Article

Modelling Toehold-Mediated RNA Strand Displacement

Petr Šulc,[1,2,*] Thomas E. Ouldridge,[2,3] Flavio Romano,[4] Jonathan P. K. Doye,[4] and Ard A. Louis[2]
[1]Center for Studies in Physics and Biology, Rockefeller University, New York, New York; [2]Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, United Kingdom; [3]Department of Mathematics, Imperial College, London, United Kingdom; and [4]Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford, United Kingdom

ABSTRACT We study the thermodynamics and kinetics of an RNA toehold-mediated strand displacement reaction with a recently developed coarse-grained model of RNA. Strand displacement, during which a single strand displaces a different strand previously bound to a complementary substrate strand, is an essential mechanism in active nucleic acid nanotechnology and has also been hypothesized to occur in vivo. We study the rate of displacement reactions as a function of the length of the toehold and temperature and make two experimentally testable predictions: that the displacement is faster if the toehold is placed at the 5′ end of the substrate; and that the displacement slows down with increasing temperature for longer toeholds.

INTRODUCTION

The emerging field of RNA nanotechnology aims to construct nanoscale structures and devices by using RNA strands (1,2). It is closely related to the more established and rapidly developing field of DNA nanotechnology, which exploits the specificity of Watson-Crick basepairing to construct impressive artificial nanoscale structures and active devices (3–10). While RNA and DNA are similar molecules, composed of a sugar-phosphate backbone to which an alphabet of four different nucleobases can attach, there are some differences. For example, in DNA the sugar is deoxyribose, and the four bases are (A, C, G, T), whereas in RNA the sugar is ribose, and the base T is replaced by U. Like DNA, RNA can form two kinds of Watson-Crick basepairs (AU or GC), but it has a higher propensity for other kinds of bonding, including the wobble basepair (GU) and other non-Watson-Crick basepairings (11) as well as numerous tertiary structure interactions.

In the cell, both molecules can store information, but whereas isolated DNA is primarily found in a B-helical double-stranded state, RNA is more versatile. RNA is typically found folded from single-stranded states into complex three-dimensional structures that contain A-form helical double-stranded segments as well as loops, bulges, and junctions. This increased structural repertoire facilitates biological functionality, so that RNA molecules can perform multiple additional roles, including catalysis, genetic regulation, structural support, and templating for molecular recognition and DNA synthesis (12–14). Because it can accomplish both storage of genetic material (like DNA) as well as catalytic activity (like proteins), it has been postulated that early life was based on RNA before DNA-based organisms appeared (the RNA-world hypothesis (15)). On the one hand, this versatility makes the prospect of using RNA nanotechnology very appealing, especially for biomedical applications. On the other hand, it makes predicting the three-dimensional structure significantly more challenging (1,16,17) than for DNA. For example, rather than designing de novo sequences that would fold into a particular functional three-dimensional structure, RNA nanotechnology often relies on functional motifs from known biologically occurring structures (18).

In this article we study the RNA equivalent of toehold-mediated strand displacement, a relatively simple dynamic reaction that has been a key component in many active devices in DNA nanotechnology. During strand displacement, a single-stranded invading strand replaces an incumbent strand that was bound in a duplex with a substrate strand. Both the invading and incumbent strands are complementary to the substrate strand, but the incumbent strand is usually a few bases shorter. When bound to the incumbent strand, the substrate strand hence has a short single-stranded overhanging region (a toehold) to which the invading strand can bind.

Systems based on DNA strand-displacement reactions have been shown to be able to perform computation (19–21) and are the basis of autonomous DNA motors (22,23). Strand-displacement reactions also have a great potential for use in RNA nanotechnology applications. For instance, a series of different reactions that involve several RNA strand-displacement steps were developed in Hochrein et al. (24). The strand displacement reaction is triggered by the presence of an mRNA strand with a particular sequence and the final product is an siRNA complex (24). One promising application in vivo is the conditional knockout of a gene by an RNA-silencing mechanism in

Submitted November 14, 2014, and accepted for publication January 26, 2015.

*Correspondence: petr.sulc@rockefeller.edu

Editor: Timothy Lohman.

0006-3495/15/03/1238/10 $2.00

http://dx.doi.org/10.1016/j.bpj.2015.01.023

the presence of mRNA created by a transcription of the triggering gene. It was recently shown in vitro that a cascade of RNA strand displacement reactions can be used for the detection of a single-stranded RNA with a particular sequence that triggers the reaction (25). Furthermore, RNA switches have been recently introduced for use in synthetic gene circuits, where the presence of a specific trigger RNA strand opens an RNA hairpin via a strand-displacement reaction. The hairpin loop contains a ribosome binding region that becomes accessible for transcription after the successful completion of the displacement. The functionality of RNA switches has been demonstrated in vivo as well as for diagnostics in vitro (26,27).

There are multiple regions of the genome transcribed into RNA molecules that are not further involved in protein production, but are themselves the final product. Thousands of these noncoding RNAs (ncRNA) have been identified and their function is a very active field of research (28–30). Some of the ncRNA interactions also involve RNA-RNA contacts and the creation of RNA double-stranded sections (31,32). It is hence plausible that RNA strand displacement might be involved in ncRNA interactions. For instance, it was hypothesized in Homann et al. (33) that RNA strand displacement reactions can occur in vivo in ribozyme-product complexes and it was shown in vitro that a single-stranded RNA can displace the cleaved strand that was bound to a ribozyme (33). More generally, the biological versatility of RNA makes it a prime candidate for many different in vivo applications of nanotechnology (1,18,34).

In this work we apply a recently developed coarse-grained model of RNA, oxRNA (35), to study the biophysics of toehold-mediated RNA strand displacement reactions. OxRNA's development followed the coarse-graining approach used in a previously developed coarse-grained model of DNA, oxDNA (36–39), which was shown to accurately reproduce the kinetics and thermodynamics of DNA strand displacement (40), including phenomena such as the effect of mismatches on displacement rates (41).

While the thermodynamics and kinetics of DNA strand displacement has been previously carefully studied both by experiments and by simulations (40,42), no systematic study is available for RNA. However, if RNA strand displacement-based devices are to realize their full potential, an understanding of the underlying mechanisms is necessary. While one might expect the general features of the reaction to be similar to that for DNA, we will also demonstrate interesting differences.

This article is organized as follows: We introduce the RNA model and the simulation methods in the next section. We then study the free-energy profile of the reaction, the rates of displacement as a function of the length and the placement (at either the 3′ or 5′ end) of the toehold, and the effect of temperature on the rates. We find that the reaction rate is increased by six orders of magnitude when the toehold length increases from 1 to 6 bases at 37°C. In contrast to DNA behavior, we find a noticeable difference in the displacement rate depending on whether the toehold is placed at the 3′ or 5′ end, with the 5′ being faster. Finally, we observe that the rate of displacement decreases with increasing temperature for a three-base toehold, but by less than expected from simply taking into account the decrease in stability of the toehold.

MATERIALS AND METHODS

A coarse-grained model of RNA

We give a brief description of the oxRNA model here, while the details of the structural, mechanical, and thermodynamical properties of the model are provided in Šulc et al. (35).

OxRNA represents each nucleotide as a single rigid body with multiple interaction sites. The rigid bodies interact with effective anisotropic interactions that are designed to capture the overall thermodynamic and structural consequence of the basepairing, stacking, and backbone interactions, as schematically shown in Fig. 1. The potential energy function of the oxRNA model is

$$V_{\mathrm{oxRNA}} = \sum_{\langle ij \rangle} \left(V_{\mathrm{backbone}} + V_{\mathrm{stack}} + V'_{\mathrm{exc}} \right) + \\ + \sum_{i,j \notin \langle ij \rangle} \left(V_{\mathrm{H.B.}} + V_{\mathrm{cross\ st.}} + V_{\mathrm{exc}} + V_{\mathrm{coaxial\ st.}} \right), \quad (1)$$

where the first sum runs over all pairs of nucleotides that are nearest neighbors on the same strand and the second sum runs over all other pairs.

The backbone interaction, V_{backbone}, is an isotropic, finitely-extensible nonlinear elastic potential and depends only on the distance between the backbone sites of the two adjacent nucleotides. This potential is used to mimic the covalent bonds in the RNA backbone that constrain the intramolecular distance between neighboring nucleotides. The nucleotides also have repulsive excluded-volume interactions V_{exc} and V'_{exc} that depend on the distance between their interaction sites, namely the backbone-backbone, stacking-stacking, and stacking-backbone distances. The excluded-volume interactions ensure that strands cannot overlap, or pass through each other in a dynamical simulation.

The duplex is stabilized by hydrogen-bonding ($V_{\mathrm{H.B.}}$), stacking (V_{stack}), and cross-stacking ($V_{\mathrm{cross\ st.}}$) interactions. These potentials are highly anisotropic and depend on the distance between the relevant interaction sites as well as the mutual orientations of the nucleotides.

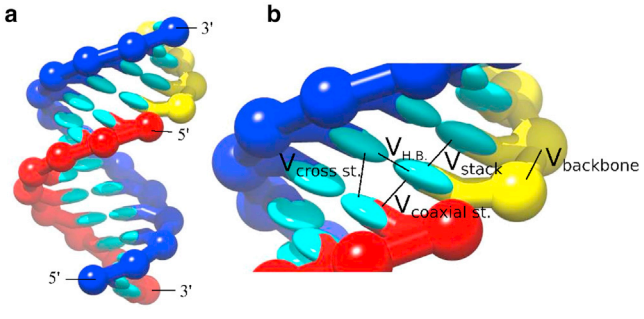

FIGURE 1 Schematic representation of (*a*) an A-RNA helix as represented by the model and (*b*) the attractive interactions in oxRNA (35). To see this figure in color, go online.

The hydrogen-bonding term $V_{\mathrm{H.B.}}$ is designed to capture the duplex-stabilizing interactions between Watson-Crick and wobble basepairs. The stacking interaction V_{stack} mimics the favorable interaction between adjacent bases, which results from a combination of hydrophobic, electrostatic, and dispersion effects.

The cross-stacking potential, $V_{\mathrm{cross\ st.}}$, is designed to capture the interactions between diagonally opposite bases in a duplex and has its minimum when the distance and mutual orientation between nucleotides corresponds to that for a nucleotide and the 3′ neighbor of the directly opposite nucleotide in an A-form helix. This interaction has been parameterized to capture the stabilization of an RNA duplex by a 3′ overhang (43). OxRNA does not include any interaction with the 5′ neighbor of the directly opposite nucleotide, as 5′ overhangs are significantly less stabilizing than 3′ overhangs (43).

The coaxial stacking potential $V_{\mathrm{coaxial\ st.}}$ represents the stacking interaction between nucleotides that are not nearest neighbors on the same strand.

As of this writing, the model does not include explicit electrostatic interactions. It was parameterized to reproduce RNA behavior at high (1-M) salt concentrations where the long-range interactions between phosphate charges on the backbone are screened. The remaining short-range electrostatics repulsion interactions are incorporated into the excluded volume potentials.

The interactions in the model were parameterized to reproduce the melting temperatures of hairpins and short oligomers as predicted by the nearest-neighbor model of RNA thermodynamics (43,44), which is the basis of most RNA secondary structure prediction tools (45–47). Two parameterizations of oxRNA are available—the sequence-averaged and the sequence-dependent version. In the sequence-dependent parameterization, the hydrogen-bonding and stacking interactions have different strengths depending on the types of the interacting bases. The sequence-averaged interaction strengths do not distinguish between the different base types, and in particular the interaction strengths of the hydrogen-bonding potentials are the same for Watson-Crick basepairs (AU and CG), and zero otherwise. The sequence-averaged interaction strengths were fitted to reproduce the thermodynamics of an averaged nearest-neighbor model, where the respective free-energy contributions were averaged over all possible basepair steps with Watson-Crick complementary basepairs.

In this work, we use the sequence-averaged parameterization, because we aim to study the displacement process for different toehold lengths and temperatures without the complications of sequence-specific effects. Our results will be comparable to experiments where average sequences are used, i.e., sequences with similar stability to the ones obtained from the averaged nearest-neighbor model for a given sequence length.

Simulation methods

The thermodynamics of the strand-displacement process is probed using the virtual move Mote Carlo (VMMC) algorithm (specifically the variant described in the Appendix of Whitelam et al. (48)) combined with umbrella sampling to help the system overcome free-energy barriers (49). The chosen order parameters were the number of bonds between the invader and the substrate strand, and bonds between the incumbent strand and the substrate. The weights assigned to the respective states were chosen by experience and then adapted by hand to ensure an efficient sampling.

We use the forward flux sampling (FFS) method (50) to estimate the rate of the strand displacement reaction. The FFS technique facilitates simulation of transition pathways by splitting up the rare transition event into multiple stages that are sampled separately. We previously used FFS simulations with the coarse-grained model of DNA, oxDNA, to study the kinetics of DNA strand displacement (40) and hybridization (51,52). In this work, we use direct FFS (described in detail in the Supporting Material of Ouldridge et al. (51)). The FFS simulations are run after an initial equilibration of 10^5 steps. The initial flux is measured through the interface at which the minimum distance between the complementary bases of the invading strand and the substrate strand is smaller than 0.84 Å. We then measured the probability of successfully crossing successive interfaces that were defined by the number of the formed bonds between the invading strand and the substrate (see the Supporting Material for details).

In our simulations, we define a basepair as being formed if the hydrogen-bonding energy ($V_{\mathrm{H.B.}}$ in Eq. 1) between the two bases is more negative than $-1.0\ k_{\mathrm{B}}T$ for $T = 300$ K, which corresponds to ~18% of typical hydrogen-bonding energies in the oxRNA model. We do not allow any misbonds, i.e., only the designed correct basepairs between the invading strand and the substrate (and the incumbent strand and the substrate) can form.

The molecular dynamics simulations for the FFS were carried out with an Andersen-like thermostat (described in the Appendix of Russo et al. (53)). All simulations are carried out with a simulation box size that corresponds to an equal strand concentration of 65.3 μM for the invading strand and substrate. Even though nucleic acid nanotechnology experiments are usually done at lower strand concentrations (approximately nanomolar or micromolar), we chose a higher concentration for our simulations in order not to spend too much time simply simulating the diffusion before the encounter and reaction of the strands. Such an approach allows us to efficiently extract the relative second-order rate constants (40,51).

The simulation step δt was chosen to be 1.53×10^{-14} s. We further used a high translational diffusion constant for an RNA nucleotide in our simulations, $D = 5.8 \times 10^{-7}\ \mathrm{m^2\ s^{-1}}$, which corresponds to a diffusion constant of $2.1 \times 10^{-8}\ \mathrm{m^2\ s^{-1}}$ for a 14-mer, larger than the experimentally measured $D_{\mathrm{exp}} = 0.92 \times 10^{-10}\ \mathrm{m^2\ s^{-1}}$ (54). We intentionally use a higher diffusion constant to speed up the diffusion of the strands in our simulations. Again, because the motion of strands is still diffusive over appreciable timescales, this should speed up the simulation without altering the basic physics of the reactions. The rotational diffusion coefficient was set to $D_{\mathrm{rot}} = 2.45 \times 10^{12}\ \mathrm{s^{-1}}$ and each nucleotide was treated as a rigid body with a diagonal inertia tensor.

One needs to be cautious in interpreting time units in coarse-grained simulations. For instance, timescales of different processes might scale with different factors (55). Furthermore, we use an artificially high diffusion constant to make simulations more efficient. Therefore, rather than trying to map our simulation results onto experimental rates, our emphasis is on computing the relative rates of similar processes (e.g., strand displacement with different lengths for the toehold), which can then be compared with experimental data should they become available. Similar methodology, applied to oxDNA, produced excellent agreement with experiment for relative rates (40,41).

RESULTS AND DISCUSSION

The different stages of the reaction are introduced in Fig. 2 as represented by the oxRNA model. We note that the respective stages of the reaction are reversible, and it is possible for the invading strand to detach before successfully completing the displacement.

The strand-displacement reaction in DNA has been studied in detail both experimentally (10,42) and theoretically with the help of a simplified one-dimensional model and models based on secondary structure as well as a previously developed coarse-grained model of DNA, oxDNA (10,40–42).

In Zhang and Winfree (42), the rate of the toehold-mediated strand displacement was measured as a function of the length of the toehold. For an average strength toehold (i.e., a sequence with roughly the same number of AT and GC bonds), it was found that the rate of the reaction increases exponentially with the length of the toehold until it saturates for a toehold of approximately six bases, where the rate is ~6.5 orders of magnitude faster than for a system with no

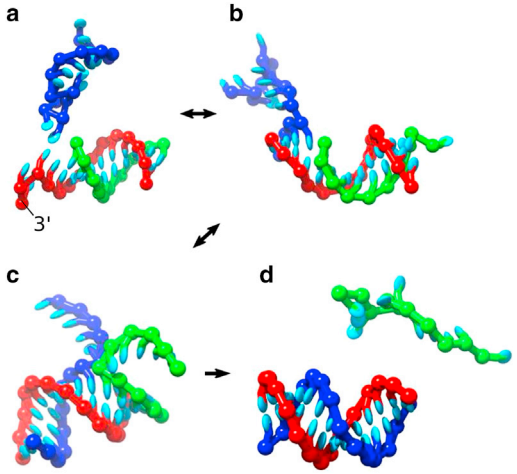

FIGURE 2 Schematic representation of the different stages of a strand displacement reaction as represented by oxRNA for a 4-nucleotide-long toehold at the 3′ end of the substrate. (*a*) The invading strand (*blue*) attaches to the toehold of the substrate (*red*). (*b*) The invading strand is fully bound to the toehold. (*c*) The invading strand displaces bonds between the incumbent (*green*) strand and the substrate. (*d*) The incumbent strand loses all bonds with the substrate. To see this figure in color, go online.

toehold. A very similar speedup was also observed in coarse-grained simulations with oxDNA (40).

We are confident that our oxRNA model is able to provide insight into the RNA strand displacement reaction, because it adopts a very similar coarse-graining approach to that developed for oxDNA, which was shown to sufficiently describe the necessary biophysical properties of DNA to capture both thermodynamic and kinetic aspects of the strand displacement reaction (40).

In the rest of this work, we will study the strand displacement reaction with toeholds of length ranging from 1 to 6, placed at either end of the substrate. The substrate has 10 basepairs with the incumbent strand. We use the oxRNA model to obtain a free-energy profile for the strand-displacement reaction, and the rates of the RNA displacement as a function of toehold length, position (3′ or 5′), and temperature.

Free-energy profile

We measure the free-energy profile of the toehold-mediated strand displacement reaction at 37°C by sampling the states of a system consisting of a 14-nucleotide substrate strand with a 4-nucleotide toehold at either the 3′ or 5′ end. A 14-nucleotide invading strand is attached to the toehold and displaces a 10-nucleotide incumbent strand. The system is illustrated in Fig. 3 *a*. We use VMMC (combined with umbrella sampling) to sample the free-energy landscape as a function of the number of the bonds between each strand and the substrate. We plot the free energy in Fig. 3 *b* as a function of the number of bonds between the invading strand and the substrate. The VMMC algorithm was run

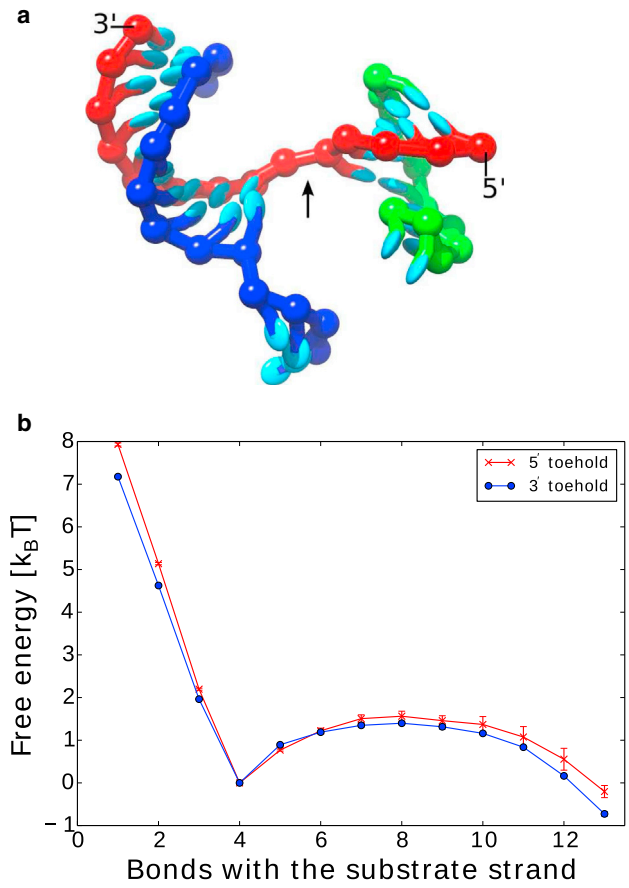

FIGURE 3 (*a*) A 14-base substrate strand (*red*) with a 14-base invader strand (*blue*) attached at its 3′ 4-base toehold and a 10-base incumbent strand (*green*). The stacking interaction at the branch migration junction (indicated by the *black arrow*) is broken, making it easier for the single-stranded overhangs of the strands to avoid each other. (*b*) Free-energy profile as a function of the number of bonds between the invading strand and the substrate for a 3′ (*blue dots*) and 5′ (*red crosses*) toehold at 37°C. The free energy has been set to 0 when the invading strand has four bonds with the substrate. The 5′ toehold has a larger free energy gain upon forming the first 4 bonds than the 3′ toehold does. To see this figure in color, go online.

for at least 5×10^{11} cluster-move attempts for each system. The errors for each point, estimated as the maximum difference from the mean from three independent simulation sets for each toehold, are at most 0.2 k_BT and are shown as error bars in the plot. In the simulation, we only allowed basepairing between the substrate and its complementary bases on the other two strands. We required that the strands always have at least one bond with the substrate to prevent dissociation. To compare the 3′ and 5′ toeholds, we plot the free-energy profiles with the free energy set to 0 when the invading strand is fully bound to the toehold (four basepairs). We note that the choice of the state where free energy is equal to 0 is the one of convenience, because only the free-energy differences between states are of importance.

The free energy decreases as more bonds are formed between the invading strand and the toehold and reaches its minimum when the invading strand is fully bound to the

toehold. There is a free-energy barrier associated with the initiation of the displacement (i.e., branch migration, where the invading strand gains additional bonds with the substrate at the expense of the incumbent strand) of ~1.6 k_BT (corresponding to ~1 kcal/mol), which saturates after ~4 bases have been displaced. A similar barrier was observed for DNA displacement (40,56), and it arises from the fact that at the branch point (the point on the substrate where the invading and incumbent strands meet) the strands are in close proximity and the single-stranded overhangs of the invading and incumbent strands cannot overlap because of the excluded volume interactions. The two strands can move away from each other, but this requires the stacking interactions at the junction to be broken, as illustrated in Fig. 3 *a*. Hence, there is a free-energy penalty associated with taking the first steps of branch migration, even though the total number of basepairs between the substrate and the invading and incumbent strands remains the same throughout the process.

After the invading strand has bound to the toehold, it will eventually either initiate branch migration or fall off. We note that binding to the 5′ toehold is more favorable than binding to the 3′ toehold, because the invading strand attached at the 5′ end has an additional cross-stacking interaction. Due to the A-form helical structure, the 3′ overhang cross-stacking interaction is stronger than the 5′ overhang, because the distance toward the 3′ overhang is smaller, hence allowing for a stronger interaction (43). The free-energy barrier for detachment of an invading strand binding to a 5′ toehold is therefore higher than for an invading strand binding to a 3′ end, as can be seen in Fig. 3.

The average extra stabilization provided by an overhang at 37°C in the nearest-neighbor model of RNA thermodynamics, which is based on experimental measurements (43), is ~−0.4 k_BT for a 5′ overhang and −1.3 k_BT for a 3′ overhang. The cross-stacking interaction in the oxRNA model was parameterized to reproduce the average melting temperatures, predicted by the averaged nearest-neighbor model, for RNA duplexes with extra unpaired overhangs at the 3′ ends of the strand. It was found that the best way to capture this effect in oxRNA was by not including any stacking interaction with the 5′ overhang, because it is significantly weaker than for the 3′ overhang.

The additional stabilization provided by the cross-stacking between the substrate and the invading strand attached to the 5′ toehold is illustrated in Fig. 4 *a*, and is ~−1.7 k_BT in our model at 37°C. Hence, the probability of falling off once the invading strand is attached is lower for an invading strand at a 5′ toehold than at a 3′ toehold.

We also note that the free energy (with respect to the free-energy of the state when invader binds to the toehold) of the states with eight or more bonds between the invader and the substrate appears to be slightly lower (up to 0.6 k_BT) for the invader at a 3′ end. This relatively small effect, which is slightly larger than statistical errors, may reflect differences in conformational freedom that the displaced strand has due to the clashes with either the minor or major groove of the invader-substrate duplex. Furthermore, toward the end of the displacement, the remaining unpaired base at the 3′ invading strand can stabilize the invader-substrate duplex because it can provide an additional cross-stacking interaction as a 30 overhang, which is not the case for the 5′ invading strand. As has been pointed out elsewhere (41), the differences in the free-energy profile close to the toehold are far more important in determining rates. We shall see in the following section that, indeed, it is the differing toehold stabilities that is the dominant factor in determining relative displacement rates.

While Fig. 3 provides information about the free-energy landscape of the strand displacement reaction as a function of the number of bonds between the invading strand and the substrate strand, it does not give us the rate of the reaction. Even though some of the neighboring states during the branch migration procedure have almost no difference in free energy, the transition between them is a complicated process that requires the loss of a basepair between the incumbent strand and the substrate and the creation of the basepair between the invading strand and a substrate.

In a previous article (40), this complex process was directly simulated for DNA with oxDNA. Furthermore, the transition rates between adjacent states were described with a simplified intuitive one-dimensional model that

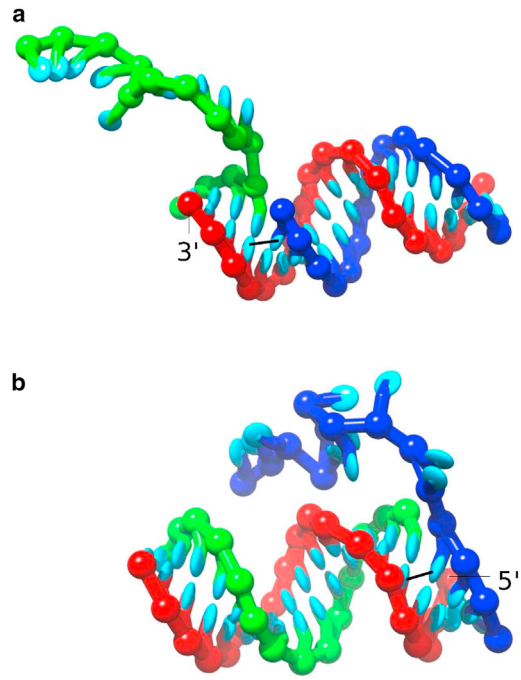

FIGURE 4 (*a*) An invading strand (*green*) at a 3′ toehold has no cross-stacking interaction with the substrate (*red*) strand at the junction. (*b*) An invading strand (*blue*) attached to a 5′ toehold. (*Black line*) The 3′ cross-stacking interaction at the junction of the invading and incumbent strand is indicated schematically. To see this figure in color, go online.

included an effective free-energy barrier height that could be fit to reproduce the rates. However, these barriers cannot be obtained without either experimental or computational input. To obtain kinetic information about the process, we therefore use FFS simulations, as described in the following sections.

Displacement rates for different toehold lengths and positions

We calculate the relative rates of strand displacement reactions using the FFS method, as outlined in the Materials and Methods. We studied systems where the incumbent strand (shown in *green* in Fig. 2) has 10 bonds with the substrate strand (shown in *red* in Fig. 2). The substrate has a toehold of length l that ranges from 1 to 6 bases. The invading (*blue*) strand is fully complementary to the substrate strand and can hence form up to $10 + l$ basepairs with the substrate strand once the incumbent strand is successfully removed. For each toehold length considered, we calculated the rates for toeholds placed at either the 3′ or 5′ ends of the substrate strand. All simulations were carried out at 37°C.

At experimental concentrations (micromolar or less), basic displacement reactions are well described by second-order kinetics. For computational convenience, we simulate much higher concentrations of strands but this can potentially lead to the breakdown of the second-order description, due to reaction intermediates that are relatively long-lived compared to the overall reaction timescales. To infer the relative second-order rate constants at low concentration from our data, we neglect the time spent in these intermediate states once binding to the toehold has occurred in our calculation of the reaction rates from the simulations. Note that we do not preclude the possibility of displacement failure after binding to the toehold; we simply ignore the time that such a process takes. More details of this approach are provided in Srinivas et al. (40) and Ouldridge et al. (51).

For short toeholds, displacement will often fail and the invading strand undergoes frequent binding and unbinding from the toehold. Because the rate of hybridization is only weakly dependent on the toehold length (40,57), it is mainly the probability of falling off of the invading strand from the toehold that will be affected by the toehold length. For long toeholds, the probability of falling off once fully bound to the toehold becomes very small, and hence it will then successfully complete displacement with a high probability. Thus we expect the rate of the overall displacement reaction to increase with increasing toehold length up to a length at which it saturates.

The rates of displacement as a function of toehold length are shown in Fig. 5. For each of the data points, we ran at least three independent sets of FFS simulations. We observe that the reaction rate of the toehold-mediated strand displacement reaction saturates for a toehold of ~5 nucleotides at a value that is ~5 orders of magnitude larger than the rate for a 1-nucleotide 5′ toehold and ~6 orders of magnitude larger than for a 1-nucleotide 3′ toehold.

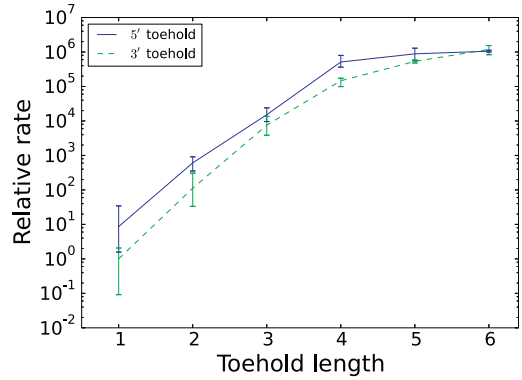

FIGURE 5 Semilogarithmic plot of mean rates of the strand displacement reaction as a function of toehold length for toeholds at the 3′ end (*dashed line*) and at the 5′ end (*solid line*) of the substrate strand. The error bars show the maximum and minimum rate obtained from all the simulations for a given length of toehold. The rates are normalized with respect to the mean rate for a 1-nucleotide toehold at the 3′ end. To see this figure in color, go online.

A similar range of relative rates was observed in DNA displacement reactions (40,42), where for the average toehold sequence, the rates saturated for a 6-nucleotide toehold at a value that is ~10^5 times larger than the rate for a 1-nucleotide toehold. We note, however, that we studied RNA displacement kinetics at 37°C, while the DNA displacement was studied at 25°C. Furthermore, no strong dependence on whether the toehold was at the 3′ or 5′ end of the substrate was observed for DNA. More generally, we expect RNA to saturate for slightly shorter toeholds because an average RNA sequence with Watson-Crick basepairs is more stable than an average DNA sequence.

We note that until the toeholds are long enough to reach the saturated rate, the mean rates for the 5′ toeholds are larger than those for the 3′ toeholds. The ratios of the mean rates for different ends of the substrate are shown in Table 1. As discussed in Free-Energy Profile, the invading strand at the 5′ end gains additional stabilization from cross-stacking interaction with the substrate strand (as shown schematically in Fig. 4), hence the probability for the invading strand to fall off from the 5′ toehold is lower and the displacement rate is increased.

TABLE 1 Ratios of the mean displacement rates for toeholds at 5′ and 3′ ends as estimated from the FFS simulations at 37°C

Toehold length	5′ rate/3′ rate
1	8.6
2	5.4
3	2.0
4	3.5
5	1.7
6	0.9

Our prediction of faster rates of displacement from a 5′ toehold is based on the experimentally verified fact that the cross-stacking with 3′ overhangs is significantly more stabilizing than the interaction with 5′ overhangs. The cross-stacking interaction in oxRNA captures the stabilization by a 3′ overhang, but the model does not include the stabilization by the 5′ overhangs, because they are significantly smaller. However, this interaction still might contribute somewhat to the stabilization of the displacement from a 3′ toehold, which would mean that our inferred speedup may be slightly overestimated.

The rates were obtained with the averaged oxRNA model, and apply to toeholds that would have stability similar to that of the averaged toehold (i.e., with approximately the same number of AU and GC bonds). However, if one designed a system with a weak toehold (AU-rich), then the saturation would be reached at a toehold length longer than 5, and similarly, for a strong (GC-rich) toehold, the saturation of the displacement speed would be reached at toehold lengths shorter than 5.

We further note that it is not only the number of AU and GC bonds that determines the free-energy stabilization for a particular sequence, it is also their order (35,39). Therefore, care needs to be taken in experiments when comparing, for instance, the rates of the 3' and 5' toeholds, because the rates will be affected by the stabilization provided by the respective toehold sequences as well as the type of bases involved in stacking and cross-stacking interactions, which also varies with sequence (43).

Furthermore, we would expect the saturated rates to be different for weak and strong toehold sequences, with the strong ones being faster, as was previously observed for DNA (42). It was observed for the oxDNA model (51) that the hybridization rates are faster for stronger sequences, because the strand has higher probability of completely binding to the toehold region after the first few basepairs were made if the basepairs created during the first contact are stronger. We expect the same sequence-dependent effects to play a role in RNA strand displacement. However, given the absence of systematic rate data on RNA strand displacement from which to compare, we focus only on the average-model description in this work.

Displacement rate at different temperatures

We further investigate the relative displacement rates for a 3-nucleotide toehold at the 3′ end for simulation temperatures ranging from 17 to 47°C. For each temperature considered, we performed at least three independent sets of FFS simulations. The mean relative rates of the displacement are shown in Fig. 6. The rates decrease with increasing temperature, with the displacement at 47°C being 5.6 times slower than at 17°C.

It is interesting to compare the temperature dependence of the rates with that for the yields of invading strands bound to the toehold only. We therefore also performed VMMC umbrella sampling simulations where the sampling of attachment and detachment of the invading strand to the 3-nucleotide toehold at the 3′ end of the substrate was performed at 37°C and extrapolated to temperatures in the range 17–47°C by a histogram reweighting method. In the simulation, the invading strand was allowed to make at most four bonds with the substrate strand. The VMMC simulations were run with the same simulation box size as the FFS simulations, corresponding to an equal strand concentration of 65.3 μM. We observed that the decrease of the displacement rate with increasing temperature is smaller than the decrease of the yield of the invading strands bound to the toehold, as shown in Table 2 and Fig. 6.

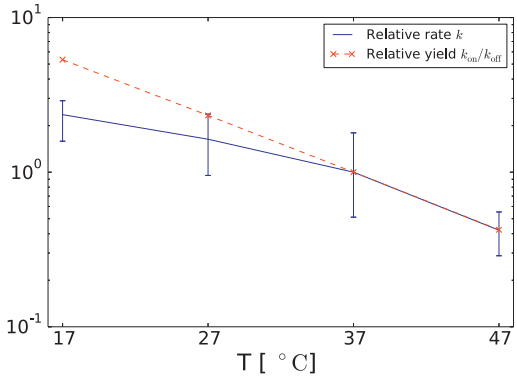

FIGURE 6 Semilogarithmic plot of mean rates of the strand displacement reaction and yields of attachment to the toehold for a 3-nucleotide toehold at the 3′ end as a function of temperature. The error bars show the maximum and minimum rate obtained from all the simulations for a given temperature. The rates and yields are normalized with respect to their values at 37°C. To see this figure in color, go online.

These results can be rationalized with a simple model for the toehold-mediated strand displacement reaction (see Srinivas et al. (40) or the Supporting Material). In the second-order limit that we consider (in which the lifetime of the three-stranded intermediate is neglected), the rate of displacement can be written as $k = k_{on} \times p_{bm|toe}$, where k_{on} is the rate of binding to the toehold and $p_{bm|toe}$ is the probability of a successful completion of branch migration once bound by the toehold. Further, $p_{bm|toe} = k_{bm}/k_{off}$, where k_{off} is the rate of unbinding from the toehold and k_{bm} is the

TABLE 2 Mean relative displacement rates *k* (*second column*) compared with the normalized yield of invading strands attached to the toehold of the substrate (*third column*)

T	Displacement rate k	k_{on}/k_{off}
17°C	2.4	5.3
27°C	1.6	2.3
37°C	1	1
47°C	0.4	0.4

The rates and yields were normalized with respect to their respective values obtained at 37°C.

rate at which the system initiates (and subsequently completes) branch migration. Thus, $k = k_{on} \times k_{bm}/k_{off}$.

The quantity k_{on}/k_{off} corresponds to the yield of invading strands attached to the toehold and is shown in the third column in Table 2. The difference in the temperature dependence of the displacement rate k and the yield of toehold-bound duplexes implies that k_{bm} must be affected by the change of temperature as well. The implied increase in k_{bm} with increasing temperature is presumably due to the disruption of basepairs involved in initiating branch migration. The passage over the initial barrier to displacement highlighted in Fig. 3 is hence made easier by increased temperature.

As the toehold gets shorter, we would expect k_{on}/k_{off} to become increasingly less dependent on temperature. Therefore, the temperature dependence of k_{bm} will be more dominant and we hence expect the slope of the graph of the displacement rate as a function of temperature to become less negative with decreasing toehold length. Indeed, for the shortest toehold lengths (and certainly for blunt-ended displacement), we would expect the displacement rate to increase with temperature, as has been reported in DNA (58).

CONCLUSIONS

We have studied toehold-mediated RNA strand displacement with a coarse-grained model of RNA, oxRNA (35). We observed behavior that was previously encountered in studies of DNA displacement (40). In particular, we showed that there is a free-energy barrier to the initiation of the displacement process that is caused by the need for the single-stranded overhangs of the invading and incumbent strand to avoid each other (entropic effect), and that in so doing they break the stacking at the branch migration junction (enthalpic effect). We found that the rate of displacement for long toeholds saturates at 5–6 orders of magnitude faster than for a single-base toehold, which is similar to observations for DNA.

We further found that for toeholds shorter than five nucleotides, i.e., before the saturation regime is reached, the displacement reaction is ~2–9 times faster at the 5′ end than at the 3′ end (at least for an average strand). The invading strand gets additional stability from a cross-stacking interaction at the 5′ end toehold that reduces the probability of falling off from the toehold, and hence increases the rate. Such an effect can be tested experimentally, and may provide an additional way to modulate displacement rates.

Finally, we studied the kinetics of the displacement reaction for a 3-nucleotide toehold at the 3′ end as a function temperature. The rates decrease by approximately a factor of 6 when the temperature is increased from 17 to 47°C. The displacement rate decreases with increasing temperature due to the destabilization of the toehold. However, although toehold dissociation is faster at higher temperature, so is the actual process of branch migration. We therefore expect that for the shortest toeholds, when dissociation will be weakly temperature-dependent, the displacement will be accelerated with increasing temperature, consistent with experimental observations for DNA (58).

The oxRNA model used for this work does not include explicit electrostatic interactions and was parameterized to RNA thermodynamics at 1-M salt. It is possible to extend our model by including a Debye-Hückel potential parameterized to reproduce thermodynamics at lower salt concentrations (59). Because the displacement involves two strands in close proximity during branch migration, it is expected that decreasing the salt concentration will have further effects on the rate of the displacement in addition to destabilizing the toehold. This problem, along with the role of mismatches in the toehold region, will be the subject of further work.

As of this writing, our model allows only Watson-Crick and wobble basepairs. It is possible that other kinds of tertiary structure contacts (such as ribose zippers or Hoogsteen basepairs) could influence the branch migration process. We have also disallowed misbonds in our simulations, meaning that no mismatched bonds between the substrate and the invading strand and incumbent strand could occur. However, sequences used in the experiments are usually designed to avoid such misbonds, and hence we would not expect misbonds to affect the conclusions drawn from our simulations. It would be interesting in the future to study the effect of misbonds and secondary structure formation on the full toehold-mediated displacement reaction. Further applications of our model might include designing RNA strands to bind to a long RNA strand with known secondary structure (60). It is possible that the designed invading strand needs to displace a double-stranded stem and hence the rate of such a process can also be influenced by the choice of the 3′ or 5′ end of the stem as the binding site.

When designing an RNA displacement reaction for applications in vivo, the number of possible sequences is greatly reduced due to biological restrictions on the sequence that have to be recognized by the interacting strands. Therefore, it is important to understand the factors that influence the kinetic rates of the underlying reaction. Our work hence provides estimates of how the rate of the reaction can be further enhanced, such as by placing the toehold at the 5′ end or increasing the toehold length.

The simulation code implementing the oxRNA model is freely available for download at dna.physics.ox.ac.uk, as a part of the oxDNA software package.

SUPPORTING MATERIAL

Supporting Materials and Methods and 22 tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(15)00117-4.

AUTHOR CONTRIBUTIONS

P.Š. designed and performed research, analyzed data, and wrote the manuscript. T.E.O., F.R., J.P.K.D., and A.A.L. designed the research, analyzed the data, and wrote the manuscript.

ACKNOWLEDGMENTS

We thank Erik Winfree, Niranjan Srinivas, and Terence Hwa for useful discussions, and Lorenzo Rovigatti for his contributions to the oxDNA simulation code development.

The authors acknowledge the Engineering and Physical Sciences Research Council and University College (Oxford) for financial support.

REFERENCES

1. Guo, P. 2010. The emerging field of RNA nanotechnology. *Nat. Nanotechnol.* 5:833–842.
2. Grabow, W. W., and L. Jaeger. 2014. RNA self-assembly and RNA nanotechnology. *Acc. Chem. Res.* 47:1871–1880.
3. Seeman, N. C. 1982. Nucleic acid junctions and lattices. *J. Theor. Biol.* 99:237–247.
4. Rothemund, P. W. K. 2006. Folding DNA to create nanoscale shapes and patterns. *Nature.* 440:297–302.
5. Douglas, S. M., H. Dietz, …, W. M. Shih. 2009. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature.* 459:414–418.
6. Douglas, S. M., I. Bachelet, and G. M. Church. 2012. A logic-gated nanorobot for targeted transport of molecular payloads. *Science.* 335:831–834.
7. Castro, C. E., F. Kilchherr, …, H. Dietz. 2011. A primer to scaffolded DNA origami. *Nat. Methods.* 8:221–229.
8. Linko, V., and H. Dietz. 2013. The enabled state of DNA nanotechnology. *Curr. Opin. Biotechnol.* 24:555–561.
9. Yurke, B., A. J. Turberfield, …, J. L. Neumann. 2000. A DNA-fueled molecular machine made of DNA. *Nature.* 406:605–608.
10. Yurke, B., and A. P. Mills, Jr. 2003. Using DNA to power nanostructures. *Genet. Program. Evol. M.* 4:111–122.
11. Leontis, N. B., J. Stombaugh, and E. Westhof. 2002. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* 30:3497–3531.
12. Elliott, D., and M. Ladomery. 2011. Molecular Biology of RNA. Oxford University Press, New York.
13. Alberts, B., A. Johnson, …, P. Walter. 2002. Molecular Biology of the Cell. Garland Science, New York.
14. Cech, T. R., and B. L. Bass. 1986. Biological catalysis by RNA. *Annu. Rev. Biochem.* 55:599–629.
15. Gilbert, W. 1986. Origin of life: the RNA world. *Nature.* 319:618.
16. Laing, C., and T. Schlick. 2010. Computational approaches to 3D modeling of RNA. *J. Phys. Condens. Matter.* 22:283101.
17. Laing, C., and T. Schlick. 2011. Computational approaches to RNA structure prediction, analysis, and design. *Curr. Opin. Struct. Biol.* 21:306–318.
18. Grabow, W., and L. Jaeger. 2013. RNA modularity for synthetic biology. *F1000Prime Rep.* 5:46.
19. Soloveichik, D., G. Seelig, and E. Winfree. 2010. DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci. USA.* 107: 5393–5398.
20. Qian, L., E. Winfree, and J. Bruck. 2011. Neural network computation with DNA strand displacement cascades. *Nature.* 475:368–372.
21. Seelig, G., D. Soloveichik, …, E. Winfree. 2006. Enzyme-free nucleic acid logic circuits. *Science.* 314:1585–1588.
22. Muscat, R. A., J. Bath, and A. J. Turberfield. 2011. A programmable molecular robot. *Nano Lett.* 11:982–987.
23. Wickham, S. F. J., J. Bath, …, A. J. Turberfield. 2012. A DNA-based molecular motor that can navigate a network of tracks. *Nat. Nanotechnol.* 7:169–173.
24. Hochrein, L. M., M. Schwarzkopf, …, N. A. Pierce. 2013. Conditional Dicer substrate formation via shape and sequence transduction with small conditional RNAs. *J. Am. Chem. Soc.* 135:17322–17330.
25. Sternberg, J. B., and N. A. Pierce. 2014. Exquisite sequence selectivity with small conditional RNAs. *Nano Lett.* 14:4568–4572.
26. Green, A. A., P. A. Silver, …, P. Yin. 2014. Toehold switches: de-novo-designed regulators of gene expression. *Cell.* 159:925–939.
27. Pardee, K., A. A. Green, …, J. J. Collins. 2014. Paper-based synthetic gene networks. *Cell.* 159:940–954.
28. Washietl, S., J. S. Pedersen, …, P. F. Stadler. 2007. Structured RNAs in the ENCODE selected regions of the human genome. *Genome Res.* 17:852–864.
29. Djebali, S., C. A. Davis, …, T. R. Gingeras. 2012. Landscape of transcription in human cells. *Nature.* 489:101–108.
30. Guttman, M., and J. L. Rinn. 2012. Modular regulatory principles of large non-coding RNAs. *Nature.* 482:339–346.
31. Hao, Y., Z. J. Zhang, …, H. Shi. 2011. Quantifying the sequence-function relation in gene silencing by bacterial small RNAs. *Proc. Natl. Acad. Sci. USA.* 108:12473–12478.
32. Gottesman, S. 2002. Stealth regulation: biological circuits with small RNA switches. *Genes Dev.* 16:2829–2842.
33. Homann, M., W. Nedbal, and G. Sczakiel. 1996. Dissociation of long-chain duplex RNA can occur via strand displacement in vitro: biological implications. *Nucleic Acids Res.* 24:4395–4400.
34. Takahashi, M. K., and J. B. Lucks. 2013. A modular strategy for engineering orthogonal chimeric RNA transcription regulators. *Nucleic Acids Res.* 41:7577–7588.
35. Šulc, P., F. Romano, …, A. A. Louis. 2014. A nucleotide-level coarse-grained model of RNA. *J. Chem. Phys.* 140:235102.
36. Ouldridge, T. E., A. A. Louis, and J. P. K. Doye. 2011. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* 134:085101.
37. Ouldridge, T. E., A. A. Louis, and J. P. K. Doye. 2010. DNA nanotweezers studied with a coarse-grained model of DNA. *Phys. Rev. Lett.* 104:178101.
38. Ouldridge, T. E. 2011. Coarse-grained modelling of DNA and DNA nanotechnology. Ph.D. thesis, Oxford University, New York.
39. Šulc, P., F. Romano, …, A. A. Louis. 2012. Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.* 137: 135101.
40. 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.
41. Machinek, R. R., T. E. Ouldridge, …, A. J. Turberfield. 2014. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nat. Commun.* 5:5324.
42. Zhang, D. Y., and E. Winfree. 2009. Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* 131:17303–17314.
43. Mathews, D. H., J. Sabina, …, D. H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* 288:911–940.
44. Xia, T., J. SantaLucia, Jr., …, D. H. Turner. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry.* 37: 14719–14735.
45. Zadeh, J. N., C. D. Steenberg, …, N. A. Pierce. 2011. NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* 32:170–173.
46. Xayaphoummine, A., T. Bucher, and H. Isambert. 2005. KINEFOLD web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.* 33:W605–W610.
47. Hofacker, I. L. 2003. Vienna RNA secondary structure server. *Nucleic Acids Res.* 31:3429–3431.

48. Whitelam, S., E. H. Feng, …, P. L. Geissler. 2009. The role of collective motion in examples of coarsening and self-assembly. *Soft Matter.* 5:1251–1262.

49. Torrie, G., and J. P. Valleau. 1977. Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J. Comput. Phys.* 23:187–199.

50. Allen, R. J., C. Valeriani, and P. Rein Ten Wolde. 2009. Forward flux sampling for rare event simulations. *J. Phys. Condens. Matter.* 21: 463102.

51. Ouldridge, T. E., P. Šulc, …, A. A. Louis. 2013. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.* 41:8886–8895.

52. Schreck, J. S., T. E. Ouldridge, …, J. P. K. Doye. 2014. DNA hairpins primarily promote duplex melting rather than inhibiting hybridization. arXiv preprint. arXiv:1408.4401.

53. Russo, J., P. Tartaglia, and F. Sciortino. 2009. Reversible gels of patchy particles: role of the valence. *J. Chem. Phys.* 131:014504.

54. Lapham, J., J. P. Rife, …, D. M. Crothers. 1997. Measurement of diffusion constants for nucleic acids by NMR. *J. Biomol. NMR.* 10:255–262.

55. Padding, J. T., and A. A. Louis. 2006. Hydrodynamic interactions and Brownian forces in colloidal suspensions: coarse-graining over time and length scales. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 74:031402.

56. Šulc, P., T. E. Ouldridge, …, A. A. Louis. 2014. Simulating a burnt-bridges DNA motor with a coarse-grained DNA model. *Nat. Comput.* 13:535–547.

57. Wetmur, J. G., and N. Davidson. 1968. Kinetics of renaturation of DNA. *J. Mol. Biol.* 31:349–370.

58. Reynaldo, L. P., A. V. Vologodskii, …, V. I. Lyamichev. 2000. The kinetics of oligonucleotide replacements. *J. Mol. Biol.* 297:511–520.

59. Wang, Q., and B. M. Pettitt. 2014. Modeling DNA thermodynamics under torsional stress. *Biophys. J.* 106:1182–1193.

60. Jakobsen, M. R., C. K. Damgaard, …, J. Kjems. 2004. A genomic selection strategy to identify accessible and dimerization blocking targets in the 5′-UTR of HIV-1 RNA. *Nucleic Acids Res.* 32:e67.

Modelling toehold-mediated RNA strand displacement: Supplementary Material

Petr Šulc,[1,2] Thomas E. Ouldridge,[2,3] Flavio Romano,[4] Jonathan P. K. Doye,[4] and Ard A. Louis[2]

[1] *Center for Studies in Physics and Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA*
[2] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford, OX1 3NP, United Kingdom*
[3] *Department of Mathematics, Imperial College, 180 Queen's Gate, London, SW7 2AZ, United Kingdom*
[4] *Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford, OX1 3QZ, United Kingdom*

1D-MODEL OF STRAND DISPLACEMENT

A simple 1-D model for the strand-displacement reaction was developed in Ref. 1 for DNA. For completeness, we repeat here the estimation of the displacement reaction rate with this simple model.

We assume that the rate at which the invading strand attaches to the toehold is k_{on}, which is concentration-dependent. The invading strands then proceeds to successfully displace the incumbent strand during branch migration with a probability $p_{\mathrm{bm|toe}}$, so the rate of the successful displacement is

$$k = k_{\mathrm{on}} \times p_{\mathrm{bm|toe}}. \tag{S1}$$

The branch migration process is initiated by displacing the first base of the incumbent strand with a rate k_{first}. The probability of initiating the branch migration rather than falling off is

$$p_{\mathrm{in}} = \frac{k_{\mathrm{first}}}{k_{\mathrm{first}} + k_{\mathrm{off}}}, \tag{S2}$$

where k_{off} is the rate for the invading strand to fall off once it is attached by the toehold.

We assume that there are b bases between the incumbent strand and the invading strand. Once the branch migration is initiated and the invading strand has displaced the first base, it has a probability $1/(b-1)$ of successfully completing the displacement, where we assumed that the branch migration is a random walk where the invading strand can gain or lose a base pair with equal probability. When the invading strand has displaced one base, the probability that the invading strand goes back to being bound just by the toehold is $1 - 1/(b-1)$. It can then again initiate displacement with probability p_{in}. We can hence approximate $p_{\mathrm{bm|toe}}$ as

$$p_{\mathrm{bm|toe}} = \frac{k_{\mathrm{first}}}{k_{\mathrm{first}} + k_{\mathrm{off}}} \left(\frac{1}{b-1} + \frac{b-2}{b-1} \times p_{\mathrm{bm|toe}} \right), \tag{S3}$$

from which we obtain

$$p_{\mathrm{bm|toe}} = \frac{k_{\mathrm{first}}}{k_{\mathrm{first}} + (b-1)k_{\mathrm{off}}} \tag{S4}$$

From Eqs. S4 and S1 we hence obtain

$$k = k_{\mathrm{on}} \times \frac{k_{\mathrm{first}}}{k_{\mathrm{first}} + (b-1)k_{\mathrm{off}}} = \frac{k_{\mathrm{on}}}{1 + (b-1)\frac{k_{\mathrm{off}}}{k_{\mathrm{first}}}}. \tag{S5}$$

In the limit of short toeholds, before the saturation regime of the displacement reaction is reached, we can assume $1 \ll (b-1)\frac{k_{\mathrm{off}}}{k_{\mathrm{first}}}$ and simplify Eq. S5 to

$$k \approx \frac{k_{\mathrm{on}}}{k_{\mathrm{off}}} \times \frac{k_{\mathrm{first}}}{b-1}, \tag{S6}$$

which is used in the main text, where we used $k_{\mathrm{bm}} = \frac{k_{\mathrm{first}}}{b-1}$.

For long toeholds, when the saturation regime is reached, we can assume $1 \gg (b-1)\frac{k_{\mathrm{off}}}{k_{\mathrm{first}}}$ and the displacement rate becomes

$$k \approx k_{\mathrm{on}}. \tag{S7}$$

FFS SIMULATIONS

The Forward Flux Sampling (FFS) algorithm [2] allows for efficient simulations of a transition from an initial (meta)stable state A (denoted as a state $Q = -2$ for our system) to a final stable state B ($Q = Q_{\max}$, where $Q_{\max}$ will be either 3 or 4 for our system, depending on the choice of order parameters as discussed in the following section). We use FFS to calculate the fluxes from a state where the invading strand is unbound to a state where the invading strand successfully binds to the substrate and removes the incumbent strand. FFS introduces a series of interfaces λ in the state space between the A and B states. First, a 'brute-force' simulation is run to estimate the flux of trajectories that leave state A and cross the first interface. Then, one selects at random a crossing point at the first interface λ^0_{-1} that was obtained from the generated trajectories and propagates it further until it either returns to the state A or reaches the second interface λ^1_0. By repeating this process one obtains an ensemble of points at the interface λ^1_0 and an estimate of the probability $P\left(\lambda^1_0|\lambda^0_{-1}\right)$ of reaching the interface λ^1_0. The procedure is repeated iteratively for the subsequent interfaces, and one can then estimate the transition rate from state A ($Q = -2$) to state B ($Q = Q_{\max}$) as

$$k_{AB} = \phi^0_{-1} \prod_{Q=0}^{Q=Q_{\max}} P\left(\lambda^{Q+1}_Q|\lambda^Q_{Q-1}\right) \tag{S8}$$

where ϕ^0_{-1} is the flux of trajectories leaving state $Q = -2$ and crossing the interface λ^0_{-1}.

Q	Definition
-2	$d > 3.36$
-1	$0.84 < d \leq 3.36$
0	$d \leq 0.84$ & $b = 0$
1	$b \geq 1$
2	$b \geq l$
3	$b \geq l + 3$
4	$b = 10 + l$ & $d_2 > 3.36$

TABLE S-1. The definition of order parameters Q for simulations with toehold lengths $l = 3$ or smaller. d is the minimum distance between complementary bases in the invading strand and the toehold on the substrate and b is the number of bonds between the invading strand and the substrate. d_2 is the minimum distance between all complementary bases of the incumbent strand and the substrate. All distances are in nm. Substrate is $10 + l$ bases long and toehold is l base long. For toehold of length $l = 1$, the order parameter $Q = 2$ is defined as $b \geq 2$.

Q	Definition
-2	$d > 3.36$
-1	$0.84 < d \leq 3.36$
0	$d \leq 0.84$ & $b = 0$
1	$b \geq 1$
2	$b \geq l$
3	$b = 10 + l$ & $d_2 > 3.36$

TABLE S-2. The definition of order parameters Q for simulations with toehold lengths l. For toehold of length $l = 1$, the order parameter $Q = 2$ is defined as $b \geq 2$. The definitions of variables in the table are the same as the ones used in Table S-1

We previously used the FFS algorithm to study the kinetics of reactions involving DNA strands with the oxDNA model, and detailed description of this approach can be found in Refs. 1, 3 and 4.

Order parameters

The definition of the order parameters Q used in the FFS simulations is provided in Tables S-1 and S-2. We originally used order parameters from Table S-2, but due to limited accuracy we then decided to run simulations for 3-nucleotide (or smaller) toeholds with order parameters as defined in Table S-1. The systems with order parameters from Table S-1 have one additional interface compared to Table S-2. The data obtained was consistent, and is included in Fig. 5 in the main text. If simulations using order parameters both from Table S-1 and Table S-2 were considered for a given system, the reported FFS results are shown in two separate tables (one for each respective choice of order parameters).

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.39\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	9000	1372515	0.0066 (0.0007)
λ_1^2	9022	71953	0.125 (0.04)
λ_2^3	922	930	0.991 (0.003)

TABLE S-3. FFS results for a 6-nucleotide toehold at the $3'$ end. The data were obtained from 3 independent sets of simulations.

FFS simulation results

The results of the forward flux sampling simulations are provided in Tables S-3 to S-19 for toeholds of lengths 1 to 6 placed at either the $3'$ or $5'$ end at simulation temperature 37 °C. The results for simulations of a 3-base toehold at the $3'$ end for temperatures 17 °C, 27 °C, and 47 °C are shown in Tables S-20, S-21, and S-22, respectively. For each system considered, at least three independent simulations were carried out. The rates shown in Figs. 5 and 6 in the main text were obtained as the average of the rates from the simulations, with errorbars showing the maximum and minimum rates encountered.

Tables S-3 to S-22 show the number of successful crossings of a given interface along with the number of trajectories launched from the previous interface, obtained by summing all trajectories from all independent simulations for a given system. The probability of successful interface crossing shown in the tables is obtained as the number of crossings divided by the number of attempts. The numbers in brackets are the maximum absolute difference of the probability (or flux) shown in the table and the respective probabilities estimated from the individual independent simulations. For each interface of the individual simulations, the sampling was run until a desired number of successful crossings was reached. Simulations from the same ensemble have the same number of successful crossings for each interface, which can be obtained as the number of crossings shown in the table divided by the number of independent simulations (up to small differences of few extra crossings for some simulations in tables S-3, S-4, S-5, and S-7 which resulted from the implementation details for a parallel architecture).

The description of the MD algorithm used for the sampling is provided in the methods section of the main text.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.40\,(0.03)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	9000	1359767	0.0066 (0.0004)
λ_1^2	9022	78126	0.115 (0.01)
λ_2^3	900	907	0.992 (0.008)

TABLE S-4. FFS results for a 6-nucleotide toehold at the 5′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.28\,(0.03)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	9000	1520553	0.006 (0.001)
λ_1^2	9015	115774	0.0779 (0.004)
λ_2^3	918	1022	0.898 (0.003)

TABLE S-5. FFS results for a 5-nucleotide toehold at the 3′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.30\,(0.02)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	9000	1363822	0.0066 (0.0006)
λ_1^2	9000	92789	0.1 (0.05)
λ_2^3	900	946	0.951 (0.008)

TABLE S-6. FFS results for a 5-nucleotide toehold at the 5′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.53\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	12000	2777034	0.0043 (0.0007)
λ_1^2	12001	128781	0.09 (0.04)
λ_2^3	903	2961	0.30 (0.01)

TABLE S-7. FFS results for a 4-nucleotide toehold at the 3′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.16\,(0.02)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	12000	2140159	0.0056 (0.0004)
λ_1^2	12000	80545	0.15 (0.04)
λ_2^3	900	1961	0.46 (0.03)

TABLE S-8. FFS results for a 4-nucleotide toehold at the 5′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.05\,(0.04)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	3432942	0.005 (0.0005)
λ_1^2	18000	288833	0.062 (0.004)
λ_2^3	900	25303	0.036 (0.004)

TABLE S-9. FFS results for a 3-nucleotide toehold at the 5′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	1730	1.06×10^{-7}	
λ	Success	Attempts	Probability
λ_0^1	18000	3432942	0.0056
λ_1^2	18000	288833	0.088
λ_2^3	900	25303	0.34
λ_3^4	900	25303	0.14

TABLE S-10. FFS results for a 3-nucleotide toehold at the 5′ end. The data were obtained from 1 set of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(1.00\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	3344529	0.0054 (0.0005)
λ_1^2	18000	282819	0.06 (0.08)
λ_2^3	900	39767	0.02 (0.01)

TABLE S-11. FFS results for a 3-nucleotide toehold at the 3′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	6000	$(1.03\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	12000	2593563	0.0046 (0.0002)
λ_1^2	12000	339889	0.035 (0.008)
λ_2^3	1000	9717	0.10 (0.01)
λ_3^4	600	2577	0.233 (0.006)

TABLE S-12. FFS results for a 3-nucleotide toehold at the 3′ end. The data were obtained from 2 sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	6000	$(0.9\,(0.003)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	12000	3304853	0.0036 (0.0004)
λ_1^2	12000	64190	0.19 (0.02)
λ_2^3	600	810861	0.0007 (0.003)

TABLE S-13. FFS results for a 2-nucleotide toehold at the 5′ end. The data were obtained from 2 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(0.87\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	3657559	0.0049 (0.0004)
λ_1^2	18000	130705	0.14 (0.19)
λ_2^3	1500	88745	0.0169 (0.004)
λ_3^4	900	15585	0.0577 (0.009)

TABLE S-14. FFS results for a 2-nucleotide toehold at the 5′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(0.78\,(0.02)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	4683647	0.0038 (0.0004)
λ_1^2	18000	219009	0.08 (0.04)
λ_2^3	1500	930597	0.0016 (0.0006)
λ_3^4	900	6234	0.144 (0.006)

TABLE S-15. FFS results for a 2-nucleotide toehold at the 3′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(0.66\,(0.07)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	4727390	0.0038 (0.0005)
λ_1^2	18000	426011	0.042 (0.015)
λ_2^3	1500	579917	0.003 (0.03)
λ_3^4	900	132335	0.007 (0.009)

TABLE S-17. FFS results for a 1-nucleotide toehold at the 5′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	6000	$(0.79\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	12000	2780446	0.004 (0.001)
λ_1^2	12000	426468	0.03 (0.09)
λ_2^3	600	1421539	0.0004 (0.0001)

TABLE S-16. FFS results for a 2-nucleotide toehold at the 3′ end. The data were obtained from 2 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(0.66\,(0.09)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	5017739	0.0036 (0.0005)
λ_1^2	18000	676294	0.03 (0.04)
λ_2^3	900	30649274	0.00003 (0.00003)

TABLE S-18. FFS results for a 1-nucleotide toehold at the 3′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(0.50\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	6935564	0.0026 (0.0002)
λ_1^2	18000	2165417	0.01 (0.02)
λ_2^3	1500	2319850	0.00065 (0.00002)
λ_3^4	900	18734	0.05 (0.03)

TABLE S-19. FFS results for a 1-nucleotide toehold at the 3′ end. The data were obtained from 3 independent sets of simulations.

λ	Crossings	Flux	
λ_{-1}^0	9000	$(0.77\,(0.02)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_0^1	18000	4119731	0.004 (0.002)
λ_1^2	1800	251974	0.07 (0.02)
λ_2^3	1500	5547	0.27 (0.06)
λ_3^4	900	3657	0.25 (0.01)

TABLE S-20. FFS results for a 3-nucleotide toehold at the 3′ end. The data were obtained from 3 independent sets of simulations at 17 °C.

λ	Crossings	Flux	
λ_{-1}^{0}	9000	$(0.87\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_{0}^{1}	18000	3843169	0.005 (0.001)
λ_{1}^{2}	1800	251058	0.07 (0.02)
λ_{2}^{3}	1500	9490	0.16 (0.01)
λ_{3}^{4}	900	3795	0.24 (0.01)

TABLE S-21. FFS results for a 3-nucleotide toehold at the $3'$ end. The data were obtained from 3 independent sets of simulations at 27 °C.

λ	Crossings	Flux	
λ_{-1}^{0}	9000	$(1.10\,(0.03)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ_{0}^{1}	18000	3766518	0.005 (0.001)
λ_{1}^{2}	1800	634145	0.028 (0.01)
λ_{2}^{3}	1500	19227	0.078 (0.03)
λ_{3}^{4}	900	3599	0.25 (0.01)

TABLE S-22. FFS results for a 3-nucleotide toehold at the $3'$ end. The data were obtained from 3 independent sets of simulations at 47 °C.

[1] N. Srinivas et al., Nucleic Acids Res. **41**, 10641 (2013).
[2] R. J. Allen, C. Valeriani, and P. R. ten Wolde, J. Phys.: Condens. Matter **21**, 463102 (2009).
[3] T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye, and A. A. Louis, Nucleic Acids Res. **41**, 8886 (2013).
[4] J. S. Schreck et al., arXiv preprint arXiv:1408.4401 (2014).