Enzyme-free nucleic acid dynamical systems

Niranjan Srinivas,[1*] James Parkin,[2] Georg Seelig,[3]
Erik Winfree,[1,2,4] and David Soloveichik[5*]

[1]Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125,
[2]Bioengineering, California Institute of Technology, Pasadena, CA 91125,
[3]Electrical Engineering, University of Washington, Seattle, WA 98195,
[4]Computer Science, California Institute of Technology, Pasadena, CA 91125,
[5]Electrical and Computer Engineering, UT Austin, 1616 Guadalupe St., Austin, TX 78701.
*To whom correspondence should be addressed;
E-mail: niranjan@dna.caltech.edu, david.soloveichik@utexas.edu

Chemistries exhibiting complex dynamics—from inorganic oscillators to gene regulatory networks—have been long known but cannot be reprogrammed at will because of a lack of control over their evolved or serendipitously found molecular building blocks. Here we show that information-rich DNA strand displacement cascades could be systematically constructed to realize complex temporal trajectories specified by an abstract chemical reaction network model. We codify critical design principles in a compiler that automates the design process, and demonstrate our approach by building a novel DNA-only oscillator. Unlike biological networks that rely on the sophisticated chemistry underlying the central dogma, our test tube realization suggests that simple Watson-Crick base pairing interactions alone suffice for arbitrarily complex dynamics. Our result establishes a basis for autonomous and programmable molecular systems that interact with and control their chemical environment.

Embedded information processing circuitry provides a powerful means for creating highly functional autonomous systems. For electromechanical machines, the past century has experienced a revolutionary advance in technological capability due to embedded control. Within living organisms, embedded control is at the heart of all cellular processes, and is often seen as the distinguishing feature between living and non-living chemistries. However, in principle, non-biological chemical systems are also capable of information processing that directs molecular behaviors. Inspired by the success of systematic approaches in electrical engineering, we seek molecular building-blocks and design rules for combining them, to systematically construct non-biological autonomous molecular systems—an approach we might call molecular programming.

Any candidate architecture for engineering chemical controllers must be capable of diverse dynamical behaviors. Since the discovery of well-mixed chemical oscillators (*1, 2*), synthetic reaction networks with complex temporal dynamics have been engineered based on small-molecule interactions, such as redox chemistries (*3*). More recently, the information-based chemistry underlying the central dogma of molecular biology has been used to create a variety of dynamical systems, such as bistable switches and oscillators in living cells (*4–7*) and in simplified cell-free systems (*8–14*) that involve a limited number of enzymes. However, the range of dynamical behaviors demonstrated by synthetic systems does not yet approach the complexity and sophistication of biological circuits (*15, 16*).

Systematically engineering a wide range of dynamical behaviors would be greatly facilitated by a "programming language" for composing relatively simple molecular building-blocks into complex dynamic networks. The language of formal chemical reaction networks (CRNs)—i.e. chemical reaction equations (with rate constants) between "formal" symbols representing species—provides a natural abstraction for specifying the diverse dynamical behaviors possible with mass-action chemical kinetics (*16–18*). Indeed, formal CRNs can be constructed to simulate arbitrary polynomial differential equations (*19, 20*), linear feedback controllers (*21*), boolean logic circuits (*22*), neural networks (*23*), distributed algorithms (*24*), and other computational models (*25*).

Dynamic DNA nanotechnology (*26*) offers an attractive molecular architecture for engineering CRNs with desired dynamic behavior. Indeed, the programmable nature of DNA-DNA interactions mediated by Watson-Crick complementarity, coupled with predictive thermodynamic models (*27*) makes it possible to rationally design molecular reaction pathways. In particular, toehold-mediated strand displacement (*28–30*) has been exploited to engineer nanoscale tweezers (*31*), enzyme-free digital logic circuits (*32, 33*), catalytic networks (*34, 35*), and dynamically self-assembled structures (*35, 36*).

Inspired by the simple yet powerful rules governing strand displacement reactions, general schemes for translating any formal CRN into a "DNA implementation" have been proposed (*37, 38*). In principle, given a CRN with formal species $\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots$, a set of DNA molecules may be designed to approximate the specified mass action kinetics with arbitrary accuracy (up to scaling rate constants and concentrations). Recently, Chen et al. (*39*) have used a general CRN-to-DNA scheme (*38*) to engineer a consensus network that compares the concentration of two DNA strands and con-

verts the "majority" into the "totality". Since the objective of the consensus network is a desired steady state, it remained unclear what new challenges would arise when designing dynamical behaviors.

Here, we demonstrate a general molecular technology for engineering enzyme-free nucleic acid dynamical systems. Designing complex temporal trajectories, rather than endpoint computations, places stringent requirements on kinetic design (Fig. 1A). As a challenging test case, we experimentally realize the rock-paper-scissors oscillator—a CRN that has been explored as a mathematical construct in theoretical biology (*40, 41*), ecology (*42, 43*), and molecular programming (*44*), but without any experimental realization in nonlinear chemistry, synthetic biology, or DNA nanotechnology.

We identify several critical design principles, both at the domain-level and at the sequence-level, and acquire improved understanding of molecular non-idealities, including means of mitigating and compensating for imperfect execution of desired reactions and spurious "leak" reactions. Our mechanistic model with measured kinetic parameters provides a "proof-by-synthesis" that designed molecular interactions are indeed sufficient for programming mass-action kinetics. Lastly, we implement a CRN-to-DNA "compiler" that incorporates our design principles: given a formal CRN, it automates the design process to provide candidate DNA sequences for implementing the desired dynamical behavior. We found that DNA sequences designed by our automated compiler led to oscillatory dynamics with no further optimization, thereby reducing our design time from 4 years to 4 weeks; we believe this tool would facilitate the use of our general technology for engineering other dynamical behaviors with DNA strand displacement.

CRN to DNA implementation scheme. Given a dynamical behavior as specified by a formal CRN program, we aim to systematically design a DNA-based implementation that approximates the specified behavior in a test tube (Fig. 1A). Each formal species in the CRN program is represented by single-stranded DNA species (signal strands). These signals are designed not to interact directly with each other. Instead, each reaction in the formal CRN is mediated by additional DNA species, including fuel species that provide both logic and free-energy for the desired reaction to occur (e.g. React complexes) and intermediate species that link the consumption of reactants and the release of products (Flux strands). Waste complexes generated as byproducts of desired pathways are designed to be incapable of further strand displacement. This strategy is modular since implementing one more reaction only requires adding the corresponding set of fuel species. The fuel species are initially present in large excess, and the DNA implementation is guaranteed to approximate the dynamics specified by the formal CRN program as long as the fuel species remain sufficiently in excess (*37*).

Our experiments have been performed in a one-pot batch-reactor, without any flow of matter or energy; therefore, the dynamics of our test-tube realizations are expected to deviate from the specified dynamics once a significant fraction of fuel species have been consumed.

The fundamental building block underlying our scheme is toehold-mediated DNA strand displacement (*28–30*), which is composed into interaction cascades. The logical design

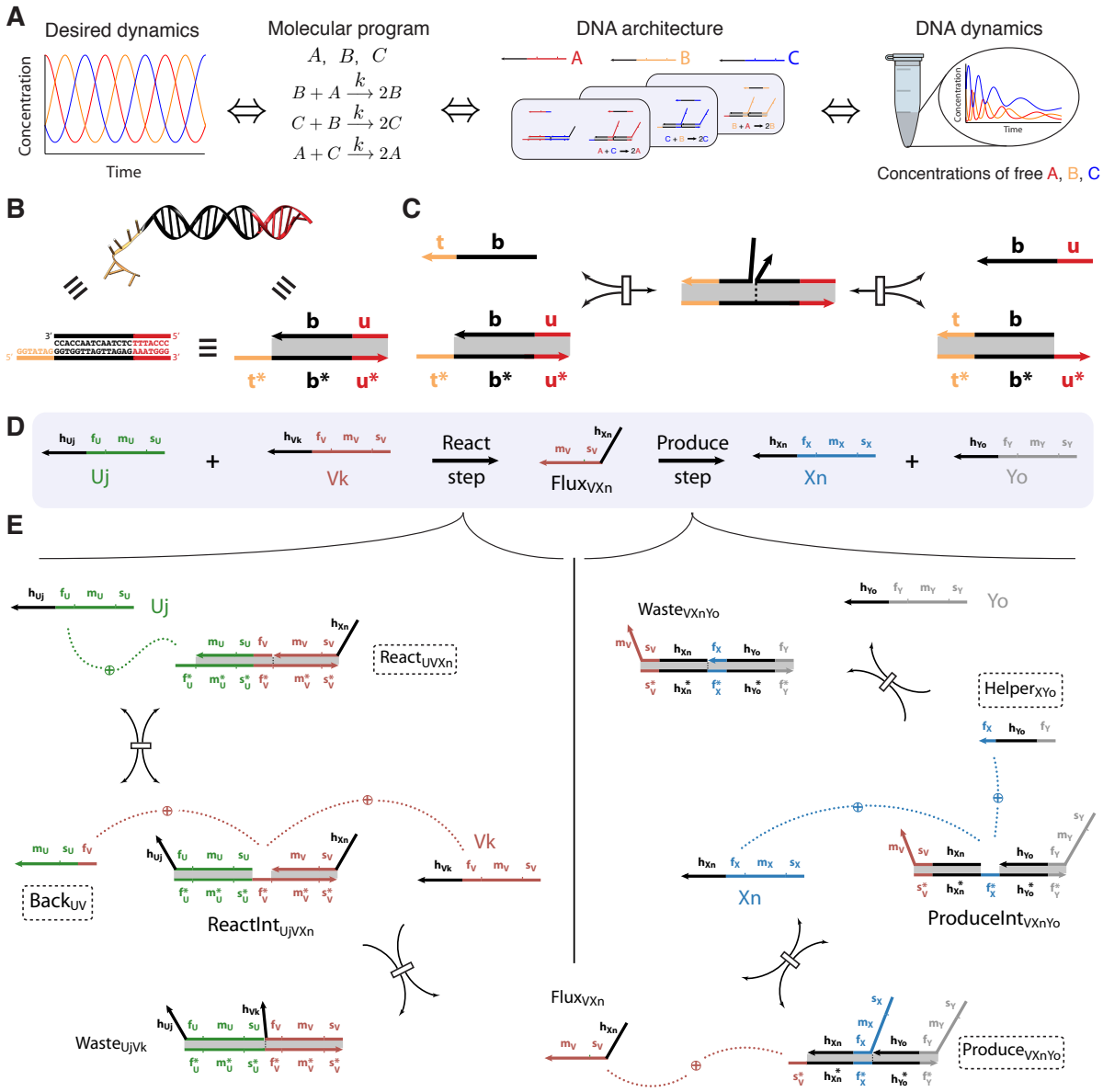

Fig. 1. (**A**) A systematic pipeline for engineering dynamical systems with DNA strand displacement. The dynamics of our closed batch reactor approximates that of the prescribed CRN as long as fuel species are in large excess. (**B**) Domain-level abstraction of a multi-strand DNA complex. Gray rectangle indicates double helix; arrows indicate 3' ends; * denotes Watson-Crick complementarity. (**C**) Reversible toehold exchange: fleeting toehold-binding facilitates strand exchange via three-way branch migration. (**D**) Implementation of the general bimolecular reaction $\mathrm{U} + \mathrm{V} \rightarrow \mathrm{X} + \mathrm{Y}$ occurs in two steps, react and produce, mediated by an intermediate Flux strand. (**E**) Individual strand displacement or toehold exchange reactions within the react and produce steps are mediated by fuel species (indicated by dashed boxes). Dotted lines illustrate toehold binding and dissociation interactions. Note that although one history domain is shown for the input signals (h_{U_j} and h_{V_k}), equivalent reactions occur with input strands with other history domains.

of strand displacement cascades is facilitated by abstracting DNA sequences into contiguous *domains* that are intended to act together as a unit (Fig. 1B). Short (5-7 nt) single-stranded domains called toeholds fleetingly co-localize the competing strands and facilitate the intramolecular exchange of base pairs (branch migration); eventually, one of the competing strands dissociates (Fig. 1C). Longer (13-25 nt) domains, called branch migration domains, bind strongly enough that spontaneous dissociation does not occur. Specificity is achieved by the choice of DNA sequence; many orthogonal strand displacement interactions can occur simultaneously in the same solution as long as sequence overlap is minimized.

Each formal species (e.g. U) is represented by a set of signal strands $(\mathrm{Ui}, \mathrm{Uj}, \ldots)$ that share three domains in common: the first toehold ($\mathrm{f_U}$), the branch migration domain ($\mathrm{m_U}$) and the second toehold ($\mathrm{s_U}$). All desired strand displacement interactions involving a given formal species (U) occur with these three domains. Additional "history" domains $(\mathrm{h_{Ui}}, \mathrm{h_{Uj}}, \ldots)$ are specific to the different signals $(\mathrm{Ui}, \mathrm{Uj}, \ldots)$ based on the location in the Produce complex that originally sequestered these strands. However, the history domains are intended to be inert; they merely facilitate the formation of the desired structure when the strands comprising the Produce complexes are annealed (Fig. S1).

The lack of direct interactions between signal strands makes it possible to independently translate formal reactions into modular sets of fuel species, which can then be simply combined to implement any desired formal CRN. As illustrated for the general bimolecular reaction $\mathrm{U} + \mathrm{V} \rightarrow \mathrm{X} + \mathrm{Y}$ in Fig. 1DE, the implementation module is conceptually divided into two steps: (1) the react step, mediated by React and Backward fuels, recognizes and consumes the reactants U and V as input, and (2) the produce step, mediated by Produce and Helper fuels, releases the products X and Y as output. These steps are linked by a Flux strand, released in the first step, that triggers the second step.

The mechanistic implementation of the react step begins with the React complex ($\mathrm{React_{VUXn}}$) reversibly consuming a signal strand representing the first formal reactant (U), releasing the Backward strand ($\mathrm{Back_{UV}}$). Because the React complex and Backward strand are both in excess (both are fuels), the resulting intermediate complex ($\mathrm{ReactInt_{VUXn}}$) will approach a pseudo-equilibrium proportional to the concentration of U. Thus, $\mathrm{ReactInt_{VUXn}}$ will interact with the second input, V, at a rate proportional to the product of their concentrations, in accordance with standard mass-action chemical kinetics for bimolecular reactions. This reaction irreversibly releases the Flux strand ($\mathrm{Flux_{VUn}}$).

In the subsequent produce step, the Flux strand initiates a pathway that releases products X and Y, which are initially bound to the Produce complex ($\mathrm{Produce_{VXnYo}}$) by their history domains and first toeholds. After the Flux strand releases the first output, a toehold is exposed that allows the Helper strand ($\mathrm{Helper_{XYo}}$) to irreversibly displace the second output. Since the history domains are specific to this formal reaction, the Flux and Helper strands between different reactions do not cross-react. Further, because each forward reaction is driven by a fuel species that is in high concentration, the produce step reactions are fast relative to the rate-limiting react step, which therefore determines the overall path-

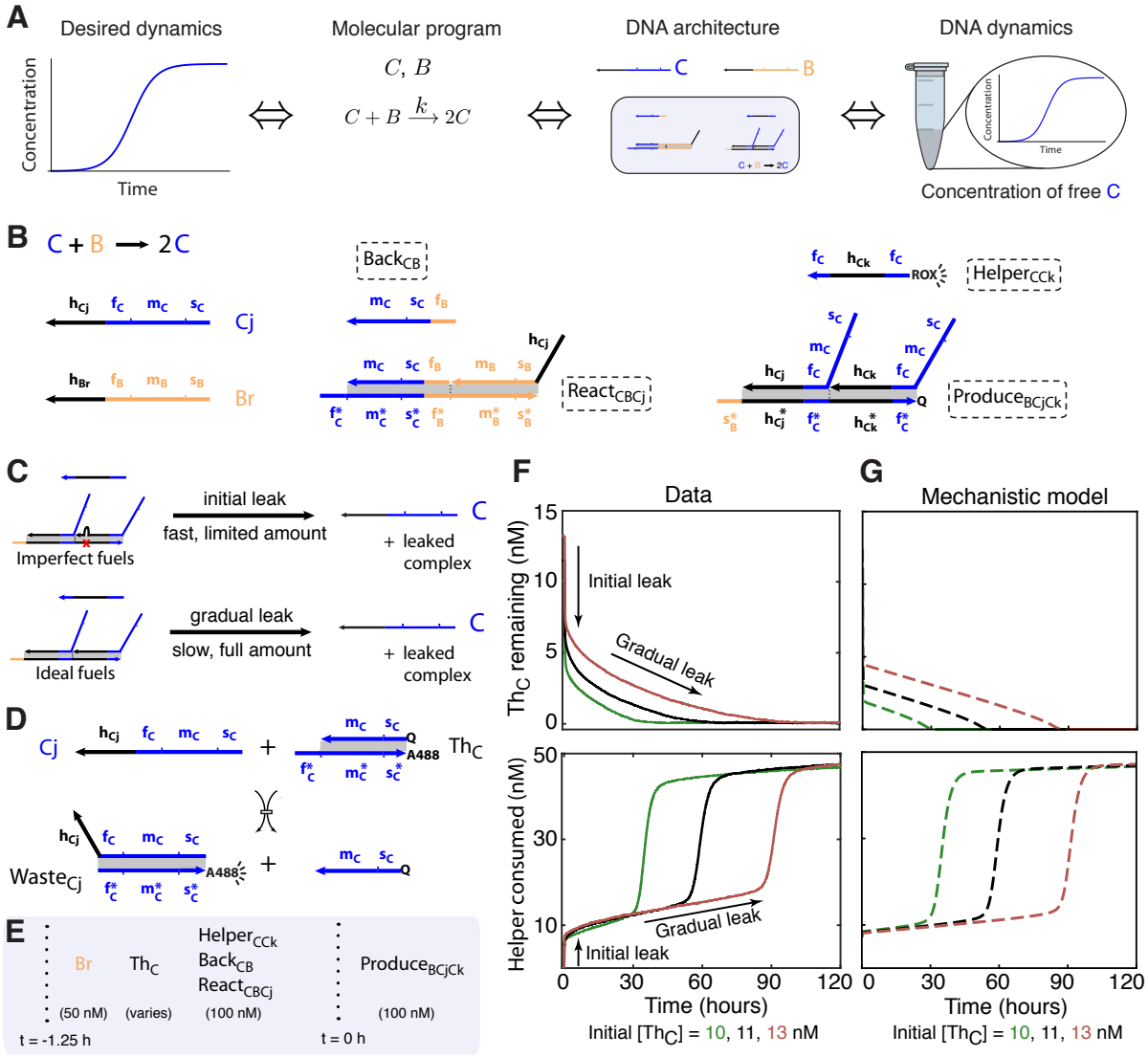

Fig. 2. (**A**) Schematic for engineering a single-reaction CRN with exponential amplification using our systematic pipeline. (**B**) Domain-level illustration of the DNA species involved (fuel species indicated by dashed boxes). (**C**) A limited amount of imperfect fuel molecules, such as those with DNA synthesis errors, release signal strands and waste products through fast spurious pathways ("initial leak"). Ideal fuel molecules release similar products through slow "gradual" leak. (**D**) A Threshold complex ($\mathrm{Th_C}$) is designed to consume leaked autocatalyst. (**E**) Experimental setup. Vertical dotted lines separate initial contents of the test tube and timed additions. Addition of Produce complexes kickstarts release of autocatalyst through initial and gradual leak. (**F**) Experimental data showing concentration of $\mathrm{Th_C}$ (top) and the amount of $\mathrm{Helper_{CCk}}$ consumed (bottom) for three independent samples with differing initial amounts of $\mathrm{Th_C}$. The progress of the reaction is monitored via fluorophores on the Helper and Threshold species shown in (B) and (D). (**G**) Mechanistic model semi-quantitatively captures the dynamics of the DNA implementation (Note S5).

way kinetics.

Our CRN-to-DNA scheme is fully general. We can construct reactions with repeated reactants or products (e.g. autocatalysis), and a different number of reactants and products (e.g. unimolecular reactions can be obtained by declaring one reactant a fuel species). Our naming scheme is both precise and general—the name and the molecule fully determine each other (Fig. S2), which facilitates automated and systematic analysis (Note S6).

Non-idealities in a single-reaction CRN. To understand the challenges in using our general CRN-to-DNA scheme for engineering dynamical behaviors, we begin with the autocatalytic single-reaction CRN $\mathrm{C} + \mathrm{B} \rightarrow 2\mathrm{C}$ (Figs. 2AB). Since exponential amplification kinetics is sensitive to both the initial concentration of the autocatalyst and the rate constants of the reactions involved, it is a stringent test of our ability to control dynamics. We obtained the domain-level specification by replacing $\mathrm{U}, \mathrm{V}, \mathrm{X}$, and Y in Fig. 1E by C, B, C, and C respectively (Figs. 2B, S3). To obtain a molecular implementation, we performed sequence design as described in the following section.

The success of an experimental realization is determined by how well the system behaves in accordance with the domain-level model. DNA strand displacement systems suffer from three main classes of molecular non-idealities: output strands can be released when they should not be ("leak"), input strands can be consumed without producing output ("substoichiometric yield"), and reactions can proceed at the wrong rate.

We observe both a limited amount of fast "initial leak", as well as a slower "gradual leak" that continues throughout the duration of the experiment (Fig. 2C,F, S16). Initial leak is thought to arise primarily from a fraction of imperfectly prepared fuel molecules—e.g. due to synthesis errors (truncations or deletions) in individual strands, or improperly folded multi-stranded complexes—that, when initially mixed together, can readily interact and release their outputs. In contrast, gradual leak cannot be avoided even with perfectly synthesized and folded molecules, since it arises from the inherent biophysics of strand displacement (*30*). Mechanistically, these reactions could initiate from invasion at the end of a helix (blunt-end) or a coaxial junction even in the absence of a toehold (*45*) (Fig. S6), or could be facilitated by spurious remote toeholds (*46*) (Fig. S19, S20).

Substochiometric yield can arise as a consequence of leak pathways: both initial and gradual leak may result in reactive complexes that can consume inputs without releasing outputs, because the outputs have already been released (Fig. S7). Substoichiometric yield may also result from other synthesis errors—e.g. truncations on toehold regions of output strands may render them nonfunctional for triggering downstream reactions (Fig. S8).

Desired reactions can take place with markedly different kinetics due to sequence differences, even when the domain-level descriptions are identical. This is partly due to the exponential dependence of strand displacement rate constants on toehold length and binding energy (*28, 29*), but can also be affected by undesired secondary structure within signal strands as well as fleeting binding between toeholds and unrelated single-stranded portions of unrelated molecules, both of which can occlude the toehold and inhibit the desired reaction.

Both types of leak are clearly evident in

the experimental implementation of the single-reaction autocatalytic CRN that converts an initial reservoir of B into C, with the fuels in excess. To prevent the immediate onset of exponential amplification due to initial leak of the autocatalyst, C, we also introduced a fast Threshold complex ($\mathrm{Th_C}$) that consumes C (Fig. 2D). Assuming that the initial Threshold concentration is greater than the initial leak of C, exponential amplification is delayed until further gradual leak of C eventually exhausts the Threshold. At this point, further gradual leak triggers exponential amplification of C, by the implemented reaction $\mathrm{C} + \mathrm{B} \rightarrow 2\mathrm{C}$, until the provided quantity of B is fully consumed. The progress of the reaction is monitored via fluorophores on the Threshold and Helper species (Note S4.1), exhibiting the expected proportional delays for three different initial amounts of Threshold (Fig. 2F). Qualitatively similar delayed amplification was also seen in two other single-reaction autocatalytic CRNs, $\mathrm{A} + \mathrm{C} \rightarrow 2\mathrm{A}$ and $\mathrm{B} + \mathrm{A} \rightarrow 2\mathrm{B}$ (Figs. S4, S5, S11, S12). The three modules had different initial and gradual leak rates (Fig. S14, Table S3, and Note S7.2), resulting in a roughly 10-fold variation in the delay times (Fig. S13).

Evidence for substoichiometric yield and non-ideal reaction rates can also be seen in Figs. 2F, S13 (e.g. the autocatalytic phase does not consume the full complement of 50 nM of Helper, and the amplification rates are not equal for the different delays), but the clearest evidence comes from measurements of individual steps of the reaction pathways. For analogous reactions, rate constants were all within a factor of 20 of each other, and mostly within a factor of 3 (Tables S1, S2). Substoichiometric yields on the order of 20% lower than ideal were observed (Fig. S22).

The essential features of the autocatalytic dynamics were captured by a quantitative mechanistic model at the level of individual strand displacement reactions (Fig 2G). For each of the three autocatalytic CRNs, all 7 rate constants for reactions shown in Figs. 1E and Fig. 2D were measured separately (Tables S1 and S2). In addition to the 21 independently measured rate constants for the desired reaction pathways, the model partially accounts for observed non-idealities (Note S5). The Produce-Helper leak (Fig. S6) was determined to be the dominant pathway and included in the model; the 3 rate constants for gradual leak were inferred from the 3 autocatalytic module experiments (Fig. S14, Table S3, and Note S7.2), while the amount of initial leak is assumed to be the same for all modules, and specified by a single parameter. Substoichiometric yield resulted from interactions with the products of leak reactions (Fig. S7) as well as an assumption that a fraction of output strands (Flux and Signal molecules) were non-functional (Fig. S8, the same fraction in all cases). Finally, to account for the concentration-dependent slow-down of toehold-mediated strand displacement due to toehold occlusion (*33*), we introduce a single parameter for the strength of toehold binding in all unproductive situations (e.g. Helper binding to the React complex, Fig. S9).

The three empirical parameters (for initial leak, non-functional output, and toehold occlusion) were fit to the full oscillator data of Figure 4B, including an additional fitting parameter for each initial signal concentration to account for imperfect pipetting and uncertainties in the initial leak. Removing any of these non-idealities from the mechanistic model could not adequately explain the data. The same three empirical parameter values were then

used when modeling the three single-reaction CRNs, using only three newly-fit parameters for the Threshold concentrations, to account again for uncertainties in the initial leak. We interpret the success of the mechanistic model to mean that we have captured the dominant effects in our system, and we expect that similar models will have considerable predictive power for future strand displacement systems.

Principles for robust molecular design. The previously discussed understanding of non-idealities in strand displacement cascades was refined over the course of four iterations of sequence design for the three autocatalytic modules, in parallel with the development of sequence design principles and experimental methods that help minimize the non-idealities (Notes S3.1, S3.2, S3, S4). The performance of Design 4, presented above, is an improvement over each previous design (Note S4).

Initial leak was reduced by several design choices and experimental methods. First, the baseline sequence design criterion was sequence symmetry minimization (*47*), which unlike purely thermodynamic approaches (*48*) is expected to help the folding process avoid being kinetically trapped in malformed conformations (*49*). Second, fuel complexes were prepared by annealing HPLC-purified oligonucleotides, followed by PAGE gel purification to minimize undesired multimers and excess single-strands (*50*). Third, because the orientation of bases on the DNA backbone ($5'$-$3'$) is known to affect the distribution of synthesis errors (*51*), we tried both orientations and achieved a three-fold reduction in initial leak by using the backbone orientation in which the toehold occurs on the $5'$ end (Note S3.5). Lastly, Threshold complexes can be used to tune the initial conditions by removing leaked signal strands from solution.

Since gradual leaks primarily arise from strand displacement through invasion at frayed blunt ends and coaxial junctions, we used 2-nt clamps at the end of React and Produce complexes and closed helices and coaxial junctions with strong (C/G) base pairs (Fig. S15). Further, we minimized spurious remote-toehold strand displacement (*46*) by avoiding even relatively weak complementarity at overhangs near coaxial junctions (Fig. 3). These strategies reduced gradual leaks as much as 15-fold relative to earlier designs (Table S5, Fig. S16, Fig. S18).

Three key design strategies were used to minimize undesired variability in rates. First, signal strands contained at most one G, reducing the propensity for undesired secondary structure; in particular, we tested for and removed secondary structure in certain single-stranded regions that are crucial for initiating strand displacement, such as toeholds and the first 3-4 bases of the branch migration domain (*30*). Secondly, toeholds were designed to be isoenergetic according to nearest-neighbor parameters (*27*) augmented with terms for coaxial stacking and protruding tails at nicks (*29, 30*). Because the same toeholds are used in different contexts—with different flanking structures and thus different energetics—we truncated 1 or 2 nucleotides in some cases to help equilibrate binding energy (specifically in the reversible toehold exchange step in the React complex pathway, which is expected to be rate determining, Fig. S22). Finally, toeholds were simultaneously designed to be as orthogonal as possible, and branch migration domain sequences were designed to be orthogonal to toeholds, in order to mitigate toehold-occlusion and the concomitant slowdown in kinetics.

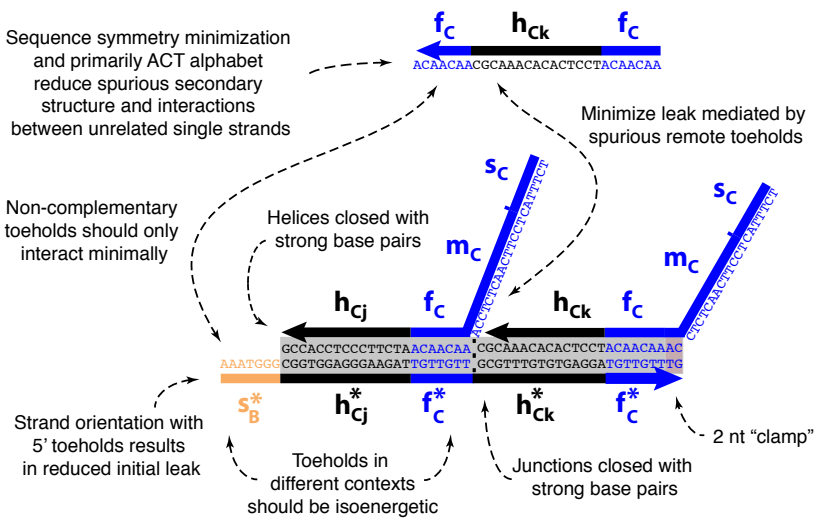

Fig. 3. Sequence design principles illustrated with a Produce complex ($\mathrm{Produce}_{\mathrm{BCjCk}}$).

To combat substoichiometric yield, we designed a modified Helper strand that, in addition to displacing the second output in the produce step, also displaces the original Flux strand that initiated the produce step, effectively enabling catalytic action by the Flux strand (Fig. S10). This "catalytic Helper" permits the Flux strand to release more outputs by initiating displacement with another Produce molecule, thereby increasing effective reaction stoichiometry. By tuning the relative concentration of the catalytic Helper strands, we may adjust reaction stoichiometry—much like the use of potentiometers for tuning local resistance in early electrical circuits.

The design principles discussed in this section involve an unusual combination of thermodynamic, kinetic, and ad-hoc criteria, which were not compatible with straightforward application of state-of-the-art sequence design tools (*52*). Therefore, custom heuristic measures for comparing candidate sequence designs were formulated and implemented as a collection of scripts that called NUPACK (*52*), Pepper (*53*), StickyDesign (*54*), and SpuriousSSM (*53*), in order to perform sequence design and analysis (Note S3). Because the three autocatalytic modules were intended to work together as an oscillator, as described in the next section, they were designed together as a single system.

A DNA strand displacement oscillator. The strand displacement oscillator (which we call the Displacillator, Fig. 4A) realizes the three reaction rock-paper-scissors CRN (*40–44*). A neutral cycle oscillator, its orbit is determined by the conservation laws for two quantities: $\mathrm{A}+\mathrm{B}+\mathrm{C}$ and $\mathrm{A}^{\mathrm{k}_2/\mathrm{k}_0}\cdot\mathrm{B}^{\mathrm{k}_3/\mathrm{k}_0}\cdot\mathrm{C}^{\mathrm{k}_1/\mathrm{k}_0}$, where $\mathrm{k}_\mathrm{i}/\mathrm{k}_0$ are unitless rate constants (*42*). Unlike most limit cycle oscillators, the rock-paper-scissors CRN oscillates for any choice of reaction rate constants and (non-steady-state) initial signal concentrations, which makes it especially suitable for implementation as a dynamical strand displacement cascade. The number of oscillation periods expected before fuels are exhausted

decreases with the amplitude of the oscillation. Thus, to reduce the amplitude resulting from initial leak, we added (more than) enough Threshold to consume all the released signal strands, yielding a quiescent metastable fuel mixture. The Displacillator was then kick-started by addition of signal strands whose effective initial concentrations were reduced by the uncertain residual Threshold amounts. The concentration of catalytic Helpers, added to compensate for substoichiometric yield, was empirically tuned to 25% of total Helper concentration. (Fig. S23 illustrates experimental set-up and calibration.)

Observing system dynamics by directly measuring free signal strand concentrations may consume or temporarily sequester a fraction of these strands. Instead, to avoid interfering with the dynamics, the net progress of each reaction can be measured by the quenching of fluorophore labeled Helper strands. Oscillatory dynamics could be clearly observed in the instantaneous consumption rates of the Helper strands (Figs. 4BC, S24), until the fuel species were depleted. The order in which the reaction rates peak and trough was consistent with the ideal rock-paper-scissors dynamics, for each of the 3 initial concentrations of signal species.

The mechanistic model (see above and Note S5) demonstrates that the emergent dynamics of the reaction mixture may be quantitatively explained by the individual strand displacement interactions that we designed, and the non-idealities that we understand. The mechanistic model was able to account for most of the measured Helper consumption dynamics, including the eventual slowdown due to fuel depletion (Fig. 4B). The model further allowed us to extrapolate signal $(\mathrm{A}, \mathrm{B}, \mathrm{C})$ concentrations that were not directly accessible to measurement (Fig. 4D). Although oscillations in reaction rates were observed via directly measured Helper fluorescence, the extrapolated signals allow tying $\mathrm{A}, \mathrm{B}, \mathrm{C}$ signal dynamics back to the ideal rock-paper-scissors CRN (Note S5.5, Fig. S26). This agreement with the design specification—the formal CRN—confirms that the Displacillator oscillates for the reasons that we intended.

A CRN-to-DNA compiler. In principle, our design strategies could be used to construct an automated pipeline for implementing any formal CRN. To do so, we integrated the sequence design tools and principles discussed previously into an end-to-end compiler, called Piperine, that accepts an arbitrary formal CRN as input and produces candidate sequences for experimental implementation. Piperine's sequence design pipeline proceeds in several stages: first, the formal CRN is translated into a set of requests for the necessary fuel molecules using the Pepper DNA design specification language (*53*); second, Pepper uses templates for each type of fuel strand or fuel complex to deduce the full set of strands and base-pairing constraints; third, toeholds are designed to be orthogonal and isoenergetic using StickyDesign (*54*); fourth, the toehold sequences and base-pairing constraints are sent to SpuriousSSM (*53*), which uses sequence symmetry minimization to obtain sequences for the long domains; fifth, proposed sequence sets are scored using heuristic criteria that make use of NUPACK (*52*) to evaluate key secondary structure and spurious binding interactions; finally, multiple independent sequence designs are compared according to these criteria, and the sequence set that scores best across the board is recommended (Note S6.3).

In order to test the compiler, we used it to

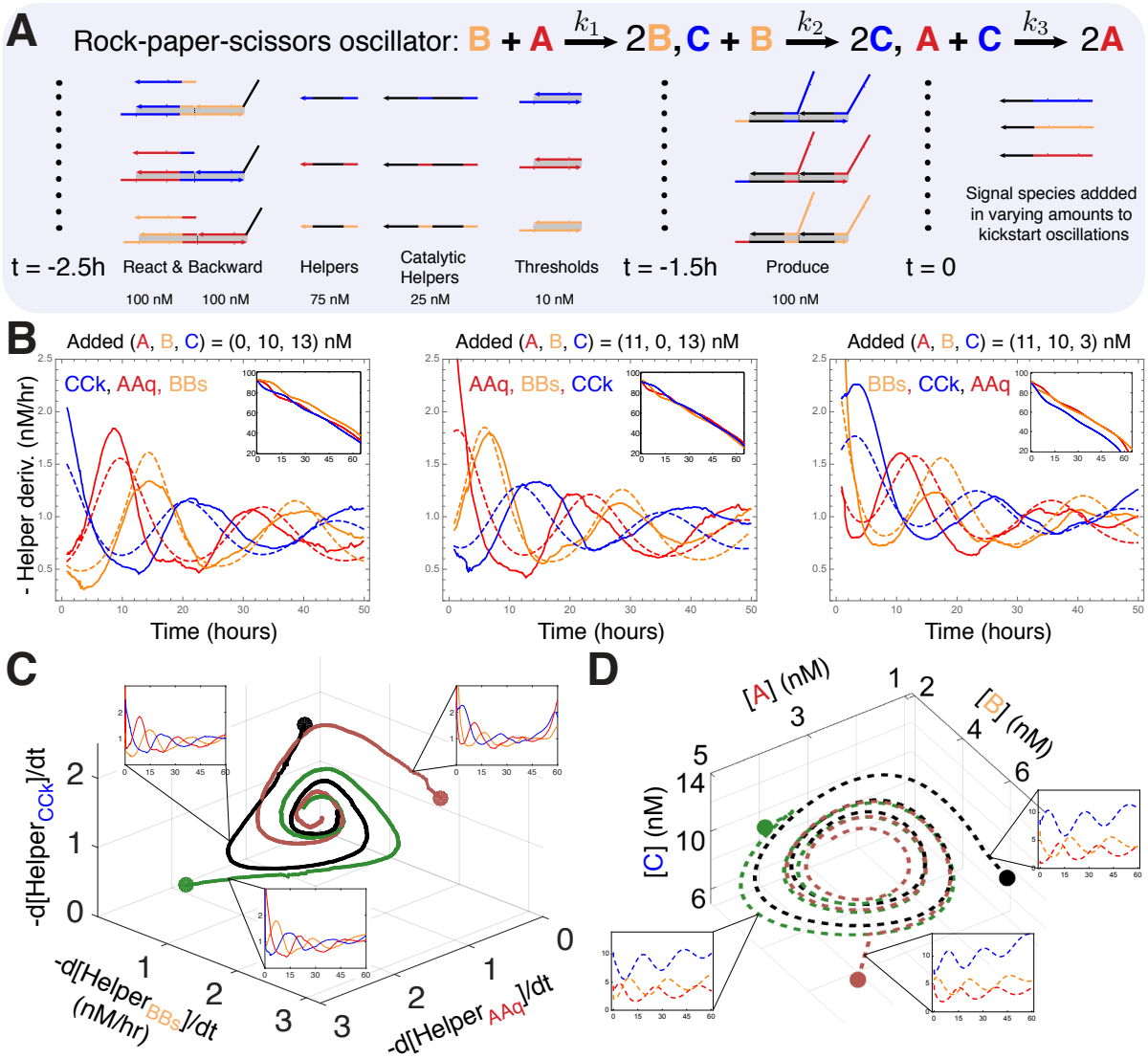

Fig. 4. (**A**) Experimental scheme for engineering the Displacillator. Vertical dotted lines separate initial contents of the test tube and timed additions. (**B**) Experimental data (solid lines) and mechanistic model fits (dashed lines) show time derivatives of the concentrations of the three Helper strands under three different initial conditions. Insets display measured Helper concentrations. (**C**) Phase plot of the experimental data shown in (**B**). Thick dots indicate initial conditions. Insets show time traces for each trajectory, as in (**B**). (**D**) Phase plot of the concentrations of the signal strands extrapolated from the mechanistic model. Insets show time traces of the signal concentrations for each trajectory.

design, from scratch, another instance of the Displacillator with completely independent sequences. We achieved a dramatic reduction in the time from initial design to observation of oscillatory behavior, from 4 years to 4 weeks, with all autocatalytic modules and the full oscillator working on the very first experiment (Fig. S33, Note S7). Applied to other CRNs of comparable size, it is reasonable to expect that Piperine will produce sequences that perform comparably well for implementing other dynamical systems. More generally, it would be straightforward to augment Piperine to compile CRNs using other translation schemes that have been proposed (*37–39, 44, 55, 56*). Indeed, the core sequence design principles used here heuristically (e.g. Fig. 3) could form the starting point for the development of rigorous sequence design methods, incorporating both thermodynamic and kinetics constraints, for an even wider variety of strand displacement cascades.

Conclusions. The development of programmable molecular technologies will require systematic architectures and automated design software. Our demonstration of a chemical oscillator using just DNA strand displacement cascades prototypes such a general technology for chemical dynamical systems. We expect that our molecular design principles and experimental methods can be generalized to implement any desired chemical kinetics, up to scaling of rate constants and concentrations. It is remarkable that such a wide range of dynamical behaviors appears attainable by utilizing no more than the principles of Watson-Crick base pairing. The well-understood molecular mechanisms underlying DNA strand displacement (*28–30*) permit detailed mechanistic design of reaction pathways, which in turn enables quantitative modeling at the level of individual strand displacement reactions.

Dynamical systems (including oscillators) instantiated in biochemistry and programmed by the choice of DNA sequence have at least a 20 year history (*8, 9, 58, 59*). A key distinguishing feature of our simple DNA architecture is that it requires no enzymes or other "black box" components that have not been rationally designed. As a concrete example, it is instructive to compare our strand displacement oscillator to other recent synthetic biochemical oscillators (Table 1). The "genelet" architecture (*9, 10, 60*) simplifies genetic regulatory networks (GRNs) by avoiding protein synthesis and using RNA to directly regulate transcription from short DNA templates; it relies on two essential enzymes, an RNA polymerase and a ribonuclease. The PEN toolbox architecture (*58*) goes further by also eliminating RNA altogether, using just a DNA polymerase, an exonuclease, and a nickase. Finally, cell-free transcription-translation (TX-TL) architectures (*59*) are sufficient for implementing many GRNs without the full complexity of living cells; whether derived from cell extract or reconstituted from purified components (*61*), over 100 essential components are involved (polymerases, ribosomes, tRNA, tRNA synthetases, amino acids, NTPs, etc). In each of these architectures, a wide variety of circuits can be implemented by introducing suitably designed DNA molecules. For the reference oscillators, we employ the number of designed nucleotides as a simple metric for design size, and the total number of base-pairs of DNA that code for enzymes as a proxy for the extent of black-box genetic information (Table 1). By these metrics, the Displacillator has

Oscillator	Design size	Black-box size	Number of Cycles	Period
DSD: Displacillator	1386 nt	0	3 (batch)	~ 20 h
Genelets: Design I (*10*)	469 nt	~ 4000 bp†	6 (batch)	~ 3 h
PEN: Oligolator (*11*)	71 nt	~ 7700 bp‡	30 (batch)	~ 1.7 h
GRNs: Repressillator (*5, 57*)	6664 bp	$>$ 100 components	$\gg 9$ (microfluidics)	~ 3 h

Table 1. Comparison to other recent synthetic cell-free biochemical oscillators. † T7 RNA polymerase, E. coli Ribonuclease H, and pyrophosphatase; ‡ Bst DNA polymerase, RecJf exonuclease, and Nt.BstNBI nickase.

the greatest fraction of rationally designed material, as well as the lowest overall design complexity when black-box components are considered. However, its relatively poor performance highlights the remaining challenges for fully rationally designed biochemical dynamical systems.

There are currently many proposals, some partially demonstrated, for implementing CRNs with DNA (*37–39, 44, 55, 56*). Each scheme makes different choices regarding the representation of signals and the implementation of desired reactions, resulting in different molecule sizes, number of additional mediating species, lengths of reaction pathways, sequence design constraints, and potential for leak reactions. Currently, it is not clear how these schemes may be compared in terms of their potential for engineering arbitrary dynamical behaviors in the test tube. Improved understanding of the biophysics of initial and gradual leak pathways, and of the sequence-dependence of kinetics for fundamental DNA mechanisms such as hybridization, branch migration, fraying, and dissociation (*30, 62, 63*), should allow molecular systems to be designed with more accurate control over kinetics and with less leak. Indeed, certain CRN-to-DNA schemes may have orders-of-magnitude lower leak (*64*), raising the prospect that higher concentrations and thus faster kinetics could be achieved reliably. Finally, providing a continuous "power supply" by replenishing fuel species and removing waste molecules (as in a continuous-flow stirred reactor (*65*)) could enable faithful dynamics on longer time scales, such as those required for controlling self-assembly or chemical reactors.

Enabling the reliable and routine use of enzyme-free nucleic acid dynamical systems as embedded chemical controllers will require integrating nucleic acid subsystems with a broad range of other chemical processes. Strand displacement cascades already have enhanced potential for modular integration with the ever-expanding range of molecular structures, machines, and devices developed in DNA nanotechnology (*66, 67*). Furthermore, nucleic acids—both DNA and RNA—are well known for their ability to bind to and sense small molecules (*68, 69*), thus providing direct mechanisms to "read" the chemical environment. Nucleic acid nanotechnology has also been applied to control chemical synthesis (*70–72*); to control the arrangement (and rearrangement) of metal nanoparticles, quantum dots, carbon nanotubes, proteins, and other molecules (*73–77*); and to control the activity of enzymes and protein motors (*78–80*). Much as genetic regulatory networks and other biochemical feedback networks control chemical and molecular functions within biological cells, it is conceivable that nucleic acid dynamical systems could serve as the information processing and control networks within complex synthetic or-

ganelles or artificial cells (*81*) that sense, compute, and respond to their chemical and molecular environment.

Acknowledgments: The authors thank C. Evans for custom modifications to sequence design software, A. Phillips for the use of unpublished software, and C. Geary, C. T. Martin, N. A. Pierce, L. Qian, P. W. K. Rothemund, S. L. Sparvath, B. Wolfe, and D. Y. Zhang for helpful discussions. NS is currently at the Department of Molecular & Cell Biology at UC Berkeley, where he is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2259-16). This work was supported by NSF Grants 0728703, 0829805, 0832824, 1317694, 1117143, the Gordon and Betty Moore Foundation's Programmable Molecular Technology Initiative, and NIGMS Systems Biology Center grant P50 GM081879.

References and Notes

1. W. C. Bray. A periodic reaction in homogeneous solution and its relation to catalysis. *J. Am. Chem. Soc*, 43:1262–1267, 1921.

2. B. P. Belousov. Periodically acting reaction and its mechanism. *Collection of Abstracts on Radiation Medicine*, 147(145 (in Russian)), 1959.

3. IR Epstein and JA Pojman. *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*. Oxford Univ Press, London, 1998.

4. Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403:339–342, 2000.

5. Michael B. Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403:335–338, January 20, 2000.

6. M. R. Atkinson, M. A. Savageau, J. T. Myers, and A. J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *escherichia coli*. *Cell*, 113:597–607, 2003.

7. J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456:516–519, 2008.

8. B. Wlotzka and J.S. McCaskill. A molecular predator and its prey: Coupled isothermal amplification of nucleic acids. *Chem Biol*, 4(1):25–33, 1997.

9. J. Kim, K. S. White, and E. Winfree. Construction of an in vitro bistable circuit from synthetic transcriptional switches. *Molecular Systems Biology*, 2(68):1–12, 2006.

10. J. Kim and E. Winfree. Synthetic *in vitro* transcriptional oscillators. *Molecular Systems Biology*, 7(1):465, 2011.

11. K. Montagne, R. Plasson, Y. Sakai, T. Fujii, and Y. Rondelez. Programming an *in vitro* DNA oscillator using a molecular networking strategy. *Molecular Systems Biology*, 7(466), 2011.

12. Adrien Padirac, Teruo Fujii, and Yannick Rondelez. Bottom-up construction of in vitro switchable memories. *Proceedings of the National Academy of Sciences*, 109(47):E3212–E3220, 2012.

13. T. Fujii and Y. Rondelez. Predator-prey molecular ecosystems. *ACS Nano*, 7(1):27–34, 2013.

14. Sergey N Semenov, Albert SY Wong, R Martijn van der Made, Sjoerd GJ Postma, Joost Groen, Hendrik WH van Roekel, Tom FA De Greef, and Wilhelm TS Huck. Rational design of functional and tunable oscillating enzymatic networks. *Nature chemistry*, 7(2):160–165, 2015.

15. Dennis Bray et al. Protein molecules as computational elements in living cells. *Nature*, 376(6538):307–312, 1995.

16. John J. Tyson, Katherine C. Chen, and Bela Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15:221–231, 2003.

17. Fritz Horn and Roy Jackson. General mass action kinetics. *Archive for rational mechanics and analysis*, 47(2):81–116, 1972.

18. P. Érdi and J. Tóth. *Mathematical models of chemical reactions: Theory and applications of deterministic and stochastic models*. Princeton Univ. Press, 1989.

19. M. Korzuhin. *Oscillatory Processes in Biological and Chemical Systems*. Nauka, Moscow (in Russian), 1967.

20. W. Klonowski. Simplifying principles for chemical and enzyme reaction kinetics. *Biophysical chemistry*, 18:73–87, 1983.

21. K. Oishi and E. Klavins. A biochemical implementation of linear I/O systems. *IET Systems Biology*, 5(4):252–260, 2011.

22. A. Hjelmfelt and J. Ross. Implementation of logic functions and computations by chemical kinetics. *Physica D: Nonlinear Phenomena*, 84:180–193, 1995.

23. A. Hjelmfelt, E. D. Weinberger, and J. Ross. Chemical implementation of neural networks and turing machines. *PNAS*, 88:10983–87, 1991.

24. J. Aspnes and E. Ruppert. *An Introduction to Population Protocols*. Middleware for Network Eccentric and Mobile Applications. Springer-Verlag, 2009.

25. A. Hjelmfelt, E. D. Weinberger, and J. Ross. Chemical implementation of finite-state machines. *PNAS*, 89:383–87, 1992.

26. D. Y. Zhang and G. Seelig. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chem.*, 3(2):103–113, 2011.

27. J. SantaLucia, Jr. and D. Hicks. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33:415–40, 2004.

28. Bernard Yurke and Allen P. Mills, Jr. Using DNA to power nanostructures. *Genetic Programming and Evolvable Machines*, 4:111–122, 2003.

29. D.Y. Zhang and E. Winfree. Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc*, 131(47):17303–17314, 2009.

30. N. Srinivas, T. E. Ouldridge, P. Sulc, J. M. Schaeffer, B. Yurke, A. A. Louis, J. P. K. Doye, and E. Winfree. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.*, 41(22):10641–10658, 2013.

31. Bernard Yurke, Andrew J. Turberfield, Allen P. Mills, Jr., Friedrich C. Simmel, and Jennifer L. Nuemann. A DNA-fuelled molecular machine made of DNA. *Nature*, 406:605–608, 2000.

32. G. Seelig, D. Soloveichik, D. Y. Zhang, and E. Winfree. Enzyme-free nucleic acid logic circuits. *Science*, 314(5805):1585–1588, 2006.

33. L Qian and E. Winfree. Scaling up digital circuit computation with DNA strand displacement cascades. *Science*, 332(6034):1196–1201, 2011.

34. D. Y. Zhang, A. J. Turberfield, B. Yurke, and E. Winfree. Engineering entropy-driven reactions and networks catlayzed by DNA. *Science*, 318:1121–1125, 2007.

35. P. Yin, H.M.-T. Choi, C. R. Calvert, and N. A. Pierce. Programming biomolecular self-assembly pathways. *Nature*, 451:318–322, 2008.

36. John P Sadowski, Colby R Calvert, David Yu Zhang, Niles A Pierce, and Peng Yin. Developmental self-assembly of a DNA tetrahedron. *ACS nano*, 8(4):3251–3259, 2014.

37. D. Soloveichik, G. Seelig, and E. Winfree. DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci. U.S.A*, 107(12):5393–5398, 2010.

38. L. Cardelli. Two domain DNA strand displacement. *Mathematical Structures in Computer Science*, 23(02):247–271, 2013.

39. Yuan-Jyue Chen, Neil Dalchau, Niranjan Srinivas, Andrew Phillips, Luca Cardelli, David Soloveichik, and Georg Seelig. Programmable chemical controllers made from DNA. *Nature Nanotech.*, 8:755, 2013.

40. M. Lachmann and G. Sella. The computationally complete ant colony: Global coordination in a system with no hierarchy. *Lecture Notes in Computer Science*, 929:784–800, 1995.

41. Luca Cardelli. Artificial biochemistry. In *Algorithmic Bioprocesses*, pages 429–462. Springer, 2009.

42. T. Reichenbach, M. Mobilia, and E. Frey. Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model. *Phys. Rev. E*, 74(051907), 2006.

43. Alexander Dobrinevski and Erwin Frey. Extinction in neutrally stable stochastic lotka-volterra models. *Physical Review E*, 85(5):051903, 2012.

44. Matthew R Lakin, Simon Youssef, Luca Cardelli, and Andrew Phillips. Abstractions for dna circuit design. *Journal of The Royal Society Interface*, 9(68):470–486, 2012.

45. Luis P Reynaldo, Alexander V Vologodskii, Bruce P Neri, and Victor I Lyamichev. The kinetics of oligonucleotide replacements. *Journal of molecular biology*, 297(2):511–520, 2000.

46. A. J. Genot, D. Yu. Zhang, J. Bath, and A. J. Turberfield. Remote toehold: A mechanism for flexible control of DNA hybridization kinetics. *2011*, 133(7):2177–2182, JACS.

47. N.C. Seeman. De novo design of sequences for nucleic acid structural engineering. *J. Biomol. Struct. Dyn.*, 8(3):573–81, 1990.

48. J.N. Zadeh, B.R. Wolfe, and N.A. Pierce. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry*, pages 439–452, 2011.

49. R. M. Dirks, M. Lin, E. Winfree, and N. A. Pierce. Paradigms for computational nucleic acid design. *Nucleic Acids Res.*, 32(4):1392–1403, 2004.

50. Materials and methods are available as supplementary materials on Science Online.

51. D.Y. Zhang and E. Winfree. Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Research*, 38:4182–4197, 2010.

52. J.N. Zadeh, C.D. Steenberg, J.S. Bois, B.R. Wolfe, M.B. Pierce, A.R. Khan, R.M. Dirks, and N.A. Pierce. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, page published online, 2010.

53. `http://www.dna.caltech.edu/DNA_Sequence_Design_Tools/`.

54. Constantine G. Evans and Erik Winfree. DNA sticky end design and assignment for robust algorithmic self-assembly. *DNA Computing and Molecular Programming LNCS*, 8141:61–75, 2013.

55. L. Cardelli. Strand algebras for DNA computing. *Nat Comput*, 10:407–428, 2011.

56. L. Qian, D. Soloveichik, and E. Winfree. Efficient Turing-universal computation with DNA polymers. *DNA Computing and Molecular Programming*, pages 123–140, 2011.

57. Henrike Niederholtmeyer, Zachary Z Sun, Yutaka Hori, Enoch Yeung, Amanda Verpoorte, Richard M Murray, and Sebastian J Maerkl. Rapid cell-free forward engineering of novel genetic ring oscillators. *Elife*, 4:e09771, 2015.

58. Alexandre Baccouche, Kevin Montagne, Adrien Padirac, Teruo Fujii, and Yannick Rondelez. Dynamic DNA-toolbox reaction circuits: A walkthrough. *Methods*, 67(2):234–249, 2014.

59. Jonathan Garamella, Ryan Marshall, Mark Rustad, and Vincent Noireaux. The all E. coli TX-TL Toolbox 2.0: a platform for cell-free synthetic biology. *ACS synthetic biology*, 5(4):344–355, 2016.

60. Jongmin Kim, John Hopfield, and Erik Winfree. Neural network computation by in vitro transcriptional circuits. In *Advances in neural information processing systems*, pages 681–688, 2004.

61. Yoshiru Shimizu, Akio Inoue, Yukihide Tomari, Tsutomu Suzuki, Takashi Yokogawa, Kazuya Nishikawa, and Takuya Ueda. Cell-free translation reconstituted with purified components. *Nature Biotechnology*, 19:751–755, 2001.

62. Joseph Malcolm Schaeffer, Chris Thachuk, and Erik Winfree. Stochastic simulation of the kinetics of multiple interacting nucleic acid strands. In *International Workshop on DNA-Based Computers*, volume LNCS 9211, pages 194–211. Springer, 2015.

63. T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye, and A. A. Louis. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucl. Acids Res.*, 41(19):8886–8895, 2013.

64. Chris Thachuk, Erik Winfree, and David Soloveichik. Leakless DNA strand displacement systems. In *DNA Computing and Molecular Programming*, pages 133–153. Springer, 2015.

65. Irving R Epstein. Complex dynamical behavior in "simple" chemical systems. *The Journal of Physical Chemistry*, 88(2):187–198, 1984.

66. David Yu Zhang, Rizal F Hariadi, Harry MT Choi, and Erik Winfree. Integrating DNA strand-displacement circuitry with DNA tile self-assembly. *Nature communications*, 4, 2013.

67. Nadrian C Seeman. *Structural DNA Nanotechnology*. Cambridge University Press, 2016.

68. M.N. Stojanovic and D.M. Kolpashchikov. Modular aptameric sensors. *J. Am. Chem. Soc.*, 126(30):9266–9270, 2004.

69. Eun Jeong Cho, Joo-Woon Lee, and Andrew D Ellington. Applications of aptamers as sensors. *Annual Review of Analytical Chemistry*, 2:241–264, 2009.

70. Shiping Liao and Nadrian C. Seeman. Translation of DNA signals into polymer assembly instructions. *Science*, 306:2072–2074, 2004.

71. Yu He and David R Liu. Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker. *Nature nanotechnology*, 5(11):778–782, 2010.

72. Wenjing Meng, Richard A Muscat, Mireya L McKee, Phillip J Milnes, Afaf H El-Sagheer, Jonathan Bath, Benjamin G Davis, Tom Brown, Rachel K O'Reilly, and Andrew J Turberfield. An autonomous molecular assembler for programmable chemical synthesis. *Nature Chemistry*, 2016.

73. Hongzhou Gu, Jie Chao, Shou-Jun Xiao, and Nadrian C Seeman. A proximity-based programmable DNA nanoscale assembly line. *Nature*, 465(7295):202–205, 2010.

74. Jinglin Fu, Minghui Liu, Yan Liu, Neal W Woodbury, and Hao Yan. Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures. *Journal of the American Chemical Society*, 134(12):5516–5519, 2012.

75. Shawn M Douglas, Ido Bachelet, and George M Church. A logic-gated nanorobot for targeted transport of molecular payloads. *Science*, 335(6070):831–834, 2012.

76. Robert Schreiber, Jaekwon Do, Eva-Maria Roller, Tao Zhang, Verena J Schüller, Philipp C Nickels, Jochen Feldmann, and Tim Liedl. Hierarchical assembly of metal nanoparticles, quantum dots and organic dyes using DNA origami scaffolds. *Nature nanotechnology*, 9(1):74–78, 2014.

77. Chao Zhou, Xiaoyang Duan, and Na Liu. A plasmonic nanorod that walks on DNA origami. *Nature communications*, 6, 2015.

78. Nathan C Gianneschi and M Reza Ghadiri. Design of molecular logic devices based on a programmable DNA-regulated semisynthetic enzyme. *Angewandte Chemie International Edition*, 46(21):3955–3958, 2007.

79. Minghui Liu, Jinglin Fu, Christian Hejesen, Yuhe Yang, Neal W Woodbury, Kurt Gothelf, Yan Liu, and Hao Yan. A DNA tweezer-actuated enzyme nanoreactor. *Nature communications*, 4, 2013.

80. Adam J M Wollman, Carlos Sanchez-Cano, Helen M J Carstairs, Robert A Cross, and Andrew J Turberfield. Transport and self-organization across different length scales powered by motor proteins and programmed by DNA. *Nature Nanotechnology*, 9:44–47, 2014.

81. Vincent Noireaux, Yusuke T Maeda, and Albert Libchaber. Development of an artificial cell, from self-organization to computation and self-reproduction. *Proceedings of the National Academy of Sciences*, 108(9):3473–3480, 2011.