J|A|C|S ARTICLES

Published on Web 11/11/2009

Controlling the Specificity of Modularly Assembled Small Molecules for RNA via Ligand Module Spacing: Targeting the RNAs That Cause Myotonic Muscular Dystrophy

Melissa M. Lee,[†] Jessica L. Childs-Disney,[‡] Alexei Pushechnikov,[†] Jonathan M. French,[†] Krzysztof Sobczak,[§] Charles A. Thornton,[§] and Matthew D. Disney*,[†]

Department of Chemistry and The Center of Excellence in Bioinformatics and Life Sciences, University at Buffalo, The State University of New York, 657 Natural Sciences Complex, Buffalo, New York 14260, Department of Chemistry and Biochemistry, Canisius College, 2001 Main Street, Buffalo, New York 14208, and Department of Neurology, University of Rochester, 601 Elmwood Avenue, Rochester, New York 14627

Received August 13, 2009; E-mail: mddisney@buffalo.edu

Abstract: Myotonic muscular dystrophy types 1 and 2 (DM1 and DM2, respectively) are caused by expansions of repeating nucleotides in noncoding regions of RNA. In DM1, the expansion is an rCUG triplet repeat, whereas the DM2 expansion is an rCCUG quadruplet repeat. Both RNAs fold into hairpin structures with periodically repeating internal loops separated by two 5′GC/3′CG base pairs. The sizes of the loops, however, are different: the DM1 repeat forms 1 × 1 nucleotide UU loops while the DM2 repeat forms 2 × 2 nucleotide 5′CU/3′UC loops. DM is caused when the expanded repeats bind the RNA splicing regulator Muscleblind-like 1 protein (MBNL1), thus compromising its function. Therefore, one potential therapeutic strategy for these diseases is to prevent MBNL1 from binding the toxic RNA repeats. Previously, we designed nanomolar inhibitors of the DM2−MBNL1 interaction by modularly assembling 6′-*N*-5-hexyonate kanamycin A (**K**) onto a peptoid backbone. The **K** ligand binds the 2 × 2 pyrimidine-rich internal loops found in the DM2 RNA with high affinity. The best compound identified from that study contains three **K** modules separated by four propylamine spacing modules and is 20-fold selective for the DM2 RNA over the DM1 RNA. Because the modularly assembled **K**-containing compounds also bound the DM1 RNA, albeit with lower affinity, and because the loop size is different, we hypothesized that the optimal DM1 RNA binder may display **K** modules separated by a shorter distance. Indeed, here the ideal DM1 RNA binder has only two propylamine spacing modules separating the **K** ligands. Peptoids displaying three and four **K** modules on a peptoid scaffold bind the DM1 RNA with K_d's of 20 nM (3-fold selective for DM1 over DM2) and 4 nM (6-fold selective) and inhibit the RNA−protein interaction with IC_{50}'s of 40 and 7 nM, respectively. Importantly, by coupling the two studies together, we have determined that appropriate spacing can affect binding selectivity by 60-fold (20- × 3-fold). The trimer and tetramer also bind ∼13- and ∼63-fold more tightly to DM1 RNAs than does MBNL1. The modularly assembled compounds are cell permeable and nontoxic as determined by flow cytometry. The results establish that for these two systems: (i) a programmable modular assembly approach can provide synthetic ligands for RNA with affinities and specificities that exceed those of natural proteins; and, (ii) the spacing of ligand modules can be used to tune specificity for one RNA target over another.

Introduction

The importance of noncoding RNAs in gene regulation and cellular expression is increasingly recognized. For example, it is estimated that noncoding pri- and pre-microRNAs account for approximately 1% of human genes yet regulate the expression of at least 10% of all genes.[1] Aberrant expression of microRNAs contributes to disease states such as cancer and viral infections.[2,3] Despite advances in understanding the roles of RNA in biology, comparatively few advances have been made in developing methods to target RNA with small molecules. Such methods could be used to design probes to understand RNA function or to modulate RNA activity. In an effort to

[†] University at Buffalo, The State University of New York.
[‡] Canisius College.
[§] University of Rochester.

(1) John, B.; Enright, A. J.; Aravin, A.; Tuschl, T.; Sander, C.; Marks, D. S. *PLoS Biol.* **2004**, *2*, e363.
(2) Garofalo, M.; Condorelli, G.; Croce, C. M. *Curr. Opin. Pharmacol.* **2008**, *8*, 661–667.
(3) Visone, R.; Croce, C. M. *Am. J. Pathol.* **2009**, *174*, 1131–1138.
(4) Aminova, O.; Paul, D. J.; Childs-Disney, J. L.; Disney, M. D. *Biochemistry* **2008**, *47*, 12670–12679.
(5) Childs-Disney, J. L.; Wu, M.; Pushechnikov, A.; Aminova, O.; Disney, M. D. *ACS Chem. Biol.* **2007**, *2*, 745–754.
(6) Disney, M. D.; Childs-Disney, J. L. *ChemBioChem* **2007**, *8*, 649–656.

10.1021/ja906877y CCC: $40.75

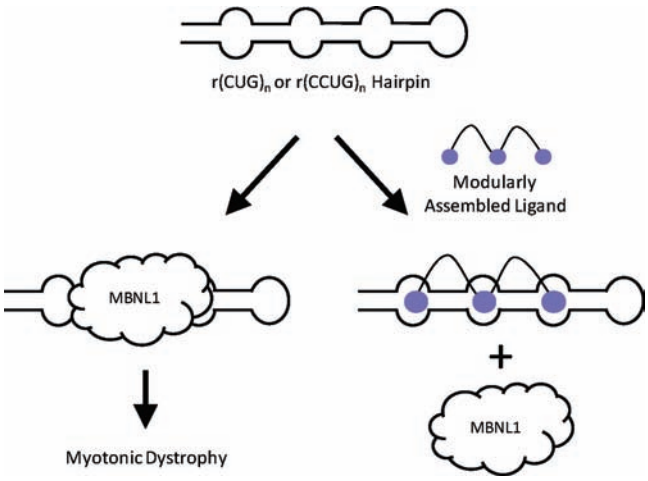

Figure 1. The expanded repeats present in the RNAs that cause myotonic dystrophy (DM) fold into hairpins with regularly repeating mismatches in the stem region. These RNAs sequester the splicing regulator MBNL1, causing misregulation of RNA splicing. A potential therapeutic strategy for DM is displacement of MBNL1 from the repeating tracts by a small molecule ligand.

develop rational methods to target RNA with small molecules, a database of RNA motif-ligand partners is being constructed.[4−7] This database could be mined against genomic sequences and RNA structures to enable a rational and modular approach to target RNA.

Myotonic dystrophy type 1 (DM1) and type 2 (DM2) are examples of human diseases caused by noncoding RNAs. These disorders are caused by expression of tandem repeats in noncoding regions of two mRNAs, resulting in a toxic gain-of-function.[8−10] DM1 is caused by an expansion of dCTG repeats in the 3′ untranslated region of the dystrophia myotonica protein kinase (*DMPK*) gene,[11] while DM2 is caused by an expansion of dCCTG tetranucleotide repeats in intron 1 of the zinc finger 9 protein (*ZNF9*) gene.[9] Despite these mutations being present in functionally unrelated genes, DM1 and DM2 have similar disease manifestations and share a common biochemical mechanism: sequestration of an essential splicing regulator by expanded RNA repeats (Figure 1).[12−14] Specifically, when the toxic dCTG and dCCTG repeats are transcribed, the RNAs fold into hairpins that display 1 × 1 nucleotide UU (DM1) or 2 × 2 nucleotide 5′CU/3′UC (DM2) internal loops (Figure 2A).[15,16] These loops provide high-affinity binding sites for Muscleblind-like 1 protein (MBNL1). Sequestration of MBNL1 by the expanded repeats interferes with its function as a splicing regulator. Because MBNL1 is responsible for the proper splicing of the muscle chloride channel[17] and insulin receptor pre-mRNAs,[18] misregulation of these transcripts leads to the abnormal muscle excitability and insulin insensitivity observed in DM patients.

The underlying biochemical cause of DM also points to a therapeutic strategy in which a small molecule would bind to the toxic repeats and inhibit formation of the DM RNA−MBNL1 complex. If inhibition or displacement were to occur in vivo, there would then be an increase in the amount of free MBNL1, which could correct the splicing defects associated with DM. This hypothesis is supported by a mouse model study in which normal splicing patterns were restored in cells that express expanded rCUG repeats by increasing the cellular concentration of MBNL1.[19] Additionally, it has recently been shown that displacement of MBNL1 from the triplet repeat RNA with an oligonucleotide corrected alternative splicing defects in a mouse myoblast cell line with r(CUG)$_n$ repeats, DM1-affected cells,[20] and in a mouse model.[21]

Previously, we used a modular assembly approach to design nanomolar inhibitors of the DM2−MBNL1 interaction.[22] Ligands were designed based on the observation that 6′-*N*-5-hexynoate kanamycin A (**K**, Figure 2B) bound most tightly to 2 × 2 nucleotide pyrimidine-rich internal loops like those found in the DM2 rCCUG repeat RNA.[5,6] A library of modularly assembled ligands was synthesized in which the valency of **K** and the distance between **K** modules were varied using a peptoid scaffold (Figure 3). The optimal distance between **K** modules, as determined by the potency for disrupting the RNA−MBNL1 complex, was afforded by four propylamine spacers separating **K** modules.[22] A peptoid containing three **K** modules was at least 20-fold specific for the DM2 RNA over other related RNAs.

Herein, we describe our studies to understand how the distance between ligand modules affects RNA binding specificity. We tested the same series of compounds used to identify potent inhibitors of the DM2 RNA−MBNL1 interaction for disruption of the DM1 RNA−MBNL1 complex. Interestingly, the optimal distance between ligand modules is shorter for the DM1 RNA than for DM2 RNA, reflective of the size difference in the respective internal loops. The optimal DM1 ligands are selective for RNAs containing rCUG repeats despite the fact

(7) 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*, 11185–11194.
(8) Mankodi, A.; Logigian, E.; Callahan, L.; McClain, C.; White, R.; Henderson, D.; Krym, M.; Thornton, C. A. *Science* **2000**, *289*, 1769–1773.
(9) Liquori, C. L.; Ricker, K.; Moseley, M. L.; Jacobsen, J. F.; Kress, W.; Naylor, S. L.; Day, J. W.; Ranum, L. P. *Science* **2001**, *293*, 864–867.
(10) Ranum, L. P.; Cooper, T. A. *Annu. Rev. Neurosci.* **2006**, *29*, 259–277.
(11) Brook, J. D.; McCurrach, M. E.; Harley, H. G.; Buckler, A. J.; Church, D.; Aburatani, H.; Hunter, K.; Stanton, V. P.; Thirion, J. P.; Hudson, T.; et al. *Cell* **1992**, *68*, 799–808.
(12) Miller, J. W.; Urbinati, C. R.; Teng-Umnuay, P.; Stenberg, M. G.; Byrne, B. J.; Thornton, C. A.; Swanson, M. S. *EMBO J.* **2000**, *19*, 4439–4448.
(13) Fardaei, M.; Rogers, M. T.; Thorpe, H. M.; Larkin, K.; Hamshere, M. G.; Harper, P. S.; Brook, J. D. *Hum. Mol. Genet.* **2002**, *11*, 805–814.
(14) Mankodi, A.; Teng-Umnuay, P.; Krym, M.; Henderson, D.; Swanson, M.; Thornton, C. A. *Ann. Neurol.* **2003**, *54*, 760–768.
(15) Tian, B.; White, R. J.; Xia, T.; Welle, S.; Turner, D. H.; Mathews, M. B.; Thornton, C. A. *RNA* **2000**, *6*, 79–87.
(16) Sobczak, K.; de Mezer, M.; Michlewski, G.; Krol, J.; Krzyzosiak, W. J. *Nucleic Acids Res.* **2003**, *31*, 5469–5482.
(17) Mankodi, A.; Takahashi, M. P.; Jiang, H.; Beck, C. L.; Bowers, W. J.; Moxley, R. T.; Cannon, S. C.; Thornton, C. A. *Mol. Cell* **2002**, *10*, 35–44.
(18) Dansithong, W.; Paul, S.; Comai, L.; Reddy, S. *J. Biol. Chem.* **2005**, *280*, 5773–5780.
(19) Kanadia, R. N.; Shin, J.; Yuan, Y.; Beattie, S. G.; Wheeler, T. M.; Thornton, C. A.; Swanson, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11748–11753.
(20) Mulders, S. A.; van den Broek, W. J.; Wheeler, T. M.; Croes, H. J.; van Kuik-Romeijn, P.; de Kimpe, S. J.; Furling, D.; Platenburg, G. J.; Gourdon, G.; Thornton, C. A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13915–13920.
(21) Wheeler, T. M.; Sobczak, K.; Lueck, J. D.; Osborne, R. J.; Lin, X.; Dirksen, R. T.; Thornton, C. A. *Science* **2009**, *325*, 336–339.
(22) Lee, M. M.; Pushechnikov, A.; Disney, M. D. *ACS Chem. Biol.* **2009**, *4*, 345–355.

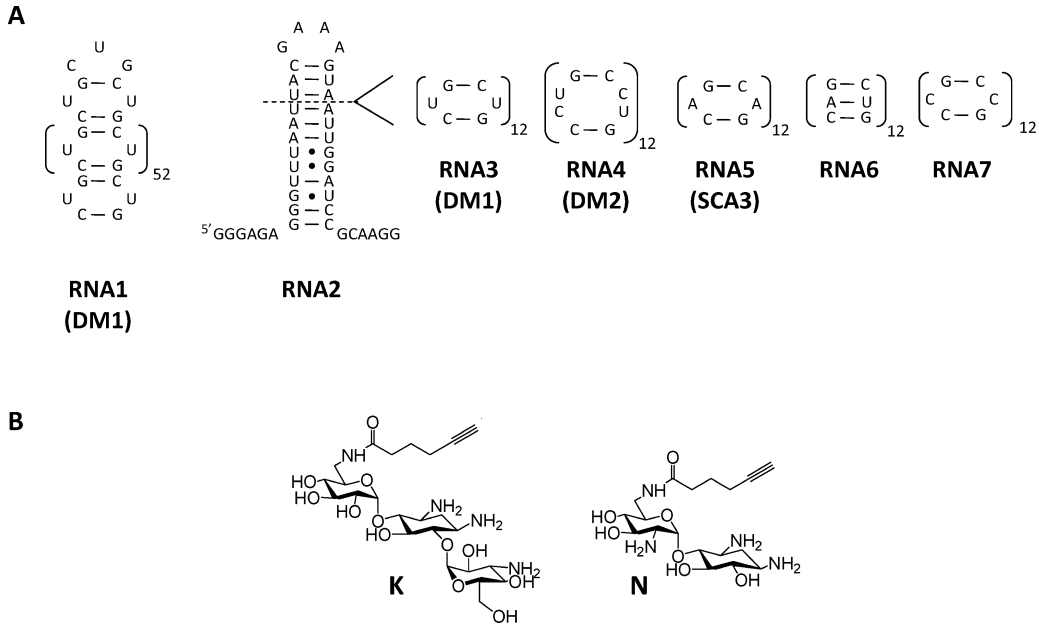

Figure 2. The ligand modules and RNAs used in this study. (A) **RNA1** displays 109 rCUG repeats that fold into a hairpin forming approximately 54 5′CUG/3′GUC internal loops. **RNA2** is the hairpin cassette into which **RNA3**−**RNA7** were inserted. **RNA3** displays 12 copies of the DM1 motif, 5′CUG/3′GUC; **RNA4** contains 12 copies of the DM2 motif, 5′CCUG/3′GUCC; **RNA5** displays the polyQ motif, 5′CAG/3′GAC, and was not expected to bind tightly to **K** or peptoids displaying multiple copies of **K**; **RNA6** is a fully paired RNA control; and **RNA7** displays 12 1 × 1 CC internal loops and is similar to loops selected to bind **K**. (B) The structures of the ligand modules: **K**, 6′-*N*-5-hexynoate kanamycin A; **N**, 6′-*N*-5-hexynoate neamine.

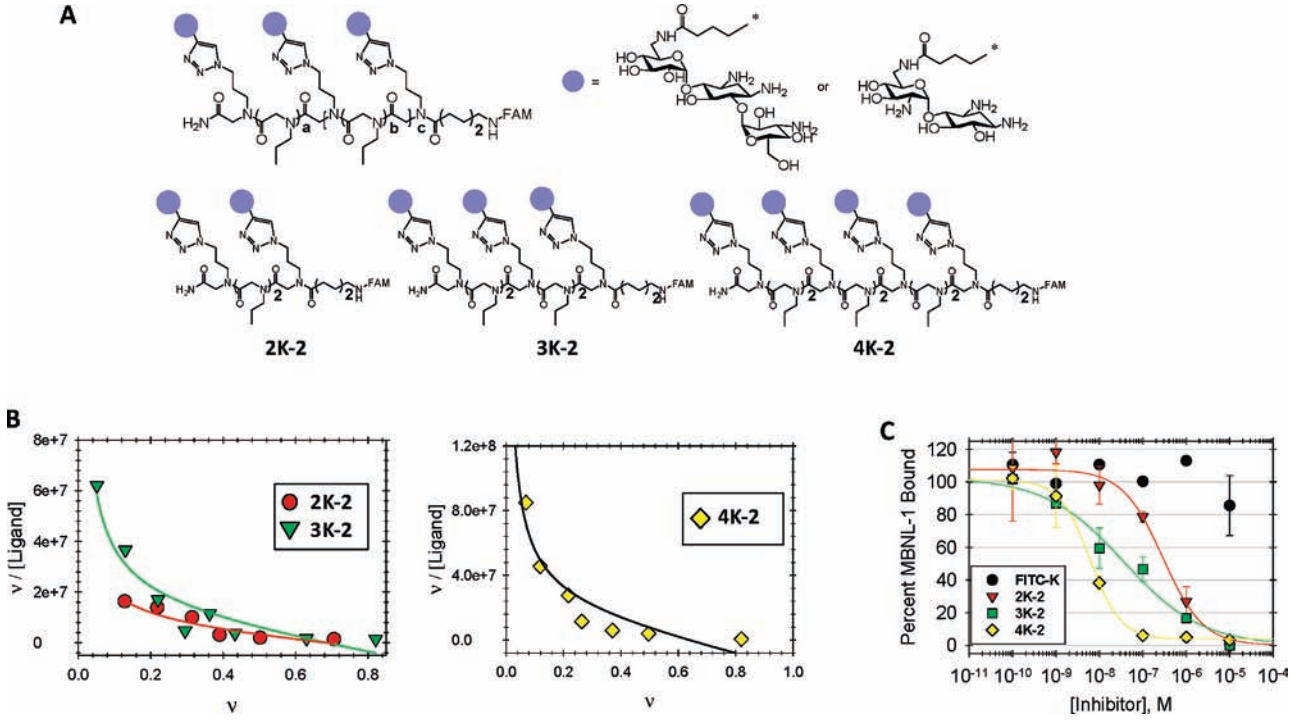

Figure 3. Structures of the peptoids used in these studies and representative inhibition and affinity assays. (A) The general structure of the peptoids studied for inhibition of DM1 RNA−MBNL1 interactions. Structures of the conjugated ligand modules are shown to the right. The most potent dimer, trimer, and tetramer (**2K-2**, **3K-2**, and **4K-2**) are drawn explicitly. The general format for peptoid nomenclature is as follows: ***n*L-*m*** where ***n*** is the valency (**c** + 2), **L** is the ligand module displayed on the peptoid, and ***m*** is the number of propylamine spacers between ligand modules (**a** and **b**). For the ligand module (**L**), **K** indicates the kanamycin derivative, **K**; and **N** indicates the neamine derivative, **N**. FAM indicates coupling of 4(5)-carboxyfluorescein. (B) Representative Scatchard plots from RNA affinity measurements fit to eq 2. (C) Representative plots of MBNL1 inhibition experiments with **RNA1** fit to eq 1.

that the **K** module binds more tightly to the DM2 internal loop. Taken together, the results from our previous study and the results reported herein show that the appropriate spacing of ligand modules can affect selectivity by as much as 60-fold. The optimal DM2 RNA trimer is 20-fold selective for the DM2 RNA over the DM1 RNA[22] while the optimal DM1 RNA trimer is 3-fold selective for the DM1 RNA over the DM2 RNA. These results aid our understanding of how both the identity of the

ligand modules and the spacing between them can be used to control the specific recognition of RNA targets by small molecules.

Experimental Section

General. All solutions were made with diethyl pyrocarbonate (DEPC)-treated, NANOpure water. Oligonucleotides were purchased from Integrated DNA Technologies (IDT).

Synthesis. The syntheses of many of the compounds used in this study have been previously described.[22] Details of synthetic procedures and characterization of new compounds are available in the Supporting Information.

RNA Transcription and Purification. RNAs were transcribed using a Stratagene RNAMaxx transcription kit per the manufacturer's standard protocol and gel purified. **RNA1** was transcribed from the corresponding plasmid[15] digested with *Xba*I. This affords an RNA transcript with a 3′ tail complementary to a DNA probe used in MBNL1 displacement assays. **RNA3**−**RNA7** were transcribed from the PCR products of the corresponding DNA templates.

Expression and Purification of MBNL1. MBNL1 was expressed and purified as described.[22] The expressed protein is fused to the lacZΩ peptide that forms functional β-galactosidase when complemented with soluble lacZα (obtained from DiscoveR$_x$ PathHunter Prolabel Detection Kit).

MBNL1 Displacement Assays. Displacement assays were completed as described[22] in black 384-well plates coated with Streptavidin (Nunc). Resorufin-β-D-galactopyranoside was used as a substrate to quantify the amount of β-galactosidase, and hence MBNL1, present. Fluorescence intensity was measured using a BioTek FLX-800 plate reader. By comparing the fluorescence intensities to wells containing no inhibitor (maximum response) and no RNA (minimum response), the percentage of MBNL1 bound can be determined. The percentage of MBNL1 bound was plotted versus ligand concentration, and the resulting curve was fit to SigmaPlot's four-parameter logistic function to determine the IC_{50} (eq 1):

$$y = D + \frac{A - D}{1 + \left(\frac{x}{IC_{50}}\right)^{\text{Hillslope}}} \quad (1)$$

where y is the percentage of MBNL1 bound, D is the minimum response plateau, A is the maximum response plateau, and x is the concentration of ligand. A and D are typically 100% and 0%, respectively. Each IC_{50} is the average of at least two measurements. To determine the multivalent effect, the IC_{50}'s were normalized for the number of ligand modules conjugated to the peptoid backbone to afford the normalized IC_{50} (NIC_{50}). The NIC_{50} was calculated by multiplying the IC_{50} by the valency. Multivalent effects were calculated by dividing the IC_{50} for **FITC-K** (monomer) by the NIC_{50} of the compound of interest. The number of moles of RNA immobilized in each well was determined using SYBR Green II as described.[23] Approximately ∼20% of the moles of RNA delivered to a well are immobilized.

RNA Binding Assays. The affinities of RNA−ligand complexes were determined as described using a fluorescence emission-based assay.[22] Briefly, RNA was folded in 1X MBNL Buffer (50 mM Tris·HCl, pH 8.0, 50 mM NaCl, 50 mM KCl, 1 mM $MgCl_2$) without $MgCl_2$ by incubating at 60 °C for 5 min followed by slowly cooling to room temperature. Next, $MgCl_2$, BSA, and the ligand of interest were added to final concentrations of 1 mM, 40 μg/mL, and 50 nM, respectively. The RNA was serially diluted in 1X MBNL buffer containing 40 μg/mL BSA and 50 nM ligand and incubated for 1 h at room temperature. Fluorescence intensity was determined using a BioTek FLX-800 plate reader. Scatchard analyses were completed to determine stoichiometry and dissociation constants, accounting for statistical effects by using a functional form of the Scatchard equation for large ligands binding to a lattice (eq 2):[24,25]

$$\frac{v}{[L]} = \frac{N(1 - lv/N)}{k}\left(\frac{1 - lv/N}{1 - (l - 1)v/N}\right)^{l-1} \quad (2)$$

where v is the moles of ligand per moles of RNA lattice, [L] is the concentration of ligand, N is the number of repeating units on the RNA, l is the number of consecutive lattice units occupied by the ligand, and k is the microscopic dissociation constant. This equation simplifies to the commonly used form of the Scatchard equation for simple systems.[24,25]

Cell Culture and Microscopy. The C2C12 (mouse myoblast) cell line was grown as a monolayer at 37 °C and 5% CO_2 in 1X DMEM supplemented with 10% FBS and 0.5% penicillin/streptomycin. For uptake experiments, cells were grown in six-well plates containing sterile glass coverslips. After 14 h of growth, the medium was removed and replaced with fresh medium containing 5 μM compound of interest. The medium was removed after 14 h, and the cells were washed with 1X DPBS (Invitrogen). The coverslip was mounted in 1X DPBS + 50% glycerol, and the cells were imaged using a Zeiss Axio-Imager Z1 Axiophot wide-field fluorescence microscope.

Flow Cytometry Analysis for Uptake and Toxicity. Cell uptake and toxicity of modularly assembled ligands were quantified using flow cytometry. Uptake assays were completed as described above except the ligand of interest was incubated with the cells for 24 or 48 h. The medium was removed, and the cells were trypsinized. After pelleting, the cells were washed with 1X DPBS, centrifuged, and the DPBS removed. The cells were resuspended in 1 mL of ice-cold 1X DPBS and placed on ice. Next, 1 μL of 1.5 mM propidium iodide was incubated with the cells on ice in the dark for 20−30 min. Analysis of 10 000 events was completed using a BD LSR II System Flow cytometer. Quadrants were set using cells that were not treated with ligand (but stained with propidium iodide) to control for autofluorescence in the fluorescein channel. The quadrants assigned ≤1% of the untreated cells as fluorescent.

Results

DM1 and DM2 are caused by tandem repeats of rCUG and rCCUG, respectively. When enough repeats are present, they fold into RNA hairpins with regularly repeating 1 × 1 and 2 × 2 nucleotide internal loops separated by two CG pairs (Figures 1 and 2A). Because the lengths of the repeating RNA sequences and the sizes of the corresponding internal loops are different, we hypothesized that the spacing between ligand modules could control molecular recognition. That is, the specificity of modularly assembled ligands could be tuned for the DM1 RNA by changing the distance between ligand modules. If true, this observation could provide insight about how to control the specificity of modularly assembled ligands that target RNA.

The ligand displayed on the modularly assembled compounds was 6′-*N*-5-hexynoate kanamycin A (**K**, Figure 2B). It was identified as a lead ligand from a previous screen that determined the internal loops that bind with the highest affinity to **K** are 2 × 2 nucleotide, pyrimidine-rich loops.[5,6] These selected internal loops are closely related to those found in DM2 RNAs. It was also determined that **K** binds the DM1 loops albeit with lower affinity.[22]

(23) Pushechnikov, A.; Lee, M. M.; Childs-Disney, J. L.; Sobczak, K.; French, J. M.; Thornton, C. A.; Disney, M. D. *J. Am. Chem. Soc.* **2009**, *131*, 9767–9779.

(24) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry*; W.H. Freeman and Co.: San Francisco, CA, 1980; Vol. 3, pp 849−886.

(25) McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1974**, *86*, 469–489.

Construction of Modularly Assembled Ligands. A library of 18 compounds with different valencies and distances between **K** modules was synthesized previously.[22] Valency was controlled by conjugating different numbers of azide-displaying modules onto which **K** modules were "clicked" via a Huisgen dipolar cycloaddition reaction (HDCR).[26] The distances between ligand modules were varied by conjugating different numbers of propylamine spacers between azides (Figure 3).[22] The nomenclature for the modularly assembled ligands is illustrated in Figure 3A and has the format "***n*L-*m***", where ***n*** is the valency (**c** + 2), **L** is the ligand module displayed on the peptoid, and ***m*** is the number of propylamine spacers between ligand modules (**a** and **b**). For the ligand module, **L**, **K** indicates the conjugation of the kanamycin derivative **K** onto azide-displaying peptoids; and **N** indicates the conjugation of the neamine derivative, **N** (Figure 2B). Modularly assembled ligands displaying **N** modules were synthesized to study the potency of polyamines. The **N** and **K** derivatives have the same number of amines. Our previous study showed that compounds displaying **N** ligand modules bind weakly to DM2 RNAs and are poor inhibitors of the DM2 RNA−MBNL1 interaction.[22] Structures of modularly assembled ligands are shown in Figure 3A.

Potency of Modularly Assembled Ligands for Inhibition of the DM1 RNA−MBNL1 Complex. The series of compounds was tested for inhibition of DM1 RNA−MBNL1 complex formation using two constructs that differ in the number of rCUG motifs displayed. One construct contains 109 rCUG repeats (**RNA1**), while the other was created by inserting 24 rCUG repeats (**RNA3**) into a hairpin cassette (**RNA2**) (Figure 2A). (**RNA2** was chosen because it has been previously shown to bind weakly to **FITC-K**[5] and modularly assembled compounds thereof.[22]) **RNA1** folds into a hairpin with up to 54 5′CUG/3′GUC internal loop motifs while **RNA3** has 12. Previously, these two RNA constructs were shown to bind MBNL1 with similar affinities, suggesting that there is no cooperativity between recognition of adjacent sites on the target RNA.[23]

Potencies (IC_{50}'s) were determined by testing compounds for inhibition of the **RNA1**-MBNL1 complex using a previously reported microtiter plate assay.[22] Initial tests were completed by equilibrating $r(CUG)_{109}$ (**RNA1**, Figure 2A), and the ligand of interest in the well of a microtiter plate for 1 h prior to addition of MBNL1 as previously described.[22] Table 1 summarizes the potency of compounds grouped by valency. In general, as valency increases, potency increases. For example, the **FITC-K** monomer has an IC_{50} of 210 μM, while IC_{50}'s range from 0.3 to ~10 μM for dimers, from 0.04 to 0.8 μM for trimers, and from 0.007 to 0.1 μM for tetramers (Table 1). Furthermore, the most potent ligand in each valency class has two propylamine spacers separating **K** modules, or **2K-2**, **3K-2**, and **4K-2** (Figure 3A and 3C).

The potency of inhibition was further analyzed by normalizing the IC_{50}'s for the number of **K** modules displayed by each inhibitor. These values, or NIC_{50}'s, were calculated by multiplying the IC_{50} by the number of **K** modules present in the compound. The NIC_{50} values range from 0.028 μM for the most potent tetramer to 210 μM for the monomer, **FITC-K** (Table 1). The multivalent effect for each modularly assembled compound was also determined. They were calculated by dividing the NIC_{50} by the IC_{50} for **FITC-K**. Multivalent effects range from ~11 for the least potent dimer (**2K-6**) to 7500 for most potent tetramer (**4K-2**). For each group of inhibitors classified by valency, multivalent effects are the greatest when there are two propylamine spacers between RNA-binding modules (Table 1). These results also show that the most potent modularly assembled compound is 7500 times more potent than the ligand module on a per mole basis of **K**.

Table 1. Inhibition of RNA−MBNL1 Complex Formation Depends on the Valency and Spacing of **K**[a]

valency	inhibitor[b]	IC_{50} (μM)	normalized IC_{50} (NIC_{50}), μM[c]	multivalent effect[d]
		RNA1[e]		
monomer	**FITC-K**	210 ± 21	210	1
dimers	**2K-0**	3.2 ± 0.78	6.4	33
	2K-1	2.8 ± 0.11	5.6	38
	2K-2*	0.3 ± 0.02	0.6	350
	2K-3	0.8 ± 0.55	1.6	131
	2K-4	1.0 ± 0.01	2	105
	2K-6	~10	~20	~11
trimers	**3K-0**	0.8 ± 0.30	2.4	88
	3K-1	0.6 ± 0.13	1.8	117
	3K-2*	0.04 ± 0.03	0.12	1750
	3K-3	0.3 ± 0.18	0.9	233
	3K-4	0.3 ± 0.09	0.9	233
tetramers	**4K-0**	0.1 ± 0.02	0.4	525
	4K-1	0.1 ± 0.01	0.4	525
	4K-2*	0.007 ± 0.003	0.028	7500
		RNA3[f]		
monomer	**FITC-K**	>200	>200	1
dimer	**2K-2***	>10	>20	ND[g]
trimer	**3K-2***	0.7 ± 0.15	2.1	>95
tetramer	**4K-2***	0.06 ± 0.010	0.24	>833

[a] The structures of the most potent dimer, trimer, and tetramer (**2K-2**, **3K-2**, and **4K-2**, respectively) are shown in Figure 3A. [b] The general format for peptoid nomenclature is shown in Figure 3A and described in the text. [c] The NIC_{50} is determined by multiplying the IC_{50} by the number of **K** modules displayed on the peptoid. [d] The multivalent effect is determined by dividing the IC_{50} for **FITC-K** by the NIC_{50} of the corresponding compound. [e] These inhibition assays were completed by adding the ligand to a well containing immobilized **RNA1** for 1 h prior to addition of MBNL1. [f] These inhibition assays were completed by adding the ligand and MBNL1 simultaneously to a well containing **RNA3**. [g] "ND" indicates that no determination could be made due to the nature of the measured values. *These structures are drawn explicitly in Figure 3A.

Control experiments were also completed to determine if the identity of the aminoglycoside displayed on the peptoid is important for inhibition of the **RNA1**−MBNL1 complex. Therefore, a tetramer displaying 6′-*N*-5-hexynoate neamine (**N**, Figures 2 and 3A), **4N-2**, was synthesized. We chose **N** over other aminoglycosides as a control for nonspecific binding because it has the same number of amino groups as **K**. Results show that **4N-2** is a weak inhibitor of the **RNA1**−MBNL1 interaction with an IC_{50} of >2 μM, or >280-fold less potent than **4K-2**. These results show that the **K** module is important for inhibition of DM1 RNA−MBNL1 complex formation.

MBNL1 displacement assays were also completed for DM1 **RNA3** (Figure 2A) and the ***n*K-2** series (Figure 3) under equilibrium conditions in which the ligand and MBNL1 were incubated simultaneously. (**RNA3** is similar to **RNA1** except that it contains only 12 copies of the DM1 motif instead of 54.) These experiments were completed to make comparisons between ligand potency and binding affinity and with previous studies with DM2 **RNA4** (Figure 2A).[22] The IC_{50}'s for **RNA3**−MBNL1 inhibition under equilibrium conditions are >200 μM for **FITC-1**; >10 μM for **2K-2**; 0.7 μM for **3K-2**; and 0.06 μM for **4K-2** (Table 1). (The time required to reach

(26) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2004**, *40*, 2021.

Table 2. Summary of Binding Affinities, Stoichiometries, and Selectivities of Ligands for DM1 RNA as Compared to MBNL1[a]

ligand	RNA3	RNA4	RNA5	RNA6	RNA7
K_d's of Ligand−RNA and MBNL1−RNA Interactions (μM);[b] (Stoichiometries)					
FITC-K	1.0; (11)	0.40; (12)[e]	>1.0	>2.0	0.35; (10.5)
2K-2	0.05; (5.3)	0.12; (4.2)	NM[d]	>0.5	0.06; (4.5)
3K-2	0.02; (3.3)	0.065; (2.3)	NM[d]	>0.5	0.02; (3.1)
4K-2	0.004; (2.6)	0.025; (2.1)	~0.25	>0.5	0.007; (2.4)
MBNL1	0.25[e]	0.12[e]	0.63[e]	>1.0	0.20
Selectivities of **4K-2** and MBNL1 for **RNA3** versus a Second RNA Target[c]					
4K-2		6	63	>140	2
MBNL1		0.5	2.5	>4	0.8

[a] The dissociation constant is reported first followed by the stoichiometry in parentheses when determined. [b] >K_d's indicate that there was either no change in emission for RNA affinity measurements or no retardation through the gel for MBNL1 affinity measurements at the given concentration. All measurements were completed at least in duplicate and have an average standard deviation of ±25%. [c] Selectivities were calculated by dividing the dissociation constant for the indicated RNA by the dissociation constant for **RNA3**. [d] "NM" indicates no measurement was made. [e] Affinities were reported previously.[22]

equilibrium was determined by altering the incubation time; the IC_{50}'s for any particular compound are the same when the equilibration times are 2, 4, and 8 h.)

Affinities, Stoichiometries, and Selectivities of Modularly Assembled Ligands. The binding affinities of **FITC-K**, and the most potent dimer (**2K-2**), trimer (**3K-2**), and tetramer (**4K-2**), were measured using a fluorescence emission assay (Table 2 and Figure 3B). Dissociation constants for these ligands binding to **RNA3** were 1, 0.050, 0.020, and 0.004 μM, respectively. Thus, modular assembly increases the affinity of ligands for binding RNA. Adding one ligand module increases affinity between 2.5- (50 nM for **2K-2** as compared to 20 nM for **3K-2**) and 20-fold (1000 nM for **K** as compared to 50 nM for **2K-2**). The stoichiometries of the RNA−ligand interactions were also determined. **RNA3** binds 11 ± 0.7 **FITC-K** modules; 5.3 ± 0.4 **2K-2** ligands; 3.3 ± 0.1 **3K-2** ligands; and 2.6 ± 0.1 **4K-2** ligands (Table 2). Thus, approximately each ligand module displayed on the peptoid chain interacts with a single 5′CUG/3′GUC motif as expected.[24,25]

The **2K-2**, **3K-2**, and **4K-2** ligands were also tested for binding to an RNA that displays 12 copies of the DM2 motif (**RNA4**) to assess specificity. The **FITC-K**, **2K-2**, **3K-2**, and **4K-2** ligands bind with K_d's of 0.40, 0.120, 0.065, and 0.025 μM, respectively. Thus, the unitary module is 2.5-fold specific for DM2 RNA, but the appropriately spaced, modularly assembled ligands are selective for DM1 rCUG repeats (**RNA3**) over DM2 rCCUG repeats (**RNA4**). The dimeric, trimeric, and tetrameric ligands are approximately 2.5-, 3-, and 6-fold specific for **RNA3** over **RNA4**, respectively. Previously, **3K-4** was shown to be 20-fold specific for **RNA4** over **RNA3**.[22] Thus, these results show that appropriate spacing of ligand modules can be used to program RNA binding specificity and can affect specificity by as much as 60-fold (20-fold × 3-fold).

Additional affinity measurements were completed to explore the selectivities of the ***n*K-2** compounds. As summarized in Table 2, the compounds are selective for **RNA3** over **RNA5** and **RNA6** (Figure 2A). **RNA5** contains 12 1 × 1 nucleotide internal loops containing AA mismatches and is a model system for spinocerebellar ataxia type 3 (SCA3). Recently, it has been shown that rCAG repeats interact with Muscleblind proteins in vivo[27] and contribute to disease pathology.[28] **RNA5** was expected to bind more weakly to compounds displaying **K** modules: (i) because 1 × 1 nucleotide AA loops did not appear in RNA sequences from previous selections with **K**;[5,6] and (ii) because of the relatively weak binding affinity of kanamycin A for the bacterial rRNA A-site, which contains an all-adenine 1 × 2 nucleotide internal loop (K_d = 18 μM).[29] As expected, **4K-2** binds relatively weakly to **RNA5** with a dissociation constant of ~250 nM. **RNA6** is fully base paired; aminoglycosides typically bind fully paired RNAs with dissociation constants in the micromolar range.[5,7,22,30,31] No binding of the tetramer was observed to **RNA6** (up to 500 nM). Thus, the tetramer is ~60-fold and >250-fold selective for the DM1 RNA over **RNA5** and **RNA6**, respectively.

The affinities of the most potent ligands for **RNA7**, which contains 12 1 × 1 nucleotide CC internal loops, were also determined. In our previous study to identify the internal loops preferred by 6′-*N*-5-hexynoate kanamycin A, a 1 × 1 nucleotide CC internal loop was also selected and binds tightly to **K**.[5] Therefore, not surprisingly, the dimer, trimer, and tetramer bind similarly to **RNA7** as they do to **RNA3** with K_d's of 0.060, 0.020, and 0.007 μM, respectively. Thus, the synthetic ligand **4K-2** exhibits only ~2-fold specificity for **RNA3** over **RNA7**. It should be noted that RNAs with regularly repeating base paired motifs are commonly found in the mammalian transcriptome. RNAs with repeats of either 1 × 1 (DM1-like) or 2 × 2 (DM2-like) nucleotide internal loops, however, are not (see http://www.rna.ccbb.utexas.edu/ for the Comparative RNA Web Site Project).[32] Therefore, the RNA that most closely mimics bystander genomic RNAs is **RNA6**. Our designed, modularly assembled ligands bind this RNA only weakly. Thus, the DM1 RNA should be preferred by the modularly assembled ligands over **RNA6** in vivo, reducing the likelihood of side effects.

Correlating Potency and Affinity. In general, as valency of the ***n*K-2** series increases, the affinity of the ligand for **RNA3** increases. For example, the affinity of **4K-2** is 250-fold higher than the affinity of **FITC-K** (Table 2). Potency for inhibiting the RNA−protein complex also increases as valency increases, but by much greater amounts, up to 7500-fold for **4K-2** (Table 1). These observations point to both surface area sequestration and affinity as important factors that affect potency. Sequestering of surface area is an important consideration because it has been shown that each molecule of MBNL1 binds approximately two or three 5′CUG/3′GUC RNA motifs.[33] Thus, effective inhibitors must block as many MBNL1 binding sites in the target RNA as possible to complete with the protein. Such observations also have been found with inhibition of the DM2−MBNL1 interaction.[22]

Comparison of Modularly Assembled Ligands and MBNL1. The affinities of the MBNL1 protein for most of the RNAs shown in Figure 2A are reported in Table 2 and include both new data and previously reported affinities.[22] In all cases,

(27) Ho, T. H.; Savkur, R. S.; Poulos, M. G.; Mancini, M. A.; Swanson, M. S.; Cooper, T. A. *J. Cell Sci.* **2005**, *118*, 2923–2933.
(28) Li, L. B.; Yu, Z.; Teng, X.; Bonini, N. M. *Nature* **2008**, *453*, 1107–1111.
(29) Wong, C. H.; Hendrix, M.; Priestley, E. S.; Greenberg, W. A. *Chem. Biol.* **1998**, *5*, 397–406.
(30) Arya, D. P.; Xue, L.; Willis, B. *J. Am. Chem. Soc.* **2003**, *125*, 10148–10149.
(31) Yoshizawa, S.; Fourmy, D.; Eason, R. G.; Puglisi, J. D. *Biochemistry* **2002**, *41*, 6263–6270.
(32) Gutell, R. R. *Nucleic Acids Res.* **1994**, *22*, 3502–3507.
(33) Warf, M. B.; Berglund, J. A. *RNA* **2007**, *13*, 2238–2251.

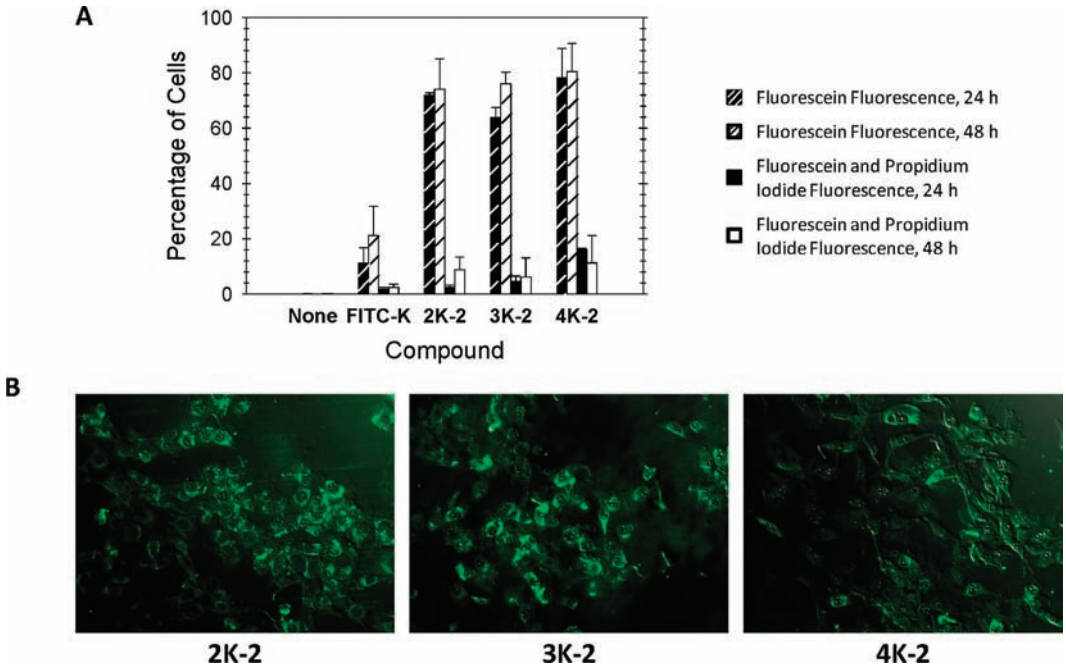

Figure 4. Uptake, toxicity, and cellular localization of **FITC-K** and modularly assembled ligands. (A) Uptake and toxicity of compounds assessed by flow cytometry. The modularly assembled ligands show considerably improved uptake in comparison to **FITC-K**. All compounds exhibit minimal toxicity. (B) Microscopic images depicting localization of compounds in C2C12 cells.

these values are in good agreement with previous reports.[33] In summary, MBNL1 binds with the highest affinity to the RNA displaying the DM2 motif, or **RNA4**, and binds about 2- and 1.7-fold more weakly to **RNA3** and **RNA7**, respectively. Binding is significantly weaker to the SCA3 repeats in **RNA5** (~5-fold) and the fully complementary **RNA6** (>8-fold).[28]

Each modularly assembled ligand with the optimal spacing (**2K-2**, **3K-2**, and **4K-2**) binds with higher affinity to **RNA3** than does MBNL1 (Table 2). For example, **4K-2** binds **RNA3** with ~63-fold higher affinity than does MBNL1. Dimeric **2K-2** and trimeric **3K-2** bind **RNA3** with ~5- and ~13-fold higher affinity, respectively. The modularly assembled ligands are also more specific **RNA3** binders than MBNL1 is. MBNL1 binds with similar affinity to **RNA3** and to **RNA7**. It is only approximately 0.5- and 2.5-fold specific for **RNA3** over **RNA4** and **RNA5**, respectively. In contrast, **4K-2** is 6- and 63-fold specific for **RNA3** over **RNA4** and **RNA5**, respectively. These results further establish that modularly assembled small molecules can be designed to have higher affinities and better selectivities for their RNA targets than natural proteins.

Uptake and Cellular Localization Properties. The uptake and cellular localization of **FITC-K** and the modularly assembled ***n*K-2** ligands in a mouse myoblast cell line (C2C12) were determined by flow cytometry and microscopy (Figure 4 and Supporting Information). The monomer, **FITC-K**, shows poor cellular uptake with low fluorescence intensities. After a 48 h incubation with the compound, only ~20% of the cells have fluorescence due to the presence of **FITC-K**. The amount of ligand taken up by the cells can be quantified by the geometric mean for fluorescence intensity (reported in Relative Fluorescence Units, RFU). This value was then normalized to the fluorescence intensity of untreated cells. The normalized geometric mean for the ligand module is 4.1 ± 0.56 RFU, corresponding to only a 4-fold increase. **FITC-K** localizes mainly in the cytoplasm but also in the nucleus. In contrast, the modularly assembled ligands show improved cell permeability. When treated for 48 h, ~75% of the cells treated with the modularly assembled compounds are fluorescent (3.75 times more than the monomer). The normalized geometric means for fluorescence intensities are also improved: 9.3 ± 1.3 RFU for **2K-2** (2.3-fold increase over **FITC-K**); 11 ± 4.4 RFU for **3K-2** (2.7-fold increase over **FITC-K**); and 13 ± 2.8 RFU for **4K-2** (3.2-fold increase over **FITC-K**). Thus, for **4K-2**, about 4 times as many cells take up ligand in approximately 3-fold higher concentrations than **FITC-K**. Evidently, the peptoid scaffold improves uptake, at least in the mouse myoblast cell line. In addition, **2K-2** and **3K-2** show low levels of toxicity as assessed by propidium iodide staining (<5% of all cells have fluorescence due to both the compound and the propidium iodide). Propidium iodide is a common staining reagent used to detect dead cells. An increase in toxicity was observed for **4K-2**, with ~12% of the cells showing positive staining for propidium iodide and fluorescence due to the compound.

Discussion

The most common method employed to identify ligands that bind RNA is high-throughput screening where an RNA target of interest is screened for binding to members of a chemical library. Although moderately successful, the hit rate for RNA targets is much lower than for protein targets. An alternative "bottom-up" approach is to identify small molecule binders for discrete secondary structure motifs, and then to search for these motifs in larger RNA targets. These interactions can then be used to design modularly assembled ligands. Critical to this approach is determining which features of the modularly assembled compounds can be used to tune RNA-binding affinity and specificity. Such features could include the identity of the ligand module, the relative orientation or spacing of ligand modules, or the identity of the spacing module. In this study, we show that RNA binding specificity can be tuned by the relative spacing of known RNA-binding ligand modules. Our previous study showed that the optimal binder of DM2 RNA,

K, also had some affinity, albeit weaker, to the DM1 RNA.[22] This prompted us to investigate the ability to tune the specificity of our binder from preferring DM2 RNA to preferring the DM1 RNA by varying the spacing between **K**. The overall results show that by modularly assembling an RNA-binding module (**K**) that is biased for binding motifs present in DM2 RNA (**RNA4**), a ligand can be designed that instead specifically recognizes DM1 RNA (**RNA3**). These investigations form a foundation to exploit both optimal and suboptimal RNA motif-ligand partners to design selective modularly assembled RNA binders.

Comparison to Previous Studies Using Modular Assembly. Modular assembly is a general strategy used in biological systems and in drug discovery to enhance binding affinity.[34,35] Factors that are important for the design of high affinity, modularly assembled compounds include the identity of the ligand modules and how they are displayed. One approach that identifies suitable ligand modules is fragment-based screening. A target of interest is screened for binding and analyzed by using either mass spectrometry[36,37] or NMR spectroscopy.[38] Often times the ligand binding sites are also determined. This approach has afforded multivalent compounds with 10- to 1000-fold enhancements in binding affinity over the individual modules.[35]

Modularly assembled compounds have been previously used to target RNA. For example, one study found that neamine dimers bind the bacterial rRNA A-site ~10-fold more tightly than the neamine monomer.[39] Other dimeric aminoglycosides inhibit the Tetrahymena ribozyme and have enhancements in potency ranging from 20- to 1000-fold relative to the monomeric aminoglycoside.[40] Herein, we observe an ~10-fold enhancement in binding for each module. It is interesting to note that the bacterial rRNA A-site mimic is a hairpin with little tertiary structure, which is similar to the RNAs used in our studies. It is likely that optimization of the peptoid backbone or use of other scaffolds that preorganize modules for binding to the RNA target could provide larger enhancements in affinity.[41] This study has shown, however, that the peptoid scaffold can program modularly assembled ligands for specific RNA binding.

Comparison to Other Ligands Targeting DM1 and DM2 RNAs. Previously, a series of modularly assembled ligands displaying **K** modules was used to disrupt the DM2 RNA−MBNL1 complex.[22] The most potent ligands from that study have four propylamine spacing modules between ligand modules. In contrast, the most potent ligands for the DM1 RNA−MBNL1 complex have two propylamine spacing modules. Despite the difference in spacing, the ligands have some similarities. For example, the average increase in affinity observed for adding one ligand module is about 10-fold. In both cases, it has also been shown that the improvement in potency afforded by increasing valency exceeds the enhancement in affinity. This observation illustrates that both the amount of the RNA target's surface area that is sequestered and the affinity are important considerations for the design of potent ligands.

The trimer and tetramer described herein are low nanomolar binders to the DM1 RNA and inhibit formation of the DM1 RNA−MBNL1 complex with nanomolar IC_{50}'s. These studies establish that modular assembly is an effective approach for designing ligands that target DM1 RNAs. Furthermore, the designed ligands are of >10-fold higher affinity and more selective than MBNL1 is. Collectively, these results are encouraging for using this strategy as a general approach to design potent ligands targeting DM1 RNAs and, perhaps, RNA in general.

Influencing Cellular Uptake and Localization Properties. Fortuitously, our modularly assembled compounds are significantly more cell permeable than the monomeric ligand module. As the valency of the small molecule increases, uptake increases. The modularly assembled compounds are taken up by more cells and in higher concentrations than the monomer (Figure 4 and Supporting Information). A series of peptoids that display a Hoechst derivative were previously studied for cell permeability.[23] Similar results were also observed for these compounds in that as valency increased, the percentage of cells that became fluorescent increased. In contrast to the results for the ***n*K-2** series, nuclear localization correlates with valency for the modularly assembled Hoechst compounds. Thus, the identity of the ligand module can alter both uptake and cellular localization. It is likely that the spacing module also affects these properties. Several studies have also shown that appropriately functionalized peptoids improved the uptake properties of cargo to which they were attached.[42,43]

Summary and Outlook

Cellular RNAs largely represent untapped potential as targets for therapeutic agents or for chemical genetics probes to investigate cellular function. In an effort to better exploit this potential, rational methods are being developed to design small molecules that bind RNA using modular assembly strategies. Critical to the effective implementation of this strategy is an understanding of the factors that can be used to control binding affinity and specificity. Herein, we have demonstrated that appropriate display of ligand modules relative to each other can convert a ligand module that is selective for DM2 RNA into a modularly assembled compound that is selective for DM1 RNA. As methods evolve to identify RNA motif-ligand interactions and to link ligand modules together, the design of ligands targeting RNA could be implemented by mining genomic sequences for secondary structure motifs known to bind small molecules. Because of recent advances in the prediction and annotation of RNA structure from sequence,[44−47] computational tools could

(34) Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2755–2794.
(35) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. *Nat. Rev. Drug Discovery* **2004**, *3*, 660–672.
(36) Seth, P. P.; Miyaji, A.; Jefferson, E. A.; Sannes-Lowery, K. A.; Osgood, S. A.; Propp, S. S.; Ranken, R.; Massire, C.; Sampath, R.; Ecker, D. J.; et al. *J. Med. Chem.* **2005**, *48*, 7099–7102.
(37) Swayze, E. E.; Jefferson, E. A.; Sannes-Lowery, K. A.; Blyn, L. B.; Risen, L. M.; Arakawa, S.; Osgood, S. A.; Hofstadler, S. A.; Griffey, R. H. *J. Med. Chem.* **2002**, *45*, 3816–3819.
(38) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531–1534.
(39) Sucheck, S. J.; Wong, A. L.; Koeller, K. M.; Boehr, D. D.; Draker, K.; Sears, P.; Wright, G. D.; Wong, C. H. *J. Am. Chem. Soc.* **2000**, *122*, 5230–5231.
(40) Michael, K.; Wang, H.; Tor, Y. *Bioorg. Med. Chem.* **1999**, *7*, 1361–1371.
(41) Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000**, *403*, 669–672.
(42) Goun, E. A.; Pillow, T. H.; Jones, L. R.; Rothbard, J. B.; Wender, P. A. *ChemBioChem* **2006**, *7*, 1497–1515.
(43) Goun, E. A.; Shinde, R.; Dehnert, K. W.; Adams-Bond, A.; Wender, P. A.; Contag, C. H.; Franc, B. L. *Bioconjugate Chem.* **2006**, *17*, 787–796.
(44) Mathews, D. H. *J. Mol. Biol.* **2006**, *359*, 526–532.
(45) Parisien, M.; Major, F. *Nature* **2008**, *452*, 51–55.

facilitate this approach. If successful, it could allow for the reliable and rational design of small molecules targeting RNA analogous to the design of modularly assembled polyamides that target DNA.[48]

Acknowledgment. We acknowledge the assistance of the Confocal Microscope and Flow Cytometry Facility in the School of Medicine and Biomedical Sciences, University at Buffalo. This work was supported by A NYSTAR JD Watson Award, The New York State Center of Excellence in Bioinformatics and Life Sciences, and the National Institutes of Health (R01-GM079235). M.D.D. is a Cottrell Scholar awarded by the Research Corporation and a Camille and Henry Dreyfus New Faculty Awardee.

Supporting Information Available: Detailed procedures for syntheses of all new compounds used in this study and their characterization; results of flow cytometry analyses including representative histograms and statistics data. This material is available free of charge via the Internet at http://pubs.acs.org.

JA906877Y

(46) Wilkinson, K. A.; Gorelick, R. J.; Vasa, S. M.; Guex, N.; Rein, A.; Mathews, D. H.; Giddings, M. C.; Weeks, K. M. *PLoS Biol.* **2008**, *6*, e96.
(47) Mathews, D. H.; Disney, M. D.; Childs, J. L.; Schroeder, S. J.; Zuker, M.; Turner, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7287–7292.
(48) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.