JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Cite This: J. Am. Chem. Soc. 2019, 141, 6776−6790
Perspective
pubs.acs.org/JACS

Targeting RNA with Small Molecules To Capture Opportunities at the Intersection of Chemistry, Biology, and Medicine

Matthew D. Disney*

Department of Chemistry, The Scripps Research Institute, Jupiter, Florida 33458, United States

ABSTRACT: The biology of healthy and disease-affected cells is often mediated by RNA structures, desirable targets for small molecule chemical probes and lead medicines. Although structured regions are found throughout the transcriptome, some even with demonstrated functionality, human RNAs are considered recalcitrant to small molecule targeting. However, targeting structured regions with small molecules provides an important alternative to oligonucleotides that target sequence. In this Perspective, we describe challenges and progress in developing small molecules interacting with RNA (SMIRNAs) to capture their significant opportunities at the intersection of chemistry, biology, and medicine. Key to establishing a new paradigm in chemical biology and medicine is the development of methods to obtain, preferably by design, bioactive compounds that modulate RNA targets and companion methods that validate their direct effects in cells and pre-clinical models. While difficult, demonstration of direct target engagement in the complex cellular milieu, along with methods to establish modes of action, is required to push this field forward. We also describe frameworks for accelerated advancements in this burgeoning area, their implications, key new technologies for development of SMIRNAs, and milestones that have led to broader acceptance of RNA as a small molecule druggable target.

1. INTRODUCTION

The First Insights into RNA Structure−Function and Its Interaction with Small Molecules. The first nucleic acid was sequenced by Robert Holley in the late 1950s, a non-coding transfer (t)RNA.[1] From this first sequence and subsequent analysis of multiple tRNAs, it was clear that they had stable structures that played essential roles in cellular function (Figure 1, top). The precise interaction of a tRNA's anticodon with a messenger (m)RNA's codon in the context of the ribosome decoded all mRNAs into protein and helped to establish the central dogma. RNA structure was a key to solving this puzzle.

Small molecule−RNA interactions also feature prominently in the early days of drug discovery. Sal Waksman and colleagues investigated compounds produced by soil microbes and how they were used as a competitive advantage. It was known that *Mycobacterium tuberculosis* was destroyed when it was transferred to soil. Thus, the hypothesis that microbes in the soil produced a substance that affected tuberculosis growth was born. Waksman's studies culminated in the isolation and identification of chemical substances of medicinal importance, including the antibiotics actinomycin, streptomycin, and neomycin, known to be effective against previously untreatable tuberculosis infections.[2] Ironically, these life-changing medicines would today be classified by some as "undrug-like". Not long after streptomycin's first clinical use, resistant strains of the virus emerged. Mapping mutations in these resistant strains became a powerful way to decipher a compound's mode of action and inform design of compounds that evade resistance. The first mutations that conferred resistance were observed in ribosomal (r)RNA and proteins associated with the ribosomal particle, particularly S12.[3] Mutations in the rRNA sequence suggested that actinomycin and aminoglycosides bound rRNA and inhibited protein synthesis.

After Holley's and Waksman's studies, Carl Woese and Harry Noller began intensive studies on the sequences and secondary structures of ribosomes from diverse organisms.[4] Their studies revealed that ribosomes had extensive structure and that organisms can be classified on the basis of rRNA sequence into three distinct domains: bacteria, archaea, and eukaryotes. The finding that archaea and eukaryotes are sister groups in the "tree of life" highlighted the importance of RNA for deciphering phylogenetic relationships and transformed our view of evolutionary history.

Using information on the rRNA sequence and chemical footprinting methods developed simultaneously by Cech[5] and Noller,[6] Noller and Moazed showed that many antibacterials directly interact with rRNA, identifying the precise binding sites within the ribosome (Figure 1, bottom).[7] Additional footprinting experiments completed by Puglisi and colleagues suggested that aminoglycosides affected the target's decoding site and the codon−anticodon interaction, stabilizing mispaired interactions and affecting protein fidelity.[8] Interestingly, the results of these footprinting experiments were confirmed by crystal structures reported by Ramakrishnan, Noller, Cate, Yonath, Moore, and Steitz, which also established new aminoglycoside binding pockets (Figure 1, bottom).[9] Thus, the binding of small molecules to RNA motifs in the context of the ribosome could dramatically affect this macromolecular machine. These studies clearly showed not only that RNA could be a target of small molecules but also that it was a target of a historically important class of life-saving medicines. Perhaps no other field has been impacted by small molecule chemical probes more than ribosome biology, due in part to its complex dynamics and the necessity to lock it into specific structures with small molecules to allow their study.

Received: December 16, 2018
Published: March 21, 2019

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

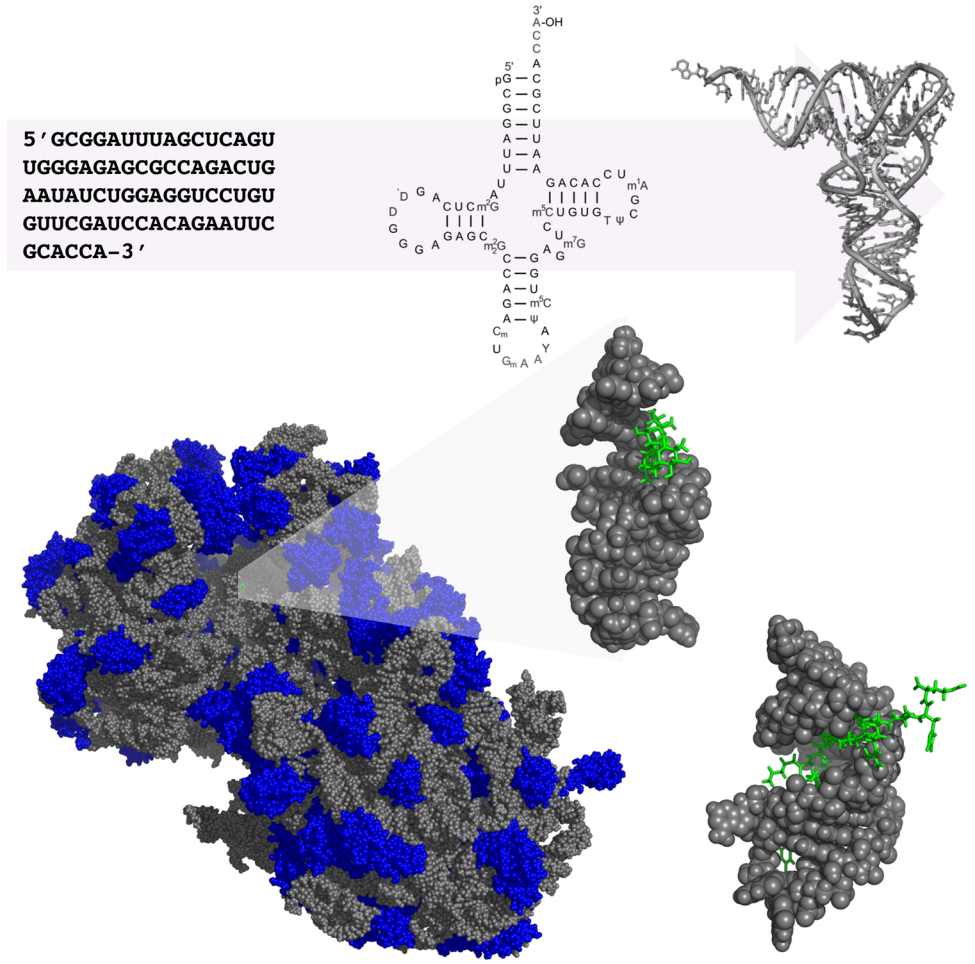

Figure 1. RNA structural hierarchy and examples of the 3D structures of small molecules bound to RNA. (Top) Sequence, secondary structure, and three-dimensional structure of tRNA. (Bottom) Left: structure of the bacterial ribosome (Protein Data Bank (PDB) ID 4V52) with protein in blue, RNA in dark gray, and neomycin B in green. Middle: binding of neomycin B to the A-site of the bacterial ribosome extracted from the structure on the left. Right: structure of a cyclic peptide bound to the HIV TAR RNA that was developed via structure-based design (PDB ID 2KDQ).

2. RNA STRUCTURE DICTATES FUNCTION THROUGHOUT BIOLOGY

The discovery of RNA catalysis by Cech (group I Introns)[10] and Altman (ribonuclease P)[11] ushered in the modern RNA World era, in which RNAs do not just encode or produce protein but have diverse functions that are dictated by their structures. The development of nucleic acid catalysts and directed evolution technologies to design aptamers[12] also suggested that RNA folds broadly control function.

In the early 2000s, a major class of RNAs were discovered that changed conformation in the presence of a small molecule to act as a genetic switch, or riboswitch.[13] Present within bacterial biosynthetic genes, riboswitches bind a metabolite, causing a conformational change that turns off or reduces translation of proteins in the biosynthetic pathway. Fortuitously, structural mimicry of the small molecule metabolite can afford antibacterials.[14] Interestingly, it was later found that the clinically used antibacterial Roseflavin[15] acted, in part, by binding to the flavin mononucleotide (FMN) riboswitch.[16] Roseflavin's mode of action was uncovered when drug-induced resistance mutations were found in the sequence encoding the riboswitch,[17] establishing that riboswitches are indeed druggable.

3. HIV TAR RNA

In the 1980s, human immunodeficiency virus (HIV) became a serious health concern. Once the virus itself was identified, an intense search to find targets for therapeutic intervention commenced. Two RNAs were identified that are critical for replication: trans-activation response (TAR) element and Rev responsive element (RRE). The interaction of TAR with the trans-activator of transcription (Tat) protein activates the long terminal repeat promoter and hence viral replication. More recently, TAR was found to produce microRNAs (miRNAs) that inhibit host cell apoptosis.[18] HIV RRE's structure is recognized by the accessory trans-activating protein Rev, allowing for nucleocytoplasmic export of the RNA. In both cases, drug discovery efforts have focused on inhibiting the interactions of TAR and RRE with their cognate proteins. Indeed, one of the first small molecule high-throughput screens against an RNA target identified inhibitors of the Tat-TAR interaction, affording the aminoglycoside neomycin B, which was also active in a cellular model.[19] A series of mutational studies first revealed that a stem-loop secondary structure was required for TAR RNA recognition by Tat, which was further refined to the bulge and loop regions of the RNA and an arginine-rich sequence in Tat. Biophysical studies revealed the

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

dynamic nature of TAR RNA, particularly the bulge interconverting between structures, clearly evident in nuclear magnetic resonance (NMR) spectroscopy studies by Puglisi and Williamson.[20]

One strategy to inhibit TAR function was to lock it into a single conformation, which fortuitously also enabled its biophysical study. Argininamide was one of the first compounds developed for this purpose.[21] Other studies soon followed, including the development of peptides and proteins that inhibit TAR.[22] Guided by structures of the Tat-TAR complex, Varani and Robinson showed that extremely high affinity ligands could be designed that inhibit viral replication (Figure 1, bottom).[23] Interestingly, a small molecule docking approach that accounted for TAR RNA's conformational flexibility identified new ligands that bind TAR and elucidated how ligands can steer conformational dynamics.[24] These studies, first realized in HIV TAR, have had large implications in targeting other RNAs, as ligand-induced dynamics is a major factor in effective small molecule binding. Collectively, the studies on the ribosome, riboswitches, and HIV Tar RNA showed that small organic ligands can affect RNAs involved in infectious disease by targeting RNA structures.

4. PHENOTYPIC SCREENS TO DRUG RNA: YOU FIND WHAT YOU ARE LOOKING FOR

One of the first phenotypic screens for an RNA target was completed by Frankel and co-workers to identify new peptides that could replace HIV Rev.[25] An RRE RNA was placed upstream of the β-galactosidase gene and introduced into *Escherichia coli* along with a plasmid encoding a peptide combinatorial library. Small peptides able to bind RRE and facilitate export could easily be deduced by blue/white colony screening.[25] These studies identified various peptides that bound RRE more avidly than Rev and elucidated various factors that were important for peptides to interact with RNA, including that α-helical peptides are optimal.[25]

The discovery that part of Roseflavin's antibacterial mode of action was binding to a riboswitch precipitated the search for new antibiotics with similar modes of action. Because riboswitches were found to be druggable targets, Howe et al. at Merck completed a phenotypic screen to identify inhibitors of the FMN riboswitch that were more selective than Roseflavin,[26] which is known to inhibit various bacterial flavoenzymes and is potentially a source of side effects, as some flavoenzymes are conserved in humans.[27,28] Merck's phenotypic screen comprised a *E. coli* strain in which the riboflavin biosynthetic pathways is conditionally essential.[26] That is, in the presence of exogenous riboflavin, the strain grows similarly to wild type, but its growth in riboflavin's absence is significantly slowed (attenuated virulence in mice). Inhibitors of FMN biosynthesis therefore slow growth in the presence of riboflavin. These studies afforded the compound Ribocil, and its on-target effects were confirmed by sequencing of Ribocil-resistant mutants.[26]

Phenotypic screens have identified novel small molecules for the treatment of spinal muscular atrophy (SMA). SMA is caused by loss of survival motor neuron 1 (SMN1) protein and hence loss of motor neurons and control of muscle movement. Fortuitously, humans have a second *SMN* gene, *SMN2*, that can replace the loss of SMN1. The *SMN1* and *SMN2* genes differ by only a single nucleotide; however, this change affects SMN alternative pre-mRNA splicing. In the case of *SMN2*, it causes exclusion of an intron, which decreases the half-life of the resulting protein. Thus, one therapeutic strategy is to alter the splicing of *SMN2* to include the skipped exon more frequently and produce a longer-lasting protein that replaces SMN1.

The phenotypic screen designed by PTC Therapeutics/Roche coupled inclusion of exon 7 to the production of luciferase in HEK293H cells.[29] The screen identified several chemical classes, which were lead optimized and studied for oral bioavailability, affording three compounds (SMN-C1, SMN-C2, and SMN-C3). All three compounds increased levels of the exon 7-containing transcript and the corresponding protein in SMA patient-derived cells at nanomolar concentrations. RNA-seq analysis revealed very few off-targets for any of the compounds, as defined by changes in abundance by a factor >2 in patient-derived cells (seven and six out of 11 725 transcripts for SMN-C1 and SMN-C3, respectively). Treatment of mouse models with mild and severe SMA with SMN-C2 or SMN-C3 increased the exon 7-containing transcript and SMN protein levels. In models of severe SMA, SMN-C2 and SMN-C3 treatments extended lifespan, improved locomotor activity, and prevented spinal cord motoneuron loss and muscular atrophy.[29]

Shortly thereafter, Novartis developed a similar luciferase-based reporter of *SMN2* exon 7 splicing in a motor neuron cell line.[30] Of the scaffolds identified that modulated splicing toward exon 7 inclusion, one that contained a pyridazine core was further explored, affording two compounds with nanomolar activity in mouse myoblasts and SMA patient-derived cells, NVS-SM1 and NVC-SM2. Owing to its more favorable pharmacokinetic properties, the activity of NVS-SM1 was confirmed in a panel of induced pluripotent stem cell (iPSC)-derived neurons, including those from SMA patients. Very few changes in the transcriptome were induced by compound treatment, with little to no effect on splicing factors or other RNA-binding proteins. NVS-C1, orally delivered to a mouse model, increased SMN protein levels in the brain in a dose-dependent fashion, increased body weight, and extended lifespan. An investigation into the mode of action indicated that NVS-C1 may form a ternary complex with *SMN2* exon 7's 5′ splice site and U1 small nuclear ribonucleoprotein (snRNP), stabilizing the interaction and facilitating exon 7 inclusion.[30]

Collectively, these examples demonstrate that phenotypic screens can indeed identify drugs that target RNA, particularly those with significant tertiary or even quaternary structure (the complex formed between SMN2 exon 7 and U1 snRNP), and that RNAs are potential drug targets. One of the reasons for RNA's rejection as a small molecule drug target has been a lack of tools to identify and validate RNA targets. In the case of Ribocil, resistance mutants defined the compound's mode of action. However, when resistance does not arise rapidly or when resistance is due to non-target-directed mutations (i.e., upregulation of a multidrug resistance pump), identifying the compound's mode of action is a larger challenge. That is, the target-agnostic nature of phenotypic screens requires tools to validate which cellular targets they directly engage to elicit the desired biological effect. Until recently, these tools had not yet been developed for RNA (described below), although many are available for proteins.

5. INTENTIONALLY TARGETING HUMAN RNAs WITH SMALL MOLECULES

Genome-wide association studies (GWAS) in concert with antisense oligonucleotides (ASOs) have enabled the identi-

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

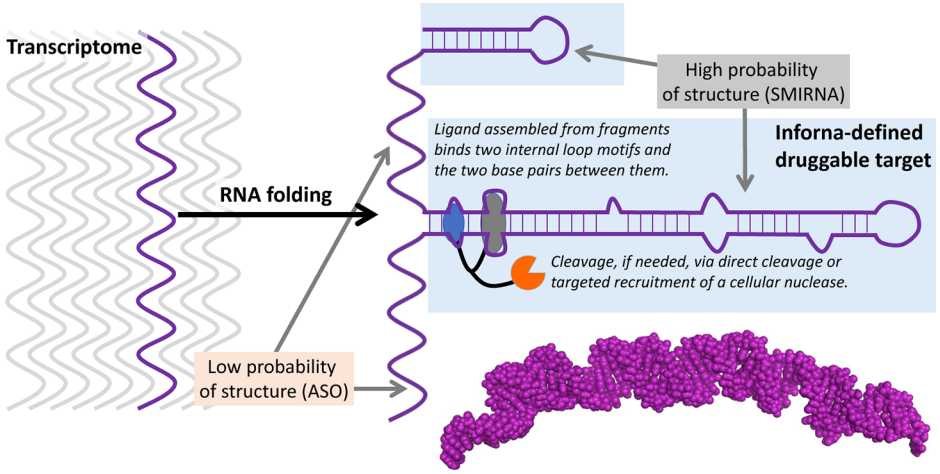

Figure 2. Identifying and drugging targetable RNA structures in the human transcriptome. Diverse RNA sequences in the human transcriptome, or the composite of RNAs made from an organism, fold into defined three-dimensional structures. Indeed, human RNAs have hubs of defined structure, some of which are evolutionarily conserved and likely functional. These regions are ideal targets for small molecules (SMIRNAs), complementary to unstructured regions targeted with oligonucleotide-based modalities. We describe the design of bioactive SMIRNAs from sequence and how these SMIRNAs have enabled target validation tools.

fication of human disease-causing RNAs. The challenge is to then intentionally target those RNAs with drug-like small molecules. High-throughput screening (HTS) has been employed against various disease-causing human targets, including miRNAs and RNA repeat expansions. MiRNAs are an interesting class of potential drug targets, as their aberrant expression or mutation has been linked to many diseases. MiRNAs, produced after two processing steps of the primary transcript, repress the amount of protein that is produced from a given mRNA via cleavage or translational repression by binding to a complementary site on the mRNA's 3′ untranslated region (UTR). Compounds have been identified from HTS campaigns that inhibit processing of a desired miRNA; however, sufficient selectivity and potency have been difficult to achieve to provide useful probes. The best hits often have, at best, high micromolar activity and limited selectivity.

RNA repeating transcripts cause or contribute to >40 microsatellite diseases,[31] none of which has a known cure. One of these diseases, myotonic dystrophy type 1 (DM1), is caused by an expanded RNA repeat in the 3′ UTR of the dystrophia myotonica protein kinase (*DMPK*) mRNA. The expanded repeat, r(CUG)^{exp}, binds to and sequesters proteins (e.g., muscleblind-like 1, MBNL1) that control pre-mRNA splicing. Sequestration of these proteins causes system-wide defects in patients, such as muscle atrophy and myotonia.[31] Various screens have been completed to identify compounds that inhibit the formation of the r(CUG)^{exp}-MBNL1 complex. For example, the NIH Molecular Libraries Initiative screened 250K compounds, affording modestly potent inhibitors.[32] Serendipitously, a potent dimeric compound emerged from this screen which formed when the small molecule was stored for prolonged periods of time.[33]

In the two cases described above, a compound could be deemed a hit if it binds to the intended RNA target or to the protein off-target (a miRNA processing enzyme or MBNL1). Considering that many compound collections are built around protein targets, a rigorous analysis of compound binding to both the RNA and the protein is essential. As mentioned, such considerations are important for phenotypic screens, which obviously makes it more difficult to construct hypotheses for potential binding partners.

Chemical probes are important in the development of small molecule RNA-directed medicines. A major area of debate is whether traditional small molecule screening against human RNA targets yields useful lead compounds, begging the question if these libraries have high-quality leads for RNA. That is, are there chemical biases in small molecules that enable them to bind RNA and are they present in small molecule libraries? Perhaps RNA has been considered undruggable because RNA-binding compounds are not present in compound collections.

6. PURPOSEFUL DESIGN OF SMALL MOLECULES THAT TARGET RNA

Above, we touched on two methods to purposefully target RNA: structure-based design of small molecules that target HIV RNAs (viral) and structural mimicry of metabolites to target riboswitches (bacterial). Given that RNA folds into composites of structures including base-paired and non-canonically paired regions with defined three dimensional (3D) structures, a broad route to target RNA with small molecules would be to define compound modules or fragments that selectively bind RNA motifs (Figure 2). Once identified from a GWAS, the RNA's structure could be experimentally determined.[34,35] The motifs present in the RNA target could then be paired to a small molecule module. Small molecules emerging from traditional screening against RNA targets typically have modest, at best, affinities and limited, if any, bioactivity. However, if privileged RNA motif–small molecule pairs were defined, then methods could be developed that are scalable and general to develop chemical probes that target RNA quickly by using modular design or fragment-based assembly. Such an information set could broadly inform small molecules that can target and affect RNA biology.

Sequence-Based Design. We have developed an approach that uses RNA sequence to enable design of lead small molecules that selectively target RNA in cells and pre-clinical animal models (Figure 2). The approach, dubbed Infora, uses the RNA's structure and a database of

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776–6790

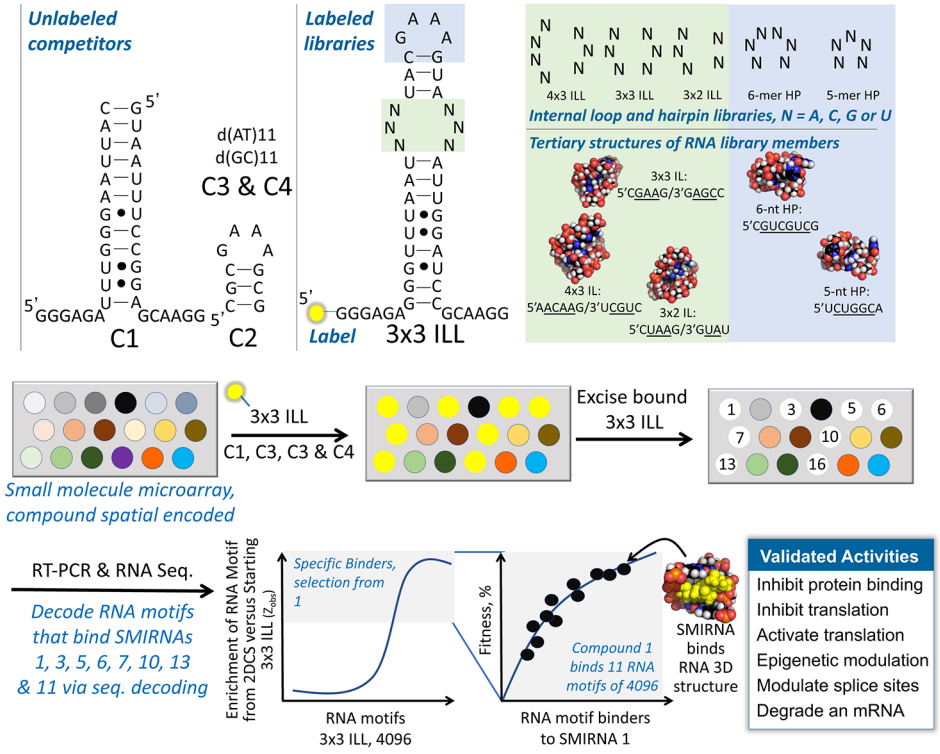

Figure 3. A library-versus-library screen, dubbed two-dimensional combinatorial screening (2DCS), defines interactions between three-dimensionally folded RNA structures and small molecules. (Top) 2DCS is completed with a library of RNA motifs that are embedded in a unimolecular, and thus amplifiable, hairpin structure. The cassette is general, as internal loop (3×3 and 4×3), bulge (3×2), hairpin (5-mer and 6-mer) and other RNA fold libraries can be studied using this approach. RNA libraries are labeled and screened for binding to small molecules in the presence of a large excess of RNAs that mimic the constant regions in the library (C1 and C2) and DNA oligonucleotides (C3 and C4). (Bottom) A microarray with an agarose surface provides a medium to spatially array and encode small molecules that can be studied for binding to RNA folds, for example by incubation with labeled 3×3 ILL in the presence of C1−C4. The three-dimensional RNA folds that bind small molecules are excised from the array and sequenced. Bioinformatic analysis is used to score the selected interactions. Briefly, by sequencing the starting library to define sequencing biases and comparing the RNA fold frequencies to those of the RNA folds selected to bind a small molecule, binding landscapes are quickly defined (via Z_{obs}). Those binders are then assigned a fitness based on the highest affinity interaction identified. These 2DCS-defined SMIRNA partners are then used to design small molecules targeting RNA and have a variety of validated activities. Atomic coordinates for the 4×3 IL, 3×3 IL, and 3×2 IL were obtained from RCSB PDB IDs 1JO7, 1HWQ, and 2LU0, respectively. Atomic coordinates for the 6-nt HP and 5-nt HP were obtained from RCSB PDB IDs 1HWQ and 2B7G, respectively.

experimentally determined small molecule−RNA partners as input.[36,37] These privileged interactions are the output of two methods: two-dimensional combinatorial screening (2DCS), a selection-based method that screens libraries of discrete RNA motifs against libraries of small molecules,[38−40] and the statistical analyses of these selections by High Throughput Structure-Activity Relationships Through Sequencing (HiT-StARTS) (Figure 3).[41]

Identifying Privileged RNA Motif−Small Molecule Interactions via 2DCS and HiT-STARTS. In 2DCS, small molecule microarrays are incubated with libraries of RNA motifs with unique 3D structures (Figure 3, top). These RNA motif libraries are modular, in that each type of motif, whether a loop, bulge, or hairpin, is displayed in the same cassette. Because of the modularity of these libraries, a wide range of RNA folds can be studied for binding ligands simultaneously. Labeled RNA motif libraries are incubated with a small molecule microarray in the presence of excess and unlabeled competitor oligonucleotides that mimic regions constant to all library members, thereby restricting binding interactions to the randomized region (Figure 3). Bound RNAs are excised from the surface, amplified, and sequenced by RNA-seq (Figure 3, bottom).

The RNA-seq data from 2DCS selections are compared to data generated from RNA-seq analysis of the starting library. The highest affinity RNA fold that binds a given small molecule is the one with the most statistically significant enrichment in the selected structures. By measuring the affinities of a small molecule for differentially enriched RNAs, a scoring function can be defined to predict the affinities of all RNA library members (Figure 3, bottom).[41,42]

These studies have defined the RNA 3D folds that bind small molecules, which comprise the Inforna database. By comparing the binding landscapes of many small molecules, we can gain insight into selectivity. For example, does a given RNA structure bind many (promiscuous) or only a few compounds? Does a compound bind many RNAs (promiscuous) or just a few (selective)? These studies also define features in small molecules that bind RNA and functional groups that can read out and bind to the three-dimensional presentation of functional groups within the RNA motif's major and minor grooves.[39] They have also shown that FDA-approved drugs, particularly kinase inhibitors, bind RNA avidly.[40]

Applying Inforna To Target Human RNAs. It is clear that human RNAs have defined hubs of 3D structure.[43,44]

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

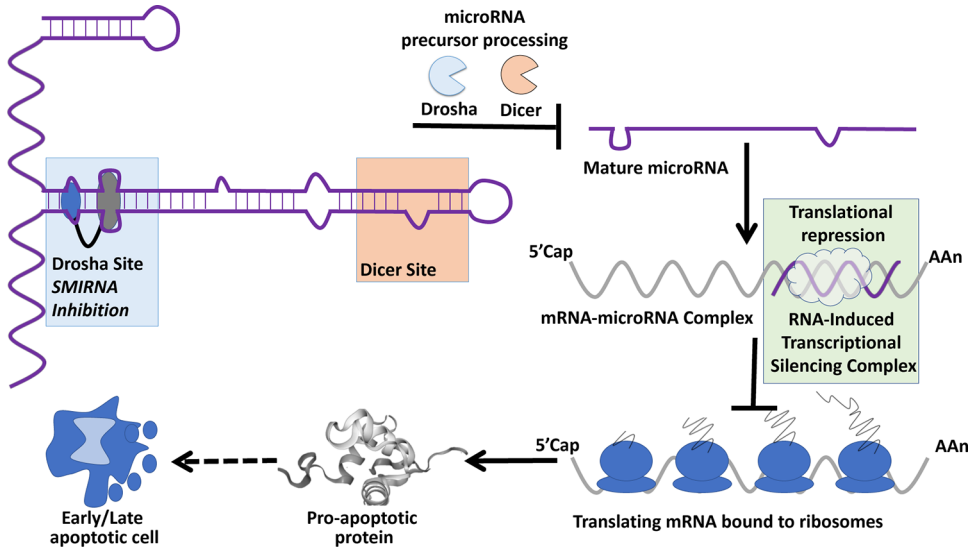

Figure 4. Non-coding microRNAs play pervasive roles in biology to repress the amount of protein translated from an mRNA. Like other cellular RNAs, microRNAs (miRNAs) are transcribed as primary transcripts (pri-miRNA) that undergo various processing steps. Indeed, pri-miRNAs are cleaved by the nuclease Drosha to liberate a precursor miRNA (pre-miRNA) that is translocated to the cytoplasm and further processed by Dicer to generate mature (functional) miRNAs. Mature miRNAs bind via base-pairing to the 3′ UTR of mRNAs with sequence complementarity and decrease the amount of protein synthesized. SMIRNAs that target nuclease processing sites inhibit biogenesis, reduce mature miRNA levels, and increase protein production of downstream targets. Many cancers aberrantly express miRNAs to repress the synthesis of pro-apoptotic proteins, and SMIRNAs have been developed against them as targeted lead medicines in several cancer indications.

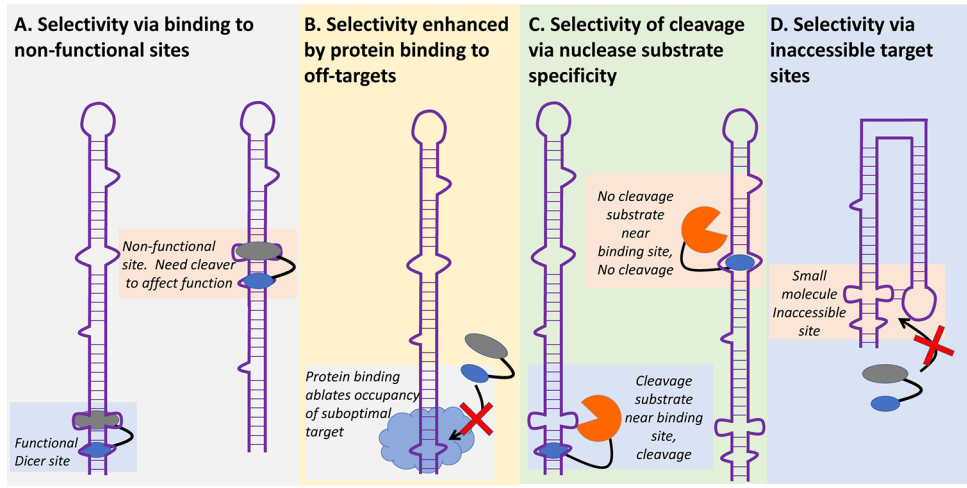

Figure 5. Factors affecting selectivity of small molecules targeting RNA. (A) Compounds can bind to RNAs but not affect biology. Targeting functional sites (i.e., nuclease processing sites) can affect RNA biology by inhibiting key processes. (B) Protein binding can enhance the selectivity of SMIRNAs. For example, sub-optimal RNA folds that bind a small molecule may be not occupied due to insufficient affinity to compete with protein binding. (C) Approaches for targeted destruction of an RNA are selective because of the inherent selectivity of the compound and the presence of a nuclease cleavage site(s) near the binding site. That is, selectivity is due to ligand binding and proper positioning of the cleaving entity. (D) Compound binding sites may not be accessible due to additional folding interactions.

Informa focuses in on highly probable regions within an RNA's structure and mines them against RNA fold−small molecule binding partners to define ligand binding sites in a robust and rational way (Figure 2). Small molecules can also be considered fragments that can be custom assembled to mimic the distance between adjacent RNA motifs, reading out their presence and enhancing affinity and selectivity.

Informa has been applied broadly to targeting miRNA precursors involved in cancers and other diseases (Figure 4), to RNA repeating transcripts that cause incurable diseases, and to mRNAs that encode undruggable proteins.[45] Although these studies have shown that RNA can be targeted with chemical probes, translating these compounds into medicines will be complex. One must be cognizant, however, that the only other modality that can be used to target human RNAs are ASOs. These high molecular weight polymers can be quickly designed from RNA sequence via Watson−Crick base-pairing and have achieved FDA approval for several indications. The road taken by these approaches was long and winding and took decades to complete. Perhaps the use of small molecules to target RNA could lead to a more accelerated path to the clinic because of the properties of the modalities themselves: they have lower molecular weight than ASOs and are less charged, and medicinal chemists have a large knowledge base on how to

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

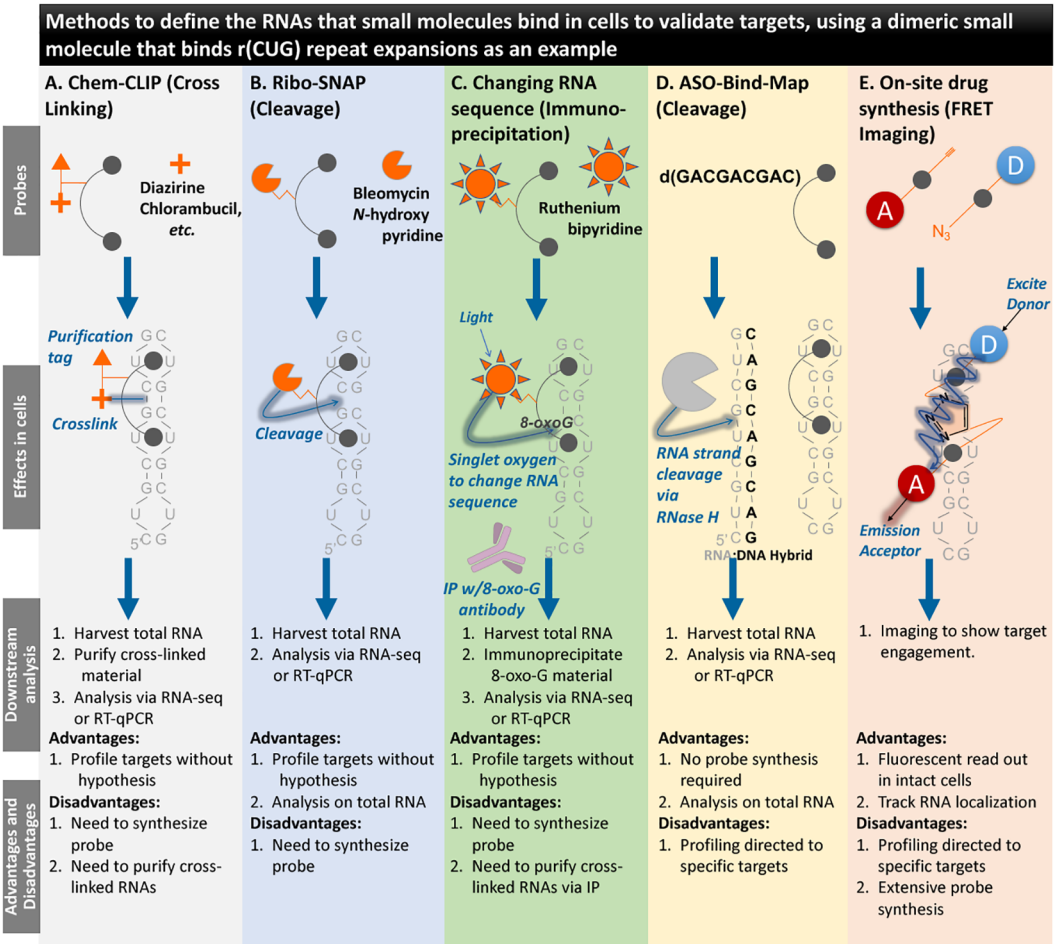

Figure 6. Target validation and profiling tools enabled by SMIRNAs. SMIRNAs have been developed that cross-link with their cellular targets (Chem-CLIP), cleave their cellular targets (RiboSNAP), change target RNA sequences, or compete with ASOs for binding (ASO-Bind-Map). The methods are complementary, and the ideal method to employ will depend on the RNA target and the inherent sequence specificity of the cross-linking, cleavage, or reactive species generated. (A) Chem-CLIP, a cross-linking approach in which small molecules bind to RNA targets and undergo a proximity-based reaction at the binding site, tagging the RNA with a purification tag. (B) Ribo-SNAP, a cleavage-based approach in which small molecules bind to RNA targets in cells and undergo a proximity-based cleavage reaction at the binding site, allowing transcriptome-wide assessment of target engagement. (C) Changing RNA sequence with a small molecule. A small molecule that targets an RNA is appended with ruthenium bipyridine. Irradiation of cells and animals with light produces reactive oxygen species that convert G to *-oxo-G. Recognition of 8-oxo-G lesions by antibodies allows immunoprecipitation of bound RNAs, which can then be analyzed. (D) ASO-Bind-Map is a competition-based experiment between ASOs and small molecules. The binding of small molecules thermodynamically stabilizes a region of defined structure and inhibits ASO binding. Inhibition of ASO cleavage indicates SMIRNA binding of the targeted mRNA. (E) On-site drug synthesis can be used to study RNA target engagement. Briefly, two pro-drugs harboring a complementary donor or acceptor bind adjacent structures in an RNA target, triggering a proximity-based click reaction (RNA is the catalyst) and producing a FRET signal. This allows imaging target engagement and also tracking of the target upon binding to the drug.

optimize small molecules. Patience is required in either case because an initial discovery could take decades to be translated to patients.

There are many key questions about 2DCS and Inforna and in the purposeful design of SMIRNAs that remain to be answered. There is a solid foundation for SMIRNAs in which compounds can be driven to have sufficient potency and selectivity by targeting multiple sites on an RNA target simultaneously. For example, dimeric compounds have been developed that target two RNA internal loop motifs, and these compounds bind to a target based on the identity of the loops that are targeted and the distance between them. These compounds are selectively bioactive and well tolerated in patient-derived cells and animal models.[46,47] As this approach evolves, linkers that read out functional groups in the RNA target, rather than simply reading out a distance, will be required for this modular strategy to reach its full potential. There is a growing information set on small molecules that bind RNA base pairs, and such modules can be incorporated into linkers.[48]

In initial investigations using 2DCS and Inforna, the focus was on providing bioactive SMIRNAs in cells and pre-clinical animal models. Now that the method has been well established, new challenges can be investigated, such as limiting the molecular weights of SMIRNAs, optimizing their medicinal properties, and other factors. Ideally, the molecular weights of the compounds targeting RNA should be as small as possible to enhance ligand efficiency and other properties. Thus, key questions that are well suited to be answered by 2DCS include not only what are the new chemotypes that confer avid and selective RNA binding but also what is the

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

lowest molecular weight fragment possible to achieve sufficient binding.

Insights into RNA Folds That Can Be Targeted by Small Molecules To Elicit a Functional Effect. There are various factors that affect whether a SMIRNA is bioactive, as binding does not guarantee a biological effect (Figure 5A). Indeed, the selectivity of a small molecule in cells can be greater than anticipated from *in vitro* studies for several reasons: (i) occupancy of a non-functional site in an off-target (Figure 5A); (ii) binding of protein to an off-target, making the small molecule binding site inaccessible (Figure 5B); (iii) for modular compounds and chimeras (discussed below), inherent selectivity due to the binding of a functional site and if a substrate for a nuclease is in close proximity to where the nuclease is positioned on an RNA target (Figure 5C); and (iv) additional structural interactions formed by the binding site (tertiary interactions or pseudoknots, for example) (Figure 5D). These same factors can also affect the druggability of a given RNA target for a small molecule.

We have shown that occupancy of a functional site is required and that bioactivity and selectivity are influenced by the avidity of the RNA−small molecule interaction and the target's expression level, among other factors (Figures 2 and 4).[49] Decades of research into RNA structure and function have provided tools that can find potential structures in an RNA target. By using cellular mapping experiments, one can identify RNA bases that are more or less likely to be in a base-paired conformation. This basic approach has been established and can be used in computational models to predict RNA structure no matter the chemical probing reagent used.[50] One must be cautious to assume that an RNA structure is functional, however, because in-depth biological investigations are required to establish function. Functional investigations need to be carefully thought out and executed in a similar manner to establishing a mode of action of a SMIRNA (see section 9).

There is some support that a combination of phylogenetic comparison and free energy minimization is able to identify evolutionarily conserved, functional structures.[51] In these approaches, one takes a portion of a given RNA and computes its propensity for structure. These structures are mined to determine if an RNA structure is evolutionarily conserved. An important analysis of RNA structures deduced from cellular mapping studies was completed to analyze whether structures are statistically significant by using a covariance analysis. These studies suggest that there is no statistically significant support for some proposed RNA structures in non-coding RNAs.[52] Thus, one must be careful in these types of analyses.

A point about RNA structure that needs to be carefully considered and that is somewhat different than for proteins is that many RNAs are unlikely to have globally well-defined structures. RNAs, however, are likely to have local hubs of defined structure. Such hubs should be high priority targets for small molecules. Thus, an RNA structure can be targeted by simple binding to affect its biology, or if one can cleave the target RNA selectively, then it could affect function even if binding was not to a functional site (see section 8).

7. VALIDATING THE RNA TARGETS OF SMALL MOLECULES IN CELLS AND ORGANISMS

One of the challenges in the developing SMIRNAs is the lack of robust methods to define ligand binding in cells and to use these results to define a mode of action. These types of tools have been invaluable in the protein targeting area, leading to assignment of new activities to enzymes and defining engaged targets. To provide methods to validate the direct interaction between RNAs and small molecules, various approaches have been developed, ranging from cross-linking and cleavage to altering RNA sequence and nuclease recruitment (Figure 6). Not only have these approaches established which RNA targets are directly engaged by small molecules, but they have also helped to define factors that influence bioactivity, as described above.

Chemical Cross-Linking and Isolation by Pull Down (Chem-CLIP). Cross-linking and purification of cross-linked products have been invaluable in the study of the interactions between proteins and small molecules[53,54] and between RNAs and proteins.[55] In Chem-CLIP (Figure 6A), a small molecule RNA binder is appended with a cross-linking module and a tag for purification. When these cell-permeable chemical probes bind to an RNA target, it brings the cross-linker into close proximity, which then reacts with the RNA. This reaction also tags the RNA with a purification module, allowing its facile isolation and identification. The RNA targets that are engaged can be deduced by RNA-seq or RT-qPCR. Binding sites can be deduced by similar approaches that include digestion of the bound RNA fragments followed by RT-qPCR analysis with a set of gene-specific primers.[47]

Chem-CLIP has answered several key questions about the molecular recognition of RNA by small molecules in cells. For example, are small molecules that target RNA structure capable of selective recognition in the context of disease-driving RNA repeat expansions vs short non-pathogenic repeats found in other transcripts? Indeed, selective recognition of long, disease-causing RNA repeats was achieved, and compounds were selectively bound to the disease-driving allele.[47]

Chem-CLIP was also used to study broadly the molecular recognition of miRNAs by small molecules, a defining factor that affects bioactivity (Figure 4). The oncogenic miR-210 is highly expressed in cancer cells transitioning from normoxia to hypoxia, or a low oxygen environment. Indeed, its expression levels are inversely proportional to probability for survival of breast cancer patients. Informa defined a highly potent and selective small molecule (**TGP-210**) that targeted pre-miR-210, inhibited its processing by Dicer in hypoxic breast cancer cells, triggered apoptosis, and slowed the growth of hypoxic tumors *in vivo*.[49] A **TGP-210** Chem-CLIP probe was used to study factors affecting RNA-targeting in living cells. Using a database of motifs found in human miRNA precursors in conjunction with Informa, we identified all miRNAs with the same motif present in pre-miR-210's Dicer site, regardless of its location within the miRNA, and all miRNAs predicted to bind **TGP-210**, albeit less optimally than pre-miR-210. These studies revealed selectivity of **TGP-210** for miR-210, as it was the most highly occupied over all types of RNAs, whether rRNA, tRNA, mRNAs, or other miRNAs. The second most highly occupied target, miR-497, was engaged 2-fold less than miR-210. Importantly, miR-497 has the exact same motif as miR-210's Dicer site. However, it is expressed at 10-fold lower levels, and the motif is not located in a processing site. Mature miR-497 levels are unaffected by treatment with **TGP-210**. These results suggest that a target's expression level affects occupancy and that binding to non-functional sites does not elicit a biological response.[49]

Chem-CLIP has also been used to define the lead compounds targeting RNA resulting from HTS campaigns

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

and lead optimization. For example, a recent study used a diazirine cross-linking module to define the binding sites for a clinical candidate to treat SMA in patient-derived cells.[56]

Mapping Binding Sites via Small Molecule Nucleic Acid Profiling by Cleavage Applied to RNA (RiboSNAP-Map). In addition to cross-linking, cleavage approaches have been used to define the RNA targets of small molecules in cells. In RiboSNAP-Map (Figure 6B), RNA-binding modules are appended with bleomycin A5 via its primary amine. Since the free amine provides a positive charge that contributes to its avidity for DNA, its conjugation to an RNA-binding module reduces DNA binding and cleavage. Thus, the RNA-binding module selectively directs the chimera to the desired RNA, bringing bleomycin A5 into close proximity and facilitating the RNA's cleavage both *in vitro* and in cells. RiboSNAP has been applied to r(CUG)^{exp}, akin to the Chem-CLIP studies described above, affording allele-selective cleavage and discrimination against shorter repeats.[47]

RiboSNAP-Map was also applied to oncogenic pri-miR-96 using a dimeric small molecule (**TGP-96**).[57] Cellular profiling of all expressed miRNAs in triple-negative breast cancer cells showed that miR-96 was most affected. Using a gene-specific primer and sequencing analysis, the exact binding site for **TGP-96** within pri-miR-96 was determined as cleavage occurred adjacent to the Drosha processing site, as expected.[57] The RiboSNAP-Map approach provides a simple-to-implement approach to map the binding sites of small molecules in cellular RNAs via RT-PCR amplification of total RNA. In contrast, Chem-CLIP requires cross-linking and purification of these adducts. The two approaches are complementary, and each may have an ideal set of targets, based on the inherent substrate preferences of the cross-linking and cleavage moieties.

Changing RNA Sequence To Determine Ligand Binding Sites. An alternative method to deduce ligand binding sites is to make small molecule-induced changes in the sequence (Figure 6C). As demonstrated in cells and in a *Drosophila* model of DM1, an RNA binder appended with tris(bipyridine)ruthenium(II) (Ru(bipy)$_3$) selectively altered the sequence of r(CUG)^{exp} upon irradiation with light.[58] Irradiation with light excites Ru(bipy)$_3$ to produce reactive oxygen species that can travel 40 Å from where they are generated to react with guanine residues, converting them to 8-oxo-7,8-dihydroguanosine (8-Oxo-G). By using an antibody that binds to RNAs with 8-oxo-G, modified RNAs can be immunoprecipitated and analyzed by RT-qPCR.[58]

Antisense Oligonucleotide Ligand Binding Site Mapping (ASO-Bind-Map). Each of the above approaches requires the synthesis of chimeric compounds comprised of an RNA-binding module and a reactive module (cross-linking, cleavage, alteration of sequence). Thus, to implement Chem-CLIP or RiboSNAP or to alter sequence, a position within the RNA-binding module that does not contribute to molecular recognition must first be identified. We therefore developed a facile approach for which the above is not required, ASO-Bind-Map (Figure 6D).[39] The ASO-Bind-Map method is a competition experiment between an RNA-binding small molecule and an ASO. By the nature of their mechanism of action, the targets of an ASO can be easily determined by depletion of RNA levels in RNA-seq data. Hybridization of ASOs to RNA targets is highly dependent on the thermal stability of the structure of the targeted RNA sequence. Small molecule binding increases the RNA's thermostability, thereby reducing the ability of an ASO to hybridize to the region where the small molecule is bound. Thus, the binding sites of small molecules within RNA targets can be determined by using a set of tiling oligonucleotides that walk down an RNA of interest. If levels of the region where the ASO binds are restored upon small molecule treatment, it indicates an overlap between the binding site for the ASO and the small molecule. The ASO-Bind-Map approach was used to profile an Inforna-derived compound that targets a functionally important region in the 3′ UTR of hepatitis C virus (HCV).[39] By using ASO-Map *in vitro* and in cells, the small molecule's binding site was validated and its mode of inhibiting viral replication confirmed.[39]

On-Site Drug Synthesis, a FRET Reporter of Target Engagement and Cellular Localization. For RNAs with two or more adjacent targetable motifs, an on-site drug synthesis approach can be used to generate a fluorescence resonance energy transfer (FRET) reporter of target engagement (Figure 6E).[47] On-site drug synthesis is enabled by equipping the small molecules with complementary chemical handles that react when they bind adjacent sites in the RNA target. That is, the RNA acts as a template or catalyst, bringing the two reactive handles, such as an azide and alkyne, within close proximity to facilitate the synthesis of a dimeric (or multimeric) compound. To generate a FRET signal, the two small molecules are also appended with donor and acceptor fluorophores that are FRET pairs. Upon on-site drug synthesis, the two fluorophores are within close proximity, and a FRET signal is observed. This approach was developed for an RNA repeat expansion, which not only verified direct target engagement and on-site drug synthesis, but was also used to image the RNA and study its cellular localization and changes thereof upon compound binding.[47]

Implications of Profiling Experiments. There are various uses for these profiling experiments depending on the question of interest. For an easy-to-implement approach to define ligand binding sites without having to complete additional chemical synthesis, an ASO-Bind-Map experiment is ideal and should be first-in-line to assess target engagement in cells for a specific target and the small molecule binding site therein. In contrast, Chem-CLIP and RiboSNAP are target agnostic and thus can be applied transcriptome-wide. The nitrogen mustard (chlorambucil), diazirine, and cleavers each have substrate preferences. Thus, using all three methods in parallel can provide a rigorous assessment of cellular targets and binding sites.

8. SMALL MOLECULE TARGETED DEGRADATION OF RNAs IN CELLS

ASOs and siRNAs are known to recruit nucleases to cleave RNAs and interface with RNA quality control (QC) pathways. Once transcribed, however, all cellular RNAs must be processed to generate their final, active forms. Such processing can include endo- and exonucleolytic cleavage of sequences at either end of the initial transcript, removal of internal sequences such as internal transcribed spacers and introns, nucleotide editing, and extensive functionalization via chemical modification. Remarkably, most cellular RNAs are subject to multiple processing reactions, with alternate pathways (e.g., alternative splicing) giving rise to distinct products that greatly increase the functional diversity of RNA and protein species encoded by individual genes. Further, there are numerous RNA QC mechanisms to eliminate incorrectly or incompletely

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

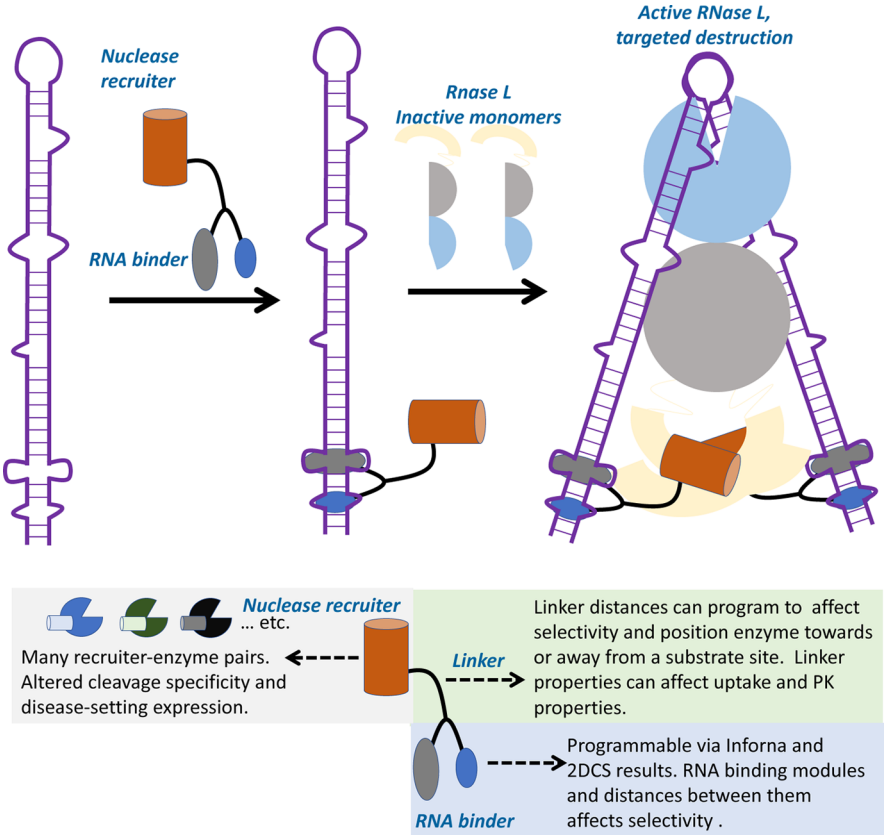

Figure 7. Ribonuclease-targeted chimeras (RIBOTACs) as an approach to cleave RNAs with endogenous nucleases. (Top) Targeted recruitment of RNase L with a small molecule. This approach has been shown to selectively and potently cleave a targeted RNA in a catalytic and sub-stoichiometric manner. (Bottom) RIBOTACs can be extended to other RNA-modifying enzymes by changing the recruiter module. Enzymes have defined substrate specificity and expression levels. As recruiters are developed for more enzymes, these factors can be used to control selectivity. Other factors that control selectivity include the linker between RNA-binding modules and the recruiter, which also influences cellular/tissue uptake and localization. Incorporation of enzyme-recruiting small molecules into Informa allows a streamlined and designer way to affect RNA biology via cleavage, akin to antisense.

processed RNAs.[59] Interestingly, the inherent activity of these enzymes and pathways could be harnessed to selectively degrade an RNA target, a potentially transformative paradigm to design drugs for the treatment of human diseases.

We therefore sought to develop a way to effect RNA cleavage with small molecules by recruiting nucleases, approach coined ribonuclease targeting chimeras (RIBOTACs; Figure 7). In a first iteration, a small molecule that bound to pri-miR-96 was appended with 2′-5′-A_4 (**TGP-96-RL**) that binds to inactive, monomeric RNase L and dimerizes it into an active nuclease.[60] RNase L is produced in all cells, and thus targeted recruitment of this enzyme would have broad implications to affect RNA lifetime with small molecules. Application of **TGP-96-RL** to triple-negative breast cancer cells reduced pri-miR-96 levels and inhibited a miR-96-mediated phenotype while having no effect on healthy breast epithelial cells.[60] A variety of important control experiments were completed to demonstrate that pri-mir-96 was cleaved by **TGP-96-RL**: (i) addition of increasing concentrations of **TGP-96** (the bioactive parent compound lacking the RNase L-recruiting module) to cells along with a constant concentration of **TGP-96-RL** ablated cleavage due to competition for the same binding site on the target; (ii) ablation of RNase L via an siRNA, but not a control siRNA, inhibited the ability of **TGP-96-RL** to cleave pri-miR-96; and (iii) immunoprecipitation of RNase L in the presence of **TGP-96-RL**, but not in its absence, pulled down pri-miR-96, supporting formation of a ternary complex.[60] One concern with activating a nuclease with a small molecule was whether the selectivity of the parent compound would be maintained. Full target profiling, however, showed exquisite selectivity for cleaving pri-miR-96. Furthermore, **TGP-96-RL** cleaved the desired target catalytically and sub-stoichiometrically.[60]

There are broad implications of this RIBOTACs approach (Figure 7). For example, other nucleases and RNA processing enzymes could be recruited to cleave an RNA, alter its sequence, etc. As each nuclease and processing enzyme has inherent substrate specificity, it is possible that a suboptimal RNA-binding module could be used. That is, the selectivity of both the targeting moiety and the recruited nuclease would define the targets of the chimera, providing significant off-target buffering (Figure 7). Further, selectivity can be garnered from the linker that tethers the RNA-binding and protein-recruiting modules, in terms of both its identity, which also influences cellular uptake and localization, and its length, which can position the enzyme toward or away from a substrate (Figure 7).

9. FUNCTIONAL VALIDATION OF RNA TARGETS IN CELLS

Given the diverse biology of RNAs, particularly non-coding (nc)RNAs, in different cells and tissues, functional validation is

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

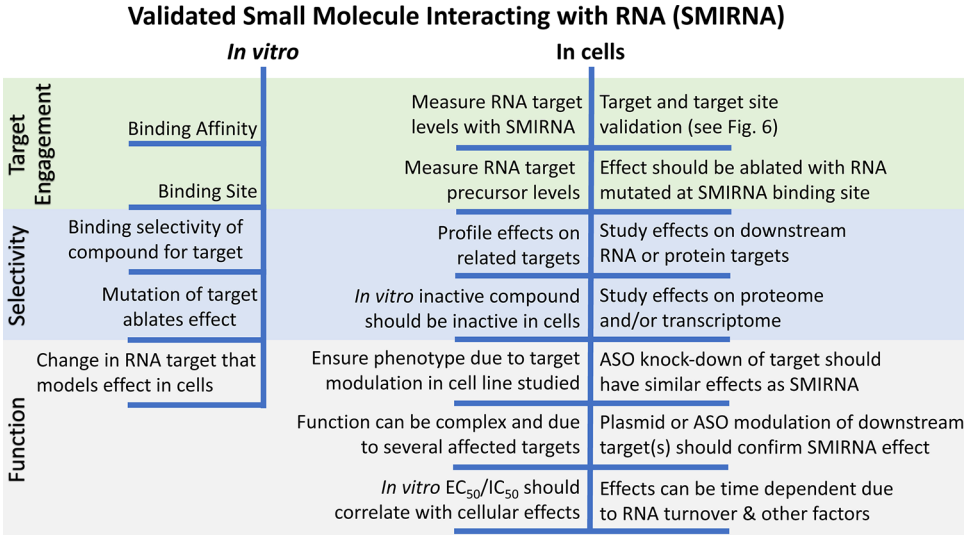

Figure 8. Experimental workflow for developing validated small molecules interacting with RNA (SMIRNA) as lead chemical probes or medicines. The design and validation of small molecules that target RNA can be challenging. Illustrated is a scheme for their comprehensive evaluation, including in vitro and cellular studies, as well as other factors that should be considered. Notably, cellular studies should comprise an assessment of the small molecule's effect on transcript and protein levels, its selectivity, direct target engagement, and its ability to reverse phenotype.

critical. It is not sufficient to demonstrate that a small molecule targets an RNA or inhibits the formation of an RNA−protein complex in vitro and presume that a phenotypic change is due to the in vitro phenomenon. Figure 8 describes various approaches that are required to establish rigorously a mode of action for a compound. A first assessment in cells should include measuring transcript levels. For compounds that inhibit miRNA biogenesis and processing, a compound with the intended mode of action should stimulate an increase in the abundance of precursors and a reduction of the mature miRNA. Reduction of both precursor and mature miRNA levels indicates transcriptional inhibition that could be due to several factors, including binding to DNA. Since many miRNAs and other ncRNAs impart their activity by affecting protein production, further validation of changes in the proteome should also be assessed via Western blotting, ELISA, etc. (Figure 4). Once a compound's mode of action has been established at both the transcript and protein levels, the inhibitor should then be assessed for reversing phenotype (Figure 8). Notably, phenotype is cell-type dependent. Take oncogenic miR-21 for example, which is highly expressed in various cancers. In breast cancer, miR-21 effects an invasive phenotype via repression of phosphatase and tensin homologue (PTEN), while in prostate cancer miR-21 effects chemoresistance via programmed cell death protein 4 (PDCD4). Thus, reversal of phenotype by a miR-21 inhibitor should be measured as reduction of invasion in breast cancer cells but enhanced chemosensitivity in prostate cancer cells. Indeed, the correct phenotype to be monitored in a particular cell or tissue type can be elucidated with ASOs or antagomiRs.

There are other important phenotype controls that should be considered (Figure 8). First, mutation of the target should decrease the binding of ligands and hence its ability to reverse phenotype. Such studies may prove difficult with miRNAs, given that the binding site for a small molecule may overlap with the sequence that binds to the 3′ UTR on a targeted mRNA. This would affect the expression in pathway-associated targets. Second, alterring the expression levels of downstream proteins should result in phenotypic changes. Overexpression of the target protein (via a plasmid) should enhance the reversal of phenotype while siRNA ablation or co-treatment with a protein inhibitor should render the SMIRNA inactive. For example, a small molecule that targeted miR-544 triggered apoptosis in hypoxic breast cancer and increased mTOR levels. The use of the mTOR inhibitor rapamycin ablated the compound's activity, supporting that mTOR pathway upregulation affects phenotype.[61] It is essential that on-target modulation of phenotype is assessed and done so rigorously.

Functional validation for compounds that target miRNAs can be complicated, as several mRNAs can be targeted by the same miRNA and several miRNAs can target the same mRNA. One way to study effects on the miRNA-mediated pathway affected by the small molecule is via siRNA ablation of the downstream mRNA target(s). Removal of the downstream target should inhibit or reduce the ability of the small molecule to effect a miRNA-directed phenotype, as we have completed for miR-96,[37] demonstrating that reversal of phenotype can indeed be traced to the effect of one miRNA on one mRNA target.

10. CHEMOTYPES IN SMALL MOLECULES THAT BIND RNA TARGETS AND DRUG-LIKENESS

The irony of much of the discussion surrounding drug-likeness is that many life-saving medicines, like streptomycin, are considered to have "undrug-like" properties. Further, oligonucleotide-based approaches, the most vetted modality to target RNA, are rather undrug-like with their high molecular weights and charge. Oligonucleotides thus can suffer from limited tissue distribution and significant off-targets in humans. Yet they have reached clinical use. Lipinski guidelines and other analyses on drug-likeness have been completed on compounds that have garnered FDA approval. For RNA-targeted small molecules, it is likely premature to assess the similarities and differences of these compounds relative to known drugs, owing to the few success stories upon which to base such analyses. In fact, the SMA drugs target an interface between an RNA and a protein and as such are not strictly RNA-targeting compounds. Furthermore, the average molecular weight of FDA-approved

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

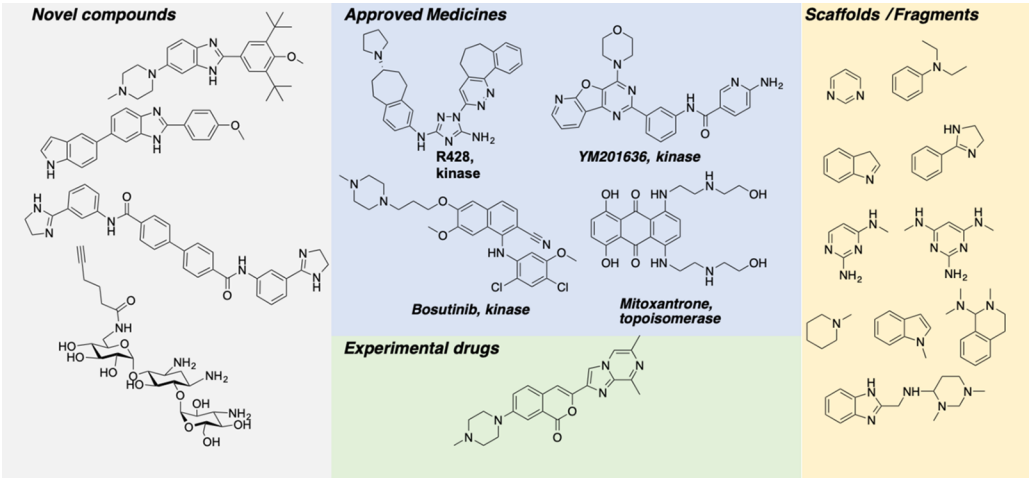

Figure 9. Chemically diverse RNA-binding small molecules and the RNA folds that interact with them have been defined by 2DCS. Novel compounds, including derivatives of nucleic acid binders to ablate inherent substrate specificity, as well as both known and experimental drugs have been studied. Analysis of RNA binders from 2DCS has shown that various fragments, both present and absent in known drugs, bind RNA avidly. The observation that known drugs bind RNA, especially kinase and topoisomerase inhibitors, suggests that RNA should be considered as both an on- and off-target of experimental and known drugs. The canonical targets of known drugs that target RNA are listed under the compound's structure.

modalities in the past several years has increased.[62] This observation perhaps calls into question the hypothesis of the existence of drug-like properties in drug discovery. Targeted protein degradation has, without a doubt, transformed the idea of using higher molecular weight modalities in the protein field, with little debate about their "drug-likeness" affecting their utility as lead medicines.

A comparison of the physicochemical properties of RNA-binding small molecules, including those from a recently reported large 2DCS study, to compounds in DrugBank revealed that they have drug-like properties (Figure 9).[39] These studies also revealed fragments or scaffolds that confer avidity and selectivity for RNA motifs (Figure 9).[39] Not surprisingly, RNA binders have an increased number of hydrogen bond donors and acceptors.[39] Caution must be exercised, however, as RNA-binding is not correlated with bioactivity in cells or in vivo. Nor is it correlated with compounds having sufficient potency and selectivity to be bioactive. Another study showed that known drugs bind RNA (Figure 9).[40] For example, kinase inhibitors bind RNA broadly and interact with certain RNA folds with good selectivity. Topoisomerase inhibitors bind to many RNAs, including the Dicer site in the pre-miR-21, with low nanomolar affinity. Cellular studies showed that the topoisomerase inhibitor that bound miR-21 modulated a miR-21-mediated phenotype and directly interacted with pre-miR-21 in cells (via a Chem-CLIP).[40] Despite these datasets, it is still currently premature to assert that drug-like space for RNA exists. There is simply not enough information at present to make such statements. Long-term we may uncover that SMIRNAs may have properties that lie outside of current "drug-like" space. For example, it would not be surprising if these molecules have more hydrogen bond donors and acceptors, a higher hydrophobic surface area, or higher molecular weights. Many areas in drug discovery are expanding the view of drug-like space outside of the Lipinski guidelines.

11. RNA AS A TARGET OF SMALL MOLECULES: CONSIDERATIONS AND PATHS FOR ACCELERATION

The drug industry rarely considers RNA as a druggable target, and phenotypic screens historically do not consider RNA as the biological target to exert an effect on phenotype. However, this might change with emerging target validation tools, the discovery of drugs that improve diseases by targeting RNA, and the observation that known drugs target RNAs and affect RNA-mediated phenotypes in a clinical setting (e.g., cancer). An essential companion dataset comprises molecules that are proven to directly engage RNA targets and robustly affect disease biology in cellular and animal models, as demonstrated with tool compounds for RNAs operating in SMA, DM1, and various cancers. There is no doubt that more successes will come in pre-clinical models, providing a rationale for sustained efforts in this area. However, direct engagement of the RNA target clearly linked to phenotype modification (see Figure 8) will be necessary to push drugging the transcriptome forward.

Although many have suggested that identifying selective small molecules for RNA targets is an insurmountable challenge, we strongly believe that many data support the notion that RNA motifs can be targeted selectively with small molecules. The above-mentioned tools (Chem-CLIP, ASO-Bind-Map, and RiboSNAP; Figure 6) have shown that compounds directly engage targets in cells, and target engagement affects disease phenotype. Kinases were once thought to be a target that could not be selectively modified with small molecules; however, myriad cases have shown that the initial proclivity to view that selectively targeting them was impossible has not been generally true. The same will likely happen with RNA, and datasets to support this hypothesis will continue to emerge. Success in an area could very easily be defined as making possible what was once viewed as being impossible.

As we discussed herein, known drugs can target RNA and also affect disease biology in a manner that is consistent with the phenotypes that drugs are known to affect in a clinical setting. Drug discovery efforts have so pinned RNA as

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

"undruggable", that RNAs are not even present in safety screens such as CEREP panels (i.e., Safety44). It is provocative to think that failures in drug safety assessment through preclinical development could be due—at least in part—to RNA off-targets. Fortuitously, RNA targets can easily be assessed by RNA-seq, providing a facile route to eliminate drugs that could have side effects on the transcriptome. Such analyses should be routine in the drug discovery process and should focus on studying side effects in tissues or human cell lines that would preclude further development, including heart, kidney, and liver. Although the resulting datasets are immense and often difficult to deconvolute, they are *de rigueur* for biology and should also emerge as such in medicinal chemistry. It is no doubt an exciting time to be drugging RNA with small molecules. The ENCODE and other projects have provided a nearly limitless number of targets, and it is likely that many of them will prove selectively and potently targetable with SMIRNAs.

AUTHOR INFORMATION

Corresponding Author

*disney@scripps.edu

ORCID

Matthew D. Disney: 0000-0001-8486-1796

Notes

The author declares the following competing financial interest(s): Co-founder of Expansion Therapeutics that is developing medicines for patients with RNA repeat expansion disorders.

ACKNOWLEDGMENTS

We thank our co-workers who have been on this journey since 2005 to develop methods to purposefully target RNA with small molecules. We thank the National Institutes of Health for funding (R01 GM97455, DP1 NS096898, P01 NS09914, and R33 NS096032 to M.D.D.).

REFERENCES

(1) Holley, R. W.; Apgar, J.; Everett, G. A.; Madison, J. T.; Marquisee, M.; Merrill, S. H.; Penswick, J. R.; Zamir, A. Structure of a Ribonucleic Acid. *Science (Washington, DC, U. S.)* **1965**, *147* (3664), 1462−5.

(2) Jones, D.; Metzger, H. J.; Schatz, A.; Waksman, S. A. Control of gram-negative bacteria in experimental animals by streptomycin. *Science* **1944**, *100* (2588), 103−105.

(3) Finken, M.; Kirschner, P.; Meier, A.; Wrede, A.; Bottger, E. C. Molecular basis of streptomycin resistance in Mycobacterium tuberculosis: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Mol. Microbiol.* **1993**, *9* (6), 1239−46.

(4) Woese, C. R.; Magrum, L. J.; Gupta, R.; Siegel, R. B.; Stahl, D. A.; Kop, J.; Crawford, N.; Brosius, J.; Gutell, R.; Hogan, J. J.; Noller, H. F. Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res.* **1980**, *8* (10), 2275−93.

(5) Inoue, T.; Cech, T. R. Secondary structure of the circular form of the Tetrahymena rRNA intervening sequence: a technique for RNA structure analysis using chemical probes and reverse transcriptase. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (3), 648−52.

(6) Stern, S.; Wilson, R. C.; Noller, H. F. Localization of the binding site for protein S4 on 16 S ribosomal RNA by chemical and enzymatic probing and primer extension. *J. Mol. Biol.* **1986**, *192* (1), 101−10.

(7) Moazed, D.; Noller, H. F. Interaction of antibiotics with functional sites in 16S ribosomal-RNA. *Nature* **1987**, *327* (6121), 389−394.

(8) Lynch, S. R.; Recht, M. I.; Puglisi, J. D. Biochemical and nuclear magnetic resonance studies of aminoglycoside-RNA complexes. *Methods Enzymol.* **2000**, *317*, 240−61.

(9) Yonath, A. Antibiotics targeting ribosomes: resistance, selectivity, synergism and cellular regulation. *Annu. Rev. Biochem.* **2005**, *74*, 649−79.

(10) Zaug, A. J.; Cech, T. R. The intervening sequence RNA of Tetrahymena is an enzyme. *Science* **1986**, *231* (4737), 470−5.

(11) Baer, M.; Altman, S. A catalytic RNA and its gene from Salmonella typhimurium A view of RNase P. *Science* **1985**, *228* (4702), 999−1002.

(12) Joyce, G. F. In vitro evolution of nucleic acids. *Curr. Opin. Struct. Biol.* **1994**, *4*, 331−6.

(13) Winkler, W.; Nahvi, A.; Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **2002**, *419* (6910), 952−956.

(14) Blount, K. F.; Wang, J. X.; Lim, J.; Sudarsan, N.; Breaker, R. R. Antibacterial lysine analogs that target lysine riboswitches. *Nat. Chem. Biol.* **2007**, *3* (1), 44−9.

(15) Otani, S.; Takatsu, M.; Nakano, M.; Kasai, S.; Miura, R.; Matsui, K. Roseoflavin, a new antimicrobial pigment from Streptomyces. *J. Antibiot.* **1974**, *27* (1), 88−89.

(16) Pedrolli, D. B.; Matern, A.; Wang, J.; Ester, M.; Siedler, K.; Breaker, R.; Mack, M. A highly specialized flavin mononucleotide riboswitch responds differently to similar ligands and confers roseoflavin resistance to Streptomyces davawensis. *Nucleic Acids Res.* **2012**, *40* (17), 8662−73.

(17) Kil, Y. V.; Mironovi, V. N.; Gorishin, I.; Kreneva, R. A.; Perumov, D. A. Riboflavin operon of Bacillus subtilis: unusual symmetric arrangement of the regulatory region. *Mol. Gen. Genet.* **1992**, *233* (3), 483−6.

(18) Klase, Z.; Winograd, R.; Davis, J.; Carpio, L.; Hildreth, R.; Heydarian, M.; Fu, S.; McCaffrey, T.; Meiri, E.; Ayash-Rashkovsky, M.; Gilad, S.; Bentwich, Z.; Kashanchi, F. HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression. *Retrovirology* **2009**, *6*, 18.

(19) Mei, H. Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. Discovery of selective, small-molecule inhibitors of RNA complexes–I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorg. Med. Chem.* **1997**, *5* (6), 1173−84.

(20) Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **1992**, *257* (5066), 76−80.

(21) Brodsky, A. S.; Williamson, J. R. Solution structure of the HIV-2 TAR-argininamide complex. *J. Mol. Biol.* **1997**, *267* (3), 624−39.

(22) Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-Ela, F.; Varani, G.; Karn, J.; Klimkait, T. An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (8), 3548−53.

(23) Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. Structural mimicry of retroviral tat proteins by constrained beta-hairpin peptidomimetics: ligands with high affinity and selectivity for viral TAR RNA regulatory elements. *J. Am. Chem. Soc.* **2004**, *126* (22), 6906−13.

(24) Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Lee, J.; Andricioaei, I.; Markovitz, D. M.; Al-Hashimi, H. M. Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble. *Nat. Chem. Biol.* **2011**, *7* (8), 553−9.

(25) Harada, K.; Martin, S. S.; Frankel, A. D. Selection of RNA-binding peptides in vivo. *Nature* **1996**, *380* (6570), 175−9.

(26) Howe, J. A.; Wang, H.; Fischmann, T. O.; Balibar, C. J.; Xiao, L.; Galgoci, A. M.; Malinverni, J. C.; Mayhood, T.; Villafania, A.; Nahvi, A.; Murgolo, N.; Barbieri, C. M.; Mann, P. A.; Carr, D.; Xia,

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

E.; Zuck, P.; Riley, D.; Painter, R. E.; Walker, S. S.; Sherborne, B.; de Jesus, R.; Pan, W.; Plotkin, M. A.; Wu, J.; Rindgen, D.; Cummings, J.; Garlisi, C. G.; Zhang, R.; Sheth, P. R.; Gill, C. J.; Tang, H.; Roemer, T. Selective small-molecule inhibition of an RNA structural element. *Nature* **2015**, *526* (7575), 672−7.

(27) Langer, S.; Hashimoto, M.; Hobl, B.; Mathes, T.; Mack, M. Flavoproteins are potential targets for the antibiotic roseoflavin in Escherichia coli. *J. Bacteriol.* **2013**, *195* (18), 4037−45.

(28) Pedrolli, D. B.; Nakanishi, S.; Barile, M.; Mansurova, M.; Carmona, E. C.; Lux, A.; Gartner, W.; Mack, M. The antibiotics roseoflavin and 8-demethyl-8-amino-riboflavin from Streptomyces davawensis are metabolized by human flavokinase and human FAD synthetase. *Biochem. Pharmacol.* **2011**, *82* (12), 1853−9.

(29) Naryshkin, N. A.; Weetall, M.; Dakka, A.; Narasimhan, J.; Zhao, X.; Feng, Z.; Ling, K. K.; Karp, G. M.; Qi, H.; Woll, M. G.; Chen, G.; Zhang, N.; Gabbeta, V.; Vazirani, P.; Bhattacharyya, A.; Furia, B.; Risher, N.; Sheedy, J.; Kong, R.; Ma, J.; Turpoff, A.; Lee, C. S.; Zhang, X.; Moon, Y. C.; Trifillis, P.; Welch, E. M.; Colacino, J. M.; Babiak, J.; Almstead, N. G.; Peltz, S. W.; Eng, L. A.; Chen, K. S.; Mull, J. L.; Lynes, M. S.; Rubin, L. L.; Fontoura, P.; Santarelli, L.; Haehnke, D.; McCarthy, K. D.; Schmucki, R.; Ebeling, M.; Sivaramakrishnan, M.; Ko, C. P.; Paushkin, S. V.; Ratni, H.; Gerlach, I.; Ghosh, A.; Metzger, F. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* **2014**, *345* (6197), 688−93.

(30) Palacino, J.; Swalley, S. E.; Song, C.; Cheung, A. K.; Shu, L.; Zhang, X.; Van Hoosear, M.; Shin, Y.; Chin, D. N.; Keller, C. G.; Beibel, M.; Renaud, N. A.; Smith, T. M.; Salcius, M.; Shi, X.; Hild, M.; Servais, R.; Jain, M.; Deng, L.; Bullock, C.; McLellan, M.; Schuierer, S.; Murphy, L.; Blommers, M. J.; Blaustein, C.; Berenshteyn, F.; Lacoste, A.; Thomas, J. R.; Roma, G.; Michaud, G. A.; Tseng, B. S.; Porter, J. A.; Myer, V. E.; Tallarico, J. A.; Hamann, L. G.; Curtis, D.; Fishman, M. C.; Dietrich, W. F.; Dales, N. A.; Sivasankaran, R. SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat. Chem. Biol.* **2015**, *11* (7), 511−7.

(31) Paulson, H. Repeat expansion diseases. *Handbook of Clinical Neurology* **2018**, *147*, 105−123.

(32) Chen, C. Z.; Sobczak, K.; Hoskins, J.; Southall, N.; Marugan, J. J.; Zheng, W.; Thornton, C. A.; Austin, C. P. Two high-throughput screening assays for aberrant RNA-protein interactions in myotonic dystrophy type 1. *Anal. Bioanal. Chem.* **2012**, *402* (5), 1889−98.

(33) Hoskins, J. W.; Ofori, L. O.; Chen, C. Z.; Kumar, A.; Sobczak, K.; Nakamori, M.; Southall, N.; Patnaik, S.; Marugan, J. J.; Zheng, W.; Austin, C. P.; Disney, M. D.; Miller, B. L.; Thornton, C. A. Lomofungin and dilomofungin: inhibitors of MBNL1-CUG RNA binding with distinct cellular effects. *Nucleic Acids Res.* **2014**, *42* (10), 6591−602.

(34) Bevilacqua, P. C.; Assmann, S. M. Technique Development for Probing RNA Structure In Vivo and Genome-Wide. *Cold Spring Harbor Perspect. Biol.* **2018**, *10* (10), a032250.

(35) Eddy, S. R. Computational analysis of conserved RNA secondary structure in transcriptomes and genomes. *Annu. Rev. Biophys.* **2014**, *43*, 433−56.

(36) Disney, M. D.; Winkelsas, A. M.; Velagapudi, S. P.; Southern, M.; Fallahi, M.; Childs-Disney, J. L. Inforna 2.0: A platform for the sequence-based design of small molecules targeting structured RNAs. *ACS Chem. Biol.* **2016**, *11* (6), 1720−8.

(37) Velagapudi, S. P.; Gallo, S. M.; Disney, M. D. Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nat. Chem. Biol.* **2014**, *10* (4), 291−7.

(38) Disney, M. D.; Labuda, L. P.; Paul, D. J.; Poplawski, S. G.; Pushechnikov, A.; Tran, T.; Velagapudi, S. P.; Wu, M.; Childs-Disney, J. L. Two-dimensional combinatorial screening identifies specific aminoglycoside-RNA internal loop partners. *J. Am. Chem. Soc.* **2008**, *130* (33), 11185−94.

(39) Childs-Disney, J. L.; Tran, T.; Vummidi, B. R.; Velagapudi, S. P.; Haniff, H. S.; Matsumoto, Y.; Crynen, G.; Southern, M. R.; Biswas, A.; Wang, Z.-F.; Tellinghuisen, T. L.; Disney, M. D. A massively parallel selection of small molecule-RNA motif binding partners informs design of an antiviral from sequence. *Chem* **2018**, *4* (10), 2384−2404.

(40) Velagapudi, S. P.; Costales, M. G.; Vummidi, B. R.; Nakai, Y.; Angelbello, A. J.; Tran, T.; Haniff, H. S.; Matsumoto, Y.; Wang, Z. F.; Chatterjee, A. K.; Childs-Disney, J. L.; Disney, M. D. Approved anticancer drugs target oncogenic non-coding RNAs. *Cell Chem. Biol.* **2018**, *25* (9), 1086−1094.

(41) Velagapudi, S. P.; Luo, Y.; Tran, T.; Haniff, H. S.; Nakai, Y.; Fallahi, M.; Martinez, G. J.; Childs-Disney, J. L.; Disney, M. D. Defining RNA-small molecule affinity landscapes enables design of a small molecule inhibitor of an oncogenic noncoding RNA. *ACS Cent. Sci.* **2017**, *3* (3), 205−216.

(42) Velagapudi, S. P.; Seedhouse, S. J.; Disney, M. D. Structure-activity relationships through sequencing (StARTS) defines optimal and suboptimal RNA motif targets for small molecules. *Angew. Chem., Int. Ed. Engl.* **2010**, *49* (22), 3816−3818.

(43) Lu, Z.; Chang, H. Y. Decoding the RNA structurome. *Curr. Opin. Struct. Biol.* **2016**, *36*, 142−8.

(44) Bevilacqua, P. C.; Ritchey, L. E.; Su, Z.; Assmann, S. M. Genome-wide analysis of RNA secondary structure. *Annu. Rev. Genet.* **2016**, *50*, 235−266.

(45) Angelbello, A. J.; Chen, J. L.; Childs-Disney, J. L.; Zhang, P.; Wang, Z. F.; Disney, M. D. Using Genome Sequence to Enable the Design of Medicines and Chemical Probes. *Chem. Rev.* **2018**, *118* (4), 1599−1663.

(46) Velagapudi, S. P.; Cameron, M. D.; Haga, C. L.; Rosenberg, L. H.; Lafitte, M.; Duckett, D. R.; Phinney, D. G.; Disney, M. D. Design of a small molecule against an oncogenic noncoding RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (21), 5898−903.

(47) Rzuczek, S. G.; Colgan, L. A.; Nakai, Y.; Cameron, M. D.; Furling, D.; Yasuda, R.; Disney, M. D. Precise small-molecule recognition of a toxic CUG RNA repeat expansion. *Nat. Chem. Biol.* **2017**, *13* (2), 188−193.

(48) Haniff, H. S.; Graves, A.; Disney, M. D. Selective small molecule recognition of RNA base pairs. *ACS Comb. Sci.* **2018**, *20* (8), 482−491.

(49) Costales, M. G.; Haga, C. L.; Velagapudi, S. P.; Childs-Disney, J. L.; Phinney, D. G.; Disney, M. D. Small molecule inhibition of microRNA-210 reprograms an oncogenic hypoxic circuit. *J. Am. Chem. Soc.* **2017**, *139* (9), 3446−3455.

(50) Mathews, D. H.; Disney, M. D.; Childs, J. L.; Schroeder, S. J.; Zuker, M.; Turner, D. H. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (19), 7287−92.

(51) Freyhult, E.; Gardner, P. P.; Moulton, V. A comparison of RNA folding measures. *BMC Bioinformatics* **2005**, *6*, 241.

(52) Rivas, E.; Clements, J.; Eddy, S. R. A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods* **2017**, *14* (1), 45−48.

(53) Simon, G. M.; Niphakis, M. J.; Cravatt, B. F. Determining target engagement in living systems. *Nat. Chem. Biol.* **2013**, *9* (4), 200−5.

(54) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; Narvaiza, I.; Saez, E.; Cravatt, B. F. Ligand and target discovery by fragment-based screening in human cells. *Cell* **2017**, *168* (3), 527−541.

(55) Ule, J.; Jensen, K. B.; Ruggiu, M.; Mele, A.; Ule, A.; Darnell, R. B. CLIP identifies Nova-regulated RNA networks in the brain. *Science* **2003**, *302* (5648), 1212−5.

(56) Wang, J.; Schultz, P. G.; Johnson, K. A. Mechanistic studies of a small-molecule modulator of SMN2 splicing. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (20), E4604−e4612.

(57) Li, Y.; Disney, M. D. Precise small molecule degradation of a noncoding RNA identifies cellular binding sites and modulates an oncogenic phenotype. *ACS Chem. Biol.* **2018**, *13* (11), 3065−3071.

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790

(58) Guan, L.; Luo, Y.; Ja, W. W.; Disney, M. D. Small molecule alteration of RNA sequence in cells and animals. *Bioorg. Med. Chem. Lett.* **2018**, *28* (16), 2794−2796.
(59) Doma, M. K.; Parker, R. RNA quality control in eukaryotes. *Cell* **2007**, *131* (4), 660−8.
(60) Costales, M. G.; Matsumoto, Y.; Velagapudi, S. P.; Disney, M. D. Small Molecule Targeted Recruitment of a Nuclease to RNA. *J. Am. Chem. Soc.* **2018**, *140* (22), 6741−6744.
(61) Haga, C. L.; Velagapudi, S. P.; Strivelli, J. R.; Yang, W. Y.; Disney, M. D.; Phinney, D. G. Small molecule inhibition of miR-544 biogenesis disrupts adaptive responses to hypoxia by modulating ATM-mTOR signaling. *ACS Chem. Biol.* **2015**, *10* (10), 2267−76.
(62) Shultz, M. D. Two Decades under the Influence of the Rule of Five and the Changing Properties of Approved Oral Drugs. *J. Med. Chem.* **2019**, *62* (4), 1701−1714.

DOI: 10.1021/jacs.8b13419
J. Am. Chem. Soc. 2019, 141, 6776−6790