Published in final edited form as:

Chembiochem. 2011 September 19; 12(14): 2143–2146. doi:10.1002/cbic.201100298.

Using Modularly Assembled Ligands To Bind RNA Internal Loops Separated by Different Distances

Jessica L. Childs-Disney^b, Pavel B. Tsitovich^c, and Matthew D. Disney^a

Matthew D. Disney: disney@scripps.edu

^aDepartment of Chemistry, The Scripps Research Institute, Scripps Florida, 130 Scripps Way, Jupiter, FL 33458 (USA)

^bCanisius College, Department of Chemistry and Biochemistry, 2001 Main Street, Buffalo, NY 14208 (USA)

^cDepartment of Chemistry, University at Buffalo, The State University of New York, Natural Sciences Complex, Buffalo, NY 14260 (USA)

Many studies have identified RNA drug or probe targets present in genomic sequence, however, most of these RNAs remain underexploited due to the difficulties in designing high affinity and specific RNA binders. Modular assembly and multivalency provide unique opportunities to tune the affinity and specificity of ligands for a target biomolecule. Herein, studies to investigate how far ligand modules should be spaced apart to bind RNA internal loops separated by different numbers of base pairs are reported. Results show that about four spacing submonomers between bis-benzimidazole modules displayed on a peptoid scaffold is optimal to span two base pairs. These results have general implications for the rational and modular design of ligands that specifically target RNA.

Previously, the assembly of a bis-benzimidazole ligand module (**H**) to bind simultaneously multiple secondary structural elements present in triplet repeating RNAs that cause myotonic muscular dystrophy type 1 (DM1), spinocerebellar ataxia type 3 (SCA3) and Huntington's disease, was reported.¹ The DM1 repeat, CUG, folds into a hairpin with periodically repeating UU mismatches separated by two base pairs in the stem (5'-CUGCUG/3'-GUCCUG). Likewise, the SCA3 and Huntington's repeats, CAG, fold into hairpins with regularly repeating AA mismatches separated by two base pairs (5'-CAGCAG/3'-GACGAC). Those studies identified an optimal distance between ligand modules to bind simultaneously these 1×1 nucleotide UU internal loops separated by two base pairs. The focus of the current study is to understand how to span longer distances between loops and to determine if the distance that is required to span two base pairs follows a simple pattern. For example, if doubling the distance that spans two base pairs would span four base pairs, etc. Such studies could serve as benchmarks to determine how to optimally target multiple motifs in an RNA by using a multivalent ligand.

A variety of other studies have utilized modular assembly approaches to target RNA. For example, bulges in HIV TAR RNA have been targeted with ethidium–arginine conjugates;² secondary-structure elements present in HIV TAR and RRE RNA, the iron responsive element, and thymidine synthase mRNA have been targeted with neomycin B–chloramphenicol conjugates;³ duplex RNAs have been targeted with peptide–quinoline conjugates;⁴ and loops in HIV RRE have been targeted with peptide–aminoglycoside

Correspondence to: Matthew D. Disney, disney@scripps.edu.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201100298>.

conjugates.⁵ Modular assembly has been shown in these cases and others⁶ to improve affinity and selectivity.

The azide-displaying bis-benzimidazole ligand module (**H**, Scheme 1) was modularly assembled onto an alkyne-displaying peptoid scaffold by using a Huisgen dipolar cycloaddition reaction.^{1a} 7 Peptoids were used as the display scaffold because both the ligand valency and the spacing between RNA binding modules (**H**) can be controlled (Scheme 1). The valency of the RNA-binding module is controlled by the number of propargylamine submonomers that are conjugated; the distance between each module, or spacing, is controlled by the number of propylamine submonomers inserted between each propargylamine submonomer (Scheme 1).^{1b} The nomenclature that is used to describe the peptoids is illustrated Scheme 1. For example, a peptoid that displays two **H** modules separated by one propylamine spacer is referred to as **2 H-1**, corresponding to values of $a=1$, $b=0$, and $c=1$ (Scheme 1). Thus, the number before the letter indicates the valency of the ligand; the letter refers to the module displayed; and, the number after the dash indicates the number of spacing submonomers between the RNA-binding modules. To complete these studies, a library of dimeric peptoids was synthesized that contain two copies of the **H** module separated by different distances (Scheme 1).^{1a}

The **H** module and the library of peptoids (**2 H-1** through **2 H-12**, and **2 H-16**) was then tested for binding to a series of RNA structures that contain two copies of the 5'-CUG/3'-GUC motif (1×1 UU internal loops) embedded in a hairpin cassette (Figure 1).⁸ These RNAs, **RNA1–RNA4**, differ in the number of base pairs between the internal loops. Binding affinity and stoichiometry were determined by using a fluorescence-based RNA binding assay in which the emission of the **H** ligand increases upon binding to an RNA target.^{1a} These data are summarized in Table 1.

The affinities and stoichiometries of the **H** ligand module for **RNA1–RNA4** were determined. The **H** module binds each RNA with similar affinity (~0.2 μM), in good agreement with previous reports.^{1a} 9 As expected, two **H** modules bind per RNA, or one **H** per 5'-CUG/3'-GUC. **H** binds herring sperm DNA (hs-DNA) with twofold higher affinity (Table 1).^{1a}

Next, each dimeric ligand was studied for binding to **RNA1–RNA4**. By testing each combination, we found that about four spacing submonomers were required to span every two base pairs in RNA structures. For example, the optimal binders for **RNA1** are **2 H-4** and **2 H-5** ($K_d=0.09 \mu\text{M}$). The **2 H-1** through **2 H-3** ligands bind much more weakly while **2 H-6** and **2 H-7** bind with 1.5- and fourfold less affinity, respectively. The optimal spacing for binding to **RNA2** (four base pairs separate the internal loops) is **2 H-8**, which has an affinity of 0.1 μM . Dimers with shorter distances between ligand modules bind much more weakly while those with longer distances bind at least 2.5-fold more weakly. **RNA3** (six base pairs separate the internal loops) binds most tightly to **2 H-12** and **2 H-16** with dissociation constants of 0.1 μM . As also observed for **RNA2**, shorter distances are not well tolerated as **2 H-1** through **2 H-9** bind sevenfold more weakly. The optimal dimer for **RNA4** (eight base pairs separate the internal loops) is **2 H-16**, although **2 H-10** through **2 H-12** bind similarly. In most cases, slightly larger than optimal distances between ligand modules are tolerated up to a certain point while shorter distances generally are not; this likely reflects flexibility in the peptoid scaffold.

The addition of **H** modules onto the peptoid scaffold increases the affinity to a target RNA by, on average, approximately twofold (Table 1). For example, the **H** module binds **RNA1** through **RNA4** with an average K_d of 225 nM while the optimal dimeric ligands for **RNA1** through **RNA4** bind with an average K_d of 120 nM (Table 1). The increase in affinity is not

as high as would be expected if the free energy for binding of the individual modules were summed and an entropy bonus was added.^{6c} There are several potential reasons why this is the case. First, the peptoid backbone could form unfavorable interactions with the negatively charged phosphodiester backbone of RNA. These could be either steric due to clashes to the propylamine side-chain spacer or electrostatic due to the peptoid backbone carbonyl being in close proximity to negatively charged RNA backbone. Second, the bis-benzimidazole Hoechst is known to form intermolecular stacked complexes in solution.¹⁰ At pH 7.0 and in 150 mM NaCl, the K_d of an intramolecular Hoechst 33258 aggregate is 200 μ M. For the intermolecular association of Hoechst derivatives that are on a single peptoid chain, the K_d is likely much higher due to the high effective concentration of ligand. Thus, for an oligomer to bind a target RNA and occupy multiple loop sites, the intramolecular complex in the peptoid will have to be broken for binding to an RNA target to occur. The net result would be an overall loss in binding affinity, however, it would increase binding specificity. The specificity would increase because there are likely only a few RNA targets to which a multivalent oligomer could bind in which the overall affinity would compensate for breaking intramolecular association. Such effects have been observed with DNA hairpins or so-called “stringency” clamps.¹¹ These effects are probably not specific to the Hoechst scaffold and are likely general features that one should consider when using modularly assembled ligands to target RNA structures, especially if the RNA binding modules are heterocycles. Third, the peptoid backbone is flexible and not likely “preorganized” for binding RNA. Binding to RNA targets would have to occur at the expense of the entropy of the unbound flexible scaffold and would decrease overall affinity.

The stoichiometries of these interactions were also determined (Table 1). RNAs that have two copies of the 5'-CUG/3'-GUC motif form interactions with two **H** modules. This indicates that the **H** modules can bind to adjacent motifs as has been previously observed.^{1a} The assembled ligands generally form a complex with two RNA molecules until they reach an optimal distance to span two 5'-CUG/3'-GUC motifs. For example, **2 H-1** to **2 H-3** form complexes with two RNAs (intermolecular binding); however, **2 H-4** forms an intramolecular stoichiometric complex; this indicates simultaneous binding to both 5'-CUG/3'-GUC motifs displayed in the same RNA. Optimal dimeric spaced ligands form complexes that bind multiple 5'-CUG/3'-GUC motifs in the same RNA simultaneously (**2 H-8**, **2 H-12**, and **2 H-16** with **RNA2**, **RNA3**, and **RNA4**, respectively).

In previously reported studies, we found that four propylamine spacing submonomers provided the most potent compounds when targeting DM1 hairpins with multiple, adjacent 5'-CUG/3'-GUC motifs.^{1a} Thus, the two studies are in good agreement. In contrast to our previous studies, however, the RNAs studied here contain loops that are separated by AU and GC pairs, not only GC pairs. Thus, about four spacers can span both homogeneous and heterogeneous alignments of base pairs in RNA structure. This observation is perhaps not surprising because the distance between internal loops separated by the same number of base pairs is not dependent on the identity of those pairs. It is also of interest to note that a dimeric ligand can span up to eight base pairs, which is the upper limit of pairs tested in this study. It is likely, however, that longer distances could be accommodated because there is only a small increase in affinity when the peptoid length is increased from **2 H-10** to the **2 H-16** ligand (Table 1).

As also observed in our previous report, the selectivity of the dimeric **H** peptoids for RNAs containing two 5'-CUG/3'-GUC motifs over hs-DNA is not improved over the **H** monomer (Table 1).^{1a} Therefore, trimeric peptoids were synthesized and studied for binding to **RNA5–RNA8**, which contain three 5'-CUG/3'-GUC motifs (Figure 1; Table 2). The average K_d of the optimal trimers for binding their respective RNAs is 72 nM. These

average affinity bonuses (over the monomer and dimers) are also similar to those reported with the DM1 hairpin study.^{1a}

Specifically, **3 H-4** forms a 1:1 complex with **RNA5** and binds with an affinity of $0.06(\pm 0.02)$ μM . As also previously observed,^{1a} selectivity over hs-DNA improves; **3 H-4** forms a 2:1 complex to the DNA and binds with an affinity of $0.48(\pm 0.07)$ μM . Analogously, **3 H-8** forms a 1:1 complex with **RNA6**, has an affinity of $0.06(\pm 0.01)$ μM , and is approximately 8.8-fold selective over hs-DNA ($K_d=0.53\pm 0.05$ μM ; ~1.5 ligands per DNA). **RNA7** forms a 1:1 complex with **3 H-12** and binds with a dissociation constant of $0.06(\pm 0.02)$ μM . Ligand **3 H-12** is approximately 8.8-fold selective for **RNA7** over hs-DNA (2:1 complex; $K_d=0.53\pm 0.03$ μM). The trimeric peptoid, **3 H-16**, binds in a 1:1 complex with **RNA8** with a K_d of $0.11(\pm 0.01)$ μM . The compound is fourfold selective for **RNA8** over hs-DNA ($K_d=0.43\pm 0.13$; 2 ligands per DNA). In order to make an additional assessment of the specificity of the optimal trimers used in this study, **3 H-4**, **3 H-8**, **3 H-12**, and **3 H-16**, were studied for their ability to bind tRNA. For each of these ligands, no change in emission was observed with up to 3 μM tRNA. This indicates that each of these compounds bind only weakly to this bystander RNA. Considering a K_d of >3 μM for these interactions, specificities for the optimal targets for these trimers relative to tRNA range from >50 - to >27 -fold. Because improved specificity is observed with the trimeric ligands, modular assembly allows for tuning the specificity of an assembled ligand for binding a target even if the ligand module lacks specificity.

In summary, we have systematically investigated how the distance between ligand modules affects binding affinity and selectivity using a library of dimeric and trimeric peptoids and RNAs that contain internal loops separated by different numbers of base pairs. Affinity and selectivity were improved by increasing the valency. Results presented here show that the assembly strategy can span different numbers of linear base pairs within RNA secondary structures. A variety of underexploited RNA drug targets contain potentially targetable motifs separated by different numbers of base pairs within a hairpin including pre- and pri-microRNAs.¹² However, the approach should also be applicable to spanning tertiary structure. Therefore, these results could aid in the modular and rational design of small molecules targeting RNAs present in genomic sequence and might enable a general approach to design ligands targeting RNA that is similar to polyamide recognition of the DNA minor groove.¹³

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding for this work was provided by the National Institutes of Health (RO1-GM079235). M.D.D. is a Camille and Henry Dreyfus New Faculty Awardee and Teacher-Scholar Awardee, a Cottrell Scholar through the Research Corporation, and a J. D. Watson Young Investigator through NYSTAR.

References

- 1a). Pushechnikov, Lee MM, Childs-Disney JL, Sobczak K, French JM, Thornton CA, Disney MD. J. Am. Chem. Soc. 2009; 131:9767–9779. [PubMed: 19552411]
- 1b). Lee MM, Pushechnikov A, Disney MD. ACS Chem. Biol. 2009; 4:345–355. [PubMed: 19348464]
- 1c). Lee MM, Childs-Disney JL, Pushechnikov A, French JM, Sobczak K, Thornton CA, Disney MD. J. Am. Chem. Soc. 2009; 131:17464–17472. [PubMed: 19904940]
2. Peytoux V, Condom R, Patino N, Guedj R, Aubertin AM, Gelus N, Bailly C, Terreux R, Cabrol-Bass D. J. Med. Chem. 1999; 42:4042–4053. [PubMed: 10514274]

3. Lee J, Kwon M, Lee KH, Jeong S, Hyun S, Shin KJ, Yu J. J. Am. Chem. Soc. 2004; 126:1956–1957. [PubMed: 14971927]
- 4a). Gooch D, Beal PA. J. Am. Chem. Soc. 2004; 126:10603–10610. [PubMed: 15327318]
- 4b). Krishnamurthy M, Gooch BD, Beal PA. Org. Lett. 2004; 6:63–66. [PubMed: 14703351]
- 4c). Carlson B, Vuyisich M, Gooch BD, Beal PA. Chem. Biol. 2003; 10:663–672. [PubMed: 12890540]
5. Park WKC, Auer M, Jaksche H, Wong CH. J. Am. Chem. Soc. 1996; 118:10150–10155.
- 6a). Gestwicki JE, Cairo CW, Strong LE, Oetjen KA, Kiessling LL. J. Am. Chem. Soc. 2002; 124:14922–14933. [PubMed: 12475334]
- 6c). Mammen M, Choi SK, Whitesides GM. Angew. Chem. 1998; 110:2908–2953.
- 6b). Göringer HU, Wagner R. Nucleic Acids Res. 1986; 14:7473–7485. [PubMed: 2429262]
- 6d). Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Science. 1996; 274:1531–1534. [PubMed: 8929414]
- 6e). Kitov PI, Sadowska JM, Mulvey G, Armstrong GD, Ling H, Pannu NS, Read RJ, Bundle DR. Nature. 2000; 403:669–672. [PubMed: 10688205]
7. Jang H, Fafarman A, Holub JM, Kirshenbaum K. Org. Lett. 2005; 7:1951–1954. [PubMed: 15876027]
- 8a). Disney MD, Childs-Disney JL. ChemBioChem. 2007; 8:649–656. [PubMed: 17394189]
- 8b). Childs-Disney JL, Wu M, Pushechnikov A, Aminova O, Disney MD. ACS Chem. Biol. 2007; 2:745–754. [PubMed: 17975888]
9. Cho J, Rando RR. Nucleic Acids Res. 2000; 28:2158–2163. [PubMed: 10773086]
10. Buurma NJ, Haq I. J. Mol. Biol. 2008; 381:607–621. [PubMed: 18617189]
11. Roberts RW, Crothers DM. Proc. Natl. Acad. Sci. USA. 1991; 88:9397–9401. [PubMed: 1946351]
- 12a). Calin GA, Croce CM. Oncogene. 2006; 25:6202–6210. [PubMed: 17028600]
- 12b). Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Science. 2001; 294:853–858. [PubMed: 11679670]
- 12c). Lau NC, Lim LP, Weinstein EG, Bartel DP. Science. 2001; 294:858–862. [PubMed: 11679671]
- 12d). Lee RC, Ambros V. Science. 2001; 294:862–864. [PubMed: 11679672]
- 12e). Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, Li M, Wang G, Liu Y. Nucleic Acids Res. 2009; 37:D98–D104. [PubMed: 18927107]
- 12f). Betel, Wilson M, Gabow A, Marks DS, Sander C. Nucleic Acids Res. 2008; 36:D149–D153. [PubMed: 18158296]
- 12g). Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. Nucleic Acids Res. 2008; 36:D154–D158. [PubMed: 17991681]
13. Dervan PB. Bioorg. Med. Chem. 2001; 9:2215–2235. [PubMed: 11553460]

Figure 1.

The secondary structures of the RNAs used in this study. Each RNA displays a linear array of two or three copies of the boxed motif in the cassette that is shown to the left.

Scheme 1.

The structure of the RNA-binding module (H) and the peptoid scaffold that allows for precise control over module valency and spacing. The nomenclature used for the assembled ligands is as follows: a ligand with a, b, c=1 is called **3 H-1**, which denotes display of three H modules each separated by one spacing module. “Ht” denotes an H module that was conjugated to the peptoid backbone.

Table 1

Measured affinities and stoichiometries of monomeric **H** and assembled dimeric (**2 H-n**) ligand-nucleic acid complexes.

	RNA1		RNA2		RNA3		RNA4		hs-DNA [b]	
	K_d [μ M]	Stoich. (RNA/lig.)	K_d [μ M]	Stoich. (RNA/lig.)	K_d [μ M]	Stoich. (RNA/lig.)	K_d [μ M]	Stoich. (RNA/lig.)	K_d [μ M]	Stoich. (DNA:lig.)
Monomer										
H	0.21±0.04	1:1.8±0.1	0.22±0.10	1:2.0±0.1	0.19±0.03	1:2.0±0.1	0.28±0.03	1:1.8±0.2	0.11±0.02	1:9±1.3
Dimers										
2H-1	> 2	2.0±0.2:1	~3	2.3±0.4:1	2.3±0.35	2.5±0.8:1	1.7±0.4	2.2±0.3:1		
2H-2	4.0±0.0	2.3±0.4:1	2.2	2.0±0.0:1	> 2	2.7±0.6:1	> 3	2.3±0.4:1		
2H-3	n.d./[a]	> 2.5:1	n.d./[a]	> 2.5:1	n.d./[a]	> 2.5:1:1	n.d./[a]	> 2.5:1		
2H-4	0.09±0.02	1.2±0.2:1	~1.7	1.9±0.1:1	n.d./[a]	> 3:1	> 1.5	2.0±0.8:1	0.05±0.01	1:8.5±0.7
2H-5	0.09±0.02	1.1±0.1:1	2.1±0.27	2.3±0.4:1	1.4±0.34	1.7±0.3:1	> 1.5	2.1±0.4:1		
2H-6	0.14±0.01	1.1±0.3:1	n.d./[a]	> 2:1	1.0±0.34	1.5±0.3:1	1.2±0.3	1.4±0.2:1		
2H-7	0.37±0.04	1.7±0.6:1	1.6±0.01	1.5±0.2:1	0.78±0.13	1.4±0.1:1	1.4±0.1	1.5±0.3:1		
2H-8	0.89±0.14	2.5±0.9:1	0.10±0.02	1.2±0.1:1	0.73±0.05	1.4±0.2:1	1.0±0.2	1.6±0.1:1	0.09±0.02	1:8.6±0.1
2H-9	0.84±0.12	2.3±0.7:1	0.64±0.13	1.2±0.2:1	0.67±0.12	1.2±0.4:1	~1.5	1.6±0.1:1		
2H-10	n.d./[a]	> 2.5:1	0.26±0.08	1.4±0.4:1	0.35±0.15	1.1±0.3:1	0.24±0.01	1.3±0.1:1	0.06±0.01	1:11.4±0.3
2H-11	n.d./[a]	> 2.5:1	0.30±0.07	1.1±0.1:1	~0.14	1.2±0.1:1	0.29±0.02	1.2±0.2:1		
2H-12	1.2±0.2	2.0±0.0:1	0.4±0.002	0.9±0.1:1	0.10±0.03	1.2±0.1:1	0.27±0.05	1.1±0.2:1	0.14±0.01	1:7.5±0.1
2H-16	2.5±0.7	2.6±0.3:1	0.3±0.02	1.0±0.1:1	0.10±0.01	1.0±0.1:1	0.19±0.08	1.0±0.1:1	0.07±0.02	1:8.3±0.7

[a] n.d.: No determination was made since the binding curves did not reach saturation;

[b] herring sperm DNA.

Table 2Dissociation constants, stoichiometries, and selectivities of trimeric peptoids (**3H-*n***).

Nucleic acid	Ligand	K_d [μ M]	Stoichiometry (DNA/RNA:lig.)	Selectivity
RNA5	3H-4	0.06 $\pm$ 0.02	1.1 $\pm$ 0.2:1	–
hs-DNA [a]	3H-4	0.48 $\pm$ 0.07	1:1.9 $\pm$ 0.1	8.0
RNA6	3H-8	0.06 $\pm$ 0.01	1.0 $\pm$ 0.2:1	–
hs-DNA	3H-8	0.53 $\pm$ 0.05	1:1.3 $\pm$ 0.1	8.8
RNA7	3H-12	0.06 $\pm$ 0.02	1.0 $\pm$ 0.2:1	–
hs-DNA	3H-12	0.53 $\pm$ 0.03	1:1.8 $\pm$ 0.3	8.8
RNA8	3H-16	0.11 $\pm$ 0.01	1.1 $\pm$ 0.2:1	–
hs-DNA	3H-16	0.43 $\pm$ 0.13	1:1.9 $\pm$ 0.1	3.9

[a] Herring sperm DNA.