Fast and Robust Strand Displacement Cascades via Systematic Design Strategies

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

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

Chris Thachuk[2]
Paul G. Allen School of Computer Science & Engineering,
University of Washington, Seattle, WA, USA

Abstract

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

2012 ACM Subject Classification Applied computing → Chemistry; Computer systems organization → Molecular computing

Keywords and phrases DNA strand displacement, Energy barriers, Kinetics

Digital Object Identifier 10.4230/LIPIcs.DNA.2022.1

Funding This work was supported by an NSF grant (CCF 2106695) and a Faculty Early Career Development Award from NSF (CCF 2143227).

1 Introduction

One goal of molecular programming is to design chemical systems that not only store and process information, but are able to manipulate matter with nanometer precision, to sense (bio-) chemical signals from their environment, to perform robust, complex and energy-efficient computation, and to actuate a physical response. This is not an easy goal. Yet, inspiring demonstrations of complex, enzyme-free circuits based on DNA strand displacement (DSD) [26] show the promise of this technology to realize the equivalent of a chemical central processing unit. While engineered DSD reactions, inspired by strand displacement in genetic recombination[14, 15], were studied as early as the 1980s [3] the groundwork for modern engineering of DSD cascades and other more complicated dynamic systems followed nearly twenty years later [19, 25].

[1] Authors contributed equally and are listed alphabetically by last name.
[2] To whom correspondence may be sent.

Sequence-independent "domain level" strategies have emerged to improve the design of high fidelity DNA strand displacement reactions and cascades. For example, recent work on *leakless* DSD cascades [18, 20, 21] have shown how these systems can be made arbitrarily robust: the rate of spurious leak reactions – that produce output in the absence of correct input – can be decreased exponentially by a linear increase to the length of the cascade. However, this increased robustness comes at a cost: designed reaction pathways have slower kinetics. This slowdown of desired kinetics can result from longer cascades, from cascades that are a series of reversible displacements when toehold-sized clamps are employed [20], and from unproductive reactions attributed to the domains shared between multiple signal strands within the same cascade (see Figure 1b).

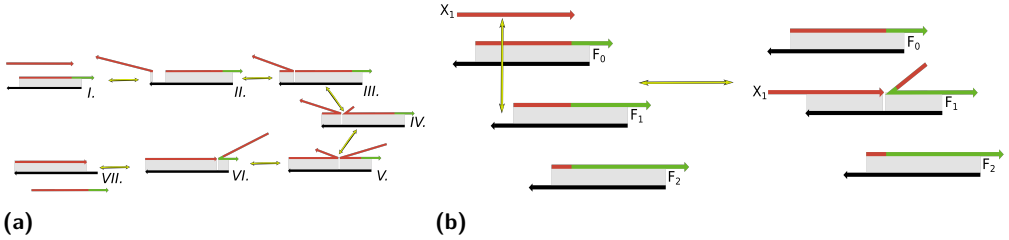


Figure 1 **(a)** A visual representation of a toehold mediated strand exchange reaction. *I.* Initial system: translator gate with invader strand (red), incumbent strand (red/green), and substrate strand (black). *II.* Initial base pair formed between translator toehold and signal strand. *III.* Toehold fully bound to signal strand. *IV.* First base pair in translator frays, incurring a free energy penalty ΔG_p for branch migration initiation [16, 10]. *V.* Branch migration occurs: base pairs between the translator strands are replaced by base pairs between the signal strand and the substrate strand (black). *VI.* Branch migration completes: incumbent strand is bound to toehold sized domain on substrate strand. *VII.* Incumbent strand dissociates from complex. **(b)** While systematic redundancy introduced in leakless cascades inhibits spurious displacement the resulting overlap in sequence space among complexes introduce new pathways for spurious invasion. Spurious invasion events do not necessarily lead to spurious displacement (*i.e.*, leak). However, these unproductive pathways sequester chemical species that are kinetically relevant to the designed reaction pathway decreasing its effective rate. The *fuel* complexes F_0, F_1 and F_2 comprise a displacement cascade that mediates the formal reaction $X_1 \rightarrow Y_1$. Illustrated is the (unproductive) spurious invasion of F_1 by the cascade input signal X_1; similarly X_1 can occlude the toehold of and spuriously invade F_2.

In parallel, sequence-dependent strategies that introduce and/or eliminate Watson-Crick-Franklin [23, 9] base-pairing violations, or "mismatches", have been studied as an effective means for tuning the kinetics and thermodynamics of DSD reactions. A study from as early as 1986 investigated the impact of a single base-pair mismatch as a parameter for engineering strand displacement reactions [3]. The effects of mismatch elimination (and introduction) in strand displacement have been widely studied [6, 1, 12] and include a thermodynamic driving force from mismatch elimination, as well as reaction rate changes that are highly dependent on the distance of the mismatch position from the site of invasion. Mismatches have been used both strategically [24, 7] and systematically [11, 13] to alter reaction kinetics [24] and thermodynamic driving forces [7, 12, 5], as a mechanism for single nucleotide polymorphism (SNP) identification [8, 17], and as a means to combat leak [13, 11, 7].

While strategies for systematic domain-level optimization of DSD circuits and those for sequence-level optimization of DSD circuits were developed primarily in parallel, they are not mutually exclusive. The drawbacks introduced to designed reaction pathways when employing domain level strategies such as leakless motifs and/or the use of clamp domains, namely slowed kinetics and/or a reduction in thermodynamic driving force, might be remedied by

sequence-level strategies. Here we demonstrate a synthesis of these two strategies to arrive at cascade designs that are robust by virtue of their domain level redundancy, yet have fast kinetics due to a systematic sequence level optimization strategy based on strategic mismatch elimination and introduction. By composing the state-of-the-art for domain level and sequence level design of DSD systems we demonstrate a rich design space for fast and robust DSD cascades. We support the efficacy of this strategy with theory based on thermodynamic modeling and preliminary experimental verification.

2 Systematic mismatch strategies for leakless cascades

Due to the domain level redundancy, introduced by *leakless* cascades in order to combat leak [18, 20], unproductive invasion is possible – signal strands present initially as input, or produced via displacement of incumbent strands, can interact with multiple fuel complexes. For example, Figure 1b illustrates how the input strand X_1 can spuriously invade fuel complex F_1 – not shown is that X_1 can also spuriously invade fuel complex F_2. While overall these unproductive invasions are unlikely to cause leak events they do create kinetic slowdown; a signal strand (transiently) bound to the wrong fuel is signal not currently propagating through the cascade and is also a source of occlusion of its bound fuel and thus a source of inhibition of other signal propagation. A confounding factor in kinetic slowdown is the use of toehold-sized clamps. This design choice has been demonstrated to suppress leak to even lower rates in leakless cascades [20], but at the cost of thermodynamic driving force since each designed reaction is based on reversible toehold exchange (see Figure 1a and Figure 5).

To combat these problems we introduce a systematic sequence level modification strategy to introduce *sentinel* positions.

▶ **Definition 1** (sentinel position). *Any position i such that an intended invader forms a Watson-Crick-Franklin base pair and every spurious invader introduces a mismatch.*

Table 1 Sequence-level strategies to meet design goals. Reasonable choices for parameters α and β are discussed in Section 2.1.

Goal	Strategy
1. Barrier to spurious toehold binding	Sentinel position in each toehold domain
2. Additional barrier to spurious invasion	Sentinel positions "close" to helix ends
3. Additional thermodynamic drive	Sentinel positions in each fuel complex
4. Avoid increasing leak	No internal loops within α nucleotides of any helix end
	No internal loops within β nucleotides of each other

Our aim is to systematically create energy barriers that must be encountered early in any unproductive invasion pathway. Furthermore, these modifications should not be at the expense of cascade robustness and, when possible, should introduce kinetic improvements and additional thermodynamic driving force for designed reactions. While the strategy is generally applicable to any system with redundant domains across a cascade we will focus on translator cascades for the sake of clarity. Consider the leakless translator cascade with toehold-sized clamps to emulate the reaction $X_1 \to Y_1$ illustrated in Figure 2 (Step 0). These additional clamps provably increase the barrier to leak reactions [22], but individual steps in the cascade are reversible; similarly, the overall cascade is reversible and thus emulates the reaction $X_1 \leftrightarrow Y_1$ with forward and reverse rate constants dependent on the strength of

forward and reverse toeholds. Can we keep the robustness benefit provided by the redundant domains *and* these additional clamps yet drive the cascade forward as if implemented as a series of "irreversible" reactions without clamps (*i.e.*, with heavily favored forward rates)? Can we significantly decrease the rates of unproductive reactions? As it turns out these goals can be simultaneously met.

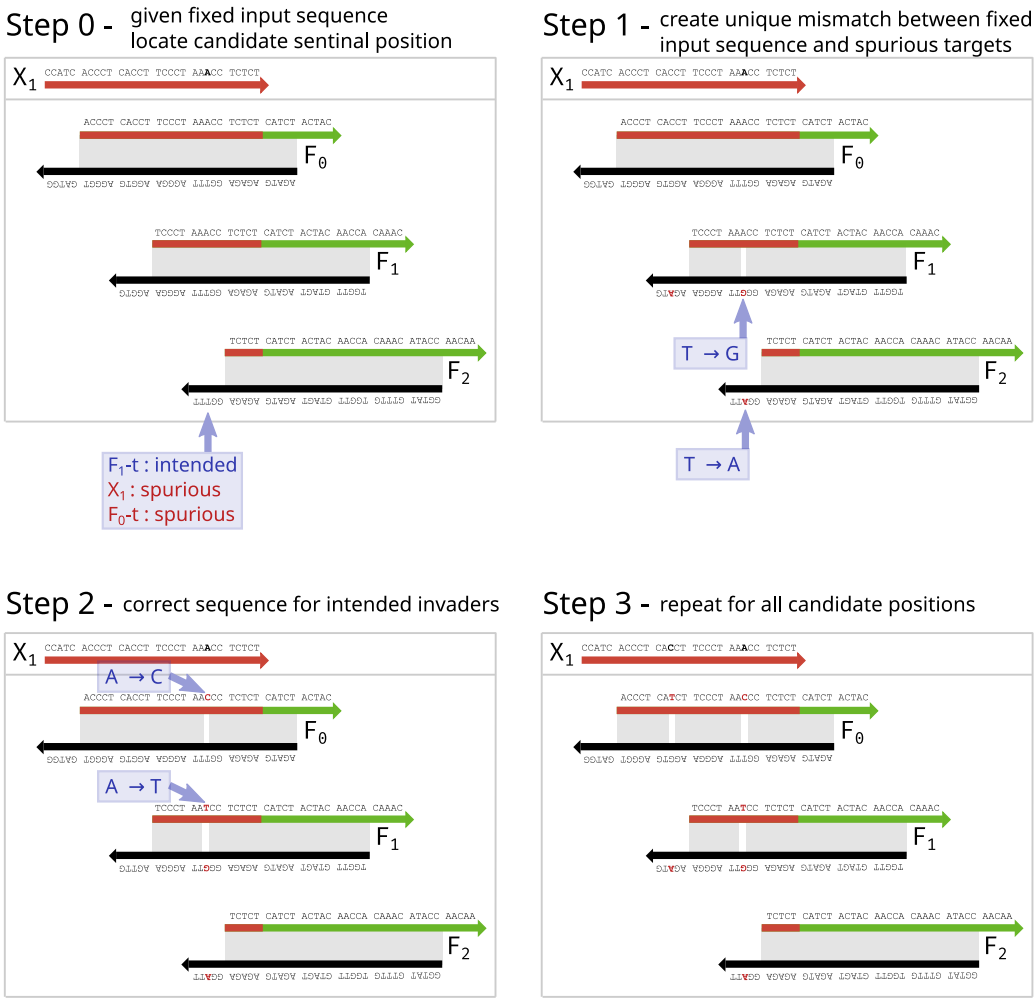


Figure 2 A general strategy for creating sentinel positions given fixed input and output sequences of a cascade. (Step 0) A candidate position i is identified. (Step 1) Unique mismatches are introduced into substrate (bottom) strands that intersect i and are not designed to interact with the fixed input. (Step 2) Sequences are corrected for intended (top strand) invaders. (Step 3) The process is repeated for candidates intersecting the fixed input. (Step 4 – not depicted) Symmetric process run to introduce sentinel positions with respect to the output sequence to guard against spurious invasion in the reverse pathway; see Figure 3 for a complete example with fixed input and output sequences.

Sentinel positions will introduce mismatches within fuel complexes. Obvious candidate positions are those intersecting toeholds within a cascade since weakening the toehold binding of spurious invaders will decrease the rate of unproductive reactions (see Section 3). However, these positions should not be chosen if they intersect fuel complexes near the end of a helix as fraying could be exacerbated by introduced mismatches and result in increased rates of leak. In those cases, positions can be chosen that intersect further away from helix ends

(*e.g.* at least a "toehold"-sized domain away). Figure 2 demonstrates how sentinel positions can be created with respect to a fixed input and output sequence of a cascade. The aim of this strategy is to internally optimize a cascade that implements a formal reaction without changing its interface to other system components (*i.e.*, the input sequence it accepts and the output sequence it produces). We note that any change to the sequences of signal strands has the potential to create spurious events with other cascades in the system, or create unwanted secondary structure. These can be incorporated as hard constraints when implementing the sentinel position design strategy.

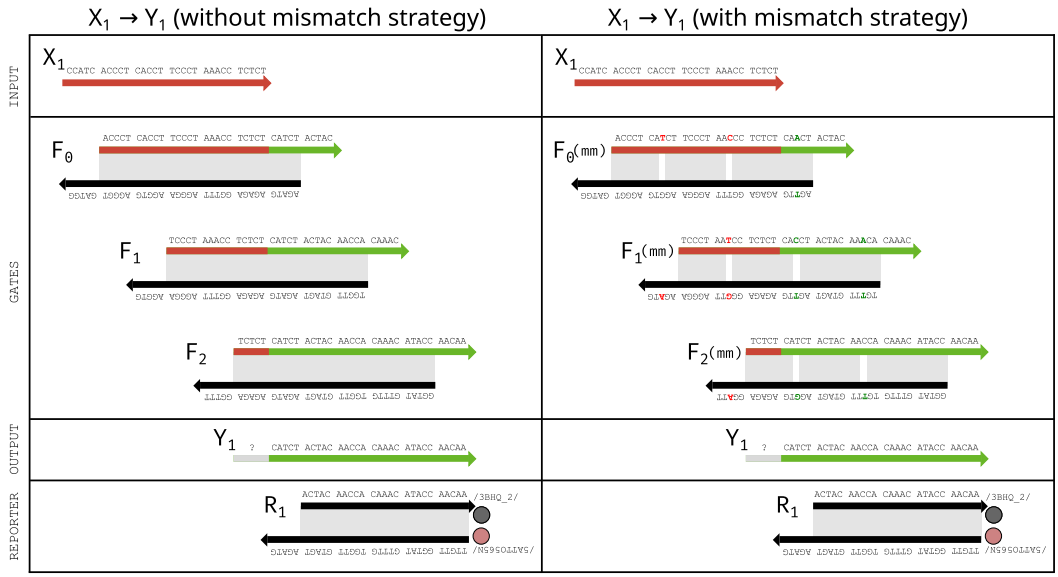


Figure 3 Domain level and sequence level designs for a $X_1 \rightarrow Y_1$ translator without mismatch strategy (left) and with mismatch strategy (right). Both systems were designed to share the same input "trigger" strand and same output strand (and reporter). Locations and sequences of mismatch positions are indicated around gaps in DNA duplexes.

Figure 3 gives a complete example of introducing sentinel positions that optimizes with respect to both the forward and (undesirable) reverse reaction pathways given a fixed input and output sequence. This system will serve as our detailed case study in subsequent sections.

2.1 Fragile regions of design space

For a collection of sequence modifications of a given DNA cascade there exist many poor sequence designs. Modification within a particular cascade may introduce unwanted secondary structure or introduce spurious interactions with species of another cascade within the same system. These conflicts can be modelled as hard or soft constraints to be evaluated when considering a modification candidate. However, there are certain design choices that are likely to yield poor candidates and can be eliminated from consideration entirely. For example, choosing a sentinel position too close to the end of a helix will result in increased fraying that in turn could lead to increased leak rates (see Figure 4a). Similarly, placing neighboring sentinel positions too close could result in the intervening duplex structure being destabilized and merging into a large internal loop, again leading to increased leak rates (see Figure 4c). We parameterize these distances in our overall design considerations.

▶ **Definition 2** (α). *The minimum distance of any sentinel position from a helix end (including nicks).*

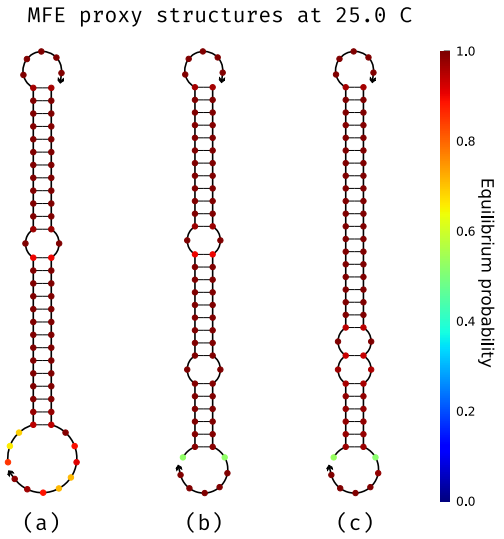


Figure 4 Stability of gate complexes and their propensity to fray at helix ends or merge neighboring mismatches into larger loops for candidate designs can depend on parameters α and β. Three candidate designs generated with (a) $\alpha = 2, \beta = 5$, (b) $\alpha = 5, \beta = 5$, and (c) $\alpha = 5, \beta = 2$.

▶ **Definition 3** (β). *The minimum distance between neighboring sentinel positions.*

Both of these distance parameters can be chosen for a particular cascade or particular reaction condition based on simple criteria: at what distance do the intended secondary structures of fuel complexes have unacceptable ensemble defect [4]?

2.2 A naive sentinel design algorithm

Beginning from reference sequences the sentinel modification strategy naturally yields a naive algorithm: for every possible combination of positions that span the interval covered by a cascade, such that no position is within distance α of a helix end and no neighboring positions are within distance β of each other, evaluate all sequence modifications that yield valid sentinel positions and disfavor each spurious invasion relative to designed reactions. Any candidate design can be rejected for violating user-defined constraints such as propensity to form undesired intra- or inter-molecular secondary structure, or by containing forbidden sequence motifs. As we will see in Section 3.3 the space of candidate designs is sufficiently small in practice that it is tractable to enumerate and evaluate each one, yet the number of candidate designs that yield improvements, relative to the unmodified cascade, is large.

3 Thermodynamic modeling

The thermodynamic landscapes of conventional strand displacement and exchange reactions based on our translator system (Figure 3) were modeled using NUPACK [4, 2] at typical conditions (25 °C, 12.5 mmol L^{-1} Mg^{2+}, 50 mmol L^{-1} Na^{+}). Free energy values (ΔG) were plotted for each step in the conventional pathways for each reaction under study and incorporated a penalty for branch migration initiation $\Delta G_{\text{p}} = 2$ kcal mol^{-1} [10, 16]. Unless otherwise noted we use the structure free energy of an unstructured invader and fuel complex with a toehold-sized clamp (F_2), both at 10 nmol L^{-1} concentration, as the reference microstate defined to have $\Delta G = 0$ kcal mol^{-1}.

3.1 Toehold-mediated strand exchange (TMSE)

A common strategy to reduce leak in DSD systems is to introduce toehold-sized clamps. With respect to our translator gate F_2, when moving from a displacement reaction (Figure 5, blue) to an exchange reaction (Figure 5, black) the structural free energy of the initial microstate is improved, but at the cost of $\approx 6\,\mathrm{kcal\,mol^{-1}}$ forward thermodynamic driving force. This difference ($\Delta\Delta G_{\text{exchange}} - \Delta\Delta G_{\text{displacement}}$) translates directly from the loss in additional base-pairs formed via displacement.

While strand exchange is typically thermodynamically neutral, strand exchange in a translator utilizing the sentinel mismatch strategy is thermodynamically favorable. Consider the specific comparison in Figure 6. The initial structural free energy of $F_2(mm)$ is higher than that of F_2, since $F_2(mm)$ has fewer paired bases, as well as mismatches destabilizing its helix. Both systems pay a concentration-dependent entropic penalty (calculated based on experimental concentration of each reactant at $10\,\mathrm{nmol\,L^{-1}}$) when base 1 of the toehold pairs (microstate II). Throughout invasion of $F_2(mm)$ by $F_1(mm)$-t (red pathway) the system drops in energy each time a mismatch is replaced by a Watson-Crick-Franklin base pair. After the last mismatch has been replaced, both strand exchange systems follow essentially the same energy path throughout the remainder of the invasion and toehold dissociation (microstates VI-VII). Without the mismatch strategy the change in structural free energy $\Delta\Delta G_{\text{no mismatch}} = -0.71\,\mathrm{kcal\,mol^{-1}}$, while with the mismatch strategy $\Delta\Delta G_{\text{mismatch}} = -10.71\,\mathrm{kcal\,mol^{-1}}$. Additionally, the reverse reaction of the translator using the mismatch strategy has a significant uphill energy activation barrier ($\Delta\Delta G_{\text{rev(mismatch)}} = 19.75\,\mathrm{kcal\,mol^{-1}}$), while the activation energy for the reverse reaction of the strand exchange system without mismatches ($\Delta\Delta G_{\text{rev(no mismatch)}} = 9.75\,\mathrm{kcal\,mol^{-1}}$) is largely due to the entropic penalty of toehold binding ($9.04\,\mathrm{kcal\,mol^{-1}}$). This additional driving force has been studied in greater detail in other work [5, 10].

When applied systematically within a cascade, as in our sentinel position design strategy, mismatches raise the energy barrier for unproductive strand invasions in leakless systems. Sequence redundancies in the leakless system allow upstream signal strands to partially invade downstream translators via toehold binding, and subsequent displacement of matching sequence (see Figure 1). The systematic mismatch strategy ensures that signal strands will mismatch with every translator except their intended target. Figure 7 shows the conventional pathways of invasion of $F_2(mm)$ by its intended invader ($F_1(mm)$-t, blue) and two other spurious invaders ($F_0(mm)$-t, green and X_1, orange). Note that without systematic mismatches the conventional invasion pathway for each spurious invader would simply be a prefix of the intended conventional invasion pathway.

We investigated the effect of the mismatch strategy in greater detail on invasion pathways of F_2 and $F_2(mm)$ when reactants were both at low concentrations ($10\,\mathrm{nmol\,L^{-1}}$) and when both were at high concentrations ($10\,\mathrm{\mu mol\,L^{-1}}$). For unproductive reactions only the most favorable spurious invader was considered – that is, the corresponding spurious invader with the largest overlap with the fuel complex; F_0-t in the control case and $F_0(mm)$-t when using the mismatch strategy. In the no mismatch case at low concentration (Figure 8a) the energy pathway for both invaders is identical until base 20 when the spurious invader completes its invasion and recovers the branch migration initiation penalty $\Delta\Delta G_{\text{p}}$ (Figure 8b, microstate IV). In contrast, when the mismatch strategy is employed at low concentration the spurious invasion pathway (i) necessarily deviates from that of the intended invasion, (ii) must overcome a higher energy barrier to invasion (Figure 8b microstate III vs microstate II), and (iii) terminates at a significantly less favorable resting microstate compared with the free energy of the intended reaction having made similar invasion progress (Figure 8b microstate IV).

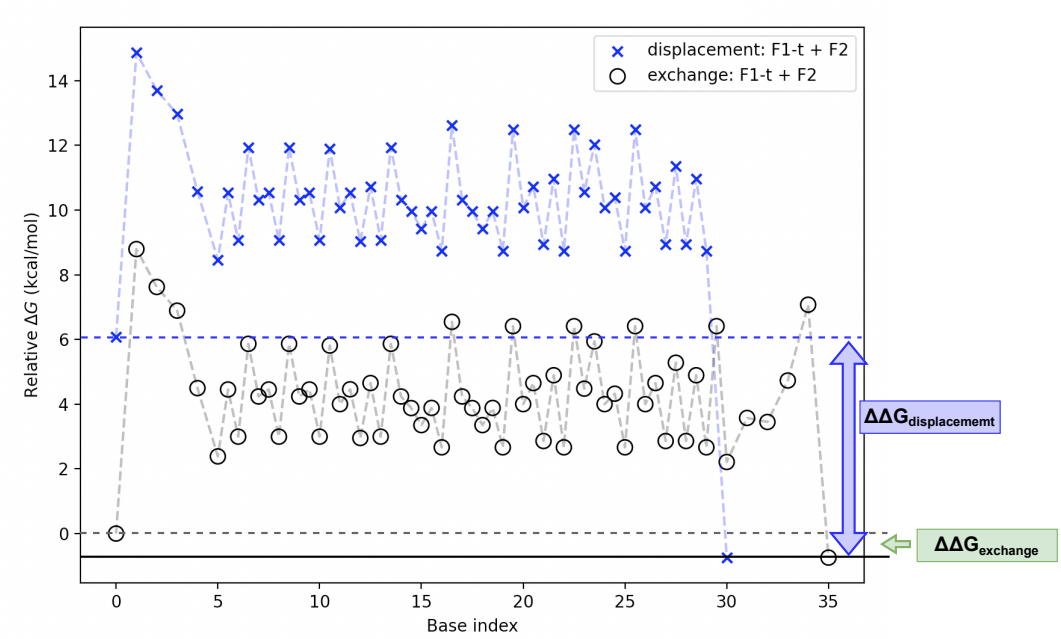


Figure 5 (black) The energy landscape for a conventional toehold mediated strand exchange (TMSE) between complex F_2 and its intended trigger strand F_1-t of Figure 3 (left). **(blue)** A similar landscape is shown for a modified F_2 complex that lacks a toehold-sized clamp thus resulting in a toehold mediated strand displacement (TMSD) reaction.

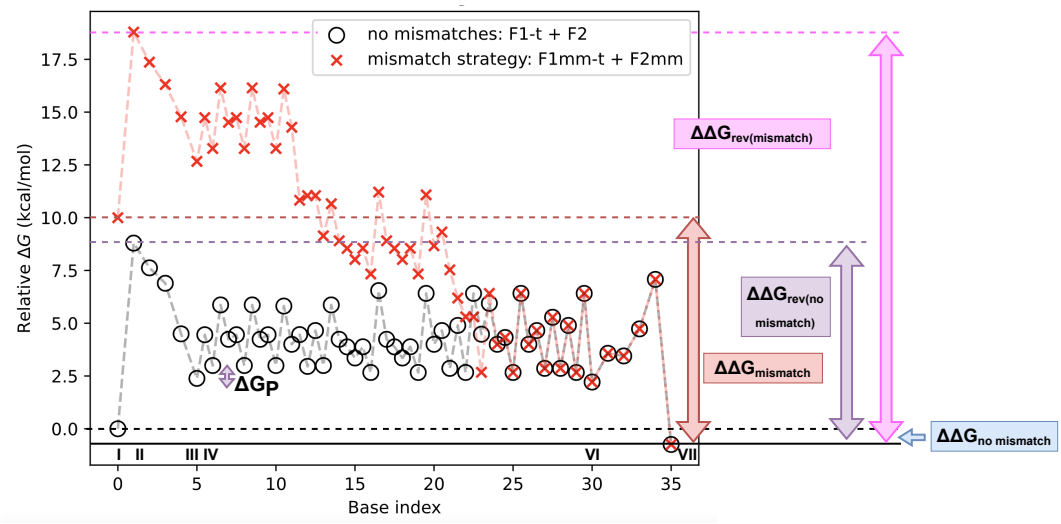


Figure 6 (red) The energy landscape for a conventional toehold mediated strand exchange (TMSE) between complex $F_2(mm)$ and its intended trigger strand $F_1(mm)$-t that results in two mismatch elimination events. **(black)** A similar landscape is shown for F_2 and invader F_1-t without a mismatch strategy. Roman numerals reference system microstates described in Figure 1a.

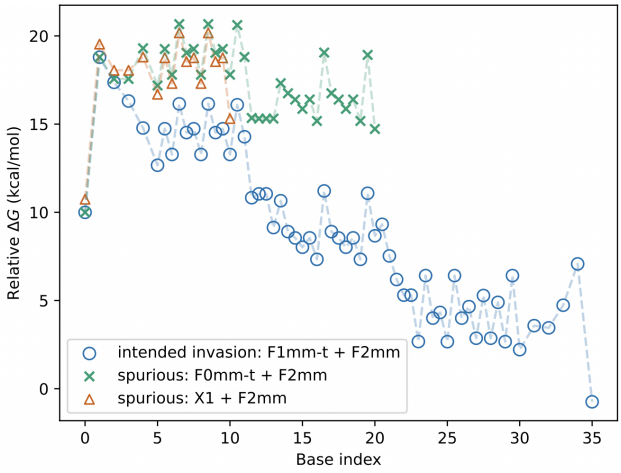
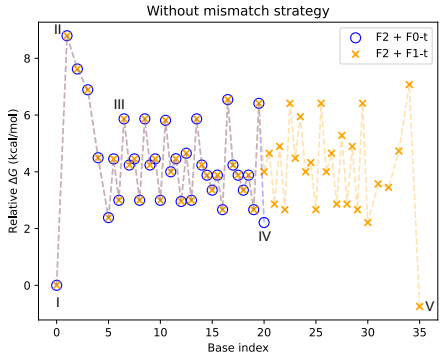
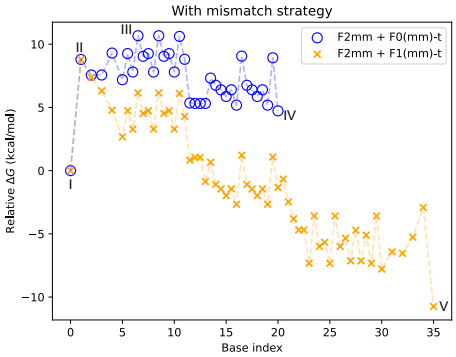


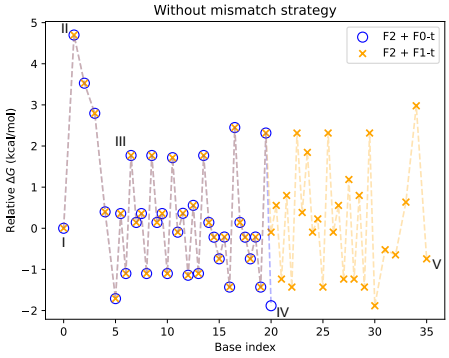
Figure 7 Mismatch strategy for TMSE in $F_2(mm)$ complex. Mismatches disfavor spurious invaders $F_0(mm)$-t (green) and X_1 (orange), and favor the intended invader $F_1(mm)$-t (blue).



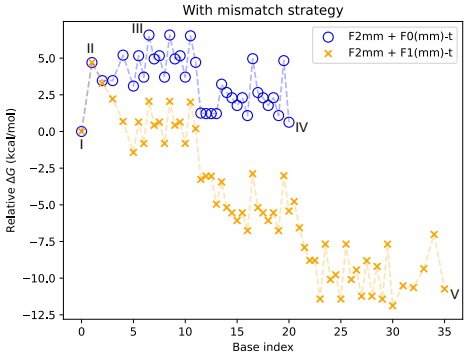
(a) Low concentration without mismatch strategy.



(b) Low concentration with mismatch strategy.



(c) High concentration without mismatch strategy.



(d) High concentration with mismatch strategy.

Figure 8 Comparison of the conventional pathways for intended and spurious invasion without and with the mismatch strategy – left column and right column, respectively – and at low concentration ($10\,\mathrm{nmol\,L^{-1}}$) and at high concentration ($10\,\mu\mathrm{mol\,L^{-1}}$) – top row and bottom row, respectively. Deviating from prior convention, note that the free energies reported in **(c)** and **(d)** are relative to structure free energy of $F_2(mm)$ and an unstructured invader.

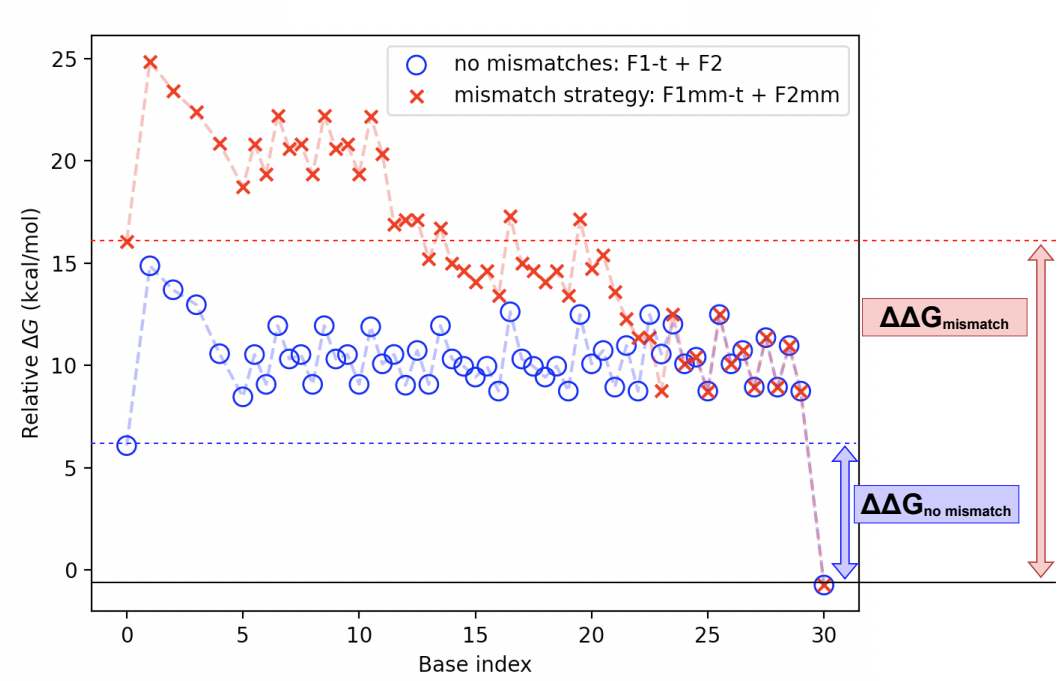


Figure 9 The thermodynamic driving force already present in TMSD **(blue)** can be significantly increased when intended invaders eliminate multiple mismatches **(red)**.

These same trends hold at higher concentrations (see Figure 8c and Figure 8d); however, there is an important distinction to consider. At the high micromolar concentrations typically used to increase the speed of leakless strand exchange ($1\,\mu\text{mol}\,\text{L}^{-1}$–$10\,\mu\text{mol}\,\text{L}^{-1}$) [22, 21] the entropic penalty of toehold binding is reduced (microstate II in Figures 8c and 8d). Modeling of translator gate F_2 at high concentration (Figure 8c) shows that not only is it equally as favorable for a spurious invader to initiate branch migration as the desired invader, but that many of the microstates during spurious invasion (*e.g.*, microstate IV) are energetically favorable to the initial configuration (*i.e.*, microstate I). Therefore, as the concentration of the non mismatch cascades increase – and thus the thermodynamic favorability of microstates with occluded toeholds and sequestered signal strands – so too do the kinetic inhibitions caused by (unproductive) spurious invasions. In contrast, even at high concentrations, the mismatch strategy remedies these issues by (i) providing an additional energy barrier to spurious invasion (Figure 8d, microstate III), (ii) ensuring all microstates of spurious invasion are unfavorable relative to the initial configuration (microstates II-IV compared with microstate I, Figure 8d), and (iii) maintaining a thermodynamic driving force ($\Delta\Delta G_{\text{mismatch}} \approx -10\,\text{kcal}\,\text{mol}^{-1}$) for the intended TMSE reaction (Figure 8d, orange, microstate I to microstate V).

3.2 Toehold mediated strand displacement (TMSD)

In Section 3.2 we consider invasions of a variant of the $F_2(mm)$ fuel complex that lacks a toehold sized clamp – the first five nucleotides from the 5′ end of strand $F_2(mm)$-b are truncated. While TMSD already has a significant driving force due to the free energy improvement of a toehold's worth of additional base pairing ($\Delta\Delta G_{\text{no mismatch}} \approx -6\,\text{kcal}\,\text{mol}^{-1}$, Figure 9) the mismatch strategy offers a means to increase this driving force further ($\Delta\Delta G_{\text{mismatch}} \approx -16\,\text{kcal}\,\text{mol}^{-1}$, Figure 9).

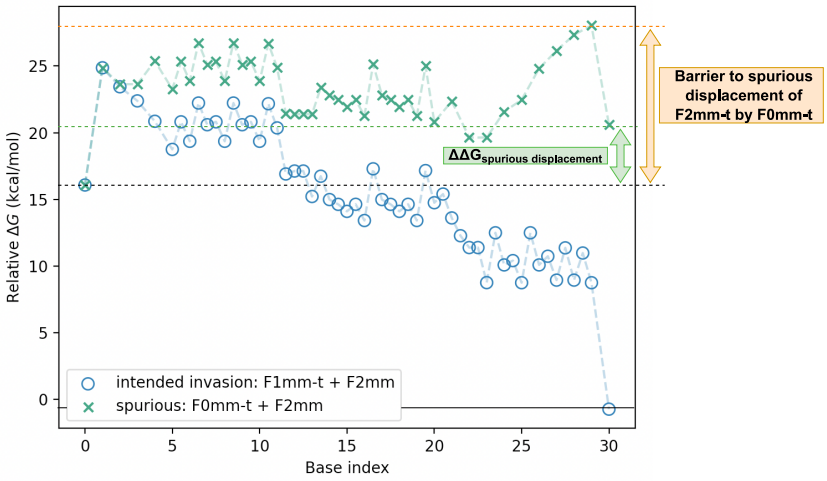


Figure 10 The sentinel mismatch strategy is compatible with leakless TMSD cascades.

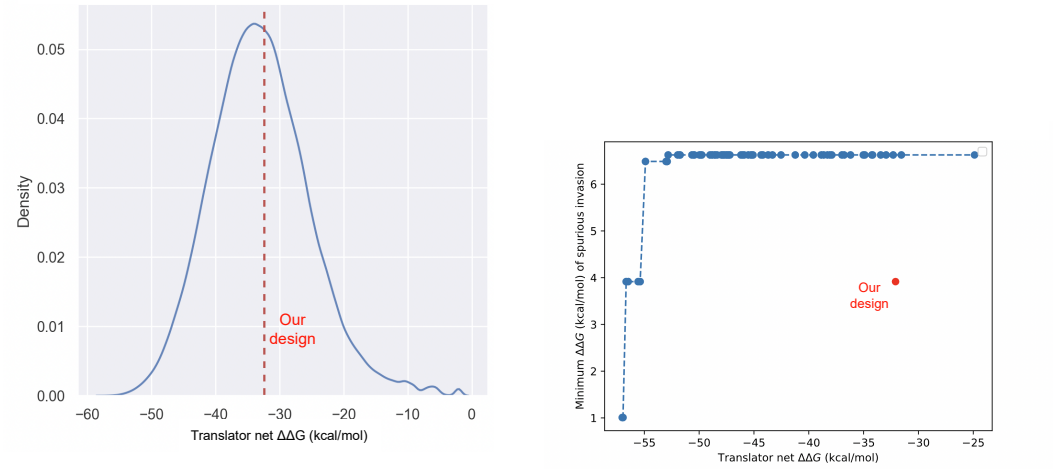
Unlike our TMSE case study, where spurious invasions are unlikely to result in leak reactions, the spurious invasion of $F_2(mm)$ by invader $F_0(mm)$-t can reach a microstate where the incumbent strand $F_2(mm)$-t is bound by only ten nucleotides (base index 25, Figure 10) and could therefore be susceptible to experimentally relevant rates of dissociation. However, the sentinel mismatch strategy can result in designs that kinetically and thermodynamically disfavor leak in TMSD.For example, in the variant considered here (i) leak is thermodynamically unfavorable ($\Delta\Delta G_{\text{spurious displacement}} = 4.52\,\text{kcal}\,\text{mol}^{-1}$), and (ii) the kinetic barrier to spurious displacement is greater than $11\,\text{kcal}\,\text{mol}^{-1}$.

3.3 Design space of sentinel mismatch strategy

Beginning from the reference sequences of Figure 5 (left) we used the naive algorithm from Section 2.2 to generate each of the 685,307 candidate sentinel designs that were valid for $\alpha = \beta = 5$. Thermodynamic properties were evaluated for each candidate including its minimum among all possible $\Delta\Delta G_{\text{spurious}}$ – the net free energy of the microstate representing the maximum progress of a spurious invasion (*e.g.*, microstate IV in Figure 8b) – and $\Delta\Delta G_{\text{translator}}$ – the net free energy of the overall cascade via designed reactions (*i.e.*, a proxy for its thermodynamic driving force). In our analysis that follows and in the annotations of Figure 11, *our design* refers to the sentinel mismatch sequences of Figure 3 (right).

The full design space visualized as a density plot of $\Delta\Delta G_{\text{translator}}$ for each candidate design is shown in Figure 11a; this includes the reference sequence design, which contain no sentinel position, as it is valid by our definition. Note that *our design*, which we investigated the thermodynamic properties of in this section and experimentally evaluated and discuss in the next section, has worse thermodynamic driving force than the average across all valid candidate designs.

Optimizing for only the driving force of the overall cascade does not necessarily result in designs that mitigate spurious invasion. In practice, one may wish to optimize for some trade-off between multiple properties. Figure 11b plots the Pareto frontier of maximizing the minimum $\Delta\Delta G_{\text{spurious}}$ and minimizing $\Delta\Delta G_{\text{translator}}$. Our design is not expected to be on this front because it contains fewer sentinel positions than are permitted by the $\alpha = \beta = 5$ constraint. Overall, the sentinel mismatch strategy allows further room for



(a) **(b)**

Figure 11 **(a)** Density distribution of all valid sentinel design candidates starting from the non mismatch reference sequences – see Figure 3 (left) – that obey $\alpha = \beta = 5$ and disfavor spurious invasion. **(b)** Pareto frontier for maximizing the minimum $\Delta\Delta G_{\text{spurious}}$ – the net free energy of any spurious invasion in a design – and minimizing $\Delta\Delta G_{\text{translator}}$ – the net free energy of the overall cascade via designed reactions.

translator optimization and can be broadly applied to generate a rich design space for any cascade with redundancy. A final design can be chosen based on multiple, user-defined criteria and constraints utilizing tools such as NUPACK [4] for evaluation.

4 Preliminary Experimental Verification

To demonstrate the efficacy of a systematic domain *and* sequence level design strategy for DSD cascades we compared a typical leakless translator as a control to a variant that incorporates sentinel positions. Our control is a cascade that requires three successive displacement reactions to translate an input to a sequence independent output and has a large entropic and enthalpic barrier to leak [20] due to its domain level redundancy and incorporation of toehold-sized clamps, making it reversible (see Figure 3 (left)). However, the redundancy within the control cascade creates a number of (unproductive) spurious invasion pathways. The variant tested incorporates the systematic mismatch strategy of sentinel positions described in Section 2 (see Figure 3 (right)). Sequences were designed using NUPACK [2, 4] to verify that all complexes were well-formed, all signal strands were unstructured, all desired toehold exchange reactions netted $\approx 0\,\text{kcal}\,\text{mol}^{-1}$ in structure free energy, and sequence motifs that might affect synthesis yield – such as homomeric repeats – were absent. To facilitate direct comparison between the translator systems with and without the sentinel strategy both systems were designed to use the same input strand as a trigger, produce identical sequences as outputs, and to use identical reporter complexes that consume this output sequence. All designed sequences were ordered from IDT with PAGE purification; annealed complexes were PAGE purified and concentrations were measured via UV absorbance and then normalized.

4.1 Leak reactions

Based on work by Wang *et al.* [21, 20], because of the levels of redundancy and presence of toehold-sized clamps, we expected negligible rates of leak in the non-mismatched cascade (our control). A recent experimental demonstration of a translator with the same domain-level design, but differing in sequence, had no significant rate of leak even at micromolar concentrations [22]. However, in order for the sentinel mismatch strategy to provide a valid remedy to the adverse kinetic effects introduced by domain level strategies that address robustness, it must itself not re-introduce significant leak. Thus we hoped to find similarly low levels of leak in both the non-mismatched and mismatched translator cascades.

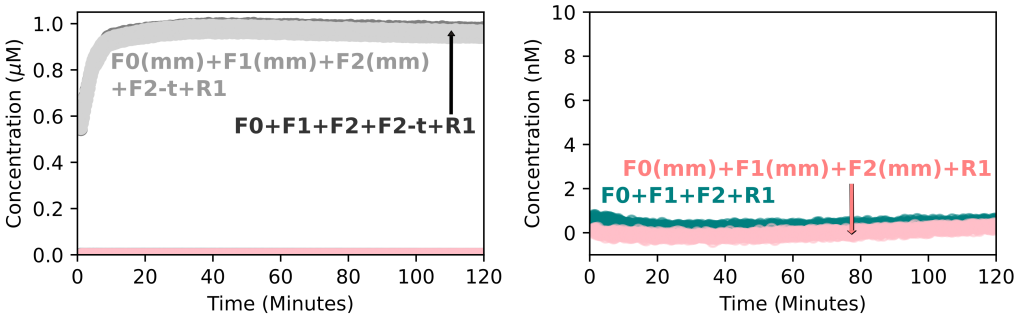


Figure 12 Testing the presence of output signal in the absence of valid input (leak) in leakless translators with and without the sentinel mismatch strategy. (Right) Time-course fluorescence results when incubating 1 µmol L^{-1} of each fuel complex in a given translator system with 1.5 µmol L^{-1} reporter and no input trigger strands are reported for the non-mismatched and mismatched translator systems in teal and pink respectively. (Left) Internal standards of the aforementioned systems spiked with final concentrations of 1 µmol L^{-1} of strand F_2-t (direct trigger to the reporter) are plotted for the same interval. All replicates are plotted for each of the experimental conditions tested. Both the mismatched and non-mismatched systems leaked a maximum of 0.054% of the total possible output in 2 hours, whereas the high controls triggered at the reporter with 1 µmol L^{-1} F_2-t reached over 50% completion in the 50 seconds between mixing of reactants and the first fluorescence reading.

To verify leak did not increase when using the mismatch strategy we prepared two sets of triplicate mixtures: one containing purified F_0, F_1, and F_2 at final nominal concentrations of 1 µmol L^{-1} and fluorescent output reporter R_1 at a final concentration of 1.5 µmol L^{-1}, and the other containing purified $F_0(mm)$, $F_1(mm)$, and $F_2(mm)$ at final concentrations of 1 µmol L^{-1} and R_1 at a final concentration of 1.5 µmol L^{-1}. Both of these sets of samples were compared to triplicate measurements of high controls of the same mixtures prepared with a final concentration of 1 µmol L^{-1} F_2-t that directly triggers the output reporter. Each reaction sample was quickly mixed and observed under time-course fluorescence measurements every 20 seconds for 2 hours. Raw fluorescence values were then normalized to concentration using internal low and high controls. With the exception of three data series which were discarded due to sample evaporation the results across all three replicates are shown in Figure 12.

From Figure 12 it appears that mismatch introduction did not significantly alter the rate of leak in our system. While the high control reactions reached above 50% completion in the 50s interval between the mixture of reaction components and the start of fluorescence measurements, the untriggered samples for both the non-mismatched and mismatched systems leaked at most 0.054% in 2 hours.

4.2 Designed reactions

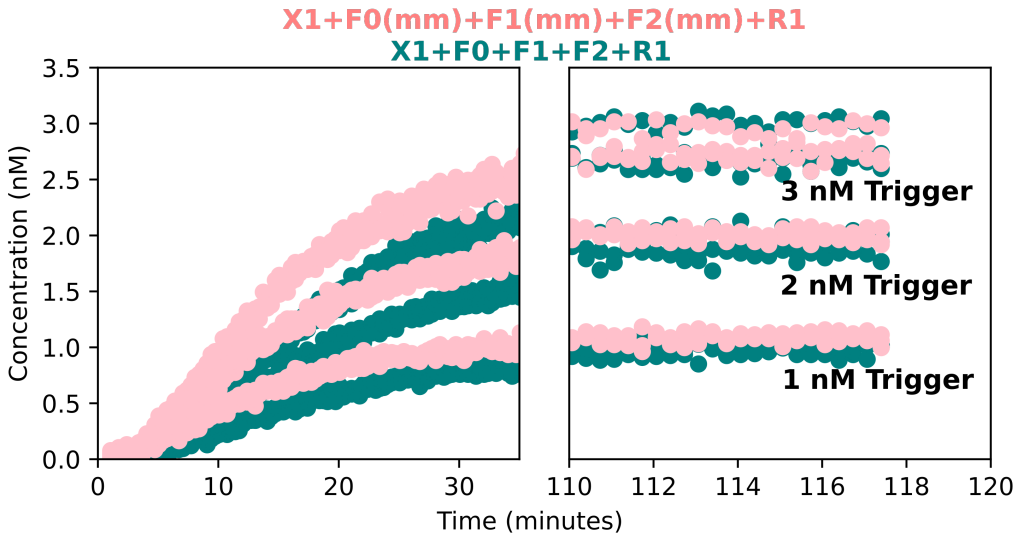


Figure 13 Comparing the triggering kinetics of leakless translator systems with and without the mismatch strategy. Each reaction was conducted with 10 nmol L^{-1} of each translator gate in the respective cascade and 15 nmol L^{-1} of output reporter R_1. The x-axis is broken to include only the beginning of the reaction and the equilibrated endpoints of the reaction. Raw fluorescence of each reaction was converted to concentration by internal normalization with a low control of untriggered cascade and high controls of cascade triggered with 10 nmol L^{-1} F_2-t at the reporter. The normalized output concentrations for each of the three replicates for each experimental condition were plotted and labeled with their reaction components as given in Figure 3. Clear differences in reaction half-times between the two systems at each concentration, even considering the variance across triplicate measurements, suggests that the introduction of mismatches increased the effective forward rate of the translator system compared to the non-mismatched variant.

In the previous section we presented evidence supporting that our strategy for introducing systematic sentinel mismatches into leakless translator cascades does not significantly increase leak relative to the non-mismatched case. This fact alone suggests some flexibility for sequence design in translator systems and supports a notion of compatibility between redundancy based leak-reduction and mismatch-based optimization strategies. However, we also theorized that the mismatch strategy would increase the observed reaction rate of a translator system by introducing kinetic barriers to unproductive reaction pathways that temporarily sequester reaction intermediates relevant to the desired reaction rate. To test this hypothesis, we measured the time-course fluorescence of triplicate samples of mismatched and non-mismatched translator cascades at 10 nmol L^{-1} gate concentration, 15 nmol L^{-1} reporter concentration, and input trigger concentrations of 1 nmol L^{-1}, 2 nmol L^{-1}, and 3 nmol L^{-1}. Additionally, we recorded triplicate measurements of high controls of each of non-mismatched and mismatched cascade with gates and reporters at the previous concentrations, and with 10 nmol L^{-1} F_2-t to directly trigger the reporter. Lastly, we measured untriggered low control cases of both the mismatched and non-mismatched translator cascades at the same concentrations previously used.

Based on the results presented in Figure 13, it appears that the sentinel mismatch strategy accomplished its goal of increasing the forward rate of the $X1 \rightarrow Y1$ translator reaction, without significantly increasing rates of leak. At each of the triggering concentrations, even

considering the variance across replicates, the reaction halftimes of the mismatched cascades were clearly lower than those in the non-mismatched cascades. This suggests that the effective rate constants of reactions when using the sentinel mismatch strategy is larger than in the non-mismatched case. Additional experiments are required to accurately quantify the effect of this systematic strategy.

5 Conclusion

We have introduced, analyzed, and given preliminary experimental validation for a systematic mismatch strategy compatible with *leakless* DNA strand displacement cascades. By promoting desired reactions and simultaneously discouraging spurious and unproductive invasions, even at high concentrations, the combination of domain and sequence level strategies unlocks a thermodynamic landscape conducive to fast, robust circuits. If this strategy proves effective broadly, as we expect, it may provide a new design standard for effective DNA strand displacement reactions and networks. The additional driving force introduced by the systematic mismatches in a toehold exchange reaction is greater than a comparable toehold-mediated strand displacement reaction without the strategy. Future experimental work will include characterization of these thermodynamic and additional kinetic differences. While we focused our strategy in the context of leakless translators it is applicable to any system with redundancy of domains between signals in a common cascade; this includes proposals of leakless architectures capable of implementing arbitrary chemical reaction networks [18]. The early success of this approach motivates the need for algorithms and design software capable of elucidating sequence level modifications that best optimize energy barriers of concern in the context of an entire strand displacement system. We presented a naive algorithm that can find an optimal design for a particular cascade, but it is unlikely to remain tractable when considering global constraints across a large system. A natural question is how to best optimize these sequence level modifications in order to improve overall kinetic rates of desired reactions without increasing undesirable reaction rates beyond a user-defined value in an overall system. Finally, what other strategies can be developed to improve strand displacement systems when combining systematic domain and sequence level design principles?

References

1 Vadim V Demidov, Michael V Yavnilovich, Boris P Belotserkovskii, Maxim D Frank-Kamenetskii, and Peter E Nielsen. Kinetics and mechanism of polyamide ("peptide") nucleic acid binding to duplex DNA. *Proceedings of the National Academy of Sciences*, 92(7):2637–2641, 1995.

2 Robert M Dirks, Justin S Bois, Joseph M Schaeffer, Erik Winfree, and Niles A Pierce. Thermodynamic analysis of interacting nucleic acid strands. *SIAM review*, 49(1):65–88, 2007.

3 MS Ellwood, M Collins, EF Fritsch, JI Williams, SE Diamond, and JG Brewen. Strand displacement applied to assays with nucleic acid probes. *Clinical chemistry*, 32(9):1631–1636, 1986.

4 Mark E Fornace, Nicholas J Porubsky, and Niles A Pierce. A unified dynamic programming framework for the analysis of interacting nucleic acid strands: Enhanced models, scalability, and speed. *ACS Synthetic Biology*, 9(10):2665–2678, 2020.

5 Natalie EC Haley, Thomas E Ouldridge, Ismael Mullor Ruiz, Alessandro Geraldini, Ard A Louis, Jonathan Bath, and Andrew J Turberfield. Design of hidden thermodynamic driving for non-equilibrium systems via mismatch elimination during DNA strand displacement. *Nature communications*, 11(1):1–11, 2020.

6 Patrick Irmisch, Thomas E Ouldridge, and Ralf Seidel. Modeling DNA-strand displacement reactions in the presence of base-pair mismatches. *Journal of the American Chemical Society*, 142(26):11451–11463, 2020.

7 Yu Sherry Jiang, Sanchita Bhadra, Bingling Li, and Andrew D Ellington. Mismatches improve the performance of strand-displacement nucleic acid circuits. *Angewandte Chemie*, 126(7):1876–1879, 2014.

8 Dmitriy A Khodakov, Anastasia S Khodakova, David M Huang, Adrian Linacre, and Amanda V Ellis. Protected DNA strand displacement for enhanced single nucleotide discrimination in double-stranded DNA. *Scientific reports*, 5(1):1–8, 2015.

9 Aaron Klug. Rosalind franklin and the discovery of the structure of DNA. *Nature*, 219(5156):808–810, 1968.

10 Hao Liu, Fan Hong, Francesca Smith, John Goertz, Thomas Ouldridge, Molly M. Stevens, Hao Yan, and Petr Šulc. Kinetics of RNA and RNA:DNA hybrid strand displacement. *ACS Synthetic Biology*, 10:3066–3073, November 2021.

11 Drew Lysne, Kailee Jones, Alma Stosius, Tim Hachigian, Jeunghoon Lee, and Elton Graugnard. Availability-driven design of hairpin fuels and small interfering strands for leakage reduction in autocatalytic networks. *The Journal of Physical Chemistry B*, 124(16):3326–3335, 2020.

12 Robert RF Machinek, Thomas E Ouldridge, Natalie EC Haley, Jonathan Bath, and Andrew J Turberfield. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nature communications*, 5(1):1–9, 2014.

13 Xiaoping Olson, Shohei Kotani, Jennifer E Padilla, Natalya Hallstrom, Sara Goltry, Jeunghoon Lee, Bernard Yurke, William L Hughes, and Elton Graugnard. Availability: A metric for nucleic acid strand displacement systems. *ACS synthetic biology*, 6(1):84–93, 2017.

14 Charles M Radding. Molecular mechanisms in genetic recombination. *Annual review of genetics*, 7(1):87–111, 1973.

15 Charles M Radding. Genetic recombination: strand transfer and mismatch repair. *Annual review of biochemistry*, 47(1):847–880, 1978.

16 Niranjan Srinivas, Thomas E. Ouldridge, Petr Šulc, Joseph M. Schaeffer, Bernard Yurke, Ard A. Louis, Jonathan P.K. Doye, and Erik Winfree. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Research*, 41:10641–10658, December 2013.

17 Weiyang Tang, Weiye Zhong, Yun Tan, Guan A Wang, Feng Li, and Yizhen Liu. DNA strand displacement reaction: a powerful tool for discriminating single nucleotide variants. *DNA Nanotechnology*, pages 377–406, 2020.

18 Chris Thachuk, Erik Winfree, and David Soloveichik. Leakless DNA strand displacement systems. In *International Workshop on DNA-Based Computers*, pages 133–153. Springer, 2015.

19 Andrew J Turberfield, JC Mitchell, Bernard Yurke, Allen P Mills Jr, MI Blakey, and Friedrich C Simmel. DNA fuel for free-running nanomachines. *Physical Review Letters*, 90(11):118102, 2003.

20 Boya Wang, Chris Thachuk, Andrew D Ellington, and David Soloveichik. The design space of strand displacement cascades with toehold-size clamps. In *International Conference on DNA-Based Computers*, pages 64–81. Springer, 2017.

21 Boya Wang, Chris Thachuk, Andrew D Ellington, Erik Winfree, and David Soloveichik. Effective design principles for leakless strand displacement systems. *Proceedings of the National Academy of Sciences*, 115(52):E12182–E12191, 2018.

22 Boya Wang, Chris Thachuk, and David Soloveichik. Speed and correctness guarantees for programmable enthalpy-neutral DNA reactions. *bioRxiv*, 2022. doi:10.1101/2022.04.13.488226.

23 James D Watson and Francis HC Crick. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–738, 1953.

24 Dongbao Yao, Tingjie Song, Xianbao Sun, Shiyan Xiao, Fujian Huang, and Haojun Liang. Integrating DNA-strand-displacement circuitry with self-assembly of spherical nucleic acids. *Journal of the American Chemical Society*, 137(44):14107–14113, 2015.

25 Bernard Yurke and Allen P Mills Jr. Using DNA to power nanostructures. *Genetic Programming and Evolvable Machines*, 4(2):111–122, 2003.

26 David Yu Zhang and Georg Seelig. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature chemistry*, 3(2):103–113, 2011.