Design Tools for Artificial DNA Structures

Erik Winfree

Department of Molecular Biology
Princeton University, and
Departments of Computer Science and
Computation and Neural Systems
California Institute of Technology
email: winfree@hope.caltech.edu

Research in DNA-based computation [1] and DNA nanotechnology [2] relies on the design of artificial DNA sequences that fulfill certain logical and structural requirements. For example, Lipton's approach to DNA-based computation [3] requires combinatorial libraries of single-stranded DNA with little secondary structure, but whose sequences contain N target domains that can be reliably and specifically bound by hybridization probes. Multi-stranded branched DNA structures used in DNA nanotechnology require sequences that are unlikely to form undesired secondary structures; a reasonable heuristic is sequence symmetry minimization [4]. The self-assembly approach to DNA-based computing [5, 6] requires sequence constraints of both types. We are developing a suite of software tools, written in MATLAB, that assist in evaluating potential DNA sequence designs according to these criteria and automatically optimizing the designs. Routines include, for example, estimates of thermodynamic stability of secondary structures, predictions of optimal secondary structure (following *mfold* [7]), and calculation of repeated subsequences. In this poster we will motivate the DNA design problem, discuss the importance of negative (as opposed to solely positive) design for DNA, and explain the criteria and algorithms we use.

References

[1] Leonard M. Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 266:1021–1024, November 11, 1994.

[2] N. C. Seeman, H. Wang, X. P. Yang, F. R. Liu, C. D. Mao, W. Q. Sun, L. Wenzler, Z. Y. Shen, R. J. Sha, H. Yan, M. H. Wong, P. Sa-Ardyen, B. Liu, H. X. Qiu, X. J. Li, J. Qi, S. M. Du, Y. W. Zhang, J. E. Mueller, T. J. Fu, Y. L. Wang, and J. H. Chen. New motifs in DNA nanotechnology. *Nanotechnology*, 9(3):257–273, 1998.

[3] Richard J. Lipton. DNA solutions of hard computational problems. *Science*, 268:542–544, 1995.

[4] Nadrian C. Seeman. *De novo* design of sequences for nucleic acid structural engineering. *Journal of Biomolecular Structure & Dynamics*, 8(3):573–581, 1990.

[5] Erik Winfree. On the computational power of DNA annealing and ligation. In Richard J. Lipton and Eric B. Baum, editors, *DNA Based Computers: DIMACS workshop, April 4, 1995*, volume 27, pages 199–221, Providence, RI, 1996. American Mathematical Society.

[6] Erik Winfree, Furong Liu, Lisa A. Wenzler, and Nadrian C. Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394:539–544, 1998.

[7] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner. Expanded sequence dependence of thermodynamic parameters provides robust prediction of RNA secondary structure. *Journal of Molecular Biology*, 288:911–940, 1999.