

similarly distinct structures appeared, representing yet earlier phases. The most notable intermediate substructure was a pairing between the 5' end of the molecule and nucleotides in the antisense target structure [Shapiro *et al.*, 2001c]. This interaction would not only prevent the translational activating interaction from forming, but it would directly prevent the antisense target structure from forming as well. Thus, unlike the other mechanisms of inactivation, this novel pairing provided a direct, steric block to both functions of the active structure. Moreover, phylogenetic analysis revealed that the interaction was extremely conserved across the entire *hok* family by significant numbers of compensatory base changes within the stem.

## 1.5 Conclusions

The massively parallel Genetic Algorithm seems to have great potential for the exploration of RNA folding pathways. When conditions for folding within the algorithm are adjusted to match those of experimental environments, the algorithm appears to report similar results to the original experiments. The strength of this approach is that, once one has correctly adjusted the computational system to match the biological system, the simulation can provide a wealth of information that would be difficult to gain from experiments alone. The algorithm offers the opportunity to catch detailed glimpses of intermediate structures that are challenging to capture and directly analyze experimentally. One could employ this capability either for verification or for prediction of the dynamic structural details of a molecule's behavior. In actuality, a combination of both approaches seems most useful. The most effective use of the algorithm is not in a vacuum, but as applied to a system about which there already exists some information. Algorithmic, experimental, and phylogenetic analyses can then mutually support and extend one another.

The methods described here illustrate the importance of having a variety of effective ways of visualizing the same data from many perspectives. The GA deals with a complex system and generates a large amount of data very rapidly. The various levels of abstraction and compression of this data into numerical and graphical representations are crucial for making sense of it all. Stem Trace has proven to provide some of the most valuable of these representations, and has been indispensable for the generation of these results.

Many hold the view that the folding process of RNA is hierarchical, that primary sequence first defines secondary structure, and that tertiary structure subsequently forms as a consequence of secondary structure. If this is the case, analyses of such systems as are described here can be carried out on each level independently, and maintain validity. However, full understanding of a system is only possible with the integration of analyses on all three levels. Thus, future directions with the GA should include an increase in its ability to consider the contributions of tertiary structure, as well as the analysis of GA-generated data by methods designed for tertiary structure analysis.

## Acknowledgment

This publication has been funded in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. N01-C0-12400.

## References

- Baumstark, T. and Riesner, D. (1995) "Only one of four possible secondary structures of the central conserved region of potato spindle tuber viroid is a substrate for processing in a potato nuclear extract." *Nucl Acids Res.* **23**(21), 4246-4254.
- Baumstark, T., Schröder, A.R.W. and Riesner, D. (1997) "Viroid processing: switch from cleavage to ligation is driven by a change from a tetraloop to a loop E conformation." *EMBO J.* **16**(3), 599-610.
- Branch, A.D., Benenfeld, B.J. and Robertson, H.D. (1985) "Ultraviolet light-induced crosslinking reveals a unique region of local tertiary structure in potato spindle tuber viroid and HeLa 5S RNA." *Proc. Natl. Acad. Sci.* **82**(10), 6590-6594.
- Branch, A.D. and Robertson, H.D. (1984) "A replication cycle for viroids and other small infectious RNA's." *Science* **223**(4635), 450-455.
- Diener, T.O. (1979a) "Viroids: structure and function." *Science* **205**(4409), 859-866.
- Diener, T.O. (1979b) *Viroids and Viroid Diseases*. John Wiley & Sons, Inc., New York.

- Diener, T.O. (1999) "Viroids and the nature of viroid diseases." *Arch. Virol. Suppl.* **15**, 203-220.
- Flores, R., Randles, J.W., Bar-Joseph, M. and Diener, T.O. (1998) "A proposed scheme for viroid classification and nomenclature." *Arch. Virol.* **14(3)**, 623-630.
- Franch, T. and Gerdes, K. (1996) "Programmed cell death in bacteria: translational control by mRNA end-pairing." *Molecular Microbiology* **21(5)**, 1049-1060.
- Franch, T., Gulyaev, A.P. and Gerdes, K. (1997) "Programmed cell death by hok/sok of plasmid R1: Processing at the hok mRNA 3' triggers structural rearrangements that allow translation and antisense RNA binding." *J. Mol. Biol.* **273**, 38-51.
- Gast, F., Kempe, D. and Sanger, H.L. (1998) "The dimerization domain of potato spindle tuber viroid, a possible hallmark for infectious RNA." *Biochemistry* **37**, 14098-14107.
- Gast, F., Kempe, D., Spieker, R.L. and Sanger, H.L. (1996) "Secondary structure probing of potato spindle tuber viroid (PSTVd) and sequence comparison with other small pathogenic RNA replicons provides evidence for central non-canonical base pairs, large A-rich loops, and a terminal branch." *J. Mol. Biol.* **262**, 652-670.
- Gerdes, K., Gulyaev, A.P., Franch, T., Pedersen, K. and Mikkelsen, N.D. (1997) "Antisense RNA-regulated programmed cell death." *Annu. Rev. Genet.* **31**, 1-31.
- Gerdes, K., Nielsen, A.K., Thursted, P. and Wagner, E.G.H. (1992) "Mechanism of killer gene reactivation: antisense RNA mediated RNaseIII cleavage ensures rapid turn-over of the hok, srnB, and pndA effector mRNAs." *J. Mol. Biol.* **226**, 637-649.
- Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H. and Sanger, H.L. (1978) "Nucleotide sequence and secondary structure of potato spindle tuber viroid." *Nature* **273**, 203-208.
- Gulyaev, A.P., Franch, T. and Gerdes, K. (1997) "Programmed cell death by hok/sok of plasmid R1: Coupled nucleotide covariations reveal a phylogenetically conserved folding pathway in the hok family of mRNAs." *J. Mol. Biol.* **273**, 26-37.
- Hecker, R., Wang, Z., Steger, G. and Riesner, D. (1988) "Analysis of RNA structures by temperature gradient gel electrophoresis: viroid replication and processing." *Gene* **72**, 59-74.

- Ishikawa, M., Meshi, T., Takeshi, O., Yoshimi, O., Teruo, S., Ichiro, U. and Shikata, E. (1984) "A revised replication cycle for viroids: The role of longer than unit length RNA in viroid replication." *Mol. Gen. Genet.* **196**(3), 421-428.
- Kasprzak, W. and Shapiro, B.A. (1999) "Stem Trace: An interactive visual tool for comparative RNA structure analysis." *Bioinformatics* **15**(1), 16-31.
- Keese, P., Visvader, J.E. and Symons, R.H. (1988) "Sequence variability in plant viroid RNAs." In *RNA Genetics*, eds. Domingo, E., Holland, J. J. and Ahlquist, P., vol. 3, p. 71-98. CRC Press, Boca Raton, FL.
- Mathews, D.H., Sabina, J., Zuker, M. and Turner, D.H. (1999) "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure." *J. Mol. Biol.* **288**, 911-940.
- Nagel, J.H.A., Gulyaev, A.P., Gerdes, K. and Pleij, C.W.A. (1999) "Metastable structures and refolding kinetics in hok mRNA of plasmid R1." *RNA* **5**, 1408-1419.
- Riesner, D. (1990) "Structure of viroids and their replication intermediates. Are thermodynamic domains also functional domains?" *Virology* **1**, 83-99.
- Riesner, D., Henco, K., Rokohl, U., Klotz, G., Kleinschmidt, A. K., Domdey, H., Jank, P., Gross, H.J. and Sanger, H.L. (1979) "Structure and structure formation of viroids." *J. Mol. Biol.* **133**, 85-115.
- Rivas, E. and Eddy, S.R. (1999) "A dynamic programming algorithm for RNA structure prediction including pseudoknots." *J. Mol. Biol.* **285**, 2053-2068.
- Shapiro, B.A., Bengali, D., Kasprzak, W. and Wu, J-C. (2001b) "Computational insights into RNA folding pathways: Getting from here to there." In *Proceedings of the Atlantic Symposium on Computational Biology and Genome Systems and Technology*.
- Shapiro, B.A., Bengali, D., Kasprzak, W. and Wu, J-C. (2001c) "RNA folding pathway functional intermediates: Their prediction and analysis." *J. Mol. Biol.* **312**(1), 27-44.
- Shapiro, B.A. and Kasprzak, W. (1996) "STRUCTURELAB: A heterogeneous bioinformatics system for RNA structure analysis." *Journal of Molecular Graphics* **14**, 194-205.
- Shapiro, B.A. and Navetta, J. (1994) "A massively parallel genetic algorithm for RNA secondary structure prediction." *The Journal of Supercomputing* **8**, 195-207.

- Shapiro, B.A. and Wu, J-C. (1996) "An annealing mutation operator in the genetic algorithm for RNA folding." *CABIOS* **12**(3), 171-180.
- Shapiro, B.A. and Wu, J-C. (1997) "Predicting RNA H-type pseudoknots with the massively parallel genetic algorithm." *Bioinformatics* **13**(4), 459-471.
- Shapiro, B.A., Wu, J-C., Bengali, D. and Potts, M.J. (2001a) "The massively parallel genetic algorithm for RNA folding: MIMD implementation and population variation." *Bioinformatics* **17**(2), 137-148.
- Steger, G., Baumstark, T., Mörchen, M., Tabler, M., Tsagris, M., Sängler, H.L. and Riesner, D. (1992) "Structural requirements for viroid processing by RNase T1." *J. Mol. Biol.* **227**, 719-737.
- Tsagris, M., Tabler, M., Mülbach, H. and Sängler, H.L. (1987) "Linear oligomeric potato spindle tuber viroid RNAs are accurately processed in vitro to the monomeric circular viroid roper when incubated with a nuclear extract form healthy potato cells." *EMBO J.* **6**(8), 2173-2183.
- Wu, J-C. and Shapiro, B.A. (1999) "A Boltzmann filter improves the prediction of RNA folding pathways in a massively parallel genetic algorithm." *Journal of Biomolecular Structure and Dynamics* **17**(3), 581-595.
- Zuker, M. (1989) "On finding all suboptimal foldings of an RNA molecule." *Science* **244**, 48-52.

## Authors' Addresses

Bruce A. Shapiro, Laboratory of Experimental and Computational Biology, The NCI Center for Cancer Research, NCI-Frederick, National Institutes of Health, Building 469, Room 150, Frederick, MD 21702, USA.  
Email: bshapiro@ncifcrf.gov.

David Bengali, Laboratory of Experimental and Computational Biology, The NCI Center for Cancer Research, NCI-Frederick, National Institutes of Health, Building 469, Room 150, Frederick, MD 21702, USA.

Wojciech Kasprzak, Intramural Research Support Program, SAIC, NCI-Frederick, Frederick, MD 21702, USA.

Jin Chu Wu, Intramural Research Support Program, SAIC, NCI-Frederick, Frederick, MD 21702, USA.