Opinion

Small molecules targeting viral RNA

Thomas Hermann

Highly conserved noncoding RNA (ncRNA) elements in viral genomes and transcripts offer new opportunities to expand the repertoire of drug targets for the development of antiviral therapy. Ligands binding to ncRNA architectures are able to affect interactions, structural stability or conformational changes and thereby block processes essential for viral replication. Proof of concept for targeting functional RNA by small molecule inhibitors has been demonstrated for multiple viruses with RNA genomes. Strategies to identify antiviral compounds as inhibitors of ncRNA are increasingly emphasizing consideration of drug-like properties of candidate molecules emerging from screening and ligand design. Recent efforts of antiviral lead discovery for RNA targets have provided drug-like small molecules that inhibit viral replication and include inhibitors of human immunodeficiency virus (HIV), hepatitis C virus (HCV), severe respiratory syndrome coronavirus (SARS CoV), and influenza A virus. While target selectivity remains a challenge for the discovery of useful RNA-binding compounds, a better understanding is emerging of properties that define RNA targets amenable for inhibition by small molecule ligands. Insight from successful approaches of targeting viral ncRNA in HIV, HCV, SARS CoV, and influenza A will provide a basis for the future exploration of RNA targets for therapeutic intervention in other viral pathogens which create urgent, unmet medical needs. Viruses for which targeting ncRNA components in the genome or transcripts may be promising include insect-borne flaviviruses (Dengue, Zika, and West Nile) and filoviruses (Ebola and Marburg). © 2016 Wiley Periodicals, Inc.

INTRODUCTION

While the large majority of approved drugs act on protein targets, a chemically diverse set of clinically important antibiotics bind selectively to the RNA components (rRNA) of the bacterial ribosome and thereby interfere with protein synthesis. These antibiotics are derived from natural products and provide a paradigm for targeting structured noncoding RNA as an approach for therapeutic intervention. Noncoding RNA (ncRNA) molecules are increasingly recognized as key regulatory players of biological processes in which they participate based on their three-dimensional structure rather than sequence. With their well-defined structure, ncRNA folds provide potentially unique interaction sites for selective small-molecule ligands that affect the RNAs’ biological function. Outside the bacterial ribosome, untranslated RNA regions as drug targets have been pursued to discover therapeutics for viral infections including those caused by the human immunodeficiency virus (HIV) and hepatitis C virus (HCV). These pathogens have single-stranded RNA genomes which include structured noncoding regions that play key roles in the viral infection cycle and harbor potential binding sites for small molecule inhibitors. The high conservation of untranslated regions in viral genomes along with the absence of

*Correspondence to: tcb@ucsd.edu

1Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, USA

2Center for Drug Discovery Innovation, University of California, San Diego, La Jolla, CA, USA

Conflict of interest: The author has declared no conflicts of interest for this article.

How to cite this article:
WIREs RNA 2016. doi: 10.1002/wrna.1373

© 2016 Wiley Periodicals, Inc.
homologous host cell RNAs render structured viral regulatory RNAs attractive targets for the development of novel antinfective drugs.7,8

Unlike typical inhibitor binding pockets in proteins, sites for ligand recognition in RNA are often shallow, highly solvated, and conformationally flexible. Therefore, structural and physicochemical properties of small molecule ligands for RNA differ markedly from drugs targeting proteins.9 Clinically used RNA-binding antibiotics are derived from natural products which, on average, are structurally more complex, more hydrophilic, and contain a larger number of hydrogen bond donors compared to drug-like compounds.10 These molecular properties create challenges in reconciling the design and optimization of RNA-binding ligands with drug-like features affecting absorption and distribution in vivo. Strategies for the design and discovery of RNA-binding ligands have been outlined long before the rise of ncRNA,11–13 and comprehensive reviews on targeting RNA with small molecules have been published regularly over the last years.14–16 Among the chemically diverse classes of RNA binders,9 recurring scaffolds are found among derivatives of aminoglycosides and planar intercalators including acridine and phenothiazine (Figure 1).

Aminoglycosides are natural products from Streptomyces soil bacteria that bind to the ribosomal decoding site and exert antibiotic activity by interfering with bacterial protein synthesis.17 Amine functional groups in aminoglycosides are protonated under physiological conditions which results in an overall cationic charge that promotes nucleic acid binding but hampers cell penetration. Aminoglycosides and their semi-synthetic derivatives have been reported to promiscuously bind numerous ncRNA targets9,14,15 including regulatory elements from HIV, which over two decades ago were the first viral RNAs to be investigated for small-molecule inhibition.18–20 It has been suggested that aminoglycosides bind to RNA folds through structural electrostatic complementarity whereby positively charged ammonium groups of the ligand recognize metal-ion-binding pockets created by the RNA architecture.21,22 Similarly, promiscuous binding to structured RNA has been observed for intercalating compounds which contain flat aromatic scaffolds such as acridine, phenazine and phenothiazine derivatives. While early ligand-binding studies of ncRNA focused on proof of concept and often involved promiscuous binders including aminoglycosides and intercalators, more recently, considerations of compounds’ drug-likeness have been prominently included in RNA targeting efforts. After all, some RNA-binding drugs such as the oxazolidinone antibiotics provide a compelling reminder that properties beneficial for RNA-targeting and drug-likeness may be reconciled within a small molecule compound (Figure 2).

Here, I focus on progress over the last 2 years in the discovery and investigation of small molecule ligands for viral RNA targets. Emphasis will be given to ligands with molecular weight <750 D. Coverage of the literature will occasionally go back further in time for previously reported compound classes if recent work has revealed new findings on mechanism of action or provided structural insight. The review of recently discovered ligands targeting viral RNA will include a discussion of compounds’ physicochemical properties that affect drug-likeness.

VIRAL RNA AS A DRUG TARGET

Antibiotics derived from natural products including aminoglycosides, macrolides, and tetracyclines have

![Figure 1](image1.png) **FIGURE 1** Aminoglycosides such as neomycin and kanamycin, and planar intercalators including derivatives of acridine, phenazine, and phenothiazine are promiscuous binders of multiple RNA targets.

![Figure 2](image2.png) **FIGURE 2** Linezolid is an oxazolidinone antibiotic (Zyvox®) that binds at the peptidyl transferase center in the 23S rRNA of the bacterial 50S ribosomal subunit and inhibits translation initiation.23 See the Box 1 for an explanation of molecular properties shown here for linezolid.
long been known to interfere with translation, and these compounds provided powerful tools in the elucidation of ribosome function. After decades of mounting evidence that ribosome-inhibiting antibiotics do so by targeting the RNA components (rRNA), Moazed and Noller demonstrated in 1987 that several classes of antibiotics bind directly to 16S rRNA. These observations spurred efforts both in academia and pharmaceutical industry to discover small molecule ligands of other structured RNAs. The HIV transactivation response (TAR) and Rev response (RRE) RNA were the first viral regulatory elements shown to bind small molecule compounds that interfered with key steps of HIV infection. The realization that the repertoire of therapeutic approaches to combat infections may be expanded to inhibitors targeting viral RNA components quickly gave birth to the notion of RNA as an emerging drug target.

The HIV transactivation response (TAR) and Rev response (RRE) RNA were the first viral regulatory elements shown to bind small molecule compounds that interfered with key steps of HIV infection. The realization that the repertoire of therapeutic approaches to combat infections may be expanded to inhibitors targeting viral RNA components quickly gave birth to the notion of RNA as an emerging drug target.

Table 1: Viral RNA Targets

<table>
<thead>
<tr>
<th>Virus</th>
<th>Family</th>
<th>Genome</th>
<th>RNA Target</th>
<th>Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human immunodeficiency virus (HIV)</td>
<td>Retrovirus</td>
<td>(+)ssRNA</td>
<td>Transactivation response (TAR) element</td>
<td>5, 32, 33, and 34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rev response element (RRE)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dimer initiation sequence (DIS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Packaging signal (Ψ)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Internal ribosome entry site (IRES)</td>
<td>6</td>
</tr>
<tr>
<td>Hepatitis C virus (HCV)</td>
<td>Flavivirus</td>
<td>(+)ssRNA</td>
<td>Frameshifting signal</td>
<td></td>
</tr>
<tr>
<td>Severe acute respiratory syndrome coronavirus (SARS CoV)</td>
<td>Coronavirus</td>
<td>(+)ssRNA</td>
<td>RNA promoter for the viral RNA-dependent RNA polymerase (RdRp)</td>
<td>35</td>
</tr>
<tr>
<td>Influenza A virus</td>
<td>Orthomyxovirus</td>
<td>(−)ssRNA</td>
<td>RNA promoter for the viral RNA-dependent RNA polymerase (RdRp)</td>
<td></td>
</tr>
</tbody>
</table>

Box 1: Molecular Properties Used to Predict and Assess Drug-Likeness

**ClogP**, is the calculated logarithm of the partition coefficient between n-octanol and water as a measure of compound hydrophilicity. Low hydrophilicity is indicated by high log P values, and values >5.0 have been implicated with poor cell membrane permeation and absorption.

**PSA**, polar surface area, is the (calculated) molecular surface created by polar atoms, primarily oxygen and nitrogen and including the attached hydrogen atoms. A PSA >140 Å² causes poor cell membrane permeation.

**HBD**, is the number of hydrogen bond donors, primarily hydrogens linked to oxygen and nitrogen atoms. Hydrogen bond donors improve solubility in water but impede cell membrane permeation. The majority of approved drugs contain 0–4 HBD (median: 1).

**BA**, bioavailability, is the experimentally determined fraction of a non-intravenously administered drug dose that reaches systemic circulation.

Like HIV, HCV carries a (+) single-strand RNA genome, which is not reverse transcribed, but directly serves as a coding mRNA for viral proteins. Translation of the HCV genome is driven by a structured RNA element in the 5' untranslated region (UTR) which serves as an internal ribosome entry site (IRES). A subdomain of the HCV IRES functions as a conformational switch which has been identified as the target site for small molecule inhibitors capturing an extended conformation of the RNA and thereby blocking viral translation. A previous review by Dibrov et al. describes the discovery of HCV IRES-directed inhibitors and structural as well as mechanism-of-action studies that established the conformational switch motif as a privileged target site for antiviral RNA-binders. In this review, I will...
outline recent progress in the development of translation inhibitors targeting the HCV IRES (Figure 4 and Table 3).

Viral ncRNA targets for small molecule ligands which have emerged more recently and have not been reviewed previously are the frameshifting pseudoknot (PK) in the genome of the SARS coronavirus (CoV) and the promoter for the viral RNA-dependent RNA polymerase in the (-)ssRNA genome of the influenza A virus. Ligand discovery efforts and structural studies of these targets will be discussed here (Figures 5 and 6; Tables 4 and 5).

**HIV Targets**

The (+) single-strand RNA genome of HIV contains multiple regulatory elements which are involved in transcriptional regulation, reverse transcription, viral protein translation, nucleocyttoplasmic transport, genome dimerization, and virion packaging. The TAR element resides within the first 59 nucleotides of the viral genome and serves as the binding site for the Tat protein. Formation of the Tat/TAR complex stimulates transcription elongation to yield full-length transcripts by a complex mechanism that also involves host cell factors. The Tat-binding site in TAR consists of a conserved RNA stem-loop with a pyrimidine-rich bulge (Figure 2) which is

![Figure 3](https://i.imgur.com/3.png) Secondary structure of the HIV-1 TAR (a) and RRE stem-loop IIB (b) RNA targets.

| TABLE 2 | Small Molecules Targeting HIV RNA |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Compound | Properties | Target | Discovery | References |
| 1 | ClogP: 4.3 PSA: 52 Å² HBD: 3 | TAR | Small-molecule microarray screening (SMM) with fluorescently labeled RNA | 36 |
| 2 | ClogP: 0.4 PSA: 75 Å² HBD: 5 | TAR | Fragment competition binding screening with fluorescently labeled peptide ligand | 37 |
| 3 | ClogP: 0.2 PSA: 99 Å² HBD: 5 | TAR | Scaffold-based design; fluorescence-based target binding confirmation | 38 |
| 4 | ClogP: 7.1 PSA: 11 Å² HBD: 0 | RRE SL-IIB | Competition binding screening with fluorescently labeled peptide ligand of FDA-approved drugs | 39 |
| 5 | ClogP: 5.2 PSA: 52 Å² HBD: 0 | RRE SL-IIB | Competition binding screening with fluorescently labeled peptide ligand of FDA-approved drugs | 39 |

1 See Box 1 for an explanation of molecular properties. ClogP and PSA were calculated with the Molsoft Molecular Properties calculator (Molsoft LLC, San Diego).
2 See Box 2 for a description of common fluorimetric ligand binding assays for RNA targets.
**FIGURE 4** | The HCV IRES subdomain IIA RNA target. (a) Secondary structure of the subdomain IIA internal loop. Residue numbering corresponds to the HCV genome sequence. (b) Crystal structure of the subdomain IIA RNA in complex with the benzimidazole IRES inhibitor 6 (Table 3). The ligand is shown in yellow stick representation. Positions of two magnesium ions are indicated by green spheres. (c) Target recognition by hydrogen bonding and partial intercalation of benzimidazole 6 in the subdomain IIA RNA complex crystal structure. Structure image was prepared from PDB coordinate file 3TZR.

**TABLE 3** | Small Molecules Targeting HCV IRES RNA

<table>
<thead>
<tr>
<th>Compound</th>
<th>Properties</th>
<th>Discovery</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>ClogP: 1.9, PSA: 47 Å², HBD: 2</td>
<td>Mass spectrometry-based screening; FRET-based target binding confirmation</td>
<td>6, 40, 41, and 42</td>
</tr>
<tr>
<td>7</td>
<td>ClogP: 3.9, PSA: 46 Å², HBD: 2</td>
<td>Scaffold-based design; FRET-based target binding confirmation</td>
<td>43 and 44</td>
</tr>
<tr>
<td>8</td>
<td>ClogP: 1.4, PSA: 53 Å², HBD: 3</td>
<td>Scaffold-based design; FRET-based target binding confirmation</td>
<td>43 and 45</td>
</tr>
<tr>
<td>9</td>
<td>ClogP: −1.9, PSA: 161 Å², HBD: 7</td>
<td>Scaffold-based design; fluorescence-based target binding confirmation</td>
<td>6 and 46</td>
</tr>
</tbody>
</table>
conformationally dynamic but adopts a stable, ordered structure in complex with Tat peptide, and small molecule ligands. Disruption of the Tat/TAR complex by competing ligands binding to the RNA target blocks efficient HIV replication and was recognized early as a potential path to the development of antiviral drugs.19,30,34 Numerous studies have focused on the discovery and design of TAR RNA-binding inhibitors, including peptides, natural products such as aminoglycosides and synthetic small molecules. These earlier efforts have been summarized in several excellent reviews.5,14,15,32,34

More recently, Schneekloth and coworkers have applied small molecule microarray (SMM) screening of a fluorescently labeled TAR hairpin model oligonucleotide to identify selective binders.36 A similar approach had been taken before by Disney and Seeberger to explore interactions of aminoglycosides with ribosomal A-site and group I intron ribozyme RNA.51 In the recent study on HIV TAR, the microarray-immobilized screening library was composed of 20,000 drug-like small molecules that were not biased for chemotypes considered to promote RNA binding. A counter-screen was applied against an unrelated RNA hairpin to exclude promiscuous binders. The SMM screen returned a low hit rate of 0.02%, excelled with a low false-positive rate, and eventually delivered a confirmed hit compound (1, Table 2) with reasonable drug-like properties, anti-HIV activity in T-lymphoblasts (EC50 value of 28 μM), as well as low cytotoxicity. The thienopyridine derivative 1 is surprisingly hydrophobic for an RNA-binding ligand and similar chemotypes have not been reported before as TAR-binders. It remains to be seen if the potency of thienopyridine derivatives

**Figure 5** Secondary structure of the SARS CoV frameshifting pseudoknot RNA.

**Figure 6** The Influenza A RNA promoter target. (a) Secondary structure of the RNA promoter. A construct used for ligand binding screening by NMR was derived by adding a stable tetraloop. Residue numbering refers to the NMR construct. (b) Model of a ligand-target complex derived by NMR spectroscopy of RNA promoter bound to compound 11 (Table 5). The ligand is shown in yellow stick representation. The added tetraloop is indicated in gray. (c) Detail view of the ligand binding site in the NMR model, showing hydrogen atoms of the ligand 11. Structure images were prepared from PDB coordinate file 2LWK.48
can be improved to furnish lead candidates for the development of antiviral drugs targeting the HIV TAR. Nevertheless, the SMM screening approach has proven to be a promising method for the discovery of drug-like ligands for RNA and fruitful application to other ncRNA targets is expected.

A fragment-based screening approach to identify ligands of the HIV TAR element has been described by Göbel and colleagues who interrogated a set of 29 small molecules that were selected to represent molecular motifs beneficial for RNA recognition.37 The fragments were rich in chemotypes that provide hydrogen bond donors and included amines, amidines, and guanidines as well as benzene rings for stacking interactions. A fluorimetric competition assay (Box 2) was used to determine ligand binding by testing for the ability of a compound to displace a dye-labeled Tat peptide from a TAR model RNA. Seven small molecules from the screening set were found to compete with Tat peptide binding at TAR with IC$_{50}$ values between 40 and 70 μM, including the quinazoline shown as an example in Table 2 (IC$_{50}$ value of 60 μM). At least two of the hit compounds had been identified previously as inhibitors of Tat/TAR complex formation and were shown to downregulate Tat transactivation in a cell-based assay.30 While cellular activity was not tested in the fragment screening approach reported by Göbel et al., one-dimensional $^1$H NMR was used to confirm ligand interaction with the RNA target by monitoring changes in imino-proton signals upon titration with compound. As a consequence of the composition bias of the screening set with small molecules rich in hydrogen bond donors, the hit compounds tended to exceed the number of hydrogen bond donors of typical drug-like molecules. However, fragment incorporation for construction of larger ligands in the future would reduce the number of hydrogen bond donor sites.

Benhida and coworkers pursued a design approach for TAR RNA-binding ligands based on amino-phenylthiazole derivatives38 which are nucleobase analogs that were proposed to interact with A–U pairs through hydrogen bonding of the amino-thiazole scaffold at the Hoogsteen edge of adenine.52 A set of 15 conjugates of the amino-phenylthiazole scaffold (termed ‘S nucleobase’; highlighted in blue in compound 3, Table 2) with different natural amino acids and dipeptides were synthesized and tested for RNA target binding as well as antiviral activity in cell culture. As a consequence of the design paradigm focusing on a scaffold for nucleobase-like interaction, the resulting compounds were hydrophilic and carried a larger number of hydrogen bond donor sites. Ligands of the TAR element were identified by monitoring fluorescent changes upon compound titration to a terminally dye-labeled TAR model oligonucleotide (Box 2). S nucleobase conjugates of arginine, lysine, and histidine were found to tightly
bind the TAR RNA. However, only the histidine derivative 3 (Table 2) was a selective TAR binder ($K_d = 17 \, \mu M$) whose affinity for the target was not diminished in the presence of other nucleic acids. The arginine and lysine conjugates showed better interaction with TAR RNA ($K_d = 2.4$ and $7.5 \, \mu M$), respectively) but binding was promiscuous and strongly affected by the presence of competitor nucleic acids. Perhaps attesting to a correlation of target selectivity and biological activity, only the histidine derivative 3 showed antiviral activity in HIV-infected human cells ($IC_{50} = 0.41 \, \mu M$). It is not clear why the viral inhibition potency of 3 exceeds the target binding affinity by 40-fold. This finding suggests, however, that compound 3 may act on other targets in addition to the TAR RNA.

Among recently described structurally complex ligands of the TAR RNA whose molecular weight exceeds 750 D are aminoglycoside-benzimidazole conjugates developed by Arya, Appella, and coworkers, and nucleobase-linked aminoglycosides synthesized by Hamasaki and colleagues. Aminoglycoside conjugates in both series are highly hydrophilic compounds which contain over 10 hydrogen bond donors and have nanomolar affinity for the TAR target. Antiviral activity of the conjugate compounds in cells has not been reported.

The HIV RRE element is located within a stretch of approximately 250 nucleotides in the second intron of the viral genomic RNA and has a more complex structure than the TAR motif. The stem-loop IIB (SL-IIB), which is formed by residues 45–75 of the HIV-1 RRE (Figure 3), serves as a high-affinity binding site for the viral Rev protein. Rev binding to RRE is required for the nucleocytoplasmic export of full-length and singly spliced viral transcripts. Disruption of the Rev–RRE interaction has been explored for over two decades as a strategy for the development of antiviral therapies. Past efforts to discover Rev competitor ligands for the RRE RNA have largely been focused on ligand-based design (summarized in comprehensive reviews). Only two studies have been published of small molecule high-throughput screens for inhibitors targeting the Rev–RRE complex. None of these approaches has yet produced bona fide inhibitors of in vitro Rev–RRE complex formation that show antiviral activity in cells. A recent study by Rana and colleagues has shown that post-transcriptional modification of HIV-1 RRE by N6-methylation of adenine bases in SL-IIB plays a critical role in the activity of the RRE/Rev complex, which serves as a poignant reminder that authentic model systems are requisite for the study of RNA targets.

Gallego and coworkers have identified inhibitors of HIV RNA biogenesis that block Rev–RRE interaction in a binding competition screen (Box 2) of 1120 FDA-approved drugs against complex formation between a fluorescently labeled Rev peptide and the RRE SL-IIB RNA target. Antiviral activity of hit compounds in cells was confirmed for...
clomiphene (4, Table 2) and cyproheptadine (5) which inhibited HIV transcription and affected levels of spliced versus unspliced viral transcripts to an extent that was consistent with their ability to disrupt Rev–RRE complex function in vitro. Clomiphene (4), a selective estrogen receptor modulator, inhibited Rev peptide binding to RRE SL-IIB with an IC<sub>50</sub> value of 3.5 μM and had antiviral activity with an EC<sub>50</sub> value of 4.3 μM in the cell-based assay. The corresponding values for cyproheptadine (5), an anti-histamine H1 receptor antagonist, were 4.2 and 17 μM (IC<sub>50</sub>/EC<sub>50</sub>). Interaction of clomiphene (4) was highly specific for the RRE SL-IIB target while target binding of cyproheptadine (5) was significantly compromised in the presence of competitor tRNA or DNA. NMR spectroscopic studies revealed the G-rich internal loop in the lower stem of SL-IIB as the interaction site for 4 and 5, which overlaps with the binding region of Rev and is consistent with the proposed mechanism of inhibition by ligand competition. Direct interaction of these hydrophobic drugs with an RNA target may appear unexpected and dispels the notion that a large number of hetero-atom hydrogen bond donors and acceptors are required to confer ‘RNA-friendly’ properties to small molecule ligands. After all, hydrogen bond donors are absent in either of the compounds 4 and 5.

**HCV IRES**

Translation of the HCV genomic RNA is driven by an IRES element that recruits ribosomes directly at the viral start codon without the need for most host cell initiation factors. The unique role of the viral IRES and its high conservation in clinical isolates has led to recognition of this ncRNA as a potential drug target. High throughput screening of over 200,000 compounds for inhibition of IRES function in cell-based reporter translation assays led to the discovery of phenazine derivatives and biaryl guanidines which had moderate selectivity for the viral translation target. While the structures of compounds from these chemical series suggest that they may interact with the IRES RNA, target binding or antiviral activity were not reported in these early studies. An important breakthrough was achieved by Seth and colleagues who used mass spectrometry-based screening of 180,000 compounds to identify 2-aminobenzimidazole derivatives as ligands binding the internal loop RNA of subdomain IIa in the HCV IRES (Figure 4(a)). Lead optimization driven by structure-activity relationship data furnished compound 6 (Table 3) which had an affinity of 0.9 μM (K<sub>d</sub>) for the IRES target (determined by mass spectrometry) and inhibited HCV in cell culture with an EC<sub>50</sub> value of 3.9 μM while showing no cytotoxicity up to 100 μM concentration. Mechanism of action studies of 2-aminobenzimidazoles revealed that these compounds act as allosteric inhibitors of an RNA conformational switch in the IRES subdomain IIa. In contrast to metabolite-sensing riboswitches, the HCV IRES subdomain IIa motif is representative of a new class of RNA conformational switches that are structurally well-defined in both ligand-free and bound states and may represent the simplest form of ligand-responsive, purely mechanical switches in nucleic acids.

The synthetic benzimidazoles are fortuitous ligands of a recognition site for guanosine and capture the subdomain IIa target in an extended conformation which blocks viral translation initiation by the IRES element. X-ray crystallography has provided structural insight into the conformational states of the subdomain IIa switch in the absence and presence of inhibitor ligand. In the extended conformation of the switch, the bound inhibitor is encapsulated by a deep solvent-excluded RNA pocket that resembles ligand interaction sites in aptamers and riboswitches (Figure 4(b) and (c)). Depth, structural complexity, and physicochemical properties of the ligand binding pocket distinguish the IRES IIa switch from other viral RNA targets and suggest that discovery of drug-like inhibitors is feasible. A FRET-based assay (Box 2) was developed to investigate ligand binding to subdomain IIa and screen for viral translation inhibitors that capture the RNA switch in an extended state. In a different assay, which monitors fluorescence changes of a nucleoside analog (Box 2) incorporated in the internal loop of subdomain IIa, diaminopiperidines such as compound 9 (Table 3) were found to bind the IRES target as well. Unlike the 2-aminobenzimidazoles, the diaminopiperidine ligands lock the RNA conformational switch in a bent state and thereby inhibit viral translation initiation. A recent review summarizes comprehensively studies on 2-aminobenzimidazole and diaminopiperidine HCV translation inhibitors.

The diaminopiperidines are hydrophilic compounds which contain numerous hydrogen bond donors, which suggests that polar interactions play an important role for RNA target recognition. As a consequence, binding affinity of compounds such as 9 for the subdomain IIa RNA target decreases with ionic strength of the medium. In contrast, 2-aminobenzimidazoles such as inhibitor 6 have more drug-like properties, but contain basic amino groups that are positively charged under physiological conditions. Moreover, construction of the pyran ring in
6 required a lengthy synthetic route that hampered a deep exploration of this chemical series.41,67 We have recently designed second generation ligands for the guanosine binding pocket in the HCV IRES subdomain IIa target which address the shortcomings of the original 2-aminobenzimidazole derivatives. In an approach to maintain the nonplanarity of benzimidazole ligands, which is introduced by the sp² carbon centers in the pyran ring of ligands such as 6, we have designed N1-coupled aryl derivatives represented by compound 7. Sterical hindrance of the N1 aryl substituent induces a twisted conformation with an overall nonplanar shape of the resulting compounds.44 In another set of analogs, we have replaced the basic imidazole ring in the first generation ligands with the less basic oxazole system.45 The resulting ligands, exemplified by compound 8, were accessible through straightforward synthetic routes and displayed good drug-like properties. However, none of the synthesized N1-coupled aryl benzimidazoles44 or oxazole derivatives45 had an affinity for the IIa RNA superior to the original 2-aminobenzimidazoles. The best ligands in the new compound series showed EC₅₀ values for target binding of 74–120 μM.44,45 A computational molecular dynamics (MD) study of 2-aminobenzimidazoles in complex with the IRES IIa RNA inspired the design of tetracyclic dipyrroloindoles which were proposed as rigid, less basic ligand candidates.68 However, synthesis or testing of the in silico designed compounds was not reported. The structural plasticity of the HCV IRES subdomain IIa illustrates the challenges facing computational approaches for RNA targets in general. Large conformational changes in the IIa RNA target upon ligand binding, accompanied by extensive rearrangement of base stacking, were unexpected and likely beyond the predictive ability of current computational approaches. The relative success of MD studies reproducing the conformational rearrangement and benzimidazole ligand binding of the IIa RNA⁶⁸,⁶⁹ may be attributed to the fact that experimental waypoints in the form of high-resolution crystal structures were available to guide the calculations. Challenges of large conformational changes in RNA targets are compounded by the dominant electrostatic interactions in the anionic polymer and concurrent impact of counter-ion binding and hydration,⁷⁰ which are recognized as the Achilles’ heel of computational approaches for RNA.⁷¹

While this review focuses on small molecule ligands with a molecular weight <750 D, brief mention should be made of copper-binding metallopeptides that have been developed to target domains of the HCV IRES and damage the RNA through metal-catalyzed cleavage.⁷²–⁷⁴ These peptides contain a conserved 3-amino acid motif of Gly–Gly–His which coordinates Cu, linked to a recognition sequence that confers RNA target binding specificity. Cu-metallopeptides with 7–27 amino acids have been reported to bind IRES domains and inhibit HCV in cell culture with submicromolar affinity.⁷²–⁷⁴

**SARS CoV Frameshifting PK**

Programmed ribosomal frameshifting is an evolutionary strategy to maximize the coding content of genomes by providing a mechanism that allows translation of overlapping reading frames.⁷⁵ Some RNA viruses, including HIV and the severe acute respiratory syndrome CoV (SARS CoV), regulate the transition of highly expressed structural proteins to viral enzymes expressed at low levels by a −1 frameshift during translation. Translational frameshifting is triggered by the interplay of two RNA motifs which include a slippery sequence where the reading frame change occurs, followed by a structured region that stalls the ribosome. The structured frameshift motif in HIV consists of a long RNA hairpin with an internal loop, which has been explored in ligand targeting approaches that were aimed at stabilizing or disrupting the RNA fold. Earlier studies on the HIV target have been discussed in recent comprehensive reviews by Le Grice, Brakier-Gingras, and coworkers.⁵,³³ In SARS CoV, translation of viral proteins required for replication including the RNA-dependent RNA polymerase is initiated by a −1 programmed frameshift which is triggered by a three-stemmed RNA PK (Figure 5).⁷⁶,⁷⁷

Park and colleagues used an in silico screening approach to identify ligands of the SARS CoV PK.⁴⁷ Since the three-dimensional structure of the target has not been determined yet, docking of compounds was performed to a model of the RNA PK. Among hit compounds identified in the screen, the drug-like 1,4-diazepane derivative 10 (‘MTDB’, Table 4) was reported as an inhibitor of translational frameshifting both in vitro and in a cell-based assay (IC₅₀ = 0.45 μM).⁴⁷ Direct interaction of 10 with the SARS CoV PK was demonstrated recently by Ritchie and coworkers who measured ligand binding to the RNA target by surface plasmon resonance (SPR) and determined a Kₐ of 210 μM.⁵⁰ They went on to perform single-molecule unfolding experiments with the PK in the absence and presence of 10, which led to the conclusion that ligand binding reduces the conformational plasticity of the RNA fold and thereby affects ribosomal frameshifting. Efficiency of frameshifting is not determined by thermodynamic stability.
of the PK or impact on ribosomal pausing but rather by the ability of the RNA fold to adopt alternate conformations and structures. Therefore, ligand binding may affect translational frameshifting through a complex mechanism that only partially relies on stabilization of the PK RNA. For some of the previously identified inhibitors of HIV translational −1 frameshifting, it has been suggested that their promiscuous RNA-binding ability may affect ribosomal RNA rather than the viral genomic frameshifting signal. Interaction with multiple targets may play a role for the inhibitory action of 10 as well, which may account for the over 450-fold higher potency of this compound as a frameshifting inhibitor in a cell-based assay compared to its RNA binding affinity.

Influenza A Virus RNA Promoter

The influenza A orthomyxovirus harbors a single-stranded RNA genome that contains eight negative-sense segments (vRNA) which encode the viral proteins and are used as templates for transcription to mRNA and replication into complementary strands (cRNA). The viral RNA-dependent RNA polymerase recognizes a partial duplex structure that forms through circularization of vRNA by hybridization between conserved nucleotides at the 5' and 3' end of each segment. The resulting RNA duplex (Figure 6(a)) is stable under physiological conditions and serves as a promoter for transcription and replication. Investigations of the RNA promoter structure by NMR revealed a bifurcated interaction of a uracil base hydrogen bonding with two consecutive adenine residues in the opposite strand. This motif in conjunction with a neighboring noncanonical AoC base pair induces widening of the major groove of the RNA promoter helix near the polymerase initiation site. Since unique structural features introduced by internal loops and noncanonical base pairs have previously been shown to provide ligand recognition sites in duplex RNA, it is conceivable that selective compound binding to the influenza A RNA promoter might interfere with polymerase interaction and prevent viral replication. Inhibitors that interfere with viral gene expression by binding to the RNA promoter would provide a novel mechanism of action for the development of anti-influenza drugs. In an earlier study, Choi and coworkers used a competition assay with a fluorescently labeled binder (Box 2) to test aminoglycoside binding at the influenza A RNA promoter. Affinity in the low micromolar range (Kd = 2.7–33 μM) was reported for various aminoglycosides but the impact of ligand binding on the promoter function was not investigated.

More recently, Varani and coworkers performed an NMR-based fragment binding screen of 4279 compounds against an oligonucleotide model of the influenza A RNA promoter (Figure 6(a)). Among seven initial hit compounds which led to changes in the RNA imino proton signals observed by NMR, the drug-like amino-quinazoline ‘DPQ’ (Table 5) had the highest affinity at a Kd of 50–60 μM. NMR spectroscopy was further used to obtain a model of the promoter in complex with (Figure 6(b) and (c)). Ligand binding is observed in a widened RNA major groove at the internal loop formed by the bifurcated U < A/A motif. The NMR model of the promoter complex suggests that interacts with the RNA target mainly by contacts of the methoxy substituents which are positioned closely to residues of the internal loop motif (Figure 6(c)). The researchers point out that methoxy group protons in are within hydrogen bonding distance to A11, A12, G13, and C22. Hydrogen bond formation involving carbon-attached donors interacting with oxygen (CH−O) and nitrogen (CH−N) acceptors is unusual but has precedent in biomolecular structures. While such interactions are weak and may be rare in small molecule recognition of RNA folds, CH−O hydrogen bonds involving an aryl-methoxy donor have been observed contributing to binding of a 6-O-methylguanine ligand in crystal structures of purine riboswitches. For several natural product antibiotics that contain aryl-methoxy groups, including anisomycin and puromycin, crystal structures of ribosome-bound complexes are available. However, the diffraction resolution of the data used to obtain these structures limit conclusions on potential hydrogen bond interactions of the methoxy groups. Emetine, an antiprotozoal drug that binds to the eukaryotic ribosome, contains aryl-o-dimethoxy-scaffolds comparable to the o-dimethoxy-quinazoline in . A recent cryo-EM structure of a Plasmodium 80S ribosome in complex with emetine shows the aryl-methoxy groups solvent-exposed and oriented away from the RNA, however, with a limited resolution at 3.2 Å.

A cell-based assay that measured cytopathic effect inhibition demonstrated antiviral activity of the RNA promoter-targeting compound against the H1N1 and H3N2 strains of influenza A as well as influenza B with EC50 values of 72, 275, and 114 μM, respectively. In a cell-based direct viral replication assay, which monitored luciferase reporter expression from a modified H1N1 influenza A virus, compound inhibited replication at an EC50 value.
of 435 μM (reported as 549 μM in a later publication). At the same time, cytotoxicity was not detected at compound concentrations over 500 μM. In a follow-up study, Bottini and colleagues synthesized and tested 16 analogs of 11. Among the second-generation amino-quinazoline analogs, a set of four structurally related compounds (12a,b,c,d; Table 5) stood out for their improved binding affinity to the influenza A RNA promoter target ($K_d = 34–44 \mu M$ by NMR) and activity as inhibitors of viral replication ($IC_{50} = 44–158 \mu M$). The authors note that a possible direct inhibition of the viral RNA-dependent RNA polymerase by the amino-quinazoline derivatives cannot be ruled out and has yet to be investigated. None of these compounds showed cytotoxicity up to the maximum tested concentration of 250 μM.

**CONCLUSION**

Viruses have compact, highly streamlined genomes which provide only a limited number of protein targets for therapeutic intervention in viral infections. The repertoire of drug targets may be expanded by targeting structured RNA elements in the genomes and transcripts of viral pathogens. In this review, I have discussed recent progress in the discovery of small molecule-targeting approaches directed at structured RNA motifs in the genomes of HIV, HCV, SARS CoV, and influenza A, which represent a variety of virus families. Regulatory sequences in viral genomes and transcripts stand out for their high barrier to resistance development against inhibitors targeting these structured RNA elements. Nevertheless, criteria of drugability have to be thoughtfully applied to select RNA targets for small molecules in the development of novel antiviral therapies. Such considerations are more important than ever in light of an unprecedented surge of new ncRNA species emerging as key participants in biological processes. Structurally well-defined, deep and solvent-excluded ligand binding pockets are rare in RNA folds but provide the most promising opportunities for discovery of drug-like inhibitors on par with drugs targeting viral proteins. Among the viral RNA motifs discussed here, the HCV IRES subdomain IIa perhaps best embodies such advantageous characteristics of a drug target, but is also unique with respect to its function as a conformational switch that harbors a proper ligand binding pocket. Similar RNA switch motifs have been discovered in over 10 other flavi- and picornaviruses, however, none of which rise to a level of threat as a human pathogen comparable to HCV. Outside viruses, comparable deeply encapsulating ligand binding pockets in RNA are found in bacterial riboswitches which are being exploited as targets for antibiotics discovery.

Lessons from efforts of ligand discovery for structured RNA elements, including regulatory motifs in the viruses discussed here, bacterial rRNA, riboswitches, and other ncRNA may inspire the future search for inhibitors targeting RNA motifs in a wider range of human-pathogenic viruses. Table 6 lists a selection of viruses with a high human disease burden for which structured RNA elements have been shown to play a role during infection.

Insect-borne flaviviruses, including Dengue, West Nile, and Zika viruses, harbor multiple-structured RNA motifs which are involved in replication, translational control, and host adaptation.

**TABLE 6** | Viruses Carrying Potential RNA Targets for Small Molecule Ligands

<table>
<thead>
<tr>
<th>Family</th>
<th>Examples</th>
<th>Genome</th>
<th>Potential RNA Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insect-borne flavivirus (arbovirus)</td>
<td>Dengue (DENV), West Nile (WNV), Yellow fever (YFV), Zika (ZIKV), Tick-borne encephalitis (TBEV)</td>
<td>(+)ssRNA</td>
<td>5' UTR (including RNA promoter in stem-loop A, SLA; RNA long-range interacting stem-loop B, SLB)(^{91,92}) Structured elements in the coding region (including capsid coding region hairpin, chp; pseudoknot C1)(^{92}) 3' UTR (including RNA long-range interacting structures)(^{91,92,94,95}) 3' UTR-derived ncRNA (including subgenomic flavivirus RNA, sfRNA)(^{96,97}) RNA promoter for the viral RNA-dependent RNA polymerase (RdRp)(^{98–100}) Structured intergenic regions (IGR) of the viral genome(^{100,101}) 5' and 3' UTR in viral transcripts(^{102,103}) Polyadenylated nuclear (PAN) noncoding RNA(^{107})</td>
</tr>
<tr>
<td>Filovirus</td>
<td>Ebola (EBOV), Marburg (MARV)</td>
<td>(−)ssRNA</td>
<td></td>
</tr>
<tr>
<td>Herpesvirus</td>
<td>Kaposi’s sarcoma associated herpesvirus (KSHV)</td>
<td>dsDNA</td>
<td>IRES in the transcript for the viral homolog of the FLICE inhibitory protein (vFLIP)(^{104–106})</td>
</tr>
<tr>
<td>Hepadnavirus</td>
<td>Hepatitis B (HBV)</td>
<td>ds/ssDNA</td>
<td>Encapsulation signal epsilon of viral pregenomic RNA (pgRNA)(^{108})</td>
</tr>
</tbody>
</table>

© 2016 Wiley Periodicals, Inc.
Several of the flavivirus RNA elements have features of conservation and structural complexity comparable to the viral RNA motifs discussed in this review and may be tractable targets for small molecule ligand discovery. A variety of potential drug targets in structured RNA elements is also found in filoviruses, which include the etiologic agents of Ebola and Marburg hemorrhagic fever. Replication of these negative-sense RNA viruses relies on a structured RNA promoter element that consists of a large hairpin involving nucleotides 1–55 in the 5′ terminus of the viral genome whose secondary structure has been established by chemical probing. The 3′ terminus of the filovirus genome does not interact with the 5′ end, unlike in the RNA promoter of influenza A virus which also has a negative-sense ssRNA genome. Filovirus genomes consist of a nonsegmented RNA that encodes reading frames for seven structural proteins which are separated by long intergenic regions (IGR). The IGR between virus proteins VP30 and VP24 has been suggested to fold into a two-armed stem-loop structure that includes a complex RNA four-way junction. Transcription of viral genes involves the VP30 nucleocapsid protein as an anti-termination factor that binds to a hairpin loop at the transcription start site. Viral transcripts have long 5′ UTR elements that derive from the IGR and contain hairpin loop structures which may play a role in transcriptional regulation and translational control. Examples of attractive RNA targets for ligand inhibition are also found in DNA viruses. Among them is the oncogenic Kaposis sarcoma associated herpesvirus (KSHV) which is a common cause of malignancies in AIDS patients. Translation of the viral homolog of the FLICE inhibitory protein (vFLIP), a key player in KSHV-induced tumorigenesis, is initiated at an IRES element that adopts a complex RNA secondary structure including a conserved segment of 252 nucleotides. Functional and chemical probing analysis of the vFLIP IRES revealed an independently folding RNA core domain linked to flexibly linked hairpin motifs which provide several structurally well-defined sites for ligand targeting. KSHV also encodes a 1077 nucleotide polyadenylated nuclear (PAN) ncRNA, which is the most abundant transcript produced from the viral genome during lytic replication. PAN RNA is essential for viral propagation, and its accumulation relies on posttranscriptional stabilization involving a cis-acting RNA motif which acts as an enhancer of nuclear retention element (ENE). The ENE is a 79 nucleotide motif in the PAN 3′ terminus which sequesters in cis the PAN poly(A) tail in a unique triple helix structure that protects the ncRNA from decay and leads to accumulation of PAN. Both, the ENE hairpin structure, which contains a uridine-rich, large internal loop, and the ENE-poly(A) triple helix complex are potential targets for small molecule ligands that may interfere with KSHV replication. One final example for a tractable RNA target in a DNA virus is the pregenomic RNA (pgRNA) of Hepatitis B virus (HBV) which serves as an intermediate template in viral replication through reverse transcription. Prior to reverse transcription, the HBV pgRNA is sequestered together with polymerase into subviral particles. Both encapsidation and initiation of reverse transcriptase require a conserved sequence in the 5′ terminal region of the pgRNA which folds into a stem loop structure with a uridine-rich internal loop referred to as the epsilon encapsidation signal. While small molecule ligands of this RNA motif have not been reported yet, RNA decoys of the epsilon sequence have been developed that sequester reverse transcriptase, which provided proof-of-principle that disruption of the pgRNA-polymerase interaction suppresses HBV replication.

In conclusion, while preventive vaccines and drugs that target well-characterized viral proteins such as polymerases and proteases are proven routes to antiviral therapy, structured RNA motifs in viral genomes and transcripts provide new opportunities to expand the repertoire of targets for the development of antinfective therapy. Selection of viral structured RNA elements for inhibitor discovery requires critical evaluation for properties that define targets for drug-like ligands. After all, just as some proteins that participate as key players in disease processes have been unyielding to the development of small molecule inhibitors, not all structured RNA elements are amenable to targeting with low molecular weight ligands. If the currently known target space for biologically active small molecules from natural product sources is of any guidance, RNA targets certainly do not appear among the low hanging fruits for drug discovery. Nevertheless, the area of antinfective drugs provides a strong record for successful examples of RNA-directed inhibitors. The bacterial ribosome is the target for over half of all natural products that are known to exhibit antibiotic activity and the majority of these compounds interact predominantly with the rRNA components. While the bacterial ribosome looms large as a paradigm of structured RNA providing a valuable target for antinfective drugs, recently, a bacterial riboswitch RNA...
has been identified in an unbiased phenotypic screen as the target for a new class of antibiotic lead compounds.\textsuperscript{119} This wonderful example of modern antibiotic discovery promises an exciting future for RNA targets emerging in the exploration of novel therapies for bacterial and viral infections.

REFERENCES

28. Park WKC, Auer M, Jaksche H, Wong C-H. Rapid combinatorial synthesis of aminoglycoside antibiotic mimetics: use of a polyethylene glycol-linked amine and a neamine-derived aldehyde in multiple component condensation as a strategy for the discovery of


