ARTICLE

ACS chemical biology

Rational and Modular Design of Potent Ligands Targeting the RNA That Causes Myotonic Dystrophy 2

Melissa M. Lee, Alexei Pushechnikov, 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

ABSTRACT Most ligands targeting RNA are identified through screening a therapeutic target for binding members of a ligand library. A potential alternative way to construct RNA binders is through rational design using information about the RNA motifs ligands prefer to bind. Herein, we describe such an approach to design modularly assembled ligands targeting the RNA that causes myotonic dystrophy type 2 (DM2), a currently untreatable disease. A previous study identified that 6′-*N*-5-hexynoate kanamycin A (**1**) prefers to bind 2×2 nucleotide, pyrimidine-rich RNA internal loops. Multiple copies of such loops are found in the RNA hairpin that causes DM2. The **1** ligand was then modularly displayed on a peptoid scaffold with varied number and spacing to target several internal loops simultaneously. Modularly assembled ligands were tested for binding to a series of RNAs and for inhibiting the formation of the toxic DM2 RNA-muscleblind protein (MBNL-1) interaction. The most potent ligand displays three **1** modules, each separated by four spacing submonomers, and inhibits the formation of the RNA-protein complex with an IC_{50} of 25 nM. This ligand has higher affinity and is more specific for binding the DM2 RNA than MBNL-1. It binds the DM2 RNA at least 30 times more tightly than related RNAs and 15-fold more tightly than MBNL-1. A related control peptoid displaying 6′-*N*-5-hexynoate neamine (**2**) is >100-fold less potent at inhibiting the RNA-protein interaction and binds to DM2 RNA >125-fold more weakly. Uptake studies into a mouse myoblast cell line also show that the most potent ligand is cell permeable.

RNA is an important yet underutilized target for small molecule intervention. Most RNA-targeting ligands have been identified *via* high-throughput screening of a ligand library for binding a therapeutically important RNA (*1*). Despite these studies, most RNA targets represent untapped potential for drug or probe design. This is in part due to the limited information on the RNA motif space that prefers to bind ligands and the chemical space of ligands that prefers to bind RNA. To better understand RNA−ligand interactions, we have been screening chemical and RNA spaces simultaneously to construct a database of RNA motif−small molecule partners (*2*–*4*). The goal of the database is to enable the rational and modular design of ligands targeting RNA (Figure 1).

Previously, several related RNA internal loops were found to bind 6′-*N*-5-hexynoate kanamycin A (Figure 1, panel a, **1**) by selecting members of an RNA internal loop library for binding **1** (*2*–*4*). In these studies, pyrimidine-rich 2×2 nucleotide internal loops (5′UU/3′UC or 5′CU/3′UU) bound **1** with the highest affinity (*2, 4*). We therefore searched the literature for disease-causing RNAs that display multiple copies of these or related motifs and found similarities with the rCCUG tetra-repeat that causes DM2. DM2 is caused when expanded repeats form a hairpin that displays multiple copies of a 2×2 5′CU/3′UC internal loop resulting in a toxic gain of function (Figure 2) (*5*). The repeating nature of the DM2 internal loops makes it an ideal system to test a modular assembly approach to target RNA. Modular assembly has been used previously as a strategy to increase the affinity of ligands for binding to a variety of targets (*6*).

The rCCUG repeat expansion that causes DM2 resides in intron 1 of the zinc finger 9 pre-mRNA; its pa-

*Corresponding author, mddisney@buffalo.edu.

Received for review January 31, 2009 and accepted April 4, 2009.

Published online April 6, 2009

10.1021/cb900025w CCC: $40.75

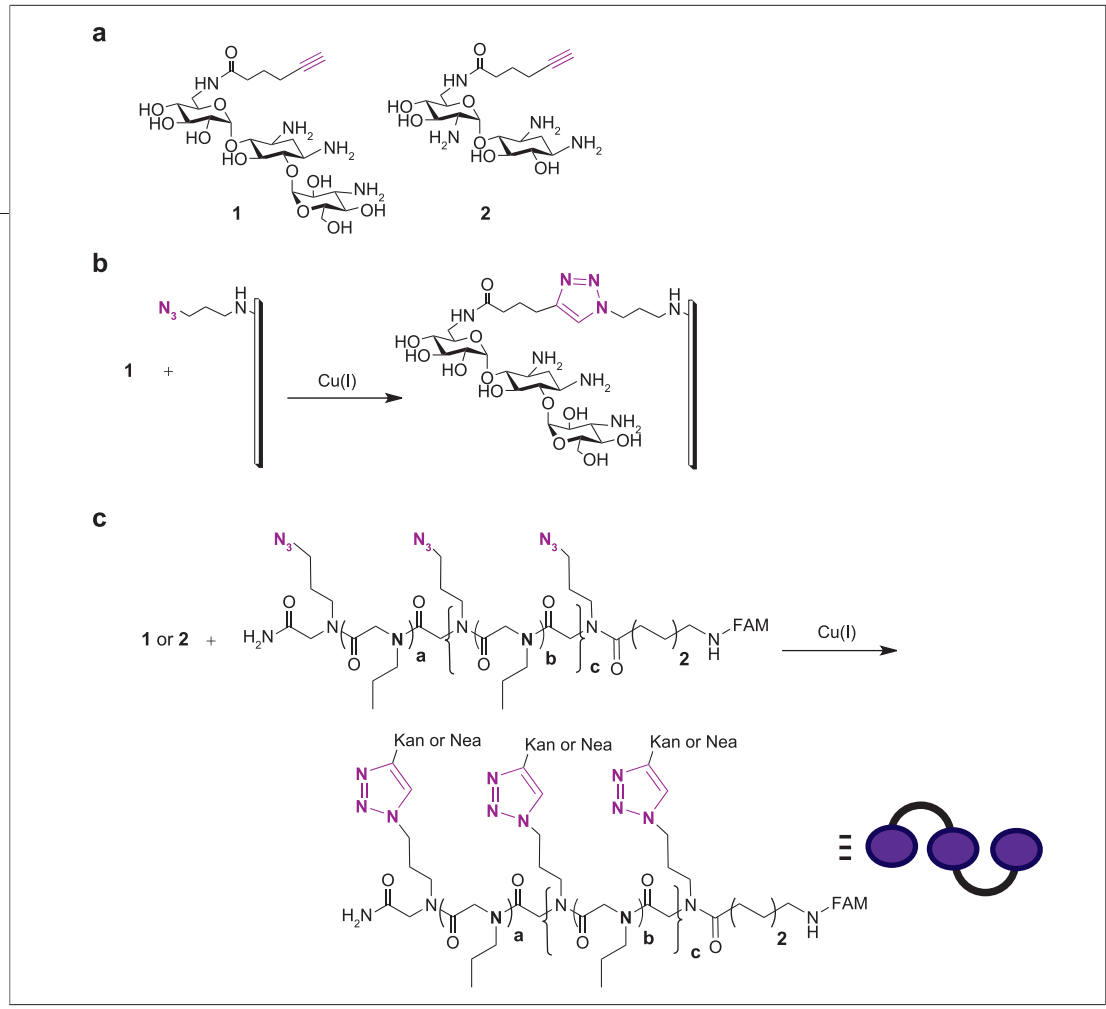

Figure 1. Peptoid modules are assembled using the same chemistry used to anchor them onto arrays to identify the RNA motifs they bind. a) 6′-*N*-5-hexynoate kanamycin A (1) and 6′-*N*-5-hexynoate neamine (2). b) A Huisgen dipolar cycloaddition reaction (HDCR) to immobilize 1 onto a microarray surface to select the RNA motifs that bind 1. c) A HDCR to display 1 or 2 (denoted Kan or Nea, respectively) on azide-functionalized peptoids. The general format for peptoid nomenclature is as follows: *n*L-*m* where *n* is the ligand valency (c + 2), L is the ligand module displayed on the peptoid, and *m* is the number of propylamine submonomers between ligand modules (a and b). The ligand modules that were conjugated to the peptoid are indicated by K, which indicates the kanamycin derivative, 1; N, which indicates the neamine derivative, 2; or P, which refers to propargylamine. Thus, 2K-4 describes a peptoid that displays two 1 modules separated by four spacing modules (a dimer), whereas 3K-4 describes a peptoid that displays three 1 modules each separated by four spacing modules (a trimer). Both of these structures are explicitly drawn in Figure 4. FAM is 4(5)-carboxyfluorescein.

thology is related to myotonic muscular dystrophy type 1 (DM1), which is caused by an expansion of rCUG in the 3′ untranslated region of the dystrophia myotonica protein kinase mRNA (*7*). Both expanded repeats bind the RNA splicing regulator muscleblind (MBNL-1), resulting in the alternative splicing of the main muscle chloride channel and the insulin receptor (*8*). Importantly, this disease model points to a therapeutic approach for treating DM: designing compounds that bind toxic RNA hairpins and inhibit MBNL-1 binding (Figure 2). In this report, we describe the first series of designed nanomolar inhibitors.

RESULTS AND DISCUSSION

Design of Compounds and RNA Constructs. From previous studies, it was determined that **1** prefers 2×2 nucleotide, pyrimidine-rich internal loops (*2*, *4*). Similar loops were identified in the RNA that causes myotonic muscular dystrophy type 2 (DM2) (*5*). In order to study ligands designed to bind the DM2 RNA and inhibit the disease-causing RNA−MBNL-1 complex, a model system was constructed. We inserted 1 or 12 copies of the internal loops present in long rCCUG repeats (5′CCUG/3′GUCC) into a hairpin cassette (*4*, *9*) to afford **RNA2** and **RNA7** (Figure 3). **RNA7** contains enough tetra-repeats so that binding affinity to MBNL-1 is the same as binding to $rCCUG_{47}$ (Supporting Information and ref 10). RNAs displaying other internal loops (Figure 3) were also designed and synthesized in order to assess the specificity of the best ligand for inhibiting the **RNA7**−MBNL-1 complex. For example, **RNA5** and **RNA12** contain 1×1 nucleotide UU internal loops (the DM1 motif),

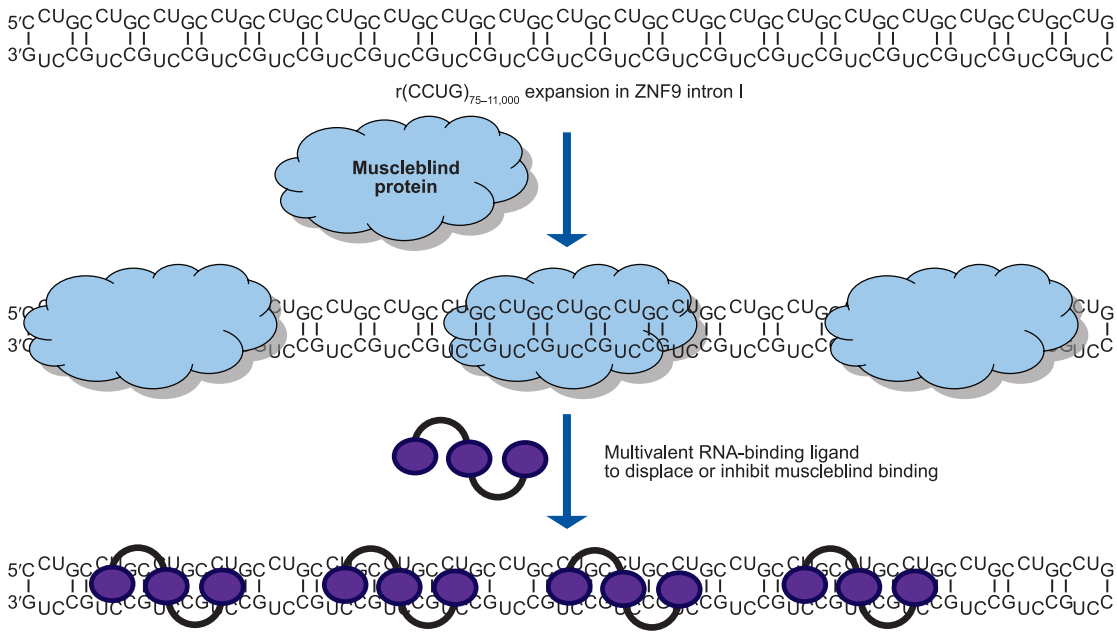

Figure 2. Schematic of the interaction of MBNL-1 with the DM2 rCCUG repeat hairpin and the use of modularly assembled ligands to inhibit this complex.

whereas **RNA6** and **RNA13** contain 1×1 nucleotide AA internal loops (the polyQ motif). **RNA3** and **RNA10** contain 2×2 5′CC/3′CC internal loops; **RNA4** and **RNA11** contain 2×2 5′UU/3′UU internal loops. **RNA8** and **RNA9** are fully base-paired RNA controls.

Modularly assembled ligands displaying **1** were synthesized by anchoring **1** onto azide-functionalized peptoids using a Huisgen dipolar cycloaddition reaction, a variant of "click chemistry" (HDCR, Figure 1, panels b and c) (*11*). This approach is attractive because a HDCR was used to anchor ligands onto microarray surfaces to identify and study the RNA motif-small molecule partners in our database (*2–4*). By using the same chemistry to anchor ligand modules onto a surface and to modularly assemble them, deleterious effects that conjugation could have on molecular recognition may be minimized. It should be noted that a HDCR was previously used to synthesize a library of 2-deoxystreptamine dimers as size-specific ligands for RNA hairpin loops (*12*).

Peptoid synthesis allows for precise control of ligand valency and spacing (*11*). Valency is controlled by the number of 3-azidopropylamine submonomers used to conjugate **1**, while spacing is controlled by inserting propylamine submonomers between azides (Figure 1, panel c). Each of the peptoids is named using the nomenclature described in Figure 1. The general format for peptoid nomenclature is as follows: *n*L-*m* where *n* is the ligand valency (**c** + 2, Figure 1), L is the ligand module displayed on the peptoid, and *m* is the number of propylamine submonomers between ligand modules (**a** and **b**, Figure 1). The ligand modules that were conjugated to the peptoid are indicated by K, which indicates the kanamycin derivative **1**; N, which indicates the neamine derivative **2**; or P, which refers to propargylamine. Thus, 2K-4 describes a peptoid that displays two **1** modules separated by four spacing modules (a dimer), and 3K-4 describes a peptoid that displays three **1** modules each separated by four spacing modules (a trimer). The full structures of dimer 2K-4 and trimer 3K-4 are shown in Figure 4. We synthesized a library of dimers and trimers with varied spacing in order to identify the optimal distance between ligand modules such that two or three adjacent loops can be targeted simultaneously.

Modularly Assembled Ligands Inhibit the DM2 RNA—MBNL-1 Interaction *in Vitro*. The library of ligands synthesized using this approach was initially studied for inhibiting formation of the **RNA7**—MBNL-1 complex

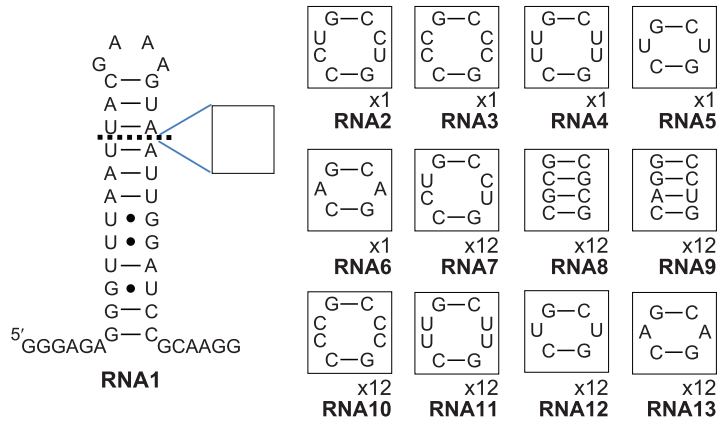

Figure 3. RNAs used to study RNA−ligand interactions. Boxed nucleotides shown to the right were inserted into RNA1. RNA2−RNA6 **contain single copies of an internal loop motif.** RNA7−RNA13 **contain 12 copies of a motif;** RNA2 **and** RNA7 **contain the DM2 motif;** RNA5 **and** RNA12 **contain the DM1 motif;** RNA6 **and** RNA13 **contain the polyQ motif;** RNA3 **and** RNA10 **contain 2×2 CC internal loops while** RNA4 **and** RNA11 **contain 2×2 UU internal loops;** RNA8 **and** RNA9 **are fully paired RNA controls.**

(Figure 2 and Table 1). The IC_{50}'s ranged from 100 μM for **FITC-1** to 1.6 nM for trivalent 3K-4. (**FITC-1** was synthesized previously (*2*) by conjugating 3-azidopropylamine to Boc-protected **1** *via* a HDCR. After reaction with fluorescein isothiocyanate, the compound was deprotected to afford **FITC-1** (structure available in the Supporting Information). Thus, **FITC-1** contains the triazole moiety that is also present in the multivalent ligands.) The range of IC_{50}'s for the dimers was 0.089−10 μM, and the range for the trimers was 0.0016−0.03 μM. The most potent dimer and trimer had four propylamine spacers separating the modules, or 2K-4 and 3K-4, and had IC_{50}'s of 80 and 1.6 nM, respectively (Figure 4). These results indicate that four propylamine spacing units provide the appropriate spacing to target adjacent loops in the DM2 hairpin. Two control trimeric peptoids with the same spacing as 3K-4 were synthesized and tested to confirm that functionalization with **1** is required for potency. The peptoid that displays propargylamine modules (3P-4) was used to determine if functionalization with a triazole moiety is sufficient for potency. The second peptoid, 3N-4, displays 6′-*N*-5-hexynoate neamine (**2**, Figure 1) and determines whether any aminoglycoside with the same number of amino groups can be used to inhibit the RNA−protein interaction. Neither compound inhibited the formation of the **RNA7**−MBNL-1 complex whatsoever at concentrations of <1 μM, indicating that display of **1** modules is required for inhibition (*13*).

To assess the enhancement in potency that modular assembly provides, the magnitude of the multivalent effects were calculated by normalizing the IC_{50}'s for the number of RNA binding modules displayed on the peptoid. These values were then compared to the IC_{50} for **FITC-1**. When the ligands were allowed to pre-equilibrate with the RNA and then MBNL-1 added, multivalent effects ranged from 8.5 to 20,000 with the trivalent ligands giving the largest enhancements. When 2K-4 or 3K-4 and MBNL-1 were added to DM2 RNA simultaneously, multivalent effects were >420 and >6700, respectively; these values are lower estimates because the IC_{50} for inhibition of the **RNA7**−MBNL-1 complex by **FITC-1** was >500 μM under these conditions (Table 1).

Modularly Assembled Ligands Bind Tightly and Specifically to DM2 RNA. Affinities and stoichiometries of RNA−ligand complexes were measured for a series of RNAs (**RNA1**−**RNA13** and bulk yeast tRNA, Figure 3 and Table 2) to **FITC-1** (*2*), 2K-4, 3K-4, and 3N-4 using a fluorescence emission assay. These RNAs were selected to probe the effect of the number of repeating motifs in the RNAs and the identity of the repeating unit. Bulk yeast tRNA was used because it is a mimic of a common bystander RNA present in cells. It should be noted that RNAs with regularly repeating base-paired motifs are commonly found in genomic RNAs; RNAs with repeats of either DM1- or DM2-like 1×1 or 2×2 nucleotide internal loops, however, are not (*14*) (see http://www.rna.ccbb.utexas.edu/ for the Comparative RNA Web Site Project).

Results show that **FITC-1**, 2K-4, and 3K-4 bind more weakly to bulk yeast tRNA and control RNAs that do not display the DM2 motif than **RNA2** or **RNA7**. Each of the three ligands form 1:1 complexes with **RNA2**, with K_d's ranging from 100−450 nM (Table 2). Increasing ligand valency increases affinity to **RNA7** more; **FITC-1**, 2K-4, and 3K-4 bind with K_d's of 400, 50, and 8 nM, respectively (Figure 4 and Table 2). The 6.25- to 8-fold increase in affinity that each RNA binding module provides is similar to enhancements observed when dimerized neamines were studied for binding a bacterial rRNA A-site mimic (*15*). Stoichiometries suggest that each ligand module displayed interacts with a single DM2 mo-

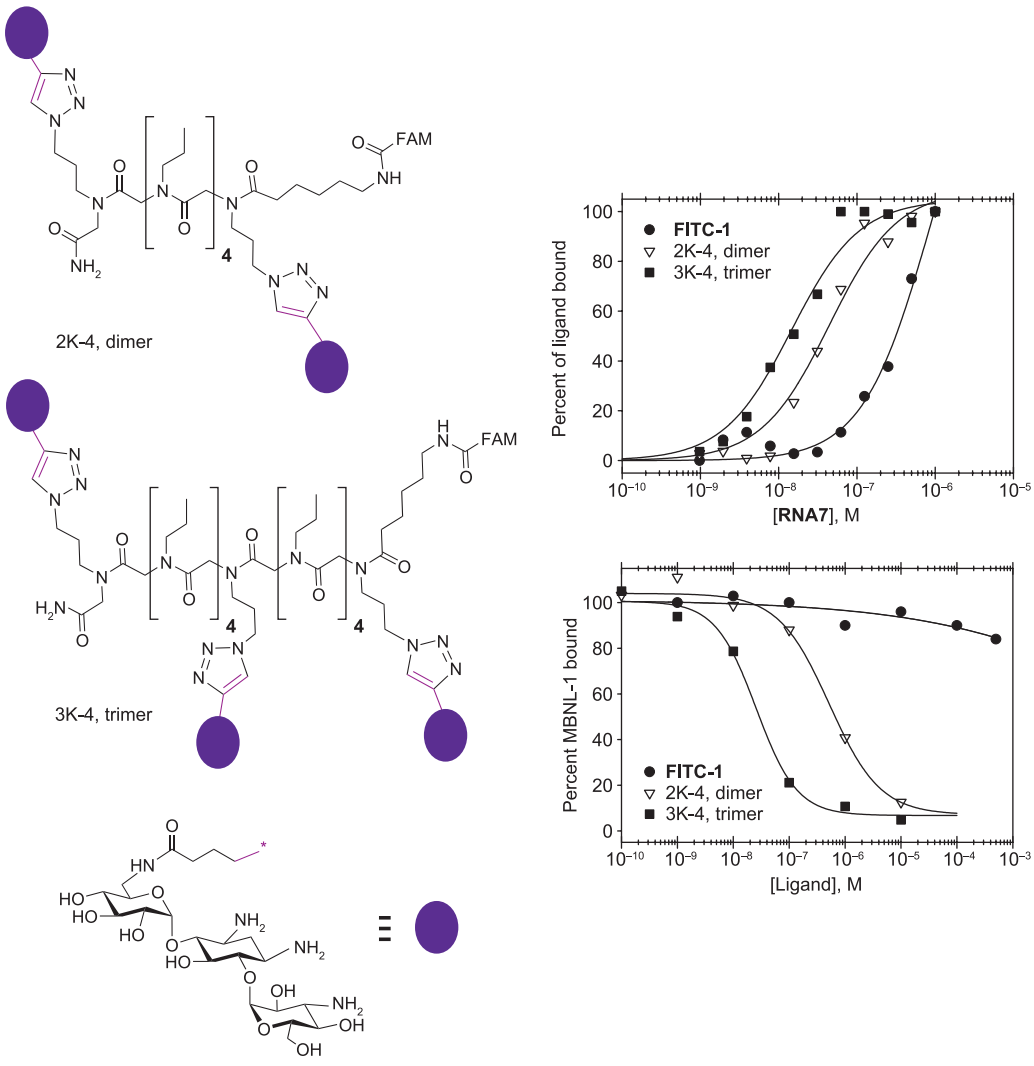

Figure 4. The primary data for and structures of the most potent dimer and trimer that target DM2 RNAs. Left) Structures of the most potent dimer 2K-4 and trimer 3K-4 that were found to inhibit the DM2−MBNL-1 interaction. Right, top) Titration of RNA7 **into a solution of** FITC-1**, 2K-4, and 3K-4. Lines connecting points are drawn for clarity. Further analysis of the binding data using eq 4 (see Methods) is given in Supporting Information. Right, bottom) Plots of data for** FITC-1**, 2K-4, and 3K-4 inhibition of DM2−MBNL-1 interactions; these data are from experiments when MBNL-1 and ligand are added simultaneously to a well containing DM2 RNA. Curves were fit to eq 1 as described in Methods.**

tif in the target RNA. For example, **FITC-1** forms a 12:1 complex with **RNA7**, whereas 2K-4 and 3K-4 form approximately 6:1 and 4:1 complexes, respectively.

To assess binding specificity, 3K-4 was tested for binding to **RNA8−RNA13**, which each displays 12 copies of a repeating motif (Table 2). Results show that this ligand binds with ≥20-fold specificity for **RNA7**. Specificity is the lowest for other closely related loops: **RNA10** that has 2×2 5′CC/3′CC loops (30-fold) and **RNA12** that displays the 1×1 UU DM1 motif (20-fold). Specificities for the other RNAs range from 110-fold to >200-fold. Additionally, 3N-4 was tested for binding to **RNA7** and **RNA8** (fully GC paired stem). It was found to bind with K_d's of >1000 and 190 nM, respectively. The speci-

TABLE 1. Inhibition of formation of the DM2 (RNA7)—MBNL-1 complex by multivalent ligands with different valencies of and distances between RNA binding modules[a]

Monomers				IC_{50}
1[b]				220
FITC-1[b]				100
FITC-1[c]				>500
Dimers				
(*n*L-*m*)[d]	IC_{50}	NIC_{50}[e]	MV[f]	
2K-0[b]	2.1	4.2	24	
2K-1[b]	5.9	11.8	8.5	
2K-2[b]	1.6	3.2	31	
2K-3[b]	0.43	0.86	120	
2K-4[b]	0.089	0.18	560	
2K-6[b]	1.6	3.2	31	
2K-8[b]	4.1	8.2	12	
2K-10[b]	5.9	12	8.5	
2K-12[b]	10	20	5	
2K-19[b]	4.4	8.8	11	
2K-4[c]	0.59	1.2	>420	
Trimers				
3K-0[b]	0.03	0.08	1200	
3K-1[b]	0.01	0.03	3300	
3K-2[b]	0.017	0.051	2000	
3K-3[b]	0.0020	0.006	17000	
3K-4[b,g]	0.0016	0.0048	20000	
3K-4[c]	0.025	0.075	>6700	

[a]All IC_{50} values are in μM. The average error in the IC_{50}'s is $\leq \pm 40\%$. [b]These inhibition assays were completed by adding the ligand to a well containing immobilized **RNA7** and then adding MBNL-1. [c]These inhibition assays were completed by adding the ligand and MBNL-1 simultaneously to a well containing **RNA7**. Data were for experiments completed with a 2 h incubation of the ligand and MBNL-1 with the immobilized RNA; longer incubations (4 or 8 h) resulted in the same IC_{50}'s, indicating that this is an equilibrium measurement. [d]Nomenclature describing the multivalent ligands is shown in Figure 1, and the structures of 2K-4 and 3K-4 are shown in Figure 4. [e]NIC_{50} is the IC_{50} normalized for the number of ligands (eq 2). [f]MV is the multivalent effect, which is calculated by dividing the IC_{50} for **FITC-1** by NIC_{50} (eq 3). [g]The 3P-4 and 3N-4 peptoids did not inhibit the formation of the DM2—MBNL-1 complex whatsoever at concentrations <1 μM.

ficity of 3N-4 to RNAs containing a continuous stretch of GC pairs is not unexpected because **2** was previously identified to bind RNA hairpins and duplexes with continuous stretches of 5′GC and 5′CG steps. The results described herein suggest that **2** could be a useful module to target stretches of GC pairs in RNA (*16*). Thus, 3K-4 is highly specific for binding **RNA7**.

Modularly Assembled Ligands Are Higher Affinity and More Specific DM2 RNA Binders than MBNL-1. To compare the affinity and specificity of 3K-4 to the *in vivo* ligand, MBNL-1, K_d's for the binding of MBNL-1 to a subset of the RNAs shown in Figure 3 were determined (Table 2). Results are in good agreement with previously published values (*10*). MBNL-1 binds most tightly to the RNA containing the DM2 motif, **RNA7**, with a K_d of 120 nM; the second highest affinity binder is **RNA12**, which contains 12 copies of the DM1 motif (250 nM). **RNA13** contains 12 copies of the polyQ motif (5′CAG/3′GAC) and has been recently implicated in binding MBNL-1 and causing spinocerebellar ataxia type 3 (*17*). MBNL-1 binds **RNA13** with a K_d of 630 nM. Weaker interactions are observed with RNAs that do not contain repeats implicated in binding MBNL-1 *in vivo*, including **RNA8** and **RNA9**, which are fully paired, and **RNA10** and **RNA11**, which contain 2×2 nucleotide 5′CC/3′CC and 5′UU/3′UU internal loops, respectively. These RNAs have K_d's that range from 920 to >2500 nM (Table 2).

The binding data for MBNL-1—RNA and multivalent ligand—RNA complexes were compared to quantify if 2K-4 and 3K-4 (Figure 3) are higher affinity and more specific for binding to **RNA7** than MBNL-1. Both 2K-4 (2.3 kDa) and 3K-4 (3.4 kDa) are higher affinity for **RNA7** than MBNL-1 (40 kDa), by 2-fold and 15-fold, respectively (Table 2). The trimer, 3K-4, is also more specific

TABLE 2. Summary of binding affinity, stoichiometry, and specificity of MBNL-1 and the ligands used in these studies[a]

	Affinities of RNA-ligand complexes												
Ligand	**RNA1**	**RNA2**  x1	**RNA3**  x1	**RNA4**  x1	**RNA5**  x1	**RNA6**  x1	**RNA7**  x12	**RNA8**  x12	**RNA9**  x12	**RNA10**  x12	**RNA11**  x12	**RNA12**  x12	**RNA13**  x12
FITC-1	>2.0	0.45; 1	1.0; 1	2.5; 1	1.2; 1	>2.5	0.4; 12						
2K-4	>2.0	0.15; 1					0.05; 6						
3K-4	>2.0	0.10; 1					0.008; 4	>2.0	>2.0	0.27; 4	0.90; 4	0.21; 4	>1.0
FITC-2	>2.0	>1.0; -					>1.0						
3N-4							>1.0	0.19; 4					
MBNL-1[b]	>2.5						0.12	1.0	>2.5	0.92	>2.5	0.25	0.63
	Specificities of 3K-4 and MBNL-1 for RNA7 **over** RNA1 **and** RNA8–RNA13												
3K-4	>200							>200	>200	30	110	30	>100
MBNL-1	>20							8.3	>20	7.7	>20	2.1	5.3

[a]Dissociation constants are reported in μM. Stoichiometries are reported after the binding affinity following the semicolon. Specificity was calculated by dividing the dissociation constant for indicated RNA by the dissociation constant for **RNA7**. Average errors in the K_d's are ≤ ±20%; K_d's reported with a > indicate that there was either no change in emission for RNA affinity measurements or no retardation through the gel for MBNL-1 affinity measurements at the given concentration. Each aminoglycoside ligand bound total yeast tRNA with a K_d > 10 μM. [b]The stoichiometries of the RNA–MBNL-1 complexes can be estimated as 6 for DM1 **RNA12** and ≥6 for DM2 **RNA7** based on a previous report (*10*) that determined MBNL-1 binds to every ≤6 base pairs in RNA.

for binding to **RNA7** than MBNL-1. For example, 3K-4 is 30-fold and >100-fold specific for binding **RNA7** over **RNA12** and **RNA13**, respectively, whereas MBNL-1 is only 2.1- and 5.3-fold specific (Table 2). Thus, despite 3K-4 having an ~12-fold lower molecular weight compared with that of MBNL-1, it is ~10-fold more specific and has 15-fold higher affinity for binding **RNA7**.

Increases in Affinity of Modularly Assembled Ligands Do Not Totally Account for Increases in Potency for Inhibition of the DM2 RNA–MBNL-1 Complex. The enhancement in binding affinity afforded by modularly assembled ligands (2K-4 and 3K-4) over the monomer (**FITC-1**) were then compared to multivalent effects observed for inhibition of the DM2 **RNA7**–MBNL-1 interaction. This comparison is only made under equilibrium conditions when the ligand and MBNL-1 were added to **RNA7** simultaneously (Table 1). Different incubation times (1, 2, 4, and 8 h) were used in order to determine when the system reached equilibrium so that multivalent effects could be correlated with binding affinities. The IC_{50}'s are the same when 2, 4, and 8 h incubation times are used, suggesting that a 2 h incubation is sufficient for the system to reach equilibrium. Results show that multivalency increases the inhibitory potency of 2K-4 and 3K-4 by >420- and >6700-fold, respectively. Affinity comparisons show that 2K-4 binds 8 times more tightly to **RNA7** than **FITC-1** and 3K-4 binds 50 times more tightly (Tables 1 and 2). Thus, increases in affinity do not account for the enhanced potency of the ligands. These results suggest that surface area effects may play a significant role in enhancing potency. Sequestration of surface area on the target RNA by modularly assembled ligands is enhanced because the peptoid linker used to conjugate modules together can also act as a steric block for MBNL-1 to bind to the target RNA.

Comparison to Other Studies That Use Modular Assembly and Multivalency To Target Biomolecules. Modular assembly or multivalency has been used as a strategy to design inhibitors for a variety of targets (*6,*

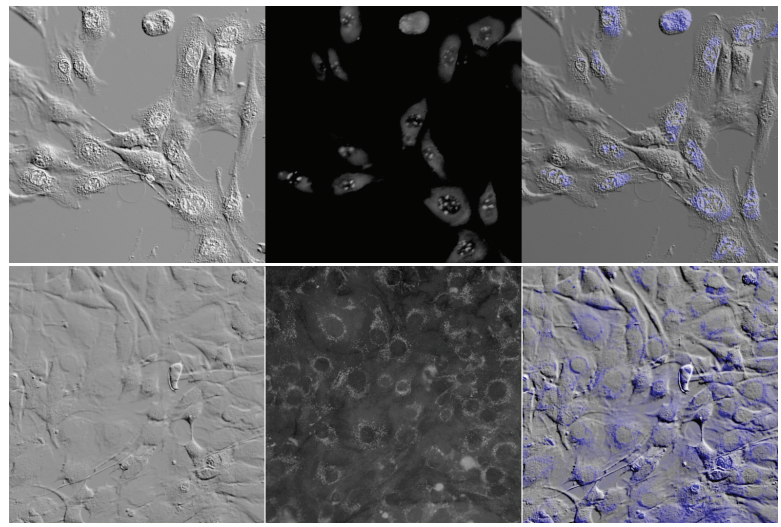

Figure 5. Uptake of 2K-4 (top) and 3K-4 (bottom) into the mouse myoblast C2C12 cell line: left) phase contrast image; middle) fluorescence image of the cells shown in the left panel; right) overlay of the two images where fluorescence from the ligand is shown in purple.

18–20). For example, STARFISH ligands displaying multiple copies of a weakly binding trisaccharide (millimolar K_d) have been developed to inhibit the Shiga-like toxins (*21*). Multivalent enhancements can be as high as 1,000,000-fold when 10 copies of the sugar are displayed on the scaffold. Such values of multivalent enhancements are some of the highest reported. Structural analysis of the ligand−Shiga-like toxin complex revealed that the STARFISH scaffold preorganized the multivalent ligands for binding to the target protein. More typical multivalent enhancements for display of ligands on random polymers are ~1,000-fold (*6*). Thus, the multivalent enhancement of >6700 for 3K-4 trimer observed herein are higher than those observed with randomly displayed polymers and most likely not as high as the designed and preorganized STARFISH ligands.

Fragment-based assembly typically uses monomers that are higher affinity binders (micromolar K_d's) to their target than the sugar monomers described above for STARFISH ligands. These assembly methods have been used in the SAR (structure−activity relationship) by NMR approach (*20*), in screening of small chemical libraries (*22, 23*), and in the modular design of polyamides targeting the DNA minor groove (*24*). In these cases, smaller numbers of ligands, dimers, for example, are most commonly used. For all of these studies, it has been shown that preorganization of the ligand modules for binding to the target has also been a critical consideration in the design of potent ligands. The peptoid scaffold used to display the RNA binding modules in this study is known to have little, if any, defined structure. Peptoids can be designed to have helical structures if chiral and/or bulky side chains are present in the molecule, but these functionalities are not part of our spacing modules (*25*). Future studies will therefore be completed to optimize the display of ligand modules in order to preorganize them to bind the DM2 RNA. Such rational design of spacing modules awaits the result of high resolution structures of these ligands bound to RNA.

Modularly Assembled Ligands Are Permeable to a Mouse Myoblast Cell Line. Both 2K-4 and 3K-4 were tested for uptake into a mouse myoblast cell line (C2C12) when simply added to the culture medium (Figure 5). After washing, cells were imaged using fluorescence microscopy, and images show that 2K-4 and 3K-4 enter mouse myoblasts without the aid of a transfection agent or vehicle. Compound 2K-4 is nucleolar and cytoplasmic, whereas 3K-4 is perinuclear with some cytoplasmic and nuclear staining. Propidium iodide staining was excluded from cells that are fluorescent due to uptake of either ligand, indicating that they are nontoxic. Previous studies have shown that guanidinylated aminoglycosides enter mammalian cells and are nontoxic (*26*).

One issue with using a modular assembly approach is that molecular weight increases as ligand valency increases. High molecular weight ligands, including oligonucleotides and oligonucleotide derivatives such as PNAs, are often not cell permeable (*27, 28*). The positive uptake results observed with these ligands may suggest that they or other related ligands could have activity in cell culture or mouse models of DM (*5, 7, 29*). Furthermore, cellular uptake and localization properties could be tuned by changing the submonomer used to space ligand modules, as appropriately functionalized peptoids (*30*) and aminoglycosides (*26*) have been used to transport cargo into cells.

Comparison to Other Ligands Targeting DM RNAs. During the course of this work, DM1−MBNL-1 inhibitors were disclosed (*31*). Inhibitors had K_i's of ~4 μM, although no ligand inhibited more than 50% of the r$(CUG)_{109}$−MBNL-1 complex at the highest concentra-

tion tested (~20 μM). In related work herein, 3K-4 binds DM1-containing **RNA12** with a K_d of 210 nM (Table 2) and inhibits the r(CUG)$_{109}$—MBNL-1 interaction with an IC_{50} of 280 nM. The 3K-4 ligand binds with a similar affinity to **RNA10** (displays 2×2 5′CC/3′CC loops; K_d of 270 nM). However, it binds more weakly to fully base-paired **RNA8** and **RNA9**, **RNA11**, which displays 2×2 5′UU/3′UU loops, and **RNA13**, which displays 1×1 AA loops (Figure 3 and Table 2). These results are encouraging because they demonstrate that modular assembly provides nanomolar inhibitors for both DM1 and DM2 systems. 3K-4 may not be an optimal DM1 binder. Four propylamine spacers between ligand modules may not be ideal for targeting DM1 repeats since the distance between loops is likely to be shorter. Further investigation of DM1 and other systems will allow the development of rules for spacing ligand modules to bind motifs separated by different distances. Results herein have already provided insight into how to space ligand modules to target simultaneously two 2×2 nucleotide internal loops separated by two base pairs. These results are likely applicable to other RNAs.

Future Cell-Based Studies with Ligands Targeting DM RNAs. It remains to be seen if a ligand that binds DM repeats and displaces MBNL-1 would be effective at curing myotonic dystrophies. For example, it should be noted that DM1 RNAs are retained in the nucleus and are not translated as effectively *in vivo* (*32*). Another report has shown, however, that many of the pre-mRNA splicing defects associated with DM1 are corrected when MBNL-1 is overexpressed in DM1-affected mouse models of the disease (*33*). Thus, increasing the free amount of MBNL-1 can allow for correction of splicing defects. The latter report suggests that displacement of MBNL-1 from DM repeats by a ligand might lead to correction of splicing defects associated with DM. The former study suggests that a ligand that targets DM RNAs in the nucleus may inhibit the RNA from being exported to the cytoplasm and translated. Only cellular studies of ligands will address these issues. Taken together, cellular uptake and localization properties of DM targeting ligands are likely critical in designing effective therapies. Modular assembly may allow for ligand uptake and localization to be tuned by using different spacing modules (*30, 34, 35*). It may be possible to identify spacing modules that localize ligands to the nucleus and other spacing modules that allow ligands to move freely between the nucleus and cytoplasm. If these ligands can be identified, then the effects of ligand localization on correction of splicing and translational defects could be more thoroughly understood.

Summary. In summary, specific and potent nanomolar inhibitors targeting the toxic RNA hairpin that causes DM2 have been rationally designed. Initial results also show that this strategy can be applied to DM1. Presently the rational design of ligands targeting RNA is difficult, and most RNA targeting approaches rely on screening RNA targets for binding to chemical libraries. A more efficient approach to design ligands targeting RNA might be the one described here, which is related to modular polyamide targeting of DNA (*24*). Such an approach would not require subjecting each new RNA target to high-throughput screening assays to identify lead ligands and could be done computationally.

METHODS

Please see Supporting Information for all details related to the synthesis of the compounds used in these studies.

Preparation of RNAs. RNA1—RNA13 were transcribed from double stranded DNA templates that were generated from PCR amplification using a primer with a T7 promoter. DNA templates and PCR primers were purchased from Integrated DNA Technologies. The r(CUG)$_{109}$ (*36*) and r(CCUG)$_{47}$, RNAs were transcribed from plasmids encoding the repeats; the r(CCUG)$_{47}$ plasmid was cloned into pUC57 by GenScript Corporation while the plasmid encoding r(CUG)$_{109}$ was a generous gift from Dr. Charles Thornton (Department of Neurology, University of Rochester). The plasmids were linearized with *Bam*HI to afford RNAs suitable for affinity measurements or XbaI to afford RNAs suitable for MBNL-1 displacement assays. RNAs shown in Figure 3 that were used in MBNL-1 displacement assays contained an additional single stranded region (3′ tail) that is complementary to a biotinylated DNA capture probe. RNAs used to determine binding affinities did not contain the 3′ tail. All oligonucleotide sequences and PCR cycling conditions used are available in Supporting Information.

MBNL-1 Displacement Assays. Recombinant MBNL-1 protein that is fused to an amino acid sequence encoding the LacZα peptide was expressed and purified as previously described (*31*). The MBNL-1 displacements assays were completed using a modified procedure based on ref 22. The average amount of RNA immobilized in the well was determined using SYBR Green II (Invitrogen) and known concentrations of RNA. On average, when 2.5 pmol of RNA was delivered to a well, 0.5 pmol was immobilized. Please see Supporting Information for detailed procedures for the MBNL-1 displacement assays and the determination of the number of moles of RNA immobilized.

The resulting data were fit to a four parameter logistic curve to determine the IC_{50}'s when the percentage of MBNL-1 bound ranged from 0—100% using the following equation:

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

where y is the percentage of MBNL-1 bound, x is the concentration of ligand, D is the minimum response plateau, and A is the maximum response plateau. A and D are typically 0% and 100%, respectively. Each IC_{50} was the average of at least two measurements, and the error is the standard deviation in those measurements. The values for the multivalent effects were computed using eqs 2 and 3:

$$\text{normalized } IC_{50}(NIC_{50}) = IC_{50} \times \text{number of displayed modules} \quad (2)$$

$$\text{multivalent effect} = \frac{IC_{50}\textbf{FITC-1}}{NIC_{50}} \quad (3)$$

Affinity Measurements. The affinity of MBNL-1 for a subset of the RNA constructs shown in Figure 3 was completed by gel shift (*10*). Please see Supporting Information for experimental details and representative autoradiograms.

For affinity measurements of ligands and RNAs, fluorescence intensity was monitored as a function of RNA concentration. Plots of [RNA]/[ligand] *versus* change in fluorescence were used to determine stoichiometries. Scatchard analysis was used to determine dissociation constants (*37*). For RNA−ligand interactions with 12 repeats, statistical effects had to be taken into account; therefore, the interaction was treated as a large ligand binding to a lattice-like chain as described (*37, 38*). As such, the resulting curves were fit to eq 4:

$$\frac{v}{[L]} = \frac{N(1 - lv/N)}{k}\left(\frac{1 - lv/N}{1 - (l - 1)v/N}\right)^{l-1} \quad (4)$$

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.

Cell Culture and Uptake Experiments. The C2C12 (mouse myoblast) cell line was maintained as a monolayer in 1X DMEM supplemented with 10% (v/v) FBS and 0.5% penicillin/streptomycin. Cells were trypsinized and added to a well of a six-well plate containing a sterile glass coverslip onto which the cells adhered. After 24 h, the medium was replaced with fresh medium containing 5 μM of either 2K-4 or 3K-4 and incubated for 14 h. After washing with 1X DPBS, the coverslip was removed and placed face down in 5 μL of 1X DPBS + 50% (v/v) glycerol. The coverslip was sealed, and cells imaged using a Zeiss photomicroscope equipped with a Princeton Micromax CCD and Scanalytics IPLab software. Procedures for propidium iodide staining are available in Supporting Information.

Acknowledgment: We thank J. L. Childs-Disney for invaluable help with RNA binding assays and for critical review of the manuscript, K. Sobczak for generous donation of MBNL-1 expression plasmid and for help with the MBNL-1 inhibition assays, and A. J. Siegel and the University at Buffalo's Microscopic Imaging Facility, Department of Biological Sciences, for obtaining all microscopy images. Financial support by A NYSTAR JD Watson Award, The Research Corporation by a Cottrell Scholar Award, The New York State Center of Excellence in Bioinformatics and Life Sciences, and the National Institutes of Health (R01-GM079235) is gratefully acknowledged.

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

REFERENCES

1. Thomas, J. R., and Hergenrother, P. J. (2008) Targeting RNA with small molecules, *Chem. Rev. 108*, 1171–1224.
2. Childs-Disney, J. L., Wu, M., Pushechnikov, A., Aminova, O., and Disney, M. D. (2007) A small molecule microarray platform to select RNA internal loop−ligand interactions, *ACS Chem. Biol. 2*, 745–754.
3. Disney, M. D., Labuda, L. P., Paul, D. J., Poplawski, S. G., Pushechnikov, A., Tran, T., Velagapudi, S. P., Wu, M., and Childs-Disney, J. L. (2008) Two-dimensional combinatorial screening identifies specific aminoglycoside-RNA internal loop partners, *J. Am. Chem. Soc. 130*, 11185–11194.
4. Disney, M. D., and Childs-Disney, J. L. (2007) Using selection to identify and chemical microarray to study the RNA internal loops recognized by 6′-*N*-acylated kanamycin A, *ChemBioChem 8*, 649–656.
5. Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., and Ranum, L. P. (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9, *Science 293*, 864–867.
6. Mammen, M., Choi, S. K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors, *Angew. Chem., Int. Ed. 37*, 2755–2794.
7. Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat, *Science 289*, 1769–1773.
8. Mankodi, A., Teng-Umnuay, P., Krym, M., Henderson, D., Swanson, M., and Thornton, C. A. (2003) Ribonuclear inclusions in skeletal muscle in myotonic dystrophy types 1 and 2, *Ann. Neurol. 54*, 760–768.
9. Bevilacqua, J. M., and Bevilacqua, P. C. (1998) Thermodynamic analysis of an RNA combinatorial library contained in a short hairpin, *Biochemistry 37*, 15877–15884.
10. Warf, M. B., and Berglund, J. A. (2007) MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T, *RNA 13*, 2238–2251.
11. Jang, H., Fafarman, A., Holub, J. M., and Kirshenbaum, K. (2005) Click to fit: versatile polyvalent display on a peptidomimetic scaffold, *Org. Lett. 7*, 1951–1954.
12. Thomas, J. R., Liu, X., and Hergenrother, P. J. (2005) Size-specific ligands for RNA hairpin loops, *J. Am. Chem. Soc. 127*, 12434–12435.
13. See Supporting Information for details.
14. Gutell, R. R. (1994) Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994, *Nucleic Acids Res. 22*, 3502–3507.
15. Sucheck, S. J., Wong, A. L., Koeller, K. M., Boehr, D. D., Draker, K., Sears, P., Wright, G. D., and Wong, C. H. (2000) Design of bifunctional antibiotics that target bacterial rRNA and inhibit resistance-causing enzymes, *J. Am. Chem. Soc. 122*, 5230–5231.
16. Aminova, O., Paul, D. J., Childs-Disney, J. L., and Disney, M. D. (2008) Two-dimensional combinatorial screening identifies specific 6′-acylated kanamycin A- and 6′-acylated neamine-RNA hairpin interactions, *Biochemistry 47*, 12670–12679.
17. Li, L. B., Yu, Z., Teng, X., and Bonini, N. M. (2008) RNA toxicity is a component of ataxin-3 degeneration in *Drosophila*, *Nature (London) 453*, 1107–1111.
18. Gestwicki, J. E., Cairo, C. W., Strong, L. E., Oetjen, K. A., and Kiessling, L. L. (2002) Influencing receptor-ligand binding mechanisms with multivalent ligand architecture, *J. Am. Chem. Soc. 124*, 14922–14933.

19. Gordon, E. J., Sanders, W. J., and Kiessling, L. L. (1998) Synthetic ligands point to cell surface strategies, *Nature (London) 392*, 30–31.
20. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR, *Science 274*, 1531–1534.
21. Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., Read, R. J., and Bundle, D. R. (2000) Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands, *Nature (London) 403*, 669–672.
22. Maly, D. J., Choong, I. C., and Ellman, J. A. (2000) Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors, *Proc. Natl. Acad. Sci. U.S.A. 97*, 2419–2424.
23. Erlanson, D. A., Braisted, A. C., Raphael, D. R., Randal, M., Stroud, R. M., Gordon, E. M., and Wells, J. A. (2000) Site-directed ligand discovery, *Proc. Natl. Acad. Sci. U.S.A. 97*, 9367–9372.
24. Dervan, P. B. (2001) Molecular recognition of DNA by small molecules, *Bioorg. Med. Chem. 9*, 2215–2235.
25. Sanborn, T. J., Wu, C. W., Zuckermann, R. N., and Barron, A. E. (2002) Extreme stability of helices formed by water-soluble poly-N-substituted glycines (polypeptoids) with α-chiral side chains, *Biopolymers 63*, 12–20.
26. Luedtke, N. W., Carmichael, P., and Tor, Y. (2003) Cellular uptake of aminoglycosides, guanidinoglycosides, and poly-arginine, *J. Am. Chem. Soc. 125*, 12374–12375.
27. Alvarez-Salas, L. M. (2008) Nucleic acids as therapeutic agents, *Curr. Top. Med. Chem. 8*, 1379–1404.
28. Lebleu, B., Robbins, I., Bastide, L., Vives, E., and Gee, J. E. (1997) Pharmacokinetics of oligonucleotides in cell culture, *Ciba Found. Symp. 209*, 47–59.
29. Kanadia, R. N., Johnstone, K. A., Mankodi, A., Lungu, C., Thornton, C. A., Esson, D., Timmers, A. M., Hauswirth, W. W., and Swanson, M. S. (2003) A muscleblind knockout model for myotonic dystrophy, *Science 302*, 1978–1980.
30. Goun, E. A., Shinde, R., Dehnert, K. W., Adams-Bond, A., Wender, P. A., Contag, C. H., and Franc, B. L. (2006) Intracellular cargo delivery by an octaarginine transporter adapted to target prostate cancer cells through cell surface protease activation, *Bioconjugate Chem. 17*, 787–796.
31. Gareiss, P. C., Sobczak, K., McNaughton, B. R., Palde, P. B., Thornton, C. A., and Miller, B. L. (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), *J. Am. Chem. Soc. 130*, 16254–16261.
32. Mastroyiannopoulos, N. P., Feldman, M. L., Uney, J. B., Mahadevan, M. S., and Phylactou, L. A. (2005) Woodchuck post-transcriptional element induces nuclear export of myotonic dystrophy 3′ untranslated region transcripts, *EMBO Rep. 6*, 458–463.
33. Kanadia, R. N., Shin, J., Yuan, Y., Beattie, S. G., Wheeler, T. M., Thornton, C. A., and Swanson, M. S. (2006) Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly-(CUG) model for myotonic dystrophy, *Proc. Natl. Acad. Sci. U.S.A. 103*, 11748–11753.
34. Yu, P., Liu, B., and Kodadek, T. (2005) A high-throughput assay for assessing the cell permeability of combinatorial libraries, *Nat. Biotechnol. 23*, 746–751.
35. Goun, E. A., Pillow, T. H., Jones, L. R., Rothbard, J. B., and Wender, P. A. (2006) Molecular transporters: synthesis of oligoguanidinium transporters and their application to drug delivery and real-time imaging, *ChemBioChem 7*, 1497–1515.
36. Tian, B., White, R. J., Xia, T., Welle, S., Turner, D. H., Mathews, M. B., and Thornton, C. A. (2000) Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR, *RNA 6*, 79–87.
37. Cantor, C. R., and Schimmel, P. R. (1980) Ligand interactions at equilibrium, in *Biophysical Chemistry*, pp 849–886, W.H. Freeman and Company, San Francisco.
38. McGhee, J. D., and von Hippel, P. H. (1974) Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice, *J. Mol. Biol. 86*, 469–489.