Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Structural Biology

Squaring theory with practice in RNA design

JP Bida[1,2] and R Das[1,3]

Ribonucleic acid (RNA) design offers unique opportunities for engineering genetic networks and nanostructures that self-assemble within living cells. Recent years have seen the creation of increasingly complex RNA devices, including proof-of-concept applications for *in vivo* three-dimensional scaffolding, imaging, computing, and control of biological behaviors. Expert intuition and simple design rules — the stability of double helices, the modularity of noncanonical RNA motifs, and geometric closure — have enabled these successful applications. Going beyond heuristics, emerging algorithms may enable automated design of RNAs with nucleotide-level accuracy but, as illustrated on a recent RNA square design, are not yet fully predictive. Looking ahead, technological advances in RNA synthesis and interrogation are poised to radically accelerate the discovery and stringent testing of design methods.

Addresses
[1] Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
[2] Department of Bioengineering, Stanford University, Stanford, CA 94305, USA
[3] Department of Physics, Stanford University, Stanford, CA 94305, USA

Corresponding author: Das, R (rhiju@stanford.edu)

Current Opinion in Structural Biology 2012, **22**:xx–yy

This review comes from a themed issue on **Engineering and design**

Edited by **Jane Clarke** and **Bill Schief**

0959-440X/$ – see front matter, Published by Elsevier Ltd.

http://dx.doi.org/10.1016/j.sbi.2012.06.003

Ribonucleic acid (RNA) is, in many respects, an ideal polymer for biomolecular design. Natural and designed RNA chains can be exquisitely functional: they code for genetic information, assemble into intricate three-dimensional structures, catalyze chemical reactions, and engage in nearly every essential biological process in living cells (extensively reviewed in [1]). Most appealingly, these RNA behaviors can be explained and designed by rules that appear strikingly simple and, in favorable cases, are quantitatively predictive. Indeed, our motivation for studying RNA design puzzles is that they offer opportunities to discover and rigorously test such predictive rules, which might, in turn, enable the modeling of any RNA 'from scratch'.

This review focuses on current and missing rules for RNA design, starting with a description of three basic heuristics in universal use. The three subsequent sections review recent work to expand each rule into predictive theories and tools, highlighting current mismatches between theory and practice. We conclude with a perspective on trends that may radically accelerate the discovery of new RNA design theories and devices.

Three heuristic design rules

All successful RNA designs to date have leveraged three basic rules, many first explored in the interrogation of natural RNA systems [3] and the design of DNA devices [4]. We illustrate these ideas with an 'RNA square' (Figure 1a–g), recently assembled from eight strands as a potential scaffold for nanoscale chemical reactions (Figure 1a) and crystallized by the Hermann lab [2••].

Rule 1. *Watson–Crick base pairs generate stable double helices.* Every RNA design made to date has taken advantage of helical stems formed as RNA strands double back on themselves or associate with other strands to form Watson–Crick base pairs. For the RNA square, its edges are four helices, each involving association of four strands (Figure 1b).

Rule 2. *RNA motifs can preserve their behavior when copied and pasted into new contexts.* Explorations of RNAs in living systems and elegant *in vitro* selections have revealed a plethora of natural RNA catalysts, sensors, and structures. Many of these molecules' functions are due to small (4–15 nucleotide) RNA motifs with non-canonical structure. These motifs can be grafted into if care is taken to avoid mispairings with flanking sequences. For example, the four 'nanocorners' of the RNA square (Figure 1c) are copies of a five-nucleotide bulge motif, drawn from a right-angle-forming structure in the Hepatitis C virus genome [5].

Rule 3. *Geometric closure ensures correct three-dimensional structure.* For 3D structures, the geometry of helical stems must successfully interconnect with noncanonical motifs — a stringent requirement for RNA designs that encompass closed 'ring' topologies (Figure 1d). For the RNA square, the choice of 10 base pairs for the edges leads to a well-ordered square assembly with no detectable alternative species. Extension to 11 base pairs precludes the closure of the square and gives higher-order assemblies [2••].

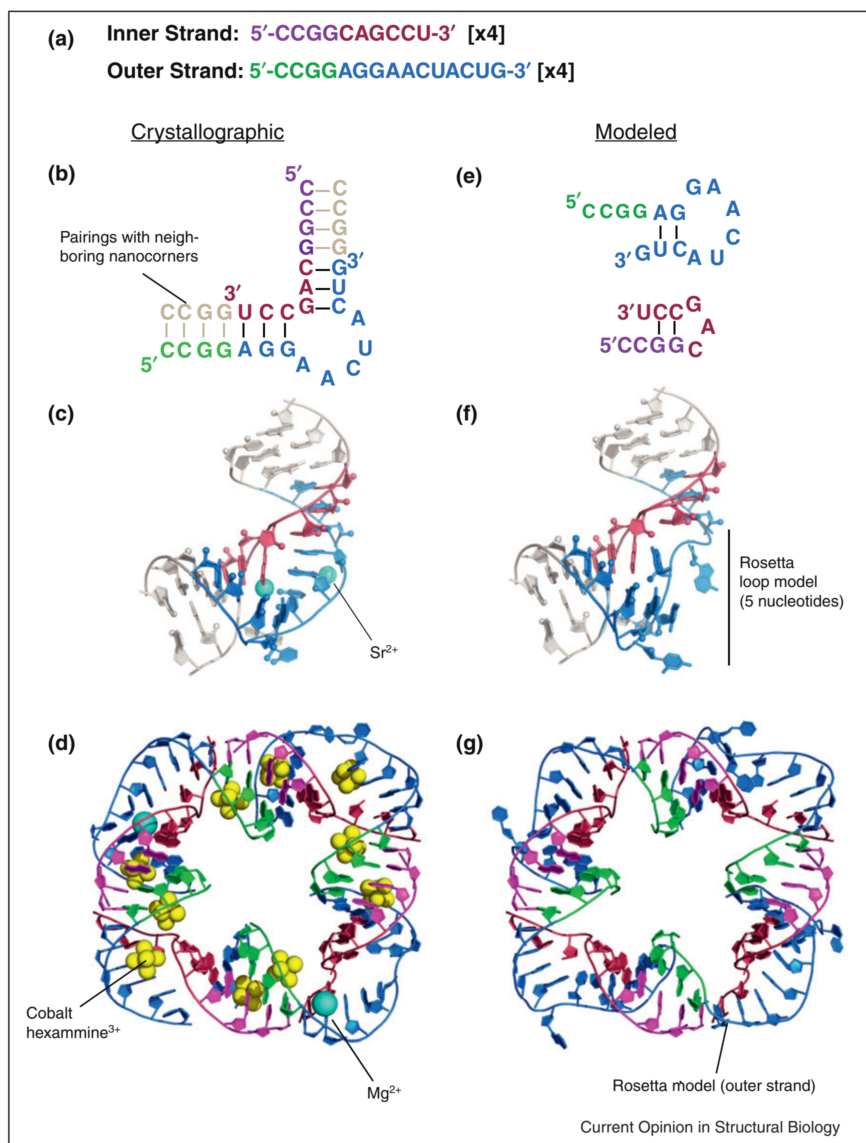
Can these intuitive rules be explained and expanded into quantitatively predictive theories for RNA design?

Expanding Rule 1: theories for RNA secondary structure design

Physical theories underlying RNA secondary structure formation are the most developed models in RNA science

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

Figure 1

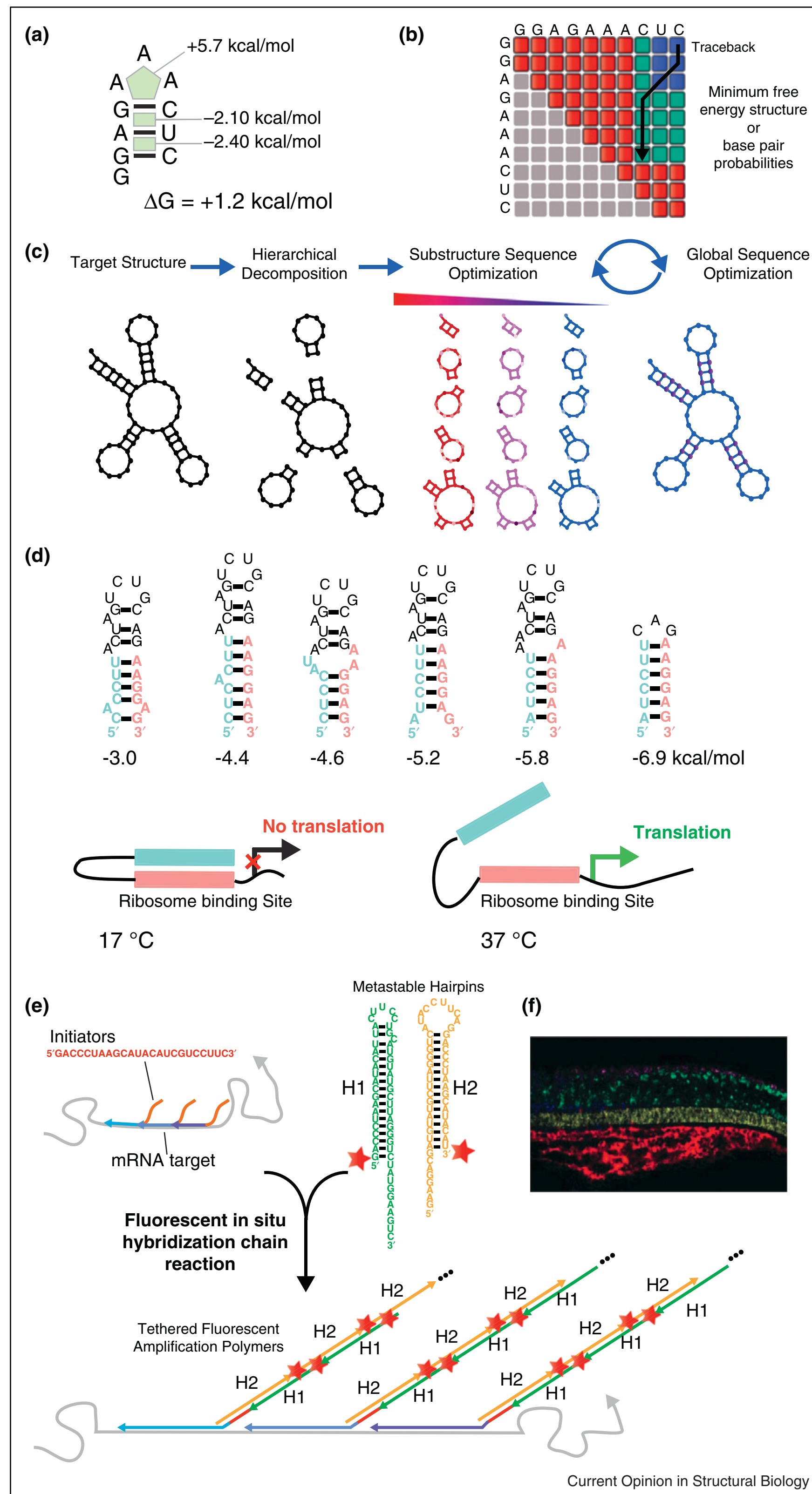


A square illustrates RNA design. **(a)** Sequences (inner and outer) determined 'by hand' by Hermann group [2••]; four copies of the two sequences form the square. **(b)** Multi-strand secondary structure involves each inner and outer strand forming a nanocorner, and four nanocorners pairing through three-nucleotide sticky ends. **(c)** Inspiration for the nanocorner, a 5-nt bulge that forms a 90 °C bend in the Hepatitis C virus internal ribosomal entry site (crystallized in PDB ID 2PN4). **(d)** Crystallographic analysis for full RNA square (PDB ID 3P59). **(e)** Secondary structure predictions for square design are inaccurate (NUPACK models shown [26]). **(f)** The *ab initio* modeling problem of rebuilding the nanocorner's 5-nt bulge — even given the rest of the coordinates — is not solvable at atomic resolution. **(g)** An RNA-Puzzle [70••]: blind prediction of the nanosquare conformation by Rosetta methods is not atomically accurate, even when given the inner strand coordinates as a constraint.

and arguably amongst the most predictive theories available in biophysics. Several decades of melt experiments on thousands of RNA sequences have been distilled into models that parameterize canonical base pair formation into two dozen 'nearest-neighbor' parameters, where the thermodynamic stability of a given base pair depends on adjacent base pairs and temperature [6,7]. Further modeling and measurements have provided approximate energetic rules (Figure 2a) for many hairpin loops, base pair mismatches, and more complex inter-helical motifs such as three-way junctions [7–13]. Given such energetic models, dynamic programming algorithms implemented in packages such as mfold/UNAfold [14], RNAstructure [15], and ViennaRNA [16] permit the comprehensive statistical mechanical description of RNA secondary structure ensembles for arbitrary sequences (Figure 2b) with the ability to recover ~70% of phylogenetically determined base pairs [7]. Modeling methods for secondary structure formation kinetics are also under exploration in packages such as Kinefold [17].

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

Figure 2



Theory and practice in RNA secondary structure design. **(a)** The free energy of a sequence in a given secondary structure conformation is calculated by summing up the free energies of nearest neighbor terms derived from extensive empirical measurements (green) [6–13]. **(b)** Dynamic programming algorithms can calculate the minimal free energy structure, base-pair probabilities, and partition function [14,15]. These methods recursively calculate

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

Most relevant for RNA design are recent approaches for predicting sequences that fold into given target secondary structures, the ''inverse'' folding problem. Methods range from simple Monte Carlo sequence searches (InverseRNA [18]) to faster hierarchical schemes (RNA-SSD [19] and INFO-RNA [20]) that attempt to solve the design problem for substructures or with simplifying assumptions before merging or optimizing solutions (Figure 2c). Most efforts have focused on finding sequences whose minimum free energy conformations recover the desired structure or, in some cases, multiple structures [16,21–23,24•]. The NUPACK software has presented optimization algorithms for a novel and intuitive target metric, the 'ensemble defect', which parameterizes the fluctuations of each nucleotide away from its desired configuration [25•,26]. Systematic experimental benchmarks of these design algorithms on novel secondary structure targets, perhaps using chemical accessibility mapping [27], would be valuable but are not yet available.

RNA secondary structure prediction and design algorithms are being widely used to help develop novel molecules. Design of RNA thermometers for temperature-controlled gene expression has made use of the nearest neighbor rules [28] (Figure 2d). Isambert, Schwalbe, and other labs have designed model systems with appealing simplicity to study cotranscriptional folding and two-state switching [29–32]. Carothers and colleagues have calculated the properties of RNA devices taking into account the folding rates of ribozymes and aptamers calculated in Kinefold [33]. Pierce and colleagues have imported the 'hybridization chain reaction' first developed in DNA engineering to create multiplexed amplifiers for *in situ* hybridization to mRNA targets in zebrafish embryos [34•] (Figure 2e and f) and to induce apoptosis of cultured human cells in response to cancer marker RNAs [35•].

While making use of RNA secondary structure modeling algorithms, these successful design efforts have still required significant insight from experts and would be challenging to generate automatically. For example, the RNA square, which was designed 'by hand' [2••], is a problem case for current secondary structure calculation methods, which predict alternative structures (Figure 1e). The squares' nanocorners are especially stable motif sequences that form noncanonical hydrogen bonds and specifically coordinate multivalent ions — features that are not treated in current RNA thermodynamic models. In principle, quantum chemical and molecular dynamics approaches should enable the *ab initio* calculation of these 'missing' parameters in arbitrary solution conditions but are still under calibration [36–43]. Further, most secondary structure prediction packages do not yet model ring-like or pseudoknot structures, although new extensions are attempting to tackle this issue [44,45,46•].

Expanding Rule 2: new RNA motifs and new combinations

A large arsenal of functional motifs is available for RNA design (Figure 3a). Quite broadly, RNA sequences that form specific three-dimensional structures, bind to small molecules or proteins, or mediate cellular localization, degradation, transcription termination, splicing, editing, or other cellular RNA processing events are widely used in biological inquiry (reviewed recently in [47,48,49]). Several groups are compiling databases with sequences and structures of existing motifs derived from natural functional RNAs or *in vitro* selections [50–54].

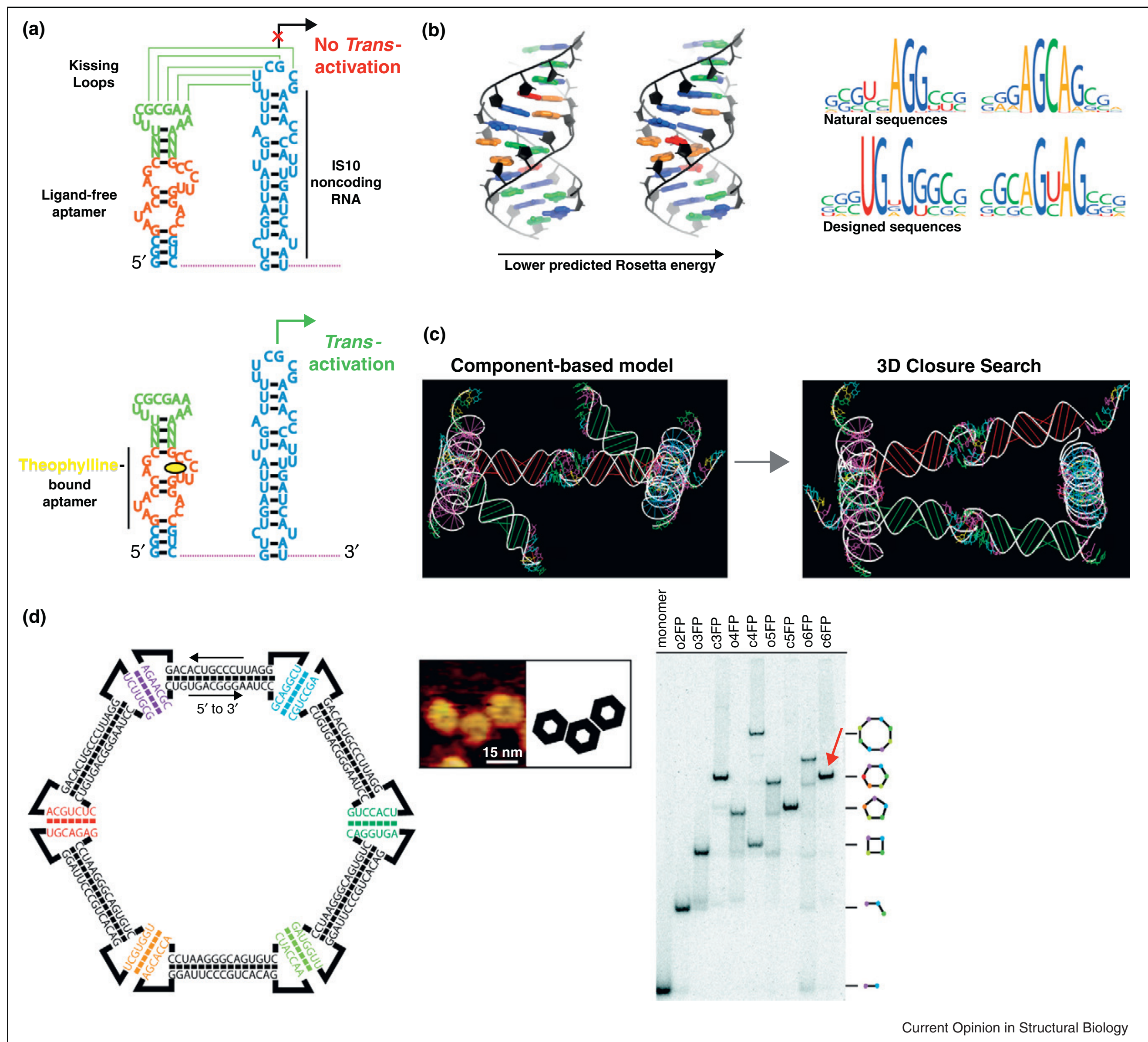
RNA designs have demonstrated the power of utilizing multiple interacting motifs, often coupled allosterically via secondary structure switches or kissing loops. Motifs embedded in such devices include small-molecule-binding 'aptamers' such as the theophylline-binding motif, catalytic modules such as the hammerhead ribozyme three-way junction, and protein-binding motifs [55–57,58•,59]. The resulting devices have been expressed in bacteria, yeast, and mammalian cells to create circuits of growing complexity, including on/off switches [60,61] (Figure 3a), RNA-based logic gates [56,62–64], cooperative behaviors [56], and bacteria that swim toward and ingest the herbicide atrazine [65]. Creative ways to discover and select new functional motifs are also being pursued, for example, through juxtaposition of randomized sequences by tertiary scaffolds [66,67] (see also below). While all the designs have leveraged the modularity of functional motifs, many successes have required trial-and-error or selection *in vivo* [57,65]. Presentation of failure cases and their in-depth dissections, as are now carried out in the protein design field [68,69], would be valuable.

Even with the growing database of known functional motifs, there may always be components for RNA design that are not naturally available, such as aptamers for specific moieties in large macromolecules. A general solution to such *de novo* design problems would be useful but has not been demonstrated. Indeed, when the RNA square crystallographic structure was offered as a blind trial to RNA modelers, no algorithm reached atomic accuracy predictions of the nanocorners even when given the coordinates of four of the eight internal strands of the motifs (Figure 1g) [70••]. Nevertheless, a recent

a substructure's properties using a matrix to store the values for each smaller substructure. **(c)** Workflow adopted by many inverse RNA structure prediction programs [19,20,25]. **(d)** Free energy calculations were used to design RNA thermosensors that inhibit translation below a target temperature [28]. **(e)** Hybridization chain reaction: metastable hairpins (green and orange) polymerize in the presence of an initiation sequence (red) to amplify the fluorescent signal during *in situ* hybridization experiments [34•]. **(f)** Zebrafish cross section with targeting TG(flk1:egfp) (red), *tpm3* (blue), *elavl3* (green), *ntla* (yellow), and *sox10* (purple) [34•].

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

Figure 3



Theory and practice in RNA tertiary structure design. **(a)** Cutting and pasting motifs: kissing loops, a theophylline aptamer, and the IS10 ncRNA were rationally combined to engineer trans-acting ncRNAs that regulates gene translation that can be turned on or off with theophylline [57]. **(b)** New tertiary motifs generated with Rosetta fixed backbone design for RNA, based on *in silico* energy minimization. The resulting designs for the most conserved domain of the signal recognition particle RNA are distinct from natural RNAs and turn out to be more stable experimentally [75•]. **(c)** RNA tectonics: combination of tertiary components and helices in three dimensions, optimized by hand or computationally [106]. **(d)** Nanoring assembled from RNAI/II inverse kissing complexes, validated by atomic force microscopy (middle) and non-denaturing gel electrophoresis [88].

explosion of activity in RNA 3D classification [71–73] and motif modeling [70••,74] (Figure 1f and g), with some algorithms achieving atomic accuracy in favorable cases [39,75•,76], suggests that *de novo* design may become a reality. As an early step, we adapted the Rosetta design method for fixed-backbone biopolymer design [77] to successfully thermostabilize a non-canonical motif in the signal recognition particle RNA through mutations previously unseen in Nature (Figure 3b) [75•].

Expanding Rule 3: theories for RNA tertiary structure design

The successful design of highly complex 3D RNA structures continues to involve connecting RNA motifs and

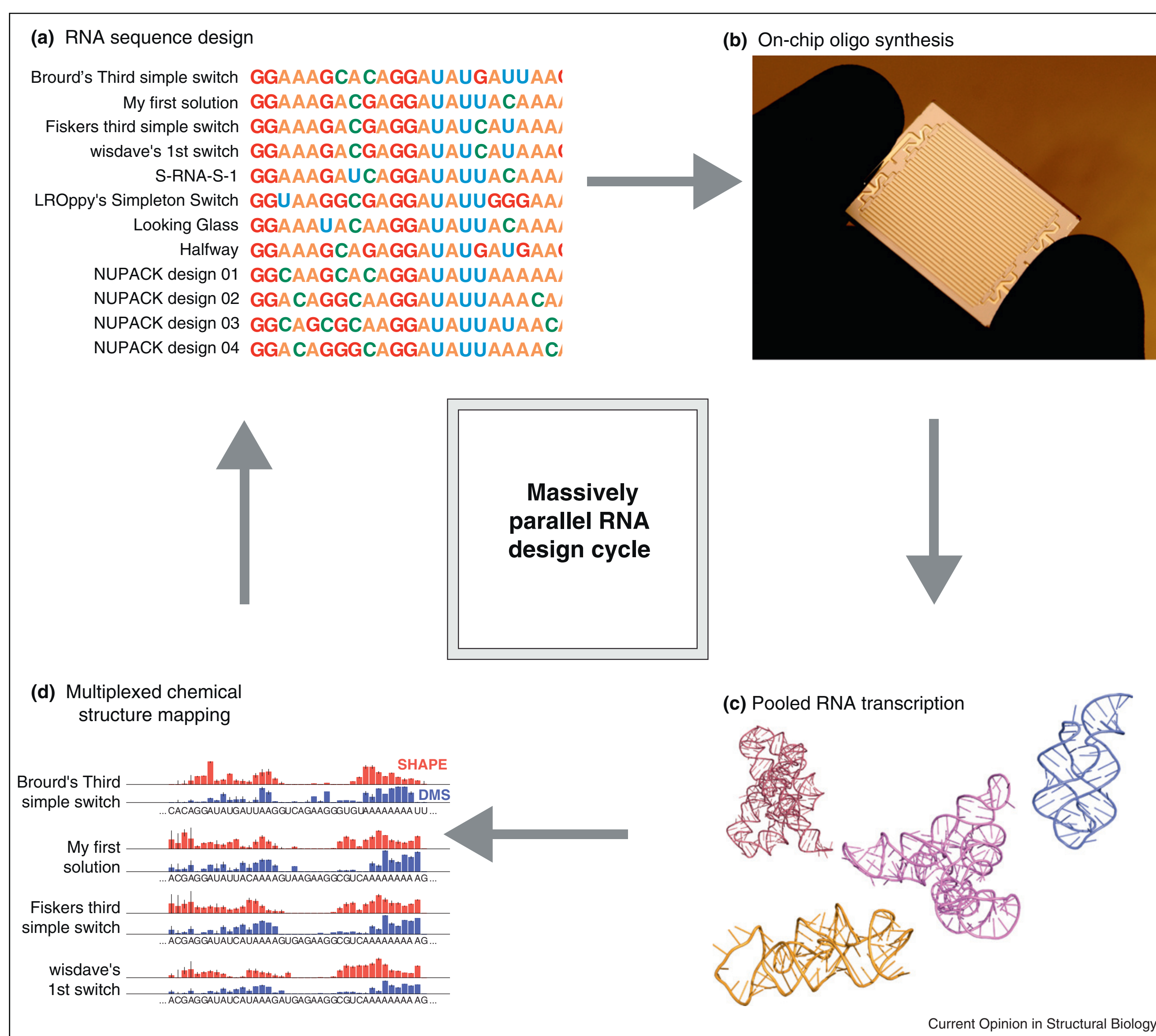
Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

helical stems, typically with additional tertiary motifs and cross-strand interactions that generate ring-like topologies (Figure 3d). This procedure, sometimes called RNA 'tectonics', is accomplished with manually directed programs such as NanoTiler [78] or Assemble [79] to put together RNA components *in silico*.

Visually compelling examples of designed 3D RNA structures have proliferated since the first 'TectoRNA', an RNA dimer stabilized by docking of tetraloops and their cognate receptors [80]. Several groups have designed and validated RNA tiles [81,82••], multimers [83], cubes [24•], prisms [84], triangles [85], fabrics [81,82••], fibrils [86,87], and rings (Figure 3d) [88,89], often using different components (tetraloop/receptors, kissing loops, tRNA-based junctions) to make similar objects [90•].

Many of these designs echo work in DNA nanotechnology [4], but RNA design permits, in principle, the deployment of nanostructures in living cells. Notably, Delebecque and colleagues [82••] have demonstrated self-assembly of nanofabrics *in vivo* and used the resulting structures to accelerate hydrogen production in bacteria

Figure 4



A massively parallel design cycle should be possible for RNA. Thousands of sequences **(a)** designed by experts, citizen scientists, and automated design tools are sent for high-throughput oligonucleotide template synthesis **(b)** [96,97], (figure adopted from LC sciences) transcribed into RNA **(c)**, and then falsified or validated at nucleotide resolution through multiplexed chemical mapping **(d)** utilizing-next generation-sequencing methods [99].

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

through the scaffolding of protein enzymes. Other applications include scaffolding gold particles for nanowires [91], encaging other molecules [84], and protecting duplexes intended for RNA silencing from degradation [88,92].

Atomic-scale validation of 3D design — as was achieved quite early for protein design [93,94] — has not yet been demonstrated for RNA nanostructures. For example, the crystallographic analysis of the four-fold sequence-symmetric RNA square showed the four nanocorners forming different structures (Figure 1d). In addition to structure modeling puzzles, an unaddressed challenge is to model the thermodynamic stabilities of RNA tertiary folds with accuracies comparable to the nearest-neighbor rules for secondary structure energetics. Improving tools for measuring helix and junction flexibility (see, e.g., [95]) should aid greatly in such efforts.

Perspective: an impending acceleration

In recent years, RNA designers have leveraged simple rules for secondary structure formation, the modularity of small RNA motifs, and three-dimensional closure to produce devices and nanostructures of growing complexity. It is particularly exciting to see these designs deployed into biological systems, suggesting future routes to biomedical devices that sense and perhaps correct cellular dysfunction. Nevertheless, challenges remain, particularly in attaining predictive theories for RNA assembly at nucleotide resolution, much less atomic resolution. Multi-strand topologies, non-canonical interactions, and energies of 3D assemblies remain unsolved challenges in RNA design theory. Both our physical models and our ability to make predictions from these models need improvement.

We are optimistic about the development of more quantitative and predictive theories for RNA structure design and eventually RNA functional design due to the explosion of RNA data expected in upcoming years. The cycle of RNA design and testing can be short, especially compared to protein engineering — arbitrary RNA sequences up to hundreds of nucleotides in length are experimentally straightforward to synthesize, purify, interrogate, and evolve on the timescale of days. Current technologies offer the parallel synthesis of thousands of arbitrary DNA templates (Figure 4a and b) [96,97], which can be transcribed into RNA *in vitro* and, in principle, *in vivo* (Figure 4c). Further, multiplexed single-nucleotide-resolution chemical mapping [98,99] and elegant selection strategies [100] can provide information-rich data on all of these molecules' structures and functions, using deep-sequencing platforms with turn-around times of hours (Figure 4d). The key challenges will then be to distribute [101] and accurately analyze these data [102–105], to update RNA modeling algorithms, and to feed back these insights into the next rounds of synthesis and mapping. The result, a massively parallel design cycle, would offer unprecedented opportunities for squaring theory with practice in RNA design.

Acknowledgments

We thank EteRNA players for sequences in Figure 4. We thank members of the Das laboratory, EteRNA players Quasispecies and Eli Fisker, J. Lee, and J. Lucks for discussions and manuscript comments. We acknowledge funding from the Burroughs-Wellcome Foundation (CASI to RD) and a Stanford Bio-X Games Fellowship (to JPB).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Gesteland RF, Cech TR, Atkins JF: *The RNA World*. Cold Spring Harbor Laboratory Press; 2006.

2. •• Dibrov SM, McLean J, Parsons J, Hermann T: **Self-assembling RNA square**. *Proc Natl Acad Sci U S A* 2011, **108**:6405-6408.
First example of high-resolution crystallography testing a three-dimensional RNA design.

3. Cruz JA, Westhof E: **The dynamic landscapes of RNA architecture**. *Cell* 2009, **136**:604-609.

4. Pinheiro AV, Han D, Shih WM, Yan H: **Challenges and opportunities for structural DNA nanotechnology**. *Nat Nanotechnol* 2011, **6**:763-772.

5. Zhao Q, Han Q, Kissinger CR, Hermann T, Thompson PA: **Structure of hepatitis C virus IRES subdomain IIa**. *Acta Crystallogr D: Biol Crystallogr* 2008, **64**:436-443.

6. Xia T, SantaLucia J Jr, Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH: **Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson–Crick base pairs**. *Biochemistry* 1998, **37**:14719-14735.

7. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH: **Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure**. *Proc Natl Acad Sci U S A* 2004, **101**:7287-7292.

8. Sheehy JP, Davis AR, Znosko BM: **Thermodynamic characterization of naturally occurring RNA tetraloops**. *RNA* 2010, **16**:417-429.

9. Turner DH, Mathews DH: **NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure**. *Nucleic Acids Res* 2010, **38**:D280-D282.

10. Davis AR, Znosko BM: **Positional and neighboring base pair effects on the thermodynamic stability of RNA single mismatches**. *Biochemistry* 2010, **49**:8669-8679.

11. Thulasi P, Pandya LK, Znosko BM: **Thermodynamic characterization of RNA triloops**. *Biochemistry* 2010, **49**:9058-9062.

12. Liu B, Diamond JM, Mathews DH, Turner DH: **Fluorescence competition and optical melting measurements of RNA three-way multibranch loops provide a revised model for thermodynamic parameters**. *Biochemistry* 2011, **50**:640-653.

13. Vanegas PL, Horwitz TS, Znosko BM: **Effects of non-nearest neighbors on the thermodynamic stability of RNA GNRA hairpin tetraloops**. *Biochemistry* 2012, **51**:2192-2198.

14. Markham NR, Zuker M: **UNAFold: software for nucleic acid folding and hybridization**. *Methods Mol Biol* 2008, **453**:3-31.

15. Reuter JS, Mathews DH: **RNAstructure: software for RNA secondary structure prediction and analysis**. *BMC Bioinformatics* 2010, **11**:129.

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

16. Lorenz R, Bernhart SH, Hoener Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL: **ViennaRNA package 2.0**. *Algorithms Mol Biol: AMB* 2011, **6**:26.

17. Xayaphoummine A, Bucher T, Isambert H: **Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots**. *Nucleic Acids Res* 2005, **33**:W605-W610.

18. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer SL, Tacker M, Schuster P: **Fast folding and comparison of RNA secondary structures**. *Monatsh Chem* 1994, **125**:167-188.

19. Andronescu M, Fejes AP, Hutter F, Hoos HH, Condon A: **A new algorithm for RNA secondary structure design**. *J Mol Biol* 2004, **336**:607-624.

20. Busch A, Backofen R: **INFO-RNA — a fast approach to inverse RNA folding**. *Bioinformatics* 2006, **22**:1823-1831.

21. Taneda A: **MODENA: a multi-objective RNA inverse folding**. *Adv Appl Bioinformatics Chem: AABC* 2011, **4**:1-12.

22. Ramlan EI, Zauner KP: **Design of interacting multi-stable nucleic acids for molecular information processing**. *Bio Systems* 2011, **105**:14-24.

23. Shu W, Liu M, Chen H, Bo X, Wang S: **ARDesigner: a web-based system for allosteric RNA design**. *J Biotechnol* 2010, **150**:466-473.

24. • Afonin KA, Bindewald E, Yaghoubian AJ, Voss N, Jacovetty E, Shapiro BA, Jaeger L: **In vitro assembly of cubic RNA-based scaffolds designed in silico**. *Nat Nanotechnol* 2010, **5**:676-682.
Thorough illustration of how to design RNA secondary structure *in silico*, build three dimensional scaffolds, and experimentally validate such designs. The resulting cubes efficiently self-assemble during transcription, suggesting the possibility of generating RNA packages inside cells.

25. • Zadeh JN, Wolfe BR, Pierce NA: **Nucleic acid sequence design via efficient ensemble defect optimization**. *J Comput Chem* 2011, **32**:439-452.
This work makes a strong case for a new intuitive metric to assess *in silico* RNA secondary structure designs and presents an elegant and fast algorithm for optimizing it.

26. Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA: **NUPACK: analysis and design of nucleic acid systems**. *J Comput Chem* 2011, **32**:170-173.

27. Weeks KM: **Advances in RNA structure analysis by chemical probing**. *Curr Opin Struct Biol* 2010, **20**:295-304.

28. Neupert J, Bock R: **Designing and using synthetic RNA thermometers for temperature-controlled gene expression in bacteria**. *Nat Protoc* 2009, **4**:1262-1273.

29. Xayaphoummine A, Viasnoff V, Harlepp S, Isambert H: **Encoding folding paths of RNA switches**. *Nucleic Acids Res* 2007, **35**:614-622.

30. Cao S, Furtig B, Schwalbe H, Chen SJ: **Folding kinetics for the conformational switch between alternative RNA structures**. *J Phys Chem B* 2010, **114**:13609-13615.

31. Furtig B, Wenter P, Reymond L, Richter C, Pitsch S, Schwalbe H: **Conformational dynamics of bistable RNAs studied by time-resolved NMR spectroscopy**. *J Am Chem Soc* 2007, **129**:16222-16229.

32. Cayrol B, Isambert H: **RNA synthetic biology inspired from bacteria: construction of transcription attenuators under antisense regulation**. *Phys Biol* 2009, **6**:025007.

33. Carothers JM, Goler JA, Juminaga D, Keasling JD: **Model-driven engineering of RNA devices to quantitatively program gene expression**. *Science* 2011, **334**:1716-1719.

34. • Choi HMT, Chang JY, Trinh LA, Padilla JE, Fraser SE, Pierce NA: **Programmable in situ amplification for multiplexed imaging of mRNA expression**. *Nat Biotechnol* 2010, **28**:1208-1212.
Amplification cascades involving RNA helix rearrangements permit high signal-to-noise imaging of up to five transcripts in vertebrate embryos.

35. • Venkataraman S, Dirks RM, Ueda CT, Pierce NA: **Selective cell death mediated by small conditional RNAs**. *Proc Natl Acad Sci U S A* 2010, **107**:16777-16782.
Fascinating demonstration of the hybridization chain reaction of designed small RNAs *in vivo*. The small RNAs are specifically triggered by cancer marker mRNAs to polymerize into long duplexes, which initiate an innate immune response that leads to cell death.

36. Ditzler MA, Otyepka M, Sponer J, Walter NG: **Molecular dynamics and quantum mechanics of RNA: conformational and chemical change we can believe in**. *Acc Chem Res* 2010, **43**:40-47.

37. Yildirim I, Stern HA, Tubbs JD, Kennedy SD, Turner DH: **Benchmarking AMBER force fields for RNA: comparisons to NMR spectra for single-stranded r(GACC) are improved by revised chi torsions**. *J Phys Chem B* 2011, **115**:9261-9270.

38. Yildirim I, Stern HA, Kennedy SD, Tubbs JD, Turner DH: **Reparameterization of RNA chi torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine**. *J Chem Theory Comput* 2010, **6**:1520-1531.

39. Lerman YV, Kennedy SD, Shankar N, Parisien M, Major F, Turner DH: **NMR structure of a 4 × 4 nucleotide RNA internal loop from an R2 retrotransposon: identification of a three purine–purine sheared pair motif and comparison to MC-SYM predictions**. *RNA* 2011, **17**:1664-1677.

40. Van Nostrand KP, Kennedy SD, Turner DH, Mathews DH: **Molecular mechanics investigation of an adenine–adenine non-canonical pair conformational change**. *J Chem Theory Comput* 2011, **7**:3779-3792.

41. Yildirim I, Kennedy SD, Stern HA, Hart JM, Kierzek R, Turner DH: **Revision of AMBER torsional parameters for RNA improves free energy predictions for tetramer duplexes with GC and iGiC base pairs**. *J Chem Theory Comput* 2012, **8**:172-181.

42. Sponer J, Sponer JE, Petrov AI, Leontis NB: **Quantum chemical studies of nucleic acids: can we construct a bridge to the RNA structural biology and bioinformatics communities?** *J Phys Chem B* 2010, **114**:15723-15741.

43. Zgarbova M, Otyepka M, Sponer J, Mladek A, Banas P, Cheatham TE 3rd, Jurecka P: **Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles**. *J Chem Theory Comput* 2011, **7**:2886-2902.

44. Cao S, Chen SJ: **Predicting structures and stabilities for H-type pseudoknots with interhelix loops**. *RNA* 2009, **15**:696-706.

45. Bindewald E, Kluth T, Shapiro BA: **CyloFold: secondary structure prediction including pseudoknots**. *Nucleic Acids Res* 2010, **38**:W368-W372.

46. • Bindewald E, Afonin K, Jaeger L, Shapiro BA: **Multistrand RNA secondary structure prediction and nanostructure design including pseudoknots**. *ACS Nano* 2011, **5**:9542-9551.
Intriguing proposal that simple, heuristic design algorithms may be powerful for designing complex 3D topologies, even if they sacrifice the thermodynamic rigor and applicability to natural RNAs of classic modeling approaches.

47. Lucks JB, Qi L, Whitaker WR, Arkin AP: **Toward scalable parts families for predictable design of biological circuits**. *Curr Opin Microbiol* 2008, **11**:567-573.

48. Liang JC, Bloom RJ, Smolke CD: **Engineering biological systems with synthetic RNA molecules**. *Mol Cell* 2011, **43**:915-926.

49. Benenson Y: **Synthetic biology with RNA: progress report**. *Curr Opin Chem Biol* 2012, **16**:1-7.

50. Klosterman P, Tamura M, Holbrook S: **SCOR: a structural classification of RNA database**. *Nucleic Acids Res* 2002, **30**:392-394.

51. Meyers L, Ellington A: **Aptamer database**. *Nucleic Acids Res* 2004, **32**:D95-D100.

52. Smolke CD: **Building outside of the box: iGEM and the BioBricks foundation**. *Nat Biotechnol* 2009, **27**:1099-1102.

53. Vanegas PL, Hudson GA, Davis AR, Kelly SC, Kirkpatrick CC, Znosko BM: **RNA CoSSMos: characterization of secondary structure motifs — a searchable database of secondary structure motifs in RNA three-dimensional structures**. *Nucleic Acids Res* 2011, **40**:D439-D444.

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

54. Bindewald E, Hayes R, Yingling Y: **RNAJunction: a database of RNA junctions and kissing loops for three-dimensional structural analysis and nanodesign**. *Nucleic Acids Res* 2008, **36**:D392-D397.

55. Matsumura S, Ohmori R, Saito H, Ikawa Y, Inoue T: **Coordinated control of a designed trans-acting ligase ribozyme by a loop-receptor interaction**. *FEBS Lett* 2009, **583**:2819-2826.

56. Win MN, Smolke CD: **Higher-order cellular information processing with synthetic RNA devices**. *Science* 2008, **322**:456-460.

57. Qi L, Lucks JB, Liu CC, Mutalik VK, Arkin AP: **Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals**. *Nucleic Acids Res* 2012, **40**:5775-5786.

58. • Lucks JB, Qi L, Mutalik VK, Wang D, Arkin AP: **Versatile RNA-sensing transcriptional regulators for engineering genetic networks**. *Proc Natl Acad Sci U S A* 2011, **108**:8617-8622.
RNA loop–loop interactions, a new addition to the *in vivo* design toolkit, can be multiplexed, work in tandem, and propagate signals via transcriptional cascades. A highlight is the quantitative and fine-grained recapitulation of a tandem RNA gate's *in vivo* properties from the careful characterization of its components.

59. Culler SJ, Hoff KG, Smolke CD: **Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins**. *Science* 2010, **330**:1251-1255.

60. Callura JM, Dwyer DJ, Isaacs FJ, Cantor CR, Collins JJ: **Tracking, tuning, and terminating microbial physiology using synthetic riboregulators**. *Proc Natl Acad Sci U S A* 2010, **107**:15898-15903.

61. Saito H, Fujita Y, Kashida S, Hayashi K, Inoue T: **Synthetic human cell fate regulation by protein-driven RNA switches**. *Nat Commun* 2011, **2**:160.

62. Yen L, Svendsen J, Lee JS, Gray JT, Magnier M, Baba T, D'Amato RJ, Mulligan RC: **Exogenous control of mammalian gene expression through modulation of RNA self-cleavage**. *Nature* 2004, **431**:471-476.

63. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y: **Multi-input RNAi-based logic circuit for identification of specific cancer cells**. *Science* 2011, **333**:1307-1311.

64. Ausländer Simon, Ausländer David, Müller Marius, Wieland Markus, Fussenegger Martin: **Programmable single-cell mammalian biocomputers**. *Nature* 2012, **487**:123-127.

65. Sinha J, Reyes SJ, Gallivan JP: **Reprogramming bacteria to seek and destroy an herbicide**. *Nat Chem Biol* 2010, **6**:464-470.

66. Geary C, Baudrey S, Jaeger L: **Comprehensive features of natural and in vitro selected GNRA tetraloop-binding receptors**. *Nucleic Acids Res* 2008, **36**:1138-1152.

67. Shiohara T, Saito H, Inoue T: **A designed RNA selection: establishment of a stable complex between a target and selectant RNA via two coordinated interactions**. *Nucleic Acids Res* 2009, **37**:e23.

68. Lassila JK, Baker D, Herschlag D: **Origins of catalysis by computationally designed retroaldolase enzymes**. *Proc Natl Acad Sci U S A* 2010, **107**:4937-4942.

69. Althoff EA, Wang L, Jiang L, Giger L, Lassila JK, Wang Z, Smith M, Hari S, Kast P, Herschlag D *et al.*: **Robust design and optimization of retroaldol enzymes**. *Protein Sci* 2012, **21**:717-726.

70. •• Cruz JA, Blanchet M-F, Boniecki M, Bujnicki JM, Chen S-J, Cao S, Das R, Ding F, Dokholyan NV, Flores SC *et al.*: **RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction**. *RNA* 2012, **18**:610-625.
First instance of a community-wide blind trial in RNA structure prediction — a compelling snapshot of what can and cannot be currently modeled.

71. Richardson JS, Schneider B, Murray LW, Kapral GJ, Immormino RM, Headd JJ, Richardson DC, Ham D, Hershkovits E, Williams LD *et al.*: **RNA backbone: consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution)**. *RNA* 2008, **14**:465-481.

72. Stombaugh J, Zirbel CL, Westhof E, Leontis NB: **Frequency and isostericity of RNA base pairs**. *Nucleic Acids Res* 2009, **37**:2294-2312.

73. Petrov AI, Zirbel CL, Leontis NB: **WebFR3D — a server for finding, aligning and analyzing recurrent RNA 3D motifs**. *Nucleic Acids Res* 2011, **39**:W50-W55.

74. Leontis NWE: *RNA 3D Structure Analysis and Prediction*. Springer Press; 2012.

75. • Das R, Karanicolas J, Baker D: **Atomic accuracy in predicting and designing noncanonical RNA structure**. *Nat Methods* 2010, **7**:291-294.
Application of core ideas in high-resolution protein structure prediction and design to RNA motifs, including a first step to 'de novo' design of new motifs.

76. Sripakdeevong P, Kladwang W, Das R: **An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling**. *Proc Natl Acad Sci U S A* 2011, **108**:20573-20578.

77. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, Kaufman K, Renfrew PD, Smith CA, Sheffler W *et al.*: **ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules**. *Methods Enzymol* 2011, **487**:545-574.

78. Bindewald E, Grunewald C, Boyle B, O'Connor M, Shapiro BA: **Computational strategies for the automated design of RNA nanoscale structures from building blocks using NanoTiler**. *J Mol Graph Model* 2008, **27**:299-308.

79. Jossinet F, Ludwig TE, Westhof E: **Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels**. *Bioinformatics* 2010, **26**:2057-2059.

80. Jaeger L, Westhof E, Leontis NB: **TectoRNA: modular assembly units for the construction of RNA nano-objects**. *Nucleic Acids Res* 2001, **29**:455-463.

81. Chworos A, Severcan I, Koyfman AY, Weinkam P, Oroudjev E, Hansma HG, Jaeger L: **Building programmable jigsaw puzzles with RNA**. *Science* 2004, **306**:2068-2072.

82. •• Delebecque CJ, Lindner AB, Silver PA, Aldaye FA: **Organization of intracellular reactions with rationally designed RNA assemblies**. *Science* 2011, **333**:470-474.
First demonstration of RNA nanostructure creation *in vivo*, based on tile sequences from an automated design algorithm. The resulting scaffolds engage with metabolic enzymes in bacteria to enhance hydrogen production.

83. Shu D, Moll WD, Deng Z, Mao C, Guo P: **Bottom-up assembly of RNA arrays and superstructures as potential parts in nanotechnology**. *Nano Lett* 2004, **4**:1717-1723.

84. Severcan I, Geary C, Chworos A, Voss N, Jacovetty E, Jaeger L: **A polyhedron made of tRNAs**. *Nat Chem* 2010, **2**:772-779.

85. Ohno H, Kobayashi T, Kabata R, Endo K, Iwasa T, Yoshimura SH, Takeyasu K, Inoue T, Saito H: **Synthetic RNA–protein complex shaped like an equilateral triangle**. *Nat Nanotechnol* 2011, **6**:116-120.

86. Nasalean L, Baudrey S, Leontis NB, Jaeger L: **Controlling RNA self-assembly to form filaments**. *Nucleic Acids Res* 2006, **34**:1381-1392.

87. Cayrol B, Nogues C, Dawid A, Sagi I, Silberzan P, Isambert H: **A nanostructure made of a bacterial noncoding RNA**. *J Am Chem Soc* 2009, **131**:17270-17276.

88. Grabow WW, Zakrevsky P, Afonin KA, Chworos A, Shapiro BA, Jaeger L: **Self-assembling RNA nanorings based on RNAI/II inverse kissing complexes**. *Nano Lett* 2011, **11**:878-887.

89. Ding F, Lu C, Zhao W, Rajashankar KR, Anderson DL, Jardine PJ, Grimes S, Ke A: **Structure and assembly of the essential RNA ring component of a viral DNA packaging motor**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**:7357-7362.

90. • Severcan I, Geary C, Verzemnieks E, Chworos A, Jaeger L: **Square-shaped RNA particles from different RNA folds**. *Nano Lett* 2009, **9**:1270-1277.
Uniquely systematic exploration of a classic puzzle in RNA design, the formation of tiles or squares.

91. Koyfman AY, Braun G, Magonov S, Chworos A, Reich NO, Jaeger L: **Controlled spacing of cationic gold nanoparticles by nanocrown RNA**. *J Am Chem Soc* 2005, **127**:11886-11887.

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003

92. Nakashima Y, Abe H, Abe N, Aikawa K, Ito Y: **Branched RNA nanostructures for RNA interference**. *Chem Commun (Cambridge, England)* 2011, **47**:8367-8369.

93. Harbury PB, Zhang T, Kim PS, Alber T: **A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants**. *Science* 1993, **262**:1401-1407.

94. Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D: **Design of a novel globular protein fold with atomic-level accuracy**. *Science* 2003, **302**:1364-1368.

95. Zhang Q, Stelzer AC, Fisher CK, Al-Hashimi HM: **Visualizing spatially correlated dynamics that directs RNA conformational transitions**. *Nature* 2007, **450**:1263-1267.

96. Zhou X, Cai S, Hong A, You Q, Yu P, Sheng N, Srivannavit O, Muranjan S, Rouillard JM, Xia Y *et al.*: **Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences**. *Nucleic Acids Res* 2004, **32**:5409-5417.

97. LeProust EM, Peck BJ, Spirin K, McCuen HB, Moore B, Namsaraev E, Caruthers MH: **Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process**. *Nucleic Acids Res* 2010, **38**:2522-2540.

98. Kladwang W, VanLang CC, Cordero P, Das R: **A two-dimensional mutate-and-map strategy for non-coding RNA structure**. *Nat Chem* 2011, **3**:954-962.

99. Lucks JB, Mortimer SA, Trapnell C, Luo S, Aviran S, Schroth GP, Pachter L, Doudna JA, Arkin AP: **Multiplexed RNA structure characterization with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq)**. *Proc Natl Acad Sci U S A* 2011, **108**:11063-11068.

100. Pitt JN, Ferre-D'Amare AR: **Rapid construction of empirical RNA fitness landscapes**. *Science* 2010, **330**:376-379.

101. Rocca-Serra P, Bellaousov S, Birmingham A, Chen C, Cordero P, Das R, Davis-Neulander L, Duncan CD, Halvorsen M, Knight R *et al.*: **Sharing and archiving nucleic acid structure mapping data**. *RNA* 2011, **17**:1204-1212.

102. Kladwang W, VanLang CC, Cordero P, Das R: **Understanding the errors of SHAPE-directed RNA structure modeling**. *Biochemistry* 2011, **50**:8049-8056.

103. Bindewald E, Wendeler M, Legiewicz M, Bona MK, Wang Y, Pritt MJ, Le Grice SF, Shapiro BA: **Correlating SHAPE signatures with three-dimensional RNA structures**. *RNA* 2011, **17**:1688-1696.

104. Washietl S, Hofacker IL, Stadler PF, Kellis M: **RNA folding with soft constraints: reconciliation of probing data and thermodynamic secondary structure prediction**. *Nucleic Acids Res* 2012, **40**:4261-4272.

105. Quarrier S, Martin JS, Davis-Neulander L, Beauregard A, Laederach A: **Evaluation of the information content of RNA structure mapping data for secondary structure prediction**. *RNA* 2010, **16**:1108-1117.

106. Kasprzak W, Bindewald E, Kim T-J, Jaeger L, Shapiro BA: **Use of RNA structure flexibility data in nanostructure modeling**. *Methods (San Diego, CA)* 2011, **54**:239-250.

Please cite this article in press as: Bida JP, Das R. Squaring theory with practice in RNA design, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.06.003