Examining tissue differentiation stability through large scale, multi-cellular pathway modeling

Richard L. Schiek and Elebeoba E. May***
Computational Sciences Department, rlschie@sandia.gov
**Computational Biology Department,* eemay@sandia.gov
Sandia National Laboratories, Albuquerque, NM, USA.

Abstract

Genetic expression and control pathways can be successfully modeled as electrical circuits. To tackle large multicellular and genome scale simulations, the massively-parallel, electronic circuit simulator, Xyce^{TM} [11], was adapted to address biological problems. Unique to this bio-circuit simulator is the ability to simulate not just one or a set of genetic circuits in a cell, but many cells and their internal circuits interacting through a common environment. Additionally, the circuit simulator Xyce can couple to the optimization and uncertainty analysis framework Dakota [2] allowing one to find viable parameter spaces for normal cell functionality and required parameter ranges for unknown or difficult to measure biological constants. Using such tools, we investigate the *Drosophila sp.* segmental differentiation network's stability as a function of initial conditions.

Introduction

Expression of a genetic code defines characteristics of a given organism. As an organism grows and adapts to its local environment specific elements of its genetic code are expressed while other elements are suppressed. Complex control mechanisms exist to regulate the expression of genes during the life of a cell. [6]–[8]

To fully appreciate how a genetic repository or genome translates into a functioning cell, one must understand the control mechanisms of genetic expression. Genetic products of a given gene can promote or suppress the further production of that gene creating a simple feedback loop. [6]–[8] Similarly, genetic products from other genes can regulate the production of a given gene creating complex feedback loops or expression cascades. Feedback loops and cascades are not limited to a single cell, but can span an entire cell culture or cellular generation influencing differentiation and development. [7]

As an abstraction to better understand genetic expression and control, genetic material and its associated control mechanisms can be viewed as a genetic switch. [1], [6]–[9] Such a switch can be modeled as an electrical circuit where a signal (the transcript from a section of DNA) is generated and altered as it interacts with other components during its propagation. This analogy is far from perfect as there are significant differences in the switching speed and signal to noise ratio of a genetic circuit versus an electric circuit. However, this analogy allows one to consider very complicated, dynamic control circuits while investigating expression stability and population dynamics. [7]

To understand and model cellular differentiation, we have simulated the *Drosophila sp.* segment polarity gene network for a 2D array of cells connected through a common diffusion limited environment. In such an environment, cells experience local concentrations of differentiation stimuli determined by neighboring cells production and consumption rates. These local stimuli effect the genetic and metabolic regulatory networks within the cell directing eventual development. Specifically, the initial conditions and history of a given cellular environment strongly influence a cell's future development. For this model problem, we have examined functionality and the systems sensitivity to initial noise by using Dakota [2]

Chemical, Biological Domain	Electrical Domain
mass	charge
mass flux	current
concentration*	voltage
stoichiometric conservation	Kirchhoff's voltage law
mass conservation	Kirchhoff's current law

Table 1: Equivalents between the chemical, biological domain and the electrical circuit modeling domain. *Note, that there is no equivalent to volume in the electrical domain. Thus for one to strictly employ concentration one must first define a consistent system volume such as a biological cell or culture plate.

to explore the systems parameter space and response functions. We are able to demonstrate that the *Drosophila sp.* differentiation network is very sensitive to noise in its initial conditions.

Framework

A biological or chemical simulation working within an electrical circuit context requires a translation framework to convert from the former domain to the latter. In electronics, a fundamental quantity is charge while in the biological domain, one is often concerned with concentration of a given chemical compound. Given a control volume, such as the volume of a cell, concentrations can be converted to mass. As a basis for a biological to electrical problem conversion, this work will equate mass of a given chemical species with charge. Each pathway or *wire* in a circuit will carry a different chemical species and the charge on that wire will denote the mass of that chemical species present in the simulation.

Continuing this analogy, electrical current which is the timed rate of change of charge is equivalent to the rate of mass change, *i.e.* how quickly a compound is used or created by the system or mass flux through the system. Voltage is a relative measure of electrical potential. The chemical or biological analog to voltage here is chemical concentration provided one is measuring voltage relative to a neutral ground. Kirchhoff's Voltage Law, KVL, which requires the voltage drop around any closed circuit to be zero [5] enforces a stoichiometric balance on chemical reactions. Kirchhoff's Current Law, KCL, which requires the current flow into a circuit node balance current flow out of that node enforces conservation of mass within the system. For reference, table 1 summarizes the analogous terms joining chemical, biological problems and electronic circuit problems. It is important to note that with such a framework in place, we are not ignoring any important kinetic or stoichiometric aspects of the biological problem just to work within an electrical circuit domain. Rather, we are using this framework to take advantage of existing simulation capabilities developed for electrical modeling problems.

Genetic Switch Application

As an example of a biologically inspired circuit, figure 1 demonstrates a genetic switch in an *E. coli* tryptophan regulation circuit. Focusing primarily on the switching, this problem is posed as a digital circuit. Thus, specific reaction rates and binding constants are ignored so that only tryptophan induced regulation is seen. Casting the problem this way allows one to study cases where little or no kinetic, stoichiometric or diffusion data are available and one wishes to test hypothetical control systems.

In figure 1, a repressor, `apoRep`, and `mRNA` are preset at constant levels and depicted as constant voltage sources. Tryptophan, `trp`, concentration oscillates and it is implemented as a time dependent

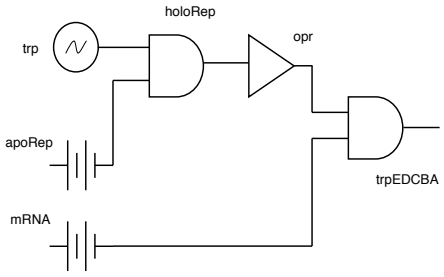

Figure 1: A tryptophan regulated switch. An oscillating input level of tryptophan, `trp`, controls the production of the gene products `trpEDCBA`.

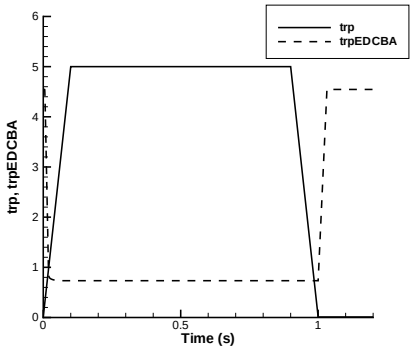

Figure 2: The tryptophan switch in action. Only when tryptophan, `trp`, levels drop are the gene products, `trpEDCBA`, available.

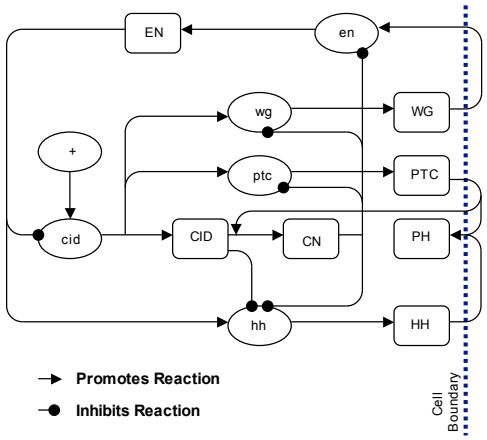

Figure 3: Developmental control circuit derived by Dassow [3] Lower case letters are mRNA (gene products) while upper case letters denote proteins. Arrows indicate places where a compound promotes the production of another compound while filled circles denote places where a compound represses the production of another species.

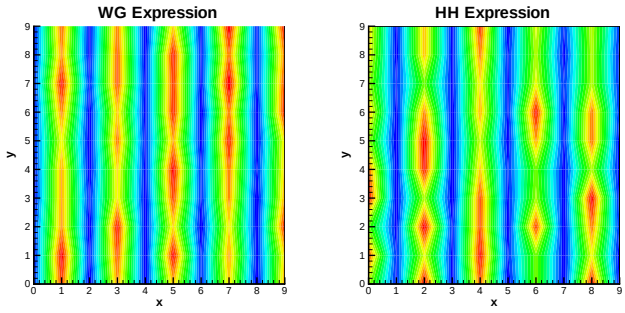

Figure 4: A 10 by 10 grid of cells starting with an initial noisy, oscillatory level of `WG` differentiates into `WG` producing and `HH` producing populations. The plot on the left depicts 5 layers of `WG` producing cells while the right contour plot depicts 5 layers of `HH` producing cells at the same time point. Initially the system was started with 10% rms. random noise in `WG` superimposed over the initial conditions.

source. If `trp` and `apoRep` are present then an activated repressor `holoRep` is formed as indicated by the AND gate. The presence of the activated repressor deactivates the operator, `opr` via a NOT gate. Without `opr`, production of the tryptophan controlled gene products, `trpEDCBA`, is shut down. Figure 2 graphically shows the genetic switch reacting to changes in the applied tryptophan level. While this example is very simple, it demonstrates that biological control concepts can be mapped into a circuit model. More importantly, this simple switch is easily embedded in a larger circuit allowing one to model complex systems from simple constituents.

Cellular Development

Development of a system from one state to another in a controlled manor usually involves feedback to assert such control. Purely chemical systems such as the Belousov-Zhabotinskii (BZ) reaction network [10] and multi-cellular networks such as *Drosophila sp.* differentiation [3] use feedback to control the system's development.

During development to of the fruit fly embryo, *Drosophila sp.* , one can image a series of bands developing along the major axis of the growing larva–a graphical indicator of the underlying cellular differentiation in progress. Figure 3 presents the control network responsible for cellular differentiation in *Drosophila sp.* [3]. Though complex, this network typically bifurcates into one of two states. If a cell is producing the gene product wg then the protein WG will likely be produced as well. The WG protein is exported into the cellular environment and picked up by neighboring cells where it can promote the expression of the gene product en. The en gene product represses the production of wg and puts the cell into a different state from a cell producing WG, specifically into a state where it is producing and expressing HH. Thus, cells will typically be producing either WG or HH with a small percentage of cells producing both of these proteins as the switch from expressing one gene to another.

Actual simulations of the *Drosophila sp.* network were carried out as follows. The network was converted to an electrical circuit using analogs for chemical reactions, material storage, promotion, repression, degradation and diffusion. Both simple kinetic and Michaelis-Menton kinetics were used in modeling the interactions within a cell. Once the circuit was created, a 10 by 10 grid of cells embedded within a diffusion limited environment was created, again as a circuit. The diffusion network in which the cells are embedded can be represented to any numerical order by a resistor-capacitor network where resistance parameters are directly proportional to chemical diffusion rates. For this work, second order accuracy was used in the spatial resolution of the diffusion network. Fundamental constants like reaction rates, enzymatic turnover rates and diffusion coefficients were parameterized within this circuit. This parameterization allows the optimization framework to alter parameters between simulation runs in order to explore the phase space for this system.

As an example of one realization in simulating the cellular development, Figure 4 depicts concentration contour plots of the species WG and HH. Initially, the system was started with zero concentration of the exported species, PH, PTC and HH and an oscillatory level of WG. This initial oscillatory state represents the initial bias that anterior-posterior, dorso-ventral patterning hierarchies initiate in the developing embryo. [4] Additionally, a 10% rms. random noise was added to the WG initial conditions to simulate disturbances of the system from an ideal starting state. Such noise was also parameterized in the circuit and varied to gauge system robustness. The simulation proceeded forward in time to model several hours of embryonic development. At the end of the simulation, the striation pattern was examined to determine if the embryo successfully developed and maintained its appropriate banded structure. Physically, the striations in concentration shown in figure 4 represent layers of cells becoming WG producing or HH producing over time an example of cellular differentiation. Incomplete or missing bands would represent a failure in differentiation.

While the full model system has 52 unknown parameters, this was reduced to a subset of 26 parameters using a model sensitivity analysis. This subset of parameters included: expression rates, enzymatic turn-over rates, reaction and diffusion rates and noise levels. A design of experiments approach was used to understand how this collection of state variables affects the resulting system. Using the Dakota optimization framework, the problem parameter space was explored using latin-hypercube sampling with over individual 50,000 simulations performed.

Figure 4 presents the probability of successfully differentiating given a specific level of initial random noise. On the leftmost side of this graph, at 0% noise, one finds the unperturbed system which successfully differentiates about 15% of the time. This 15% probability is relative to the search space used for unknown parameter estimation in this problem. As we used a large search space only a small fraction of that space proved viable for embryonic development. When noise was added to the system, the volume of viable parameter sets leading to successful embryonic development decreased quickly. Noise

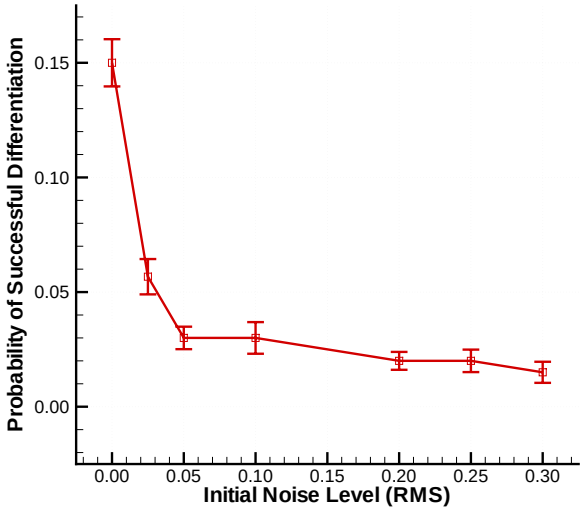

Figure 5: Noise in the initial WG concentration field significantly reduces the probability of successful differentiation. Error bars represent 99% confidence intervals around the mean probability of successful differentiation based on a binomial distribution (i.e. a pass, fail experiment).

levels of 0.05 and 0.10 are realistic concentration fluctuations on the time scale of differentiation; higher levels were explored to fully understand the trend.

Conclusion

Though still in development, this biological circuit simulator has the potential to handle large and complex problems. Depending on the type of data available, one can cast problems as digital or analog circuits and easily simulate many replica of a single circuit interacting with a collection of others. Through the coupling to an optimization framework, one can explore the dynamics of multiple cellular networks or of entire cell cultures elucidating governing parameters as well. Here, mulit-cellular coupling demonstrated that the *Drosophila sp.* differentiation network is very sensitive to initial noise.

References

[1] Arkin, A.P., "Synthetic cell biology," *Current Opinion in Biotechnology* 12, pp.638–644 (2001).

[2] Dakota home page: http://software.sandia.gov

[3] Dassow, George von, Meir, E., Munro, E. & Odell, G., "The segment polarity network is a robust development module," *Nature* 406, pp. 188–192 (2000).

[4] Ghabrial Amin, Lusching, S., Metzstein, M. & Krasnow, M., "Branching Morphogenesis of the *Drosophila* Tracheal System," *Annu. Rev. Cell Dev. Biol.* 19, pp. 623–47 (2003).

[5] Horowitz, Paul & Hill, Winfield, "The Art of Electronics," 2nd ed., Cambridge University Press, Cambridge (1989).

[6] McAdams, Harley H. & Shapiro, L., "Circuit Simulation of Genetic Networks," *Science* 269, pp. 650–656 (1995).

[7] McAdams, Harley H. & Arkin, A., "Simulation of Prokaryotic Genetic Circuits," *Annu. Rev. Biophys. Biomol. Struct.* 27, pp. 199–224 (1998).

[8] McAdams, Harley H. & Arkin A., "Gene regulation: Towards a circuit engineering discipline," *Current Biology* 10, R318–320 (2000).

[9] Palsson, Bernhard, "The Challenges of in silico biology," *Nature Biotechnology* 18, pp. 1147–1150 (2000).

[10] Tyson, J.J, "A Quantitative Account of Oscillations, Bistability and Traveling Waves in the Belousov-Zhabotinskii Reaction," in *Oscillations and Traveling Waves in Chemical Systems* editors R.J. Field and M. Burger, John Wiley & Sons (1985).

[11] Xyce home page: http://www.cs.sandia.gov/xyce

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.