Supplemental Information for:

In silico design and enzymatic synthesis of functional RNA nanoparticles

Kirill A. Afonin¹, Wojciech K. Kasprzak², Eckart Bindewald², Maria Kireeva³, Mathias Viard¹,², Mikhail Kashlev³, Bruce A. Shapiro¹#

¹ Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA
² Basic Science Program, Leidos Biomedical Research, Inc., NCI Center for Cancer Research, Frederick National Laboratory for Cancer Research, Frederick, MD, USA
³ Gene Regulation and Chromosome Biology Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

# - correspondence: Bruce Shapiro, phone/fax 301-846-5536/301-846-5598; shapirbr@mail.nih.gov
Nanostructure Modeling: NanoTiler and RNA2D3D

There are many programs for RNA 3D modeling that can be utilized to varying degrees in modeling of nanostructures/nanoparticles (RNA NPs). These include, for example, FARNA, FARFAR, iFOLDRNA, MC-Sym, NAST, RNAComposer, RSIM, Discovery Studio, Swiss PDB viewer, Assemble\(^1\). We mostly rely on our own programs, NanoTiler and RNA2D3D, when modeling RNA NPs (see Figure 4)\(^10,11\). NanoTiler has been developed specifically for NP model building, and it can implement all three strategies described in the main text (helix-driven, junction-driven, and rules-driven automated shape discovery). The program has a user graphical interface or GUI (Figure 4a) for interactive work; however, we predominantly use its script interpreting capability, which allows for computationally intensive tasks to be performed automatically\(^12,13\). Two examples of scripted applications are tecto-square closure exploration based on junction samples from molecular dynamics trajectories (MD) combined with linker helices and optimization of helix placement and corner bridging in the nano-cube models\(^12,13\). Importantly, using the helix-driven strategy, NanoTiler can generate junctions \textit{de novo} when no available database structures can satisfy the design. This capability was used in the modeling of nanocubes\(^14\). The program can uniformly apply distortions to the helices in order to achieve structure closure. In the tecto-square characterization study we verified that the helix distortions introduced by NanoTiler in order to affect the tectosquare closure were within the range of distortions generated in the MD simulation for the corresponding regions\(^12\). An additional advantage of using the NanoTiler-generated helix distortions in this study was that the dynamic state search space was not limited to the output of the relatively short MD simulations of the tecto-square building blocks that, in principle, could have yielded no dynamic states resulting in the full structure closure. Once the 3D structure design is considered satisfactory (validated structurally), the NanoFolder program (and web server) can be used to perform sequence design for the multiple RNA strands comprising the full structure, with the aim of optimizing their experimental self-assembly\(^15\). Once the sequences are designed, NanoTiler can be used again to perform 3D mutations on the initial dummy sequence-based structure model to create its final version that can be further computationally characterized, as described in the Nanostructure Characterization section in the main text\(^16-21\).

RNA2D3D is an interactive program with a flexible GUI (see Figure 4b)\(^11\). It accepts an RNA sequence and a corresponding secondary structure descriptor, which may include pseudoknots, and automatically and quickly generates a first-pass, approximate three-dimensional model based on an idealized geometry approach, including heuristic stacking of pseudoknots with adjacent helices. The program also allows one to define interactions between multiple building blocks, such as the monomers forming a tecto-square or multiple tecto-squares forming a mesh\(^11,12,22\). Model refinements usually need to follow this step due to the inherent limitations of the purely geometric approach. However, the first pass model is sufficiently interpretable to give the user clear indications of how to proceed further; for example, how to disentangle structurally colliding sub-domains\(^11,12,23,24\). The RNA2D3D tools facilitate 3D manipulations of selected parts. The program’s graphical interface allows the user to interact with the model via its 2D or 3D depictions, displayed side-by-side (see Figure 4b) or one at-a-time, whichever makes it easier to perform operations. Selected helices can be coaxially stacked, individual base pairs opened or added, and the helical geometry extended into internal (bulge) and multi-branch loops (an operation called “compactification”). Structural motifs from databases, such as RNAJunction or the PDB, for example, or single-stranded fragments can be
substituted in place of equivalent model sub-domains. Modeling of kissing loops, even in the absence of equivalent motifs in a database, can be also performed by the program, and we have taken advantage of this option in our NP and natural structure modeling. The whole 3D model, its user-defined sub-domains or elements, such as single stranded fragments, can be subjected to energy minimizations and short molecular dynamics runs in order to correct and refine the model. In general, however, beyond generating the initial 3D coordinates, the program provides no automated, knowledge-based model development, and, as such, it is a tool meant for a user with some modeling expertise.

Both, NanoTiler and RNA2D3D, as well as many other software packages developed in our laboratory, are available at: http://binkley2.ncifcrf.gov/users/bshapiro/software.html.

Synergy between the Computational and the Experimental Aspects of the NP Design Process

An important aspect of the RNA NPs design process (the pipeline is under development) is the feedback between the computational and the experimental parts, which allows refining both the algorithmic methodologies and the experimental protocols.

One example comes from the work on nano-cubes characterization. The experimental results indicated that the assembled cubes’ sizes (hydrodynamic radii measured with DLS) were larger than those of the cube models. Given this information, we first explored the maximum stable sizes of the models (no changes in base pairing after energy minimization) with the aid of NanoTiler, but the resized models were still smaller than the experimental measurements. Therefore we explored the models’ flexibilities with the aid of Anisotropic Network Model simulations (ANM). The apparent size changes due to the distortions predicted by the ANM brought the computational and the experimental NP size measurements into agreement (see Figure 5). The ANM results also offered more insight into the observed experimental assembly yield differences and the measured melting temperature differences for the cube variants.

Another example is the development of the sequence design program. The importance of constraining designed sequences by more than thermodynamic folding predictions alone became only clear after discussions with the experimental collaborators of the project. This led to a set of scoring terms such as the avoidance of repetition of the same type of nucleotide and the favoring of G:C base pairs at helical ends. The evolved set of rules is incorporated in our NanoFolder program.

The final example of the synergy between our computational an experimental research involves the exploration of the interactions of the bolaamphiphile delivery agents and siRNAs. The computational results obtained by molecular dynamics of the bolaamphiphiles (GLH-19 and GLH-20, both of which have positively charged head groups) alone or in conjunction with siRNAs explained and enhanced our understanding of how these entities interacted with each other. For example, the experimentally determined better protected siRNAs afforded by bola GLH-19 compared to GLH-20 can be explained by the higher binding affinity of GLH-19 to the siRNA partly due to the increased number of chloride ions found around GLH-20 using molecular dynamics. The head groups of GLH-20 are more deeply buried than GLH-19, thereby...
trapping the ions. This in turn neutralized some of the positive charges associated with GLH-20 thus reducing the affinity of GLH-20 to the siRNA.
Acknowledgements

This publication has been funded in part with federal funds from the Frederick National Laboratory for Cancer Research, National Institutes of Health, under Contract HHSN261200800001E. This research was additionally supported in part by the Intramural Research Program of the National Institutes of Health, Center for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

1. Das, R.; Baker, D., Automated De Novo Prediction of Native-Like RNA Tertiary Structures. Proc Natl Acad Sci U S A 2007, 104, 14664-9.
2. Das, R.; Karanicolas, J.; Baker, D., Atomic Accuracy in Predicting and Designing Noncanonical RNA Structure. Nat Methods 2010, 7, 291-4.
3. Parisien, M.; Major, F., The Mc-Fold and Mc-Sym Pipeline Infers RNA Structure from Sequence Data. Nature 2008, 452, 51-5.
4. Sharma, S.; Ding, F.; Dokholyan, N. V., Ifoldrna: Three-Dimensional RNA Structure Prediction and Folding. Bioinformatics 2008, 24, 1951-2.
5. Jonikas, M. A.; Radmer, R. J.; Laederach, A.; Das, R.; Pearlman, S.; Herschlag, D.; Altman, R. B., Coarse-Grained Modeling of Large RNA Molecules with Knowledge-Based Potentials and Structural Filters. RNA 2009, 15, 189-99.
6. Popenda, M.; Szachniuk, M.; Antczak, M.; Purzycka, K. J.; Lukasiak, P.; Bartol, N.; Blazewicz, J.; Adamiak, R. W., Automated 3d Structure Composition for Large Rnas. Nucleic Acids Res 2012, 40, e112.
7. Bida, J. P.; Maher, L. J., 3rd, Improved Prediction of RNA Tertiary Structure with Insights into Native State Dynamics. RNA 2012, 18, 385-93.
8. Kaplan, W.; Littlejohn, T. G., Swiss-Pdb Viewer (Deep View). Brief Bioinform 2001, 2, 195-7.
9. Jossinet, F.; Ludwig, T. E.; Westhof, E., Assemble: An Interactive Graphical Tool to Analyze and Build RNA Architectures at the 2d and 3d Levels. Bioinformatics 2010, 26, 2057-9.
10. Bindewald, E.; Grunewald, C.; Boyle, B.; O'Connor, M.; Shapiro, B. A., Computational Strategies for the Automated Design of RNA Nanoscale Structures from Building Blocks Using Nanotiler. J Mol Graph Model 2008, 27, 299-308.
11. Martinez, H. M.; Maizel, J. V., Jr.; Shapiro, B. A., Rna2d3d: A Program for Generating, Viewing, and Comparing 3-Dimensional Models of RNA. J Biomol Struct Dyn 2008, 25, 669-84.
12. Kasprzak, W.; Bindewald, E.; Kim, T. J.; Jaeger, L.; Shapiro, B. A., Use of RNA Structure Flexibility Data in Nanostructure Modeling. Methods 2011, 54, 239-50.
13. Afonin, K. A.; Kasprzak, W. K.; Bindewald, E.; Puppala, P. S.; Diehl, A. R.; Hall, K. T.; Kim, T. J.; Zimmermann, M. T.; Jernigan, R. L.; Jaeger, L., et al., Computational and Experimental Characterization of RNA Cubic Nanoscaffolds. Methods 2013.
14. Afonin, K. A.; Bindewald, E.; Yaghoubian, A. J.; Voss, N.; Jacovetty, E.; Shapiro, B. A.; Jaeger, L., In Vitro Assembly of Cubic RNA-Based Scaffolds Designed in Silico. *Nat Nanotechnol* 2010, 5, 676-82.
15. Bindewald, E.; Afonin, K.; Jaeger, L.; Shapiro, B. A., Multistrand RNA Secondary Structure Prediction and Nanostructure Design Including Pseudoknots. *ACS Nano* 2011, 5, 9542-51.
16. Case, D. A.; Darden, T. A.; Cheatham, T. E., 3rd; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M., et al., Amber12. In University of California: San Francisco, 2012.
17. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J Chem Phys* 1995, 103, 8577-93.
18. Bahar, I.; Erman, B.; Haliloglu, T.; Jernigan, R. L., Efficient Characterization of Collective Motions and Interresidue Correlations in Proteins by Low-Resolution Simulations. *Biochemistry* 1997, 36, 13512-23.
19. Wang, Y.; Rader, A. J.; Bahar, I.; Jernigan, R. L., Global Ribosome Motions Revealed with Elastic Network Model. *J Struct Biol* 2004, 147, 302-14.
20. Wang, Y.; Jernigan, R. L., Comparison of Trrna Motions in the Free and Ribosomal Bound Structures. *Biophys J* 2005, 89, 3399-409.
21. Zimmermann, M. T.; Kloczkowski, A.; Jernigan, R. L., Mavens: Motion Analysis and Visualization of Elastic Networks and Structural Ensembles. *BMC Bioinformatics* 2011, 12, 264.
22. Chworos, A.; Severcan, I.; Koyfman, A. Y.; Weinkam, P.; Oroudjev, E.; Hansma, H. G.; Jaeger, L., Building Programmable Jigsaw Puzzles with RNA. *Science* 2004, 306, 2068-72.
23. Gao, F.; Gulay, S. P.; Kasprzak, W.; Dinman, J. D.; Shapiro, B. A.; Simon, A. E., The Kissing-Loop T-Shaped Structure Translational Enhancer of Pea Enation Mosaic Virus Can Bind Simultaneously to Ribosomes and a 5' Proximal Hairpin. *J Virol* 2013, 87, 11987-2002.
24. Kasprzak, W. K.; Shapiro, B. A., Computational Prediction and Modeling Aid in the Discovery of a Conformational Switch Controlling Replication and Translation in a Plus-Strand Virus. In *RNA 3d Structure Analysis and Prediction, Nucleic Acids and Molecular Biology*, Leontis, N.; Westhof, E., Eds. Springer: New York, 2012; pp 119-42.
25. Kim, T.; Afonin, K.A.; Viard, M.; Koyfman, A.Y.; Sparks, S.; Heldman, E.; Grinberg, S.; Linder, C.; Blumenthal, R.P.; Shapiro, B.A., In Silico, In Vitro, and In Vivo Studies Indicate the Potential Use of Bolaamphiphiles for Therapeutic siRNAs Delivery. *Mol Ther Nucleic Acids* 2013, 2, e80.