nature nanotechnology

ARTICLES

https://doi.org/10.1038/s41565-019-0517-8

Computational design of three-dimensional RNA structure and function

Joseph D. Yesselman[1], Daniel Eiler[2], Erik D. Carlson[3,4,5], Michael R. Gotrik[1], Anne E. d'Aquino[3,5,6], Alexandra N. Ooms[7], Wipapat Kladwang[1], Paul D. Carlson[8], Xuesong Shi[1], David A. Costantino[2], Daniel Herschlag[1,9,10], Julius B. Lucks[3,4,5,6], Michael C. Jewett[3,4,5,6], Jeffrey S. Kieft[2] and Rhiju Das[1,11]*

RNA nanotechnology seeks to create nanoscale machines by repurposing natural RNA modules. The field is slowed by the current need for human intuition during three-dimensional structural design. Here, we demonstrate that three distinct problems in RNA nanotechnology can be reduced to a pathfinding problem and automatically solved through an algorithm called RNAMake. First, RNAMake discovers highly stable single-chain solutions to the classic problem of aligning a tetraloop and its sequence-distal receptor, with experimental validation from chemical mapping, gel electrophoresis, solution X-ray scattering and crystallography with 2.55 Å resolution. Second, RNAMake automatically generates structured tethers that integrate 16S and 23S ribosomal RNAs into single-chain ribosomal RNAs that remain uncleaved by ribonucleases and assemble onto messenger RNA. Third, RNAMake enables the automated stabilization of small-molecule binding RNAs, with designed tertiary contacts that improve the binding affinity of the ATP aptamer and improve the fluorescence and stability of the Spinach RNA in cell extracts and in living *Escherichia coli* cells.

RNA-based nanotechnology is an emerging field that harnesses RNA's unique structural properties to create new nanostructures and machines[1,2]. Perhaps more so than for other biomolecules, the RNA tertiary structure is composed of discrete and recurring components known as tertiary 'motifs'[3]. Along with the helices that they interconnect, many of these structural motifs appear highly modular; that is, each motif folds into a well-defined three-dimensional (3D) structure in a broad range of contexts[2,4–6]. By exploiting symmetry, motif repetition, expert modelling and computational tools for visualization and modelling flexibility, these motifs have been assembled into new polyhedra, sheets and cargo-carrying nanoparticles for biomedical use[7–10]. Despite these advances, current methods still rely on human intuition and the field cannot yet generate RNAs as sophisticated as natural RNA machines, which are asymmetric, too large to be modelled by 3D RNA structure prediction methods and composed of vast repertoires of distinct interacting motifs, most of which are not yet well characterized[11–13].

We present here a new approach to 3D RNA design based on the recognition that numerous recurring problems in the field can be cast into the same 'pathfinding' problem (Fig. 1). First, a founding problem of RNA nanotechnology involves designing a compact nanostructure that aligns the two parts of the tetraloop/tetraloop receptor (TTR) so that they can form a tertiary contact upon RNA chain folding (Fig. 1a). This task requires finding RNA sequences that interconnect the 5′ and 3′ ends of the tetraloop (orange) to the 3′ and 5′ ends of the tetraloop receptor, respectively (blue in Fig. 1a). The problem has previously been solved through a combination of expert manual modelling and the symmetric assembly of multiple chains[5,14]. In all cases, an important guiding principle—sometimes called RNA architectonics[4]—is to design the intermediate RNA chains so that they form RNA modules previously seen in nature, including both canonical double-stranded helices and non-canonical RNA motifs that twist and translate between two desired helical end-points at the tetraloop and the receptor. We call this design task the 'RNA motif pathfinding problem'. The general complexity of this pathfinding task has prevented the design of asymmetric, single-chain solutions to the TTR stabilization problem.

A second problem is highly analogous to the TTR stabilization problem but is more difficult. Efforts to select engineered ribosomes with messenger RNA decoding, polypeptide synthesis and protein excretion functions optimized for new substrates might be dramatically accelerated through the design of integrated ribosomes. An important step towards this goal involves tethering the two 23S and 16S ribosomal RNAs (rRNAs) of the ribosome into a single RNA strand that supports *Escherichia coli* growth[15–18]. 3D designs for the tether require solving the RNA motif pathfinding problem over >100 Å distances and avoiding steric collisions with the ribosome's RNA and protein components (blue and orange strands in Fig. 1b). Even after the identification of appropriate helix end-points, this difficult design challenge previously took more than a year to solve using in vivo assays based on trial-and-error refinement[16,17] or the ad hoc combination of non-canonical motifs without explicit 3D modelling[15,18].

[1]Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA. [2]Department of Biochemistry and Molecular Genetics, University of Colorado Denver School of Medicine, Aurora, CO, USA. [3]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA. [4]Chemistry of Life Processes Institute, Northwestern University, Evanston, IL, USA. [5]Center for Synthetic Biology, Northwestern University, Evanston, IL, USA. [6]Interdisciplinary Biological Sciences Graduate Program, Northwestern University, Evanston, IL, USA. [7]Department of Cancer Genetics & Genomics, Stanford University School of Medicine, Stanford, CA, USA. [8]Robert F. Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY, USA. [9]Department of Chemistry, Stanford University School of Medicine, Stanford, CA, USA. [10]Stanford ChEM-H (Chemistry, Engineering, and Medicine for Human Health), Stanford University, Stanford, CA, USA. [11]Department of Physics, Stanford University, Stanford, CA, USA. *e-mail: rhiju@stanford.edu

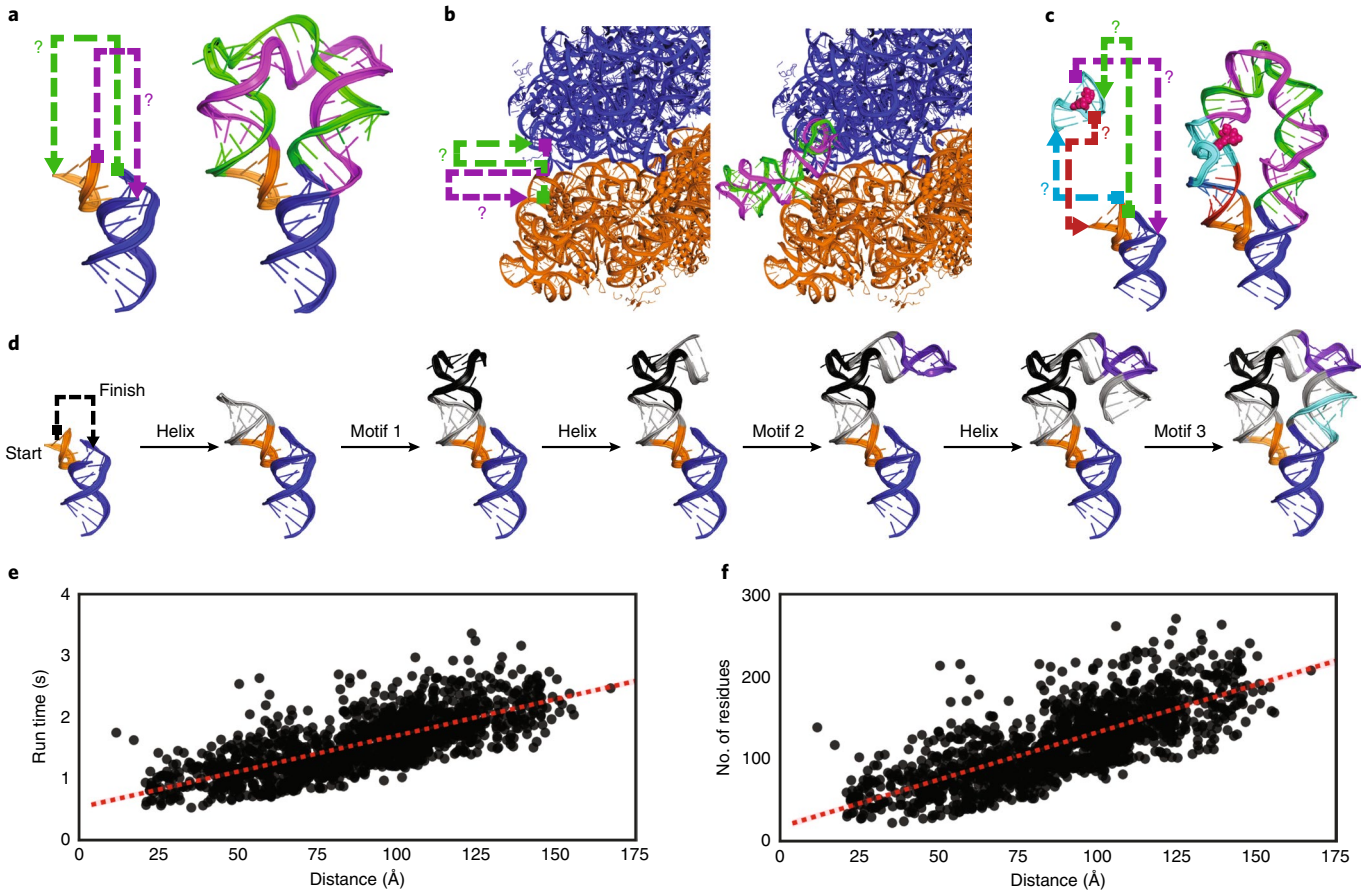


Fig. 1 | Problems in RNA nanotechnology reduced to RNA motif pathfinding problems and solved by RNAMake. a, 'MiniTTRs' require two strands (green and purple) between a tetraloop (orange) and tetraloop receptor (blue). **b**, Tethered ribosomes require two strands (green and purple) to link the small subunit (orange) to the large subunit (blue). **c**, 'Locking' a small-molecule binding aptamer (cyan) by designing four strands (green, purple, teal and red) to a peripheral tertiary contact (orange and blue). Red spheres, ATP molecules. **d**, Demonstration of the RNAMake design algorithm, which builds an RNA path via the successive addition of motifs and helices from a starting base pair to the ending base pair. **e**,**f**, Computational efficiency for RNAMake to design connections between each pair of hairpins on the 50S *E. coli* ribosome. The run time scales linearly with problem size, as measured by the translational distance between helical end-points (**e**) or the number of residues required for segments (**f**) (higher order junctions are utilized in Supplementary Fig. 14).

A third problem involves a more complex instance of two RNA motif pathfinding problems (green, purple, red and teal lines in Fig. 1c). A ubiquitous task in RNA nanotechnology is the selection of 'aptamer' RNAs that sense or carry target small molecules, such as adenosine 5′-triphosphate or fluorophores[19]. Despite recent progress[20,21], improving aptamers requires numerous rounds of tedious selections, with few design tools available to guide consistent improvements. The desired stabilizations might be achieved by peripheral tertiary contacts that extend out of either end of the aptamers and encircle them, bracing them into their functional 3D arrangements (Fig. 1c), analogous to the tertiary contacts that 'lock' natural riboswitch aptamers[22]. However, such rational design has not been carried out due to the difficulty of finding the required four strands that interconnect a given aptamer structure and a tertiary contact.

Here, we present a 3D RNA design algorithm, RNAMake, that solves all three cases of the RNA motif pathfinding problem described above. Gauntlets of structural and functional measurements test that these computationally designed nanostructures, ribosomes and ATP and fluorescent RNA aptamers achieve their design goals, without the need for any further rounds of trial and error.

The RNAMake algorithm and motif library

RNAMake uses a 3D motif library drawn from all unique, publicly deposited crystallographic RNA structures and an efficient algorithm to discover combinations of these motifs and helices that solve the RNA motif pathfinding problem (Methods and Supplementary Table 1). The final set of non-canonical motifs contained 461 unique two-way junctions, 61 higher-order junctions, 290 variable-length hairpins and 89 tertiary contacts. The pathfinding algorithm assembles canonical helical segments that range from 1 to 22 base pairs with these non-canonical structural motifs, step-by-step in a depth-first search (Fig. 1d and Methods). The canonical helical segments are idealized and sequence invariant[23]; after completion of the 3D structural designs, they are filled in with sequences that best match the target secondary structure and minimize alternative secondary structures[24]. Owing to its efficient algorithmic implementation, RNAMake is able to find solutions rapidly; the run time scales linearly with the problem size, and the discovery of exceptionally long double-stranded RNA paths that snake around the entire ribosome takes less than 3 s (run on a Macbook Pro 2016, 2.9 GHz Intel Core i7) (Fig. 1e,f).

RNAMake TTR designs achieve high stability

The problem of creating a well-folded RNA nanostructure was first solved two decades ago by repurposing the well-characterized TTR tertiary contact to bring together two separate RNA chains[5], analogous to the P4–P6 domain of the *Tetrahymena* group I self-splicing intron and other natural functional RNAs. Although later

RNA nanotechnology studies used the TTR module and other structural motifs to design different nanostructures, the original and later designs are all multichain assemblies[25–29]. We chose to test RNAMake on the TTR problem because of the prospect of achieving de novo single-chain solutions to this fundamental problem, which we hypothesized might also help crystallization. We generated 16 diverse single-chain solutions with RNAMake, which we called 'miniTTR' designs.

Standard biochemical and biophysical assays for the RNA structure confirmed folding for the majority of the miniTTR designs. We tested the miniTTR RNAs for the correct secondary structure and tertiary contact formation with single nucleotide resolution chemical mapping (SHAPE (selective 2′-hydroxyl acylation analysed by primer extension) and DMS (dimethyl sulfate) (Supplementary Fig. 1b); Fig. 2a gives DMS at the tetraloop and receptor nucleotides), for compact folds through native gel electrophoresis and mutational analysis (Fig. 2b and Supplementary Fig. 1c) and for tertiary stability through Mg^{2+} binding curves (Fig. 2c,d and Supplementary Table 2). Overall, 11 of the 16 designs passed these experimental screens (details given in Supplementary results and Supplementary Table 3). Several miniTTR constructs required less than 1 mM Mg^{2+} to fold stably, similarly to or better than reported midpoints for natural TTR-containing RNA nanostructures. Indeed, miniTTR 2 and miniTTR 6 exhibited folding stabilities better than that of the P4–P6 RNA in side-by-side assays (Fig. 2c). Furthermore, miniTTR 6 had a much sharper Mg^{2+} dependence than P4–P6 with an apparent Hill coefficient of over ten (Fig. 2c). The stability of the RNAMake designs was particularly notable given that P4–P6 and other natural TTR-containing RNAs are larger than the miniTTR designs and have additional stabilizing tertiary contacts[30–32] and that other attempts to make artificial minimized TTR constructs gave significantly worse stabilities[33].

After the gel-based and chemical mapping tests, we tested whether the RNAMake designs might allow crystallization and thereby enable high-resolution characterization of the structural accuracy of the designs. After small-angle X-ray scattering (SAXS) measurements confirmed a monomeric structure even at high RNA concentrations (>1 μM) (Fig. 2e,f, Supplementary results and Supplementary Fig. 2a,b), we were able to obtain crystals of miniTTR 6 that diffracted at 2.55 Å resolution ($1/\sigma$ of 1.0) (Fig. 2g and Supplementary Table 10). The crystal structure and the RNAMake model agreed with an all-heavy-atom root mean squared deviation (r.m.s.d.) of 4.2 Å, better than the nanometre-scale accuracy typically sought in RNA nanotechnology. The primary discrepancy between our modelled 3D structure and the crystal structure was a single motif, a triple mismatch drawn from the large ribosomal subunit (Fig. 2h, right). This motif formed multiple consecutive non-canonical base pairs with high B factors in our miniTTR 6 crystal instead of the conformation found in the ribosomal structure, which involved flipped-out adenosines (residues O2360–O2363 and O2424–O2426 (Protein Data Bank (PDB) 1S72)) (Fig. 2h, left). Other motifs in the design achieved near-atomic accuracy, including the TTR tertiary contact (r.m.s.d. 0.45 Å (Fig. 2i)), a kink-turn variant drawn from the archaeal 50S ribosomal subunit (r.m.s.d. 2.0 Å (Fig. 2j))[34] and a 'right-angle turn' drawn from a viral internal ribosomal entry site domain (r.m.s.d. 1.28 Å (Fig. 2k))[25].

Automated 3D design of tethered ribosomal subunits

After testing RNAMake's performance in designing compact RNA nanostructures, we evaluated whether it might solve a practical problem involving nanostructures that must traverse long distances (compare Fig. 1a,b). The ribosome is a ribonucleoprotein machine dominated by two extensive RNA subunits, the 16S and 23S rRNAs. In previous work, we constructed a tethered ribosome called Ribo-T in which the large and small subunit rRNAs were connected by an RNA tether to form a single subunit ribosome[17]. In that work, the major bottleneck involved more than a year of numerous trial-and-error iterations to identify RNA tethers that were not cleaved by ribonucleases in vivo when wild-type ribosomes were replaced in the Squires strain of *E. coli*[17]. The Squires strain cells lack genetic rRNA alleles, surviving off plasmids that can be exchanged using positive and negative selections. Early failure rounds that involved ribosomes from our and other studies are shown in Fig. 3a,b and the success with Ribo-T in Fig. 3c. Nevertheless, the current tethers in Ribo-T are unstructured and unlikely to remain stable if other modules are incorporated (Fig. 3c). We hypothesized that an automated design by RNAMake might give structured, chemically stable tethers for this design problem.

RNAMake generated 100 designs (RM-Tethers), which contained either four or five non-canonical structural motifs each (Methods gives the motif selections), to tether the H101 helix on a circularly permuted 23S rRNA to the h44 helix on the 16S rRNA (Fig. 1b and Supplementary Fig. 3b). Of the nine diverse solutions we tested (RM-Tether 1–9), DNA templates for seven could be synthesized, and the transformation of these DNA templates into the Squires strain allowed us to assay whether the RNAMake designs could replace wild-type ribosomes deleted from growing bacteria. One of these seven constructs, RM-Tether 4, led to the viable growth of bacterial colonies. DNA sequencing confirmed that these colonies harboured the correct RM-Tether 4 plasmid; and RNA electrophoresis confirmed the presence of a single dominant RNA species with the same length as Ribo-T, with no detectable products that corresponded to separate 16S or 23S rRNA lengths or other cleavage products (Fig. 3d). Although the growth rate of this strain was low (Supplementary Fig. 4d), we were able to confirm independently that the ribosomes loaded on messenger RNA in vitro, using integrated synthesis, assembly, and translation (iSAT) in ribosome-free S150 extracts[35,36]. Similar to Ribo-T[16], we detected 70S/monosome[37] and polysomes (and no 30S or 50S subunits) by separation of the iSAT-prepared RM-Tether 4 ribosomes on a sucrose gradient (Fig. 3e and Methods). Electrophoresis of the polysome fraction confirmed that it contained an uncleaved rRNA the same size as Ribo-T (Fig. 3f). In addition, SHAPE-Seq mapping on this rRNA confirmed that the RM-Tether 4 can be reverse transcribed from one ribosomal subunit to the other across both strands of the tether and highlights a chemical reactivity profile consistent with the design, with one region of flexibility around the middle junction (Supplementary Fig. 5). Taken together, these data demonstrate that RNAMake-designed ribosomes with structured, chemically stable tethers can replace wild-type ribosomes in vivo and more than one such ribosome can be loaded onto a single message in vitro. RNAMake obviates the repeated rounds of trial and error that were previously required to achieve these design goals.

RNAMake stabilizes small-molecule binding aptamers

As a final series of tests, we evaluated whether RNAMake could solve 3D design problems whose complexity precluded prior progress even with trial and error or large-scale library selections. Small molecules can be bound and sensed by artificially selected RNA aptamers. Unfortunately, these molecules often exhibit weakened binding affinities or instability in biological environments and additional rounds of selection to improve aptamers typically give diminishing returns[38–40]. By expanding RNAMake to allow the design of interconnections between multiple pairs of helices (Fig. 1c), we tested the hypothesis that the computational design of peripheral tertiary contacts might 'lock' these artificial aptamers into their bound conformation even in the absence of a ligand. By reducing the number of alternative structures available in the unbound state, such locking contacts could selectively increase the free energy of the unbound state and thereby improve the free-energy difference to the bound state, and so lead to a better affinity to small molecule targets.

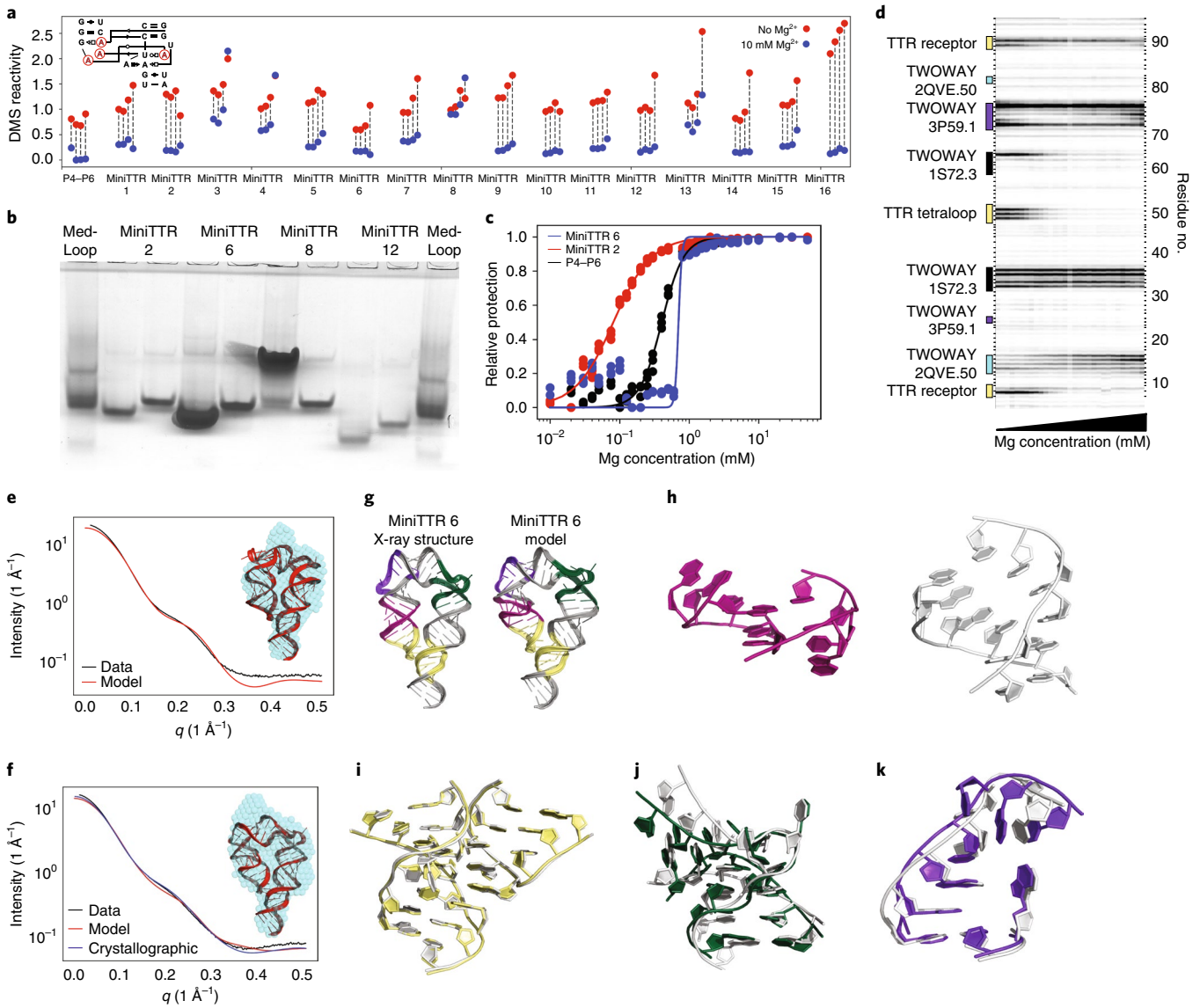


Fig. 2 | Solving the TTR design problem. a, Quantification of DMS reactivity in the absence of Mg^{2+} and with 10 mM Mg^{2+} for RNAMake-designed miniTTR constructs and the P4–P6 domain of the *Tetrahymena* ribozyme as a large natural RNA comparison. Inset: the four adenosines in the TTR (red circles) that undergo protection on Mg^{2+}-dependent folding. **b**, Native gel assays testing whether mutation of the GAAA tetraloop (left lane) to the UUCG mutant (right lane) disrupts the miniTTR tertiary fold and slows its mobility. 'MedLoop' lanes are a control RNA with similar size. Supplementary Fig. 1c gives other constructs. **c**, Quantification of the miniTTR folding stability based on Mg^{2+} binding curves read out by DMS mapping for miniTTR 6, miniTTR 2 and P4–P6. **d**, Raw data from the Mg^{2+} titration (concentrations: 0.00, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00, 1.25, 1.50, 1.75, 2.00, 2.50, 3.00, 3.50, 4.00, 5.00, 7.50, 10.00 mM) of miniTTR 2 highlight the change in DMS reactivity in the TTR and the motifs used in the design. **e**,**f**, SAXS analysis: experimental intensity versus scattering amplitude and low-resolution reconstruction derived from experimental scattering profiles (blue beads, inset) overlaid on the RNAMake-designed model (cartoon, inset) for miniTTR 2 (**e**) and miniTTR 6 (**f**). In **f**, the SAXS prediction from the miniTTR crystal structure is also shown (blue line). **g**–**l**, X-ray crystal structure of miniTTR 6 tests the accuracy of the RNAMake model at atomic resolution: overall RNAMake and X-ray structure (**g**), magnified views of a triple mismatch motif (right) from the ribosome (left) (**h**), TTR (**i**), kink-turn motif (**j**) and right-angle turn (**k**). In **h**–**k**, the crystal structures are white.

First, we sought to stabilize a classic aptamer for adenosine 5′-triphosphate and adenosine 5′-monophosphate (ATP and AMP, respectively), which is in wide use in RNA nanotechnology but whose binding has not been appreciably improved since its discovery in 1993[41–45]. In total, we tested ten ATP aptamers embedded by RNAMake into scaffolds with tetraloop–receptor contacts, which we called ATP-TTR designs (Fig. 4a and Supplementary Fig. 6; Methods describes the modelling of helix flexibility used for these designs). Chemical mapping confirmed that four of these RNAs formed the TTR and also retained their ability to bind to ATP, as assessed by the DMS protection of aptamer nucleotides A13 and A14 (Supplementary Table 4 and Supplementary Fig. 7). Titrations of ATP read out through chemical mapping (Fig. 4d and Supplementary Table 4) showed that three designs achieved better ATP dissociation constants (K_d of 1.5, 4.1 and 1.4 μM) than that of the isolated ATP aptamer under the same conditions (K_d = 16.2 μM), improvements by up to an order of magnitude. Three of the ATP-TTRs gave ligand-free DMS reactivity profiles in the aptamer regions similar to that of the ligand-bound aptamer, which suggests that they preform the structure needed for ATP binding rather than

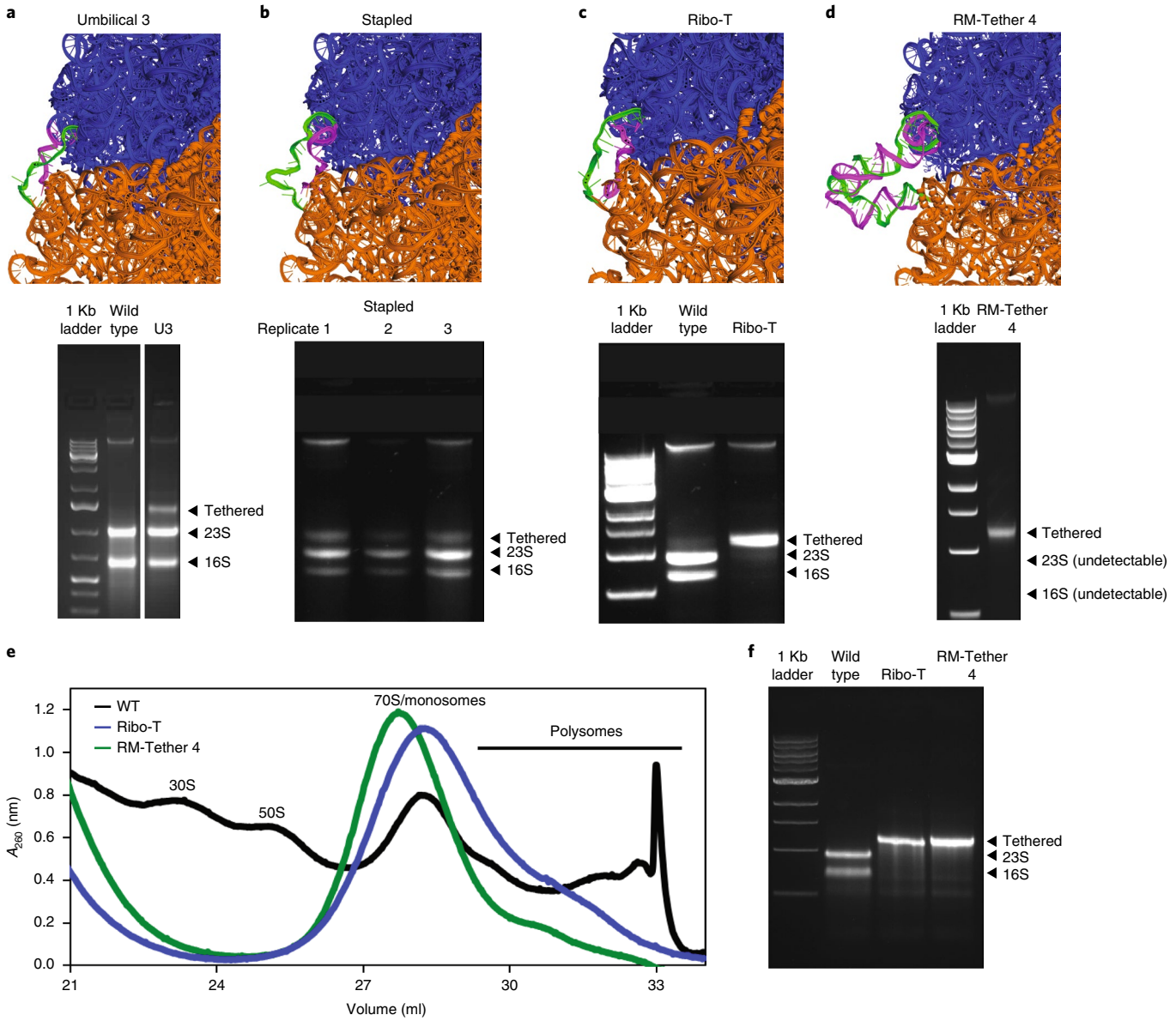


Fig. 3 | The ribosome tethering problem. a–d, Modelling (top panels) of tethers (green and magenta strands) that connect 16S and 23S rRNA into a single rRNA, and agarose gel electrophoresis (bottom panels) of RNA extracted from *E. coli* (Squires strain) in which wild-type ribosomes were completely replaced with the designed molecules (top): an early U3 tether (Umbilical 3) designed by the Jewett lab, cleaved into two bands in vivo (**a**), a stapled ribosomes developed by a separate group[15] (three replicates shown), also cleaved in this plasmid context and strain in vivo (**b**), the successful Ribo-T design developed after a year of manual trial and error to withstand cleavage in vivo (**c**) and RM-Tether 4, a design automatically generated by RNAMake, which also presents as a single band in vivo (**d**). **e**, Sucrose gradient fractionation prepared from in vitro iSAT reactions that express wild-type ribosomes, Ribo-T version 1.0 and RM-Tether 4. Peaks correspond to small subunits (30S), large subunits (50S), monosomes/70S and polysomes (standard assignments of the peaks are given in, for example, Underwood et al.[37]). **f**, Agarose electrophoresis analysis confirms that the polysome fraction of **e** is composed of tethered ribosomes. Full gels are given in Supplementary Figs. 5 and 6.

require conformational rearrangements observed in the isolated ATP aptamer (Fig. 4b,c and Supplementary Table 4)[41]. These results demonstrate that the TTR peripheral contact efficiently couples to enhance the binding of ATP in the aptameric region, as desired. As a further test of this coupling, we confirmed that the Mg^{2+} requirements to form the TTR were reduced in the presence compared to the absence of the small-molecule ligand in these constructs (Supplementary results and Supplementary Fig. 7b).

As a second test of aptamer stabilization, we assessed whether RNAMake could stabilize the Spinach RNA, which binds an analogue of the green fluorescent protein chromophore (*Z*)-4-(3,5-difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one (DFHBI) within a G-quadruplex. Binding to Spinach enhances the fluorescence of DFHBI by ~1,000-fold relative to that of unbound ligand, which makes this RNA useful for biological interrogations[39,46], although its binding affinity, brightness, folding efficiency and biological stability remain poor even after extensive efforts to discover improvements, such as the minimized Spinach and Broccoli aptamers[47–50]. We characterized 16 'Spinach-TTR molecules' designed by RNAMake to embed the Spinach aptamer into scaffolds with tetraloop–receptor contacts (Fig. 4e and Supplementary Fig. 8). SHAPE chemical mapping confirmed that

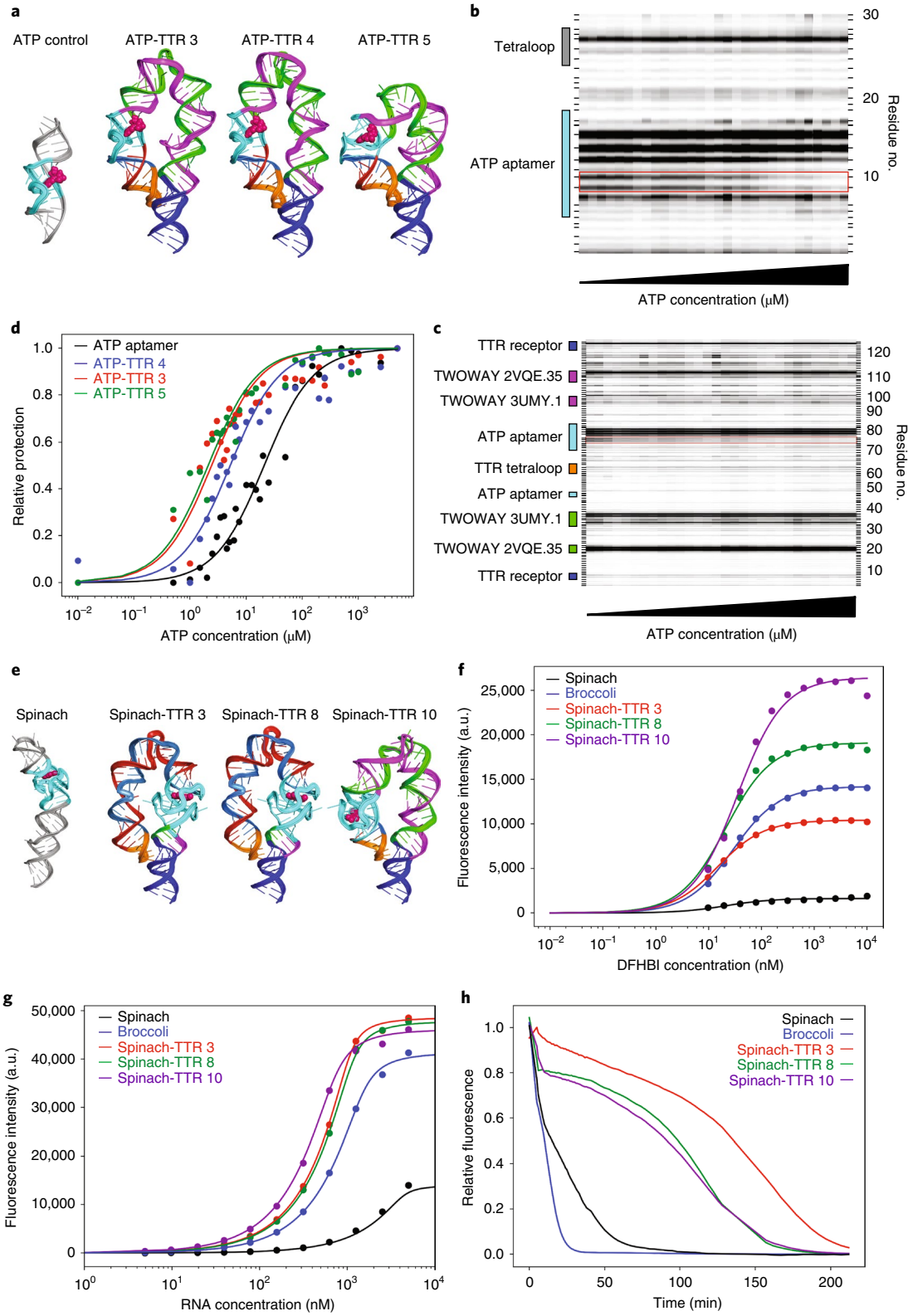


Fig. 4 | Stabilizing aptamers for ATP and light-up fluorophores through designer tertiary contacts. a, 3D models of the ATP aptamer alone and of ATP-TTR 3, 4 and 5. **b**, DMS probing of ATP titration of the ATP aptamer; red box denotes the two adenines that become protected on addition of ATP. **c**) DMS probing of ATP titration of ATP-TTR 4; the red box denotes the two adenines that become protected on the addition of ATP. **d**, Quantified DMS protection as a function of ATP concentration for the ATP aptamer compared to those of ATP-TTR 3, 4 and 5. **e**, The crystal structure of the Spinach aptamer bound to DFHBI (left) and 3D models of Spinach-TTR 3, 8 and 10. **f**, Fluorescence measurements of a DFHBI titration at constant RNA concentrations (200 nM) of Spinach, Broccoli and Spinach-TTR 3, 8 and 10. **g**, Fluorescence measurements of an RNA titration with constant DFHBI concentrations (400 nM) of Spinach, Broccoli and Spinach-TTR 3, 8 and 10. **h**, Fluorescence in 20% *X. laevis* egg extract as compared to the construct in buffer for Spinach, Broccoli and Spinach-TTR 3, 8 and 10. a.u., arbitrary units.

Table 1 | RNAMake designs and component motifs generated for all four design problems

Design problem	Designs	Successful designs	Motifs used	Experimentally validated motifs
MiniTTR	16	11	23	20
Ribosome tether	9	1	14	3
ATP-TTR	10	3	16	6
Spinach-TTR	16	13	26	20
All	51	28	56[a]	41[a]

[a]Some motifs were used in multiple design problems.

these molecules form tetraloop–receptor contacts in 13 of the 15 cases that could be tested (Supplementary Fig. 9 and Supplementary Table 5). By carrying out fluorescence assays that titrated both RNA and DFHBI concentrations, we evaluated the dissociation constants, brightness and folding efficiency of these designs. Seven of the 16 Spinach-TTR designs exhibited a twofold brighter fluorescence than that of the original Spinach as well as a brighter Broccoli aptamer (Fig. 4f, Supplementary Fig. 10 and Supplementary Tables 5 and 6). Two of these constructs, Spinach-TTR 3 and 8, were not only brighter but also gave a higher affinity and improved folding efficiency relative to that of Broccoli and a minimized Spinach construct, Spinach-min (Fig. 4g and Supplementary Tables 5 and 6).

We hypothesized that these improvements to in vitro stability measures might also lead to an improved stability in a harsh biological environment. When the DFHBI-bound RNAs were challenged with 20% whole cell lysate extracted from the eggs of *Xenopus laevis*, six of the seven Spinach-TTR constructs exhibited fluorescence for a longer time than the control Spinach and Broccoli sequences (Methods). Spinach-TTR 3 exhibited a particularly high stability (Fig. 4h), with a time to half fluorescence of 131 minutes, compared to <20 minutes for Spinach, Spinach-min and Broccoli (Supplementary Table 6 and Supplementary Fig. 11). This same robust fluorescence of the Spinach-TTRs was observed in 20% *E. coli*. lysate, which suggests a general stabilization in biological environments (Supplementary Fig. 12). We finally sought to assess the ability of the Spinach-TTR constructs to activate fluorescence in cells, using *E. coli* as a test bed. Six Spinach-TTR designs were cloned into a plasmid for T7 RNA polymerase-driven expression (Methods). Each Spinach-TTR variant was able to significantly activate expression above the background, and several designs exceeded the fluorescence observed for both Spinach and Broccoli in vivo (Supplementary Fig. 13).

Conclusions

As RNA nanotechnology seeks to create artificial molecules closer in sophistication to natural RNA molecules, the design of tertiary structures that are as complex, asymmetric and diverse as natural RNAs becomes an important goal. Here, we hypothesized that several distinct tasks in designing complex RNA tertiary structures might be reduced to instances of a single RNA motif pathfinding problem and developed the algorithm RNAMake to solve the pathfinding task (Table 1 and Supplementary results). For the miniTTR nanostructure design problem, 11 of 16 molecules exhibited the correct tertiary fold in nucleotide-resolution chemical mapping and electrophoresis assays, and we achieved a crystal structure of one design that confirmed its accuracy at a high resolution. For the problem of tethering *E. coli* 16S and 23S rRNAs into a single RNA molecule, one of nine RNAMake-designed molecules replaced ribosomes in vivo and was confirmed to translate in polysomes in cell-free translation reactions. For the problem of stabilizing aptamers through locking tertiary contacts, 3 of 10 RNAMake-designed ATP-TTR molecules achieved an improved affinity to ATP compared to that of the starting aptamer, and seven of 16 Spinach-TTR designs maintained their binding affinity for the DFHBI fluorophore while achieving improvements in fluorescence and folding efficiency in vitro, and in stability in extracto and in vivo. In each task, RNAMake achieved its design objectives in a single round of tests that involved the parallel synthesis of 8–16 constructs, without further trial-and-error iteration.

As RNAMake is applied to more problems, we expect its success rate to improve further. Accumulating knowledge as to which structural motifs recur in successful versus failing designs may allow an empirical scoring for the modularity of each motif; inferences for some motifs, such as A–A mismatches, are already possible (Supplementary discussion). Second, the incorporation of motifs that are known to sample at least two conformations (for example, the triple mismatch in miniTTR6 herein or kink turns) may allow an improved design of such machines as the ribosome, and improved cryogenic electron microscopy methods may provide more detailed feedback on such distinct states[18,51]. Third, natural structured RNAs often contain multiple tertiary contacts and multibranched junctions and we have extended RNAMake's pathfinding method to design such motifs (Supplementary Fig. 14). Finally, we expect RNAMake's computational design approach to be complementary to library selection and high-throughput screening methods[52], especially for larger problems that require numerous non-canonical motifs. By distributing RNAMake as a source code and a server, we hope to encourage these applications and extensions of computational RNA design.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41565-019-0517-8.

Received: 26 July 2017; Accepted: 24 June 2019;
Published online: 19 August 2019

References

1. Guo, P. The emerging field of RNA nanotechnology. *Nat. Nanotechnol.* **5**, 833–842 (2010).
2. Grabow, W. W. & Jaeger, L. RNA self-assembly and RNA nanotechnology. *Acc. Chem. Res.* **47**, 1871–1880 (2014).
3. Leontis, N. B., Lescoute, A. & Westhof, E. The building blocks and motifs of RNA architecture. *Curr. Opin. Struct. Biol.* **16**, 279–287 (2006).
4. Jaeger, L. & Chworos, A. The architectonics of programmable RNA and DNA nanostructures. *Curr. Opin. Struct. Biol.* **16**, 531–543 (2006).
5. Jaeger, L. & Leontis, N. B. Tecto-RNA: one-dimensional self-assembly through tertiary interactions. *Angew. Chem. Int. Ed.* **39**, 2521–2524 (2000).
6. Zhang, H. et al. Crystal structure of 3WJ core revealing divalent ion-promoted thermostability and assembly of the Phi29 hexameric motor pRNA. *RNA* **19**, 1226–1237 (2013).
7. Weizmann, Y. & Andersen, E. S. RNA nanotechnology—the knots and folds of RNA nanoparticle engineering. *MRS Bull.* **42**, 930–935 (2017).
8. Jasinski, D., Haque, F., Binzel, D. W. & Guo, P. Advancement of the emerging field of RNA nanotechnology. *ACS Nano* **11**, 1142–1164 (2017).
9. Afonin, K. A. et al. Computational and experimental characterization of RNA cubic nanoscaffolds. *Methods* **67**, 256–265 (2014).
10. Jossinet, F., Ludwig, T. E. & Westhof, E. Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. *Bioinformatics* **26**, 2057–2059 (2010).
11. Wimberly, B. T. et al. Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339 (2000).
12. Nguyen, T. H. D. et al. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* **523**, 47–52 (2015).
13. Miao, Z. et al. RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA* **23**, 655–672 (2017).
14. Nasalean, L., Baudrey, S., Leontis, N. B. & Jaeger, L. Controlling RNA self-assembly to form filaments. *Nucleic Acids Res.* **34**, 1381–1392 (2006).
15. Fried, S. D., Schmied, W. H., Uttamapinant, C. & Chin, J. W. Ribosome subunit stapling for orthogonal translation in *E. coli*. *Angew. Chem. Int. Ed.* **54**, 12791–12794 (2015).

16. Orelle, C. et al. Protein synthesis by ribosomes with tethered subunits. *Nature* **524**, 119–124 (2015).
17. Carlson, E. D. Creating Ribo-T: (design, build, test)^n. *ACS Synth. Biol.* **4**, 1173–1175 (2015).
18. Schmied, W. H. et al. Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature* **564**, 444–448 (2018).
19. Famulok, M. Oligonucleotide aptamers that recognize small molecules. *Curr. Opin. Struct. Biol.* **9**, 324–329 (1999).
20. Porter, E. B., Polaski, J. T., Morck, M. M. & Batey, R. T. Recurrent RNA motifs as scaffolds for genetically encodable small-molecule biosensors. *Nat. Chem. Biol.* **13**, 295–301 (2017).
21. Gotrik, M. et al. Direct selection of fluorescence-enhancing RNA aptamers. *J. Am. Chem. Soc.* **140**, 3583–3591 (2018).
22. Montange, R. K. & Batey, R. T. Riboswitches: emerging themes in RNA structure and function. *Annu. Rev. Biophys.* **37**, 117–133 (2008).
23. Macke, T. J. & Case, D. A. in *Molecular Modeling of Nucleic Acids* (eds Leontis, N. B. & SantaLucia, J.) 379–393 (American Chemical Society, 1997).
24. Lee, J. et al. RNA design rules from a massive open laboratory. *Proc. Natl Acad. Sci. USA* **111**, 2122–2127 (2014).
25. Dibrov, S. M., McLean, J., Parsons, J. & Hermann, T. Self-assembling RNA square. *Proc. Natl Acad. Sci. USA* **108**, 6405–6408 (2011).
26. Afonin, K. A. et al. Multifunctional RNA nanoparticles. *Nano Lett.* **14**, 5662–5671 (2014).
27. Khisamutdinov, E. F. et al. Fabrication of RNA 3D nanoprisms for loading and protection of small RNAs and model drugs. *Adv. Mater.* **28**, 10079–10087 (2016).
28. Bindewald, E., Grunewald, C., Boyle, B., O'Connor, M. & Shapiro, B. A. Computational strategies for the automated design of RNA nanoscale structures from building blocks using NanoTiler. *J. Mol. Graph Model* **27**, 299–308 (2008).
29. Huang, L. & Lilley, D. M. J. A quasi-cyclic RNA nano-scale molecular object constructed using kink turns. *Nanoscale* **8**, 15189–15195 (2016).
30. Wu, L., Chai, D., Fraser, M. E. & Zimmerly, S. Structural variation and uniformity among tetraloop–receptor interactions and other loop–helix interactions in RNA crystal structures. *PLoS ONE* **7**, e49225 (2012).
31. Frederiksen, J. K., Li, N.-S., Das, R., Herschlag, D. & Piccirilli, J. A. Metal-ion rescue revisited: biochemical detection of site-bound metal ions important for RNA folding. *RNA* **18**, 1123–1141 (2012).
32. Rangan, P., Masquida, B., Westhof, E. & Woodson, S. A. Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA* **100**, 1574–1579 (2003).
33. Fiore, J. L. & Nesbitt, D. J., An RNA folding motif: GNRA tetraloop–receptor interactions. *Q Rev. Biophys.* **46**, 223–264 (2013).
34. Klein, D. J., Schmeing, T. M., Moore, P. B. & Steitz, T. A. The kink-turn: a new RNA secondary structure motif. *EMBO J.* **20**, 4214–4221 (2001).
35. Jewett, M. C., Fritz, B. R., Timmerman, L. E. & Church, G. M. In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Mol. Syst. Biol.* **9**, 678 (2013).
36. Fritz, B. R., Jamil, O. K. & Jewett, M. C. Implications of macromolecular crowding and reducing conditions for in vitro ribosome construction. *Nucleic Acids Res.* **43**, 4774–4784 (2015).
37. Underwood, K. A., Swartz, J. R. & Puglisi, J. D. Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. *Biotechnol. Bioeng.* **91**, 425–435 (2005).
38. Carothers, J. M., Oestreich, S. C. & Szostak, J. W. Aptamers selected for higher-affinity binding are not more specific for the target ligand. *J. Am. Chem. Soc.* **128**, 7929–7937 (2006).
39. Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **333**, 642–646 (2011).
40. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
41. Jiang, F., Kumar, R. A., Jones, R. A. & Patel, D. J. Structural basis of RNA folding and recognition in an AMP–RNA aptamer complex. *Nature* **382**, 183–186 (1996).
42. Huang, Z. & Szostak, J. W. Evolution of aptamers with a new specificity and new secondary structures from an ATP aptamer. *RNA* **9**, 1456–1463 (2003).
43. Sassanfar, M. & Szostak, J. W. An RNA motif that binds ATP. *Nature* **364**, 550–553 (1993).
44. Sazani, P. L., Larralde, R. & Szostak, J. W. A small aptamer with strong and specific recognition of the triphosphate of ATP. *J. Am. Chem. Soc.* **126**, 8370–8371 (2004).
45. Geary, C., Chworos, A., Verzemnieks, E., Voss, N. R. & Jaeger, L. Composing RNA nanostructures from a syntax of RNA structural modules. *Nano Lett.* **17**, 7095–7101 (2017).
46. Kellenberger, C. A., Chen, C., Whiteley, A. T., Portnoy, D. A. & Hammond, M. C. RNA-based fluorescent biosensors for live cell imaging of second messenger cyclic di-AMP. *J. Am. Chem. Soc.* **137**, 6432–6435 (2015).
47. Strack, R. L., Disney, M. D. & Jaffrey, S. R. A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat. Methods* **10**, 1219–1224 (2013).
48. Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **136**, 16299–16308 (2014).
49. Ketterer, S., Fuchs, D., Weber, W. & Meier, M. Systematic reconstruction of binding and stability landscapes of the fluorogenic aptamer Spinach. *Nucleic Acids Res.* **43**, 9564–9572 (2015).
50. Song, W., Strack, R. L., Svensen, N. & Jaffrey, S. R. Plug-and-play fluorophores extend the spectral properties of Spinach. *J. Am. Chem. Soc.* **136**, 1198–1201 (2014).
51. Shi, X., Huang, L., Lilley, D. M. J., Harbury, P. B. & Herschlag, D. The solution structural ensembles of RNA kink-turn motifs and their protein complexes. *Nat. Chem. Biol.* **12**, 146–152 (2016).
52. Buenrostro, J. D. et al. Quantitative analysis of RNA–protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nat. Biotechnol.* **32**, 562–568 (2014).

Acknowledgements

We thank S. Bonilla for assistance in performing the native gel assays and A. Watkins for discussions about ribosome tether design. We thank the Straight lab for graciously providing *X. laevis* whole cell lysate. This work was supported by the National Institutes of Health, through NIGMS MIRA R35 GM122579 (R.D.), R01 GM121487 (R.D.), New Innovator Award 1DP2GM110838 (J.B.L.), Ruth L. Kirschstein National Research Service Award Postdoctoral Fellowships GM112294 (J.D.Y.) and GM100953 (D.E.), P01 GM066275 (D.H.) and R35 GM118070 (J.S.K.), a Stanford School of Medicine Discovery Innovation Award (R.D.), Army Research Office W911NF-16-1-0372 (M.C.J.), the National Science Foundation through MCB-1716766 (M.C.J.), Career Award 1452441 (J.B.L.) and Graduate Research Fellowship DGE-1324585 (A.E.D.), the David and Lucile Packard Foundation (M.C.J.) and the Camille Dreyfus Teacher-Scholar Program (M.C.J.).

Author contributions

R.D. and J.D.Y. conceived the study. J.D.Y. developed RNAMake and generated the models and sequences used throughout the study. A.N.O., W.K. and J.D.Y. performed the chemical mapping, titrations and native gel assays for miniTTR constructs. X.S. and D.H. designed and performed the SAXS on miniTTR 2 and 6. D.E. solved the miniTTR crystal structure assisted by D.A.C. and J.S.K. in preparing the RNA and analysis. E.D.C., A.E.D. and M.C.J. made and tested the RNAMake-designed ribosomes P.D.C. and J.B.L. carried out the SHAPE-seq and in vivo tests of Spinach-TTRs. M.R.G. performed fluorescence and lysate experiments on Spinach-TTRs. J.D.Y. and R.D. wrote the paper, with participation by all the authors.

Competing interests

Stanford University is filing a patent on aptamer stabilization with RNAMake.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41565-019-0517-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to R.D.

Peer review information: *Nature Nanotechnology* thanks Peixuan Guo, Nils (G) Walter and other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Software availability. All the software and source code used in this work are freely available for non-commercial use. RNAMake software and documentation are at http://rnamake.stanford.edu. An RNAMake server to perform scaffolding and aptamer stabilization is available at http://rnamake.stanford.edu. EteRNAbot secondary structure design is available at https://software.eternagame.org.

Sequences and primers. All the sequences and primers used in this study are given in Supplementary sequences and Supplementary primers, respectively.

Building the motif library. To build a curated motif library of all the RNA structural components, we obtained the set of non-redundant RNA crystal structures managed by the Leontis and Zirbel groups[53] (version 1.45, http://rna.bgsu.edu/rna3dhub/nrlist/release/1.45). This set specifically removes redundant RNA structures that are identical to previously solved structures, such as ribosomes crystallized with different antibiotics. We processed each RNA structure to extract every motif with Dissecting the Spatial Structure of RNA (DSSR)[54] with the following command:

```
x3dna-dssr -i file.pdb -o file_dssr.out
```

We manually checked each extracted motif to confirm that it was the correct type, as DSSR sometimes classifies tertiary contacts as higher-order junctions and vice versa. For each motif collected from DSSR, we ran the X3DNA `find_pair` and `analyze` programs to determine the reference frame for the first and last base pair of each motif to allow for the alignment between motifs:

```
find_pair file.pdb 2> /dev/null stdout | analyze stdin
>& /dev/null
```

The naming convention for each motif involves the motif classification, the originating PDB accession code and a unique number to distinguish from other motifs of the same type, all separated by periods. For example, TWOWAY.1GID.2, is a two-way junction from the PDB 1GID and is the third two-way junction to be found in this structure. All the motifs retain their original residue numbering, chain identifications and relative position compared to their originating structure.

In addition to the motifs derived from the PDB, we also utilized the make-na web server (http://casegroup.rutgers.edu/nab.html) to generate idealized helices of between 2 and 22 base pairs in length[23]. All the motifs in these generated libraries are bundled with RNAMake and are grouped together by type (junctions, hairpins and so on) in sqlite3 databases in the RNAMake directory RNAMake/resources/motif_libraries_new.

Automatically building new RNA segments. RNAMake seeks a path for RNA helices and non-canonical motifs that can connect two base pairs separated by a target translation and rotation. We developed a depth-first search algorithm to discover such RNA paths. The algorithm is guided by a heuristic cost function f inspired by prior manual design efforts[2,25] and is composed of two terms:

$$f(\text{path}) = h(\text{path}) + g(\text{path}) \tag{1}$$

The first term, h(path), describes how close the last base pair in the path is to the target base pair; h(path) = 0 corresponds to a perfect overlap in translation and rotation. The functional form for h(path) depends on the spatial position of each base pair's centroid d and an orthonormal coordinate frame R that defines the rotational orientation of each base pair[46]:

$$h(\text{path}) = |\vec{d_1} - \vec{d_2}| + W\left(|\vec{d_1} - \vec{d_2}|\right) \sum_{i}^{3} \sum_{j}^{3} \text{abs}\left(R_{1ij} - R_{2ij}\right) \tag{2}$$

Here $W(d)$ is:

$$W(d) = \begin{cases} 0, & \text{if } d > 150 \\ \log \frac{150}{d}, & \text{if } 1.5 < d < 150 \\ 2, & \text{if } 1.5 > d \end{cases} \tag{3}$$

where d is measured in ångströms. The weight $W(d)$ reduces the importance of the current base pair and the target base pair with a similar alignment when they are spatially far apart. This term conveys the intuition that aligning the two coordinate frames becomes important only as the path of the motif and helices approaches the target base pair. RNAMake readily allows for the exploration of alternative forms of the cost function terms (2) and (3), which include more standard rotationally invariant metrics to define rotation matrix differences[55] or base-pair-to-base-pair r.m.s.d. values based on quaternions[56], but these were not tested in the current study.

The second term in the cost function (1) is g(path), which parameterizes the properties of the non-canonical RNA motifs and helices that comprise the path at each stage of the calculation:

$$g(\text{path}) = \frac{S_{ss}(\text{path})}{2} + 2N_{\text{motifs}} \tag{4}$$

where S_{ss} is a secondary structure score for all the motifs and helices in the path. This S_{ss} term favours longer canonical helices as well as motifs with frequently recurring base pairs, as follows. All the base pairs found in the RNA motif are scored based on their relative occurrences in all the high-resolution crystal structures; all the unpaired residues receive a penalty and Watson–Crick base pairs receive an additional bonus score (Supplementary Table 9). Values were derived based on logarithms of the frequencies of these elements in the crystallographic database, that is the inverse Boltzmann approximation[57], so that that the frequency of the elements in RNAMake designs was similar to that seen in natural RNA tertiary structures. In addition to the secondary structure score, N_{motifs} penalizes the total number of motifs in the path, here taken as the number of non-canonical motifs plus the number of helices (independent of the helix length).

The search adds motifs and helices to the path in a depth-first manner, and as the total cost function f(path) decreases, the back-tracking of f(path) increases. Any solutions with h(path) less than five, that is, overlap at approximately nucleotide resolution between the path's last base pair and the target base pair, are accepted into a list of final designs. The balance between g(path) and h(path) allows RNAMake to reduce the number of motif combinations considered and find most solutions in a few seconds. For each solution, we then used EteRNAbot, a secondary structure optimization algorithm that has undergone extensive empirical tests[24], to fill in the helix sequences.

Proteins that are included in the coordinates supplied to RNAMake are represented as steric beads centred at the Cα atom of each amino acid. This representation allows RNAMake to avoid steric clashes with proteins, particularly for the ribosome tethering problems.

Design, synthesis and experimental testing of miniTTR constructs. RNAMake designs of miniTTR constructs, in vitro synthesis and experimental testing are given in the Supplementary methods.

Design, construction and experimental testing of ribosome tether constructs. RNAMake designs of ribosome tether constructs, cloning and replacing wild-type ribosome and experimental are given in the Supplementary methods.

Design, synthesis and experimental testing of aptamer-stabilizing constructs. RNAMake designs of ATP- and Spinach-stabilized constructs and experimental testing by RNA structure probing, fluorescence measurements, Spinach-TTR stability assay in *Xenopus* egg extract and in vivo Spinach aptamer testing are given in the Supplementary methods.

Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request. Furthermore, all of our chemical mapping data are available on https://rmdb.stanford.edu, and a detailed table of the accession identifications is given in the Supplementary Information.

References

53. Petrov, A. I., Zirbel, C. L. & Leontis, N. B. Automated classification of RNA 3D motifs and the RNA 3D Motif Atlas. *RNA* **19**, 1327–1340 (2013).
54. Lu, X.-J., Bussemaker, H. J. & Olson, W. K. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res.* **43**, e142 (2015).
55. Huynh, D. Q. Metrics for 3D rotations: comparison and analysis. *J. Math. Imaging Vis.* **35**, 155–164 (2009).
56. Karney, C. F. F. Quaternions in molecular modeling. *J. Mol. Graph Model* **25**, 595–604 (2007).
57. Finkelstein, A. V., Badretdinov, A. Ya & Gutin, A. M. Why do protein architectures have Boltzmann-like statistics? *Proteins* **23**, 142–150 (1995).

nature nanotechnology

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41565-019-0517-8

In the format provided by the authors and unedited.

Computational design of three-dimensional RNA structure and function

Joseph D. Yesselman[1], Daniel Eiler[2], Erik D. Carlson[3,4,5], Michael R. Gotrik[1], Anne E. d'Aquino[3,5,6], Alexandra N. Ooms[7], Wipapat Kladwang[1], Paul D. Carlson[8], Xuesong Shi[1], David A. Costantino[2], Daniel Herschlag[1,9,10], Julius B. Lucks[3,4,5,6], Michael C. Jewett[3,4,5,6], Jeffrey S. Kieft[2] and Rhiju Das[1,11]*

[1]Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA. [2]Department of Biochemistry and Molecular Genetics, University of Colorado Denver School of Medicine, Aurora, CO, USA. [3]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA. [4]Chemistry of Life Processes Institute, Northwestern University, Evanston, IL, USA. [5]Center for Synthetic Biology, Northwestern University, Evanston, IL, USA. [6]Interdisciplinary Biological Sciences Graduate Program, Northwestern University, Evanston, IL, USA. [7]Department of Cancer Genetics & Genomics, Stanford University School of Medicine, Stanford, CA, USA. [8]Robert F. Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY, USA. [9]Department of Chemistry, Stanford University School of Medicine, Stanford, CA, USA. [10]Stanford ChEM-H (Chemistry, Engineering, and Medicine for Human Health), Stanford University, Stanford, CA, USA. [11]Department of Physics, Stanford University, Stanford, CA, USA. *e-mail: rhiju@stanford.edu

Supplemental Information:

Title: Computational Design of Three-dimensional RNA Structure and Function

Authors: Joseph D. Yesselman, Daniel Eiler, Erik D. Carlson, Michael R. Gotrik, Anne E. d'Aquino, Alexandra N. Ooms, Wipapat Kladwang, Xuesong Shi, David A. Costantino, Paul Carlson, Julius B. Lucks, Daniel Herschlag, Michael C. Jewett, Jeffrey S. Kieft, Rhiju Das*

*Correspondence should be addressed to R.D. (rhiju@stanford.edu)

Table of Contents:

Abbreviation Legend:

Abbreviation	Meaning
n.d.	not determined
n.a.	not applicable

Supplemental Results

Detailed analysis of chemical mapping of RNAMake miniTTR designs

To probe the structures of the miniTTR designs generated by RNAMake, we performed quantitative chemical mapping with selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) and dimethyl sulfate (DMS) (Supplementary Figure 1b) (1). For all 16 designs, we compared the SHAPE and DMS reactivity of each miniTTR RNA to its respective secondary structure. Of the 1386 nucleotides in the sixteen miniTTR constructs, 1367 (98.7%) were either reactive at target unpaired regions or protected at target helical residues, supporting the predicted secondary structures. All 19 outliers occurred at helix edges (*i.e.,* flanking base pairs of motifs). These data supported the formation of the expected secondary structures for all miniTTR designs (Supplementary Figure 1, Supplementary Table 3).

To evaluate the formation of tertiary structure, we investigated the change in DMS reactivity of both tetraloop and tetraloop-receptor adenines as a function of Mg^{2+} concentration. Previous studies have demonstrated that TTR formation in the P4-P6 domain is strongly stabilized by Mg^{2+} (2–4). As a control for the unfolded state, we measured the DMS reactivities of the tetraloop and tetraloop-receptor adenines of the TTR of the P4-P6 domain without Mg^{2+} (A248, A151, A152, and A153, see Figure 2a). The adenines exhibited reactivities of 1.27, 0.72, 0.70, and 0.90, respectively. Here and below, the values are normalized to the reactivity of the reference hairpin loops that flank each design (1). Upon the addition of 10 mM Mg^{2+}, the adenines involved in the TTR became protected from DMS modification in the P4-P6 control (Figure 2a). As with this folding control, for 12 of the 16 designs (miniTTRs 1, 2, 5-7, 9-12 and 14-16), we observed a more than two-fold decrease in the reactivity of the TTR adenine residues

(Figure 2a, Supplementary Table 3). These results were consistent with Mg^{2+}-dependent TTR formation. The remaining designs (miniTTRs 3, 4, 8 and 13) did not demonstrate significant changes in DMS reactivity upon addition of 10 mM Mg^{2+}, indicating that the TTR interaction did not form.

Native gel electrophoresis of miniTTR designs

As an independent test of miniTTR folding, we replaced each RNA's GAAA tetraloop with a UUCG tetraloop, which does not form the sequence-specific TTR tertiary contact (5) and is predicted to reduce the RNA's mobility in non-denaturing polyacrylamide gel electrophoresis, as observed for the P4-P6 domain (6). Of the 16 miniTTR constructs tested, 12 designs displayed mobility shifts consistent with the formation of the TTR tertiary contact (Figure 2b, Supplementary Figure 1c, Supplementary Table 3). Constructs 4 and 15 exhibited mobility shifts that were inconsistent with our chemical mapping results. The UUCG mutant of miniTTR design 4 displayed a mobility shift, but it did not demonstrate a full two-fold decrease in TTR DMS reactivity, suggesting partial folding. Compared to its UUCG mutant, miniTTR design 15 in the wild-type form (GAAA tetraloop) exhibited a wide, slow-mobility band. In all other cases, the electrophoretic mobility measurements were concordant with our quantitative SHAPE and DMS chemical mapping data, supporting the formation of the TTR and a compact tertiary fold.

Quantitative analysis of SAXS data agreement with miniTTR structural models

The observed scattering profiles of miniTTR 2 and miniTTR 6 agreed well with the profiles predicted from their corresponding RNAMake models, with χ^2/d.o.f. values of 13 and 27, respectively (Figures 2e and 2f). These values are near the values of 2–8 obtained from comparisons between predictions of RNA crystallographic models and

scattering data (7–9), suggesting similar overall folds with some local differences, an expectation confirmed for miniTTR 6 by X-ray crystallography (see main text).

Detailed description of in vivo tests of RNAMake ribosome tether constructs

To test RNAMake's accuracy in designing functional RNAs, each RM-Tether design (Supplementary Figure 3, Supplementary Figure 4a-c for secondary structure diagrams) was cloned into the pRibo-T plasmid (10) and used to replace the wild-type ribosomal rRNA plasmid in the SQ171fg strain (10). After 48 hours at 37 °C, no colonies were visible on the plates. Fresh plates were replica plated and incubated for a further 72 hours at 37°C, after which colonies appeared on the plate for RM-Tether design 4. Eight colonies were picked and checked for the loss of the wild-type rRNA plasmid. Growth curves were generated in liquid culture at 37°C (Supplementary Figure 4d).

These constructs exhibited doubling times of 1.5 ± 0.8 days and a maximum OD_{600} of 0.7 ± 0.3 (Supplementary Figure 4d). While slower than wild-type *E. coli*, the first successful version of Ribo-T was also slower growing before mutational optimization (10). After five days of growth, total RNA was extracted and analyzed by gel electrophoresis. Seven of the eight clones showed a clean Ribo-T-sized band and no detectable wild-type 23S and 16S rRNA bands (Supplementary Figure 4e), indicating the formation of stable tethered ribosomes, as previously demonstrated for Ribo-T (10).

Mg^{2+} titration of ATP-TTR provides independent evidence for coupling of designed tertiary structure and ATP binding

The strongest evidence for successful stabilization of ATP aptamers in the ATP-TTR designs came from measurements showing improved binding affinities to ATP,

described in the main text. As an independent test of whether the RNAMake-designed tertiary folding coupled to ATP binding, we examined the Mg^{2+}-dependent formation of the tertiary contact TTR with and without the nucleotide ligand. If the TTR scaffold positively contributes to ATP binding, then this construct should exhibit less dependence on magnesium when bound to ATP, relative to the same construct in the absence of ATP, due to the cooperative effects of binding. To test this prediction, we performed a 32-point Mg^{2+} titration (0.05 to 10 mM) on ATP-TTR 3 with and without 50 μM adenosine 5'-monophosphate (AMP). We switched the ligand from ATP to AMP as the two ligands have indistinguishable affinities to the aptamer and we wished to avoid Mg^{2+} chelation effects that occur with ATP. We observed that AMP did indeed couple into folding of the TTR. The Mg^{2+} -dependent reactivity midpoint ($[Mg^{2+}]_{1/2}$) for the TTR was 70 μM in the presence of 50 μM AMP, a 50% improvement relative to measurements without AMP, which gave $[Mg^{2+}]_{1/2}$ = 110 μM; see Supplementary Figure 7b. Taken with the ATP titration experiments, these results provide strong evidence that RNAMake-designed tertiary structure couples tertiary folding to ATP/AMP binding in ATP-TTR molecules.

Diversity of motifs in validated designs

At the beginning of this study we hypothesized that RNAMake might enable structural designs that were previously impossible due to the limited dictionary of non-canonical motifs previously amenable to human-optimized design. Supporting this idea, RNAMake used 54 noncanonical structural motifs across all the molecules tested in this study, and 43 of these appeared once or more in molecules that achieved their folding or functional design objectives (Table 1). Interestingly, each of the problems had a mostly distinct set of motifs used to solve their respective problem with little overlap between them (Supplementary Table 7). This observation suggests that it would be infeasible to solve

all of these problems with a significantly smaller set of motifs. Further supporting this view, we attempted to find algorithmic solutions to these problems with a smaller list of 9 'classic' structural motifs used previously in RNA design, including kink-turn motifs, right-angle bends, and stable three-way junctions (Supplementary Table 8). Achieving solutions to the miniTTR problem required us to manually expand the length limit for RNAMake, and RNAMake was not able to find any solutions to the ribosome design or either of the two aptamer stabilization problems. Taken together, these results indicate that use of a large dictionary of noncanonical structural motifs is critical for solving complex RNA 3D design problems.

Supplemental Discussion

Detailed discussion of motifs introduced into RNAMake designs

The miniTTR designs comprised 23 unique motifs (Supplementary Figure 1a), and we noted that they could be classified into distinct motif categories. Every miniTTR design contained at least one of the following motifs to create the near-180° turn necessitated by the design challenge: a large (>10 residue) bend such as a kink-turn (11), J5/5a from the P4-P6 domain (12), or an S-turn (13) (see, e.g., TWOWAY.1S72.20, TWOWAY.1GID.2 and TWOWAY.3BNQ.6; Supplementary Table 1). Second, each design included a near-helical motif; an example is a set of three consecutive non-canonical base pairs (see, e.g., TWOWAY.1S72.51 in Supplementary Table 1). Finally, some designs contained small motifs, such as a single adenine bulge or an A-A mismatch, used to make fine structural adjustments (see, e.g., TWOWAY.1S72.90 and TWOWAY.1S72.49 in Supplementary Table 1). Previous work on RNA design, which was based on manual modeling by RNA experts, did not test these latter two categories of motifs because they are difficult to model without automatic tools but appear to be

necessary in natural RNAs for generating asymmetric structures for which small refinements of helical twists are required.

One motif, an A-A mismatch (TWOWAY.1S72.49), appeared in all four miniTTR designs that did not form TTRs (miniTTR 3, miniTTR 4, miniTTR 8, miniTTR 13). Unlike other A-A mismatches that form a non-canonical base pair (14, 15), the adenines in the TWOWAY.1S72.49 structure are not paired and may reflect a high-energy structure that should not be used to design stable assemblies. The other motif with an interesting structural deviation was a triple mismatch (TWOWAY.1S72.62). Chemical mapping suggested that this motif retained its anticipated conformation at nucleotide resolution in miniTTR 6, but crystallographic analysis showed the motif's structure to be incorrect at atomic resolution (Figure 2h). TWOWAY.1S72.62's alternative conformation is likely due to the lack of the tertiary contacts provided by its parent ribosomal context but not the miniTTR design. Interestingly, this structural change was still compatible with the global folding of the miniTTR 6 design. This result suggests that residual uncertainties in RNAMake's motif library will not preclude the consistent design of asymmetric structures at nucleotide resolution.

As discussed in the main text, motifs used in the other problems studied here had little overlap with the miniTTR constructs (see, e.g., Supplementary Table 7). For example, of the fourteen motifs used in the RM-Tether designs, only three were also in the miniTTRs (TWOWAY.1NUV.3, TWOWAY.1S72.29, and TWOWAY.1S72.39). Nevertheless, the RM-Tether designs contained the same three distinct motif categories as the miniTTR designs. First, all designs included two large (>10 residues) bends (see, e.g., TWOWAY.1S72.29, TWOWAY.1S72.39. and TWOWAY.1S72.42 in Supplementary

Figure 3a). Furthermore, each design included at least one near-helical motif (see, e.g., TWOWAY.1NUV.3, TWOWAY.2VQE.46, TWOWAY.3LOA.0 in Supplementary Figure 3a) and one small motif (see, e.g., TWOWAY.1S72.100, TWOWAY.1J9H.3 in Supplementary Figure 3a).

Supplemental Methods

RNAMake design of miniTTR molecules

To generate miniTTR designs, we first extracted the coordinates from the X-ray crystal structure of a TTR from the P4-P6 domain of the *Tetrahymena* ribozyme (residues 146-157, 221-246, and 228-252 from PDB 1GID) (16). Second, we used RNAMake to build structural segments composed of two-way junctions and helices spanning the last base pair of the hairpin (146,157) to base pair (221,252) of the tetraloop-receptor, thus connecting the TTR into a single continuous strand (Figure 1d). Of 200,000 RNA segments generated, sixteen were selected based on two criteria: 1) the fewest number of motifs used in the solution (i.e. only three unique tertiary motifs); and 2) the tightest predicted atom-wise alignment of the miniTTR to its target spatial and rotational orientations. These computational designs ranged from 75 to 102 nucleotides in size (for full sequences, see Supplemental Document: Sequences.xlsx), significantly shorter than the 157 nucleotides of the natural P4-P6 domain RNA.

An example RNAMake command line to discover miniTTR's is the following:

```
design_rna -pdb p4p6_short.pdb -start_bp A146-A157 -end_bp
A221-A252 -designs 1000000
```

Here, `p4p6_short.pdb` is P4-P6 with residues 233 to 241 removed. This command runs the RNAMake design algorithm to build a new RNA segment between the base pair consisting of nucleotides 146 and 157 and the base pair consisting of nucleotides 221 and 252, also on chain A. The `design_rna` application automatically removes the nucleotides between these two ends, leaving only the two segments of the TTR remaining. This application also underlies the RNAMake server.

Preparation of RNA for SAXS and X-ray crystallography.

RNA used for crystallization was transcribed with T7 RNA polymerase from PCR-generated double-stranded DNA templates as described in (17). These templates were ordered from IDT as gBlocks with, in the 5´ to 3´ direction, a T7 promoter sequence, hammerhead ribozyme, miniTTR sequence of interest, and HDV ribozyme. RNA transcripts were ethanol precipitated overnight, washed with 70% ethanol and dissolved in water. RNA transcripts were purified from ribozymes and uncleaved products using PAGE purification. RNAs were eluted overnight at 4 °C, concentrated, and buffer-exchanged three times into water using Amicon Ultra concentrators (10K cutoff, Millipore). RNA was quantified and then stored at −20 °C until use.

Preparation of RNA for chemical mapping, native gels, and fluorescence experiments

DNA oligonucleotides were designed with Primerize (18), ordered from IDT (Integrated DNA Technologies) and used to generate double-stranded DNA templates using PCR assembly. For a full list of sequences and primers used in this study see Supplemental Material: Sequences.xlsx, Primers.xlsx. DNA template and RNA transcript preparation and quality checks were carried out as previously described (1, 19). For miniTTR and ATP-TTR chemical mapping, RNA transcribed with T7 RNA polymerase (New England

Biolabs) was purified using the RNA Clean & Concentrator 5 kit (Zymo Research). For native gels and fluorescence experiments, RNA was transcribed using a TranscriptAid T7 High Yield Transcription Kit (Thermo Scientific) and following the manufacturer's protocol. After DNase treatment for 30 minutes, the recovered RNA was purified using a Zymo RNA purification kit and resuspended in TE buffer (10mM Tris HCl, 1mM EDTA, pH 7.5).

Native gel shift assays

Native gel shift assays were conducted using a BioRad Criterion™ Cell gel cassette. Gels were cast using 11% acrylamide and 1X THEM buffer (pH 7.2) (10 mM $MgCl_2$) and polymerized by adding 10% ammonium persulfate (300 µL) and TEMED (30 µL) to 30 mL of gel mix. Following polymerization, the gel apparatus was set up in a 4 °C cold room and fully immersed in an ice bath until the gel and buffer apparatus were cooled to 4 °C. Then, RNA constructs were prepared for folding by incubating them in folding buffer consisting of 100 mM Na-HEPES (pH 7.5), 10 mM $MgCl_2$, and 0.5 mM spermidine. 100 µg of RNA was mixed with folding buffer in the absence of spermidine and incubated at 65 °C for 3 min. The solution was cooled at room temperature for 10 min, and then 0.5 mM spermidine was added. The RNAs were then vortexed for 10 seconds and centrifuged at 4,000 rpm for 10 min. Immediately following folding, loading dye was added to the RNA solutions, and the samples were directly loaded onto the cooled gel apparatus. The gel was run at 15 watts for 1.5 hours, and the temperature of the gel apparatus was strictly monitored to avoid overheating. After electrophoresis, the gel was removed from the cassette, carefully placed in a glass pan and incubated with 150 mL of Stains-All solution (Sigma-Aldrich) on an orbital shaker for 15 min. For de-staining, the Stains-All was removed, and the gel was rinsed with

deionized water and subsequently incubated in fresh deionized water on an orbital shaker for 20 min; the gel was then immediately imaged on a document scanner.

RNA structural probing and data analysis

Chemical mapping (DMS and SHAPE) was performed as previously described (1, 19). Briefly, modification reactions were performed in a 20 µL volume containing 1.2 pmol of RNA and 50 mM Na-HEPES (pH 8.0). For the ATP-TTRs, 2 µL of 50 mM ATP (final concentration: 5 mM) was added instead of 2 µL of H_2O to probe the bound conformation. For the Spinach-TTRs, 2 µL of 33 µM DFHBI was added instead of 2 µL of H_2O to probe the bound conformation. Before $MgCl_2$ or the chemical modifier was added, the RNA was heated to 90 °C for 3 minutes, then left on the bench top to cool to room temperature and then folded for 20 min in 10 mM $MgCl_2$ and 50 mM Na-HEPES (pH 8.0). To chemically modify the RNA, either 5 µL of DMS (1% v/v in 10% ethanol) or 1M7 (5 mg/mL in anhydrous DMSO) were added to each reaction to a total volume of 25 µL. After 5 minutes of incubation at room temperature, the reactions were quenched with 0.5 M Na-MES (pH 6.0). After quenching, poly(dT) magnetic beads (Ambion) and FAM-labeled Tail2-A20 primers were added for reverse transcription. Samples were separated and purified using magnetic stands, washed twice with 70% ethanol, and air-dried. The beads were resuspended in ddH_2O and reverse transcription mix, then incubated at 48 °C for 30 min. RNA was degraded by adding 1 volume of 0.4 M NaOH and incubating at 90 °C for 3 minutes; the sample was then cooled and neutralized with an additional volume of acid quench (prepared as 2 volumes of 5 M NaCl, 2 volumes of 2 M HCl, and 3 volumes of 3 M sodium acetate, pH 5.2). Fluorescently labeled cDNA was recovered by magnetic bead separation, rinsed twice with 70% ethanol and air-dried. The beads were resuspended in Hi-Di formamide containing ROX-350 ladder

(Applied Biosystems), then loaded on a capillary electrophoresis sequencer (ABI3130, Applied Biosystems).

The HiTRACE 2.0 package was used to analyze the CE data (20). Electrophoretic traces were aligned and baseline subtracted using linear and nonlinear alignment routines as previously described (20). Reactivities were determined by fitting these traces to sums of Gaussian peaks, followed by background subtraction, signal attenuation correction, and normalization to flanking reference hairpins (1).

Mg^{2+} titrations were carried out at 8 $MgCl_2$ concentrations (0, 0.1, 0.2, 0.5, 1.0, 2.0, 5.0 and 10 mM) for the miniTTRs and the P4-P6 domain, and at 32 $MgCl_2$ concentrations (0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 3.5, 4, 5, 7.5, 10 mM) for the ATP-TTR experiments. To estimate Mg^{2+} titration midpoints, the relative protection values (f_j^i) for each residue *j* in the TTR at each Mg^{2+} concentration *i* were calculated. The quantitative DMS reactivity of the folded and unfolded state of each TTR residue was taken from P4-P6.

$$f_j^i = \frac{A_j^i - A_j^{unfold}}{A_j^{fold} - A_j^{unfold}}$$

These values were then fit to the Hill equation:

$$f_j^i = \frac{(\lceil Mg^{2+} \rceil_i / K)^n}{1 + (\lceil Mg^{2+} \rceil_i / K)^n}$$

For each data set, global Hill fits were generated using scipy's curve_fit function, and errors were estimated through bootstrapping.

ATP 32 point titrations where performed at ATP concentrations of 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 6.5, 7.5, 10, 12.5, 15, 17.5, 20, 25, 50, 75, 100, 150, 200, 250, 300, 500, 750, 1000, 2500, 5000 μM. Dissociation constants were fit in MATLAB using LIFFT (21) using the following commands:

```
gp = [1:32];
N = size(area_peak, 1);
whichres = [33 34];
plotres = whichres;
conc = [0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 6.5, 7.5,
10, 12.5, 15, 17.5, 20, 25, 50, 75, 100, 150, 200, 250, 300, 500,
750, 1000, 2500, 5000];
lifft( area_peak( :, gp ), conc( gp ), [N: -1 : 1], K, n,
whichres, 'hill', [], plotres, 1, [], 0.01 );
```

Here `gp` gives which columns to include in the fit (here, there are 32 conditions); `N` is the size of the data matrix; `whichres` gives the residues used in the fit; and `conc` is the ATP concentration in μM at each column.

All chemical mapping data are available to download from rmdb.stanford.edu at the following accession IDs:

Accession ID	Data
MTTR1_DMS_0001	miniTTR 1 DMS/SHAPE
MTTR2_DMS_0001	miniTTR 2 DMS/SHAPE
MTTR3_DMS_0001	miniTTR 3 DMS/SHAPE
MTTR4_DMS_0001	miniTTR 4 DMS/SHAPE
MTTR5_DMS_0001	miniTTR 5 DMS/SHAPE
MTTR6_DMS_0001	miniTTR 6 DMS/SHAPE
MTTR7_DMS_0001	miniTTR 7 DMS/SHAPE
MTTR8_DMS_0001	miniTTR 8 DMS/SHAPE
MTTR9_DMS_0001	miniTTR 9 DMS/SHAPE
MTTR10_DMS_0001	miniTTR 10 DMS/SHAPE
MTTR11_DMS_0001	miniTTR 11 DMS/SHAPE
MTTR12_DMS_0001	miniTTR 12 DMS/SHAPE
MTTR13_DMS_0001	miniTTR 13 DMS/SHAPE
MTTR14_DMS_0001	miniTTR 14 DMS/SHAPE
MTTR15_DMS_0001	miniTTR 15 DMS/SHAPE
MTTR16_DMS_0001	miniTTR 16 DMS/SHAPE
MTTR1_MGTI_0001	miniTTR 1 Mg^{2+} Titration
MTTR2_MGTI_0001	miniTTR 2 Mg^{2+} Titration

MTTR3_MGTI_0001	miniTTR 3 Mg^{2+} Titration
MTTR4_MGTI_0001	miniTTR 4 Mg^{2+} Titration
MTTR5_MGTI_0001	miniTTR 5 Mg^{2+} Titration
MTTR6_MGTI_0001	miniTTR 6 Mg^{2+} Titration
MTTR7_MGTI_0001	miniTTR 7 Mg^{2+} Titration
MTTR8_MGTI_0001	miniTTR 8 Mg^{2+} Titration
MTTR9_MGTI_0001	miniTTR 9 Mg^{2+} Titration
MTTR10_MGTI_0001	miniTTR 10 Mg^{2+} Titration
MTTR11_MGTI_0001	miniTTR 11 Mg^{2+} Titration
MTTR12_MGTI_0001	miniTTR 12 Mg^{2+} Titration
MTTR13_MGTI_0001	miniTTR 13 Mg^{2+} Titration
MTTR14_MGTI_0001	miniTTR 14 Mg^{2+} Titration
MTTR15_MGTI_0001	miniTTR 15 Mg^{2+} Titration
P4P61_MGTI_0001	P4-P6 Mg^{2+} Titration
ATPCON_DMS_0001	ATP aptamer DMS
ATTR03_DMS_0001	ATP-TTR 3 DMS
ATTR04_DMS_0001	ATP-TTR 4 DMS
ATTR05_DMS_0001	ATP-TTR 5 DMS
ATTR06_DMS_0001	ATP-TTR 6 DMS
ATTR07_DMS_0001	ATP-TTR 7 DMS
ATTR09_DMS_0001	ATP-TTR 9 DMS
ATPCON_TITR_0001	ATP aptamer ATP Titration
ATTR03_TITR_0001	ATP-TTR 3 ATP Titration
ATTR04_TITR_0001	ATP-TTR 4 ATP Titration
ATTR05_TITR_0001	ATP-TTR 5 ATP Titration
AMGNA3_DMS_0001	ATP-TTR 3 Mg^{2+}Titration no ATP
AMGA3_DMS_0001	ATP-TTR 3 Mg^{2+}Titration 50 μM ATP
SCONTR_1M7_0001	Spinach aptamer SHAPE
STTR01_1M7_0001	Spinach-TTR 6 SHAPE
STTR02_1M7_0001	Spinach-TTR 10 SHAPE
STTR03_1M7_0001	Spinach-TTR 11 SHAPE
STTR04_1M7_0001	Spinach-TTR 13 SHAPE
STTR06_1M7_0001	Spinach-TTR 6 SHAPE
STTR07_1M7_0001	Spinach-TTR 10 SHAPE
STTR08_1M7_0001	Spinach-TTR 11 SHAPE
STTR09_1M7_0001	Spinach-TTR 13 SHAPE
STTR10_1M7_0001	Spinach-TTR 6 SHAPE
STTR11_1M7_0001	Spinach-TTR 10 SHAPE
STTR12_1M7_0001	Spinach-TTR 11 SHAPE
STTR13_1M7_0001	Spinach-TTR 13 SHAPE
23SRRNA_H101_0001	23S H101, iSAT SHAPE-Seq
16SRRNA_H44_0001	16S h44, iSAT SHAPE-Seq
RIBOT_TE1_0001	Ribo-T v1.0 strand #1, iSAT SHAPE-Seq
RIBOT_TE2_0001	Ribo-T v1.0 strand #2, iSAT SHAPE-Seq
RNAMKT4_TE1_0001	RM-Tether 4, strand #1, iSAT SHAPE-Seq
RNAMKT4_TE2_0001	RM-Tether 4, strand #2, iSAT SHAPE-Seq

SAXS measurements, analysis and modeling

RNA transcripts were purified using an Agilent 1260 Infinity HPLC using a gradient of 13–23% buffer B (100 mM TEAA (pH 7.0) and 50% acetonitrile) in buffer A (100 mM TEAA (pH 7.0) and 0.2% sodium azide) over a Varian PLRP-S 1000 Å 8 µm 150 × 7.5 mm column. Fractions containing miniTTR constructs were pooled, concentrated, and buffer-exchanged three times into water using Amicon Ultra concentrators (Millipore). The RNA was then quantified and stored at −20 °C until use.

Small-angle X-ray scattering measurements were carried out at Bio-SAXS beamline BL4-2 at the Stanford Synchrotron Radiation Lightsource (SSRL). Scattering data were collected with a 1.7 m sample-to-detector distance and a beam energy of 11 keV (wavelength of 1.127 Å). RNA samples were first buffer-exchanged into running measurement buffer solution consisting of 70 mM Tris-HCl (pH 7.4), 160 mM NaCl, 10 mM $MgCl_2$ and 5 mM DTT using Amicon Ultra centrifugal filters (10K cutoff, Millipore). Approximately 50 µL of the buffer-exchanged RNA (5 mg/mL) was then loaded onto a 24 mL Superdex 200 size-exclusion column (GE Healthcare) that had been pre-equilibrated with the running measurement buffer solution, then run at a flow rate of 0.05 mL/min using an AKTA Ettan FPLC (GE Healthcare). The elution was directed to the sample flow path for immediate SAXS data collection every 5 seconds, with a 2 second exposure time. The SAXS images were processed using the SASTOOL program. The first 100 images were used to create the buffer scattering profiles. The segment of the main elution peak with constant, scale-independent scattering profiles was used to calculate the sample scattering profiles.

The SAXS profiles of the miniTTR 2 and 6 RNAMake models were predicted and compared with the experimental profiles (Figure 2e-f) using FoXS (22). 3D bead models

of miniTTR 2 and 6 were generated using DAMMIF and DAMMIN (23) and overlaid with their corresponding RNAMake models in PyMOL.

MiniTTR crystallization

Purified miniTTR 6 RNA diluted in buffer A (30 mM HEPES (pH 7.5), 20 mM $MgCl_2$, and 100 mM KCl) was incubated at 65 °C for 2 min, centrifuged at 13,000 rpm for 2 min, and snap-cooled on ice for approximately 5 min before moving to 25 °C to set up crystallization trays. Within 2-4 weeks, miniTTR 6 crystallized at 25 °C as plates or clusters of plates via sitting-drop vapor diffusion by mixing 2 µL of miniTTR 6 at a concentration of 100 µM with 3 µL of crystallization solution containing 40 mM sodium cacodylate (pH 5.5), 20 mM $MgCl_2$, 2 mM cobalt hexammine, and 40% 2-methyl-2,4-pentanediol (MPD). Crystals of miniTTR 6 grew to maximum dimensions of 700 x 700 x 20 µm and were stabilized and cryogenically protected by increasing the MPD to a final concentration of 44%. Crystals were flash-frozen by plunging into liquid nitrogen.

Diffraction data were collected at 100 K using synchrotron X-ray radiation at beamline 4.2.2 of the Advanced Light Source, Lawrence Berkeley National Laboratory (Berkeley, CA). The data were processed and scaled using X-ray Detector Software (XDS) (24). The scaled data were handled using Collaborative Computational Project programs (17).

Structure determination and refinement

The initial structural determination of the miniTTR 6 in the C2 space group was carried out from molecular replacement (MR) in Phaser (CCP4) searching for one copy of a 31-nucleotide model of only the tetraloop and receptor with the identical sequence (25). The rotational and translational Z-scores were somewhat low, 4.6 and 5.9 respectively, but

the maps were of sufficient quality to enable the iterative building of all the residues into the $2F_o$-F_c and F_o-F_c maps. Composite omit maps in PHENIX were used to help confirm the model and reduce model bias from the initial MR solution. The models were built using COOT (25) and refined using REFMAC5 and PHENIX (17). The final model was refined in REFMAC5 and ERRASER (17, 26), and the overall R_{work} and R_{free} were refined to 22.9% and 27.4%, respectively. The structure derived from the miniTTR was refined to 2.55 Å against a data set scaled to an overall I/σ of 1.0 at the highest resolution shell with 98.5% completeness. Final crystallographic statistics can be found in Supplementary Table 10. The crystal structures of miniTTR 6 have been deposited in the PDB, ID 6DVK. All structural figures were prepared using PyMOL (http://www.pymol.org/).

Design of RNAMake generated tethers for ribosome RNAs.

For ribosome tether designs, we used PDB coordinates 3R8T and 4GD2 for the 50S and 30S ribosomal subunit structures respectively. From the 50S coordinates, we removed residues A2854-A2863 and, from the 30S, we removed residues A1445-A1457. Due to the large size we had to generate an RNAMake motif graph file to parse the structures into a format easier for RNAMake to read on command line. The script is available as part of RNAMake (rnamake/bin/ribosome_tethering.py). This generated file was called 'start_pose.mg'. To generate the designs, we then ran

```
design_rna -mg start_pose.mg -designs 10000
```

Over 100 designs were generated and 9 were selected to maximize the number of unique motifs utilized.

Modeling existing ribosome tethers

To generate models of previous ribosome tethers, we used Rosetta fragment assembly of RNA with Full Atom Refinement (27). For Umbilical 3 (U3), removed residues A1448-A1452 and A1455-A1456 (PDB: 3R8T), and A2853-A2855 and A2858-A2864 (PDB: 4GD2). We combined these residues and renumbered them into a starting PDB-formatted file U3_template.pdb. We also generated a fasta file U3.fasta with the residues in the start pdb and the tether sequence:

```
> tether strand 1
ccuucgaaaaaagcgaugcg

> tether strand 2
cgcaaaaaaagga
```

and ran the following command using FARFAR to generate the U3 model:

```
rna_denovo -s U3_template.pdb -fasta U3.fasta -minimize_rna
true
```

Similarly for the stapled design construct we removed residues A1447-A1449, A1454-A1457 from the 30S ribosome pdb coordinates in 3R8T and A2861-A2865, A2853-A2856 from the 50S ribosome pdb coordinates in 4GD2 into a common start pdb we denoted stapled_template.pdb. We also generated a fasta file stapled.fasta with the residues in the start pdb and the tether sequence:

```
> tether strand 1
accggggucaacagccguucagugcgu

> tether strand 2
cgcacugacggacaugguccuggag
```

And ran the corresponding FARFAR command:

```
rna_denovo -s stapled_template.pdb -fasta stapled.fasta
-minimize_rna true
```

Lastly for the Ribo-T construct we removed residues A1454-A1455, A1449-A1453 from 3R8T and residues A2855-A2857, A2858-A2862 from 4GD2 and combined them into a common start pdb we denoted ribo_t_template.pdb. We also generated a fasta file ribo_t.fasta with the residues in the start pdb and the tether sequence:

```
> tether strand 1
cuucgaaaaaaaacgaug

> tether strand 2
cagaaaaaaaaagg
```

And ran the corresponding FARFAR command:

```
rna_denovo -s ribo_t_template.pdb -fasta ribo_t.fasta
-minimize_rna true
```

Design and cloning of novel tethers for a ribosome with tethered subunits.

The designed tethers were cloned into plasmid pRibo-T-A2058G (10). The backbone was generated for each design using forward (f) and reverse (r) primer pairs (noted with "bb") in Primers.xlsx in separate PCR reactions using plasmid pRibo-T as a template (10), Phusion polymerase (NEB), and 3% DMSO. PCR cycling was as follows: 98 °C for 3 min; 25 cycles of 98 °C for 30 sec, 55 °C for 30 sec, 72 °C for 2 min; and 72 °C for 10 min. Circularly permuted 23S ribosomal RNA (rRNA) was generated with forward and reverse primer pairs (noted with "23S" in Supplemental Document: Primers.xlsx), the pRibo-T template, and the same PCR conditions as described above. Each PCR reaction was purified by gel extraction from a 0.7% agarose gel with an E.Z.N.A. gel extraction kit (Omega). Each purified backbone (50 ng) was assembled with the respective 23S insert in 3-fold molar excess using Gibson assembly (28). Assembly reactions were transformed into POP2136 cells, and the cells were grown at 30 °C overnight. Colonies were picked and plasmids were isolated using an E.Z.N.A. miniprep kit (Omega) and confirmed with full plasmid sequencing by ACGT, Inc.

Replacement of wild type ribosomes with RNAMake tethered ribosomes

Each purified plasmid (100 ng) was separately transformed into electrocompetent SQ171fg cells containing pCSacB (10). Cells were recovered in 1 mL of SOC media at 37 °C with shaking for 1 hour. Fresh SOC (1.85 mL) supplemented with 50 µg/mL carbenicillin and 0.25% sucrose was inoculated with 250 µL of recovered cells and incubated overnight at 37 °C with shaking. Cultures (10% and 90%) were plated on LB agar plates supplemented with 50 µg/mL carbenicillin, 5% sucrose and 1 mg/mL erythromycin and incubated at 37 °C.

After 48 hours with no visible colonies, the plates were replica plated onto fresh LB agar plates supplemented with 50 µg/mL carbenicillin, 5% sucrose and 1 mg/mL erythromycin and incubated at 37 °C. After 72 additional hours, colonies appeared on the plate containing RM-Tether design 4. Eight colonies were streaked onto LB agar supplemented with 50 µg/mL carbenicillin and 1 mg/mL erythromycin and LB agar supplemented with 30 µg/mL kanamycin (to confirm loss of the pCSacB plasmid) and were also used to inoculate 5 mL of LB supplemented with 50 µg/mL carbenicillin and 1 mg/mL erythromycin. Plates were incubated at 37 °C, and cultures were incubated at 37 °C with shaking. The OD_{600} of the cultures was tracked to generate growth curves (Biochrom Libra S4 spectrophotometer). After 5 days at 37 °C, total RNA was extracted using an RNA extraction kit from Qiagen. Total RNA was analyzed by gel electrophoresis on a 1% agarose gel with GelRed. Total plasmid was extracted from saturated 5 mL cultures with an E.Z.N.A. miniprep kit (Omega) and sequenced to confirm the correct RM-Tether design 4 sequence.

In vitro construction, testing, and characterization of ribosomes

For *in* vitro characterization of ribosomes, all constructs (wild type, Ribo-T v1.0, and RM-Tether 4) were cloned to be under control of a T7 promoter. The T7 promoter was introduced into primers, and amplified using the wild type, Ribo-T v1.0, and RM-Tether 4 plasmids as templates for PCR amplification. PCR products were blunt end ligated, transformed into DH5α *E. coli* cells using electroporation, and plated onto LB-agar/ampicillin plates at 37°C. Plasmid was recovered from resulting clones and sequence confirmed.

In vitro ribosome synthesis, assembly, and translation (iSAT) reactions were set-up as previously described (29). Briefly, eight 15 µL reactions were prepared and incubated for 2 hours at 37 °C, then pooled together.

Sucrose gradients were prepared from buffer C (10 mM Tris-OAc (pH = 7.5 at 4 °C), 60 mM NH_4Cl, 7.5 mM $Mg(OAc)_2$, 0.5 mM EDTA, 2 mM DTT) with 10 and 40% sucrose in SW41 polycarbonate tubes using a Biocomp Gradient Master. Gradients were placed in SW41 buckets and chilled to 4 °C. 120 µL of pooled iSAT reactions were loaded onto the gradients. The gradients were ultra-centrifuged at 22,500 rpm for 17 hours at 4 °C, using an Optima L-80 XP ultracentrifuge (Beckman-Coulter) at medium acceleration and braking (setting of 5 for each). Gradients were analyzed with a BR-188 density gradient fractionation system (Brandel) by pushing 60% sucrose into the gradient at 0.75 mL/min (at normal speed). Traces of A254 readings versus elution volumes were obtained for each gradient. Gradient fractions were collected and analysed for rRNA content by gel electrophoresis in 1% agarose and imaged in a GelDoc Imager (Bio-Rad). Ribosome profile peaks were identified based on the rRNA content as representing 30S or 50S subunits, 70S ribosomes, or polysomes.

Fractions containing 70S ribosomes and polysomes were collected and pooled. These fractions were recovered as previously described (29), with pelleted iSAT ribosomes resuspended in iSAT buffer, aliquoted, and flash-frozen. These pelleted fractions were re-run on a 1% agarose gel and imaged in a GelDoc Imager to confirm tethering in monosome and polysome peaks.

in vitro SHAPE-seq chemical mapping on tethered ribosomes.

For SHAPE-seq*, in vitro* ribosome synthesis, assembly, and translation reactions were set-up as previously described (29, 30). Briefly, 15 µL iSAT reactions each possessing wild type, Ribo-T, or RM-40 were prepared in triplicate, incubated for 2 hours at 37 °C, then placed on ice. To perform SHAPE modification, samples were warmed to 37 °C for 5 minutes, and 7.5 µL of each sample was added to 0.83 µL of 65 mM 1-methyl-7-nitroisatoic anhydride (1M7) or 0.83 µL DMSO (control solvent). Reactions were incubated for 2 minutes, then all samples were Trizol extracted, ethanol precipitated, washed twice with 70% ethanol, and resuspended in 10 µL water. Subsequent library preparation steps were performed as described previously (31) with one exception: 2 custom reverse transcription primers were used to simultaneously probe the regions containing T1 (5′ – GGTTAAGCCTCACGG – 3′) and T2 (5′ – CCCTACGGTTACCTTGTTACGAC – 3′). Following 2 x 75bp paired-end Illumina sequencing, SHAPE reactivities were calculated as described by Yu *et al.* (32), mapping both modification-induced stops and mutations. Raw reactivities were calculated using Spats v1.9.8 (https://github.com/LucksLab/spats), and were then linearly re-scaled to account for estimated differences in SHAPE probe concentration between replicates. Specifically, one replicate was first selected as the reference. Reactivities for the other

datasets were divided by the reference at each position, then the median value of this ratio was taken as the scale factor. Reactivities across each dataset were divided by their scale factor. The same experimental replicate was used to scale reactivities, and reactivities are presented as the average value over these re-scaled replicates.

Generation of RNAMake aptamer stabilization designs

Although our original RNAMake algorithm produced solutions for both the ATP and DFHBI stabilization, these simulations exhibited low motif diversity. To produce more solutions, we added more helix 'motifs'. Similar to prior approaches seeking to leverage flexibility expected of RNA motifs (33, 34), we expanded our helix conformation to an ensemble of motifs that were similar but not identical to a single helix conformation. Instead of having a single idealized helix for each base pair length, we produced 100 diversified helices for given base pair length. We prepared these models by generating ensembles for each base pair step from crystallographic models and generating random helices of a given length, we took conformations that were most divergent from the idealized structure (for more details, see reference (35)). This set of 100 helices increased the number of solutions for the aptamer stabilization design problems by allowing better closure of RNA motif paths.

Starting with PDB 1AM0 we removed residues A6-A18 and A33-A35 to achieve a minimal ATP aptamer flanked by single Watson-Crick base pairs. We moved these residues into a new PDB 'ATP_min.pdb'. We then ran `apt_stablization`, an executable in RNAMake to generate stabilized aptamers given a pdb of the aptamer.

```
apt_stablization -pdb ATP_min.pdb -designs 1000000
```

In all, 5210 designs were generated. As with previous construct designs, designs were selected that maximized motif usage and minimized the chain closure score or how close the optimized sequence is to the target base pair.

Starting with PDB 6B14 we removed residues R19-R31 and R49-R66 to achieve the minimal DFHBI binding aptamer (Spinach_min.pdb). We then ran `apt_stablization`, as follows.

```
apt_stablization –pdb Spinach_min.pdb –designs 1000000
```

In all, 697 designs were generated, and a subset were again chosen to maximize number of motifs tested and the chain closure score (how close the designed RNA sequence is to overlay with its target base pair).

Fluorescence characterization of Spinach-TTRs

A stock of DFHBI (Sigma) was prepared in PBSMKT (1X phosphate buffered saline, 5 mM $MgCl_2$, 100 mM KCl, 0.01% Tween-20, pH 7.2) and its absorbance measured using a UV spectrophotometer (NanoDrop, Thermo Scientific). The DFHBI concentration was calculated using an extinction coefficient of 30,100 cm^{-1}/M at 423 nm as previously reported (36). A DFHBI titration was performed in half area, flat-bottomed black 96-well plates (Corning) at a final RNA concentration of 200 nM with DFHBI concentration ranging from 10 μM to 10 nM prepared in a 1:2 dilution series. After mixing, the plates were covered with an adhesive film to prevent evaporation and temperature-cycled from room temperature to 4 ˚C twice over the course of 1 hour to allow aptamer-target

equilibration while minimizing magnesium-dependent self-cleavage. Measurements were acquired at room temperature and wells were excited at 462±10 nm and emission was measured at 504±15 nm using a Tecan M1000 plate reader. A fluorescence background was obtained at each DFHBI concentration in the absence of RNA and subtracted from the corresponding wells. The corrected signal for each aptamer at every DFHBI concentration was then least-squares fit using a custom MATLAB script using a 1:1 complexation model according to the following equation:

$$F = B_{max} * \frac{[T]}{[T] + K_d}$$

Here, [*T*] is the concentration of DFHBI, K_d is the dissociation constant of the given aptamer, and B_{max} is the maximum brightness obtained for the given concentration of aptamer.

Next, we prepared an RNA titration assay using identical measurement, equilibration, and buffer conditions, except with the amount of DFHBI constant at 400 nM and RNA concentrations ranging from 5 μM down to 5 nM prepared in a 1:2 dilution series. A background fluorescence was obtained at 400 nM DFHBI in the absence of RNA and subtracted from each well. The corrected signal was then least-squares fit using a custom MATLAB script using a 1:1 complexation model according to the following equation:

$$F = F_{max}\left(\frac{[A] * f + DT + K_d - \sqrt{([A] * f + DT + K_d)^2 - 4 * [A] * f * DT}}{2 * DT}\right)$$

Where [*A*] was the concentration of aptamer, *f* is the folding efficiency, *DT* is the DFHBI concentration (400 nM), K_d is the dissociation constant calculated for each sequence above, and F_{max} is the maximum fluorescence signal at dye-binding saturation. Quantum

yields were obtained through direct comparison of F_{max} with the literature value for Broccoli (QY = 0.72; Supplementary Table 6) (37).

Xenopus egg lysate stability assay for Spinach-TTRs

Each TTR Spinach aptamer was prepared in 60 µL PBSMKT containing 1.66 µM total RNA and 30 µL of this was added to 50 µL of 5 µM DFHBI in PBSMKT in two wells per aptamer. Next, 20 µL of PBSMKT was added to one well per aptamer to give a final concentration of 500 nM RNA and 2.5 µM DFHBI in order to provide a baseline fluorescence. Next, 20 µL of 100% frog egg lysate prepared 4 hours earlier (38) and stored at 4 ˚C, was added to each well and pipet mixed. (Higher lysate concentrations were too optically absorbent to allow fluorescence measurements). Fluorescence measurements were then obtained for every well every 1 minute for 30 minutes, then every 3 minutes for 1 hour, and after every 5 minutes for an additional hour. For evaluation of times to half-fluorescence, the fluorescence of each aptamer in wells containing lysate was normalized to the same aptamer's fluorescence in PBSMKT at every time point in order to account for photobleaching.

Spinach-TTR stability assay in 15% E. coli lysate

Each TTR Spinach aptamer was prepared in PBSMK (1X PBS pH 7.2, 5 mM $MgCl_2$, 100 mM KCl) containing 1 µM RNA and 2.5 µM DFHBI. The RNA/DFHBI mixture was equilibrated on ice for 30 minutes before aliquoting 50 µL into 4 wells per RNA species. As control reactions, 50 µL of PBSMK containing 2.5 µM DFHBI was added to one of these wells per RNA. Immediately prior to use, PBSMLK (1X PBS pH 7.2, 5 mM $MgCl_2$, 40% *E. coli* S30 lysate[47], 100 mM KCl) containing 2.5 µM DFHBI was prepared and 50 µL of this mixture was added to each well to give final concentrations of 500 nM RNA,

2.5 μM DFHBI, and 20% *E. coli* lysate. Immediately upon addition of PBSMLK, fluorescence intensities were obtained for every well and repeated every 30 s for 8 hours using a Tecan M1000 plate reader.

In vivo Spinach aptamer testing

To test the *in vivo* fluorescence of Spinach-TTR variants, designed sequences were cloned between a T7 promoter and T7 terminator in a plasmid harboring carbenicillin resistance and a ColE1 origin of replication (Supplementary Figure 13a). Plasmids were transformed into chemically competent *E. coli* strain BL21*(DE3) (*F- ompT hsdSB (rB- mB-) gal dcm me131 [DE3]*), plated on Difco LB+Agar plates containing 100 μg/mL carbenicillin, and grown overnight at 37 °C. A cellular autofluorescence control containing a blank plasmid (39) was also included. Individual colonies were grown overnight in LB containing 100 μg/mL carbenicillin, then diluted 1:50 into fresh LB. After 1 h, Isopropyl-β-D-thiogalactoside (IPTG) was added at a final concentration of 100 μM to induce expression of T7 RNA polymerase. After 4.5 h of additional shaking, cells were diluted 1:200 into 1x Phosphate Buffered Saline (PBS) containing 2 mg/mL kanamycin and 200 μM (Z)-4-(3,5-Difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one (DFHBI), then incubated at 37 °C for 5 minutes. A BD Accuri C6 Plus flow cytometer fitted with a high-throughput sampler was then used to measure fluorescence of at least 50,000 events for each sample. Measurements were taken for 4 biological replicates.

Flow cytometry data analysis was performed using FlowJo (v10.4.1). Cells were gated by FSC-A and SSC-A, and the same gate was used for all samples. The geometric mean fluorescence was calculated for each sample, then all fluorescence measurements were converted to Molecules of Equivalent Fluorescein (MEFL) using

CS&T RUO Beads (BD). The average fluorescence (MEFL) of cells expressing blank plasmid (pJBL002) in the presence of DFHBI was then subtracted from each measured fluorescence value.

RNAMake Name	Common Name	Designs Included	Motif Type	Translation (Å)	Angle (degrees)
TWOWAY.1D4R.2	A-C/C-A double mismatch	1	Near Helical	11.7	17.2
TWOWAY.1GID.1		1	Near Helical	13.6	33.1
TWOWAY.1GID.2	J5/5A from P4-P6	6	Large Turn	13.3	129.6
TWOWAY.1J9H.3	single U bulge	1	Small Motif	4.7	8.89
TWOWAY.1JBR.0	Sarcin-ricin	2	Near Helical	21.3	18.0
TWOWAY.1LNG.3		2	Near Helical	18.6	15.9
TWOWAY.1MWL.0	U-U mismatch	1	Small Motif	7.9	4.1
TWOWAY.1NUV.0		3	Near Helical	13.2	23.0
TWOWAY.1OOA.1		1	Large Turn	19.5	73.5
TWOWAY.1Q96.1	Sarcin-ricin	2	Near Helical	21.7	14.2
TWOWAY.1S72.100	AA bulge	4	Small Motif	7.6	14.0
TWOWAY.1S72.12	Sarcin-ricin	1	Near Helical	18.9	17.6
TWOWAY.1S72.17		2	Near Helical	14.8	17.7
TWOWAY.1S72.20	Kink-Turn	2	Large Turn	19.0	100.9
TWOWAY.1S72.23	REV element IL	3	Near Helical	13.9	37.1
TWOWAY.1S72.29	Kink-Turn	8	Large Turn	15.6	121.7
TWOWAY.1S72.3	Kink-Turn variant	3	Large Turn	15.3	108.3
TWOWAY.1S72.34		1	Large Turn	21.6	114.4
TWOWAY.1S72.38	Kink-Turn variant	3	Large Turn	16.7	137.2
TWOWAY.1S72.39		3	Large Turn	12.1	50.2
TWOWAY.1S72.42	Kink-Turn variant	2	Large Turn	13.4	132.7
TWOWAY.1S72.47		1	Small Motif	8.2	7.9
TWOWAY.1S72.49		4	Small Motif	9.4	77.0
TWOWAY.1S72.51	5S Loop E	3	Near Helical	14.9	19.1
TWOWAY.1S72.6		3	Large Turn	20.5	61.8
TWOWAY.1S72.62		1	Near Helical	5.0	3.7
TWOWAY.1S72.86		2	Small Motif	7.6	34.6
TWOWAY.1S72.90		3	Near Helical	6.6	40.6
TWOWAY.2FQN.2		1	Near Helical	11.4	17.0
TWOWAY.2HW8.0	Reverse Kink-Turn	3	Large Turn	11.8	107.6
TWOWAY.2PN4.2		2	Large Turn	13.1	103.9
TWOWAY.2VPL.0		2	Large Turn	11.9	112.0
TWOWAY.2VQE.10	Kink-Turn	2	Large Turn	10.5	120.0
TWOWAY.2VQE.13		1	Near Helical	12.6	1.4
TWOWAY.2VQE.15	Kink-Turn Variant	1	Large Turn	18.8	113.0
TWOWAY.2VQE.18		2	Near Helical	6.5	16.9
TWOWAY.2VQE.26		1	Near Helical	13.3	13.7
TWOWAY.2VQE.35		1	Near Helical	8.8	5.9
TWOWAY.2VQE.45	GGC bulge	3	Small Motif	9.73	9.4
TWOWAY.2VQE.50		1	Small Motif	7.2	27.8
TWOWAY.2VQE.6	Kink-Turn Variant	1	Large Turn	15.1	116.3
TWOWAY.2VQE.9		1	Near Helical	23.1	54.4
TWOWAY.2ZY6.0		1	Near Helical	11.2	9.3

TWOWAY.3BNQ.2		1	Near Helical	11.6	33.9
TWOWAY.3BNQ.6	Double S-turn motif	4	Near Helical	11.6	38.3
TWOWAY.3DVZ.0	Sarcin-ricin	1	Near Helical	18.8	12.2
TWOWAY.3GX5.0	Kink-Turn Variant	4	Large Turn	15.9	113.6
TWOWAY.3LOA.0		1	Near Helical	12.5	28.9
TWOWAY.3P59.1	HCV IRES domain IIa	4	Large Turn	13.3	98.8
TWOWAY.3RW6.0	Kink-Turn Variant	6	Large Turn	14.3	106.1
TWOWAY.3UMY.1	Kink-Turn Variant	6	Large Turn	15.4	113.8
TWOWAY.4BW0.0	Kink-Turn Variant	3	Large Turn	16.6	107.0
TWOWAY.4K27.3		1	Near Helical	11.3	18.4
TWOWAY.4OO8.0		2	Large Turn	18.8	58.2

Supplementary Table 1: All motifs used in this study.

Table of all motifs selected out of structural motif library by RNAMake to design 16 miniTTR molecules, 9 RM-Tethers, 10 ATP-TTRs and 16 Spinach-TTRs. Note: Additional motifs are also available in the library but did not appear in the molecules tested in this study.

Construct	Mg^{2+} Midpoint (mM)	Apparent Hill coefficient	ΔG (1 mM Mg^{2+})	ΔG (10 mM Mg^{2+})
P4-P6	0.43 +0.04/-0.04	2.55 +0.58/-0.54	-1.3	-4.8
miniTTR 1	1.12 +0.34/-0.24	1.08 +0.32/-0.23	0.1	-1.4
miniTTR 2	0.08 +0.01/-0.01	1.83 +0.12/-0.14	-2.7	-5.2
miniTTR 3	n.d.	n.d.	n.d.	n.d.
miniTTR 4	n.d.	n.d.	n.d.	n.d.
miniTTR 5	1.64 +0.32/-0.22	1.05 +0.19/-0.12	0.3	-1.1
miniTTR 6	0.74 +0.0/-0.02	>10.00[a]	-1.8[a]	-15.4[a]
miniTTR 7	3.31 +0.79/-0.55	1.13 +0.21/-0.20	0.8	-0.7
miniTTR 8	n.d.	n.d.	n.d.	n.d.
miniTTR 9	0.84 +0.11/-0.11	2.19 +0.77/-0.49	-0.2	-3.2
miniTTR 10	0.74 +0.08/-0.06	1.41 +0.16/-0.12	-0.3	-2.2
miniTTR 11	0.87 +0.13/-0.10	1.04 +0.11/-0.09	-0.1	-1.5
miniTTR 12	0.50 +0.05/-0.03	1.07 +0.07/-0.09	-0.4	-1.9
miniTTR 13	n.d.	n.d.	n.d.	n.d.
miniTTR 14	0.44 +0.02/-0.01	1.99 +0.16/-0.14	-1.0	-3.7
miniTTR 15	0.95 +0.14/-0.14	1.17 +0.19/-0.17	0.0	-1.6
miniTTR 16	0.24 +0.08/-0.02	1.22 +0.03/-0.04	-1.0	-2.7

[a] Mg^{2+} dependence for miniTTR 6 was sharper than curve with apparent Hill coefficient of 10.0; fits herein assume 10.0.

Supplementary Table 2: Folding stabilities of each miniTTR estimated by Mg^{2+}-induced folding.

Data are derived from DMS mapping experiments as a function of $MgCl_2$, in background of 50 mM Na-HEPES, pH 8.0, at room temperature (24 °C). MiniTTR 16 could not be fit due to problems with electrophoresis. Folding stabilities were estimated based on the approximation of $\Delta G = -nk_B T \ln([Mg^{2+}] / K_d)$, where n is the apparent Hill coefficient, k_B is the Boltzmann constant, T is the temperature, and K_d is the $[Mg^{2+}]$ at the folding midpoint.

Construct	SHAPE and DMS support Secondary Structure[a]	TTR DMS Reactivity Fold Change[b]	Native Gel Mobility Shift (cm)[c]	Mg^{2+} Folding Midpoints[d]
miniTTR 1	95.2%	3.01	0.205	1.12 +0.34/-0.24
miniTTR 2	94.2%	6.94	0.247	0.08 +0.01/-0.01
miniTTR 3	96.6%	1.63	0.055	>10
miniTTR 4	96.6%	1.74	0.204	>10
miniTTR 5	98.1%	4.1	0.236	1.64 +0.32/-0.22
miniTTR 6	95.5%	3.39	0.382	0.74 +0.01/-0.02
miniTTR 7	97.2%	2.66	0.226	3.31 +0.79/-0.55
miniTTR 8	98.5%	1.16	-1.117	>10
miniTTR 9	98.5%	6.18	0.348	0.84 +0.11/-0.11
miniTTR 10	98.5%	6.59	0.405	0.74 +0.08/-0.06
miniTTR 11	96.7%	4.79	0.282	0.87 +0.13/-0.10
miniTTR 12	96.4%	5.3	0.406	0.50 +0.05/-0.03
miniTTR 13	94.2%	1.72	-0.066	>10
miniTTR 14	98.6%	5.21	0.408	0.44 +0.02/-0.01
miniTTR 15	94.2%	3.79	-0.108	0.95 +0.14/-0.14
miniTTR 16	96.2%	14.47	0.456	0.24 +0.08/-0.02

[a]Percent of helical residues that have SHAPE and DMS reactivities < 0.5 reactivity units, suggesting they are in base pairs.
[b]For DMS chemical mapping with and without 10 mM Mg^{2+}, a 2-fold reduction in mean DMS reactivity at the four TTR adenines was considered to pass screen (green).
[c]Distance traveled in gel of RNA compared to mutant with tetraloop GAAA changed to UUCG. Positive numbers correspond to faster gel mobility (more compact fold) with wild type tetraloop, as expected for correctly folded RNA.
[d]RNA that was more than half folded with [Mg^{2+}]< 10 mM was considered to pass screen

Supplementary Table 3: Biochemical assay results for each miniTTR design.

SHAPE, DMS, native gel, and Mg^{2+} titration assays for the miniTTR constructs. Coloring of green or red annotates whether construct passed screen or not, as described in footnote.

Design	Reactivity DMS Change of A9 and A10 upon ATP binding[a]	Mean DMS reactivity at TTR without ATP[b]	Formed TTR with ATP (fold change in DMS reactivity)[c]	K_d for ATP, µM[d]
ATP-TTR 1*	n.d.	n.d.	n.d.	n.d.
ATP-TTR 2*	n.d.	n.d.	n.d.	n.d.
ATP-TTR 3	-0.24	0.04	1.00	1.5 +0.51/-0.38
ATP-TTR 4	-0.24	0.09	1.46	4.1 +1.30/-0.96
ATP-TTR 5	-0.27	0.17	1.94	1.4 +0.46/-0.35
ATP-TTR 6*	0.02	0.14	2.28	n.d.
ATP-TTR 7*	0.04	0.27	1.85	n.d.
ATP-TTR 8	-0.11	1.28	1.16	n.d.
ATP-TTR 9	-0.71	0.28	2.84	n.d.
ATP-TTR 10	-0.22	1.26	0.90	n.d.
ATP aptamer	-0.41	n.a.	n.a.	16.2 +5.70/-4.00

[a] Decrease in reactivity beyond 0.2 exceeds experimental error and considered evidence for ATP binding at ATP aptamer. Values normalized to mean DMS reactivity of single-stranded adenosines in reference GAGUA hairpins flanking design.
[b] Mean DMS reactivity less than 0.5 taken as evidence for tetraloop/tetraloop-receptor (TTR) formation.
[c] Fold change in DMS reactivity with and without ATP. If both the mean reactivity is under 0.5 and the fold change is under 2 it is considered a success.
[d] K_d lower than reference ATP aptamer demonstrated successful stabilization of ATP aptamer.
[e] Chemical mapping data for ATP-TTR 1 and 2 could not be processed due to strong stops on the capillary electrophoresis readout.
* Construct had strong stops in capillary electrophoresis, making data too weak to be reliable

Supplementary Table 4: Summary of DMS chemical mapping experiments for all ATP-TTR constructs.

Coloring of green or red annotates whether construct passed screen or not, as described in footnotes.

Design	SHAPE TTR reactivity[a]	DFHBI K_d (nM)	DFHBI B_{max} (AU)
Spinach-TTR 1	0.11	76.7±41.0	17909±13
Spinach-TTR 2	0.03	103.7±19.2	31910±92
Spinach-TTR 3	0.02	37.5±52.9	28195±35
Spinach-TTR 4	0.07	847.2±215.3	38576±27
Spinach-TTR 5	n.d.	n.d.	n.d.
Spinach-TTR 6	0.06	211.7±71.9	48657±33
Spinach-TTR 7	0.51	83.9±15.0	23577±60
Spinach-TTR 8	0.04	25.7±20.2	31861±17
Spinach-TTR 9	0.03	49.3±35.5	25260±19
Spinach-TTR 10	0.19	49.3±14.2	41381±13
Spinach-TTR 11	0.03	61.9±26.6	47226±25
Spinach-TTR 12	0.06	52.5±51.5	33131±36
Spinach-TTR 13	0.03	43.6±49.4	39419±44
Spinach-TTR 14	0.29	81.7±53.2	35107±32
Spinach-TTR 15	0.03	61.7±6.6	31691±41
Spinach-TTR 16	0.63	102.9±46.8	34933±25
Spinach	n.a.	58.4±50.1	6473±65
Spinach min	n.a.	91.7±34.0	39544±22

[a] The average SHAPE reactivity of the 3 Adenines in the tetraloop under 0.5 is considered protected.

Supplementary Table 5: Initial screen of Spinach-TTR molecules.

SHAPE reactivity at tetraloop/receptors; and the binding affinity and brightness of the Spinach-TTRs compared to controls, based on titrations of the light-up fluorophore DFHBI.

Design	K_d (nM)	Brightness, relative to Broccoli	Quantum Yield	Folding Efficiency	time to ½ signal in lysate (min)
Spinach-TTR 3	14.3±1.6	1.17	0.84	38.3±1.6	131.1
Spinach-TTR 6	110.1±12.7	1.07	0.77	53.3±2.0	41.2
Spinach-TTR 8	20.5±4.1	1.17	0.84	36.9±1.2	92.9
Spinach-TTR 9	26.2±2.8	1.07	0.77	29.3±0.5	68.2
Spinach-TTR 10	33.8±6.6	1.1	0.79	60.8±3.1	86.8
Spinach-TTR 11	37.8±7.5	1.11	0.8	59.2±2.3	11.0
Spinach-TTR 13	28.9±5.5	1.1	0.79	49.7±2.0	58.5
Spinach	14.6±1.3	0.38	0.27	9.5±0.2	17.4
Spinach-min	48.7±7.3	1.03	0.74	39.5±3.1	9.5
Broccoli	27.8±2.7	1	0.72	28.5±1.8	10.5

Supplementary Table 6: Summary for fluorescent properties of Spinach-TTRs compared to controls, based on DFHBI and RNA titrations (see Methods).

Design problem	miniTTR	Ribosome tether	ATP-TTR	Spinach-TTR
miniTTR	23	3	5	4
Ribosome tether	3	14	2	3
ATP-TTR	5	2	16	4
Spinach-TTR	4	3	4	26

Supplementary Table 7: Overlap of motifs selected by RNAMake to solve the four different design problems.

Gray boxes give total number of motifs used for each problem, including those shared across problems; white boxes show number of shared motifs between problems.

Motif	Example study using motif in 3D design
IRES domain IIa	Dibrov et al. PNAS 2011 (40)
AAG/GA junction	Geary et al. NAR 2011 (41)
CGA/UAAG junction	Geary et al. NAR 2011 (41)
A-bulge	Geary et al. Nano Lett. 2017 (42)
Kink-Turn	Huang and Lilley. Nanoscale 2016 (43)
tRNA 4-way junction	Geary et al. Nano Lett. 2017 (42)
UA_h 3WJ	Geary et al. Nano Lett. 2017 (42)
pRNA 3WJ	Zhang et al. RNA 2013 (44)
16S 3WJ	Bindewald et al. ACS Nano. 2012 (45)

Supplementary Table 8: Structural motifs used in previous RNA design studies, used to test impact of much larger size of RNAMake motif library.

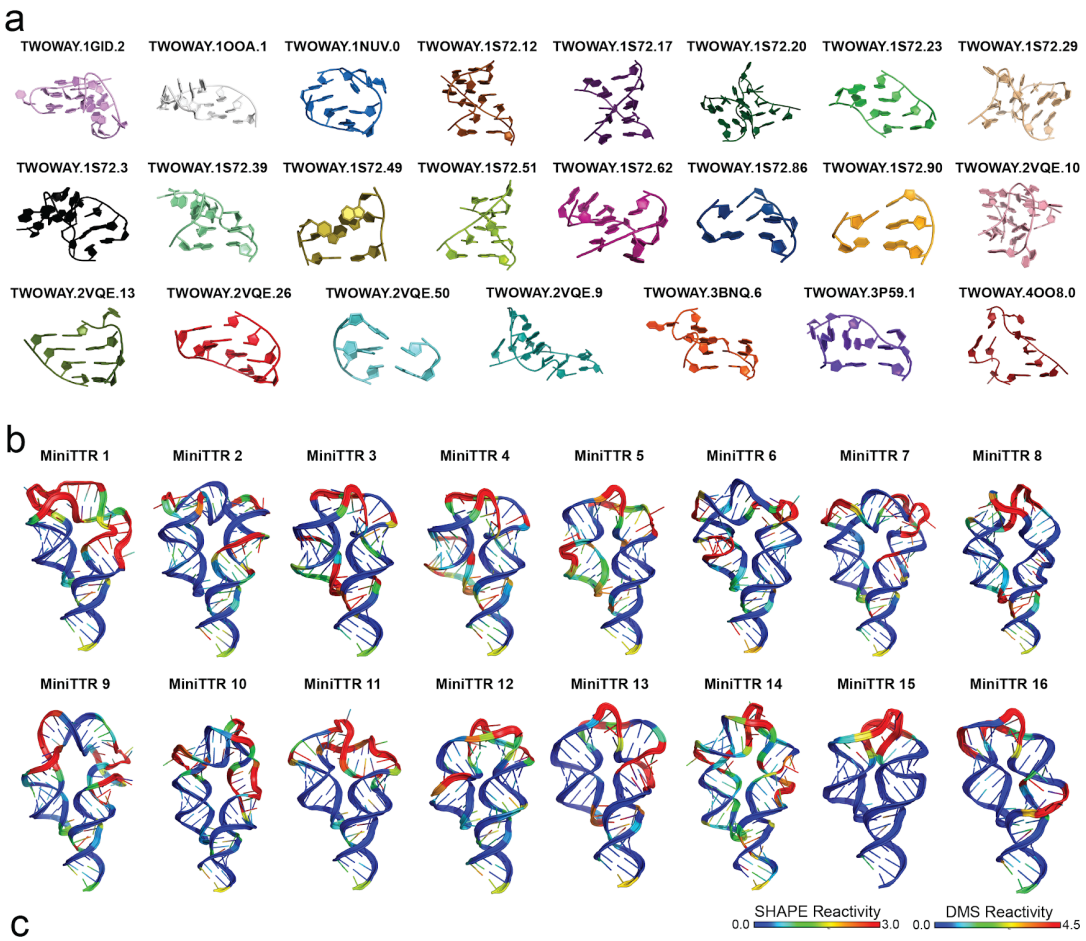
X3DNA bp Type	Leontis-Westhof	Energetic Penalty
cm-	N/A	6.11
cM-M	tHH	6.11
tW+W	tWW	3.11
c.+M	N/A	5.69
.W+W	N/A	6.11
tW-M	tWH	2.42
tm-M	tSH	2.72
cW+M	cWH	3.33
.W-W	N/A	4.33
cM+.	N/A	6.11
c.-m	N/A	6.11
cM+W	cHW	4.40
tM+m	N/A	6.11
tM-W	tHW	3.02
cm-m	cSS	5.12
cM-W	tHW	6.11
cW-W	cWW	-2.00
c.-M	N/A	5.44
cm+M	cSH	2.71
cm-M	tSH	3.23
....	N/A	4.18
cm-W	cSW	4.37
tM-m	tSH	2.84
c.-W	N/A	6.11
cM+m	cHS	5.69
cM-m	tSH	3.12

Supplementary Table 9: RNAMake scoring penalties for each base pair type.

Penalties applied based on frequencies of each base pair type observed in RNA crystallographic database, as $-k_B T \ln(\text{frequency})$. Each residue not in a base pair adds an additional 2.0 to the total penalty for a motif.

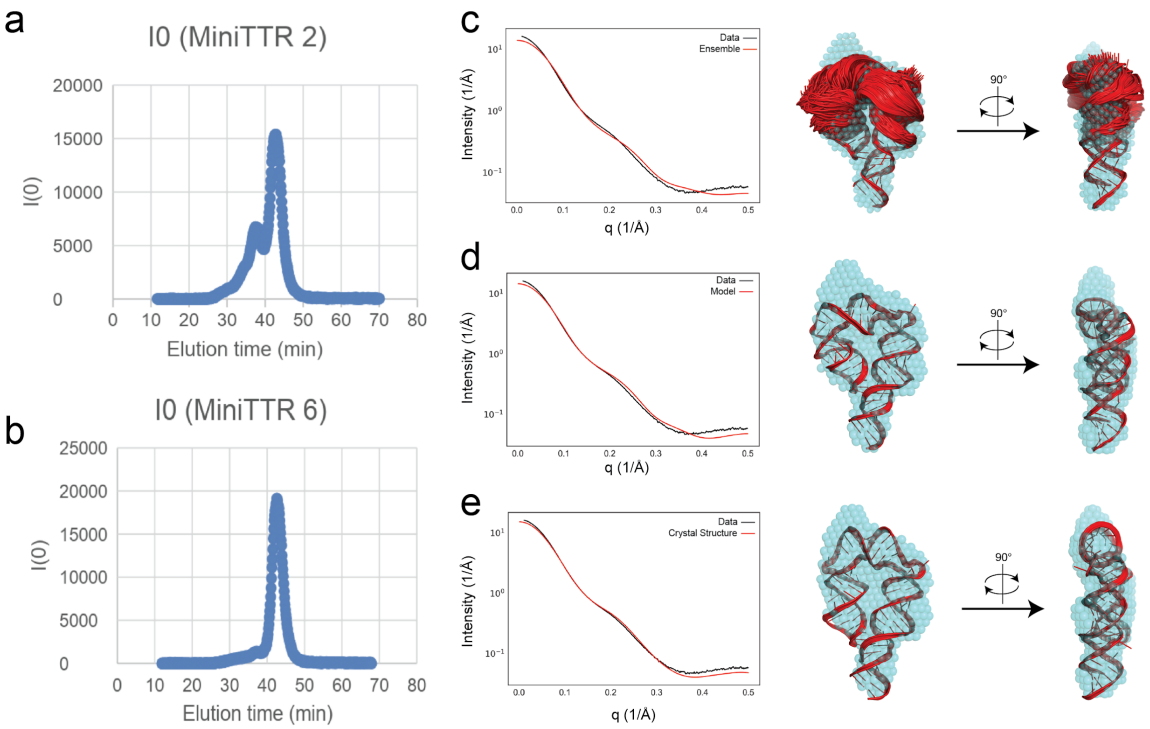
	Native
Space group	C2
Cell dimensions	
a, *b*, *c* (Å)	233.372, 25.358, 42.861
α, β, γ (°)	90.0, 99.7, 90.0
Wavelength	1.58954
Resolution (Å)	2.55-115 (2.55-2.75)
R_{sym} or R_{merge}	13.5(72.7)
I / σI	6.445(1.0)
Completeness (%)	94.3(88.9)
Redundancy	1.555(1.376)
Refinement	
Resolution (Å)	2.55-50.0
No. reflections	8467
R_{work} / R_{free}	23.2/26.9
No. atoms	2025
RNA	2008
Ligand/ion	15
Water	2
B-factors	
RNA	41.0
Ligand/ion	32.8
Water	5.9
R.m.s deviations	
Bond lengths (Å)	0.0056
Bond angles (°)	1.367

Supplementary Table 10: Data collection, phasing and refinement statistics for miniTTR 6.


a
TWOWAY.1GID.2
TWOWAY.1OOA.1
TWOWAY.1NUV.0
TWOWAY.1S72.12
TWOWAY.1S72.17
TWOWAY.1S72.20
TWOWAY.1S72.23
TWOWAY.1S72.29
TWOWAY.1S72.3
TWOWAY.1S72.39
TWOWAY.1S72.49
TWOWAY.1S72.51
TWOWAY.1S72.62
TWOWAY.1S72.86
TWOWAY.1S72.90
TWOWAY.2VQE.10
TWOWAY.2VQE.13
TWOWAY.2VQE.26
TWOWAY.2VQE.50
TWOWAY.2VQE.9
TWOWAY.3BNQ.6
TWOWAY.3P59.1
TWOWAY.4OO8.0
b
MiniTTR 1
MiniTTR 2
MiniTTR 3
MiniTTR 4
MiniTTR 5
MiniTTR 6
MiniTTR 7
MiniTTR 8
MiniTTR 9
MiniTTR 10
MiniTTR 11
MiniTTR 12
MiniTTR 13
MiniTTR 14
MiniTTR 15
MiniTTR 16
SHAPE Reactivity
0.0
3.0
DMS Reactivity
0.0
4.5
c
Med Loop
MiniTTR 1
MiniTTR 3
MiniTTR 4
MiniTTR 5
MiniTTR 7
MiniTTR 9
Med Loop
Med Loop
MiniTTR 10
MiniTTR 11
MiniTTR 13
MiniTTR 14
MiniTTR 15
MiniTTR 16
Med Loop

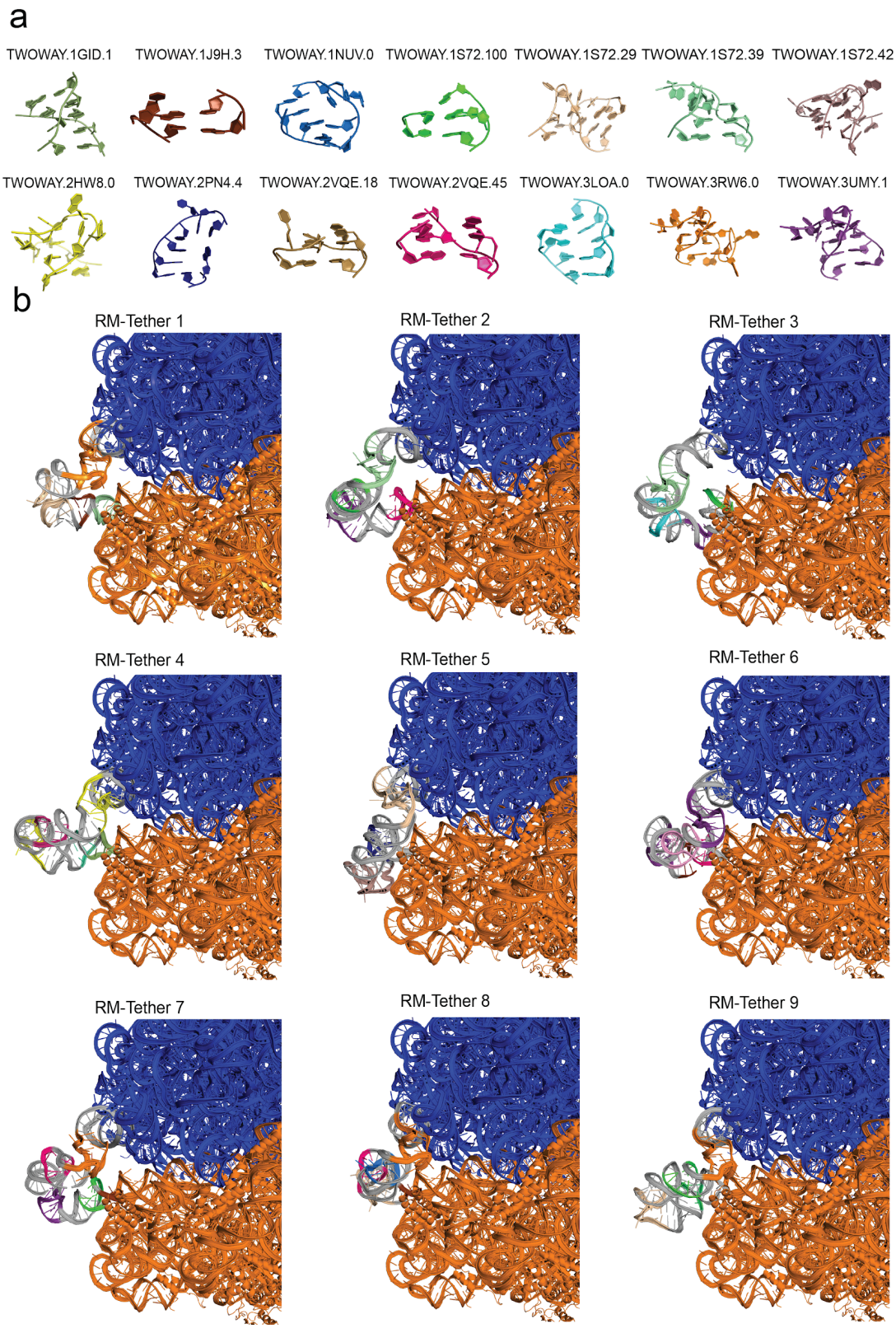
Supplementary Figure 1: Motif analysis, SHAPE and DMS reactivities and native gels for miniTTR constructs.

a) All motifs that appear in the 16 miniTTR constructs. b) Chemical mapping for each design, measured in 10 mM $MgCl_2$, 50 mM Na-HEPES, pH 8.0, at ambient temperature (24 °C). The backbone is colored based on the SHAPE reactivity, and the sticks are colored according to the DMS reactivity of each residue. c) Native gel mobility assays for all miniTTR constructs excluding 2, 6, 8 and 12 (see Figure 2)



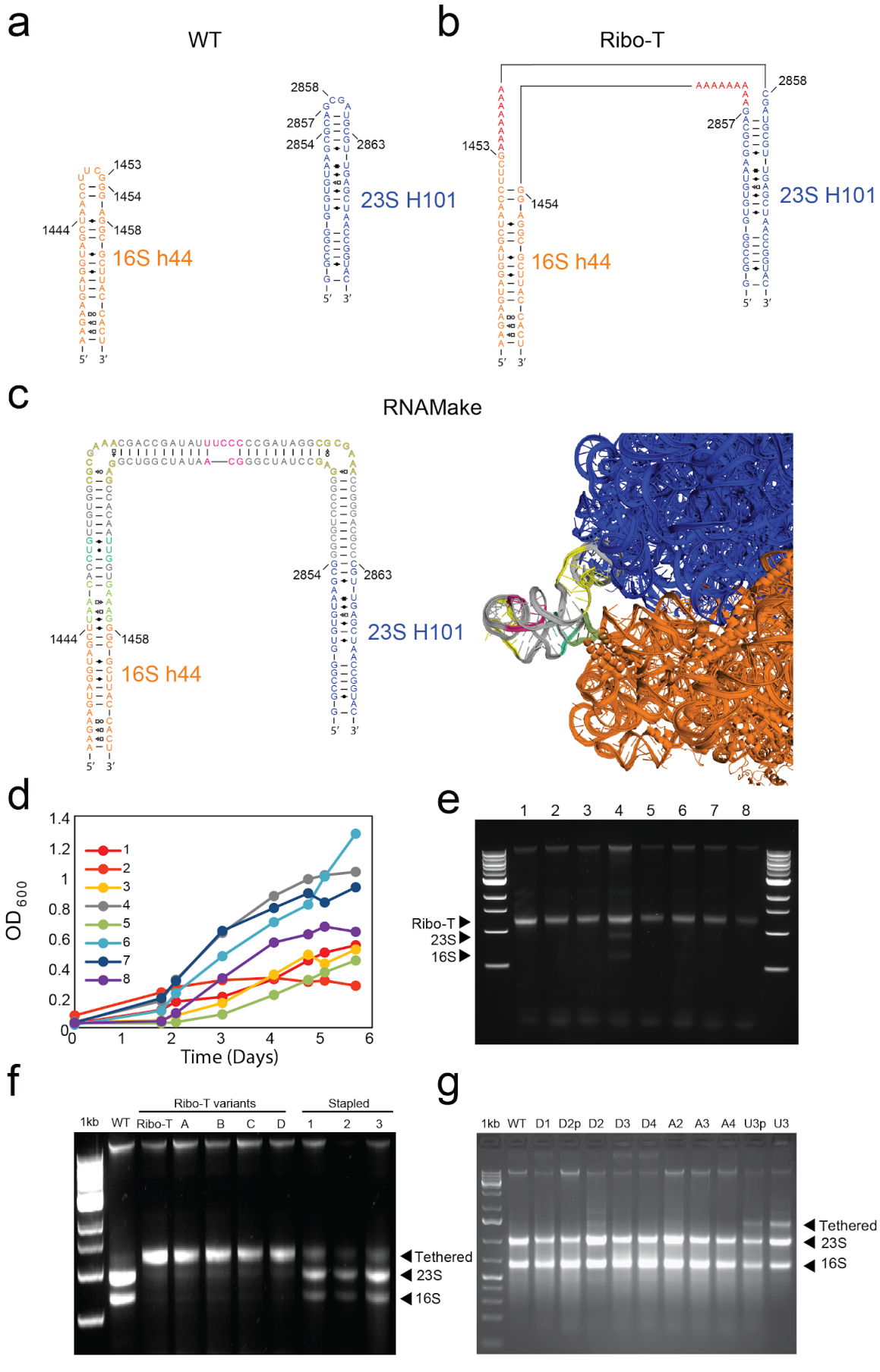
Supplementary Figure 2: SAXS-monitored size-exclusion FPLC of miniTTR 2 and miniTTR 6.

Size-exclusion FPLC of miniTTR designs 2 (a) and 6 (b) were monitored by SAXS (see Methods). The I(0) of each elution fraction, assuming a single component, is proportional to both the concentration and molecular mass of the component. Thus, the ratio of the molar fraction of a dimer peak to that of a monomer peak will be approximately half the ratio of their peak areas. The main elution peak at approximately 43 min and the minor peak at approximately 37.5 min were identified as monomers and dimers, respectively, using a method that can directly determine molecular mass from the actual SAXS profiles (1). Integration of I(0) curves confirm that <15% and <5% of miniTTRs 2 and 6, respectively, occur in higher-order structures. MiniTTR 6 compared to predicted profiles and RNAMake 3D models for (c) estimated thermal ensemble of the RNAMake model, (d) the modeled centroid structure from the thermal ensemble, and (e) the crystal structure.


a
TWOWAY.1GID.1
TWOWAY.1J9H.3
TWOWAY.1NUV.0
TWOWAY.1S72.100
TWOWAY.1S72.29
TWOWAY.1S72.39
TWOWAY.1S72.42
TWOWAY.2HW8.0
TWOWAY.2PN4.4
TWOWAY.2VQE.18
TWOWAY.2VQE.45
TWOWAY.3LOA.0
TWOWAY.3RW6.0
TWOWAY.3UMY.1
b
RM-Tether 1
RM-Tether 2
RM-Tether 3
RM-Tether 4
RM-Tether 5
RM-Tether 6
RM-Tether 7
RM-Tether 8
RM-Tether 9

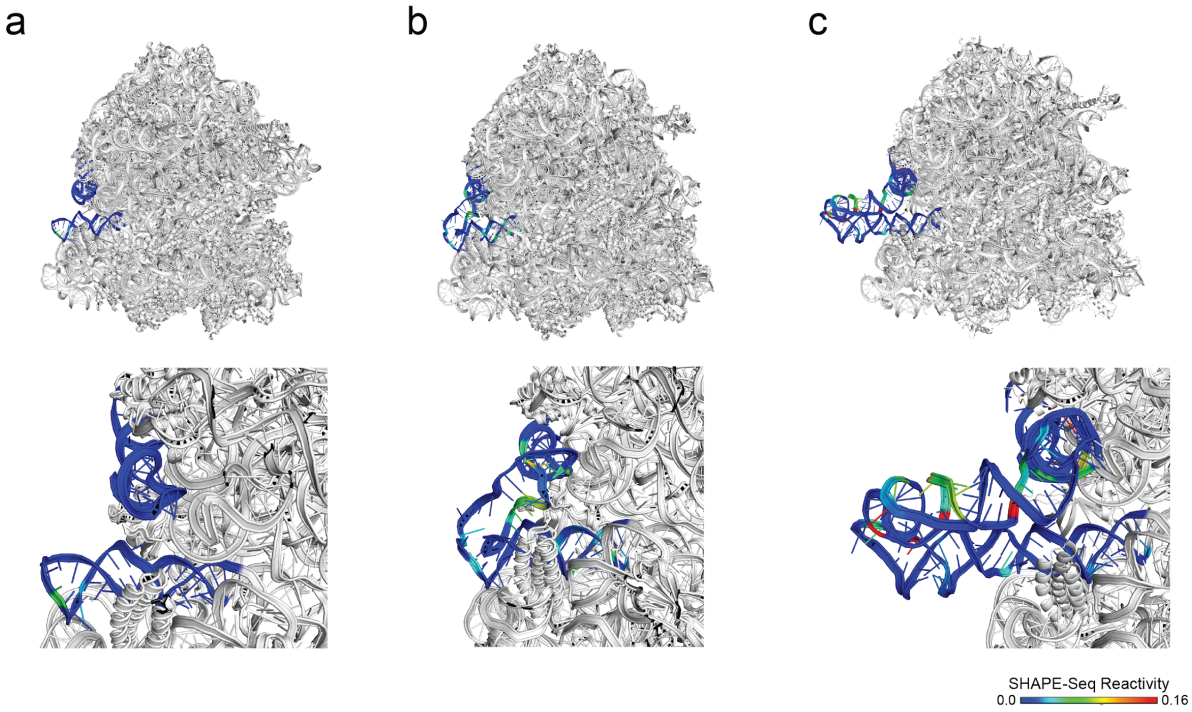
Supplementary Figure 3: Summary of RM-Tether designs.

a) all motifs incorporated into 9 RM-Tether designs by RNAMake, to bridge between h44 and H101 in small and large ribosomal RNAs, respectively. b) All 9 RNAMake RM-Tether designs, colored according to the motifs that were used to generate it, helices (gray), small subunit (orange), and large subunit (blue).

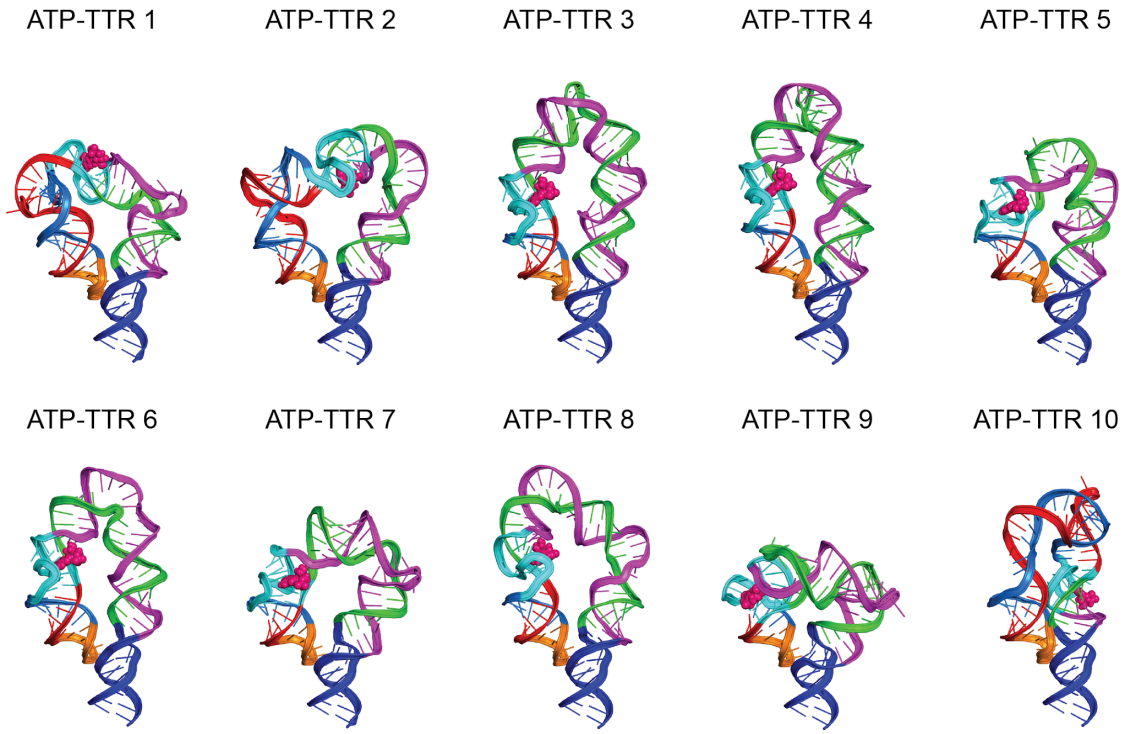

a
WT
1453
1454
1444
1458
16S h44
2858
2857
2854
2863
23S H101
b
Ribo-T
AAAAAAA
1453
1454
16S h44
2858
2857
23S H101
c
RNAMake
1444
1458
16S h44
2854
2863
23S H101
d
OD600
Time (Days)
e
1 2 3 4 5 6 7 8
Ribo-T
23S
16S
f
Ribo-T variants
Stapled
1kb WT Ribo-T A B C D 1 2 3
Tethered
23S
16S
g
1kb WT D1 D2p D2 D3 D4 A2 A3 A4 U3p U3
Tethered
23S
16S

Supplementary Figure 4: Secondary structure comparison and full gels for ribosome tethering studies

a) Secondary structure of 16S helix 44 (h44) and wild-type 23S helix 101 (H101). b) Published Ribo-T secondary structure (2). c) (left) Secondary structure of RM-Tether 4 designed with RNAMake. Each motif is colored as in the 3D structure model (right). d) Culture density of *E. coli* cells containing RNAMake RM-Tether design 4. e) Gel assay of 8 distinct colonies, verifying the RNAMake-designed Ribo-T. In one case, lane 4, we observed very faint 16S and 23S-like rRNA bands, possibly reflecting linker cleavage either in the cell or during ribosome isolation. f) Agarose gel electrophoresis of RNA extracted from *E. coli* (Squires strain SQ171fg) in which wild type ribosomes were completely replaced with the indicated ribosomes. The double band under the wild type lane indicates the 16S and 23S rRNA in the two separate subunits of the ribosome. The Ribo-T lane represents the original Ribo-T construct published by Orelle *et al.* (10). The Ribo-T variants (a-d) represent tether variants of Ribo-T (unpublished data from the Jewett lab) that remain tethered, and the stapled ribosomes (1-3) represent tethers generated from published work by Fried, *et al.* (46). In our system and constructs, stapled tethers do not maintain a single-subunit entity, resulting in a 3-band pattern. g) U3 tether (Umbilical 3) designed by the Jewett lab, presents as three bands *in vivo* (last lane).

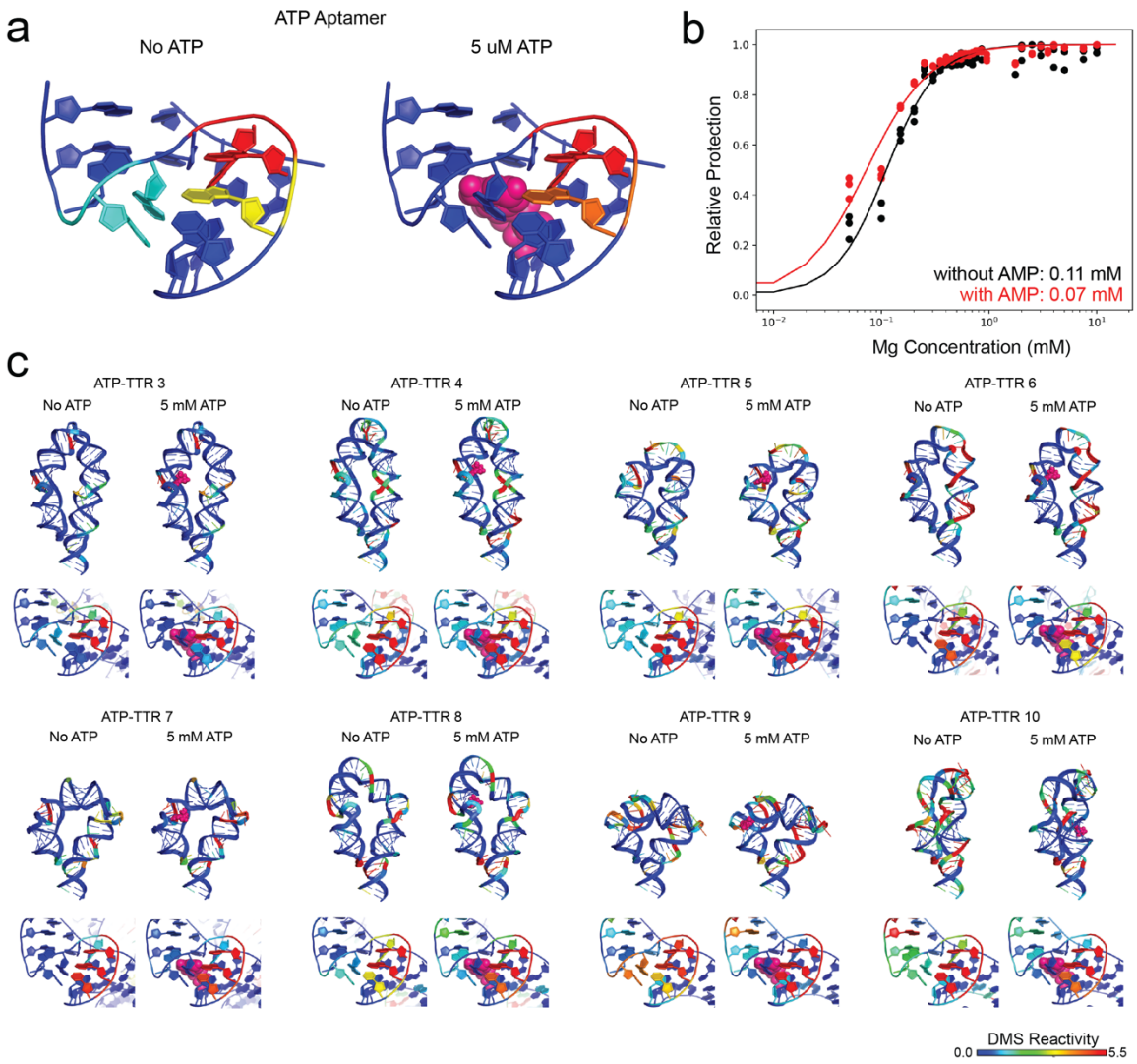


Supplementary Figure 5: SHAPE-Seq chemical mapping of tethered ribosomes, colored on backbones of (a) WT ribosome, (b) Ribo-T tethered ribosome, and (c) RM-Tether design 4.



Supplementary Figure 6: Models of 10 ATP-TTRs designed by RNAMake to 'lock' the ATP aptamer through a bracing tertiary contact.

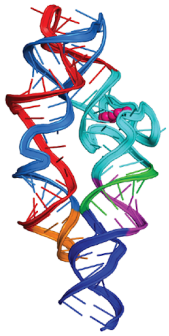
The small-molecule binding aptamer (cyan; ATP molecule in pink) connected by four strands (green, purple, teal, magenta) to a peripheral tertiary contact (orange, blue).



Supplementary Figure 7: Chemical mapping for the ATP-TTR designs

a) DMS chemical mapping of the ATP aptamer in isolation with and without 5 mM ATP, in 10 mM $MgCl_2$, 50 mM Na-HEPES, pH 8.0, at ambient temperature (24 °C). b) 32 point Mg^{2+} titration of ATP-TTR 3 with and without 50 μM AMP, assessed by DMS reactivities at tetraloop and receptor (AMP used instead of ATP to avoid Mg^{2+} chelation effects). c) in each panel, (top) DMS chemical mapping without and with 5 mM ATP, (bottom) zoomed and aligned ATP aptamer section of each design.

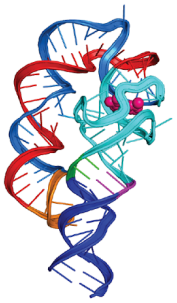
Spinach-TTR 1



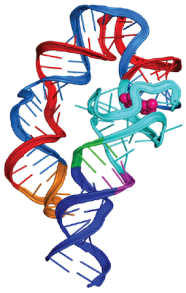
Spinach-TTR 2

Spinach-TTR 3

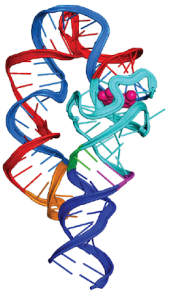
Spinach-TTR 4



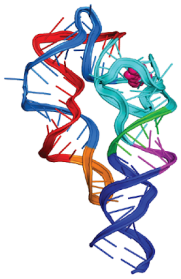
Spinach-TTR 5



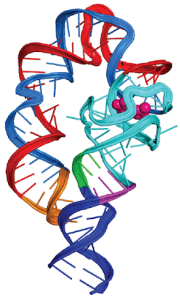
Spinach-TTR 6



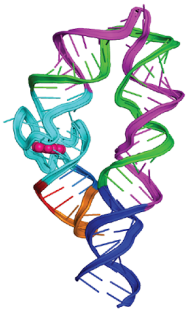
Spinach-TTR 7



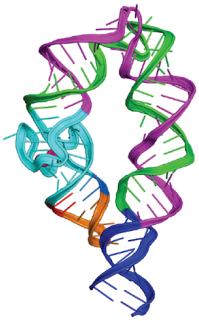
Spinach-TTR 8



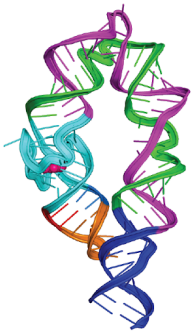
Spinach-TTR 9



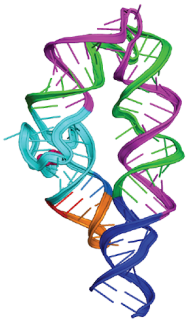
Spinach-TTR 10



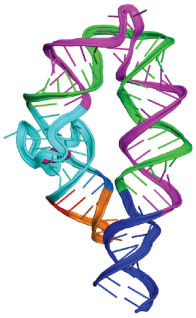
Spinach-TTR 11



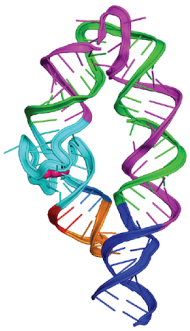
Spinach-TTR 12



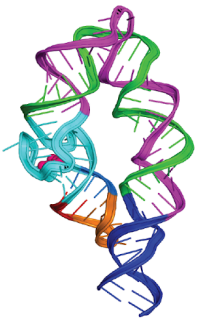
Spinach-TTR 13



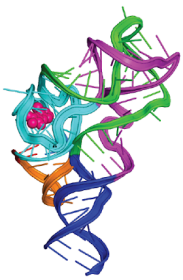
Spinach-TTR 14



Spinach-TTR 15

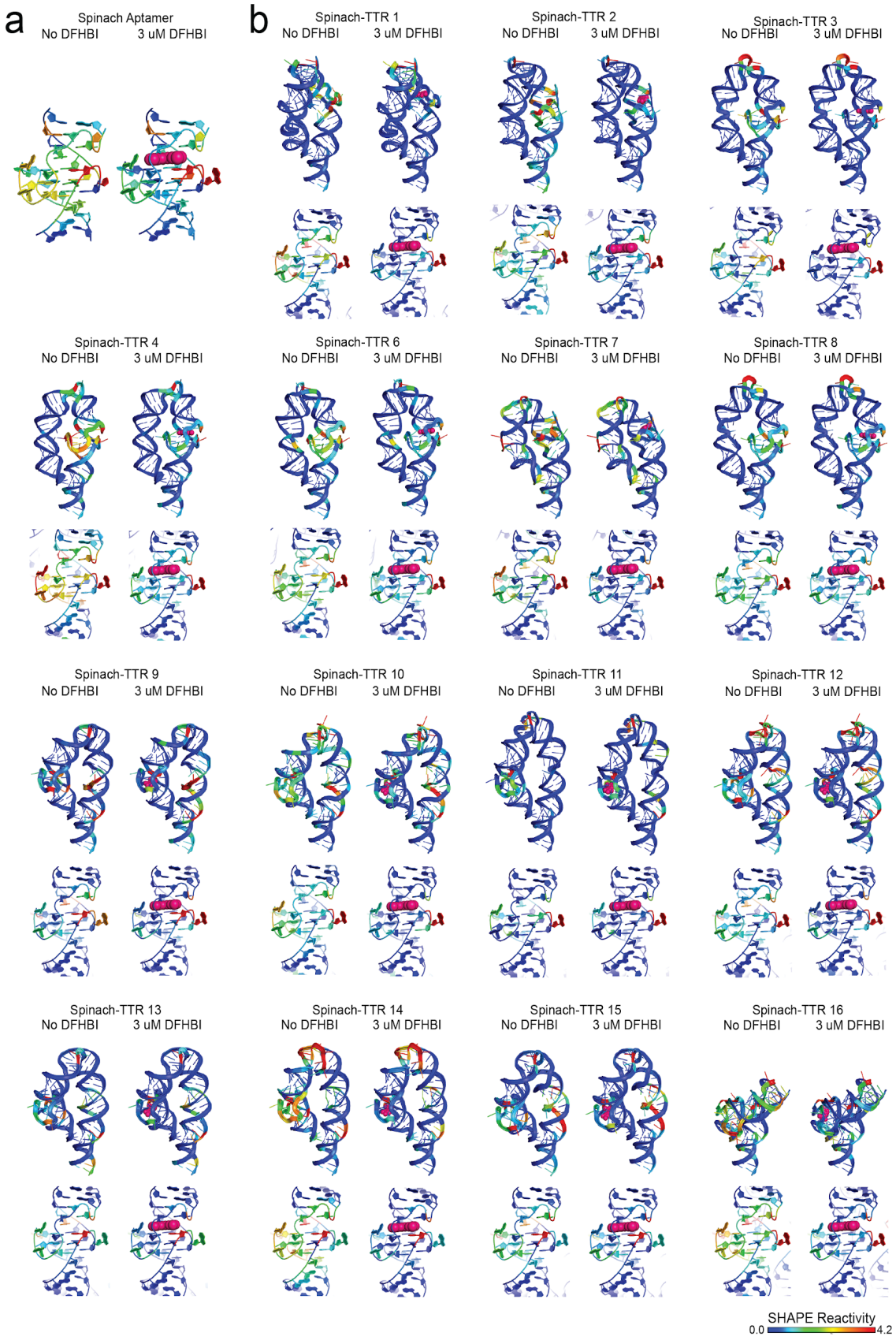


Spinach-TTR 16



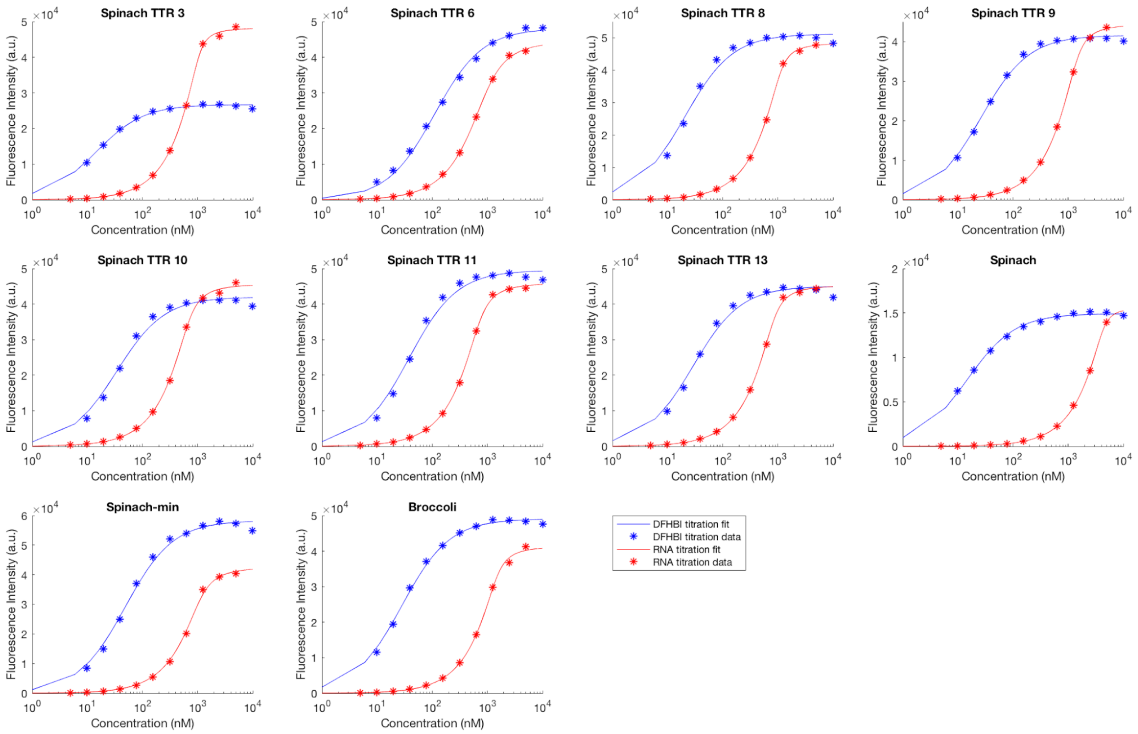
Supplementary Figure 8: Models of 16 Spinach-TTRs designed by RNAMake to 'lock' the Spinach aptamer for the DFHBI fluorophore through a bracing tertiary contact.

The small-molecule binding Spinach aptamer (cyan; DFHBI molecule in pink) connected by four strands (green, purple (for the first two strands); teal, magenta (for the second two strands)) to a peripheral tertiary contact (orange, blue).

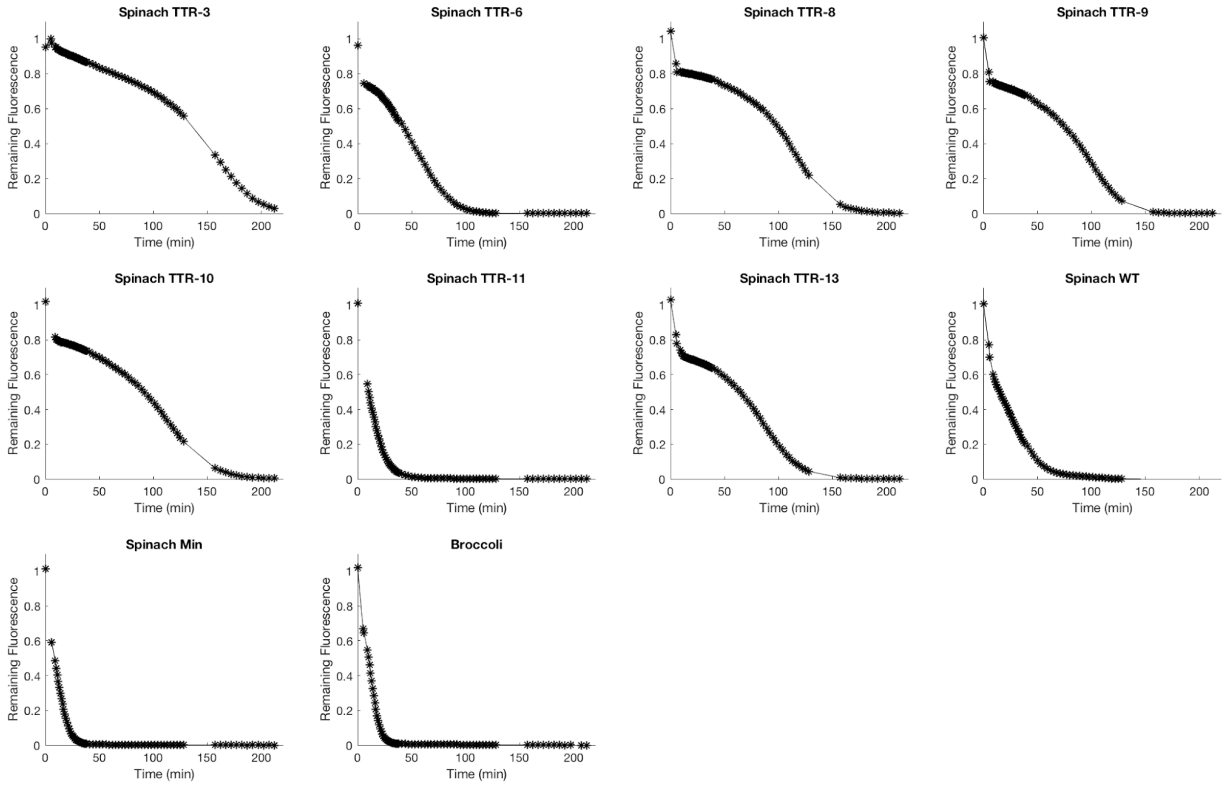

a
Spinach Aptamer
No DFHBI
3 uM DFHBI
b
Spinach-TTR 1
Spinach-TTR 2
Spinach-TTR 3
Spinach-TTR 4
Spinach-TTR 6
Spinach-TTR 7
Spinach-TTR 8
Spinach-TTR 9
Spinach-TTR 10
Spinach-TTR 11
Spinach-TTR 12
Spinach-TTR 13
Spinach-TTR 14
Spinach-TTR 15
Spinach-TTR 16
SHAPE Reactivity
0.0
4.2

Supplementary Figure 9: SHAPE chemical mapping for Spinach aptamer with and without 3 μM DFHBI.

a) SHAPE chemical mapping of Spinach aptamer with and without 3 μM DFHBI, in 10 mM $MgCl_2$, 1 M KCl, 50 mM Na-HEPES, pH 8.0, at ambient temperature (24 °C). b) (top) SHAPE chemical mapping for each of the Spinach-TTR constructs without and with 3 μM DFHBI, (bottom) zoomed and aligned Spinach aptamer section of each design.

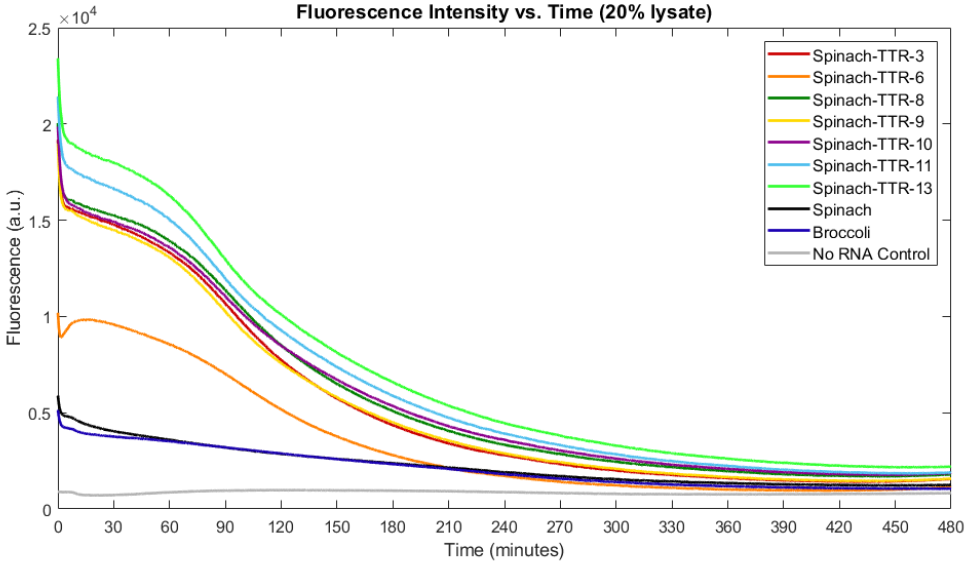


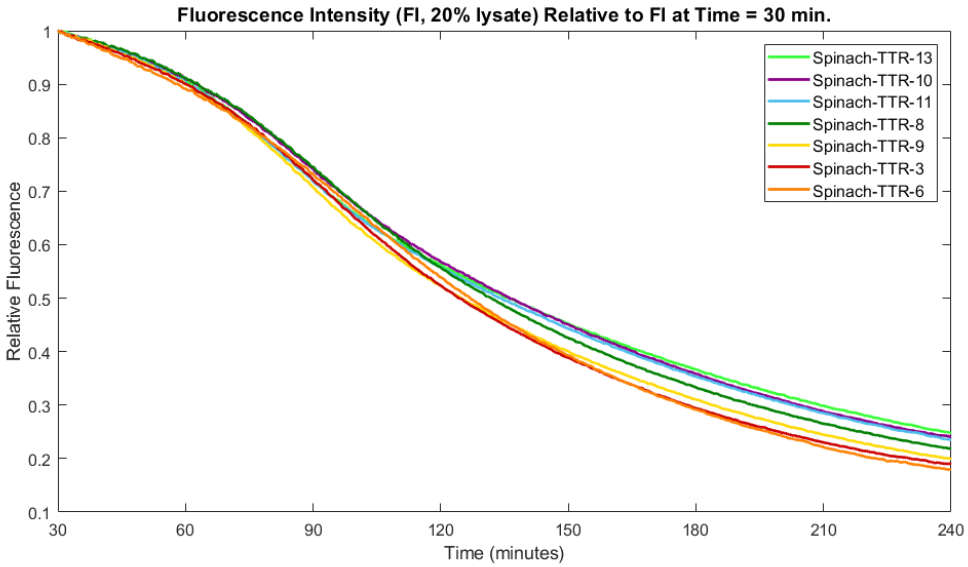
Supplementary Figure 10: Fits for RNA and DFHBI titrations for Spinach-TTRs. Titrations were performed for each Spinach-TTR to determine binding affinities and fluorescence quantum yields. DFHBI titrations (blue) were performed at 200 nM of RNA with the DFHBI concentration varying from 10 nM to 10μM in a 1:2 dilution series. RNA titrations (red) were performed at 400 nM of DFHBI with the RNA concentration for each species ranging from 5 nM to 5μM in a 1:2 dilution series.



Supplementary Figure 11: Summary of Spinach-TTR/DFHBI fluorescence after challenge with 20% *Xenopus* egg lysate.

RNA for each species was prepared in duplicate in 80 µL of 1X PBSMKT (1X phosphate buffered saline pH 7.2, 5 mM $MgCl_2$, 100 mM KCl, 0.01% Tween-20) buffer and measured at t=0 to confirm the duplicates had comparable brightness. Then, 20 µL of PBSMKT or 20 µL of 100% fresh *Xenopus* egg lysate was added to each replicate and pipette mixed to give a final RNA concentration of 500 nM and DFHBI concentration of 2.5 µM. Measurements were taken every 1 minute for minutes 5-40, then every 3 minutes for minutes 40-130. After a 30 minute resting period, measurements were performed every 5 minutes for minutes 160-210.

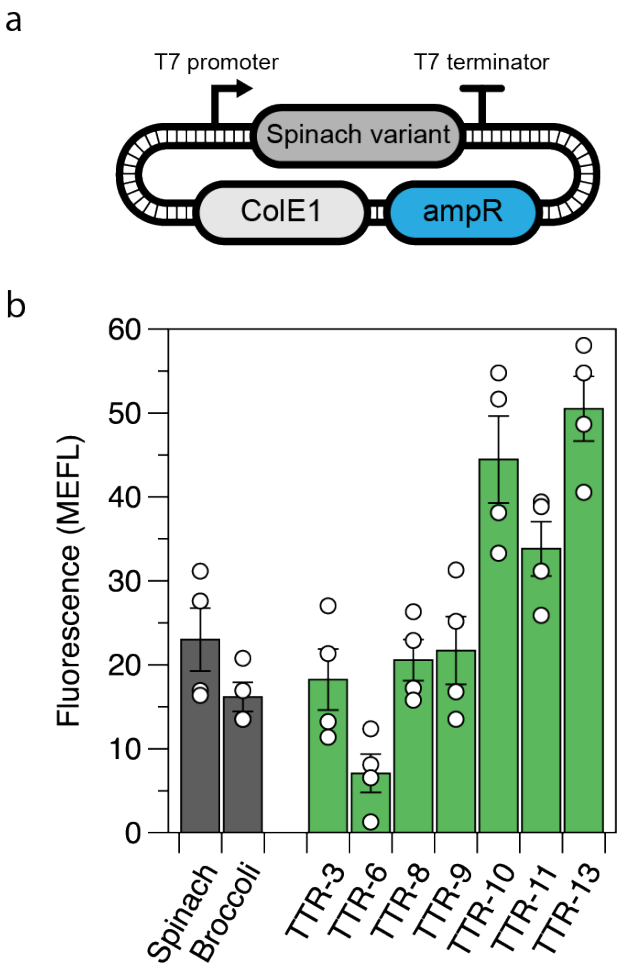




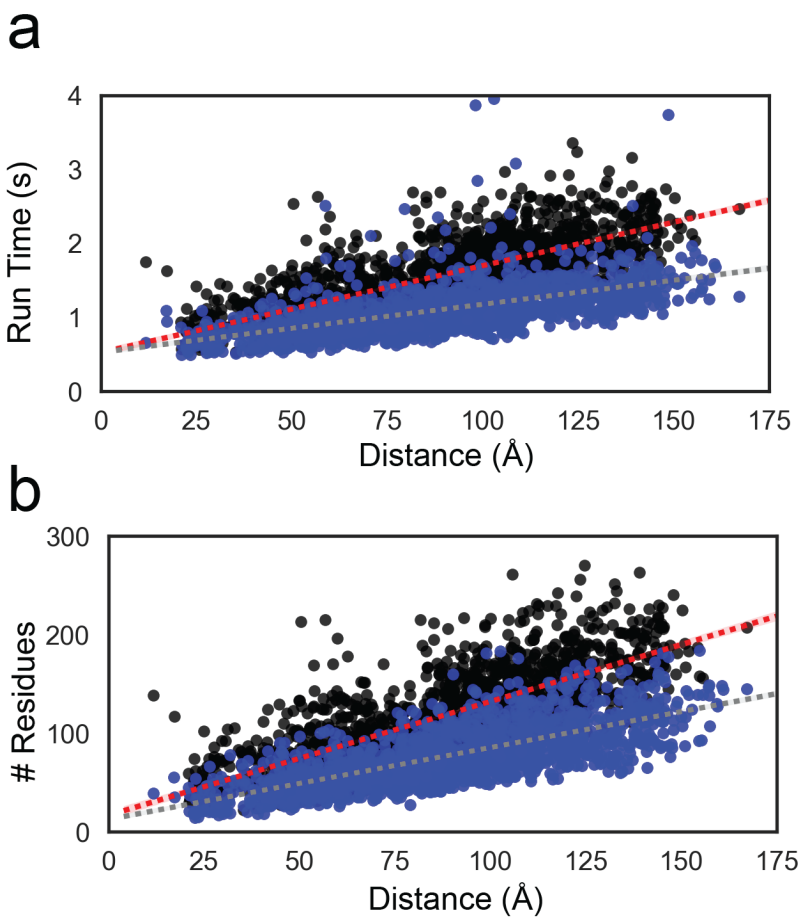
Supplementary Figure 12: Summary of Spinach-TTR/DFHBI fluorescence after challenge with 20% *E. coli* lysate.

a) Fluorescence intensity (FI) of each Spinach-TTR construct, RNA Spinach, and RNA Broccoli immediately after addition of lysate-containing buffer (see Methods) to yield final

concentrations of 500 nM RNA, 2.5 μM DFHBI, and 20% *E. coli* lysate. Fluorescence intensities (FI) were measured every 30 s for 8 hours (excitation/emission: 462±9 nm/504±20 nm). Samples were prepared and measured in triplicate and the average values of the three measurements are shown. b) Long-term stabilities of the TTR constructs in 20% *E. coli* lysate estimated by normalizing the FI of each construct to the same construct's FI 30 minutes after addition of lysate.



Supplementary Figure 13: *E. coli in vivo* performance of Spinach-TTR constructs. a) Plasmid schematic for *E. coli in vivo* testing of Spinach-TTR constructs. b) Measured fluorescence of Spinach-TTR constructs following subtraction of background fluorescence. Spinach and Broccoli are shown in grey, and designed variants are shown in green. Error bars represent the standard error of the mean (SEM) of four biological replicates, and circles represent each measurement.



Supplementary Figure 14: Benchmarking the use of higher order junctions

Computational efficiency of RNAMake for designing connections between each pair of hairpins on the 50S *E. coli* ribosomal subunit with (blue) and without multiway junctions (black); compare to Figures 1e-f. Even including multiway junctions, run time continues to scale linearly with problem size, as measured by (a) translational distance between helical endpoints or (b) number of residues required for segments. Solutions containing multiway junctions are on average produced faster and contain fewer residues.

Supplementary References

1. Kladwang, W., Mann, T. H., Becka, A., Tian, S., Kim, H., Yoon, S., and Das, R. (2014) Standardization of RNA chemical mapping experiments. *Biochemistry* **53**, 3063–3065
2. Murphy, F. L., and Cech, T. R. (1993) An independently folding domain of RNA tertiary structure within the Tetrahymena ribozyme. *Biochemistry* **32**, 5291–5300
3. Young, B. T., and Silverman, S. K. (2002) The GAAA tetraloop-receptor interaction contributes differentially to folding thermodynamics and kinetics for the P4-P6 RNA domain. *Biochemistry* **41**, 12271–12276
4. Frederiksen, J. K., Li, N.-S., Das, R., Herschlag, D., and Piccirilli, J. A. (2012) Metal-ion rescue revisited: biochemical detection of site-bound metal ions important for RNA folding. *RNA* **18**, 1123–1141
5. Takamoto, K., Das, R., He, Q., Doniach, S., Brenowitz, M., Herschlag, D., and Chance, M. R. (2004) Principles of RNA compaction: insights from the equilibrium folding pathway of the P4-P6 RNA domain in monovalent cations. *J. Mol. Biol.* **343**, 1195–1206
6. Buchmueller, K. L., Webb, A. E., Richardson, D. A., and Weeks, K. M. (2000) A collapsed non-native RNA folding state. *Nat. Struct. Biol.* **7**, 362–366
7. Lipfert, J., Chu, V. B., Bai, Y., Herschlag, D., and Doniach, S. (2007) Low-resolution models for nucleic acids from small-angle X-ray scattering with applications to electrostatic modeling. *J Appl Crystallogr* **40**, s229–s234
8. Rambo, R. P., and Tainer, J. A. (2010) Improving small-angle X-ray scattering data for structural analyses of the RNA world. *RNA* **16**, 638–646
9. Lipfert, J., Herschlag, D., and Doniach, S. (2009) Riboswitch conformations revealed by small-angle X-ray scattering. *Methods Mol. Biol.* **540**, 141–159
10. Orelle, C., Carlson, E. D., Szal, T., Florin, T., Jewett, M. C., and Mankin, A. S. (2015) Protein synthesis by ribosomes with tethered subunits. *Nature* **524**, 119–124
11. Klein, D. J., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J.* **20**, 4214–4221
12. Szewczak, A. A., and Cech, T. R. (1997) An RNA internal loop acts as a hinge to facilitate ribozyme folding and catalysis. *RNA* **3**, 838–849
13. Hermann, T., and Patel, D. J. (1999) Stitching together RNA tertiary architectures. *J. Mol. Biol.* **294**, 829–849
14. Kiliszek, A., Kierzek, R., Krzyzosiak, W. J., and Rypniewski, W. (2010) Atomic resolution structure of CAG RNA repeats: structural insights and implications for the trinucleotide repeat expansion diseases. *Nucleic Acids Res.* **38**, 8370–8376
15. Tawani, A., and Kumar, A. (2015) Structural Insights Reveal the Dynamics of the Repeating r(CAG) Transcript Found in Huntington's Disease (HD) and Spinocerebellar Ataxias (SCAs). *PLoS One* **10**, e0131788
16. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273**, 1678–1685

17. Eiler, D., Wang, J., and Steitz, T. A. (2014) Structural basis for the fast self-cleavage reaction catalyzed by the twister ribozyme. *Proc. Natl. Acad. Sci. USA* **111**, 13028–13033

18. Tian, S., Yesselman, J. D., Cordero, P., and Das, R. (2015) Primerize: automated primer assembly for transcribing non-coding RNA domains. *Nucleic Acids Res.* **43**, W522-6

19. Cordero, P., Kladwang, W., VanLang, C. C., and Das, R. (2012) Quantitative dimethyl sulfate mapping for automated RNA secondary structure inference. *Biochemistry* **51**, 7037–7039

20. Yoon, S., Kim, J., Hum, J., Kim, H., Park, S., Kladwang, W., and Das, R. (2011) HiTRACE: high-throughput robust analysis for capillary electrophoresis. *Bioinformatics* **27**, 1798–1805

21. Das, R. (2018) LIkelihood-based Fits of Folding Transitions (LIFFT) for Biomolecule Mapping Data. *BioRxiv* 294041

22. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., and Sali, A. (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105**, 962–974

23. Franke, D., and Svergun, D. I. (2009) DAMMIF , a program for rapidab-initio shape determination in small-angle scattering. *J Appl Crystallogr* **42**, 342–346

24. Kabsch, W. (2010) XDS. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 125–132

25. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **67**, 235–242

26. Chou, F.-C., Echols, N., Terwilliger, T. C., and Das, R. (2016) RNA Structure Refinement Using the ERRASER-Phenix Pipeline. *Methods Mol. Biol.* **1320**, 269–282

27. Cheng, C. Y., Chou, F.-C., and Das, R. (2015) Modeling complex RNA tertiary folds with Rosetta. *Meth. Enzymol.* **553**, 35–64

28. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345

29. Jewett, M. C., Fritz, B. R., Timmerman, L. E., and Church, G. M. (2013) In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Mol. Syst. Biol.* **9**, 678

30. Fritz, B. R., Jamil, O. K., and Jewett, M. C. (2015) Implications of macromolecular crowding and reducing conditions for in vitro ribosome construction. *Nucleic Acids Res.* **43**, 4774–4784

31. Watters, K. E., Abbott, T. R., and Lucks, J. B. (2016) Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq. *Nucleic Acids Res.* **44**, e12

32. Yu, A. M., Evans, M. E., and Lucks, J. B. (2018) Estimating RNA structure chemical

probing reactivities from reverse transcriptase stops and mutations. *BioRxiv* 292532

33. Kasprzak, W., Bindewald, E., Kim, T.-J., Jaeger, L., and Shapiro, B. A. (2011) Use of RNA structure flexibility data in nanostructure modeling. *Methods* **54**, 239–250

34. Afonin, K. A., Kasprzak, W., Bindewald, E., Puppala, P. S., Diehl, A. R., Hall, K. T., Kim, T. J., Zimmermann, M. T., Jernigan, R. L., Jaeger, L., and Shapiro, B. A. (2014) Computational and experimental characterization of RNA cubic nanoscaffolds. *Methods* **67**, 256–265

35. Yesselman, J. D., Denny, S. K., Bisaria, N., Herschlag, D., Greenleaf, W. J., and Das, R. (2019) RNA tertiary structure energetics predicted by an ensemble model of the RNA double helix. *Proc. Natl. Acad. Sci. U.S.A.,* in proofs.

36. Song, W., Strack, R. L., Svensen, N., and Jaffrey, S. R. (2014) Plug-and-play fluorophores extend the spectral properties of Spinach. *J. Am. Chem. Soc.* **136**, 1198–1201

37. Paige, J. S., Wu, K. Y., and Jaffrey, S. R. (2011) RNA mimics of green fluorescent protein. *Science* **333**, 642–646

38. Desai, A., Murray, A., Mitchison, T. J., and Walczak, C. E. (1999) The use of Xenopus egg extracts to study mitotic spindle assembly and function in vitro. *Methods Cell Biol.* **61**, 385–412

39. Lucks, J. B., Qi, L., Mutalik, V. K., Wang, D., and Arkin, A. P. (2011) Versatile RNA-sensing transcriptional regulators for engineering genetic networks. *Proc. Natl. Acad. Sci. USA* **108**, 8617–8622

40. Dibrov, S. M., McLean, J., Parsons, J., and Hermann, T. (2011) Self-assembling RNA square. *Proc. Natl. Acad. Sci. USA* **108**, 6405–6408

41. Geary, C., Chworos, A., and Jaeger, L. (2011) Promoting RNA helical stacking via A-minor junctions. *Nucleic Acids Res.* **39**, 1066–1080

42. Geary, C., Chworos, A., Verzemnieks, E., Voss, N. R., and Jaeger, L. (2017) Composing RNA Nanostructures from a Syntax of RNA Structural Modules. *Nano Lett.* **17**, 7095–7101

43. Huang, L., and Lilley, D. M. J. (2016) A quasi-cyclic RNA nano-scale molecular object constructed using kink turns. *Nanoscale* **8**, 15189–15195

44. Zhang, H., Endrizzi, J. A., Shu, Y., Haque, F., Sauter, C., Shlyakhtenko, L. S., Lyubchenko, Y., Guo, P., and Chi, Y.-I. (2013) Crystal structure of 3WJ core revealing divalent ion-promoted thermostability and assembly of the Phi29 hexameric motor pRNA. *RNA* **19**, 1226–1237

45. Bindewald, E., Afonin, K., Jaeger, L., and Shapiro, B. A. (2011) Multistrand RNA secondary structure prediction and nanostructure design including pseudoknots. *ACS Nano* **5**, 9542–9551

46. Fried, S. D., Schmied, W. H., Uttamapinant, C., and Chin, J. W. (2015) Ribosome subunit stapling for orthogonal translation in e. coli. *Angew. Chem. Int. Ed. Engl.* **54**, 12791–12794

47. Silverman, A.D., Kelley-Loughnane, Lucks, J.B., and Jewett, M.C. (2019) Deconstructing cell-free extract preparation for *in vitro* activation of transcriptional genetic circuitry. *ACS Synth. Bio.* **8**, 403-414.