

Harnessing Computational Approaches for RNA-Targeted Drug Discovery

Yuanzhe Zhou¹, Shi-Jie Chen^{1,2,*}

¹Department of Physics and Astronomy, University of Missouri, Columbia, MO 65211, USA; ²Department of Biochemistry, MU Institute for Data Science and Informatics, University of Missouri, Columbia, MO 65211, USA

ABSTRACT

RNA molecules have emerged as promising therapeutic targets due to their diverse functional and regulatory roles within cells. Computational modeling in RNA-targeted drug discovery presents a significant opportunity to expedite the discovery of novel small molecule compounds. However, this field encounters unique challenges compared to protein-targeted drug design, primarily due to limited experimental data availability and current models' inability to adequately address RNA's conformational flexibility during ligand recognition. Despite these challenges, several studies have successfully identified active RNA-targeting compounds using structure-based approaches or quantitative structure-activity relationship (QSAR) models. This review offers an overview of recent advancements in modeling RNA-small molecule interactions, emphasizing practical applications of computational methods in RNA-targeted drug discovery. Additionally, we survey existing databases that catalog nucleic acid-small molecule interactions. As interest in RNA-small molecule interactions grows and curated databases expand, the field anticipates rapid development. Novel computational models are poised to enhance the identification of potent and selective small-molecule modulators for therapeutic needs.

Keywords: RNA-ligand interaction, RNA-targeted drug discovery, virtual screening, molecular docking, QSAR

ABBREVIATIONS: ncRNAs, non-coding RNAs; miRNAs, microRNAs; lncRNAs, long non-coding RNAs; circRNAs, circular RNAs; rRNAs, ribosomal RNAs; SMN2, survival of motor neuron 2; SMA, spinal muscular atrophy; FSE, frameshifting stimulation element; SARS-CoV, severe acute respiratory syndrome coronavirus; ENE, element for nuclear expression; KSHV, Kaposi's sarcoma-associated herpesvirus; PAN, polyadenylated nuclear; HIV-1, human immunodeficiency virus type 1; TAR, trans-activation response; HBV, hepatitis B virus; pgRNA, pre-genomic RNA; QSAR, quantitative structure-activity relationship; ML, machine learning; VS, virtual screening; DL, deep learning; NA, nucleic acid; SM, small molecule; RS, riboswitch; G4, G-quadruplex; iM, i-Motif; SELEX, systematic evolution of ligands by exponential enrichment; ALIS, automated ligand identification system; SMM, small molecule microarray; EBVS, ensemble-based virtual screening; RDC, residual dipolar coupling; RF, ribosomal frameshifting; UTR, untranslated region; mRNA, messenger RNA; AUROC, area under the ROC curve; AIDS, acquired immunodeficiency syndrome; HCV, hepatitis C virus; IRE, iron responsive element; LASSO, least absolute shrinkage and selection operator; MLR, multiple linear regression; CNN, convolutional neural network; PTC, peptidyl transferase center

*Please send correspondence to Shi-Jie Chen at chenshi@missouri.edu.

Copyright: © 2024 Yuanzhe Zhou *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Received July 8, 2024; **Accepted** December 21, 2024

INTRODUCTION

Recent years have seen growing interest in RNA-targeted drug discovery due to RNA's diverse functional and regulatory roles in gene expression, protein synthesis, and viral replication¹⁻¹². Identifying small molecules that bind to specific RNA targets could lead to novel therapeutic agents for treating diseases associated with dysregulated RNA functions. Compared to the protein-targeted approaches, targeting RNA significantly broadens the druggable genome and may provide more effective disease treatments due to RNA's critical roles in gene expression and regulation^{2, 3, 6, 13, 14}. It is estimated that ~85% of the human genome is transcribed into RNAs^{2, 15}, while only approximately 1.5% encodes proteins^{3, 6, 13, 14, 16-20}. The vast majority of the transcribed RNAs are non-coding RNAs (ncRNAs), which do not encode proteins but rather serve crucial enzymatic, structural, and regulatory functions in gene expression. Recent studies have shown that many ncRNAs, including microRNAs (miRNAs), long non-coding RNAs (lncRNAs), and circular RNAs (circRNAs), have been found to be associated with human diseases such as cancer and neurodegenerative disorders, making them potential new drug targets²¹⁻²⁶. The discovery of ribocil and its derivatives²⁷⁻³⁰ - potent synthetic ligands that bind to FMN riboswitches with chemically distinct scaffold compared to the natural ligand - demonstrates the possibility of using riboswitches as the antibacterial drug targets³¹, alternative to the traditional ribosomal RNAs (rRNAs) targets^{32, 33}, to combat bacterial infections. Other RNA targets, such as survival of motor neuron 2 (SMN2) pre-mRNA in spinal muscular atrophy (SMA)³⁴, ribosomal frameshifting stimulation element (FSE) within the severe acute respiratory syndrome coronavirus 1/2 (SARS-CoV-1/2) RNA genome³⁵⁻³⁹, element for nuclear expression (ENE) triple helix presented in Kaposi's sarcoma-associated herpesvirus (KSHV) polyadenylated nuclear (PAN) lncRNA [40], human immunodeficiency virus type 1 (HIV-1) trans-activation response (TAR) element^{41, 42}, a flexible "priming loop" region of the *cis*-acting regulatory element (ϵ) located at the 5'-ends of the hepatitis B virus (HBV) pre-genomic RNA (pgRNA)^{43, 44}, have also been shown to be targetable by small molecule compounds. Notably, FDA's approval of an RNA splicing modulator, Evrysdi (risdiplam) in August 2020 for treating SMA represents the first small molecule drug designed to directly target non-ribosomal RNA, paving the way for further advancements in RNA-targeted drug discovery^{11, 45-47}.

Despite the growing interest in RNA-targeted drug discovery, computational modeling of RNA-small

molecule interactions is still in its infancy. The *in silico* process of identifying lead compounds for specific RNA targets varies based on the availability of RNA structural information and computational resources, ranging from quantitative structure-activity relationship (QSAR)-based approaches to structure-based approaches. QSAR-based approaches⁴⁸⁻⁵¹ build simple analytic models to establish the relationships between a set of predefined molecular descriptors (e.g., physicochemical properties such as molecular weight, polarizability, surface area, number of H-bond donors/acceptors, and 3D shape) and binding-related quantities for experimentally identified bound ligands. These QSAR models can then be used to quantify RNA-small molecule associations for given small molecule compound libraries. In contrast, structure-based approaches⁵²⁻⁶⁵, such as molecular docking, leverage prior knowledge of RNA at various structural levels to gain detailed insights into drug-target interaction mechanisms. These methods provide a more comprehensive understanding of the binding process but typically require more extensive computational resources and structural data. See Fig. 1 for examples of QSAR and molecular docking workflows.

However, RNA flexibility and data scarcity pose significant challenges in RNA-targeted drug discovery. Unlike proteins, RNA molecules are highly dynamic, adopting multiple, rapidly inter-converting conformations^{1, 14, 66-68}. This inherent flexibility complicates the design of small molecules that can reliably bind to specific RNA structures. Recent comparative studies have shown that current docking software struggles to model the conformational changes of RNA receptors induced by ligand binding^{69, 70}. Accurate modeling of RNA-ligand interactions requires knowledge of RNA conformational ensemble. Furthermore, the limited availability of high-quality structural data on RNA-small molecule complexes further complicates the drug design process, especially for data-driven approaches involving machine learning (ML)^{71, 72}. In the following sections, we will focus on available database resources and practical applications of computational tools that facilitate the early stages of RNA-targeted small molecule drug discovery. We will not delve into the detailed methodologies employed by various computational models, as these have been covered in comprehensive reviews elsewhere^{66, 71, 73, 74}.

DATABASES WITH KNOWN RNA-LIGAND INTERACTIONS

RNA-targeted small molecule drug discovery is still

in its early stage. The development of robust predictive models for RNA-small molecule interactions is currently hindered by the limited availability of high-quality, experimentally validated data. While several databases have emerged in recent years, many lack substantial amount of 3D structural information for RNA-small molecule systems. Data-driven models, particularly deep learning (DL) models — a subset of ML models that leverages deep neural networks to perform complex tasks — often comprise a significant number of trainable parameters. These models require sufficient high-quality training data to mitigate common issues such as overfitting. A notable example is AlphaFold 3⁷⁵, which can model interactions between diverse biomolecules, including RNA-small molecule interactions. To further improve current DL models, there is a pressing demand for carefully curated benchmark datasets that encompass detailed 3D structural information and experimentally measured properties, such as binding affinity. For example, in the most recent version of PDBbind (version 2021), there are 22,920 protein-ligand complexes with experimentally measured binding affinity data compared to only 171 nucleic acid-ligand complexes with affinity data⁷⁶. By offering a standardized training and testing dataset, a well-curated database of RNA-small molecule interactions would be an invaluable resource for the scientific community, providing a solid foundation for advancing various computational approaches to understanding RNA biology and RNA-targeted therapeutics. In the following, we briefly overview currently available databases containing curated nucleic acid-small molecule interactions published after 2016. See Table 1 for a list of reviewed databases.

Databases without 2D/3D structural information of nucleic acid

The Nucleic Acid Ligand Database (NALDB)⁷⁷ is a database providing detailed experimental information on 3,610 ligands that target several types of nucleic acid structures. The information includes target sequence, ligand structures and chemical properties, along with their K_i , K_D , and IC_{50} values, etc. The ligands are broadly categorized into six classes based on their corresponding binding nucleic acid partners: double-stranded DNA/RNA, G-quadruplex DNA/RNA, DNA/RNA aptamers, and special DNA/RNA structures. The NALDB web portal offers search functionality that enables users to filter the database based on drug-like properties or similarities to user-uploaded query molecules.

NoncoRNA⁸⁰ is a manually curated database providing experimental information on interactions between various ncRNAs and drugs. It contains 8,233 entries involving 5,568 ncRNAs and 154 drugs across 134 cancer types. The database offers a web interface for browsing, searching, and downloading the data. Each entry in NoncoRNA contains detailed information on the ncRNAs, drugs, and related cancers, including ncRNA expression patterns, experimental detection techniques, drug responses, and literature references, etc.

To investigate the druggability of RNAs with small molecules, Rizvi *et al.*⁵ employed the automated ligand identification system (ALIS) affinity-selection mass spectrometry platform to screen 42 RNAs across diverse RNA classes against approximately 60,000 compounds. This screening generated millions of target-compound interaction data points. From the identified RNA binders,

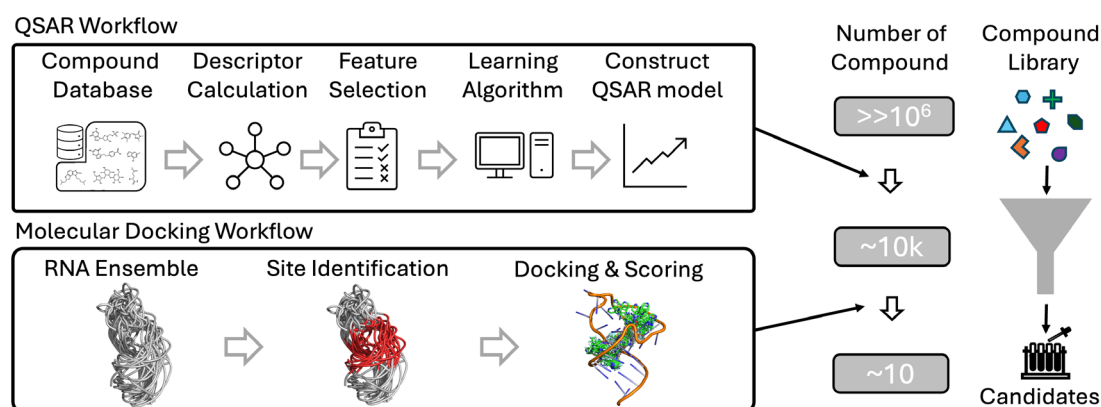


Fig. 1. Examples of QSAR and molecular docking workflows for RNA-targeted virtual screening (VS). The illustrated VS combines QSAR and molecular docking in a consecutive manner, effectively reducing computational costs while retaining similar screening power.

the researchers constructed an RNA-focused library of roughly 3,700 small molecules with enhanced RNA-targeting capabilities. Several RNA targets in the database can be accessed by modifying existing experimentally solved 3D structures in the RCSB PDB database⁸⁹. Unfortunately, the authors did not provide public access to the detailed screening results.

Recently, Yazdani *et al.*⁵⁰ reported the Repository of BInders to Nucleic acids (ROBIN), which contains small molecule microarray (SMM) screening results for 36

individual nucleic acids (27 RNA and 9 DNA) against a library of 24,572 small molecules. The database identified 2,003 RNA-binding small molecules targeting various RNA motifs, including RNA hairpins, DNA/RNA G-quadruplexes, RNA triple helices, RNA pseudoknots, and RNA three-way junctions. Notably, ROBIN is currently the largest public database for nucleic acid-small molecule interactions. Although lacking RNA 3D structure information, many RNA targets in ROBIN can be obtained by modifying related experimentally solved

Table 1. Databases containing RNA-small molecule interactions published after 2016

Name	Year	NA Type	# of NA-SM Interactions	Focus and Strength	Web Portal	Ref
NALDB	2016	RNA/DNA	3,610	SM-NA association (contains K_i , K_D , IC_{50})	Yes	77
Inforna 2.0	2016	RNA	1,936	SM-RNA association (contains RNA 2D info)	Yes ^c	78
RiboD	2019	RS	1,907 ^a	information of RS and gene/operon it regulates	Yes	79
NoncoRNA	2020	ncRNA	8,233	ncRNAs and drug associations	Yes	80
Rizvi <i>et al.</i>	2020	RNA	~2,300,000 ^b	SM-RNA association (large-scale screening)	No	5
G4LDB 2.2	2021	G4/iM	28,500	SM-G4/iM association	Yes	81
RSwitch	2021	RS	215 ^a	information of pathogenic bacterial RS	Yes	82
PDBbind v.2021	2021	RNA/DNA	171	SM-NA association (contains 3D structures and binding properties)	Yes	76
RNALigands	2022	RNA	841	SM-RNA association (contains RNA 2D info)	Yes	83
R-BIND 2.0	2022	RNA	1,494	SM-RNA association (contains RNA 2D info)	Yes	84
HARIBOSS	2022	RNA	1,471	SM-RNA association (contains 3D structures)	Yes	85
R-SIM	2023	RNA	2,501	SM-RNA binding affinities	Yes	86
ROBIN	2023	RNA/DNA	1,627,072 ^b	SM-NA association (large-scale screening; publicly available)	No ^d	50
Ribocentre-switch	2023	RS	89,591 ^a	information of RS and related ligand binding	Yes	87
AptaDB	2024	RNA/DNA	393	information of NA aptamers from SELEX-based screening	Yes	88

^aNumber of riboswitch sequences. ^bDistinct from other studies, the full results of probed binding interactions were reported, including binders and non-binders. ^cAvailable upon request. ^dNo official website but the dataset is downloadable through the publisher's website. NA, nucleic acid; SM, small molecule; RS, riboswitch; ncRNA, non-coding RNA; G4, G-quadruplex; iM, i-Motif; SELEX, systematic evolution of ligands by exponential enrichment.

3D structures available in the RCSB PDB database⁸⁹, making it a valuable benchmark dataset for developing effective nucleic acid-specific virtual screening tools.

Databases with 2D/3D structural information of nucleic acid

Inforna 2.0⁷⁸ is a template database that contains 1,936 pairs of known RNA secondary structure motif-ligand bound complexes, comprising 244 unique small molecules and 1,331 RNA motifs. The data was collected from an in-house two-dimensional combinatorial screen against a panel of RNA secondary structures, including hair-pins, symmetric and asymmetric internal loops, and bulges, etc. The database can identify RNA structural motifs similar to the user-provided RNA secondary structures and perform similarity searches in large compound libraries using the native bound ligands as lead compounds.

RNA Ligands⁸³ is a database focused on collecting RNA secondary structural motifs and their bound small molecule ligands. It contains 841 RNA-ligand pairs gathered from three sources: RCSB PDB database (386 pairs)⁸⁹, R-BIND (67 pairs)⁹⁰, and Inforna 2.0 (388 pairs)⁷⁸. By integrating a motif alignment algorithm, RNA Ligands can identify secondary structure motifs from user input that are similar to those stored in the database. The matched RNA motifs and their bound small molecule ligands can provide valuable information for identifying lead compounds.

RNA-targeted Bioactive ligand Database (R-BIND) 2.0⁸⁴ is the biennial update of the original R-BIND database⁹⁰, designed to facilitate the discovery of RNA-targeted chemical probes. It contains comprehensive information on reported chemical probes that target non-rRNA and exhibit biological activity. The database includes 1,494 recorded RNA-small molecule interactions for 153 unique small molecules, where experimental details for *in vitro* assays, cell-based assays, and/or animal model studies conducted after the primary screen for each ligand can be found. The web portal of the database also offers search functionality based on RNA structure. The algorithm uses an input RNA sequence (or structure of interest), conducts RNA structure prediction, and then returns ligands in the database that bind that motif (if any) based on RNA secondary structure and size.

Several other databases also contain 3D structural information for nucleic acid-small molecule complexes. PDBbind v. 2021⁷⁶ database contains 171 nucleic acid-ligand complexes with solved structures and measured binding affinities. RNA-Small molecule Interaction Miner (R-SIM)⁸⁶ database includes 2,501 experimentally

validated RNA-small molecule interactions, covering 461 unique RNA targets and 1,288 unique small molecules. HARIBOSS (Harnessing RIBOnucleic acid - Small molecules Structures)⁸⁵ database comprises 311 compounds, 862 RNA-small molecule complexes, and 1471 pockets collected from RCSB PDB database⁸⁹.

Databases focusing on small molecules targeting specific nucleic acid types

Currently, several databases have been curated for small molecules interacting with specific nucleic acid types, including riboswitches^{79, 82, 87}, nuclear acid aptamers⁸⁸, and nucleic acids with special topologies (e.g., G-quadruplex and i-Motif⁸¹). Three databases focus on riboswitches are RiboD⁷⁹, RSwitch⁸², and Ribocentre-switch⁸⁷. Among these, Ribocentre-switch is the most recent and comprehensive database encompassing 56 riboswitches and 26 orphan riboswitches from over 430 references, with a total of 89,591 sequences. It provides information on target sequences, structures, functions, ligand binding pockets, and biological applications. AptADB⁸⁸ is an aptamer database derived from experimentally validated aptamer-target interaction data collected from the literature. It includes information such as aptamer structure and experimental binding affinities. Currently, AptADB contains 1,350 experimentally validated aptamer-target interactions, of which 393 involve small molecules. G4LDB 2.2⁸¹ focuses on nucleic acid structures with special topologies, such as G-quadruplex and i-Motif. It currently contains more than 3,200 G-quadruplex/i-Motif ligands, with ~28,500 activity entries.

COMPUTATIONAL MODELING OF RNA-SMALL MOLECULE INTERACTIONS AND THEIR APPLICATIONS IN RNA-TARGETED DRUG DISCOVERY

Over the years, researchers have developed various computational models to streamline the early stages of RNA-targeted drug development. These models aim to reduce the time and cost associated with experimental testing by facilitating critical processes such as binding site detection, binding mode prediction, and ultimately, lead identification and optimization^{66, 71–74}. See Table 2 for a list of applications reviewed here for RNA-targeted drug discovery.

Structure-based virtual screening identifies RNA-targeting small molecules

Structure-based approaches, such as RNA-ligand

Table 2. Applications of RNA-targeted drug discovery

RNA Target	Year	Computational Method	Ref
SARS-CoV-1 RF	2011	Docking	[37]
Purine RS	2011	Docking	[92]
HIV-1 TAR	2011/2018	Docking	[41, 42]
PreQ ₁ RS	2022	Docking	[93]
HBV ϵ	2023	Docking	[44]
SARS-CoV-2 RF	2024	Docking	[100]
Inforna	2017/2020	Template-based scoring	[115, 117, 119]
RNA hairpin ^a	2022	QSAR	[51]

^aHairpin 91 within ribosomal peptidyl transferase center (PTC) of *Mycobacterium tuberculosis*. RF, ribosomal frameshifting; RS, riboswitch; TAR, trans-activation response.

molecular docking, provide detailed insights into binding modes and the structural basis of RNA-small molecule recognition, crucial for designing potent and selective drugs. Unlike protein-ligand docking, software specifically designed for RNA-ligand interactions is limited⁶⁶. Available tools include all-in-one docking software (AutoDock⁵², DOCK 6⁵³, rDock⁵⁴, RLDOCK⁵⁵,⁵⁶, NLDOCK⁵⁷) and standalone scoring functions ranging from physics-based to knowledge-based and machine-learning models (DrugScoreRNA⁵⁸, LigandRNA⁵⁹, SPA-LN⁶⁰, ITScore-NL⁶¹, AnnapuRNA⁶², RNAPosers⁶³, RNAmigos⁶⁵, and RNAPosers-ssMD⁶⁴). While significant progress has been made in predicting the binding of a flexible ligand and a rigid *holo*-form RNA structure with limited side chain flexibility⁶⁶, capturing the ligand-modulated RNA conformational dynamics remains challenging. Small molecule recognition is sensitive to local binding pocket conformations, and even minor deviations from the *holo*-conformation could render current docking methods ineffective. This issue is far more pronounced for RNA than protein, as RNA ligand binding site conformations are more flexible^{1, 14, 91}. Most current molecular docking studies employ an ensemble-based virtual screening (EBVS) workflow to address this issue. In this approach, a candidate compound library is docked into an ensemble of RNA structures rather than a single rigid conformation. This structural ensemble can be obtained through either computational method, such as MD simulation, or experiments data, such as those from NMR residual dipolar coupling (RDC). However, the EBVS approach is typically more computationally expensive and may not provide significant improvements

over single-conformation docking without experimentally informed structural information. Consequently, only a limited number of successful molecular docking-based virtual screenings against RNA targets have been reported to date^{35–37, 41–44, 92, 93}.

SARS-CoV-1 and SARS-CoV-2 are two coronavirus strains responsible for the SARS outbreak and COVID-19 outbreak, respectively⁹⁴. These viruses utilize ribosomal frameshifting (RF), also known as translational frameshifting, a conserved mechanism that allows multiple proteins to be produced from a single mRNA sequence⁹⁵. During translation, RF occurs when the ribosome shifts its reading frame, typically by one or two nucleotides, resulting in the synthesis of a protein different from what the standard reading frame would produce. This process enables efficient use of genetic material and is crucial for viral replication and gene expression regulation. Studies have shown that the RNA pseudoknot at the -1 RF site effectively stimulates -1 RF in SARS-CoV-1^{35, 36}. Disrupting -1 RF by targeting this pseudoknot with a small molecule could potentially inhibit viral infectivity and production. Park *et al.*³⁷ explored this possibility by using DOCK 4.0⁹⁶ to screen 80,000 compounds against a computationally predicted and optimized SARS-pseudoknot structure. The structure was initially predicted using PSEUDOVIEWER^{97, 98} and the SYBYL 6.9 system (Tripos, Inc.), then optimized with AMBER 8⁹⁹. Based on DOCK scoring function and visual inspection, 58 highly ranked compounds were selected for further evaluation. *In vitro* and *cell-based* -1 RF assays on these compounds led to the identification of a novel ligand capable of dramatically inhibiting SARS-CoV-

1's -1 RF. Recently, Mathez *et al.*¹⁰⁰ employed a similar molecular docking-based approach using the Schrodinger Suite version 2023-2¹⁰¹ to virtually screen the ROBIN dataset⁵⁰, where 2,003 RNA-targeting compounds were docked against an ensemble of MD-refined crystal structures of the pseudoknot involved in the -1 RF. After performing high-throughput, standard precision (SP), and extra precision (XP) docking, the final poses of each compound were refined using Prime MM-GBSA. Twenty-one molecules binding to at least 3 clusters were visually inspected and experimentally validated against the wild-type virus. Several compounds demonstrated antiviral activity with varying EC₅₀ values, and one compound induced a significant reduction in frameshift efficiency. Further investigation into analogues of this compound revealed one analogue with more potent antiviral activity and an improved ability to reduce frameshift efficiency.

Riboswitches are naturally occurring RNA segments located in the untranslated region (UTR) of messenger RNA (mRNA) molecules. They contain a ligand-binding aptamer domain where specific ligands can bind and induce alternative folding of the downstream expression platform, regulating transcription, translation, or other gene expression process^{102, 103}. With over 55 distinct classes of natural riboswitches discovered that selectively sense small molecules or elemental ions, and many more predicted to exist^{102, 103}, riboswitches present an alternative target for antibacterial drug development. This approach could help combat emerging multidrug resistant bacteria, complementing traditional rRNAs targeting strategy³¹⁻³³. Daldrop *et al.*⁹² utilized DOCK software to screen 2,615 unique compounds (including eight experimentally verified hits) against the adenine riboswitch (AR) and a related guanine riboswitch carrying a point mutation (GRA). Using a force-field-based scoring function modified with RNA-specific parameters, Daldrop *et al.* achieved an area under the receiver operation characteristic curve (AUROC) of 0.98. Further molecular docking and experimental investigation of top-ranked commercially available compounds identified four hits that bind to GRA, with two exhibiting novel scaffolds compared to the known ligands [92]. More recently, Kallert *et al.*⁹³ investigated the capabilities of protein-based docking programs to reproduce native binding modes and discriminate known binders from decoys. Kallert *et al.* conducted structure-based virtual screening by docking 12,507 compounds into the *Bacillus subtilis* preQ₁ riboswitch aptamer domain using protein-ligand docking software FRED¹⁰⁴, HYBRID¹⁰⁵, and FlexX¹⁰⁶. Subsequent microscale thermophoresis assays identified

six active compounds out of 23 tested VS hits, with K_D values ranging from 29.5 nM to 11.0 μM. Notably, the VS results indicated that combining docking score-based ranking and pharmacophore hypothesis-based binding mode interpretation (i.e., human expert intervention) is superior to the docking score-based ranking alone, providing guidance for hit selection.

HIV, a subgroup of retrovirus, can cause acquired immunodeficiency syndrome (AIDS) after infection. Of the two types of HIV (i.e., HIV-1/2), HIV-1 is the one that was first discovered and is the more dominant and pathogenic strain¹⁰⁷. Studies have shown that the trans-activation response (TAR) element of HIV-1, a well folded stem-bulge-loop RNA in the 5' UTR of the viral genome, is essential for HIV replication. It interacts with the Tat protein, helping recruit host transcriptional machinery to facilitate efficient viral DNA transcription¹⁰⁸. This central role underscores TAR's potential as a therapeutic target. To account for TAR's large conformational flexibility upon ligand binding, Stelzer *et al.*⁴¹ adopted an EBVS approach. Stelzer *et al.* used structural data, such as NMR residual dipolar coupling, to construct an RNA dynamic ensemble for molecular docking studies. Using the Internal Coordinate Mechanics (ICM, Molsoft¹⁰⁹) docking program, They screened ~51,000 small molecules against each of the 20 conformers in the TAR ensemble. Fluorescence-based assays of the top 57 commercially available hits experimentally validated six active small molecules that bind TAR and inhibit its interaction with Tat, with K_D and K_i values ranging from 55 nM to 122 μM and from 710 nM to 169 μM, respectively. Ganser *et al.*⁴² employed a similar approach to validate ability of EBVS to discriminate between known nonhits and hits targeting TAR. Ganser *et al.* docked ~100,000 drug-like molecules, including 177 experimentally verified hits, against the RDC-informed TAR conformational ensemble¹¹⁰ using ICM¹⁰⁹. EBVS enriched the compound library with an AUROC of 0.88 and identified 42% of hits after screening only 2% of the compounds, which corresponds to an enrichment factor of 21. By varying ensemble size and using TAR ensemble generated computationally without considering experimental structural information, the authors observed a significant decrease in enrichment performance demonstrating the importance of an accurate, large RNA conformational ensemble.

HBV infection causes liver disease and accounts for over 600,000 deaths annually¹¹¹. Recent experimental studies by LeBlanc *et al.*⁴³ identified several small molecules, including Raloxifene, Bazedoxifene, and a *de novo* derivative, that selectively target a flexible "priming

loop” region of the *cis*-acting regulatory element (ϵ) located at the 5' ends of the HBV pre-genomic RNA (pgRNA). To further investigate the potential of HBV ϵ as a promising drug target to complement current anti-HBV therapies⁴³, Oleginski *et al.*⁴⁴ carried out virtual screening against HBV ϵ . Oleginski *et al.* docked 1,604 FDA-approved compounds into NMR-solved RNA ensemble structures with AutoDock Vina¹¹². Using the predicted affinity of the known ϵ -targeting ligand Raloxifene as a cutoff, 66 top-ranked compounds were selected based on commercial availability and potential adverse effects. Among these, three anti-Hepatitis C virus (HCV) compounds (Simeprevir, Ledipasvir, and Daclatasvir) were found to bind with full-length HBV ϵ , showing approximate EC₅₀ values of 298, 145, and 62 μ M, respectively. Further dye-displacement assays with additional RNAs containing structural elements similar to HBV ϵ validated Daclatasvir as a selective ϵ -targeting compound⁴⁴. This work supports the possibility that targeting ϵ dynamics may be an effective anti-HBV therapeutic strategy.

While molecular docking is widely used in structure-based virtual screening, approaches solely relying on 2D structural information have also found successfully applications in RNA-targeted drug discovery. One such model is Inforna¹¹³, a template-based method capable of identifying hit compounds for given RNA target based on their similarities to the known compound-interacting RNA secondary structure motifs such as hairpins, symmetric and asymmetric internal loops, and bulges. The latest iteration, Inforna 2.0⁷⁸, contains 1,936 pairs of known RNA secondary structure motif-ligand complexes. For a given RNA target, Inforna first identifies its secondary structure motifs and matches them against the template database. For any matching motif, it then finds the corresponding ligand partners and calculates fitness scores. These scores provide a measure of RNA-ligand binding affinity as well as the selectivity of the RNA motif against other small molecules. RNAs with high selectivity and affinity fitness scores are more likely to be druggable. Using the top-scored small molecules as lead compounds, chemical similarity screening against compound libraries can identify potential potent binders. Compared to traditional molecular docking-based approaches, Inforna and Inforna 2.0 take advantage of known interactions between small molecules and RNA structural motifs. The similarity-based approach can provide a massive speedup when performing large-scale compound screening. Additionally, Inforna and Inforna 2.0 do not require 3D structural information of the target, which makes them suitable for tasks where only 2D structural information

is available. In practice, Inforna and Inforna 2.0 can complement molecular docking, where one can first use Inforna and Inforna 2.0 to reduce the compound library to a reasonable size and then proceed with molecule docking that provides a refined compound ranking and reveals details of the interaction mechanism. Inforna has proven successful in various studies^{6, 114}, such as identifying small molecules that target oncogenic miRNA precursors^{115, 116}, an A bulge in the iron responsive element (IRE)¹¹⁷ of the SNCA mRNA related to Parkinson's disease¹¹⁸, and a miRNA-200 family member precursor associated with type 2 diabetes¹¹⁹, etc.

Quantitative Structure-Activity Relationship (QSAR) models for RNA-small molecule binding

Studies employing QSAR models have demonstrated their effectiveness in prioritizing drug candidates and optimizing lead compounds^{47–51, 72, 120–123}. Compared to structure-based approaches, QSAR models are computationally efficient, allowing researchers to screen large compound libraries. These models leverage statistical and machine learning techniques to analyze existing data on known compounds, identifying patterns that link structural characteristics to specific biological effects or physicochemical properties such as binding affinity. However, QSAR models have limitations. They are typically trained on small molecule datasets for specific RNA targets, which makes them difficult to generalize to other RNA types. Additionally, they usually cannot provide detailed descriptions of the microscopic interactions involved in binding. Despite QSAR's success in protein-based drug design, only a few QSAR studies have been conducted for RNA targets. In the following paragraph, we discuss several selected applications of QSAR in RNA-targeted drug discovery. Interested readers are directed to the references^{47–51, 72, 120–123} in addition to those reviewed in the main text.

Cai *et al.*⁴⁹ developed a workflow utilizing QSAR to connect molecular descriptors of a given ligand with its binding profiles, including binding affinity K_D and kinetic rate constants (k_{on} and k_{off}), against a specific RNA. To test the proposed workflow, the authors chose the HIV-1 TAR element as the RNA target, since TAR is a well-studied antiviral target with many experimentally verified bound compound candidates for training^{108, 124–126}. The training set comprised 48 compounds, including 29 reported TAR-binding ligands and 19 compounds with known RNA targeted scaffolds. For each ligand, 435 calculated descriptors, ranging from topological to electrostatic terms, along with their measured binding parameters,

were used to train the regression model. The least absolute shrinkage and selection operator (LASSO) was employed prior to training to select an optimal set of molecular descriptors. Comparison of predictive performance on the test set between multiple linear regression (MLR) model and various ensemble tree methods (e.g., random forest and boosting trees) showed that MLR gave overall best correlation between predicted and measured binding parameters. One possible reason is the utilization of the LASSO for descriptor selection prior to MLR. Lasso has been widely used in QSAR studies to control model complexity by applying a penalty constraint to the loss function during modeling. This helps identify the most relevant descriptors and reduce noise, which might have been particularly effective for this dataset.

A common practice after identifying lead compounds involves optimization cycles where certain chemical functional groups are replaced to improve the potency of the candidates. With the availability of large virtual screening libraries of small molecules and experimentally validated activities, the principles governing ligand binding can be learned by machine learning models to guide drug design. Grimberg *et al.*⁵¹ compared several machine learning-based models, including LASSO regression, decision tree classifier, and convolutional neural network (CNN), for optimizing phenylthiazole-containing small-molecule inhibitors targeting the RNA hairpin 91 within the ribosomal peptidyl transferase center (PTC) of *Mycobacterium tuberculosis*. By training machine learning models on a dataset containing 791 drug-sized molecules with the 2-phenylthiazole moiety and their predicted binding scores for the RNA target (generated by the molecular docking program AutoDock⁵²), chemical features important for small molecule binding were identified and subsequently used as guidelines for synthesizing new inhibitors. Functional validation was conducted after synthesizing ten small molecules, of which four were found to be potent inhibitors that target hairpin 91 in the ribosomal PTC of *M. tuberculosis* and stop translation.

CONCLUSIONS

RNA-targeted drug discovery is a rapidly evolving field with immense potential for addressing various human pathologies, from bacterial and viral infections to cancer. Although structure-based approaches such as molecular docking and QSAR modeling have already demonstrated success in identifying novel RNA-targeting compounds^{6, 37, 41, 42, 44, 51, 92, 93, 115, 117}, significant challenges still remain,

particularly in accounting for RNA's inherent flexibility and the relative scarcity of high-quality structural and binding data. The development of more advanced molecular tools for RNA-based therapeutics^{127–129}, together with more comprehensive and well-curated databases of RNA-small molecule interactions, as discussed in this review, will be crucial for advancing the field. These resources will provide the necessary data to train more sophisticated machine learning models and improve our understanding of the principles governing RNA-ligand recognition. With major pharma companies converging upon this field and forging various biotech-pharma alliances¹¹, we anticipate an unprecedented increase in the demand for efficient and accurate computational methods for modeling RNA-small molecule interactions. This may include the development of RNA-specific scoring functions, improved modeling of RNA dynamics upon ligand recognition, and integration of experimental data derived from NMR, X-ray crystallography, and cryo-EM to generate more accurate structural ensembles. Additionally, the application of deep learning approaches is likely to yield powerful new predictive models as more data becomes available. As computational methods for RNA-small molecule binding continue to advance, we can realistically anticipate a notable rise in RNA-targeted small molecule therapeutics entering clinical trials in the coming years.

AUTHOR CONTRIBUTIONS

S-J.C. and Y.Z. conceived the project. S-J.C. acquired the funding. S-J.C. supervised the project. Y.Z. performed literature analysis. Y.Z. and S-J.C. wrote the manuscript. All authors have read and approved the manuscript.

FUNDING

This work was supported by the National Institutes of Health under Grants R35-GM134919 and U54-AI170660, and the National Science Foundation under Grant Number CHE-2154924.

CONFLICTS OF INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available in the manuscript or in the supplementary materials.

DECLARATION OF GENERATIVE AI IN SCIENTIFIC WRITING

The large language model (LLM), Claude, developed by Anthropic was used for language refinement only.

REFERENCES

1. Thomas Hermann. Small molecules targeting viral rna. *WIREs RNA*. 2016;7(6):726–743.
2. Colleen M. Connelly, Michelle H. Moon, and John S. Schneckloth. The emerging role of rna as a therapeutic target for small molecules. *Cell Chemical Biology*. 2016;23(9):1077–1090.
3. Katherine Deigan Warner, Christine E. Hajdin, and Kevin M. Weeks. Principles for targeting rna with drug-like small molecules. *Nature Reviews Drug Discovery*. 2018;17(8):547–558.
4. Judy Lieberman. Tapping the rna world for therapeutics. *Nature Structural & Molecular Biology*. 2018;25(5):357–364.
5. Noreen F. Rizvi, Jr. John P. Santa Maria, Ali Nahvi, Joel Klappenbach, Daniel J. Klein, *et al.* Targeting rna with small molecules: Identification of selective, rna-binding small molecules occupying drug-like chemical space. *SLAS DISCOVERY: Advancing the Science of Drug Discovery*. 2020;25(4):384–396. PMID: 31701793.
6. Matthew G. Costales, Jessica L. Childs-Disney, Hafeez S. Haniff, and Matthew D. Disney. How we think about targeting rna with small molecules. *Journal of Medicinal Chemistry*. 2020;63(17):8880–8900. PMID: 32212706.
7. Ai-Ming Yu, Young Hee Choi, and Mei-Juan Tu. Rna drugs and rna targets for small molecules: Principles, progress, and challenges. *Pharmacological Reviews*. 2020;72(4):862–898.
8. Samantha M. Meyer, Christopher C. Williams, Yoshihiro Akahori, Toru Tanaka, Haruo Aikawa, *et al.* Small molecule recognition of disease-relevant rna structures. *Chemical Society Reviews*. 2020;49:7167–7199.
9. Yanqiu Shao and Qiangfeng Cliff Zhang. Targeting rna structures in diseases with small molecules. *Essays in Biochemistry*. 2020;64(6):955–966.
10. Aline Umuhire Juru and Amanda E. Hargrove. Frameworks for targeting rna with small molecules. *Journal of Biological Chemistry*. 2021;296.
11. Cormac Sheridan. First small-molecule drug targeting rna gains momentum. *Nature Biotechnology*. 2021;39(1):6–8.
12. Jessica L. Childs-Disney, Xueyi Yang, Quentin M. R. Gibaut, Yuquan Tong, *et al.* Targeting rna structures with small molecules. *Nature Reviews Drug Discovery*. 2022;21(10):736–762.
13. Joseph C. Somody, Stephen S. MacKinnon, and Andreas Windemuth. Structural coverage of the proteome for pharmaceutical applications. *Drug Discovery Today*. 2017;22(12):1792–1799.
14. Francesca Tessaro and Leonardo Scapozza. How ‘protein-docking’ translates into the new emerging field of docking small molecules to nucleic acids? *Molecules*. 2020;25(12).
15. Matthew J. Hangauer, Ian W. Vaughn, and Michael T. McManus. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding rnas. *PLOS Genetics*. 2013;9(6):1–13.
16. Andrew L. Hopkins and Colin R. Groom. The druggable genome. *Nature Reviews Drug Discovery*. 2002;1(9):727–730.
17. John P. Overington, Bissan Al-Lazikani, and Andrew L. Hopkins. How many drug targets are there? *Nature Reviews Drug Discovery*. 2006;5(12):993–996.
18. Michele Clamp, Ben Fry, Mike Kamal, Xiaohui Xie, James Cuff, *et al.* Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences*. 2007;104(49):19428–19433.
19. Scott J Dixon and Brent R Stockwell. Identifying druggable disease-modifying gene products. *Current Opinion in Chemical Biology*. 2009;13(5):549–555. Omics/Biopolymers/Model Systems.
20. Iakes Ezkurdia, David Juan, Jose Manuel Rodriguez, Adam Frankish, *et al.* Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Human Molecular Genetics*. 2014;23(22):5866–5878.
21. Masayuki Matsui and David R. Corey. Non-coding rnas as drug targets. *Nature Reviews Drug Discovery*. 2017;16(3):167–179.
22. Zhiwen Fu, Liu Wang, Shijun Li, Fen Chen, *et al.* Microrna as an important target for anticancer drug development. *Frontiers in Pharmacology*. 12, 2021.
23. Luisa Statello, Chun-Jie Guo, Ling-Ling Chen, and Maite Huarte. Gene regulation by long non-coding rnas and its biological functions. *Nature Reviews Molecular Cell Biology*. 2021;22(2):96–118.
24. Tim R. Mercer, Trent Munro, and John S. Mattick. The potential of long noncoding rna therapies. *Trends in Pharmacological Sciences*. 2022;43(4):269–280.
25. John S. Mattick, Paulo P. Amaral, Piero Carninci, Susan Carpenter, *et al.* Long non-coding rnas: definitions, functions, challenges and recommendations. *Nature Reviews Molecular Cell Biology*. 2023;24(6):430–447.

26. Kinga Nemeth, Recep Bayraktar, Manuela Ferracin, and George A. Calin. Non-coding rnas in disease: from mechanisms to therapeutics. *Nature Reviews Genetics*. 2024;25(3):211–232.
27. John A. Howe, Hao Wang, Thierry O. Fischmann, Carl J. Balibar, *et al.* Selective small-molecule inhibition of an rna structural element. *Nature*. 2015;526(7575):672–677.
28. Stephen E. Motika, Rebecca J. Ulrich, Emily J. Geddes, Hyang Yeon Lee, Gee W. Lau, and Paul J. Hergenrother. Gram-negative antibiotic active through inhibition of an essential riboswitch. *Journal of the American Chemical Society*. 2020;142(24):10856–10862. PMID:32432858.
29. Thierry O. Fischmann, Hao Wang, Haifeng Tang, Artjohn Villafania, *et al.* Atomic resolution mechanistic studies of ribocil: A highly selective unnatural ligand mimic of the e. coli fmn riboswitch. *RNA Biology*. 2016;13(10):946–954. PMID: 27485612.
30. Hao Wang, Paul A. Mann, Li Xiao, Charles Gill, *et al.* Dual-Targeting Small-Molecule Inhibitors of the Staphylococcus aureus FMN Riboswitch Disrupt Riboflavin Homeostasis in an Infectious Setting. *Cell Chemical Biology*. 2017;24(5):576–588.e6.
31. Vipul Panchal and Ruth Brenk. Riboswitches as drug targets for antibiotics. *Antibiotics*. 2021;10(1).
32. Jinzhong Lin, Dejian Zhou, Thomas A. Steitz, Yury S. Polikanov, and Matthieu G. Gagnon. Ribosome-targeting antibiotics: Modes of action, mechanisms of resistance, and implications for drug design. *Annual Review of Biochemistry*. 2018;87(1):451–478. PMID: 29570352.
33. Laiying Zhang, Jun He, Lang Bai, Shihua Ruan, Tao Yang, and Youfu Luo. Ribosome-targeting antibacterial agents: Advances, challenges, and opportunities. *Medicinal Research Reviews*. 2021;41(4):1855–1889.
34. Nikolai A. Naryshkin, Marla Weetall, Amal Dakka, Jana Narasimhan, *et al.* SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science*. 2014;345(6197):688–693.
35. Ewan P Plant, Gabriela C P'erez-Alvarado, Jonathan L Jacobs, Bani Mukhopadhyay, Mirko Hennig, and Jonathan D Dinman. A three-stemmed mrna pseudoknot in the sars coronavirus frameshift signal. *PLOS Biology*. 2005;3(6).
36. Mei-Chi Su, Chung-Te Chang, Chiu-Hui Chu, Ching-Hsiu Tsai, and Kung-Yao Chang. An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. *Nucleic Acids Research*. 2005;33(13):4265–4275.
37. So-Jung Park, Yang-Gyun Kim, and Hyun-Ju Park. Identification of rna pseudoknot-binding ligand that inhibits the -1 ribosomal frameshifting of sars-coronavirus by structure-based virtual screening. *Journal of the American Chemical Society*. 2011;133(26):10094–10100. PMID: 21591761.
38. Dustin B. Ritchie, Jingchuan Soong, William K. A. Sikkema, and Michael T. Woodside. Anti-frameshifting ligand reduces the conformational plasticity of the sars virus pseudoknot. *Journal of the American Chemical Society*. 2014;136(6):2196–2199. PMID: 24446874.
39. Mo Yang, Feyisola P. Olatunji, Curran Rhodes, Sumirtha Balaratnam, *et al.* Discovery of small molecules targeting the frameshifting element rna in sars-cov-2 viral genome. *ACS Medicinal Chemistry Letters*. 2023;14(6):757–765.
40. Monalisa Swain, Abeer A Ageeli, Wojciech K Kasprzak, Mi Li, *et al.* Dynamic bulge nucleotides in the kshv pan ene triple helix provide a unique binding platform for small molecule ligands. *Nucleic Acids Research*. 2021;49(22):13179–13193.
41. Andrew C. Stelzer, Aaron T. Frank, Jeremy D. Kratz, Michael D. Swanson, *et al.* Discovery of selective bioactive small molecules by targeting an rna dynamic ensemble. *Nature Chemical Biology*. 2011;7(8):553–559.
42. Laura R. Ganser, Janghyun Lee, Atul Rangadurai, Dawn K. Merriman, *et al.* High-performance virtual screening by targeting a high-resolution rna dynamic ensemble. *Nature Structural & Molecular Biology*. 2018;25(5):425–434.
43. Andrew P. Longhini, Lukasz T. Olenginski, Fardokht Abulwerdi, Stefano Ginocchio, *et al.* Structural insights of the conserved “priming loop” of hepatitis b virus pre-genomic rna. *Journal of Biomolecular Structure and Dynamics*. 2022;40(20):9761–9773. PMID: 34155954.
44. Lukasz T. Olenginski, Wojciech K. Kasprzak, Solomon K. Attionu, Bruce A. Shapiro, and Theodore K. Dayie. Virtual screening of hepatitis b virus pre-genomic rna as a novel therapeutic target. *Molecules*. 2023;28(4).
45. Hasane Ratni, Martin Ebeling, John Baird, Stefanie Bendels, *et al.* Discovery of risdiplam, a selective survival of motor neuron-2 (smn2) gene splicing modifier for the treatment of spinal muscular atrophy (sma). *Journal of Medicinal Chemistry*. 2018;61(15):6501–6517. PMID: 30044619.
46. Theodora Markati, Gemma Fisher, Sithara Ramdas, and Laurent Servais. Risdiplam: an investigational survival motor neuron 2 (smn2) splicing modifier for spinal muscular atrophy (sma). *Expert Opinion on Investigational Drugs*. 2022;31(5):451–461. PMID: 35316106.
47. Greta Bagnolini, TinTin B. Luu, and Amanda E. Hargrove. Recognizing the power of machine learning and other computational methods to accelerate progress in small molecule targeting of rna. *RNA*. 2023;29(4):473–488.

48. Jonathan M. Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, *et al.* A deep learning approach to antibiotic discovery. *Cell*. 2020;180(4):688–702.e13.
49. Zhengguo Cai, Martina Zafferani, Olanrewaju M. Akande, and Amanda E. Hargrove. Quantitative structure-activity relationship (qsar) study predicts small-molecule binding to rna structure. *Journal of Medicinal Chemistry*. 2022;65(10):7262–7277. PMID: 35522972.
50. Kamyar Yazdani, Deondre Jordan, Mo Yang, Christopher R. Fullenkamp, *et al.* Machine learning informs rna-binding chemical space**. *Angewandte Chemie International Edition*. 2023;62(11):e202211358.
51. Hadar Grimberg, Vinay S. Tiwari, Benjamin Tam, Lihi Gur-Arie, Daniela Gingold, Lea Polachek, and Barak Akabayov. Machine learning approaches to optimize small-molecule inhibitors for rna targeting. *Journal of Cheminformatics*. 2022;14(1):4.
52. Ruth Huey, Garrett M. Morris, Arthur J. Olson, and David S. Goodsell. A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*. 2007;28(6):1145–1152.
53. P. Therese Lang, Scott R. Brozell, Sudipto Mukherjee, Eric F. Pettersen, *et al.* Dock 6: Combining techniques to model rna-small molecule complexes. *RNA*. 2009;15(6):1219–1230.
54. Sergio Ruiz-Carmona, Daniel Alvarez-Garcia, Nicolas Foleppe, A. Beatriz Garmendia-Doval, *et al.* rdock: A fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLOS Computational Biology*. 2014;10(4):1–7.
55. Li-Zhen Sun, Yangwei Jiang, Yuanzhe Zhou, and Shi-Jie Chen. Rldock: A new method for predicting rna-ligand interactions. *Journal of Chemical Theory and Computation*. 2020;16(11):7173–7183. PMID: 33095555.
56. Yangwei Jiang and Shi-Jie Chen. Rldock method for predicting rna-small molecule binding modes. *Methods*. 2021.
57. Yuyu Feng, Keqiong Zhang, Qilong Wu, and Sheng-You Huang. Nldock: a fast nucleic acid-ligand docking algorithm for modeling rna/dna-ligand complexes. *Journal of Chemical Information and Modeling*. 2021;61(9):4771–4782. PMID: 34468128.
58. Patrick Pfeffer and Holger Gohlke. Drugscorernaknowledge-based scoring function to predict rna-ligand interactions. *Journal of Chemical Information and Modeling*. 2007;47(5):1868–1876. PMID: 17705464.
59. Anna Philips, Kaja Milanowska, Grzegorz Lach, and Janusz M. Bujnicki. Ligandrna: computational predictor of rna-ligand interactions. *RNA*. 2013;19(12):1605–1616.
60. Zhiqiang Yan and Jin Wang. Spa-lin: a scoring function of ligand-nucleic acid interactions via optimizing both specificity and affinity. *Nucleic Acids Research*. 2017;45(12):e110–e110.
61. Yuyu Feng and Sheng-You Huang. Itscore-nl: An iterative knowledge-based scoring function for nucleic acid-ligand interactions. *Journal of Chemical Information and Modeling*. 2020;60(12):6698–6708. PMID: 33291885.
62. Filip Stefaniak and Janusz M. Bujnicki. Annapurna: A scoring function for predicting rna- small molecule binding poses. *PLOS Computational Biology*. 2021;17(2):1–31.
63. Sahil Chhabra, Jingru Xie, and Aaron T. Frank. Rnaposers: Machine learning classifiers for ribonucleic acid-ligand poses. *The Journal of Physical Chemistry B*. 2020;124(22):4436–4445. PMID: 32427491.
64. Yichen Liu and Aaron T. Frank. Using selectively scaled molecular dynamics simulations to assess ligand poses in rna aptamers. *Journal of Chemical Theory and Computation*. 2022;18(9):5703–5709. PMID: 35926894.
65. Carlos Oliver, Vincent Mallet, Roman Sarrazin Gendron, Vladimir Reinharz, Nicolas Hamilton, William Waldispuhl. Augmented base pairing networks encode RNA-small molecule binding preferences. *Nucleic Acids Research*. 2020;48(14):7690–7699.
66. Yuanzhe Zhou, Yangwei Jiang, and Shi-Jie Chen. Rna-ligand molecular docking: Advances and challenges. *WIREs Computational Molecular Science*. 2022;12(3):e1571.
67. Chia-Chieh Chu, Raphael Plangger, Christoph Kreutz, and Hashim M Al-Hashimi. Dynamic ensemble of HIV-1 RRE stem IIB reveals non-native conformations that disrupt the Rev-binding site. *Nucleic Acids Research*. 2019;47(13):7105–7117.
68. Laura R. Ganser, Megan L. Kelly, Neeraj N. Patwardhan, Amanda E. Hargrove, and Hashim M. Al-Hashimi. Demonstration that small molecules can bind and stabilize low- abundance short-lived rna excited conformational states. *Journal of Molecular Biology*. 2020;432(4):1297–1304.
69. Dejun Jiang, Huifeng Zhao, Hongyan Du, Yafeng Deng, *et al.* How good are current docking programs at nucleic acid–ligand docking? a comprehensive evaluation. *Journal of Chemical Theory and Computation*. 2023;19(16):5633–5647. PMID: 37480347.
70. Rupesh Agarwal, Rajitha Rajeshwar T, and Jeremy C. Smith. Comparative assessment of pose prediction accuracy in rna–ligand docking. *Journal of Chemical Information and Modeling*. 2023;63(23):7444–7452. PMID: 37972310.
71. Yuanzhe Zhou and Shi-Jie Chen. Advances in machine-learning approaches to rna-targeted drug design. *Artificial*

- Intelligence Chemistry*. 2024;2(1):100053.
72. Ella Czarina Morishita and Shingo Nakamura. Recent applications of artificial intelligence in rna-targeted small molecule drug discovery. *Expert Opinion on Drug Discovery*. 2024;19(4):415–431. PMID: 38321848.
 73. Jacopo Manigrasso, Marco Marcia, and Marco De Vivo. Computer-aided design of rna- targeted small molecules: A growing need in drug discovery. *Chem*. 2021;7(11):2965–2988.
 74. Mattia Bernetti, Riccardo Aguti, Stefano Bosio, Maurizio Recanatini, Matteo Masetti, and Andrea Cavalli. Computational drug discovery under rna times. *QRB Discovery*. 2022;3:e22.
 75. Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, *et al*. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*. 2024;630(8016):493–500.
 76. Zhihai Liu, Minyi Su, Li Han, Jie Liu, *et al*. Forging the basis for developing protein–ligand interaction scoring functions. *Accounts of Chemical Research*. 2017;50(2):302–309. PMID: 28182403.
 77. Subodh Kumar Mishra and Amit Kumar. NALDB: nucleic acid ligand database for small molecules targeting nucleic acid. *Database*. 2016;2016:baw002.
 78. Matthew D. Disney, Audrey M. Winkelsas, Sai Pradeep Velagapudi, Mark Southern, *et al*. Inforna 2.0: A platform for the sequence-based design of small molecules targeting structured rnas. *ACS Chemical Biology*. 2016;11(6):1720–1728. PMID: 27097021.
 79. Sumit Mukherjee, Sukhen Das Mandal, Nikita Gupta, Matan Drory-Retwitzer, *et al*. RiboD: a comprehensive database for prokaryotic riboswitches. *Bioinformatics*. 2019;35(18):3541–3543.
 80. Lulu Li, Pengfei Wu, Zhenyu Wang, Xiangqi Meng, *et al*. Noncorna: a database of experimentally supported non-coding rnas and drug targets in cancer. *Journal of Hematology & Oncology*. 2020;13(1):15.
 81. Yu-Huan Wang, Qian-Fan Yang, Xiao Lin, Die Chen, *et al*. G4LDB 2.2: a database for discovering and studying G-quadruplex and i-Motif ligands. *Nucleic Acids Research*. 2021;50(D1):D150–D160.
 82. Robert Penchovsky, Nikolett Pavlova, and Dimitrios Kaloudas. Rswitch: A novel bioinformatics database on riboswitches as antibacterial drug targets. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. 2021;18(2):804–808.
 83. Saisai Sun, Jianyi Yang, and Zhaolei Zhang. Rnaligands: a database and web server for rna-ligand interactions. *RNA*. 2022;28(2):115–122.
 84. Anita Donlic, Emily G. Swanson, Liang-Yuan Chiu, Sarah L. Wicks, *et al*. R-bind 2.0: An updated database of bioactive rna-targeting small molecules and associated rna secondary structures. *ACS Chemical Biology*. 2022;17(6):1556–1566. PMID: 35594415.
 85. F P Panei, R Torchet, H M'enager, P Gkeka, and M Bonomi. HARIBOSS: a curated database of RNA-small molecules structures to aid rational drug design. *Bioinformatics*. 2022;38(17):4185–4193.
 86. Sowmya Ramaswamy Krishnan, Arijit Roy, and M. Michael Gromiha. R-sim: A database of binding affinities for rna-small molecule interactions. *Journal of Molecular Biology*. 2023;435(14):167914. Computation Resources for Molecular Biology.
 87. Fan Bu, Xiaowei Lin, Wenjian Liao, Zhizhong Lu, *et al*. Ribocentre-switch: a database of riboswitches. *Nucleic Acids Research*. 2023;52(D1):D265–D272.
 88. Long Chen, Zhuohang Yu, Zengrui Wu, Moran Zhou, *et al*. Aptadb: a comprehensive database integrating aptamer–target interactions. *RNA*. 2024;30(3):189–199.
 89. Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, *et al*. The Protein Data Bank. *Nucleic Acids Research*. 2000;28(1):235–242.
 90. Brittany S. Morgan, Bilva G. Sanaba, Anita Donlic, Diane B. Karloff, *et al*. Hargrove. R-bind: An interactive database for exploring and developing rna-targeted chemical probes. *ACS Chemical Biology*. 2019;14(12):2691–2700. PMID: 31589399.
 91. Jiaying Luo, Wanlei Wei, J'érôme Waldispuhl, and Nicolas Moitessier. Challenges and current status of computational methods for docking small molecules to nucleic acids. *European Journal of Medicinal Chemistry*. 2019;168:414–425.
 92. Peter Daldrop, Francis E. Reyes, David A. Robinson, Colin M. Hammond, *et al*. Novel ligands for a purine riboswitch discovered by rna-ligand docking. *Chemistry & Biology*. 2011;18(3):324–335.
 93. Elisabeth Kallert, Tim R. Fischer, Simon Schneider, Maike Grimm, *et al*. Protein-based virtual screening tools applied for rna–ligand docking identify new binders of the preq1-riboswitch. *Journal of Chemical Information and Modeling*. 2022;62(17):4134–4148. PMID: 35994617.
 94. Ben Hu, Hua Guo, Peng Zhou, and Zheng-Li Shi. Characteristics of sars-cov-2 and covid-19. *Nature Reviews Microbiology*. 2021;19(3):141–154.
 95. Pramod R. Bhatt, Alain Scaiola, Gary Loughran, Marc Leibundgut, *et al*. Structural basis of ribosomal frameshifting during translation of the sars-cov-2 rna genome. *Science*. 2021;372(6548):1306–1313.

96. Todd J.A. Ewing, Shingo Makino, A. Geoffrey Skillman, and Irwin D. Kuntz. Dock 4.0: Search strategies for automated molecular docking of flexible molecule databases. *Journal of Computer-Aided Molecular Design*. 2001;15(5):411–428.
97. Yanga Byun and Kyungsook Han. PseudoViewer: web application and web service for visualizing RNA pseudoknots and secondary structures. *Nucleic Acids Research*. 2006;34(suppl.):W416–W422.
98. Kyungsook Han and Yanga Byun. PseudoViewer2: visualization of RNA pseudoknots of any type. *Nucleic Acids Research*. 2003;31(13):3432–3440.
99. David A. Case, Thomas E. Cheatham III, Tom Darden, *et al.* The amber biomolecular simulation programs. *Journal of Computational Chemistry*. 2005;26(16):1668–1688.
100. Gregory Mathez, Andrea Brancale, and Valeria Cagno. Novel inhibitors of sars-cov-2 rna identified through virtual screening. *Journal of Chemical Information and Modeling*. 2024;64(15):6190–6196. PMID: 39037082.
101. Schrodinger release 2023-2; maestro, schrodinger, llc new york, ny, 2023.
102. Lukasz T. Olenginski, Savannah F. Spradlin, and Robert T. Batey. Flipping the script: Understanding riboswitches from an alternative perspective. *Journal of Biological Chemistry*. 2024;300(3):105730.
103. Kumari Kavita and Ronald R. Breaker. Discovering riboswitches: the past and the future. *Trends in Biochemical Sciences*. 2023;48(2):119–141.
104. Mark McGann. Fred pose prediction and virtual screening accuracy. *Journal of Chemical Information and Modeling*. 2011;51(3):578–596. PMID: 21323318.
105. Mark McGann. Fred and hybrid docking performance on standardized datasets. *Journal of Computer-Aided Molecular Design*. 2012;26(8):897–906.
106. Matthias Rarey, Bernd Kramer, Thomas Lengauer, and Gerhard Klebe. A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology*. 1996;261(3):470–489.
107. Steven G. Deeks, Julie Overbaugh, Andrew Phillips, and Susan Buchbinder. Hiv infection. *Nature Reviews Disease Primers*. 2015;1(1):15035.
108. Fardokht A. Abulwerdi, Matthew D. Shortridge, Joanna Sztuba-Solinska, Robert Wilson, *et al.* Schneekloth. Development of small molecules with a noncanonical binding mode to hiv-1 trans activation response (tar) rna. *Journal of Medicinal Chemistry*. 2016;59(24):11148–11160. PMID: 28002966.
109. Ruben Abagyan, Maxim Totrov, and Dmitry Kuznetsov. Icm - a new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*. 1994;15(5):488–506.
110. Loic Salmon, Gavin Bascom, Ioan Andricioaei, and Hashim M. Al-Hashimi. A general method for constructing atomic-resolution rna ensembles using nmr residual dipolar couplings: The basis for interhelical motions revealed. *Journal of the American Chemical Society*. 2013;135(14):5457–5466. PMID: 23473378.
111. Jeffrey D. Stanaway, Abraham D. Flaxman, Mohsen Naghavi, Christina Fitzmaurice, *et al.* The global burden of viral hepatitis from 1990 to 2013: findings from the global burden of disease study 2013. *The Lancet*. 2016;388(10049):1081–1088.
112. Oleg Trott and Arthur J. Olson. Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient timization. *Journal of Computational Chemistry*. 2010;31(2):455–461.
113. Sai Pradeep Velagapudi, Steven M. Gallo, and Matthew D. Disney. Sequence-based design of bioactive small molecules that target precursor micrnas. *Nature Chemical Biology*. 2014;10(4):291–297.
114. Blessy M. Suresh, Amirhossein Taghavi, Jessica L. Childs-Disney, and Matthew D. Disney. Fragment-based approaches to identify rna binders. *Journal of Medicinal Chemistry*. 2023;66(10):6523–6541. PMID: 37186875
115. Sai Pradeep Velagapudi, Yiling Luo, Tuan Tran, Hafeez S. Haniff, *et al.* Defining rna- small molecule affinity landscapes enables design of a small molecule inhibitor of an oncogenic noncoding rna. *ACS Central Science*. 2017;3(3):205–216. PMID: 28386598.
116. Matthew D. Disney, Sai Pradeep Velagapudi, Yue Li, Matthew G. Costales, and Jessica L. Childs-Disney. Chapter three - identifying and validating small molecules interacting with rna (smirnas). In Amanda E. Hargrove, editor, *RNA Recognition*, volume 623 of *Methods in Enzymology*. Academic Press. 2019; pages 45–66.
117. Peiyuan Zhang, Hye-Jin Park, Jie Zhang, Eunsung Junn, *et al.* Translation of the intrinsically disordered protein α -synuclein is inhibited by a small molecule targeting its structured mrna. *Proceedings of the National Academy of Sciences*. 2020;117(3):1457–1467.
118. Woojin Scott Kim, Katarina K'agedal, and Glenda M. Halliday. Alpha-synuclein biology in lewy body diseases. *Alzheimer's Research & Therapy*. 2014;6(5):73.
119. Hafeez S. Haniff, Xiaohui Liu, Yuquan Tong, Samantha M. Meyer, *et al.* A structure-specific small molecule inhibits a mirna-200 family member precursor and reverses a type 2 diabetes phenotype. *Cell Chemical Biology*.

- 2022;29(2):300–311.
120. Piotr Setny and Joanna Trylska. Search for novel aminoglycosides by combining fragment- based virtual screening and 3d-qsar scoring. *Journal of Chemical Information and Modeling*. 2009;49(2):390–400. PMID: 19173606.
 121. Y. Nataraja Sekhar, M. Ravi Shashi Nayana, N. Sivakumari, Muttineni Ravikumar, and S.K. Mahmood. 3d-qsar and molecular docking studies of 1,3,5-triazene-2,4-diamine derivatives against r-rna: Novel bacterial translation inhibitors. *Journal of Molecular Graphics and Modelling*. 2008;26(8):1338–1352.
 122. Iwona Maciagiewicz, Shu Zhou, Stephen C. Bergmeier, and Jennifer V. Hines. Structure–activity studies of rna-binding oxazolidinone derivatives. *Bioorganic and Medicinal Chemistry Letters*. 2011;21(15):4524–4527.
 123. Sowmya R Krishnan, Arijit Roy, and M Michael Gromiha. Reliable method for predicting the binding affinity of RNA-small molecule interactions using machine learning. *Briefings in Bioinformatics*. 2024;25(2):bbac002.
 124. Houn-Yau Mei, Mei Cui, Andrea Heldsinger, Shannon M. Lemrow, *et al.* Inhibitors of protein-rna complexation that target the rna: Specific recognition of human immunodeficiency virus type 1 tar rna by small organic molecules. *Biochemistry*. 1998;37(40):14204–14212. PMID: 9760258.
 125. Lei Zeng, Jiaming Li, Michaela Muller, Sherry Yan, *et al.* Selective small molecules blocking hiv-1 tat and coactivator pcaf association. *Journal of the American Chemical Society*. 2005;127(8):2376–2377. PMID:15724976.
 126. Joanna Sztuba-Solinska, Shilpa R. Shenoy, Peter Gareiss, Lauren R. H. Krumpe, *et al.* Schneekloth. Identification of biologically active, hiv tar rna-binding small molecules using small molecule microarrays. *Journal of the American Chemical Society*. 2014;136(23):8402–8410. PMID: 24820959.
 127. Morgan Chandler, Brittany Johnson, Emil Khisamutdinov, Marina A. Dobrovolskaia, *et al.* The international society of rna nanotechnology and nanomedicine (isrnn): The present and future of the burgeoning field. *ACS Nano*. 2021;15(11):16957–16973. PMID: 34677049.
 128. Lewis Rolband, Damian Beasock, Yang Wang, Yao-Gen Shu, *et al.* Biomotors, viral assembly, and rna nanobiotechnology: Current achievements and future directions. *Computational and Structural Biotechnology Journal*. 2022;20:6120– 6137.
 129. Yelizza I. Avila, Laura P. Rebolledo, Elizabeth Skelly, Renata de Freitas Saito, *et al.* Cracking the code: Enhancing molecular tools for progress in nanobiotechnology. *ACS Applied Bio Materials*. 2024;7(6):3587–3604. PMID: 38833534.