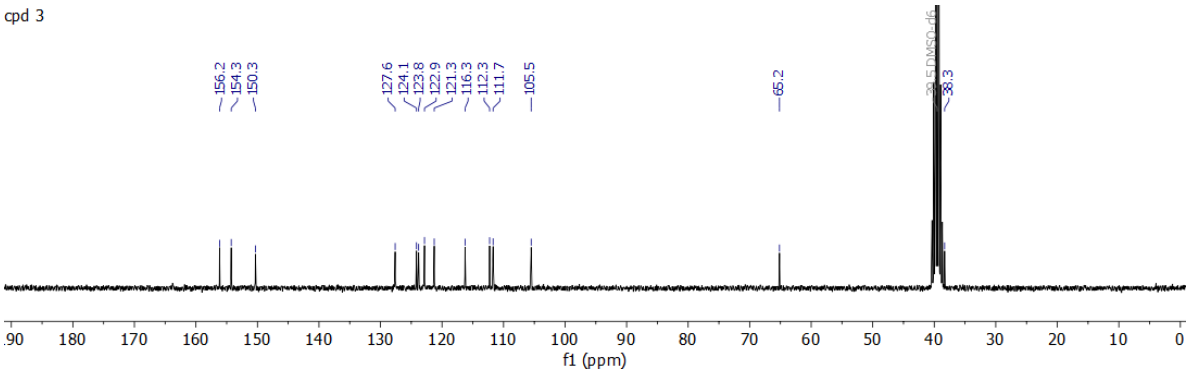
Ligand	k_{on} [1/Ms]	k_{off} [1/s]	Chi^2
$PreQ_1$	$(4.10 \pm 0.37) \cdot 10^4$	$(1.07 \pm 0.10) \cdot 10^{-3}$	0.61
	$(8.33 \pm 0.71) \cdot 10^4$	$(2.39 \pm 0.20) \cdot 10^{-3}$	1.37
	$(9.59 \pm 0.75) \cdot 10^4$	$(1.41 \pm 0.11) \cdot 10^{-3}$	1.83
$PreQ_1$ Average	**$(7.34 \pm 2.88) \cdot 10^4$**	**$(1.62 \pm 0.69) \cdot 10^{-3}$**	
$PreQ_0$	$(3.41 \pm 0.13) \cdot 10^4$	$(3.01 \pm 0.12) \cdot 10^{-3}$	1.12
	$(2.64 \pm 0.08) \cdot 10^4$	$(2.16 \pm 0.06) \cdot 10^{-3}$	2.52
	$(3.80 \pm 0.30) \cdot 10^4$	$(3.26 \pm 0.26) \cdot 10^{-3}$	0.71
$PreQ_0$ Average	**$(3.28 \pm 0.59) \cdot 10^4$**	**$(2.81 \pm 0.58) \cdot 10^{-3}$**	

4. Analytical data compounds 3–8



3





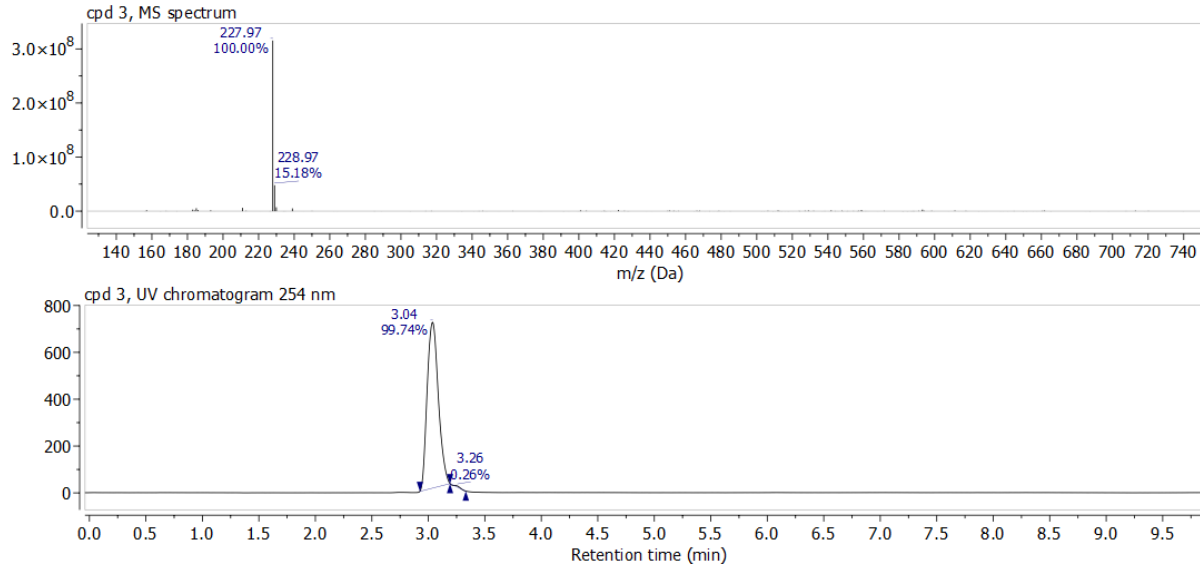
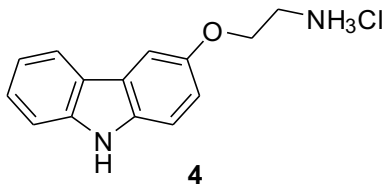
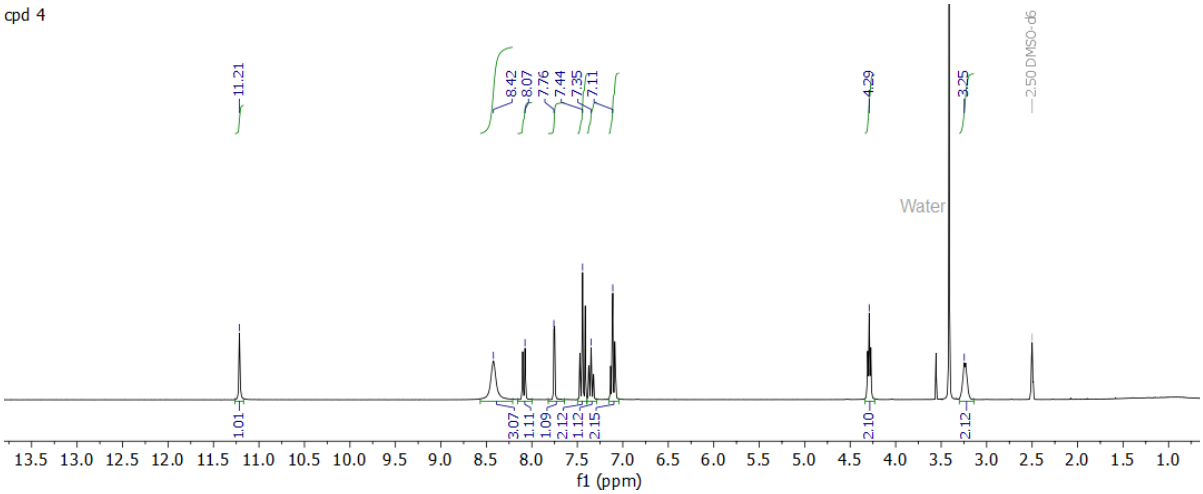
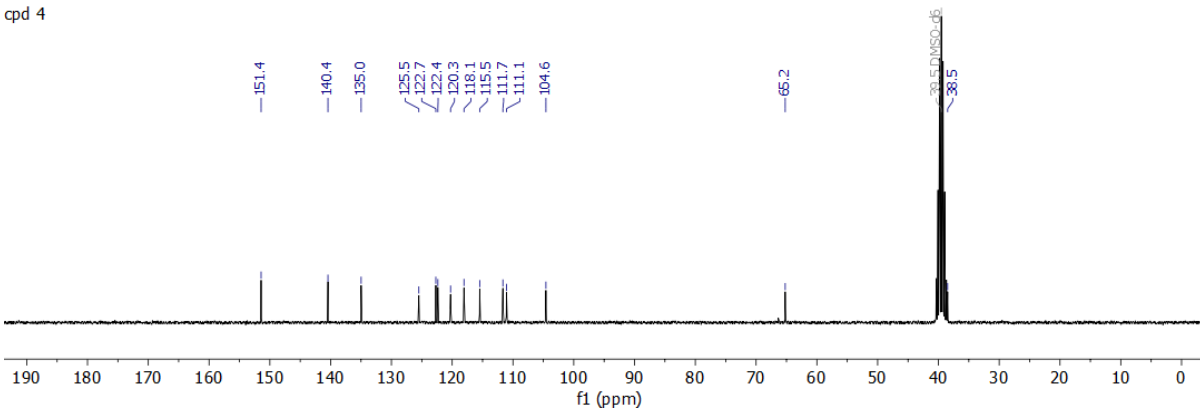


Figure S18. From top: ^{1}H-, ^{13}C-NMR spectra, ESI-MS spectrum and LC UV chromatogram at 254 nm of compound **3**. LC-ESI-MS measured with Zorbax column, isocratic 40% acetonitrile in water + 0.1% formic acid over 10 min.


O
NH3Cl
N
H

4


cpd 4
11.21
8.42
8.07
7.76
7.44
7.35
7.11
4.29
3.25
2.50 DMSO-d6
Water
1.01
3.07
1.11
1.09
2.12
1.12
2.15
2.10
2.12
13.5 13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
f1 (ppm)


cpd 4
151.4
140.4
135.0
125.5
122.7
122.4
120.3
118.1
115.5
111.7
111.1
104.6
65.2
39.5 DMSO-d6
38.5
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

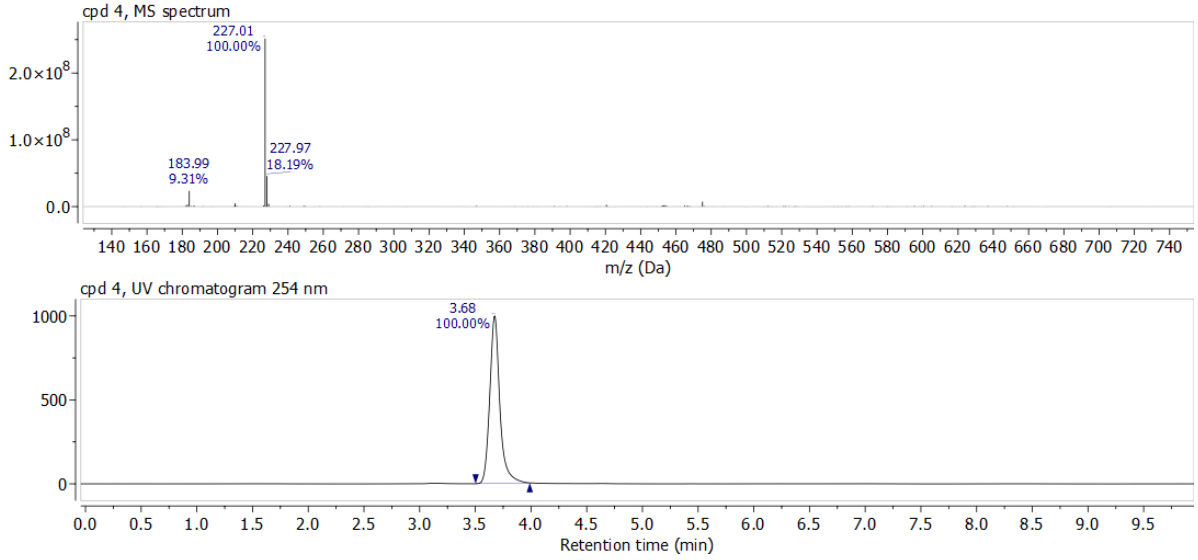
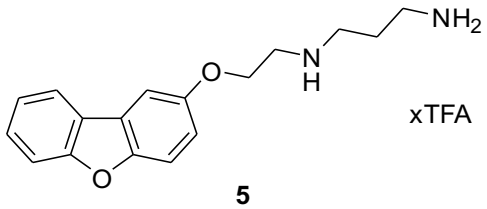
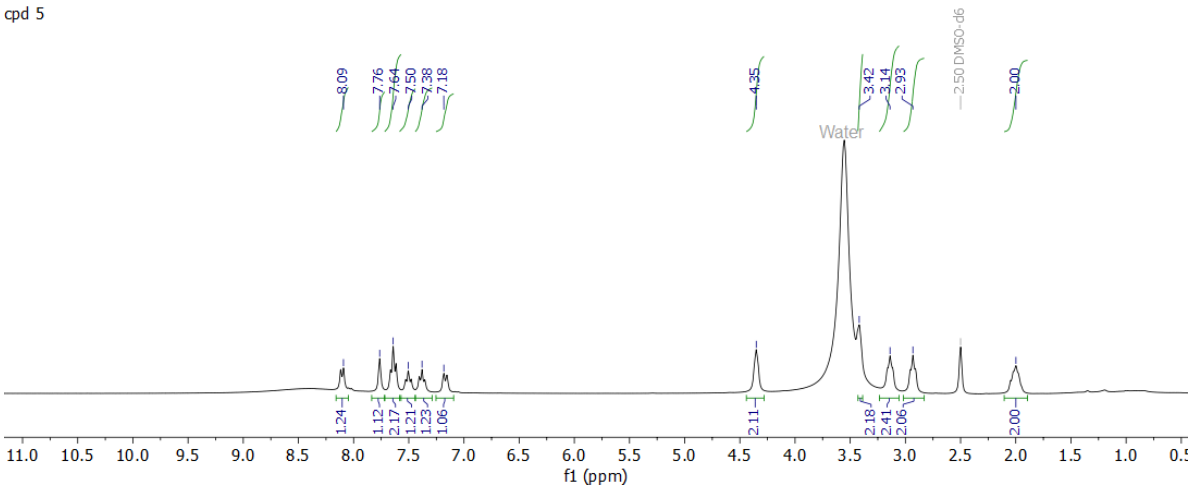
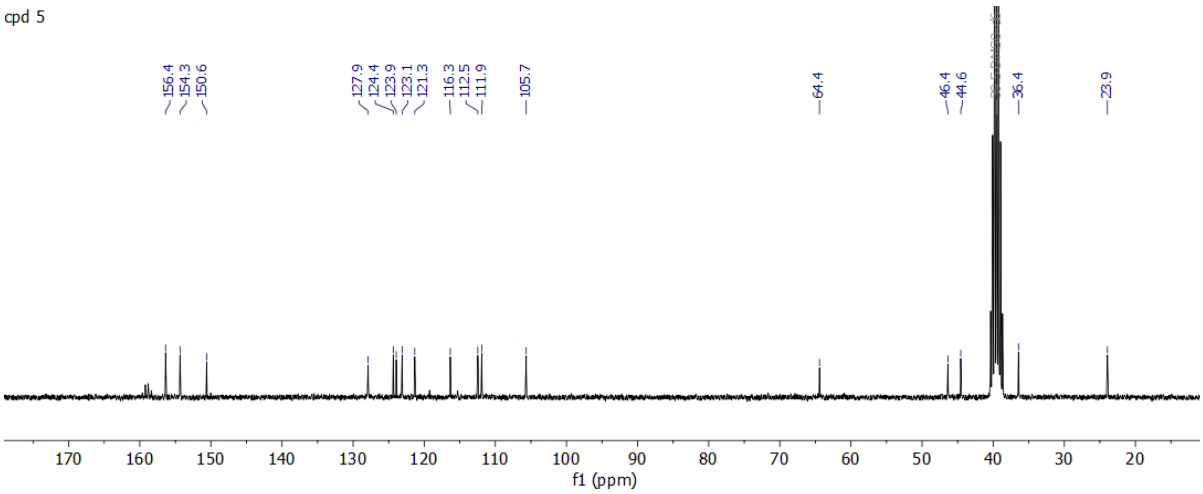


Figure S19. From top: ^{1}H-, ^{13}C-NMR spectra, ESI-MS spectrum and LC UV chromatogram at 254 nm of compound **4**. LC-ESI-MS measured with Zorbax column, isocratic 30% acetonitrile in water + 0.1% formic acid over 10 min.


NH2
N
H
O
O
xTFA

5


cpd 5
8.09
7.76
7.64
7.50
7.38
7.18
4.35
3.42
3.14
2.93
2.50 DMSO-d6
2.00
Water
1.24
1.12
2.17
1.21
1.23
1.06
2.11
2.18
2.41
2.06
2.00
f1 (ppm)


cpd 5
156.4
154.3
150.6
127.9
124.4
123.9
123.1
121.3
116.3
112.5
111.9
105.7
64.4
46.4
44.6
36.4
23.9
f1 (ppm)

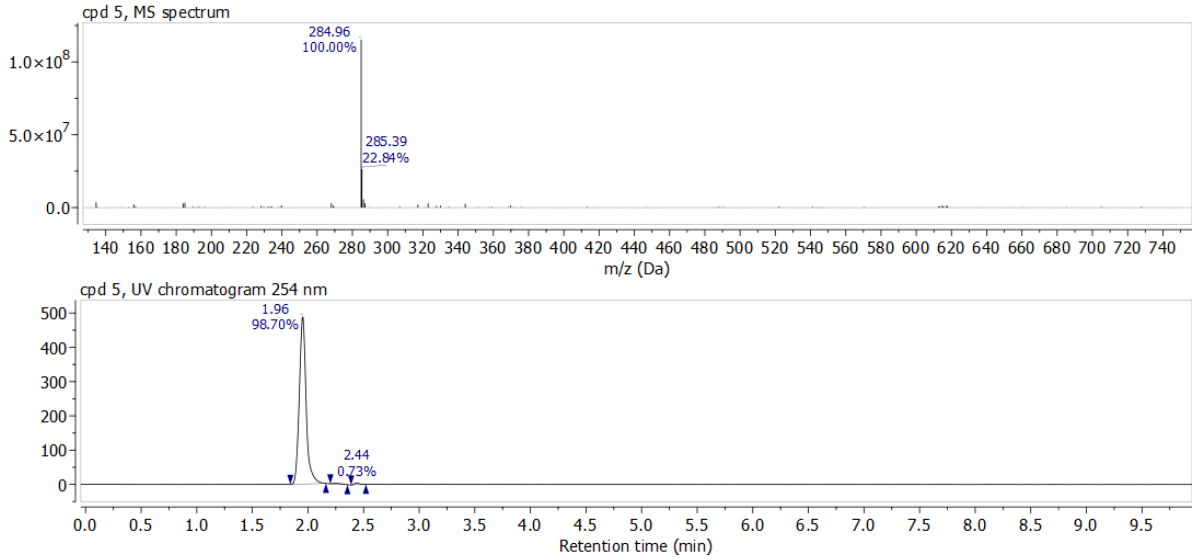
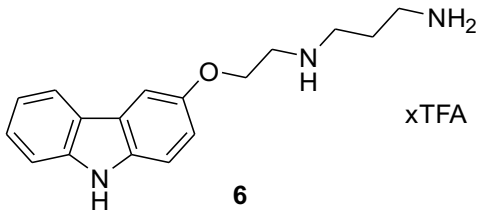
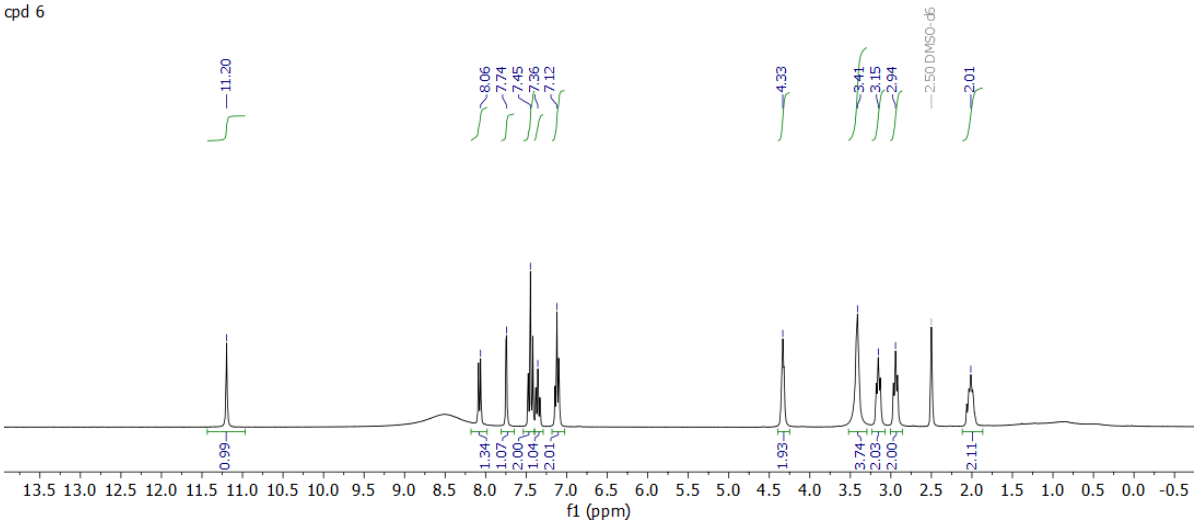
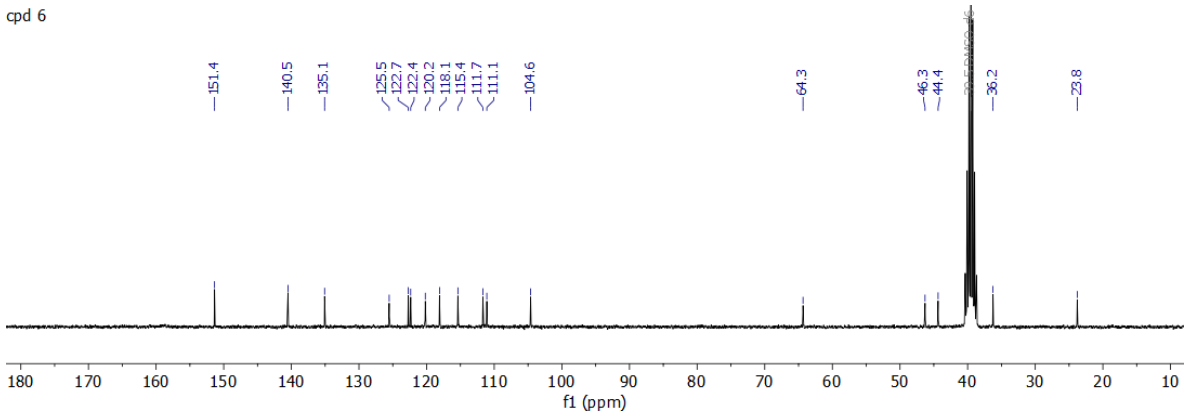


Figure S20. From top: ^{1}H-, ^{13}C-NMR spectra, ESI-MS spectrum and LC UV chromatogram at 254 nm of compound **5**. LC-ESI-MS measured with Zorbax column, isocratic 40% acetonitrile in water + 0.1% formic acid over 10 min.


NH2
N
H
O
N
H
xTFA

6


cpd 6
11.20
8.06
7.74
7.45
7.36
7.12
4.33
3.41
3.15
2.94
2.50 DMSO-d6
2.01
0.99
1.34
1.07
2.00
1.04
2.01
1.93
3.74
2.03
2.00
2.11
f1 (ppm)


cpd 6
151.4
140.5
135.1
125.5
122.7
122.4
120.2
118.1
115.4
111.7
111.1
104.6
64.3
46.3
44.4
39.5 DMSO-d6
36.2
23.8
f1 (ppm)

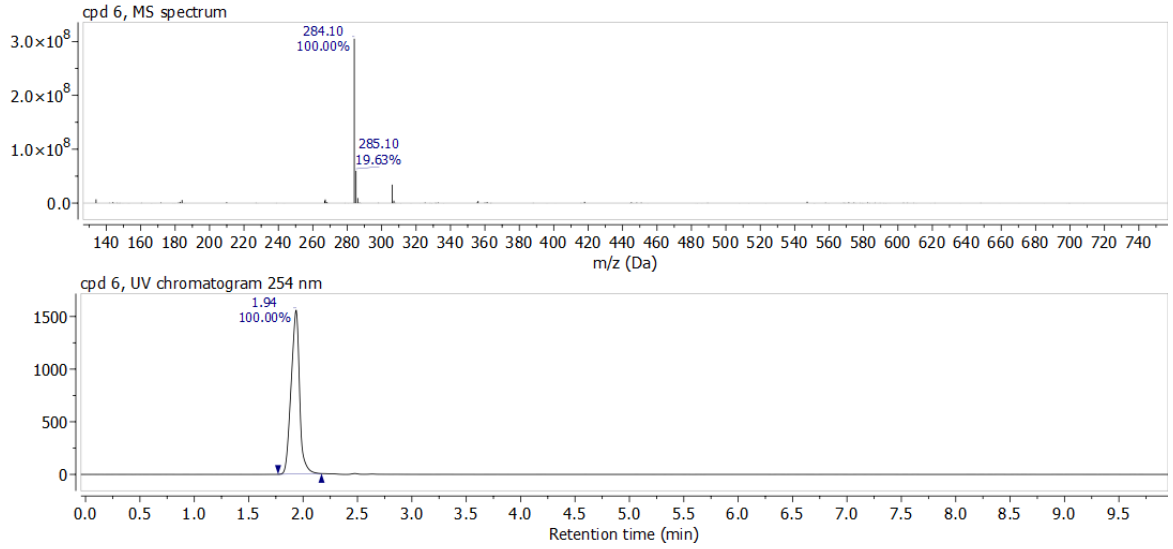
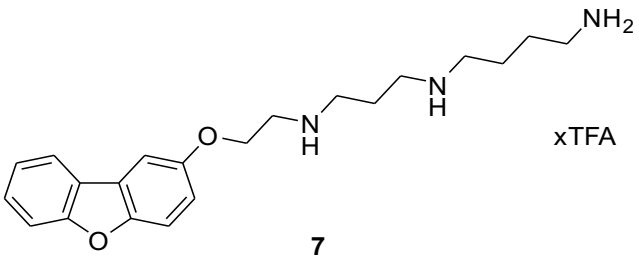
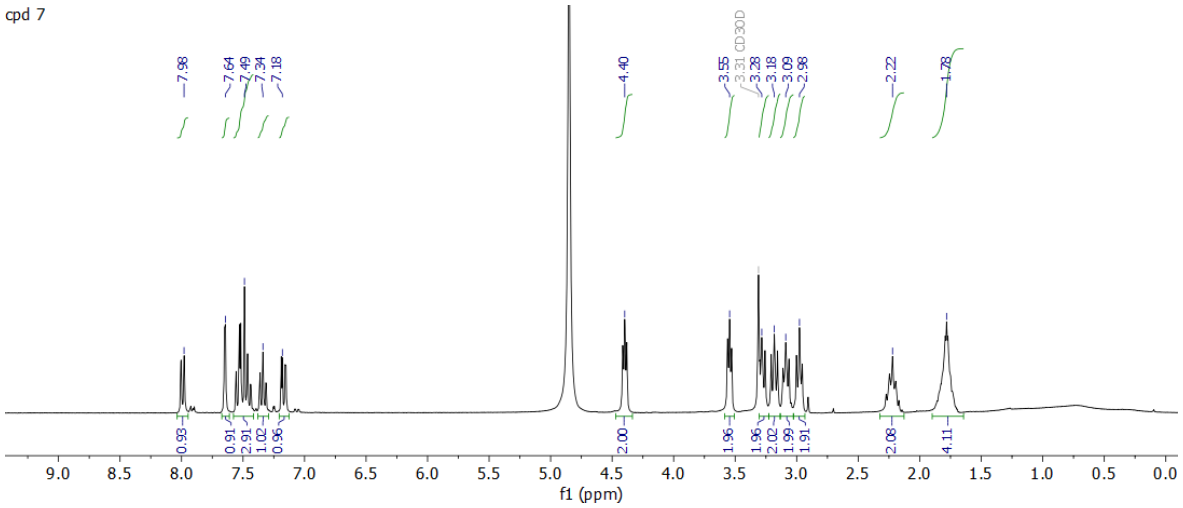
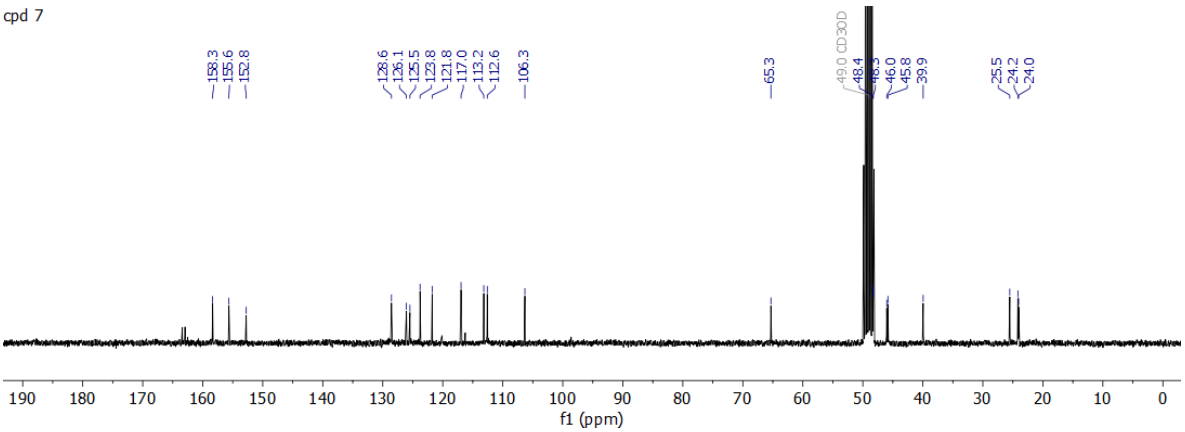


Figure S21. From top: ^{1}H-, ^{13}C-NMR spectra, ESI-MS spectrum and LC UV chromatogram at 254 nm of compound **6**. LC-ESI-MS measured with Zorbax column, isocratic 40% acetonitrile in water + 0.1% formic acid over 10 min.


NH2
N
H
N
H
O
O
xTFA

7


cpd 7
7.98
7.64
7.49
7.34
7.18
4.40
3.55
3.31 CD3OD
3.28
3.18
3.09
2.98
2.22
1.78
0.93
0.91
2.91
1.02
0.96
2.00
1.96
1.96
2.02
1.99
1.91
2.08
4.11
f1 (ppm)


cpd 7
158.3
155.6
152.8
128.6
126.1
125.5
123.8
121.8
117.0
113.2
112.6
106.3
65.3
49.0 CD3OD
48.4
46.0
45.8
39.9
25.5
24.2
24.0
f1 (ppm)

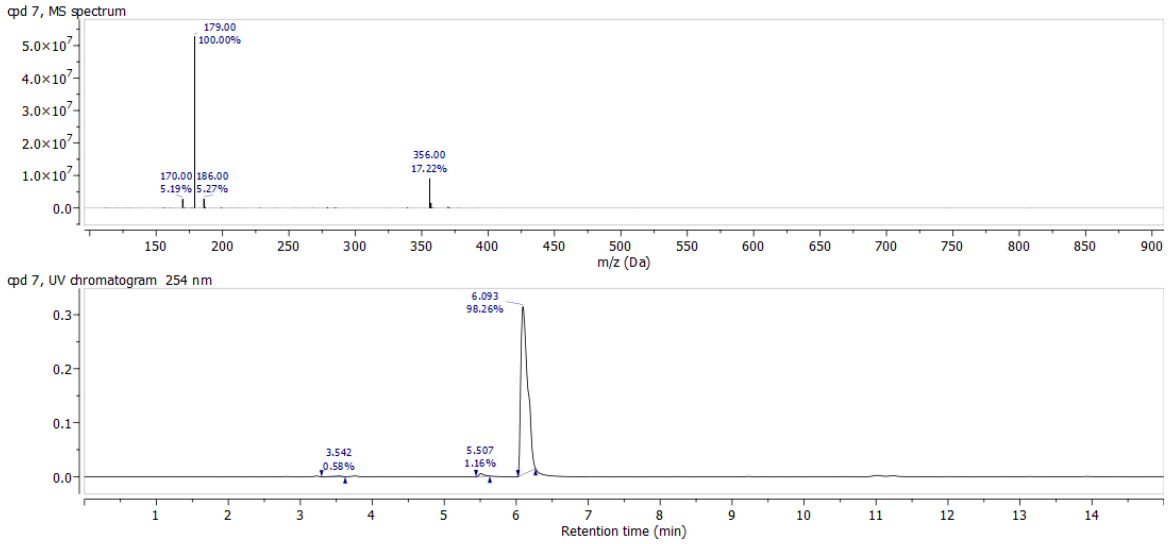
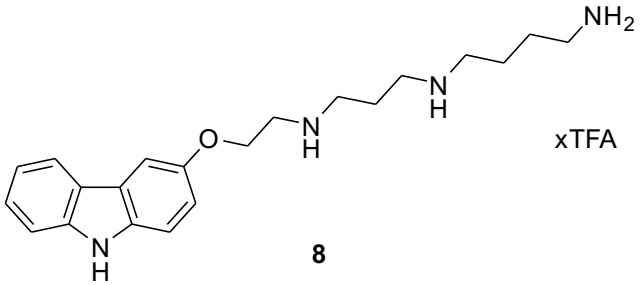
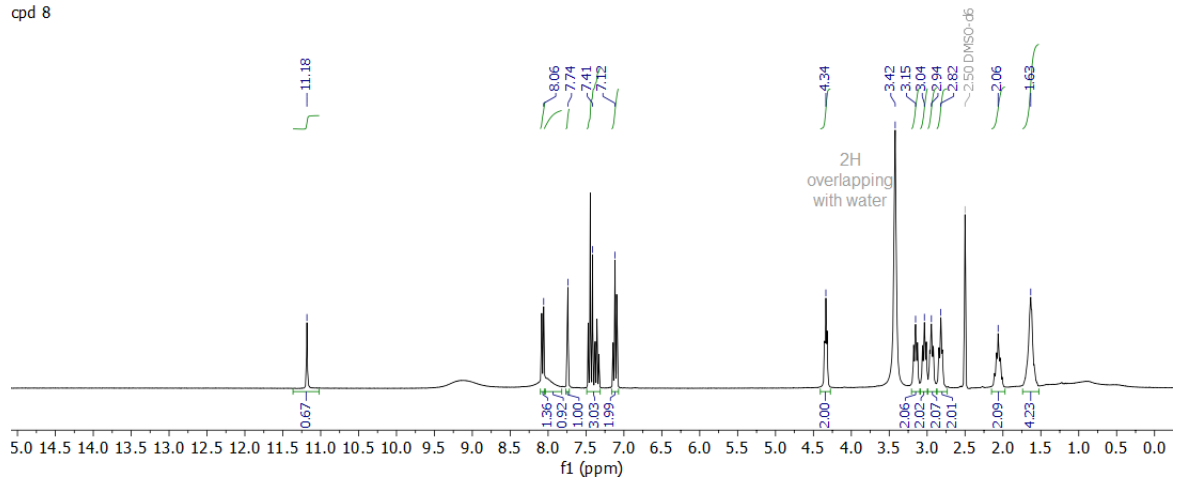
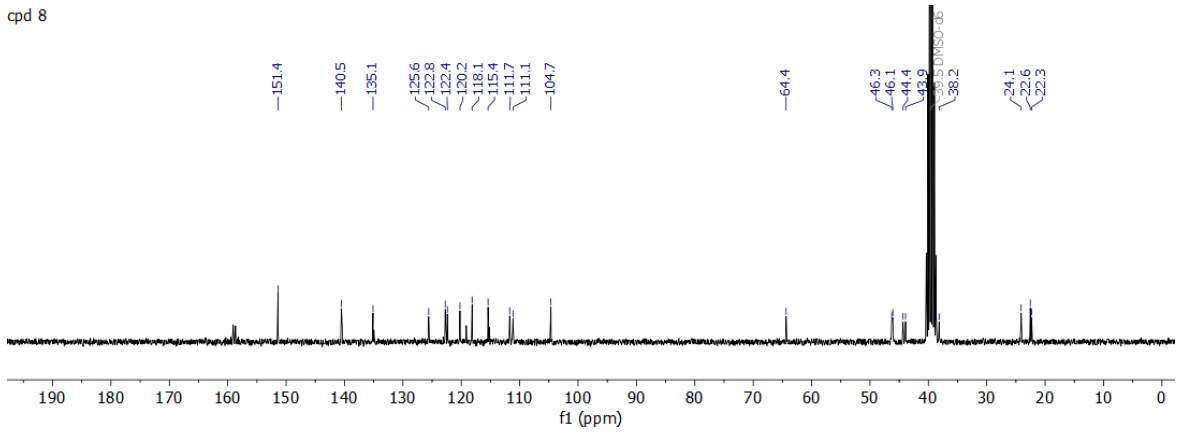


Figure S22. From top: ^{1}H-, ^{13}C-NMR spectra, ESI-MS spectrum and LC UV chromatogram at 254 nm of compound **7**. LC-ESI-MS measured with MZ-Aqua column, gradient 10–90% acetonitrile in water + 0.1% formic acid over 15 min.


NH2
N
H
O
N
H
xTFA
N
H

8


cpd 8
11.18
8.06
7.74
7.41
7.12
4.34
3.42
3.15
3.04
2.94
2.82
2.50 DMSO-d6
2.06
1.63
2H
overlapping
with water
0.67
1.36
0.92
1.00
3.03
1.99
2.00
2.06
2.02
2.07
2.01
2.09
4.23
f1 (ppm)


cpd 8
151.4
140.5
135.1
125.6
122.8
122.4
120.2
118.1
115.4
111.7
111.1
104.7
64.4
46.3
46.1
44.4
43.9
39.5 DMSO-d6
38.2
24.1
22.6
22.3
f1 (ppm)

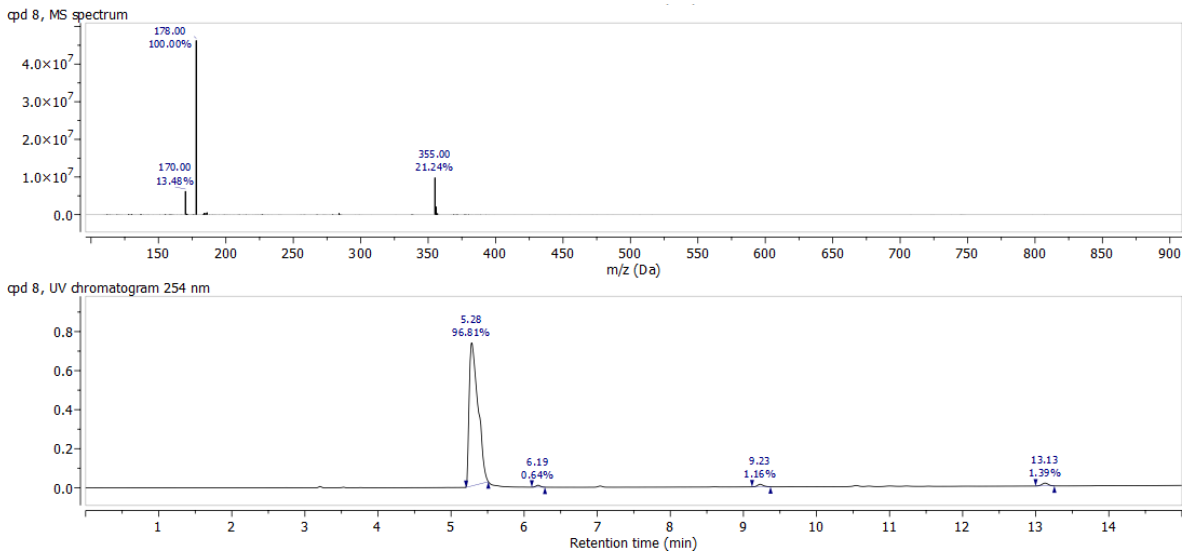


Figure S23. From top: ^{1}H-, ^{13}C-NMR spectra, ESI-MS spectrum and LC UV chromatogram at 254 nm of compound **8**. LC-ESI-MS measured with MZ-Aqua column, gradient 10–90% acetonitrile in water + 0.1% formic acid over 15 min.

5. Miscellaneous

Labelled RNA	Sequence
5'-biotin-*Tte* preQ$_1$ riboswitch	Biotin-GGAGAGGUUCUAGCUACACCUCUAUAAAAA ACUAA
5'-Cy5-*Tte* preQ$_1$ riboswitch	Cy5-GGAGAGGUUCUAGCUACACCUCUAUAAAAAACU AA
5'-biotin-*Ba* SAM-VI riboswitch	Biotin-GGCAUUGUGCCUCGCAUUGCACUCCGCGGG GCGAUAAGUCCUGAAAAGGGAUGUC
5'-Cy5-*Ba* SAM-VI riboswitch	Cy5-GGCAUUGUGCCUCGCAUUGCACUCCGCGGG GCGAUAAGUCCUGAAAAGGGAUGUC

Table S6. Sequences of the 5'-Cy5- and 5-biotin-labelled riboswitches.

DNA oligo	Sequence
Ba SAM-VI riboswitch template	GACATCCCTTTTCAGGACTTATCGCCCCGCGGAGTG CAATGCGAGGCACAATGCCTATAGTGAGTCGTATTA
Ba SAM-VI riboswitch forward primer	CGCGCGAAGCTTAATACGACTCACTATA
Ba SAM-VI riboswitch reverse primer	GACATCCCTTTTCAGGACTTATCGCCCCG

Table S7. Sequences of the DNA oligos used for the synthesis of the *Ba* SAM-VI riboswitch via PCR and IVT.

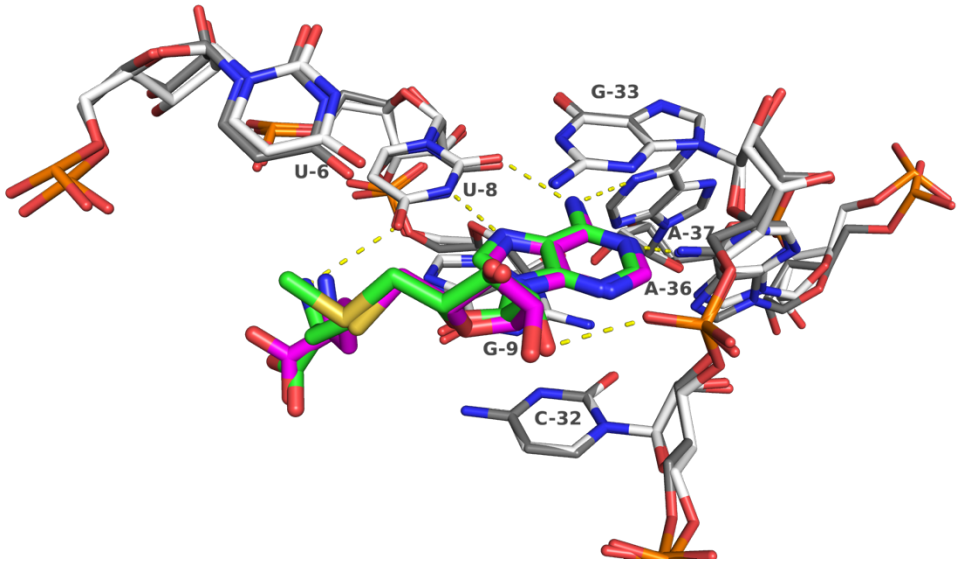


Figure S24. Superposition of the *Ba* SAM-VI riboswitch in complex with SAM (PDB-ID: 6LAS, green SAM carbon atoms, white RNA carbon atoms) and SAH (PDB-ID: 6LAU, magenta SAH carbon atoms, grey RNA carbon atoms).

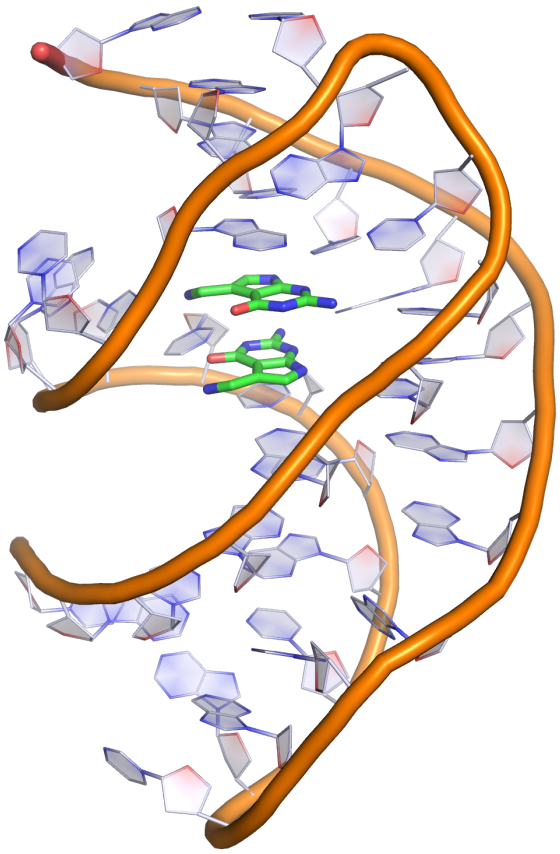


Figure S25. Crystal structure of *Carnobacterium antarcticum* $preQ_1$ type 1 riboswitch in complex with two $preQ_1$ molecules (green carbon atoms) (PDB-ID: 8FB3).

7. SAR Study on the SAM-I and SAM-VI Riboswitches – Preliminary Data

7.1. Introduction

The necessity for prospective novel antibiotics has steadily grown in recent years due to the emerge of multi-drug resistant pathogens.[15,117,123] The discovery of bacterial riboswitches two decades ago[147–150] led to the advent of a novel approach to targeting relevant pathogens such as those from the ESKAPE group.[9,170] Riboswitches have been the subject of extensive study with the aim of elucidating their regulatory mechanisms, ligand recognition and distribution among different species.[8,136,137,151,160,163,164,209] One of the most pervasive riboswitch families is the SAM-sensing family, which encompasses both the SAM-I riboswitch and the SAM-VI riboswitch. [160,163,209] These riboswitches are involved in a number of pathways within bacterial sulfur metabolism, including the biosynthesis of SAM which is an important methyl donor.[198,200,209] The structure of the *Tte* SAM-I riboswitch was elucidated in 2006[179], after which by various studies were conducted with the objective of in obtaining a ligand recognition pattern.[210,211,320] The sequence of this riboswitch exhibits significant similarity to portions of the *S. aureus* genome, indicating a probable structural and ligand recognition similarity.[214] Consequently, the well-studied *Tte* SAM-I riboswitch[179,210,211,320] represents an ideal model system for the investigation of small molecule binding to this structure.

The *Ba* SAM-VI riboswitch was identified in 2018.[161] These bacteria are not known to be human pathogens, but there exist some case reports of *Bfidobacteria* involved in pneumonia, prosthetic joint infections and sepsis in immunocompromised patients.[321–323] Interestingly, the sequence of the SAM-VI riboswitch family was found to share a high degree of similarity with that of SAM-III riboswitches.[161] The SAM-VI riboswitch thus represents a model system for investigation of small molecules that might bind in a similar manner to the SAM-III riboswitch which is found in multiple human pathogens like *E. faecalis*.[177,178]

The molecular recognition patterns of the SAM-I and SAM-VI riboswitches for SAM and SAH, respectively, exhibit distinct differences due to their structural differences.

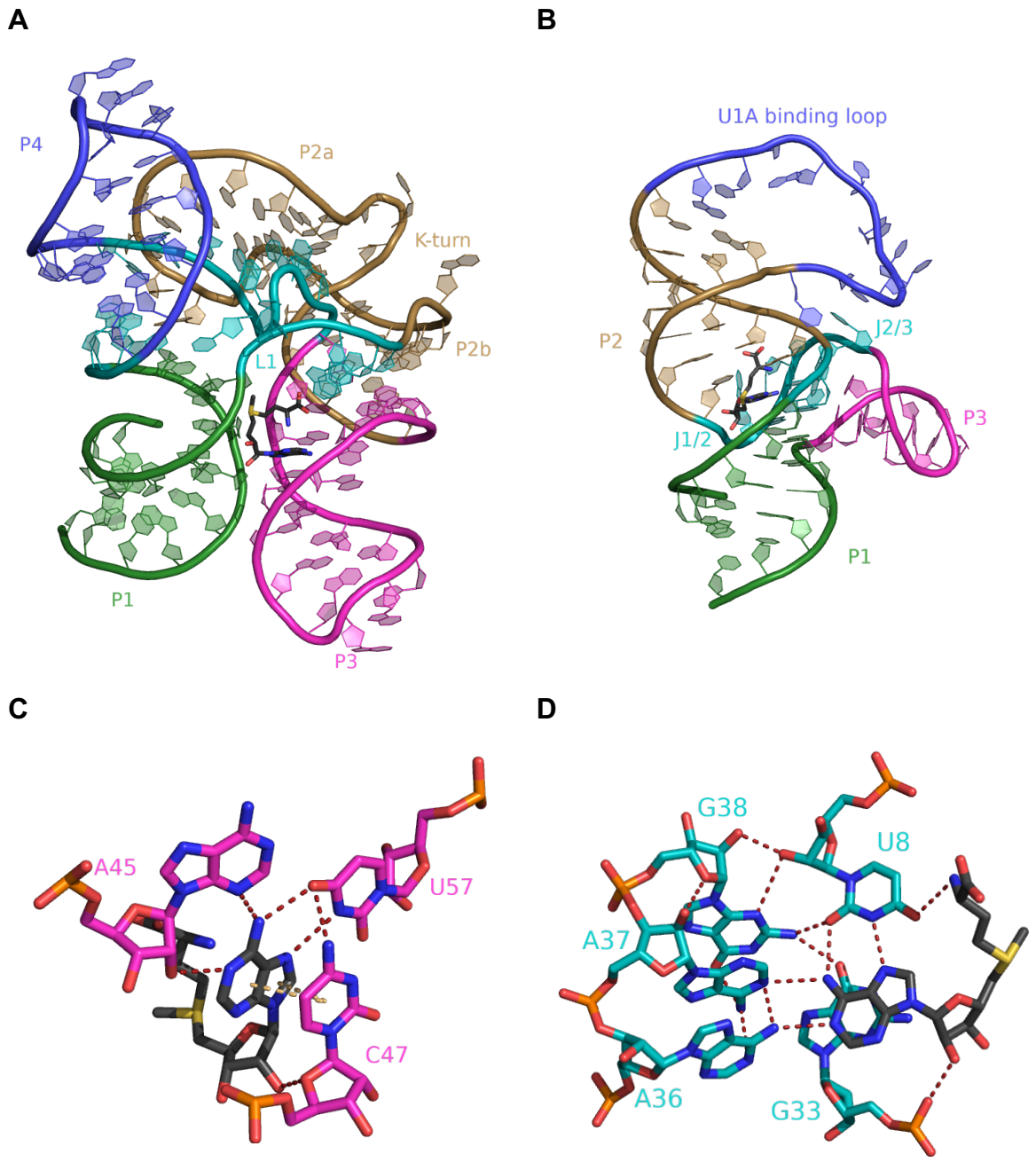


Figure 25: Structures and molecular recognition of the *Tte* SAM-I riboswitch (PDB: 2GIS) and the *Ba* SAM-VI riboswitch (PDB: 6LAS). **A**, Tertiary structure of the *Tte* SAM-I riboswitch (shown as cartoon) complexed with SAM (shown as sticks). **B**, Tertiary structure of the *Ba* SAM-VI riboswitch (shown as cartoon) complexed with SAM (shown as sticks). **C**, Stacking and hydrogen bonding interactions between SAM and P3 bases in the SAM-I riboswitch binding pocket. **D**, Recognition of the adenosine moiety of SAM by junctional bases in the SAM-VI riboswitch binding pocket. This figure was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

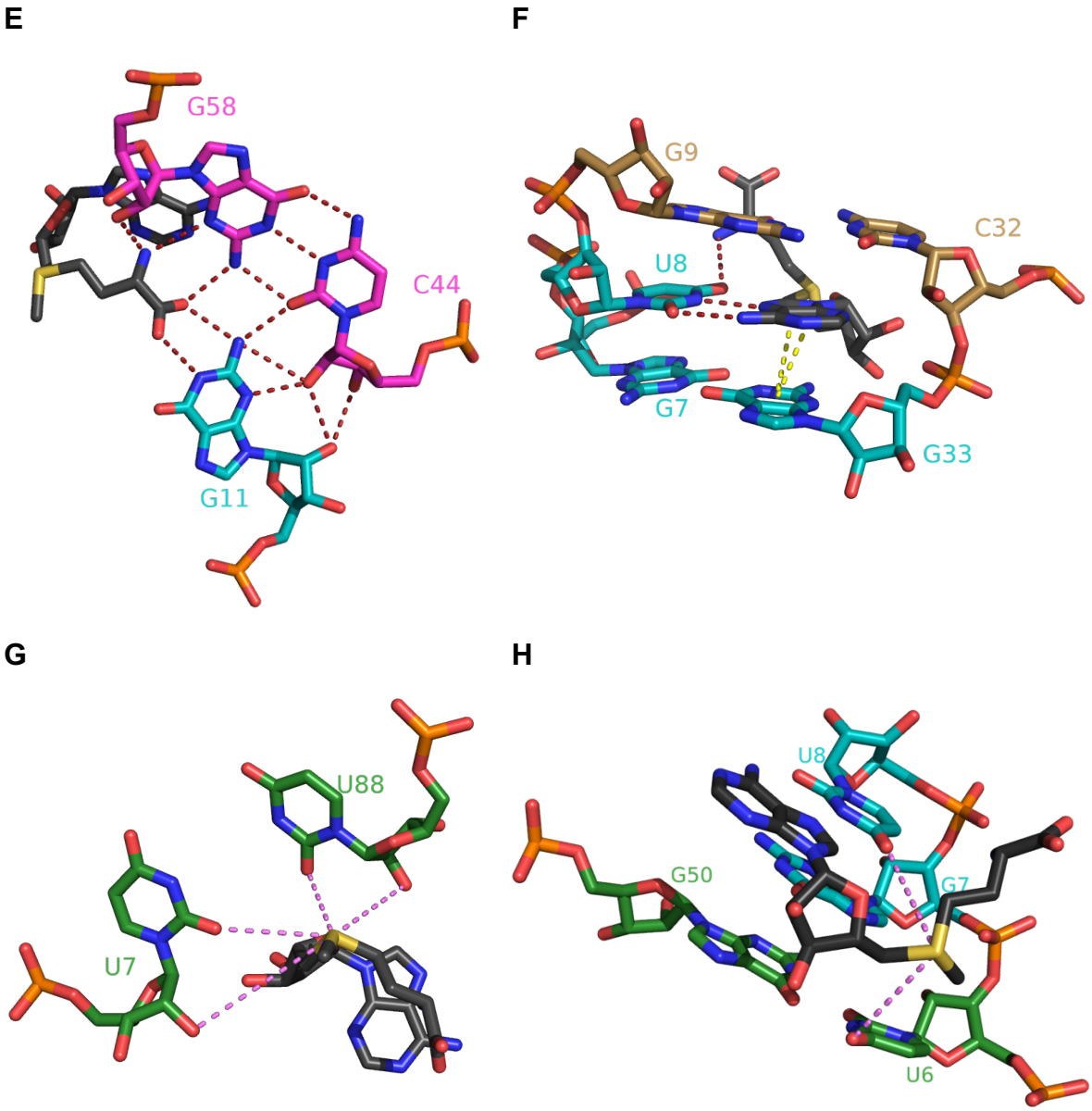


Figure 25: Structures and molecular recognition of the *Tte* SAM-I riboswitch (PDB: 2GIS) and the *Ba* SAM-VI riboswitch (PDB: 6LAS). **E**, Anchoring of the methionine tail of SAM by hydrogen bonding interactions with the base triple (G58–C44)·G11 in the SAM-I riboswitch binding pocket. **F**, Hydrogen bonding interactions between the adenosine moiety of SAM and U8 and stacking interactions with G33 in the SAM-VI riboswitch binding pocket. **G**, Electrostatic interaction of the sulfonium moiety with the carbonyl oxygens of U7 and U88 in the SAM-I riboswitch binding pocket. **H**, Electrostatic interaction of the sulfonium moiety with the carbonyl oxygens of U6 and U8 in the SAM-VI riboswitch binding pocket. Note: Hydrogen bonding interactions are colored red, stacking interactions are colored yellow, and electrostatic interactions are colored violet. This figure was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

Both the SAM-I and SAM-VI riboswitches belong to the junctional riboswitch class.[160,165] The *Tte* SAM-I riboswitch consists of four stems (P1-P4) which are connected by a three-way junction (Figure 25 A). The SAM binding pocket lies between the minor grooves of P1 and P3. The pocket is capped by the junctional region J1/2 of

the central loop L1. SAM adopts a compact conformation when binding. The ligand is recognized by several residues within this pocket: The adenosine moiety forms hydrogen-bonds with A45, C47, and U57 and π-π stacking with C47 from P3 (Figure 25 C). The methionine tail is anchored by hydrogen-bonding to the triple base pair (G58–C44)·G11 of P3 and J1/2 (Figure 25 E). The positively charged sulfonium moiety interacts with the carbonyl oxygens of U7 and U88 of P1 (Figure 25 G).[179,209,211] The *Ba* SAM-VI riboswitch consists of three stems (P1-P3) connected via a three-way junction (Figure 25 B). The SAM binding pocket is located in the intersection of P1, P2 and P3 stabilized by J1/2 and J2/3. In contrast to SAM-I riboswitches, the methionine tail of SAM binding to SAM-VI riboswitches is oriented outwards generating a extend ligand conformation. The adenosyl moiety is held in place by π-π stacking with G33 of J2/3 and hydrogen-bonding with U8, G33, A36, and A37 of J1/2 and J2/3 (Figure 25 D, F). The sulfonium moiety forms electrostatic interactions with the carbonyl oxygens of U6 and U8 of P1 and J1/2 (Figure 25 H).[161,162,209] In both riboswitches, these electrostatic interactions are found to be crucial for SAM affinity and selectivity. SAH adopts similar conformations to SAM; however, the loss of electrostatic interactions due to the absence of a positively charged sulfonium results in a significant reduction of affinity.[209]

The aim of this SAR study was to ascertain the influence of distinct structural motifs within SAM and SAH, respectively, on their binding affinity to both riboswitches. To this end, an MST (affinity determination) and an ITC (thermodynamic and kinetic properties of further developed compounds) assay for the purpose of testing a series of SAM analogue compounds against the SAM-I and SAM-VI riboswitches were established. This encompasses the adenine- or adenosine moiety, the amino acid moiety, and the positively charged sulfonium moiety, which is the primary factor responsible for SAM/SAH discrimination. Consequently, adenine, adenosine, methionine, ornithine, and diamino butyric acid (DABA, Table 6) were evaluated as partial structures of SAM and SAH. Sinefungin (SFG) was tested as an aza-analog of SAM to evaluate the impact of the positive charged sulfonium moiety. Furthermore, a series of SAM analogues (Table 6) with variations in the side chain[324] (synthesized by Dr. Marvin Schwickert and Sabrina Hoba of the Schirmeister group at the University of Mainz) were tested.

Table 6: Structures of the compounds tested in this SAR study

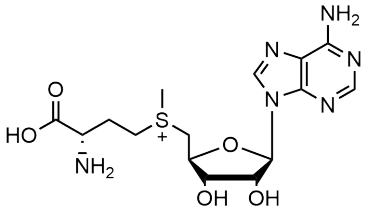
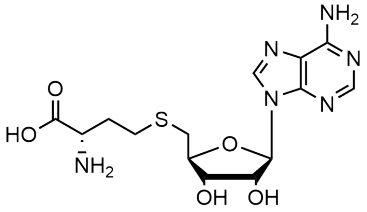
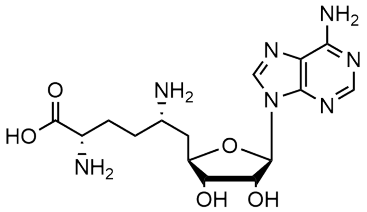
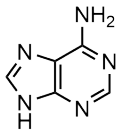
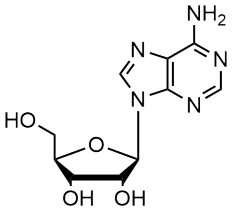
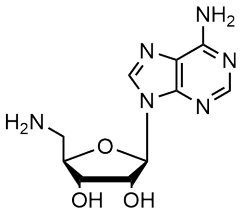
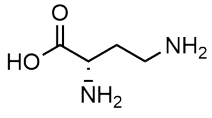
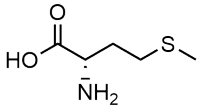
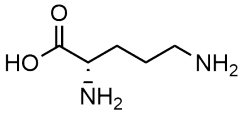
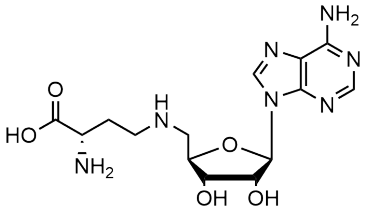
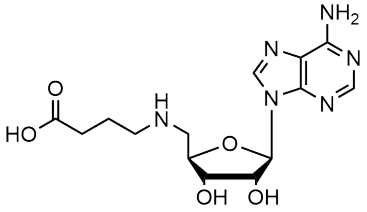
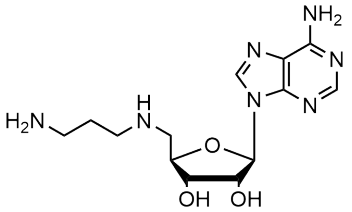
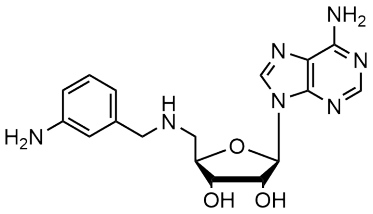
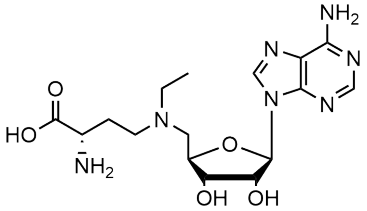
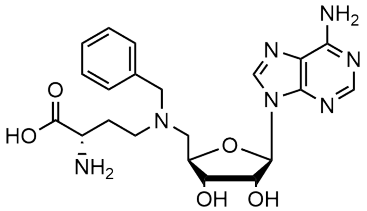
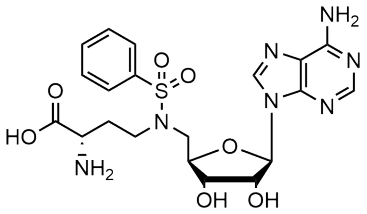
Compound	
1 (SAM)	
2 (SAH)	
3 (SFG)	
4 (adenine)	
5 (adenosine)	
6 (5'-ADA)	
7 (*L*-DABA)	
8 (*L*-methionine)	
9 (*L*-ornithine)	

Table 6: Structures of the compounds tested in this SAR study

	Compound
10	
11	
12	
13	
14	
15	
16	

This SAR study provides the starting point of a project aiming the investigation of small molecules distinct from the SAM/SAH motif that bind with high affinity and specificity to the SAM-I and SAM-VI riboswitches. The continuation of this project involves an ultra-large library virtual screening in accordance to previous approaches[325–328] of both small molecule and fragment libraries based on the results of this SAR study. Subsequently, the hit compounds and fragments, respectively, will be tested in the established assays to obtain a deeper understanding of the binding properties. Promising binders will be synthetically derivatized and optimized for further development of lead structures in early-stage antibiotic drug discovery.

7.2. Results and Discussion

Table 7: Binding affinities of compound **1**–**16** tested via MST against the *Tte* SAM-I and *Ba* SAM-VI riboswitches. n. b.: non-binder.

Compound	K_D **[µM]**	
	SAM-I Riboswitch	**SAM-VI Riboswitch**
1 (SAM)	0.103 ± 0.043	0.699 ± 0.185
2 (SAH)	84.7 ± 7.33	34.4 ± 12.0
3 (SFG)	21.0 ± 4.97	4.22 ± 0.585
4 (adenine)	n.b.	≥ 350
5 (adenosine)	n.b.	136 ± 19.2
6 (5'-ADA)	292 ± 90.9	20.8 ± 3.69
7 (*L*-DABA)	n.b.	n.b.
8 (*L*-methionine)	n.b.	n.b.
9 (*L*-ornithine)	n.b.	n.b.
10	1.92 ± 0.748	55.5 ± 13.3
11	5.18 ± 1.95	93.5 ± 11.3
12	n.b.	9.75 ± 4.51
13	n.b.	≥ 800
14	n.b.	10.4 ± 5.78
15	28.6 ± 2.31	n.b.
16	≥ 400	n.b.

Table 8: Thermodynamic profile of SAM, SAH, **10** and **12** binding to the SAM-VI riboswitch measured by ITC.

Compound	K_D [µM]	ΔG [kJ/mol]	ΔH [kJ/mol]	–TΔS [kJ/mol]
1 (SAM)	0.712 ± 0.022	–35.1 ± 0.06	–67.2 ± 3.39	32.1 ± 3.42
2 (SAH)	16.0 ± 5.09	–27.5 ± 0.757	–35.5 ± 1.63	8.00 ± 2.29
10	3.86 ± 0.406	–30.9 ± 0.250	–108 ± 13.0	77.2 ± 13.2
12	1.73 ± 0.511	–33.0 ± 0.839	–87.5 ± 1.76	54.5 ± 1.18

Table 9: Thermodynamic profile of SAM binding to the SAM-I riboswitch measured by ITC.

Compound	K_D [µM]	ΔG [kJ/mol]	ΔH [kJ/mol]	–TΔS [kJ/mol]
1 (SAM)	0.469 ± 0.006	–36.2 ± 0.058	–62.1 ± 1.53	25.9 ± 1.53

The affinities of SAM (K_D = 0.699 ± 0.185 µM) and SAH (K_D = 34.4 ± 12.0 µM) to the *Ba* SAM-VI riboswitch measured by MST were in good agreement with the literature.[162] The 50-fold decrease in affinity of SAH compared to SAM highlights the importance of the positive charge for tight binding due to electrostatic interactions with U6. This electrostatic interaction is a so-called charge-assisted chalcogen bond, a type of non-covalent interaction involving a chalcogen atom as electron acceptor that is influenced by the presence of a charge. The sulfur atom can interact with electron donors because it has regions of positive electrostatic potential (σ-holes) opposite to its covalent bond. The interaction strength is significantly enhanced when the sulfur carries a positive charge because the electrostatic attraction between the positive charged σ-hole and the electron donor (in this case the lone pair of electrons of O2 in U6) is stronger.[212] The SAM isostere SFG which adopts the same chirality showed a 6-fold decreased affinity (K_D = 4.22 ± 0.585 µM), while the aza-analogue **10** (K_D = 55.5 ± 13.3 µM) loses affinity by a factor of 80. Both small molecules possess positively charged amino groups under assay and physiological conditions (pH 7.4). A reason for the decreased affinity of SFG and **10** might be that their positive charge is non-permanent. The slightly higher affinity of SFG compared to **10** might be due to the closer proximity of the charged amino group to U6 and potential additional interactions with G7 or G50. Compound **14** (K_D = 10.4 ± 5.78 µM), the *N*-ethyl analogue of **10**, binds to the SAM-VI riboswitch with a 5-fold increased affinity compared to **10**. Adenosine (**5**) and 5'-ADA (**6**) showed binding affinities of 136 ± 19.2 µM and 20.8 ± 3.69 µM, respectively. The

adenine rings of both molecules can form hydrogen-bonds with U8, A36, and A38 as well as π-π stacking with G33 in the same manner as SAM. Their ribose 2'- and 3'- hydroxyl groups also interact with G33 via hydrogen-bonding. The positively charged 5'-amino group of **6** might interact with the oxygens of U6 and U8 and therefore lead to a 6-fold improved affinity compared to adenosine. Adenine (**4**) only interacts with U8, A36, A38 (hydrogen-bonds) and G33 (π-π stacking). The absence of a positively charged moiety as well as an extended "tail" leads to a significant decreased affinity with a K_D value of ≥ 350 µM. The amino acids DABA (**7**), methionine (**8**) and ornithine (**9**) showed no binding. This circumstance highlights the importance of a ring system with a positively charged moiety in distinct proximity to form crucial polar interactions (hydrogen-bonds and stacking) in the binding pocket. The influence of an amino acid tail was investigated with compounds **11** and **12**: The presence of a carboxyl moiety (**11**) instead of an amino acid leads to a decrease in affinity (K_D = 93.5 ± 11.3 µM), while a primary amino group alone leads to an improved affinity (K_D = 9.75 4.51 µM). This indicates that a second positive charge is favored over an uncharged or negatively charged moiety in this position, as observed in the electrostatic anchoring study (Chapter 6). In the case of **11**, the negatively charged carboxyl moiety interacts repulsively with the also negatively charged phosphate backbone and lacks hydrogen-bonding with U8. In contrast, the terminal amino group of **12** can interact with U8 and is attracted by the phosphate backbone. Previous studies showed that the association rate of **12** carrying two net positive charges increases compared to the uncharged SAH and therefore leading to an improved affinity. This hypothesis is supported by the behavior of **13**, which carries a terminal aniline moiety. This aniline structure is neutral and sterically more challenging which might be the reason for the dramatically decreased affinity to ≥ 800 µM. Sterically also challenging are the benzylamine- and phenyl sulfonamide-moieties of **15** and **16**, respectively. Both small molecules did not show binding to the SAM-VI riboswitch. Although the tertiary amine of **15** is positively charged, its interaction with U6 and U8 might be prevented by the phenyl ring. Both structures potentially lead to base clashes which only can be avoided by adopting an energetically unfavored conformation. Consequently, these two small molecules do not enter the binding pocket.

The ITC measurements showed K_D values of 0.712 ± 0.022 µM and 16.0 ± 5.09 µM for SAM and SAH, respectively. With a comparable Gibb's free energy $\Delta G < 0$, both

binding events are spontaneous. Compared to SAH, the binding of SAM is more driven by both enthalpy and entropy contributions, but its greater exothermic nature ($-\Delta H_{SAM} > -\Delta H_{SAH}$) makes it more thermodynamically favored despite the higher entropy cost ($-T\Delta S_{SAM} > -T\Delta S_{SAH}$). This indicates that the binding of SAM to the SAM-VI riboswitch involves stronger interactions such as the charge-assisted chalcogen-bonding which is not present in the SAH-SAM-VI complex. Compounds **10** (K_D = 3.86 ± 0.406 µM) and **12** (K_D = 1.73 ± 0.511 µM) also showed affinities in ITC experiments comparable to those determined by MST. The Gibb's free energies of both compounds showed similar values compared to SAM and SAH, indicating a comparable spontaneity of these binding events. **10** shows the most enthalpy-driven interaction, while also exhibiting the highest entropy-cost ($-T\Delta S$ = 77.2 ± 13.2 kJ/mol). This suggests a very strong but also highly ordering binding. **12** strikes a balance with strong enthalpy and entropy contributions similar to SAM, but less thermodynamically favored. SAM, **10** and **12** carry a positive charge in 5'-position and are more thermodynamically favored than the neutral SAH. This highlights once again the impact of the positive charge in 5'-position on an affine and strong binding to the SAM-VI riboswitch.

In summary, the complex network of hydrogen bonds and intercalation between bases of the SAM-VI riboswitch and the adenosyl-moiety seems to have a huge impact on ligand binding. Therefore, an exchange of this partial structure only might lead to similar binding affinities in the case of approximately identically hydrogen-donor/acceptor profiles. However, this exchange has not been studied and needs to be validated. A positive charge in a specific distance to the aforementioned structure is also crucial for binding as well as the polarity of this cationic charge. Different substituents in this position revealed that sterically challenging motifs such as aromatic substituents prevent the essential electrostatic interactions and must be avoided. The "tail" structure is also favored when carrying an amino acid structure or positively charged group that can form polar interactions to anchor the ligand in the binding pocket.

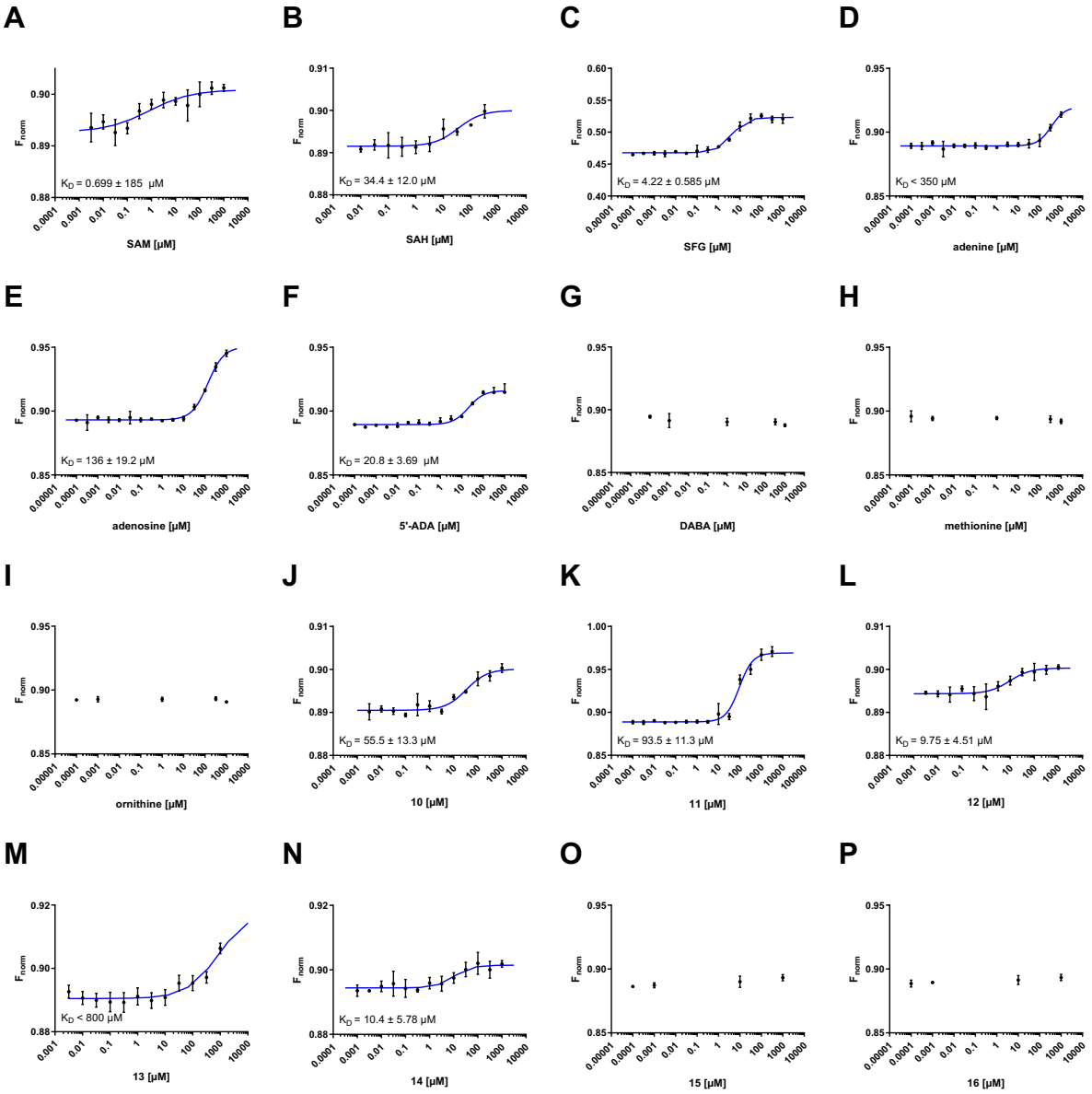


Figure 26: MST dose response results for SAM (**A**), SAH (**B**), SFG (**C**), adenine (**D**), adenosine (**E**), 5'-ADA (**F**), DABA (**G**), methionine (**H**), ornithine (**I**), **10** (**J**), **11** (**K**), **12** (**L**), **13** (**M**), **14** (**N**), **15** (**O**), and **16** (**P**) against the *Ba* SAM-VI riboswitch. DABA, methionine, ornithine, **15**, and **16** did not show a significant change in normalized fluorescence.

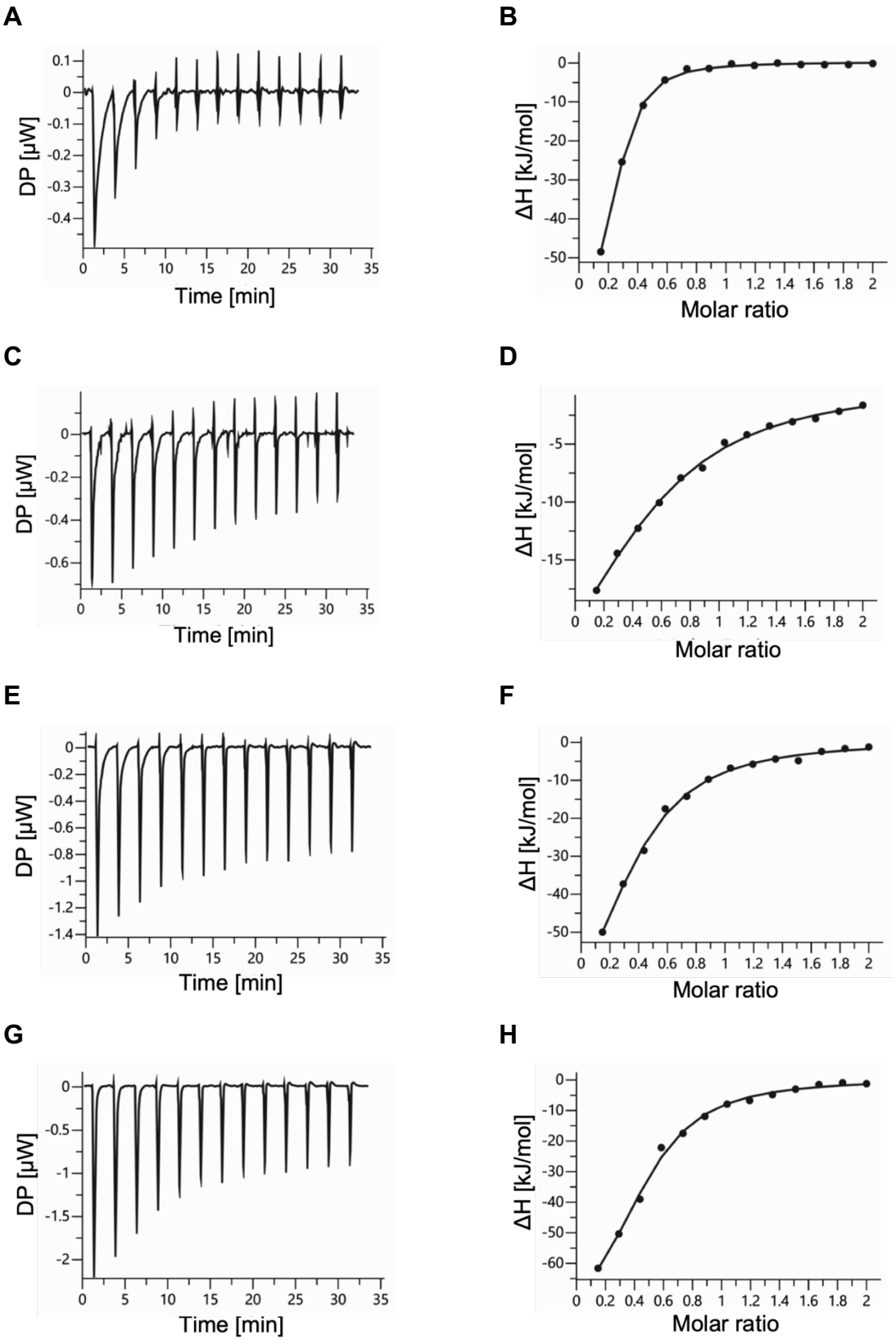


Figure 27: ITC evaluation of one representing experiment of SAM (**A**, **B**), SAH (**C**, **D**), **10** (**E**, **F**) and **12** (**G**, **H**) to the *Ba* SAM-VI riboswitch each. **A**, **C**, **E**, **G**: Injection raw data. **B**, **D**, **F**, **H**: Resulting isotherms by plotting the integrated heat against the molar ratio.

The *Tte* SAM-I riboswitch discriminates more strongly between the natural ligands SAM (MST: K_D = 0.103 ± 0.043 µM; ITC: K_D = 0.469 ± 0.006 µM) and SAH (K_D = 84.4 ± 7.33 µM) than the *Ba* SAM-VI riboswitch.[210,211] This decrease in affinity indicates once again the key role of the positive charge for ligand binding. SFG (K_D = 21.0 ± 4.97 µM) and **10** (K_D = 1.92 ± 0.748 µM) exhibit decreased affinities compared to SAM. The 200-fold decreased affinity of SFG is in good agreement with the previously reported relative affinity by in-line probing assay.[210] The position of the positively charged amino group might affect the binding affinity and therefore leading to a 10-fold improved affinity of **10** towards SFG. According to this hypothesis, **14** was expected to bind the *Tte* SAM-I riboswitch with similar potency. However, binding was not observed for this compound. A possible reason for this circumstance is that a substitution of the amino group leads to a loss of interactions with U7 and U88. **15** and **16** also carry substituents in this position, with **15** being positively charged and **16** neutral due to the sulfonamide structure. **15** (K_D = 28.6 ± 2.31 µM) binds with similar affinity as SFG to the RNA, while **16** did not show binding. This indicates that in addition to the positive charge, an aromatic substitution might be beneficial for binding. Aromatic ring systems such as in **15**, possibly form stacking interactions with bases of P1 and therefore anchor the ligand in the binding pocket. To confirm this hypothesis, the binding behavior of a small molecule set with aliphatic and aromatic substituted nitrogen atoms needs to be investigated. As the ligand conformation in the SAM-I pocket is compact, the nature of the terminal “tail” structure was evaluated with compounds **11** (negatively charged carboxyl moiety) and **12** (positively charged primary amine moiety) in addition to **10** (zwitterionic amino acid moiety). Compound **11** (K_D = 5.18 ± 1.95 µM) was found to bind with similar affinity as **10** (K_D = 1.92 ± 0.748 µM), while **12** did not show binding. Consequently, it can be assumed that the interaction of the terminal carboxylate with G11 and G58 are more important for ligand binding than the interaction of the amino group with G58; in fact, the absence of those terminal oxygens seems to be the reason for the loss of binding affinity. The non-binding of **13** (terminal aniline moiety) supports this hypothesis. Furthermore, the binding behavior of partial structural elements of the SAM/SAH motif was investigated. The positively charged amino acids DABA and ornithine did not show binding to the SAM-I riboswitch, indicating that an aromatic structure for stacking and hydrogen-bonding interactions is crucial for binding. The neutral methionine as part of

SAH also did not show binding which also suggests that the adenosyl moiety of SAH has a huge impact on binding affinity. Adenosine, adenine (both non-binder) and 5'-ADA (K_D = 292 ± 90.9 µM) permit the conclusion that the adenosyl moiety with a positive charge in 5'-position is crucial for binding as all binding small molecules (except SAH) feature this motif.

Additionally, Stoddard *et al*.[320] demonstrated that the ligand-free *Tte* SAM-I riboswitch binding pocket conformation differs from the ligand-bound state. In the apo-form, the internal loop is defined by the base triple (A45–U57)·A46 occupying the binding site of the adenosyl-moiety (Figure 28 A). When SAM is binding, A46 is flipped out of the A45–U57 plane and placed adjacent to C47–G56 (Figure 28 B). The fact that adenine and adenosine did not show binding suggests that the electrostatic interaction of a positively charged moiety is required for unlocking the binding pocket.

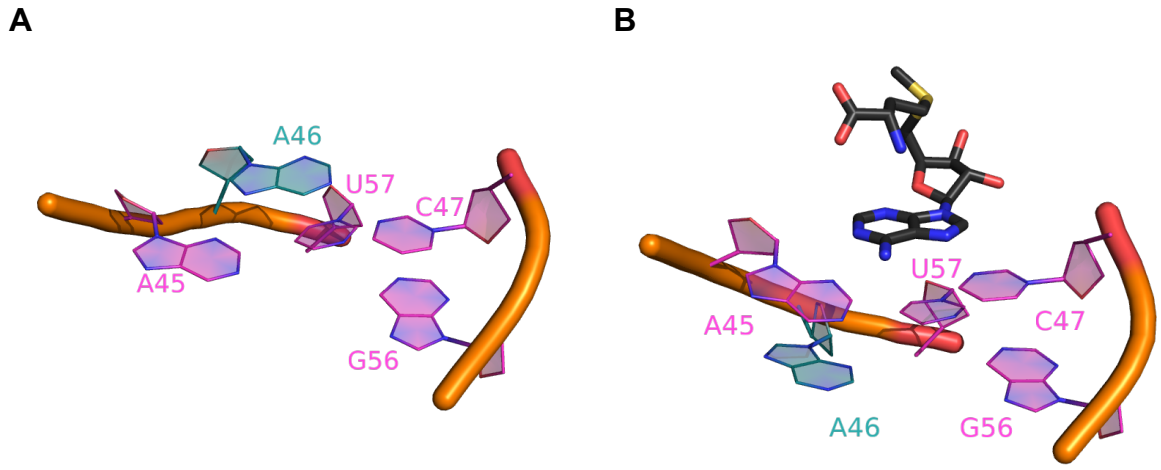


Figure 28: Comparison of the *Tte* SAM-I riboswitch binding pocket in the ligand-free (apo) and bound state (bases shown as cartoon). **A**, Apo state (PDB: 3IQP) with A46 (teal) occupying the adenosyl-binding site of SAM as part of a base triple with A45 and U57. **B**, SAM- (shown as sticks) bound state (PDB: 3IQR) with A46 (teal) flipped out of the A45–U57 plane. This figure was created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

In summary, the most important feature for ligand recognition seems to be the positive charge in 5'-position. The sulfonium moiety can form charge-assisted chalcogen bonds with U7 and U88. The additional electron donor compared to the SAM-VI riboswitch might explain the higher affinity of SAM to the SAM-I riboswitch. The distinct hydrogen-donor/acceptor pattern of the adenosyl moiety also seems to strongly influence the binding affinity – an exchange of this moiety might only be beneficial when imitating this pattern. A substitution of the positively charged atom might not be necessary; however only aromatic substituents seem to be accepted. A terminal carboxyl moiety

was also found to be crucial for affinity – the possibility of adjusting the length of this tail structure needs to be determined.

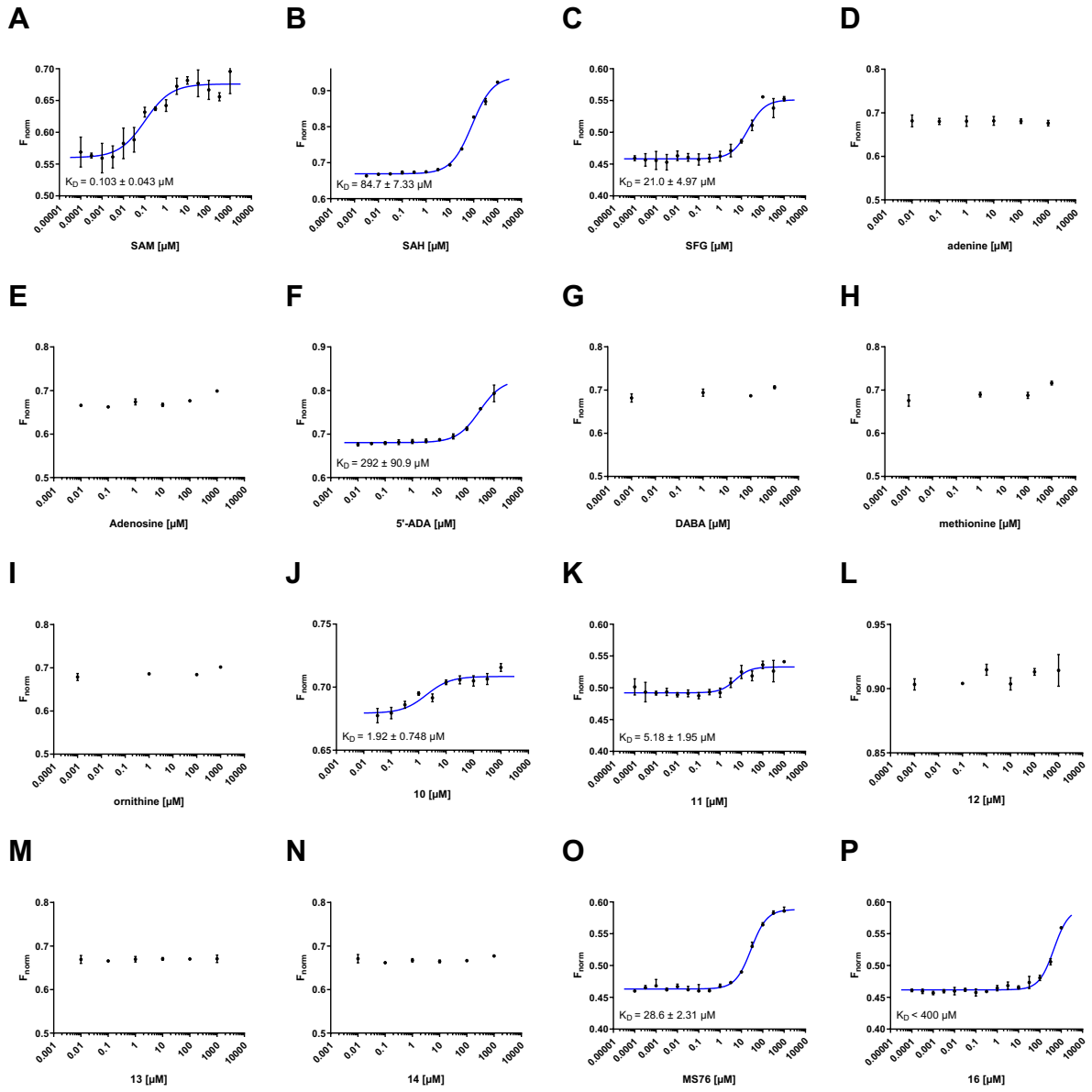


Figure 29: MST dose response results for SAM (**A**), SAH (**B**), SFG (**C**), adenine (**D**), adenosine (**E**), 5'-ADA (**F**), DABA (**G**), methionine (**H**), ornithine (**I**), **10** (**J**), **11** (**K**), **12** (**L**), **13** (**M**), **14** (**N**), **15** (**O**), and **16** (**P**) against the *Tte* SAM-I riboswitch. Adenine, adenosine, DABA, methionine, ornithine, **12**, **13**, and **14** did not show a significant change in normalized fluorescence.

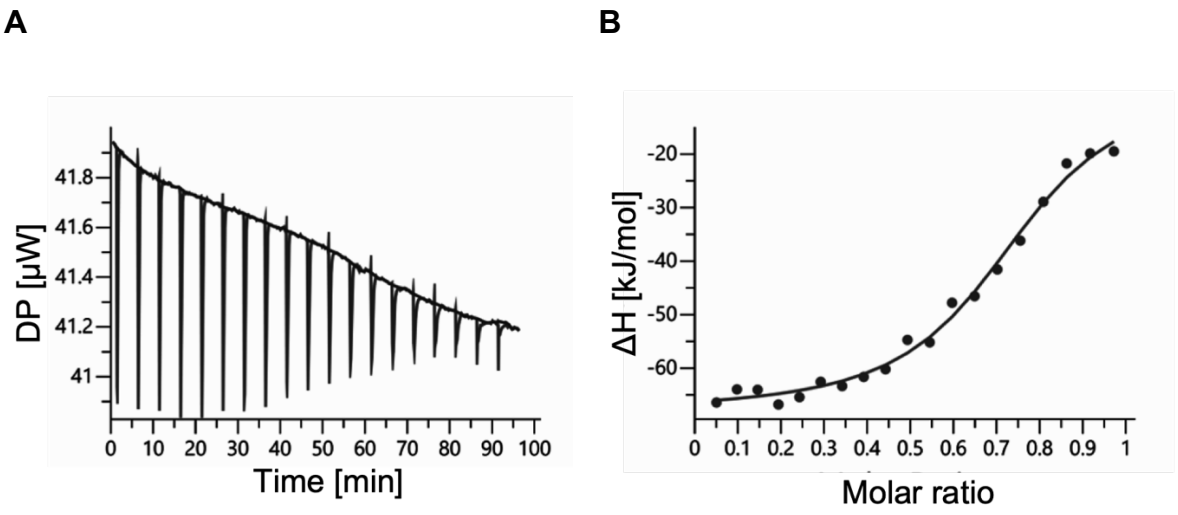


Figure 30: ITC evaluation of one representing experiment of SAM binding to the *Tte* SAM-I riboswitch. **A**, Injection raw data. **B**, Resulting isotherm of the integrated heat plotted against the molar ratio.

7.3. Experimental Section

RNA/DNA oligos

The DNA oligos for the *Tte* SAM-I riboswitch preparation were purchased from Integrated DNA Technologies (IDT, Coralville, USA). The *Ba* 5'-Cy5-SAM-VI riboswitch was purchased from Eurofins Genomics (Ebersberg, Germany).

Table 10: Sequences of the purchased DNA/RNA oligos

DNA Oligo ID	**Sequence**
SAM-I riboswitch template	5'-CGGCTCATCTTTCAACGTTTCCGCTGCAGGAATTGG CACCATTTCTGGTTGCCGGGGTTCTTCGCGCCAGTC GCTCCCCCTCTCTTGATAAGCCTATAGTGAGTCGTATTA-3'
SAM-I riboswitch forward primer	5'-CGCGCGAAGCTTAATACGACTCACTATA-3
SAM-I riboswitch reverse primer	5'-CGGCTCATCTTTCAACGTTTCCG-3'
RNA Oligo ID	**Sequence**
5'-Cy5-SAM-VI riboswitch	5'-Cy5-GGCAUUGUGCCUCGCAUUGCACUCCGCGGGGCG AUAAGUCCUGAAAAGGGAUGUC-3'

Table 11: Sequences of the *Tte* SAM-I and the *Ba* SAM-VI riboswitch

Riboswitch	Sequence
SAM-I riboswitch	5'-GGCUUAUCAAGAGAGGGGGAGCGACUGGCGCGA AGAACCCCGGCAACCAGAAAUGGUGCCAAUUCCUGC AGCGGAAACGUUGAAAGAUGAGCCG-3'
SAM-VI riboswitch	5'-GGCAUUGUGCCUCGCAUUGCACUCCGCGGGGCG AUAAGUCCUGAAAAGGGAUGUC-3'

RNA preparation

PCR. The DNA template for the IVT were amplified using PCR. The PCR reaction mixture contained 10 nM of template DNA, 2 µM of each forward and reverse primer, 400 µM dNTP Mix (New England BioLabs (NEB), Ipswich, USA), 3 mM $MgCl_2$, 0.05 U mL^{-1} Taq DNA Polymerase (NEB, Ipswich, USA) and 1× standard Taq reaction buffer (NEB, Ipswich, USA) in a final volume of 200 µL. The reaction was performed using a ^{3}Prime thermal cycler (Cole-Parmer, Stone, UK) with the following cycling conditions: initial denaturation (95 °C, 30 s), followed by 30 cycles of denaturation (95 °C, 30 s), annealing (temperature depends on the primer sequence and was determined by NEB TM calculator, 30 s) and extension (72 °C, 45 s), and a final extension (72 °C, 4 min). The PCR product homogeneity was checked on a 10% native PAGE gel to confirm the size and purity.

IVT. The crude PCR products were used as templates for IVT using the T7 polymerase system. The IVT reaction mixture contained 400 µL PCR product, (without further purification), 7.5 mM of each NTP (Jena Bioscience, Jena, Germany), 10 mM DTT, 30 mM $MgCl_2$, 5 µL HighYield T7 RNA polymerase mix (Jena Bioscience, Jena, Germany) and 1× HighYield T7 reaction buffer (Jena Bioscience, Jena, Germany) in a final volume of 1 mL. The reaction was incubated at 37 °C for 4 hours.

RNA extraction and purification. The formed pyrophosphate was removed by centrifugation at 14,000 g for 5 min directly after transcription and the cleared supernatant was transferred to a micro reaction tube for further sample processing. To remove the PCR- and IVT-enzymes, phenol/chloroform extraction was performed twice using ROTI® phenol/chloroform/isoamyl alcohol (Carl Roth, Karlsruhe,

Germany) and the standard protocol from the supplier. For further purification and concentration of the RNA the RNA Clean & Concentrator-5 kit (Zymo Research, Irvine, USA) including DNase I was used. The RNA product was checked on a 10% denaturing PAGE gel to confirm size and purity.

Microscale thermophoresis. MST experiments were performed on a Monolith NT.115 (NanoTemper Technologies, Munich, Germany) using standard uncoated capillaries. All steps involving labelled RNA were conducted under light protection. For refolding, the unlabelled and 5'-Cy5-labelled RNA was heated to 75 °C for 5 min and cooled down to room temperature for 20 min. The 5'-Cy5-SAM-VI riboswitch was diluted to 20 nM in MST buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl, 25 mM $MgCl_2$) and supplemented with ligands in appropriate concentrations typically ranging from 1 mM to 0.1 nM in 3.16-fold (semi-logarithmic) dilution series. The SAM-I riboswitch was diluted to 100 nM in MST buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl, 25 mM $MgCl_2$), labelled with 1 µM Syto59 or 1 × SybrGold, and supplemented with ligands in appropriate concentrations typically ranging from 1 mM to 0.1 nM in 3.16-fold (semi-logarithmic) dilution series. SAM and SAH were used as positive controls. For each experiment, one DMSO negative control was carried out. Final DMSO concentration was always 2%. All MST measurements were analyzed using the NT analysis software (version 1.5.41) and exported for statistical analysis and plotting in GraphPad Prism (version 8.0.1, GraphPad Software, Boston, MA, USA, https://www.graphpad.com).

Isothermal titration calorimetry. ITC experiments were performed on an automated MicroCal PEAQ-ITC calorimeter (Malvern Panalytical, Malvern, UK) at 20 °C. Prior to titration, the RNA was desalted using Zebra spin desalting columns (ThermoFisher, Waltham, USA) following the standard protocol from the supplier and lyophilized. The solid RNA was resolved in ITC buffer (50 mM HEPES pH 7.5, 40 mM KCl, 2 mM $MgCl_2$ for *Tte* SAM-I riboswitch and 40 mM HEPES pH 7.0, 50 mM KCl, 10 mM $MgCl_2$ for *Ba* SAM-VI riboswitch). For refolding, the RNA was heated to 75 °C for 5 min and cooled down to room temperature for 20 min.
For the *Tte* SAM-I riboswitch experiments, the RNA was diluted to 20 µM in ITC buffer, and SAM was diluted to 100 µM in ITC buffer. For each experiment, 19 ligand injections

in a volume of 3.0 µL, with a 0.5 µL s^{-1} rate, 150 s intervals between injections and a reference power of 41.9 µW were performed.

For the *Ba* SAM-VI riboswitch experiments, the RNA was diluted to 15 µM (SAM, compounds **10** and **12** measurements) and 30 µM (SAH measurements) in ITC buffer. SAM, **10** and **12** were diluted to 150 µM in ITC buffer, SAH was diluted to 300 µM in ITC buffer. For each experiment, 13 or 19 ligand injections in a volume of 3.0 µL, with a 0.5 µL s^{-1} rate, 150 s intervals between injections and a reference power of 41.9 µW were performed.

All experiments contained 0.2% DMSO. Integrated heat data were analyzed using a one-site binding model via the MicroCal PEAQ-ITC Analysis Software (version 1.21), provided by the manufacturer.

8. Conclusion and Outlook

The objective of this doctoral thesis was to characterize diverse interactions between small molecules and RNA targets, focusing on bacterial riboswitches – specifically $preQ_1$, SAM-I, and SAM-VI – as well as the viral HCV IRES. This research contributes significantly to the field of structure-based drug design targeting RNA in pathogenic organisms, aligning with contemporary advancements in RNA-small molecule interactions studies.

RNA has emerged as a critical target for novel therapeutic development, particularly in the context of bacterial and viral infections. However, identifying RNA targets that are both disease-specific and druggable presents significant challenges.[1,2,5,100] Like proteins, which have traditionally been the primary focus of drug design, RNA displays distinctive structural features and dynamic conformations that are essential for the survival of pathogens.[1,5,51] However, these same attributes can also render RNA challenging to target selectively. This selectivity challenge is particularly pronounced as many RNA structures such as hairpins are widespread among different species. This lack of diversity is due to the only four building blocks of RNA and the anionic backbone that influence RNA's secondary structure.[1,2] Consequently, the uniqueness of RNA structures contributes significantly to the selectivity challenge. While many RNA motifs are conserved among different species due to their essential biological functions, unique structures with specialized functionalities exist in specific organisms. Among those structures are bacterial riboswitches[7,9,14,136,137,152,163,166] and viral RNA elements[16] such as the already targeted HCV and enterovirus 71 IRES[17–19,21,40,230,329] and HIV TAR[22,330–334]. Especially for small molecules that bind to bacterial riboswitches, it has been considered that some human proteins use the same metabolites as the bacterial RNA, for instance, SAM and TPP. Commonly, the binding mechanism are distinct, so the rational design of bacterial riboswitch directed small molecules with chemical dissimilarity to the natural metabolite is possible.[14,155,180] Achieving selective targeting is crucial to avoid off-target effects that could disrupt host cellular processes and lead to toxicity.[5,124]

Moreover, druggability is another key hurdle. Many RNA targets lack well-defined pockets or grooves that can accommodate small molecules with high affinity and

specificity, making it difficult to design effective drugs. The structural plasticity and dynamic nature of RNA further complicate this issue, as RNA may adopt multiple conformations that influence ligand binding.[1,2,5]

The development and application of robust biophysical methods is crucial for understanding RNA-small molecule interactions, particularly in evaluating the binding affinity, kinetics, and specificity of potential therapeutics.[2,6,100] Assays like FRET, ITC, MST, and SPR are widely applied in drug discovery, particularly in screening campaigns, SAR studies, and hit-to-lead optimization.[2,6]

MST enables the detection of molecular interactions by measuring the changes in the movement of molecules under a temperature gradient. The movement of the molecules are detected by fluorescence. Consequently, the target RNA must be labelled with a fluorescent probe. MST provides medium-low throughput and might struggle with very weak or strong interactions, but benefits from minimal sample consumption.[62,266] ITC measures the heat change during molecular interactions and provides detailed thermodynamic profiles[236,293] and, in some cases, kinetic profiles[272,291]. This method is often used to refine SAR by elucidating binding enthalpy and entropy contributions, and to guide lead optimization campaigns.[335–338] Within fragment optimization processes, LE, LLE, and FQ besides the affinity play a crucial role.[103,104] These metrics can be determined by ITC experiments (equations (1)–(3)).[274] However, ITC has lower throughput and requires higher sample concentrations, making it less suitable for large-scale screenings.[6,337] SPR detects real-time binding kinetics via changes in refractive index at a sensor surface, offering high sensitivity and the ability to measure kinetic parameters, but may be limited by non-specific binding and the need for immobilization of one interaction partner.[70,339,340] The throughput depends strongly on the complexity of the experiments but is typically medium, so this method, like ITC and MST, is not a primary screening method.[341] FRET assays rely on energy transfer between fluorophores and are invaluable in studying molecular interactions in a cellular context[271], with the potential for higher throughput in real-time monitoring.[50] However, FRET can be limited by the need for careful fluorophore labelling, potential photobleaching, and the requirement for proximity of interacting molecules.[6,271]

Collectively, these four techniques offer complementary insights into binding kinetics, thermodynamics, and molecular interactions, making them indispensable in the

iterative process of finding and optimizing drug candidates, though each has its throughput limitations and specific challenges. These methods were applied in the projects of this thesis for studying the interactions of small molecules with bacterial and viral RNA targets.

The bacterial riboswitches studied in this thesis – preQ$_1$, SAM-I, and SAM-VI (as model system for SAM-III, Chapter 1.5.5) – are key regulatory elements that control gene expression in response to small molecule metabolites.[8,136,137,151,183,209] These riboswitches are highly conserved among bacterial species, making them attractive targets for antibiotic development in the age of multi-drug resistance.[7,9,14] By targeting these RNA elements, it is possible to inhibit the expression of essential bacterial genes, offering a novel approach to combat antibiotic resistance.[9,14] The selectivity challenge is mitigated here, as riboswitches have distinct structures that are not typically found in human RNA, reducing the risk of off-target toxicity.[8,9,152]

Similar to bacterial riboswitches, the HCV IRES is a crucial viral RNA element that facilitates the translation of viral proteins necessary for replication. Inhibiting the HCV IRES could effectively disrupt the viral life cycle, offering a potential strategy for antiviral therapy for viruses carrying an IRES element. Given the high mutation rates of viruses like HCV, RNA-targeted therapies that focus on conserved structural motifs of viral RNA may prove advantageous by reducing the likelihood of resistance development.[16,40,232] However, ensuring that these targets are druggable requires innovative strategies in small molecule design, as explored in this thesis.

The development of a novel, effective MST method for studying RNA-ligand interactions was a significant advancement of this work. This method, employed in our SAR study on the SAM-I riboswitch, provided detailed insights into ligand recognition, which are critical for optimizing drug candidates. The accessibility and versatility of this method could facilitate future research efforts to identify and optimize small molecule binders for the SAM-I riboswitch, which is present in priority pathogens such as *S. aureus*.[166] These efforts should focus on structures distinct from the SAM/SAH motif to avoid off-target interactions with human SAM-dependent methyltransferases. This structural motif can be divided into the adenine core moiety, the ribose, and the amino acid (methionine) tail. Within our SAR study, the tail structure carrying a terminal carboxylate and the charged sulfur/nitrogen atom was found to be crucial for binding to the SAM-I riboswitch. However, the length of this “tail” might be adjustable.

Consequently, the adenine/adenosine moiety needs to be further elucidated as the “tail” is now well explored. Fragment docking studies conducted by Laura Almena Rodriguez, suggest that the adenine moiety could be replaced by 2-aminopyrimidine or picolinamide fragments (Figure 31 A, C) that mimic adenine’s interaction pattern with A45 and U57 in the binding site (Figure 31 A–F) as they are predicted to be positioned similarly to SAM in the binding pocket (Figure 31 H). The 2-aminopyrimidine fragment exhibits the same hydrogen-bonding pattern (Figure 31 B) as the adenine moiety (Figure 31 F) of the natural ligand SAM (Figure 31 E). The picolinamide fragment also exhibits similar hydrogen-bonding interactions with A45 and U57 like SAM. However, instead of the second nitrogen (SAM, 2- aminopyrimidine), the amide oxygen serves ad hydrogen-bond acceptor for the ribose-2’-OH of A45 (Figure 31 D). Fragment growing can begin at the 5-position for the 2-aminopyrimidine moiety and the 4-position for the picolinamide fragments. It might be advisable to introduce a charged nitrogen or sulfur atom to address electrostatic interactions with U7 and U88 (Figure 31 G) in the binding site. As our SAR study showed, a terminal carboxylate proved to be crucial for binding due to its interactions with G11 and G58 (Chapter 7). Thus, this moiety should be attached at the terminus of the “tail”-substructure.

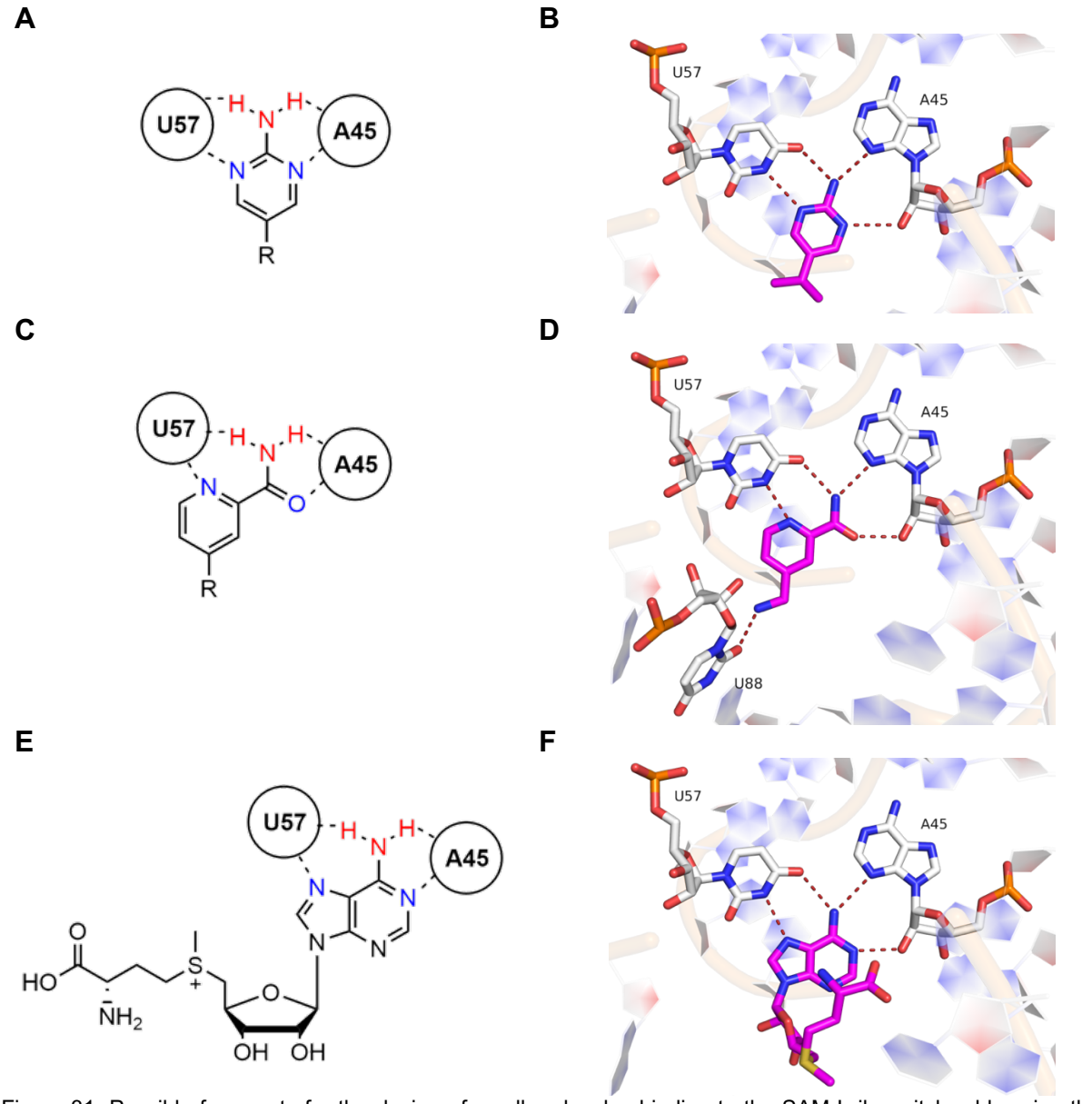


Figure 31: Possible fragments for the design of small molecules binding to the SAM-I riboswitch addressing the adenine subpocket, their interaction pattern and predicted position in the binding site in comparison to SAM. **A**, 2- Aminopyrimidine fragment structure and 2D interaction pattern with A45 and U57. **B**, Predicted position and hydrogen bonding with A45 and U57 of the 2-aminopyrimidine fragment (showed as sticks, magenta) in the binding site (showed as sticks, white). **C**, Picolinamide fragment structure and 2D interaction pattern with A45 and U57. **D**, Predicted position and hydrogen bonding with A45 and U57 of the picolinamide fragment (showed as sticks, magenta) in the binding site (showed as sticks, white). **E**, SAM structure and 2D interaction pattern with A45 and U57. **F**, Position and hydrogen bonding with A45 and U57 of SAM (showed as sticks, magenta) in the binding site (showed as sticks, white). **G**, SAM (shown as sticks, magenta) interaction pattern in the binding site (shown as sticks, white). **H**, Position overlay of SAM (shown as sticks, purple), the 2-aminopyrimidine (shown as sticks, yellow), and picolinamide (shown as sticks, teal) fragments in the binding site (shown as sticks, white). Note: H-bond acceptors are colored blue, and H-bond donors are colored red (panels **A**, **C**, **E**). H-bonding interactions are showed as red, pi-interactions as yellow and electrostatic interactions as violet dotted lines. Panels **A**, **C**, and **E** were created with ChemDraw 20, panels **B**, **D**, **F**, **G**, and **H** were created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.based on PDB: 5FJC342.

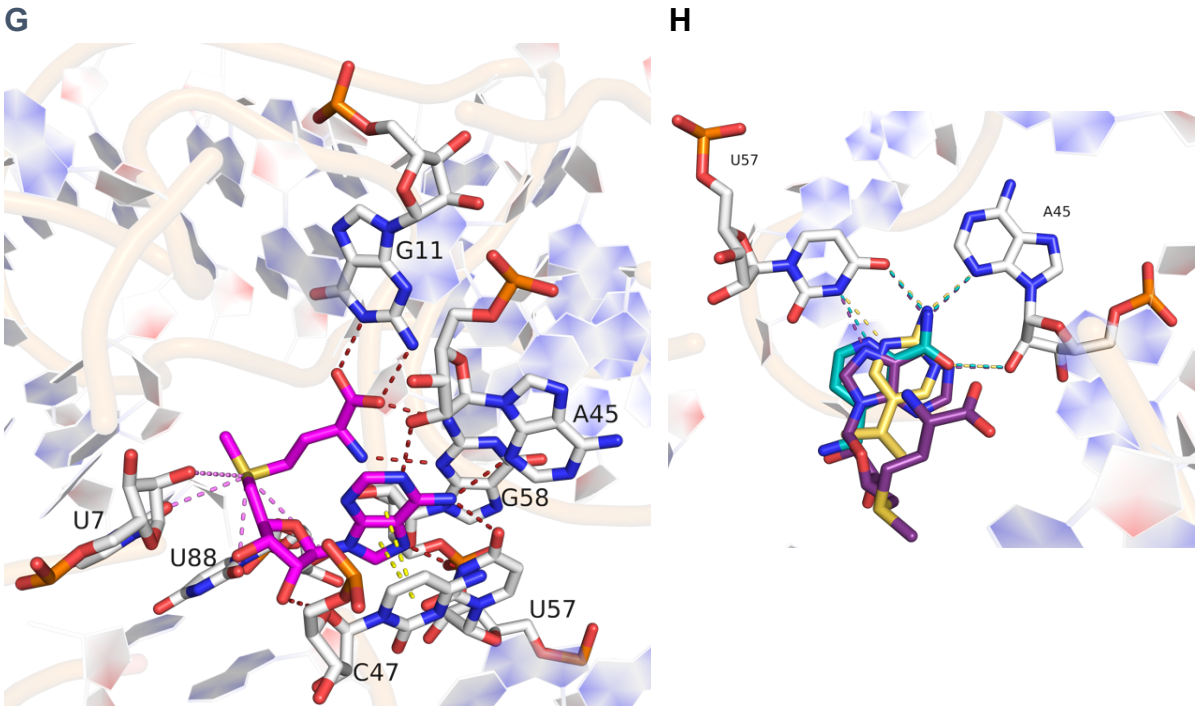


Figure 31: Possible fragments for the design of small molecules binding to the SAM-I riboswitch addressing the adenine subpocket, their interaction pattern and predicted position in the binding site in comparison to SAM. **A**, 2- Aminopyrimidine fragment structure and 2D interaction pattern with A45 and U57. **B**, Predicted position and hydrogen bonding with A45 and U57 of the 2-aminopyrimidine fragment (showed as sticks, magenta) in the binding site (showed as sticks, white). **C**, Picolinamide fragment structure and 2D interaction pattern with A45 and U57. **D**, Predicted position and hydrogen bonding with A45 and U57 of the picolinamide fragment (showed as sticks, magenta) in the binding site (showed as sticks, white). **E**, SAM structure and 2D interaction pattern with A45 and U57. **F**, Position and hydrogen bonding with A45 and U57 of SAM (showed as sticks, magenta) in the binding site (showed as sticks, white). **G**, SAM (shown as sticks, magenta) interaction pattern in the binding site (shown as sticks, white). **H**, Position overlay of SAM (shown as sticks, purple), the 2-aminopyrimidine (shown as sticks, yellow), and picolinamide (shown as sticks, teal) fragments in the binding site (shown as sticks, white). Note: H-bond acceptors are colored blue, and H-bond donors are colored red (panels **A**, **C**, **E**). H-bonding interactions are showed as red, pi-interactions as yellow and electrostatic interactions as violet dotted lines. Panels **A**, **C**, and **E** were created with ChemDraw 20, panels **B**, **D**, **F**, **G**, and **H** were created with the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.based on PDB: 5FJC342.

In our study on the HCV IRES (Chapter 5), we identified four virtual screening hits (compounds **8**, **8a**–**c**) featuring a 2-guanidino-quinazoline moiety (Figure 32) as binders for this RNA target similar to previous reported inhibitors.[228] This compound class represents a promising starting point for further optimization to enhance drug-like properties as the LE values are already in the optimal range of > 3 kcal/mol per heavy atom.[105] Furthermore, Leeson and Springthorpe[106] propose LLE values > 5 for potential drug candidates. Among our hit compounds, **8b** was found to be the most promising candidate for optimization, as it showed the highest potency (EC_{50} = 61 µM

in FRET assay, K_D = 88 µM in MST assay), a LE value of 0.31 kcal/mol and moderate lipophilicity (logP 1.1[343], LLE 2.9).[104] Potential improvements should focus on enhancing the binding affinity while not significantly altering the lipophilicity as this value is already within an acceptable range. Modifying this compound by adding polar interactions may improve affinity and LLE. The fragment-like molecular weight of 246 g/mol can be increased during the optimization process[92]. However, it is advisable to avoid a significant reduction in lipophilicity and the introduction of bulky moieties that would result in a massive increase in molecular weight during this process. Consequently, only subtle changes may be favorable to achieve this goal.

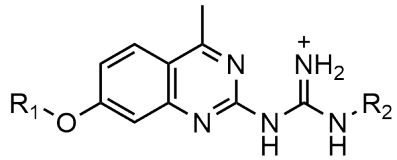


Figure 32: 2-Guanidino-quinazoline motif of HCV IRES subdomain IIa binders **8**, **8a**–**c** identified by virtual screening and validated by FRET and MST assays. Compound **8** (R_1: Et, R_2: Me), compound **8a** (R_1: Et, R_2: H), compound **8b** (R_1: Me, R_2: Me), compound **8c** (R_1: Me, R_2: H) (Chapter 5). This figure was created with Chemdraw 20.

While the concept of electrostatic anchoring with charged small molecules, such as aminoglycosides, is not novel, issues like non-selectivity and toxicity remain significant challenges.[5,10,93,99] However, our research on the preQ$_1$ and SAM-VI riboswitches demonstrates that is possible to engineer selective, high-affinity charged molecules with drug-like properties, which offers a promising strategy for overcoming these limitations. This approach could be particularly useful in optimizing validated small molecule RNA-binders with high selectivity but poor association rates. By attaching a charged “tail”-moiety, the association rate with the target RNA could be enhanced through electrostatic attraction, improving the binding profile of the small molecule. This strategy was successfully applied to the SAM-VI riboswitch model system, and a similar benefit could be derived for future optimization of SAM-I riboswitch-binding small molecules. Nevertheless, the use of electrostatic anchoring might not be useful as general design principle in RNA-ligand design. However, during ligand optimization processes, the introduction of specific interactions might be helpful to obtain drug candidates with excellent drug-like profiles and selectivity.

Although computational methodologies for studying RNA-small molecule interactions are still evolving, our findings from the preQ$_1$ riboswitch and HCV IRES studies prove

this structure-guided approach as valid alternative to the traditional HTS for RNA targets with natural ligands (e.g., riboswitches) and those without (e.g., IRES elements, lncRNAs). The integration of computer-based methodologies and biophysical methods presents a powerful approach to overcoming the challenges of selectivity and druggability in modern RNA-targeted drug design. Our research in this field provides valuable insights that could inform the future design of RNA-targeted therapeutics. Nevertheless, this strategy is contingent upon the knowledge of RNA tertiary structures, which is currently limited to approximately 2,000 RNA-related structures compared to approximately 200,000 protein-related structures.[153] As a result of the development of sophisticated methodologies for achieving a deeper understanding of RNA structure and dynamics, this limitation is gradually being reduced, thereby creating opportunities for the utilization of computer-aided approaches in structure-based drug design.

However, the recognition of RNA as a potential drug target represents a significant paradigm shift in the development of novel therapeutics, particularly for diseases such as cancer and neurodegenerative disorders, where treatment options remain limited.[1] Recently, with the approval of the splicing modulator risdiplam for the treatment of SMA, a remarkable success of this concept was achieved.[29–31] Regarding severe human diseases, three relevant classes of RNA are of particular interest: lncRNAs, short tandem repeat expansions, and miRNAs. These RNAs are associated with the regulation various cellular processes and thus, when dysregulated, can lead to a variety of pathological conditions.[344] Typically, lncRNAs are involved in multiple functions such as protein translation[345,346], and splicing[347–349]. Mutations and/or altered expression of lncRNAs results in a dysregulation of their functionality and can lead to a variety of severe diseases such as cancer (e.g., breast, lung-, and ovarian cancer)[344,350,351] and neurodegenerative diseases (e.g., Alzheimer's, Parkinson's, and Huntington's disease, amyotrophic lateral sclerosis (ALS))[352,353]. In cancer, lncRNAs typically act as oncogenes (upregulated) or tumor suppressors (downregulated) to enable tumor development and progression.[344] One prominent tumor-related lncRNA is the metastasis-associated lung adenocarcinoma transcript-1 (MALAT1), which acts as an oncogene in non-small cell lung cancer (NSCLC). MALAT1 is involved in multiple cellular functions, including the regulation of gene expression and pre-mRNA splicing, enabling cell proliferation, survival, and migration. In many NSCLC patients, this

lncRNA is upregulated, thereby enabling tumor growth and metastasis, which ultimately results in a poor prognosis for affected patients.[351,354] MALAT1 lncRNA possesses a triple helix structural motif, which was already successfully targeted with small molecules using a combination of various methods: SMM for identification, *in vitro* FRET and ITC assays, cellular assays, and molecular docking.[351] This once again highlights the promising strategy of combining different approaches in RNA-targeted drug design.

Another potential target for future research approaches in RNA-targeting small molecules is the β-site amyloid precursor protein cleaving enzyme 1-antisense (BACE1-AS) lncRNA. This conserved antisense lncRNA, which is transcribed from the opposite strand of the BACE1 gene, exhibits high expression levels in the brains of patients diagnosed with Alzheimer's disease. This encoded protein, β-secretase BACE1, is involved in the production of amyloid-β plaques, which are associated with neurodegeneration, memory loss, and dementia.[353,355] The BACE1-AS lncRNA has been successfully targeted by siRNA, which led to an improvement of the cognitive function due to reduced formation of amyloid-β plaques *in vivo*.[356] By tertiary structure determination, potential binding sites for small molecules could be explored and further investigated within a structure-based approach. Selectively targeting this RNA structure could be an important step forward in the fight against this highly progressive neurodegenerative disease.

RNA short tandem repeat (STR) expansions also hold great potential for future therapeutic development due to the fact that they possess multiple identical binding sites for small molecules.[1] Currently, there is no FDA-approved drug targeting toxic STR expansions.[357] As these structures are associated with serious diseases such as myotonic dystrophy type 1 (DM1, CUG expansions), fragile X-associated tremor/ataxia syndrome (FXTAS, CGG expansions), or Huntington's disease (CAG expansions) for which there is usually no alternative therapy, the approach of targeting the repeat expansions with small molecules represents a promising strategy to address this unmet medical need.[357,358] Normally, STRs (1–6 nucleotide repeats, e.g., CUG, CCUG, and AUUCU) occur frequently in many RNAs and are associated with functional and regulatory processes. However, when those repeats abnormally expand, they can lead to transcription termination and production of loss-of-function proteins.[358,359] For instance, in myotonic dystrophy type 1 (DM1), CUG expansions ($r(CUG)^{exp}$) are

present in the 3'-UTR of the dystrophia myotonica protein kinase (DMPK) mRNA[5]. This r(CUG)^{exp} can bind to splicing factor proteins and thus, causing their loss of function. Consequently, pre-mRNA splicing is negatively affected, and inappropriate proteins are translated.[358–360] By targeting these toxic CUG repeat expansions with small molecules, the resulting defects can be improved significantly.[360–362] Rzuczek *et al.* investigated the interaction between small molecules and r(CUG)^{exp}. For this purpose, they used a combination of computational small molecule design, biochemical assays, covalent binding techniques for target validation and fluorescence imaging.[360] The results demonstrate that addressing STR expansions with small molecules using a broad spectrum of methodologies in combination represents a viable approach that can be further advanced to develop effective small molecule drugs for the treatment of various diseases associated with STR expansions. Recently, DEL technology in combination with various *in vitro* and *in vivo* assays was used to identify new small molecule binder for the DM1-related CUG repeats. Subsequently, the lead compound was linked to bleomycin to achieve cleavage of the toxic CUG repeats.[362] In a fragment-based approach, bleomycin was also used as initiator for RNA degradation.[361] The concept of linking bleomycin as known compound that induces direct DNA and RNA cleavage[361], to selective RNA-binding small molecules might be useful for other RNA targets. However, this cleavage strongly depends on the selectivity of the RNA-binding small molecule which is linked to bleomycin.[100,360,362] Another upcoming strategy that leads to RNA degradation is the development of ribonuclease targeting chimeras (RiboTACs, Figure 33). In this case, an RNase L recruiting molecule is linked to an RNA-binding small molecule. Similar to the bleomycin induced cleavage concept, the specificity of the RiboTAC depends on the selectivity and affinity of the RNA-binder.[100] Recently, this concept was exploited to targeting the SARS-CoV-2 RNA. For this purpose, small molecule hit compounds were identified and further examined by AbsorbArray, an SMM approach, and competitive screening. The target RNA structure, the SARS-CoV-2 attenuator hairpin, which is responsible for frameshifting, was validated by Chem-CLIP. The most potent small molecule was then developed as RIBOTAC.[363] The development of RiboTACs is still ongoing against various RNA-related diseases, such as cancer[364,365], amyotrophic lateral sclerosis (ALS)[366], and Parkinson's disease[367], thereby using a variety of methods within the SBD approach. Recently, Meyer *et al*.[368] reported a DEL screening

to identify novel and more potent RNase L recruiter for the next generation of RiboTACs.

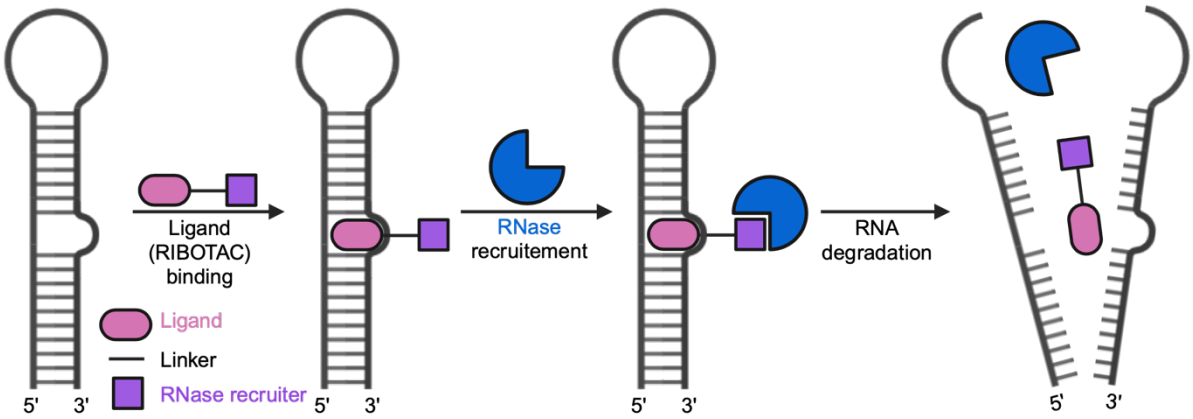


Figure 33: Simplified principle of a RiboTAC. The ligand (pink), which is linked (black line) to an RNase L recruiter (violet), binds to the RNA of interest. Then the RNase L (blue) is recruited and binds to the recruiter, resulting in RNA cleavage and degradation. This figure was created with BioRender.com. according to Costales *et al.*.[366]

All these approaches offer considerable promise for a future in which disease-relevant RNA molecules are effectively addressed by highly potent small molecules, and diseases with poor prognoses, such as cancer, neurodegenerative diseases, and those caused by multi-resistant bacteria and viruses[1], are effectively managed. Notwithstanding these achievements, no RNA-targeted small molecule has been approved to date, with the exception of risdiplam[29–31]. The development of this drug provides clear evidence of the feasibility of this approach, yet significant challenges remain. The integration of the full range of available methods in combination will be pivotal in advancing modern drug development and lay the foundation for the next generation of drugs. In this regard, this thesis not only advances our understanding of RNA-small molecule interactions but also contributes to the broader field of structure-based drug design. By addressing the dual challenges of selectivity and druggability, and leveraging cutting-edge biophysical tools, this research opens new avenues for the development of therapeutics against bacterial and viral pathogens as well as providing general strategies for RNA-targeted small molecule drug design, addressing critical challenges in global health.

9. Bibliography

1. Falese, J. P., Donlic, A. & Hargrove, A. E. Targeting RNA with Small Molecules: From Fundamental Principles Towards the Clinic. *Chem Soc Rev* **50**, 2224–2243 (2021).
2. Childs-Disney, J. L. *et al.* Targeting RNA Structures with Small Molecules. *Nat Rev Drug Discov* **21**, 736–762 (2022).
3. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How Many Drug Targets are There? *Nat Rev Drug Discov* **5**, 993–996 (2006).
4. Dixon, S. J. & Stockwell, B. R. Identifying Druggable Disease-Modifying Gene Products. *Curr Opin Chem Biol* **13**, 549–555 (2009).
5. Warner, K. D., Hajdin, C. E. & Weeks, K. M. Principles for Targeting RNA with Drug-Like Small Molecules. *Nat Rev Drug Discov* **17**, 547–558 (2018).
6. Haniff, H. S., Knerr, L., Chen, J. L., Disney, M. D. & Lightfoot, H. L. Target-Directed Approaches for Screening Small Molecules against RNA Targets. *SLAS Discovery* **25**, 869–894 (2020).
7. Lünse, C. E., Schüller, A. & Mayer, G. The Promise of Riboswitches as Potential Antibacterial Drug Targets. *International Journal of Medical Microbiology* **304**, 79–92 (2014).
8. Kavita, K. & Breaker, R. R. Discovering Riboswitches: The Past and the Future. *Trends Biochem Sci* **48**, 119–141 (2023).
9. Panchal, V. & Brenk, R. Riboswitches as Drug Targets for Antibiotics. *Antibiotics* **10**, 45 (2021).
10. Di Giorgio, A. & Duca, M. Synthetic Small-Molecule RNA Ligands: Future Prospects as Therapeutic Agents. *Medchemcomm* **10**, 1242–1255 (2019).
11. Krause, K. M., Serio, A. W., Kane, T. R. & Connolly, L. E. Aminoglycosides: An Overview. *Cold Spring Harb Perspect Med* **6**, (2016).
12. Avent, M. L., Rogers, B. A., Cheng, A. C. & Paterson, D. L. Current Use of Aminoglycosides: Indications, Pharmacokinetics and Monitoring for Toxicity. *Intern Med J* **41**, 441–449 (2011).
13. Hussen, N. H. A. *et al.* Long-Term Toxicity of Fluoroquinolones: A Comprehensive Review. *Drug Chem Toxicol* **47**, 1–12 (2023).

14. Blount, K. F. & Breaker, R. R. Riboswitches as Antibacterial Drug Targets. *Nat Biotechnol* **24**, 1558–1564 (2006).
15. De Oliveira, D. M. P. *et al.* Antimicrobial Resistance in ESKAPE Pathogens. *Clin Microbiol Rev* **31**, e00181-19 (2020).
16. Hermann, T. Small Molecules Targeting Viral RNA. *Wiley Interdiscip Rev RNA* **7**, 726–743 (2016).
17. Dibrov, S. M. *et al.* Hepatitis C Virus Translation Inhibitors Targeting the Internal Ribosomal Entry Site. *J. Med. Chem* **57**, 1694–1707 (2014).
18. Zhou, S., Rynearson, K. D., Ding, K., Brunn, N. D. & Hermann, T. Screening for Inhibitors of the Hepatitis C Virus Internal Ribosome Entry Site RNA. *Bioorg Med Chem* **21**, 6139–6144 (2013).
19. Parsons, J. *et al.* Conformational Inhibition of the Hepatitis C Virus Internal Ribosome Entry Site RNA. *Nat Chem Biol* **5**, 823–825 (2009).
20. Paulsen, R. B. *et al.* Inhibitor-Induced Structural Change in the HCV IRES Domain IIa RNA. *Proc Natl Acad Sci U S A* **107**, 7263–7268 (2010).
21. Seth, P. P. *et al.* SAR by MS: Discovery of a New Class of RNA-Binding Small Molecules for the Hepatitis C Virus: Internal Ribosome Entry Site IIA Subdomain. *J Med Chem* **48**, 7099–7102 (2005).
22. Sztuba-Solinska, J. *et al.* Identification of Biologically Active, HIV TAR RNA-Binding Small Molecules Using Small Molecule Microarrays. *J Am Chem Soc* **136**, 8402–8410 (2014).
23. Pascale, L. *et al.* Deciphering Structure-Activity Relationships in a Series of Tat/TAR Inhibitors. *J Biomol Struct Dyn* **34**, 2327–2338 (2016).
24. Tompa, D. R., Immanuel, A., Srikanth, S. & Kadhirvel, S. Trends and Strategies to Combat Viral Infections: A Review on FDA Approved Antiviral Drugs. *Int J Biol Macromol* **172**, 524–541 (2021).
25. Nemudryi, A. *et al.* CRISPR-Based Engineering of RNA Viruses. *Sci Adv* **9**, eadj8277 (2023).
26. Bennett, C. F., Krainer, A. R. & Cleveland, D. W. Antisense Oligonucleotide Therapies for Neurodegenerative Diseases. *Annu Rev Neurosci* **42**, 385–406 (2019).

27. Ranasinghe, P., Addison, M. L., Dear, J. W. & Webb, D. J. Small Interfering RNA: Discovery, Pharmacology and Clinical Development – An Introductory Review. *Br J Pharmacol* **180**, 2697–2720 (2023).
28. Kingwell, K. First CRISPR Therapy Seeks Landmark Approval. *Nat Rev Drug Discov* **22**, 339–341 (2023).
29. Ratni, H., Scalco, R. S. & Stephan, A. H. Risdiplam, the First Approved Small Molecule Splicing Modifier Drug as a Blueprint for Future Transformative Medicines. *ACS Med Chem Lett* **12**, 874–877 (2021).
30. Ratni, H. *et al.* Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (SMN2) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA). *J Med Chem* **61**, 6501–6517 (2018).
31. Paik, J. Risdiplam: A Review in Spinal Muscular Atrophy. *CNS Drugs* **36**, 401–410 (2022).
32. Yu, A. M., Choi, Y. H. & Tu, M. J. RNA Drugs and RNA Targets for Small Molecules: Principles, Progress, and Challenges. *Pharmacol Rev* **72**, 862–898 (2020).
33. U. S. Food and Drug Administration, Center for Drug Evaluation and Research. *Evrysdi NDA 213535 Approval Letter, August 7, 2020.* (2020).
34. Markati, T., Fisher, G., Ramdas, S. & Servais, L. Risdiplam: An Investigational Survival Motor Neuron 2 (SMN2) Splicing Modifier for Spinal Muscular Atrophy (SMA). *Expert Opin Investig Drugs* **31**, 451–461 (2022).
35. Wilson, D. N. *et al.* The Oxazolidinone Antibiotics Perturb the Ribosomal Peptidyl-Transferase Center and Effect tRNA Positioning. *Proc Natl Acad Sci U.S.A.* **105**, 13339–13344 (2008).
36. Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M. & Yusupova, G. A New Understanding of the Decoding Principle on the Ribosome. *Nature* **484**, 256–259 (2012).
37. Brodersen, D. E. *et al.* The Structural Basis for the Action of the Antibiotics Tetracycline, Pactamycin, and Hygromycin B on the 30S Ribosomal Subunit. *Cell* **103**, 1143–1154 (2000).
38. Fourmy, D. *et al.* Structure of the A Site of Escherichia Coli 16S Ribosomal RNA Complexed with an Aminoglycoside Antibiotic. *Science (1979)* **274**, 1367–1371 (1996).

39. Blondeau, J. M., DeCarolis, E., Metzler, K. L. & Hansen, G. T. The Macrolides. *Expert Opin Investig Drugs* **11**, 189–215 (2002).
40. Dibrov, S. M. *et al.* Structure of a Hepatitis C Virus RNA Domain in Complex with a Translation Inhibitor Reveals a Binding Mode Reminiscent of Riboswitches. *Proceedings of the National Academy of Sciences (PNAS)* **109**, 5223–5228 (2012).
41. Howe, J. A. *et al.* Selective Small-Molecule Inhibition of an RNA Structural Element. *Nature* **526**, 672–677 (2015).
42. Howe, J. A. *et al.* Atomic Resolution Mechanistic Studies of Ribocil: A Highly Selective Unnatural Ligand Mimic of the E. Coli FMN Riboswitch. *RNA Biol* **13**, 946–954 (2016).
43. Rizvi, N. F. & Smith, G. F. RNA as a Small Molecule Druggable Target. *Bioorg Med Chem Lett* **27**, 5083–5088 (2017).
44. Kaur, J. *et al.* RNA-Small-Molecule Interaction: Challenging the “Undruggable” Tag. *J Med Chem* **67**, 4259–4297 (2024).
45. Velagapudi, S. P. *et al.* Design of a Small Molecule Against an Oncogenic Noncoding RNA. *Proc Natl Acad Sci U S A* **113**, 5898–5903 (2016).
46. Li, Y. & Disney, M. D. Precise Small Molecule Degradation of a Noncoding RNA Identifies Cellular Binding Sites and Modulates an Oncogenic Phenotype. *ACS Chem Biol* **13**, 3065–3071 (2018).
47. Tadesse, K. & Benhamou, R. I. Targeting MicroRNAs with Small Molecules. *Noncoding RNA* **10**, 17 (2024).
48. Costales, M. G. *et al.* Small Molecule Inhibition of microRNA-210 Reprograms an Oncogenic Hypoxic Circuit. *J Am Chem Soc* **139**, 3446–3455 (2017).
49. Ganser, L. R., Kelly, M. L., Herschlag, D. & Al-Hashimi, H. M. The Roles of Structural Dynamics in the Cellular Functions of RNAs. *Nat Rev Mol Cell Biol* **20**, 474–489 (2019).
50. McGovern-Gooch, K. R. & Baird, N. J. Fluorescence-Based Investigations of RNA-Small Molecule Interactions. *Methods* **167**, 54–65 (2019).
51. Costales, M. G., Childs-Disney, J. L., Haniff, H. S. & Disney, M. D. How We Think about Targeting RNA with Small Molecules. *J Med Chem* **63**, 8880–8900 (2020).

52. Bush, J. A. *et al.* Systematically Studying the Effect of Small Molecules Interacting with RNA in Cellular and Preclinical Models. *ACS Chem Biol* **16**, 1111–1127 (2021).
53. Yang, W. Y., Wilson, H. D., Velagapudi, S. P. & Disney, M. D. Inhibition of Non-ATG Translational Events in Cells Via Covalent Small Molecules Targeting RNA. *J Am Chem Soc* **137**, 5336–5345 (2015).
54. Velagapudi, S. P., Li, Y. & Disney, M. D. A Cross-Linking Approach to Map Small Molecule-RNA Binding Sites in Cells. *Bioorg Med Chem Lett* **29**, 1532–1536 (2019).
55. Suresh, B. M. *et al.* A General Fragment-Based Approach to Identify and Optimize Bioactive Ligands Targeting RNA. *Proceedings of the National Academy of Sciences (PNAS*= **117**, 33197–33203 (2020).
56. Balaratnam, S. *et al.* A Chemical Probe Based on the PreQ1 Metabolite Enables Transcriptome-Wide Mapping of Binding Sites. *Nat Commun* **12**, (2021).
57. Costales, M. G., Childs-Disney, J. L., Haniff, H. S. & Disney, M. D. How We Think about Targeting RNA with Small Molecules. *J Med Chem* **63**, 8880–8900 (2020).
58. Ritchie, T. J. & Macdonald, S. J. F. How Drug-Like are 'Ugly' Drugs: Do Drug-Likeness Metrics Predict ADME Behaviour in Humans? *Drug Discov Today* **19**, 489–495 (2014).
59. Barbachyn, M. R. & Ford, C. W. Oxazolidinone Structure-Activity Relationships Leading to Linezolid. *Angewandte Chemie - International Edition* **42**, 2010–2023 (2003).
60. Rizvi, N. F. *et al.* Discovery of Selective RNA-Binding Small Molecules by Affinity-Selection Mass Spectrometry. *ACS Chem Biol* **13**, 820–831 (2018).
61. Wicks, S. L. & Hargrove, A. E. Fluorescent Indicator Displacement Assays to Identify and Characterize Small Molecule Interactions with RNA. *Methods* **167**, 3–14 (2019).
62. Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. & Duhr, S. Molecular Interaction Studies Using Microscale Thermophoresis. *Assay Drug Dev Technol* **9**, 342–353 (2011).
63. Connelly, C. M. *et al.* Synthetic ligands for PreQ 1 riboswitches provide structural and mechanistic insights into targeting RNA tertiary structure. *Nat Commun* **10**, (2019).

64. Rizvi, N. F. & Nickbarg, E. B. RNA-ALIS: Methodology for Screening Soluble RNAs as Small Molecule Targets Using ALIS Affinity-Selection Mass Spectrometry. *Methods* **167**, 28–38 (2019).
65. Brenner, S. & Lerner, R. A. Encoded Combinatorial Chemistry. *Proc Natl Acad Sci U.S.A.* **89**, 5381–5383 (1992).
66. Peterson, A. A. & Liu, D. R. Small-Molecule Discovery Through DNA-Encoded Libraries. *Nat Rev Drug Discov* **22**, 699–722 (2023).
67. Chen, Q. *et al.* Expanding the DNA-Encoded Library Toolbox: Identifying Small Molecules Targeting RNA. *Nucleic Acids Res* **50**, E67 (2022).
68. Benhamou, R. I. *et al.* DNA-Encoded Library Versus RNA-Encoded Library Selection Enables Design of an Oncogenic Noncoding RNA Inhibitor. *Proceedings of the National Academy of Sciences (PNAS)* **119**, e2114971119 (2021).
69. Nguyen, H. H., Park, J., Kang, S. & Kim, M. Surface Plasmon Resonance: A Versatile Technique for Biosensor Applications. *Sensors* **15**, 10481–10510 (2015).
70. Englebienne, P., Hoonacker, A. Van & Verhas, M. Surface Plasmon Resonance: Principles, Methods and Applications in Biomedical Sciences. *Spectroscopy* **17**, 255–273 (2003).
71. Butcher, S. E. & Pyle, A. M. The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks. *Acc Chem Res* **44**, 1302–1311 (2011).
72. Jones, C. P. & Ferré-D'Amaré, A. R. RNA Quaternary Structure and Global Symmetry. *Trends Biochem Sci* **40**, 211–220 (2015).
73. Nachtergaele, S. & He, C. The Emerging Biology of RNA Post-Transcriptional Modifications. *RNA Biol* **14**, 156–163 (2017).
74. Wilkinson, E., Cui, Y. H. & He, Y. Y. Context-Dependent Roles of RNA Modifications in Stress Responses and Diseases. *Int J Mol Sci* **22**, 1–17 (2021).
75. Cappannini, A. *et al.* MODOMICS: A database of RNA Modifications and related Information. 2023 Update. *Nucleic Acids Res* **52**, D239–D244 (2024).
76. Deng, J. *et al.* RNA Structure Determination: From 2D to 3D. *Fundamental Research* **3**, 727–737 (2023).

77. Zhang, K. *et al.* Cryo-EM Structure of a 40 kDa SAM-IV Riboswitch RNA at 3.7 Å Resolution. *Nat Commun* **10**, (2019).
78. Wehler, T. & Brenk, R. Structure-Based Discovery of Small Molecules Binding to RNA. in *RNA Therapeutics* (ed. Garner, A. L.) 47–77 (Springer Verlag, 2018).
79. Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E. & Weeks, K. M. RNA Motif Discovery by SHAPE and Mutational Profiling (SHAPE-MaP). *Nat Methods* **11**, 959–965 (2014).
80. Smola, M. J., Rice, G. M., Busan, S., Siegfried, N. A. & Weeks, K. M. Selective 2'-Hydroxyl Acylation Analyzed by Primer Extension and Mutational Profiling (SHAPE-MaP) for Direct, Versatile and Accurate RNA Structure Analysis. *Nat Protoc* **10**, 1643–1669 (2015).
81. Weeks, K. M. & Mauger, D. M. Exploring RNA Structural Codes with SHAPE Chemistry. *Acc Chem Res* **44**, 1280–1291 (2011).
82. Watkins, A. M., Rangan, R. & Das, R. FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. *Structure* **28**, 963-976.e6 (2020).
83. Freire, F. *et al.* A Simple NMR Analysis of the Protonation Equilibrium That Accompanies Aminoglycoside Recognition: Dramatic Alterations in the Neomycin-B Protonation State Upon Binding to a 23-mer RNA Aptamer. *Chemical Communications* **2**, 174–176 (2007).
84. Guilbert, C. & James, T. L. Docking to RNA via Root-Mean-Square-Deviation-Driven Energy Minimization with Flexible Ligands and Flexible Targets. *J Chem Inf Model* **48**, 1257–1268 (2008).
85. Spyrakis, F. *et al.* The Roles of Water in the Protein Matrix: A Largely Untapped Resource for Drug Discovery. *J Med Chem* **60**, 6781–6828 (2017).
86. Carothers, J. M., Oestreich, S. C., Davis, J. H. & Szostak, J. W. Informational Complexity and Functional Activity of RNA Structures. *J Am Chem Soc* **126**, 5130–5137 (2004).
87. Draper, D. E., Grilley, D. & Soto, A. M. Ions and RNA Folding. *Annu Rev Biophys Biomol Struct* **34**, 221–243 (2005).
88. Chu, V. B., Bai, Y., Lipfert, J., Herschlag, D. & Doniach, S. A Repulsive Field: Advances in the Electrostatics of the Ion Atmosphere. *Curr Opin Chem Biol* **12**, 619–625 (2008).

89. Lipfert, J., Doniach, S., Das, R. & Herschlag, D. Understanding Nucleic Acid-Ion Interactions. *Annu Rev Biochem* **83**, 813–841 (2014).
90. Draper, D. E. RNA Folding: Thermodynamic and Molecular Descriptions of the Roles of Ions. *Biophys J* **95**, 5489–5495 (2008).
91. Manning, G. S. The Molecular Theory of Polyelectrolyte Solutions with Applications to the Electrostatic Properties of Polynucleotides. *Quarterly Reviews of Biophysics n* **2**, 179–246 (1978).
92. Lipinski, C. A. Lead- and Drug-Like Compounds: The Rule-of-Five Revolution. *Drug Discov Today Technol* **1**, 337–341 (2004).
93. Sengupta, R. N. & Herschlag, D. Enhancement of RNA/Ligand Association Kinetics via an Electrostatic Anchor. *Biochemistry* **58**, 2760–2768 (2019).
94. Gleitsman, K. R., Sengupta, R. N. & Herschlag, D. Slow Molecular Recognition by RNA. *RNA* **23**, 1745–1753 (2017).
95. Zhang, Z., Witham, S. & Alexov, E. On the Role of Electrostatics in Protein-Protein Interactions. *Phys Biol* **8**, (2011).
96. Lo Conte, L., Chothia, C. & Janin, J. The Atomic Structure of Protein-Protein Recognition Sites. *J Mol Biol* **285**, 2177–2198 (1999).
97. Hemsath, L., Dvorsky, R., Fiegen, D., Carlier, M. F. & Ahmadian, M. R. An Electrostatic Steering Mechanism of Cdc42 Recognition by Wiskott-Aldrich Syndrome Proteins. *Mol Cell* **20**, 313–324 (2005).
98. Schreiber, G., Haran, G. & Zhou, H. X. Fundamental Aspects of Protein - Protein Association Kinetics. *Chem Rev* **109**, 839–860 (2009).
99. Aradi, K., Di Giorgio, A. & Duca, M. Aminoglycoside Conjugation for RNA Targeting: Antimicrobials and Beyond. *Chemistry - A European Journal* **26**, 12273–12309 (2020).
100. Kovachka, S. *et al.* Small Molecule Approaches to Targeting RNA. *Nat Rev Chem* **8**, 120–135 (2024).
101. Koehn, J. T., Felder, S. & Weeks, K. M. Innovations in Targeting RNA by Fragment-Based Ligand Discovery. *Curr Opin Struct Biol* **79**, 102550 (2023).
102. Lundquist, K. P., Panchal, V., Gotfredsen, C. H., Brenk, R. & Clausen, M. H. Fragment-Based Drug Discovery for RNA Targets. *ChemMedChem* **16**, 2588–2603 (2021).

103. Schultes, S. *et al.* Ligand Efficiency As a Guide in Fragment Hit Selection and Optimization. *Drug Discov Today Technol* **7**, e157–e162 (2010).
104. Hopkins, A. L., Keserü, G. M., Leeson, P. D., Rees, D. C. & Reynolds, C. H. The Role of Ligand Efficiency Metrics in Drug Discovery. *Nat Rev Drug Discov* **13**, 105–121 (2014).
105. Hopkins, A. L., Groom, C. R. & Alex, A. Ligand Efficiency: A Useful Metric For Lead Selection. *Drug Discov Today* **9**, 430–431 (2004).
106. Leeson, P. D. & Springthorpe, B. The Influence of Drug-Like Concepts on Decision-Making in Medicinal Chemistry. *Nat Rev Drug Discov* **6**, 881–890 (2007).
107. Reynolds, C. H., Bembenek, S. D. & Tounge, B. A. The Role of Molecular Size in Ligand Efficiency. *Bioorg Med Chem Lett* **17**, 4258–4261 (2007).
108. Reynolds, C. H., Tounge, B. A. & Bembenek, S. D. Ligand Binding Efficiency: Trends, Physical Basis, and Implications. *J Med Chem* **51**, 2432–2438 (2008).
109. Nissink, J. W. M. Simple Size-Independent Measure of Ligand Efficiency. *J Chem Inf Model* **49**, 1617–1622 (2009).
110. Orita, M., Ohno, K. & Niimi, T. Two 'Golden Ratio' Indices in Fragment-Based Drug Discovery. *Drug Discov Today* **14**, 321–328 (2009).
111. Zeller, M. J. *et al.* Subsite Ligand Recognition and Cooperativity in the TPP Riboswitch: Implications for Fragment-Linking in RNA Ligand Discovery. *ACS Chem Biol* **17**, 438–448 (2022).
112. Cressina, E., Chen, L., Abell, C., Leeper, F. J. & Smith, A. G. Fragment Screening Against the Thiamine Pyrophosphate Riboswitch thiM. *Chem Sci* **2**, 157–165 (2011).
113. Suresh, B. M. *et al.* Low-Molecular Weight Small Molecules Can Potently Bind RNA and Affect Oncogenic Pathways in Cells. *J Am Chem Soc* **144**, 20815–20824 (2022).
114. Binas, O. *et al.* 19F NMR-Based Fragment Screening for 14 Different Biologically Active RNAs and 10 DNA and Protein Counter-Screens. *ChemBioChem* **22**, 423–433 (2021).
115. Lee, M. K. *et al.* A Novel Small-Molecule Binds to the Influenza A Virus RNA Promoter and Inhibits Viral Replication. *Chemical Communications* **50**, 368–370 (2014).

116. Tam, B. *et al.* Discovery of Small-Molecule Inhibitors Targeting the Ribosomal Peptidyl Transferase Center (PTC) of: M. tuberculosis. *Chem Sci* **10**, 8764–8767 (2019).
117. Wohlleben, W., Mast, Y., Stegmann, E. & Ziemert, N. Antibiotic Drug Discovery. *Microb Biotechnol* **9**, 541–548 (2016).
118. Katz, L. & Baltz, R. H. Natural Product Discovery: Past, Present, and Future. *J Ind Microbiol Biotechnol* **43**, 155–176 (2016).
119. Bérdy, J. Thoughts and Facts about Antibiotics: Where we Are Now and Where We Are Heading. *Journal of Antibiotics* **65**, 385–395 (2012).
120. Monciardini, P., Iorio, M., Maffioli, S., Sosio, M. & Donadio, S. Discovering New Bioactive Molecules from Microbial Sources. *Microb Biotechnol* **7**, 209–220 (2014).
121. Lewis, K. Platforms for Antibiotic Discovery. *Nat Rev Drug Discov* **12**, 371–387 (2013).
122. Nathan, C. Antibiotics at the Crossroads. *Nature* **431**, 899–902 (2004).
123. Nathan, C. & Cars, O. Antibiotic Resistance — Problems, Progress, and Prospects. *New England Journal of Medicine* **371**, 1761–1763 (2014).
124. Gajdács, M. The Concept of an Ideal Antibiotic: Implications for Drug Design. *Molecules* **24**, 892 (2019).
125. Cars, O. Securing Access to Effective Antibiotics for Current and Future Generations. Whose Responsibility? *Ups J Med Sci* **119**, 209–214 (2014).
126. World Health Organization (WHO). *Prioritization of Pathogens to Guide Discovery, Research and Development of New Antibiotics for Drug-Resistant Bacterial Infections, Including Tuberculosis*. (2017).
127. Spellberg, B. The Future of Antibiotics. *Crit Care* **18**, 1–7 (2014).
128. Projan, S. J. Why is Big Pharma Getting Out of Antibacterial Drug Discovery? *Curr Opin Microbiol* **6**, 427–430 (2003).
129. U. S. Food and Drug Administration, Center for Drug Evaluation and Research. *Zevtera NDA 218275 Approval Letter, April 3, 2024*.
130. Farrell, D. J., Flamm, R. K., Sader, H. S. & Jones, R. N. Ceftobiprole Activity Against Over 60,000 Clinical Bacterial Pathogens Isolated in Europe, Turkey, and Israel from 2005 to 2010. *Antimicrob Agents Chemother* **58**, 3882–3888 (2014).

131. Shi, Z. *et al.* A Comprehensive Overview of the Antibiotics Approved in the Last Two Decades: Retrospects and Prospects. *Molecules* **28**, (2023).
132. Matzner, D. & Mayer, G. (Dis)similar Analogues of Riboswitch Metabolites as Antibacterial Lead Compounds. *J Med Chem* **58**, 3275–3286 (2015).
133. Deigan, K. E. & Ferré-D'Amaré, A. R. Riboswitches: Discovery of Drugs That Target Bacterial Gene-Regulatory RNAs. *Acc Chem Res* **44**, 1329–1338 (2011).
134. Mandal, M. & Breaker, R. R. Gene Regulation by Riboswitches. *Nat Rev Mol Cell Biol* **5**, 451–463 (2004).
135. Serganov, A. The Long and the Short of Riboswitches. *Curr Opin Struct Biol* **19**, 251–259 (2009).
136. Serganov, A. & Nudler, E. A Decade of Riboswitches. *Cell* **152**, 17–24 (2013).
137. Pavlova, N., Kaloudas, D. & Penchovsky, R. Riboswitch Distribution, Structure, and Function in Bacteria. *Gene* **708**, 38–48 (2019).
138. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of Gene Expression by a Natural Metabolite-Responsive Ribozyme. *Nature* **428**, 281–286 (2004).
139. Lee, E. R., Baker, J. L., Weinberg, Z., Sudarsan, N. & Breaker, R. R. An Allosteric Self-Splicing Ribozyme Triggered by a Bacterial Second Messenger. *Science (1979)* **329**, 845–848 (2010).
140. Chen, A. G. Y., Sudarsan, N. & Breaker, R. R. Mechanism for Gene Control by a Natural Allosteric Group I Ribozyme. *RNA* **17**, 1967–1972 (2011).
141. Wickiser, J. K., Winkler, W. C., Breaker, R. R. & Crothers, D. M. The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch. *Mol Cell* **18**, 49–60 (2005).
142. Wickiser, J. K., Cheah, M. T., Breaker, R. R. & Crothers, D. M. The Kinetics of Ligand Binding by an Adenine-Sensing Riboswitch. *Biochemistry* **44**, 13404–13414 (2005).
143. Guedich, S. *et al.* Quantitative and Predictive Model of Kinetic Regulation by E. Coli TPP Riboswitches. *RNA Biol* **13**, 373–390 (2016).
144. Lemay, J. F. *et al.* Comparative Study Between Transcriptionally- and Translationally-Acting Adenine Riboswitches Reveals Key Differences in Riboswitch Regulatory Mechanisms. *PLoS Genet* **7**, (2011).

145. Uhm, H., Kang, W., Ha, K. S., Kang, C. & Hohng, S. Single-Molecule FRET Studies on the Cotranscriptional Folding of a Thiamine Pyrophosphate Riboswitch. *Proc Natl Acad Sci U S A* **115**, 331–336 (2017).
146. Hua, B. *et al.* Real-Time Monitoring of Single ZTP Riboswitches Reveals a Complex and Kinetically Controlled Decision Landscape. *Nat Commun* **11**, (2020).
147. Nahvi, A. *et al.* Genetic Control by a Metabolite Binding mRNA. *Chem Biol* **9**, 1043–1049 (2002).
148. Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine Derivatives Bind Messenger RNAs Directly to Regulate Bacterial Gene Expression. *Nature* **419**, 952–956 (2002).
149. Winkler, W. C., Cohen-Chalamish, S. & Breaker, R. R. An mRNA Structure That Controls Gene Expression by Binding FMN. *Proceedings of the National Academy of Science (PNAS)* **99**, 15908–15913 (2002).
150. Mironov, A. S. *et al.* Sensing Small Molecules by Nascent RNA: A Mechanism to Control Transcription in Bacteria. *Cell* **111**, 747–756 (2002).
151. Breaker, R. R. Riboswitches and the RNA World. *Cold Spring Harb Perspect Biol* **4**, (2012).
152. McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E. & Breaker, R. R. Riboswitch Diversity and Distribution. *Rna* **23**, 995–1011 (2017).
153. Berman, H. M. *et al. The Protein Data Bank*. *Nucleic Acids Research* vol. 28 http://www.rcsb.org/pdb/status.html (2000).
154. Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R. & Patel, D. J. Structural Basis for Gene Regulation by a Thiamine Pyrophosphate-Sensing Riboswitch. *Nature* **441**, 1167–1171 (2006).
155. Kang, M., Peterson, R. & Feigon, J. Structural Insights into Riboswitch Control of the Biosynthesis of Queuosine, a Modified Nucleotide Found in the Anticodon of tRNA. *Mol Cell* **33**, 784–790 (2009).
156. Jenkins, J. L., Krucinska, J., McCarty, R. M., Bandarian, V. & Wedekind, J. E. Comparison of a PreQ1 riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *Journal of Biological Chemistry* **286**, 24626–24637 (2011).

157. Liberman, J. A., Salim, M., Krucinska, J. & Wedekind, J. E. Structure of a Class II PreQ1 Riboswitch Reveals Ligand Recognition by a New Fold. *Nat Chem Biol* **9**, 353–355 (2013).
158. Ren, A., Rajashankar, K. R. & Patel, D. J. Fluoride Ion Encapsulation by Mg2+ Ions and Phosphates in a Fluoride Riboswitch. *Nature* **486**, 85–89 (2012).
159. Pleij, C. W. A. Pseudoknots: A New Motif in the RNA Game. *Trends Biochem Sci* **15**, 143–147 (1990).
160. Price, I. R., Grigg, J. C. & Ke, A. Common Themes and Differences in SAM Recognition Among SAM Riboswitches. *Biochim Biophys Acta* **1839**, 931–938 (2014).
161. Mirihana Arachchilage, G., Sherlock, M. E., Weinberg, Z. & Breaker, R. R. SAM-VI RNAs Selectively Bind S-Adenosylmethionine and Exhibit Similarities to SAM-III Riboswitches. *RNA Biol* **15**, 371–378 (2018).
162. Sun, A. *et al.* SAM-VI riboswitch structure and signature for ligand discrimination. *Nat Commun* **10**, (2019).
163. Breaker, R. R. The Biochemical Landscape of Riboswitch Ligands. *Biochemistry* **61**, 137–149 (2022).
164. Barrick, J. E. & Breaker, R. R. The Distributions, Mechanisms, and Structures of Metabolite-Binding Riboswitches. *Genome Biol* **8**, R239 (2007).
165. Peselis, A. & Serganov, A. Structure and Function of Pseudoknots Involved in Gene Expression Control. *Wiley Interdiscip Rev RNA* **5**, 803–822 (2014).
166. Pavlova, N. & Penchovsky, R. Genome-Wide Bioinformatics Analysis of FMN, SAM-I, GlmS, TPP, Lysine, Purine, Cobalamin, and SAH Riboswitches for Their Applications as Allosteric Antibacterial Drug Targets in Human Pathogenic Bacteria. *Expert Opin Ther Targets* **23**, 631–643 (2019).
167. Wood, S., Ferré-D'Amaré, A. R. & Rueda, D. Allosteric Tertiary Interactions Preorganize the c-di-GMP riboswitch and Accelerate Ligand Binding. *ACS Chem Biol* **7**, 920–927 (2012).
168. Sudarsan, N. *et al.* Riboswitches in Eubacteria Sense the Second Messenger Cyclic di-GMP. *Science (1979)* **321**, 411–413 (2008).
169. Smith, K. D., Lipchock, S. V., Livingston, A. L., Shanahan, C. A. & Strobel, S. A. Structural and Biochemical Determinants of Ligand Binding by the c-di-GMP Riboswitch. *Biochemistry* **49**, 7351–7359 (2010).

170. Pavlova, N. & Penchovsky, R. Bioinformatics and Genomic Analyses of the Suitability of Eight Riboswitches for Antibacterial Drug Targets. *Antibiotics* **11**, (2022).
171. Baker, J. L. *et al.* Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science (1979)* **335**, 233–235 (2012).
172. Batey, R. T., Gilbert, S. D. & Montange, R. K. Structure of a Natural Guanine-Responsive Riboswitch Complexed with the Metabolite Hypoxanthine. *Nature* **432**, 411–415 (2004).
173. Kang, M., Eichhorn, C. D. & Feigon, J. Structural Determinants for Ligand Capture by a Class II PreQ1 Riboswitch. *Proc Natl Acad Sci U S A* **111**, (2014).
174. Lenkeit, F., Eckert, I., Hartig, J. S. & Weinberg, Z. Discovery and Characterization of a Fourth Class of Guanidine Riboswitches. *Nucleic Acids Res* **48**, 12889–12899 (2020).
175. Regulski, E. E. *et al.* A Widespread Riboswitch Candidate That Controls Bacterial Genes Involved in Molybdenum Cofactor and Tungsten Cofactor Metabolism. *Mol Microbiol* **68**, 918–932 (2008).
176. Meyer, M. M., Roth, A., Chervin, S. M., Garcia, G. A. & Breaker, R. R. Confirmation of a Second Natural PreQ1 Aptamer Class in Streptococcaceae Bacteria. *RNA* **14**, 685–695 (2008).
177. Lu, C. *et al.* Crystal Structures of the SAM-III/SMK Riboswitch Reveal the SAM-Dependent Translation Inhibition Mechanism. *Nat Struct Mol Biol* **15**, 1076–1083 (2008).
178. Gong, S., Wang, Y., Wang, Z., Wang, Y. & Zhang, W. Reversible-Switch Mechanism of the SAM-III Riboswitch. *Journal of Physical Chemistry B* **120**, 12305–12311 (2016).
179. Montange, R. K. & Batey, R. T. Structure of the S-adenosylmethionine Riboswitch Regulatory mRNA Element. *Nature* **441**, 1172–1175 (2006).
180. Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Regulation of Riboflavin Biosynthesis and Transport Genes in Bacteria by Transcriptional and Translational Attenuation. *Nucleic Acids Res* **30**, 3141–3151 (2002).
181. Roth, A. *et al.* A riboswitch Selective for the Queuosine Precursor PreQ1 Contains an Unusually Small Aptamer Domain. *Nat Struct Mol Biol* **14**, 308–317 (2007).

182. Klein, D. J., Edwards, T. E. & Ferré-D'Amaré, A. R. Cocrystal Structure of a Class I PreQ1 Riboswitch Reveals a Pseudoknot Recognizing an Essential Hypermodified Nucleobase. *Nat Struct Mol Biol* **16**, 343–344 (2009).
183. McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. Structural, Functional, and Taxonomic Diversity of Three PreQ1 Riboswitch Classes. *Chem Biol* **21**, 880–889 (2014).
184. Liberman, J. A. *et al.* Structural Analysis of a Class III PreQ1 Riboswitch Reveals an Aptamer Distant from a Ribosome-Binding Site Regulated by Fast Dynamics. *Proc Natl Acad Sci U S A* **112**, E3485–E3494 (2015).
185. Kerpedjiev, P., Hammer, S. & Hofacker, I. L. Forna (Force-Directed RNA): Simple and Effective Online RNA Secondary Structure Diagrams. *Bioinformatics* **31**, 3377–3379 (2015).
186. Schroeder, G. M. *et al.* Analysis of a PreQ1-I Riboswitch in Effector-Free and Bound States Reveals a Metabolite-Programmed Nucleobase-Stacking Spine That Controls Gene Regulation. *Nucleic Acids Res* **48**, 8146–8164 (2020).
187. Eichhorn, C. D., Kang, M. & Feigon, J. Structure and Function of PreQ1 Riboswitches. *Biochim Biophys Acta* **1839**, 939–950 (2014).
188. Iwata-Reuyl, D. Biosynthesis of the 7-Deazaguanosine Hypermodified Nucleosides of Transfer RNA. *Bioorg Chem* **31**, 24–43 (2003).
189. Phillips, G. *et al.* Biosynthesis of 7-Deazaguanosine-Modified tRNA Nucleosides: A New Role for GTP Cyclohydrolase I. *J Bacteriol* **190**, 7876–7884 (2008).
190. McCarty, R. M. & Bandarian, V. Biosynthesis of Pyrrolopyrimidines. *Bioorg Chem* **43**, 15–25 (2012).
191. Suddala, K. C. *et al.* Single Transcriptional and Translational PreQ1 Riboswitches Adopt Similar Pre-Folded Ensembles That Follow Distinct Folding Pathways Into the Same Ligand-Bound Structure. *Nucleic Acids Res* **41**, 10462–10475 (2013).
192. Schroeder, G. M., Kiliushik, D., Jenkins, J. L. & Wedekind, J. E. Structure and Function Analysis of a Type III PreQ1-I Riboswitch from Escherichia Coli Reveals Direct Metabolite Sensing by the Shine-Dalgarno Sequence. *Journal of Biological Chemistry* **299**, (2023).

193. Liberman, J. A., Bogue, J. T., Jenkins, J. L., Salim, M. & Wedekind, J. E. ITC Analysis of Ligand Binding to PreQ1 Riboswitches. in *Methods in Enzymology* (ed. Burke-Aguero, D. H.) vol. 549 435–450 (Elsevier, 2014).
194. Moon, M. H., Hilimire, T. A., Sanders, A. M. & Schneekloth, J. S. Measuring RNA-Ligand Interactions with Microscale Thermophoresis. *Biochemistry* **57**, 4638–4643 (2018).
195. Wang, Y., Parmar, S., Schneekloth, J. S. & Tiwary, P. Interrogating RNA-Small Molecule Interactions with Structure Probing and Artificial Intelligence-Augmented Molecular Simulations. *ACS Cent Sci* **8**, 741–748 (2022).
196. Sekowska, A., Kung, H.-F. & Danchin, A. Sulfur Metabolism in Escherichia Coli and Related Bacteria: Facts and Fiction. *J Mol Microbiol Biotechnol* **2**, 145–177 (2000).
197. Batey, R. T. Recognition of S-Adenosylmethionine by Riboswitches. *Wiley Interdiscip Rev RNA* **2**, 299–311 (2011).
198. Fontecave, M., Atta, M. & Mulliez, E. S-Adenosylmethionine: Nothing Goes to Waste. *Trends Biochem Sci* **29**, 243–249 (2004).
199. Struck, A. W., Thompson, M. L., Wong, L. S. & Micklefield, J. S-Adenosyl-Methionine-Dependent Methyltransferases: Highly Versatile Enzymes in Biocatalysis, Biosynthesis and Other Biotechnological Applications. *ChemBioChem* **13**, 2642–2655 (2012).
200. Wang, J. X. & Breaker, R. R. Riboswitches that sense S-Adenosylmethionine and S-Adenosylhomocysteine. *Biochemistry and Cell Biology* **86**, 157–168 (2008).
201. Wang, J. X., Lee, E. R., Morales, D. R., Lim, J. & Breaker, R. R. Riboswitches that Sense S-Adenosylhomocysteine and Activate Genes Involved in Coenzyme Recycling. *Mol Cell* **29**, 691–702 (2008).
202. Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E. & Breaker, R. R. An mRNA Structure That Controls Gene Expression by Binding S-Adenosylmethionine. *Nat Struct Biol* **10**, 701–707 (2003).
203. Corbino, K. A. *et al.* Evidence for a Second Class of S-Adenosylmethionine Riboswitches and Other Regulatory RNA Motifs in alpha-Proteobacteria. *Genome Biol* **6**, (2005).

204. Fuchs, R. T., Grundy, F. J. & Henkin, T. M. The SMK Box is a New SAM-Binding RNA for Translational Regulation of SAM Synthetase. *Nat Struct Mol Biol* **13**, 226–233 (2006).
205. Weinberg, Z. *et al.* Comparative Genomics Reveals 104 Candidate Structured RNAs from Bacteria, Archaea, and Their Metagenomes. *Genome Biol* **11**, R31 (2010).
206. Weinberg, Z. *et al.* The Aptamer Core of SAM-IV Riboswitches Mimics the Ligand-Binding Site of SAM-I Riboswitches. *RNA* **14**, 822–828 (2008).
207. Weinberg, Z. *et al.* Identification of 22 Candidate Structured RNAs in Bacteria Using the CMfinder Comparative Genomics Pipeline. *Nucleic Acids Res* **35**, 4809–4819 (2007).
208. Poiata, E., Meyer, M. M., Ames, T. D. & Breaker, R. R. A Variant Riboswitch Aptamer Class for S-Adenosylmethionine Common in Marine Bacteria. *RNA* **15**, 2046–2056 (2009).
209. Zheng, L. *et al.* Structure-Based Insights Into Recognition and Regulation of SAM-Sensing Riboswitches. *Sci China Life Sci* **66**, 31–50 (2023).
210. Lim, J., Winkler, W. C., Nakamura, S., Scott, V. & Breaker, R. R. Molecular-Recognition Characteristics of SAM-Binding Riboswitches. *Angewandte Chemie* **118**, 978–982 (2006).
211. Montange, R. K. *et al.* Discrimination Between Closely Related Cellular Metabolites by the SAM-I Riboswitch. *J Mol Biol* **396**, 761–772 (2010).
212. Piña, M. D. L. N., Frontera, A. & Bauza, A. Charge Assisted S/Se Chalcogen Bonds in SAM Riboswitches: A Combined PDB and ab Initio Study. *ACS Chem Biol* **16**, 1701–1708 (2021).
213. Grundy, F. J. & Henkin, T. M. The S Box Regulon: A New Global Transcription Termination Control System for Methionine and Cysteine Biosynthesis Genes in Gram-Positive Bacteria. *Mol Microbiol* **30**, 737–749 (1998).
214. Burgold-Voigt, S. *et al.* Characterisation and Molecular Analysis of an Unusual Chimeric Methicillin Resistant Staphylococcus Aureus Strain and its Bacteriophages. *Front Genet* **12**, (2021).
215. Tamura, M. & Holbrook, S. R. Sequence and Structural Conservation in RNA Ribose Zippers. *J Mol Biol* **320**, 455–474 (2002).

216. Adamson, C. S. *et al.* Antiviral Drug Discovery: Preparing for the Next Pandemic. *Chem Soc Rev* **50**, 3647–3655 (2021).
217. Li, G. & De Clercq, E. Overview of Antiviral Drug Discovery and Development: Viral Versus Host Targets. in *Drug Discovery* vol. 80 1–27 (2022).
218. Thomas, S. J. & Yoon, I. K. A Review of Dengvaxia®: Development to Deployment. *Hum Vaccin Immunother* **15**, 2295–2314 (2019).
219. Fatima, K. & Syed, N. I. Dengvaxia Controversy: Impact on Vaccine Hesitancy. *J Glob Health* **8**, (2018).
220. De Clercq, E. Strategies in the Design of Antiviral Drugs. *Nat Rev Drug Discov* **1**, 13–25 (2002).
221. Marzi, M., Vakil, M. K., Bahmanyar, M. & Zarenezhad, E. Paxlovid: Mechanism of Action, Synthesis, and in Silico Study. *Biomed Res Int* **2022**, (2022).
222. Nakamoto, S., Kanda, T., Wu, S., Shirasawa, H. & Yokosuka, O. Hepatitis C Virus NS5A Inhibitors and Drug Resistance Mutations. *World J Gastroenterol* **20**, 2902–2912 (2014).
223. World Health Organization. *Global Hepatitis Report 2024: Action for Access in Low- and Middle-Income Countries*. (2024).
224. Appel, N., Schaller, T., Penin, F. & Bartenschlager, R. From Structure to Function: New Insights Into Hepatitis C Virus RNA Replication. *Journal of Biological Chemistry* **281**, 9833–9836 (2006).
225. Dibrov, S. M., Johnston-Cox, H., Weng, Y. & Hermann, T. Functional Architecture of HCV IRES Domain II Stabilized by Divalent Metal Ions in the Crystal and in Solution. *Angewandte Chemie* **119**, 230–233 (2007).
226. Ji, H., Fraser, C. S., Yu, Y., Leary, J. & Doudna, J. A. Coordinated Assembly of Human Translation Initiation Complexes by the Hepatitis C Virus Internal Ribosome Entry Site RNA. *Proc Natl Acad Sci U. S. A.* **101**, 16990–16995 (2004).
227. Otto, G. A. & Puglisi, J. D. The Pathway of HCV IRES-Mediated Translation Initiation. *Cell* **119**, 369–380 (2004).
228. Jefferson, E. A. *et al.* Biaryl Guanidine Inhibitors of in Vitro HCV-IRES Activity. *Bioorg Med Chem Lett* **14**, 5139–5143 (2004).
229. Wang, W. *et al.* Hepatitis C Viral IRES Inhibition by Phenazine and Phenazine-Like Molecules. *Bioorg Med Chem Lett* **10**, (2000).

230. Carnevali, M., Parsons, J., Wyles, D. L. & Hermann, T. A Modular Approach to Synthetic RNA Binders of the Hepatitis C Virus Internal Ribosome Entry Site. *ChemBioChem* **11**, 1364–1367 (2010).
231. Charrette, B. Design and Synthesis of Ligands Targeting the Hepatitis C Virus Internal Ribosome Entry Site. (University of California, 2016).
232. Davis, D. R. & Seth, P. P. Therapeutic Targeting of HCV Internal Ribosomal Entry Site RNA. *Antivir Chem Chemother* **21**, 117–128 (2011).
233. Payne, S. Virus Interactions with the Cell. in *Viruses* 23–35 (Elsevier, 2017).
234. Martinez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J. & Embarek, A. M. Insights Into Structural and Mechanistic Features of Viral IRES Elements. *Front Microbiol* **8**, (2018).
235. Gilbert, S. D. & Batey, R. T. Monitoring RNA-Ligand Interactions Using Isothermal Titration Calorimetry. in *Riboswitches - Methods and Protocols* 97–114 (2009).
236. Salim, N. N. & Feig, A. L. Isothermal Titration Calorimetry of RNA. *Methods* **47**, 198–205 (2009).
237. Nguyen, B., Tanious, F. A. & Wilson, W. D. Biosensor-Surface Plasmon Resonance: Quantitative Analysis of Small Molecule-Nucleic Acid Interactions. *Methods* **42**, 150–161 (2007).
238. Hendrix, M., Priestley, E. S., Joyce, G. F. & Wong, C.-H. Direct Observation of Aminoglycoside-RNA Interactions by Surface Plasmon Resonance. *J. Am. Chem. Soc* **119**, 3641–3648 (1997).
239. Regulski, E. E. & Breaker, R. R. In-Line Probing Analysis of Riboswitches. in *Post-Transcriptional Gene Regulation* (ed. Wilusz, J.) 53–67 (2008).
240. Wang, B., Wilkinson, K. A. & Weeks, K. M. Complex Ligand-Induced Conformational Changes in tRNAAsp Revealed by Single-Nucleotide Resolution SHAPE Chemistry. *Biochemistry* **47**, 3454–3461 (2008).
241. Yamauchi, T. *et al.* Roles of Mg2+ in TPP-Dependent Riboswitch. *FEBS Lett* **579**, 2583–2588 (2005).
242. Mayer, M. & James, T. L. Detecting Ligand Binding to a Small RNA Target via Saturation Transfer Difference NMR Experiments in D2O and H2O. *J Am Chem Soc* **124**, 13376–13377 (2002).

243. Kulshina, N., Edwards, T. E. & Ferré-D'Amaré, A. R. Thermodynamic Analysis of Ligand Binding and Ligand Binding-Induced Tertiary Structure Formation by the Thiamine Pyrophosphate Riboswitch. *RNA* **16**, 186–196 (2010).
244. Lünse, C. E. & Mayer, G. Reporter Gene-Based Screening for TPP Riboswitch Activators. in *Antibiotics – Methods and Protocols* (ed. Sass, P.) 227–236 (Springer New York, New York, NY, 2017).
245. Shimizu, Y., Kanamori, T. & Ueda, T. Protein Synthesis by Pure Translation Systems. *Methods* **36**, 299–304 (2005).
246. Bains, J. K., Qureshi, N. S., Ceylan, B., Wacker, A. & Schwalbe, H. Cell-free transcription-translation system: a dual read-out assay to characterize riboswitch function. *Nucleic Acids Res* **51**, E82–E82 (2023).
247. Voyvodic, P. L. *et al.* Plug-and-Play Metabolic Transducers Expand the Chemical Detection Space of Cell-Free Biosensors. *Nat Commun* **10**, (2019).
248. Sato, W., Rasmussen, M., Deich, C., Engelhart, A. E. & Adamala, K. P. Expanding Luciferase Reporter Systems for Cell-Free Protein Expression. *Sci Rep* **12**, (2022).
249. Enterina, J. R., Wu, L. & Campbell, R. E. Emerging Fluorescent Protein Technologies. *Curr Opin Chem Biol* **27**, 10–17 (2015).
250. Mishin, A. S., Belousov, V. V., Solntsev, K. M. & Lukyanov, K. A. Novel Uses of Fluorescent Proteins. *Curr Opin Chem Biol* **27**, 1–9 (2015).
251. Li, Y. & Breaker, R. R. Kinetics of RNA Degradation by Specific Base Catalysis of Transesterification Involving the 2γ-Hydroxyl Group. *J Am Chem Soc* **121**, 5364–5372 (1999).
252. Breslow, R. Kinetics and Mechanism in RNA Cleavage. *Proc. Natl. Acad. Sci. USA* **90**, 1208–1211 (1993).
253. AbouHaidar, M. G. & Ivanovb, I. G. Non-Enzymatic RNA Hydrolysis Promoted by the Combined Catalytic Activity of Buffers and Magnesium Ions. *Zeitschrift für Naturforschung C* **54**, 542–548 (1999).
254. Fordyce, S. L., Kampmann, M.-L., Van Doorn, N. L., Thomas, M. & Gilbert, P. Long-Term RNA Persistence in Postmortem Contexts. *Investig Genet* **4**, 7 (2013).
255. Schoenmaker, L. *et al.* mRNA-Lipid Nanoparticle COVID-19 Vaccines: Structure and Stability. *Int J Pharm* **601**, 120586 (2021).

256. Pogocki, D., Scho¨neich, C. & Scho¨neich, S. Chemical Stability of Nucleic Acid-Derived Drugs. *J Pharm Sci* **89**, 443–456 (2000).
257. Zhang, K., Hodge, J., Chatterjee, A., Moon, T. S. & Parker, K. M. Duplex Structure of Double-Stranded RNA Provides Stability Against Hydrolysis Relative to Single-Stranded RNA. *Environ Sci Technol* **55**, 8045–8053 (2021).
258. Cheng, F. *et al.* Research Advances on the Stability of mRNA Vaccines. *Viruses* **15**, (2023).
259. Tjernberg, A., Markova, N., Griffiths, W. J. & Hallén, D. DMSO-related effects in protein characterization. *J Biomol Screen* **11**, 131–137 (2006).
260. Strauss, J. H. & Kelly, R. B. Denaturation of RNA with Dimethyl Sulfoxide. *Biopolymers* **6**, 793–807 (1968).
261. Crawford, J. E., Chan, S. I. & Schweizer, M. P. NMR Studies of Organic Solvent Denatured Yeast Phenylalanyl Transfer RNA at 220 MHz. *Biochem Biophys Res Commun* **44**, 1–7 (1971).
262. Lee, J., Vogt, C. E., McBrairty, M. & Al-Hashimi, H. M. Influence of Dimethylsulfoxide on RNA Structure and Ligand Binding. *Anal Chem* **85**, 9692–9698 (2013).
263. Umuhire Juru, A., Patwardhan, N. N. & Hargrove, A. E. Understanding the Contributions of Conformational Changes, Thermodynamics, and Kinetics of RNA-Small Molecule Interactions. *ACS Chem Biol* **14**, 824–838 (2019).
264. Vo, T., Paul, A., Kumar, A., Boykin, D. W. & Wilson, W. D. Biosensor-Surface Plasmon Resonance: A Strategy to Help Establish a New Generation RNA-Specific Small Molecules. *Methods* **167**, 15–27 (2019).
265. Asmari, M. *et al.* Studying Molecular Interactions via Capillary Electrophoresis and Microscale Thermophoresis: A Review. *Electrophoresis* **44**, 1114–1142 (2023).
266. El Deeb, S. *et al.* Microscale Thermophoresis As a Powerful Growing Analytical Technique for the Investigation of Biomolecular Interaction and the Determination of Binding Parameters. *Methods Appl Fluoresc* **10**, 042001 (2022).
267. Xie, Y., Dix, A. V. & Tor, Y. FRET Enabled Real Time Detection of RNA-Small Molecule Binding. *J Am Chem Soc* **131**, 17605–17614 (2009).

268. Chen, C. Z. *et al.* Two High-Throughput Screening Assays for Aberrant RNA-Protein Interactions in Myotonic Dystrophy Type 1. *Anal Bioanal Chem* **402**, 1889–1898 (2012).
269. Chen, J. L. *et al.* Design, Optimization, and Study of Small Molecules That Target Tau Pre-mRNA and Affect Splicing. *J Am Chem Soc* **142**, 8706–8727 (2020).
270. Shin, Y. H. *et al.* A Novel Time-Resolved Fluorescence Resonance Energy Transfer Assay for the Discovery of Small-Molecule Inhibitors of HIV-1 Tat-Regulated Transcription. *Int J Mol Sci* **24**, (2023).
271. Algar, W. R., Hildebrandt, N., Vogel, S. S. & Medintz, I. L. FRET as a Biomolecular Research Tool — Understanding Its Potential While Avoiding Pitfalls. *Nat Methods* **16**, 815–829 (2019).
272. Burnouf, D. *et al.* KinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *J Am Chem Soc* **134**, 559–565 (2012).
273. Velázquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S. & Freire, E. Isothermal Titration Calorimetry. *Curr Protoc Cell Biol* **23**, 17.8.1-17.8.24 (2004).
274. Hammerschmidt, S. J., Barthels, F., Weldert, A. C. & Kersten, C. Advanced Isothermal Titration Calorimetry for Medicinal Chemists with ITCcalc. *J Chem Educ* **101**, 1086–1095 (2024).
275. Duhr, S. & Braun, D. Why Molecules Move Along a Temperature Gradient. *Proc Natl Acad Sci U.S.A.* **103**, 19678–19682 (2006).
276. Napolitano, V. *et al.* Structure-Based Design, Synthesis and Evaluation of a Novel Family of PEX5-PEX14 Interaction Inhibitors Against Trypanosoma. *Eur J Med Chem* **243**, (2022).
277. Zimmermann, R. A. *et al.* Chemical Space Virtual Screening Against Hard-to-Drug RNA Methyltransferases DNMT2 and NSUN6. *Int J Mol Sci* **24**, (2023).
278. Acquah, F. A. & Mooers, B. H. M. Targeting RNA Structure to Inhibit Editing in Trypanosomes. *Int J Mol Sci* **24**, (2023).
279. Avivi, M. Y. *et al.* Nucleic Acid Hybridization-Based Detection of Pathogenic RNA Using Microscale Thermophoresis. *Journal of Biological Chemistry* **300**, (2024).
280. Garner, A. L. Contemporary Progress and Opportunities in RNA-Targeted Drug Discovery. *ACS Med Chem Lett* **14**, 251–259 (2023).

281. Nowak, J. S. *et al.* Microscale Thermophoresis (MST) and Spectral Shift (SpS) in Drug Discovery. *Trends in Analytical Chemistry* **176**, 117716 (2024).
282. Ludwig, C. Difusion Zwischen Ungleich Erwärmten Orten Gleich Zusammengesetzter Lösungen. *Sitz Math Naturwiss Classe Kaiserlichen Akad Wiss* **20**, 539 (1856).
283. Platten, J. K. The Soret effect: A Review of Recent Experimental Results. *Journal of Applied Mechanics and Transactions* **73**, 5–15 (2006).
284. Rahman, M. A. & Saghir, M. Z. Thermodiffusion or Soret Effect: Historical Review. *Int J Heat Mass Transf* **73**, 693–705 (2014).
285. Kohonen, M. M., Karaman, M. E. & Pashley, R. M. Debye Length in Multivalent Electrolyte Solutions. *Langmuir* **16**, 5749–5753 (2000).
286. Jerabek-Willemsen, M. *et al.* MicroScale Thermophoresis: Interaction Analysis and Beyond. *J Mol Struct* **1077**, 101–113 (2014).
287. Entzian, C. & Schubert, T. Studying Small Molecule-Aptamer Interactions Using MicroScale Thermophoresis (MST). *Methods* **97**, 27–34 (2016).
288. Chramiec-Głąbik, A., Rawski, M., Glatt, S. & Lin, T.-Y. Electrophoretic Mobility Shift Assay (EMSA) and Microscale Thermophoresis (MST) Methods to Measure Interactions Between tRNAs and Their Modifying Enzymes. in *RNA-Protein Complexes and Interactions Methods and Protocols Second Edition* (ed. Lin, R.-J.) vol. 2 29–53 (Springer US, New York, NY, 2023).
289. Seidel, S. A. I. *et al.* Microscale Thermophoresis Quantifies Biomolecular Interactions Under Previously Challenging Conditions. *Methods* **59**, 301–315 (2013).
290. Kolbeck, P. J. *et al.* Molecular Structure, DNA Binding Mode, Photophysical Properties and Recommendations for Use of SYBR Gold. *Nucleic Acids Res* **49**, 5143–5158 (2021).
291. Dumas, P. *et al.* Extending ITC to Kinetics with kinITC. in *Methods in Enzymology* (ed. Feig, A. L.) vol. 567 157–180 (Academic Press Inc., 2016).
292. Dumas, P. *et al. Implementation of KinITC Into AFFINImeter*. (Malvern Instruments Limited, 2015).
293. Feig, A. L. Review: Applications of Isothermal Titration Calorimetry in RNA Biochemistry and Biophysics. *Biopolymers* **87**, 293–301 (2007).

294. Meschel, S. V. A Brief History of Heat Measurements by Calorimetry with Emphasis on the Thermochemistry of Metallic and Metal-Nonmetal Compounds. *Calphad* **68**, (2020).
295. Danielsson, I., Nelander, B., Sunner, S. & Wadsö, I. An Isothermal Jacket Titration Calorimeter. *Acta Chem Scand* **18**, 995–998 (1964).
296. Johnston, H. D. The thermodynamics (log K, [Delta]H°, [Delta]S°, [Delta]Cp°) of Metal Ligand Interaction in Aqueous Solution.|nI.|p Design and Construction of an Isothermal Titration Calorimeter.|nII.|pThe Interaction of Cyanide Ion with Bivalent Nickel, Zinc, Cadmium and Mercury.|nIII.|pThe Interaction of Glycinate Ion with Bivalent Manganese, Iron, Cobalt, Nickel, Copper, Zinc and Cadmium. (Brigham Young University, Provo, 1968).
297. Christensen, J. J., Izatt, R. M. & Hansen, L. D. New Precision Thermometric Titration Calorimeter. *Review of Scientific Instruments* **36**, 779–783 (1965).
298. Christensen, J. J., Johnston, H. D. & Izatt, R. M. An Isothermal Titration Calorimeter. *Review of Scientific Instruments* **39**, 1356–1359 (1968).
299. Muñoz, E. & Piñeiro, A. AFFINImeter Software: From its Beginnings to Future Trends – A Literature Review. *J Appl Bioanal* **4**, 124–139 (2018).
300. Falconer, R. J., Schuur, B. & Mittermaier, A. K. Applications of Isothermal Titration Calorimetry in Pure and Applied Research from 2016 to 2020. *Journal of Molecular Recognition* **34**, (2021).
301. Feig, A. L. Studying RNA-RNA and RNA-Protein Interactions by Isothermal Titration Calorimetry. *Methods Enzymol* **468**, 409–422 (2009).
302. Weickhmann, A. K. *et al.* The Structure of the SAM/SAH-Binding Riboswitch. *Nucleic Acids Res* **47**, 2654–2665 (2019).
303. Su, H. & Xu, Y. Application of ITC-Based Characterization of Thermodynamic and Kinetic Association of Ligands with Proteins in Drug Design. *Front Pharmacol* **9**, 1133 (2018).
304. Medintz, I. & Hildebrandt, N. *FRET - Förster Resonance Energy Transfer from Theory to Applications*. (Wiley-VCH, 2013).
305. Liedberg, B. O., Nylander, C. & Lundstrom, I. Surface Plasmon Resonance for Gas Detection and Biosensing. *Sensors and Actuators* **4**, 299–304 (1983).

306. Arney, J. W. & Weeks, K. M. RNA-Ligand Interactions Quantified by Surface Plasmon Resonance with Reference Subtraction. *Biochemistry* **61**, 1625–1632 (2022).
307. Šípová, H. & Homola, J. Surface Plasmon Resonance Sensing of Nucleic Acids: A Review. *Anal Chim Acta* **773**, 9–23 (2013).
308. Wang, Q. *et al.* Research Advances on Surface Plasmon Resonance Biosensors. *Nanoscale* **14**, 564–591 (2022).
309. Ghayoor, R., Zangenehzadeh, S. & Keshavarz, A. Design of High-Sensitivity Surface Plasmon Resonance Sensor Based on Nanostructured Thin Films for Effective Detection of DNA Hybridization. *Plasmonics* **17**, 1831–1841 (2022).
310. Fabini, E. & Danielson, U. H. Monitoring Drug–Serum Protein Interactions for Aarly ADME Prediction Through Surface Plasmon Resonance Technology. *J Pharm Biomed Anal* **144**, 188–194 (2017).
311. Sun, B., Xu, J., Liu, S. & Li, Q. X. Characterization of Small Molecule–Protein Interactions Using SPR Method. in *Protein–Protein Interactions: Methods and Protocols* (ed. Mukhtar, S.) 149–159 (Springer US, New York, NY, 2023).
312. Douzi, B. Surface Plasmon Resonance: A Sensitive Tool to Study Protein–Protein Interactions. in *Bacterial Secretion Systems: Methods and Protocols* (eds. Journet, L. & Cascales, E.) 363–382 (Springer US, New York, NY, 2023).
313. Madeira, A. *et al.* Identification of Protein-Protein Interactions by Surface Plasmon Resonance Followed by Mass Spectrometry. *Curr Protoc Protein Sci* **1**, (2011).
314. Philip, A. & Kumar, A. R. The Performance Enhancement of Surface Plasmon Resonance Optical Sensors Using Nanomaterials: A Review. *Coord Chem Rev* **458**, 214424 (2022).
315. Yesudasu, V., Pradhan, H. S. & Pandya, R. J. Recent Progress in Surface Plasmon Resonance Based Sensors: A Comprehensive Review. *Heliyon* **7**, e06321 (2021).
316. Sun, A. *et al.* SAM-VI riboswitch structure and signature for ligand discrimination. *Nat Commun* **10**, 5728 (2019).
317. Connelly, C. M. *et al.* Synthetic ligands for PreQ1 riboswitches provide structural and mechanistic insights into targeting RNA tertiary structure. *Nat Commun* **10**, 1501 (2019).

318. Han, G., Tamaki, M. & Hruby, V. J. Fast, efficient and selective deprotection of the tert-butoxycarbonyl (Boc) group using HCl/dioxane (4 M). *The Journal of Peptide Research* **58**, 338–341 (2001).
319. Schwickert, M. *et al.* Discovery of Inhibitors of DNA Methyltransferase 2, an Epitranscriptomic Modulator and Potential Target for Cancer Treatment. *J Med Chem* **65**, 9750–9788 (2022).
320. Stoddard, C. D. *et al.* Free State Conformational Sampling of the SAM-I Riboswitch Aptamer Domain. *Structure* **18**, 787–797 (2010).
321. Lee, Y. M. & Kim, Y. J. Two Cases of Bacteremic Pneumonia Caused by Bifidobacterium Species. *Anaerobe* **73**, (2022).
322. Bush, L. M., Kleper, Þ., De Almeida, N. F., Martin, G. & Perez, M. T. Probiotic-Associated Bifidobacterium Septic Prosthetic Joint Arthritis. *Infectious Diseases in Clinical Practice* **22**, e39–e41 (2014).
323. Yim, G. H., Heon Yang, C., Kim, H. & Chong, Y. Case of Sepsis Caused by Bifidobacterium Longum. *J Clin Microbiol* **37**, 1227–1228 (1999).
324. Schwickert, M. *et al.* Covalent S-Adenosylhomocysteine-Based DNA Methyltransferase 2 Inhibitors with a New Type of Aryl Warhead. *ACS Med Chem Lett* **14**, 777–787 (2023).
325. Sadybekov, A. A. *et al.* Synthon-Based Ligand Discovery in Virtual Libraries of Over 11 Billion Compounds. *Nature* **601**, 452–459 (2022).
326. Müller, J. *et al.* Magnet for the Needle in Haystack: 'Crystal Structure First' Fragment Hits Unlock Active Chemical Matter Using Targeted Exploration of Vast Chemical Spaces. *J Med Chem* **65**, 15663–15678 (2022).
327. Beroza, P. *et al.* Chemical Space Docking Enables Large-Scale Structure-Based Virtual Screening to Discover ROCK1 Kinase Inhibitors. *Nat Commun* **13**, (2022).
328. Lyu, J. *et al.* Ultra-Large Library Docking for Discovering New Chemotypes. *Nature* **566**, 224–229 (2019).
329. Davila-Calderon, J. *et al.* IRES-Targeting Small Molecule Inhibits Enterovirus 71 Replication Via Allosteric Stabilization of a Ternary Complex. *Nat Commun* **11**, (2020).
330. Mousseau, G., Mediouni, S. & Valente, S. T. Targeting HIV Transcription: The Quest For a Functional Cure. *Curr Top Microbiol Immunol* **389**, 121–145 (2015).

331. Bannwarth, S. & Gatignol, A. HIV-1 TAR RNA: The Target of Molecular Interactions Between the Virus and its Host. *Curr HIV Res* **3**, 61–71 (2005).
332. Blond, A., Ennifar, E., Tisné, C. & Micouin, L. The Design of RNA Binders: Targeting the HIV Replication Cycle As a Case Study. *ChemMedChem* **9**, 1982–1996 (2014).
333. Khatkar, P. *et al.* HIV-1 Transcription Inhibition Using Small RNA-Binding Molecules. *Pharmaceuticals* **17**, (2024).
334. Davidson, A., Patora-Komisarska, K., Robinson, J. A. & Varani, G. Essential Structural Requirements for Specific Recognition of HIV TAR RNA by Peptide Mimetics of Tat Protein. *Nucleic Acids Res* **39**, 248–256 (2011).
335. Ferenczy, G. G. & Keserű, G. M. Thermodynamic Profiling for Fragment-Based Lead Discovery and Optimization. *Expert Opin Drug Discov* **15**, 117–129 (2020).
336. Holdgate, G., Fisher, S. & Ward, W. The Application of Isothermal Titration Calorimetry to Drug Discovery. in *Biocalorimetry 2: Applications of Calorimetry in the Biological Sciences* (eds. Ladbury, J. E. & Doyle, M. L.) 59–79 (John Wiley & Sons Ltd, Chrichester, 2004).
337. Holdgate, G. A., Scott, A. & Davies, G. Thermodynamics in Lead Optimization. in *Thermodynamics and Kinetics of Drug Binding* (eds. Keserü, G. M. & Swinney, D. C.) 107–135 (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2015).
338. Freire, E. Isothermal Titration Calorimetry: Controlling Binding Forces in Lead Optimization. *Drug Discov Today Technol* **1**, 295–299 (2004).
339. Das, S., Singh, S., Chawla, V., Chawla, P. A. & Bhatia, R. Surface Plasmon Resonance As a Fascinating Approach in Target-Based Drug Discovery and Development. *TrAC - Trends in Analytical Chemistry* **171**, 117501 (2024).
340. Chavanieu, A. & Pugnière, M. Developments in SPR Fragment Screening. *Expert Opin Drug Discov* **11**, 489–499 (2016).
341. Löfås, S. Technology Review Optimizing the Hit-to-Lead Process Using SPR Analysis. *Assay Drug Dev Technol* **2**, 407–415 (2004).
342. Huang, L., Wang, J. & Lilley, D. M. J. A Critical Base Pair in k-Turns Determines the Conformational Class Adopted, and Correlates With Biological Function. *Nucleic Acids Res* **44**, 5390–5398 (2016).
343. Wildman, S. A. & Crippen, G. M. Prediction of Physicochemical Parameters by Atomic Contributions. *J Chem Inf Comput Sci* **39**, 868–873 (1999).

344. Mattick, J. S. *et al.* Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol* **24**, 430–447 (2023).
345. Zucchelli, S. *et al.* SINEUPs: A New Class of Natural and Synthetic Antisense Long Non-Coding RNAs That Activate Translation. *RNA Biol* **12**, 771–779 (2015).
346. Carrieri, C. *et al.* Long Non-Coding Antisense RNA Controls Uchl1 Translation Through an Embedded SINEB2 Repeat. *Nature* **491**, 454–457 (2012).
347. Morrissy, A. S., Griffith, M. & Marra, M. A. Extensive Relationship Between Antisense Transcription and Alternative Splicing in the Human Genome. *Genome Res* **21**, 1203–1212 (2011).
348. Romero-Barrios, N., Legascue, M. F., Benhamed, M., Ariel, F. & Crespi, M. Survey and Summary – Splicing Regulation by Long Noncoding RNAs. *Nucleic Acids Res* **46**, 2169–2184 (2018).
349. Pisignano, G. & Ladomery, M. Epigenetic regulation of Alternative Splicing: How lncRNAs Tailor the Message. *Noncoding RNA* **7**, 1–19 (2021).
350. Nandwani, A., Rathore, S. & Datta, M. LncRNAs in Cancer: Regulatory and Therapeutic Implications. *Cancer Lett* **501**, 162–171 (2021).
351. Abulwerdi, F. A. *et al.* Selective Small-Molecule Targeting of a Triple Helix Encoded by the Long Noncoding RNA, MALAT1. *ACS Chem Biol* **14**, 223–235 (2019).
352. Li, K. & Wang, Z. lncRNA NEAT1: Key Player in Neurodegenerative Diseases. *Ageing Res Rev* **86**, 101878 (2023).
353. Li, D. *et al.* Insights Into lncRNAs in Alzheimer's Disease Mechanisms. *RNA Biol* **18**, 1037–1047 (2020).
354. Guo, F., Guo, L., Li, Y., Zhou, Q. & Li, Z. MALAT1 is an Oncogenic Long Non-Coding RNA Associated With Tumor Invasion in Non-Small Cell Lung Cancer Regulated by DNA Methylation. *Int J Clin Exp Pathol* **8**, 15903–15910 (2015).
355. Doxtater, K., Tripathi, M. & Khan, M. Recent Advances on the Role of Long Non-Coding RNAs in Alzheimer's Disease. *Neural Regen Res* **15**, 2253–2254 (2020).
356. Zhang, W., Zhao, H., Wu, Q., Xu, W. & Xia, M. Knockdown of BACE1-AS by siRNA Improves Memory and Learning Behaviors in Alzheimer's Disease Animal Model. *Exp Ther Med* **16**, 2080–2086 (2018).

357. Pascual-Gilabert, M., Artero, R. & López-Castel, A. The Myotonic Dystrophy Type 1 Drug Development Pipeline: 2022 Edition. *Drug Discov Today* **28**, 1–11 (2023).
358. Ciesiolka, A., Jazurek, M., Drazkowska, K. & Krzyzosiak, W. J. Structural Characteristics of Simple RNA Repeats Associated With Disease and Their Deleterious Protein Interactions. *Front Cell Neurosci* **11**, 97 (2017).
359. Wojciechowska, M. & Krzyzosiak, W. J. Cellular Toxicity of Expanded RNA Repeats: Focus on RNA Foci. *Hum Mol Genet* **20**, 3811–3821 (2011).
360. Rzuczek, S. G. *et al.* Precise Small-Molecule Recognition of a Toxic CUG RNA Repeat Expansion. *Nat Chem Biol* **13**, 188–193 (2017).
361. Gibaut, Q. M. R. *et al.* Transcriptome-Wide Studies of RNA-Targeted Small Molecules Provide a Simple and Selective r(CUG)exp Degrader in Myotonic Dystrophy. *ACS Cent Sci* **9**, 1342–1353 (2023).
362. Gibaut, Q. M. R. *et al.* Study of an RNA-Focused DNA-Encoded Library Informs Design of a Degrader of a r(CUG) Repeat Expansion. *J Am Chem Soc* **144**, 21972–21979 (2022).
363. Haniff, H. S. *et al.* Targeting the SARS-COV-2 RNA Genome With Small Molecule Binders and Ribonuclease Targeting Chimera (RiboTAC) Degraders. *ACS Cent Sci* **6**, 1713–1721 (2020).
364. Kargbo, R. B. RIBOTACs: Small Molecules Selectively Destroy Cancer-Associated RNA. *ACS Med Chem Lett* **12**, 1872–1873 (2021).
365. Costales, M. G., Matsumoto, Y., Velagapudi, S. P. & Disney, M. D. Small Molecule Targeted Recruitment of a Nuclease to RNA. *J Am Chem Soc* **140**, 6741–6744 (2018).
366. Bush, J. A. *et al.* Ribonuclease Recruitment Using a Small Molecule Reduced c9ALS/FTD r(G4C2) Repeat Expansion in Vitro and in Vivo ALS Models. *Sci. Transl. Med* **13**, eabd5991 (2021).
367. Tong, Y. *et al.* Decreasing the Intrinsically Disordered Protein a-Synuclein Levels by Targeting its Structured mRNA With a Ribonuclease-Targeting Chimera. *Proc Natl Acad Sci U S A* **121**, (2024).
368. Meyer, S. M. *et al.* DNA-Encoded Library Screening To Inform Design of a Ribonuclease Targeting Chimera (RiboTAC). *J Am Chem Soc* **144**, 21096–21102 (2022).