Abstracts for the

28th International Conference on DNA Computing and Molecular Programming (DNA28)

August 8-12, 2022

University of New Mexico

Albuquerque, NM, USA

SANTA FE INSTITUTE

Foreword

This is the book of abstracts for the 28th International Conference on DNA Computing and Molecular Programming, held on August 8-12, 2022 at the University of New Mexico in Albuquerque, NM, USA.

Research in DNA computing and molecular programming draws together mathematics, computer science, physics, chemistry, biology, and nanotechnology to address the analysis, design, and synthesis of information-based molecular systems. This annual meeting is the premier forum where scientists with diverse backgrounds come together with the common purpose of using interdisciplinary approaches to advance molecular computation. Continuing this tradition, the 28th International Conference on DNA Computing and Molecular Programming (DNA28), under the auspices of the International Society for Nanoscale Science, Computation, and Engineering (ISNSCE), will focus on the most recent experimental and theoretical results that promise the greatest impact.

As this is the first in-person iteration of the DNA conference since 2019, the local organizers of DNA28 would like to express their gratitude to all contributors, reviewers, attendees, and members of the steering and program committees for making this gathering possible. We would also like to thank our industrial and institutional sponsors: Integrated DNA Technologies, Illumina, ISNSCE, NSF, New Mexico INBRE, the New Mexico Consortium, the Santa Fe Institute, UNM Department of Computer Science, UNM Vice President for Research, and UNM School of Engineering.

Enjoy the conference!

Darko Stefanovic
Matthew Lakin
DNA28 Organizing Committee Co-Chairs

Monday August 8th

8am	Check-in opens
8:30-9am	Breakfast
9–9:30am	Main conference welcome and introduction
9:30–10:30am	KEYNOTE TALK: Models of self-assembling systems: varying levels of abstraction and objectives (pg 2) Matthew Patitz (University of Arkansas, USA)
10:30–11am	Break
11–11:30am	Covered DNA core tiles for robust tuning of spurious nucleation (pg 3) Trent Rogers, Constantine Evans and Damien Woods
11:30am–12pm	SAT-assembly: Designing minimum complexity building blocks to assemble arbitrary shapes (pg 4) Joakim Bohlin, Ard Louis, Andrew Turberfield and Petr Sulc
12–2pm	Lunch (on your own)
2–2:30pm	On Turedo Hierarchies and Intrinsic Universality (pg 5) Samuel Nalin and Guillaume Theyssier (eligible for best student presentation award)
2:30–3pm	Computing Real Numbers with Large-Population Protocols Having a Continuum of Equilibria (pg 6) Xiang Huang and Rachel Huls
3–3:30pm	Break
3:30–3:45pm	Fridge Compiler: automated synthesis of molecular circuits (pg 7) Lancelot Walthieu, Gus Smith, Luis Ceze and Chris Thachuk
3:45–4pm	Single-pass transmembrane double-stranded DNA with functional toeholds for non-destructive intra- and extravesicular nucleic acid target recognition (pg 8) Gde Bimananda Mahardika Wisna, Nirbhik Acharya, Tarushyam Mukherjee, Ranjan Sasmal, Hao Yan and Rizal Hariadi
4–4:30pm	Fast and robust strand displacement cascades via systematic design strategies (pg 9) Tiernan Kennedy, Cadence Pearce and Chris Thachuk (eligible for best student presentation award)
4:30–5pm	A cooperative DNA catalyst (pg 10) Dallas Taylor, Samuel Davidson and Lulu Qian (eligible for best student presentation award)
5:30–7:30pm	Poster session 1

Tuesday August 9th

8:30-9am	Breakfast
9–10am	KEYNOTE TALK: Dynamic control of biomolecular phase separation (pg 12) Elisa Franco (University of California Los Angeles, USA)
10–10:30am	Exploring Material Design Space with a Deep-Learning Guided Genetic Algorithm (pg 13) Kuan-Lin Chen and Rebecca Schulman (eligible for best student presentation award)
10:30–11am	Break
11–11:30am	Protecting heterochiral DNA nanostructures against exonuclease-mediated degradation (pg 14) Tracy Mallette and Matthew R. Lakin (eligible for best student presentation award)
11:30–12pm	DNA strand-displacement temporal logic circuits (pg 15) Anna Lapteva, Namita Sarraf and Lulu Qian (eligible for best student presentation award)
12–2pm	Lunch (on your own)
2–2:15pm	Scaling up reusable DNA circuits using heat as a universal energy source (pg 16) Tianqi Song and Lulu Qian
2:15–2:30pm	Highly Parallel Synthesis, Computation, and Readout of DNA Logic Gates using High-Throughput Synthesis and Sequencing (pg 17) Chandler Petersen, Samantha Borje, Gourab Chatterjee, Yuan-Jyue Chen and Georg Seelig
2:30–3:30pm	KEYNOTE TALK: Hourglass emergence + collective computation in nature (pg 18) Jessica Flack (Santa Fe Institute, USA)
3:30–4pm	Break
4–4:30pm	Growth dynamics: Precisely controlled self-assembly order of DNA tile nanostructures (pg 19) Constantine Evans, David Doty and Damien Woods
4:30–5pm	Universal Shape Replication Via Self-Assembly With Signal-Passing Tiles (pg 20) Andrew Alseth, Matthew Patitz and Daniel Hader
5–5:30pm	Algorithmic DNA origami: Scaffolded DNA computation in one and two dimensions (pg 21) Tristan Stérin, Abeer Eshra and Damien Woods (eligible for best student presentation award)
5:30–7:30pm	Poster session 2

Wednesday August 10th

8:30-9am	Breakfast
9–10am	KEYNOTE TALK: Strand exchange reactions for analysis and computation (pg 24) Andrew Ellington (University of Texas at Austin, USA)
10–10:30am	Toehold-Mediated Strand Displacement in Random Sequence Pools (pg 25) Thomas Mayer, Lukas Oesinghaus and Friedrich C. Simmel (eligible for best student presentation award)
10:30–11am	Parallel molecular computation on digital data stored in DNA (pg 26) Boya Wang, Siyuan S Wang, Cameron Chalk, Andrew D Ellington and David Soloveichik
11–11:30am	Break
11:30am–12pm	Rule-of-thumb-free geometry-driven design of arbitrary complex curved DNA origami with ENSnano (pg 27) Nicolas Levy, Allan Mills, Gaétan Bellot and Nicolas Schabanel (eligible for best student presentation award)
12–12:30pm	A Coupled Reconfiguration Mechanism for Single-Stranded DNA Strand Displacement Systems (pg 28) Hope Amber Johnson and Anne Condon
12:30–1pm	Supervised learning in DNA-based winner-take-all neural networks (pg 29) Kevin Cherry and Lulu Qian (eligible for best student presentation award)
1–1:30pm	Conference group photograph
1:30–6pm	Excursion to Old Town Albuquerque
6–10pm	Conference dinner at Explora

Thursday August 11th

8:30-9am	Breakfast
9–10:30am	TULIP AWARD KEYNOTE TALK: Computational Design of Nucleic Acid Circuits (pg 32) Andrew Phillips
10:30–11am	Break
11–11:30am	Metal-Mediated Molecular Programming in DNA (pg 33) Simon Vecchioni, Brandon Lu, Yoel Ohayon, Karol Woloszyn, Chengde Mao, James Canary, Nadrian Seeman and Ruojie Sha
11:30am–12:30pm	SPECIAL PANEL: A Tribute to Ned Seeman
12:30–2pm	Lunch (on your own)
2–2:30pm	The Structural Power of Reconfigurable Circuits in the Amoebot Model (pg 34) Andreas Padalkin, Christian Scheideler and Daniel Warner
2:30–3pm	Fault-Tolerant Shape Formation in the Amoebot Model (pg 35) Daniel Warner, Christian Scheideler and Irina Kostitsyna
3–3:30pm	Break
3:30–4pm	Modelling and optimisation of a DNA stack nano-device using probabilistic model checking (pg 36) Bowen Li, Neil Mackenzie, Ben Shirt-Ediss, Natalio Krasnogor and Paolo Zuliani
4–4:30pm	NUPACK: Molecular Programming in the Cloud (pg 37) Mark E. Fornace, Jining Huang, Cody T. Newman, Nicholas J. Porubsky, Marshall B. Pierce and Niles A. Pierce
4:30–4:45pm	Main conference closing

Friday August 12th (Computational Modeling Workshop Day)

9:15–9:30am	Workshop welcome and introduction
9:30–10:00am	The Tall, Thin Molecular Programmer (pg 40) Erik Winfree
10:00–10:30am	Automated Leak Analysis of Nucleic Acid Circuits (pg 41) Andrew Phillips
10:30–11am	Break
11am–1pm	NUPACK tutorial: analysis and design of nucleic acid systems with the all-new NUPACK cloud web app (pg 42) Mark Fornace, Jining Huang, Cody T. Newman, and Niles A. Pierce
1–2:30pm	Lunch (on your own)
2:30–3pm	A roadmap for the future development of ENSnano (pg 43) Nicolas Levy
3–3:30pm	The oxDNA ecosystem (pg 44) Petr Sulc
3:30–4:30pm	Open format panel / discussion
4:30–4:45pm	Workshop closing

Poster Session 1 - Monday August 8th, 5:30-7:30pm

#42 Multi-Objective Sequence Selection for Scaffolded Origami Nanostructures (pg 48)
Ben Shirt-Ediss, Jordan Connolly, Emanuela Torelli, Silvia Navarro, Juan Elezgaray, Jaume Bacardit and Natalio Krasnogor (eligible for best student / postdoc poster award)

#47 Single-pass transmembrane double-stranded DNA with functional toeholds for non-destructive intra- and extravesicular nucleic acid target recognition (pg 49)
Gde Bimananda Mahardika Wisna, Nirbhik Acharya, Tarushyam Mukherjee, Ranjan Sasmal, Hao Yan and Rizal Hariadi (eligible for best student / postdoc poster award)

#48 Digital Nucleic Acid Memory 2.0: Advances in storage capacity and ease-of-use (pg 50)
Sarah Kobernat and Eric Hayden (eligible for best student / postdoc poster award)

#49 Design and analysis of compound nanostructures using oxView (pg 51)
Michael Matthies, Joakim Bohlin, Erik Poppleton, Jonah Procyk and Petr Šulc

#50 Using oxDNA simulations to characterize structural consequences of design modifications on a DNA origami leaf-spring engine (pg 52)
Erik Poppleton, Mathias Centola, Michael Famulok and Petr Šulc (eligible for best student / postdoc poster award)

#51 Operant conditioning of stochastic chemical reaction networks (pg 53)
David Arredondo and Matthew R. Lakin (eligible for best student / postdoc poster award)

#52 Computational analysis of conformational changes of nucleic acid nanostructures induced by overhangs (pg 54)
Michael Matthies, Matthew Sample, Lu Yu and Petr Sulc (eligible for best student / postdoc poster award)

#54 Unsupervised and Interpretable Machine Learning for Aptamer Analysis (pg 55)
Jonah Procyk and Petr Sulc (eligible for best student / postdoc poster award)

#56 Molecular-responsive DNA nanopore for biomolecular measurements (pg 56)
Hiromu Akai and Kan Shoji (eligible for best student / postdoc poster award)

#61 Dendric DNA Origami for Efficient DDS Carrier (pg 57)
Akinori Kuzuya

#64 Multi-gate Boolean processors in a cell free transcriptional platform (pg 58)
Judee Sharon, Kate Adamala, Chelsea Dasrath, Aiden Fujiwara, Alessandro Synder, Mace Blank and Sam O'Brien (eligible for best student / postdoc poster award)

#67 Developmental self-assembly of a DNA ring with stimulus-responsive size and growth direction (pg 59)
Allison Glynn, Samuel Davidson and Lulu Qian (eligible for best student / postdoc poster award)

Poster Session 2 - Tuesday August 9th, 5:30-7:30pm

#14 Storing digital data in DNA-based tape (pg 72)
Afsaneh Sadremomtaz, Robert Glass, Jorge Guerrero, Micheal Brandon Reed, Eric Josephs and Reza Zadegan

#20 Nucleic Acid Memory: Super Resolution Microscopy enhances novel approach to DNA data storage (pg 73)
Luca Piantanida, George Dickinson, Golam Mortuza, William Clay, Christopher Green, Chad Watson, Eric Hayden, Timothy Andersen, Wan Kuang, Elton Graugnard, Reza Zadegan and William Hughes

#43 Nanopore decoding for DNA-computed over and under-expression of microRNA patterns (pg 74)
Sotaro Takiguchi and Ryuji Kawano (eligible for best student / postdoc poster award)

#44 Kinetic approach to suppress DNA hybridization error (pg 75)
Hiroyuki Aoyanagi, Simone Pigolotti, Shinji Ono and Shoichi Toyabe (eligible for best student / postdoc poster award)

#45 A geometric framework for reaction enumeration in computational nucleic acid devices (pg 76)
Sarika Kumar and Matthew R. Lakin (eligible for best student / postdoc poster award)

#46 Study of DNA nanostructures for delivering gene templates for homology directed repair (pg 77)
Diana Lopez, Wolfgang Pfeifer and Carlos Castro (eligible for best student / postdoc poster award)

#55 Scalable design and construction of DNA truss structures (pg 78)
Yudai Yamashita, Ibuki Kawamata and Satoshi Murata (eligible for best student / postdoc poster award)

#57 NanoFrame: A web-based DNA wireframe design tool for 3D structures (pg 79)
Samson Petrosyan and Grigory Tikhomirov (eligible for best student / postdoc poster award)

#58 small: A modular, extensible nanostructure design framework (pg 80)
Durham Smith and Grigory Tikhomirov (eligible for best student / postdoc poster award)

#60 DNA nanochannels leading to an artificial single-cell communication system (pg 81)
Hiroki Koiwa and Kan Shoji (eligible for best student / postdoc poster award)

#62 A refined shape annealing algorithm for the optimal generation of DNA origami designs (pg 82)
Bolutito Babatunde, Jonathan Cagan and Rebecca Taylor (eligible for best student / postdoc poster award)

#65 Distinguishing Molecular Circuit Input Pulses via a Pulse Detector (pg 83)
Colin Yancey and Rebecca Schulman (eligible for best student / postdoc poster award)

#66 A loser-take-all DNA circuit (pg 84)
Kellen Rodriguez, Namita Sarraf and Lulu Qian (eligible for best student / postdoc poster award)

#68 Rate-Independent DNA-Based Binary-Weight ReLU Neural Networks (pg 85)
Cameron Chalk, Boya Wang, Marko Vasic and David Soloveichik (eligible for best student / postdoc poster award)

#69 A minimal DNA Origami for Seeding Tiled DNA Nanotube Bundles (pg 86)
Sarah Webster and Deborah Kuchnir Fygenson (eligible for best student / postdoc poster award)

#70 Efficient Elementary Step Implementations in Nucleic Acid Kinetics Simulators (pg 87)
Boyan Beronov, Jordan Lovrod, Chengwei Zhang and Anne Condon (eligible for best student / postdoc poster award)

#72 Scaling up reusable DNA circuits using heat as a universal energy source (pg 88)
Tianqi Song and Lulu Qian (eligible for best student / postdoc poster award)

#74 The Transformation of Real-Valued Concentrations into Binary Signals (pg 89)
James Lathrop, Dawn Nye and Hugh Potter (eligible for best student / postdoc poster award)

#75 Qslib: Python control of qPCR machines for molecular programming experiments (pg 90)
Constantine Evans

#78 Compact and Efficient Chemical Boltzmann Machines (pg 91)
Inhoo Lee, William Poole and Erik Winfree (eligible for best student / postdoc poster award)

#80 Highly Parallel Synthesis, Computation, and Readout of DNA Logic Gates using High-Throughput Synthesis and Sequencing (pg 92)
Chandler Petersen, Samantha Borje, Gourab Chatterjee, Yuan-Jyue Chen and Georg Seelig (eligible for best student / postdoc poster award)

#81 Growing arbitrary patterns with neural reaction-diffusion (pg 93)
Salvador Buse and Erik Winfree (eligible for best student / postdoc poster award)

#83 Counting the number of input addition by DNA reaction system driven by DNA polymerase (pg 94)
Ibuki Kawamata, Motokazu Furuya and Satoshi Murata

#84 Facet TAM: An abstract tile assembly model with facet growth (pg 95)
Ahmed Shalaby, Constantine Evans and Damien Woods (eligible for best student / postdoc poster award)

#85 Formal Semantics for Stochastic Chemical Reaction Networks (pg 96)
Andres Ortiz-Munoz (eligible for best student / postdoc poster award)

Monday August 8th

8am	Check-in opens
8:30-9am	Breakfast
9–9:30am	Main conference welcome and introduction
9:30–10:30am	KEYNOTE TALK: Models of self-assembling systems: varying levels of abstraction and objectives (pg 2) Matthew Patitz (University of Arkansas, USA)
10:30–11am	Break
11–11:30am	Covered DNA core tiles for robust tuning of spurious nucleation (pg 3) Trent Rogers, Constantine Evans and Damien Woods
11:30am–12pm	SAT-assembly: Designing minimum complexity building blocks to assemble arbitrary shapes (pg 4) Joakim Bohlin, Ard Louis, Andrew Turberfield and Petr Sulc
12–2pm	Lunch (on your own)
2–2:30pm	On Turedo Hierarchies and Intrinsic Universality (pg 5) Samuel Nalin and Guillaume Theyssier (eligible for best student presentation award)
2:30–3pm	Computing Real Numbers with Large-Population Protocols Having a Continuum of Equilibria (pg 6) Xiang Huang and Rachel Huls
3–3:30pm	Break
3:30–3:45pm	Fridge Compiler: automated synthesis of molecular circuits (pg 7) Lancelot Walthieu, Gus Smith, Luis Ceze and Chris Thachuk
3:45–4pm	Single-pass transmembrane double-stranded DNA with functional toeholds for non-destructive intra- and extravesicular nucleic acid target recognition (pg 8) Gde Bimananda Mahardika Wisna, Nirbhik Acharya, Tarushyam Mukherjee, Ranjan Sasmal, Hao Yan and Rizal Hariadi
4–4:30pm	Fast and robust strand displacement cascades via systematic design strategies (pg 9) Tiernan Kennedy, Cadence Pearce and Chris Thachuk (eligible for best student presentation award)
4:30–5pm	A cooperative DNA catalyst (pg 10) Dallas Taylor, Samuel Davidson and Lulu Qian (eligible for best student presentation award)
5:30–7:30pm	Poster session 1

KEYNOTE TALK: Models of self-assembling systems: varying levels of abstraction and objectives

Matthew J. Patitz[1]

[1]University of Arkansas, Fayetteville, AR, USA

A model is an abstract representation of an object or system and is a simplification of what it represents, often utilizing smaller scale and/or lesser complexity. Despite such simplifications, models are typically designed so that they capture one or more essential aspects of the what they are modeling, allowing the targeted aspects to be studied more easily. Scientists use models to better understand complex systems and to give them predictive power. In the field of DNA nanotechnology, accurate models are especially important since many of the systems being studied are too small to directly observe throughout their evolution. In this talk, I will discuss the development and use of a variety of models of DNA-based self-assembling systems. I will cover models that seek to capture the essential features of systems at varying levels of abstraction, and explore the objectives as well as the limitations of such models. The goal is to show how models over a broad spectrum of abstraction can be useful, how connections can be made to mathematics and computer science, and where future work in the development, and refinement, of such models may be helpful to the DNA nanotechnology community.

Covered DNA core tiles for robust tuning of spurious nucleation

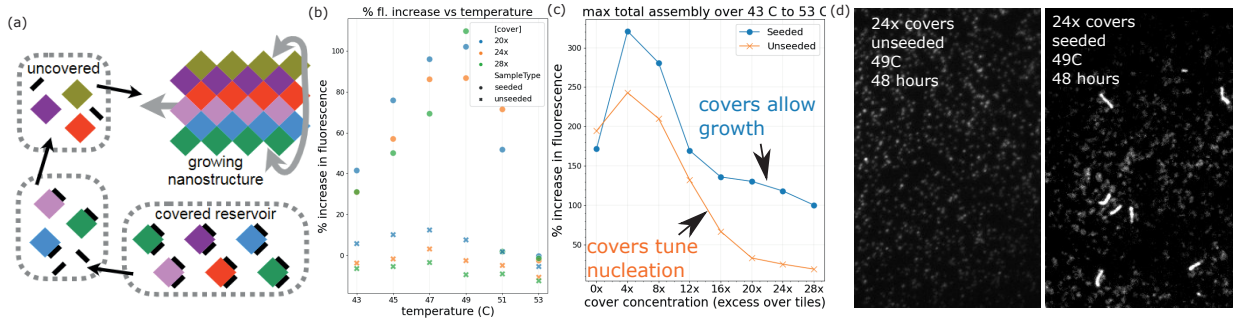
Trent Rogers, Constantine Evans, Damien Woods
Hamilton Institute, Department of Computer Science, Maynooth University

Despite decades of work, and several major successes, unintended and often high-rate spurious (unseeded) nucleation remains a central challenge for the design of DNA tile self-assembly systems.

Proposal. We propose a simple, active, strand motif, called a cover, as a means to control the rate of spurious nucleation in DNA tile self-assembly systems. The cover motif is a short domain that is complementary to one of the binding domains of a DNA tile, as shown below. The idea is to take any DNA tile system, and add a number of short cover strands, giving the experimenter the ability to tune the rate of spurious nucleation via cover concentration, without preventing growth. Our two main hypotheses are: (1) The rate of spurious nucleation is tuned by cover concentration (meaning nucleation is suppressed via increased cover concentration). (2) That tuning is achieved while maintaining a positive, and measurable, nanotube growth rate.

Results. We experimentally characterise the nucleation rates achieved over wide cover concentration (0–28x) and temperature (10 °C) ranges, using a model 12-helix DNA nanotube system. First, bulk fluorescence measurements (via qPCR machine) on 12-helix DNA nanotubes with covers at 0x to 28x, in increments of 4x, over (1x) tile concentration, show that the rate of spurious nucleation rapidly decreases with increasing cover concentration. Second, this is backed up by a thorough fluorescence microscopy analysis (with 36 distinct (temperature, concentration) conditions) with similar results, but shows reasonable nanotube lengths are achieved, showing that growth does not stall. Third, we show that seeded growth is achievable at a range of cover concentrations (20x to 28x) where spurious nucleation is almost completely suppressed.

Conclusion. Our data provides evidence in favour of hypotheses (1) and (2), at least for our novel core tile motif. Given what we know about existing (coverless) tile systems, our work suggests that covers may act to tune/suppress spurious nucleation beyond what is achievable by mere uniform tile-concentration and temperature adjustments (e.g. single-stranded tiles at 100 nM achieve a mere 2.5 °C seeded growth window, here we widen that to 6 °C). Systems to date either have uncontrolled nucleation, which causes poor-quality structures, or require labourious purification techniques, or completely suppress nucleation via sensitive temperature-hold protocols requiring time-consuming calibration steps, or else complete change the tile motif. Here, we open the door to (A) custom tuned nucleation rates, but with (B) a general and simple motif that can, at least in principle, be 'spiked' into any existing DNA tile-based self-assembly system to potentially gain nuanced, tunable, control over the nucleation rate, while allowing (seeded or unseeded) growth.



(a) Hypotheses for cover mechanism: fully covered tiles do not participate in self-assembly, fully uncovered tiles do. Increasing the cover concentration, reduces the "effective tile concentration" which decreases the spurious nucleation rate much more than the growth rate. (b) A plot of the percentage increase in fluorescence over a 96 hour hold for 36 combinations of conditions. Unseeded samples show little signal increase indicating very limited tube formation while seeded samples show large signal increase from 45 °C to 51 °C. (c) A plot showing the max fluorescence increase over 43 °C to 53 °C for unseeded and seeded samples at a variety of cover concentrations. (d) Fluorescence microscopy data of samples consisting of 24x cover concentration.

SAT-assembly: Designing minimum complexity building blocks to assemble arbitrary shapes

Joakim Bohlin[1,2], Andrew J. Turberfield[2], Ard A. Louis[2], and Petr Šulc[1]

[1] *School of Molecular Sciences, Arizona State University, Tempe, USA*
[2] *Department of Physics, Oxford University, Oxford, UK*

We present a new method, "*SAT-assembly*", to design self-assembling 2D or 3D shapes from minimum number of distinct building blocks. As opposed to "fully addressable" solutions, where each individual particle is its own unique species, our method tries to find the smallest number of distinct particle types that is required to assemble target structure in high yield. We have previously introduced this approach for self-assembling pyrochlore and cubic diamond lattices [1,2], and here we develop a new automated design pipeline to find minimum set that can assemble an arbitrary 3D structure. The main advantage of this technique is that fewer species are required.

Our method relies on formulating the inverse design problem (i.e. finding the set of particles and their interactions to self-assemble a given target) as a Boolean satisfiability problem, in terms of binary variables that have to satisfy a particular set of logical clauses. We can then harness available highly efficient algorithms that can find solution that forms the target structure as a free-energy minimum in few seconds. We then introduce an automated 3D stochastic assembly solver that runs a lattice-based simulation of assembly of these designed blocks. If they aggregate or misassemble into a competing structure, a new solution is generated that avoids the undesired structure. The yield and kinetics are studied by molecular dynamics simulations of a patchy particle model that abstracts self-assembling particles, such as DNA origamis, proteins, or DNA-coated nanoparticles.

We use our design pipeline to find minimum building blocks for a range of distinct finite-size 3D shapes, include symmetric and asymmetric ones. For certain examples, we find that the target shape can be assembled in high yield with less than 10% of the number of building blocks needed for fully addressable assembly. For very rigid interactions (that would correspond e.g. to shape-complementary DNA origamis), the minimum assembly kits assemble faster than the fully addressable solution. We further introduce an automated design tool that constructs target 3D shapes from DNA origami building blocks, allowing for easy setup of coarse-grained simulations with the oxDNA model to verify and computationally characterize the designs prior to experimental realization. Finally, we also introduce a version of the inverse design pipeline for multifarious assemblies, where we find a minimum set of building blocks that can assemble into multiple different target designs.

Our method can be used for design of self-assembling systems, with applications in biotemplated manufacturing and assembly of multicomponent structures, be it at protein, DNA or DNA-functionalized particles level.

References

[1] Russo, J., Romano, F., Kroc, L., Sciortino, F., Rovigatti, L., & Šulc, P. (2022). SAT-assembly: A new approach for designing self-assembling systems. Journal of Physics: Condensed Matter.
[2] Romano, F., Russo, J., Kroc, L., & Šulc, P. (2020). Designing patchy interactions to self-assemble arbitrary structures. Physical Review Letters, 125(11), 118003.

On Turedo Hierarchies and Intrinsic Universality

Samuel Nalin[1] and Guillaume Theyssier[2]

[1]Univ. Orléans, INSA Centre Val de Loire, LIFO EA 4022, FR-45067 Orléans, France, [2]I2M, CNRS, Université Aix-Marseille, France

This paper is about turedos, which are Turing machines whose head can move in the plane (or in a higher-dimensional space) but only in a self-avoiding way, by putting marks (letters) on visited positions and moving only to unmarked, therefore unvisited, positions. The turedo model has been introduced recently as a useful abstraction of oritatami systems, which where established a few years ago as a theoretical model of RNA co-transcriptional folding. The key parameter of turedos is their lookup radius: the distance up to which the head can look around in order to make its decision of where to move to and what mark to write. In this paper we study the hierarchy of turedos according to their lookup radius and the dimension of space using notions of simulation up to spatio-temporal rescaling (a standard approach in cellular automata or self-assembly systems). We establish that there is a rich interplay between the turedo parameters and the notion of simulation considered. We show in particular, for the most liberal simulations, the existence of 3D turedos of radius 1 that are intrinsically universal for all radii, but that this is impossible in dimension 2, where some radius 2 turedo are impossible to simulate at radius 1. Using stricter notions of simulation, intrinsic universality becomes impossible, even in dimension 3, and there is a strict radius hierarchy. Finally, when restricting to radius 1, universality is again possible in dimension 3, but not in dimension 2, where we show however that a radius 3 turedo can simulate all radius 1 turedos.

Computing Real Numbers with Large-Population Protocols Having a Continuum of Equilibria

Xiang Huang[1] and Rachel Huls[2]

[1]Department of Computer Science, University of Illinois Springfield, USA, [2]Department of Mathematics, University of Illinois Springfield, USA

Bournez, Fraigniaud, and Koegler defined a number in [0,1] as computable by their Large-Population Protocol (LPP) model, if the proportion of agents in a set of marked states converges to said number over time as the population grows to infinity. The notion, however, restricts the ordinary differential equations (ODEs) associated with an LPP to have only finitely many equilibria. This restriction places an intrinsic limitation on the model. As a result, a number is computable by an LPP if and only if it is *algebraic*, namely, not a single transcendental number can be computed under this notion.

In this paper, we *lift* the finitary requirement on equilibria. That is, we consider systems with a continuum of equilibria. We show that essentially all numbers in [0,1] that are computable by bounded general-purpose analog computers (GPACs) or chemical reaction networks (CRNs) can also be computed by LPPs under this new definition. This implies a rich series of numbers (e.g., the reciprocal of Euler's constant, $\pi/4$, Euler's γ, Catalan's constant, and Dottie number) are all computable by LPPs. Our proof is constructive: We develop an algorithm that transfers bounded GPACs/CRNs into LPPs. Our algorithm also fixes a gap in Bournez et al.'s construction of LPPs designed to compute any arbitrary algebraic number in [0,1].

Fridge Compiler: automated synthesis of molecular circuits

Lancelot Wathieu[1], Gus Smith[1], Luis Ceze[1], and Chris Thachuk[1]

[1]Paul G. Allen School of Computer Science & Engineering, University of Washington, USA

Rationally designed molecular circuits describable by well-mixed chemical reaction kinetics can realize arbitrary Boolean function computation yet differ significantly from their electronic counterparts: (i) NOT gates are often infeasible thus requiring the use of dual-rail encoding of inputs and (ii) spatial locality is absent necessitating that gates are distinguishable (*i.e.*, their design dictates possible interactions with other components). Given a Boolean function and a "fridge" inventory of distinguishable components, which subset should be mixed to realize the desired computation? This molecular circuit synthesis question is often solved manually and *ad hoc*. Existing synthesis tools intended for electronic circuits can improve upon manual determination, but can still result in non-optimal circuits or ones that cannot be built with the current inventory. This work overcomes these limitations by introducing the *Fridge Compiler* molecular circuit synthesis tool backed by the solver-aided programming language Rosette [1].

The Fridge Compiler was designed specifically for optimality, flexibility, and ease-of-use. Users can design optimal circuits by first selecting from a list of common DNA strand displacement (DSD) architectures and then specifying (i) a cost function (*e.g.*, component count), (ii) arbitrary constraints (*e.g.*, avoid mixing gates *K* and *L*), (iii) a fridge inventory, and (iv) the Boolean function to compute. Power users can define new molecular architectures (e.g., protein circuits), specify custom constraints and cost functions, and explore the space of valid circuits.

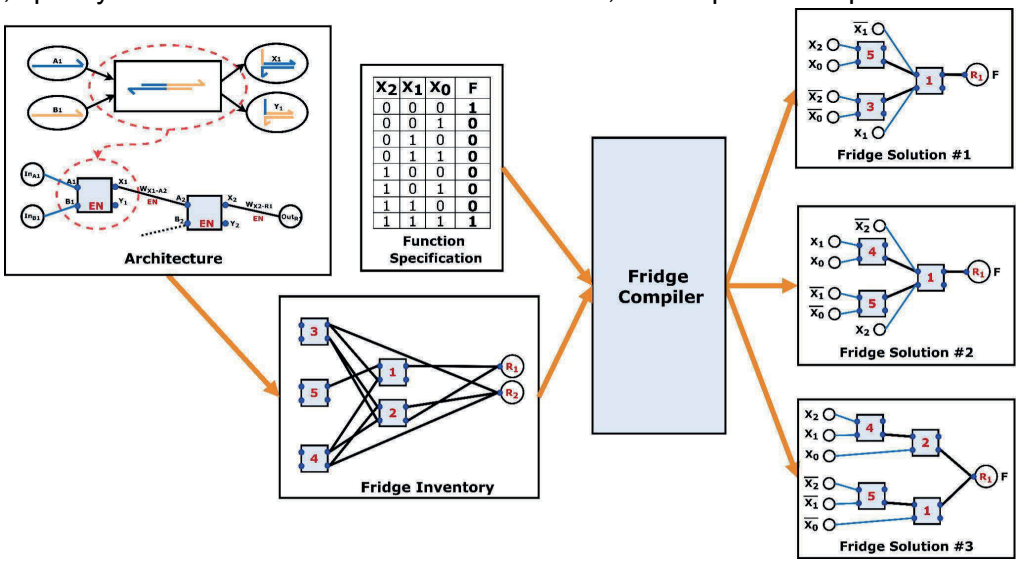


Fig. 1: Given a inventory and function specification the *Fridge Compiler* can synthesize all valid circuit solutions (#1, #2, #3, …), (size) optimal solutions (#1, #2), and optimal solutions that satisfy hard constraints (*#1; "don't mix gates 4 and 5, nor gates 1 and 2").*

The Fridge Compiler is particularly well-suited for molecular computing where the cost and effort of designing, preparing and purifying new circuit parts is often a significant barrier. Empirical results demonstrate its flexibility to synthesize arbitrary Boolean functions using three different DSD architectures. All 3-bit and 4-bit predicate functions were synthesized from a large custom inventory on the order of seconds in the worst case on a modern laptop. Importantly, and absent in previously known approaches specific to molecular circuits, the Fridge Compiler is logically sound, complete, and optimal for the user specified cost function.

[1] Torlak, E., & Bodik, R. (2013, October). Growing solver-aided languages with Rosette. In *Proceedings of the 2013 ACM international symposium on New ideas, new paradigms, and reflections on programming & software* (pp. 135-152).

Single-pass transmembrane double-stranded DNA with functional toeholds for non-destructive intra- and extravesicular nucleic acid target recognition

Gde Bimananda M. Wisna[1,2], Nirbhik Acharya[2], Tarushyam Mukherjee[2], Ranjan Sasmal[2], Hao Yan[2,3], Rizal F. Hariadi[1,2]

[1]Department of Physics, Arizona State University, USA, [2]Biodesign Institute, Arizona State University, USA, [3]School of Molecular Sciences, Arizona State University, USA

Cells express nucleic acid with genetic information specific to cell types and their states. Additionally, the interaction between cells and extracellular ligands, as well as intracellular biomarkers, tune cellular activities. These biomarkers, often in the form of exRNA produced inside and secreted by other cells, facilitate intercellular signaling and communications. Therefore, it is important to recognize intra- and extracellular nucleic acids simultaneously while keeping the cells intact for further studies and cell classification. A relatively straightforward path to achieve recognition of nucleic acid targets is to use DNA nanostructures that are complementary to the nucleic acid targets. In addition, mimicking membrane proteins that span the membrane provides an attractive way for intra- and extracellular target recognition and to transduce information across the membranes. Here, we study cholesterol-modified double-stranded DNA (dsDNA) sensors spanning the lipid membrane whose toeholds can recognize nucleic acids target inside and outside synthetic lipid bilayer. The dsDNA sensors have three common domains consisting of extracellular, transmembrane, and intracellular domains, mimicking single-pass transmembrane proteins. As a proof-of-concept, we utilize synthetic 10–20 μm giant unilamellar vesicles (GUVs) as our cell-like lipid bilayer compartment model. The two-strand design provides a simple approach significantly minimizing the stoichiometric issue of structure formation. We show that with the two cholesterols modification only on one of the duplex strands, the structure still spans the membrane and binds to the target inside GUVs due to the duplex helicity. Our preliminary data with mammalian cells (HEK 293T) show that dsDNA sensors successfully anchor on the plasma membrane and achieve extracellular DNA target recognition. The dsDNA sensor keeps the cells and vesicles intact during recognition. It paves a way toward a more complex signal transmission process across lipid bilayers to enable DNA signal amplification and manipulations in the outer space of vesicles or cells, providing target binding from the inside.

Fast and Robust Strand Displacement Cascades via Systematic Design Strategies

Tiernan Kennedy[1], Cadence Pearce[1], and Chris Thachuk[1]

[1]Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, WA, USA

A barrier to wider adoption of molecular computation is the difficulty of implementing arbitrary chemical reaction networks (CRNs) that are robust and replicate the kinetics of designed behavior. DNA Strand Displacement (DSD) cascades have been a favored technology for this purpose due to their potential to emulate arbitrary CRNs and known principles to tune their reaction rates. Progress on *leakless* cascades have demonstrated that DSDs can be arbitrarily robust to spurious “leak” reactions when incorporating systematic domain level redundancy. These improvements in robustness result in slower kinetics of designed reactions. Existing work has demonstrated the kinetic and thermodynamic effects of sequence mismatch introduction and elimination during displacement. We present a systematic, sequence modification strategy for optimizing the kinetics of leakless cascades without practical cost to their robustness. An in-depth case study explores the effects of this optimization when applied to a typical leakless translator cascade. Thermodynamic analysis of energy barriers and kinetic experimental data support that DSD cascades can be fast and robust.

A cooperative DNA catalyst

Dallas N. Taylor[1,3†], Samuel R. Davidson[2†] and Lulu Qian[1,2,3⋆]
[1] *Computation and Neural Systems,* [2] *Bioengineering, and* [3] *Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA
[†] *Equal contribution,* [⋆] *e-mail: luluqian@caltech.edu*

Abstract. DNA catalysts are fundamental building blocks for diverse molecular information-processing circuits. Allosteric control of DNA catalysts has been developed to activate desired catalytic pathways at desired times. Here we introduce a new type of DNA catalyst that we call a cooperative catalyst: a pair of reversible reactions are employed to drive a catalytic cycle, where two signal species, which can be interpreted as an activator and input, both exhibit catalytic behavior for output production. We demonstrate the role of a dissociation toehold in controlling the kinetics of the reaction pathway and the significance of a wobble base pair in promoting the robustness of the activator. We show near-complete output production with one tenth of input and activator concentrations. The system involves just a double-stranded gate species and a single-stranded fuel species, as simple as the seesaw DNA catalyst, which has no allosteric control. The simplicity and modularity of the design make the cooperative DNA catalyst an exciting addition to strand-displacement motifs for general-purpose computation and dynamics.

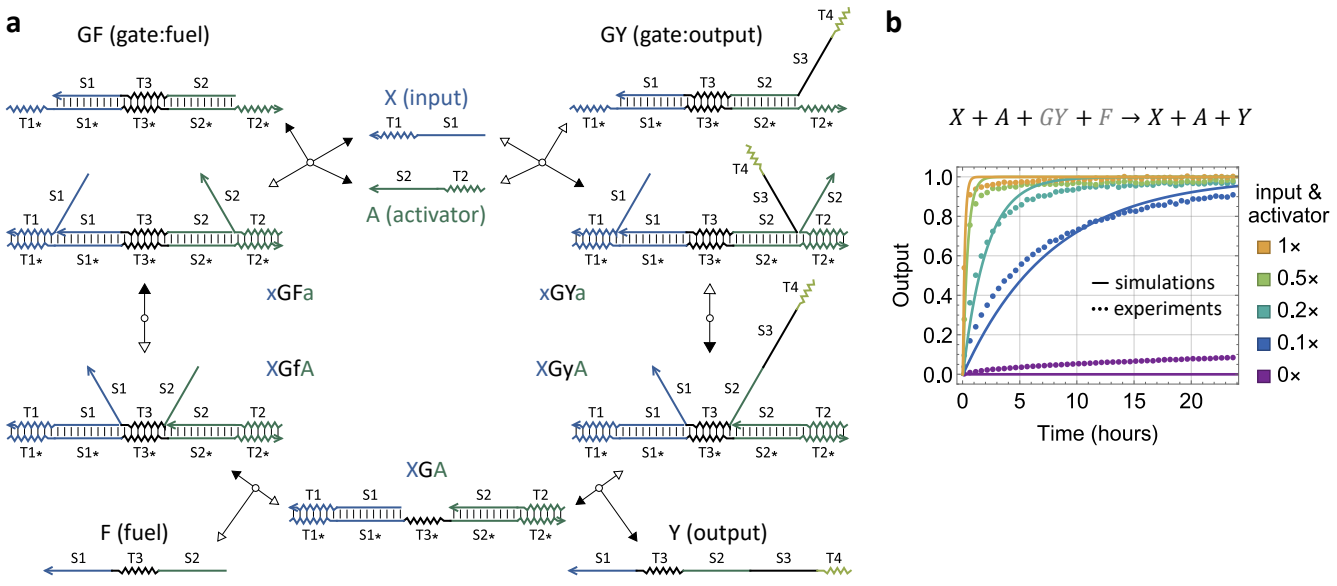


Figure 1: **DNA strand-displacement implementation of a cooperative catalyst. (a)** Catalytic reaction pathway. **(b)** Simulation and fluorescence kinetics data with varying input and activator concentrations.

Tuesday August 9th

8:30-9am	Breakfast
9–10am	KEYNOTE TALK: Dynamic control of biomolecular phase separation (pg 12) Elisa Franco (University of California Los Angeles, USA)
10–10:30am	Exploring Material Design Space with a Deep-Learning Guided Genetic Algorithm (pg 13) Kuan-Lin Chen and Rebecca Schulman (eligible for best student presentation award)
10:30–11am	Break
11–11:30am	Protecting heterochiral DNA nanostructures against exonuclease-mediated degradation (pg 14) Tracy Mallette and Matthew R. Lakin (eligible for best student presentation award)
11:30–12pm	DNA strand-displacement temporal logic circuits (pg 15) Anna Lapteva, Namita Sarraf and Lulu Qian (eligible for best student presentation award)
12–2pm	Lunch (on your own)
2–2:15pm	Scaling up reusable DNA circuits using heat as a universal energy source (pg 16) Tianqi Song and Lulu Qian
2:15–2:30pm	Highly Parallel Synthesis, Computation, and Readout of DNA Logic Gates using High-Throughput Synthesis and Sequencing (pg 17) Chandler Petersen, Samantha Borje, Gourab Chatterjee, Yuan-Jyue Chen and Georg Seelig
2:30–3:30pm	KEYNOTE TALK: Hourglass emergence + collective computation in nature (pg 18) Jessica Flack (Santa Fe Institute, USA)
3:30–4pm	Break
4–4:30pm	Growth dynamics: Precisely controlled self-assembly order of DNA tile nanostructures (pg 19) Constantine Evans, David Doty and Damien Woods
4:30–5pm	Universal Shape Replication Via Self-Assembly With Signal-Passing Tiles (pg 20) Andrew Alseth, Matthew Patitz and Daniel Hader
5–5:30pm	Algorithmic DNA origami: Scaffolded DNA computation in one and two dimensions (pg 21) Tristan Stérin, Abeer Eshra and Damien Woods (eligible for best student presentation award)
5:30–7:30pm	Poster session 2

KEYNOTE TALK: Dynamic control of biomolecular phase separation

Elisa Franco[1,2]

[1]Department of Mechanical and Aerospace Engineering, University of California, Los Angeles, USA. [2]Department of Bioengineering, University of California, Los Angeles, USA.

Phase separation is a widespread phenomenon in living cells, that takes advantage of spontaneous aggregation of certain proteins and RNA species to spatially and dynamically organize a variety of molecules and reactions. Taking inspiration from nature, our group is developing methods to build condensates and patterns using artificial DNA molecules, that are relevant to the synthesis of self-regulated, autonomous biomaterials. I will present our recent results on controlling the rate of condensation and dissolution of artificial DNA motifs using chemical reactions. I will also discuss mathematical models that support and guide our experiments by capturing specific and non-specific interactions among the motifs.

Exploring Material Design Space with a Deep-Learning Guided Genetic Algorithm

Kuan-Lin Chen[1] and Rebecca Schulman[1,2,3]

[1]Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, USA, [2]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA, [3]Department of Chemistry, Johns Hopkins University, Baltimore, MD, USA

Designing complex, dynamic yet multi-functional materials and devices is challenging because the design spaces for these materials have numerous interdependent and often conflicting constraints. Taking inspiration from advances in artificial intelligence and their applications in material discovery, we propose a computational method for designing metamorphic DNA-co-polymerized hydrogel structures. The method consists of a coarse-grained simulation and a deep learning-guided optimization system for exploring the immense design space of these structures. Here, we develop a simple numeric simulation of DNA-co-polymerized hydrogel shape change and seek to find designs for structured hydrogels that can fold into the shapes of different Arabic numerals in different actuation states. We train a convolutional neural network to classify and score the geometric outputs of the coarse-grained simulation to provide autonomous feedback for design optimization. We then construct a genetic algorithm that generates and selects large batches of material designs that compete with one another to evolve and converge on optimal objective-matching designs. We show that we are able to explore the large design space and learn important parameters and traits. We identify vital relationships between the material scale size and the range of shape change that can be achieved by individual domains and we elucidate trade-offs between different design parameters. Finally, we discover material designs capable of transforming into multiple different digits in different actuation states.

Protecting heterochiral DNA nanostructures against exonuclease-mediated degradation

Tracy L. Mallette[1] and Matthew R. Lakin[1,2,3]

[1]Center for Biomedical Engineering, University of New Mexico, USA, [2]Department of Computer Science, University of New Mexico, USA, [3]Department of Chemical & Biological Engineering, University of New Mexico, USA

Heterochiral DNA nanotechnology employs nucleic acids of both chiralities to construct nanoscale devices for applications in the intracellular environment. Interacting directly with cellular nucleic acids can be done most easily using D-DNA of the naturally occurring right-handed chirality, however, D-DNA is more vulnerable to degradation than enantiometric left-handed L-DNA. Here we report a novel combination of D-DNA and L-DNA nucleotides in triblock heterochiral copolymers, where the L-DNA domains act as protective caps on D-DNA domains. We demonstrate that the D-DNA components of strand displacement-based molecular circuits constructed using this technique resist exonuclease-mediated degradation during extended incubations in serum-supplemented media more readily than similar devices without the L-DNA caps. We show that this protection can be applied to both double-stranded and single-stranded circuit components. Our work enhances the state of the art for robust heterochiral circuit design and could lead to practical applications such as in vivo biomedical diagnostics.

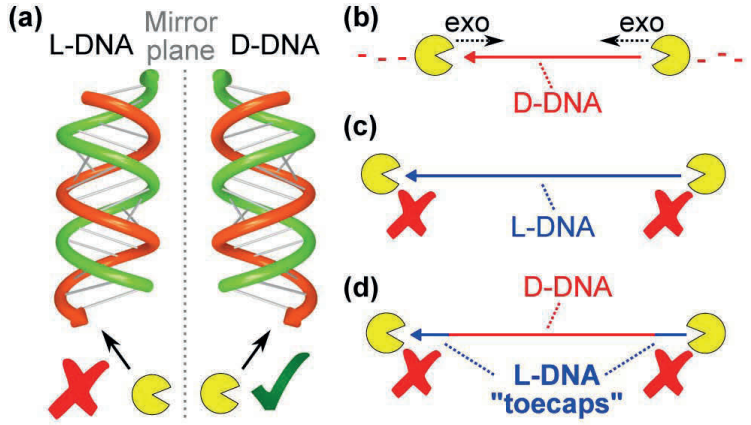


Fig. 1. Heterochiral DNA nanotechnology for robust molecular computing devices. (a) L-DNA is the chiral mirror image of naturally occurring D-DNA. (b) D-DNA domains are expected to be degraded by exonucleases present in model biological fluids. (c) Since the chirality of the backbone sugars is flipped, exonucleases that have evolved to recognized-nucleic acid strand termini do not readily recognize L-nucleic acids, enabling L-DNA to survive for extended periods of time in biological fluids. (d) Here we exploit this property to protect robust heterochiral DNA nanostructures by adding L-DNA “toecaps” to the strand termini of D-DNA domains, thereby protecting them against recognition by exonucleases.

Full paper: ACS Synthetic Biology Article ASAP, DOI: 10.1021/acssynbio.2c00105

DNA strand-displacement temporal logic circuits

Anna P. Lapteva[1†], Namita Sarraf[1†] and Lulu Qian[1,2⋆]
[1] *Bioengineering, and* [2] *Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA
[†] *Equal contribution,* [⋆] *e-mail: luluqian@caltech.edu*

Abstract. Molecular circuits capable of processing temporal information are essential for complex decision making in response to both the presence and history of a molecular environment. A particular type of temporal information that has been recognized to be important is the relative timing of signals. Here we demonstrate the strategy of temporal memory combined with logic computation in DNA strand-displacement circuits capable of making decisions based on specific combinations of inputs as well as their relative timing. The circuit encodes the timing information of inputs in a set of memory strands, which allows for the construction of logic gates that act on current and historical signals. We show that mismatches can be employed to reduce the complexity of circuit design, and that shortening specific toeholds can be useful for improving the robustness of circuit behavior. We also show that a detailed model can provide critical insights for guiding certain aspects of experimental investigations that an abstract model cannot. We envision that the design principles explored in this study can be generalized to more complex temporal logic circuits and incorporated into other types of circuit architectures, including DNA-based neural networks, enabling the implementation of timing-dependent learning rules and opening up new opportunities for embedding intelligent behaviors into artificial molecular machines.

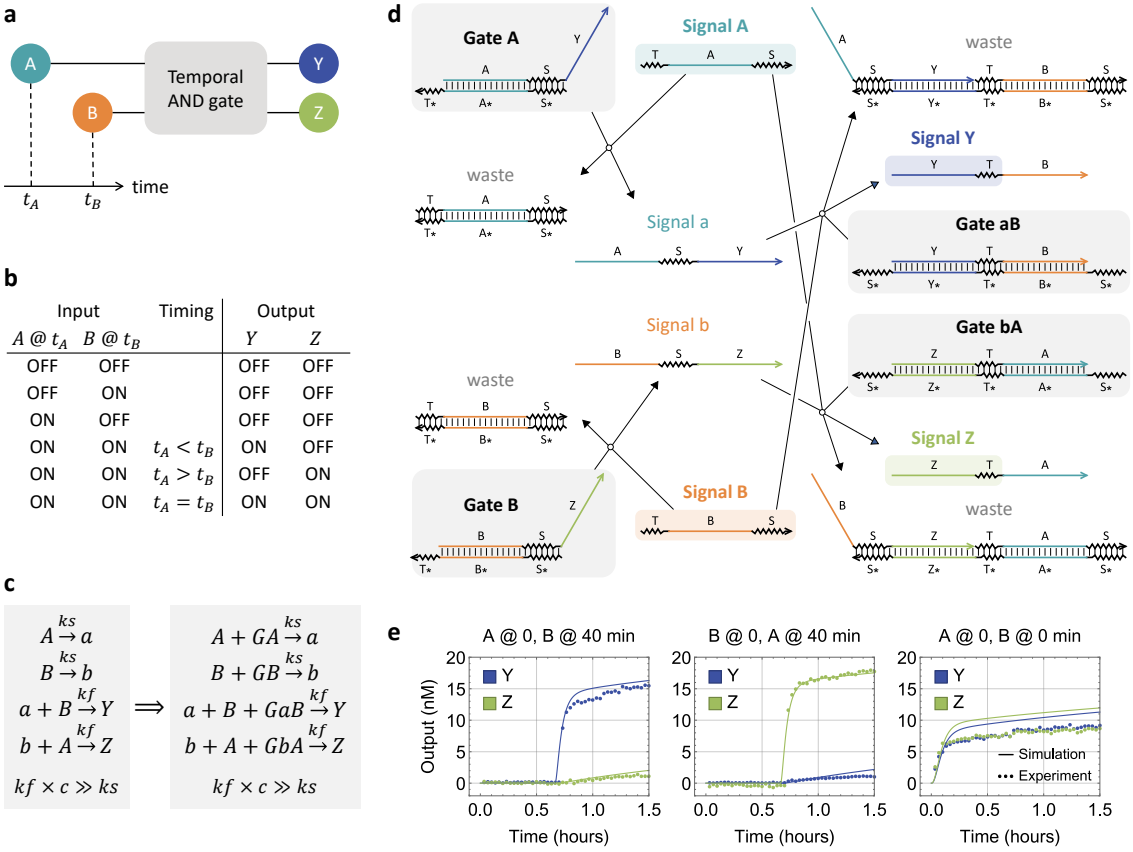


Figure 1: **A temporal AND gate. (a)** Abstract circuit diagram. **(b)** Truth table. **(c)** Chemical reaction network implementation. **(d)** DNA strand-displacement implementation. **(e)** Simulation and fluorescence kinetics data.

Scaling up reusable DNA circuits using heat as a universal energy source

Tianqi Song[1] and Lulu Qian[1,2]
[1] *Bioengineering,* [2] *Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA

State-of-the-art enzyme-free nucleic-acid circuits cannot remain active and respond to changing input signals – they reach equilibrium states when some molecular components are used up. How can these circuits stay active and process time-varying information in a molecular environment? Several types of chemical power supplies have been explored to reinvigorate the system, but none of these approaches appear to be scalable. Inspired by recent examples of recycling DNA circuits by heat, and building on the observation of nucleic-acid strands with strong secondary structures can reach kinetically trapped states during thermal annealing, here we develop a general-purpose circuit architecture where all molecular components can be recharged (restored to their initial states) using heat as a universal energy source after the input is inactivated. We show that an exceptionally simple motif can be designed to respond to a quick temperature ramp and hundreds of distinct molecules composing a complex DNA neural network can spontaneously reconfigure to their original structures within minutes, allowing at least ten rounds of computation with various input signals. Our approach ensures that arbitrarily complex circuits can be powered by the same energy source, no debugging is necessary for the power supply itself, and the circuit performance remains roughly constant over time. This work introduces a scalable approach for enabling sustained operation of enzyme-free molecular circuits, and establishes the foundation for future demonstrations of unsupervised learning in chemical systems.

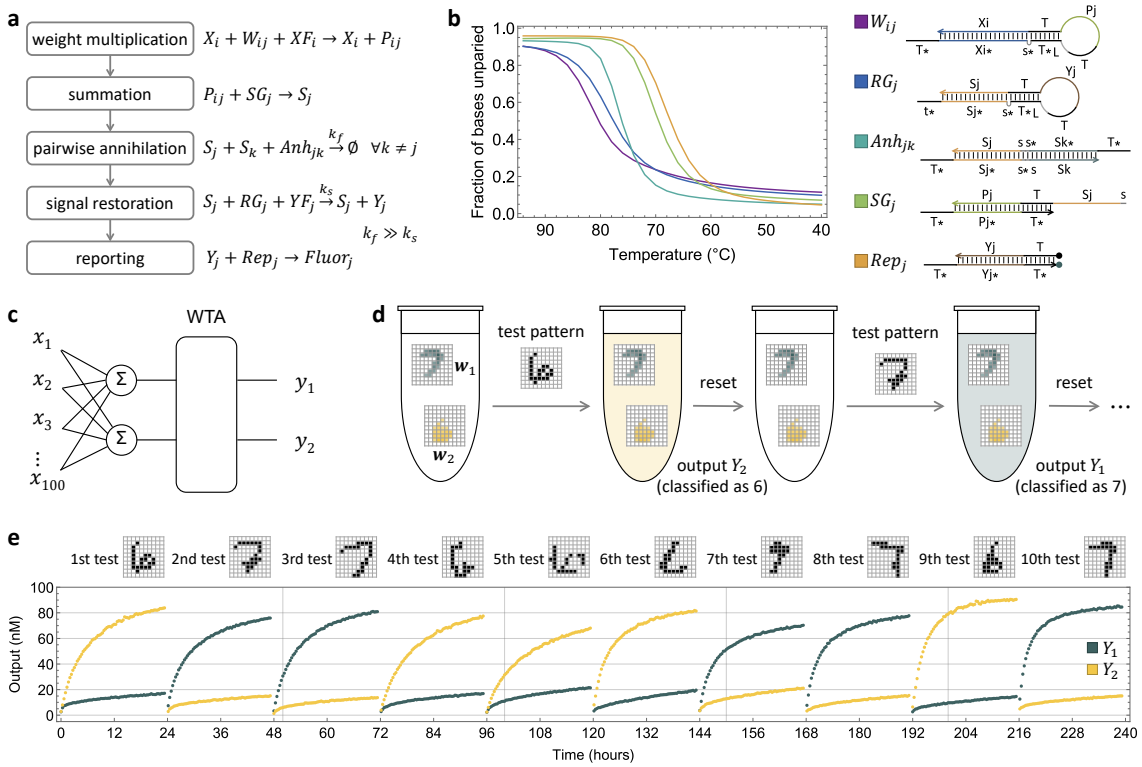


Figure 1: **A reusable DNA neural network. a,** Chemical reaction network implementation. **b,** DNA strand-displacement implementation with a melting profile that allows for the system to reach a kinetically trapped state during thermal annealing, which resets the circuit to its initial state after the input is inactivated. **c,** Circuit diagram of a 100-bit, 2-memory winner-take-all (WTA) neural network. **d,** Steps in testing the reusability of the circuit. **e,** Fluorescence kinetics data of classifying ten MNIST patterns sequentially added to the same test tube. Each reset was performed by adding complementary input strands to inactivate the previous test pattern and heating up to 95 °C for 5 minutes and cooling down to 20 °C in 1 minute.

Highly Parallel Synthesis, Computation, and Readout of DNA Logic Gates using High-Throughput Synthesis and Sequencing

Chandler Petersen[1], Samantha Borje[1], Gourab Chatterjee[1], Yuan-Jyue Chen[1,2], and Georg Seelig[1]

[1]University of Washington, USA, [2]Microsoft Research, Seattle, WA

DNA computing's strength lies in its ability to process information while still in molecular form, allowing computation and programmable control of biological matter at the nanoscale. Implementations of Boolean logic gates [4, 5] have been particularly promising for a variety of applications. However, full realization of their potential will require overcoming existing limitations of molecular logic circuits. The size of current DNA logic circuits is limited for several reasons. First, the cost of column-synthesized DNA oligonucleotides is expensive. Second, DNA gates typically require purification to remove errors in stoichiometry, but this is very time-consuming and not multiplexable, causing significant challenge to scale up. Finally, the reading bandwidth of conventional fluorescence reporter readout is limited due to spectral overlaps. To address these limitations, here we present multiplexable methods that enables highly parallel preparation and reading of nicked double-stranded DNA (ndsDNA) gates [2, 3] by using array-based DNA synthesis and next-generation sequencing (NGS). Each ndsDNA gate consists of a join and fork gate. For AND logic, the join gate accepts two input signals, which triggers a strand displacement cascade resulting in an intermediate private co-signal to the fork gate, producing a final output signal with the same domain structure [2, 3]. Additionally, each fork gate performs catalytic amplification of the private co-signal, providing signal restoration local to each gate [6]. This modular design pairs computation with amplification to address signal degradation inherent to multi-layered circuits. Each gate contains flanking "barcode" regions that are used for selective PCR and sequencing of reacted gates. With these modifications, complex circuits with potentially thousands of gates may be read out in a single sequencing reaction [1]. This is the first demonstration of large-scale synthesis and sequencing of DNA gates, with the potential to realize more complex circuits and programmable logic on a scale not yet seen.

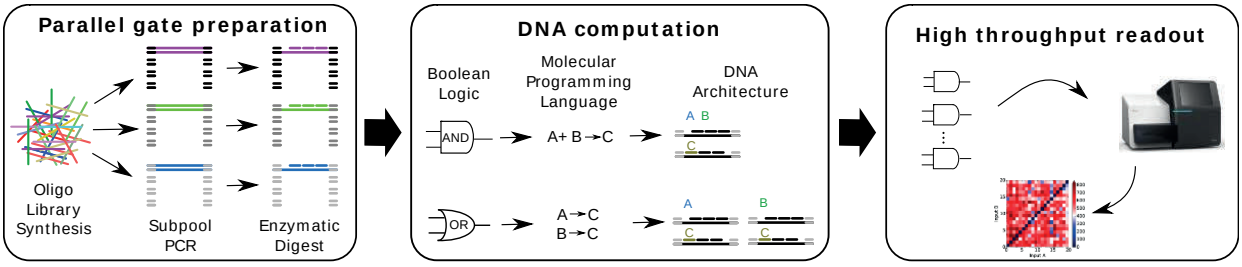


Figure 1. Overview of circuit synthesis, computation, and read out. Left: Gates are derived from an array synthesized DNA pool, selectively amplified with PCR, and enzymatically processed to create ndsDNA gates. Middle: Computation occurs in pooled reactions upon addition of inputs. Right: End point results of Boolean logic computation are read out in parallel using NGS.

[1] Cardelli L 2021 *arXiv*
[2] Cardelli L 2010 *EPTCS* **26** pp. 47-61
[3] Chen Y-J et al 2013 *Nat. Nanotechnol* **8** 755-762
[4] Seelig G et al 2006 *Science* **314** pp. 1585-1588
[5] Qian L and Winfree E 2011 *Science* **332** pp. 1196-1201
[6] Zhang D et al 2007 *Science* **318** pp. 1121-1125

KEYNOTE TALK: Hourglass Emergence + Collective Computation in Nature

Jessica Flack[1]

[1]Santa Fe Institute, Santa Fe, NM, USA

Examples of computation in nature include the expansion and contraction of heart cells by intracellular nanotube circuits, gene regulatory networks controlling gene expression contributing to the embryo body plan, "microbial power grids" shuffling electrons to compute metabolic functions, neural circuits computing decisions at the whole organism level, social insect colonies adaptively adjusting nest architectural properties in response to environmental changes, primate social circuits computing aspects of social structure like the distribution of power, and human voting designed to identify the consensus view of the best presidential candidate. Like the fact that life evolves, the fact that biological systems compute functional outputs is a truism, yet we are still a long way from understanding how nature computes. A test of such understanding might be the development of a formal language for biological computation grounded in biology’s universal collective property, information processing mechanisms, and thermodynamic constraints, that permits rigorous, comparative study of energy and information transformations across a wide range of systems. In this talk I will discuss the challenges as I see them, touching on some of the basic elements and mechanics of computation in biological systems, as well as the implications of re-framing biology as collective computation for theories of emergence.

Growth dynamics: Precisely controlled self-assembly order of DNA tile nanostructures*

Constantine Evans† David Doty‡ Damien Woods†

The field of DNA nanostructures has primarily concerned itself with its namesake: making precise nanoscale *structures* out of DNA. Here, our goal is to precisely control self-assembly *dynamics*, or order of assembly, of DNA tile nanostructures. The primary motivation is fundamental: we wish gain step-by-step control, in as far as possible, of the process of self-assembly, by careful molecular design choices together with a universal approach that will enable programming of almost any self-assembly dynamics. Our approach takes inspiration from how theorists think about a computation driving the formation of a structure (so-far relegated to theorem-proving papers), and stands in stark contrast to traditional crystallographic approaches of tuning macro variables like temperature and concentration for mere coarse-grained control of crystal nucleation and bulk growth rates.

Design. We take inspiration from the basic insight that controlled growth order requires control of facet growth initiation, and provide a universal growth dynamics design schema that enables systematic control of growth order at various spatial scalings. The fundamental components are a almost standard single-stranded tile, two novel double-tile designs for driving growth in any of the four cardinal directions, and a systematic way, using lattice holes, to prevent growth in unwanted directions. Using this, we design a set of 4,096 DNA tiles that we call a *palette*; the experimenter may then 'paint' any desired growth order onto a 32×32 canvas.

Results. We demonstrate 10 growth orders at two spatial block resolutions, $\sim 10 \times 10$ and $\sim 6 \times 6$, including spirals, combs, and S-shapes. Time-lapse atomic force microscopy shows intended growth order without unintended facet growth, even over several days. Despite the large number of tiles, our method typically requires less than 30 minutes of bench work mixing strands, followed by an anneal and temperature hold to grow any desired block growth order.

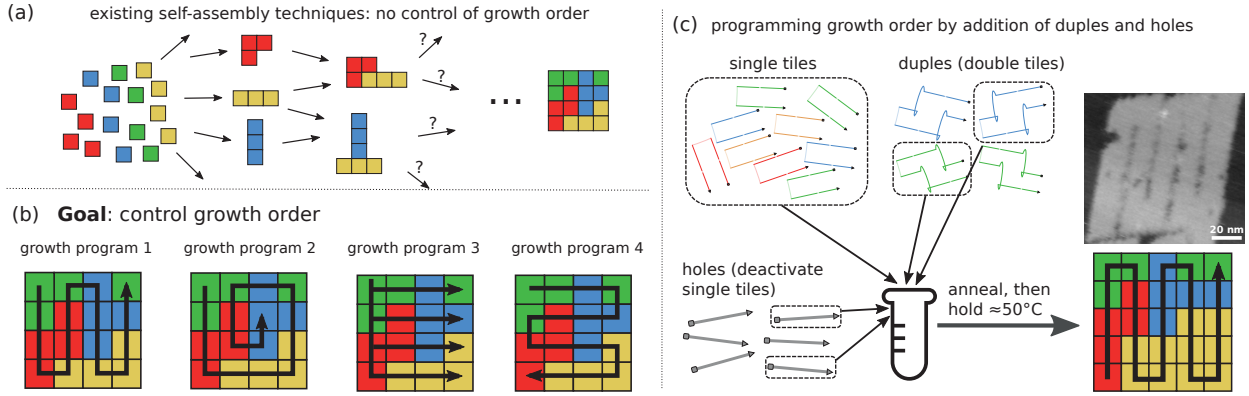


Figure 1: **Goal:** control of the order in which monomers attach to self-assemble a structure. **(a)** Prior tile structure self-assembly techniques do not control the order in which tiles attach. **(b)** Our goal is to enable precise control over the order of attachment. **(c)** Our strategy is to design a single library of DNA strands such that, based on the order in which we want tiles to attach, mixing an appropriate subset of strands from the library leads to self-assembly in the desired order. All "single" tiles are included, as well as "duples" to encourage growth in some directions, and "holes" (strands complementary to single tiles to deactivate them) to discourage growth in some directions. (top right) Atomic force microscope image of sample 5×5 square growth in "vertical zig-zag" order.

Applications include a pathway to high-yield (low error) tile nanostructures (by completely disallowing undesired growth pathways), as well as a stepping stone to the kinds of algorithmic self-assembly that have only appeared in theory papers presented in this conference series.

*Supported by ERC grant # 772766, Active-DNA), and SFI under grant numbers 18/ERCS/5746 and 20/FF-P-P/8843, and NSF awards 1900931 and 1844976. †Hamilton Institute, Department of Computer Science, Maynooth University, Ireland ‡Department of Computer Science, University of California, Davis, USA

Universal Shape Replication Via Self-Assembly With Signal-Passing Tiles (extended abstract)

Andrew Alseth, Daniel Hader, and Matthew J. Patitz

University of Arkansas, USA

In this paper, we investigate shape-assembling power of a tile-based model of self-assembly called the Signal-Passing Tile Assembly Model (STAM). In this model, the glues that bind tiles together can be turned on and off by the binding actions of other glues via ``signals''. In fact, we prove our positive results in a version of the model in which it is slightly more difficult to work (where tiles are allowed to rotate) but show that they also hold in the standard STAM. Specifically, the problem we investigate is ``shape replication'' wherein, given a set of input assemblies of arbitrary shape, a system must construct an arbitrary number of assemblies with the same shapes and, with the exception of size-bounded junk assemblies that result from the process, no others. We provide the first fully universal shape replication result, namely a single tile set capable of performing shape replication on arbitrary sets of any 3-dimensional shapes without requiring any scaling or pre-encoded information in the input assemblies. Our result requires the input assemblies to be composed of signal-passing tiles whose glues can be deactivated to allow deconstruction of those assemblies, which we also prove is necessary by showing that there are shapes whose geometry cannot be replicated without deconstruction. Additionally, we modularize our construction to create systems capable of creating binary encodings of arbitrary shapes, and building arbitrary shapes from their encodings. Because the STAM is capable of universal computation, this then allows for arbitrary programs to be run within an STAM system, using the shape encodings as input, so that any computable transformation can be performed on the shapes.

Acknowledgments: This work was supported in part by NSF grant CAREER-1553166.

Algorithmic DNA origami:
Scaffolded DNA computation in one and two dimensions*

Tristan Stérin† Abeer Eshra† Damien Woods
(†equal contribution)
Hamilton Institute and Department of Computer Science, Maynooth University, Ireland.

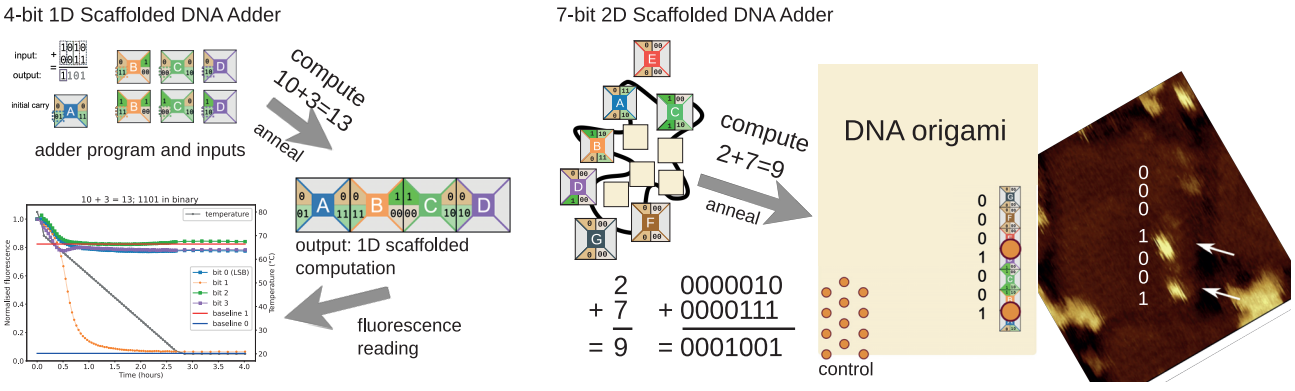
DNA origami has enormous thermodynamic and kinetic barriers to off-target structure formation, throughout a wide range of experimental conditions and without need for careful DNA-sequence design. In contrast, current designs for *algorithmic* DNA nanostructures present nucleation and algorithmic error-handling challenges, partially due to inherently non-equilibrium design.

Proposal. Here, we propose a Scaffolded DNA Computer in 1D and 2D. The idea is simple: a DNA origami staple location implements a computational element (logic gate) using information-encoding toeholds, and the scaffold itself acts as a non-computational substrate that tethers gates together to form a circuit. A small number of input staples trigger the computation, algorithmically driving which staple goes where. Computation may happen during, or even after, an anneal, and the approach lends itself naturally to algorithmic reconfiguration of DNA nanostructures.

Design challenges. This simple-sounding proposal presents several major design challenges, including DNA sequence design of computing elements that tether to adversarial, almost hardcoded, (M13) scaffold sequence, and strand displacement through partially or totally formed helical structures.

Design and results. Theoretically, our approach is framed around a simple theoretical model called the Scaffolded DNA Computer. We take a principled approach to the design challenges by designing a small set of (universal) toeholds strong enough to handle such structural displacement. We implement a computationally expressive 1D Scaffolded DNA Computer, demonstrating a wide range of computations, including: bit-copying, a 4-bit adder, a 4-bit parity checker, a complex 3-state NFA, and programs that multiply by 3 (in base 2) and divide by 2 (in base 3). We report a total of 77 experiments in 1D and we measure an estimated yield of 91 %. We then implement a 7-bit copier and a 7-bit adder using a full 2D DNA origami scaffold, pushing computation through 14 helices. We report 5 additions and 2 bit-copy computations in 2D.

Conclusions. The range of computations carried out, and the testing of both linear and 2D scaffolds, present strong evidence that full-scale 2D, or even 3D, DNA origami based computation is realisable in the not-so-distant future. Such thermodynamically-favoured computation holds much promise as it has the potential to inherit the energetic and structural favourability of arguably the most successful technology of the field, DNA origami, while wielding the power and expressiveness provided by molecular computation.



*Research supported by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 772766, Active-DNA project), and Science Foundation Ireland (SFI) under grant numbers 18/ERCS/5746 and 20/FFP-P/8843.

Wednesday August 10th

8:30-9am	Breakfast
9–10am	KEYNOTE TALK: Strand exchange reactions for analysis and computation (pg 24) Andrew Ellington (University of Texas at Austin, USA)
10–10:30am	Toehold-Mediated Strand Displacement in Random Sequence Pools (pg 25) Thomas Mayer, Lukas Oesinghaus and Friedrich C. Simmel (eligible for best student presentation award)
10:30–11am	Parallel molecular computation on digital data stored in DNA (pg 26) Boya Wang, Siyuan S Wang, Cameron Chalk, Andrew D Ellington and David Soloveichik
11–11:30am	Break
11:30am–12pm	Rule-of-thumb-free geometry-driven design of arbitrary complex curved DNA origami with ENSnano (pg 27) Nicolas Levy, Allan Mills, Gaétan Bellot and Nicolas Schabanel (eligible for best student presentation award)
12–12:30pm	A Coupled Reconfiguration Mechanism for Single-Stranded DNA Strand Displacement Systems (pg 28) Hope Amber Johnson and Anne Condon
12:30–1pm	Supervised learning in DNA-based winner-take-all neural networks (pg 29) Kevin Cherry and Lulu Qian (eligible for best student presentation award)
1–1:30pm	Conference group photograph
1:30–6pm	Excursion to Old Town Albuquerque
6–10pm	Conference dinner at Explora

KEYNOTE TALK: Strand exchange circuits for biomedical applications

Andrew D. Ellington[1]

[1]Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, TX, 78703

One of the remarkable features of DNA nanotechnology is the possibility of carrying out logical operations via DNA strand exchange, itself a unique mechanism with almost no equivalents in chemistry or biology. The Ellington lab has employed strand exchange circuitry to improve the surety and information processing capabilities of molecular diagnostics, and a review of this work will provide insights into translational and real-world applications of DNA nanotechnology, including relative to the ongoing (and perhaps next) pandemic. In addition, though, new developments in the ability to dynamically store and exchange information based on strand exchange (SIMD) makes manifest a potential connection between the ability to do diagnostics at scale, and the ability to store information from such assays directly, in a DNA database. The congruence of information processing and information storage in a potential biomedical application hopefully reveals new frontiers for DNA nanotechnology.

Toehold-Mediated Strand Displacement in Random Sequence Pools

Thomas Mayer, Lukas Oesinghaus, and Friedrich C. Simmel*

Physics Department, TU Munich, D-85748 Garching, Germany

Abstract: Toehold-mediated strand displacement (TMSD) has been used extensively for in vitro sensing and computing as well as for the realization of molecular circuits in vivo. In contrast to in vitro applications using only purified components, the reaction kinetics for circuits in total RNA or in vivo environments are altered by a variety of mechanisms. One such mechanism is the interaction of the TMSD circuit components with interfering strands of other nucleic acids in the environment, e.g., the transcriptome in the case of a living cell. Here, we investigate the impact of a background of random sequences on the kinetics of TMSD circuits. We begin by studying individual interfering strands of defined length and structure and use the obtained data to build an empirical model that estimates the impact of novel interfering strands on kinetics. As expected, the accessibility of the toehold domain turns out to be more critical to the kinetics than that of the branch migration domain. We then investigate the influence of pools of random sequence nucleic acid strands and compare their behavior to pools of defined interfering strands. This results in the following insights: First, we find that the kinetics are most significantly influenced by only a small subpopulation of strongly interacting strands and not by the bulk of the random pool. Consequently, the kinetics inside a random pool can be approximately mimicked by the addition of a much smaller pool of 100 strands representing only the 99th percentile of its strongest binders. Second, sample preparation crucially determines the resulting kinetic behavior because of the equilibration process of the invader strands with the background sequences. The difference in the reaction speed between an unequilibrated and an equilibrated experiment is found to be up to one order of magnitude. Third, we find a surprising difference between the impact of a random DNA and a random RNA background on DNA TMSD circuit kinetics, where the DNA pool appears to slow down the displacement reaction much more compared to RNA. Finally, we compared two established and a novel technique that help to speed up TMSD reactions in the presence of a sequence background: Using a three letter alphabet, protecting the toehold domain by intramolecular secondary structure in the invader strand, and protecting the toehold by an additional blocking strand with multiple small internal loops. While all of these techniques are found to be useful, only the latter can be used without strong constraints on the design space. We expect that our insights will be useful for the construction of displacement circuits that are robust to molecular noise without introducing design constraints.

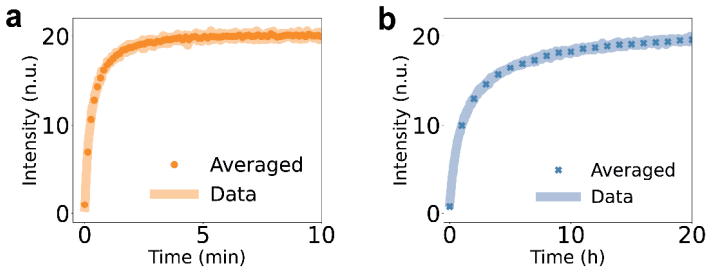


Figure 1. Kinetic curves of a TMSD circuit without molecular noise (a) and with a random sequence pool (b). The reaction speed is shifted from minutes to hours.

Parallel molecular computation on digital data stored in DNA

Boya Wang[+], Siyuan S. Wang[+], Cameron Chalk, Andrew D. Ellington, and David Soloveichik
University of Texas at Austin, USA [+]Equal contribution

DNA is a promising data storage medium due to its high density, stability, and longevity [1]. Typical DNA storage schemes do not allow in-memory computation, and instead computation requires DNA sequencing, in silico computation, followed by synthesis of new DNA. In contrast to this slow and expensive loop, our SIMD||DNA model [2] achieves massively parallel in-memory computation on DNA data storage. In the SIMD||DNA paradigm, a single register storing a binary string is represented by a multi-stranded DNA complex. Information is encoded in the patterns of nicks and exposed single-stranded regions. An instruction (a set of DNA strands) applied to registers with magnetic beads manipulates the information by updating the strand compositions and binding patterns of the registers. To engineer SIMD||DNA in practice, we constructed the registers using the M13 plasmid. We introduced mismatches to the SIMD||DNA design for a secondary encoding, making it compatible with sequencing-based readout. Like a computer's memory, multiplexed information stored in the SIMD||DNA paradigm is compatible with random access and erasure. Experimental results on 4-bit registers show computation for parallel binary counting and elementary cellular automaton Rule 110 computation. Scaling up the computational power, we show that the registers can be stored and reused after a given set of computations by conducting multiple rounds of computation. This is so far the largest strand displacement system using naturally-occurring DNA sequences with 122 distinct strand displacement steps.

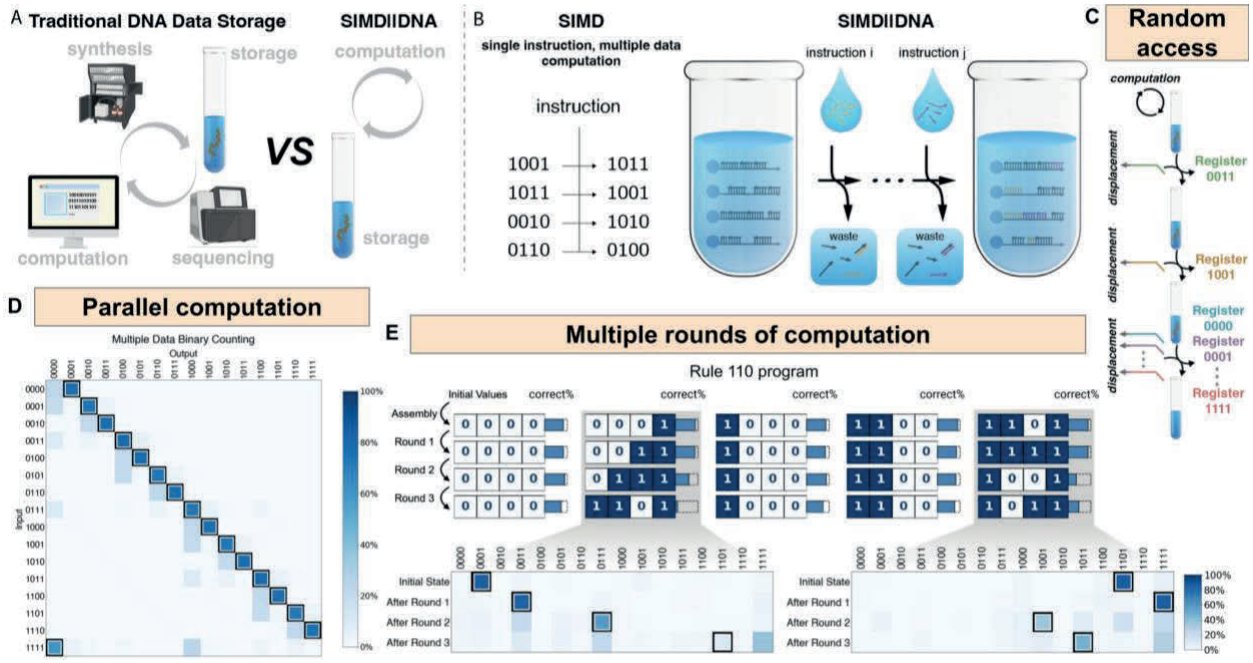


Fig. (A) Traditional DNA data storage versus the SIMD||DNA data storage. SIMD||DNA allows (B) parallel, in-memory computation and (C) random-access memory. (D) Results for the binary counting program performed on 16 registers in parallel show that all inputs were incremented in parallel. (E) Results for multiple rounds of computation for the Rule 110 program.

[1] Church G, Gao Y, and Kosuri S 2012 *Science* **337** 1628–1628.
[2] Wang B, Chalk C, and Soloveichik D 2019 *LNCS* **11648** 219-235.

Rule-of-thumb-free geometry-driven design of arbitrary complex curved DNA origami with ENSnano *(Abstract)*

Nicolas Levy[1] Allan Mills[4] Gaëtan Bellot[*,4] Nicolas Schabanel[*,3,1,2]

1. LIP (UMR 5668), École Normale Supérieure de Lyon, France.
2. IXXI, École Normale Supérieure de Lyon, France.
3. CNRS, France.
4. U. de Montpellier, CBS, INSERM, CNRS, France.

* Co-corresponding authors.

Abstract. The key step in DNA origami design is the positioning of crossovers which constrains the structure geometrically. Various crossover positioning rules of thumb (e.g. [1,2,3]) have been established that allow to reliably produce DNA origamis with the desired shape.

Here, we propose a new method for designing curved origamis that deviates radically from the pattern-based previous approaches in [1,2]. We have developed a new model for DNA double helices curved in 3D that allows us to directly position the DNA double helices constituting the desired shape in the 3D interface of our software ENSnano [4]. The crossovers' positions are then simply deduced from the 3D positions of the nucleotides, as predicted by our model. This geometry-based interactive approach shortcuts the tedious process of manually coming up with a pattern suited for the desired curvature, and furthermore allows to deal transparently with structure whose curvature varies continuously. We also propose an innovative 2D representation synchronizing curved parallel double helices without relying on insertions or deletions, by automatically adapting the cell width for each nucleotide in the array representation (Fig. 1c).

We provide experimental data validating our curvy DNA model by successfully annealing two DNA origamis conceived thanks to two new DNA design methods. The first origami (Fig. 1) consists in a 6-helices bundle following an interactively created bezier curve ⌘ whose curvature gets as low as 4.7nm. The second is an asymmetrical Möbius torus (Fig. 2) whose DNA strands are routed along 2 spiraling helices covering its whole surface. This new spiraling technique, allowed by our DNA curvy model, enables to grasp xovers within a continuous range which results in an easier-to-design and smoother surface. Both of our designs folded as is, without any need to redesign their xover schemes.

Bibliography extract. 1. Dietz *et al*, Science, 2009; 2. Han *et al*, Science, 2011; 3. Zhang *et al*, Nature nano., 2015; 4. Levy *et al*, DNA27, 2021.

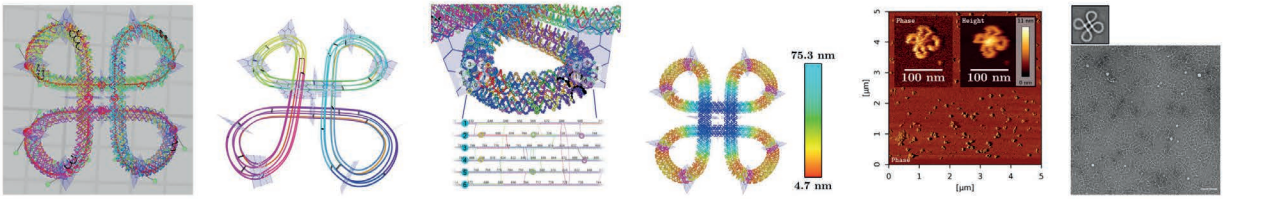


Figure 1: Preliminary data on the Bezier curve ⌘ origami.

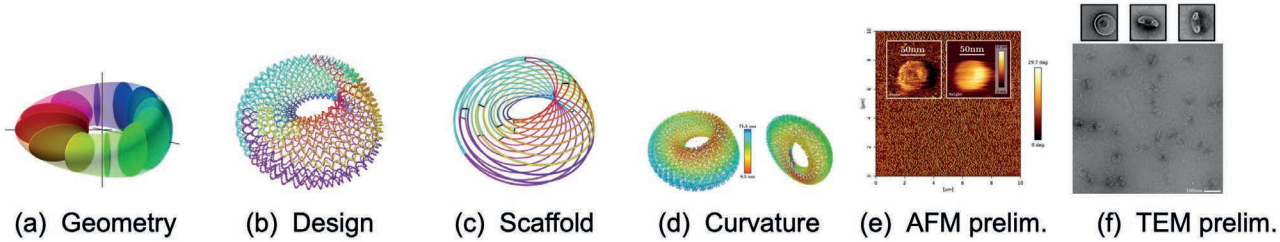


Figure 2: Preliminary data on the Möbius torus origami.

A Coupled Reconfiguration Mechanism for Single-Stranded DNA Strand Displacement Systems

Hope Amber Johnson[1] and Anne Condon[1]

[1]The University of British Columbia, Vancouver, BC, Canada

DNA Strand Displacement (DSD) systems model basic reaction rules, such as toehold-mediated strand displacement and 4-way branch migration, that modify complexes of bound DNA strands. DSD systems have been widely used to design and reason about the correctness of molecular programs, including implementations of logic circuits, neural networks, and Chemical Reaction Networks. Such implementations employ a valuable toolkit of mechanisms—sequences of basic reaction rules—that achieve catalysis, reduce errors (e.g., due to leak), or simulate simple computational units such as logic gates, both in solution and on surfaces. Expanding the DSD toolkit of DSD mechanisms can lead to new and better ways of programming with DNA.

Here we introduce a new mechanism, which we call *controlled reconfiguration*. We describe one example where two single-stranded DSD complexes interact, changing the bonds in both complexes in a way that would not be possible for each independently on its own via the basic reaction rules allowed by the model. We use *coupled reconfiguration* to refer to instances of controlled reconfiguration in which two reactants change each other in this way. We note that our DSD model disallows pseudoknots and that properties of our coupled reconfiguration construction rely on this restriction of the model.

A key feature of our coupled reconfiguration example, which distinguishes it from mechanisms (such as 3-way strand displacement or 4-way branch migration) that are typically used to implement molecular programs, is that the reactants are single-stranded. Leveraging this feature, we show how to use coupled reconfiguration to implement Chemical Reaction Networks (CRNs), with a DSD system that has both single-stranded signals (which represent the species of the CRN) and single-stranded fuels (which drive the CRN reactions). Our implementation also has other desirable properties; for example it is capable of implementing reversible CRNs and uses just two distinct toeholds. We discuss drawbacks of our implementation, particularly the reliance on pseudoknot- freeness for correctness, and suggest directions for future research that can provide further insight on the capabilities and limitations of controlled reconfiguration.

Supervised learning in DNA-based winner-take-all neural networks

Kevin M. Cherry[1] and Lulu Qian[1,2]
[1]*Bioengineering,* [2]*Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA

Learning allows living systems to be born simple but develop into endless diversity and richness; it also accelerates evolution by altering the shape of the fitness landscape – known as the Baldwin effect. Here we show that a test tube of DNA molecules can be programmed to perform supervised learning in winner-take-all neural networks, where the system is exposed to examples of what it may encounter and examples of the desired response; this information is used to improve its capability for handling similar situations in the future. Unlike in previous demonstrations where a DNA neural network could only perform a particular type of pattern classification task once the molecules are mixed together in a test tube, in this work the weight molecules encoding the memories are initially inactive; once trained, an appropriate collection of them will become active. These adaptive weights allow the system to learn from a molecular environment and develop its memories for performing any desired pattern classification tasks within a specified complexity (i.e. number of bits in a pattern and number of classes of patterns).

We demonstrated supervised learning followed by pattern classification. Over the course of learning, input strands representing distinct training patterns were added to the test tube every 24 hours. Together with a label strand, each set of input strands triggered a response of the neural network to adjust its weight activators. We showed that the system correctly classified a representative set of 9-bit patterns after training. Experiments from 4-bit to 9-bit learning and testing were used to evaluate the essential aspects of the system behavior that determines the scalability of the molecular implementation. The understanding was used to guide us through several design iterations, until arriving at the current design, which we believe will enable systems of increasing complexity.

Overall, this work begins to address the question of how chemical systems could gain the capability to learn from an unknown environment, using information encountered earlier to develop memories for processing similar information at a later point. Learning is proven to be powerful and pervasive in biology; demonstrating learning at the molecular level in DNA-based artificial neural networks helps establish a deeper connection between natural algorithms and engineered molecular systems, laying out the foundation for future applications in programmable active materials.

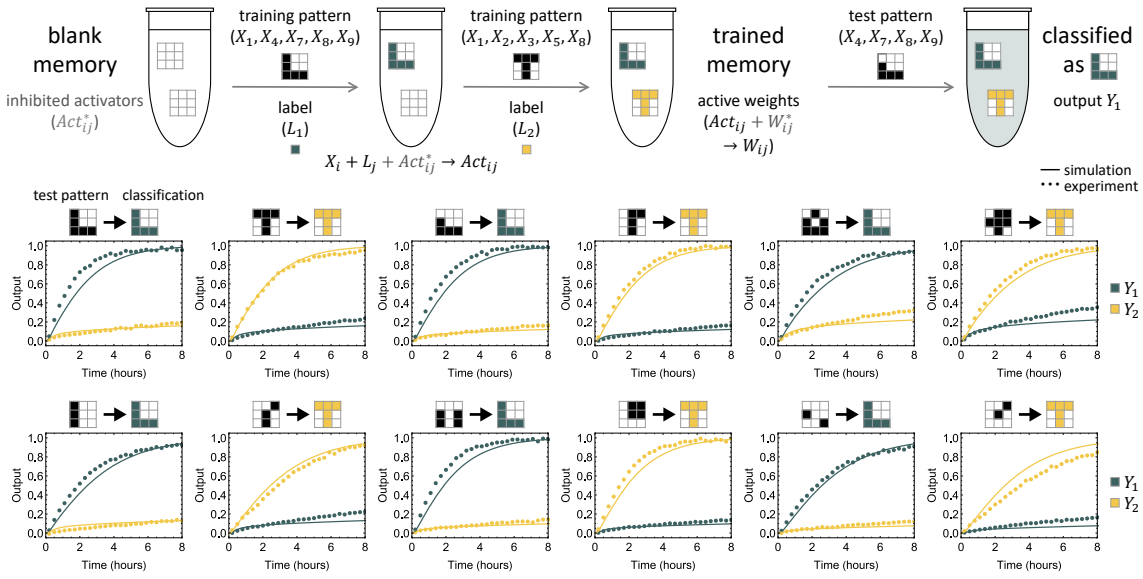


Figure 1: **Pattern classification in a molecularly trained 9-bit 2-memory DNA neural network.**

Thursday August 11th

8:30-9am	Breakfast
9–10:30am	TULIP AWARD KEYNOTE TALK: Computational Design of Nucleic Acid Circuits (pg 32) Andrew Phillips
10:30–11am	Break
11–11:30am	Metal-Mediated Molecular Programming in DNA (pg 33) Simon Vecchioni, Brandon Lu, Yoel Ohayon, Karol Woloszyn, Chengde Mao, James Canary, Nadrian Seeman and Ruojie Sha
11:30am–12:30pm	SPECIAL PANEL: A Tribute to Ned Seeman
12:30–2pm	Lunch (on your own)
2–2:30pm	The Structural Power of Reconfigurable Circuits in the Amoebot Model (pg 34) Andreas Padalkin, Christian Scheideler and Daniel Warner
2:30–3pm	Fault-Tolerant Shape Formation in the Amoebot Model (pg 35) Daniel Warner, Christian Scheideler and Irina Kostitsyna
3–3:30pm	Break
3:30–4pm	Modelling and optimisation of a DNA stack nano-device using probabilistic model checking (pg 36) Bowen Li, Neil Mackenzie, Ben Shirt-Ediss, Natalio Krasnogor and Paolo Zuliani
4–4:30pm	NUPACK: Molecular Programming in the Cloud (pg 37) Mark E. Fornace, Jining Huang, Cody T. Newman, Nicholas J. Porubsky, Marshall B. Pierce and Niles A. Pierce
4:30–4:45pm	Main conference closing

TULIP AWARD KEYNOTE TALK: Computational Design of Nucleic Acid Circuits

Andrew Phillips[1]

[1]AstraZeneca, UK

Information processing circuits made of nucleic acids show great potential for enabling a broad range of biotechnology applications, including high-precision biosensing, diagnostics and therapeutics. This diversity of applications is supported by a range of implementation strategies, including nucleic acid strand displacement, localisation to substrates, and the use of enzymes to process information. This talk presents a programming language that allows such nucleic acid circuits to be designed and analysed. The language extends traditional logic programming with an equational theory to express nucleic acid molecular motifs, and uses logic predicates to automatically generate computational models of circuit behaviour. We also demonstrate how a similar approach can be used to design genetic circuits for reprogramming cell behaviour. More generally, we anticipate that the integration of nucleic acid programming languages with computational modelling, lab automation and machine learning will accelerate the development of future biotechnology applications.

Metal-Mediated Molecular Programming in DNA

Simon Vecchioni,[1] Brandon Lu,[1] Yoel Ohayon,[1] Karol Woloszyn,[1] Chengde Mao,[2] James W. Canary,[1] Nadrian C. Seeman,[1] Ruojie Sha[1]

[1]Department of Chemistry, New York University, New York, NY 10003, USA.

[2]Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA.

DNA as a vehicle for molecular programming relies on the predictability and information storage capability of the canonical Watson-Crick base pairs, A:T and G:C. Since the inception of DNA nanotechnology 40 years ago, designer self-assembly in DNA architectures has relied on the structure-function relationship of the Watson-Crick paradigm. With only four letters, it has become clear that there are many limitations on the diversity and ensuing complexity of self-assembling nanostructures. To address this limitation, we present an expanded DNA alphabet based on metal-mediated DNA (mmDNA) base pairing in which we substitute hydrogen bonds between pyrimidine nucleobases with metal ion coordination. Previous work in the field has identified the superlative thermostability[1,2] of C:Ag^+:C and T:Hg^{2+}:T base pairs: fully-metalated 11mer duplexes melt near 100 °C,[3] but a systematic structural comparison has not been carried out.

To that end, we have characterized 30 unique mmDNA base pairs *via* x-ray diffraction by leveraging the crystalline nature of the 3D tensegrity triangle lattice.[4] In addition to developing a set of rules for metal-base coordination, we have elucidated generalized design parameters for mmDNA nanostructures. These data demonstrate that mmDNA base pairs are amenable to programmable self-assembly, and thereby expand the lexicon of DNA. As a "semantomorphic" (*shape encoding*) tool for molecular programming, mmDNA not only gives rise to topology for branched DNA structures, but also increases the functionality of DNA itself by introducing bioinorganic behavior to DNA through enhanced electron conduction.[5] In this study, we perform calculations to compare energy levels of the base pairs in our library with insights toward molecular electronics. We anticipate the integration of this expanded mmDNA alphabet as a means of programming specific interactions into DNA nanostructures and the exploitation of their electronic, magnetic, and catalytic properties for years to come.

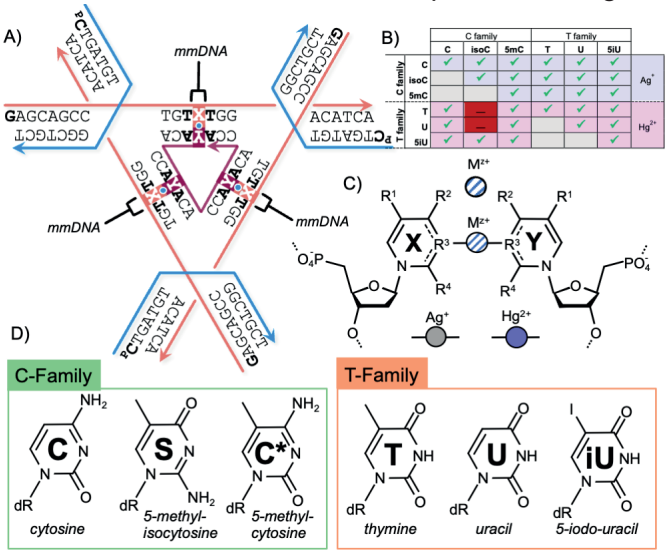


Figure 1. mmDNA design schematic used in this study. **A)** The DNA tensegrity triangle motif contains a pyrimidine-pyrimidine (X:Y) mismatch at the center of each asymmetric unit. **B)** 30 mmDNA base pair combinations of natural and unnatural pyrimidine bases (green checks). Two combinations preclude crystallization (red boxes). Isomeric combinations are omitted (grey boxes). **C)** The generalized pyrimidine-pyrimidine pair is shown with the functional groups modified in this study. **D)** C-like and T-like pyrimidines.

[1] Torigoe, H. *et al. Biochimie* **94**, 2431-2440, (2012).
[2] Wang, Y., Ritzo, B. & Gu, L.-Q. *RSC Advances* **5**, 2655-2658, (2015).
[3] Vecchioni, S. *et al. Scientific Reports* **9**, 6942, (2019).
[4] Zheng, J. *et al. Nature* **461**, 74-77, (2009).
[5] Toomey, E. *et al. J Phys Chem C* **120**, 7804-7809, (2016).

The Structural Power of Reconfigurable Circuits in the Amoebot Model

Andreas Padalkin[1], Christian Scheideler[1], and Daniel Warner[1]

[1]Universität Paderborn, Germany

The *amoebot model* [Derakhshandeh et al., SPAA 2014] has been proposed as a model for programmable matter consisting of tiny, robotic elements called *amoebots*. We consider the *reconfigurable circuit extension* [Feldmann et al., JCB 2022] of the geometric (variant of the) amoebot model that allows the amoebot structure to interconnect amoebots by so-called *circuits*. A circuit permits the instantaneous transmission of signals between the connected amoebots. In this paper, we examine the structural power of the reconfigurable circuits. We start with some fundamental problems like the *stripe computation problem* where, given any connected amoebot structure S, an amoebot u in S, and some axis X, all amoebots belonging to axis X through u have to be identified. Second, we consider the *global maximum problem*, which identifies an amoebot at the highest possible position with respect to some direction in some given amoebot (sub)structure. A solution to this problem can then be used to solve the *skeleton problem*, where a (not necessarily simple) cycle of amoebots has to be found in the given amoebot structure which contains all boundary amoebots. A canonical solution to that problem can then be used to come up with a canonical path, which provides a unique characterization of the shape of the given amoebot structure. Constructing canonical paths for different directions will then allow the amoebots to set up a spanning tree and to check symmetry properties of the given amoebot structure. The problems are important for a number of applications like rapid shape transformation, energy dissemination, and structural monitoring. Interestingly, the reconfigurable circuit extension allows polylogarithmic-time solutions to all of these problems.

Fault-Tolerant Shape Formation in the Amoebot Model

Irina Kostitsyna[1], Christian Scheideler[2], and Daniel Warner[2]

[1]Department of Mathematics and Computer Science, TU Eindhoven, The Netherlands,
[2]Department of Computer Science, Universität Paderborn, Germany

The amoebot model is a distributed computing model of programmable matter. It envisions programmable matter as a collection of computational units called amoebots or particles that utilize local interactions to achieve tasks of coordination, movement and conformation. In the geometric amoebot model the particles operate on a hexagonal tessellation of the plane. Within this model, numerous problems such as leader election, shape formation or object coating have been studied. One area that has not received much attention so far, but is highly relevant for a practical implementation of programmable matter, is fault tolerance. The existing literature on that aspect allows particles to crash but assumes that crashed particles do not recover. We proposed a new model in which a crash causes the memory of a particle to be reset and a crashed particle can detect that it has crashed and try to recover using its local information and communication capabilities. We present an algorithm that solves the hexagon shape formation problem in our model if a finite number of crashes occur and a designated leader particle does not fail. At the heart of our solution lies a fault-tolerant implementation of the spanning forest primitive, which, since other algorithms in the amoebot model also make use of it, is also of general interest.

Modelling and Optimisation of a DNA Stack Nano-Device Using Probabilistic Model Checking

Bowen Li[1], Neil Mackenzie[1], Ben Shirt-Ediss[1], Natalio Krasnogor[1], and Paolo Zuliani[1]

[1]Interdisciplinary Computing and Complex bioSystems (ICOS) Research Group, School of Computing, Newcastle University, Newcastle upon Tyne, UK

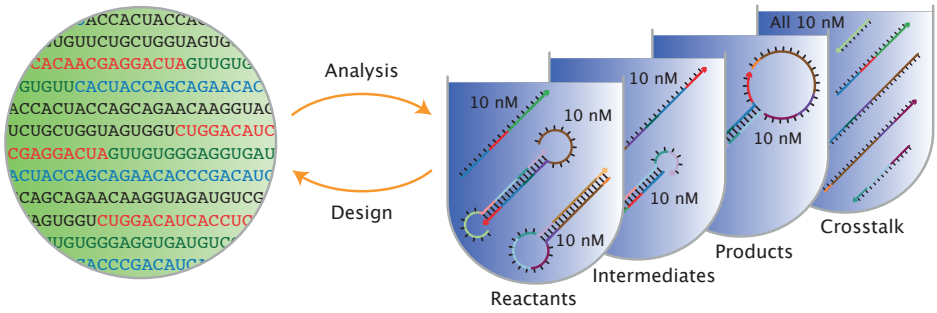
A DNA stack nano-device is a bio-computing system that can read and write molecular signals based on DNA-DNA hybridisation and strand displacement. *In vitro* implementation of the DNA stack faces a number of challenges affecting the performance of the system. In this work, we apply probabilistic model checking to analyse and optimise the DNA stack system. We develop a model framework based on *continuous-time Markov chains* to quantitatively describe the system behaviour. We use the PRISM probabilistic model checker to answer two important questions: 1) What is the minimum required incubation time to store a signal? And 2) How can we maximise the yield of the system? The results suggest that the incubation time can be reduced from 30 minutes to 5-15 minutes depending on the stack operation stage. In addition, the optimised model shows a 40% increase in the target stack yield.

NUPACK: Molecular Programming in the Cloud

Mark E. Fornace,[1,2,#] Jining Huang,[1,#] Cody T. Newman,[1,#] Nicholas J. Porubsky,[2] Marshall B. Pierce, and Niles A. Pierce[1,3,*]

[1]Division of Biology & Biological Engineering, Caltech, [2]Division of Chemistry & Chemical Engineering, Caltech, [3]Division of Engineering & Applied Science, Caltech, [#]Authors contributed equally. *Corresponding author: niles@caltech.edu

NUPACK is a growing software suite for the analysis and design of nucleic acid structures, devices, and systems serving the needs of researchers in the fields of molecular programming, nucleic acid nanotechnology, synthetic biology, and across the life sciences. NUPACK algorithms are unique in treating complex and test tube ensembles containing arbitrary numbers of interacting strand species, providing crucial tools for capturing concentration effects essential to analyzing and designing the intermolecular interactions that are a hallmark of these fields. The all-new NUPACK web app has been re-architected for the cloud, leveraging a cluster that scales dynamically in response to user demand to enable rapid job submission and result inspection even at times of peak user demand. The web app exploits the all-new NUPACK 4 scientific code base as its backend, offering enhanced physical models (coaxial and dangle stacking subensembles), dramatic speedups (e.g., 20-120x for test tube analysis), and increased scalability for large complexes. NUPACK 4 algorithms can also be run locally using the all-new NUPACK Python module.



Friday August 12th (Computational Modeling Workshop Day)

9:15–9:30am	Workshop welcome and introduction
9:30–10:00am	The Tall, Thin Molecular Programmer (pg 40) Erik Winfree
10:00–10:30am	Automated Leak Analysis of Nucleic Acid Circuits (pg 41) Andrew Phillips
10:30–11am	Break
11am–1pm	NUPACK tutorial: analysis and design of nucleic acid systems with the all-new NUPACK cloud web app (pg 42) Mark Fornace, Jining Huang, Cody T. Newman, and Niles A. Pierce
1–2:30pm	Lunch (on your own)
2:30–3pm	A roadmap for the future development of ENSnano (pg 43) Nicolas Levy
3–3:30pm	The oxDNA ecosystem (pg 44) Petr Sulc
3:30–4:30pm	Open format panel / discussion
4:30–4:45pm	Workshop closing

The Tall, Thin Molecular Programmer

Erik Winfree[1,2,3]

[1]Computer Science, [2]Computation and Neural Systems, [3]Bioengineering, California Institute of Technology, Pasadena, CA, USA

The period from 1950 to 1980 roughly spans the invention of the transistor to the general purpose home computer. During that time, the complexity of electrical circuits exploded, and computer engineers established the foundations for designing such systems — thereby enabling the million-fold increase of complexity we've experienced since then. DNA nanotechnology has grown from tens of nucleotides in the early 1980s to system designs specifying millions of nucleotides today, which is a roughly comparable rate of growth. What lessons do the early years of the computer revolution offer to our field, especially with respect to how computational models can be used to help master often bewildering complexity?

Automated Leak Analysis of Nucleic Acid Circuits

Andrew Phillips[1]

[1]AstraZeneca, UK

Nucleic acids are a powerful engineering material that can be used to implement a broad range of computational circuits at the nanoscale, with potential applications in high-precision biosensing, diagnostics, and therapeutics. However, nucleic acid circuits are prone to leaks, which result from unintended displacement interactions between nucleic acid strands. Such leaks can grow combinatorially with circuit size, are challenging to mitigate, and can compromise circuit behaviour. Here we present a method for the automated leak analysis of nucleic acid circuits, as described in [1]. Our method extends the logic programming functionality of the Visual DSD language, developed for the design and analysis of nucleic acid circuits, with predicates for leak generation and for excluding low probability leak reactions, together with a leak reaction enumeration algorithm. We use our method to identify leak reactions affecting the performance of control circuits and to analyse leak mitigation strategies. We design control circuits with reduced leakage, which can in turn serve as building blocks for more complex circuits. By integrating our method within an open-source nucleic acid circuit design tool, we enable the leak analysis of a broad range of circuits, as a step toward facilitating robust and scalable nucleic acid circuit design.

[1] Zarubiieva, I., Spaccasassi, C., Kulkarni, V., & Phillips, A. (2022). Automated Leak Analysis of Nucleic Acid Circuits. ACS Synthetic Biology, 11(5), 1931-1948. https://doi.org/10.1021/acssynbio.2c00084

NUPACK tutorial: analysis and design of nucleic acid systems with the all-new NUPACK cloud web app

Mark E. Fornace,[1,2,#] Jining Huang,[1,#] Cody T. Newman,[1,#] and Niles A. Pierce[1,3,*]

[1]Division of Biology & Biological Engineering, Caltech, [2]Division of Chemistry & Chemical Engineering, Caltech, [3]Division of Engineering & Applied Science, Caltech, [#]Authors contributed equally. *Corresponding author: niles@caltech.edu

This hands-on tutorial will cover analysis and design of nucleic acid systems using the all-new NUPACK cloud web app. Live demos will illustrate dramatic performance enhancements using the all-new NUPACK 4 backend. The analysis tutorial will cover elucidation of the properties of complex and test tube ensembles. The design tutorial will provide a detailed introduction to multistate test tube design with application to reaction pathway engineering (and large-scale structural engineering), including coverage of reaction pathway specification, target test tube specification, the ensemble defect, hard constraints, soft constraints, and defect weights. Participants will be able to formulate and run their own analysis and design jobs of interest during the tutorial.

A roadmap for the future development of ENSnano

Nicolas Levy[1]

[1]LIP (UMR5668), École Normale Supérieure de Lyon, France

ENSnano is a software for designing DNA nanostructures that aims to offer the best possible user experience, without compromising on design capabilities. The early development of ENSnano has been focused on establishing a sound geometric model of DNA, as well as designing efficient user interfaces to manipulate DNA nanostructures. We believe that these goals have been reached and that the next step is to expand ENSnano's capabilities beyond the design of DNA origamis, with the long term hope of providing a general-purpose DNA nanostructures design software. In this presentation we will make an overview of the current practices and needs in DNA nanotechnology, and identify what ENSnano is missing to integrate in the existing workflows. We will discuss possible solutions as well as their future integration in ENSnano. More precisely, we will discuss planed features that will improve ENSnano's capabilities in the following areas: Automation, planing of experiments, design of dynamical structures, design of parametrised structures, and sequence design.

The oxDNA ecosystem

Petr Sulc[1]

[1]School of Molecular Sciences and Center for Molecular Design and Biomimetics, The Biodesign Institute, Arizona State University, 1001 South McAllister Avenue, Tempe, AZ 85281, USA

Here, we give a general overview of cyberinfrasructure built for computational analysis of DNA/RNA and protein-DNA hybrid nanotechnology. We will discuss the recent upgrades to the oxDNA code, which implements the coarse-grained models oxDNA, oxRNA and ANM-oxDNA for computational studies of nanostructures and has recently been extended to include new Python interface for simulation and analysis. We will next discuss oxView.org [http://oxview.org/], a browser-based DNA nanotechnology editing tool, which can be used to setup and analyze oxDNA simulations, as well as work as standalone design tool with scripting interface. We will showcase oxDNA.org [http://oxdna.org/], a free public webserver for oxDNA simulations, and finally we will also show Nanobase.org [http://nanobase.org/], a recently launched online repository for sharing DNA/RNA protein-DNA hybrid nanostructures design, with support for conversion between different file formats.

Poster Session 1 - Monday August 8th, 5:30-7:30pm

#42 Multi-Objective Sequence Selection for Scaffolded Origami Nanostructures (pg 48)
Ben Shirt-Ediss, Jordan Connolly, Emanuela Torelli, Silvia Navarro, Juan Elezgaray, Jaume Bacardit and Natalio Krasnogor (eligible for best student / postdoc poster award)

#47 Single-pass transmembrane double-stranded DNA with functional toeholds for non-destructive intra- and extravesicular nucleic acid target recognition (pg 49)
Gde Bimananda Mahardika Wisna, Nirbhik Acharya, Tarushyam Mukherjee, Ranjan Sasmal, Hao Yan and Rizal Hariadi (eligible for best student / postdoc poster award)

#48 Digital Nucleic Acid Memory 2.0: Advances in storage capacity and ease-of-use (pg 50)
Sarah Kobernat and Eric Hayden (eligible for best student / postdoc poster award)

#49 Design and analysis of compound nanostructures using oxView (pg 51)
Michael Matthies, Joakim Bohlin, Erik Poppleton, Jonah Procyk and Petr Šulc

#50 Using oxDNA simulations to characterize structural consequences of design modifications on a DNA origami leaf-spring engine (pg 52)
Erik Poppleton, Mathias Centola, Michael Famulok and Petr Šulc (eligible for best student / postdoc poster award)

#51 Operant conditioning of stochastic chemical reaction networks (pg 53)
David Arredondo and Matthew R. Lakin (eligible for best student / postdoc poster award)

#52 Computational analysis of conformational changes of nucleic acid nanostructures induced by overhangs (pg 54)
Michael Matthies, Matthew Sample, Lu Yu and Petr Sulc (eligible for best student / postdoc poster award)

#54 Unsupervised and Interpretable Machine Learning for Aptamer Analysis (pg 55)
Jonah Procyk and Petr Sulc (eligible for best student / postdoc poster award)

#56 Molecular-responsive DNA nanopore for biomolecular measurements (pg 56)
Hiromu Akai and Kan Shoji (eligible for best student / postdoc poster award)

#61 Dendric DNA Origami for Efficient DDS Carrier (pg 57)
Akinori Kuzuya

#64 Multi-gate Boolean processors in a cell free transcriptional platform (pg 58)
Judee Sharon, Kate Adamala, Chelsea Dasrath, Aiden Fujiwara, Alessandro Synder, Mace Blank and Sam O'Brien (eligible for best student / postdoc poster award)

#67 Developmental self-assembly of a DNA ring with stimulus-responsive size and growth direction (pg 59)
Allison Glynn, Samuel Davidson and Lulu Qian (eligible for best student / postdoc poster award)

#71 Throttling enables rate-insensitive chemical reaction networks (pg 60)
Dominic Scalise and Lulu Qian

#73 Simple software to design recipes for complicated sample mixes (pg 61)
Constantine Evans, David Doty and Damien Woods

#76 Stochastic chemical reaction networks for graph coloring (pg 62)
Philippa Richter, Salvador Buse and Erik Winfree (eligible for best student / postdoc poster award)

#77 Tile-displacement-based shape reconfiguration in DNA origami tile assemblies (pg 63)
Namita Sarraf, Kellen Rodriguez and Lulu Qian (eligible for best student / postdoc poster award)

#79 Stochastic surface chemistry can solve hard problems! (pg 64)
Mohini Misra, Jean-Sebastien Paul, Salvador Buse and Erik Winfree (eligible for best student / postdoc poster award)

#82 Neural Computation in Boltzmann Liquids, Hopfield Droplets, and Murugan Condensation (pg 65)
Salvador Buse, Arvind Murugan and Erik Winfree (eligible for best student / postdoc poster award)

#86 Molecular computation using DNA-based synthetic condensates (pg 66)
Sungho Do, Chanseok Lee, Taehyun Lee, Do-Nyun Kim and Yongdae Shin (eligible for best student / postdoc poster award)

#87 Multi-Fidelity Parameter Inference for an Arrhenius Model of DNA Elementary Step Kinetics (pg 67)
Jordan Lovrod, Boyan Beronov, Anne Condon and Erik Winfree (eligible for best student / postdoc poster award)

#88 Fridge Compiler: automated synthesis of molecular circuits (pg 68)
Lancelot Walthieu, Gus Smith, Luis Ceze and Chris Thachuk (eligible for best student / postdoc poster award)

#89 Predicting accurate ab initio DNA electron densities with equivariant neural networks (pg 69)
Alex Lee, Joshua Rackers, and William Bricker

Multi-Objective Sequence Selection for Scaffolded Origami Nanostructures

Ben Shirt-Ediss[1], Jordan Connolly[1], Emanuela Torelli[1], Silvia Adriana Navarro[1], Juan Elezgaray[2], Jaume Bacardit[1] and Natalio Krasnogor[1]

[1]Interdisciplinary Computing and Complex Biosystems Research Group, School of Computing, Newcastle University, Newcastle-upon-Tyne, NE4 5TG, UK
[2]CRPP, UMR5031, 33600 Pessac, France

Given a 2D or 3D scaffolded origami design, we report a new multi-objective algorithm for selecting appropriate scaffold and staple sequences (Figure 1). The algorithm processes a large pool of scaffold/staple sequence sets realising the same origami and each sequence set is scored across four metrics M1, M2, M3, M4. Each metric represents a different kind of off-target interaction potentially able to prohibit origami self-assembly. The sequence sets are mapped to a 3D or 4D objective space where the pareto front of non-dominated sequence sets is calculated. The pareto set is further subject to Multi-Criteria Decision Making methods to automatically rank solutions. Filters are also applied to signal if certain pareto origamis have local problems like G-quadruplexes in staples, or staples that are able to hybridise with perfect complementarity to the designed scaffold domains of other staples. On evaluating this summary information, a decision maker performs the final choice of which sequence set to order. Initial computational results for seven 2D and seven 3D origamis reveal that sequence sets often exist which have significantly less off-target interactions that the population average (typically a 10%-40% relative decrease in each metric). The selector can also be used to identify optimal DeBruijn sequences [1]. We are testing two 2D origamis in the lab to verify if low/high metric scores translate to good/poor origami assembly yield. Our hypothesis is that low metric scores should give good assembly yield even under fast temperature ramps.

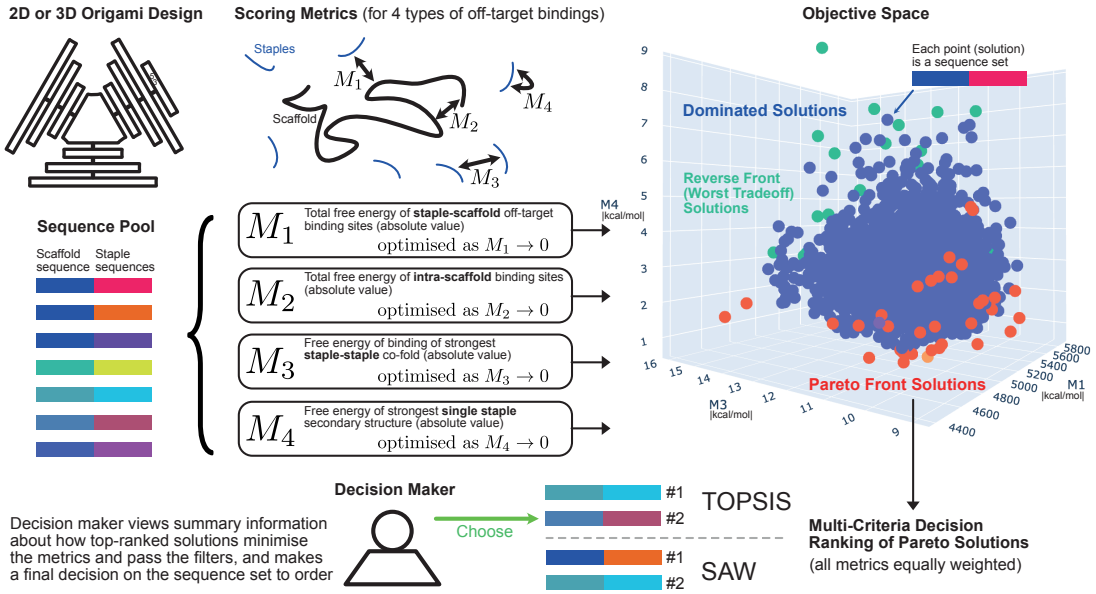


Figure 1: Multi-objective sequence selection process applied to an existing origami design.

This work was supported by: EPSRC EP/N031962/1; European Union's Horizon 2020 research and innovation programme under grant agreement no. 899833; Royal Society International Exchange Grant IES\R1\180080; Royal Academy of Engineering Chair in Emerging Technologies to NK.

[1] J. Kozyra et al. 2017, *ACS Synth. Biol* **6**, **7**, 1140–1149

Single-pass transmembrane double-stranded DNA with functional toeholds for non-destructive intra- and extravesicular nucleic acid target recognition

Gde Bimananda M. Wisna[1,2], Nirbhik Acharya[2], Tarushyam Mukherjee[2], Ranjan Sasmal[2], Hao Yan[2,3], Rizal F. Hariadi[1,2]

[1]Department of Physics, Arizona State University, USA, [2]Biodesign Institute, Arizona State University, USA, [3]School of Molecular Sciences, Arizona State University, USA

Cells express nucleic acid with genetic information specific to cell types and their states. Additionally, the interaction between cells and extracellular ligands, as well as intracellular biomarkers, tune cellular activities. These biomarkers, often in the form of exRNA produced inside and secreted by other cells, facilitate intercellular signaling and communications. Therefore, it is important to recognize intra- and extracellular nucleic acids simultaneously while keeping the cells intact for further studies and cell classification. A relatively straightforward path to achieve recognition of nucleic acid targets is to use DNA nanostructures that are complementary to the nucleic acid targets. In addition, mimicking membrane proteins that span the membrane provides an attractive way for intra- and extracellular target recognition and to transduce information across the membranes. Here, we study cholesterol-modified double-stranded DNA (dsDNA) sensors spanning the lipid membrane whose toeholds can recognize nucleic acids target inside and outside synthetic lipid bilayer. The dsDNA sensors have three common domains consisting of extracellular, transmembrane, and intracellular domains, mimicking single-pass transmembrane proteins. As a proof-of-concept, we utilize synthetic 10–20 μm giant unilamellar vesicles (GUVs) as our cell-like lipid bilayer compartment model. The two-strand design provides a simple approach significantly minimizing the stoichiometric issue of structure formation. We show that with the two cholesterols modification only on one of the duplex strands, the structure still spans the membrane and binds to the target inside GUVs due to the duplex helicity. Our preliminary data with mammalian cells (HEK 293T) show that dsDNA sensors successfully anchor on the plasma membrane and achieve extracellular DNA target recognition. The dsDNA sensor keeps the cells and vesicles intact during recognition. It paves a way toward a more complex signal transmission process across lipid bilayers to enable DNA signal amplification and manipulations in the outer space of vesicles or cells, providing target binding from the inside.

Digital Nucleic Acid Memory 2.0: Advances in storage capacity and ease-of-use

Sarah E. Kobernat[1], Amanda Wolf[2], Ben Balzer[2], Luca Piantanida[3], George Dickinson[3], Golam Mortuza[4], Tim Anderson[4], Wan Kuang[5], Will Hughes[3], Eric Hayden[2]

1 Biomolecular Sciences graduate programs, Boise State University, USA
2 Department of Biological Sciences, Boise State University, USA
3 Micron School of Materials Science and Engineering, Boise State University, USA
4 Department of Computer Science, Boise State University, USA
5 Department of Electrical and Computer Engineering, Boise State University, USA

It is predicted that in the next 10-15 years large data storage will be prohibitively expensive or entirely unavailable [1]. While the rate of data generation worldwide is quickly outpacing our current storage capacity, DNA is a promising medium to address this deficit because of its longevity and data density. The Nucleic Acid Memory Institute at Boise State University recently published a prototype of digital nucleic acid memory (dNAM) that uses DNA origami as a breadboard upon which fluorescent probes transiently bind at addressable locations in an encodable matrix format [2]. However, the prototype was limited in size by the commercially available scaffold used and hundreds of liquid handling steps required to synthesize each origami. To overcome these challenges, we designed a 60% larger node, and cloned and produced a custom 11,054 nt scaffold. We also validated staple strands that were synthesized as oligo pools, eliminating hundreds of liquid handling steps. The larger custom origami was confirmed by AFM and super resolution microscopy. The results indicate that we can replicate our prototype using 6 different origami instead of 15, with orders of magnitude fewer liquid handling steps. These advances improve the scalability of dNAM through higher data capacity and more efficient synthesis, moving the platform toward real world applications.

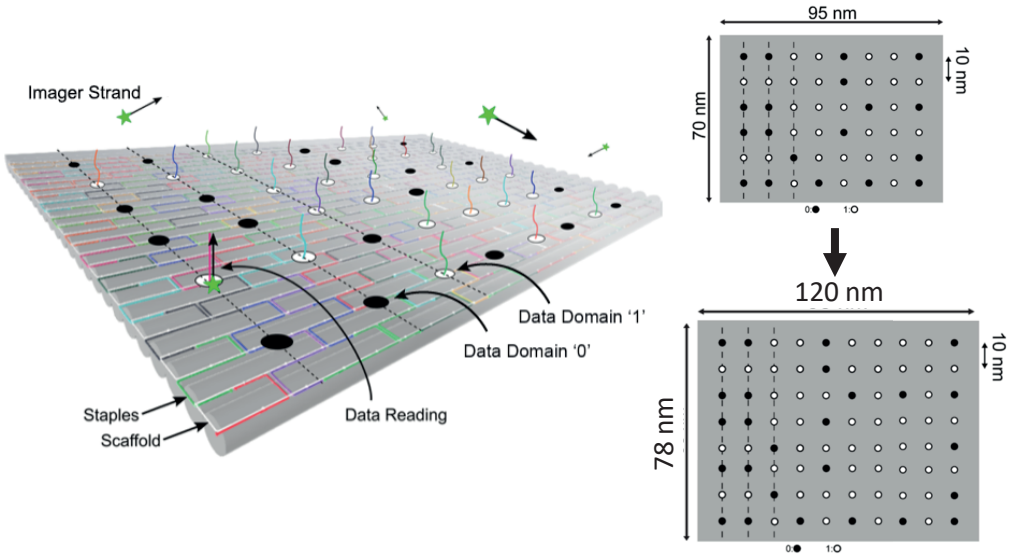


Figure 1. Left- schematic of dNAM. Right- old dNAM design vs new design with 60% more data points using an 11KB scaffold. Adapted from [2]

References:

1. Victor, Z. 2018 Semiconductor Synthetic Biology Roadmap. 1-36 (2018). https://doi.org/10.13140/RG.2.2.34352.40960.
2. Dickinson, G.D., Mortuza, G.M., Clay, W. *et al.* An alternative approach to nucleic acid memory. *Nat Commun* **12,** 2371 (2021). https://doi.org/10.1038/s41467-021-22277-y

Design and analysis of compound nanostructures using oxView

Michael Matthies[1], Joakim Bohlin[1,2], Erik Poppleton [1], Jonah Procyk [1], Petr Šulc[1]

1. School of Molecular Sciences and Center for Molecular Design and Biomimetics, The Biodesign Institute, Arizona State University, 1001 South McAllister Avenue, Tempe, AZ 85281, USA

2. Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK.

Molecular simulation has become an integral part of the DNA/RNA nanotechnology research pipeline. In particular, understanding the dynamics of structures and single-molecule events has improved the precision of nano scaffolds and diagnostic tools. We present oxView [1], [2], a design tool for visualization, design, editing and analysis of simulations of DNA, RNA and nucleic acid–protein nanostructures. oxView provides an accessible software platform for designing novel structures, tweaking existing designs, preparing them for simulation in the oxDNA/RNA molecular simulation engine [3], [4], [5], [6], [7] and creating visualizations of simulation results.

A particular strength of oxView is the preparation of compound designs consisting of multiple nano structures potentially prepared in different design tools. In this poster we show some of the ways these compound structures can be manipulated in oxView. Including the shape complementary bricks [8] [Fig 1] as well as scripting examples preparing 3d crystal lattices of provided origami.

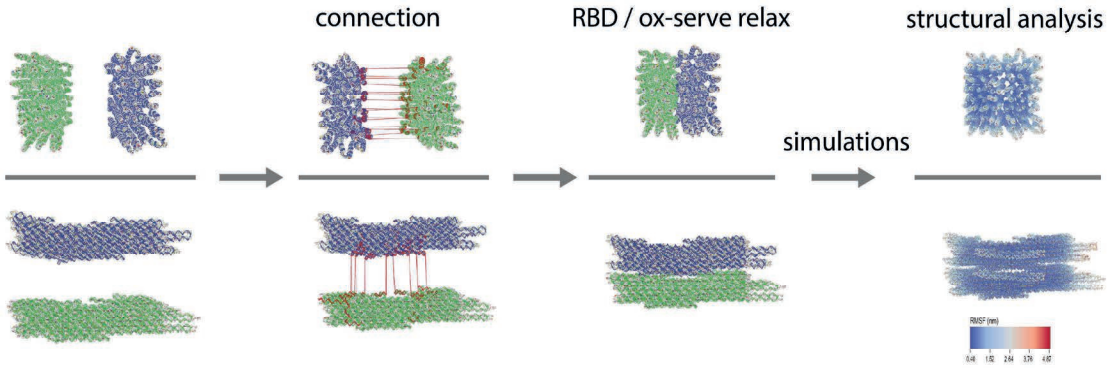


Figure1: Illustration of compound structure assembly using oxView. [8]

These procedures provide a practical basis from which researchers, including experimentalists with limited computational experience, can integrate simulation and 3D visualization into their existing research programs.

1. E. Poppleton, J. Bohlin, M. Matthies, S. Sharma, F. Zhang, P. Šulc, NAR (2020) (DOI: 10.1093/nar/gkaa417)

2. J. Bohlin, M. Matthies, E. Poppleton, J. Procyk, A. Mallya, H. Yan, P. Šulc accepted Nat. Prot. 2022 (DOI: 10.1038/s41596-022-00688-5)

3. B. E. K. Snodin, F. Randisi, M. Mosayebi, P. Šulc, J. S. Schreck, F. Romano, T. E. Ouldridge, R. Tsukanov, E. Nir, A. A. Louis, J. P. K. Doye, J. Chem. Phys. (2015) (DOI: 10.1063/1.4921957)

4. P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye, A. A. Louis, J. Chem. Phys. (2012) (DOI: 10.1063/1.475413)

5. L. Rovigatti, P. Šulc, I. Z. Reguly, F Romano, J. Comput. Chem. (2015) (DOI: 10.1002/jcc.23763)

6. T. E. Ouldridge, A. A. Louis and J. P. K. Doye, J. Chem. Phys, (2011) (DOI: 10.1063/1.3552946)

7. P. Šulc, F. Romano, T. E. Ouldridge, J. P. K. Doye, A. A. Louis: A nucleotide-level coarse-grained model of RNA, J. Chem. Phys. (2014) (DOI: 10.1063/1.4881424)

8. T. Gerling, K.F. Wagenbauer, A.M. Neuner, H. Dietz, Science (2015) (DOI: 10.1126/science.aaa5372)

Using oxDNA simulations to characterize structural consequences of design modifications on a DNA origami leaf-spring engine

Erik Poppleton[1], Mathias Centola[2], Michael Famulok[2], and Petr Šulc[1]

[1]Arizona State University, USA, [2]University of Bonn, Germany

We present here a series of oxDNA[1-4] simulations performed in support of experimental realization of a chemically-driven leaf-spring nanoengine (NE)[5]. The NE is composed of two rigid DNA bundles connected at the vertex by a two-layered flexure. Motion is generated by attaching a T7 RNA polymerase to one of the arms and upon introduction of NTPs, the polymerase binds to, processes along, and detaches from a double-stranded bridge, creating a rhythmic flapping motion. Simulations rationalized the effects of structural motifs and design choices, including the removing the double-stranded bridge (NB), optimizing the sequence of single-stranded regions (NS), introducing nicks in the bridge (nNE), and coupling motion of a driver structure to a passive follower.

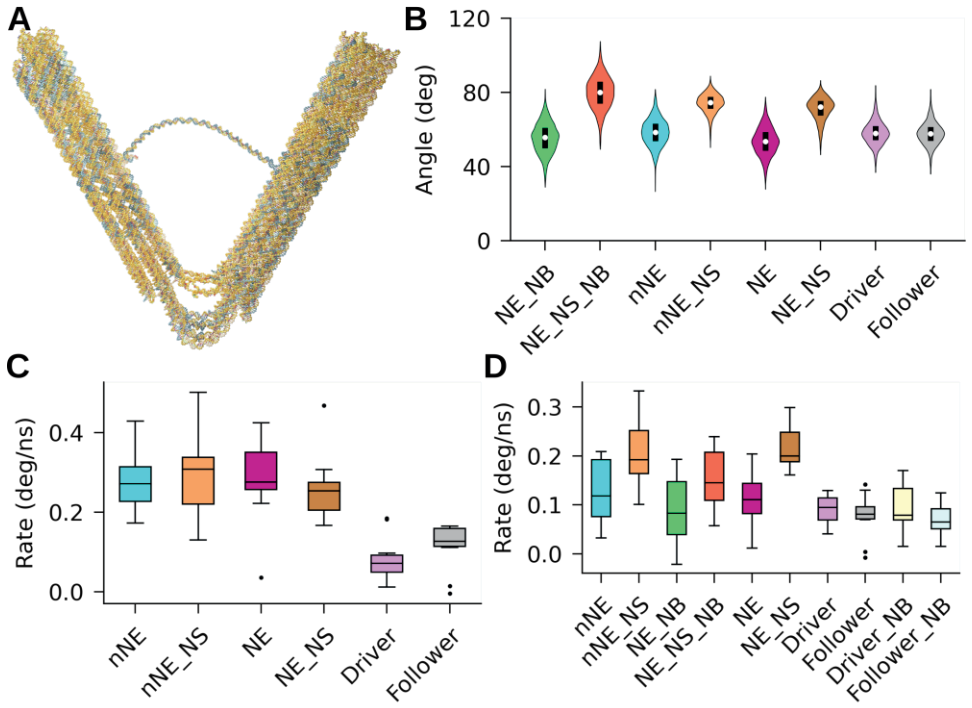


Fig 1. (A) The mean structure of the leaf spring engine during an oxDNA simulation. **(B)** The simulated angle distribution of each design at equilibrium. **(C)** The rate of closing for each design under 16 pN applied external force. **(D)** The rate of opening for each design equilibrated in the closed state and then released.

References:

[1] Ouldridge T E et al 2011 *J. Chem. Phys*. **134**, 085101.
[2] Šulc P et al 2012 *J. Chem. Phys*. **137**, 135101.
[3] Snodin B E K et al 2015 *J. Chem. Phys*. **142**, 234901
[4] Rovigatti L et al 2015 *J. Comput. Chem* **36**, **1**
[5] Centola et al. 2021 *BioRxiv* 2021.12.22.473833

Operant conditioning of stochastic chemical reaction networks

David Arredondo[1], and Matthew R. Lakin[1]

[1]University of New Mexico, USA

Adapting one's behavior to environmental conditions and past experience is a key trait of living systems. In the biological world, there is evidence for adaptive behaviors such as learning even in naturally occurring, non-neural, single-celled organisms. In the bioengineered world, advances in synthetic cell engineering and biorobotics have created the possibility of implementing lifelike systems engineered from the bottom up. This will require the development of programmable control circuitry for such biomimetic systems that is capable of realizing such non-trivial and adaptive behavior, including modification of subsequent behavior in response to environmental feedback. To this end, we report the design of novel stochastic chemical reaction networks capable of probabilistic decision-making in response to stimuli. We show that a simple chemical reaction network motif can be tuned to produce arbitrary decision probabilities when choosing between two or more responses to a stimulus signal. We further show that simple feedback mechanisms from the environment can modify these probabilities over time, enabling the system to adapt its behavior dynamically in response to positive or negative reinforcement based on its decisions. This system thus acts as a form of operant conditioning of the chemical circuit, in the sense that feedback provided based on decisions taken by the circuit form the basis of the learning process. Our work thus shows that simple chemical systems can be used to implement lifelike behavior in engineered biomimetic systems.

Computational analysis of conformational changes of nucleic acid nanostructures induced by overhangs.

Matthew Sample[1], Michael Matthies[1], Lu Yu[1], Petr Šulc[1]

[1] School of Molecular Sciences and Center for Molecular Design and Biomimetics, The Biodesign Institute, Arizona State University, 1001 South McAllister Avenue, Tempe, AZ 85281, USA

A widely accepted strategy to functionalize nucleic acid nanostructures is through the usage of single or double-stranded (ss/ds) overhangs. For this, oligonucleotides at specific positions get extended by a sequence which is accessible to other interacting elements modified with the complementary sequence. Commonly used design tools don't provide an easy way to access the implications these extensions have on the conformations of the nanostructure. Here we present a detailed study on the effects the overhangs have on a rectangular single and double layered nanostructure using the popular molecular dynamics simulation framework oxDNA[1,2,3,4,5]. We show that overhangs can deform an underlying structure significantly (Fig. 1) and find that the change can be rationalized through the change in conformational entropy. Since many functionalized designs assume flat structures and don't account for these effects, our computational results can be used to adopt existing designs towards required applications.

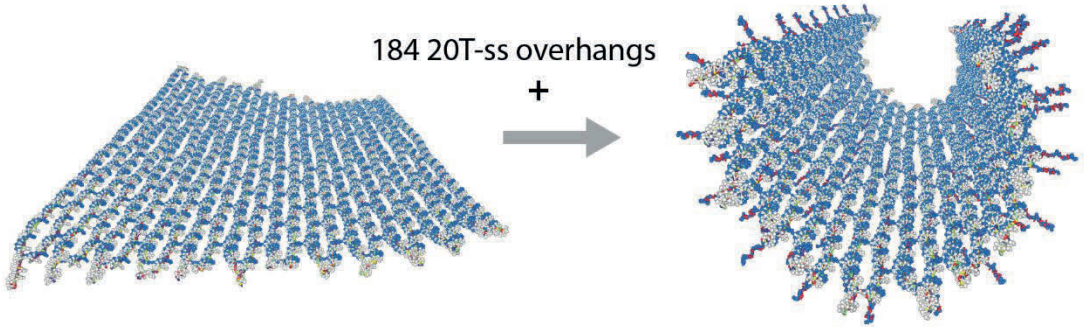


Figure 1. Effects of the addition of 184 ss-overhangs of length 20 to a flat sheet origami. Both structures represent mean structures of an equilibrium sampled trajectory of the corresponding design.

[1] Ouldridge T E et al 2011 *J. Chem. Phys*. **134**, 085101.
[2] Šulc P et al 2012 *J. Chem. Phys*. **137**, 135101.
[3] Snodin B E K et al 2015 *J. Chem. Phys*. **142**, 234901
[4] Rovigatti L et al 2015 *J. Comput. Chem* **36**, **1**
[5] Poppleton E. et al 2020 *NAR* **48, 72**

Unsupervised and Interpretable Machine Learning for Aptamer Analysis

Jonah Procyk[1] and Petr Šulc[1]

[1]Arizona State University, USA

Selection protocols such as SELEX, where molecules are selected over multiple rounds for their ability to bind to a target molecule of interest, are popular methods for obtaining binders for diagnostic and therapeutic purposes. With the increasing amount of such high-throughput experimental data available, machine learning techniques have become increasingly popular for molecular datasets analysis. We apply a two-layer neural network architecture, a Restricted Boltzmann Machine, to generalize the sequence data. Due to the unique architecture of our model[1], we obtain a latent representation that is directly interpretable as well as generative. We show our model can be trained on sequence ensembles from SELEX experiments for DNA aptamers and used to estimate the fitness of the sequences obtained through the experimental protocol. To confirm our findings', generated sequences were verified experimentally.

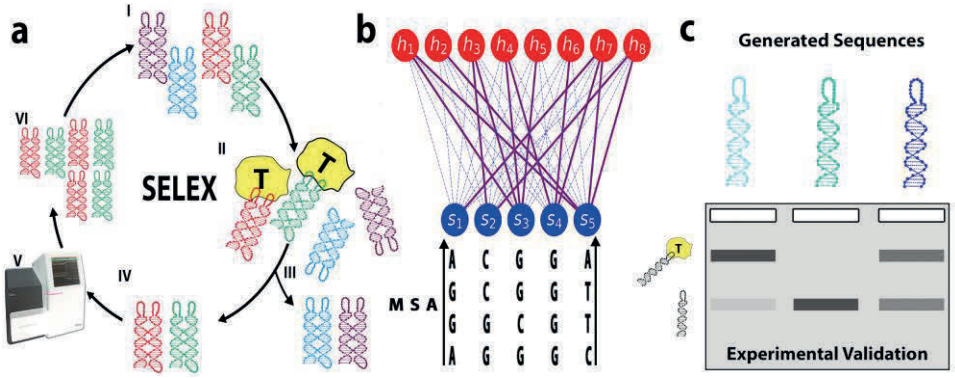


Figure 1: Schematic view of the SELEX experiment and the RBM-based analysis.
a: The SELEX procedure used to obtain DNA aptamers that bind to thrombin consists of the following steps: I) We start with an initial library of DNA sequences. II) DNA aptamers compete with each other to bind to thrombin. III) Sequences that are unbound (or bound too weakly) are washed away. IV) Remaining bound sequences dissociate after the sample is heated up. V) Binding sequences are sequenced. VI) Using polymerase chain reaction (PCR), multiple copies are made of the remaining sequences, resulting into a new library of aptamers for the next round of selection. b: The sequenced aptamers from respective rounds of the SELEX protocol are used to train the parameters of the Restricted Boltzmann Machine model. In this unsupervised neural network architecture, a layer of visible units carry the aptamer sequence, while the layer of hidden units extract representations. The weighted connections between the two layers are learned through maximization of the log-likelihood of the sequences obtained through SELEX. c: Single loop sequences generated using the Restricted Boltzmann Machine model are experimentally validated using gel assays.

[1] Tubiana, J., Cocco, S., & Monasson, R. (2019). Learning protein constitutive motifs from sequence data. *ELife*, *8*.

Molecular-responsive DNA nanopore for biomolecular measurements

Hiromu Akai, and Kan Shoji

Nagaoka University of Technology, Japan

In this study, we developed a DNA nanopore that can repeatedly open and close in response to ATP. The presence of ATP (30 μM) in electrolyte solution was detected by electrically observing the open-close behavior of the DNA nanopore. In the future, the concentration of ATP would be measured by analyzing the open-close frequency of the DNA nanopore.

1. Introduction - Nanopore sensors using DNA-based nanopores potentially offer high-sensitive single molecular analysis because of the high designability and robustness of DNA nanopores [1]. However, it is challenging to add molecular selectivity. Here, we propose a molecular-responsive DNA nanopore that can repeatedly open and close in response to ATP. We demonstrate that the molecular-responsive DNA nanopore sensor can detect ATP (Fig. 1a).

2. Methods – The molecular-responsive DNA nanopore which consists of a tubular DNA nanostructure and an ATP-binding DNA aptamer was designed. Channel current measurements of DNA nanopores with and without ATP were conducted using gold-needle-supported lipid bilayers [2]. DNA nanopores were immobilized on the gold needle, and the lipid bilayer was formed by inserting the needle into a layered bath solution of an electrolyte and an oil/lipid mixture. The DNA nanopores were spontaneously inserted into the bilayer (Fig. 1b).

3. Results and discussion – Different current signals were obtained with and without ATP. Without ATP, the fluctuated current signals were observed, indicating that aptamers fluctuated due to competition between the hydrogen binding force of aptamers and the electrophoretic force from a gold needle electrode (Fig. 1c). On the other hand, open-close-like signals were observed in the presence of ATP (Fig. 1d). This result suggests that the binding force between the aptamer and ATP becomes dominant. In conclusion, the ability to detect ATP with nanopore sensors using the molecular-responsive DNA nanopore was demonstrated. In the future, we will investigate the ability to measure ATP concentration with the nanopore sensor, and experimental conditions such as electrolyte concentration and applied voltage for high-sensitive ATP detection.

4. References

[1] Diederichs T et al 2019 *Nat. Commun.* **10** 5018
[2] Shoji K, Kawano R and White R J 2019 *ACS Nano* **13** 2

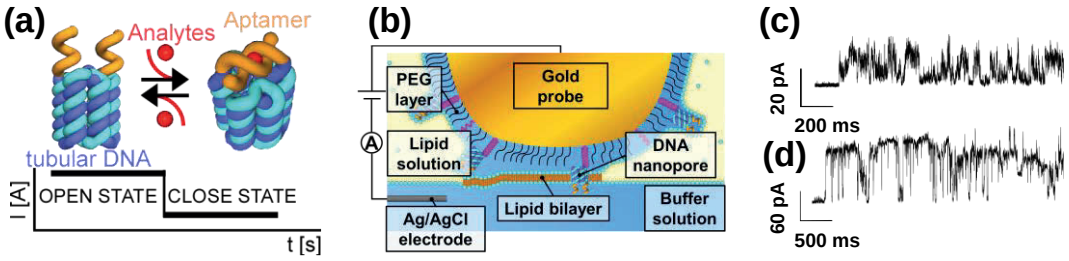


Fig. 1 (a) An open-close principle of a molecular-responsive DNA nanopore. (b) Lipid bilayer formation using a DNA probe system. (c, d) Recorded channel currents of DNA nanopores (c) without and (d) with ATP (30 μM).

Dendric DNA Origami for Efficient DDS Carrier

Akinori Kuzuya
Department of Chemistry and Materials Engineering, Kansai University
3-3-35 Yamate, Suita, Osaka 564-8680, Japan
E-mail: kuzuya@kasnsai-u.ac.jp

In this study, DNA origami dendrimer has been designed and successful formation of the structure was confirmed using atomic force microscope (AFM), transmission electron microscopy (TEM), dynamic light scattering method (DLS), etc. Decoration of the structure with streptavidin proteins was also examined by modifying appropriate staple strands. Doxorubicin-loaded DNA origami dendrimer was efficiently up-taken by cells.

Today, DNA materials are attracting broad interests for their versatile utilities in medical applications. Particularly, DNA origami nanostructures are extensively studied recently. We also have constructed various DNA origami nanostructures including nanomechanical devices. In this study, we aimed to design a new DNA origami for drug delivery systems (DDS). Among various possible nanocarriers used in DDS such as micelles, liposomes, and nanogels, we chose dendrimers to represent with DNA origami. We have designed DNA origami dendrimers with as much as 108branches in fourth generation, and confirmed successful formation of desired structure using AFM, DLS, TEM, and agarose gel electrophoresis, etc. We also investigated biotin-mediated coverage of the structure with streptavidin, by modifying the biotinylated DNA on the periphery. The resistance to degradation in FBS was compared before and after coverage, and cellular-uptake by HeLa was investigate.

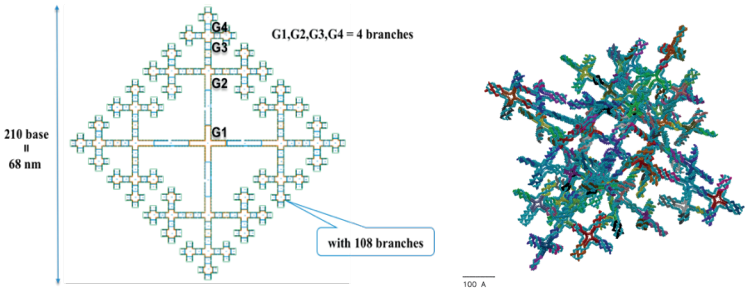


Fig.1. Design of DNA origami dendrimer, and typical simulated 3D structure.

The designed structure is shown in Fig.1. Four generations of four-way junctions are formed in the structures giving 108 branches in total. Each branch has 105 base pairs from the center to the end, and thus the expected diameter becomes 68 nm. Fig.2 shows an AFM image of prepared DNA origami dendrimer on mica. A few objects with expected diameter were observed. TEM and DLS also gave similar particle sizes. Agarose gel electrophoresis showed clear band shift before and after streptavidin coverage. The observed objects in AFM measurements of streptavidin-DNA origami dendrimer complexes were 5-nm higher than the unmodified dendrimers, which is almost the expected diameter of streptavidin tetramers. When HeLa cells were incubated in the presence of fluorescein-modified DNA origami dendrimer, significant green fluorescence was observed in cytosol. Similarly, addition of doxorubicin-load DNA origami dendrimer resulted in efficient doxorubicin delivery into the nucleus.

In summary, we have successfully constructed DNA origami dendrimer. This structure can be prepared in one step, and any molecule or substance can be easily modified on the surface of the structure in abundance, so that application to DDS materials is expected.

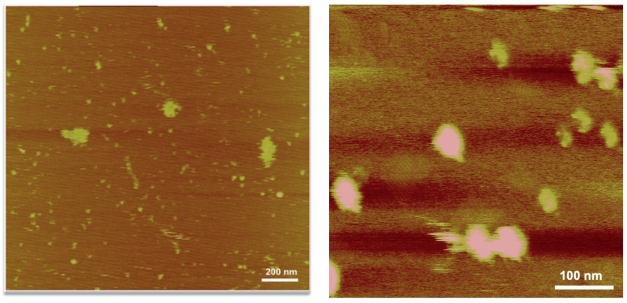


Fig.2. Typical AFM images before and after streptavidin decoration.

Multi-gate Boolean processors in a cell free transcriptional platform

Judee Sharon, Kate Adamala, Chelsea Dasrath, Aiden Fujiwara, Alessandro Synder, Mace Blank and Sam O'Brien

While traditional electronic computing tools become smaller and cheaper, there is a physical limit to the applications and flexibility of silica-based platform. There is a need to overcome these limitations in biological interfacing technologies (including implants and biodegradable sensors), as well as a need to develop solutions with capacities characteristic of natural systems (like regeneration and autonomous processing). The solution to some of these problems might lie in biologically-inspired computing platforms. However, most of the biocomputing technologies developed to date have their own unique challenges. Live cell-based biocomputing platforms are characterized by high noise and low reproducibility, while non-living biochemical platforms have low reliability and lack signal amplification1. Here we present a platform that attempts to reconcile some of the limitations present in other biocomputing solutions.

We have designed Boolean logic gates that perform operations in a cell free environment using enzymes and an RNA aptamer transcriptional output. The core architecture of the NAND, NOT, and NOR logic gates include a single-strand DNA gate template encoding a T7 RNA polymerase promoter sequence, a restriction enzyme cut site, and an RNA aptamer sequence. The inputs are smaller complementary strands of DNA that hybridize with the restriction enzyme cut site region. The corresponding restriction enzyme facilitates a transformation of the gate template and the cell free transcriptional platform outputs an RNA aptamer fluorescent signal depending on the binary result (0 = low/auto fluorescence and 1 = high fluorescence). The architecture of an AND and OR gate depend on DNA polymerases and a similar fluorescent RNA aptamer output. The inversely truncated sense and antisense strands of these gate templates are completed using DNA polymerization with cell free transcription following the reaction to display the binary result of the AND or OR operation. To aid in the design of each gate, we created a web platform – trumpet.bio – that will allow a user to select a gate and what is encoded in each gate template. Each Trumpet-generated gate template is analyzed with Nupack2 RNA prediction software to ensure that the transcribed RNA aptamer will fold accurately into its secondary structure when acting as an output to a gate.

Of these basic logic gates, we chose to create a multi-gate processor using the NAND gate because of its universality in creating all the other basic gates. Singular NAND gates of our design perform orthogonal operations when multiplexed. A unique gate template can be mixed with one matching input pair and many mismatching input pairs and still result in a successful logical operation. NAND gates can also be designed to output small single-strand DNA that can act as inputs to downstream NAND gates. Using two unique NAND gates that output one single-strand DNA each, a third subsequent NAND gate is transformed into the final output of a fluorescent RNA aptamer. This 3-gate processor is an example of an OR gate made up entirely of NAND gates. These cell free Boolean logic gates are the foundation for creating a simple, yet reliable biocomputing platform for future biological interfacing technologies.

Developmental self-assembly of a DNA ring with stimulus-responsive size and growth direction

Allison T. Glynn[1†], Samuel R. Davidson[1†] and Lulu Qian[1,2⋆]
[1]*Bioengineering and* [2]*Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA
[†]*Equal contribution,* [⋆]*e-mail: luluqian@caltech.edu*

Abstract. Developmental self-assembly of DNA nanostructures provides an ideal platform for studying the power and programmability of kinetically controlled structural growth in engineered molecular systems. Triggered initiation and designated sequencing of assembly and disassembly steps have been demonstrated in structures with branches and loops. Here we introduce a new strategy for selectively activating distinct subroutines in a developmental self-assembly program, allowing structures with distinct properties to be created in response to varying molecular signals. We demonstrate this strategy in triggered self-assembly of a DNA ring, the size and growth direction of which are responsive to a key molecule. We articulate that reversible assembly steps with slow kinetics at appropriate locations in a reaction pathway could enable multiple populations of structures with stimulus-responsive properties to be simultaneously created in one developmental program. These results open up a broad design space for molecular self-assembly with adaptive behaviors, toward advanced control in synthetic materials and molecular motors.

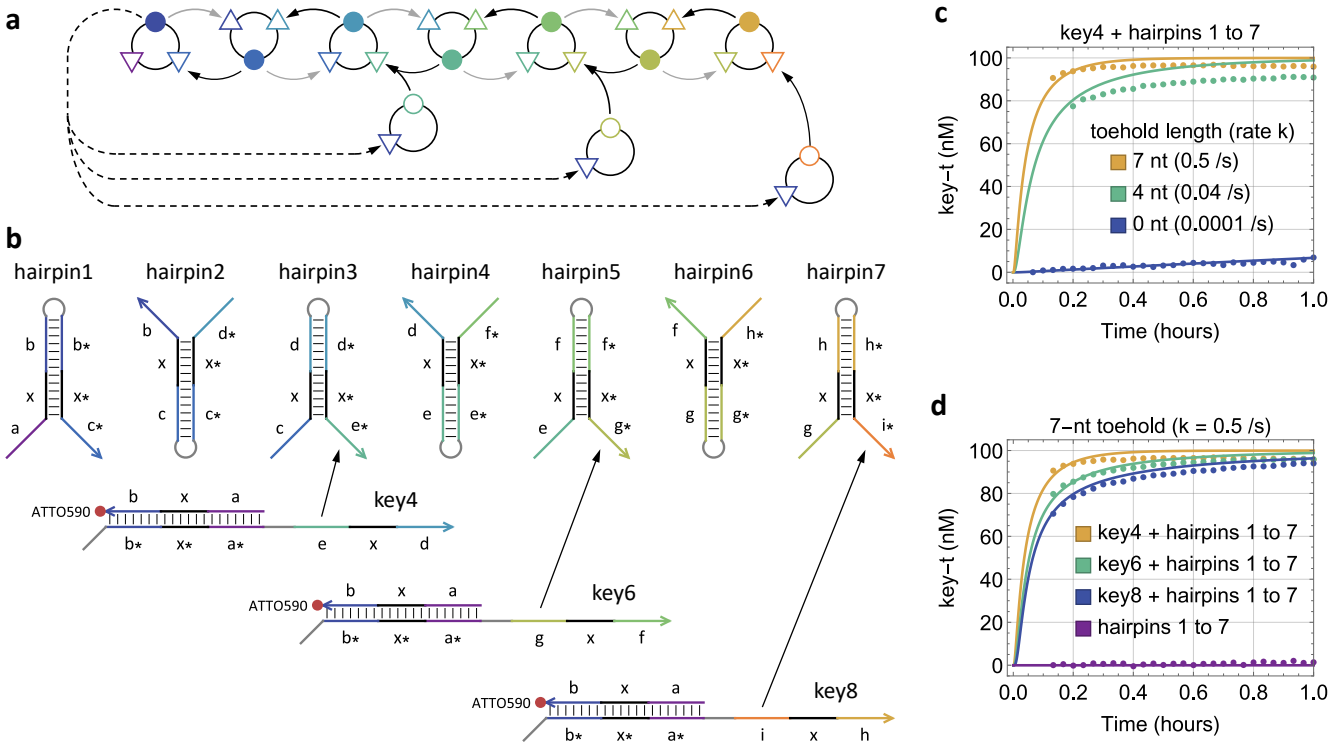


Figure 1: **Developmental self-assembly of a DNA ring with stimulus-responsive size.** **(a)** Abstract reaction graph. **(b)** Domain-level strand diagrams. Simulation and fluorescence kinetics data of **(c)** 4-stranded ring formation with varying toehold sizes, and **(d)** 4, 6, and 8-stranded ring formation with varying keys.

Throttling enables rate-insensitive chemical reaction networks.

Dominic Scalise[1], and Lulu Qian[1]

[1]California Institute of Technology, USA

DNA-based chemical reaction networks (CRNs) exhibit up to an order of magnitude uncertainty in their reaction rate constants [1,2]. This can impair the intended behavior of DNA circuits and can require significant tuning of experiments [2].

Rate independent CRNs [3] are designed to perform the same behavior regardless of their reaction rate constants. However, the class of functions computable with rate independent CRNs remains severely limited [3]. In contrast to strict rate independence, rate or concentration *insensitive* CRNs only require that rate constants and concentrations are specified to within an order of magnitude [4,5]. For example, DNA seesaw circuits [5] require that thresholding reactions occur with an order of magnitude higher rate constant than amplification reactions, but can accommodate arbitrarily higher thresholding rate constants above this design specification.

We outline a method for designing rate-insensitive mass action CRNs for the class of networks shown in reference [6], where an output species is continuously produced and degraded. In our construction, all rate constants can independently vary by an order of magnitude without altering the CRN behavior. Our technique (Fig. 1) specifies a pair of reference reactions that produce and degrade [6,7] a reference species R at rate constants r_P and r_D, defined to be unitary values. All other reactions are normalized to the reference through catalytic “throttle” species that govern the rate of each reaction. Increasing throttle concentrations speeds up the effective rate constants for reactions that are slower than the reference, while decreasing throttle concentrations slows down reactions that are faster than the reference. This technique suggests ways to design DNA CRNs that are robust to experimental uncertainty in rate constants.

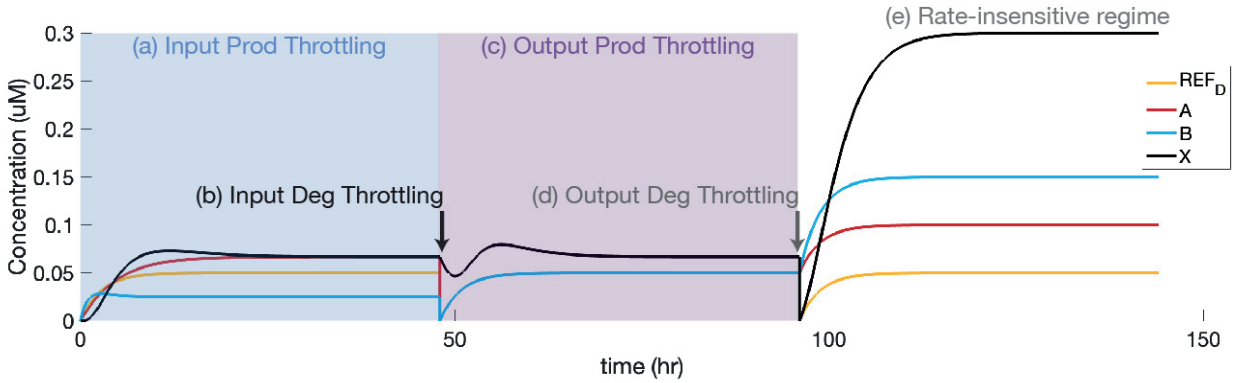


Figure 1. Throttling a multiplication CRN [6] with uncertain rate constants. Reference (yellow) is held at an arbitrary unit concentration. Autonomous steps measure the offsets of other reactions from the reference, and throttle (a,b) the input layer and (c,d) output layer of the CRN. (e) After throttling, the output of the multiplier correctly reads 2x3=6, with a unit concentration of 50 nM.

[1] Zhang D and Winfree E 2009 *J. Am. Chem. Soc.* **131**.47
[2] Thubagere A and Qian L 2017 *Nat. Comm.* **8**.1
[3] Chen H, Doty D and Soloveichik D 2014 *Proc.5th conf. Innov. Theor. Comp. Sci.*
[4] Shinar G and Feinberg M 2010 *Science* **327**.5971
[5] Qian L and Winfree E 2011 *Science* **332**.6034
[6] Buisman H *et al* 2009 *Artificial Life* **15**.1
[7] Scalise D, Dutta N and Schulman R 2018 *J. Am. Chem. Soc.* **140**.38

Simple software to design recipes for complicated sample mixes*

Constantine Evans† David Doty‡ Damien Woods†

Scaling DNA computing systems to large numbers of components presents the practical problem of actually combining those components correctly in the lab. Systems could require thousands of individual strands to be mixed together, with accurate concentrations being important and a single missing component having the potential to completely alter behavior. Experiments could involve many samples with varying, overlapping subsets of components, created through multiple layers of intermediate mixes. The complexity of keeping track of components spread across many plates, at different concentrations, and layers of mixes incorporating them, poses a challenge for actually implementing complex designs in the lab.

As part of the Growth Dynamics project[1], we developed a Python library to systematically, efficiently, and safely design recipes for mixes (https://github.com/cgevans/mixes). Designed to be used in Jupyter notebooks, users specify mixes in Python code as a series of actions (for example, desired concentrations or volumes) on sets of components, which can be automatically drawn from references such as DNA synthesis order information. Mixes themselves can be used as components in other mixes, with the library recursively keeping track of individual components, and carrying out checks such as ensuring that volumes of intermediate mixes are sufficient. The library generates tables of mix instructions including components in plates arranged for efficient pipetting, along with maps of component locations on plates for each mix.

Specifying mixes with code allows them to be defined programmatically, significantly simplifying the construction of large systems: this could be as simple as a for-loop selecting individual strands to include, or as complex as code integrating with an online spreadsheet interface allowing the user to select locations to include on a molecular canvas of DNA tiles, as used in the Growth Dynamics project. By recursively keeping track of individual components and their final concentrations, series of mix specifications can also be checked for correctness: for example, for DNA tile systems specified within the Alhambra tile system designer[2], kinetic Tile Assembly Model simulations can be run based on the components and concentrations in a mix.

*Supported by ERC grant 772766, Active-DNA, SFI grants 18/ERCS/5746 and 20/FFP-P/8843, and NSF awards 1900931 and 1844976.

†Hamilton Institute, Department of Computer Science, Maynooth University, Ireland

‡Department of Computer Science, University of California, Davis, USA

[1]C. Evans, D. Doty, and D. Woods, DNA 28, 2022

[2]C. Evans and E. Winfree, 2018, and https://github.com/DNA-and-Natural-Algorithms-Group/alhambra

Stochastic chemical reaction networks for graph coloring

Philippa Richter[1], Salvador Buse[2] and Erik Winfree[1,2⋆]
[1] *Computer Science, and* [2] *Bioengineering*
California Institute of Technology, Pasadena, CA 91125, USA
⋆ *e-mail: winfree@caltech.edu*

Abstract. Stochastic chemical reaction networks (CRNs) can simulate powerful deterministic models such as Boolean circuits and space-bounded Turing machines by carefully ensuring that the stochasticity does no damage to the progress of the computation [1]. But can the inherent stochasticity be exploited as a resource that makes the computation faster, more efficient, or more compact? Recently it was shown that the WalkSAT algorithm for the NP-complete formula satisfiability problem [2] has a natural analog in stochastic CRNs [3]. Here, we demonstrate an even more straightforward CRN for finding solutions to a second NP-complete problem, the graph k-coloring problem. For a graph $G = (V, E)$ with $n = |V|$ vertices and $m = |E|$ edges, the WalkColoringCRN contains species $v[i, c]$ whose presence or absence indicates whether vertex i should be colored c, species $v[i]$ to indicate vertex i is uncolored, and $tryv[i, c]$ to indicate an attempt to recolor the vertex that may be rejected if it introduces a conflict. The reactions are:

$$v[i] \xrightarrow{k_1} tryv[i, c] \qquad \forall i \in V, c \in \{1, \ldots, k\}$$
$$v[i, c] + v[j, c] \xrightarrow{k_2} v[i] + v[j, c] \qquad \forall (i, j) \in E, c \in \{1, \ldots, k\}$$
$$tryv[i, c] + v[j, c] \xrightarrow{k_3} v[i] + v[j, c] \qquad \forall (i, j) \in E, c \in \{1, \ldots, k\}$$

Unlike the previous WalkSATCRN, which required more complex trimolecular reactions, WalkColoringCRN requires no more than binary reactions, and yet demonstrates comparable efficiency for coloring graphs with 3 or 4 colors. Three other CRN algorithms for graph coloring are also compared.

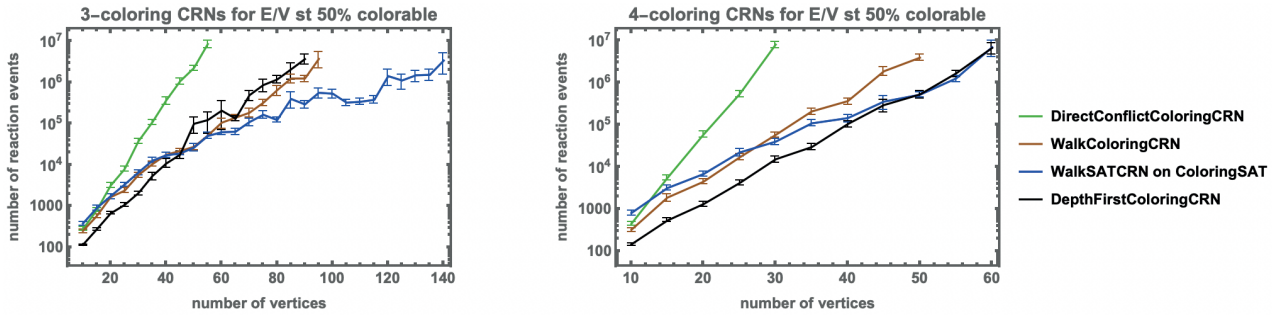


Figure 1: **CRN algorithm efficiency on random hard coloring problems.** The average number of reaction steps to find a solution, as a function of problem size, for four CRN algorithms. Typically-hard random graph-coloring problems are generated by choosing, for each graph size, the number of vertices such that about half the generated graphs are colorable (by MiniSAT). Each CRN algorithm is challenged on the first 100 colorable graphs thus generated. **(Left)** Three-coloring problems up to graphs of 115 vertices. **(Right)** Four-coloring problems up to graphs of 60 vertices.

[1] Neil D Jones, Lawrence H Landweber, and Y Edmund Lien. Complexity of some problems in Petri nets. *Theoretical Computer Science*, 4:277–299, 1977.

[2] Bart Selman, Henry A Kautz, and Bram Cohen. Noise strategies for improving local search. In *Proceedings of AAAI-94*, pages 337–343, 1994.

[3] Erik Winfree. Chemical reaction networks and stochastic local search. In *DNA Computing and Molecular Programming (Lecture Notes in Computer Science)*, volume 11648, pages 1–20. Springer, 2019.

Tile-displacement-based shape reconfiguration in DNA origami tile assemblies

Namita Sarraf[1], Kellen R. Rodriguez[2,3,4], and Lulu Qian[1,4]
[1]*Bioengineering,* [2]*Business Economics & Management,* [3]*Astrophysics, and* [4]*Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA

DNA origami tile arrays have been envisioned to provide spatial organization for complex molecular information-processing circuits and swarm molecular robots. Toward adaptive behaviors in these complex systems, we have developed a tile displacement mechanism that allows information-based autonomous reconfiguration in DNA origami tile arrays. Experimental demonstration of tile displacement systems has so far been limited to simple logic and swap behaviors in arrays of up to 3 by 3 in size, where the initial and final structures were of the same shape and the invader structures consisted of one or two tiles. Could tile displacement allow for more complex reconfigurations where not only the patterns on individual tiles but also the shapes of tile assemblies are responsive to resources available in a molecular environment?

Here we explore a new class of tile displacement systems where an initial DNA origami tile assembly is reconfigured into a different shape facilitated by multi-tile invaders. Using an example of a sword reconfiguring into a snake, we show how to control the length distribution of DNA origami polymers by tuning the ratio of distinct types of monomers responsible for growing and capping, and to grow 3D tubes with origami tiles and open them up into 2D structures via tile displacement. We develop design principles for improving the yield of multi-tile invaders by increasing the gap between the melting temperatures of origami monomers and their interactions between external and internal edges of an assembly. Furthermore, we demonstrate the robustness of tile displacement with varying toehold and branch migration domain configurations, and identify a type of edge design that results in significantly slower kinetics.

This work extends the demonstration of tile displacement from simple swap behaviors in fixed arrays to shape reconfiguration of tile assemblies with more complex invaders. Unlike tile assembly and disassembly, tile displacement allows for sequestered domains on tile edges to be activated only upon the arrival of invader tiles, which is key for cascading of reactions and complex network behaviors, and for the kinetics of reactions to be controlled by the strength of the toehold domains. In theory, tile displacement is Turing universal and can simulate arbitrary two-dimensional synchronous block cellular automata, where each transition rule for updating the state of a 2 by 2 neighborhood is implemented by just a single tile. We hope that the ability of shape reconfiguration now opens up new theoretical questions on the computational power of tile displacement.

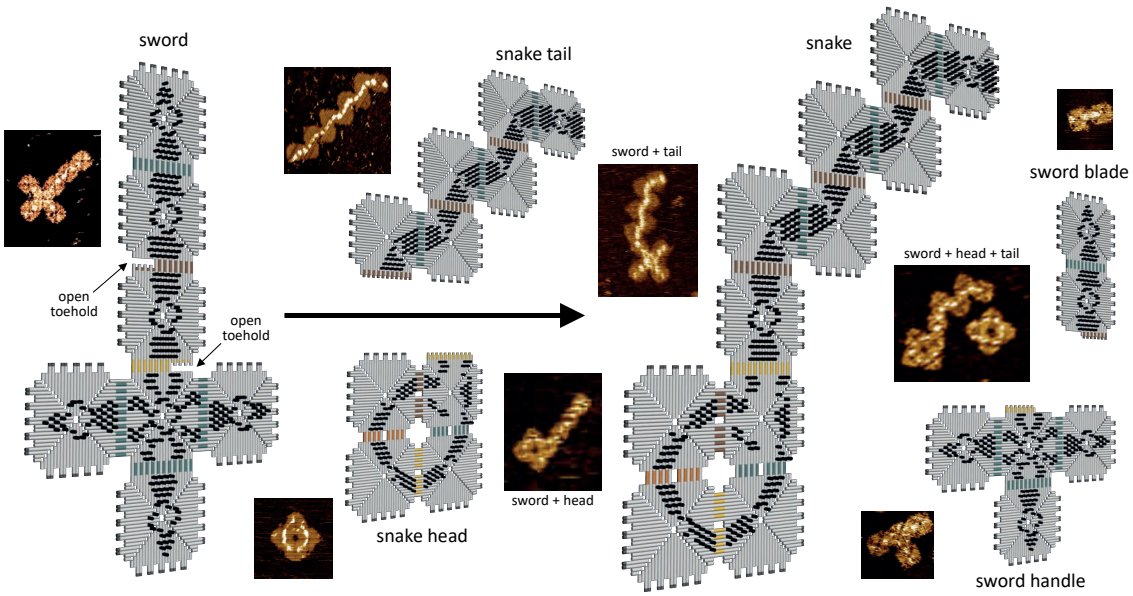


Figure 1: **A seven-tile sword shape reconfiguring into a snake shape with a variable tail length.**

Stochastic surface chemistry can solve hard problems!

Mohini Misra[1], Jean-Sebastien Paul[2], Salvador Buse[3], and Erik Winfree[1,3⋆]
[1]*Computer Science,* [2]*Information Science & Technology, and* [3]*Bioengineering*
California Institute of Technology, Pasadena, CA 91125, USA
⋆*e-mail: winfree@caltech.edu*

Abstract. Chemistry is stochastic at the smallest size scales. Rather than being an impediment to computation, stochastic Chemical Reaction Networks (CRNs) can exploit their inherent randomness to carry out natural and efficient stochastic local search algorithms for solving hard constraint satisfaction problems such as formula satisfiability (3-SAT) and graph vertex coloring (3-coloring) [1, 2]. While these constructions are both remarkable and insightful, there is a limitation due to the well-mixed nature of the CRNs: a larger CRN, with more species and more reactions, is needed to color a larger graph or solve a larger formula. Might it be possible for a single CRN to be uniformly capable of solving arbitrarily-sized problems using stochastic local search? For this, we turn to a model of spatially-organized chemical reactions, the surface CRN model [3]. We present a surface CRN with 7 species and 24 reactions that can solve any solvable 3-coloring problems, and a surface CRN with 11 species and 31 reactions that can solve any satisfiable 3-SAT formula. The systems are designed such that each reaction recognizes a constraint satisfaction violation and attempts to fix it; they can halt only when a correct solution is reached, and they provably will eventually reach a solution if one exists. Unfortunately, these constructions do not find solutions quickly. Using ideas for effectively biasing random search [4, 5], we construct only somewhat larger surface CRNs that can effectively solve random planar graphs with over 100 vertices and random 3-SAT formulas with over 100 variables. The principles used in these constructions (global constraints from local constraints, rejection of problematic random search steps, directional propagation of signals over distances) may be of relevance more generally for other types of stochastic molecular algorithms that involve spatial organization, such as pattern formation and self-healing.

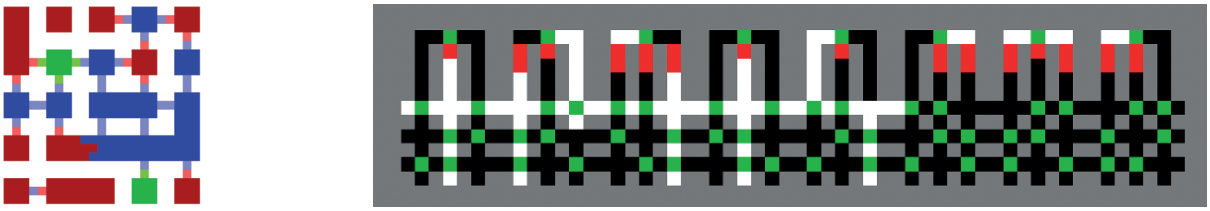


Figure 1: **Snapshots of surface CRNs performing graph coloring and formula satisfaction. (Left)** Finding a 3-coloring solution for a randomly-generated planar graph. Vertices are represented by the thick shapes, colored darkly. Edges are represented by the 2-cell-long thin connectors joining these, colored lightly. In this snapshot, the surface is still in the process of ensuring no two connected vertices have the same color, hence some vertex regions remain inconsistently colored and some edge connectors do not join distinctly-colored vertices. **(Right)** Finding a solution to a 3-variable, 8-clause 3-SAT problem. Horizontal lines correspond to variables; vertical lines correspond to literals in clauses. Intersections between the two usually cross over each other (green), but some do not, signifying a variable being connected to a clause. Black represents false and white represents true; red NOT gates near the clauses negate literals. The CRN is in the process of making sure each clause has at least one true incoming literal.

[1] Erik Winfree. Chemical reaction networks and stochastic local search. In *DNA Computing and Molecular Programming (Lecture Notes in Computer Science)*, volume 11648, pages 1–20. Springer, 2019.

[2] Philippa Richter, Salvador Buse, and Erik Winfree. Constraint satisfaction and homeostasis in stochastic chemical reaction networks: a case study on graph coloring. In preparation.

[3] Samuel Clamons, Lulu Qian, and Erik Winfree. Programming and simulating chemical reaction networks on a surface. *Journal of the Royal Society Interface*, 17:20190790, 2020.

[4] AD Petford and DJA Welsh. A randomised 3-colouring algorithm. *Discrete Mathematics*, 74:253–261, 1989.

[5] Bart Selman, Henry A Kautz, and Bram Cohen. Noise strategies for improving local search. In *Proceedings of AAAI-94*, pages 337–343, 1994.

Neural Computation in Boltzmann Liquids and Hopfield Droplets

Salvador Buse[3], Arvind Murugan[4], and Erik Winfree[1,2,3⋆],
[1]*Computer Science,* [2]*Computation & Neural Systems, and* [3]*Bioengineering*
California Institute of Technology, Pasadena, CA 91125, USA
[4]*Department of Physics*
University of Chicago, Chicago, IL 60637, USA
⋆*e-mail: winfree@caltech.edu*

Abstract. The principles of neural computation may provide valuable guidance for understanding high-dimensional molecular systems that arise in biology and in molecular engineering. DNA nanotechnology has already given us programmable strand displacement cascades that compute like neural networks [1, 2] and programmable self-assembly whose nucleation kinetics computes like neural networks [3] — physically, these correspond respectively to systems in gas phase (e.g., a dilute solution where particles of interest only rarely interact) and systems in solid phase (e.g., arranged in a crystalline lattice where each molecule is in constant contact with the same neighbors over its lifetime). Do the principles of neural computation also extend to the liquid phase, where each molecule is always in contact with other molecules but always different molecules as positions change? Programmable DNA liquids provide a potential concrete implementation [4], if theoretical connections to neural computation can be identified. Here we present three such theoretical connections. First, we consider an equilibrium thermodynamic view wherein a multicomponent liquid on an Ising-like lattice encodes a probability distribution, can perform inference, and can learn. This "Boltzmann Liquid" model expands on the gas-phase "Chemical Boltzmann Machine" model [5] by introducing an interaction energy matrix in addition to the species energy vector that characterizes equilibrium in chemical reaction networks. The mechanisms for performing inference by clamping and for learning the energy matrix via Hebbian/anti-Hebbian wake-sleep cycles generalize naturally from classical Boltzmann machines [6]. Second, we consider the energy landscape for multicomponent liquid droplets that are internally well-mixed, now illustrating a connection to Hopfield network associative memories. This "Hopfield Droplet" model can be examined from a thermodynamic perspective: basins in the energy landscape dictate possible stable ratios of the species for droplets that mature by Ostwald ripening, yielding the recall of a pattern vector corresponding to the closest memory. Finally, we consider the kinetics of Hopfield droplet condensation when it is triggered in response to temperature or concentration changes. The interaction matrix of the multicomponent liquid in this non-equilibrium context controls the nucleation barriers for forming different memory droplets and can perform sophisticated pattern recognition – here the concepts are adapted from the study of multifarious self-assembly of crystalline structures [7, 8, 3]. These fundamental investigations into neural computation within multicomponent programmable liquids will augment our toolbox for engineering cell-scale molecular robotic systems and may even shed light on why biomolecular liquid condensates are so pervasive within cell biology [9].

[1] Lulu Qian, Erik Winfree, and Jehoshua Bruck. Neural network computation with DNA strand displacement cascades. *Nature*, 475:368–372, 2011.

[2] Kevin M Cherry and Lulu Qian. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature*, 559:370–376, 2018.

[3] Constantine Glen Evans, Jackson O'Brien, Erik Winfree, and Arvind Murugan. Pattern recognition in the nucleation kinetics of non-equilibrium self-assembly. In preparation.

[4] Byoung-jin Jeon, Dan T Nguyen, and Omar A Saleh. Sequence-controlled adhesion and microemulsification in a two-phase system of DNA liquid droplets. *The Journal of Physical Chemistry B*, 124:8888–8895, 2020.

[5] William Poole, Thomas Ouldridge, Manoj Gopalkrishnan, and Erik Winfree. Detailed balanced chemical reaction networks as generalized Boltzmann machines. *arXiv preprint arXiv:2205.06313*, 2022.

[6] Geoffrey E Hinton, Terrence J Sejnowski, and David H Ackley. Boltzmann machines: Constraint satisfaction networks that learn. Technical report, Carnegie-Mellon University, Department of Computer Science Pittsburgh, PA, 1984.

[7] Arvind Murugan, Zorana Zeravcic, Michael P Brenner, and Stanislas Leibler. Multifarious assembly mixtures: Systems allowing retrieval of diverse stored structures. *Proceedings of the National Academy of Sciences*, 112:54–59, 2015.

[8] Weishun Zhong, David J Schwab, and Arvind Murugan. Associative pattern recognition through macro-molecular self-assembly. *Journal of Statistical Physics*, 167:806–826, 2017.

[9] Andrew S Lyon, William B Peeples, and Michael K Rosen. A framework for understanding the functions of biomolecular condensates across scales. *Nature Reviews Molecular Cell Biology*, 22:215–235, 2021.

Molecular computation using DNA-based synthetic condensates

Sungho Do[1], Chanseok Lee[3], Taehyun Lee[1], Do-Nyun Kim[1, 3, 4] and Yongdae Shin[1, 2, 3*]

[1]Department of Mechanical Engineering, Seoul National University, Republic of Korea,
[2]Interdisciplinary Program in Bioengineering, Seoul National University, Republic of Korea,
[3]Institute of Advanced machines and Design, Seoul National University, Republic of Korea,
[4]Institute of Engineering Research, Seoul National University, Republic of Korea

E-mail address: ydshin@snu.ac.kr

Biomolecular condensates have recently been shown to play key roles in diverse cellular activities including gene regulation, cell signalling and stress survivals [1]. Inspired by the organization of intracellular condensates, we design synthetic condensates using DNA that exhibit programmable phase behaviors as well as functions. The material properties of DNA-based condensates vary depending on environmental factors, ranging from liquid-like droplets to solid-like gels. We demonstrate selective partitioning of certain clients into cognate condensates using different structural motifs for phase separation and composition control. We also show that the dramatic acceleration of DNA-based molecular computation can be achieved by selective partitioning of components into condensates. We anticipate that our platform will be highly useful in realizing more life-like artificial systems with high-order functionalities.

Reference list

[1] Yongdae Shin, Clifford P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).

Multi-fidelity parameter inference for an Arrhenius model of DNA elementary step kinetics

Jordan Lovrod[1], Boyan Beronov[1], Anne Condon[1], and Erik Winfree[2]

[1]University of British Columbia, Canada, [2]California Institute of Technology, US

Reaction pathways of interacting nucleic acid strands can be modeled using detailed-balance continuous-time Markov chains (CTMCs), whose discrete states and Markovian transitions correspond to, respectively, secondary structures and elementary base pair changes [1]. In addition to the states' free energies captured by a thermodynamic model, a kinetic model describes the non-equilibrium dynamics of a CTMC via elementary transition rates. In the Metropolis kinetic model, step-wise energy barriers consist only of free energy changes [1], whereas the Arrhenius model also includes activation energies which account for local interactions during an elementary step [2]. A common kinetic model can be calibrated over a range of DNA reactions by assuming an inverse relationship between the mean first passage time (MFPT) of CTMCs and experimentally observed reaction rates. For most reactions of interest, estimating the MFPT requires a state space truncation approximately preserving the dominant pathways [2, 3].

We build on an existing Bayesian approach to this inverse problem for the Metropolis and Arrhenius kinetic models [2]. In this context, Bayesian *posteriors* are often more practically useful than *maximum likelihood* or *maximum a posteriori* point estimates, because posterior distributions better capture the uncertainty over and the coupling between parameters, and because model averaging can yield more diverse predictions of reaction pathways. Posterior inference on manually designed truncated state spaces was previously performed using an affine invariant MCMC algorithm, which, while simple and generic, imposed strong limitations on the dataset size and the fidelity of each state space approximation. We therefore introduce a multi-fidelity MCMC method for joint inference in the Metropolis and Arrhenius models, using a specialised proposal kernel that leverages their structural relationship.

Intuitively, this allows the simpler Metropolis parameter space to constrain and accelerate inference in the more complex Arrhenius parameter space, and is expected to substantially improve statistical efficiency. Formally, we devise a multi-level MCMC algorithm over the joint parameter space, with proposal kernels that couple the rate matrices and the MFPTs of the two models. This is an improvement only to the inference algorithm, leaving unchanged the CTMC truncation and the current choice of Bayesian likelihood, i.e., synthetic Gaussian noise around the log-rate corresponding to the simulated MFPT. Nevertheless, the more general algorithmic structure motivates us to redesign the existing CTMC simulator – taking inspiration from the probabilistic programming literature – towards a domain-specific language for representing and computing with state spaces, kinetic models and probability spaces over them.

References

[1] J.M. Schaeffer. "Stochastic Simulation of the Kinetics of Multiple Interacting Nucleic Acid Strands". PhD thesis. California Institute of Technology, 2013.

[2] S. Zolaktaf et al. "Inferring Parameters for an Elementary Step Model of DNA Structure Kinetics with Locally Context-Dependent Arrhenius Rates". In: *DNA Computing and Molecular Programming, Lecture Notes in Computer Science*. 2017, pp. 172–187.

[3] S. Zolaktaf et al. *The pathway elaboration method for mean first passage time estimation in large continuous-time Markov chains with applications to nucleic acid kinetics*. 2021. URL: https://arxiv.org/abs/2101.03657.

Fridge Compiler: automated synthesis of molecular circuits

Lancelot Wathieu[1], Gus Smith[1], Luis Ceze[1], and Chris Thachuk[1]

[1]Paul G. Allen School of Computer Science & Engineering, University of Washington, USA

Rationally designed molecular circuits describable by well-mixed chemical reaction kinetics can realize arbitrary Boolean function computation yet differ significantly from their electronic counterparts: (i) NOT gates are often infeasible thus requiring the use of dual-rail encoding of inputs and (ii) spatial locality is absent necessitating that gates are distinguishable (*i.e.*, their design dictates possible interactions with other components). Given a Boolean function and a "fridge" inventory of distinguishable components, which subset should be mixed to realize the desired computation? This molecular circuit synthesis question is often solved manually and *ad hoc*. Existing synthesis tools intended for electronic circuits can improve upon manual determination, but can still result in non-optimal circuits or ones that cannot be built with the current inventory. This work overcomes these limitations by introducing the *Fridge Compiler* molecular circuit synthesis tool backed by the solver-aided programming language Rosette [1].

The Fridge Compiler was designed specifically for optimality, flexibility, and ease-of-use. Users can design optimal circuits by first selecting from a list of common DNA strand displacement (DSD) architectures and then specifying (i) a cost function (*e.g.*, component count), (ii) arbitrary constraints (*e.g.*, avoid mixing gates *K* and *L*), (iii) a fridge inventory, and (iv) the Boolean function to compute. Power users can define new molecular architectures (e.g., protein circuits), specify custom constraints and cost functions, and explore the space of valid circuits.

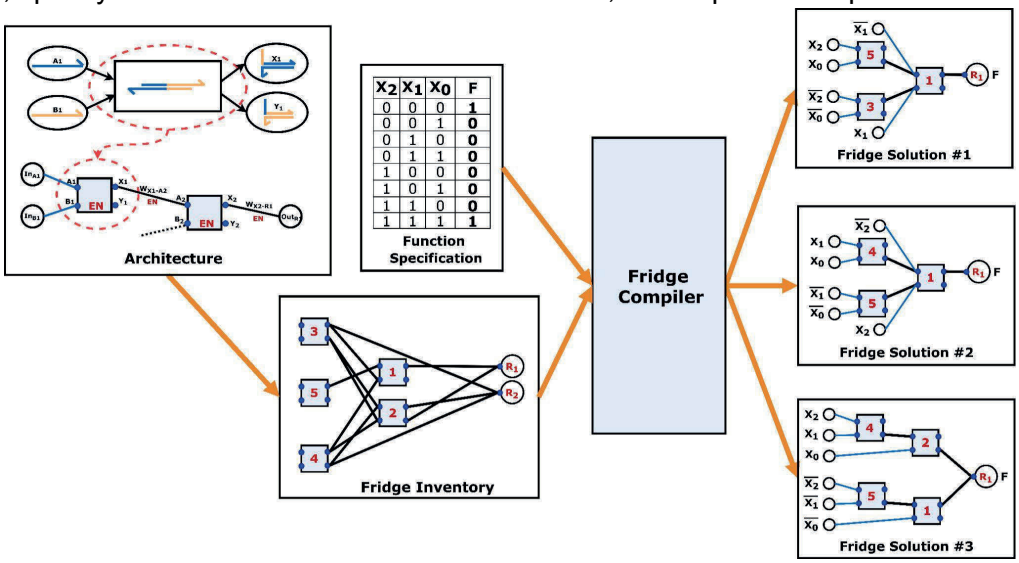


Fig. 1: Given a inventory and function specification the *Fridge Compiler* can synthesize all valid circuit solutions (#1, #2, #3, …), (size) optimal solutions (#1, #2), and optimal solutions that satisfy hard constraints (*#1; "don't mix gates 4 and 5, nor gates 1 and 2").*

The Fridge Compiler is particularly well-suited for molecular computing where the cost and effort of designing, preparing and purifying new circuit parts is often a significant barrier. Empirical results demonstrate its flexibility to synthesize arbitrary Boolean functions using three different DSD architectures. All 3-bit and 4-bit predicate functions were synthesized from a large custom inventory on the order of seconds in the worst case on a modern laptop. Importantly, and absent in previously known approaches specific to molecular circuits, the Fridge Compiler is logically sound, complete, and optimal for the user specified cost function.

[1] Torlak, E., & Bodik, R. (2013, October). Growing solver-aided languages with Rosette. In *Proceedings of the 2013 ACM international symposium on New ideas, new paradigms, and reflections on programming & software* (pp. 135-152).

Predicting accurate *ab initio* DNA electron densities with equivariant neural networks

Alex Lee,[1] Joshua Rackers,[2] and William Bricker[1]

[1]Department of Chemical and Biological Engineering, University of New Mexico, Albuquerque, NM 87131, USA

[2]Center for Computing Research, Sandia National Laboratories, Albuquerque, NM 87123, USA

One of the fundamental limitations of accurately modeling biomolecules like DNA is the inability to perform quantum chemistry calculations on large molecular structures. We present a machine learning model based on an equivariant Euclidean Neural Network framework to obtain accurate *ab initio* electron densities for arbitrary DNA structures that are much too large for conventional quantum methods. The model is trained on representative B-DNA base pair steps that capture both base pairing and base stacking interactions. The model produces accurate electron densities for arbitrary B-DNA structures with typical errors of less than 1%. Crucially, the error does not increase with system size, which suggests that the model can extrapolate to large DNA structures with negligible loss of accuracy. The model also generalizes reasonably to other DNA structural motifs such as the A- and Z-DNA forms, despite being trained on only B-DNA configurations. The model is used to calculate electron densities of several large-scale DNA structures, and we show that the computational scaling for this model is essentially linear. We also show that this machine learning electron density model can be used to calculate accurate electrostatic potentials for DNA. These electrostatic potentials produce more accurate results compared to classical force fields and do not show the usual deficiencies at short range.

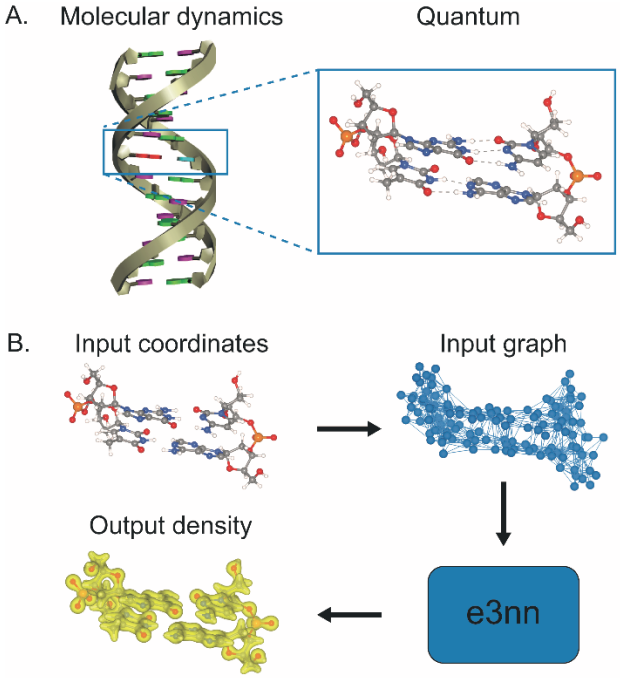


Figure 1 - Schematic for training a machine learning model for DNA. (A) Molecular dynamics simulations are run on B-DNA 12-mers. Quantum calculations are performed on representative snapshots extracted from the central two base pairs to obtain training densities. (B) The e3nn machine learning model takes atomic coordinates as input, converts them into a graph which is fed through a neural network, and outputs the electron density for the given input coordinates.

Poster Session 2 - Tuesday August 9th, 5:30-7:30pm

#14 Storing digital data in DNA-based tape (pg 72)
Afsaneh Sadremomtaz, Robert Glass, Jorge Guerrero, Micheal Brandon Reed, Eric Josephs and Reza Zadegan

#20 Nucleic Acid Memory: Super Resolution Microscopy enhances novel approach to DNA data storage (pg 73)
Luca Piantanida, George Dickinson, Golam Mortuza, William Clay, Christopher Green, Chad Watson, Eric Hayden, Timothy Andersen, Wan Kuang, Elton Graugnard, Reza Zadegan and William Hughes

#43 Nanopore decoding for DNA-computed over and under-expression of microRNA patterns (pg 74)
Sotaro Takiguchi and Ryuji Kawano (eligible for best student / postdoc poster award)

#44 Kinetic approach to suppress DNA hybridization error (pg 75)
Hiroyuki Aoyanagi, Simone Pigolotti, Shinji Ono and Shoichi Toyabe (eligible for best student / postdoc poster award)

#45 A geometric framework for reaction enumeration in computational nucleic acid devices (pg 76)
Sarika Kumar and Matthew R. Lakin (eligible for best student / postdoc poster award)

#46 Study of DNA nanostructures for delivering gene templates for homology directed repair (pg 77)
Diana Lopez, Wolfgang Pfeifer and Carlos Castro (eligible for best student / postdoc poster award)

#55 Scalable design and construction of DNA truss structures (pg 78)
Yudai Yamashita, Ibuki Kawamata and Satoshi Murata (eligible for best student / postdoc poster award)

#57 NanoFrame: A web-based DNA wireframe design tool for 3D structures (pg 79)
Samson Petrosyan and Grigory Tikhomirov (eligible for best student / postdoc poster award)

#58 small: A modular, extensible nanostructure design framework (pg 80)
Durham Smith and Grigory Tikhomirov (eligible for best student / postdoc poster award)

#60 DNA nanochannels leading to an artificial single-cell communication system (pg 81)
Hiroki Koiwa and Kan Shoji (eligible for best student / postdoc poster award)

#62 A refined shape annealing algorithm for the optimal generation of DNA origami designs (pg 82)
Bolutito Babatunde, Jonathan Cagan and Rebecca Taylor (eligible for best student / postdoc poster award)

#65 Distinguishing Molecular Circuit Input Pulses via a Pulse Detector (pg 83)
Colin Yancey and Rebecca Schulman (eligible for best student / postdoc poster award)

#66 A loser-take-all DNA circuit (pg 84)
Kellen Rodriguez, Namita Sarraf and Lulu Qian (eligible for best student / postdoc poster award)

#68 Rate-Independent DNA-Based Binary-Weight ReLU Neural Networks (pg 85)
Cameron Chalk, Boya Wang, Marko Vasic and David Soloveichik (eligible for best student / postdoc poster award)

#69 A minimal DNA Origami for Seeding Tiled DNA Nanotube Bundles (pg 86)
Sarah Webster and Deborah Kuchnir Fygenson (eligible for best student / postdoc poster award)

#70 Efficient Elementary Step Implementations in Nucleic Acid Kinetics Simulators (pg 87)
Boyan Beronov, Jordan Lovrod, Chengwei Zhang and Anne Condon (eligible for best student / postdoc poster award)

#72 Scaling up reusable DNA circuits using heat as a universal energy source (pg 88)
Tianqi Song and Lulu Qian (eligible for best student / postdoc poster award)

#74 The Transformation of Real-Valued Concentrations into Binary Signals (pg 89)
James Lathrop, Dawn Nye and Hugh Potter (eligible for best student / postdoc poster award)

#75 Qslib: Python control of qPCR machines for molecular programming experiments (pg 90)
Constantine Evans

#78 Compact and Efficient Chemical Boltzmann Machines (pg 91)
Inhoo Lee, William Poole and Erik Winfree (eligible for best student / postdoc poster award)

#80 Highly Parallel Synthesis, Computation, and Readout of DNA Logic Gates using High-Throughput Synthesis and Sequencing (pg 92)
Chandler Petersen, Samantha Borje, Gourab Chatterjee, Yuan-Jyue Chen and Georg Seelig (eligible for best student / postdoc poster award)

#81 Growing arbitrary patterns with neural reaction-diffusion (pg 93)
Salvador Buse and Erik Winfree (eligible for best student / postdoc poster award)

#83 Counting the number of input addition by DNA reaction system driven by DNA polymerase (pg 94)
Ibuki Kawamata, Motokazu Furuya and Satoshi Murata

#84 Facet TAM: An abstract tile assembly model with facet growth (pg 95)
Ahmed Shalaby, Constantine Evans and Damien Woods (eligible for best student / postdoc poster award)

#85 Formal Semantics for Stochastic Chemical Reaction Networks (pg 96)
Andres Ortiz-Munoz (eligible for best student / postdoc poster award)

Storing digital data in DNA-based tape

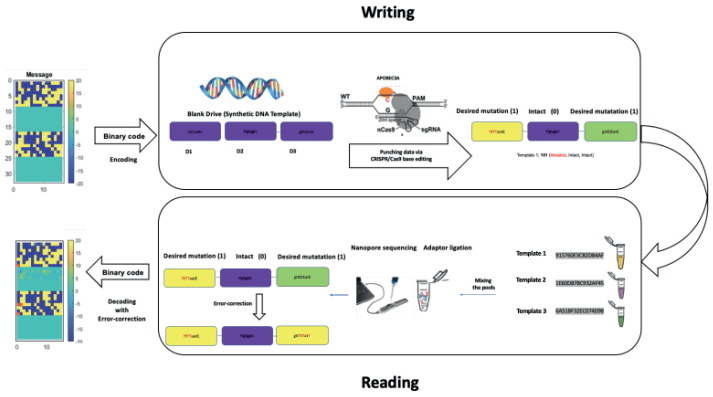
Afsaneh Sadremomtaz[1,3], Robert Glass[2,3], Jorge Guerrero[1,3], Micheal Brandon Reed [1,3], Eric Josephs[2,3]*, Reza Zadegan[1,3]*

[1] Department of Nanoengineering, NC A&T State University, Greensboro, NC, USA

[2] Department of Nanoscience, UNC Greensboro, Greensboro, NC, USA

[3] Joint School of Nanoscience and Nanoengineering, Greensboro, NC, USA

Global digital data will surpass 1.4 Yottabits by 20251. While demands for data storage increases, the conventional digital storage systems are reaching their physical limits2. Therefore, alternative information storage materials including DNA storage systems that could provide viable energetics, spatial capacity, data retention, and economics solutions are becoming a trend[2]. DNA as an alternative for data storage potentially enables great physical density and scalability[1,2]. Additionally, rapid development of DNA synthesis and sequencing has helped the community to consider DNA as an alternative information storage medium[2]. Considerable challenges that need to be addressed for DNA to become a main-stream DNA storage include sustainability, latency, and scalability. While majority of the community are attempting to address scalability, we believe de novo DNA synthesis –which in most cases results in large amounts of toxic waste– is a major bottleneck[1,2]. In this report, we used DNA Mutational Overwriting Storage (DMOS) to overwrite the sequence content of greenly synthesized template DNA domains (bits) to write the digital data. Our DMOS DNA templates contain 16-bit domains and addressing strings. We wrote the digital information by allowing or prohibiting localized mutations the DNA templates via guided synthetic chimeric enzymes (hereby called punchers). Our puncher enzyme (APOBEC3A) recognizes the desired domain by aid of our predefined library of ribonucleic protein complexes (CRISPR-dead Cas9 and guide RNAs)[3]. The developed cell-free punchers perform cytidine deamination of single-stranded DNA strands, resulting in base substitution mutations of base-C (intact state) to base-T (mutated state) in the targeted domains[3]. Also, we developed error-correction and DNA sequencing codes that enable high data retention in our DMOS system. As schematically demonstrated in Fig.1, we performed mutational edits to write the data, added addressing strings to the encoded DNA pools and combined them to store the data, and performed nanopore sequencing to read the data[4].



"This work was funded in part by the National Science Foundation (MCB 2027738)"

1- Zhirnov, V., Zadegan, RM., Sandhu, GS., Church, GM., Hughes, WL. Nucleic acid memory. *Nat Mater*. 2016;15(4):366-370. doi:10.1038/nmat4594.

2- Zhrinov, V., Rasic, D. 2018 Semiconductor Synthetic Biology Roadmap. (Retrieved on 03/14/2020 from https://www.src.org/library/publication/p095387/p095387.pdf).

3- Marshall, R., Maxwell, C. S., Collins, S. P., Jacobsen, T., Luo, M. L., Begemann, M. B., Gray, B. N., January, E., Singer, A., He, Y., Beisel, C. L. & Noireaux, V. Rapid and Scalable Characterization of CRISPR Technologies Using an E. coli Cell-Free TranscriptionTranslation System. Molecular cell 69, 146-157.e143 (2018), doi:10.1016/j.molcel.2017.12.007.

4- Timp, W., Comer, J. & Aksimentiev, A. DNA base-calling from a nanopore using a Viterbi algorithm. Biophysical Journal 102, L37-L39 (2012), doi:10.1016/j.bpj.2012.04.009.

Nucleic Acid Memory

Super Resolution Microscopy enhances novel approach to DNA data storage

Luca Piantanida[1], George D. Dickinson[1], Golam M. Mortuza[2], William Clay[1], Christopher M. Green[1,5], Chad Watson[1], Eric J. Hayden[3], Timothy Andersen[2], Wan Kuang[4], Elton Graugnard[1], Reza Zadegan[1,6] and William L. Hughes[1]

[1]Micron School of Materials Science and Engineering, Boise State University, Boise, ID, USA. [2]Department of Computer Science, [3]Department of Biological Sciences, [4] Department of Electrical and Computer Engineering, Boise State University, Boise, ID, USA. Present address: [5]Center for Bio/Molecular Science and Engineering, U.S. Naval Research Laboratory, Washington, DC, USA; [6]Department of Nanoengineering, Joint School of Nanoscience & Nanoengineering, North Carolina A&T State University, USA.

The ever-present connectivity in our lives, and the data storage demands that come with it, is growing exponentially. The projected material supply for silicon-based memory technologies is unable to satisfy future demand, therefore, alternative memory materials are being explored in academia and industry[1]. DNA is analogous to a biological hard drive. It carries and transfers information with exceptional density, stability, and energy efficiency, making it a compelling alternative to current non-volatile information storage technologies.

The Nucleic Acid Memory (NAM) Institute at Boise State exploits DNA as a programmable material to engineer emerging data storage technologies. Here we present our first prototype, digital Nucleic Acid Memory (dNAM), which spatially encodes and retrieves small datasets using only DNA as the material[2].

dNAM is made from DNA origami assembly technique, it is structurally characterized using Atomic Force Microcopy (AFM, Fig.1b) and it is read using Super Resolution Microscopy (SRM, Fig.1c). The origami structure serves as a breadboard where short dye-labelled DNA strands are the fluorescent imager probes (Fig.1a). These probes transiently hybridize with short protruding single strands periodically positioned on the origami design implementing the DNA PAINT technique on its surface. Every time a hybridization event occurs, a blinking signal is recorded. Associating the blinking signal to a “1” and no signal to a “0”, dNAM resembles a molecular version of Lite-Brite toy where distinct patterns of pegs encode different digital data. Processing the SRM reading with a custom error-correction algorithm, dNAM provides an areal data density of 330 Gbit/cm^2 and is able to use a subset of origami to retrieve the encoded message 100% of the time. Unlike other approaches to DNA-based data storage, reading dNAM does not require DNA sequencing and so, the synthesis of new custom DNA strands every time new data is encoded, reducing significantly the costs. As such, this research can provide a valuable path to DNA data storage applications for the next-generation of digital memory materials.

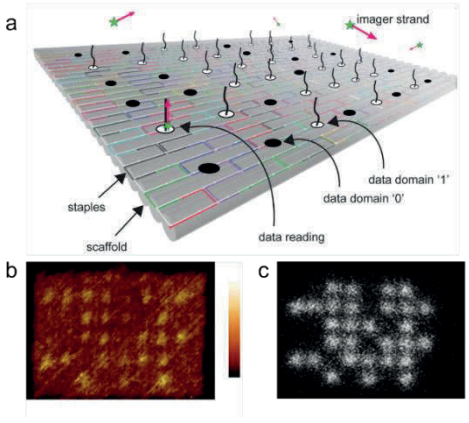


Figure 1: (a) dNAM design, (b) structure (AFM, color scale 2nm) and (c) DNA PAINT reading (SRM).

References

[1] Zhirnov, V., Zadegan R., Sandhu G. et al., *Nat. Mat.*, 2016, **15**(4) 366-370.
[2] Dickinson, G.D., Mortuza G., Clay W., Piantanida L., et al., *Nat. Commun.*, 2021, **12**(1) 2371.

Nanopore decoding for DNA-computed over and under-expression of microRNA patterns

Sotaro Takiguchi, and Ryuji Kawano

Tokyo University of Agriculture and Technology, Japan

DNA computing has been expected to expand the research field from mathematical computations to diagnostic applications due to the biocompatibility and sequence-programmability of DNA molecules. However, decoding the output information to a human-recognizable signal generally requires multiple-step procedures or fluorescence detection. We have previously proposed simple and label-free decoding of oligonucleotides in mathematical DNA computation with nanopore technology [1-2]. Here, as a real-life application of our proposed method, we report a nanopore decoding for DNA computing-assisted pattern recognition of microRNAs, which are promising biomarkers for early cancer diagnosis. The microRNA profile in bodily fluids shows over and under-expression with cancer type-specificity and its simultaneous monitoring is challenging. In this study, utilizing the sequence design of DNA-based parallel computations [3], we designed the diagnostic DNAs to enable simultaneous recognition of two individual over and under-expressed microRNAs (Fig. 1). According to thermodynamic simulation, the DNA-computed output resulted in three types of duplexes, in which the abundance ratio changed depending on the four types of microRNA expression patterns (Fig. 2). In nanopore measurement, the output DNA/microRNA duplex undergoes unzipping whilst passing through an α-hemolysin, a pore-forming protein, nanopore, showing a current blocking signal. We confirmed that each output duplex was distinguishable in nanopore measurement as a specific current blocking signal. As the results of nanopore decoding, the calculated abundance ratio was consistent with the results of the simulation, indicating that we successfully demonstrated the simultaneous monitoring of over and under-expression patterns of microRNAs. Furthermore, we confirmed that this system can work with target specificity and correctly work in the case of low concentrations of microRNA. We believe that our method will be applicable as a simple cancer diagnostic tool.

References

[1] R. Kawano 2018 *Biotechnol. J.* 1800091

[2] S. Takiguchi and R. Kawano 2021 *Nanoscale* **13** 6192-6200

[3] L. M. Adleman 1994 *Science* **226** 1021-102

Acknowledgement

This work was partially supported by KAKENHI (No. 19H00901) from MEXT.

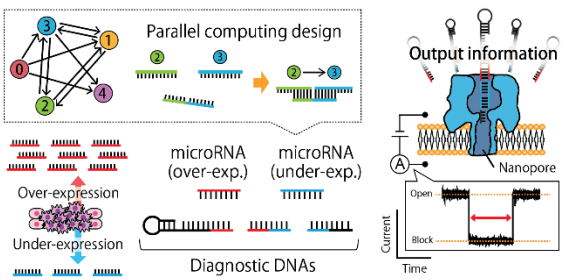


Fig. 1 The design of diagnostic DNAs and the illustration of nanopore decoding.

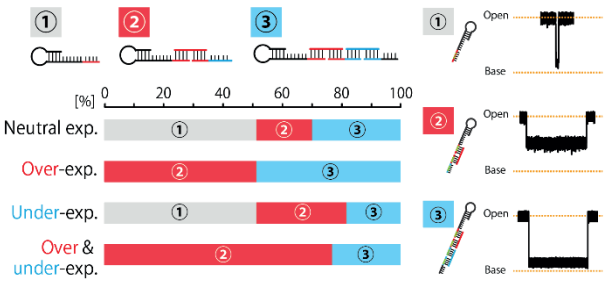


Fig. 2 The duplex structures and abundance ratio of the DNA-computed output.

Kinetic approach to suppress DNA hybridization error

Hiroyuki Aoyanagi[1], Simone Pigolotti[2], Shinji Ono[1], and Shoichi Toyabe[1]

[1]Tohoku University, Japan, [2]Okinawa Institute of Science and Technology Graduate University, Japan

Introduction The specificity of DNA hybridization is crucial in various technologies such as PCR, genome editing, and DNA origami. Energetic suppression is a conventional methodology for decreasing the hybridization error (e.g., Tm tuning), which relies on the free energy difference ΔG (Fig. 1a). However, such approaches have a fundamental limitation in performance [1]. In contrast, we can avoid such a limitation by kinetic suppression by sculpting a kinetic barrier ΔG_a (Fig. 1b). In this study, we show that we decreased the hybridization error by a simple method and that the decrease is the consequence of kinetic suppression.

Methods We set up a PCR-based experimental system. We mixed a single side of the primer (P), two types of templates (R, W), and polymerase (Fig. 1a). The primer (P) hybridizes to the template (R) complementary but to the template (W) with a base mismatch. We quantified the elongated products $(\overline{\mathrm{R}},\ \overline{\mathrm{W}})$ by qPCR. We defined the error rate η as a fraction of $\overline{\mathrm{W}}$ against all the products $\overline{\mathrm{R}} + \overline{\mathrm{W}}$. Our approach to decrease η is to add a blocker strand $(\mathrm{B_W})$ which strongly hybridizes to the primer region of the template W (Fig. 1b).

Results Without the blocker, we had to sacrifice the reaction efficiency to decrease η, which shows the limitation of the energetic suppression [1]. On the other hand, we could decrease η by about 80 % compared to the no-blocker condition without sacrificing the efficiency by adding the blocker $\mathrm{B_W}$. Moreover, we could also increase η by adding another blocker that preferably hybridizes to R. These results qualitatively indicate that a kinetic barrier was exactly built by the addition of the blocker.

We created a mathematical model and successfully reproduced our experimental results. From the model, we obtained a representation of ΔG_{a} when we added the blocker and quantitatively proved that this simple controlling method is the consequence of kinetic suppression.

Importance A physics-based approach for binding error suppression has not been explored so much in biotechnology. This quantitative characterization enables applications of this blocker method to various technologies other than PCR.

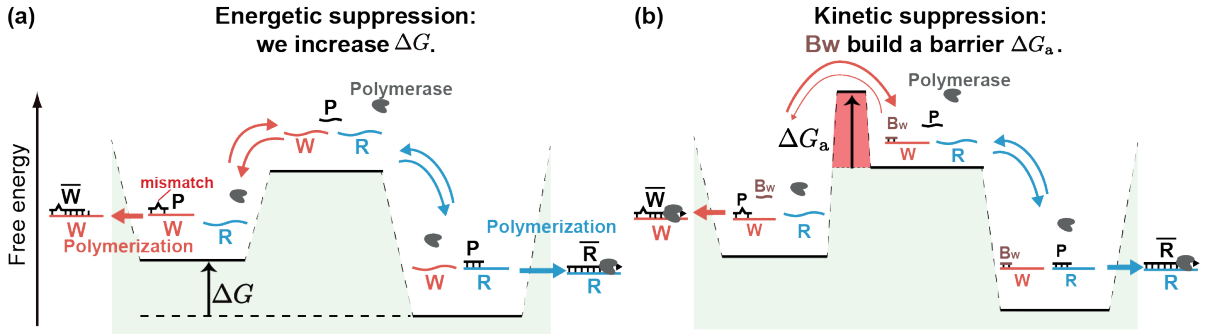


Figure 1. (a) Energetic suppression and the experimental setup. (b) Kinetic suppression by the blocker method.

[1] Sartori P and Pigolotti S 2013 *Phys. Rev. Let.* **110** 188101

A geometric framework for reaction enumeration in computational nucleic acid devices

Sarika Kumar[1], Matthew R. Lakin[1,2,3]
[1]Dept. of Computer Science, University of New Mexico, Albuquerque, NM, USA
[2]Dept. of Chemical & Biological Engineering, University of New Mexico, Albuquerque, NM, USA
[3]Center for Biomedical Engineering, University of New Mexico, Albuquerque, NM, USA

Abstract: Cascades of DNA strand displacement reactions enable the design of potentially large circuits with complex behavior. Computational modeling of such systems is desirable to enable rapid design and analysis. In previous work [1], graph theory was used to enumerate reactions implementing strand displacement between a wide range of complex structures. However, to cope with the rich variety of possible graph-based structures required enumeration rules with complicated side-conditions. This work presents an alternative approach to tackle the problem of enumerating reactions at domain level involving complex structures by integrating with a geometric constraint solving algorithm. The rulesets from previous work are simplified by replacing side-conditions with a general check on the geometric plausibility of structures generated by the enumeration algorithm. This produces a highly general geometric framework for reaction enumeration. Here, we instantiate this framework to solve geometric constraints by a structure sampling approach in which we randomly generate sets of coordinates and check whether they satisfy all the constraints. We demonstrate this system by applying it to several examples from literature where molecular geometry plays an important role, including DNA hairpin reactions and remote toehold reactions. This work therefore opens a path to the integration of reaction enumeration and molecular structural modeling, leading to new tools for rapid and precise domain-level modeling of DNA strand displacement reaction networks.

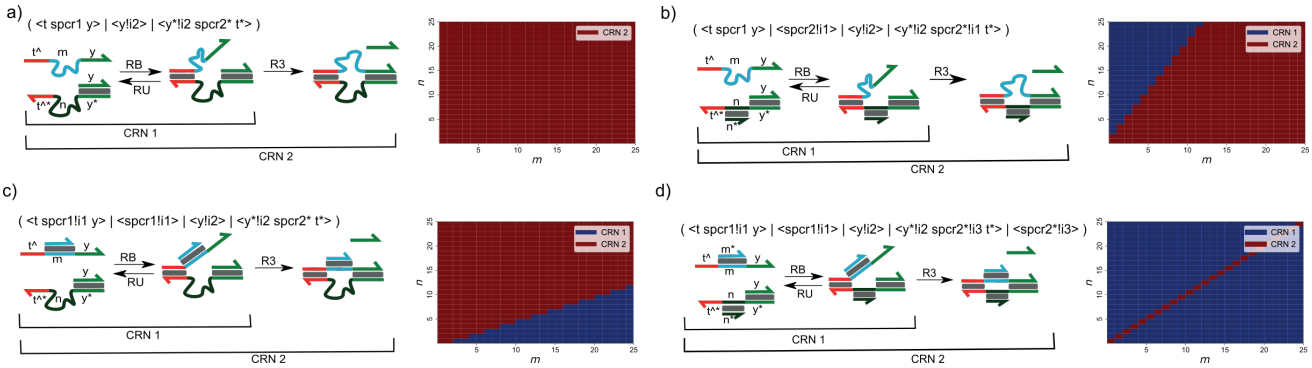


Figure: Geometric enumeration results of remote toehold reaction [2] using structure sampling approach. CRN 1 represents the toehold binding reaction and CRN 2 represents both toehold binding and strand displacement reaction. The effect of modulating the length and changing the nature of spacers m and n on the plausibility of CRN 1 and CRN 2 is shown graphically. In the plots, blue color region represents the plausibility of CRN 1 and red color region represents the plausibility of CRN 2. This demonstrates that the feasibility of a CRN can be determined based on its geometric property.

References:

[1] Petersen R L, Lakin M R, Phillips A 2016 *Theor. Comput. Sci.* **632** 43–73
[2] Genot AJ, Zhang DY, Bath J, Turberfield AJ 2011 JACS 133(7):2177–2182.

Study of DNA nanostructures for delivering gene templates for homology directed repair

Diana Lopez, Wolfgang Pfeifer and Carlos Castro

Despite major advances in gene editing technologies, currently it remains a major challenge to deliver genes more than a few kilobases in length. DNA origami structures, where long single-stranded DNA is folded into a compact nanostructure, present an attractive approach to package large genes for delivery into cell nuclei; however, delivery of genetic material that has been folded into a compact structure into cell nuclei has remained largely unexplored. This work focuses on studying the use of DNA nanostructures to package genes into DNA nanostructures for the delivery of templates for homology directed repair (HDR) into live cells. HDR relies on homology arms whose sequence mimics regions of genome to enable incorporation of genes at specific target sites mediated by CRISPR-Cas9 proteins. Folding the HDR template DNA into a compact structure provides the advantage of holding the homology arms close together, which may increase the efficiency of incorporation. We implemented the oxDNA model to simulate and guide the design process of different versions of CRISPR–Cas9 HDR templates. To assess functional differences between the unstructured HDR templates and the origami versions, we also performed simulations of all unstructured versions and computed the average distances between each DNA template termini. Consistent with AFM characterization of the HDR templates, the simulation results predicted that terminal homology sites are farther apart in the unstructured template compared to the folded nanostructure, which we believe plays a key role in the gene integration of the CRISPR–Cas9 HDR mechanism.

Scalable design and construction of DNA truss structures

Yudai Yamashita[1], Ibuki Kawamata[1, 2], Satoshi Murata[1]

[1]Tohoku University, Japan [2]Ochanomizu University, Japan

DNA nanostructures constructed by single-stranded tile (SST) method [1] can obtain high addressability because each position can correspond to a unique sequence of DNA, and in addition, the structures can be directly fabricated in one-pot assembly. Those structures are expected to contribute to various nanotechnology applications that need precise arrangement of various molecules. To scale up the methodology and construct larger DNA structures, recent technology to synthesize numerous DNAs with arbitrary sequences is made use of [2]. One of the directions to further scale up the size of the DNA structures is improving the space efficiency of the tile while maintaining geometric accuracy and the material rigidity. When the number of strands becomes large, it is also necessary to carefully design the base sequences of DNAs, by which non-specific hybridization can be prevented.

In this work, we propose a novel method to create large-scale DNA nanostructures consisting of 4-domain DNAs (Fig. 1a) synthesized by an oligo-pool technology [3]. Introducing a truss-wireframe geometry, we aim to design rigid and scalable structures (Fig. 1b). The DNA nanostructure with trusses aligned between two layers of a square lattice can be broken down into a structural unit of 8-arm junction (Fig. 1c). Each single-stranded DNA has one poly-A domain and three different domains whose sequence can be computationally designed to be orthogonal (Fig. 1d). Based on the method, we designed several sizes of structures, one of which has 1870 strands (Fig. 1e, f), and demonstrated the scalability and versatility of the method. The latest experimental results will be presented at the conference along with coarse-grained simulation results using oxDNA [4]. Our method may provide a platform for a variety of applications that require precise and wide-area arrangement of molecules.

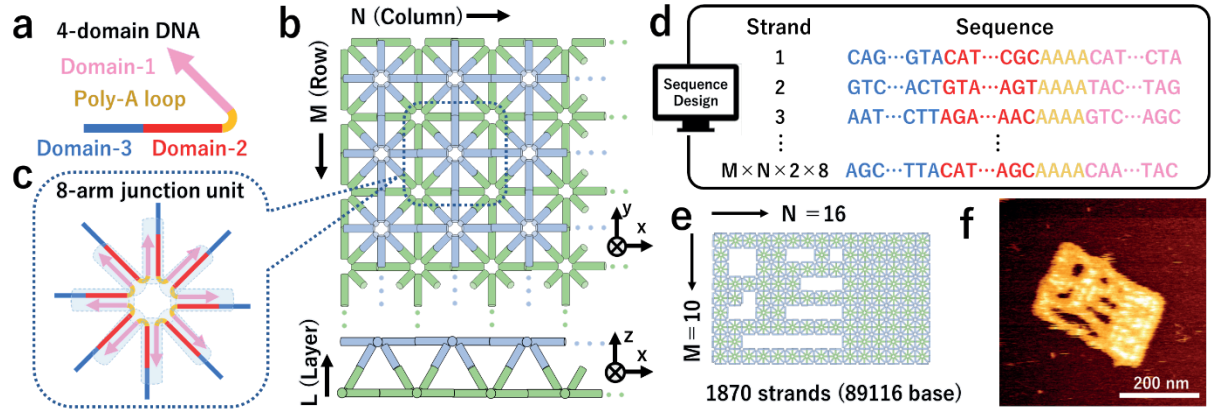


Figure 1: Design and experimental results. (a) Four domains of DNA. (b) Design of a M×N×2-unit DNA truss structure. (c) 8-arm junction unit. (d) List of designed orthogonal sequences. (e) Design of a DNA truss structure with a specific pattern of cavities. (f) AFM image.

[1] Wei, B., et al. (2012) *Nature*, **485**(7400), 623-626.
[2] Ong, L. L., et al. (2017) *Nature*, **552**(7683), 72-77.
[3] Tian, J., et al. (2004) *Nature*, **432**(7020), 1050-1054.
[4] Ouldridge, T.E., et al. (2011) *J. Chem. Phys,* **134**(8), 02B627.

NanoFrame: A web-based DNA wireframe design tool for 3D structures

Samson Petrosyan, Grigory Tikhomirov

May 30, 2022

1 Abstract

NanoFrame.org is a web-based DNA nanostructure design tool for wireframe origami cages. The synthesized shapes can be combined together in NanoFrame's Playground to construct more complicated multi-origami structures. We focus on wireframes because they afford to cover a larger surface area compared to close-packed conventional Rothemund's DNA origami with the same amount of DNA. This is important for embedding molecular recognition capability properties of DNA into non-DNA materials to guide their assembly. To achieve this goal, we make the cages openable at the desired temperature by synthesizing structures with the flexibility to wrap around nanoparticles.

We developed algorithms for scaffold routing, staple breaking, and cage opening. In addition to generating polyhedra with a graphical user interface, users can also program shapes using NanoFrame's JSON-like NFR format. To enable cross-platform integration with other software, we provide tools for converting NFR files into other nucleic acid file formats.

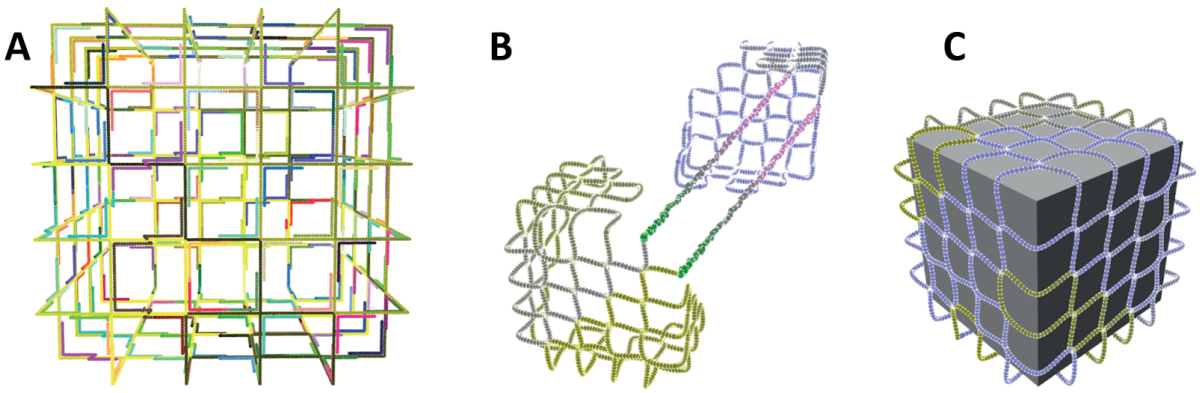


Figure 1: Wireframe suit designed with our software NanoFrame.org showing closed DNA cage (A), open form (B), and closed form with a nanoparticle inside (C).

`small`: A modular, extensible nanostructure design framework

Durham Smith and Grigory Tikhomirov

Electrical Engineering and Computer Sciences, University of California, Berkeley

Structural DNA nanotechnology has advanced to enable the creation of incredibly complex structures which are able to integrate a wide variety of non-DNA molecules. These advances are partly due to the development of design and simulation tools for creating DNA nanostructures and partly due to chemists' ingenuity in creating means to functionalize DNA with other chemical species. Modeling and simulation the properties of these hybrid nanomaterials may go beyond the capabilities of current structural DNA nanostructure design tools. For example for modeling the electromagnetic field along a zero-mode waveguide made of gold nanoparticles organized on a DNA breadboard would require both simulation of the structural integrity of the DNA breadboard, using a molecular dynamics, and of the electromagnetic fields around the structure, using a Maxwell's equation solver. Such simulation tools are typically stand-alone programs, requiring users to both describe the system under simulation in each program and be familiar with the intricacies of each program's use. A framework that allows the integration of such tools would streamline the designs in DNA nanotechnology and enable extending DNA nanotechnology principles to construct high performance materials and devices from non-DNA components.

Here we present `small` a framework for designing arbitrary hybrid DNA nanostructures and integrating with third party simulation software. We provide examples of `small` integration with oxDNA for MD simulation of DNA nanostructures, oxView, VMD and ChimeraX for visualization, and MEEP for the simulation of Maxwell's equations. Furthermore, `small` is able to export DNA sequences of the structures designed in a format directly able to be ordered from IDT as well as produces pick lists required to automate the synthesis of these on the Echo 525 acoustic liquid handler.

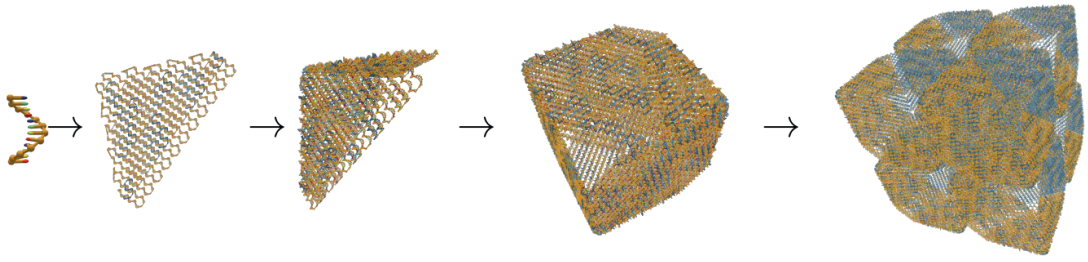


Figure 1: A $2 \times 2 \times 2$ array of DNA cubes, each comprised of 6 DNA origami tiles, designed using `small`.

DNA nanochannels leading to an artificial single-cell communication system

Hiroki Koiwa and Kan Shoji

Nagaoka University of Technology, Japan

This paper describes the development of a DNA nanochannel that has enough length to penetrate two cell membranes. The DNA nanochannel was constructed by a six-helix bundle (6HB) DNA structure, and the length of the channel was 21 nm. In this study, we performed the channel current measurement of DNA nanochannels to evaluate the ion permeability of nanochannels.

1. Introduction - Gap junctions (GJs) play a significant role in cell-cell communications by directly connecting two cells with nanochannels. Thus, if an artificial nanochannel that works as GJ systems can be developed and inserted into target cells, an artificial single-cell communication system will be developed. In this study, we propose a transmembrane DNA nanochannel with enough length to connect two cell membranes and evaluate the ability to form ion channels by channel current measurements of the DNA nanochannel.

2. Methods – We designed the 6HB DNA nanostructure with 21 nm length as the transmembrane DNA nanochannel. Then, we conducted channel current measurements of the DNA nanochannels using gold-needle-supported lipid bilayers [1]. DNA nanochannels were pre-modified on the gold needle, and the bilayer was formed by inserting the needle into a layered bath chamber of an electrolyte and an oil/lipid mixture. DNA nanochannels were spontaneously inserted into the bilayer, and the ion currents were recorded with a patch-clamp amplifier (Fig. 1a).

3. Results and Discussion – Step-like current increases were observed, indicating insertions of DNA nanochannels into the lipid bilayer (Fig. 1b). From these step-like signals, we calculated the channel conductance, and conductance values between 0.1 to 0.5 nS were mainly obtained (Fig. 1c). Since the theoretical conductance value of the single DNA nanochannel can be calculated as 0.3 nS using Hille equation [2], we concluded that single nanochannel insertions were obtained in the channel current measurements. From these results, our DNA nanochannel penetrates the lipid bilayer and functions as an ion channel. As a next step, we will evaluate the penetration ability of the DNA nanochannel through two lipid bilayers by constructing the lipid bilayer formation system that can form two lipid bilayers.

4. References

[1] Shoji K et al 2019 *ACS Nano* **13** 2606-2614
[2] Hille B 2001 *Ion channels of excitable membranes 3rd ed.* (Sunderland, MA: Sinauer)

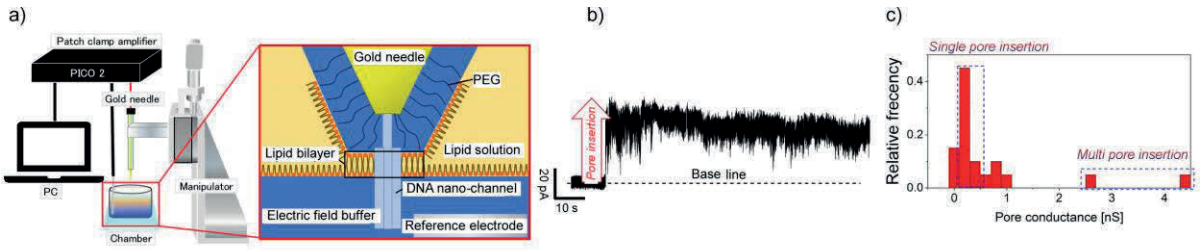


Fig. 1 a) A scheme of an ion current measurement system. b) A recorded channel current of a DNA nanochannel. c) A histogram of pore conductance calculated from current signals.

A refined shape annealing algorithm for the optimal generation of DNA origami designs

Bolutito Babatunde[1], Jonathan Cagan[1] and Rebecca Taylor[1,2,3]

[1]Department of Mechanical Engineering, [2]Department of Biomedical Engineering, [3]Department of Electrical and Computer Engineering, Carnegie Mellon University, USA

In the recent years, the field of structural DNA nanotechnology has advanced rapidly due to transformative design tools. Although these tools have been revolutionary, they still bear two key limitations. Base-level design tools push users towards conceptually simpler extrusion-style designs due to the complexity of manually constructing more irregular geometries. This limitation gave rise to the development of powerful automated and semi-automated tools for generating wireframe [1] and solid extrusion-based systems [2]. However, these new tools still limit users to fully conceptualize their design and are therefore unable to assist with practical challenges encountered in DNA origami design such as efficient use of the scaffold.

We have previously shown that generative optimization strategies can be used to design DNA origami nanostructures through automation [3]. In this work, we show how this method can be used to implement coating-type designs for arbitrary input shapes that use differing amounts of scaffold. This is accomplished by employing a shape annealing algorithm [4], which creates a language of DNA nanostructures with shape grammars [5] and drives designs from the language towards a desired configuration using simulated annealing [6]. Shape grammars use a formal rule-based method to define the relationship between geometric shapes while simulated annealing is a stochastic optimization technique. This formal approach addresses the limitations from current automated and semi-automated tools by generating complex DNA origami designs in an optimally driven way only with a set of boundaries and desired characteristics. In previous work, this new approach was demonstrated by directing the scaffold routing pattern to fill or coat the walls of distinct polyhedral meshes as boundaries. Here, we present, a more refined approach that addresses the quality, thickness, and scaffold utilization of generated coatings.

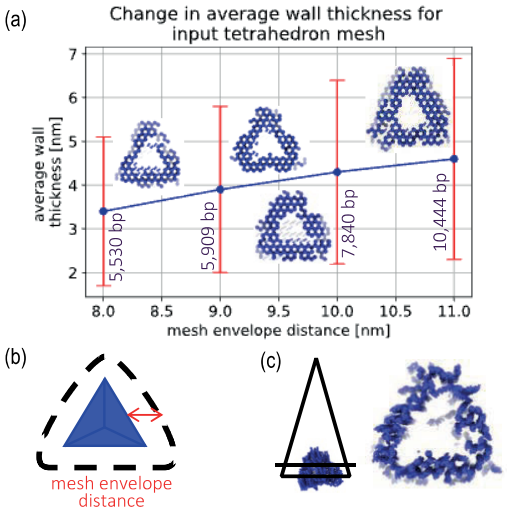


Fig 1. Results from refined method using a tetrahedral mesh envelope as bounds of the design space: (a) Plot of average change in wall thickness [nm] with increasing average scaffold length and mesh envelope distance (MED) [nm] including an example of slice view of shape generated per distance; (b) MED is the set distance to mesh for scaffold generation; (c) Slice view example of fully relaxed configuration from MED of 10 nm after oxDNA simulation with total scaffold length of 7,063 bp. Isometric view showing exact slice location (left).

Improvements are achieved through changes in the simulated annealing algorithm by adding a Hustin move set algorithm [7] and refining the articulation of coating behavior. The Hustin move set algorithm dynamically adjusts scaffold route selection probabilities based on previous performance so that routes that improve the overall scaffold pattern have a higher probability of being selected. The result in Fig 1(a) shows an improved control of generated scaffold route towards the edge of the walls of the structure and optimal scaffold length usage with increasing average wall thickness. Fig 1(c) is an example of the shape generated which shows fair maintenance of structural integrity with slight unraveling after oxDNA simulation. The results in Fig 1 illustrate the potential of this approach in coding coating behavior.

[1] Jun H et al 2019 *ACS Nano* **13** 2083, [2] Huang C-M et al 2020 *bioRxiv*, [3] Babatunde B, Cagan J, and Taylor R E 2021 *Appl. Sci.* **11** 2950, [4] Cagan J and Mitchell W J 1993 *Environ. Plann. B* **20** 5, [5] Stiny G 1980 *Environ. Plann. B* **7** 343, [6] Kirkpatrick S, Gelatt C D, and Vecchi M P 1983 *Science* **220** 671, [7] Hustin S and Sangionvanni-Vincetelli A 1987 *IEE Physical Design Workshop on Placement and Floorplanning (Hilton Head, SC*)

Distinguishing Molecular Circuit Input Pulses via a Pulse Detector

Colin Yancey ✉
Department of Chemical and Biomolecular Engineering, Johns Hopkins University, MD, USA

Rebecca Schulman ✉
Department of Chemical and Biomolecular Engineering, Johns Hopkins University, MD, USA
Department of Computer Science, Johns Hopkins University, MD, USA
Department of Chemistry, Johns Hopkins University, MD, USA

Abstract

Chemical systems have the potential to direct the next generation of dynamic materials if they can be integrated with a material while acting as the material's own regulatory network. Molecular circuits that use DNA and RNA strand displacement coupled with RNA synthesis, such as genelets, are promising chemical systems for this role. Genelets can produce a range of dynamic behaviors that respond to unique sets of environmental inputs. However, the general design of genelet circuits for this task remains a challenge: a circuit needs an information space large enough to accommodate an assortment of computational features, but there are few general strategies for designing circuit parameters in such a complex system. Here we seek to address this challenge by implementing a specific type of circuit design. While a number of networks that generate specific types of time-varying outputs have been developed, there are fewer examples of networks that recognize specific types of time-varying inputs. We design genelet networks that respond to such specific types of time-varying inputs and test their performance using an *in silico* model of circuit behavior. Specifically, the genelet circuits we design act as bandpass filters for their input, so that the circuits recognize input pulses of specific sizes. These networks' sizes are commensurate with the sizes of circuits that have been synthesized previously. The circuits also use reasonable concentrations of circuit elements. These factors suggest that they could be built and characterized in the laboratory.

2012 ACM Subject Classification Hardware → Biology-related information processing; Computing methodologies → Modeling and simulation

Keywords and phrases DNA nanotechnology, Information processing, Molecular circuits, DNA computing

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

Funding The authors are grateful to NSF award CIF-2107246 for support of this work.

42nd Conference on Very Important Topics (CVIT 2016).
Editors: John Q. Open and Joan R. Access; Article No. 23; pp. 23:1–23:1

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

A loser-take-all DNA circuit

Kellen R. Rodriguez[1,2,4†], Namita Sarraf[3†] and Lulu Qian[3,4⋆]
[1]*Business Economics & Management,* [2]*Astrophysics,* [3]*Bioengineering, and* [4]*Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA
[†]*Equal contribution,* [⋆]*e-mail: luluqian@caltech.edu*

Abstract. DNA-based neural networks are a type of DNA circuit capable of molecular pattern recognition tasks. Winner-take-all DNA networks have been developed to scale up the complexity of molecular pattern recognition with a simple molecular implementation. This simplicity was achieved by replacing negative weights in individual neurons with lateral inhibition and competition across neurons, eliminating the need for dual-rail representation. Here we introduce a new type of DNA circuit that is called loser-take-all: an output signal is ON if and only if the corresponding input has the smallest analog value among all inputs. We develop a DNA strand-displacement implementation of loser-take-all circuits that is cascadable without dual-rail representation, maintaining the simplicity desired for scalability. We characterize the impact of effective signal concentrations and reaction rates on the circuit performance, and derive solutions for compensating undesired signal loss and rate differences. Using these approaches, we successfully demonstrate a three-input loser-take-all circuit with nine unique input combinations. Complementary to winner-take-all, loser-take-all DNA circuits could be used for recognition of molecular patterns based on their least similarities to a set of memories, allowing classification decisions for patterns that are extremely noisy. Moreover, the design principle of loser-take-all could be more generally applied in other DNA circuit implementations including k-winner-take-all.

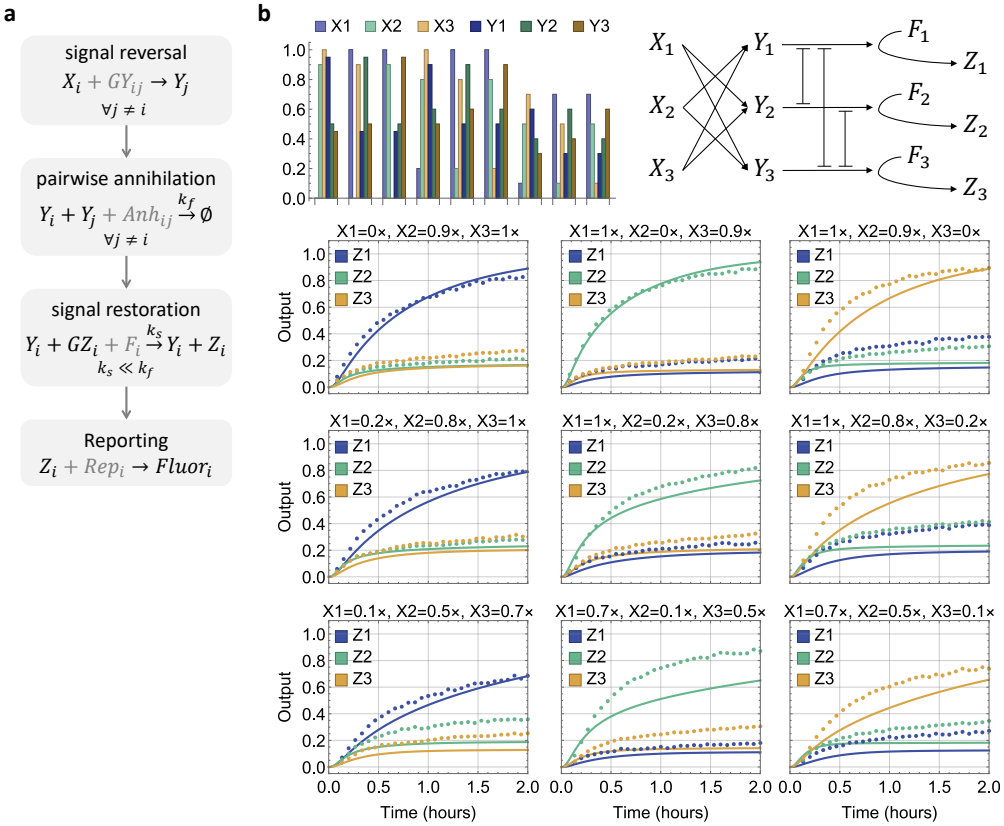


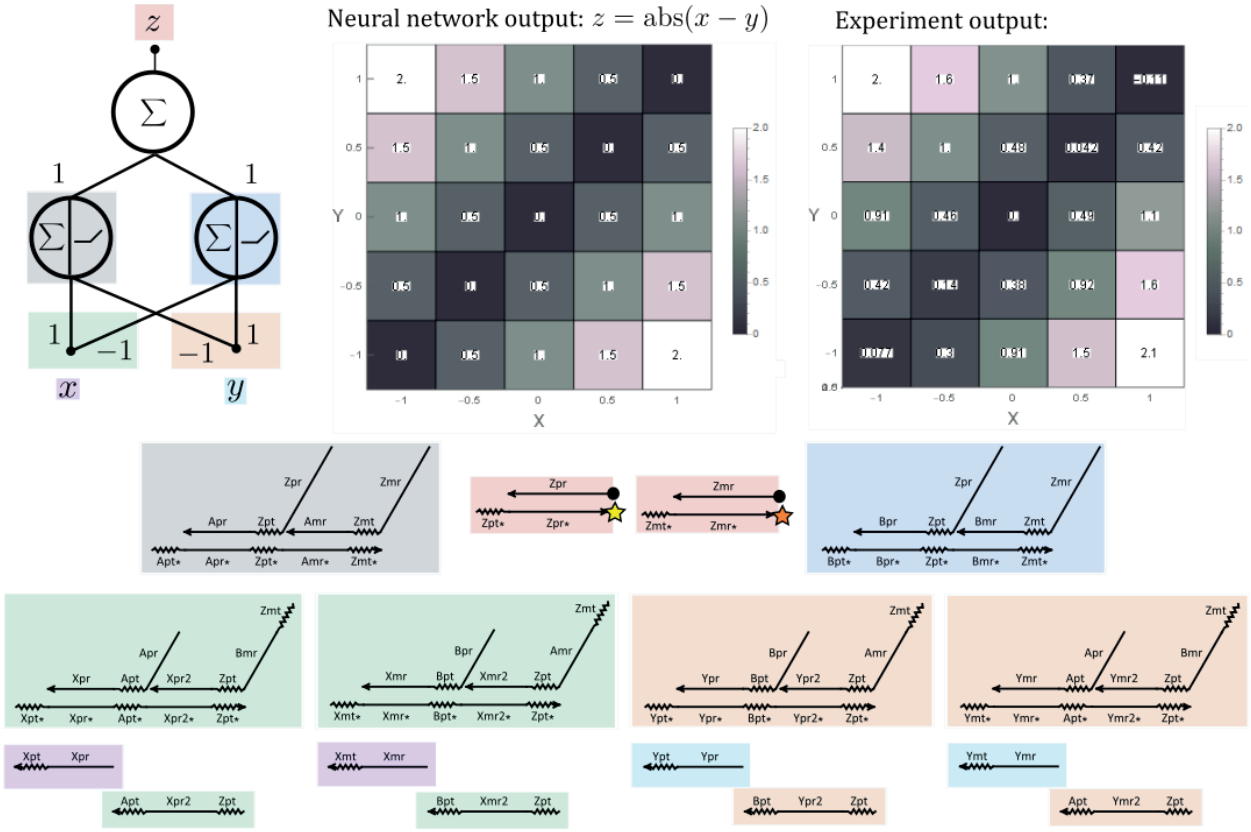
Figure 1: **DNA strand-displacement implementation of a three-input loser-take-all circuit. (a)** Chemical reaction network implementation. **(b)** Simulation and fluorescence kinetics data with nine combinations of input concentrations. Abstract reaction diagram indicates the reactions involved in the experiments. Bar chart shows all input values and expected reversed signal values.

Rate-Independent DNA-Based Binary-Weight ReLU Neural Networks

Cameron Chalk[1], Boya Wang[1], Marko Vasić[1], David Soloveichik[1]

[1]The University of Texas at Austin, USA

Rate-independent chemical reaction networks (RICRNs) [1] guarantee equilibria that are independent of reaction rate constants or kinetic rate law. In other words, they reach a static equilibrium that is independent of the order in which reactions occurred. Prior work [2] shows how RICRNs can compactly implement binary-weight ReLU neural networks, allowing a pipeline from *in-silico* training of a neural network to compilation of an RICRN which computes the learned function, addressing the problem of how best to program CRNs. We show the resultant RICRNs have compact strand displacement implementations. Having been derived from RICRNs, the strand displacement cascades do not need precisely tuned rates and promise robustness to non-well-mixedness and other deviations from mass action, except with respect to leak and substoichiometric completion often seen in strand displacement. We give preliminary experimental validation of the strand displacement cascades, showing that the output closely matches that of the original binary-weight ReLU neural network. The tested network is a toy, two-neuron network computing a continuous generalization of XOR. Future work will test larger networks and robustness properties derived from the cascades' RICRN roots.



[1] Chen H, Doty D, and Soloveichik D. 2014. *ITCS.* 313-326.
[2] Vasić M., Chalk C, Luchsinger A, Khurshid S, and Soloveichik D. 2022. To appear in *PNAS*. Preprint available at https://arxiv.org/abs/2109.11422.

A minimal DNA origami for seeding tiled DNA nanotube bundles

Sarah Webster[1], and Deborah Kuchnir Fygenson[1,2]

[1]Physics Department and [2]Biomolecular Science & Engineering, University of California, Santa Barbara, CA, US 93106

Using DNA origami, a self-assembly method for building arbitrary and highly addressable structures out of DNA, we designed what we believe to be the shortest possible seed for nucleating tiled DNA nanotubes. This minimal seed, made with a 384 nucleotide scaffold, is 75% the size of the smallest seed described in the literature[1]. At this small size, we believe the seed can only support nucleation from one end. To enable nanotube nucleation from both ends, and to create additional sticky end positions to support seed assemblies, we used 586 and 768 nucleotide long scaffolds.

All of our seeds form with high yield and nucleate nanotubes equally well as the larger control seed (3024 nucleotide long scaffold). In principle, their small size makes them less costly to produce and easier to incorporate into hierarchical assemblies: smaller seeds should diffuse faster and therefore interact more often, allowing for faster assembly and higher yields. Additionally, the small number of staple strands required allows for inexpensive design of new hierarchical patterns using the same scaffold.

Here we explore strategies for building seed clusters for nucleating nanotube bundles using these minimal seed designs. By introducing 4 or 6 sticky ends (90 or 60 degrees apart from one another) around the outside perimeter of the seeds, we designed both square and hexagonally packed lattices of nucleation sites. We then demonstrated their nucleation of nanotube bundles from hexagonally packed lattices.

We further designed trimers, which consist of 3 seed monomers in the shape of a triangle, and tetramers, which consist of 4 seed monomers in the shape of a square.These were functionalized with only 2 sticky ends 60 and 90 degrees apart respectively.

[1] Jorgenson T, Mohammed A, Agrawal D, and Shulman R 2017 *ACS Nano* **11** 1927-1936

Efficient Elementary Step Implementations in Nucleic Acid Kinetics Simulators

Boyan Beronov, Jordan Lovrod, Chenwei Zhang, and Anne Condon[1]

[1]University of British Columbia

Abstract

Nucleic acid kinetics simulators such as Kinfold [1], Multistrand [2,3], or KFOLD [4] are useful for predicting reaction rates and designing DNA programs. Starting from an initial secondary structure, these simulators generate a stochastic trajectory of pseudoknot free structures in elementary steps, or moves, that add or remove a base pair. Fast probabilistic move generation that adheres to a biophysically-motivated kinetic model is important to overall simulator efficiency. Building on previous work [2,3,4], we propose an O(N)-time elementary move implementation, improving the worst-case runtime performance over previous work by a factor of N, where N is the total number of bases in the interacting strands. The implementation is O(log N) worst-case time for the Kawasaki model. Our implementation leverages standard assumptions about energy and kinetic models, and assumes that the number of interacting strands is O(1), and that multiloops have O(1) closing base pairs (i.e., branches).

Background and related work. Flamm et al.'s Kinfold uses O(N^2) time per elementary step. The Multistrand and KFOLD simulators achieve O(log N) worst case time per step to *sample* a move by incorporating new algorithms and data structures. While they also achieve O(log N) time to *update* the state (i.e., secondary structure) in typical cases—where not only the number of branches, but also the number of unpaired bases of a loop is O(1) —the time per step can still be $\mathbf{\Theta}(N^2)$ in the worst case, e.g., when a large internal loop with $\mathbf{\Theta}(N)$ unpaired bases is created.

Our contributions. We propose a method for sampling a move and updating the state that takes O(N) time even for large loops. We assume that the free energy of a secondary structure is the sum of loop energies, with the energy of a large loop being a function of the number of unpaired bases and the local contexts of closing base pairs. By local context of a base pair, we mean the identities of its bases and their immediate neighbors (if any). We leverage the fact that, for standard kinetic models (e.g., Metropolis, Kawasaki, or Arrhenius), the rate of a move from structure x to structure y can be expressed as a function of two arguments: the free energy difference $\Delta G(x) - \Delta G(y)$, and the local contexts of the closing base pairs which delimit the newly created and removed loops.

To convey key ideas of our approach, consider an elementary move that adds a base pair to a loop L. We partition all possible such moves into O(N) *types*. Moves of the same type have the same rate, because they have identical local contexts and the free energy of the resulting secondary structures are the same. In our representation of L, we store the total number of moves of each type, and their associated rate. We represent the state as a balanced tree, with loop representations at the leaves and internal nodes storing the sum of rates of descendant loops. With this representation, a move can be sampled in O(log N) time. Then the state is updated, removing loop L and adding the representations of two newly-created loops. Computing these representations takes O(N) time, using tables that are pre-computed in O(N^2) time which provide crucial information about move types in prefixes of the strand.

References.

[1] Flamm C et al 2000 *RNA* **6**(3):325-338
[2] Schaeffer et al 2015 DNA 21:194-211
[3] Schaeffer PhD thesis. Caltech 2012
[4] Dykeman 2015 *NAR* **43**(12):5708-5715

Scaling up reusable DNA circuits using heat as a universal energy source

Tianqi Song[1] and Lulu Qian[1,2]
[1] *Bioengineering,* [2] *Computer Science*
California Institute of Technology, Pasadena, CA 91125, USA

State-of-the-art enzyme-free nucleic-acid circuits cannot remain active and respond to changing input signals – they reach equilibrium states when some molecular components are used up. How can these circuits stay active and process time-varying information in a molecular environment? Several types of chemical power supplies have been explored to reinvigorate the system, but none of these approaches appear to be scalable. Inspired by recent examples of recycling DNA circuits by heat, and building on the observation of nucleic-acid strands with strong secondary structures can reach kinetically trapped states during thermal annealing, here we develop a general-purpose circuit architecture where all molecular components can be recharged (restored to their initial states) using heat as a universal energy source after the input is inactivated. We show that an exceptionally simple motif can be designed to respond to a quick temperature ramp and hundreds of distinct molecules composing a complex DNA neural network can spontaneously reconfigure to their original structures within minutes, allowing at least ten rounds of computation with various input signals. Our approach ensures that arbitrarily complex circuits can be powered by the same energy source, no debugging is necessary for the power supply itself, and the circuit performance remains roughly constant over time. This work introduces a scalable approach for enabling sustained operation of enzyme-free molecular circuits, and establishes the foundation for future demonstrations of unsupervised learning in chemical systems.

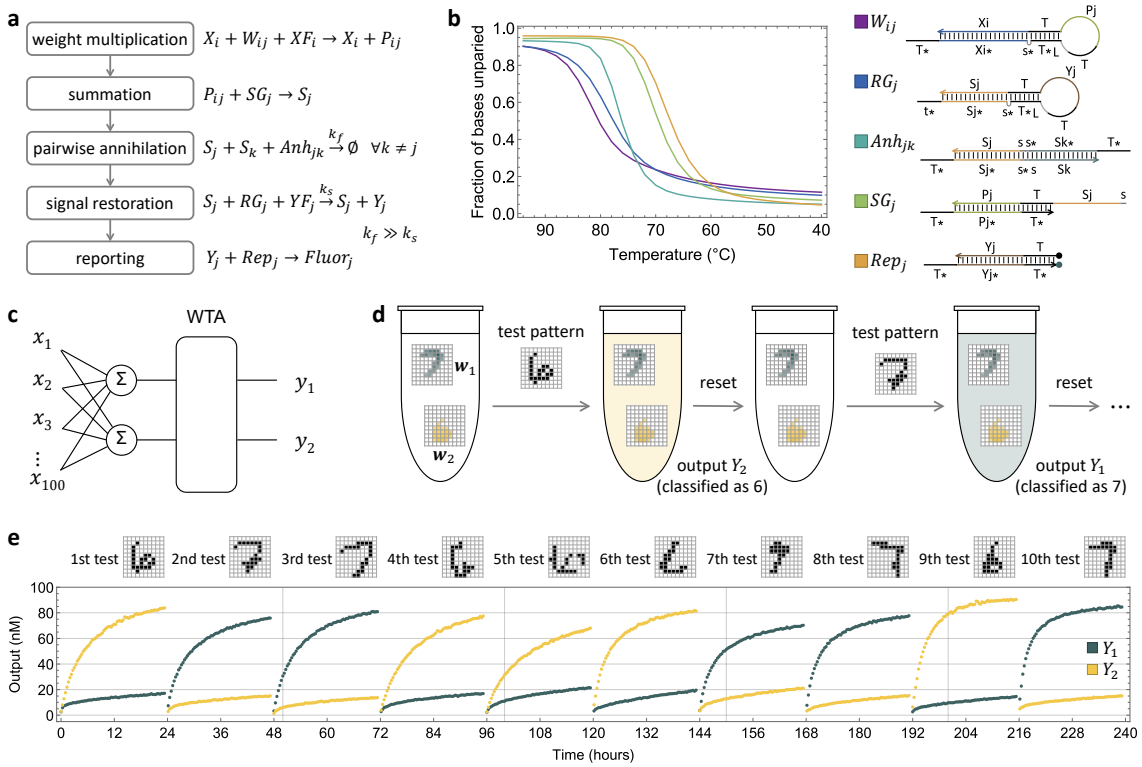


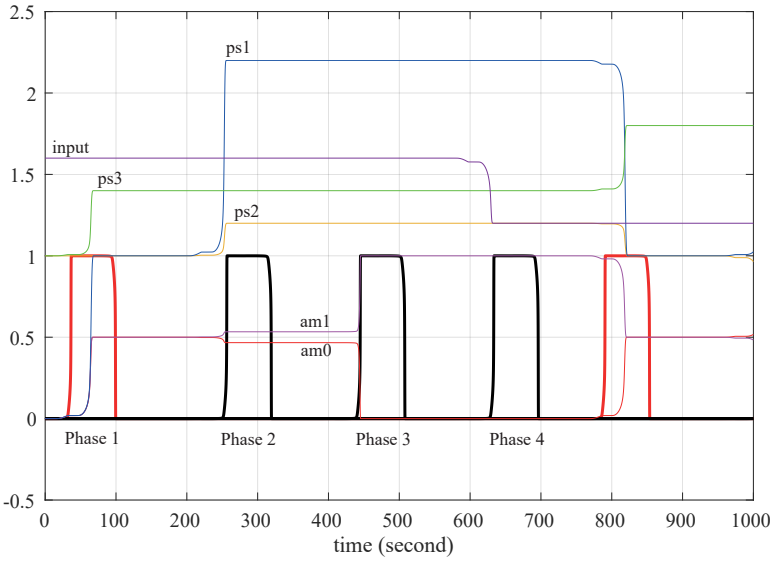
Figure 1: **A reusable DNA neural network. a,** Chemical reaction network implementation. **b,** DNA strand-displacement implementation with a melting profile that allows for the system to reach a kinetically trapped state during thermal annealing, which resets the circuit to its initial state after the input is inactivated. **c,** Circuit diagram of a 100-bit, 2-memory winner-take-all (WTA) neural network. **d,** Steps in testing the reusability of the circuit. **e,** Fluorescence kinetics data of classifying ten MNIST patterns sequentially added to the same test tube. Each reset was performed by adding complementary input strands to inactivate the previous test pattern and heating up to 95 °C for 5 minutes and cooling down to 20 °C in 1 minute.

Transformation of Real-Valued Concentrations into Binary Signals*

James I. Lathrop, Dawn A. Nye, Hugh D. Potter
Iowa State Univeristy

Abstract

Chemical Reaction Networks (CRNs) are a rich and well-studied model for molecular computation. We have seen many interesting developments in deterministic CRNs in two areas in particular: emulating discrete calculations and computing and analyzing analog functions. The interplay between these two paradigms provides promising opportunities of both practical and theoretical interest. We consider the problem of designing a CRN to transform an input chemical signal into its binary encoding; the output bitstring is represented as a time series of high and low pulses in a designated output species. Several impossibility theorems demonstrate unavoidable limitations in performing this transformation, including the inability to output the exact bits of an arbitrary input concentration. Additionally, we provide a CRN capable of solving a restricted version of the analog-to-digital problem, and an example simulation of a single output bit is shown below.



- **input.** This species is the input to the analog-to-digital converter in the range of $[1, 2)$. Note that only the fractional part of the input is converted. Also note that the input is destroyed as it is converted in this CRN. It is trivial to copy an approximation of the input value to another species if the concentration of the input must be retained.
- **am0 and am1.** $am0$ and $am1$ form an approximate majority computation that determines if a signal is above or below 0.5.
- **ps1.** This species is the value the input should be set to if the input is less than 0.5.
- **ps2.** This species is the value the input should be set to if the input is greater than 0.5.
- **ps3.** This species computes $3 - input$ for use in setting the starting values of $am0$ and $am1$.

*This research supported in part by National Science Foundation Grant 1900716.

Qslib: Python control of qPCR machines for molecular programming experiments*

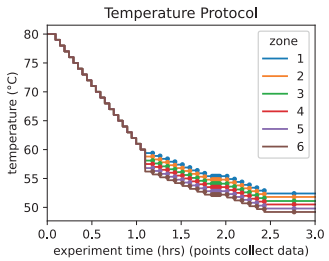
Constantine Evans†

Real-time quantitative PCR (qPCR) machines are well-suited for fluorescence measurements of molecular programming systems: they can precisely and programmably control the temperature of many samples while taking multiple-color, real-time fluorescence readings, all in parallel. However, their software is often narrowly focused on qPCR and other specific uses, making some potential molecular programming applications of the hardware less convenient or, in some cases, infeasible.

Qslib (https://github.com/cgevans/qslib) is a Python library, used for three upcoming presentations at DNA28[1], that adapts the QuantStudio 5 (Applied Biosystems) qPCR machine for more flexible and convenient use as a general-purpose programmable temperature control and fluorescence measurement device. As a Python library rather than a GUI, qslib is designed both to allow everyday use through Jupyter notebooks, and to be readily extended for more complex applications. Protocols are specified as code, which allows for precise descriptions, simple changes, and protocols that would be time-consuming or impossible to program in the machine's usual software. Fluorescence, temperature, and timing data beyond that normally available on the machine can be accessed and conveniently plotted.

During runs, qslib supports extensive remote monitoring, with real-time exports to time-series databases and dashboard software, mid-run data access, and complete remote control, allowing runs to be started, paused, and ended, and allowing quick physical access to samples for mid-run additions or removals. It also allows arbitrary mid-run changes to temperature schedules and data collection parameters, impossible with the software for many qPCR machines: users, for example, can watch fluorescence and adjust temperatures to quickly find where particular dynamics begin, or can adjust for monomer depletion as growth of a self-assembling system slows. This control can even be made automatic, with temperatures controlled by feedback from real-time fluorescence data.

(a)
```
Stage.hold_at("80°C", "5 minutes"),
Stage.stepped_ramp(from_temperature="80°C",
                   to_temperature="60°C",
                   total_time="20 minutes",
                   n_steps=20,
                   collect=True)
```



(b)

```
exp.sync_from_machine()
exp.pause_now()
exp.machine.drawer_open()
# Take samples for imaging.
exp.machine.drawer_close()
exp.resume()
exp.change_protocol_from_now(
    [Stage.hold_at("50°C", "72 hours",
    "10 minutes", collect=True)])
```

(c)

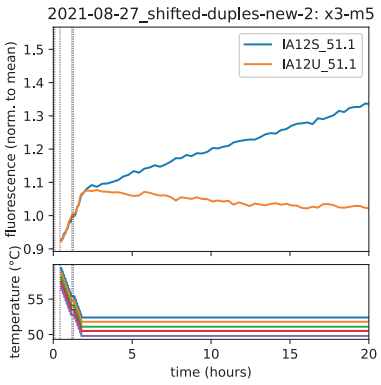


Figure 1: (a) Protocols are specified as code and can be easily plotted, as for the temperature ramp with six offset temperatures for different plate columns shown here. (b) Real-time data can be sent to open source dashboard software, such as Grafana, shown here. During runs, data can be imported and plotted, samples can be removed for imaging, and temperature schedules can be changed arbitrarily: for example, to quickly find a growth temperature in exploratory experiments. (c) Qslib offers convenient plotting of fluorescence data, and one-second-resolution temperature data, unavailable in the manufacturer's software. Here, a seeded DNA tile ribbon with a FRET pair probe grows at 51.1°C (blue), while an unseeded ribbon does not (orange); similar samples at six temperatures were run in parallel on the machine.

*Supported by ERC grant 772766, Active-DNA, and SFI grants 18/ERCS/5746 and 20/FFP-P/8843. †Hamilton Institute, Department of Computer Science, Maynooth University, Ireland[1] C. Evans et al, DNA 28, 2022; T. Rogers et al, DNA 28, 2022; T. Stérin and A. Eshra et al, DNA 28, 2022.

Compact and Efficient Chemical Boltzmann Machines

Inhoo Lee[1], William Poole[2], and Erik Winfree[1,2,3⋆]
[1] *Computer Science,* [2] *Computation & Neural Systems, and* [3] *Bioengineering*
California Institute of Technology, Pasadena, CA 91125, USA
⋆ *e-mail: winfree@caltech.edu*

Abstract. The classical Boltzmann machine (BM) is an energy-based generative model for machine learning [1] with a surprisingly natural learning rule. Given weights $w_{i,j}$ and thresholds θ_i, asynchronous probabilistic updates at temperature T set $x_i \leftarrow 1$ with probability $p_i = 1/(1+\exp(a_i/T))$ where $a_i = \sum_j w_{i,j}x_j - \theta_i$ is the accumulated input to binary neuron i. At steady-state, $Prob(x) = \frac{1}{Z}\exp(-E(x)/T)$ where energy $E(x) = -\frac{1}{2}\sum_{i,j} w_{i,j}x_ix_j + \sum_i \theta_i x_i$ and partition function $Z = \sum_x \exp(-E(x)/T)$. Given partial knowledge of a subset of variables, inference of the conditional probability distribution is generated by clamping the known neurons while letting the unclamped neurons update and reach steady state. Surprisingly, stochastic chemical reaction networks (CRNs) can exactly simulate Boltzmann machines [2] and in fact detailed balanced CRNs can be viewed as a generalization of Boltzmann machines to "neurons" with integer count states [3]. Indeed, the Boltzmann machine learning rule, which adjust weights according to clamped vs free-running averages, translates perfectly to adjustments of species energies in detailed balanced CRNs. However, in those works, CRN constructions that exactly matched the Boltzmann machine distributions, and thus had the capacity to represent complex probabilistic information, required $O(N2^d)$ reactions where N is the number of neurons and d is their degree, i.e., how many other neurons they receive input from. Thus large fully-connected networks are not feasible using existing techniques. Here, we demonstrate a non-detailed-balanced construction that uses just $O(N^2)$ reactions yet accurately matches Boltzmann machine distributions. The CRN has binary dual-rail species X_i^0 and X_i^1 representing neuron states, accumulator species A_i^0 and A_i^1 whose counts represent their summed input, and input species C_i^0 and C_i^1 are used to clamp units to known values during inference; these species may also be roughly interpreted as a genetic regulatory network whose genes X_i^b produce proteins A_i^b that regulate gene activity.

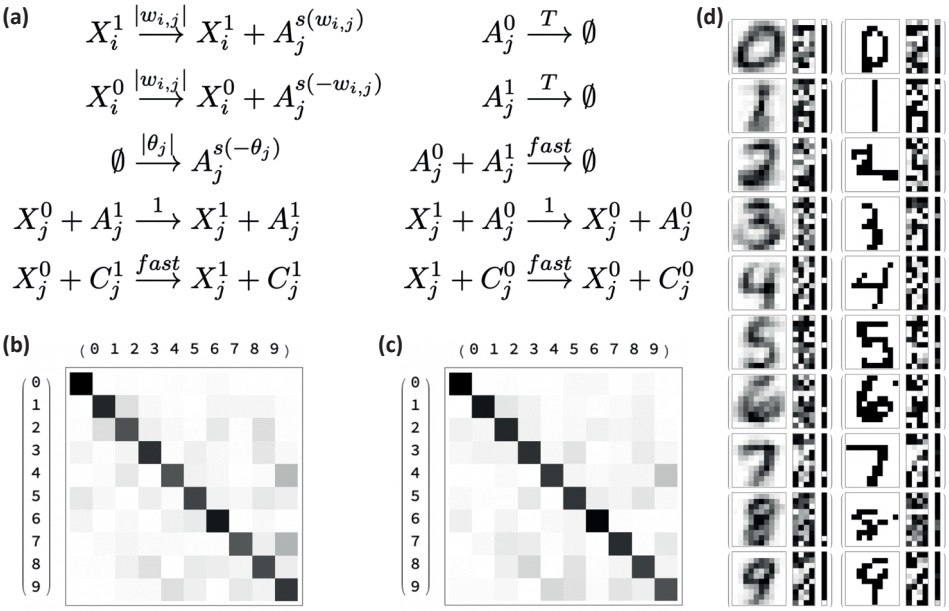


Figure 1: **The Titration Chemical Boltzmann Machine on** 10×10 **MNIST images. (a)** Reactions have rate constants that depend on the BM weights and thresholds, with products possibly depending on the sign of the parameter where $s(v) = 1$ if $v > 0$ and 0 otherwise. **(b)** The confusion matrix for a trained classical BM where rows indicate the true class and columns indicate the inferred class. 100 image units, 40 hidden units, 10 class units. **(c)** The confusion matrix for the corresponding Titration CBM. **(d)** Titration CBM state averages when clamping class units (left) and image units (right).

[1] Geoffrey E Hinton, Terrence J Sejnowski, and David H Ackley. Boltzmann machines: Constraint satisfaction networks that learn. Technical report, Carnegie-Mellon University, Department of Computer Science, Pittsburgh, PA, 1984.

[2] William Poole, Andrés Ortiz-Muñoz, Abhishek Behera, Nick S Jones, Thomas E Ouldridge, Erik Winfree, and Manoj Gopalkrishnan. Chemical Boltzmann machines. In *DNA Computing and Molecular Programming (Lecture Notes in Computer Science)*, volume 10467, pages 210–231. Springer, 2017.

[3] William Poole, Thomas Ouldridge, Manoj Gopalkrishnan, and Erik Winfree. Detailed balanced chemical reaction networks as generalized Boltzmann machines. *arXiv preprint arXiv:2205.06313*, 2022.

Highly Parallel Synthesis, Computation, and Readout of DNA Logic Gates using High-Throughput Synthesis and Sequencing

Chandler Petersen[1], Samantha Borje[1], Gourab Chatterjee[1], Yuan-Jyue Chen[1,2], and Georg Seelig[1]

[1]University of Washington, USA, [2]Microsoft Research, Seattle, WA

DNA computing's strength lies in its ability to process information while still in molecular form, allowing computation and programmable control of biological matter at the nanoscale. Implementations of Boolean logic gates [4, 5] have been particularly promising for a variety of applications. However, full realization of their potential will require overcoming existing limitations of molecular logic circuits. The size of current DNA logic circuits is limited for several reasons. First, the cost of column-synthesized DNA oligonucleotides is expensive. Second, DNA gates typically require purification to remove errors in stoichiometry, but this is very time-consuming and not multiplexable, causing significant challenge to scale up. Finally, the reading bandwidth of conventional fluorescence reporter readout is limited due to spectral overlaps. To address these limitations, here we present multiplexable methods that enables highly parallel preparation and reading of nicked double-stranded DNA (ndsDNA) gates [2, 3] by using array-based DNA synthesis and next-generation sequencing (NGS). Each ndsDNA gate consists of a join and fork gate. For AND logic, the join gate accepts two input signals, which triggers a strand displacement cascade resulting in an intermediate private co-signal to the fork gate, producing a final output signal with the same domain structure [2, 3]. Additionally, each fork gate performs catalytic amplification of the private co-signal, providing signal restoration local to each gate [6]. This modular design pairs computation with amplification to address signal degradation inherent to multi-layered circuits. Each gate contains flanking "barcode" regions that are used for selective PCR and sequencing of reacted gates. With these modifications, complex circuits with potentially thousands of gates may be read out in a single sequencing reaction [1]. This is the first demonstration of large-scale synthesis and sequencing of DNA gates, with the potential to realize more complex circuits and programmable logic on a scale not yet seen.

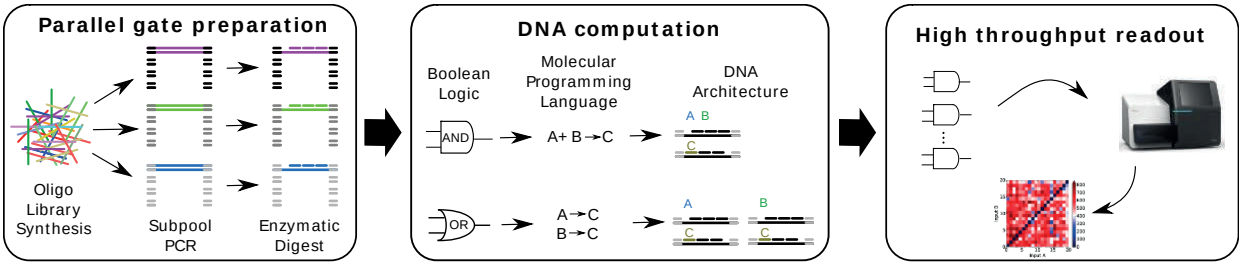


Figure 1. Overview of circuit synthesis, computation, and read out. Left: Gates are derived from an array synthesized DNA pool, selectively amplified with PCR, and enzymatically processed to create ndsDNA gates. Middle: Computation occurs in pooled reactions upon addition of inputs. Right: End point results of Boolean logic computation are read out in parallel using NGS.

[1] Cardelli L 2021 *arXiv*
[2] Cardelli L 2010 *EPTCS* **26** pp. 47-61
[3] Chen Y-J et al 2013 *Nat. Nanotechnol* **8** 755-762
[4] Seelig G et al 2006 *Science* **314** pp. 1585-1588
[5] Qian L and Winfree E 2011 *Science* **332** pp. 1196-1201
[6] Zhang D et al 2007 *Science* **318** pp. 1121-1125

Growing arbitrary patterns with neural reaction-diffusion

Salvador Buse[3] and Erik Winfree[1,2,3⋆]
[1] *Computer Science,* [2] *Computation & Neural Systems, and* [3] *Bioengineering*
California Institute of Technology, Pasadena, CA 91125, USA
⋆ *e-mail: winfree@caltech.edu*

Abstract. Since the seminal work of Turing, reaction-diffusion has been a mechanism thought to underlie some biological pattern formation. Molecular programming technology has advanced to the point where we can employ these principles for the synthesis of structured materials, rather than just as a way to explain pattern formation in nature. In particular, DNA strand displacement cascades can be designed to implement arbitrary chemical reaction network (CRN) dynamics, including in the reaction-diffusion regime. Earlier theoretical work showed how to form an arbitrary complex spatial pattern starting from simple initial conditions by first forming a coordinate space and then computing a function that indicates the pattern's color [1]. This has been extended such that the function's computation is implemented by a CRN neural network that can be trained in silico [2]. Here we demonstrate an alternative approach that does not involve an explicit coordinate grid, and which makes use of fully recurrent CRN neural networks rather than a feedforward architecture. Our inspiration comes from neural cellular automata [3], which uses differentiable programming on a model similar to reaction-diffusion neural networks to achieve robustly self-healing pattern formation. We adapt their techniques to construct explicit reaction-diffusion neural network models based on the components and reactions developed in prior experimental work [4].

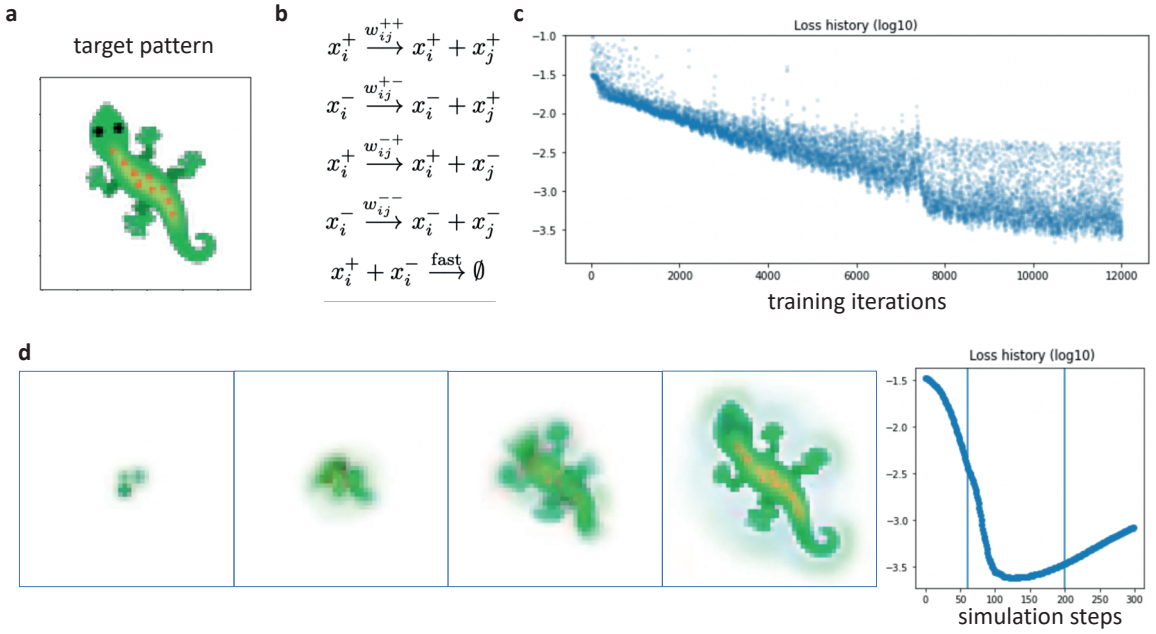


Figure 1: **Differentiable programming of reaction-diffusion pattern formation. a,** A target image in RGB-α, of 60×60 pixels. **b,** The CRN architecture. Indices range from 1 to N, so there are $2N$ species. **c,** Training error during optimization. **d,** Snapshots from a growth movie using a fully-trained network, with the loss function computed at each simulation step.

[1] Dominic Scalise and Rebecca Schulman. Designing modular reaction-diffusion programs for complex pattern formation. *Technology*, 2:55–66, 2014.

[2] Marko Vasić, Cameron Chalk, Austin Luchsinger, Sarfraz Khurshid, and David Soloveichik. Programming and training rate-independent chemical reaction networks. *Proceedings of the National Academy of Sciences*, 119:e2111552119, 2022.

[3] Alexander Mordvintsev, Ettore Randazzo, Eyvind Niklasson, and Michael Levin. Growing neural cellular automata. *Distill*, 2020. DOI: 10.23915/distill.00023.

[4] Kevin Cherry and Lulu Qian. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature*, 559:370–376, 2018.

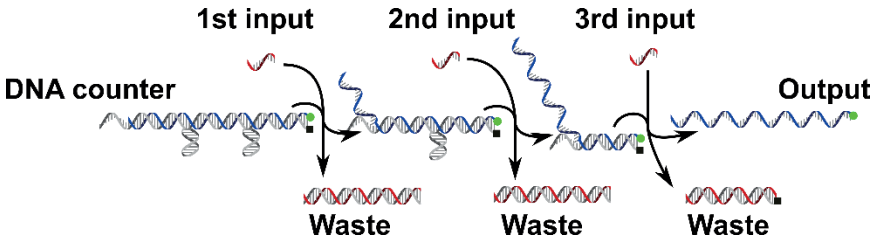
Counting the number of input addition by DNA reaction system driven by DNA polymerase

Ibuki Kawamata[1,2], Motokazu Furuya[1] and Satoshi Murata[1]

[1]Tohoku University, Japan, [2]Ochanomizu University Japan

By programming the interactions among enzymes and DNA strands with carefully designed base sequences, it is possible to demonstrate dynamical functionalities such as oscillation [1], bistable circuit [2], and sequential strand generation [3]. Here, we propose a system composed of a DNA complex and DNA polymerase, which can undergo state transitions upon the addition of input DNAs with identical base sequence (Fig. 1a). For example, the output DNA is released from the system only after three times of input additions with sufficient time intervals (Fig. 1b). Since a similar mechanism takes place in each round of state transition, it is possible to change the number by changing the size of initial DNA complex. To function the system as expected, it is important to implement a mechanism to detect only the number of input addition (*e. g.* 3 times of single amount of input *vs* 1 times of triple amount of input). We hope that the proposed system can be used as a counter to record the number of molecular events. In the poster presentation, the latest results will be discussed.

(a)



(b)

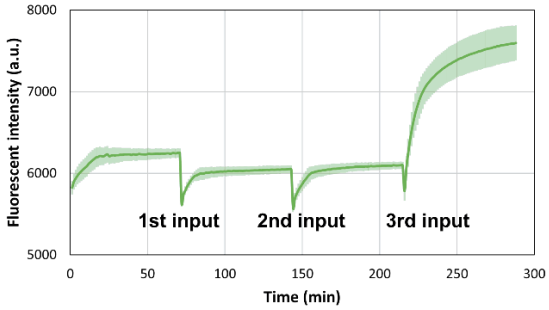


Figure 1. (a) Schematics of the reactions of DNA counter. The DNA counter changes its state upon 1st, 2nd, and 3rd input DNAs, all of which have identical base sequences. All reactions are driven by a strand-displacement activity of DNA polymerase. Only after the 3rd input, the output DNA is released as single-stranded and becomes fluorescent. (b) Preliminary experimental result of the DNA counter.

[1] T. Fujii and Y. Rondelez, 2013, *ACS Nano*, **7**, 27
[2] S. W. Schffter and R. Schulman, 2018, *Nat. Chem.*, **11**, 829
[3] I. Kawamata, S.-i. M. Nomura and S. Murata, 2022, *New Gener. Comput.*

Facet TAM: An abstract tile assembly model with facet growth*

Ahmed Shalaby Constantine Evans Damien Woods

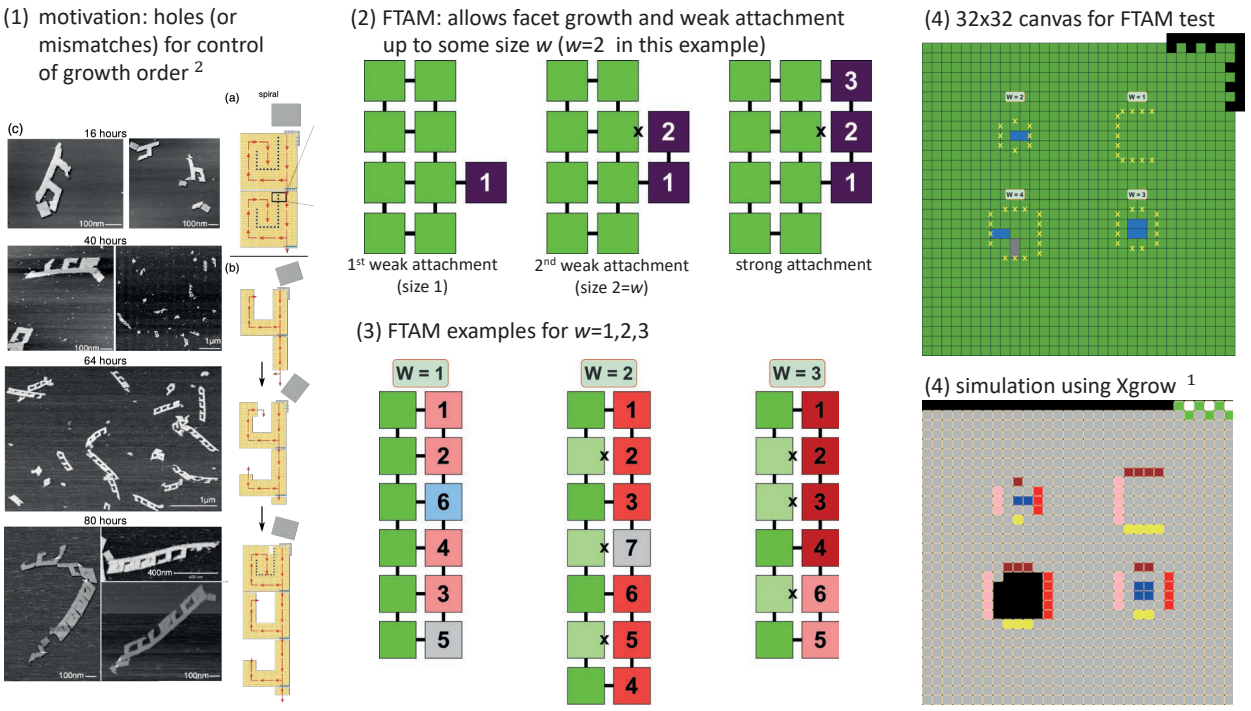
Hamilton Institute, Department of Computer Science, Maynooth University, Ireland

Introduction. The abstract and kinetic Tile Assembly Models (aTAM and kTAM)[1] are molecular programming models intended to guide the design and implementation of DNA tile-based algorithmic self-assembly systems. The aTAM is for programming and theoretical analyses, and the kTAM facilitates stochastic simulation of candidate experimental DNA tile systems.

Our recent work on DNA self-assembly[2] exploits kTAM-feasible phenomena, specifically facet growth and weak attachment of subassemblies, that is not permitted in the aTAM. The kTAM expresses these features but in a complex model more suited to simulation than programming and theoretical analysis, in particular due to complexities around reversible/unbinding steps. Similar phenomena are also studied in papers on snaked proofreading[3] and size-dependent self-assembly[4]. Here, we propose the Facet Tile Assembly Model, or FTAM.

FTAM. We augment the aTAM with an additional parameter: $w \in \mathbb{N}$ that denotes the maximum permissible size of any *weakly bound subassembly* (WBS), defined as follows using aTAM terminology: Given an assembly α and a temperature $\tau \in \mathbb{N}$, if α has a glue-cut of weight (sum of glue strengths) $< \tau$ where the subassembly α_{stable} on one side of the cut is τ-stable and the subassembly α_{WBS} on the other side is of size $\leq w$, then we say that α_{WBS} is a WBS. The WBS α_{WBS} may or may not be stable, but its attachment to α_{stable} is not stable (weak).

The simplest version of our model permits at most one WBS and uses (aTAM) temperature 2, although we are also considering natural generalisations that allow an assembly to have multiple, parallel-growth, WBSs. We will use the FTAM to design DNA self-assembly experiments that probe the capabilities of controlled facet-based growth.



*Research Supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement grant # 772766, Active-DNA), and Science Foundation Ireland (SFI) under grant numbers 18/ERCS/5746 and 20/FFP-P/8843. [1] Winfree. PhD thesis. 1999 [2] Evans, Doty, Woods. DNA28 Track B. 2022 [3] Chen, Goel. Proc. DNA10. 2004; and Chen, Schulman, Goel, Winfree. Nano Letters. 2019 [4] Fekete, Schweller, Winslow. ISAAC 2015

Formal Semantics for Stochastic Chemical Reaction Networks

Andrés Ortiz-Muñoz

Santa Fe Institute

April 29, 2022

Abstract

We define the semantics of stochastic chemical reaction networks (SCRNs) in terms of formal power series, also known as generating functions. Under this formulation the familiar chemical master (CME) equation becomes a formal partial differential equation on the probability generating function of a SCRN. We focus on this generating function formulation of the CME in order to derive general expressions for the stochastic dynamics of the system. For that purpose we define a class of *regular* solutions to the CME as ones where the probability of infinite paths vanishes at all finite times. Equivalently, regular solutions to the CME are those that can be expressed as the exponential of an infinitesimal stochastic operator. The formal structure of these solutions reflects the combinatorial nature of trajectories in state space. The remaining material focuses on stationary solutions to the CME, which are time-independent. We show that complex-balanced SCRNs are precisely those that admit an exponential power series as a stationary solution. Interestingly, combinatorially, exponential power series are generating functions of unordered sets. We define factorial moments and their generating function and derive a simple relationship between the factorial moment generating function and the probability generating function. We speculate about the combinatorial interpretation of the relationship between probability and factorial moment generating functions. Finally, we define assembly systems, which are complex-balanced SCRNs where each species has a composition, and for which all reactions preserve composition. We derive expressions for the factorial moments of assembly systems.

This the third chapter of my Ph.D. thesis, which can be accessed through this link.