Probabilistic switching circuits in DNA
Supporting information

Daniel Wilhelm, Jehoshua Bruck and Lulu Qian
*correspondence to: luluqian@caltech.edu

Contents

1 Materials and methods

1.1 DNA sequence design

All domain sequences (Table S1) were generated with a three letter code (A, C and T), using the method described in ref. [1]. Sequences of strands were simply generated by composing the domains together (Table S2). A one-nucleotide clamp, complementary to the first nucleotide in the tail of the gate species, was used in all gate bottom strands to reduce undesired gate-gate interactions [1].

1.2 DNA oligonucleotide synthesis

DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT). The gate strands were purchased unpurified (standard desalting). The input and reporter strands were purchased purified (HPLC). All strands were purchased as dry powder, and stored at 100 μM in Milli-Q water (Millipore) at 4 °C.

1.3 Annealing protocol and gel purification

Gate species were annealed at 20 μM, with equal stoichiometry of top and bottom strands. Reporter species were annealed at 20 μM with a 20% excess of top strands. All DNA complexes were annealed in 1× TE buffer with 12.5 mM Mg^{2+}. Annealing was performed in a thermal cycler (Eppendorf), first heating up to 95 °C for 2 minutes, and then slowly cooling down to 20 °C at the rate of 6 sec per 0.1 °C. All annealed complexes were stored at 4 °C. After annealing, the gate species were purified using 15% polyacrylamide gel electrophoresis (PAGE).

1.4 Fluorescence spectroscopy

Fluorescence kinetics data were collected every minute in a spectrofluorimeter (Fluorolog-3, Horiba). Experiments were performed with 1.5 mL reaction mixture per cuvette, in fluorescence cuvettes (Hellma #119.004F-QS) at 25 °C. The excitation/emission wavelengths were set to 588/608 nm for ROX and 524/541 nm for TET. Both excitation and emission bandwidths were set to 2 nm, and the integration time was 10 seconds for all experiments.

1.5 Modeling and simulations

Simulations were performed in Mathematica, using a CRNSimulator package [2] that converts a set of chemical reactions to ordinary differential equations (ODEs) and solves the equations to produce mass-action kinetics of the given reactions. For example, the simulations shown in Fig. 2C were performed using the four listed reactions, with a common rate constant $k = 6.5 \times 10^5$ /M/s specified in the figure caption, the initial concentrations of all signal, gate and reporter species specified in Fig. 1D ($I_3 = 1\times$, $Gate(I_5 \rightarrow I_6) = 1\times$, $Gate(I_5 \rightarrow \emptyset) = 1\times$, $Rep_6 = 2\times$), the two listed initial concentrations of switching signal S_1 ($Gate(I_3 \rightarrow I_5) = 0\times$ and $1\times$), and the labeled standard concentration $1\times = 50$ nM. The time delay for adding the input and deterministic switching signals (i.e. the gap in the collected data points) varied among experiments, due to the difference in the number of molecules added and the variance in the pipetting speed, and was modeled as a manually tuned variable. For example, the delay time in the simulations shown Fig. 2C was tuned to be 6.5 min to fit the data.

2 Supplementary data, analysis and design

2.1 Toehold problem

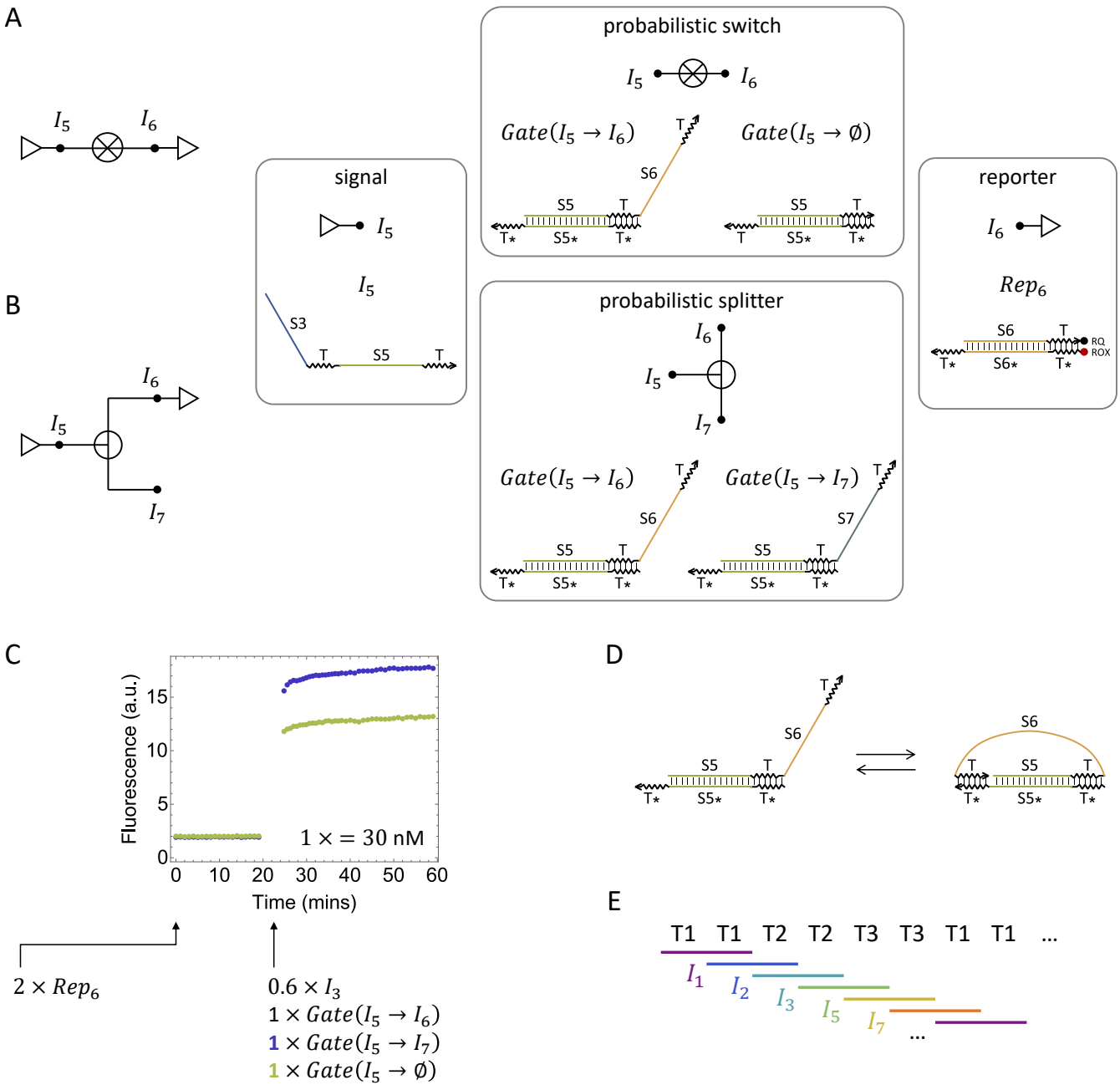


Fig. S1. Toehold problem. Circuit diagram and DNA species of (*A*) a $1/2$ probabilistic switch (pswitch), and (*B*) a $1/2$ probabilistic splitter. (*C*) Fluorescence kinetics experiments of the pswitch and splitter. (*D*) A hypothesis for why the pswitch yielded less output signal than the splitter: The uncovered toehold in the tail of the gate species $Gate(I_5 \rightarrow I_6)$ binds to the complementary toehold in the gate bottom strand and forms a loop structure, inhibiting the gate species from interacting with the input signal, thus resulting in a slower reaction rate compared to the gate species without a tail — $Gate(I_5 \rightarrow \emptyset)$. (*E*) A solution using three distinct toeholds, and an example toehold assignment for the three-bit UPG.

2.2 Data normalization

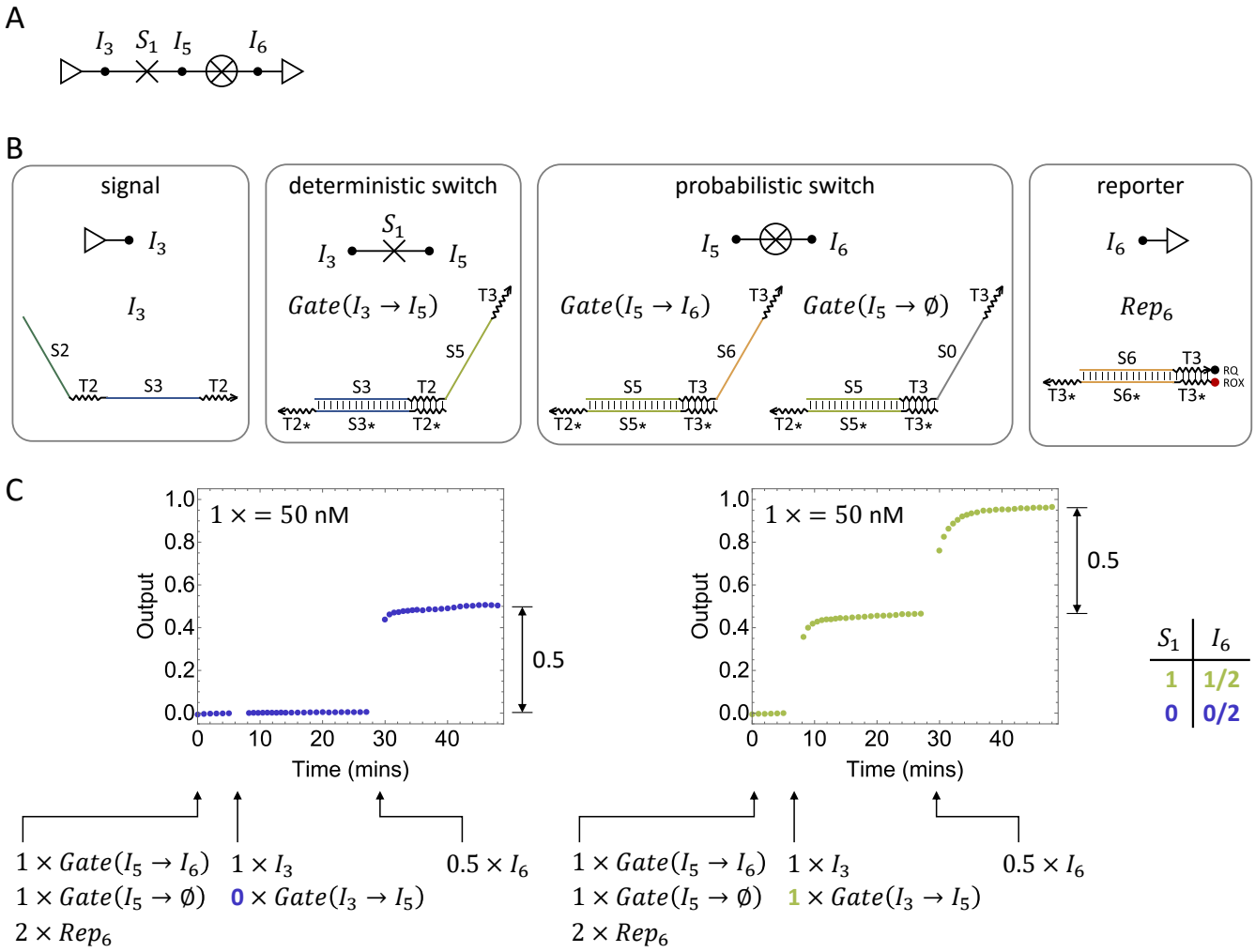


Fig. S2. Data normalization. (*A*) Circuit diagram, (*B*) DNA species, and (*C*) fluorescence kinetics experiments of a one-bit universal probability generator. At the end of each experiment, $0.5\times$ output signal I_6 was introduced to trigger a direct fluorescence signal change, which we refer to as a post-experiment triggering step. The data was then normalized using the average of the first five data points as 0 and two times the difference between the average of the last five data points before and after post-experiment triggering as 1.

2.3 A two-bit universal probability generator

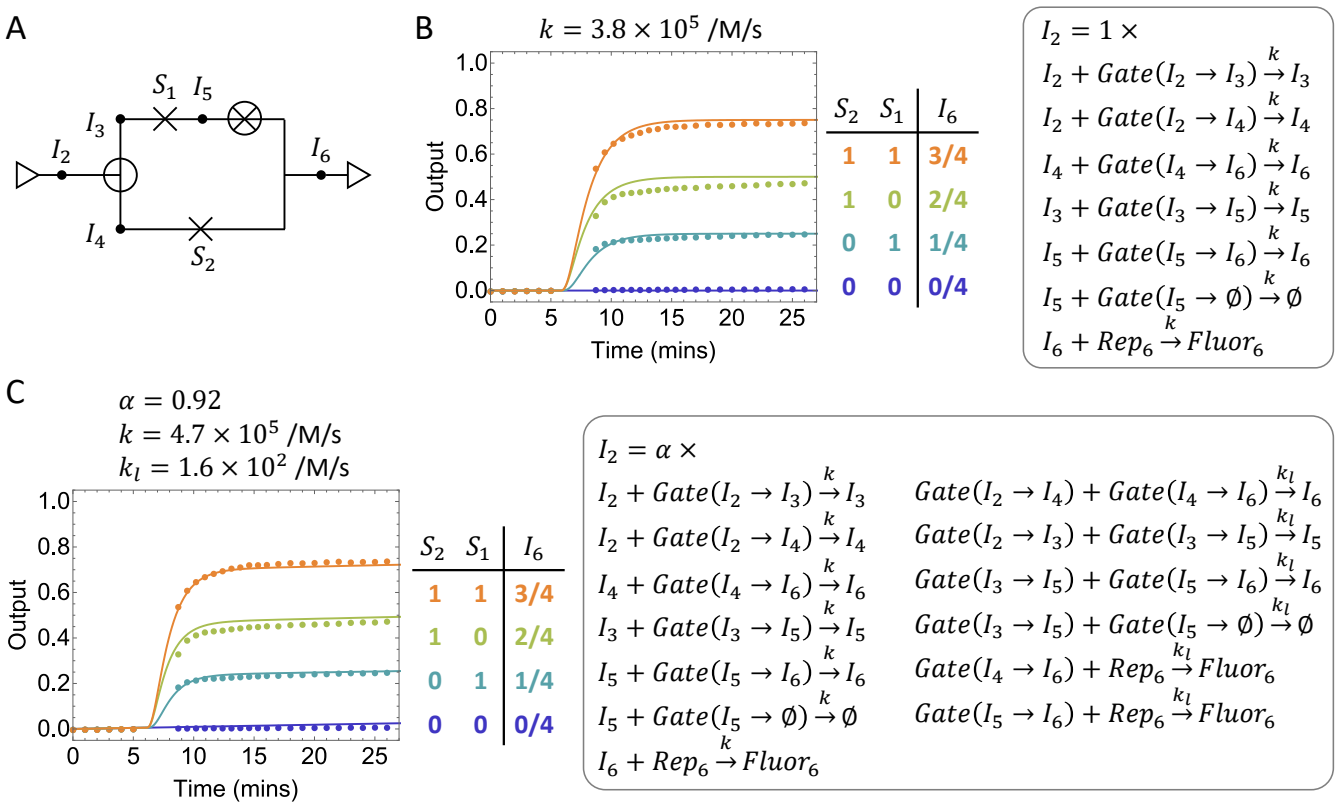


Fig. S3. Simulations of the two-bit universal probability generator. (*A*) Circuit diagram. (*B*) Simulations of the desired reactions overlaid with fluorescence kinetics data. A single rate constant $k = 3.8 \times 10^5$ /M/s provided a good fit to the data. (*C*) Simulations of the desired reactions and leak reactions overlaid with fluorescence kinetics data. Leak reactions are zero-toehold strand displacement reactions between an upstream and downstream gate species (or between an upstream gate and a downstream reporter), initiated by DNA blunt end stacking (see supplementary note S8 of ref. [1]). Because the nominal concentration of a DNA molecule can be higher than its effective concentration [3], an additional parameter α in the model was introduced to allow an up to 10% inaccuracy of the input concentration ($0.9 \leq \alpha \leq 1$), taking into consideration the signal loss caused by synthesis errors in the DNA strands. $\alpha = 0.92$, $k = 4.7 \times 10^5$ /M/s (rate constant of the desired reactions), and $k_l = 1.6 \times 10^2$ /M/s (rate constant of the leak reactions) provided a good fit to the data.

2.4 Effective concentration problem

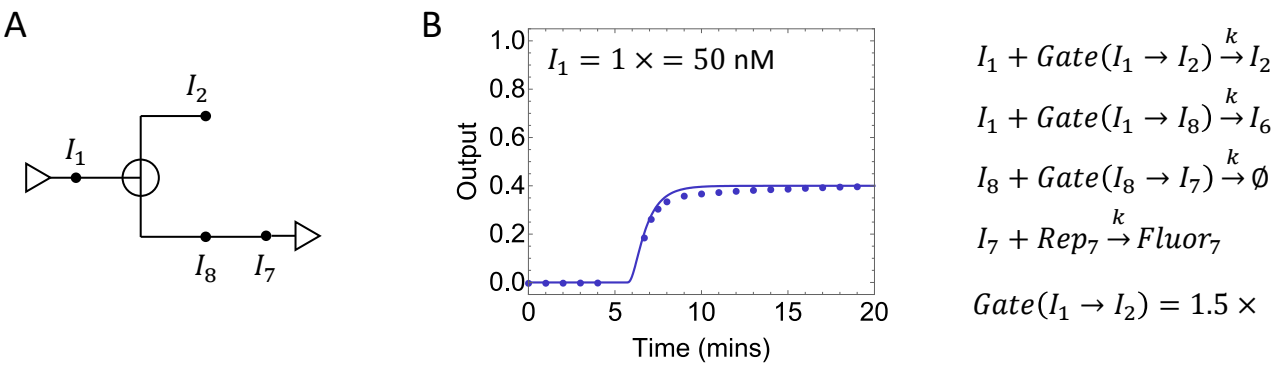


Fig. S4. Effective concentration problem. (*A*) Circuit diagram, (*B*) simulation and fluorescence kinetics experiment of a splitter in the three-bit UPG. The splitter yielded roughly 0.4 instead of the desired 0.5 output. We hypothesized that the effective concentration of $Gate(I_1 \rightarrow I_2)$ was 50% higher than that of $Gate(I_1 \rightarrow I_8)$, as used in the simulation.

2.5 A three-bit universal probability generator

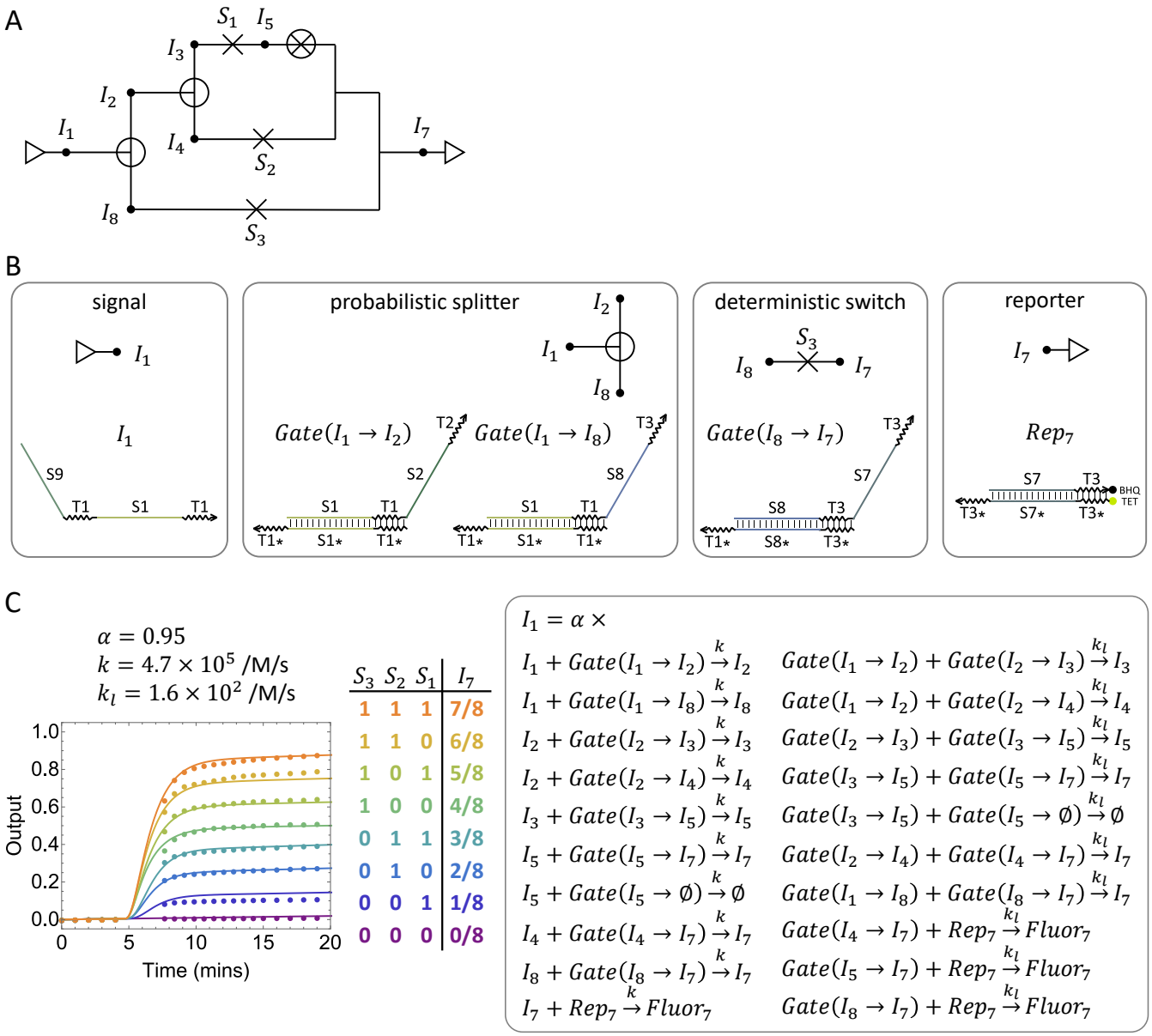
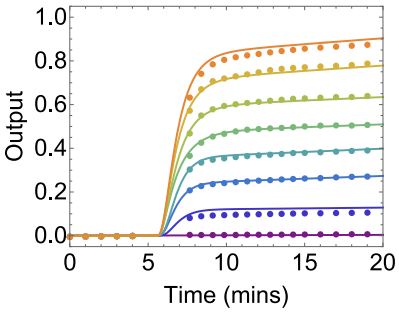


Fig. S5. A three-bit universal probability generator. (*A*) Circuit diagram. (*B*) DNA species in addition to those used in the two-bit UPG (shown in Fig. 2B and E). (*C*) Simulations of the desired reactions and leak reactions overlaid with fluorescence kinetics data. The same values of k and k_l from modeling the two-bit UPG were directly applied in the simulations. To better fit the data, α was adjusted to 0.95, since the input signal here (I_1) is different from that in the two-bit UPG (I_2) and can have a different effective concentration.

$\alpha = 0.93$

$k_1 = 5.8 \times 10^5$ /M/s
$k_2 = 1.0 \times 10^6$ /M/s
$k_3 = 1.0 \times 10^6$ /M/s
$k_{l2} = 1.6 \times 10^2$ /M/s
$k_{l3} = 1.6 \times 10^2$ /M/s
$k_{l4} = 1.3 \times 10^3$ /M/s
$k_{l5} = 1.6 \times 10^2$ /M/s
$k_{l8} = 3.2 \times 10^3$ /M/s
$k_{lr} = 3.2 \times 10^1$ /M/s



S_3	S_2	S_1	I_7
1	1	1	7/8
1	1	0	6/8
1	0	1	5/8
1	0	0	4/8
0	1	1	3/8
0	1	0	2/8
0	0	1	1/8
0	0	0	0/8

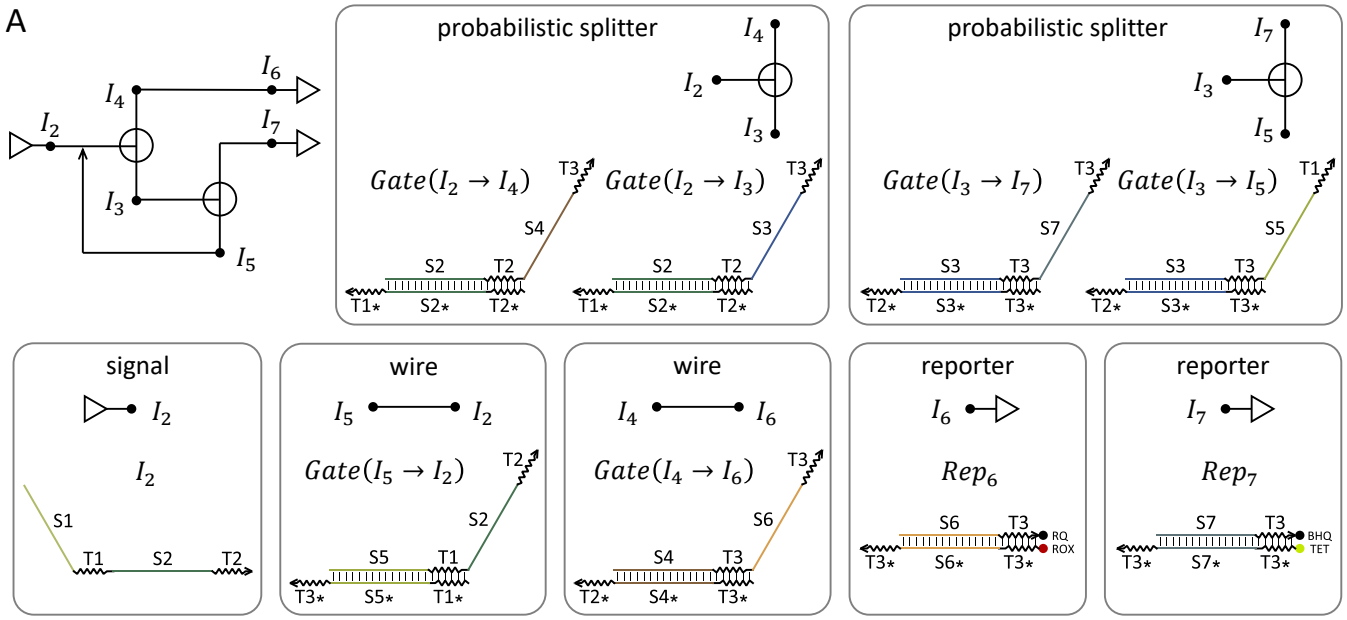
$I_1 = \alpha \times$

$I_1 + Gate(I_1 \to I_2) \xrightarrow{k_1} I_2$
$I_1 + Gate(I_1 \to I_8) \xrightarrow{k_1} I_8$
$I_2 + Gate(I_2 \to I_3) \xrightarrow{k_1} I_3$
$I_2 + Gate(I_2 \to I_4) \xrightarrow{k_1} I_4$
$I_3 + Gate(I_3 \to I_5) \xrightarrow{k_2} I_5$
$I_5 + Gate(I_5 \to I_7) \xrightarrow{k_2} I_7$
$I_5 + Gate(I_5 \to \emptyset) \xrightarrow{k_2} \emptyset$
$I_4 + Gate(I_4 \to I_7) \xrightarrow{k_2} I_7$
$I_8 + Gate(I_8 \to I_7) \xrightarrow{k_1} I_7$
$I_7 + Rep_7 \xrightarrow{k_3} Fluor_7$

$Gate(I_1 \to I_2) + Gate(I_2 \to I_3) \xrightarrow{k_{l2}} I_3$
$Gate(I_1 \to I_2) + Gate(I_2 \to I_4) \xrightarrow{k_{l2}} I_4$
$Gate(I_2 \to I_3) + Gate(I_3 \to I_5) \xrightarrow{k_{l3}} I_5$
$Gate(I_3 \to I_5) + Gate(I_5 \to I_7) \xrightarrow{k_{l5}} I_7$
$Gate(I_3 \to I_5) + Gate(I_5 \to \emptyset) \xrightarrow{k_{l5}} \emptyset$
$Gate(I_2 \to I_4) + Gate(I_4 \to I_7) \xrightarrow{k_{l4}} I_7$
$Gate(I_1 \to I_8) + Gate(I_8 \to I_7) \xrightarrow{k_{l8}} I_7$
$Gate(I_4 \to I_7) + Rep_7 \xrightarrow{k_{lr}} Fluor_7$
$Gate(I_5 \to I_7) + Rep_7 \xrightarrow{k_{lr}} Fluor_7$
$Gate(I_8 \to I_7) + Rep_7 \xrightarrow{k_{lr}} Fluor_7$

Fig. S6. Simulations of the three-bit universal probability generator. Simulations of the desired reactions and leak reactions are overlaid with fluorescence kinetics data. Three different rate constants (k_1 to k_3) are used for the three types of desired reactions that each has a different toehold (T1 to T3). Five different rate constants (k_{l2}, k_{l3}, k_{l4}, k_{l5}, and k_{l8}) are used for the five types of leak reactions, between two gate species, that each has a different branch migration domain (S2, S3, S4, S5 and S8). Another rate constant is used for leak reactions between a gate and the reporter. Keeping as many rate constants the same as in the previous simulations (Fig. S4C), but allowing rate adjustments that led to a better fit to the data, the simulations quantitatively agreed with the experiments. It is reasonable that the rate constant of the leak reactions between a gate and a reporter ($k_{lr} = 3.2 \times 10^1$ /M/s) is smaller than that of the leak reactions between two gate species ($k_{l2}, \cdots, k_{l8} \geq 1.6 \times 10^2$ /M/s), because the reporter strands were purchased purified and thus the reporters may have fewer synthesis errors compared to the gate species.

2.6 A feedback circuit



$\alpha = 0.94$
$k_1 = 5.8 \times 10^5$ /M/s
$k_2 = 1.0 \times 10^6$ /M/s
$k_3 = 1.0 \times 10^6$ /M/s
$k_{l2} = 1.6 \times 10^2$ /M/s
$k_{l3} = 2.0 \times 10^3$ /M/s
$k_{l4} = 1.3 \times 10^3$ /M/s
$k_{l5} = 1.6 \times 10^2$ /M/s
$k_{lr} = 3.2 \times 10^1$ /M/s

$I_6 = \frac{1}{2}$, $I_7 = \frac{1}{4}$

$I_6 = \frac{2}{3}$, $I_7 = \frac{1}{3}$

$I_2 = \alpha \times$

$I_2 + Gate(I_2 \to I_3) \xrightarrow{k_1} I_3$
$I_2 + Gate(I_2 \to I_4) \xrightarrow{k_1} I_4$
$I_3 + Gate(I_3 \to I_5) \xrightarrow{k_2} I_5$
$I_3 + Gate(I_3 \to I_7) \xrightarrow{k_2} I_7$
$I_4 + Gate(I_4 \to I_6) \xrightarrow{k_2} I_6$
$I_5 + Gate(I_5 \to I_2) \xrightarrow{k_3} I_2$
$I_6 + Rep_6 \xrightarrow{k_3} Fluor_6$
$I_7 + Rep_7 \xrightarrow{k_3} Fluor_7$

$Gate(I_5 \to I_2) + Gate(I_2 \to I_3) \xrightarrow{k_{l2}} I_3$
$Gate(I_5 \to I_2) + Gate(I_2 \to I_4) \xrightarrow{k_{l2}} I_4$
$Gate(I_2 \to I_3) + Gate(I_3 \to I_5) \xrightarrow{k_{l3}} I_5$
$Gate(I_2 \to I_3) + Gate(I_3 \to I_7) \xrightarrow{k_{l3}} I_7$
$Gate(I_2 \to I_4) + Gate(I_4 \to I_6) \xrightarrow{k_{l4}} I_6$
$Gate(I_3 \to I_5) + Gate(I_5 \to I_2) \xrightarrow{k_{l5}} I_2$
$Gate(I_4 \to I_6) + Rep_6 \xrightarrow{k_{lr}} Fluor_6$
$Gate(I_3 \to I_7) + Rep_7 \xrightarrow{k_{lr}} Fluor_7$

Fig. S7. A feedback circuit. (*A*) Circuit diagram and DNA species. (*B*) Simulations of the desired reactions and leak reactions overlaid with fluorescence kinetics data. Three different rate constants (k_1 to k_3) are used for the three types of desired reactions that each has a different toehold (T1 to T3). Four different rate constants (k_{l2} to k_{l5}) are used for the four types of leak reactions, between two gate species, that each has a different branch migration domain (S2 to S5). Another rate constant is used for leak reactions between a gate and a reporter. Using the same rate constants in the previous simulations (Fig. S5), except for k_{l3}, the simulations quantitatively agreed with the experiments. It is reasonable that k_{l3} is different from that in the three-bit UPG, since wiring of the feedback circuit resulted in a different toehold adjacent to S3.

2.7 More robust pswitch and splitter designs

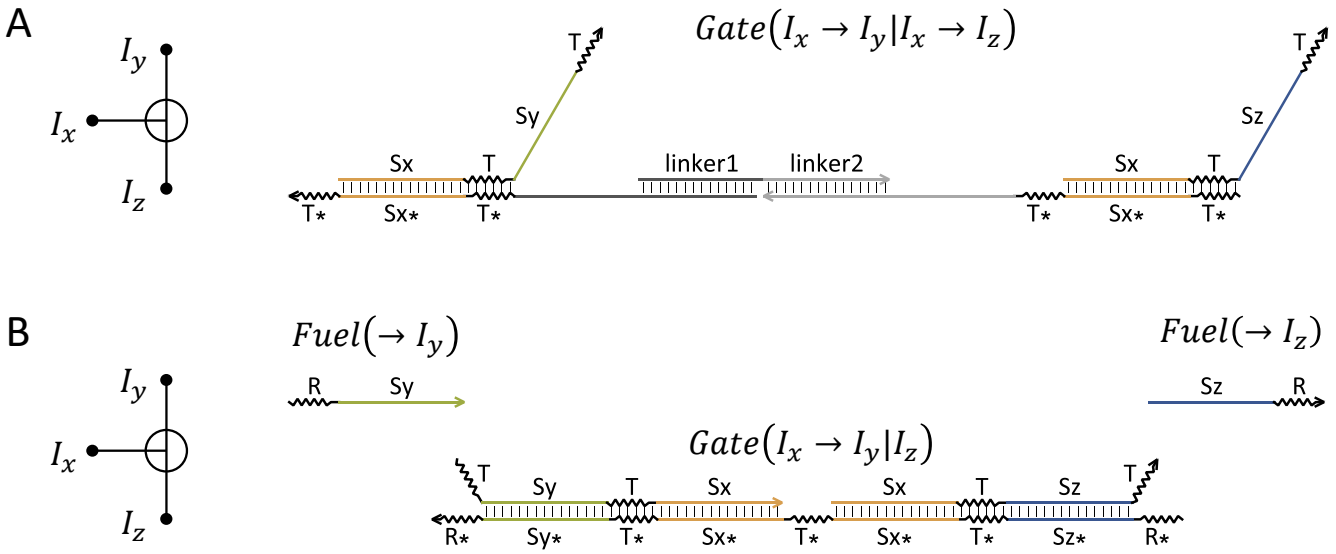


Fig. S8. More robust pswitch and splitter designs. (*A*) To reduce the effect of concentration and pipetting errors, two gate species in a $1/2$ pswitch or splitter could be linked together by another strand that contains two unique linker domains. To avoid two copies of the same signal strand binding to the two gate bottom strands in one dual-gate complex, the two gate species could be annealed separately and then incubated together with the linker strand. Complexes that include both gates could then be gel purified. This way, the concentration of the gates would have to be equal. (*B*) Alternatively, two gate species could be linked together to share one common toehold in the middle, forcing the input strand to interact with one or the other gate but not both. In this design, the Sy and Sz domains in the joint gate should not be simultaneously exposed with a toehold domain, otherwise they would directly interact with a downstream joint gate without the input strand being present. Thus, the bottom strand of the gate is extended to cover up the Sy and Sz domains, allowing reversible strand displacement reactions with two fuel strands, initiated by a different toehold R. Only when the input strand is present, cooperative hybridization [4] involving the input and one fuel strand should take place and stochastically produce one output signal.

2.8 A more composable deterministic switch design

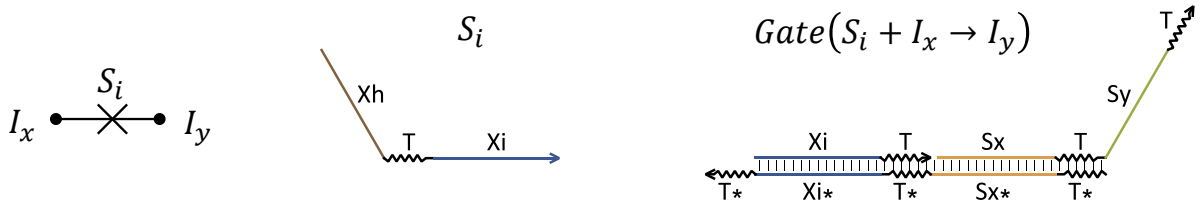


Fig. S9. A more composable deterministic switch design. To allow the probabilistic switching circuits to be composed together with other kinds of circuits such as DNA-based logic circuits, the gate species implementing deterministic switches could be extended to interact with a single-stranded switching signal.

3 DNA sequences

3.1 Domain sequences

Table S1: Domain sequences.

Domain name	Sequence	Complementary domain	Complementary sequence
T1	CTTACC	T1*	GGTAAG
T2	ACACAC	T2*	GTGTGT
T3	CTCCTC	T3*	GAGGAG
S0	AAAAAAAAAAAAAAA		
S1	CAAAATCCAAAACCT	S1*	AGGTTTTGGATTTTG
S2	CATCCATTCCACTAT	S2*	ATAGTGGAATGGATG
S3	CACCATCAAATAACT	S3*	AGTTATTTGATGGTG
S4	CTCAATAACATCTCT	S4*	AGAGATGTTATTGAG
S5	CCAAACAAAACCTAT	S5*	ATAGGTTTTGTTTGG
S6	AACCACCAAACTTAT		
S7	CCTAACACAATCACT		
S8	CACCCTAAAATCTAT	S8*	ATAGATTTTAGGGTG
S9	TCAAAACCAACTACT		

3.2 Strand sequences

Table S2: Strand sequences.

Strand name	Domains	Sequence
I1	S9 T1 S1 T1	TCAAAACCAACTACT CTTACC CAAAATCCAAAACCT CTTACC
I2/G(I1->I2)-t	S1 T1 S2 T2	CAAAATCCAAAACCT CTTACC CATCCATTCCACTAT ACACAC
G(I1->I8)-t	S1 T1 S8 T3	CAAAATCCAAAACCT CTTACC CACCCTAAAATCTAT CTCCTC
G(I1->)-b	G T1* S1* T1*	G GGTAAG AGGTTTTGGATTTTG GGTAAG
I3/G(I2->I3)-t	S2 T2 S3 T2	CATCCATTCCACTAT ACACAC CACCATCAAATAACT ACACAC
GF(I2->I3)-t	S2 T2 S3 T3	CATCCATTCCACTAT ACACAC CACCATCAAATAACT CTCCTC
G(I2->I4)-t	S2 T2 S4 T3	CATCCATTCCACTAT ACACAC CTCAATAACATCTCT CTCCTC
G(I2->)-b	G T2* S2* T1*	G GTGTGT ATAGTGGAATGGATG GGTAAG
G(I3->I5)-t	S3 T2 S5 T3	CACCATCAAATAACT ACACAC CCAAACAAAACCTAT CTCCTC
G(I3->I5)-b	G T2* S3* T2*	G GTGTGT AGTTATTTGATGGTG GTGTGT
GF(I3->I5)-t	S3 T3 S5 T1	CACCATCAAATAACT CTCCTC CCAAACAAAACCTAT CTTACC
G(I3->I7)-t	S3 T3 S7 T3	CACCATCAAATAACT CTCCTC CCTAACACAATCACT CTCCTC
G(I3->)-b	G T3* S3* T2*	G GAGGAG AGTTATTTGATGGTG GTGTGT
G(I4->I6)-t	S4 T3 S6 T3	CTCAATAACATCTCT CTCCTC AACCACCAAACTTAT CTCCTC
G(I4->I6)-b	T T3* S4* T2*	T GAGGAG AGAGATGTTATTGAG GTGTGT
G(I4->I7)-t	S4 T3 S7 T3	CTCAATAACATCTCT CTCCTC CCTAACACAATCACT CTCCTC
G(I4->I7)-b	G T3* S4* T2*	G GAGGAG AGAGATGTTATTGAG GTGTGT
G(I5->I0)-t	S5 T3 S0 T3	CCAAACAAAACCTAT CTCCTC AAAAAAAAAAAAAAA CTCCTC
G(I5->I2)-t	S5 T1 S2 T2	CCAAACAAAACCTAT CTTACC CATCCATTCCACTAT ACACAC
G(I5->I2)-b	G T1* S5* T3*	G GGTAAG ATAGGTTTTGTTTGG GAGGAG
G(I5->I6)-t	S5 T3 S6 T3	CCAAACAAAACCTAT CTCCTC AACCACCAAACTTAT CTCCTC
G(I5->)-b	T T3* S5* T2*	T GAGGAG ATAGGTTTTGTTTGG GTGTGT
G(I5->I7)-t	S5 T3 S7 T3	CCAAACAAAACCTAT CTCCTC CCTAACACAATCACT CTCCTC
G(I5->I7)-b	G T3* S5* T2*	G GAGGAG ATAGGTTTTGTTTGG GTGTGT
G(I8->I7)-t	S8 T3 S7 T3	CACCCTAAAATCTAT CTCCTC CCTAACACAATCACT CTCCTC
G(I8->I7)-b	G T3* S8* T1*	G GAGGAG ATAGATTTTAGGGTG GGTAAG
Rep6-t	S6 T3 RQ	AACCACCAAACTTAT CTCCTC /3IAbRQSp/
Rep6-b	ROX T3* S6* T3*	/56-ROXN/ GAGGAG ATAAGTTTGGTGGTT GAGGAG
Rep7-t	S7 T3 BHQ	CCTAACACAATCACT CTCCTC /3BHQ-1/
Rep7-b	TET T3* S7* T3*	/5TET/ GAGGAG AGTGATTGTGTTAGG GAGGAG

G(Ii->Ij)-b is a gate bottom strand used specifically with G(Ii->Ij)-t to create the gate species. G(Ii->)-b is used in more than one gate species, each with a different signal strand G(Ii->Ij)-t, G(Ii->Ik)-t, etc. GF(Ii->Ij)-t is used in the feedback circuit; it has a different toehold compared to G(Ii->Ij)-t.

References

[1] Lulu Qian and Erik Winfree. Scaling up digital circuit computation with DNA strand displacement cascades. *Science*, 332(6034):1196–1201, 2011.

[2] D. Soloveichik. CRNSimulator. `http://users.ece.utexas.edu/~soloveichik/crnsimulator.html`, 2009.

[3] Anupama J Thubagere, Chris Thachuk, Joseph Berleant, Robert F Johnson, Diana A Ardelean, Kevin M Cherry, and Lulu Qian. Compiler-aided systematic construction of large-scale DNA strand displacement circuits using unpurified components. *Nature Communications*, 8:14373, 2017.

[4] David Yu Zhang. Cooperative hybridization of oligonucleotides. *Journal of the American Chemical Society*, 133(4):1077–1086, 2010.