# Rule-based modeling of DNA multi-strand dynamics

Zoltan Karoly[1], Frankie Spencer[2], Iris Tusa[3], Alexandru Amarioarei[3], Gefry Barad[3], Corina Itscus[3], Andrei Paun[3], Mihaela Paun[3] and Eugen Czeizler[2,3]

[1] *Department of Computer Science, Tallinn University of Technology, Tallinn, Estonia*
[2] *Department of Information Technology, Åbo Akademi University, Turku, Finland*
[3] *National Institute for Research and Development of Biological Sciences, Bucharest, Romania*
*Corresponding Author's Email: eugen.czeizler@abo.fi*

The dynamics of nucleic-acids dynamical systems is intrinsically based on local interaction. The major acting mechanisms are that of Watson-Crick complementarity on one-hand, generating binding events, and thermal energy on the other, generating random motion and un-binding. It is thus predictable that such systems could be captured successfully by computational modeling paradigms based on local interactions, such as the rule-based modeling methodology [1, 2].

Using the BioNetGen Language (BNGL) formalism [3] and the NFsim computational platform [4], we created a computational modeling framework for DNA multi-strand dynamics. The system takes as input populations of isolated (or partially bounded) oriented single-stranded DNA molecules (ssDNAs), and simulates their binding and dissociation reactions. At the core of this model implementation lie 9 binding and un-binding local interaction rules, each with its own kinetic rate constant, and each implemented through one or several rule-based reactions. Using these reactions/rules, we capture the dynamics of the DNA-dynamical system, by modeling: - the initial binding of short toeholds, here implemented as length-3 short complementary and opposite-oriented sub-sequences, - the "breathing" dynamics in-between bounded ssDNAs, - random un-binding events (between one pair of nucleotides), - as well as the un-binding of loosely connected ssDNAs.

Since the BNGL formalism is not suited for tracking and reporting the global mapping of the components within a heterogeneous complex, as it focuses on the count of local patterns, we have developed distinct subroutines for this purpose. Thus, the content of the dynamical system is unloaded at various time-points within the simulation, and it is re-assembled for visualization and further numerical analysis. A simple 2D graphical representation of the assembled complexes is also generated, where one can track the ssDNAs within the complex as well as all the binding interactions. Note, the strands are not displayed as per their relative positions, but rather listed one under another. Nevertheless, the visualization is highly useful for assessing the overall assembly status and the stability of one complex.

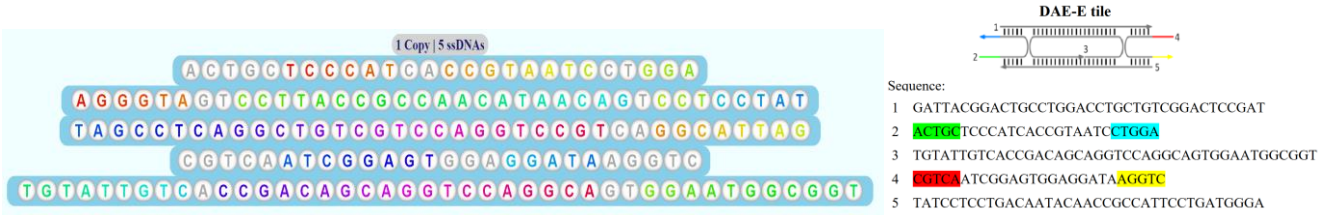


**Figure 1**. Left: fraction of the generated output of a DNA dynamical system simulation consisting of 5 ssDNAs which assemble into a DAE-E tile; nucleotides that are bound by hydrogen bonds are colored similarly. Right: the design scheme and nucleotide sequences of the 5 strands assembling the DNA tile; taken from [5].

Different from other computational modelling frameworks for DNA strand assembly, we can modify some of the systems parameters, such as the temperature of the system, during the system simulation. Thus, we can model also an entire annealing process for the formation of a DNA structure. Many other parameters can also be specified (or adjusted mid-simulation); e.g., we can define specific binding/un-binding kinetic rates for each different length-3 subsequence.

## References

[1] A. Mohammed E. Czeizler, E. Czeizler, Theoret. Comp. Sci. 2017, 701, 203-215.
[2] A. Amarioarei et al. Int. J. of Unconv. Comp. 2017, 13, 399-415.
[3] L. A. Harris, et al. Bioinformatics, 2016, 32, 3366–3368 (2016).
[4] MW Sneddon, JR Faeder and T Emonet, Nature Methods, 2011, 8(2), 177-83.
[5] S. Jiang, F. Hong, H. Hu, H Yan, Y. Liu, ACS Nano, 2017, 11(9), 9370–9381.