DOI: http://dx.doi.org/10.7551/978-0-262-31050-5-ch058

Modeling Scalable Pattern Generation in DNA Reaction Networks

Peter B. Allen, Xi Chen, Zack B. Simpson and Andrew D. Ellington*

University of Texas at Austin
Institute for Cell and Molecular Biology
andy.ellington@mail.utexas.edu

Abstract

We have developed a theoretical framework for developing patterns in multiple dimensions using controllable diffusion and designed reactions implemented in DNA. This includes so-called strand displacement reactions in which one single-stranded DNA hybridizes to a hemi-duplex DNA and displaces another single stranded DNA, reversibly or irreversibly. These reactions can be designed to proceed with designed rate and molecular specificity. By also controlling diffusion by partial complementarity to a stationary, cross-linked DNA, we can generate predictable patterns. We demonstrate this with several simulations showing deterministic, arbitrary shapes in space.

Introduction

Pattern formation is biologically and technologically important. Biomimetic methods for moving from top-down to bottom-up formation of designed patterns and materials have the potential to revolutionize manufacturing by dramatically reducing costs. These approaches include biomimetic molecular recognition(Chen et al. 2011) leading to self-assembled, folded structures made from block-copolymers,(Murnen et al. 2010) biopolymers(Rothemund 2006) or patterned microparticles. Yet none of these techniques have recapitulated the "algorithmic" assembly used by complex organisms to create macroscopic structures.(Peter and Davidson 2009) Very precise submicroscopic structures have been generated using deterministic DNA assembly in so-called DNA origami, but this is at the molecules' own size scale and is not scalable to cellular length scales (Rothemund 2006). Longer-range ordering has been accomplished with DNA assembled nanoparticle crystals, but the definition of the pattern is limited to repetitive patterns(Macfarlane et al. 2011).

Biological patterns are often an outgrowth of the behavior of reaction-diffusion networks, as first described by Alan Turing(Turing 1952). Mathematical models of reaction-diffusion networks have been shown to be capable of generating complex and beautiful patterns resembling everything from leopards' spots to variegated pigmentation in sea shells. That said, the first actual demonstration of a biological Turing mechanism occurred almost 40 years after the theoretical description, (Castets et al. 1990) illustrating how difficult these systems are to study, let alone engineer.

One of the aims of synthetic biology is to standardize the engineering of biology. Being able to rationally program spatial-temporal organization would be a great accomplishment, but requires the ability to algorithmically set down biological molecules and superstructures in specific times and places. While no scalable, programmatic pattern formation has yet been demonstrated, we now describe a practical approach that should allow for arbitrary pattern formation from bottom-up principles. Our approach appropriately rests on having programmable chemical reaction networks (CRNs) unfold in time and space.

While complex chemical reaction-diffusion systems (e.g., the well known B-Z reaction)(Vanag and Epstein 2001) are known, they are far from programmable. We will instead rely upon implementing CRNs with programmable DNA circuits(Yin et al. 2008; Phillips and Cardelli 2009). Soloveichik et al.(Soloveichik et al. 2010) have previously described a method by which CRNs can be implemented in DNA, and some of that system's predictions have been verified in vitro(Zhang and Winfree 2009). However, this work focused solely on the implementation of DNA CRNs in time, rather than in space. We now hope to design DNA CRNs that are inhomogeneous in space. We will initially focus on small, modular DNA reaction networks that can be treated as "primitives," meaning that the basic reaction can be duplicated, modified, and run in parallel. These primitives can then be the basis for the design of more complex CRNs in algorithmic pattern generators.

Results

Arbitrary reaction networks can be designed and implemented in DNA

In order to form predictable patterns, we require interacting reaction networks. DNA strand displacement reactions can be used to construct individual reactions with predictable kinetics.[11] In the strand displacement reaction, a single-stranded DNA molecule (ssDNA) binds to a hemi-duplex DNA molecule via specific Watson-Crick pairings (toehold). This toehold then initiates strand displacement to form a longer, more stable duplex (dsDNA), with concomitant release of a second single strand (**Figure 1A**). Reversible strand displacement reactions can be similarly designed. Because progression of the reaction is only favorable for complementary DNA strands, parallel reactions occurring concurrently in solution can be designed to be chemically

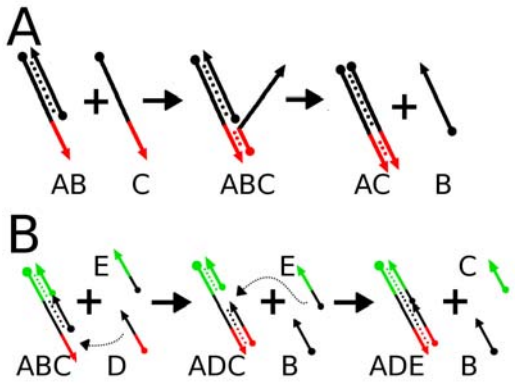

Figure 1: DNA-DNA reactions including. (A) shows a strand displacement and (B) strand displacement chain

"orthogonal," as eloquently described by Phillips and Cardelli(Phillips and Cardelli 2009).

These strand displacement reactions can be further coupled in arbitrary networks(Soloveichik et al. 2010). Most importantly for our purposes, the individual strand displacement reactions yield single-stranded products that are potential inputs for additional strand displacement reactions. Such coupled reactions can obviously be used to create CRNs. See **Figure 1B** for a schematic of this process. The great advantage of these DNA-based CRNs is that they are rationally programmable, unlike (for instance) a kinase which modifies a transcription factor via a relatively idiosyncratic rule-set in the context of a metabolic CRN. The modularity of the DNA components can be seen both in terms of the flexibility of sequence design, and in terms of the ready combination of components to create the network.

Simulation in MATLAB

The elements of the simulator are diffusion, reactivity with bimolecular kinetics, and a system for displaying the results. These elements can be implemented in either 1D or 2D. Diffusion and chemical kinetics are defined in terms of first order ODE, and can be solved using MATLAB's ODE45 solver or a simple Euler method solver.

Sequence-mediated diffusion control

In order to create patterns with a reaction-diffusion system, we must control both reactivity and diffusion. We can slow the diffusion of any given component of a CRN system by altering its sequence and affixing antisense oligonucleotides to a hydrogel (for example, by co-polymerizing antisense molecules terminating with an acrylic moiety, an acrydite). **Figure 2A** shows how the DNA may be anchored into the hydrogel superstructure. Depending on the design of a given DNA substrate, gate, or product, some single-stranded DNAs may have partial complementarity to the immobile antisense strand, and others less or no complementarity. This will lead to controllable, differential diffusion through the hydrogel. Diffusion parameters for a given DNA can be altered from fully diffusible to completely fixed depending on the number and strength of the base-pairs formed. This is not unlike chromatography where the equilibrium between bound analyte and unbound analyte determines the retention time.

To work towards the simulation of CRNs in a space where different molecular species will have differential diffusion, we first examine mobile DNA species A and B which are presumed to have equal diffusion coefficients ($D=4\times10^{-5}$ cm^2 sec^{-1}) in the gel. In the presence of an immobilized, complementary species *C we compare the predicted diffusion of A and B. We further assume that species A has no significant interactions with *C, but that B does. We can implement this latter reaction as a simple equilibration:

$$B + *C \leftrightarrow *BC$$

Both C and BC have zero diffusion (noted with asterisk, above) because C is covalently linked to the gel. This slows the effective diffusion of B relative to A (which does not form a complex with *C). Thus the relative diffusion rates of the species differ despite otherwise identical size. To illustrate this, we compare the case where the reaction above is performed under conditions where fast equilibration makes [B] = [*BC] such that species B spends half of its time in a non-diffusive complex. In the case for species A, $K_{eq}=0$ so that species A diffuses freely without interacting significantly with *C. **Figure 2C** shows the results of this simulation. A and B have very different concentration profiles at the same time point. A second modular design element is a short, linear "tail" on the end of other DNA components that partially hybridizes to a stationary, cross-linked molecule to change its diffusion.

Beyond this simple simulation, the diffusion of an

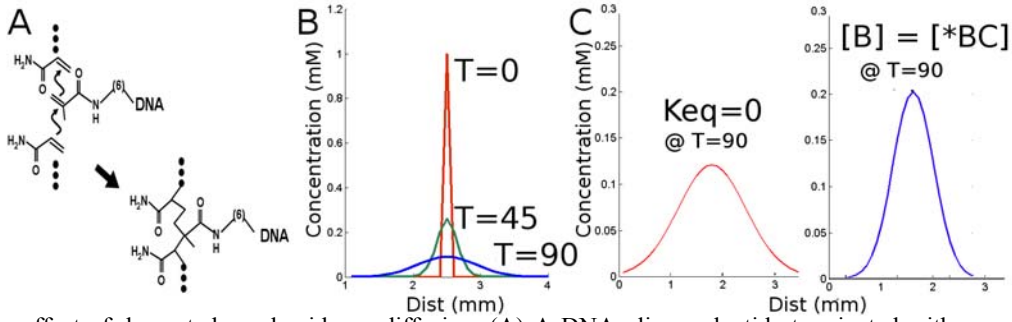

Figure 2: The effect of decorated acrylamide on diffusion. (A) A DNA oligonucleotide terminated with a acrydite moiety is incorporated into a growing acrylamide polymer. (B) Starting from a narrow distribution, DNA spreads through a gel by free diffusion. In (C) free diffusion (left, Keq=0) is compared to a species that interacts significantly with the immobile DNA (right, Keq=0.5), which diffuses more slowly.

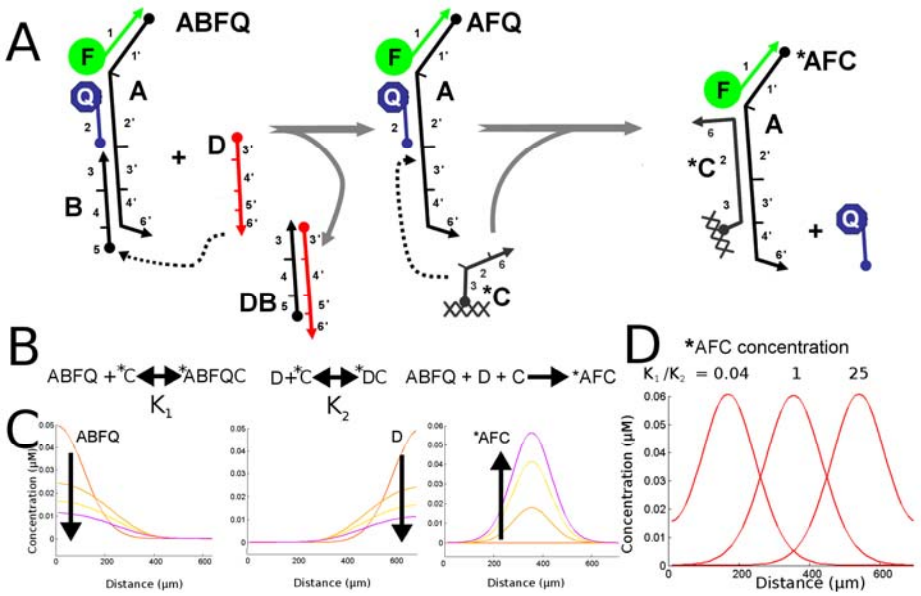

Figure 3: (A) shows the reaction governing the formation of the final, fluorescent product *AFC. DNA sub sequences are numbered with their complementary sub-sequences denoted with an apostrophe. (B) shows the three reactions occurring simultaneously; ABFQ and D are in fast equilibrium with *C to form immobile products (denoted by asterisks) with equilibrium constants K_1 and K_2. ABFQ and D form an immobile product *AFC when they are co-localized. (C) shows the concentrations of reactants ABF, Q, and D through time as well as the evolution of *AFC. (D) shows the final concentrations of product *AFC as a function of the ratio of the two equilibrium constants K_1 and K_2.

oligonucleotide is known to be influenced by its size and conformation. This provides opportunities to engineer a given DNA strand's diffusion. We will show that we can implement a spatially controlled reaction by controlling the reactivity and the diffusion of a set of properly designed DNA constructs.

Dynamic modification of diffusion and fluorescence of CRN components

The state of a DNA molecule (i.e. conformation, hybridization) can be transduced to optical information by strategic placement of fluorophores and quenchers, in a manner analogous to a molecular beacon. By having the fluorescence of DNA substrates change as they diffuse, react, and are immobilized we can potentially create dynamic, observable patterns. We will initially illustrate this by having two rapidly diffusing molecules react to form a local, immobilized fluorescent product.

From a historical perspective, this is similar to Ouchterlony double-diffusion experiments(Ouchterlony 1958). In these experiments an antigen and a mixture of antibodies are allowed to diffuse toward each other through a gel matrix. Depending on the diffusion constants of the antigen and antibody, a region of visible immuno-precipitation will occur at some location between the starting locations. Thus, Ouchterlony experiments could be used to infer intermolecular interactions by observing the location of a reaction product.

A similar approach in DNA can be engineered to produce a detectable product at a specific location. The strategy for implementing this is shown in **Figure 3A**. We set up a simulation modeling two diffusing DNA reagents, ABFQ and D. These reagents interact transiently with an immobilized DNA strand, *C (immobile species are denoted with an asterisk). This slows their diffusion by a predictable degree as shown above. When they meet at a location between their starting regions, they react and develop a fluorescent product. Fluorogenesis is accomplished by releasing a fluorescent product from its proximity to a quencher. Because the fluorescent product is also complementary to the cross-linked DNA, it is locked in place as it is generated.

Specifying a feature's location by modulating interactions with a DNA-gel

Adjusting interactions with a gel, as we have seen, can change the effective diffusion of a mobile DNA. By tuning the interaction strength, diffusion rates can be specified. These interactions are shown in **Figure 3B** with their equilibrium constants. By adjusting these equilibrium constants, we can control the location where the fluorescent product is produced. **Figure 3C** shows the cross section of the fluorescence pattern that would be generated in a gel when both species diffuse at equal rates; product evolution occurs in the center. **Figure 3D** shows three cases of that result from different ratios of the equilibrium constants, K_1 to K_2. It should be noted that although the position of product, *AFC, is only affected by the ratio of K_1 to K_2 rather than the absolute values of K_1 and K_2, these absolute values affect the time required for the pattern to develop.

This clearly shows that the location of the reaction can be varied by changing the relative equilibrium constants which are determined by the degree of complementarity to the immobilized oligonucleotide.

Using complementarity to adjust diffusion

It is computationally expensive to model an equilibrium between fixed and mobile states for each species in the CRN. To simplify and speed up the simulation, we implement an effective diffusion coefficient D_{eff} for each species depending on its assumed complementarity to the fixed DNA. Fast equilibrium with a fixed, immobile state constitutes a time-average of the diffusion at normal rate (corresponding to diffusion coefficient D) and zero. From the concentrations and standard free energy of the DNA-DNA reaction, (Nakano et al. 1999)we can calculate K_{eq}, the standard equilibrium constant for the reaction. And from that, we can derive the fraction of time spent in the fixed state, *BC.

First, we calculate the dissociation constant and the concentration of the reactants from their initial concentrations, B_0 and C_0.

$$K_d = 1/K_{eq} \quad (1)$$
$$[B] = B_0 - [BC] \quad (2)$$
$$[C] = C_0 - [BC] \quad (3)$$

Taking the definition of the equilibrium constant:

$$K_{eq} = [BC] / [B][C] \quad (4)$$

And defining the fraction of B bound at any given time as follows:

$$F_{bound} = [BC] / B_0 \quad (5)$$

We can substitute and simplify using the quadratic equation to express the Fbound of BC in terms of B_0, C_0 and K_d:

$$F_{bound} = \frac{(B_0 + C_0 + K_d) - \sqrt{(B_0 + C_0 + K_d)^2 - 4B_0C_0}}{2B_0} \quad (6)$$

We can therefore express the effective diffusion coefficient with the following relationship:

$$D_{eff} = (1-F_{bound}) \times D \quad (7)$$

For complex simulations, we will use D_{eff} in lieu of modeling an equilibrium between a mobile and immobile state. To predict the diffusion from sequence, we use an estimated Keq of B + *C ↔ *BC based on calculated ΔG from widely used base-pair stacking energies(Breslauer et al. 1986).

Specifying location in 2-dimensions using coupled reactions

The ability to specify the location of a reaction product is also expandable into multiple dimensions. The products of two separate, non interaction reactions then proceed to create a third product. In other words, two Ouchterlony line generators can be designed and aligned such that only at the intersection will a final product be evolved. This takes the form of the reactions shown in **Figure 4A-C**. The system shown in **Figure 4A** allows species AC and B to diffuse horizontally and where they meet they produce species C in a vertical line. Likewise **Figure 4B** shows a system that produces a horizontal line of product G. At the intersection of these two lines, a products C and G react sequentially with the immobilized fluorogenic construct *FPQ to form a central

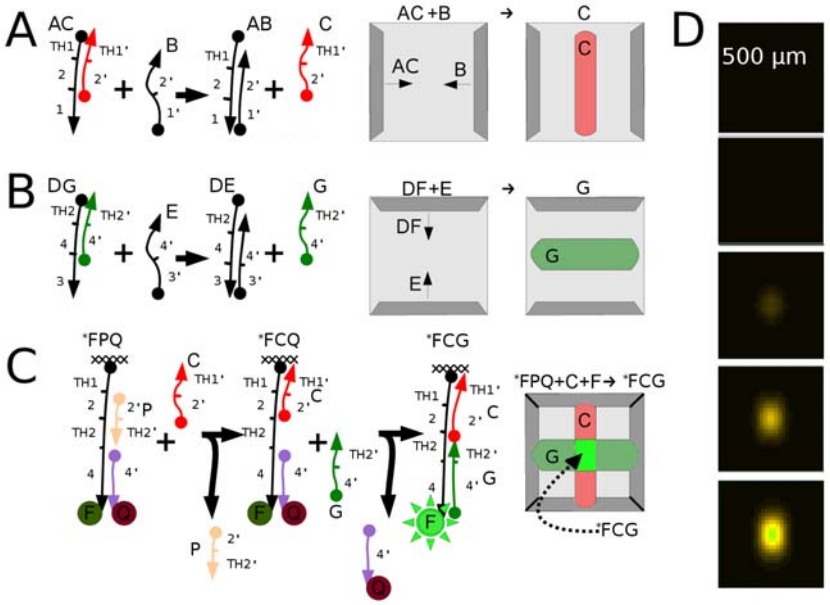

Figure 4: (A) shows a strand displacement that results in product C located in the vertical line. (B) shows a second strand displacement with a different toe-hold that results in product G located in a horizontal line. (C) shows the fluorogenic strand displacement in which immobilized FPQ becomes immobilized fluorescent product FCG only after reacting with *both* C and G. This produces a single fluorescent region located in the center of the gel. (D) shows images from our simulator showing the evolution of product FCG over time; time points are evenly spaced from 10 to 30 hours.

fluorescent spot of the final product, *FCG. In essence, we treated the line generator defined above as a module and applied it again in a second dimension to create a point generator.

Figure 4D shows the results of the simulation of this system. We made the simplifying assumption of using the effective diffusion according to equation 7, alleviating the need for additional terms to account for transiently bound species. The slight asymmetry in the final frame of **Figure 4D** is due to the sequential reaction of X with C then F; the delay introduced by requiring the first reaction to be complete allows for more progress in the reaction originating from the top and bottom reaction and, thus, a taller spot of final product, *FCG.

Reactions can be made orthogonal and used to generate composite patterns

Because DNA-DNA interactions can be designed to be very specific, it is possible to build reaction networks where each reaction is chemically orthogonal; the reactions will not aberrantly interact. Parallel expansion of the 'point generation' program described above allows for the generation of pre-specified, arbitrary, complex patterns via designing the sequence of the interacting DNA molecules. In other words, multiple instances of the type of addressable point generator described above can run within the same gel at the same time and thus form more complicated patterns.

We present an example in which we selectively de-quench immobilized fluorophore in multiple regions where separate, lateral and vertical reaction-diffusion system overlap. Each system, lateral (row) and vertical (column), has a pair of reagents with defined characteristic that determine the final position of the developed feature. A feature therefore can be developed as a “pixel” at an arbitrary position. We call the seven instances of the line generation module A through G. The gel homogeneously includes two immobilized fluorophore-quencher pair species, Xa and Xb. These each require two separate strand displacement reactions in order to become fluorescent. These two toehold regions for the two displacement reactions are shown as TH1 and TH2 in **Figure 4C**.

The process works as follows for reactions systems A, D and G: immobilized Xa has a version of TH1 that responds only to the product of reaction system A. Thus a “primed” column is generated in which TH2 is open only where A reacts. See left column in **Figure 5A**. Products D and G react specifically with TH2 on product Xa (products E, F and H do not). Thus two specific regions of the primed Xa column are fully de-quenched (and turn green). A system of eight such strand displacement reactions per the prototype shown in **Figure 4** can be designed to construct a five point design in 2-dimensions.

With the appropriate diffusion coefficients and reactivities, five regions are induced to fluoresce. The intended regions are shown schematically in **Figure 5A** and the results of the simulation with these parameters is shown in **Figure 5B.**

An interesting property of this system is its scalability. The topology of the pattern will be generated without regard to the dimensions of the gel slab (although the time and material necessary to achieve the result will increase with the size of the gel). This is shown in **Figure 5C** where a simulation was

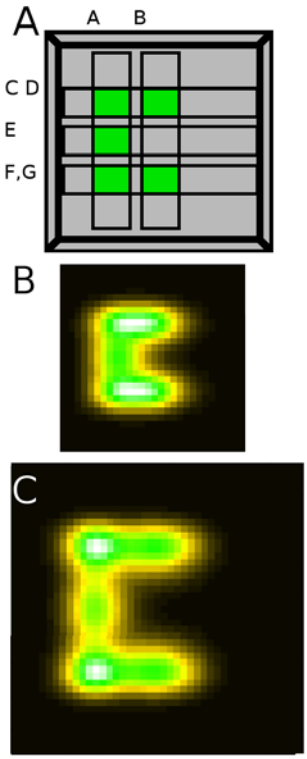

Figure 5: (A) shows the regions activated by the reaction systems described in Table 1. The labeled regions indicate intended location the products of reaction systems A-H should appear. (B) shows a simulation of the same system showing an X pattern. (C) Shows the same simulation with a larger space, all else remaining constant. The pattern scales with the space

run with all parameters consistent to that in **Figure 5B** except that the size of the simulated region was enlarged.

Scalability and resolution limits

The minimal size of the features generated by this system scales with the overall size of the gel in which the reactions are occurring. Minimal features are generated when reactants diffuse only a short distance into each others’ territory before reacting. In other words, for sharp features the effective diffusion rate must not greatly exceed the effective reaction rate. Whether a given reaction is diffusion- or reaction-limited can be characterized in terms of the Thiele modulus. Reeves et al. (Reeves et al. 2006)conclude that for effective patterning using diffusing signal molecules, the Thiele modulus must be approximately 1. At this value the influences of reaction and diffusion are balanced.

This can be best illustrated with a thought experiment. We can take a gel of width 600 μm and embed a reaction where two DNA molecules are diffusing towards one another (as shown above in **Figure 3**). They will react with a rate coefficient of 10^6 mol^{-1} sec^{-1}(Soloveichik et al. 2010). If we take the diffusion rate to be extremely slow, the advance edges of the DNA samples will yield a low, broad concentration profile and a correspondingly broad feature (see **Figure 6A**). In the opposite extreme, if we consider a diffusion rate that is very fast, such that the molecules diffuse

across the entire gel before they find a partner and react, then this clearly produces a broad feature (see **Figure 6B**).

In order to find an optimal diffusion constant that will produce a narrow line we can use a numerical simulation. A typical DNA substrate has a native diffusion rate of 4.6×10^{-8} cm^2 sec^{-1} in a 5% acrylamide gel (experimental estimate, data not shown), a value that can be further modified by interactions with oligonucleotides immobilized in the gel. Using the reasonable estimates of the values of diffusion and reaction coefficients above, we estimate that the smallest full-width, half-max (FWHM) feature that can be generated in a 600 µm gel is approximately 63 µm or ~11% of the width of the gel (**Figure 6C**). This corresponds to a hybridization length of 10 residues between substrate and gel, which reduces diffusion by 58% to ca. 2×10^{-8} cm^2 sec^{-1}. The topology (and hence the resolution) of the feature size should scale with the outer dimensions of the gel. If we decrease the width of the gel to 150 µm, (see **Figure 6D**) the optimized diffusion coefficient produces a feature of width ~13 µm (again ~9% of the width), a modification of the diffusion rate that corresponds to ~11 bases of hybridization.

In principle, so long as diffusion can be limited to match the overall size of the gel, there is no limit to the smallest feature size that can be generated except that it will be minimally about 10% of the width of the gel itself. It may be that biological systems that utilize reaction-diffusion systems for spatial organization may be limited in their precision by this same minimal relative feature width. From a practical, experimental standpoint, this implies that to be able to make very small features, one must be able to manipulate increasingly smaller samples. In addition, there will be a

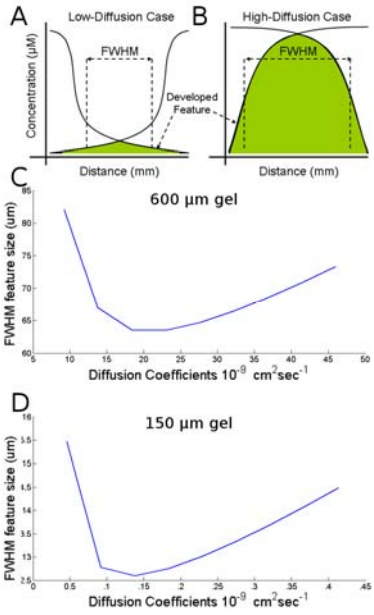

Figure 6: Thought experiment illustration showing that in both the high-diffusion extreme (A) and the low-diffusion extreme (B), the resulting feature is broad. (C) Shows the relationship between the diffusion coefficients of reagents and the final full-width half-max (FWHM) size of the generated feature in the case with a gel of 600 µm in width and (D) 150 µm in width.

breakdown of the relationship between increased hybridization and lower D_{eff} as a given interaction becomes strong enough to affix DNA strands semi-permanently, so that lateral motion cannot be modeled by simple diffusion. This breakpoint occurs at a k_{off} of ca. 10^{-2} sec^{-1} at room temperature, or approximately 15 base-pairs of interaction between substrate and gel (Robelek et al. 2006). This practical limitation sets the minimal, controllable feature size at about 10 µm for the types of reagents and timescales described here.

From a theoretical standpoint, this work shows that a chemical system can develop an arbitrary feature in space using only chemically defined parameters. There is a resolution limit to systems functioning by passive diffusion. This limit should apply to any CRN that develops spatial order from a homogenous system by such reaction-diffusion mechanisms. Thus, this limitation may be relevant to natural as well as synthetic systems.

Spatial segregation to control chemical reactions

Spatial organization can potentially be used to control chemical reactivity. For example, oligonucleotides linked to small, reactive organic compounds can be organized by templating, which will in turn help to control the order and regiospecificity of the reactions that the small molecules undergo(Li and Liu 2004; Kleiner et al. 2010). By implementing such templated reactions in the context of CRNs it should be possible to add new levels of spatial and temporal control to such assemblies.

We first suggest that control over diffusion can direct the creation of a specific product. Two oligonucleotides carrying reactive chemical species (A and B) can be formed into lines in a gel using procedures those described above (see **Figure 7A**). In this example, the diffusion of precursors AA_0 and BB_0 (see **Figure 7B**) are controlled by partial hybridization of an immobilized oligonucleotide to domains 5 and 6 on the chimeric oligonucleotide precursors. Upon immobilization, strand displacement reactions (e.g. $AA_0+A_1 \rightarrow A+waste$) 'activate' A and B to become substrates for additional hybridization and reaction. The diffusion of a third reactant, DNA species D, is similarly adjusted by complementarity to domain 0. Species D diffuses slowly so that it can react with already immobilized, activated A and B, forming either DA or DB. However, since reactive species D diffuses through the activated lines of A and B sequentially this mediates the order of reaction among the small molecule cargoes. Only one of the two possible products shown in **Figure 7C**, DAB, should be generated. We simulate the relative production of DAB relative to DBA in **Figure 7D.**

Similar ordered reactions have been performed by programmed DNA nanorobots(Yurke 2007). However, as with many aspects of the amorphous computations described herein, the scalability of 'classic' DNA nanotechnology is doubtful, especially for the production of chemicals in bulk. Gel-based separations are already common, and thus the concept of controlled, gel-based reactions is more amenable to scaling. Moreover, the process of chemical assembly could occur continuously in the gel, with new reactants constantly diffusing, being activated, and assembling in an ordered fashion. The system is eminently programmable, and changing only the immobilized DNA sequence should change

the order and kinetics of compound activation, which would in turn alter both the nature and efficiency of production of the final chemical product. Simplistically, depending on the sequences of the chimeras, DBA rather than DAB could be produced with high specificity. More importantly, though, should a given reaction prove inefficient, the width of the self-assembled DNA bands could be increased by simply altering hybridization and diffusivity, allowing more time for the reaction to occur. It is reasonable to think that this system could be combined with programmable nucleic acid reactions[10] to realize a fully programmable, algorithmic system for chemical construction.

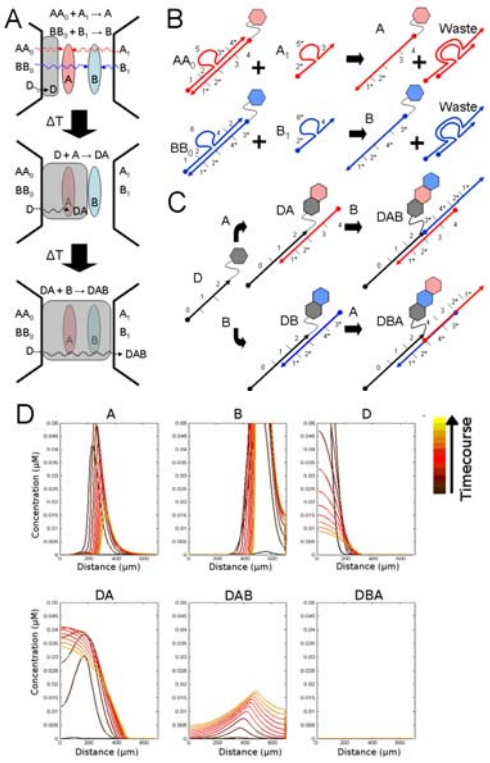

Figure 7: (A) Species A and B develop as lines in a gel through which species D diffuses. (B) Schematic for reactions developing A and B. AA_0 meets A_1 and develops A (and likewise for BB_0 and B). (C) Diffusing DNA species D carries a reactive small molecule (hexagram) slowly through lines of A and B. D and A or B react to form a different product depending on relative locations. (D) Numerical CRN simulation results for all relevant species shows DAB production is greater than DBA.

Discussion

The diverse forms of self-organization in living systems develop from ostensibly simple homogeneity. This has fascinated humans since antiquity(Aristotle 2004). We have suggested a engineerable system that can create spatial patterns from chemical information. Biology excels at this feat but the methods by which it is accomplished are idiosyncratic and not as amenable to engineering as the methods presented here.

Our system is at some level inspired by Alan Turning's seminal paper formulating a set of conditions for pattern formation including plausible kinetic equations with symmetry-breaking properties. Turing speculated that such reaction-diffusion systems could be the basis of embryonic morphogensis(Turing 1952). His work made clear that specific properties of reactivity and diffusivity are necessary conditions for generating self-organized patterns.

By developing concepts for programming both diffusion and reactivity using nucleic acid sequence information, we provide a path forward for better understanding, mimicking, and ultimately exploiting the CRNs that elude chemists and underlie biology. Biological reaction-diffusion systems are hypothesized to regulate key biological pathways as a general model for the formation of complex patterns(Maini and Othmer 2001). However, despite the long history of research into biological reaction-diffusion systems, most studies focus on either the understanding of natural pattern-formation systems or theoretical possibilities to generate stochastic patterns. Many biological phenomena might be re-imagined in the context of our designable, modular, reaction-diffusion system. Examples might include recapitulating the mechanisms of *Drosophila* development wherein diffusible signals and feedback pathways generate the initial polarization of the embryo. We note that our Turing-inspired simulations predict a resolution limit of 10% of the width of the system. This is indeed the value observed in *Drosophila* development(Gregor et al. 2007). Visibly patterned phenomena such as skin pigmentation may be developed from a reaction-diffusion type Turing mechanism (Nakamasu et al. 2009) and might also be demonstrated with our system. Nucleic acid pattern generators are not yet applicable *in vivo*. Organisms are chemical reaction networks capable of self-replication given appropriate substrates and inputs. Given that there is no adequate definition of life, much less artificial life, our attempts to generate programmable chemical reaction networks can be seen as a first step towards creating synthetic organisms.

Beyond fomenting better understanding of biology, these CRNs should allow entirely new applications in chemistry and materials science. Self-organizing chemistry has previously been experimentally demonstrated in what is now known as the Belousov-Zhabotinsky reaction. This reaction, like Turing's hypothetical reactions, has specific diffusion rates that affect the appearance of patterns(Field and Noyes 1974). However, as was the case with biological development, such reactions cannot be readily elaborated or engineered. New deterministic and algorithmic patterns can potentially lead to the generation of "smart" materials whose bulk architectures are structured down to the nanoscale. For example, Janus particles, whose surfaces are two differentially patterned hemispheres, can be used to generate complex topologies(Chen et al. 2011). It stands to reason that particles with more complex surfaces generated by internal reaction-diffusion systems could generate more complex, patterned associations. Additionally, a reaction-diffusion system might allow for a macroscopic positioning of other DNA structures such as DNA origami(Rothemund 2006). A meso-scale pattern might be etched into a medium by selectively melting a polymer gel cross-linked by self-assembled DNA helices(Zhu et al. 2010).

In order to rationally design even more complex, algorithmic, developmental programs for later applications, we need to now develop the equivalent of a chemical "compiler" and a test bed for its programs. Complex reaction networks should be specified at a modular level and then rationally rendered into constituent chemicals capable of running the specified reactions. This is only realistic if the design can be made rational and generalized. There must be explicit, computable relationships between sequence and inter-reactivity and between sequence and diffusivity. The thermodynamic properties of nucleic acid hybridization are well known(Nakano et al. 1999). Linear strands with specific energies of hybridization to a immobile strand can thus be computationally designed to specify diffusion. This can be combined with the powerful set of DNA modules that have been shown to be modular has been demonstrably "compiled" into large circuits useful to computation(Qian and Winfree 2011). We envision that the species in such circuits (including amplifiers, thresholds and logic gates) could will dynamically modulate diffusivity by alternately exposing or hiding diffusion-modifying sequences to the fixed medium.

Ultimately, modularity should prove very important in developing such self-organizing systems, as will abstraction and encoding. There is evidence that modularity has emerged from natural evolution as well(Ravasz et al. 2002). By analogy to computer science, implementing a system of modules as an 'operating system' for CRNs should be like a high-level computer language. A computer programmer need not know the deepest workings of the hardware (e.g., machine code, register shifts, memory addresses, etc.) in order to write useful software. The work presented herein is a step toward such a CRN language and compiler.

Acknowledgments

This work was funded by the National Institute of Health (1 R01 GM094933 and 1 F32 GM095280), and the National Security Science and Engineering Faculty Fellowship (FA9550-10-1-0169). The content are solely the responsibility of the authors and do not necessarily represent the official views of the sponsors.

References

Aristotle (2004). On the Generation of Animals, Kessinger Publishing.

Breslauer, K. J., R. Frank, H. Blacker and L. A. Marky (1986). *Proc. Nat. Acad. Sci.* **83**(11): 3746-3750.

Castets, V., E. Dulos, J. Boissonade and P. De Kepper (1990). *Phys. Rev. Let.* **64**(24): 2953-2953.

Chen, Q., J. K. Whitmer, S. Jiang, S. C. Bae, E. Luijten and S. Granick (2011). *Science* **331**(6014): 199-202.

Field, R. J. and R. M. Noyes (1974). *J. Am. Chem. Soc.* **96**(7): 2001-2006.

Gregor, T., D. W. Tank, E. F. Wieschaus and W. Bialek (2007). *Cell* **130**(1): 153-164.

Kleiner, R. E., C. E. Dumelin, G. C. Tiu, K. Sakurai and D. R. Liu (2010). *Journal of the American Chemical Society* **132**(33): 11779-11791.

Li, X. and D. R. Liu (2004). *Angew Chem Int Ed Engl* **43**(37): 4848-70.

Macfarlane, R. J., B. Lee, M. R. Jones, N. Harris, G. C. Schatz and C. A. Mirkin (2011). *Science* **334**(6053): 204-208.

Maini, P. K. and H. G. Othmer (2001). Mathematical models for biological pattern formation, Springer.

Murnen, H. K., A. M. Rosales, J. N. Jaworski, R. A. Segalman and R. N. Zuckermann (2010). *J. Am. Chem. Soc.* **132**(45): 16112-16119.

Nakamasu, A., G. Takahashi, A. Kanbe and S. Kondo (2009). *Proc. Natl. Acad. Sci. U. S. A.* **106**(21): 8429-34.

Nakano, S., M. Fujimoto, H. Hara and N. Sugimoto (1999). *Nuc. Acids Res.* **27**(14): 2957-65.

Ouchterlony, O. (1958). *Prog. Allergy* **5**: 1-78.

Peter, I. S. and E. H. Davidson (2009). *FEBS Letters* **583**(24): 3948-3958.

Phillips, A. and L. Cardelli (2009). *J. Roy. Chem. Soc.* **6**(Suppl 4): S419-S436.

Qian, L. and E. Winfree (2011). *Science* **332**(6034): 1196-1201.

Ravasz, E., A. L. Somera, D. A. Mongru, Z. N. Oltvai and A. L. Barabasi (2002). *Science* **297**(5586): 1551-5.

Reeves, G. T., C. B. Muratov, T. Schupbach and S. Y. Shvartsman (2006). *Dev. Cell.* **11**(3): 289-300.

Robelek, R., F. D. Stefani and W. Knoll (2006). *Phys. Status Solidi (A)* **203**(14): 3468-3475.

Rothemund, P. W. (2006). *Nature* **440**(7082): 297-302.

Soloveichik, D., G. Seelig and E. Winfree (2010). *Proc. Nat. Acad. Sci. U.S.A.* **107**(12): 5393-5398.

Turing, A. M. (1952). *Philos. T. Roy. Soc. B* **237**(641): 37-72.

Vanag, V. K. and I. R. Epstein (2001). *Phys. Rev.* **87**(22): 228301.

Yin, P., H. M. T. Choi, C. R. Calvert and N. A. Pierce (2008). *Nature* **451**(7176): 318-322.

Yurke, B. (2007). *Controlled Nanoscale Motion* **711**: 331-347.

Zhang, D. Y. and E. Winfree (2009). *J. Am. Chem. Soc.* **131**: 17303-17314.

Zhu, Z., C. Wu, H. Liu, Y. Zou, X. Zhang, H. Kang, C. J. Yang and W. Tan (2010). *Angew. Chem. Int. Ed.* **49**(6): 1052-1056.