

Modeling DNA-Strand Displacement Reactions in the Presence of Base-Pair Mismatches

Patrick Irmisch, Thomas E. Ouldrige, and Ralf Seidel*

Cite This: *J. Am. Chem. Soc.* 2020, 142, 11451–11463

Read Online

ACCESS |



Metrics & More

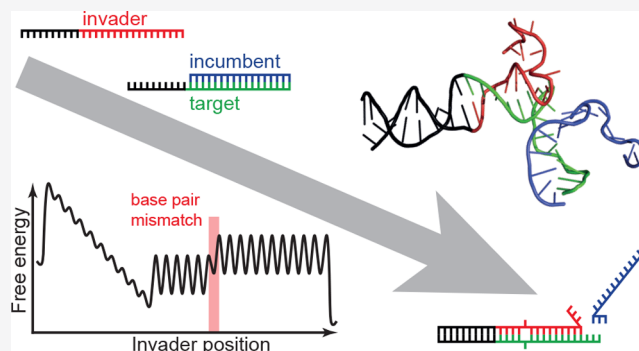


Article Recommendations



Supporting Information

ABSTRACT: Toehold-mediated strand displacement is the most abundantly used method to achieve dynamic switching in DNA-based nanotechnology. An “invader” strand binds to the “toehold” overhang of a target strand and replaces a target-bound “incumbent” strand. Here, the complementarity of the invader to the single-stranded toehold provides the free energy bias of the reaction. Despite the widespread use of strand displacement reactions for realizing dynamic DNA nanostructures, variants on the basic motif have not been completely characterized. Here we introduce a simple thermodynamic model, which is capable of quantitatively describing the kinetics of strand displacement reactions in the presence of mismatches, using a minimal set of parameters. Furthermore, our model highlights that base pair fraying and internal loop formation are important mechanisms when involving mismatches in the displacement process. Our model should provide a helpful tool for the rational design of strand-displacement reaction networks.



INTRODUCTION

Detailed insight into biological building blocks and their assembly principles has inspired the field of nanotechnology; particular interest has been paid to DNA, in which specific molecular interactions can be easily encoded in a nucleotide sequence.^{1–3} The major mechanism to achieve dynamic interactions in DNA-based systems is toehold-mediated strand displacement.⁴ A bimolecular association between a target and an incumbent strand is reversed by addition of an invader oligonucleotide, which displaces the incumbent (Figure 1). This reaction is facilitated by unpaired toehold nucleotides on the target strand to which the invader binds when initiating the displacement process. The additional toehold interaction makes the binding of the invader free-energetically favorable and provides the driving force for the reaction. DNA strand displacement reactions have been employed to realize dynamic devices such as molecular motors^{5,6} and guidable robots,^{7–11} chemical reaction networks^{12–15} that allow for complex computations,^{16–19} and biosensors^{20–22} operating under physiological conditions.

To fine-tune the performance of dynamic devices and chemical reaction networks, it is desirable to adjust the kinetics of a strand-displacement reaction without affecting the thermodynamic equilibrium of the system. This task can be fulfilled by incorporating mismatches that lead either to increased²³ (in case of incumbent–target mismatches) or to decreased²⁴ (in case of invader–target mismatches) reaction rates. By simply changing the mismatch position, the completion time of the reaction can be altered over several

orders of magnitude without changing the energetics of the system.^{23,24} Mismatches are therefore an attractive “non-invasive” possibility to program the kinetic parameters of strand-exchange reactions.²⁵ When applied in complex molecular circuitry, designing the whole system response requires knowledge of the individual reactions. Modeling of mismatch-based strand displacement reactions²⁴ has hitherto required, however, computationally demanding coarse-grained molecular dynamics simulations.^{26,27}

The aim of this work is to establish a simple and easily applicable model framework, which allows quantitative prediction of the kinetics of strand displacement reactions in the presence of mismatches, including position-dependent effects. We base our model on a previously introduced Markov chain²⁸ but expand it with the following essential contributions: (i) spontaneous incumbent dissociation prior to full displacement,²⁴ (ii) the incorporation of mismatches into the energy landscape of the Markov chain, and (iii) the fraying of DNA segments near mismatches. Despite the increased complexity of the model, the strand displacement process is still mapped onto a simple one-dimensional free-energy

Received: March 20, 2020

Published: June 4, 2020



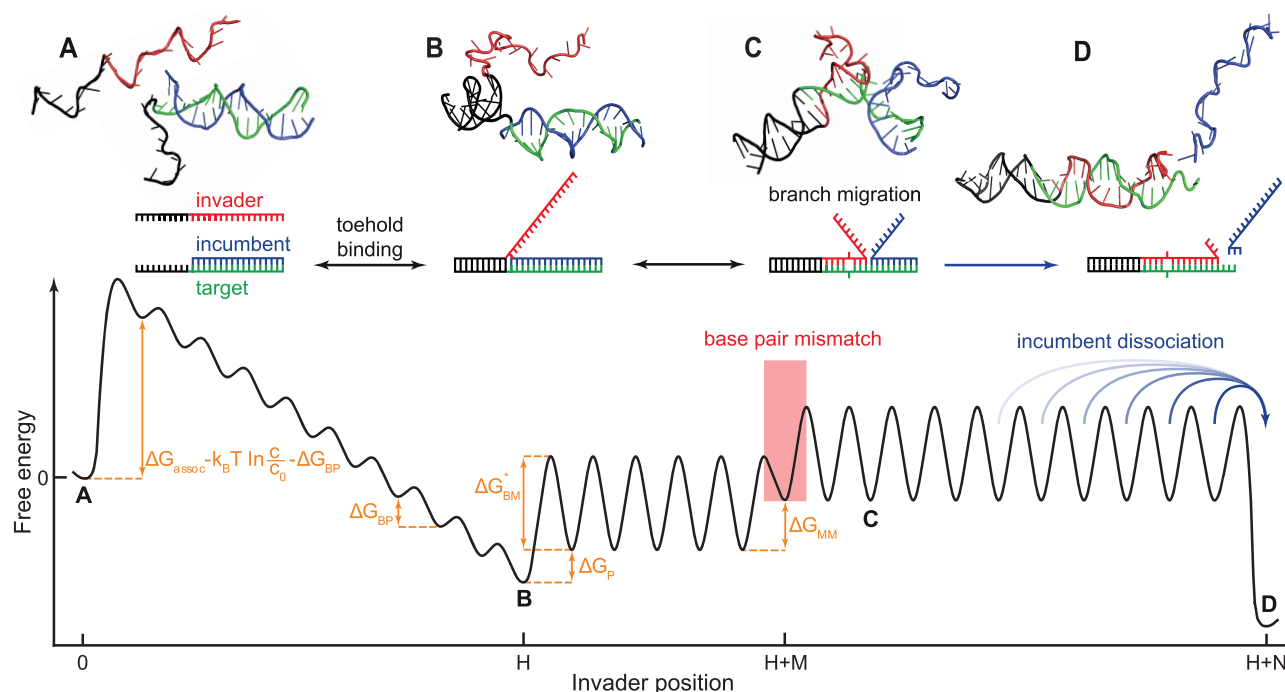


Figure 1. Scheme and simplified free-energy landscape of the toehold-mediated strand displacement reaction. (A) Start configuration. An N nucleotide-long incumbent strand (shown in blue) is hybridized to an $H + N$ nt long target DNA strand (shown in black/green) leaving a H nt ssDNA overhang called toehold. (B) Toehold binding. An invader strand that is complementary to the target binds first to the toehold. (C) Branch migration. The invader (red part) replaces the incumbent strand sequentially but reversibly in single base pair steps. It uses the fraying of the incumbent bases at the fork to hybridize with transiently unpaired target strand bases (green part of target). Mismatches between invader and target introduce an energetic penalty, which are assumed to increase mainly the backward rate of the branch migration step at the mismatch position. (D) Incumbent dissociation. Incumbent dissociation that completes the reaction, occurs when only few incumbent–target base pairs are left. The plot at the bottom shows a simplified energy landscape of this process as a function of the furthest incumbent–target base pair. The energy parameters that describe the landscape are shown in orange.

landscape. Using a minimal set of adaptable parameters, our model successfully describes the complex dependence of the displacement kinetics on the mismatch position and the toehold length for invader–incumbent mismatches. Once parametrized on this data set, the model correctly predicts the displacement kinetics in cases of incumbent–target mismatches as well as in the presence of two invader–target mismatches. For the latter, our modeling revealed that unpaired DNA loop formation between the mismatches can alleviate the total mismatch penalty, which allows a facilitated crossing of two mismatches in close proximity. Overall, our model provides a first basis to understand and predict the influence of structural alterations in strand-displacement reactions. It also serves as a model case for mechanism-based off-target predictors for CRISPR-Cas enzymes,²⁹ since these enzymes discriminate between on- and off-targets by an analogous strand-displacement reaction.^{30–32}

RESULTS

One-Dimensional Markov-Chain Model for Strand-Displacement with Mismatches. The basic strand displacement reaction starts with the binding of the invader strand to the toehold of the incumbent–target duplex (Figure 1A,B). Subsequently, the incumbent is successively displaced from the duplex by the invader, in a process called branch migration (Figure 1C). Once a sufficiently low number of incumbent–target base pairs are left, the incumbent can spontaneously dissociate (Figure 1D). As shown previously,²⁴ invader–target mismatches reduce the rate of the displacement reaction in a

complex manner that depends on the position of the mismatch and the length of the toehold (see also below). To describe this behavior theoretically, we adapted a previously established Markov-chain model for mismatch-free strand displacement reactions.²⁸ To predict a rate constant, it is sufficient to consider a single target in a bath of invader molecules at concentration c . We consider the association of the invader with the target to occur in a sequential manner starting from position 1 of the toehold with further invader–target base pairs forming in single base-pair steps. This way one can devise a simple one-dimensional energy landscape with the furthest incumbent–target base pair as reaction coordinate (Figure 1).

Position zero denotes the state of the unbound invader (state A in Figure 1) to which we assign a free energy of zero. The free energies of all subsequent states shall be based on a minimal set of free parameters that are all taken as positive values. To establish the first invader–target base pair, the free energy of the system changes due to the loss of translational and orientational degrees of freedom of the combined molecule, the uptake of an invader out of the unbound invader pool (equaling the negative chemical potential of the unbound invaders), and the establishment of a single pair of base-pairing interactions (energy gain of $-\Delta G_{BP}$). The total free-energy change can then be written as

$$G(1) = \Delta G_{\text{assoc}} - k_B T \ln(c/c_0) - \Delta G_{BP} \quad (1)$$

where c is the free invader concentration and ΔG_{assoc} comprises the cost for the reduced conformational freedom and the

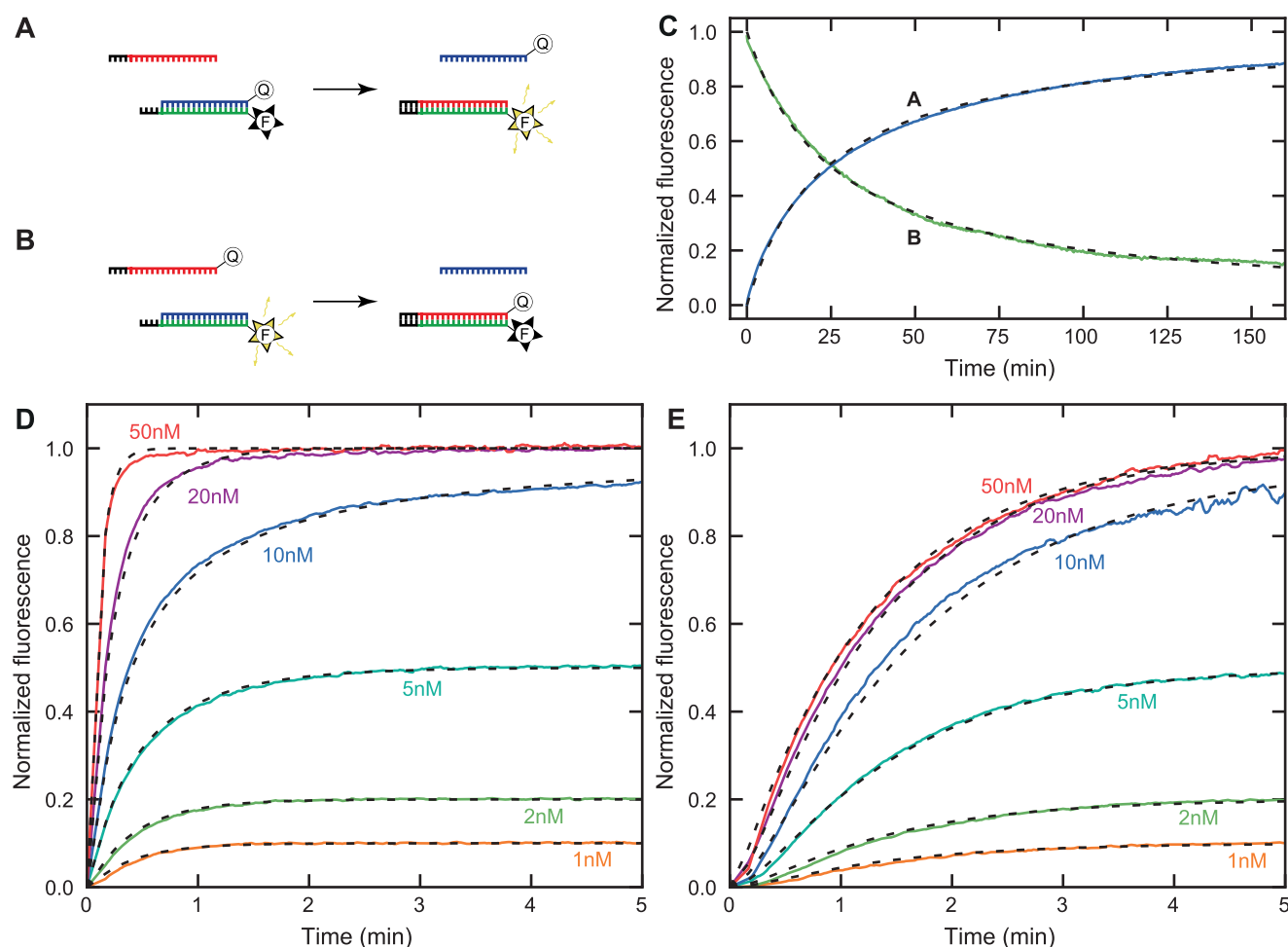


Figure 2. Measurement principle and fits to the reaction kinetics. Measurement scheme for strand displacement reactions using a target tagged with a fluorophore (F) and either (A) an incumbent tagged with a quencher (Q) or (B) an invader tagged with a quencher. (C) Fluorescence signal over time measured for a perfect matching invader with a 4 nucleotide toehold using the two alternative quencher placements (solid lines). The concentrations of all strands was 10 nM. Fitting the data to second order kinetics (dashed lines) yielded rate constants of $k_A = (7.2 \pm 0.2) 10^4/\text{Ms}$ and $k_B = (6.5 \pm 0.2) 10^4/\text{Ms}$. (D) Displacement kinetics measured for different concentrations of a perfectly matching invader with a 10 nt toehold invader (solid lines). A global fit to the data using second-order kinetics (see eq 9) is shown as black dashed lines. (E) Displacement kinetics measured for different concentrations of an invader with a mismatch at position 5 and a 10 nt toehold (solid lines). The reaction was described by a two-step process (dashed lines, see eq 11). The concentration of incumbent and target strands was 10 nM for all measurements in D and E.

standard chemical potential of the invader (at a standard concentration of $c_0 = 1 \text{ M}$).

For the subsequent states, the invader progressively base-pairs (zips-up) with the toehold bases such that the system's free energy decreases by $-\Delta G_{\text{BP}}$ per step until all H toehold bases are paired and state B is reached (Figure 1). We neglect any sequence dependence of ΔG_{BP} and rather apply a constant average base-pairing free energy. Beyond position H , the invader starts to displace the incumbent. Initiation of the displacement process requires an additional free-energetic penalty ΔG_p (between H and $H+1$)²⁸ due to the formation of an additional ssDNA overhang at the incumbent strand. For subsequent strand displacement (branch migration) steps, invader–target base pairs can only be formed by breaking incumbent–target base pairs (state C in Figure 1) such that no net free energy change occurs. This changes when the invader forms a mismatch with the target at position $H + M$ (position M of the incumbent–target complex), where a mismatch penalty ΔG_{MM} needs to be introduced (Figure 1). The subsequent strand displacement states are again isoenergetic with the mismatch state. When the invader approaches the

target end (state D, state position $H + N$), the incumbent can dissociate. Dissociation provides a free-energy gain in which the initial costs of the invader association (eq 1) are liberated.

To obtain a kinetic model of the strand displacement reaction, one needs to define the rate constants for all single base-pair steps of the reaction. By the principle of detailed balance, the ratio between the forward (indicated by +) and the backward (indicated by −) rate constants of subsequent positions is related to the free-energy difference $G_{n+1} - G_n$ between these positions:

$$\frac{k_n^+}{k_{n+1}^-} = e^{-(G_{n+1} - G_n)/k_B T} \quad (2)$$

Let k_{bp} be the rate constant at which a bound invader forms an additional base pair with the toehold. Detailed balance provides then the opening rate (fraying rate) of a formed base pair:

$$k_{1 < n < H}^- = k_{\text{bp}} e^{-G_{\text{bp}}/k_B T} \quad (3)$$

This relation can, to a first approximation, be extended to the dissociation of the first base pair, which should be limited by single base-pair fraying.²⁸ Detailed balance provides then for the bimolecular association of the invader to the toehold:

$$k_{\text{bind}} = k_{\text{bp}} e^{-\Delta G_{\text{assoc}}/k_{\text{B}}T} \frac{c}{c_0} \quad (4)$$

During branch migration, the establishment of a new invader–target base-pair requires first the breakdown of the corresponding incumbent–target base pair and afterward rearrangements of all involved strands. The branch migration rate for a single base pair step is thus considerably slower than k_{bp} .²⁸ This can be seen as an increased transition barrier by ΔG_{BM} between subsequent strand displacement steps compared to the simple hybridization of invader base pairs to the toehold (see local energy barriers in Figure 1). The branch migration rate is thus given as

$$k_{\text{bm}} = k_{\text{bp}} e^{-\Delta G_{\text{BM}}/k_{\text{B}}T} \quad (5)$$

For the branch migration step that establishes the mismatch, the rate ratio between forward and backward steps is reduced by the mismatch penalty. We assume here that the forward step that establishes the mismatch occurs with unchanged rate constant ($k_{H+M-1}^+ = k_{\text{bm}}$) since an intact incumbent–target base pair needs to be broken. The backward rate constant that removes the mismatch thus increases according to

$$k_{H+M}^- = k_{\text{bm}} e^{\Delta G_{\text{MM}}/k_{\text{B}}T} \quad (6)$$

This choice reflects the facilitated fraying of the mismatched base pairs,³³ allowing a facilitated incumbent progression (seen as a reduced transition barrier for the backward step in Figure 1). In contrast, for initiation of branch migration that occurs at the expense of ΔG_{p} , we assume the forward rate to be reduced and the backward rate to be unchanged (normal branch migration step).

Based on these considerations, the rate constants for all single steps of this initial one-dimensional Markov chain model are fully parametrized by ΔG_{assoc} , ΔG_{BP} , ΔG_{p} , ΔG_{BM} , ΔG_{MM} , and k_{bm} (with the latter defining the absolute time). Additionally, we have to consider that the incumbent can dissociate before the invader reaches the last target base pair,²⁴ due to spontaneous fraying of the incumbent base pairs. Such a possibility can be incorporated by implementing “short-cuts” that terminate strand-displacement at earlier positions (blue arrows in Figure 1). The rate constants of the short-cut steps are estimated as

$$k_n^{\text{off}} \approx k_{\text{bp}} e^{-n\Delta G_{\text{bp}}/k_{\text{B}}T} \quad (7)$$

which is position-dependent and decreases approximately exponentially with the number of remaining incumbent–target base pairs (see note S2 for the estimation of short-cut rates).

To allow comparison of the model with experimental data, one has to calculate the global rate at which strand displacement occurs. This rate is provided by the reciprocal mean-first-passage time, i.e., the mean time at which state *D* (incumbent dissociation) is reached when initially starting at state *A* (unbound invader). The parametrized model allows us to derive an expression for the mean first-passage time by introducing boundaries for a transmission mode under the assumption of a stationary process (see Figure S1 and note S1). The global strand displacement rate of the model can then be compared with experimentally measured rates.

Mean Reaction Times from Experimental Data. To experimentally measure strand displacement kinetics, we labeled the target-strand with a fluorophore and depending on the experiment, either the incumbent (Figure 2A) or the invader strand (Figure 2B) with a quencher. This way, either a fluorescence increase or a fluorescence decrease is measured as the reaction proceeds. We first verified that both labeling strategies produce the same reaction kinetics for perfectly matching invaders (Figure 2C). For economic reasons we used quencher-labeled incumbent-strands to characterize invader mismatches and quencher-labeled invader-strands to characterize incumbent mismatches. We next fit the fluorescent curves using a simplified three-state ODE model (see below). From the fitted parameters, we inferred the mean reaction time assuming a constant invader concentration as simulated by the model. Within the three-state model, the invader and the incumbent–target complex form first a reversible intermediate followed by an irreversible displacement step



with *I* being the invader, *Q* the incumbent, and *T* the target. To avoid overfitting of the kinetics with three individual rates, we distinguished two limiting cases for the reaction:

A low population of the three-strand intermediate: This limit is obtained for fast displacement reactions with $k_1 \ll k_2$ (e.g., systems without mismatches) and/or fast toehold dissociation with $k_1 \ll k_{-1}$ (short toeholds). In this case the reaction can be simplified to^{28,34}



This reaction is first order in *I*, first order in *QT*, and second order overall, with *k* being an effective second-order rate constant. Solving the rate equation set (see the Methods) allowed fitting of experimentally measured data. Good agreement was found over a large range of different invader concentrations (Figure 2D). From the second-order rate constant *k* obtained in the fit, the mean reaction time in the limit of constant invader concentration can be calculated from

$$\Delta t_{\text{mean}} = 1/(ck_1) \quad (10)$$

A high population of the three-strand intermediate: This case is obtained if ck_1 is faster than the two other rate constants that describe the dissociation of the intermediate. It is seen for invaders with long toeholds and toehold-proximal mismatches.²⁴ The simplified rate equation has in this case to include the intermediate complex



where k_1 is a second order rate constant and k_2 is a first order rate constant. Solving the rate equation set (see Methods) allowed fitting of experimentally measured data over a large range of varying invader concentrations (Figure 2E). The observed lag phase in the beginning is typical for consecutive reactions with more than one rate limiting step. At higher invader concentrations, the reaction kinetics saturated in agreement with the first-order character of the second reaction step. For this two-step process, the mean reaction time in the limit of constant invader concentration is given by

$$\Delta t_{\text{mean}} = \Delta t_1 + \Delta t_2 = \frac{1}{ck_1} + \frac{1}{k_2} \quad (12)$$

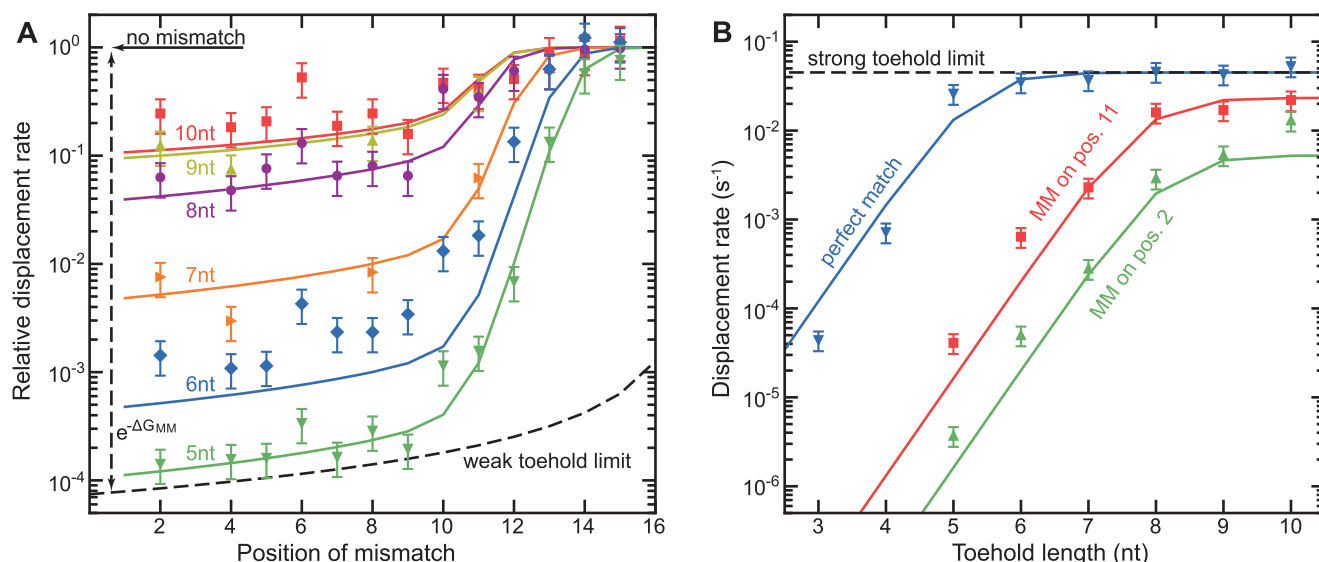


Figure 3. Strand displacement rates for single invader–target mismatches together with global fits of the Markov chain model. (A) Relative strand displacement rate (compared to the rate of a perfectly matching invader) vs mismatch positions for different toehold length (as indicated in the plot). Using relative rates removed residual variability between the five different target sequences used. Experimentally obtained rates are shown as solid scatter symbols. Results of a global fit to the whole data set are shown as solid lines. (B) Measured absolute strand displacement rate vs toehold length plotted for a perfectly matching invader as well as single mismatches at positions 2 and 11 (solid scatter symbols). Results of a global fit to the data using the energy parameters from A but an adjustable branch migration rate k_{bm} are shown as solid lines.

Table 1. Parameter Sets Used to Describe the Strand Displacement Reactions^a

data set	ΔG_{BP} ($k_B T$)	ΔG_{assoc} ($k_B T$)	ΔG_P ($k_B T$)	ΔG_{BM} ($k_B T$)	ΔG_{MM} ($k_B T$)	k_{bm} ($10^3/s$)	k_{bp} ($10^6/s$)
this study	2.52	2.5 ± 0.2	3.5 ± 0.2	7.4 ± 0.2	9.5 ± 0.2	36 ± 2	59 ± 11
Machinek ²⁴	2.51	4.0 ± 0.2	5.4 ± 0.4	8.5 ± 0.3	8.5 ± 0.2	53 ± 4	206 ± 6
ref		3.2^{35}	3.4^{28}	8.9^{28}	$9.5/9.6$	$10-20^{28}$	75^{28}

^aThe quantities in the first two rows were obtained from fits to the indicated experimental data sets, except ΔG_{BP} and k_{bp} . ΔG_{BP} and the reference value for ΔG_{MM} were calculated using DINAMelt.³⁶ k_{bp} was calculated from k_{bm} and ΔG_{BM} according to eq 5. Other reference values were taken from literature as indicated.

In the subsequent experiments, one of the two simplified reaction schemes could, within error, describe the measured kinetics and was correspondingly applied.

Modeling Strand Displacement Reactions with Single Invader–Target Mismatches. To test whether our Markov-chain model can describe displacement reactions in the presence of invader–target mismatches, we recorded the strand displacement kinetics for a variety of different DNA substrates all containing a 17 nt branch migration domain and extracted the corresponding global strand displacement rates. Compared to a previous measurement,²⁴ we used an extended range of toehold lengths (3–10 nts) such that model fitting should provide a fine-tuned parameter set. In addition, we used symmetric AGA trinucleotides in the incumbent strand to introduce C–C invader–target mismatches that were flanked on either side by A–T base pairs. All introduced mismatches had thus the same mismatch penalty within the nearest neighbor approximation.³⁵ Locating multiple AGA trinucleotides on an incumbent allowed us to probe several mismatch positions with one sequence but also to place two mismatches at defined distances from 2 to 13 bp. To cover all mismatch positions ranging from $M = 2-15$, five different incumbent–target sequences were required in total (Table S1).

We first determined the relative displacement rate as a function of the mismatch position for the different toehold lengths (Figure 3A), which resembled previous results.²⁴ Toehold-proximal mismatches exhibited more strongly re-

duced rates than more distal mismatches. Mismatches close to the toehold-distal end barely differed from the perfectly matching system in agreement with incumbent dissociation prior to mismatch enclosure.

We next carried out a global fit of the obtained relative displacement rates for all combinations of mismatch position and toehold length. To this end, the base-pairing free energy ΔG_{BP} for the toehold binding was fixed to an average value of $\Delta G_{BP} = 2.52 k_B T$ as obtained from the DINAMelt web server³⁶ (see Table S1). The remaining energy parameters of the random walk model, i.e., ΔG_{assoc} , ΔG_P , ΔG_{BM} , and ΔG_{MM} were taken as free parameters (see Table 1). The fit curves reproduced the measured complex behavior of the strand displacement rates depending on the mismatch position and the toehold length remarkably well (Figure 3A, Figure S2A). Allowing incumbent dissociation before full displacement was essential for good agreement with the measurements (Figure S2B).

Furthermore, we ensured that the model prediction was rather insensitive to how the mismatch penalty was introduced. A decrease of the forward rate versus an increase of the backward rate (as used here) at the mismatch position provided similar results (Figure S3A,B). Overall, these findings suggest that our random walk model captures correctly the main features of the strand displacement process in the presence of mismatches. This conclusion is supported by the reasonable free-energy parameters that were obtained from the

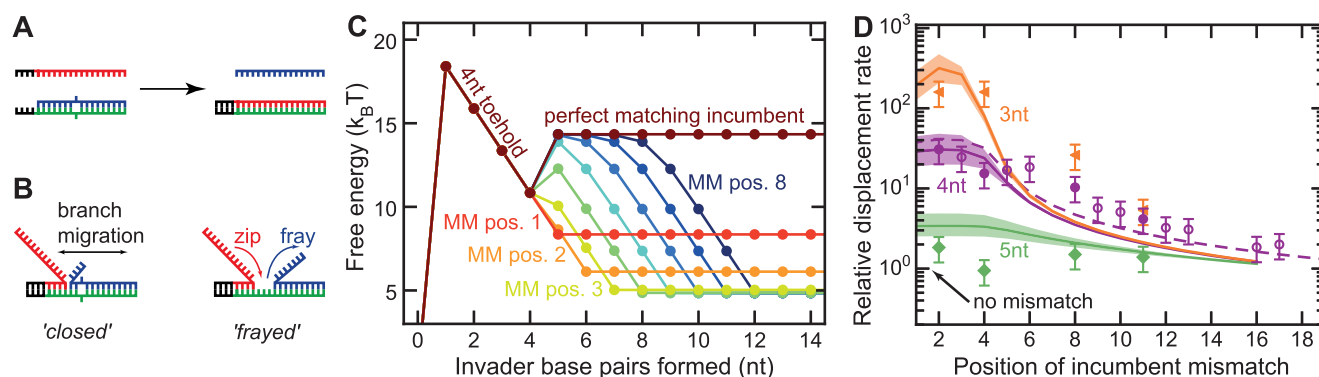


Figure 4. Strand displacement in case single of incumbent–target mismatches. (A) Scheme of a reaction with a single mismatch on the incumbent strand which gets “repaired” during the displacement reaction. (B) Two possible states of an incumbent–target mismatch. In the “closed” state, the mismatch is surrounded on both sides by paired incumbent–target bases. In the “frayed” state, the incumbent–target base pairs opened up toward the invader by fraying. (C) Simplified free-energy landscapes of the strand exchange reaction when calculating the free-energy gain of the “mismatch repair” by combining the two states depicted in B. Shown are landscapes for the perfectly matching incumbent as well as for single mismatches placed at positions 1 to 8. Free-energy parameters were taken from the global fit shown in Figure 3. (D) Measured relative strand displacement rates vs mismatch position for toehold lengths of 3, 4, and 5 nt (filled and open symbols, which are for 17 nt and 22 nt long displacement domains, respectively). Predictions of the Markov chain model for a 17nt long displacement domain are shown as solid lines and for a 22nt long displacement domain as dashed line. Colored areas around the predictions represent the confidence interval based on the errors of the energy parameters.

fit of $\Delta G_{\text{assoc}} = (2.5 \pm 0.2)k_{\text{B}}T$, $\Delta G_{\text{P}} = (3.5 \pm 0.2)k_{\text{B}}T$, $\Delta G_{\text{BM}} = (7.4 \pm 0.2)k_{\text{B}}T$, and $\Delta G_{\text{MM}} = (9.5 \pm 0.2)k_{\text{B}}T$ (see Table 1 and Discussion). Most importantly, the obtained mismatch penalty of $\Delta G_{\text{MM}} = (9.5 \pm 0.2)k_{\text{B}}T$ was in good agreement with the value of $9.5k_{\text{B}}T$ calculated with DINAMelt.³⁶

The branch migration rate k_{bm} that sets the absolute time constant of the model was obtained from fitting the measured displacement rates as function of the toehold length for single target sequences (Figure 3B). This fit used the previously obtained energy parameters with k_{bm} being the only free parameter. A global fit provided again a good agreement with the measurements. The obtained value of $k_{\text{bm}} = (36 \pm 2)10^3/\text{s}$ was in good agreement with the value assumed by Srinivas et al.²⁸ of up to $20 \times 10^3/\text{s}$.

Using an analytical solution of a simplified version of our model that does not include incumbent dissociation before full displacement (see note S3) shows that in the absence²⁸ and presence of a mismatch the displacement rate scales for short toeholds exponentially with the toehold length, which is also seen in the measurements (Figure 3B). Furthermore, the free energy penalty of a mismatch provides an absolute limit for the relative reduction in displacement rate (dashed curve in Figure 3A). The largest rate reduction relative to the mismatch-free system of $e^{-\Delta G_{\text{MM}}/k_{\text{B}}T}$ is obtained for toehold-proximal mismatches. Lower reductions are obtained for more distal mismatches in agreement with the data. In the long toehold limit, the displacement rate saturates with increasing toehold length, since the displacement reaction itself, rather than toehold occupancy becomes limiting (Figure 3B, note S3).

As an independent test, we applied the random walk model to the published data set recorded by Machinek et al.²⁴ The base-pairing energy was fixed to $\Delta G_{\text{BP}} = 2.51k_{\text{B}}T$. Again, a good agreement between data and fit was obtained (Figure S4). The best-fit parameters of $\Delta G_{\text{assoc}} = (4.0 \pm 0.2)k_{\text{B}}T$, $\Delta G_{\text{P}} = (5.4 \pm 0.4)k_{\text{B}}T$, $\Delta G_{\text{BM}} = (8.5 \pm 0.3)k_{\text{B}}T$, $\Delta G_{\text{MM}} = (8.5 \pm 0.2)k_{\text{B}}T$, and $k_{\text{bm}} = (5.3 \pm 0.4) \times 10^4/\text{s}$ were very similar to the parameters obtained for our data (see Table 1). Overall, the good agreement with the two experimental data sets suggests

that our model can describe displacement reactions in the presence of invader mismatches.

Predicting Strand-Displacement Rates for Incumbent–Target Mismatches. We next tested whether our parametrized random walk model can be applied to predict the influence of preexisting incumbent–target mismatches on the displacement reaction (Figure 4A). This mismatch type recently gained considerable interest since it allows acceleration of the reaction as well as the establishment of hidden thermodynamic drives.²³ Instead of being a penalty during the displacement process, a mismatch in the incumbent–target duplex acts rather as a benefit. The invader is “repairing” the mismatch during the displacement, such that the backward reaction is disfavored and in turn the forward direction, i.e., incumbent displacement, becomes favored (Figure 4D).

We first measured strand displacement rates for incumbent target mismatches using the same mismatch type as before. In addition to a displacement domain of 17 nt we also tested a 22 nt domain to better explore the length range at which incumbent mismatches act. In agreement with previous data,²³ incorporation of incumbent mismatches was found to accelerate the displacement rates. This effect was again most pronounced for short toeholds and toehold-proximal mismatches. To model this behavior, we adapted our random walk model by introducing a negative mismatch penalty. It was incorporated by increasing the forward rate at the mismatch position, i.e., at the step that leads to mismatch elimination, in analogy with our previous strategy. The modeling described the general trend, i.e., an accelerated reaction for early mismatches and short toeholds. The rate increase in the model decayed, however, much faster with increasing distance of the mismatch from the toehold than experimentally observed (Figure S5). To explain this difference, we considered fraying of the incumbent–target base pairs when the branch point approaches the mismatch as seen in coarse-grained simulations.²³ In this case, the base pairs preceding the mismatch can open as the branch migration point draws near to mitigate the mismatch penalty. We modeled this effect by considering a combined macrostate of a “frayed” state, with

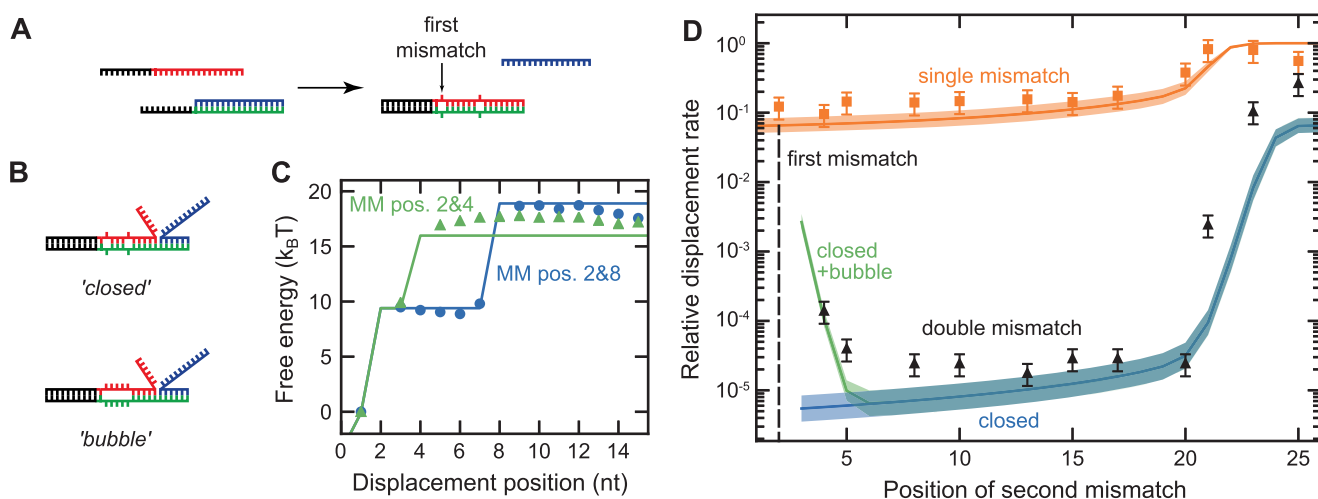


Figure 5. Strand displacement in case of two invader–target mismatches. (A) Scheme of the reaction in which two mismatches are created. (B) Two possible states in case of two invader–target mismatches. The “closed” state comprises only paired bases around the mismatches, whereas for the “bubble” state the bases between the mismatches and the mismatches themselves are unpaired. (C) Simplified free-energy landscape for two invaders with a first mismatch at position 2 and a second mismatch at position 4 (green line) or position 8 (blue line) compared to free energies obtained by oxDNA simulations (scatter symbols—sampling error is smaller than symbol size). Free-energy values are relative to the free-energy of the first displaced base of the incumbent. (D) Measured relative strand displacement rates vs position of the second mismatch (solid black triangles). The first mismatch was located at position 2 and the toehold length was 10 nt. Theoretical predictions are shown for the different models of the energy penalty of double mismatches considering either the closed state only (blue line) or the closed and the bubble states (green line). Single mismatch data together with the corresponding prediction are shown in orange. Colored areas around the model predictions represent the confidence interval based on the errors of the energy parameters.

unpaired bases, and a “closed” state, with paired bases before the mismatch (Figure 4B).

The free-energy penalties of each of the states when introducing a mismatch into a DNA duplex are then given as

$$\Delta G_{\text{FRAY}}(n) = (M - n)\Delta G_{\text{BP}} \quad (13)$$

$$\Delta G_{\text{CLOSED}} = \Delta G_{\text{MM}} \quad (14)$$

where n and M are the positions of the branch point and of the mismatch with respect to the fully bound toehold. Fraying is thus energetically more probable the more the branch point approaches the mismatch position. The position-dependent free energy is obtained from the combined macrostate. For the free energy of the states after mismatch elimination ($n \geq M$ relative to the state with fully bound toehold we then get (see note S4):

$$G(H + n) = G(H + M) \quad (15)$$

$$G(H + n) = G(H) + k_B T \ln(e^{-\Delta G_{\text{FRAY}}(0)/k_B T} + e^{-(\Delta G_{\text{CLOSED}} - \Delta G_P)/k_B T}) \quad (16)$$

The free energy for the displacement positions before the mismatch ($0 < n < M$) is given as

$$G(H + n) = G(H + M) - k_B T \ln(e^{-\Delta G_{\text{FRAY}}(n)/k_B T} + e^{-\Delta G_{\text{CLOSED}}/k_B T}) \quad (17)$$

The resulting free energy profiles exhibit an approximately linear energy decrease starting approximately 3 bp before the mismatch position in agreement with coarse-grained molecular simulation results.²³ Using this energy landscape for the modeling shows accelerated rates also for mismatches that are more distant from the toehold in agreement with the measurements. In addition, the magnitude of the rate increase as function of the toehold length is correctly described. For

toehold-proximal incumbent mismatches, the frayed state is preferred, which effectively abolishes the initiation penalty for mismatches up to position four. High displacement rates are obtained over this entire range, since the fraying effectively elongates the toehold.

Overall, we note that our model can correctly describe increased displacement rates in case of incumbent–target mismatches by simply using the parameters obtained from displacement reactions with invader–target mismatches. The incorporated fraying for incumbent–target mismatches can be neglected for invader–target mismatches, where a distribution of the mismatch penalty over multiple positions has only a minor influence (Figure S3C,D).

Predicting Rates for Two Invader–Target Mismatches. To test our model further, we characterized and modeled the strand displacement kinetics in the presence of two invader–target mismatches for a 27 nt displacement domain. To this end, we used invader strands with 10 nt toeholds, with a first mismatch located at position 2 and a second mismatch located at varying positions (4–25) along the displacement domain. The rates for the corresponding single-mismatch constructs exhibited less than a 10-fold reduction compared to the mismatch free systems (orange data in Figure 5D).

In contrast, the strand displacement rates in the presence of a second mismatch were significantly reduced by up to 5 orders of magnitude (Figure 5D). Notably, our model, which was only parametrized on the single-mismatch data, could successfully predict such a strong rate reduction. Widely separated double mismatches showed a much better agreement with the model expectation than very close double mismatches, which showed enhanced displacement rates.

To understand the facilitated displacement in case of two nearby mismatches, we considered an additional bubble formation between proximal mismatches similar to the fraying

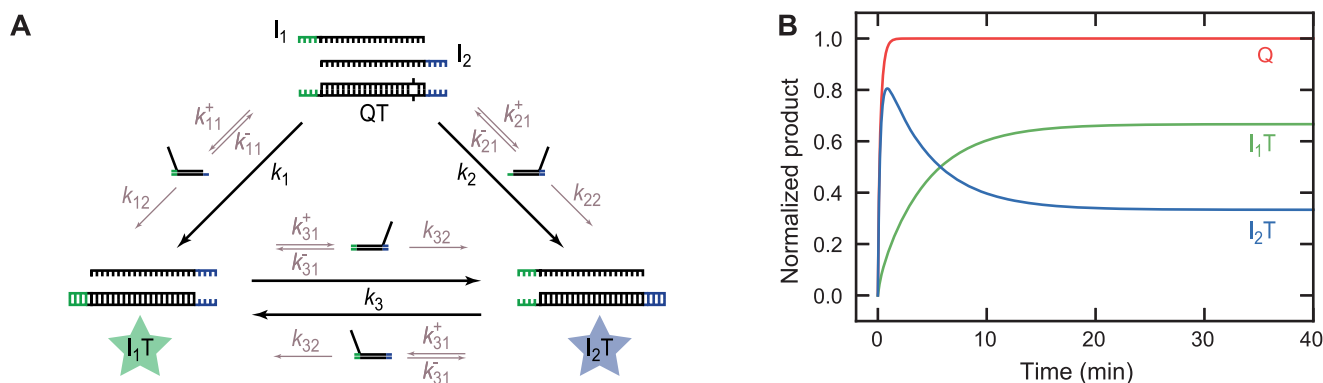


Figure 6. Modeling pulse-generation from two competitive displacement reactions using incumbent–target mismatches. (A) Scheme of the two-toehold competitive displacement reactions used to create a pulse signal.²³ Each strand displacement reaction is subdivided into two steps—the reversible formation of a three-strand intermediate and the irreversible actual displacement. Production of I_1T or I_2T provides a respective blue or green fluorescent signal. (B) Numerical solution of the pulse-generating competitive displacement reactions using the predicted rate constants from the first passage model. Due to an asymmetric placement of the mismatch, I_2T is strongly favored over I_1T on short time scales, while on long time scales the thermodynamic equilibrium is reached such that $[I_2T]$ decreases again. Initial concentrations were $[I_1] = 40$ nM, $[I_2] = 20$ nM and $[QT] = 10$ nM.

before. A single mismatch introduces to either nearest neighbor base pair a structural distortion, which constitutes a high penalty to introduce this new structural phase in-between a stretch of B-form duplex DNA.

The formation of a denaturation bubble between two nearby mismatches would avoid introducing the mismatch penalty twice such that it may, on short distances, be favored over two isolated mismatches with paired bases in-between. To test this possibility, we used oxDNA simulations²⁷ to probe the free-energy profile of strand displacement with mismatches at positions 2 and 4 or 2 and 8 of the branch migration domain. When the oxDNA potential was augmented to more accurately reflect the destabilization of duplexes by single mismatches, we observed that base pairs tended to form between mismatches at positions 2 and 8, but not between 2 and 4. Moreover, we observed a decrease in the overall free-energy penalty for incorporating the two mismatches of approximately $1k_B T$ for mismatches at positions 2 and 4. This observation supports the hypothesis that bubble formation can facilitate the incorporation of two nearby mismatches during strand displacement (see Figure 5C, Figure S6C).

To test this hypothesis further, we demonstrated that the displacement reactions were not dependent on base-pairing in between two mismatches separated by 2 bp by introducing an additional mismatch at this position (see below).

To model bubble formation, we considered a combined macrostate composed of a closed state with two isolated mismatches and the bubble state. The free energy for bubble formation comprises an initiation penalty for disrupting the helix phase, the disruption of the base pairs energy of $\Delta G_{BP}/b_p$ and an entropic loop penalty for confining the ssDNA strands of $\Delta G_L/b_p$.³⁷ We estimated the initiation penalty combined with the free energy for disrupting the first base pair to be on the order of ΔG_{MM} . The free energies of the two states relative to the perfectly matching invader are then given as

$$\Delta G_{\text{CLOSED}} = 2\Delta G_{\text{MM}} \quad (18)$$

$$\Delta G_{\text{BUBBLE}}(d) = \Delta G_{\text{MM}} + d(\Delta G_{BP} + \Delta G_L) \quad (19)$$

where d is the distance between the two mismatches. We set ΔG_L to $0.8k_B T$ in agreement with predictions from DINAmelt³⁶ and NUPACK³⁸ (Figure S6B). Assuming that

introduction of the first mismatch introduces a full mismatch penalty while the penalty of the second mismatch is reduced according to the total free energy of the macrostate; we get for the position dependent free energies for $M1 \leq n < M2$:

$$G(H + n) = G(H) + \Delta G_p + \Delta G_{\text{MM}} \quad (20)$$

and for $n \geq M2$:

$$G(H + n) = G(H) + \Delta G_p + k_B T \ln(e^{-\Delta G_{\text{CLOSED}}/k_B T} + e^{-\Delta G_{\text{BUBBLE}}(d)/k_B T}) \quad (21)$$

with $M1$ and $M2$ being the positions of the first and the second mismatch, respectively. The resulting free-energy profile reflects the overall shape predicted by oxDNA simulations (Figure 5C). Considering bubble formation in the random-walk model finally allowed the reproduction of the increased rates seen for very close double mismatches (Figure 5D). The exact distribution of the free energy penalties between the two mismatches had a negligible impact on the predicted reaction rates (Figure S7). Overall, our augmented model is able to predict the kinetics of strand displacement reactions in the presence of double mismatches.

We additionally tested strand displacement on a shorter, 17 nt long displacement domain in the presence of double mismatches. Also in this case, a strong rate reduction due to the second mismatch was observed (Figure S8A). However, it was—except for very widely separated mismatches—smaller than predicted by the model including only the closed state. Considering an additional bubble state could describe the rates for very close mismatches. The hypothesis of bubble formation was supported by testing systems with an additional mismatch in-between two nearby mismatches, which did not alter the displacement rate (Figure S8D,E). The experimental rates for mismatches at intermediate separations were, however, still higher than the model prediction. A possible explanation for this discrepancy could be that a toehold-bound invader additionally started invasion from the toehold-distal end in case of short displacement domains and long reaction times (see “backdoor invasion” scheme in Figure S9A). Alternatively, the displaced incumbent could loop back and base-pair with the nucleotides in the bubble between two mismatches, such that a pseudoknot-like structure would form (Figure S9A).

Additional experiments did not support a “backdoor invasion” (Figure S8B,C,E), while pseudoknot-like structures were not found to be stable in simulations. Incorporating a bubble-pairing mechanism (see note S5), could describe the experimental data measured for the short displacement domains somewhat better. Notably, such additional effects are, however, not relevant for longer displacement domains (Figure S9B).

Modeling a Strand Displacement Reaction Network.

After successfully describing single displacement reactions with mismatches, we show in the following how our predictions can be applied to model a small DNA reaction network. In particular, we use the pulse-generation reaction of a competitive two-toehold displacement system of Haley et al.²³ (see Figure 6A). It consists of an incumbent–target complex with a displacement domain of 20 bp and 4 nt toeholds on either side of the target strand. Rate differences for the incumbent displacement from one vs the other toehold were achieved by an asymmetrically placed incumbent–target mismatch at position 3 relative to one of the toeholds. When adding the two corresponding invaders I_1 and I_2 , an initial strongly favored displacement for I_2 is achieved (in Figure 6B). Over long time scales I_1 can displace I_2 and thermodynamic equilibrium is reached, such that the initial rapid increase in $[I_2T]$ is followed by a decrease providing a pulse pattern (Figure 6B).

To predict the kinetics of the pulse-generation reaction, we modeled the kinetics of each of the four displacement reactions using the three-state approximation (eq 8) where first a three-strand intermediate is formed by a reversible step, with second-order association rate constant k_1^+ and first-order dissociation rate constant k_1^- , followed by the actual displacement step with first-order rate constant k_2 . This approximation is also motivated by the observation that the experimental displacement kinetics can be described by such a scheme (Figure 2). For a given displacement reaction, these rates can be conveniently obtained from the first passage model (see note S6).

To model the pulse generation of the two-toehold system, we considered four displacement reactions (see Figure 6A). For the initial displacement reactions, a displacement domain of 20 bp and an incumbent–target mismatch at either position 3 or 18 was considered for each invader. For the subsequent displacement reactions, in which one invader displaces the other, a 24 bp displacement domain with four terminal invader–target mismatches was modeled equally for both displacement directions. To calculate the three rate constants for each displacement reaction we applied the parameter set of this study. The rate constants were used in a set of ordinary differential equations that described the reaction scheme (see note S7). A numeric solution of the equation set provided the time dependencies of incumbent release and formation of the invader target complexes I_1T and I_2T (Figure 6B). The obtained kinetics was in good qualitative agreement with the experimental data from Haley et al.²³ In particular, the pulse pattern of I_2T could be reproduced. Thus, the model predictions can be applied to describe displacement reaction networks including the different types of mismatches. We note, however, that compared to Haley et al.²³ the modeled kinetics was faster by more than 1 order of magnitude. In part, this discrepancy may be due to the interference of fluorophore and quencher labels in the experiment of Haley et al.;²³ however, sequence-dependent variations of this magnitude are known

for strand displacement reactions³⁴ (see the Discussion below).

DISCUSSION AND CONCLUSION

The introduction or the elimination of mismatches in strand displacement processes is a convenient tool to tune the rate of these widely applied reactions³⁹ by orders of magnitude.^{23,24} In this study, we provided a comprehensive characterization of mismatch-based strand displacement. Most importantly, we developed and parametrized a quantitative model that is able to explain and predict the rates of these reactions. We emphasize that a unique parameter set is able to describe single and double invader mismatches as well as incumbent mismatches. Our model produces reliable results without the need for time-consuming molecular dynamics simulations. Furthermore, we could successfully apply the model predictions to simulate a simple displacement reaction network that included mismatch elimination and generation (see Figure 6). To this end, each displacement reaction was approximated by a two-step process. The corresponding rate constants were a direct output of the random walk model and software code to calculate the rate constants is made available to the public.⁴⁰ We expect that, using this approach, other and more complex reaction networks can be conveniently simulated.

The introduced random-walk model was parametrized using displacement reactions with single invader mismatches. Two independent data sets could be described with a similar set of parameters (Table 1). The obtained penalties for the used C–C mismatches of $(9.5 \pm 0.2)k_B T$ and $(8.5 \pm 0.2)k_B T$ were in agreement with the mismatch penalties obtained from DNA melting studies as used by DINAmelt³⁶ and NUPACK.³⁸ To describe reactions with other mismatches, it should therefore be possible to use published mismatch penalties from melting studies.

The value of the branch-migration step time in strand displacement reactions is currently still under intense discussion, since it is difficult to access this parameter experimentally. Values ranging from tens^{41–43} to hundreds^{28,34,44} of microseconds have been reported for comparable experimental conditions. We obtained step times of $28 \pm 2 \mu s$ and $19 \pm 2 \mu s$ for the current data set and the data sets of Machinek et al.,²⁴ respectively. This suggests that the average branch-migration step time is in the order of 10^{-5} – 10^{-4} s. The obtained values for ΔG_{BM} of $(7.4 \pm 0.2)k_B T$ and $(8.5 \pm 0.3)k_B T$ were similar to the values of $6.1k_B T$ and $8.9k_B T$ inferred by Srinivas et al.²⁸ for two different energy landscape models. Linked to k_{bm} and ΔG_{BM} is the base-pair formation rate k_{bp} , for which we found values in the range of $10^7/s$ to $10^8/s$, again comparable to Srinivas et al.²⁸ (see Table 1). The values that were obtained for the association energy ΔG_{assoc} of $(2.5 \pm 0.2)k_B T$ and $(4.0 \pm 0.2)k_B T$ were close to the value of $3.2k_B T$ from DNA melting experiments.³⁵ Furthermore, the value for the initiation penalty ΔG_p obtained for our data set of $(3.5 \pm 0.2)k_B T$ was similar to the value of $3.4k_B T$ estimated by Srinivas et al.²⁸

While the obtained parameters are similar to previously published values, differences beyond the confidence intervals of our data persist. On the one hand, our model provides a reasonable but not perfect description (given the error bars) of the experimental data, such that the fit errors of the parameters are likely an underestimate. On the other hand, strand displacement rates are known to exhibit a considerable sequence dependence and can vary by more than 1 order of

magnitude for given lengths of toehold and displacement domain.³⁴ This sequence dependence is caused by several factors, such as a sequence-dependent free energy of base pairing and as a consequence considerable sequence dependencies for the rates of base-pair opening and branch-migration.⁴⁵ Furthermore, fluorescent labels that can possess hydrophobic interactions with the DNA bases can impact the measured kinetics. While for the present analysis performed on a number of different sequences, sequence-dependencies seem to play a minor role, caution should be taken when using unusually slow or fast strand displacement systems.³⁴

A systematic investigation into how the local sequence influences toehold binding as well as branch migration rates is, unfortunately, not available yet. Further work is certainly warranted to reveal the sequence-dependence of the local branch migration rates and all other model parameters.

To cope with sequence dependencies, we suggest to always apply the base-pairing free energy, as inferred from SantaLucia et al.,³⁵ for the particular toehold and mismatch penalty of interest. Furthermore, we note that in the weak toehold limit, at which many displacement reactions are operated, the mismatch penalty sets a well-defined (position-dependent) rate reduction compared to the mismatch-free system (Figure 3A). It should therefore in many cases be sufficient to characterize experimentally the mismatch-free system and rescale correspondingly the time scale of the prediction, i.e., k_{bm} . This approach could, e.g., successfully be applied to the pulse generating reaction (Figure 6). In general, we recommend the parameters obtained for the data set of this study (see Table 1), since it included a more comprehensive range of toehold lengths compared to Machinek et al.²⁴

Our modeling allowed us to extract values for the branch migration rate as well as all free-energy parameters such that no assumptions on the parameters had to be made, in contrast to modeling mismatch free systems.²⁸ It could thus also be used for a detailed study of the sequence dependence of the various parameters. We note that, recently, a similar model was established to describe branch migration rates for single invader mismatches.⁴⁶ This approach yielded a 4 orders of magnitude larger branch-migration step time of ~100 ms and suggested that mismatches form an unsurmountable barrier for the invader such that strand displacement only occurs by incumbent dissociation. In contrast, our model implies that mismatches are readily overcome during the displacement process in agreement with molecular dynamics simulations²⁴ and the experimentally observed impact of additional mismatches.

After parametrizing the model, we could also successfully predict the displacement kinetics for mismatches in the incumbent–target duplex that promote increased reaction rates. These mismatches are of interest since they can provide a free-energetic bias in reaction networks without increasing the toehold length and thus keep leak reaction pathways suppressed.²³ We found that a localized mismatch penalty placed at the mismatch position could not fully describe the experimental data. Rather the fraying of the incumbent–target base pairs had to be considered, such that the impact of the mismatch penalty extended by up to four base pairs toward the toehold. Our improved free-energy landscape (Figure 4C) reproduced the result of a previous coarse-grained simulation study²³ and provided a better description of the data. The strongest rate increase is observed for toehold-proximal mismatches (Figure 4D). In this case, fraying dominates the

base pairing right from the beginning of the displacement process, leading to an effectively elongated toehold. This effect may again open the way for leak reactions such that the usage of toehold proximal mismatches may be problematic for such applications. We further note that a rate increase is not obtained for toeholds longer than four nucleotides where the probability for incumbent displacement is already high.

Our parametrized model could also describe the large rate reductions observed for invaders with double mismatches. For proximal double mismatches, loop formation with unpaired bases between the mismatches had to be considered, which is equivalent to the base pair fraying for incumbent mismatches. We restricted the experiments of double mismatches to long toeholds, where in agreement with the model, the mean displacement time was increased from seconds to days. For the given mismatch strength, the mean-displacement times for short toeholds would be increased from days to years, which is experimentally not feasible. Importantly, this suggests that in reaction networks a high specificity can be obtained for targets that differ with respect to each other by just two base pairs. According to our model, the double-mismatch kinetics for long toeholds is not limited by the invader binding but rather by the intrinsic branch migration across the mismatches and can thus not be influenced by the invader concentration.

In summary, the established modeling of strand displacement reactions in the presence of invader and incumbent mismatches provides a simple and convenient tool to estimate strand displacement rates instead of using elaborate experimental characterization or computationally demanding simulations. Our model can be applied to model displacement reaction networks but also to predict the specificities of invaders in such systems. We think that with correspondingly chosen free-energetic penalties also other structural alterations, such as loops⁴⁷ or junctions⁴⁸ in invader– or incumbent–target duplexes can be described by the model. Beyond strand displacement systems based on purified nucleic acids, our modeling will also provide a useful reference for more complex systems such as CRISPR-Cas effector complexes. Developing rational, mechanism-based target predictors²⁹ may help to avoid off targeting by these enzymes in biotechnological and medical applications.

METHODS

DNA Oligonucleotides. In order to avoid strong sequence-dependent effects, we followed some simple design principles for the strand displacement systems. The free-energetic properties of the employed strands were chosen to be similar. Thus, the toehold sequence was chosen to be equal for all experiments and to possess a hybridization free-energy that decreased nearly linearly with the number of base pairs. The target-incumbent duplex was designed similarly using the same criterion. Mismatches were introduced between the same nearest neighbor base pairs. In particular, we chose the symmetric trinucleotide sequence 5'-AGA in the incumbent strand to introduce C–C mismatches. This allowed us to place mismatches at a distance of 2 nt using the sequence 5'-AGAGA. Furthermore, the chosen mismatch was the least stable symmetric mismatch.⁴⁹ To avoid unintended secondary structure formation that perturbs the reaction kinetics,⁵⁰ we selected sequences where the formation of alternative dimer structures between any of the three strands, and the formation of intramolecular and hairpin stems, was limited to 5 and 3 bp, respectively. The sequence design was carried out using a python code that evaluated randomly generated sequences based on these criteria.

All oligonucleotide sequences used are shown in Supplementary Tables 1–4. In order to test all desired invader mismatch positions for

the 17 bp displacement domain (including double mismatches), five different incumbent–target systems with 17 bp long displacement domains were used. For testing the desired incumbent mismatches, three invader–target systems with 17 bp displacement domains were employed. To reveal the impact of incumbent fraying for toehold distal mismatches, we additionally employed two systems with 22nt long displacement domains. For measurements of double-invader mismatches using a 27 bp displacement domain, two additional systems were designed.

Unlabeled, HPLC-purified DNA oligonucleotides were purchased from Sigma-Aldrich, and labeled HPLC-purified oligonucleotides (3'-BHQ1 or 5'-6FAM) were acquired from IBA. Oligonucleotide concentrations were evaluated from the measured absorbance at 260 nm⁵¹ using a P-330 NanoPhotometer (Implen). Incumbent–target duplexes were hybridized using 1 μ M of the individual strands in buffer containing 10 mM Tris–HCl (pH of 8.0), 50 mM NaCl, and 1 mM EDTA and slow cooling from 95 to 4 °C at 1 K/min.

Fluorescence Measurements. All measurements of the reaction kinetics were performed in buffer containing 10 mM Tris–HCl (pH 8.0), 50 mM NaCl, and 10 mM MgCl₂ using 10 nM of all involved strands unless indicated (Figure 2). Fast reaction kinetics ($\Delta t_{\text{meas}} > 12$ h) were measured in a temperature-controlled PerkinElmer LS55 fluorimeter at 25 °C in 3500 μ L cuvettes (Hellma Analytics). The excitation was set to 495 nm, and the emission was recorded at 519 nm, both with a slit width of 7 nm. A 950 μ L portion of the incumbent–target duplex solution was always added first and allowed to reach a stable temperature for 10 min. Subsequently, the measurement was started by quickly adding 950 μ L of invader solution, which was pre-equilibrated at 25 °C. Data acquisition started 10 s after the reaction start in intervals of 1 s. Slow reaction kinetics ($\Delta t_{\text{meas}} > 12$ h) were measured in a Synergy H1 Hybrid Multi-Mode Microplate Reader at 25 °C in 24 well μ -plates (ibidi). The excitation path used a 485/20 nm filter. Emitted fluorescent light was separated using a 510 nm dichroic mirror and a 528/20 nm filter. A 750 μ L portion of incumbent–target duplex solution was added first into the wells and allowed to equilibrate for 5 min. Subsequently, 750 μ L of invader solution was quickly added and covered with 400 μ L mineral oil (Sigma-Aldrich) to prevent evaporation. Data points were acquired every 5 min. Additionally, the plate was slightly shaken every 24 h to sustain mixing.

Fluorescence Trajectory Fitting. The kinetics of the system were then modeled using rate equations according to the reaction schemes. For a low population of the three-strand intermediate (eq 9) the following set of differential equations was used:

$$\frac{d[I]}{dt} = \frac{d[QT]}{dt} = -k[I][QT] \quad (22)$$

$$\frac{d[Q]}{dt} = \frac{d[IT]}{dt} = +k[I][QT] \quad (23)$$

The rates for a high population of the three-strand intermediate (eq 11) are described by

$$\frac{d[I]}{dt} = \frac{d[QT]}{dt} = -k_1[I][QT] \quad (24)$$

$$\frac{d[IQT]}{dt} = +k_1[I][QT] - k_2[IQT] \quad (25)$$

$$\frac{d[Q]}{dt} = \frac{d[IT]}{dt} = +k_2[IQT] \quad (26)$$

Prior to fitting, the ground level was subtracted from the fluorescence traces. For conversion between product concentration $[IT]$ and fluorescent signal F , a linear scaling parameter A was used. Reaction kinetics for labeled incumbent–target duplexes were fitted using

$$F = A \cdot [IT] \quad (27)$$

and for labeled target–invader using

$$F = A \cdot ([QT]_0 - [IT]) \quad (28)$$

The statistical error of the measured mean displacement times (global reaction rates) was $\sim 25\%$ as obtained from repeat measurements (Figure S11). The given errors for relative displacement rates was thus $\sqrt{2} \times 25\% = 35\%$.

Data Analysis. All data was analyzed using python 3.6. To obtain rate constants, a set of ordinary differential equations was numerically integrated using “odeint” and then used for fitting using “curve_fit”, both of which are modules from the SciPy package. The integrator “odeint” was always used on default settings except for *rtol* and *atol*, which determine the error control performed by the solver, being set to 10^{-12} . The “curve_fit” utility, which uses nonlinear least-squares to fit a function to data, was always used on default settings. The predictions of the random walk model were also fitted to the data using “curve_fit”, by defining the model as a function that returns the inverse of the mean-first-passage time in dependence of toehold length and mismatch position. To obtain the energy parameters, a global fit, weighted by the statistical error, of the displacement rate was performed with the initial parameter guess being $\Delta G_{\text{assoc}}^{\text{ini}} = 3k_B T$, $\Delta G_{\text{p}}^{\text{ini}} = 3k_B T$, $\Delta G_{\text{BM}}^{\text{ini}} = 8k_B T$ and $\Delta G_{\text{MM}}^{\text{ini}} = 9k_B T$. The absolute time constant of the model was obtained by an unweighted global fit with the initial parameter guess $\Delta G_{\text{bm}}^{\text{ini}} = 10^4/\text{s}$. All plots were generated using the Matplotlib library.

oxDNA Simulations. Simulations were performed with the standalone oxDNA simulation package, which is available for download from <https://dna.physics.ox.ac.uk/>. For simplicity, simulations were all performed using an augmented version of the sequence-dependent parametrization of the original oxDNA model,²⁷ which neglects major/minor grooving and is optimized to fit experimental data at high monovalent salt concentration. All simulations were performed using the VMMC algorithm⁵² at 25 °C, with umbrella sampling⁵³ to improve convergence. Detailed simulation inputs, enabling the simulations to be rerun, along with raw and processed output data, are available to download from DOI: [10.5281/zenodo.3655218](https://doi.org/10.5281/zenodo.3655218).

oxDNA is known to slightly underestimate the destabilizing effect of base-pair mismatches.^{24,25} Moreover, mismatches in oxDNA provide little disruption to the duplex structure. To model the experimental system, which has a particularly strong purine–purine mismatch, we therefore added a repulsion between C–C nucleotides. This repulsion was implemented by assigning a negative $\epsilon = -5.00$ hydrogen-bonding strength (in simulation units) between pairs of C–C nucleotides. The strength was tuned to approximately reproduce the $\Delta G_{\text{MM}} = (9.5 \pm 0.2)k_B T$ obtained from the fitting to the single-mismatch system. The mismatch strength was verified by simulating a 15 bp duplex in three settings: perfectly matched; with a mismatch but no additional C–C repulsion; and with a mismatch and with additional C–C repulsion. In all cases, five independent simulations were run for approximately 5×10^9 total steps after an initialization period of 9×10^8 steps. Umbrella sampling was performed as a function of the total number of base pairs, with the complete separation of the strands forbidden. The total free energy of all configurations with more than a single base pair was compared to the free energy of the state with a single base pair as an estimate of the destabilizing effect of the mismatch. Relative to the perfectly matched duplex, the original oxDNA model exhibited a $\Delta\Delta G = 7.33 \pm 0.44 k_B T$, and the augmented model $\Delta\Delta G = 9.31 \pm 0.17 k_B T$, with the uncertainty given as standard errors on the mean.

Free-energy profiles for displacement with two invader–target mismatches were performed for sequences corresponding to the experiments with mismatches at positions 2, 4 and 2, 8. Umbrella sampling was performed as a function of the number of invader–target and incumbent–target base pairs, with an additional constraint that the first base pair of the toehold and the final base pair of the incumbent–target duplex were prohibited from breaking. This strategy improved sampling by preventing the formation of off-pathway traps. For each system, 10 independent simulations were run for approximately 6.5×10^{10} VMMC steps each, following 7.1×10^7 initialization steps. The free energy profile reported is obtained by

combining the sample frequency in all simulations; the error is obtained as the standard error on the mean across profiles inferred from individual simulations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.0c03105>.

Derivation of the model, application to an independent data set, control measurements, model tests, error estimation and used sequences (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Ralf Seidel – Peter Debye Institute for Soft Matter Physics, Universität Leipzig, 04103 Leipzig, Germany; orcid.org/0000-0002-6642-053X; Email: ralf.seidel@uni-leipzig.de

Authors

Patrick Irmisch – Peter Debye Institute for Soft Matter Physics, Universität Leipzig, 04103 Leipzig, Germany

Thomas E. Ouldrige – Imperial College Centre for Synthetic Biology and Department of Bioengineering, Imperial College London, London SW7 2AZ, United Kingdom

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/jacs.0c03105>

Notes

The authors declare no competing financial interest.

Software code allowing for a given displacement system, the calculation of the strand displacement rate, and the rate constants for the three-rate approximation as well as the code for data fitting and plotting is provided at <https://github.com/pirmisch/SDmismatches>. oxDNA simulation inputs and output data are available at DOI: [10.5281/zenodo.3655218](https://doi.org/10.5281/zenodo.3655218).

■ ACKNOWLEDGMENTS

We acknowledge Petr Šulc and Flavio Romano for providing modified oxDNA code allowing a repulsion between non-complementary nucleotides. We thank Philine Hietschold and Mareike Zink for granting us access to their microplate reader. This work was supported by a consolidator grant of the European Research Council (724863) to R.S. T.O. is supported by a Royal Society University Research Fellowship.

■ REFERENCES

- (1) Seeman, N. C.; Sleiman, H. F. DNA nanotechnology. *Nature Reviews Materials* **2018**, *3*, 17068.
- (2) Kim, J.; Jang, D.; Park, H.; Jung, S.; Kim, D. H.; Kim, W. J. Functional-DNA-Driven Dynamic Nanoconstructs for Biomolecule Capture and Drug Delivery. *Adv. Mater.* **2018**, *30*, 1707351.
- (3) Ijaes, H.; Nummelin, S.; Shen, B.; Kostainen, M.; Linko, V. Dynamic DNA Origami Devices: from Strand-Displacement Reactions to External-Stimuli Responsive Systems. *Int. J. Mol. Sci.* **2018**, *19*, 2114.
- (4) Guo, Y.; Wei, B.; Xiao, S.; Yao, D.; Li, H.; Xu, H.; Song, T.; Li, X.; Liang, H. Recent advances in molecular machines based on toehold-mediated strand displacement reaction. *Quant. Biol.* **2017**, *5*, 25–41.
- (5) Yurke, B.; Turberfield, A. J.; Mills, A. P.; Simmel, F. C.; Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **2000**, *406*, 605–608.

- (6) Venkataraman, S.; Dirks, R. M.; Rothmund, P. W. K.; Winfree, E.; Pierce, N. A. An autonomous polymerization motor powered by DNA hybridization. *Nat. Nanotechnol.* **2007**, *2*, 490–494.
- (7) Shin, J.-S.; Pierce, N. A. A Synthetic DNA Walker for Molecular Transport. *J. Am. Chem. Soc.* **2004**, *126*, 10834–10835.
- (8) Lund, K.; Manzo, A. J.; Dabby, N.; Michelotti, N.; Johnson-Buck, A.; Nangreave, J.; Taylor, S.; Pei, R.; Stojanovic, M. N.; Walter, N. G.; Winfree, E.; Yan, H. Molecular robots guided by prescriptive landscapes. *Nature* **2010**, *465*, 206.
- (9) Wickham, S. F. J.; Bath, J.; Katsuda, Y.; Endo, M.; Hidaka, K.; Sugiyama, H.; Turberfield, A. J. A DNA-based molecular motor that can navigate a network of tracks. *Nat. Nanotechnol.* **2012**, *7*, 169.
- (10) Sherman, W. B.; Seeman, N. C. A Precisely Controlled DNA Biped Walking Device. *Nano Lett.* **2004**, *4*, 1203–1207.
- (11) Thubagere, A. J.; Li, W.; Johnson, R. F.; Chen, Z.; Doroudi, S.; Lee, Y. L.; Izatt, G.; Wittman, S.; Srinivas, N.; Woods, D.; Winfree, E.; Qian, L. A cargo-sorting DNA robot. *Science* **2017**, *357*, No. eaan6558.
- (12) Lakin, M. R.; Youssef, S.; Cardelli, L.; Phillips, A. Abstractions for DNA circuit design. *J. R. Soc., Interface* **2012**, *9*, 470–486.
- (13) Chen, Y.-J.; Dalchau, N.; Srinivas, N.; Phillips, A.; Cardelli, L.; Soloveichik, D.; Seelig, G. Programmable chemical controllers made from DNA. *Nat. Nanotechnol.* **2013**, *8*, 755–762.
- (14) Soloveichik, D.; Seelig, G.; Winfree, E. DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5393–5398.
- (15) Meijer, L. H. H.; Joesaar, A.; Steur, E.; Engelen, W.; van Santen, R. A.; Merck, M.; de Greef, T. F. A. Hierarchical control of enzymatic actuators using DNA-based switchable memories. *Nat. Commun.* **2017**, *8*, 1117.
- (16) Qian, L.; Winfree, E. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **2011**, *332*, 1196–1201.
- (17) Qian, L.; Winfree, E.; Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **2011**, *475*, 368–372.
- (18) Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E. Enzyme-Free Nucleic Acid Logic Circuits. *Science* **2006**, *314*, 1585–1588.
- (19) Genot, A. J.; Bath, J.; Turberfield, A. J. Combinatorial Displacement of DNA Strands: Application to Matrix Multiplication and Weighted Sums. *Angew. Chem., Int. Ed.* **2013**, *52*, 1189–1192.
- (20) Wang, S.; Yang, F.; Jin, D.; Dai, Q.; Tu, J.; Liu, Y.; Ning, Y.; Zhang, G.-J. Toehold Mediated One-Step Conformation-Switchable “Signal-On” Electrochemical DNA Sensing Enhanced with Homogeneous Enzymatic Amplification. *Anal. Chem.* **2017**, *89*, 5349–5356.
- (21) Zhang, Z.; Hejesen, C.; Kjelstrup, M. B.; Birkedal, V.; Gothelf, K. V. A DNA-Mediated Homogeneous Binding Assay for Proteins and Small Molecules. *J. Am. Chem. Soc.* **2014**, *136*, 11115–11120.
- (22) Hu, P.; Li, M.; Wei, X.; Yang, B.; Li, Y.; Li, C.-Y.; Du, J. Cooperative Toehold: A Mechanism To Activate DNA Strand Displacement and Construct Biosensors. *Anal. Chem.* **2018**, *90*, 9751–9760.
- (23) Haley, N. E. C.; Ouldrige, T. E.; Mullor Ruiz, I.; Geraldini, A.; Louis, A. A.; Bath, J.; Turberfield, A. J. Design of hidden thermodynamic driving for non-equilibrium systems via mismatch elimination during DNA strand displacement. *Nat. Commun.* **2020**, *11*, 2562.
- (24) Machinek, R. R. F.; Ouldrige, T. E.; Haley, N. E. C.; Bath, J.; Turberfield, A. J. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nat. Commun.* **2014**, *5*, 5324.
- (25) Jiang, Y. S.; Bhadra, S.; Li, B.; Ellington, A. D. Mismatches Improve the Performance of Strand-Displacement Nucleic Acid Circuits. *Angew. Chem., Int. Ed.* **2014**, *53*, 1845–1848.
- (26) Ouldrige, T. E.; Louis, A. A.; Doye, J. P. K. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* **2011**, *134*, 085101.
- (27) Sulc, P.; Romano, F.; Ouldrige, T. E.; Rovigatti, L.; Doye, J. P. K.; Louis, A. A. Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.* **2012**, *137*, 135101.
- (28) Srinivas, N.; Ouldrige, T. E.; Sulc, P.; Schaeffer, J. M.; Yurke, B.; Louis, A. A.; Doye, J. P. K.; Winfree, E. On the biophysics and

kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.* **2013**, *41*, 10641–10658.

(29) Klein, M.; Eslami-Mossallam, B.; Arroyo, D. G.; Depken, M. Hybridization Kinetics Explains CRISPR-Cas Off-Targeting Rules. *Cell Rep.* **2018**, *22*, 1413–1423.

(30) Szczelkun, M. D.; Tikhomirova, M. S.; Sinkunas, T.; Gasiunas, G.; Karvelis, T.; Pschera, P.; Siksnys, V.; Seidel, R. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 9798–9803.

(31) Rutkauskas, M.; Sinkunas, T.; Songailiene, I.; Tikhomirova, M. S.; Siksnys, V.; Seidel, R. Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. *Cell Rep.* **2015**, *10*, 1534–1543.

(32) Sternberg, S. H.; Redding, S.; Jinek, M.; Greene, E. C.; Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **2014**, *507*, 62.

(33) Bhattacharya, P. K.; Cha, J.; Barton, J. K. ¹H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches. *Nucleic Acids Res.* **2002**, *30*, 4740–4750.

(34) Zhang, D. Y.; Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **2009**, *131*, 17303–17314.

(35) SantaLucia, J.; Hicks, D. The Thermodynamics of DNA Structural Motifs. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440. PMID: 15139820.

(36) Markham, N. R.; Zuker, M. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* **2005**, *33*, W577–W581.

(37) Tran, T.; Cannon, B. Differential Effects of Strand Asymmetry on the Energetics and Structural Flexibility of DNA Internal Loops. *Biochemistry* **2017**, *56*, 6448–6459.

(38) Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A. NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.* **2011**, *32*, 170–173.

(39) Simmel, F. C.; Yurke, B.; Singh, H. R. Principles and Applications of Nucleic Acid Strand Displacement Reactions. *Chem. Rev.* **2019**, *119*, 6326–6369.

(40) <https://github.com/pirmisch/SDmismatches>.

(41) Green, C.; Tibbetts, C. Reassociation rate limited displacement of DNA strands by branch migration. *Nucleic Acids Res.* **1981**, *9*, 1905–1918.

(42) Radding, C. M.; Beattie, K. L.; Holloman, W. K.; Wiegand, R. C. Uptake of homologous single-stranded fragments by superhelical DNA. *J. Mol. Biol.* **1977**, *116*, 825–839.

(43) Zolaktaf, S.; Dannenberg, F.; Rudelis, X.; Condon, A.; Schaeffer, J. M.; Schmidt, M.; Thachuk, C.; Winfree, E. *Lecture Notes in Computer Science*; Springer International Publishing, 2017; pp 172–187.

(44) Panyutin, I. G.; Hsieh, P. The kinetics of spontaneous DNA branch migration. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 2021–2025.

(45) Karymov, M. A.; Bogdanov, A.; Lyubchenko, Y. L. Single Molecule Fluorescence Analysis of Branch Migration of Holliday Junctions: Effect of DNA Sequence. *Biophys. J.* **2008**, *95*, 1239–1247.

(46) Broadwater, D.; Kim, H. The Effect of Basepair Mismatch on DNA Strand Displacement. *Biophys. J.* **2016**, *110*, 1476–1484.

(47) Genot, A. J.; Zhang, D. Y.; Bath, J.; Turberfield, A. J. Remote Toehold: A Mechanism for Flexible Control of DNA Hybridization Kinetics. *J. Am. Chem. Soc.* **2011**, *133*, 2177–2182.

(48) Chen, X. Expanding the rule set of DNA circuitry with associative toehold activation. *J. Am. Chem. Soc.* **2012**, *134*, 263–271.

(49) Peyret, N.; Seneviratne, P. A.; Allawi, H. T.; SantaLucia, J. Nearest-Neighbor Thermodynamics and NMR of DNA Sequences with Internal AA, CC, GG, and TT Mismatches. *Biochemistry* **1999**, *38*, 3468–3477.

(50) Gao, Y. Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. *Nucleic Acids Res.* **2006**, *34*, 3370–3377.

(51) Mulhardt, C. *Molecular Biology and Genomics*; Elsevier Science, 2010.

(52) Whitelam, S.; Feng, E. H.; Hagan, M. F.; Geissler, P. L. The role of collective motion in examples of coarsening and self-assembly. *Soft Matter* **2009**, *5*, 1251–1262.

(53) Torrie, G.; Valleau, J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.