Published in final edited form as:

ACS Chem Biol. 2015 December 18; 10(12): 2706–2715. doi:10.1021/acscchembio.5b00430.

Studying a Drug-like, RNA-Focused Small Molecule Library Identifies Compounds That Inhibit RNA Toxicity in Myotonic Dystrophy

Suzanne G. Rzuczek[†], Mark R. Southern[‡], and Matthew D. Disney^{*,†}

[†]Department of Chemistry, The Scripps Research Institute, Scripps Florida, 130 Scripps Way #3A1, Jupiter, Florida 33458, United States

[‡]Translational Research Institute, The Scripps Research Institute, Scripps Florida, 130 Scripps Way #3A1, Jupiter, Florida 33458, United States

Abstract

There are many RNA targets in the transcriptome to which small molecule chemical probes and lead therapeutics are desired. However, identifying compounds that bind and modulate RNA function in cellulo is difficult. Although rational design approaches have been developed, they are still in their infancies and leave many RNAs “undruggable”. In an effort to develop a small molecule library that is biased for binding RNA, we computationally identified “drug-like” compounds from screening collections that have favorable properties for binding RNA and for suitability as lead drugs. As proof-of-concept, this collection was screened for binding to and modulating the cellular dysfunction of the expanded repeating RNA (r(CUG)^{exp}) that causes myotonic dystrophy type 1. Hit compounds bind the target in cellulo, as determined by the target identification approach Competitive Chemical Cross-Linking and Isolation by Pull-down (C-ChemCLIP), and selectively improve several disease-associated defects. The best compounds identified from our 320-member library are more potent in cellulo than compounds identified by high-throughput screening (HTS) campaigns against this RNA. Furthermore, the compound collection has a higher hit rate (9% compared to 0.01–3%), and the bioactive compounds identified are not charged; thus, RNA can be “drugged” with compounds that have favorable pharmacological properties. Finally, this RNA-focused small molecule library may serve as a useful starting point to identify lead “drug-like” chemical probes that affect the biological (dys)function of other RNA targets by direct target engagement.

RNA’s diverse and essential biological functions have cemented it as an important class of targets for therapeutics and chemical probes. For example, long, noncoding RNAs, microRNAs, riboswitches, and antisense transcripts function to regulate gene expression.(1–6) Oligonucleotides, which can be designed by simple Watson–Crick base pairing rules, are

* Disney@scripps.edu.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscchembio.5b00430.

- Compound structures, fluorescence, chemical similarity, chemical property, toxicity analyses (PDF)

The authors declare no competing financial interest.

commonly employed to target RNA. Indeed, antisense and RNAi-based oligonucleotides have been used successfully to drug malfunctioning RNAs in both cells and animals, showing that RNAs are indeed viable therapeutic targets.(7–10) Unfortunately, oligonucleotides are generally not cell permeable or drug-like, and efficient, general delivery systems have been elusive.

As an alternative to oligonucleotide-based probes, various approaches have been developed to identify small molecules that target RNA. We have designed small molecules to bind an RNA of interest by examining its secondary structural motifs and comparing them to a database of known and annotated RNA motif-small molecule interactions.(11) Structure-based approaches include mimicking natural substrates for riboswitches,(12–14) designing small molecules to interact with hydrogen bond donors and acceptors in RNA grooves,(15–17) and mimicking interactions between RNAs and proteins.(18–20) Lastly, computational experiments have been used to explore RNA–small molecule interactions by docking validated binders into RNA landscapes.(21) Despite these advances, targeting many RNA targets is still intractable due to limited available data such as 3-dimensional structures, chemotypes that engender RNA binding affinity or selectivity, and RNA structural elements that form small molecule binding sites. Screening of small molecule libraries for binding RNA targets could generate data about the latter two points. However, few small molecule-screening collections are enriched in RNA binders; in fact, most are biased with compounds that bind to proteins.

To develop an RNA-focused small molecule screening collection, we computationally analyzed small molecules within The Scripps Research Institute's compound collection to identify a subset of compounds that have favorable features for binding RNA ($n = 320$). The compounds were then screened for binding to the r(CUG) repeating RNA (r(CUG)^{exp}) that causes myotonic dystrophy type 1 (DM1), although this library could be used to screen other biologically relevant RNA targets. r(CUG)^{exp} is located in the 3' untranslated region (UTR) of the dystrophin myotonia protein kinase (*DMPK*) mRNA.(22) Healthy individuals have 5–37 repeats in the *DMPK* RNA, while DM1-affected individuals have 50–2000 repeats,(23) with disease severity increasing as a function of repeat length.(24) r(CUG)^{exp} causes disease via a gain-of-function mechanism in which the RNA binds to and sequesters proteins that are involved in RNA biogenesis.(25–29) Protein sequestration results in the deregulation of alternative pre-mRNA splicing(26–28) and decreased nucleocytoplasmic transport and translation of the *DMPK* mRNA.(30) By screening this focused collection of small molecules, we identified a variety of small molecules that bind this RNA and improve disease-associated pre-mRNA splicing defects in cell culture models. Competitive Chemical Cross-Linking and Isolation by Pull Down (C-ChemCLIP) experiments show that these designer compounds bind to the intended target in cells. Analysis of the hit compounds revealed favorable and unfavorable chemotypes that affect bioactivity and selectivity. Our RNA-focused library may have broader use in the identification of small molecules that bind to and modulate the activity of other cellular RNAs.

Results and Discussion

Collection of the RNA-Focused Small Molecule Library

In an effort to understand chemotypes that confer affinity for RNA, we have investigated the binding of small RNA motifs to various classes of small molecules.(31–35) Among others, these studies suggested that a benzimidazole scaffold is privileged for binding RNA.(32) Furthermore, several studies have shown that these compounds bind selectively to RNA motifs present in the transcriptome, particularly 1×1 nucleotide internal loops.(11, 36, 37) Thus, we utilized chemical similarity searching to identify small molecules that have bis-benzimidazole or similar core structures based on the position of structural features such as rings, positive charge, polar, and hydrophobic groups, among others.(38, 39)

The compound collection was refined based on “drug-likeness” and potential to interact with RNA, which is of considerable interest as many small molecules that are known to bind RNA do not follow traditional drug-likeness guidelines. For example, many RNA binders have >5 hydrogen bond donors or acceptors and have molecular weights >500 Da. They likely accommodate many favorable contacts with RNA and have increased surface areas to interact with the RNA’s secondary structure. Our resulting collection of 320 compounds is both structurally diverse and chemically similar (Figure S-1). Moreover, they are drug-like, as assessed by Lipinski’s Rule of 5 (RO5), Lilly MedChem Rules, and Pan Assay Interference Compounds (PAINS) analyses.(40–42) RO5 is a subset of predictors for solubility and permeability of small molecules based on analysis of the physicochemical properties of over 2000 drugs and drug candidates.(40) The Lilly MedChem Rules were designed to remove compounds that are likely to give invalid or misleading results from HTS screens.(42) These include compounds that are unstable, reactive, promiscuously active for multiple target families, toxic, or metabolically unstable, or that interfere with activity measurements. PAINS analysis consists of a group of substructure filters designed to remove compounds that are promiscuous or interfere with many biochemical high-throughput screening (HTS) assays.(41) For this RNA-focused library, 83% of the compounds are in compliance with RO5 guidelines; the remainders generally fail only one rule and fail by only a small amount. Additionally, the compounds hold up well when analyzed by the Lilly MedChem Rules and PAINS analysis.(41, 42) Thus, our studies investigated if traditional drug-like guidelines can also be used to drug RNA, as applied to the RNA that causes DM1.

In Vitro Screening of Compounds for Disruption of an RNA–Protein Complex That Causes DM1

As mentioned above, DM1 is caused by an RNA gain-of-function mechanism in which $r(\text{CUG})^{\text{exp}}$ binds and inactivates proteins involved in RNA biogenesis.(25–29) One protein, muscleblind-like 1 (MBNL1), regulates the alternative splicing of a subset of RNAs that are deregulated in DM1.(27, 43–45) We therefore screened our library of small molecules (100 μM concentration) for disrupting the $r(\text{CUG})_{12}$ –MBNL1 complex in vitro using a previously described time-resolved fluorescence resonance energy transfer (TR-FRET) assay.(46) Hit compounds were defined as disrupting the $r(\text{CUG})_{12}$ –MBNL1 complex with a potency that is ≥ 3 times the standard deviation observed for all compounds. This analysis afforded 28 hit compounds, or a hit rate of 9% (Figure 1A). The most potent inhibitors disrupted >70% of

the complex at 100 μ M concentration and included compounds **1**, **2**, and **3** (Figure 1B). The majority of hits inhibited 50% of the complex at this concentration (structures shown in Figure S-2). The IC_{50} of each hit compounds was measured, and the results are shown in Figure 1B. The IC_{50} values for these compounds range from $\sim$ 30 to $\sim$ 150 μ M. Typically, lead compounds contained multiple substructures that can interact with RNA.(35)

Hit compounds were then subjected to substructure analysis as described previously.(32, 35) The most common substructures contained substituted pyridyl, benzimidazole, or imidazole ring systems. Specific functionalities derived from these substructures include benzyl imidazoles, benzyl benzimidazoles, and pyridyl benzimidazoles; although they occur $\sim$ 2 times more frequently in hit compounds than in the 320-member library, they are not statistically significant (Figure 1C). These chemotypes have been found in other RNA binders and have comprised structures in RNA-focused compound collections.(35) Similarity analysis of the hit compounds compared to the Scripps Drug Discovery Library revealed that 98% of library compounds have a similarity to the actives of 0.45 or less, and 99.9% of the Molecular Libraries Small Molecule Repository (MLSMR) have a similarity value of 0.75 or less (Figure S-5).

Bioactivity: Improving DM1-Associated Alternative Pre-mRNA Splicing Defects

The biological activity of hit compounds was studied in a DM1 cellular model. Of the 28 leads identified from our in vitro screen, 22 were commercially available and further evaluated. Small molecules were tested for their ability to improve DM1-associated pre-mRNA splicing defects caused by sequestration of MBNL1 (Figure 2).(47) In particular, we co-transfected a DM1 mini-gene that contains 960 r(CUG) repeats and a mini-gene that reports on alternative splicing of cardiac troponin T (cTNT) exon 5, which is deregulated in DM1.(25) In the presence of r(CUG)₉₆₀, $\sim$ 90% of exon 5 is included in the mature mRNA, while the amount of mature mRNA containing exon 5 in healthy cells is $\sim$ 55% (absence of r(CUG)₉₆₀) (Figure 2A, B). Thus, if the compound targets r(CUG)₉₆₀ and frees MBNL1 *in cellulo*, then a decrease in exon 5 inclusion should be observed. In these experiments, 12 of the tested compounds improved cTNT alternative pre-mRNA splicing patterns toward wild type (Figure 2B).

The specificity of the compounds was then assessed by monitoring changes in cTNT splicing patterns in cells that do not express r(CUG)₉₆₀. Compounds that are selective for the DM1 RNA should have no effect on cTNT pre-mRNA splicing in healthy cells that do not express r(CUG)₉₆₀. Nine compounds significantly exacerbated cTNT splicing in healthy cells, increasing exon 5 inclusion, while three compounds significantly decreased the amount of exon 5 included. Therefore, these compounds were deemed nonselective and were not further pursued. On the basis of potency and selectivity, the most promising lead compounds were **1**, **16**, and **17** (Figures 1B and 2). Interestingly, these three compounds do not have a positive charge at physiological pH. One goal of drug discovery efforts for RNA targets is to find uncharged, bioactive compounds. These molecules appear to meet this goal, *vide infra*.

A dose response was completed for compounds **16** and **17**, which retained modest activity at 10 μ M for improving the cTNT pre-mRNA splicing defect (Figure 2C). The toxicity of

compounds **1**, **16**, and **17** was assessed in this cellular system by using a WST-1 assay and the compounds do not affect cell viability (Figure S-6). The selectivities of these molecules were further studied *in vitro* by competition dialysis and *in cellulo* for target engagement using a method developed in our laboratory called Chemical Cross-Linking and Isolation by Pull-down (ChemCLIP).(48, 49)

In Vitro Selectivity of Bioactive Small Molecules **1**, **16**, and **17**

The selectivities of compounds **1**, **16**, and **17** were measured by competition dialysis using various RNA and protein targets. RNA targets included a mimic of the human rRNA A-site, the MBNL1-binding site in the cTNT pre-mRNA; an RNA with GC base pairs, and RNAs with different numbers of r(CUG) repeats (r(CUG)_{1×2}, r(CUG)_{6×2}, and r(CUG)₁₀₉, an expanded repeat of a length causative of disease(23)) (Figure 3A). By studying these RNAs, we can gain insight into compound selectivity and if a compound binds cooperatively to RNAs with multiple 5'CUG/3'GUC motifs (secondary structure displayed by r(CUG)^{exp}; Figure 3A). Binding to MBNL1 was also evaluated to assess the propensity of the compounds to bind MBNL1 over r(CUG)^{exp}.

All three bioactive small molecules bind the desired target, r(CUG)₁₀₉, to the greatest extent as compared to the other RNAs and MBNL1 (Figure 3B). The small molecule with the highest loading onto r(CUG)₁₀₉ was also the most bioactive, compound **17**. Interestingly, although r(CUG)₁₀₉ has ~9 times more 5'CUG/3'GUC binding sites than r(CUG)_{6×2}, the loading of **17** is ~14-fold higher on r(CUG)₁₀₉ than on r(CUG)_{6×2}. This suggests that **17** binds cooperatively to RNAs with longer repeats, which may be an important factor for target recognition and bioactivity. It has been previously observed that compounds can bind to r(CUG)^{exp} with positive cooperativity.(50) Compound **16** may not benefit from cooperative binding as it has the lowest loading onto r(CUG)₁₀₉ yet still selectively improves DM1 pre-mRNA splicing defects. This effect likely becomes significant as the repeat length increases. In our *in vitro* TR-FRET screen, the repeat length is 12, while the pathogenic repeat length used in cellular studies is 960.

In Cellulo Selectivity of Bioactive Small Molecules **1**, **16**, and **17**

In some previously reported studies, small molecules improve DM1-associated defects not by direct target engagement but rather by decreasing levels of r(CUG)^{exp} (*DMPK*) at the transcriptional level.(51) Therefore, to validate the mechanism of action for the compounds described herein, we studied if **1**, **16**, and **17** bind r(CUG)^{exp} *in cellulo* as well as their effects on the abundance of *DMPK* mRNA (Figure 4). At bioactive concentrations, neither **16** nor **17** had any effect on the abundance of *DMPK* mRNA, while compound **1** diminished *DMPK* abundance by ~40% (Figure 4A).

The biological targets of the small molecules were further investigated using Competitive ChemCLIP (C-ChemCLIP) (Figure 4B).(48, 49) In ChemCLIP, an RNA-binding compound is appended with (i) a reactive module (chlorambucil; **CA**) that cross-links to the RNAs that the small molecule binds *in cellulo*; and (ii) a biotin moiety for facile pull-down and isolation of those targets.(49) The pulled-down fraction is then analyzed by qRT-PCR. In this particular application, we used C-ChemCLIP using a previously studied compound,

2H-4-CA-Biotin, that binds and cross-links to r(CUG)^{exp} *in cellulo*.⁽⁴⁹⁾ If our lead compounds also engage r(CUG)^{exp} *in cellulo*, then the amount of the RNA pulled down by **2H-4-CA-Biotin** should decrease as a function of **1**, **16**, or **17** concentration. In the presence of **1**, **16**, and **17**, the amount of r(CUG)^{exp} pulled down by **2H-4-CA-Biotin** is significantly reduced and indicates that the small molecules bind r(CUG)^{exp} (Figure 4C). Taken together, the bioactivity of **16** and **17** can be directly attributed to binding r(CUG)^{exp}; the bioactivity of compound **1** may have a mixed mode of inhibition in which the compound inhibits transcription of the r(CUG)^{exp}-containing mRNA or causes its degradation and binds the repeats, all of which could lead to improvement of r(CUG)^{exp}-associated dysfunction.

Chemotype Analysis of Bioactive Compounds

Analysis of the physicochemical properties of selective bioactive compounds and bioactive compounds with nonspecific effects was completed to determine if certain chemotypes positively or negatively contribute to cellular selectivity. The results are summarized in Table 1. Collectively, physicochemical properties of the nonspecific compounds are not statistically different from the starting library. The majority of these compounds are benzimidazoles with either a hydrogen or nitrogen at the 5 position and an aromatic or heteroaromatic ring at the 2 position. Conversely, compounds that are selectively bioactive have several physicochemical properties that are statistically different from the starting library. On average, selectively bioactive compounds have a larger topological polar surface area (TPSA) ($101 \pm 33 \text{ \AA}^2$ vs 75 ± 25 , $p = 0.002$) compared to the starting library. Also, this subset of compounds has more hydrogen bond donors (3 ± 1 vs 2 ± 1 , $p = 0.02$) and acceptors (4 ± 1 vs 3 ± 1 , $p = 0.05$) than the library. The majority of these compounds are benzimidazoles with a para-substituted phenyl ring at the 2 position.

The nonselective bioactive compounds can be divided into two groups: those that exacerbate splicing toward a DM1 phenotype (increase exon 5 inclusion; $n = 9$) and those that exacerbate wild type splicing (decrease exon 5 inclusion; $n = 3$) in the absence of r(CUG)^{exp}. Compounds that exacerbate DM1 splicing defects have a lower logP (2.4 ± 1.2 vs 3.4 ± 1.5 , $p = 0.04$) and higher TPSA ($92 \pm 22 \text{ \AA}^2$ vs $75 \pm 25 \text{ \AA}^2$, $p = 0.04$) compared to the entire screening library. A common structural feature found in over half of this subset is a 5-aminobenzimidazole with a nitrogen-containing heterocycle at the 2 position. Compounds that exacerbate wild type splicing are not statistically different from the starting library of compounds, although two of the three compounds are 6-aminobenzimidazoles with a halogenated N-phenyl benzamidine.

Lead compounds **1**, **16**, and **17** also have statistically significant properties different from the initial 320 compounds used in the HTS screen. These compounds have a greater molecular weight (398 ± 27 vs 317 ± 75 amu, $p = 3 \times 10^{-10}$) and a larger TPSA (111 ± 39 vs $75 \pm 25 \text{ \AA}^2$, $p = 0.01$) than the screening library. Analysis of the structures revealed that **16** and **17** both contain 2-phenyl substituted benzimidazoles with a furan carboxamide attached at the 4 position of the phenyl ring. Interestingly, these were the only two library members that contained this scaffold, which may be important to include in future RNA-focused library screens.

Summary and Outlook

RNA has become an important drug target as its dysfunction and aberrant expression have been associated with a wide variety of diseases. Finding selective small molecules that target the structures of disease-associated RNAs has been challenging, as many screening collections are biased toward protein targets. The identification of features that favor RNA binding could enable the compilation of RNA-focused libraries, although such data thus far are limited.

Herein, we have compiled an RNA-focused library for HTS screening using a selective RNA-binding small molecule that targets the toxic RNA found in DM1 as a lead compound for a chemical similarity search. Overall, our library obeyed most of the traditional medicinal chemistry rules designed for HTS screening collections. Screening this RNA-focused library against the intended target resulted in hit rate of 9%. Previously reported HTS screens using traditional libraries against RNA targets had hit rates ranging from 0.01% to 3%.^(46, 52–57) Furthermore, the hits from our RNA-focused library were more potent *in cellulo* than previous hit compounds identified by our group using similar screening methods.⁽⁴⁹⁾ Interestingly, a screen of the Molecular Libraries Probe Production Centers Network (MLPCN) library against the 5' UTR of α -synuclein resulted in the identification of a probe with similar chemical features, such as the number of rotatable bonds, number of hydrogen bond acceptors and TPSA as our most bioactive compounds.⁽⁵⁶⁾ Substructure analysis of the hits showed that r(CUG)^{exp} favors compounds containing substituted pyridyl, benzimidazole, or imidazole ring systems. Analysis of hit compounds in a DM1 cellular model revealed features that favor selective bioactivity and identified particular scaffolds that should either be included or avoided in future library screens. Our RNA-focused small molecule library can serve as a starting point to identify lead compounds for other disease-associated RNA targets.

Methods

RNA-Focused Small Molecule Library Selection

The TSRI small molecule library has approximately 1 million compounds available for screening efforts, comprising both diverse and focused sublibraries. The drug discovery collection forms the largest part and consists of over 600000 unique and drug-like compounds. It is constructed from commercial sources as well as past and current internal medicinal chemistry/drug discovery efforts. Compound management is via a customized chemical registration/search system that is built on ISIS/Host technology (Biovia, formerly Accelrys). A molecular similarity search, using the ISIS:Base client was performed to find compounds that were structurally similar to the reference compound. The degree of similarity is dependent on the percentage of searchable keys that the query compound has in common to a compound stored in the database. (ISIS sets approximately 960 searchable keys, defining structural features such as ring structure or heteroatom arrangement.) No compounds were found to be greater than 75% similar, although a workable number (815) were found to be greater than 65%. Of these, 621 were available from the screening center from which the 320 compounds that are most chemically similar to the *bis*-benzimidazole lead (H1)(32) were selected. This enabled the assay to be performed in a 384 well plate with

space for controls. Chemical descriptors were then calculated in Pipeline Pilot (Biovia), including Lipinski's Rule of Five,(40) PAINS,(41) and Lilly MedChem Rules(42) via customized components that were created by the Scripps Molecular Screening Center. The final compounds were selected based on similarity score, availability from the Scripps Library (replenishment rules), and for purity and drug-likeness (see Results and Discussion).

Chemo-Informatic Analysis of the Small Molecule Library

Chemical substructures of hit compounds were determined using NGCG Scaffold Hopper software R group analysis (<https://tripod.nih.gov/chembl/v17>).(58) LogP, quantification of hydrogen bond donors and acceptors, TPSA, and rotatable bonds were calculated using Instant JChem software (JChem version 5.11.4, ChemAxon, www.chemaxon.com). The Scripps Drug Discovery library was compared to the active compounds using a fingerprint-based approach. Extended connectivity fingerprints with a path length of 6 (ECFP-6) and Tanimoto similarities were calculated with Biovia Pipeline Pilot. The closest pairwise similarities between each of the library compounds and any of the actives were then plotted with the histogram tool in Microsoft Excel's Data Analysis add-in.

High-Throughput Time-Resolved Fluorescence Resonance Energy Transfer (TR-FRET) Screening

The in vitro activity of library compounds was assessed by disruption of the r(CUG)₁₂-MBNL1 complex using a previously reported TR-FRET assay.(32, 46) Biotinylated RNA was folded by incubating at 60 °C in 1× Folding Buffer (20 mM HEPES, pH 7.5, 100 M KCl, and 10 mM NaCl) followed by slowly cooling to RT. The buffer was then adjusted to 1× Assay Buffer (1× Folding Buffer supplemented with 2 mM MgCl₂, 2 mM CaCl₂, 5 mM DTT, 0.1% BSA, 0.05% Tween-20) and MBNL1-His6 was added. The final concentrations of RNA and MBNL1-His6 were 80 and 62 nM, respectively. After incubation at RT for 15 min, 8 µL aliquots were then dispensed into each well of low volume, 384-well plates using a flying reagent dispenser. A 400 nL aliquot of small molecule stock solution (2.5 mM) was added using a Beckman Coulter Biomek NXp Laboratory Automation Workstation, affording a final small molecule concentration of ~100 µM. The samples were allowed to equilibrate at RT for 15 min, after which 1 µL of antibody solution (1:1 mixture of 8.8 ng/µL Anti-His6-Tb and 800 nM streptavidin XL-665) was dispensed into the wells. Control wells for maximum TR-FRET (100% complex formation) contained 9 µL of 1× Assay Buffer and 1 µL water. Controls for minimum TR-FRET (no complex formation) contained 9 µL of 1× Assay Buffer, 1 µL water, and no RNA or protein. Plates were incubated at RT for 1 h, and then TR-FRET was measured using a Molecular Devices SpectraMax M5 plate reader using an excitation wavelength of 345 nm and a 420 nm cutoff. Fluorescence was measured at 545 and 665 nm. Library compounds were screened for potential intrinsic fluorescence under screening conditions. No library compounds were fluorescent at 665 nm and two nonhit compounds have low fluorescence at 545 nm (Figures S-3 and S-4).

To calculate the percent inhibition of complex formation, the ratio of fluorescence intensity at 545 and 665 nm in the presence of compound was compared to the ratio in the absence of small molecule (100% r(CUG)₁₂-MBNL1 complex formation) and in the absence of RNA

and protein (no complex formation). Hit compounds were defined as those that had a percent inhibition 3 times the standard deviation of the percent inhibition of the entire plate.

Improvement of DM1-Associated Splicing Defects in a DM1 Cell Culture Model Using RT-PCR

To determine bioactivity of hit compounds as assessed by improvement of DM1-associated alternative pre-mRNA splicing defects, a previously reported method was utilized.(25, 59) Briefly, HeLa cells were grown as monolayers in 96-well plates in growth medium (1× DMEM, 10% FBS, and 1× Glutamax (Invitrogen)). Once the cells reached ~80% confluence, they were transfected with 200 ng of total plasmid using Lipofectamine 2000 (Invitrogen) according to the manufacturer's standard protocol. Equal amounts of plasmid expressing a DM1 mini-gene with 960 CTG repeats(25) and mini-gene that reports on cTNT alternative splicing (regulated by MBNL1)(25) were used. After 5 h, the transfection cocktail was removed and replaced with growth medium containing the compound of interest. After 20–24 h, the cells were lysed, and total RNA was harvested using a Zymo Quick RNA miniprep kit. An on-column DNA digestion was completed per the manufacturer's recommended protocol. A sample of RNA was subjected to RT-PCR as previously described(59) using ~150 ng of total RNA.

The RT-PCR primers for the cTNT mini-gene were 5'-GTTTACAACCATCTAAAGCAAGATG (forward) and 5'-GTTGCATGGCTGGTGCAGG (reverse). The forward primer was radiolabeled with γ -32P ATP using T4 polynucleotide kinase. RT-PCR products were separated using a denaturing 5% polyacrylamide gel run at 50 W for 2 h in 1× TBE buffer. Gels were imaged using a Molecular Dynamics Typhoon 9410 variable mode imager.

Competition Dialysis

Competition dialysis was performed using Pierce Slide-A-Lyzer MINI dialyzer units with a 3500 Da molecular weight cutoff. The RNAs evaluated were folded by heating an 8 μ M solution of RNA in 199 μ L of 1× Dialysis Buffer (8 mM NaH₂PO₄, pH 7.2, 185 mM NaCl, and 1 mM EDTA) at 65 °C for 5 min followed by slowly cooling to RT on the benchtop. Then, 1 μ L of 1 mM small molecule was added (5 μ M final concentration), and the solution was transferred to the dialyzer units (100 μ L samples in duplicate). The dialyzer units were then placed in 150 mL of 5 μ M small molecule in 1× Dialysis Buffer, and the mixture was stirred at 200 rpm at 4 °C for 48 h. Following equilibration, 90 μ L of each sample was transferred to a microcentrifuge tube and treated with 10 μ L of 10% SDS to dissociate the ligand. The total ligand concentration (C_t) within each dialysis unit was quantified spectrophotometrically using the appropriate absorbance wavelength and extinction coefficient for each compound. The free ligand concentration (C_f) was determined from a sample of the dialysate solution, which did not vary significantly from the initial concentration. The bound ligand concentration (C_b) was then determined using eq 1:

$$C_b = C_t - C_f \quad (1)$$

where C_b , C_t , and C_f are concentrations of bound, total, and free ligand, respectively.

Target Pull-Down by C-ChemCLIP and Analysis by qRT-PCR

HeLa cells were grown as monolayers to ~80% confluence in 100 mm³ plates in growth medium. Cells were transfected with 10 μ g of plasmid expressing the DM1 mini-gene using Lipofectamine 2000 (Invitrogen) according to the manufacturer's standard protocol. After 5 h, the transfection cocktail was removed and replaced with growth medium containing 100 nM **2H-4-CA-Biotin** and 10 μ M bioactive small molecule. After 20–24 h, the cells were lysed and total RNA was harvested using Trizol reagent (Ambion) according to the manufacturer's protocol. Approximately 30 μ g of total RNA was used for pull-down using 15 nmol of streptavidin–agarose beads (Sigma-Aldrich; 15 μ g/mL biotin loading) as previously described.⁽⁴⁹⁾ The lysate was incubated with the beads for 1 h at RT with shaking at 500 rpm. The solution was removed, and the beads were washed with 500 μ L aliquots of 1 $\times$ PBS. Bound material was released from the beads by heating at 95 °C for 5 min in 30 μ L of 95% formamide containing 10 mM EDTA, pH 8.2.

Reverse transcription reactions were carried out using qScript cDNA synthesis kit by adding ~60 ng of either total RNA or captured RNA according to the manufacturer's protocol. Then, 30% of each cDNA sample was used for quantitative real time PCR (qPCR) analysis for each primer set. qPCR was performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems) using Power SYBR Master Mix (Life Technologies). The PCR primers for the DM1/*DMPK* mini-gene mRNA were 5'-CGTGCAAGCGCCAG (forward) and 5'-CTCCACCAACTTACTGTTTCATTCT (reverse). The PCR primers for 18S rRNA were 5'-GTAACCCGTTGAACCCATT (forward) and 5'-CCATCCAATCGGTAGTAGCG (reverse).

Analysis of the Effects of Small Molecules on *DMPK* mRNA Abundance by qRT-PCR

HeLa cells were grown as monolayers to ~80% confluence in 6-well plates in growth medium. Cells were transfected with 2.5 μ g/well of total plasmid using Lipofectamine 2000 (Invitrogen) according to the manufacturer's standard protocol. Equal amounts of plasmids expressing the DM1/*DMPK* and cTNT mini-genes were used. After 5 h, the transfection cocktail was removed and replaced with growth medium containing 100 μ M small molecule. After 20–24 h, total RNA was isolated as described above. Reverse transcription reactions were carried out using a qScript cDNA synthesis kit by adding ~200 ng of total RNA according to the manufacturer's protocol. The qRT-PCR analyses were carried out as described above.

Toxicity Analysis of Bioactive Compounds

The toxicities of **1**, **16**, and **17** were assessed using our model system for DM1 pre-mRNA splicing defects as described above. After treatment with compound for 24 h, the cells were washed with 1 $\times$ DPBS and treated with 100 μ L of a 10% solution of WST-1 reagent (Roche) in growth medium. Cells were incubated with this solution for 30 min at 37 °C and then 60 μ L aliquots were transferred to a 96-well plate. The absorbance at 450 and 690 nm was measured using a Molecular Devices SpectraMax M5 plate reader.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. St Laurent G, Faghihi MA, Wahlestedt C. Non-coding RNA transcripts: Sensors of neuronal stress, modulators of synaptic plasticity, and agents of change in the onset of Alzheimer's disease. *Neurosci Lett*. 2009; 466:81–88. DOI: 10.1016/j.neulet.2009.08.032 [PubMed: 19699259]
2. Qureshi IA, Mehler MF. Long non-coding RNAs: novel targets for nervous system disease diagnosis and therapy. *Neurotherapeutics*. 2013; 10:632–646. DOI: 10.1007/s13311-013-0199-0 [PubMed: 23817781]
3. Calin GA, Croce CM. MicroRNAs and chromosomal abnormalities in cancer cells. *Oncogene*. 2006; 25:6202–6210. DOI: 10.1038/sj.onc.1209910 [PubMed: 17028600]
4. Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. *Nat Biotechnol*. 2006; 24:1558–1564. DOI: 10.1038/nbt1268 [PubMed: 17160062]
5. Blount KF, Wang JX, Lim J, Sudarsan N, Breaker RR. Antibacterial lysine analogs that target lysine riboswitches. *Nat Chem Biol*. 2007; 3:44–49. DOI: 10.1038/nchembio842 [PubMed: 17143270]
6. Alvarez-Salas LM. Nucleic acids as therapeutic agents. *Curr Top Med Chem*. 2008; 8:1379–1404. DOI: 10.2174/156802608786141133 [PubMed: 18991725]
7. Wheeler TM, Sobczak K, Lueck JD, Osborne RJ, Lin X, Dirksen RT, Thornton CA. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science*. 2009; 325:336–339. DOI: 10.1126/science.1173110 [PubMed: 19608921]
8. Wheeler TM, Lueck JD, Swanson MS, Dirksen RT, Thornton CA. Correction of *CIC-1* splicing eliminates chloride channelopathy and myotonia in mouse models of myotonic dystrophy. *J Clin Invest*. 2007; 117:3952–3957. DOI: 10.1172/JCI33355 [PubMed: 18008009]
9. Mulders SAM, van den Broek WJAA, Wheeler TM, Croes HJE, van Kuik-Romeijn P, de Kimpe SJ, Furling D, Platenburg GJ, Gourdon G, Thornton CA, Wieringa B, Wansink DG. Triplet-repeat oligonucleotide-mediated reversal of RNA toxicity in myotonic dystrophy. *Proc Natl Acad Sci U S A*. 2009; 106:13915–13920. DOI: 10.1073/pnas.0905780106 [PubMed: 19667189]
10. Sobczak K, Wheeler TM, Wang W, Thornton CA. RNA interference targeting CUG repeats in a mouse model of myotonic dystrophy. *Mol Ther*. 2013; 21:380–387. DOI: 10.1038/mt.2012.222 [PubMed: 23183533]
11. Velagapudi SP, Gallo SM, Disney MD. Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nat Chem Biol*. 2014; 10:291–297. DOI: 10.1038/nchembio.1452 [PubMed: 24509821]
12. Sudarsan N, Wickiser JK, Nakamura S, Ebert MS, Breaker RR. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev*. 2003; 17:2688–2697. DOI: 10.1101/gad.1140003 [PubMed: 14597663]
13. Kubodera T, Watanabe M, Yoshiuchi K, Yamashita N, Nishimura A, Nakai S, Gomi K, Hanamoto H. Thiamine-regulated gene expression of *Aspergillus oryzae* thiA requires splicing of the intron containing a riboswitch-like domain in the 5'-UTR. *FEBS Lett*. 2003; 555:516–520. DOI: 10.1016/S0014-5793(03)01335-8 [PubMed: 14675766]
14. Sudarsan N, Cohen-Chalamish S, Nakamura S, Emilsson GM, Breaker RR. Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. *Chem Biol*. 2005; 12:1325–1335. DOI: 10.1016/j.chembiol.2005.10.007 [PubMed: 16356850]
15. Fourmy D, Recht MI, Blanchard SC, Puglisi JD. Structure of the A site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science*. 1996; 274:1367. doi: 10.1126/science.274.5291.1367 [PubMed: 8910275]
16. Hermann T, Westhof E. Aminoglycoside binding to the hammerhead ribozyme: a general model for the interaction of cationic antibiotics with RNA. *J Mol Biol*. 1998; 276:903–912. DOI: 10.1006/jmbi.1997.1590 [PubMed: 9566195]
17. Yoshizawa S, Fourmy D, Puglisi JD. Structural origins of gentamicin antibiotic action. *EMBO J*. 1998; 17:6437–6448. DOI: 10.1093/emboj/17.22.6437 [PubMed: 9822590]

18. Niu Y, Jones AJ, Wu H, Varani G, Cai J. γ -AApeptides bind to RNA by mimicking RNA-binding proteins. *Org Biomol Chem*. 2011; 9:6604–6609. DOI: 10.1039/c1ob05738c [PubMed: 21826330]
19. Davidson A, Leeper TC, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson JA, Varani G. Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc Natl Acad Sci U S A*. 2009; 106:11931–11936. DOI: 10.1073/pnas.0900629106 [PubMed: 19584251]
20. Kwon Y, Arndt HD, Mao Q, Choi Y, Kawazoe Y, Dervan PB, Uesugi M. Small molecule transcription factor mimic. *J Am Chem Soc*. 2004; 126:15940–15941. DOI: 10.1021/ja0445140 [PubMed: 15584709]
21. Lind KE, Du Z, Fujinaga K, Peterlin BM, James TL. Structure-based computational database screening, in vitro assay, and NMR assessment of compounds that target TAR RNA *Chem Biol*. 2002; 9:185–193. DOI: 10.1016/S1074-5521(02)00106-0 [PubMed: 11880033]
22. Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, Sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Jones C, Juvonen V, Johnson K, Harper PS, Shaw DJ, Housman DE. 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*. 1992; 68:799–808. DOI: 10.1016/0092-8674(92)90154-5 [PubMed: 1310900]
23. Tsiflidis C, MacKenzie AE, Mettler G, Barcelo J, Korneluk RG. Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nat Genet*. 1992; 1:192–195. DOI: 10.1038/ng0692-192 [PubMed: 1303233]
24. Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, Cannon SC, Thornton CA. Expanded CUG repeats trigger aberrant splicing of *CIC-1* chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell*. 2002; 10:35–44. DOI: 10.1016/S1097-2765(02)00563-4 [PubMed: 12150905]
25. Philips AV, Timchenko LT, Cooper TA. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science*. 1998; 280:737–741. DOI: 10.1126/science.280.5364.737 [PubMed: 9563950]
26. Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol*. 1995; 128:995–1002. DOI: 10.1083/jcb.128.6.995 [PubMed: 7896884]
27. Miller JW, Urbinati CR, Teng-Ummuay P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS. Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J*. 2000; 19:4439–4448. DOI: 10.1093/emboj/19.17.4439 [PubMed: 10970838]
28. Charlet BN, Savkur RS, Singh G, Philips AV, Grice EA, Cooper TA. Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol Cell*. 2002; 10:45–53. DOI: 10.1016/S1097-2765(02)00572-5 [PubMed: 12150906]
29. Fardaei M, Larkin K, Brook JD, Hamshire MG. In vivo co-localisation of MBNL protein with *DMPK* expanded-repeat transcripts. *Nucleic Acids Res*. 2001; 29:2766–2771. DOI: 10.1093/nar/29.13.2766 [PubMed: 11433021]
30. Amack JD, Paguio AP, Mahadevan MS. Cis and trans effects of the myotonic dystrophy (DM) mutation in a cell culture model. *Hum Mol Genet*. 1999; 8:1975–1984. DOI: 10.1093/hmg/8.11.1975 [PubMed: 10484765]
31. Disney MD, Labuda LP, Paul DJ, Poplawski SG, Pushechnikov A, Tran T, Velagapudi SP, Wu M, Childs-Disney JL. Two-dimensional combinatorial screening identifies specific aminoglycoside-RNA internal loop partners. *J Am Chem Soc*. 2008; 130:11185–11194. DOI: 10.1021/ja803234t [PubMed: 18652457]
32. Parkesh R, Childs-Disney JL, Nakamori M, Kumar A, Wang E, Wang T, Hoskins J, Tran T, Housman DE, Thornton CA, Disney MD. Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database & chemical similarity searching. *J Am Chem Soc*. 2012; 134:4731–4742. DOI: 10.1021/ja210088v [PubMed: 22300544]
33. Velagapudi SP, Seedhouse SJ, French J, Disney MD. Defining the RNA internal loops preferred by benzimidazole derivatives via two-dimensional combinatorial screening and computational analysis. *J Am Chem Soc*. 2011; 133:10111–10118. DOI: 10.1021/ja200212b [PubMed: 21604752]

34. Velagapudi SP, Pushechnikov A, Labuda LP, French JM, Disney MD. Probing a 2-aminobenzimidazole library for binding to RNA internal loops via two-dimensional combinatorial screening. *ACS Chem Biol.* 2012; 7:1902–1909. DOI: 10.1021/cb300213g [PubMed: 22958065]
35. Tran T, Disney MD. Identifying the preferred RNA motifs and chemotypes that interact by probing millions of combinations. *Nat Commun.* 2012; 3:1125.doi: 10.1038/ncomms2119 [PubMed: 23047683]
36. Cho J, Rando RR. Specific binding of Hoechst 33258 to site 1 thymidylate synthase mRNA. *Nucleic Acids Res.* 2000; 28:2158–2163. DOI: 10.1093/nar/28.10.2158 [PubMed: 10773086]
37. Pushechnikov A, Lee MM, Childs-Disney JL, Sobczak K, French JM, Thornton CA, Disney MD. Rational design of ligands targeting triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3. *J Am Chem Soc.* 2009; 131:9767–9779. DOI: 10.1021/ja9020149 [PubMed: 19552411]
38. Willett P. Similarity searching using 2D structural fingerprints. *Methods Mol Biol.* 2011; 672:133–158. DOI: 10.1007/978-1-60761-839-3_5 [PubMed: 20838967]
39. Hawkins PC, Skillman AG, Nicholls A. Comparison of shape-matching and docking as virtual screening tools. *J Med Chem.* 2007; 50:74–82. DOI: 10.1021/jm0603365 [PubMed: 17201411]
40. Leeson P. Drug discovery: chemical beauty contest. *Nature.* 2012; 481:455–456. DOI: 10.1038/481455a [PubMed: 22281594]
41. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem.* 2010; 53:2719–2740. DOI: 10.1021/jm901137j [PubMed: 20131845]
42. Bruns RF, Watson IA. Rules for identifying potentially reactive or promiscuous compounds. *J Med Chem.* 2012; 55:9763–9772. DOI: 10.1021/jm301008n [PubMed: 23061697]
43. Paul S, Dansithong W, Kim D, Rossi J, Webster NJ, Comai L, Reddy S. Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant *IR* splicing. *EMBO J.* 2006; 25:4271–4283. DOI: 10.1038/sj.emboj.7601296 [PubMed: 16946708]
44. Cardani R, Mancinelli E, Rotondo G, Sansone V, Meola G. Muscleblind-like protein 1 nuclear sequestration is a molecular pathology marker of DM1 and DM2. *Eur J Histochem.* 2006; 50:177–182. [PubMed: 16920640]
45. Jiang H, Mankodi A, Swanson MS, Moxley RT, Thornton CA. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum Mol Genet.* 2004; 13:3079–3088. DOI: 10.1093/hmg/ddh327 [PubMed: 15496431]
46. Chen CZ, Sobczak K, Hoskins J, Southall N, Marugan JJ, Zheng W, Thornton CA, Austin CP. Two high-throughput screening assays for aberrant RNA-protein interactions in myotonic dystrophy type 1. *Anal Bioanal Chem.* 2012; 402:1889–1898. DOI: 10.1007/s00216-011-5604-0 [PubMed: 22218462]
47. Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. *Genes Dev.* 2003; 17:419–437. DOI: 10.1101/gad.1048803 [PubMed: 12600935]
48. Yang WY, Wilson HD, Velagapudi SP, Disney MD. Inhibition of non-ATG translational events in cells via covalent small molecules targeting RNA. *J Am Chem Soc.* 2015; 137:5336–5345. DOI: 10.1021/ja507448y [PubMed: 25825793]
49. Guan L, Disney MD. Covalent small-molecule-RNA complex formation enables cellular profiling of small-molecule-RNA interactions. *Angew Chem, Int Ed.* 2013; 52:10010–10013. DOI: 10.1002/anie.201301639
50. Childs-Disney JL, Stepniak-Konieczna E, Tran T, Yildirim I, Park H, Chen CZ, Hoskins J, Southall N, Marugan JJ, Patnaik S, Zheng W, Austin CP, Schatz GC, Sobczak K, Thornton CA, Disney MD. Induction and reversal of myotonic dystrophy type 1 pre-mRNA splicing defects by small molecules. *Nat Commun.* 2013; 4:2044.doi: 10.1038/ncomms3044 [PubMed: 23806903]
51. Coonrod LA, Nakamori M, Wang W, Carrell S, Hilton CL, Bodner MJ, Siboni RB, Docter AG, Haley MM, Thornton CA, Berglund JA. Reducing levels of toxic RNA with small molecules. *ACS Chem Biol.* 2013; 8:2528–2537. DOI: 10.1021/cb400431f [PubMed: 24028068]
52. Ketley A, Chen CZ, Li X, Arya S, Robinson TE, Granados-Riveron J, Udosen I, Morris GE, Holt I, Furling D, Chaouch S, Haworth B, Southall N, Shinn P, Zheng W, Austin CP, Hayes CJ, Brook JD.

High-content screening identifies small molecules that remove nuclear foci, affect MBNL distribution and CELF1 protein levels via a PKC-independent pathway in myotonic dystrophy cell lines. *Hum Mol Genet.* 2014; 23:1551–1562. DOI: 10.1093/hmg/ddt542 [PubMed: 24179176]

53. O'Leary DA, Vargas L, Sharif O, Garcia ME, Sigal YJ, Chow SK, Schmedt C, Caldwell JS, Brinker A, Engels IH. HTS-compatible patient-derived cell-based assay to identify small molecule modulators of aberrant splicing in myotonic dystrophy type 1. *Curr Chem Genomics.* 2010; 4:9–18. DOI: 10.2174/1875397301004010009 [PubMed: 20502647]
54. Blount K, Puskarz I, Penchovsky R, Breaker R. Development and application of a high-throughput assay for glmS riboswitch activators. *RNA Biol.* 2006; 3:77–81. DOI: 10.4161/rna.3.2.3102 [PubMed: 17114942]
55. Bandyopadhyay S, Ni J, Ruggiero A, Walshe K, Rogers MS, Chattopadhyay N, Glicksman MA, Rogers JT. A high-throughput drug screen targeted to the 5' untranslated region of Alzheimer amyloid precursor protein mRNA. *J Biomol Screening.* 2006; 11:469–480. DOI: 10.1177/1087057106287271
56. Ross, N.; Metkar, SR.; Le, H.; Burbank, J.; Cahill, C.; Germain, A.; MacPherson, L.; Bittker, J.; Palmer, M.; Rogers, J.; Schreiber, SL. Identification of a Small Molecule That Selectively Inhibits Alpha-Synuclein Translational Expression. National Center for Biotechnology Information (US); Bethesda, MD: 2011.
57. Garcia-Lopez A, Monferrer L, Garcia-Alcover I, Vicente-Crespo M, Alvarez-Abril MC, Artero RD. Genetic and chemical modifiers of a CUG toxicity model in *Drosophila*. *PLoS One.* 2008; 3:e1595.doi: 10.1371/journal.pone.0001595 [PubMed: 18270582]
58. Clark AM, Labute P. Detection and assignment of common scaffolds in project databases of lead molecules. *J Med Chem.* 2009; 52:469–483. DOI: 10.1021/jm801098a [PubMed: 19093885]
59. Warf MB, Nakamori M, Matthys CM, Thornton CA, Berglund JA. Pentamidine reverses the splicing defects associated with myotonic dystrophy. *Proc Natl Acad Sci U S A.* 2009; 106:18551–18556. DOI: 10.1073/pnas.0903234106 [PubMed: 19822739]

Figure 1.

Results from screening the RNA-focused small molecule collection for inhibition of the r(CUG)₁₂-MBNL1 complex. (A) *In vitro* activity of hit compounds for disruption of the r(CUG)₁₂-MBNL1 complex as determined by a TR-FRET assay. Compounds were designated as hits if they disrupted the complex at a level 3 times the standard deviation of all compounds evaluated. Compounds were screened at 100 μM concentration. (B) Structures of compounds that have demonstrated activity *in vitro* and *in cellulo* with IC_{50} values determined by a TR-FRET assay. Compounds **1**, **2**, and **3** were the most active *in*

vitro, while **1**, **16**, and **17** demonstrated the best activity in cells. (C) Substructure analysis of hit compounds identified from the in vitro screen shown in (A). RNA-binding modules identified by the screen include 2-phenyl benzimidazoles, substituted pyridines, and N-phenylbenzamides.

Figure 2.

Bioactivity of hit compounds for improving DM1-associated pre-mRNA splicing defects. (A) Schematic of the pre-mRNA splicing patterns observed for the cTNT mini-gene in the presence and absence of $r(\text{CUG})_{960}$. (B) Quantification of splicing analysis with representative gel images for the most bioactive compounds, **1**, **16**, and **17**. These compounds were selected because they improve DM1 pre-mRNA splicing defects and do not alter cTNT pre-mRNA splicing in the absence of $r(\text{CUG})_{960}$. Compounds were evaluated at 100 μM concentration. The green line indicates the percentage of exon 5 inclusion

observed in healthy, non-DM1 cells, while the pink line indicates the percentage of exon 5 inclusion observed in DM1-affected cells. Asterisk (*) indicates $p < 0.05$ and double asterisk (**) indicates $p < 0.01$. (C) Dose responses for bioactive hit compounds **16** and **17** for improving the cTNT pre-mRNA splicing defect in cells expressing r(CUG)₉₆₀.

Figure 3.

Results of competition dialysis experiments of bioactive small molecules with various RNAs and MBNL1. (A) Secondary structures of the RNAs evaluated by competition dialysis. (B) The compounds evaluated selectively bound RNAs containing r(CUG) repeats compared to other RNAs. Targets were tested at 8 μM concentration. Compound 17 had the highest loading on r(CUG)₁₀₉. Asterisk (*) indicates $p < 0.05$ and double asterisk (**) indicates $p < 0.01$.

Figure 4.

Target validation of bioactive small molecules. (A) Effects of small molecules on *DMPK* mRNA levels derived from the DM mini-gene (100 μ M concentration). Compounds **16** and **17** had no significant effect on *DMPK* levels, while compound **1** causes a decrease in *DMPK* levels as determined by a two-tailed Student t test. Triple asterisk (***) indicates $p < 0.001$. (B) Schematics of Chemical Cross-Linking and Isolation by Pull-down (ChemCLIP) and Competitive ChemCLIP (C-ChemCLIP). Cells containing r(CUG)₉₆₀ are co-treated with a small molecule known to bind and cross-link to r(CUG)^{exp} *in cellulo*, **2H-4-CA-Biotin**, and a bioactive small molecule. Following affinity purification, *in cellulo* binding to the target RNA is assessed by qRT-PCR. (C) Results of C-ChemCLIP with bioactive small molecules **1**, **16**, and **17**. These compounds bound r(CUG)₉₆₀ to a similar extent in cells.

Table 1

Summary of Physicochemical Properties That Influence Bioactivity^a

Compound Class	Molecular Weight	Rotatable Bonds	logP	TPSA (Å ²)	H Bond Donors	H Bond Acceptors	Common Substructure ^{b, c, d}
Starting Library	317 ± 75	3 ± 2	3.4 ± 1.5	75 ± 25	2 ± 1	3 ± 1	
Non-Selectively Bioactive; Affect pre-mRNA Splicing in the absence of r(CUG) ^{exp}							
All Bioactives in Absence of r(CUG) ^{exp}	319 ± 71	3 ± 2	3.0 ± 1.5	85 ± 23	2 ± 1	4 ± 1	
Increases DM1 Splicing Defects	302 ± 73	2.5 ± 2	2.4 ± 1.2 [*]	92 ± 22 [*]	3 ± 1	4 ± 1
Decreases DM1 Splicing Defects	371 ± 37	4 ± 1	4.7 ± 0.8	63 ± 5	2 ± 1	3 ± 1
Selectively Bioactive; Do not Affect pre-mRNA Splicing in the absence of r(CUG) ^{exp} but Affect pre-mRNA in the presence of r(CUG) ^{exp}							
All Bioactives in Presence of r(CUG) ^{exp}	352 ± 68	3.9 ± 2	3.2 ± 1.9	101 ± 33 ^{**}	3 ± 1 ^{**}	4 ± 1 [*]
Lead Compounds	398 ± 27 ^{***}	5 ± 1	4.1 ± 1.3	111 ± 39 ^{**}	2 ± 1	4 ± 2

^a Statistically significant differences observed from the starting library are denoted as^{*} p 0.05;^{**} p 0.01; ^{***} p 0.001.^b Fifty-six percent of compounds that nonselectively increase cTNT exon 5 inclusion in cells without r(CUG)^{exp} are aminobenzimidazoles with a nitrogen heterocycle at the 2 position.^c A common structural feature of compounds that decrease cTNT exon 5 inclusion is 2-phenylbenzimidazole linked to a halogenated benzamide (X = Cl, Br).^d Compounds that have no effect on cTNT splicing in the absence of r(CUG)^{exp} commonly contain a 2-phenylbenzimidazole where R is a secondary amide.