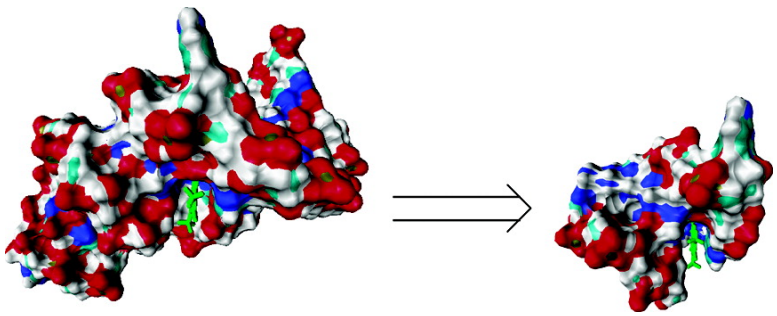
Communication

Unusually Short RNA Sequences: Design of a 13-mer RNA that Selectively Binds and Recognizes Theophylline

Peter C. Anderson, and Sandro Mecozzi

J. Am. Chem. Soc., **2005**, 127 (15), 5290-5291• DOI: 10.1021/ja0432463 • Publication Date (Web): 26 March 2005



J|A|C|S
COMMUNICATIONS

Published on Web 03/26/2005

Unusually Short RNA Sequences: Design of a 13-mer RNA that Selectively Binds and Recognizes Theophylline

Peter C. Anderson[§] and Sandro Mecozzi*[,§,‡]

School of Pharmacy and Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53705

Received November 9, 2004; E-mail: smecozzi@pharmacy.wisc.edu

The ability of small RNAs to recognize and bind selectively to ligand targets is now recognized as critical in gene regulation.[1] MicroRNAs (miRNAs) and short interfering RNAs (siRNAs), both ~22 nt in length, are key players in eukaryotic RNA-based gene regulation, specifically, in the process of RNA interference (RNAi).[1−4] The RNAi pathway can be found in a vast range of eukaryotes, including various mammalian cells.[5] In RNAi, miRNAs and siRNAs are incorporated into effector complexes (the RNA-induced silencing complex; RISC). The RISC then proceeds to direct either cleavage or translational repression of a target mRNA that is complementary to the small RNA.[1] Knowledge of the mechanisms of RNAi is being applied to the development of therapeutic treatments for HIV, cancer, viral hepatitis, and genetic diseases.[3]

Another class of small RNA molecules is likewise being exploited for its ligand recognition ability. Aptamers are short (generally ~30 to a few hundred nucleotides) nucleic acids developed by *in vitro* evolution techniques that bind to specific ligands with high affinity and selectivity.[6,7] The abundance of three-dimensional structures that RNA can adopt has allowed aptamers to be developed against a wide array of target molecules, including small molecules, peptide sequences, proteins, and oligosaccharides, often with affinities in the low nanomolar range.[8] Aptamers are currently being investigated in clinical trials for their use as therapeutics against HIV and macular degeneration, and preclinical studies of aptamers for other disease states are being conducted.[9]

As nucleic acids increase in length, they have a greater chance of getting locked into metastable inactive conformations due to unexpected interactions as they sample larger conformational spaces. The free energy of stacking of RNA nucleotides allows inactive conformations to be separated by large energy barriers from active conformations, which are often much larger than those in proteins.[10] This suggests that shorter aptamers may have a therapeutic advantage over their longer counterparts by being less prone to becoming inactive.

To evaluate whether an RNA aptamer's nucleotide sequence can be shortened while still allowing the RNA to bind selectively to its target ligand, we have developed a methodology that combines molecular dynamics (MD) simulations, thermodynamic integration (TI) calculations, and in vitro binding assays. A 33 nt RNA aptamer[11] that binds to the bronchodilator theophylline (**1**) with a K_a of ~3.3×10^6 M^{-1} was selected as a simple model system due to its small size and planar ligand (Figure 1). This aptamer (PDB code 1EHT) is able to discriminate against caffeine (**2**), which differs from theophylline only by a methyl group at the N7 position, with a 10^4-fold difference in binding affinity.[11,12]

The methodology that we have developed to shorten an RNA involves iterative computational deletion of nucleotides outside of the binding pocket, followed by MD simulations of the resulting truncated RNA−theophylline complex. The process is repeated until MD simulations no longer predict stability of the binding pocket structure and show either weakening of the complex or departure of the theophylline molecule from the pocket.

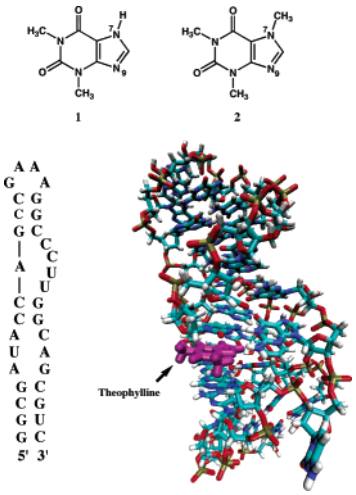


Figure 1. Secondary structure of the 33-mer aptamer for theophylline (left) and a three-dimensional view of the aptamer complexed with theophylline (right).[12]

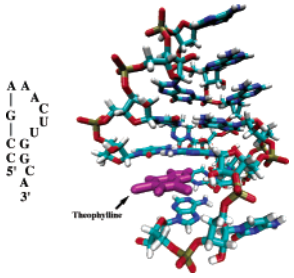


Figure 2. Secondary structure of the 13-mer aptamer truncation (left) and a three-dimensional view of the truncated RNA complexed with theophylline (right). This is the shortest nucleotide sequence predicted by MD simulations to bind to theophylline.

Our methodology was used to remove computationally all nucleotides outside of the binding pocket, with the exception of a GAAA tetraloop and two additional scaffolding nucleotides (C9 and C22), forming a 13-mer (Figure 2). This 13-mer structure is the shortest truncation that we developed that is both predicted computationally and observed experimentally to bind theophylline.

Molecular dynamics simulations 4 ns in length with explicit water solvent were performed on the 13-mer−theophylline complex as well as on the 13-mer−caffeine complex using the AMBER 7 suite of programs,[13] and the trajectories predicted structural stability of the RNA and binding of the theophylline. Thermodynamic integration calculations predicted binding energies, ΔG_{bind}, of theophylline and caffeine to the 13-mer as −5.8 kcal/mol (K_a = 16 800 M^{-1})

[§] School of Pharmacy.
[‡] Department of Chemistry.

10.1021/ja0432463 CCC: $30.25

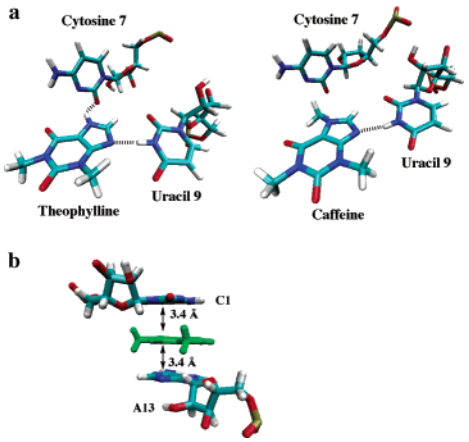


Figure 3. Interactions in the binding pocket of the 13-mer RNA. (a) Theophylline forms two hydrogen bonds (dashed lines) to C7 and U9 (left) and caffeine forms one hydrogen bond to U9 only (right). (b) Theophylline stacks between C1 and A13 (corresponding to C8 and A28 in the 33-mer), separated from both nucleotide rings by 3.4 Å in both the 33-mer aptamer and the 13-mer truncation.

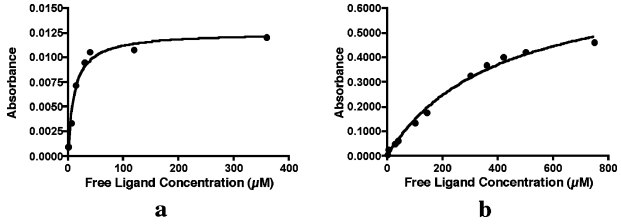


Figure 4. Binding regression curves for theophylline binding (a) and caffeine binding (b) to the 13-mer RNA. All measurements were done at a temperature of 23 °C.

and −5.0 kcal/mol (K_a = 4400 M^{-1}), respectively. In addition, the MD trajectories indicate that the theophylline molecule forms two hydrogen bonds with the 13-mer binding pocket, whereas caffeine forms only one hydrogen bond with a greater donor−acceptor distance (Figure 3a). Moreover, structural analysis of the theophylline trajectory suggests that the 13-mer binding pocket has stacking interactions with theophylline similar to those present in the original aptamer (Figure 3b).

The predicted binding energies and selectivity were confirmed experimentally by in vitro binding assays of the 13-mer RNA with both theophylline and caffeine. Increasing concentrations of ligand were added to solutions of constant RNA concentration, and UV absorbances of the mixtures were fit to a 1:1 nonlinear binding isotherm[14] (Figure 4).

The regression results indicate values of ΔG_{bind} as −6.7 ± 0.1 kcal/mol (K_a = 76 100 ± 15 800 M^{-1}) and −4.4 ± 0.1 kcal/mol (K_a = 1700 ± 180 M^{-1}) for theophylline and caffeine, respectively.

Our assay methodology was verified by a positive-control assay employing the 33-mer aptamer. The 33-mer was transcribed from a 50 nt antisense DNA with a 17 nt promoter, and transcription was carried out using T7 polymerase. The transcribed 33-mer was evaluated for its binding affinity to theophylline using the same UV assay methodology. Nonlinear regression yields a ΔG_{bind} of −8.9 ± 0.5 kcal/mol (K_a = 3.3 × 10^6 ± 2.0 × 10^5 M^{-1}) for the 33-mer, which is also that reported by use of liquid scintillation counting.[12]

The 13-mer structure represents the minimal structure capable of binding theophylline. The 11-mer structure formed by removal of the C2−G10 base pair of the 13-mer was unable to bind theophylline both in vitro and computationally. Moreover, the 13-mer exhibits a slightly smaller binding energy than that of the 15-mer truncation, from which it is directly derived. Thermodynamic integration calculations predict that theophylline binds to the 15-mer with a ΔG_{bind} of −6.5 kcal/mol (K_a = 54 400 M^{-1}). Finally, the average root-mean-squared deviations (RMSDs) of RNA atomic positions from MD trajectories of unbound 33-mer and 13-mer in explicit solvent are 2.38 and 2.57 Å, respectively, suggesting that the 33-mer in its unbound state is more stable and has less overall conformational flexibility than the unbound 13-mer, presumably due to the latter's smaller number of stabilizing hydrogen bonds between base pairs. Similarly, the average RMSD of the 11-mer in explicit solvent is 3.10 Å, consistent with a further decrease in stability in going from the 13-mer to the 11-mer. This markedly decreased stability of the 11-mer is putatively the reason it fails to bind theophylline. The melting temperatures of the 33-mer, the 13-mer, and the 11-mer are 70, 37, and 23 °C, respectively. These values support the observed trend in binding energies.

In summary, we have shown that a 33-mer RNA aptamer with high affinity and selectivity for theophylline can be truncated to an exceptionally short 13-mer RNA containing only the original binding pocket and six additional scaffolding nucleotides while still retaining selective binding. The fact that the binding ability of the original theophylline 33-mer RNA aptamer can be mimicked by a 13-mer oligoribonucleotide suggests that even very short RNAs can, in principle, carry out previously unrecognized tasks in a cellular environment. To our knowledge, this is the first demonstration of the ability of a 13-mer RNA to be able to bind selectively to a ligand target. This suggests the possibility of using very short RNAs for therapeutic intervention.

Acknowledgment. We thank Prof. Charles Lauhon for useful discussions and help in preparing the 1EHT RNA aptamer, Prof. Kenneth Connors for suggestions regarding the in vitro binding assay, and Prof. James Dahlberg for reviewing and commenting on this manuscript. We would also like to thank the University of Wisconsin−Madison and the WARF Foundation for funding.

Supporting Information Available: Experimental details of 13-mer, 33-mer, and 11-mer RNA binding assays, 15-mer structure from which 13-mer was derived, absorption spectrum of theophylline, rationale for the use of UV−vis assay, and thermal denaturation conditions. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) He, L.; Hannon, G. J. *Nat. Rev. Genet*. **2004**, *5*, 522−531.
(2) Carrington, J. C.; Ambros, V. *Science* **2003**, *301*, 336−338.
(3) Hannon, G. J.; Rossi, J. J. *Nature* **2004**, *431*, 371−378.
(4) Bartel, D. P. *Cell* **2004**, *116*, 281−297.
(5) Hannon, G. J. *Nature* **2002**, *418*, 244−251.
(6) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505−510.
(7) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818−822.
(8) Hermann, T.; Patel, D. J. *Nature* **2000**, *287*, 820−824.
(9) Rimmele, M. *ChemBioChem* **2003**, *4*, 963−971.
(10) Uhlenbeck, O. C. *RNA* **1995**, *1*, 4−6.
(11) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J. P.; Pardi, A. *Nat. Struct. Biol.* **1997**, *4*, 644−649.
(12) Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425−1429.
(13) Case, D. A. et al. *AMBER 7*; University of California: San Francisco, CA, 2002.
(14) Connors, K. A. *Binding Constants: The Measurement of Molecular Complex Stability*; John Wiley & Sons: New York, 1987; pp 141−156.

JA0432463