Effective design principles for leakless strand displacement systems

Boya Wang[a], Chris Thachuk[b], Andrew D. Ellington[c], Erik Winfree[b,d,e], and David Soloveichik[a,1]

[a]Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712; [b]Computer Science, California Institute of Technology, Pasadena, CA 91125; [c]Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712; [d]Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125; and [e]Bioengineering, California Institute of Technology, Pasadena, CA 91125

Edited by Ronald R. Breaker, Yale University, New Haven, CT, and approved November 9, 2018 (received for review April 20, 2018)

Leaks

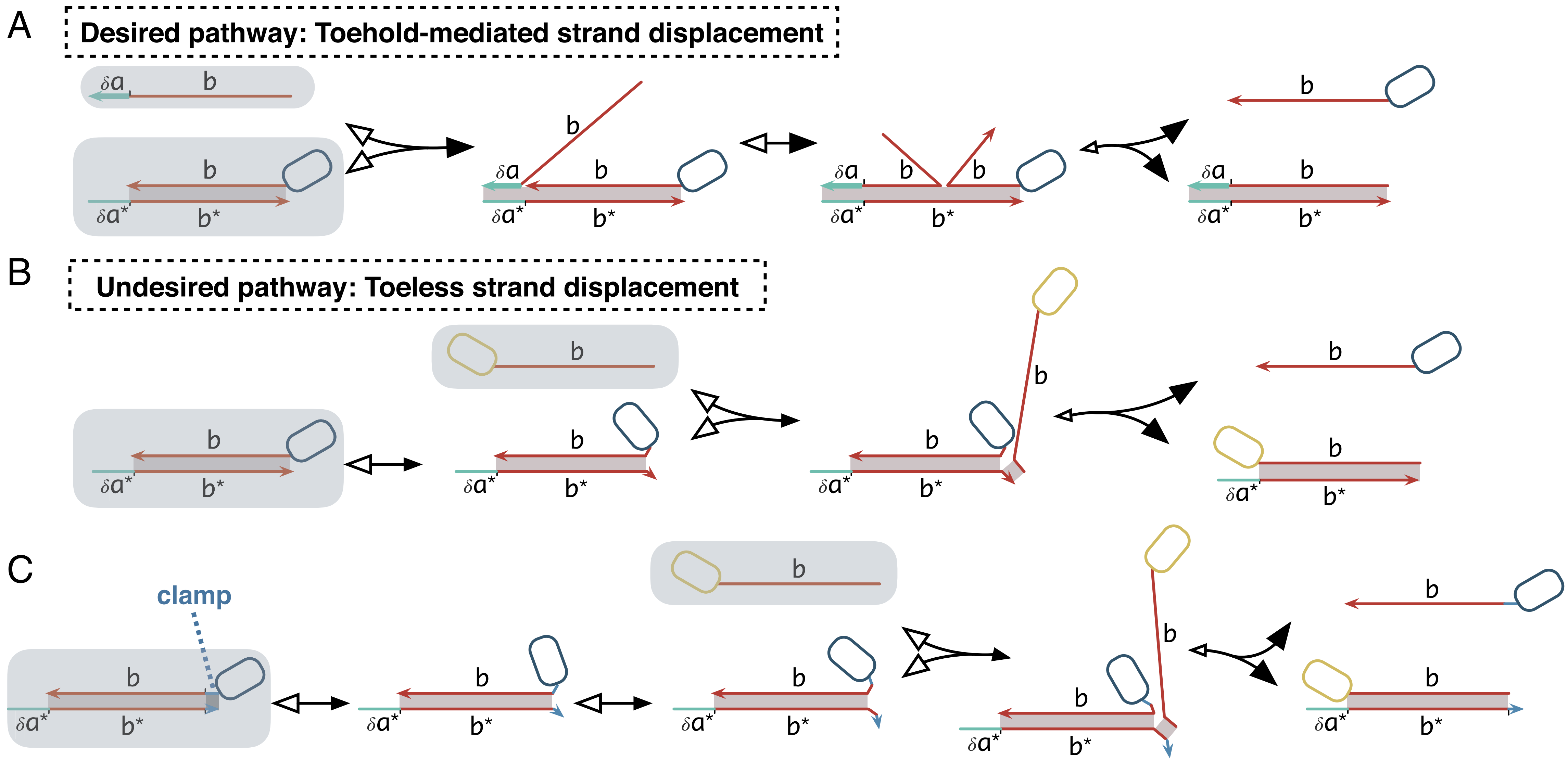


Fig. 1. Desired and leak strand displacement reactions. Letter labels represent domains, which are contiguous bases that logically act as a unit. The domains with complementary sequences are labeled with asterisks. The black arrowheads show the direction of forward reactions, and the white arrowheads show the direction of backward reactions relative to the illustrated reaction pathway. Small arrowheads indicate reactions that are expected to be slow relative to other steps. Shaded background indicates molecular species that are initially present. Toehold domains are labeled by the symbol δ and a thicker strand line. (*A*) Intended strand displacement reactions are initiated by binding of a toehold (domain δa) followed by the displacement of the incumbent strand by the matching portion of the invader strand (domain *b*). The participating strands can have other domains on either side of the involved displacing region; for example, the dark blue rounded rectangle box represents the part of the released strand on the 5′ end not involved in this interaction. (*B*) Leak is hypothesized to be caused by toeless strand displacement reactions, which start from the fraying of the end of a DNA helix where no neighboring base pairs can help stabilize the structure. (Here, both ends of the helix could fray. The figure only shows one scenario.) After they are frayed, the opened nucleotides can be the initiation point for the undesired strand displacement. The invading strand can be a part of a larger complex, which is indicated by the yellow rounded rectangle box. (*C*) Short clamp domains (indicated by a darker shaded helix) help to reduce leak. The toeless displacement on the right side of the helix now requires larger fraying as shown. (Note that the clamp shown does not inhibit toeless displacement on the left side of the helix.)

Leaks in Single Long Domain Design

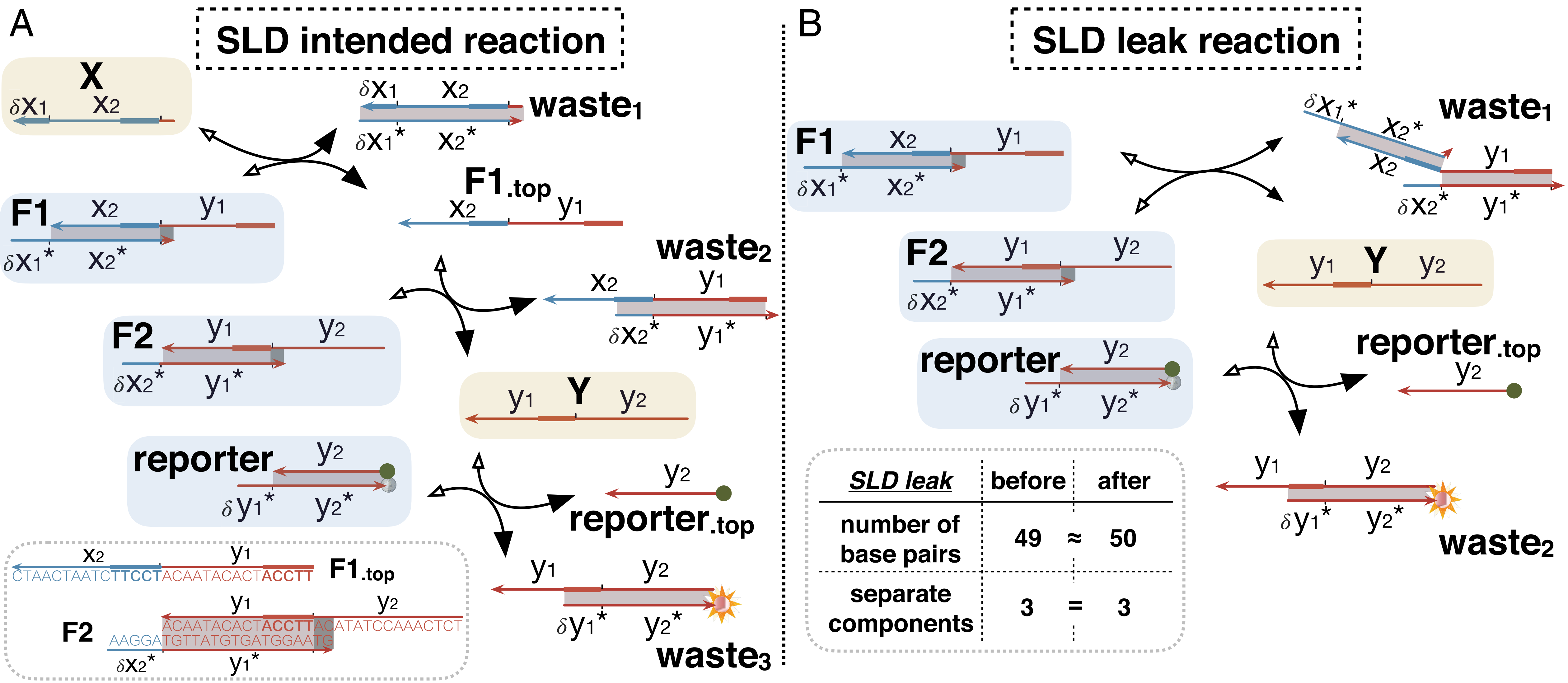


Fig. 2. The typical SLD translator system. The symbol δ denotes a toehold-size subsection of a long domain (e.g., δx_2^* is the toehold-size 5′ end of x_2^*). The faint blue background indicates the fuel and reporter species; the faint yellow background indicates the signal species. Domains that logically belong to input X and output Y are colored blue and red, respectively. (*A*) The intended pathway of the translator $X \rightarrow Y$ that translates strand X to strand Y. Per standard practice, every fuel complex contains a clamp domain to help reduce leak. The input strand X first interacts with the fuel complex *F1*, displacing an intermediate strand $F1_{.top}$ with the toehold domain δx_2 and the long domain y_1 exposed, which then reacts with *F2* releasing the output strand Y. The output strand Y can be detected by a downstream reporter. (*Inset*) The sequences of $F1_{.top}$ and fuel *F2* used in our experiments. (*B*) The leak reaction of the translator in the absence of the input strand X. After the clamp in *F2* opens, the unbound domain y_1 in the fuel complex *F1* can interact with the sequestered domain y_1^* in *F2* through toeless strand displacement and then, produce the output strand Y. (*Inset*) Thermodynamic analysis suggests that there is very little energy difference between the leaked and unleaked states. Each full long domain is 15 bases, domains with the symbol δ and toehold domains are 5 bases, and clamp domains are 2 bases. The change in the number of base pairs due to the net leak reaction is equal to $\{\text{toehold size}\} - 2 \times \{\text{clamp size}\} = 1$. As long as this number is small, it does not play an important role in the thermodynamics of the system. There is no difference in the number of separate components.

Leaks in Double Long Domain Design

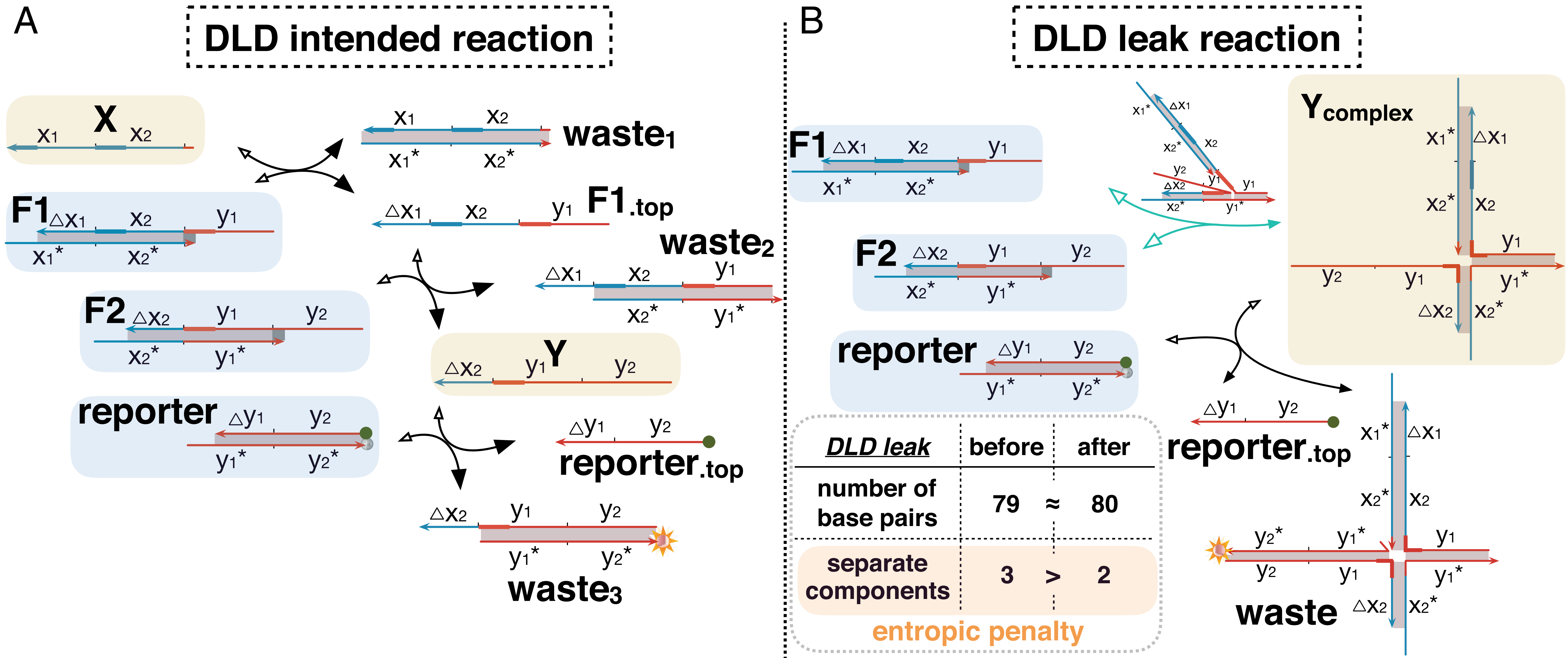
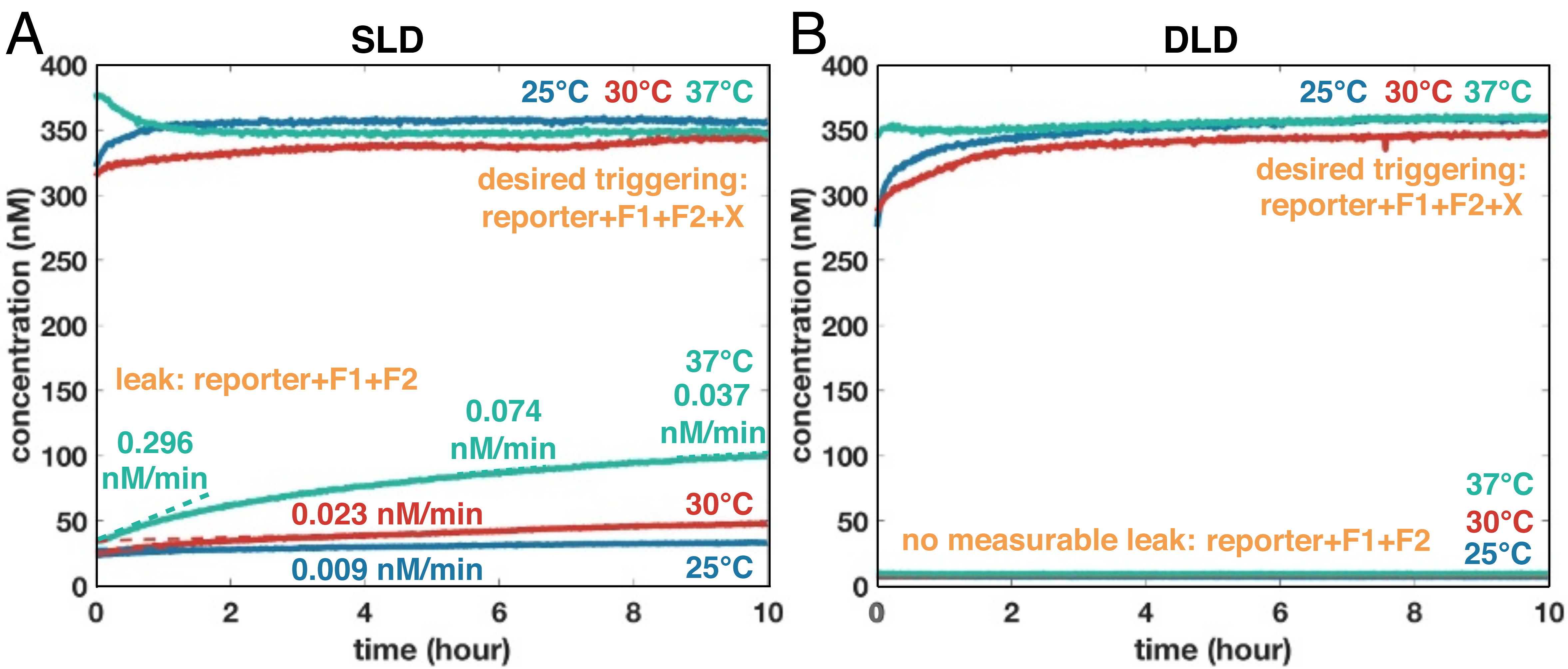
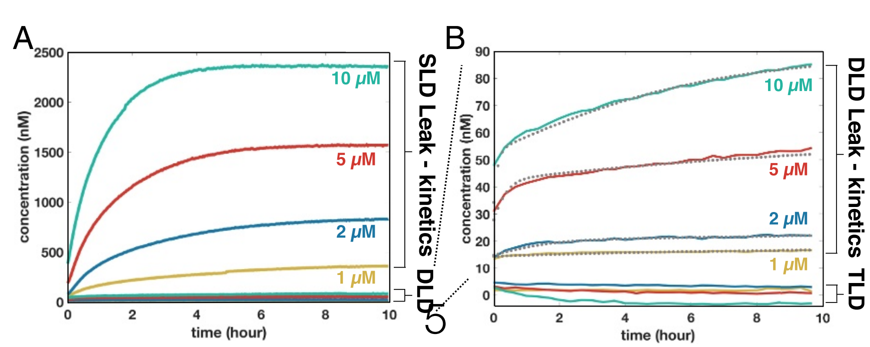


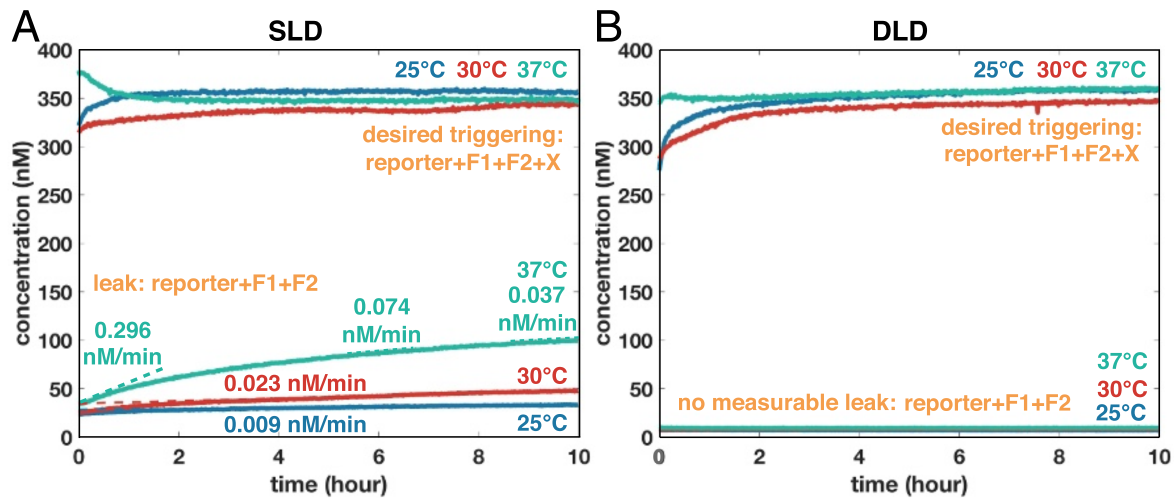
Fig. 3. The DLD translator system. The symbol $\triangle$ denotes an almost full domain, which is a 10-base subsection of a full domain. The faint yellow background indicates the signal species in the intended pathway (input X and output Y) and the species that can trigger the reporter in the leak pathway ($Y_{complex}$). (*A*) The intended pathway of the translator $X \rightarrow Y$ that translates strand X to strand Y, which is similar to that of the SLD translator. (*B*) The proposed leak pathway of the DLD translator in the absence of the input strand X. After the clamp domain in *F2* is open, the unbound domain y_1 in fuel *F1* displaces the y_1 domain in fuel *F2* through toeless strand displacement (the intermediate is shown in the pathway above the arrow), resulting in a short-lived species ($Y_{complex}$). Since the Δx_2 domain is still bound, $Y_{complex}$ can quickly revert back to *F1* and *F2* via a unimolecular reaction (green). False-positive signal requires $Y_{complex}$ to react with the downstream reporter before dissociating. Note that $Y_{complex}$ can also isomerize to other configurations (*SI Appendix*, Fig. S7). (*Inset*) Thermodynamic analysis suggests that the leaked state has higher thermodynamic energy (one unit of entropic penalty) than the unleaked state. The change in the number of base pairs due to the net leak reactions is equal to $\{\text{toehold size}\} - 2 \times \{\text{clamp size}\} = 1$. The number of separate components decreases by one after leak. Under experimental conditions, one entropy penalty is worth roughly 8 bp (in the text), and therefore, the thermodynamics is dominated by the entropic cost of leak.

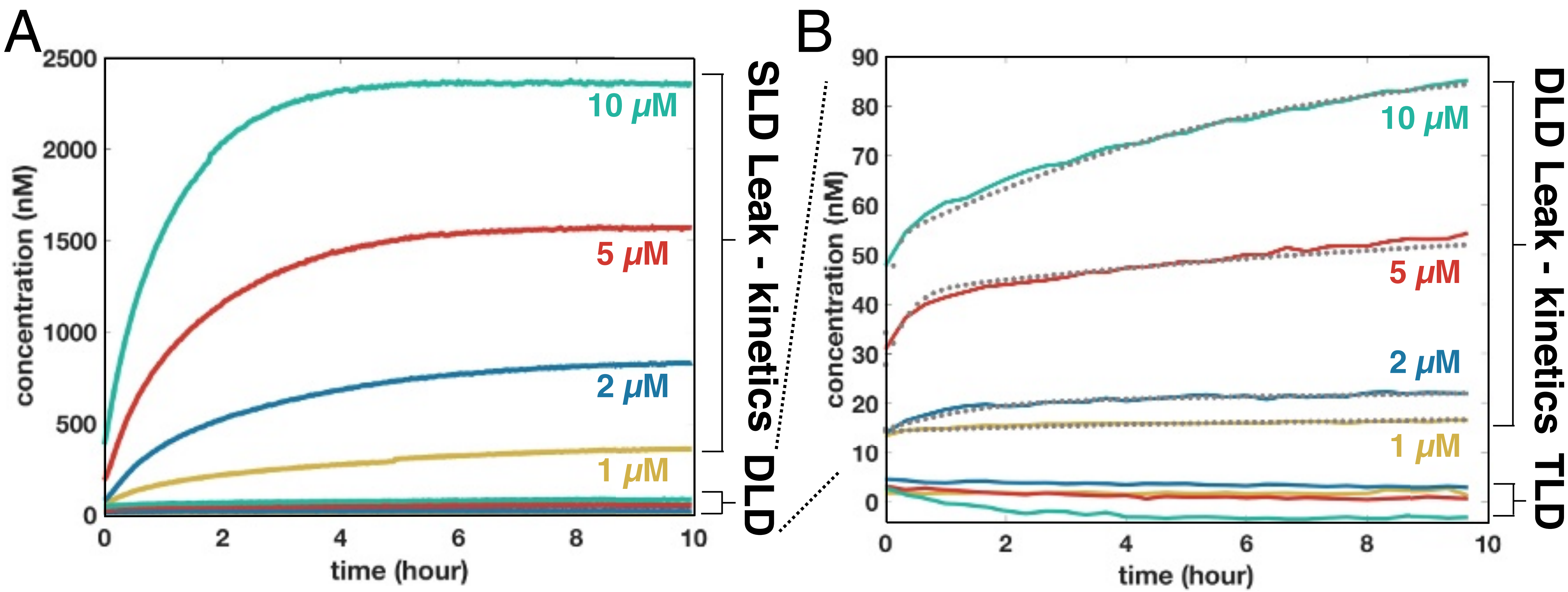
Leaks in SLD vs DLD





Leaks in SLD vs DLD





Leaks in Triple Long Domain Design

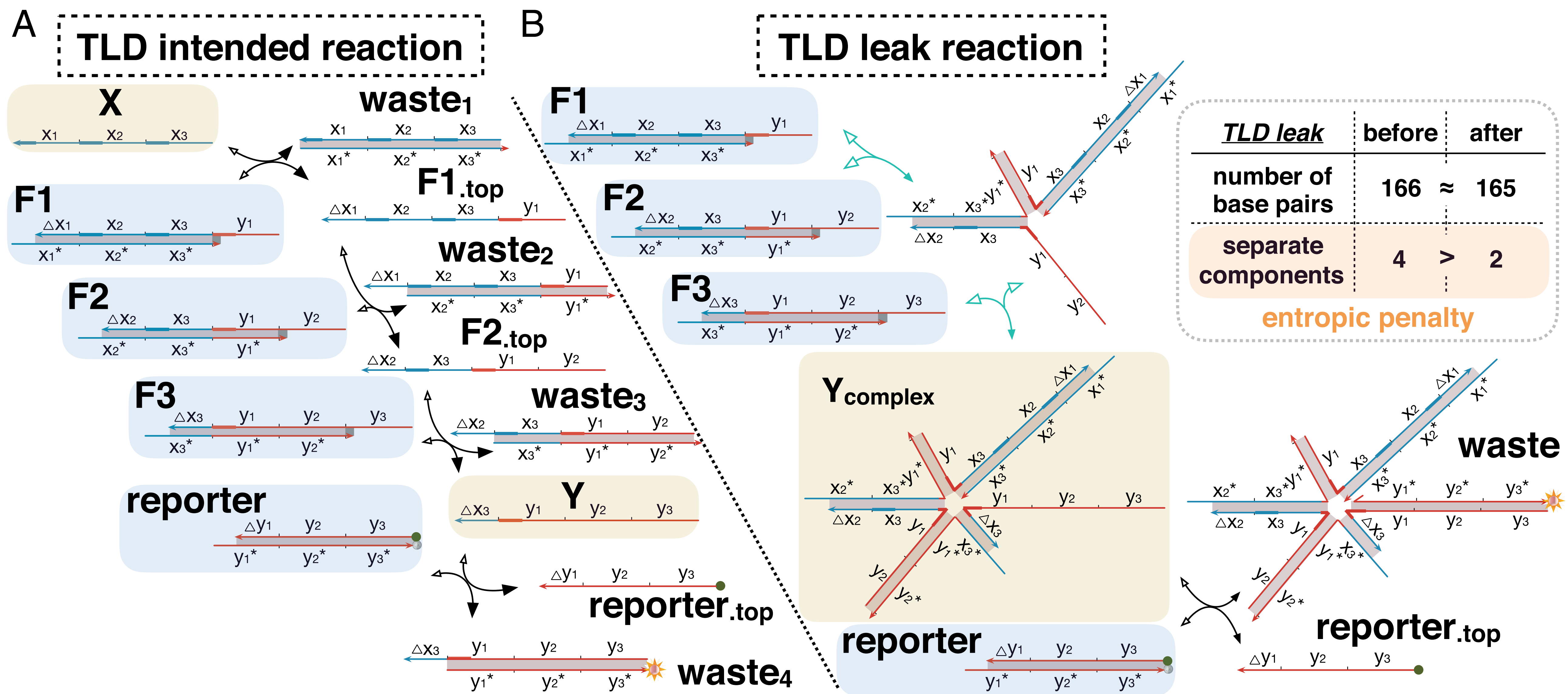


Fig. 4. The TLD translator system. (*A*) The intended pathway of the translator $X \rightarrow Y$ that translates strand X to strand Y. (*B*) One possible leak pathway of the TLD translator in the absence of the input strand X. After the clamp domain in *F2* is open, the unbound domain y_1 in fuel *F1* displaces the y_1 domain in fuel *F2* through toeless strand displacement, resulting in a short-lived four-stranded complex, which could quickly reverse (unimolecular reaction; green) to the original configuration. Then, after the clamp domain in *F3* is open, *F3* reacts with this four-stranded complex and forms a six-stranded species $Y_{complex}$, which can similarly quickly reverse. Finally, the downstream reporter needs to capture this transient six-stranded complex before the reverse reaction occurs. This pathway suggests that the leak rate in the TLD translator should be even slower than that in the DLD translator, since this leak pathway requires a longer series of slow reactions to occur, and each transient in the leak pathway is more likely to quickly reverse toward the nonleak state. (*Inset*) Thermodynamic analysis shows that leak has an energy penalty of two units of entropy, which is one unit of entropy more than in the DLD design. The change in the number of base pairs due to the net leak reactions is equal to $\{\text{toehold size}\} - 3 \times \{\text{clamp size}\} = -1$. The number of separate components decreases by two after leak.

DLD cascade: Wires

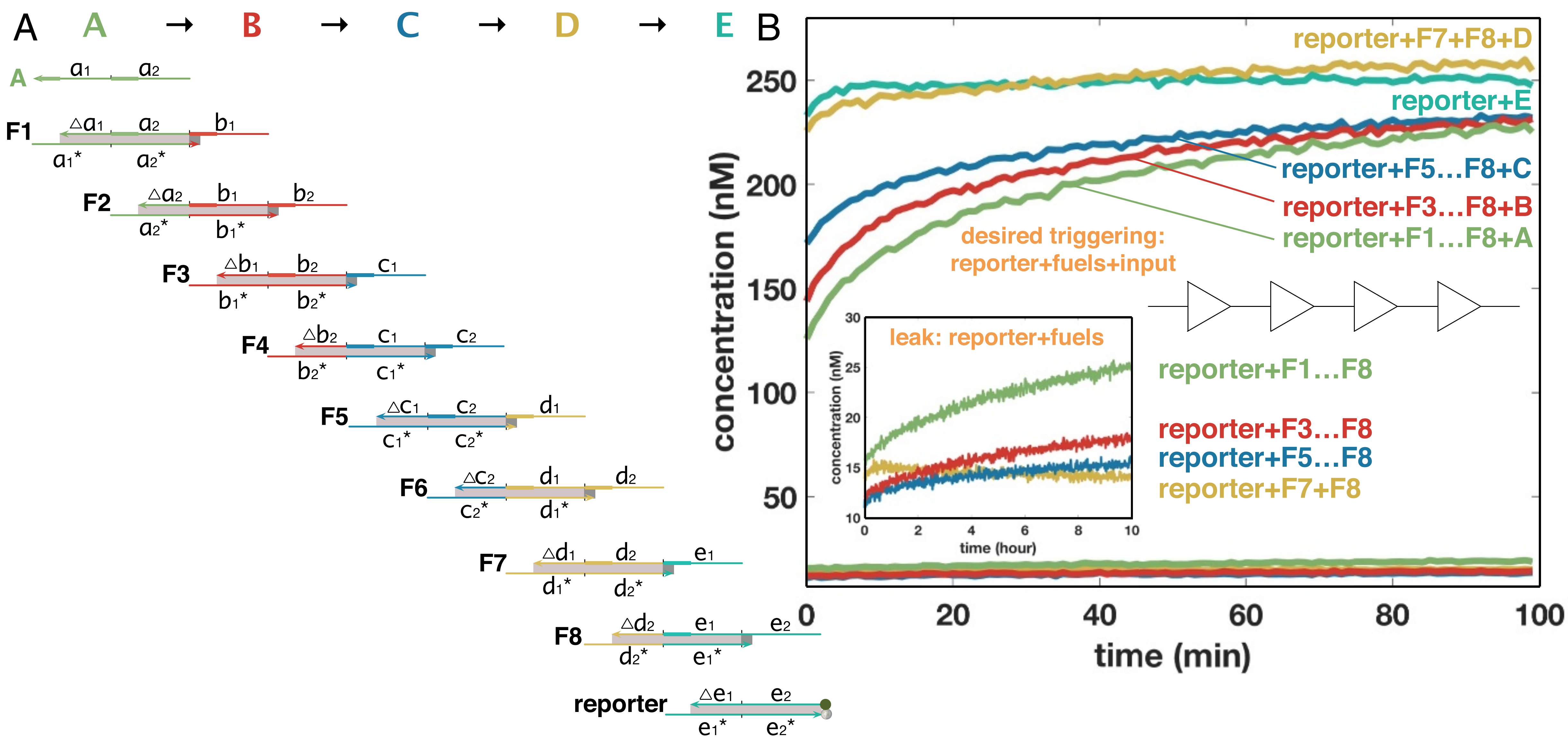


Fig. 8. DLD linear translator cascade. (*A*) Scheme of a four-layer translator cascade with nine steps of strand displacement implemented with eight fuel complexes and a reporter. (*B*) Kinetics of the desired reactions in the presence of input and the leak reactions in the absence of input. We tested different depths of the cascade by including a different subset of fuels. In each case, the desired reaction reached half-completion at the first measured data point. In the timeframe of 100 min, the leak fraction ([leak]/[initial fuel]) is less than 4%. (*Inset*) Leak of the translator cascade over 10 h. The largest leak fraction is roughly 5%: [reporter] = [fuels] = 500 nM, [input] = 250 nM, 25 °C. The inputs are added no more than 3 min before measurement.

DLD cascade: OR Gates

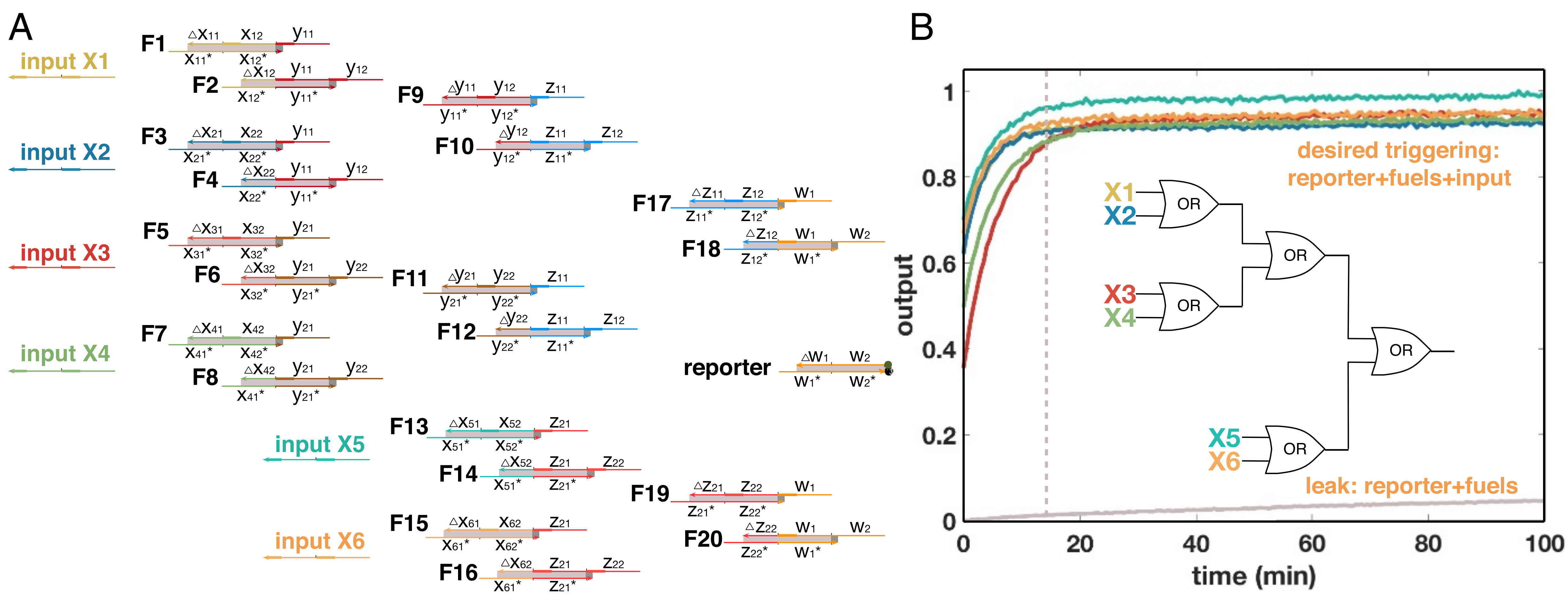


Fig. 9. DLD *OR* circuit. (*A*) Scheme of a three-layer DLD translator cascade with seven layers of strand displacement implemented with 20 fuels and a reporter, accepting six independent inputs. (*B*) Kinetics of the *OR* circuit with and without input signals. Output value 1 corresponds to fully triggering the reporter. Five of six desired triggering signals reach half-completion at the first measured data point, and all of six desired triggering signals reach 90% completion within 15 min (vertical dashed line). Leak is less than 1.5% in the first 15 min [input] = [fuels] = 1 μM, [reporter] = 500 nM, 25 °C. The inputs are added no more than 3 min before measurement.