

Chapter 13

Design and Characterization of Topological Small RNAs

Jack Hassall*, Paul MacDonald*, Teresa Cordero,
William Rostain, and Alfonso Jaramillo

Abstract

RNA can self-assemble into complex structures through base pairing, as well as encode information and bind with proteins to induce enzymatic activity. Furthermore, RNA can possess intrinsic enzymatic-like (ribozymatic) activity, a property that, if necessary, can be activated only upon the binding of a small molecule or another RNA (as is the case in aptazymes). As such, RNA could be of use in nanotechnology as a programmable polymer capable of self-assembling into complex topological structures. In this chapter we describe a method for designing advanced topological structures using self-circulating RNA, exemplified by three tiers of topologically manipulated self-assembling synthetic RNA systems. The first tier of topological manipulation, the RNA knot is a physically locked structure, formed by circularizing one monomer of knotted single-stranded RNA left with loose ends (an “open” knot). The second tier, a two interlocking ring system, is made by interlocking two circular RNA components: a circular RNA target, and an RNA lasso designed to intercalate the target before circularizing. The third tier naturally extends this system into a string of topologically locked circular RNA molecules (an RNA chain). We detail the methodology used for designing such topologically complex RNAs, including computational predictions of secondary structure, and where appropriate, RNA-RNA interactions, illustrated by examples. We then describe the experimental methods used for characterizing such structures, and provide sequences of building blocks that can be used for topological manipulation of RNA.

Key words Topological manipulation, RNA lasso, Permuted intron–exon ribozyme, RNA knot, Interlocking RNA

1 Introduction

RNA is a powerful construction tool in nanotechnology for enabling the production of flexible self-assembling structures. Various strategies have been used for their creation as reviewed by Gao [1]. One strategy is to employ secondary structure prediction software to create self assembling RNA of a defined shape for use as building blocks in more complex architectures. Another is to use natural sequences capable of ribozymatic activity. Since ribozymes usually

*Authors contributed equally to this work.

rely on tertiary structure interactions for activity, they cannot be designed *de novo* using current technology. However, they can be incorporated within larger designs, permitting the creation of custom RNA architectures. The aim of this chapter is to describe computational design techniques, developed with homage to methods established in the literature [2–6], which could be used to build topologically linked RNAs using ribozymes. We consider circular RNA as our building blocks, since they are more stable than their linear counterparts [7], and have previously been successfully utilized in self-assembling systems (e.g., as riboregulators [8]). The complex architectures designed here emerge from the topological manipulation of structurally simpler circular RNA molecules. Three tiers of topological manipulation are presented, from the knotting of a single stranded RNA to the production of a chain of multiple topologically interlocked RNA molecules. We provide insight into the sophisticated structures and scaffolds one can create through current RNA modelling techniques and delve into the methods used for the detection and characterization of these structures.

The first tier of topological manipulation is seen in the self-catalytic knotting of a single RNA strand. Knots are ubiquitous in their nature; as one can observe in an inexplicably tangled set of headphones, or by glancing down at their own laced shoes. Such knots can be considered “open,” as their ends remain loose and unconnected. The sites at which each strand crosses or overlaps itself are referred to as crossing points, and whether a strand crosses over itself or under itself at each crossing point, and how many crossing points a knot has, defines that knot. The topological knots designed here are none too dissimilar; differing in that the ends are joined together so that they cannot be undone. One can achieve this by selecting previously found “open knots” in the literature [9]—or discovering more through analysis the 3D structure of known pseudoknots—and then circularizing them (Fig. 1). In 1992, Putturaju and Been developed the Permuted Intron–Exon (PIE) method to circularize a given sequence of RNA and theoretically any Group I intron can be used to achieve this (Fig. 2) [10, 11]. A particular region of Group I introns has been shown to be peripheral (referred to as the P6 loop) and so can be removed from the Group I intron without affecting its function leaving the Group I intron in two halves [12, 13]. The two halves of the intron are then permuted, that is, so that the second half now precedes the first half whilst retaining the 5′–3′ orientation of the sequences. As the structure and function of the intron is unchanged, *cis-splicing* occurs as before, only now the 3′ and 5′ exons circularize. Sequences of our choosing can then be inserted in between the 3′ and 5′ exon, permitting their circularization so long as they do not interfere with the group I intron architecture to a significant degree, and consequently enhancing their stability, whilst also hampering degradation by exonucleases [14, 15].

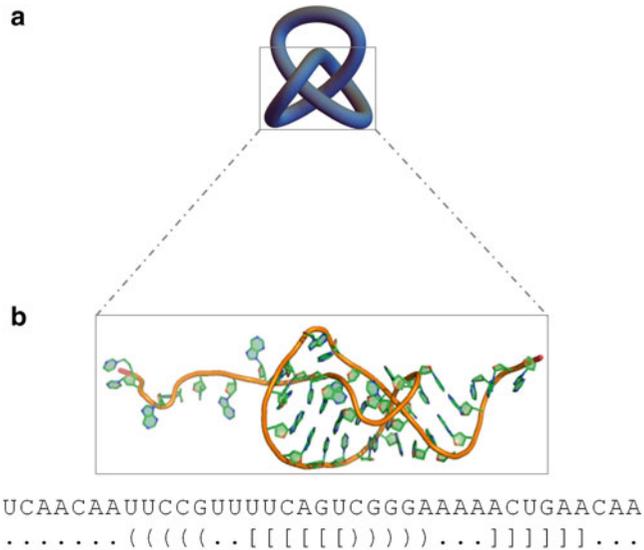


Fig. 1 Circularization of RNA “open” knots that produce topological knots. **(a)** An idealized schematic of a simple topological knot (in this instance, a trefoil knot). **(b)** A 3D visualization of an “open” knot of RNA produced in PyMol from estimates of 3D structure by RNAcomposer using the secondary structure delineated in *dot bracket* notation *below*. Circularization of such an open knot using the PIE method would result in a trefoil knot

The next tier of manipulation is the topological linking of two circular RNA, which has been previously described in literature as a two interlocking ring system. A study in 2008 [16] saw the creation of synthetic, functional RNA’s (termed RNA lassos) that could bond with a chosen target RNA molecule, and then circularize. The lassos are a novel class of antisense agents, consisting of two key features: a Hairpin Ribozyme, and an antisense target sequence. The antisense target region consists of short section of bases that are complementary to that of the target RNA molecule, encouraging the binding of the lasso to the target, and enabling the ribozyme to latch on to a chosen RNA molecule before circularizing around it. The lasso was originally designed for use as an antisense suppressor, targeting a selected nucleic acid sequence to inhibit interactions with other biomolecules, and providing a wide variety of uses within the gene expression market [12, 16]. Here however, it will be utilized as a tool to construct RNA architectures. For this design, the RNA lasso will target a self-assembled circular RNA, in this instance a circular product made using the PIE method, resulting in the formation of the two interlocking RNA rings. Prior experiments using an RNA lasso to target a circular RNA product synthetically formed using DNA ligase have proved successful, justifying the use of RNA lassos as building blocks for the higher tier systems [16].

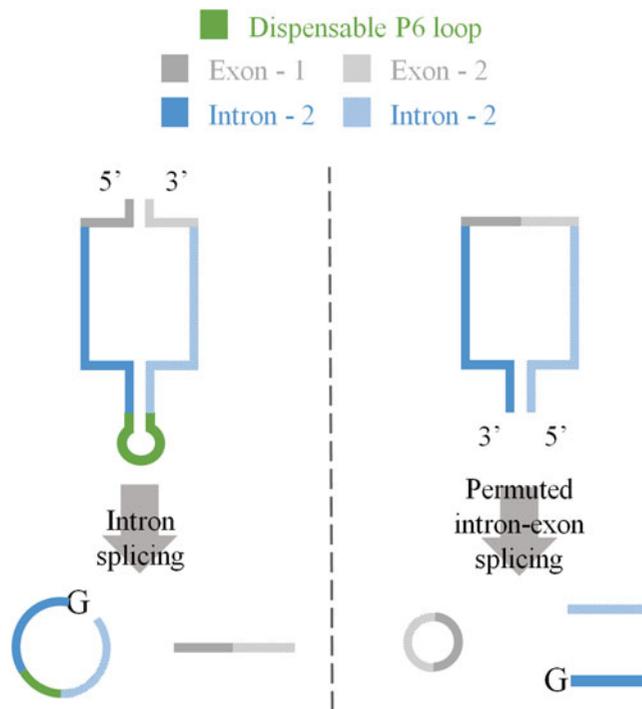


Fig. 2 Comparison of group I intron splicing and permuted intron–exon splicing. Normal intron splicing (*cis*-splicing) results in the ligation of the exonic regions, and can induce intron circularization. The permuted intron–exon (PIE) method rearranges the RNA sequences (in a fashion described as a cyclic permutation) so that the second half of the sequence precedes the first (with the dispensable P6 loop removed). Splicing of this sequence now produces a circularized exon, as desired. Note that these intron-splicing reactions occur spontaneously in the presence of a guanosine cofactor (such as GTP) and Mg^{2+} ions

The last tier of topological manipulation extends the two interlocking ring system using two RNA lassos to produce a string of topologically interlocked circles (an RNA chain) (depicted in Fig. 3a). Both lassos will have two antisense target regions, which will each be complementary to those on the remaining lasso. This will enable the lassos to interact in an alternating manner, assembling a chain of circular RNA. The RNA lasso used in the previous tier was redesigned for this system.

Previously the lasso comprised of three main components: the Hairpin ribozyme, a spacer, and the antisense target. This spacer provided the necessary flexibility for the lasso to bind to its target, circularize around it, and is engineered using a CAA nucleotide repeat [17]. The redesigned lasso, however, contains two spacers, and two different targets regions (which we refer to as TNF- α and PIE-target). TNF- α is a murine tumor necrosis factor that was

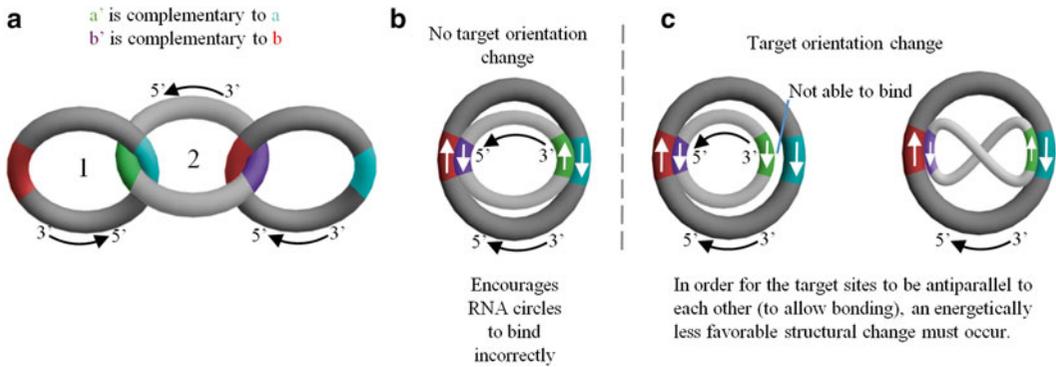


Fig. 3 Lasso interaction. (a) Demonstrates the way in which the two components of the chain system will interact, through an alternating target and lasso manor. (b) Depicts how double bonding of two lassos can occur when there is no target orientation change. Double bonding is when both targets of one lasso bind to both the targets of another. (c) Demonstrates how double bonding can be avoided through a target orientation change. Note *white arrows* depict target orientation to one another

originally targeted by an RNA lasso in the work by Dallas et al. [16] and the PIE-target is the same as target sequence used in the two interlocking ring system. A second spacer is included to avoid the possibility of one bound target sequence sterically interfering with the binding of the second target sequence [17]. If both target regions are not freely accessible the chain system will be less likely to form, and subsequently the two lassos are engineered so that only one target region will interact with one lasso. The lassos are designed with this constraint in mind, so as to prevent a situation where both targets on one lasso, bond to another lasso (Fig. 3b, c). To reduce the likelihood of this occurring, one of the target sequences runs in a different (5'–3') orientation to the rest, in other words, its sequence is reversed. This change of orientation will disable double target binding, since one of the lassos would need to change to an energetically less favorable structure in order to form hydrogen bonds between the target sequences. It should be noted that this alteration would not impact the chain structure, as the molecules are free entities, and so can freely orientate themselves.

The RNA chain could theoretically form a necklace structure, should the terminating lasso bind to the first. Designing some lassos to contain more than two target sites, which we term “junction” lassos, would enable the creation of bifurcating structures. By altering the ratio of lassos and junction lassos, different end products could be obtained, opening up the possibility of a vast array of new nanostructures.

2 Materials

2.1 RNA

Computational Design

Computational modelling should be used throughout the design process to assess the viability of the sequences. This modelling falls into two related categories: secondary structure and RNA–RNA interaction prediction. The latter is especially needed when considering the upper two tiers of topological manipulation, in pursuance of modelling the interaction of multiple RNA molecules.

2.1.1 Secondary Structure Prediction

Three webserver prediction tools were used to predict secondary structure (*see Note 1*):

1. NUPACK [18].
2. Vienna RNAfold program [19].
3. Kinefold [20].

2.1.2 RNA Interaction Prediction

The online server based RNA–RNA interaction models:

1. NUPACK [18].
2. Ribomaker [21].
3. Vienna's RNAcofold package [19].

Applying these in combination provides a means by which to validate results.

2.2 RNA Synthesis

Synthesize DNA of the designed sequence and perform *in vitro* transcription, using an appropriate kit. Tables 1 and 2 list the RNA sequences of parts used in our example constructs.

Restriction enzymes to linearize plasmid for in vitro transcription:
For interlocking lasso systems

1. PstI (FastDigest Thermo Scientific) with 10× FastDigest Buffer, knot
2. EcoRI (FastDigest Thermo Scientific) with 10× FastDigest Buffer, and for
3. PIE—Hind-III (BioLabs) using CutSmart Buffer

In vitro transcription:

TranscriptAid T7 High Yield kit (Thermo Scientific) (*see Note 2*), using the following buffer and reagents (Quantities given per 1 µg of DNA in 8 µL DEPC water): 2 µL TranscriptAid Buffer, 8 µL rNTPs mix (ATP, UTP, GTP, and CTP), 2 µL TranscriptAid Enzyme Mix.

2.3 RNA

Characterization

2D-Denaturing Urea PAGE (Polyacrylamide Gel electrophoresis) involves running two gels:

2.3.1 2D-Denaturing Urea PAGE

1. 1D gel: 6 % polyacrylamide gel: 1× TBE, 8 M urea, 30 % acrylamide (19:1). Use 200 µL 10 % APS, and 10 µL TEMED to set gel.

Table 1
RNA sequences of lasso parts.

Name	Sequence	Ref.
Lasso	<u>AGGUGUCCUCGUC</u> CGGUAGACGAGAGAAAGCCUACCAGAGAAACACACGACGUAAGUCGUGGUACA <u>UU</u> ACCUGGUACA *Spacer and target region(s) *CAAGGCUGUCCUCGUC <u>Underlined</u> regions are spliced out when the lasso circularizes.	[16]
TNF- α target	GUUCUCUUCAAGGGACAAGGCU	[16]
PIE-target	AUAAUUUGCCAUCUGGGU	[29]
Lasso Spacer	AACAACAACAAGAACACAACAAC	[16]
Lasso 1 spacers ^a	1. AACACAACAACAACAACAACAACAACAACA 2. AACACAACAACAACAACAACAACAACAACA	[16]
Lasso 2 spacers ^a	1. AACUAACAACAACAACAACAACAACAACAACA 2. AACUAACAACAACAACAACAACAACAACAACA	[16]
PIE (using group I intron found in T4)	GGUUCUACAUA <u>AAU</u> GCCUAACGACUAUCCUUUGGGAGUAGGGUCAAGUAGUCGAAACGAUAG ACAACUUGCUUUAACAAGUUGGAGUAUAAGUCUCUGCAUGGUGACAUGCAGCUGGUAUA AUCCGGGUAAGAUAAACGACCUUAUCUGAACAUAAU <u>GC</u> UACCGUUAAU <u>UU</u> GGAUCC. *Aptamer* <u>CUGCAGAU</u> GUUUUCUUGGGUAAUAGGGCCUGAGUAUAAGGUGACUUAACUUGU AAUCUAUCUAAACGGGAAACCUCUCUAGUAAGACCAAUCCCGUGCUAAA <u>UU</u> GUAGGACUAAAGCUC UCGAG <u>Underlined</u> region denotes spliced out exon.	[29]
Streptavidin aptamer	CGACCAGAAUCAUGCAAGUGCGUAAGAUAGUCGGGGCCGG	[29]

Bold regions indicate regions where different parts can be inserted into the sequence

^aModified version of original lasso spacer

Table 2
RNA sequences of knot parts

Name	Sequence	Ref.
PIE (using group I intron found in <i>Graphium penicillioides</i>)	GGGGCCGUCCACAGACUAAGUGGUUGUGGGUAGGACGUCU GUUCUAUCUAAGAUUAUAGUCGGGCCUUGCGAGAGAUCGC AGGGGUUUCUGCGUCCGUA* Aptamer *AAGUCGUAACAAG <u>GUUUUGCUAAGCAACAGUCCUUCUGUAGCAGAGCUUAA</u> GGAGACCUUGCGCCAAAAGGCGACUAUAAAUAUUCGGCAG CAAGUCAGUAUCACUGGCGACACAAUCGAAUUGCGGGGAC GCUUUAAGCUCGACGGUACCAACCGCUGGAAAUCCAGC GGGVCUAGCAUGCUGAGAGUCAGAACUCGUCGAGAUGUA CAAUAAGCAAUCCGCAGCGGUCCCC <u>Underlined</u> region denotes spliced out exon.	[29]
Knot aptamer ^a	UCAACAAUCCGUUUUCAGUCGGGAAAAACUGA ACAA (<i>see</i> Fig. 1)	[23, 24]

Bold regions indicate regions where different parts can be inserted into the sequence

^aFifth base change to a C (from a G) to reduce the likelihood of it binding with the intron

- 2D gel: 10 % acrylamide gel: 1× TBE, 8 M urea, 30 % acrylamide (19:1). Use 200 µL 10 % APS, and 10 µL TEMED to set gel.
- Urea gel loading buffer (Fisher Scientific).
- Ethidium bromide (EtBr).
- SCIE PLAS TV50 Single Gel Vertical Electrophoresis Unit.
- BioDocAnalyze Live (BioDoc Analyze Biometra) Gel Documentation Unit.

When characterizing the interlocking RNA systems the following buffers will be used:

- Magnesium Buffer—10 mM MgCl₂, 50 mM Tris-HCl (pH 7.5), and 20 % formamide. Make a 10× buffer containing MgCl₂ and Tris-HCl, then add formamide with the sample so that the concentration of formamide is 20 %.
- EDTA Buffer—10 mM EDTA, 50 mM Tris-HCl (pH 7.5), and 20 % formamide. Make in a similar fashion to magnesium buffer.

2.3.2 RNase R Treatment

RNase R (Epicentre) distributed with associated 10× RNase R Reaction Buffer. RNase R is a 3' → 5' exoribonuclease that digests linear RNAs but not circular or lariat RNA structures [25, 26] (*see* Note 3).

2.3.3 Reverse-Transcriptase-PCR (RT-PCR)

- RT: RNA sample, 2 µL (40 µM) Primer A (IDT) and 4 µL (10 mM) dNTPs (Thermo Scientific), 4 µL M-MuLV Reverse Transcriptase (New England Biolabs), 0.5 µL associated 10× M-MuLV buffer and 0.5 µL Ribolock (Thermo Scientific).

2. PCR: 2 μL RT product, 2 μL of Primers A and B (IDT), 0.4 μL (10 mM) dNTPs (ThermoScientific), 0.6 μL DMSO (BioLabs), and 4 μL Phusion Polymerase (BioLabs).
3. Thermocycling done using TProfessional TRIO Thermocycler.
4. Gel Extraction: DNA Clean and Concentrator (Zymo Research) and ADB (Zymo Research).
5. Software ClusterW2 was used to perform sequence alignments.

3 Methods

3.1 RNA

Computational Design

3.1.1 Tier 1:

Topological Knot

1. To find a suitable topological knot one can download over 1000 Group I sequences from the Group I Intron Sequence and Structure Database (GISSD) (categorized by subgroup here at <http://www.rna.whu.edu.cn/gissd/sequence.html>) [22]. Exon sequences are listed in lowercase characters and intron sequences in uppercase characters.
2. Amalgamate the each group I intron subgroup fasta file into one contiguous fasta file. Develop a script to convert this file into a format with the following fields:
 - (a) Intron name
 - (b) 5' exon sequence
 - (c) Intron sequence
 - (d) 3' exon sequence
3. To ensure that the “open” knot both forms and circularizes correctly, we aim to minimize interaction between the open knot sequence and the group I intron used to circularize it (via the PIE method). To select the most suitable intron for this task, use the following “indicators” (listed in order of importance):
 - (a) Flanking exons should not bind to the open knot sequence, as this could inhibit knot formation as well as interfere with the *cis*-splicing mechanism.
 - (b) The open knot sequence should not bind to the intron, as this too could interfere with knot formation, and prevent exon circularization.
4. To assess these criteria, use ViennaRNA package’s RNAcifold program [19]:
 - (a) Download and install/compile the ViennaRNA Package from: <http://www.tbi.univie.ac.at/RNA/index.html>
 - (b) The RNAcifold program is employed with the command: RNAcifold [OPTIONS]
Where instead of [OPTIONS] one inserts the two RNA sequences that interact, concatenated by the ampersand symbol (&), for example, CAUAA&GUGUA.

Multiple concatenated sequences can be imported into the program at once, and the results can be piped into a text file for further manipulation, for example, for Mac/Unix users:

```
RNAcofold CAUAA&GUGUA → exampleTextFile.txt
```

The software takes two RNA sequences and calculates the structure that has the lowest free energy (*see* **Note 4**).

5. Calculate the sum of the minimum free energies associated with: the 5' exon binding to the open knot, and the 3' exon binding to the open knot. This will serve as a numerical proxy for the first criterion. Similarly, adherence to the second criterion can be estimated by calculating the minimum free energy associated with the binding of the open knot sequence to the intron.
6. Select the group I intron with the best indicator value—the lowest free energies—then proceed to further evaluation.
7. Gauge whether both the group I intron and the open knot sequences fold correctly by checking the structure of these sequences, both independently and in combination, with the Vienna RNAfold package [19] and the ILM webserver (Kinefold) [18]. To avoid hybridization between the knot and the circularizing scaffold, model their interaction with RNAcofold. Should the knot and circularizing scaffold interact, restart the process with another circularizing scaffold from **step 5** (or consider choosing another knot sequence/making suitable point mutations).

3.1.2 Tier 2 and 3: Interlocking Structures

1. Model the secondary structure of the lasso with NUPACK:
 - (a) Access and adjust “RNA energy parameters” to “Matthews et al. 1999”.
 - (b) Adjust “Number of strand species” to 2 and “Maximum complex size” to amount required.
 - (c) Follow given instructions in help section of server.
2. Verify the following criteria:
 - The predicted secondary structure of the ribozyme should remain relatively intact, to increase the probability that it folds correctly.
 - The antisense target region should have least four bases accessible to act as a “Toehold”—regions that are single stranded in the target species [20].
 - Should a target region fold into a conformation where some of its base pairs are bound, the total predicted equilibrium probability of this binding should be small (at least below 0.5 base pair probability).

3. Validate results using Ribomaker:
 - (a) Access online server at: <https://absynth.issb.genopole.fr/Bioinformatics/tools/Ribomaker/>.
 - (b) Follow on screen instructions.

3.2 RNA Synthesis

3.2.1 *In Vitro* Transcription

1. Linearize plasmid using a suitable restriction enzyme for your construct (for our constructs, *see* Subheading 2 for a suitable enzyme).
2. Produce RNA *in vitro* using an appropriate transcription kit (e.g., TranscriptAid T7 High Yield kit from Thermo Scientific) (*see* **Note 5**).
3. Purify RNA by phenol–chloroform precipitation extraction [27]:
 - (a) Add to 20 μ L of transcription product 105 μ L DEPC water, and 15 μ L NaAC. Next add 140 μ L phenol–chloroform. Centrifuge for 5 min at 14,100 rcf.
 - (b) Extract top (aqueous) layer; add to this extraction 140 μ L chloroform. Centrifuge at 14,100 rcf for 5 min.
 - (c) Extract top layer, and add to this 280 μ L Ethanol (99.8 %). Store this overnight at -20 °C.
4. Centrifuge for 30 min and discard resulting supernatant. Wash pellet with 500 μ L cold Ethanol (70 %).
5. Resuspend pellet in 25 μ L DEPC water.

3.3 Characterization of Circular RNA

3.3.1 2D Urea PAGE

2D Denature Urea PAGE (Polyacrylamide gel Electrophoresis) takes advantage of the change in migration speed seen in non-linear topologies of nucleic acid when passing through differing polyacrylamide concentrations. Various topologies of nucleic acids can run on a one-dimensional gel without necessarily being distinguishable on the 1D gel alone (Fig. 6 gives example 1D gel result). Running the gel a second time, at a higher concentration of polyacrylamide, separates any circular RNA from the diagonal. This is because the resistance migrating through the polyacrylamide mix is higher for circular RNA than linear strands. Hence, linear RNA migrates approximately at the same speed as in the 1D gel, causing them to move into a diagonal line across the gel, leaving any circular RNA to protrude out from the diagonal (Fig. 4 depicts the trajectory taken by linear and topological RNAs on a 2D PAGE). An example of this procedure is outlined in Fig. 7.

1. Make 6 % polyacrylamide gel and run as normal (*see* **Notes 6** and **7**).
2. Once 1D gel is complete cut out column of interest using a scalpel.

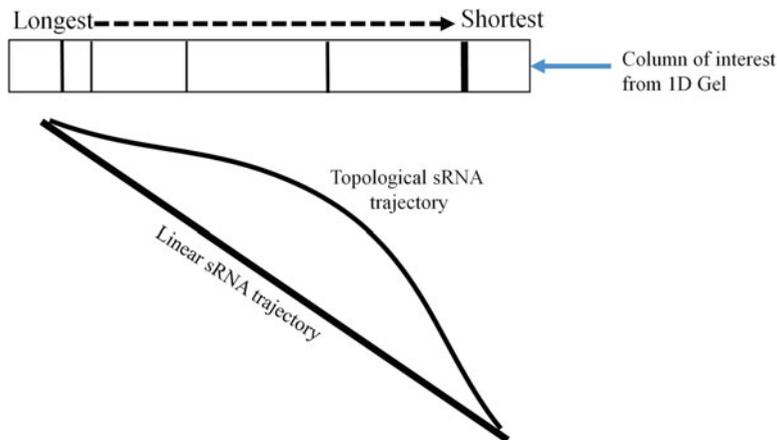


Fig. 4 Denaturing PAGE with RNAs from in vitro transcription of PIE sequence with streptavidin aptamer. Samples were separated by denaturing PAGE and the gel was stained with ethidium bromide. Lane 1, RNA marker with the sizes of the standards indicated on the *left* in nt; lane 2, RNA sample from in vitro transcription of PIE sequence with streptavidin aptamer, DNA plasmid expressing a self-circularized RNA [29]; and lane 3, RNA sample from in vitro transcription of PIE sequence with streptavidin aptamer treated with RNase R. The band containing circular RNA is indicated on the *right* with a *blue arrowhead*

3. Make second gel (10 % polyacrylamide) with room to place column horizontally above (*see* Fig. 4 and **Note 8**). Ensure there are no bubbles between the column and the gel before running.
4. Run the second gel; look for characteristic arc of nonlinear RNA (Figs. 4 and 7).

3.3.2 RNase R Treatment

Using this enzyme to treat the samples loaded into the gels above provides a clearer indication of the conformation of the RNA (Figs. 6 and 7).

1. Treat sample with RNase R, using a non-treated sample as a control. The amount of RNase R required is proportional to the amount of RNA you have; in 10 min at 37 °C one unit converts 1 µg of poly-r(A) into acid-soluble ribonucleotides.
2. To compare samples use a Urea PAGE—6 % polyacrylamide gel (using the same method outlined above) (Fig. 6).
3. The resulting columns from this gel can be carried forth into 2D gel (following the protocol outlined previously). In the second gel, the arc of the nonlinear RNA, and the differing brightness between the bands of linear and nonlinear RNA, indicate the presence of circular RNA (Fig. 7).

3.3.3 RT-PCR

Reverse-transcriptase PCR (RT-PCR) can be used to independently validate circular RNA detection via 2D PAGE. Comparison of two sequences can be achieved using ClusterW2 from <http://www.ebi.ac.uk/Tools/msa/clustalw2/> (Fig. 8). After the RT reaction, a PCR using divergent primers will only yield a band if the RNA tested is circular (Fig. 5 and *see* **Note 9**). Repetition of

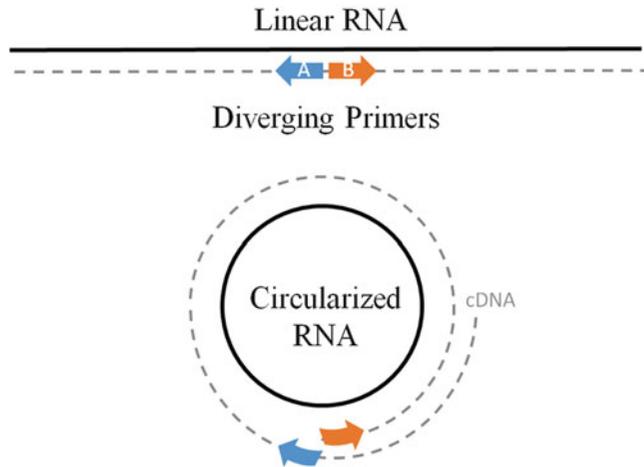


Fig. 5 Trajectories of circular and linear RNA in a 2D Denaturing Urea PAGE [7]

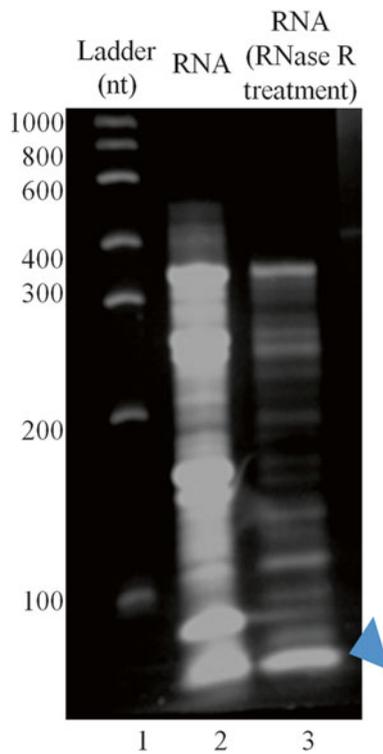


Fig. 6 2D Denaturing PAGE with RNAs from in vitro transcription of PIE sequence with streptavidin aptamer. Columns of 1D gel were separated by denaturing PAGE and the gel stained with ethidium bromide. (a) RNA sample from in vitro transcription of PIE sequence with streptavidin aptamer. (b) RNA sample from in vitro transcription of PIE sequence with streptavidin aptamer treated with RNase R. The positions of the circular RNA are indicated with a *blue arrowhead*

3. Add to solution Ribolock, M-MuLV Reverse Transcriptase and its associated buffer. Using a Thermocycler, heat the solution to 50 °C for 1 min (*see Note 10*), after which raise to 85 °C for 5 min, and finally 65 °C for 20 min [29].
4. Run a standard PCR with reagents outlined in Subheading 2.
5. Product can be visualized on a 1 % agarose gel; each band will correspond to the number of DNA repeats of the circular RNA.
6. The bands can be extracted from the gel and sequenced with either of the primers to ensure circularization and correct ligation site. Extraction Protocol:
 - (a) Weigh gel and add 100 mg of gel for every 300 μL ADB.
 - (b) Heat to 55 °C for 10 min (shake after 5 min).
 - (c) Transfer solution to a spin column and spin down for 1 min. Discard supernatant.
 - (d) Add 200 μL Washing buffer to the column and spin down for 1 min. Discard supernatant and repeat this step.
 - (e) Spin for 1 min to remove residual wash. Transfer column to a new 1.5 mL tube.
 - (f) Add 9 μL Elution buffer, then incubate for 2 min.
 - (g) Spin solution down and dispose of column.

3.4 Characterization of Two Interlocking Rings and Chain System

3.4.1 1D Urea PAGE

RNA lassos needs Mg^{2+} ions to pass between linear and circular conformations, and can be tested in both buffers (one with magnesium and one with EDTA) to characterize the interlocking RNA complexes. The EDTA acts as a sink removing any magnesium present in the solution.

There are four conditions that the lasso and target should operate under to be confident in the characterization of the interlocking system. Run each of the four conditions in separate wells, for both buffers, on a 1D PAGE (Subheading 3.3.1).

1. *Without the target present:* 7 μL lasso, 2 μL buffer, 4 μL formamide (100 %), and 7 μL H_2O . This demonstrates the location of the different lasso forms (linear, half processed and circular) when run on 1D PAGE, so that they can be ruled out in later lanes. For the chain system, run the two lassos in separate lanes in the specified solution.
2. *Lasso with the target (or other lasso) present:* 7 μL lasso, 7 μL target (or other lasso), 2 μL buffer, and 4 μL formamide (100 %). After the 1D gel has run, some bands will be seen at different molecular weights for the two buffer solutions. At this stage it will not be clear which bands contain the interlocking rings, and which are simply molecular dimers (or multimers).
3. *Lasso with the target (or other lasso) present following heat treatment:* 7 μL lasso, 7 μL target (or other lasso), 2 μL buffer, and

4 μL formamide (100 %). Before the Urea loading buffer is added, heat the solutions to 95 °C. Any of the complexes formed between the molecules that are not interlocked will be lost, supporting the case for one band containing interlocking complexes over another.

4. *Lasso with the target (or other lasso) present in RNase R*: 7 μL lasso, 7 μL target (or other lasso), 2 μL buffer, and 4 μL formamide (100 %). Before Urea loading buffer is added, treat with RNase R (*see* Subheading 3.3.2). After this treatment only the interlocking complexes and circular RNA should be present, enabling the distinction of the interlocking system.

3.4.2 Reverse Transcription-PCR (RT-PCR)

Once detected in the gel-based tests, the band believed to be the two-ring interlocking system can be validated using RT-PCR. Two RT-PCRs can be conducted on the sample to verify the presence of both components (of the two interlocking rings or chain system). If both RT-PCRs indicate the presence of the component, it follows that the eluted product is an interlocking complex.

1. Elute out the band presumed to containing the two interlocking rings or chain system.
2. As in the previous RT-PCR design, divergent primers are essential (*see* Subheading 3.3.3 and Fig. 5). Primers should be specific to each lasso/circular target.
3. Run RT-PCR as before, with each set of primers run separately.
4. If both RT-PCRs indicate the presence of the component, it follows that the eluted product is an interlocking complex.

4 Notes

1. NUPACK and RNAfold are based on deterministic models and generate results for shorter sequences almost instantaneously. However, they exclude the possibility of pseudoknots when generating minimum free energy structures for sequences smaller than 100 bp, and Xayaphoummine, Bucher and Isambert [20] found that there is practically no correlation between predictions for low energy structures including and excluding pseudoknots. Kinofold predictions are stochastically generated, and the webserver can handle 100+ bp sequences that may fold into pseudoknots, but it is currently more difficult to extract secondary structure information from. It is advised to run simulations multiple times using different random seeds in order to validate results.
2. If a T7 or SP6 promoter is used, remember to add a G at the 5' of the RNA during secondary structure, as it is part of the

promoter but transcribed by the polymerase. T7 and SP6 may also add a base at the 3' of the RNA.

3. Ribonuclease R (RNase R) is an exoribonuclease that digests linear RNAs. It does not however digest lariat or circular RNA structures, or double stranded RNA with 3' overhangs shorter than seven nucleotides. Most cellular RNAs will be digested completely by RNase R, with the exception of tRNAs, 5S RNA, and intron lariats [30]
4. At the time of writing, RNAcofold appears to estimate the free binding energy between two separate RNA strands by concatenating the strand into one longer strand before finding the minimum free energy of this new structure using the RNAfold algorithm. If this is the case, RNAcofold is inappropriate for investigating interactions between small sequences, but still serves as a good indicator of binding energies for longer sequences [28], especially in consideration of the ease with which data can be imported and exported using RNAcofold. Subsequently, it is advised that steps are taken to validate these binding energies using alternative algorithms following the initial filtering process.
5. PIE products are incubated for 30 min at 55 °C after the transcription to allow the products to circularize. This heating encourages self-ligation of the flanking exons during splicing [29].
6. The PAGE gels can sometimes run out from the bottom of plates when setting. To avoid this, one can set the bottom of the gel by taking 500 μ L gel mix solution and adding to this 1 μ L TEMED, and 8 μ L APS. This will quickly begin to polymerize, so immediately add the solution between the gel plates. This solution will set much faster than the main gel, and so will seal the bottom of the plates before the PAGE gel slips out.
7. To improve the quality of PAGE gels, a pre-run (running the gel with no sample) is recommended as it reduces the "smiling effect" sometimes seen in these gels.
8. Placing the gel column horizontally on top of the new gel can be difficult. If the column gets stuck, or shows signs of resistance when being moved, add water to lubricate gel.
9. Owing to the elaborate secondary structure of the molecules tested with RT-PCR, the yield of the RT-PCR reaction can be low. This is because the secondary structure of these molecules can be complicated, with many loops, resulting in the RNA polymerase sometimes falling off the RNA, reducing the overall yield.
10. Longer incubation times during reverse transcription using RNase H deficient reverse transcriptases will result in a higher number of repeats of the sequence within the resulting DNA. This is due to the transcriptase looping around the RNA circle, but producing linear DNA.

Acknowledgements

Work funded by the grants PROMYS (FP7-KBBE-613745) and EVOPROG (FP7-ICT-610730) to A.J. J.H. is funded by EPSRC and PM additionally acknowledges EPSRC for a PhD studentship through the MOAC Doctoral Training Centre grant number EP/F500378/1. W.R. is funded by a DGA-DSTL PhD fellowship. We thank Dr. So Umekage for all his help and the donation of the PIE T4 with Streptavidin aptamer plasmid. We also thank Ricardo Marco and Jiaye He for their help and support with the bioinformatics.

References

- Guo P (2010) The emerging field of RNA nanotechnology. *Nat Nanotechnol* 5(12):833–842
- Rodrigo G, Landrain TE, Jaramillo A (2012) De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proc Natl Acad Sci* 109:15271–15276. doi:10.1073/pnas.1203831109
- Yin P, Choi HMT, Calvert CR, Pierce NA (2008) Programming biomolecular self-assembly pathways. *Nature* 451(17):318–322
- Rodrigo G, Landrain TE, Majer E, Daròs J-A, Jaramillo A (2013) Full design automation of multi-state RNA devices to program gene expression using energy-based optimization. *PLoS Comput Biol* 9(8):e1003172
- Rostain W, Landrain TE, Rodrigo G, Jaramillo A (2014) Regulatory RNA design through evolutionary computation and strand-displacement. *Methods Mol Biol* (in this edition)
- Hansch C, Maloney P, Fujita T, Muir R (1962) Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. *Nature* 194:178–180
- Jeck W, Sharpless N (2014) Detecting and characterizing circular RNAs. *Nat Biotechnol* 32(5):453–461
- Rostain W, Shen S, Cordero T, Rodrigo G, Jaramillo A (2014) Engineering a circular riboregulator in *Escherichia coli*. bioRxiv. doi:10.1101/123456
- Burton AS (2010) Characterization of novel functions and topologies in RNA. Portland State University PhD Dissertation. Dissertations and theses. Paper 363
- Nielsen H, Fiskaa T, Birgisdóttir A, Haugen P, Einvik C, Johansen S (2003) The ability to form full-length intron RNA circles is a general property of nuclear group I introns. *RNA* 9(12):1464–1475
- Puttaraju M, Been M (1992) Group I permuted intron-exon (PIE) sequences self-splice to produce circular exons. *Nucleic Acids Res* 20(20):5357–5364
- Stahley M, Strobel S (2006) RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis. *Curr Opin Struct Biol* 16(3):319–326
- Salvo J, Coetzee T, Belfort M (1990) Deletion-tolerance and trans-splicing of the bacteriophage T4 td intron. *J Mol Biol* 211(3):537–549
- Bohjanen PR, Colvin RA, Puttaraju M, Been MD, Garcia-Blanco MA (1996) A small circular TAR RNA decoy specifically inhibits Tat-activated HIV-1 transcription. *Nucleic Acids Res* 24:3733–3738
- Nott A, Tsai L-H (2013) The Top3 β way to untangle RNA. *Nat Neurosci* 16:1163–1164
- Dallas A, Balatskaya S, Kuo T, Ilves H, Vlassov A, Kaspar R, Kisich K, Kazakov S, Johnston B (2008) Hairpin ribozyme-antisense RNA constructs can act as molecular lassos. *Nucleic Acids Res* 36(21):6752–6766
- Ge Q, Ilves H, Dallas A, Kumar P, Shorestein J, Kazakov S, Johnston B (2010) Minimal-length short hairpin RNAs: the relationship of structure and RNAi activity. *RNA* 16(1):106–117
- Zadeh J, Steenberg C, Bois J, Wolfe B, Pierce M, Khan A, Dirks R, Pierce N (2011) NUPACK: analysis and design of nucleic acid systems. *J Comput Chem* 32(1):170–173
- Lorenz R, Bernhart S, Höner zu Siederdisen C, Tafer H, Flamm C, Stadler P, Hofacker I (2011) ViennaRNA package 2.0. *Algorithms Mol Biol* 6(1):26
- Xayaphoummine A, Bucher T, Isambert H (2005) Kinefold web server for RNA/DNA folding path and structure prediction including

- pseudoknots and knots. *Nucleic Acids Res* 33(Web Server):W605–W610
21. Rodrigo G, Jaramillo A (2014) RiboMaker: computational design of conformation-based riboregulation. *Bioinformatics* 1:2–4
 22. Zhou Y, Lu C, Wu Q, Wang Y, Sun Z, Deng J, Zhang Y (2007) GISSD: group I intron sequence and structure database. *Nucleic Acids Res* 36(Database):D31–D37
 23. Taufer M et al (2009) PseudoBase++: an extension of PseudoBase for easy searching, formatting and visualization of pseudoknots. *Nucleic Acids Res* 37(Database):D127–D35
 24. Jaeger J, Restle T, Steitz TA (1998) The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor. *EMBO J* 17(15):4535–4542
 25. Suzuki H, Zuo Y, Wang J, Zhang M, Malhotra A, Mayeda A (2006) Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res* 34:63
 26. Vincent HA, Deutscher MP (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. *J Biol Chem* 281:29769–29775
 27. Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162:156–159
 28. Rodrigo G, Prakash S, Shen S, Majer E, Daròs JA, Jaramillo A (2014) Cooperative riboregulation in living cells through allosteric programming of toehold activation. *bioRxiv*. doi:10.1101/009688
 29. Umekage S, Kikuchi Y (2009) In vitro and in vivo production and purification of circular RNA aptamer. *J Biotechnol* 139(4):265–272
 30. Epicentre (An Illumina Company) (2012) Ribonuclease R, *E. coli*. EPILIT 266