Published in final edited form as:

ACS Chem Biol. 2012 March 16; 7(3): 496–505. doi:10.1021/cb200413a.

Chemical correction of pre-mRNA splicing defects associated with sequestration of muscleblind-like 1 protein by expanded r(CAG) transcripts

Amit Kumar[†], Raman Parkesh[†], Lukasz J. Sznajder[‡], Jessica Childs-Disney[†], Krzysztof Sobczak^{‡,*}, and Matthew D. Disney^{†,*}

[†]Department of Chemistry, The Scripps Research Institute, Scripps Florida, 130 Scripps Way 3A1, Jupiter, FL 33458 [‡]Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznan, Poland

Abstract

Recently, it was reported that expanded r(CAG) triplet repeats (r(CAG)^{exp}) associated with untreatable neurological diseases cause pre-mRNA mis-splicing likely due to sequestration of muscleblind-like 1 (MBNL1) splicing factor. Bioactive small molecules that bind the 5'CAG/3'GAC motif found in r(CAG)^{exp} hairpin structure were identified by using RNA binding studies and virtual screening/chemical similarity searching. Specifically, a benzylguanidine-containing small molecule was found to improve pre-mRNA alternative splicing of MBNL1-sensitive exons in cells expressing the toxic r(CAG)^{exp}. The compound was identified by first studying the binding of RNA 1×1 nucleotide internal loops to small molecules known to have affinity for nucleic acids. Those studies identified 4',6-diamidino-2-phenylindole (DAPI) as a specific binder to RNAs with the 5'CAG/3'GAC motif. DAPI was then used as a query molecule in a shape- and chemistry alignment-based virtual screen to identify compounds with improved properties, which identified 4-guanidinophenyl 4-guanidinobenzoate as small molecule capable of improving pre-mRNA splicing defects associated with the r(CAG)^{exp}-MBNL1 complex. This compound may facilitate the development of therapeutics to treat diseases caused by r(CAG)^{exp} and could serve as a useful chemical tool to dissect the mechanisms of r(CAG)^{exp} toxicity. The approach used in these studies, defining the small RNA motifs that bind known nucleic acid binders and then using virtual screening to optimize them for bioactivity, may be generally applicable for designing small molecules that target other RNAs in human genomic sequence.

INTRODUCTION

RNA forms a variety of structures that impart diverse functions.(1, 2) For example, RNAs catalyzes chemical reactions (3), riboswitches and microRNAs regulate complex networks of gene expression (4, 5), and RNA encodes and synthesizes protein.(6) Thus, RNA is an important target for chemical genetics probes or therapeutics.(7, 8) One advantage of using RNA as a drug target is that structural information can be accurately deduced from sequence. For example, secondary structural information, which includes the motifs that comprise an RNA, can be obtained by free energy minimization (9, 10) or phylogenetic

* authors to whom correspondence should be addressed. Krzysztof Sobczak, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznan, Poland, ksobczak@wp.pl, Matthew D. Disney, Department of Chemistry, The Scripps Research Institute, Scripps Florida, 130 Scripps Way 3A1, Jupiter, FL 33458, disney@scripps.edu, Phone: 561-228-2203, Fax: 561-228-2147.

Supporting Information Available: This material is free via the Internet at <http://pubs.acs.org>.

comparison.(11) Once the ensemble of motifs present in an RNA target are known, modular assembly strategies can be applied to target multiple motifs within a target RNA simultaneously to improve affinity and specificity.(12–14) Most RNAs are underutilized as targets, however, because of a general lack of information that is available on small molecule-RNA interactions.

In order to more effectively design compounds that target RNA, more information is needed on the RNA motifs that can be targeted by small molecules. One class of small molecules that binds to RNA motifs is DNA-binding agents.(15, 16) Of particular interest is the binding of these DNA-binding agents to 1×1 nucleotide RNA internal loops because they are present in expanded triplet repeat RNAs that contribute to neurological and neuromuscular disorders such as myotonic dystrophy type 1 (DM1), Huntington's disease (HD), spinocerebellar ataxia type 3 (SCA3), and fragile X-associated tremor-ataxia syndrome (FXTAS). All CNGs form similar hairpin architectures; the only difference is the identity of the internal loop nucleotides (U/U, A/A or G/G internal loops).

Thus, competition dialysis and spectroscopic methods were employed to determine the selectivity of DNA-binding agents to 1×1 nucleotide RNA internal loops (Figure 1) in an effort to identify small molecule leads. Such leads could be further optimized for bioactivity. From these investigations, it was determined that 4',6-diamidino-2-phenylindole (DAPI) binds specifically to RNAs containing the 5'CAG/3'GAC motif (where the underlined A's indicate the internal loop) found in expanded r(CAG)^{exp} repeats (r(CAG)^{exp}) that are associated with HD and SCA3. A virtual screen was completed in order to identify other compounds that are similar to DAPI in shape and chemistry alignment, or the positioning of functional groups, rings, hydrophobic groups, charge, etc. The most potent small molecule, 4-guanidinophenyl 4-guanidinobenzoate, corrects pre-mRNA splicing defects due to sequestration of the MBNL1-r(CAG)^{exp} complex.(17) Such studies provide useful leads for the development of chemical probes to study r(CAG)^{exp} toxicity as well as potential therapeutics that target r(CAG)^{exp}.

RESULTS & DISCUSSION

Studying the binding of small molecules to 1×1 nucleotide RNA internal loops

Previous studies have shown that cell-permeable DNA-binding agents interact with RNAs containing 1×1 nucleotide internal loops.(15, 16) Interestingly, 1×1 nucleotide internal loops are found multiple times in transcripts with expanded triplet repeats (18) and are associated with disorders such as DM1, HD, SCA3 and FXTAS. We therefore used competition dialysis (19) to evaluate the binding selectivity of a variety of cell permeable DNA-binding agents for RNA 1×1 nucleotide internal loops and other nucleic acids. The small molecules used in competition dialysis studies include mitoxantrone (MTX), acridine orange (AO), DAPI, proflavin, daunomycin (DNM), quinacrine (QCN), 9-amino-6-chloro-2-methoxy acridine (ACMA), and netropsin (NTSP) (Figure 1). AO, proflavin, DNM, MTX, QCN, and ACMA bind DNA by intercalation.(20–23) MTX binds a hairpin structure in a *tau* (*MAPT*) pre-mRNA (24), which has been implicated in Alzheimer's and other neurodegenerative diseases. A structure of MTX bound to *tau* hairpin has been reported.(25) Derivatives of AO and proflavin have been used as components of small molecules that bind RNAs as these structures intercalate into stretches of GC pairs.(26) DNM is related in structure to doxorubicin, which has been shown to bind to the HIV slippery sequence and inhibit frameshifting.(27) DAPI and netropsin either bind in the minor groove or have a mixed binding mode that includes minor groove binding and intercalation.(28–30) DAPI has been shown to bind to and inhibit group I intron self-splicing.(31, 32)

The nucleic acids used in these studies include each possible 1×1 nucleotide internal loop closed by two GC pairs (Figure 2).(33–35) The internal loops were embedded in a cassette containing a GNRA-type hairpin that ensures proper folding.(10, 36) The 5'CUG/3'GUC, 5'CAG/3'GAC, and 5'CGG/3'GGC loops are present in the RNAs that cause DM1, HD and SCA3, and FXTAS, respectively.(37) A fully paired RNA was used as a control for ligands that bind to canonically paired RNAs helices.(38)

The results of competition dialysis studies are shown in Figure 3. Two ligands, proflavin and AO, show little preference for binding to a fully paired RNA or any of the RNAs with 1×1 nucleotide internal loops. This is perhaps not surprising since these ligands would most likely interact with RNA by intercalating between the consecutive GC pairs (22, 39) common to all RNAs studied (Figure 2).

MTX, NTSP, and ACMA show modest binding preferences. MTX has a slight preference for 5'CCG/3'GAC while it discriminates against 5'CUG/3'GCC. In a previously reported NMR spectroscopy study of MTX bound to a *tau* pre-mRNA hairpin, MTX binds at the base of the stem between two GC pairs that flank a bulged A.(25) There is electrostatic complementarity between the MTX side chains and the major groove, which could also occur in the RNA with the 5'CCG/3'GAC motif. NTSP has a slight preference for 5'CCG/3'GUC over 5'CAG/3'GAC, 5'CGG/3'GAC, 5'CGG/3'GCC, and 5'CCG/3'GAC (which are greater than the remaining loops studied). It discriminates against 5'CAG/3'GGC. ACMA binds preferentially to the fully paired RNA and the RNA displaying 5'CCG/3'GCC slightly over 5'CCG/3'GUC

Clear binding preferences are observed for DNM, QNC, and DAPI. DNM prefers to bind 5'CCG/3'GAC with a stoichiometry of ~2:1 while QNC prefers to bind 5'CAG/3'GCC with an ~2:1 stoichiometry (based on the amount of bound ligand relative to the amount of RNA target). DAPI preferentially binds to 5'CAG/3'GAC and 5'CAG/3'GCC in stoichiometric amounts. DAPI has been previously shown to bind the minor groove of AT steps in DNA and to intercalate at AU steps in RNA.(28) There is also evidence from NMR spectroscopy studies that DAPI intercalates between CG steps in DNA (28) and binds to the minor groove in DNAs containing T/T mismatches.(29)

Most interestingly, DAPI displayed selectivity for the 5'CAG/3'GAC motif. This motif periodically repeats in transcripts with r(CAG)^{exp} (18) that are associated with HD, SCA3, and Kennedy's disease. Like expanded r(CUG) repeats (r(CUG)^{exp}), r(CAG)^{exp} have recently been shown to bind and sequester muscleblind-like 1 protein leading to pre-mRNA splicing defects.(17) Therefore, the affinity of DAPI for RNAs containing various 1×1 nucleotide internal loops and a fully paired RNA was determined using fluorescence anisotropy (Table 1). In good agreement with the results of competition dialysis, DAPI binds most tightly to the RNA containing the 5'CAG/3'GAC motif ($0.22 \pm 0.06 \mu\text{M}$). It binds approximately 2.5-fold more weakly to RNAs containing 5'CUG/3'GUC ($0.55 \pm 0.2 \mu\text{M}$) and 5'CGG/3'GGC ($0.50 \pm 0.1 \mu\text{M}$) and ~5-fold more weakly to an RNA with the 5'CCG/3'GCC motif ($0.99 \pm 0.2 \mu\text{M}$). The K_D for binding to a fully paired RNA is $> 3 \mu\text{M}$.

Identifying DAPI-like small molecules that specifically bind 1×1 nucleotide A/A internal loops using virtual screening

Since DAPI shows modest selectivity and reasonable binding affinity for an RNA containing a single copy of the 5'CAG/3'GAC motif that is associated with HD and SCA3, we aimed to identify chemically similar compounds with improved affinities and selectivities. Thus, virtual screening was used to identify ligands that have similar shape and chemistry alignment as DAPI. Chemistry alignment (40) accounts for the positioning of

hydrogen bond donors, hydrogen bond acceptors, anions, cations, hydrophobic groups and rings.

A 3D computational shape comparison between DAPI and a 250,000-member compound library available from the National Cancer Institute (NCI) was completed using the rapid overlay of chemical structures (ROCS) shape algorithm.(41) ROCS treats atoms as Gaussian functions and thus the overlap between two (or more) atoms are all Gaussian functions. This algorithm implemented in ROCS speeds up the 3-dimensional shape-based calculations, as it calculates the maximal intersection of the volumes of two molecules. The degree of shape similarity is quantified by the shape Tanimoto coefficient (42) that ranges in values from 0 (no similarity) to 1 (complete shape similarity). Chemistry alignment is specified by the color force field (Implicit Mills Dean (40)) and is reported as the color score that has a range of 0 (no overlap) to 1 (complete chemistry alignment overlap). The combo score is the sum of the shape Tanimoto coefficient and the color score. A series of the top hits from the virtual screen were selected for testing. Nine compounds were selected for further investigation based on their chemical diversity as determined by visual inspection (Table 3).

Determining the Selectivities of DAPI-like Compounds

Two of the nine compounds identified from the virtual screen have a single absorption band near 260 nm, making them incompatible with competition dialysis experiments. Therefore, the binding and selectivity of the remaining compounds were further investigated (Figure 4). Compounds D1, D2, and D3 show little binding preference as they bind a fully paired RNA and the RNAs with different 1×1 nucleotide internal loops similarly. D4 has the following binding preferences: 5'CCG/3'GAC > 5'CAG/3'GAC >> other RNAs; it binds approximately stoichiometrically to 5'CCG/3'GAC and 5'CAG/3'GAC. D7 is the most selective for 5'CAG/3'GAC and binds stoichiometrically. D9 shows little preference for most RNAs and discriminates against 5'CAG/3'GAC.

Optical Melting Experiments

The binding of DAPI and the compounds identified from the virtual screen to an RNA with the 5'CAG/3'GAC motif was also studied by optical melting (Table 2). In all cases, simple monophasic melting curves were obtained. The presence of DAPI and D2 increase the melting temperature to the greatest extent compared to the other compounds when one equivalent of ligand is added (6 °C and 5 °C, respectively). The T_m increases more modestly for compounds D1, D4, D6, and D7 (2–3 °C). Greater increases in the T_m are observed when two equivalents of the ligand are added: 4–5 °C for D1 and D4, 8–10 °C for DAPI, D2, D6, and D7. The increases in T_m indicate that the small molecules bind to the RNA and increase its thermodynamic stability.

Inhibition of the r(CAG)^{exp}-MBNL1 Complex by Small Molecules

A recent report showed that r(CAG)^{exp} mimics r(CUG)^{exp} in the misregulation of alternative splicing.(17) r(CUG)^{exp} cause myotonic dystrophy type 1 (DM1) via an RNA gain-of-function mechanism.(37) Specifically, the toxic r(CUG)^{exp} present in the 3' untranslated region of the dystrophin myotonia protein kinase (DMPK) mRNA (37, 43, 44) bind to and inactivates MBNL1. Sequestration of MBNL1 by r(CUG)^{exp} causes mis-splicing of various pre-mRNAs in muscles including cardiac troponin T (cTNT), the insulin receptor (INSR), and the muscle-specific chloride ion channel (CLCN1).(45) Each of these splicing defects explains physiological phenotypes of DM1 such as insulin insensitivity and myotonia.

Several other studies have shown that r(CAG)^{exp} cause toxicity by the same gain-of-function mechanism.(46–49) In particular: (i) r(CAG)^{exp} interact with MBNL1 and form nuclear foci in cells and animal models (17, 50–52); (ii) recombinant MBNL1 binds

r(CAG)^{exp} and r(CUG)^{exp} with similar affinities (14, 17); and, (iii) r(CAG)^{exp} has been found to be toxic in *Drosophila* models when an untranslated r(CAG)^{exp} is expressed.(53)

Therefore, we determined if DAPI and the DAPI-like compounds could disrupt the r(CAG)^{exp}-MBNL1 complex using a time-resolved FRET (TR-FRET) assay. Briefly, 5'-biotinylated r(CAG)₁₂ and MBNL1-His₆ are pre-equilibrated, and incubated with the compound of interest. Next, streptavidin-XL665 (binds to the biotinylated RNA oligonucleotide) and Tb-Anti-His₆ (binds to MBNL1) are added. If the compound does not disrupt the r(CAG)₁₂-MBNL1 interaction, then XL665 and Tb are within close enough proximity to form a FRET pair and TR-FRET can be measured. Conversely, TR-FRET is not observed if the compound inhibits formation of the r(CAG)₁₂-MBNL1 complex. It should be noted that the RNA-protein complex is pre-formed prior to addition of the compound of interest. Therefore, in order to be scored as an active compound, the small molecule must displace MBNL1 from r(CAG) repeats.

The IC₅₀'s measured for the DAPI-like compounds range from >1000 to 3 μM (Table 3). The least potent compounds have the least amount of charge and hydrophobic surface area with which to interact with RNA while the most potent compounds contain the most charge and the largest hydrophobic surface area. The most potent compound, D6, is at least 20-fold more potent than the other compounds studied and contains two benzylguanidine side chains linked by an ester (Table 3). D6 binds to an RNA with a single copy of the 5'CAG/3'GAC motif with a dissociation constant of 60 nM as determined by fluorescence anisotropy (Table 1 and Supporting Information).

Compounds that Bind 5'CAG/3'GAC and Inhibit the r(CAG)-MBNL1 Complex Improve pre-mRNA Splicing Defects In Cell-Based Assays

A recent report showed that r(CAG)^{exp} mimics the misregulation of pre-mRNA alternative splicing that is observed for r(CUG)^{exp} in different cellular model systems.(17) We determined that the pre-mRNAs of the *NCOR2* and *ATP2A1* genes are mis-spliced in both DM1- and HD-derived fibroblast cells expressing different lengths of r(CUG)^{exp} and r(CAG)^{exp}, respectively (Supplemental Figure S-11A and S-12). Moreover, the alternative splicing of all the transcripts is MBNL1-sensitive in HEK293, as was determined by overexpression of MBNL1 (Supplemental Figure S-11C & D). Therefore these splicing markers were used to study the bioactivity of DAPI and DAPI-like compounds in HD patient-derived (r(CAG)69) and control fibroblasts. We also used an artificial model system of HD in which a neuroblastoma cell line stably expresses r(CAG)^{exp} (r(CAG)74). (17) In this cell line, the splicing of *NCOR2*, *ATP2A1* and *MAPT* pre-mRNAs is abnormally shifted towards a DM-like pattern (Figure 5C and D and Supplemental Figure S-12).

As an initial screen, several DAPI-like compounds were incubated at the same concentration with both cellular models (100 μM), and their effects on cell morphology (an estimate of toxicity) and on *ATP2A1* splicing (an estimate of bioactivity) were determined (Supplemental Figures S-9 and S-10). D6 showed the least amount of toxicity as assessed by cell morphology (Figure 5A) and most significantly improved *ATP2A1* mis-splicing (Figure 5B). In contrast, DAPI showed the strongest morphological changes (Supporting Information) and exacerbated splicing abnormalities of *ATP2A1* (Supplemental Figures S-9 and S-10). Interestingly, the structure of D6 is somewhat similar to pentamidine, which has been shown to improve DM1-associated pre-mRNA splicing defects in cellular and animal models of myotonic dystrophy.(54)

Next, we determined if D6 could improve other pre-mRNA splicing defects in HD fibroblasts (express r(CAG)69) and the neuroblastoma model system (expresses r(CAG)74). Dose responses were completed for improvement of the mis-splicing of *NCOR2* and

ATP2A1 in HD fibroblasts and of *NCOR2*, *ATP2A1*, and *MAPT* in the neuroblastoma cell line (Figure 5B, C, and D). Statistically significant improvement of *NCOR2* mis-splicing was observed when HD fibroblasts were treated with 50–300 μ M D6. Likewise, improvement in *ATP2A1* mis-splicing was observed when cells were treated with 100–300 μ M D6. In good agreement with the results observed in the HD fibroblasts, statistically significant improvement in the mis-splicing of *NCOR2* and *ATP2A1* was also observed in the neuroblastoma cell line (Figure 5C). Gratifyingly, D6 was found to be the most potent inhibitor of the MBNL1-r(CAG)^{exp} complex, however, *in vitro* potency may not correlate with bioactivity due to factors that may include compound toxicity and cellular uptake.

Unexpectedly, D6 slightly worsens the splicing of the *MAPT* pre-mRNA toward the DM-like phenotype in the neuroblastoma cell line express 74 r(CAG) repeats (Figure 5D). A more pronounced effect is observed when a normal neuroblastoma cell line is treated with D6 (Figure 5D). This may be due to D6 binding the pre-mRNA and affecting the splicing outcome. Importantly, D6 does not affect the splicing of pre-mRNAs of other tested transcripts in control cell lines that do not express toxic r(CAG)^{exp} (Supplemental Figure S-12). Taken together, these results suggest that D6 inhibits the r(CAG)^{exp}-MBNL1 complex *in vivo* with at minimum modest specificity, increasing the concentration of free MBNL1 and improving splicing defects.

As is observed in DM1-affected cells,(55–58) the presence of r(CAG)^{exp} leads to formation of nuclear RNA foci (Figure 5E).(17) Therefore, we also determined if D6 could disrupt nuclear RNA foci formation in r(CAG)^{exp}-expressing cells. To measure this phenotype, fluorescence *in situ* hybridization (FISH) was employed using a fluorescently labeled hybridization probe that binds to r(CAG)^{exp}.(17) As shown in Figure 5E, D6 does not significantly affect the number or the intensity of nuclear r(CAG)^{exp} foci even though D6 improves splicing defects. Other studies, however, have described similar results in which nuclear foci and normal splicing patterns are observed in cell lines expressing r(CUG)^{exp}.(59, 60) These results suggest that D6 displaces enough MBNL1 to improve splicing defects but not enough MBNL1 or other proteins (50, 55, 57, 58) to disrupt nuclear RNA foci.

SUMMARY AND OUTLOOK

We identified a bioactive small molecule that improves splicing defects associated with expanded r(CAG) repeats. The compound was identified by investigating the RNA-binding preferences for a series of small molecules with known affinity for nucleic acids. A lead molecule that binds 5'CAG/3'GAC, DAPI, was optimized by virtually screening a database of small molecules for compounds that are similar in shape and chemistry alignment. Specifically, compound D6 (Table 3) inhibits the r(CAG)-MBNL1 complex formation *in vitro* and improves a subset of pre-mRNA splicing defects that are associated with formation of the r(CAG)^{exp}-MBNL1 complex *in vivo*. This approach could yield small molecules that specifically target other RNAs in general and other repeating transcripts in particular. The potency and specificity of the compounds could be improved using a modular assembly strategy.(13, 14, 16, 61)

Although many studies have shown that r(CAG)^{exp} cause splicing defects and toxicity due to an RNA gain of function (17, 50, 53), other studies, however, have disregarded the pathogenic role of r(CAG)^{exp} transcripts on nuclear foci formation (62) or mis-splicing.(50) Thus, the identification of a small molecule that targets r(CAG)^{exp}, freeing MBNL1 and improving splicing defects would provide further evidence of a dominant RNA gain-of-function mechanism. This compound could also provide a chemical probe to better understand the mechanisms associated with r(CAG)^{exp} toxicity.

METHODS

RNA Preparation and Purification

The RNA used in TR-FRET assays (5'-biotinylated-(CAG)₁₂) was purchased from Dharmacon. The ACE protecting groups were cleaved using Dharmacon's deprotection buffer. The sample was lyophilized, dissolved in milliQ water and desalted using a PD-10 gel filtration column (GE Healthcare). RNAs used in competition dialysis and fluorescence anisotropy experiments were prepared via run off-transcription using synthetic DNA templates (Integrated DNA Technologies, Inc.) and a Stratagene RNAMaxx transcription kit. They were purified by denaturing polyacrylamide gel electrophoresis as previously described (34) and dissolved in 1× Competition Dialysis Buffer (8 mM NaH₂PO₄, 185 mM NaCl and 1 mM EDTA; pH 7.2).

Oligonucleotide concentrations were determined by absorbance at 260 nm (90 °C for the RNA used in TR-FRET assays; room temperature for RNAs used in competition dialysis) using a Beckman Coulter DU800 UV-Vis spectrophotometer equipped with a Peltier temperature controller unit. Oligonucleotide extinction coefficients were determined using HyTher version 1.0 (Nicolas Peyret and John SantaLucia Jr., Wayne State University, Detroit, MI).^(63, 64) The parameters used by HyTher were calculated using information on the extinction coefficients of nearest neighbors in RNA.⁽⁶⁵⁾

Ligands

All ligands viz. Mitoxantrone (MTX), Acridine Orange (AO), Proflavin, 4',6-diamidino-2-phenylindole (DAPI), Netropsin (NTSP), Daunomycin (DNM), Quinacrine (QNC) and 9-Amino 6-Chloro 2-Methoxy Acridine (ACMA) were dissolved in milliQ water. The concentrations of the ligands were determined by UV absorbance spectroscopy using their respective extinction coefficients at their absorbance maxima (608 nm, 430 nm, 296 nm, 353 nm, 444 nm, 480 nm, 424 nm, and 430 nm for MTX, AO, NTSP, DAPI, proflavin, DNM, QNC and ACMA, respectively). DAPI-like compounds were procured from NCI (Table 3) and were purified by high performance liquid chromatography before use. The spectrometric properties for the compounds are provided in the Supporting Information (SI Table 1).

Competition Dialysis

Pierce Slide-A-Lyzer MINI dialyzer units (Pierce Biotechnology, Inc.) were dialyzed against water for 24 h in order to ensure that the units did not leak. Competition dialysis was completed as previously described by Chaires.⁽²²⁾ Briefly, RNAs were folded by heating at 60 °C for 5 min followed by slowly cooling to room temperature on the bench top. Dialysis units containing 0.1 mL of 1.6 μM RNA were placed in 200 mL of 1 μM ligand in 1× Competition Dialysis Buffer. The samples were allowed to equilibrate with the dialysate by stirring at 200 rpm for 24 h at room temperature (20–22 °C). Previous studies have shown that this is sufficient time for the sample to reach equilibrium.⁽¹⁹⁾

At the end of the equilibration period, 67.5 μL of each sample was carefully removed from the dialyzer unit and transferred to microcentrifuge tube. To each sample, 7.5 μL of 10% (w/v) sodium dodecyl sulfate (SDS) was added to give a final concentration of 1% (w/v) SDS, which is sufficient to dissociate the ligand. The SDS step was completed to ensure accurate determination of the concentration of bound ligand as the spectroscopic properties of the bound ligand could be different from that of unbound. The total ligand concentration (C_l) within each dialysis unit was determined spectrophotometrically using an appropriate absorbance wavelength and extinction coefficient for each compound. Appropriate corrections were made for the small dilution resulting from the addition of SDS.

The free ligand concentration (C_f) was determined from an aliquot of the dialysate solution, which did not vary appreciable from the initial 1 μ M. The bound ligand concentration (C_b) was then determined using Equation 1:

$$C_b = C_t - C_f \quad (\text{eq. 1})$$

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

Affinity Measurements Using Fluorescence Anisotropy

Fluorescence anisotropy measurements were performed on a Molecular Devices SpectraMax M5 microplate reader using an excitation wavelength of 368 nm and emission wavelength of 500 nm. Measurements were performed in a buffer solution containing 20 mM HEPES, pH 7.5, 5 mM KCl, 150 mM NaCl, and 1 mM MgCl_2 (1 $\times$ Affinity Buffer). The RNA of interest was folded by heating the RNA at 60 $^{\circ}\text{C}$ for 5 min in 1 $\times$ Affinity Buffer followed by slowly cooling to room temperature on the bench top.

For direct measurements, the RNA was titrated into 50 nM compound of interest in 1 $\times$ Affinity Buffer with a 5 min equilibration time between additions. Anisotropy was measured as a function of RNA concentration and fit to Equation 2:

$$A = A_0 + 0.5\Delta \in \left\{ ([L]_0 + [RNA]_0) + K_t - (([L]_0 + [RNA]_0 + K_t)^2 - 4[L]_0[RNA]_0)^{0.5} \right\} \quad (\text{eq. 2})$$

where A is the observed anisotropy, A_0 is the fluorescence anisotropy in the absence of RNA, $\Delta \in$ is the difference between the fluorescence intensity in the absence of RNA and the fluorescence anisotropy in the presence of infinite RNA concentration, $[L]_0$ is the concentration of ligand, $[RNA]_0$ is the concentration of the RNA of interest, and K_t is the dissociation constant.

For competition measurements, the RNA of interest was added to 50 nM DAPI until saturation was reached. Then, the compound of interest was titrated into the solution, and the anisotropy was measured after a 5 min equilibration period. The fraction of ligand bound as a function of compound was fit to equation 3:

$$\theta = \frac{1}{2[L]_0} \left[K_t + \frac{K_t}{K_d} [C_t] \right] + [RNA] + [L] - \sqrt{\left(K_t + \frac{K_t}{K_d} [C_t] + [RNA] + [L] \right)^2 - 4[L][RNA]} \quad (\text{eq. 3})$$

where θ is the fraction of D2 bound, K_t is the dissociation constant for D2 as determined from direct measurements, K_d is the dissociation constant for the competing DAPI-like compound, $[C_t]$ is the concentration of competing DAPI-like compound, $[L]$ is the concentration of D2, $[RNA]$ is the concentration of RNA required to reach saturation.

Optical Melting Experiments

Optical melting curves were obtained using a Beckman Coulter DU 800 spectrometer. The temperature of the cell holder was regulated by using a peltier heater. The RNAs were first denatured in 1 $\times$ Affinity Buffer by heating to 90 $^{\circ}\text{C}$ and cooling to 15 $^{\circ}\text{C}$. The absorbance at 260 nm as a function of temperature was recorded using a rate of 1 $^{\circ}\text{C}/\text{min}$. The melting temperature, T_m , was calculated using Equation 4:

$$T_m = \frac{\Delta H^{\circ} \times 1000}{\Delta S^{\circ}} - 273.15 \quad (\text{eq. 4})$$

where ΔH° and ΔS° are the changes in enthalpy and entropy, respectively.

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

The TR-FRET assay used to identify lead inhibitors of the r(CAG)₁₂-MBNL1 complex is based on PubChem BioAssay AID 2675. Briefly, 5'-biotinylated r(CAG)₁₂ was folded in 1× Folding Buffer (20 mM HEPES, pH 7.5, 110 mM KCl, and 10 mM NaCl) by heating at 60 °C followed by slowly cooling to room temperature on the bench top. The buffer was adjusted to 1× Assay Buffer (20 mM HEPES, pH 7.5, 110 mM KCl, 10 mM NaCl, 2 mM MgCl₂, 2 mM CaCl₂, 5 mM DTT, 0.1% BSA, and 0.5% Tween-20) and MBNL1-His₆ was added. The final concentrations of RNA and MBNL1 were 80 nM and 60 nM, respectively. The sample was allowed to equilibrate at room temperature for 5 min, and then the compound of interest was added. After 15 min, streptavidin-XL665 and anti-His₆-Tb antibody were added to final concentrations of 40 nM and 0.44 ng/μL, respectively, in a total of 10 μL. The samples were incubated for 1 h at room temperature and then transferred to a well of a black 96-well plate. Time-resolved fluorescence was measured on a Molecular Devices SpectraMax M5 plate reader. Fluorescence was first measured using an excitation wavelength of 345 nm and an emission wavelength of 545 nm (fluorescence due to Tb). TR-FRET was then measured by using an excitation wavelength of 345 nm, an emission wavelength of 665 nm, a 200 μs evolution time, and a 1.5 ms integration time. The ratio of fluorescence intensity of 545 nm and 665 nm as compared to the ratios in the absence of ligand and in the absence of RNA were used to determine IC₅₀'s.

Virtual Screening of NCI Database for DAPI-like Compounds

Virtual screening was performed using a Mac OSX 10.6 operating system with a 2.8GHz Quad-Core Intel Xeon W3530 "Nehalem" processor and 6 GB of memory. As a starting point for the virtual screen, Chem3D (Cambridge software) was used to energy-minimize (MMFF94force field) a 3-dimensional conformer of DAPI. Omega 2.3.2 (v 2.02) was then used to generate 100 conformers of the query molecule using an energy window of 10 kcal and an RMSD of 0.8 Å. Up to 100 conformers of each molecule in the NCI database were also generated. 3D-shape comparisons (shape Tanimoto coefficient) were completed using rapid overlay of chemical structure (ROCS, version 3.0.0) while the chemistry alignment overlap (color score) was completed by using the color force field (Implicit Mills Dean (40)) from the OpenEye software package. The combo score is the sum of the shape Tanimoto coefficient and the color score.

Cell culture

Fibroblasts derived from an unaffected individual (GM07492 line), HD patient (GM04281 line expressing *HTT* gene with normal repeat tract, r(CAG)₁₇, and expanded repeats, r(CAG)₆₉) as well as myotonic dystrophy type 1 probands (GM03987, GM04033 and GM03989 lines expressing mutant *DMPK* transcript containing ~500, ~1000 and ~2000 CUG repeat tracts, respectively) were obtained from the Coriell Cell Repositories. Human Neuronal Epithelioma cell line, SK-N-MC (neuroblastoma) and SK-N-MC with stable expression of 74 CAG repeats in 3'UTR of EGFP reporter gene (neuroblastoma (CAG)₇₄) (17) were a gift from Professor Włodzimierz J. Krzyżosiak (Institute of Bioorganic Chemistry, Polish Academy of Sciences in Poznań, Poland). All fibroblast cell lines were maintained in MEM medium (Invitrogen) supplemented with 10% FBS (Sigma), 1× antibiotic antimycotic solution (Sigma) and 1× MEM non-essential amino acid solution (Sigma). Neuroblastoma and HEK293 cells were grown in 1× DMEM medium (Invitrogen) supplemented with 10% FBS and 1× antibiotic antimycotic solution. Overexpression of MBNL1 protein was performed in HEK293 cells grown on 4 cm plates (~60% confluence) transfected with 1.5 μg of plasmid coding EGFP-MBNL1 fusion protein using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. After 48 h, cells were harvested and RNA was isolated.

Treatment of Cell Lines with Small Molecules

Fibroblast and SK-N-MC cells were seeded in 6-well plates in complete medium (as above) and grown to ~60% confluence. Then, fresh medium containing the small molecule of interest was added at different final concentration: 50 μ M – 300 μ M. For untreated cells, the same volume of water was added. After 48 h, cells were harvested and total RNA was isolated.

RT-PCR splicing assays

Total RNA was isolated from cells using TRI Reagent (Sigma) according to the manufacturer's instructions. RNA concentration was measured by NanoDrop spectrophotometer. Approximately 1 μ g of total RNA was reverse transcribed using a Superscript III kit (Invitrogen) and random primers (Promega) according to manufacturer's recommendations. All cDNAs were diluted 5-fold with water prior to PCR. GoTaq Flexi DNA Polymerase (Promega) and 0.1 mM dNTP mix, 1.5 mM MgCl₂, 1 μ M of each primer were used for RT-PCR assays. Please see the Supporting Information for the primers and cycling conditions used to determine the splicing products for each transcript.

The products of PCR amplification were separated using a 2% agarose gel in 1 $\times$ TBE buffer. The gels were stained with 0.5 μ g/ml ethidium bromide and imaged. Each cDNA sample was assayed in three or six independent RT-PCR reactions. The fractions of alternative splicing isoforms were calculated by dividing the intensity of PCR product band corresponding to the normal splicing variant (opposite of the DM isoform) by the total intensity of both splicing forms. Statistical significance was assessed using the average percentage ($\pm$ SD) of the normal isoform obtained from at least six independent experiments in comparison the average percentage ($\pm$ SD) of untreated cells. Statistical significance was determined by a two-tailed t-test using Microsoft Excel (* for $P < 0.05$ and ** for $P < 0.01$).

Fluorescence In Situ Hybridization (FISH)

For RNA fluorescence *in situ* hybridization (FISH) experiments, treated and untreated fibroblasts were grown on cover slips and then fixed in 4% PFA/PBS at 4 $^{\circ}$ C and washed three times with 1 $\times$ PBS. The cells were pre-hybridized with 30% formamide and 2 $\times$ SSC for 10 min, followed by hybridization in buffer containing 30% formamide, 2 $\times$ SSC, 0.02% BSA, 66 mg/ml yeast tRNA (Ambion), 480 U/ml RNasin (Promega) and 2 ng/ml DNA/LNA probes (CTG)₆-CA. The probe was labeled at the 5'-end with Cy3 and modified at positions 2, 5, 8, 13, 16 and 19 with LNA residues. Post-hybridization washing was completed in 30% formamide and 2 $\times$ SSC at 45 $^{\circ}$ C for 30 min. followed by 1 $\times$ SSC at 37 $^{\circ}$ C for the next 30 min. FISH images were acquired on a Zeiss Axioskop 2 Plus microscope (zoom 630 $\times$) equipped with a AxioCam HRm camera and processed with AxioVision Release 4.6 software.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ABBREVIATIONS

AMCA	9-amino-6-chloro-2-methoxy acridine
AO	acridine orange
ATP2A1	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1
DAPI	4',6-diamidino-2-phenylindole

DM1	myotonic dystrophy type 1
DMPK	dystrophia myotonica protein kinase
DNM	Daunomycin
FISH	fluorescence <i>in situ</i> hybridization
FXS	Fragile X Syndrome
FXTAS	Fragile X-associated Tremor Ataxia Syndrome
HD	Huntington's disease
IC₅₀	half-maximal inhibitory concentration
K_d	dissociation constant
MAPT	microtubule-associated protein tau
MBNL1	muscleblind-like 1 protein
MTX	Mitoxantrone
NCI	National Cancer Institute
NCOR2	nuclear receptor corepressor 2
NTSP	Netropsin
QCN	Quinacrine
r(CAG)^{exp}	expanded r(CAG) transcripts
r(CUG)^{exp}	expanded r(CUG) transcripts
ROCS	rapid overlay of chemical structures
SCA3	spinocerebellar ataxia type 3

Acknowledgments

This work was funded by the National Institutes of Health (3R01GM079235-02S1 and 1R01GM079235-01A2 to MDD) and the Polish Ministry of Science and Higher Education grant N302 260938. MDD is a Camille & Henry Dreyfus New Faculty Awardee, a Camille & Henry Dreyfus Teacher-Scholar, and a Research Corporation Cottrell Scholar.

REFERENCES

1. Doudna JA. Structural genomics of RNA. *Nat. Struct. Biol.* 2000; 7 Suppl.:954–956. [PubMed: 11103998]
2. Batey RT, Rambo RP, Doudna JA. Tertiary Motifs in RNA Structure and Folding. *Angew. Chem. Int. Ed. Engl.* 1999; 38:2326–2343. [PubMed: 10458781]
3. Zaug AJ, Cech TR. The intervening sequence RNA of Tetrahymena is an enzyme. *Science.* 1986; 231:470–475. [PubMed: 3941911]
4. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science.* 2001; 294:853–858. [PubMed: 11679670]
5. Winkler W, Nahvi A, Breaker RR. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature.* 2002; 419:952–956. [PubMed: 12410317]
6. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The structural basis of ribosome activity in peptide bond synthesis. *Science.* 2000; 289:920–930. [PubMed: 10937990]
7. Gallego J, Varani G. Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges. *Acc. Chem. Res.* 2001; 34:836–843. [PubMed: 11601968]

8. Hamy F, Felder ER, Heizmann G, Lazdins J, Aboul-ela F, Varani G, Karn J, Klimkait T. An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc. Natl. Acad. Sci. U. S. A.* 1997; 94:3548–3553. [PubMed: 9108013]
9. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U. S. A.* 2004; 101:7287–7292. [PubMed: 15123812]
10. Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* 1999; 288:911–940. [PubMed: 10329189]
11. Woese CR, Magrum LJ, Gupta R, Siegel RB, Stahl DA, Kop J, Crawford N, Brosius J, Gutell R, Hogan JJ, Noller HF. Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res.* 1980; 8:2275–2293. [PubMed: 6159576]
12. 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]
13. 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]
14. 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]
15. Cho J, Hamasaki K, Rando RR. The binding site of a specific aminoglycoside binding RNA molecule. *Biochemistry.* 1998; 37:4985–4992. [PubMed: 9538017]
16. 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]
17. Mykowska A, Sobczak K, Wojciechowska M, Kozlowski P, Krzyzosiak WJ. CAG repeats mimic CUG repeats in the misregulation of alternative splicing. *Nucleic Acids Res.* 2011 in press.
18. 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]
19. Chaires, JB.; Ragazzon, PA.; Garbett, NC. *Curr. Protoc. Nucleic Acid Chem.* 2003. A competition dialysis assay for the study of structure-selective binding to nucleic acids. p Unit 8.3.
20. Li N, Ma Y, Yang C, Guo L, Yang X. Interaction of anticancer drug mitoxantrone with DNA analyzed by electrochemical and spectroscopic methods. *Biophys. Chem.* 2005; 116:199–205. [PubMed: 15893412]
21. Lyles MB, Cameron IL. Interactions of the DNA intercalator acridine orange, with itself, with caffeine, and with double stranded DNA. *Biophys. Chem.* 2002; 96:53–76. [PubMed: 11975993]
22. Neidle S, Pearl LH, Herzyk P, Berman HM. A molecular model for proflavine-DNA intercalation. *Nucleic Acids Res.* 1988; 16:8999–9016. [PubMed: 3174439]
23. Wang AH, Ughetto G, Quigley GJ, Rich A. Interactions between an anthracycline antibiotic and DNA: molecular structure of daunomycin complexed to d(CpGpTpApCpG) at 1.2-Å resolution. *Biochemistry.* 1987; 26:1152–1163. [PubMed: 3567161]
24. Liu Y, Peacey E, Dickson J, Donahue CP, Zheng S, Varani G, Wolfe MS. Mitoxantrone analogues as ligands for a stem-loop structure of tau pre-mRNA. *J. Med. Chem.* 2009; 52:6523–6526. [PubMed: 19839622]
25. Zheng S, Chen Y, Donahue CP, Wolfe MS, Varani G. Structural basis for stabilization of the tau pre-mRNA splicing regulatory element by novantrone (mitoxantrone). *Chem. Biol.* 2009; 16:557–566. [PubMed: 19477420]
26. Carlson CB, Vuyisich M, Gooch BD, Beal PA. Preferred RNA binding sites for a threading intercalator revealed by *in vitro* evolution. *Chem. Biol.* 2003; 10:663–672. [PubMed: 12890540]

27. Marcheschi RJ, Mouzakis KD, Butcher SE. Selection and characterization of small molecules that bind the HIV-1 frameshift site RNA. *ACS Chem. Biol.* 2009; 4:844–854. [PubMed: 19673541]
28. Tanious FA, Veal JM, Buczak H, Ratmeyer LS, Wilson WD. DAPI (4',6-diamidino-2-phenylindole) binds differently to DNA and RNA: minor-groove binding at AT sites and intercalation at AU sites. *Biochemistry.* 1992; 31:3103–3112. [PubMed: 1372825]
29. Trotta E, D'Ambrosio E, Ravagnan G, Paci M. Evidence for DAPI intercalation in CG sites of DNA oligomer [d(CGACGTCG)]₂: a ¹H NMR study. *Nucleic Acids Res.* 1995; 23:1333–1340. [PubMed: 7753623]
30. Trotta E, Paci M. Solution structure of DAPI selectively bound in the minor groove of a DNA T.T mismatch-containing site: NMR and molecular dynamics studies. *Nucleic Acids Res.* 1998; 26:4706–4713. [PubMed: 9753740]
31. Disney MD, Childs JL, Turner DH. Hoechst 33258 selectively inhibits group I intron self-splicing by affecting RNA folding. *ChemBiochem.* 2004; 5:1647–1652. [PubMed: 15532034]
32. Miletto KE, Leibowitz MJ. Pentamidine inhibition of group I intron splicing in *Candida albicans* correlates with growth inhibition. *Antimicrob. Agents Chemother.* 2000; 44:958–966. [PubMed: 10722497]
33. 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]
34. Disney MD, Childs-Disney JL. Using selection to identify and chemical microarray to study the RNA internal loops recognized by 6'-N-acylated kanamycin A. *ChemBiochem.* 2007; 8:649–656. [PubMed: 17394189]
35. Bevilacqua JM, Bevilacqua PC. Thermodynamic analysis of an RNA combinatorial library contained in a short hairpin. *Biochemistry.* 1998; 37:15877–15884. [PubMed: 9843393]
36. Serra MJ, Turner DH. Predicting thermodynamic properties of RNA. *Methods Enzymol.* 1995; 259:242–261. [PubMed: 8538457]
37. Caskey CT, Pizzuti A, Fu YH, Fenwick RG Jr, Nelson DL. Triplet repeat mutations in human disease. *Science.* 1992; 256:784–789. [PubMed: 1589758]
38. Dervan PB. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* 2001; 9:2215–2235. [PubMed: 11553460]
39. Arambula JF, Ramisetty SR, Baranger AM, Zimmerman SC. A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBLN protein binding. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:16068–16073. [PubMed: 19805260]
40. Mills JE, Dean PM. Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J. Comput. Aided Mol. Des.* 1996; 10:607–622. [PubMed: 9007693]
41. Grant JA, Gallardo MA, Pickup BT. A fast method of molecular shape comparison. A simple application of a Gaussian description of molecular shape. *J. Comput. Chem.* 1996; 17:1653–1666.
42. Haigh JA, Pickup BT, Grant JA, Nicholls A. Small molecule shape-fingerprints. *J. Chem. Inf. Model.* 2005; 45:673–684. [PubMed: 15921457]
43. 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]
44. 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]
45. Day JW, Ranum LP. RNA pathogenesis of the myotonic dystrophies. *Neuromuscul. Disord.* 2005; 15:5–16. [PubMed: 15639115]
46. Li LB, Bonini NM. Roles of trinucleotide-repeat RNA in neurological disease and degeneration. *Trends Neurosci.* 2010; 33:292–298. [PubMed: 20398949]
47. Wojciechowska M, Krzyzosiak WJ. Cellular toxicity of expanded RNA repeats: focus on RNA foci. *Hum. Mol. Genet.* 2011 in press.
48. Shieh SY, Bonini NM. Genes and pathways affected by CAG-repeat RNA-based toxicity in *Drosophila*. *Hum. Mol. Genet.* 2011 in press.

49. Tsoi H, Lau CK, Lau KF, Chan HY. Perturbation of U2AF65/NXF1-mediated RNA nuclear export enhances RNA toxicity in polyQ diseases. *Hum. Mol. Genet.* 2011; 20:3787–3797. [PubMed: 21725067]
50. Ho TH, Savkur RS, Poulos MG, Mancini MA, Swanson MS, Cooper TA. Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. *J. Cell. Sci.* 2005; 118:2923–2933. [PubMed: 15961406]
51. Hsu RJ, Hsiao KM, Lin MJ, Li CY, Wang LC, Chen LK, Pan H. Long tract of untranslated CAG repeats is deleterious in transgenic mice. *PLoS One.* 2011; 6:e16417. [PubMed: 21283659]
52. Wang LC, Chen KY, Pan H, Wu CC, Chen PH, Liao YT, Li C, Huang ML, Hsiao KM. Muscleblind participates in RNA toxicity of expanded CAG and CUG repeats in *Caenorhabditis elegans*. *Cell. Mol. Life. Sci.* 2011; 68:1255–1267. [PubMed: 20848157]
53. Li LB, Yu Z, Teng X, Bonini NM. RNA toxicity is a component of ataxin-3 degeneration in *Drosophila*. *Nature.* 2008; 453:1107–1111. [PubMed: 18449188]
54. 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]
55. Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshire MG, Harper PS, Brook JD. Three proteins, MBNL, MBLN and MBXL, co-localize *in vivo* with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum. Mol. Genet.* 2002; 11:805–814. [PubMed: 11929853]
56. 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]
57. 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]
58. Miller JW, Urbinati CR, Teng-Umuay 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]
59. 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]
60. Gomes-Pereira M, Cooper TT, Gourdon G. Myotonic Dystrophy mouse models: towards rational therapy development. *Trends in Mol. Med.* 2011; 17:506–517. [PubMed: 21724467]
61. Childs-Disney JL, Tsvitov PB, Disney MD. Using modularly assembled ligands to bind RNA internal loops separated by different distances. *ChemBiochem.* 2011
62. Yu Z, Wang AM, Robins DM, Lieberman AP. Altered RNA splicing contributes to skeletal muscle pathology in Kennedy disease knock-in mice. *Dis Model Mech.* 2009; 2:500–507. [PubMed: 19692580]
63. Peyret N, Seneviratne PA, Allawi HT, SantaLucia J Jr. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry.* 1999; 38:3468–3477. [PubMed: 10090733]
64. SantaLucia J Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A.* 1998; 95:1460–1465. [PubMed: 9465037]
65. Puglisi JD, Tinoco I Jr. Absorbance melting curves of RNA. *Methods Enzymol.* 1989; 180:304–325. [PubMed: 2482421]

Mitoxantrone (MTX)

Acridine Orange (AO)

4',6-Diamidino-2-Phenylindole (DAPI)

Daunomycin (DNM)

Proflavin

Quinacrine (QNC)

Netrospin (NTSP)

9-Amino-6-Chloro-2-Methoxy Acridine (AMCA)

Figure 1.

Structures of the small molecules known to have affinity for nucleic acids that were initially studied for binding to RNA 1×1 nucleotide internal loops via competition dialysis and various spectroscopic methods.

Figure 2.

The secondary structures of the RNAs used to study the binding and selectivity of DNA-binding agents and related compounds. The nomenclature used for each of these RNAs is indicated below the secondary structure.

Figure 3.

The data from competition dialysis used to study the binding of the ligands shown in Figure 1 to RNA structures shown in Figure 2. Competition dialysis experiments were completed by incubating one ligand with the RNAs indicated.

Figure 4.

The data from competition dialysis used to study the compounds from the NCI library showing preference of RNA loop over a fully paired RNA. Only a subset of these compounds could be tested for binding using this method due to a lack of an absorbance band that is distinct from that of RNA.

Figure 5.

D6 corrects pre-mRNA splicing defects associated with sequestration of MBNL1 by r(CAG)^{exp}. A, microscopic images that show cell morphology is not affected by increasing concentration of D6. B, D6 for improves pre-mRNA splicing defects HD patient-derived cell line. Statistically significant improvement of splicing is observed for both pre-mRNAs tested when the cells are treated with 100–300 μ M D6. C, D6 improves pre-mRNA splicing defects in a neuroblastoma cells that stably expresses r(CAG)74. Statistically significant improvement of splicing is observed for both pre-mRNAs tested when the cells are treated with 100–300 μ M D6. D, D6 shifts the splicing of *MAPT* pre-mRNA towards the DM phenotype in both r(CAG)-expressing and normal neuroblastoma cells. E, fluorescence *in situ* hybridization (FISH) imaged using fluorescence microscopy show that D6 does not disrupt the formation of nuclear r(CAG)^{exp} foci in HD-derived fibroblasts (upper panels). Cell nuclei were stained with DAPI (lower panels). A * indicates $P < 0.05$ while ** indicates $P < 0.01$. RT-PCR bands representing both normal (N) and DM-like (DM) splicing isoforms are indicated on the side of gel images.

Table 1

Binding constants for DAPI and DAPI-like ligands for nucleic acid targets. Dissociation constants are reported in μM .

Ligand	5'CAAG/3'GUC	5'CAAG/3'GAC	5'CUUG/3'GUC	5'CGGG/3'GGC	5'CCGG/3'GCG
DAPI ^a	>2	0.22±0.06	0.55±0.2	0.50±0.1	0.99±0.2
D2 ^a	>2	0.70±0.08			
D6 ^b		0.06±0.03			

^aThese K_d 's were measured by using a direct binding assay since the compounds are fluorescent.

^bThese K_d 's were measured by using a competitive binding assay in which D2 was used as the fluorescent indicator.

Table 2

The change in melting temperature, ΔT_m , of an RNA displaying 5'CAG/3'GAC motif in the presence of various ligands.

Compound	ΔT_m (1 equiv.)	ΔT_m (2 equiv.)
DAPI	6	10
D1	2	4
D2	5	10
D4	2	5
D6	3	9
D7	3	8

Table 3

Structures of DAPI-like compounds identified from a virtual screen and the corresponding shape Tanimoto coefficients, color scores, combo scores, and IC_{50} 's.

Compound ID	Structure	Shape Tanimoto	Colored Score	Combo Score	IC_{50} (μ M)
D1		0.78	0.48	1.22	60
D2		0.86	0.78	1.64	>125
D3		0.90	0.63	1.53	ND ^d
D4		0.77	0.55	1.32	60
D5		0.75	0.42	1.17	500
D6		0.87	0.57	1.44	3

Compound ID	Structure	Shape Tanimato	Colored Score	Combo Score	IC ₅₀ (μM)
D7		0.88	0.60	1.35	60
D8		0.82	0.43	1.25	ND ^a
D9		0.87	0.40	1.27	>1000

^aThe IC₅₀ could not be measured because the compound interfered with the FRET measurements.