12-2006

Generating New Specific Rna Interaction Interfaces Using C-loops

Kirill A. Afonin

Neocles B. Leontis
Bowling Green State University, leontis@bgsu.edu

J|A|C|S
ARTICLES
Published on Web 12/01/2006

Generating New Specific RNA Interaction Interfaces Using C-Loops

Kirill A. Afonin and Neocles B. Leontis*

Contribution from the Department of Chemistry and Center for Bimolecular Sciences, Bowling Green State University, Bowling Green, Ohio 43402

Received June 17, 2006; E-mail: leontis@bgsu.edu

Abstract: New RNA interaction interfaces are reported for designing RNA modules for directional supramolecular self-assembly. The new interfaces are generated from existing ones by inserting C-loops between the interaction motifs that mediate supramolecular assembly. C-Loops are new modular motifs recently identified in crystal structures that increase the helical twist of RNA helices in which they are inserted and thus reduce the distance between pairs of loop or loop-receptor motifs from 11 to 9 base-stacking layers while maintaining correct orientation for binding to cognate interaction interfaces. Binding specificities of C-loop-containing molecules for cognate molecules that also have inserted C-loops were found to range up to 20-fold. Binding affinities for most C-loop-containing molecules were generally equal or higher than those for the parent molecules lacking C-loops.

Introduction

Modular RNA motifs that mediate tertiary interactions have been used to engineer artificial self-assembling RNA molecules that form nano- and mesoscopic structures, including closed cooperative oligomeric complexes, long straight fibers,[1,2] and two-dimensional arrays.[3] Individual RNA tertiary interactions, such as the ubiquitous hairpin loop/receptor motifs,[4,5] are relatively weak and readily reversible, and thus generally exhibit fast on- and off-rates. This makes them suitable components for designing RNA molecular machines that cycle through different conformational states. However, to achieve intermolecular self-assembly at submicromolar concentrations, loop/receptor motifs must be used in pairs. Each molecule must present two loop or receptor motifs, properly oriented for interaction with the cognate motifs located in the partner molecule. Using two different binding loops, GAAA (L1) and GGAA (L2), and their cognate receptors, R1 and R2, one can engineer only two unique, non-self-associating, intermolecular interfaces suitable for directional assembly.[2] One interacting interface comprises R1 and L2 motifs in the first molecule and L1 and R2 motifs in the second molecule. Alternatively, the first molecule presents R1 and R2 and the second molecule L1 and L2. To be correctly oriented for interaction, the motifs must be separated by an integral number of helical turns. Strongest binding was observed with loop and receptor motifs separated by one helical turn (about 11 basepairs). A spacer comprising two helical turns was found to give weaker binding.[6] One way to increase the combinatorial possibilities for specific RNA self-assembly is to find additional, loop/receptor interaction motifs that interact with the same geometry and that exhibit orthogonal specificity to existing motifs. The L1/R1 pair was identified in the 3D structure of a group I intron, while the L2/R2 pair was obtained by in vitro selection methods.[4,10]

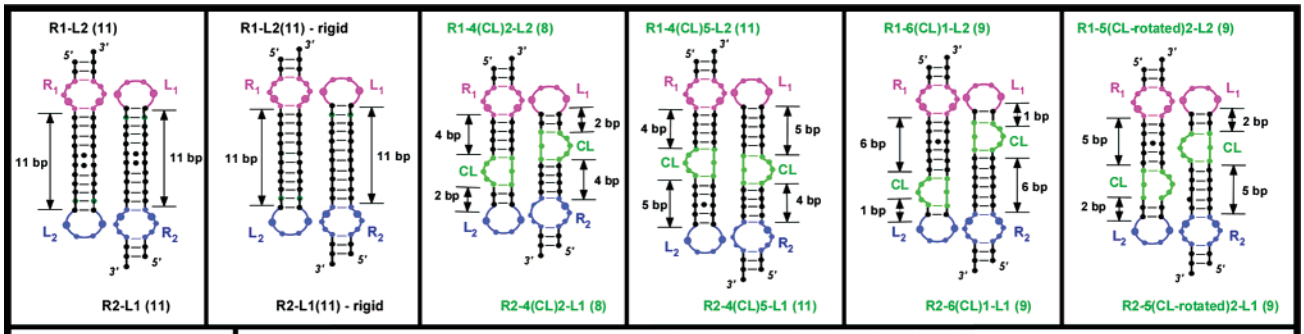
Here, we explore a second strategy for generating additional interaction interfaces, suggested by the identification in crystal structures of the recurrent RNA "C-loop" motif, which locally increases the helical twist of any RNA helix in which it is embedded.[7] C-Loops are internal loops that consist of two base-triples formed by two stacked Watson−Crick basepairs interacting with loop bases. A helix containing one embedded C-loop completes one helical turn in about nine base stacking layers, that is, seven Watson−Crick basepairs and the two base-triples belonging to the C-loop, as compared to about 11 basepairs for a canonical helix. Inserting a C-loop thus shortens the helix by about 7 Å and suggests that the same loop/receptor motif pairs can be used with C-loops to generate molecules that associate preferentially with other C-loop-containing molecules as compared to molecules having the same motifs positioned by 11 basepair helical spacers lacking C-loops.

Materials and Methods

Design of the C-Loop-Containing Tectonics. We employed computer-aided modular 3D modeling to design RNA molecules to test experimentally. As a starting 3D model, we used the NMR structure of a previously designed self-assembling RNA molecule composed of R1 and L1 motifs separated by 11 Watson−Crick basepairs (PDB

(1) Jaeger, L.; Leontis, N. *Angew. Chem., Int. Ed.* **2000**, *39*, 2521−2524.
(2) Nasalean, L.; Baudrey, S.; Leontis, N. B.; Jaeger, L. *Nucleic Acids Res.* **2006**, *36*, 1381−92.
(3) Chworos, A.; Severcan, I.; Koyfman, A. Y.; Weinkam, P.; Oroudjev, E.; Hansma, H. G.; Jaeger, L. *Science* **2004**, *306*, 2068−2072.
(4) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Szewczak, A. A.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1696−9.
(5) Costa, M.; Michel, F. *EMBO J.* **1995**, *14*, 1276−85.
(6) Jaeger, L.; Westhof, E.; Leontis, N. B. *Nucleic Acids Res.* **2001**, *29*, 455−463.
(7) Lescoute, A.; Leontis, N. B.; Massire, C.; Westhof, E. *Nucleic Acids Res.* **2005**, *33*, 2395−2409.

Chart 1. (Upper Panel) Schematic Diagrams of Representative Tecto-RNA Molecules Indicating Positions of C-Loop Motifs (CL) and Interacting Receptor (R1, R2) and Loop Motifs (L1, L2); (Lower Panel) Summary of RNA Assembly Experiments[a]



Yes - more then 60% of dimer formed

No - less then 40% of dimer formed

R2/L1			R1/L2	Without CL			With CL								
				11 bp		9 bp	11bp	8 bp	9 bp						
				#1	#17	#3	#14	#4	#6	#8	#10	#12	#18	#20	#22	#24
Without CL	11 bp		#2	Yes	Yes	No										
			#16	Yes	47	No			No	320	275	67	180	290	375	250
With CL	11bp		#15	No	No	No	No									
	8 bp		#5	No	No	No		No								
	9 bp		#7	No	No	No			No	Yes	No	No	No	Yes	No	Yes
			#9	Yes	Yes	No			Yes	65	No	Yes	Yes	Yes	Yes	Yes
			#11	Yes	Yes	No			No	No	29	Yes	Yes	Yes	Yes	Yes
			#13	Yes	Yes	No			No	Yes	Yes	22	Yes	Yes	Yes	Yes
			#19	Yes	Yes								55	Yes	Yes	Yes
			#21	Yes	Yes									25	Yes	Yes
			#23	Yes	Yes										19	Yes
			#25	Yes	Yes											21

[a] R1/L2 molecules were present at 300 nM, and R2/L1 molecules were 3′-radio-labeled and present at 0.5 nM. Numbers represent measured K_d's in nanomolar. "Yes" and "no" refer to whether association of the indicated pair was observed at these concentrations. Where measured, K_d's replace "yes". Green numbers indicate complexes for which both molecules contain C-loops.

code: 2adt).[1,8] Keeping the interacting loops and receptors intact, C-loop modules were inserted symmetrically in each molecule, in both orientations of the C-loop and at different positions along the helical spacer separating the R1 and L1 motifs. The specific C-loop used was from Helix 50 of the 23S rRNA of *H. marismortui* (PDB code: 1s72).[9] The modeled 3D structures of C-loop-containing RNAs were converted into secondary structures (Chart 1). Sequences were designed iteratively using Mfold to screen for ambiguous sequences possibly folding into undesired conformations.

RNA Preparation. RNA molecules were prepared by runoff transcription of PCR amplified DNA templates. Synthetic DNA molecules coding for the antisense sequence of the desired RNA were purchased from IDT DNA (www.idtdna.com) and amplified by PCR using primers containing the T7 RNA polymerase promoter. PCR products were purified using the QiaQuick PCR purification kit (Qiagen Sciences, MD 20874), and RNA molecules were prepared by in vitro transcription using T7 RNA polymerase (TAKARA BIO INC. http://www.takaramirusbio.com) and purified on denaturing urea gel (PAGE) (15% acrylamide, 8 M urea). The RNA was eluted from gel slices overnight at 4 °C into buffer containing 300 mM NaCl, 10 mM Tris pH 7.5, 0.5 mM EDTA, ethanol precipitated, and rinsed twice with 80% ethanol, dried, and dissolved in water.

pCp Labeling of RNA Molecules. T4 phosphokinase (T4PK) (New England BioLabs Inc.) was used to transfer the ^{32}P-gamma phosphate of ATP to the 5′-end of 3′-cytidine monophosphate (Cp), resulting in the formation of radio-labeled pCp. T4 RNA ligase (New England BioLabs Inc.) was used to label the 3′-ends of RNA molecules by

(8) Davis, J. H.; Tonelli, M.; Scott, L. G.; Jaeger, L.; Williamson, J. R.; Butcher, S. E. *J. Mol. Biol.* **2005**, *351*, 371−382.
(9) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steiz, T. A. *Science* **2000**, *289*, 878−9.

attaching [^{32}P]Cp (10−20 pmol). Labeled material was purified on denaturing polyacylamide gels (12% acrylamide, 8 M urea).

Assembly Experiments. All of the assembly experiments reported in this study were analyzed on 7% (15:1) nondenaturing polyacrylamide native gels containing 15 mM $Mg(OAc)_2$ and run at 4 °C with constant recycling of the running buffer (89 mM Tris-borate, pH 8.3/15 mM $Mg(OAc)_2$). Prior to the addition of the buffer and $Mg(OAc)_2$, the RNA samples containing a fixed amount (0.1 nM) of 3′-end labeled [P^{32}]-Cp-labeled RNA and a sufficient concentration of its cognate partner (300 nM) were heated to 90 °C for 1 min and immediately snap-cooled on ice to minimize the formation of intermolecular base pairing. Tris-borate buffer (89 mM, pH 8.3) was added, and the samples were incubated at 30 °C for 5 min, before addition of $Mg(OAc)_2$ to 15 mM and continuing incubation for 30 min. Equal volume of loading buffer (same buffer with 0.01% bromphenol blue, 0.01% xylene cyanol, 50% glycerol) was added to each sample before loading on native gel. Gels were run for 3 h, at 50 mA with constant buffer recirculation, dried under vacuum, placed on a phosphor storage screen for 16 h, and scanned using a Storm phosphoimager (Amersham, Storm 860, http://www.gehealthcare.com).

Dissociation Constants (K_d) Determination. RNA samples containing a fixed amount (0.1 nM) of 3′-end labeled [P^{32}]Cp-labeled RNA and increasing concentrations of the cognate partner molecules were assembled as described above. Monomers and dimers were quantified using the ImageQuant software. Equally sized boxes were drawn around bands corresponding to dimers and monomers. The percentage of dimer-forming complexes was calculated by dividing the corresponding quantified values for dimers by the total sum of monomer and dimer values in the corresponding lane. K_d's were determined as the concentration at which one-half of the RNA molecules were dimerized.

Lead (Pb^{2+})-Induced Cleavage. RNA samples at 300 nM concentration (including a fixed amount, 1 nM, of cognate [P^{32}]Cp-3′-end labeled RNA) were treated as indicated above. After addition of 500 mM NaOAc, lead cleavage was induced by adding 60 mM $Pb(OAc)_2$ and stopped after 60 min by adding 100 mM EDTA followed by ethanol precipitation. RNA fragments were electrophoresed for 5 h at room temperature on denaturing gels for 5 h. The gels were washed with 5% CH_3COOH, 30% C_2H_5OH for 5 min and dried as described above. Untreated RNA was run as a control and alkaline treated (pH 9, 90 °C, 3 min), and RNase T1 digested RNA samples were used as sequence markers.

Results

Design and Nomenclature of C-Loop-Containing Tecto-RNA Molecules. The nomenclature for RNA molecules used in this study is explained in the upper panel of Chart 1. We will refer to tecto-RNA molecules comprising one hairpin loop (L) and one loop-receptor (R) as R-L(*n*) molecules where *n* indicates the length of the helical spacer in stacking layers. L1 and L2 indicate the hairpin tetraloops 5′-GAAA and 5′-GGAA, and R1 and R2 are their cognate receptors as previously described.[2,10] Thus, the first molecule in the upper panel of Chart 1 is named "R1-L2(11)". This indicates that it contains receptor R1 and loop L2 separated by 11 basepairs (including two U·U wobble pairs). The molecule in the second panel is labeled R1-L2(11)-rigid to indicate that the helix contains canonical Watson−Crick basepairs in place of the U·U basepairs in R1-L2(11) and R2-L1(11). When present, C-loops are indicated by "CL", and their positions in molecules relative to the loop and receptor motifs are given as "R-*n*(CL)*m*-L", where *n* and *m* are the number of basepairs separating the C-loop from the R and L motifs, respectively. For example, the molecule designated "R1-6(CL)1-L2(9)" has the C-loop located six basepairs from R1 and one basepair from L2, providing nine "stacking layers" between the R and L motifs. The stacking layers include seven Watson−Crick basepairs and two base-triples from the C-loop. Sequences and secondary structures of all 25 RNA molecules studied are provided in Chart 2. An asterisk indicates the first and last basepairs of the helix connecting the loops and receptors. These basepairs are included when determining how many basepairs or stacking layers the helix comprises.

C-Loops inserted in R-L tecto-RNA molecules with the 5′-GCACU-3′ sequence of the C-loop located 5′- to the hairpin loop will be referred to as having the "standard" orientation, while the inverted orientation will be designated "rotated". Preliminary 3D modeling indicated that, to avoid steric hindrance during self-assembly, the C-loop in the rotated orientation should be positioned at least two basepairs from loop or receptor motifs. For the "standard" orientation, modeling indicated it could be feasible to position the C-loop just one basepair from the hairpin loop motif.

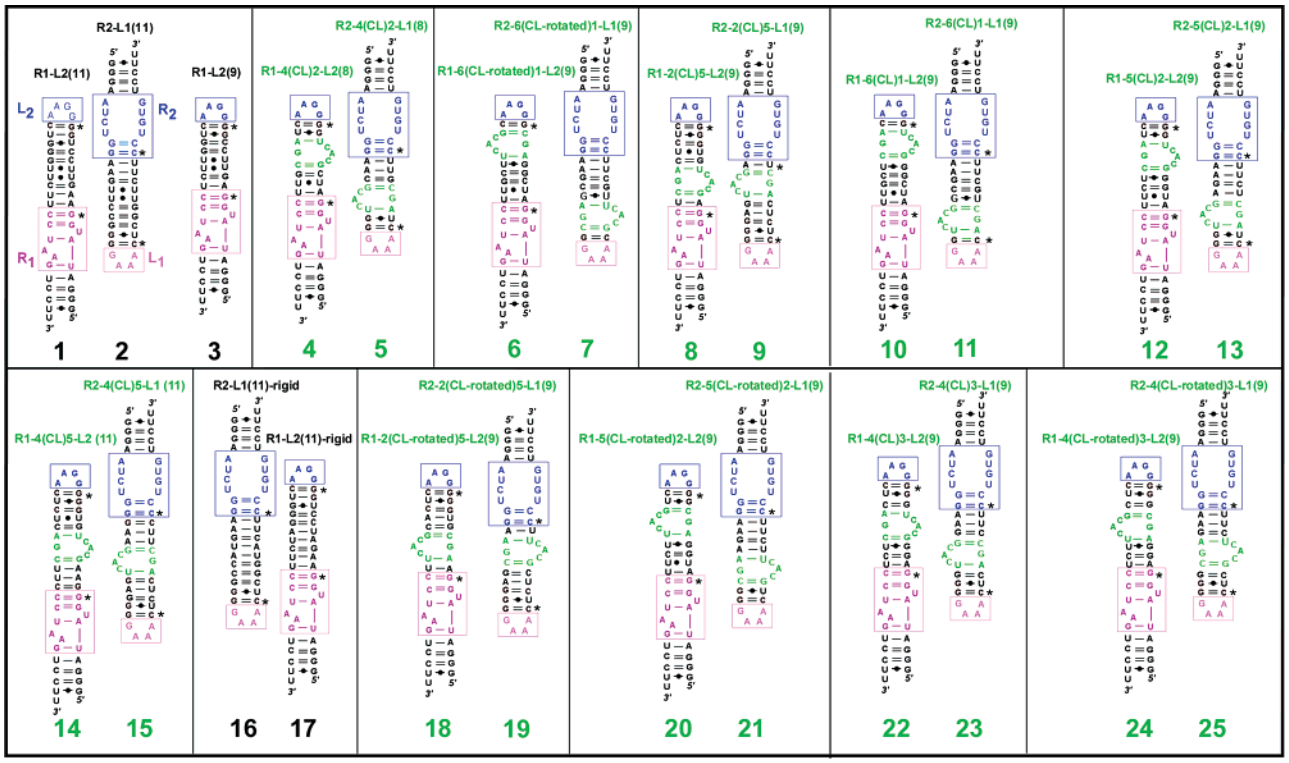
Assembly Experiments and Dissociation Constants (K_d's). Assembly experiments were carried out using gel electrophoresis to determine binding affinities of cognate molecules, with and without C-loops (Figure 1A). Radio-labeled R2-L1 C-loop-containing molecules were mixed with cognate R1-L2 partners with and without C-loops and analyzed by electrophoresis as described in the Materials and Methods. Results are summarized in Chart 1 (lower panel).

On survey experiments, the concentration of the radio-labeled molecule was ∼0.5 nM and the unlabeled molecule was 300 nM. As expected, molecules **1** (R1-L2(11)) and **2** (R2-L1(11)) assemble under these conditions and served as mobility markers. Furthermore, molecules **1** and **16**, **2** and **17**, and **16** and **17** also assemble. Molecules **16** and **17** are more rigid versions of **1** and **2**. Although molecule **3** has motifs that are cognate to those of molecules **2** and **16**, it fails to assemble with either molecule because it only has nine basepairs separating its R1 and L2 motifs. This result also agrees with previous work.[6] Furthermore, molecule **3**, which has no C-loop, does not assemble with any of the C-loop-containing molecules tested (molecules **5**, **7**, **9**, **11**, **13**, and **15**).

Molecules **14** (R1-4(CL)5-L2(11)) and **15** (R2-4(CL)5-L1-(11)) are cognate molecules containing identically positioned C-loops and a total of 11 stacking layers separating the R and L motifs (nine Watson−Crick basepairs and two base-triples belonging to the C-loop), but they do not assemble. Moreover, molecule **15** does not bind to **1**, **17**, or **3**. Likewise, C-loop-containing cognate molecules **4** (R1-4(CL)2-L2(8)) and **5** (R2-4(CL)2-L1(8)), each of which only has eight stacking layers, also do not assemble. However, cognate molecules **8** and **9** (R-2(CL)5-L(9)), which have nine stacking layers and identically positioned C-loops, do assemble, as do also the cognate pairs **10** and **11** (R-6(CL)1-L(9)), **12** and **13** (R-5(CL)2-L(9)), **18** and **19** (R-2(CL-rotated)5-L(9)), **20** and **21** (R-5(CL-rotated)2-L(9)), **22** and **23** (R-4(CL)3-L(9)), and **24** and **25** (R-4(CL-rotated)3-L(9)). All of these molecules have nine stacking layers but differ in the position and orientation of the C-loop relative to the R and L motifs. Each of these matched pairs of C-loop-containing cognate molecules produce sharp dimer bands on native electrophoresis gels under the conditions of Figure 1A, which

(10) Bates, A. D.; Callen, B. P.; Cooper, J. M.; Cosstik, R.; Geary, C.; Glidle, A.; Jaeger, L.; Pearson, J. L.; Proipin-Perez, M.; Xu, C.; Cumming, D. R. S. *Nano Lett.* **2006**, *6*, 445−448.

Chart 2. Secondary Structures of Cognate RNA Molecules Designed for This Study[a]



[a] Each molecule is named as described in the text. Matched molecules are oriented to align cognate loop and receptor motifs. C-Loops are indicated in green, interacting R1/L1 motifs are in red, and R2/L2 motifs are in blue.

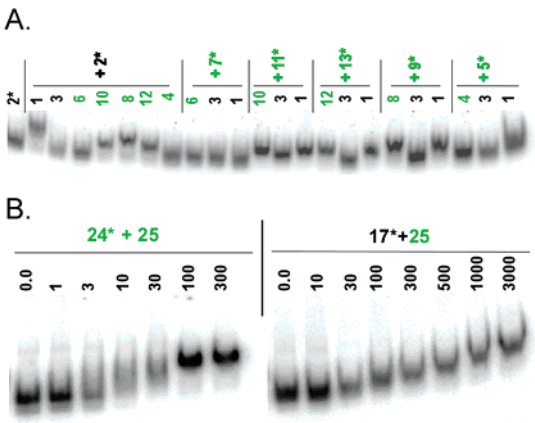


Figure 1. (A) Autoradiogram of a representative assembly experiment carried with native PAGE as described in the Materials and Methods. The radio-labeled molecule in each lane (~0.5 nM) is indicated with an asterisk. Unlabeled molecules were added to 300 nM. (B) Representative autoradiagrams of native gels used to determine K_d's for assembly of C-loop-containing molecules. Left panel: Association of C-loop-containing molecules **24** and **25**. Right panel: Association of 25 with 17, which lacks C-loop and contains 11 basepairs in the linker helix. RNA concentrations are in nanomolar. C-Loop-containing molecules are labeled with green numerals; molecules lacking C-loops are labeled with black numerals.

shows one representative set of experiments. Dissociation constants (K_d's) were measured for each matched pair of molecules (**8**/**9**, **10**/**11**, **12**/**13**, **16**/**17**, **20**/**21**, **22**/**23**, **24**/**25**). All of these K_d's are below 65 nM (Chart 1, main diagonal of lower panel). Interestingly, most C-loop-containing pairs showed higher binding affinities than the R-L(11)-rigid pair **16** and **17**.

The cognate molecules **6** and **7** (R-6(CL-rotated)1-L), each of which have the rotated C-loop located just one basepair from hairpin loop, do not assemble. Molecule **7** (R-6(CL-rotated)1-L) also fails to assemble with molecule **12** (R-5(CL)2-L), and likewise molecule **6** does not assemble with **13** (R-5(CL)2-L). Both **12** and **13** have the C-loop just two basepairs from the hairpin loop. However, **7** does assemble with **8** (R-2(CL)5-L), and correspondingly **6** assembles with **9** (R-2(CL)5-L). Both **8** and **9** have the C-loop close to the receptor motif. This result should be contrasted with the failure of **11** (R-6(CL)1-L) to bind to **8** (R-2(CL)5-L) and correspondingly of **10** (R-6(CL)1-L) to bind to **9**. Molecules **11** and **10** are identical to **7** and **6** except that the C-loops are in "standard" orientation. While **7** binds to **8** but not to **12**, **11** shows the opposite behavior, binding to **12** but not to **8**. Likewise, **6** binds to **9** but not to **13**, while **10** binds **13** but not **9**. These unexpected specificities are discussed further in the Discussion.

We observed that, under the same conditions, C-loop-containing molecules **9**, **11**, **13**, **19**, **21**, **23**, and **25**, all of which have nine baselayers, also assembled under the survey conditions with molecules **1** and **17**, both of which lack C-loops and have 11 basepairs in the linker helix. To determine the selectivity in binding of molecules with and without C-loops, we compared K_d's for C-loop-containing pairs (**8** and **9**, **10** and **11**, **12** and **13**, **18** and **19**, **20** and **21**, **22** and **23**, **24** and **25**) with K_d's measured between **16**, which has 11 basepairs and no C-loop with the cognate C-loop-containing molecules **8**, **10**, **12**, **18**, **20**, **22**, and **24**, all of which have nine stacking layers and inserted C-loops. All measured K_d's (nM) are presented in the table in Chart 1 (lower panel). The K_d measurements showed that C-loop-containing molecules bind with higher affinities to other C-loop-containing molecules with nine stacking layers than to molecule **16**, which comprises 11 stacking layers and no C-loop. The ratio of corresponding K_d's ranged from 3- to 15-fold.

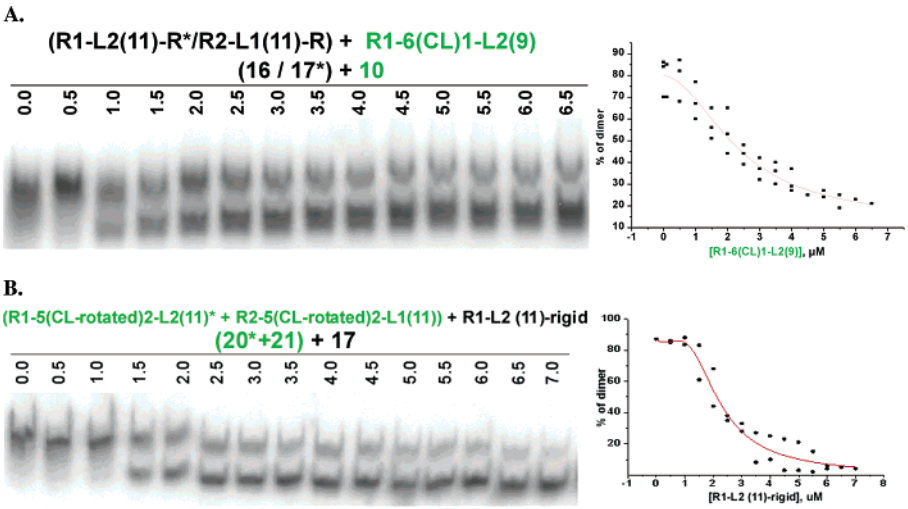


Figure 2. (A) Competition experiment in which radio-labeled **17** (0.10 μM) and the C-loop-containing molecule **10** (0.0−6.5 μM) compete for binding to molecule **16** (0.11 μM). Right panel shows fitted data from three separate experiments. (B) Competition between radio-labeled molecule **20** (0.10 μM) and **17** (0−7.0 μM) for binding to molecule **21** (0.11 μM). Right panel shows fitted data from two separate experiments.

Competition Experiments. To measure the selectivity of the binding more directly and accurately, we carried out competition experiments for representative C-loop-containing molecules. First, complexes consisting of fixed amounts of **16** (0.11 μM) and 3′-end-labeled **17** (0.1 μM) were titrated in separate experiments with increasing concentrations of unlabeled molecules **8**, **10**, or **12** over the range 0−1.5 μM. Each of these molecules competes with **17** for binding to **16**, which contains cognate loop and receptor motifs. It was found that at least 0.6 μM of molecules **8** or **12** and 1.5 μM of **10** are required to displace 50% of **17** from **16**, indicating a 6-fold selectivity for binding of **8** to **9** or **12** to **13** and 15-fold selectivity for **10** and **11** over binding to **16**. This is consistent with the K_d data. The titration of the **16**/**17** dimer with molecule **10** is shown in Figure 1. As the concentration of **10** is increased, labeled **17** is displaced to the monomer band, but some dimer is observed even at 6.5 μM of **10**. The competition experiment was repeated three times with similar results, and the fraction of **17** bound to **16** is plotted against the concentration of **10** in the right panel of Figure 2A. One-half of labeled molecule **17** was displaced at about 1.8 μM concentration of competing C-loop-containing molecule **10**, indicating about 18-fold selectivity of **16** for **17** versus **10** for **17**.

A competition experiment was also carried out for **20** and **21**. Labeled **20** (0.1 μM) bound to **21** (0.11 μM) was titrated with **17**. Approximately 2.0 μM of **17** was required to displace 50% of **20** from **21** (shown in Figure 2B).

Monitoring C-Loop Formation by Pb(II)-Induced Cleavage. Pb(II) is widely used as a conformational probe because it preferentially cleaves the phosphodiester backbone in non-canonically paired motifs or flexible regions of RNA molecules.[6,11−13] Pb(II) cleavage experiments were carried out to confirm that RNA molecules containing C-loops were folding properly and that the C-loops were maintained on assembly. Figure 3 shows Pb(II) probing data for radio-labeled molecule **24**, in the monomer form and bound to an excess of **25**. In the monomer state, molecule **24** was cleaved by Pb(II) at nucleotides 5, 6, 7, 36, 37, and 38, all of which belong to the receptor (R1); nucleotides 13, 14, 15, and 29, which belong to the C-loop; and nucleotides 21−24 in the hairpin loop (L2). In the dimer state, where molecule **24** is bound to **25**, only nucleotides 7 and 36 in R1 and 13, 14, 15 in the C-loop are subject to cleavage. The fact that the cleavages in the C-loop occur in the

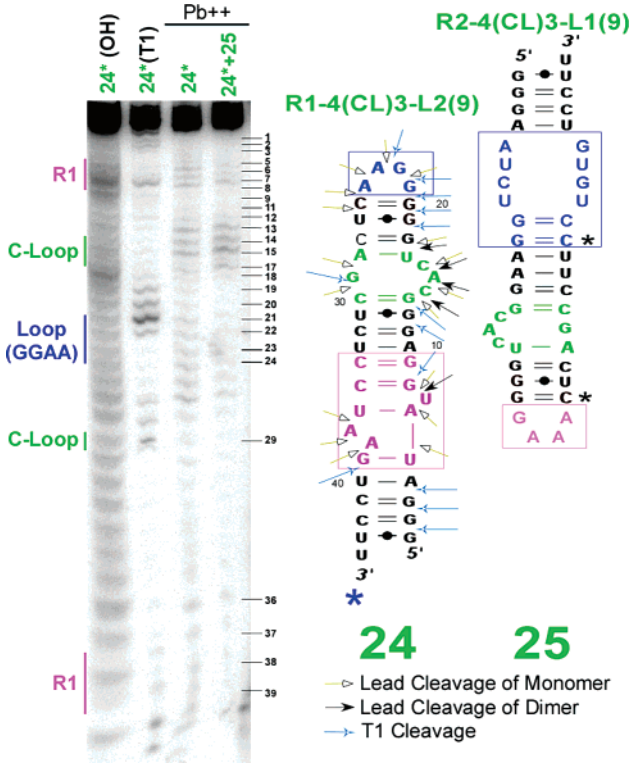


Figure 3. Lead(II) probing of radio-labeled molecule **24** (10 nM) probed in the monomer state and bound to a molar excess of molecule **25** (300 nM). Lanes 1 and 2 are hydroxide and RNase T1 treatments. Lead(II)-induced cleavage (Pb++) was performed as described in the Materials and Methods. In the right panel, the secondary structures of molecule **24** and **25** are shown to indicate cognate motifs that interact on dimerization. The cleavage sites in **24** in the monomer and dimer states are indicated with open and solid arrows. Arrows indicate different cleavages within molecule **24**.

(11) Gornicki, P.; Baudin, F.; Romby, P.; Wiewiorowski, M.; Kryszosiak, W.; Ebel, J. P.; Ehresmann, C.; Ehresmann, B. J. *Biomol. Struct. Dyn.* **1989**, *6*, 971−984.
(12) Lindell, M.; Romby, P.; Wagner, E. G. *RNA* **2002**, 534−541.
(13) Ciesiolka, J.; Michalowski, D.; Wrzesinski, J.; Krajewski, J.; Krzyzosiak, W. J. *J. Mol. Biol.* **1998**, *275*, 211−220.

Table 1. Summary of Binding Selectivities Calculated from K_d Ratios or Measured Directly with Competition Experiments[a]

CL-containing molecules	position of CL	selectivity
Low Selectivity in Assembly (Less than 10-fold)		
8, 9	R-2(CL)5-L	4.9 (7.5)
12, 13	R-5(CL)2-L	3.0 (5)
18, 19	R-2(CL-rotated)5-L	3.2 (6)
High Selectivity in Assembly (More than 10-fold)		
10, 11	R-6(CL)1-L	9.4 (18)
20, 21	R-5(CL-rotated)2-L	11.6 (20)
22, 23	R-4(CL)3-L	19.7
24, 25	R-4(CL-rotated)3-L	11.9

[a] For competition experiments, the radio-labeled even-numbered C-loop-containing molecule of each matched pair competed with R-L(11) molecule **17** for binding to the respective odd-numbered C-loop-containing molecule. Selectivities calculated from K_d's (Chart 2) are given in the third column with selectivities obtained from competition experiments in parentheses.

dimer as in monomer provides evidence the C-loop is intact in the dimer. Other sites in the loop and receptor are protected by dimer formation.

Discussion

Molecules **14** and **15**, which contain C-loops but comprise 11 stacking layers, do not assemble with each other or bind to **1** or **17** (11 basepairs, no C-loop) or **3** (9 bapairs, no C-loop). In these molecules, the motifs, while separated by approximately the same distance as in **1** or **17**, are expected to have the loop and receptor motifs incorrectly aligned for assembly, due to the presence of the C-loop. The results indicate they behave as predicted.

Molecules **4** and **5**, with only eight stacking layers, are also expected to have the loop and receptor motifs incorrectly aligned for productive interactions. They also fail to assemble. In addition, **5** fails to bind to molecule **3**, which has nine basepairs.

From the experiments with molecules **6**−**13**, a surprising result is obtained. To summarize the results, molecules containing the rotated C-loop positioned one basepair from the hairpin loop (i.e., **6** and **7**) fail to bind to each other or with molecules having the C-loop in standard orientation, and positioned near the hairpin loop, either one (**10** and **11**) or two basepairs (**12** and **13**) from it. However, they do bind to molecules in which the C-loop is located close to the receptor motif (**8** and **9**). This is made even more significant by the fact that the corresponding molecules with the C-loop in the "standard" orientation (**10** and **11**) have the opposite specificity. Molecule **10** binds **13** but not **9**, and molecule **11** binds **12** but not **8**. We attribute these results to steric clashes that C-loops can generate or avoid, depending on how they are oriented and positioned. Thus, the C-loop positioned close to the hairpin loop in the rotated orientation appears to interfere with the loop/receptor interaction by clashing with helical residues near the receptor of the partner molecule, unless the C-loop in the second molecule is positioned close to the receptor as in molecules **8** and **9**. The presence of the second C-loop appears to allow the molecules to avoid the steric clash, but only when the C-loop is in the standard orientation. When it is rotated (molecule **18**), **7** does not bind. This result will be further investigated to determine whether specific interfaces can be obtained using C-loop steric clashes and their avoidance.

The binding selectivities for molecules with the new C-loop interfaces are summarized in Table 1. The selectivity was calculated by dividing the K_d values for the binding of C-loop-containing molecules to molecule **16** by the K_d's for the binding to matched C-loop-containing molecules. The R-2(CL)5-L (**8** and **9**) and R-2(CL-rotated)5-L (**18** and **19**) designs gave the highest K_d's (65 and 55 nM, respectively) and also showed the lowest selectivity. R-5(CL)2-L (**12** and **13**) gave lower K_d (22 nM) as well as low selectivity (∼3 fold), while the same design, differing only in the C-loop orientation, R-5(CL-rotated)2-L (**20** and **21**), had comparable K_d (25 nM) but ∼12 fold selectivity relative to binding to RL(11)-rigid. By placing the C-loop five basepairs from the receptor and two basepairs from the loop, one obtains molecules about 4 times more selective. These molecules have the "rotated" C-loop conformation (**20**, **21**) instead of the "standard" orientation as in **12** and **13**. R-6(CL)1-L (**10** and **11**) showed comparably low K_d (22 nM) as molecules **20** and **21** (25 nM) and about the same 10-fold selectivity. Competition experiments showed that C-loop-containing molecules **10** and **11** bind with each other with about 18-fold selectivity as compared to **16** and **17** (Figure 2A). Comparable results were obtained in competition experiments for molecules **20** and **21**, which have the C-loop rotated and moved one basepair further from the hairpin (Figure 2B). The R-4(CL)3-L (**22** and **23**) and R-4(CL-rotated)3-L (**24** and **25**) molecules gave comparably low K_d's (19 and 25 nM, respectively) and higher selectivity with respect to binding to R-L(11)-rigid (∼20- and ∼12-fold). These results show there is considerable flexibility in the positioning and orientation of C-loops to achieve high affinity binding with good selectivity.

The binding of C-loop-containing molecules having nine stacking layers to cognate molecules with 11 basepairs lacking C-loops, while initially surprising, can be rationalized as due to unfolding of the C-loop followed perhaps by alternative basepairing that increases the distance between loop and receptor motifs to more closely match the 11-base pair spacing in molecules **16** and **17**. Alternatively, binding may induce a curvature in the C-loop-containing molecule so as to produce a convex surface that splays the loop and receptor motifs of the C-loop-containing molecule outward to match a concave induced bending of the R-L(11) molecules that orients the cognate motifs inward to optimize binding. In this case, C-loop unfolding may not be necessary. If C-loop unfolding is occurring, the use of more stable C-loops should increase the selectivity for the desired C-loop-containing cognate interacting interface and decrease binding to R-L(11) molecules, and thus increase selectivity and improve the usefulness of C-loop-containing interfaces for combinatorial assembly. The relative stabilities of different C-loops have not been determined. Sequence analysis shows that a number of sequence variants are available for future work. In addition, in vitro selection (SELEX) can be applied to select directly for more stable C-loops.

From the Pb(II)-induced cleavage experiments, we inferred that the C-loop remains folded after dimerization, at least when bound to other C-loop-containing molecules, because Pb(II) cleavage in the C-loop was identical for monomer and dimer. This is consistent with the low K_d's obtained for most combinations of C-loop-containing molecules as compared to molecules lacking C-loops. Dimerization protects the nucleotides of the L2(GGAA) hairpin and the R1 receptor from Pb(II) cleavage on a dimer formation, in agreement with previous results using R1-L1(11) molecules.[6]

Conclusion

This study demonstrates that by incorporating C-loop motifs into tecto-RNA molecules we can create new interaction interfaces with comparable and, in some cases, better binding affinities and moderate to good (up to 20-fold) selectivity using previously described receptor/loop cognate motifs. It may be possible to increase these selectivities by identifying more stable C-loops. The affinities of C-loop-containing molecules depend on the locations and orientations of C-loops within cognate partners. The molecules with the C-loops giving the highest affinities and binding specificities were determined to be R-6(CL)1-L, R-4(CL)3-L, R-5(CL-rotated)2-L, and R-4(CL-rotated)3-L. Furthermore, insertion of the C-loop motif makes possible self-assembly using pairs of loop/receptor motifs with only nine stacking layers separating the interaction motifs. The results suggest that a possible role for C-loops in biological molecules is to modulate the orientation at which interacting motifs are presented by rigid RNA elements to optimize tertiary interactions. Finally, this work suggests that it may be possible to create new specificities by judiciously positioning and orienting the C-loop close to the interacting motifs to generate and specifically relieve steric clashes.

Acknowledgment. We thank Jesse Stombaugh for assistance with figure preparation and for critical reading of this manuscript. This work was supported by grants from the American Chemical Society (ACS PRF# 42357-AC 4) and the National Institutes of Health (2 R15GM055898-03) to N.B.L.

JA064289H