Localized Catalytic DNA Circuits for Integrated Information Processing in Molecular Machines

Thesis by
Samuel Ryan Davidson

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy, Bioengineering

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2025
Defended October 25th, 2024

ACKNOWLEDGEMENTS

My advisor, Lulu Qian, has consistently supported me as a graduate student in every conceivable manner. She has met with me weekly or biweekly to discuss my research and provide insights. She cultivated a healthy and happy lab culture, providing pleasant spaces and times to congregate as a community. She took seriously the department's initiative to ask each lab to appoint a Diversity, Equity, and Inclusion ambassador who should present on a DEI-related topic quarterly (a role that initially challenged me and ultimately strengthened me).

Lulu (and Erik) are not shy about their lack of interest in participating in the biotechnology industry through marketable applications or startup spin-offs. Yet, Lulu has encouraged my interest in an industry career by happily allowing my participation in Frances Arnold's Biotechnology Leadership Program, as well as the full-time industrial internship I pursued which interrupted my PhD progress from March to July of 2020—although surely I would not have made much experimental progress since I would not have been able to physically go to campus during this time per Caltech's response to the pandemic.

Speaking of which, I must take immunosuppressants because I have an autoimmune disease (Crohn's), so I am at all times moderately immunocompromised. When pre-vaccine COVID-19 could have threatened my life, Lulu ensured the safety of all lab members by providing resources and regulations that prevented disease transmission. Throughout the post-vaccine era, Lulu has continued to support me in all of my personal decisions to mitigate risk.

Lulu is adept in all her roles, as a scientist, a teacher, and a mentor. If occasionally she fails to meet her own standards for herself, it is only because they are impossibly high for any human being. But Lulu's humanity filled my graduate student experience with light, so I implore her to celebrate any[1] ways in which she is uniquely herself.

Erik Winfree supported my doctoral work primarily through the joint group meetings where I regularly heard his insights and feedback. Ever-ready to challenge any presenter with questions about claims in their slides that they had not fully considered, Erik inspired me to think more carefully about my assumptions and stay true to the skeptical values that any good scientist ought to have. Furthermore, because Erik challenges presenters out of love for science and the pursuit of truth, he sets a standard of supportiveness through constructive criticism that I see reflected in the way our lab community members treat each other. I never feel like I'm being attacked if a fellow lab member points out something I've said or done wasn't quite as scientifically rigorous as it should have been; to the contrary, I know that I'm being supported to be my best self.

Much of my practical knowledge about how to accomplish my lab goals and how to succeed as a graduate student came from Kevin Cherry and Namita Sarraf, both of whom provided excellent grounded, realistic perspectives right when I needed to hear such things.

I have at least one more thing to thank Namita for. Since 2020, volunteers from throughout the field of molecular programming have collaborated to slowly write and edit The Art of Molecular Programming textbook. As lead editor for the circuits section, I witnessed how this project helped to unite our field and build a virtual international community. I'm eternally grateful to Namita and Dominic Scalise (another former labmate) for initiating this project, and to all the members of the circuits team for their amazing work. While producing the textbook itself has taken many years, I am glad that I had the opportunity to quickly pick the brains of a wide subset of the field at any time by simply pinging the general channel in our Slack workspace. I learned a great deal from my collaborators there. Because of this connectedness, I was able to get feedback from Josie Kishi and Nikhil Gopalkrishnan on the section of this thesis pertaining to three-letter code scaffolds. Thanks to you both as well.

It was a big help to have the feedback of Spencer Winter, Matthew Plazola, and Martin Holmes on the writing found past this acknowledgements section. I couldn't ask for better people to carry on the proud traditions I have devised and enforced in recent years.

Producing over 280 different DNA origami samples would have been a literal pain in the neck[2] if it hadn't been for the support of Miki Yun, Richard Murray, and all the members of the Murray Lab who regularly maintain the Echo acoustic liquid handling robot and make it available to many other labs at Caltech. Their generosity and spirit of collaboration is felt and appreciated.

The hero of Keck, the building that houses the Qian Lab, is Christy Jenstad. She is willing to go above and beyond to keep everything functioning smoothly so that we rarely are forced to stop working by any situation pertaining to fires, floods, electrical issues, or pipes leaking, all of which seems (from my perspective) to be the fault of those pesky labs upstairs. What are they working on up there? It had better be important! Christy would know, since she's somehow friends with everyone in the building due to being the friendliest person at Caltech.

I would be remiss if I failed to express any gratitude toward:

- each and every teaching assistant who helped take a load off of me when I was lead TA of BE/CS 196a
- Pamela Bjorkman, who supported my work as lab DEI ambassador and shared my slides with other BBE DEI representatives
- Richard Murray, who supported my initiative to connect genderqueer folks at Caltech with off-campus resources
- Andrey Shur, who was the first person to train me at Caltech and also uploaded to a public repository his designs for 3D-printed tube racks which space tubes to fit with every other multichannel tip
- TechHub Assistant Arleen Hom, who enabled me to print my modified versions of Andrey's designs
- Christopher T. Yeh, a fellow graduate student who added his voice to support my request for bike-friendly speed bumps, which I now happily bike through on my way to and from lab
- John S. Onderdonk, Assistant Vice President for Facilities Operations and Services, by whose authority those bike-friendly speed bumps were constructed
- ChatGPT, which suggested some reasonable improvements to a few passages in this thesis

Finally, I thank my friends and family, all of whom played a very large role in my sanity staying largely intact throughout this process. In particular, I thank my friend Marissa Padilla for being the funniest person I’ve ever met and insisting that I stop working now and then, and my partner Pumpkin Jackson, who is incredibly supportive and patient, and who proves that the best way to survive any difficulty is to remain true to one’s self.

ABSTRACT

This thesis supports the long-term goal of engineering molecular devices with computational complexity akin to cells. Like cells, artificial molecular devices can benefit from integrating multiple computational modalities.

To that end, this thesis advances molecular computing systems in three modalities: dynamic molecular assembly, well-mixed circuits, and spatially-organized cascades. Specifically, it introduces methods to enhance control over DNA structural assembly, well-mixed DNA circuits, and DNA circuits localized to a DNA origami surface.

As DNA structural assembly grows increasingly complex, so too grows the potential for off-target structures. This issue can be addressed through developmental self-assembly, where components join a growing structure in a programmed sequence under controlled kinetics. The scope of developmental self-assembly is here expanded by a method enabling specific pathway selection among multiple encoded options.

Well-mixed DNA circuits require catalytic motifs for signal restoration and amplification. A catalytic motif is presented where two input strands cooperate to control catalysis. This motif could enhance AND gates and thresholding, and could enable adaptive memories and learning behaviors in DNA-based neural networks.

Localized DNA circuits lack cascadable catalytic mechanisms for signal restoration and amplification. Two designs for a localized catalytic mechanism are presented. Each omits any intermediate diffusible species to support nanodevices compatible with uncontrolled environments, as in biomedical contexts. This constraint leads to design lessons; principally, we respond to leak in the first design through geometric constraints in the second design.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Glynn, A. T., Davidson, S. R. & Qian, L. Developmental Self-Assembly of a DNA Ring with Stimulus-Responsive Size and Growth Direction. *Journal of the American Chemical Society* **144,** 10075–10079. DOI: 10.1021/jacs.2c03853 (June 8, 2022).
2. Taylor, D. N., Davidson, S. R. & Qian, L. A Cooperative DNA Catalyst. *Journal of the American Chemical Society* **143,** 15567–15571. DOI: 10.1021/jacs.1c07122 (Sept. 29, 2021).

- **[1]**

A.T.G. and S.R.D. contributed equally. A.T.G. and L.Q. developed the original concept. A.T.G. and S.R.D. collaborated to design experiments. S.R.D. ran all experiments while undergraduates were not allowed to come in person to lab during the COVID19 pandemic and refined experimental methods. A.T.G. ran experiments to complete the project and to address reviewer comments. All three authors met weekly to analyze data and adapt experimental plans accordingly. All three authors wrote the manuscript.

- **[2]**

D.N.T. and S.R.D. contributed equally. D.N.T. and L.Q. developed the original concept. D.N.T. and S.R.D. collaborated to design experiments. S.R.D. ran all experiments while undergraduates were not allowed to come in person to lab during the COVID19 pandemic and refined experimental methods. D.N.T. ran experiments to complete the project and to address reviewer comments. All three authors met weekly to analyze data and adapt experimental plans accordingly. All three authors wrote the manuscript.

UNPUBLISHED CONTENT AND CONTRIBUTIONS

Work in Chapters 4 and 5 is not yet submitted for publication. Potential authors would include Samuel R. Davidson, Spencer Winter, and Lulu Qian.

L.Q. developed the original concept for the “pins-and-needles” design, and S.R.D. developed the original concept for the “latch” design. All experiments were designed by S.R.D., with the exception of the experiment to detect dangling staples, which was designed in collaboration between S.R.D. and S.W. All reactions were run and analyzed by S.R.D. Result interpretation took place weekly as a collaboration between L.Q. and S.R.D. A conference submission (unpublished) served as the original draft of Chapters 4 and 5, and was written collaboratively by L.Q. and S.R.D. The version included here was modified and expanded by S.R.D. All appendices were written by S.R.D.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Number *Page*

LIST OF TABLES

Number *Page*

NOMENCLATURE

Connectivity. A property that describes whether two circuit components are "wired together," i.e., designed to interact.

Leak. The spurious production of a signal molecule under unintended conditions, typically through the interaction of two or more species which are not intended to interact.

Reachability. A property of components in spatially-organized DNA molecular circuits that plays a role in determining their connectivity—whether the complementary binding sites on the two components can physically reach each other.

Spatially-organized. Describes molecules which are arranged on a structure at functionally meaningful locations.

Tether. An inert domain in a localized component which serves a geometric purpose, as opposed to a reactive purpose.

Well-mixed. Describes molecules that are roughly evenly distributed throughout a solution; their concentration is not related to their position.

Chapter 1

INTRODUCTION

The long-term vision guiding the work presented in this thesis is to engineer molecular devices with computational complexity comparable to cells. Living cells are autonomous molecular machines, naturally occurring nanorobots, which process information in many different ways depending on context. Construction of artificial autonomous molecular machines, human-made nanorobots, might be achieved most readily by developing rationally-designed systems which emulate the molecular computational modalities underlying living cells.

Specifically, this thesis presents work aiming to enhance our ability to build molecular computing systems within three computational modalities: dynamic molecular assembly, well-mixed circuits, and spatially-organized cascades. These modalities can be quickly understood through biological examples.

Some cells, such as those of the *Amoeba* genus, can quickly transform the shape of their bodies. Despite appearing to move fluidly, their shape is at all times maintained by a rigid cytoskeleton. Fluid-like motion is possible for cells because they can grow and shrink their cytoskeleton's structural components in a matter of seconds [3]. Cells perform these actions in response to stimuli to affect their environments. The underlying mechanisms are carefully controlled dynamic molecular events that assemble large structures from many smaller molecules. This computational modality, dynamic molecular assembly, is the first computational modality advanced in this thesis (Chapter 2).

The second computational modality advanced in this thesis (Chapter 3) is the well-mixed molecular circuit. Many such systems exist inside cells; for example, inside a nucleus, or inside a bacterial cell, many genetic regulatory molecules freely diffuse and interact. These well-mixed components form genetic regulatory networks, performing computation by suppressing or promoting gene expression.

The third and final computational modality advanced in this thesis (Chapters 4 and 5) is the spatially-organized cascade. Molecular spatial organization in cells takes diverse forms; components which interact beneficially are often co-localized inside organelles or within a phase-separated region of the cytoplasm. The type of spatially-organized cascade advanced in this thesis is most analogous to scaffolded signal pathways. Cells produce scaffold proteins to organize many smaller elements so that they can perform computation more quickly than can take place through free diffusion. These systems can rapidly amplify or inhibit a signal depending on the context.

Methods have been established and developed to rationally construct molecular systems from DNA which function within each of the three computational modalities discussed above. To understand how and why we advanced each of these types of systems, it is critical to understand how DNA nanotechnology enables powerful methods for structures and circuits, how structures and circuits can be combined to achieve systems with unique and exciting properties, and how recent advances allow us to envision a route toward technology previously limited to the realm of science fiction.

The following sections provide separate histories of DNA structures and DNA computing with two brief intermissions in 2008 and 2012 to discuss projects which combine concepts from both areas.

1.1 Foundations of DNA Nanotechnology

Researchers in the field of DNA nanotechnology typically consider Ned Seeman and Len Adleman to be the original pioneers of engineering DNA structures and DNA-based computing, respectively.

DNA Structures

Prior to 2004, it was not yet known that neutrophils (a type of immune cell) naturally use DNA to construct extracellular nets to trap and kill bacteria [4].

Therefore, when in 1982 Ned Seeman designed lattices made of branched DNA, his concept represented a radical departure from typical DNA research [5]. In 1998, under Ned's mentorship, Erik Winfree constructed DNA tiles that could stick together and form designed patterns [6]. Made of one or a few short strands, DNA tiles have been used for algorithmic computation [7] and as building blocks in complex structures [8].

In turn, while a postdoctoral fellow in Erik Winfree's group, Paul Rothemund developed DNA origami in 2006 [9]. This material is constructed from one long single-stranded scaffold strand of DNA, which is folded into arbitrary

2D or 3D shapes by a mixture of short staple strands. Each staple strand is typically composed of 2-4 domains, each of which can bind to a different region on the scaffold. Thus, just by mixing the scaffold and staples, DNA origami will spontaneously form through Brownian diffusion as various hybridization events take place. The most useful aspect of DNA origami is that it is addressable, meaning that molecules can be arranged in specific desired patterns on its surface. This trait allows nanoscale precision in assembly and measurement, prompting NIST to take an interest in advancing DNA origami technology [10].

Other types of nucleic acid structures have proliferated since the invention of DNA origami, in part thanks to the popular software package NUPACK which was first released around the same time [11]. Central to NUPACK are algorithms for efficiently predicting the most likely conformation that one or more DNA strands will take under a given set of conditions. It also predicts the probability that any given base will be paired or unpaired, and allows sequence optimization to ensure that designed DNA structures form as intended. It can be used for predictions and designs involving just one “complex ensemble"—the set of all possible ways an ordered list of DNA strands might interact—or involving one or more “test tube ensembles", each of which contains all possible complex ensembles for a given set of DNA strands. In the latter case, if initial concentrations are given for each strand, NUPACK will predict the final concentrations of each complex in the test tube at equilibrium. A limitation of this prediction is that NUPACK does not account for kinetic traps which may prevent certain complexes from readily forming. As of this writing, NUPACK is limited in that it cannot predict complexes involving pseudoknots (structures which cannot be plotted as a planar graph where no lines representing basepairs cross each other).

DNA Circuits

In 1994, Len Adleman demonstrated a method for solving an NP-complete problem using DNA strands and enzymes [12]. While this method was not the most efficient way to solve the problem, Adleman’s works suggested that the inherently parallel nature of chemical reactions could enable powerful computation. This work served as inspiration for many researchers to begin thinking of better ways to perform computation using molecules.

In 2000, Yurke *et. al.* showed that it was possible to use toehold-mediated strand displacement (TMSD) to switch a DNA construct between two states, dependent on the presence of two trigger strands [13]. The majority of DNA circuits since then have relied on TMSD as the core mechanism enabling construction of molecular computers.

In 2004, Dirks and Pierce invented the Hybridization Chain Reaction (HCR), which demonstrated conditional nucleic acid self-assembly, a principle enabling control over the order of events in nucleic acid assembly pathways. HCR is also a powerful method for amplifying signals which led to many patents and a successful company focusing on quantitative imaging [14]. This work further expanded interest in engineering systems that take advantage of DNA TMSD.

Inspired by this trend and Len Adleman's vision for DNA-based computing, in 2006, three future professors in the Winfree group constructed Boolean logic circuits based on DNA TMSD [15]. They thereby set several standards for molecular computation. Building blocks in molecular computers should be cascadable, meaning that their outputs should be able to act as inputs to another copy of the same building block. Each computational unit should be modular, meaning that they can be rearranged to produce unique systems with different overall functions. Finally, robust molecular computational systems must include signal restoration to combat signal loss. Because a signal represents order, whereas noise is a form of disorder, signal loss affects all systems which seek to transmit signals.

Computer scientists began developing the theory of using chemicals to perform computation in the 1960s [16]. By the 1990s, chemical reaction networks (CRNs) had become a well-established computational model. Like Turing machines and cellular automata, CRNs were valuable for theoretical computer science, but physical implementation was challenging. This situation reversed in the 2000s. Because DNA TMSD reactions enable programmable reaction connectivity and kinetics, a wide range of arbitrary CRNs could be implemented physically with far greater ease than in decades past [17]. Ongoing work in CRN theory continues to inform and guide development in DNA circuits.

1.2 Intermission 1: 2008

In 2008, Yin *et. al.* established a mechanism for developmental self-assembly (the first computational modality discussed above) [18]. With this system, a trigger molecule initiates a cascade of isothermal assembly steps programmed to take place in a specific order with controlled kinetics. In other words, the execution of the DNA circuits presented in this work results in the assembly of DNA structures. This system was originally applied to construct linear and branched 2D structures, but later work showed it could also be used to construct looping 2D structures and wireframe 3D structures [19]. These concepts will be discussed in greater detail in Chapter 2.

1.3 Select Advances in DNA Nanotechnology 2009-2012

Advances in Structures 2009-2012

DNA origami has many geometric constraints due to the sturdy helical twist of DNA, so designing it was labor-intensive prior to the 2009 release of Cadnano, software specifically built for the purpose of origami design [20].

In 2011, a different useful software package for modeling DNA structures became available. Compared to NUPACK, OxDNA took a very different approach, modeling each nucleotide as a bead on a chain [21]. While computationally intensive, simulating physical thermodynamic and mechanical properties with OxDNA allows predictions to be made about much more complicated structures involving DNA pseudoknots, such as DNA origami. OxDNA has therefore allowed more complex DNA origami structures to be built, while also providing insights about the shape of 3D DNA origami structures, which can be difficult to image.

Advances in Circuits 2009-2012

In 2009, David Soloveichik developed the CRNSimulator package, which has been widely used to model CRNs and help predict the behaviors of DNA circuits [22].

As DNA circuits grew in size, the potential for undesired reactions increased. These undesired reactions increased the loss of desired signals, and also produced undesired signals through leak. To improve the scalability of DNA logic gates, it was necessary to improve methods for combating signal loss with signal restoration, and for preventing leak. In 2011, Lulu Qian and Erik Winfree developed catalysts (ancestral to the catalytic systems presented in Chapters 3, 4, and 5) and thresholds for DNA circuits that performed signal restoration at every step of a circuit [23]. This work also prevented stacking-bond leak between gates by incorporating clamps into every toehold.

Also in 2011, an important characterization study quantified reaction rates involving remote toeholds [24]. Knowing these rates can help to provide a bound on predicted rates between reacting components on a surface (as in Chapters 4 and 5), since toehold hybridization between neighbors can be compared both to a DNA hairpin closing and to remote toeholds.

1.4 Intermission 2: 2012

In 2012, Douglas *et. al.* demonstrated a DNA origami nanorobot which could bind to specific blood cancer cell types with high precision, subsequently release a payload, and prevent the cancer cells from growing and multiplying [25].

This proof-of-concept took place in cell culture only, although no spatial constraint would be expected to prevent transfer to an organism. The nanorobot was roughly 40 nm in diameter. If it were injected into the human bloodstream, it likely could readily circulate through all blood vessels, since even the thinnest capillaries have an inner diameter roughly 100 times larger at 4 μm [26]. However, in 2012, it was not yet known whether DNA origami could be safely injected into living animals or how long it would last before being degraded or excreted.

Still, the nanorobot could sense its environment, react to a stimulus, and exert a meaningful effect on cells. All of the energy it needed was stored within it or came from its environment. For context, consider the smallest batteries ever made as of August 2024 [27]. These batteries were 10 μm in diameter and therefore would not fit inside a capillary. A microscopic robot powered by these would be still larger, requiring insulating material, a sensor, and an actuator. By these metrics, the field of electronics has not yet approached what DNA nanotechnology could already achieve twelve years ago: sub-micron-scale self-powered devices capable of reacting to stimuli and affecting their environment programmatically.

The proof-of-concept DNA nanorobot [25] built on all of the work that had come before it. Advanced though it was, its computational power (AND gate logic) was less complex than what could be achieved in well-mixed circuits at the time. In fact, embedding computational power in a DNA nanorobot designed to explore a biologically-relevant environment remains a nontrivial task. Chapters 4 and 5 of this thesis cover our latest work to address some of the remaining hurdles: the need for signal restoration in localized DNA circuits, designs which can carry everything they need with them to new environments, and challenges in ensuring correct assembly.

1.5 Select Advances in DNA Nanotechnology Since 2012

Many developments since 2012 have made DNA nanorobots more feasible and promising, fostered in part by a global community of researchers whose mutual respect, collaborative spirit, and regular interactions enable rapid progress.

Advances in Structures

Our group has made substantial strides in terms of the modularity of DNA origami structures, developing methods for constructing larger objects from multiple DNA origami tiles [28, 29], and exploring methods for rearranging DNA origami tiles through the process of tile displacement [30, 31]. Other groups have made inspirational reconfigurable DNA origami structures as well, such as origami that changes in length [32] or can bend or twist [33] in response to environmental stimulus. These developments allow us to envision nanorobots that can reconfigure with much more complexity than a box which simply opens or closes.

As DNA origami structures have diversified, so too has the software available to design them. In 2019, an alternative to Cadnano was released. Scadnano [34] has several key advantages over Cadnano which have made designing DNA origami structures even easier. Scadnano works in any web browser, allows Python scripting (its name is short for scriptable Cadnano), and is more flexible than Cadnano in what types of designs are allowed. This last point is critical, since it opens the future of DNA structures to the full creativity of researchers.

Another valuable tool is ENSNano, software released in 2021 which bridges the gap between Cadnano/Scadnano and OxDNA by quickly converting 2D blueprints to 3D models that can be manually adjusted, simplifying the design process and allowing designs to be refined before running computationally-intensive OxDNA simulations [35].

In keeping with the collaborative nature of the field, a cloud-based repository of DNA origami designs and models (Nanobase) has allowed researchers to better share their work and quickly build on each others' accomplishments [36].

Meanwhile, the toolbox for imaging DNA origami has expanded. Where previously it was only possible to image 2D DNA origami structures using an AFM, high resolution 3D origami images are now typically generated using cryo-EM [37, 38].

Advances in Circuits

Some critical first steps have been taken in terms of embedding computation on an origami surface (discussed in detail in Chapter 4), as well as in terms of modeling surface-based CRNs [39, 40].

Too many advances have been made in well-mixed DNA circuits to cover here, but some which I found inspiring include:

- the invention of an annihilator motif which allows two signals to cancel each other for winner-take-all neural network computation [41]
- the development of allosteric toeholds, which enable an invader strand to bind to the top strand of a gate rather than the bottom—where the bottom is the strand with an open toehold [42]
- improvements in minimizing leak by using two or more branch migration domains [43, 44]

Knowing about these diverse design motifs opened my mind to consider more creative possibilities for reaction types, which in turn allowed the development of the latch motif discussed in Chapter 5.

Advances in Biological Compatibility

Since 2012, researchers have improved the biocompatibility of DNA nanotechnology across numerous metrics.

In 2018, researchers announced a method to stabilize DNA origami against a wide variety of conditions, such as distilled water, fetal bovine serum, and temperatures up to 90°C. Origami stabilized in this manner lasted longer than conventional origami under physiological conditions, and during exposure to nucleases. The method, internal crosslinking via UV point welding, essentially adds more covalent bonds to DNA after exposure to 310 nm UV light for 30 minutes to two hours [45].

Mechanisms for delivering DNA nanostructures into living systems vary. The simplest mechanism may be the one used to deliver nucleic acids into plants. The plant cell wall excludes particles larger than 20 nm, but DNA nanostructures smaller than this will infiltrate plant cells naturally after cutting a leaf with a razor and injecting a sample with a syringe [46]. Mechanisms for delivering DNA origami into cultured human cells and living mice are slightly more complicated.

DNA origami can be electroporated into human cell lines such as HEK293T. If desired, DNA origami delivered in such a way can be used for gene expression, because cellular mechanisms will unfold and linearize DNA origami. Transcription of genes within the unfolded scaffold will take place as if the DNA had been delivered in a more traditional format. It does not matter whether the coding or template strand for the gene is used as the scaffold—cellular mechanisms will replace the staples with the conventional reverse complement of the scaffold. However, internal crosslinking via UV point welding stabilizes DNA origami such that it will not denature inside HEK293T cells, suppressing gene expression [47].

DNA origami delivered intravenously into mice was found to accumulate in their kidneys. For healthy mice, no effect was observed, whereas for mice which already had acute kidney disease, the DNA origami actually protected kidney function. Additionally, the authors checked for markers of toxicity and immunogenicity, and did not find any, suggesting that DNA origami may be biocompatible by default [48]. In 2024, lipid nanoparticles—much like the ones used for COVID vaccines—were demonstrated to serve well as a mechanism for delivering DNA origami into cells of specified mouse organs. The origami had to be made compatible with low pH conditions through UV cross-linking. When it was co-delivered with luciferase mRNA, it enhanced gene expression compared to lipid nanoparticles containing luciferase mRNA alone, because (for reasons not yet understood) it increased the total amount of mRNA delivered to the cells. This work also found that UV cross-linked origami is not toxic [49].

In 2013, DNA-RNA hybrid origami was developed [50, 51] and in 2014, RNA origami was demonstrated as well [52]. Since then, it has become possible to produce RNA origami through transcription, which could enable applications where origami structures are produced inside cells rather than delivered to a biological environment [38].

Other work with alternate nucleic acids has further strengthened the position that DNA nanotechnology could lead to major advances in biology, medicine, and diagnostics. Making a version of a DNA circuit or structure that is immune to degradation by a nuclease, and generally orthogonal to biological systems, can be as simple as rebuilding it in the mirror form, L-DNA. Simple double-stranded L-DNA, coupled to a chemotherapy drug, has been shown to reduce the systemic toxicity of the drug in mice, while efficiently delivering the drug to tumors to stop their growth [53]. L-DNA circuits have been shown to function inside mammalian cells with less leak, faster kinetics, and greater stability than D-DNA circuits [54]. Even simply adding L-DNA caps to D-DNA greatly enhances its longevity inside living human cells by limiting its degradation to endonucleases only [55]. Furthermore, DNA circuits which translate signals between D-DNA and L-DNA show that naturally-occurring molecular information can inform an orthogonal L-DNA molecular computing system which can then act upon its environment by translating its outputs back to D-DNA in a final step [56]. Therefore, only the initial and final components in a DNA computing system need ever be rigorously designed for compatibility with an environment full of natural nucleic acids. Recent advances in mirror-image proteins could lead to nanorobots which combine multiple types of orthogonal biomolecules [57].

Mechanisms for DNA nanorobots have also improved. In 2021, a DNA nanorobot was demonstrated which used antibodies to release a payload in response to two antigen molecules with AND gate logic [58]. Another DNA nanorobot consists of a swarm of building blocks that form a shell around viruses, neutralizing their ability to infect human cells [59].

All of these developments are inspiring to us, and led to specific design constraints we applied to the work on localized DNA circuits described in Chapters 4 and 5 of this thesis. Additionally, because many of the methods allow facile conversion of a non-biocompatible DNA nanodevice to a biocompatible version, we were able to focus purely on engineering DNA nanotechnology that is only proven to work in a lab setting while maintaining confidence that a broader scope is within reach. It has been established that within certain bounds, anything we develop can be adapted for applications later if desired.

The Present

DNA nanotechnology has already found use in industrial applications. For example, Illumina is a DNA sequencing company that uses a technology called Sequencing by Synthesis (SBS) to quickly sequence many strands of DNA in parallel [60]. SBS was originally developed at Cambridge in the late 1990s [61]. This technology bears some superficial resemblance to some of the concepts discussed in Chapter 4 and 5, in that DNA strands are localized to a surface where their signals are amplified. However, in the case of SBS, the position of the strands on the inorganic surface is random and the amplification of signals is performed by DNA polymerase in a localized variation on PCR. In contrast, the work in this thesis involves DNA localized to designated positions on a DNA origami structure, and signals which are amplified by a DNA catalytic circuit without the use of enzymes.

As another example, since 2018, the company Molecular Instruments has used DNA nanotechnology (specifically HCR) to image tissues and organisms, where target nucleic acids and proteins are color-coded such that their distribution throughout the sample can be quantified and visually understood. This technology has applications in research and diagnostics [62].

Illumina and Molecular Instruments apply DNA nanotechnology to generate unprecedented quantities of useful data about biomolecules. Considering all of the recent advancements in biological compatibility, DNA nanotechnology is now also ripe for therapeutic applications. Additionally, advances in DNA origami could lead to industrial developments where molecules are arranged in intentional layouts, rather than dispersed randomly across a surface.

In fact, in 2023, several new startups featuring DNA origami in their core technological offerings began operating with government funding in Germany [63]. For example, Nanogami aims to arrange DNA origami on biochips with the goal of developing diagnostic devices with unprecedented value, achieving relatively complex sensing capabilities with a relatively low production cost. Plectonic reports development of a DNA origami device using AND logic which irreversibly binds tumor cells and reversibly binds healthy cells. It only changes conformation upon binding tumor cells, revealing a molecule which triggers a response from immune cells. They further claim that this device reduces the presence of tumors in animal models with fewer side effects than existing cancer immunotherapy strategies. CPTx claims to be able to develop custom antiviral drugs based on origami structures which can capture and inactivate any arbitrary virus. They have developed proof of concept variants of their technology which can capture the following viruses: AAV2, polio 3, dengue, norovirus, HPV16, SARS 2, chikunguya, adenovirus, influenza, and rubella. They claim that for human cell cultures infected with each of ten types of virus, the presence of their technology prevented further viral infection such that the human cell cultures transitioned to a healthy state.

These industrial developments are promising, even though biotechnology startups are risky endeavors. However, even if all of these startups succeed in their goals, more fundamental research is still needed to help reach the full potential of DNA nanotechnology. We have yet to develop a biocompatible DNA nanodevice with computing power comparable to cascades of Boolean logic gates. The more computational power we can embed in nanorobots, the more specificity they could have in delivering medicine to the correct cell type at the right time, and the less likely they will be to produce false positives and negatives if applied toward diagnostics.

1.6 This Thesis

Despite the focus of this introduction, this thesis does not seek to develop any technology with immediate commercial applications, nor is any new DNA nanorobot presented here. However, this collection of fundamental research was performed with the understanding that any work in this field could enable or inspire translatable innovation that directly benefits individual humans and humanity as a whole. Therefore, this thesis was written with the singular hope of expanding our capabilities as molecular engineers to build a better future.

Chapter 2 describes a method to trigger a series of DNA strands to join each other in a prescribed sequence. Beginning their existence as hairpins, these strands do not interact with each other until a specific key complex is added that initiates their assembly. While the information contained in the hairpins encodes possible assembly pathways, only the key complex contains the final piece of information needed to determine the precise set of reactions which take place, and in what order. As DNA structural assembly grows increasingly complex, the potential for off-target structures and mixed populations also grows. Projects such as this one, which elucidate ways to exert more control over the events that take place during structural assembly, could be essential for expanding the possible nanodevices that we can build.

Chapter 3 combines the mechanism of the Seesaw Catalyst [23] with cooperative strand displacement to achieve a catalyst which relies on two inputs rather than just one. The concentration of each input only affects the kinetics of output production, but not its final completion level. If either input is entirely missing, however, the reaction will not take place. This motif could enable advances in AND gates and thresholding. The data presented here can also help with predictive modeling, since it illustrates how kinetics change when a toehold is located at the center of a complex rather than on an end as in the typical well-mixed case. Future advances in localized circuits may rely on our ability to model this type of toehold, as discussed in Chapter 4.

Chapter 4 presents a flawed design for a localized catalyst. This variant on the Seesaw Catalyst concept is tethered to DNA origami, and builds on past work by other groups using hairpins as gates on origami surfaces. These designs worked well previously because they either did not include any localized ssDNA fuel, or because each reaction step on the origami surface involved a diffusible intermediate. In the first case, leak between two double-stranded complexes, particularly hairpins, is known to be especially slow. In the latter case, it was not possible for the localized components to leak until the DNA origami was mixed with diffusible fuels. Even then, localized components typically could not leak with each other without involving some diffusible element. Our design sought to eliminate diffusible intermediate components, and furthermore did include an ssDNA fuel. These changes allowed localized components to directly leak with each other. Therefore, in a textbook case of Murphy's law, they did.

In Chapter 5, we present a leak mitigation strategy which resolved the problem introduced in Chapter 4. This strategy could only work in circuits that are localized to a rigid structure because it relies on geometric constraints. Rather than changing the DNA sequences or domains, we take advantage of the surface to restrict the locations of gates and fuels such that they physically cannot reach each other to leak. Only the intended reaction pathway can change the location of the gate such that the fuel can reach it. Therefore, much like in Chapter 2, we present a method for more tightly controlling the sequence of events in a pathway of possible molecular reactions, although our approach is entirely different. While the leak reduction mechanism achieved its goals, we encountered some difficulties in ensuring that this new localized circuit always assembled as designed. Some recent advances which could greatly improve assembly of these circuits are discussed in Chapter 6, which also concludes this thesis.

Chapter 2

DEVELOPMENTAL SELF-ASSEMBLY OF A DNA RING WITH RESPONSIVE SIZE

1. Glynn, A. T., Davidson, S. R. & Qian, L. Developmental Self-Assembly of a DNA Ring with Stimulus-Responsive Size and Growth Direction. *Journal of the American Chemical Society* **144,** 10075–10079. DOI: 10.1021/jacs.2c03853 (June 8, 2022).

2.1 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.

2.2 Introduction

Molecular self-assembly is key to the functionality of living cells, allowing lipids, nucleic acids, and proteins to organize themselves into structures with desired shapes and properties. Understanding the principles of self-assembly in engineered molecular systems has fundamental importance to controlling the behavior of biomolecules for technological advances. DNA self-assembly is one of the most well-studied areas of engineered molecular self-assembly [64]. Complex shapes with up to gigadalton sizes have been created with nanometer precision [9, 65–67]. Moreover, a self-assembly process could be designed to carry out complex computation and algorithms [68, 69].

Most of the DNA self-assembly processes investigated so far take place spontaneously during thermal annealing, but some exhibit isothermal behavior in response to a triggering signal [14, 18, 70]. Triggered self-assembly processes allow desired structures to grow at desired times, while the isothermal property allows for applications where temperature changes are undesired, for example in a biological environment [71, 72]. Similar to how the kinetics of growth in multicellular development is orchestrated by genetic programs, the kinetic pathway of triggered self-assembly can be controlled by molecular programs encoded in DNA; this type of behavior has been referred to as developmental self-assembly [19]. The kinetic control was achieved by toehold sequestration and toehold-mediated DNA strand displacement [73]: an initiator strand reacts with a hairpin motif by toehold binding and branch migration, which reveals a previously sequestered toehold for subsequent reactions. Using this mechanism, dendritic structures [18] and a tetrahedron [19] have been created with a prescribed sequence for every self-assembly step.

The prior investigations raised an important challenge regarding the design space of triggered self-assembly: as seen in biological systems, development can be influenced by changing environmental conditions throughout the entire growth process rather than just within the initiation step. What new design principles can be established to enable self-assembly of DNA nanostructures with stimuli-responsiveness more deeply embedded within the growth process? To begin answering this question, here we show that distinct signal molecules can be designed to selectively activate a subroutine (e.g. a subset of steps) in a developmental self-assembly program, resulting in structures of varying sizes or growth directions assembled from the same set of building blocks. In these examples, the signal molecules encode both the start and stop conditions of a growth process, paving the way for future explorations involving more complex conditions.

Various strategies for growing DNA nanostructures with programmable sizes have been developed, for example using an increasing number of strands for creating DNA tubes with increasing circumferences [74] or using distinct connector strands with a specific offset to enforce how wide a sheet must grow before rolling up into a tube [75]. Due to the nature of spontaneous

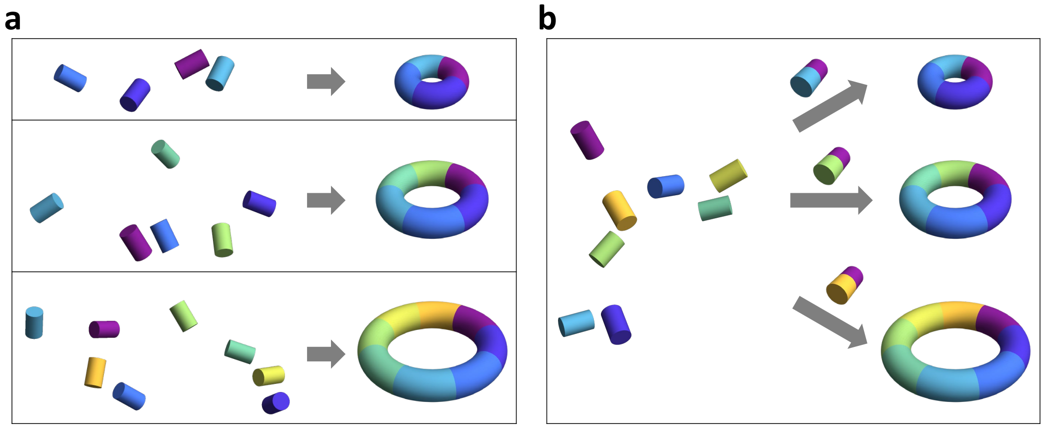


Figure 2.1: **Concept of (a) spontaneous and (b) developmental self-assembly that create DNA rings with distinct sizes.**

self-assembly, these strategies lead to immediate growth of partial target structures once the DNA strands are mixed together (Figure 2.1a). Seeded growth is possible in tile self-assembly systems, allowing the width of DNA ribbons to be controlled by using a DNA origami structure as an information-bearing seed [76]. In that case, nucleation only occurs when the seed is present. However, once the growth begins, DNA tiles will spontaneously bind to each other as they are all by default activated. By contrast, developmental self-assembly utilizes hairpin motifs that are activated one at a time. This unique property makes it possible to design a system where the entire growth process is inhibited through kinetic traps, and parts of which can be selectively activated upon specific signals; these signals could then alter the outcome of self-assembly (Figure 2.1b).

2.3 Results and Discussion

To demonstrate the concept of developmental self-assembly with stimulus-responsive properties, we designed a set of hairpins that can be triggered to form a DNA ring with varying sizes depending on the identity of a key that functions as both an initiator and a terminator (Figure 2.2a). As shown in the abstract reaction graph (notation explained in the figure caption), a unique toehold comprising the output port in each of the hairpins is initially sequestered, preventing them from interacting with each other when all keys are absent. A key reacts specifically with one of the hairpins, activating its output port for the next assembly step. A cascade of assembly reactions occurs until a hairpin has been activated whose output port matches the input port on the key. A disassembly reaction then takes place to release a strand from the key and close the ring, completing the self-assembly pathway. As shown in Figure 2.2b, the seven unique hairpins each consists of two exposed toeholds (input ports), a common branch migration domain (colored in black), and a sequestered toehold (output port). Three unique key molecules each consists of two strands, one of which opens up the first hairpin and the other will be released by the last hairpin in the designated self-assembly pathway. The released strand is labeled with a fluorophore for detection of pathway completion.

Each assembly step that opens up a hairpin is a reversible strand displacement reaction (Figure 2.3). The forward reaction is driven by additional base pairs in the toehold domain while the backward reaction is driven by the entropic gain of one free molecule. Toehold sequences were designed in NUPACK [77] to ensure that each forward reaction is at approximately the same rate as the backward reaction based on the estimated equilibrium constant with 100 nM reactant concentrations. The final disassembly step that closes the ring is an irreversible strand displacement reaction driven forward by entropy and possibly additional base pairs. A short loop domain in each hairpin provides desired structural flexibility for ring formation. The loop domain in the last assembled hairpin (hairpin1) also provides an option for controlling the kinetics of ring closure by serving as a toehold for initiating displacement of the fluorophore-labeled strand (key-t) in a key complex. Once key-t is released, no open toehold will be available for further reactions and thus the self-assembly pathway will be completed.

We first characterized the ring formation by gel electrophoresis. For control purposes, we designed a set of non-hairpin DNA strands that spontaneously self-assemble into a ring or a linear structure upon thermal annealing, with sizes similar to each of the triggered rings (Figure 2.4a) or the largest linear structures before ring closure. When the seven hairpins were mixed together without any keys (Figure 2.4b, lane 11), no products of larger sizes were observed; when key4, key6, or key8 was present, products with increasing sizes were formed (lanes 12 to 14), comparable to the 4-stranded, 6-stranded, and 8-stranded annealed ring controls (lanes 6, 8, and 10). In addition to SYBR Gold staining, the same gel was also imaged with ATTO590 fluorescence

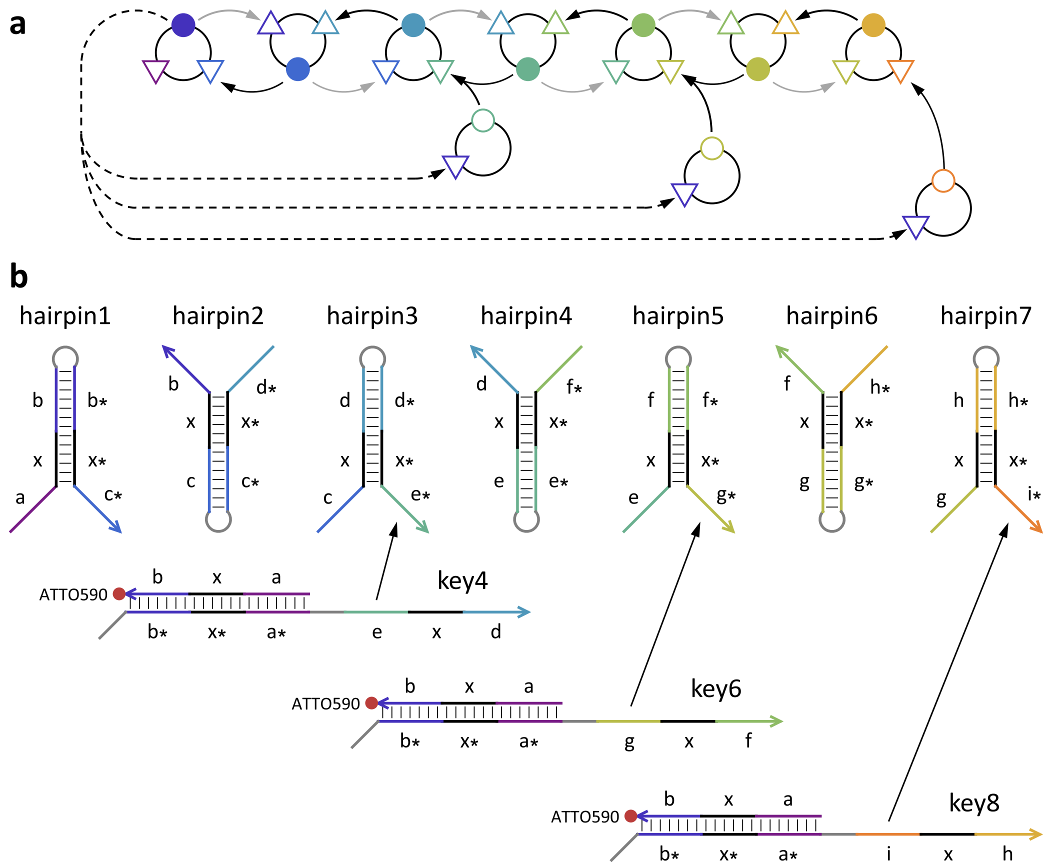


Figure 2.2: **Design of a DNA ring with stimulus-responsive size.** (a) Abstract reaction graph. Each hairpin is represented as a node with three ports, and each key as a node with two ports. Triangles and circles indicate input and output ports, respectively, while their open or filled representation corresponds to an exposed or sequestered toehold. Solid and dashed black arrows indicate assembly and disassembly reactions, respectively. Gray arrows indicate possible reactions that are not used in the designed ring formation. (b) Domain-level strand diagrams. Each unique toehold and branch migration domain is labeled with a distinct letter. Asterisks indicate sequence complementary. Three keys for triggering the self-assembly of DNA rings with 4, 6, and 8 strands are labeled as key4, key6, and key8, respectively.

(Figure 2.4c). All keys were consumed (comparing with lane 3), and the fluorophore-labeled key-t strand was released (comparing with lane 4) in all lanes containing hairpins and a key, indicating successful ring formation. Unlike the annealed controls of linear structures (lanes 5, 7, and 9), no key-t strands were observed in the products (lanes 12 to 14), confirming the difference between intermediates and products.

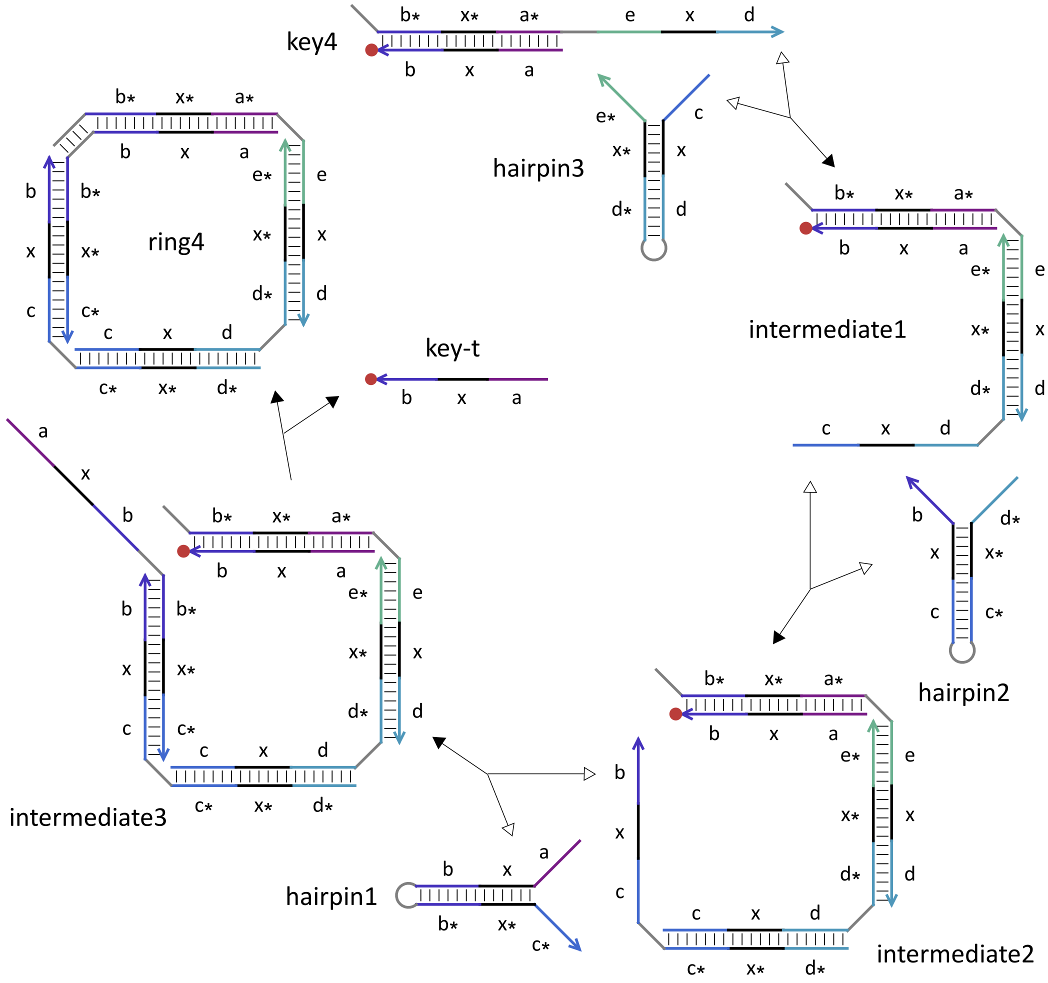


Figure 2.3: **Reaction pathway of a 4-stranded ring formation.** Forward and backward reactions are indicated by filled and open arrowheads, respectively. Each toehold, branch migration, and loop domain has 7, 6, and 4 nucleotides, respectively.

Intermediate steps of ring formation were also studied on a gel (Figure S6). When key8 was mixed together with an increasing number of hairpins 7 through 2, various sized products with an increasingly wider spectrum were observed (lanes 6 to 11). These products appeared both with SYBR Gold staining and with ATTO590 fluorescence, indicating that they contained the key-t strand. The spectrum of products confirmed that each assembly step is reversible, while the darkest band in each lane represents the most common product at equilibrium. Only when all seven hairpins were present (lane 12), a product band appeared at a position similar to the 8-stranded annealed ring control, while all products with smaller sizes largely disappeared. The disappearance of smaller-sized products confirmed that the final disassembly step is largely irreversible.

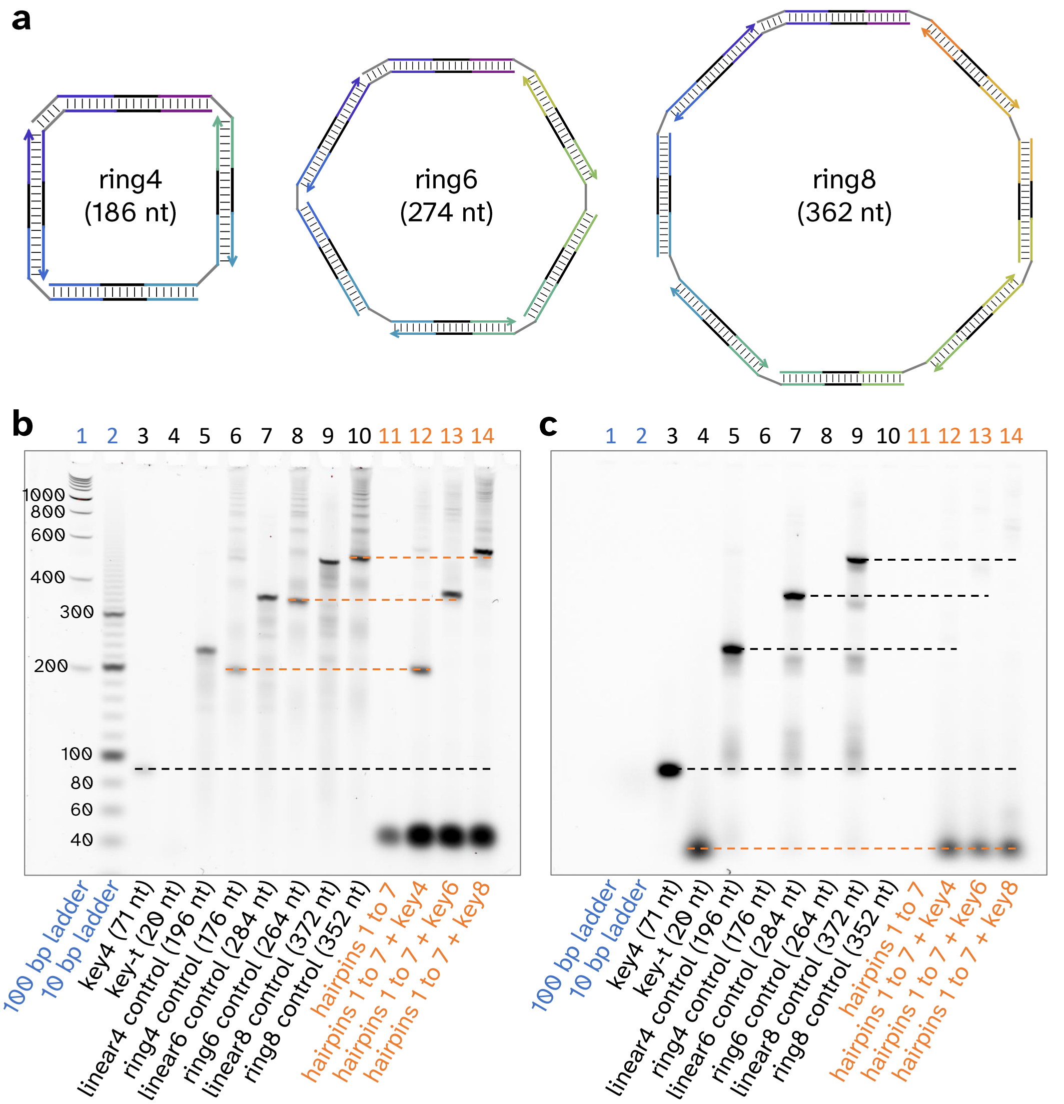


Figure 2.4: **Formation of DNA rings.** (a) Rings of three distinct sizes. Gel electrophoresis with (b) SYBR Gold staining and (c) ATTO590 fluorescence. Lanes containing DNA ladders, control structures, and hairpins without or with a key are labeled in blue, black, and orange, respectively. Dashed black or orange lines indicate comparison between controls and reactants/intermediates or products, respectively. Hairpins, keys, and control structures were at 150, 100, and 100 nM, respectively. Samples of hairpins and a key were incubated at room temperature for roughly one hour before loading onto the gel.

Interestingly, the gels also revealed the growth of polymers larger than the target ring size as well as double-sized rings. This is because in competition with ring closure, a linear polymer (e.g. intermediate3 in Figure 2.3) could also react with a second copy of the key molecule (e.g. key4) or any other linear polymers containing a key (e.g. intermediates 1 through 3) to form a longer polymer. A linear polymer with twice as many strands in the target ring could then close up to form a double-sized ring. The ring closure is an unimolecular reaction and the growth of a longer polymer is a bimolecular reaction, and thus the competition should favor the desired ring formation at a sufficiently low concentration. Indeed, insignificant amounts of larger polymers and double-sized rings were observed at 100 nM (Figure S4) compared to 1 μM (Figure S5), agreeing with simulation predictions (Figure S7).

Next, for quantitative understandings of reaction kinetics and completion of the ring formation, we performed fluorescence kinetics experiments at 100 nM. A reporter molecule was used to detect the amount of released key-t strand (Figure 2.5b). Change of fluorescence in bulk was normalized to concentration based on control experiments (Figure S8). To investigate the range of kinetics that can be controlled by a toehold on the key (Figure 2.5a), we varied the toehold length in a set of experiments on 4-stranded ring formation (Figure 2.5c). A simple model was developed for simulating the expected system behavior (SI note S2, Modeling and simulation). The forward rate of each assembly step was estimated based on effective strand displacement rate quantified in previous studies [78, 79]. The backward rate was calculated based on the equilibrium constant from NUPACK analysis (Figure S2a). Comparing simulation with experimental data, we then estimated the unimolecular rate constant (k) of strand displacement that occurs in the disassembly step that leads to ring closure, assuming that the rate scales inversely proportional to the ring size. Kinetics over a range of $\approx$four orders of magnitude was observed with 0 to 7-nt toeholds (Figure 2.5c), where the lower bound is comparable to the rate of strand displacement with a remote toehold and a double-stranded spacer [80].

To investigate the impact of ring size on the overall assembly kinetics, we used the same toehold on three distinct keys separately added to a mixture of seven hairpins. Formation of 4, 6, and 8-stranded rings all approached near completion within one hour, while fluorescence change without a key was hardly detectable (Figure 2.5d). Much longer experiments over 20 hours allowed the rate of leak between a key and reporter to be estimated (Figure S9a), which is comparable to strand displacement rate without a toehold [79]. Sim-

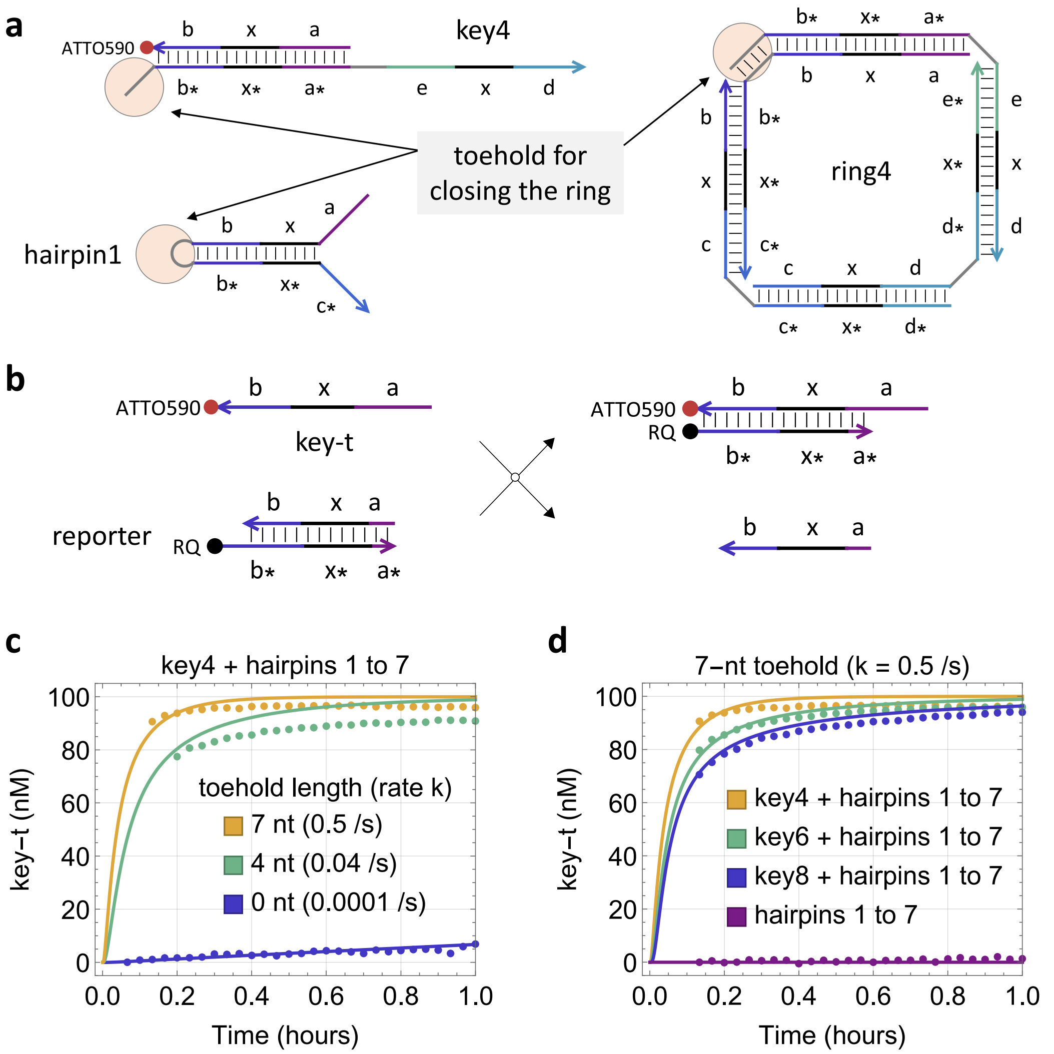


Figure 2.5: **Kinetics of ring formation.** (a) Toehold design for controlling the kinetics of the disassembly step that leads to ring closure. (b) Reporting mechanism that detects the amount of released key-t strand. Simulation (solid trajectories) and fluorescence kinetics data (dotted trajectories) of (c) 4-stranded ring formation with varying toehold sizes, and (d) ring formation with varying keys. Initial concentrations of keys, hairpins, and the reporter were 100, 150, and 150 nM, respectively.

ilarly, leak between hairpin1 and a key was also estimated (Figure S9b). In this case, because the loop domain in haripin1 was used as a toehold for displacing the key-t strand while closing the ring (Figure 2.5a), a four-way strand displacement leak reaction is possible when the hairpin is closed. With a 4-nt loop/toehold, the leak was not measurable. With a 7-nt loop/toehold, the leak was still lower than four-way strand displacement rate with a single exposed toehold [81].

Finally, we demonstrated a second type of stimulus-responsive growth where two unique keys were designed to control self-assembly in opposite directions of a reaction graph (Figure 2.6a). This is possible because each hairpin has two input ports (i.e. two exposed toeholds), enabling subsequent assembly steps involving different hairpins (Figure 2.6b). Fluorescence data suggested desired ring formation when each of the two keys was present (Figures 2.6c and S11b). Additionally, gel electrophoresis confirmed that both rings were of the correct size (Figure S11a).

2.4 Conclusions

To summarize, we have shown that the information of both start and stop conditions of a developmental self-assembly pathway can be encoded in a key molecule, allowing structures with distinct properties to grow in response to stimuli. The same strategy could be generalized to self-assembled structures with stimulus-responsive shapes—for example, linear, ring, and tree structures could be created as distinct subroutines of the same developmental program by choosing appropriate start and stop conditions within one self-assembly reaction graph. We have also demonstrated that an overall self-assembly process can be composed of reversible and irreversible steps, where the kinetics of a certain step can be tuned to be substantially slower than others. Exploiting this strategy, populations of structures with responsive growth properties could be created. For example, when the disassembly step of ring closure was removed, linear structures with varying sizes were created. In more complex reaction graphs with reversible steps, multiple slow assembly steps could be employed at desired locations, driving the growth of multiple populations of structures with distinct properties.

In general, developmental self-assembly is not only relevant for growing static structures, but also important for controlling the kinetics of dynamical structures that function as motors [18, 82]. Understanding the full range of behavior that can be programmed in developmental self-assembly is critical for embedding sophisticated control within materials such as hydrogels [83–85] and engineered molecular devices such as synthetic membrane channels [86, 87], enabling them to exhibit adaptive properties in response to a changing molecular environment.

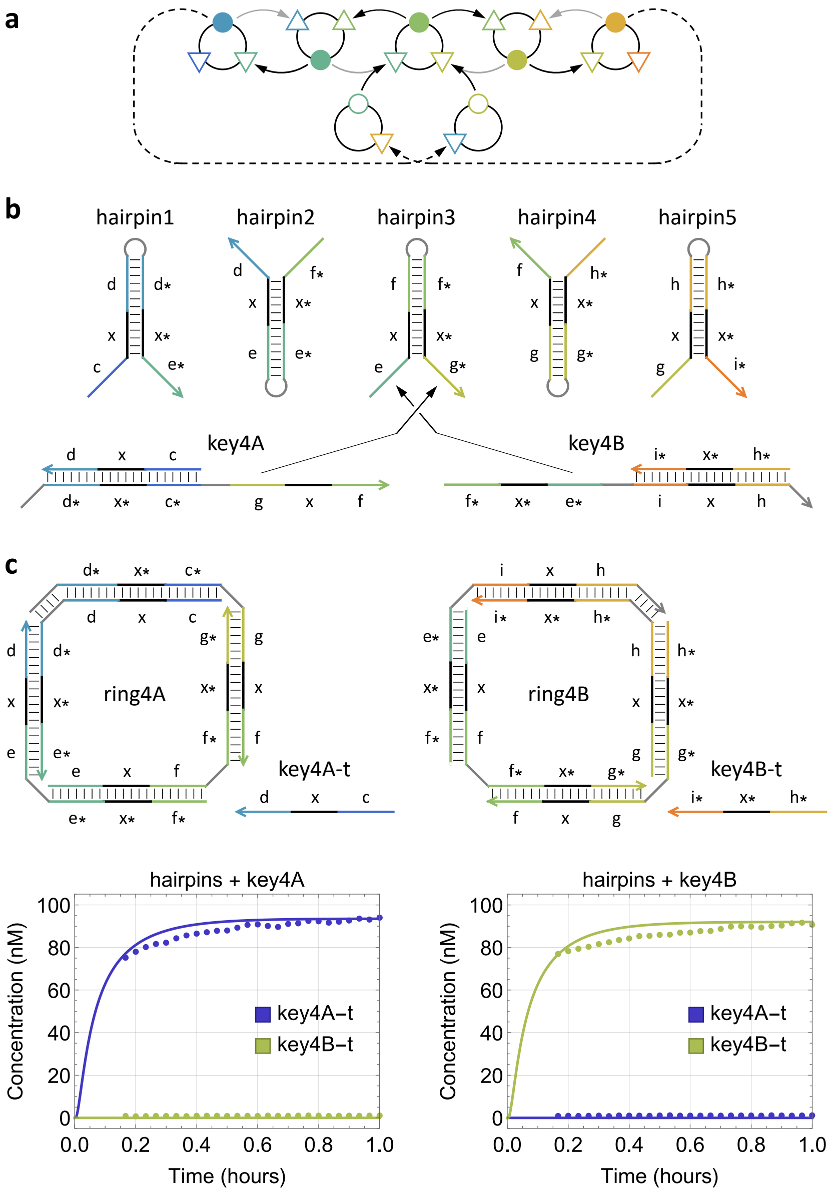


Figure 2.6: **Stimulus-responsive growth direction in developmental self-assembly.** (a) Abstract reaction graph. (b) Domain-level strand diagrams. The toehold for ring closure is 4-nt long in each of the two keys, complementary to the loop domain in hairpin1 and hairpin5. (c) Simulation and fluorescence kinetics data of ring formation with identical size but opposite growth directions controlled by two distinct keys. Initial concentrations of keys and hairpins were 100 and 150 nM, respectively. Reporting mechanism is shown in Figure S10.

2.5 Acknowledgements

A.T.G. and S.R.D. were supported by Caltech internal funds for BE/CS 196, a course on design and construction of programmable molecular systems.

S.R.D. was also supported by a NIH/BLP training grant (5 T32 GM 112592-5). L.Q. was supported by a NSF award (1813550).

Chapter 3

A COOPERATIVE DNA CATALYST

1. Taylor, D. N., Davidson, S. R. & Qian, L. A Cooperative DNA Catalyst. *Journal of the American Chemical Society* **143,** 15567–15571.
DOI: 10.1021/jacs.1c07122 (Sept. 29, 2021).

3.1 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.

3.2 Introduction

Catalytic reactions play essential roles in chemical and biological systems, underlying a wide range of enzymatic activities involving RNA and proteins. In engineered molecular systems, DNA has also been shown to carry out non-covalent catalytic reactions [18, 78, 88–91], opening up possibilities for embedding more sophisticated control within chemistry, materials, and medicine. For example, a DNA catalyst can be used for signal amplification in molecular diagnostics [92, 93]. When composed together with other motifs, a DNA catalyst can enable a variety of functions including signal restoration that is critical for scaling up the complexity of DNA-based digital logic circuits [23] and weight multiplication that is a basic component in DNA-based neural networks [94].

Prior work has demonstrated that several properties of DNA catalysts can be well controlled. Utilizing the mechanism of DNA strand displacement [73], the reaction rate of a DNA catalyst can be controlled by the length of a toehold [78]. A key property of a catalyst is that it serves as an input signal to trigger the production of an output signal without being consumed itself (Figure 3.1a): a small amount of input X can result in a much larger amount of output Y, facilitated by gate GY and fuel F. Moreover, allosteric control can be introduced by designing the gate to be initially inhibited and only react with the input when a consumable activator signal A is present [95] (Figure 3.1b). It has been articulated that this type of activatable catalyst is central to the realization of adaptive memories and learning behaviors in DNA-based neural networks [41, 96].

3.3 Results

In this work, we introduce a new type of catalyst that we call a cooperative catalyst: two signal species X and A cooperatively and catalytically produce output Y (Figure 3.1c). The additional signal can be viewed as an activator that provides allosteric control for $X \rightarrow X+Y$ without being consumed itself. This type of catalyst extends the functionality of DNA catalysts from unimolecular to bimolecular, satisfies the criteria for being used as a composable motif, and is sufficiently simple for building robust and scalable systems (SI Note S2, Concept of three types of catalysts and Figure S1).

Similar to the seesaw DNA catalyst [23, 97], a pair of reversible reactions are employed to entropically drive a catalytic cycle:

$$X + A + GY \rightleftharpoons XGA + Y$$

$$XGA + F \rightleftharpoons X + A + GF$$

Output Y is initially inhibited in a gate GY. Input X and activator A cooperatively react with the gate to release the output while becoming bound to the gate themselves. A fuel F then reacts with the input-activator-bound gate XGA, freeing up both input and activator while generating a waste product

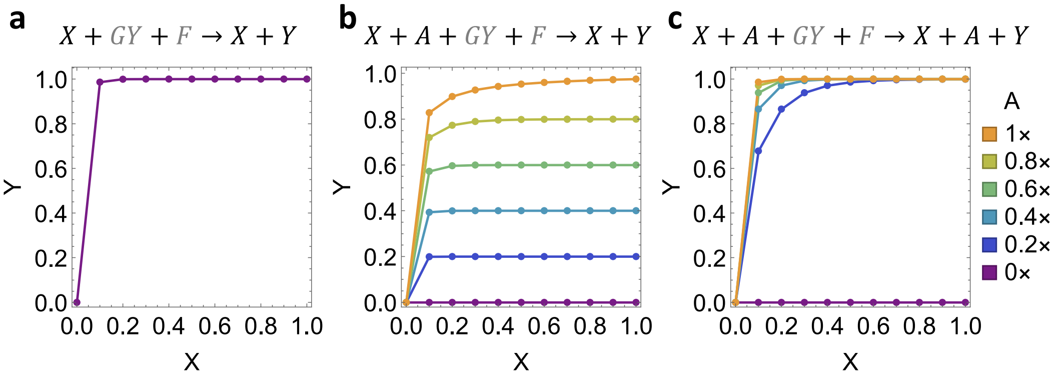


Figure 3.1: **Simplified overall reaction and characteristic simulation of a (a) basic catalyst, (b) allosteric catalyst, and (c) cooperative catalyst.** Signal and auxiliary species are colored in black and gray, respectively. Each plot shows the relative concentration of output Y at 1 hour versus that of input X. Standard concentration $1\times = 100$ nM. Initial concentrations of gate and fuel are $GY = 1\times$ and $F = 2\times$, respectively. Initial concentration of activator A is shown in the legend.

GF. An excess amount of fuel continues to drive the reaction forward for output production.

Unlike the seesaw DNA catalyst, each reversible reaction here is a trimolecular instead of a bimolecular reaction, and is implemented with cooperative hybridization [98] instead of toehold exchange [78] (Figure 3.2a). The input and activator strands each consists of a short toehold (T1 or T2) and a longer branch migration domain (S1 or S2). They each bind to an open toehold on one side of a double-stranded gate:output complex, and compete with the output strand for binding to the gate strand in a branch migration process. When both strands reach the end of branch migration and become fully double-stranded, the output strand is attached to the gate by only a short toehold domain (T3) which can spontaneously dissociate. The input-activator-bound gate has a center toehold open, which now allows the fuel strand to bind. The fuel strand has two branch migration domains flanking a toehold, each initiating a competition with the input or activator for binding to the gate. Eventually both input and activator strands are released, completing a catalytic cycle.

In cooperative hybridization, each toehold binding (with rate constant k_f), branch migration (with rate constant k_b), and toehold dissociation step (with rate constant k_r) can occur independently rather than simultaneously as a pair. Thus, besides the representative states shown in Figure 3.2a, additional

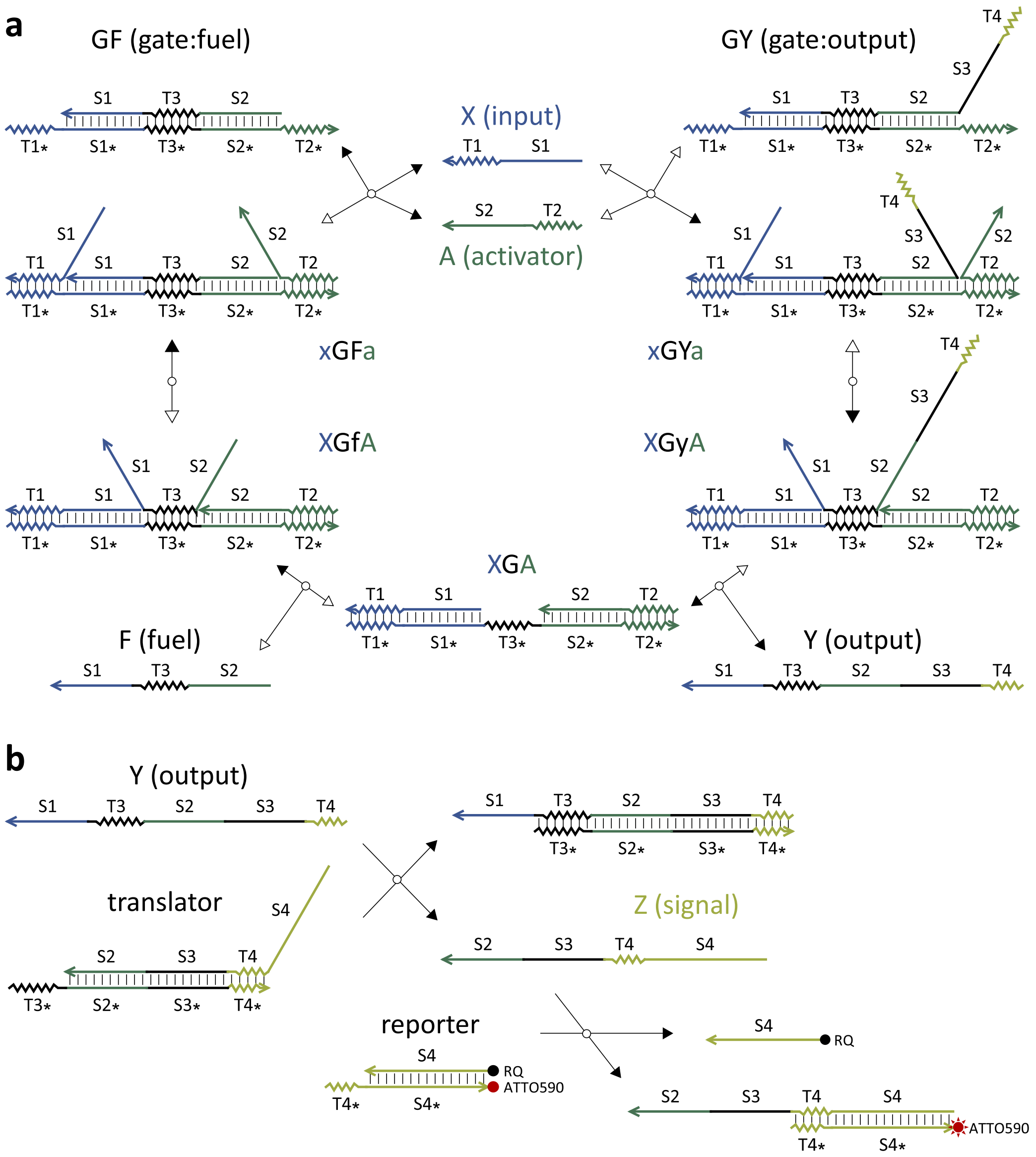


Figure 3.2: **DNA strand displacement implementation of a cooperative catalyst.** (a) Catalytic reaction pathway. (b) Signal translation and reporting. Single strands are named with a single letter. Multi-stranded complexes are named with multiple letters indicating the strands that they consist of, where a capital letter indicates a strand that is bound by at least one branch migration domain and a small letter indicates a strand that is bound by only a toehold. Forward and backward reactions are indicated by filled and open arrowheads, respectively. For simplicity, a clamp domain is not shown here but illustrated in Figure S2.

states are involved for understanding the behavior of the molecules, half of which are shown here (the other half is shown in SI Note S3, Modeling and simulation of cooperative catalyst):

$$X + GY \underset{k_{r1}}{\overset{k_f}{\rightleftarrows}} xGY \underset{k_b}{\overset{k_b}{\rightleftarrows}} XGY$$

$$A + GY \underset{k_{r2}}{\overset{k_f}{\rightleftarrows}} GYa \underset{k_b}{\overset{k_b}{\rightleftarrows}} GYA$$

$$\begin{tikzcd}
 & & X + GYA \arrow[dr, shift left, "k_f"] & & & \\
X + GYa \arrow[dr, shift left, "k_f"] & & & xGYA \arrow[ul, shift left, "k_{r1}"] \arrow[dll, shift left, "k_b"] \arrow[dr, shift left, "k_b"] & & \\
 & xGYa \arrow[ul, shift left, "k_{r1}"] \arrow[dl, shift left, "k_{r2}"] \arrow[urr, shift left, "k_b"] \arrow[drr, shift left, "k_b"] & & & XGyA \arrow[ul, shift left, "k_b"] \arrow[dl, shift left, "k_b"] \arrow[r, shift left, "k_{r3}"] & XGA + Y \arrow[l, shift left, "k_f"] \\
A + xGY \arrow[ur, shift left, "k_f"] & & & XGYa \arrow[ull, shift left, "k_b"] \arrow[ur, shift left, "k_b"] \arrow[dl, shift left, "k_{r2}"] & & \\
 & & A + XGY \arrow[ur, shift left, "k_f"] & & &
\end{tikzcd}$$

Clearly, when only input X or activator A is present, no output Y will be produced.

To differentiate the output strand from the fuel, a third branch migration domain (S3) is needed—this allows the output strand to participate in downstream reactions that require a toehold (T3) and two adjacent branch migration domains (S2 and S3). It is often desired that input and output signals have the same format and independent sequences so that distinct DNA strand displacement motifs can be composed together for more complex system behavior. To achieve that, a translator can be designed to react with the output strand Y and produce a signal strand Z that includes a toehold and a single branch migration domain (T4 and S4); this reaction can be designed to be irreversible by including the toehold (T4) in the output strand (Figure 3.2b). Importantly, the fuel strand can only react with the translator reversibly without producing any signal Z. After the signal has been translated, a previously developed reporter [41, 99] can then be employed for fluorescence signal readout (Figure 3.2b).

With the above design, we set off to characterize the circuit behavior using fluorescence kinetics experiments (SI Note S1, Materials and methods) and to gain a quantitative understanding by comparing the data with simulation (SI Note S3, Modeling and simulation of cooperative catalyst). We first demonstrated the cooperativity between input and activator for output production. The output signal reached near completion within 1 hour when the input and activator strands were both present, and remained low when either strand was absent (Figure 3.3).

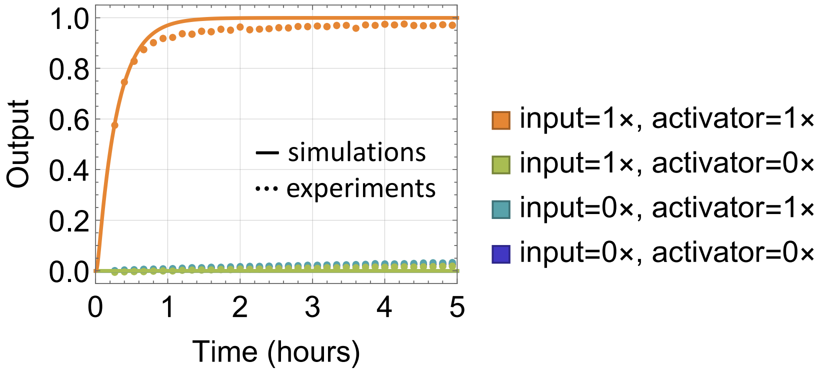


Figure 3.3: **Cooperativity between input and activator.** The bottom three trajectories overlap. Here and in later figures, standard concentration $1\times = 100$ nM; initial concentrations of gate:output, fuel, translator, and reporter are $1\times$, $2\times$, $1.5\times$, and $1.5\times$, respectively.

Next, we investigated the catalytic property of the input when the activator is present. Previous studies on cooperative hybridization focused on irreversible reactions [41, 98]. Here, reversibility is central to the catalytic behavior, and thus we must understand how the toehold responsible for reversibility (T3) controls the kinetics of the overall reaction pathway. Specifically, the rates of two reactions depend on the length and sequence of T3 (Figure 3.4a): When the input and activator are both bound to the gate and fully branch-migrated, the dissociation rate of the output strand depends on T3. After the release of the output strand and when the fuel binds to the input-activator-bound gate, the probability of the fuel successfully initiating a branch migration before dissociation also depends on T3. A shorter T3 would make the output dissociation rate faster but the success probability of fuel displacing input and activator smaller, and thus it is undesired for this toehold to be either too long or too short. Moreover, it is known that the strand displacement rate in toehold exchange reactions largely depends on the initiation toehold when its length is no shorter than the dissociation toehold [78]. Based on these two considerations, we expected that a 7-nt T3 would result in fast kinetics of the overall system when both T1 and T2 are 7 nucleotides.

However, the experimental observation was surprising (Figure 3.4b): the kinetics was very slow, roughly the same as simulations if the dissociation rate

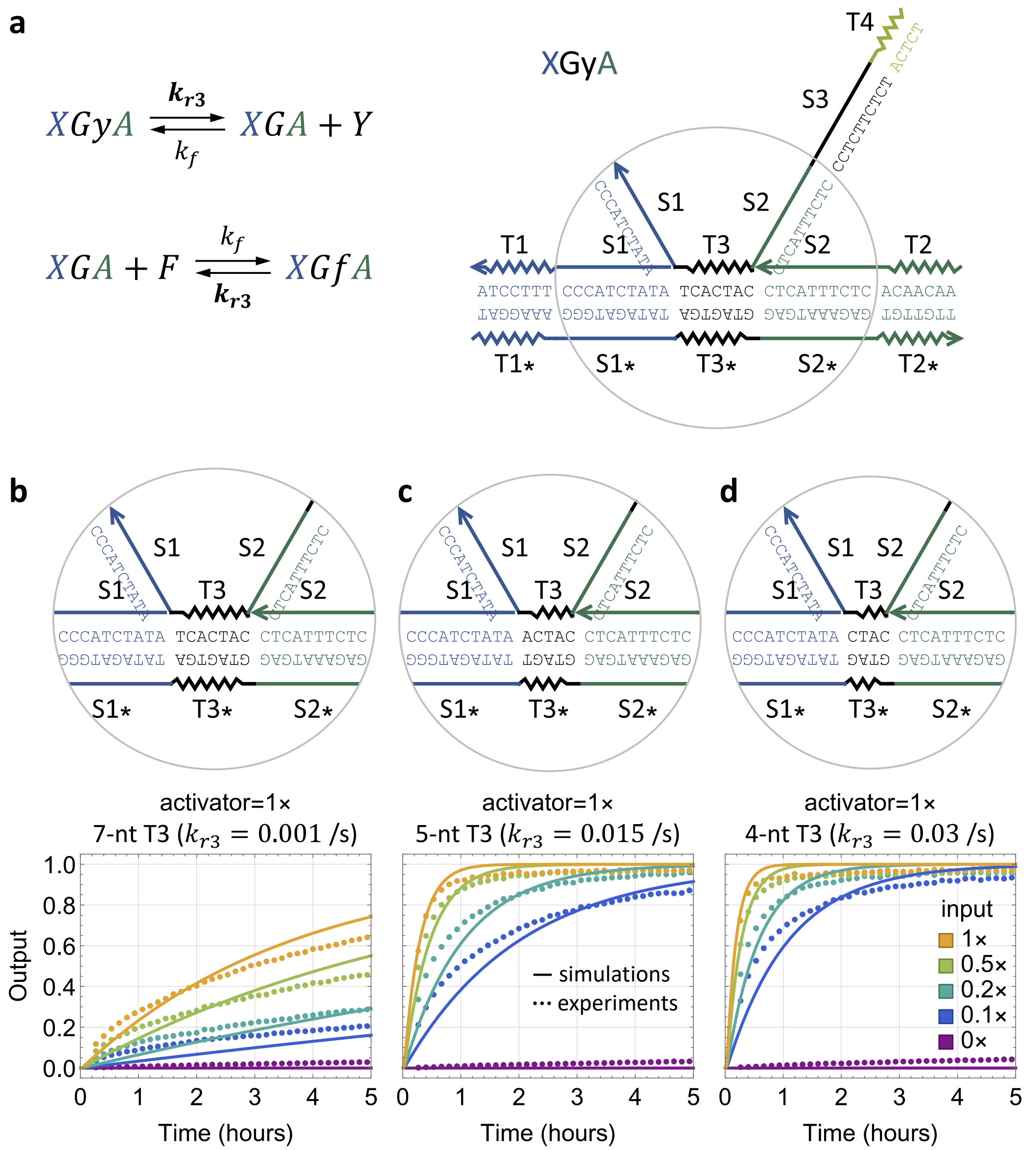


Figure 3.4: **Catalytic property of the input.** (a) Reaction rates (shown in bold) that depend on the sequence of toehold domain T3. Simulation and fluorescence kinetics data with varying input concentration, and with (b) 7-nt, (c) 5-nt, and (d) 4-nt T3 domains.

of T3 (k_{r3}) is set to 0.001 /s. As toehold dissociation is commonly estimated as $10^{(6-L)}$ /s for average sequences [79], where L is the length of the toehold, this value of k_{r3} would correspond to a 9-nt toehold. A possible explanation here is the stacking energy between the ends of two DNA helices [100]. While the role of coaxial base-stacking has been well studied in understanding the kinetics of simpler strand displacement reactions [79], the structure of the four-stranded molecule shown in Figure 3.4a is more complex: there are two stacking bonds between toehold T3 and two branch migration domains S1 and S2, both of which contribute to the reversibility of cooperative hybridization. This situation is related to but distinct from an internal toehold within a three-stranded complex, where toeholds with 4 to 7 nucleotides have been applied [91, 101].

To explore whether faster kinetics could be achieved with a shorter T3, we reduced the toehold length to 5 and 4 nucleotides. Approximately 15 to 30-fold speedup was observed (Figure 3.4cd), suggesting that strand displacement rate in reversible cooperative hybridization depends on both initiation and dissociation toeholds, even when the initiation toeholds are longer. With a 4-nt T3, the catalytic property of the input was clearly demonstrated; with $0.1\times$ input, the output reached near completion within 5 hours (Figure 3.4d).

We then showed that, like the input, a small amount of activator (e.g. $0.1\times$) was sufficient for catalyzing output production (Figure 3.5b). Besides tuning toehold lengths, we investigated whether a wobble [102] (G·T base pair) in branch migration domains could be exploited for promoting input or activator recycling without slowing down output production. Specifically, if there is a wobble in the S2 domain when the activator is bound to the gate, the process of activator release by the fuel would be biased forward (Figure 3.5a, first reaction, $k_b > k'_b$); if the wobble also exists when the output is bound to the gate, branch migration involving the activator and output could also be biased to favor output release (Figure 3.5a, second reaction, $k_b > k''_b$) depending on the position of the wobble [103] (SI Note S5, Effect of a wobble or mismatch for promoting activator recycling).

Experiments showed improvement in the catalytic property of the activator, indicated by the faster kinetics of the output production when activator was $0.1\times$ (Figure 3.5c). Interestingly, the improvement was particularly significant with unpurified gate:output and translator complexes (Figure S3), either with a wobble or with a mismatch [104] (non-Watson–Crick and non-wobble base pair). The impurity led to a behavior similar to the suppression of the activator by a threshold, where the output production appeared much slower compared to experiments with purified complexes. Introducing a wobble or mismatch in the S2 domain promoted the robustness of the activator, allowing it to catalyze output production even when its concentration is much lower than anticipated.

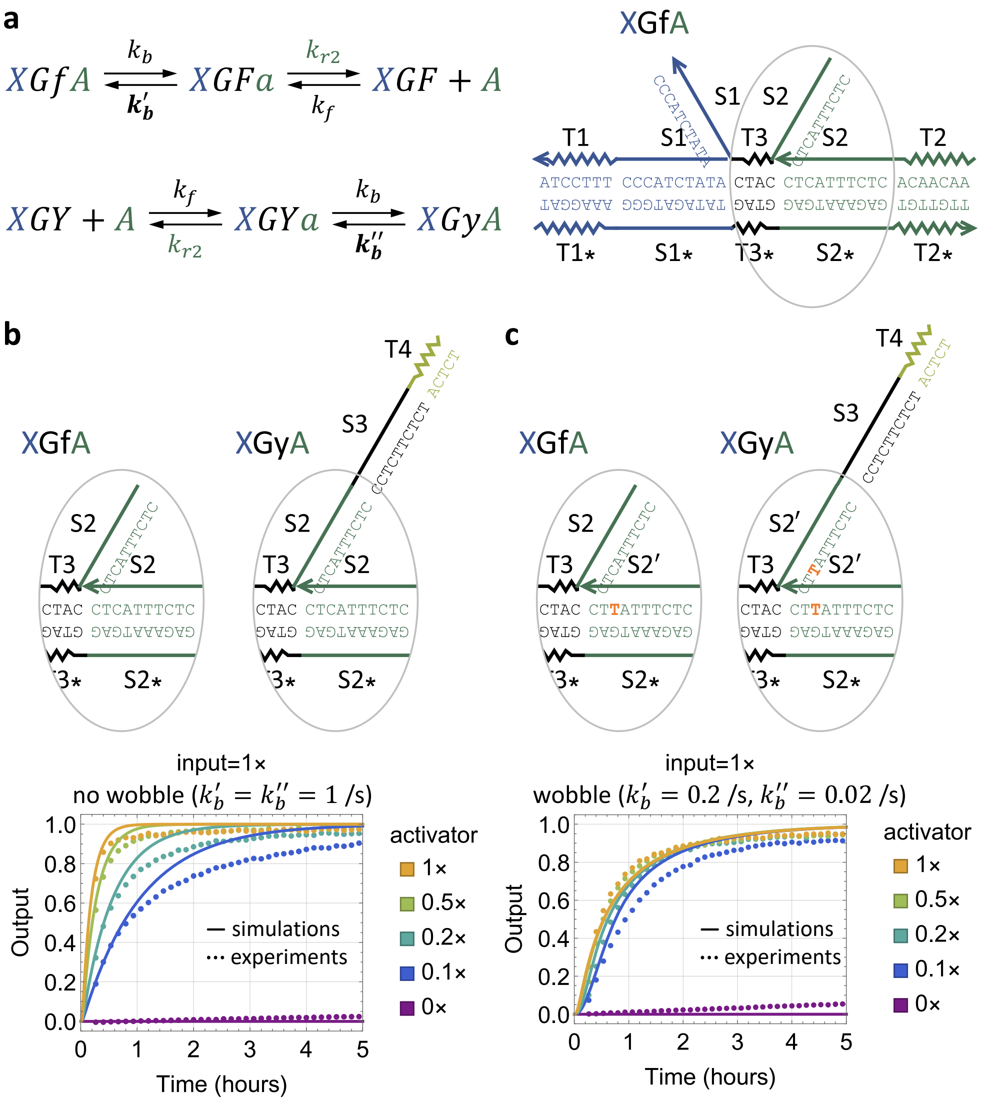


Figure 3.5: **Catalytic property of the activator.** (a) Reaction rates (shown in bold) that depend on the sequence of branch migration domain S2. Simulation and fluorescence kinetics data with varying activator concentration, and (b) without and (c) with a wobble base pair in the S2 domain.

Comparing simulation with data, we estimated that branch migration rate with a wobble is 0.2 to 0.02 /s (Figures 3.5c and S3), roughly 5 to 50-fold slower than branch migration rate with no wobble or mismatch [79]. This is consistent with the previous studies on the kinetics of strand displacement with mismatches [105–107].

Finally, we demonstrated cooperative catalytic behavior with low concentrations of both input and activator. The output production was near completion within 24 hours even when input and activator were both at $0.1\times$ (Figure 3.6a). It is clear that the output concentration is not limited by either input or activator concentrations so long as they are both present (Figure 3.6b), illustrating a key property of a cooperative catalyst.

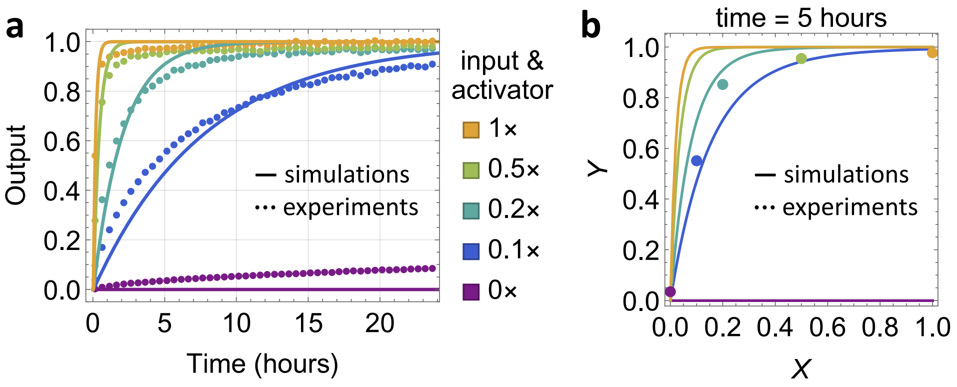


Figure 3.6: **Demonstration of the cooperative catalyst.** (a) Simulation and fluorescence kinetics data with varying input and activator concentrations. (b) Simulation and data of the output concentration at 5 hours versus the input concentration. T3 is 4 nt. There is no wobble in the S2 domain.

3.4 Discussion and Conclusions

The cooperative catalyst that we developed here will enable many advances in DNA circuits. For example, it could be used to build an AND gate with near-perfect signal restoration (SI Note S6 and Figure S5) and a better threshold mechanism that combines the advantages of sequential [15] and competitive [23] thresholding (SI Note S7 and Figure S6). Like the seesaw gate used for a basic catalyst, the cooperative gate GY is also two-stranded; this structural simplicity is particularly important for maintaining the robustness of DNA circuits when synthesis errors are inevitable [99].

In contrast to gate activation [95], allosteric control of a DNA catalyst could also be accomplished by input activation [108, 109]. However, these approaches require sequence dependence between input and activator strands, and the inhibited input signal cannot participate in other reaction pathways; both properties suggest limitations for composability with other DNA strand displacement motifs. Importantly, our approach provides the same format of signal species without any sequence dependence between them, making it particularly suitable for further empowering DNA circuits with general-purpose computation and dynamics [101, 110, 111].

3.5 Acknowledgments

D.N.T. and S.R.D. were supported by Caltech internal funds for BE/CS 196, a course on design and construction of programmable molecular systems. D.N.T. was also supported by the Summer Undergraduate Research Fellowships (SURF) program at Caltech and a NSF award (1908643). S.R.D. was also supported by a NIH/BLP training grant (5 T32 GM 112592-5). L.Q. was supported by a NSF award (1908643). The authors thank E. Winfree for reading and commenting on the manuscript.

Chapter 4

LOCALIZED CATALYTIC DNA CIRCUITS FOR SCALABLE MOLECULAR COMPUTATION ON A SURFACE
PART 1: LESSONS FROM A FLAWED DESIGN

4.1 Introduction

The identity of each component in an electronic circuit is determined by its function as well as its physical connectivity, that is, the set of other circuit components to which it is wired. This property of electronic circuits enables their massive scalability: an electronic circuit may be built from identical copies of the same components. Furthermore, undesired interactions or “short circuits" are easily prevented by simply not adding a wire to connect two components that should not interact. Conversely, in a well-mixed DNA circuit [15], communication may occur when any two molecules collide through Brownian diffusion. Because connectivity cannot be constrained spatially, it must instead be constrained by reaction specificity. Therefore, all molecular components must have different DNA sequences[3] even if they carry out the same function in different parts of the circuit.

Electronic circuits can be scaled up without increasing the number of distinct components, and the number of possible errors will only increase linearly. Well-mixed DNA circuits scale much worse: the number of distinct components increases linearly with circuit size, whereas the number of possible errors increases quadratically. Moreover, because there is a maximum total concentration of molecules, the speed of each reaction decreases linearly with circuit size as the concentration of reactants decreases.

Spatial organization is the property of electronics which confers its advantages over well-mixed molecular circuits. However, spatial organization is not unique to electronic circuits. It evolved separately at many different scales in biological systems. For example, in the brain, neurons interact with each other through the physical contact of synapses; all memory and learning functions rely on the proper wiring. In the cell, scaffold proteins assemble multiple components of a signaling pathway, increasing the efficiency and specificity of signal transduction.

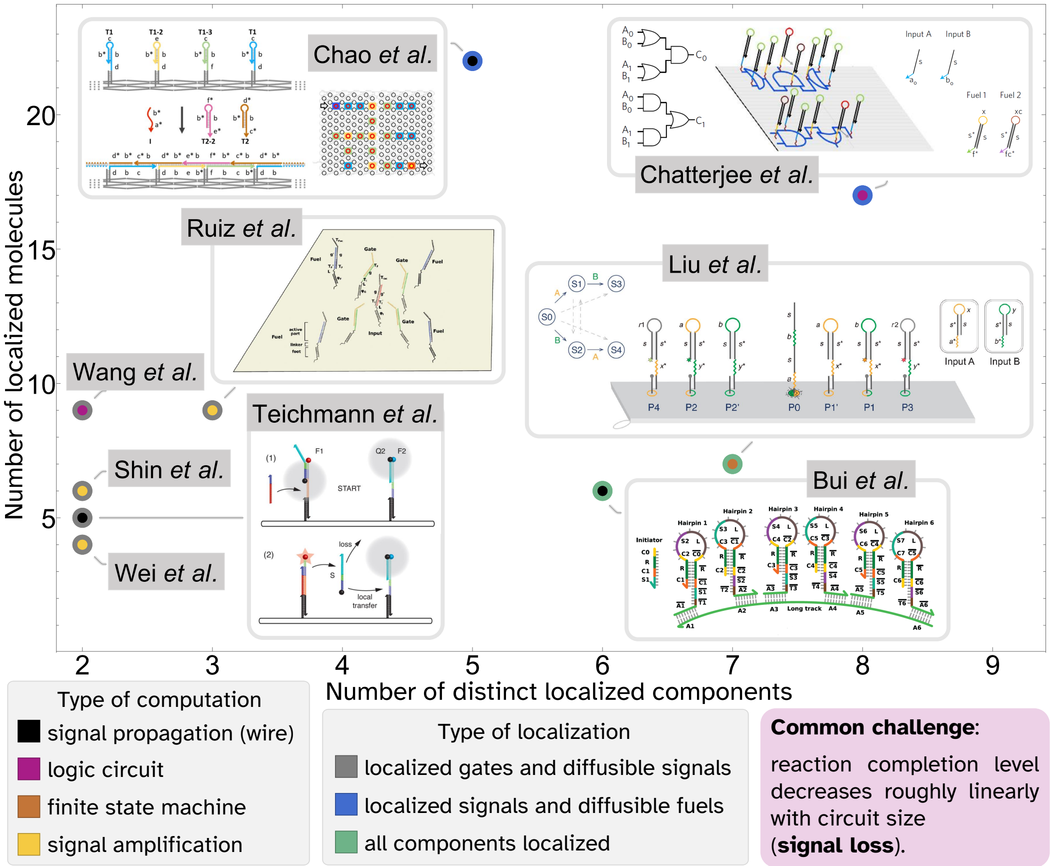
Physical arrangement can be introduced in DNA circuits. For example, one can use DNA origami [9] as a circuit breadboard to organize molecular components at fixed locations on a surface so that they may only interact with neighboring molecules. With this approach, DNA circuits of arbitrary sizes could in principle be built from a constant number of distinct components, where possible errors only increase linearly with circuit size, and the local concentration and thus individual reaction rate will remain constant [39]. In other words, spatial organization of DNA circuits could enable scaling factors comparable to electronics.

These advantages must be weighed against the challenges inherent in spatial organization. Compared to well-mixed circuits, spatially-organized circuits are more laborious and costly to assemble. The former requires only mixing circuit components, while the latter additionally requires assembly processes which may involve mixing, annealing, and purification of structural components. Ensuring correct structural assembly is essential, but not trivial. Furthermore, the unimolecular reactions of localized systems, which confer the benefit of fast computing independent of concentrations, may also present difficulties in cases where slow reaction rates are desirable. Finally, the maximum size of localized systems can be limited by the size of their organizing structure, although the ability to tile origami structures could ultimately make this limit irrelevant in practice.

Despite these drawbacks, the benefits of spatial organization have motivated a number of works involving DNA circuits localized to DNA origami. It has been shown that hybridization chain reactions [14] can be localized on a track strand [112] (Figure 4.1, Bui *et al.*) or on the surface of a two-dimensional DNA origami structure [113], carrying out signal propagation through junctions [114] (Figure 4.1, Chao *et al.*) and logic circuit computation [115]. The most complex circuit to date consisted of six two-input AND and OR gates, implemented with 17 localized hairpins [115] (Figure 4.1, Chatterjee *et al.*). Six of these hairpins received four unique input signals and produced two unique output signals. Two common hairpins represented an intermediate signal and a threshold, respectively; these hairpins were reused in logic gates at various locations in the circuit. Overall, eight distinct localized components were used to build several logic circuits with different functions. The circuits produced output signals in roughly ten minutes [115], which was at least ten times faster than well-mixed circuits that computed similar functions [23]. The localized circuits were present in solution at just 5 nM with diffusible reporter, input, and fuel species included at concentrations ranging from 40 to 200 nM; the well-mixed circuits were run with gates at a standard concentration of 100 nM with some gate, threshold, fuel, and reporter species included at concentrations ranging from 60 to 200 nM.

A limitation of the localized logic circuits shown in Figure 4.1, Chatterjee *et al.* is that diffusible fuel molecules were used to facilitate localized signal processing. When the fuel molecules were at a relatively high concentration (200 nM fuel vs. 5 nM origami) and the number of reaction steps was small (three to seven), the bimolecular reaction rate involving a free-floating fuel could be similar to or even faster than the unimolecular reaction rate involving co-localized molecules on a surface [115]. However, significant slowdown will occur for larger circuits. For example, the bimolecular reaction rate will become ten times slower for a circuit with 36 reaction steps as the computation approaches completion because the overall fuel concentration will decrease over time. This limitation could be addressed by decreasing origami concentration and increasing fuel concentration, but the former is limited by the capacity of the readout mechanism to distinguish signal from noise, while the latter may be limited by cost, leak, or crowding effects depending on the circumstances and whether leakless design [43] is applied. Moreover, signal loss limited the scalability of these localized DNA circuits. For example, Chatterjee *et al.* observed that reaction completion level decreased roughly linearly with circuit size, approaching 70% signal loss in their largest circuit (Supplementary Figure S29 in Reference [115]). Presumably, this signal loss was due to the lack of full incorporation of localized components (see Appendix A.1, Staple Incorporation Background).

Signal loss in well-mixed circuits [15] was overcome by incorporating a catalytic mechanism. [23] Catalytic mechanisms, including catalytic hairpin assembly reactions [18, 92] and toehold exchange reactions [23, 78], have been explored to improve the signal loss in localized DNA circuits [118, 120, 121] (Figure 4.1, Ruiz *et al.*, Wei *et al.*, Shin *et al.*). However, in these systems, the signals are diffusible. When a signal is released from a gate, it can either be captured by a nearby downstream gate or diffuse away [117, 119] (Figure 4.1, Teichmann *et al.*). A diffusible intermediate signal is more likely to be lost to



Localized gates and diffusible signals

Challenge: leak between localized gates is substantial if the distance is small, and signals are mostly lost to solution if the distance is large.

Localized signals and diffusible fuels

Challenge: reaction rate depends on fuel concentration, and rates will slow in larger circuits as fuels are consumed.

All components localized

Challenge: leak between localized molecules must be minimized before reusing components for larger circuits.

Figure 4.1: **Previous work on localized DNA strand displacement circuits.** Works are arranged by localized species set size (x axis) and localized system size (y axis). The outer dot color indicates type of localization and the inner dot color indicates type of computation. To demonstrate that components can be reused in localized DNA circuits, it is desired to build large systems (top of plot) with interesting computational behaviors (inner circle) using few distinct components (left side of plot).

In no set order, works referenced are acknowledged as follows. All permissions conveyed through Copyright Clearance Center, Inc. Liu *et al.*: [116] (CC BY-NC 4.0). Teichmann *et al.*: reprinted with permission from [117] (©2014 American Chemical Society). Bui *et al.*: reprinted with permission from [112] (©2017 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim). Ruiz *et al.*: reproduced from [118] with permission from the Royal Society of Chemistry. Chao *et al.*: reproduced from [114] with permission from Springer Nature. Chatterjee *et al.*: reproduced from [115] with permission from Springer Nature. Wang *et al.*: [119]. Shin *et al.*: [120]. Wei *et al.*: [121].

solution than to interact with its intended target (Figure 3 in Reference [117]). Furthermore, when an intermediate signal diffuses away, it may be captured by a downstream gate on a different origami surface, which will cause errors in parallel computation. To prevent diffusion of intermediate signals, decreasing the distance between co-localized gates was tried, but then substantial leak between them was observed (Supplementary Figure S5 in Reference [117]).

Given this observation, leak caused by toeless strand displacement between co-localized components should be considered carefully. In well-mixed systems, toeless reactions can often be neglected since they are orders of magnitude slower than reactions including toeholds; without a toehold, the first step of branch migration requires the bimolecular joining of two species, while with a toehold, the first step of branch migration is unimolecular. In contrast, for co-localized components, the first step of branch migration is unimolecular whether or not a toehold is present. For any reaction to take place, rather than needing to overcome the relatively large entropic barrier represented by two molecules joining together, the molecules need only overcome the relatively small entropic barrier represented by two free ends of co-localized species joining together to form a loop on the surface. Therefore, while localized toeless strand displacement should still be expected to be much slower than localized toehold-mediated strand displacement, the ratio between the two rates may not correspond to that observed for well-mixed systems. Furthermore, in well-mixed systems, all reactions begin roughly simultaneously at the time all species are mixed. Conversely, in localized systems, toeless reactions might begin at the time of localized system assembly. Especially given that localized systems involving DNA origami typically require purification steps after initial assembly, it is desirable to produce localized systems which do not begin to leak immediately after they are assembled.

One design for a localized DSD reaction system employed co-localized gate and fuel molecules, but no quantification of leak between these species was reported [118]. Given that toeless leak between single-stranded and double-stranded species is faster than toeless leak between two double-stranded species in well-mixed systems, we expect co-localized gate-fuel leak to be worse than co-localized gate-gate leak. Confirming this hypothesis was a secondary goal for this work.

In well-mixed DNA circuits, leak reduction mechanisms such as doubled long domains [43] and clamps [23, 44] can be applied to reduce gate-gate leak; these mechanisms could be investigated in localized DNA circuits. Several approaches have been reported for reducing gate-fuel leak in well-mixed

DNA circuits, including mismatches [122, 123] and interfering strands [124]. However, the improvement in leak was limited and the performance of the catalyst was simultaneously reduced.

In addition to diffusible fuels and diffusible signals, another type of diffusible intermediate species was utilized in work demonstrating a DNA finite state machine with all components localized on a DNA origami surface [116] (Figure 4.1, Liu *et al.*). In this case, each state transition involved a reaction with a new input signal. By definition, input signals are received from a molecular environment and thus must be diffusible. For systems involving a new input at each step, computation speed is limited by input concentration, and will decrease with an increasing number of sequential input signals.

Apart from the motives to avoid diffusible signals and fuels outlined above, we envision a long-term aim that will not be possible if diffusible components are involved. Inspired by living cells, spatially organized molecular computation could eventually enable the construction of synthetic autonomous molecular machines. To achieve their full potential, such machines must be able to maintain their function while exploring a variable environment. They therefore must be designed to carry everything they need with them. This constraint could be achieved by encapsulating diffusible components within a closed volume such as a lipid nanoparticle or an origami box; however, such an approach requires mechanisms for transmitting signals between the environment and the interior. To avoid this complication, in our designs, we do not allow diffusible components apart from one or more initial trigger(s) representing environmental input, and one or more final reporter(s) representing an environmental target.

In contrast to all examples of localized circuits discussed previously, there exist demonstrations of circuits with all signal propagation steps involving localized components only [112, 113] (Figure 4.1, Bui *et al.*). In these examples, all circuit components are unique and no components are reused at distinct locations, which poses a challenge for scaling complexity: all toehold sequences must be carefully designed to encode specificity and avoid leak. We aim to develop circuits where components with the same sequences can be reused at distinct locations. In addition, these examples did not employ any localized signal restoration mechanism, and would thus be expected to suffer from signal loss with increased length. They did not seek to measure this property, and did not vary the number of localized gates to explore it. Nevertheless, given the signal loss observed in other systems [115, 117], the need for signal restoration is clear.

To summarize, catalytic mechanisms are essential for addressing signal loss, and excluding diffusible intermediate components from localized DNA circuits is important for the speed, parallelism, and environmental orthogonality of computation. Both of these criteria will improve the scalability of localized DNA circuits and allow for a demonstration of robust circuit computation with increasing complexity built from the same small number of components. Hence, these were the original primary goals of this work.

Our aforementioned secondary goal took precedence after our hypotheses were better informed by experimental data: to achieve localized signal amplification, a quantitative understanding of leak between localized gate and fuel molecules must be obtained and an effective leak reduction mechanism must be identified.

4.2 Results and Discussion

Pins-and-needles circuits

We developed a design for a localized catalytic DNA circuit nicknamed “pins-and-needles" which involved reactions between hairpins and single-stranded extensions (Figure 4.2).

The basic reaction mechanism is similar to a seesaw motif [23] except that the two-stranded gate complexes are converted to hairpins by linking the 5’ end of the gate strand to the 3’ end of the signal strand in a hairpin loop (Figure 4.2a). The gate strand is then extended with a 20-nucleotide domain to hybridize to the DNA origami surface via staple extensions. This property enables the gates to change position, like DNA walkers, as will be discussed below. The fuel strand also has an extended domain complementary to the extension on the gate strand and is directly localized on origami surface as staple extensions. When an input trigger is added, it reacts with the input gate to activate a localized input signal, changing the circuit from state S0 to state S1 (Figure 4.2b). In state S1, the input signal reacts with the output gate and activates the output signal, resulting in S2. In state S2, the fuel strand reversibly binds to the open toehold on the gate, undergoes strand displacement in two directions, and relocates the still-active output signal while

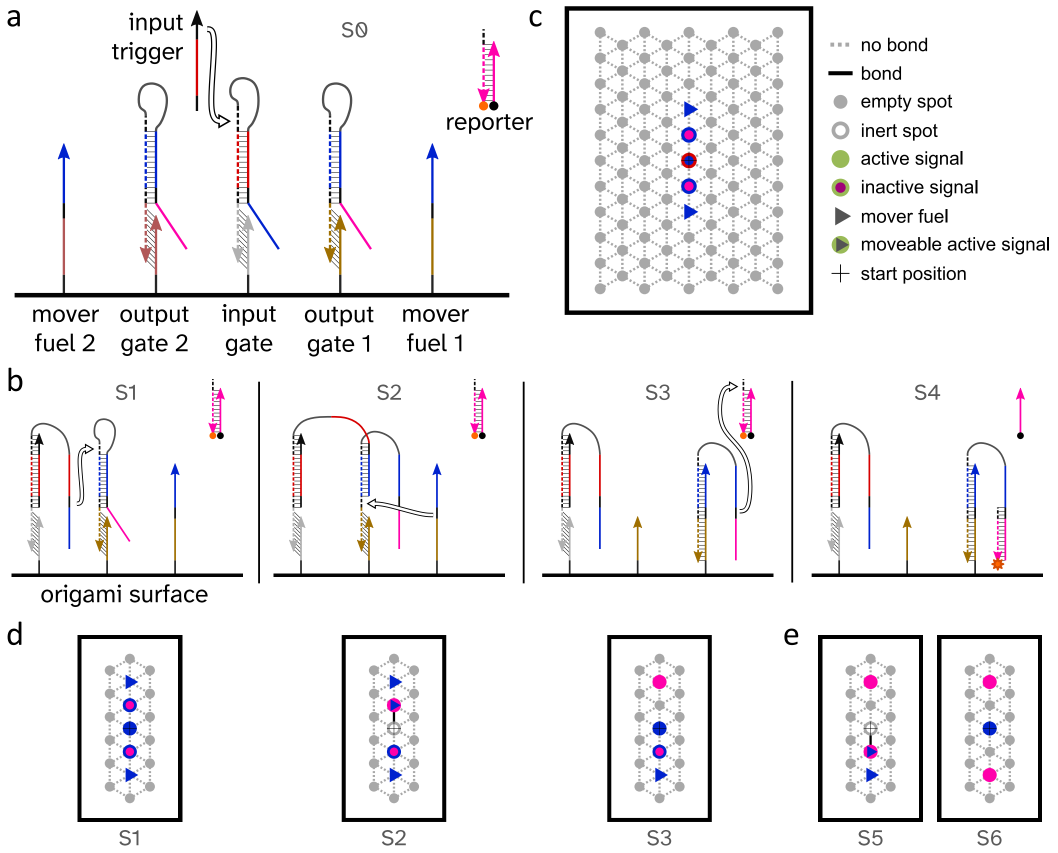


Figure 4.2: **"Pins-and-needles" design for a localized catalytic DNA circuit. a,** Domain diagram of a two-pathway catalytic circuit in its initial state S0. Colors indicate distinct domain sequences and their complements. Solid and dotted lines indicate domains with `ATC` and `ATG` sequences, respectively. **b,** Domain diagrams for the reaction steps in a single pathway that produces one output signal. The two pathways generating two output signals can take place in either order. **c,** Abstract diagram of the circuit shown as a top-down view of the DNA origami surface. Each symbol represents a distinct circuit component. Colors indicate the identities of signals and fuels. The legend applies to all abstract diagrams in this figure. **d,** Abstract diagrams that correspond to states S1 through S3. **e,** Abstract diagrams for the reaction steps in the second pathway, demonstrating the production of two output signals without consuming the input itself. The abstract diagrams in d and e show a partial area of the same origami shown in c.

simultaneously releasing and reactivating the input. These interactions produce state S3. In state S3, the output signal triggers the reporter, increasing fluorescence and changing the circuit to state S4. All reactions are reversible except for the reporting reaction. The reporting reaction can be approximated as irreversible, but for an in-depth discussion, see Appendix G, Latch With

Modified Reporter. The reporting reaction can also occur at state S2, leading to an alternative reaction pathway with the same final state. This alternative reaction pathway does not meaningfully affect the function of the system, as discussed in Appendix C, Alternative Reaction Pathway for Pins-and-Needles.

After a single reaction pathway reaches S3, the output signal from this pathway is activated on the surface and moved to a new location. The released input signal remains at its original location and can activate a second output. After two reaction pathways, the circuit will generate two copies of the output signal, doubling the signal strength compared to the input and preserving the input for participation in other tasks.

We developed an abstraction to facilitate more complex circuit designs (Figure 4.2c). The abstraction includes components arranged on a hexagonal grid because the tethering sites are thus arranged on the origami structure we utilized. We chose to use the double-layer DNA origami that was previously used for a cargo-sorting robot [125] because it is rigid enough to prevent crosstalk between distant components while still maximizing usable surface area and providing a total of 95 tethering sites. Each active signal is represented as a solid circle. When a signal is bound to a gate, it is inactive and represented as a two-layer circle in which the color of the outer circle indicates the identity of the gate and the color of the inner circle indicates the identity of the signal. Each fuel is represented as a solid triangle. It is called a “mover fuel" because it moves an active output signal to its own location while releasing and reactivating an input signal. Each movable active signal is represented as a triangle inside a circle in which the color of the circle indicates the identity of the signal and the color of the triangle indicates the identity of the fuel with which the signal-tethering gate could react. The start position where an input gate is localized is represented by a cross. After the input trigger is added, no open toehold will be available on the gate and thus it is considered inert and represented as a hollow circle. When an active signal reacts with an adjacent gate, a bond will be made between the two locations, which is represented by a solid line.

The benefit of movable components—referring to the ability of the output gate to physically shift from its original position to that of the mover fuel—is that the circuit architecture is naturally compatible with both computation and molecular robotics. It has been shown both in theory [40, 126, 127] and in experiments [125, 128] that DNA walkers can carry out nanomechanical tasks such as molecular transportation while processing information and making decisions in response to a molecular environment. Potentially, the mover fuel in the pins-and-needles circuit allows for a signal to simultaneously function as a robot, moving around on a surface while seeking and transporting cargo molecules. Because the circuit is catalytic, one robot could activate multiple other robots, recruiting them for more complex tasks.

Correct assembly of the initial state shown in Figure 4.2a is critical for the localized circuit to perform the desired function. The hairpin gates will naturally form at a high temperature during annealing and reach their kinetically trapped states without undesired interactions with each other [18]. However, the fuel strands would compete with the intended staple extensions for binding to the hairpin gates, resulting in undesired output activation without the presence of input. Thus, we had to anneal the origami without the fuels and utilize an isothermal hybridization step after annealing to insert the fuels into the origami structure. For the M13mp18 scaffold, which forms significant secondary structure at room temperature, the efficiency of insertion is known to be lower than that of staple incorporation during annealing [125]. We therefore expected that 10 to 20% of the origami would have a fuel missing.

To understand the expected behavior of the circuit design, we developed a simple model to estimate the localized reaction rates. The bimolecular hybridization rate for diffusible DNA molecules is commonly estimated as 2×10^6 per molar per second [79]. To estimate the unimolecular hybridization rate for tethered DNA molecules on a surface, we exploited the method for calculating the hybridization rate of hairpin closing based on the thermodynamic measurements of hairpin loops [104] (Figure 4.3a). In this model, the entire structure connecting the two ends of a binding site is treated as a hairpin loop:

$$n = (\sum_i n_i^{ss} \times 0.43 + \sum_i n_i^{ds} \times 0.34 + d)/0.43$$

where n_i^{ss} is the number of nucleotides in a single-stranded domain, n_i^{ds} is the number of nucleotides in a double-stranded domain, and d is the distance between two tether locations on origami. Here, 0.43 nm per nucleotide and 0.34 nm per base pair are used to estimate the length of single- and double-a

Bimolecular hybridization rate : $k_{h2} \approx 2 \times 10^6$ /M/s

Entropy of a hairpin loop of length n:

$$\Delta S_{\text{hp}}(n) \approx \Delta G_{37,\text{hp}} \times \frac{1000}{310.15} = (6.3 + 2.44 \times R/1000 \times 310.15 \times \ln[n/30]) \times \frac{1000}{310.15}$$

Unimolecular hybridization rate: $k_{h1}(n) \approx k_{h2} \times e^{-\frac{\Delta S_{\text{hp}}(n)}{R}}$ /s

Dissociation rate of a double-stranded domain of length l: $k_d(l) \approx 10^{6-l}$ /s

Branch migration rate: $k_b \approx 1$ /s

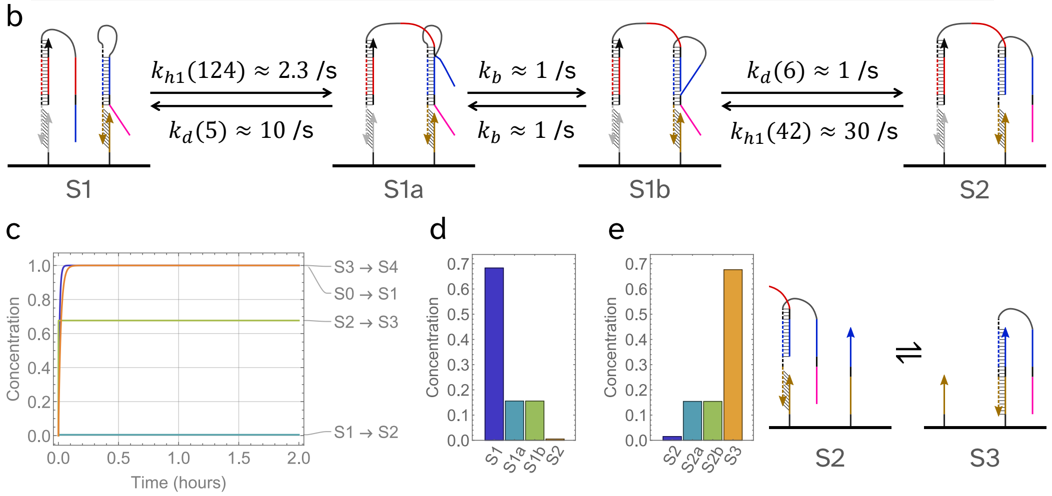


Figure 4.3: **Modeling the pins-and-needles circuit. a,** Formulas for the estimation of unimolecular hybridization, dissociation, and branch migration rates. The entropy of a hairpin loop of length n is calculated based on Equation 7 and Table 4 in SantaLucia and Hicks [104], where $n > 4$. The ideal gas constant $R = 1.9872$ cal/mol·K. The unit of $\Delta G_{37,hp}$ is kcal/mol. The unit of ΔS_{hp} is cal/mol·K. **b,** Estimated rates for the intermediate reaction steps of the localized input activating an output. **c,** Simulations of the four reactions shown in Figure 4.2 with mass-action kinetics. All simulations in this work were performed by solving ordinary differential equations using the CRN Simulator [22]. **d,** Steady-state concentrations of the intermediate states shown in b. **e,** Steady-state concentrations of the intermediate states from S2 to S3, in which a fuel moves the output to a new location while reactivating the input. The concentrations are relative to a standard concentration of 1× = 3 nM.

stranded DNA, respectively (see Appendix B, DNA Lengths, for the source of these values).

This is an extremely rough estimation as realistically single strands, double strands, and origami structures have very different structural flexibility and should not be treated the same. Additionally, the formula used to estimate the entropy of hairpin loops is used in SantaLucia and Hicks [104] to extrapolate from their original dataset for hairpins of loop size 3–30; a formula based on dataset(s) including larger hairpins would be expected to generate more accurate values. Nonetheless, this simple model allows us to make qualitative and semi-quantitative predictions about certain properties of a circuit design, to compare different designs, and to guide experiments.

For example, the unimolecular hybridization rate for the input binding to an output gate was calculated as 2.3 per second based on a loop size of 124, whereas that for the output rebinding to the gate was calculated as 30 per second based on a loop size of 42 (Figure 4.3b). Because the dissociation rate of a 5-nucleotide toehold (10 per second) is faster than the input binding rate, and the output rebinding rate is faster than the dissociation rate of a 6-nucleotide toehold (1 per second), taking the additional stacking bond into consideration, the overall reaction step of input activating an output is expected to be biased backward. Simulations confirmed that the step of input activating output (S1 to S2) would be biased backward whereas the step of fuel reactivating input (S2 to S3) would be biased forward (Figures 4.3c-e). This observation agrees with the fact that there is an entropic cost going from state S1 to S2 as a bond is introduced between the two tether locations (analogous to one fewer free molecule in a well-mixed system) and an enthalpic cost of losing one stacking bond. Simulations also suggested that the bimolecular reaction of output triggering a diffusible reporter would be slower than the unimolecular reaction steps on a surface (Figure 4.3c).

To investigate the behavior of the two-pathway catalytic circuit (Fig. 4.2), we performed fluorescence experiments with a negative control and three positive controls (Figure 4.4a). The negative control without an input trigger was used to quantify leak between circuit components (trajectory 5). A positive control was designed for each reaction pathway (Figure 4.4b), where a unique toehold was added to the 3’ end of each output gate and an output trigger was used to activate the output while detaching the gate from the surface (trajectories 3 and 4). Crosstalk between each positive control reaction was inhibited by using unique branch migration domain sequences for the region tethering the output gates to the origami (the brown domain in Fig. 4.4b adjacent to the additional toehold). A third positive control combined the two output triggers to fully activate both pathways directly (trajectory 2). By comparing the circuit behavior (trajectory 1) to the three positive controls,

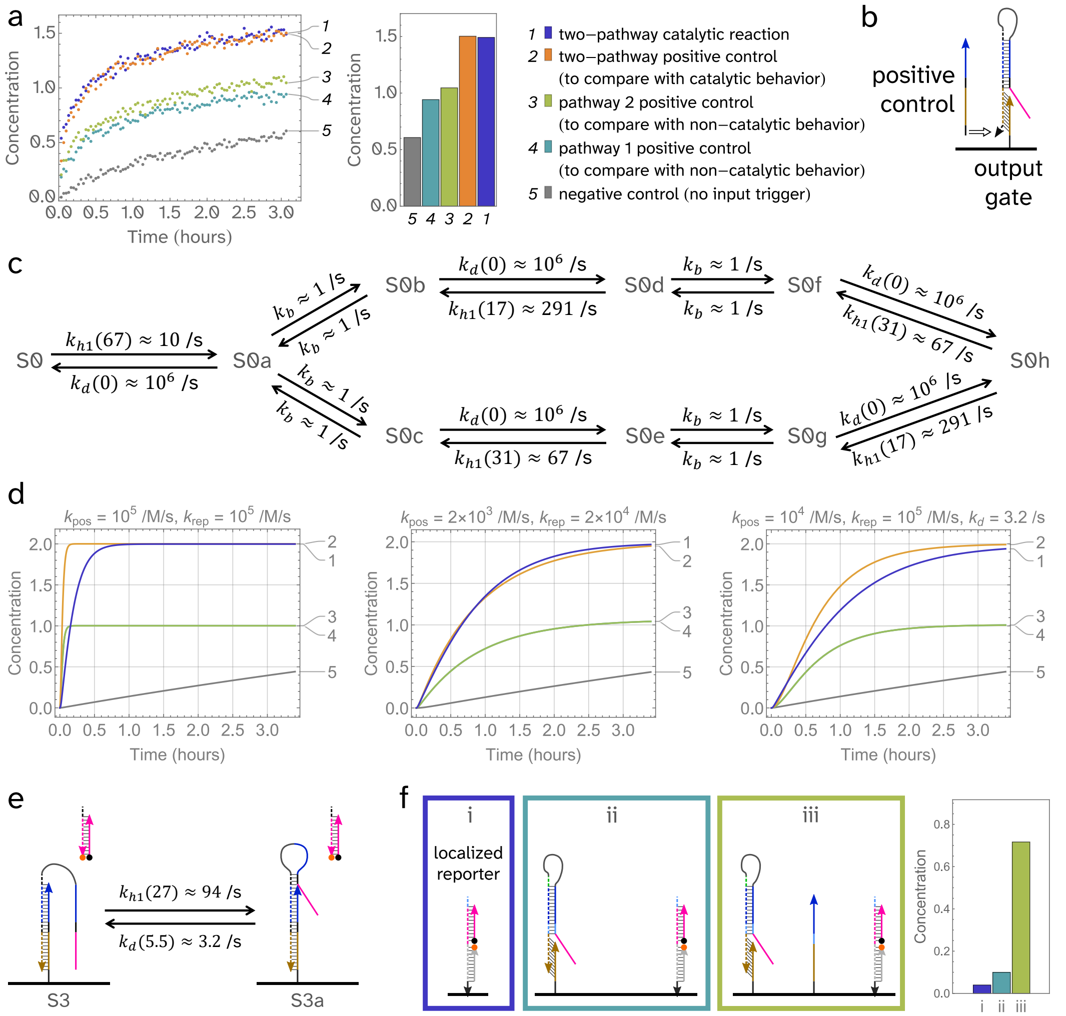


Figure 4.4: **Experiments and analysis for the pins-and-needles circuit. a,** Fluorescence kinetics experiments. Raw fluorescence was normalized to relative concentration based on positive controls at 24 h. Standard concentration 1× =3 nM. Data at 3.4 h are shown in the bar chart. The concentrations of DNA origami, input trigger, and positive control triggers were 1×, 50×, and 50×, respectively. **b,** Positive control design including an additional toehold. **c,** Estimated rates for the leak reaction between localized fuel and output gate. **d,** Simulations with three sets of rate constants for explaining the slow kinetics observed in the experiments. Trajectories 3 and 4 are identical. k_{pos}, k_{rep}, and k_d are the rate constants for the positive control triggering, reporting, and toehold dissociation shown in c, respectively. **e,** Toehold occlusion that explains the slow reporting reaction. **f,** Investigating the source of the leak using a localized reporter and no diffusible species. The concentration of DNA origami was 1×(3 nM). Raw fluorescence (t=1h) was normalized to relative concentration based on reporter calibration with varying amounts of diffusible output.

we would understand if the input is capable of functioning as a catalyst and producing two output signals.

As desired, when the input trigger was present, the output trajectory reached a similar signal level compared to the two-pathway positive control, which was roughly 64% higher than the input signal as well as the single-pathway positive controls (Figure 4.4a). However, a considerable amount of leak was observed in the negative control, which also influenced all other trajectories. Furthermore, the kinetics of the circuit, including the positive controls, were much slower than expected. These undesired traits were also present in a version of the system involving three mismatches in the central toehold of the input gate, discussed in Appendix D, Triple Mismatch Variant of Pins-and-Needles.

To understand the source of leak, we modeled the toeless strand displacement reaction between the localized fuel and output gate (Figure 4.4c). Simulations suggested that the reaction was indeed much faster than gate-fuel leak in a well-mixed system [23] (Figure 4.4d, left plot). To explain the slow kinetics of desired reactions, we explored two hypotheses in simulation. First, the reporting reaction (Fig. 4.2b S3-S4) was 5 times slower than expected and the output triggering (Fig. 4.4b) was 50 times slower than expected (Figure 4.4d, middle plot). This could be due to the electrostatic repulsion between origami and small diffusible molecules like the reporter and trigger strand. All DNA is negatively charged, and like charges repel each other. The impact of the collective negative charge of a DNA origami structure may be strongest near its surface. If so, we can explain why the trigger strand was affected more than the reporter: the added toehold on the output gate for the positive control was much closer to the origami surface than the relevant toehold on the output strand. Second, the activated output can still bind to the open toehold on the gate (Figure 4.4e), and this unimolecular toehold occlusion is much faster than the bimolecular reporting reaction. Assuming only one of the two hypotheses was true, we found the second hypothesis to be more likely, because it allowed for more realistic reporting and triggering rates (Figure 4.4d, right plot). However, it remains possible that some combination of the two hypothesized phenomena affected the system.

We also performed fluorescence experiments to further investigate leak using a localized reporter that provided a fast irreversible drain (Figure 4.4f). In this case, all reactions take place on the surface without a diffusible trigger, diffusible reporter, or any other diffusible species. Fluorescence kinetics trajectories were flat over several hours, and thus data at 1 hour are shown in the bar chart. See Appendix E, Localized Reporter, for the full trajectories.

In this set of experiments, a negative control was performed using the localized reporter by itself; this establishes the background. The second experiment served to quantify what leak, if any, can directly occur between the output gate and the localized reporter. They are expected to reach each other geometrically, but the absence of a toehold and the tendency of hairpins to close should normally not allow a reaction. Indeed, a small 6% leak above the background was observed. The third experiment quantified the leak between the output gate and the mover fuel. A large 68% leak above the background was observed, which was likely the entire population of origami with all three components (fuel, output gate, and reporter) incorporated (Appendix A.1, Staple Incorporation Background).

A natural approach to diminish the leak would be to increase the distance between the mover fuel and the output gate. However, the increased distance would also affect the desired reaction of fuel moving the active output while reactivating the input; in this circuit design, the physical distances involved in the leak and desired reactions are unavoidably coupled. Certainly, there is no way to ensure that the molecules involved in the leak cannot physically reach each other without causing the same change in the molecules involved in the desired reaction.

In other words, there is no way to change the reachability of one pair without changing the reachability of the other, where reachability is the property of two circuit components that describes whether their complementary binding domains can physically come into contact.[4] In a robust design for a localized DNA circuit, any two components that are not designed to interact should not be able to reach each other; this property is not achieved in pins-and-needles.

Given our expectation that any modification to the pins-and-needles design would represent a trade-off resulting in marginal improvements at best, we developed an alternative circuit design that specifically changes the reachability of localized circuit components at different reaction states.

Chapter 5

LOCALIZED CATALYTIC DNA CIRCUITS FOR SCALABLE MOLECULAR COMPUTATION ON A SURFACE PART 2: LESSONS FROM AN IMPROVED DESIGN

5.1 Introduction

Chapter 4 detailed our efforts to construct a localized catalytic DNA circuit using a Pins-and-needles design, which we found suffered from problematic leak. We transitioned away from the Pins-and-needles design in 2022 after identifying the cause of the leak and predicting that any modifications to the design to slow the leak reactions would equally slow the desired reactions.

Despite these challenges, we remained dedicated to the original goals of the project: to develop a catalytic mechanism for localized DNA circuits, and to do so without utilizing any diffusible intermediate components. Contemporaneous advances establishing the biological compatibility of DNA nanomachines (discussed in Chapter 1) reinforced our convictions that artificial autonomous molecular machines, or DNA nanorobots, might be made possible with advances in DNA circuits localized to DNA origami.

Thus, in order to further this vision, we set out to implement a localized catalytic circuit incorporating a robust leak reduction mechanism. To that end, we developed a geometrically-constrained design; this design superficially resembles a swinging bar door latch, and thus we refer to it as the “latch" circuit.

5.2 Results

Latch circuits

In the latch circuit, the gate complex is not a hairpin as in pins-and-needles, but rather a two-stranded complex as in the well-mixed seesaw motif (Figure 5.1a). Also unlike the pins-and-needles circuit, the gate and output strands are tethered to two distinct sites. The toeless leak reaction between the fuel and the output gate is not possible due to a geometric limitation: for branch migration to complete, the 5’ end of the fuel strand, which is very close to the tethering site, would need to reach the 3’ end of the gate strand, which is approximately two sites away. Furthermore, given our typical tether lengths of 0 for the fuels and outputs, the latch design prevents the toehold of the fuel from reaching the closed toehold on the gate as it transiently breathes open, such that toeless strand displacement cannot even initiate. However, when the localized input (for clarity of presentation, shown as a diffusible molecule) displaces the output strand from the gate complex, it breaks the bond between the gate and output while forming an input:gate complex fully within the reachable region of the fuel. Only once the second state has been reached can the fuel interact with the gate, which is when such an interaction is desirable. At this point the fuel can displace the input, reactivating it while forming a bond with the gate. The active input can then proceed to trigger a second pathway in the catalytic circuit.

Without the use of hairpins, the gate strand would not exclusively bind to the output strand during annealing, as it is equally likely to form a two-stranded complex with the input or fuel strand. To ensure correct assembly of the circuit, inhibitor strands were used to cover up some of the fuel and input domains (Figure 5.1b and Appendix M, Detailed Latch Diagrams). Specifically, the inhibitor strands covered all domains which were not the branch migration domain shared by the input, fuel, and output (the blue domain); had they covered this domain as well, they would have been more likely to bind to the wrong strand during annealing. Additionally, the activator strands that would have been needed to remove them would have been capable of triggering leak reactions. With the inhibitors, the designed initial state of the circuit is the same as the expected state at equilibrium. Assuming our annealing protocol was slow enough, this state should therefore be achieved during annealing. In addition to allowing for one-pot annealing in circuit assembly, the fuel inhibitor also makes it easier to measure the kinetics of leak between fuel and output gate by introducing a fuel activation step at any desired time point during an experiment.

Unlike the pins-and-needles circuit, modeling of the latch circuit predicted that the forward and backward rates for all localized reaction steps are similar to each other (Figure 5.2a). This prediction makes sense because there is approximately[5] no entropic or enthalpic costs going from one state to another as the number of bonds between localized components remains constant and no stacking bond is gained or lost throughout the entire reaction pathway

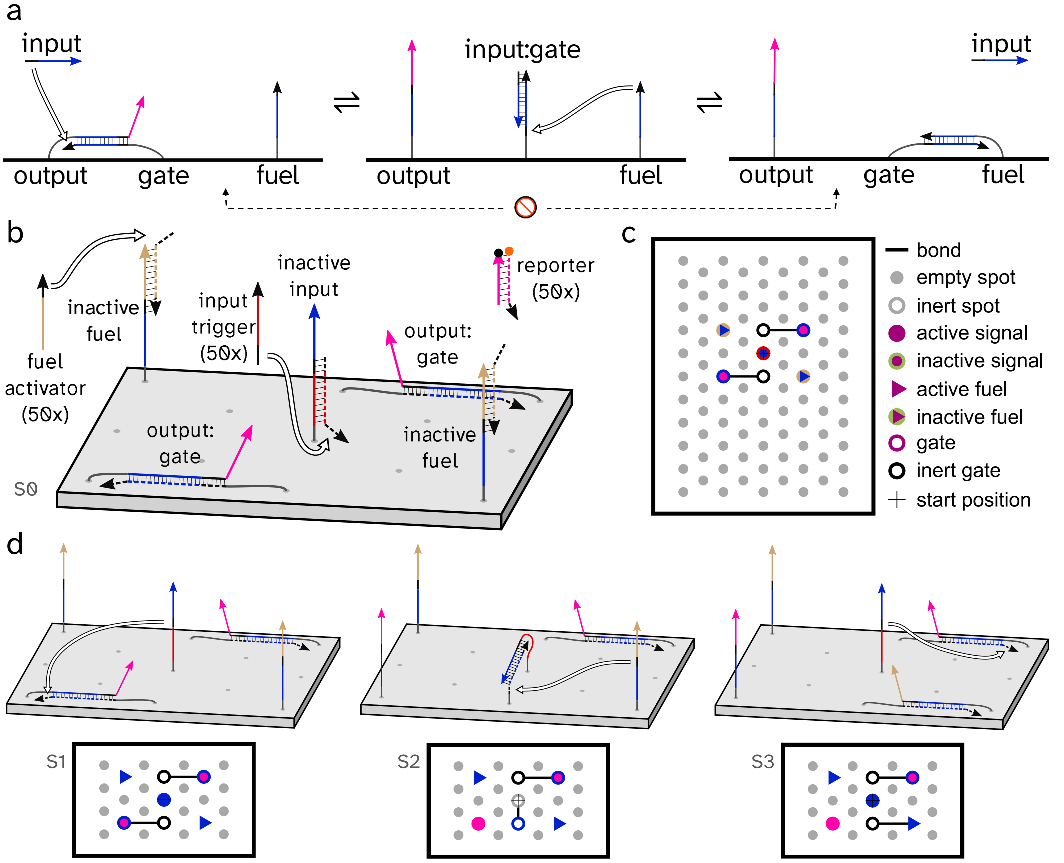


Figure 5.1: **“Latch" design for a two-pathway localized catalytic DNA circuit. a,** Reaction steps in a single catalytic pathway. As a simplification for clarity, input is shown as diffusible here, even though in reality it is localized. Tethers in these diagrams, the grey domains that directly extend from the origami, are all depicted as the same length, but in all experiments covered in this thesis, input and gates use 13 nt tethers while the outputs and fuels do not have any tethers. **b,** Domain diagram of a two-pathway catalytic circuit in its initial state S0. Input is now shown as localized on the origami surface. From state S0 to S1, the input trigger activates the input, and the fuel activator removes the fuel inhibitor. **c,** Abstract diagram of the circuit shown as a top-down view of the DNA origami surface. Each symbol represents a distinct circuit component. Colors indicate the identities of signals and fuels. The legend is reused for all abstract diagrams in this figure. **d,** Domain and abstract diagrams for the full set of reactions in one pathway. In state S1, the input activates an output from one pathway. In state S2, the fuel releases the input from the gate, reactivating it. In state S3, the reactivated input can trigger the second pathway. The two pathways can take place in either order. All reactions are reversible. The abstract diagrams in d show a partial area of the same origami shown in c.

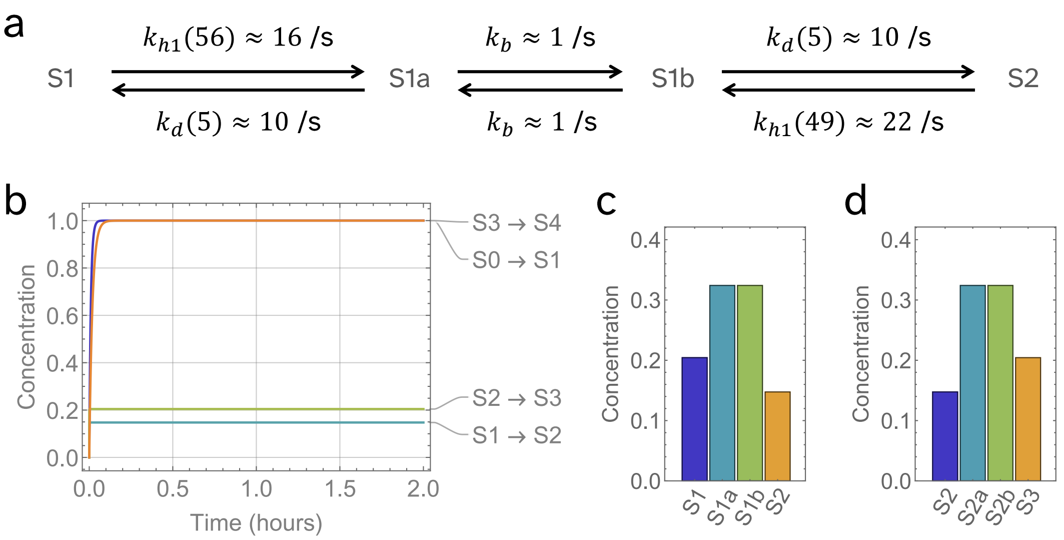


Figure 5.2: **Modeling the latch circuit. a,** Estimated rates for the intermediate reaction steps of the localized input activating an output (state S1 to S2 shown in Figure 5.1). For each hybridization step, the entire structure connecting the two ends of the binding site is treated as a hairpin loop. **b,** Simulations of the four reactions shown in Figure 5.1 with mass-action kinetics. **c,** Steady-state concentrations of the intermediate states from S1 to S2, in which the input releases an output while becoming bound to a gate strand. **d,** Steady-state concentrations of the intermediate states from S2 to S3, in which a fuel binds to the gate while reactivating the input. The concentrations are relative to a standard concentration of 1× = 3 nM.

(Figure 5.1d). Steady state concentrations of active output at state S2 (Fig. 5.1d) and active input at state S3 (Fig. 5.1d) are still relatively low in simulations because over 60% of the molecules are at the intermediate states of branch migration (Figures 5.2b-d). This result is consistent with the experimental observation in a toehold-exchange-based DNA walker on origami [129].

Fluorescence kinetics experiments were performed to quantify the behavior of the latch circuit (Figure 5.3). The fuel activator was added at 3 hours after the input trigger was added so that we could probe the catalytic characteristic of the two-pathway circuit. Before the fuel activator was added, the input on each origami could only activate one of the two outputs. Each output was designed to trigger the same reporter. Therefore, the collective signal level of the origami population was not expected to exceed the signal level of the input; in other words, we would expect the signal level before 3 hours to be similar to the output of a single-pathway circuit. This was true in the experimental data (Figure 5.3c, 0 to 3 hours, comparing to Figures 5.3a and b).

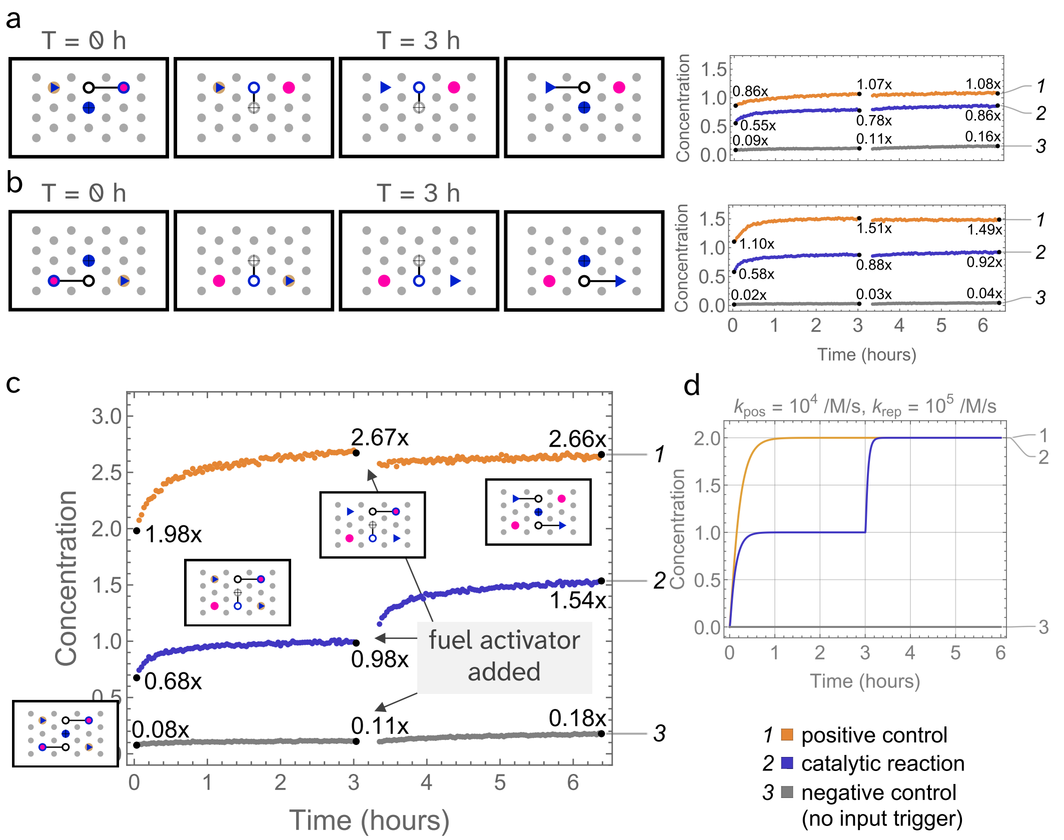


Figure 5.3: **Experiments and analysis for the latch circuit. a-c,** Abstract diagrams and fluorescence kinetics experiments for the first (a) second (b), and both (c) reaction pathways. Raw fluorescence was normalized to relative concentration based on positive controls at 24 hours (Appendix F.1, Latch Normalization). Standard concentration $1\times \approx 3.0$, 3.7, and 2.4 nM, respectively, in a, b, and c. The concentration of DNA origami was $1\times$ and that of input trigger, fuel activator, and output trigger (used for positive controls, see Fig. M.1) was $50\times$. **d,** Simulations of the ideal circuit behavior for comparing to the experimental data. k_{pos} and k_{rep} are the rate constants for the output triggering (used for positive control, see Fig. M.1) and reporting reactions, respectively. The concentrations are relative to a standard concentration of $1\times = 3$ nM.

After the fuel activator was added, the output signal increased by roughly 1.57-fold (from $0.98\times$ to $1.54\times$), demonstrating the catalytic behavior of the circuit. A more rigorous metric of 60.9% signal amplification takes leak into account. See Appendix F.2, Quantitative Analysis of Latch Experimental Data, for an explanation of how this metric and others were calculated. Even though the increased output signal ($1.54\times$, Fig. 5.3c) was not as high as the positive control ($2.66\times$, Fig. 5.3c and Fig. M.1), it was only 13% below the sum of the two output signals produced in the two single-pathway experiments ($0.86\times + 0.92\times = 1.78\times$, Fig. 5.3a and 5.3b).

A modest improvement in these values was obtained by using strands ordered PAGE-purified, as discussed in Appendix H, Latch With Purified Components. Signal amplification increased to 67.3%, and the agreement between the output signal and the positive control was improved (+12.8%).

Unlike for the pins-and-needles design, simulated kinetics for the catalytic reaction agreed well with the experimental data (Figure 5.3d), suggesting that the reporting reaction was indeed much faster without intramolecular toehold occlusion in the output strand.

Importantly, the leak was significantly reduced in the latch circuit compared to the pins-and-needles circuit. Roughly 2.7% (Appendix F.2, Quantitative Analysis of Latch Experimental Data) of leak above background in the negative control experiment was detected 3 hours after the fuel activator was added (Figure 5.3c). Interestingly, the leak and reaction completion levels were different for the two single-pathway circuits (Figures 5.3a and b). One possible explanation for this is that staples at different locations in an origami structure are known to have different incorporation efficiencies, and that staple extensions occasionally become sequestered within an origami structure (Appendix A, Staple Incorporation Rates). Missing staples and sequestered extensions can also explain why the circuit output in the two-pathway circuit failed to double after adding the fuel activator. Both fuel strands are needed to ensure full catalysis takes place, so the possibility of a missing or sequestered fuel could compound the possibility of a missing or sequestered input.

If low staple incorporation rates were the only possible explanation for these deviations from simulation, then we would estimate that in the single-pathway circuits, 30 to 40% of the origami either had a missing or sequestered input. This value falls within an established range, [130] so on first inspection, it may appear like a full explanation.

On closer inspection, however, two peculiar behaviors cannot be explained by the absence or sequestration of any staple extension:

1. The completion level of the positive control curves always exceeded the

expected value of 2× origami concentration.

2. The kinetics of the positive control curves are much slower than expected in all cases, and differed between the top and bottom pathways in the single-pathway circuits.

To verify that these behaviors were not just an artifact of this set of experiments, we performed more experiments with the latch relocated to different positions on the origami (Appendix I, Latch Relocated), which corroborated this finding.

We performed several investigations to explain these behaviors. First, we checked our assumption that a large concentration of our reporter was a sufficient sink to drive forward all reversible reactions to completion. We confirmed that these were valid assumptions experimentally (Appendix G, Latch With Modified Reporter).

Next, knowing that industrial DNA synthesis is often imperfect, we checked a choice we had made based on information provided by our source of DNA, Integrated DNA Technologies (IDT). We assumed we could order unpurified DNA while avoiding truncation in staple extensions by extending staples only at their 3' ends with circuit components. We confirmed this design decision was sound, in agreement with IDT's claims (Appendix K, Latch With Staple Orientation Reversed). Additionally, while using strands ordered PAGE-purified did confer some meager improvement, it did not truly address the unexpected behaviors of the latch (Appendix H, Latch With Purified Components). Therefore, we reasoned that the underlying issue was unrelated to strand purity.

We then checked whether the issue was instead related to origami purity. This work so far has only involved Amicon spin purification of origami, which eliminates excess staples and any other short DNA strands, but does not remove origami aggregates. In Thubagere *et al.* [125], a non-negligible fraction of the double-layer DNA origami structure was observed to form undesired aggregates during annealing. Agarose gel purification was applied to remove the aggregates, resulting in significant improvement in desired reaction completion (Supplementary Figure S4 in Thubagere *et al.* [125]). However, since that work, poly(`C`) extensions have been added to the edge staples of the double-layer origami [131]. Poly(`C`) extensions prevent aggregation by eliminating stacking bond interactions at origami edges. With this modification, we did not see as much evidence for aggregation in our agarose gels, and agarose purification did not improve the latch function (Appendix L, Agarose Purification of Origami).

Considering whether the problem could be related to an interaction between strands on different origami surfaces, we performed a set of experiments to measure inter-origami interactions (Appendix R, Quantifying Inter-Origami Leak). We did not observe any detectable degree of inter-origami interaction, suggesting that the origami concentrations are low enough such that any inter-origami interactions are negligible.

Ultimately, one experiment aiming to improve the slow kinetics of the positive control curves pointed toward the most likely explanation for the latch's flawed performance. Experiments with a modified positive control with an additional toehold domain demonstrated dramatic increases in kinetics and completion level (Appendix N, Modified Positive Control and Duplicated Complexes). The best explanation for this phenomenon is the duplication of some complexes on the surface (Appendix N.2, Complex Duplication). Unfortunately, because of the secondary structure of staples with homology to M13mp18, we were unable to directly quantify these duplicated complexes (Appendix O, Attempted Measurement of Duplicated Complexes with Staple Translators).

Duplicated complexes explain the initial rapid kinetics of the positive control which take place before plate reading measurements begin, followed by slow kinetics that vary by latch location. The fast kinetics correspond to the portion of the gate:output complexes that formed correctly, added to the portion of gate:output complexes that are tethered only by the staple on the output strand. The slow kinetics correspond to the fraction of gate:output complexes which are tethered only by the staple on the gate strand; the toehold on this complex is occluded by the dangling staple of the output strand (Fig. N.4). The leak which takes place after the fuel activator is added is likely caused by this type of malformed structure as well. Although pins-and-needles demonstrated that toeless leak can be fast on a surface, this particular reaction is worse than a toeless leak because of the interactions between the dangling output staple and the “open" toehold. This occlusion also explains why the % fuel leak varied by location on the origami surface (see Appendix I, Latch Relocated).

All attempts to eliminate the possibility of duplicated complexes failed (Appendix P, Attempts to Address Duplicated Complexes), primarily due to the constraints of the double-layer origami structure. While we were able to apply geometric constraints to eliminate the type of leak encountered in the pins-and-needles design, since the double-layer structure forms at a temperature lower than the hybridization temperature of the circuit components, it is not possible to apply geometric constraints to ensure proper assembly. More critically, since the scaffold for the structure is a viral genome and does not use a three-letter code, it has substantial secondary structure that can inhibit insertion of circuit components at low temperatures.

Complex duplication represents an assembly problem which must be resolved before the latch design can be expected to be cascadable. With each new component added to the origami surface, the number of possible duplicated complexes increases quadratically. Furthermore, duplicated complexes are not tethered to the surface at two locations, so they are not geometrically constrained to prevent leak; when cascading the latch, each component has more neighbors with which it can leak if the designed geometric constraints are missing.

New methods enable facile production of custom scaffolds, in stark contrast to their inaccessibility the onset of this project (discussed in greater detail in Chapter 6, Conclusions and Discussion). Taking advantage of this development could facilitate efforts to prevent complex duplication. Assuming this issue is resolved, the next step will be development of an irreversible translator; all of the localized reactions in this design for the latch are reversible, which is expected to pose a problem for scalability. To understand why, it is best to compare the circumstances in well-mixed and localized circuits. Unlike well-mixed enthalpy-neutral DNA circuits that can be driven forward by high concentration of gates, the localized enthalpy-neutral DNA circuits that we have demonstrated so far were mainly driven forward because of the high-concentration reporter that served as a downstream drain. Furthermore, whereas well-mixed enthalpy-neutral DNA circuits have a lower bound of output concentration in arbitrarily long cascades [44], we would expect that spatially-organized enthalpy-neutral circuits would slow down dramatically with longer cascades for the following reason. Because the number of gates in each layer is limited by the 2D geometry, using entropy to drive the reactions forward would consume physical space and reduce the complexity of computation which can be achieved within a given area. Instead, localized irreversible reaction steps will be required to facilitate the construction of complex circuits with long cascades.

Once irreversible steps are added, it will be possible to demonstrate robust signal propagation by incorporating signal restoration in linear cascades (Figure 5.4a). Note that the latch layouts in these hypothetical cascades are oriented in the *cis* rather than *trans* configuration; initial experiments showed that the latch performed somewhat worse in the cis layout (see Appendix J, Latch Re-Oriented), though that effect may vanish once duplicated complexes can be eliminated.

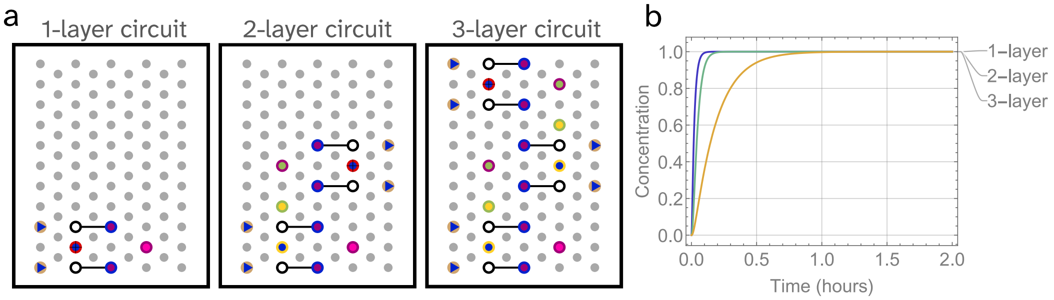


Figure 5.4: **Design and simulation of linear cascades in latch circuits. a,** Abstract diagrams. **b,** Simulations. The concentrations are relative to a standard concentration of $1\times = 3$ nM.

In the one-layer circuit (Fig. 5.4a), a latch has two outputs that feed into the same wire (implemented by an irreversible translator), such that the signal will be preserved even if only one of the outputs is activated. In the two- and three-layer circuits (Fig. 5.4a), the two upstream outputs feed into a wire leading to the downstream input. Ideally, the same completion level is shown in simulation regardless of the number of circuit layers (Figure 5.4b). Realistically, we expect that even if the catalytic circuits still suffer from signal decay in longer cascades, they will produce higher signals than non-catalytic cascades with the same number of signal propagation steps.

It will also be possible to demonstrate enhanced signal amplification and fan-out through branching cascades.

In a potential layout of a two-layer circuit (Figure 5.5a), an input is placed at the center of the origami (colored in blue at state S1). Each of the two

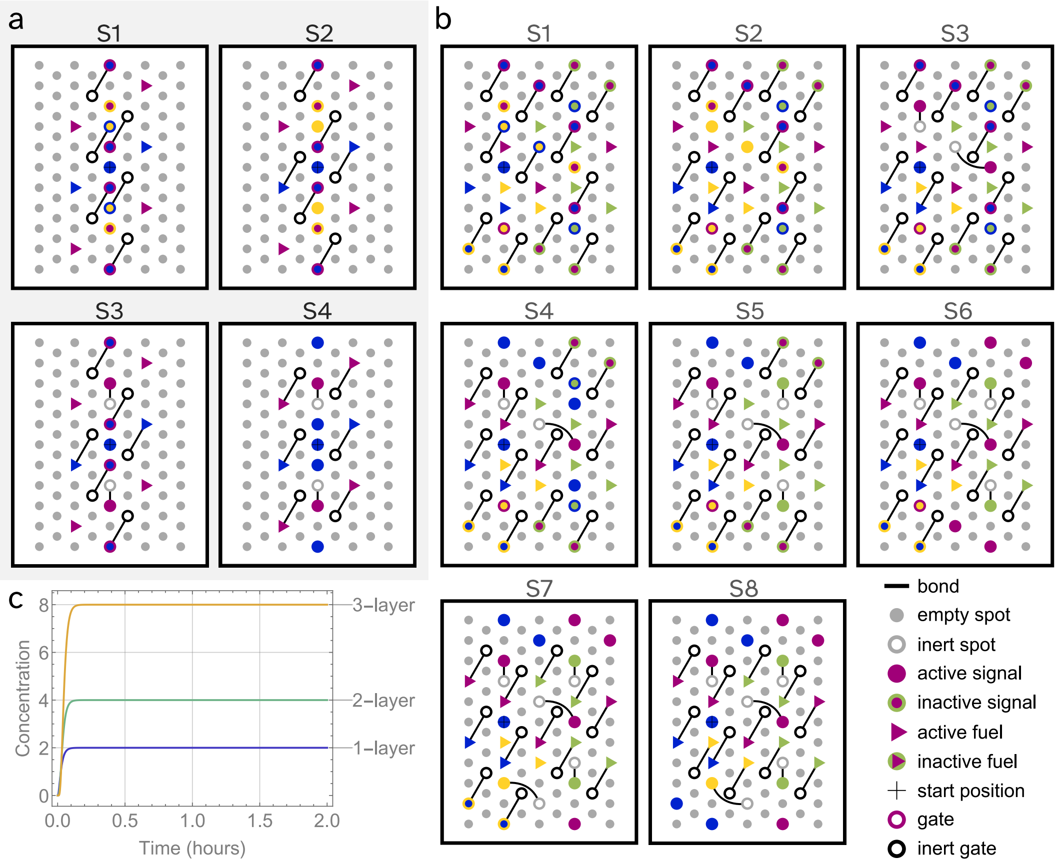


Figure 5.5: **Design and simulation of branching cascades in latch circuits.** **a,** Abstract diagrams of a two-layer branching cascade. Four representative reaction states are shown here. **b,** Abstract diagrams of a four-layer branching cascade, where the first to fourth layers have a fan-out of 2, 4, 4, and 2, respectively. Eight representative reaction states are shown here. **c,** Simulations of branching cascades with an increasing number of layers. The concentrations are relative to a standard concentration of $1\times = 3$ nM.

outputs of the upstream latch (colored in yellow at state S2) directly and irreversibly activates an input of the downstream latch (colored in purple at state S3). Two downstream latches then each produce two outputs that have the same identity as the input, totaling 5 blue signals and 2 purple signals at state S4. This layout involves three interweaving latches of two types, which requires adjacent gates and fuels to have distinct sequences in order to avoid undesired leak and crosstalk. Other layouts that use the space less efficiently will allow for circuits with the same function to be built with fewer distinct latches. Similarly, we designed a four-layer circuit that produces 13 output signals of 4 types (Figure 5.5b). Ideally, branching cascades can produce exponentially more signals with increasing layers (Figure 5.5c), but practically the degree of signal amplification and fan-out will be limited by what can fit on a given 2D origami surface.

5.3 Discussion

We have shown that the latch design is effective for reducing leak reactions in localized catalytic DNA circuits. However, it does not have movable components like the pins-and-needles design, and thus does not support molecular robotics in addition to circuit computation. Furthermore, more work is needed to establish a robust way to prevent complex duplication, and to enable detection of any dangling staples. Nonetheless, without involving intermediate diffusible species, both designs demonstrated localized signal amplification that is useful for improving the robustness and scalability of molecular information processing on a surface.

Well-mixed catalytic DNA circuits are primarily time-constrained, whereas localized catalytic DNA circuits are primarily space-constrained. Consider amplifying a single copy of a DNA strand to an amount that is measurable on a plate reader. For a well-mixed catalytic circuit, even if there is absolutely no leak, the time that it will take to amplify the signal is bounded by the bimolecular hybridization rate constant and the maximum concentration of molecules. In contrast, for a localized catalytic circuit, except for the first reaction step where the signal must land on the origami surface, the speed of reaction will not be limited by the concentration of the origami. However, the degree of amplification that can be achieved is limited by the size of the origami structure. A 600,000-fold amplification demonstrated in well-mixed catalytic DNA circuits [93] would require DNA origami structures of sizes 15 times larger than the current largest uniquely addressable origami [132]. Fortunately, the size of DNA origami structures has been doubling roughly every 2.2 years [133], allowing us to dream about such possibilities in ten years. To achieve more complexity sooner, future work may take advantage of origami tiling to provide space for more complex circuits.

There would be another major advantage to moving toward circuits arranged on origami tiles. In previous work, our group showed that tile displacement allows for modular reconfiguration in DNA origami assemblies [30, 31]. We envision that this mechanism could be used to create dynamic behaviors in artificial cells (or cell-scale molecular machines), allowing them to reprogram their own functions by replacing parts of their structures and better perform distinct tasks in response to a changing molecular environment. A key challenge is that the DNA origami tiles in these reconfigurable systems have yet to be functionalized with localized DNA circuits that control how the molecular machines make decisions. Signal insulation across DNA origami tiles is needed in these localized DNA circuits such that after one tile has been replaced with another, the signals on neighboring tiles will remain active to propagate and compute a different function. The catalytic circuits demonstrated in this work can provide the desired signal insulation and open doors to new developments in reconfigurable molecular machines.

5.4 Materials and Methods

Annealing

Unless otherwise noted, annealing for all DNA complexes and origami was performed in an Eppendorf Thermocycler. The lid was held at 105°C. The samples were heated to 90°C, held for 5 minutes, then slowly cooled to 20°C at a rate of 0.1 degree every six seconds. For complexes, annealing took place inside Eppendorf LoBind 0.5 mL tubes.

Nanodropping

The only measurement we used from our Nanodrop was the A260 reading. After taking three measurements of one droplet, or three measurements of three droplets (9 total) when accuracy was critical, the A260 values were averaged. This value was divided by the absorbance path reading of 1 cm and an extinction coefficient calculated for each strand or complex. The output of this calculation was a molar concentration.

DNA Origami Preparation

Scaffold M13mp18 was ordered from Bayou Biolabs and used without any further purification. Unmodified staple strands were ordered unpurified from IDT in an Echo-qualified 384-well source plate. Modified staple strands and staple strands with extensions were ordered from IDT and added to a separate Echo-qualified 384-well source plate (Labcyte Cat. P-05525). Staple mixes were generated by using an Echo 525 acoustic liquid handling robot to transfer from these two source plates to a destination plate (typically either

Eppendorf 96-well Twin.tec LoBind PCR plates, Skirted, or Thermo Scientific Armadillo PCR plates, 96-well, Skirted).

The source plates were stored at 4°C and sealed with Thermo polyolefin plastic plate seals (232701). After removing them from refrigeration, they were immediately spun at 1500 rcf for 5 minutes to allow them to come to room temperature while minimizing evaporative loss by condensation sticking to the plate seals. After removing the plate seal to add any additional components, a new plate seal was applied, followed by another 2 minute spin at 1500 rcf before Echo liquid transfer.

After Echo liquid transfer, plate seals were immediately applied to source and destination plates. The destination plate was immediately spun in a small tabletop plate centrifuge. Subsequently, scaffold was added to each relevant well in the destination plate, as well as necessary buffers to achieve a final annealing buffer of 10mM tris and 1mM EDTA at expected pH 8.0 with 12.5 mM Mg^{2+}. The total volume per well was always 50 μL. During annealing, scaffold concentration was typically $1\times =$58 nM, staple concentration was typically $5\times$, and concentration of any other species such as inactivators was typically $6\times$. A multichannel was used to pipette up and down to mix. Ahead of annealing, an aluminum plate seal (Thermo 232699) was used for the destination plate, which was spun down at 1500 rcf for one minute.

Unless otherwise noted, the annealing protocol for origami was the same as the one described above used for annealing complexes.

Following annealing, origami were immediately separated from excess staples and any other small DNA strands by filtration in Amicon Ultra .5 mL Ultracel 100K. The origami were transferred from the destination plate to the Amicon spin filter columns using a multichannel, as well as a 3D-printed tube rack designed to space the Amicon spin filter columns to every other multichannel tip. The multichannel, with its tips containing residual droplets of origami, was placed on a rack. A different multichannel was used to transfer 50 μL of annealing buffer to the just-emptied destination wells to collect any remaining droplets of origami. The original multichannel, with its original tips, was used to transfer the buffer containing any residual origami to the Amicon spin filter columns. An Eppendorf repeater pipette was used to transfer 400 additional μL of annealing buffer into the Amicon spin filter columns. Then, the Amicon spin filter columns were spun at 2500 RCF for 12 minutes for consistency with prior work. The effluent waste was removed and 500 μL of annealing buffer was added to wash the origami in the Amicon spin filter columns. The same spin conditions were used, the effluent was removed, the wash was repeated, and finally the inner spin columns were removed from the waste collection tubes and placed upside-down in elution tubes. Elution took place at 2500 RCF for 4 minutes.

DNA origami could then be nanodropped, with typical results of 40 nM calculated using an estimated extinction coefficient of 113,493,900 $L/(mol * cm)$.

Agarose purification

Unless otherwise noted, origami was used without any further purification. When agarose purification was used, we closely followed the same protocol from Reference [125].

Fluorescence kinetics experiments

Buffers for master mixes were prepared first, including 0.01 percent by volume Tween-20 to prevent DNA from sticking to tube walls. We did not evaluate whether a higher concentration of Tween-20 would provide better protection, or whether a lower concentration would be sufficient, and this aspect of our methods remains an open question. It would be desirable to decrease the concentration of Tween-20 if possible, because its soapy consistency increases the chance of pipetting bubbles. After preparing these buffers inside Greiner Bio-One Sapphire PCR tubes (Item No. 683201), the tubes were vortexed for 10 seconds using a Thermo Scientific Vortex Maxi Mix II at its maximum speed setting to allow the Tween-20 to coat the walls prior to adding any DNA. The tubes were spun down in an Eppendorf MiniSpin Plus benchtop centrifuge for 15 seconds at 8.0*10^3 rcf. Reporter was added to each master mix. Next, buffers and controls were added to Corning™ Low-Volume 384-Well Clear-Bottom Polystyrene Plates. Origami samples were added to the master mixes (still in tubes), which were then vortexed and spun down. The master mixes were distributed to wells in the 384-well plate, followed by positive control strands and finally trigger strands. In all cases, concentrations were achieved such that 20 μL of master mix was added to the wells and 10 μL of buffer (in the negative control case), positive control, or trigger strand. These values were selected to minimize pipetting error by using a 20 μL pipette and 10 μL pipette at the upper end of their ranges. A 10 μL 12-channel pipette set to 10 μL was used to pipette the wells up and down 20 times. The same polyolefin plate seals used for storing Echo source plates (Thermo polyolefin plastic plate seals (232701)) were used for sealing plates during reactions. The plates were spun down at 1500 RCF for 1 minute before beginning measurements.

Fluorescence values were measured on the same Biotek plate readers referenced in Chapters 2 and 3 using a gain of 100. The Biotek was set to record 100 measurements per data point—ten times our group's typical value—to account for the relatively low 1× concentration of 3 nM. This low concentration was selected to conserve materials, since origami requires more DNA to be used for every experiment compared to well-mixed systems.

Chapter 6

CONCLUSIONS AND DISCUSSION

The most difficult part of getting a latch to work effectively is ensuring its correct assembly. A major reason that latch assembly is so challenging is the lack of control we have over the order in which staples join an origami structure. If a large addressable structure could instead be produced by developmental self-assembly, as discussed in Chapter 2, then it would be much easier to guarantee that the correct circuit complexes form on the structure. Each component would join the structure in a desired sequence, resulting in a homogeneous product without any duplicated gates.

6.1 Model Comparison

In Chapters 4 and 5, we presented a simple model for predicting reaction rates on an origami surface. Much more advanced models exist, and each of these correspondingly takes more time to use. Had we spent time modeling our systems with one of these, we would not have been able to run as many experiments, but those we did run may have been better informed. The main reason we did not take this approach was that our problems were not closely related to the kinetics of the localized reactions, but rather to the completion levels and to the positive control kinetics (which pertained to bimolecular reactions). To help determine whether our simple model gives reasonable estimates for rates, I here compare the results it gives to rates produced in other works.

Our simple model does align well with both theory and experiment for DNA walkers modeled using molecular dynamics simulations (Figure 3c in Reference [134]) when the DNA walker must take a relatively short step size of 5-20 nm. For steps of 25 nm or larger, our model quickly diverges by orders of magnitude from both dynamic modeling and experimental results. While dynamic modeling and experiment both show the rate decreasing by orders of magnitude as the spacing increases, our simple model vastly overestimates the reaction rate. This initial match followed by divergence is easily explained. At short step sizes, the flexible portion on the surface is very large relative to the distance it must cross. Our simple model is based on hairpin closing, and a hairpin does not have any distance it must cross at all. Therefore, as the distance increases, the reaction stops resembling a closing hairpin. At a sufficiently large distance, the reaction will become impossible because the domains with homology will no longer be able to reach each other.

This insight dictates that our simple model might be reasonably applied for designs which are similar to the pins-and-needles motif, where the distances between each component is short compared to the lengths of the domains which must hybridize. Conversely, this model should not be expected to work well for designs which operate more like the latch, where the need for geometric constraints dictates that any toehold on an invading strand cannot reach much farther than its target open toehold on a gate. As a result, our simulations for the latch system should be considered only a qualitative foundation for a more rigorous simulation that is still needed. A different simple simulation that more accurately models systems similar to the latch would also be worth developing.

Our simple model does not give results of the same order of magnitude as those provided in past work using calculations of local concentrations [135] or structure sampling [136]. In the case of the older work using local concentrations, the authors clearly articulated that it was only a starting point that would need to be parameterized against experimental data. Therefore, it is not surprising that it did not match well, even when using their example calculations for a system with relatively long DNA and relatively short surface distances to cross. In the case of the newer work involving structure sampling, the system in question only involves relatively similar distances for what the toehold can reach and how far it needs to travel to find its match. Therefore, as discussed above, our simple model is not a good fit for the circumstances of the reactions.

Ultimately, I would not encourage any model to be parameterized based on our results with the latch system until the issue of duplicated complexes can be addressed. At a minimum, it will be necessary to be able to quantify duplicated complexes, which will require staples with less secondary structure. The best way to go about achieving that goal is discussed in the next section regarding origami scaffolds.

6.2 Origami Scaffold Choice

In our lab's experience, some seemingly minor details can have a major impact on whether an entire system functions correctly. For example, a winner-take-all DNA neural network depends on an annihilator with carefully balanced toehold energies [41], a cargo-sorting robot on a 2D surface will "teleport" through the third dimension if the surface is too flexible [125], and a design for indefinite array of triangular origami tiles will inadvertently form rhombic triacontahedra instead if one edge identity is used instead of two [29]. At a project's onset, we can anticipate some potential issues and employ careful design adjustments to resolve them. Other issues only become apparent the hard way: through numerous experiments with results that individually explain only a little, but collectively provide a new understanding about molecular behavior. That was certainly the case for the issue of leak in the pins-and-needles design, and for the issue of duplicated complexes in the latch design. To address leak in pins-and-needles, a major design overhaul was required that sent us back to the drawing board[6]. To address the issue in the latch, conversely, a different type of paradigm shift is required. Just as cargo-sorting DNA robots suffered from leak until a more rigid origami platform was developed, manufacturing localized catalytic DNA circuits without diffusible intermediates will remain challenging until a DNA origami platform with greater assembly control is available. Ideally, a structure that forms at a temperature higher than the melting point of circuit complexes would allow adding geometric constraints to the assembly process, such that only neighboring components may hybridize. A proposed design for such a structure is discussed in Appendix Q, A Staple Layout for a Proposed 14-Helix Origami Structure that Forms at High Temperatures. Additionally, use of a custom scaffold with less secondary structure would be invaluable for improving assembly and could convey additional advantages facilitating development of DNA nanorobots.

The implementation of a custom scaffold using a three-letter code to minimize secondary structure, often used in strand displacement systems [137], has been successfully demonstrated [138–140]. Some of the challenges I encountered—such as slow staple insertion rates, inconsistent Latch formation varying by location, and failed detection of dangling staples by translators—could potentially be alleviated by employing a custom scaffold specifically designed to minimize self-interaction. In a prior work from our group, it was suggested that tile displacement rates may be hindered by occlusion which could be addressed by a three-letter code [31].

Despite the advantages offered by three-letter codes, the M13mp18 viral genome remains widely used due to its low cost and the abundance of existing origami designs that utilize it. Custom-designed scaffolds are not commercially available and must be produced in-house, typically by transfecting bacteria such as *E. coli* with plasmids or phagemids [141].

For our project, we utilized the M13mp18 scaffold with the established double-layer origami design due to its convenience, our inability to foresee the challenges that would arise, and the limited availability of effective methods for producing custom scaffolds at the project's onset in early 2019. Although Kishi *et al.* demonstrated origami formed using a custom scaffold with a three-letter code in 2017, their scaffold was sourced directly from IDT [138]. As of 2024, ordering a gBlock for such a scaffold costs approximately $283 for ≈1 μg, whereas 500 μg of M13mp18 is available from Bayou Biolabs for just $150. This striking price difference—the custom scaffold costs ≈1000 times more per microgram despite being much shorter (1394 nt vs. 7249 nt)—compelled Kishi *et al.* to amplify the scaffold using PCR, a method not viable for long-term use due to the risk of accumulating point mutations.

Recent advancements have enabled the relatively inexpensive and efficient production of custom ssDNA scaffolds [141]. Wu *et al.* leveraged these techniques to produce scaffolds using all four possible three-letter codes, three of which (all but `ACG`) resulted in successful origami folding under the typical annealing process of slowly cooling from a high temperature [140]. Furthermore, the `ATC` scaffold was capable of folding under isothermal conditions at room temperature, confirming claims made in a patent filed in 2017 about a 1000 nt `ATC` scaffold [139]. The authors suggest that scaffolds containing `G` may have more secondary structure due to the possibility of forming `G`-quadruplexes, which might explain why only the `ATC` variant could form isothermally at room temperature. Regardless, this remarkable outcome highlights the significant reduction in secondary structure for the three-letter code scaffold. Secondary structure is the primary motive for annealing by slowly cooling. Additionally, the folding times were impressively short at room temperature, with a 1131 nt scaffold folding in under 5 minutes and a 5481 nt scaffold folding within 30 minutes. These results suggest that the primary challenge in isothermal staple insertion stems from the secondary structure of the standard M13mp18 scaffold, though further research with three-letter-code scaffolds is necessary to confirm this hypothesis.

The custom scaffolds developed by Wu *et al.* do not adhere to some of the desirable design constraints applied in Kishi *et al.*'s scaffold, which does not permit repetition of subsequences longer than 8 nucleotides to ensure correct staple binding. Although there is no explicit mention of additional design constraints in Kishi *et al.*, their avoidance of long homopolymeric runs (e.g., no poly(`A`) or poly(`T`) longer than 5, and no poly(`G`) longer than 3) likely improves synthesis quality and facilitates proper annealing, especially in the context of localized DNA circuits that use poly(`T`) tethers. Furthermore, any subsequence of length 8 or longer contains at least one `G`, ensuring that short staple domains are firmly bound. In contrast, Wu *et al.*'s 5481 nt scaffold contains runs of up to 10 poly(`C`), 7 poly(`T`), and 10 poly(`A`) sequences, as well as duplicated subsequences of up to 14 nt.

Using the constraints from Kishi *et al.*, I have generated a three-letter-code scaffold sequence of 3038 nt that begins with their existing 1394 nt design, as well as a 3206 nt sequence generated from scratch. These are not upper limits but rather represent the result of basic sequence exploration. By relaxing the rule against repeated subsequences longer than 8 nt to a more lenient threshold of 11 nt, I was able to design a 9151 nt scaffold (included in attached files)—longer than those tested by Wu et al. and exceeding the length of M13mp18. I strongly recommend that future projects aimed at developing localized circuits on DNA origami begin by producing rigid structures like the double-layer tile with a new three-letter-code scaffold. I encourage the use of any sequences that I attached here, or modified variants of them.

Beyond the issues pertaining to secondary structure, there are practical reasons to move away from M13mp18. For applications of DNA nanomachines in complex biological environments—such as in animal models, human tissue samples, or clinical diagnostics—it is critical to avoid immunogenic sequences, which are often present in viral genomes like M13mp18 [141]. Even though M13mp18 is the genome of a bacteriophage and therefore cannot infect mammalian cells, it has been found that bacteriophage DNA can still trigger an immune response in mice [142]. Additionally, custom scaffolds can be designed to be much larger than M13mp18. One study demonstrated the creation of a 51,466 nt scaffold, which enabled the folding of single-layer origami with a surface area seven times larger than that of M13mp18 [143]. A more complex circuit involving cascades of Latches would benefit from larger origami surfaces. Furthermore, an increased surface area would allow a sparser grid for staple extensions, enabling the design of staples with only long binding domains (see Appendix Q, A Staple Layout for a Proposed 14-Helix Origami Structure that Forms at High Temperatures for an example). Such staples would hybridize fully to the scaffold at high temperatures before their extensions hybridize to each other to form circuit components. On the origami surface, unlike in solution, geometric constraints can prevent the formation of undesired complexes.

6.3 Reconfigurable Molecular Machines

Despite the challenges we encountered, I remain confident that a solution can be developed to achieve cascadable localized catalytic DNA circuits. Eliminating the type of leak we encountered in pins-and-needles seemed insurmountable before we developed the latch system. Once duplicated complexes are eliminated, a reasonable next goal would be to apply localized catalytic circuits to achieve signal restoration in scalable spatially-organized computing systems. Constructing an OR gate, for example, will be trivial—all that is needed are two possible outputs that can activate the same downstream gate. If the downstream gate is placed at a location that is reachable to both, then OR computation will take place. Other types of gates may be much more challenging to construct. An AND gate may rely on a design similar to that of the cooperative catalyst discussed in Chapter 3, such that two inputs are needed to release one output. Beyond that, designing NOT gates will likely be avoidable through dual-rail logic, as in well-mixed systems [23].

Work is ongoing to allow origami tile displacement to be controlled by other strands, either well-mixed or localized to the tiles themselves. Once established, this technology can be combined with catalytic localized circuits to take advantage of their signal-insulating property: after catalyzing downstream reactions, the input to a catalytic system remains active and unbound to any other gate. Therefore, if a subset of its downstream gates were to be replaced with a different subset of downstream gates, the same input could act again in a new context to produce different results. Consequently, self-reconfigurable circuits could be built—see Figure 6.1 for a conceptual example.

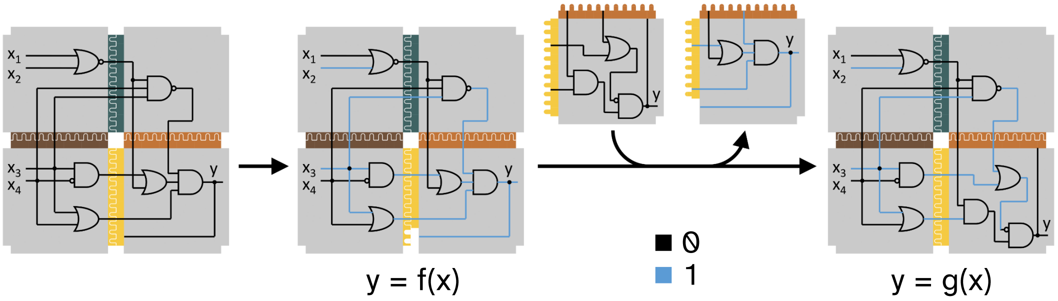


Figure 6.1: **Conceptual diagram of a self-reconfigurable circuit.** Initially, a circuit is laid out across a collection of DNA origami tiles. Upon exposure to a diffusible input, they compute function $f(x)$. In this case, the function causes a tile toehold to be exposed, allowing tile displacement to take place. The new tile has different circuit components, and since the signal in the upstream circuits was insulated, it remains active to act upon the circuit components in the new tile. As a result, the overall system now computes a new function $g(x)$.

This example might not seem consistent with the rest of this thesis upon first consideration. A nanorobot will not naturally encounter new origami tiles to replace the old in a biological context; however, just as in Figure 5.1a, this conceptual diagram is a simplification of our ultimate vision. Nanorobots may carry additional tiles into their new environments by flexible tethers, or they may be able to reorganize their own tiles from within one larger assembly. Whatever the mechanism, self-reconfigurable nanorobots will be adaptable, allowing them to perform a sequence of context-dependent functions.

This thesis has argued for a specific path forward to achieve self-reconfigurable nanorobots. If the reader shares our goals, but differs in their view of the best route toward them, that is wonderful. Despite our best efforts to be comprehensive, we have surely overlooked considerations which will be obvious to researchers with different backgrounds. Additionally, the applications of the work presented here, and the many past accomplishments that led to it, extend far beyond our vision for self-reconfigurable nanorobots. Whatever possibilities the reader[7] envisions will determine the future that lies ahead.

BIBLIOGRAPHY

1. Glynn, A. T., Davidson, S. R. & Qian, L. Developmental Self-Assembly of a DNA Ring with Stimulus-Responsive Size and Growth Direction. *Journal of the American Chemical Society* **144,** 10075–10079. DOI: 10.1021/jacs.2c03853 (June 8, 2022).
2. Taylor, D. N., Davidson, S. R. & Qian, L. A Cooperative DNA Catalyst. *Journal of the American Chemical Society* **143,** 15567–15571. DOI: 10.1021/jacs.1c07122 (Sept. 29, 2021).
3. Zwetsloot, A. J., Tut, G. & Straube, A. Measuring microtubule dynamics. *Essays in Biochemistry* **62** (eds Wakefield, J. G. & Moores, C. A.) 725–735. DOI: 10.1042/EBC20180035 (Dec. 7, 2018).
4. Brinkmann, V. *et al.* Neutrophil Extracellular Traps Kill Bacteria. *Science* **303,** 1532–1535. DOI: 10.1126/science.1092385 (Mar. 5, 2004).
5. Seeman, N. C. Nucleic acid junctions and lattices. *Journal of Theoretical Biology* **99,** 237–247. DOI: 10.1016/0022-5193(82)90002-9 (Nov. 21, 1982).
6. Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394,** 539–544. DOI: 10.1038/28998 (Aug. 1998).
7. Woods, D. *et al.* Diverse and robust molecular algorithms using reprogrammable DNA self-assembly. *Nature* **567,** 366–372. DOI: 10.1038/s41586-019-1014-9 (Mar. 2019).
8. Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485,** 623–626. DOI: 10.1038/nature11075 (May 2012).
9. Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302. DOI: 10.1038/nature04586 (2006).
10. Majikes, J. M. & Liddle, J. A. DNA Origami Design: A How-To Tutorial. *Journal of Research of the National Institute of Standards and Technology* **126,** 126001. DOI: 10.6028/jres.126.001 (Jan. 8, 2021).
11. Dirks, R. M., Bois, J. S., Schaeffer, J. M., Winfree, E. & Pierce, N. A. Thermodynamic Analysis of Interacting Nucleic Acid Strands. *SIAM Review* **49,** 65–88. DOI: 10.1137/060651100 (Jan. 2007).

12. Adleman, L. M. Molecular Computation of Solutions to Combinatorial Problems. *Science* **266,** 1021–1024.
DOI: 10.1126/science.7973651 (Nov. 11, 1994).

13. Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406,** 605–608.
DOI: 10.1038/35020524 (Aug. 2000).

14. Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences* **101,** 15275–15278.
DOI: 10.1073/pnas.0407024101 (2004).

15. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314,** 1585–1588.
DOI: 10.1126/science.1132493 (2006).

16. Aris, R. Prolegomena to the rational analysis of systems of chemical reactions. *Archive for Rational Mechanics and Analysis* **19,** 81–99.
DOI: 10.1007/BF00282276 (Jan. 1, 1965).

17. Soloveichik, D., Seelig, G. & Winfree, E. DNA as a universal substrate for chemical kinetics. *Proceedings of the National Academy of Sciences* **107,** 5393–5398.
DOI: 10.1073/pnas.0909380107 (Mar. 23, 2010).

18. Yin, P., Choi, H. M., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451,** 318–322.
DOI: 10.1038/nature06451 (2008).

19. Sadowski, J. P., Calvert, C. R., Zhang, D. Y., Pierce, N. A. & Yin, P. Developmental self-assembly of a DNA tetrahedron. *ACS Nano* **8,** 3251–3259.
DOI: 10.1021/nn4038223 (2014).

20. Douglas, S. M. *et al.* Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Research* **37,** 5001–5006.
DOI: 10.1093/nar/gkp436 (Aug. 1, 2009).

21. Ouldridge, T. *Coarse-grained modelling of DNA and DNA self-assembly* (Oxford University, UK, 2011). `https://ora.ox.ac.uk/objects/uuid:b2415bb2-7975-4f59-b5e2-8c022b4a3719`.

22. Soloveichik, D. *CRNSimulator* `http://users.ece.utexas.edu/~soloveichik/crnsimulator.html`. 2009.

23. Qian, L. & Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332,** 1196–1201.
DOI: 10.1126/science.1200520 (2011).

24. Genot, A. J., Zhang, D. Y., Bath, J. & Turberfield, A. J. Remote Toehold: A Mechanism for Flexible Control of DNA Hybridization Kinetics. *Journal of the American Chemical Society* **133,** 2177–2182. DOI: 10.1021/ja1073239 (Feb. 23, 2011).

25. Douglas, S. M., Bachelet, I. & Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **335,** 831–834. DOI: 10.1126/science.1214081 (Feb. 17, 2012).

26. Müller, B. *et al*. *High-resolution tomographic imaging of microvessels Developments in X-Ray Tomography VI* Developments in X-Ray Tomography VI. **7078** (SPIE, Sept. 18, 2008), 89–98. DOI: 10.1117/12.794157.

27. Zhang, G. *et al*. High energy density picoliter-scale zinc-air microbatteries for colloidal robotics. *Science Robotics* **9,** eade4642. DOI: 10.1126/scirobotics.ade4642 (Aug. 14, 2024).

28. Tikhomirov, G., Petersen, P. & Qian, L. Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature* **552,** 67–71. DOI: 10.1038/nature24655 (Dec. 2017).

29. Tikhomirov, G., Petersen, P. & Qian, L. Triangular DNA origami tilings. *Journal of the American Chemical Society* **140,** 17361–17364. DOI: 10.1021/jacs.8b10609 (2018).

30. Petersen, P., Tikhomirov, G. & Qian, L. Information-based autonomous reconfiguration in systems of interacting DNA nanostructures. *Nature Communications* **9,** 5362. DOI: 10.1038/s41467-018-07805-7 (2018).

31. Sarraf, N., Rodriguez, K. R. & Qian, L. Modular reconfiguration of DNA origami assemblies using tile displacement. *Science Robotics* **8,** eadf1511. DOI: 10.1126/scirobotics.adf1511 (2023).

32. Song, J. *et al*. Reconfiguration of DNA molecular arrays driven by information relay. *Science* **357,** eaan3377. DOI: 10.1126/science.aan3377 (July 28, 2017).

33. Wang, D. *et al*. Programmable Transformations of DNA Origami Made of Small Modular Dynamic Units. *Journal of the American Chemical Society* **143,** 2256–2263. DOI: 10.1021/jacs.0c10576 (Feb. 10, 2021).

34. Doty, D., Lee, B. L. & Stérin, T. *scadnano: A Browser-Based, Scriptable Tool for Designing DNA Nanostructures* 26th International Conference on DNA Computing and Molecular Programming (DNA 26) (2020) (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020). DOI: 10.4230/LIPIcs.DNA.2020.9.

35. Levy, N. & Schabanel, N. ENSnano: A 3D Modeling Software for DNA Nanostructures. *LIPIcs, Volume 205, DNA 27* **205.** In collab. with Lakin, M. R. & Šulc, P., 5:1–5:23.
DOI: 10.4230/LIPICS.DNA.27.5 (2021).

36. Poppleton, E., Mallya, A., Dey, S., Joseph, J. & Šulc, P. Nanobase.org: a repository for DNA and RNA nanostructures. *Nucleic Acids Research* **50,** D246–D252.
DOI: 10.1093/nar/gkab1000 (D1 Jan. 7, 2022).

37. Kube, M. *et al.* Revealing the structures of megadalton-scale DNA complexes with nucleotide resolution. *Nature Communications* **11,** 6229.
DOI: 10.1038/s41467-020-20020-7 (Dec. 4, 2020).

38. McRae, E. K. S. *et al.* Structure, folding and flexibility of co-transcriptional RNA origami. *Nature Nanotechnology* **18,** 808–817.
DOI: 10.1038/s41565-023-01321-6 (July 2023).

39. Qian, L. & Winfree, E. *Parallel and scalable computation and spatial dynamics with DNA-based chemical reaction networks on a surface DNA Computing and Molecular Programming: 20th International Conference, DNA 20, Kyoto, Japan, September 22-26, 2014. Proceedings 20* (2014), 114–131.
DOI: 10.1007/978-3-319-11295-4_8.

40. Clamons, S., Qian, L. & Winfree, E. Programming and simulating chemical reaction networks on a surface. *Journal of the Royal Society Interface* **17,** 20190790.
DOI: 10.1098/rsif.2019.0790 (2020).

41. Cherry, K. M. & Qian, L. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* **559,** 370–376.
DOI: 10.1038/s41586-018-0289-6 (2018).

42. Yang, X., Tang, Y., Traynor, S. M. & Li, F. Regulation of DNA Strand Displacement Using an Allosteric DNA Toehold. *Journal of the American Chemical Society* **138,** 14076–14082.
DOI: 10.1021/jacs.6b08794 (Oct. 26, 2016).

43. Wang, B., Thachuk, C., Ellington, A. D., Winfree, E. & Soloveichik, D. Effective design principles for leakless strand displacement systems. *Proceedings of the National Academy of Sciences* **115,** E12182–E12191.
DOI: 10.1073/pnas.1806859115 (2018).

44. Wang, B., Thachuk, C. & Soloveichik, D. Speed and correctness guarantees for programmable enthalpy-neutral DNA reactions. *ACS Synthetic Biology*.
DOI: 10.1021/acssynbio.2c00356 (2023).

45. Gerling, T., Kube, M., Kick, B. & Dietz, H. Sequence-programmable covalent bonding of designed DNA assemblies. *Science Advances* **4,** eaau1157.
DOI: 10.1126/sciadv.aau1157 (Aug. 17, 2018).

46. Zhang, H. *et al. DNA Nanostructures Coordinate Gene Silencing in Mature Plants* Feb. 2, 2019.
DOI: 10.1101/538678.

47. Kretzmann, J. A. *et al.* Gene-encoding DNA origami for mammalian cell expression. *Nature Communications* **14,** 1017.
DOI: 10.1038/s41467-023-36601-1 (Feb. 23, 2023).

48. Jiang, D. *et al.* DNA origami nanostructures can exhibit preferential renal uptake and alleviate acute kidney injury. *Nature Biomedical Engineering* **2,** 865–877.
DOI: 10.1038/s41551-018-0317-8 (Nov. 2018).

49. Liu, Y. *et al.* Organ-Specific Gene Expression Control Using DNA Origami-Based Nanodevices. *Nano Letters* **24,** 8410–8417.
DOI: 10.1021/acs.nanolett.4c02104 (July 10, 2024).

50. Endo, M. *et al.* RNA-templated DNA origami structures. *Chemical Communications* **49,** 2879–2881.
DOI: 10.1039/C3CC38804B (Mar. 12, 2013).

51. Wang, P., Ko, S. H., Tian, C., Hao, C. & Mao, C. RNA–DNA hybrid origami: folding of a long RNA single strand into complex nanostructures using short DNA helper strands. *Chemical Communications* **49,** 5462–5464.
DOI: 10.1039/C3CC41707G (May 21, 2013).

52. Endo, M., Takeuchi, Y., Emura, T., Hidaka, K. & Sugiyama, H. Preparation of Chemically Modified RNA Origami Nanostructures. *Chemistry – A European Journal* **20,** 15330–15333.
DOI: 10.1002/chem.201404084 (2014).

53. Lee, A.-Y., Kim, K.-R., Yu, J. H. & Ahn, D.-R. L-DNA linear duplex: An efficient drug delivery carrier with a simple structure. *Journal of Industrial and Engineering Chemistry* **74,** 187–192.
DOI: 10.1016/j.jiec.2019.02.025 (June 25, 2019).

54. Zhong, W. & Sczepanski, J. T. Direct Comparison of d-DNA and l-DNA Strand-Displacement Reactions in Living Mammalian Cells. *ACS Synthetic Biology* **10,** 209–212.
DOI: 10.1021/acssynbio.0c00527 (Jan. 15, 2021).

55. Mallette, T. L., Lidke, D. S. & Lakin, M. R. Heterochiral modifications enhance robustness and function of DNA in living human cells. *ChemBioChem* **25,** e202300755.
DOI: 10.1002/cbic.202300755 (2024).

56. Mallette, T. L. & Lakin, M. R. Protecting Heterochiral DNA Nanostructures against Exonuclease-Mediated Degradation. *ACS Synthetic Biology* **11,** 2222–2228.
DOI: 10.1021/acssynbio.2c00105 (July 15, 2022).

57. Harrison, K., Mackay, A. S., Kambanis, L., Maxwell, J. W. C. & Payne, R. J. Synthesis and applications of mirror-image proteins. *Nature Reviews Chemistry* **7,** 383–404.
DOI: 10.1038/s41570-023-00493-y (June 2023).

58. Engelen, W., Sigl, C., Kadletz, K., Willner, E. M. & Dietz, H. Antigen-Triggered Logic-Gating of DNA Nanodevices. *Journal of the American Chemical Society* **143,** 21630–21636.
DOI: 10.1021/jacs.1c09967 (Dec. 29, 2021).

59. Sigl, C. *et al*. Programmable icosahedral shell system for virus trapping. *Nature Materials* **20,** 1281–1289.
DOI: 10.1038/s41563-021-01020-4 (Sept. 2021).

60. *Next-Generation Sequencing (NGS) | Explore the technology* `https://www.illumina.com/science/technology/next-generation-sequencing.html` (2024).

61. *History of Illumina Sequencing* `https://web.archive.org/web/20141012151855/http://technology.illumina.com/technology/next-generation-sequencing/solexa-technology.html` (2024).

62. *About Molecular Instruments, Inc*. Molecular Instruments. `https://www.molecularinstruments.com/about` (2024).

63. 1E9. *Disrupting medicine: DNA Origami as the new cornerstone of nano robotics* Sept. 25, 2023. `https://www.youtube.com/watch?v=2WxNtRBwpfM`.

64. Seeman, N. C. & Sleiman, H. F. DNA nanotechnology. *Nature Reviews Materials* **3,** 1–23.
DOI: 10.1038/70123 (2017).

65. Ong, L. L. *et al*. Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature* **552,** 72–77.
DOI: 10.1038/nature24648 (2017).

66. Wagenbauer, K. F., Sigl, C. & Dietz, H. Gigadalton-scale shape-programmable DNA assemblies. *Nature* **552,** 78–83.
DOI: 10.1038/nature24651 (2017).

67. Tikhomirov, G., Petersen, P. & Qian, L. Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature* **552,** 67–71.
DOI: 10.1038/nature24655 (2017).

68. Rothemund, P. W., Papadakis, N. & Winfree, E. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol* **2,** e424.
DOI: 10.1371/journal.pbio.0020424 (2004).

69. Woods, D. *et al.* Diverse and robust molecular algorithms using reprogrammable DNA self-assembly. *Nature* **567,** 366–372.
DOI: 10.1038/s41586-019-1014-9 (2019).

70. Zhang, D. Y., Hariadi, R. F., Choi, H. M. & Winfree, E. Integrating DNA strand-displacement circuitry with DNA tile self-assembly. *Nature communications* **4,** 1–10.
DOI: 10.1038/ncomms2965 (2013).

71. Choi, H. M. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature Biotechnology* **28,** 1208–1212.
DOI: 10.1038/nbt.1692 (2010).

72. Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333,** 470–474.
DOI: 10.1126/science.1206938 (2011).

73. Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406,** 605–608.
DOI: 10.1038/35020524 (2000).

74. Yin, P. *et al.* Programming DNA tube circumferences. *Science* **321,** 824–826.
DOI: 10.1126/science.1157312 (2008).

75. Zhang, Y. *et al.* Programming DNA tube circumference by tile offset connection. *Journal of the American Chemical Society* **141,** 19529–19532.
DOI: 10.1021/jacs.9b08921 (2019).

76. Barish, R. D., Schulman, R., Rothemund, P. W. & Winfree, E. An information-bearing seed for nucleating algorithmic self-assembly. *Proceedings of the National Academy of Sciences* **106,** 6054–6059.
DOI: 10.1073/pnas.0808736106 (2009).

77. Zadeh, J. N. *et al.* NUPACK: analysis and design of nucleic acid systems. *Journal of Computational Chemistry* **32,** 170–173.
DOI: 10.1002/jcc.21596 (2011).

78. Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *Journal of the American Chemical Society* **131,** 17303–17314.
DOI: 10.1021/ja906987s (2009).

79. Srinivas, N. *et al.* On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Research* **41,** 10641–10658. DOI: 10.1093/nar/gkt801 (2013).

80. Genot, A. J., Zhang, D. Y., Bath, J. & Turberfield, A. J. Remote toehold: a mechanism for flexible control of DNA hybridization kinetics. *Journal of the American Chemical Society* **133,** 2177–2182. DOI: 10.1021/ja1073239 (2011).

81. Dabby, N. L. *The kinetics of toehold-mediated four-way branch migration in Synthetic molecular machines for active self-assembly: prototype algorithms, designs, and experimental study* PhD thesis (California Institute of Technology, 2013). DOI: 10.7907/T0ZG-PA07.

82. Venkataraman, S., Dirks, R. M., Rothemund, P. W., Winfree, E. & Pierce, N. A. An autonomous polymerization motor powered by DNA hybridization. *Nature Nanotechnology* **2,** 490–494. DOI: 10.1038/nnano.2007.225 (2007).

83. Wang, J. *et al.* Clamped hybridization chain reactions for the self-assembly of patterned DNA hydrogels. *Angewandte Chemie International Edition* **56,** 2171–2175. DOI: 10.1002/ange.201610125 (2017).

84. Cangialosi, A. *et al.* DNA sequence–directed shape change of photopatterned hydrogels via high-degree swelling. *Science* **357,** 1126–1130. DOI: 10.1126/science.aan3925 (2017).

85. Fern, J. & Schulman, R. Modular DNA strand-displacement controllers for directing material expansion. *Nature Communications* **9,** 1–8. DOI: 10.1038/s41467-018-06218-w (2018).

86. Langecker, M. *et al.* Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **338,** 932–936. DOI: 10.1126/science.1225624 (2012).

87. Burns, J. R., Seifert, A., Fertig, N. & Howorka, S. A biomimetic DNA-based channel for the ligand-controlled transport of charged molecular cargo across a biological membrane. *Nature Nanotechnology* **11,** 152–156. DOI: 10.1038/nnano.2015.279 (2016).

88. Turberfield, A. J. *et al.* DNA fuel for free-running nanomachines. *Physical Review Letters* **90,** 118102. DOI: 10.1103/PhysRevLett.90.118102 (2003).

89. Bois, J. S. *et al.* Topological constraints in nucleic acid hybridization kinetics. *Nucleic Acids Research* **33,** 4090–4095. DOI: 10.1093/nar/gki721 (2005).

90. Seelig, G., Yurke, B. & Winfree, E. Catalyzed relaxation of a metastable DNA fuel. *Journal of the American Chemical Society* **128,** 12211–12220. DOI: 10.1021/ja0635635 (2006).

91. Zhang, D. Y., Turberfield, A. J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318,** 1121–1125. DOI: 10.1126/science.1148532 (2007).

92. Li, B., Ellington, A. D. & Chen, X. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Research* **39,** e110–e110. DOI: 10.1093/nar/gkr504 (2011).

93. Chen, X., Briggs, N., McLain, J. R. & Ellington, A. D. Stacking nonenzymatic circuits for high signal gain. *Proceedings of the National Academy of Sciences* **110,** 5386–5391. DOI: 10.1073/pnas.1222807110 (2013).

94. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475,** 368–372. DOI: 10.1038/nature10262 (2011).

95. Yang, X., Tang, Y., Traynor, S. M. & Li, F. Regulation of DNA strand displacement using an allosteric DNA toehold. *Journal of the American Chemical Society* **138,** 14076–14082. DOI: 10.1021/jacs.6b08794 (2016).

96. Lakin, M. R. & Stefanovic, D. Supervised learning in adaptive DNA strand displacement networks. *ACS Synthetic Biology* **5,** 885–897. DOI: 10.1021/acssynbio.6b00009 (2016).

97. Qian, L. & Winfree, E. A simple DNA gate motif for synthesizing large-scale circuits. *Journal of the Royal Society Interface* **8,** 1281–1297. DOI: 10.1098/rsif.2010.0729 (2011).

98. Zhang, D. Y. Cooperative hybridization of oligonucleotides. *Journal of the American Chemical Society* **133,** 1077–1086. DOI: 10.1021/ja109089q (2011).

99. Thubagere, A. J. *et al.* Compiler-aided systematic construction of large-scale DNA strand displacement circuits using unpurified components. *Nature Communications* **8,** 14373. DOI: 10.1038/ncomms14373 (2017).

100. Protozanova, E., Yakovchuk, P. & Frank-Kamenetskii, M. D. Stacked–unstacked equilibrium at the nick site of DNA. *Journal of Molecular Biology* **342,** 775–785. DOI: 10.1016/j.jmb.2004.07.075 (2004).

101. Srinivas, N., Parkin, J., Seelig, G., Winfree, E. & Soloveichik, D. Enzyme-free nucleic acid dynamical systems. *Science* **358**. DOI: 10.1126/science.aal2052 (2017).

102. Allawi, H. T. & SantaLucia, J. Thermodynamics and NMR of internal G·T mismatches in DNA. *Biochemistry* **36,** 10581–10594. DOI: 10.1021/bi962590c (1997).

103. Haley, N. E. *et al.* Design of hidden thermodynamic driving for non-equilibrium systems via mismatch elimination during DNA strand displacement. *Nature Communications* **11,** 1–11. DOI: 10.1038/s41467-020-16353-y (2020).

104. SantaLucia Jr, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annual Review of Biophysics and Biomolecular Structure* **33,** 415–440. DOI: 10.1146/annurev.biophys.32.110601.141800 (2004).

105. Zhang, D. Y. & Winfree, E. Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Research* **38,** 4182–4197. DOI: 10.1093/nar/gkq088 (2010).

106. Machinek, R. R., Ouldridge, T. E., Haley, N. E., Bath, J. & Turberfield, A. J. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nature Communications* **5,** 1–9. DOI: 10.1038/ncomms6324 (2014).

107. Irmisch, P., Ouldridge, T. E. & Seidel, R. Modeling DNA-strand displacement reactions in the presence of base-pair mismatches. *Journal of the American Chemical Society* **142,** 11451–11463. DOI: 10.1021/jacs.0c03105 (2020).

108. Zhang, D. Y. & Winfree, E. Dynamic allosteric control of noncovalent DNA catalysis reactions. *Journal of the American Chemical Society* **130,** 13921–13926. DOI: 10.1021/ja803318t (2008).

109. Chen, X. Expanding the rule set of DNA circuitry with associative toehold activation. *Journal of the American Chemical Society* **134,** 263–271. DOI: 10.1021/ja206690a (2012).

110. Soloveichik, D., Seelig, G. & Winfree, E. DNA as a universal substrate for chemical kinetics. *Proceedings of the National Academy of Sciences* **107,** 5393–5398. DOI: 10.1007/978-3-642-03076-5_6 (2010).

111. Chen, Y.-J. *et al.* Programmable chemical controllers made from DNA. *Nature nanotechnology* **8,** 755–762. DOI: 10.1038/nnano.2013.189 (2013).

112. Bui, H. *et al*. Design and analysis of localized DNA hybridization chain reactions. *Small* **13,** 1602983.
DOI: 10.1002/smll.201602983 (2017).

113. Bui, H. *et al*. Localized DNA hybridization chain reactions on DNA origami. *ACS Nano* **12,** 1146–1155.
DOI: 10.1021/acsnano.7b06699 (2018).

114. Chao, J. *et al*. Solving mazes with single-molecule DNA navigators. *Nature Materials* **18,** 273–279.
DOI: 10.1038/s41563-018-0205-3 (2019).

115. Chatterjee, G., Dalchau, N., Muscat, R. A., Phillips, A. & Seelig, G. A spatially localized architecture for fast and modular DNA computing. *Nature Nanotechnology* **12,** 920–927.
DOI: 10.1038/nnano.2017.127 (2017).

116. Liu, L. *et al*. A localized DNA finite-state machine with temporal resolution. *Science Advances* **8,** eabm9530.
DOI: 10.1126/sciadv.abm9530 (2022).

117. Teichmann, M., Kopperger, E. & Simmel, F. C. Robustness of localized DNA strand displacement cascades. *ACS Nano* **8,** 8487–8496.
DOI: 10.1021/nn503073p (2014).

118. Ruiz, I. M. *et al*. Connecting localized DNA strand displacement reactions. *Nanoscale* **7,** 12970–12978.
DOI: 10.1039/C5NR02434J (2015).

119. Wang, D. *et al*. Molecular logic gates on DNA origami nanostructures for microRNA diagnostics. *Analytical Chemistry* **86,** 1932–1936.
DOI: 10.1021/ac403661z (2014).

120. Shin, S. W., Ahn, S. Y., Lim, Y. T. & Um, S. H. Improved Sensitivity of Intramolecular Strand Displacement Based on Localization of Probes. *Analytical Chemistry* **91,** 14808–14811.
DOI: 10.1021/acs.analchem.9b03173 (2019).

121. Wei, Q. *et al*. A DNA nanowire based localized catalytic hairpin assembly reaction for microRNA imaging in live cells. *Chemical Science* **9,** 7802–7808.
DOI: 10.1039/C8SC02943A (2018).

122. Jiang, Y. S., Bhadra, S., Li, B. & Ellington, A. D. Mismatches improve the performance of strand-displacement nucleic acid circuits. *Angewandte Chemie* **126,** 1876–1879.
DOI: 10.1002/ange.201307418 (2014).

123. Olson, X. *et al*. Availability: A metric for nucleic acid strand displacement systems. *ACS Synthetic Biology* **6,** 84–93.
DOI: 10.1021/acssynbio.5b00231 (2017).

124. Lysne, D. *et al.* Availability-driven design of hairpin fuels and small interfering strands for leakage reduction in autocatalytic networks. *The Journal of Physical Chemistry B* **124,** 3326–3335. DOI: 10.1021/acs.jpcb.0c01229 (2020).

125. Thubagere, A. J. *et al.* A cargo-sorting DNA robot. *Science* **357,** eaan6558. DOI: 10.1126/science.aan6558 (2017).

126. Dannenberg, F., Kwiatkowska, M., Thachuk, C. & Turberfield, A. J. DNA walker circuits: computational potential, design, and verification. *Natural Computing* **14,** 195–211. DOI: 10.1007/s11047-014-9426-9 (2015).

127. Boemo, M. A., Lucas, A. E., Turberfield, A. J. & Cardelli, L. The formal language and design principles of autonomous DNA walker circuits. *ACS Synthetic Biology* **5,** 878–884. DOI: 10.1021/acssynbio.5b00275 (2016).

128. Wickham, S. F. *et al.* A DNA-based molecular motor that can navigate a network of tracks. *Nature nanotechnology* **7,** 169–173. DOI: 10.1038/nnano.2011.253 (2012).

129. Li, J. *et al.* Exploring the speed limit of toehold exchange with a cartwheeling DNA acrobat. *Nature Nanotechnology* **13,** 723–729. DOI: 10.1038/s41565-018-0130-2 (2018).

130. Strauss, M. T., Schueder, F., Haas, D., Nickels, P. C. & Jungmann, R. Quantifying absolute addressability in DNA origami with molecular resolution. *Nature Communications* **9,** 1600. DOI: 10.1038/s41467-018-04031-z (Apr. 23, 2018).

131. Sarraf, N. *Towards integrated molecular machines: structural, mechanical, and computational motifs* PhD thesis (California Institute of Technology, 2023). DOI: 10.7907/cdwp-c709.

132. Wintersinger, C. M. *et al.* Multi-micron crisscross structures grown from DNA-origami slats. *Nature Nanotechnology* **18,** 281–289. DOI: 10.1038/s41565-022-01283-1 (2023).

133. Winfree, E. *The Tall Thin Molecular Programmer 2023 IEEE International Solid-State Circuits Conference (ISSCC)* (2023), 1–3. DOI: 10.1109/ISSCC42615.2023.10067743.

134. Khara, D. C. *et al.* DNA bipedal motor walking dynamics: an experimental and theoretical study of the dependency on step size. *Nucleic Acids Research* **46,** 1553–1561. DOI: 10.1093/nar/gkx1282 (2018).

135. Dalchau, N., Chandran, H., Gopalkrishnan, N., Phillips, A. & Reif, J. Probabilistic analysis of localized DNA hybridization circuits. *ACS Synthetic Biology* **4,** 898–913.
DOI: 10.1021/acssynbio.5b00044 (2015).

136. Kumar, S., Weisburd, J. M. & Lakin, M. R. Structure sampling for computational estimation of localized DNA interaction rates. *Scientific Reports* **11,** 12730.
DOI: 10.1038/s41598-021-92145-8 (2021).

137. Qian, L. & Winfree, E. A simple DNA gate motif for synthesizing large-scale circuits. *Journal of The Royal Society Interface* **8,** 1281–1297.
DOI: 10.1098/rsif.2010.0729 (Feb. 4, 2011).

138. Kishi, J. Y., Schaus, T. E., Gopalkrishnan, N., Xuan, F. & Yin, P. Programmable autonomous synthesis of single-stranded DNA. *Nature Chemistry* **10,** 155–164.
DOI: 10.1038/nchem.2872 (Feb. 2018).

139. Gopalkrishnan, N., Schaus, T. E. & Yin, P. *US Patent* 11235972B2. `https://patents.google.com/patent/US11235972B2/en`(2022).

140. Wu, H. *et al.* Expanding DNA Origami Design Freedom with De Novo Synthesized Scaffolds. *Journal of the American Chemical Society* **146,** 16076–16084.
DOI: 10.1021/jacs.4c03148 (June 12, 2024).

141. Bush, J. *et al.* Synthesis of DNA Origami Scaffolds: Current and Emerging Strategies. *Molecules* **25,** 3386.
DOI: 10.3390/molecules25153386 (Jan. 2020).

142. Podlacha, M. *et al.* Interactions of Bacteriophages with Animal and Human Organisms—Safety Issues in the Light of Phage Therapy. *International Journal of Molecular Sciences* **22,** 8937.
DOI: 10.3390/ijms22168937 (Aug. 19, 2021).

143. Marchi, A. N., Saaem, I., Vogen, B. N., Brown, S. & LaBean, T. H. Toward Larger DNA Origami. *Nano Letters* **14,** 5740–5747.
DOI: 10.1021/nl502626s (Oct. 8, 2014).

144. Schneider, F., Möritz, N. & Dietz, H. The sequence of events during folding of a DNA origami. *Science Advances* **5,** eaaw1412.
DOI: 10.1126/sciadv.aaw1412 (May 3, 2019).

145. Derr, N. D. *et al.* Tug-of-War in Motor Protein Ensembles Revealed with a Programmable DNA Origami Scaffold. *Science* **338,** 662–665.
DOI: 10.1126/science.1226734 (Nov. 2, 2012).

146. Record Jr., M. T., Anderson, C. F. & Lohman, T. M. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Quarterly Reviews of Biophysics* **11,** 103–178.
DOI: 10.1017/S003358350000202X (May 1978).

147. Smith, S. B., Cui, Y. & Bustamante, C. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science* **271,** 795–799.
DOI: 10.1126/science.271.5250.795 (Feb. 9, 1996).

148. McCauley, M. J. & Williams, M. C. Mechanisms of DNA binding determined in optical tweezers experiments. *Biopolymers* **85,** 154–168.
DOI: 10.1002/bip.20622 (2007).

149. Rodriguez, K. R., Sarraf, N. & Qian, L. A loser-take-all DNA circuit. *ACS Synthetic Biology* **10,** 2878–2885.
DOI: 10.1021/acssynbio.1c00318 (2021).

150. Lapteva, A. P., Sarraf, N. & Qian, L. DNA strand-displacement temporal logic circuits. *Journal of the American Chemical Society* **144,** 12443–12449.
DOI: 10.1021/jacs.2c04325 (2022).

151. *Oligonucleotide synthesis: Coupling efficiency and quality control* Integrated DNA Technologies. `https://www.idtdna.com/pages/education/decoded/article/oligo-synthesis-why-idt-leads-the-oligo-industry` (2024).

152. *Which type of oligo purification should I choose?* Integrated DNA Technologies. `https://www.idtdna.com/pages/education/decoded/article/which-type-of-purification-should-i-choose` (2024).

153. Gopinath, A. *et al.* Absolute and arbitrary orientation of single-molecule shapes. *Science* **371,** eabd6179.
DOI: 10.1126/science.abd6179 (2021).

154. Thubagere, A. J. *Programming Complex Behavior In DNA-based Molecular Circuits And Robots* PhD thesis (California Institute of Technology, 2017).
DOI: 10.7907/Z9WD3XMS.

155. Kopperger, E. *et al.* A self-assembled nanoscale robotic arm controlled by electric fields. *Science* **359,** 296–301.
DOI: 10.1126/science.aao4284 (2018).

156. Fornace, M. E. *et al.* *NUPACK: Analysis and Design of Nucleic Acid Structures, Devices, and Systems* Nov. 10, 2022.
DOI: 10.26434/chemrxiv-2022-xv98l.

157. Saisuk, W., Suksamai, C., Srisawat, C., Yoksan, S. & Dharakul, T. The helper oligonucleotides enable detection of folded single-stranded DNA by lateral flow immunoassay after HCR signal amplification. *Talanta* **248,** 123588.
DOI: 10.1016/j.talanta.2022.123588 (Oct. 2022).

158. Dunn, K. E. *et al.* Guiding the folding pathway of DNA origami. *Nature* **525,** 82–86.
DOI: 10.1038/nature14860 (2015).

159. Aksel, T., Navarro, E. J., Fong, N. & Douglas, S. M. *Design principles for accurate folding of DNA origami* Mar. 19, 2024.
DOI: 10.1101/2024.03.18.585609.

160. Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M. & Pierce, N. A. Constrained multistate sequence design for nucleic acid reaction pathway engineering. *Journal of the American Chemical Society* **139,** 3134–3144.
DOI: 10.1021/jacs.6b12693 (2017).

161. Qian, L., Soloveichik, D. & Winfree, E. *Efficient Turing-universal computation with DNA polymers International Workshop on DNA-Based Computers* (2010), 123–140.
DOI: 10.1007/978-3-642-18305-8_12.

162. Cardelli, L. Two-domain DNA strand displacement. *Mathematical Structures in Computer Science* **23,** 247–271.
DOI: 10.1017/S0960129512000102 (2013).

163. Johnson, R. F. & Qian, L. *Simplifying chemical reaction network implementations with two-stranded DNA building blocks International Conference on DNA Computing and Molecular Programming* (eds Geary, C. & Patitz, M. J.) **174** (Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020), 2:1–2:14.
DOI: 10.4230/LIPIcs.DNA.2020.2.

164. Vasić, M., Soloveichik, D. & Khurshid, S. CRN++: Molecular programming language. *Natural Computing* **19,** 391–407.
DOI: 10.1007/978-3-030-00030-1_1 (2020).

165. Vasic, M., Chalk, C., Khurshid, S. & Soloveichik, D. *Deep molecular programming: a natural implementation of binary-weight ReLU neural networks Proceedings of the 37th International Conference on Machine Learning* (eds III, H. D. & Singh, A.) **119** (PMLR, July 2020), 9701–9711. `proceedings.mlr.press/v119/vasic20a.html`.

Appendix A

STAPLE INCORPORATION RATES

A.1 Staple Incorporation Background

When origami is annealed, not every staple is typically incorporated into the structure. Forming origami is a stochastic and cooperative process, where at any time point during annealing, each staple has a different probability of being incorporated based on which other staples were incorporated previously [144]. For a single-layer rectangular origami, Strauss *et al.* found that the staple incorporation rate varies from 48% in staples near the edges of the structure to 95% in the middle of the structure [130]. That study used the same type of folding buffer as ours (1×TE, 12.5 mM Mg2+), the same type of scaffold (M13mp18), and a similar staple pattern with maximum continuous domains of 16bp. They used a somewhat longer cooling regimen (ours: 90°C for 5 min, then linearly cooled to 20°C over 70 minutes; theirs: 80°C for 5 min, rapid cooling to 60°C, then a linear cooling ramp to 4°C over the subsequent 180 minutes). While there is undoubtedly some difference in the incorporation rate for staples in the double-layer origami structure compared to a single-layer structure, we believe the range of values measured in this study is a reasonable approximation for expected staple rates in our work. Incorporation rates derived from other papers are given in Supplementary Table 2 of reference [130], ranging from 76.5% to 95%.

In addition, Strauss *et al.* was able to measure a rate of extension sequestration of roughly 7% by seeking to detect extensions at both the 5’ and 3’ ends of staples. Specifically, some staples could occasionally be detected at only their 5’ end or their 3’ end, signifying that staple extensions are not always accessible to perform their intended function. The mechanism for this issue was not established, but it is reasonable to expect that a similar effect could occur in our work. Perhaps during annealing, staple extensions are occasionally trapped on the wrong side of the surface due to steric hindrance.

A.2 Attempts to improve incorporation rates and reduce sequestration of extensions for the latch

We did not directly measure staple incorporation rates in any experiment; however, given that some authors have reported annealing time [145] and staple:scaffold ratio [130, 145] affecting staple incorporation rates, the following experiments were conducted to assess whether annealing time and staple:scaffold ratio might have an effect on latch function.

Some origami structures are known to form poorly when annealing takes place too quickly potentially as a result of secondary structure or steric hindrance. In particular, we read about one structure, a star-shaped 12-helix bundle, which required a 19h annealing protocol [145]. Experiments involving longer, slowing annealing protocols were explored in order to determine whether any undesired aspects of latch function resulted from excessively fast annealing (Fig. A.1a). The resulting differences in metrics fell within the range of experimental noise (Table A.1, DO.2.b vs DO.2b.s), so a long anneal time did not appear to confer any benefit. It is possible that the effects of complex duplication, which will be discussed in Appendix N, Modified Positive Control and Duplicated Complexes, obfuscate any beneficial effects of a longer anneal.

One source [130] recommended using a staple:scaffold ratio of 50, whereas our typical ratio was only 5–7. Another source [145] used a staple:scaffold ratio of 90 for some staples with extensions. To determine whether any undesired latch properties resulted from an undesired staple:scaffold ratio, experiments were performed using a range of ratios. At the low end, we tried using a staple:scaffold ratio of 2 (Appendix P.1, Annealing with Low Concentration Components, DO.LC.c), which resulted in a small decrease of ≈2% in signal amplification, but improved % completion by ≈8%. At the high end, we tried a staple:scaffold ratio of 20 (Fig. A.1b and Table A.1), which decreased % signal amplification and % completion by ≈28% and ≈46% respectively. The best explanation for this result is complex duplication, which will be discussed in Appendix N, Modified Positive Control and Duplicated Complexes.

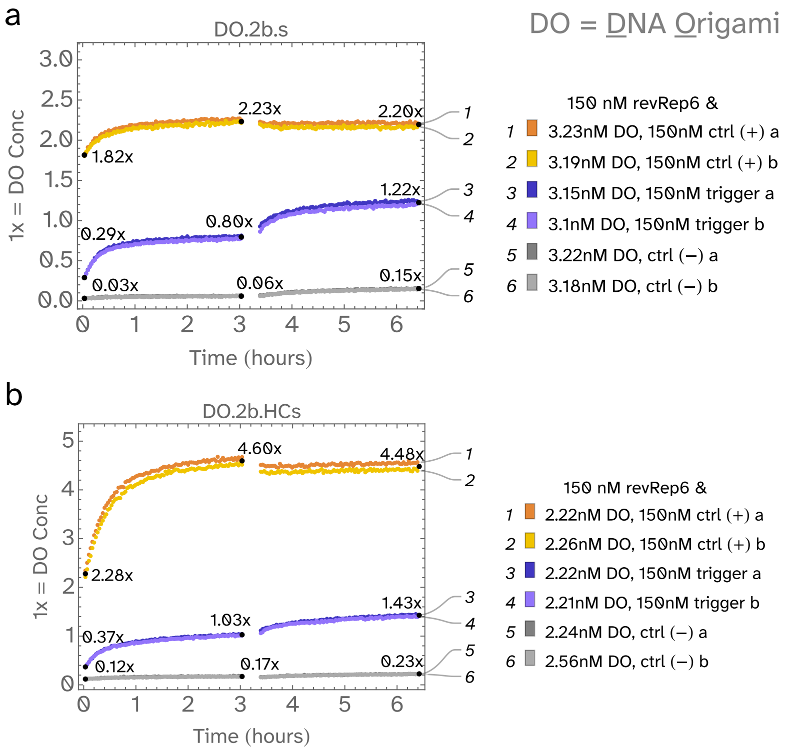


Figure A.1: **Effects of annealing slowly and with high circuit component concentrations.** Metrics given in Table A.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Latch annealed slowly (over ≈ 19h rather than 110 minutes as usual) **b,** Latch annealed slowly, and with circuit components added at 20–21× scaffold concentration rather than 5–7×.

Metric	DO.2.b	DO.2b.s	DO.2b.HCs
% Signal Amplification	61.5	56.1	44
% Completion	52.9	53.8	28.8
% Fuel leak	4.8	4.4	1.2
% Initial leak	2.3	1.5	2.6
Annealing time (h)	≈1.8	≈19	≈19
Staple:scaffold ratio	5–7	5–7	20–21

Table A.1: **Effect of varying annealing time and including high concentrations of circuit components on metrics. 2.b,** Latch annealed under standard conditions (110 minutes, staple:scaffold ratio of 5–7) **2b.s,** Latch annealed slowly (≈ 19h) **2b.HCs,** Latch annealed slowly (≈ 19h), and with circuit components added at 20–21× scaffold concentration. The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data.

Appendix B

DNA LENGTHS

For dsDNA, a contour length of .34 nm/bp is well-established [146–148]. Based on the salt dependence of the melting temperature of DNA, a contour length for ssDNA of .43 nm/nt was once estimated [146], and this value was used in the paper establishing DNA origami in 2006 [9]. However, published in the same year, work using optical tweezers to measure the length of ssDNA measured a contour length of .56 nm/nt [148]. This value is consistent with prior work finding that dsDNA, under sufficient tension, will undergo a sharp transition to a contour length of .58 nm/bp [147], signifying force-induced denaturation [148]. Furthermore, when ssDNA alone is placed under tension, its contour length was found to reach values between .6 and .65 nm/nt [148].

For the purposes of establishing reachability, that is, which DNA molecules on an origami surface can physically reach each other, I believe it is appropriate to use the longer .65 nm/nt value to provide a margin of error sufficient to prevent leak. Conversely, for the purpose of estimating reaction rates based on the hairpin closing model, I consider the relaxed length of .56 nm/nt to be more reasonable than .65 nm/nt.

For historic reasons, while working on the pins-and-needles system and the latch, we used the older .43 nm/nt value while running our simulations. We did not change the value in our exploratory simulations because they were purely meant to be qualitative, not quantitative; they were typically not parameterized with experimental data, and the use of our simple model for localized reactions was a known limitation.

The original equation for approximating a hairpin length in the simple model:

$$n = (\sum_i n_i^{ss} \times 0.43 + \sum_i n_i^{ds} \times 0.34 + d)/0.43$$

An updated equation that should be used going forward, assuming that this simple model is a good approximation for localized reactions:

$$n = (\sum_i n_i^{ss} \times 0.56 + \sum_i n_i^{ds} \times 0.34 + d)/0.56$$

Appendix C

ALTERNATIVE REACTION PATHWAY FOR PINS-AND-NEEDLES

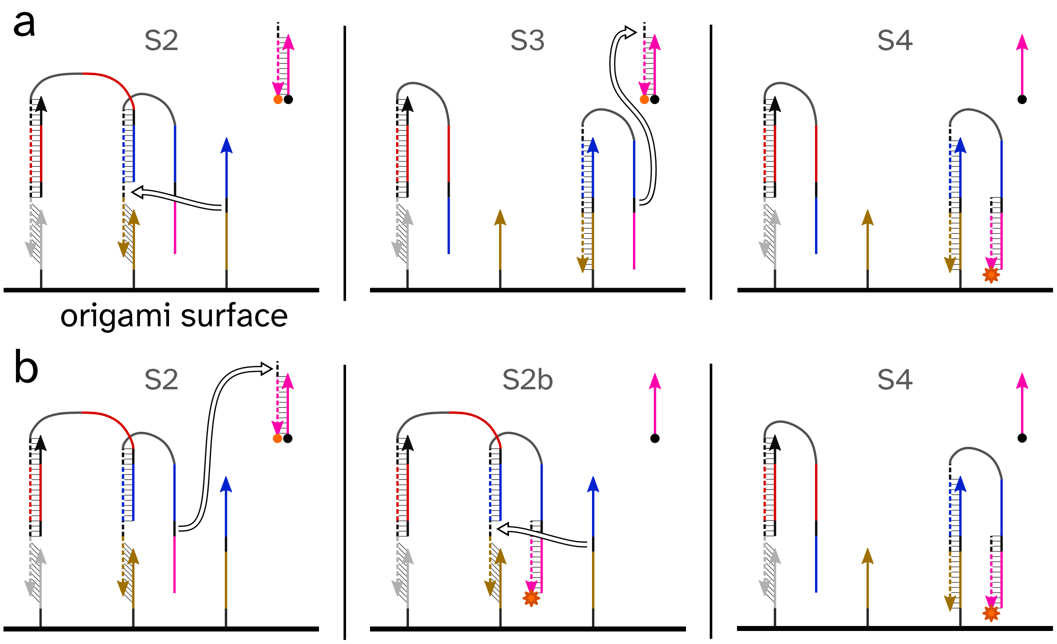


Figure C.1: **Alternative pins-and-needles reaction pathway.** **a,** The standard reaction pathway for comparison, reproduced from the main text (Fig. 4.2b S2-S4). **b,** The alternative reaction pathway.

The reporter may react with the system at state S2 rather than at state S3. If so, the system will reach state S2b (Figure C.1). From S2b, the system will not be able to return to S1 because the toehold on the output gate will be irreversibly bound to one of the reporter strands. The only possible reaction which could next take place is essentially the same reaction that changes state S2 to S3, although in this case it would change state S2b directly to state S4.

Depending on the concentrations of all species and the relative rates of local reactions and reactions involving diffusible components, this alternative reaction pathway could potentially speed up the pins-and-needles system. If the leak in the pins-and-needles system were addressed, enabling cascading the system with a series of local reactions, it should be expected that the final pins-and-needles catalytic system in the pathway would operate somewhat more quickly than all of the upstream copies; while all copies could operate by the standard reaction pathway, only the final copy could additionally follow the alternative reaction pathway discussed here.

Appendix D

TRIPLE MISMATCH VARIANT OF PINS-AND-NEEDLES

A variant of the pins-and-needles input gate containing three mismatches in the central toehold was used in a set of fluorescence experiments, providing some additional insights into the behavior of the pins-and-needles system. For this experiment, three mismatches are also present in the input trigger. For convenience of comparison, the results for the standard version of the system are provided here first (Fig.D.1a), with the results for the version of the system containing the three mismatches below (Fig.D.1c). The locations of the mismatches are diagrammed in Fig.D.1d and e.

One difference in the two experiments that is unrelated to the mismatches must be discussed first to avoid confusion. More noise is apparent in the fluorescence curves of Fig.D.1a than in Fig.D.1c. This discrepancy is purely the result of using different settings in the plate reader (only 100 measurements per data point for Fig.D.1a, but 255 measurements per data point for Fig.D.1c).

Another difference in the two experiments that is unrelated to the mismatches is the sequence of the double-stranded linker domain tethering the input gate to the origami surface. The change in this sequence was made to optimize the orthogonality between the three linker domains, but as the three sequences are drawn from a set of sequences that are all designed to be orthogonal to each other, this change is not expected to have any effect.

The mismatches would not be expected to affect the positive control reactions, which do not involve the input gate at all, or the negative control reaction, since the mismatches do not provide an obvious source of additional leak between the inactive input gate and any downstream circuit components. Indeed, none of these curves appear remarkably different between Fig.D.1 a and c.

The expected consequences of these mismatches primarily center around the reversible transition between states S1 and S2 and the transition from state S2 to S3. The first forward reaction (S1 to S2) would be expected to be slowed, since the hybridizing toehold is now essentially a 2-bp remote toehold with a 3-nt spacer. Additionally, the reverse reaction (from S2 back to S1) would be expected to be faster, since the dissociating toehold includes the three mismatches. The expected effect of these changes in the S1-S2 transition on the overall catalytic reaction is slower kinetics; however, faster kinetics would be expected for the transition from state S2 to S3 since the dissociating toehold has the three mismatches. Therefore, it is not obvious what to expect for the overall effect on the system kinetics, but experimental results indicating slower kinetics would suggest that the effects on the S1-S2 transition are much more important, while faster kinetics would suggest that the effects on the S2-S3 transition are much more important.

The experimental results demonstrate a slowdown for the first half hour; whereas in the standard system the two-pathway positive control and catalytic curves demonstrate very similar kinetics, in the system with the mismatches, the catalytic curve initially lags behind the two-pathway positive control curve. The results for the system with three mismatches also show faster kinetics later on, with an apparent shift in behavior by t=0.5h. As in the standard system, both the positive control and catalytic reactions slow down visibly by t=0.5h, but the catalytic curve takes longer to slow such that it overtakes the two-pathway positive control curve before t=1.0h. Around t=0.5h, the catalytic curve appears to have a steeper slope than the two-pathway positive control curve.

These results suggest that the expected effects on the S1-S2 transition and the S2-S3 transition might be similarly important. Slower kinetics in the catalytic curve at first may be the result of the effects on the S1-S2 transition, while faster kinetics in the catalytic curve subsequently could reflect the effects on the S2-S3 transition.

Another insight can be gleaned from the fact that the rate of the catalytic reaction exceeds that of the positive control at any point at all. Ignoring for now the leak reactions evident in the negative control curve, which can be assumed to contribute equivalently for the two-pathway positive control and the catalytic case, it might be intuited that the catalytic reaction might not ever exceed the two-pathway positive control curve since both cases involve bimolecular trigger steps and bimolecular reporting steps but only the catalytic reaction involves the added complexity of additional local steps. This intuition ignores two factors: firstly, the two-pathway positive control involves two bimolecular trigger steps (one for each pathway) while the catalytic case

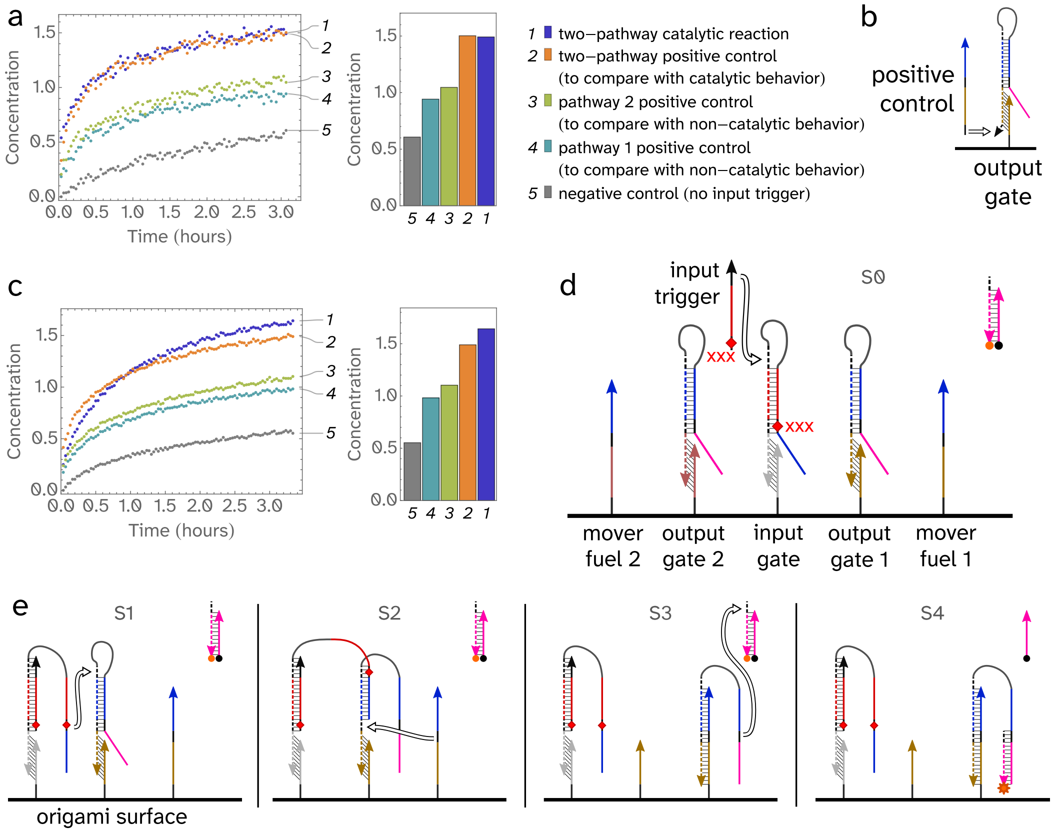


Figure D.1: **Comparison of the pins-and-needles circuit without and with three mismatches. a,** Fluorescence kinetics experiments for the standard system, reproduced from Fig. 4.4a. Raw fluorescence was normalized to relative concentration based on positive controls at 24 h. Standard concentration 1× =3 nM. Data at 3.4 h are shown in the bar chart. The concentrations of DNA origami, input trigger, and positive control triggers were 1×, 50×, and 50×, respectively. **b,** Positive control design including an additional toehold, reproduced from Fig. 4.4b. **c,** Fluorescence kinetics experiments for the system involving three mismatches. Apart from the mismatches, all other conditions are the same. **d,** Domain diagrams for the location of the three mismatches at the center of a five-bp toehold. Apart from the mismatches, this diagram is reproduced from Fig. 4.2a. **e,** Domain diagrams for the location of the three mismatches throughout a single reaction pathway. Apart from the mismatches, this diagram is reproduced from Fig. 4.2b.

involves only one for the whole system; secondly, the hybridizing toehold for the positive control trigger reaction is located closer to the origami surface than that of the input trigger. More experiments are needed to explore whether and to what extent toehold proximity to the origami surface impacts reaction rates.

Appendix E

LOCALIZED REPORTER

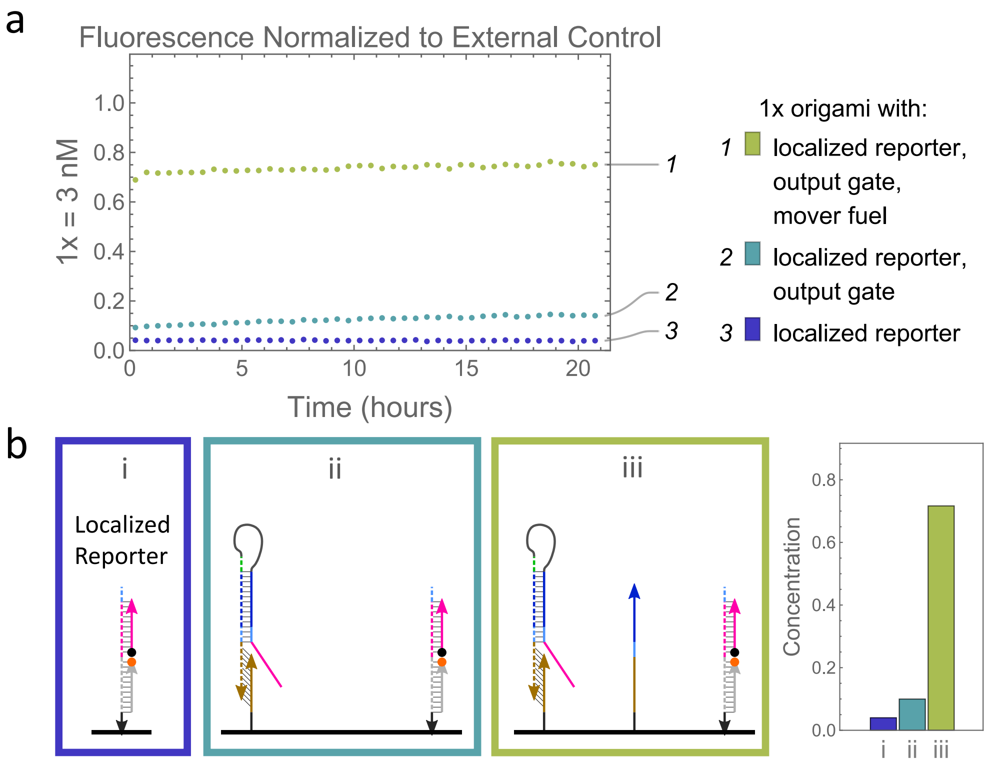


Figure E.1: **Pins-and-needles control experiment involving localized reporter. a,** Kinetics for a pins-and-needles control experiment involving a localized reporter. The number of components included varies, but in no case is the input gate present. A rapid leak reaction is shown to take place between the output gate and the mover fuel. **b,** For convenience, the same figures from Figure 4.4f are reproduced here.

While very slow, a clear gradual leak reaction is observed in curve 2 between the output gate and the localized reporter. These molecules are within the range of reachability. Given that the pink domain on the output gate is protected by a 2-bp clamp at its 3’ end, we find that protecting against toeless leak between a dsDNA complex and an ssDNA domain by traditional methods is insufficient to prevent observable leak on a DNA origami surface.

Fluorescence measurement in these experiments began ≈2 hours after the origami manufacturing process ended without an intermediate storage phase at 4°C. Curve 2 indicates that localized circuits constructed in this manner might slowly leak in storage at 4°C. If this concern were confirmed experimentally, it would indicate that these types of localized circuits should be used promptly after assembly (unless an alternate storage method such as freezing or lyophilization is used).

Appendix F

LATCH DATA PROCESSING

F.1 Latch Normalization

Unless otherwise noted, all normalization for the latch datasets took place as follows.

First, a raw dataset was produced (Fig. F.1a). Typically, a data point was collected every 2 minutes during the first three hours, as well as during the next three hours after fuel activation. After these six hours, kinetics tended to be well-captured by data points collected every 10 minutes, so this interval was used for the subsequent four hours to reduce wear on the plate reader. After this point, data tended to slowly flatten, and any further changes occurred slowly enough to be captured by data points collected every half hour.

Raw fluorescence values are known to vary with the age of the bulb in a plate reader, as well as the model of plate reader being used. Some additional variation may take place resulting from the use of a different plate or plate seal. Therefore, to ensure consistency despite any varying experimental conditions, a simple external control (Fig. F.1b) was run alongside each experiment. This control included a master mix containing the reporter(s) used in the experiment, at the same concentration used in the experiment, divided among at least two wells in the plate reader. Only buffer was added to the first of these two wells; this well served as a negative control to establish background fluorescence. To the second well was added a molecule with close similarity to the expected output of the experiment. To act as a simple positive control for calibration, this molecule was added at the experiment's defined 1× concentration. For consistency, when 1 μL of fuel activator in TE/Mg^{2+} buffer was added to all reactions containing the latch at t=3h, 1 μL of TE/Mg^{2+} buffer was added to the external control reactions. A first round of normalization was applied based on this external control (Fig. F.1c).

We quantified the concentration of origami present in each experiment by including a fluorophore-labeled staple (Fig. F.1f), which fluoresced at a different wavelength from the fluorophore on our reporter. In order to convert raw fluorescence values in this wavelength to origami concentration values, a

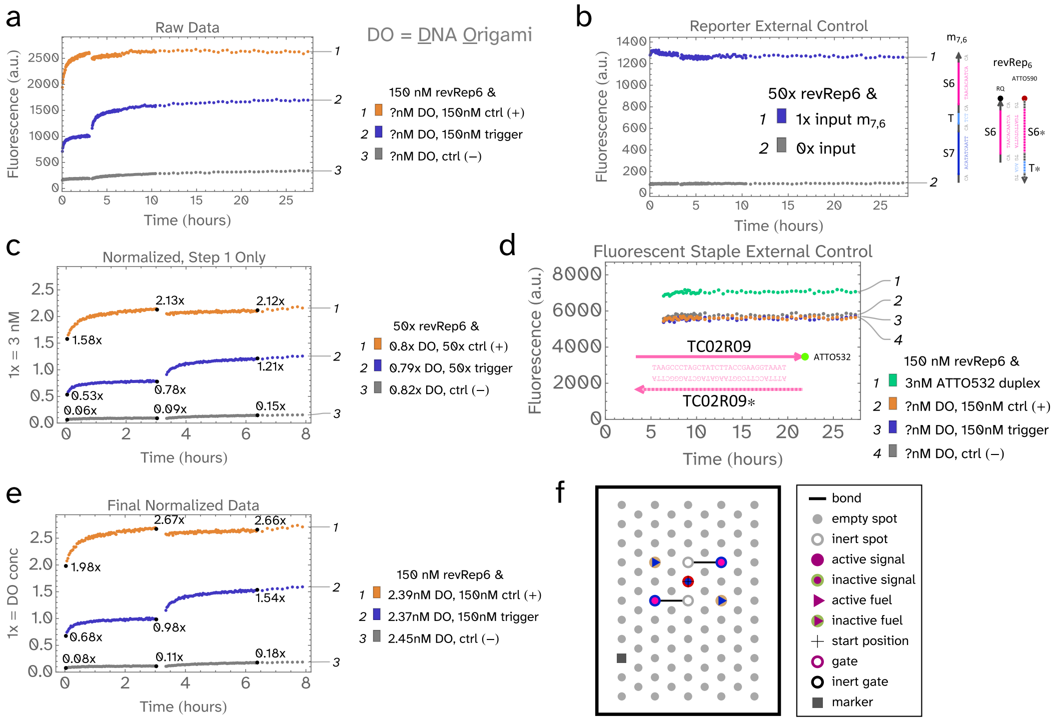


Figure F.1: **Normalization process for latch circuit experiments.** Data from the two-pathway experiment presented in Fig. 5.3 is used throughout this figure as an example. Abbreviations used in this figure: origami (ori), positive control (pos ctrl), reverse-orientation variant of Rep6 from [23] (revRep6). **a,** Raw data for a two-pathway latch. **b,** The external reporter control run alongside the two-pathway latch samples. **c,** Data for this two-pathway latch normalized based on the reporter control only. The concentrations are relative to a standard concentration of 1$\times$ $=$ 3 nM. **d,** The fluorescent staple control run alongside the two-pathway latch samples. The raw data in this plot, generated using different wavelength settings than in all other plots, measures ATTO532 rather than ATTO590. **e,** Data for this two-pathway latch normalized twice, first based on the reporter control, then by the fluorescent staple control. **f,** Layout for this two-pathway latch, showing the position of the staple bearing the fluorescent marker.

second external control was run alongside each experiment (Fig. F.1d). This external control consisted only of a duplex of the modified staple annealed to its reverse complement, present at 1× concentration of the experiment (typically 3nM). Fluorophores are known to behave differently on ssDNA vs. dsDNA, so the use of a duplex allowed us to best ensure a close comparison between the origami and the control. As in the case of the other external control described above, for consistency, when 1 μL of fuel activator in TE/Mg^{2+}

buffer was added to all reactions containing the latch at t=3h, 1 μL of TE/Mg^{2+} buffer was added to this external control.

While the target concentration of origami was typically 3nM, we observed non-negligible variation, best ascribed to inaccuracy in origami quantification on a Nanodrop coupled with typical pipetting error. Since the completion level of any curve in our experiments was directly related to the origami concentration, it was necessary to perform a second normalization in order to compare the results of different experiments (Fig. F.1e). Without this second round of normalization, plots were visually misleading, occasionally leading to false conclusions. I believe it is appropriate to normalize raw data by two external controls, as was done here, for any projects involving localized circuits on DNA origami. At minimum, a staple bearing a fluorophore should always be included to verify that origami is present at the intended concentration.

F.2 Quantitative Analysis of Latch Experimental Data

After final normalization, six data points are of particular interest for quantifying latch performance. These points are individually labeled in Figure F.2, and are used to calculate four metrics by equations listed in Table F.1.

The first metric is percent signal amplification. This quantity might be considered the most important, since it is a fundamental goal of this project to amplify a signal above $1\times$. The second metric is percent completion. This value corresponds to the discrepancy between the catalytic signal and the positive control. We would never expect this value to reach 100 percent due to sources of signal loss such as missing staples, but it remains desirable to maximize it. An excessively large discrepancy might be indicative of a positive control curve with an unexpectedly high completion level. As an aside, we penalize this metric slightly based on the amount of leak which takes place in the negative control curve before fuel activator is added (i.e., we subtract value ii from value v rather than subtracting value i from v). We assume that the same type of leak occurring in this portion of the negative control curve could also take place in the inactive pathway for the catalytic experimental case before the fuel activator is added. We subtract value i from value vi on the assumption that the positive control reaction pathway dominates the system, such that the leak pathway indicated in the negative control curve does not meaningfully contribute to the positive control curve.

The third metric is percent fuel leak, measuring the amount of leak which takes place after the fuels are activated. Since geometry in the latch prevents leak between the fuel and the gate:output complex, any fuel leak must either take place between different origami structures (R, Quantifying Inter-Origami Leak) or between the fuel and a malformed structure with geometry other than what was designed (L, Agarose Purification of Origami; N, Modified Positive Control and Duplicated Complexes). The fourth metric is percent initial leak, measuring any leak which takes place quickly in the period of time between mixing the samples and the start of plate reading (typically about 5 minutes). This type of leak could be caused by inter-origami interactions, poor synthesis quality, or any origami structures bearing an unbound output strand. Its typical low value throughout all latch experiments demonstrates that the latch design overcame the issue at the core of the pins-and-needles design.

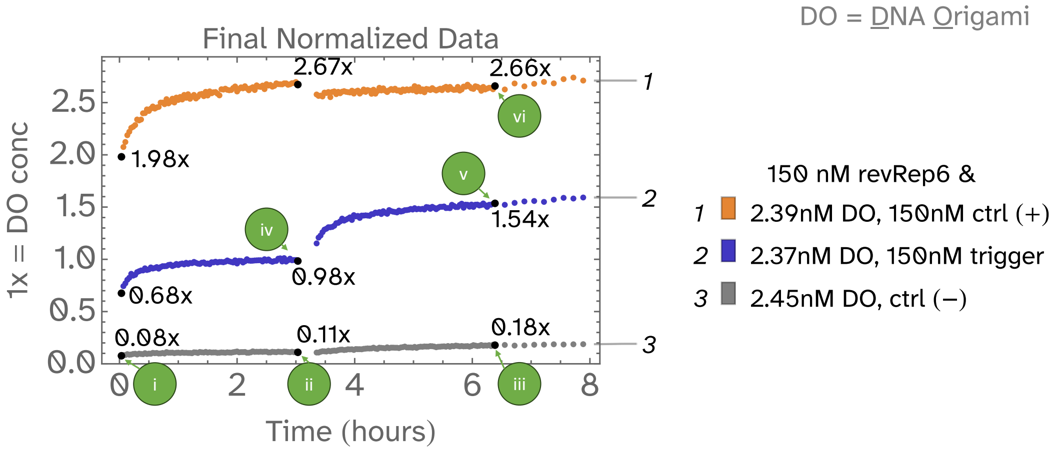


Figure F.2: **Data points used in quantitative analysis.** The final normalized data shown in Figure F.1 is reproduced here with six individual data points labeled i-vi, which are used to generate four metrics that can be used to quantify latch performance, explained in detail in Table F.1. No replicated wells were run for the experiment depicted here. In other experiments, wells were replicated; for these experiments, an average of the two corresponding data points is used as inputs to the same equations. As in other experiments, trigger and positive control strands were added at t=0h, and fuel activator was added to all reactions containing the latch at t=3h.

Metric	Equation	Example Value
% Signal Amplification	$(v - iv)/(iv - i)$	60.9
% Completion	$(v - ii)/(vi - i)$	55.2
% Fuel leak	$(iii - ii)/(vi - i)$	2.7
% Initial leak	$(i)/(vi)$	2.9

Table F.1: **Metrics used in quantitative analysis.** Metrics, and the equations used to produce them, are given here. The example metrics are calculated based on the six data points labeled in Figure F.2.

Appendix G

LATCH WITH MODIFIED REPORTER

A reporter typically used in past work from our group [1, 2, 23, 41, 94, 99, 149, 150] is called Rep6 (Figure G.1a). Initially [23, 94], the ROX fluorophore was used for this reporter, but subsequently ATTO590 was used instead. For convenience, the same molecule (Rep6) has been referred to by other names, such as Rep[1] [41], Rep1 [149], or ReporterY [150]. This reporter was used in Ch. 2 for the bidirectional ring formation, where it was referred to as reporterB. It was also used in Ch. 3 for the cooperative catalyst and Ch. 4 for the Pins-and-Needles design, where it is referred to simply as reporter.

Rep6 was not typically used with Ch. 5's Latch design for reasons discussed in Appendix K, Latch With Staple Orientation Reversed. Instead, a version of Rep6 with the strand orientation reversed was used (Figure G.1b). This modified reporter is here referred to as revRep6. Note that "rev" stands for reverse orientation, not reversible.

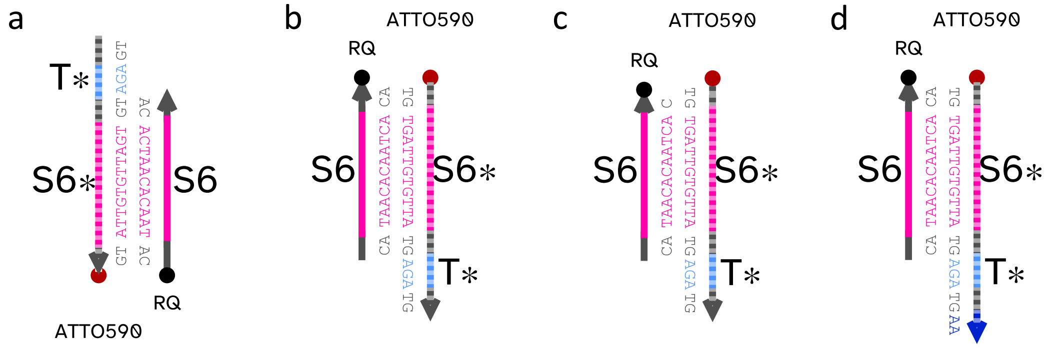


Figure G.1: **Reporter sequence-level diagrams. a,** Rep6 sequence-level diagram. **b,** revRep6 sequence-level diagram. **c,** irRep6 sequence-level diagram. **d,** frRep6 sequence-level diagram.

Recently, Rep6 was found to react reversibly with the fluorophore and quencher acting like a 3 nt toehold (unpublished work from our group). Given the high degree of similarity between Rep6 and Reverse Orientation Rep6 (revRep6), we suspect that revRep6 may also react reversibly with the fluorophore and quencher acting like a 3 nt toehold. Since the catalytic reactions of the Latch are also reversible, we added a large excess (50×) of reporter to ensure that all the desired reactions would be driven forward to near-completion. Our simulations supported this assumption, and furthermore, we tested it experimentally by using a recently developed method for making our Rep6 react irreversibly (unpublished work from our group). This method for an irreversible version of Rep6 works by deleting one nucleotide adjacent to the quencher; thus, given the high degree of similarity between Rep6 and revRep6, we assumed with reasonable confidence that the same irreversibility would be achieved by the same alteration in revRep6. Thus, we produced a likely irreversible variant of revRep6 (Figure G.1c) with a deletion at the nucleotide adjacent to the quencher, naming this reporter “irreversible” reverse orientation Rep6 or irRep6.

The same two-pathway latch variant described in Figure 5.3, also called DO.2.a in Appendix I, Latch Relocated, was used. Drawing from the same sample of origami, the reaction was repeated two days later, with the only change being the reporter. No obvious change in terms of kinetics or completion levels resulted (Fig. G.2), and furthermore the % signal amplification worsened slightly (perhaps due to experimental noise), so we concluded that using a large excess (50×) of the original revRep6 was sufficient to drive forward the reactions on the surface to completion.

Metric	DO.2.a	DO.2.a+IR	DO.2.ar	DO.2.ar+FR
% Signal Amplification	60.9	49.8	54.6	50.8
% Completion	55.2	55.7	51.4	50.3
% Fuel leak	2.7	2.5	1.8	1.7
% Initial leak	2.9	2.1	3.3	3.5
Origami batch	1	1	2	2
Reporter	revRep6	irRep6	revRep6	frRep6

Table G.1: **Effect of irreversible and fast reporters on metrics.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data.

Having determined that the possible reversibility of revRep6 had no observable effect on the system’s kinetics, we wanted to determine whether the forward reporting reaction rate was a limiting factor. To explore this possibility, a variant of revRep6 (Figure G.1d) with a 7nt toehold (rather than 5nt) was produced by adding 2nt of homology (`AA`) to the upstream branch migration domain S7. For consistency with the prior experiments involving the irreversible reporter, it would have been ideal to combine this new reporter with

DO.2.a, but the remaining volume of DO.2.a was too low. Therefore, DO.2.ar was produced, a new batch that is otherwise identical to DO.2.a. To confirm that the two batches perform similarly, the original reaction was replicated with DO.2.ar (Fig. G.2c). Indeed, the two origami batches produced similar results under the same conditions, demonstrating some experimental noise in terms of metrics (Table G.1). Simultaneously, an experiment combining DO.2.ar with the new "fast" revRep6 (frRep6) with the elongated toehold demonstrated no remarkable changes in terms of completion level, raising our confidence level in our prior conclusion that using a large excess (50×) of the original Reverse Orientation Rep6 was sufficient to drive forward the reactions on the surface. A moderate change in the initial reaction kinetics for the output signal may be observed upon close scrutiny (Fig. G.2c vs. d), but the result was ambiguous.

For consistency, the original revRep6 was used in all other reactions unless otherwise noted.

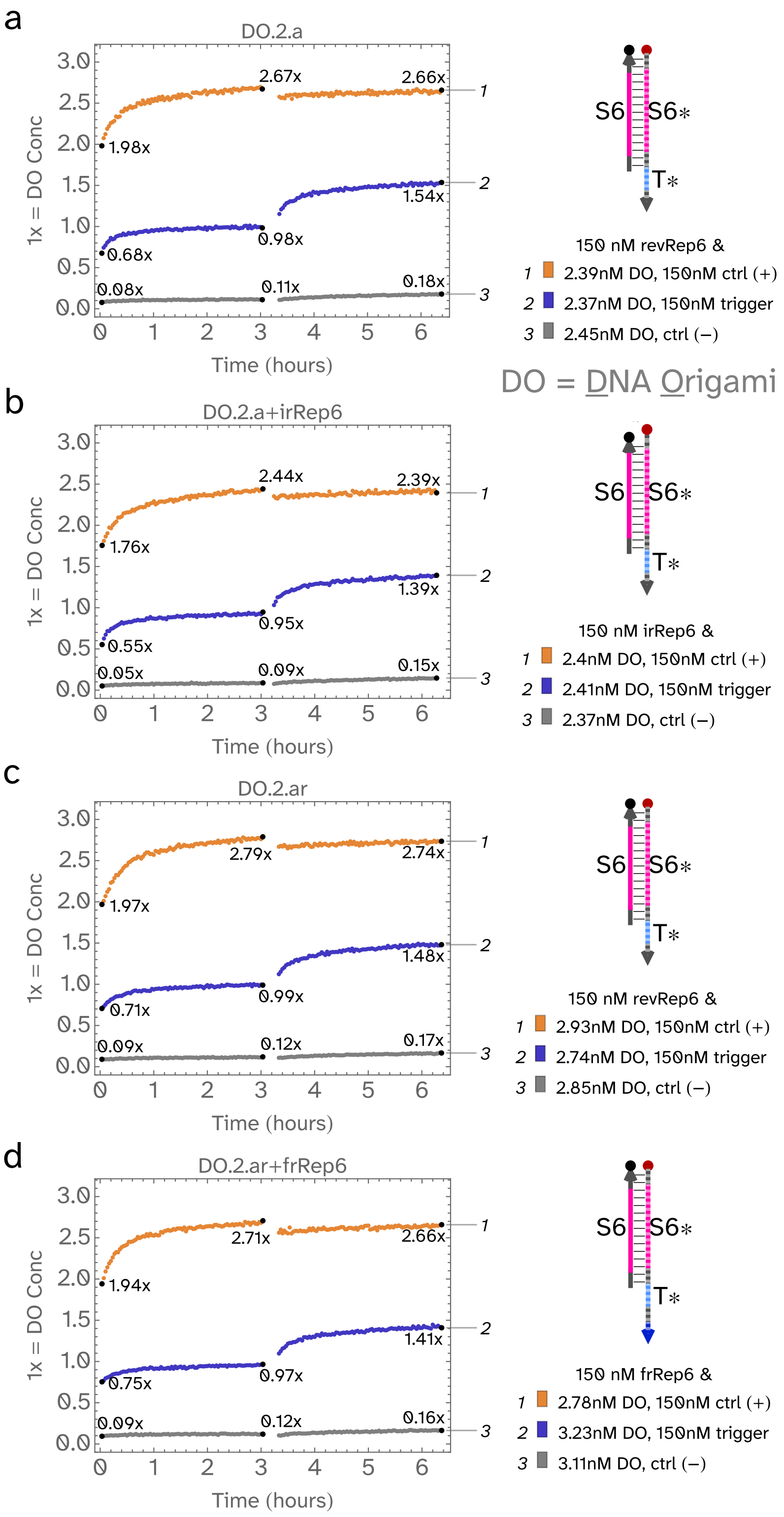


Figure G.2: **Effect of irreversible and fast reporters.** Metrics given in Table G.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Original reaction with revRep6 (DO.2.a). **b,** Reaction involving DO.2.a and irRep6. **c,** DO.2.ar is a remade batch of DO.2.a, here reacted with original revRep6. **d,** DO.2.ar reacted with frRep6.

Appendix H

LATCH WITH PURIFIED COMPONENTS

We ordered all of our strands from Integrated DNA Technologies (IDT), a DNA synthesis company. IDT's DNA synthesis process results in substantial contamination by truncated strands which increases with strand length. IDT claims that the percentage of an unpurified sample which is the intended full-length product is ≈80-90% for strands of 20-40 nt, but falls to ≈60-65% for strands of 70-90 nt [151].

The extended staple strands that serve as circuit components in the latch fall within this latter range. Nevertheless, due to the expensive nature of purified strands, we typically ordered our strands unpurified, apart from those in the reporter. To quantify the contribution of truncated contaminants on latch performance, we ordered one set of latch components PAGE-purified from IDT. IDT claims their PAGE purification process typically results in full-length oligo percentages greater than 85% [152].

The latch variants involving PAGE-purified components were located at the same position of the latch in DO.2.b, as discussed in Appendix I, Latch Relocated. Therefore, results for DO.2.b are included in Figure H.1 and Table H.1 for comparison.

Origami variants bearing latches including PAGE-purified components are here named DO.P. In DO.P.a (Fig. H.1b), only the input strand is PAGE-purified. In DO.P.b (Fig. H.1c), conversely, all components are PAGE-purified except for the input strand. Finally, in DO.P.c (Fig. H.1d), all components including the input strand are PAGE-purified.

As seen in Table H.1, all metrics but the initial leak were improved when all circuit components were PAGE-purified. However, the effects of including some PAGE-purified components and some unpurified components were mixed. Additionally, the cases with more PAGE-purified components saw an unexpected increase in the positive control output, which already exceeded 2×.

These issues are best explained by the presence of duplicated complexes, as discussed in Appendix O, Attempted Measurement of Duplicated Complexes with Staple Translators. Without any duplicated complexes, it would not be possible for the trigger curve to exceed 1× before fuel activation at T=3h, as occurred in Fig. H.1c and d. Overall, we were able to rule out the use of unpurified strands as the major factor causing undesired latch behavior, although it was confirmed to be a minor contributor.

Metric	DO.2.b	DO.P.a	DO.P.b	DO.P.c
% Signal Amplification	61.5	48.1	78.2	67.3
% Completion	52.9	56.7	60.2	65.7
% Fuel leak	4.8	3.5	6	2.7
% Initial leak	2.3	2	6.3	3.8
input PAGE purified	no	yes	no	yes
all else PAGE purified	no	no	yes	yes

Table H.1: **Metrics for latch variants without and with PAGE-purified components.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data.

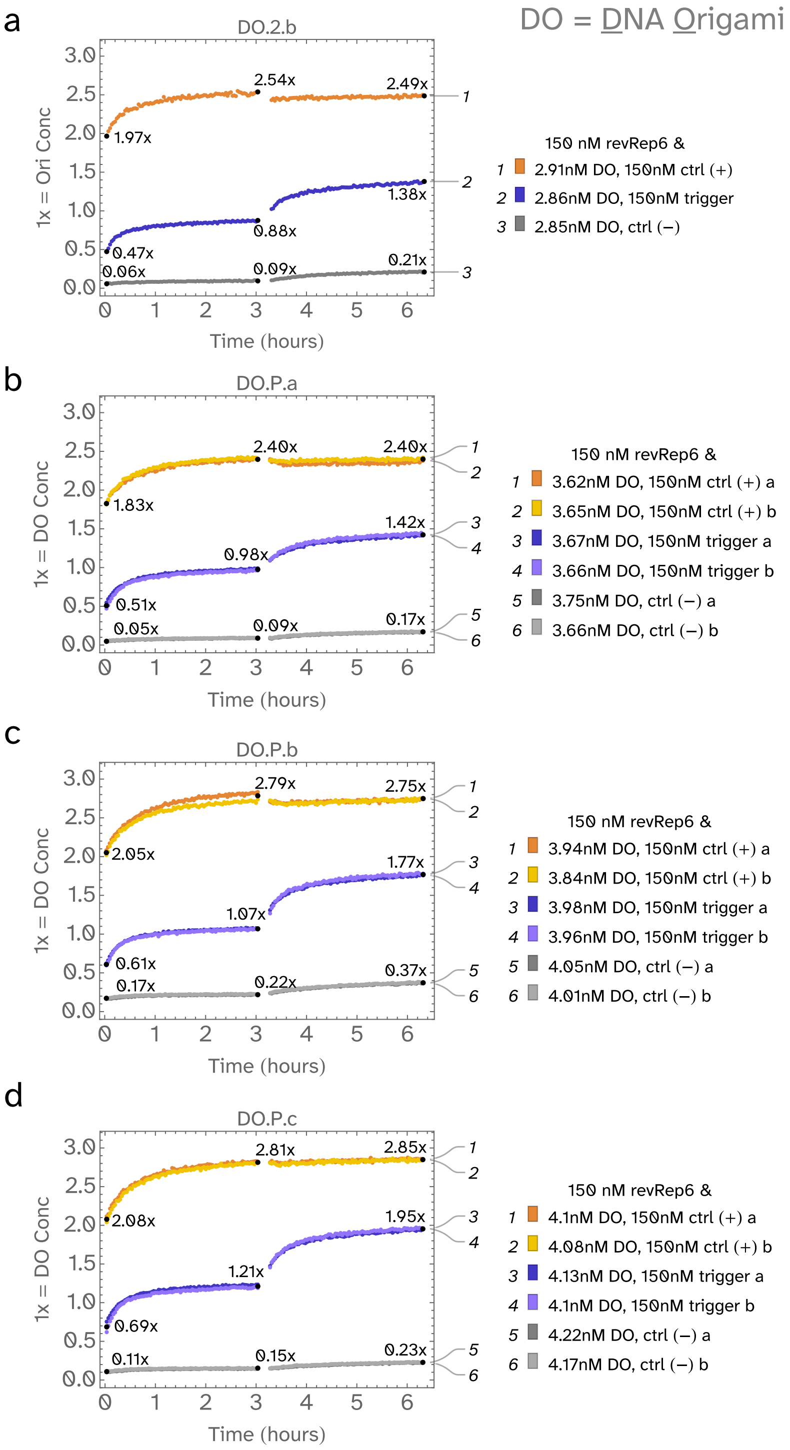


Figure H.1: **Effects of using PAGE-purified components.** Metrics given in Table H.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** No components PAGE-purified. **b,** Input PAGE-purified. **c,** All circuit components except input PAGE-purified. **d,** All circuit components PAGE-purified.

Appendix I

LATCH RELOCATED

Experiments to determine whether the latch position on the origami affects its function are summarized here. Potential causes of different results include the established inconsistency in staple incorporation rate at different positions [130] and slight variations in surface geometry. In addition, if some staples were not fully incorporated into the origami, their sequences would vary at each location, resulting in location-dependent spurious interactions with circuit components.

To distinguish each of these similar origami variants, each is given a name containing a number (the number of pathways) and a letter (corresponding to their figure label). For example, the top and bottom one-pathway variants described in Fig. 5.3 are named DO.1.a and DO.1.b respectively, and the two-pathway variant previously described in Fig. 5.3 is here named DO.2.a (Fig. I.1a).

We originally planned to assemble a cascading series of multiple latches on one origami surface, branching out from the center. To that end, we explored a variant where the latch was shifted down by one row, flipped from a left spiral to a right spiral, and rotated clockwise by 60° (Fig. I.1b). Unlike DO.2.a, which only involves extensions from staples that are fully located in the top layer, DO.2.b involves a gate strand extended from a staple that joins the two layers of double-layer origami. The distinction between the two staples is diagrammed in Figure I.2, where green staples are internal staples located entirely within one layer, yellow staples are internal staples that connect two layers, and black staples are edge staples that connect two layers. Edge staples that are fully located within one layer have no valid sites for extensions and are thus omitted from Fig. I.2, but they are included in the full Scadnano diagram (Fig. I.6) where they are colored red. A valid extension site must face outward, promoting extensions that project orthogonally from the origami surface and ostensibly impeding extension sequestration (see Appendix A.1, Staple Incorporation Background). Positions at the left and right edge of this origami structure do not face outward; instead, they face in-plane

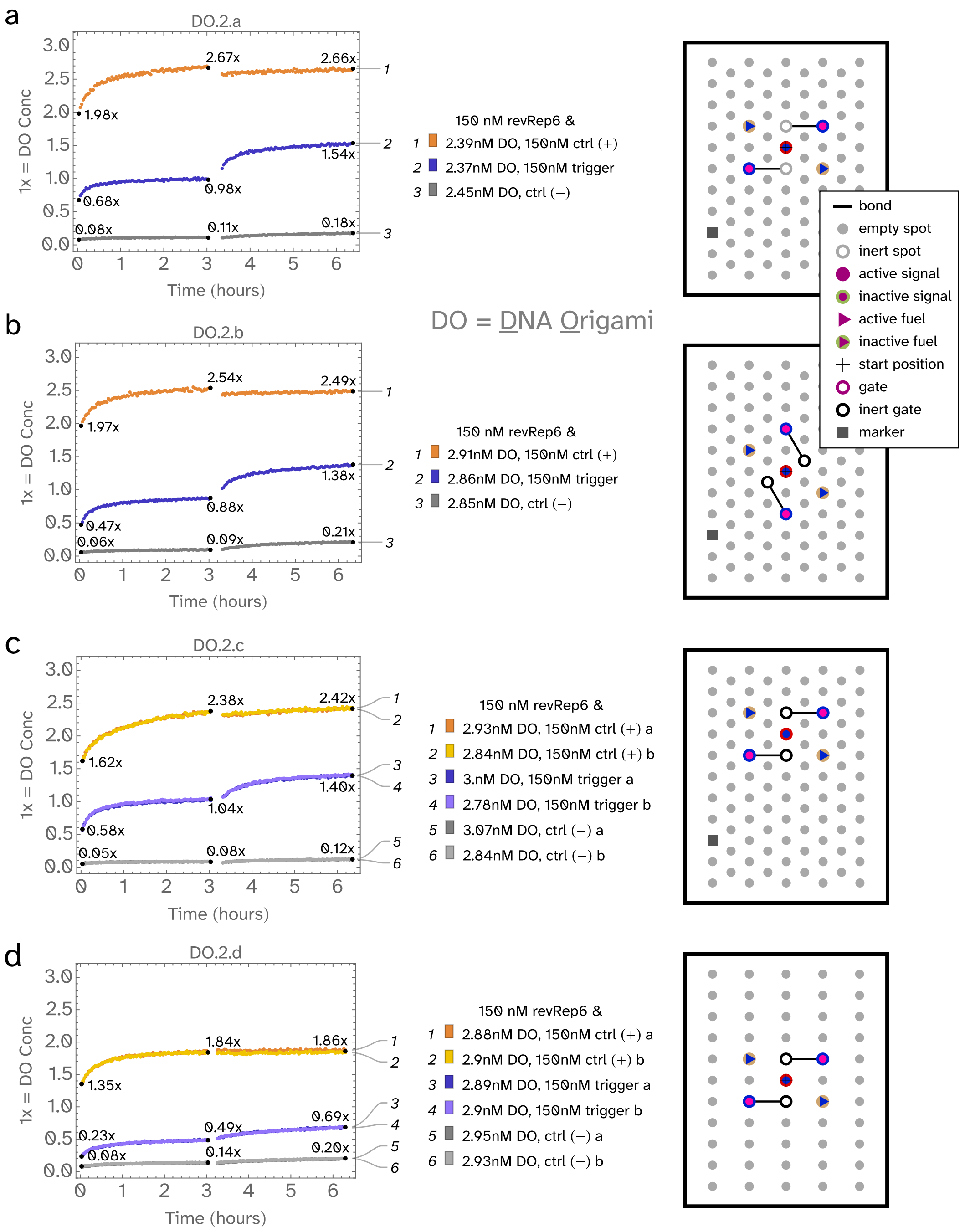


Figure I.1: **Kinetics of latch placed near the center.** Metrics given in Table I.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Original latch location. **b,** Latch shifted down one row and rotated 60° clockwise. **c,** Latch shifted up one row compared to original. **d,** Latch on the bottom layer.

with the origami surface, comparable to sites for staple crossovers. Thus, they are not valid extension sites. Additionally, they are modified with 4(C) extensions to inhibit stacking bonds [131], thereby reducing undesired origami tile aggregation. Even if these positions were valid extension sites in terms of geometry, modifying them with extensions bearing circuit components rather than 4(`C`) could increase the risk of tile aggregation.

Metric	DO.2.a	DO.2.b	DO.2.c	DO.2.d
% Signal Amplification	60.9	61.5	36.6	49.2
% Completion	55.2	52.9	55.5	30.8
% Fuel leak	2.7	4.8	1.6	3.7
% Initial leak	2.9	2.3	2.1	4.2
Input row,column	5,5	6,5	4,5	6,5
Angle	0°	60°	0°	0°
Handedness	L	D	L	L
Origami side	top	top	top	bottom

Table I.1: **Metrics for two-pathway latch variants located near the origami center.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data. Handedness is determined by defining the positive z-axis by the input projecting out of the origami surface toward the viewer, defining the positive x-axis by a gate, and considering whether that gate is bound to an output in the positive or negative y-direction (corresponding to D or L respectively). **a,** Original latch location. **b,** Latch shifted down one row and rotated 60° clockwise. **c,** Latch shifted up one row compared to original. **d,** Latch on the bottom.

Despite the difference in types of staples used, the behavior of DO.2.a and DO.2.b are very similar, yielding metrics (Table I.1, DO.2.a vs DO.2.b) that fall within the range of experimental noise observed between different batches of the same origami (Table G.1, DO.2.a vs DO.2.ar). Without considering DO.2.c or 2.d, this observation suggests that the system behavior is fairly consistent when placed near the center of the origami, and that extending staples that join the two layers does not meaningfully affect the latch formation. This latter conclusion, if true, is initially surprising. We might expect the origami folding process to have some effect upon the correct formation of the latch's two pathways; during the annealing process, the gate from one pathway may be equally likely to bind the output from the intended pathway or an output from the opposite pathway if this binding event occurs before the staples have joined the origami structure. Additionally, if such undesired combinations form and bind the scaffold by one staple, the gate:output domain is not physically long enough to allow both staples to remain bound once the complete origami structure has formed. We might therefore expect these undesired output:gate combinations to be displaced away by the correct combination of output:gate since only the latter can physically reach all of the positions on the scaffold once all staples have joined; however, whether such a corrective reaction will take place before the end of annealing might depend on its kinetics and thus the particular types of staples involved. The yellow staple that bears DO.2.b's gate strand, the one which joins the two layers of double-layer origami, would be expected to bind to the bottom layer of origami first during the origami formation process, since that is where it has the longest continuous domain. Intuition would suggest that the loop closure entropy that any staple must overcome to join two scaffold domains within one layer of this origami structure should intuitively be much smaller than the entropy of closing the loop between the two layers, since the latter represents a much larger distance on the scaffold; however, while this type of intuition likely holds true for the staples which first join a folding structure, it has been shown that loop closure entropy for loops which have not yet been closed rapidly shrinks as a result of early loop closure events (see Fig. 5C in [144]). As a result, the sequence of loop closure events cannot be readily predicted based on the loop sizes measured on the unfolded scaffold (see Fig. 5B in [144]). In the absence of a sufficiently detailed simulation for modeling the sequence of events during origami folding (which will be discussed in Appendix P.2, Approach: Seed Staples and Autobreak), it is impossible to state with any confidence whether the 3' terminus of the yellow staple bearing one of the gates in DO.2.b will be folded later than any of the 3' termini of the green staples in DO.2.a or DO.2.b. As a result, the similar performance of DO.2.a and DO.2.b does not allow us to draw any conclusions about the effect of the origami folding process upon the formation or function of the latch.

For reasons that will be explained later in this thesis (Appendix O, Attempted Measurement of Duplicated Complexes with Staple Translators), we wished to explore placing the latch at a site where staple secondary structure is minimal. We found that an optimal position (explained in detail in Appendix O.1, Selecting a Location with Minimal Secondary Structure) for limiting staple secondary structure was one row above the original latch position, and we tested this variant as well (Fig. I.1c).

These experiments were sensitive. If a negative control well became contaminated with an aerosolized droplet of a neighboring experimental or positive

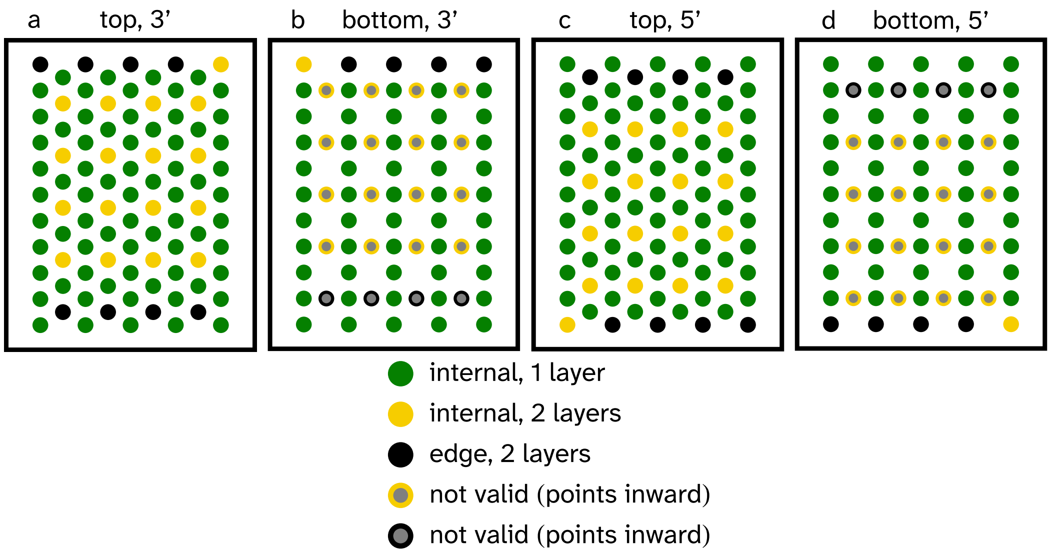


Figure I.2: **Double layer origami staple layout.** Color-coded diagrams of staples, where green dots correspond to internal (not edge) staples fully located within one layer, yellow dots correspond to internal staples crossing between two layers, and black dots correspond to edge staples crossing between two layers. Solid dots represent valid extension sites. Staple extensions placed at the location of dots with a gray center would be expected to point from the bottom layer inward toward the top layer, so these are not valid extension sites. They are included here to highlight the potential for confusion which could result from exclusively using the Scadnano diagram to design localized circuit layouts. **a,** Top layer extension sites, 3’ end. Commonly used throughout this thesis. **b,** Bottom layer extension sites, 3’ end. Used for the latch variant in Fig. I.1d. **c,** Top layer extension sites, 5’ end. Used for the latch variant in Fig. K.1b. **d,** Bottom layer extension sites, 5’ end. Not used in this thesis.

control well (as might conceivably take place during mixing via multichannel pipette), that negative control well would show radically different results from what would have occurred in the absence of contamination. To avoid the possibility of drawing false conclusions based on single wells affected by this hypothesized contamination, starting with the experiment corresponding to Fig. I.1c, it became standard practice to make a double volume of master mix to spread across two replicates of each experiment. Curves from both row a and b are shown, but since the results were nearly identical, the curves from one replicate often obscure the other at a typical zoom.

Strangely, for DO.2.c, the % signal amplification (Table I.1) is strikingly worse than that of DO.2.a and DO.2.b, while its other metrics fall within the range of experimental noise. One possible explanation is that DO.2.c is placed one row nearer to the origami edge than DO.2.a. As will be shown later in this section, latches placed near the origami edge perform worse. Identifying the underlying cause of this phenomenon would require experiments and simulations beyond the scope of this thesis.

For free-floating DNA origami in a test tube, there is no intrinsic top or bottom. In solution, all origami tiles are expected to be oriented randomly. On the mica surface used in AFM imaging, some proportion—biased by factors such as extensions and tile shape—of a population of origami tiles typically lands with a given side facing up, though it is possible to design origami such that it will land on silica surfaces oriented at a desired angle with a desired side facing up [153]. For the double-layer origami structure, extensions from either layer point outward, away from the origami structure, as desired for arranging decorative elements, functional molecules, or circuitry. The only important distinction between the two layers is that one has fewer sites available for extensions. The layer with more sites available was typically used to arrange DNA robots in past work, so to align with human intuition of arranging objects on top of a surface, the layer with more extension sites available was named the "top layer", and the layer with fewer extension sites was named the "bottom layer" [154]. Fortunately, there are enough sites in the proper locations on the bottom layer to support at least one copy of the Latch. While we would not expect to normally arrange circuits on the bottom layer, we wanted to confirm our basic assumption that a Latch placed on the bottom layer would behave no differently from a Latch placed on the top layer. Therefore, I produced an origami variant bearing a latch located at a position similar to that of DO.2.a, but on the bottom side of the origami rather than the top (Fig. I.7). The resulting origami variant, DO.2.d (Fig. I.1d), performed worse in terms of % signal amplification and % completion when compared to DO.2.a and DO.2.b (Table I.1), beyond what we would typically expect for the natural experimental variation in origami samples. Unlike DO.2.c, DO.2.d is not located closer to the origami edge, so that line of reasoning can be discarded here. Additionally, while DO.2.d has more neighboring staples of the yellow type (connecting the two origami layers) than any of DO.2.a–c, the notion that the presence of these neighboring staples could explain the poor performance of DO.2.d is challenging to support given the strong performance of DO.2.b which extends one gate strand from a yellow staple. All of DO.2.d’s staples are of the green (within one origami layer) variety. Nonetheless, this type of explanation remains plausible and worthy of future exploration. As stated above, the sequence of staples joining the origami structure would be expected to affect proper formation of the latch, and since it that sequence is currently unknown, there is a reasonable possibility that one layer folds substantially earlier than the other. The only other differences being in the staple sequences, the most likely alternative explanation is that the rates of forming undesired complexes on the surface increased due to interactions between the staples and their extensions during annealing. Another possibility is worse synthesis quality in these staples compared to the originals (both ordered unpurified). To eliminate that possible factor, the experiments would have to be redone with staples ordered purified.

Our original plan for a branching cascade of multiple latches would have required making full use of the origami surface. We therefore explored the effect of relocating the latch to the edges (Fig. I.3), where we observed much worse performance (Table I.2) as compared to the center. These results are particularly likely to be caused by worse staple incorporation rates (Appendix A, Staple Incorporation Rates), assuming that the double-layer structure has worse staple incorporation near the edges as was shown for single-layer rectangles [130].

Metric	DO.2.e	DO.2.f	DO.2.g	DO.2.h
% Signal Amplification	51.6	42.5	92.4	48.8
% Completion	28.2	20.7	10.5	15.3
% Fuel leak	6.2	3.3	5	3.8
% Initial leak	5.3	6.3	6.5	5.3
Input row,column	3,5	3,5	9,5	9,5
Handedness	D	L	D	L

Table I.2: **Metrics for two-pathway latch variants located near the origami edge.** These values correspond to the plots in Fig. I.3. Note that variants with D handedness have an output in an edge position and variants with L handedness have a fuel in an edge position. The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data.

Each two-pathway experiment involves summed effects. To better distinguish individual contributors to our results, we performed related one-pathway experiments, as shown in Figure I.4 and Figure I.5.

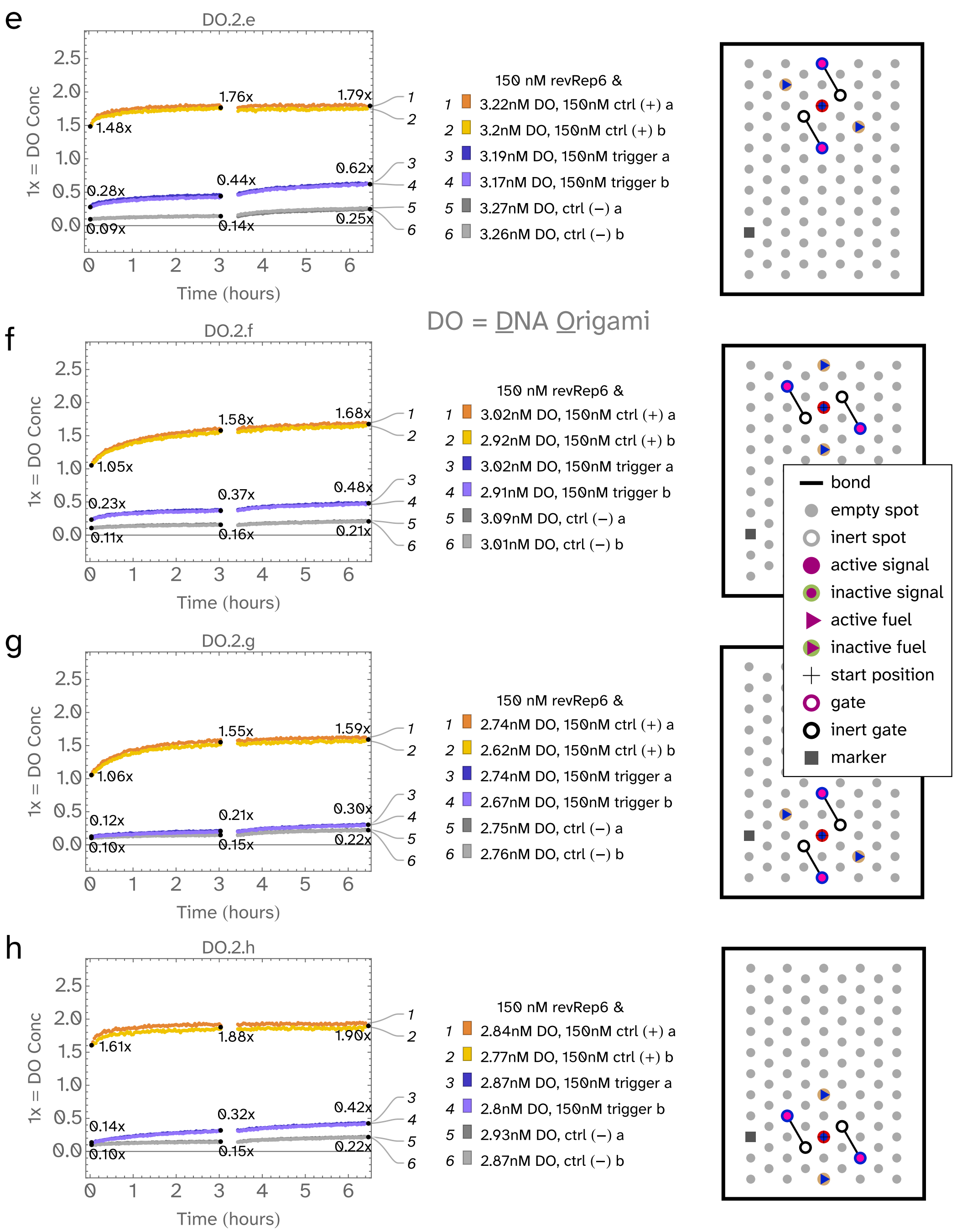


Figure I.3: **Kinetics of latch placed near edge.** Metrics given in Table I.2. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **e,** An output in the first row. **f,** A fuel in the first row. **g,** An output in the last row. **h,** A fuel in the last row.

While similar kinetics are observed across all of the one-pathway cases when the input is activated, the rate of the reaction involving the positive control varies drastically at different positions. These different positive control rates

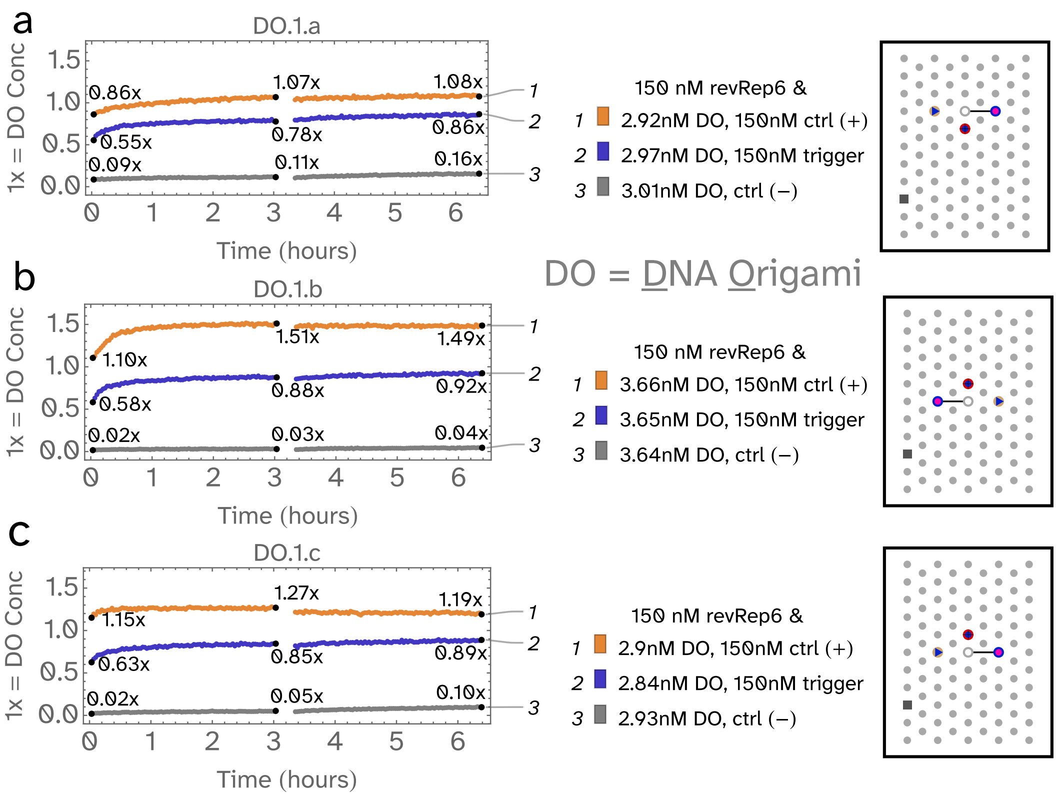


Figure I.4: **Kinetics of one-pathway latches.** Metrics given in Table I.2. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** The top pathway of DO.2.a. **b,** The bottom pathway of DO.2.a. **c,** A reversed variation on the bottom pathway of DO.2.b.

are largely obscured in the two-pathway variants, since those plots can only show the summed result of two different positive control reaction pathways. This phenomenon cannot be explained by missing staples or by extensions being sequestered on the wrong side of the structure. It is strong evidence for a different mode of circuit malformation, complex duplication, which will be discussed in detail in Appendix N, Modified Positive Control and Duplicated Complexes.

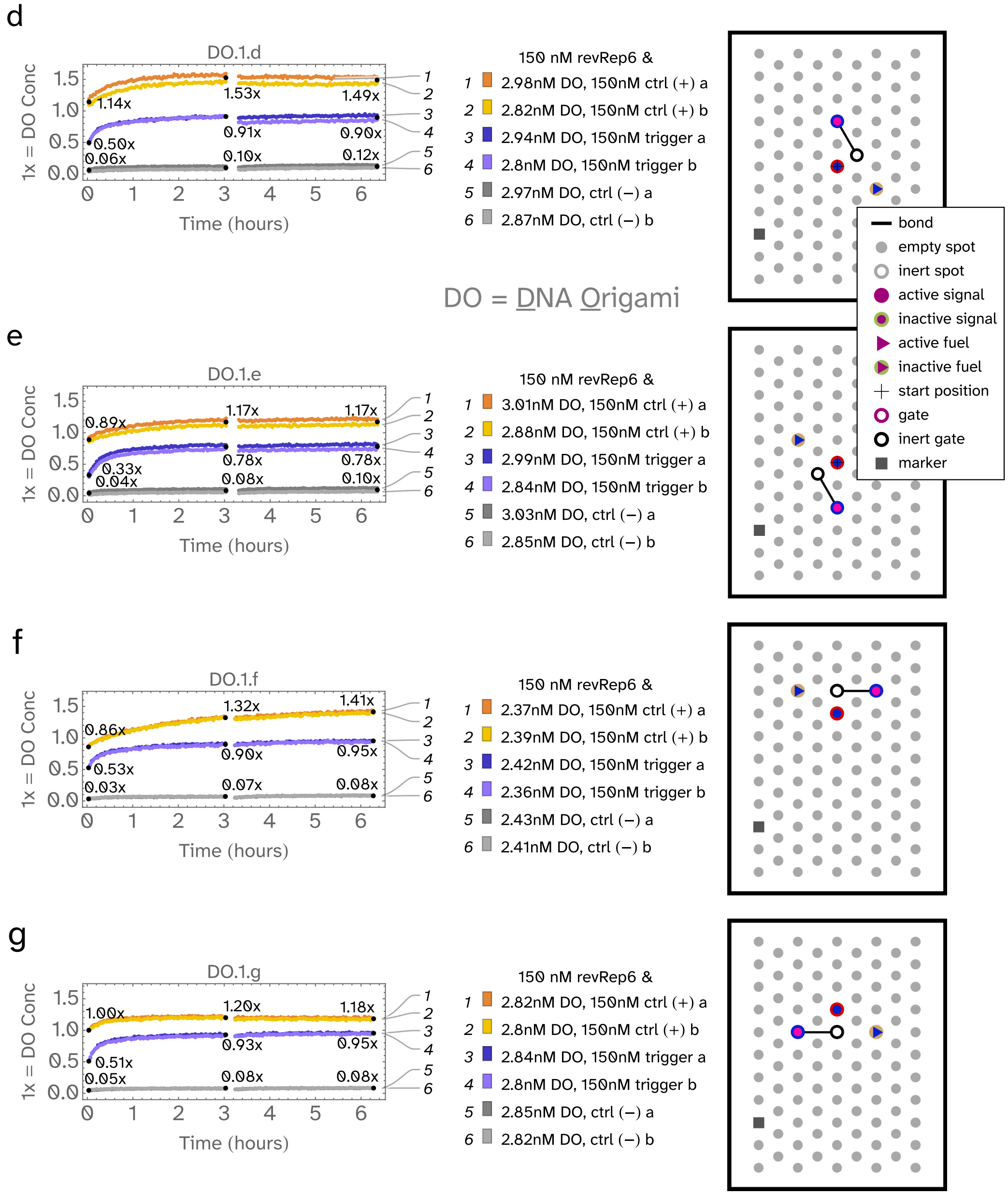


Figure I.5: **Kinetics of one-pathway latches, continued.** Metrics given in Table I.2. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **d,** The top pathway of DO.2.b. **e,** The bottom pathway of DO.2.b. **f,** The top pathway of DO.2.c. **g,** The bottom pathway of DO.2.c.

Metric	DO.1.a	DO.1.b	DO.1.c	DO.1.d	DO.1.e	DO.1.f	DO.1.g
% Completion	75.8	60.8	71.4	55.7	61.9	63.8	76.7
% Fuel leak	4.2	3.7	1.0	1.4	1.1	0.8	0.1
% Initial leak	7.9	1.6	1.1	3.9	3.5	2.3	3.9
Input row	5	5	5	6	6	4	4
Input column	5	5	5	5	5	5	5
Angle	0°	180°	180°	60°	240°	0°	180°
Handedness	L	L	D	D	D	L	L

Table I.3: **Metrics for one-pathway variants.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data. % Signal Amplification values are not given for one-pathway experiments because signal amplification cannot take place with only one pathway.

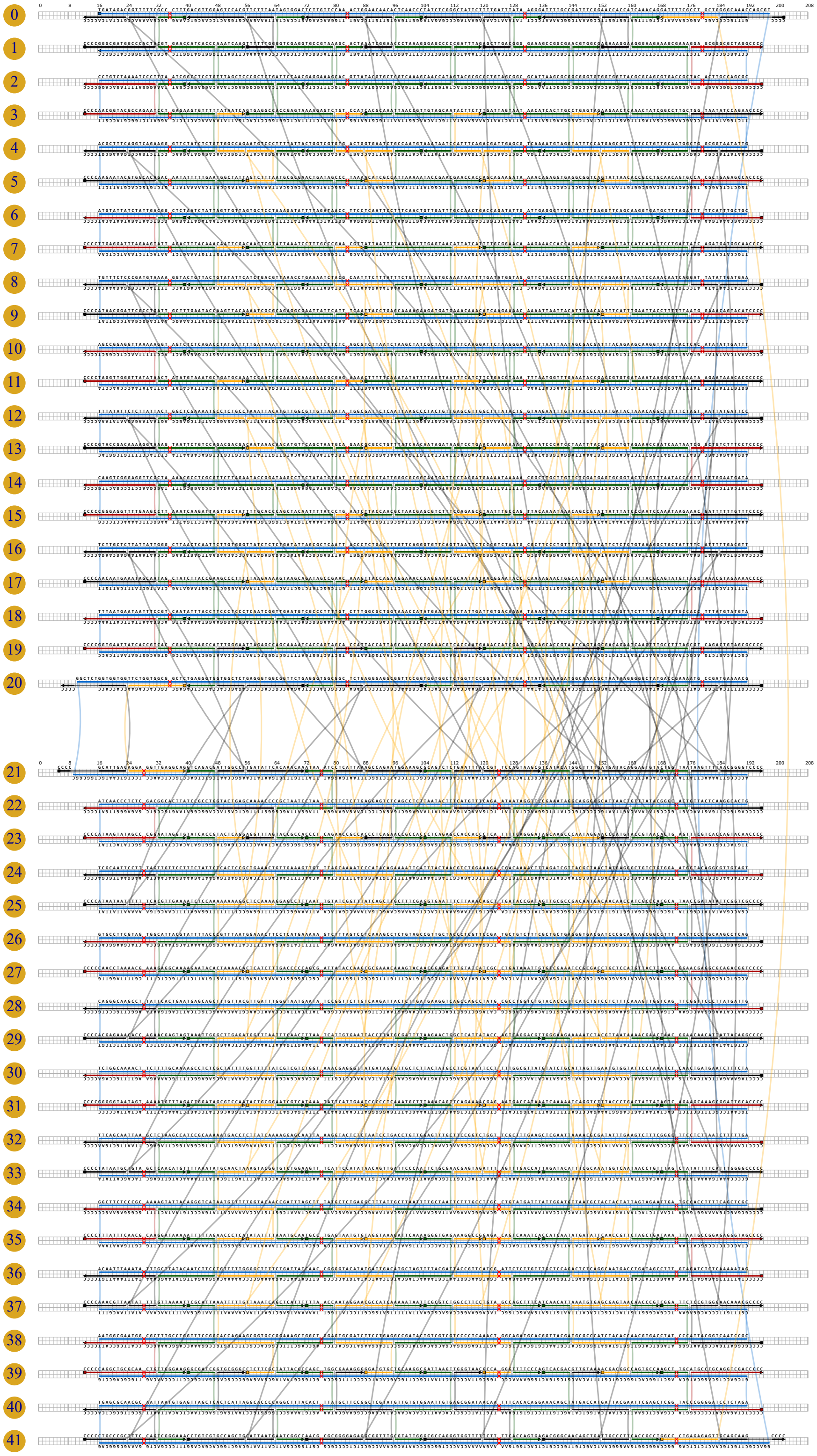


Figure I.6: **Double-layer origami Scadnano diagram.** The top layer spans helices 0-20. In reality, the bottom layer (helices 21-41) is rotated 90° counterclockwise relative to its appearance in this diagram.

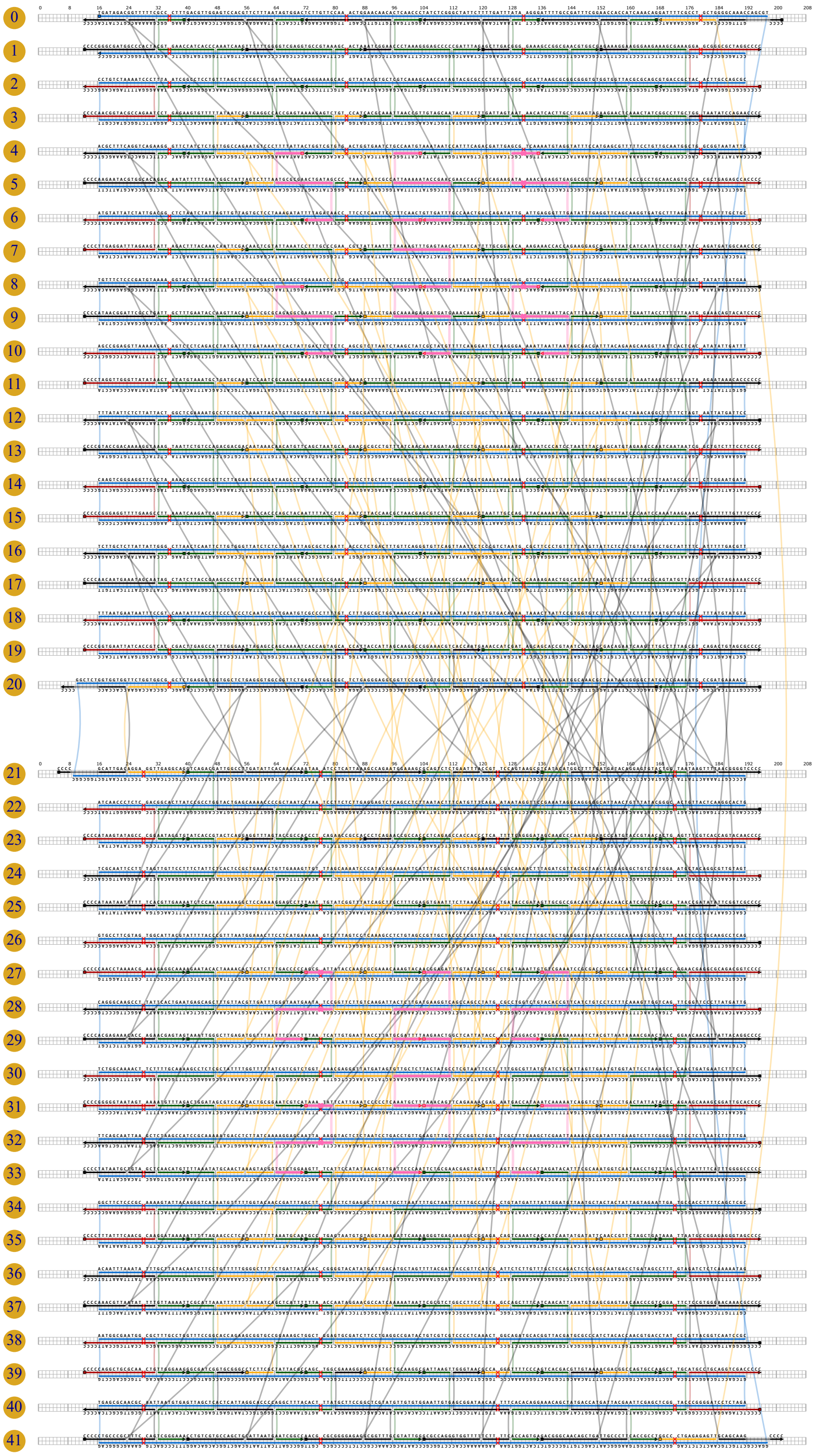


Figure I.7: **Scadnano diagram with highlights.** In pink: staples on the top layer used for the latch in Fig. 5.3 (DO.2.a), and staples on the bottom layer used in DO.2.d.

Appendix J

LATCH RE-ORIENTED

All latch variants discussed elsewhere in this thesis are laid out in the trans orientation, meaning that the two pathways face in opposite directions. This orientation is ideal for a branching cascade, where the latch is applied toward fan-out. Conversely, for the purposes of a linear cascade and signal restoration, it would be best to use a cis orientation where the two pathways face in the same direction. In cis, the two pathways could ostensibly trigger the same downstream translators.

fall within the range of experimental noise observed between different batches of the same origami (Table G.1, DO.2.a vs DO.2.ar)

Focusing on variants located near the center of the origami structure, we found that the cis-oriented latch, here named DO.Cis (Fig. J.1a, Table J.1), behaved considerably worse than trans-oriented latch variants such as DO.2.a and DO.2.b in I.1. Specifically, its catalytic property (% signal amplification) was much lower, at 36.1 compared to DO.2.a's value of 60.9. DO.Cis should be expected to behave similarly to DO.2.a, since the only difference between the two is the position of the output and fuel in the bottom pathway. We can only speculate as to why this loss in function took place, since the one-pathway results for the bottom pathways in trans and cis are fairly similar (compare Fig. I.4 b and c).

When the cis-oriented latch was relocated to the corner of the latch (here named DO.Cis.b), its output signal dropped (Fig. J.1b), much in the same way as the trans-oriented latch lost output signal when moved near an edge (Table I.2). Additionally, while DO.Cis.b performed worse in terms of signal amplification than any of the trans-oriented variants located near the origami edge (DO.2.e–h), this comparison is less meaningful since DO.Cis.b is placed in a corner and may suffer from two edge effects rather than just one. For the purposes of comparison with DO.2.e–h, it would have been more appropriate to use a variant located at their same position, but the corner location was chosen to evaluate whether a particular design for a latch cascade would be feasible. That proposed cascade layout was discarded due to the poor performance of DO.Cis.b.

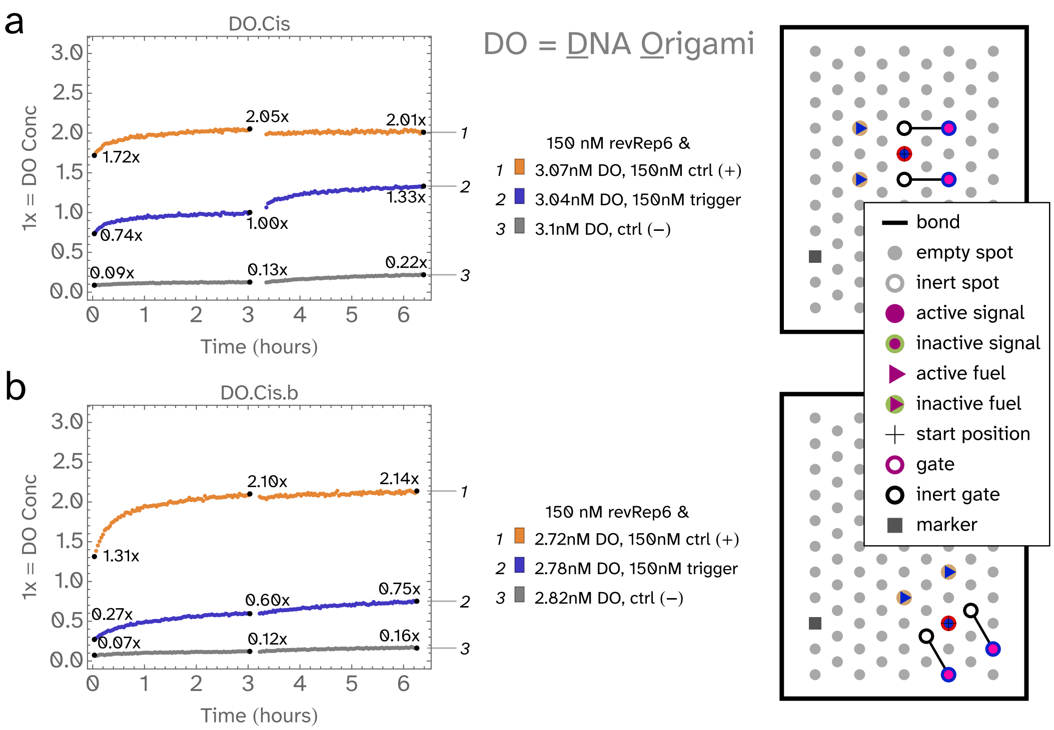


Figure J.1: **Kinetics of cis-oriented latch variants.** Metrics given in Table I.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Cis-oriented latch in the same position of DO.2.a, the original latch location. **b,** Cis-oriented latch relocated to a corner.

Metric	DO.Cis	DO.Cis.b
% Signal Amplification	36.1	30.2
% Completion	62.7	30.6
% Fuel leak	4.8	2
% Initial leak	4.3	3.4
Input row,column	5,5	9,7
Angle	0°	60°

Table J.1: **Metrics of cis-oriented latch variants.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data. **DO.Cis,** Cis-oriented latch in the same position of DO.2.a, the original latch location. **DO.Cis.b,** Cis-oriented latch relocated to a corner.

Appendix K

LATCH WITH STAPLE ORIENTATION REVERSED

IDT synthesizes its strands in the direction 3’ to 5’, the opposite of natural DNA synthesis by DNA polymerase [151]. IDT adds each base in a strand one at a time, but each base has some small probability of failure to attach to the growing oligo. IDT claims that their method avoids deletions by capping any strand which failed to include a base, such that samples are contaminated by strands truncated at the 5’ end. Staples for origami are typically ordered unpurified, yet origami structures form anyway, suggesting that incorporating some staples with truncation is acceptable for forming the correct structure or that full staples can displace truncated staples with the unbound portion of the scaffold acting as a toehold.

Therefore, to facilitate the use of unpurified strands as extended staples, we designed the latch to be extended from the 3’ end of staples only. We expected that any strands with truncation at the 5’ end will either be missing the entire staple domain and thus fail to be incorporated into the origami structure, or will be missing only part of the staple domain. The latter possibility would not be problematic for the reasons discussed above, and because any unincorporated staples or circuit components are washed away during the Amicon filtration step (5.4, Materials and Methods).

We explored one variation on the latch with a reversed orientation, such that all strands are extended from the 5’ end. We expected this variant to perform worse, since the circuit components would include some truncation. Indeed, as demonstrated in Figure K.1, the output signal fell far below typical levels. This experiment establishes a design technique: assuming strands in localized DNA circuits are ordered from DNA synthesis companies which produce their strands from 3’ to 5’, all staple extensions should be placed at the 3’ end of staples only, as was done in the majority of our experiments.

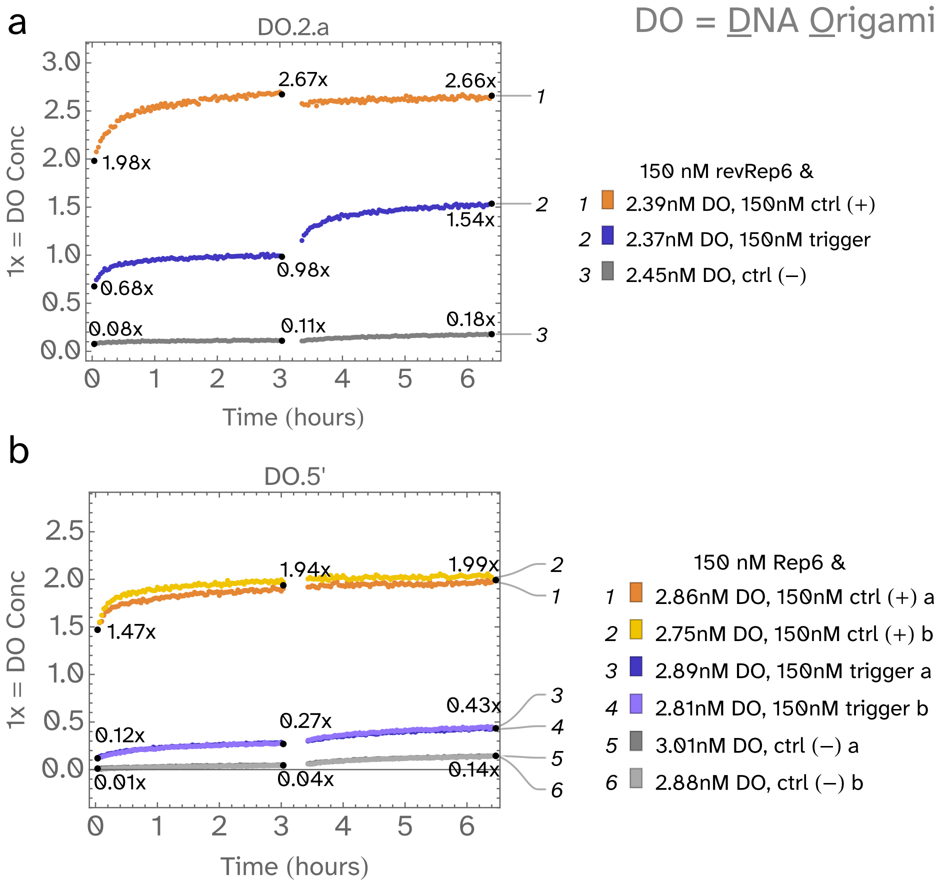


Figure K.1: **Effects of extending staples from their 5' ends rather than 3'.** Metrics given in Table K.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Original version with components extended from staple 3' end. **b,** Variant with components extended from staple 5' end.

Metric	DO.2.a (3')	DO.5'
% Signal Amplification	60.9	63.3
% Completion	55.2	19.6
% Fuel leak	2.7	5
% Initial leak	2.9	0.5

Table K.1: **Metrics for latch variants extended from the 3' vs 5' end of staples.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data.

Appendix L

AGAROSE PURIFICATION OF ORIGAMI

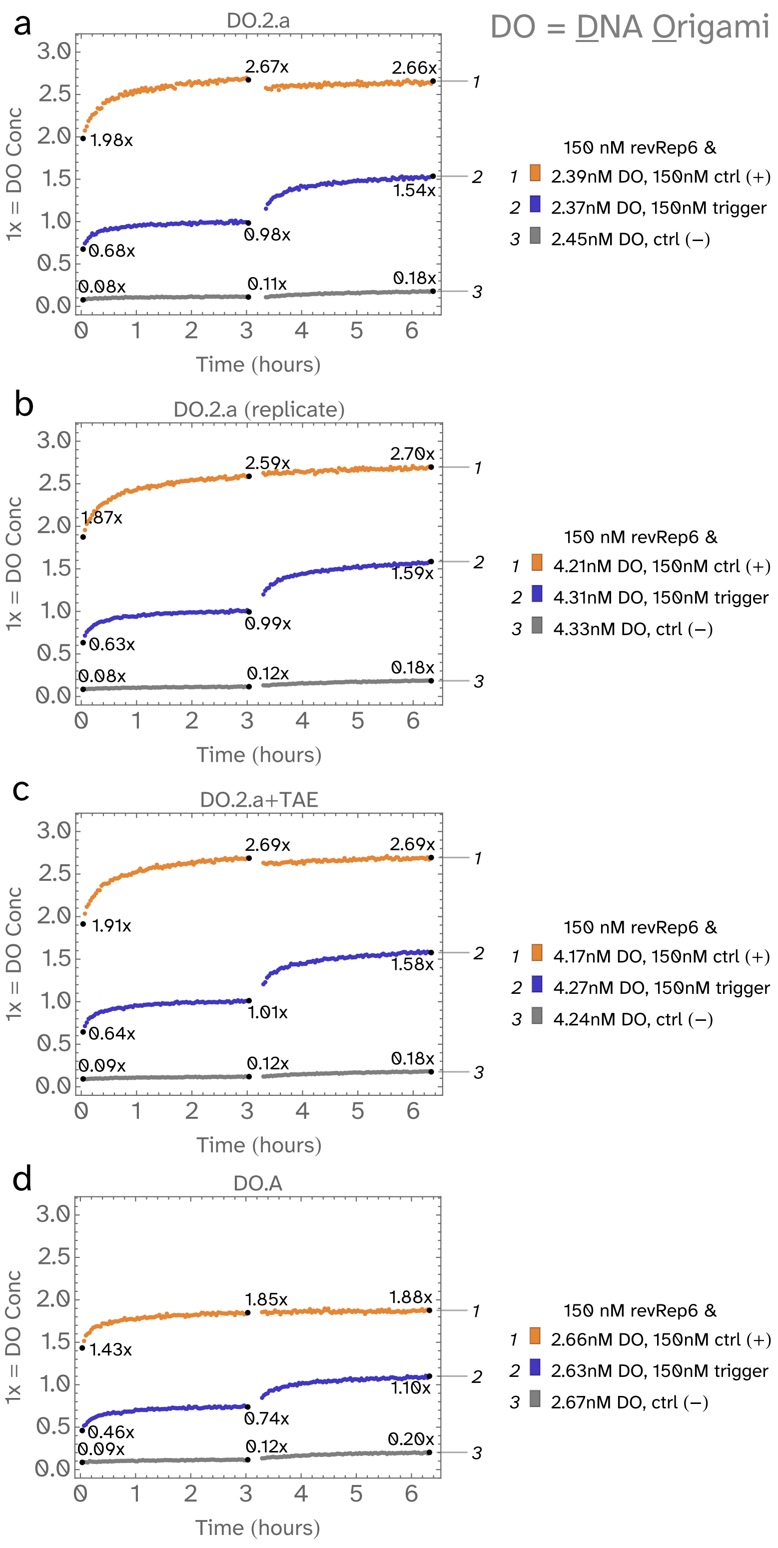
Agarose purification was essential to obtain valid results in past work on cargo-sorting DNA robots, which involved the same origami structure [125]. Thus, we investigated whether agarose purification was important for latch functionality. The experiment included a replicate of the original latch sample with nothing changed in the experimental procedure (DO.2.a), which illustrated that the latch performed almost exactly the same after storage at 4°C for one month. Alongside this replicate, two other samples were run simultaneously: a replicate of the original latch with an amount of TAE included in the buffer as would be expected had the origami sample been agarose-purified, and finally the agarose-purified variant (DO.A).

When comparing DO.A with DO.2.a or the more directly comparable DO.2.a+TAE, none of the metrics we typically use to evaluate latch performance were improved (table L.1); however, the completion level of the positive control curve in this case (Fig. L.1d) falls below 2× unlike that of other experiments (Fig. L.1a–c). This unique result could represent an improvement if it resulted from the reduction or elimination of malformed structures (i.e. the complex duplication discussed in Appendix N.2, Complex Duplication); however, given that the completion level for the experimental curve dropped by 30.4% between DO.2.a+TAE and DO.A, nearly equal to the 30.1% drop observed for the positive control curve completion level, it appears likely that this apparent improvement is a coincidence. Possibly, agarose purification affected the normalization process by eliminating some undesired structure which bears the latch but not the staple modified with ATTO532, such that the apparent origami concentration measured by fluorescence is higher for the agarose-purified variant. The validity of this hypothesis relies upon the potential discrepancy in fluorescence between the ATTO532-modified staple incorporated into the origami structure and the control where it is part of a duplex.

It is possible that agarose purification was more important for the double-layer origami tile when it was used in 2017 for cargo-sorting robots [125] because edge staples were not modified with 4(`C`) extensions, a method established in 2018 [155] for preventing undesired origami tile aggregation by inhibiting stacking bonds. Such extensions were added to the double-layer origami tile [131] before the inception of this project, however.

Metric	DO.2.a	DO.2.a (r)	DO.2.a+TAE	DO.A
% Signal Amplification	60.9	65	61.5	55.7
% Completion	55.2	56.3	56	55
% Fuel leak	2.7	2.7	2.2	4.9
% Initial leak	2.9	3.1	3.4	4.5
Batch age (days)	1	30	30	3
TAE	no	no	yes	yes
Agarose purification	no	no	no	yes

Table L.1: **Effect of agarose purification on metrics.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data. **DO.2.a,** Original unpurified latch. **DO.2.a (r),** Replicate experiment with same batch of DO.2.a, still unpurified. **DO.2.a+TAE,** DO.2.a, unpurified, but run with same amount of TAE in buffer as expected from agarose purification. **DO.A,** Agarose-purified version of DO.2.a.



DO = DNA Origami

Figure L.1: **Effect of agarose purification on kinetics.** Metrics given in Table L.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Original unpurified latch (DO.2.a). **b,** Replicate experiment with same batch of DO.2.a, still unpurified. **c,** DO.2.a, unpurified, but run with same amount of TAE in buffer as expected from agarose purification. **d,** Agarose-purified version of DO.2.a.

Appendix M

DETAILED LATCH DIAGRAMS

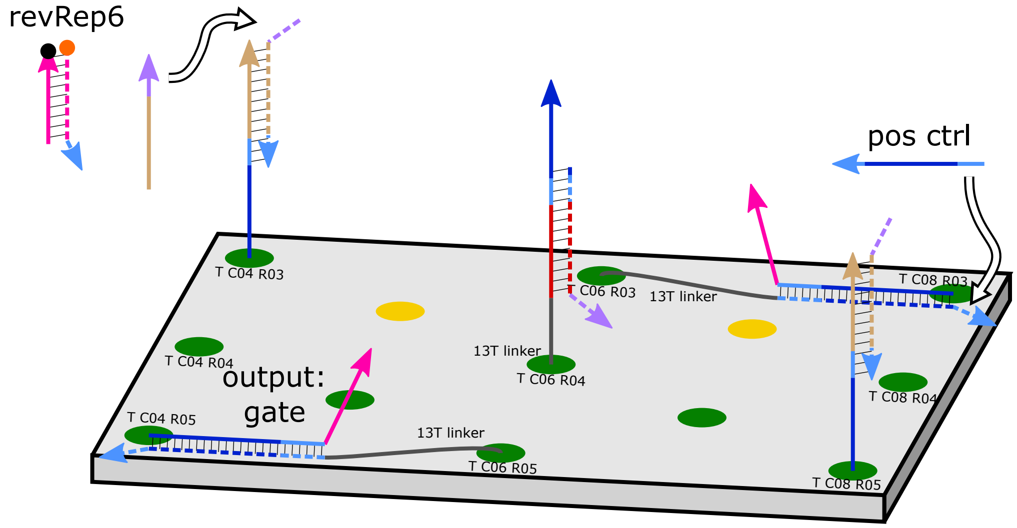


Figure M.1: **Detailed diagram of DO.2.a in state S0 with positive control.**

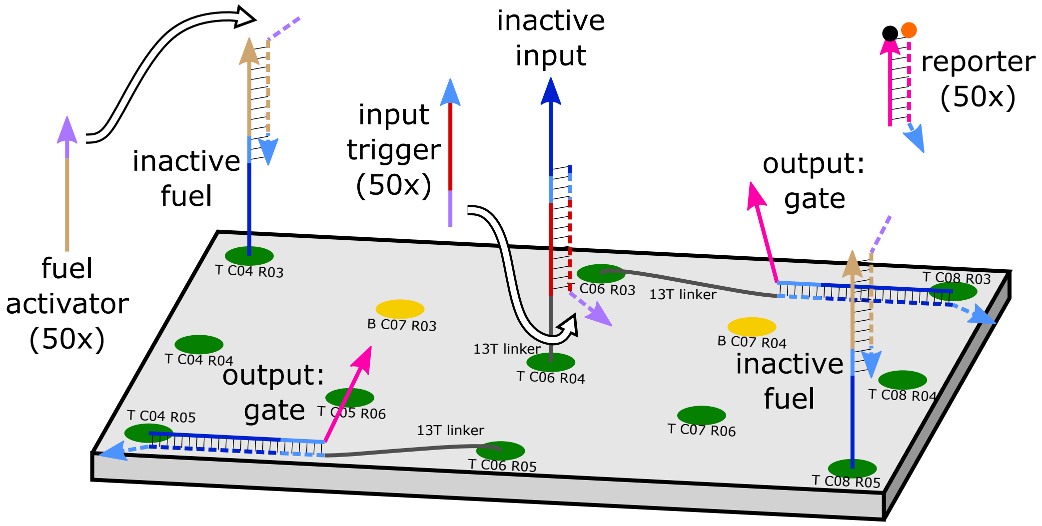


Figure M.2: **Detailed diagram of DO.2.a in state S0 with input trigger.**

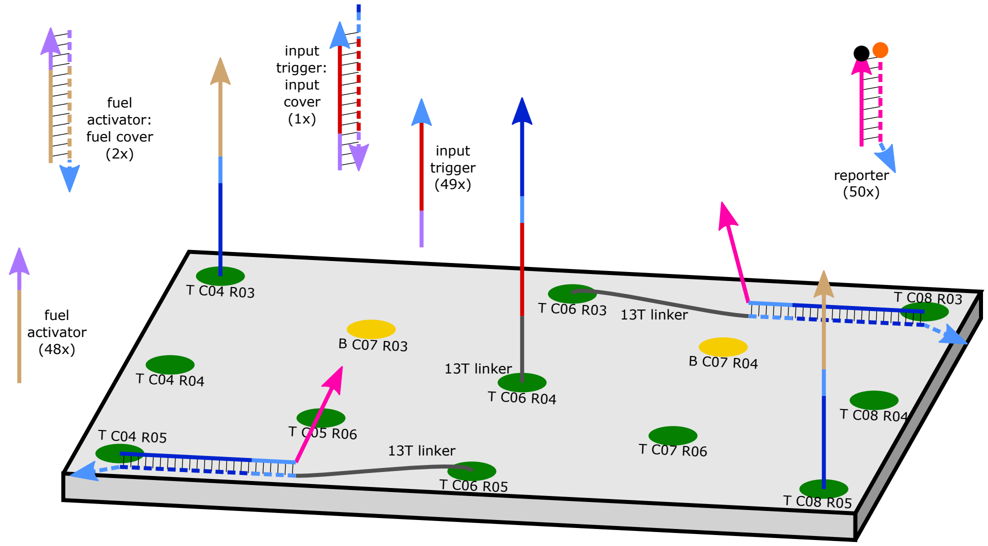


Figure M.3: **Detailed diagram of DO.2.a in state S1.**

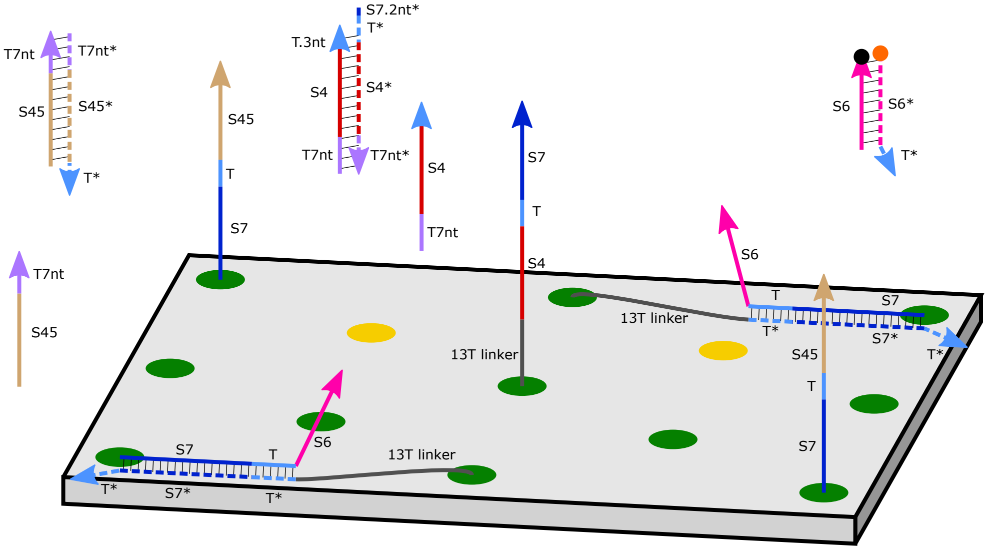


Figure M.4: **Detailed diagram of DO.2.a in state S1 with domain labels.** Other than T7nt, all domain names are consistent with Reference [23].

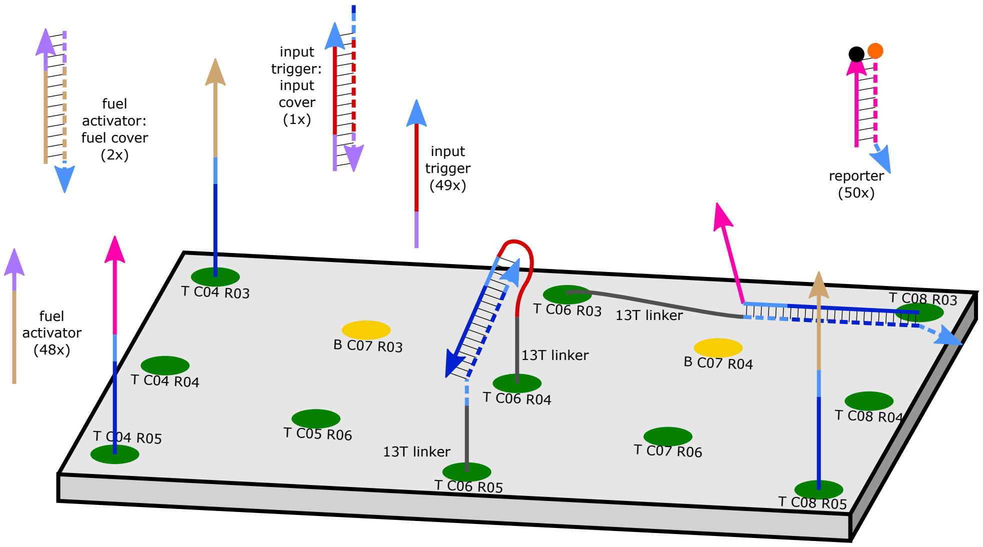


Figure M.5: **Detailed diagram of DO.2.a in state S2.**

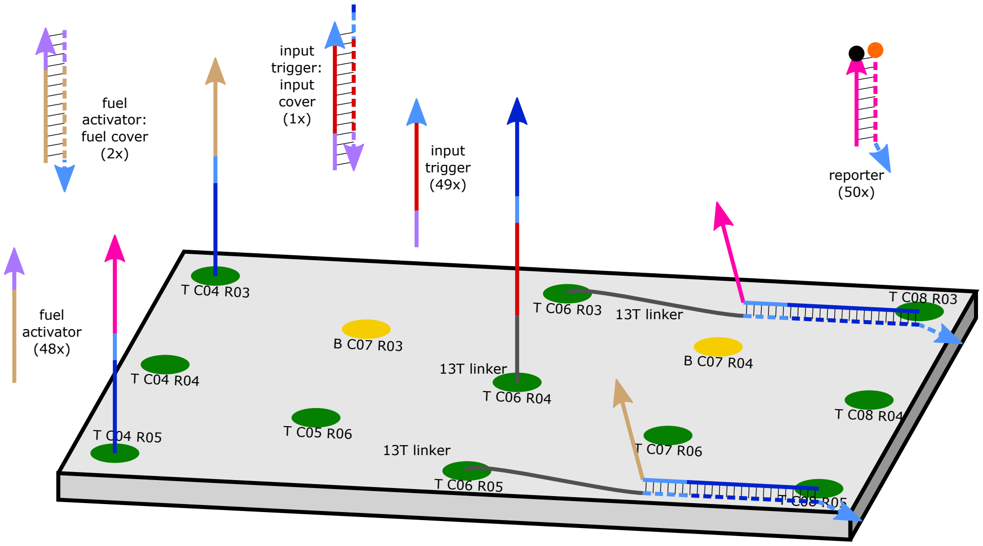


Figure M.6: **Detailed diagram of DO.2.a in state S3.**

Appendix N

MODIFIED POSITIVE CONTROL AND DUPLICATED COMPLEXES

N.1 Investigation of Positive Control Kinetics

None of our troubleshooting experiments discussed thus far could explain the slow kinetics of the green positive control curves. Typically, simple reaction systems such as a single strand displacement followed by reporting take place much more quickly than what we consistently observed. To investigate this mystery, we added a fourth reaction condition to our typical three: the presence of a "fast" positive control molecule with an extra toehold added to the opposite end from the original invading toehold. The extra toehold is 7 adenines (`AAAAAAA`) added for homology to the poly(`T`) tether on the gate strand. The fast positive control is depicted alongside the latch in Figure N.1.

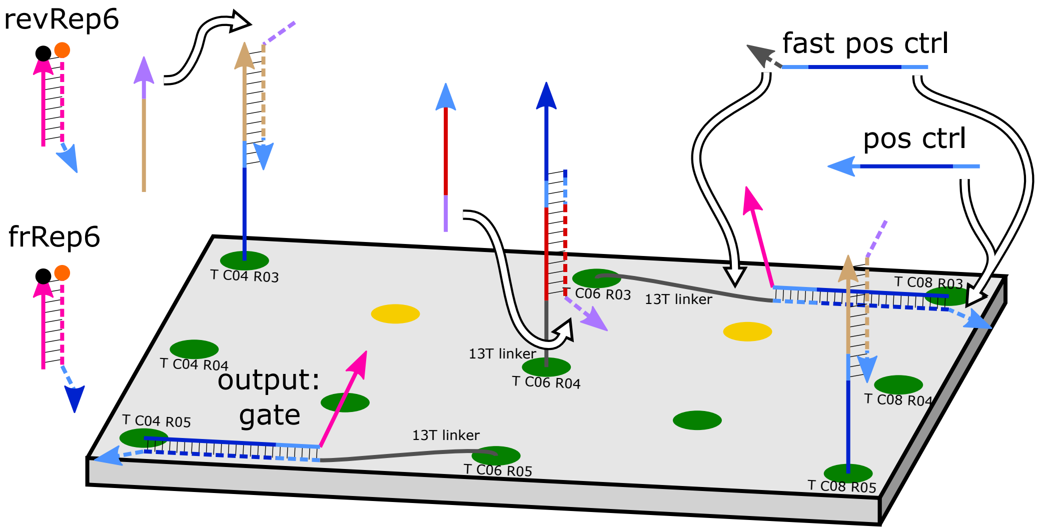


Figure N.1: **Detailed diagram of DO.2.a with fast reporter and fast positive control variants.** As in all other latch experiments, diffusible components are added at 50× origami concentration. Green and yellow positions on the origami surface correspond to color-coding diagrammed in Fig. I.2a and used in Fig. I.6.

One result of this change (Fig. N.2a) was as expected. The new positive control achieved faster kinetics than that observed with the original positive control.

Within the ≈10 minutes between reaction assembly and the start of fluorescence measurement, the new positive control achieved ≈86% completion, while the original positive control only achieved ≈71% completion.

These reactions were run alongside the experiments comparing the original revRep6 to a "fast" variant of revRep6 with a longer toehold called frRep6, as explained in Appendix G, Latch With Modified Reporter. Whereas frRep6 did not produce an obvious impact on reaction kinetics for the typical three reaction curves, it does appear to enhance kinetics for the case with the fast positive control. Compared to the ≈86% and ≈71% values discussed above, the relevant figures here are ≈93% and ≈72%. This increase in kinetics for the fast positive control curve in the presence of frRep6 suggests that reporting could be a rate-limiting step in the fast positive control case only.

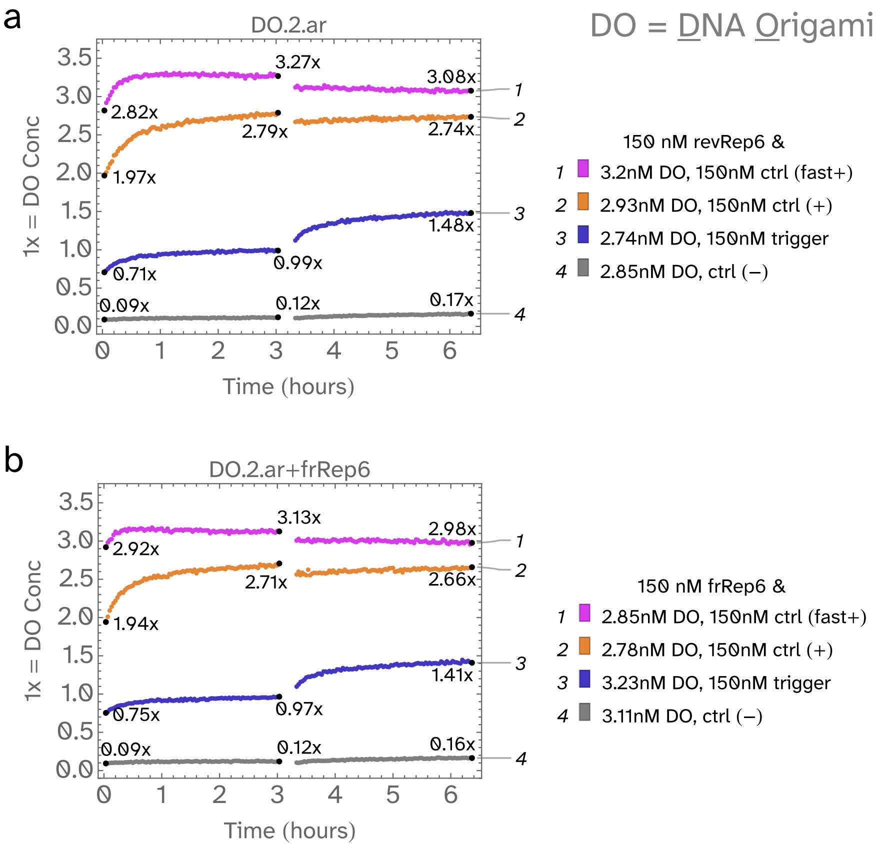


Figure N.2: **Effect of fast positive control.** No new metrics are calculated using the curve with the fast positive control, but original metrics given in Table G.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** DO.2.ar is a remade batch of DO.2.a, here reacted with original revRep6. **b,** DO.2.ar reacted with frRep6.

Other changes in behavior between the original fast positive controls were not as expected. The fast positive control achieved ≈12% higher completion levels than the original. Additionally, comparing the shape of the fast positive control to that of the original positive control illuminates a feature of the original positive control curves that was not obvious previously. The original positive control curves appear to exhibit biphasic kinetics, with a fast reaction dominating the first ≈30 minutes followed by a slow reaction taking place over the subsequent 2.5 hours. By contrast, the fast positive control curves have already reached their completion levels at t=30 min, followed be a subsequent gradual decline which could result from factors such as materials lost to the plate walls, evaporation, or photobleaching.

At first, given the complexity of the latch system, many hypotheses to explain the difference in positive control trigger behaviors may seem equally likely. The following control experiment (Figure N.3) narrowed down the list of possible explanations by demonstrating that the same phenomenon can be observed in a minimal system where only the output:gate complex is included. In this case, since there is no input, the input trigger is not expected to affect the results; indeed, the experimental results for the negative control case and the input trigger case are nearly identical.

Given the result shown in Fig. N.3, it is possible to eliminate any hypotheses reliant upon gate:output interactions with the input gate and the fuel, interactions between the gate:output complexes of the two pathways, or malformed structures that involve the input gate, the fuel, or combinations of the gate from one pathway with the output from the other pathway.

N.2 Complex Duplication

The simplest explanation for the difference in behavior between the two positive control cases is that some or all of the origami tiles bear malformed structures involving duplicated gate:output complexes, as in Figure N.4. The expected result would be at least three different possible reactions involving the positive control strand: (1) the intended reaction with the desired gate:output complex, (2) the unintended reaction with the complex formed from an untethered gate strand and an tethered output strand, and (3) the unintended reaction with the complex formed from a tethered gate strand and an untethered output strand. Each of these reactions might have different

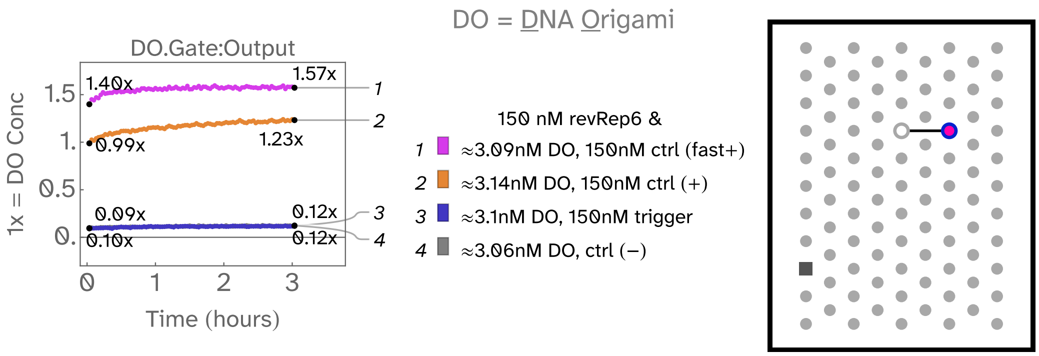


Figure N.3: **Effect of fast positive control in a minimal system.** When only one copy of the output:gate complex is included on the origami surface, the completion level for the case with the fast positive control trigger is higher than the completion level with the original positive control trigger, and the kinetics for the fast positive control remain faster than the kinetics for the case with the original positive control. Origami concentrations given in the legend are estimates, as they were not measured during the first three hours of this experiment; this experiment was run as part of a set of experiments explained in Appendix R, Quantifying Inter-Origami Leak.

rates, but we would expect the third to be much slower than the other two, since in only that case the free toehold in the complex would be expected to be partially sequestered by interactions with the unbound staple domain on the output strand. The degree to which the staple domain slows the positive control reaction would be determined by the sequence of the staple domain, meaning that this effect would be site-specific. This effect likely explains the location dependence observed in the positive control kinetics.

For the unintended reaction with the complex formed from a tethered gate strand and an untethered output strand, the fast positive control would be expected to eliminate the slowdown caused by the interaction between the staple domain and gate toehold; the fast positive control trigger can initiate strand displacement from the opposite end, where only very weak interactions would be expected between the output branch migration domain S6 and the poly(T) tether. This hypothesis also explains why the original positive control displays biphasic kinetics, while the fast positive control doesn't—the original positive control reacts quickly with gate:output complexes that formed correctly or where only the output is tethered until these types of gate:output complexes run out after about 30 minutes, while it simultaneously

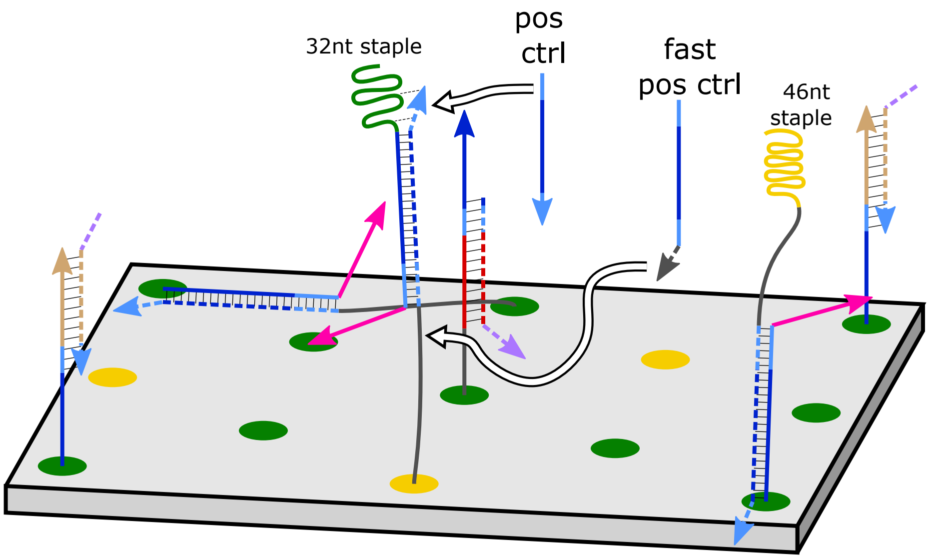


Figure N.4: **Detailed diagram of DO.2.b with duplicated complexes.** In this example, the output:gate complex in only one pathway is duplicated. We should expect a mixed population, where some latches will be formed correctly, while others will have a varying number of duplicated complexes. DO.2.b is chosen as an example for this diagram because it more clearly indicates how each duplicated complex would bear a dangling staple corresponding to a position other than where it is tethered. When the original positive control reacts with the duplicated complex with the dangling 32nt staple, it experiences slowed kinetics due to interactions between the staple and the toehold. In contrast, the fast positive control can either react by this slowed reaction pathway or an alternate faster pathway involving the poly(`T`) tether.

reacts slowly with the portion of gate:output complexes where only the gate is tethered over the first three hours. In contrast, the fast positive control quickly reacts with all three types of gate:output complex over a period somewhat shorter than 30 minutes.

To explain the higher completion level caused by the fast positive control requires some fraction of the gate toeholds to be fully inhibited by interactions with the unbound staple, which might depend on the degree of truncation in the unbound staple. The staple strands were unpurified in this experiment. Potentially, running an experiment combining the fast positive control and the PAGE-purified strands might allow the two positive control curves to reach a more similar completion level.

The hypothesis of duplicated complexes also helps explain the the leak in the negative control curve which takes place after the fuel activator is added (%

fuel leak). This leak can likely be attributed to the toeless reaction between the tethered fuel and the complex that forms between a tethered gate and an untethered output. With the output untethered, there is no longer a geometric constraint preventing this toeless reaction from taking place. This reaction might be expected to take place more quickly, as was observed for the toeless leak reaction between the pins-and-needles output gate and the tethered reporter (Fig. E.1a, second curve), but this reaction should be less entropically favorable than the reaction involving the tethered reporter. While the reaction involving the tethered reporter begins and ends with two dangling extensions projecting from the origami surface, this hypothesized fuel leak reaction involves the formation of a loop on the surface from two dangling extensions. Additionally, this leak reaction would be slowed by the same mechanism which might slow the positive control reaction: interaction between the unbound output staple and the gate toehold.

Appendix O

ATTEMPTED MEASUREMENT OF DUPLICATED COMPLEXES WITH STAPLE TRANSLATORS

Considering each step in the annealing sequence elucidates how duplicated complexes might form. During annealing, it might be expected that the longest continuous domains of dsDNA form first. Since the longest continuous domain in most staples is 16 bp, and the minimal continuous domain in circuit complexes is 20bp, we would expect that circuit complexes form first – especially given that the circuit domains use a three-letter code and thus have less internal secondary structure than the scaffold and the staples. The gate strand could conceivably bind to the output, fuel, or input strand, but it would be doing so in competition with the fuel and input inhibitor strands. Therefore, assuming annealing takes place slowly enough to reach equilibrium, the gate should almost entirely bind to the output, as predicted by NUPACK (see Fig. O.1).

The result is a gate:output complex. As the temperature decreases, it would be expected that either staple of the gate:output complex would be equally likely to bind the scaffold first. In other words, the staple extended with the output might bind the scaffold at the output staple site first, or the staple extended with the gate might bind the scaffold at the gate staple site first. For clarity, we will refer to whichever staple is bound first as staple 1, and the site on the scaffold where it is bound as site 1.

Imagine now that there is a forming origami structure where a gate:output complex is bound by staple 1 at site 1 on the scaffold. The gate:output complex has a dangling staple 2, and the scaffold has an open site 2. In solution, there are more copies of gate:output complex with unbound staple 2. So, there are now two possible reactions:

1. the unimolecular hybridization reaction between site 2 on the scaffold and staple 2 of the complex already localized at site 1
2. the bimolecular hybridization reaction between site 2 on the scaffold and staple 2 on a different copy of the complex in solution

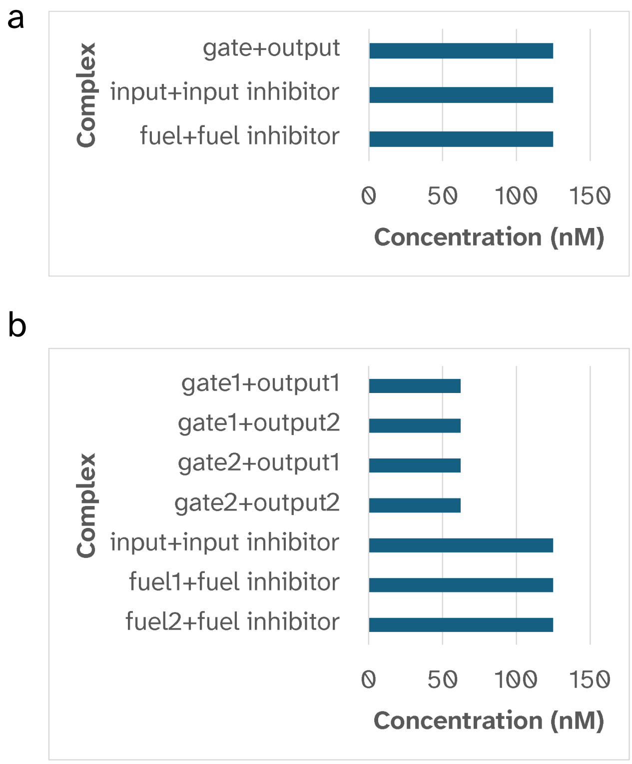


Figure O.1: **NUPACK analysis of latch annealing equilibrium.** These analyses serve to illustrate the expected annealing behavior for the latch components in the absence of the scaffold or other staples, and can help provide insights into annealing events which might take place before staples begin to form the origami structure. NUPACK settings: DNA, 25°C, 0.0125 M Mg^{2+}. Initial species concentrations approximate the true conditions during annealing. **a,** NUPACK results for a one-pathway latch. All concentrations initially set to 125 nM. **b,** NUPACK results for a two-pathway latch. Initial fuel inhibitor concentration is 250 nM, such that enough fuel inhibitor is present to bind both fuels. All other initial concentrations are set to 125 nM.

The first reaction is desirable, resulting in the intended latch design. The second reaction is undesirable, resulting in a duplicated complex (Fig. N.4). We originally expected that the first reaction would be fast enough to outcompete the second reaction, since the first is unimolecular. However, since many of our results can be explained well by the presence of duplicated complexes, this assumption was likely invalid. This conclusion is especially well-supported by our results for cases with only one pathway (Appendix I, Latch Relocated), and thus only one possible combination of output and gate.

In the case of experiments involving both pathways, there are two different types of gate and output, each corresponding to the top and bottom pathways.

The gate in the top pathway (for example) should be equally likely to bind the output from the top or bottom pathways, assuming the hybridization takes place before the origami has formed. If an origami structure were developed that formed at a higher temperature than circuit complexes, we would expect geometry to ensure the correct output hybridizes the correct gate (Fig. O.2a).

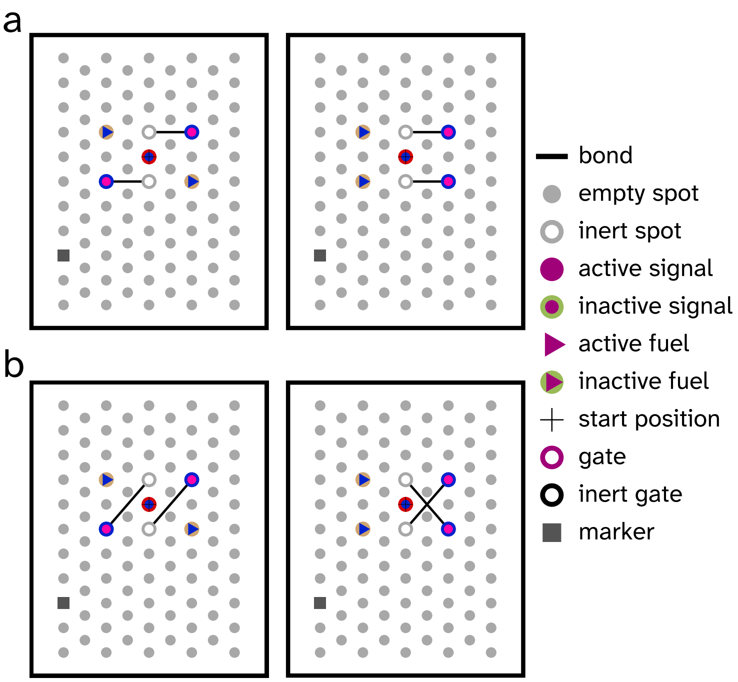


Figure O.2: **Valid vs. invalid gate:output combinations. a,** DO.2.a and DO.Cis, reproduced for comparison. The "bond" lines cover a distance of ≈10.4 nm as calculated in equation O.2. **b,** Invalid gate:output combinations for DO.2.a and DO.Cis. Assuming formation of the full double-layer origami structure, these combinations are expected not to form as diagrammed here due to geometric constraints. The "bond" lines cover a distance of ≈15.9 nm as calculated in equation O.3.

The maximum reachable length of the output:gate complex is 15.25 nanometers. The distance between any two adjacent points on the hex grid of valid staple extension sites on the top layer of the double layer origami structure is ≈6 nm. Therefore, the distance between a valid combination of output and gate tether points is ≈10.4 nm, such that a valid output:gate complex can form while tethered to both points. The distance between an invalid (Fig. O.2b) combination of output and gate tether points is ≈15.9 nm, such that an invalid output:gate complex would not be able to form while tethered to both points. Calculations to support these values are given in the following equations (O.1–O.3):

$$(.65nm/nt)(13nt) + (0.34nm/bp)(20bp) = 15.25nm \tag{O.1}$$

$$2(6nm * \frac{\sqrt{3}}{2}) = 10.4nm \tag{O.2}$$

$$\sqrt{(6nm * 2)^2 + (10.4nm)^2} = 15.9nm \tag{O.3}$$

See appendix B, DNA Lengths, for a discussion of appropriate values to use when calculating lengths of DNA. See appendices P.2, Raising Origami Formation Temperature, and Q, A Staple Layout for a Proposed 14-Helix Origami Structure that Forms at High Temperatures, for more discussion on the topic of designing origami that forms at a higher temperature.

In the absence of origami that fully forms at a temperature higher than gates of length 20bp, it should be expected that half of the gates are bound to the wrong output by the time staples join the growing origami structure. For the half of gates which are bound to the correct output, the expected reactions which may take place are the same as those considered above for the latch with only one pathway. Thus, complex duplication is already expected to occur without considering the issue of invalid output:gate combinations. Additionally, when an invalid combination of output:gate binds by staple 1 at site 1 on the scaffold, there are two additional possible reactions:

1. the unimolecular hybridization reaction between site 2 on the scaffold and staple 2 of the complex already localized at site 1
2. the bimolecular hybridization reaction between site 2 on the scaffold and staple 2 on a different copy of the complex in solution (complex duplication)

In the case of the first reaction for the invalid output:gate, a structure is formed which is incompatible with the geometry of the full double-layer origami structure, as discussed above. Incorporation of all other staples would require this reaction to be reversed, although the complete incorporation of all staples should not be expected to take place (Appendix A, Staple Incorporation Rates).

To summarize, our data for experiments involving either one pathway or two pathways agreed with the complex duplication hypothesis, and hybridization pathways leading to the formation of duplicated complexes during annealing can be conceptualized. Operating under the assumption that duplicated complexes form, we wished to quantify various classes of duplicated complex by taking advantage of the only feature distinguishing them from complexes that formed correctly: dangling staples. See Fig. N.4 for examples of 32nt and 46nt dangling staples.

Since dangling staples are the only part of a duplicated complex that is not present in the desired structure, we needed to develop translators that could detect them to discern duplicated complexes from intended complexes. We aimed to produce translator gates enabling the duplicated complexes to serve as inputs to a strand displacement reaction. This plan requires a single-stranded region of the staple to function as a toehold to initiate the reaction for detecting a dangling staple. Strand displacement can be inhibited by sequences with high secondary structure, which is one reason for the common use of three-letter code domains in past work from our group [41, 125, 137]. In this case, however, all staples use a four-letter code with relatively high secondary structure, although the secondary structure varies with the individual staple sequence. While the gate:output complexes were the most likely to be duplicated on the surface, we ultimately envisioned using translators to quantify all 7 possible dangling staples. Therefore, it was important that no staple had a particularly high secondary structure.

O.1 Selecting a Location with Minimal Secondary Structure

To select a location for a Latch with staples lacking problematic secondary structure, we produced a heat map of the top layer using NUPACK's [156] predicted complex free energy of each staple as a proxy for secondary structure. Green indicates staples with less secondary structure, and red indicates staples with more secondary structure. In the second diagram, all staples with a complex free energy worse than -4.445 kcal/mol are colored black, indicating that these should be avoided.

A new position for a Latch was selected based on this constraint, depicted in the third diagram. The origami bearing a Latch at this location was previously described in Appendix I, Latch Relocated, where it was named DO.2.c. A relevant possible malformed variant of DO.2.c with duplicated complexes is depicted in Figure O.4.

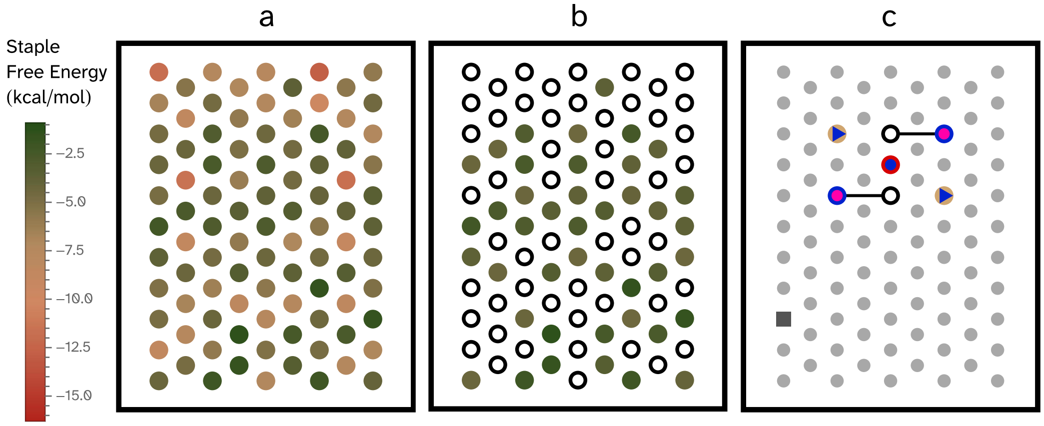


Figure O.3: **Process to relocate the latch to a position avoiding secondary structure. a,** Heat map of complex free energy in top layer. **b,** The same heat map, but all positions with a complex free energy worse than -4.445 kcal/mol are colored black. **c,** Position for DO.2.c.

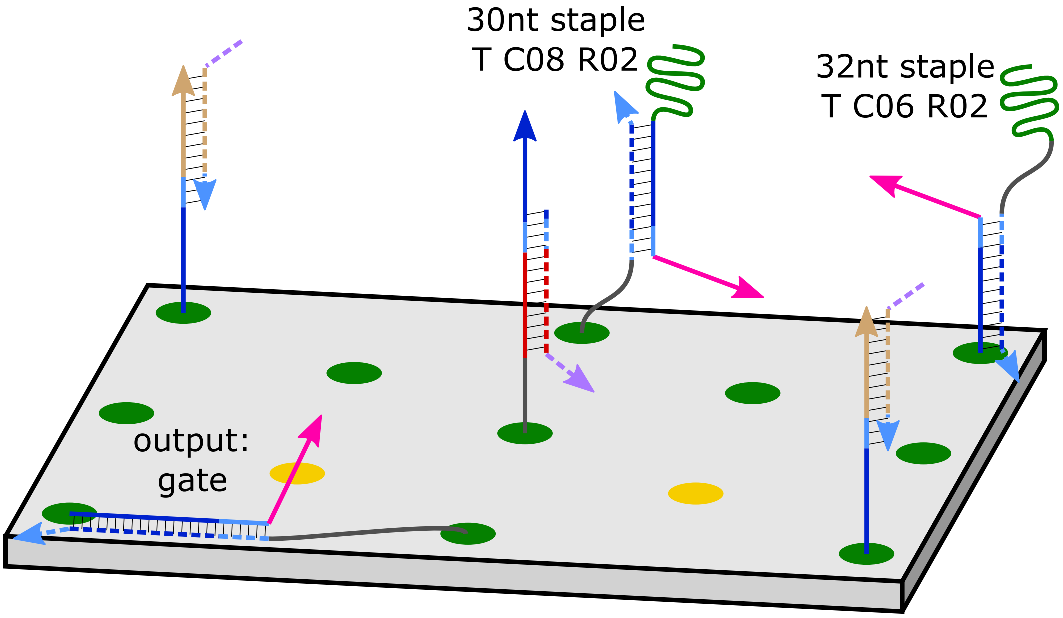


Figure O.4: **DO.2.c with a duplicated complex.**

O.2 Staple Translator Experiment

We began by designing a translator (Fig. O.5a) to detect just one of the potential dangling staples of DO.2.c. For convenience, we aimed to use the translator to activate an existing reporter. Since the output strand is designed to react only with revRep6, we designed the reporter to trigger Rep6 (the reporter in the original orientation). While the use of this reporter was a suboptimal design choice given the increased potential for toeless strand displacement, success in this initial experiment would have justified ordering another reporter alongside a full set of translators. Unfortunately, a control complex bearing the dangling staple failed to displace the top strand of the translator at a useful rate.

The control complex (Fig. O.5b) was designed to be as similar as possible to the hypothetical malformed structure where the gate:output complex is only tethered to the origami surface by the gate strand. Despite all staples at this position having been selected for their relatively low secondary structure, NUPACK [156] analysis of the control complex (Fig. O.5c) shows a non-negligible amount of secondary structure in the staple domain, forming a hairpin with a weak stem. It was not immediately obvious without experimental evidence whether this degree of secondary structure would be excessively problematic.

We ran this control experiment (Fig. O.5d) in a manner consistent with typical fluorescence kinetics reactions involving origami. 1× was 3nM. When present, the dangling staple control molecule was added at 1×. The translator and Rep6 were present in all cases at 50×. The translator was PAGE-purified. Three conditions were tested:

1. negative control condition 1 with only reporter and translator
2. negative control condition 2 with reporter and an output:gate control molecule with a staple (T C04 R04) that does not match the translator
3. the experimental condition including the reporter, translator, and the output:gate control that matches the translator

Curves are normalized to an external positive control: the reporter at 50× and the top strand of the translator at 1×).

Under typical Seesaw gate conditions, the negative controls would be expected to leak slowly by toeless strand displacement, which is observed. Negative control condition 2 would be expected to leak more quickly since it has an additional source of toeless displacement, and this is also observed. However, the experimental case would be expected to quickly reach a completion level of 1×, but instead it only achieves a rapid initial jump of roughly .1× above background, followed by a curve with kinetics similar to the negative

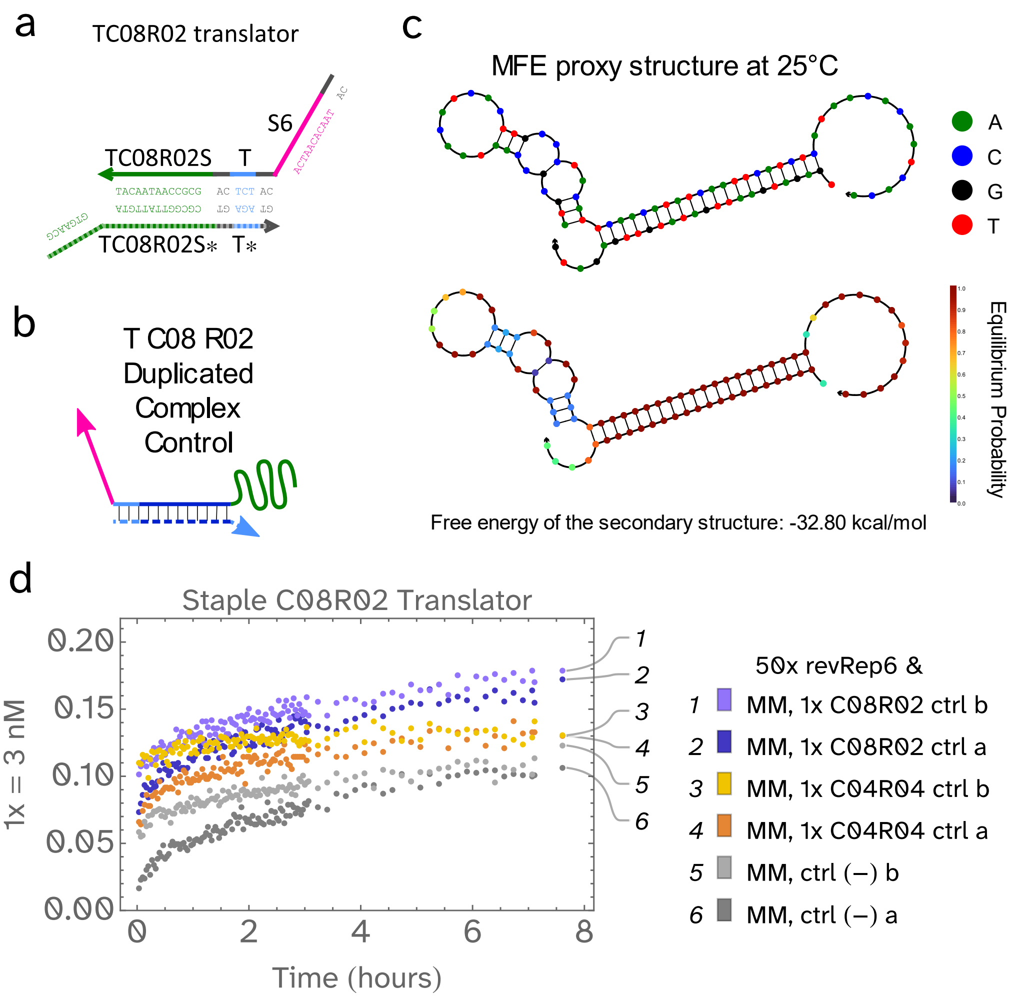


Figure O.5: **Experimental design and results for detection of a dangling staple on a well-mixed control complex. a,** Sequence-level diagram for the translator designed to detect a dangling staple. **b,** Domain-level diagram for the well-mixed control complex bearing a dangling staple. **c,** NUPACK [156] analysis of the control complex shows that some secondary structure in the dangling staple is expected, despite our attempt to avoid secondary structure. **d,** Results. The design was not effective.

controls. This outcome suggests that the secondary structure of the staple prevents it from strand displacing at a rate above a toeless reaction. Most likely, the .1× jump corresponds to variants where the staple domain in the control is truncated (see Appendix H, Latch With Purified Components) just enough to reduce the secondary structure without truncating away the portion of the staple that can react with the translator. Regardless of the reason, we determined that it would not be straightforward to reliably detect dangling staples by this method. We considered unfolding the secondary structure of dangling staples by adding helper strands, as has been achieved in other work [157]; however, designing and verifying helper strands for every staple would require a significant time investment. We chose to prioritize other experiments. Future experiments aiming to detect dangling staples would greatly benefit from the use of a three-letter-code scaffold, as discussed in the conclusion of this thesis.

Appendix P

ATTEMPTS TO ADDRESS DUPLICATED COMPLEXES

P.1 Annealing with Low Concentration Components

A natural first approach to mitigating the possibility of duplicated complexes is to add all circuit components at a lower concentration relative to the scaffold during annealing. We explored this strategy, adding components at 2× the scaffold concentration rather than our typical 5-6×, but the results were not substantially different from previous experiments (Fig. P.1, Table P.1). Arguably, the best performance was achieved in the case where all circuit components were added at 2× the scaffold concentration, where the % completion value increased from 52.9 to 57.2 and the completion level of the positive control curve fell closer to the expected value of 2×.

Metric	DO.2.b	DO.LC.a	DO.LC.b	DO.LC.c
% Signal Amplification	61.5	54.4	50.5	60.4
% Completion	52.9	47.1	58.9	57.2
% Fuel leak	4.8	4.1	8.1	5.1
% Initial leak	2.3	2.9	8.2	3.9
Input:scaffold ratio	5	2	5	2
Input inhibitor:scaffold ratio	5	2	5	2
Other Latch parts:scaffold ratio	5	5	2	2

Table P.1: **Metrics for latch variants exploring addition of circuit components at 2× scaffold concentration rather than 5×.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data. **DO.2.b,** All components added at 5× scaffold concentration. **DO.LC.a,** Input and input inhibitor added at 2×, and all else added at 5×. **DO.LC.b,** All circuit components except input and input inhibitor added at 2×. **DO.LC.c,** All circuit components added at 2×.

P.2 Raising Origami Formation Temperature

Motivation and Background

If the temperature of origami formation could be increased, it would be expected to reduce the formation of duplicated complexes, especially in the two-pathway case. The best way to understand why this would take place is to imagine the sequence of events during annealing, starting from the assump-

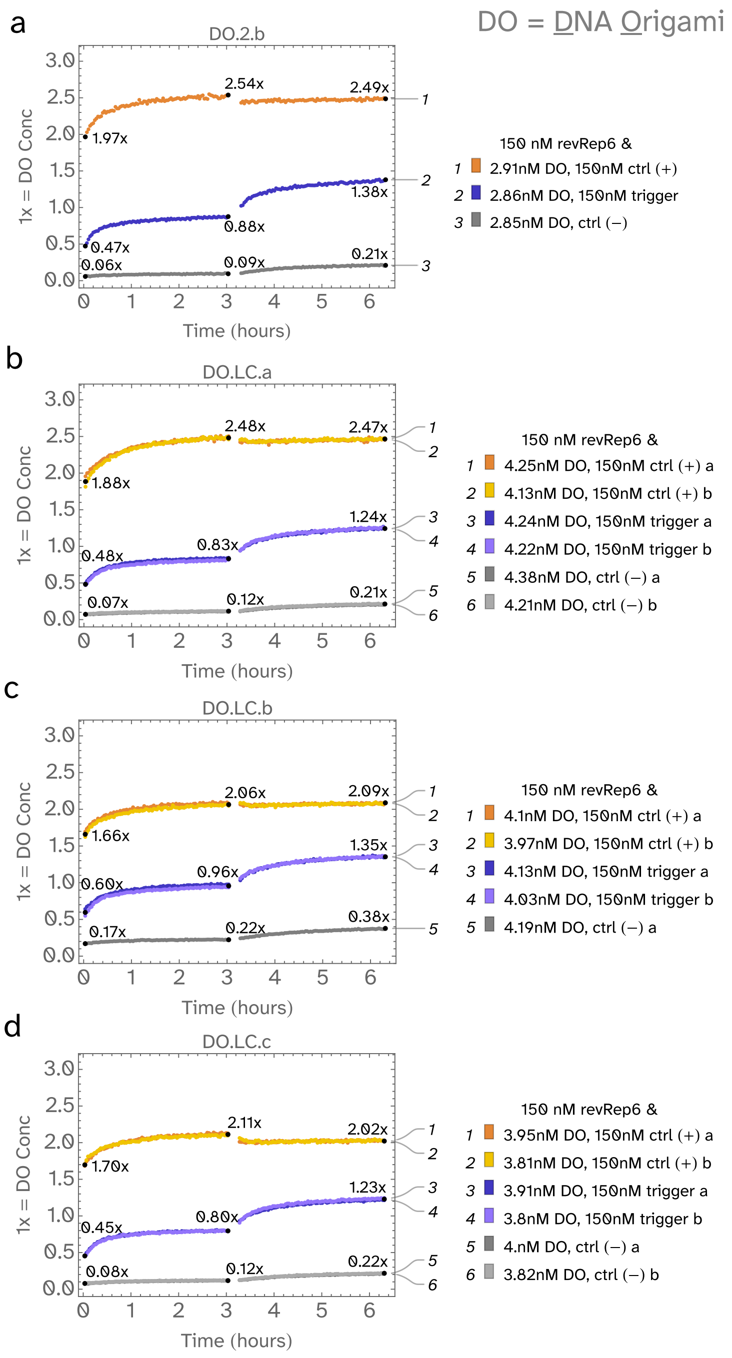


Figure P.1: **Effects of addition of circuit components at 2× scaffold concentration rather than 5×.** Metrics given in Table P.1. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** All components added at 5× scaffold concentration. **b,** Input and input inhibitor added at 2×, and all else added at 5×. **c,** All circuit components except input and input inhibitor added at 2×. **d,** All circuit components added at 2×.

tion that the origami formation temperature has been raised considerably such that the overwhelming majority of cases, the origami structure will fully form before the latch complexes hybridize. In this case, we can envision an intermediate state in which the origami structure is present with each staple localized, but all extensions are unbound. As the temperature slowly decreases, there will be a state where the gate strand can reversibly bind to the input, fuel, and output strands. The gate in the top pathway will not be able to reach the output or fuel from the bottom pathway. This geometric constraint fundamentally reduces the number of possible duplicated complexes which could form for the two-pathway case. During this state of reversible binding, the fuel inhibitor and the input inhibitor will compete with the gate strand to bind the fuel and the input. Since no such inhibitor exists for the output strand, the gate strand will preferentially bind the output strand. If annealing takes place slowly enough to allow the system to reach equilibrium, then in the ideal case, the gate strand will only bind the output strand.

Motivated to find a method for raising the folding temperature for the double-layer origami structure, we examined the process of origami folding. For a continuous domain of dsDNA, the free energy of hybridization is primarily a function on the length of the domain and its GC content. When an origami structure folds, the most likely first step is a hybridization event between the scaffold and a staple segment with a relatively high free energy of hybridization. However, it has been shown that following this step, any subsequent steps are challenging to predict [144]. The free energy of hybridization of staple strands is not strongly correlated with how early they bind the scaffold. Furthermore, it is not possible to make any predictions based on the 3D structure of the final origami, since a recognizable compact structure will not form until the majority of staples have fully bound the scaffold. Many different factors beyond influence the course of folding events, including the distance between domains along the scaffold which must be brought together (more entropy is lost when longer distances are bridged), any off-target interactions which might take place, and the availability of stacking bonds provided by neighboring staples. Often, after one staple joins that bridges a large distance, a cluster of related staples will rapidly anneal to the scaffold because they no longer have a large entropic penalty to overcome [144]. The use of staples that bridge large distances has been shown to influence folding pathways, which is a property that can be used to design a large set of staples where the addition or removal of a much smaller set of staples will cause dramatically different origami structures to form [158]. Additionally, it has been shown that an origami structure took much longer to fold (360 minutes rather than 45 minutes) after deleting the 20 fastest-incorporating terminal segments of staples, while deleting the 20 slowest-incorporating terminal segments of staples caused a less substantial delay, reaching an equivalent point in the folding process after 240 minutes rather than 45 or 360 (see Fig. 3F in [144]).

Approach: Seed Staples and Autobreak

Given these considerations, we hypothesized that staples with extended continuous domains with unusually high free energies of hybridization could act like seeds that nucleate crystal growth. In other words, we hoped that we could design “seed staples” which would enable all staples in an origami structure to join sooner during annealing than they otherwise would have. It would have been necessary to measure the sequence of staple folding events as was done previously[144] to confirm whether this hypothesis was valid; however, since this hypothesis was not the main focus of our efforts, we only measured whether including our design for seed staples would improve the function of the latch. I encourage a study further investigating this hypothesis; whatever the results, such a study would clarify the mechanism of origami folding and inform improved DNA origami designs.

Software for optimizing DNA origami designs to fold readily has recently been made available [159]. Called Autobreak, this software includes a scoring function, which at first glance appears to predict the temperature at which each staple incorporates into an origami structure.

However, since Autobreak only aims to find an optimal staple route, rather than accurately predict the sequence of events during origami folding, it does not model folding kinetics (including each individual hybridization, dissociation, and displacement event) or staple cooperativity. Instead, it treats each staple as if it were the first to join the structure, furthermore treating each staple as if it bound the scaffold at all sites all at once rather than one continuous domain at a time, and estimates a temperature at which the staple would do so. Modeling each staple as if it joins all at once is a major simplification of reality; it has been shown that the incorporation times of staple 5’ and 3’ termini are independent [144]. Therefore, while this

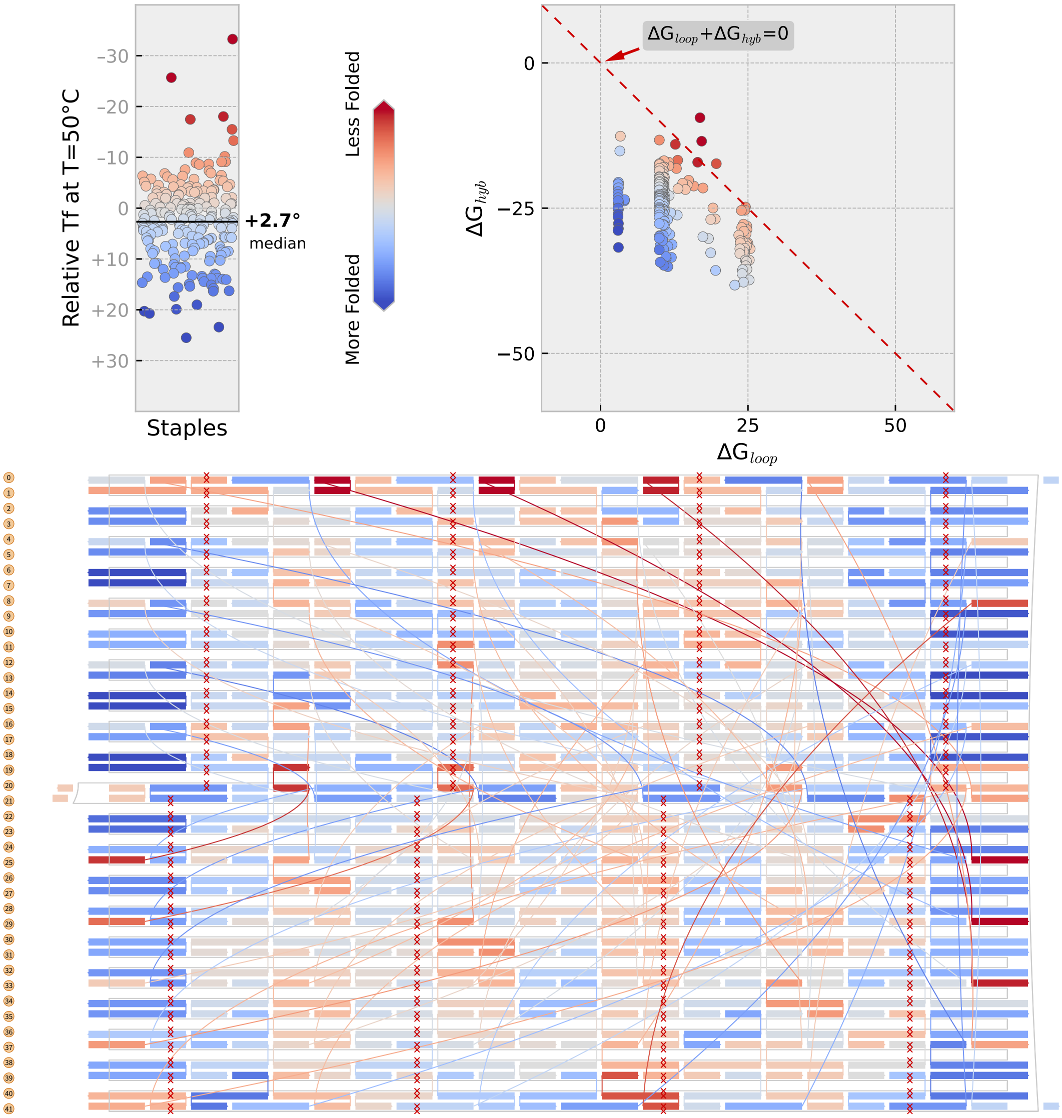


Figure P.2: **Autobreak results for the double layer origami without seed staples.** These plots were generated using the Autobreak Colab notebook [159].

software's optimization function could be invaluable to researchers designing new origami structures, and furthermore will likely produce origami structures which fold at higher temperatures than usual, it is not sufficient for accurately predicting the temperature at which a given origami structure will be fully folded. A model which accurately predicts folding temperature of each staple, or more accurately each contiguous staple domain, could use Autobreak for inspiration.

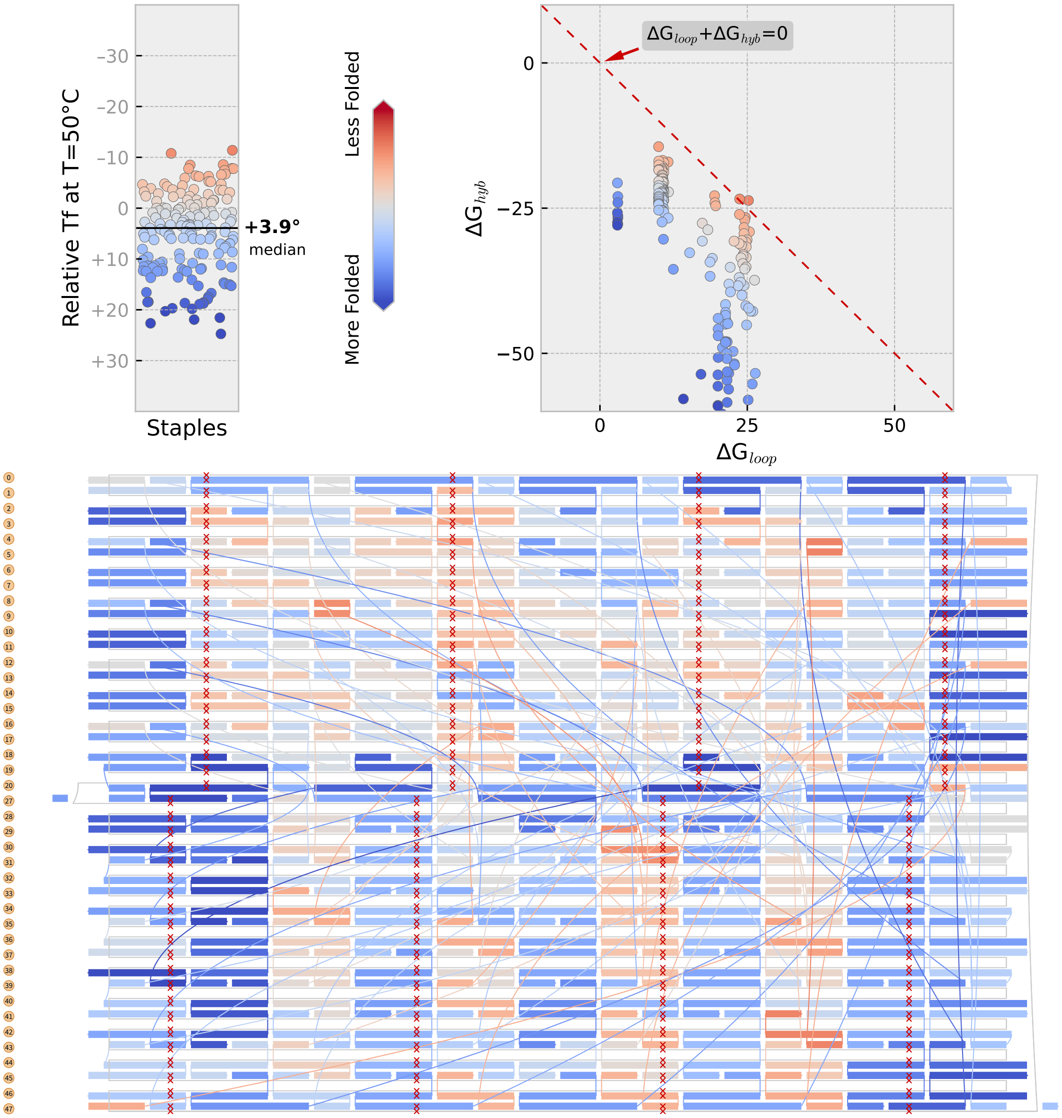


Figure P.3: **Autobreak results for the double layer origami with seed staples.** Seed staples are present at all edges, near the latch site in the top layer, and in alternating columns of the bottom layer. These plots were generated using the Autobreak Colab notebook [159].

While Autobreak could not predict whether seed staples would change the temperature at which other staples fold, it could estimate whether our designs for seed staples would fold at a higher temperature than they otherwise would have. We designed various types of seed staples for double-layer origami by joining existing staples and analyzed them with Autobreak. All of the seed staples were predicted to fold more readily than the original staples they were built from. Some were predicted to fold at a similar temperature to the best, most readily-folding staples in the original design (compare Figure P.2 and Figure P.3).

Metric	DO.2.b	DO.SS.a	DO.SS.b	DO.SS.c	DO.SS.d
% Signal Amplification	61.5	55	45.2	44.4	46.6
% Completion	52.9	47.6	44.3	44.4	45.1
% Fuel leak	4.8	5	3.9	2.9	3.6
% Initial leak	2.3	3.1	2.7	3.6	3.1
Internal seed staples	no	yes	no	no	yes
First+last row edge seeds	no	no	yes	yes	yes
Other edge seed staples	no	no	no	yes	yes

Table P.2: **Effect of seed staples on two-pathway latch metrics.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data. **SS.a,** Internal seed staples only. **SS.b,** First and last row edge seed staples only. **SS.c,** All available edge seed staples. **SS.d,** All available seed staples, both edge and internal.

Since no model existed to truly predict the folding temperature of origami, we could not predict whether these particular seed staples would be sufficient to raise the overall origami folding temperature high enough to ensure that the circuit components hybridize after origami folding. To indirectly answer this question, we tested whether the latch function was improved by the presence of our seed staples, testing them with the latch located at the position of DO.2.b (see Fig. I.1). Unfortunately, no improvement was observed, with all latch variants performing moderately worse than what had been previously achieved (Fig. P.4, Table P.2). It is difficult to explain why the latches performed worse, and easier to attribute this result to experimental noise. If the seed staples succeeded in raising the temperature at which the origami folds, and the negative effect on latch function is real, then this set of experiments indicates an unresolved mystery about the latch formation process.

P.3 Insertion of Latch Complexes

To prevent mismatched gate:output complexes from forming, we annealed the gate:output complexes separately from the rest of the origami and inserted them subsequently at a low temperature. Aiming to maximize the odds of

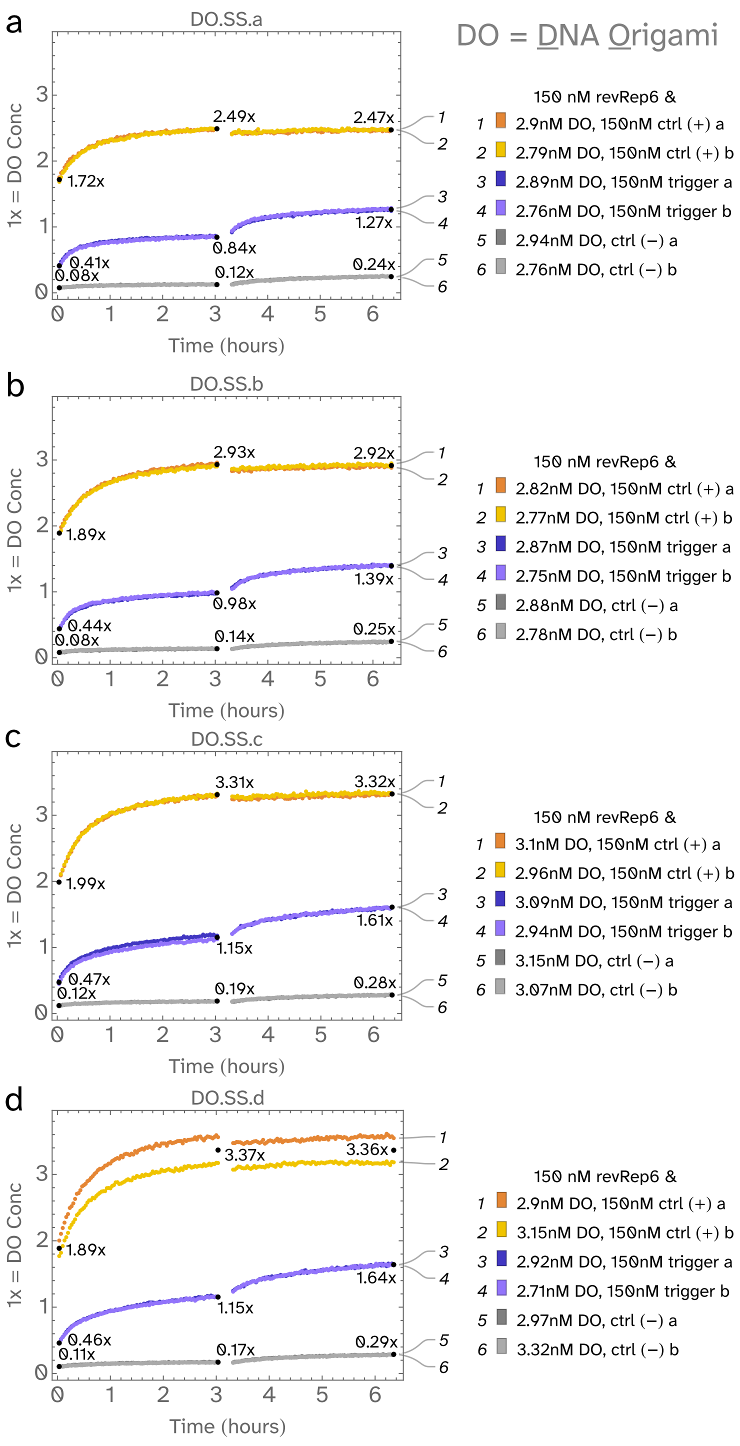


Figure P.4: **Effect of seed staples on two-pathway latch kinetics.** Metrics given in Table P.2. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Internal seed staples only. **b,** First and last row edge seed staples only. **c,** All available edge seed staples. **d,** All available seed staples, both edge and internal.

successful insertion, we used the position of DO.2.c, the location of the latch expected to have the least secondary structure in its staples (see Appendix O.1, Selecting a Location with Minimal Secondary Structure, for an explanation).

This approach was not expected to eliminate the possibility of duplicated complexes entirely, but unlike the annealing method, it would ensure that the gate strand from the top path would bind the output from the top path, and that the gate from the bottom path would bind the output from the bottom path. Unfortunately, our 1-path results indicated that only the bottom gate:output complex could be fully inserted, so we did not attempt a 2-path version of this experiment (which is the only condition where we would expect materially different results). Results with the top pathway appeared to show that only partial insertion took place, possibly only at one staple site.

Metric	DO.I.a	DO.I.b	DO.I.c	DO.I.d
% Completion	81.6	72.3	67.1	57.8
% Fuel leak	7.2	1.9	1.6	1.8
% Initial leak	6.1	1.3	8.7	7.5
Gate:output insertion	yes	yes	no	no
Inactive input, fuel insertion	no	no	yes	yes
Input row,column	4,5	4,5	4,5	4,5
Angle	0°	180°	0°	180°
Handedness	L	L	L	L

Table P.3: **Effect of insertion on one-pathway latch metrics.** The equations used to produce these metrics are given in Appendix F.2, Quantitative Analysis of Latch Experimental Data.

The secondary structure of each staple, and each part of the scaffold, varies dramatically (see Fig. O.3a)—therefore it is unsurprising that insertion is more challenging at some locations than at others; however, the secondary structures of the staples of the gate:output complex in the top pathway are not predicted to be worse than the secondary structures of the staples of the gate:output complex in the bottom pathway (see Fig. O.3b), so this explanation is not satisfying. A more likely explanation for this discrepancy in insertion performance is found when performing NUPACK [156] analysis on each pair of gate and output staples without including the gate and output circuit domains. In this case, an interaction between the gate and output staples in the top pathway is predicted (Fig. P.6a), whereas no such interaction is predicted for the staples of the bottom pathway (Fig. P.6b).

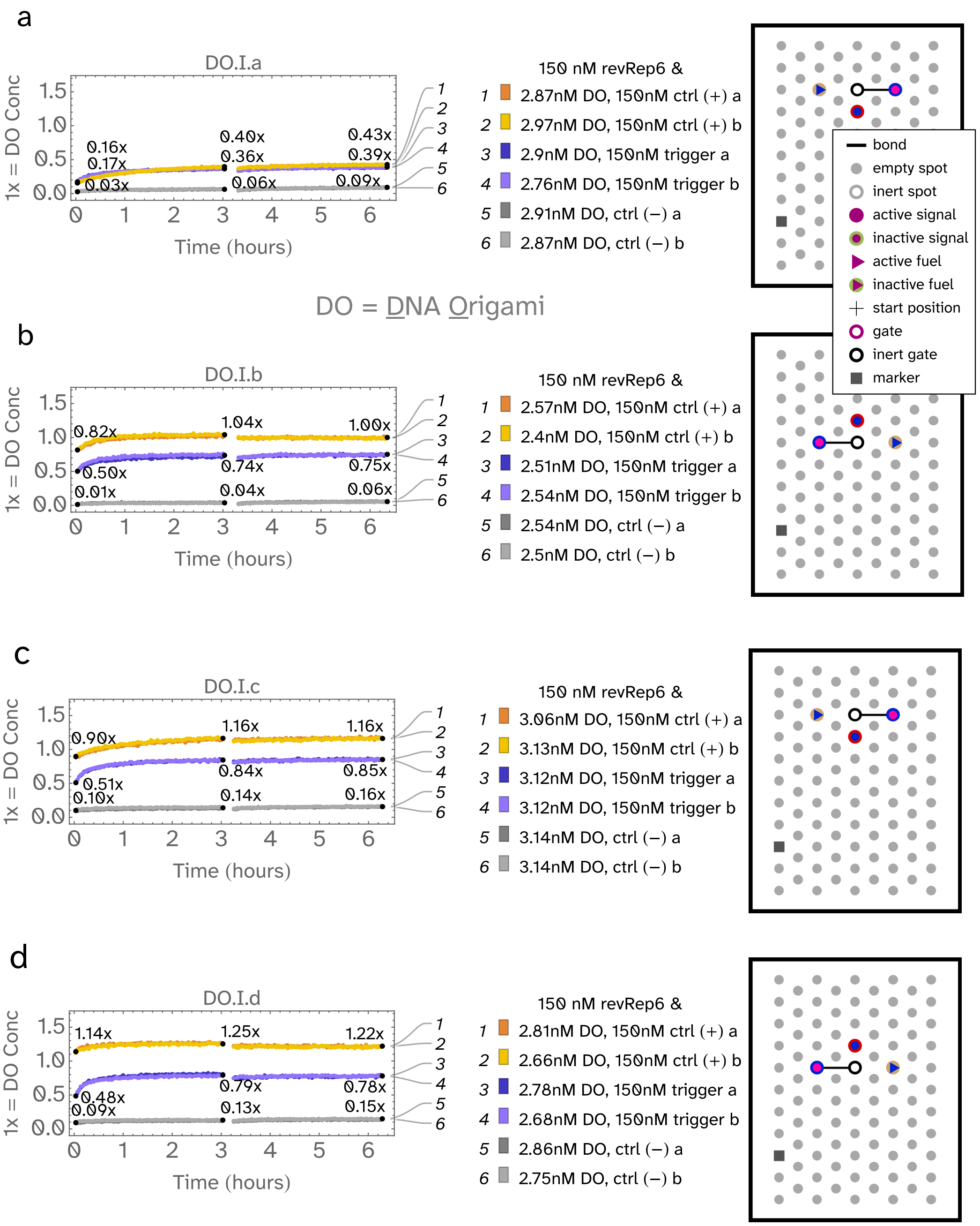


Figure P.5: **Effect of isothermal insertion on one-pathway latch kinetics.** Metrics given in Table P.3. Trigger and positive control strands added at t=0h. Fuel activator added to all reactions at t=3h. **a,** Insertion of gate:output, top pathway. **b,** Insertion of gate:output, bottom pathway. **c,** Insertion of inactive input and fuel, top pathway. **d,** Insertion of inactive input and fuel, bottom pathway.

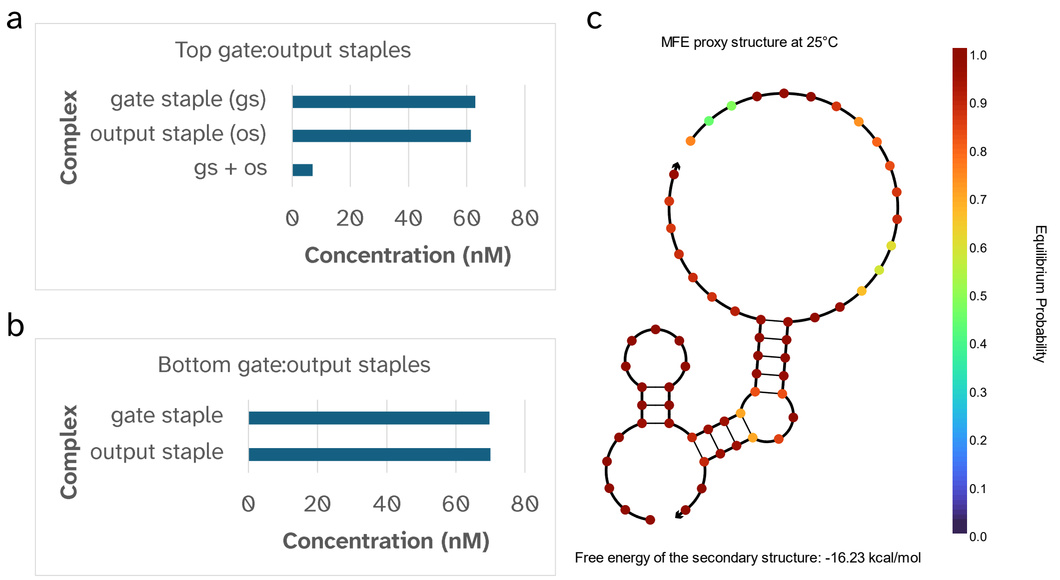


Figure P.6: **NUPACK analysis of staple-staple interactions for the inserted gate:output complexes.** NUPACK [156] settings: DNA, 25°C, 0.0125 M Mg^{2+}. Initial species concentrations approximate the true conditions during insertion. **a,** Equilibrium prediction for the gate and output staples in the top pathway. **b,** Equilibrium prediction for the gate and output staples in the bottom pathway. No meaningful interaction is predicted. **c,** Minimum free energy (MFE) complex predicted to form between the gate and output staples in the top pathway only. This structure corresponds to the entry gs + os in **a**.

Due to the 13nt linker domain of the gate, interactions such as the one shown in Fig. P.6c between the staple domains can take place without needing to consider geometry; however, even if the linker domain were omitted, the interaction in Fig. P.6c could still take place for this particular structure. The length between the 3' ends of the interacting staples (from which the circuit components are extended) is greater than the length of the gate:output complex ($9 * .65 + 9 * .34 = 8.91 > 6.8 = 20 * .34$). Therefore, the reachability requirements are satisfied.

While the interaction predicted in Fig. P.6a is not strong enough to allow the resulting complex to dominate NUPACK's predicted equilibrium with each staple initially present at 70nM, the configuration with the two staples interacting might dominate the real unimolecular equilibrium. Since NUPACK is not able to consider pseudoknots [156], we missed this staple-staple interaction during our original NUPACK analysis of the full gate:output complexes including the circuit domains. Therefore, I believe it is insufficient to consider only the individual staple secondary structures when determining whether it is possible to insert a complex involving two staple extensions. I believe it is appropriate to also consider the potential interactions between two staples.

Moving beyond the issues in DO.I.a, results from the bottom pathway (DO.I.b), as well as results from inserting the inactivated input and inactivated fuel instead of the output:gate complexes (DO.I.c–d), all appeared fairly similar to the results when annealing (compare to Fig. I.5f and g). One possible explanation for this outcome is that the relative rates of duplicated complex formation are similar for both insertion and annealing under the conditions explored. Here, insertion was performed over 16 hours at 40°C with an origami concentration of $\approx$50 nM and inserted component concentration of $\approx$70 nM. Potentially, complex duplication might be reduced by performing the insertion with lower origami and staple concentrations, favoring the unimolecular reaction leading to the correct latch structure. The origami concentration could subsequently be restored to a useful value through Amicon filtration.

As discussed in the conclusion of this thesis, custom scaffolds with a three-letter code enable origami to fold rapidly under isothermal room temperature conditions [139, 140], meaning that entire origami structures can be readily formed by isothermal insertion rather than annealing. This development indicates that the most effective path for future exploration of isothermal staple insertion is with custom scaffolds. Additionally, staples with three-letter codes would be much less likely to interact with each other, as may have inhibited insertion of the top gate:output pathway in DO.I.a.

Another possible approach could be to use UV crosslinking [45], as discussed in the introduction of this thesis, to stabilize an origami structure with some staples missing, and subsequently insert the missing staples at a high temperature. This approach may fail if the unbound thymidines in the regions of scaffold missing a staple become covalently bound to each other during UV crosslinking; furthermore, UV crosslinking will not prevent staple-staple interactions (again, referring to what may have caused issues for DO.I.a). Thus, I believe is is most appropriate to prioritize the custom scaffold approach.

Appendix Q

A STAPLE LAYOUT FOR A PROPOSED 14-HELIX ORIGAMI STRUCTURE THAT FORMS AT HIGH TEMPERATURES

I designed a theoretical DNA structure which might better support the latch and other localized circuits. The following criteria guided my design:

1. The structure must be more rigid than a single-layer 2D structure. It is not known how rigid a DNA origami structure must be so that reachability can be reliably constrained, but it has been established that a single-layer 2D structure is insufficiently rigid, whereas a double-layer rectangle is sufficiently rigid [125].
2. The structure must have one long dimension and one narrow dimension to facilitate a long cascade of circuit components. The narrow dimension must not be too narrow—a latch cannot fit on the outside of a six-helix bundle because of spacing requirements. However, an upper bound on the narrow dimension is required because the short lengths of available origami scaffolds are limiting.
3. The structure must either lie flat, or if it is three-dimensional, must have a high degree of symmetry. A cylinder is ideal.
4. Each staple must possess at least one domain substantially longer than 20bp, to ensure that the staples hybridize to the scaffold at a high temperature. The staples will therefore have just one long domain and one short domain to avoid excessively long staples that suffer from poor synthesis quality. It is not possible to use a 3D structure where each strand may connect to 3 or more other strands, such as a strand that is at the center of a triangular arrangement of 3 other strands—see reference [145] for an example of such a structure.
5. Each staple should be as similar to every other staple as possible, such that moving circuit components from one position to another will be less likely to affect their function. Ideally, therefore, there should be no need for deletions or insertions to account for twist.

I considered whether to use one of the two typical grids for origami strand layouts: square and hexagonal.

Double-stranded DNA has 1 turn every ≈10.5 basepairs. In order for a DNA origami structure to lie flat, each successive staple crossover must occur after a whole number of turns plus one half turn, which necessarily fails to give a whole number of basepairs. The number of basepairs between staple crossovers in a flat origami structure is given by equation Q.1:

$$bpBetweenCrossovers_{flat} = 10.5n + \frac{1}{2}10.5 = 10.5n + 5.25 \tag{Q.1}$$

Eq. Q.1 may give either a whole number plus .25, or plus .75. In reality, basepairs are indivisible units; thus, DNA origami structures that arrange their helices on a square grid must use deletions to avoid producing a twisted structure. This deletion process is called twist correction.

To avoid the complication of twist correction, many DNA origami structures use a hex grid to arrange their scaffolds instead. In this case, a staple crossover must take place every whole number of turns plus 1/3 or 2/3 turns, as represented by equations Q.2 and Q.3.

$$bpBetweenCrossovers_{hex1} = 10.5n + \frac{1}{3}10.5 = 10.5n + 3.5 \tag{Q.2}$$

$$bpBetweenCrossovers_{hex2} = 10.5n + \frac{2}{3}10.5 = 10.5n + 7 \tag{Q.3}$$

By restricting n to odd numbers in eq. Q.2 and even numbers in eq. Q.3, whole numbers can be guaranteed. While easier to design in the sense that no twist correction is required, origami arranged using this method does not lie flat because there are no straight lines on a hex grid.

Hex grids benefit from this numerical advantage over square grids for DNA origami because 21, the number of basepairs in two turns, is divisible by 3. Noting this feature, I determined that a similar advantage could be conferred by the fact that 21 is also divisible by 7. While there is no related grid that enables construction of arbitrary shapes, an approximately cylindrical tube of DNA origami can be constructed as a tetradecagon (14-gon) without the need for any twist correction. This tube would easily satisfy design criteria 2 and 3, and a natural staple layout for this structure meets criterion 4. A tube will certainly be more rigid than single-layer origami structure, so this design

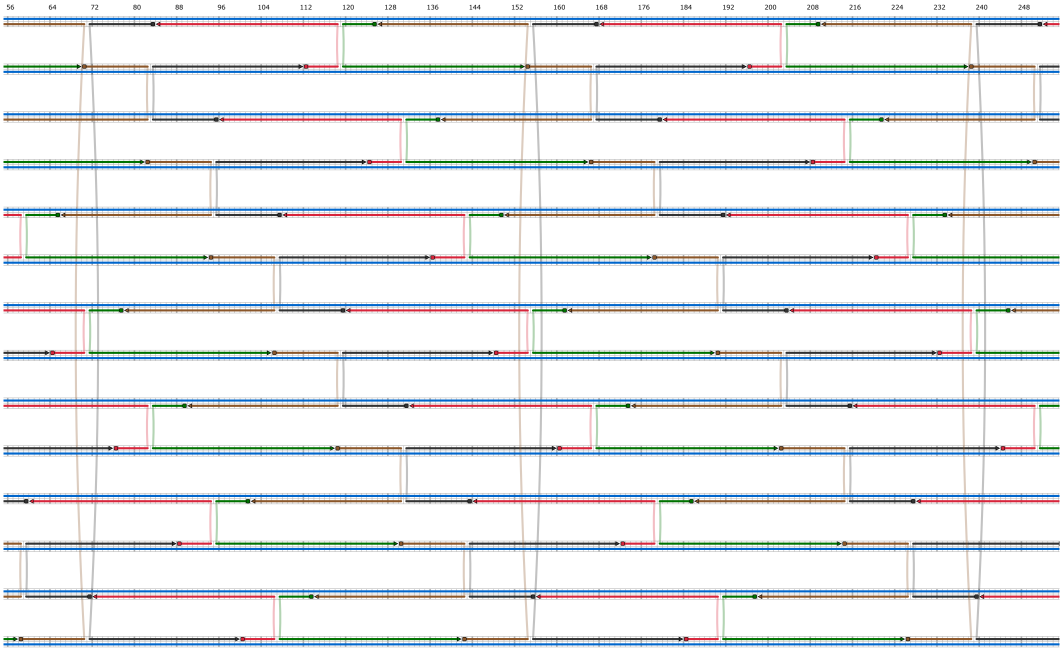


Figure Q.1: **Staple layout for a proposed 14-helix origami structure that forms at high temperatures.** Using scaffold M13p8064 as an example, the length of the tube can be at minimum 545 bp, roughly six times as long as the portion shown here.

meets criterion 1, although a potential concern is the wide spacing of staple crossovers that was required to allow high-temperature folding.

Using the Autobreak software provided by reference [159], even a version of this structure lacking edge staples was predicted to form at significantly higher temperatures than double-layer origami (compare Figure P.2 and Figure Q.2); however, as noted previously (Appendix P.2, Approach: Seed Staples and Autobreak), Autobreak was not designed to accurately predict the temperature at which each staple will join because it does not take into account cooperativity or kinetics.

Determining whether this structure is sufficiently rigid to constrain reachability, and measuring the temperature at which it folds compared to other structures, would be enough work on its own to constitute a project. Therefore, due to time constraints, it was not pursued further. Given its potential for enabling correct assembly of localized circuits, I encourage future exploration of this 14-helix tube.

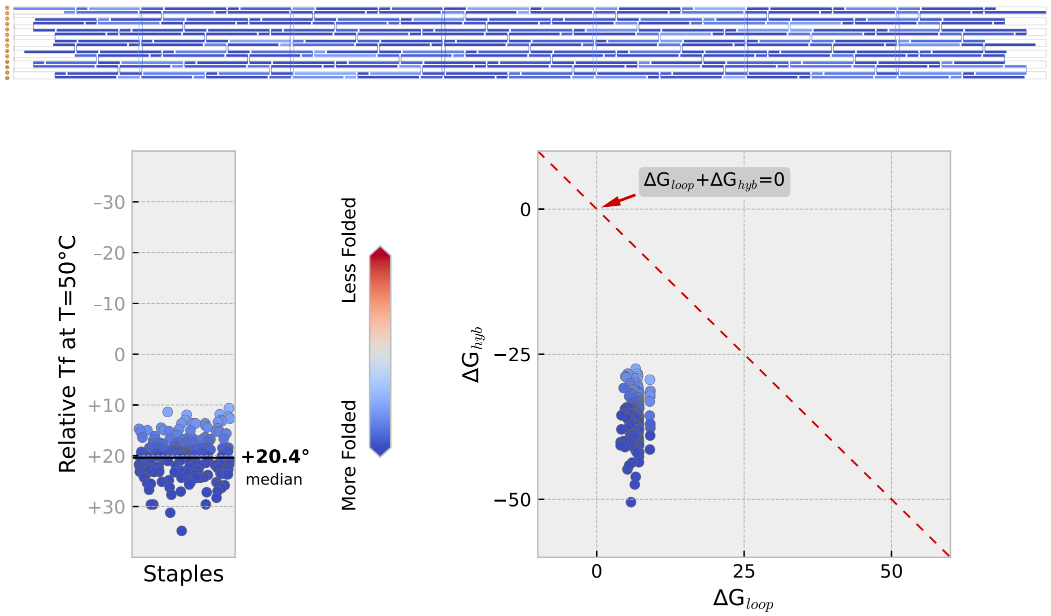


Figure Q.2: **Autobreak results for a proposed 14-helix origami structure that forms at high temperatures.** These plots were generated using the Autobreak Colab notebook [159].

Appendix R

QUANTIFYING INTER-ORIGAMI LEAK

Four experiments were run to measure various forms of inter-origami leak (Fig. R.1). No inter-origami leak was detected in any of the four. All four experiments included the fast positive control that was explained in Appendix N, Modified Positive Control and Duplicated Complexes, and the first three hours of Fig. R.1a were shown in Appendix N as Fig. N.3.

Unlike any other experiment presented in this thesis, these involved multiple types of origami variant in the same master mix. Therefore, the origami concentration measured is a summed value of every origami species present. The first two and last experiment (Fig. R.1a,b,d) were expected to have an overall origami concentration of 6 nM since two types of origami were both added at 3 nM. The third experiment (Fig. R.1c) was expected to have an overall origami concentration of 9 nM since three types of origami were added with each at 3 nM. As usual, some experimental noise produced values that deviated from these ideals somewhat.

In the first three experiments (Fig. R.1a–c), an origami with an activated fuel is only added after 3 hours. No apparent effect is produced in any experiment when this origami is added, showing that in all other experiments, % fuel leak is entirely attributable to localized leak reactions.

In the fourth experiment (Fig. R.1d), all origami variants shown are initially present. After 3 hours, the fuel activator is added, which still has no observed effect. This experiment supports the conclusion about fuel leak determined in the other three experiments by demonstrating that no cooperative leak mechanism takes place that depends on the presence of the fuel activator strand.

The third inter-origami leak experiment (Fig. R.1c) further demonstrates that no leak takes place between the input on one origami and the gate:output complex on another origami. The output curve is identical to the negative control curve.

One final conclusion can be drawn from these experiments that is unrelated to inter-origami leak. Since in all cases the fast positive control curve has faster kinetics and a higher completion level than the original positive control curve, just as it did in prior experiments involving full latch pathways, the conclusions previously drawn pertaining to this phenomenon are corroborated. Having simplified the system down to its bare essentials, particularly in the case of the first three hours of (Fig. R.1a), alternate hypotheses to explain our observations become much more challenging to formulate.

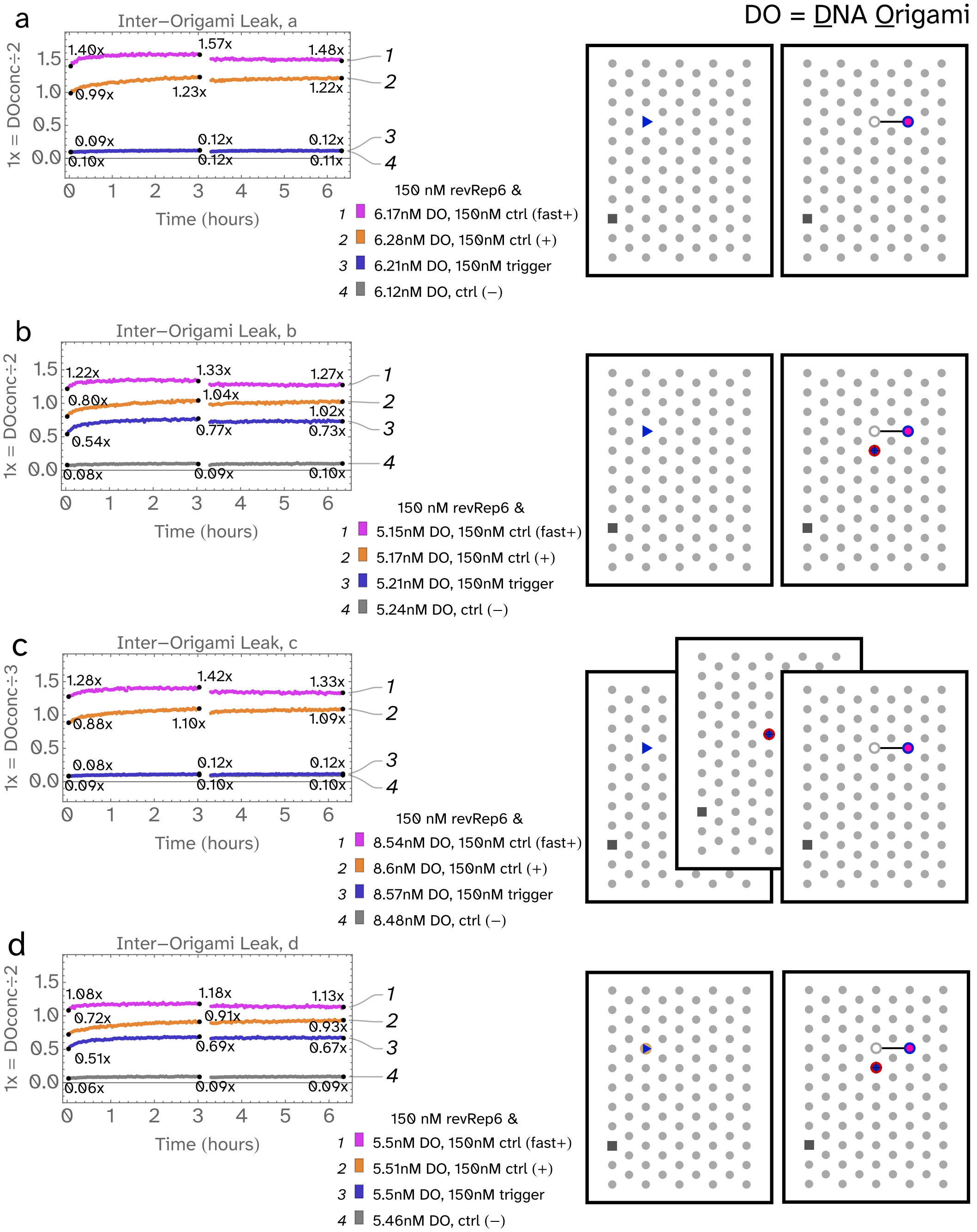


Figure R.1: **Kinetics of inter-origami leak experiments. a,** Origami with an output:gate complex. At t=3h, origami with an active fuel is added. **b,** Origami with all components of a one-pathway latch except fuel. At t=3h, origami with an active fuel is added. **c,** One origami with an input and a second origami with an output:gate complex. At t=3h, origami with an active fuel is added. **d,** One origami with an inactive fuel only and a second origami with all other components of a one-pathway latch. Both origamis are initially present and the fuel activator is added at t=3h.

Appendix S

SI FOR CHAPTER 2

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*

Supporting information

S1 Methods

Sequence design

As shown in Figure S3, each hairpin consists of two exposed 7-nt toeholds, a conserved 6-nt branch migration domain (x), a sequestered 7-nt toehold, and a 4-nt loop domain (s = `TTTT`). For implementation of a DNA ring with 4, 6, or 8 strands, nine unique toeholds (a through i) were needed in seven hairpins. Sequences of domains a, x, and b were predetermined so as to utilize a previously designed reporter complex [41, 99] for fluorescence readout. The remaining seven toeholds (c through i) were designed using the multi-state concentration-based design tool in NUPACK [160]. A three-letter code (`A`, `T`, and `C` only) [23, 105] was utilized for all asterisk domains ($a*$, $b*$, etc.), which reduces undesired secondary structures in the single-stranded domains functioning as an input to open up a hairpin (e.g. c x b in hairpin2 for reacting with hairpin1, c x d in hairpin3 for reacting with haripin2, etc.). However, this strategy does not address spurious interactions between the two exposed toeholds in a hairpin, as they have opposite three-letter codes (e.g. `A`, `T`, `C` for toehold $c*$ and `A`, `T`, `G` for toehold a in hairpin1). Thus, it was important that target hairpin structures were specified in NUPACK. Moreover, to achieve roughly the same forward and backward reaction rates of each reversible assembly step in the ring formation, we set the target concentrations of each hairpin at opened and closed states both to be 50 nM. An example NUPACK script and preview for designing toehold c are shown below:

```
#
# design material, temperature (C), and trials
#
material = DNA
temperature = 25.0
trials = 5
#
# target structures
#
structure hairpin1 = U7 D13 U4 U7
structure input1 = U20
structure hairpin1input1 = U24 D20 +
#
# sequence domains
#
domain a = TGAGATG
domain x = TGATTG
domain b = TGTTATG
domain s = T4
domain c = D7
#
# strands
#
strand hairpin = a x b s b* x* c*
strand input = c x b
#
# thread strands onto target structures
#
hairpin1.seq = hairpin
input1.seq = input
hairpin1input1.seq = hairpin input
#
# target test tubes
#
tube HItube = hairpin1 input1 hairpin1input1
#
```

```
# target concentrations for target structures in test tubes
#
HItube.hairpin1.conc[nM] = 50
HItube.input1.conc[nM] = 50
HItube.hairpin1input1.conc[nM] = 50
#
# design against all off-target ordered complexes of up to this
# number of strands (design against monomers, dimers, and trimers)
#
HItube.maxsize = 3
#
# stop conditions for normalized ensemble defect
#
HItube.stop[%] = 2.0
```

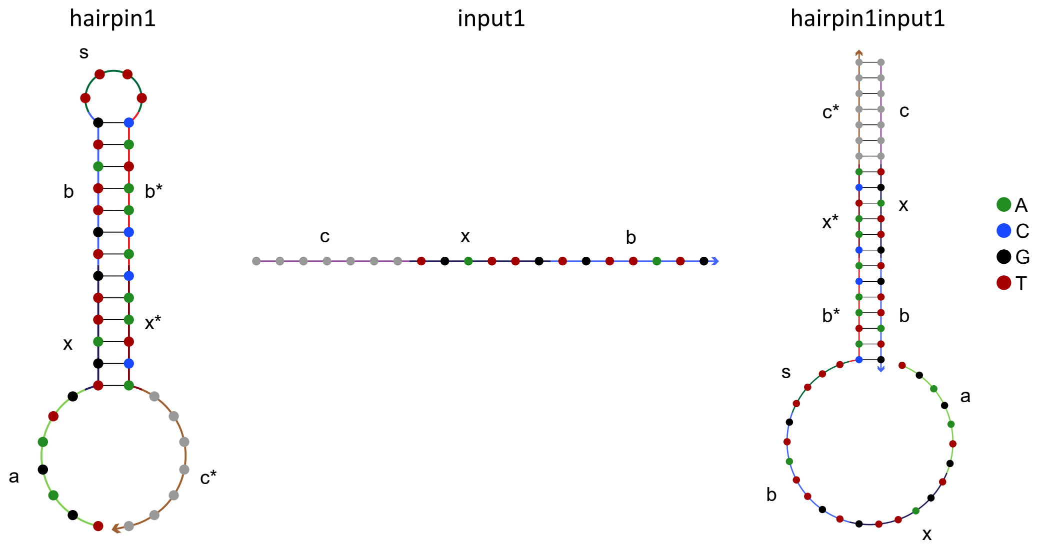


Figure S1: **NUPACK preview for designing the sequence of toehold c.**

Toeholds c to i were designed sequentially, and every new toehold sequence was selected by comparing it to the existing toeholds and ensuring that no pairs of toeholds share more than three continuous nucleotides. Once all toehold sequences were obtained, hairpin structures at closed and opened states were analyzed with the NUPACK analysis tool [77] (Figure S2). Linear structures involved in the designed self-assembly pathways were also analyzed to verify that intermediate products are expected to assemble correctly and that no significant spurious interactions are predicted among system components.

a

	1 uM	**100 nM**
hairpin1 + input1	0.84 uM	58 nM
hairpin2 + input2	0.81 uM	52 nM
hairpin3 + input3	0.80 uM	51 nM
hairpin4 + input4	0.81 uM	51 nM
hairpin5 + input5	0.80 uM	50 nM
hairpin6 + input6	0.76 uM	43 nM
hairpin7 + input7	0.79 uM	49 nM
Average equilibrium constant (K_c)	**4**	**1**

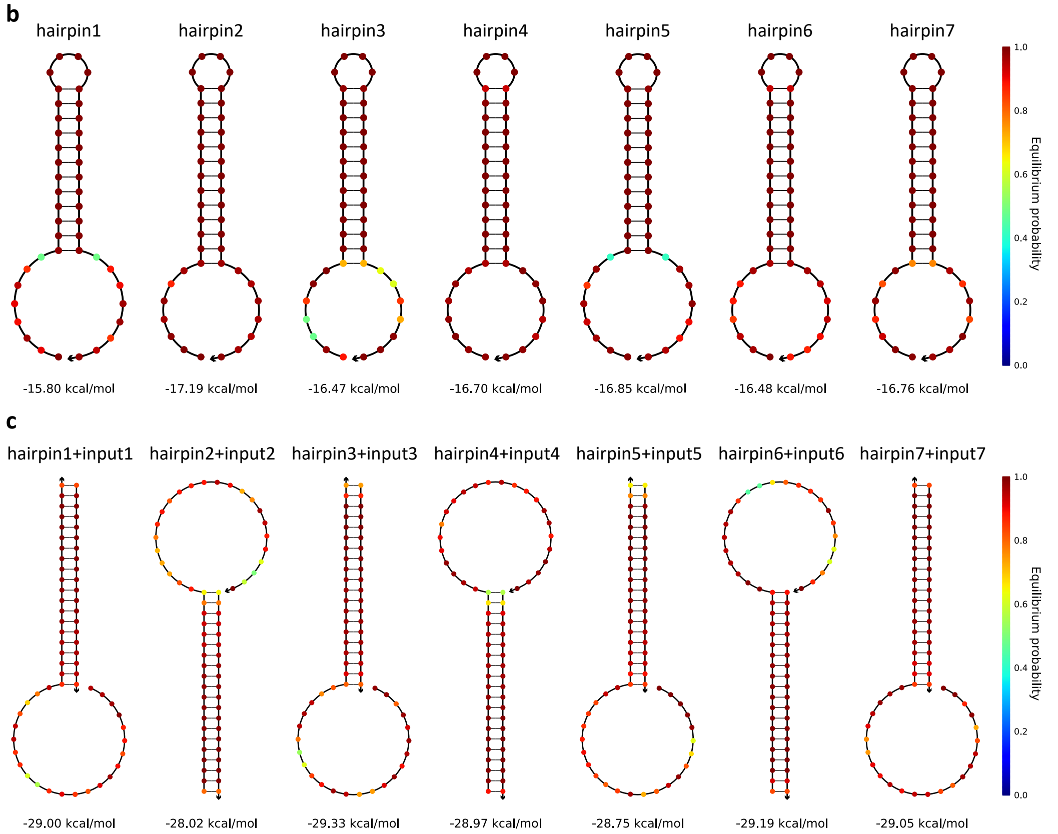


Figure S2: **NUPACK analysis of hairpin structures and equilibrium conditions.** (a) Equilibrium concentration of each hairpin:input complex with 1 μM or 100 nM of each strand. (b) Minimum free energy (MFE) structure of each hairpin at 25 °C. (c) Minimum free energy (MFE) structure of each hairpin:input complex at 25 °C.

Sequences of the key complexes were simply obtained by composing the domain sequences together. For understanding the kinetics of ring formation, an additional 7-nt toehold (t) was designed and used in hairpin1 and keys (Figure S3).

For realizing the reporting mechanism shown in Figure 2.5b, the fluorophore and quencher-labeled strands were directly taken from a previously developed reporter [41, 99]. A cover strand partially complementary to the quencher-labeled strand was designed to facilitate the reporting reaction, which is initiated by a 5-nt effective toehold consisting of a 2-nt single-stranded domain in the partially exposed $b*$ domain together with the fluorophore-quencher interaction which was experimentally measured to be equivalent to $\approx$3 nucleotides. All DNA sequences of strands for creating a ring and detecting the ring formation are listed in Table S1.

Eleven non-hairpin strands used in formation gel controls were designed with random domain sequences satisfying the three-letter code (Table S2). Each strand consists of four unique 7-nt domains, two unique 6-nt domains, and a 4-nt `TTTT` domain equivalent to the toeholds, branch migration domains, and loop domain in the hairpins. Strands 1 through 3, together with strand4-linear and key4-F were designed to form a 5-stranded linear structure similar to that shown in Figure 2.3, but the single-stranded part of the structure does not have the same sequence as key4-F and thus cannot displace it. Strands 1 through 3, together with strand4-ring were designed to form a 4-stranded ring structure similar to that shown in Figure 2.3, but all double-stranded domains have unique sequences so that the only product that should form after annealing is a ring. Similarly, strands 5 through 8 were designed as controls for the 6 and 8-stranded ring formation.

Three trigger strands used in control experiments for fluorescence data normalization (Figure S8a) were designed by taking the existing domain sequences and shortening them to create an 8-nt toehold for releasing the key-F strand in each of the three key complexes (Table S3).

Sample preparation

DNA oligonucleotide synthesis. Custom single-stranded DNA oligos were manufactured by Integrated DNA Technologies (IDT). Strands modified with a fluorescent dye or dark quencher were delivered HPLC-purified while unmodified strands were delivered with standard desalting only. With formulation service LabReady, all strands were provided at ≈100 μM in IDTE, i.e., IDT-brand Tris-EDTA (TE) buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0). Before use, concentration of each strand was confirmed on a NanoDrop (Thermo Fisher) by averaging three measurements of absorbance at 260 nm for a 1 μL droplet. All strands were stored at 4 °C.

Annealing protocol and buffer condition. TE buffer with 10×$MgCl_2$ was added to mixtures of strands intended to form double-stranded complexes or individual strands intended to form hairpins to produce samples that are ready for annealing in a final buffer of TE with 12.5 mM $MgCl_2$. The two strands in the keys and the reporter were mixed at an equimolar ratio and annealed at 10 μM. The linear and circular control strands were mixed at equimolar ratios and annealed at 3.5 μM. Each hairpin strand was annealed at 90 μM. Reaction mixtures were all held at 90 °C for 2 minutes, then ramped down to 20 °C by 0.1 °C per 6 seconds on a thermal cycler (Eppendorf).

Purification. Polyacrylamide gel electrophoresis (PAGE) was used to purify annealed hairpins on a 12% gel made with, and run in, TAE buffer with 12.5 mM magnesium acetate tetrahydrate at 150 volts for 9 hours. A single desired band for each complex was excised, minced, and left idle for at least 24 hours at room temperature (≈24 °C) in TE buffer with 12.5 mM $MgCl_2$, during which period DNA transferred from the gel to the buffer by diffusion. After buffer recovery, concentrations of purified hairpins were measured on a NanoDrop (Thermo Fisher).

Formation gel experiments

Native PAGE was used to separate DNA on a 6% gel made with, and run in, TBE buffer with 12.5 mM magnesium acetate tetrahydrate at 100 volts for 112 minutes. Gels were imaged using the Bio-Rad ChemiDoc MP Imaging System and Image Lab software. For imaging gels stained with SYBR Gold, the default Image Lab settings for SYBR Gold were used : Standard Filter (Filter 1) and UV trans illumination. For imaging strands with fluorophore ATTO590, the 695/55 filter and red epi LED light source were used.

Fluorescence kinetics experiments

A 96-well plate (Corning) was used to organize 110 μL of each sample, mixed at a standard concentration 1× $=$100 nM. Fluorescence levels were taken in 2-minute intervals on a microplate reader (Synergy H1, Biotek) at room temperature (≈22 °C). 598 nm excitation and 629 nm emission wavelengths were used for fluorophore ATTO590.

2 μM of a 20-nt poly(T) strand (referred to as 20T) was added to each master mix to coat the walls of pipette tips and tubes prior to adding any other strands [105]. 20T, along with ubiquitous use of LoRetention tips and DNA LoBind tubes (Eppendorf), mitigated the loss of DNA strands to surfaces.

In each set of fluorescence kinetics experiments, a master mix was made with 20× 20T, 1× key, and 1.5× reporter. The master mix was transferred to one well per sample on a microplate, and various combinations of trigger strands or hairpins at $1.5\times$ were added to each of the wells. For negative controls, TE was added to maintain the same target concentration. Employing a master mix rather than preparing each sample individually is known to result in data with better reproducibility.

Data normalization

As shown in Figure S8, each set of fluorescence kinetics experiments included a negative control with just a key and reporter as well as a positive control with an excess trigger strand added to directly release the fluorophore-labeled strand from the key complex without the need for any hairpins. The maximum raw fluorescence was calculated by averaging the initial five measurements of the negative control, and the minimum raw fluorescence was calculated by averaging the final five measurements of the positive control. These values were respectively taken as 0× and 1× to convert fluorescence data to concentration data. In this data normalization method, 0 nM (0×) is interpreted as the highest fluorescence of the reaction mixture when no key-F strand has been released, and 100 nM (1×) is interpreted as the lowest fluorescence of the reaction mixture after key-F has been fully released from the key complex and become quenched by the reporter.

S2 Modeling and simulation

All simulations were performed with mass-action kinetics using CRNSimulator [22].

The following reactions were used to model the assembly steps:

$$\begin{aligned} key_i + hairpin_{i-1} &\underset{kr}{\overset{kf}{\rightleftharpoons}} KH_{i-1} \quad i = 4, 6, 8 \\ KH_i + H_{i-1} &\underset{kr}{\overset{kf}{\rightleftharpoons}} KH_{i-1} \quad i = 2, 3, \ldots, 7 \end{aligned} \tag{S.1}$$

where $kf = 2 \times 10^6$ /M/s and $kr = 0.2$ /s. Each species KH_i indicates a linear polymer that starts with a key and ends with hairpin$_i$. kf was estimated based on the effective strand displacement rate with a 7-nt toehold [78, 79]. kr was calculated as follows:

$$kr = \frac{kf \times c}{K_c} \tag{S.2}$$

where standard concentration $c = 100$ nM, and $K_c = 1$ was obtained from NUPACK analysis (Figure S2a).

The following reaction was used to model the disassembly step:

$$KH_1 \xrightarrow{4k/i} ring + keyt \tag{S.3}$$

where $k = 0.0001$ /s, 0.04 /s, and 0.5 /s for 0, 4, and 7-nt toeholds, respectively, were estimated based on experiments shown in Figure S9c. Here we assume that the unimolecular strand displacement rate scales inversely proportional to the ring size ($i = 4, 6, 8$).

The following reaction was used to model the reporting step for creating a ring with responsive size:

$$keyt + reporter \xrightarrow{krep} wasteFQ + waste \tag{S.4}$$

where $krep = 10^5$ /M/s was estimated based on control experiments shown in Figure S8c.

The following reactions were used to model the reporting step for creating a ring with responsive growth direction:

$$\begin{aligned} key4At + translatorA &\xrightleftharpoons[kt]{kt} signalA + wasteA \\ key4Bt + translatorB &\xrightleftharpoons[kt]{kt} signalB + wasteB \\ signalA + reporterA &\xrightleftharpoons[kFQ1]{krep1} wasteF1 + wasteQ1 \\ signalB + reporterB &\xrightleftharpoons[kFQ2]{krep2} wasteF2 + wasteQ2 \end{aligned} \tag{S.5}$$

where $kt = 2 \times 10^6$ /M/s was estimated based on the effective strand displacement rate with a 7-nt toehold on the translators, $krep1 = krep2 = 10^5$ /M/s were estimated based on the effective strand displacement rate with a 5-nt toehold on the reporters, and $kFQ1 = 1.5 \times 10^3$ /M/s and $kFQ2 = 2 \times 10^3$ /M/s were estimated based on experiments shown in Figure 2.6c. The reverse rates of the reporters ($kFQ1$ and $kFQ2$) are due to the interaction between a fluorophore and a quencher, and the estimates here are consistent with previous estimates using the same reporters.

The following reactions were used to model leak:

$$\begin{aligned} key_i + reporter &\xrightarrow{klrep} wasteFQ + waste \quad i = 4, 6, 8 \\ key_i + hairpin_1 &\xrightarrow{kl} keyt \quad i = 4, 6, 8 \end{aligned} \tag{S.6}$$

where $klrep = 4$ /M/s, and $kl = 0$ /M/s, 0 /M/s, and 3 /M/s for 0, 4, and 7-nt toeholds, respectively, were estimated based on experiments shown in Figures S9a and b.

The following reactions were used to model the assembly steps in growing a linear polymer longer than the target ring size:

$$\begin{aligned} KH_1 + key_i &\xrightarrow{kp} KKH_i + keyt \quad i = 4, 6, 8 \\ KH_1 + KH_i &\xrightarrow{kp} KKH_i + keyt \quad i = 1, 2, \ldots, 7 \\ KKH_i + H_{i-1} &\xrightleftharpoons[kr]{kf} KKH_{i-1} \quad i = 2, 3, \ldots, 8 \end{aligned} \tag{S.7}$$

where kf and kr are the same as above and $kp = 10^6$ /M/s was estimated based on gel electrophoresis data shown in Figures S4 and S5.

The following reaction was used to model the disassembly step in a linear polymer closing up to form a double-sized ring:

$$KKH_1 \xrightarrow{4k/i} ring2 + keyt \tag{S.8}$$

where k is the same as above and $i = 8, 12, 16$ were used for the doubled ring sizes.

S3 Supplementary design diagrams, simulations, and experiments

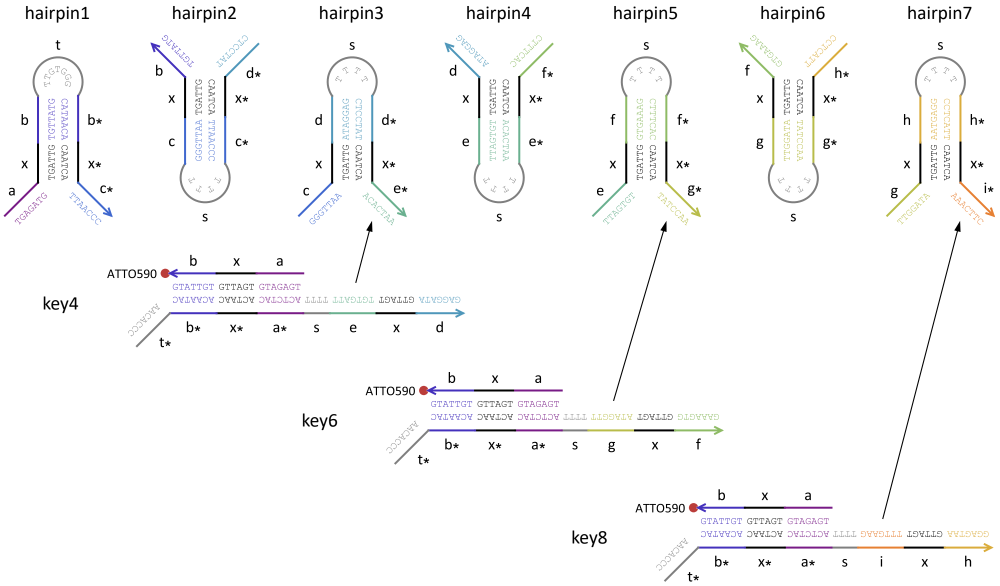


Figure S3: **Sequence-level design diagrams of the DNA ring with responsive size.** A 7-nt toehold t in hairpin1 and its complementary toehold t$*$ in the keys are shown here. For experiments on the kinetics of ring closure with 0 and 4-nt toeholds (Figure 2.5), the t domain was `TTTT` in hairpin1 (same as the s domain in other hairpins) and the t$*$ domain was nonexistent or `AAAA` in key4.

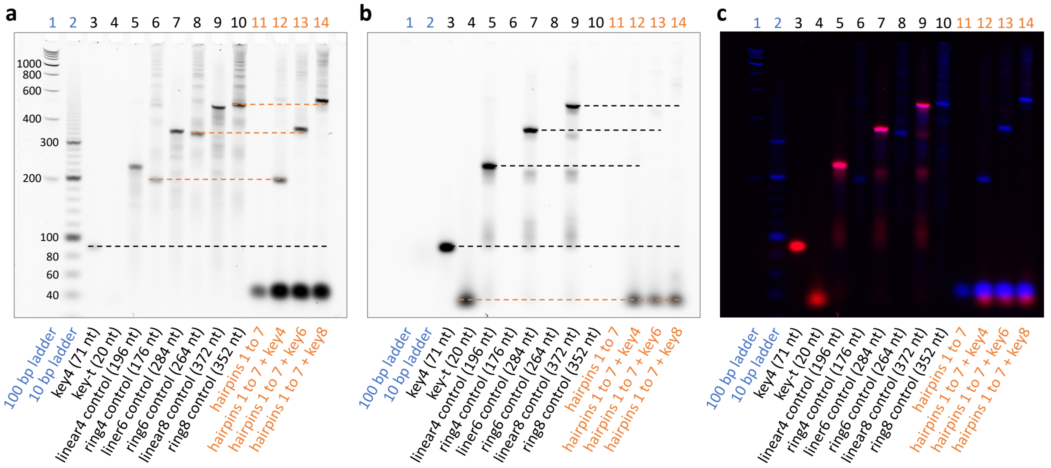


Figure S4: **Gel electrophoresis of ring formation with (a) SYBR Gold staining (same image as in Figure 2.4b), (b) ATTO590 fluorescence (same image as in Figure 2.4c), and (c) the two channels overlaid.** Lanes containing DNA ladders, control structures, and hairpins without or with a key are labeled in blue, black, and orange, respectively. Dashed black and orange lines indicate a comparison between controls and reactants or intermediates, and between controls and products, respectively. Hairpins, keys, and control structures were at 150, 100, and 100 nM, respectively. Samples of hairpins and a key were incubated at room temperature for roughly one hour before they were loaded onto the gel.

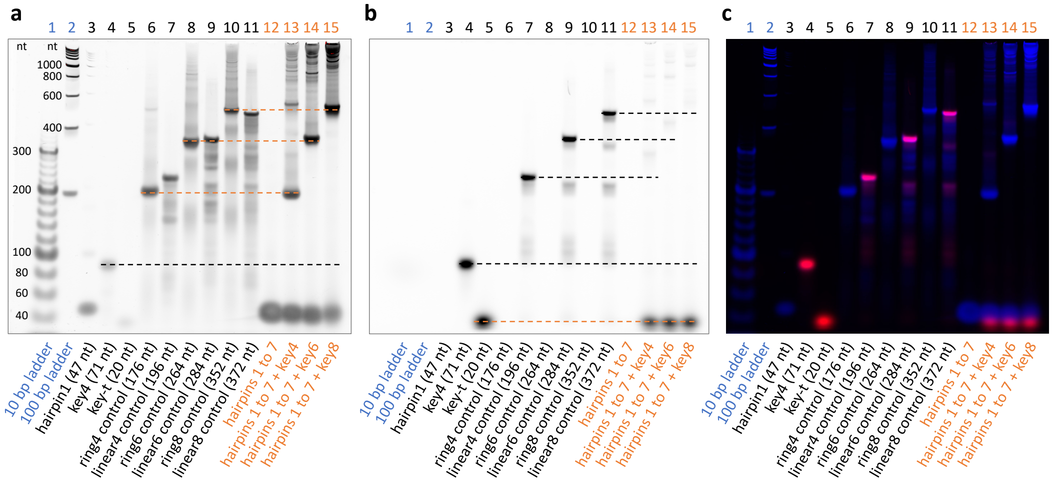


Figure S5: **Gel electrophoresis of ring formation with (a) SYBR Gold staining, (b) ATTO590 fluorescence, and (c) the two channels overlaid.** All conditions were the same as in Figure S4, except that hairpins and keys were at 1 and 1.5 μM concentrations, respectively.

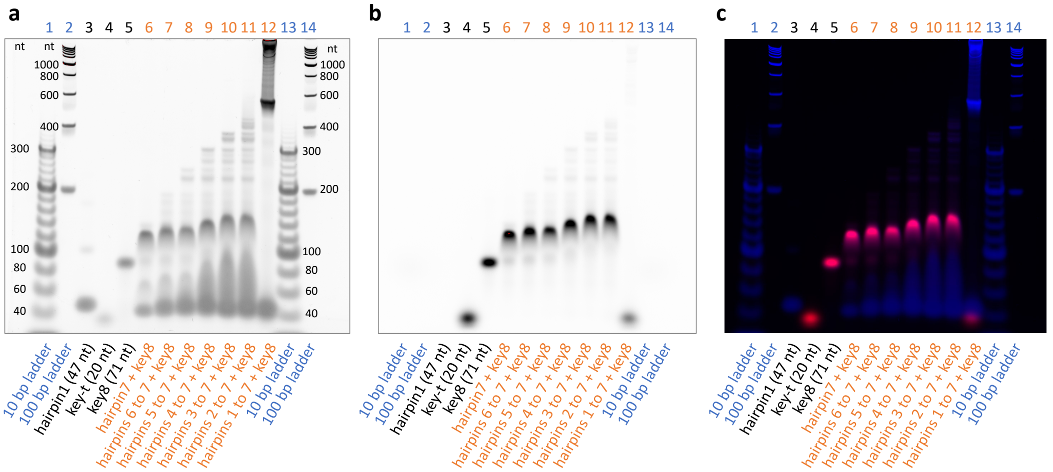


Figure S6: **Gel electrophoresis of 8-stranded ring formation including intermediate steps with (a) SYBR Gold staining, (b) ATTO590 fluorescence, and (c) the two channels overlaid.** Lanes containing DNA ladders, control structures, and hairpins without or with a key are labeled in blue, black, and orange, respectively. Hairpins and key8 were at 1 and 1.5 μM concentrations, respectively. All ring and linear control structures (sequences shown in Table S3) were at 1 μM concentration. Samples of key8 with hairpin(s) were incubated at room temperature for roughly one hour before they were loaded onto the gel.

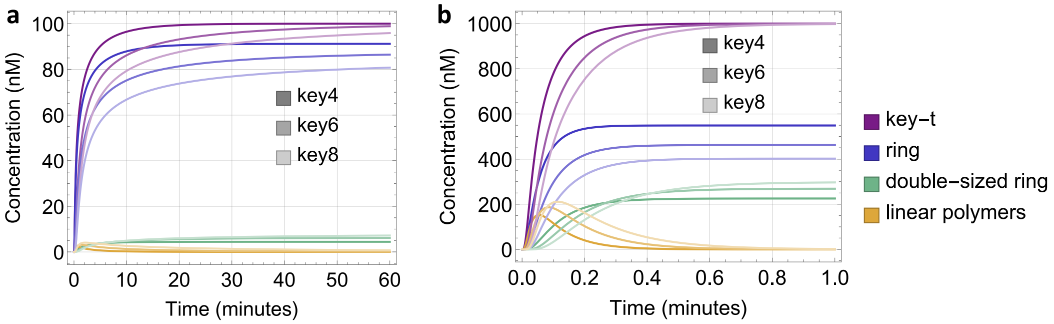


Figure S7: **Simulation of ring formation at (a) 100 nM and (b) 1 μM concentrations, considering the formation of linear polymers that are longer than the target ring size and one such polymer closing up into a double-sized ring.** Reactions and rate constants are described in SI Note S2, Modeling and simulation.

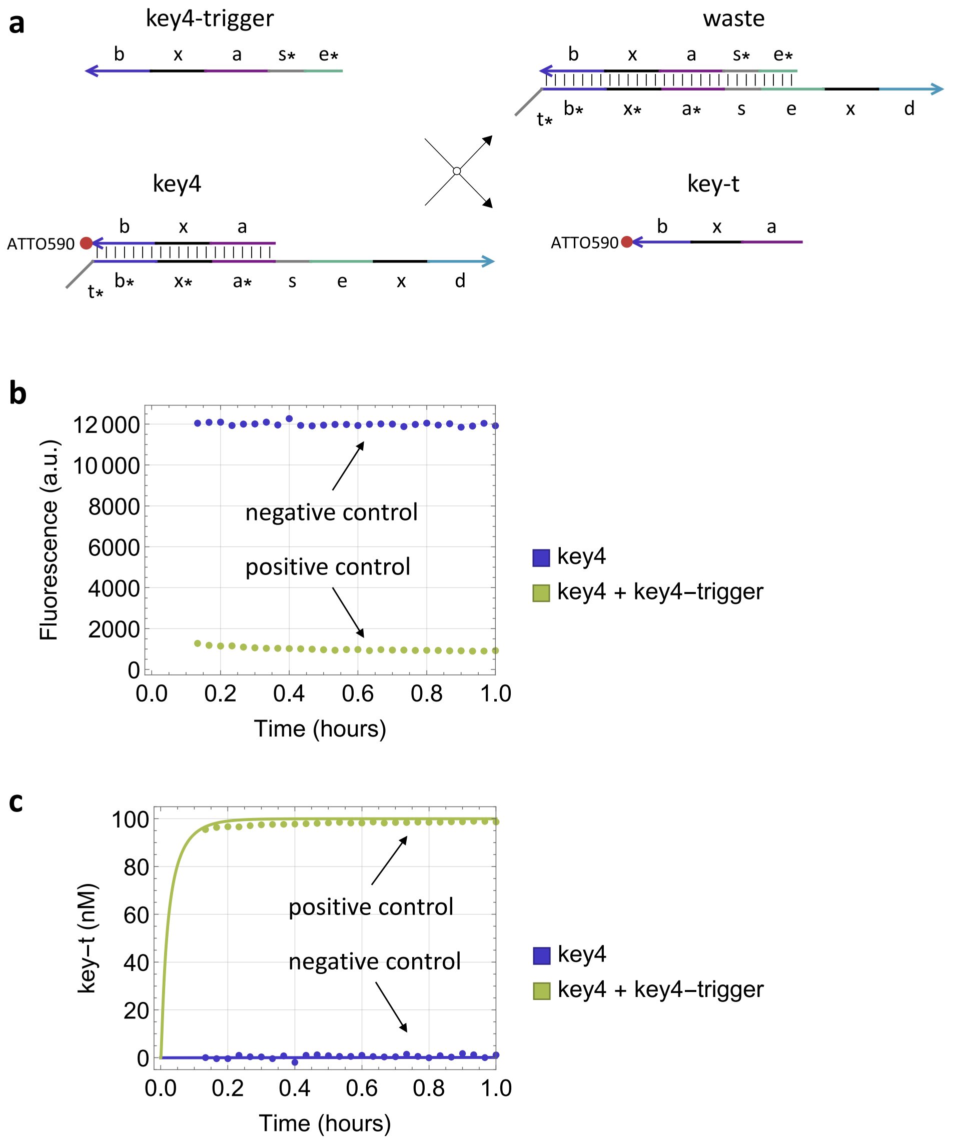


Figure S8: **Control experiments for fluorescence data normalization.** (a) Design of a triggering mechanism that releases the key-t strand. The reaction is initiated by an 8-nt toehold consisting of the 4-nt s∗ domain and the first 4 nucleotides on the 3’ end of the e∗ domain. (b) Raw fluorescence data of negative and positive controls. (c) Simulation and normalized data of negative and positive controls. Simulation and data are shown as solid and dotted trajectories, respectively. Initial concentrations of key4, key4-trigger, and reporter were 100, 150, and 150 nM, respectively. Raw fluorescence was converted to concentration by using the average of the first five data points in the negative control as 0 and that of the last five data points in the positive control as 100 nM. All fluorescence kinetics experiments in this paper were performed with their negative and positive controls using corresponding trigger strands (sequences shown in Table S4).

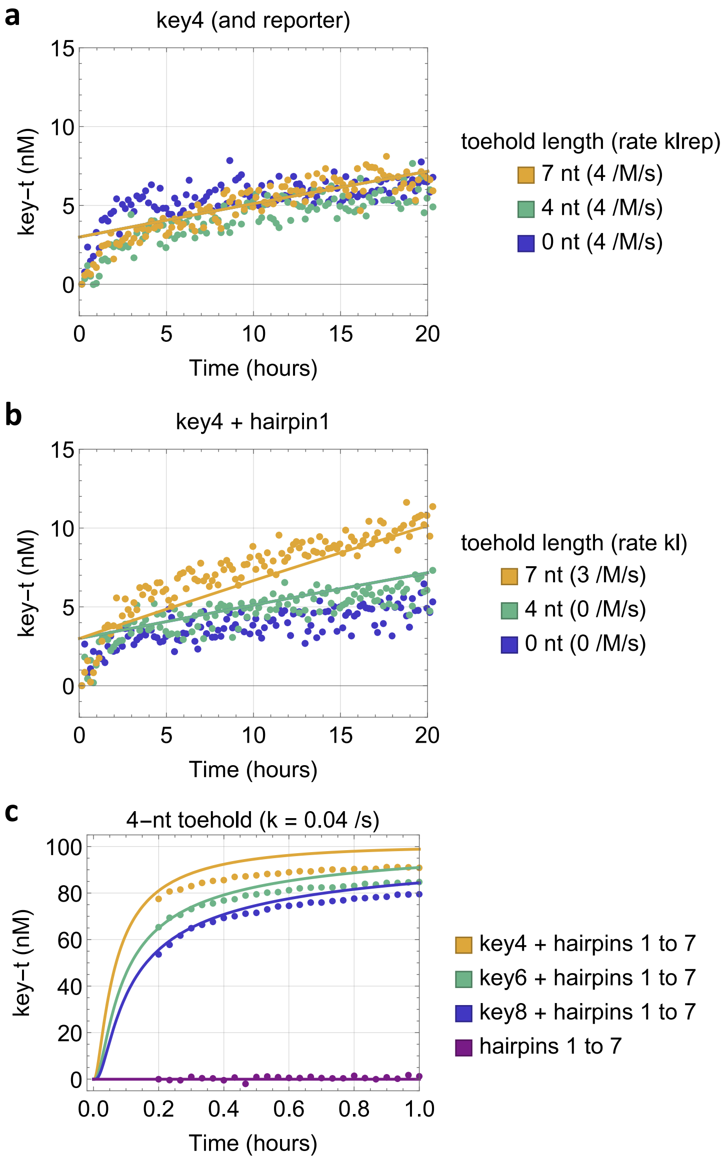


Figure S9: **Fluorescence kinetics of ring formation.** (a) Leak reaction between a key and reporter. (b) Leak reactions between hairpin1 and a key with various toehold lengths. (c) Ring formation with a 4-nt toehold. Simulation and data are shown as solid and dotted trajectories, respectively. Initial concentrations of keys, hairpins, and the reporter were 100, 150, and 150 nM, respectively. Simulation of leak reaction was adjusted with an initial offset in order to fit the long-term slope of the data for rate estimation. The difference between simulation and experimental data in c indicates that the concentrations of keys were roughly 8% below that in the positive controls, which is within the range of common experimental errors.

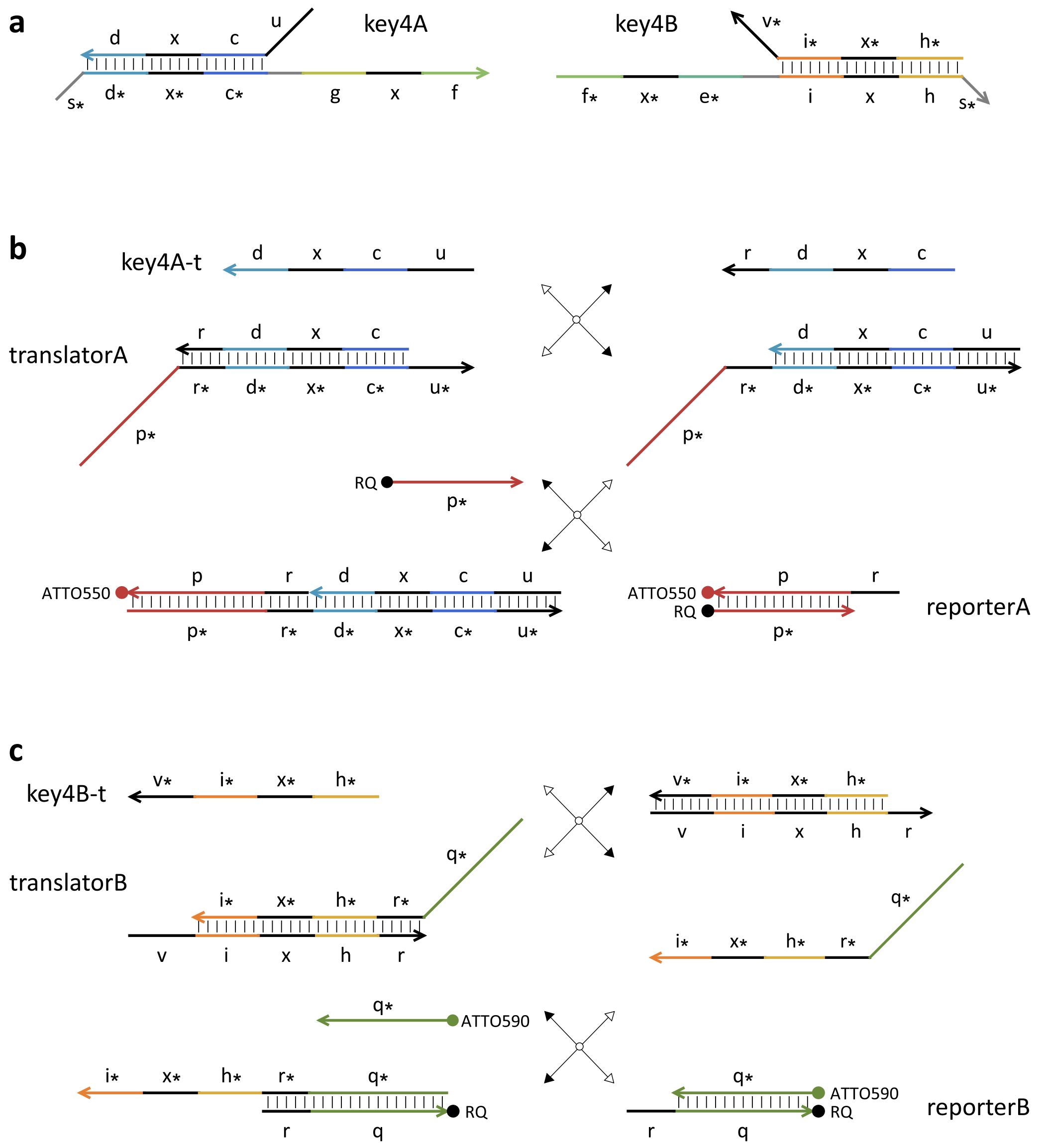


Figure S10: **Reporting mechanism for bidirectional ring formation.** (a) Each of the two keys is modified with an additional toehold in the top strand (u and v∗, respectively). The toehold for ring closure is 4-nt long in each of the two keys (s∗, complementary to the loop domain in hairpin1 and hairpin5 shown in Figure 2.6b). (b, c) Once released, each top strand can react with a translator to yield a signal that can be detected by a standard reporter [41, 99]. The reaction with a reporter is typically modeled as an irreversible reaction. However, when the upstream reaction is reversible and not driven forward catalytically (e.g. the reaction with a translator here), it is more accurate to model the reporting reaction as a reversible reaction, because the fluorophore and quencher can bind to each other with a strength similar to a 3-nt toehold (SI Note S2, Modeling and simulation).

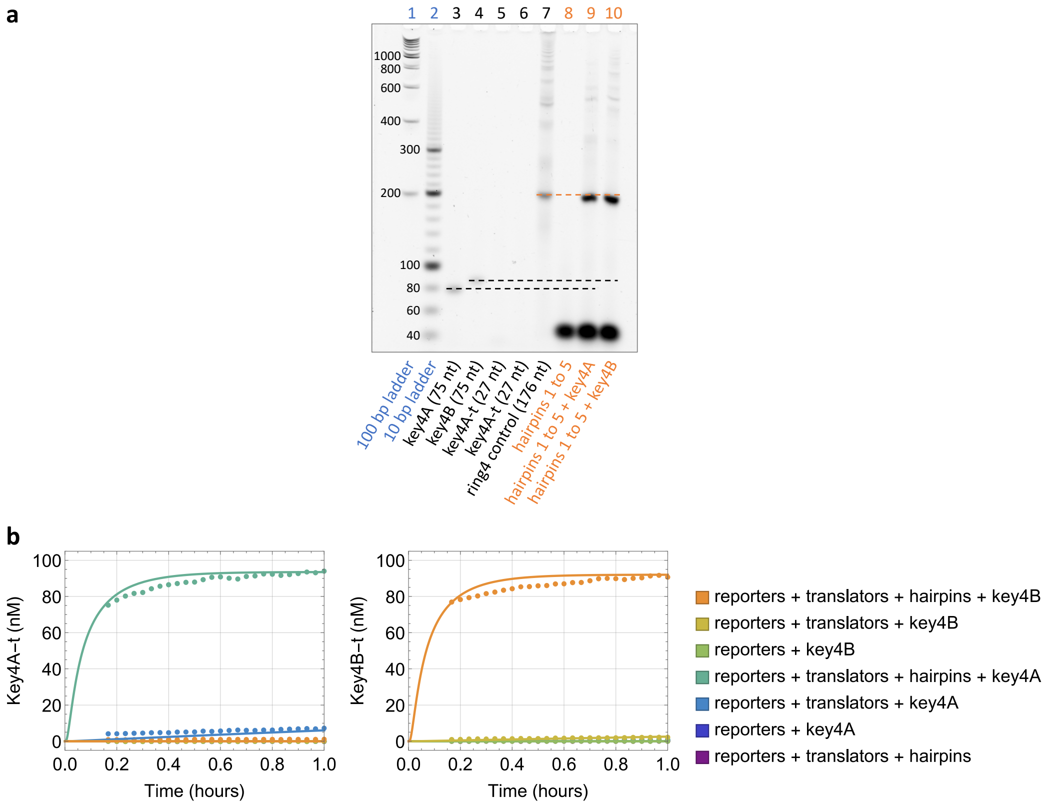


Figure S11: **Experiments for bidirectional ring formation.** (a) Gel electrophoresis with SYBR Gold staining. Lanes containing DNA ladders, control structures, and hairpins without or with a key are labeled in blue, black, and orange, respectively. Dashed black and orange lines indicate a comparison between controls and reactants or intermediates, and between controls and products, respectively. Hairpins, keys, and the control structure were at 150, 100, and 100 nM concentrations, respectively. Samples of hairpins and a key were incubated at room temperature for at roughly one hour before they were loaded onto the gel. (b) Simulation and fluorescence kinetics data. Simulation and data are shown as solid and dotted trajectories, respectively. The two pairs of trajectories for reporters + translators + hairpins + key4A and reporters + translators + hairpins + key4B are the same as the trajectories shown in Figure 2.6c. Initial concentrations of keys, hairpins, translators, and reporters were 100, 150, 150, and 150 nM, respectively. A slow leak reaction was observed between a key and a translator, with an estimated rate of 120 /s and 50 /s, respectively, for key4A and key4B. These estimates are similar to the previously measured rate of four-way strand displacement with two toeholds of lengths 6 and 0 nt [81].

S4 DNA sequences

Table S1: DNA sequences of strands for creating a ring with responsive size.

Name	Sequence
Hairpin1	TGAGATGTGATTGTGTTATGTTTTCATAACACAATCATTAACCC
Hairpin1-7ntToe	TGAGATGTGATTGTGTTATGTTGTGGGCATAACACAATCATTAACCC
Hairpin2	CTCCTATCAATCATTAACCCTTTTGGGTTAATGATTGTGTTATG
Hairpin3	GGGTTAATGATTGATAGGAGTTTTCTCCTATCAATCAACACTAA
Hairpin4	CTTTCACCAATCAACACTAATTTTTTAGTGTTGATTGATAGGAG
Hairpin5	TTAGTGTTGATTGGTGAAAGTTTTCTTTCACCAATCATATCCAA
Hairpin6	CCTCTTACAATCATATCCAATTTTTTGGATATGATTGGTGAAAG
Hairpin7	TTGGATATGATTGTAAGAGGTTTTCCTCTTACAATCAAAACTTC
Key-t	TGAGATGTGATTGTGTTATG /3ATTO590N/
Key4-b-0ntToe	CATAACACAATCACATCTCATTTTTTAGTGTTGATTGATAGGAG
Key4-b-4ntToe	AAAACATAACACAATCACATCTCATTTTTTAGTGTTGATTGATAGGAG
Key6-b-4ntToe	AAAACATAACACAATCACATCTCATTTTTTGGATATGATTGGTGAAAG
Key8-b-4ntToe	AAAACATAACACAATCACATCTCATTTTGAAGTTTTGATTGTAAGAGG
Key4-b-7ntToe	CCCACAACATAACACAATCACATCTCATTTTTTAGTGTTGATTGATAGGAG
Key6-b-7ntToe	CCCACAACATAACACAATCACATCTCATTTTTTGGATATGATTGGTGAAAG
Key8-b-7ntToe	CCCACAACATAACACAATCACATCTCATTTTGAAGTTTTGATTGTAAGAGG
Reporter-cover	ATGTGATTGTGTT

Reporter-quencher	/5IAbRQ/ CATAACACAATCACA

Table S2: DNA sequences of strands for creating a ring with responsive growth direction.

Name	Sequence
Hairpin1	GGGTTAATGATTGATAGGAGTTTTCTCCTATCAATCAACACTAA
Hairpin2	CTTTCACCAATCAACACTAATTTTTTAGTGTTGATTGATAGGAG
Hairpin3	TTAGTGTTGATTGGTGAAAGTTTTCTTTCACCAATCATATCCAA
Hairpin4	CCTCTTACAATCATATCCAATTTTTTGGATATGATTGGTGAAAG
Hairpin5	TTGGATATGATTGTAAGAGGTTTTCCTCTTACAATCAAAACTTC
Key4A-t	AGAGTGAGGGTTAATGATTGATAGGAG
Key4A-b	AAAACTCCTATCAATCATTAACCCTTTTTTGGATATGATTGGTGAAAG
Key4B-t	CACCTCTTACAATCAAAACTTCACCACAA
Key4B-b	CTTTCACCAATCAACACTAATTTTGAAGTTTTGATTGTAAGAGGAAAA
TranslatorA-t	GGGTTAATGATTGATAGGAGTGAGATG
TranslatorA-b	CACTCATCCTTTACATCTCACTCCTATCAATCATTAACCCTCACTCT
TranslatorB-t	CATAACACAATCACATCTCACCTCTTACAATCAAAACTTC
TranslatorB-b	TTGTGGTGAAGTTTTGATTGTAAGAGGTGAGATG
ReporterA-fluorophore	TGAGATGTAAAGGATGAGTG /3ATTO550N/
ReporterA-quencher	/5IAbRQ/ CACTCATCCTTTACA
ReporterB-fluorophore	TGAGATGTGATTGTGTTATG /3ATTO590N/
ReporterB-quencher	/5IAbRQ/ CATAACACAATCACA

Hairpins 1 through 5 in Table S2 are the same as hairpins 3 through 7 in Table S1. ReporterB-fluorophore and reporterB-quencher in Table S2 are the same as key-t and reporter-quencher in Table S1.

Table S3: DNA sequences of strands used in gel controls.

Name	Sequence
Strand1	CATAACACAATCACATCTCATTTTTTAACCCTCTTCACTCCTAT
Strand2	TTAGTGTGTATGTGTGAAAGTTTTATAGGAGTGAAGAGGGTTAA
Strand3	CTTTCACACATACACACTAATTTTTATCCAATCACCTCCTCTTA
Strand4-Linear	GAAGTTTGAGAGTATATGTATTTTTAAGAGGAGGTGATTGGATA
Strand4-Ring	TGAGATGTGATTGTGTTATGTTTTTAAGAGGAGGTGATTGGATA
Strand5	TACATATACTCTCAAACTTCTTTTCACCATTACAATCCACAATC
Strand6-Linear	TTGTTATGTGTGAGATTTTATTTTGATTGTGGATTGTAATGGTG
Strand6-Ring	TGAGATGTGATTGTGTTATGTTTTGATTGTGGATTGTAATGGTG
Strand7	TAAAATCTCACACATAACAATTTTCCACACACTACCCATCTTCA
Strand8-Linear	TGGATTTATGAAGTGAAGTATTTTTGAAGATGGGTAGTGTGTGG
Strand8-Ring	TGAGATGTGATTGTGTTATGTTTTTGAAGATGGGTAGTGTGTGG

Table S4: DNA sequences of key trigger strands used in fluorescence experiments.

Name	Sequence
Key4-Trigger	CTAAAAAATGAGATGTGATTGTGTTATG
Key6-Trigger	CCAAAAAATGAGATGTGATTGTGTTATG
Key8-Trigger	CTTCAAAATGAGATGTGATTGTGTTATG
Key4A-Trigger	AAAAAAGGGTTAATGATTGATAGGAG
Key4B-Trigger	CCTCTTACAATCAAAACTTCAAAATTA

Appendix T

SI FOR CHAPTER 3

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*

Supporting information

S1 Materials and methods

Sequence design

DNA sequences of the cooperative catalytic circuit were designed following the principles outlined in Supplementary Note S1.1 of the previous work on seesaw DNA circuits [23] and are listed here in Table S1.

To ensure that all double-stranded domains are stably bound in all initial, intermediate, and final species, we chose the length of domains S1, S2, and S3 to be 10 nucleotides: when T1, T2, T3, and T4 are 5 nucleotides or longer, the shortest double-stranded domains in the intermediate species will be at least 15 nucleotides. For example, these double-stranded domains include T3+S2 when the input is fully bound to the gate (XGY), S1+T3 when the activator is fully bound to the gate (GYA), and S3+T4 when the fuel is fully bound to the translator (FGZ).

Sequences of S1, S2, and S3 were drawn from a pool of 10-nt domains generated by the software developed for seesaw DNA circuits [23], and sequences of T1, T2, and T3 were drawn from a list of experimentally verified 7-nt toeholds designed for a DNA strand displacement oscillator [101]. In cases when T3 was shortened from 7 to 5 or 4 nucleotides, the desired number of nucleotides were simply removed from the 3’ end of the toehold sequence in the gate and fuel species while the translator was kept unchanged. Sequences of T4 and S4 were predetermined for compatibility with an existing reporter [41, 99].

To reduce spurious interactions between T2$*$ and S3 on the gate and promote availability of the open toehold, T2 was selected to contain `A` and `T` only while S3 was selected to contain `T` and `C` only. To reduce toeless leak reaction between the gate and fuel, S1 and S2 were selected to contain a `C` at their 3' and 5' ends, respectively; less breathing takes place at the terminus of a double helix ending in a `C`-`G` base pair rather than an `A`-`T` base pair.

When investigating the catalytic behavior of the activator, we used wobbles and mismatches to bias branch migration and favor the release of the activator for completing a catalytic cycle. Specifically, a wobble was introduced to the 8th nucleotide and mismatches to the 7th and 8th nucleotides on the 5' end of the S2 domain in the activator but not in the fuel. The same wobble and mismatches were also employed in the output initially bound to the gate (named output/gate-t in Table S1, where gate-t indicates gate top strand) so that branch migration in output production remains unbiased. Wobble or mismatch positions closer to the 5' end of S2 were not explored because they would reduce the stability of the double-stranded domain and encourage undesired leak reaction between the gate and fuel.

Similar to the clamp design in seesaw DNA circuits [23], a 2-nt clamp domain was embedded in the translator and reporter for inhibiting leak reaction between the two that could be initiated by a stacking bond between the ends of the two helices. Since the clamp domain is not present in the output strand, the translation reaction is effectively reversible, while the forward reaction should be at least 100 times faster than the backward reaction because of the length difference between the initiation and dissociation toeholds.

NUPACK [77] analysis was performed for all designed species to verify their equilibrium concentrations and minimum free energy (MFE) structures at 25 °C with 100 nM of each strand. In particular, we confirmed that intermediate species XGY and GYA should be sufficiently stable even when T3 is 4 nucleotides, and the gate should be sufficiently stable with wobbles and mismatches at the chosen positions.

Sample preparation

DNA oligonucleotide synthesis. All strands were chemically synthesized by Integrated DNA Technologies (IDT). Strands covalently bonded to a fluorophore or quencher were ordered HPLC-purified while unmodified strands were ordered with standard desalting only. Formulation service LabReady was selected so that all strands were delivered at ≈100 μM in IDTE, IDT's formulation of Tris-EDTA (TE) buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0). On arrival, concentration of each strand was verified on a NanoDrop (Thermo Fisher) by averaging three measurements of absorbance at 260 nm for a 1 μL droplet. DNA was stored at 4 °C.

Annealing protocol and buffer condition. The two strands in the gate and translator were mixed in a 1:1 ratio and annealed at 45 μM. The reporter was annealed at 20 μM with the top quencher strand and bottom fluorophore strand in a 6:5 ratio—the concentration of the reporter was determined by the concentration of the bottom strand; excess top strand was used here because it should not react with any other molecules in the circuit and it helps ensure that all copies of the fluorophore strand should be bound to a quencher strand. The buffer for all double-stranded complexes was TE with 12.5 mM Mg^{2+}. To anneal, the following protocol was run on a thermal cycler (Eppendorf): hold at 90 °C for 2 minutes, ramp down to 20 °C by 1 °C per 60 seconds.

Purification. Following annealing, polyacrylamide gel electrophoresis (PAGE) was used to purify gate and translator on a 12% gel. A single desired band for each complex was cut, diced, and incubated for at least 24 hours at room temperature (≈24 °C) in TE buffer with 12.5 mM Mg^{2+}, during which period DNA should diffuse out of the gel and into the buffer. Gel fragments were then discarded. Concentrations of purified complexes were measured on a NanoDrop (Thermo Fisher).

Fluorescence kinetics experiments

In each set of experiments, 110 μL of each sample was mixed at a standard concentration (1×) of 100 nM in a 96-well plate (Corning). Fluorescence levels were read every 2 minutes using a microplate reader (Cytation 5, Biotek) at 25°C. Excitation and emission wavelengths were set to 598 and 629 nm, respectively, for fluorophore ATTO590.

To reduce the loss of DNA to the surfaces of pipette tips and tubes, 2 μM of a 20-nt poly(`T`) strand [91] (referred to as 20`T` below) was added first to each sample, and LoRetention tips and DNA LoBind tubes (Eppendorf) were used in all experiments.

In each set of fluorescence kinetics experiments with varying input and activator concentrations, a master mix was made with $20\times$ 20T, $1\times$ gate, $2\times$ fuel, $1.5\times$ translator, and $1.5\times$ reporter. The master mix was transferred to a number of wells on a plate, and distinct amounts of input and activator were then added to each of the wells—compared to preparing each sample separately, the master mix allows for improved consistency across experiments.

Data normalization

Each set of fluorescence kinetics experiments included a negative control with $0\times$ input and $0\times$ activator as well as a positive control with $1\times$ input and $1\times$ activator. The minimum raw fluorescence was determined by the averaging the initial five measurements of the negative control, and the maximum raw fluorescence was determined by averaging the final five measurements of the positive control. The minimum and maximum raw fluorescence were then used as $0\times$ and $1\times$, respectively, to convert fluorescence data to concentration data. In this data normalization method, 0 nM $(0\times)$ is interpreted as the background fluorescence of the reaction mixture before any signal has been detected, and 100 nM $(1\times)$ is interpreted as the highest fluorescence of the reaction mixture after the output strand Y has been fully released from the gate, translated to signal Z, and detected by the reporter.

S2 Concept of three types of catalysts

In theory, DNA strand displacement (DSD) systems are capable of implementing arbitrary chemical kinetics [110], including any desired type of catalyst. The underlying principle is that for any given chemical reaction that involves a set of signal species (e.g. X and Y shown in Figure S1), auxiliary species (e.g. GY and F) can be designed to facilitate the desired interactions between signals. To achieve systematic implementation of chemical reaction networks (CRNs) where each signal species can participate in multiple distinct reactions, two criteria are necessary. First, there should be no sequence dependence across signals. For example, if the sequence of Y depends on the sequence of X, then $X1 \rightarrow X1 + Y$ and $X2 \rightarrow X2 + Y$ could result in a conflict where no sequence of Y satisfying both reactions exists. On the other hand, the sequences of auxiliary species do depend on the signal species—for example, GY designed to facilitate the production of Y when X is present must incorporate two domains of sequences that are determined by the sequences of X and Y, respectively. The sequence dependence between signal and auxiliary species does not affect the composability of individual reactions, because auxiliary species designed for one reaction should not participate in any other reactions and thus can be viewed as internal to each reaction. Second, all signals should have the same format (e.g. a toehold followed by a branch migration domain). If reactant X and product Y have different formats in $X \rightarrow X + Y$, then $Y \rightarrow Y + Z$ cannot use the same implementation scheme due to a conflict of Y having the format of a product in one reaction and that of a reactant in the other. Note that product species commonly contain history domains that do not participate in downstream reactions—in chemical reaction networks that consist of irreversible reactions (such as $X \rightarrow X + Y$), these history domains do not need to be considered for comparing the format of signal species.

Various CRN-to-DSD implementation schemes have been proposed [110, 161, 162], yet successful experimental demonstrations have so far been limited to relatively simple systems involving up to three non-catalytic bimolecular reactions [101, 111]. If behaviors as general as arbitrary chemical kinetics or Turing-universal computation are not required, DNA strand displacement implementations of catalysts alone can be scalable. Systems with dozens of catalytic reactions in the form of $X \rightarrow X + Y$ have been demonstrated for performing digital logic and neural network computation [23, 41]. Previous work has shown that the scalability of DNA strand displacement systems depends on the simplicity of the motifs—when only one and two-stranded molecules were involved in synthesizing a system (at its initial state before any input signal arrives), the system behavior was robust to oligonucleotide synthesis errors and potential malformation of annealed structures [99]. Thus we are motivated to understand whether catalysts with enhanced functionality can be implemented with simple DNA strand displacement motifs where each initial molecule consists of no more than two strands [163].

First, $X \to X + Y$ can be extended with allosteric control $X + A \to X + Y$, where output Y is only produced when a consumable activator A is present. DNA strand displacement implementation for this type of catalyst has been developed using two auxiliary species including a two-stranded gate and a single-stranded fuel (named GY and F in Figure S1) [95]. In that implementation, the activator strand is consumed in forming a two-stranded product AY, which is functionally equivalent to Y because the activator strand only covers up a portion of the output strand not involved in downstream reactions. The input X and activator A have sequence dependence (a toehold complementary to each other) and different formats (a toehold followed by a branch migration domain vs. two concatenated toeholds), both of which pose some challenges for the implementation to be used in larger systems.

Next, $X \to X + Y$ can be extended from unimolecular to bimolecular $X + A \to X + A + Y$, where in addition to catalyst X, a second catalyst A is required for the production of Y. This second catalyst can be viewed as an activator that provides allosteric control for $X \to X + Y$ without being consumed itself. More generally, both X and A are signal species that cooperatively and catalytically produce Y. In this work, we develop a DNA strand displacement implementation for the cooperative catalyst. As simple as the previous examples of basic DNA catalyst and allosteric DNA catalyst discussed above, the cooperative DNA catalyst also uses two auxiliary species including a two-stranded gate and a single-stranded fuel. Unlike the allosteric DNA catalyst discussed above and other more complex variations previously developed [108, 109], the cooperative DNA catalyst that we show here requires no sequence dependence between two input signals X and A. With a simple, two-stranded translator, output Y can be converted to the same format as either input with independent sequence (Figures S5b and S5c).

The difference in the abstract chemical reactions of the three types of catalyst determines their distinct usages in various scenarios. For example, when auxiliary species are in large excess, their concentrations can be treated as roughly constant. In this scenario, a constant rate of output production can be controlled by the concentration(s) of the catalyst(s). For the allosteric catalyst with a consumable activator, the system behavior is similar to the basic catalyst if the activator is in large excess, otherwise the rate of output production will decrease as the activator is used up. For the cooperative catalyst, the rate of output production scales linearly with both catalysts, which could be used for computing the multiplication of two real numbers when combined with a degradation reaction [164]. Alternatively, in a second scenario, the concentrations of auxiliary species can be used to control the completion level of output production, while the presence or absence of the catalyst(s) are used to turn ON or OFF output production. The concentration of a consumable activator in the allosteric catalyst can also be used to determine output completion, so long as it is less than the concentrations of auxiliary species—this behavior is useful for learning an analog weight in chemical neural networks [96]. By contrast, with the cooperative catalyst, the output completion solely depends on the concentrations of auxiliary species—this behavior is useful for learning a binary weight in chemical neural networks [165]. While the simulations of these two scenarios shown in Figure S1 provide a characteristic understanding for certain example behaviors, the exact kinetics for each type of catalyst is a tunable variable. As the rate of DNA strand displacement reactions can be well controlled by the length and sequence of toeholds [78, 79], desired kinetics can be tailored for different usages of the catalysts by altering the toehold designs.

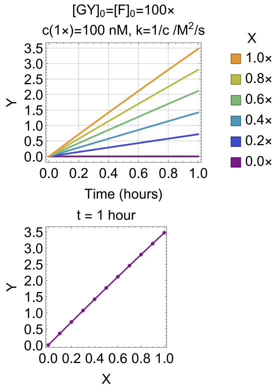
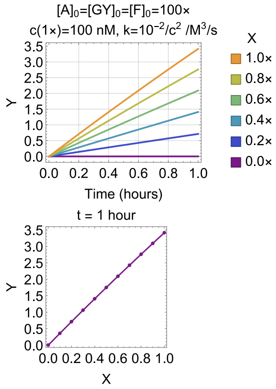
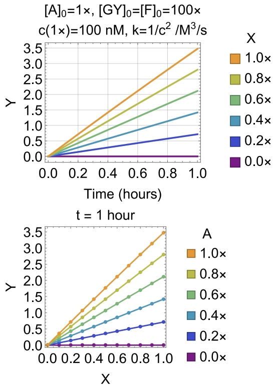
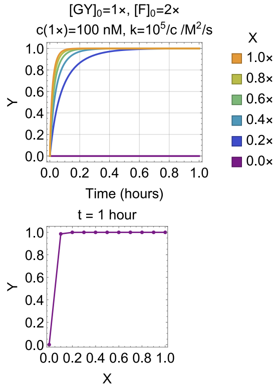
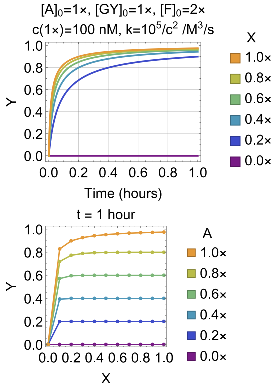
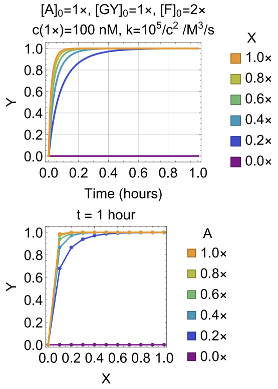
	Basic catalyst	**Allosteric catalyst**	**Cooperative catalyst**
Abstract chemical reaction	$X \rightarrow X + Y$	$X + A \rightarrow X + Y$	$X + A \rightarrow X + A + Y$
DNA strand-displacement implementation	$X + GY \rightleftharpoons XG + Y$ $XG + F \rightleftharpoons X + GF$	$A + GY \rightleftharpoons AGY$ $X + AGY \rightleftharpoons XG + AY$ $XG + F \rightleftharpoons X + GF$	$X + A + GY \rightleftharpoons XGA + Y$ $XGA + F \rightleftharpoons X + A + GF$
Simplified overall reaction	$X + GY + F \xrightarrow{k} X + Y$	$X + A + GY + F \xrightarrow{k} X + Y$	$X + A + GY + F \xrightarrow{k} X + A + Y$
Steady-state output concentration	$[Y]_\infty = \min([GY]_0, [F]_0)$ if $[X]_0 > 0$	$[Y]_\infty = \min([A]_0, [GY]_0, [F]_0)$ if $[X]_0 > 0$	$[Y]_\infty = \min([GY]_0, [F]_0)$ if $[X]_0, [A]_0 > 0$
Example scenario 1 with controlled constant rate of output production	$[GY]_t, [F]_t \approx$ constant $[Y]_t = \alpha [X]_0 \cdot t$ 	$[A]_t, [GY]_t, [F]_t \approx$ constant $[Y]_t = \alpha [X]_0 \cdot t$ 	$[GY]_t, [F]_t \approx$ constant $[Y]_t = \alpha [X]_0 [A]_0 \cdot t$ 
Example scenario 2 with controlled completion level of output production	$[F]_0 \geq [GY]_0$ $[Y]_\infty = [GY]_0$ if $[X]_0 > 0$ 	$[F]_0 \geq [GY]_0 \geq [A]_0$ $[Y]_\infty = [A]_0$ if $[X]_0 > 0$ 	$[F]_0 \geq [GY]_0$ $[Y]_\infty = [GY]_0$ if $[X]_0, [A]_0 > 0$ 

Figure S1: **Concept of three types of catalysts.** In DNA strand displacement implementations, signal species are colored in black, auxiliary species are colored in gray, and intermediate and waste products are colored in blue. Characteristic simulations for example scenarios were performed using the simplified overall reaction.

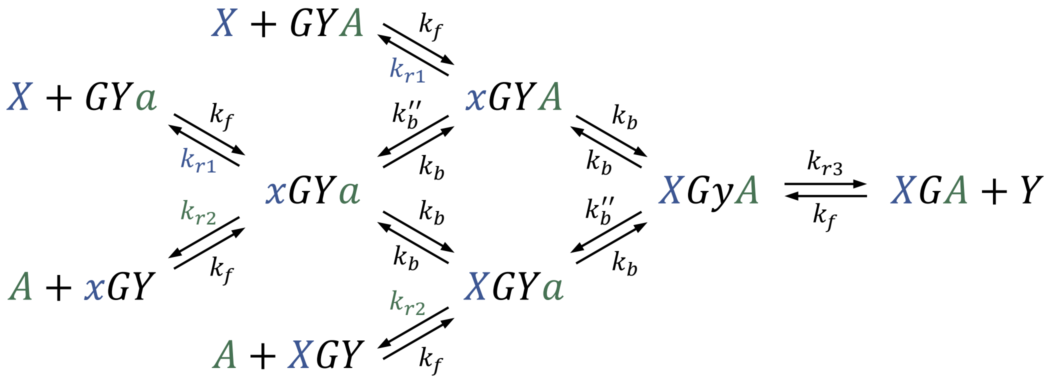
S3 Modeling and simulation of cooperative catalyst

All simulations were performed with mass-action kinetics using CRNSimulator [22].

The following reactions were used to model reversible cooperative hybridization $X + A + GY \rightleftharpoons XGA + Y$, where k_b'' was introduced for understanding the effect of wobble and mismatch:

$$X + GY \underset{k_{r1}}{\overset{k_f}{\rightleftharpoons}} xGY \underset{k_b}{\overset{k_b}{\rightleftharpoons}} XGY$$

$$A + GY \underset{k_{r2}}{\overset{k_f}{\rightleftharpoons}} GYa \underset{k_b''}{\overset{k_b}{\rightleftharpoons}} GYA$$



$X + A + GF \rightleftharpoons XGA + F$ was modeled similarly, where k_b' was introduced for understanding the effect of wobble and mismatch:

$$X + GF \underset{k_{r1}}{\overset{k_f}{\rightleftharpoons}} xGF \underset{k_b}{\overset{k_b}{\rightleftharpoons}} XGF$$

$$A + GF \underset{k_{r2}}{\overset{k_f}{\rightleftharpoons}} GFa \underset{k_b}{\overset{k'_b}{\rightleftharpoons}} GFA$$

$$\begin{tikzcd}
 & X + GFA \arrow[dr, shift left, "k_f"] & & & \\
X + GFa \arrow[dr, shift left, "k_f"] & & xGFA \arrow[ul, shift left, "k_{r1}"] \arrow[dl, shift right, "k_b"'] \arrow[dr, shift left, "k_b"] & & \\
 & xGFa \arrow[ul, shift left, "k_{r1}"] \arrow[dl, shift left, "k_{r2}"'] \arrow[ur, shift right, "k'_b"'] \arrow[dr, shift left, "k_b"] & & XGfA \arrow[ul, shift left, "k_b"] \arrow[dl, shift left, "k_b"'] \arrow[r, shift left, "k_{r3}"] & XGA + F \arrow[l, shift left, "k_f"] \\
A + xGF \arrow[ur, shift left, "k_f"'] & & XGFa \arrow[ul, shift left, "k_b"] \arrow[ur, shift right, "k'_b"'] \arrow[dl, shift right, "k_{r2}"'] & & \\
 & A + XGF \arrow[ur, shift right, "k_f"'] & & &
\end{tikzcd}$$

Output Y reacting with translator GZ was modeled as follows, where k'_{r3} is dissociation rate of T3 binding to the translator and k_{rc} is dissociation rate of a 2-nt clamp:

$$Y + GZ \underset{k'_{r3}}{\overset{k_f}{\rightleftharpoons}} yGZ \underset{k_b}{\overset{k''_b}{\rightleftharpoons}} YGZ \underset{k_f}{\overset{k_{rc}}{\rightleftharpoons}} YG + Z$$

Fuel F can also react with translator GZ but the reaction is reversible without producing signal Z:

$$F + GZ \underset{k'_{r3}}{\overset{k_f}{\rightleftharpoons}} fGZ \underset{k_b}{\overset{k_b}{\rightleftharpoons}} FGZ$$

Signal Z reacting with reporter R was modeled as follows, where k_s is effective rate of an irreversible strand displacement reaction with a 5-nt toehold:

$$Z + R \xrightarrow{k_s} O$$

The following rate constants were used in all simulations:

$k_f = 2 \times 10^6$ /M/s

$k_{r1} = 0.1$ /s

$k_{r2} = 0.2$ /s

$k_{r3} = 0.001$ /s, 0.015 /s, and 0.03 /s for T3 = 7, 5, and 4 nucleotides, respectively.

$k'_{r3} = 0.33$ /s, 5 /s, and 10 /s for T3 = 7, 5, and 4 nucleotides, respectively.

$k_b = 1$ /s

$k'_b = 1$ /s and 0.2 /s for S2 domain without and with a wobble, respectively.

$k''_b = 1$ /s and 0.02 /s for S2 domain without and with a wobble, respectively.

$k_{rc} = 10^4$ /s

$k_s = 10^5$ /M/s

k_{r1} (the dissociation rate of toehold T1) is smaller than k_{r2} (the dissociation rate of toehold T2), because the observed kinetics with $1\times$ input and $0.1\times$ activator (Figure 3.5b) was slightly slower than that with $0.1\times$ input and $1\times$ activator (Figure 3.4d). Presumably, this could be due to the sequence difference in toeholds T1 and T2 or the structural asymmetry of the gate:output molecule—the single-stranded S3 and T4 domains could partially inhibit the open toehold T2$*$ through spurious binding.

Understanding the biophysics and kinetics of reversible cooperative hybridization merits future study. The model in this work does not fully explain the behavior of the system. For example, a 30-fold difference in dissociation rate k_{r3} was estimated comparing a 7-nt and 4-nt toehold, which is significantly less than the 1000-fold difference suggested in previous studies [78, 79]. Moreover, taking the additional stacking bond into consideration as roughly 1-nt energy, the 4-nt toehold dissociation rate (0.03 /s) is still much smaller than expected (10 /s). A more detailed model at the base-pair level would be particularly useful for a better understanding of the effect of a wobble [107].

S4 Sequence-level design diagrams

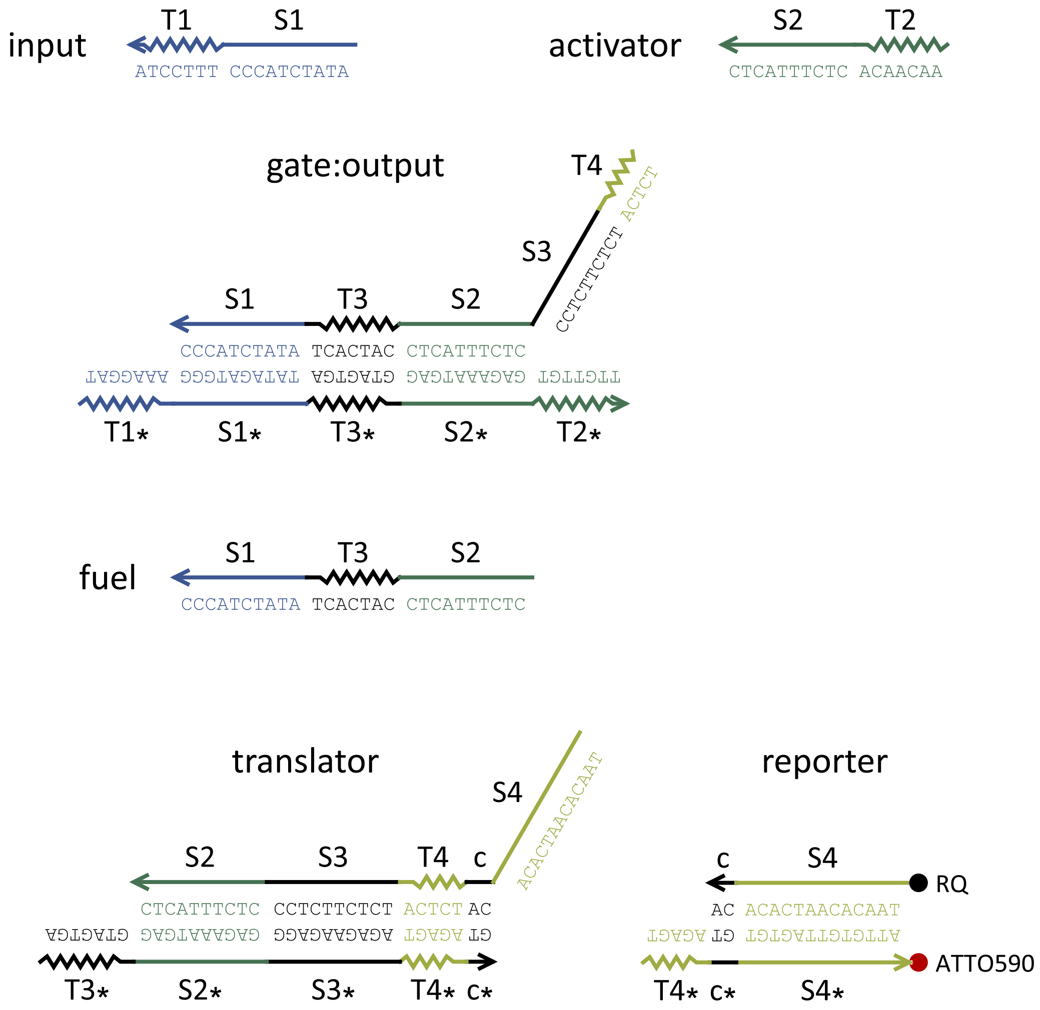


Figure S2: **Sequence-level diagrams.** A 2-nt clamp domain is labeled as c. A 7-nt toehold T3 and branch migration domain S2 without any wobbles or mismatches are shown here. Sequences for varying lengths of toehold T3 and branch migration domain S2 with a wobble or mismatch are listed in Table S1.

S5 Effect of a wobble or mismatch for promoting activator recycling

Previous studies have shown that a mismatch (i.e. non-Watson–Crick base pairs) in the branch migration domain of the invader strand has an effect of slowing down strand displacement, the kinetics of which depends on the position and nucleotide sequence of the mismatch for a given toehold [105–107]. This effect is useful for improving the performance of DNA catalysts. For example, leak between a gate complex and a fuel strand can be reduced by introducing mismatch modifications in the fuel strand [123]. Desired output production can be sped up by the introduction of mismatch modifications in the input and output strands, which favors the process of fuel releasing the input [103].

Here we investigate whether the above principle of mismatches can be applied to improve the performance of the cooperative catalyst. Specifically, we aim to use a G·T wobble [102], which is known to be more stable than other types of mismatches [104], to promote activator recycling without increasing leak.

Unlike the experiments shown in Figure 3.5c, unpurified gate:output and translator complexes were used in the following experiments that we initially performed for evaluating the effect of wobbles compared with mismatches. Without gel purification, it is possible that the double-stranded complexes contained excess single strands due to stoichiometry errors or otherwise contained a small fraction of malformed structures. This impurity led to a type of behavior similar to thresholding, where the output production appeared much slower compared to experiments with purified complexes, as if the activator was at a much lower concentration than expected. For example, 0.001× activator was used in simulations (Figure S3d) in order to roughly agree with the experiments with 0.1× activator (Figure S3c). Thus, the impurity allowed us to evaluate the robustness of the activator promoted by a wobble.

We introduced a wobble modification (i.e. changing a C to a T) at varying positions within the S2 domain in the activator and output strands, while keeping the fuel perfectly complementary to the gate strand (Figure S3a). When the activator but not the output strand was modified, the output production became slower (Figure S3b), indicating that the biased branch migration in output release played a more significant role than the biased branch migration in activator recycling (i.e. activator release by fuel). However, when the activator and output strands were both modified, the output production became faster, especially for wobble modification at position 8 (Figure S3b). In this case, only two out of the four possible positions were explored experimentally because more fraying is expected in the gate:output complex for modification at positions 1 and 3 in the output strand, which will likely result in increased leak between the complex and fuel.

We then compared the wobble modification at position 8 with a mismatch modification at the same and an adjacent position (Figure S3c). Interestingly, the output production was faster with the wobble than the mismatches (Figure S3c, left plot), while the leak was slower (Figure S3c, right plot).

To better understand the observed system behavior, we used simulations to estimate the bias in branch migration rates affected by the wobble at position 8 (Figure S4). We found that a bias in branch migration involving the activator and fuel alone cannot explain the faster output production in experiments (Figure S4a). It is necessary that a bias also exists in branch migration involving the activator and output (Figure S4b). This bias is reasonable because position 8 is near the end of branch migration for the activator and near the beginning of branch migration for the output, while branch migration is expected to slow down more significantly for non-Watson-Crick base pairs near the beginning [106, 107]. Because the bias in branch migration involving the activator and fuel still has some effect on the overall rate of output production, it is desirable that the wobble is not too close to the end of branch migration for the activator (e.g. position 10). Moreover, simulations suggested that a bias in branch migration involving the input and output, together with the above biases, could further speed up output production (Figure S4c)—this would be useful for future developments such as an activator-producing threshold discussed in SI Note S7.

While a wobble in branch migration domain S2 resulted in faster kinetics when the activator concentration was low (Figure S3c), it resulted in slower overall kinetics when the activator concentration was relatively high (Figure 3.5c). This is because the branch migration rate of the output reacting with the translator also slowed down with the wobble (SI Note S3, Modeling and simulation of cooperative catalyst). This slowdown in the translation reaction explains why too much increase in k_b/k''_b decreases the overall signal production in simulations (Figure S4). If needed, the translator can be redesigned with $S2'$ and $S2'*$ domain sequences in the top and bottom strands, respectively, to match the modified domain sequence S2′ in the output strand.

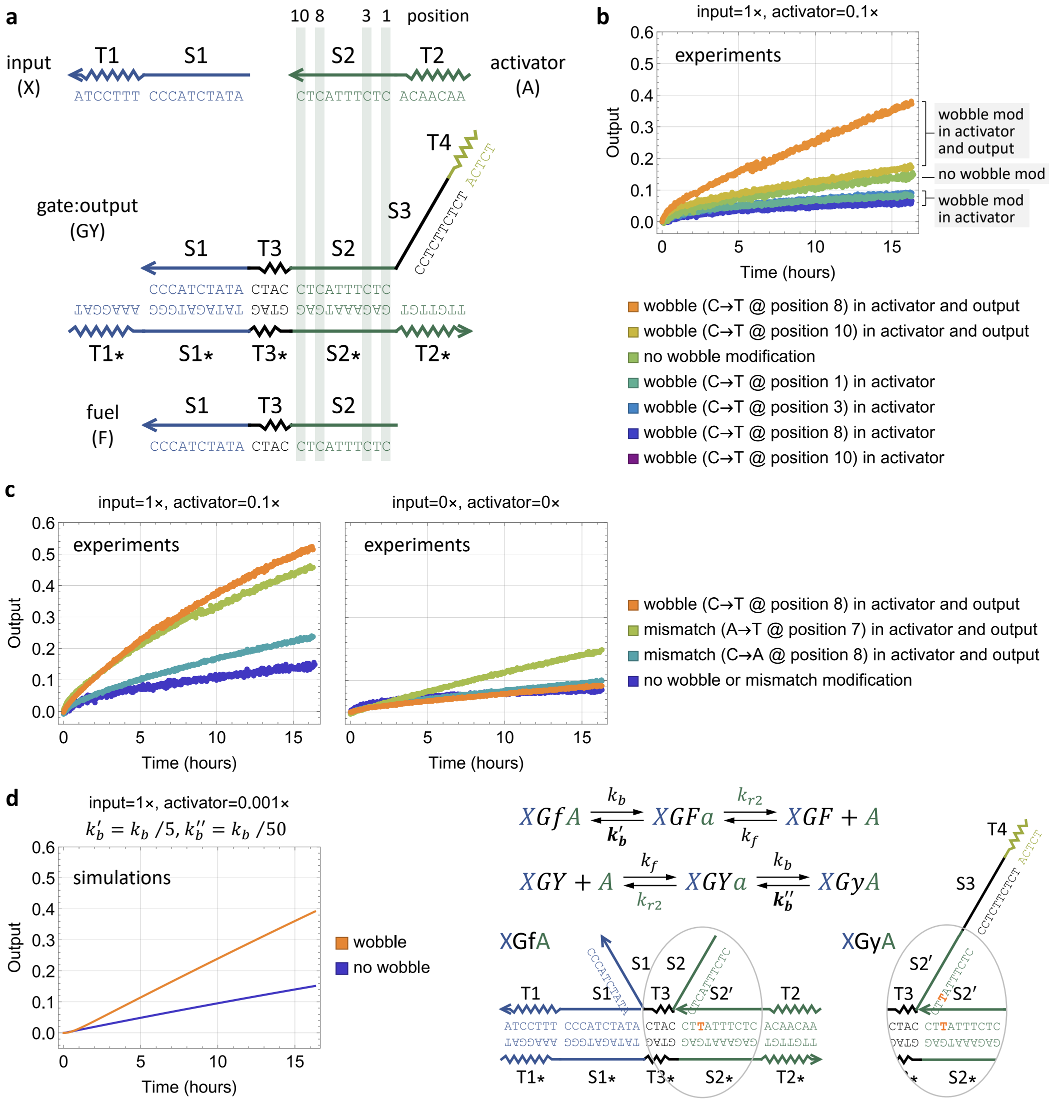
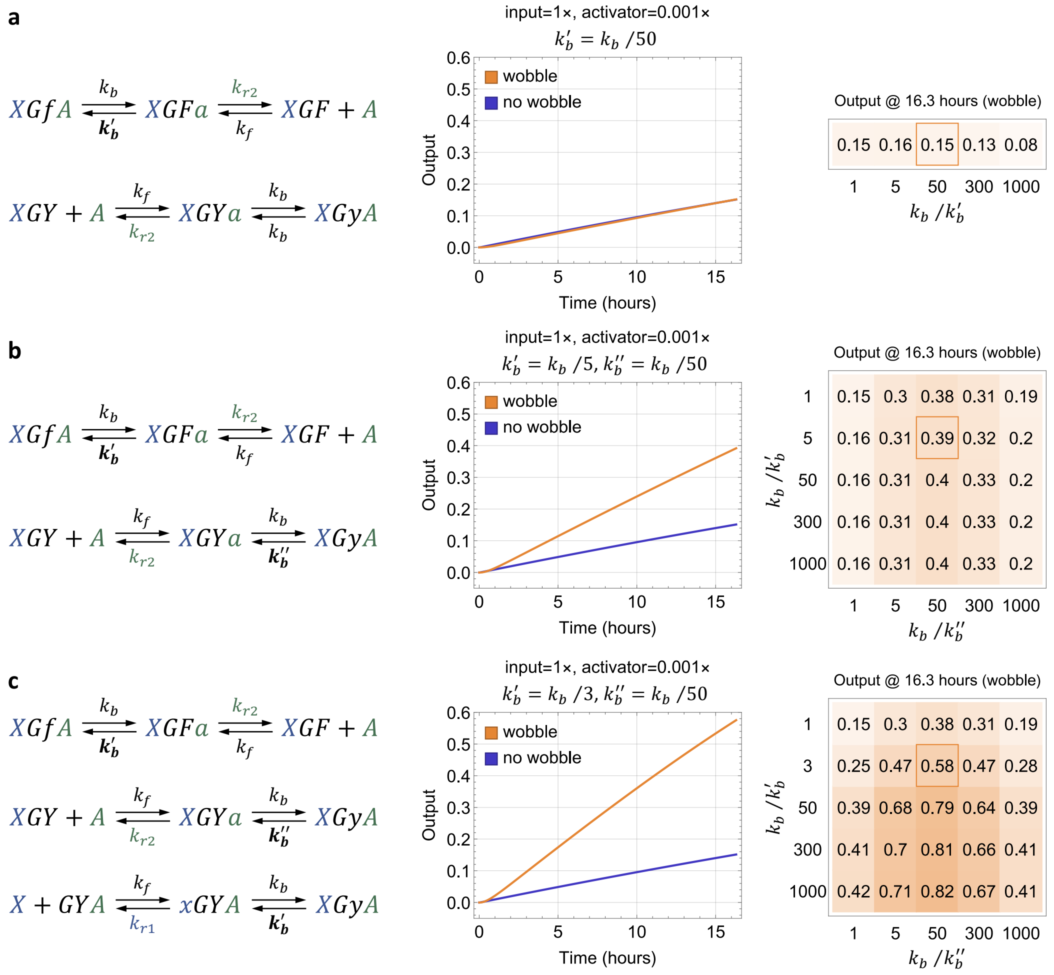


Figure S3: **Robustness of the activator with unpurified gates. (a)** Sequence-level diagrams with labeled nucleotide positions in the S2 domain. **(b, c)** Fluorescence kinetics data for the cooperative catalyst without or with a (b) wobble or (c) mismatch modification in the activator and output strands. **(d)** Simulations without or with modified branch migration rate constants attributed to a wobble. Standard concentration $1\times = 100$ nM. Initial concentrations of gate:output, fuel, translator, and reporter are $1\times$, $2\times$, $1.5\times$, and $1.5\times$, respectively. Output is shown as a relative concentration to $1\times$. The experiment for wobble modification at position 8 in activator and output shown in b and c are two repeats of the same experiment performed on different days; the difference in kinetics is likely due to experimental noise.



Output @ 16.3 hours (wobble)

k_b/k'_b	1	5	50	300	1000
	0.15	0.16	0.15	0.13	0.08

Output @ 16.3 hours (wobble)

k_b/k'_b \ k_b/k''_b	1	5	50	300	1000
1	0.15	0.3	0.38	0.31	0.19
5	0.16	0.31	0.39	0.32	0.2
50	0.16	0.31	0.4	0.33	0.2
300	0.16	0.31	0.4	0.33	0.2
1000	0.16	0.31	0.4	0.33	0.2

Output @ 16.3 hours (wobble)

k_b/k'_b \ k_b/k''_b	1	5	50	300	1000
1	0.15	0.3	0.38	0.31	0.19
3	0.25	0.47	0.58	0.47	0.28
50	0.39	0.68	0.79	0.64	0.39
300	0.41	0.7	0.81	0.66	0.41
1000	0.42	0.71	0.82	0.67	0.41

Figure S4: **Simulation analysis of the effect of wobble. (a)** Simulations with a bias in branch migration involving the activator and fuel (first reaction, k_b: fuel displacing activator, k'_b: activator displacing fuel, $k_b \geq k'_b$). **(b)** Simulations with the above bias and a bias in branch migration involving the activator and output (second reaction, k_b: activator displacing output, k''_b: output displacing activator, $k_b \geq k''_b$). **(c)** Simulations with the above biases and a bias in branch migration involving the input and output (third reaction, k_b: input displacing output, k'_b: output displacing input, $k_b \geq k'_b$). Representative reactions are shown here, and the full list of reactions involving the modified branch migration rates are shown in SI Note S3. Standard concentration $1\times = 100$ nM. Initial concentrations of gate:output, fuel, translator, and reporter are $1\times$, $2\times$, $1.5\times$, and $1.5\times$, respectively. Output is shown as a relative concentration to $1\times$. Kinetics trajectories for one example ratio of branch migration rates are shown in the plot, while the relative concentration of the output at 16.3 hours (time of the last data point in the kinetics plot) is shown in an array with varying ratios of branch migration rates. The orange box in the array highlights the example shown in the kinetics plot.

S6 An AND gate with near-perfect signal restoration

The cooperative catalyst could be used to build a better AND gate. It is important to embed signal restoration within DNA circuits for cleaning up noise and resolving signal decay. An implementation of DNA logic gates has been proposed with the aim of processing input signals with substantial noise (OFF $= 0 - 0.4\times$ and ON $= 0.6 - 1\times$) [97], but the complexity of the design has inhibited successful experimental demonstration (a two-input AND gate requires 5 gates, 5 thresholds, and 2 fuel strands). With the cooperative catalyst, the desired logic function could be implemented with a much simpler design (1 gate, 2 thresholds, and 1 fuel strand as shown in Figure S5a). Because the two input strands in the cooperative catalyst have independent domain sequences, two distinct threshold complexes can be designed without any concern of threshold crosstalk [23].

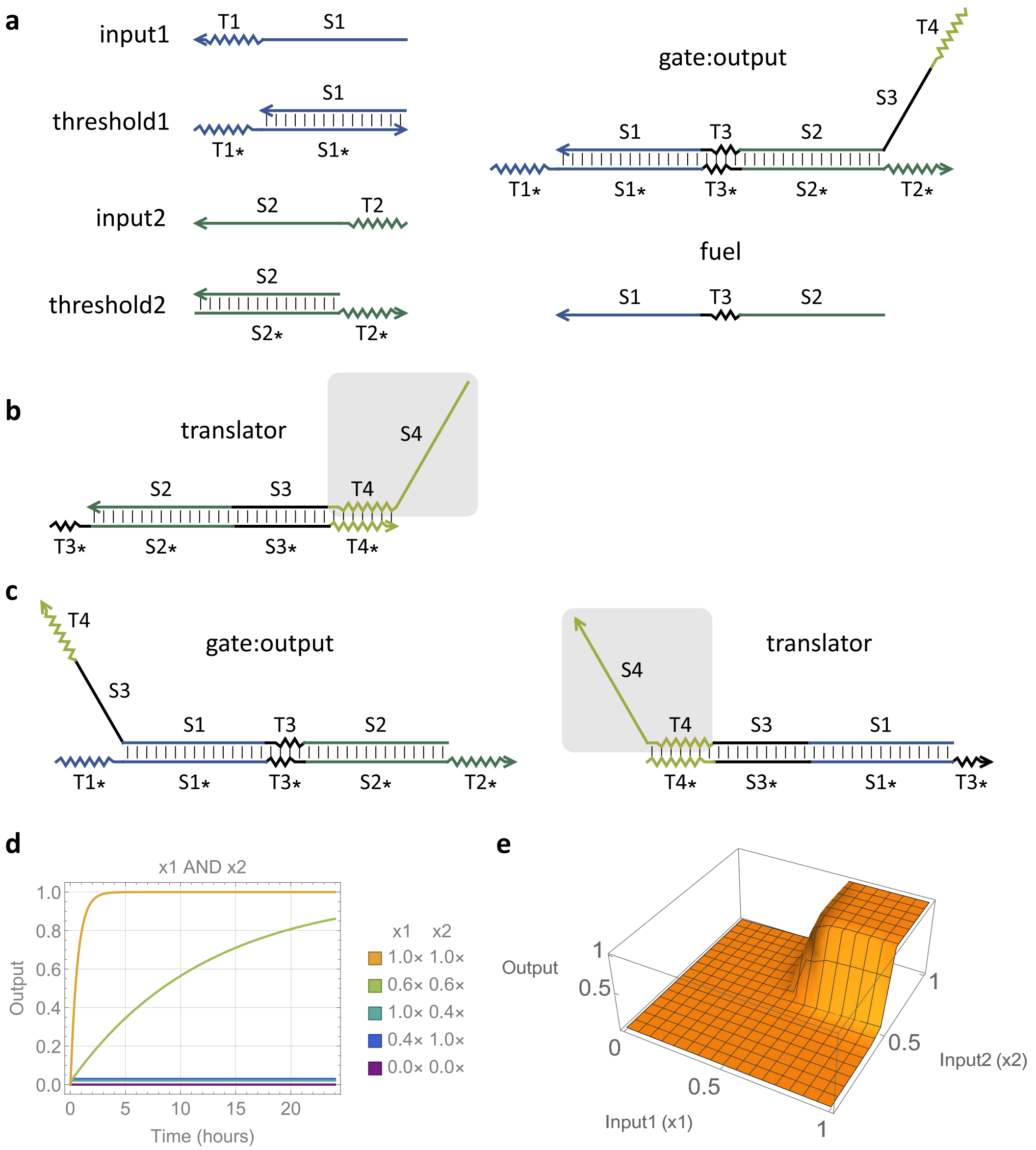


Figure S5: **An AND gate with near-perfect signal restoration. (a, b)** Design diagrams of (a) an AND gate and (b) a translator producing a signal strand that has the same format as input1. **(c)** An alternative gate and translator design for producing a signal with the same format as input2. One of the two designs can be used for each AND gate in a logic circuit, depending on which input format its downstream gate requires. **(d)** Simulated kinetics of the AND gate with five example input combinations. **(e)** Simulated input-output relationship at $t = 24$ hours. Standard concentration $1\times = 100$ nM.

S7 An activator-producing threshold

The cooperative catalyst could be used to develop a better threshold mechanism that combines the advantages of sequential [15] and competitive [23] thresholding. In sequential thresholding ($X+Th \rightarrow GY$ and $X+GY+F \rightarrow Y$), the threshold (Th) concentration must be the same as the gate (GY) concentration, which introduces an undesired constraint on how much noise can be suppressed in the input signal versus how much output signal can be produced. In competitive thresholding ($X+Th \xrightarrow{k_f} \emptyset$ and $X+GY+F \xrightarrow{k_s} X+Y$, $k_f \gg k_s$), there are no constraints between the threshold and gate concentrations, but the thresholding reaction must be much faster than the catalyst. With the cooperative catalyst, thresholding and signal amplification can be implemented in two sequential steps: $X+Th \rightarrow A$ and $X+A+GY+F \rightarrow X+A+Y$. This implementation (Figure S6a) neither requires a significant rate difference nor does there exist any dependence between threshold and output concentrations. The kinetics of the cooperative catalyst simulated here (Figure S6c) is not as fast as it could be. With further tuning of toehold sequences, possibly introducing wobbles or mismatches in both S1 and S2 domains, and using a faster translator, much faster signal restoration could potentially be achieved.

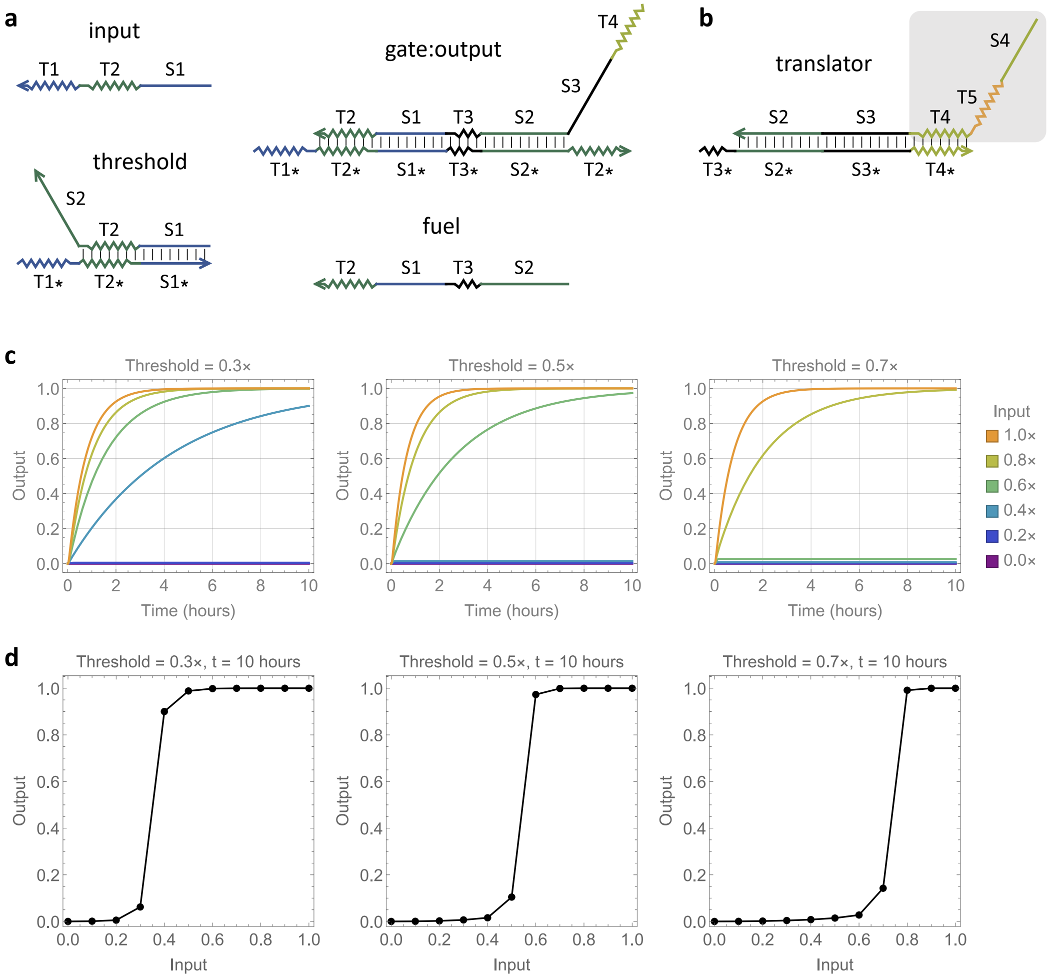


Figure S6: **An activator-producing threshold. (a, b)** Design diagrams of (a) a signal restoration circuit using an activator-producing threshold, and (b) a translator producing a signal strand that has the same format as the input. **(c)** Simulated kinetics of the signal restoration circuit with three example threshold and six example input values. **(d)** Simulated input-output relationship at $t = 10$ hours. Standard concentration $1\times = 100$ nM.

S8 DNA sequences

Table S1: DNA sequences of the cooperative catalyst. All sequences are listed from 5’ to 3’.

Name	Sequence
output/gate-t (7ntT3)	`TCTCA TCTCTTCTCC CTCTTTACTC CATCACT ATATCTACCC`
output/gate-t (5ntT3)	`TCTCA TCTCTTCTCC CTCTTTACTC CATCA ATATCTACCC`
output/gate-t (4ntT3)	`TCTCA TCTCTTCTCC CTCTTTACTC CATC ATATCTACCC`
output/gate-t (4ntT3 and wobble10)	`TCTCA TCTCTTCTCC CTCTTTACTT CATC ATATCTACCC`
output/gate-t (4ntT3 and wobble8)	`TCTCA TCTCTTCTCC CTCTTTATTC CATC ATATCTACCC`
output/gate-t (4ntT3 and mismatch8)	`TCTCA TCTCTTCTCC CTCTTTAATC CATC ATATCTACCC`
output/gate-t (4ntT3 and mismatch7)	`TCTCA TCTCTTCTCC CTCTTTTCTC CATC ATATCTACCC`
gate-b (7ntT3)	`TAGGAAA GGGTAGATAT AGTGATG GAGTAAAGAG TGTTGTT`
gate-b (5ntT3)	`TAGGAAA GGGTAGATAT TGATG GAGTAAAGAG TGTTGTT`
gate-b (4ntT3)	`TAGGAAA GGGTAGATAT GATG GAGTAAAGAG TGTTGTT`
input	`ATATCTACCC TTTCCTA`
activator	`AACAACA CTCTTTACTC`
activator (wobble10)	`AACAACA CTCTTTACTT`
activator (wobble8)	`AACAACA CTCTTTATTC`
activator (wobble3)	`AACAACA CTTTTTACTC`
activator (wobble1)	`AACAACA TTCTTTACTC`
activator (mismatch8)	`AACAACA CTCTTTAATC`
activator (mismatch7)	`AACAACA CTCTTTTCTC`
fuel (7ntT3)	`CTCTTTACTC CATCACT ATATCTACCC`
fuel (5ntT3)	`CTCTTTACTC CATCA ATATCTACCC`

fuel (4ntT3)	`CTCTTTACTC CATC ATATCTACCC`
signal/translator-t	`CATAACACAATCACA TCTCA` `TCTCTTCTCC CTCTTTACTC`
translator-b	`AGTGATG GAGTAAAGAG` `GGAGAAGAGA TGAGA TG`
reporter-t	`/5IAbRQ/ CATAACACAATCACA`
reporter-b	`TG AGA TGTGATTGTGTTATG` `/3ATTO590N/`

Appendix U

LIST OF SUPPLEMENTARY FILES

1. Spreadsheet of DNA Sequences
2. OxDNA Simulation of Double-Layer Origami Dynamics (GIFs)
3. Designs for Three-Letter Code Scaffolds
4. Scadnano and Cadnano Files for Partial Design of a 14-Helix Tube Origami Structure

INDEX

ENDNOTES

1 Yes, even the ways that are not related to a pursuit of a more beautiful world.

2 After hours of pipetting, I often develop a tension headache that is best resolved with a neck and shoulder massage. I keep a massage hook in my office for this purpose.

3 Sequences for toeholds and branch migration domains may be reused in different molecular components, but each unique molecular component must at least have a unique ordered set of domains.

4 The reachability of a pair of molecules is generally determined by the positions of their tether points on the origami structure and the length of all domains that connect their tether point to their reactive elements. The relatively rigid double-layer origami structure was originally designed to ensure that cargo-sorting robots could only reach neighboring tracks [125]; on a flexible single-layer origami structure, it is not possible to reliably constrain reachability.

5 The loop size and shape changes between different states. When the gate is bound to either the output or the fuel, the inter-tether distance is ≈10.4 nm and there is only one 13 nt ssDNA tether. This loop is relatively tight and constrained. There is an entropic advantage when the input is bound to the gate strand. In that case, the inter-tether distance is only 6 nm, there are two 13 nt ssDNA tethers, and there is even an extra branch migration domain forming part of the total loop length. This loop is relatively loose and flexible.

6 The first time I drew the concept for the latch on Lulu's office whiteboard is visually recorded in my memory.

7 I count myself among the readership because I will have to refer to this thesis once I have forgotten precisely what was written here.