Biochemistry

Viewpoint

Cite This: Biochemistry 2018, 57, 6129−6131

pubs.acs.org/biochemistry

High-Throughput Explorations of RNA Structural Modularity

Mark A. Boerneke* and Kevin M. Weeks*

Department of Chemistry, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

All levels of RNA structure influence the diverse biological functions of these biomolecules. The primary sequence of an RNA folds into a base-paired secondary structure, which may then form a complex intramolecular tertiary structure, which in turn may interact with an RNA, protein, or small-molecule ligand partner. Much is known about the folding and unfolding energetics of RNA secondary structure, but our understanding of the thermodynamic stability and folding of RNA tertiary structure remains incomplete.

That our understanding of RNA tertiary folding remains so incomplete is perplexing. The principles that underlie RNA folding ought to be simple.[1] There are only four primary RNA nucleotide building blocks, and there are relatively few basic kinds of secondary structures (helices, loops, and junctions). Moreover, the number of ways that RNAs interact to form complex higher-order structures, as visualized to date in high-resolution structures, is also fairly limited. Pioneering biochemical and biophysical studies investigating RNA tertiary structure and folding[1,2] have established numerous principles regarding energetics and folding. These foundational studies have generally focused on one or several variations in sequence, topology, or ionic conditions within a model higher-order RNA structure such as tRNA, ribozymes, other noncoding RNAs, or the small and large ribosomal subunits. Stochastic and molecular dynamics simulations of model RNAs have further defined how topological constraints influence tertiary structure folding and dynamics.[3,4]

A major insight from all of these studies is that RNA structure is modular, meaning that one local structure can be readily substituted with another. For example, a helix can be substituted with another of a different sequence as long as the substituted element is roughly the same "size". In addition, different kinds of through-space tertiary interactions can often be substituted for one another as long as steric and connectivity features are compatible with the overall RNA fold. Indeed, this substitutability is the foundation for the phylogenetic studies that led to the first, largely accurate, models of the ribosomal RNAs. But how modular is RNA? How many different structural assemblies can be formed by a given class of RNA? How is the number of distinct local sequence combinations (which scales as N^4 and can rapidly become very large) related to the total number of distinct structural conformations? To what extent can diverse RNA sequences and topologies accomplish similar tertiary folds?

Recently, Denny et al.[5] began to address these fundamental questions and to develop a predictive model for the modularity of RNA structure by creating a high-throughput biophysical screen to examine the folding energetics of RNA two-way junctions (Figure 1, box). Two-way junctions, also called bulges or internal loops, are key contributors to RNA dynamic behavior and comprise the simplest level of structure more complex than a helix. The approach taken by Denny et al. uses a cleverly repurposed high-throughput sequencing instrument to measure the conformational effects of diverse RNA junction variations on the thermodynamic binding equilibria between two structured RNAs. The authors made use of a known pseudosymmetric tertiary assembly comprised of two RNA tetraloops and their two tetraloop receptors (Figure 1, blue). To form, the two sets of tertiary interactions must be connected by an intervening, mostly helical, linker of the right geometry.

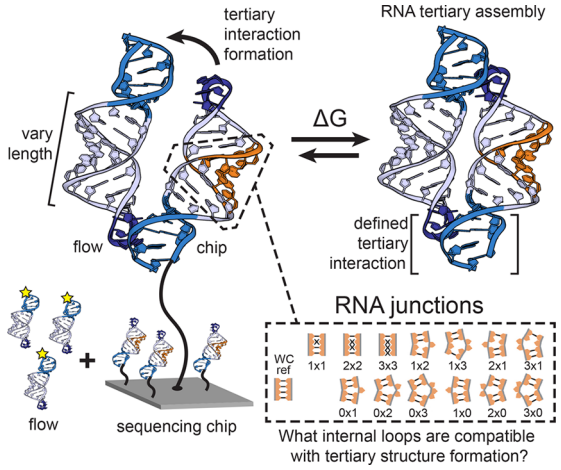

Figure 1. Junctions in RNA tertiary structure formation. The conformational behavior of diverse two-way RNA junctions (boxed) was assessed by inserting junctions into the helix of a known RNA tertiary assembly, comprised of two RNA tetraloops (dark blue) and their two tetraloop receptors (light blue), and measuring the binding equilibrium of tertiary interaction formation. Junctions are defined as *N*x*N*, where *N* indicates the number of noncanonical or unpaired nucleotides. The chip RNAs were attached to a repurposed sequencing chip, and flow RNAs were in solution. Thermodynamic fingerprints (not shown) were generated by studying junctions in multiple structural contexts involving different helix lengths in both flow and chip RNAs and different inserted two-way junctions in the chip helix (orange). RNA junction secondary structure illustrations (boxed) were adapted from ref 5.

To test the effect of RNA two-way junctions on tertiary structure formation, the authors inserted various junction sequences into one half of the tetraloop−receptor motif (Figure 1, orange), which was immobilized on a surface. A second tetraloop−receptor element was free in solution and modified with a fluorophore to enable measurement of the extent of assembly formation. Three RNA helix lengths were used on the "flow" side of the tertiary structure assembly and, notably, ∼1700 RNA junctions were evaluated on the "chip"

Received: September 19, 2018
Published: October 15, 2018

DOI: 10.1021/acs.biochem.8b00999
Biochemistry 2018, 57, 6129−6131

side. The authors then evaluated the thermodynamic effects of inserted junctions in the "chip" RNAs on the stability of their tertiary interactions with the "flow" RNAs. These "thermodynamic fingerprints" were then clustered to identify junctions with similar properties. In some cases, clusters could also be compared with previously determined high-resolution RNA structures to build generalized computational models for local two-way junction structure.

Comparative analysis of the junction thermodynamic fingerprints confirmed known and revealed new principles for how RNA junction sequence and structure govern three-dimensional tertiary folding. For example, junctions with shared topological constraints generally have similar conformational preferences that can also be significantly influenced by primary sequence. Notably, diverse junction sequences and topologies are compatible with the formation of the tetraloop-receptor tertiary interaction. In addition, structural flexibility in a junction can promote tertiary assembly in structurally perturbed contexts.

As expected, junctions with similar numbers and arrangements of unpaired residues clustered into a relatively small number of distinct conformational classes. For the 1 × 1 junction topology, all 112 possible distinct sequence combinations were studied (inclusive of an adjacent Watson−Crick pair), and these clustered into only seven unique conformational classes (Figure 2), revealing strong modularity and interchangeability. For these seven clusters, there are distinct subclasses. Those enriched in purine-purine mismatches (GG, AA, GA, and AG) are distinct from clusters enriched in pyrimidine-pyrimidine mismatches (CC, UU, CU, and UC), and purine-pyrimidine mismatches (AC and CA) are conformationally most similar to Watson−Crick base pairs. Similarly, all possible 144 sequence combinations of the 2 × 2 junctions cluster into only 20 conformational classes, and the 3 × 3 junctions (1728 distinct sequence combinations) cluster into only 49 conformational classes (Figure 2). Within each class of junction, the number of total structures collapses into a much smaller number of modular elements (the clusters) that are characterized by sequence- and position-dependent effects.

A few junctions have conformational behavior highly divergent from that of their simple topological (*NxN*) class. Conversely, numerous junctions from different topological classes had highly similar conformational behavior. These cases highlight the advantages of a high-throughput thermodynamics approach to interrogation of RNA junction conformational behavior and emphasize the limitations of using a topology-only model. Some RNA junctions likely achieve complex folds in yet to be fully understood ways, highlighting the need for further structural studies.

The authors used junction thermodynamic fingerprint data to develop a computational model to predict the folding energetics of previously uncharacterized junctions in tertiary assemblies. The key idea behind this model is that the conformational behavior of specific topological classes might be best represented by a thermodynamically related group of structures, an ensemble, rather than a single structure. Ensembles for this model were constructed on the basis of crystal structures of junctions with similar thermodynamic fingerprints (Figure 2). This ensemble model predicted binding of tertiary assembly partners containing uncharacterized junctions more accurately than models derived from a single static conformation based on a crystal structure. These results suggest that thermodynamics-derived structural ensembles robustly capture the conformational behavior of an RNA junction and will be useful for guiding future efforts to model RNA folding.

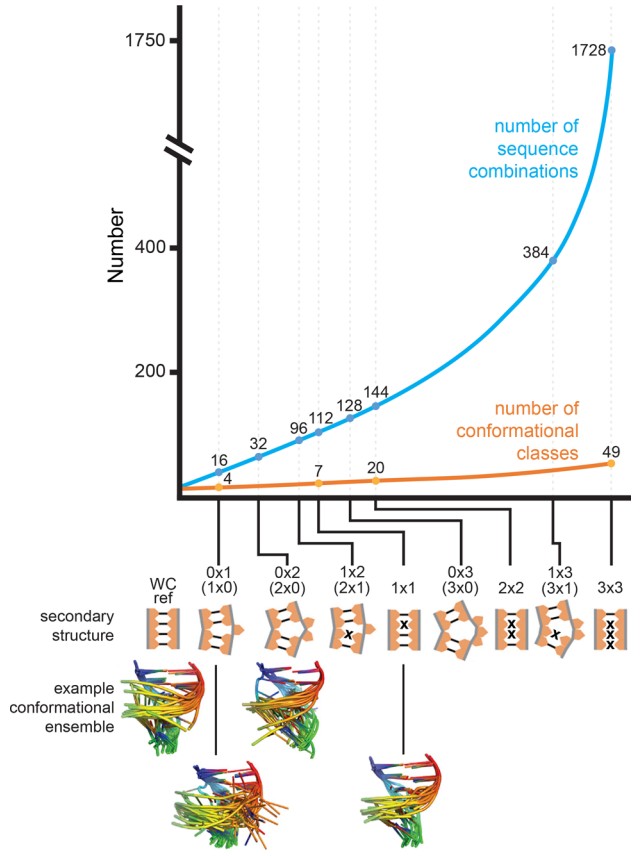

Figure 2. Diverse RNA junction sequence compositions cluster into a small number of distinct conformational classes. The number of sequence combinations is shown for each junction topology (blue). Cluster analysis, performed on select junction topologies, revealed that the large number of sequence combinations converges into relatively few conformational classes with unique thermodynamic and structural properties (orange). The secondary structure is shown for each junction topology; mirror image topologies are shown in parentheses. Select conformational ensembles are shown beneath their corresponding junction topology. RNA junction secondary structure illustrations and example conformational ensembles were adapted from ref 5.

This work demonstrates how using massively parallel thermodynamic measurements of RNA−RNA binding, clustering based on thermodynamic fingerprints, and modeling these groupings as ensembles provides a path forward toward predicting how primary RNA sequences encode the formation of biologically critical RNA tertiary structures. Both RNA secondary and tertiary structure modeling algorithms might be improved by incorporating restraints from thermodynamics-derived ensembles. A welcome result of the analysis by Denny et al. is that, consistent with prior work focusing on just a few motifs, RNA structure is highly modular and there are many fewer conformational classes than sequence classes (Figure 2).

This work investigated the simplest case of nonhelical structures, two-way junctions. Natural next steps will be to study more complex tertiary folds, such as three-way and other complex multihelix junctions, and to evaluate the influence of protein binding partners on RNA conformational behavior. These studies will quickly become very complex, and new

DOI: 10.1021/acs.biochem.8b00999
Biochemistry 2018, 57, 6129−6131

frameworks will likely be needed to keep the total number of conformational classes both representative and manageable. RNA conformational preferences for recognizing small molecules could also be investigated by these methods, enhancing our understanding of structural ensembles that might be best therapeutically targetable in dynamic RNAs. These strategies have the potential to reshape how we think about RNA tertiary structure evolution, conformational dynamics, and ligand binding.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: mboernek@email.unc.edu.
*E-mail: weeks@unc.edu.

ORCID

Mark A. Boerneke: 0000-0003-0092-3203
Kevin M. Weeks: 0000-0002-6748-9985

Funding

M.A.B. is supported by a Ruth L. Kirschstein postdoctoral fellowship (F32 GM128330). K.M.W. is supported by Grants R01 AI068462 (RNA structure methods development) and R35 GM122532 (RNA function discovery).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the many authors that we could not cite due to reference limits.

REFERENCES

(1) Tinoco, I., and Bustamante, C. (1999) How RNA folds. *J. Mol. Biol. 293*, 271−281.

(2) Brion, P., and Westhof, E. (1997) Hierarchy and Dynamics of RNA Folding. *Annu. Rev. Biophys. Biomol. Struct. 26*, 113−137.

(3) Chu, V. B., Lipfert, J., Bai, Y., Pande, V. S., Doniach, S., and Herschlag, D. (2009) Do conformational biases of simple helical junctions influence RNA folding stability and specificity? *RNA 15*, 2195−205.

(4) Mustoe, A. M., Brooks, C. L., and Al-Hashimi, H. M. (2014) Topological constraints are major determinants of tRNA tertiary structure and dynamics and provide basis for tertiary folding cooperativity. *Nucleic Acids Res. 42*, 11792−11804.

(5) Denny, S. K., Bisaria, N., Yesselman, J. D., Das, R., Herschlag, D., and Greenleaf, W. J. (2018) High-Throughput Investigation of Diverse Junction Elements in RNA Tertiary Folding. *Cell 174*, 377−390.e20.

DOI: 10.1021/acs.biochem.8b00999
Biochemistry 2018, 57, 6129−6131