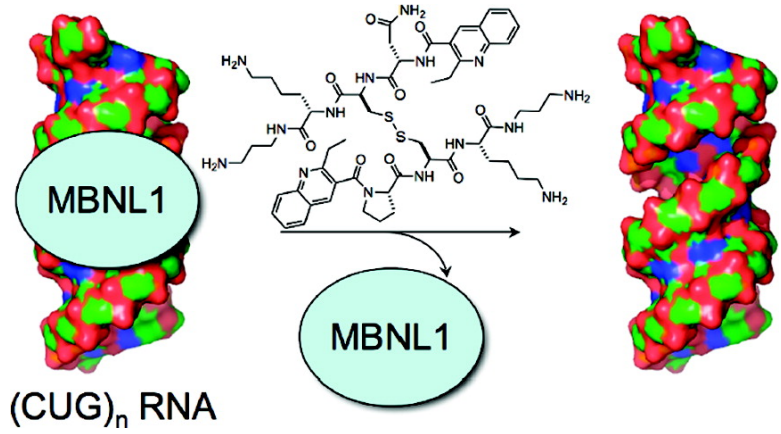
Article

Dynamic Combinatorial Selection of Molecules Capable of Inhibiting the (CUG) Repeat RNA#MBNL1 Interaction In Vitro: Discovery of Lead Compounds Targeting Myotonic Dystrophy (DM1)

Peter C. Gareiss, Krzysztof Sobczak, Brian R. McNaughton, Prakash B. Palde, Charles A. Thornton, and Benjamin L. Miller

J. Am. Chem. Soc., **2008**, 130 (48), 16254-16261 • DOI: 10.1021/ja804398y • Publication Date (Web): 08 November 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 11/08/2008

Dynamic Combinatorial Selection of Molecules Capable of Inhibiting the (CUG) Repeat RNA−MBNL1 Interaction In Vitro: Discovery of Lead Compounds Targeting Myotonic Dystrophy (DM1)

Peter C. Gareiss,[‡,⊥] Krzysztof Sobczak,[||] Brian R. McNaughton,[§,⊥] Prakash B. Palde,[‡,⊥] Charles A. Thornton,[||] and Benjamin L. Miller*[,†,‡,⊥]

Departments of Dermatology, Biochemistry and Biophysics, Chemistry, and Neurology, and The Center for Future Health, University of Rochester, Rochester, New York 14642

Received June 10, 2008; E-mail: benjamin_miller@urmc.rochester.edu

Abstract: Myotonic dystrophy type 1 (DM1), the most common form of muscular dystrophy in adults, is an RNA-mediated disease. Dramatically expanded (CUG) repeats accumulate in nuclei and sequester RNA-binding proteins such as the splicing regulator MBNL1. We have employed resin-bound dynamic combinatorial chemistry (RBDCC) to identify the first examples of compounds able to inhibit MBNL1 binding to (CUG) repeat RNA. Screening an RBDCL with a theoretical diversity of 11 325 members yielded several molecules with significant selectivity for binding to (CUG) repeat RNA over other sequences. These compounds were also able to inhibit the interaction of GGG-$(CUG)_{109}$-GGG RNA with MBNL1 in vitro, with K_i values in the low micromolar range.

Introduction

Myotonic dystrophy type 1 (DM1) is the most common form of muscular dystrophy in adults, affecting 1 in 8000 people.[1] DM1 is characterized by multisystemic symptoms, including myotonia, wasting of the muscle, testicular atrophy, cataracts, and cardiac defects. Unlike typical genetic diseases, which follow the traditional central dogma (a mutated gene is transcribed and translated to an altered encoded protein which affects cellular function), DM1 is governed by an RNA-mediated mechanism.[2] Specifically, DM1 is caused by expansion of CTG repeats located in the 3′ untranslated region of the *DMPK* (DM protein kinase) gene on chromosome 19q.[3] Transcription produces toxic mRNA containing hundreds to thousands of (CUG) repeats, which form long and stable hairpin structures.[4] The (CUG) repeat RNA accumulates in nuclear foci and sequesters RNA-binding proteins such as the MBNL (muscleblind) family of splicing regulators.[5] (CUG) repeat sequestration of these splicing regulators causes misregulated and aberrant splicing of a variety of gene products including the chloride channel 1, which is a major cause of myotonia in DM (Figure 1).[6] Small molecules and peptides capable of binding (CUG) repeat RNA and disrupting its interaction with splicing proteins are highly desirable as potential therapeutic agents to restore normal splicing.

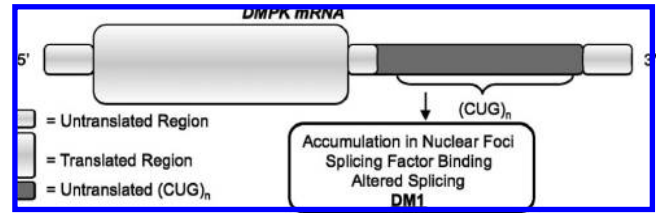


Figure 1. Transcript giving rise to DM1 pathology.

The development and synthesis of small molecules and peptides able to bind RNA with high affinity and selectivity continues to be a major focus of bioorganic chemistry.[7,8] Recent advances include the application of aminoglycosides, aminoglycoside derivatives,[9] helix threading peptides,[10] and other novel compounds to the molecular recognition of RNA sequences involved in several disease states. Unlike DNA,[11] there is not

[†] Department of Dermatology.
[‡] Department of Biochemistry and Biophysics.
[§] Department of Chemistry.
[||] Department of Neurology.
[⊥] The Center for Future Health.

(1) Machuca-Tzili, L.; Brook, D.; Hilton-Jones, D. *Muscle Nerve* **2005**, *32*, 1.
(2) Wheeler, T. M.; Thornton, C. A. *Curr. Opin. Neurology* **2007**, *20*, 572.
(3) Brook, J. D.; McCurrach, M. E.; Harley, H. G.; Buckler, A. J.; Church, A. D.; Aburatani, H.; Hunter, K.; Stanton, V. P.; Thirion, J.-P.; Hudson, T.; Sohn, R.; Zemelman, B.; Snell, R. G.; Rundle, S. A.; Crow, S.; Davies, J.; Shelbourne, P.; Buxton, J.; Jones, C.; Juvonen, V.; Johnson, K.; Harper, P. S.; Shaw, D. J.; Housman, D. E. *Cell* **1992**, *68*, 799.
(4) (a) Michalowski, S.; Miller, J. W.; Urbinati, C. R.; Paliouras, M.; Swanson, M. S.; Griffith, J. *Nucleic Acids Res.* **1999**, *27*, 3534. (b) Tian, B.; White, R. J.; Xia, T.; Welle, S.; Turner, D. H.; Mathews, M. B.; Thornton, C. A. *RNA* **2000**, *6*, 79.
(5) Lin, X.; Mankodi, A.; Kanadia, R. N.; Swanson, M. S.; Thornton, C. A. *Hum. Mol. Genet.* **2006**, *15*, 2087.
(6) (a) Mankodi, A.; Takahashi, M. P.; Jiang, H.; Beck, H. L.; Bowers, W. J.; Moxley, R. T.; Cannon, S. C.; Thornton, C. A. *Mol. Cell* **2002**, *20*, 35. (b) Wheeler, T. M.; Lueck, J. D.; Swanson, M. S.; Dirksen, R. T.; Thornton, C. A. *J. Clin. Invest.* **2007**, *117*, 3952.
(7) Thomas, J. R.; Hergenrother, P. A. *Chem. Rev.* **2008**, *108*, 1171.
(8) (a) Tor, Y. *Chembiochem* **2003**, *4*, 998. (b) Foloppe, N.; Matassova, N.; Aboul-Ela, F. *Drug Discovery Today* **2006**, *11*, 21. (c) Hermann, T. *Biopolymers* **2003**, *70*, 4.

10.1021/ja804398y CCC: $40.75

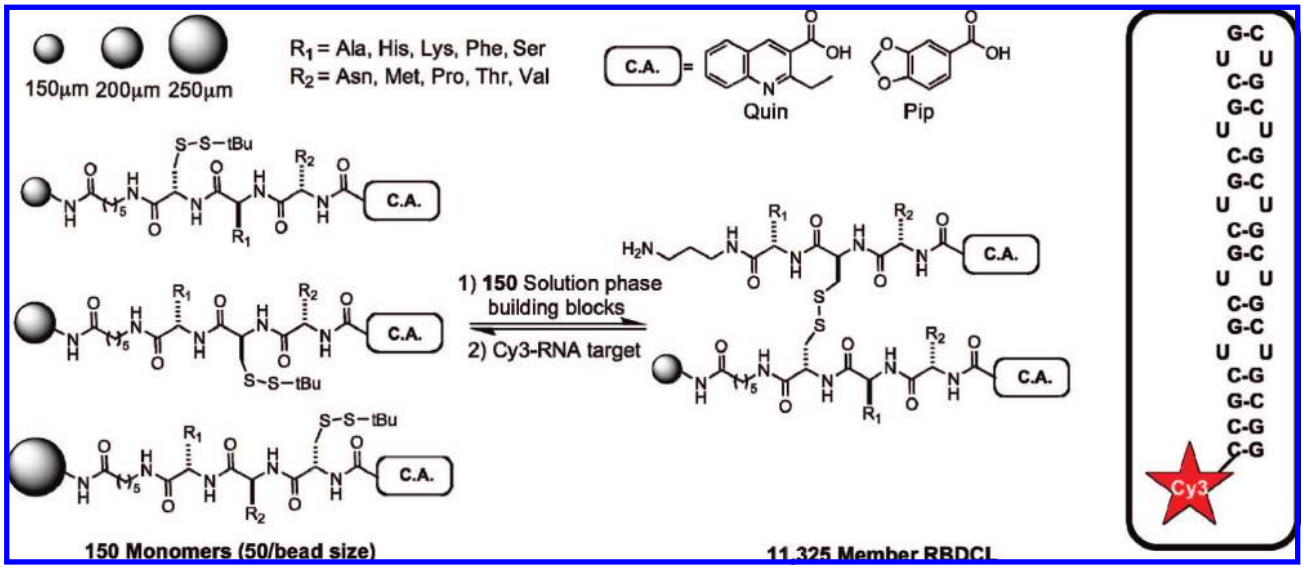


Figure 2. Design of the 11 325-member resin-bound dynamic combinatorial library. Cysteine position is encoded by bead size (normalized for equivalent compound loading), while amino acids are selected to provide unique peptide masses. Boxed: the 5′ Cy-3 labeled 5′-CCG-$(CUG)_{10}$-CGG-3′ RNA (sequence **A**) screened.

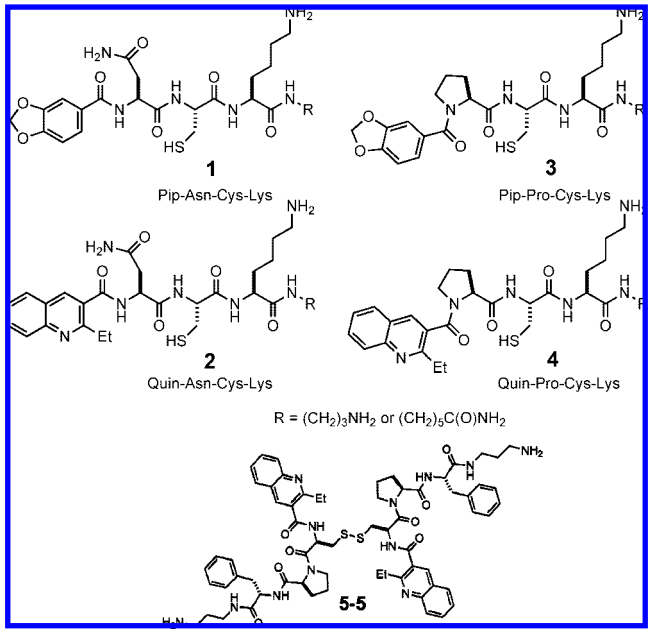
yet a canonical set of "rules" which one can follow directly relating nucleotide sequence to the design of a selective binder. Thus, combinatorial strategies are particularly attractive for rapidly identifying and testing new RNA-binding structures,[12] although directed design strategies have also found success.[13] We recently reported a novel dynamic combinatorial[14] screening method, termed resin-bound dynamic combinatorial chemistry (RBDCC),[15] and its application to the identification of a ligand for an RNA hairpin involved in HIV-1 replication.[16] While a recent report from Ludlow and Otto indicates that analysis of large dynamic combinatorial libraries is feasible in some cases,[17] the analytical challenges associated with such analyses, much like those associated with the analysis of "static" mixture libraries,[18] have kept the size of the vast majority of DCLs well below 100 compounds. RBDCC circumvents traditional DCC constraints on library size by spatial segregation of selected (active) compounds from large diverse libraries. To target RNA binding, a resin-bound dynamic combinatorial library (RBDCL) with a theoretical size of 11 325 members was created from 150 resin-attached, cysteine-containing building blocks and an identical set of solution-phase building blocks. When allowed to undergo disulfide exchange (a robust, mild, and widely used method for DCL equilibration[19]) under thermodynamic control in the presence of a fluorescently labeled target RNA, components of high-affinity binders are readily identified by physical removal of fluorescent resin beads and subsequent mass spectral analysis. Importantly, RBDCC is an innately competitive assay since resin-bound library constituents must compete for binding with compounds in solution. We reasoned that screening this RBDCL against a (CUG) repeat RNA (Figure 2) could provide novel lead compounds for DM1. In addition, comparison with the our previous results would constitute an effective test of the ability of the 11 325-member (theoretical) RNA-targeted RBDCL to yield a different ligand when screened against a different sequence from that originally targeted.

Results and Discussion

In order to ensure only products resulting from disulfide exchange were selected in the screen, we first determined a set of screening, washing, and fluorescence microscopy exposure conditions that prevented misidentification of resin-bound

(9) (a) Staple, D.; Venditti, V.; Niccolai, N.; Elson-Schwab, L.; Tor, Y.; Butcher, S. *ChemBioChem* **2007**, *9*, 93. (b) Agnelli, F.; Sucheck, S. J.; Marby, K. A.; Rabuka, D.; Yao, S. L.; Sears, P. S.; Liang, F. S.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2004**, *43*, 1562. (c) Haddad, J.; Kotra, L. P.; Llano-Sotelo, B.; Kim, C.; Azucena, E. F., Jr.; Liu, M.; Vakulenko, S. B.; Chow, C. S.; Mobashery, S. *J. Am. Chem. Soc.* **2002**, *124*, 3229. (d) Rao, Y.; Venot, A.; Swayze, E. E.; Griffey, R. H.; Boons, G. J. *Org. Biomol. Chem.* **2006**, *4*, 1328. (e) Simonsen, K. B.; Ayida, B. K.; Vourloumis, D.; Winters, G. C.; Takahashi, M.; Shandrick, S.; Zhao, Q.; Hermann, T. *ChemBioChem* **2003**, *4*, 886.
(10) (a) Krishnamurthy, M.; Gooch, B. D.; Beal, P. A. *Org. Lett.* **2004**, *6*, 63. (b) Krishnamurthy, M.; Simon, K.; Orendt, A. M.; Beal, P. A. *Angew. Chem., Int. Ed.* **2007**, *46*, 7044.
(11) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
(12) Disney, M. D.; Labuda, L. P.; Paul, D. J.; Poplawski, S. G.; Pushechnikov, A.; Tran, T.; Velagapudi, S. P.; Wu, M.; Childs-Disney, J. L. *J. Am. Chem. Soc.* **2008**, *130*, 11185.
(13) (a) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2005**, *127*, 12434. (b) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *Biochemistry* **2006**, *45*, 10928. (c) Blount, K. F.; Tor, Y. *ChemBioChem* **2006**, *7*, 1612.
(14) For recent reviews of dynamic combinatorial chemistry, see: (a) Ladame, S. *Org. Biomol. Chem.* **2008**, *6*, 219. (b) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652.
(15) McNaughton, B. R.; Miller, B. L. *Org. Lett.* **2006**, *8*, 1803.
(16) McNaughton, B. R.; Gareiss, P. C.; Miller, B. L. *J. Am. Chem. Soc.* **2007**, *129*, 11306.
(17) Ludlow, R. F.; Otto, S. *J. Am. Chem. Soc.* **2008**, *130*, 12218.
(18) (a) Carell, T.; Wintner, E. A.; Bashir-Hashemi, A.; Rebek, J. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 2059. (b) Carell, T.; Wintner, E. A.; Rebek, J. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 2061. (c) Carell, T.; Wintner, E. A.; Sutherland, A. J.; Rebek, J.; Dunayevskiy, Y. M.; Vouros, P. *Chem. Biol.* **1995**, *2*, 171. (d) Dunayevskiy, Y. M.; Vouros, P.; Wintner, E. A.; Shipps, G. W.; Carell, T.; Rebek, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6152.
(19) A subset of the large body of work in this area includes: (a) Hioki, H.; Still, W. C. *J. Org. Chem.* **1998**, *63*, 904. (b) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2000**, *122*, 12063. (c) Nicolaou, K. C.; Hughes, R.; Cho, S. Y.; Winssinger, N.; Smethurst, C.; Labischinski, H.; Endermann, R. *Angew. Chem., Int. Ed.* **2000**, *39*, 3828. (d) Leclaire, J.; Vial, L.; Otto, S.; Sanders, J. K. M. *Chem. Commun.* **2005**, 1959. (e) Hotchkiss, T.; Kramer, H. B.; Doores, K. J.; Gamblin, D. P.; Oldham, N. J.; Davis, B. G. *Chem. Commun.* **2005**, 4264. (f) Liénard, B. M. R.; Hüting, R.; Lassaux, P.; Galleni, M.; Frère, J.-M.; Schofield, C. J. *J. Med. Chem.* **2008**, *51*, 684.

monomer (present as S-*t*-butyl disulfides) as "hits". RBDCL screening was then performed as previously described in PBS pH 7.2 + 1 mM $MgCl_2$ at 22 °C for 72 h with 1 μM Cy3-CCG-$(CUG)_{10}$-CGG RNA (sequence **A**).[16] Electrophoretic RNA analysis confirmed that no RNA degradation occurred during the experiment. Post-screen fluorescence microscopy identified four beads exhibiting significant fluorescence. These four beads, representing components critical to high-affinity ligands, were removed via syringe and washed with PBS followed by methanol, and compounds were photolytically cleaved from the resin (50 μL 4:1 MeOH/H_2O, 365 nm, 24 h). The RBDCC screen is conducted under conditions such that exchange with bead-bound monomers is favored over "nonproductive" exchange in solution: each resin-bound monomer is in a concentration of 2.58 μM (0.86 μM per resin bead, and theoretically three resin beads per monomer), while solution-phase monomers are present at a concentration of 200 nM. As a result, we only observe monomer thiol-S-*t*-butyl disulfides in the mass spectrum; in this case, we identified library monomers **1**−**4** as the selected components. The four selected components had a high degree of sequence similarity: (Quin/Pip)-(Asn/Pro)-Cys-Lys. Importantly, these differ from the components obtained from screening this library against the HIV-1 frameshift stimulating RNA stemloop; in that case, **5-5** ((Quin-Cys-Pro-Phe)$_2$) was identified as the highest affinity binder following a secondary screen and independent binding analysis by surface plasmon resonance.



The identities of monomers **1**−**4** allow for 10 unique possible homo- and heterodisulfides, neglecting terminus differentiation due to bead immobilization; taking this into account raises the number of possible combinations to 16. However, in evaluating the results of an RBDCC screen, we assume that binding will be commutative for a "linear" DCL member (i.e., A-S-S-B will have the same affinity as B-S-S-A). As the disulfide linking the two monomers A and B is neither directional nor chiral, this will hold absolutely true if the termini of A and B are identical and will hold in practice if the termini are different but do not contribute significantly to binding to the target. Because the linker and terminal functionality (a propyl amine linker remaining from the synthesis of solution-phase monomers) are different in the library at hand, but are common to all library members, our assumption here is that both of these termini have only a limited, and uniform, role in binding. We recognize that this assumption may be an oversimplification and hope to examine this issue in greater detail in future efforts. However, it is important to note that secondary screening, coupled with resynthesis of "hit" compounds and solution-phase affinity testing, controls for these assumptions.

In order to confirm mass spectrometry results, and potentially narrow down the number of library hit structures, a secondary screen was conducted in which all 16 possibilities were individually synthesized on bead and exposed in parallel to 1 μM Cy3-CCG-$(CUG)_{10}$-CGG RNA. Note that, in this case, in contrast to the RBDCC screen, each bead carries only a single compound type, and no equilibration occurs. After 3 h, beads were washed and imaged by fluorescence spectroscopy. All beads bearing homo- or heterodimers fluoresced, in contrast to beads bearing thiol-S-*t*-butyl monomers alone, validating both the RBDCC process and mass spectral analysis (see selected examples in Figure 3). However, in contrast to our previous work, no clear differences in fluorescence intensity were observed among the 16 sets of resin beads, suggesting that all combinations of these four monomers could have similar affinities for RNA sequence **A**, at least in their resin bead-bound forms. Therefore, the 10 disulfides were individually synthesized in order to permit solution-phase binding analysis. Preparative scale synthesis was carried out on Wang resin with a diaminopropane linker, as in the synthesis of solution-phase monomers. This means that the compounds evaluated in solution bear one more basic group than the resin-bound compounds (4 rather than 3), a systematic change that should affect the binding ability of all compounds similarly.

A filter binding assay system was utilized to provide a rapid initial assessment of binding affinity.[20] All binding studies were performed in PBS pH 7.2 + 1 mM $MgCl_2$ at 22 °C. Various concentrations of peptides were incubated with 10 nM FAM-labeled RNA (sequence **B**: 5′-GGG-$(CUG)_{109}$-GGG-3′) in a total volume of 50 μL for 20 min. A slot blot apparatus was then assembled with a wet nitrocellulose filter, on top of a wet nylon filter, on top of filter paper (Figure 4, left). Forty microliters of each binding mixture was then loaded into an individual well of the apparatus and allowed to penetrate the filters for 10 min. This assay is predicated on the ability of nitrocellulose to retain peptides, while allowing RNA to pass through. The peptidic disulfides selected from the DCL efficiently bind to the nitrocellulose filter, and in contrast, the FAM-labeled RNA strands penetrate the nitrocellulose filter and are bound by the nylon filter. As such, RNA bound by the peptidic library members remains bound on the nitrocellulose filter, and unbound RNA passes to the nylon filter.

Densitometric analysis of the ratio of labeled RNA on the nitrocellulose to nylon filters allows quantification of binding (Figure 4, left). All binding isotherms were fit to the logistic binding model.[21] The repetitive nature of the target sequences most likely results in binding stoichiometries that are more complex than a simple 1:1 interaction; however, as all measurements were performed at [RNA] $\ll K_D$, this does not affect the validity of the reported K_D values. As seen in Figure 4, the 10 possible disulfides identified in the screen all bind GGG-$(CUG)_{109}$-GGG RNA (sequence **B**). A suite of experiments run with a ^{32}P-labeled $(CUG)_{56}$ transcript provided similar results; these are included in the Supporting Information. Compounds **3-3**, **4-4**, **2-4**, and **3-4** exhibited the highest affinity (~2 μM) in these experiments, and as such, we chose to focus on these

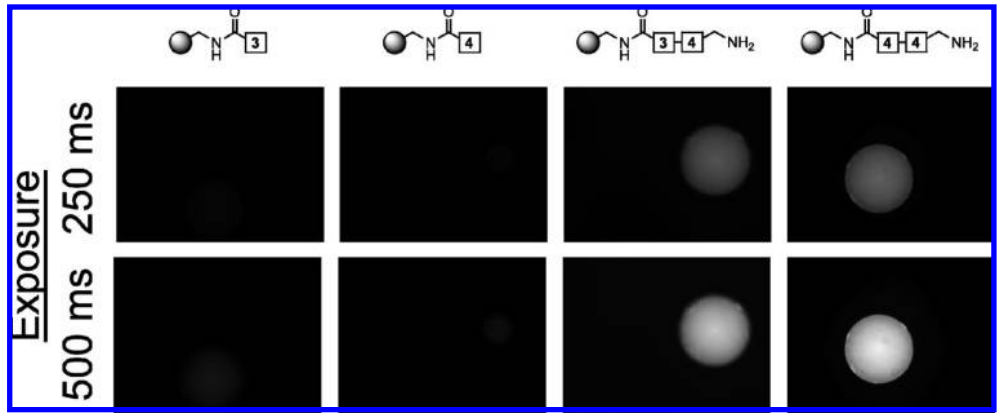


Figure 3. On-bead confirmation of binding between Cy3-CCG-$(CUG)_{10}$-CGG RNA and selected dimers as visualized by fluorescence microscopy. Two exposures are shown in order to highlight differences between strong RNA binding to bead-bound dimers, versus minimal binding to bead-bound monomers.

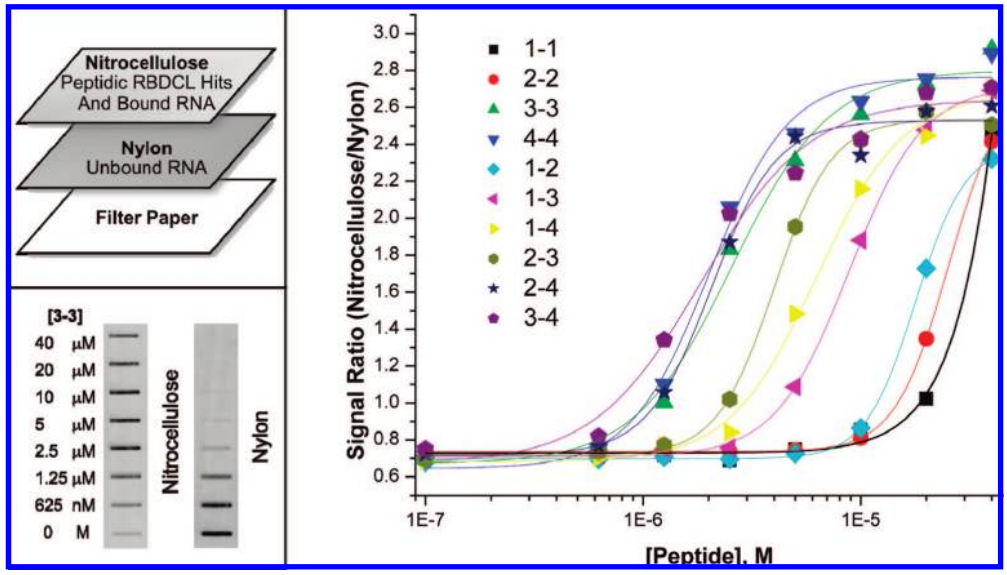


Figure 4. Top left: assembly of filters for filter binding assay. Bottom left: Raw filter images of binding analysis of compound **3-3** and RNA sequence **B** (GGG-$(CUG)_{109}$-GGG). Right: Binding curves of the 10 possible selected disulfide RBDCL ligands for sequence **B**.

compounds for further study. To test specificity and to confirm the determined binding affinities, filter binding assays and fluorescence titrations were performed with target 5′-FAM-CCG-$(CUG)_{10}$-CGG-3′ RNA (sequence **A**), an RNA incorporating the (CUG) repeat only in the stem (5′-FAM-CCG-$(CUG)_3$-GGCAAC-$(CUG)_3$-CGG-3′, sequence **C**), a (CUG) loop RNA (5′-FAM-CGCGCUGCUGCGCG-3′, sequence **D**), a (CUG)-(CAG) complementary RNA hairpin (5′-FAM-CCAGCUG-GCAACAGCUGG-3′, sequence **E**), and the HIV-1 frameshift stimulatory sequence (FSS) RNA hairpin used in our previous study (5′-FAM-GGCCUUCCCACAAGGGAAGGCC-3′, sequence **F**). Competition experiments utilizing target RNA (sequence **A**) and a 20-fold molar excess of total yeast tRNA were performed to measure nonspecific RNA binding.[22] Compound **5-5**, the HIV FSS RNA ligand, served as a negative control.

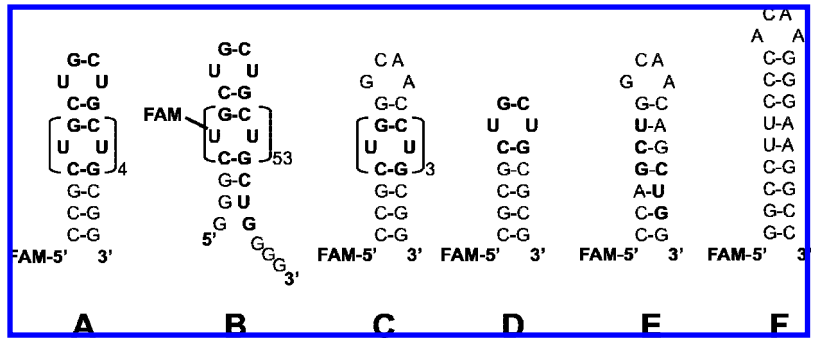
As seen in Table 1, compounds **3-3**, **4-4**, **2-4**, and **3-4** all bind (CUG) repeat RNA with similar affinities. Addition of a 40-fold base (20-fold molar) excess of total yeast tRNA results in only an approximately 2-fold loss of affinity to sequence **A**, as measured by both filter binding assay (FBA) and fluorescence titration (FT). This provides an important demonstration of selectivity for $(CUG)_n$ binding.[22] Binding to sequences **A**, **B**, and **C** is of the same order of magnitude for all four compounds, as one would expect based on the identically repeating secondary structures predicted for these sequences.[4] Curiously, although FBA gave binding constants for sequence **D** that were similar to those measured for sequences **A**−**C**, FT measurements with this sequence did not provide standard saturation profiles, likely due to differences in the two titration methods (vide infra). The importance of the U−U mismatch was assessed by measuring binding to sequence **E**; FBA indicated a roughly 5-fold decrease in affinity for all compounds, while FT indicated a 10-fold loss for compound **4-4** and >10-fold loss or no saturable binding for other compounds. Sequence **F**, the HIV-1 frameshift stimulatory sequence employed in our previous RBDCC screen,[16] provided another measure of sequence selectivity and also allowed cross-correlation with the affinity of compound

(20) (a) Tai, N.; Ding, Y.; Schmitz, J. C.; Chu, E. *Nucleic Acids Res.* **2002**, *30*, 4481. (b) Nilsson, P.; Henriksson, N.; Niedzwiecka, A.; Balatsos, N. A.; Kokkoris, K.; Eriksson, J.; Virtanen, A. *J. Biol. Chem.* **2007**, *282*, 32902.

(21) $y = B + [(A - B)/(1 + (x/x_0)^p)]$ where A = min, B = max, x = [compound], $x_0 = K_D$, and p = power.

(22) Luedtke, N. W.; Liu, Q.; Tor, Y. *Biochemistry* **2003**, *42*, 11391.

Table 1. Top: RNA Sequences Employed in Binding Constant Measurements. Bottom: K_D (μM ± standard deviation, where relevant) Values for Selected Ligands **3-3**, **4-4**, **2-4**, and **3-4** Determined by Filter Binding Assay (FBA) and Fluorescence Titration (FT)[a]



method	sequence	**3-3** K_D (μM)	**4-4** K_D (μM)	**2-4** K_D (μM)	**3-4** K_D (μM)	**5-5** K_D (μM)
FBA	**A**	5.4 ± 0.6	6.7 ± 0.2	4.5 ± 0.6	4.1 ± 0.2	ND
FBA	**A** + 20× tRNA	14 ± 4	18 ± 2.1	10 ± 0.7	9.6 ± 0.7	ND
FBA	**B**	2.5	2.1	2.1	1.9	ND
FBA	**C**	9.3 ± 0.4	7.1 ± 1.1	6.4 ± 1.0	4.7 ± 0.1	ND
FBA	**D**	7.5 ± 0.9	8.5 ± 0.1	4.6 ± 0.1	6.0 ± 0.1	ND
FBA	**E**	21 ± 1.4	>40	19.5 ± 1	21 ± 1.8	ND
FBA	**F**	41 ± 9	24 ± 5	22 ± 2	16 ± 4	ND
FT	**A**	1.4	2.2	2.1	2.1	>40
FT	**A** + 20× tRNA	2.4	3.4	2.9	4.2	ND
FT	**C**	2.1	1.7	1.6	2.6	ND
FT	**D**	NS	NS	NS	NS	ND
FT	**E**	NS	17	NS	NS	ND
FT	**F**	ND	>34	ND	>62	0.35 ± 0.11

[a] FBA were performed in triplicate, except for sequence **B**, which was done once; NS = not saturable; ND = not determined.

5-5. This molecule binds sequence **F** with a K_D of 0.35 ± 0.11 μM (as measured by fluorescence titration), while displaying little to no affinity for sequence **A** (K_D > 40 μM). Conversely, compounds **4-4** and **3-4** show very limited affinity for sequence **F** (K_D > 34 and 62 μM, respectively).

It is important to note that both filter binding assays and fluorescence titrations have their own idiosyncrasies, and these may contribute to the different levels of sequence selectivity for binding constants obtained using the two methods. Filter binding assays are not true equilibrium methods,[23] a complicating factor in their use for the extraction of equilibrium binding constants. Fluorescence titrations, on the other hand, rely on binding-induced changes in a fluorophore tag that may be distant from the actual binding site. Despite these issues, the data clearly show that all four compounds bind, and fluorescence titrations (the preferred method here because it is a true equilibrium technique) show significant selectivity for the target (CUG) repeat sequence.

It has been previously demonstrated that the aggregation of compounds can occasionally confound analysis of library screening results.[24] To ensure that aggregation was not a contributing factor in the observed binding, compounds **3-3**, **4-4**, **2-4**, and **3-4** were examined by dynamic light scattering (DLS). As shown in Table 2, only compound **3-3** showed any evidence of aggregation, albeit at concentrations much higher than those used in our experiments (100 μM and 1 mM in PBS).

Congo Red, a dye known to form aggregates,[24] served as a positive control in our experiments. At a concentration of 750 μM in PBS buffer, the dye formed aggregates with a hydrodynamic radius (H_r) of 53 nm. This was clearly evident from the high counts per second (kcps) compared to buffer, and the decay in the autocorrelation function over a 10 to 10,000 μs scale. In contrast, compounds **4**−**4**, **3**−**4** and **2**−**4** did not show any aggregation at 1 mM concentration in PBS buffer, as they yielded an autocorrelation function lacking a well-defined decay over time and very low intensities (kcps) close to the intensity value of the PBS buffer alone. However, compound **3-3** showed aggregation at a concentration of 1 mM in PBS buffer. This decreased somewhat with the addition of 0.01% Triton-X 100, as evident from the intensities, but was not completely removed. When tested at a reduced concentration (100 μM), compound **3-3** still appeared to aggregate. It was difficult to detect aggregation at concentrations lower than 100 μM due to low intensities.

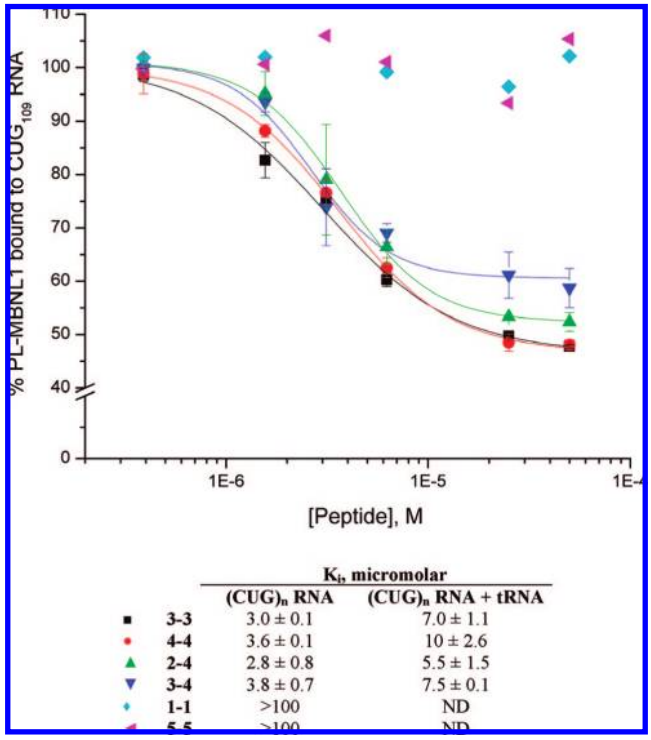
Table 2. Results of Dynamic Light Scattering (DLS) Measurements[a]

compound	concentration (μM)	H_r (nm)	intensity (kcps)
1 × PBS buffer		no particles	6.9 ± 0.51
Congo Red	750	69.25	242 ± 10
Congo Red + Triton-X 100 (0.01%)	750	40.72	190 ± 24
3-3	1000	35.54	160 ± 60
3-3 + Triton-X 100 (0.01%)	1000	30.64	78.5 ± 8.8
3-3	100	29.83	16.5 ± 4.9
4-4	1000	1.06	8.5 ± 0.72
3-4	1000	0.98	7.3 ± 1.2
2-4	1000	0.87	9.6 ± 1.8

[a] All compounds were evaluated in 1× PBS, pH 7.2. Hydrodynamic radius (H_r) values were obtained by regularization and are approximate for compounds **4**-**4**, **3**-**4**, and **2**-**4**, as the low intensity produced by non-aggregators is a significant source of error.

We next tested the ability of the selected ligands to inhibit binding of MBNL1 protein to (CUG) repeat RNA, an interaction implicated in type 1 myotonic dystrophy. Peptide-mediated

(23) Wong, I.; Lohman, T. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5428.
(24) (a) McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 4265. (b) Feng, B. Y.; Toyama, B. H.; Wille, H.; Colby, D. W.; Collins, S. R.; May, B. C. H.; Prusiner, S. B.; Weissman, J.; Shoichet, B. K. *Nat. Chem. Biol.* **2008**, *4*, 197.



		K_i, micromolar	
		$(CUG)_n$ RNA	$(CUG)_n$ RNA + tRNA
■	3-3	3.0 ± 0.1	7.0 ± 1.1
●	4-4	3.6 ± 0.1	10 ± 2.6
▲	2-4	2.8 ± 0.8	5.5 ± 1.5
▼	3-4	3.8 ± 0.7	7.5 ± 0.1
◆	1-1	>100	ND
◀	5-5	>100	ND

Figure 5. Inhibition of the $(CUG)_{109}$ RNA−MBNL1 protein interaction. Reported K_i values are an average of ≥3 independent experiments ± SD. ND = Not Determined. For clarity, only data for $(CUG)_n$ RNA experiments are plotted; $(CUG)_n$ RNA + tRNA data are shown in the Supporting Information.

inhibition of $(CUG)_{109}$-MBNL1 binding was determined using an enzyme fragment complementation (EFC) assay.[25] Briefly, this was performed by first immobilizing $(CUG)_{109}$ RNA to a 96-well plate. Next, the immobilized RNA was incubated with recombinant PL-MBNL1 fusion (where "PL" is the commercial "ProLabel" enzyme donor peptide)[26] in the presence of varying concentrations of inhibitor. The "Enzyme Acceptor" (EA-β-galactosidase, or EA-β-Gal) complement was then added, and the activity of the resulting plate, $(CUG)_{109}$ bound (PL-MBNL1)-(EA-β-Gal) complex, was monitored via a chemiluminsecent substrate. Only the EA-β-Gal bound to the PL-MBNL1 is capable of performing the luminescent reaction, and as such, any luminescence correlates to the amount of MBNL1 bound to the $(CUG)_{109}$ immobilized on the plate. Wells lacking $(CUG)_{109}$ RNA served as a background measure of nonspecific luminescence and were subtracted from each experiment to yield a 0% bound value. Wells containing no peptide inhibitor added served as 100% bound. Values of percent bound PL-MBNL1 versus peptide concentration were plotted, and data were fit to the logistic equation[21] to allow extraction of K_i values (Figure 5).

We were pleased to observe that several of the selected compounds inhibit the $(CUG)_n$−MBNL1 interaction with K_i values in the same range as their measured dissociation constants (K_D). Importantly, the selected compounds are able to inhibit the $(CUG)_n$−MBNL1 interaction in the presence of ∼40-fold base excess of yeast tRNA with only ∼2−3-fold loss in K_i. Compounds **1-1** and **5-5** do not show any inhibitory effect, as expected based on their lack of affinity for $(CUG)_n$ RNA. A maximum 50% total inhibition was observed, which may be explained by the ability of MBNL1 to bind short sequences of (CUG) RNA and the possibility that the compounds do not mask all possible MBNL1 binding sites in the $(CUG)_{109}$ hairpin. It is important to note that small changes in levels of splicing factors, such as MBNL1 sequestration by (CUG) RNA, have large effects on splicing,[27] and thus even modest inhibition of the MBNL1−(CUG) RNA interaction may be therapeutically useful.

Conclusions

In summary, screening an RBDCL with a theoretical size of 11 325 members provided ligands with good affinity and selectivity for $(CUG)_n$ repeat RNA, a causative agent of type 1 myotonic dystrophy (DM1). Importantly, the selected ligands are the first examples of compounds able to inhibit the (CUG) repeat RNA−MBNL1 protein interaction. Inhibition of this interaction in vitro for four selected library members was found to occur with low micromolar K_i values, consistent with measured K_D values. These lead compounds provide an excellent platform for ongoing SAR studies aimed at increasing affinity and specificity for (CUG) repeat RNA, as well as efforts to generate compounds suitable for in vivo studies. Finally, these results confirm the utility of the RBDCC method, in general, and specifically as a strategy for the rapid generation of sequence-selective RNA-binding compounds.

Experimental Section

General. RNA sequences **A**, **C**, **D**, **E**, and **F** were purchased from IDTDNA with RP-HPLC purification. Sequence **B** was prepared by in vitro transcription as previously described.[28] All CUG RNA variants were dissolved in 1× PBS (pH 7.2) + 1 mM $MgCl_2$ and renatured by heating to 80 °C for 2 min followed by slow cooling to room temperature (see Supporting Information for details). Total yeast tRNA was purchased (Fluka 83853, ∼20 A_{260}/mg), dissolved in 1× PBS (pH 7.2) + 1 mM $MgCl_2$ and renatured by heating to 80 °C for 2 min followed by slow cooling to room temperature.

RBDCC Screen. Synthesis of the 150 library building blocks on-bead has been described previously.[16] Following library synthesis, library screening proceeded as follows: first, control experiments were conducted to ensure that monomeric resin-bound building blocks do not bind the target RNA to any appreciable extent. In triplicate, in 1.5 mL solid-phase reaction vessels (BioRad Econo-Pac chromatography columns), 450 beads (150 of each size, 3 beads/compound, 387 μM total) were incubated with Cy3-labeled $(CUG)_{10}$ RNA sequence **A** (1 μM) in 1× PBS + 1 mM $MgCl_2$, pH 7.2. This solution was agitated for 3 h on a LabQuake rotator. The solution was then drained from the resin by vacuum, and the beads in each vessel were washed three times with buffer (1 mL, 1 min each). After washing, the resin was suspended in 2 mL of buffer and transferred to a Petri dish. Imaging the resin under a fluorescence microscope equipped with a Cy3 filter showed that no resin beads exhibited a fluorescence signal, suggesting that none of the library building block monomers (thiol-S-*t*-butyl protected) bound the target RNA.

Next, the RBDCL was screened against the target RNA sequence **A**. A heterogeneous mixture of solution-phase library constituents (30 μM based on average molecular weight) was incubated with resin-bound library constituents (450 beads; 150 of each size, 3

(25) Sobczak, K.; Thornton, C. A. Manuscript in preparation.
(26) DiscoveRx PathHunter ProLabel Detection kit: (a) Eglen, R. M. *Assay Drug Dev. Technol.* **2002**, *1*, 97. (b) Olson, K. R.; Eglen, R. M. *Assay Drug Dev. Technol.* **2007**, *5*, 137.
(27) Black, D. L. *Annu. Rev. Biochem.* **2003**, *72*, 291.
(28) Yuan, Y.; Compton, S. A.; Sobczak, K.; Stenberg, M. G.; Thornton, C. A.; Griffith, J. D.; Swanson, M. S. *Nucleic Acids Res.* **2007**, *35*, 5474.

beads per compound, 387 μM total concentration of resin-bound monomer based on average molecular weight) and Cy3-labeled $(CUG)_{10}$ RNA sequence **A** (1 μM) in 1× PBS + 1 mM $MgCl_2$, pH 7.2. Libraries were equilibrated in quadruplet for 72 h, a period of time shown by HPLC to be sufficient for equilibrium to be reached. The resin was then drained, washed four times with PBS + 1 mM $MgCl_2$ for 90 s each, plated with 2 mL of buffer in a Petri dish, and analyzed by fluorescence microscopy.

After screening, four beads exhibited significant fluorescence and were removed via syringe and washed five times with 1 mL of 1× PBS, pH 7.2, for 5 min, then five times with 1 mL of methanol for 5 min. The size of each selected bead was determined to be medium, thus identifying each of the bead-bound compounds had cysteine at the second amino acid position (Figure 2). Then, each bead was subjected to photolytic cleavage (365 nm) for 24 h in Eppendorf tubes containing 100 μL of acetonitrile/methanol (4:1) to allow identification of binding library members. The resulting solution, which contained the putative RNA binding library members cleaved from the resin, was analyzed by mass spectrometry.

Synthesis of Ligands Identified by RBDCL Screen. Selected DCL monomers **1**, **2**, **3**, and **4** were synthesized on solid phase using standard FMOC main chain and Boc/Trt side chain protecting chemistry. Briefly, each monomer was synthesized on Wang resin (100−200 mesh size, 1 mmol/g loading, 500 mg, 0.5 mmol). First, the resin was activated through the addition of 1−1′-carbonyl diimidizole (1620 mg, 5 mmol, 10 equiv) in 12 mL of DMF. This suspension was rotated on a LabQuake rotator for 12 h. The vessel was then evacuated and washed three times with 15 mL of DCM. Propane diamine (421 μL, 5 mmol, 10 equiv) was added in 12 mL of DMF and rotated for an additional 12 h. The resin was then washed six times with DCM and six times with DMF. FMOC-Lys(Boc)-OH (702.5 mg, 1.5 mmol, 3 equiv), HBTU (570 mg, 1.5 mmol, 3 equiv), and DIPEA (424 μL, 2.5 mmol, 5 equiv) in 10 mL of DMF were added to each batch of resin, rotated for 1 h, and the resin was washed. Then FMOC was removed (20% piperidine /DMF, 30 min), and resin was washed. FMOC-Cys(Trt)-OH (878 mg, 1.5 mmol, 3 equiv), HBTU (570 mg, 1.5 mmol, 3 equiv), and DIPEA (424 μL, 2.5 mmol, 5 equiv) in 10 mL of DMF were added to each batch of resin, rotated for 1 h, and the resin was washed. FMOC was removed with (20% piperidine /DMF, 30 min), and resin was washed. Then either FMOC-Asn(Trt)-OH (895 mg, 1.5 mmol, 3 equiv) for compounds **1** and **2** or FMOC-Pro-OH (506 mg, 1.5 mmol, 3 equiv) for compounds **3** and **4**, HBTU (570 mg, 1.5 mmol, 3 equiv), and DIPEA (424 μL, 2.5 mmol, 5 equiv) in 10 mL of DMF were added, rotated for 1 h, and the resin was washed. FMOC was removed with (20% piperidine/DMF, 30 min), and resin was washed. Finally, piperonylic acid (250 mg, 1.5 mmol, 3 equiv) for compounds **1** and **3** or 3-carboxy-2-ethyl-3-quinolinium chloride (353 mg, 1.5 mmol, 3 equiv) for compounds **2** and **4**, HBTU (570 mg, 1.5 mmol, 3 equiv), and DIPEA (424 μL, 2.5 mmol, 5 equiv) in 10 mL of DMF were added and rotated for 1 h, and the resin was washed. Final products were cleaved from the resin, and Boc/Trt groups were removed by treatment with 10 mL of a 1% triethyl silane/50% TFA solution in DCM for 2 h. Products were purified by precipitation in chilled ether (−20 °C). Solids were concentrated by centrifugation (2500 rpm, 10 min), the solution was removed, and fresh ether was added. The solution was mixed by vortex, and solids were again concentrated by centrifugation. This series was repeated five times. After the last washing step, the solids were dried by lyophilization, resulting in off white powders.

DCL disulfides **1-1**, **1-2**, **1-3**, **1-4**, **2-2**, **2-3**, **2-4**, **3-3**, **3-4**, and **4-4** were prepared by mixing equimolar amounts of **1**, **2**, **3**, or **4** with **1**, **2**, **3**, or **4** in water and allowing them to undergo oxidative disulfide formation for a period of 7 days. Disulfide formation was monitored by HPLC. When disulfide formation had reached completion, the resulting desired disulfides were separated and purified by preparative reverse-phase HPLC using a 0 to 40% acetonitrile/water (0.1% TFA) gradient.

Fluorescence Titrations. Fluorescence titrations were performed using a Varian Cary Eclipse spectrophotometer. Two microliters of 50 or 500 μM compound was titrated into 400 μL of FAM-labeled RNA sequences and allowed to equilibrate for at least 10 min, or until no change in fluorescence spectra was observed. Changes in fluorescence emission at 518 nm (excitation at 490 nm) were measured. Raw data were corrected for dilution-dependent changes, and Em_{518} was plotted against peptide concentration and fit to the one site binding equation $y = (b_{max} \times x)/(K_D + x)$.

Dynamic Light Scattering. Dynamic light scattering (DLS) data were collected on a DynaPro 801 Molecular Sizing Instrument (Protein Solutions Inc.). All measurements were recorded at 22 °C. The compounds were dissolved in 1× PBS buffer (pH 7.2). Prefiltered samples were injected into a 12 μL cell and illuminated by a 25 mW laser at 750 nm. The detector angle was 90°. The hydrodynamic radius (H_r) was calculated using the dynamic software package version 4.0. Ten repeat measurements were performed on each sample, and the data were filtered based on sum of squares (SOS) and baseline error calculations. Each reported intensity value is an average of three or more measurements.

Evaluation of Small Molecule (CUG)−MBNL1 Inhibition. A DiscoveRx PathHunter ProLabel Enzyme Fragment Complementation assay was employed to test the ability of selected ligands to disrupt CUG repeat RNA−MBNL1 protein interaction in vitro.

To coat 96-well plates with $(CUG)_{109}$, 50 μL of 0.5 μM 5′-biotinylated DNA (5′-TTTTAATTTTAGGATCCCCCCAG-3′) in 50 mM Na_2HPO_4 (pH 8.5) + 1 mM EDTA was immobilized on white streptavidin-coated 96-well plates (Pierce 15502) for 18 h, and the plate was aspirated and washed. To this oligo was hybridized 50 μL of a 0.1 μM solution of a $(CUG)_{109}$ transcript 5′-GGG$(CUG)_{108}$CUGGGGGGAUCCUAAAAUUAAAA)-3′ (underlined sequence is complementary to biotinylated DNA) in buffer A (50 mM tris, pH 8.0, 50 mM NaCl, 50 mM KCl, 1 mM $MgCl_2$) for 2 h, and the plate was aspirated and washed one time with buffer A. This process immobilizes ∼0.5 pmol of $(CUG)_{109}$ RNA per well.

Various concentrations of compounds were then added with or without 40-fold base excess of yeast tRNA in a total volume of 60 μL in buffer A and incubated for 30 min. For competition experiments, 100 pmol of tRNA was added to the binding mixture (∼40-fold base excess, or ∼20-fold molar excess). Then, 10 μL of 1 μM (buffer A + 0.05% tween) recombinant MBNL1 protein containing an N-terminal β-Gal (PL) (PL = ProLabel) enzyme donor sequence (PL-MBNL1 fusion protein) was added, incubated for 1 h, and the plate was aspirated and washed one time with buffer A + 0.05% tween. At this step, any PL-MBNL1 protein bound to the $(CUG)_{109}$ RNA remains bound to the 96-well plate.

The enzyme fragment complementation (EFC) assay utilizes the DiscoveRx PathHunter ProLabel Detection kit. Fifty microliters of 50× diluted solution of the supplied β-Gal enzyme acceptor in PBS pH 7.2 was added and incubated for 1 h. Then, 10 μL of 100× diluted solution of the supplied β-Gal substrate (Galacton + Emerald luminescence enhancer, DiscoveRx) was added and incubated for 1 h.

Only the β-Gal enzyme acceptor bound to the PL-MBNL1 enzyme donor is capable of producing the luminescent product, and as such, any luminescence correlates to the amount of $(CUG)_{109}$ bound MBNL1. Luminescence of each well was then read. Wells without added $(CUG)_{109}$ RNA served as a background measure of nonspecific luminescence and were subtracted from each experiment to yield a 0% bound value. Wells containing no peptide inhibitor added served as 100% bound. Percent bound PL-MBNL1 versus peptide concentration was plotted, and data were fit to a logistic binding model to allow extraction of K_i values.

Acknowledgment. Financial support from Paul D. Wellstone Muscular Dystrophy Cooperative Research Center 08U54NS048843-

05 and NS58345 is gratefully acknowledged. P.C.G. and B.R.M. were supported by NIH T32AR007472. We thank Prof. Joseph E. Wedekind for the use of his dynamic light scattering instrument, and Jolanta Krucinska for assistance with conducting DLS measurements. We also thank Dr. Alice Bergmann (University of Buffalo) for acquiring mass spectrometry data (NSF award CHE0091977).

Supporting Information Available: Full spectral characterization of compounds, filter binding assays, fluorescence titrations, inhibition studies, and experimental details. This material is available free of charge via the Internet at http://pubs.acs.org.

JA804398Y