Biophysical Journal
Article

Biophysical Society

First passage time study of DNA strand displacement

D. W. Bo Broadwater, Jr.,[1] Alexander W. Cook,[1] and Harold D. Kim[1,*]
[1]School of Physics, Georgia Institute of Technology, Atlanta, Georgia

ABSTRACT DNA strand displacement, in which a single-stranded nucleic acid invades a DNA duplex, is pervasive in genomic processes and DNA engineering applications. The kinetics of strand displacement have been studied in bulk; however, the kinetics of the underlying strand exchange were obfuscated by a slow bimolecular association step. Here, we use a novel single-molecule fluorescence resonance energy transfer approach termed the "fission" assay to obtain the full distribution of first passage times of unimolecular strand displacement. At a frame time of 4.4 ms, the first passage time distribution for a 14-nucleotide displacement domain exhibited a nearly monotonic decay with little delay. Among the eight different sequences we tested, the mean displacement time was on average 35 ms and varied by up to a factor of 13. The measured displacement kinetics also varied between complementary invaders and between RNA and DNA invaders of the same base sequence, except for T → U substitution. However, displacement times were largely insensitive to the monovalent salt concentration in the range of 0.25–1 M. Using a one-dimensional random walk model, we infer that the single-step displacement time is in the range of ∼30–300 μs, depending on the base identity. The framework presented here is broadly applicable to the kinetic analysis of multistep processes investigated at the single-molecule level.

SIGNIFICANCE DNA strand displacement occurs when a single nucleic acid strand invades and replaces another nearly identical strand in a duplex. This process is ubiquitous in biology and is fundamental to the field of DNA nanotechnology. Previous kinetic studies of strand displacement either used DNA strands much longer than those found in practical applications or were obscured by a rate-limiting bimolecular step known as toehold formation. In this study, we introduce a new, single-molecule scheme that enables direct measurement of the strand displacement first passage time. Our observed kinetics demonstrate highly nontrivial sequence dependence as well as surprising differences between RNA and DNA invaders.

INTRODUCTION

Nucleic acids' ability to form hydrogen bonds between complementary Watson-Crick bases allows for a rich set of complicated, multistep kinetic behaviors such as duplex hybridization (1) and dehybridization (2), Holliday junction structural dynamics (3,4), and strand invasion (5). In particular, strand displacement, which is the exchange of bases between two competing nucleic acid strands of identical sequence, occurs in homologous recombination (6,7), DNA replication (8), and RNA transcription (9), as well as CRISPR/Cas (10) and the related Cascade complex (11). In addition to fundamental genomic processes, DNA nanotechnology exploits strand displacement to create nanoscale gadgets (12–14) and computational circuits (15–18). Strand displacement also aids in the development of quantitative assays for detection of nucleic acid (19–21) and enzymatic activity (22,23) with improved probe specificity (17,24,25).

For practical applications, strand displacement is implemented with the "invader" strand and a partial duplex composed of the "incumbent" strand and the "substrate" strand (Fig. 1; (18)). The partial duplex has two distinct domains: 1) the single-stranded overhang called the toehold, which is critical to the speed and efficiency of the reaction (26), and 2) the duplex region called the displacement domain. Toehold-mediated strand displacement is initiated when the invader strand anneals to the toehold in a bimolecular reaction. Once a stable toehold interaction is formed,

Submitted May 28, 2020, and accepted for publication January 12, 2021.
*Correspondence: harold.kim@physics.gatech.edu
D. W. B. Broadwater and Alexander W. Cook contributed equally to this work.
Editor: Jason Kahn.
https://doi.org/10.1016/j.bpj.2021.01.043

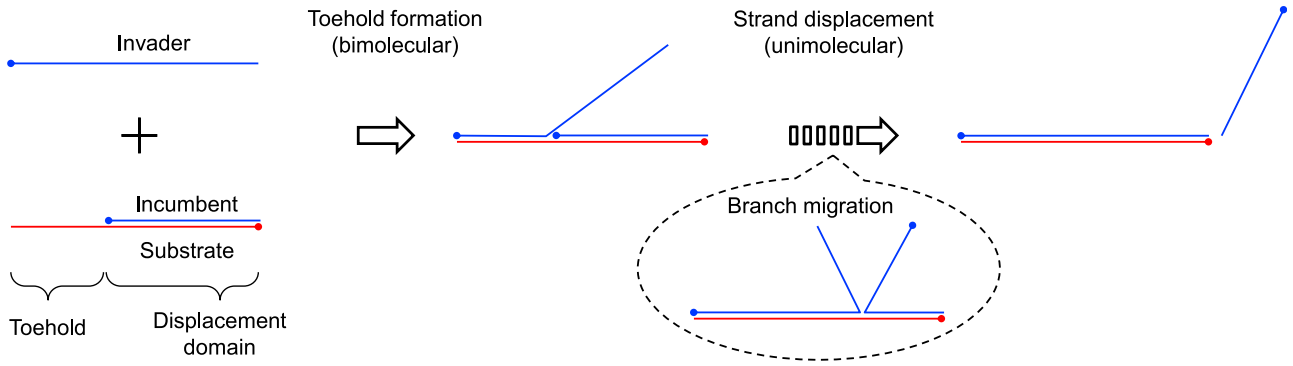


FIGURE 1 Toehold-mediated DNA strand displacement. Each line represents a single-stranded DNA with the 5′ end marked by a dot. Blue and red colors represent complementary sequences. In bulk studies, the apparent reaction kinetics are dominated by the slow bimolecular step. In this study, we focus on the unimolecular strand displacement, which is thought to occur through branch migration. To see this figure in color, go online.

the incumbent can be displaced by the dangling strand of the invader through spontaneous opening of a basepair between substrate and incumbent and closing of a basepair between substrate and invader. This unimolecular strand displacement is also called branch migration (27–29).

Great attention has been paid to the kinetics of toehold-mediated strand displacement (26,29–32). However, these kinetics have been measured mostly in bulk, in which the reaction kinetics are limited by bimolecular toehold association. Therefore, the measured kinetics do not shed light on the unimolecular branch migration. Other studies using long (∼1 kbp) DNA estimated the branch migration time per basepair step to be ∼10 μs (27,28), but strand displacement in those studies took place in a D-loop geometry that restricted the movement of dangling strands during branch migration, unlike the geometry of current interest. Also, ∼1 kbp scale extends far beyond the length scales of interest for DNA nanotechnology. To understand whether and how sequence can be used to control displacement kinetics, we require experimental studies on unimolecular displacement of short oligos that can be modeled at the single base level.

In this study, we introduce a DNA "fission" assay to study toehold-mediated DNA strand displacement kinetics. The fission assay employs single-molecule fluorescence resonance energy transfer (smFRET) to directly measure the first passage displacement time (33,34) for the unimolecular reaction that occurs between toehold formation and incumbent dissociation. Using a wide-field total internal reflection fluorescence microscope, we measured the displacement kinetics for a 14-nucleotide displacement domain of eight different sequences. The mean displacement time varied by more than 10-fold between the slowest and fastest sequence and was on average ∼ 35 ms, and the histograms of displacement times obtained at 4.4-ms resolution showed a monotonic decay with little to no lag. We found that the displacement kinetics depend on the base sequence and the nucleic acid type (DNA versus RNA) of the invader but not on monovalent salt concentration. We analyzed the first passage time histograms of strand displacement using a symmetric random walk model to extract single basepair step times. The best fit to the histograms was obtained with ∼33, ∼200, ∼250, and ≲33 μs for A, C, T, and G, respectively. Our study reports the displacement rates of short DNA oligos and reveals biophysical mechanisms that govern DNA strand displacement kinetics.

MATERIALS AND METHODS

Sample preparation

Custom DNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA). The 26-nt substrate was internally labeled near the end distal to the toehold with a Cy3 fluorophore. The 24-nt invader molecule was labeled with a Biotin-TEG linker at the end proximal to the toehold for surface immobilization. The 14-nt incumbent sequences were labeled with a Cy5 fluorophore at the end distal to the toehold. All oligos were purified by the manufacturer with high-performance liquid chromatography. The specific sequences are in Tables S1–S3. Partial duplexes were constructed by combining substrate and incumbent at a 1:10 ratio (0.5 μM substrate, 5.0 μM incumbent) in buffer at pH 7 containing 1 M NaCl, 10 mM Tris, and 1 mM EDTA. The excess of incumbent strands was meant to minimize the number of single-stranded substrates in solution; unpaired substrates can compete with the partial duplexes for binding with the surface-bound invaders, whereas lone incumbent strands do not bind to the invaders and will not fluoresce on their own. The mixture was heated to 95°C and slowly cooled for 3 h to 4°C to ensure the partial duplex was fully annealed.

Experimental setup

Molecules were observed with an objective-type total internal reflection fluorescence microscope assembled on a commercial microscope body (IX81; Olympus, Tokyo, Japan). Fluorophores were excited by a 532 nm laser (BWN-532-50E; B&W Tek, Newark, DE). Images were 2 × 2 binned and captured with an EMCCD (DU-897ECS0-#BV; Andor Technology, Belfast, UK), and images were recorded at 228 fps with 3.96-ms exposure time using Micro-Manager software (35). This high frame rate was achieved by cropping the image height to 64 superpixels. Experiments were performed on flow cells constructed as previously described in Le and Kim (36), whereas flow volume and flow rate (900 μL min^{-1}) were controlled by a syringe pump (NE-1000; New Era Pump Systems, East Farmingdale, NY).

The surface was passivated with polyethylene glycol to minimize nonspecific binding (36). After NeutrAvidin coating, the biotin-containing invader molecules were immobilized by flowing in at a concentration of 1 nM. Next, 20 μL of partial duplexes were pumped into the flow cell at 2.5 nM in an oxygen-scavenging imaging buffer (37), which contained 1 nM 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox), 5 mM protocatechuic acid, 100 nM protocatechuate 3,4-dioxygenase, and 100 mM Tris-HCl (pH 7).

An appearance of a high-FRET signal marked the formation of the toehold. A low-FRET signal appeared as strand displacement concluded. The FRET signal time series was recorded and analyzed using in-house MATLAB software (The MathWorks, Natick, MA). The lifetime of the high-FRET state was observed for many molecules to collect a distribution of displacement times.

Statistics of displacement times

Here, we provide an analytical expression we used to fit the histograms of displacement times. We model strand displacement as a one-dimensional random walk:

$$0 \underset{r_1}{\overset{f_0}{\rightleftarrows}} 1 \underset{r_2}{\overset{f_1}{\rightleftarrows}} 2 \underset{r_3}{\overset{f_2}{\rightleftarrows}} \dots \underset{r_n}{\overset{f_{n-1}}{\rightleftarrows}} n. \tag{1}$$

In this model, each state is denoted by i, the number of displaced bases, and the measured displacement time corresponds to the first passage time from the reflecting state 0 on the left boundary to the absorbing state n on the right boundary. The forward rate and reverse rate from state i are denoted as f_i and r_i, respectively. Because state n is the absorbing state, $r_n = 0$. The time dependence of the system is governed by the master equation:

$$\frac{\partial|\psi(t)\rangle}{\partial t} = \mathbf{L}|\psi(t)\rangle, \tag{2}$$

where $\mathbf{L}$ is the transition matrix operator, and the ket vector $|\psi(t)\rangle$ represents the system state. The probability amplitude to be in the absorbing state $|n\rangle$ at some time t after starting in $|0\rangle$ is then given by (38)

$$P(n,t\,|\,0) = \langle n\,|\,e^{\mathbf{L}t}|0\rangle. \tag{3}$$

The experimentally accessible datapoints in single-molecule experiments are the number of displacement events ΔN detected during a short time interval or bin time Δt. These numbers form the so-called dwell time or survival-time histogram. For a sufficiently large number of total events N_0, ΔN in the i-th bin is related to the probability amplitudes according to

$$\begin{aligned}\frac{\Delta N(i)}{N_0} &= P(n,i\Delta t|0) - P(n,(i-1)\Delta t\,|\,0)\\ &= \left\langle n\,\middle|\,e^{\mathbf{L}i\Delta t}\left(1-e^{-\mathbf{L}\Delta t}\right)\,\middle|\,0\right\rangle.\end{aligned} \tag{4}$$

This can be expanded using the left and right eigenvectors of $\mathbf{L}$, $\langle\varphi_k^L|$ and $|\phi_k^R\rangle$ that satisfy $\langle\phi_k^L\,|\,\phi_k^R\rangle = \delta_{mn}$, and their corresponding eigenvalue $-\mu_k$:

$$\frac{\Delta N(i)}{N_0} = \sum_k \left\langle n\middle|\phi_k^R\right\rangle\left\langle\phi_k^L\,\middle|\,0\right\rangle e^{-\mu_k i\Delta t}\left(1-e^{\mu_k\Delta t}\right). \tag{5}$$

In the limit of $\Delta t \to 0$, Eq. 5 yields the first passage time density $f(t)$:

$$\begin{aligned}f(t) \equiv \lim_{\Delta t\to 0}\tfrac{1}{N_0}\tfrac{\Delta N(i)}{\Delta t} &= P(n,t\,|\,0)\\ &= \sum_k \left\langle n\,\middle|\,\phi_k^R\right\rangle\left\langle\phi_k^L\,\middle|\,0\right\rangle(-\mu_k)e^{-\mu_k t}.\end{aligned} \tag{6}$$

We used Eq. 5 to fit the measured histograms of displacement times with a fixed Δt that corresponds to the frame time of 4.4 ms. In the representation of $|0\rangle \to (1, 0, 0, \dots, 0)^{\mathrm{T}}$, $\mathbf{L}$ is an asymmetric tridiagonal matrix:

$$\mathbf{L} = \begin{bmatrix} -f_0 & r_1 & 0 & \cdots & 0 & 0\\ f_0 & -(f_1+r_1) & r_2 & \cdots & 0 & 0\\ 0 & f_1 & -(f_2+r_2) & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & -(f_{n-1}+r_{n-1}) & 0\\ 0 & 0 & 0 & \cdots & f_{n-1} & 0 \end{bmatrix}, \tag{7}$$

whose left and right eigenvectors and eigenvalues can be obtained using MATLAB.

We also present here the expression we use to analyze the mean first passage time τ (39). τ can be computed using Eq. 6 as

$$\tau = \int_0^\infty tf(t)dt. \tag{8}$$

A more useful expression can be obtained in terms of an invertible submatrix of $\mathbf{L}$, which we term $\mathbf{A}$:

$$\mathbf{L} = \begin{bmatrix} [\mathbf{A}] & \begin{matrix}0\\ \vdots\\ 0\end{matrix}\\ 0\cdots f_{n-1} & 0 \end{bmatrix}. \tag{9}$$

Using the normalization of probability amplitude

$$P(n,t|0) = 1 - \sum_{i=0}^{n-1} P(i,t\,|\,0), \tag{10}$$

we can express τ in terms of the inverse of $\mathbf{A}$

$$\begin{aligned}\int_0^\infty tf(t)dt &= -\int_0^\infty td\left[\sum_{i=0}^{n-1}\left\langle i\,\middle|\,e^{-\mathbf{A}t}\,\middle|\,0\right\rangle\right]\\ &= \int_0^\infty \sum_{i=0}^{n-1}\left\langle i\,\middle|\,e^{-\mathbf{A}t}\,\middle|\,0\right\rangle dt\end{aligned}$$

$$= \sum_{i=0}^{n-1} \langle i \,|\, \mathbf{A}^{-1} \,|\, 0 \rangle. \tag{11}$$

In the matrix presentation, the inverse matrix $\mathbf{A}^{-1}$ is related to matrix cofactors by

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \mathbf{C}^{\mathrm{T}}. \tag{12}$$

Plugging Eq. 12 into Eq. 11,

$$\tau = \frac{1}{\det(\mathbf{A})} \sum_{j=1}^{n} C_{1j}. \tag{13}$$

This sum of cofactors can be equated to the determinant of matrix $\mathbf{A}'$, which replaces the first row of $\mathbf{A}$ with 1. Hence, the mean first passage time is given by the ratio of two matrix determinants (32):

$$\tau = \frac{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}}, \tag{14}$$

where a_{ij} represents the matrix elements of $\mathbf{A}$. Eq. 14 can also be expressed in terms of the bias factor $\alpha_i = f_{i-1}/r_i$ as (32)

$$\tau = \frac{1}{f_0} + \frac{1+\alpha_1}{\alpha_1} \cdot \frac{1}{f_1} + \frac{1+\alpha_1+\alpha_2\alpha_1}{\alpha_2\alpha_1} \cdot \frac{1}{f_2} + \cdots. \tag{15}$$

RESULTS

To focus on the unimolecular kinetics of strand displacement, we took a surface-based smFRET approach (Fig. 2). In this approach, the invader is immobilized on the glass surface of a flow chamber, and the partial duplex between the donor (Cy3)-labeled substrate and the acceptor (Cy5)-labeled incumbent are perfused into the chamber. The toehold length (10-bp) is chosen so that toehold formation is practically irreversible throughout the experiment. Upon toehold formation, a diffraction-limited spot emerges out of the diffusive background in the Cy5 channel. Upon incumbent dissociation, the spot changes fluorescence emission from the Cy5 channel to the Cy3 channel. We termed this experimental scheme "fission" because the duplex labeled with the FRET pair is split as a result of strand displacement.

Partial duplexes were constructed by annealing Cy3-labeled substrate molecules and Cy5-labeled incumbent molecules. Invader molecules were biotinylated near the end containing the toehold sequence and immobilized onto the surface (see Fig. 2 *A*). As shown in Figs. 2 *B* and S10, high-FRET signals started to appear in the field of view after partial duplexes were flowed into the chamber.

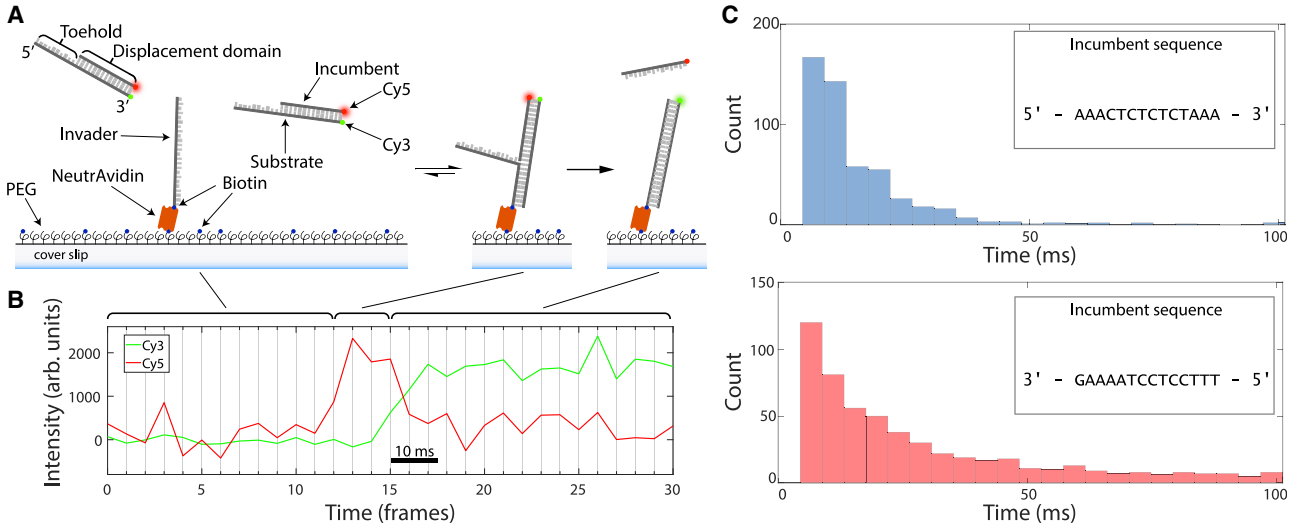


FIGURE 2 An overview of strand displacement data acquisition. (*A*) Experimental scheme of fission assay. Invader strands were immobilized on a polyethylene glycol-passivated coverslip surface via a biotin-NeutrAvidin linker. Partial duplexes were labeled with a FRET pair (Cy5, incumbent; Cy3, substrate) and were drawn into the flow cell. After a diffusive search process, the partial duplex binds to the toehold, and the incumbent strand is displaced. (*B*) Sample acceptor and donor time traces. Acceptor (Cy5) signal increases upon toehold binding, and high-FRET signal is sustained until displacement occurs, which is indicated by a low-FRET signal. The displacement time is identified as the high-FRET lifetime. (*C*) Displacement time distributions for different sequences. Displacement time distributions show sequence dependence in the displacement domain (incumbent sequence shown). Displacement lifetimes are collected for many traces and assembled into a distribution. The binning size is the single frame time (4.4 ms). To see this figure in color, go online.

The average time at which spots appeared became shorter at a higher concentration of partial duplexes, and the transition of FRET from high to low only occurred in the presence of the matching displacement domain. Without the matching displacement domain, the high-FRET spots remained until they photobleached, which confirms that dissociation of the 10-bp toehold is much slower than the typical minute-long observation period. The red signal jumped to a high level in one or two frames, which suggests that toehold formation is much faster than our time resolution, and can therefore be considered instantaneous for analysis purposes. This high-level red signal lasted for variable periods of time from trace to trace, but the eventual transition back to low-FRET always occurred in one or two frames. Simultaneously with the disappearance of the signal from the Cy5 channel, a new signal appeared in the Cy3 channel, consistent with the fission scheme (Fig. 2 *A*). Based on these observations, the first arrival of a high-FRET spot was attributed to toehold formation, and the transition from high- to low-FRET was attributed to completion of strand displacement. Hence, the dwell time in the high-FRET state (Fig. 2 *B*) represents the displacement time.

By performing the fission assay multiple times, we could record hundreds of strand displacement events for one particular displacement system and build a histogram of displacement times. To investigate the sequence dependence of strand displacement kinetics, we tested eight unique strand displacement systems, each with a different sequence in the displacement domain. We obtained these histograms at the finest bin width of 4.4 ms, two of which are shown in Fig. 2 *C*. Note that displacement events faster than the exposure time do not produce a clear signal in the acceptor channel, and therefore, the first bin of the histogram starts from 4.4 ms. For comparison of the histogram across all eight different sequences, we also present the histograms as a normalized heat map in Fig. 3. The salient feature of these histograms is that they decay monotonically with little or no delay. Six out of eight sequences show decay from the first bin; only two sequences show more events in the second than in the first bin. Nonetheless, we find a significant difference in the characteristic decay time among the tested sequences (*black dots*, Fig. 3). The fastest mean displacement time is 8 ms, whereas the slowest is 107 ms. The average over all sequences is 35 ms.

To ensure that the observed difference between different sequences is not due to the uncertainties of the histograms, we need to establish the baseline uncertainties in the empirical histograms. As explained above, each histogram is obtained by combining displacement events taken from multiple runs of the fission assay in 1 day using the same reagents and flow cell. Hence, each histogram possesses statistical uncertainty because of the finite number of events and empirical uncertainty because of fluctuations in the experimental conditions. To estimate these uncertainties, we randomly sampled 80% of the events collected on the same day and re-evaluated the mean displacement time (Fig. S11). The spread of the mean values is nonuniform among different sequences. For example, sequences 4 and 6 have a similar total mean but show different uncertainties. Nonetheless, the uncertainty in the mean for each sequence

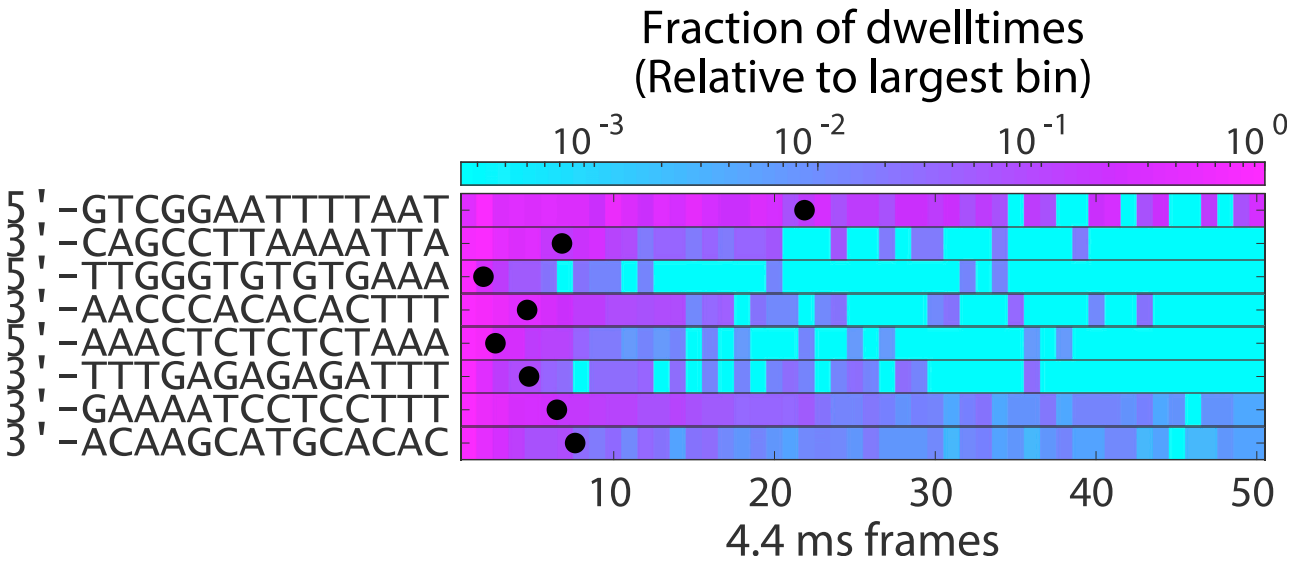


FIGURE 3 A heat map demonstrating the distribution of displacement times within the first 50 frames of a typical experiment for each of the eight featured sequences. The sequences of the incumbent strands are listed to the left. The color map is set to a logarithmic scale to better represent the exponentially decaying distributions. Here, each experiment is scaled such that the most populous bin is set to unity. The second bin is the most populous in experiments with incumbent strands 5′-GTCGGAATTTTAAT and 3′-CAGCCTTAAAATTA; in all other experiments, the first bin is the most populous. The mean of all recorded displacement times is marked by a black dot for each experiment. From top to bottom, the number of events recorded for each experiment is 273, 751, 143, 890, 522, 169, 627, and 609. To see this figure in color, go online.

is much narrower than the variation among different sequences. We also documented the variability of the histogram means obtained at different times over a 4-year span by two users (Fig. S12). This empirical variability is much higher than the statistical variability because of reasons that are not completely clear. For transparency, we present these individual mean values in Fig. 4.

In addition to the basepair sequence of the displacement domain, the base sequence of the invader can also affect the displacement kinetics. As shown in Fig. 4 *A*, the same displacement domain can be invaded using a toehold extended from either the 5′ end or the 3′ end of the displacement domain. We refer to these complementary invasions as top and bottom invasions. The mean displacement times of top and bottom invasions are clearly different for all three displacement domains we tested. No particular invasion side was consistently faster; for sequence 1, bottom invasion is faster, but for sequence 2, top invasion is. Interestingly, when RNA with an identical sequence except for T → U substitution was used as an invader in place of DNA, the faster side was switched (Fig. 5). All of these results suggest that the displacement kinetics are not completely determined by the basepair sequence or the thermodynamic stability of the displacement domain, but rather that the measured displacement kinetics are sensitive down to the chemical makeup of invading bases.

Last, we investigated the salt dependence of displacement kinetics. Monovalent salt can screen the negative charges on the phosphate backbone and alter the thermodynamics and kinetics of basepairing (40). However, its effect on the kinetics of branch migration is less clear because branch migration involves the basepairing dynamics of two competing strands. As shown in Fig. 6, the mean displacement time shows little change from 250 mM to 1 M [NaCl].

DISCUSSION

Using the fission assay, we measured the unimolecular branch migration kinetics in toehold-mediated DNA strand displacement. Using wide-field total internal reflection fluorescence microscopy and subregion readout of an electron-multiplying charge-coupled device camera, we were able to record thousands of strand displacement events at a 4.4 ms frame rate. Our fission assay begins in a dark field of view with unlabeled invader strands immobilized on the surface and monitors displacement events through the appearance and disappearance of the FRET signal on the surface. The experimental design permits us to use high-excitation intensity to detect fast displacement events at high signal/noise; strong excitation of fluorescent molecules begins only at the start of branch migration. Hence, the undesirable effect of photobleaching is eliminated.

Our fission assay produces data that could not be obtained to date. It separates out the bimolecular toehold formation step from the rest so that the apparent displacement time truly reflects a unimolecular process. In the language of

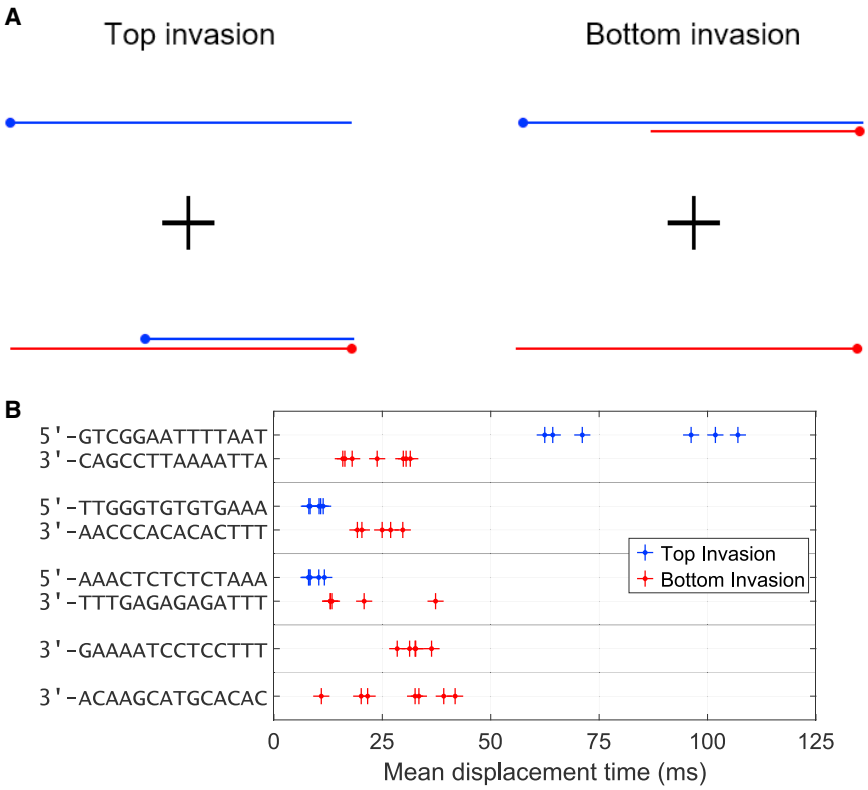


FIGURE 4 Kinetics of various displacement systems. (*A*) Schematic of invasion designs. Top invasion is defined by an invader with a 5′ toehold. In bottom invasion, all strands are replaced by their reverse complement. Bottom invasion is defined by an invader with a 3′ toehold. These systems are highly related as duplex basepairing is identical in both systems. (*B*) Mean displacement times. The mean displacement time is calculated from data below the 95th percentile. The data collection process was entirely repeated several times for each invasion system. The mean times can change by an order of magnitude depending on the sequence. The sequences are the invader nucleotides that are beyond the toehold with the nt most proximal to the toehold written first. The top-invasion sequences therefore correspond to the conventional 5′ to 3′ direction, whereas the bottom invasion sequences are listed 3′ to 5′. To see this figure in color, go online.

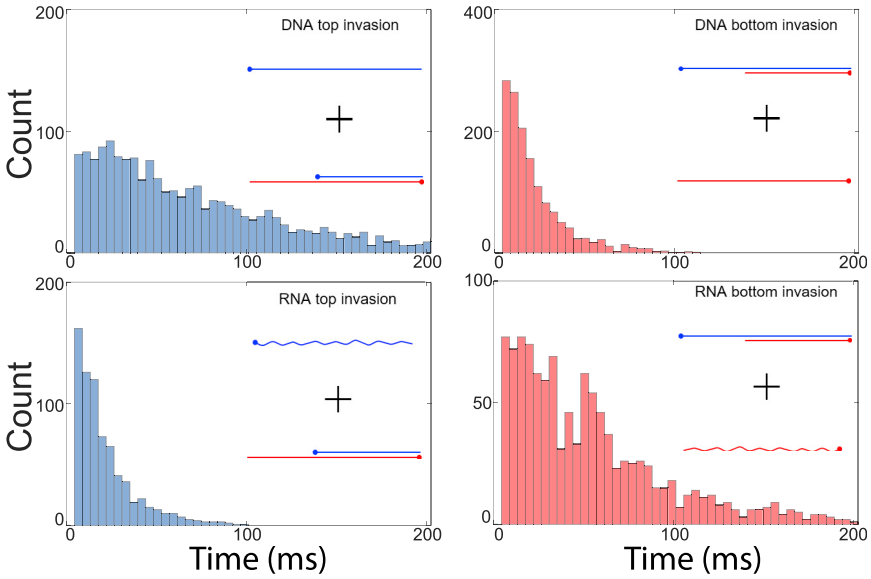


FIGURE 5 Comparison between RNA and DNA invasion. We measured displacement times for a pair of complementary DNA invaders that exhibited the largest difference between sides of invasion (*blue*, top invasion; *red*, bottom invasion) in comparison to all other pairs of invaders we observed (*top row*, straight invader). We measured displacement times for RNA versions of the invaders that were identical in sequence except for a T → U substitution (*bottom row*, wavy invader). DNA partial duplexes were employed in all cases. Again, we found a difference in passage time depending on the side of invasion. Further, we noticed that the relative times switched with respect to the side of invasion. To see this figure in color, go online.

stochastic processes, the displacement time represents the first passage time: the time taken for the branch point to start from the first position and reach the last for the first time. The fission scheme allows access to the full distribution of individual displacement times, which is more informative than just the average values. Below, we use the first passage time analysis to extract single-step migration rates from the measured histograms and discuss potential microscopic mechanisms that may control these rates.

The most elementary model to describe DNA strand displacement is a one-dimensional random walk among states defined by the number of displaced basepairs (Eq. 1). Displacement is initiated after the invader hybridizes to the toehold and continues until the incumbent loses all basepairs with the substrate to the invader. Any intermediate state during this process can be envisioned as two dangling strands branching off from the duplex stem (Fig. 7). At the junction or the branch point, an incumbent (invader) base can spontaneously break away from the substrate base, and the most adjacent invader (incumbent) base can base-pair with the substrate base. As a result, the branch point can move by one base in either direction. The branch point,

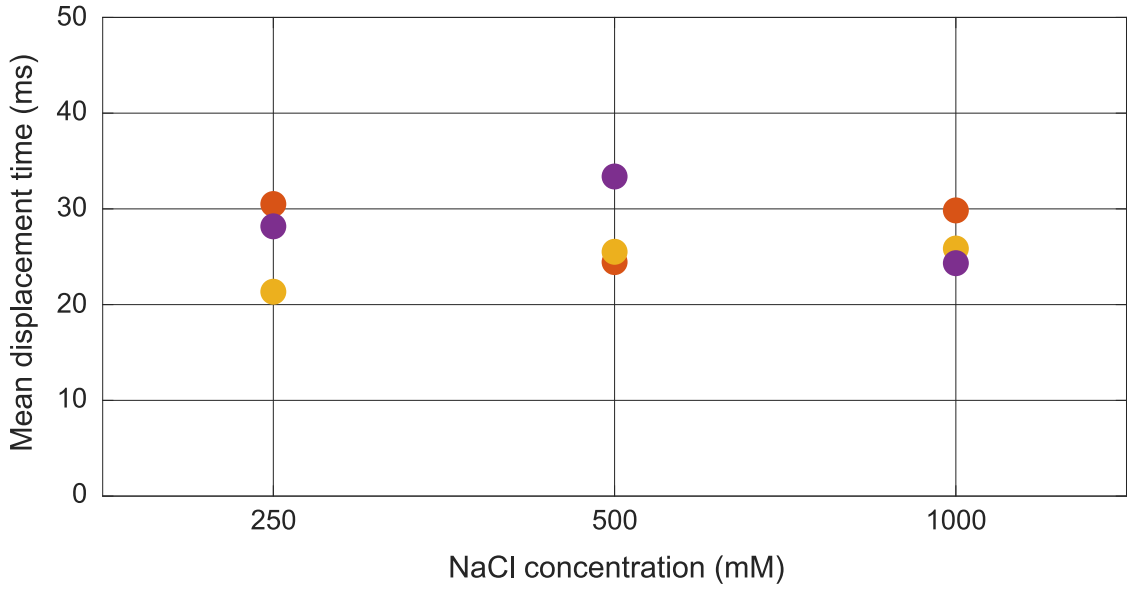


FIGURE 6 Salt dependence of mean displacement time. Data points sharing the same color were collected on the same day using the same reagents, aside from partial duplex solutions containing different concentrations of NaCl. The mean displacement time does not appear to depend strongly on the NaCl concentration. The incumbent sequence used for these measurements is 5′-ATTAAAATTCCGAC-3′. To see this figure in color, go online.

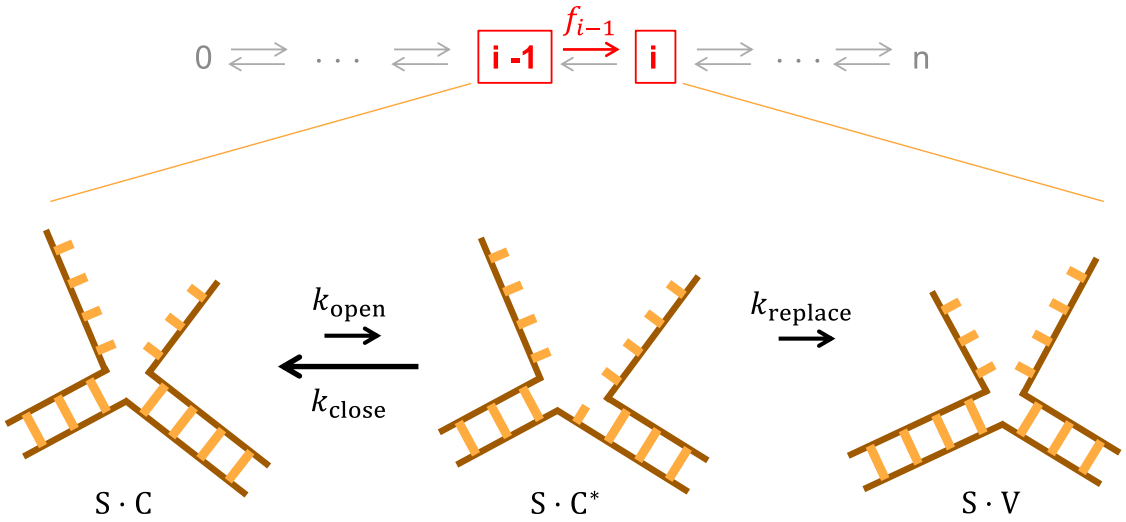


FIGURE 7 A closer look at the single basepair step transition. Any forward arrow in the one-dimensional Markov chain involves opening of a basepair between the substrate (S) and the incumbent (C) and closing of a basepair between the substrate and the invader (V). For a reverse arrow, the roles of incumbent and invading bases are simply flipped. Hence, each step can be modeled as a first passage from an initial state S · C to a final state S · V through an intermediate state S · C* in which a single basepair is transiently open. The transiently unbound incumbent base can rebind the substrate base at a rate of k_{close} or be replaced with the invading base at a rate of $k_{replace}$. To see this figure in color, go online.

however, cannot recede into the toehold region because the incumbent is shorter than the substrate. Therefore, branch migration can be modeled as a one-dimensional random walk with single base steps from a reflecting boundary on one end (state 0) to an absorbing boundary on the other (state n).

It is straightforward to derive the first passage time statistics from a Markov chain like Eq. 1. The simplest model is a uniform random walk in which all transition rates are equal ($\{f_i, r_i\} = k$). Such a model can be represented by a free energy landscape shown in Fig. 8 with troughs separated by equal height barriers. Based on Eq. 15, the mean first passage time (τ) is given by

$$\tau = \frac{1}{2k} n(n+1). \tag{16}$$

Using Eq. 16, $n = 14$, and the measured mean first passage time of 30 ms, we can estimate the single-step migration time (k^{-1}) to be ~286 μs. This estimate is also consistent with the measured histogram of displacement times. If single-step migration occurs more slowly than the time resolution, the histogram of displacement times must exhibit a strong delay or lag in the early times (Fig. S13). However, our measured histograms at 4.4 ms bin width show little or no lag, which points to a single-step migration time much shorter than 4.4 ms.

However, this estimated time of ~286 μs per step is likely to be longer than the true value because displacement events faster than the ~4-ms exposure time are not included in our measurement. To extract the single-step migration rates in a more accurate, unbiased way despite this missing fraction of events, we fit the analytical solution Eq. 5 to all eight histograms with four shared parameters representing rates for A, G, C, and T. This global fitting procedure looks for the best set of rates that describe all eight histograms in the least-squares sense, excluding the missing first bin. It also implies a nonuniform symmetric random walk (Fig. 8) in which the single-step migration rate depends only on the identity of the base to be displaced. Therefore, each step has the same forward and reverse rates ($f_{i-1} = r_i$). The extracted step times for A, C, and T bases are ~33, ~200, and ~250 μs, respectively. The step time for the G base did not converge, but the goodness of fit increased with faster values. Thus, we estimate the step time for G to be $\lesssim$33 μs. As predicted, these times obtained by fitting histograms in their entirety are all faster than ~286 μs obtained from the mean values that omit fast events.

Our estimated single-step migration rates (0.003 μs^{-1} to 0.03 μs^{-1}) appear to be much slower than the rate of basepair fraying or base flipping ($\geq 1\ \mu s^{-1}$) (41–44). Similarly, a previous study by Srivinas et al. (29) also inferred the single-step migration rate to be much slower than the fraying rate. This discrepancy suggests that a single basepair opening event does not always lead to single-step branch migration. As shown in Fig. 7, a basepair between the substrate and the incumbent can transiently open and close with rate constants of k_{open} and k_{close}, respectively. When the substrate base is transiently unbound (S · C*), the invading base can basepair with the substrate base and replace the incumbent base at a rate of $k_{replace}$. We can safely assume that k_{close} is much faster than k_{open} based on the known basepair stability (45). Coarse-grained molecular dynamics simulations (29) show that the branch migration intermediate frequently

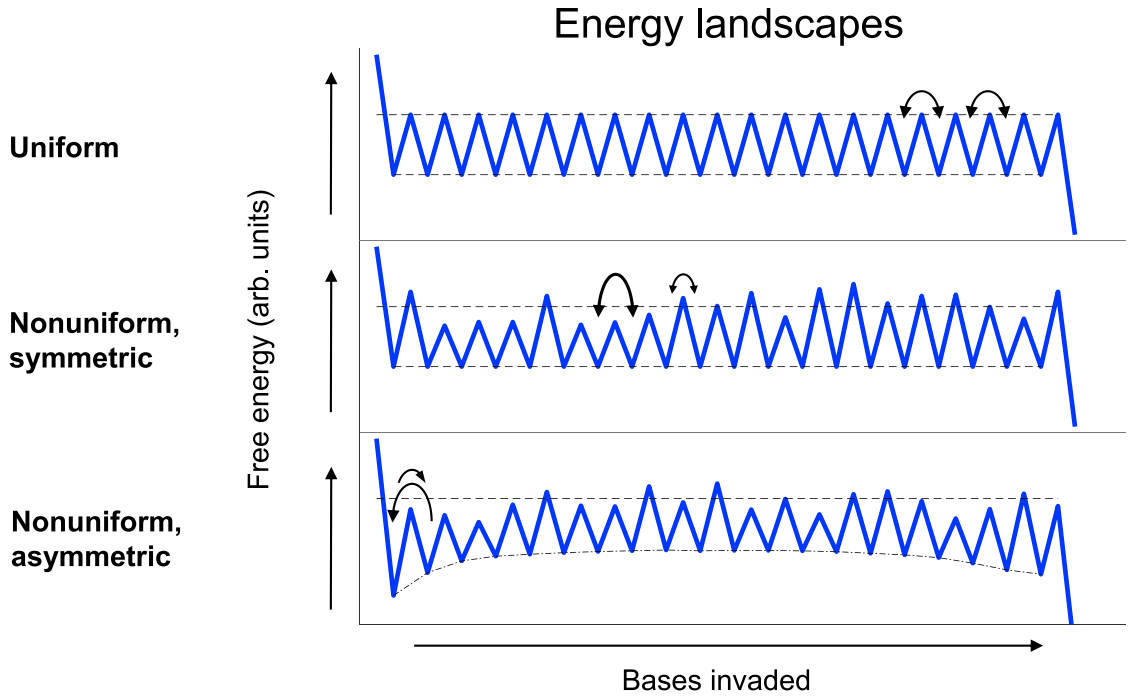


FIGURE 8 Free energy landscapes. All landscapes contain a reflecting boundary to the left and an absorbing boundary to the right. The uniform landscape is characterized by one uniform rate for transitions in any direction. The nonuniform, symmetric landscape allows for variation in rates so long as the forward rate from one state to another is set equal to the corresponding reverse rate. The nonuniform, asymmetric landscape additionally allows forward and reverse rates to differ. The relative free energies of the states in the nonuniform, asymmetric model are drawn from Srinivas et al. (29). To see this figure in color, go online.

adopts a coaxially unstacked state in which a transiently open incumbent base would be closer to the substrate base than the invading base (S · C*, Fig. 7). Therefore, we reason that k_{close} is also much faster than k_{replace}. Given $k_{\text{close}} \gg k_{\text{open}}, k_{\text{replace}}$, S · V will appear at the rate of

$$k \approx \frac{1}{[\text{S} \bullet \text{V}]} \frac{d[\text{S} \bullet \text{V}]}{dt} \approx \frac{k_{\text{open}}}{k_{\text{close}}} \bullet k_{\text{replace}} \ll k_{\text{open}}. \qquad (17)$$

Hence, the single-step migration rate k is expected to be much slower than the single basepair opening rate.

We stress that a symmetric random walk is an oversimplification of strand displacement. As shown in Fig. S14, the symmetric random walk model significantly underrepresents the range of observed displacement times; the fastest observed histogram and the slowest observed histogram are outside the range represented by the fitted curves. Therefore, the observed sequence dependence calls for a more complicated model. We list below several microscopic mechanisms that indicate strand displacement is more properly described as an asymmetric random walk ($f_{i-1} \neq r_i$).

First, displacement of the first basepair is energetically less favorable than the rest because it creates steric exclusion between dangling bases (29,46). Srinivas et al. (29) measured the thermodynamic penalty for the steric exclusion to be 2.0 kcal/mol at 25°C, which corresponds to ~30-fold slower f_0 than all other rates (k). According to Eq. 15, a bias in the first step ($\alpha_1 = f_0/r_1$) alters the mean first passage time to

$$\tau = \frac{1}{2k} n \left(n + \frac{2}{\alpha_1} - 1 \right). \qquad (18)$$

With a strong reverse bias ($\alpha_1 = 1/30$) in the first step, the single-step time (k^{-1}) is estimated to be 69 μs, faster than our previous estimate of 286 μs based on a completely symmetric random walk. 69 μs per step also falls well within the range (53–103 μs) inferred by Srinivas et al. (29). Second, the stability of a basepair is highly influenced by its nearest neighboring basepair, which would render the basepair opening rate direction dependent. For example, let us consider an A base in two adjacent branch migration intermediates $\text{G}^{\vee}\text{AC}$ and $\text{GA}^{\vee}\text{C}$, where "$^{\vee}$" refers to the branch point. In $\text{G}^{\vee}\text{AC}$, A is stacked more closely on C, whereas in $\text{GA}^{\vee}\text{C}$, A is stacked more closely on G. Therefore, the rate of A flipping out would be different between forward and reverse transitions. Third, the incumbent and the invader base at the branch point carry dangling strands of variable lengths depending on the state. These dangling strands will inevitably affect the diffusion rates of the bases at the branch point. To demonstrate this idea, we performed the fission assay with an invader extended by five nucleotides at the 3′ end. As shown in Fig. 9, the displacement kinetics become significantly slower even with the same

displacement domain. This result is consistent with the idea that a base with a longer dangling strand invades more slowly. As strand displacement progresses, the dangling part of the invader becomes shorter, and the dangling part of the incumbent becomes longer. Hence, the forward rate should become faster ($f_{i-1} < f_i$) and the reverse rate slower ($r_{i-1} > r_i$). These dangling strand-dependent rates produce asymmetric barriers in the free energy landscape, causing the basins to follow a concave curve (Figs. 8 and S15). Previous oxDNA simulations also predicted a concave free energy landscape (29,46). In the asymmetric random walk model, single-step rates are not only base dependent but also position dependent. Determining these rates would require measurements at a much larger scale, which is beyond the scope of this study.

We assumed that the number of steps is equal to the number of basepairs for modeling purposes. This could raise concern that this number may not accurately reflect the number of actual branch migration steps taken because the incumbent can spontaneously dissociate near the end of migration. In our previous work (32), we estimated spontaneous dissociation of a 2-bp incumbent to be ~10 μs, which would be comparable to the branch migration step rate we measured in this study. This means that the last few steps can occur either via branch migration or spontaneous dissociation, and the rate would be dominated by the faster of the two. Regardless, our proposed asymmetric branch migration model, in which forward rates become faster and reverse rates become slower, would effectively account for such an effect.

We made an interesting observation that RNA invasion and DNA invasion occur at very different rates even with the same invader sequence (except for T to U substitution). For the one sequence we tested, RNA invaded faster than DNA (Fig. 5) from one side but more slowly from the other side. Several factors may contribute to this finding. Structurally, an RNA-DNA hybrid duplex adopts an A-form helix (47–49), whereas a DNA-DNA duplex adopts a B-form helix. The thermodynamic stability difference between RNA-DNA and DNA-DNA duplexes depends on the sequence (50), with purine (AG)-rich substrate favoring RNA-DNA hybrid duplexes (51). Directional differences in stacking between DNA and RNA are known to persist even in the single-stranded form (52) and could contribute to this inversion of side dependence. In a similar vein, a recent study shows that coaxial stacking between an RNA-DNA hybrid duplex and a DNA-DNA homoduplex is stronger when the interhelical junction contains a 5′ RNA end than when it contains a 3′ RNA end (53). This effect may partially contribute to the faster top invasion by RNA shown in Fig. 5. However, another RNA sequence we tested exhibited faster bottom invasion than top invasion, suggesting that the base sequence is a stronger determinant of the displacement rate than the invasion polarity. RNA invasion of a DNA duplex in particular is a fundamental feature of the CRISPR-Cas system (10,54,55). R-loop formation appears to be the rate-limiting step for DNA cleavage (56,57) and is highly sequence dependent (56,58,59) but proceeds much more slowly (~1 s) than the spontaneous displacement rate we measured in this study. It will be thus interesting to investigate whether the sequence dependence is preserved between spontaneous and enzyme-mediated displacement reactions in the future.

The lack of salt dependence of the measured displacement kinetics was at first surprising to us because salt has a substantial effect on basepairing thermodynamics (60). Experimental measurements of salt-dependent opening and closing rates of a single basepair are scarce, but we can still infer their salt dependence from molecular

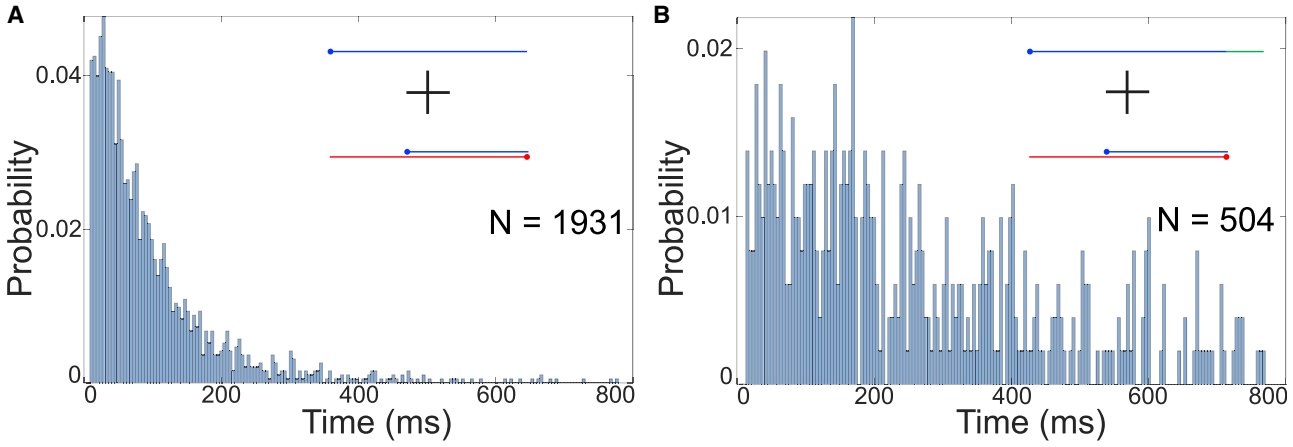


FIGURE 9 An overhang on the invader strand increases displacement times. (*A*) A top-invasion experiment with incumbent sequence 5′-GTCGGAATTT-TAAT-3′. The distribution of displacement times forms an exponential decay with little lag. After invasion has completed, every base on the invader strand is bound to the substrate. (*B*) An experiment conducted with a modified invader possessing a 5T overhang on the toehold-distal end (*green*). These additional bases affect the kinetics of displacement but should not bind to the substrate. The invader overhang increases overall displacement times and introduces a clear lag to the distribution. To see this figure in color, go online..

dynamics study (61) and hybridization and dissociation measurements of short oligos (62,63). These studies show that monovalent cations stabilize basepairing mainly by increasing the rate of basepair closing instead of decreasing the rate of basepair opening. Despite the strong salt dependence of basepair closing (k_{close} and $k_{replace}$), our proposed three-state model for branch migration (Eq. 17; Fig. 7) predicts that salt dependences of k_{close} and $k_{replace}$ will cancel each other out and render step migration rates, f and r, largely salt independent.

Even in the case in which fs and rs all carry a weak salt dependence through k_{open}, we can show that the overall salt dependence of the mean displacement time remains weak. Based on an experimental study (63), we assume a simple power law dependence of k_{open} on $[Na^+]$ ($k_{open} \sim [Na^+]^{\alpha}$) so that all fs and rs change by the same factor c upon changing $[Na^+]$. According to Eq. 14, the mean first passage time is equal to the ratio of two matrix determinants:

$$\tau = \frac{\det(\mathbf{A}')}{\det(\mathbf{A})}. \tag{19}$$

Because $\det(\mathbf{A}') \sim c^{n-1}$ and $\det(\mathbf{A}) \sim c^n$, $\tau \sim c^{-1}$. Hence, the overall displacement of the n basepair domain follows the weak salt dependence of k_{open}. Either way, we are able to rationalize the weak salt dependence of the mean displacement time (Fig. 6).

We hope that our results will be beneficial to the field of DNA nanotechnology. Our work has provided sequence-specific branch migration step times that could be used to rationally design sequences with desired kinetics. For example, our results could aid in the design of complex interaction networks between competing reactions with specifically tuned kinetics. Further, our fission assay opens the door to understanding branch migration kinetics in more reaction conditions than we studied here (e.g., buffers, pH, and temperature).

In this study, we assumed that strand displacement proceeds through one-dimensional branch migration, but it is possible that other mechanisms are at play. The invader might invade through the end distal to the toehold when terminal basepairs fray or through internal basepairs that spontaneously open up. Although internal invasion is highly unlikely for the short displacement domain we used here, it would be more probable for longer displacement domains. We also cannot rule out direct swapping between segments of invader and incumbent (64), invasion through triplex formation (65,66), or concurrent dissociation of a weakly bound incumbent strand (31,32). All these processes can occur in parallel, which makes it difficult to predict the strand displacement rate for any given sequence. In this regard, a future study on a much larger set of displacement domain sequences would help us to attain more accurate phenomenological models for explaining the sequence dependence of strand displacement kinetics.

CONCLUSIONS

We developed a novel smFRET assay that we call fission to study the timing of the unimolecular reaction that occurs during toehold-mediated strand displacement. Our fission assay separates the timescales between the slower toehold formation step and the faster displacement step and enabled us to tally displacement first passage times distributions for 11 separate invasion schemes. We found nontrivial sequence dependence in the distributions, whereas the mean first passage times varied by an order of magnitude. Further, we highlighted significant differences between the "side" of invasion, which suggest the kinetics are not completely determined by basepair sequence alone. Curiously, we showed that DNA and RNA invaders can behave drastically differently despite having identical sequences (apart from a T → U substitution). Finally, we demonstrated that displacement times were relatively unchanged over a wide range of salt concentrations. Motivated by these results, we developed a one-dimensional random walk model described above to estimate single base displacement times. This model is widely relevant to multistep processes, and we anticipate our analysis to be highly important to an array of biological reactions.

APPENDIX: ASYMMETRIC RANDOM WALK MODEL

We attempt to expand the random walk model to explain the 3.3-fold slower mean strand displacement time with the longer invader. This result suggests that the forward transition barrier would decrease with position as the overhang length of the invader becomes shorter. Likewise, the reverse transition barrier should increase with position. Although the exact position dependence of the barrier height is not known, we assume a simple analytical form (Hill equation or exponential function) for forward and reverse barrier heights (Eqs. SS.1 and SS.2).

$$\Delta G_f^*(i) = \frac{8.5RT}{1 + i/K} \tag{S.1}$$

$$\Delta G_r^*(i) = \frac{8.5RT}{1 + (14 - i)/K}. \tag{S.2}$$

Here, i represents the position along the energy landscape ($i \in \{0, 1, \ldots, 14\}$), the numerator comes from a previous estimate (29), and the factor K in the denominator determines the concavity of the landscape. For example, by substituting $K \approx 240$, we obtain a concave energy landscape as shown in Fig. S15, whose midpoint (position 7) of the landscape is about 1 kcal/mol higher than the end states (position 0 or 14). This concavity is similar to that seen in the landscape constructed from simulations (29). If the invader is lengthened by five extra bases, all forward barriers become higher:

$$\Delta G_f^*(i) = \frac{8.5RT}{1 + (i - 5)/K}, \tag{S.3}$$

whereas the reverse barriers are unchanged (Eq. SS.3). Hence, the energy landscape becomes tilted (*dotted line*, Fig. S6), and the corresponding mean first passage time is predicted to be slower by 3.3-fold based on Eq. 14 of the main text. This crude exercise shows that our observed overhang effect is semiquantitatively consistent with a simple asymmetric random walk model.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2021.01.043.

AUTHOR CONTRIBUTIONS

D.W.B.B. and H.D.K. designed research. D.W.B.B. and A.W.C. performed experiments. D.W.B.B., A.W.C., and H.D.K. contributed analytical tools. D.W.B.B. and A.W.C. analyzed data. D.W.B.B., A.W.C., and H.D.K. wrote the manuscript.

ACKNOWLEDGMENTS

The authors thank the current and past members of the Kim laboratory for critical discussions during the research project and helpful comments on the manuscript.

This work was supported by the National Institutes of Health (R01GM112882) and the National Science Foundation (1517507).

REFERENCES

1. Ouldridge, T. E., P. Šulc, …, A. A. Louis. 2013. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.* 41:8886–8895.
2. Sanstead, P. J., and A. Tokmakoff. 2018. Direct observation of activated kinetics and downhill dynamics in DNA dehybridization. *J. Phys. Chem. B.* 122:3088–3100.
3. McKinney, S. A., A.-C. Déclais, …, T. Ha. 2003. Structural dynamics of individual Holliday junctions. *Nat. Struct. Biol.* 10:93–97.
4. Bugreev, D. V., O. M. Mazina, and A. V. Mazin. 2006. Rad54 protein promotes branch migration of Holliday junctions. *Nature.* 442:590–593.
5. Wright, W. D., S. S. Shah, and W.-D. Heyer. 2018. Homologous recombination and the repair of DNA double-strand breaks. *J. Biol. Chem.* 293:10524–10535.
6. Chen, Z., H. Yang, and N. P. Pavletich. 2008. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature.* 453:489–494.
7. Savir, Y., and T. Tlusty. 2010. RecA-mediated homology search as a nearly optimal signal detection system. *Mol. Cell.* 40:388–396.
8. Mi, C., S. Zhang, …, H. Zhang. 2020. Strand displacement DNA synthesis by DNA polymerase gp90 exo— of Pseudomonas aeruginosa phage 1. *Biochimie.* 170:73–87.
9. KIreeva, M., C. Trang, …, M. Kashlev. 2018. RNA–DNA and DNA–DNA base-pairing at the upstream edge of the transcription bubble regulate translocation of RNA polymerase and transcription rate. *Nucleic Acids Res.* 46:5764–5775.
10. Singh, D., S. H. Sternberg, …, T. Ha. 2016. Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun.* 7:12778.
11. Ivančić-Baće, I., J. Al Howard, and E. L. Bolt. 2012. Tuning in to interference: R-loops and cascade complexes in CRISPR immunity. *J. Mol. Biol.* 422:607–616.
12. Andersen, E. S., M. Dong, …, J. Kjems. 2009. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature.* 459:73–76.
13. Thubagere, A. J., W. Li, …, L. Qian. 2017. A cargo-sorting DNA robot. *Science.* 357:eaan6558.
14. Chang, Y., Z. Wu, …, R. Yuan. 2019. Simply constructed and highly efficient classified cargo-discharge DNA robot: a DNA walking nanomachine platform for ultrasensitive multiplexed sensing. *Anal. Chem.* 91:8123–8128.
15. Zhang, D. Y., and G. Seelig. 2011. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.* 3:103–113.
16. Cherry, K. M., and L. Qian. 2018. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature.* 559:370–376.
17. Oh, J.-M., C. C. Venters, …, G. Dreyfuss. 2020. U1 snRNP regulates cancer cell migration and invasion in vitro. *Nat. Commun.* 11:1.
18. Simmel, F. C., B. Yurke, and H. R. Singh. 2019. Principles and applications of nucleic acid strand displacement reactions. *Chem. Rev.* 119:6326–6369.
19. Wang, K., M.-Q. He, …, Y.-L. Yu. 2018. Autonomous DNA nanomachine based on cascade amplification of strand displacement and DNA walker for detection of multiple DNAs. *Biosens. Bioelectron.* 105:159–165.
20. Li, Y., H. Chen, …, J. Zhang. 2019. Cellular interface supported toehold strand displacement cascade for amplified dual-electrochemical signal and its application for tumor cell analysis. *Anal. Chim. Acta.* 1064:25–32.
21. Sapkota, K., A. Kaur, …, S. Dhakal. 2019. Single-step FRET-based detection of femtomoles DNA. *Sensors (Basel).* 19:3495.
22. Cui, Y.-X., X.-N. Feng, …, D.-M. Kong. 2019. An integrated-molecular-beacon based multiple exponential strand displacement amplification strategy for ultrasensitive detection of DNA methyltransferase activity. *Chem. Sci. (Camb.).* 10:2290–2297.
23. Lee, C. Y., H. Kim, …, H. G. Park. 2019. Fluorescent S1 nuclease assay utilizing exponential strand displacement amplification. *Analyst (Lond.).* 144:3364–3368.
24. Figg, C. A., P. H. Winegar, …, C. A. Mirkin. 2020. Controlling the DNA hybridization chain reaction. *J. Am. Chem. Soc.* 142:8596–8601.
25. Tang, W., W. Zhong, …, Y. Liu. 2020. DNA strand displacement reaction: a powerful tool for discriminating single nucleotide variants. *Top. Curr. Chem. (Cham).* 378:10.
26. Zhang, D. Y., and E. Winfree. 2009. Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* 131:17303–17314.
27. Radding, C. M., K. L. Beattie, …, R. C. Wiegand. 1977. Uptake of homologous single-stranded fragments by superhelical DNA. IV. Branch migration. *J. Mol. Biol.* 116:825–839.
28. Green, C., and C. Tibbetts. 1981. Reassociation rate limited displacement of DNA strands by branch migration. *Nucleic Acids Res.* 9:1905–1918.
29. Srinivas, N., T. E. Ouldridge, …, E. Winfree. 2013. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.* 41:10641–10658.
30. Qian, L., and E. Winfree. 2011. Scaling up digital circuit computation with DNA strand displacement cascades. *Science.* 332:1196–1201.
31. Machinek, R. R. F., T. E. Ouldridge, …, A. J. Turberfield. 2014. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nat. Commun.* 5:5324.
32. Broadwater, D. W. B., Jr., and H. D. Kim. 2016. The effect of basepair mismatch on DNA strand displacement. *Biophys. J.* 110:1476–1484.
33. Chou, T., and M. R. D'Orsogna. 2014. First passage problems in biology. *In* First-Passage Phenomena and their Applications. R. Metzler, S. Redner, and G. Oshanin, eds. World Scientific, pp. 306–345.
34. Polizzi, N. F., M. J. Therien, and D. N. Beratan. 2016. Mean first-passage times in biology. *Isr. J. Chem.* 56:816–824.

35. Edelstein, A. D., M. A. Tsuchida, …, N. Stuurman. 2014. Advanced methods of microscope control using μManager software. *J. Biol. Methods.* 1:e10.

36. Le, T. T., and H. D. Kim. 2014. Studying DNA looping by single-molecule FRET. *J. Vis. Exp* 51667.

37. Aitken, C. E., R. A. Marshall, and J. D. Puglisi. 2008. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* 94:1826–1835.

38. Hartich, D., and A. Godec. 2019. Interlacing relaxation and first-passage phenomena in reversible discrete and continuous space Markovian dynamics. *J. Stat. Mech.* 2019:024002.

39. Kim, S. K. 1958. Mean first passage time for a random walker and its application to chemical kinetics. *J. Chem. Phys.* 28:1057.

40. Huguet, J. M., C. V. Bizarro, …, F. Ritort. 2010. Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. USA.* 107:15431–15436.

41. Andreatta, D., S. Sen, …, M. A. Berg. 2006. Ultrafast dynamics in DNA: "fraying" at the end of the helix. *J. Am. Chem. Soc.* 128:6885–6892.

42. Banavali, N. K. 2013. Partial base flipping is sufficient for strand slippage near DNA duplex termini. *J. Am. Chem. Soc.* 135:8274–8282.

43. Zgarbová, M., M. Otyepka, …, P. Jurečka. 2014. Base pair fraying in molecular dynamics simulations of DNA and RNA. *J. Chem. Theory Comput.* 10:3177–3189.

44. Lindahl, V., A. Villa, and B. Hess. 2017. Sequence dependency of canonical base pair opening in the DNA double helix. *PLoS Comput. Biol.* 13:e1005463.

45. Frank-Kamenetskii, M. D., and S. Prakash. 2014. Fluctuations in the DNA double helix: a critical review. *Phys. Life Rev.* 11:153–170.

46. Šulc, P., T. E. Ouldridge, …, A. A. Louis. 2015. Modelling toehold-mediated RNA strand displacement. *Biophys. J.* 108:1238–1247.

47. Gyi, J. I., A. N. Lane, …, T. Brown. 1998. Solution structures of DNA.RNA hybrids with purine-rich and pyrimidine-rich strands: comparison with the homologous DNA and RNA duplexes. *Biochemistry.* 37:73–80.

48. Conn, G. L., T. Brown, and G. A. Leonard. 1999. The crystal structure of the RNA/DNA hybrid r(GAAGAGAAGC). d(GCTTCTCTTC) shows significant differences to that found in solution. *Nucleic Acids Res.* 27:555–561.

49. de Oliveira Martins, E., V. B. Barbosa, and G. Weber. 2019. DNA/RNA hybrid mesoscopic model shows strong stability dependence with deoxypyrimidine content and stacking interactions similar to RNA/RNA. *Chem. Phys. Lett.* 715:14–19.

50. Sugimoto, N., S. Nakano, …, M. Sasaki. 1995. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry.* 34:11211–11216.

51. Huppert, J. L. 2008. Thermodynamic prediction of RNA-DNA duplex-forming regions in the human genome. *Mol. Biosyst.* 4:686–691.

52. Isaksson, J., S. Acharya, …, J. Chattopadhyaya. 2004. Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern. *Biochemistry.* 43:15996–16010.

53. Cofsky, J. C., D. Karandur, …, J. A. Doudna. 2020. CRISPR-Cas12a exploits R-loop asymmetry to form double-strand breaks. *eLife.* 9:e55143.

54. Mulepati, S., A. Héroux, and S. Bailey. 2014. Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science.* 345:1479–1484.

55. Hong, F., and P. Šulc. 2019. An emergent understanding of strand displacement in RNA biology. *J. Struct. Biol.* 207:241–249.

56. McCoy, D. E., T. Feo, …, R. O. Prum. 2018. Structural absorption by barbule microstructures of super black bird of paradise feathers. *Nat. Commun.* 9:1.

57. Gong, S., H. H. Yu, …, D. W. Taylor. 2018. DNA unwinding is the primary determinant of CRISPR-cas9 activity. *Cell Rep.* 22:359–371.

58. Szczelkun, M. D., M. S. Tikhomirova, …, R. Seidel. 2014. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. USA.* 111:9798–9803.

59. Zeng, Y., Y. Cui, …, J. Lou. 2018. The initiation, propagation and dynamics of CRISPR-SpyCas9 R-loop complex. *Nucleic Acids Res.* 46:350–361.

60. SantaLucia, J., Jr., and D. Hicks. 2004. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 33:415–440.

61. Wang, Y., T. Liu, …, W. Zhang. 2020. Salt effect on thermodynamics and kinetics of a single RNA base pair. *RNA.* 26:470–480.

62. Cisse, I. I., H. Kim, and T. Ha. 2012. A rule of seven in Watson-Crick base-pairing of mismatched sequences. *Nat. Struct. Mol. Biol.* 19:623–627.

63. Dupuis, N. F., E. D. Holmstrom, and D. J. Nesbitt. 2013. Single-molecule kinetics reveal cation-promoted DNA duplex formation through ordering of single-stranded helices. *Biophys. J.* 105:756–766.

64. Paramanathan, T., D. Reeves, …, J. Gelles. 2014. A general mechanism for competitor-induced dissociation of molecular complexes. *Nat. Commun.* 5:5207.

65. Lee, I.-B., S.-C. Hong, …, A. Johner. 2012. Kinetics of the triplex-duplex transition in DNA. *Biophys. J.* 103:2492–2501.

66. Chen, J., Q. Tang, …, J. Yan. 2017. Parallel triplex structure formed between stretched single-stranded DNA and homologous duplex DNA. *Nucleic Acids Res.* 45:10032–10041.

Biophysical Journal, Volume *120*

Supplemental information

First passage time study of DNA strand displacement

D.W. Bo Broadwater Jr., Alexander W. Cook, and Harold D. Kim

SUPPORTING MATERIAL

Oligonucleotide sequences used in this study

Supplementary Table S1. Invader sequences. The first 8 sequences correspond to the sequences in Figure 4 with top invaders preceding bottom invaders. The RNA sequences are identical to the first two sequences of this table but with a T→U substitution and correspond to partial duplexes formed from the first two rows of Tables S2 and S3. The final invader is identical to the first but with a 5T extention on the toehold-distal side, as seen in Fig. 9.

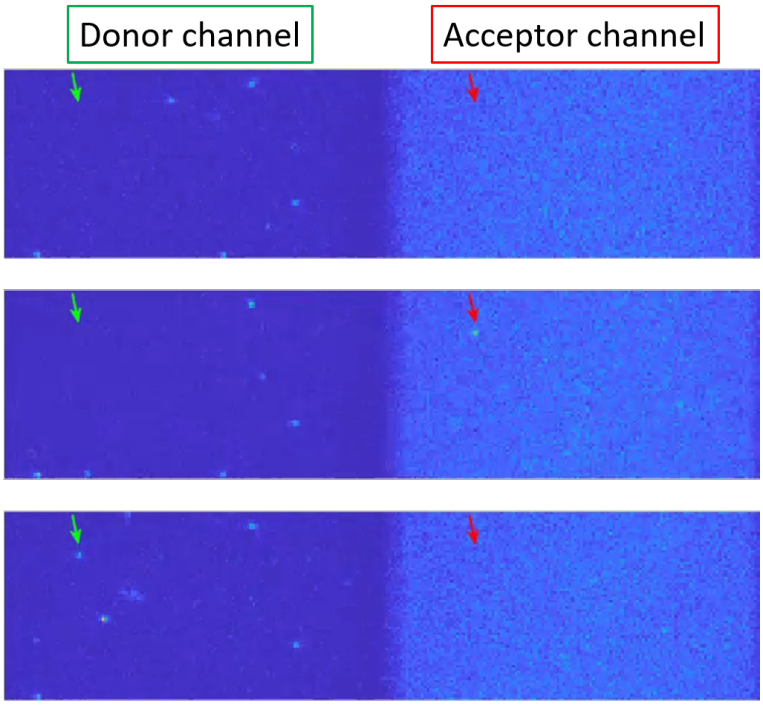
5′-/BioTEG/ACCTGGTGTTGTCGGAATTTTAAT-3′
5′-ATTAAAATTCCGACAACACCAGGT/BioTEG/-3′
5′-/BioTEG/GGTTATTGGGTTGGGTGTGTGAAA-3′
5′-TTTCACACACCCAACCCAATAACC/BioTEG/-3′
5′-/BioTEG/CAATCAAATAAAACTCTCTCTAAA-3′
5′-TTTAGAGAGAGTTTTATTTGATTG/BioTEG/-3′
5′-TTTCCTCCTAAAAGACCACCACCT/BioTEG/-3′
5′-CACACGTACGAACAAACACCAGGT/BioTEG/-3′
5′-/BioTEG/ACCUGGUGUUGUCGGAAUUUUAAU-3′
5′-AUUAAAAUUCCGACAACACCAGGU/BioTEG/-3′
5′-/BioTEG/ACCTGGTGTTGTCGGAATTTTAATTTTTT-3′

Supplementary Table S2. Incumbent sequences. The order corresponds to Figure 4 with top invaders preceding bottom invaders.

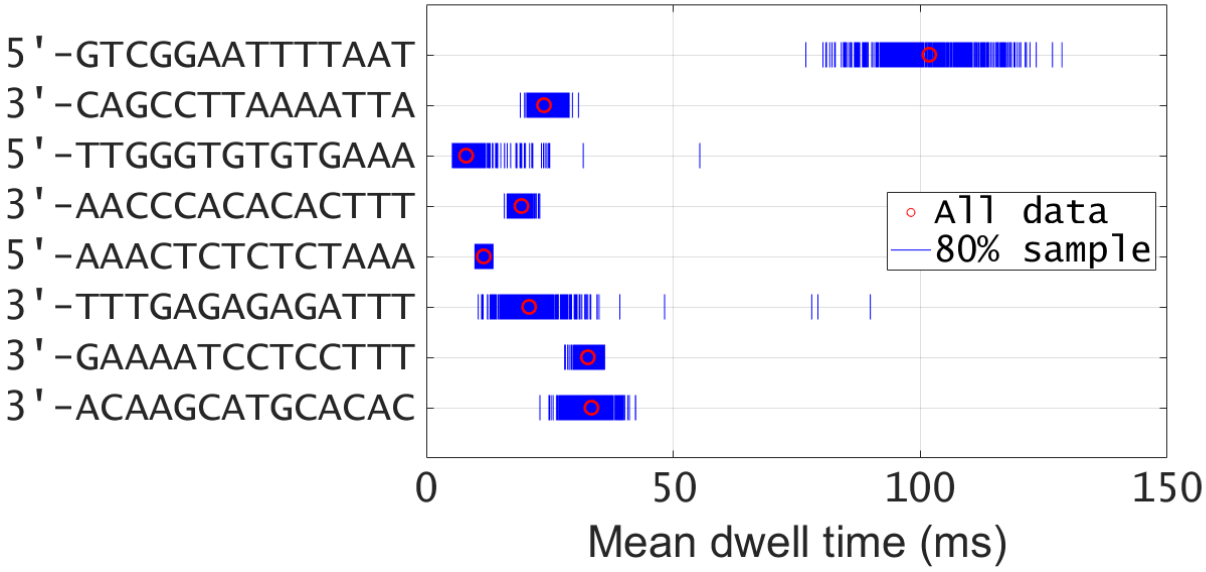
5′-GTCGGAATTTTAAT/Cy5/-3′
5′-/Cy5/CACACGTACGAACA-3′
5′-TTGGGTGTGTGAAA/Cy5/-3′
5′-/Cy5/TTTCACACACCCAA-3′
5′-AAACTCTCTCTAAA/Cy5/-3′
5′-CTTTTAGGAGGAAA/Cy5/-3′
5′-/Cy5/GACGACTATAGTTC-3′
5′-TGTTCGTACGTGTG/Cy5/-3′

Supplementary Table S3. Substrate sequences. The order corresponds to Figure 4 with top invaders preceding bottom invaders.

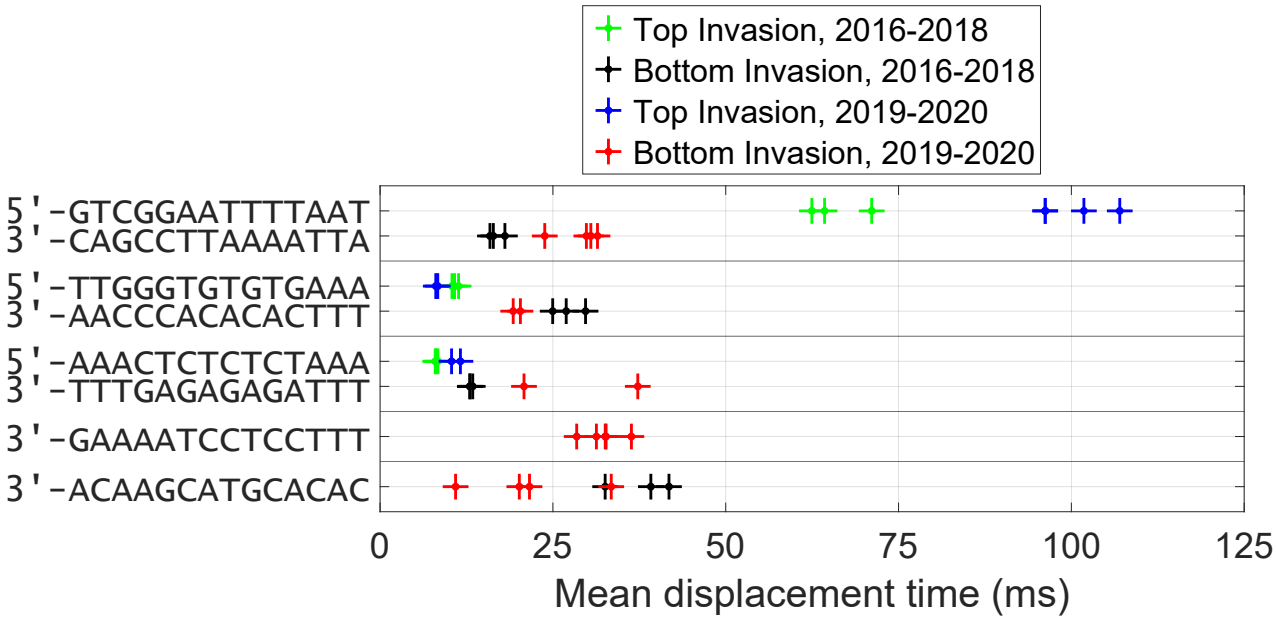
5′-CA/Cy3/ATTAAAATTCCGACAACACCAGGT-3′
5′-ACCTGGTGTTGTCGGAATTTTAAT/Cy3/AC-3′
5′-CA/Cy3/TTTCACACACCCAACCCAATAACC-3′
5′-GGTTATTGGGTTGGGTGTGTGAAA/Cy3/AC-3′
5′-CA/Cy3/TTTAGAGAGAGTTTTATTTGATTG-3′
5′-CAATCAAATAAAACTCTCTCTAAA/Cy3/AC-3′
5′-AGGTGGTGGTCTTTTAGGAGGAAA/Cy3/AC-3′
5′-ACCTGGTGTTTGTTCGTACGTGTG/Cy3/AC-3′



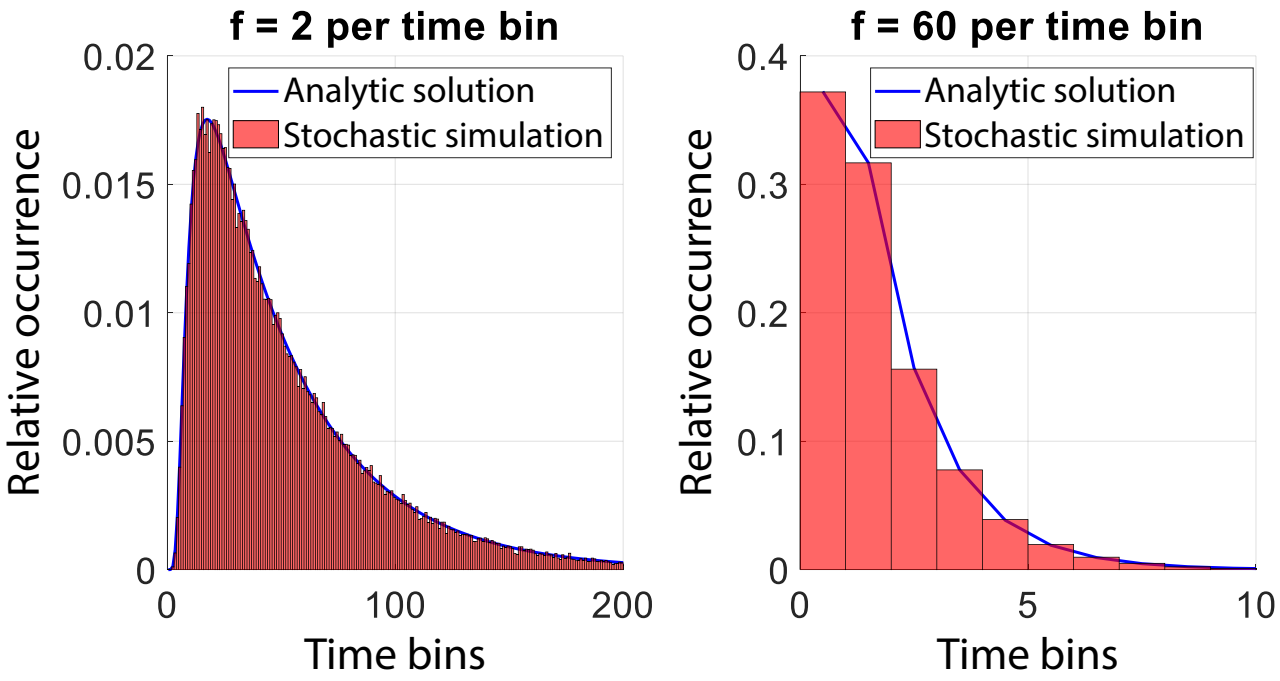
Supplementary Figure S1. Unprocessed single frame data acquired by the EMCCD with an exposure time of 3.96 ms. The green and red arrows track a single displacement event in the donor and acceptor channels, respectively. The top image shows no fluorescence signal. The middle image shows a fluorescent spot in the acceptor channel due to binding of the partial duplex to the invader. The bottom image shows a fluorescent spot in the donor channel due to the dissociation of the incumbent strand.



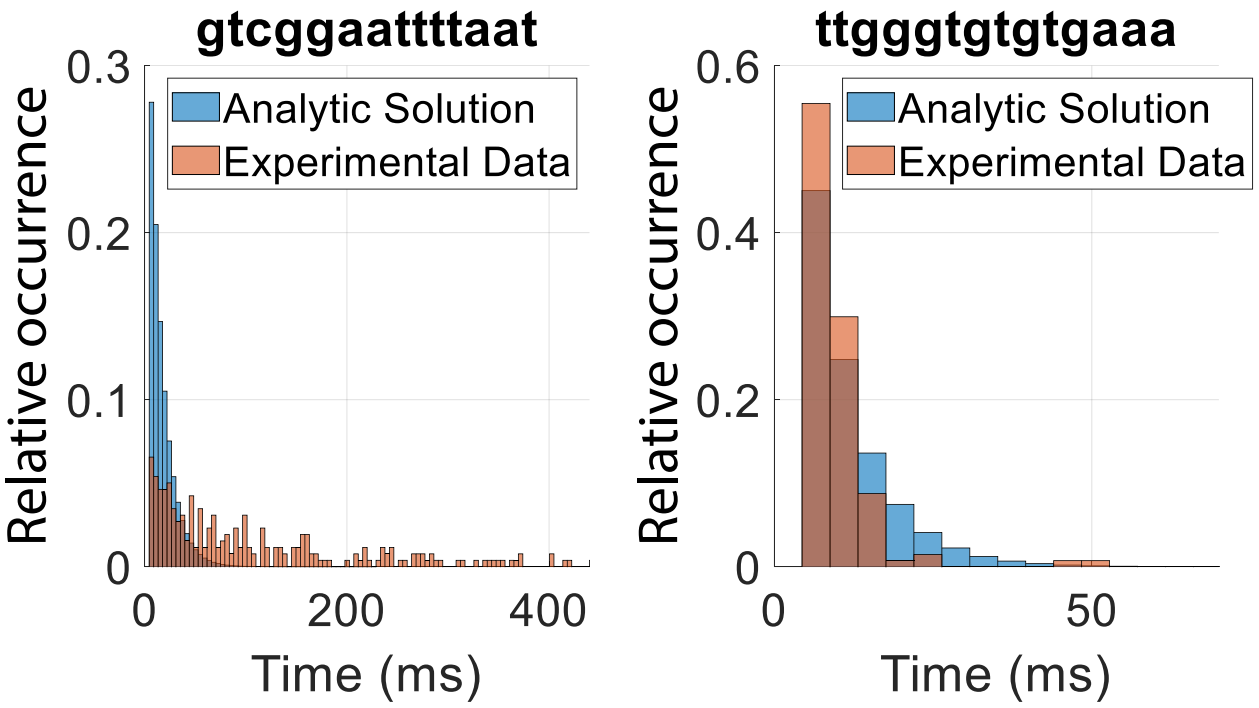
Supplementary Figure S2. A representation of the variance of mean displacement times. The mean displacement time is shown in red for one series of experiments using each of the eight sequences employed in this study. Each blue line represents the mean of a random sample of the data points within the experiment. Each sample is 80% the size of the total data set and allows for replacement between each selected point; 500 such samples have been taken for each sequence. Before each mean is calculated, the longest 5% of dwell times are discarded. This result demonstrates that the separation in mean dwell times between different sequences is not merely the result of insufficiently robust statistics.



Supplementary Figure S3. A comparison between mean dwell times calculated over the course of four years. Two different graduate students worked on this project. Student 1 collected data between 2016 and 2018, while Student 2 collected data from 2019 to 2020. The sequences of the corresponding incumbent strands are labelled on the left.



Supplementary Figure S4. Dwell-time histogram vs. bin width. Both plots show the normalized histograms of first passage times of a homogeneous random walker on a lattice with 14 sites. The red histograms are produced from a stochastic simulation (n = 10000), while the blue lines represent the analytic solution given by Eq. 5. A key feature of multi-step processes is a noticeable lag in the first passage time distribution (left). The reason for this is that some time must pass for a random walker to reach the end if there are multiple steps. However, if the binning time is large compared to the single step rate, then the lag can disappear (right). Further, for large binning times, the mean first passage time (measured in bins) must equal about half the number of lattice sites which implies the distribution will have a very short tail.



Supplementary Figure S5. Analytic solutions compared to experimental results. Both histograms show comparisons between experimentally obtained distributions of first passage times (orange histograms) and the solution yielded by Eq. 5 using the best-fit A, C, and T rates (blue histograms). The slowest-displacing sequence (left) and the fastest (right) are shown here. Both plots are labelled with the sequence of the invasion domain in the $5'$ to $3'$ direction.

Asymmetric random walk model

We attempt to expand the random walk model to explain the 3.3-fold slower mean strand displacement time with the longer invader. This result suggests that the forward transition barrier would decrease with position as the overhang length of the invader becomes shorter. Likewise, the reverse transition barrier should increase with position. Although the exact position dependence of the barrier height is not known, we assume a simple analytical form (Hill equation or exponential function) for forward and reverse barrier heights (SEqs. S1 and S2).

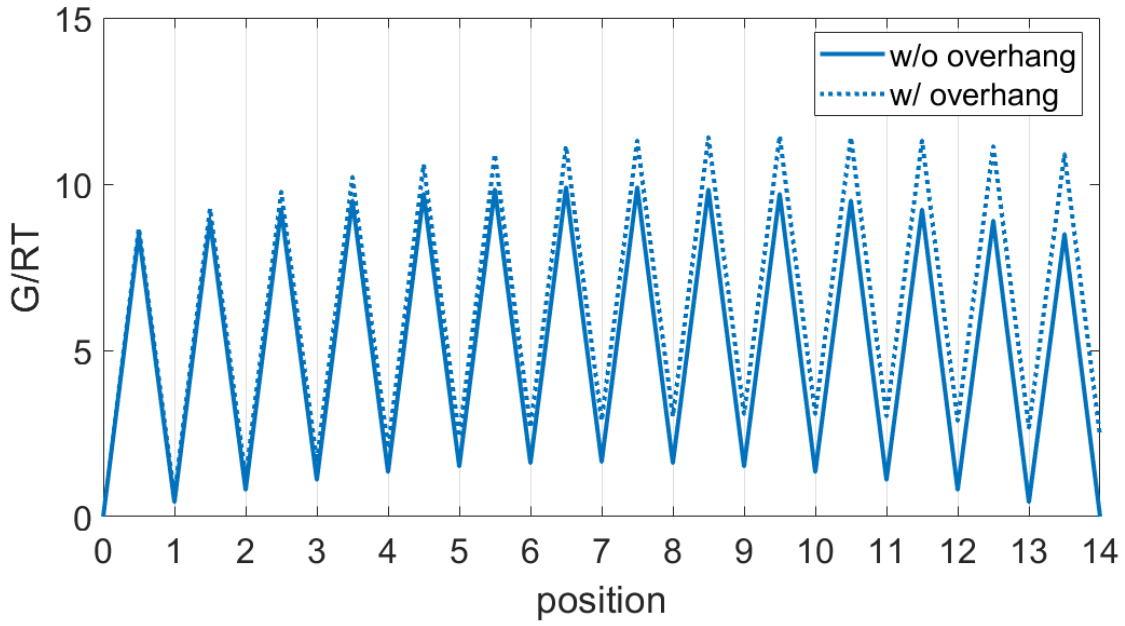
$$\Delta G_f^*(i) = \frac{8.5RT}{1 + i/K} \tag{S1}$$

$$\Delta G_r^*(i) = \frac{8.5RT}{1 + (14 - i)/K} \tag{S2}$$

Here, i represents the position along the energy landscape ($i \in \{0, 1, ..., 14\}$), the numerator comes from a previous estimate [30], and the factor K in the denominator determines the concavity of the landscape. For example, by substituting $K \approx 240$, we obtain a concave energy landscape as shown in SFig. S6, whose midpoint (position 7) of the landscape is about 1 kcal/mol higher than the end states (position 0 or 14). This concavity is similar to that seen in the landscape constructed from simulations [30]. If the invader is lengthened by 5 extra bases, all forward barriers become higher:

$$\Delta G_f^*(i) = \frac{8.5RT}{1 + (i - 5)/K} \tag{S3}$$

whereas the reverse barriers are unchanged (SEq. S3). Hence, the energy landscape becomes tilted (dotted line, SFig. S6), and the corresponding mean first passage time is predicted to be slower by 3.3-fold based on Eq. 14 of the main text. This crude exercise shows that our observed overhang effect is semi-quantitatively consistent with a simple asymmetric random walk model.



Supplementary Figure S6. Free energy landscape for asymmetric random walk. As strand displacement progresses, the length of the invader becomes shorter while the length of the incumbent becomes longer. Therefore, we expect the forward rate to become faster, and the reverse rate to become slower with the position along the free energy landscape. A free energy landscape that is consistent with this kinetic model is represented by the solid line. Extending the invader by an overhang of five bases causes the landscape to tilt, as shown by the dotted line.