ACS Chemical Biology

Articles

pubs.acs.org/acschemicalbiology

A Small Molecule That Targets r(CGG)^{exp} and Improves Defects in Fragile X-Associated Tremor Ataxia Syndrome

Matthew D. Disney,*,† Biao Liu,† Wang-Yong Yang,† Chantal Sellier,§ Tuan Tran,†,‡ Nicolas Charlet-Berguerand,§ and Jessica L. Childs-Disney†

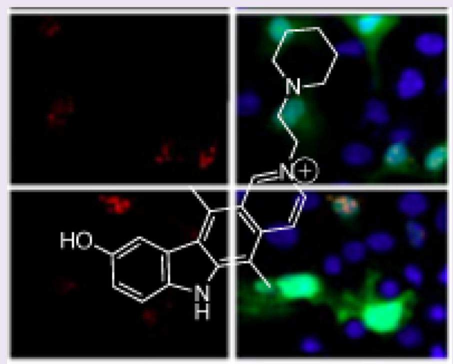
†Department of Chemistry, The Kellogg School of Science and Engineering, The Scripps Research Institute, Scripps Florida, 130 Scripps Way 3A1, Jupiter, Florida 33458, United States

‡Department of Chemistry, University at Buffalo, Buffalo, New York 14620, United States

§Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM) U964, Centre National de la Recherche Scientifique (CNRS) UMR7104, University of Strasbourg, Illkirch, France

Supporting Information

ABSTRACT: The development of small molecule chemical probes or therapeutics that target RNA remains a significant challenge despite the great interest in such compounds. The most significant barrier to compound development is defining which chemical and RNA motif spaces interact specifically. Herein, we describe a bioactive small molecule probe that targets expanded r(CGG) repeats, or r(CGG)^{exp}, that causes Fragile X-associated Tremor Ataxia Syndrome (FXTAS). The compound was identified by using information on the chemotypes and RNA motifs that interact. Specifically, 9-hydroxy-5,11-dimethyl-2-(2-(piperidin-1-yl)ethyl)-6*H*-pyrido[4,3-b]carbazol-2-ium binds the 5′C<u>G</u>G/3′G<u>G</u>C motifs in r(CGG)^{exp} and disrupts a toxic r(CGG)^{exp}-protein complex *in vitro*. Structure−activity relationship studies determined that the alkylated pyridyl and phenolic side chains are important chemotypes that drive molecular recognition of r(CGG)^{exp}. Importantly, the compound is efficacious in FXTAS model cellular systems as evidenced by its ability to improve FXTAS-associated pre-mRNA splicing defects and to reduce the size and number of r(CGG)^{exp}-containing nuclear foci. This approach may establish a general strategy to identify lead ligands that target RNA while also providing a chemical probe to dissect the varied mechanisms by which r(CGG)^{exp} promotes toxicity.



RNA plays diverse and important roles in biological processes.[1] Aberrant RNA function causes many severe diseases.[2] For example, microRNA dysregulation can contribute to cancer,[3] and single nucleotide mutations in mRNAs cause β-thalassemia and inherited breast cancer.[4] RNA trinucleotide repeat expansions cause or contribute to various neurological disorders[5] including Fragile X-associated Tremor Ataxia Syndrome (FXTAS), myotonic dystrophy type 1 (DM1), and Huntington's disease (HD)[6] and may contribute to Fragile X Syndrome (FXS).[7]

Although RNA transcripts with expanded repeats cause the diseases mentioned above, the physiological response to the repeats, and thus the causes of disease, are quite different. Differences are mainly due to the location of the expanded repeats in a given mRNA. For example, HD is caused by an expansion of r(CAG) in the coding region of huntingtin mRNA. In the most well established mechanism of HD, disease is caused when expanded r(CAG) repeats are translated into a toxic polyQ version of huntingtin.[8] Thus, HD is caused by a gain-of-function at the protein level. In FXS, >200 copies of d(CGG) in the 5′ untranslated region (UTR) of the fragile X mental retardation 1 (*FMR1*) gene causes disease by initiating local methylation, transcriptional silencing, and loss of fragile x mental retardation protein (FMRP).[9] Thus, FXS is caused by a loss-of-function by a protein. Lastly, FXTAS and DM1 are caused when expanded repeats present in UTR's sequester proteins that are involved in pre-mRNA splicing regulation.[10,11] Sequestration of these proteins causes the aberrant splicing of a variety of pre-mRNAs, leading to the expression of defective proteins. Thus, FXTAS and DM1 are caused by an RNA gain-of-function.

Despite the contribution of expanded RNA repeats to diseases, there are few compounds that target these RNAs in particular and non-ribosomal RNAs in general. Our group recently reported two approaches to design small molecules[12,13] and modularly assembled compounds[14] that bind RNA and modulate its function *in vivo*. In particular, we have used information about RNA motif−small molecule interactions[15−17] and chemical similarity searching[18−21] to design bioactive ligands that target r(CUG)^{exp} and r(CAG)^{exp}, which cause DM1 and HD, respectively.[12−14] These approaches were thus applied to develop bioactive small molecules that target r(CGG)^{exp}.

Received: March 22, 2012
Accepted: July 30, 2012
Published: September 4, 2012

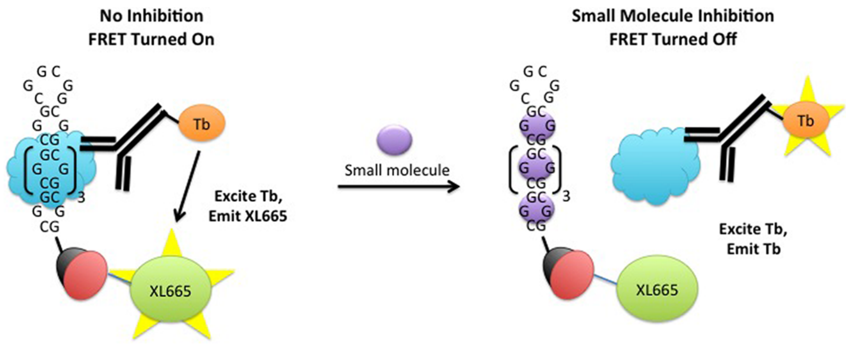


Figure 1. Schematic of the protein displacement assay that was used to identify small molecule inhibitors of the r(CGG)$_{12}$−DGCR8Δ interaction and to determine their potencies. The r(CGG)$_{12}$ oligonucleotide is labeled with a 5′-biotin, while DGCR8Δ (blue cloud) contains a histidine (His) tag. Left, in the absence of inhibitor, DGCR8Δ binds to r(CGG)$_{12}$. Binding is quantified by using two antibodies that form a FRET pair, an anti-His antibody labeled with Tb that binds to DGCR8Δ and streptavidin labeled with XL665 that binds to r(CGG)$_{12}$. The two fluorophores are within close enough proximity to form a FRET pair. Tb is excited at 345 nm; the resulting emission (~545 nm) excites XL665, which emits at 665 nm. Right, in the presence of inhibitor, the r(CGG)$_{12}$−DGCR8Δ interaction is disrupted, and the two fluorophores are not within close enough proximity to form a FRET pair. Therefore, emission is observed only at 545 nm (due to Tb). XL665 emission is not observed.

All RNA trinucleotide repeats fold into a hairpin with periodically repeating non-Watson−Crick pairs (1×1 nucleotide internal loops) in the stem.[10] We therefore probed an RNA-focused small molecule library enriched with chemotypes that bind RNA 1×1 nucleotide internal loops, such as the 1×1 nucleotide GG internal loop in r(CGG)^{exp} (Figure 1).[13] It was previously determined that Hoechst 33258[22] and a related derivative[16] bind 1×1 nucleotide GG internal loops. Therefore, one of the chemotypes used as a basis for our small molecule library was Hoechst.

Since FXTAS is caused by sequestration of proteins that regulate pre-mRNA splicing, a high-throughput protein-displacement assay was used to screen for inhibitors. From this library, a designer small molecule, 9-hydroxy-5,11-dimethyl-2-(2-(piperidin-1-yl)ethyl)-6*H*-pyrido[4,3-*b*]carbazol-2-ium, was identified. The compound binds tightly to RNAs containing non-Watson−Crick GG pairs and is efficacious in cell culture models of FXTAS. Specifically, it improves pre-mRNA splicing defects and reduces the size and number of r(CGG)^{exp} nuclear foci. Thus, this compound may serve as a chemical probe to understand how r(CGG)^{exp} causes FXTAS and may contribute to FXS. Collectively, these studies suggest that small molecules targeting traditionally recalcitrant RNA targets can be developed.

RESULTS AND DISCUSSION

FXTAS is caused by a pathogenic mechanism in which there is a gain-of-function by an expanded r(CGG) repeat, or r(CGG)^{exp}, present in the 5′ UTR of the *FMR1* gene.[10] r(CGG)^{exp} folds into a hairpin structure with regularly repeating non-Watson-Crick GG pairs closed by GC base pairs, or 5′C<u>G</u>G/3′G<u>G</u>C motifs (Figure 1).[23] FXTAS patients are carriers of premutation alleles (55−200 repeats) and have increased *FMR1* mRNA levels and normal or moderately low FMRP protein expression levels.[24,25] Evidence for RNA gain-of-function comes from animal models and cell-based assays. For example, insertion of untranslated r(CGG)^{exp} (of the length that causes FXTAS) into mice and *Drosophila* cause deleterious effects like those observed in humans that have FXTAS.[26,27] In particular, it has been shown that there is genetic interaction between r(CGG)^{exp} and Purα that mediates neurodegeneration.[28] In cell-based models, r(CGG)^{exp} forms nuclear inclusions, and the size of inclusions scales with the length of the repeat and the age of death from the disease.[29,30]

A more detailed mechanism for the RNA gain-of-function has recently been elucidated from studies of patient-derived cell lines. In studies by the Charlet group, it was shown that r(CGG)^{exp} sequesters and inactivates the Src-Associated substrate during mitosis of 68 kDa protein (Sam68).[10] Sam68 is a nuclear RNA-binding protein involved in alternative splicing regulation,[32] and the sequestration of Sam68 by r(CGG)^{exp} leads to the pre-mRNA splicing defects observed in FXTAS patients.[10] The RNA−protein complex is a scaffold for the assembly of other proteins such as muscleblind-like 1 protein (MBNL1) and heterogeneous nuclear ribonucleoprotein (hnRNP). Additional data indicate that Sam68 does not bind r(CGG)^{exp} directly, but rather the interaction is mediated by another protein. The Charlet group has recently determined that this protein is DGCR8 (Sellier, C. (IGBMC, University of Strasbourg), Freyermuth, F. (IGBMC, University of Strasbourg), Tabet, R. (IGBMC, University of Strasbourg), Tran, T. (The Scripps Research Institute), He, F. (University of Michigan), Ruffenach, F. (IGBMC, University of Strasbourg), Alunni, V. (IGBMC, University of Strasbourg), Moine, H. (IGBMC, University of Strasbourg), Thibault, C. (IGBMC, University of Strasbourg), Page, A. (IGBMC, University of Strasbourg), Tassone, F. (University of California, Davis), Willemsen, R. (Erasmus MC, Rotterdam), Disney, M. D. (The Scripps Research Institute), Todd, P.K. (University of Michigan), Hagerman, P. (University of California, Davis), and Charlet-Berguerand, N. (IGBMC, University of Strasbourg); unpublished data and ref 31). Taken together, targeting r(CGG)^{exp} with a small molecule to inhibit protein binding is an attractive treatment for FXTAS. We therefore screened a library enriched in small molecules that are biased, or focused, for binding RNA to identify lead ligands that bind r(CGG)^{exp}.

Probing an RNA-Focused Small Molecule Library To Identify Inhibitors of the r(CGG)$_{12}$−DCGR8Δ Complex. In order to construct a library of small molecules that is enriched in compounds that have the potential to recognize RNA 1×1 nucleotide internal loops like the ones that are displayed in r(CGG)^{exp} (Figure 1), previously reported chemical similarity searches were employed.[12,13] Those searches identified compounds that are similar to the bis-benzimidazole Hoechst 33258, 4′,6-diamidino-2-phenylindole (DAPI), and pentam-

dx.doi.org/10.1021/cb300135h | *ACS Chem. Biol.* 2012, 7, 1711−1718

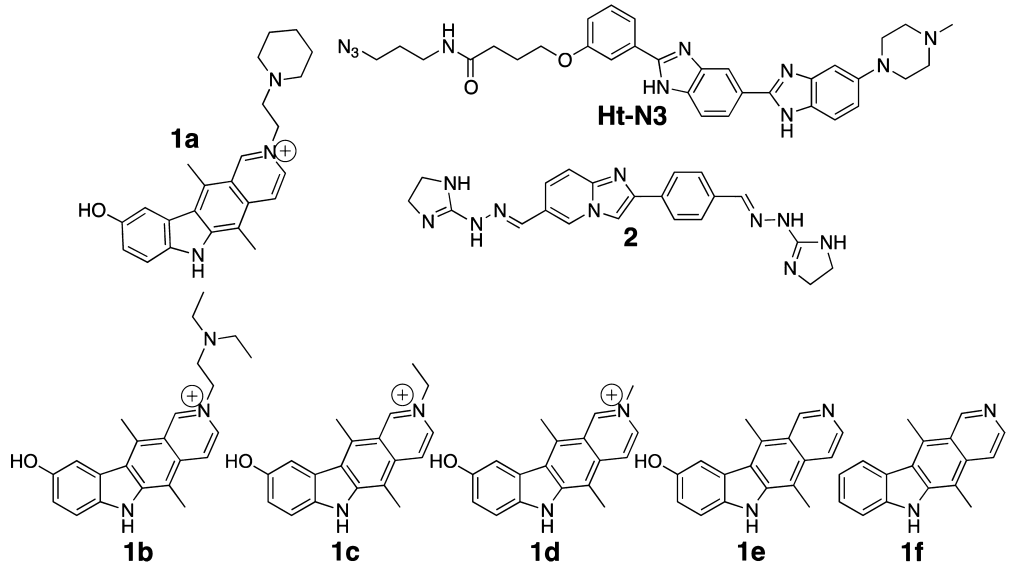


Figure 2. Structures of the small molecules identified from an RNA-focused library that inhibit the $r(CGG)_{12}$−DGCR8Δ interaction and derivatives of the most potent compound (**1a**). **1b**−**1f** were used to construct structure−activity relationships and define the active pharmacophore. Inhibition is markedly decreased for derivatives **1e** and **1f** (Table 1).

Table 1. Potencies of 1a−1f for Disruption of the $r(CGG)_{12}$−DGCR8Δ Complex and Corresponding Affinities for an RNA Containing One 5′C<u>G</u>G/3′G<u>G</u>C Motif[a]

	1a	1b	1c	1d	1e	1f
% displacement at 25 μM	85 ± 1	91 ± 5	96 ± 9	87 ± 5	46 ± 5	0
IC_{50}, μM	13 ± 0.4	8 ± 0.3	13 ± 0.2	7 ± 0.2	∼25	ND[b]
K_d, nM	76 ± 4	38 ± 1	69 ± 5	50 ± 18	NM[c]	NM[c]

[a]The potencies of the compounds are reported as IC_{50}'s as determined from the TR-FRET assay. [b]ND denotes that no determination could be made. [c]NM denotes that no measurement was made.

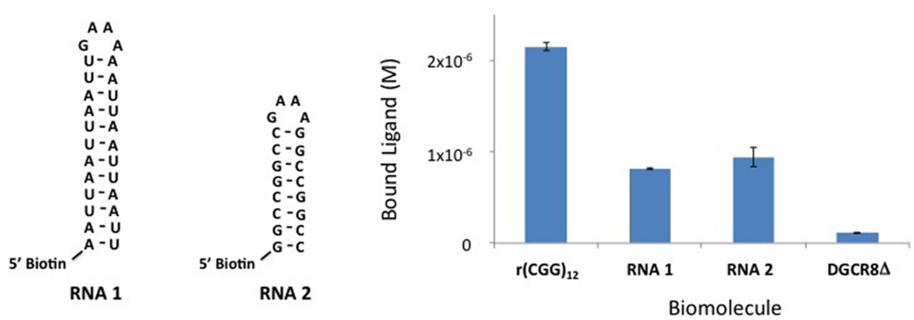


Figure 3. Results of competition dialysis experiments used to assess the specificity of **1a** for $r(CGG)_{12}$. Left, the secondary structures of two fully paired RNAs used in competition dialysis. Right, plot of the amount of ligand bound to various RNAs and DGCR8Δ.

idine. This RNA-focused collection of small molecules contains two small molecules that improve defects that are associated with $r(CAG)^{exp}$ and $r(CUG)^{exp}$ in cell culture models of HD and DM1, respectively.[12,13] Thus, Hoechst-, pentamidine-, and DAPI-like compounds were screened to identify inhibitors of the $r(CGG)_{12}$−DCGR8Δ protein complex.

Screening was completed using a time-resolved FRET assay that has been previously described for identifying inhibitors of the $r(CUG)^{exp}$−MBNL1 and $r(CAG)^{exp}$−MBNL1 complexes (Figure 1).[12,13] Briefly, a 5′-biotinylated RNA oligonucleotide containing 12 r(CGG) repeats is incubated with His-tagged DGCR8Δ. The ligand of interest is then added, followed by addition of two antibodies that recognize the RNA (streptavidin-XL665) or DGCR8Δ (Tb labeled anti-His). If the compound does not displace DGCR8Δ, then Tb and XL-665 are within close enough proximity to form a FRET pair. If, however, the ligand displaces the protein, then no FRET is observed.

From this screen, three compounds (Figure 2) were identified that disrupt the $r(CGG)^{exp}$−DGCR8Δ complex in the low to mid micromolar range. (Either no or very slight inhibition was observed for all other compounds at 100 μM.) They include compounds **1a**, **2**, and **Ht-N$_3$** (Figure 2). Interestingly, all three compounds were derived from the Hoechst or *bis*-benzimidazole chemical similarity search. Dose−response curves show that **1a** and **Ht-N$_3$** disrupt the $r(CGG)_{12}$−DGCR8Δ complex with IC_{50}'s of 12 and 33 μM, respectively. Compound **2**, however, disrupts only ∼25% of the $r(CGG)_{12}$−DGCR8Δ complex at 100 μM.

Molecular Recognition of r(CGG)^{exp} by 1a. To further investigate the chemotypes in compound **1a** that allow effective recognition of r(CGG)^{exp} and inhibition of the r(CGG)$_{12}$–DGCR8Δ complex, a series of derivatives were studied (Figure 2). These compounds probe the role of (i) the identity of the alkylated pyridyl side chain, (ii) the phenolic side chain, and (iii) the positive charge. Table 1 summarizes the IC$_{50}$'s and the percentage of protein displaced from r(CGG)$_{12}$ by 25 μM of each compound. The IC$_{50}$'s for inhibition of protein binding for **1b**, **1c**, and **1d** are similar to that of **1a** (5–12 μM). The IC$_{50}$ of compounds **1e** is ~25 μM, while **1f** has no effect on protein binding at 25 μM. Taken together, the positive charge due to the alkylated pyridyl side chain and the exocyclic hydroxyl group are required for compound potency.

In the protein displacement assay, inhibition occurs if the small molecule binds the protein or the RNA. Therefore, we used competition dialysis[33] to assess the selectivity of **1a**. A series of RNA targets, including two base paired RNAs, r(CGG)$_{12}$ (a mimic of r(CGG)^{exp} used in the displacement assay; Figure 1), and DGCR8Δ were used (Figure 3). The results of these studies show that **1a** binds tightly to r(CGG)$_{12}$, while very little binding is observed to DGCR8Δ. Although some binding is observed to fully paired RNAs, less than half of the amount of ligand that partitioned into r(CGG)$_{12}$ partitioned into these samples. Thus, **1a** binds preferentially to r(CGG)$_{12}$ over the other targets tested. The binding affinities of **1a–1d** for an RNA with a non-Watson–Crick GG pair were also determined. The measured K_d's are similar for all four compounds and range from ~40 – 75 nM (Table 1), as expected on the basis of their similar potencies.

Biological Activity of 1a in Model Cellular Systems of FXTAS. In order to assess the bioactivity of **1a**, a model cellular system of FXTAS was used.[10] Previously, it has been shown that pre-mRNA splicing defects are observed in survival of motor neuron 2 (*SMN2*) and B-cell lymphoma x (*Bcl-x*) mRNAs when cells express r(CGG)^{exp}.[10] These pre-mRNA splicing defects are due to sequestration of Sam68 by r(CGG)^{exp}; Sam68 directly regulates the alternative splicing of *SMN2* and *Bcl-x*.[10] Specifically, exon 7 of the *SMN2* mRNA is included too frequently in FXTAS model systems; ~70% of *SMN2* mRNA contains exon 7 when r(CGG)^{exp} is expressed, while exon 7 is included in only ~30% of *SMN2* mRNA in cells that do not express r(CGG)^{exp} (Figure 4, top). Likewise, there are two isoforms of *Bcl-x* mRNA, *Bcl-xL* and *Bcl-xS*. In FXTAS cellular model systems, 60% of the *Bcl-x* mRNA is the *Bcl-xL* isoform. In healthy cells, only 40% of the mRNA is the *Bcl-xL* isoform (Figure 4, bottom).

When cells that express r(CGG)$_{60}$ are treated with **1a**, improvement in *SMN2* and *Bcl-x* pre-mRNA splicing defects is observed (Figure 4). Improvement of *SMN2* splicing defects can be observed when cells are treated with as little as 20 μM **1a**. *SMN2* mis-splicing is further improved at higher concentrations: treatment with 100 μM **1a** improves pre-mRNA splicing levels to approximately 70% of wild type (absence of r(CGG)^{exp}), while treatment with 500 μM restores pre-mRNA splicing to wild type levels (Figure 4). **1a** also improves dysregulation of *Bcl-x* splicing. Statistically significant improvement is observed when cells are treated with 100 μM of **1a**, while restoration of wild type splicing patterns are observed at 500 μM (Figure 4). No statistically significant effect on *SMN2* or *Bcl-x* splicing was observed when cells that do not express r(CGG)$_{60}$ are treated with **1a**. This suggests that the improvement of pre-mRNA splicing defects is due to **1a** displacing proteins from r(CGG)$_{60}$.

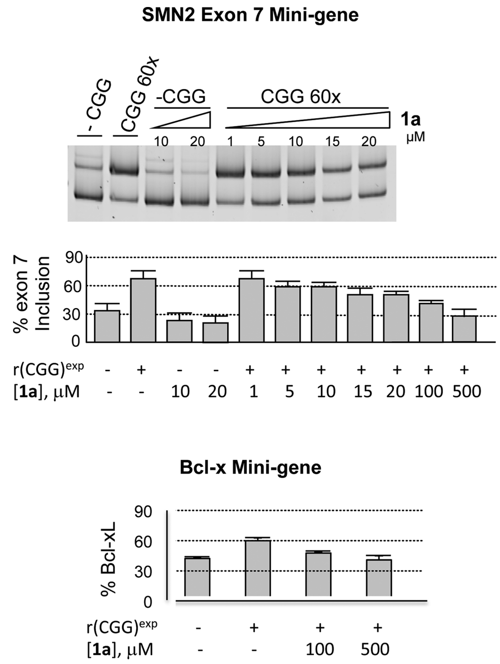


Figure 4. *In vivo* efficacy of **1a** against FXTAS as assessed by improvement in pre-mRNA splicing defects. Briefly, COS7 cells were transfected with an *SMN2* or *Bcl-x* mini-gene in the presence or absence of a mini-gene that express 60 r(CGG) repeats (CGG 60X). The cells were then treated with **1a**. Total RNA was harvested, and the alternative splicing of the *SMN2* exon 7 or *Bcl-x* exon 2 was determined by RT-PCR. (Top) **1a** improves *SMN2* pre-mRNA splicing defects. (Bottom) **1a** improves *Bcl-x* pre-mRNA splicing defects.

Control experiments were also completed to determine the specificity of **1a**, that is, if it affects the splicing of RNAs not controlled by Sam68. In these experiments, a *PLEKHH2*[17] or cardiac troponin T (cTNT)[34] mini-gene was used, as their alternative splicing is not regulated by Sam68. The addition of **1a** (500 μM) did not affect *PLEKHH2* or cTNT alternative splicing (Supplementary Figures S-4 and S-5). Thus, the effect of **1a** on pre-mRNA splicing appears to be specific to the splicing of pre-mRNAs regulated by Sam68.

Another phenotype of cells expressing r(CGG)^{exp} is the formation of nuclear foci. Additional studies were therefore completed to determine if **1a** can decrease the number or size of foci by using a fluorescence *in situ* hybridization (FISH) assay. As can be observed in Figure 5, **1a** reduces the size and the number of foci. Collectively, the improvement in the formation of foci and in the dysregulation of pre-mRNA alternative splicing show that **1a** binds r(CGG)^{exp} in cellular systems and displaces bound proteins that are then free to complete their normal physiological functions.

Comparison to Other Small Molecules That Target RNA. A few bioactive small molecules have been shown to bind to expanded triplet repeats *in vivo* and to improve associated defects.[12−14,17] For example, a *bis*-benzimidazole,[13] pentamidine,[17] and modularly assembled *bis*-benzimidazoles [14] target the r(CUG)^{exp} that causes DM1. Each improves pre-mRNA splicing defects. In general, modularly assembled ligands that bind simultaneously multiple 5′C<u>U</u>G/3′G<u>U</u>C motifs in r(CUG)^{exp} are more potent inhibitors. For example, a

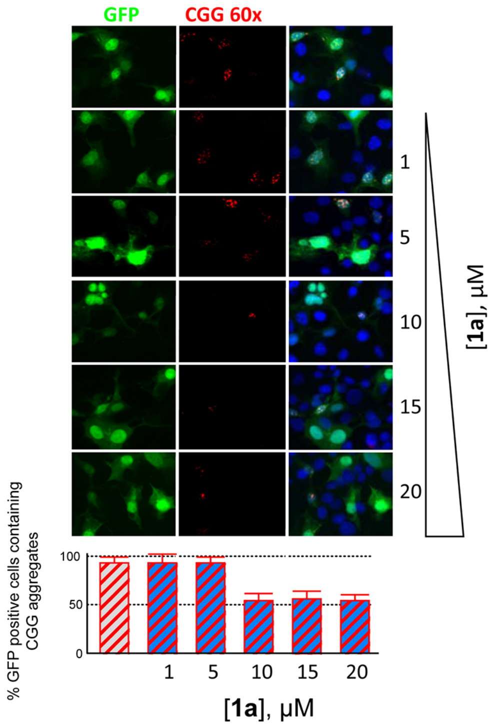


Figure 5. **1a** decreases r(CGG)^{exp}-containing nuclear foci as assessed by fluorescence *in situ* hybridization (FISH). Briefly, COS7 cells were co-transfected with a plasmid encoding 60 r(CGG) repeats and a plasmid encoding GFP. Cells were then treated with **1a** and probed with 5′(CCG)$_8$-Cy3 DNA oligonucleotide probe. Only cells that are GFP-positive were analyzed for the presence of nuclear foci. (Top) Confocal microscopy images of cells treated with different concentrations of **1a**. For all panels: left, GFP fluorescence (indicates transfected cells); middle, Cy3 fluorescence (indicates r(CGG)^{exp}); right, overlay of GFP, Cy3, and DAPI (indicates nuclei) fluorescence images. (Bottom) Plot of the number of r(CGG)^{exp}-containing nuclear foci as a function of the concentration of **1a**.

monomeric *bis*-benzimidazole (**H1**) improves pre-mRNA splicing defects in DM1 model systems to wild type levels when 2000 μM compound is used. A dimeric modularly assembled compound that displays two copies of a bis-benzimidazole, **2H-4**, improves pre-mRNA splicing levels back to wild type when cells are treated with 50 μM compound. This represents a greater than 40-fold enhancement in bioactivity provided by a modular assembly approach even though the assembled compounds are of higher molecular weight and not classically "drug-like." The improved bioactivity of the modularly assembled compound could be due to the increased surface area occupied by the compound, residence time on the RNA target, and the affinity and selectivity of modularly assembled ligands for r(CUG)^{exp}.[14,16]

In order to synthesize second-generation modularly assembled compounds that target r(CGG)^{exp}, a site that can be used to conjugate **1a**-like compounds onto an assembly scaffold must be identified. Fortuitously, our SAR studies showed that the side chain that emerges from the pyridyl group can be altered since it does not affect potency. Thus, this site is ideal for the addition of reactive groups that can be anchored onto an assembly scaffold.

Implications. The identification of a bioactive small molecule that modulates r(CGG)^{exp} provides lead compounds that could become therapies not only for FXTAS but also for other disorders that are mediated by r(CGG)^{exp}. Notably, this includes Fragile X Syndrome (FXS), the most common single gene cause of autism.[35] FXS is caused by hypermethylation of the *FMR1* locus, leading to transcriptional silencing and loss of FMRP protein.[9] There is some evidence that an RNAi-like mechanism may also contribute in which r(CGG)^{exp} is cleaved into small RNAs that enable transcriptional silencing.[7] If r(CGG)^{exp} does indeed contribute to FXS, then a small molecule that targets r(CGG)^{exp} could help differentiate modes of toxicity and/or disease mechanisms.

Lastly, the ability of a small molecule to target r(CGG)^{exp} in cellular models of FXTAS and reverse pre-mRNA splicing defects provides further evidence for an RNA gain-of-function mechanism. Since this study is another example of a small molecule that targets a non-ribosomal RNA that causes disease, it shows that small molecules can be developed to target noncoding regions in RNA.

■ METHODS

Small Molecules. All small molecules were procured from the National Cancer Institute (NCI), Sigma Aldrich, or The Scripps Research Institute. The purities of the compounds used in additional studies (IC_{50}'s, affinities, *etc.*) were determined by HPLC, and their masses were confirmed by ESI mass spectrometry. All compounds were ≥95% pure. These data are available in the Supporting Information (Figure S-6 and Table S-2).

Oligonucleotide Preparation and Purification. The RNAs used in the protein displacement assay (5′-biotin-(CGG)$_{12}$) and competition dialysis were purchased from Dharmacon. The ACE protecting groups were cleaved using Dharmacon's deprotection buffer (100 mM acetic acid, adjusted to pH 3.8 with TEMED) by incubating at 60 °C for 2 h. The samples were lyophilized, resuspended in water, and desalted using a PD-10 gel filtration column (GE Healthcare). The concentrations were determined by absorbance at 90 °C using a Beckman Coulter DU800 UV−vis spectrophotometer equipped with a Peltier temperature controller unit. Extinction coefficients (at 260 nm) were calculated using the HyTher server,[36,37] which uses nearest neighbors parameters.[38]

DGCR8Δ Expression and Purification. His-tagged DGCR8Δ was expressed in *Escherichia coli* BL21 cells *via* induction with 1 mM IPTG for 4 h. Cells were lysed in 50 mL of Lysis Buffer (50 mM Tris-Cl pH 8.0, 150 mM NaCl, 2 mM 2-mercaptoethanol, 10 mM imidazole, 0.1% (v/v) Igepal, 2 mg mL^{-1} lysozyme, and 1 mM PMSF) for 30 min on ice. DNase I was then added to a final concentration of 1 U mL^{-1}, and cells were sonicated (60% power for 9 × 10 s). The DGCR8Δ protein was purified *via* FPLC (Akta Explore, GE Healthcare) using a HiTrap Ni-column (GE Healthcare), followed by a cation exchange column (HiTrap SP FF, GE Healthcare) and a Superdex 75 size exclusion column. The protein was concentrated and dialyzed in a Vivaspin 15 centrifugal concentrator (Sartorius Stedim Biotech) into Storage Buffer (10 mM Tris-Cl pH 7.6, 200 mM NaCl, 1 mM EDTA, and 5 mM DTT, and 30% (v/v) glycerol) and stored at −20 °C.

Determination of Compound Potency *via* a Protein Displacement Assay. The protein displacement assay used to identify inhibitors of the r(CGG)$_{12}$−DGCR8Δ complex is based on PubChem BioAssay AID 2675 (Figure 1), which utilizes TR-FRET between antibodies that bind the RNA and the protein. The assay was conducted in 1X TR-FRET Assay Buffer (20 mM HEPES pH 7.5, 110 mM KCl, 110 mM NaCl, 0.1% (w/v) BSA, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.05% (v/v) Tween-20, and 5 mM DTT) with 5 μM yeast extract bulk tRNA (Roche Diagnostics), 160 nM RNA, 154.5 nM His-tagged DGCR8Δ, 40 nM Streptavidin-XL665 (HTRF, Cisbio Bioassays) and 4.4 ng μL^{-1} Anti-His$_6$-Tb (HTRF, Cisbio Bioassays).

The RNA was folded by incubation at 60 °C for 5 min in 1X Folding Buffer (20 mM HEPES, pH 7.5, 110 mM KCl, and 110 mM NaCl) followed by slow cooling to RT. Then, DGCR8Δ and the other buffer components specified above were added to the folded RNA. After incubating for 15 min at RT, 9 μL of the mixture was transferred

dx.doi.org/10.1021/cb300135h | ACS Chem. Biol. 2012, 7, 1711−1718

to a microcentrifuge tube containing 1 μL ligand at varying concentrations. A 9 μL aliquot of this final mixture was transferred to a well of a 384-well white plate (Greiner) and incubated for 1 h at RT. To exclude ligands that perturb F545/F665, a 9 μL control solution containing antibodies and different ligand concentrations in 1X TR-FRET Assay Buffer but no RNA or protein was also transferred to the plate.

The time-resolved fluorescence at 545 and 665 nm was measured using a SpectraMax M5 plate reader (Molecular Devices, Inc.) with excitation wavelength of 345 nm, cutoff at 420 nm, 200 μs delay, and 1500 μs integration time. The ratio of fluorescence intensities at 545 and 665 nm (F545/F665) for a series of ligand dilutions were fit to eq 1:

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

where y is the percentage of DGCR8Δ displacement, B is the percentage of DGCR8Δ displacement in the absence of ligand (0%), A is the maximum percentage displacement of DGCR8Δ (typically 100%), and the IC_{50} is the concentration of ligand where half of the protein is displaced from the RNA.

Competition Dialysis. Competition dialysis was completed as previously described.[33] Briefly, 2 μM RNA or protein was transferred into Slide-a-Lyzer MINI dialysis units with a molecular weight cutoff of 2,000 (Thermo Scientific), and the units were placed into a solution of 0.7 μM ligand. Two blank units containing only buffer were used to monitor equilibration by checking the absorbance at the peak wavelength. After the blank units reached equilibrium, sodium dodecyl sulfate (SDS) was added to a final concentration of 1%, and the absorbance was measured. This absorbance was used to determine total ligand concentration (C_t). The concentration of the dialysate (free ligand concentration, C_f) was determined analogously. The bound ligand concentration (C_b) was then determined using eq 2:

$$C_b = C_t - C_f \qquad (2)$$

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

RNA-Binding Assays *via* Dye Displacement. Dissociation constants were determined using an in-solution, fluorescence-based assay.[39−47] RNA was folded in 1X DNA buffer (8 mM Na_2HPO_4, pH7.0, 185 mM NaCl, 0.1 mM EDTA, 40 μg mL^{-1} BSA) at 60 °C for 5 min and allowed to slowly cool to RT. The annealed RNA was then titrated into 1X DNA buffer containing 1000 nM Hoechst 33258. Fluorescence signal was recorded using a Bio-Tek FLX-800 plate reader, which was equipped with excitation filter at 360/40 nm and emission filter at 460/40 nm. The change in fluorescence intensity as a function of RNA concentration was fit to the following equation:[39,48]

$$I = I_0 + 0.5\Delta\varepsilon\{([\mathrm{Ht}]_0 + [\mathrm{RNA}]_0 + K_t) - (([\mathrm{Ht}]_0 + [\mathrm{RNA}]_0 + K_t)^2 - 4[\mathrm{Ht}]_0[\mathrm{RNA}]_0)^{0.5}\} \qquad (3)$$

where I is the observed fluorescence intensity, I_0 is the fluorescence intensity in the absence of RNA, $\Delta\varepsilon$ is the difference between the fluorescence intensity in the absence of RNA and in the presence of infinite RNA concentration and is in units of M^{-1}, $[\mathrm{Ht}]_0$ is the concentration of Hoechst 33258, $[\mathrm{RNA}]_0$ is the concentration of the selected internal loop or control RNA, and K_t is the dissociation constant.

Ligands **1a**−**1d** were then added to compete for binding to the RNA (1 μM) in the presence of Hoechst 33258 (1 μM). Reduction in fluorescence of Hoechst 33258 was measured using a Bio-Tek FLX-800 plate reader as a function of ligand concentration (**1a**−**1d**) and was fit to the following equation:[43]

$$\Theta = \frac{1}{2[\mathrm{Ht}]_0}\left[K_t + \frac{K_t}{K_d}[\mathrm{C_t}]_0 + [\mathrm{RNA}]_0 + [\mathrm{Ht}]_0 - \sqrt{\left(K_t + \frac{K_t}{K_d}[\mathrm{C_t}]_0 + [\mathrm{RNA}]_0 + [\mathrm{Ht}]_0\right)^2 - 4[\mathrm{Ht}]_0[\mathrm{RNA}]_0}\right] + A \qquad (4)$$

where Θ is the fraction bound of Hoechst 33258, K_t is the dissociation constant for Hoechst 33258, K_d is the dissociation constant of the competing ligand, $[\mathrm{Ht}]_0$ is the total concentration of the Hoechst 33258, $[\mathrm{C_t}]_0$ is the total concentration of the competing ligand, A is the fraction bound of Hoechst 33258 at infinite concentration of the competing ligand, and $[\mathrm{RNA}]_0$ is the total concentration of RNA.

Improvement of Splicing Defects in a Cell Culture Model Using RT-PCR. In order to determine if **1a** improves FXTAS-associated splicing defects *in vivo*, a cell culture model system was used. Briefly, COS7 cells were grown as monolayers in 24- or 96-well plates in growth medium (1X DMEM, 10% FBS, and 1X GlutaMax (Invitrogen)). After the cells reached 90−95% confluency, they were transfected using Lipofectamine 2000 reagent (Invitrogen) or FuGENE HD (Roche) per the manufacturer's standard protocol. Equal amounts of a plasmid expressing a 60 CGG repeats and a mini-gene of interest (*SMN2* or *Bcl-x*) were used. Approximately 5 h post-transfection, the transfection cocktail was removed and replaced with growth medium containing **1a**. After 16−24 h, the cells were lysed in the plate, and total RNA was harvested with a Qiagen RNAEasy kit or a GenElute kit (Sigma). An on-column DNA digestion was completed per the manufacturer's recommended protocol.

A sample of RNA was subjected to reverse transcription-polymerase chain reaction (RT-PCR) using 5 units of AMV Reverse Transcriptase from Life Sciences or Superscript II (Invitrogen). Approximately 300 ng was reverse transcribed, and 150 ng was subjected to PCR. RT-PCR products were observed after 25−30 cycles of 95 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 10 min. The products were separated by polyacrylamide or agarose gel electrophoresis, stained, and imaged using a Typhoon phosphor-imager. The splicing isoforms were quantified using QuantityOne software (BioRad). Table S-2 lists the RT-PCR primers used for each mini-gene construct.

Two sets of control experiments were completed: (i) COS7 cells were co-transfected with a control plasmid that does not contain CGG repeats and the *SMN2* or *Bcl-x* mini-gene as described above, and (ii) COS7 cells were co-transfected with the mini-gene that expresses 60 r(CGG) repeats and a mini-gene that encodes a pre-mRNA whose splicing is not controlled by Sam68 (*PLEKHH2* or cTNT).[49]

Disruption of Nuclear Foci Using Fluorescence In Situ Hybridization (FISH). FISH experiments were completed as previously described.[10] Briefly, COS7 cells were plated onto glass coverslips and co-transfected with plasmids encoding for $r(CGG)_{60}$ and GFP. After 2 h, compound was added and incubated with the cells for 24 h. The cells were fixed in 4% paraformaldehyde in PBS (pH 7.4) for 15 min and washed three times with PBS. Then, they were permeabilized with 0.5% Triton X-100 in PBS. Prior to addition of the FISH probe, the cells were prehybridized in a 2X SSC buffer containing 40% formamide and 10 mg mL^{-1} BSA for 30 min. The coverslips were hybridized for 2 h in 2X SSC buffer supplemented with 40% formamide, 2 mM vanadyl ribonucleoside, 60 μg mL^{-1} tRNA, 30 μg mL^{-1} BSA, and 0.75 μg $(CCG)_8$-Cy3 DNA oligonucleotide probe. The cells were washed twice in 2X SSC containing 50% formamide and then twice in 2X SSC. Following FISH, the coverslips were incubated for 10 min in 2X SSC containing 1 μg mL^{-1} DAPI and rinsed twice in 2X SSC. The coverslips were then mounted in Pro-Long media and examined using either a simple fluorescence microscope (Leica) or a Leica DM4000 B confocal microscope.

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Disney@scripps.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank R. Parkesh for chemical similarity searches. This work was funded by the National Institutes of Health (3R01GM079235-02S1 and 1R01GM079235-01A2 to MDD), INSERM AVENIR funding (to N.C.B.), ANRGENOPAT grant P007942 (to N.C.B.), and by The Scripps Research Institute. M.D.D. is a Camille & Henry Dreyfus New Faculty Awardee, a Camille & Henry Dreyfus Teacher-Scholar, and a Research Corporation Cottrell Scholar.

■ REFERENCES

(1) Atkins, J. F., Gesteland, R. F., and Cech, T. R., Eds. (2011) *RNA Worlds: From Life's Origins to Diversity in Gene Regulation*, 3rd ed., Cold Spring Harbor Laboratory Press, New York.

(2) Cooper, T. A., Wan, L. L., and Dreyfuss, G. (2009) RNA and Disease. *Cell 136*, 777−793.

(3) Calin, G. A., and Croce, C. M. (2006) MicroRNAs and chromosomal abnormalities in cancer cells. *Oncogene 25*, 6202−6210.

(4) Wilton, S. D., and Fletcher, S. (2005) RNA splicing manipulation: strategies to modify gene expression for a variety of therapeutic outcomes. *Curr. Gene Ther. 5*, 467−483.

(5) Orr, H. T., and Zoghbi, H. Y. (2007) Trinucleotide repeat disorders. *Annu. Rev. Neurosci. 30*, 575−621.

(6) Mykowska, A., Sobczak, K., Wojciechowska, M., Kozlowski, P., and Krzyzosiak, W. J. (2011) CAG repeats mimic CUG repeats in the misregulation of alternative splicing. *Nucleic Acids Res. 39*, 8938−8951.

(7) Jin, P., Alisch, R. S., and Warren, S. T. (2004) RNA and microRNAs in fragile X mental retardation. *Nat. Cell Biol. 6*, 1048−1053.

(8) Bates, G. (2003) Huntingtin aggregation and toxicity in Huntington's disease. *Lancet 361*, 1642−1644.

(9) Jin, P., and Warren, S. T. (2000) Understanding the molecular basis of fragile X syndrome. *Hum. Mol. Genet. 9*, 901−908.

(10) Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R. K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D. J., Hagerman, P. J., and Charlet-Berguerand, N. (2010) Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO J. 29*, 1248−1261.

(11) Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science 289*, 1769−1773.

(12) Kumar, A., Parkesh, R., Sznajder, L. J., Childs-Disney, J. L., Sobczak, K., and Disney, M. D. (2012) Chemical correction of pre-mRNA splicing defects associated with sequestration of muscleblind-like 1 protein by expanded r(CAG)-containing transcripts. *ACS Chem. Biol. 7*, 496−505.

(13) Parkesh, R., Childs-Disney, J. L., Nakamori, M., Kumar, A., Wang, E., Wang, T., Hoskins, J., Tran, T., Housman, D. E., Thornton, C. A., and Disney, M. D. (2012) 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. 134*, 4731−4742.

(14) Childs-Disney, J. L., Hoskins, J., Rzuczek, S., Thornton, C., and Disney, M. D. (2012) Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive. *ACS Chem. Biol. 7*, 856−862.

(15) Cho, J., and Rando, R. R. (2000) Specific binding of Hoechst 33258 to site 1 thymidylate synthase mRNA. *Nucleic Acids Res. 28*, 2158−2163.

(16) Pushechnikov, A., Lee, M. M., Childs-Disney, J. L., Sobczak, K., French, J. M., Thornton, C. A., and Disney, M. D. (2009) 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. 131*, 9767−9779.

(17) Warf, M. B., Nakamori, M., Matthys, C. M., Thornton, C. A., and Berglund, J. A. (2009) Pentamidine reverses the splicing defects associated with myotonic dystrophy. *Proc. Natl. Acad. Sci. U.S.A. 106*, 18551−18556.

(18) Bostrom, J., Greenwood, J. R., and Gottfries, J. (2003) Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graph. Model. 21*, 449−462.

(19) Grant, J. A., Gallardo, M. A., and Pickup, B. T. (1996) A fast method of molecular shape comparison. A simple application of a Gaussian description of molecular shape. *J. Comput. Chem. 17*, 1653−1666.

(20) Haigh, J. A., Pickup, B. T., Grant, J. A., and Nicholls, A. (2005) Small molecule shape-fingerprints. *J. Chem. Inf. Model. 45*, 673−684.

(21) Mills, J. E., and Dean, P. M. (1996) Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J. Comput. Aided Mol. Des. 10*, 607−622.

(22) Cho, J., Hamasaki, K., and Rando, R. R. (1998) The binding site of a specific aminoglycoside binding RNA molecule. *Biochemistry 37*, 4985−4992.

(23) Sobczak, K., Michlewski, G., de Mezer, M., Kierzek, E., Krol, J., Olejniczak, M., Kierzek, R., and Krzyzosiak, W. J. (2010) Structural diversity of triplet repeat RNAs. *J. Biol. Chem. 285*, 12755−12764.

(24) Tassone, F., Hagerman, R. J., Loesch, D. Z., Lachiewicz, A., Taylor, A. K., and Hagerman, P. J. (2000) Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of *FMR1* messenger RNA. *Am. J. Med. Genet. 94*, 232−236.

(25) Tassone, F., Hagerman, R. J., Taylor, A. K., and Hagerman, P. J. (2001) A majority of fragile X males with methylated, full mutation alleles have significant levels of *FMR1* messenger RNA. *J. Med. Genet. 38*, 453−456.

(26) Willemsen, R., Hoogeveen-Westerveld, M., Reis, S., Holstege, J., Severijnen, L. A., Nieuwenhuizen, I. M., Schrier, M., van Unen, L., Tassone, F., Hoogeveen, A. T., Hagerman, P. J., Mientjes, E. J., and Oostra, B. A. (2003) The *FMR1* CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. *Hum. Mol. Genet. 12*, 949−959.

(27) Jin, P., Zarnescu, D. C., Zhang, F., Pearson, C. E., Lucchesi, J. C., Moses, K., and Warren, S. T. (2003) RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. *Neuron 39*, 739−747.

(28) Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T. C., Liu, H., Feng, Y., and Warren, S. T. (2007) Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a *Drosophila* model of fragile X tremor/ataxia syndrome. *Neuron 55*, 556−564.

(29) Tassone, F., Hagerman, R. J., Garcia-Arocena, D., Khandjian, E. W., Greco, C. M., and Hagerman, P. J. (2004) Intranuclear inclusions in neural cells with premutation alleles in fragile X associated tremor/ataxia syndrome. *J. Med. Genet. 41*, e43.

(30) Greco, C. M., Berman, R. F., Martin, R. M., Tassone, F., Schwartz, P. H., Chang, A., Trapp, B. D., Iwahashi, C., Brunberg, J., Grigsby, J., Hessl, D., Becker, E. J., Papazian, J., Leehey, M. A., Hagerman, R. J., and Hagerman, P. J. (2006) Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain 129*, 243−255.

(31) Sellier, C., Hagerman, P., Willemsen, R., Charlet-Berguerand, N. DROSHA/DGCR8 sequestration by expanded CGG repeats leads to

global micro-RNA processing alteration in FXTAS patients [abstract]. 12th International Fragile X Conference, Detroit, MI.

(32) Chawla, G., Lin, C. H., Han, A., Shiue, L., Ares, M., Jr., and Black, D. L. (2009) Sam68 regulates a set of alternatively spliced exons during neurogenesis. *Mol. Cell. Biol. 29*, 201−213.

(33) Chaires, J. B., Ragazzon, P. A., and Garbett, N. C. (2003) A competition dialysis assay for the study of structure-selective ligand binding to nucleic acids, *Curr. Protoc. Nucleic Acid Chem. Chapter 8*, Unit 8 3.

(34) Philips, A. V., Timchenko, L. T., and Cooper, T. A. (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science 280*, 737−741.

(35) McLennan, Y., Polussa, J., Tassone, F., and Hagerman, R. (2011) Fragile x syndrome. *Curr. Genomics 12*, 216−224.

(36) Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J., Jr. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry 38*, 3468−3477.

(37) SantaLucia, J., Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A. 95*, 1460−1465.

(38) Puglisi, J. D., and Tinoco, I., Jr. (1989) Absorbance melting curves of RNA. *Methods Enzymol. 180*, 304−325.

(39) Disney, M. D., Labuda, L. P., Paul, D. J., Poplawski, S. G., Pushechnikov, A., Tran, T., Velagapudi, S. P., Wu, M., and Childs-Disney, J. L. (2008) Two-dimensional combinatorial screening identifies specific aminoglycoside-RNA internal loop partners. *J. Am. Chem. Soc. 130*, 11185−11194.

(40) Tran, T., and Disney, M. D. (2010) Two-dimensional combinatorial screening of a bacterial rRNA A-site-like motif library: defining privileged asymmetric internal loops that bind aminoglycosides. *Biochemistry 49*, 1833−1842.

(41) Aminova, O., Paul, D. J., Childs-Disney, J. L., and Disney, M. D. (2008) Two-dimensional combinatorial screening identifies specific 6′-acylated kanamycin A- and 6′-acylated neamine-RNA hairpin interactions. *Biochemistry 47*, 12670−12679.

(42) Tran, T., and Disney, M. D. (2011) Molecular recognition of 6′-*N*-5-hexynoate kanamycin A and RNA 1×1 internal loops containing CA mismatches. *Biochemistry 50*, 962−969.

(43) Childs-Disney, J. L., Wu, M., Pushechnikov, A., Aminova, O., and Disney, M. D. (2007) A small molecule microarray platform to select RNA internal loop-ligand interactions. *ACS Chem. Biol. 2*, 745−754.

(44) Childs-Disney, J. L., and Disney, M. D. (2008) A simple ligation-based method to increase the information density in sequencing reactions used to deconvolute nucleic acid selections. *RNA 14*, 390−394.

(45) Paul, D. J., Seedhouse, S. J., and Disney, M. D. (2009) Two-dimensional combinatorial screening and the RNA Privileged Space Predictor program efficiently identify aminoglycoside-RNA hairpin loop interactions. *Nucleic Acids Res. 37*, 5894−5907.

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

(47) Velagapudi, S. P., Seedhouse, S. J., French, J., and Disney, M. D. (2011) Defining the RNA internal loops preferred by benzimidazole derivatives via 2D combinatorial screening and computational analysis. *J. Am. Chem. Soc. 133*, 10111−10118.

(48) Wang, Y., and Rando, R. R. (1995) Specific binding of aminoglycoside antibiotics to RNA. *Chem. Biol. 2*, 281−290.

(49) Warf, M. B., and Berglund, J. A. (2007) MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA 13*, 2238−2251.