Supporting Information

Soloveichik et al. 10.1073/pnas.0909380107

SI Text

The Formal Model of Chemical Reaction Networks (CRNs). Consider a system with n species $X_1, \ldots, X_n$. Let $\mathscr{U}$ be the set of unimolecular reactions and $\mathscr{B}$ be the set of bimolecular reactions. A unimolecular reaction $i \in \mathscr{U}$ is defined by the index $r_i \in \{1, \ldots, n\}$ of its single reactant species, real-valued rate constant $k_i > 0$, and integer product coefficients for each species $c_{i,j} \geq 0$:

$$X_{r_i} \xrightarrow{k_i} c_{i,1} \cdot X_1 + \cdots + c_{i,n} \cdot X_n. \quad \text{[S1]}$$

A bimolecular reaction $i \in \mathscr{B}$ is defined by the indexes $r_{i,1}, r_{i,2} \in \{1, \ldots, n\}$ of its two reactant species, real-valued rate constant $k_i > 0$, and integer product coefficients for each species $c_{i,j} \geq 0$:

$$X_{r_{i,1}} + X_{r_{i,2}} \xrightarrow{k_i} c_{i,1} \cdot X_1 + \cdots + c_{i,n} \cdot X_n. \quad \text{[S2]}$$

The concentrations $[X_j] \geq 0$ of species X_j are governed by the reactions according to the ordinary differential equations:

$$\dot{[X_j]} = -\sum_{i \in \mathscr{U} | r_i = j} k_i [X_j] - \sum_{i \in \mathscr{B} | r_{i,1} = j} k_i [X_j][X_{r_{i,2}}] - \sum_{i \in \mathscr{B} | r_{i,2} = j} k_i [X_{r_{i,1}}][X_j] + \sum_{i \in \mathscr{U}} c_{i,j} k_i [X_{r_i}] + \sum_{i \in \mathscr{B}} c_{i,j} k_i [X_{r_{i,1}}][X_{r_{i,2}}], \quad \text{[S3]}$$

with initial concentrations $[X_j](0) \geq 0$.

Procedure for Compiling a Formal CRN to DNA-based Chemistry. In this section we summarize the procedure for compiling an arbitrary CRN of unimolecular and bimolecular reactions into strand displacement DNA-based chemistry. By assuming perfect strand displacement reactions (e.g., no leak reactions; see main text), the target system can be simulated with arbitrary accuracy over an arbitrarily long period, up to a uniform scaling in time and a uniform scaling of the concentrations of the formal species.

Given a CRN, we first formulate the corresponding system of reactions implementable with strand displacement chemistry by using the modules described in the main text. Choosing parameter $C_{\max}$, we simulate the system, increasing $C_{\max}$ as needed to obtain a desired level of accuracy. As shown in *Proof of Convergence of the DNA Implementation to the Target CRN as* $C_{\max} \to \infty$, perfect accuracy is obtained in the limit $C_{\max} \to \infty$. Then we rescale the resulting implementation into a regime with realistic concentrations and strand displacement rate constants while preserving the same dynamic behavior. Last, we design DNA sequences implementing the desired strand displacement reactions, under reasonable assumptions about DNA hybridization and branch migration behavior.

Consider $C_{\max}$ to be an accuracy parameter to be determined later. We first choose $q_{\max}$, which after rescaling time will equal the maximum rate constant achievable by a DNA strand displacement reaction in the implementation. To set $q_{\max}$ note that our construction requires $q_i, qs_j \leq q_{\max}$ in reactions 2, 7, and 12 of the main text (Figs. 2–4). Consequently, with buffer cancellation, the smallest

$$q_{\max} = \max\{\max_{i \in \mathscr{U}} k_i / C_{\max} + \sigma, \max_{i \in \mathscr{B}} k_i + \sigma, \max_j (2\sigma - \sigma_j)\},$$

where $\sigma_j = \Sigma_{i \in \mathscr{B} | r_{i,1} = j} k_i$ and $\sigma = \max_j \{\sigma_j\}$. σ_j indicates the overall rate constant for formal species X_j participating as the left reactant in *any* bimolecular reaction; when translated to DNA, this will determine how much DNA species X_j is buffered. We can use a larger $q_{\max}$ but that will result in a slower implementation after rescaling.

Let $\gamma^{-1} = q_{\max}(q_{\max} - \sigma)^{-1}$ be the buffering-scaling factor, indicating the scale-up to rate constants necessary to cancel the effect of buffering of the signal species. Every unimolecular reaction $i \in \mathscr{U}$ corresponds to the reactions:

$$X_{r_i} + G_i \xrightarrow{q_i} O_i, \quad \text{[S4]}$$

$$O_i + T_i \xrightarrow{q_{\max}} c_{i,1} \cdot X_1 + \cdots + c_{i,n} \cdot X_n, \quad \text{[S5]}$$

where $q_i = \gamma^{-1} k_i C_{\max}^{-1}$. Every bimolecular reaction $i \in \mathscr{B}$ corresponds to the reactions

$$X_{r_{i,1}} + L_i \underset{q_{\max}}{\overset{q_i}{\rightleftharpoons}} H_i + B_i, \quad \text{[S6]}$$

$$X_{r_{i,2}} + H_i \xrightarrow{q_{\max}} O_i, \quad \text{[S7]}$$

$$O_i + T_i \xrightarrow{q_{\max}} c_{i,1} \cdot X_1 + \cdots + c_{i,n} \cdot X_n, \quad \text{[S8]}$$

where $q_i = \gamma^{-1} k_i$ and $B_i = B_{i'}$ if $X_{r_{i,1}} = X_{r_{i',1}}$ and $X_{r_{i,2}} = X_{r_{i',2}}$. (This equality arises because the sequence for DNA species B_i depends only on the identity of reactants $r_{i,1}$ and $r_{i,2}$ and does not contain any domain unique to reaction i.) Furthermore, for every species X_j for which $\sigma_j < \sigma$ we formulate the reactions

$$X_j + LS_j \underset{q_{\max}}{\overset{qs_j}{\rightleftharpoons}} HS_j + BS_j, \quad \text{[S9]}$$

where $qs_j = \gamma^{-1}(\sigma - \sigma_j)$. These reactions serve no purpose other than to increase the buffering load of DNA species that are not already maximally buffered, so that in the end all DNA species are buffered equally. [Similar techniques are useful for balancing networks of reactions with Michaelis–Menten kinetics (1).] The initial concentrations of G_i, T_i, L_i, B_i, LS_j, and BS_j are $C_{\max}$, and the initial concentration of the signal species $[X_j](0) = \gamma^{-1}\widehat{[X_j]}(0)$, where $\widehat{[X_j]}(0)$ are the initial concentrations in the target CRN.

We choose $C_{\max}$ large enough to attain a desired level of accuracy ascertained by numerical simulations of [**S4**–**S9**]. Once the desired level of accuracy is achieved over a sufficiently long time period of the behavior of the target reaction network, we choose an appropriate time-scaling factor α and concentration-scaling factor β to be compatible with the limitations in concentrations and rate constants for the experimental regime. In particular, for fastest experimentally realistic implementation of desired dynamics we can rescale the DNA implementation so that the largest strand displacement rate constant $q_{\max}$ scales to 10^6/M/s and the largest concentration $C_{\max}$ scales to 10^{-5} M.

We design DNA sequences on the basis of the modules illustrated in Figs. 2–4 (main text) as follows (see the main text for a discussion of the design of unique sequences). For each species X_j we use a unique species identifier and a species-specific domain for the buffering module (black domain in Fig. 4, main text). For each reaction we use unique reaction-specific domains

(black domains in Figs. 2–3, main text). Strand displacement reactions with rate constant $q_{\max}$ should use a full complementary toehold domain for maximum reaction rate. Strand displacement rate constants $q < q_{\max}$ are obtained by using partially complementary toehold domains x_q^* (see Fig. S1) on the relevant DNA complexes, with the rate constant predicted as in ref. 2. In the design of T_i complexes, the product coefficients $c_{i,j} \geq 1$ are implemented by having $c_{i,j}$ copies of the X_j product released by the T_i complex.

In the experimental protocol, each DNA complex is prepared separately by mixing the component strands, annealing, and purifying as in ref. 3. All the auxiliary species as well as the initial signal species are then mixed together to begin the experiment. Fluorescent markers may be used to measure the concentration of signal species X_j directly or indirectly (2).

We automated the above compilation procedure up to the sequence design step. *Mathematica* code that enumerates the strand displacement reactions in our DNA implementation of a given target CRN, scales the system into an experimentally realistic regime, and simulates the ideal target system together with the DNA implementation can be downloaded from http://dna.caltech.edu/SupplementaryMaterial/DNA_for_CRNs/.

Proof of Convergence of the DNA Implementation to the Target CRN as $C_{\max} \to \infty$. In this section we use singular perturbation theory (4, 5) to prove that our DNA implementation (with buffer cancellation) approaches the modeled target system with increasing $C_{\max}$. (Without buffer cancellation, the limit $C_{\max} \to \infty$ is not sufficient, and the additional limit $q_{\max} \to \infty$ is required for arbitrarily high accuracy.) Intuitively, singular perturbation theory provides a rigorous way of separating our DNA implementation into fast and slow dynamics, with the separation increasing as $C_{\max}$ get larger. Then the slow dynamics approach the desired target system in the limit of fast dynamics reaching instant pseudoequilibrium. For simplicity, we assume a fixed concentration of reservoir species G_i, T_i, L_i, B_i, LS_j, and BS_j equal to $C_{\max}$ at all time. This approximation is reasonable because $C_{\max}$ is large (4).

The ordinary differential equations we obtain from the unscaled DNA implementation [**S4–S9**] are as follows. For every species X_j:

$$\begin{aligned}[\dot{X_j}] = &- \sum_{i \in \mathscr{U} | r_i = j} \gamma^{-1} k_i [X_j] - \sum_{i \in \mathscr{B} | r_{i,1} = j} (\gamma^{-1} k_i C_{\max} [X_j] - q_{\max} C_{\max} [H_i]) \\ &- \sum_{i \in \mathscr{B} | r_{i,2} = j} q_{\max} [X_j][H_i] + \sum_{i \in \mathscr{U} \cup \mathscr{B}} c_{i,j} q_{\max} C_{\max} [O_i] \\ &- \gamma^{-1} (\sigma - \sigma_j) C_{\max} [X_j] + q_{\max} C_{\max} [HS_j], \end{aligned} \quad \textbf{[S10]}$$

and

$$[\dot{HS_j}] = \gamma^{-1} (\sigma - \sigma_j) C_{\max} [X_j] - q_{\max} C_{\max} [HS_j]. \quad \textbf{[S11]}$$

For every unimolecular reaction $i \in \mathscr{U}$:

$$[\dot{O_i}] = \gamma^{-1} k_i [X_{r_i}] - q_{\max} C_{\max} [O_i]. \quad \textbf{[S12]}$$

For every bimolecular reaction $i \in \mathscr{B}$:

$$[\dot{O_i}] = q_{\max} [X_{r_{i,2}}][H_i] - q_{\max} C_{\max} [O_i], \quad \textbf{[S13]}$$

$$[\dot{H_i}] = \gamma^{-1} k_i C_{\max} [X_{r_{i,1}}] - q_{\max} C_{\max} [H_i] - q_{\max} [X_{r_{i,2}}][H_i]. \quad \textbf{[S14]}$$

By using Tikhonov's theorem (Theorem 9.1 of ref. 5), we will prove the following statement of convergence. Let $[X_j](t)$ be solutions to the DNA implementation [**S10–S14**] and $\widehat{[X_j]}(t)$ be solutions of the target abstract system of chemical reactions [**S3**]. We start at $t = 0$ with $[X_j](0) = \gamma^{-1} \widehat{[X_j]}(0)$ and zero concentrations of $[HS_j]$, $[O_i]$, and $[H_i]$. Let $0 < t_1 < t_2$ be two time points such that $\widehat{[X_j]}(t)$ is finite [some systems of chemical reactions can achieve infinite concentrations infinite time (e.g., the system $2X_1 \to 3X_1$)] on $t \in [0, t_2]$. Then there is a $C^*_{\max}$ such that $\forall\ C_{\max} > C^*_{\max}$, at any time point $t \in [t_1, t_2]$,

$$|[X_j](t) - \widehat{[X_j]}(t)| = O(1/C_{\max}). \quad \textbf{[S15]}$$

Note that, whereas the error of $[X_j]$ at a fixed time scales linearly with $1/C_{\max}$, the scaling of $C_{\max}$ as a function of time to achieve some desired error is dependent on the target system.

In order to apply singular perturbation theory, we first make the variable substitutions $o_i = C_{\max}[O_i]$ and $x_j = [X_j] + \sum_{i \in \mathscr{B} | r_{i,1} = j} [H_i] + [HS_j]$ in [**S10–S14**]. Then we can write our system of differential equations in standard form for the singular perturbation method with small parameter $\epsilon = 1/C_{\max}$, and with x_j as the slow variables, and $[X_j]$, $[HS_j]$, o_i, and $[H_i]$ as the fast variables:

$$\begin{aligned}\dot{x_j} = &- \sum_{i \in \mathscr{U} | r_i = j} \gamma^{-1} k_i [X_j] - \sum_{i \in \mathscr{B} | r_{i,1} = j} q_{\max} [X_{r_{i,2}}][H_i] \\ &- \sum_{i \in \mathscr{B} | r_{i,2} = j} q_{\max} [X_j][H_i] + \sum_{i \in \mathscr{U} \cup \mathscr{B}} c_{i,j} q_{\max} o_i, \end{aligned} \quad \textbf{[S16]}$$

$$\begin{aligned}\epsilon[\dot{X_j}] = &-\epsilon \sum_{i \in \mathscr{U} | r_i = j} \gamma^{-1} k_i [X_j] - \sum_{i \in \mathscr{B} | r_{i,1} = j} (\gamma^{-1} k_i [X_j] - q_{\max} [H_i]) \\ &- \epsilon \sum_{i \in \mathscr{B} | r_{i,2} = j} q_{\max} [X_j][H_i] + \epsilon \sum_{i \in \mathscr{U} \cup \mathscr{B}} c_{i,j} q_{\max} o_i \\ &- \gamma^{-1} (\sigma - \sigma_j) [X_j] + q_{\max} [HS_j], \end{aligned} \quad \textbf{[S17]}$$

$$\epsilon[\dot{HS_j}] = \gamma^{-1} (\sigma - \sigma_j) [X_j] - q_{\max} [HS_j], \quad \textbf{[S18]}$$

for every $i \in \mathscr{B}$:

$$\epsilon \dot{o_i} = \gamma^{-1} k_i [X_{r_i}] - q_{\max} o_i, \quad \textbf{[S19]}$$

and for every $i \in \mathscr{B}$:

$$\epsilon \dot{o_i} = q_{\max} [X_{r_{i,2}}][H_i] - q_{\max} o_i, \quad \textbf{[S20]}$$

$$\epsilon[\dot{H_i}] = \gamma^{-1} k_i [X_{r_{i,1}}] - q_{\max} [H_i] - \epsilon q_{\max} [X_{r_{i,2}}][H_i]. \quad \textbf{[S21]}$$

The adjoined system is constructed from the differential equations for the fast variables [**S17–S21**] by making the substitution $\epsilon d(\cdot)/dt = d(\cdot)/d\tau$ with a new time variable τ, setting $\epsilon = 0$ on the right-hand side, and fixing the slow variables x_j. After simplification we obtain the following adjoined system:

$$d[X_j]/d\tau = q_{\max} x_j - q_{\max} \gamma^{-1} [X_j], \quad \textbf{[S22]}$$

$$d[HS_j]/d\tau = \gamma^{-1} (\sigma - \sigma_j) [X_j] - q_{\max} [HS_j], \quad \textbf{[S23]}$$

for every $i \in \mathscr{U}$:

$$do_i/d\tau = \gamma^{-1} k_i [X_{r_i}] - q_{\max} o_i, \quad \textbf{[S24]}$$

and for every $i \in \mathscr{B}$:

$$do_i/d\tau = q_{\max}[X_{r_{i,2}}][H_i] - q_{\max}o_i, \quad \textbf{[S25]}$$

$$d[H_i]/d\tau = \gamma^{-1}k_i[X_{r_{i,1}}] - q_{\max}[H_i]. \quad \textbf{[S26]}$$

To obtain Eq. **S22** note that $q_{\max}[X_j] + \sum_{i\in\mathscr{B}|r_{i,1}=j}\gamma^{-1}k_i[X_j] + \gamma^{-1}(\sigma - \sigma_j)[X_j] = q_{\max}[X_j] + \gamma^{-1}[X_j]\sigma = q_{\max}\gamma^{-1}[X_j]$, first by using the definition $\sigma_j = \sum_{i\in\mathscr{B}|r_{i,1}=j}k_i$ and then $\gamma^{-1} = q_{\max}(q_{\max} - \sigma)^{-1}$.

In the adjoined system **[S22–S26]** the differential equations for the different formal reactions i are decoupled because x_j are fixed. The adjoined system has the following steady-state values:

$$\widehat{[X_j]} = \gamma x_j, \quad \textbf{[S27]}$$

$$\widehat{[HS_j]} = q_{\max}^{-1}(\sigma - \sigma_j)x_j, \quad \textbf{[S28]}$$

for every $i \in \mathscr{U}$:

$$\widehat{o_i} = q_{\max}^{-1}k_i x_{r_i}, \quad \textbf{[S29]}$$

for every $i \in \mathscr{B}$:

$$\widehat{o_i} = \gamma q_{\max}^{-1}k_i x_{r_{i,1}}x_{r_{i,2}}, \quad \textbf{[S30]}$$

$$\widehat{[H_i]} = q_{\max}^{-1}k_i x_{r_{i,1}}. \quad \textbf{[S31]}$$

Because the different reactions i are decoupled, it can be easily shown that the adjoined system variables $[X_j]$, $[HS_j]$, o_i, and H_i converge to their steady-state values exponentially quickly, uniformly in x_j. Specifically, there exist positive constants d_1 and d_2 such that for any positive x_j,

$$|[X_j](\tau) - \widehat{[X_j]}| \le |[X_j](0) - \widehat{[X_j]}| \cdot d_1 e^{-d_2\tau}$$

with equivalent bounds holding for $[HS_j]$, o_i, and $[H_i]$. This convergence fulfills the nontrivial precondition of Tikhonov's theorem (Theorem 9.1 of ref. 5). (The continuity preconditions and the uniqueness of the steady state of the adjoined system, as well as the uniqueness of the solution of the reduced system, can be trivially established.) The reduced system can now be defined by assuming that the fast variables are always at their steady-state value, including in the differential equations for the slow variables. Using steady-state values **[S27–S31]** in **[S16]** results in the reduced system for $[X_j]$ of the desired target form **[S3]**. Then Tikhonov's theorem implies error bounds **[S15]**.

Additional Information about the Examples. In this section we provide additional information on the example CRNs and their DNA implementations from the main text. For each example, the complete target reaction network together with the corresponding DNA reactions is presented in the accompanying *Mathematica* notebook available at http://dna.caltech.edu/SupplementaryMaterial/DNA_for_CRNs/.

Lotka–Volterra oscillator. Figs. S2 and S3 expand on the Lotka—Volterra oscillator example from the main text, Fig. 5*B*. By using the domain notation of Figs. 1–4 of the main text, Fig. S2 diagrams the DNA molecules that must be prepared and Fig. S3 illustrates the resulting strand displacement reactions.

Oregonator Limit-Cycle Oscillator. The Oregonator limit-cycle oscillator is a classic simplified model of the Belousov–Zhabotinsky reaction (6). We used unscaled rate constants corresponding to kinetic parameters of ref. 6, equation (II), with the following modifications: (*i*) k_{M5} was increased to 446 to speed up the oscillations; f was decreased to 0.994 to accommodate oscillations in this regime (see ref. 6, figure 6). (*ii*) A and B (corresponding to the reservoir concentration of BrO_3^-) were increased to 0.065 from 0.06 M to speed up the approach to the limit cycle. The target formal CRN, rate constants, and the reactions corresponding to our DNA implementation are shown in Fig. S4.

Rössler Chaotic System. The chaotic system of Willamowsky and Rössler is a classic example of chaos under mass-action kinetics (7). The formal reactions and unscaled rate constants were taken from ref. 8. The target formal CRN, rate constants, and the reactions corresponding to our DNA implementation are shown in Fig. S5.

2-bit Pulse Counter Digital Circuit. Formal chemical kinetics can emulate digital logic circuits in numerous ways (e.g., refs. 9 and 10). For this example, we constructed AND gates as shown in Fig. 6*C* of the main text on the basis of a dual-rail signal representation: Every signal line corresponds to two molecular species representing its 0 and 1 value, respectively (10). Any of the inputs to an AND gate can be negated simply by reversing the roles of the two signal species corresponding to this input. Every AND gate was augmented with a dynamic threshold to sharpen the on/off behavior of the output signal, preventing the buildup of error (10). Such gates can be composed in arbitrary feed-forward and feedback configurations of their electronic counterparts. Two edge-triggered master–slave flip-flops were constructed in the standard manner and composed to produce a ripple counter (11). More direct chemical implementations of flip-flops and latches are possible that do not reduce to digital logic gates (10). The resulting chemical system consists of 182 reactions between 118 species. See the accompanying *Mathematica* notebook for the complete CRN.

Chemical systems whose behavior does not depend on small variations in reaction rates, but rather effectively on the possibility or impossibility of certain reactions, do not require buffering modules in their DNA implementation for qualitatively correct behavior. Digital circuits constructed as above are examples of such systems. Omitting buffering modules may significantly simplify their implementations. Further, recall that the rate constants in the DNA implementation are scaled to ensure $q_i, qs_j \le q_{\max}$, where q_i is the strand displacement rate constant controlling the rate of reaction i and qs_j is the strand displacement rate constant controlling the buffering of species X_j. If $\max_j\{qs_j\} \gg \max_i\{q_i\}$, the dynamics of the system may be sped up significantly by excluding the buffering modules and requiring only that $q_i \le q_{\max}$. The accompanying *Mathematica* notebook includes two versions of the 2-bit pulse counter construction: with buffering modules (plotted in Fig. 6*C* of the main text) and without.

Incrementer State Machine. Uniform computation models such as register machines constitute a contrasting computation style to digital circuits (12). Uniform computation in chemical systems has been explored recently, mostly for stochastic chemical kinetics (13, 14). The state machine example shown in Fig. 6*D* of the main text is based on these constructions. The clock governing the state transitions of the machine is a generalization of the oscillator from ref. 15) and was suggested to us by D. Dotty (personal communication) for the mass-action implementation of state machines. Although only 3 clock states are shown in Fig. 6*D* of the main text, the implemented clock cycles through 12 states to ensure the temporal separation of the 3 "active" states. The

ideal CRN continues to increment the number of green peaks until it eventually errs. The maximum count attained before erring is determined by the initial concentration of the "off-cycle" clock species. The CRN consists of 27 reactions between 25 species. See the accompanying *Mathematica* notebook for the complete CRN.

As with the previous example, the DNA implementation does not require buffering modules for qualitatively correct behavior (see accompanying *Mathematica* notebook).

Discussion of Experimental Considerations. ***Sequence design.*** We require the design of a unique species identifier for each formal species as well as one or more unique long domains for each reaction, with sequences distinct enough to effectively eliminate undesired reactions. There appear to be enough such sequences for implementing even large formal CRNs: The number of long domain sequences satisfying combinatorial criteria intended to ensure specific strand displacement reactivity appears to grow exponentially with domain length, with over 3,000 sequences available for typical lengths (16). This sequence design problem is made easier because in our construction only "positive" long domains (e.g., 1, 2, 3, etc.) appear single-stranded, while all "negative" long domains (e.g., 1^*, 2^*, 3^*, etc.) appear only in double-stranded form. (The designs in ref. 17, but not those of ref. 3, also had this property.) Thus, rather than the stringent "strong sequence design" criteria that every domain x binds exclusively with x^*, it is sufficient to ensure that no two positive domains can hybridize in solution and that no positive domain other than x can effect strand displacement of a x:x^* double-stranded domain. The former criterion can be addressed by designing long domains using a three-letter alphabet, such as just A, T, and C, a simple technique that has proved effective in minimizing unwanted interactions (17). The latter criterion can be addressed by designing long domains such that every pair of distinct domains has a sufficient number and distribution of mismatches; because branch migration proceeds only if the long domain matches (18, 19), a relatively small number of mismatches is expected to be sufficient (16).

Whereas the number of distinct toehold sequences is limited because these sequences must be short—and thus a large network must necessarily reuse toeholds between species—it is straightforward to modify our construction so that some or all toehold sequences are identical. Because branch migration cannot complete and result in strand displacement unless the long domain sequences match, off-pathway binding due to reused toehold sequences will result only in transient unproductive associations that may slightly slow down the kinetics but will not result in erroneous strand displacement operations.

Whereas these heuristic principles appear to be effective in practice so far, DNA sequence design strategies remain an active area of research (20, 21), and a more rigorous approach would be desirable to establish the true constraints sequence design places on strand displacement cascade scalability.

Granularity of effective rate constants. Because we set rate constants by the binding strength of toehold domains, and there are a finite number of short sequences available for toeholds, a limited number of exact rate constants can be achieved this way—although they are distributed over many orders of magnitude. However, there are many other degrees of freedom available, such as the relative concentrations of complexes L_i and B_i for bimolecular reactions. Practically, toehold domains would likely provide an order-of-magnitude control over formal rate constants, whereas rates would be fine-tuned by adjusting auxiliary species concentrations.

Network architecture optimization. Whereas we provide a systematic algorithm for compiling a set of chemical reactions into DNA, in practice it may often be preferable to reduce the complexity by optimizing the construction for the particular system of interest. For example, in many cases complete sequence independence between strands may be unnecessary, and some T_i gates may be eliminated. Although impossible in general, sometimes certain formal species X_j may be mapped directly to DNA complexes G_i rather than signal strands for a more concise implementation of some bimolecular reactions. Such optimizations will often be desirable for practical implementation.

Leak reactions and flow reactors. Reactions involving multistranded DNA complexes are often leaky; i.e., an output is released even if no input is present. This leak is not yet fully understood and could be because of poorly synthesized or assembled molecules and/or because of unintended 3-way or 4-way branch migration pathways. With appropriate purification of the DNA molecules, the rate of such leaky output production is many orders of magnitude smaller than that for the desired reactions (17, 22). Even so, if the leak process involves very high-concentration complexes such as the auxiliary complexes in our construction, it may strongly affect the overall reaction kinetics. It is unavoidable that such effects will lead to more rapid signal degradation than expected from an idealized theoretical model. Still, a naive experimental implementation today of our construction for the Lotka–Volterra system would likely yield several oscillations before running down. Furthermore, running times could be extended significantly by future advances that reduce leak with redesigned molecular motifs or compensation reactions, whereas running times could be extended indefinitely by providing auxiliary complexes at low concentrations in a continuous-flow stirred-tank reactor (23). Because lower auxiliary complex concentrations would be necessary in a flow reactor, much faster dynamics could also be obtained by scaling up the concentrations of signal species.

1. Kim J, Hopfield JJ, Winfree E (2004) Neural Network Computation by in Vitro Transcriptional Circuits. *Neural Information Processing Systems*, eds Saul LK, Weiss Y, Bottou L (MIT Press, Cambridge, MA), Vol 17, pp 681–688.
2. Zhang DY, Winfree E (2009) Control of DNA strand displacement kinetics using toehold exchange. *J Am Chem Soc* 131(47):17303–17314.
3. Seelig G, Soloveichik D, Zhang DY, Winfree E (2006) Enzyme-free nucleic acid logic circuits. *Science* 314(5805):1585–1588.
4. Klonowski W (1983) Simplifying principles for chemical and enzyme reaction kinetics. *Biophys Chem* 18(2):73–87.
5. Khalil HK (1996) *Nonlinear Systems* (Prentice Hall, Englewood Cliffs, NJ).
6. Field RJ, Noyes RM (1974) Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction. *J Chem Phys* 60:1877–1884.
7. Willamowski KD, Rössler OE (1980) Irregular oscillations in a realistic abstract quadratic mass action system. *Z Naturforsch A* 35:317–318.
8. Gaspard P (2005) *Encyclopedia of Nonlinear Science* (Routledge, London), pp 808–811.
9. Hjelmfelt A, Weinberger ED, Ross J (1991) Chemical implementation of neural networks and Turing machines. *Proc Natl Acad Sci USA* 88(24):10983–10987.
10. Magnasco MO (1997) Chemical kinetics is Turing universal. *Phys Rev Lett* 78(6):1190–1193.
11. Marcovitz AB (2004) *Introduction to Logic Design* (McGraw–Hill, New York), 2nd Ed.
12. Sipser M (1996) *Introduction to the Theory of Computation* (PWS, Boston).
13. Liekens AML, Fernando CT (2007) *Turing Complete Catalytic Particle Computers*, Lecture Notes in Computer Science (Springer, Berlin), Vol 4648, pp 1202–1211.
14. Soloveichik D, Cook M, Winfree E, Bruck J (2008) Computation with finite stochastic chemical reaction networks. *Nat Comp* 7(4):615–633.
15. Lachmann M, Sella G (1995) The computationally complete ant colony: Global coordination in a system with no hierarchy. *Advances in Artificial Life*, Lecture Notes in Computer Science (Springer, Berlin), Vol 929, p 784–800.
16. Qian L, Winfree E (2009) A simple DNA gate motif for synthesizing large-scale circuits. *DNA Computing 14*, Lecture Notes in Computer Science (Springer, Berlin), Vol 5347, pp 70–89.
17. Zhang DY, Turberfield AJ, Yurke B, Winfree E (2007) Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* 318(5853):1121–1125.
18. Panyutin IG, Hsieh P (1993) Formation of a single base mismatch impedes spontaneous DNA branch migration. *J Mol Biol* 230(2):413–424.
19. Panyutin IG, Hsieh P (1994) The kinetics of spontaneous DNA branch migration. *Proc Natl Acad Sci USA* 91(6):2021–2025.
20. Tulpan D, et al. (2005) Thermodynamically based DNA strand design. *Nucleic Acids Res* 33(15):4951–4964.

21. Kitajima T, Takinoue M, Shohda K, Suyama A (2008) Design of code words for DNA computers and nanostructures with consideration of hybridization kinetics. *Lect Notes Comp Sci* 4848:119–129.
22. Seelig G, Yurke B, Winfree E (2006) Catalyzed relaxation of a metastable DNA fuel. *J Am Chem Soc* 128(37):12211–12220.
23. Epstein IR, Pojman JA (1998) *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford Univ Press, London).

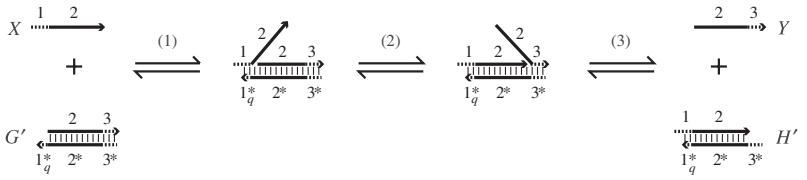

Fig. S1. The strand displacement molecular primitive with a partially complementary toehold. This figure is the analog of Fig. 1 of the main text, showing complex G with a truncated toehold domain 1^*. The effective forward strand displacement rate constant of $X + G' \rightarrow Y + H'$ is reduced compared to the full complement, with the subscript q on partial toehold 1^*_q indicating the resulting rate constant. Alternatively, domain 1^*_q may be of the same length as 1 but contain some mismatching bases.

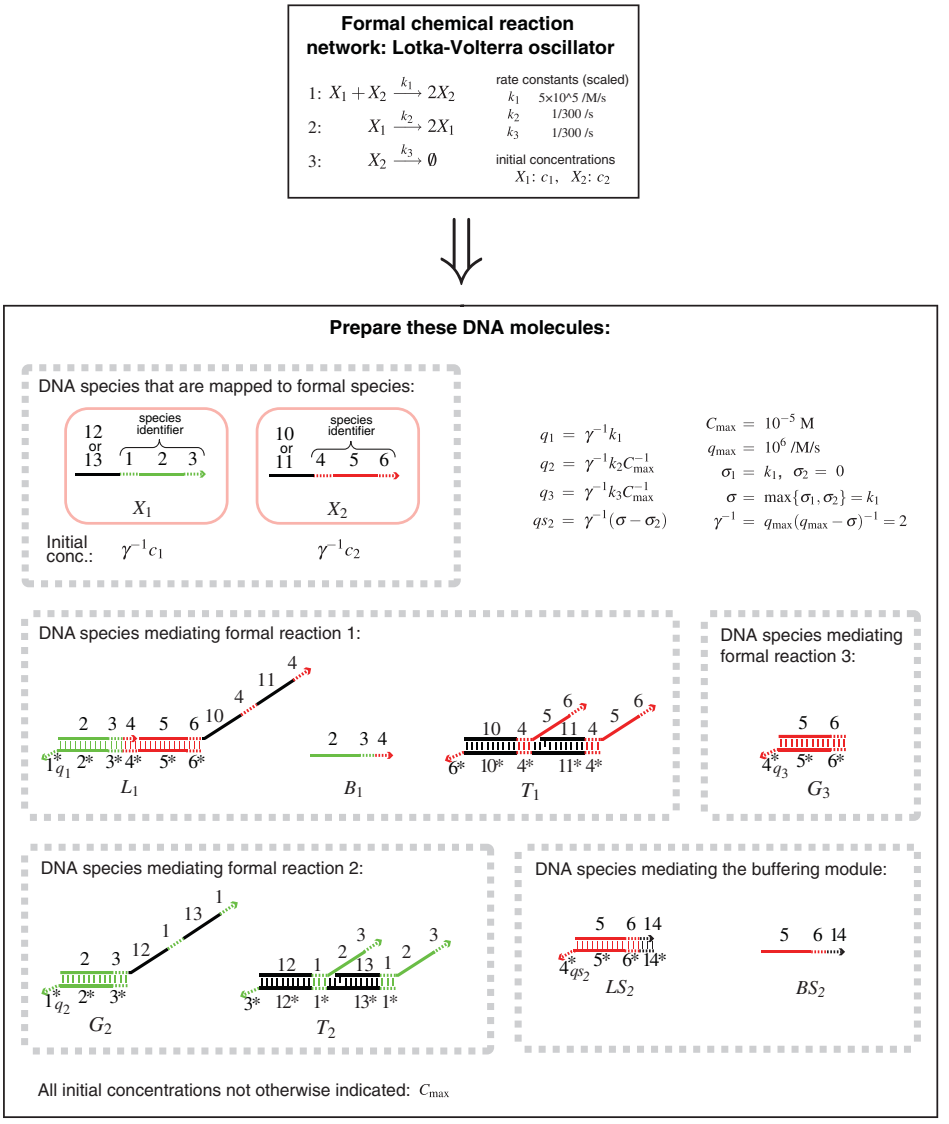

Fig. S2. In order to implement the Lotka–Volterra oscillator as described in the main text, we need to prepare the indicated DNA species. The strand displacement reactions corresponding to the desired formal reactions are shown in Fig. S3.

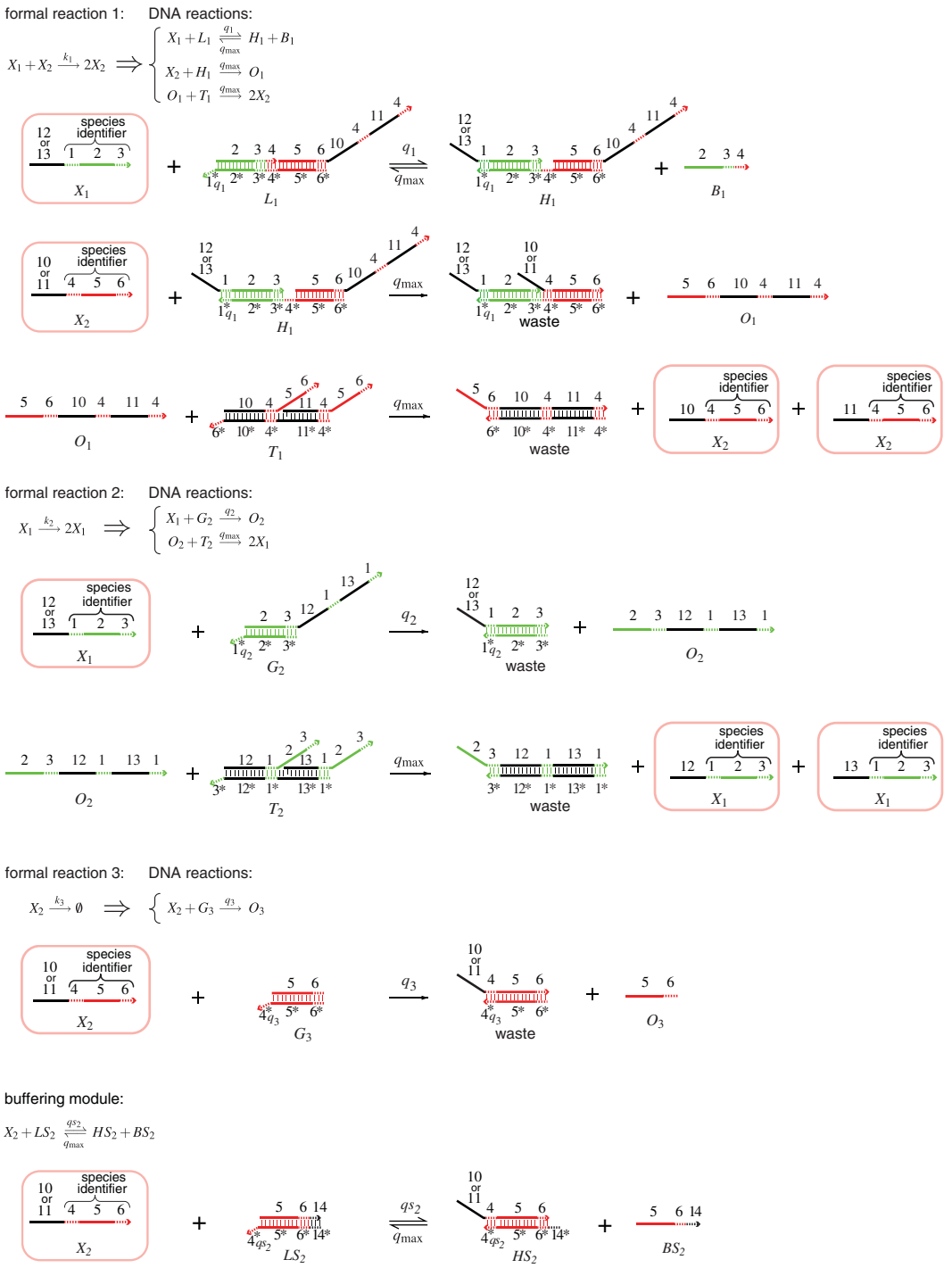

Fig. S3. DNA strand displacement reactions corresponding to the DNA implementation of the Lotka–Volterra oscillator shown in Fig. 5 of the main text and Fig. S2.

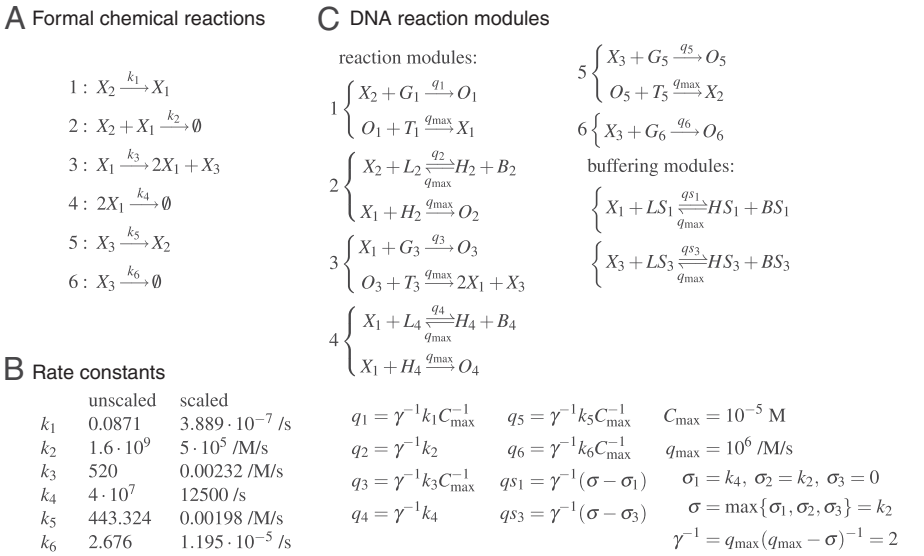

Fig. S4. Details of the implementation of the Oregonator limit-cycle oscillator shown in main text, Fig. 6*A*.

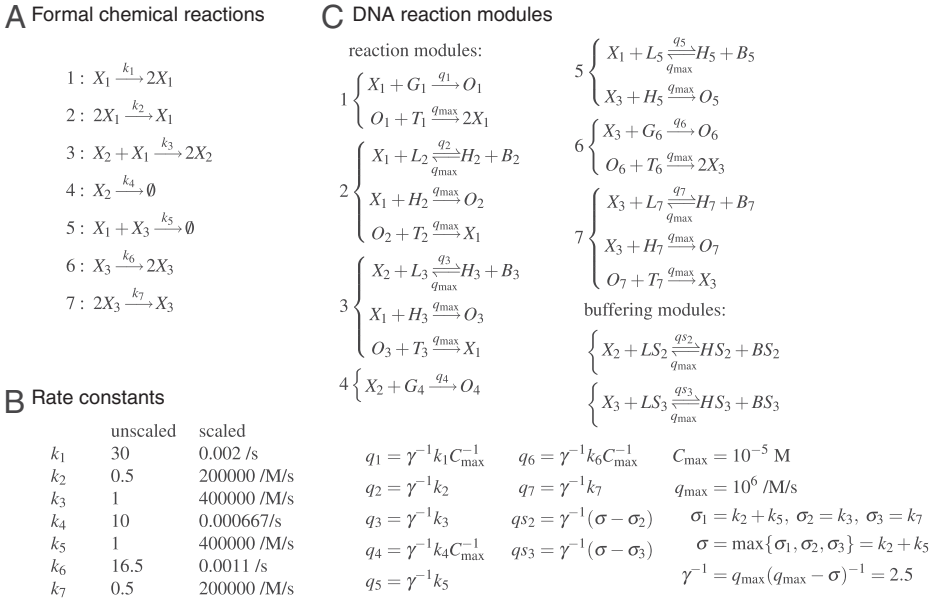

Fig. S5. Details of the implementation of the Rössler chaotic system shown in main text, Fig. 6*B*.