Kinetic and Thermodynamic Analysis of RNA Binding by Tethered Oligonucleotide Probes: Alternative Structures and Conformational Changes

Sharon T. Cload, Paul L. Richardson, Yaz-Hong Huang, and Alanna Schepartz*

Contribution from the Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Connecticut 06511-8118

Received January 4, 1993

Abstract: Binding of tethered oligonucleotide probes (TOPs) to the *Leptomonas collosoma* spliced leader RNA (SL RNA) was characterized using equilibrium, kinetic, and chemical structure-mapping methods. Equilibrium dissociation constants for TOP–SL RNA complexes are in the nanomolar range, with binding $\Delta G°_{obs}$ between –8.3 and –9.6 kcal mol^{-1} (10 mM $MgCl_2$, 10 mM NaCl, 0.1 mM EDTA, 25 °C). The highest SL RNA affinities are observed with TOPS **1** and **4**, which contain the shortest tethers. Formation of the complex between TOP **4** and the SL RNA is characterized by a second-order rate constant (k_{on}) of 2.2×10^3 M^{-1} s^{-1} (10 mM $MgCl_2$, 10 mM NaCl, 0.1 mM EDTA, 25 °C). Dissociation of the complex between TOP **4** and the SL RNA is characterized by a first-order rate constant (k_{off}) of 4.3×10^{-4} s^{-1} under similar conditions, corresponding to an average complex lifetime of 40 min. Chemical structure-mapping experiments performed on the uncomplexed SL RNA support a secondary structure model proposed recently by LeCuyer and Crothers. Chemical structure-mapping experiments performed on the SL RNA–TOP **1** complex demonstrate that TOP **1** shifts the SL RNA from the LeCuyer–Crothers secondary structure into a structure that is not well populated in the absence of TOP **1**.

The past decade has been marked by research directed toward the design and synthesis of compounds capable of binding with high affinity to specific DNA sequences[1,2] and conformations.[3] Despite considerable progress in this area, our understanding of the chemical principles that govern noncovalent interactions is still insufficient to allow the construction of a non-natural molecule whose precision in the recognition of DNA can rival that of a Watson–Crick complement.[4] Consequently, the search for selective modulators of gene expression at the DNA or RNA level has focused on the use of complementary oligonucleotides.[2,5,6] The use of complementary oligonucleotides for the recognition of RNA,[7-10] unfortunately, is exceedingly complicated.[11] Although formally single stranded, RNA molecules base pair intramolecularly, and the resultant stem loops can fold to generate molecules with greater structural complexity than duplex DNA.[12] The ability of individual RNA molecules to fold into stable, globular, three-dimensional structures that are essential to their function has been verified experimentally for the tRNA family[13] and for the *Tetrahymena* IVS,[14] and it is likely that many other RNAs exhibit the same property. The increased structural complexity that results from RNA folding complicates recognition because it limits the accessibility of long single-stranded regions that must be differentiated for specific binding.[15] Importantly, the functional significance of tertiary folding in RNA molecules may be exploited to provide a novel approach toward gene regulation: molecules that disrupt RNA structure should be extremely effective at inhibiting RNA function.[16]

Recently we reported two studies on a family of oligonucleotide-based molecules designed to recognize RNA in a way that acknowledges the complexity of RNA structure.[17,18] Tethered oligonucleotide probes (TOPs **1–5**, Figure 1) recognize RNA on the basis of both sequence and structure. Each TOP consists of two short oligonucleotides joined by a tether whose length and structure may be controlled by chemical synthesis. The two oligonucleotides in each TOP hybridize to two noncontiguous regions on the RNA, and the tether traverses the distance between them (Figure 2). Because TOPs should in theory be capable of binding simultaneously to two short single-stranded regions in a target RNA, rather than one long region which may not be fully

* To whom correspondence should be addressed.

(1) Egholm, M.; Burchardt, O.; Nielsen, P. E.; Berg, R. *Science* **1991**, *254*, 1497. Helene, C.; Thuong, N. T. *Nucleic Acids Mol. Biol.* **1988**, *2*, 105. Dervan, P. B. *Science* **1986**, *232*, 464. Sigman, D. S. *Acc. Chem. Res.* **1986**, *19*, 180.

(2) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543.

(3) Pyle, A. M.; Barton, J. K. *Prog. Inorg. Chem.* **1990**, *38*, 413.

(4) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

(5) Maher, L. J., III; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725.

(6) Francois, J.-C.; Saison-Behmoaras, T.; Thuong, N. T.; Helene, C. *Biochemistry* **1989**, *28*, 9617.

(7) Brown, R. S.; Dewan, J. C.; Klug, A. *Biochemistry* **1985**, *24*, 4785.

(8) Carter, B. J.; de Vroom, E.; Long, E. C.; van der Marel, G. A.; van Boom, J. H.; Hecht, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9373.

(9) Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839. Chow, C. S.; Hartmann, K. M.; Rawlings, S. L.; Huber, P. W.; Barton, J. K. *Biochemistry* **1992**, *31*, 3534. Chow, C. S.; Barton, J. K. *Biochemistry* **1992**, *31*, 5423. Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972.

(10) Zuckerman, R. N.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1766.

(11) Ecker, D. J.; *et al. Science* **1992**, *257*, 958.

(12) Draper, D. E. *Acc. Chem. Res.* **1992**, *25*, 201. Tinoco, I., Jr.; Puglisi, J. D.; Wyatt, J. R. *Nucleic Acids Mol. Biol.* **1990**, *4*, 205. Wyatt, J. R.; Puglisi, J. D.; Tinoco, I., Jr. *BioEssays* **1989**, *11*, 100. Tinoco, I., Jr.; Davis, P. W.; Hardin, C. C.; Puglisi, J. D.; Walker, G. T.; Wyatt, J. *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *87*, 135.

(13) Kim, S. H.; Sussman, J. L.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Wang, A. H. J.; Seeman, N. C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4970. Quigley, G. J.; Rich, A. *Science* **1976**, *194*, 796. Woo, N. H.; Roe, B. A.; Rich, A. *Nature* **1980**, *286*, 346. Westof, E.; Dumas, P.; Moras, D. *J. Mol. Biol.* **1985**, *184*, 119. Rould, M. A.; Perona, J. J.; Soll, D.; Steitz, T. A. *Science* **1989**, *246*, 1135. tRNA NMR: Patel, D. J.; Shapiro, L.; Hare, D. *Q. Rev. Biophys.* **1987**, *20*, 25. Cheong, C.; Varani, G.; Tinoco, I., Jr. *Nature* **1990**, *346*, 680. Puglisi, J. D.; Wyatt, J. R.; Tinoco, I., Jr. *Biochemistry* **1990**, *29*, 4215. Varani, G.; Wimberly, B.; Tinoco, I., Jr. *Biochemistry* **1989**, *28*, 7760.

(14) Latham, J. A.; Cech, T. R. *Science* **1989**, *245*, 276. Celander, D. W.; Cech, T. R. *Biochemistry* **1990**, *29*, 1355. Wang, J. F.; Cech, T. R. *Science* **1992**, *256*, 526.

(15) Kean, J. M.; Murakami, A.; Blake, K. R.; Cushman, C. D.; Miller, P. S. *Biochemistry* **1988**, *27*, 9113.

(16) Salunkhe, M.; Wu, T.; Letsinger, R. L. *J. Am. Chem. Soc.* **1992**, *114*, 8768.

(17) Richardson, P.; Schepartz, A. *J. Am. Chem. Soc.* **1991**, *113*, 5109.

(18) Cload, S. T.; Schepartz, A. *J. Am. Chem. Soc.* **1991**, *113*, 6324.

0002-7863/93/1515-5005$04.00/0

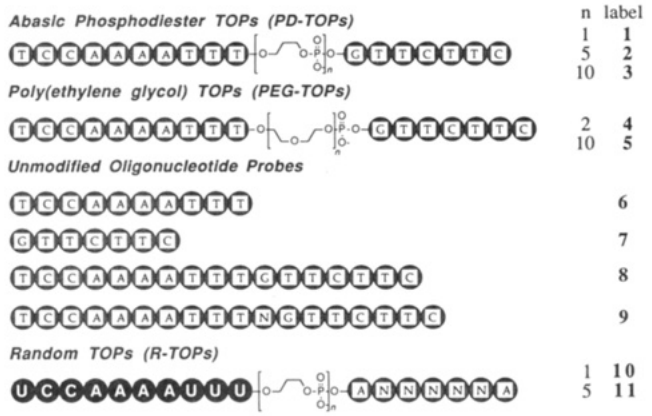

Figure 1. Structures of TOPs and other oligonucleotides used in this study. The 5′-terminus of each probe is shown on the left. TOPs **1–9** contain deoxyribonucleotides (shaded circles). TOPs **10** and **11** contain both deoxyribonucleotides and ribonucleotides (solid circles) as indicated. N represents an approximate equimolar mixture of all four possible bases (see text).

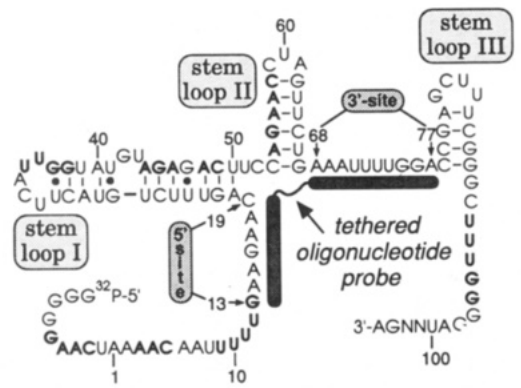

Figure 2. Secondary structure of the *L. collosoma* SL RNA showing originally proposed base-pairing pattern and the locations of the 3′- and 5′-sites. Residues in boldface type denote sequences that are partially complementary to TOPs **1–5**.

accessible, we predicted that TOPs might display very high RNA affinities and specificities.

Two classes of TOPs have been described (Figure 1).[17,18] One class is characterized by a tether composed of a flexible poly(oxyphosphinicooxy-1,3-propanediol) backbone[19] designed to mimic the backbone of natural DNA, while the other is characterized by a neutral poly(ethylene glycol) tether designed to minimize possible electrostatic effects.[20–22] Both TOP classes were designed to recognize the spliced leader RNA (SL RNA) from *Leptomonas collosoma*, a noninfectious trypanosomatid closely related to the infectious *Trypanosoma brucei* and *Trypanosoma cruzi*.[23,24] TOPs **1–5** contain the DNA sequences that complement two noncontiguous regions of the SL RNA secondary structure drawn in Figure 2: the ten bases spanning A68–A77 (3′-site) and the seven bases spanning G13–C19 (5′-site). Although TOPs **1–5** are partially complementary to several other regions of the SL RNA (shown in bold type in Figure 2), RNase H mapping experiments indicate that TOPs **1–5** bind only their targeted sites to form monomeric complexes when the SL RNA is in excess and that the two oligonucleotides within each TOP bind simultaneously.[17,18] Herein we describe experiments that evaluate quantitatively the kinetics and thermodynamics of TOP binding to the SL RNA. Our initial results prompted us to perform chemical structure-mapping experiments to evaluate the effect of TOP binding on the structure of the SL RNA itself. These experiments revealed that, in the absence of TOP **1**, the secondary structure drawn in Figure 2 does not accurately represent the base-pairing pattern of the 5′-half of the SL RNA. Remarkably, TOP **1** is capable of shifting the 5′-half of the SL RNA from its preferred secondary structure into the one shown in Figure 2 without a large sacrifice in K_d. This suggests that TOPs might represent a general strategy for controlling the folded structure of a globular RNA.

Experimental Section

Synthesis of Tethered Oligonucleotide Probes. Oligonucleotides were synthesized using the solid-phase phosphoramidite method[25] on an Applied Biosystems 380B Synthesizer. The syntheses of TOPs **1–5** have been described previously.[17,18] Deoxyoligonucleotides were cleaved from the solid support (NH_4OH, 1 h, 25 °C), deprotected (NH_4OH, 8 h, 55 °C), and purified by use of a denaturing 20% (20:1 cross-linked) polyacrylamide gel.[26] Randomized tethered oligonucleotides (R-TOPs) **10** and **11** were prepared by reacting equimolar mixtures of the four nucleoside phosphoramidites during the appropriate synthesis cycles. Ribonucleoside phosphoramidites were purchased from Peninsula. Deoxyribonucleoside phosphoramidites were purchased from ABI. R-TOPs were cleaved from the solid support (3:1 EtOH/NH_4OH, 1 h, 25 °C), deprotected (3:1 EtOH/NH_4OH, 8 h, 55 °C), desilylated (1M (n-Bu)$_4$NF, THF, 6 h, 55 °C), and purified by use of a denaturing 20% (20:1 cross-linked) polyacrylamide gel.[26] Oligonucleotide homogeneity was assessed by high-resolution gel electrophoresis of 5′-end-labeled ([γ-^{32}P]ATP, T4 polynucleotide kinase) material.[27] All TOPs were >95% homogeneous by this criterion.

Preparation of SL RNA. SL RNA was prepared by run-off transcription of a *Bam* H1 linearized plasmid[28] and purified by gel electrophoresis (15% acrylamide, 20:1 cross-linked, 7 M urea). SL RNA was dephosphorylated using calf intestinal phosphatase (Boeringer Mannheim) and labeled on the 5′-end with [γ-^{32}P]ATP using T4 polynucleotide kinase. The sequence was confirmed by enzymatic[29] and dideoxy[30,31] sequencing.

Determination of Dissociation Constants. Equilibrium dissociation constants of the TOP–SL RNA complexes were measured by use of an electrophoretic mobility shift assay.[32–34] In a typical procedure, SL RNA (0.0025–2.4 μM in 1.5X incubation buffer) was heated at 75–80 °C for 5 min and cooled to 25 °C over 30 min before the addition of ~0.2 nM 5′-end-labeled TOP dissolved in water. Binding reactions were incubated at 25 °C in incubation buffer (50 mM Tris pH 7.5, 10 mM $MgCl_2$, 10 mM NaCl, 0.1 mM EDTA, 0.05% xylene cyanole, 5% sucrose, final pH 7.3) and loaded immediately onto a 0.75 × 140 × 150 mm 10% (107:1 cross-linked) polyacrylamide gel prepared with running buffer (33.3 mM Tris base, 66.7 mM HEPES acid, 10 mM $MgCl_2$, 0.1 mM EDTA, pH 7.5). Gels were maintained at 3–4 °C during electrophoresis by immersion in a buffer cooled by a circulating, temperature controlled water bath. The amounts of complexed and free TOP were quantified on a Betagen 605 Blot Analyzer (Betagen Inc., Waltham, MA). The time necessary to establish equilibrium between a TOP and the SL RNA was determined for each TOP by plotting the fraction of TOP bound versus time. At 25

(19) Seela, F.; Kaiser, K. *Nucleic Acids Res.* **1987**, *15*, 3113.
(20) Letsinger, R. L.; Singman, C. N.; Histand, G.; Salunkhe, M. *J. Am. Chem. Soc.* **1988**, *110*, 4470.
(21) Miller, P. S.; Agris, C. H.; Aurelian, L.; Blake, K. R.; Murakami, A.; Reddy, M. P.; Spitz, S. A.; Ts'o, P. O. P. *Biochimie* **1985**, *67*, 769.
(22) Durand, M.; Chevrie, K.; Chassignol, M.; Thuong, N. T.; Maurizot, J. C. *Nucleic Acids Res.* **1990**, *18*, 6353.
(23) Milhausen, M.; Nelson, R. G.; Sather, S.; Selkirk, M.; Agabian, N. *Cell* **1984**, *38*, 721.
(24) Bruzik, J. P.; Van Doren, K.; Hirsh, D.; Steitz, J. A. *Nature* **1988**, *335*, 559.
(25) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859.
(26) Maniatis, T.; Fritsch, E. F.; Sambrook, B. *Molecular Cloning*; Cold Spring Harbor: Plainview, NY, 1982.
(27) Vournakis, J. N.; Celantano, J.; Finn, M.; Lockar, R. E.; Mitra, T.; Pavlakis, G.; Troutt, A.; van den Berg, M.; Wurst, R. M. *Gene Amplif. Anal.* **1981**, *2*, 267.
(28) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783.
(29) Donis-Keller, H. *Nucleic Acids Res.* **1980**, *8*, 3133.
(30) Sanger, F.; Nicklen, S.; Coulsen, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5463.
(31) Krol, A.; Carbon, P. *Methods Enzymol.* **1989**, *180*, 212.
(32) Fried, M.; Crothers, D. M. *Nucleic Acids Res.* **1981**, *9*, 6505.
(33) Garner, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047.
(34) Pyle, A. M.; McSwiggen, J. A.; Cech, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8187. Pyle, A. M.; Cech, T. R. *Nature* **1991**, *350*, 628.

Table I. Equilibrium Dissociation Constants of SL RNA–TOP Complexes at 25 °C[a]

probe	K_d (nM)	$\Delta G°_{obs}$ (kcal mol^{-1})[b]
PD-TOP **1**	99 ± 6[c]	−9.5
PD-TOP **2**	263 ± 11	−8.9
PD-TOP **3**	833 ± 75	−8.3
PEG-TOP **4**	88 ± 9	−9.6
PEG-TOP **5**	276 ± 25	−8.9
6	22 200 ± 2200[c]	−6.3
7	112 700 ± 2250[c]	−5.4
8	171 ± 10	−9.2
9	238 ± 12	−9.0

[a] Experimental protocols described in text. [b] Calculated from the relationship $\Delta G°_{obs} = -RT \ln (1/K_d)$ where $R = 0.00198$ kcal mol^{-1} K^{-1} and $T = 298$ K. [c] Error represents the relative standard error of estimate $(S_y) = 100[\sum(\Theta_T - \Theta_E)^2/n]^{1/2}/\langle\Theta_E\rangle$ where Θ_T and Θ_E are the fractions bound as determined from theory and experiment, respectively, and n is the number of data points.[35] [c] Estimated.

°C, TOPs **1**–**5** required a 60–80-min incubation with the SL RNA to reach equilibrium. Therefore, TOPs were incubated with the SL RNA for 2 h prior to nondenaturing gel electrophoresis.

Careful analysis of autoradiograms revealed that SL RNA–TOP complexes migrate as two very tightly spaced bands of approximately equal intensity upon nondenaturing gel electrophoresis. The mobility of the doublet is nearly identical to the mobility of monomeric 5′-end-labeled SL RNA (also a tight doublet) under these conditions and significantly greater than the mobility of an SL RNA dimer formed at concentrations higher than 3 μM. Since the SL RNA migrates as a single band upon denaturing gel electrophoresis, the appearance of two bands under native conditions signals the existence of alternative SL RNA conformations. Both conformations exhibit approximately equal affinities for the TOPs. These two bands may represent alternative conformations of stem loop II or subtle changes in tertiary structure (vide infra). There is a large difference in mobility between the tight doublet and the SL RNA dimer, indicating that TOPs do not cross-link two SL RNA molecules when they bind.

K_d values were determined from best fits of the data to the binding curve defined by the equation

$$\Theta = [\text{SL RNA}]_T/[K_d + [\text{SL RNA}]_T] \quad (1)$$

where Θ = fraction of TOP bound = cpm bound TOP/(cpm bound TOP + cpm free TOP). Radioactivity was quantified using a Betagen 605 Blot Analyzer (Betagen Co., Waltham, MA). All TOPs bound to >87% completion and fit the binding equation with errors $(S_y) < 10\%$. The relative standard error of estimate $(S_y) = 100[\sum(\Theta_T - \Theta_E)^2/n]^{1/2}/\langle\Theta_E\rangle$ where Θ_T and Θ_E are the fractions bound as determined from theory and experiment, respectively, and n is the number of data points.[35] The K_d obtained varied by <15% from day to day (except for **3**, which varied by 21%). The K_ds for **6** and **7** were too high to measure directly and were estimated by competition: SL RNA (0.3 μM in 2× incubation buffer) was renatured as above before the addition of ~0.2 nM 5′-end-labeled **5** and 0–24 μM unlabeled **6** or 0–50 μM unlabeled **7** dissolved in water. After 2 h the components were resolved by nondenaturing gel electrophoresis and the amount of complexed and free **5** was quantified as above. K_d values for **6** and **7** were estimated from best fits of the data to the equation[36]

$$\Theta = K_{RO} + \frac{K_{RO}}{K_{RT}}T_T + R_T + O_T - \left[\left(K_{RO} + \frac{K_{RO}}{K_{RT}}T_T + R_T + O_T\right)^2 - 4O_TR_T\right]^{1/2}/2O_T \quad (2)$$

Θ represents the fraction of TOP **5** bound, and R_T, O_T, and T_T represent respectively the total concentration of SL RNA, TOP **5**, and **6** or **7**. K_{RT} represents the dissociation constant for the SL RNA–TOP **5** complex and K_{RO} represents the dissociation constant for the SL RNA complex with **6** or **7**. Curve fitting was performed using Kaleidagraph 2.1.2 (Abelbeck Software, Reading, PA). Dissociation constants estimated by nonlinear least squares analysis of the data are shown in Table I.

(35) Devore, J. L. *Probability and Statistics for Engineering and the Sciences*; Brooks/Cole: Monterey, CA, 1987.

(36) Lin, S.-Y.; Riggs, A. D. *J. Mol. Biol.* **1972**, *72*, 671. Weeks, K. M.; Crothers, D. M. *Biochemistry* **1992**, *31*, 10281.

Kinetic Analysis of Binding. All reactions were performed at 25 °C in a binding buffer composed of 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM NaCl, 0.1 mM EDTA, and 5% sucrose, final pH 7.3. The SL RNA used in kinetic experiments was renatured by heating at 75–80 °C in 1.5× binding buffer for 4 min followed by cooling to 25 °C over 30 min. The tethered probe used in kinetic experiments, TOP **4**, was 5′-phosphorylated with [γ-^{32}P] ATP (New England Nuclear) and T4 polynucleotide kinase (New England Biolabs). Binding reactions were analyzed on 16 × 14 × 0.15 cm 10% (107:1 cross-linked polyacrylamide) gels run at 300 V for 1–2 h. Gels were maintained at 3–6 °C during electrophoresis by immersion in a buffer cooled by a circulating, temperature-controled water bath. The electrophoresis buffer contained 33.3 mM Tris base, 66.7 mM HEPES acid pH 7.5, 10 mM $MgCl_2$, and 0.1 mM EDTA. Ratios of free and bound TOP were quantified using a Betagen Blot Analyzer.

Tethered Oligonucleotide Probe Association Reactions. Association reactions were composed of 50–400 nM SL RNA and approximately 2 nM TOP **4** and were initiated by combining the two components in binding buffer at 25 °C. After various times, the association reactions were halted upon addition of a 10 000-fold excess of unlabeled TOP **4** (final concentration 20 μM). Quenched reactions were incubated an additional minute at 25 °C, then applied directly to a running 10% polyacrylamide gel. Although the quenching procedure prevents formation of additional ^{32}P-labeled TOP–SL RNA complex, it does not cause significant dissociation of the pre-existing complex during the time necessary (<1 min) for the complex to enter the gel matrix. As described below, at [SL RNA] < 400 nM and in the presence of 20 μM unlabeled TOP **4**, $k_{off} = 4.7 \times 10^{-4}$ s^{-1}. Thus the average solution lifetime of the SL RNA–TOP complex $(1/k_{off})$ is approximately 40 min. Association reactions were performed at least twice at each SL RNA concentration.

Tethered Oligonucleotide Probe Dissociation Reactions. Dissociation rates were measured at [SL RNA_0] = 50, 100, 200, and 400 nM. SL RNA and 2 nM TOP **4** were incubated at 25 °C in binding buffer for a minimum of 90 min to allow the binding reaction to reach equilibrium. Dissociation was initiated by addition of unlabeled TOP **4** to a final concentration of 20 μM. After various times, the reaction mixtures were applied directly to a running 10% polyacrylamide gel, as described under Association Reactions. Dissociation reactions were performed at least twice at each RNA concentration.

Chemical Modification Experiments. All reactions contained SL RNA (150 nM) and RNasin (0.1 unit/μL, Promega). Modification reactions were performed in either CMCT buffer (50 mM sodium borate, 10 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, pH 8.0), DMS buffer (200 mM HEPES–KOH, 10 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, pH 8.0), or kethoxal buffer (160 mM potassium cacodylate, 10 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, pH 7.2). In a typical procedure, SL RNA (300 nM) was pre-equilibrated with or without 0.15–15 μM TOP **1** or 1.5 mM **6** or **7** in 100 μL of 2× buffer for 2 h at 25 °C prior to the addition of the modifying reagent. CMCT (1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide *p*-toluenesulfonate, Sigma) (42 mg/mL dissolved fresh in sterile water), neat DMS (dimethyl sulfoxide, Sigma), or kethoxal (2-keto-3-ethoxybutyraldehyde, ICN Biochemicals) (37 mg/mL dissolved fresh in 20% ethanol/water) was then added to give a final concentration of 28 mM, 120 μM, or 28 mM, respectively, upon dilution to 200 μL. The reactions were mixed gently and incubated at 25 °C for either 10 min (CMCT and DMS) or 30 min (kethoxal). The DMS reactions were halted by the addition of 50 μL of DMS stop solution (1 M sodium acetate, 1 M Tris–acetate, 1 M β-mercaptoethanol, pH 7.5) followed by incubation on ice for 10 min. The kethoxal reactions were stopped by the addition of 16 μL of 0.5 M potassium borate, pH 7.0, followed by 50 μL of 1 M sodium acetate, pH 7.0. CMCT reactions were halted by the addition of 50 μL of 1 M sodium acetate, pH 7.0. The RNA was precipitated with 3 volumes of ethanol, washed with cold 70% ethanol, and dried *in vacuo*. The pellets were dissolved in 100 μL of 200 mM NaCl, 25 mM potassium borate, pH 7.0, and extracted twice with 24:1 chloroform/isoamyl alcohol. The reactions were precipitated twice with ethanol, washed with 70% ethanol, and dried.

The extent of SL RNA modification was then analyzed using reverse transcriptase analysis.[31,37] A synthetic 5′-end-labeled primer (ATCCCCAAAGCCCGA, 100 000 cpm/μL, ~33 nM) complementary to SL RNA nucleotides U86–U100 was annealed to the modified RNA (2.5 μM) in 25 mM potassium borate, pH 7.0, by heating the mixture to 80 °C for 1 min and allowing it to cool slowly to 25 °C over 10 min

(37) Moazed, D.; Stern, S.; Noller, H. F. *J. Mol. Biol.* **1986**, *187*, 399. Inoue, T.; Cech, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 648.

in a final volume of 6-μL. Primer extension was performed at 37 °C for 45 min in reverse transcriptase buffer (50 mM Tris-HCl, 8 mM $MgCl_2$, 30 mM KCl, 50 mM DTT, 5 mM potassium borate, pH 7.0) containing avian myeloblastosis virus reverse transcriptase (0.2 unit/μL, Promega), RNasin (0.4 unit/μL, Promega), and 170 μM each dNTP. The reactions were quenched with NaCl (to 200 mM) and carrier tRNA (to 0.125 mg/mL). The reactions were ethanol precipitated twice, washed with 70% ethanol, dried, and analyzed by high-resolution gel electrophoresis.

Assay for Binding Specificity of Random TOPs. RNase H hydrolyzes RNA contained in RNA–DNA hybrids.[38] Therefore, incubation of RNase H with 5′-end-labeled SL RNA and a TOP permits the TOP binding sites to be determined when the reaction products are subjected to high-resolution gel electrophoresis and autoradiography.[17,18] RNase H reactions (15 μL) contained 13 nM 5′-end-labeled SL RNA, 1.3–13 μM R-TOP, 4 units of RNasin (Promega), and 0.8 units of RNase H (Pharmacia) in RNase H buffer (57 mM Tris-HCl pH 8.0, 12.7 mM $MgCl_2$, 1.6 mM DTT, 135 mM KCl, 0.9 mM EDTA, 0.05 μg/μL BSA). In a typical procedure, SL RNA was renatured in 2.5× RNase H buffer by heating to 70 °C and cooling to 25 °C over a 30-min period before addition of R-TOP, RNase H, and RNasin. Reactions were incubated at 25 °C for 2 h and halted by the addition of EDTA to a final concentration of >17 mM. Cleavage products were resolved on a denaturing 15% (20:1 cross-linked) polyacrylamide gel alongside enzymatic sequencing reactions and visualized by autoradiography. The RNase H digestion products contain 3′-hydroxy termini[39] and do not have the same mobility as enzymatic or alkaline digestion products containing 3′-phosphate termini. By excising a single RNase H cleavage product, subjecting it to alkaline hydrolysis, and resolving the products on a high-resolution gel, it was established that an RNase H product migrates slightly slower than an enzymatic sequencing product that is one nucleotide longer.

Competition of Random-TOP Binding. End-labeled SL RNA (13 nM, 100 000 cpm) was incubated for 2 h with RNase H (0.8 units), R-TOP **10** or **11** or the oligonucleotide ANNNNNA (13 μM), and 6.5 or 65 μM r(UCCAAAAUUU) (**12**) in 15 μL of RNase H buffer. The R-TOP and competitor oligonucleotide were added simultaneously and the reactions performed as described.[17,18]

Results

TOPs **1–5** (Figure 1) contain the DNA sequences designed to complement two noncontiguous, single-stranded regions of the SL RNA secondary structure drawn in Figure 2: the ten bases spanning A68–A77 (3′-site) and the seven bases spanning G13–C19 (5′-site). This secondary structure was proposed on the basis of a relative free energy calculation carried out with the FOLD program[40] and the observation of a conserved folding pattern among several different SL RNAs.[24] Single-site probes **6** and **7** are complementary to only the 3′-site or the 5′-site, respectively. Oligodeoxynucleotides **8** and **9** contain the 3′- and 5′-site-specific oligonucleotides separated by a single phosphodiester or nucleotide, respectively. R-TOPs **10** and **11** contain the RNA sequence complementary to the 3′-site (A68–A77) joined through one or five abasic phosphodiester monomers to the DNA sequence ANNNNNA, where N represents an approximate equimolar mixture of all four possible nucleotides. Thus each random TOP represents a pool of 1024 different molecules. All oligonucleotides were synthesized on an Applied Biosystems automated oligonucleotide synthesizer.

Dissociation Constants of SL RNA–TOP Complexes. Equilibrium dissociation constants of the SL RNA–TOP complexes were measured in order to quantify the effects of bivalent hybridization[41] and the influence of tether length and structure on SL RNA recognition. Binding curves representing association of TOPs **1–5** with the SL RNA at 25 °C are shown in Figure 3, and K_ds are tabulated in Table I. In each case, a plot of the fraction TOP bound versus SL RNA concentration provides a good fit to an equation (1) that describes formation of a 1:1 complex. Examination of the K_ds reveal three trends: (1) The

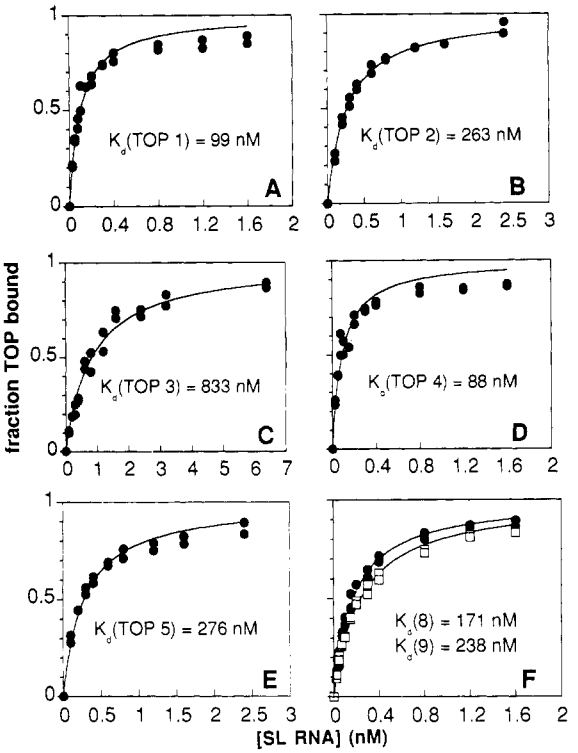

Figure 3. Equilibrium binding of TOPs **1–5** and **8** (●) and **9** (□) to the *L. collosoma* SL RNA. Solid lines represent the best fit of the data to eq 1, which describes binding one TOP or oligonucleotide to a single molecule of SL RNA and saturates at 100% bound, as described in the Experimental Section. Note that the *x*-axis varies throughout the figure.

SL RNA affinities of the anionic and neutral TOPs are similar. PD-TOP **2** and PEG-TOP **5** contain tethers of approximately equal length, but the tether in PD-TOP **2** is a polyanion and that in PEG-TOP **5** is uncharged. The two molecules, however, exhibit equal affinities for the SL RNA, with K_ds of ~270 nM under these conditions. (2) Binding affinity depends modestly on tether length. $\Delta G°_{obs}$ becomes less favorable by 0.6–0.7 kcal mol^{-1} as the maximum tether length increases from ~7 to ~35 Å (compare PD-TOPs **1** and **2** and PEG-TOPs **4** and **5**). $\Delta G°_{obs}$ becomes less favorable by 0.6 kcal mol^{-1} as the tether length is increased to 70 Å in TOP **3**. (3) TOPs **1** and **4**, with ~7-Å tethers, exhibit the highest affinities for the SL RNA, with $\Delta G°_{obs}$ of −9.5 and −9.6 kcal mol^{-1} at 25 °C, respectively. These values are approximately 3 kcal mol^{-1} more favorable than the $\Delta G°_{obs}$ estimated for single-site probe **6** ($\Delta G°_{obs}$ = −6.5 kcal mol^{-1}) or **7** ($\Delta G°_{obs}$ = −5.3 kcal mol^{-1}), neither of which is capable of binding the 3′- and 5′-sites simultaneously. Not only do TOPs **1** and **4** display a higher SL RNA affinity than single-site probes **6** and **7**, they also display a 2-fold higher affinity than oligonucleotides consisting of the 3′- and 5′-site-specific oligonucleotides linked directly (as in **8**, K_d = 171 ± 5 nM, $\Delta G°_{obs}$ = −9.2 kcal mol^{-1}) or through a single nucleotide (as in **9**, K_d = 238 ± 6 nM, $\Delta G°_{obs}$ = −9.0 kcal mol^{-1}).

TOP Association Kinetics. A reversible bimolecular binding reaction is described by the following equilibrium expression:

$$R + T \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} RT \qquad (3)$$

R, T, and RT refer to the concentration of free SL RNA, free TOP, and SL RNA–TOP complex, respectively. The kinetic constants k_{on} and k_{off} represent the second-order association rate constant and first-order dissociation rate constant, respectively. We emphasize that the binding reaction is likely to proceed in several steps. Using our gel electrophoresis assay, we can detect only the TOP and the SL RNA–TOP complex. A complete

(38) Donis-Keller, H. *Nucleic Acids Res.* **1979**, *7*, 179.
(39) Miller, H. I.; Riggs, A. D.; Gill, G. N. *J. Biol. Chem.* **1973**, *248*, 2621.
(40) Zucker, M. *Science* **1989**, *244*, 48.
(41) Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4046.

understanding of the microscopic binding steps leading to the complex shown in Figure 2 will require alternative methods.[42] In the presence of excess SL RNA relative to TOP, the SL RNA concentration remains effectively constant during the course of the association reaction and may be approximated by the initial SL RNA concentration, $[R_0]$. Under these conditions the kinetics of reversible complex formation are described by the integrated rate equation

$$\ln\left[\frac{[F_t] - [F_{eq}]}{[F_0] - [F_{eq}]}\right] = -k_{obs}t \qquad (4)$$

where

$$k_{obs} = k_{on}[R_0] + k_{off} \qquad (5)$$

The fraction of free TOP (F) at any given time is obtained from the cpm in the band corresponding to free TOP divided by the sum of the cpm in the bands corresponding to free and bound TOP. F_0, F_t, and F_{eq} represent the fraction of free TOP at time zero, at time t, and at equilibrium, respectively. The observed rate of association at each $[R_0]$, k_{obs}, is obtained from the negative slope of a plot of $\ln(([F_t] - [F_{eq}])/([F_0] - [F_{eq}]))$ versus time. A plot of k_{obs} versus $[R_0]$ provides values for both k_{on} and k_{off}.

To measure the association rate of the TOP–SL RNA complex, freshly renatured SL RNA (50–400 nM) and labeled TOP **4** (2 nM) were mixed in binding buffer and incubated at 25 °C. Binding reactions were quenched at appropriate time intervals with excess unlabeled TOP **4**. Bound and free labeled TOP were then resolved by nondenaturing gel electrophoresis and quantified. Figure 4A illustrates the results of experiments with $[R_0]$ = 50, 67, 83, 100, 150, 200, and 400 nM and [TOP] = 2 nM plotted as the fraction of TOP **4** bound ($1 - F_t$) versus time. Figure 4B shows the first-order plots that were used to obtain k_{obs}. The value of F_{eq} at each $[R_0]$ was calculated from the K_d of 90 nM measured for the SL RNA–TOP **4** complex (Table I). The linearity of the plots in Figure 4B demonstrates that the experimental data are consistent with a bimolecular complexation mechanism. A value for k_{on} = $(2.2 \pm 0.2) \times 10^3$ M^{-1} s^{-1} is obtained from the slope of the plot of k_{obs} versus $[R_0]$ in Figure 4C. This plot is linear between $[R_0]$ concentrations of 100 and 400 nM. The intercept at $[R_0] = 0$ in Figure 4C is equal to $k_{off} = (3.6 \pm 0.3) \times 10^{-4}$ s^{-1}. This value agrees well with the value for $k_{off} = (4.3 \pm 1.0) \times 10^{-4}$ s^{-1} (open symbol at $[R_0] = 0$) determined in separate experiments described below, suggesting that the deviation from linearity seen at low [SL RNA] is not significant.

TOP Dissociation Kinetics. An irreversible unimolecular dissociation reaction is described by the expression

$$RT \xrightarrow{k_{off}} R + T \qquad (6)$$

where k_{off} represents the rate constant for dissociation of the complex RT. The kinetics of such a dissociation reaction are described by the integrated first-order rate equation

$$\ln\left[\frac{(1 - F_t)}{(1 - F_0)}\right] = -k_{off}t \qquad (7)$$

where F_0 and F_t represent the fraction of free TOP at times zero and t, respectively.

To determine the rate constant for dissociation of the TOP **4**–SL RNA complex, we used a technique similar to that described for the determination of the association rate constant. SL RNA–TOP **4** complexes were formed at a molar ratio of between 25:1 and 200:1 in binding buffer. Dissociation was initiated by addition of excess unlabeled TOP **4**. Figure 5A shows the results of experiments with $[SL\ RNA_0]$ = 50, 100, 200, and 400 nM plotted as the fraction of free TOP, F_t, versus time. Figure 5B shows the first-order plots of $-\ln((1 - F_t)/(1 - F_0))$ versus time that were used to obtain k_{off}. The plots are linear, consistent with a mechanism involving first-order dissociation of the SL RNA–TOP complex. Table II lists the values of k_{off} measured at each $[SL\ RNA_0]$. The similarity of the measured values at [SL RNA] < 400 nM indicates that the dissociation rate is independent of SL RNA concentration in this range. The average value of $(4.3 \pm 1.0) \times 10^{-4}$ s^{-1} agrees well with the k_{off} value determined from the intercept of a plot of the observed association rate constant k_{obs} versus [SL RNA], $(3.6 \pm 0.3) \times 10^{-4}$ s^{-1} (Figure 4C). Thus the complex of the SL RNA with TOP **4** has an average solution lifetime of 40 min under our experimental conditions.

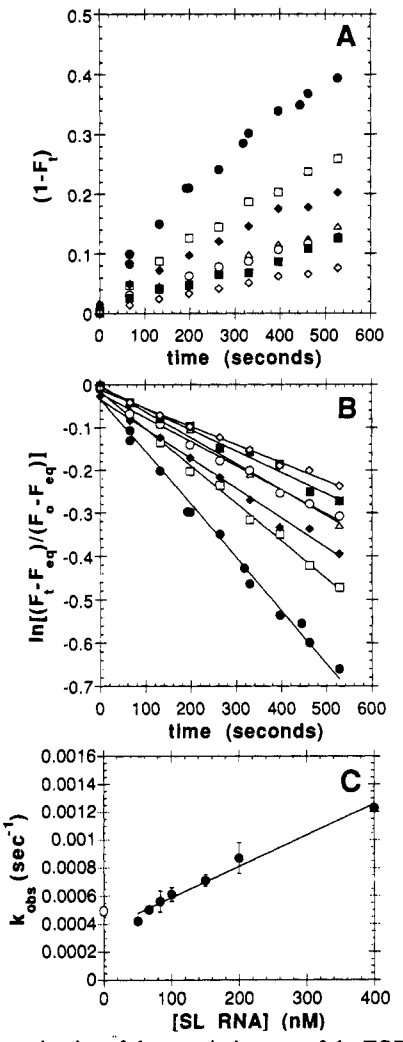

Figure 4. Determination of the association rate of the TOP **4**–SL RNA complex. (**A**) Plot of $(1 - F_t)$, the fraction TOP bound as a function of time, t, in the presence of $[SL\ RNA_0]$ = 50 (◇), 67 (■), 83 (○), 100 (△), 150 (◆), 200 (□), and 400 (●), nM. (**B**) Determination of k_{obs}. Plot of $\ln[(F_t - F_{eq})/(F_0 - F_{eq})]$ versus time for $[SL\ RNA_0]$ = 50 (◇), 67 (■), 83 (○), 100 (△), 150 (◆), 200 (□), and 400 (●) nM. F_{eq} was calculated from the concentration of SL RNA and the K_d (TOP **4**–SL RNA) = 90 nM (see Figure 3). The solid line represents the best linear least squares fit of the data to eq 4. (**C**) Determination of k_{on} and k_{off}. Plot of the k_{obs} from Part B against $[SL\ RNA_0]$. The open point at [SL RNA] = 0 is from an independent measurement of k_{off} (see Table II). The solid line represents the best linear least squares fit of the data to eq 5.

Measurement of values for k_{off} and k_{on} allows calculation of the equilibrium constant for TOP **4** binding to the SL RNA according to the relation $K_{eq} = k_{off}/k_{on}$. The equilibrium constant calculated in this way is 164 nM, which agrees within a factor of 2 with the dissociation constant of 88 nM measured under equilibrium conditions.

Chemical Structure-Mapping Experiments. The effects of TOP

(42) Bevilacqua, P. C.; Kierzek, R.; Johnson, K. A.; Turner, D. H. *Science* **1992**, *258*, 1355.

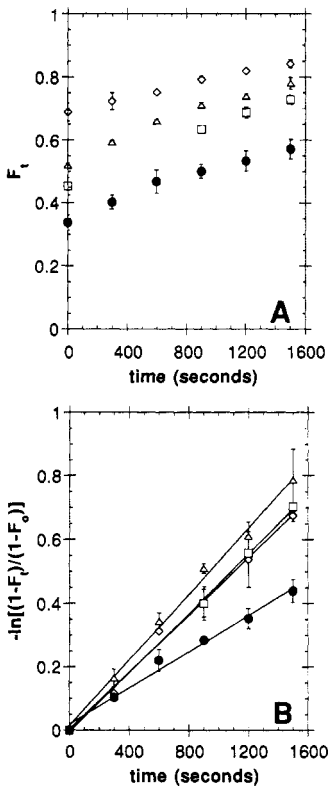

Figure 5. Determination of the rate of dissociation of the TOP **4**–SL RNA complex. (**A**) Plot of F_t, the fraction TOP **4** remaining unbound as a function of time, t, for [SL RNA_0] = 50 (◇), 100 (△), 200 (□), and 400 nM (●). (**B**) Determination of k_{off}. Plot of $-\ln[F_t/F_o]$ versus time for [SL RNA_0] = 50 (◇), 100 (△), 200 (□), and 400 nM (●). The solid line represents the best linear least squares fit of the data to eq 7.

Table II. Kinetic Parameters for Dissociation of the SL RNA–TOP **4** Complex[a]

[SL RNA_0] (nM)	$10^4 k_{off}$ (s^{-1})[b]	[SL RNA_0]	$10^4 k_{off}$ (s^{-1})[b]
50	4.5 ± 0.4[c]	200	4.6 ± 0.1
100	5.2 ± 0.5	400	2.9 ± 0.2

[a] Experimental protocols described in text. [b] Based on least squares analysis of dissociation rate data. [c] Error represents the range of two independent determinations.

binding of SL RNA structure were investigated by comparing the reactivity of SL RNA nucleotides to base-specific chemical probes in the presence and absence of TOP **1**.[43] PD-TOP **1** was chosen for these studies since it (along with PEG-TOP **4**) displayed the highest SL RNA affinity. Three chemical probes were used: CMCT reacts with RNA at N3 of U, and to a lesser extent at N1 of G;[44] DMS reacts at N1 of A, N3 of C, N7 of G,[31] and to a lesser extent, N7 of A;[45] and kethoxal reacts at N1 and N2 of G.[46] Reaction occurs only if these atoms are accessible to solution and not involved in hydrogen-bonding interactions. Modified bases are identified by use of a reverse transcriptase assay in which the chemically modified RNA is used as a template for reverse transcriptase catalyzed DNA synthesis. With the exception of N7-methyl-G and N7-methyl-A, all of the chemically modified bases listed above cause reverse transcriptase to pause or stop one nucleotide prior to the modified base. If reverse transcription is primed with an end-labeled oligonucleotide, the sites of modification can be determined from the lengths of the transcribed DNA when the products are subjected to high-resolution gel electrophoresis alongside dideoxy sequencing lanes.

(43) Ehresmann, B.; et al. *Nucleic Acids Res.* **1987**, *15*, 9109.
(44) Gilham, P. T. *J. Am. Chem. Soc.* **1962**, *84*, 687.
(45) Peattie, D. A.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4679.
(46) Shapiro, R.; Hachmann, J. *Biochemistry* **1966**, *5*, 2799. Shapiro, R.; Cohen, B. I.; Shiuey, S. J.; Maurer, H. *Biochemistry* **1969**, *8*, 238.

We first analyzed the reactivity of SL RNA nucleotides in the absence of TOP **1** to establish a baseline for subsequent comparisons and to provide experimental verification of the proposed secondary structure. The data are shown in Figure 6. All three modifying reagents react with the SL RNA to some extent in the absence of TOP **1** (lane 2 in Figure 6A,C, lane 3 in Figure 6B). Reaction with CMCT (Figure 6A) occurs moderately at U22–U24 and weakly at U8–U10, U36, U39, U50–U51, and U60. Reaction with kethoxal (Figure 6B) occurs strongly at G37–G38, G45, and G75–G76 and weakly at G42, G47, and G67. Reaction with DMS (Figure 6C) occurs strongly at A48, A54, A56–A57, and A68–A69 and weakly at C5, A6, A17, A29, C33–A34, and C49. The locations of these sites are mapped on the proposed SL RNA secondary structure in Figure 7A. Significant reaction is observed at nucleotides located within the 3′-site, between stem loops II and III. This region of the SL RNA is proposed to be single stranded, and the observed reactivity supports this assignment. The pattern of reactivity observed throughout much of the rest of the SL RNA was not expected. In particular, extensive reaction is observed at nucleotides contained *within* stem loops I and II. These regions are proposed to be base paired and should be unreactive.

Not only is significant reaction observed in putative double stranded regions, but little reaction is observed in putative single stranded regions. For example, little reaction is observed at the 5′-site although our RNase H mapping experiments indicate that TOPs **1**–**5** bind at this site. Low reactivity can result from an involvement in hydrogen-bonding interactions or solvent/reagent inaccessibility, whereas significant reactivity is indicative of a single-stranded or non-Watson–Crick base-paired structure. In summary, chemical modification experiments carried out on the free SL RNA support the proposed base-pairing pattern between C58 and the 3′-terminus but indicate that alternative base-pairing schemes should be considered between the 5′-end and A57. An alternative base-pairing scheme for the free SL RNA will be discussed later.

Next we carried out chemical modification experiments on the SL RNA under conditions where all the SL RNA is bound to TOP **1**. Lanes 3–5 in Figure 6A,C and lanes 4–6 in Figure 6B illustrates that the SL RNA nucleotides that react with modifying reagents change significantly under these conditions. The presence of TOP **1** enhances the reactivity of some SL RNA nucleotides and protects others. Significant enhancement is observed at U-1, A7–U10, C25–U26, U31–U32, U35–U36, and U51 (toward CMCT), at C5–A6 and C33–A34 (toward DMS), and at G42 (toward kethoxal). Surprisingly, three pyrimidines, C5, C33, and U60, show significant reactivity toward kethoxal in the presence of TOP **1**. Significant protection is observed at U22–U24 (toward CMCT), at A17, A29, A48, C49, A54, A56–A57, and A68–69 (toward DMS), and at G45, G47, and G75–G76 (toward kethoxal). Figure 7B illustrates the sites of TOP-induced enhancement and protection mapped on the proposed SL RNA secondary structure. Note that with the exception of the 3′-site, all of the increases and decreases in reactivity are localized to nucleotides to the 5′-side of U60 and are found primarily in stem loops I and II. Notably these changes occur where the free SL RNA shows unanticipated reactivity. One adenine in the 5′-binding site is protected in the presence of TOP **1**. Neither single-site probe **6** or **7** produces the extensive changes in reactivity observed with TOP **1**, even at concentrations significantly higher than their K_d (lanes 6 and 7 in Figure 6A,C, lanes 7 and 8 in Figure 6B).

Sequence Preferences of Random Probes. The sequence preferences of R-TOPs **10** and **11** were evaluated by use of an RNase H assay and 5′-end-labeled SL RNA (Figure 8).[17,18] Incubation of SL RNA with RNase H and a 100- or 1000-fold excess of R-TOPs **10** or **11** (analogous to PD-TOPs **1** and **2**, respectively) leads to cleavage of the SL RNA at several sites.

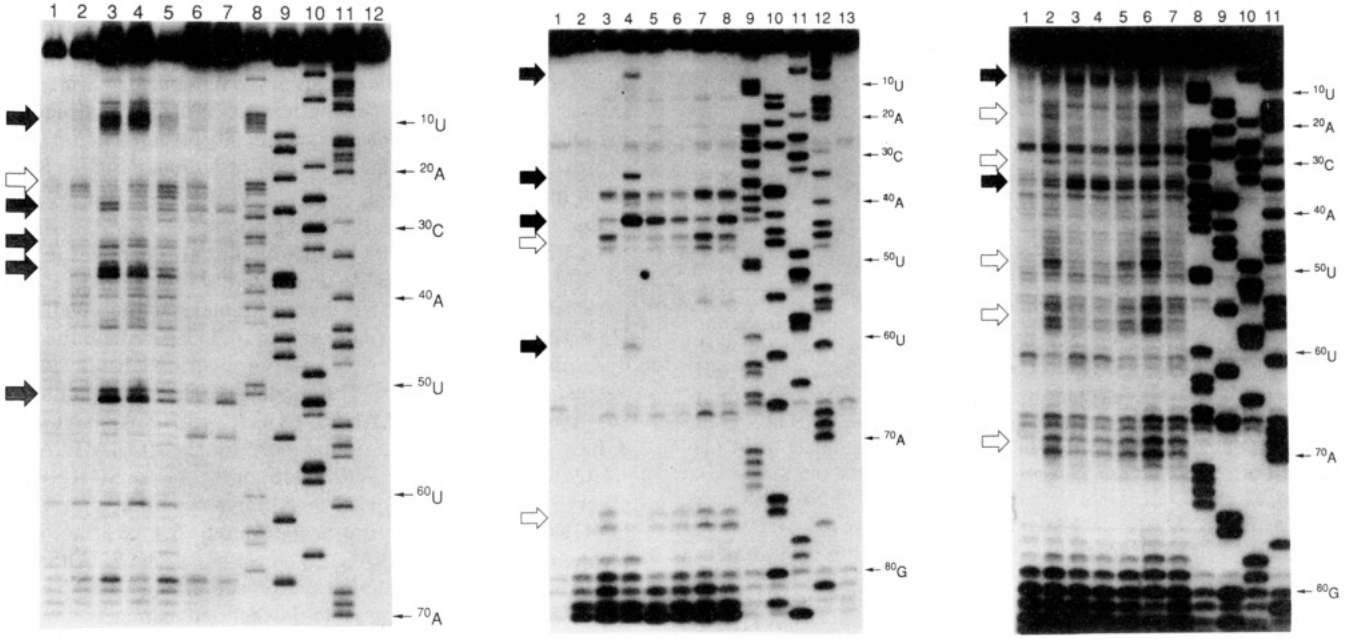

Figure 6. Autoradiogram illustrating modification of the SL RNA by CMCT, DMS, and kethoxal in the presence and absence of TOP **1** and oligonucleotides **6** and **7**. **(A, left) CMCT.** Unless otherwise indicate, all reactions contained 150 nM SL RNA, 0.1 unit/μL RNasin, 50 mM sodium borate (pH 8.0), 10 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, and 11.8 mg/mL CMCT. Lane 1, RNA alone (no CMCT); lane 2 no oligonucleotide; lane 3, 15 μM **1**; lane 4, 1.5 μM **1**; lane 5, 150 nM **1**; lane 6, 1.5 mM **6**; lane 7, 1.5 mM 7; lane 8, ddA sequencing; lane 9, ddC; lane 10, ddG; lane 11, ddT; lane 12, intact RNA. **(B, center) Kethoxal.** Unless otherwise indicated, all reactions contained 150 nM SL RNA, 0.1 unit/μL RNasin, 160 mM potassium cacodylate (pH 7.2), 10 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, 2% ethanol, and 3.7 mg/mL kethoxal. Lanes 1 and 2, intact RNA; lane 3, no oligonucleotide; lane 4, 15 μM **1**; lane 5, 1.5 μM **1**; lane 6, 150 nM **1**; 7, 1.5 mM **6**; lane 8, 1.5 mM **7**; lane 9, ddA sequencing; lane 10, ddC; lane 11, ddG; lane 12, ddT; lane 13, intact RNA. **(C, right) DMS.** Unless otherwise indicated, all reactions contained 150 nM SL RNA, 0.1 unit/μL RNasin, 200 mM HEPES–KOH (pH 8.0), 10 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, and 0.015 mg/mL DMS. Lane 1, RNA alone (no DMS); lane 2, no oligonucleotide; lane 3, 15 μM **1**; lane 4, 1.5 μM **1**; lane 5, 150 nM **1**; lane 6, 1.5 mM **6**; lane 7, 1.5 mM **7**; 8 ddA sequencing; 9, ddC; lane 10, ddG; lane 11, ddT; lane 12, intact RNA.

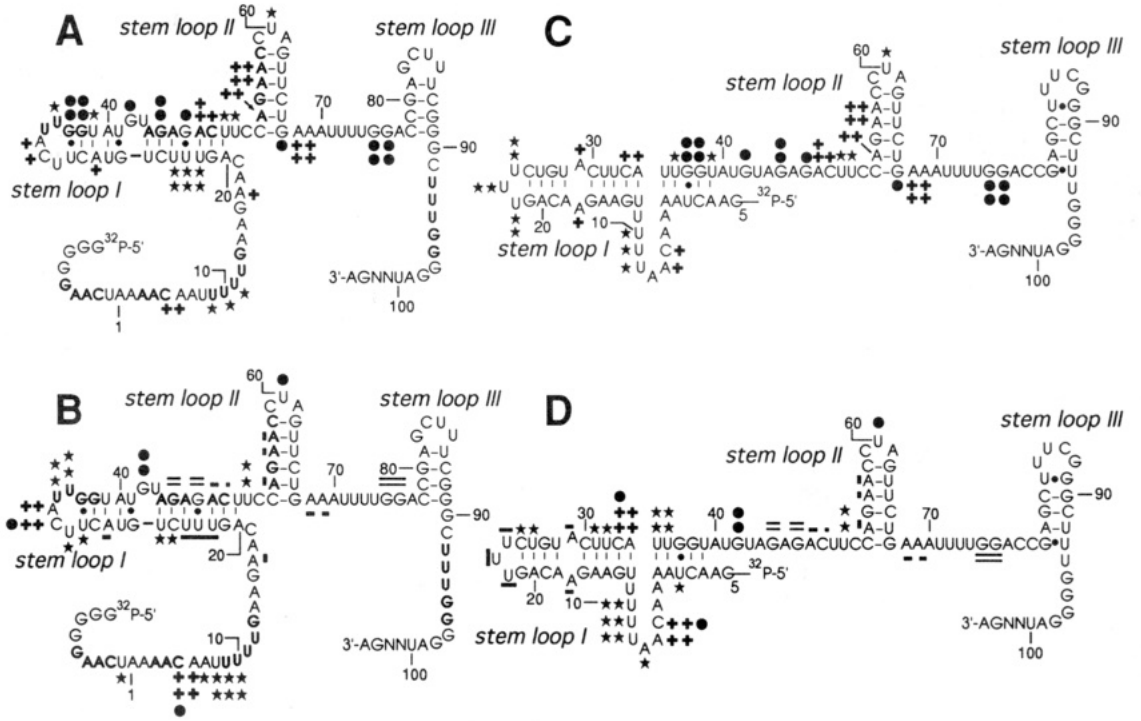

Figure 7. (**A**) Schematic illustration of the *L. collosoma* SL RNA depicting sites modified by DMS, kethoxal, and CMCT in the absence of TOP **1**. Symbols indicate the site and extent of attack by CMCT (★), DMS (+), or kethoxal (●). Highly reactive nucleotides are indicated with two symbols and weakly reactive nucleotides with one. (**B**) Schematic illustration of the *L. collosoma* SL RNA depicting the changes in reactivity observed in the presence of TOP **1**. Enhancements in reactivity are indicated by ★ (CMCT), + (DMS), or ● (kethoxal). Diminished reactivity is indicated by – (CMCT), - - (DMS), or = (kethoxal). Extensive changes are indicated with two symbols. (**C**) Schematic illustration of the LeCuyer–Crothers structure for the *L. collosoma* SL RNA depicting sites modified by DMS, kethoxal, and CMCT in the absence of TOP **1**. (**D**) Schematic illustration of the LeCuyer–Crothers structure for the *L. collosoma* SL RNA depicting the changes in reactivity of SL RNA bases observed in the presence of TOP **1**.

Cleavage is observed (lanes 9–12) at A48, U50, C59–U60, and U66. No cleavage is observed at the 5′-site (G13–C19) even though the R-TOP pool contains molecules complementary to this region and TOPs **1**–**5** generate significant cleavage at this

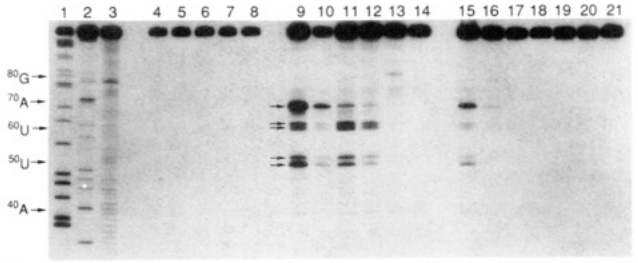

Figure 8. Autoradiogram illustrating cleavage of 5′-end-labeled SL RNA by R-TOPs **10** and **11** and RNase H. Lanes 1–3 represent G, A, and OH^- sequencing lanes, respectively. Unless otherwise indicated, each lane contains 5′-end-labeled SL RNA (13 nM, 100 000 cpm), 57 mM Tris-HCl pH 8.0, 12.7 mM $MgCl_2$, 1.6 mM DTT, 135 mM KCl, 0.9 mM EDTA, 0.05 μg/μL BSA, 4 units of RNasin (Promega), and 0.8 units of RNas H (Pharmacia) in a 15-μL volume. Lane 4, intact RNA (no RNase H); lane 5, RNase H control; lane 6, 13 μM **10** (no RNase H); lane 7, 13 μM **11** (no RNase H); lane 8, 13 μM ANNNNNA (no RNase H); lane 9, 13 μM **10**; lane 10, 1.3 μM **10**; lane 11, 13 μM **11**; lane 12, 1.3 μM **11**; lane 13, 13 μM ANNNNNA; lane 14, 1.3 μM ANNNNNA; lane 15, 13 μM **10**, 6.5 μM **12**; lane 16, 13 μM **10**, 65 μM **12**; lane 17, 13 μM **11**, 6.5 μM **12**; lane 18, 13 μM **11**, 65 μM **12**; lane 19, 13 μM ANNNNNA, 6.5 μM **12**; lane 20, 13 μM ANNNNNA, 65 μM **12**; lane 21, 130 μM **12**.

site when present at 10^3–10^4 lower relative concentration.[17,18] No cleavage is seen at the 3′-site (bases A68–A77), since an RNase H-insensitive RNA–RNA hybrid is formed here when an R-TOP binds. The RNase H cleavage patterns generated by the two R-TOPs differ slightly: the product of cleavage at U66 is generated preferentially with R-TOP **10** (lane 9) whereas it is a minor product with R-TOP **11** (lane 11). Cleavage is inhibited when the RNase H reaction is performed in the presence of a competitor oligoribonucleotide complementary to only the 3′-site and an R-TOP (lanes 15–18). Moreover, incubation of the SL RNA with the oligoribonucleotide ANNNNNA elicits no RNase H cleavage at any site listed above (lanes 13 and 14). We conclude that the cleavage bands observed in lanes 9–12 are generated from a complex in which both ends of the R-TOP are bound simultaneously to the SL RNA. The RNase H cleavage in the 5′-site (G13–C19) indicates that this site is not a preferred second binding site for a TOP anchored to the 3′-site. The preferred sites are located within and to the 5′-side of stem II.

Discussion

Each of the TOPs in Figure 1 is comprised of two short oligonucleotides, **6** and **7**, joined by a flexible tether of variable length and composition. We have shown previously that TOPs **1**–**5** bind the SL RNA in a highly specific manner and that both oligonucleotides are hybridized simultaneously to their prescribed binding sites in the complex.[17,18] Although several regions of the SL RNA are partially complementary to TOPs **1**–**5**, RNase H mapping experiments indicated that TOPs bind to only their designated 3′- and 5′-sites, which are single stranded in the originally proposed SL RNA secondary structure. We found previously that the extent of RNase H cleavage at each site varied with tether length for both the PD-TOPs and the PEG-TOPs.[17,18] However, RNase H mapping experiments do not provide an accurate measure of relative binding affinity, since the rate of enzyme-catalyzed RNA hydrolysis may be sequence and structure dependent and differ between SL RNA–TOP complexes. Consequently, equilibrium dissociation constants of the SL RNA–TOP complexes were measured to evaluate quantitatively the relative affinities of TOPs for their prescribed sequences within the SL RNA, as well as the effect of tether length and structure on the observed K_d. These experiments demonstrated that TOPs **1**–**5** do not bind the SL RNA as well as would be predicted from the affinities of single-site probes **6** and **7**. Subsequent kinetic experiments revealed that TOP **4** binds the SL RNA with a second-order rate constant significantly lower than that expected for the bimolecular association of two single stranded oligonucleotides. In order to understand the origins of the diminished affinity and the slow kinetics, we carried out extensive chemical structure-mapping experiments. These experiments revealed that TOP **4** causes a significant change in the conformation of the SL RNA when it binds.

Equilibrium Studies on the Recognition of RNA within an SL RNA Context. We cannot predict binding free energies for the SL RNA complexes because a complete set of thermodynamic parameters for RNA–DNA hybrid duplexes is not available. However, we can compare the values we measure with thermodynamic parameters for RNA–RNA hybrids[47] based on the nearest neighbor model.[48] This model predicts a $\Delta G°_{25}$ of −11.6 kcal mol^{-1} for the r(AAAUUUUGGA)–r(UCCAAAAUUU) duplex and a $\Delta G°_{25}$ of −8.7 kcal mol^{-1} for the r(GAAGAAC)–r(GUUCUUC) duplex. Within the context of the SL RNA we observe a $\Delta G°_{25}$ of −6.3 kcal mol^{-1} for recognition of r(AAA-UUUUGGA) by single-site probe **6** and −5.4 kcal mol^{-1} for recognition of r(GAAGAAC) by single-site probe **7**. Although our values are measured in a solution containing 10 mM NaCl and 10 mM $MgCl_2$ in contrast to the thermodynamic parameters derived from data obtained at 1 M NaCl, the comparison should be reasonably valid.[49] If so, the DNA–RNA duplexes are consistently 0.5 kcal mol^{-1} less stable per base pair than the corresponding RNA–RNA duplexes. This result is consistent with recent findings of Hall and McLaughlin, who observed a similar destabilization (0.3–0.4 kcal mol^{-1} per base pair) of pentamer RNA–DNA duplexes relative to the RNA–RNA duplex.[50] Whether the destabilization observed with **6** and **7** is a general characteristic of RNA–DNA hybrids[51] or one specific to these systems must await determination of the full set of thermodynamic parameters.

The equilibrium dissociation constants measured for TOPs **1**–**5** indicate that covalently linking oligonucleotides **6** and **7** into a single molecule with a synthetic tether has a significant effect on the stability of the resultant SL RNA complex. The complexes of TOPs **1** and **4** exhibit the highest stabilities, with $\Delta G°_{obs} > 3$ kcal mol^{-1} more favorable than those for the complexes of **6** or **7**. As discussed by Jencks, the "economics" of linking two ligands A and B into one ligand A–B may be evaluated in terms of the binding energies of ligands A and B ($\Delta G°_A$ and $\Delta G°_B$), and a "connection free energy" ΔG_s that describes the entropic and enthalpic changes brought about by linkage: $\Delta G°_{obs} = \Delta G°_A + \Delta G°_B - \Delta G_s$.[41] Comparison of the binding free energies for **4** ($\Delta G°_{obs} = -9.6$ kcal mol^{-1}) with those for **6** ($\Delta G°_{obs} = -6.3$ kcal mol^{-1}) and **7** ($\Delta G°_{obs} = -5.4$ kcal mol^{-1}) gives $\Delta G_s = -2.1$ kcal mol^{-1}. This indicates that although TOPs **1** and **4** bind the SL RNA significantly better than oligonucleotides **6** and **7**, they bind less well than the sum of the $\Delta G°_{obs}$ for **6** and **7** by approximately 2 kcal mol^{-1}.

Kinetic Studies on the Recognition of RNA within an SL RNA Context. Why is the connection free energy negative for binding TOPs to the SL RNA? During determination of the dissociation constants for the SL RNA–TOP complexes we observed that TOPs **1**–**5** required approximately 1 h to reach equilibrium with the SL RNA. Equilibrium should be reached within minutes if

(47) Freier, S. M.; Kierzek, R.; Jaeger, J. A.; Sugimoto, N.; Caruthers, M. H.; Neilson, T.; Turner, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 9373.

(48) Tinoco, I., Jr.; Uhlenbeck, O. C.; Levine, M. D. *Nature* (*London*) **1971**, *230*, 362. Timoco, I., Jr.; Borer, P. N.; Dengler, B.; Levine, M. D.; Uhlenbeck, O. C.; Crothers, D. M.; Gralla, J. *Nature* (*New Biology*) **1973**, *246*, 40. Borer, P. N.; Dengler, B.; Tinoco, I., Jr.; Uhlenbeck, O. C. *J. Mol. Biol.* **1974**, *86*, 843. Gralla, J.; Crothers, D. M. *J. Mol. Biol.* **1973**, *85*, 497. Gralla, J.; Crothers, D. M. *J. Mol. Biol.* **1973**, *78*, 301.

(49) Jaeger, J. A.; Zucker, M.; Turner, D. H. *Biochemistry* **1990**, *29*, 10147.

(50) Hall, K. B.; McLaughlin, L. W. *Biochemistry* **1991**, *30*, 10606.

(51) Dahm, S. C.; Uhlenbeck, O. C. *Biochemie* **1990**, *72*, 819. Martin, F. H.; Tinoco, I., Jr. *Nucleic Acids Res.* **1980**, *8*, 2295.

the TOPs bound the SL RNA with the rate constant expected for hybridization of short complementary nucleic acid single strands ($\sim 1.4 \times 10^6$ M^{-1} s^{-1}; at [SL RNA] = 5 nM, $t_{1/2} \leq 100$ s).[52] Slow association kinetics signal that binding requires a change in RNA structure.[53-55] If the binding of TOPs **1–5** was to force the SL RNA into a less stable structure, then the free energy lost would partially offset the free energy gained through Watson–Crick base pairing with the TOP and a negative connection free energy could result.

To investigate fully the kinetics of TOP binding within the context of the SL RNA, we measured the rate constants for association and dissociation of the SL RNA–TOP **4** complex. The second-order rate association constant is $(2.2 \pm 0.2) \times 10^3$ M^{-1} s^{-1}. This value is 1000 times smaller than the average second-order rate constant for hybridization of two short single-stranded oligonucleotides.[52] Slow association rates are not a general characteristic of binding oligonucleotides to globular RNAs: the second-order rate constants for binding substrate to the *Tetrahymena* ribozyme,[42,54] or binding short oligonucleotides to tRNA[53] or to a single-stranded loop in a 47-nucleotide transcript of the Ha-*ras* gene,[55] are approximately 10^6–10^7 M^{-1} s^{-1}. The k_{on} we measure is comparable to that measured by Maher *et al.* for binding a polypyrimidine sequence to duplex DNA to form a triplex[56] and is 100 times faster than that measured for hybridization of an oligonucleotide to a structured stem in the same Ha-*ras* transcript.[55]

Evidence for an Alternative SL RNA Secondary Structure. We examined the extent of SL RNA base pairing and accessibility in the presence and absence of TOP **1** using an extensive series of chemical modification experiments. We reasoned that if TOP binding occurred with no significant change in SL RNA conformation, then the changes in SL RNA reactivity should be limited to protection of the 3′- and 5′-sites. If TOP binding requires significant structural reorganization, changes in reactivity would be observed at sites other than the 3′- and 5′-sites. Our initial experiments performed in the absence of TOP **1** were inconsistent with certain features of the proposed SL RNA secondary structure and will be discussed first.

The reactivity of the SL RNA to modifying reagents in the absence of TOP **1** (Figure 7A) is consistent with structural features proposed for the 3′-half but not the 5′-half of the molecule. Significant modification occurs in the 3′-site, consistent with single-stranded character in this region. However, significant modification occurs within stem loops I and II, which are predicted to be Watson–Crick base paired and which, therefore, should be unreactive. Little modification occurs at the 5′-site, indicating that this site is either not single stranded or single stranded and relatively inaccessible in the folded RNA. Thus, the chemical modification data suggest that in the absence of TOP **1** the 5′-half of the SL RNA, nucleotides −9 to +57, is folded into a secondary structure other than that shown in Figure 7A.

Although our data are inconsistent with part of the originally proposed structure, they support a revised model for the 5′-half of the SL RNA proposed by LeCuyer and Crothers based on T-jump and mutagenesis experiments.[57] The revised secondary structure superimposed with our chemical modification data is shown in Figure 7C. Now, most reactive nucleotides map to single-stranded domains: the A3–U11 loop, the A17,A29 bulge, the U22–24 loop, the central single-stranded domain between G42 and U50, and the 3′-site. The extensive reaction observed between C33 and G38 argues against the presence of stable canonical base pairing in this stem and may have biological relevance.[58]

It is notable that three ostensibly double-stranded adenosines in stem loop II react readily with dimethyl sulfate. The enhanced DMS reactivity of adenines base paired within short, AU-rich stems is not without precedent[59] but is usually observed at positions destabilized by limited stacking interactions (adjacent loop junctions, wobble pairs). As the adenines in question do not fall in this category, we conclude that a simple stem loop does not accurately represent the structure of this region.

Evidence for a TOP-Induced Change in SL RNA Structure. Chemical modification experiments carried out on the SL RNA–TOP **1** complex indicate that significant changes in base-pairing patterns or solvent accessibility occur upon TOP binding. Nucleotides that display enhanced or diminished reactivity in the complex are identified on the original SL RNA structure in Figure 7B and on the LeCuyer–Crothers structure in Figure 7D. Enhancements and protections are observed throughout the first 80 nucleotides of the SL RNA and are not limited to protection of the 3′- and 5′-sites. The large number of SL RNA nucleotides whose reactivity changes upon TOP binding indicates that the complex displays a different set of bases to the chemical probes than the free SL RNA.

The most striking observation is that, with few exceptions, those nucleotides that show *enhanced* reactivity to single-strand-specific chemical probes upon TOP binding map to *single-stranded* regions of the secondary structure shown in Figure 7B and those nucleotides that show *diminished* reactivity toward single-strand-specific chemical probes upon TOP binding (except those at the 3′- and 5′-sites) map to *double-stranded* regions of this model. The only exceptions are C25 and U26, which show a moderate increase in CMCT reactivity in the presence of TOP **1**. In constrast, when the data are superimposed on the LeCuyer–Crothers structure (Figure 7D), many nucleotides that show enhanced reactivity in the complex map to double-stranded regions and many that show diminished reactivity map to single-stranded regions. These data are consistent with a model in which TOP **1** shifts the SL RNA from the LeCuyer–Crothers structure (Figure 7B and D) into the structure shown in Figure 7A and C. The fact that single site probes **6** and **7** fail to produce these extensive changes in reactivity (even at concentrations significantly higher than their K_d) indicates that these simple oligonucleotides are unable to fully induce the conformational switch. The observation that TOPs **1–5** exhibit similar affinities for the SL RNA (Table I) indicates that ΔG^s is dominated by the change in the SL RNA conformation.

Support of the Accessibility of Nucleotides 54–59 by R-TOPs. R-TOPs contain the oligoribonucleotide complementary to the 3′-site joined through one or five oxyphosphinicooxy-1,3-propanediol units to the deoxyoligonucleotide sequence ANNNNNA. Since R-TOPs are in theory capable of binding the SL RNA everywhere, the sites they select reveal those which are both close and accessible to a TOP anchored to the 3′-site. RNase H mapping experiments indicate that the 5′-site of the SL RNA is not a preferred site. Preferred sites include the single-stranded region preceeding stem loop II in the LeCuyer–Crothers structure and stem loop II itself. The absence of significant RNase H reaction at the 5′-site indicates that the 3′-site-specific oligonucleotide alone is insufficient to induce the conformational change observed with TOP **1**. This is consistent with chemical modification results that show minimal changes in SL RNA conformation in the presence of single-site probe **6**.

(52) Cantor, C. R.; Schimmel, P. R. *The Behavior of Biological Macromolecules*; Freeman: New York, 1971; p 1217.
(53) Freier, S. M.; Tinoco, I. C., Jr. *Biochemistry* **1975**, *14*, 3310. Uhlenbeck, O. C. *J. Mol. Biol.* **1972**, *65*, 25. Yoon, K.; Turner, D. H.; Tinoco, I. C., Jr. *J. Mol. Biol.* **1975**, *99*, 507. Fedor, M. J.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1668. Herschlag, D.; Cech, T. R. *Biochemistry* **1990**, *29*, 10172.
(54) Herschlag, D.; Cech, T. R. *Biochemistry* **1990**, *29*, 10159.
(55) Lima, W. F.; Monia, B. P.; Ecker, D. J.; Freier, S. M. *Biochemistry* **1992**, *31*, 12055.
(56) Maher, J. L., III; Dervan, P. B.; Wold, B. J. *Biochemistry* **1990**, *29*, 8820.
(57) LeCuyer, K.; Crothers, D. M. Submitted for publication.
(58) Steitz, J. A. *Science* **1992**, *257*, 888.
(59) Rairkar, A.; Rubino, H. M.; Lockard, R. E. *Biochemistry* **1988**, *27*, 582.

Conclusions

A combination of kinetic, thermodynamic, and chemical methods was used to analyze the binding of tethered oligonucleotide probes to the SL RNA of the trypanosomatid *L. collosoma*. Equilibrium experiments reveal that TOPs recognize the SL RNA with dissociation constants in the nanomolar range. Kinetic experiments reveal that formation of the TOP **4**–SL RNA complex is characterized by an association rate of 2.2×10^3 M^{-1} s^{-1} and a dissociation rate of 4.3×10^{-4} s^{-1}. The average solution lifetime of the SL RNA–TOP **4** complex is 40 min under physiologically relevant conditions of salt, pH, and temperature. Chemical modification experiments support a model for the SL RNA secondary structure proposed recently by LeCuyer and Crothers.[57] Kinetic and chemical structure-mapping data are used to demonstrate that TOP **1** (and presumably TOPs **2**–**5**) are capable of shifting the SL RNA from the LeCuyer–Crothers conformation into a conformation that is not well populated (and is presumably higher in energy) in the absence of a TOP. This suggests that TOPs could represent a general strategy for disrupting structure, and therefore function, in a globular RNA. Finally, we note that R-TOPs could be used in a general way to map accessible sites within the context of local RNA structure.

Acknowledgment. We are grateful to J. Steitz and J. Bruzik for the SL RNA construct, the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University for oligonucleotide synthesis, and J. Hartwig and D. Crothers for comments on the manuscript. S.T.C. and P.L.R. thank the Department of Education for Predoctoral Fellowships. This work was supported by the National Institutes of Health (Grant GM43501), with additional support provided by Merck & Co., Inc., Eli Lilley & Co., Rohm and Haas Co., Sterling Winthrop, Inc., Pfizer, Inc., and Hoffmann-LaRoche, Inc. A.S. is an Eli Lilley Biochemistry Fellow, a Fellow of the David and Lucile Packard Foundation, a Camille and Henry Dreyfus Teacher–Scholar, and an NSF Presidential Young Investigator.