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

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

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

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

where $k_{\rm off}$ 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 $p_{\rm in}$. We can hence approximate $p_{\rm bm|toe}$ as

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

from which we obtain

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

From Eqs. S4 and S1 we hence obtain

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

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

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

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

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

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

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 ($Q = Q_{\rm max}$, where $Q_{\rm max}$ 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 ($Q = Q_{\rm max}$) as

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

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^0_{-1}	9000	$(0.87\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ^1_0	18000	3657559	0.0049 (0.0004)
λ^2_1	18000	130705	0.14 (0.19)
λ^3_2	1500	88745	0.0169 (0.004)
λ^4_3	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^0_{-1}	9000	$(0.78\,(0.02)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ^1_0	18000	4683647	0.0038 (0.0004)
λ^2_1	18000	219009	0.08 (0.04)
λ^3_2	1500	930597	0.0016 (0.0006)
λ^4_3	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^0_{-1}	9000	$(0.66\,(0.07)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ^1_0	18000	4727390	0.0038 (0.0005)
λ^2_1	18000	426011	0.042 (0.015)
λ^3_2	1500	579917	0.003 (0.03)
λ^4_3	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^0_{-1}	6000	$(0.79\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ^1_0	12000	2780446	0.004 (0.001)
λ^2_1	12000	426468	0.03 (0.09)
λ^3_2	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^0_{-1}	9000	$(0.66\,(0.09)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ^1_0	18000	5017739	0.0036 (0.0005)
λ^2_1	18000	676294	0.03 (0.04)
λ^3_2	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^0_{-1}	9000	$(0.50\,(0.01)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ^1_0	18000	6935564	0.0026 (0.0002)
λ^2_1	18000	2165417	0.01 (0.02)
λ^3_2	1500	2319850	0.00065 (0.00002)
λ^4_3	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^0_{-1}	9000	$(0.77\,(0.02)) \times 10^{-7}$	
λ	Success	Attempts	Probability
λ^1_0	18000	4119731	0.004 (0.002)
λ^2_1	1800	251974	0.07 (0.02)
λ^3_2	1500	5547	0.27 (0.06)
λ^4_3	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).