Supporting Information

Availability: A Metric for Nucleic Acid Strand Displacement Systems

Xiaoping Olson†, Shohei Kotani†, Jennifer E. Padilla†, Natalya Hallstrom†, Sara Goltry†,
Jeunghoon Lee†,‡, Bernard Yurke†,§, William L. Hughes*,†, and Elton Graugnard*,†

† Micron School of Materials Science & Engineering,
‡ Department of Chemistry & Biochemistry,
§ Department of Electrical & Computer Engineering,
Boise State University, 1910 University Drive, Boise, ID, 83725, USA.

*willhughes@boisestate.edu & eltongraugnard@boisestate.edu
Contents

S1. Strand sequences and reporter schematic ................................................................. 3
  Table S1. Substrate and catalyst sequences ................................................................. 3
  Table S2. Reporter sequences ....................................................................................... 3
  Table S3. Fuel sequences .............................................................................................. 4

S2. Extinction coefficient calculation .............................................................................. 5

S3. Reaction kinetics models .......................................................................................... 5

S4 Transient binding of the substrate and fuel ................................................................. 8

S5 Minimum free energy structures of the fuel strands .................................................. 9

S6. Availabilities of fuel strand bases .............................................................................. 10

S7. Analysis in the context of the Intuitive Energy Landscape (IEL) model ..................... 12
  S7.1 The intuitive energy landscape model for leakage pathways ................................ 12

S8. Multiple location fuel modifications ......................................................................... 17

S9. Analysis of catalytic rates ......................................................................................... 21

S10. Rate constants and mutual availability .................................................................. 23
  Table S4. Rate constants and total mutual availability ................................................. 23

S11. Mutual availability and rate constants in a hairpin system ...................................... 24

Supporting Information – 2
S1. Strand sequences and reporter schematic

Table S1. Substrate and catalyst sequences and schematic

| Name               | Sequence (5’ to 3’)                                      |
|--------------------|----------------------------------------------------------|
| Signal             | CCACATACATCATATTCCCTCATTCAATACCCTACG                     |
| Output             | CTACTTTACACCTAGTCTCCAAACTAATTTACGG                      |
| Backbone           | TGGAGACGTAAGGATTTGAAATGAGGGCGTAAGTTAGTTGGAGACGTAAGG     |
| Catalyst           | CATTCAATACCTACGTCTCCATACTTTAGCC                         |
| 1 nt deletion catalyst | ATTCAATACCTACGTCTCCATACTTTAGCC                        |

Table S2. Reporter sequences and schematic

| Name      | Sequence (5’ to 3’)                                      |
|-----------|----------------------------------------------------------|
| TET       | /5TET/ CCACATACATCATATTCCCT                              |
| Quencher  | TTGAATGAGGAATATGATGTAGTTGGGG /3IABkFQ/                  |

Figure S1. Schematics of (a) the substrate complex, and (b) the network for fluorescence reporting the presence of displaced Signal strands.
| Fuel | Sequence (5 to 3) | MFE (kcal/mol) |
|------|------------------|----------------|
| **Unmodified** | | |
| Original | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| **5 end modifications** | | |
| A5 | ACTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| T5 | TCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| G5 | GCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| A5T5 | ATAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| A5G5 | AGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| T5T5 | TTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| G5T5 | GTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| G5 | GTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| G5T5 | GTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| **Nick location modifications** | | |
| A5d | CCTAGCTCCACTAATTACGGACCTCATTTAATACCCCTACG | -2.20 |
| T5d | CCTAGCTCCACTAATTACGTGCCTCATTTAATACCCCTACG | -2.51 |
| C5d | CCTAGCTCCACTAATTACGGACCCTCATTTAATACCCCTACG | -1.47 |
| A5dA5d | CCTAGCTCCACTAATTACGGAAACCTCATTTAATACCCCTACG | -2.20 |
| A5dT5d | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -2.20 |
| A5dT5d | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -2.20 |
| T5dT5d | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -2.99 |
| C5dC5d | CCTAGCTCCACTAATTACGGACCTCATTTAATACCCCTACG | -2.66 |
| C5dG5d | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -3.04 |
| C5dA5d | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -1.47 |
| C5dT5d | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -1.47 |
| C5d | CCTAGCTCCACTAATTACGGACCTCATTTAATACCCCTACG | -1.81 |
| T5d | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -1.81 |
| G5 | CCTAGCTCCACTAATTACGGATCTCATTTAATACCCCTACG | -1.89 |
| **3 end modifications** | | |
| A10dA4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAA | -1.81 |
| A10dT4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAT | -1.81 |
| A10dC4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAC | -1.81 |
| T10dA4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAT | -1.81 |
| T10dT4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAT | -1.81 |
| T10dC4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAT | -1.81 |
| G10dA4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGA | -1.81 |
| G10dT4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGT | -1.81 |
| G10dC4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGC | -1.81 |
| A4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAC | -1.81 |
| T4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAC | -1.81 |
| C4d | CCTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAC | -1.81 |
| **Multiple-location modifications** | | |
| G10dA25 | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -1.81 |
| G10dG25 | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -1.96 |
| G10dT25A25 | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTACG | -6.71 |
| G10dT25G25 | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -2.80 |
| G10dT25C25 | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -1.81 |
| G10dG25T25 | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -1.81 |
| T10dT25A4d | CTTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGT | -2.99 |
| G10dT25G4d | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -1.81 |
| G10dT25C4d | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -2.51 |
| G10dG25T4d | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -2.80 |
| G10dG25T4d | GGTAGCTCCACTAATTACGGCCCTACTTAATACCCCTAGG | -4.49 |
S2. Extinction coefficient calculation

The extinction coefficients at 260 nm for single-stranded DNA components were provided by integrated DNA technologies. The extinction coefficients at 260 nm for the reporter and the substrate were calculated by summing up the single and double-stranded regions, as following $e_{\text{total}} = e_{\text{single}} + e_{\text{double}}$, where $e_{\text{double}} = e_{\text{top strand}} + e_{\text{bottom strand}} - 3200 \text{ (units)} N_{\text{AT}} + 2000 \text{ (units)} N_{\text{GC}}$, where $N_{\text{AT}}$ and $N_{\text{GC}}$ are respectively the number of AT pairs and GC pairs in the double-stranded regions.

S3. Reaction kinetics models

We consider the following model for the catalytic system:

$$F + C + S \overset{k_{\text{cat}}}{\longrightarrow} W + C + OB + SB$$
(1)

$$F + S \overset{k_{\text{leak}}}{\longrightarrow} W + OB + SB$$
(2)

$$SB + R \overset{k_{\text{TET}}}{\longrightarrow} \text{Fluorescence}$$
(3)

Equation (1) is the catalyzed reaction for the catalytic system in Fig. 1. In this reaction $F$ is the fuel, $C$ is the catalyst, and $S$ is the substrate. In this reaction the fuel $F$, catalyst $C$, and substrate $S$ combine to form waste product $W$ and release catalyst $C$, output $OB$, and signal $SB$. The rate constant is denoted $k_{\text{cat}}$.

Equation (2) is the leakage reaction for the catalytic system in Fig. 1. In this reaction the fuel $F$ and the substrate $S$ combine to form the waste product $W$ and release output $OB$ and signal $SB$. This reaction proceeds with the rate constant $k_{\text{leak}}$.

Equation (3) is the reaction of signal $SB$ and reporter $R$ with a rate constant $k_{\text{TET}} = 8 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$, which is much faster than Equations (1) and (2). As a consequence,
there is no reporter delay and we treat the fluorescence as a direct measure of the SB concentration.

Equations (1-3) do not include the reverse reactions, that is, the reverse reactions are considered to be negligible.

Leakage and catalyzed reactions have the following rate equation:

\[
\frac{d[SB]}{dt} = k_{cat}[F][C][S] + k_{leak}[F][S]
\]  

Mass balance equations are:

\[
[F] = [F]_0 - [SB]
\]  

\[
[S] = [S]_0 - [SB]
\]  

\[
[C] = [C]_0
\]

Equations (4) through (7) yield:

\[
\frac{d[SB]}{dt} = k_{cat} ([F]_0 - [SB])[C]_0 ([S]_0 - [SB]) + k_{leak} ([F]_0 - [SB]) ([S]_0 - [SB])
\]

One obtains:

\[
[SB] = \frac{[F]_0 [S]_0 (1 - \exp\left[t (k_{cat}[C]_0 + k_{leak}) ([S]_0 - [F]_0)\right])}{[F]_0 - [S]_0 \exp[t (k_{cat}[C]_0 + k_{leak}) ([S]_0 - [F]_0)]}
\]

When \(t \to \infty\) and \([S]_0 > [F]_0\)

\[
[SB]_\infty = [F]_0
\]

\[
[SB]_{cat} = \frac{[SB]}{[SB]_\infty} = \frac{[S]_0 (1 - \exp[t (k_{cat}[C]_0 + k_{leak}) ([S]_0 - [F]_0)])}{[F]_0 - [S]_0 \exp[t (k_{cat}[C]_0 + k_{leak}) ([S]_0 - [F]_0)]}
\]

The fluorescence concentration is approximately equal to the signal strand concentration. Rate constants of catalytic reaction can be obtained by fitting the fluorescence versus time data using the equation (11), as illustrated in Fig. 6b.

For leakage reaction \([C]_0 = 0\), thus
\[ \text{[SB]} = \frac{[F]_0[S]_0(1-\exp[t^{\text{kleak}}([S]_0-[F]_0)])}{[F]_0-\exp[t^{\text{kleak}}([S]_0-[F]_0)]} \] (12)

When \( t \to \infty \) and \( [S]_0 < [F]_0 \)

\[ \text{[SB]}_{\infty} = [S]_0 \] (13)

\[ \text{[SB]}_{\text{leak}} = \frac{[SB]}{[SB]_{\infty}} = \frac{[F]_0(1-\exp[t^{\text{kleak}}([S]_0-[F]_0)])}{[F]_0-\exp[t^{\text{kleak}}([S]_0-[F]_0)]} \] (14)

This is the equation to which the fluorescence data is fit in order to extract the leakage rate constant from the fluorescence versus time data of the leakage experiments, as illustrated in Fig. 6c.
S4 Transient binding of the substrate and fuel

Figure S2. The transient binding between the substrate toehold and the fuel. The y domain of the fuel strand and $y'$ domain of the substrate toehold can hybridize as shown in reaction I. The x domain on the 5' end of the fuel will further displace the x domain of the signal strand via 3-way branch migration (reaction II). This interaction might interfere with the 3' end of the fuel binding with the substrate and thus affect the leakage rate constants.
S5 Minimum free energy structures of the fuel strands

Figure S3. Minimum free energy structures of fuels with single-location modifications as calculated by NUPACK. The Gibbs free energy of each structure is shown in kcal/mol.
S6. Availabilities of fuel strand bases

![Figure S4](image-url)

**Figure S4.** Leakage rate constants in descending order and corresponding base availabilities of fuel strands. Each base of the fuel strand is represented as a solid circle colored according to the NUPACK calculated availability. Mismatched bases are circled in gray.
Figure S5. Catalyzed reaction rate constants in descending order and corresponding fuel base availabilities. Each base of the fuel strand is represented as a solid circle colored according to NUPACK calculated availability. The four bases complementary to the toehold are highlighted with the gray box. Mismatched bases are circled in gray.
S7. Analysis in the context of the Intuitive Energy Landscape (IEL) model

Recently, Srinivas et al. published a comprehensive analysis of branch migration and strand displacement and proposed an intuitive energy landscape (IEL) model to describe the experimental rates of strand displacement reactions. In an effort to gain additional insight into leakage, we analyzed the leakage reaction pathways in the framework of the IEL by incorporating the effects of fuel sequence mismatch energy penalties. From the IEL model, we can predict the leakage pathways with the lowest thermodynamic energy barriers (Figs. S7a,b), yet the most significant fuel modifications were those that affected the high-energy pathways (Fig. S6a) where the backbone base availabilities were highest. In IEL model, five essential rates and energy parameters $k_{hi}$, $k_{uni}$, $\Delta G_s$, $\Delta G_p$, and $\Delta G_{bp}$ were used to describe strand displacement reaction kinetics, which includes rates of hybridization, fraying, branch migration and branch migration initiation. These parameters were derived or fitted for the case of minimal secondary structures for the single-stranded components and stable duplexes. However, to accurately predict reaction rates, including leakage, base availability from the secondary structures of single strands and unstable duplexes, especially the blunt ends and nick locations, may be an essential part of the IEL model. Future work will include incorporating availability and mutual availability into the IEL model to gain a greater understanding, and thus greater control, of leakage reaction mechanisms.

S7.1 The intuitive energy landscape model for leakage pathways

Figure S6 and S7 shows four possible leakage pathways analyzed in the context of the intuitive energy landscape (IEL) of Srinivas et al. In Fig. S6a, state A is the initial state in which the fuel ($G_1T_2$) and substrate are separated, and the Gibbs free energy is taken as zero. For simplicity, all energy parameters were taken from Ref. 4. From state A to B, one end of the substrate frays at the cost of one base pair stacking $|\Delta G_{bp}|$. Three base pairs of the substrate have to fray to form the first base pair with the $G_1T_2$ due to two nucleotide modifications (indicated in the red dots). This results in a higher energy barrier for the pathway illustrating leakage by the fuel 5’ end compared to the other three. There is an intermediate step in which the fuel and substrate come close together in correct registry at
the cost of $|\Delta G_{\text{init}}|$. State C is the formation of the first base pair between the fuel and substrate with the energy gain of $|\Delta G_{\text{bp}}|$, and in this process two overhangs protrude from both sides of the nick with the energy penalty of $|\Delta G_{\text{p}}|$. In branch migration process, three base pairs of the fuel have to fray due to the hairpin structure and the energy cost for each base is $|\Delta G_{\text{bp}}|$. State D is when the fuel hairpin opened and form new base pairs with substrate. The sawtooth pattern represents branch migration steps and the top of the sawtooth is the intermediate transition state with energy penalty of $|\Delta G_{\text{s}}|$. In state E, the fuel has completely displaced the output strand, decreasing the system energy by $|\Delta G_{\text{init}}|$ and $|\Delta G_{\text{p1}}|$ (energy difference from no overhang to one overhang in the nick). In the final state F, both the output and signal strands are completely displaced by the fuel, and the gain of the system energy is $|\Delta G_{\text{init}}|$ and $|\Delta G_{\text{p2}}|$ (energy difference from no overhang to two overhangs in the nick). In Fig. S6a, the final state F is higher than the point Orig. (no mismatches in the fuel strand).
Figure S6. Intuitive energy landscape (IEL) for four proposed leakage pathways for G1T2 fuel modification. Red states indicate mismatch penalties. (a) Leakage initiated at 3’ end of the fuel (base 1), (b) Leakage initiated at nick left of the fuel (base 24). Yellow circles denote fraying locations. States A-G are described in the text. For illustration, the values of $\Delta G_{p1}$, $\Delta G_{p2}$, $\Delta G_{s}$, and $\Delta G_{\text{init}}$ were taken from Srinivas et al. $^d$ $|\Delta G_{p1}|$ is the energy difference from no overhang to one overhang in the nick, $|\Delta G_{p2}|$ is the energy difference from no overhang to two overhangs in the nick, and $|\Delta G_{p3}|$ is the energy difference from one overhang to two overhangs in the nick. For comparison, point “Orig.” shows the final state energy when the original fuel is used.
Figure S7. Intuitive energy landscape (IEL) for four proposed leakage pathways for G1T2 fuel modification. Red states indicate mismatch penalties. (a) Leakage initiated at nick right of the fuel (base 25), (b) Leakage initiated at 3’ end of the fuel (base 44). States A-G are described in the text. Yellow circles denote fraying locations. For illustration, the values of ΔG_{bp}, ΔG_p, ΔG_{in} and ΔG_{init} were taken from Srinivas et al.4 |ΔG_{p1}| is the energy difference from no overhang to one overhang in the nick, |ΔG_{p2}| is the energy difference from no overhang to two overhangs in the nick, and |ΔG_{p3}| is the energy difference from one overhang to two overhangs in the nick. For comparison, point “Orig.” shows the final state energy when the original fuel is used.

The leakage rate constant can be expressed by $k \approx \frac{2e^{-\frac{|\Delta G_{bp}|}{RT}}k_{bi}}{\rho + 2\gamma e^{(\Delta G_{in} + \Delta G_{p})/RT}}$, where $b$ is the number of bases in the branch migration region, and $\rho$ is defined as $\rho = \frac{k_{uni}}{k_{bi}}$. $\gamma$ is defined as $\gamma = e^{-(\Delta G_{bp} - \Delta G_{assoc})/RT}u_0$, and $\Delta G_{init} = \Delta G_{volume} + \Delta G_{assoc}$, where $\Delta G_{volume} = RT\ln\left(\frac{u_0}{u}\right)$, and $\Delta G_{assoc}$ is the free energy cost of association at a standard concentration of
\(u_0 = 1 M\) due to reduction in the entropy caused by lost translational and orientation degrees of freedom.\(^4\) Unimolecular and hybridization rate constants are denoted by \(k_{uni}\) and \(k_{bi}\) respectively, where \(R\) is the universal gas constant and \(T\) is the temperature.

From the above leakage rate constant equation, the availabilities of the fuel and substrate bases can be incorporated into the IEL model. First, the fuel and substrate collide in the nucleation step, which can be described by the bimolecular rate constant \(k_{eff}\). When the fuel has secondary structures and substrate frays at blunt ends and nick sites, base availabilities of fuel and substrate backbone at nucleation sites must be included in the \(k_{eff}\) to account for the probability of successful nucleation. Then branch migration proceeds after nucleation and this process is a unimolecular reaction, which can be described by \(k_{uni}\). In the branch migration process, the sawtooth amplitude (\(\Delta G_s\)) may need to be adjusted to account for secondary structures of the fuel strand. Therefore base availabilities of the fuel could be quantitatively incorporated into local free energy maxima along the sawtooth or through a sequence-dependent base-pairing energy and/or \(k_{uni}\). By considering base availability of the fuel and substrate backbone, the leakage rate constant can be modified through IEL model to precisely predict reaction rates. The biophysics of nucleation and branch migration steps could be further understood by utility of availability concept in the future.
S8. Multiple location fuel modifications

Multiple location fuel modifications have a stronger leakage suppression than the single location modifications, as seen in Fig. S8. The MFE structures for the fuel strands with sequence modifications at multiple locations are shown in Fig. S9. For the multiple location modifications, the leakage rate dropped by 100-fold to an almost undetectable level by introducing mismatches at all the four vulnerable locations. As more mismatches are introduced, leakage reduction can be attributed to an increase in secondary structure of the fuel strand, consistent with the MFE structures seen in Fig. S9, and a decrease in driving force. As the number of mismatches introduced to the fuel strand increases, the availability of the fuel strand decreases for this system. The changes in availability not only affect the nucleation sites but also appear in regions that can disrupt branch migration. For example, the point circled in yellow in Fig. S8b with a M of ~0.083 and ln($k_{\text{leak}}$) of ~0, is discussed further below (Fig. S10).

Multiple site fuel modifications also showed lower catalytic reaction rates, as expected from their decreased complementarity to the substrate backbone. As an attempt to recover the rate of the catalytic reaction, we increased the toehold length from 4 nt to 5 nt on domain $3^*$ of the intermediate 3 (I3) by deleting one nucleotide on the 5' end of the catalyst. This had the desired effect only on the toehold mismatch modifications, while demonstrating the opposite effect on the other modifications (Fig. S8). There exists a trade-off between slowing the dissociation of the signal strand (SB) and accelerating fuel binding to intermediate 3 (I3) at toehold domain $3^*$. This trade-off is advantageous in the case that the fuel strand has a mismatch in the toehold region (base 25), while it is disadvantageous in the case that the fuel is unmodified in the toehold location. Some multiple location modifications achieve a larger $k_{\text{cat}}/k_{\text{leak}}$ ratio compared with the original design, however none of these modification exceeded the best of the single location modifications (the ratio of 4, Fig.S8d).
**Figure S8.** (a) Leakage rate constants of multiple location fuel modifications. (b) Natural log plot of the leakage rate constant versus the mutual availability between the fuel strand and the backbone on the substrate. Representative error bars of selected samples are shown, indicating that the scatter of the data is greater than the experimental error. The red line is the fit for all modifications combined. (c) Catalytic rate constants of multiple location fuel modifications. Gray bars represent substrate and fuel reacted with original catalyst and red bars represent substrate and fuel reacted with catalyst with 1 nt deletion at 5’ end. (d) The ratio $k_{\text{cat}}/k_{\text{leak}}$ for multiple location fuel modifications. Gray bars represent substrate and fuel reacted with original catalyst and red bars represent substrate and fuel reacted with catalyst with 1 nt deletion at 5’ end.
Figure S9. Minimum free energy structures of fuels with multiple-location modifications as calculated by NUPACK. The Gibbs free energy for each structure is shown in kcal/mol.
Figure S10. (a) Minimum free energy (MFE) structures of original fuel and G1G1T24A25 modification as calculated by NUPACK and the Gibbs free energy of each structure. (b) Base availabilities for the original fuel and fuel modification G1G1T24A25. (c) An alternative reaction pathway to Fig. 1b between the modified fuel (G1G2T24A25) and substrate.

Fuel modification G1G2T24A25 formed a stronger hairpin structure compared with the original fuel (Fig S10a) yet yielded a high leakage rate given its low total mutual availability with the substrate backbone (circled in yellow in Fig. S8b). The fuel base availabilities are shown in Fig S10b. However the base availabilities are higher in domain 4a, x and part of domain 3 for the modified fuel, which are corresponding to the double helix region between the signal and backbone of the substrate. An alternative reaction pathway is proposed in Fig. S10c. Fuel (G1G2T24A25) can initiate reaction with the substrate through the 3’ end of the fuel, which displaces x domain of the signal due to substrate fraying and then completely displaces signal strand through branch migration. In addition, the toehold domain y of the substrate is less likely to stick with y domain of the fuel since it is sequestered in an stable hairpin structure. The x domain of the fuel is also unlikely to further displace x domain of the signal due to a 4-way branch migration. Thus, the right-side of the substrate are more vulnerable for fuel invasion and high leakage rate. This analysis offers a plausible explanation to the anomalously high leakage rate for a strand with an overall low availability. This also highlights the utility of the concept of base availability and minimum free energy (MFE) structures for analyzing leakage reaction pathways and yielding insight on the leakage reaction rate.
S9. Analysis of catalytic rates

The effect of mismatches on the catalyzed reaction can be explained via the reaction mechanism. The mismatch positions play a very important role in the catalyzed reaction. Trends can be observed by grouping the mismatch positions of the fuel strand as bases 1, 2 and 24, base 25, and bases 43 and 44. This catalytic cycle can be simplified to four reactions modeled by Zhang et al. as shown below.\(^3\)

\[
S + C \xrightleftharpoons[k_{-1}]{k_1} I3 + SB \tag{15}
\]

\[
I3 + F \xrightarrow[k_2]{k_2} I5 + OB \tag{16}
\]

\[
I5 \xrightleftharpoons[k_{-3}]{k_3} C + W \tag{17}
\]

\[
SB + R \xrightleftharpoons[k_{Fluorescence}]{k_{Fluorescence}} Fluorescence \tag{18}
\]

The reaction between the substrate and the catalyst and the signal reacting with the reporter are not affected by introducing fuel mismatch modifications. Equation 16 is a toehold mediated strand displacement and its rate constant \(k_2\) is mainly determined by the toehold domain 3 of the fuel strand (Fig. 1). Thus, this reaction is strongly affected by toehold modifications at base 25. Equation 17 is the release of catalyst from the last intermediate (I5) to produce the waste product. Mismatch modifications at bases 43 and 44 slow down the catalyst release.

Base 25 is in domain 3, which is used by the fuel (F) to bind the intermediate 3 (I3). Single mismatches at base 25 of the fuel strand slow down the overall catalytic reaction from one order of magnitude to over two orders of magnitude (Fig. 6). Mismatches at the toehold position are the most detrimental to the catalytic rate.
Bases 1, 2 and 24 modifications had the least impact on the catalyzed rate. To explain this, we look at where they are encountered during strand displacement. According to Machinek et al., the last few base-pairs at the end of the branch migration spontaneously melt during strand displacement. Bases 1 and 2 are involved at the end of the strand displacement in Eqn. 16. The last few bases of the strand OB spontaneously fall off from intermediate 4 before fuel mismatches at base 1 and 2 are involved in the reaction. However, base 24 is right next to toehold domain 3 where it is encountered at the beginning of branch migration according to the pathway depicted in Fig. 1a. An alternative reaction pathway can proceed as following: the fuel strand releases the catalyst first instead of releasing strand OB due to higher activation energy barrier caused by mismatches at base 24. Mismatches at base 24 is encountered after the fuel establishes a long toehold with the strand backbone. This agrees with Machinek’s result that once a sufficiently long toehold has been established, mismatches in the branch migration region do not significantly impede strand displacement.

Equation 17 is a first order reaction, in which 6 base pairs in the catalyst strand spontaneously dissociate from intermediate (I5). For the 3’ site fuel modifications, the catalyst has to spontaneously detach seven bases for modifications at base 44 and eight bases for modifications at bases 44 and 43. Catalytic reaction rates of fuel modifications at base 44 are slower than the unmodified fuel strand, which can be attributed to a slower release of the catalyst from intermediate (I5). The effect of delaying catalyst release is even stronger in bases 43 and 44 fuel modifications.
## S10. Rate constants and mutual availability

A complete summary of leakage and catalytic rate constants, performance ratios, and total mutual availabilities are provided in Table S4 for both single and multiple location fuel modifications.

### Table S4. Rate constants and total mutual availability

| Fuel | $k_{cat} (M^2 s^{-1})$ | $k_{cat} (M^2 s^{-1})$ | $k_{cat}/k_{stat} (M^2)$ | $\ln(k_{cat})$ | $M$ |
|------|-------------------------|-------------------------|--------------------------|-----------------|------|
| **Unmodified** | | | | | |
| Original | 8.120 | 2.856E+13 | 3.512E+12 | 2.094 | 0.63336 |
| 5’ end modifications | | | | | |
| A1 | 8.320 | 2.433E+13 | 2.924E+12 | 2.119 | 0.63252 |
| T1 | 7.794 | 2.850E+12 | 3.656E+12 | 2.053 | 0.63431 |
| G1 | 5.272 | 1.800E+13 | 3.414E+12 | 1.662 | 0.61041 |
| A1A2 | 5.866 | 1.700E+13 | 2.897E+12 | 1.769 | 0.63366 |
| A1T2 | 4.908 | 1.700E+13 | 3.464E+12 | 1.591 | 0.63329 |
| A1G2 | 2.415 | 1.398E+13 | 5.790E+12 | 0.882 | 0.59486 |
| T1A2 | 5.509 | 1.683E+13 | 3.055E+12 | 1.706 | 0.63375 |
| T1T2 | 4.823 | 1.397E+13 | 2.896E+12 | 1.573 | 0.63426 |
| T1G2 | 3.452 | 1.392E+13 | 4.032E+12 | 1.239 | 0.61562 |
| G1A2 | 5.372 | 1.665E+13 | 3.099E+12 | 1.681 | 0.62457 |
| G1T2 | 2.010 | 2.456E+13 | 1.216E+13 | 0.698 | 0.57031 |
| G1G2 | 1.822 | 8.769E+12 | 4.813E+12 | 0.600 | 0.56708 |
| **Nick location modifications** | | | | | |
| A24 | 7.774 | 2.283E+13 | 2.937E+12 | 2.05 | 0.64705 |
| T24 | 7.240 | 2.167E+13 | 2.992E+12 | 1.980 | 0.47923 |
| C24 | 13.590 | 3.16E+13 | 2.320E+12 | 2.609 | 0.66726 |
| A24A25 | 5.721 | 9.355E+11 | 1.635E+11 | 1.744 | 0.47431 |
| A24T25 | 5.779 | 4.381E+11 | 7.580E+10 | 1.754 | 0.64776 |
| A24G25 | 4.029 | 3.650E+11 | 9.058E+10 | 1.394 | 0.57029 |
| T24A25 | 3.568 | 7.983E+11 | 2.237E+11 | 1.727 | 0.39064 |
| T24T25 | 3.830 | 5.755E+11 | 1.502E+11 | 1.343 | 0.44876 |
| T24G25 | 2.944 | 3.109E+11 | 1.055E+11 | 1.080 | 0.4381 |
| C24A25 | 6.768 | 3.534E+12 | 5.222E+11 | 1.912 | 0.65833 |
| C24T25 | 6.304 | 1.215E+12 | 1.926E+11 | 1.841 | 0.66309 |
| C24G25 | 5.390 | 1.518E+12 | 2.818E+11 | 1.685 | 0.5717 |
| A25 | 3.730 | 4.829E+12 | 1.294E+12 | 1.316 | 0.5302 |
| T25 | 4.128 | 6.103E+11 | 1.478E+11 | 1.418 | 0.50664 |
| G25 | 3.243 | 3.694E+11 | 1.138E+11 | 1.176 | 0.36412 |
| **3’ end modification** | | | | | |
| A44A44 | 6.146 | 8.039E+12 | 1.307E+12 | 1.816 | 0.6314 |
| A44T44 | 5.737 | 7.995E+12 | 1.393E+12 | 1.747 | 0.63224 |
| A44G44 | 5.812 | 8.506E+12 | 1.463E+12 | 1.760 | 0.63252 |
| T44A44 | 6.633 | 1.020E+13 | 1.538E+12 | 1.892 | 0.63188 |
| T44T44 | 5.723 | 8.556E+12 | 1.494E+12 | 1.745 | 0.62999 |
| T44C44 | 6.275 | 1.258E+13 | 2.004E+12 | 1.836 | 0.63264 |
| G44A44 | 5.170 | 1.900E+13 | 3.673E+12 | 1.643 | 0.63356 |
| G44T44 | 5.156 | 1.180E+13 | 2.287E+12 | 1.640 | 0.6313 |
| G44C44 | 5.798 | 8.707E+12 | 1.501E+12 | 1.757 | 0.6201 |
| A44 | 8.507 | 1.700E+13 | 1.998E+12 | 2.141 | 0.63267 |
| T44 | 7.262 | 1.439E+13 | 1.981E+12 | 1.983 | 0.63299 |
| C44 | 7.938 | 1.427E+13 | 1.797E+12 | 2.072 | 0.63157 |
| **Multiple-location modifications** | | | | | |
| G1T2A25 | 1.404 | 1.97E+12 | 1.407E+12 | 0.336 | 0.47453 |
| G1G1G43 | 0.704 | 6.12E+12 | 8.698E+12 | -0.352 | 0.57542 |
| G1T2G1G43 | 1.019 | 4.28E+10 | 4.200E+10 | 0.019 | 0.08283 |
| G1G2G4G43 | 0.151 | 4.55E+10 | 3.021E+11 | -1.893 | 0.36378 |
| G1T2G4T44 | 1.049 | 8.63E+12 | 8.298E+12 | 0.039 | 0.57801 |
| G1G2G4T1T44 | 0.750 | 5.09E+12 | 6.783E+12 | -0.287 | 0.5646 |
| T2G1G1T1G44 | 0.894 | 1.55E+10 | 1.735E+10 | -0.112 | 0.4084 |
| G1T2A2G4T44 | 0.259 | 1.85E+11 | 7.143E+11 | -1.351 | 0.47614 |
| G1T2G4T4T44 | 0.581 | 8.14E+10 | 1.401E+11 | -0.543 | 0.40441 |
| G1G2G4T1T44 | 0.183 | 4.54E+10 | 2.486E+11 | -1.700 | 0.31731 |
| G1G2T2G4G4T44 | 0.082 | 1.12E+10 | 1.359E+11 | -2.496 | 0.12528 |
S11. Mutual availability and rate constants in a hairpin system

Figure S11. (a) Leakage rate constants for each hairpin 2 modification plotted versus total mutual availability between hairpin 1 and hairpin 2. The leakage rate for the original hairpins is shown in black while the 1 nt, 2 nt, and 3 nt hairpin 2 modifications are shown in blue, red and dark yellow, respectively. (b) Natural log plot of the leakage rate constant versus the total mutual availability. The solid gray line is the fit for all modifications combined with an adjusted R-squared 0.82.

The total mutual availability $M$ was calculated in NUPACK for Jiang et al.’s hairpin design CircA. In Figure 3 and 4 of (6), the final intensity for 50 nM H1 is about 4300 relative fluorescence units (RFU) for the original hairpin system, and the original leakage 6 RFU/min is corresponding to 0.0697 nM/min. Thus the original leakage rate constant can be calculated as $0.0697 \text{ nM min}^{-1}/(50 \text{ nM} \times 50 \text{ nM}) = 465 \text{ M}^{-1}\text{s}^{-1}$. The data of total mutual availability and natural log of leakage rate constants was fit with a linear line. The fit for the dataset has a slope of 4.29 with an adjusted $R^2$ of 0.82. In NUPACK, the concentration for each hairpin is set to 50 nM, temperature is 25 °C, dangle is set to all and salt concentrations are 0.145 M Na$^+$ and 0 M Mg$^{2+}$.

References:

1. Qian, L., and Winfree, E. (2011) Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades, *Science* 332, 1196-1201.
2. Puglisi, J. D., and Tinoco Jr, I. (1989) [22] absorbance melting curves of RNA, In *Methods in Enzymology* (James E. Dahlberg, J. N. A., Ed.), pp 304-325, Academic Press.
3. Zhang, D. Y., Turberfield, A. J., Yurke, B., and Winfree, E. (2007) Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA, *Science* 318, 1121-1125.
4. Srinivas, N., Ouldridge, T. E., Sulc, P., Schaeffer, J. M., Yurke, B., Louis, A. A., Doye, J. P. K., and Winfree, E. (2013) On the biophysics and kinetics of toehold-mediated DNA strand displacement, *Nucleic Acids Research* 41, 10641-10658.

5. Machinek, R. R., Ouldridge, T. E., Haley, N. E., Bath, J., and Turberfield, A. J. (2014) Programmable energy landscapes for kinetic control of DNA strand displacement, *Nat Commun* 5.

6. Jiang, Y. S., Bhadra, S., Li, B., and Ellington, A. D. (2014) Mismatches Improve the Performance of Strand-Displacement Nucleic Acid Circuits, *Angewandte Chemie* 126, 1876-1879.