# Molecular Recognition of a RNA:DNA Hybrid Structure

Jinsong Ren,[†] Xiaogang Qu,[†] Nanibhushan Dattagupta,[‡] and Jonathan B. Chaires*,[†]

*Department of Biochemistry*
*University of Mississippi Medical Center*
*2500 North State Street, Jackson, Mississippi 39216-4505*
*Applied Gene Technologies, Inc.*
*6190 Cornerstone Court East, Suite 101*
*San Diego, California 92121*

*Received February 8, 2001*
*Revised Manuscript Received May 29, 2001*

RNA:DNA hybrids are important structures that are transiently formed in the course of many vital biological processes, including DNA replication,[1] telomere replication by telomerase,[2] and the replication of HIV (and other retroviruses) by reverse transcription.[1] The rate of replication is often elevated in malignant cells, making selective inhibition of replication a fundamental goal in cancer chemotherapy.[3] Since RNA:DNA hybrid structures play key roles in these biological processes, compounds that selectively target the hybrid structure would represent a new type of potential therapeutic agent.

Much effort is underway to rationally design small molecules that will selectively bind to unique nucleic acid structures such as triplexes,[4−6] tetraplexes,[7,8] and left-handed DNA.[9] However, few attempts have been made to target RNA:DNA hybrids. Inhibition of telomerase was attempted by using RNA or DNA oligonucleotides to target regions of the hTR RNA component of the enzyme, blocking the formation of the RNA:DNA hybrid structure that normally initiates telomere replication.[10] Bisdistamycins were explored as selective binding agents to a model Okazaki fragment that contained both DNA duplex and RNA:DNA hybrid regions.[11]

The discovery of small molecules that selectively bind to the RNA:DNA hybrid has been hampered by the lack of a convenient and rapid screening assay for compounds that might recognize the structure. We recently described a thermodynamically sound competition dialysis assay for rapidly screening structurally selective nucleic acid binding agents.[12−15] In the competition dialysis assay, 19 different nucleic acid structures (at identical concentration of 75 $\mu$M) are dialyzed against a common test ligand solution (at 1 $\mu$M concentration). Following equilibration, the amount of ligand bound to each nucleic acid structure is determined by absorbance or fluorescence measurements. Structures included in the assay range from single strands, through a variety of duplex forms, to multistranded triplex and tetraplex structures. Full details of the assay have been published.[12−15] Our assay offers a new perspective on the structural and sequence selectivity of ligand−nucleic acid interactions since it allows for simultaneous comparison of binding to a large number of competing structures under identical solution conditions. Described here is an application of the competition dialysis assay that allowed us to identify several compounds that preferentially bind to a particular RNA:DNA hybrid (poly rA:poly dT). These compounds represent potential lead molecules for the design and synthesis of even more selective agents.

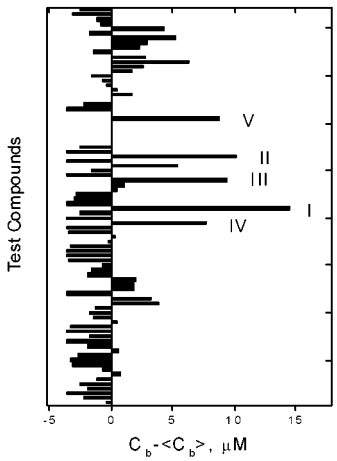


**Figure 1.** Ligand binding to poly rA:poly dT. Results for binding to the hybrid structure were culled from separate competition dialysis studies of 84 compounds. Data are presented as the difference between the amount of each compound bound and the average amount bound to the hybrid structure, $\langle C_b \rangle$. $\langle C_b \rangle$ is the arithmetic mean for the binding of all 84 compounds, and equals 3.7 ± 5.0 $\mu$M.

Results obtained for ellipticine using the competition dialysis assay are shown in Figure S1 (Supporting Information).[16] That data show that ellipticine binds preferentially to the poly rA:poly dT hybrid duplex. Binding to the RNA:DNA hybrid is nearly 2-fold greater than that to any duplex DNA structure in the assay. Ellipticine binding to the RNA:DNA hybrid represents a structural preference rather than a sequence preference for AT base pairs, since its binding to duplex DNA forms is characterized by a distinct GC base pair preference.

The assay allows rapid screening of large numbers of potential lead compounds. Figure 1 shows the binding of 84 different compounds to the poly rA:poly dT hybrid. These results were culled from complete competition dialysis assays (such as shown in Figure S1) done for each compound.[17] Figure 1 shows the difference between the amount of a specific compound bound and the average amount bound for all ligands. The plot emphasizes those compounds with higher than average binding affinity for the DNA:RNA hybrid structure. Figure 1 shows that a small subset of compounds binds to the hybrid structure in amounts that exceed one standard deviation from the mean. These compounds are ellipticine (**I**), ethidium (**II**), coralyne (**III**), propidium (**IV**), and TAS103 (**V**). The chemical structures of these compounds are shown in Figure 2. These compounds represent

* To whom correspondence should be sent: Telephone − (601) 984-1523 Facsimile − (601) 815-1171 Email − jchaires@biochem.umsmed.edu.
[†] University of Mississippi Medical Center.
[‡] Applied Gene Technologies, Inc.

(1) Mathews, C. K.; van Holde, K. E.; Ahern, K. G. *Biochemistry*; Benjamin/Cummings: San Francisco, 2000; Vol. 3, p 876ff.
(2) McEachern, M. J.; Krauskopf, A.; Blackburn, E. H. *Annu. Rev. Genet.* **2000**, *34*, 331.
(3) Pratt, W. R.; Ruddon, R.; Ensminger, W. D.; Maybaum, J. *The Anticancer Drugs*; Oxford University Press: Oxford, 1994; Vol. 2, pp 306−309.
(4) Mergny, J. L. et al. *Science* **1992**, *256*, 1681.
(5) Escude, C. et al. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 3591.
(6) Jenkins, T. C. *Curr. Med. Chem.* **2000**, *7*, 99.
(7) Perry, P. J.; Jenkins, T. C. *Exp. Opin. Invest. Drugs* **1999**, *8*, 1981.
(8) Han, H.; Hurley, L. H. *Trends Pharm. Sci.* **2000**, *21*, 136.
(9) Qu, X.; Trent, J. O.; Fokt, I.; Priebe, W.; Chaires, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12032.
(10) Glukhov, A. I.; Zimnik, O. V.; Gordeev, S. A.; Severin, S. E. *Biochem. Biophys. Res. Commun.* **1998**, *248*, 368.
(11) (a) Gmeiner, W. H., et al. *J. Biomol. Struct. Dyn.* **1999**, *17*, 507. (b). Gmeiner, W. H., et al. *Nucleosides, Nucleotides, Nucleic Acids* **2000**, *19*, 1365.
(12) Ren, J.; Chaires, J. B. *Biochemistry* **1999**, *38*, 16067.
(13) Ren, J.; Chaires, J. B. *J. Am. Chem. Soc.* **2000**, *122*, 424.
(14) Ren, J.; Bailly, C.; Chaires, J. B. *FEBS Lett.* **2000**, *470*, 355.
(15) Ren, J.; Chaires, J. B. *Methods Enzymol.* **2001**, *340*, in press.

(16) Full details of the competition dialysis method, including a detailed description of nucleic acid samples, may be found in refs 12−15. Solution conditions employed for the competition dialysis assay were the following: 6 mM $Na_2HPO_4$, 2 mM $NaH_2PO_4$, 1 mM $Na_2EDTA$, 0.185 M NaCl, pH 7.0.
(17) The structures of the 84 compounds used in this comparative study were provided to reviewers and are available from the authors upon request.

10.1021/ja015649y CCC: $20.00 
Published on Web 06/15/2001

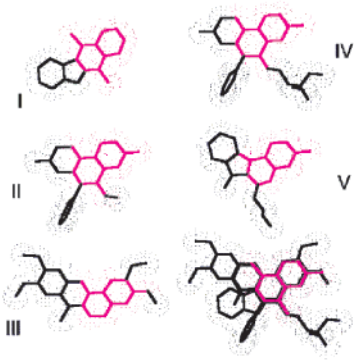


**Figure 2.** Structures of compounds with higher than average binding to poly rA:poly dT. The compounds are ellipticine (**I**), ethidium (**II**), coralyne (**III**), propidium (**IV**) and TAS103 (**V**). A common structural motif in these compounds is highlighted in color. The image in the lower right-hand corner is a manual superposition of compounds **I**−**IV**. The structures shown are energy minimized conformations obtained using molecular mechanics computations in the HyperChem software package, employing the $MM^+$ force field.[2]

(to the best of our knowledge) the first small molecules identified that uniquely recognize the poly rA:poly dT structure.

The three-dimensional structures of the compounds that preferentially bind to the hybrid structure share a common structural motif (Figure 2). This common motif is a planar aromatic ring system with a "bay" region, and can be clearly seen in the superposition of compounds **I**−**IV** in the lower right image in Figure 2. This common motif represents a possible pharmacophore for RNA:DNA hybrid recognition, a supposition that will clearly require more testing for verification.

The compounds that bind preferentially to the RNA:DNA hybrid inhibit an important enzymatic activity, digestion of the duplex by RNase H.[18] RNase H activity is necessary in DNA replication, and an RNase H activity is associated with all known reverse transcriptases (except telomerase). RNase H degrades the RNA stand of the hybrid. Figure 2S (Supporting Information) shows that ellipticine, ethidium, and propidium are all effective inhibitors of RNase H, presumably by their interaction with the DNA:RNA hybrid substrate and blocking of either enzyme binding or the cleavage reaction. Ellipticine is an effective inhibitor at micromolar concentrations, consistent with its tight binding to the hybrid structure.

Detailed binding studies[19] were done for ellipticine, ethidium, and propidium under solution conditions that support RNase H activity, with the results shown in Table 1 (Supporting Information). The binding constants for these compounds are directly proportional to the concentrations of each required to inhibit RNase H activity, as expected. The complete thermodynamic profiles for the binding of these compounds to the hybrid do not, however, reveal a common pattern. Binding of ellipticine to the hybrid is entropically driven, whereas binding of ethidium and propidium is driven by a large, negative enthalpy contribution. We assume that these three compounds bind by a common mode, most probably intercalation, but that supposition requires more direct testing by hydrodynamic methods.[20]

High-resolution structures for a number of model RNA:DNA hybrid structures were obtained by NMR and X-ray crystallography.[21−24] The general conclusion from these studies is that hybrids adopt unique secondary structures that are distinct from canonical DNA and RNA structures. The RNA strand of the hybrid generally adopts and A-form conformation while the DNA strand generally is similar to the B-form conformation. Major and minor groove widths and geometries in hybrid structures differ from those found in either duplex DNA or RNA. We cannot yet specify the structural basis for the molecular recognition of hybrid structures by the compounds discovered here. Our results point to the need for high-resolution structural and computational studies directed toward understanding a new type of structural recognition.

The results of the experiments described here identify several compounds that preferentially target a particular RNA:DNA hybrid duplex, a structure that is intimately involved in several important biological processes. These small molecules can effectively inhibit RNase H activity at micromolar concentrations. A potential pharmacophore was identified that represents a promising lead for the design of new agents for more stringent recognition of the DNA:RNA hybrid structure. Nucleic acids are polymorphic, and their unique structures offer unique opportunities for selective molecular recognition of conformations that participate in biological functions. Small molecules with such molecular recognition may prove effective as chemotherapeutic agents. A recent survey revealed that current pharmaceutics are directed toward only 500 cellular targets, and that only 2% of these targets are nucleic acids.[25] Our opinion is that the dearth of nucleic acid targets arises from a failure to appreciate the key functional roles of transient, unique DNA or RNA structures in gene replication and expression. The results shown here demonstrate that an appropriate search using an appropriate assay can reveal small molecules with novel molecular recognition capabilities.

**Acknowledgment.** Supported by NCI grant CA35635.

**Supporting Information Available:** Figures showing results of competition dialysis and RNase H assays, table of thermodynamic binding data, and table of structures used in the studies shown in Figure 2 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA015649Y

(18) RNase H (Promega, Madison, WI) activity was measured using absorbance measurements at 260 nm, following exactly the assay described by: Raschke, T. M.; Kho, J.; Marquese, S. *Nature Struct. Biol.* **1999**, *6*, 825. Activity was measured in a buffer containing 50 mM Tris, pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, at 20 °C. The substrate (poly rA:poly dT) concentration was 5 $\mu$M bp and the RNase H concentration was 81 nM.

(19) Complete thermodynamic profiles for ligand binding to poly rA:poly dT were determined by fluorescence titration and isothermal titration calorimetry, using experimental procedures developed in this laboratory that are fully described in the following: Qu, X.; Chaires, J. B. *Methods Enzymol.* **2000**, *321*, 353. Haq, I.; Jenkins, T. C.; Chowdhry, B. Z.; Ren, J.; Chaires, J. B. *Methods Enzymol.* **2000**, *323*, 373. All binding parameters were determined in a buffer containing 50 mM Tris, pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, at 20 °C.

(20) Suh, D.; Chaires, J. B. *Bioorg. Med. Chem.* **1995**, *3*, 723.
(21) Schmitz, U.; Blocker, F. J. H.; James, T. L. In *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press: New York, 1999; Chapter 8.
(22) Gyi, J. L.; Lane, A. N.; Conn, G. L.; Brown, T. *Biochemistry* **1998**, *37*, 73.
(23) Conn, G. L.; Brown, T.; Leonard, G. A. *Nucleic Acids Res.* **1999**, *27*, 555.
(24) Gmeiner, W. H., et al. *Biochemistry* **1999**, *38*, 1166.
(25) Drews, J. *Science* **2000**, *287*, 1960.