Published in final edited form as:

ACS Chem Biol. 2013 October 18; 8(10): . doi:10.1021/cb400265y.

Features of Modularly Assembled Compounds That Impart Bioactivity Against an RNA Target

Suzanne G. Rzuczek¹, Yu Gao¹, Zhen-Zhi Tang², Charles A. Thornton², Thomas Kodadek¹, and Matthew D. Disney^{1,*}

¹Department of Chemistry, The Scripps Research Institute, Scripps Florida, 130 Scripps Way #3A1, Jupiter, FL 33458

²Department of Neurology, School of Medicine and Dentistry, University of Rochester, Rochester, NY 14642

Abstract

Transcriptomes provide a myriad of potential RNAs that could be the targets of therapeutics or chemical genetic probes of function. Cell permeable small molecules, however, generally do not exploit these targets, owing to the difficulty in the design of high affinity, specific small molecules targeting RNA. As part of a general program to study RNA function using small molecules, we designed bioactive, modularly assembled small molecules that target the non-coding expanded RNA repeat that causes myotonic dystrophy type 1 (DM1), r(CUG)^{exp}. Herein, we present a rigorous study to elucidate features in modularly assembled compounds that afford bioactivity. Different modular assembly scaffolds were investigated including polyamines, α -peptides, β -peptides, and peptide tertiary amides (PTAs). Based on activity as assessed by improvement of DM1-associated defects, stability against proteases, cellular permeability, and toxicity, we discovered that constrained backbones, namely PTAs, are optimal. Notably, we determined that r(CUG)^{exp} is the target of the optimal PTA in cellular models and that the optimal PTA improves DM1-associated defects in a mouse model. Biophysical analyses were employed to investigate potential sources of bioactivity. These investigations show that modularly assembled compounds have increased residence times on their targets and faster on rates than the RNA-binding modules from which they were derived and faster on rates than the protein that binds r(CUG)^{exp}, the inactivation of which gives rise to DM1-associated defects. These studies provide information about features of small molecules that are programmable for targeting RNA, allowing for the facile optimization of therapeutics or chemical probes against other cellular RNA targets.

INTRODUCTION

RNA is an important target for small molecule probes of function or lead therapeutics. Yet, very few RNAs have been exploited as such. Validated targets include ribosomal RNA,^{1, 2} which constitutes 80–90% of total cellular RNA,^{3–5} and riboswitches, that have known metabolite binders that can be mimicked to aid inhibitor design.^{6, 7} Compounds targeting the ribosome and riboswitches have been extremely useful probes to help understand RNA function. One major challenge in RNA chemical biology is how to exploit other RNAs in

* author to whom correspondence should be addressed; disney@scripps.edu.

ASSOCIATED CONTENT

Supporting Information Available: Synthetic procedures, compound characterization, experimental procedures, representative gel images, representative flow cytometry analyses, and Supplementary Tables and Figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

the transcriptome similarly. This is a considerable challenge given the low cellular abundance of these RNAs⁸ and the lack of lead small molecules.⁹

In an effort to exploit other potential RNA targets in the transcriptome with small molecules, we have developed a “bottom-up” strategy to design small molecules that bind an RNA of interest. That is, we define the small, discrete RNA secondary structural elements that are privileged for binding small molecules;^{10–14} the interactions are then deposited into a database. The secondary structural elements in our selection studies are kept intentionally small such that they are likely components of larger cellular RNAs. The secondary structure of an RNA target is compared to our database of interactions, providing lead compounds. Leads can be optimized using various strategies including chemical similarity searching^{15, 16} and/or modular assembly.^{13, 17–22} One application of this strategy has been the rational design of bioactive small molecules that target the RNA that causes myotonic dystrophy type 1 (DM1).^{15, 17, 18}

DM1 is a presently incurable neuromuscular disease caused by a r(CUG) expansion, r(CUG)^{exp}, in the 3' untranslated region (UTR) of the dystrophin myotonia protein kinase (*DMPK*) mRNA.^{23, 24} The RNA folds into a hairpin structure that displays regularly repeating 1×1 nucleotide internal loops (5' CUG/3' GUC motifs; Figure 1)^{25, 26} that are conformationally flexible.^{8, 27} The loops are high affinity binding sites for muscleblind-like 1 protein (MBNL1), and sequestration of MBNL1 causes its inactivation and subsequent dysregulation of alternative pre-mRNA splicing.^{28–31} Formation of the r(CUG)^{exp}-MBNL1 complex causes various disease-associated defects including (Figure 1): (i) pre-mRNA splicing defects;^{28, 32–34} (ii) formation of nuclear foci that consist of r(CUG)^{exp}-protein complexes;^{35–37} and, (iii) translational defects of *DMPK* mRNA due to poor nucleocytoplasmic transport.^{38–40}

Since the root cause of DM1 is r(CUG)^{exp}, a variety of strategies have been employed to disrupt r(CUG)^{exp}-MBNL1 complexes, thus releasing MBNL1 and restoring regulation of alternative splicing. Oligonucleotides that target r(CUG)^{exp} improve DM1-associated defects upon injection into DM1 mouse models.^{33, 41, 42} Small molecules have also been developed that target r(CUG)^{exp} including pentamidine, *bis*-benzimidazoles, naphthyl pyridines, and triazines.^{13, 15, 17–19, 43–46} The most potent are modularly assembled compounds that target the repeating nature of r(CUG)^{exp}, binding multiple 5' CUG/3' GUC motifs simultaneously.^{13, 17–19} These compounds are composed of a modular assembly scaffold that displays multiple copies of an RNA-binding module on a single chain.^{13, 17–19}

Herein, we describe detailed studies aimed at understanding features in modular assembly scaffolds that provide bioactivity. It is shown that the nature of the scaffold significantly affects potency and bioactivity. By studying different scaffolds including polyamines, α -peptides, β -peptides, and peptide-tertiary amides (PTAs), we determined that the PTA scaffold is optimal based on bioactivity in two different cellular assays, stability against proteases, cellular permeability, and toxicity. These studies define features that can be programmed into small molecules to afford bioactivity and can be applied to other potential RNA targets in the transcriptome.

RESULTS AND DISCUSSION

Design of Modularly Assembled Compounds

We previously reported the design of modularly assembled compounds that inhibit the r(CUG)^{exp}-MBNL1 complex *in vitro*¹³ and *in vivo*¹⁸. The compounds are comprised of a peptoid backbone that displays multiple copies of the bis-benzimidazole **Ht** separated by four spacing modules, or **nH-4** compounds (Figure 2). The **2H-4** compound is the most

efficacious in DM1 cell culture models, improving various DM1-associated defects (Figure 1).¹⁸ In an effort to understand features in modularly assembled compounds that afford high affinity binding to r(CUG)^{exp} and bioactivity, we designed and synthesized a series of second-generation compounds in which the assembly scaffold was varied. In particular, we studied polyamine, α -peptide, β -peptide, peptoid, and peptide tertiary amide (PTA) assembly scaffolds (Figure 2). These scaffolds allowed us to study of the effect of: (i) forming favorable interactions with a negatively charged phosphodiester backbone (polyamines); (ii) amide bond spacing (α - and β -peptides); (iii) *N*-alkylation (α - and β -peptides, peptoids, and PTAs); and, (iv) chirality of the C- α carbon atom (PTAs and peptoids). PTAs are conformationally restricted, which could pre-organize them for recognition of an RNA target (Figure 2). Since the most potent first generation compound was a dimer (**2H-4**),¹⁸ we focused on dimers of each assembly scaffold. Compounds were synthesized on solid-phase resin and coupled to an **Ht** derivative that contains a carboxylate. (**Ht** binds 5' CUG/3' GUC, and thus is the RNA-binding module.¹³) It should be noted that all second-generation compounds are soluble in water at 10 mM concentration and 100 μ M concentration in cell culture medium.

Second generation compounds are named **2H-X** where **2H** indicates two **H** RNA-binding modules and **X** is specific for each scaffold. The nomenclature for each scaffold is: (i) polyamines: **2H-P** where **P** indicates the particular polyamine (ED is ethylenediamine; BD is butanediamine; HD is hexanediamine; SPD is spermidine; SPM is spermine); (ii) α -peptides: **2H-nG** where **nG** indicates the number of glycine residues; (iii) β -peptides: **2H-n β A** where **n β A** indicates the number of β -alanine residues; and (iv) PTAs: **2H-nNX** where **NX** indicates the particular *N*-alkylated spacing module and **n** indicates the number of modules.

Effect of the Assembly Scaffold on *In Vitro* Potency

As mentioned above, DM1 is caused by sequestration of MBNL1 by r(CUG)^{exp}.^{28, 32–37} Therefore, *in vitro* potency was determined by measuring displacement of MBNL1 from a r(CUG)₁₂-MBNL1 complex using a time-resolved FRET (TR-FRET) assay.^{15, 47} Each dimeric compound disrupts the r(CUG)₁₂-MBNL1 complex with IC₅₀'s between ~3 and ~300 μ M (Figure 3). The most potent compounds are: **2H-SPM**, **2H-3G**, and **2H-4G**. Polyamines and α -peptides (polyglycines) are on average more potent than PTAs and peptoids. This may be due to the presence of imino protons that can favorably interact with the phosphodiester backbone of RNA (Figure 2).

The spacing between **H** RNA-binding modules affects *in vitro* potency by more than an order of magnitude in almost all scaffold classes (Figure 3), as previously observed.^{13, 19, 21} For example, **2H-SPM** has an IC₅₀ of ~3 μ M while **2H-ED** has an IC₅₀ of ~30 μ M. The most potent α -peptides are **2H-3G** and **2H-4G** (IC₅₀ $\approx$ 3 μ M) while **2H-6G** is the least potent (IC₅₀ $\approx$ 30 μ M). Similar differences are observed for **2H-n β A**, **2H-nNMe**, and **2H-KnNPr**. In summary, the most potent compound in each assembly scaffold class is: polyamines, **2H-SPM**; peptides, **2H-3G** $\approx$ **2H-4G** (only **2H-3G** was further investigated); PTAs, **2H-3NPr**, **2H-3NMe**, and **2H-K4NPr**.

Effect of the Assembly Scaffold on Proteolytic Stability

(Supplementary Table S-1). We investigated the proteolytic stability of the most potent compound in each class by measuring its susceptibility to pronase digestion at pH 7.8 and pH 9.0 for 18 h.⁴⁸ The most proteolytically stable compound is **2H-SPM** (no observed degradation) while **2H-3NMe** is the least stable (completely degraded). Intermediate stability against pronase digestion was observed for the remaining compounds.

Effect of the Assembly Scaffold on Bioactivity

The best compound from each scaffold class was studied for improving DM1-associated defects in two DM1 cellular models (Figure 4A & C and S-1). As mentioned above, sequestration of MBNL1 by r(CUG)^{exp} causes alternative pre-mRNA splicing defects.^{28, 32–34} Therefore, we studied the MBNL-regulated alternative splicing of exon 5 of cardiac troponin T (*cTNT*) pre-mRNA.^{49, 50} HeLa cells were transiently transfected with a DM1 mini-gene that encodes r(CUG)^{exp} and a *cTNT* mini-gene^{45, 51} followed by treatment with 10 μ M compound. In the absence of r(CUG)^{exp}, the exon 5 inclusion rate is ~45% while in the presence of r(CUG)^{exp}, the exon 5 inclusion rate is ~85% (Figure 4A & B). Interestingly, not all of the compounds tested improve the *cTNT* pre-mRNA splicing defect (Figure 4B). There is some correlation between *in vitro* potency and improvement of splicing defects (Figures 3 & 4B). Based on *in vitro* potency, one would predict the following *in vivo* potencies: **2H-3G** $\approx$ **2H-SPM** > **2H-3NPr** > **2H-3NMe** $\approx$ **2H-K4NPr**. Although these trends are generally observed, **2H-K4NPr** is an outlier as it is actually the most potent *in vivo* for improving the *cTNT*-splicing defect.

Next, compounds were studied for improving translational defects associated with DM1 (due to poor nucleocytoplasmic transport of r(CUG)^{exp}-containing transcripts).^{38–40} In particular, a cellular model system was employed in which the C2C12 cell line stably expresses r(CUG)₈₀₀ embedded in the 3' UTR of firefly luciferase.¹⁸ Akin to *DMPK* mRNA in DM1-affected cells, the expanded repeat impairs nucleocytoplasmic transport of luciferase mRNA and thus decreases luciferase expression. Compounds that bind r(CUG)^{exp} and disrupt the r(CUG)^{exp}-MBNL1 complex may stimulate cytoplasmic transport of the luciferase mRNA and thus translation of luciferase (Figure 4C). A significant increase in luciferase activity (and hence improvement of the DM1-associated translational defect) is observed for all compounds (Figure 4D). There is little difference between the potencies of the compounds, with **2H-SPM** having less improvement than the others. Importantly, the compounds do not affect luciferase activity of cells that stably express firefly luciferase mRNA without r(CUG)₈₀₀ (10 μ M compound).

Effect of the Assembly Scaffold on Cellular Permeability and Toxicity

Next, we studied the effect of the assembly scaffold on cellular permeability and toxicity by flow cytometry (Figure S-4). Permeability was monitored by using the intrinsic fluorescence of the **H** module while toxicity was measured by propidium iodide (PI) staining. Cells were treated with 1 or 10 μ M compound for 24 h and then analyzed. There is little difference in the cellular permeability or toxicity of second-generation compounds when dosed with 10 μ M compound (Figure S-4), with the exception of **2H-SPM**, which shows significant toxicity. Cellular permeability does not change at 1 μ M dosage, except for **2H-SPM**, which is reduced by 50%. Taken together, permeability and toxicity data suggest that polyamines may be suboptimal modular assembly scaffolds. With the exception of **2H-SPM**, all compounds are more cell permeable and less toxic than the parent compound, **2H-4**, when cells are treated with 1 μ M compound (Figure S-4). It has been shown that *N*-alkylated glycines such as peptoids and PTAs enter cells in greater amounts than peptides^{52, 53} and that structurally constrained peptides are more cell permeable than their unstructured counterparts.^{54–57}

Effect of *N*-alkyl Side Chains in PTAs

Considering all *in vivo* data, the optimal compound is **2H-K4NPr** (Figures 3, 4, and S-4). It most effectively improves splicing defects (Figure 4B), has comparable activity for improving translational defects (Figure 4D), is highly cell permeable and has low toxicity (Figure S-4), and is proteolytically (and likely metabolically⁴⁸) stable. Therefore, we

investigated the impact of the spacing module in **2H-K4NPr** on bioactivity. Specifically, the influences of *N*-alkylation and α -carbon substitution were studied (Figure 5A). Previous studies have shown that the spacing module affects affinity, potency, cellular permeability, and localization.²⁰

In vitro potencies were determined using the TR-FRET assay described above. IC₅₀ values ranged from 20 to 55 μ M. The *N*-methyl PTA, **2H-K4NMe**, is most potent while **2H-K4NiBu** is least potent (Figure 5B). This suggests that bulky substituents may hinder tight association between r(CUG)₁₂ and the small molecule. Interestingly, **2H-K4NH** and **2H-K4NiBu** are proteolytically unstable (Supplementary Table S-1).

The compounds were assessed for improving DM1-associated pre-mRNA splicing defects when cells are dosed at 10 μ M. The most efficacious compounds are the parent **2H-K4NPr** and **2H-K4NH** (no tertiary amide) followed by **2H-K4NMe** and **2H-K4** (no substituent on the α -C); **2H-K4NiBu** is inactive. The **2H-K4NPr** derivatives were also evaluated for improving translational defects using the luciferase reporter system. All derivatives significantly increase luciferase expression to a similar extent when cells are treated with 10 μ M compound.

Next, the cellular permeability and toxicity of each **2H-K4NPr** derivative were determined via flow cytometry (Figure 5E). As observed for second-generation compounds, there is little difference in permeability when cells are dosed with 10 μ M compound with the exception of **2H-K4**, a peptoid. A marked difference is observed for **2H-K4** and third generation compounds at the 1 μ M dosage as the other derivatives (peptides and PTAs) are ~4 times more permeable. In general, little toxicity is observed for **2H-K4NPr** derivatives with the exception of **2H-K4NiBu**, which has significant toxicity at both 1 and 10 μ M, and **2H-K4**, which shows some toxicity when cells are treated with 10 μ M compound. (Figure 5E).

In summary, **iBu** substitution is disadvantageous for bioactivity, rendering the compound inactive in both cellular assays (Figure 5C & D) and increasing cellular toxicity (Figure 5E). Substitution of **Pr** with **H** or **Me** is better tolerated, although **Me** substitution decreases improvement of splicing defects (Figure 5C). In contrast, **2H-K4NH** has comparable activity as **2H-K4NPr** as assessed by restoring pre-mRNA splicing patterns. Cellular permeability and toxicity are similar for **2H-K4NPr**, **2H-K4NMe**, and **2H-K4NH**. Removal of the methyl group on the α -carbon, affording **2H-K4**, negatively affects cellular permeability (Figure 5E).

Further Analysis of **2H-K4NPr** and **2H-K4NMe**

We further investigated **2H-K4NPr** and **2H-K4NMe** based on their abilities to improve DM1-associated defects and because of their greater cell permeabilities (Figure 5E) and proteolytic/metabolic stabilities (Supplementary Table S-1) than **2H-K4NH** and **2H-K4**. First, a dose response for improvement of pre-mRNA splicing defects was measured. Both compounds improve the *cTNT* splicing defect at 5 μ M (Supplementary Figure S-3) but not at 1 μ M (data not shown). Interestingly, **2H-K4NMe** is slightly better than **2H-K4NPr** at this concentration (Supplementary Figure S-3A). Neither compound affects *cTNT* splicing in cells that do not express r(CUG)^{exp} (Supplementary Figure S-3). A dose response for improving translational defects was also completed for **2H-K4NMe**. Significant activity was observed when cells were treated with as little as 1 μ M compound (Supplementary Figure S-3B). Both **2H-K4NPr** and **2H-K4NMe** appear to improve alternative pre-mRNA splicing and translational defects at concentrations somewhat less than their *in vitro* IC₅₀s. These observations could be due to: (i) *in vitro* and *in vivo* assays measure very different

phenomena: inhibition of the r(CUG)-MBNL1 interaction (*in vitro*) as compared to the activity of free MBNL1 (*in vivo*); (ii) a small effect on the translational defect is also observed for the PTA backbone (**K4NMe**; Supplementary Figure S-3B); and (iii) a recent study by Jog et al. reported that once a critical threshold of MBNL1 concentration is achieved, relatively small changes in the amount of active MBNL1 significantly affects the number and severity of splicing defects.⁵⁸ For example, when MBNL1 levels are unaffected, the percentage of cTNT exon 5 inclusion is 49.5%. When MBNL1 levels are decreased by 79%, 87%, 97%, and 98% via siRNA knock-down, the percentages of exon inclusion are 48.2% (no change as compared to cells treated with a scrambled siRNA), 54.3%, 81.5%, and 88.4%.⁵⁸ It should be noted that these effects are unique for each pre-mRNA substrate.⁵⁸ Taking together activity, stability, toxicity, and ease of synthesis, **2H-K4NMe** is more ideal than **2H-K4NPr**.

Additional control experiments were completed with **2H-K4NMe**. To ensure that alternative splicing of pre-mRNAs not regulated by MBNL1 are unaffected, we studied the alternative splicing of pleckstrin homology domain containing, family H member 2 (*PLEKHH2*) using a mini-gene construct.⁴⁵ Splicing of *PLEKHH2* was not affected at 10 μ M (Supplementary Figure S-2), demonstrating **2H-K4NMe**'s selectivity for r(CUG)^{exp} and alternative pre-mRNA splice events regulated by MBNL1. Moreover, the unconjugated PTA backbone, **K4NMe**, shows little activity in the translational assay (Supplementary Figure S3-B).

The binding of **2H-K4NMe** to various RNAs including r(CUG)₁₂, r(CAG)₁₂, r(AU)₁₂ and r(GC)₈ (Figure S-5) was studied by surface plasmon resonance (SPR) spectrometry. In summary, **2H-K4NMe** is selective for r(CUG)₁₂ by ~30–100-fold over the other RNAs studied (Table 1). In general, it has a faster association rate (k_{on}) and a slower dissociation rate (k_{off}) when binding to r(CUG)₁₂ compared to the other RNAs.

The kinetic parameters for **2H-K4NMe** binding to r(CUG)₁₂ show that this compound has advantages over the monomeric RNA binding module and MBNL1 (Table 1). For example, **H** binds r(CUG)₁₂ with a K_d of 2000 nM, or 15-fold weaker than **2H-K4NMe**. Previous studies show that **2H-4** only had an affinity bonus of 1.5-fold compared to monomer.¹³ The increased affinity of **2H-K4NMe** could be due to improved interactions of the PTA scaffold with the RNA or to the more pre-organized nature of the PTA backbone. PTAs have limited flexibility due to 1,3-allylic strain introduced by the substituents on both the α -carbon and the iminonitrogen and thus sample less conformational space when finding the optimal binding mode for r(CUG)₁₂. Additionally, the k_{on} and k_{off} rates are 10-fold faster and 10-fold slower for **2H-K4NMe** than **H**, suggesting a hinge effect (Supplementary Figure S-5) provides benefits for both the rates of complex formation and complex dissociation for modularly assembled compounds.

2H-K4NMe binds 5-fold more tightly to r(CUG)₁₂ than MBNL1 does. High affinity binding is advantageous for compounds that target r(CUG)^{exp} as lower concentrations of compound are likely to be required for bioactivity, thus minimizing potential non-specific effects at efficacious doses. **2H-K4NMe** also has kinetic advantages for binding r(CUG)₁₂ over MBNL1 as its on rate is 79-fold faster than MBNL1's (Table 1).

These binding kinetic observations may be due to the nature in which the small molecule and MBNL1 bind to RNA targets. As a regulator of alternative splicing, MBNL1 must interact specifically with RNA and does so via four zinc finger (ZnF) domains. A crystal structure of the ZnF3/4 domain complexed with two copies of single stranded r(CGCUGU) was recently reported.⁵⁹ Both zinc fingers interact with one molecule of RNA with ZnF3 forming contacts to the 5'GC step and ZnF4 forming contacts to 5'GCU. The RNA molecules are oriented antiparallel to each other. These results suggest that MBNL1 binding

induces a chain-reversal trajectory and unzipping of the bound RNA. It is likely that MBNL1 opens the r(CUG) hairpin stem to afford two single stranded regions, which is supported by the fact that MBNL1 binds weakly to fully base-paired RNAs.⁶⁰ The increased on rate for the designer small molecule onto r(CUG)^{exp} may be a consequence of such molecular recognition. That is, the small molecule binds the native hairpin structure of r(CUG)^{exp} and stabilizes it whereas MBNL1 initially interacts with r(CUG)^{exp}, then disrupts or remodels the hairpin structure to afford the final bound conformation.

Identifying the Cellular Targets of 2H-K4NMe

The cellular targets of the optimal r(CUG)^{exp}-targeting compound were identified using a modified version of a pull down experiment (Figure 6). Briefly, a biotinylated **2H-K4NMe** derivative (**2H-K4NMe-Biotin**) was synthesized and anchored onto streptavidin resin to afford an affinity purification matrix. Total RNA was extracted from a DM1 cellular model and incubated with the affinity matrix. After extensive washing, northern blotting revealed that r(CUG)^{exp} is highly enriched in the eluted RNA bound by **2H-K4NMe-Biotin** (Figure 6). Interestingly, r(CUG)^{exp} was not pulled down by streptavidin beads that display an oligonucleotide complementary to r(CUG)^{exp}, 5'-biotind(CAG)₁₂. This is likely due to the highly structured nature of the r(CUG)^{exp} target.^{8, 25, 27, 60–62} Since oligonucleotides recognize sequence, formation of a duplex between r(CUG)^{exp} and a complementary oligonucleotide can only occur after the intramolecular structure is unfolded, a significant barrier for complex formation. Small molecules that recognize structure would not have this barrier for binding to highly structured RNA targets.

2H-K4NMe Improves Splicing Defects in a Mouse Model of DM1

2H-K4NMe was selected for animal studies based on overall performance in cell-based assays, increased proteolytic stability as compared to **2H-KNH** and decreased structural complexity compared to **2H-K4NPr**. A mouse model of DM1 has been reported in which 250 rCUG repeats are expressed using an actin promoter (human skeletal actin long repeat, HSA^{LR}).²⁴ The presence of these repeats results in dysregulation of alternative splicing in the muscle-specific chloride ion channel (*Clcn1*) and the sarcoplasmic/endoplasmic reticulum calcium ATPase 1 (*Sercaf/Atp2a1*) pre-mRNAs. *Clcn1* exon 7a is excluded ~100% in mRNA from normal adult mice; DM1 mice have an exclusion rate of ~50%. After DM1 mice were treated for 7 days with 100 mg/kg/d of **2H-K4NMe**, the exclusion rate of exon 7A was restored to 71% (*p*-value = 0.007) (Figure 7 and S-6). Mis-splicing of *Sercaf* exon 21 was also partially rescued. The exon 22 inclusion rate is ~100% in normal adult mice and ~11.8% in DM1 mice. After treatment with **2H-K4NMe**, the inclusion rate increases to 21.5% (*p*-value = 0.008). Importantly, **2H-K4NMe** did not affect the alternative splicing of *Itgb1* and *CapZB* mRNAs, which are not regulated by MBNL1 (Supplementary Figure S-7).

Summary and Outlook

In this study we investigated the role of the modularly assembly scaffold on *in vitro* and *in vivo* activities of compounds designed to target the RNA that causes myotonic dystrophy type 1. Based on a variety of criteria, including *in vitro* and *in vivo* potency, proteolytic stability, cellular permeability, and toxicity, the optimal scaffold is a PTA. Importantly, a kinetics investigation showed that modularly assembled small molecules have faster on rates and slower off rates than monomeric small molecules or MBNL1 protein, perhaps owing to a hinge effect to difference in the nature of the binding of small molecule and protein to r(CUG)^{exp}. These kinetic advantages likely explain the potent biological effects of our divalent compounds, as the RNA-binding modules from which they are composed of by

themselves are inactive. Such effects can be programmed into designer small molecules targeting other cellular RNAs.

In addition, we leveraged the chemistry allowed by a modular assembly approach to incorporate a biotin tag. This modification allowed the pull down of cellular targets of a small molecule using streptavidin beads and demonstrated that r(CUG)^{exp} is indeed a target of our designed compounds. This is the first study that identified r(CUG)^{exp} as a cellular target of small molecules that improve DM1-associated defects and further supports our design strategy. Other tags can be incorporated into modularly assembled compounds including warheads for targeted cleavage or covalent cross-linking.^{63, 64}

Importantly, our optimal compound improves DM1-associated alternative splicing defects in a mouse model and thus has therapeutic potential. Our future studies will focus on identifying off-targets, and thus potential modes of toxicity, in cellular models and DM1 fibroblasts using RNA-seq. Such information will be essential to engineer enhanced selectivity of our compounds, which may be afforded by further altering the modular assembly scaffold or by altering the RNA-binding module. Since the number of RNAs that contribute to disease is quickly expanding from genome sequencing and transcriptome-wide studies, it is likely that our investigations form a foundation for programming properties into small molecules that engender bioactivity.

METHODS

Detailed experimental descriptions for compound synthesis and characterization, RT-PCR analysis, flow cytometry, SPR experiments, target pull down and Northern blotting can be found in the Supporting Information.

Time-Resolved Fluorescence Resonance Energy Transfer (TR-FRET) Assay

Initial *in vitro* activity of modularly assembled compounds was assessed by disruption of the r(CUG)^{exp}-MBNL1 complex using a previously reported TR-FRET assay.^{15, 47} The ratio of fluorescence intensity of 545 and 665 nm in the presence and absence of compound were used to calculate percent inhibition. The resulting curves were fit to the following equation to determine IC₅₀ values:

$$y = B + \frac{A - B}{1 + \left(\frac{IC_{50}}{x}\right)^{hillslope}}$$

where y is the ratio of fluorescence intensities at 545 nm and 665 nm (F₅₄₅/F₆₆₅), x is the concentration of small molecule, B is the F₅₄₅/F₆₆₅ ratio reflective of the maximum amount of FRET observed (absence of small molecule); A is the F₅₄₅/F₆₆₅ ratio reflective of the minimum amount of FRET observed (absence of RNA, protein, and small molecule; only antibodies are present); and the IC₅₀ is the concentration of small molecule where half of the protein is displaced by small molecule.

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

DM1-associated alternative pre-mRNA splicing defects were assessed using a previously reported model system.⁴⁵ Briefly, HeLa cells were grown as monolayers in 96-well plates in growth medium (1X DMEM, 10% FBS, and 1X Glutamax (Invitrogen)). Cells were transfected at 90–95% confluency 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 interrupted CTG repeats⁵¹ and a mini-gene of interest (*cTNT*⁵¹ or *PLEKHH2*⁴⁵) 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 Sigma-Aldrich GeneElute total mammalian 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.¹⁵

Improvement of DM1-Associated Translational Defects Using a Luciferase Model System

C2C12 cell lines expressing 800 or 0 CTG repeats in the 3' UTR of luciferase were grown as monolayers in 96-well plates in growth medium (1X DMEM, 10% FBS, 1X Glutamax, (Invitrogen) and 1X penicillin/streptomycin (MP Biomedicals LLC)). After 24 h, the compound of interest was added in 50 μ L of growth medium. Cells were treated with compound for 24 h. Luciferase activity was determined as previously described.¹⁸

Treatment in Mice

All experimental procedures, mouse handling, and husbandry were completed in accordance with the Association for Assessment and Accreditation of Laboratory Animal Care. A mouse model for DM1, HSA^{LR} in line 20b, was used.²⁴ HSA^{LR} mice express human skeletal actin RNA with 250 CUG repeats in the 3' UTR. Age- and gender-matched HSA^{LR} mice were injected intraperitoneally with 100 mg/kg **2H-K4NMe** in water for treatment or 0.9% NaCl for control once per day for 7 days. Mice were sacrificed one day after the last injection, and the vastus muscle was obtained. RNA was extracted from the vastus tissue, and cDNA was synthesized as previously described.⁶⁵

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank K. Lowe for assistance with flow cytometry experiments and J. Childs-Disney for assistance with SPR data analysis and manuscript review. This work was funded by a postdoctoral fellowship from the Myotonic Dystrophy Foundation (to SGR), by the Wellstone Muscular Dystrophy Cooperative Research Center (U54NS48843 to CAT), and by the US National Institutes of Health (R01-GM079235 to MDD).

REFERENCES

- Poehlsgaard J, Douthwaite S. The bacterial ribosome as a target for antibiotics. *Nat. Rev. Microbiol.* 2005; 3:870–881. [PubMed: 16261170]
- Projan SJ. New (and not so new) antibacterial targets - from where and when will the novel drugs come? *Curr. Opin. Pharmacol.* 2002; 2:513–522. [PubMed: 12324252]
- Berg, JM.; Tymoczko, JL.; Stryer, L. *Biochemistry*, Sixth Edition. New York: W.H. Freeman and Company; 2007.
- Johnson LF, Abelson HT, Penman S, Green H. The relative amounts of the cytoplasmic RNA species in normal, transformed and senescent cultured cell lines. *J. Cell. Physiol.* 1977; 90:465–470. [PubMed: 192738]
- Johnson LF, Williams JG, Abelson HT, Green H, Penman S. Changes in RNA in relation to growth of the fibroblast. III. Posttranscriptional regulation of mRNA formation in resting and growing cells. *Cell.* 1975; 4:69–75. [PubMed: 1078787]
- Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. *Nat. Biotechnol.* 2006; 24:1558–1564. [PubMed: 17160062]

7. Blount KF, Wang JX, Lim J, Sudarsan N, Breaker RR. Antibacterial lysine analogs that target lysine riboswitches. *Nat. Chem. Biol.* 2007; 3:44–49. [PubMed: 17143270]
8. Kumar A, Park H, Fang P, Parkesh R, Guo M, Nettles KW, Disney MD. Myotonic dystrophy type 1 RNA crystal structures reveal heterogeneous 1 × 1 nucleotide UU internal loop conformations. *Biochemistry.* 2011; 50:9928–9935. [PubMed: 21988728]
9. Guan L, Disney MD. Recent advances in developing small molecules targeting RNA. *ACS Chem. Biol.* 2012; 7:73–86. [PubMed: 22185671]
10. Childs-Disney JL, Wu M, Pushechnikov A, Aminova O, Disney MD. A small molecule microarray platform to select RNA internal loop-ligand interactions. *ACS Chem. Biol.* 2007; 2:745–754. [PubMed: 17975888]
11. Velagapudi SP, Seedhouse SJ, Disney MD. Structure-activity relationships through sequencing (StARTS) defines optimal and suboptimal RNA motif targets for small molecules. *Angew. Chem. Int. Ed. Engl.* 2010; 49:3816–3818. [PubMed: 20397174]
12. 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. [PubMed: 18652457]
13. 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. [PubMed: 19552411]
14. Velagapudi SP, Seedhouse SJ, French J, Disney MD. Defining the RNA internal loops preferred by benzimidazole derivatives via 2D combinatorial screening and computational analysis. *J. Am. Chem. Soc.* 2011; 133:10111–10118. [PubMed: 21604752]
15. 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. [PubMed: 22300544]
16. Kumar A, Parkesh R, Sznajder LJ, Childs-Disney JL, Sobczak K, Disney MD. Chemical correction of pre-mRNA splicing defects associated with sequestration of muscleblind-like 1 protein by expanded r(CAG)-containing transcripts. *ACS Chem. Biol.* 2012; 7:496–505. [PubMed: 22252896]
17. Childs-Disney JL, Parkesh R, Nakamori M, Thornton CA, Disney MD. Rational design of bioactive, modularly assembled aminoglycosides targeting the RNA that causes myotonic dystrophy type 1. *ACS Chem. Biol.* 2012; 7:1984–1993. [PubMed: 23130637]
18. Childs-Disney JL, Hoskins J, Rzuczek SG, Thornton CA, Disney MD. Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potentially bioactive. *ACS Chem. Biol.* 2012; 7:856–862. [PubMed: 22332923]
19. Lee MM, Childs-Disney JL, Pushechnikov A, French JM, Sobczak K, Thornton CA, Disney MD. Controlling the specificity of modularly assembled small molecules for RNA via ligand module spacing: targeting the RNAs that cause myotonic muscular dystrophy. *J. Am. Chem. Soc.* 2009; 131:17464–17472. [PubMed: 19904940]
20. Lee MM, French JM, Disney MD. Influencing uptake and localization of aminoglycoside-functionalized peptoids. *Mol. Biosyst.* 2011; 7:2441–2451. [PubMed: 21611644]
21. Lee MM, Pushechnikov A, Disney MD. Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2. *ACS Chem. Biol.* 2009; 4:345–355. [PubMed: 19348464]
22. Disney MD, Lee MM, Pushechnikov A, Childs-Disney JL. The role of flexibility in the rational design of modularly assembled ligands targeting the RNAs that cause the myotonic dystrophies. *Chembiochem.* 2010; 11:375–382. [PubMed: 20058255]
23. 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. [PubMed: 1310900]
24. Mankodi A, Logigian E, Callahan L, McClain C, White R, Henderson D, Krym M, Thornton CA. Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science*. 2000; 289:1769–1773. [PubMed: 10976074]
25. Tian B, White RJ, Xia T, Welle S, Turner DH, Mathews MB, Thornton CA. Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR. *RNA*. 2000; 6:79–87. [PubMed: 10668800]
26. Sobczak K, Michlewski G, de Mezer M, Kierzek E, Krol J, Olejniczak M, Kierzek R, Krzyzosiak WJ. Structural diversity of triplet repeat RNAs. *J. Biol. Chem*. 2010; 285:12755–12764. [PubMed: 20159983]
27. Parkesh R, Fountain M, Disney MD. NMR spectroscopy and molecular dynamics simulation of r(CCGCUGCGG)₂ reveal a dynamic UU internal loop found in myotonic dystrophy type 1. *Biochemistry*. 2011; 50:599–601. [PubMed: 21204525]
28. Miller JW, Urbinati CR, Teng-Umnay 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. [PubMed: 10970838]
29. 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. [PubMed: 16946708]
30. 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]
31. 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. [PubMed: 15496431]
32. Kanadia RN, Johnstone KA, Mankodi A, Lungu C, Thornton CA, Esson D, Timmers AM, Hauswirth WW, Swanson MS. A muscleblind knockout model for myotonic dystrophy. *Science*. 2003; 302:1978–1980. [PubMed: 14671308]
33. 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. [PubMed: 19608921]
34. Orengo JP, Ward AJ, Cooper TA. Alternative splicing dysregulation secondary to skeletal muscle regeneration. *Ann. Neurol*. 2011; 69:681–690. [PubMed: 21400563]
35. 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. [PubMed: 7896884]
36. Mankodi A, Urbinati CR, Yuan QP, Moxley RT, Sansone V, Krym M, Henderson D, Schalling M, Swanson MS, Thornton CA. Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet*. 2001; 10:2165–2170. [PubMed: 11590133]
37. Wojciechowska M, Krzyzosiak WJ. Cellular toxicity of expanded RNA repeats: focus on RNA foci. *Hum. Mol. Genet*. 2011; 39:8938–8951.
38. Mastroiannopoulos NP, Feldman ML, Uney JB, Mahadevan MS, Phylactou LA. Woodchuck post-transcriptional element induces nuclear export of myotonic dystrophy 3' untranslated region transcripts. *EMBO Rep*. 2005; 6:458–463. [PubMed: 15832171]
39. Amack JD, Mahadevan MS. The myotonic dystrophy expanded CUG repeat tract is necessary but not sufficient to disrupt C2C12 myoblast differentiation. *Hum. Mol. Genet*. 2001; 10:1879–1887. [PubMed: 11555624]
40. 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. [PubMed: 10484765]
41. Lee JE, Bennett CF, Cooper TA. RNase H-mediated degradation of toxic RNA in myotonic dystrophy type 1. *Proc. Natl. Acad. Sci. U. S. A*. 2012; 109:4221–4226. [PubMed: 22371589]

42. Kanadia RN, Shin J, Yuan Y, Beattie SG, Wheeler TM, Thornton CA, Swanson MS. Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:11748–11753. [PubMed: 16864772]
43. Gareiss PC, Sobczak K, McNaughton BR, Palde PB, Thornton CA, Miller BL. Dynamic combinatorial selection of molecules capable of inhibiting the (CUG) repeat RNA-MBNL1 interaction in vitro: Discovery of lead compounds targeting myotonic dystrophy (DM1). *J. Am. Chem. Soc.* 2008; 130:16254–16261. [PubMed: 18998634]
44. Arambula JF, Ramisetty SR, Baranger AM, Zimmerman SC. A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:16068–16073. [PubMed: 19805260]
45. 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. [PubMed: 19822739]
46. Garcia-Lopez A, Llamusi B, Orzaez M, Perez-Paya E, Artero RD. In vivo discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:11866–11871. [PubMed: 21730182]
47. 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. [PubMed: 22218462]
48. Hook DF, Gessier F, Noti C, Kast P, Seebach D. Probing the proteolytic stability of beta-peptides containing alpha-fluoro- and alpha-hydroxy-beta-amino acids. *Chembiochem.* 2004; 5:691–706. [PubMed: 15122642]
49. Orenge JP, Bundman D, Cooper TA. A bichromatic fluorescent reporter for cell-based screens of alternative splicing. *Nucleic Acids Res.* 2006; 34:e148. [PubMed: 17142220]
50. Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. *Genes Dev.* 2003; 17:419–437. [PubMed: 12600935]
51. Philips AV, Timchenko LT, Cooper TA. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science.* 1998; 280:737–741. [PubMed: 9563950]
52. Yu P, Liu B, Kodadek T. A high-throughput assay for assessing the cell permeability of combinatorial libraries. *Nat. Biotechnol.* 2005; 23:746–751. [PubMed: 15908941]
53. Kwon YU, Kodadek T. Quantitative evaluation of the relative cell permeability of peptoids and peptides. *J. Am. Chem. Soc.* 2007; 129:1508–1509. [PubMed: 17283989]
54. Verdine GL, Hilinski GJ. Stapled peptides for intracellular drug targets. *Methods Enzymol.* 2012; 503:3–33. [PubMed: 22230563]
55. Moellering RE, Cornejo M, Davis TN, Del Bianco C, Aster JC, Blacklow SC, Kung AL, Gilliland DG, Verdine GL, Bradner JE. Direct inhibition of the NOTCH transcription factor complex. *Nature.* 2009; 462:182–188. [PubMed: 19907488]
56. Bautista AD, Appelbaum JS, Craig CJ, Michel J, Schepartz A. Bridged beta(3)-peptide inhibitors of p53-hDM2 complexation: correlation between affinity and cell permeability. *J. Am. Chem. Soc.* 2010; 132:2904–2906. [PubMed: 20158215]
57. Daniels DS, Schepartz A. Intrinsically cell-permeable miniature proteins based on a minimal cationic PPII motif. *J. Am. Chem. Soc.* 2007; 129:14578–14579. [PubMed: 17983240]
58. Jog SP, Paul S, Dansithong W, Tring S, Comai L, Reddy S. RNA splicing is responsive to MBNL1 dose. *PLoS One.* 2012; 7:e48825. [PubMed: 23166594]
59. Teplova M, Patel DJ. Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat. Struct. Mol. Biol.* 2008; 15:1343–1351. [PubMed: 19043415]
60. Warf MB, Berglund JA. MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA.* 2007; 13:2238–2251. [PubMed: 17942744]
61. Mooers BH, Logue JS, Berglund JA. The structural basis of myotonic dystrophy from the crystal structure of CUG repeats. *Proc. Natl. Acad. Sci. U S A.* 2005; 102:16626–16631. [PubMed: 16269545]

62. Sobczak K, de Mezer M, Michlewski G, Krol J, Krzyzosiak WJ. RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.* 2003; 31:5469–5482. [PubMed: 14500809]
63. Guan L, Disney MD. Small molecule-mediated cleavage of RNA in living cells. *Angew. Chem. Int. Ed. Engl.* 2013; 52:1462–1465. [PubMed: 23280953]
64. Guan L, Disney MD. Covalent small molecule-RNA complex formation in living cells enables cellular profiling of small molecule-RNA interactions. *Angew. Chem. Int. Ed. Engl.* 2013 in press.
65. Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, Moxley RT, Swanson MS, Thornton CA. Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.* 2006; 15:2087–2097. [PubMed: 16717059]

Figure 1.

Schematic for the disease pathology of DM1. An expansion of r(CUG) repeats (r(CUG)^{exp}) located in the 3' UTR of the *DMPK* mRNA folds into a hairpin structure that binds and sequesters muscleblind-like 1 protein (MBNL1), a pre-mRNA splicing regulator. This decrease in functional MBNL1 results in dysregulation of alternative splicing of pre-mRNAs controlled by MBNL1. Other DM1-associated defects include decreased translation of the *DMPK* pre-mRNA and the formation of nuclear foci. Modularly assembled small molecules that target r(CUG)^{exp} have the potential of improving defects by displacing MBNL1.

Figure 2.

Structures of first and second generation modularly assembled compounds that bind $r(\text{CUG})^{\text{exp}}$. First generation compound **2H-4** contains a peptoid assembly scaffold. Classes of second generation compounds consist of polyamines, α -peptides, β -peptides, and peptide tertiary amides (PTAs).

Figure 3.

IC₅₀'s of second generation modularly assembled compounds for disrupting the r(CUG)₁₂-MBNL1 complex. Potencies were determined using a TR-FRET assay.^{15, 47} The most potent compounds from each class are: **2H-SPM**, **2H-3G** $\approx$ **2H-4G**, and **2H-3NPr**, **2H-3NMe**, and **2H-K4NPr** ($\approx$ **2H-K1NPr**).

Figure 4.

Effect of the modular assembly scaffold on improving DM1-associated pre-mRNA splicing and translational defects. (A) Representation of the pre-mRNA splicing pattern observed for the *cTNT* mini-gene in the presence and absence of the DM1 mini-gene. Cells were treated post-transfection with 10 μ M compound in growth medium. Alternative splicing was assessed by RT-PCR. (B) Quantification of RT-PCR analysis. Second generation compounds were compared to the parent compound, **2H-4**. The most active compound in this assay is **2H-K4NPr**, which has a PTA assembly scaffold ($p = 0.03$; indicated by “*”). (C) Schematic of the cell model used to study translational defects. A stably transfected C2C12 cell line expresses firefly luciferase mRNA with r(CUG)₈₀₀ in the 3' UTR. r(CUG)₈₀₀ causes the transcript to be mostly retained in the nucleus and thus not efficiently translated. If a small molecule binds to r(CUG)₈₀₀ and displaces or inhibits MBNL1 binding, then the transcript is more efficiently exported from the nucleus and translated in the cytoplasm. (D) Quantification of the effect of second generation compounds on the DM1 translational defect as measured by luciferase activity. Compounds were tested at 10 μ M and compared to the parent compound, **2H-4**. Enhancement in luciferase activity is relative to untreated cells. (No change in luciferase activity is equal to 0%.) Second generation compounds were less effective than **2H-4** at improving translational defects. All second generation compounds have similar activities.

Figure 5.

Bioactivity of modularly assembled third generation compounds. Derivatives of the most bioactive second generation compound, **2H-K4NPr**, were designed to investigate the effect of the substituents on the tertiary amide. (A) Structures of third generation compounds. Derivatives of **2H-K4NPr** were prepared consisting of peptide, PTA and peptoid scaffolds. (B) IC_{50} 's of third generation modularly assembled compounds for disrupting the r(CUG)₁₂-MBNL1 complex. **2H-K4NMe**, which has an *N*-methyl PTA scaffold, is the most potent. "*" indicates $p < 0.05$ as determined by a two-tailed Student *t*-test. (C) Quantification of the improvement of DM1-associated alternative pre-mRNA splicing defects by third generation compounds. **2H-K4NPr** improves dysregulation of the *cTNT* pre-mRNA splicing defect to the greatest extent. (D) Quantification of the improvement of DM1-associated translational defects by third generation compounds. Compounds demonstrated comparable activity to the parent compound, **2H-4**. (E) Cellular permeability and toxicity of third generation modularly

assembled compounds. “+” indicates Hoechst 33258, which was employed as a positive control for gating.

Figure 6.

Pull down experiments to identify the cellular RNA targets of designer small molecule **2H-K4NMe-Biotin**. Top, structure of the compound that was anchored onto streptavidin functionalized agarose beads to provide an affinity matrix. Bottom, left, SYBR gold stained denaturing agarose gel. Bottom right, Northern blot of the SYBR gold stained gel with a probe for r(CUG)_{exp}. Lane 1, 1 Kb ladder; the top four bands (from top to bottom) are: 3 Kb, 2 Kb, 1.5 Kb, and 1 Kb; Lane 2, total RNA harvested from cells; Lane 3, empty; Lane 4, pull down by using 5'-Biotin d(CAG)₁₀; Lane 5, **K4NMe-Biotin** bound fraction (control

small molecule that should not bind r(CUG)^{exp}; Lane 6, **2H-K4NMe-Biotin** bound fraction; Lane 7, final wash from resin prior to elution of the bound RNA from lane 6.

Figure 7.

2H-K4NMe improves DM1-associated alternative pre-mRNA splicing defects in a DM1 mouse model. The DM1 mouse model expresses the human skeletal actin (HSA) transgene containing 250 CTG repeats (HSA^{LR}; where LR indicates long repeats). Wild type (WT) refers to FVB mice. (A) Top, schematic of *Clcn1* alternative splicing in WT and DM1 mice. Bottom, analysis of *Clcn1* alternative splicing by RT-PCR when mice are treated for each day for 7 days with 100 mg/kg of **2H-K4NMe**. (B) Top, schematic of *Serca1* (*Atp2a1*) alternative splicing in WT and DM1 mice. Bottom, analysis of *Serca1* alternative splicing by RT-PCR when mice are treated with 100 mg/kg/d of **2H-K4NMe**. “***” indicates p<0.01 as determined by a two-tailed Student t-test.

Table 1

Characterization of the ligands binding to various RNAs as assessed by SPR.^a

RNA	2H-K4NMe						Selectivity for r(CUG) ₁₂
	k _{off} (1) (s ⁻¹)	k _{off} (2) (s ⁻¹)	k _{on} (1) (M ⁻¹ s ⁻¹)	k _{on} (2) (M ⁻¹ s ⁻¹)	K _{obs} (1) (M)		
r(CUG) ₁₂ ^b	3.43×10 ⁻³	5.00×10 ⁻²	2.61×10 ⁵	1.87×10 ³	1.31×10 ⁻⁸		-
r(CAG) ₁₂ ^b	6.95×10 ⁻²	8.47×10 ⁻²	8.38×10 ⁴	1.63×10 ³	8.29×10 ⁻⁷		63
r(AU) ₁₂ ^b	1.65×10 ⁻²	3.59×10 ⁻¹	4.13×10 ⁴	1.98×10 ⁴	4.00×10 ⁻⁷		30
r(GC) ₈ ^b	8.22×10 ⁻²	1.05×10 ⁻¹	5.86×10 ⁴	2.48×10 ⁴	1.40×10 ⁻⁶		107
H monomer							
r(CUG) ₁₂ ^c	3.69×10 ⁻²	-	1.76×10 ⁴	-	2.10×10 ⁻⁶		-
MBNL1							
r(CUG) ₁₂ ^c	2.03×10 ⁻³	-	3.28×10 ³	-	6.19×10 ⁻⁷		-

^aThe standard deviations for each of the measurements is ±10% and the secondary structures of the RNAs are provided in Figure S-5.

^bData were fit best to a two component binding isotherm.

^cData were fit best to a one component binding isotherm.