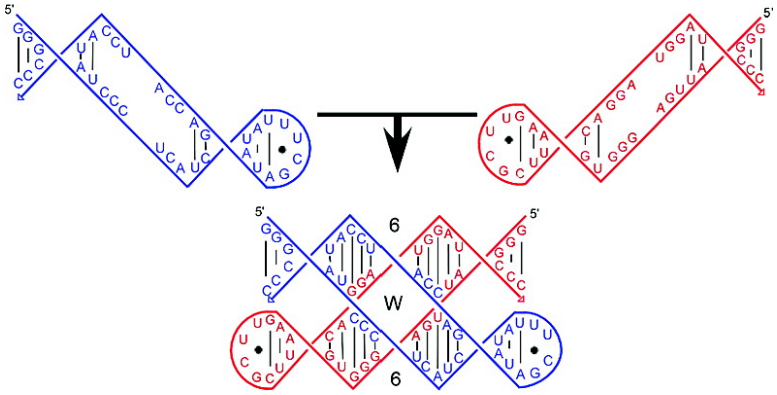
Article

Specific RNA Self-Assembly with Minimal Paranemic Motifs

Kirill A. Afonin, Dennis J. Cieply, and Neocles B. Leontis

J. Am. Chem. Soc., **2008**, 130 (1), 93-102 • DOI: 10.1021/ja071516m



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 12/12/2007

Specific RNA Self-Assembly with Minimal Paranemic Motifs

Kirill A. Afonin,[†] Dennis J. Cieply,[†] and Neocles B. Leontis*[,†,‡]

Department of Chemistry and Center for Photochemical Sciences, Bowling Green State University, Bowling Green, Ohio 43403, and Department of Chemistry and Center for Biomolecular Sciences, Bowling Green, State University, Bowling Green, Ohio 43403

Received March 24, 2007; E-mail: leontis@bgsu.edu

Abstract: The paranemic crossover (PX) is a motif for assembling two nucleic acid molecules using Watson−Crick (WC) basepairing without unfolding preformed secondary structure in the individual molecules. Once formed, the paranemic assembly motif comprises adjacent parallel double helices that crossover at every possible point over the length of the motif. The interaction is reversible as it does not require denaturation of basepairs internal to each interacting molecular unit. Paranemic assembly has been demonstrated for DNA but not for RNA and only for motifs with four or more crossover points and lengths of five or more helical half-turns. Here we report the design of RNA molecules that paranemically assemble with the minimum number of two crossovers spanning the major groove to form paranemic motifs with a length of three half turns (3HT). Dissociation constants (K_d's) were measured for a series of molecules in which the number of basepairs between the crossover points was varied from five to eight basepairs. The paranemic 3HT complex with six basepairs (3HT_6M) was found to be the most stable with $K_d = 1 \times 10^{-8}$ M. The half-time for kinetic exchange of the 3HT_6M complex was determined to be ~100 min, from which we calculated association and dissociation rate constants $k_a = 5.11 \times 10^3$ $M^{-1}s^{-1}$ and $k_d = 5.11 \times 10^{-5}$ s^{-1}. RNA paranemic assembly of 3HT and 5HT complexes is blocked by single-base substitutions that disrupt individual intermolecular Watson−Crick basepairs and is restored by compensatory substitutions that restore those basepairs. The 3HT motif appears suitable for specific, programmable, and reversible tecto-RNA self-assembly for constructing artificial RNA molecular machines.

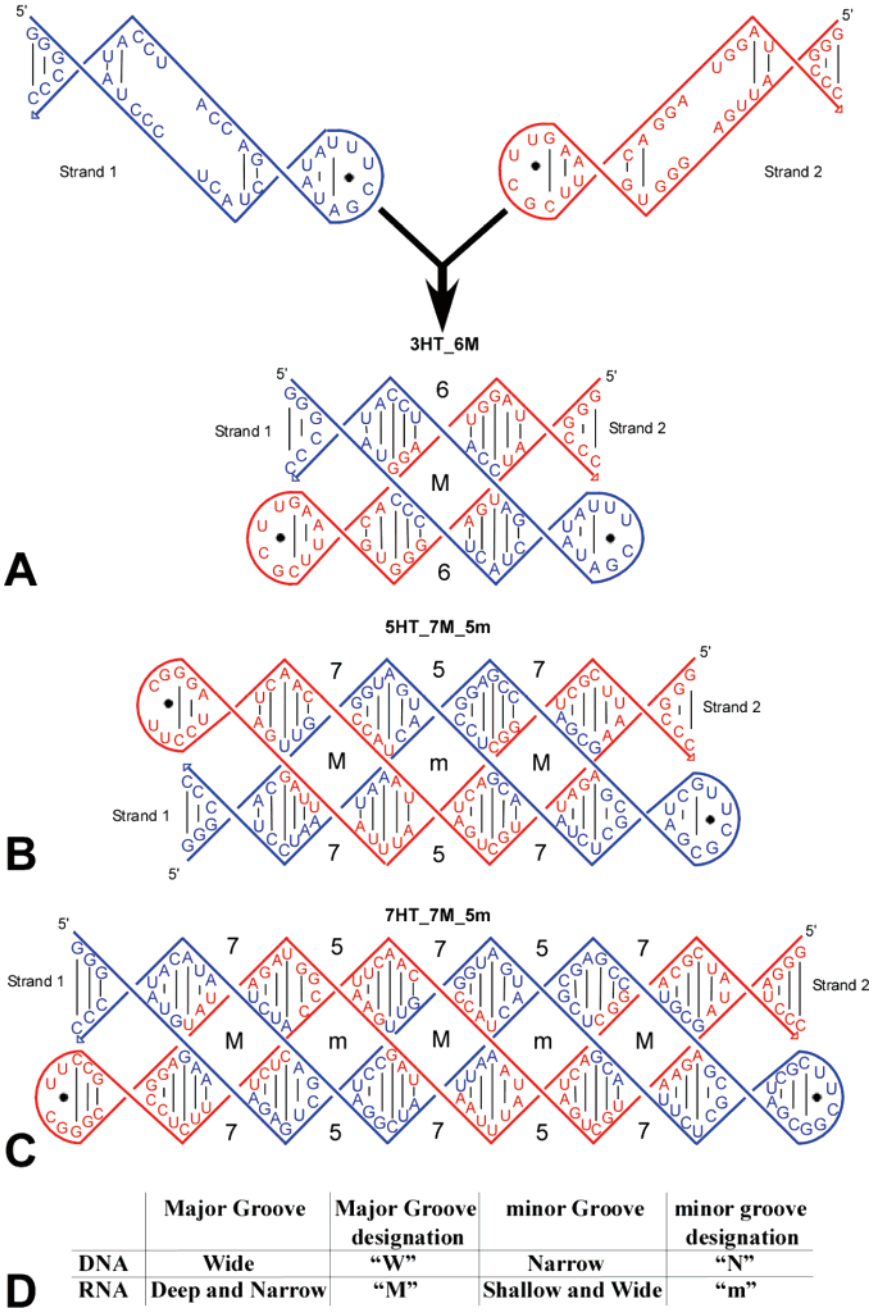
Introduction

In biomimetic nanotechnology, one aims to construct complex, nanoscale supramolecular structures from modular units using a "bottom-up" approach inspired by biological systems. The preformed modular units are designed to undergo controlled and reversible self-assembly without external manipulation of individual molecules.[1,2] This approach imitates the self-assembly of complex objects in biological systems, which produce large complex structures such as ribosomes and splicesomes from smaller modular structures by hierarchical folding and assembly.[3] Moreover, such biological structures are dynamic, comprise moving parts, and successfully bind and release dissociable factors during functional cycles.

The use of RNA as a medium for nanotechnology, has been called "RNA tectonics" and involves three steps:[4,5] (1) Conceptual, modular design at the level of 3D structure using computer modeling techniques, (2) realization of the 3D design as the requisite supporting secondary structure, and (3) detailed design of uniquely folding sequence.[2,6−8] Recent work in RNA nanotechnology has resulted in the design of artificial RNA units that assemble to form oriented filaments, closed complexes, and 2D arrays.[4,9−11] The programmed assembly of RNA monomers ("tecto-RNAs") requires specific tertiary interactions. These can be identified in atomic-resolution 3D structures or selected *in vitro* using SELEX methods.[12] "Loop−receptor" interactions, which form between specific hairpin or internal loops and cognate receptor motifs, comprise an important type of tertiary motif that occurs recurrently in large biological RNA molecules. These interactions are found in all large biological RNAs. They are sufficiently weak to be readily reversible and occur in all large biological RNA structures.[4,5,13] Loop−receptor interactions avoid plectonemic braiding of individual RNA stem−loops, which would entangle the interacting units and therefore require unfolding of secondary structure to form. Therefore they can be considered a simple form of paranemic binding motif. The diversity of artificial RNA self-assembling modules ("RNA

[†] Center for Photochemical Sciences.
[‡] Center for Biomolecular Sciences.

(1) Ball, P. *Nanotechnology* **2002**, *13*, R15−R28.
(2) Seeman, N. C.; Belcher, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (Suppl 2), 6451−6455.
(3) Goodsell, D. S. *Bionanotechnology: Lessons from Nature*; Wiley-Liss: Hoboken, NJ, 2004.
(4) Jaeger, L.; Leontis, N. B. *Angew. Chem., Int. Ed.* **2000**, *39*, 2521−2524.
(5) Jaeger, L.; Westhof, E.; Leontis, N. B. *Nucleic Acids Res.* **2001**, *29*, 455−463.
(6) Lehn, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4763−4768.
(7) Roco, M. C. *Curr. Opin. Biotechnol.* **2003**, *14*, 337−346.
(8) Whitesides, G. M.; Boncheva, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4769−4774.
(9) Chworos, A.; Severcan, I.; Koyfman, A. Y.; Weinkam, P.; Oroudjev, E.; Hansma, H. G.; Jaeger, L. *Science* **2004**, *306*, 2068−2072.
(10) Hassan, B.; Afonin, K. A.; Nasalean, L.; Jaeger, L.; Leontis, N. B. In preparation.
(11) Nasalean, L.; Baudrey, S.; Leontis, N. B.; Jaeger, L. *Nucleic Acids Res.* **2006**, *34*, 1381−1392.
(12) Jaeger, L.; Chworos, A. *Curr. Opin. Struct. Biol.* **2006**, *16*, 531−543.
(13) Jaeger, L.; Michel, F.; Westhof, E. *J. Mol. Biol.* **1994**, *236*, 1271−1276.



D

	Major Groove	Major Groove designation	minor Groove	minor groove designation
DNA	Wide	"W"	Narrow	"N"
RNA	Deep and Narrow	"M"	Shallow and Wide	"m"

Figure 1. Design of RNA molecules for paranemic assembly. The cohesion of two paranemic strands designed to interact over (A) 3 half turns (**36M1** and **36M2**), (B) 5 half turns (**5751** and **5752**), and (C) 7 half turns (**7751** and **7752**). In panels A−C, "M" indicates the major groove and "m" indicates the minor groove. **(**D**)** Correspondence between designations used for major and minor grooves spanned by crossovers in DNA and RNA paranemic complexes. The major groove is designated "W" in DNA and "M" in RNA paranemics, whereas the minor groove is designated "N" in DNA and "m" in RNA. A half turn corresponds to between 5 and 6 basepairs. Basepairs are indicated by vertical lines connecting bases on paired RNA strands. The sugar−phosphate backbone of each strand is indicated by the thick red or blue lines.

tectons") is limited by the availability and specificity of receptor−loop motifs. Alternate binding motifs that allow for more programmability, while maintaining similar geometries, are desirable. This led to our investigating the suitability of the paranemic crossover motif for RNA self-assembly.

DNA nanotechnology generally utilizes stably branched DNA molecules that are assembled with structurally well-defined cohesion methods, most usually sticky-end Watson−Crick basepairing.[14] This approach has been applied to construct regular 2D and 3D geometric molecular objects such as planar squares and triangles[15] and polyhedra such as cubes[16] and truncated octahedral.[17] Linear 2D and 3D arrays have also been developed using double- and triple-crossover junctions, both based on stable Holliday junctions.[14,18,19] Recently, nanomechanical devices have been developed such as stress gauges,[20]

(14) Seeman, N. C. *Chem. Biol.* **2003**, *10*, 1151−1159.

(15) Seeman, N. C. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 225−248.
(16) Chen, J. H.; Seeman, N. C. *Nature* **1991**, *350*, 631−633.
(17) Zhang, S.; Seeman, N. C. *J. Mol. Biol.* **1994**, *238*, 658−668.
(18) Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, *394*, 539−544.
(19) LaBean, T. J. *J. Am. Chem. Soc.* **2000**, *122*, 1848−1860.
(20) Shen, W.; Bruist, M. F.; Goodman, S. D.; Seeman, N. C. *Angew. Chem., Int. Ed.* **2004**, *43*, 4750−4752.

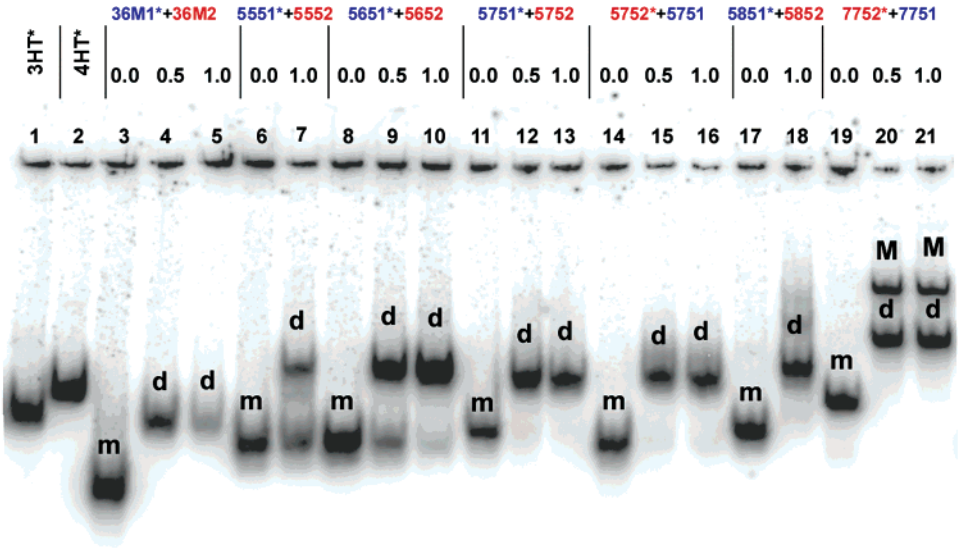


Figure 2. Native gel electrophoresis assay of formation of paranemic RNA complexes with 3, 5, and 7 half turns (3HT, 5HT, and 7HT complexes). Lanes 1 and 2 are size markers 3HT and 4HT. Concentrations of unlabeled molecules are given in μM units. The radio-labeled strand in each lane is indicated by (*) and is present at ~0.5 nM. Each series of experiments is separated by vertical bars. The first lane in each series contains only the labeled monomer strand. The monomer bands are indicated with "**m**", dimers are indicated with "**d**", and multimers are indicated by "**M**".

molecular switches,[21,22] and DNA walking devices.[23] Both single and double paranemic cohesions have been utilized to assemble large objects into 2D arrays.[24] The development of new cohesion motifs including paranemic cohesion has played an important role in the development of DNA nanotechnology. It has been suggested that paranemic DNA structures occur also in living systems, participating in genes committed to or undergoing copying processes during replication.[25]

The assembly of two RNA molecules to form the simplest possible paranemic motif, comprising two crossovers and three double-helical half turns (3HT), is shown in Figure 1A. Figure 1B shows a paranemic motif with four crossovers and five half turns (5HT), and Figure 1C shows another paranemic motif comprising six crossovers and seven half turns (7HT). As shown in Figure 1, when paranemic motifs form, the component strands crossover at every possible point and form intermolecular Watson−Crick basepairs using all the bases not involved in intramolecular basepairs. Moreover, intramolecular Watson−Crick basepairs are not disturbed.[26,27] All Watson−Crick basepairs in Figure 1 are shown with vertical lines of varying lengths to depict the twist of the double helix. Dots indicate the non-Watson−Crick basepair that forms between the first U and the G of stable UUCG hairpin loops. For the interacting molecules forming the paranemic motif to be reversibly separable without denaturation, only an even number of crossovers and an odd number of half turns is permissible.[27] Because it is not possible to assemble two separate strands to create a paranemic motif comprising an even number of half turns, the smallest paranemic cohesion motif comprises three helical half turns. Figure 1 illustrates the alternating arrangement of major (M) and minor (m) grooves between crossover points in paranemic motifs. Figure 1D shows the correspondence between designations used for DNA and RNA paranemic assembly. Previous workers chose to use the letters "W" (for wide) and "N" (for narrow) to designate the major groove and minor grooves in DNA paranemic assembly. These designations are not appropriate for RNA, because the double helix has a narrow but deep major groove and a wide but shallow minor groove. Therefore, we chose "M" to designate the major groove and "m" for the minor groove, following another convention. As 3HT molecules have only two crossover points, they can be designed so the crossovers span either the major or minor groove. For this work we chose to span the major groove as this allows for a larger number of base pairs between non-self-strands. molecular dynamic simulations of paranemic crossovers in DNA have indicated that the most structurally stable arrangement for the paranemic motif entails 6 basepairs in the major groove and 5 in the minor groove.[28] We shall refer to such complexes as "6W_5N" for DNA (consistent with previous work) and "6M_5m" for RNA.

The paranemic crossover as a cohesion motif is an attractive avenue for tecto-RNA because of the programmability of Watson−Crick basepairing and the reversibility of binding without strand entanglement or disruption of internal helices. Here we demonstrate the feasibility of RNA paranemic assembly and characterize its specificity, binding affinity, and reversibility. We show that the minimal 3HT paranemic motif is feasible with RNA.

Materials and Methods

Sequence Design. The 3HT molecules were designed with the help of two programs, RNAsoft[29] and Mfold.[30] We used the "RNA Designer" module of RNAsoft to produce over 200 possible sequences that fold

(21) Mao, C.; Sun, W.; Shen, Z.; Seeman, N. C. *Nature* **1999**, *397*, 144−146.
(22) Yan, H.; Zhang, X.; Shen, Z.; Seeman, N. C. *Nature* **2002**, *415*, 62−65.
(23) Yin, P.; Yan, H.; Daniell, X. G.; Turberfield, A. J.; Reif, J. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 4906−4911.
(24) Constantinou, P. E.; Wang, T.; Kopatsch, J.; Israel, L. B.; Zhang, X.; Ding, B.; Sherman, W. B.; Wang, X.; Zheng, J.; Sha, R.; Seeman, N. C. *Org. Biomol. Chem.* **2006**, *4*, 3414−3419.
(25) Maiti, P. K.; Pascal, T. A.; Vaidehi, N.; Heo, J.; Goddard, W. A., 3rd. *Biophys. J.* **2006**, *90*, 1463−1479.
(26) Holliday, R. *Genetics* **1964**, *50*, 323−335.
(27) Shen, Z.; Yan, H.; Wang, T.; Seeman, N. C. *J. Am. Chem. Soc.* **2004**, *126*, 1666−1674.
(28) Yagil, G. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 475−559.
(29) Andronescu, M.; Aguirre-Hernandez, R.; Condon, A.; Hoos, H. H. *Nucleic Acids Res.* **2003**, *31*, 3416−3422.

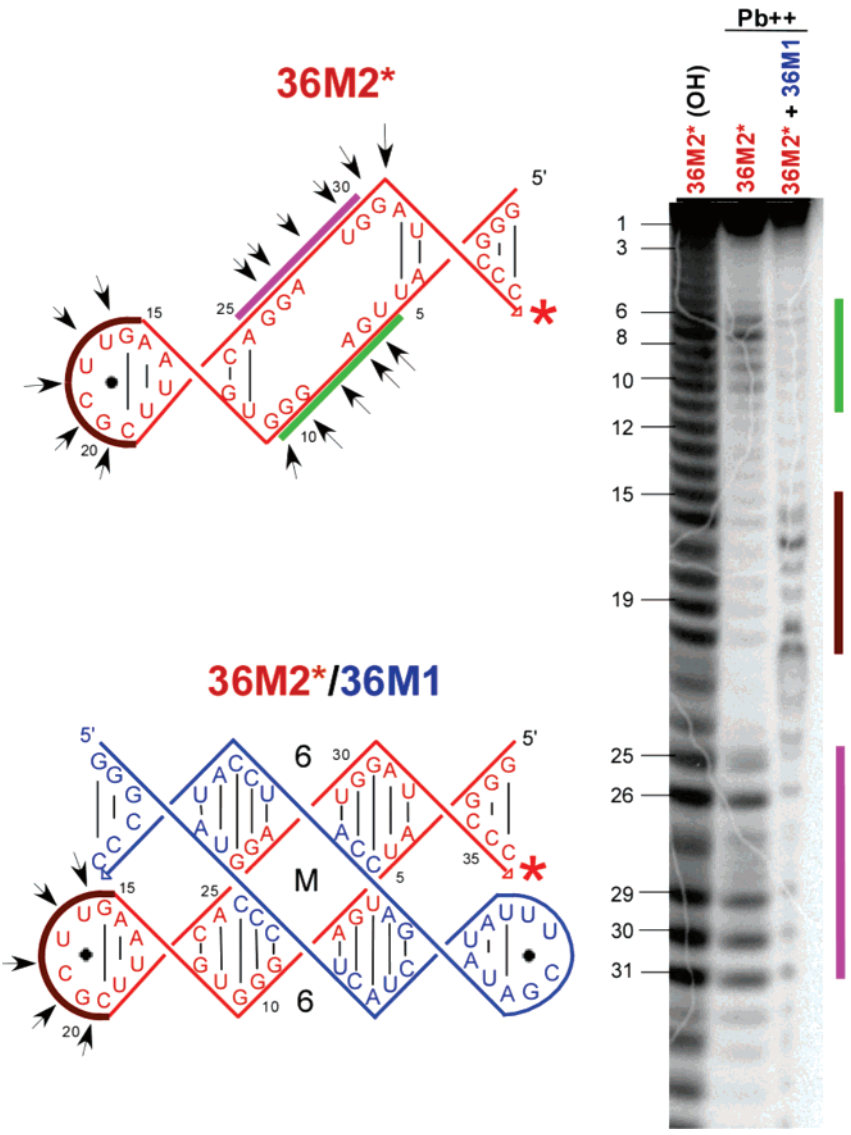


Figure 3. Pb^{2+} probing of **36M2*** in the monomer form and complexed with **36M1**. **(**Right) Denaturing gel analysis of Pb^{2+} probing of radiolabeled **36M2** (10 nM) in the monomer state and complexed to **36M1** (0.3 μM). "OH" indicates alkali-treated RNA. (Left) Arrows indicate lead(II)-induced cleavages within molecule **36M2** in the monomer state (upper left) and complexed to **36M1** (lower left). The red star indicates the labeled 3′-end of **36M2**.

into the desired secondary structure (showing Figure 1A), which comprises a stable five basepair terminal stem, a 6 × 6 unpaired internal loop composed of nucleotides intended to form the PX motif with a partner molecule, and a five basepair stem flanking a stable 3′-UUCG-5′ hairpin loop. The desired secondary structure was input into RNA Designer using bracket-dot notation with sequence constraints for the UUCG hairpin loop and for the 6 × 6 internal loop, which was taken from a paranemic crossover used in DNA. The target GC content was set to 70%. We used Mfold to check whether each strand, when folded alone, produced the desired structure and to identify sequences with relatively high folding free energies. Sequences with numerous adjacent A/U basepairs were eliminated to arrive at the sequence used. The sequence obtained is designated **36M1** ($\Delta G = -7.4$ kcal/mol). This procedure was repeated to obtain the partner molecule **36M2** ($\Delta G = -11.6$ kcal/mol). In summary, the design strategy aimed at favoring the paranemic binding of **36M1** and **36M2** by decreasing the stability of each strand when folded alone.

The 5HT and 7HT paranemic RNA molecules were based on previously reported DNA PX sequences,[31,32] replacing thymine (T) with uracil (U), and using stable UUCG RNA tetraloops to convert DNA four-stranded systems to RNA two-strand systems.[33] The RNA molecules resemble PX dumbbell (DBs) molecules except that they are closed only on one end.[27] The sequences of all molecules reported in this study are provided in the figures.

RNA Preparation. RNA molecules were prepared by runoff transcription of PCR amplified DNA templates as previously described.[34] Synthetic DNA molecules coding for the anti-sense 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, Maryland 20874). RNA molecules were prepared by *in vitro* transcription using T7 RNA polymerase (Takara Bio Inc., http://www.takara-bio.com) and purified on denaturing polyacrylamide gels (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, rinsed twice with 80% ethanol, dried, and dissolved in water.

(30) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406−3415.
(31) Seeman, N. C. *J. Theor. Biol.* **1982**, *99*, 237−247.
(32) Seeman, N. C. *J. Biomol. Struct. Dyn.* **1990**, *8*, 573−581.
(33) Antao, V. P.; Lai, S. Y.; Tinoco, I., Jr. *Nucleic Acids Res.* **1991**, *19*, 5901−5905.
(34) Afonin, K. A.; Leontis, N. B. *J. Am. Chem. Soc.* **2006**, *128*, 16131−16137.

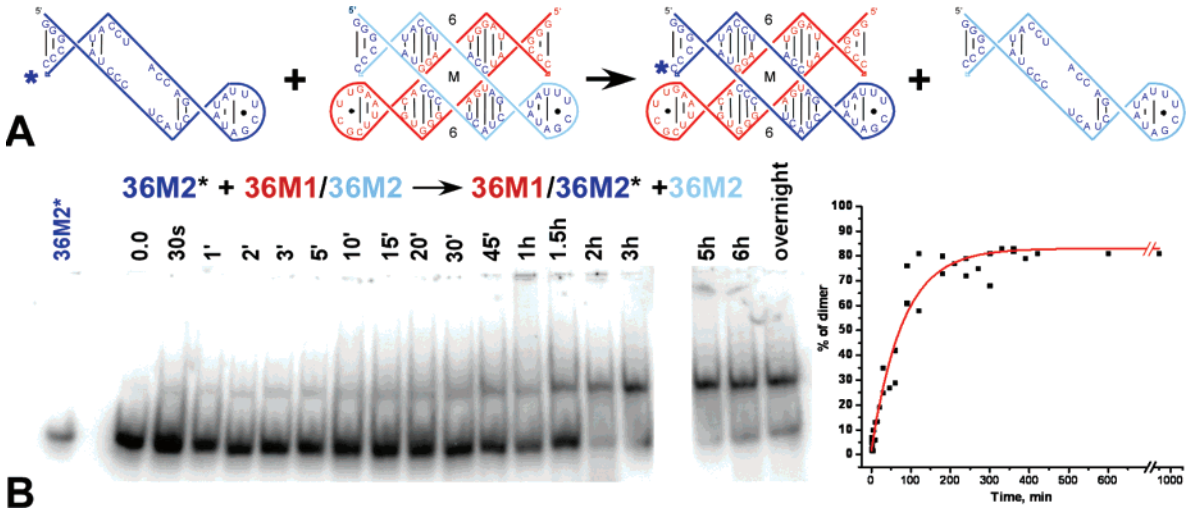


Figure 4. Kinetic exchange of **36M1/36M2** complexes. (A) Schematic representation of the kinetic exchange which results from dissociation of the preformed, unlabeled complex composed of **36M1** (in red) and **36M2** (in light blue) and formation of the labeled complex composed of **36M1** bound to labeled **36M2*** (dark blue), displacing unlabeled **36M2** (light blue). (B, Left) Radiolabeled **36M2*** (0.5 nM) was added (at time = 0) to the preformed complex of unlabeled **36M2** (0.4 μM) and **36M1** (0.3 μM). Aliquots were withdrawn at the indicated times and frozen immediately on dry ice. Samples were thawed and loaded on native gels for analysis in reverse order. (B, Right) The data from two separate gels were combined and fitted with a single exponentially decaying curve, giving the exchange half-time $t_{1/2}$= 100 ± 12 min.

Radio-Labeling of RNA Molecules. T4 phosphokinase (T4PK from New England BioLabs Inc.) was used to transfer the ^{32}P-gamma phosphate of ATP to the 5′-end of 3′-Cytidine monophosphate (Cp) to form radio-labeled pCp. T4 RNA ligase (New England BioLabs Inc.) was used to attach radio-labeled pCP to the 3′-ends of RNA molecules (10−20 pmol). Labeled material was purified on denaturing polyacylamide gels (12% acrylamide, 19:1 bis:monomer, 8 M urea).

Assembly Experiments. Prior to the addition of the buffer and $Mg(OAc)_2$, RNA samples containing a fixed amount (~0.5 nM) of 3′-end labeled RNA and the unlabeled binding partner molecule (300 nM) were heated to 90 °C for 1 min and immediately snap cooled on ice to avoid intermolecular basepairing. Tris-borate buffer (89 mM, pH 8.3) was added, and the samples were incubated at 30 °C for 5 min. Then $Mg(OAc)_2$ was added to 15 mM and incubation continued for 30 min. An equal volume of loading buffer (same buffer with 0.01% bromphenol blue, 0.01% xylene cyanol, 50% glycerol) was added to each sample for analysis on 7% (15:1) 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 $Mg(OAc)_2$). 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).

Determination of Dissociation Constants (K_d). A fixed, small amount (~0.5 nM) of 3′-end labeled RNA was mixed with increasing concentrations of the partner RNA molecule to assemble as described above and analyzed on native gels. Autoradiography was carried out using a Molecular Dynamics phosphoimaging system. Monomers and dimers were quantified using the ImageQuant software, and the percentage of dimer was calculated and plotted as a function of the RNA concentration. The data were subjected to nonlinear curve fitting using the MATLAB **nlinfit** routine to determine K_d and the **nlpredci** routine to determine the 95% confidence intervals for K_d. The dissociation constant for the reaction AB → A + B can be represented as:

$$K_d = ([A][B])/[AB] = ([A_0] - ([AB]) \times ([B_0] - [AB]))/[AB] \quad \text{(Eq. 1)}$$

where $[A_0]$ and $[B_0]$ are the initial (or total) concentrations of A and B. The way we conduct the experiments to determine K_d, the total concentration of the radiolabeled molecule, $[A_0]$, is kept fixed and much lower than the expected K_d and all values of $[B_0]$. Therefore $[B] \approx [B_0]$. Setting $r = [AB]/[A_0]$, which is the percentage of complex AB measured from the gel as a function B_0, eq 1 simplifies to eq 2:

$$K_d = (1 - r)*[B_0]/r \quad \text{(Eq. 2)}$$

This equation was solved for r and fit nonlinearly to obtain K_d.

Lead (Pb^{2+})-Induced Cleavage. RNA samples at 300 nM concentration, including a fixed amount (1 nM) of cognate 3′-end labeled RNA, were prepared as described above. After the 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 on denaturing gels for 5 h at room temperature. The gels were washed with 5% CH_3-COOH, 30% C_2H_5OH for 5 min and dried as described above. Untreated labeled RNA was run as a control, and labeled RNA samples were treated with alkali (pH 9, 90 °C, 3 min) or RNase T1 to produce sequence ladders.

Results

Designs of RNA Molecules for Paranemic Assembly. First we explain the abbreviations used to specify the designed molecules. The abbreviations reflect the characteristics of the PX complexes they are intended to form. These characteristics include the length of the paranemic motif in half turns, the number of basepairs between crossover points in the major (M) groove or minor (m) groove (when relevant), and the strand number (either S1 or S2). For example, the molecule designated 5HT_7M_5m_S1 (or **5751**), is the first strand of a pair of complementary molecules designed to form a motif of five half turns, with seven basepairs between crossovers in the major groove and five basepairs in the minor groove. It is designed to bind paranemically to **5752**. An additional lowercase letter after some sequence numbers was used to indicate sequences with point mutations. The number of half turns in each helical domain flanking the central dyad axis is achieved by successful formation of the paranemic complex.[27] The central dyad axis lies horizontally between the assembled strands, passing through each crossover point (see Figure 1). For 3HT molecules, only one number is needed to indicate the number of basepairs

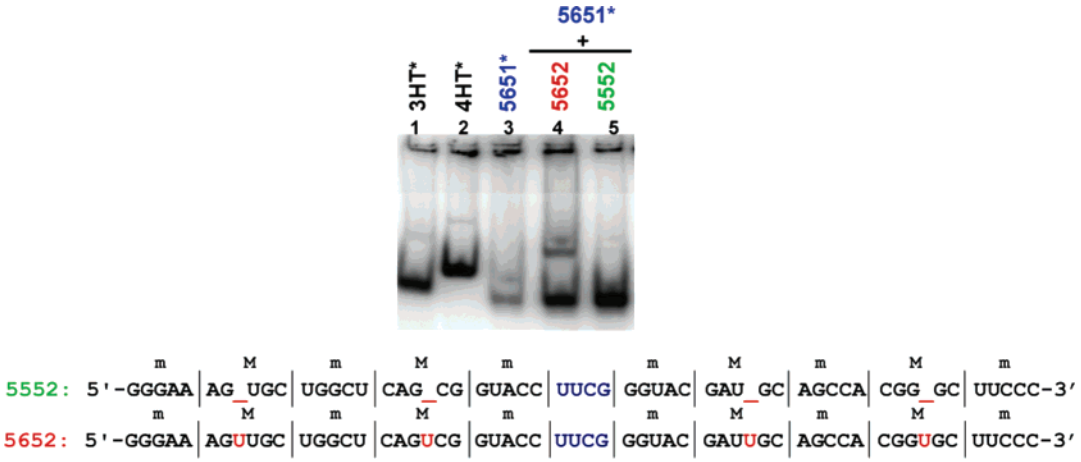


Figure 5. Molecule **5651** assembles paranemically with **5652** but not **5552**. Molecules **5651** and **5652** form six basepairs in the major groove, whereas **5552** can only form five. (Top) Lanes 1 and 2: size markers; lanes 3 to 5: **5651*** (0.5 nM) alone; with **5652** (1.0 μM); with **5552** (1.0 μM). (Bottom) The sequences of **5552** and **5652** are compared. Red letters indicate insertions; dark blue letters indicate the terminal UUCG tetraloops.

between crossover points, as there are only two. In the present work, the 3HT RNA molecules were designed to crossover the major groove and so no minor groove designation is needed. For example, molecule 3HT_6M_S1 (or **361**) indicates strand one of the molecule with 6 basepairs between crossover points spanning the major groove.

Paranemic RNA Self-Assembly. Paranemic assembly was assayed using native polyacrylamide gel electrophoresis as described in the Methods. For each pair of paranemic molecules, a small amount of the radio-labeled strand was mixed with increasing amounts of the complementary unlabeled strand. A representative gel is shown in Figure 2. For each set of paranemic molecules, the first lane on the gel contains only the radio-labeled strand at a concentration of ~0.5 nM, whereas in the adjacent lanes the radio-labeled strand (~0.5nM) is titrated with increasing amounts of the second, unlabeled strand to assay formation paranemic complexes. Formation of complexes was indicated by the appearance of a new band on the gel with the mobility expected for a dimer as determined by comparison to mobility controls, 3HT and 4HT (lanes 1 and 2 of Figure 2). Molecules 3HT (65 nucleotides) and 4HT (90 nucleotides) are H-shaped RNA molecules composed of a single 4-way junction connecting two adjacent helices. Thus, molecule 3HT is equivalent in size and shape to the 3HT paranemic complexes while 4HT is intermediate in size between paranemic 3HT and 5HT complexes.[10] Additional bands with mobilities slower than dimer complexes were only observed in significant amounts for 7HT complexes and are attributed to multimer complexes.

Figure 2 shows assays for the formation of RNA paranemic complexes 3HT_6M (lanes 3−5), 5HT_5M_5m (lanes 6−7), 5HT_6M_5m (lanes 8−10), 5HT_7M_5m (lanes 11−13), 5HT_8M_5m (lanes 17−18) and 7HT_7M_5m (lanes 19−21). Lanes 3−5 demonstrate the successful assembly of a 3HT complex by molecules **36M1*** and **36M2**. Lanes 6 and 7 show that even with addition of **5552** at 1 μM concentration the assembly of **5551*** and **5552** is not complete. Lanes 8−10 show that **5651*** and **5652** assemble at 0.5 μM concentration to form dimers with only a trace amount of monomer still present. Lanes 11−16 show that assembly of **5751** and **5752** is complete at 0.5 μM and does not depend on which strand is labeled, **5751** or **5752**. Lanes 17−18 show that the complex **5851/5852** forms readily at 1.0 μM concentration, but a faint band occurs above the dimer that can be attributed to a multimer complex. Lanes 19−21 show the assembly of the 7HT complex, **7751/7752**, which forms a significant amount of a multimer band in addition to the dimer band.

Using Pb(II) to Probe the Formation of Paranemic Complexes. Pb(II) is widely used as a conformational probe for RNA because it preferentially cleaves the phosphodiester backbone in non-canonically paired motifs or flexible regions of RNA molecules.[34] Pb(II) cleavage experiments were carried out to confirm that in the binding of **36M1** and **36M2**, which was shown to occur using native electrophoresis (see Figure 2), the predicted intermolecular basepairs do in fact form and that 3HT paranemic assembly does indeed occur when **36M1** binds to **36M2**. Radio-labeled **36M2** was treated with Pb(II) in the monomer form and in the presence of an excess of **36M1** as described in Materials and Methods, and the results are shown in Figure 3. The products of the cleavage reaction are run on denaturing gels, which separate the strands by length with nucleotide resolution. Only fragments bearing the intact 3′-end that contains the radio-label are visible on the autoradiogram shown in Figure 3, allowing one to determine where the cleavage occurs. Figure 3 lane 2 shows that molecule **36M2**, in the monomer state, was cleaved by Pb(II) at nucleotides 6−11 and 26−31, which correspond to the position of the internal loop that is intended to form the paranemic interaction, and at nucleotides 17−20, which correspond to the UUCG hairpin loop. The gel data are summarized in the schematic diagrams, which show the correspondence between bands on the gel and nucleotides in the sequence. Thus, as expected, nucleotides 6−11 and 26−31, which are not expected to base pair in the monomer, are sensitive to Pb(II) cleavage. When bound to **36M1**, however, **36M2** is preferentially cleaved by Pb(II) in the hairpin loop, nucleotides 17−20, and not at the nucleotides of the internal loop (Figure 3, lane 3). Because 36M2 is radioactively end-labeled, RNA fragments that result from more than one cut are not visualized by autoradiography. The results shown in Figure 3 thus provide evidence that assembly does in fact occur by formation of the desired paranemic motif.

Kinetic Exchange. One of the salient characteristics of paranemic cohesion is that interacting molecules bind to each other without unfolding internal helices and thus are not braided plectonemically. One therefore expects binding to be reversible.

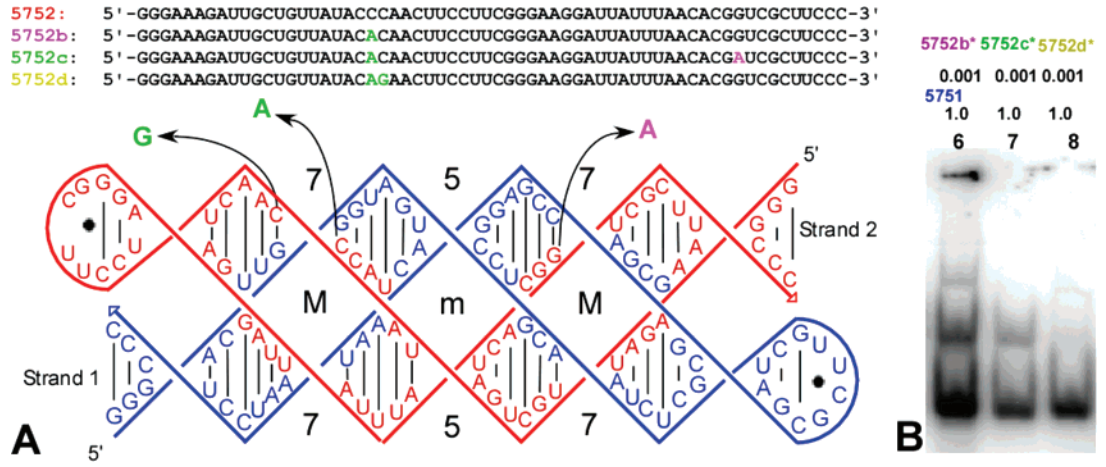


Figure 6. Specificity of paranemic assembly for the 5HT system. (A) Point mutations introduced in **5752** to test binding specificity. In **5752b** one substitution is present, in **5752c** two substitutions are present in the same groove, and in **5752d** two substitutions are present in different grooves. (B) Sequence specificity assembly experiments in which unlabeled **5751** is mixed with labeled **5752b***, **5752c***, and **5752d***.

In other words, the paranemically bound strands should be able to dissociate without the need to raise the temperature to denature the internal structure of the monomers. To test this, we carried out kinetic exchange experiments on the 36M complex at 30 °C. Figure 4A provides the scheme for the exchange reaction employed in these experiments.

First, a stoichiometric excess (0.4 μM) of unlabeled **36M2** (shown in light blue) was mixed with 0.3 μM unlabeled **36M1** (shown in red) at 30 °C in the presence of 15 mM Mg^{2+} as described in Methods to form the paranemic complex. At time = 0, a small amount (~0.5 nM) of radio-labeled **36M2*** (shown in dark blue) was added to initiate the exchange reaction between the unlabeled **36M2** strands in the complex and the free, radio-labeled **36M2*** monomers in solution. Aliquots (10 μL) were taken at successive time intervals, mixed with loading buffer (10 μL, 0 °C) containing 15 mM Mg^{2+}, and immediately placed on dry ice to stop the reaction. Samples were loaded on native analysis gels in reverse order of withdrawal to minimize the time between loading and running for aliquots with short incubation times. A representative autoradiogram of an exchange experiment is shown in Figure 4B.

As shown in the scheme in Figure 4, panel A, radio-labeled **36M2*** can only bind to **36M1** molecules that have first dissociated from unlabeled **36M2.** Upon complex formation, the radio-labeled monomers make the detection of the exchange reaction products possible and thus provide a tool for monitoring the exchange kinetics until equilibrium is again reached, at which point the radioactive **36M2** is distributed between monomer and dimer in a ratio of ~1:3. This ratio is set by the 0.1 μM excess in the total concentration of **36M2** (0.4 μM) compared to the total **36M1** concentration (0.3 μM). Thus the maximum concentration of dimer is limited by the concentration of **36M1** to no more than 0.3 μM in which case the monomer concentration is 0.1 μM. A complete analysis of the exchange kinetics is provided in the Supporting Information.

Lane 1 is the monomer **36M2*** as a control. Lanes 2−19 are aliquots taken at the indicated times after addition of **36M2***. Within 30 s, a noticeable amount of free **36M2*** has incorporated into the complex, and by about 2 h equilibrium is reached. The data were analyzed by nonlinear fitting to single-exponential curves. A pseudo-first-order rate constant of 9.8×10^{-3} min^{-1} (corresponding $t_{1/2} = 100 \pm 12$ min) was obtained from the fitting. Dissociation and association constants k_d and k_a were calculated as described in the Supporting Information and found to be $k_d = 5.11 \times 10^{-5}$ s^{-1} and $k_a = 5.11 \times 10^3$ $M^{-1}s^{-1}$. From the value of k_d, an exchange half time ($T_{1/2}$) for the **36M1/36M2** complex was determined to be $T_{1/2} = 220$ min.

The kinetic exchange experiment reveals that RNA paranemic complexes do in fact disassemble and reassemble at significant rates below the expected melting temperatures. This shows that the 3HT paranemic complex is dynamic yet also relatively stable kinetically, with a half-life of almost 4 h.

Sequence Specificity. To be useful for RNA nanotechnology, paranemic assembly motifs must display sufficient sequence specificity and programmability. We examined two aspects of specificity − matching of the number of bases between crossover points and Watson−Crick complementarity of bases programmed to pair. Therefore, we attempted to assemble pairs of strands that differed either in the number of expected base pairs between crossover points or in Watson−Crick base complementarity for intermolecular basepairs. An experiment involving 5HT molecules is shown in Figure 5. In these experiments, all labeled molecules are present at 0.5 nM. Lanes 1 and 2 are the mobility markers 3HT and 4HT. In lanes 3 through 5, labeled **5651*** is run alone, in the presence of its complement, **5652**, and with **5552**. Molecule **5551** is the complement of molecule **5552** and has two fewer nucleotide pairs than **5651** in the major (M) groove. The sequence of **5552** differs from that of **5652** by the elimination of one basepair from each wide groove crossover segment, and there are two wide grooves for 5HT paranemic assemblies. These sequence alterations are illustrated in the lower part of Figure 5. Lanes 3−5 show that **5651** only assembles in the presence of its complement **5652**.

To demonstrate the basepairing specificity of a paranemic strand for its complement, molecules **5752b**, **5752c**, and **5752d** were prepared, which differ from **5752** by one or two base substitutions. These molecules are identical to **5752** except for the base changes shown in Figure 6A. These changes occur only in the major groove between crossover points and do not affect the free energy of folding of individual strands, in agreement with folding predictions carried out for the individual strands using Mfold, which showed that the mutations did not change the pre-folding of the monomers. The results of the gel

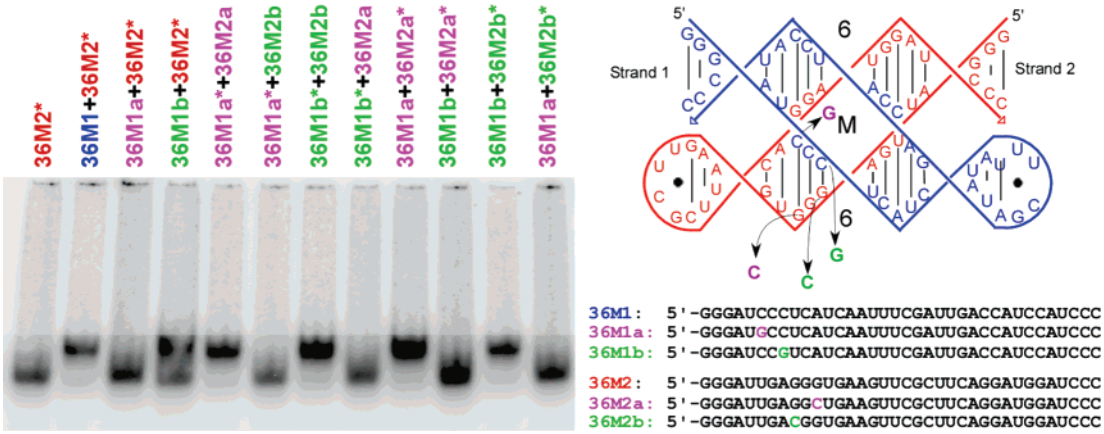


Figure 7. Specificity of RNA paranemic assembly with 3HT molecules. (Left) Native PAGE gel shows that cognate paranemic molecules **36M1**+**36M2**, **36M1a**+**36M2a**, and **36M1b**+**36M2b** all form dimer complexes, whereas noncognate pairs **36M1a**+**36M2**, **36M1a**+**36M2b**, **36M1b**+**36M2**, and **36M1b**+**36M2a** do not. Radio-labeled molecules are indicated by * and are at 0.5 nM, whereas unlabeled molecules are at 50 nM. (Right) the mutations (indicated by colors) introduced in **36M1** and **36M2** to test the specificity of binding. Molecules **36M1a** and **36M2a** contain compensating mutations, as do **36M1b** and **36M2b**.

in Figure 6B show that when molecule **5751** is mixed with **5752b*** (lane 6), **5752c*** (lane 7), or **5752d*** (lane 8), the monomer is favored. The results also show that a single WC mismatch (**5752b**) compromises dimer assembly, and two WC mismatches (**5752c** and **5752d**) largely prevent it, even at 1.0 μM concentration. Given the large number of basepairs formed when 5HT complexes form, it is not surprising that some assembly occurs when there is a single mismatch, as in the **5751**/**5752b** complex. Therefore, further studies were carried out on the 3HT system.

Figure 7 shows experiments in which single nucleotide substitutions and compensating changes were made in molecules **36M1** and **36M2**. New molecules **36M1a** and **36M2a** have compensating mutations at the crossover point whereas **36M1b** and **36M2b** have compensating mutations one basepair away from the crossover point. Assembly experiments shown in the left panel of Figure 7 demonstrated that for any combination of molecules having one basepair mismatch, paranemic assembly does not occur. This includes **36M1/36M2a**, **36M1/36M2b**, **36M1a/36M2**, **36M1a/36M2b**, and **36M1b/36M2a**; however, **36M1b/36M2** forms some dimer but not in a clear fashion. Dimer formation is restored in the **36M1a/36M2a** and **36M1b/36M2b** pairs, which contain compensating substitutions that restore intermolecular Watson−Crick basepairing.

We also studied the effect of major (M) groove size on the paranemic assembly of 3HT molecules (Figure 8). New pairs of molecules **3*i*M1** and **3*i*M2**, with $i = 4$, 5, 6, 7, and 8, were designed by modification of **36M1** and **36M2** by insertion or deletion of complementary nucleotides. The sequences of all of these molecules are shown in Figure 8B. Native PAGE gel experiments show that paranemic molecules **38M1/385M2**, **37M1/37M2**, **35M1/35M2**, and **37M1/38M2**, all form dimer complexes. **37M1/38M2** is the only complex formed by two molecules that differ in the number of bases between crossover points. Interestingly, 34M1 does not assemble with any other molecule, not even 34M2.

Determination of Dissociation Constant for 3HT Complexes. We determined the binding affinity of the 3HT paranemic cognate complexes **35M1/35M2**, **36M1/36M2**, **37M1/37M2** (Figure 9), **38M1/38M2**, and the noncognate complex **38M1/37M2**. The fraction of dimer was determined by native

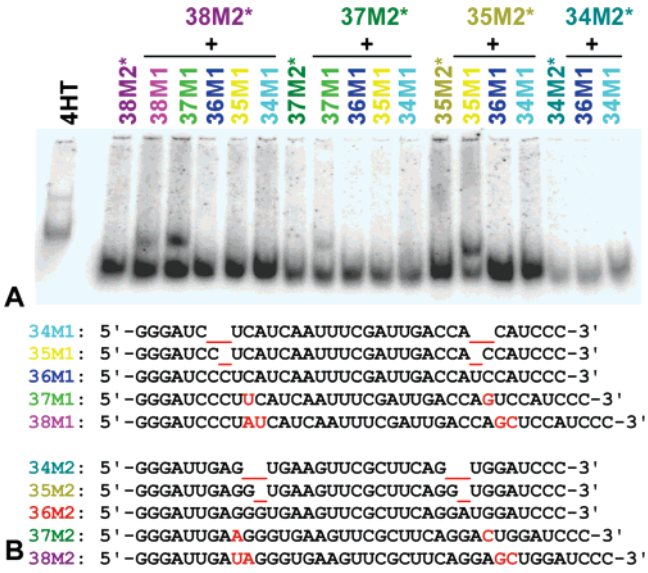


Figure 8. Effect of a groove size on paranemic assembly with 3HT molecules. (A) Native PAGE gel shows that only combinations of **38M1**+**38M2**, **37M1**+**37M2**, **35M1**+**35M2**, and **37M1**+**38M2** form paranemic dimer complexes. Radiolabeled molecules are indicated by * and are at 0.5nM, while unlabeled molecules are at 200 nM. (B) Sequences of all molecules used for this experiment. Red letters indicate insertion mutations. Red gaps indicate deletion mutations.

gel electrophoresis using a fixed amount of radiolabeled **35M2**, **36M2**, **37M2**, and **38M2** (∼0.5nM). The concentrations of added unlabeled molecules **35M1**, **36M1**, **37M1**, and **38M1** were varied from 1 nM to 3.0 μM. The relative concentrations of monomer and dimer bands were determined by integrating the autoradiograms. The percentage of dimer was plotted as a function of total concentration of the added, unlabeled molecule, and the data were subjected to nonlinear fitting as described in Methods to obtain the dissociation constants (K_d) and 95% confidence ranges for the nonlinear fitting. The K_d's, presented in Table 1, show that the most stable 3HT complexes are obtained with 6 basepairs in the major groove between the crossover points.

Discussion

The results presented show that specific, reversible, yet high-affinity RNA self-assembly can be achieved with minimal 3HT

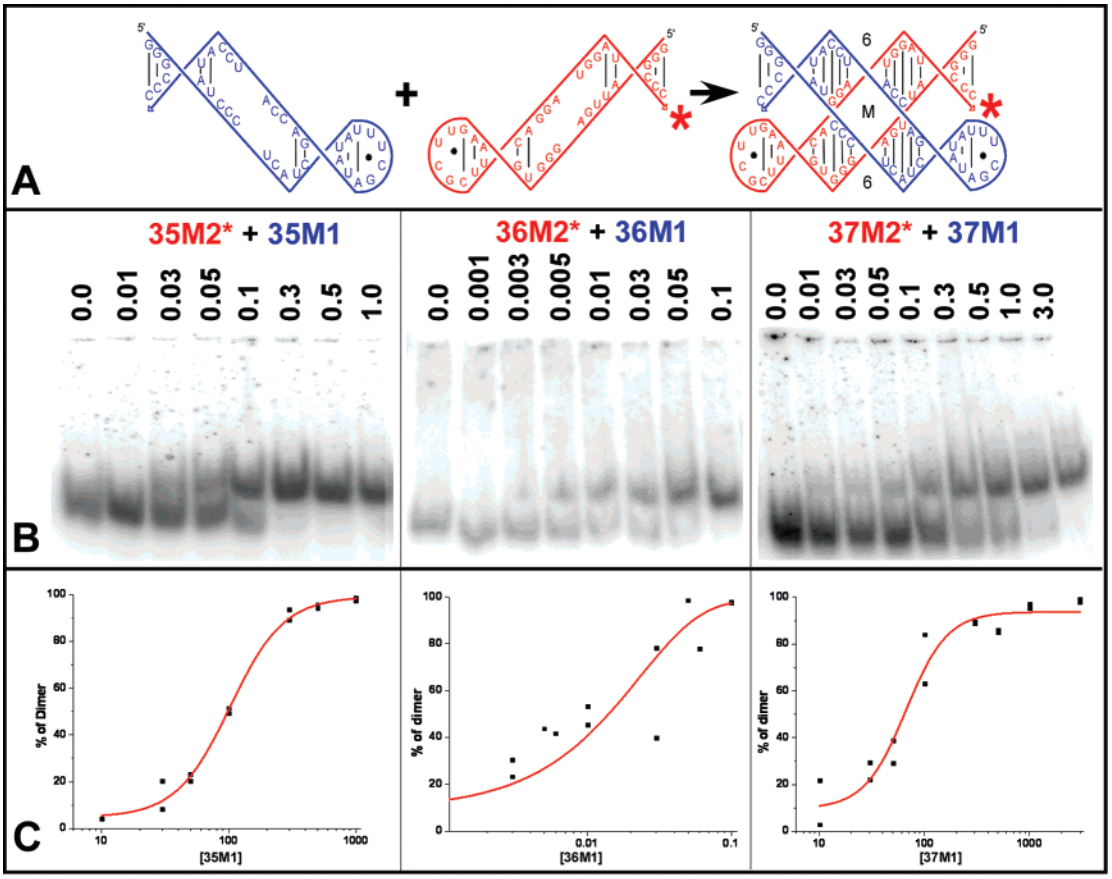


Figure 9. Determination of the dissociation constants (K_d's) for **35M1/35M2, 36M1/36M2,** and **37M1/37M2** dimers. (A) Schematic representation of the assembly process for **36M1** and radiolabeled **36M2*** paranemic molecules. (B) Radiolabeled **35M2***, **36M2***, and **37M2*** (~0.5nM) were titrated with increasing amounts on **35M1, 36M1,** and **37M1 (**indicated in μM units). (C) Quantification of gel data for K_d determinations, each represents combined fitting of 2 datasets. The Axes are marked in units of nM.

Table 1. Dissociation Constants (K_d's) Determined for Selected 3HT Complexes[a]

	K_d's in nM (95% confidence limits)		
complexes	experiment 1	experiment 2	combined data
35M1/35M2	99 (40−159) nM	94 (40−148) nM	97 (63−131) nM
36M1/36M2	16 (4−27) nM	7 (5−10) nM	10 (6−14) nM
37M1/37M2	74 (47−102) nM	49 (25−73) nM	61 (43−78) nM
38M1/38M2	169 (81−257) nM		
38M1/37M2	183 (92−274) nM		

[a] K_d's were obtained by nonlinear data fitting of native gel data as described under Methods. Fitted K_d values are reported with 95% confidence intervals.

paranemic motifs in which the two crossover points span the major groove. The K_d determinations clearly show that 3HT_6M motifs, which comprise six basepairs between the crossover points, are most stable ($K_d \approx 10$ nM) by almost an order of magnitude compared to the next most stable, the 3HT_5M motifs ($K_d \approx 100$ nM) and the 3HT_7M motifs ($K_d \approx 60$ nM) of similar sequence. Nonetheless the 3HT_5M and 3HT_7M paranemic motifs are worth studying further, as the binding affinities can be expected to vary considerably depending on the sequence. Moreover, these three motif families appear to be largely orthogonal to each other, showing little binding affinity for each other, even when they share almost identical sequences. The one exception is the binding of 3HT_8M to 3HT_7M. The K_d measurements show that 3HT_8M1 binds 3HT_8M2 and 3HT_7M2 with comparable affinity. The gel experiment in Figure 8 shows further that 3HT_8M2 binds 3HT_7M1. This suggests that all the inter-strand basepairs may not be forming in the 3HT_8M complexes. Further work will be needed to clarify this point.

Twelve intermolecular Watson−Crick basepairs, six in each helix, form during 3HT_6M paranemic assembly. The average number of nucleotide pairs per helical turn is ~11 for RNA, whereas the number of pairs in the 5HT complexes obtained in this work is 11 basepairs for the 6:5 (M:m) complexes, 12 for 7:5 complexes, and 13 for 8:5 complexes. In DNA, molecular dynamics (MD) simulations have suggested that the extra nucleotide pairs are accommodated primarily by a change in the writhing of the helix axis.[25] This writhing, which can lead to twisting of the individual helices around the central dyad access, may be essential for paranemic binding, and is therefore accommodated by the excess of basepairs over the normal 11 basepairs per turn.

This work shows that a single base mismatch disrupts assembly in 3HTM complexes at relevant concentrations (nano- to micromolar) and that compensating base substitutions that restore Watson−Crick basepairing restore binding affinity. These results are consistent with the Pb^{2+}-cleavage data that show that the internal loop that participates in intermolecular basepairing is susceptible to hydrolysis in the monomer state, but is protected in the dimer state. Likewise, the kinetic exchange data lends further support to the conclusion that the desired paranemic complexes form.

The 3HTM paranemic crossover motif thus offers a repertoire of complementary receptors that bind through programmable

Watson−Crick base pairing but are completely unlinked topologically, facilitating reversible assembly and complex and dynamic functional cycles. The incorporation of stable UNCG tetraloops to link pairs of RNA strands favors pre-folding of each monomer in the desired stem−loop structure required for paranemic assembly.

Conclusion

The interest in RNA as a medium for nanotechnology resides in its ability to form complex, dynamic, protein-like structures, stabilized by extensive tertiary interactions, with the design advantages and programmability of DNA. Important molecular machines in the cell, such as ribosomes and splicesomes, are RNA-based. Biological RNA molecules rely on long-range, tertiary interactions to initiate, guide, and stabilize the compactly folded 3D structures required for their functioning. RNA molecular machines rely on reversible tertiary and quaternary interactions to cycle between functional states. In the "bottom-up" approach to nanotechnology, large complex structures are produced by self-assembly of smaller, pre-folded units presenting interaction motifs in the appropriate orientations. This approach requires an extensive library of high affinity, high specificity interaction motifs. RNA−RNA interaction motifs, identified in 3D structures of biological RNAs, including kissing hairpins, pseudoknots, and loop−receptor interactions, have been used in the design of self-assembling RNA molecules. Additional motifs have been identified in sequence libraries or have been selected by *in vitro* combinatorial methods. However the number of tertiary RNA interacting motifs that is currently available is still limited and many are not as specific as desirable for nanotechnology applications. There is therefore a clear need for additional interaction motifs for RNA nanotechnology, ones that are programmable, specific, reversible, and that have adequate (and tunable) binding affinities. The minimal paranemic 3HT motifs explored in this work meet these criteria.

The kinetic studies show that RNA PX assembly with 3HT is reversible. As for DNA, RNA PX assembly does not require plectonemic braiding and complexes can form from preannealed hairpin molecules without disrupting internal basepairs. This result opens avenues for RNA design that exploit the specificity and programmability of the PX motif to design molecular switches that do not require denaturation to undock, as has been partially explored with the DNA PX/JX switching machine.[24]

The similarity between Watson−Crick RNA paranemic cohesion explored in this paper and the loop−receptor interactions that are commonly observed in large RNA molecules is noteworthy. Therefore, this system will enhance tecto-RNA design, taking advantage of the same principles that make loop−receptor interactions appealing but with the additional benefit of programmability to provide greater variety and flexibility of molecular design.

Acknowledgment. We thank Irina V. Novikova and Jennifer L. Benson for technical assistance, Professor Michael A. J. Rodgers of the Ohio Laboratory for Kinetic Spectrometry for helpful discussions, Professor Nadrian Seeman of New York University for encouragement to do this project, and the anonymous reviewers for useful suggestions. This work was supported by a grant from the National Institutes of Health (2R15GM055898-04).

Supporting Information Available: Detailed analysis of the calculation of association and dissociation rate constants (k_a and k_d) for the complex **36M1**/**36M2** from the kinetic and affinity data. This material is available free of charge via the Internet at http://pubs.acs.org.

JA071516M