J|A|C|S
COMMUNICATIONS

Published on Web 08/17/2005

Size-Specific Ligands for RNA Hairpin Loops

Jason R. Thomas, Xianjun Liu, and Paul J. Hergenrother*
Department of Chemistry, Roger Adams Laboratory, University of Illinois, Urbana, Illinois 61801

Received March 16, 2005; E-mail: hergenro@uiuc.edu

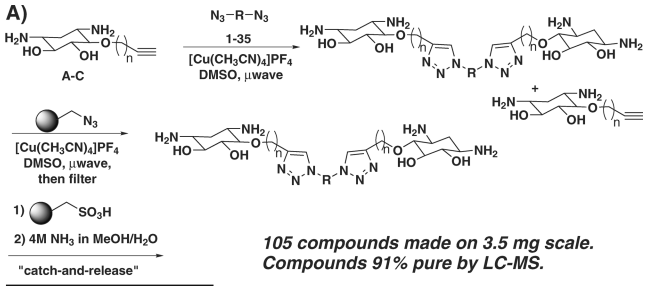
Of the proteins encoded by the ~30,000 human genes, it is estimated that only 10−14% have appropriate binding sites for small molecules.[1] As such, large swaths of the proteome have resisted traditional drug discovery efforts. Although targeting the individual mRNAs that code for the remaining 86−90% of the proteins has been considered,[2] the examples of exogenously added small molecules that target mRNA are vanishingly few.[3] Targeting one mRNA of the transcriptome requires a small molecule ligand to distinguish between many RNAs of similar sequence and structure. Unlike DNA, which exists in the fully base paired, double helical form in vivo, single stranded mRNA adopts a variety of secondary structures, including hairpin loops and bulges. Thus, small molecules that specifically recognize certain RNA secondary structures, especially comparatively rare motifs such as octaloops and three-base bulges, could be useful for targeting individual RNAs in vivo. Herein we report the discovery of compounds that specifically recognize RNA hairpin octaloops or selectively bind RNA hairpin tetraloops.

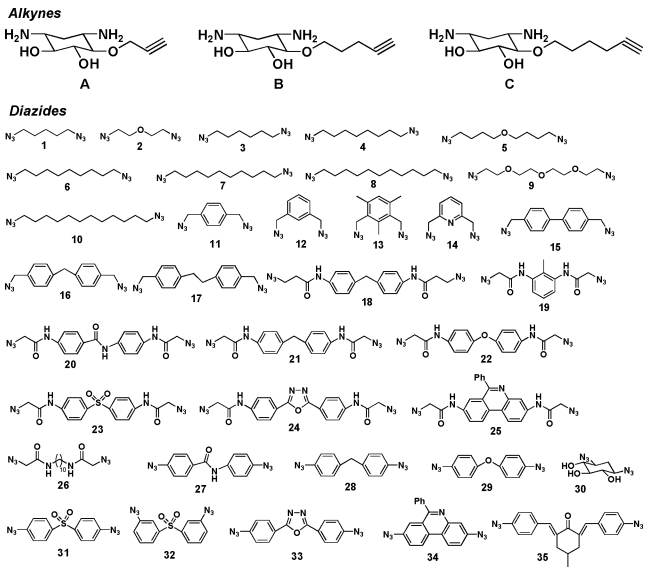
We have recently disclosed a class of compounds, dimers of deoxystreptamine, which bind tightly to RNA hairpin loops.[4] Although this is one of the few classes of small molecules that will recognize hairpin loops, in general these compounds do not discriminate in their binding and show nearly equal affinities for tetra-, hexa-, hepta-, and octaloops. To identify small molecules that are able to recognize and bind one loop size over all others, a combinatorial library of 105 deoxystreptamine dimers was synthesized in parallel, and the library members were subsequently evaluated for their hairpin loop binding properties.

A facile protocol for the dimerization of deoxystreptamine was developed by utilizing the copper-catalyzed variant of the Huisgen 1,3-dipolar cycloaddition of alkynes and azides (Scheme 1A).[5,6] It was found that 2.2 equiv of deoxystreptamine (modified to contain an alkyne) reacted rapidly with 1 equiv of a diazide to give the bis-triazole deoxystreptamine dimer. Under microwave heating, this reaction proceeded almost quantitatively in 40 s. A two-step purification scheme was utilized that involved the sequential addition of azide beads (to scavenge the remaining alkyne) and acidic resin (to "catch" the product and allow the removal of the copper salts). As test cases indicated that the protocol depicted in Scheme 1A provided excellent yields of highly pure products, three enantiopure deoxystreptamine-alkynes (**A**−**C**, Scheme 1) and 35 diazides (**1**−**35**, Scheme 1) were synthesized (see Supporting Information). These building blocks were selected to vary both the length and chemical nature of the linkage between the deoxystreptamine monomers. Once in hand, these alkynes and azides were coupled in parallel to produce 105 bis-triazole products; all compounds were synthesized on a 0.01 mmol scale to provide approximately 3.5 mg of each product. Analysis of the 105 reaction products by LC−MS indicated that the library had an average purity of 91%, and the molecular ion peak was confirmed for all library members (see Supporting Information for exact purity of individual compounds).

All 105 compounds were evaluated versus tetra-, hexa-, hepta-, and octahairpin loops of RNA (structures **I**−**IV** in Table 1).[7] To assess ligand binding, the RNA loops were labeled at their 3′ end

Scheme 1. Parallel Library of Deoxystreptamine Dimers





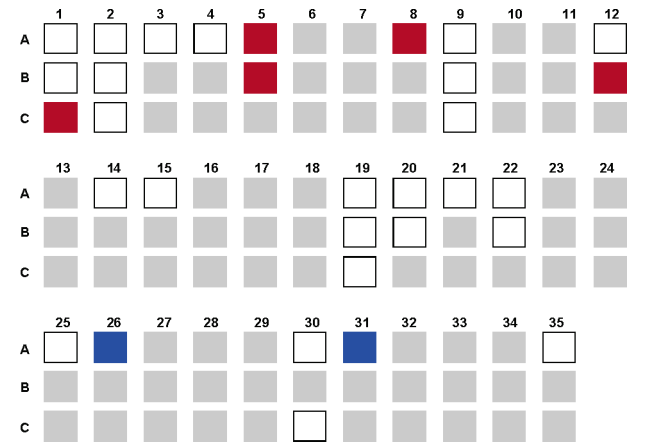
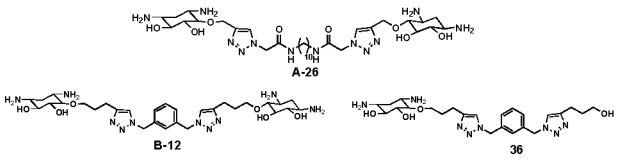


Figure 1. The 105 compounds synthesized were each assessed for their binding to RNA hairpin loops **I**, **II**, **III**, **IV** (shown in Table 1). Compounds **A-5**, **A-8**, **B-5**, **B-12**, and **C-1** (red boxes) had a greater than 10-fold selectivity for octaloop **IV**, whereas compounds **A-26** and **A-31** (blue boxes) had selectivity for tetraloop **I**. Gray boxes indicate compounds that bound the RNAs nonselectivity, while white boxes denote compounds that had binding constants greater than 1.5 μM for all RNAs. See Supporting Information for full binding data.

10.1021/ja051685b CCC: $30.25

Table 1. Dissociation Constants (in μM) for Small Molecule−RNA Interactions

Compound	I	II	III	IV	V	VI	VII	VIII	IX
A-26	0.30	6.49	1.60	4.82	>25.0	5.54	0.10	0.07	>25.0
B-12	>25.0	10.03	9.87	0.32	>25.0	>25.0	1.36	1.54	0.15
36	>25.0	>25.0	>25.0	>25.0	1.05	9.60	>25.0	4.34	>25.0



with fluorescein.[8] The four RNA loops were evaluated at five or more concentrations of the 105 ligands to obtain an estimate of the dissociation constant. As shown in Figure 1, two compounds showed selectivity for the RNA tetraloop (**A-26** and **A-31**; blue squares in Figure 1), and five compounds showed octaloop specificity (**C-1**, **A-5**, **B-5**, **A-8**, **B-12**; red squares in Figure 1).

These compounds that showed reasonable selectivity were re-tested to obtain dose-dependent curves from which K_d values could be calculated (Table 1). Compound **A-26** showed a strong affinity ($K_d = 0.3\ \mu$M) for RNA tetraloop **I**, which was >10-fold tighter than its binding to hexaloop **II** and octaloop **IV**, and ~5-fold tighter than the binding to heptaloop **III**. Among the octaloop selective ligands, compound **B-12** showed both excellent affinity for octaloop **IV** ($K_d = 0.31\ \mu$M) as well as substantial specificity (>30-fold in all cases).

On the basis of their selectivity profiles, the RNA binding specificity of compounds **A-26** and **B-12** were further evaluated. In the interest of identifying compounds that are truly specific for the RNA hairpin loop motif, these size-selective ligands were also tested against both a RNA duplex and a RNA single-base bulge. Both **A-26** and **B-12** appear to be selective for hairpin loops; the dissociation constant for **B-12** with duplex RNA **V** and bulge-containing RNA **VI** is greater than 25 μM, while **A-26** binds to the RNA bulge ~12-fold weaker than it binds to the tetraloop.

The site of ligand binding was determined for **A-26** and **B-12** using RNase footprinting; these footprinting experiments confirmed that the compounds bind to the loop region of the hairpin loops (see Supporting Information for footprints). The binding of these compounds to RNA hairpin loops was unaffected by high salt concentrations, or presence of competitor tRNAs[9] (see Supporting Information). In addition, the binding specificity of **A-26** and **B-12** was independently confirmed using UV-melting experiments (see Supporting Information). Finally, the importance of the dimeric nature of the deoxystreptamine was confirmed through the synthesis and evaluation of the mono-deoxystreptamine compound **36**, which was not a general RNA loop binder (Table 1).

To assess the binding of **A-26** to tetraloops of unrelated sequence, the GNRA (**VII**) and the UUCG (**VIII**) tetraloop sequences were utilized. These tetraloops are thermodynamically stable and widely observed in vivo.[10] In addition, the UUCG loop is present in the A-site of the 16S ribosomal RNA. Binding assays indicate that **A-26** binds these biologically relevant tetraloops **VII** and **VIII** with a strong affinity (70−100 nM, Table 1).

In contrast to tetraloops, RNA hairpin octaloops are much less common. The RNA octaloop present in the hepatitis C IRES was chosen to assess the primary sequence requirements for the binding of compound **B-12**.[11] Again, this compound binds octaloop **IX** (whose sequence is unrelated to octaloop **IV**) very tightly, with a $K_d = 150$ nM. The combined binding data for compounds **A-26** and **B-12** are consistent with the notion that these ligands are selective for the size of the RNA hairpin loop, rather than its precise primary sequence.

Years of research have resulted in the development of a paradigm for small molecule−duplex DNA binding based on the primary sequence of the DNA.[12] In contrast, the development of general methods for small-molecule−RNA binding has been limited. However, RNA adopts unique secondary structures that might be targeted with appropriate small molecule "modules" specific for various sizes of RNA hairpin loops, bulges, or internal loops. Described herein is a compound that specifically binds to RNA tetraloops and another that binds to octaloops; both bind with a high degrees of specificity over other loop sizes and other secondary structures. These are the first compounds reported to have such discrimination between RNA hairpin loops of various sizes. This stands in contrast to common RNA ligands (such as aminoglycosides) that have been shown to bind promiscuously to a variety of RNA secondary structures.[13] In addition, the strength of the small molecule−RNA interactions described herein (70−320 nM) compares favorably with most other RNA−ligand complexes, which typically fall in the low micromolar range.[14] It is remarkable that small perturbations in the linker can lead to compounds of substantial specificity; the structural basis for this discrimination is under active investigation and will be reported in due course.

Acknowledgment. This work was supported by the National Institutes of Health (NIHGMS R01-GM68385) and the Office of Naval Research (N00014-02-1-0390). P.J.H. is a fellow of the Alfred P. Sloan foundation.

Supporting Information Available: Full experimental protocols, binding data, RNase I footprints, and characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hopkins, A. L.; Groom, C. R. *Nat. Rev. Drug. Discovery* **2002**, *1*, 727−730.
(2) For reviews on targeting RNA see: (a) Sucheck, S. J.; Wong, C.-H. *Curr. Opin. Chem. Biol.* **2000**, *4*, 678−686. (b) Hermann, T. *Angew. Chem., Int. Ed.* **2000**, *39*, 1890−1904. (c) Michael, K.; Tor, Y. *Chem. Eur. J.* **1998**, *4*, 2091−2098.
(3) Examples of targeting mRNA with small molecules: (a) Werstuck, G.; Green, M. R. *Science* **1998**, *282*, 296−298. (b) DeNap, J. C.; Thomas, J. R.; Musk, D. J.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2004**, *126*, 15402−15404.
(4) Liu, X.; Thomas, J. R.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2004**, *126*, 9196−9167.
(5) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596−2599.
(6) For other dimeric RNA binding small molecules, see: (a) 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. (b) Tok, J. B.; Huffman, G. R. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1593−1595.
(7) Stable RNA stem loops were designed using mfold (Mathews, D. H.; Sabrina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911−940). The corresponding pentaloop showed multiple structures in mfold and was not utilized in this study.
(8) Fluorescently end-labeled RNA to determine small molecule−RNA K_d's: (a) Llano-Sotelo, B.; Chow, C. S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 213−216. (b) Llano-Sotelo, B.; Azucena, E. F.; Kotra, L. P.; Mobashery, S.; Chow, C. S. *Chem. Biol.* **2002**, *9*, 455−463. (c) Thomas, J. R.; DeNap, J. C.; Wong, M. L.; Hergenrother, P. J. *Biochemistry* **2005**, *44*, 6800−6808.
(9) Luedtke, N. W.; Liu, Q.; Tor, Y. *Biochemistry* **2003**, *42*, 11391−11403.
(10) Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379−404.
(11) Klinck, R.; Westhof, E.; Walker, S.; Afshar, M.; Collier, A.; Aboul-Ela, F. *RNA* **2000**, *6*, 1423−1431.
(12) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284−299.
(13) Verhelst, S. H.; Michiels, P. J.; van der Marel, G. A.; van Boeckel, C. A.; van Boom, J. H. *ChemBioChem* **2004**, *5*, 937−942.
(14) For another high affinity small molecule−RNA interaction see: Evans, J. M.; Turner, B. A.; Bowen, S.; Ho, A. M.; Sarver, R. W.; Benson, E.; Parker, C. N. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 993−996.

JA051685B