Drugging the RNA World

Matthew D. Disney, Brendan G. Dwyer, and Jessica L. Childs-Disney

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

Correspondence: disney@scripps.edu

SUMMARY

Although we live in the remnants of an RNA world, the world of drug discovery and chemical probes is firmly protein-centric. Developing highly selective small molecules targeting RNA is often considered to be an insurmountable challenge. Our goal is to demystify the design of such compounds. In this review, we describe various approaches to design small molecules that target RNA from sequence and the application of these compounds in RNA biology, with a focus on inhibition of human RNA–protein complexes. We have developed a library-versus-library screening approach to define selective RNA–small-molecule binding partners and applied them to disease-causing RNAs, in particular noncoding oncogenic RNAs and expanded RNA repeats, to modulate their biology in cells and animals. We also describe the design of new types of small-molecule probes that could broadly decipher the mysteries of RNA in cells.

Outline

1 Introduction	4 Concluding remarks
2 Targeting RNA with small molecules	References
3 Using RNA structure to design selective small molecules targeting RNA	

1 INTRODUCTION

1.1 Sequence-Structure: The Tree of Life in the RNA World

Robert Holley and colleagues were the first to sequence a biomolecule, transfer RNA (tRNA)^{Ala}, after purifying modest amounts to study its role in protein synthesis. Phylogenetic sequence alignment and enzymatic probing suggested that tRNA adopted a cloverleaf structure (Holley 1965). Indeed, all tRNAs were found to adopt this general structure (Dudock et al. 1969). These insights were critical to the fundamental understanding of translation, an RNA-driven process. The interaction between the codon and anticodon, formed when the tRNA adopts its cloverleaf structure, decodes a messenger RNA (mRNA) into a polypeptide.

Although these investigations occurred more than 60 years ago, their reverberations are still felt today. The work of Holley and colleagues stimulated great efforts into developing improved methods for RNA sequencing (RNA-seq). These studies marched from (1) the plus and minus method of Sanger sequencing, to (2) Gilbert–Maxam sequencing, to (3) dideoxy sequencing, and, finally, to (4) the next-generation sequencing technology of today that can sequence an entire genome in a day for <\$1000. Perhaps one of the most interesting aspects of sequencing has been the recognition that noncoding RNAs are numerous in cells, laying the foundation for future studies to explore their functions.

Holley's studies showed that the RNA sequence is only half of the equation; the structure adopted by tRNAs was essential to carry out its function in translation. Carl Woese, another giant in the history of science, was interested in deducing structure from sequence. Indeed, he obtained the sequence of ribosomal RNAs (rRNAs) from diverse organisms and used sequence alignment to deduce their secondary structures (Woese and Pace 1993). In particular, Woese was interested in identifying “molecular fossils” to rigorously trace the origins and associations of secondary structures. Associations of these molecular fossils to each other defined the tree of life, a controversial view at the time. Before Woese's work, organisms were simply classified by having or not having a nucleus. We now recognize that the phylogenetic tree of life can be classified into bacteria, archaea, and eukarya, not simply prokaryotes and eukaryotes. Holley's and Woese's studies as well as the discovery of catalytic RNAs (Stark et al. 1978; Zaug and Cech 1986) advanced the concept of an RNA world.

The studies described above inspired a keen interest in developing approaches to experimentally map RNA structure (Woese et al. 1980). The same Gilbert–Maxam reagents for DNA sequencing also react with RNA and can be tuned to identify unpaired bases (Peattie and Gilbert 1980).

Studies by Cech (Inoue and Cech 1985) and Noller (Stern et al. 1986) used these reagents and showed that reverse transcriptase (RT) cannot proceed through modified bases, allowing one to use sequencing methods to read out unpaired bases. Computational approaches have also emerged to annotate RNA structure from sequence. These studies use the energetics of base pairing to predict the thermodynamically most stable RNA structure from its sequence and can also predict suboptimally folded (i.e., less energetically stable) structures. Indeed, computational algorithms have accurately predicted RNA structure as shown in the work of Turner (Jaeger et al. 1989) and Zuker (Zuker 1989). These advancements were followed by efforts to combine experimental and computational approaches to further refine these predictions, including those generated by cell-permeable chemical modification reagents that map RNA structure in cells (Wells et al. 2000; Mathews et al. 2004). The combination of computation and experiment has elucidated the structures of entire viral genomes (Deigan et al. 2009; Watts et al. 2009) and identified patterns in RNA structure within the human transcriptome (Bevilacqua et al. 2016; Lu and Chang 2016). Mapping the sequence and structure of RNAs depicted the tree of life in the RNA world and established the foundation for recent advancements in drugging RNA as we will discuss later in this review.

2 TARGETING RNA WITH SMALL MOLECULES

2.1 Ribosomal RNA and Riboswitches as Targets of Small Molecules

Streptomycin was the first-in-line treatment for tuberculosis infections and was isolated from bacteria, along with other natural products such as neomycin, in 1943 by Albert Schatz while he was a Ph.D. student in the laboratory of Selman Waksman (Jones et al. 1944). Both aminoglycoside antibiotics have served as invaluable drugs to treat otherwise fatal bacterial infections by inhibiting protein synthesis. Interestingly, streptomycin ushered in several historical advances in treatment and clinical trials. For example, it was used in the first double-blind trial ever completed to gain approval for use in humans. However, resistance of tuberculosis to streptomycin treatment was observed shortly after its introduction.

Noller and Moazed applied chemical probing methods to gain insight into the binding of antibiotics, including aminoglycosides, to the ribosome (Moazed and Noller 1987). Several features were identified among ligand binding sites, most notably (1) resistance correlated with mutations in the rRNA-binding site, and (2) these mutations were often located in functional sites, namely, the tRNA-binding site and the decoding site.

Figure 1. The ribosome is a well-known target of small molecules. (A) Various compounds bind to specific sites in the bacterial ribosome and protect them from chemical modification. High-resolution structures of the ribosomal subunits have been determined in the presence and absence of antibiotics. Shown here is the 30S subunit complexed with the aminoglycoside puromycin; the ribosome subunit is shown in purple and puromycin is shown in Corey, Pauling, and Kolton (CPK) colors. (B) Common secondary structural elements found in cellular RNAs.

Footprinting and biophysical analysis by Puglisi and colleagues further showed that the binding of antibiotics to the rRNA aminoacyl (A)-site affected decoding by stabilizing the binding of noncognate tRNAs (Yoshizawa et al. 1999). Subsequently, high-resolution structures of the rRNA particle in complex with antibiotics by the Ramakrishnan group confirmed Puglisi's studies completed with the A-site model (Fig. 1) (Carter et al. 2000). Interestingly, an A-site model was also used to show that antibiotic potency is correlated with altering A-site dynamics, not affinity (Kaul et al. 2006). Collectively, these studies established the sites of antibacterial action and showed that smaller RNA models are effective tools for studying small-molecule binding and can provide insight into therapeutic action.

Riboswitches have also emerged as tried-and-true targets of small molecules affecting RNA. Henkin and Breaker identified these small regulatory RNAs, primarily found in bacteria, that switch conformation in the presence of sufficient concentrations of small-molecule metabolites (Winkler et al. 2002). This conformational switch inhibits the synthesis of proteins responsible for production of the metabolite (Nahvi et al. 2002; Grundy and Henkin 2004). Structural mimicry approaches have been used to design riboswitch inhibitors that short-circuit metabolic synthesis pathways (Blount et al. 2007) and hence serve as antibacterials. Screening has also been used to identify inhibitors of riboswitches. For example, scientists at Merck identified ribocil as targeting bacterial riboflavin riboswitches by using a phenotypic screen (Howe et al. 2015). Studies show that the compound acts as a structurally distinct synthetic mimic of the natural ligand, flavin mononucleotide. Taken together, studies of the ribosome and riboswitches showed that RNA is a bona fide target of small molecules and that ligands can indeed bind discrete sites in RNA.

3 USING RNA STRUCTURE TO DESIGN SELECTIVE SMALL MOLECULES TARGETING RNA

3.1 The Transcriptome as a Source of Druggable RNAs: Potential and Perceived Risks

Although antibacterial agents targeting the ribosome have had an enormous impact on biomedical research and have provided invaluable chemical tools to understand biology, progress in targeting other types of RNAs has been modest to date. A large body of work has generated small molecules that target viral RNAs, particularly those from human immunodeficiency virus (HIV) and hepatitis C, among others (Hamy et al. 1997; Parsons et al. 2009), squarely cementing the applicability of small-molecule treatment to target RNAs involved in infectious disease.

There are few examples, however, of small molecules that target human RNAs associated with disease. Recently, Novartis and PTC/Roche identified compounds that modulate the splicing of a precursor mRNA (pre-mRNA) for the treatment of spinomuscular atrophy (Naryshkin et al. 2014; Palacino et al. 2015). Novartis' compound is thought to bind and stabilize an RNA structure at an exon-intron junction, thereby affecting exon inclusion. Other small molecules have been found to target pre-mRNAs and affect splice site selection (Luo and Disney 2014) or stimulate readthrough of premature termination codons in various diseases, particularly Duchenne muscular dystrophy (Welch et al. 2007), although the mechanism of action has been controversial (Auld et al. 2009).

Advancing small-molecule targeting of human RNAs is particularly timely as it is now appreciated that there are many more potential RNA targets than protein targets. Recent estimates suggest that 1%–2% of the human genome is translated into protein, yet >50% is transcribed into RNA (Clamp et al. 2007; Gerstein et al. 2007). Further, only ~15% of proteins are in families that are druggable

(Hopkins and Groom 2002). Although many RNAs may not be functional, these Encyclopedia of DNA Elements (ENCODE) studies have revealed that there is a large amount of druggable space in the transcriptome. The studies described herein are changing the perceived notion that RNA is undruggable.

Why has RNA been considered undruggable with small molecules in the past? The most highly explored and studied targets have complex three-dimensional structures with “protein-like” folds, which were thought to be required for selective small-molecule binding. Most RNAs, including microRNAs (miRNAs) and mRNAs, have defined secondary structure (see Fig. 1B for examples of secondary structures) but limited tertiary structure and were thus considered untargetable and not suitable for drug discovery efforts. It is an often-missed irony that many known drugs are not “drug-like” based on Lipinski guidelines (Lipinski 2004), which are often over- and/or misstated. Considering RNA as undruggable is of the same nature.

3.2 RNA Secondary Structure and Disease: RNA-Binding Proteins

RNA secondary structure regulates many cellular processes including (1) pre-mRNA alternative splicing (recognition of intron–exon junctions by splicing regulators), (2) translation (ribosomal loading controlled by secondary structure, including G-quadruplexes, and 5′ and 3′ untranslated regions [UTRs]), (3) cellular trafficking and delivery (by binding to RNA-binding proteins [RBPs]), (4) RNA editing, and (5) miRNA biogenesis. Notably, disease occurs when RNA secondary structure is altered by mutation or repeated too frequently, as observed in microsatellite (RNA repeat expansion) disorders. Thus, targeting RNA–protein interactions that lead to cellular dysfunction, in particular those that involve defined RNA secondary structure, could provide a therapeutic strategy. For a comprehensive review of how RNA secondary structure influences cellular processes and disease, please see Bernat and Disney (2015). Herein, we focus on small-molecule targeting of human RNAs that interact with proteins. Many excellent reviews are available that discuss targeting of other RNAs (Thomas and Hergenrother 2008; Guan and Disney 2012), including specialized reviews on targeting the bacterial ribosome (Yonath 2005; Blanchard et al. 2010), riboswitches (Blount and Breaker 2006), and viruses (Hermann 2016).

3.3 Identifying Selective RNA Motif–Small Molecules via Two-Dimensional Combinatorial Screening (2DCS)

Collectively, the studies that target mRNAs described above suggest that (1) targeting RNA secondary structure

is relevant to a wide variety of diseases, particularly those in which a secondary structure binds a protein, and (2) human RNAs with extensive secondary structure but limited tertiary structure can indeed be targeted with small molecules. To do so by rational design, two types of data are required: the RNA’s secondary structure, or composite of motifs, and a database of small molecules and the RNA motifs that are privileged for binding them. As described in the Introduction, accurate models of RNA secondary structure can be generated by a combination of free-energy minimization, experimental constraints, and phylogenetic alignment. Unfortunately, there has been a dearth of information about privileged chemical space for RNA and targetable RNA motifs. To provide data of this type, a library-versus-library screen approach dubbed two-dimensional combinatorial screening (2DCS) was developed (Fig. 2) (Disney et al. 2008). The 2DCS approach allows the assignment of the RNA motif–binding landscapes for small molecules (Velagapudi et al. 2010, 2017).

In a 2DCS experiment, a library of small molecules is immobilized onto an agarose-coated microarray. The agarose substrate was carefully chosen as it allows (1) spatial encoding of the small molecules, (2) increased loading of the compounds on the surface, and (3) excision of bound RNAs. The small-molecule microarrays are then incubated with an RNA motif library that displays a randomized region in the pattern of a discrete secondary structure, whether a bulge, internal loop, or hairpin (Fig. 1B). Importantly, incubations are completed under highly stringent conditions; that is, a large excess of oligonucleotides that mimic regions common to all members of the RNA library is included during hybridization. These competitor oligonucleotides are either single-stranded or duplex RNAs that cannot be amplified by primers selective for members of the RNA motif library. Bound RNAs are excised from the array surface, amplified, and sequenced.

Selected RNA motif–small-molecule interactions are assigned a fitness of binding via a scoring function, a method named structure–activity relationships through sequencing (StARTS) and its high-throughput variant using RNA-seq, HiT-StARTS (Velagapudi et al. 2010, 2017). A scoring function is defined by sequencing the starting library and comparing the frequency of each of the individual motifs with the frequency of the same motif in the selected RNA pool. This pooled population comparison affords a Z_{obs} score for each RNA motif–small-molecule pair. RNA–small-molecule motifs with larger Z -scores bind with higher affinity than ones with lower ones, defining a binding landscape that affords the scoring function for each selection. Further, relative Z_{obs} scores between

Figure 2. Sequence- and secondary structure-based approaches developed for small molecules targeting RNA. (A) Schematic of two-dimensional combinatorial screening (2DCS), a library-versus-library screen. A small-molecule library is immobilized onto a microarray and hybridized with an RNA motif library. The selected RNAs are sequenced and analyzed to form a chemical code for selectively targeting RNA. (B) Schematic of Inforna, a computational method to identify “druggable” RNA targets from sequence. The selected RNA motif–small-molecule partners from 2DCS are mined against the transcriptome to agnostically identify RNAs that can be targeted with these small molecules.

RNA motif–small-molecule partners can rapidly define ligand selectivity.

Collectively, 2DCS studies have shown that a myriad of ligands can bind RNA with affinities that range from low nanomolar to mid-micromolar (Disney et al. 2008; Velagapudi et al. 2010, 2017; Tran and Disney 2012). Furthermore, compound classes that bind RNAs range from aminoglycosides to heterocyclic and drug-like small molecules. Last, a broad swath of RNA motifs can selectively bind small molecules, including symmetric and asymmetric internal loops, bulges, and hairpins (Fig. 2). Thus, there are tremendous opportunities to exploit both a wide variety of chemistries and a diverse array of RNA structural motifs when drugging RNA, an irony given the aforementioned long history of skepticism concerning the “druggability” of RNA.

3.4 Merging the Results of 2DCS with Bioinformatics: Inforna, a Sequence-Based Approach to Design Small Molecules Targeting RNA

To design small molecules that specifically target RNA structures based on its sequence, we developed a computational approach dubbed Inforna (Velagapudi et al. 2014; Disney et al. 2016). Inforna integrates RNA motif–small-molecule interactions identified via 2DCS and structural data of target RNAs. That is, it generates lead compounds for an RNA of interest by comparing the motifs found in the RNA target’s structure with the RNA motif–small-molecule interactions in our database. Because all RNAs have secondary structures, the approach is broadly applicable to identify rationally selective small-molecule chemical probes and lead medicines targeting RNA.

3.5 Targeting Protein-Binding Sites in Disease-Causing miRNAs

miRNAs are small, noncoding RNAs that regulate gene expression by binding to the 3' UTRs of mRNAs with complementary sequences (Fire et al. 1998; Lagos-Quintana et al. 2001; Lee and Ambros 2001). Transcribed in a primary (pri-) form, miRNAs are processed in two steps:

1. Cleavage of 5' and 3' sequences by the nuclear nuclease Drosha to afford pre-miRNAs that are exported to the cytoplasm.
2. Cleavage of the pre-miRNA's hairpin by the cytoplasmic nuclease Dicer (Hammond et al. 2000).

Nuclease processing sites represent functional motifs, and inhibition of cleavage by Drosha or Dicer prevents biogenesis of the miRNA. Below, we describe the development of small molecules that have effectively inhibited maturation of disease-causing miRNAs by binding to Drosha or Dicer processing sites.

In the first broad application of Inforna (Velagapudi et al. 2014), all human pre-miRNAs were folded and compared with the RNA folds to which small-molecule-binding partners had been identified via 2DCS. Thus, the approach is target-agnostic rather than target-centric, the latter occurring when a single or a few target biomolecules are screened against a chemical library. Inforna identified compounds that target nearly two dozen disease-associated miRNAs that have broad therapeutic implications. Nearly half of these interactions were bioactive without lead optimization. Activities of the compounds ranged from triggering apoptosis in breast cancer cells (Velagapudi et al. 2014), short-circuiting hypoxic signaling in breast cancer cells (Costales et al. 2017), and inhibiting the invasive properties of hepatocellular carcinoma (Childs-Disney and Disney 2016).

The first target vetted by Inforna was a small molecule that targets the Drosha site of pri-miR-96 (Velagapudi et al. 2014). In triple-negative breast cancer, the miR-96-182-183 cluster is up-regulated. Two of the miRNAs, miR-96 and miR-182, repress the pro-apoptotic transcription factor forkhead box protein O1 (FOXO1). FOXO1 controls transcription of proapoptotic members of the B-cell lymphoma 2 (Bcl-2) family. In vitro, the lead compound bound the Drosha site with a 2 μM dissociation constant and inhibited Drosha processing of a model pri-miRNA transcript (Velagapudi et al. 2014). Treatment of Michigan Cancer Foundation-7 (MCF-7) breast cancer cells with 40 μM compound decreased levels of mature miR-96 by 90% but did not significantly affect any of the other miRNAs in the miR-96-182-183 cluster. Additionally, the compound increased the amount of the pri-miR-96, indicating that inhibition of

Drosha processing of the desired target is the mode of action. Application of the compound (40 μM) derepressed FOXO1 protein, increasing its levels by approximately two-fold (Velagapudi et al. 2014).

Importantly, small-molecule inhibition of miR-96 had a functional effect, triggering apoptosis of MCF-7 breast cancer cells (Velagapudi et al. 2014). To investigate whether apoptosis was triggered via the FOXO1-miR-96 pathway, the FOXO1 mRNA was knocked down via a small interfering RNA (siRNA), which should ablate compound activity. Indeed, knockdown of FOXO1, but not glyceraldehyde 3-phosphate dehydrogenase (GAPDH), mRNA decreased the compound's activity by $\sim 70\%$, as measured by induction of apoptosis. Collectively, these studies indicate that the small molecule short-circuits the FOXO1-miR-96 pathway (Velagapudi et al. 2014).

The data described above have important biological implications.

1. It is well known that multiple miRNAs can affect a single mRNA and single miRNAs can affect multiple mRNAs. Yet, in the case of miR-96 and cancer, it appears that inhibition of a single miRNA targeting FOXO1 is sufficient to reverse a disease-associated phenotype. Such effects would be difficult to determine using antisense oligonucleotides because the sequences of miR-96 and miR-183 are so similar that an oligonucleotide would target both.
2. We compared the selectivity of the small molecule with an antisense agent for reducing the level of mature miRNAs, and the small molecule is at least as selective, if not more selective at this level (Velagapudi et al. 2014).

Although our anti-miR-96 small molecule selectively inhibited biogenesis of miR-96, its activity was not sufficient to test in vivo. Thus, we used Inforna to identify compounds that bind to sites near the Drosha site of pri-miR-96 (Velagapudi et al. 2016). By linking the two compounds together, a higher affinity binder was designed named Targaprimir-96 (Fig. 3A). This compound had a >200 -fold enhancement in avidity for the target sites in the pri-miR-96 with a K_d of 20 nM. Application of the Targaprimir-96 to MDA-MB-231 triple-negative breast cancer cells selectively inhibited production of miR-96 and boosted pri-miR-96 levels (Velagapudi et al. 2016). The molecule selectively induced apoptosis in breast cancer cells in which miR-96 is up-regulated but did not affect healthy breast cells in which miR-96 is expressed at normal levels. Importantly, this dimeric molecule slowed tumor growth in a xenograft mouse model, reduced mature miR-96 levels, and increased FOXO1 expression in tumors (Velagapudi et al. 2016). This is the first example of a small molecule that can inhibit tumor growth by targeting a noncoding RNA.

Figure 3. Examples of targeting RNA with small molecules that have been designed by using Infora. (A) Targaprimir-96 targets the Drosha site in the microRNA-96 (miR-96) hairpin precursor. The compound inhibits miR-96 biogenesis and triggers apoptosis via a miR-96-forkhead box protein O1 (FOXO1) circuit. (B) Targapremir-210 silences miR-210, induces apoptosis of hypoxic, but not normoxic, cells, and reduces tumor burden in a xenograft mouse model.

miRNAs can also be up-regulated in response to changes in the cellular environment, for example when cancer cells are in a low oxygen (hypoxic) environment, which leads to resistance to chemotherapeutics. During the transition from normoxia to hypoxia, miR-210 is up-regulated and forms a hypoxic circuit with hypoxia-induc-

ible factor 1- α (HIF-1 α) (Semenza 2003). In particular, miR-210 stimulates a hypoxic response as it targets the 3' UTR of glycerol-3-phosphate dehydrogenase 1-like enzyme (GPD1L); GPD1L stabilizes prolyl hydroxylase (PHD). Under normoxic conditions, PHD hydroxylates prolines on HIF-1 α that causes it to be proteolyzed by the proteasome.

In the absence of stable PHD, HIF-1 α dimerizes with HIF-1 β , which transcriptionally regulates genes that contribute to cancer metastasis (Semenza 2003).

Infora identified a lead compound, Targapremir-210, that bound to the Dicer site of pre-miR-210 (Costales et al. 2017). The compound reduced levels of mature miR-210 with concomitant increase in the amount of pre-miR-210. Expected changes in GDP1L (increased) and HIF-1 α (decreased) levels were observed as well (Costales et al. 2017) to extents similar to those observed on anti-miR-210 antagomir treatment (an antagomir is an antisense oligonucleotide). Further, application of Targapremir-210 to hypoxic cells, but not to normoxic cells, caused the cells to undergo apoptosis (Fig. 3B). Profiling experiments showed that the small molecule was as selective as the oligonucleotide. Furthermore, forced overexpression of pre-miR-210 ablated the ability of Targapremir-210 to stimulate apoptosis (Costales et al. 2017). Thus, small molecules can target one portion of a biological circuit, in particular in a hypoxic circuit, and affect biological outcome, in this case stimulation of apoptosis.

One of the fastest routes to push a compound into clinical development is to repurpose a known drug with established tolerability and toxicological profiles. Thus, 2DCS was used to profile the RNA-binding abilities of known drug derivatives to infer new targets. Interestingly, 5'-azido neomycin B binds the Drosha processing site in pre-miR-525 (Velagapudi et al. 2014; Childs-Disney and Disney 2016). This miRNA is up-regulated in ~60% of liver cancers and confers invasive properties to hepatocellular carcinoma (HCC) cells by silencing zinc finger protein 395 (ZNF395). Application of 5'-azido neomycin B and its Food and Drug Administration (FDA)-approved derivative neomycin B to

liver cancer cells silenced production of mature miR-525 and enhanced production of ZNF395 (Childs-Disney and Disney 2016). The compound also effectively modulated the invasive phenotype mediated by miR-525. Thus, FDA-approved drugs can inhibit miRNA-associated phenotypes in designed and programmable ways. Interestingly, neomycin B is a second-line agent for hepatic encephalopathy (HE) and bacterial infections due to cirrhosis.

3.6 Profiling RNA–Small-Molecule Interactions in Cells: Chemical Cross-Linking and Isolation by Pull-Down (Chem-CLIP)

One major void in the development of small molecules that target RNA is established cellular target recognition profiling methods. Work of the Darnell group has developed cross-linking and immunoprecipitation (CLIP), which cross-links proteins to RNA to identify binding partners (Ule et al. 2003, 2018). In CLIP, ultraviolet (UV) light is used to cross-link protein–RNA complexes, which are purified with antibody-functionalized beads and analyzed by sequencing. CLIP has been an invaluable tool used to study multiple pathways, including identification of protein-binding sites on pre-mRNAs that influence splicing outcomes.

As light can limit the applicability of a CLIP-based approach, we developed Chem-CLIP in which the cross-linking event is a proximity-based chemical reaction; that is, light is not required (Fig. 4) (Guan and Disney 2013). Small molecules that target RNA have been appended with a nucleic acid–reactive module (chlorambucil) and a biotin purification tag, or Chem-CLIP probes. The RNA-binding modules drive binding to their RNA targets in cells, effec-

Figure 4. Schematic of chemical cross-linking isolated by pull-down (Chem-CLIP), an approach to identify the RNA targets of small molecules in cells. (A) Small molecules targeting RNA are appended with cross-linking and biotin modules. Application of Chem-CLIP probes to cells cross-links the small molecule to its cellular RNA targets. The biotin tag allows the targets to be purified from cells, which are identified by quantitative reverse transcription polymerase chain reaction (RT-qPCR) or RNA sequencing (RNA-seq). (B) In competitive Chem-CLIP (C-Chem-CLIP), cells are cotreated with the Chem-CLIP probe and the parent compound. Bona fide targets are depleted in the pulled-down fraction.

tively biotinylating them upon cross-linking. By using streptavidin, the small molecule–RNA complexes can be isolated and analyzed by using RNA-seq or quantitative reverse transcription polymerase chain reaction (RT-qPCR). Chem-CLIP not only identifies bound RNAs but also can identify ligand-binding sites within the transcripts in cells (Yang et al. 2015; Rzuczek et al. 2017). In fact, Chem-CLIP has shown that small molecules can be allele-selective when designed to target expanded RNA repeats (Rzuczek et al. 2017).

Chem-CLIP has been broadly applied to study the druggability of noncoding miRNA precursors. For example, the Targapremir-210 Chem-CLIP probe selectively reacted with pre-miR-210 in vitro, whereas a control compound lacking the RNA-binding module did not (Costales et al. 2017). In cells, the compound pulled down miR-210 from hypoxic triple-negative breast cancer cells, as detected by an enrichment in the pulled-down fraction as compared with unpurified lysate. To study the selectivity of Targapremir-210, we studied ligand occupancy via Chem-CLIP of various abundant noncoding RNAs, abundant mRNAs, hypoxia-associated miRNAs, and miRNAs that have potential ligand-binding sites for Targapremir-210. The latter we dub “RNA isoforms” or the RNA targets that have motifs predicted by Inforna to bind the lead small molecule, including even those that might have lower affinities. Importantly, these Chem-CLIP experiments showed that miR-210 is the most enriched RNA in the pulled-down fractions. For the hypoxia-associated miRNAs, miR-210 was the most pulled down, by at least threefold (Costales et al. 2017). Further, 14 other miRNA precursors contained potential binding sites for Targapremir-210, including miR-497, which has the same motif as miR-210 in its Dicer site. None of these miRNAs was enriched in the pull-down fraction (Costales et al. 2017).

To assess the selectivity of Targapremir-210 on a transcriptome-wide scale, we studied its effect on 92 highly abundant transcripts in the pulled-down fraction including rRNAs, small RNAs (sRNAs), tRNAs, and mRNAs. Only six RNAs were enriched in the pulled-down fractions (Costales et al. 2017). Competitive-Chem-CLIP (C-Chem-CLIP) experiments were completed to assess if any of these targets were pulled down owing to some lack of selectivity of the cross-linker. Bona fide targets will be depleted in the pull-down fraction when an excess of a competitive, unreactive ligand is added; targets that non-selectively reacted with the cross-linker will not be competed off by the parent compound. Indeed, a few of the enriched targets were not competed off, suggesting some nonspecific binding/cross-linking due to chlorambucil or biotin modules of the Chem-CLIP probe (Costales et al. 2017). Collectively, these results suggest that bioactivity of a small molecule is influenced by (1) its affinity for an RNA

target (of note, the affinity of Targapremir-210 for miR-210's Dicer site is 165 nM, whereas its affinity for other motifs ranged from 2.5 to 11 μ M); (2) the expression level of the target, as miR-497 is expressed at 10-fold lower levels than miR-210; and (3) whether the small-molecule binding site is in the Dicer or Drosha processing site.

3.7 Targeting Protein-Binding Sites in RNA Repeat Expansions

Microsatellite diseases are a class of neurological and neuromuscular disorders caused by repeat expansions (Orr and Zoghbi 2007). Repeat expansions are located in (1) open reading frames (ORFs; Huntington's disease [HD]), (2) UTRs (myotonic muscular dystrophy type 1 [DM1], fragile X-associated tremor ataxia syndrome [FXTAS], and fragile X syndrome [FXS]), and (3) introns (amyotrophic lateral sclerosis [ALS] and myotonic dystrophy type 2 [DM2]). For those diseases caused by sequences that lie in noncoding regions, a major driver of disease-associated toxicity can be the RNA transcribed from the repeat expansion, in particular its sequestration of RNA-binding proteins. In the cases of DM1, DM2, and FXTAS, the RNA repeat expansions bind and inactivate proteins that regulate alternative pre-mRNA splicing (Fardaei et al. 2002; Disney et al. 2012; Sellier et al. 2013). It is likely that the RNA repeat expansion that causes ALS also sequesters a splicing regulator; however, it is currently thought that toxicity is mainly driven by production of toxic dipeptide repeats via repeat-associated non-ATG (RAN) translation (Mori et al. 2013; Zu et al. 2013; Zhang et al. 2014) and defects in nuclear pore transport (Zhang et al. 2015). This disease pathomechanism, sequestration of RNA-binding proteins, provides a therapeutic strategy in which a small molecule binds the RNA repeat expansion and inhibits sequestration of RNA-binding proteins such that they can complete their normal cellular functions.

Various groups have tackled targeting RNA repeat expansions with small molecules, whether by design or screening. Indeed, compounds have been developed that improve disease-associated defects in cells and in *Drosophila* and mouse models. The most well-studied RNA repeat expansion is the trinucleotide repeat that causes DM1, r(CUG)^{exp}, which binds and sequesters the RBP splicing regulator muscleblind-like 1 (MBNL1), among others (Fardaei et al. 2002). Located in the 3' UTR of the dystrophin myotonia protein kinase (DMPK) mRNA (Brook et al. 1992), r(CUG)^{exp} folds into a hairpin structure with periodically repeated 1 $\times$ 1 UU internal loops, the binding sites for MBNL1. Compounds designed by various laboratories have focused on engineering selective binding to these loops. For example, using a structure of r(CUG) repeats, the Zimmerman lab-

oratory designed a small molecule comprised of an acridine (well-known DNA intercalators) and a triaminotriazine that interacts with the RNA via Janus-wedge hydrogen bonding (Arambula et al. 2009). Optimization of this first-generation compound afforded three bioactive compounds that improved missplicing and formation of nuclear foci in a DM1 cellular model at concentrations between 50 and 100 μM ; one second-generation compound improved defects in a *Drosophila* model (Jahromi et al. 2013).

Using Inforna, our laboratory has developed various compounds that improve DM1-associated defects in cellular (including DM1 patient-derived cells) and mouse models (Childs-Disney et al. 2012; Parkesh et al. 2012; Rzuczek et al. 2013, 2017). First-generation compounds were based on a *bis*-benzimidazole RNA-binding module (H) that were multivalently displayed on a peptoid scaffold; the optimal compounds display H modules at the precise distance that separated UU loops in $r(\text{CUG})^{\text{exp}}$, or $n\text{H}-4$ (in which n indicates the number of H RNA-binding modules and 4 indicates that four propylamine spacing modules separate the Hs). Indeed, 2H-4 ameliorated a DM1-associated splicing defect with a half maximal inhibitory concentration (IC_{50}) of $\sim 10 \mu\text{M}$; note the H module itself was inactive (Childs-Disney et al. 2012). We also studied the ideal backbone for display of H RNA-binding modules and identified that the peptide tertiary amide 2H-K4NMeS was superior (Rzuczek et al. 2013). The backbone was further improved to enhance metabolic stability to afford 2H-K4NMeS, a 100 nM inhibitor of DM1-associated toxicity in patient-derived cells (Rzuczek et al. 2017). Likewise, the H RNA-binding module was optimized via chemical similarity searching to afford H1, which improved DM1-associated defects in cellular and mouse models (Parkesh et al. 2012).

By far the most potent small molecule that targets $r(\text{CUG})^{\text{exp}}$ reacts to form multivalent compounds upon binding cellular $r(\text{CUG})^{\text{exp}}$ with an IC_{50} of $\sim 10 \text{ nM}$ in DM1 patient-derived cells (Rzuczek et al. 2017). That is, 2H-K4NMeS was modified to contain an alkyne and azide (Aak-2H-K4NMeS- N_3). When two or more Aak-2H-K4NMeS- N_3 bind adjacent sites in $r(\text{CUG})^{\text{exp}}$, the azides and alkynes undergo a proximity-based click reaction; that is, the RNA serves as the catalyst, and no copper is required. This *in situ* click approach was first shown using the RNA that causes DM2.

Potency is likely driven by affinity of the small molecule for the RNA, the large amount of RNA's surface area occupied by the ligand, and the small molecule's k_{on} and k_{off} rates, particularly in comparison with the RBP. Importantly, Aak-2H-K4NMeS- N_3 is selective for $r(\text{CUG})^{\text{exp}}$ more than other RNAs containing short, nonpathogenic $r(\text{CUG})$ repeats including the wild-type (WT) DMPK allele

(Rzuczek et al. 2017). Selectivity is likely mainly derived from Aak-2H-K4NMeS- N_3 's recognition of the $r(\text{CUG})^{\text{exp}}$ structure; short $r(\text{CUG})$ repeats generally do not fold into an array of 1×1 UU internal loops. These factors that influence potency and selectivity are likely general for all RNA repeat expansions that sequester RNA-binding proteins.

3.8 Chemical Techniques to Image RNAs and to Affect Subcellular Localization and Function

Green fluorescent proteins (GFPs) have revolutionized the manner in which protein function and localization are studied (Stearns 1995). Various genetic approaches have been used to engineer RNA GFP mimics, the most successful of which appends an RNA of interest with a fluorescent aptamer (Paige et al. 2011). Chemical approaches to image cellular RNAs in their natural context, without genetic modification, are highly desirable. To this end, the on-site, click chemistry approach described above was modified to synthesize a turn-on fluorescence resonance energy transfer (FRET) sensor upon binding to $r(\text{CUG})^{\text{exp}}$ (Rzuczek et al. 2017). To allow for the synthesis of the fluorescent probes on-site, two derivatives of Aak-2H-K4NMeS- N_3 were synthesized: Aak-2H-K4NMeS-FAM (in which FAM is fluorescein) and TAMRA-2H-K4NMeS- N_3 (in which TAMRA is 5-carboxytetramethylrhodamine). Thus, on binding $r(\text{CUG})^{\text{exp}}$, Aak-2H-K4NMeS-FAM and TAMRA-2H-K4NMeS- N_3 click together such that the two dyes are close together and FRET is observed (a readout of reaction). Robust reaction to form the FRET sensor was observed in DM1 patient-derived cells, whereas FRET was not observed in cells that did not express $r(\text{CUG})^{\text{exp}}$ as such (i.e., FRET was $r(\text{CUG})^{\text{exp}}$ -dependent). When cells were incubated with Aak-2H-K4NMeS-FAM and 2H-K4NMeS- N_3 (lacking FRET acceptor), no FRET was observed, showing that stochastic alignment of the compounds on the target did not result in FRET (Fig. 5) (Rzuczek et al. 2017). In DM1, sequestration of RBPs by $r(\text{CUG})^{\text{exp}}$ disallows the RNA's export from the nucleus and thus decreases translation of DMPK. Importantly, the FRET sensor shows that the small molecule alleviated this nucleocytoplasmic transport defect, trafficking the mutant DMPK RNA harboring $r(\text{CUG})^{\text{exp}}$ to the cytoplasm and increasing its translation (Rzuczek et al. 2017). Thus, small molecules can affect subcellular localization of mRNAs and impact biological function, such as translation.

3.9 Targeting Protein-Binding Sites in Human Telomerase RNA

First discovered by Greider and Blackburn (Greider and Blackburn 1985), telomerase is a ribonucleoprotein complex (RNP) responsible for addition of telomeric DNA to

Figure 5. An example of on-site click chemistry to create a fluorescence resonance energy transfer (FRET) reporter to image RNA in cells without genetic modification. (A) Structures of an RNA-binding module (represented by a blue circle), 2H-K4NMeS, Aak-2H-K4NMeS-FAM (a green circle represents a fluorescein FRET donor), and TAMRA-2H-K4NMeS-N₃ (a purple circle represents TAMRA [5-arboxytetramethylrhodamine] FRET acceptor). (B) Scheme of on-site FRET pair synthesis that occurs after FRET pair precursors bind to r(CUG)^{exp} and undergo an RNA-catalyzed click reaction, allowing imaging of RNA in cells. (C) Images of myotonic muscular dystrophy type 1 (DM1) (CUG)₅₀₀ fibroblasts and healthy (CUG)₁₅ fibroblasts with Aak-2H-K4NMeS-FAM only (left), Aak-2H-K4NMeS-FAM with TAMRA-2H-K4NMeS-N₃ (middle), and Aak-2H-K4NMeS-FAM with TAMRA-2H-K4NMeS (lacks N³ click partner) (right). (D) On-site click chemistry generates FRET reporters to show 2H-K4NMeS selectively binds to r(CUG)₅₀₀ in DM1 fibroblasts. ***P < 0.001, as determined by a two-tailed Student's *t*-test. (A and B are adapted, whereas C and D are directly reproduced, with permission, from Rzuczek et al. 2017.)

chromosome ends, providing genome stability. Telomerase RNA adopts a complex secondary structure, comprising the template for telomeric DNA synthesis and elements that bind telomerase proteins (Romero and Blackburn 1991). Interestingly, telomerase is highly expressed in stem cells and cancer cells but has very low expression and activity in somatic cells (Feng et al. 1995). Collectively, these data

suggest that a small molecule that inactivates telomerase in cancer cells could serve as a chemotherapeutic and could do so selectively, that is, healthy cells would not be affected as the target is absent. Conversely, underexpression of telomerase leads to dyskeratosis congenita (DKC) and is often due to mutations in the RNA component (Theimer et al. 2003) that inhibit its binding to telomerase proteins.

Although DKC often manifests itself as changes in skin pigmentation and nail formation, it is often fatal as a result of bone marrow failure. Thus, an effective treatment for DKC could be achieved by stabilizing the functional fold of telomerase RNA with an oligonucleotide or small molecule.

Several laboratories have identified small molecules that bind regions within telomerase RNA to prevent binding to telomerase protein or to the telomeric DNA substrate. The James laboratory docked more than 3000 FDA-approved drugs into the structure of the telomerase P2b region (Theimer et al. 2003), affording 20 potential binders (Pinto et al. 2008). Thirteen compounds bind the RNA, as evidenced by nuclear magnetic resonance (NMR) spectroscopy. By using a combination of chemical similarity searching and docking, 24 additional small molecules were identified that selectively bound telomerase RNA. The selective small molecules fall into four main categories: chained aromatics, 4-aminoquinolines, phenothiazines, and 6-7-6 ring systems (Pinto et al. 2008), which could be further medicinally optimized. Other small molecules that bind telomerase RNA and inhibit protein or DNA binding include Hoechst 33258 (Dominick et al. 2004), various aminoglycosides (Dominick et al. 2004), Rivanol (DNA/RNA hybrid) (West et al. 2001), and acridines (DNA/RNA hybrid) (West et al. 2001).

3.10 Targeting Protein-Binding Sites in pre-mRNAs to Affect Splicing Outcomes

Many diseases are caused by aberrant alternative splicing of pre-mRNAs. For those caused by mutation of the RNA itself, a small molecule could offer a therapeutic option. For example, inherited frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17) is an incurable disease that is directly caused by increased inclusion of exon 10 in microtubule-associated protein τ (MAPT) mRNA (Jiang et al. 2000). Various studies have shown that mutations within a hairpin at the MAPT exon 10-intron junction destabilize the RNA's thermodynamic stability, increasing both binding to U1 small nuclear ribonucleoprotein complex (snRNP) and hence exon 10 inclusion (Jiang et al. 2000; Donahue et al. 2006). Therefore, small molecules were designed that bind and stabilize this mutant MAPT and improve deregulation of alternative splicing (Luo and Disney 2014), the first small molecules known to affect exon inclusion/exclusion in cells by binding directly to the RNA target. Furthermore, the compounds were selective, as introduction of a single nucleotide to the small-molecule-binding site eliminated binding and bioactivity (Luo and Disney 2014).

Spinal muscular atrophy (SMA) is caused by a loss of function of the survival motor neuron 1 protein (SMN1).

Interestingly, the sequence of the survival motor neuron 2 (SMN2) gene differs from SMN1 by a single nucleotide; this single nucleotide change increases the rate of exon 7 exclusion. The absence of exon 7 decreases SMN2 mRNA's half-life by twofold, as compared with SMN1. Thus, increasing exon 7 inclusion in SMN2 mRNA with a small molecule could be therapeutically beneficial. Indeed, this strategy is viable as first explored with antisense oligonucleotides by Adrian Krainer's laboratory (Cartegni and Krainer 2003). Since then, PTC Therapeutics/Roche and Novartis have separately developed small molecules that increase exon 7 inclusion in SMN2 mRNA and improve SMA-associated defects, including motor function, in a mouse model (Naryshkin et al. 2014; Palacino et al. 2015). Palacino et al. (Novartis) showed that the small molecule stabilized the interaction between SMN2 pre-mRNA and U1 snRNP in a sequence-selective manner, forming a ternary complex (Palacino et al. 2015). RNA-seq analysis showed that the compounds developed collaboratively by PTC/Roche affected the alternative splicing of <0.04% of genes (Naryshkin et al. 2014). Collectively, these studies suggest that small molecules can be developed to alter splicing outcomes, which could be a therapeutic strategy for many diseases. Thus, it is of great import to understand how to drive the molecular recognition of RNA motifs by small molecules from a fundamental perspective. In concert with structural data, small molecules could be designed to disrupt or enhance RNA-protein interactions.

4 CONCLUDING REMARKS

A myriad of roles for RNA in cellular biology have been uncovered. Notably RNA's interaction with RBPs is essential for many biological processes. Aberrant interactions with proteins can lead to disease, as observed in the RNA gain-of-function mechanism that causes microsatellite disorders and the enhanced binding of U1 snRNP to a MAPT mutation that causes FTDP-17. Importantly, to exact a therapeutic effect in such RNA gain-of-function diseases, a small-molecule drug would need to bind the mutant RNA to prevent protein binding. Binding to the protein itself would inhibit its normal cellular function and likely exacerbate disease. Conversely, in the case of loss-of-function diseases that decrease protein binding, a therapeutic small molecule needs to enhance protein binding, as shown for drugs that bind to SMN2 pre-mRNA and alter its alternative splicing.

These studies exemplify that RNA is indeed a viable target for drug-like small molecules, and they thus argue for a fundamental understanding of RNA-small-molecule interactions. The impact of such small-molecule probes could be transformative, as these tools can be used to define

the pathways controlled by regulatory small RNAs (e.g., miRNAs) or to define the effects of long noncoding RNAs and toxic RAN peptides on transcription, gene silencing, cell fate, proliferation, survival, and transformation. Indeed, novel and transformative technologies have been developed that could enable the selective targeting of any RNA molecule by chemical probes. Proof-of-principle studies have established that this new generation of RNA-targeting chemical probes have tremendous potential to transform medicine for the treatment of a broad class of disorders and diseases.

ACKNOWLEDGMENTS

This work was funded by the National Institutes of Health (NIH; R01 GM097455; DP1 NS096898; P01 NS099114) and the Muscular Dystrophy Association (Grant ID 380467).

REFERENCES

*Reference is also in this collection.

- Arambula JF, Ramisetty SR, Baranger AM, Zimmerman SC. 2009. A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding. *Proc Natl Acad Sci* **106**: 16068–16073.
- Auld DS, Thorne N, Maguire WF, Inglese J. 2009. Mechanism of PTC124 activity in cell-based luciferase assays of nonsense codon suppression. *Proc Natl Acad Sci* **106**: 3585–3590.
- Bernat V, Disney MD. 2015. RNA structures as mediators of neurological diseases and as drug targets. *Neuron* **87**: 28–46.
- Bevilacqua PC, Ritchey LE, Su Z, Assmann SM. 2016. Genome-wide analysis of RNA secondary structure. *Ann Rev Genet* **50**: 235–266.
- Blanchard SC, Cooperman BS, Wilson DN. 2010. Probing translation with small-molecule inhibitors. *Chem Biol* **17**: 633–645.
- Blount KF, Breaker RR. 2006. Riboswitches as antibacterial drug targets. *Nat Biotechnol* **24**: 1558–1564.
- Blount KF, Wang JX, Lim J, Sudarsan N, Breaker RR. 2007. Antibacterial lysine analogs that target lysine riboswitches. *Nat Chem Biol* **3**: 44–49.
- Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, et al. 1992. Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**: 799–808.
- Cartegni L, Krainer AR. 2003. Correction of disease-associated exon skipping by synthetic exon-specific activators. *Nat Struct Biol* **10**: 120–125.
- Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. 2000. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**: 340–348.
- Childs-Disney JL, Disney MD. 2016. Small molecule targeting of a microRNA associated with hepatocellular carcinoma. *ACS Chem Biol* **11**: 375–380.
- Childs-Disney JL, Hoskins J, Rzuczek SG, Thornton CA, Disney MD. 2012. Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive. *ACS Chem Biol* **7**: 856–862.
- Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, Kellis M, Lindblad-Toh K, Lander ES. 2007. Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci* **104**: 19428–19433.
- Costales MG, Haga CL, Velagapudi SP, Childs-Disney JL, Phinney DG, Disney MD. 2017. Small molecule inhibition of microRNA-210 reprograms an oncogenic hypoxic circuit. *J Am Chem Soc* **139**: 3446–3455.
- Deigan KE, Li TW, Mathews DH, Weeks KM. 2009. Accurate SHAPE-directed RNA structure determination. *Proc Natl Acad Sci* **106**: 97–102.
- Disney MD, Labuda LP, Paul DJ, Poplawski SG, Pushechnikov A, Tran T, Velagapudi SP, Wu M, Childs-Disney JL. 2008. Two-dimensional combinatorial screening identifies specific aminoglycoside-RNA internal loop partners. *J Am Chem Soc* **130**: 11185–11194.
- Disney MD, Liu B, Yang WY, Sellier C, Tran T, Charlet-Berguerand N, Childs-Disney JL. 2012. A small molecule that targets r(CGG)^{exp} and improves defects in fragile X-associated tremor ataxia syndrome. *ACS Chem Biol* **7**: 1711–1718.
- Disney MD, Winkelsas AM, Velagapudi SP, Southern M, Fallahi M, Childs-Disney JL. 2016. Informa 2.0: A platform for the sequence-based design of small molecules targeting structured RNAs. *ACS Chem Biol* **11**: 1720–1728.
- Dominick PK, Keppler BR, Legassie JD, Moon IK, Jarstfer MB. 2004. Nucleic acid-binding ligands identify new mechanisms to inhibit telomerase. *Bioorg Med Chem Lett* **14**: 3467–3471.
- Donahue CP, Muratore C, Wu JY, Kosik KS, Wolfe MS. 2006. Stabilization of the exon 10 stem loop alters pre-mRNA splicing. *J Biol Chem* **281**: 23302–23306.
- Dudock BS, Katz G, Taylor EK, Holley RW. 1969. Primary structure of wheat germ phenylalanine transfer RNA. *Proc Natl Acad Sci* **62**: 941–945.
- Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshire MG, Harper PS, Brook JD. 2002. Three proteins, MBNL, MBL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum Mol Genet* **11**: 805–814.
- Feng J, Funk WD, Wang SS, Weinrich SL, Avilion AA, Chiu CP, Adams RR, Chang E, Allsopp RC, Yu J, et al. 1995. The RNA component of human telomerase. *Science* **269**: 1236–1241.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806–811.
- Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbel JO, Emanuelsson O, Zhang ZD, Weissman S, Snyder M. 2007. What is a gene, post-ENCODE? History and updated definition. *Genome Res* **17**: 669–681.
- Greider CW, Blackburn EH. 1985. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* **43**: 405–413.
- Grundy FJ, Henkin TM. 2004. Regulation of gene expression by effectors that bind to RNA. *Curr Opin Microbiol* **7**: 126–131.
- Guan L, Disney MD. 2012. Recent advances in developing small molecules targeting RNA. *ACS Chem Biol* **7**: 73–86.
- Guan L, Disney MD. 2013. Covalent small-molecule-RNA complex formation enables cellular profiling of small-molecule-RNA interactions. *Angew Chem Int Ed Engl* **52**: 10010–10013.
- Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**: 293–296.
- Hamy F, Felder ER, Heizmann G, Lazdins J, Aboul-ela F, Varani G, Karn J, Klimkait T. 1997. An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc Natl Acad Sci* **94**: 3548–3553.
- Hermann T. 2016. Small molecules targeting viral RNA. *Wiley Interdisc Rev RNA* **7**: 726–743.
- Holley RW. 1965. Structure of an alanine transfer ribonucleic acid. *JAMA* **194**: 868–871.
- Hopkins AL, Groom CR. 2002. The druggable genome. *Nat Rev Drug Discov* **1**: 727–730.
- Howe JA, Wang H, Fischmann TO, Balibar CJ, Xiao L, Galgoci AM, Malinverni JC, Mayhood T, Villafania A, Nahvi A, et al. 2015. Selective small-molecule inhibition of an RNA structural element. *Nature* **526**: 672–677.

- Inoue T, Cech TR. 1985. 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* **82**: 648–652.
- Jaeger JA, Turner DH, Zuker M. 1989. Improved predictions of secondary structures for RNA. *Proc Natl Acad Sci* **86**: 7706–7710.
- Jahromi AH, Fu Y, Miller KA, Nguyen L, Luu LM, Baranger AM, Zimmerman SC. 2013. Developing bivalent ligands to target CUG triplet repeats, the causative agent of myotonic dystrophy type 1. *J Med Chem* **56**: 9471–9481.
- Jiang Z, Cote J, Kwon JM, Goate AM, Wu JY. 2000. Aberrant splicing of τ pre-mRNA caused by intronic mutations associated with the inherited dementia frontotemporal dementia with parkinsonism linked to chromosome 17. *Mol Cell Biol* **20**: 4036–4048.
- Jones D, Metzger HJ, Schatz A, Waksman SA. 1944. Control of gram-negative bacteria in experimental animals by streptomycin. *Science* **100**: 103–105.
- Kaul M, Barbieri CM, Pilch DS. 2006. Aminoglycoside-induced reduction in nucleotide mobility at the ribosomal RNA a-site as a potentially key determinant of antibacterial activity. *J Am Chem Soc* **128**: 1261–1271.
- Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. 2001. Identification of novel genes coding for small expressed RNAs. *Science* **294**: 853–858.
- Lee RC, Ambros V. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**: 862–864.
- Lipinski CA. 2004. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov Today Technol* **1**: 337–341.
- Lu Z, Chang HY. 2016. Decoding the RNA structurome. *Curr Opin Struct Biol* **36**: 142–148.
- Luo Y, Disney MD. 2014. Bottom-up design of small molecules that stimulate exon 10 skipping in mutant MAPT pre-mRNA. *ChemBiochem* **15**: 2041–2044.
- Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci* **101**: 7287–7292.
- Moazed D, Noller HF. 1987. Interaction of antibiotics with functional sites in 16S ribosomal-RNA. *Nature* **327**: 389–394.
- Mori K, Arzberger T, Grasser FA, Gijssels I, May S, Rentzsch K, Weng SM, Schludi MH, van der Zee J, Cruts M, et al. 2013. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol* **126**: 881–893.
- Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR. 2002. Genetic control by a metabolite binding mRNA. *Chem Biol* **9**: 1043–1049.
- Naryshkin NA, Weetall M, Dakka A, Narasimhan J, Zhao X, Feng Z, Ling KK, Karp GM, Qi H, Woll MG, et al. 2014. Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* **345**: 688–693.
- Orr HT, Zoghbi HY. 2007. Trinucleotide repeat disorders. *Annu Rev Neurosci* **30**: 575–621.
- Paige JS, Wu KY, Jaffrey SR. 2011. RNA mimics of green fluorescent protein. *Science* **333**: 642–646.
- Palacino J, Swalley SE, Song C, Cheung AK, Shu L, Zhang X, Van Hoeser M, Shin Y, Chin DN, Keller CG, et al. 2015. SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat Chem Biol* **11**: 511–517.
- Parkesh R, Childs-Disney JL, Nakamori M, Kumar A, Wang E, Wang T, Hoskins J, Tran T, Housman D, Thornton CA, et al. 2012. Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching. *J Am Chem Soc* **134**: 4731–4742.
- Parsons J, Castaldi MP, Dutta S, Dibrov SM, Wyles DL, Hermann T. 2009. Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat Chem Biol* **5**: 823–825.
- Peattie DA, Gilbert W. 1980. Chemical probes for higher-order structure in RNA. *Proc Natl Acad Sci* **77**: 4679–4682.
- Pinto IG, Guilbert C, Ulyanov NB, Stearns J, James TL. 2008. Discovery of ligands for a novel target, the human telomerase RNA, based on flexible-target virtual screening and NMR. *J Med Chem* **51**: 7205–7215.
- Romero DP, Blackburn EH. 1991. A conserved secondary structure for telomerase RNA. *Cell* **67**: 343–353.
- Rzuczek SG, Gao Y, Tang ZZ, Thornton CA, Kodadek T, Disney MD. 2013. Features of modularly assembled compounds that impart bioactivity against an RNA target. *ACS Chem Biol* **8**: 2312–2321.
- Rzuczek SG, Colgan LA, Nakai Y, Cameron MD, Furling D, Yasuda R, Disney MD. 2017. Precise small molecule recognition of a toxic CUG RNA repeat expansion. *Nat Chem Biol* **13**: 188–193.
- Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, et al. 2013. Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. *Cell Rep* **3**: 869–880.
- Semenza GL. 2003. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* **3**: 721–732.
- Stark BC, Kole R, Bowman EJ, Altman S. 1978. Ribonuclease P: An enzyme with an essential RNA component. *Proc Natl Acad Sci* **75**: 3717–3721.
- Stearns T. 1995. Green fluorescent protein. The green revolution. *Curr Biol* **5**: 262–264.
- Stern S, Wilson RC, Noller HF. 1986. Localization of the binding site for protein S4 on 16 S ribosomal RNA by chemical and enzymatic probing and primer extension. *J Mol Biol* **192**: 101–110.
- Theimer CA, Finger LD, Trantirek L, Feigon J. 2003. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci* **100**: 449–454.
- Thomas JR, Hergenrother PJ. 2008. Targeting RNA with small molecules. *Chem Rev* **108**: 1171–1224.
- Tran T, Disney MD. 2012. Identifying the preferred RNA motifs and chemotypes that interact by probing millions of combinations. *Nat Commun* **3**: 1125.
- Ule J, Jensen KB, Ruggiu M, Mele A, Darnell RB. 2003. CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302**: 1212–1215.
- * Ule J, Hwang HW, Darnell RB. 2018. CLIP identifies Nova-regulated RNA networks in the brain. *Cold Spring Harb Persp Biol* **10**: a032243.
- Velagapudi SP, Seedhouse SJ, Disney MD. 2010. Structure–activity relationships through sequencing (StARTS) defines optimal and suboptimal RNA motif targets for small molecules. *Angew Chem Int Ed Engl* **49**: 3816–3818.
- Velagapudi SP, Gallo SM, Disney MD. 2014. Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nat Chem Biol* **10**: 291–297.
- Velagapudi SP, Cameron MD, Haga CL, Rosenberg LH, Lafitte M, Duckett DR, Phinney DG, Disney MD. 2016. Design of a small molecule against an oncogenic noncoding RNA. *Proc Natl Acad Sci* **113**: 5898–5903.
- Velagapudi SP, Luo Y, Tran T, Haniff HS, Nakai Y, Fallahi M, Martinez GJ, Childs-Disney JL, Disney MD. 2017. Defining RNA–small molecule affinity landscapes enables design of a small molecule inhibitor of an oncogenic noncoding RNA. *ACS Cent Sci* **3**: 205–216.
- Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW Jr, Swanson R, Burch CL, Weeks KM. 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**: 711–716.
- Welch EM, Barton ER, Zhuo J, Tomizawa Y, Friesen WJ, Trifillis P, Paushkin S, Patel M, Trotta CR, Hwang S, et al. 2007. PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **447**: 87–91.
- Wells SE, Hughes JM, Igel AH, Ares M Jr. 2000. Use of dimethyl sulfate to probe RNA structure in vivo. *Methods Enzymol* **318**: 479–493.
- West C, Francis R, Friedman SH. 2001. Small molecule/nucleic acid affinity chromatography: Application for the identification of telomerase inhibitors which target its key RNA/DNA heteroduplex. *Bioorg Med Chem Lett* **11**: 2727–2730.

- Winkler W, Nahvi A, Breaker RR. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**: 952–956.
- Woese C, Pace N. 1993. Probing RNA structure, function, and history by comparative analysis. In *The RNA world* (ed. Gesteland RF, Cech TR, Atkins JF), pp. 91–117. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Woese CR, Magrum LJ, Gupta R, Siegel RB, Stahl DA, Kop J, Crawford N, Brosius J, Gutell R, Hogan JJ, et al. 1980. Secondary structure model for bacterial 16S ribosomal RNA: Phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res* **8**: 2275–2293.
- Yang WY, Wilson HD, Velagapudi SP, Disney MD. 2015. Inhibition of non-ATG translational events in cells via covalent small molecules targeting RNA. *J Am Chem Soc* **137**: 5336–5345.
- Yonath A. 2005. Antibiotics targeting ribosomes: Resistance, selectivity, synergism and cellular regulation. *Ann Rev Biochem* **74**: 649–679.
- Yoshizawa S, Fourmy D, Puglisi JD. 1999. Recognition of the codon-anticodon helix by ribosomal RNA. *Science* **285**: 1722–1725.
- Zaug AJ, Cech TR. 1986. The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science* **231**: 470–475.
- Zhang Y-J, Jansen-West K, Xu Y, Gendron TF, Bieniek KF, Lin W-L, Sasaguri H, Caulfield T, Hubbard J, Daugherty L, et al. 2014. Aggregation-prone c9FTD/ALS poly(GA) RAN translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol* **128**: 505–524.
- Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, Daley EL, Miller SJ, Cunningham KM, Vidensky S, et al. 2015. The *C9orf72* repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**: 56–61.
- Zu T, Liu Y, Banez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, et al. 2013. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci* **110**: E4968–E4977.
- Zuker M. 1989. On finding all suboptimal foldings of an RNA molecule. *Science* **244**: 48–52.

Cold Spring Harbor Perspectives in Biology

Drugging the RNA World

Matthew D. Disney, Brendan G. Dwyer and Jessica L. Childs-Disney

Cold Spring Harb Perspect Biol 2018; doi: 10.1101/cshperspect.a034769

Subject Collection [RNA Worlds](#)

Alternate RNA Structures

Marie Teng-Pei Wu and Victoria D'Souza

Approaches for Understanding the Mechanisms of Long Noncoding RNA Regulation of Gene Expression

Patrick McDonel and Mitchell Guttman

Principles and Practices of Hybridization Capture Experiments to Study Long Noncoding RNAs That Act on Chromatin

Matthew D. Simon and Martin Machyna

Linking RNA Sequence, Structure, and Function on Massively Parallel High-Throughput Sequencers

Sarah K. Denny and William J. Greenleaf

Extensions, Extra Factors, and Extreme Complexity: Ribosomal Structures Provide Insights into Eukaryotic Translation

Melanie Weisser and Nenad Ban

Nascent RNA and the Coordination of Splicing with Transcription

Karla M. Neugebauer

Combining Mass Spectrometry (MS) and Nuclear Magnetic Resonance (NMR) Spectroscopy for Integrative Structural Biology of Protein–RNA Complexes

Alexander Leitner, Georg Dorn and Frédéric H.-T. Allain

Structural Biology of Telomerase

Yaqiang Wang, Lukas Susac and Juli Feigon

Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes

Berthold Kastner, Cindy L. Will, Holger Stark, et al.

What Are 3' UTRs Doing?

Christine Mayr

Single-Molecule Analysis of Reverse Transcriptase Enzymes

Linnea I. Jansson and Michael D. Stone

CRISPR Tools for Systematic Studies of RNA Regulation

Jesse Engreitz, Omar Abudayyeh, Jonathan Gootenberg, et al.

Relating Structure and Dynamics in RNA Biology

Kevin P. Larsen, Junhong Choi, Arjun Prabhakar, et al.

Beyond DNA and RNA: The Expanding Toolbox of Synthetic Genetics

Alexander I. Taylor, Gillian Houlihan and Philipp Holliger

For additional articles in this collection, see <http://cshperspectives.cshlp.org/cgi/collection/>

**All Modifications and
Oligo Types Synthesized**

Long Oligos • Fluorescent • Chimeric • DNA • RNA • Antisense

Oligo Modifications?

Your wish is our command.

Discovering and Mapping the Modified Nucleotides That Comprise the Epitranscriptome of mRNA

Bastian Linder and Samie R. Jaffrey

Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast

Clemens Plaschka, Andrew J. Newman and Kiyoshi Nagai

For additional articles in this collection, see <http://cshperspectives.cshlp.org/cgi/collection/>

All Modifications and Oligo Types Synthesized

Oligo Modifications?
Your wish is our command.

Long Oligos • Fluorescent • Chimeric • DNA • RNA • Antisense