Classifying enthalpy-neutral DNA strand displacement cascades: trading off desired and undesired signal propagation

Boya Wang[a], **Chris Thachuk**[b], **and David Soloveichik**[a,1]

[a]Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712, USA; [b]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, 98195

Artificially designed molecular systems made of DNA can exhibit complex and programmable behaviors. One kind of building block, the enthalpy-neutral DNA strand displacement cascade, propagates signal without gaining or losing overall base pairs. This flexible and programmable component has been used for constructing molecular logic circuits, smart structures and devices, for systems with complex autonomously generated dynamics, and for diagnostics. Limiting their utility, however, these systems are susceptible to the spurious release of output even in the absence of the proper combination of inputs—so-called leak. In addition, the speed of these systems is usually affected by undesired reaction pathways such as reversible unproductive binding (toehold occlusion) or spurious displacement. We systematize the properties of the simplest enthalpy-neutral strand displacement cascades (logically linear topology). We use combinatorial arguments to show that enthalpy-neutral linear cascades can be engineered with stronger thermodynamic guarantees to leak than previously studied systematic ways to reduce leak in non-enthalpy-neutral designs. Depending on a few basic parameters, we develop a taxonomy for the desired and undesired properties and trade-offs between them. Further, we experimentally confirm the theoretical prediction of the properties on different designs. These results provide an effective guide for future designs of molecular systems.

molecular programming | DNA nanotechnology | strand displacement | leak

Nature evolved sophisticated molecular systems with diverse functional behaviors, ranging from simple ON/OFF chemical switches to vast gene regulatory networks. Inspired by nature, artificial molecular systems have shown great promise in approaching this level of complexity. Guided by ideas from computer science and electrical engineering, molecular systems engineered from first principles are finding applications in chemistry, material sciences and medicine. Toehold-mediated DNA strand displacement is a powerful mechanism for molecular information processing and dynamics (2, 3). A strand displacement reaction is the process where two strands hybridize with each other and displace a pre-hybridized strand. This process initiates when the invading strand binds to a *toehold*—an unbound single-stranded region (usually 5 to 8 nt) of the pre-hybridized complex. The displaced strand could then serve, in turn, as the displacing strand for downstream strand displacement events. In strand displacement cascades, single-stranded DNA typically fulfills the role of *signals* that carry information, while pre-hybridized DNA complexes drive their interaction (and are consequently called *fuels*).

In a type of strand displacement called *toehold-exchange*, releasing the pre-hybridized strand breaks a previously bound toehold-size region (4). We say that such a reaction is enthalpy-neutral if the formed and broken toehold bonds have (approximately) equal free-energy. Enthalpy-neutral strand displacement cascades have played an important role in prior work. Driven forward solely by the configurational entropy of releasing molecules, molecular signals can get amplified exponentially (5). Enthalpy-neutral strand displacement is also used as an important catalytic module in chemical circuits for performing square-root computation (6), and neural network computation (7, 8). Designing molecular robots that can move along a track and sort cargo (9), implementing chemical reaction networks (10), and engineering enzyme-free dynamical behavior such as oscillation (11), all rely on the enthalpy-neutral strand displacement module. In enthalpy-neutral strand displacement, mismatches between the invading strand and its target can significantly affect the equilibrium, allowing the differentiation of single-nucleotide mismatches with high specificity (12). This principle has been applied in detecting low concentrations of mismatched targets in human genomic DNA (13, 14).

Even the simplest kind of enthalpy-neutral strand displacement cascades—a cascade of translators, which are logically equivalent to repeater gates—can already exhibit complex and useful behavior. In a single translator, a signal strand serves as the input and through a series of strand displacement reactions,

Significance Statement

Molecular control circuits embedded within chemical systems to direct molecular events have transformative applications in synthetic biology, medicine, and other fields. However, due to the complexity of molecular interactions, it is challenging to understand and formally reason about the collective functions of components within larger systems. Focusing on systems composed of one type of programmable DNA strand displacement reaction, we develop a formal framework to analyze desired and undesired properties affecting speed and correctness, and experimentally confirm the predictions. The properties are derived from structural design parameters based on combinatorial arguments (counting bonds and separate complexes). This approach, analogous to the rigorous proofs of algorithm correctness in computer science, can guide the future engineering of robust and efficient molecular algorithms.

B.W., C.T. and D.S. designed the research; C.T. proved the theorems for kinetic properties; B.W. and D.S. proved the other theorems; B.W. performed experiments and analyzed data; B.W., C.T., D.S. wrote the paper.

Preliminary versions of some of the results appeared in conference proceedings (1). The authors declare no conflict of interests.

[1]To whom correspondence should be addressed. E-mail: david.soloveichik@utexas.edu

an output signal strand is released whose sequence is independent of the input strand (Figure 1). Chaining translators allows us to build systems of translators. Such linear translator cascades have been used in a molecular automata system that can selectively target cellular surfaces (15). Translators can also be composed to compute the logic *OR* of many distinct inputs translated to a common output. Enthalpy-neutral translators (in the form of seesaw gate complexes) have been used to construct digital logic circuits (6). A catalytic system can be constructed if a translator chain's output is the same as its input.

Although the DNA strand displacement mechanism has shown great promise for programming molecular systems, the current scale of these systems remains limited. The main obstacle is *leak*, which occurs when undesired reactions get spontaneously triggered in the absence of the correct input. Fuels are designed not to have interacting toeholds; however, fraying at the end of a double helix can spontaneously expose a nucleation point for subsequent strand displacement even without the input. This *toeless strand displacement* is thought to be the primary source of leak (16).

To combat leak, a number of approaches have been tested at the enthalpic level and the entropic level. At the entropic level, since leak results from a spurious interaction of the fuel complexes, decreasing fuel concentrations increases the entropic cost of leak. Thus to combat leak, fuels are typically kept at low concentration. However this limits the general speed of strand displacement cascades—instead of seconds, complex cascades often take hours (6). At the enthalpic level, adding a threshold species that gains more enthalpic energy than the desired reaction can consume the leaked signal at a faster rate than it propagates to downstream components (6, 17). To reduce the enthalpic energy gain of spurious reactions, mismatches (18, 19) or adding 1 to 3 nucleotides as the *clamp* (6, 11, 17) have been introduced. Leak can only occur after the entire clamp and one or more additional nucleotides that are adjacent to the clamp fray. Clamps, in some form, are now commonly used in the majority of strand displacement systems.

Recently, a leak reduction method ("*NLD* design") was proposed where leak requires binding of N separate fuels, where the choice of N can be made arbitrarily large (16, 20). The chance of this happening decreases exponentially with N; equivalently, the energy barrier to leak can be made arbitrarily large by means of the entropy component.

In this work, we develop a systematic approach to design enthalpy-neural strand displacement (*EN*) cascades and rigorously prove properties related to desired and undesired signal propagation. We focus on translator cascades, and show that by adjusting two parameters which describe the length of the double-stranded region of the fuels (defined by N) and the length of the sequence that overlaps between two consecutive fuels (defined by *shift*), a variety of designs can be achieved, each with unique properties. Among the properties likely to hinder the kinetics of the intended reaction pathway, we characterize the parameter choices that lead to toehold occlusion (6, 11)—blocking the toehold for a desired reaction, and spurious strand displacement—(partial) displacement of a strand on a fuel by a spurious invader.

To analyze undesired signal (leak), we use a simple thermodynamic model which assumes that "enthalpy" and "entropy" are the dominating factors deciding whether a configuration is favorable or not (21–23). More specifically, we assume the main contribution of enthalpy is the number of bonds and that of entropy is the number of separate complexes.* Thus when comparing two configurations, if the number of bonds is the same while one configuration has n fewer separate complexes than the other, then the first configuration is considered unfavorable as it incurs n units of entropic penalty relative to the second. Similarly, all else being equal, a configuration with n fewer bonds than another configuration is unfavorable and has a relative enthalpic penalty of n units.† We show a thermodynamic penalty to leak consisting of an enthalpy penalty of $\frac{N}{shift}$, augmented by an entropy penalty of $\frac{N-r}{(shift-1)} - 1$ even if r bonds break. Interestingly, since no design satisfies all the properties we want—high energy penalty to leak, no toehold occlusion and no spurious displacement—understanding the taxonomy of designs is necessary to make proper design choices.

In a linear enthalpy-neutral cascade, every reaction is reversible, and thus the signal tends to spread out across the cascade more than in linear designs driven enthalpically by the bonding of toeholds. However, we show that the spreading does not limit the capability of enthalpy-neutral cascades to effectively propagate signal. Specifically, we prove that the system completion level does not decrease arbitrarily with the depth of the cascade, allowing long cascades to be constructed. Since the leak penalty increases with larger N (keeping *shift* fixed), while the level of desired output is bounded from below, the "signal-to-noise" ratio can be made arbitrarily high.

Compared with the *NLD* design discussed above, we achieve a stronger thermodynamic guarantee of correct behavior. In the *NLD* design, a leaked upstream signal can start a cascade which gains enthalpy from toehold binding for every downstream strand displacement step, which could be overall favorable for long cascades. Thus, the prior work corresponds to a kinetic barrier in our model. In contrast, in our *EN* design, a leaked upstream signal cannot be driven by the enthalpy of forming new bonds downstream. In fact, leak requires the overall *loss* of bonds in our design (enthalpic penalty), and leak in our designs is always thermodynamically penalized. Even more, in the typical experimental regime, compared with the height of the kinetic barrier in the *NLD* design, our thermodynamic penalty is twice as high, and designs with small *shift* are also more efficient in terms of total DNA used.

We conclude with experimental demonstrations of the *EN* designs with different sets of parameters. The laboratory results confirm our theoretical categorization of designs with regard to leak and undesirable kinetic properties of toehold occlusion and spurious displacement. We show that a translator design categorized as having a high energy penalty to leak has less leak than a comparable design without such a penalty. Further, a design with toehold occlusion reaches reaction completion slower than a comparable one without.

* For simplicity, we use "enthalpy" as a shorthand to refer to bond formation, although we acknowledge both the enthalpic and entropic contributions of bond formation in real physical chemistry (which can be physically distinguished based on their temperature dependence). Similarly our simple model ignores other components of entropy such as conformational entropy and entropy of mixing.

† Roughly speaking, "one unit of enthalpic penalty" corresponds to an average of $l \cdot 1.5$ kcal/mol, where l is the length of the domain (typically 5-10 nucleotides for a toehold). "One unit of entropic penalty" at concentration C M corresponds to $\Delta G^{\circ}_{\text{assoc}} + RT\ln(1/C) \approx 1.96 + 0.6\ln(1/C)$ kcal/mol (24). With these numbers, at roughly 650 nM concentration, binding an additional $l = 7$ domain is equal to one unit of entropy. At low concentrations the entropic penalty becomes dominant, while the enthalpic penalty prevails at high concentrations.

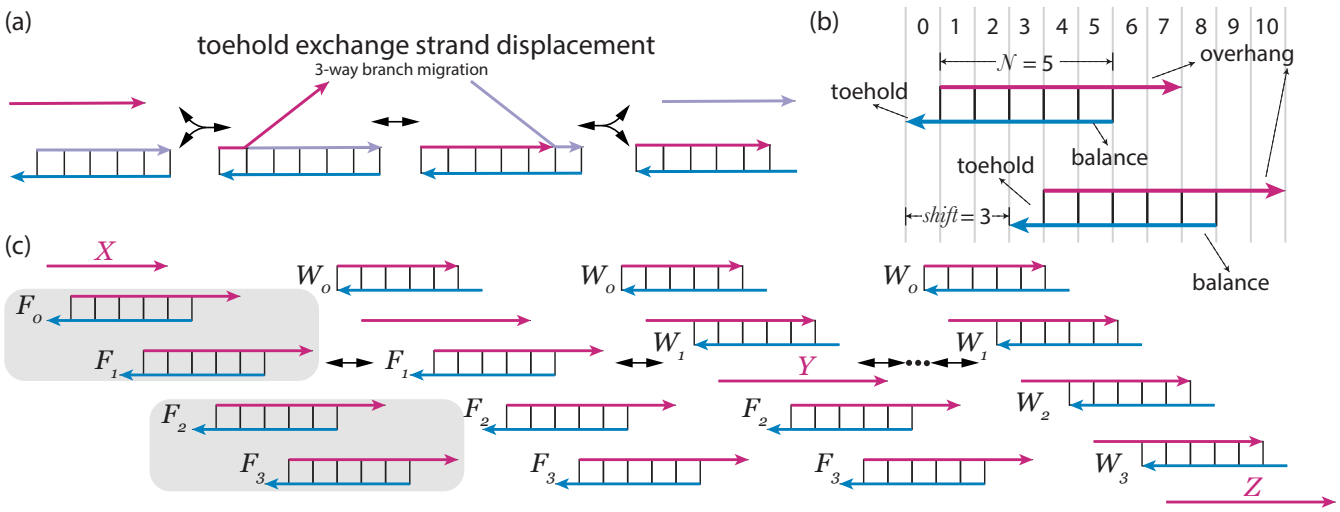


Fig. 1. (a) The fundamental reaction step we consider. (b) The conventions of the *EN* design. (c) Desired reaction pathway of a 2-translator cascade ($X \to Y \to Z$). After 4 strand displacement steps, the signal strand X is translated to signal strand Z.

We also confirm the spurious displacement reaction pathways that could in principle affect desired reaction kinetics. The leak reduction we observe compares favorably with the previously proposed *NLD* designs, in terms of the kinetic leak rate, the relative triggered signal to leak ratio and the maximum amount of leak generated.

1. Design space of enthalpy-neutral strand displacement cascades

A. System description. We first introduce the conventions used in this paper. We use *top strand* to indicate the strand at the top of a fuel complex and *bottom strand* to indicate the strand at the bottom of a fuel complex in our illustrations. We use the domain level abstraction for DNA strands: Strands are divided into domains, where a domain represents a concatenation of DNA bases treated as a unit. Domains can hybridize to or dissociate from a complementary domain. Here all domains are assumed to be short enough that if a complex is bound only by a single domain it can dissociate (see Figure 1a). (Such domains are typically said to be of toehold-length.) If a complex is bound by two or more domains then it cannot dissociate. We assume that all domains are orthogonal (no cross-talk). Note that this is a strong assumption because the size of a domain is restricted to that of a toehold and there are a limited number of distinct toeholds that could be designed (see Discussion). The desired pathway consists entirely of toehold exchange strand displacement reactions. We assume that fraying cannot open a whole domain.

In a particular complex, a domain can be *single-stranded* (a.k.a *open*) or *double-stranded*. The domains in top strands are *top domains*, and the domains in bottom strands are *bottom domains*. The single-stranded top domains on a fuel together are called an *overhang*. Single-stranded bottom domains are also called *toeholds*. The domains located at the right end of a double-stranded helix are called *balance domains*, since they need to break after strand displacement steps and they ensure equal enthalpy before and after the reaction. Note that the *balance domain* also serves as a clamp. Historically, clamps were added to previous strand displacement designs to "clamp-down" the ends of helices to reduce leak. Without the balance domain on the second fuel in Figure 1b, the overhang of the first fuel can initiate toeless strand displacement.

We align the unique domains in a row and represent their identities using the horizontal positions (i.e., numbering on top of Figure 1b). Domains aligned vertically have the same or complementary sequence: Top domains aligned vertically are the same, and bottom domains aligned vertically are the same; Top domains are complementary to the bottom domains aligned vertically. When we refer to a *position*, we mean all domains that have the same or complementary sequences, and are drawn vertically aligned in our figures. Note that as is typical of at-scale strand displacement experiments, there are many copies of each fuel although only one copy is drawn.

B. Linear strand displacement cascades. In this work we concentrate on linear strand displacement cascades. When the input signal strand is present, it reacts with the first fuel displacing the top strand, which then serves as the input to trigger the downstream fuel (Figure 1c). Such a cascade can continue if subsequent downstream fuels are available. A linear strand displacement cascade can be used to build *translators*. A translator effectively converts an input signal strand to an output signal strand of independent sequence. Note that the Figure 1c shows two translators of two fuels each (since Y is sequence independent from X, and Z is sequence independent from Y).

To design a linear, enthalpy-neutral strand displacement cascade, two parameters are necessary and sufficient. We use parameter N ($N \geqslant 2$) to represent the number of double-stranded domains in a fuel and parameter *shift* to represent the distance between consecutive single-stranded toeholds (see Figure 1b). The parameter *shift* should be between 1 and N to ensure signal can be transmitted to the next fuel. In Figure 2, we illustrate the diversity of designs when $N = 6$ for all valid values of *shift*.

Once the parameters N and *shift* are determined, the types of each domain can be assigned. Using 0 indexing, we start from the 0th fuel which is responsible for the first strand displacement step (e.g. Figure 1c). The toehold of the 0th fuel

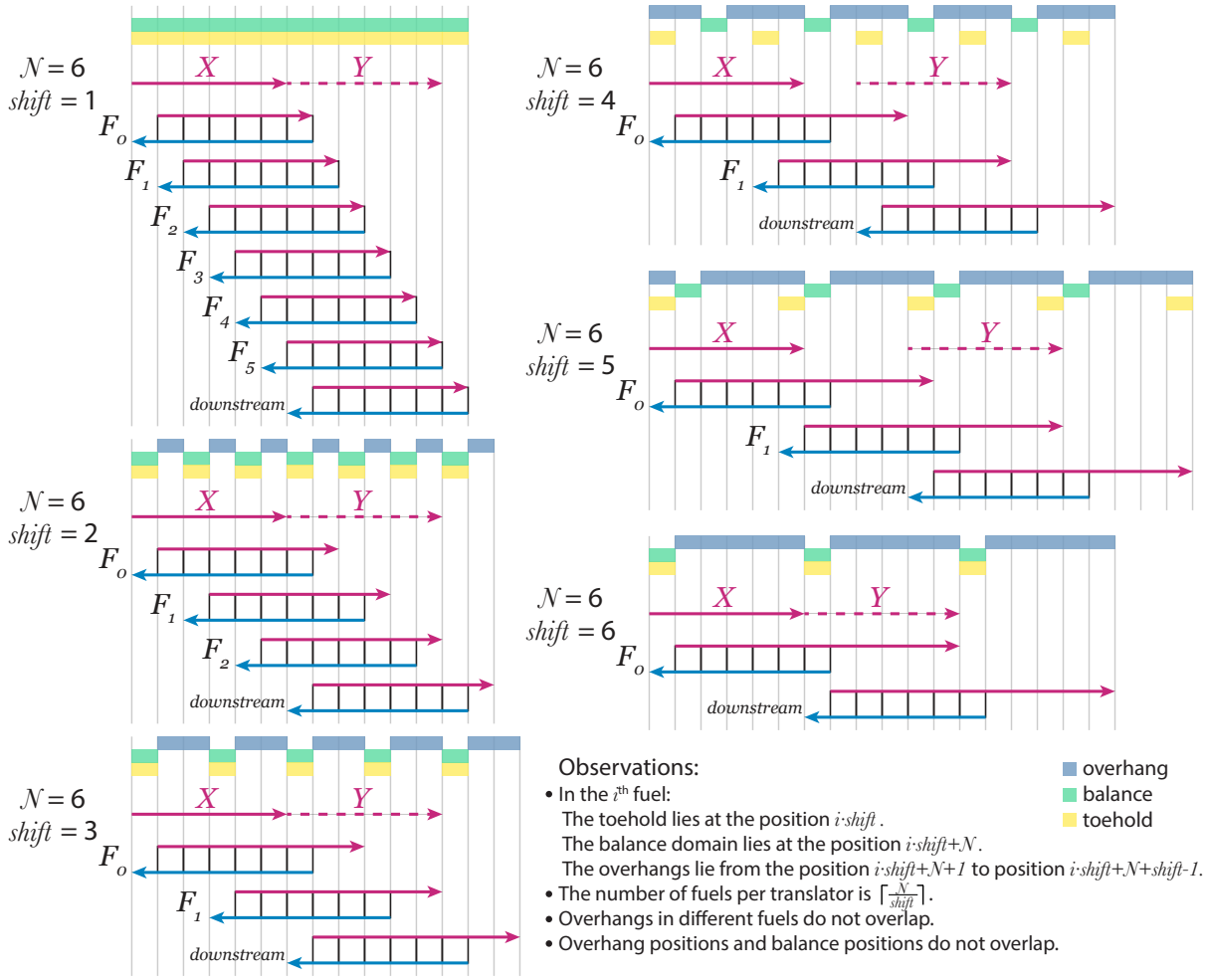


Fig. 2. Initial configurations for linear strand displacement cascades with $N = 6$ and all values of *shift*. For each combination of parameters, enough fuels are shown to output a signal of independent sequence from the input X (i.e., a full translator), as well as an additional (downstream) fuel. The dashed strand Y indicates the positions of the output signal strand that the downstream fuel takes as input. The positions are colored on top of each design according to their position types (see legend). Top strands are colored in purple and bottom strands are colored in blue. Note that although drawn as a single copy, we consider a regime with multiple copies of each fuel. The observations are justified and explained in Section S2.

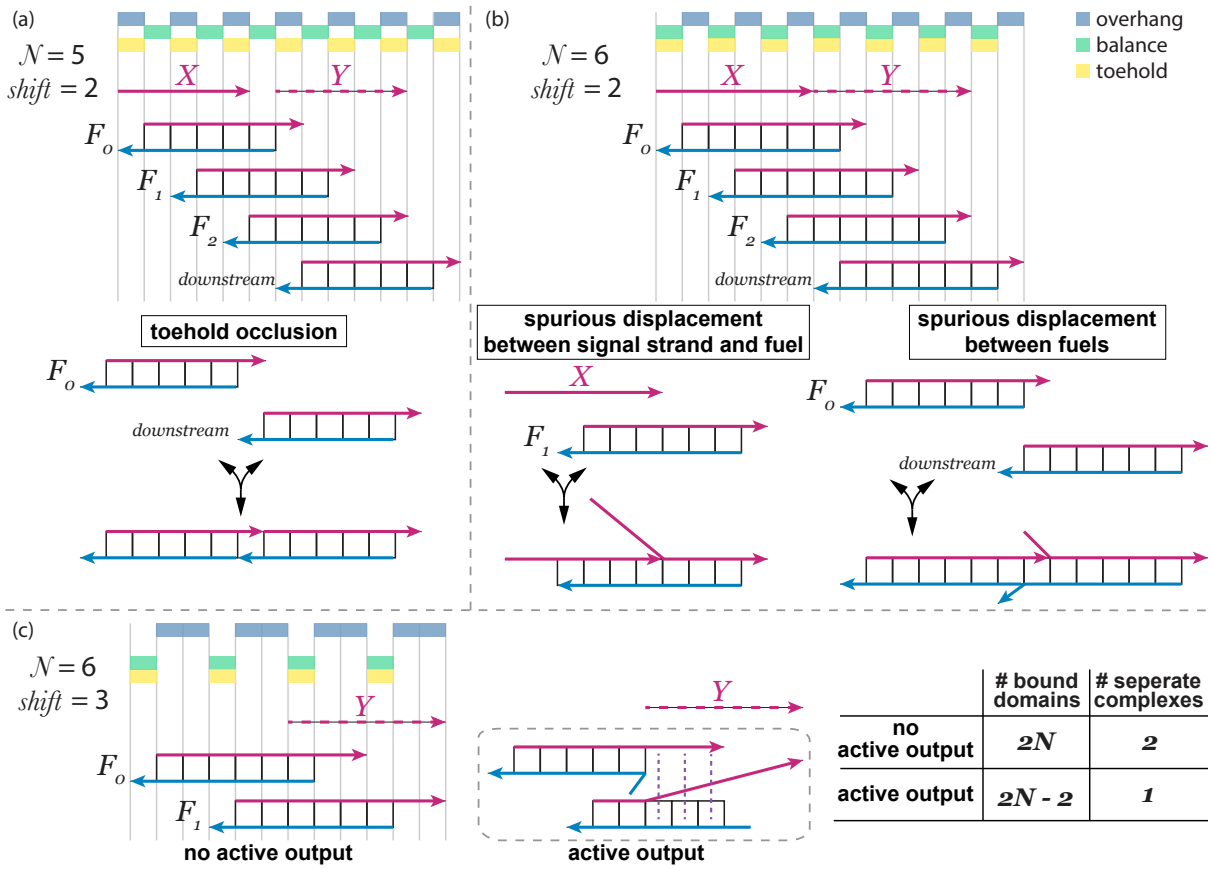
is located at position 0.

A position can have different position types: *toehold*, *overhang*, and *balance*. We assume the linear strand displacement cascade is extended both upstream and downstream, so that every bottom strand has exactly the same layout of position types on all the domains. If there exists a certain domain in a position, that position is assigned a position type according to the domain it contains. For example, the position with a toehold domain is a toehold position.

The coloring of the positions refers to its position type (Figure 2). In particular, the color yellow indicates a toehold position, and the color green indicates a balance position, and the color blue indicates an overhang position. If a position has multiple types, it is colored in both. Basic properties and observations for the enthalpy-neutral strand displacement cascade are shown in Figure 2 and listed in Section S2.

C. Asymptotic completion level of translator cascades. With a cascade of effectively irreversible strand displacement reactions (every reaction gains the energy of binding a toehold), it is safe to assume that most of the input signal should propagate through to the end. However, with a cascade of reversible reactions such as those we necessarily obtain with enthalpy-neutral strand displacement, signal might decrease with the length of the cascade if the signal "spreads out" across the layers. Does this mean that enthalpy-neutral strand displacement reactions cannot be composed into long cascades? In this section we prove a lower-bound on the amount of final signal output by a chain of enthalpy-neutral strand displacement reactions that is *independent* of the length of the cascade, which shows that long cascades are indeed feasible.

To analyze a system with a cascade of enthalpy-neutral strand displacement reactions, we simplify each reaction to be a bimolecular reversible toehold exchange reaction $X + F \rightleftharpoons Y + W$, where X is the input signal, F is the fuel, Y is the displaced strand of this reaction and W is the waste species. Assuming that the two toeholds in a toehold exchange reaction (i.e., toehold and balance domains in our nomenclature) have the same thermodynamic binding strength, the net reaction of a translator has $\Delta G^o \approx 0$ and the equilibrium constant of each reaction can be treated as 1. Thus for a single translator, if the initial concentration for the reactants are $[X]_0 = \alpha$,



	# bound domains	# seperate complexes
no active output	2N	2
active output	2N - 2	1

Fig. 3. Examples of the properties of linear enthalpy-neutral strand displacement cascades. (a) Toehold occlusion in the design $N = 5$, *shift* $= 2$. (b) Spurious displacement in the design $N = 6$, *shift* $= 2$. (c) Energy penalty for active output in the design $N = 6$, *shift* $= 3$. The dashed gray box indicates the strands in the box are in one complex. The dash purple lines indicate bound domains. The shown active output configuration incurs both an enthalpy (two fewer bound domains) and entropy (one fewer separate complex) penalty and is thus thermodynamically unfavorable.

$[F]_0 = 1$, at chemical equilibrium, the concentration of output strand Y will be $\frac{\alpha}{\alpha+1}$ [‡].

We then ask how much output signal will be produced if multiple reactions are cascaded together. Suppose we have a n-layer reaction system:

$$X_1 + F_1 \rightleftharpoons X_2 + W_1$$
$$\dots$$
$$X_i + F_i \rightleftharpoons X_{i+1} + W_i$$
$$\dots$$
$$X_n + F_n \rightleftharpoons X_{n+1} + W_n$$

The system starts with all the fuels (F_i, $i = 1, 2, \dots, n$) at concentration 1 and X_1 at α. The reactions are driven forward by the imbalance between F_i and W_i as described by the following theorem.

Theorem 1.1 *Given a linear enthalpy-neutral strand displacement cascades with n ($n \geqslant 2$) layers, and the fuel concentration 1 and the input concentration α, the equilibrium concentration of the output signal is at least $\alpha\frac{1-2\alpha}{1-\alpha}$ independent of n.*

[‡] This is obtained by solving the equilibrium concentrations: Assume the equilibrium concentration of Y is $[Y]$, and the equilibrium concentration for the other species are $[W] = [Y]$, $[X] = \alpha - [Y]$, $[F] = 1 - [Y]$. According to the equilibrium expression for this reaction $K = \frac{[Y][W]}{[X][F]}$, where K is the equilibrium constant; in this case, $K = 1$. Solve the equation for $[Y]$, and we get $[Y] = \frac{\alpha}{\alpha+1}$.

If we look at the lower bound on the ratio of output to input, $\frac{1-2\alpha}{1-\alpha}$ approaches 1 for small α—i.e., almost all of the input is converted to output. On the other hand, for $\alpha \geqslant \frac{1}{2}$ the theorem does not provide a positive lower bound on the output and thus is not meaningful (although, of course, the output signal is higher for larger α). Overall, since our lower bound does not change with the number of layers n, it implies that concatenating the enthalpy-neutral strand displacement reactions generates at least a constant fraction of output signal independent of the number of layers.

2. Kinetic properties

In this section, we show that designs are susceptible to undesirable properties such as *toehold-occlusion* or *spurious strand displacement* to varying degrees. Certain designs preclude some of these phenomena entirely. We prove kinetic properties of designs based on their parametrization of N and *shift*. Many of our arguments rely on showing whether regularly spaced positions with certain domains (such as toeholds) can intersect with other regularly spaced positions or intervals with different domains (such as overhangs). Those arguments are based on some useful lemmas in Section S1.

A. Toehold occlusion. In strand displacement systems, reaction kinetics can be controlled by the strength of a toehold (4). Stronger toeholds can enable faster reaction kinetics; however,

if the toehold strength is too strong, undesired binding of a toehold can slow down reaction kinetics since to undo the undesired binding, toehold dissociation can become a rate limiting step. Since fuels are typically present in high concentration, this is especially problematic when overhangs of fuel can bind to toeholds of other fuel. This creates so-called *toehold occlusion* (6, 11), which can significantly slow down the intended reaction kinetics in the presence of input signal. Toehold occlusion exists when there are overlapping overhang and toehold positions (Figure 3a).

Theorem 2.1 *Toehold occlusion is not possible in a linear enthapy-neutral strand displacement cascades if and only if N is a multiple of* shift.

B. Spurious strand displacement. Spurious strand displacement events are unproductive reactions that can slow down the intended kinetics of the system. Spurious displacement occurs when any proper prefix or suffix of a fuel's double-stranded helix is displaced by a spurious invader. A spurious invader of a fuel is any complex different from its intended input strand. We partition our analysis into two categories as shown in Figure 3b: (i) spurious displacement between fuels, and (ii) spurious displacement between signal strand and fuels.

Spurious displacement between fuels can become increasingly problematic with respect to the intended kinetics of a cascade, as the concentration of the system is increased. The intent for increasing concentration is to speed up the desired reaction; however, higher concentration increases the fraction of fuels spuriously bound to each other, making them unavailable for their intended reactions with signal strands. Lemma S4.3 and Lemma S4.4 discuss designs with spurious displacement in the absence of input.

The second type of spurious displacement is when a free signal strand (including the input), can act as a spurious invader of a fuel other than its designed target. When the input concentration is significantly lower than fuel, as is typical, signal strands can become involved in numerous unproductive reactions thus slowing (possibly significantly) signal propagation through every layer of the cascade. Lemma S4.5 discusses spurious displacement in the presence of input.

The following theorem shows when either kind of spurious displacement is avoided.

Theorem 2.2 *Spurious displacement is not possible in a linear enthapy-neutral strand displacement cascade if and only if* $\text{shift} = N - 1$.

3. Thermodynamic properties: Thermodynamic penalty to leak

An effective strand displacement cascade ideally has a high signal-to-leak ratio, even at thermodynamic equilibrium. To study the thermodynamics of leak, we need to answer two questions about the system in the absence of input: What is thermodynamically preferred and what is the thermodynamic penalty to leak? Our thermodynamic arguments are based on counting the number of bonds formed (more bonds are more favorable) and the number of separate complexes (more complexes are more favorable) (21, 23). An example of analyzing the energy penalty to leak is shown in Figure 3c. The penalty to leak is quantified with respect to these two thermodynamic factors. We show that by varying design parameters N and *shift*, we vary the thermodynamic penalty to leak. Note that all of our combinatorial arguments are valid both in the single-molecule regime or when the fuels have multiple copies. Different fuels may also be present in different amounts.

For the purposes of the arguments in this section, we say the *ith-output region* (for any $i > 0$) consists of N positions $i \cdot \mathit{shift}$ to $i \cdot \mathit{shift} + N - 1$. We say a configuration has an *active output* if some complex has open top domains in all N positions of some output region. These positions represent the domains in a top strand that can displace a strand of the downstream fuel. A *configuration* is a matching between corresponding top and bottom domains, where each match is one bond. We use *initial configuration* to represent the configuration where complexes are prepared separately—with one top strand and one bottom strand in a complex—and then mixed together without any interaction (e.g. the configurations shown in Figure 2).

We first show that in the absence of input, the initial configurations are the thermodynamically preferred configurations for designs without toehold occlusion (Section 3.A). Thus the thermodynamic penalty to produce an active output in the absence of input is based on the comparison between the active-output configuration and the initial configuration. An active output has N unbound domains. At some of these positions, to produce an active output in the absence of input, it is sufficient to bring some complexes together without decreasing the overall number of bonds. This decrease in the number of separate complexes results in an entropic penalty. However, at other positions, no matter how many complexes are brought together, producing an active output requires decreasing the number of bonds, which results in an enthalpic penalty (as discussed in section 3.B). Section 3.C discusses the trade off between the enthapic and entropic penalty.

Note that the arguments in this section only apply to designs without toehold occlusion, because for these designs, in the absence of input, the thermodynamically preferred configurations are the initial configurations. However, for designs with toehold occlusion, depending on the parameters of the system (strength of toehold, concentration, temperature, etc.), the thermodynamically preferred configuration is variable and toeholds can be either occluded or not. (If occluding toeholds is preferred for a given reaction condition, the thermodynamically preferred configuration has the maximum number of possible bonds and the results about the enthalpic penalty in section 3.B still hold.)

A. Initial configuration is thermodynamically preferred without toehold occlusion. We first show that for designs without toehold occlusion, in the absence of input, the initial configuration is thermodynamically preferred. We simplify the thermodynamic properties of a system by analyzing the contributions of "enthalpy"—number of bonds—and "entropy"—number of separate complexes. There might be tradeoffs between enthalpy and entropy, but if we find a configuration that simultaneously maximizes both, it is thermodynamically preferred. In our case, we will show that the initial configuration simultaneously maximizes both unless we are willing to lose a lot (N) of bonds. Thus it is the thermodynamically preferred configuration in a wide range of experimental regimes.

In our proof of the following theorem, we gradually restrict the possible complexes of the thermodynamically preferred configuration through Lemma S5.2, Lemma S5.3, and finally by showing that they must consist of exactly one top and one

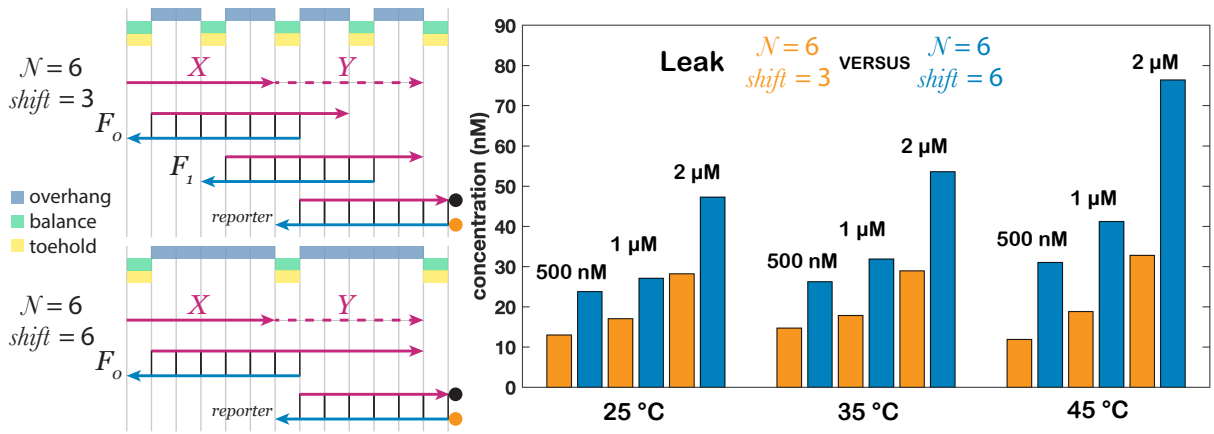


Fig. 4. Comparison of designs with different energy penalties to leak. Signals at different thermodynamic equilibrium (25 °C, 35 °C and 45 °C) after a slow annealing from 90 °C (0.06 °C/min) without input strand. The design with parameter set $N = 6$, *shift* $= 3$ has less relative leak. The top strand of the reporter is labeled with a quencher (black circle). The bottom strand of the reporter is labeled with a fluorophore (orange circle). The concentrations for fuels and reporters are labeled in the figure.

bottom strand (Lemma S5.4). Combined with an analysis of the tradeoff between bonds and separate complexes later developed in Theorem 3.3, this yields the following result:

Theorem 3.1 *For designs without toehold occlusion, given a linear enthalpy-neutral strand displacement cascade of arbitrary depth, in the absence of input, we have: (i) Among the configurations with maximum bonding, the initial configuration is the unique configuration that maximizes the number of separate complexes. (ii) The initial configuration has at least N more bonds than any configuration with more separate complexes.*

B. Enthalpic penalty to leak. Our arguments about the enthalpic penalty are based solely on counting the difference in bonds between the thermodynamically preferred configuration and the configuration with active output. We show that the enthalpic penalty to leak can be increased arbitrarily by tuning the parameters N and *shift*.

To have an active output, it is necessary to break some bonds if there is no free top domain among the output region. Thus it is sufficient to count the number of positions among the output region that do not have excess top domains (top-limiting positions; Lemma S5.6).

Theorem 3.2 *Given a linear enthalpy-neutral strand displacement cascade of arbitrary depth, in the absence of input, any configuration having active output has at least $\lceil \frac{N}{\text{shift}} \rceil$ fewer bonds than the maximum-bond configuration.*

This theorem implies that in the absence of input there is an enthalpic penalty of $\lceil \frac{N}{shift} \rceil$ bonds to generate an active output from a maximum-bond configuration. Note that in the designs without toehold occlusion, the initial configuration maximizes the number of bonds and so can act as the unique point of comparison (via Theorem 3.1). The theorem still applies to designs with toehold occlusion but the thermodynamically preferred configurations are not necessarily maximum-bonded so the relevance of the point of comparison is not clear.

In contrast, when the input is present, the signal can be propagated all the way with no net enthalpic cost. Thus by increasing N we can enlarge the enthalpic penalty to leak without increasing the enthalpic penalty to correct output. The enthalpic penalty holds even at high concentrations when there is little entropic penalty of joining two complexes into one, and thus sets the lower bound on the required energy for leak independent of concentration.

The enthalpic penalty shows that leak results in fewer bonds compared with maximum bonding. Moreover, Theorem 3.1(ii) shows that leak cannot compensate for the loss of bonds with the thermodynamic benefit of creating additional separate complexes (unless N bonds break). Indeed, in the next section we show that leak results in a *decrease* in the number of separate complexes, and that this entropic penalty can be made arbitrarily large by varying N and shift.

C. Entropic penalty to leak. In this section we show that no matter how strands are rearranged in our enthalpy-neutral strand displacement cascades, any configuration with active output has fewer separate complexes—thus it incurs an entropy penalty. As the point of comparison, we count the difference in the number of separate complexes between the thermodynamically preferred configuration as captured by Theorem 3.1 and the active-output configuration.

By Theorem 3.2, we know that the enthalpic penalty has a lower bound of $\lceil \frac{N}{shift} \rceil$. Thus the arguments below assume that the configuration with active output has r bonds fewer than maximum, where $r \geqslant \lceil \frac{N}{shift} \rceil$. Beyond this lower-bound on r, the entropy penalty can be traded-off for an decreased enthalpy penalty by increasing r. Intuitively, forcing fewer bonds to form gives more flexibility to produce active output without bringing as many complexes together.

To count the difference in the number of separate complexes, we can equivalently count the number of top strands associated in a complex in the active-output configuration (since the thermodynamically preferred configuration has only one top strand in every complex by Lemma S5.4). Then by a combinatorial argument based on counting the types of positions in top and bottom strands (Lemma S5.9, Lemma S5.10, Lemma S5.11 and Lemma S5.12), we show that for active output there must exist some complex with at least $\frac{N-r}{shift-1}$ top strands (Lemma S5.11 and Lemma S5.12), implying the following theorem:

Theorem 3.3 *For designs with* shift $\neq 1$ *and no toehold occlusion, given a linear enthalpy-neutral strand displacement*

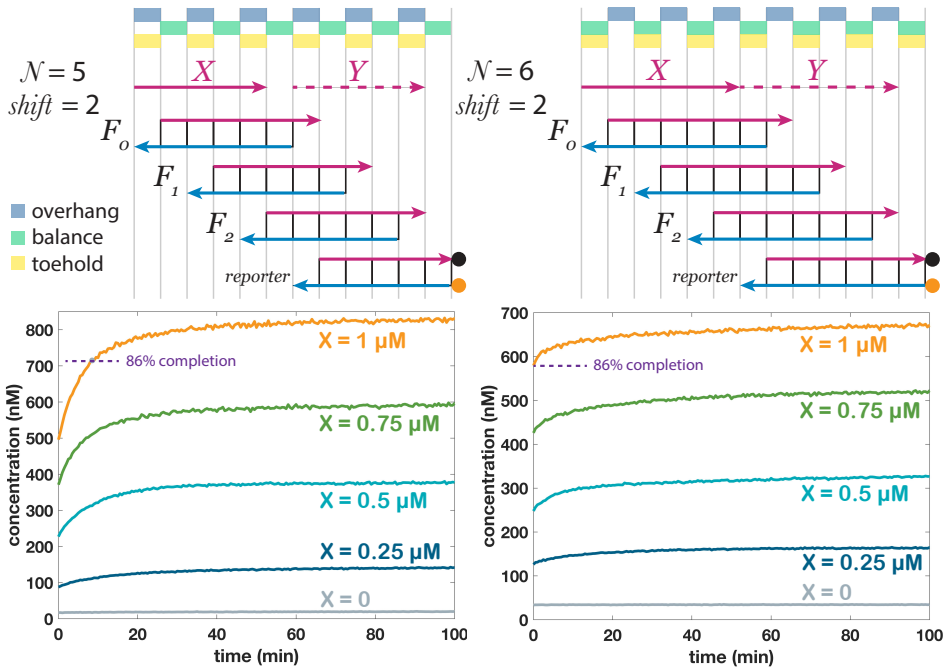


Fig. 5. Comparison of kinetics of designs with ($N = 5$, *shift* $= 2$) and without ($N = 6$, *shift* $= 2$) toehold occlusion. The design with the parameter set $N = 6$, *shift* $= 2$ is significantly faster. The concentrations of reporter and fuels were $2\ \mu$M. The reaction temperature was $25\ ^\circ$C. As in all kinetic curves in this work, time 0 represents the first data point measured, which is about 2 minutes after samples were mixed.

cascade of arbitrary depth, in the absence of input, any configuration with active output and r ($r \in \mathbb{N}, r \geqslant \frac{N}{\text{shift}}$) bonds away from maximum has at least $\frac{N-r}{\text{shift}-1} - 1$ fewer separate complexes than the initial configuration.

This theorem implies that in the absence of input there is an entropic penalty of $\frac{N-r}{shift-1} - 1$ units to produce an active output. In the case of *shift* $= 1$, active output requires an enthalpy penalty of $r = N$, but can result in one additional separate complex.

As in the case of enthalpy, when the input is present, the signal can be propagated all the way with no net entropic cost. Combined with the results of the previous sections, if we fix *shift* and increase N, both the enthalpic (Theorem 3.2) and entropic penalty (Theorem 3.3) to leak increases, while the desired triggering signal reaches an asymptote (Theorem 1.1).

The thermodynamic penalties to leak in this and the previous section improve on the *NLD* design. As discussed before, *NLD* leak may be overall thermodynamically favorable in long cascades due to the bonding of additional toeholds after active output is generated. This is not, by itself, enough to necessarily prefer the *EN* designs to *NLD* designs because a high kinetic barrier might practically prevent leak better than a relatively low thermodynamic penalty. Thus, we want to quantitatively compare the two. Both the *NLD* kinetic barrier and the enthalpy-neutral thermodynamic penalty can be made arbitrarily large by proper choice of parameters, so we compare the size of both as a common function of the complexity of the system as captured by the number of fuel complexes and the total amount of DNA. In Section S7 we show that, for the same number of fuel complexes, our end-state thermodynamic penalty to leak is $\beta+1$ the size of the *NLD* kinetic barrier when the free energy of forming a new bond (toehold) is β times as favorable as a new separate complex. Thus our penalty to leak is always larger, and that when enthalpy and entropy are balanced[§] ($\beta = 1$) our penalty is twice as high. For larger β (e.g., especially high concentrations), the advantage of our design is even larger (any *shift*). Designs with small *shift* are also more efficient in terms of total DNA used.

4. Trade-offs between the properties

We have shown that different choices of N and *shift* yield *EN* designs with varying thermodynamic and kinetic properties. We desire a large thermodynamic penalty to leak, which requires a large ratio $N/$*shift* (Theorems 3.2 and 3.3). However, only designs with *shift* $= N - 1$ can avoid spurious strand displacement 2.2, which is incompatible with a large $N/$*shift* ratio. In fact, designs with the largest thermodynamic penalty to leak also have the most potential spurious interactions between fuels and signal strands. (Compare the spurious interactions of the X input signal with fuels other than its intended target for different *shift* in Figure 2.)

Besides spurious displacement, the other undesired kinetic property is toehold occlusion. While the absence of toehold occlusion is compatible with a large thermodynamic penalty (Theorem 2.1), there is a fundamental incompatibility between the two kinetic properties and only one design can avoid both:

Corollary 4.0.1 *The only design that avoids both toehold occlusion and spurious displacement is* $N = 2$, shift $= 1$.

In summary, there is no best design with respect to all thermodynamic and kinetic properties studied here. Instead, the entire taxonomy we develop informs the choice of translator design, and one should be chosen based on the expected

[§] It is experimentally desirable to be close to the regime where forming a new bond (toehold) or a new separate complex is equally favorable. This regime balances high concentration (reaction speed) with toehold occlusion and spurious signal-fuel binding.

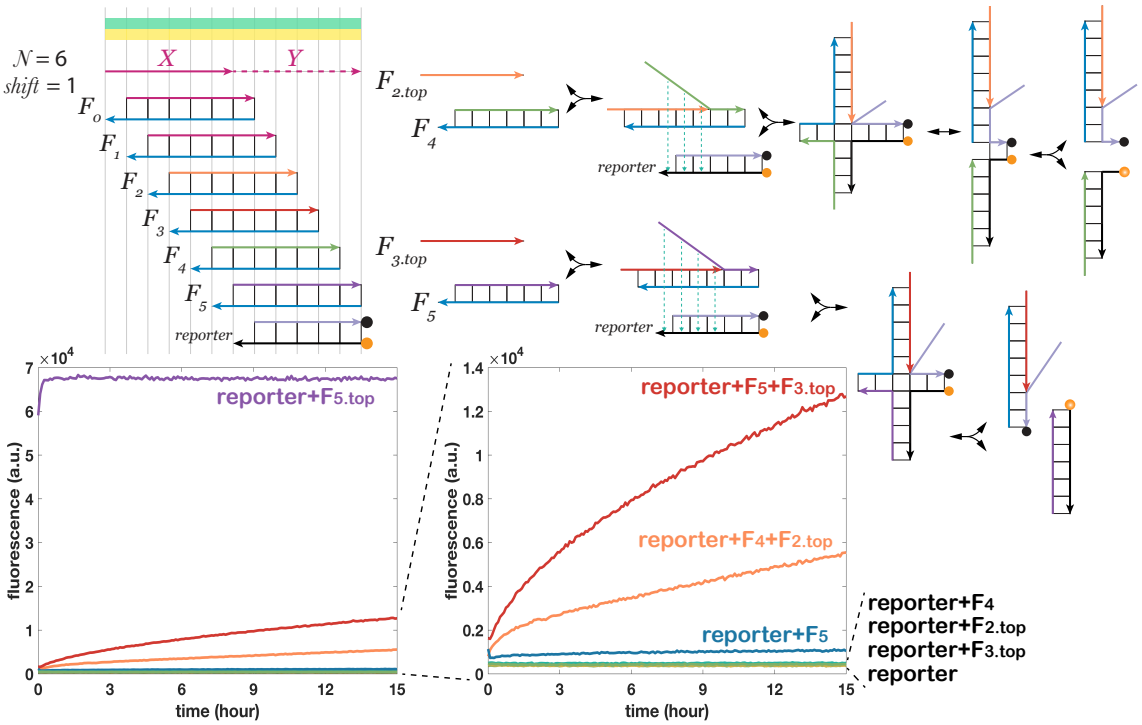


Fig. 6. Spurious displacement pathways and kinetics of spurious interactions at the design $N = 6$, *shift* $= 1$. The concentrations of reporter and fuels were 500 nM. The input at each layer was 100 nM. The reaction temperature was $25\ ^\circ$C.

conditions of its planned use. For example, high concentration conditions may be best served by a design with no toehold occlusion and a balance between its enthalpic penalty to leak and its potential number of spurious displacement reactions.

5. Experiment results

In this section we experimentally confirm the theoretical results we develop in prior sections. In Section 5.A we experimentally compare two sets of parameters of N and *shift*, and the one that is predicted to have a larger penalty is less leaky. In Section 5.B, we confirm the theoretically identified parameters for toehold occlusion—an undesired property that may lead to slower kinetics. Besides spurious toehold binding, another undesired property is spurious displacement wherein a partially displaced strand can trigger a series of undesired reactions. We identify one such spurious pathway in Section 5.C, supporting the relevance of the theoretical results in Section 2.B. Theoretically, the enthalpy-neutral strand displacement cascades introduces an additional enthalpic energy penalty to leak compared to prior work which only focuses on entropy (16, 20). We experimentally demonstrate the comparison in Sections 5.D. We conclude by testing desired triggering and leak in longer cascades in Section 5.E, showing small overall leak but robust desired triggering in accord with the theory of Section 1.C.

A. Designs with different energy penalties to leak. We experimentally tested the energy penalties to leak among two designs with different parameters (Figure 4). The $N = 6$, *shift* $= 3$ design has 2 unit of enthalpic cost and 1 unit of entropic cost to leak. The $N = 6$, *shift* $= 6$ design has only 1 unit of enthalpic cost, and no entropic cost.

Experiments demonstrate that with various temperatures, the $N = 6$, *shift* $= 3$ design has less leak than the $N = 6$, *shift* $= 6$ design (Figure 4). The leak concentration of the $N = 6$, *shift* $= 6$ design increases with temperature, since the leak penalty is only contributed by enthalpy and increasing temperature reduces the enthalpic penalty. In contrast to enthalpy, the effect of entropy on the equilibrium is independent of temperature. In contrast, the $N = 6$, *shift* $= 3$ design does not have significant temperature effect on leak concentration. This could be due to the combination of both the enthalpic and entropic contributions to the leak penalty.

At various concentrations, the $N = 6$, *shift* $= 3$ design also has less leak than the $N = 6$, *shift* $= 6$ design. Since the energy penalty to leak of the $N = 6$, *shift* $= 6$ design is independent of entropy, the leak concentration is supposed to be proportional to the initial concentration of the reacting species. However, the experimental leak concentration does not double when the initial concentration doubles. This deviation could be caused by some non-entropy-neutral reactions. Note that the design does not have the undesired property toehold occlusion, and the non-entropy-neutral reactions could be strong and spurious association or the fluorophore-quencher interaction holding the two strands of the reporter.

In the presence of input signal, the two designs produce similar amounts of output signals (Figure S3). Indeed, the low-leak $N = 6$, *shift* $= 3$ design has slightly higher output.

B. Designs with and without toehold occlusion. Our theoretical results classify designs with toehold occlusion (Theorem 2.1), which can hinder the kinetics of desired strand displacement. Here, we experimentally compare the designs with ($N = 5$, *shift* $= 2$) and without ($N = 6$, *shift* $= 2$) toehold occlusion. As shown in Figure 5, at various concentrations of input, the design without toehold occlusion ($N = 6$, *shift* $= 2$) reaches completion faster than the design with toehold occlu-

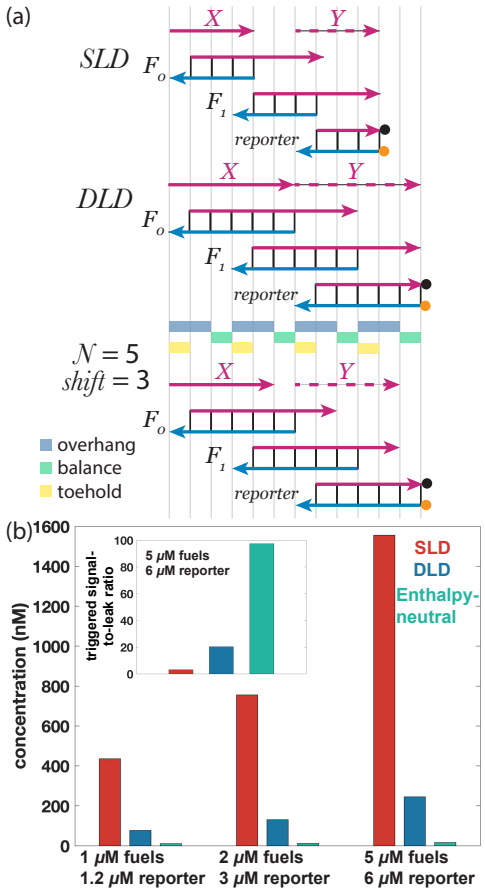


Fig. 7. Comparison of leak between (a) the *SLD*, *DLD* and *EN* design ($N = 5$, *shift* $= 3$) from the perspective of (b) the thermodynamic equilibrium. The experiments measuring the leak concentration at the equilibrium do not contain the input signal strands. Fuels and reporter were slowly annealed (0.06 °C/min) from 90 °C to 25 °C within 18 h. (Inset) The triggered signal-to-leak ratios of the *NLD* design and the *EN* design. The *SLD* and the *DLD* designs were assumed to reach full completion. The triggered signal for the *EN* design was shown in Figure S4. The concentrations for fuels and input were 5 μM, and the reporter 6 μM. The reaction temperature was 25 °C. For simplicity, the clamp domains in the *SLD* and *DLD* designs (which extend the bottom strand by 2 nt) are not shown here but are illustrated in Figure S8. Note that if the clamps were extended to toehold size, these two designs could also be categorized as *EN* designs (see Discussion).

sion ($N = 5$, *shift* $= 2$). We define completion as the signal level at 100 min. When the input concentration is 1 μM, the design without toehold occlusion reaches 86% of completion before the first measured data point, but it takes nine minutes for the design with toehold occlusion to reach the same completion level. Note that these two designs share the same toeholds, but there are slight differences between the energies of the toehold and balance domains. These thermodynamic differences actually favor the design with toehold occlusion, and thus cannot account for the kinetic slowdown.

C. Spurious displacement pathway with signal strands. Spurious displacement may result in the formation of complex multi-stranded structures or intermediates, which could interfere with desired kinetics. Theorem 2.2 classifies designs with and without spurious displacement overall, and Lemmas S4.4 and S4.5 separate the cases of spurious displacement with the input absent or present. We chose the design with $N = 6$, *shift* $= 1$, which falls into the category of spurious displacement with input, to test the predicted spurious displacement reaction pathway. This parameter set represents the extreme case of spurious displacement since every input to a fuel has multiple invading sites with other fuels.

Figure 6 shows two possible spurious displacement pathways involved with signal strands and the fuels that are not the designed targets of the signal strands. The signal strand first binds and partially displaces the top strand of the fuel, exposing a single-stranded portion of the top strand. The single-stranded portion can then interact with other downstream species. We use a reporter to represent the downstream species and detect the reaction pathway. After a 3-way branch migration and then a 4-way branch migration process, the fluorophore and quencher in the reporter separate. The gradual increase of fluorescence only in the presence of both an upstream fuel and a signal that can spuriously react with the fuel is consistent with the described spurious displacement pathway.

D. Leak reduction compared to designs with only an entropy barrier. Prior work based leak reduction on a purely entropic barrier wherein multiple separate complexes needed to co-localize to generate leak (16, 20). All of our *EN* designs have an additional enthalpic barrier to leak requiring the decrease in the number of bonds. To experimentally compare leak reduction of our *EN* design to the prior work, we chose one of the parameter pairs $N = 5$ and *shift* $= 3$. This parameter set has the desired property that leak requires at least two units of enthalpic penalty (breaking two bonds) compared with the maximum-bond configuration (Theorem 3.2).

We compare with the typical leaky translator (*SLD* design) and the previously reported leak reduction method based exclusively on an entropic penalty (*DLD* design) (16, 20) (Figure 7a). Note that the *DLD* design refers to the *NLD* design when $N = 2$.

In our experiments, these three translator designs share the same sequence space. All the fuels have the same toeholds. The *DLD* and the *EN* design share the same reporter. The detailed comparison for the designs at the sequence level is shown in Figure S8. The fuel concentrations used here are 10–100 times larger than the typical concentration used in strand displacement systems. Figure 7b compares the leak concentration at thermodynamic equilibrium in the absence of input signal, which sets an upper bound of the total leak for an isolated translator. To achieve thermodynamic equilibrium, fuels and reporter were slowly annealed together from 90 °C to 25 °C within 18 h. The *EN* design has an order of magnitude less leak than the *DLD* design and about two orders of magnitude less than the *SLD* design.

Although enthalpy-neutral strand displacement is not driven by the formation of new bonds as are the *SLD* and *DLD* designs, we observed that the triggered signal is only around three times lower than full triggering. Thus the signal-to-leak ratio of the *EN* design is 30 times higher than of the *SLD* design and four times higher than of the *DLD* design (Figure 7b inset).

E. Desired triggering and leak in longer cascades. Beyond a single translator, we wanted to know (1) how much leak (without input signal) and (2) desired output signal (with input signal) a translator cascade can generate as the number of translators increases.

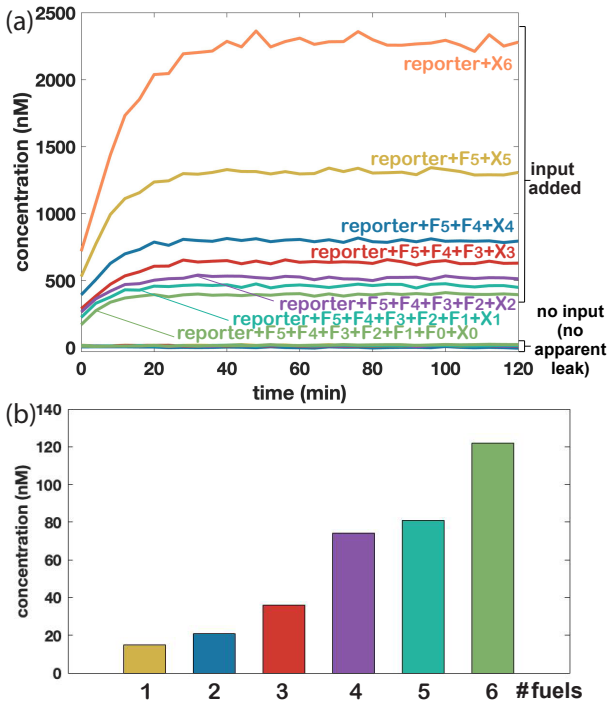


Fig. 8. Kinetics and thermodynamic equilibrium of translator cascades with the *EN* design ($N = 5$, *shift* $= 3$). (a) Kinetic behaviour in the presence and absence of input signal, for cascades of different length. (b) The total amount of leak in the absence of input signal at thermodynamic equilibrium measured after annealing, for cascades of different length. The concentrations of the reporter and the fuels were around $5\ \mu$M. The concentration of each input was $2.5\ \mu$M. The reaction temperature was $37\ ^\circ$C. Note that X_i in the figure refers to the input to fuel F_i, and X_6 is the input to the reporter. (Note that the actual reporter used is one domain shorter than the $N = 5$ design and the overhand on F_5 was shortened to compensate (Figure S9), which is not expected to affect the completion level.)

Figure 8a shows that in the time period of the experiment, in the absence of input signal strand, the translator cascades of 1 to 6 fuels (1 to 3 translators) all show no apparent leak. In the presence of input signal, the completion level decreases with the number of layers. However, as more layers are added, the completion level does not decrease linearly, and indeed seems to approach an asymptote—behavior consistent with the theoretical prediction of Section 1.C.

Figure 8b studies the leak of translator cascades of varying depth at thermodynamic equilibrium. The leak at equilibrium increases as the number of fuels increases. However, even if there are 6 fuels (3 translators), the total leak concentration is still less than 3% of the fuel concentration. Indeed, it is about 5 times less than the leak concentration for a *single* translator in the *DLD* design, and it is an order of magnitude less than that in the *SLD* design.

These results suggest that the absolute leak concentration of the *EN* design is significantly less than the previously reported *DLD* design. Although the absolute amount of triggered signal of the *EN* design is lower than that of the *DLD* design, the relative triggered signal-to-leak ratio of the *EN* design is still higher than the *DLD* design. Our results suggest that the *EN* design could be preferable, especially when a system requires absolutely smaller leak, such as when concatenating the translators with downstream catalytic or auto-catalytic systems.

6. Discussion

The full diversity of enthalpy-neutral linear-topology strand displacement cascades is accessible by varying two parameters N and *shift*, with rigorous guarantees on the enthalpic and entropic penalties to leak determined by these parameters. In contrast to previously reported methods for arbitrarily decreasing leak which rely on entropic barrier arguments, we describe how the enthalpic and entropic penalties to leak can be raised arbitrarily. The enthalpic penalty argument is particularly germane for the high concentration regime where the entropy penalty for joining complexes is smaller. We further prove that certain parameter values result in other desirable properties like no spurious displacement and no toehold occlusion. Since no design satisfies all desired properties, understanding the taxonomy of designs is necessarily to tune the tradeoffs and make proper design choices.

If the clamps in the *NLD* design (16, 20) are extended to toehold size, the *NLD* design can be categorized into one class of the *EN* design. For example, the *SLD* design would be categorized as $N = 4$ and *shift* $= 3$, and the *DLD* design would be categorized as $N = 6$ and *shift* $= 3$. According to the taxonomy developed here, this extended *NLD* design for any redundancy N_r has toehold occlusion and spurious strand displacement. In this sense, the *EN* design is broader than the *NLD* design with toehold-size clamps, allowing for more flexibility in balancing desired and undesired properties.

In addition to theoretical analysis, we experimentally tested designs with different energy penalties to leak, confirming different degrees of leak reduction and improvement over the previous *NLD* design. Our experiments also corroborated that slower kinetics due to toehold occlusion in otherwise similar designs. Finally, in identifying spurious displacement reaction pathways, the experiments confirm the rationale for studying spurious displacement.

The scale of current strand displacement systems exceeds the capabilities of thermodynamic modeling tools (25) and (even more so) kinetic simulators (26). In contrast, combinatorial arguments like those espoused in this work apply to infinitely many, arbitrarily large systems.

A. Orthogonal toeholds. Our argument relies on the strong assumption that all domains are orthogonal. In reality, given the limited size of a domain (that of a toehold), as the number of distinct domains increases, it is not possible to make all the domains orthogonal. Nonetheless we note that in certain cases having the same sequence in multiple domain positions seems to pose no problem. (For example, in designs where N is not a multiple of *shift*, all the balance domains could have the same sequences.) Future work could further explore how to assign the same domains without undesired interactions and how the number of orthogonal domains needed scales with the length of double-stranded region N.

B. Reconfiguration. Spurious strand displacements can be more complicated when multiple species are involved. Some may result in the formation of complex multi-stranded structures. In Figure 9, we show contrasting examples of possible reconfiguration after spurious strand displacement. The first example requires a bimolecular reaction to undo, while unimolecular reconfiguration is sufficient for the second. Note that to the first approximation, even at the "high" concen-

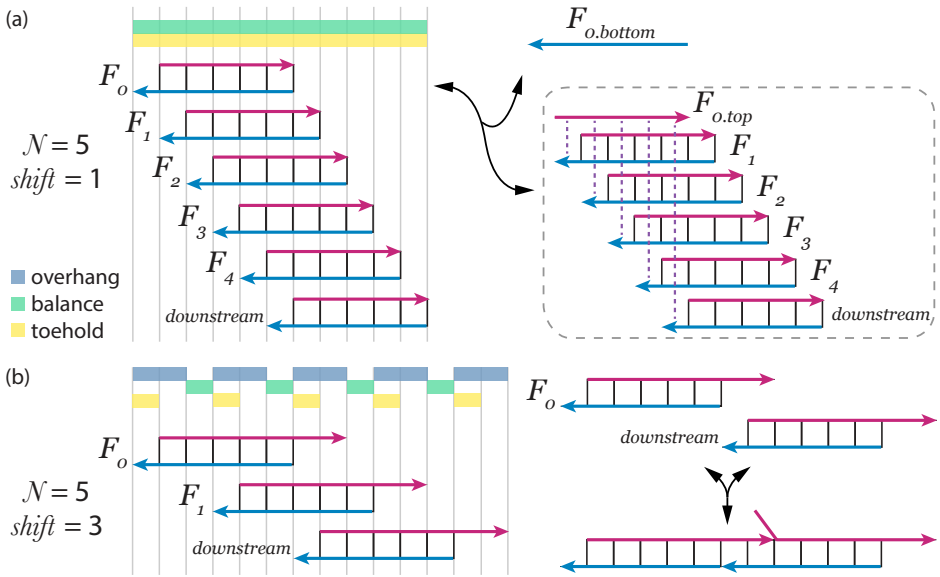


Fig. 9. Examples of configurations reachable with spurious strand displacement. (a) Spurious strand displacement results in a configuration that requires a bimolecular reaction to undo. In the absence of input signal, multiple spurious displacement events could result in all of the bottom domains of one fuel being displaced by toeholds of other fuels. This results in a free (unbound) bottom strand. Although the multiple toeless displacement events to cause this reconfiguration are unlikely, once formed, it requires a slow bimolecular reaction to undo. The dashed gray box indicates the strands in the box are in one complex. The dash purple lines indicate bound domains. (b) Spurious strand displacement results in a configuration that requires a unimolecular reaction to undo.

trations used here, bimolecular reactions are relatively slower than unimolecular reactions.

These two examples suggest that there may be a separation between designs with hard-to-reverse spurious reconfiguration (requiring bimolecular steps to undo) and those in which any spurious reconfiguration can be quickly undone (unimolecular steps are sufficient). Indeed, in certain cases, we can prove that such hard-to-reverse displacement is thermodynamically penalized by connecting it to active output (see Section S6). Unfortunately, our argument holds only when *shift* is large leading to a small thermodynamic penalty; obtaining the penalty for arbitrary *shift* remains an area for further research.

Materials and Methods

The proofs are in SI. All DNA sequences are listed in SI. DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT). Fuel and reporter species were annealed in TAE/Mg^{2+} buffer. After annealing, the fuel species were PAGE purified. Fluorescence measurements were taken after mixing all the species.

ACKNOWLEDGMENTS. The authors thank Cameron Chalk for reading and commenting on the manuscript. B.W. and D.S. were supported by NSF grants CCF-1618895 and CCF-1652824. C.T. was supported by NSF grants CCF-1317694 and CCF-2106695.

1. B Wang, C Thachuk, AD Ellington, D Soloveichik, The design space of strand displacement cascades with toehold-size clamps in *DNA Computing and Molecular Programming*, eds. R Brijder, L Qian. (Springer International Publishing, Cham), pp. 64–81 (2017).
2. DY Zhang, G Seelig, Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.* **3**, 103–113 (2011).
3. FC Simmel, B Yurke, HR Singh, Principles and applications of nucleic acid strand displacement reactions. *Chem. reviews* **119**, 6326–6369 (2019).
4. DY Zhang, E Winfree, Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* **131**, 17303–17314 (2009).
5. DY Zhang, AJ Turberfield, B Yurke, E Winfree, Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA. *Science* **318**, 1121–1125 (2007).
6. L Qian, E Winfree, Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**, 1196–1201 (2011).
7. L Qian, E Winfree, J Bruck, Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368–372 (2011).
8. KM Cherry, L Qian, Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* **559**, 370–376 (2018).
9. AJ Thubagere, et al., A cargo-sorting DNA robot. *Science* **357**, eaan6558 (2017).
10. YJ Chen, et al., Programmable chemical controllers made from DNA. *Nat. Nanotechnol.* **8**, 755–762 (2013).
11. N Srinivas, J Parkin, G Seelig, E Winfree, D Soloveichik, Enzyme-free nucleic acid dynamical systems. *Science* **358** (2017).
12. DY Zhang, SX Chen, P Yin, Optimizing the specificity of nucleic acid hybridization. *Nat. Chem.* **4**, 208–214 (2012).
13. JS Wang, DY Zhang, Simulation-guided DNA probe design for consistently ultraspecific hybridization. *Nat. Chem.* **7**, 545–553 (2015).
14. JS Wang, YH Yan, DY Zhang, Modular probes for enriching and detecting complex nucleic acid sequences. *Nat. Chem.* **9**, 1222–1228 (2017).
15. M Rudchenko, et al., Autonomous molecular cascades for evaluation of cell surfaces. *Nat. Nanotechnol.* **8**, 580–586 (2013).
16. B Wang, C Thachuk, AD Ellington, E Winfree, D Soloveichik, Effective design principles for leakless strand displacement systems. *Proc. Natl. Acad. Sci. United States Am.* **115**, E12182–E12191 (2018).
17. G Seelig, D Soloveichik, DY Zhang, E Winfree, Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–1588 (2006).
18. YS Jiang, S Bhadra, B Li, AD Ellington, Mismatches Improve the Performance of Strand-Displacement Nucleic Acid Circuits. *Angewandte Chemie* **126**, 1876–1879 (2014).
19. X Olson, et al., Availability: A metric for nucleic acid strand displacement systems. *ACS Synth. Biol.* **6**, 84–93 (2017).
20. C Thachuk, E Winfree, D Soloveichik, Leakless DNA strand displacement systems in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* (Springer Verlag), Vol. 9211, pp. 133–153 (2015).
21. D Doty, TA Rogers, D Soloveichik, C Thachuk, D Woods, Thermodynamic Binding Networks in *DNA Computing and Molecular Programming*, eds. R Brijder, L Qian. (Springer International Publishing, Cham) Vol. 10467, pp. 249–266 (2017).
22. K Breik, C Chalk, D Doty, D Haley, D Soloveichik, Programming substrate-independent kinetic barriers with thermodynamic binding networks. *IEEE/ACM Transactions on Comput. Biol. Bioinforma.* **18**, 283–295 (2019).
23. K Breik, C Thachuk, M Heule, D Soloveichik, Computing properties of stable configurations of thermodynamic binding networks. *Theor. Comput. Sci.* **785**, 17–29 (2019).
24. J SantaLucia, D Hicks, The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 415–440 (2004).
25. JN Zadeh, et al., NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173 (2011).
26. JM Schaeffer, C Thachuk, E Winfree, Stochastic simulation of the kinetics of multiple interacting nucleic acid strands in *International Workshop on DNA-Based Computers.* (Springer), pp. 194–211 (2015).