DNA Reaction Networks Fueled by Strand Displacement

Hieu Bui
Outline

• Yurke-Tuberfield's DNA Tweezers: the invention of toe-hold mediated strand displacement
• Yurke: Catalytic Tweezers
• Winfree's Seesaw Gates
• Zhang's DNA Reaction Networks
• Zhang’s Allosteric DNA Catalytic Reactions
• Zhang’s DNA Catalytic Reactions
• Soloveichik's DNA Chemical Kinetics
• Cardelli's DNA Strand Algebracircuits
Toehold mediated strand displacement

Binding starts via toehold domains → Branch migration → Strand displacement completes
Example of Toehold mediated strand displacement
Details of Toehold mediated strand displacement

1. Toehold binding

2. Branch migration

3. Strand dissociation
Toehold-mediated strand displacement
vs toehold-mediated strand exchange
(where original strand melts off)
Example of Toehold mediated strand exchange
Multiple States of Toehold mediated strand displacement
Kinetics of Toehold mediated strand displacement and exchange (Zhang and Winfree)
Strand Displacement

\[ G^\circ_{\text{ABC}}, \quad G^\circ_{\text{rABC}}, \quad G^\circ_{\text{lABC}} \]

\[ \Delta G^\circ_r = G^\circ_{\text{rABC}} - G^\circ_{\text{ABC}} \]

\[ \Delta G^\circ_l = G^\circ_{\text{lABC}} - G^\circ_{\text{ABC}} \]

Nearest neighbor model

\[ P_r \propto \exp(-\Delta G^\circ_r / RT) \]

\[ P_l \propto \exp(-\Delta G^\circ_l / RT) \]
Yurke-Tuberfield's DNA Tweezers: the invention of toe-hold mediated strand displacement


Construction and operation of the molecular tweezers. a, Molecular tweezer structure formed by hybridization of oligonucleotide strands A, B and C. b, Closing and opening the molecular tweezers. Closing strand F hybridizes with the dangling ends of strands B and C (shown in blue and green) to pull the tweezers closed. Hybridization with the overhang section of F (red) allows F strand to remove F from the tweezers, forming a double-stranded waste product FF and allowing the tweezers to open. Complementary sections of B, C, F and F that hybridize to close and open the tweezers are coloured as in Fig. 1.
Energetics of Yurke-Tuberfield's DNA Tweezers

Closed Position

Open Position

Fuel

Fuel Waste
Rotary device using PX–JX2 junctions controlled by toe-hold mediated strand displacement

Yan, H., Zhang, X. P., Shen, Z. Y & Seeman, N. C.
A robust DNA mechanical device controlled by hybridization topology.
Seesaw Gates: use of toe-hold mediated strand exchange
Details of Seesaw Gates

The DNA motif for ‘seesaw’ gates. (a) Abstract gate diagram. Red numbers indicate initial concentrations. (b) The DNA gate motif and reaction mechanism. S1, S2, S3 and S4 are the recognition domains; T is the toehold domain; T’ is the Watson–Crick complement of T, etc. Arrowheads mark the 3’ ends of strands. Signal strands are named by their domains from 3’ to 5’, i.e. from left to right, so the input is S1TS2; gate base strands and threshold bottom strands are named by their domain from 5’ to 3’. All reactions are reversible and unbiased; solid lines indicate the dominant flows for the initial concentrations shown in (a), while the reverse reactions are dotted. (c) The threshold motif and reaction mechanism. The toehold is extended by a few bases (s’1, the complement of the first few 5′ bases of S1), providing an increased rate constant relative to the gate itself. Branch migration intermediate states are omitted from the diagram. (d) Example sequences. Gate complexes and signal molecules are shown at the domain level (second column) and at the sequence level (third column). Here, recognition domain sequences are 15 nt, the toehold domain sequence is 5 nt, and the toehold is extended by 3 nt for the threshold. Other lengths are possible, so long as they ensure that recognition domains will not spontaneously dissociate, toehold exchange is fast, and thresholding is sufficiently faster.
Abstract diagrams for seesaw gate circuits. (a) The general form of a gate node. Each gate \( i \) may be connected to many wires on each side, potentially all \( N \) nodes in the network, including itself. For each wire from the right side of gate \( i \) to the left side of gate \( j \), the initial concentration of the free signal \( w_{i,j} \) may be written above the wire, and the initial concentrations of gate complex \( g_{j,i:i} \) (\( w_{j,i} \) bound to gate \( i \)) and \( g_{i:i,j} \) (\( w_{i,j} \) bound to gate \( i \)) may be written within the node at the ends of the corresponding wires. Gate concentrations are simply omitted if they are zero. Initial concentrations of \( \text{th}_{j,i:i} \) (the threshold for \( w_{j,i} \) arriving at gate \( i \)) and \( \text{th}_{i:i,j} \) (the threshold for \( w_{i,j} \) arriving at gate \( i \)) may be written in the same locations as \( g_{j,i:i} \) and \( g_{i:i,j} \), respectively, but as negative numbers—or omitted if they are zero. (b) The general form of a wire. Each wire is specifically connected on its left end to the right side of a gate node, and connected on its right end to the left side of a gate node. (c) An example circuit with five realized gates (numbered circles), five virtual gates (numbers at ends of wires), and 11 wires. Each wire is identified by the two gates it connects; thus the virtual gates serve to provide full names (and sequences) to their incident wires. Note that circuit diagrams may be drawn without providing gate numbers, as they are not relevant to circuit function.
Seesaw Gates

Logical OR Gate

Logical AND Gate

Circuit diagrams and input/output behaviour of boolean logic gates. Output wires with arrowheads indicate that a downstream load is assumed, which consumes signal strands as they are released. (a–b) A two-input OR gate and a two-input AND gate using, respectively, 1 and 4 seesaw gates, the ‘1-4 scheme’. Circuits constructed using the 1-4 scheme are not clean, and thus would perform worse if threshold crosstalk and threshold inhibition were modelled. (c–d) A two-input OR gate and a two-input AND gate using two seesaw gates each, the ‘2-2 scheme’. Circuits constructed using the 2-2
Seesaw Gates for Logic OR and AND
Seesaw Gates

Compiling boolean logic circuits. (a) A sample circuit with six gates. (b) Translation into an equivalent dual-rail circuit with 12 gates. (c,d) Translation into an equivalent seesaw gate circuit with 32 gates (1-4 scheme) and 26 gates (2-2 scheme). (e,f) Simulation results for all 32 possible input vectors in the 1-4 scheme and in the 2-2 scheme. The concentrations of all four dual-rail output species are shown as a function of time. Delays vary with the input, depending the shortest decision path through the network. Simulations were run using the concentration $1x = 50$ nM, with ON inputs at 0.9x and OFF inputs at 0.1x. For the 1-4 scheme, the simulated reaction equations were augmented to also model threshold crosstalk, which degrades the performance of OR gates—but the system still works.
A 74L85 standard 4-bit magnitude comparator (four layers deep) and its seesaw circuit simulation, with $1x = 50$ nM. (a) The digital logic circuit diagram. The corresponding seesaw circuit has roughly 100 seesaw gates. (b) Seesaw circuit simulation with selected input vector of A greater than B. (c) Seesaw circuit simulation with selected input vector of A smaller than B. (d) Seesaw circuit simulation with selected input vector of A equal to B.
Seesaw Gates

Implementation of relay circuits. (a) A simple circuit with current source (battery) and controlled device (denoted by a resistor), the corresponding seesaw gate circuit, and its simulation using $1x = 50 \text{ nM}$. Shaded and unshaded sides of seesaw gates assist checking that a wire always connects different sides of two seesaw gates as required by node polarity, i.e. each wire connects the shaded side of one seesaw gate to the unshaded side of another. Switching signal A is provided at 1x if ON, or else 0.1x if OFF. Input current signal was provided at 10x; to verify that no output signal is produced when the current input is OFF, a 1x signal was provided. (b) AND logic. (c) OR logic. (d) A more complex circuit. Overlapping trajectories (orange and light blue) were shifted to the left by 100 s to make them visible. (e) Switching signal fan-out, current signal fan-out and current signal fan-in.
The seesaw gate motif and the construction of linear threshold gates.

Analogue time-domain circuits. (a) A catalytic cascade that exhibits initially quadratic growth, with $\alpha \approx 125 \text{ h}^{-2}$. Temporal trajectories are shown for a series of exponentially decreasing initial input concentrations. (b) A positive feedback circuit that exhibits initially exponential growth, with $\beta \approx 17 \text{ h}^{-1}$. The same series of exponentially decreasing input concentrations now yields a series of trajectories with linearly increasing half-completion times. (c) A pulse-generating circuit. Pulse amplitude depends on the input concentration. Here, we use a linear series of input concentrations between 0x and 1x. All simulations use 1x = 50 nM.
Parallel preparation of seesaw gates as hairpin precursors. Using DNA microarray synthesis technologies, each gate, threshold and fuel is made as a single strand. After cleavage from the surface, the mixture is annealed to form hairpins. Restriction enzymes then cleave the hairpins to form gate and threshold complexes. Relative concentrations (red numbers) are set based on the number of DNA chip spots dedicated to a given strand; in this example, one spot produces 0.25x.
Catalysis is the increase in the rate of a chemical reaction due to the participation of an additional substance called a catalyst. A reaction with a Catalysis is a Catalytic Reaction.
DNA Catalytic Cascades

**Catalytic DNA Systems**

1. Catalytic system powered by increase in the number of bases paired.
   

2. Catalytic system entropically driven by increase in the number of DNA strands.
   
Making a metastable fuel complex

Seelig, et al., JACS 90, 12211 (2006)
DNA Catalytic Cascades

Catalytic speedup of fuel-complex decay

Seelig, et al., JACS 90, 12211 (2006)
DNA Catalytic Cascades

Catalytic speedup: 5000
Turnover: 40

Seelig, et al., JACS 90, 12211 (2006)
DNA Catalytic Cascades

Entropy drive catalyst

Substrate S

DNA Catalytic Cascades

Entropy drive catalyst


Catalytic cycle

Domain 3 is four bases long.

Readout scheme

Fluorescence
DNA Catalytic Cascades

Entropy driven catalyst

Catalytic speedup: $1.9 \times 10^4$

DNA Catalytic Cascades

Using entropy to go uphill energetically

Substrate S

Waste W

DNA Catalytic Cascades

A catalytic cascade

[C0] is constant with time
For short times
[OB0] is proportional to $t$
[OB1] grows as $t^2$

DNA Catalytic Cascades

After 12 hours we can reliably distinguish between 0 pM and 1 pM of catalyst. 1pM of catalyst generated 900 pM of reporter. This is 900 fold amplification.

1 pM corresponds to about one molecule per eukaryotic cell.

DNA Catalytic Cascades

An autocatalytic system

DNA Catalytic Cascades

Exponential growth with saturation

DNA Reaction Networks
DNA Reaction Networks

\[ k_0 = 2.3 \times 10^1 \text{ M}^{-1} \text{s}^{-1} \]
\[ k_1 = 6.5 \times 10^5 \text{ M}^{-1} \text{s}^{-1} \]
\[ k_2 = 4.2 \times 10^5 \text{ M}^{-1} \text{s}^{-1} \]
\[ k_{\text{ROX}} = 4 \times 10^5 \text{ M}^{-1} \text{s}^{-1} \]
The entropy-driven reaction. (A) System components. Number labels denote functional domains, which are continuous stretches of DNA that act as units in binding. Domain $\bar{x}$ is the complement of (and will hybridize to) domain $x$. (B) The proposed catalytic pathway. Reverse reactions are also present and modeled (with the exception of $\text{I5} + \text{OB} \rightarrow \text{I4}$, which occurs at a negligible rate). (C) Analysis by PAGE (12% native gel) of the reaction mechanism. Unless otherwise noted, all experiments were performed at 25°C in tris-acetate (TE) buffer supplemented with 12.5 mM MgCl$_2$. Here, $[S] = [F] = 200$ nM. $[C] = 200$ nM, except where $C^*$ denotes 20 nM. “ann.” denotes that species were annealed; “30 m” denotes that the reaction occurred for 30 min. See fig. S5 for the full gel, including control lanes. (D) Fluorescent reporter strategy. ROX denotes the carboxy-X-rhodamine fluorophore, and RQ denotes the Iowa Black Red Quencher. Domain 2 is subdivided into 2a, 2b, and 2c; 2ab consists of 2a and 2b (Table 1). (E) Demonstration of catalysis. Different amounts of C were introduced into the system at $t \approx 0$. Here, $[S] = 10$ nM = $1 \times$, $[F] = 13$ nM, and $[OR] = 30$ nM. Fluorescence (in all figures) is reported in units such that 0.0 is the background fluorescence of the quenched reporter and 1.0 is the fluorescence of $\sim 10$ nM of triggered reporter. The control trace (black) shows the reaction with no substrate S and no catalyst C. Dotted lines show curves calculated with the reduced reaction model. sim, simulated.
A two-layer cascaded network. (A) Schematic. See table S2 for sequences of new domains. (B) Kinetics. Indicated amounts of initial catalyst C0 were added at $t \approx 0$. Fluorescence derives from reporter complex OR (Fig. 1D) at 30 nM. Dotted lines show simulated traces; see SOM text S8 for details on reaction rates and modeling. a.u., arbitrary units. (Inset) Response to 0.0010$x$, 0.0003$x$, and 0.0001$x$ catalyst. The asterisk indicates that three independent reaction traces are shown. 1.0 fluorescence units correspond to $\approx 10$ nM of triggered reporter.
Allosteric enzymes are enzymes that change their conformational ensemble upon binding of an effector, resulting in a change in binding affinity at a different ligand binding site.

This provides "action at a distance" through binding of one ligand affecting the binding of another at a distant site.
Allosteric Enzymes

http://resources.schoolscience.co.uk/unilever/16-18/proteins/protch6pg5.html
Allosteric DNA Catalytic Reactions

is the complement of (and will hybridize to) domain x. The sequences of 11 and 12 are illustrative of the domain concept and not in use for the allosteric catalyst design. (B) The allosteric catalyst (AC). There are three mechanistically important states that the AC can adopt: AC-OFF, AC-free, and AC-ON. Of these three, AC-ON is the most thermodynamically favored (due to the lengths of the hairpin stems; see Table 1), and AC-free is the least thermodynamically favored. The hairpin stem in AC-OFF is designed to be short enough that it can spontaneously open, causing AC to adopt the AC-free state. Then AC will quickly and with high probability fold into AC-ON. The predicted abundances of each state at equilibrium, for the sequences in Table 1, are shown. Domain 4t is a short two nucleotide domain that is complementary to the 3'-most two nucleotides of domain 4. The presence of domain 4t helps ensure that the AC-OFF is catalytically inactive. (C) The catalytic cycle. The AC binds to the substrate S via domain 5 and displaces the strand 6·3·4. The newly exposed 3 domain allows the fuel F to bind, displace output OB, and finally displace AC-ON. Domains 4 and 5 on the AC must be single-stranded in order for the catalytic cycle to proceed.

Table 1. Domain Sequences

<table>
<thead>
<tr>
<th>domain</th>
<th>sequence</th>
<th>length (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5'-CTTCTCTACA-3'</td>
<td>10</td>
</tr>
<tr>
<td>2a</td>
<td>5'-CTTACTGAC-3'</td>
<td>6</td>
</tr>
<tr>
<td>2b</td>
<td>5'-CTTACTCA-3'</td>
<td>6</td>
</tr>
<tr>
<td>2c</td>
<td>5'-ACTTACCTACGG-3'</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>5'-CCCT-3'</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5'-CATTCATACCTACCTAG-3'</td>
<td>16</td>
</tr>
<tr>
<td>4t</td>
<td>5'-CACTAC-3'</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5'-TCTCCACGAG-3'</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>5'-CCACATACATCATATT-3'</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>5'-TC-3'</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>5'-CTGACTTAC-3'</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>5'-GTATCTG-3'</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>5'-GTCTACTCTCTAAATG-3'</td>
<td>14</td>
</tr>
</tbody>
</table>

D.Y. Zhang et al. JACS 2008, 130, 13921
Allosteric DNA Catalytic Reactions

Figure 2. State changing of the AC. (A) State changing by addition of inhibitor (Inh) and activator (Act). Act binds to AC-ON to form the Inh:AC-ON complex. Free Act binds to Inh:AC-OFF to release AC-OFF and duplex waste product Inh:Act. AC-OFF then spontaneously converts to AC-ON. (B) Analysis by PAGE (12% native gel) of state changing. Here, [AC] = 200 nM, [Inh] = 250 nM, and [Act] = 300 nM. “(ann)” denotes that species were annealed; “(30 m)” denotes that the reaction proceeded for 30 min. Lanes 1 through 5 show slight smearing because the allosteric catalyst can dynamically switch between its two states over the course of the gel running. (C) The production of OB is quantitated via stoichiometric reaction with reporter complex OR to yield increased fluorescence. ROX denotes the carboxy-X-rhodamine fluorophore (attached to the DNA molecule via an NHS ester), and RQ denotes the Iowa Black Red Quencher. Domain 2 is subdivided into 2a, 2b, and 2c; 2ab consists of 2a and 2b (Table 1). The concentration of the OR reporter complex was always in excess of S to ensure that reporting delay time was approximately consistent. (D) Catalytic activity of the allosteric catalyst. Fluorescence (in all figures) is normalized so that 1 normalized unit (n.u.) of fluorescence corresponds to 1 nM of unquenched fluorophore-labelled strand 1-2a. Various reagents were added at $t = 0$. Serving as a control, catalyst C from ref 1 has sequence 4-5. Inh:AC-OFF was prepared by annealing AC with a 3× excess (2.7 nM final concentration) of Inh.

D.Y. Zhang et al. JACS 2008, 130, 13921
Figure 3. Allosteric catalyst behavior. (A) Dependence of catalytic behavior on the balance of [Act] and [Inh]. [S] = 30 nM = 1x, [F] = 40 nM = 1.3x, [AC] = 0.9 nM, [Inh] = 9 nM. Various amounts of Act were added at $t = 0$ min, with the number label shown being the value of $\alpha$, the stoichiometric excess of Act. Red shows where the AC is expected to be OFF (Inh:AC-OFF); green shows where AC is expected to be ON (AC-ON). (B) Sigmoidal activation curve. The total catalytic activity over 6 h (from (A)) is plotted against $\alpha$. Blue trace denotes the expected behavior when AC, Inh, and Act are in equilibrium. (C) Dynamic switching. Initially, 0.9 nM AC is annealed with 2.7 nM Inh. Reagents were added to cause concentration changes as follows: 20 nM Act at $t = 2$ h, 40 nM Inh at $t = 4$ h, 60 nM Act at $t = 7$ h, 80 nM Inh at $t = 9$ h, 100 nM Act at $t = 12$ h, 120 nM Inh at $t = 14$ h, 140 nM Act at $t = 19$ h. The fluorescence level was adjusted for dilution by multiplying the dilution factor where appropriate. (C, inset) Rate fitting for observed activity between 7 and 12 h. The production rate of OB was fit to be 3:0 nM/h for AC-ON, and 0:18 nM/h for Inh:AC-OFF (blue traces).

Table 2. Strand Sequences

<table>
<thead>
<tr>
<th>strand</th>
<th>domain</th>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>4 5 7 10/7 4:10 9</td>
<td>CATTCAATACCTACG TCTCCA TCTGCTACTCCTAATG GA TGGAGA CG CTATTGAGTAGAC CTAGATAC</td>
</tr>
<tr>
<td>Inh</td>
<td>8 9 10</td>
<td>CTTGACTG GTATCTAG GTCTACCTCCTAATG</td>
</tr>
<tr>
<td>Act</td>
<td>10/9 8</td>
<td>CATTGAGTAGTAGAC CTAGATAC GATGCAAG</td>
</tr>
<tr>
<td>F</td>
<td>2 3 4</td>
<td>CTCAGCTCCCTAATCCTACG GCCCT</td>
</tr>
<tr>
<td>OB</td>
<td>1 2</td>
<td>CATTCAATACCCCTACG</td>
</tr>
<tr>
<td>SB (part of S)</td>
<td>6 3 4</td>
<td>CCACATACCTATTT CCTCATTCAATCCTACG</td>
</tr>
<tr>
<td>LB (part of S)</td>
<td>5 4 3 2</td>
<td>TGGAGA CTAGGGTATGAAATG AGGG</td>
</tr>
<tr>
<td>OF (part of OR)</td>
<td>1 2a</td>
<td>CCGTAAATGTTGAGAAGAGTACG</td>
</tr>
<tr>
<td>OQ (part of OR)</td>
<td>2b 3a 1</td>
<td>TGGAGA CGTAGG TGTAGGAAG</td>
</tr>
</tbody>
</table>

Table 3. Catalytic Activity Based on State

<table>
<thead>
<tr>
<th>time (h)</th>
<th>state</th>
<th>OB production rate (nM/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–2</td>
<td>OFF</td>
<td>0.00</td>
</tr>
<tr>
<td>2–4</td>
<td>ON</td>
<td>3.5</td>
</tr>
<tr>
<td>4–7</td>
<td>OFF</td>
<td>0.24</td>
</tr>
<tr>
<td>7–9</td>
<td>ON</td>
<td>3.0</td>
</tr>
<tr>
<td>9–12</td>
<td>OFF</td>
<td>0.18</td>
</tr>
<tr>
<td>12–14</td>
<td>ON</td>
<td>1.6</td>
</tr>
<tr>
<td>14–19</td>
<td>OFF</td>
<td>0.08</td>
</tr>
</tbody>
</table>
Allosteric DNA Catalytic Reactions

D.Y. Zhang et al. JACS 2008, 130, 13921
DNA Catalytic Reactions

Figure 1. A non-covalent strand displacement reaction catalyzed a target ssDNA molecule C [adapted from (9)]. (A) DNA abstraction. The double-helix DNA molecule (top) is typically abstracted as two directional lines, one for each strand, with base identities shown (middle). Here, we abstract the DNA molecule one step further by grouping contiguous nucleotides into domains, functional regions of DNA that act as a unit in binding (bottom). Domains are labeled by numbers. Domain $\bar{x}$ is the complement of (and will hybridize to) domain $x$. The strands OP, SP and SL form the three-stranded DNA complex S. The DNA molecule in the top panel was drawn using Nanoengineer, a free DNA visualization software by Nanorex. (B) The designed mechanism of catalytic function. (C) Fluorescent reporter complex. Output product OP reacts stoichiometrically with reporter complex R to yield a fluorescent strand. ROX denotes the carboxy-3,5-dihydrothiazole fluorophore (attached to the DNA molecule via an NHS ester), and RQ denotes the Iowa Black Red Quencher. This indirect reporter complex was used because of the thermodynamic effects of fluorophore-quencher binding (20). From (9), $k_{\text{ROX}}$, the second-order rate constant of reaction between OP and R, was measured to be $4 \times 10^7 \text{M}^{-1}\text{s}^{-1}$. In experiments, the concentration of the reporter $R$ was in excess of the concentration of the fuel $F$ and substrate $S$ to minimize the reporter delay (no more than 2 min for $[R] = 30 \text{nM}$). (D) Experimental and simulation results from (9). Dotted lines show ordinary differential equation (ODE) simulation results according to the model in Table 2.
DNA Catalytic Reactions

Figure 2. Catalytic turnover. (A) Raw data for turnover experiments. Traces showed significantly more noise than typical; possibly, this is due to lamp and temperature instability. (B) Turnover plotted as a function of time. Turnover is calculated as the excess normalized fluorescence above leak (0× trace) divided by concentration of catalyst; e.g. $T_{0.001\%}(t) = F_{0.001\%}(t) - F_\text{0}(t)/0.001\% - 10\text{nM}$. ‘Old sim’ denotes simulations using the model presented in (9), which was fitted only to the 10 nM data shown in Figure 1D. The ‘new sim’ simulations use the model and rate constants fitted to the data presented in this article. The relative ordering and the quantitative differences between the ‘new sim’ simulations and the experimental results are not considered significant—all traces are considered to be within experimental error of one another. (C) Our new model that accounts for $F$ subpopulations. A small fraction (fitted to be 1.0%) of $F$ exists as $F\beta$, with deletions in domain 4. These react with $I\alpha$ to yield $X$, from which $C$ cannot dissociate.

Table 2. Original model

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Rate Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S + C \overset{k_1}{\underset{k_2}{\rightleftharpoons}} I\alpha + SP$</td>
<td>$K_1 = 6.5 \times 10^3 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$I\alpha + F \overset{k_4}{\underset{k_3}{\rightleftharpoons}} I\alpha + OP$</td>
<td>$K_2 = 4.2 \times 10^3 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$I\alpha \overset{k_5}{\underset{k_6}{\rightleftharpoons}} W + C$</td>
<td>$K_3 = 4 \times 10^{-3} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$S + F \overset{k_7}{\underset{k_8}{\rightleftharpoons}} OP + SP + W$</td>
<td>$K_0 = 2.3 \times 10^1 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
</tbody>
</table>

Table 3. Revised model

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Rate Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S + C \overset{k_1}{\underset{k_2}{\rightleftharpoons}} I\alpha + SP$</td>
<td>$k_1 = 2.7 \pm 0.5 \times 10^5 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$I\alpha + F\beta \overset{k_4}{\underset{k_3}{\rightleftharpoons}} I\alpha + OP$</td>
<td>$k_2 = 1.1 \pm 0.7 \times 10^6 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$I\alpha \overset{k_5}{\underset{k_6}{\rightleftharpoons}} W + C$</td>
<td>$k_3 = 1.1 \pm 0.5 \times 10^{-2} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$S + F\beta \overset{k_7}{\underset{k_8}{\rightleftharpoons}} OP + SP + W$</td>
<td>$k_0 = 5 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$I\alpha + F\beta \overset{k_1}{\underset{k_2}{\rightleftharpoons}} X + OP$</td>
<td>$k_1 = 1.0 \pm 0.3 \times 10^{-2}$</td>
</tr>
<tr>
<td>$S + F\beta \overset{k_1}{\underset{k_2}{\rightleftharpoons}} OP + SP + W2$</td>
<td>$k_0 = 5 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$C + W2 \overset{k_1}{\underset{k_2}{\rightleftharpoons}} X$</td>
<td>$k_0 = 5 \text{M}^{-1} \text{s}^{-1}$</td>
</tr>
</tbody>
</table>
DNA Catalytic Reactions

Figure 3. Catalytic function and model results for (A) 100 nM, (B) 30 nM, (C) 10 nM and (D) 1 nM substrate S concentration. The new model that accounts for catalyst inactivation (Table 3) fits the experimental data better than the old model (Table 2). Red traces denote 0.1× catalyst, whereas blue traces denote catalyst concentrations of ~1 nM.
DNA Catalytic Reactions

**Figure 4.** Effects of overhangs on catalytic activity. Schematics for catalyst C with a (A) 3’ dsDNA overhang, (B) 3’ssDNA overhang, (C) 5’dsDNA overhang, (D) 5’ssDNA overhang and (E) 5’ssDNA overhang (mostly poly-T). (F) Catalytic activity of catalyst molecules with various overhangs. The magenta traces show that the 7 and 8 strands possess no catalytic activity on their own, and that the increased catalytic activity of (A) and (C) over (B) and (D) is only due to the single/double-stranded state of the overhang.

**Figure 5.** 5’/3’ inverted catalyst and substrate. (A) Schematic and (B) results. This system uses the same reporter complex as the original catalytic reaction (shown in Figure 1C). Substrate S2 does not react significantly with reporter complex R. Note that the single-stranded domain 2 on S2 could hybridize to the complementary domain 2b on reporter R. From there, it is possible to initiate a four-way branch migration process that could result in the release of the fluorophore-labeled DNA. However, four-way branch migration processes are significantly slower than three-stranded branch migration, and the hybridization of domain 2b is transient enough that this unintended pathway does not seem to be significant at our experimental conditions.
DNA Catalytic Reactions

Figure 6. Effects of single base catalyst C mutations. (A) Mismatches. The trajectory labels show the position and new identity of the mutated base. For example, 'm-17-A' denotes that the 17th base of catalyst C (from the 5’-end), was mutated from thymine to a adenine. (B) Insertions. The inserted base is inserted before (5’ of) the position denoted. (C) Deletions. (D) Summary of suppression by various catalyst mutations. The suppression factor is calculated as the initial slope of activity by the standard trace divided by the initial slope of activity by the mutated catalysts:

\[ S.F. = \frac{F(t_{10\text{min}}) - F(0\text{min})}{(F_m(t_{3\text{h}}) - F_m(0\text{h}))/(3\text{h})} \]

where \( F_m(t) \) is the fluorescence value due to a mutated catalyst at time \( t \). (D. inset) The sequence of the catalyst molecule C is shown with the positions at which mutations were performed.
DNA Catalytic Reactions

Figure 7. Effects of single-base fuel $F$ mutations. (A) Mismatches. (B) Insertions. The inserted base is inserted before (5') of the position denoted. (C) Deletions. (D) Summary of the catalytic activity using mutant fuel molecules. (D. inset) Sequence of the fuel molecule $F$ and positions of mutations.
DNA Catalytic Reactions

Figure 8. Behavior of the catalytic reaction using fuel, substrate and catalyst oligonucleotides with no post-synthesis strand purification. (A) Effects of using unpurified DNA on catalytic activity. Uppercase ‘F’ denotes fuel $F$ purified commercially by HPLC, while lowercase ‘f’ denotes unpurified fuel. Similarly, uppercase ‘S’ and ‘C’ denote that the strands in $S$ and the catalyst $C$ were purified. Note that though the substrate $s$ used unpurified DNA strands, it was still manually purified by PAGE to ensure correct stoichiometry. (B) Effects of unpurified DNA on the uncatalyzed reaction rate. The dotted rate shows simulation results for $k_0 = 5 \text{ M}^{-1} \text{s}^{-1}$. (C) Denaturing gel of purified and unpurified strands and complexes. Lanes 2 and 3 show the PAGE-purified substrate prepared from unpurified and purified $OP$, $SP$ and $SL$ strands, respectively. This shows the degree to which truncated strands are present in PAGE-purified $S$ and $s$. Lanes 4 and 5 show the unpurified and purified catalyst, respectively. Lanes 6 and 7 show the unpurified and purified fuel, respectively. (D) Substrate $S$ purity. Gel band intensities are displayed in arbitrary units (a.u.). Solid blue lines denote the (arbitrarily chosen) limits of the correct-length $SL$ strand, while dotted blue lines denote the (arbitrarily chosen) limits of the truncated $SL$. The black horizontal line denotes the background intensity of the gel. The integrated intensity of the correct-length $SL$ above the background is divided by the sum of it and the integrated intensity of truncated $SL$ to yield the correct length fraction (summarized in Table 4). Similarly, solid magenta lines denote the limits of correct-length $OP$ and $SP$, while dotted magenta lines denote the limits of truncated $OP$ and $SP$. (E) Catalyst $C$ purity. (F) Fuel $F$ purity.
DNA Catalytic Reactions

Figure 10. Catalytic design using four-letter alphabets. (A) Schematic of catalytic design using four-letter base sequences. Domain sequences are given in Table 5. (B) Fluorescence characterization of the kinetics of the reporter complex. The best fit value of $k_{44}$ (that produced the simulation traces shown as dotted lines) was $8.2 \times 10^5$ M$^{-1}$ s$^{-1}$. (C) Fluorescence characterization of the kinetics of the four-letter catalyst system. (D) Schematic of hybrid catalytic design using four-letter base sequences in the output domains S1 and S2, and a three-letter base sequences in the catalytic domains 63–66. Domain sequences are given in Table 5. ‘RG’ on the reporter complex denotes Rhodamine Green. ‘FQ’ on the reporter complex denotes the proprietary Iowa Black Fluorescence Quencher. (E) Fluorescence characterization of the kinetics of the reporter complex. The best fit value of $k_{44}$ (that produced the simulation traces shown as dotted lines) was $1.4 \times 10^5$ M$^{-1}$ s$^{-1}$. (F) Fluorescence characterization of the kinetics of the hybrid sequence catalyst system.

Table 5. Domain sequences for four-letter designs

<table>
<thead>
<tr>
<th>Dom.</th>
<th>Sequence</th>
<th>Length (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>5'-TGTTACTGGCTCTGAT-3'</td>
<td>16</td>
</tr>
<tr>
<td>42a</td>
<td>5'-GACC-3'</td>
<td>4</td>
</tr>
<tr>
<td>42b</td>
<td>5'-AAATGAAT-3'</td>
<td>7</td>
</tr>
<tr>
<td>42c</td>
<td>5'-ACCGTTAC-3'</td>
<td>9</td>
</tr>
<tr>
<td>42d</td>
<td>5'-GACCAATGATACCCTTAC-3'</td>
<td>20</td>
</tr>
<tr>
<td>43</td>
<td>5'-GAAAG-3'</td>
<td>5</td>
</tr>
<tr>
<td>44</td>
<td>5'-GCACATTAAGTCTAC-3'</td>
<td>15</td>
</tr>
<tr>
<td>45</td>
<td>5'-CAATGTTC-3'</td>
<td>8</td>
</tr>
<tr>
<td>46</td>
<td>5'-ATAAGCTATGGT-3'</td>
<td>15</td>
</tr>
<tr>
<td>51</td>
<td>5'-ATAAGCTTGATAGC-3'</td>
<td>15</td>
</tr>
<tr>
<td>52a</td>
<td>5'-GAGAC-3'</td>
<td>5</td>
</tr>
<tr>
<td>52b</td>
<td>5'-CTAGCA-3'</td>
<td>7</td>
</tr>
<tr>
<td>52c</td>
<td>5'-CCTGAACAC-3'</td>
<td>10</td>
</tr>
<tr>
<td>52d</td>
<td>5'-GAGACCTAGCAACCTGAAACCA-3'</td>
<td>22</td>
</tr>
<tr>
<td>53</td>
<td>5'-CCCTC-3'</td>
<td>5</td>
</tr>
<tr>
<td>64</td>
<td>5'-ATACAATACC-3'</td>
<td>11</td>
</tr>
<tr>
<td>65</td>
<td>5'-TCACCATG-3'</td>
<td>8</td>
</tr>
<tr>
<td>66</td>
<td>5'-TTTTTTTTTT-3'</td>
<td>10</td>
</tr>
</tbody>
</table>
DNA Catalytic Reactions

Figure 11. Effects of impurity and 5'3' orientation on maximum turnover. (A) Schematic of the original catalytic pathway when the fuel strand possesses an unintended 5'-Transesterification. 5'-Transected fuel strands 5't are unable to quickly displace catalyst 5m from intermediate 5m. Catalyst cannot be regenerated to catalyze other reactions. So turnover is severely limited. (C) Maximum turnover of the original catalyst design. Catalyst C is unmodified, and substrate complex S is PAGE purified from unpurified strands. Fuel molecules F are unmodified (None' trace), HPLC purified (HPLC), or dual HPLC/PAGE-purified ("Dual"). HPLC and dual HPLC/PAGE-purified fuel allows a maximum turnover of over 50, while unmodified fuel allows a maximum turnover of less than 10. Maximum turnover is calculated as in Figure 2B: the plotted turnover is the excess fluorescence signal of an experiment with unmodified concentration of S/over that of an experiment lacking C, divided by (C). (D) The reporter complex for the mirror catalytic process, the strands for this reporter complex likewise mirror the sequence of the strands for the original catalytic system. (E) Maximum turnover measurement of the mirror catalytic design. Catalyst 5m is unmodified, and substrate complex 5m is PAGE purified from unpurified strands. Fuel molecules 5m are unmodified (None' trace), HPLC purified (HPLC) or dual HPLC/PAGE-purified ("Dual"). The maximum turnover of the mirrored catalytic system using dual HPLC/PAGE-purified fuel is seen to be roughly 45, while the HPLC purified and unpurified fuels allows a maximum turnover of no more than 10. The spike near t ~ 0 is due to the subtractive nature of the method for calculating turnover, and is likely an artifact (this also exists in Figure 2B).

Table 6. Domain sequences of mirrored sequences (Figure 11B)

<table>
<thead>
<tr>
<th>Dom.</th>
<th>Sequence</th>
<th>Length (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1m</td>
<td>5'-ACATCTTTTCT-3'</td>
<td>10</td>
</tr>
<tr>
<td>2am</td>
<td>5'-GGCATTTCAATCA-3'</td>
<td>12</td>
</tr>
<tr>
<td>2bm</td>
<td>5'-ACCTCT-3'</td>
<td>6</td>
</tr>
<tr>
<td>2cm</td>
<td>5'-GCATCC-3'</td>
<td>6</td>
</tr>
<tr>
<td>2m</td>
<td>5'-GGCATTTCAATCAACCTCTGCATCC-3'</td>
<td>24</td>
</tr>
<tr>
<td>3m</td>
<td>5'-TCCC-3'</td>
<td>4</td>
</tr>
<tr>
<td>4m</td>
<td>5'-GCATCCATAACTTAC-3'</td>
<td>16</td>
</tr>
<tr>
<td>5m</td>
<td>5'-ACCTCT-3'</td>
<td>6</td>
</tr>
<tr>
<td>6m</td>
<td>5'-TTATACTACATACACC-3'</td>
<td>16</td>
</tr>
</tbody>
</table>
DNA Chemical Kinetics

Fig. 1. Strand displacement molecular primitive. Domains are labeled by numbers, with * denoting Watson–Crick complementarity. Multiple elementary steps are indicated: (1) binding of toeholds 1 and 1*; (2) branch migration, a random walk process where domain 2 of strand 2–3 is partially displaced by domain 2 of strand 1–2; (3) the separation of toeholds 3 and 3*. (Inset) The single-reaction model of strand displacement.

Fig. 2. Unimolecular module: DNA implementation of the formal unimolecular reaction $X_1 \rightarrow X_2 + X_3$ with reaction index $i$. Orange boxes highlight the DNA species that correspond to the formal species $X_1$ (species identifier 1-2-3), $X_2$ (4-5-6), and $X_3$ (7-8-9). Domains identical or complementary to species identifiers for $X_1$, $X_2$, and $X_3$ are colored red, green, and blue, respectively. Black domains (10 and 11) are unique to this formal reaction. To reduce rate constant $q_i$, toehold domain 1*$_i$ may not be a full complement of domain 1. (A) Complex $G_i$ undergoes a strand displacement reaction with strand $X_1$, with $X_1$ displacing strand $O_i$. (B) $O_i$ displaces $X_2$ and $X_3$ from complex $T_i$. Without buffer cancellation, $q_i = k_i C_{max}^{-1}$, with buffer cancellation, $q_i = \gamma^{-1} k_i C_{max}^{-1}$. Reaction equations of type 2–3 are used in simulations (Figs. 5 and 6); simplified reaction equations 4–5 are useful for analysis.
DNA Chemical Kinetics

Fig. 3. Bimolecular module: DNA implementation of the formal bimolecular reaction $X_1 + X_2 \rightarrow X_3$ with reaction index $i$. The black domain (12) is unique to this formal reaction. (A) $X_1$ reversibly displaces $B_i$ from complex $L_i$, producing complex $H_i$. (B) $X_2$ displaces $O_i$ from complex $H_i$. Occurrence of reaction $B$ precludes the backward reaction of $A$. (C) $O_i$ displaces $X_3$ from complex $T_i$. Without buffer cancellation, $q_i = k_i$; with buffer cancellation, $q_i = \gamma^{-1}k_i$. Reaction equations of type 7–9 are used in simulations (Figs. 5 and 6); simplified reaction equations 10–11 are useful for analysis.

Fig. 4. Buffering module: DNA implementation of the buffering module used to cancel out the buffering effect. A buffering module is needed for each formal species $X_j$ for which $q_j < \alpha$. The buffering module for species $X_2$ is shown. $X_2$ reversibly displaces $BS_2$ from complex $LS_2$ to produce complex $HS_2$, similarly to the first reaction of the bimolecular module. The black domain (13) is unique to this buffering module for species $X_2$. Set $q_{S_j} = \gamma^{-1}(\alpha - \alpha_j)$. Reaction equations 12 are used in simulations (Figs. 5 and 6); simplified reaction equations 13 are useful for analysis.
DNA Kinetics for Lotka-Volterra Chemical Oscillator

(A) Ideal chemical reactions
1. $X_1 + X_2 \xrightarrow{k_1} 2X_2$
2. $X_1 \xrightarrow{k_2} 2X_1$
3. $X_2 \xrightarrow{k_3} \emptyset$

<table>
<thead>
<tr>
<th>unscaled</th>
<th>scaled</th>
<th>$q_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>1.5</td>
<td>$5 \times 10^3$ M/s</td>
</tr>
<tr>
<td>$k_2$</td>
<td>1</td>
<td>1/300 s</td>
</tr>
<tr>
<td>$k_3$</td>
<td>1</td>
<td>1/300 s</td>
</tr>
</tbody>
</table>

(B) DNA reaction modules
1. $X_1 + L_1 \xrightarrow{q_1 / q_{\text{max}}} H_1 + B_1$
2. $X_1 + H_1 \xrightarrow{q_{\text{max}}} O_1$
3. $O_1 + T_1 \xrightarrow{q_{\text{max}}} 2X_2$

Buffering module:
4. $X_2 + L_S \xrightarrow{q_{\text{max}}} H_S + B_S$

5. $X_2 \xrightarrow{q_{\text{max}}} 2X_1$

6. $X_2 + G_2 \xrightarrow{q_2 / q_{\text{max}}} O_2$

7. $O_2 + T_2 \xrightarrow{q_{\text{max}}} 2X_1$

8. $X_2 + G_3 \xrightarrow{q_3 / q_{\text{max}}} O_3$

(C) Simulation of ideal and DNA reactions

Fig. 5. Lotka-Volterra chemical oscillator example. (A) The formal chemical reaction system to be implemented with original (unscaled) and scaled rate constants. Desired initial concentrations of $X_1$ and $X_2$ are 2 and 1 unscaled and 20 and 10 nM scaled. (B) Reactions modeling our DNA implementation. Each formal reaction corresponds to a set of DNA reactions as indicated. Species $X_2$ requires a buffering module because $\sigma_2 < \sigma$ ($\sigma = \sigma_1 = k_1$ and $\sigma_2 = 0$). Maximum strand displacement rate constant $q_{\text{max}} = 10^6$ M$^{-1}$ s$^{-1}$ and initial concentration of auxiliary species $G_i$, $T_i$, $L_i$, $B_i$, $L_S$, and $B_S$ is $C_{\text{max}} = 10$ μM. Buffering-scaling factor $\gamma^{-1} = q_{\text{max}}(q_{\text{max}} - \sigma)^{-1} = 2$. The initial concentrations of strands $X_1$ and $X_2$ introduced into the system is $\gamma^{-1} 20$ nM = 40 nM and $\gamma^{-1} 10$ nM = 20 nM. (C) Plot of the concentrations of $X_1$ (Red curve) and $X_2$ (Green curve) for the ideal system (Dashed line) and the corresponding DNA species (Solid line).
DNA Chemical Kinetics

Fig. 6. Examples showing more complex behavior. In all the maximum strand displacement rate constant $q_{\text{max}} = 10^6 \text{ M}^{-1} \text{s}^{-1}$ and the initial concentration of auxiliary species $C_{\text{max}} = 10 \mu\text{M}$. See Figs. S4 and S5 and SI Text for the rate constants used in A–D, as well as the details of C and D. Plots show the ideal CRN (Dashed lines) and the DNA reactions (Solid lines).