Title: Computational Design of Asymmetric Three-dimensional RNA Structures and Machines.

Authors: Joseph D. Yesselman¹, Daniel Eiler², Erik D. Carlson^{3,4, 5}, Alexandra N. Ooms⁶, Wipapat Kladwang¹, Xuesong Shi¹, David A. Costantino², Daniel Herschlag^{1,7, 8}, Michael C. Jewett^{3,4,5}, Jeffrey S. Kieft², Rhiju Das^{1,9,*}

¹ Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305, USA.

² Department of Biochemistry and Molecular Genetics, University of Colorado Denver School of Medicine, Aurora, Colorado 80045, USA.

³ Department of Chemical and Biological Engineering, Northwestern University, 2145 Sheridan Road, Tech E-136, Evanston, Illinois 60208, USA.

⁴ Chemistry of Life Processes Institute, Northwestern University, 2170 Campus Drive, Evanston, Illinois 60208, USA.

⁴ Center for Synthetic Biology, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA.

⁶ Department of Cancer Genetics & Genomics, Stanford University School of Medicine, Stanford, CA 94305, USA.

⁷ Department of Chemistry, Stanford University School of Medicine, Stanford, CA 94305, USA.

⁸ Stanford ChEM-H (Chemistry, Engineering, and Medicine for Human Health), Stanford University, Stanford, California 94305, USA.

⁹ Department of Physics, Stanford University, Stanford, CA 94305, USA.

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

Abstract:

The emerging field of RNA nanotechnology seeks to create nanoscale 3D machines by repurposing natural RNA modules, but successes have been limited to symmetric assemblies of single repeating motifs. We present RNAMake, a suite that automates design of RNA molecules with complex 3D folds. We first challenged RNAMake with the paradigmatic problem of aligning a tetraloop and sequence-distal receptor, previously only solved via symmetry. Single-nucleotide-resolution chemical mapping, native gel electrophoresis, and solution x-ray scattering confirmed that 11 of the 16 ‘miniTTR’ designs successfully achieved clothespin-like folds. A 2.55 Å diffraction-resolution crystal structure of one design verified formation of the target asymmetric nanostructure, with large sections achieving near-atomic accuracy (< 2.0 Å). Finally, RNAMake designed asymmetric segments to tether the 16S and 23S rRNAs together into a synthetic single-stranded ribosome that remains uncleaved by ribonucleases and supports life in *Escherichia coli*, a challenge previously requiring several rounds of trial-and-error.

Introduction:

RNA-based nanotechnology is an emerging field that harnesses RNA's unique structural properties. Perhaps more so than for other biomolecules, RNA's tertiary structure is composed of discrete and recurring components known as tertiary 'motifs' (1–4). Along with the helices that they interconnect, many of these structural motifs appear highly modular; that is, each motif folds into a common, well-defined 3D structure in a large range of contexts (5–7). By exploiting symmetry, motif repetition, and expert modeling, these motifs have been assembled into rationally designed structures, including polygons, cubes, and sheets (8–14). Nevertheless, the RNA molecules generated to date are mostly symmetrical and structurally simple, lacking the complexity observed in natural RNA machines, which contain a larger repertoire of distinct interacting motifs. The importance of this structural diversity is implied by the wide range of functionality observed in naturally occurring RNAs (15–18). Rationally designing RNAs with increased functional diversity will require methods to generate asymmetric folds with a larger set of motifs (19).

Including additional structural motifs in rational 3D RNA design is challenging for two reasons. First, introducing new motifs demands rigorous testing for modular behavior, i.e., determining if the motifs maintain their 3D folds when installed into new RNAs. Second, with current methods and software, the rational 3D design of RNA can be laborious, even with a limited set of motifs, as the designer is required to intuit the correct helical separation between motifs, ensure that tertiary contacts are in the proper orientation, and guarantee the correct formation of secondary structure.

Here, we address these issues by introducing a method to design continuous chains of RNA motifs that twist and translate between any two desired helical endpoints that can

harbor functional elements or tertiary contacts (Figure 1). We then demonstrate two proofs of concept using design problems that have been challenging in prior studies. First, to allow for stringent structural tests, we tackle the classic problem of designing an RNA that aligns the two parts of the tetraloop/tetraloop-receptor (TTR) so that they can form the TTR tertiary contact upon RNA chain folding (Figure 1b). This 3D design challenge was the first problem solved in RNA nanotechnology (7), but all solutions to date have either relied on symmetric design principles to guide dimer, fiber, or higher-order assembly (20) or have resulted in a tertiary structure with minimal stability (21). In addition, these prior solutions required manual modeling by experts. Second, we demonstrate that computational 3D design can aid in difficult alignment challenges encountered in the development of functional RNA machines (22–24). For example, design challenges involving ribosomal subunit tethering (22,23) previously proved difficult to address using manual modeling, as they require complex RNA segments to correctly orient RNA nanostructures relative to one another in 3D space. To demonstrate the broad application of computational 3D design, we build and test novel RNA segments that tether and align the small and large subunits of the ribosome together into a single RNA strand *in vivo*; this design problem previously required more than a year to solve using manual design and trial-and-error refinement based on *in vivo* assays (25).

Both the new minimal TTR (miniTTR) and ribosome tether constructs were generated using a fully automated toolkit for motif-based design that we call RNAMake. This toolkit contains a library of hundreds of unique motifs and a path-finding algorithm to connect any two points in 3D space at a desired relative orientation using an RNA segment. Our tests demonstrate that purely computational design can make use of multiple structurally diverse motifs within a single strand of RNAs to generate asymmetric, monomeric structures that fold correctly *in vitro* and to reengineer machines that support life.

Furthermore, our results confirm the modularity of dozens of novel non-canonical motifs not previously utilized in design.

Results:

Generation and Use of a Curated Structural Motif Library.

To expand the library of available motifs for rational design, we collected motifs from all unique RNA structures that had been deposited in a crystallographic database (see Methods). The final set contained two-way junctions, higher-order junctions, hairpins, and tertiary contacts with 461, 61, 290, and 89 unique motifs respectively. To efficiently sample over the large set of motif modules in our library, we developed a path-finding search algorithm. This algorithm models canonical helical segments of 1 to 22 base pairs that interconnect noncanonical structural motifs (see Methods). The canonical helical segments are idealized and sequence invariant (26), but after the completion of 3D structural design, they are filled in with sequences that best match the target secondary structure and minimize alternative secondary structures. This optimization of the helical sequences is carried out with EteRNAbot, developed previously through internet-scale crowdsourcing and tested through chemical mapping (27).

RNAMake Automates RNA 3D Design That Includes Diverse Motifs.

The assessment of the structural accuracy of RNAMake's design ability is the first step toward using it as a reliable tool to generate functional RNA machines. We assessed RNAMake's ability to generate structurally accurate 3D RNA structures by challenging it to align the tetraloop and receptor parts of the TTR in their bound conformation using a designed RNA segment. To generate these miniTTR constructs, we first extracted the coordinates from the X-ray crystal structure of the TTR from the P4-P6 domain of the *Tetrahymena* ribozyme (residues 146-157, 221-246, and 228-252 from PDB 1GID) (28).

Second, we used RNAMake to build structural segments composed of two-way junctions and helices spanning the last base pair of the hairpin (A146-U157) to base pair U221-A252 of the tetraloop-receptor, connecting the TTR into a single continuous strand (Figure 1b). Of 200,000 RNA segments generated, sixteen were selected based on two criteria: 1) fewest motifs per solution (three unique tertiary motifs) and 2) tightest alignment of the two segments of the TTR to their target spatial and rotational orientations. These computational designs ranged from 75 to 102 nucleotides in size (for full sequences, see SI Table 1), significantly shorter than the 157 nucleotides of the natural P4-P6 domain RNA.

The miniTTR designs included 23 unique motifs (Figure 1a and 1c) encompassing three distinct motif categories. First, each design included 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 (29), J5/5a from the P4-P6 domain (30), or an S-turn (31) (see, e.g., TWOWAY.1S72.20, TWOWAY.1GID.2 and TWOWAY.3BNQ.6 in Figure 1a; SI Table 2). Second, each design included at least one near-helical motif that was approximately, but not exactly, 'straight'; an example is a set of three consecutive non-canonical base pairs (see, e.g., TWOWAY.1S72.51 in Figure 1a). 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 Figure 1a). Previous work on RNA design, which was based on manual modeling by RNA experts, did not test these latter two categories of motifs, which are difficult to model without automatic tools, but appear to be necessary in natural RNAs for generating asymmetric structures in which small refinements of helical twists are required.

Chemical Mapping of MiniTTRs to Probe TTR Formation.

To probe the structures of the miniTTR designs generated by RNAMake, we first performed selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) and quantitative chemical mapping with dimethyl sulfate (DMS) (SI Figure 1) (32). 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, SI Table 3). These data supported the formation of the expected secondary structures for all miniTTR designs (SI Figure 1).

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+} (33–35). As a control for the unfolded state, the DMS reactivities of the tetraloop and tetraloop-receptor adenines of the TTR in the P4-P6 domain (A248, A151, A152 and A153) were measured and found to be 1.27, 0.72, 0.70, and 0.90, respectively, in 50 mM Na-HEPES, pH 8.0 buffer without Mg^{2+} . Here and below, the values are normalized to the reactivity of the reference hairpin loops that flank each design (32). 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 2d). 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 2d). 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.

To more precisely assess and compare folding stabilities, we determined how much Mg^{2+} was required to fold each of the miniTTRs that passed the screen described above (36). The miniTTR designs exhibited midpoints ranging from 0.12 to 7.0 mM Mg^{2+} (SI Table 4). The most stable designs were miniTTR 2 (Figure 2e and 2f) and miniTTR 6 (Figure 2e), with midpoints of 0.94 ± 0.39 and 0.12 ± 0.03 mM Mg^{2+} respectively. Their midpoints were similar to the Mg^{2+} midpoint measured for P4-P6 (0.38 ± 0.06 mM, Figure 2e). These observations suggested that at least two of the 16 automated designs achieved stabilities similar to that of the natural P4-P6 RNA fold, which is stabilized by the same tertiary contact, as well as by an additional tertiary contact (a metal core/metal core receptor) (28).

Native Gel Mobility Shifts Due to TTR Formation.

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 (37) and is predicted to reduce the RNA's mobility in non-denaturing polyacrylamide gel electrophoresis, as observed for the P4-P6 domain (38). Of the 16 miniTTR constructs tested, 12 designs displayed mobility shifts consistent with the formation of the TTR tertiary contact (Figure 2a-c, also see Table 1 for a comparison to chemical mapping). 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 exhibited a wide, slow-mobility band, suggesting multimer formation. 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.

Small-angle X-ray Scattering Suggests Predominantly Stable Monomers.

We carried out small-angle X-ray scattering (SAXS) measurements on our most stable designs, miniTTR 2 and miniTTR 6 (Figure 3). The observed scattering profiles of miniTTR 2 and miniTTR 6 agreed reasonably well with the profiles predicted from their corresponding RNAMake models, with $\chi^2/\text{d.o.f.}$ values of 13 and 27, respectively (Figure 3a). These values are near values of 2–8 obtained from comparisons between predictions of RNA crystallographic models and scattering data (39–41), suggesting similar overall folds but with some local differences, an expectation confirmed for miniTT6 by X-ray crystallography (see next section).

To compare the coarse shapes of miniTTR designs 2 and 6 in solution to the predicted shapes, we generated low-resolution 3D representations of the experimental X-ray scattering profiles through the automated arrangement of ‘dummy atoms’ guided solely by the data (see Methods). For both constructs, the overall shapes of these unbiased shape reconstructions agreed well with the global folds of the RNAMake models (Figure 3b and 3c). Finally, size exclusion chromatography monitored by SAXS confirmed that <15% and <5% of miniTTRs 2 and 6, respectively, occur in higher-order structures (SI Figure 2). These results confirm that both the tested miniTTRs are primarily stable monomers, even at the RNA concentrations used for SAXS (>1 μM), again consistent with our RNAMake models.

X-ray Crystallographic Structure Demonstrates Near-atomic Accuracy of Design.

As the most stringent test of the RNAMake design algorithm, we sought to compare our modeled structures to atomic-resolution experimental structures. After targeting several of the miniTTR constructs for crystallization, we were able to obtain crystals of miniTTR 6 that diffracted at 2.55 Å resolution (I/σ of 1.0) (Table 2), and these data allowed for an atomic-level reconstruction. The SAXS profile predicted from these coordinates agreed with the SAXS data ($\chi^2/\text{d.o.f.} = 9$, Figure 3a), supporting the assumption that the crystal captured the solution conformation. The global fit between the crystallographic model and the structure predicted by RNAMake gave a heavy-atom RMSD of 4.2 Å. This value was somewhat less accurate than the heavy-atom RMSD of 3.4 ± 1.0 Å estimated from modeling the RNA's thermal fluctuations (SI Figures 3 and 4), and the main source of deviation could be traced to a single feature: a triple mismatch (TWOWAY.1S72.62) (Figure 4e). The 3D model of this triple mismatch was derived from a ribosome crystal structure in which multiple bases flipped out from the motif to form A-minor interactions with flanking helices. In the context of the miniTTR, this motif did not flip out the bases; instead, it formed multiple consecutive non-canonical base pairs with high B-factors, presumably due to the absence of tertiary interaction partners. However, the other motifs of the design achieved near-atomic accuracy. Aligning the TTR from the model to the crystal structure (Figure 4b) gave a heavy-atom RMSD of 0.45 Å. Another motif selected by RNAMake in miniTTR 6 was TWOWAY.1S72.20, a kink-turn variant from the archaeal 50S ribosomal subunit (42). The RNAMake model agreed with the crystal structure with a heavy-atom RMSD of 2.0 Å. The overall topologies and interactions were nearly identical, and the flipped-out residues were also in the same orientation (Figure 4c). The last motif came from a viral internal ribosomal entry site domain incorporated into an RNA nanosquare (11). This motif exhibited a heavy-atom RMSD of 1.28 Å between the RNAMake model and the crystal structure (Figure 4d). The only difference involved two unpaired residues that were stacked in the RNAMake model but

flipped out in the miniTTR 6 crystal structure to form water-mediated contacts with other symmetry-related copies of the RNA in the crystal lattice. These changes were minor and did not affect the relative positioning or orientation of the helices connected by the motif (Figure 4d).

Tethered Ribosomes Designed Using RNAMake Support *Escherichia coli* Growth.

After RNAMake's success in the miniTTR structure design challenge, we sought to assess the algorithm's ability to design molecules for function in a complex biological context. As a proof of concept, we designed RNA segments to tether the ribosome. The ribosome is a ribonucleoprotein machine dominated by two extensive RNA subunits. Previously we described the first successful construction of a ribosome with covalently tethered subunits that could support bacterial growth even in the absence of wild type ribosomes (23). In that work, several iterations of design were necessary to identify RNA tethers that support life in *E. coli* and were not cleaved by ribonucleases *in vivo*.

Here, we used RNAMake to computationally design RNA segments that tether the 50S and 30S ribosomal subunits together (SI Figure 5). The tethers were built between the H101 helix on the 50S subunit and the h44 helix on the 30S subunit (H101_h44_Tether), similar to the previous Ribo-T design (23). We designed nine unique constructs with RNAMake (Figure 5b) and successfully cloned seven. Each design utilizes between four and five motifs. The overlap between the motifs utilized in the miniTTRs and the H101_h44_Tether designs is low. Of the fourteen motifs used in the H101_h44_Tether designs, only three were also in the miniTTRs (TWO WAY.1NUV.3, TWO WAY.1S72.29 and TWO WAY.1S72.39). Nevertheless, like the miniTTRs, the H101_h44_Tether designs contained the same three distinct categories of motifs. First, all designs contained two large (>10 residue) bends (see, e.g., TWO WAY.1S72.29,

TWOWAY.1S72.39 and TWOWAY.1S72.42 in Figure 5a). This characteristic differed from that of the miniTTR designs, where it was possible to only have one bend. Furthermore, unlike the miniTTRs, each design contained at least one near-helical motif (see, e.g., TWOWAY.1NUV.3, TWOWAY.2VQE.46, TWOWAY.3LOA.0 in Figure 5a) and one small motif (see, e.g., TWOWAY.1S72.100, TWOWAY.1J9H.3 in Figure 5a).

To test RNAMake's accuracy in designing functional RNA, each H101_h44_Tether design (Figure 5a) was cloned into the pRibo-T plasmid (23) and used to replace the wild-type ribosomal rRNA plasmid in the SQ171fg strain (23). 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 H101_h44_Tether design plate. 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 (Figure 5c). After 5 days of growth, total RNA was extracted and analyzed by gel electrophoresis. All clones (except clone 4) showed a clean Ribo-T-sized band and no detectable wild-type 23S and 16S rRNA bands (Figure 5d), indicating the formation of stable tethered ribosomes as previously demonstrated (23). Plasmids were extracted from these colonies, and their sequences were each confirmed as the correct H101_h44_Tether design 4 plasmid. The growth curves of the 7 successful clones (excluding clone 4) exhibited doubling times of 1.5 ± 0.8 days and a maximum OD₆₀₀ of 0.7 ± 0.3 . While slower than wild type *E. coli*, the first successful version of Ribo-T was also slower growing before mutational optimization (23). Taken together, these data demonstrate that tethered ribosomes with automatically designed tethers using RNAMake can fully support *E. coli* growth in the absence of wild-type ribosomes.

Discussion:

Automated 3D Design with a High Success Rate.

We have demonstrated that RNAMake can automate the design of new RNA structures, requiring only the start and end points of the desired structure. Taking a classic and common problem in RNA 3D design, the translational and rotational alignment of partner segments to form a target tertiary contact, RNAMake generated 11 out of 16 successful designs that passed our chemical mapping and native gel mobility tests. Furthermore, miniTTR 2 and 6 achieved stability at low $[Mg^{2+}]$, similar to or better than the natural TTR-containing P4-P6, despite using fewer nucleotides (97 and 94 nts respectively, compared to the 157 nts in the P4-P6 domain) and containing fewer tertiary contacts (the P4-P6 domain contains a metal-core/metal-core receptor in addition to a TTR). One possible reason why these two miniTTR designs were more stable than the P4-P6 domain at lower $[Mg^{2+}]$ is that the motifs in the miniTTR designs were chosen rationally to achieve a given orientation, while the motifs in the P4-P6 domain were naturally selected to support the broader folding landscape and multi-step functional cycle of the larger *Tetrahymena* self-splicing group I intron. Interestingly, miniTTRs have shallower Mg^{2+} dependence than the P4-P6 RNA (Figure 2); designing sharper Mg^{2+} dependences, which are characteristic of many natural tertiary RNA structures, will be an important future challenge. Our results suggest that it is possible to computationally generate RNAs that are just as or more stable than naturally occurring RNAs, analogous to results in protein design (43,44).

Expanding the Experimentally Vetted ‘Building Blocks’ of RNA 3D Design.

In RNA nanotechnology efforts to date, utilizing a previously untested motif in 3D design has required trial-and-error testing to verify that the RNA remains in a well-defined 3D structure in new contexts. Due to the amount of human and experimental effort required, the inclusion of novel motifs has been slow, with researchers generally drawing from a

limited set of motifs previously verified to be modular and reliable, and these motifs are typically deployed one at a time within highly repetitive arrangements rather than being combined into complex asymmetric combinations. This process has generated simple RNAs that lack the complexity of natural RNA machines.

In our miniTTR study, automated computational design resulted in RNAs containing a total of 23 motifs, 20 of which had not been used previously in the 3D rational design of RNA. The motifs ranged from a single bulged A to complex motifs such as kink-turn variants and S-turns (Figure 1a, SI Table 2). Of these new motifs, a large majority (18) appeared in designs validated by our biochemical experiments, but two others gave informative discrepancies. One motif, an A-A mismatch (TWOWAY.1S72.49, Figure 1a), appeared in all four miniTTR designs that did not form TTRs (3, 4, 8, 13). Unlike other A-A mismatches that form a non-canonical base pair (45,46), 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, Figure 1a). 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 4e). 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, which, in fact, was our most stable design, as assessed by Mg^{2+} titration. This result suggests that residual uncertainties in RNAMake's motif library will not preclude the consistent design of asymmetric structures at nucleotide resolution. Our experimental tests of the remaining 18 modular motifs,

including two verified at atomic resolution, expand both the number and diversity of motifs available for RNA nanotechnology.

RNAMake Accelerates the Design of *in vivo* RNA Nano-machines.

The generation of a hybrid ribosomal RNA containing the sequences of both the small and large rRNA subunit sequences linked together by an engineered two-stranded tether significantly expanded the promise of synthetic biology. Yet, without the assistance of 3D modeling software such as RNAMake, the time invested in its development was substantial, requiring over a year of experimenting with different tether sequences (25). Utilizing RNAMake, we greatly simplified the design process. In our first attempt, we generated nine unique tether designs and successfully cloned seven. Of these, one design was able to support *E. coli* growth in the absence of wild-type ribosomes, while remaining uncut *in vivo*. Although its growth rate was slow (doubling every 1.5 ± 0.8 days), achieving success *in vivo* on the first attempt suggests that RNAMake can significantly accelerate design of RNA machines. As in the case of the first Ribo-T, we can now use both rational design and selection techniques to further optimize this tether for the multiple states in the ribosome's functional cycle. RNAMake enables the computational design of diverse molecules to solve new problems, raising the prospect of increasing RNA design complexity beyond what has been achieved with manual methods.

Acknowledgments:

We thank S. Bonilla for assistance in performing native gel assays and A. Watkins for discussions about ribosome tether design. This work was supported by National Institutes of Health grant R01 GM100953 (to R.D.), Ruth L. Kirschstein National Research Service Award Postdoctoral Fellowships GM112294 (to J.D.Y.) and

GM100953 (to D.E.), National Institute of General Medical Sciences grant R35
GM118070 (to J.S.K.), Army Research Office W911NF-16-1-0372 (to M.C.J.), National
Science Foundation grant MCB-1716766 (to M.C.J.), the David and Lucile Packard
Foundation (to M.C.J.) and National Institutes of Health grant P01 GM066275 (to D.H.).

Contributions:

R.D. and J.D.Y. conceived of the study. J.D.Y. developed RNAMake and generated the
models and sequences used throughout the study. A.N.O. and W.K. performed the
experimental chemical mapping, titrations and native gel assays. J.D.Y. analyzed and
processed the results from these experiments. X.S. performed the small-angle X-ray
scattering on miniTTRs 2 and 6. D.E. solved the crystal structure of miniTTR 6 and was
assisted by D.C. and J.S.K. in preparing the RNA and in the analysis. E.D.C. and M.C.J.
made and tested the RNAMake-designed Ribo-T tethers. J.D.Y. and R.D. wrote the
paper, with participation by all authors.

Methods:

Software Availability.

All software and source code used in this work are freely available for non-commercial
use. RNAMake software and documentation are at
<https://github.com/jyesselm/RNAMake>. EteRNAbot is available at
<https://github.com/EteRNAgame/EteRNABot>.

Building the Motif Library.

To build a curated motif library of all RNA structural components, we obtained the set of
non-redundant RNA crystal structures managed by the Leontis and Zirbel groups (18)
(version 1.45: [http://rna.bgsu.edu/rna3dhub/nrlist/release/1.45\(weblink\)](http://rna.bgsu.edu/rna3dhub/nrlist/release/1.45(weblink))). 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) (47) 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 alignment between motifs:

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

In addition to the motifs derived from the PDB, we also utilized the make-na web server (<http://structure.usc.edu/make-na/server.html>) to generate idealized helices of between 2 and 22 base pairs in length (26). All motifs in these generated libraries are bundled with RNAMake and are grouped together by type (junctions, hairpins, etc.) in sqlite3 databases in the directory `RNAMake/rnamake/resources/motif_libraries/(weblink)`.

Automatically Building New RNA Segments.

RNAMake seeks a path for RNA helices and noncanonical 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 (5,11) and is composed of two terms:

$$f(path) = h(path) + g(path) \quad (1)$$

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 defining the rotational orientation of each base pair (48):

$$h(path) = |\vec{d}_1 - \vec{d}_2| + W(|\vec{d}_1 - \vec{d}_2|) \sum_i^3 \sum_j^3 abs(R_{1ij} - R_{2ij}) \quad (2)$$

Here, $W(d)$ is:

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

The weight $W(d)$ reduces the importance of the current base pair and the target base pair with 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 in (2) and (3), including more standard rotationally invariant metrics to define rotation matrix differences (49) or base-

pair-to-base-pair RMSDs based on quaternions (50), 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 comprising the path at each stage of the calculation

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

where S_{ss} is a secondary structure score for all the motifs and helices in the path. This S_{ss} favors longer helices as well as motifs with frequently recurring base pairs, as follows. All base pairs found in the RNA motif are scored based on their relative occurrences in all high-resolution crystal structures; all unpaired residues receive a penalty, and Watson-Crick base pairs receive an additional bonus score (values are given in SI Table 5). 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 helix length).

The search adds motifs and helices to the path in a depth-first manner, while the total cost function $f(path)$ decreases, back-tracking if $f(path)$ increases. Any solutions with $h(path)$ less than 5, i.e., 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, finding most solutions in a few seconds. For each

solution, we then used EteRNAbot, a secondary structure optimization algorithm that has undergone extensive empirical tests (27), to fill in helix sequences.

An example RNAMake command line 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.

RNA Structural Probing and miniTTR Data Analysis.

DNA oligonucleotides were designed with Primerize (51), ordered from IDT (Integrated DNA Technologies) and used to generate double-stranded DNA templates using PCR assembly. DNA template and RNA transcript preparation and quality checks were carried out as previously described (32,52).

Chemical mapping (DMS and SHAPE) was performed as previously described (32,52). Briefly, modification reactions were performed in a 20 μ L volume containing 1.2 pmol of RNA and 50 mM Na-HEPES (pH 8.0). Before the chemical modifier was added, the RNA was heated to 90 $^{\circ}$ 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 μ L 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 μ L.
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 ddH₂O 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; it 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 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, available as a MATLAB
toolbox at <https://github.com/hitrace> (53) and a web server at <http://hitrace.org> (54).

Electrophoretic traces were aligned and baseline subtracted using linear and nonlinear
alignment routines as previously described (53). 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 (32).

To estimate Mg²⁺ titration midpoints, the relative protection values (f_i^j) for each residue j
in the TTR at each Mg²⁺ concentration i were calculated. The quantitative DMS reactivity
of the folded and unfolded state of each TTR residue was taken from P4-P6.

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

521

522 These values were then fit to the Hill equation:

$$f_j^i = \frac{([Mg^{2+}]_i/K)^n}{1 + ([Mg^{2+}]_i/K)^n}$$

523

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

526

527 **Native Gel Shift Assays.**

528 Native gel shift assays were conducted using a BioRad Criterion™ Cell gel cassette.

529 Gels were cast using 11% acrylamide and 1X THEM buffer (pH 7.2) (10 mM MgCl₂) and

530 polymerized by adding 10% ammonium persulfate (300 μL) and TEMED (30 μL) to 30

531 mL of gel mix. Following polymerization, the gel apparatus was set up in a 4 °C cold

532 room and fully immersed in an ice bath until the gel and buffer apparatus were cooled to

533 approximately 4 °C. Then, RNA constructs were prepared for folding by incubating them

534 in folding buffer consisting of 100 mM Na-HEPES (pH 7.5), 10 mM MgCl₂, and 0.5 mM

535 spermidine. 100 μg of RNA was mixed with folding buffer in the absence of

536 spermidine and incubated at 65 °C for 3 min. The solution was cooled at room

537 temperature for 10 min, and then 0.5 mM spermidine was added. The RNAs were then

538 vortexed for 10 seconds and centrifuged at 4,000 rpm for 10 min. Immediately following

539 folding, loading dye was added to the RNA solutions, and the samples were directly

540 loaded onto the cooled gel apparatus. The gel was run at 15 watts for 1.5 hours, and

541 the temperature of the gel apparatus was strictly monitored to avoid overheating. After

542 electrophoresis, the gel was removed from the cassette, carefully placed in a glass pan

543 and incubated with 150 mL of Stains-All 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.

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 MgCl₂ 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 3a) using FoXS (55). 3D bead models of miniTTR 2 and 6 were generated using DAMMIF and DAMMIN (56,57) and overlaid with their corresponding RNAMake models in PyMOL.

Materials and Methods for X-ray Crystallography.

RNA used for crystallization was transcribed with T7 RNA polymerase from PCR-generated double-stranded DNA templates as described in (58). 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.

MiniTTR Crystallization.

Purified miniTTR 6 RNA diluted in buffer A (30 mM HEPES (pH 7.5), 20 mM MgCl₂, 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 MgCl₂, 2 mM cobalt hexammine, and 40% 2-methyl-2,4-pentenediol (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) (59). The scaled data were handled using Collaborative Computational Project programs (60).

Structure Determination and Refinement.

The initial structural determination of the miniTTR 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 (60). 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 (61). The models were built using COOT (62) and refined using REFMAC5 and PHENIX (60). The final model was refined in REFMAC5 (60), and the overall R_{work} and R_{free} were refined to 22.9% and 27.7%, 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 Table 2. The crystal structures of miniTTR 6 have been deposited in PDB, ID 5VOQ. All structural figures were prepared using PyMOL (<http://www.pymol.org/>).

Design and Cloning of Novel Tethers for a Ribosome with Tethered Subunits.

The designed tethers (SI Table 6) were cloned into plasmid pRibo-T-A2058G (23) using the primers in SI Table 7. The backbone was generated for each design using forward (f) and reverse (r) primer pairs (noted with “bb”) in SI Table 7 in separate PCR reactions using plasmid pRibo-T as a template (23), 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 SI Table 7), 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 (63). 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 Ribo-T.

Each purified plasmid (1 µL) was separately transformed into electrocompetent SQ171fg cells containing pCSacB (23). 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.

646 After 48 hours with no visible colonies, the plates were replica plated onto fresh LB agar
647 plates supplemented with 50 µg/mL carbenicillin, 5% sucrose and 1 mg/mL erythromycin
648 and incubated at 37 °C. After 72 additional hours, colonies appeared on the plate
649 containing H101_h44_Tether design 4. Eight colonies were streaked onto LB agar
650 supplemented with 50 µg/mL carbenicillin and 1 mg/mL erythromycin and LB agar
651 supplemented with 30 µg/mL kanamycin (to confirm loss of the pCSacB plasmid) and
652 were also used to inoculate 5 mL of LB supplemented with 50 µg/mL carbenicillin and 1
653 mg/mL erythromycin. Plates were incubated at 37 °C, and cultures were incubated at 37
654 °C with shaking. The OD₆₀₀ of the cultures was tracked to generate growth curves
655 (Biochrom Libra S4 spectrophotometer). After 5 days at 37 °C, total RNA was extracted
656 using an RNA extraction kit from Qiagen. Total RNA was analyzed by gel
657 electrophoresis on a 1% agarose gel with GelRed. Total plasmid was extracted from
658 saturated 5 mL cultures with an E.Z.N.A. miniprep kit (Omega) and sequenced to
659 confirm the correct H101_h44_Tether design 4 sequence.

Figures:

Figure 1: a) All unique motifs included in the miniTTR designs (See SI Table 2 for common names where available). b) Demonstration of RNAMake's design algorithm, which builds an RNA path via the successive addition of motifs and helices from the starting base pair to the ending base pair. c) Ensemble of all RNAMake miniTTR designs. Each is colored according to the motifs (panel (a)) that were used to generate it.

Figure 2: a-c) Native gel assay for each miniTTR design. Each design contains both a wild-type RNA with a GAAA tetraloop and a UUCG mutant that is unable to form the TTR interaction. d) Quantification of DMS reactivity with and without 10 mM Mg^{2+} for P4-P6 and all tested miniTTR constructs. e) Quantification of the change in DMS reactivity of adenosines involved in the TTR as a function of Mg^{2+} for miniTTR 6 (blue), miniTTR 2 (red) and P4-P6 (black). f) Raw data from the Mg^{2+} titration of miniTTR 2, highlighting the change in DMS reactivity in the TTR and the motifs used in the design.

Figure 3: SAXS analysis of miniTTRs 2 and 6. a) Experimental intensity versus scattering vector for miniTTR 2 (top) and miniTTR 6 (bottom) compared to predictions based on models. b, c) Bead model derived from experimental scattering profiles overlaid with RNAMake models for miniTTR 2 (b) and miniTTR 6 (c).

Figure 4: Comparison between the RNAMake model and X-ray crystallographic structure of MiniTTR 6. a) Overviews of both the entire RNAMake model, X-ray crystallographic and secondary structure of miniTTR 6. b) Overlay of the tetraloop/tetraloop-receptor (TTR) region. c) Overlay of the first motif selected by RNAMake. d) Overlay of the second motif selected by RNAMake. e) Comparison between the RNAMake model (left) and the X-ray crystal structure (right) of the last motif selected by RNAMake. In (b)-(e), the motif from the X-ray structure is in white, and the RNAMake model is in color.

701 **Figure 5:** Summary of the RNAMake-based design of a functional tether between the 50S and
 702 30S ribosomal subunits. a) All unique motifs included in the H101_h44_Tether designs. b)
 703 Ensemble of all RNAMake H101_h44_Tether designs. Each is colored according to the motifs
 704 (panel (a)) that were used to generate it. c) Culture density of *E. coli* cells containing RNAMake
 705 H101_h44_Tether design 4. d) Gel assay of 8 distinct colonies, verifying that all but clone 4 have
 706 a Ribo-T rRNA-sized band with loss of wild-type 23S and 16S rRNA bands.

Construct	Chemical Mapping	Mobility Assay	Mg ²⁺ Titration
miniTTR 1	Success	Success	Success
miniTTR 2	Success	Success	Success
miniTTR 3	Partial	Failure	Failure
miniTTR 4	Partial	Success	Failure
miniTTR 5	Success	Success	Success
miniTTR 6	Success	Success	Success
miniTTR 7	Success	Success	Success
miniTTR 8	Failure	Failure	Failure
miniTTR 9	Success	Success	Success
miniTTR 10	Success	Success	Success
miniTTR 11	Success	Success	Success
miniTTR 12	Success	Success	Success
miniTTR 13	Failure	Failure	Failure
miniTTR 14	Success	Success	Success
miniTTR 15	Success	Multimer	Success
miniTTR 16	Success	Success	Success

Table 1: Biochemical assay results for each miniTTR.

710
711

	Native
Space group	C2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (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)
<i>R</i> _{sym} or <i>R</i> _{merge}	13.5(72.7)
<i>I</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
<i>R</i> _{work} / <i>R</i> _{free}	22.9/27.66
No. atoms	3071
RNA	3049
Ligand/ion	10
Water	4
<i>B</i> -factors	
RNA	40.915
Ligand/ion	33.284
Water	11.047
R.m.s deviations	
Bond lengths (Å)	0.0077
Bond angles (°)	1.762

712
713
714
715

Table 2: Data collection, phasing and refinement statistics for miniTTR 6.

References:

1. Butcher SE, Pyle AM. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc Chem Res.* 2011;44(12):1302–11.
2. Hendrix DK, Brenner SE, Holbrook SR. RNA structural motifs: building blocks of a modular biomolecule. *Q Rev Biophys.* 2005;38(3):221–43.
3. Leontis NB, Westhof E. The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol.* 2006;16(3):279–87.
4. Leontis NB, Westhof E. Analysis of RNA motifs. *Curr Opin Struct Biol.* 2003;13(3):300–8.
5. Grabow WW, Jaeger L. RNA self-assembly and RNA nanotechnology. *Acc Chem Res.* 2014;47(6):1871–80.
6. Jaeger L, Chworos A. The architectonics of programmable RNA and DNA nanostructures. *Curr Opin Struct Biol.* 2006;16(4):531–43.
7. Jaeger L, Leontis NB. Tecto-RNA: One-dimensional self-assembly through tertiary interactions. Vol. 39, *Angewandte Chemie - International Edition*. Wiley Online Library; 2000. p. 2521–4.
8. Grabow WW, Zakrevsky P, Afonin KA, Chworos A, Bruce A. Self-assembling RNA nanorings based on RNAI/II inverse kissing complexes. 2012;11(2):878–87.
9. Severcan I, Geary C, Verzemnieks EJ, Chworos A, Jaeger L. Square-Shaped RNA Particles From Different RNA Folds. 2012;100(2):130–4.
10. Afonin KA, Viard M, Koyfman AY, Martins AN, Kasprzak WK, Panigaj M, et al. Multifunctional RNA nanoparticles. *Nano Lett.* 2014;14(10):5662–71.
11. Dibrov SM, McLean J, Parsons J, Hermann T. Self-assembling RNA square. *Proc Natl Acad Sci U S A.* 2011;108(16):6405–8.
12. Geary C, Rothmund PWK, Andersen ES. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science.* 2014;345(6198):799–804.
13. Guo P. The Emerging Field of RNA Nanotechnology. *Nat Nanotechnol.* 2010;221(1):18–34.
14. Khisamutdinov EF, Jasinski DL, Li H, Zhang K, Chiu W, Guo P. Fabrication of RNA 3D Nanoprisms for Loading and Protection of Small RNAs and Model Drugs. *Adv Mater.* 2016;28(45):10079–87.
15. Wimberly BT, Wimberly BT, Brodersen DE, Brodersen DE, Jr WMC, Jr WMC, et al. Structure of the 30S ribosomal subunit. *Nature.* 2000;327–39.

- 751 16. Nguyen THD, Galej WP, Bai X, Savva CG, Newman AJ, Scheres SHW, et al. The
752 architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature*. 2015;523(7558):47–
753 52.
- 754 17. Woodson SA. Structure and assembly of group I introns. *Curr Opin Struct Biol*.
755 2005;15(3 SPEC. ISS.):324–30.
- 756 18. Petrov AI, Zirbel CL, Leontis NB. Automated classification of RNA 3D motifs and
757 the RNA 3D Motif Atlas. 2013;1327–40.
- 758 19. Leontis NB, Westhof E. Self-assembled RNA nanostructures. *Science*.
759 2014;345(6198):732–3.
- 760 20. Nasalean L, Baudrey S, Leontis NB, Jaeger L. Controlling RNA self-assembly to
761 form filaments. *Nucleic Acids Res*. 2006;34(5):1381–92.
- 762 21. Jucker FM, Heus HA, Yip PF, Moors EH, Pardi A. A network of heterogeneous
763 hydrogen bonds in GNRA tetraloops. Vol. 264, *Journal of molecular biology*.
764 Elsevier; 1996. p. 968–80.
- 765 22. Fried SD, Schmied WH, Uttamapinant C, Chin JW. Ribosome subunit stapling for
766 orthogonal translation in *E. coli*. *Angew Chem - Int Ed*. 2015;54(43):12791–4.
- 767 23. Orelle C, Carlson ED, Szal T, Florin T, Jewett MC, Mankin AS. Protein synthesis by
768 ribosomes with tethered subunits. *Nature*. 2015;doi: 10.1038/nature14862.
- 769 24. Delebecque CJ, Aldaye FA. Organization of Intracellular Reactions with Rationally
770 Designed RNA Assemblies. 2011;470.
- 771 25. Carlson ED. Creating Ribo-T: (Design, Build, Test). *ACS Synth Biol*. 2015 Nov
772 20;4(11):1173–5.
- 773 26. Macke TJ, Case DA. Modeling Unusual Nucleic Acid Structures BT - Molecular
774 Modeling of Nucleic Acids. *Mol Model Nucleic Acids*. 2009;682:379–93.
- 775 27. Lee J, Kladwang W, Lee M, Cantu D, Azizyan M, Kim H, et al. RNA design rules
776 from a massive open laboratory. 2014;111(6).
- 777 28. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, et al. Crystal
778 Structure of a Group I Ribozyme Domain : Principles of RNA Packing.
779 1996;273(September).
- 780 29. Klein DJ, Schmeing TM, Moore PB, Steitz TA. The kink-turn: A new RNA
781 secondary structure motif. *EMBO J*. 2001;20(15):4214–21.
- 782 30. Szewczak AA, Cech TR. An RNA internal loop acts as a hinge to facilitate ribozyme
783 folding and catalysis. Vol. 3, *Rna*. 1997. p. 838–49.
- 784 31. Hermann T, Patel DJ. Stitching together RNA tertiary architectures. *J Mol Biol*.
785 1999;294(4):829–49.

- 786 32. Kladwang W, Mann TH, Becka A, Tian S, Kim H, Yoon S, et al. Standardization of
787 RNA chemical mapping experiments. *Biochemistry (Mosc)*. 2014;53(19):3063–5.
- 788 33. Murphy FL, Cech TR. An independently folding domain of RNA tertiary structure
789 within the Tetrahymena ribozyme. *Biochemistry (Mosc)*. 1993;32:5291–300.
- 790 34. Young BT, Silverman SK. The GAAA tetraloop-receptor interaction contributes
791 differentially to folding thermodynamics and kinetics for the P4-P6 RNA domain.
792 *Biochemistry (Mosc)*. 2002;41(41):12271–6.
- 793 35. Frederiksen JK, Li N, Das R, Herschlag D, Piccirilli JA. Metal-ion rescue revisited :
794 Biochemical detection of site-bound metal ions important for RNA folding. 2012;
- 795 36. Lipfert J, Doniach S, Das R, Herschlag D. Understanding Nucleic Acid – Ion
796 Interactions. Vol. 83, *Annual Reviews of Biochemistry*. 2014. 813-841 p.
- 797 37. Takamoto K, Das R, He Q, Doniach S, Brenowitz M, Herschlag D, et al. Principles
798 of RNA compaction: Insights from the equilibrium folding pathway of the P4-P6
799 RNA domain in monovalent cations. *J Mol Biol*. 2004;343(5):1195–206.
- 800 38. Buchmueller KL, Webb AE, Richardson DA, Weeks KM. A collapsed non-native
801 RNA folding state. *Nat Struct Biol*. 2000;7(5):362–6.
- 802 39. Lipfert J, Chu VB, Bai Y, Herschlag D, Doniach S. Low-resolution models for
803 nucleic acids from small-angle X-ray scattering with applications to electrostatic
804 modeling. *J Appl Crystallogr*. 2007;40(SUPPL. 1):229–34.
- 805 40. Rambo RP, Tainer JA. Improving small-angle X-ray scattering data for structural
806 analyses of the RNA world. *Rna*. 2010;16(3):638–646.
- 807 41. Lipfert J, Herschlag D, Doniach S. Riboswitch Conformations Revealed by Small-
808 Angle X-Ray Scattering. 2009;540:141–59.
- 809 42. Klein DJ, Moore PB, Steitz TA. The roles of ribosomal proteins in the structure
810 assembly, and evolution of the large ribosomal subunit. *J Mol Biol*.
811 2004;340(1):141–77.
- 812 43. Watters AL, Deka P, Corrent C, Callender D, Varani G, Sosnick T, et al. The Highly
813 Cooperative Folding of Small Naturally Occurring Proteins Is Likely the Result of
814 Natural Selection. *Cell*. 2007;128(3):613–24.
- 815 44. Huang P-S, Feldmeier K, Parmeggiani F, Fernandez Velasco DA, Höcker B, Baker
816 D. De novo design of a four-fold symmetric TIM-barrel protein with atomic-level
817 accuracy. *Nat Chem Biol*. 2016;12(November):29–34.
- 818 45. Kiliszek A, Kierzek R, Krzyzosiak WJ, Rypniewski W. Atomic resolution structure of
819 CAG RNA repeats: Structural insights and implications for the trinucleotide repeat
820 expansion diseases. *Nucleic Acids Res*. 2010;38(22):8370–6.

- 821 46. Tawani A, Kumar A. Structural Insights Reveal the Dynamics of the Repeating
822 r(CAG) Transcript Found in Huntington's Disease (HD) and Spinocerebellar Ataxias
823 (SCAs). *PloS ONE*. 2015;10(7):e0131788.
- 824 47. Lu XJ, Bussemaker HJ, Olson WK. DSSR: an integrated software tool for
825 dissecting the spatial structure of RNA. *Nucleic Acids Res*. 2015;43(21):e142.
- 826 48. Lu XJ, Olson WK. 3DNA: A software package for the analysis, rebuilding and
827 visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res*.
828 2003;31(17):5108–21.
- 829 49. Huynh DQ. Metrics for 3D rotations: Comparison and analysis. *J Math Imaging Vis*.
830 2009;35(2):155–64.
- 831 50. Karney CFF. Quaternions in molecular modeling. *J Mol Graph Model*.
832 2007;25(5):595–604.
- 833 51. Tian S, Yesselman JD, Cordero P, Das R. Primerize: Automated primer assembly
834 for transcribing non-coding RNA domains. *Nucleic Acids Res*. 2015;43(W1):W522–
835 6.
- 836 52. Cordero P, Kladwang W, Vanlang CC, Das R. Quantitative dimethyl sulfate
837 mapping for automated RNA secondary structure inference. *Biochemistry (Mosc)*.
838 2012;51(36):7037–9.
- 839 53. Yoon S, Kim J, Hum J, Kim H, Park S, Kladwang W, et al. HiTRACE: High-
840 throughput robust analysis for capillary electrophoresis. *Bioinformatics*.
841 2011;27(13):1798–805.
- 842 54. Kim H, Cordero P, Das R, Yoon S. HiTRACE-Web : an online tool for robust
843 analysis of high-throughput capillary electrophoresis. 2013;41(June):492–8.
- 844 55. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. Accurate SAXS Profile
845 Computation and its Assessment by Contrast Variation Experiments. *Biophys J*.
846 2013 Aug 20;105(4):962–74.
- 847 56. Svergun DI. Advanced solution scattering data analysis methods and their
848 applications. *J Appl Crystallogr*. 2000 Jun;33(3 Part 1):530–534.
- 849 57. Franke D, Svergun DI. It DAMMIF, a program for rapid it ab-initio shape
850 determination in small-angle scattering. *J Appl Crystallogr*. 2009 Apr;42(2):342–
851 346.
- 852 58. Eiler D, Wang J, Steitz TA. Structural basis for the fast self-cleavage reaction
853 catalyzed by the twister ribozyme. *Proc Natl Acad Sci U S A*. 2014;111(36):13028–
854 33.
- 855 59. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr*. 2010;66(2):125–32.

- 856 60. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al.
857 Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol*
858 *Crystallogr.* 2011;67(4):235–42.
- 859 61. Zhou XE, Gao X, Barty A, Kang Y, He Y, Liu W, et al. X-ray laser diffraction for
860 structure determination of the rhodopsin-arrestin complex. *Sci Data.*
861 2016;3:160021.
- 862 62. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot.
863 *Acta Crystallogr D Biol Crystallogr.* 2010;66(4):486–501.
- 864 63. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO.
865 Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat*
866 *Methods.* 2009 May;6(5):343–5.
- 867