UNIVERSITY OF CALIFORNIA

Santa Barbara

Stochasticity and Synchrony in the Mammalian Circadian Network

A Thesis submitted in partial satisfaction

of the requirements for the degree of

Master of Science

in

Chemical Engineering

by

John H. Abel

Committee in charge:

Professor Francis J. Doyle III, Co-Chair

Professor Linda R. Petzold, Co-Chair

Professor Michelle A. O'Malley

Professor Baron Peters

September 2015

The Thesis of John H. Abel is approved.

Professor Baron Peters

Professor Michelle A. O'Malley

Professor Linda R. Petzold, Committee Co-Chair

Professor Francis J. Doyle III, Committee Co-Chair

July 2015

Stochasticity and Synchrony

in the Mammalian Circadian Network

John H. Abel

Acknowledgement

Foremost I would like to thank Professors Frank Doyle and Linda Petzold for their overwhelming support for this research, and for allowing me great freedom in this project. I would also like to thank my committee, Professors Baron Peters and Michelle O'Malley for their thoughtful commentary and assistance in my work. The Doyle and Petzold lab groups have been instrumental in this work. I would like to especially thank Dr. Peter St. John for constant guidance, support, and this LaTeX thesis style; and also Dr. Brian Drawert, Ben Bales, Dr. Kirsten Meeker, and Kelsey Dean. I would finally like to thank our experimental collaborators, Prof. Erik Herzog and his laboratory group (Dr. Daniel Granados-Fuentes, Thomas Wang, Dr. Vania Carmona-Alcocer, and Anne Sun), and our theory collaborators (Prof. Stephanie Taylor, Prof. Joerg Stelling, and Lukas Widmer). It has been a privilege to work with them.

On a more personal note, the UC Santa Barbara Department of Chemical Engineering has been a fantastic place to conduct research, and it is with heartache that I will be continuing graduate studies at Harvard. Specifically thank you to the Las Vegas crew and fantasy football league, and anyone who shared a conversation about research on the bluffs. Above all the rest, I would like to thank Faith Blake for 1. teaching me all the genetics I never learned, and 2. agreeing to move to Cambridge, MA. The adventure has only begun.

I would certainly be remiss if I did not also thank my funding: this work was supported in part by the National Institutes of Health under grant 1R01GM096873-01, and the Institute for Collaborative Biotechnologies under grant W911NF-09-0001 from the U.S. Army Research Office. The content of the information does not necessarily reflect the position of the policy of the Government, and no official endorsement should be inferred. I would also like to thank the UC Santa Barbara Mellichamp Cluster for my fellowship, which has allowed me great freedom in travel and research, and without which I would not have been exposed to so many new ideas and discoveries.

Abstract

Stochasticity and Synchrony in the Mammalian Circadian Network

John H. Abel

A vast majority of life on Earth exists in an environment where resource availability and environmental conditions are temporally periodic. There is therefore an evolutionary advantage for organisms to partition behavior into certain times of day. Circadian rhythms, endogenous near-24 hour oscillations in gene expression, perform this task. These rhythms exert control over a large fraction of biological processes, and as such are implicated in a wide range of diseases, especially metabolic and mental disorders. Circadian rhythms are generated at a single-cell level through a complex set of interlocked genetic feedback loops. Individual components of the circadian network are considered "sloppy" due to stochastic noise, and it is only through the interaction of cellular oscillators at a network level that precise rhythms are generated. Medically treating or reverse-engineering this complicated genetic architecture necessitates mathematical understanding at multiple physical scales, from cells to tissue.

This thesis seeks to describe the complex dynamics and hierarchical organization of circadian rhythms in mammals through systems dynamics and mathematical approaches. The overarching theme of this work will be the interplay between stochasticity and synchronization in circadian rhythms. Stochastic noise and precise oscillation are not completely at odds, however. In this thesis, I first develop a model of the circadian oscillator which incorporates the core negative feedback loop and an important neuropeptide coupling pathway. I use this model to investigate claims about the roles of *Cryptochrome* isoforms within the core circadian clock, and show that despite seemingly-different roles, experimental data is consistent with a parallel

role for *Cryptochrome* isoforms. Next, I present a method for inferring functional connections within the suprachiasmatic nucleus (SCN), the mammalian "master clock," and describe the network structure within the SCN. Finally, I examine growth and development of the SCN *in utero*.

Contents

1 Introduction **1**
- 1.1 Circadian Model Organisms . . . 2
- 1.2 Organization of Mammalian Circadian Rhythms . . . 3
- 1.3 Mathematical Approaches . . . 6
- 1.4 Biological Methods . . . 8
- 1.5 Significance Within Chemical Engineering . . . 9

2 The Role of Cryptochrome in the Cell-Autonomous Oscillator **13**
- 2.1 Background . . . 13
- 2.2 Model Construction and Parameter Identification . . . 14
- 2.3 Results and Discussion . . . 24

3 Functional Network Structure of the Suprachiasmatic Nucleus **28**
- 3.1 Background . . . 28
- 3.2 Experimental Design and Inference Methodology . . . 29
- 3.3 Results and Discussion . . . 32

4 Kinetic Monte Carlo Simulations in Python with GillesPy **40**
- 4.1 Background . . . 40
- 4.2 Using Gillespy . . . 41
- 4.3 Distribution . . . 42

5 Future Directions **43**
- 5.1 Temporal Separation in SCN Signal Processing . . . 43
- 5.2 Growth and Development of the SCN . . . 47

Appendix **58**
- S1 Example for Constructing and Simulating a GillesPy Model . . . 58

List of Figures

1.1 The mammalian cell-autonomous circadian clock 5
1.2 Intercellular signalling pathways in the suprachiasmatic nucleus 7
1.3 Example of a mouse actogram 10
1.4 Schematic of the *Per2::Luc* fusion 11

2.1 VIP model diagram and wild-type simulation results 15
2.2 Experimental and VIP model *Cry* knockout phenotypes 26
2.3 VIP model correctly predicts fraction of oscillatory energy in circadian range 27

3.1 Experimental resynchronization protocol for network inference 30
3.2 MIC identifies strongest connections within each SCN 33
3.3 Functional node degree distribution within the SCN 35
3.4 Heatmap of neuronal node degree in each SCN 36
3.5 Area under the ROC curve (AUC) for three simulated network types, inferred by MIC 37
3.6 Area under the ROC curve (AUC) for the inferred network types. . . . 39

5.1 Signal processing in the SCN 45
5.2 Development of the SCN 48

S1 Gillespy simulation example 60

List of Tables

2.1 Ordinary differential equations comprising the coupled circadian model 21
2.2 Parameter descriptions for the coupled circadian ODE model 22
2.3 Components of the model fitness function for optimizing a parameter set 23

3.1 SCN Network Characteristics 35

Chapter 1

Introduction

Nearly all life on Earth exists in a temporally periodic environment. The 24-hour day-night cycle brings with it a wide variety of cyclic effects, including changes in temperature, light, and radiation. Circadian rhythms function as a biological feed-forward controller, adapting the physiology and behavior of an organism to these predictable changes in its environment. Endogenous circadian rhythms have been observed in nearly all species, from simple prokaryotic cyanobacteria, to plants, insects, and mammals [1–4]. While the genetic mechanism controlling these rhythms varies, circadian rhythms share three characteristics: they are endogenous, they are temperature-compensated, and they are entrainable [5]. Importantly, the endogenous or self-sustaining nature of circadian rhythms indicates that they are not merely responses to environmental cues, but instead are generated by underlying biological processes.

Circadian rhythms are thought to have first evolved to protect metabolites in simple organisms from harsh solar ultraviolet radiation [6]. In complex organisms, circadian rhythms control gene expression across a wide portion of the genome. Approximately 10% of all transcripts expressed in human tissue are circadian-regulated,

affecting nearly every biochemical pathway [7]. As such, circadian regulation impacts many biological processes including cell cycles, body temperature, metabolism, insulin sensitivity, and activity patterns [8–11].

Developing a quantitative understanding of circadian rhythms is desirable from both a medical and an engineering perspective. With the emergence of a 24-hour society, perturbations to natural circadian rhythms are common. Chronic disturbances, such as shift work, have been associated with metabolic disorders [12]. Irregular circadian rhythms have also associated with neurodegenerative and psychiatric disease [13]. Additionally, due to metabolic differences, medicines taken at different phases of the circadian cycle differ in effect or effectiveness. Medically, it is desirable to develop pharmacological or therapeutic interventions to strengthen circadian rhythms and improve overall health. To do so requires a strong understanding of the circadian clock. For an engineering standpoint, the circadian network is a complex yet precise gene network. Understanding how these networks are organized is desirable in order to create synthetic genetic networks for genetic engineering. The circadian system provides an interesting and physiologically-relevant case study for a mathematical understanding of gene networks.

1.1 Circadian Model Organisms

As with much of biology, circadian rhythms are studied through the use of model organisms. These model organisms include: *Drosophila melanogaster* (fruit fly, insect), *Neurospora crassa* (spore mold, fungus), *Arabidopsis thaliana* (thale cress, plant), *Synechococcus elongatus* (cyanobacteria, bacteria), and *Mus musculus* (common lab mouse, mammal). Notably missing from this list is *Escherichia coli*, as these bacteria are not thought to have circadian rhythms. In many organisms (such as fruit flies and

mice), the circadian clock is conserved to a degree. Additionally, eukarotes have been shown to use transcription-translation feedback loops to generate circadian rhythms, whereas the few prokaryotic oscillators (such as that in cyanobacteria) use progressive-phosphorylation oscillators. Interestingly, this means that the KaiABC oscillator of *S. elongatus* may be reconstituted in a test tube. In this thesis, I focus on the mammalian circadian oscillator, through the model organism *Mus musculus*.

1.2 Organization of Mammalian Circadian Rhythms

Circadian rhythms are generated at a single-cellular level, and are considered to be cell-autonomous. In multicellular organisms, cell-autonomous clocks must interact at tissue and system levels to establish a coherent phase across the organism. For mammals this task is accomplished by a "master clock," the suprachiasmatic nucleus (SCN) [14]. The SCN is a small region of the hypothalamus, consisting of approximately 20,000 neurons. It receives light input from the optic chasm to entrain to external day-night cycles. Signals from the SCN synchronize peripheral cellular oscillators which control gene expression in tissues throughout the body.

Single-Cell Rhythms

The mammalian cell-autonomous circadian oscillator is comprised of interlocking genetic transcription-translation feedback loops (TTFLs), as shown in Fig. 1.1 [4]. The central negative feedback loop driving rhythmicity is the *Period-Cryptochrome* (*Per-Cry*) loop, in which *Per* and *Cry* are transcribed and translated, form heterodimers, and re-enter the nucleus to repress their own transcription through sequestration of CLOCK-BMAL1 Enhancer box (E-box) activators. PER-CRY heterodimers are then slowly degraded [15]. As transcription repressors are degraded, transcription re-

activates, creating a cycle with a period of approximately 24 hours. Clock-controlled genes are then regulated by participants in this core loop, through transcription regulators including E-boxes, DBP/E4BP4 binding elements (D-boxes), and RevErbA/ROR binding elements (RREs) [16]. This circadian network has been identified in many cell types including fibroblast, adipose, neuronal, liver, and skeletal cells [17–20].

Despite this complex systems of feedback loops, single cells do not oscillate with precise rhythms [21, 22]. Rather, extrinsic and intrinsic cellular noise lead to variability in period length at a single-cell level [23]. Extrinsic sources of noise include environmental conditions and differences in the physical makeup of each cell. Intrinsic sources of noise include low copy numbers of oscillator components and diffusion of biochemical species within the cell. Because each cell within an organism is individually "sloppy," synchronization between cells and an organizational hierarchy is required to achieve precise rhythms at an organism-level.

Organism-Wide Rhythms

In the suprachiasmatic nucleus, cells exchange neuropeptides, including VIP, AVP, and GABA, to maintain synchrony of the oscillator population [14, 24]. Fig. 1.2 shows many of the signaling pathways thought to play a role in maintaining synchrony in the SCN. Explanted SCN maintain precise rhythms for over a month when plated in the absence of any external cues. Oscillators in peripheral tissues are thought to lack paracrine signaling, and as such do not maintain coordinated rhythms in the absence of an entraining signal. Entraining cues from the SCN as well as time-dependent feeding, rather than intercellular coupling, are required to maintain precise, coordinated rhythms in peripheral tissue. Likewise, destruction of the SCN in otherwise-healthy animals results in loss of rhythmic behavior, while SCN transplants restore

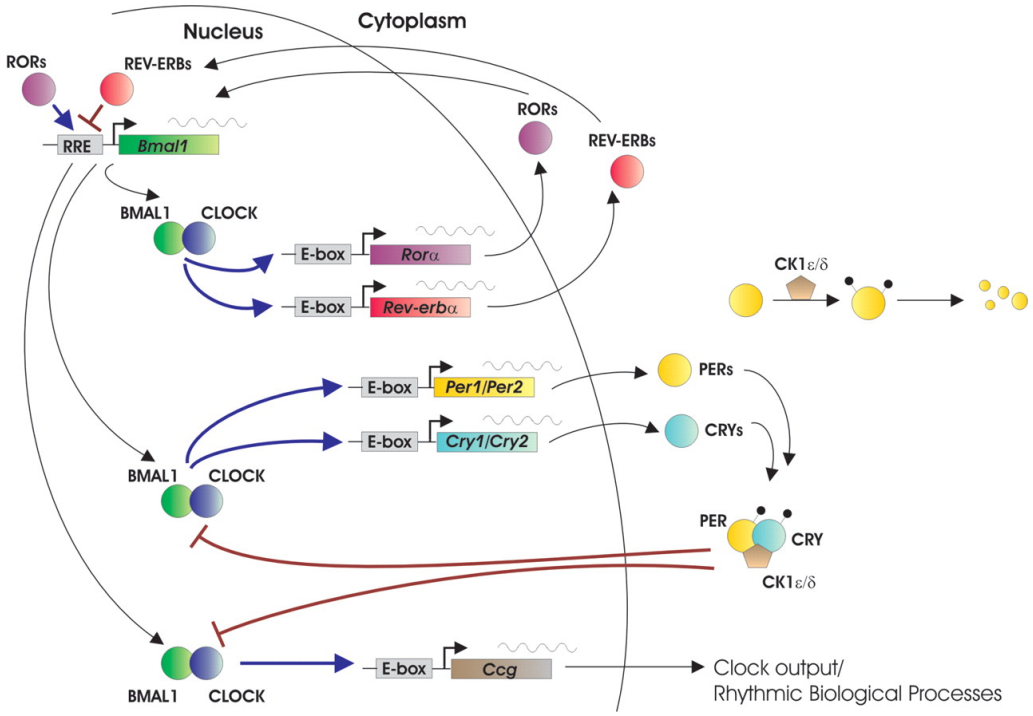

Figure 1.1: The mammalian cell-autonomous circadian clock. The mammalian clock is comprised of interlocking genetic feedback loops. The core negative loop is the *Per-Cry* loop. *Ccg* represents a family of clock-controlled genes, here symbolically controlled by an E-box region. Figure from [4].

rhythms across the organism [25].

1.3 Mathematical Approaches

The inherent complexity in circadian genetic networks necessitates a quantitative approach to understanding the underlying dynamics. Dynamics of gene expression are frequently modeled using ordinary differential equation (ODE) approaches. When coupled, these systems of ODEs can be used to represent and understand complex genetic networks. Historically, ODE modeling is a well-established approach to understanding circadian rhythm dynamics [26–29]. ODE models are of the form:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t), p), \tag{1.1}$$

where $\mathbf{f}$ is comprised of rate equations for biochemical states $\mathbf{x}$, and is parameterized by p. It is also common to treat the circadian gene network as an attractive *limit cycle oscillator*, in which Eqn. 1.1 satisfies:

$$\lim_{t\to\infty}[\mathbf{x}(t) - \mathbf{x}(t+\tau)] = 0 \tag{1.2}$$

indicating that concentrations $\mathbf{x}$ oscillate with period τ. ODE models are easily constructed, but are not easily parameterized as limit cycle oscillators. While highly useful, ODE approaches cannot capture intrinsic molecular fluctuations which play an important role in circadian behavior.

Stochastic approaches, through Gillespie (or kinetic Monte Carlo) algorithms, are computationally expensive, but directly capture intrinsic molecular noise due to low copy numbers. Stochastic approaches have also been previously applied to circadian rhythms at a single-cell level [30, 31]. Rather than using a deterministic set of

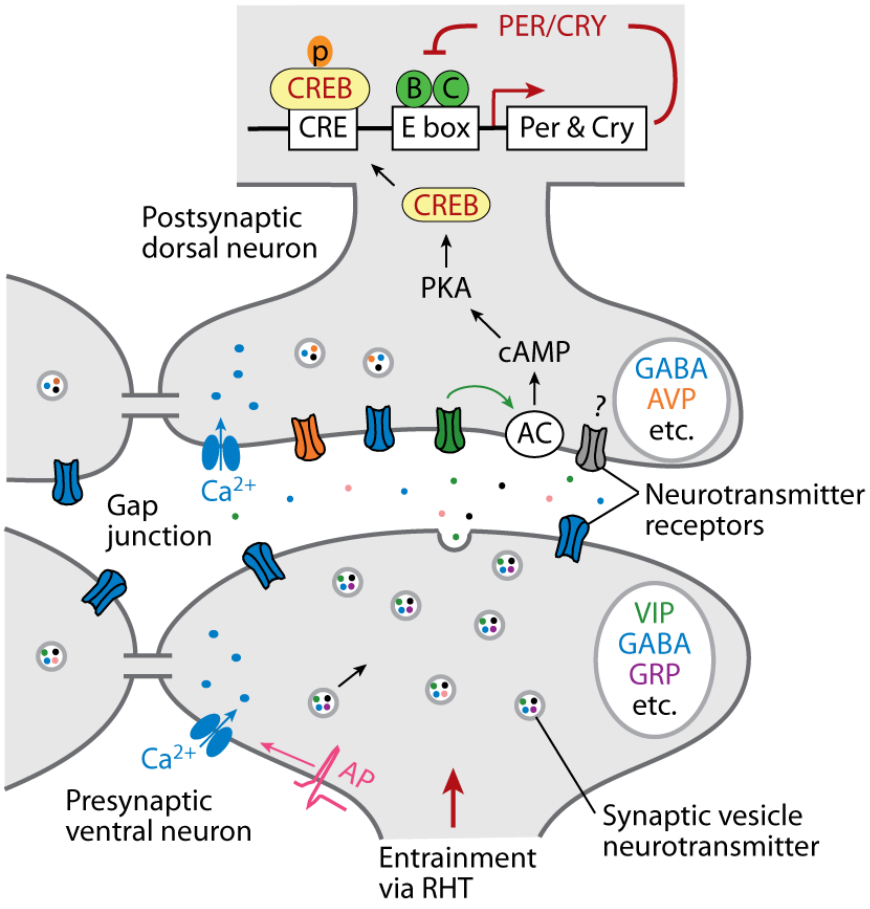

Figure 1.2: Intercellular signalling pathways in the suprachiasmatic nucleus. Figure from [14].

rate equations, stochastic simulation algorithms solve the chemical master equation (CME), of the form:

$$\frac{d\mathbf{s}}{dt} = \mathbf{A}(t)\mathbf{s}, \tag{1.3}$$

where $\mathbf{s}$ is a vector of n states s_i and $\mathbf{A}$ is an $n \times n$ matrix of propensities. As $\mathbf{A}$ may depend on $\mathbf{s}(t)$, the resulting Markov process is nonstationary.

For this work, both ODE and stochastic approaches will be used, as they are each suitable for specific situations. Additionally, understanding circadian data requires mathematical tools from a variety of fields, including sensitivity analysis, discrete signal processing, and statistics; these methods will be introduced *ad hoc*.

1.4 Biological Methods

The research presented in this thesis is purely computational, however mathematical understanding necessitates understanding of experiment.

The earliest method for understanding circadian behavior in mammals is through recordings of actograms. Actograms are generally recordings of wheel activity of mice or other rodents, but have also been used to monitor circadian behavior of larger animals including bears. As shown in Fig. 1.3, when under 12:12 light:dark (LD) conditions, mouse wheel running starts at the same time each day. When in constant darkness (DD), the intrinsic period of the mouse clock is slightly shorter than 24 hours, leading to the slanted rhythms shown. Genetic or environmental perturbations may affect this wheel running behavior, and these phenotypes are captured via actograms. Actogram still often used to record behavioral reactions to genetic perturbation. In Chapter 2, actogram recordings are among the data used in fitting a model of the circadian clock to long or short period knockout phenotypes.

Aside from common biological techniques, circadian research often uses the *Pe-*

riod2::Luciferase bioluminescent reporter to observe circadian gene expression [32]. This fusion allows the real-time observation of *Period* production in individual cells and also in cells within a tissue. A gene for destabilized *Luciferase* (*dLuc*) is attached to the *Period* propoter region (Fig. 1.4). The destabilized protein produced is degraded rapidly, so the output state of the *Period* gene is near-instantaneously observable. These methods have largely replaced wheel-running behavior, and are currently the state-of-the-art method for analyzing circadian phase, as they provide increased information about the state of the core circadian oscillator. *Per2::Luc* bioluminescence traces are used extensively to monitor single-cell rhythms in Chapters 2 and 3.

1.5 Significance Within Chemical Engineering

Chemical engineering is based on the fundamental principles of thermodynamics, transport phenomena, and reaction kinetics. This thesis is not constrained to those topics alone, but is grounded in them. Most prominently, the equations describing genetic network dynamics are equivalent to, and based on, chemical kinetic equations. The rate equations governing biochemical species take familiar forms: mass-action, Michaelis-Menten, and Hill type kinetics. Mathematically, these equations may be solved identically to the more traditional reactor design equations. Here, control theory can be applied as well–concepts such as frequency domain analysis and transfer function model representations may yield significant insight into the dynamics of biochemical pathways. When these systems of reactions are broken down into stochastic processes, statistical mechanics concepts apply as well. The Gillespie algorithm, commonly used in stochastic systems biology, is effectively a kinetic Monte Carlo scheme [34]. As in statistical mechanics, fluctuations can play a significant role in system behavior, leading to dynamics that cannot be well-approximated by deterministic

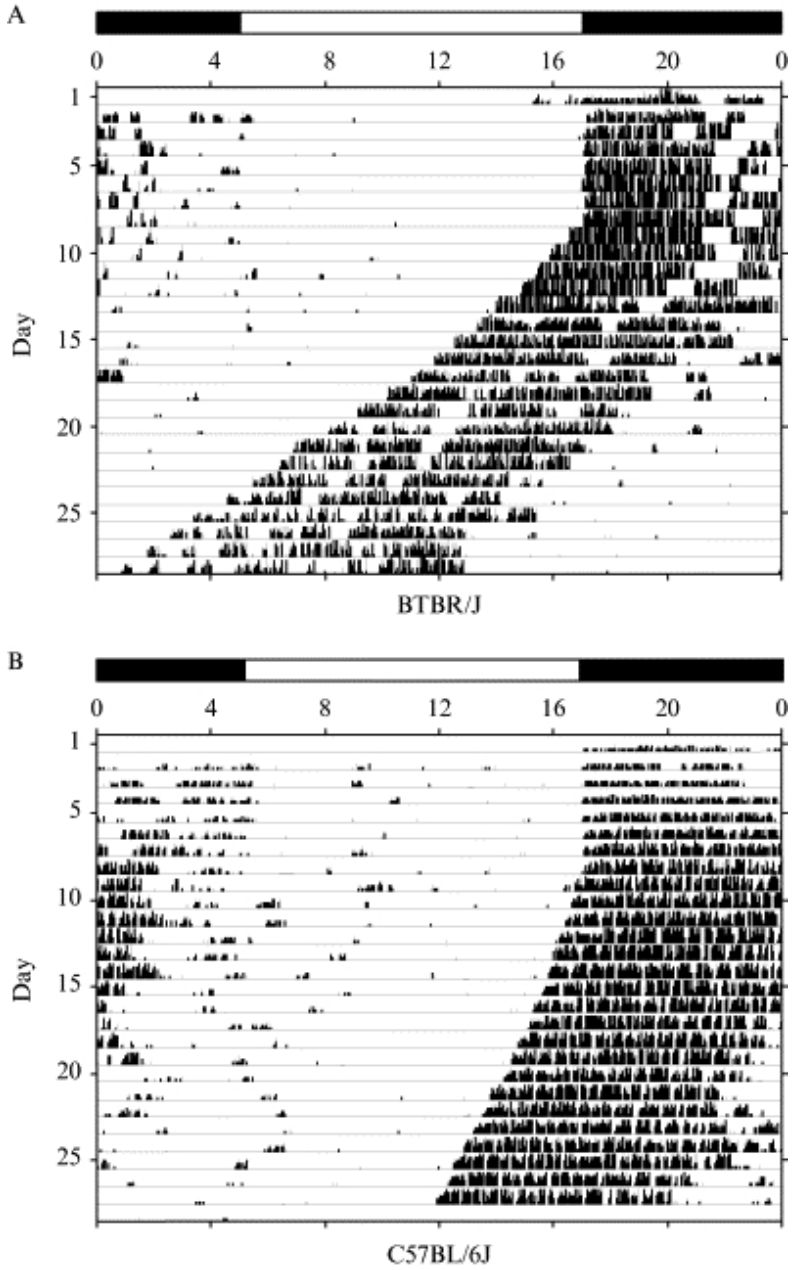

Figure 1.3: Example of a mouse actogram. Figure from [33]. Two mouse genotypes are shown in their free-running (DD) periods. (A) BTBR/J mouse. (B) C57BL/6J mouse. The difference in period length is due only to the different mouse strains shown. In LD conditions, period is precisely 24 hours.

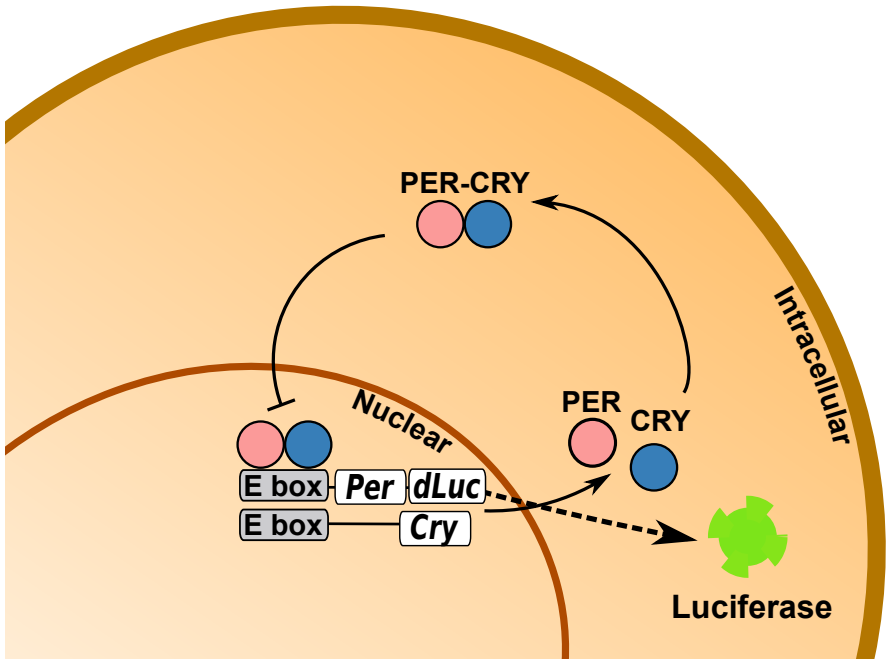

Figure 1.4: Schematic of the *Per2::Luc* fusion. Destabilized luciferase degrades rapidly (half-life 30 hours), and therefore provides the ability to monitor *Per* expression in real-time.

solutions.

Transport phenomena plays a significant role in intra- and intercellular movement of biomacromolecules. To this point, mathematical research in circadian biology has primarily considered the interior of the cell to be spatially homogeneous. While this is physically incorrect, it is a useful approximation. Future models may account for transport between compartments of the cell with simple rate equations, using either partial differential equation or spatial stochastic solvers [35]. A strong understanding of diffusive processes can assist in determining the validity of these approximations and building physical understanding of cellular dynamics. Although intercellular transport is thought to be primarily synaptic, diffusive connections are thought to be the primary communication between the SCN and peripheral oscillators [25].

Mathematics from chemical engineering may also arise unexpectedly. For example, the phase of uncoupled stochastic oscillators was recently captured with a reaction-diffusion equation [36].

Chapter 2

The Role of Cryptochrome in the Cell-Autonomous Oscillator [1]

2.1 Background

Within the mammalian feedback loop, there exist multiple isoforms of core clock genes *Period* and *Cryptochrome*: *Per1,2,3* and *Cry1,2* [4, 37]. Particular interest has been shown in the redundant *Cry*s, due to differing effects on period length and single-cell rhythmicity when knocked out in mice [18, 38]. $Cry1^{-/-}$ mice display shortened circadian periods in constant darkness (DD), and individual neurons from $Cry1^{-/-}$ mice show arrhythmic oscillation. $Cry2^{-/-}$ mice have longer periods in DD, and retain rhythmicity in individual neurons. Importantly, whole-tissue SCN explants in $Cry1^{-/-}$ mice retained rhythmicity, indicating that intercellular coupling plays

[1] Portions of Section 2 appear in J. H. Abel, L. A. Widmer, P. C. St. John, J. Stelling, and F. J. Doyle III, "A Coupled Stochastic Model Explains Differences in Circadian Behavior of Cry1 and Cry2 Knockouts," *IEEE Life Sciences Letters*, vol. 1, no. 1, pp. 3-6, 2015.

a role in rescuing rhythms despite single cell arrhythmicity [18, 39]. Recent work has proposed that *Cry1* and *Cry2* play a parallel role in the oscillator [15], however, other studies have suggested that parallel roles are inconsistent with the observed knockout rhythmicity [40]. Alternately, phase delays or strength of repression have been suggested as possible causes for these differences [41, 42].

In this section, the parallel role of *Cry1* and *Cry2* is examined mathematically through the construction of a coupled stochastic model of the *Per-Cry* feedback loop. Results indicate that stochastic noise and relative protein abundances are sufficient to account for the difference in single-cell rhythmicity. This confirms that observed knockout behavior is in fact consistent with a parallel role for *Cry* isoforms.

2.2 Model Construction and Parameter Identification

Model Components

A model of the core oscillator was constructed to capture essential components of the *Per-Cry* feedback loop, and intercellular coupling through the neurotransmitter vasoactive intestinal peptide, *VIP* (Fig. 2.1A). The model explicitly includes states: *Per*, *Cry1*, *Cry2*, and *VIP* mRNA; PER, CRY1, CRY2, and VIP protein; and PER-CRY1, PER-CRY2 and CREB transcription factors. CRY1 and CRY2 have parallel roles in the model, and differ in degradation rates and abundances, as experimentally shown [15, 43]. This resulted in an 11-state model with 34 kinetic parameters. For a full list of equation and parameters, please see Tables 2.1 and 2.2.

Rate Laws

Kinetic rates were captured by either mass-action terms (for translation and dimerization reactions) or Michaelis-Menten terms (for enzyme-mediated degradation, shared

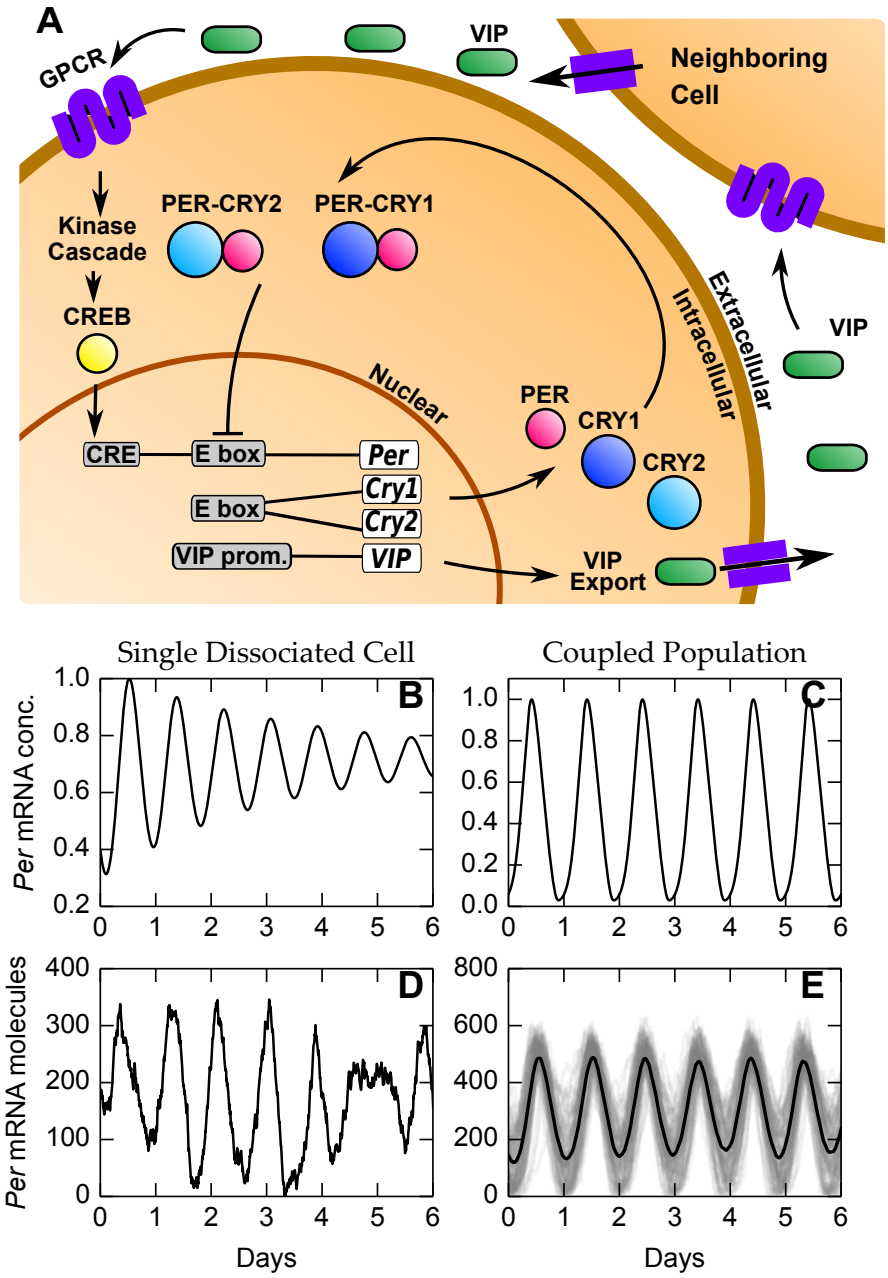

Figure 2.1: VIP model diagram and wild-type simulation results. (A) A schematic of model components and interactions. Results of deterministic (B-C) and stochastic (D-E) solution of a single dissociated cell (center) and coupled model (right). *Period* mRNA is the state shown. Model cells are decoupled by setting VIP concentration to 0.

enzyme-mediated degradataion, or transcription factors). The model did not require nonphysical Hill terms (those with cooperativity greater than unity) to form a limit-cycle oscillator, in part due to the effects of the positive feedback VIP loop [44].

A "shared degradation" term occurs when a single enzyme degrades two species competitively. Here, a single enzyme (FBXL3, denoted "E") is considered to have primary responsibility for nuclear degradation of CRYs (denoted "S1" and "S2") [45]:

$$[\mathrm{E}] + [\mathrm{S1}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\mathrm{ES1}] \xrightarrow{k_{\mathrm{d},1}} [\mathrm{E}]$$

$$[\mathrm{E}] + [\mathrm{S2}] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [\mathrm{ES2}] \xrightarrow{k_{\mathrm{d},2}} [\mathrm{E}].$$

As in [15], identical equilibrium constants are assumed for CRY1 and CRY2 binding with FBXL3. However, $k_{d,1}$ and $k_{d,2}$ differ. Upon simplification and use of the pseudo-steady state hypothesis (as in [15]), this yields shared degradation terms of:

$$r_{d,1} = \frac{V_{\mathrm{max},1}[\mathrm{S1}]}{K_M + [\mathrm{S1}] + [\mathrm{S2}]}$$

$$r_{d,2} = \frac{V_{\mathrm{max},2}[\mathrm{S2}]}{K_M + [\mathrm{S1}] + [\mathrm{S2}]}.$$

Similarly, we derive terms for the co-repression of *Cry* expression, assuming equal affinity of nuclear CRY1-PER and CRY2-PER (repressors R1 and R2) for the E-box binding site, resulting in transcription rates of the form:

$$r_{trans} = \frac{V_{\mathrm{max}}}{K_M + [\mathrm{R1}] + [\mathrm{R2}]}.$$

Derivation of an Activation Co-repression Rate

Unlike *Cry* expression, *Per* transcription is additionally impacted by CREB through the C-regulatory element box, a transcriptional promoter. This requires the derivation of an activation and co-repression equation, where CREB is not necessary for gene expression, but increases the rate at which transcription occurs. Because the CRY-PER transcription repressors act at the E-box, we assume they do not interfere with binding of CREB. Thus, with nuclear CRY-PER repressors R1 and R1, with CREB activator A1, with *Per* gene G, and *Per* mRNA M, the set of possible reactions at the

Per gene is as follows:

$$[\mathrm{G}] \xrightarrow{k_\mathrm{t}} [\mathrm{G}] + [\mathrm{M}]$$

$$[\mathrm{G}] + [\mathrm{R1}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\mathrm{GR1}]$$

$$[\mathrm{G}] + [\mathrm{R2}] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [\mathrm{GR2}]$$

$$[\mathrm{GA}] \xrightarrow{k_\mathrm{prom}} [\mathrm{GA}] + [\mathrm{M}]$$

$$[\mathrm{G}] + [\mathrm{A}] \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} [\mathrm{GA}]$$

$$[\mathrm{GA}] + [\mathrm{R1}] \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} [\mathrm{GAR1}] \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} [\mathrm{GR1}] + [\mathrm{A}]$$

$$[\mathrm{GA}] + [\mathrm{R2}] \underset{k_{-6}}{\overset{k_6}{\rightleftharpoons}} [\mathrm{GAR2}] \underset{k_{-7}}{\overset{k_7}{\rightleftharpoons}} [\mathrm{GR2}] + [\mathrm{A}]$$

There are ultimately two mRNA-producing reactions here, involving the gene without activator or repressor bound, or the gene with only the activator bound. Here, we make simplifying assumptions about repressor binding, namely that R1 and R2 bind with equal affinity, and activator and repressor binding do not affect each other. For repressor binding and dissociation:

$$k_1 = k_2 = k_4 = k_6 \equiv k_{rb}, \tag{2.1}$$

$$k_{-1} = k_{-2} = k_{-4} = k_{-6} \equiv k_{rd}; \tag{2.2}$$

and for activator binding and dissociation:

$$k_3 = k_{-5} = k_{-7} \equiv k_{ab}, \tag{2.3}$$

$$k_{-3} = k_5 = k_7 \equiv k_{ad}. \tag{2.4}$$

Furthermore, we define equilibrium constants K_R and K_A in their standard form:

$$K_R \equiv \frac{k_{rd}}{k_{rb}} \tag{2.5}$$

$$K_A \equiv \frac{k_{ad}}{k_{ab}} \tag{2.6}$$

By invoking the equilibrium approximation, we can assume that equilibrium is formed between the gene and its transcription factors much more rapidly than gene expression. Thus, we have two mRNA-producing equations:

$$r_g = v_g[\mathrm{G}]$$

$$r_{ga} = v_{ga}[\mathrm{GA}].$$

Additionally, by performing a balance on G (that is, noting that [G] + [GA] + [GR1] + [GR2] + [GAR1] + [GAR2] = [G_0]), we can find the concentrations [G] and [GA]:

$$[\mathrm{G}] = [\mathrm{G}_0] \left(\frac{1}{1 + \frac{[\mathrm{R1}]}{K_R} + \frac{[\mathrm{R2}]}{K_R}} \right) \left(1 - \frac{\frac{[\mathrm{A}]}{K_A}}{1 + \frac{[\mathrm{A}]}{K_A}} \right) \tag{2.7}$$

$$[\mathrm{GA}] = [\mathrm{G}_0]\left(\frac{1}{1+\frac{[\mathrm{R1}]}{K_R}+\frac{[\mathrm{R2}]}{K_R}}\right)\left(\frac{\frac{[\mathrm{A}]}{K_A}}{1+\frac{[\mathrm{A}]}{K_A}}\right). \tag{2.8}$$

By assuming that K_A is large, that is $k_{ad} \gg k_{ab}$, and refactoring $[\mathrm{G}_0]$, K_A, and K_R into rates v_A and v_R, we ultimately arrive at our simplified activation co-repression transcription rate equation:

$$r_{trans} = \frac{v_A[\mathrm{A}] + v_R}{K_R + [\mathrm{R1}] + [\mathrm{R2}]}. \tag{2.9}$$

This equation is ultimately used for *Per* mRNA production only.

Model Fitting

To model a single dissociated neuron, external VIP concentration was set to 0 to mimic the lack of physical connections in plates of mechanically-dissociated neurons. To model a population, VIP was shared between the nearest four neighboring cells on a 15-by-15 2D grid, with periodic (toroidal) boundaries. Parameter optimization was performed by fitting the data to experimental mRNA and protein stoichoimetry, relative amplitudes, phase relationships, and RNAi knockdown experiments (for a full list, see Table 2.3). Additionally, the model was fit at a population rather than a single-cell level, with the VIP coupling pathway active. The coupled cell was required to be a limit cycle oscillator Parameters and equations were converted to stochastic propensities through a volume parameter Ω, fit by the desynchronization rate of uncoupled oscillators [36, 46]. The optimization and simulation was performed with the CasADi computer algebra package, the SUNDIALS ODE solvers suite, the DEAP evolutionary algorithm toolkit, and the StochKit2 stochastic simulation algorithm package [47–50].

Table 2.1: Ordinary differential equations comprising the coupled circadian model. [51]

State Variable	Symbol	Model Equation
Per mRNA	**p**	$\frac{d\mathbf{p}}{dt} = \frac{v_{1pp}\mathbf{CREB} + v_{2pr}}{K_{1p} + \mathbf{C1P} + \mathbf{C2P}} - \frac{v_{3p}\mathbf{p}}{K_{2dp} + \mathbf{p}}$
Cry1 mRNA	**c1**	$\frac{d\mathbf{c1}}{dt} = \frac{v_{4c1r}}{K_{3c} + \mathbf{C1P} + \mathbf{C2P}} - \frac{v_{5c1}\mathbf{c1}}{K_{4dc} + \mathbf{c1}}$
Cry2 mRNA	**c2**	$\frac{d\mathbf{c2}}{dt} = \frac{v_{6c2r}}{K_{3c} + \mathbf{C1P} + \mathbf{C2P}} - \frac{v_{7c2}\mathbf{c2}}{K_{4dc} + \mathbf{c2}}$
VIP mRNA	**vip**	$\frac{d\mathbf{vip}}{dt} = \frac{v_{8vr}}{K_{5v} + \mathbf{C1P} + \mathbf{C2P}} - \frac{v_{9v}\mathbf{vip}}{K_{6dv} + \mathbf{vip}}$
Per Protein	**P**	$\frac{d\mathbf{P}}{dt} = k_{1p}\mathbf{p} - \frac{v_{10P}\mathbf{P}}{K_{8dP} + \mathbf{P}} - v_{11aCP}\mathbf{P} \times \mathbf{C1} - v_{11aCP}\mathbf{P} \times \mathbf{C2} + v_{12dCP}\mathbf{C1P} + v_{12dCP}\mathbf{C2P}$
Cry1 Protein	**C1**	$\frac{d\mathbf{C1}}{dt} = k_{2c}\mathbf{c1} - \frac{v_{13C1}\mathbf{C1}}{K_{9dC} + \mathbf{C1}} - v_{11aCP}\mathbf{P} \times \mathbf{C1} + v_{12dCP}\mathbf{C1P}$
Cry2 Protein	**C2**	$\frac{d\mathbf{C2}}{dt} = k_{2c}\mathbf{c2} - \frac{v_{14C2}\mathbf{C2}}{K_{9dC} + \mathbf{C2}} - v_{11aCP}\mathbf{P} \times \mathbf{C2} + v_{12dCP}\mathbf{C2P}$
VIP Protein	**VIP**	$\frac{d\mathbf{VIP}}{dt} = k_{3v}\mathbf{vip} - v_{15V}\mathbf{VIP}$
CRY1-PER Dimer	**C1P**	$\frac{d\mathbf{C1P}}{dt} = v_{11aCP}\mathbf{P} \times \mathbf{C1} - v_{12dCP}\mathbf{C1P} - \frac{v_{16C1P}\mathbf{C1P}}{K_{10dCn} + \mathbf{C1P} + \mathbf{C2P}}$
CRY2-PER Dimer	**C2P**	$\frac{d\mathbf{C2P}}{dt} = v_{11aCP}\mathbf{P} \times \mathbf{C2} - v_{12dCP}\mathbf{C2P} - \frac{v_{17C2P}\mathbf{C2P}}{K_{10dCn} + \mathbf{C1P} + \mathbf{C2P}}$
CREB Protein	**CREB**	$\frac{d\mathbf{CREB}}{dt} = \frac{v_{18V}\mathbf{VIP}}{K_{11V} + \mathbf{VIP}} - \frac{v_{19CR}\mathbf{CREB}}{K_{12dCR} + \mathbf{CREB}}$

Table 2.2: Parameter descriptions for the coupled circadian ODE model. [51]

Parameter	Description	Value	Units
v_{1pp}	CREB-induced *Per* mRNA promotion	0.235	[-]/hr
v_{2pr}	*Per* mRNA transcription	0.415	$[-]^2$/hr
v_{3p}	*Per* mRNA degradation	0.478	[-]/hr
v_{4c1r}	*Cry1* mRNA transcription	0.350	$[-]^2$/hr
v_{5c1}	*Cry1* mRNA degradation	1.44	[-]/hr
v_{6c2r}	*Cry2* mRNA transcription	0.124	[-]/hr
v_{7c2}	*Cry2* mRNA degradation	2.28	[-]/hr
v_{8vr}	*VIP* mRNA transcription	0.291	$[-]^2$/hr
v_{9v}	*VIP* mRNA degradation	1.35	[-]/hr
v_{10P}	*Per* protein degradation	13.0	[-]/hr
v_{11aCP}	PER-CRY dimer formation	0.493	$([-]\times \text{ hr})^{-1}$
v_{12dcp}	PER-CRY dimer dissociation	0.00380	1/hr
v_{13C1}	*Cry1* protein degradation	4.12	[-]/hr
v_{14C2}	*Cry2* protein degradation	0.840	[-]/hr
v_{15V}	*VIP* protein degradation	0.723	1/hr
v_{16C1P}	PER-CRY1 dimer degradation	0.0306	[-]/hr
v_{17C2P}	PER-CRY2 dimer degradation	0.0862	[-]/hr
v_{18V}	CREB activation by VIP receptors	0.789	[-]/hr
v_{19CR}	CREB deactivation	1.27	[-]/hr
k_{1p}	PER translation	7.51	1/hr
k_{2c}	CRY translation	0.572	1/hr
k_{3v}	VIP translation	5.50	1/hr
K_{1p}	*Per* transcription constant	0.264	[-]
K_{2dp}	*Per* degradation constant	0.00795	[-]
K_{3c}	*Cry* transcription constant	0.156	[-]
K_{4dc}	*Cry* degradation constant	1.94	[-]
K_{5v}	*VIP* transcription constant	0.115	[-]
K_{6dv}	*VIP* degradation constant	0.110	[-]
K_{7dP}	*Per* protein degradation constant	0.0372	[-]
K_{8dC}	*Cry* protein degradation constant	4.23	[-]
K_{9dCn}	PER-CRY dimer degradation constant	0.0455	[-]
K_{10V}	CREB protein activation constant	1.46	[-]
K_{11CR}	CREB protein deactivation constant	1.01	[-]

Table 2.3: Components of the model fitness function for optimizing a parameter set. [51] Criteria 1-11, 14-18 from [43, 52]; criteria 12, 13 from [18, 38, 43]; criterion 20 from [53]; and criteria 19, 21 from [54].

Index	Description	Weight	Desired	Result
1	*Per* mRNA Peak-trough ratio	0.5	Large	Large
2	*Cry1* mRNA Peak-trough ratio	0.5	2.16	2.30
3	*Cry2* mRNA Peak-trough ratio	0.5	2.24	2.20
4	*Per* protein Peak-trough ratio	3	Large	Large
5	*Cry1* protein Peak-trough ratio	3	3.247	2.41
6	*Cry2* protein Peak-trough ratio	3	1.98	1.60
7	Fraction PER of total PER, CRY1, CRY2	3	0.10	0.06
8	Fraction CRY1 of total PER, CRY1, CRY2	3	0.56	0.63
9	Fraction CRY2 of total PER, CRY1, CRY2	3	0.34	0.31
10	*Cry1* siRNA sensitivity	5	< 0	< 0
11	*Cry2* siRNA sensitivity	5	> 0	> 0
12	*Cry1* knockout period	12	< 0.95	0.89
13	*Cry2* knockout period	12	> 1.15	1.16
14	Fraction PER-CRY1 of total CRY1	1	0.40	0.22
15	Fraction PER-CRY2 of total CRY2	1	0.35	0.10
16	t_{max} mRNAs - t_{max} complexes	3	0.75	0.81
17	t_{max} cytocsolic protein - t_{max} mRNAs	3	0.25	0.01
18	t_{max} PER-CRY - t_{max} cytosolic protein	3	0	0.18
19	t_{max} CREB - t_{max} mRNAs	8	> 0.80	0.93
20	*VIP* protein peak-trough ratio	5	3.00	3.48
21	CREB peak-trough ratio	5	3.00	2.76
22	Deterministic coupled cells must synchronize	20	True	True

2.3 Results and Discussion

In the wild-type (WT) single-cell deterministic case (Fig. 2.1B), the model shows damped oscillations. This may appear to conflict with self-sustained oscillations observed in WT individual neurons, however, damped oscillators driven by noise cannot be experimentally distinguished from noisy limit cycles [55]. Here as well, single-cell stochastic simulation displays sustained oscillation (Fig. 2.1D). When coupled, the single cell oscillators cross a Hopf bifurcation to form a deterministic limit cycle (Fig. 2.1C), and maintain synchrony in the stochastic case (Fig. 2.1E).

When compared to experimental results from [18] (Fig. 2.2), both single-cell (A-F) and population (G-L) simulations show good agreement with experimental data. In single dissociated cells, WT and *Cry2*$^{-/-}$ appear significantly more rhythmic than *Cry1*$^{-/-}$. Furthermore, *Cry* knockouts display increased *Per* expression (309% of WT for *Cry1*$^{-/-}$, 104% of WT for *Cry2*$^{-/-}$) in both experiment and simulation. This is consistent with a parallel role for *Cry* isoforms, in which *Per* expression must compensate for decreased *Cry* expression in order to repress transcription. Coupling is shown to maintain rhythms and synchrony for all genotypes in experiment and in simulation, as it causes the cells to cross a Hopf bifurcation.

When using a quantitative rhythmicity metric, the fraction of energy in the circadian discrete wavelet bin (Fig. 2.3), model oscillatory behavior closely agrees with experiment. It is important to note that this stochastic behavior was not tuned or fit, but rather is a natural consequence of parallel *Cry* roles with more abundant and stable *Cry1*. Ultimately, this indicates that parallel roles for *Cry* are consistent with experimental evidence. This also underlines the essential role stochastic noise and mathematical analysis plays in understanding the circadian oscillator, as this consistency is not explicitly evident from experiment. Lastly, this result suggest that single

dissociated cell circadian gene networks may not be limit cycle oscillators, but rather fixed points with noise driven oscillation.

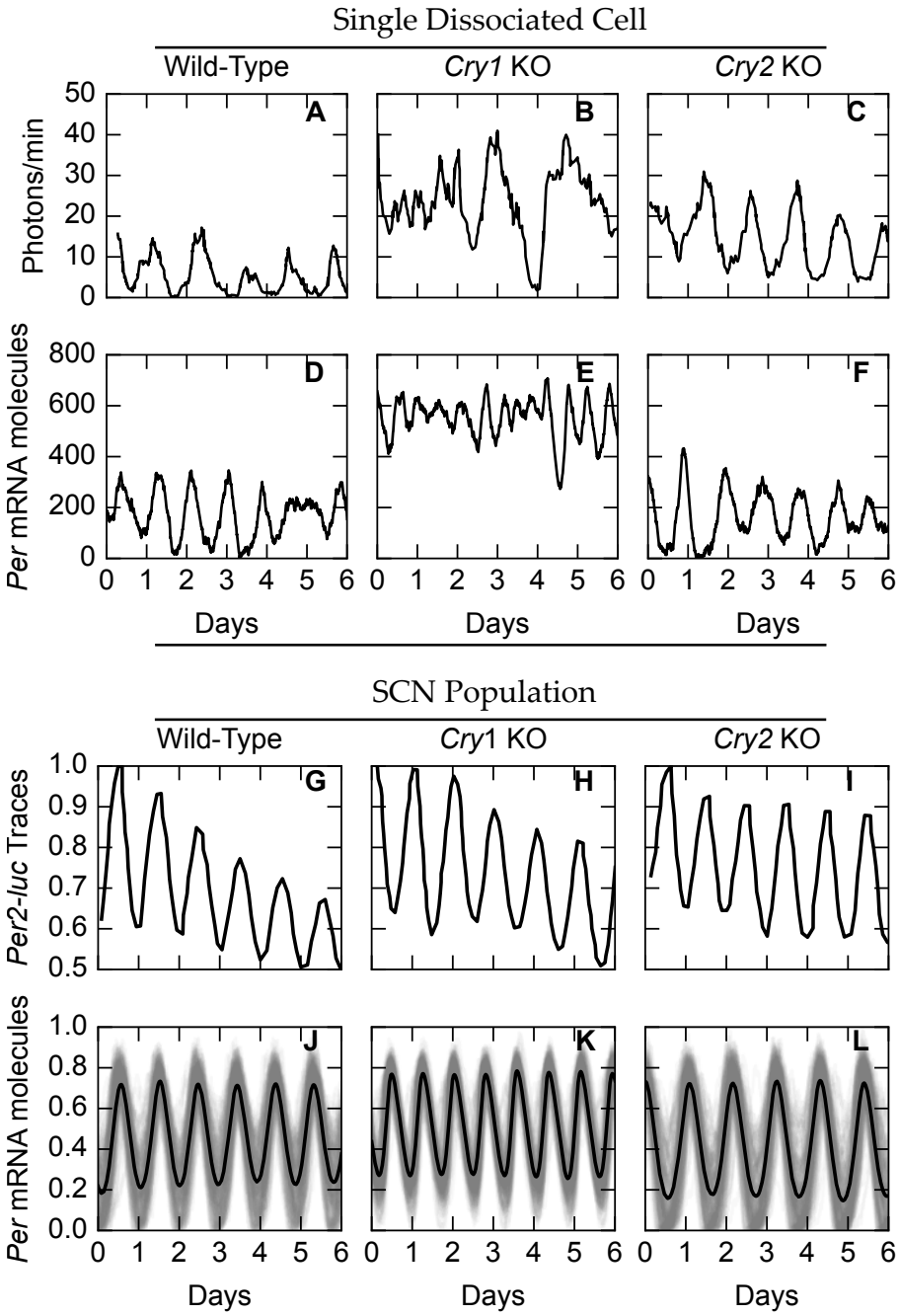

Figure 2.2: Experimental and VIP model *Cry* knockout phenotypes. Experimental *Per2::Luc* bioluminescence traces (A-C, G-I) and simulation *Per* mRNA trajectories (D-F, J-L) are shown. In single cells (A-F), *Cry1* knockout leads to decreased rhythmicity. Coupling (G-L) restores rhythms and synchrony for all genotypes. Experimental data (from [18]) and simulation show good agreement.

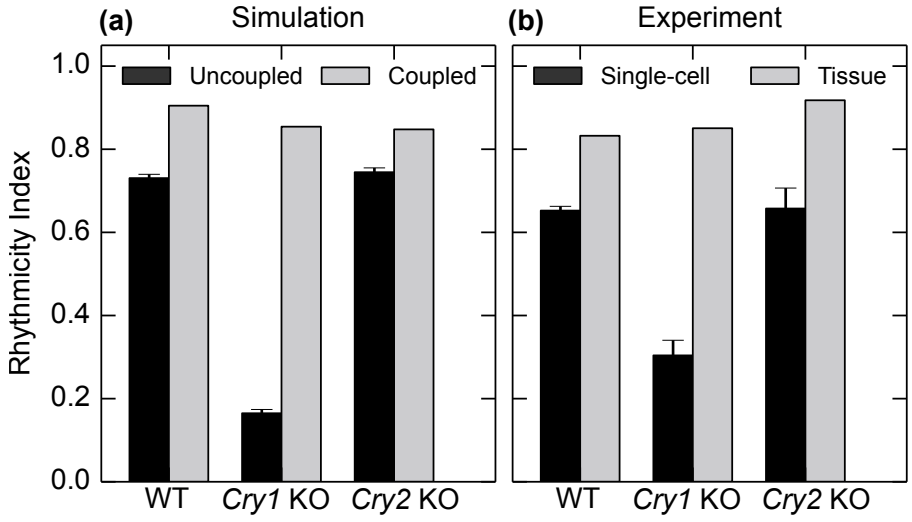

Figure 2.3: VIP model correctly predicts fraction of oscillatory energy in circadian range. Single uncoupled cell error bars represent variance. Coupled cell and tissue energy was calculated from mean of individual cell trajectories within the coupled population.

Chapter 3

Functional Network Structure of the Suprachiasmatic Nucleus [1]

3.1 Background

The suprachiasmatic nucleus is the mammalian "master clock," responsible for synchronizing the body to external light cues, and coordinating rhythms across systems in the body. Collective dynamics of neurons within the SCN is contingent upon properties of the underlying cell-autonomous oscillator, communication pathways through neuropeptides such as VIP and γ-aminobutyric acid (GABA), and the overall structure of the network connecting these neurons [44, 56–59]. The cell-autonomous oscillator and neuropeptide coupling effects are relatively well-understood, however, the network structure has remained elusive. Previous works (including that in Section 2) have assumed nearest-neighbor, small-world, or mean-field networks on a semi-arbitrary basis [27, 59–62]. Recent work has focused on fast-scale coupling of

[1] Portions of Section 3 are in preparation for submission as J. H. Abel[†], K. Meeker[†], D. Granados-Fuentes, P. C. St. John, T. Wang, B. Bales, F. J. Doyle III, E. D. Herzog, and L. R. Petzold, "Functional Network Inference of the Mammalian Suprachiasmatic Nucleus." All experimental procedures were performed by members of the Herzog lab.

neuron firing rather than the slow-scale biochemical coupling tied to synchronization [63].

Inferring slow-scale neuropeptidergic coupling in the SCN presents significant challenges. Cells within the SCN display tightly-synchronized low-frequency non-stationary oscillations in gene expression, preventing the use of common techniques such as transfer entropy, directed transfer functions, or between-sample analysis of connectivity (BSAC) [63–66]. In this section, the functional network within the SCN is examined through a novel tetrodotoxin (TTX)-driven resynchronization experiment, along with the application of the maximal information coefficient (MIC) [67]. *Functional* connections are mathematical (non-physical) connections between correlated nodes. These connections are not necessarily indicative of underlying physical connections, but rather aim to examine how the network functions. MIC is ideally suited to this network inference, as it can detect nonlinear relationships between noisy oscillatory states. Results indicate that the SCN is functionally connected with an exponentially-distributed small-world network. The hubs of this network are largely constrained to the ventrolateral SCN ("core") region. This structure is thought to be optimally robust to random error, and rapidly synchronize to day-night cycles.

3.2 Experimental Design and Inference Methodology

In the suprachiasmatic nucleus, oscillation is highly synchronized. To identify connections within the SCN, it is necessary to perturb those oscillations away from synchrony. Tetrodotoxin (TTX), a neurotoxin carried in several species of fish, provides a unique means of weakening and desynchronizing these connections. TTX blocks Na+ channels and represses circadian output [69]. Within several days of TTX washout, phase relationships between SCN cells are restored to match pre-TTX phases [70].

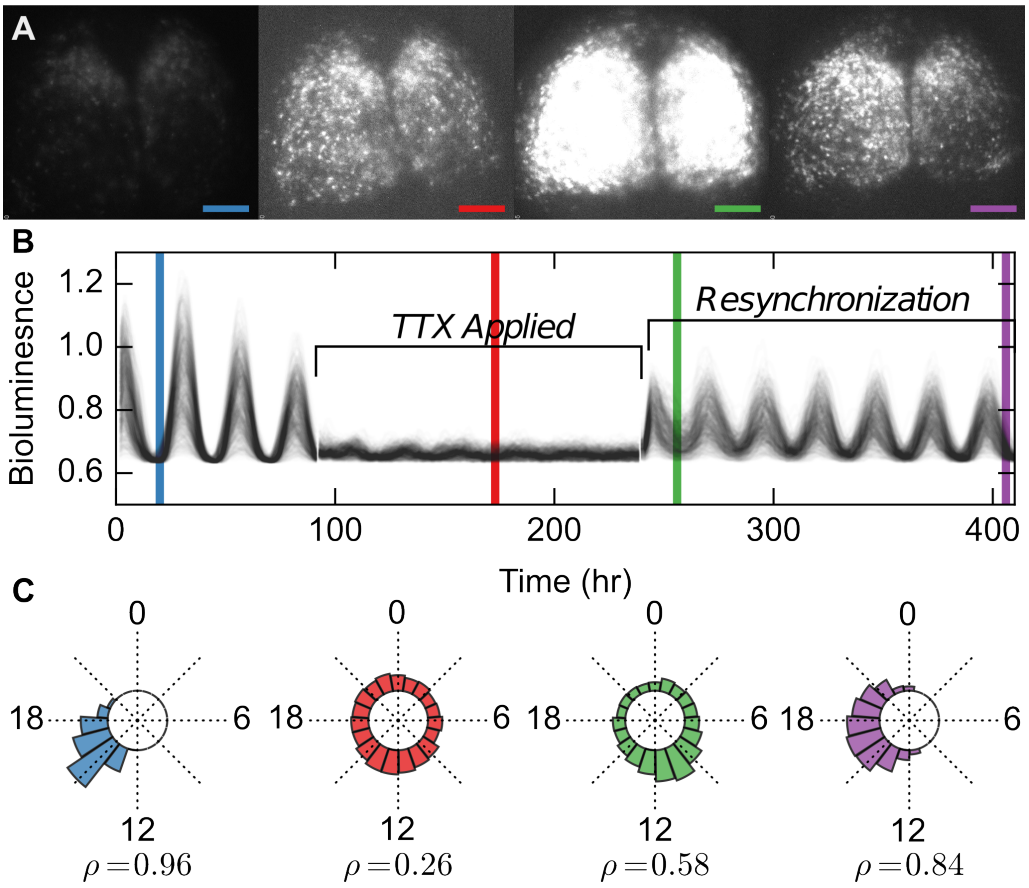

Figure 3.1: Experimental resynchronization protocol for network inference . (A) Microscope images taken at four points during the protocol. (B) Bioluminescence traces from neurons identified within the SCN. Units are count $\times 10^3$ per hour. (C) Radial histogram of phase, and calculated Kuramoto parameter ρ at each of the four locations highlighted in (A) and (B) [68].

This brief resynchronization allows for differentiation between tightly-connected cells and more loosely connected cells, which take longer to synchronize.

For this experiment, SCN were explanted from 7 day old homozygous *Per2:Luc* mice housed in 12:12 L:D conditions. TTX is applied, as described in [71], and remains in the meduim for six days of imaging. Examples of this imaging are shown as Fig. 3.1A. The TTX is then washed out by full-volume medium exchange. Bioluminescence is then recorded for an additional 8-12 days, as the networks reform initial phase relationships. The full time course for SCN $\mathcal{A}$ is shown as Fig. 3.1B. Kuramoto parameter ρ and a radial histogram of phases at four points along the time course (corresponding to points demarcated in A and B) are shown in Fig. 3.1C.

The maximal information coefficient (MIC) applies an efficient pairwise binning algorithm to maximize the mutual information for a pair of trajectories. Mutual information $I(X;Y)$ is intuitively understood as the reduction in entropy of a discrete random variable X, $H(X)$, given by observing Y. That is,

$$I(X;Y) = H(X) - H(X|Y). \tag{3.1}$$

The entropy $H(X)$ is defined as

$$H(X) = -\sum_{x\in X} P_X(x)\log P_X(x), \tag{3.2}$$

and the conditional entropy $H(X|Y)$ is defined as

$$H(X|Y) = \sum_{y\in Y} P_Y(y)\Big[-\sum_{x\in X} P_{X|Y}(x|y)\log\left(P_{X|Y}(x|y)\right)\Big]. \tag{3.3}$$

Recalling that $P_{X|Y}(x|y) = P_{XY}(x,y)/P_Y(y)$, the mutual information takes the com-

mon form

$$I(X;Y) = \sum_{x \in X, y \in Y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x) P_Y(y)}. \tag{3.4}$$

MIC applies a gridding algorithm to X and Y to maximize $I(X;Y)$ with respect to partitioning. This removes the arbitrary nature of selecting a bin size for finding the mutual information of sampled continuous random variables, as bin selection affects raw mutual information score. A limit to bin numbers on X and Y, n_x and n_y, is applied to prevent the trivial binning of individual data points. This limit is set to

$$n_x \times n_y \leq N^{0.6}, \tag{3.5}$$

where N is the number of samples, as suggested by [67]. The maximum mutual information score possible for random variables X and Y given the above binning, is $\log \min(n_x, n_y)$. The MIC is therefore normalized by this value to restrict its range to $[0, 1]$. For this study, X and Y are bioluminescence traces from single neurons within the SCN. Although MIC has received significant criticism for reduced statistical power, all neurons within the SCN are correlated due to direct or indirect connections[72]. As such, the p-values associated with MIC are not used, rather, a between-sample analysis of correlation is used to identify a critical MIC threshold.

3.3 Results and Discussion

The resulting MIC connectivity matrices must be thresholded to provide an adjacency matrix of connections within each SCN. A threshold was selected by inferring connections between cells from every SCN (Fig. 3.2). Connections within the same SCN (Fig. 3.2A) are considered valid, while connections between SCNs from different animals (Fig. 3.2B) are invalid. The strongest connections identified by MIC are

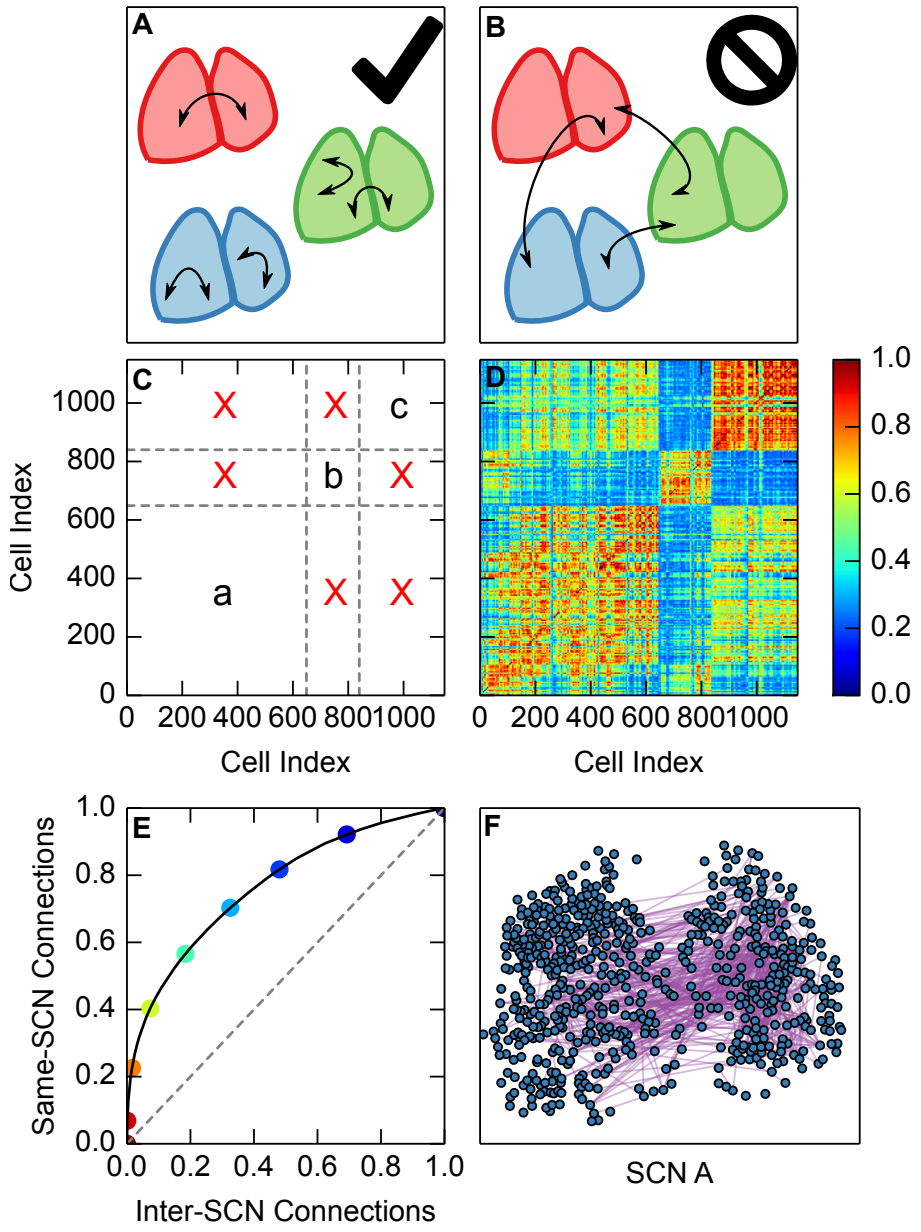

Figure 3.2: MIC identifies strongest connections within each SCN. (A) Connections identified within a single SCN are biologically feasible, whereas connections between distinct SCNs (B) are physically impossible. (C) An adjacency matrix for all SCNs combined identifies valid regions (a, b, c) and invalid regions (red X). (D) MIC scores for connections between cells in each SCN. (E) A pseudo-ROC curve shows that cells in feasible regions are preferentially identified as connected by MIC. Colors along this curve correspond to heatmap values in (D). This curve is useful in selecting a critical MIC threshold. (F) Example network of connections in SCN $\mathcal{A}$ identified with a threshold of $m_c = 0.98$.

primarily in "valid" regions (Fig. 3.2C,D). A pseudo-ROC of valid vs. invalid connections was then generated (Fig. 3.2E), and a threshold was selected to give a false positive rate of 0.005. This threshold was then scaled to result in an equal average node degree for each SCN.

Resulting network properties for each SCN are summarized in Table 3.1. The average path length is approximately equal to that of a randomly-generated network, and the clustering coefficient is much greater than that of an equivalent random network. Thus, the network may be classified as "small-world." Small-world networks commonly arise in biological systems, as they provide more rapid communication across larger spatial distances than nearest-neighbor networks.

In Fig. 3.3, it is shown the node degree distribution is best fit with an exponential fit. The exponential network is similar to a scale-free network in that it has many nodes with low node degree, and relatively fewer "hubs" of very high node degree. This network type is highly resistant to random error, as these errors are more likely to arise in nodes of low importance [73].

When spatially identified, hubs of these SCN networks are preferentially located in the ventrolateral SCN, also called the "core" (Fig. 3.4). This region has previously been identified as the primary producer of VIP, and the region which receives light input to entrain the clock [14]. This core-shell orientation, with hubs primarily in the core, may result in enhanced synchronization and entrainment properties for the SCN [60, 76].

This inference method identifies functional connections. In Fig. 3.5, the relationship between functional and physical connections is examined. The SCN resynchronization experiment was performed *in silico*, using networks of various topologies and node degrees, and two coupled models of the mammalian circadian network [51, 77]. The functional network was then inferred and compared back to the phys-

Table 3.1: SCN Network Characteristics

SCN	n_{cells}	L_{SCN}	L_{rand}	C^{Δ}_{SCN}	C^{Δ}_{rand}	S^{Δ}
$\mathcal{A}$	649	4.63	2.67	0.422	0.027	8.98
$\mathcal{B}$	191	2.88	2.13	0.497	0.098	3.77
$\mathcal{C}$	308	4.00	2.36	0.400	0.058	3.93

Characteristic path length L and clustering coefficient C^{Δ} [73, 74], and S^{Δ} [75] compared between the inferred network and equivalent randomly-generated networks indicate that the SCN has a small-world topology.

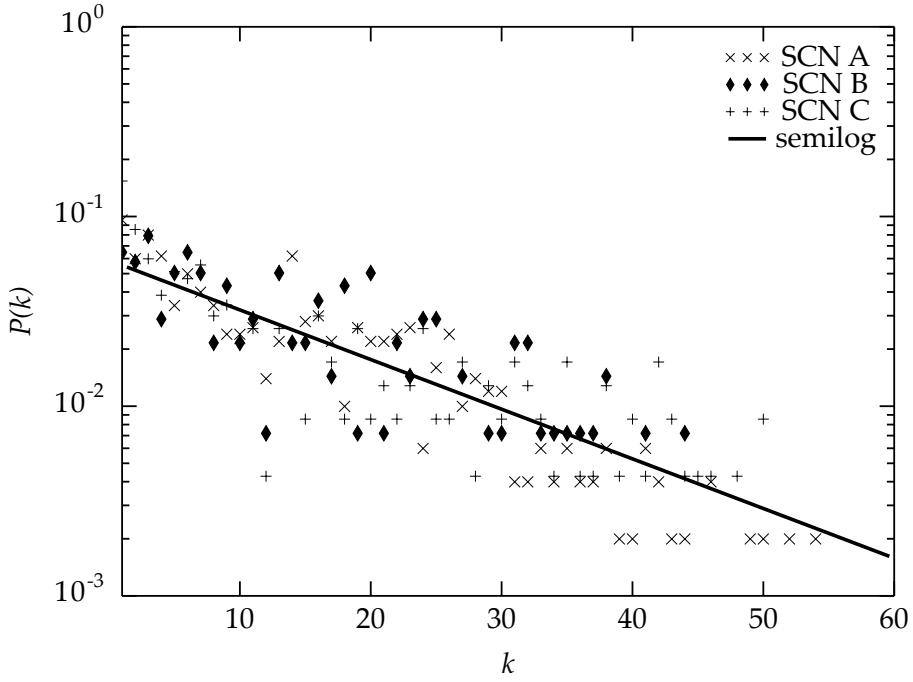

Figure 3.3: Functional node degree distribution within the SCN . Node degree distribution is best approximated with an exponential fit. Probability $P(k)$ of a node having degree k is plotted vs. node degree k. Linear fit on semilog plot $R^2 = 0.69$.

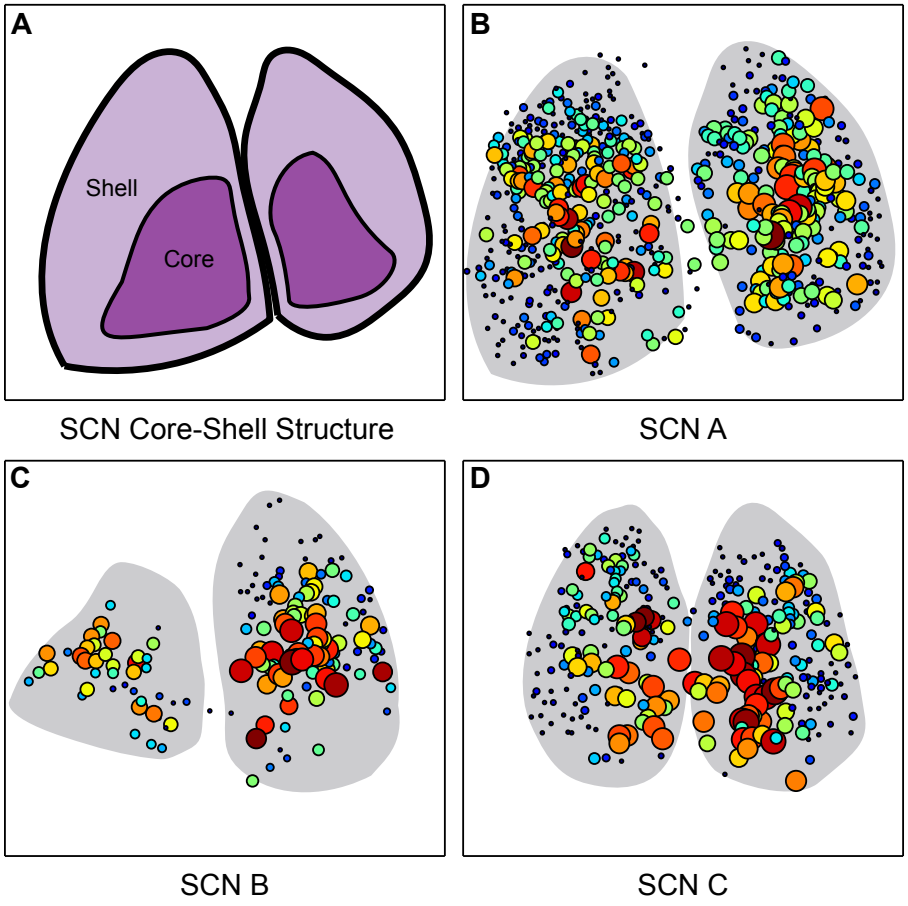

Figure 3.4: Heatmap of neuronal node degree in each SCN . (A) The SCN has been previously understood to have a core-shell structure. (B-D) Heatmaps show neurons with many connections are preferentially located in the central regions of the SCN. Node size and color are proportional to $\log k$.

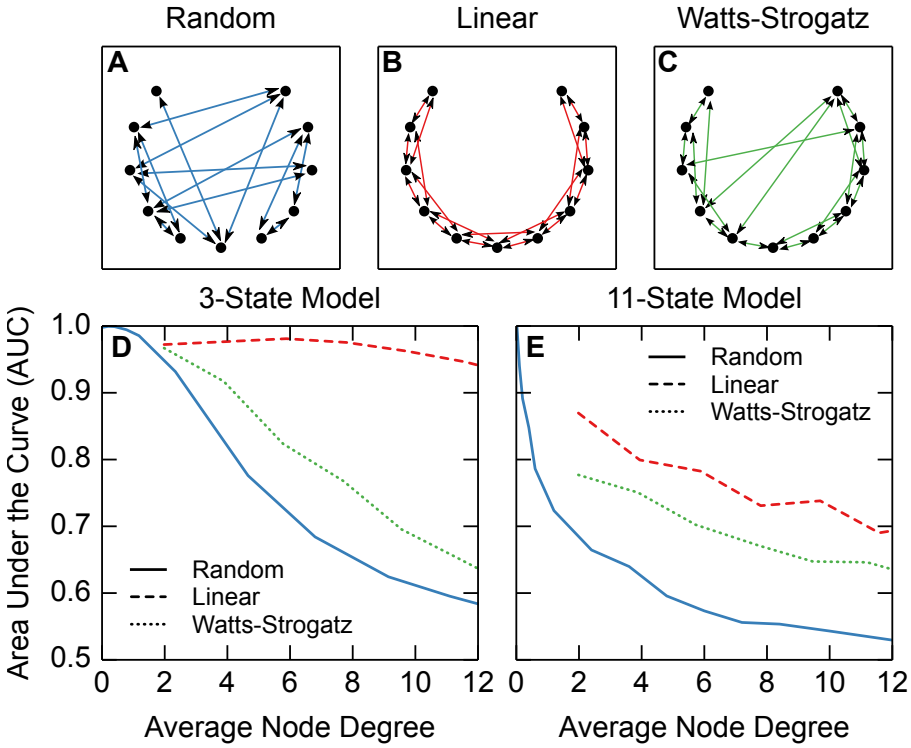

Figure 3.5: Area under the ROC curve (AUC) for three simulated network types, inferred by MIC . (A) Random, (B) linear, and (C) Watts-Strogatz $\beta = 0.1$ small-world networks were simulated with stochastic coupled circadian models, then inferred. Average AUC values are shown for a three state circadian model [77], and the VIP model as described in Section 2, for 100-cell networks with varying average node degree. This demonstrates limits on the ability of MIC to relate functional connections to underlying physical connections.

ical network. Area under the ROC curve (AUC) is used as a metric to determine ability to infer connections in these networks. Results show that the method loses the ability to detect physical connections in random and highly-connected networks. In these networks, indirect physical coupling is often identified as direct functional connections. The SCN networks $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ were also simulated, and inferred via *in silico Per* mRNA traces. Average AUC values for these networks were 0.89 and 0.78 for the models from [77] and [51], respectively (Fig. 3.6).

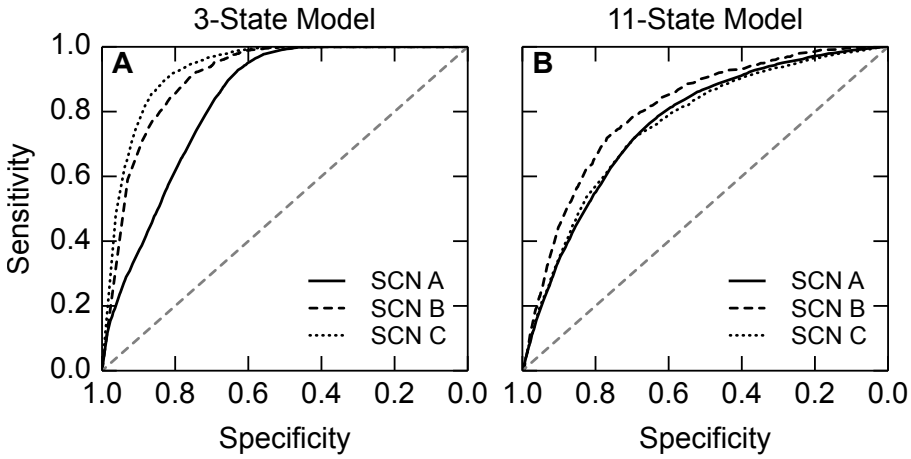

Figure 3.6: Area under the ROC curve (AUC) for the inferred network types. . The inferred networks were simulated with two models, (A) from [77], and (B) from [51]. Moderately high ROC values indicate that our functional connections are largely consistent with the underlying physical connections.

Chapter 4

Kinetic Monte Carlo Simulations in Python with GillesPy [1]

4.1 Background

Low gene and protein copy numbers within cells has been shown to result in intrinsic stochastic fluctuations in concentration. These fluctuations have been implicated in a variety of unique cellular constructs, such as noise-induced oscillators or toggle switches, and have been shown to play a role in cell differentiation and decision making [78, 79]. Stochastic noise cannot be captured by traditional ODE models, which accurately describe molecular dynamics only in the high-concentration limit. Instead of integrating ODEs, stochastic simulation algorithms (SSAs) solve a chemical master equation (CME) consisting of probabilistic propensities, and return an ensemble of solution trajectories, each of which represents a single possible solution to the evolution of the system over time.

[1] Portions of Section 4 will appear as: J. H. Abel[†], B. Drawert[†], A. Hellander, and L. R. Petzold, "GillesPy: A Python Package for Stochastic Model Building and Simulation," in *Proceedings of Foundations of Systems Biology in Engineering 2015*, Boston, MA, August 2015.

Since the definition of the CME, there have been many algorithms developed to simulate or approximately simulate the CME. The first and simplest of these, the "direct method," was first presented by D.T. Gillespie in 1977 [34]. Since then, additional methods have been developed to reduce the high computational cost of explicit simulation of the CME, including the optimized direct method (ODM), the next-reaction method (NRM), the composition-rejection method, and τ-leaping [80–83]. The StochKit2 suite of C-based SSAs provides a command-line interface to these solvers, and simulates stochastic models defined in .xml format [50].

In the process of model building, fitting, and simulation, it became necessary to write a Python module for simple and efficient stochastic model construction and simulation. It was decided that this module would be expanded for general release as the high-level interface to the StochKit SSAs (StochKit has a low-level command line interface, and StochSS provides a GUI). GillesPy (a portmanteau of *Gilles*pie and *Py*thon) is a resulting package for building and simulation of stochastic biochemical models in Python.

4.2 Using Gillespy

GillesPy is designed in a straightforward, object-oriented format in order to simply and efficiently create models. Model construction in GillesPy is performed using the `gillespy.model` object. Parameters (with values), species (with initial conditions), and reactions (either mass-action, or with a custom propensity function) are then attached to the model object. The model object may be output as .xml to use with StochKit at the command line, or instead simulated directly in Python through the `gillespy.StochKitSolver` object. This object is created with a model, the time of simulation, and any of the other StochKit-available simulation options (includ-

ing interval, realizations, specified solver). A simulation is performed by calling `StochKitSolver.run(model,**kwargs)`. The trajectories are returned in the form of a Numpy array for further processing. This allows stochastic simulations in GillesPy to be included in larger computational projects with ease, and allows the use of other third-party data analysis or modeling tools. An example for using this package in a simple dimer formation and degradation system is included in the Appendix.

4.3 Distribution

This package has since been incorporated into the PySB Python framework for modeling in systems biology as the method for performing stochastic simulations [84]. GillesPy is designed for use with Linux or Mac OS X, and is freely available under GPL version 3 at `http://github.com/GillesPy/gillespy`.

Chapter 5

Future Directions

5.1 Temporal Separation in SCN Signal Processing

Intercellular coupling in the mammalian brain is achieved through a variety of means, from rapid electrical stimuli, to slow neuropeptide signaling cascades. The differences in these physical processes results in a temporal separation in signal transmission. Electrical signals are transmitted in seconds, while neuropeptide signals may take minutes to even hours to process. This separation also results in frequencies propagating at varying amplification, depending on the dynamics of the coupling pathway and the frequency of the input signal. It is therefore possible to determine the temporal scale at which a neuronal signalling pathway is active by applying a frequency domain analysis.

This is evident from a simple thought experiment, in which we treat each step along a signalling pathway as a continuous stirred tank reactor (CSTR). Mathematically, a coupling pathway with two reactions and a pair of CSTRs are equivalent, as the representation for input x_1, intermediate x_2, output x_3 and kinetic rate parameters

k would be:

$$\begin{aligned} \frac{dx_2}{dt} &= f(x_1, x_2, \mathbf{k}) \\ \frac{dx_3}{dt} &= f(x_2, x_3, \mathbf{k}) \end{aligned} \tag{5.1}$$

where f consists of kinetic rate equations. For linear f, these equations form a linear time-invariant (LTI) system, and solution of this system is trivial for a wide range of signals via transfer function formulation [85]. Nonlinear systems may be linearized about a steady-state in order to apply equivalent analyses with reasonable accuracy. As with a cascade of CSTRs, this hypothetical biochemical signaling pathway forms a low-pass filter. In real biological systems, signaling cascades are more complex. Nonlinear Hill-type or Michaelis-Menten kinetic equations are often used to capture these cellular dynamics [86]. Initial studies of biochemical pathways have highlighted the importance of signal filtering properties in understanding the dynamics of cellular processes, however, few systems have been studied in any depth [87, 88].

Circadian systems present an additional challenge: they are not time-invariant, invalidating the steady-state assumption. Circadian systems could instead be linearized about pseudo-steady states at a range of phases, under the assumption that the clock moves slowly relative to the speed of the pathway. To validate that this transfer function linearization corresponds to the actual dynamics of the limit cycle oscillator, we can compare the linearized analytic solution to the results of simulating a perturbation in the system. The frequency response can then be determined by applying a Fourier transform at perturbation frequency to the perturbed input and the system output, and constructing a Bode magnitude plot.

In Fig. 5.1, we compare these results for two simple circadian models. Fig. 5.1A uses a simple first-order LTI system to demonstrate the accuracy of constructing a Bode plot via simulation and applying the Fourier transform. Because the system

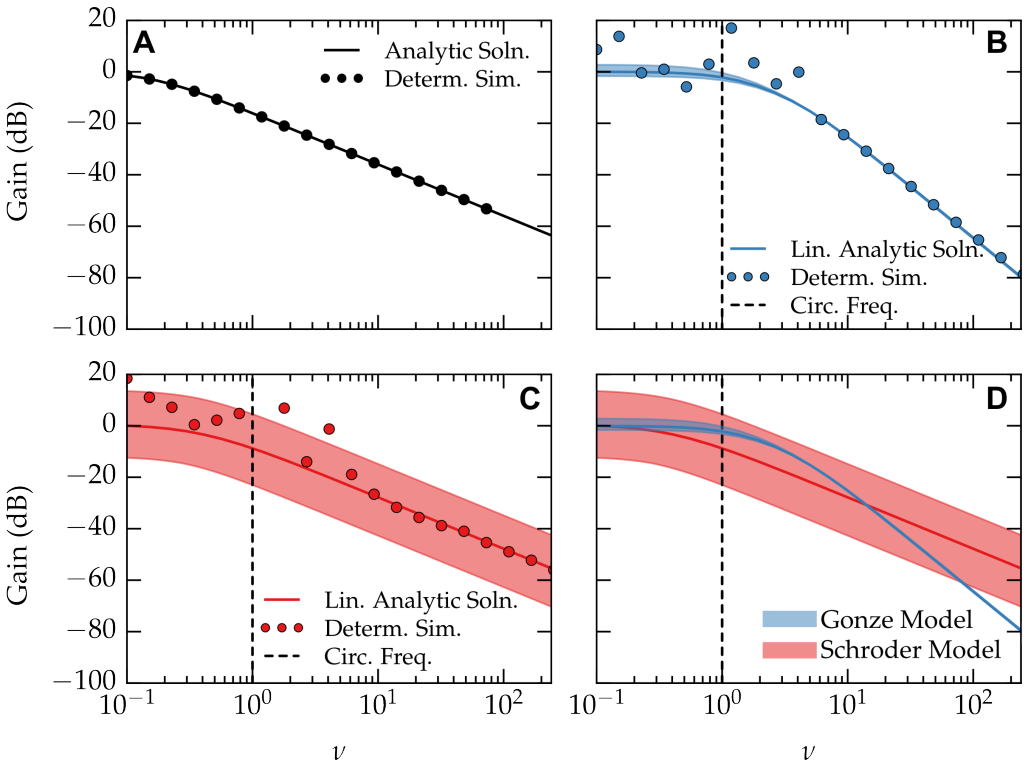

Figure 5.1: Signal processing in the SCN. Frequency domain analysis provides a method of analyzing coupling dynamics in the SCN. (A) a simple first-order LTI model validates the Fourier transform method of finding the gain. (B) Normalized transfer function dynamics of intercellular coupling in the model from [89]. The linearized analytic solution is plotted as a range of values, as it varies with phase. Weighted mean is shown as a darker line. This model is second-order. Dashed line represents the natural frequency of the oscillator. (C) Transfer function dynamics of intercellular coupling in the model from [77]. This model is first-order. (D) Comparison of analytic solution regions for these models, showing low-pass filter shape.

is LTI, the analytical solution is exact. The deterministic simulation results at 20 frequencies lie on the analytic solution, confirming that this method is accurate. In Fig. 5.1B and C, the analytic and deterministic simulation solutions for two models of coupling in the circadian clock are compared. The shaded regions represent the range of analytic solutions, which are phase dependent, and the solid line is the weighted average of these regions. The natural (circadian) frequency of each oscillator is normalized to 1.0 and identified with a dashed line. Fig. 5.1B is the solution to a second-order model from Gonze *et al.*, 2005 [89], and Fig. 5.1C is the solution to a first-order model from Schroder *et al.*, 2012 [77]. Both models predict that coupling in the SCN forms a low-pass filter, in which slow scale perturbations are propagated at a much higher amplitude. The analytic approximation and simulation results also show reasonable agreement at frequencies ~ 5 times that of the oscillator. At these frequencies, the pseudo-steady state hypothesis is more accurate. In Fig. 5.1D, the analytical solution regions are compared for models from B and C. Although the transfer function models of the system are of different orders, there is significant overlap and consistency in overall frequency response properties. Most importantly, this result provides clear support for the damping of high-frequency signal along the SCN synchronization pathway, which may be important in filtering errors and stochastic noise in single-cell rhythms.

This analysis could theoretically be expanded to additional models and pathways, and also experimental data, to determine the role of each coupling pathway between cells in the SCN by temporal function. Furthermore, this could be applied to any limit cycle oscillator model, which are common in models of neuron populations. Additional analyses should also be performed to determine a theoretical bound on error in linearization of the coupling pathway, and for limit-cycle systems in general. Furthermore, this methodology could be applied to stochastic systems and interest-

ing non-circadian biochemical pathways.

5.2 Growth and Development of the Suprachiasmatic Nucleus [1]

In continued collaboration with the Herzog lab, work has begun to examine the onset of rhythmicity and synchrony in the SCN during development. The gestation period of the common mouse is approximately 20 days, designated E(mbryonic day)1-20. Little is known regarding how the SCN develops and circadian rhythms are generated during this period. It has been recently shown that rhythms begin around day 14-15.

Fig. 5.2 contains initial results for developmental days E13-E17, a time thought to be critical in development of the SCN. Over this four-day period cells within the SCN become both rhythmic and synchronized, with the onset of rhythmicity around day E14.5-E15.5, and synchrony and period length stabilized by day E15.5. SCNs explanted from fetal mice were too small to identify individual neurons, instead, a pixel-based analysis was performed. Pixel regions of interest (ROI) were determined to be rhythmic if the detrended trajectory had its largest Lomb-Scargle periodogram peak, $p < 0.05$, for a period between 18 and 32 hours [90]. Detrending was performed via discrete wavelet transform by removing the lowest frequency (trend) bin (periods $>$ 128hr). A cosinor analysis was alternately performed, and showed comparable results. Phase ϕ, and Kuramoto synchrony parameter ρ, were calculated via a discrete wavelet detrend, restricting signal to the 16-32hr period range, then applying a Hilbert transform [68]. This initial result is highly interesting, but clearly does not

[1] This work was performed in collaboration with A. Sun, V. Carmona, and E. D. Herzog, Department of Biology, Washington University in St. Louis. All experimental work was performed by the Herzog lab.

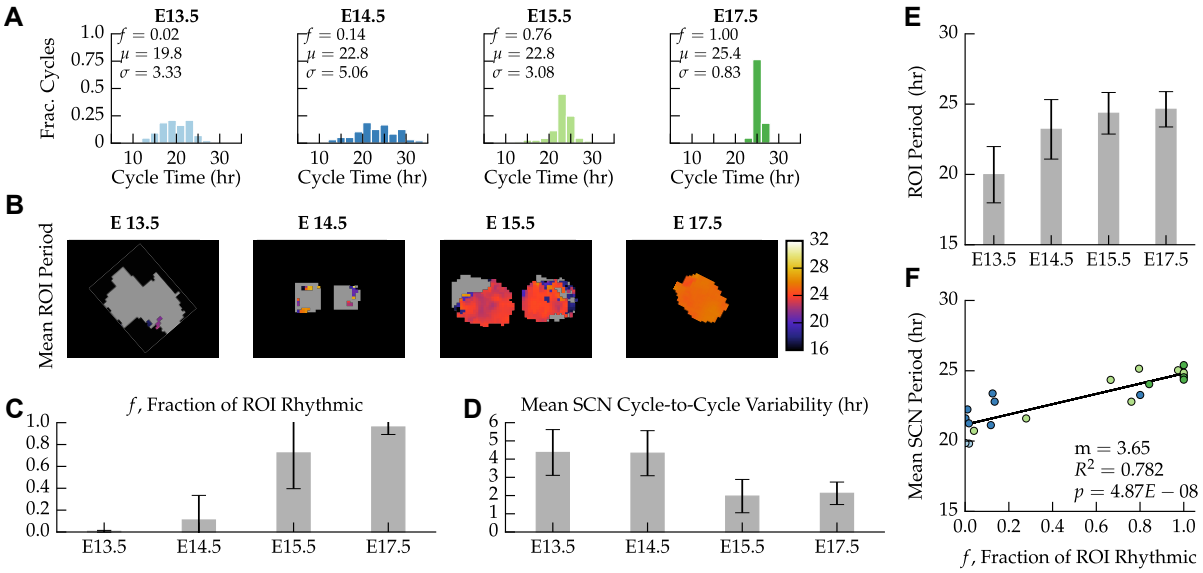

Figure 5.2: Development of the SCN. (A) Distribution of cycle lengths for all ROIs for the example SCNs shown in (B). The tightening of the distributions for E15.5 and E17.5 is indicative of oscillations stabilizing potentially due to intercellular communication. (B) Mean period of each ROI within the SCN. As cells begin to communicate, the period across the SCN stabilizes. (C) Fraction of ROI rhythmic at each developmental day (mean $\pm$ S.D.). Most SCNs become rhythmic at E15.5, but with significant variation due to different development rates of different samples. It is not until E17.5 that all ROI become rhythmic. (D) Also in day E15.5, mean SCN cycle-to-cycle variability decreases (mean $\pm$ S.D.), and (E) mean ROI period length reaches its final value (mean $\pm$ S.D.). (F) Mean SCN period is shown to increase with the fraction of ROI rhythmic (f) until approaching a final value of approximately 24.5 hours.

provide a complete explanation of network development.

These preliminary results will be expanded both in theory and experiment. Theoretically, the synchrony of each SCN will be examined, and tests for changes while the SCN is plated (growth under observation) will be performed. Experimentally, staining for neuropeptides will be performed, in order to implicate a mechanism for achieving synchrony of the fetal SCN. In the future, the network structure should also be examined, although a new method will be required, since TTX experiments are likely too harsh for the fetal SCN. This work provides a strong starting point for understanding the development of the SCN.

Future goals for this project include determining period-to-period variability, amplitude, and precision of oscillation throughout the stages of development. Staining for the presence of VIP and potentially GABA will be used to relate the onset of rhythmicity and synchrony to the production of specific neuropeptides implicated in circadian coupling. A more ambitious goal is inferring the network structure as the fetus develops, to determine how neuronal networks grow. It is thought that initial synchrony in the network is achieved before functional synapses are fully developed, and that multiple, possibly-redundant modes of synchronization are involved [91]. However, no detailed single-cell examination of the developing SCN network has been performed to date.

Bibliography

[1] M. Ishiura, S. Kutsuna, S. Aoki, H. Iwasaki, C. R. Andersson, A. Tanabe, S. S. Golden, C. H. Johnson, and T. Kondo, "Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria," *Science*, vol. 281, no. 5382, pp. 1519–1523, 1998.

[2] C. R. McClung, "Plant circadian rhythms," *Plant Cell*, vol. 18, no. 4, pp. 792–803, 2006.

[3] N. R. Glossop, L. C. Lyons, and P. E. Hardin, "Interlocked feedback loops within the Drosophila circadian oscillator," *Science*, vol. 286, no. 5440, pp. 766–768, 1999.

[4] C. H. Ko and J. S. Takahashi, "Molecular components of the mammalian circadian clock," *Hum. Mol. Genet.*, vol. 15, no. 2, pp. 271–277, 2006.

[5] J. C. Dunlap, J. J. Loros, and P. J. DeCoursey, *Chronobiology*. Sinauer Associates, Inc., 2004.

[6] S. S. Nikaido and C. H. Johnson, "Daily and circadian variation in survival from ultraviolet radiation in Chlamydomonas reinhardtii," *Photochem. Photobiol.*, vol. 71, no. 6, pp. 758–765, 2000.

[7] P. L. Lowrey and J. S. Takahashi, "Mammalian circadian biology: elucidating genome-wide levels of temporal organization," *Annu. Rev. Genomics Hum. Genet.*, vol. 5, no. 47, pp. 407–441, 2004.

[8] R. Refinetti and M. Menaker, "The circadian rhythm of body temperature," *Physiol. Behav.*, vol. 51, no. 3, pp. 613–637, 1992.

[9] J. Bieler, R. Cannavo, K. Gustafson, C. Gobet, D. Gatfield, and F. Naef, "Robust synchronization of coupled circadian and cell cycle oscillators in single mammalian cells," *Mol. Syst. Biol.*, vol. 10, no. 7, p. 739, 2014.

[10] P. Sassone-Corsi, "Molecular clocks: mastering time by gene regulation," *Nature*, vol. 392, no. 6679, pp. 871–874, 1998.

[11] J. Bass and J. S. Takahashi, "Circadian integration of metabolism and energetics," *Science*, vol. 330, no. 6009, pp. 1349–1354, 2010.

[12] E. Haus and M. Smolensky, "Biological clocks and shift work: Circadian dysregulation and potential long-term effects," *Cancer Causes Control*, vol. 17, no. 4, pp. 489–500, 2006.

[13] K. Wulff, S. Gatti, J. G. Wettstein, and R. G. Foster, "Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease," *Nat. Rev. Neurosci.*, vol. 11, no. 8, pp. 589–599, 2010.

[14] D. K. Welsh, J. S. Takahashi, and S. A. Kay, "Suprachiasmatic nucleus: cell autonomy and network properties," *Annu. Rev. Physiol.*, vol. 72, pp. 551–577, 2010.

[15] T. Hirota, J. W. Lee, P. C. St. John, M. Sawa, K. Iwaisako, T. Noguchi, P. Y. Pongsawakul, T. Sonntag, D. K. Welsh, D. A. Brenner, F. J. Doyle III, P. G. Schultz, and S. A. Kay, "Identification of small molecule activators of cryptochrome," *Science*, vol. 337, no. 6098, pp. 1094–1097, 2012.

[16] H. R. Ueda, S. Hayashi, W. Chen, M. Sano, M. Machida, Y. Shigeyoshi, M. Iino, and S. Hashimoto, "System-level identification of transcriptional circuits underlying mammalian circadian clocks," *Nat. Genet.*, vol. 37, no. 2, pp. 187–192, 2005.

[17] H. Ando, H. Yanagihara, Y. Hayashi, Y. Obi, S. Tsuruoka, T. Takamura, S. Kaneko, and A. Fujimura, "Rhythmic messenger ribonucleic acid expression of clock genes and adipocytokines in mouse visceral adipose tissue," *Endocrinology*, vol. 146, no. 12, pp. 5631–5636, 2005.

[18] A. C. Liu, D. K. Welsh, C. H. Ko, H. G. Tran, E. E. Zhang, A. A. Priest, E. D. Buhr, O. Singer, K. Meeker, I. M. Verma, F. J. Doyle III, J. S. Takahashi, and S. A. Kay, "Intercellular coupling confers robustness against mutations in the SCN circadian clock network," *Cell*, vol. 129, no. 3, pp. 605–616, 2007.

[19] K. A. Lamia, U. M. Sachdeva, L. DiTacchio, E. C. Williams, J. G. Alvarez, D. F. Egan, D. S. Vasquez, H. Juguilon, S. Panda, R. J. Shaw, C. B. Thompson, and R. M. Evans, "AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation," *Science*, vol. 326, no. 5951, pp. 437–440, 2009.

[20] A. C. Zambon, E. L. McDearmon, N. Salomonis, K. M. Vranizan, K. L. Johansen, D. Adey, J. S. Takahashi, M. Schambelan, and B. R. Conklin, "Time- and exercise-dependent gene regulation in human skeletal muscle," *Genome Biol.*, vol. 4, no. 10, p. R61, 2003.

[21] N. Barkai and S. Leibler, "Circadian clocks limited by noise," *Nature*, vol. 403, no. 6767, pp. 267–268, 2000.

[22] E. D. Herzog, S. J. Aton, R. Numano, Y. Sakaki, and H. Tei, "Temporal precision in the mammalian circadian system: a reliable clock from less reliable neurons," *J. Biol. Rhythms*, vol. 19, no. 1, pp. 35–46, 2004.

[23] M. Scott, B. Ingalls, and M. Kaern, "Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks," *Chaos*, vol. 16, no. 2, p. 026107, 2006.

[24] S. J. Aton and E. D. Herzog, "Come together, right...now: synchronization of rhythms in a mammalian circadian clock," *Neuron*, vol. 48, no. 4, pp. 531–534, 2005.

[25] R. Silver, J. LeSauter, P. A. Tresco, and M. N. Lehman, "A diffusible coupling signal from the transplanted suprachiasmatic nucleus controlling circadian locomotor rhythms," *Nature*, vol. 382, no. 6594, pp. 810–813, 1996.

[26] J.-C. Leloup and A. Goldbeter, "Toward a detailed computational model for the mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, pp. 7051–7056, 2003.

[27] T.-L. To, M. A. Henson, E. D. Herzog, and F. J. Doyle III, "A molecular model for intercellular synchronization in the mammalian circadian clock," *Biophys. J.*, vol. 92, no. 11, pp. 3792–3803, 2007.

[28] M. J. Rust, J. S. Markson, W. S. Lane, D. S. Fisher, and E. K. O'Shea, "Ordered phosphorylation governs oscillation of a three-protein circadian clock," *Science*, vol. 318, no. 5851, pp. 809–812, 2007.

[29] H. P. Mirsky, A. C. Liu, D. K. Welsh, S. A. Kay, and F. J. Doyle III, "A model of the cell-autonomous mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 27, pp. 11 107–11 112, 2009.

[30] D. B. Forger and C. S. Peskin, "Stochastic simulation of the mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 2, pp. 321–324, 2005.

[31] K. Meeker, R. Harang, A. B. Webb, D. K. Welsh, F. J. Doyle III, G. Bonnet, E. D. Herzog, and L. R. Petzold, "Wavelet measurement suggests cause of period instability in mammalian circadian neurons," *J. Biol. Rhythms*, vol. 26, no. 4, pp. 353–362, 2011.

[32] S.-H. Yoo, S. Yamazaki, P. L. Lowrey, K. Shimomura, C. H. Ko, E. D. Buhr, S. M. Siepka, H.-K. Hong, W. J. Oh, O. J. Yoo, M. Menaker, and J. S. Takahashi, "PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 15, pp. 5339–5346, 2004.

[33] S. M. Siepka and J. S. Takahashi, "Methods to record circadian rhythm wheel running activity in mice," *Methods Enzymol.*, vol. 393, pp. 230–239, 2005.

[34] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2361, 1977.

[35] B. Drawert, S. Engblom, and A. Hellander, "URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries," *BMC Syst. Biol.*, vol. 6, no. 1, p. 76, 2012.

[36] P. C. St. John, S. R. Taylor, J. H. Abel, and F. J. Doyle III, "Amplitude Metrics for Cellular Circadian Bioluminescence Reporters," *Biophys. J.*, vol. 107, no. 11, pp. 2712–2722, 2014.

[37] P. J. van der Spek, K. Kobayashi, D. Bootsma, M. Takao, A. P. Eker, and A. Yasui, "Cloning, tissue expression, and mapping of a human photolyase homolog with similarity to plant blue-light receptors," *Genomics*, vol. 37, no. 2, pp. 177–182, 1996.

[38] G. T. van der Horst, M. Muijtjens, K. Kobayashi, R. Takano, S. Kanno, M. Takao, J. de Wit, A. Verkerk, A. P. Eker, D. van Leenen, R. Buijs, D. Bootsma, J. H. Hoeijmakers, and A. Yasui, "Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms," *Nature*, vol. 398, no. 6728, pp. 627–630, 1999.

[39] J. A. Evans, H. Pan, A. C. Liu, and D. K. Welsh, "Cry1-/- circadian rhythmicity depends on SCN intercellular coupling," *J. Biol. Rhythm.*, vol. 27, no. 6, pp. 443–452, 2012.

[40] E. V. McCarthy, J. E. Baggs, J. M. Geskes, J. B. Hogenesch, and C. B. Green, "Generation of a novel allelic series of cryptochrome mutants via mutagenesis reveals residues involved in protein-protein interaction and CRY2-specific repression," *Mol. Cell. Biol.*, vol. 29, no. 20, pp. 5465–5476, 2009.

[41] M. Ukai-Tadenuma, R. G. Yamada, H. Xu, J. A. Ripperger, A. C. Liu, and H. R. Ueda, "Delay in feedback repression by cryptochrome 1 is required for circadian clock function," *Cell*, vol. 144, no. 2, pp. 268–281, 2011.

[42] S. K. Khan, H. Xu, M. Ukai-Tadenuma, B. Burton, Y. Wang, H. R. Ueda, and A. C. Liu, "Identification of a novel cryptochrome differentiating domain required for feedback repression in circadian clock function," *J. Biol. Chem.*, vol. 287, no. 31, pp. 25 917–25 926, 2012.

[43] C. Lee, J.-P. Etchegaray, F. R. Cagampang, A. S. Loudon, and S. M. Reppert, "Posttranslational Mechanisms Regulate the Mammalian Circadian Clock," *Cell*, vol. 107, no. 7, pp. 855–867, 2001.

[44] B. Ananthasubramaniam, E. D. Herzog, and H. Herzel, "Timing of Neuropeptide Coupling Determines Synchrony and Entrainment in the Mammalian Circadian Clock," *PLoS Comput. Biol.*, vol. 10, no. 4, 2014.

[45] E. E. Zhang, A. C. Liu, T. Hirota, L. J. Miraglia, G. Welch, P. Y. Pongsawakul, X. Liu, A. Atwood, J. W. Huss, J. Janes, A. I. Su, J. B. Hogenesch, and S. A. Kay,

"A genome-wide RNAi screen for modifiers of the circadian clock in human cells," *Cell*, vol. 139, no. 1, pp. 199–210, 2009.

[46] J. Rougemont and F. Naef, "Dynamical signatures of cellular fluctuations and oscillator stability in peripheral circadian clocks," *Mol. Syst. Biol.*, vol. 3, no. 93, p. 93, 2007.

[47] J. Andersson, J. Akesson, and M. Diehl, *Recent Advances in Algorithmic Differentiation*. Springer Berlin Heidelberg, 2012, vol. 87.

[48] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward, "SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers," *ACM Trans. Math. Softw.*, vol. 31, no. 3, pp. 363–396, 2005.

[49] F. Fortin and D. Rainville, "DEAP: Evolutionary algorithms made easy," *J. Mach. Learn. Algorithms*, vol. 13, pp. 2171–2175, 2012.

[50] K. R. Sanft, S. Wu, M. Roh, J. Fu, R. K. Lim, and L. R. Petzold, "StochKit2: software for discrete stochastic simulation of biochemical systems with events," *Bioinformatics*, vol. 27, no. 17, pp. 2457–2458, 2011.

[51] J. H. Abel, L. A. Widmer, P. C. St. John, J. Stelling, and F. J. Doyle III, "A Coupled Stochastic Model Explains Differences in Circadian Behavior of Cry1 and Cry2 Knockouts," *IEEE Life Sci. Lett.*, vol. 1, no. 1, 2015.

[52] Y. Lee, R. Chen, H.-m. Lee, and C. Lee, "Stoichiometric relationship among clock proteins determines robustness of circadian rhythms," *J. Biol. Chem.*, vol. 286, no. 9, pp. 7033–7042, 2011.

[53] A. Morin, L. Denoroy, and M. Jouvet, "Daily variations in concentration of vasoactive intestinal polypeptide immunoreactivity in discrete brain areas of the rat," *Brain Res.*, vol. 538, no. 1, pp. 136–140, 1991.

[54] K. Obrietan, S. Impey, D. Smith, J. Athos, and D. R. Storm, "Circadian Regulation of cAMP Response Element-mediated Gene Expression in the Suprachiasmatic Nuclei," *J. Biol. Chem.*, vol. 274, no. 25, pp. 17 748–17 756, 1999.

[55] P. O. Westermark, D. K. Welsh, H. Okamura, and H. Herzel, "Quantification of circadian rhythms in single cells," *PLoS Comput. Biol.*, vol. 5, no. 11, p. e1000580, 2009.

[56] S. J. Aton, C. S. Colwell, A. J. Harmar, J. Waschek, and E. D. Herzog, "Vasoactive intestinal polypeptide mediates circadian rhythmicity and synchrony in mammalian clock neurons," *Nat. Neurosci.*, vol. 8, no. 4, pp. 476–483, 2005.

[57] C. H. Ko, Y. R. Yamada, D. K. Welsh, E. D. Buhr, A. C. Liu, E. E. Zhang, M. R. Ralph, S. A. Kay, D. B. Forger, and J. S. Takahashi, "Emergence of noise-induced oscillations in the central circadian pacemaker," *PLoS Biol.*, vol. 8, no. 10, p. e1000513, 2010.

[58] C. Vasalou, E. D. Herzog, and M. A. Henson, "Multicellular model for intercellular synchronization in circadian neural networks," *Biophys. J.*, vol. 101, no. 1, pp. 12–20, 2011.

[59] S. An, R. Harang, K. Meeker, D. Granados-Fuentes, C. A. Tsai, C. Mazuski, J. Kim, F. J. Doyle III, L. R. Petzold, and E. D. Herzog, "A neuropeptide speeds circadian entrainment by reducing intercellular synchrony," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 46, pp. E4355–E4361, 2013.

[60] C. Vasalou, E. Herzog, and M. Henson, "Small-world network models of intercellular coupling predict enhanced synchronization in the suprachiasmatic nucleus," *J. Biol. Rhythms*, vol. 24, no. 3, pp. 1–19, 2009.

[61] S. Bernard, D. Gonze, B. Cajavec, H. Herzel, and A. Kramer, "Synchronization-induced rhythmicity of circadian oscillators in the suprachiasmatic nucleus," *PLoS Comput. Biol.*, vol. 3, no. 4, p. e68, 2007.

[62] D. DeWoskin, W. Geng, A. R. Stinchcombe, and D. B. Forger, "It is not the parts, but how they interact that determines the behaviour of circadian rhythms across scales and organisms," *Interface Focus*, vol. 4, no. 3, p. 20130076, 2014.

[63] G. M. Freeman, R. M. Krock, S. J. Aton, P. Thaben, and E. D. Herzog, "GABA networks destabilize genetic oscillations in the circadian pacemaker," *Neuron*, vol. 78, no. 5, pp. 799–806, 2013.

[64] M. Garofalo, T. Nieus, P. Massobrio, and S. Martinoia, "Evaluation of the performance of information theory-based methods and cross-correlation to estimate the functional connectivity in cortical networks," *PLoS One*, vol. 4, no. 8, p. e6482, 2009.

[65] L. Bettencourt, G. Stephens, M. Ham, and G. Gross, "Functional structure of cortical neuronal networks grown in vitro," *Phys. Rev. E*, vol. 75, no. 2, p. 021915, 2007.

[66] M. Kaminski, M. Ding, W. A. Truccolo, and S. L. Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biol. Cybern.*, vol. 85, no. 2, pp. 145–157, 2001.

[67] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, "Detecting novel associations in large data sets," *Science*, vol. 334, no. 6062, pp. 1518–1524, 2011.

[68] Y. Kuramoto, *Chemical oscillations, waves, and turbulence*. Courier Corporation, 2003.

[69] W. J. Schwartz, R. A. Gross, and M. T. Morton, "The suprachiasmatic nuclei contain a tetrodotoxin-resistant circadian pacemaker," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 84, no. 6, pp. 1694–1698, 1987.

[70] S. Yamaguchi, H. Isejima, T. Matsuo, R. Okura, K. Yagita, M. Kobayashi, and H. Okamura, "Synchronization of cellular clocks in the suprachiasmatic nucleus," *Science*, vol. 302, no. 5649, pp. 1408–1412, 2003.

[71] A. B. Webb, N. Angelo, J. E. Huettner, and E. D. Herzog, "Intrinsic, nondeterministic circadian rhythm generation in identified mammalian neurons," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 38, pp. 16 493–16 498, 2009.

[72] J. B. Kinney and G. S. Atwal, "Equitability, mutual information, and the maximal information coefficient," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 9, pp. 3354–3359, 2014.

[73] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[74] M. E. Newman, C. Moore, and D. J. Watts, "Mean-field solution of the small-world network model," *Phys. Rev. Lett.*, vol. 84, no. 14, pp. 3201–3204, 2000.

[75] M. D. Humphries and K. Gurney, "Network 'small-world-ness': a quantitative method for determining canonical network equivalence," *PLoS One*, vol. 3, no. 4, p. e0002051, 2008.

[76] M. Hafner, H. Koeppl, and D. Gonze, "Effect of network architecture on synchronization and entrainment properties of the circadian oscillations in the suprachiasmatic nucleus," *PLoS Comput. Biol.*, vol. 8, no. 3, 2012.

[77] S. Schroder, E. D. Herzog, and I. Z. Kiss, "Transcription-based oscillator model for light-induced splitting as antiphase circadian gene expression in the suprachiasmatic nuclei," *J. Biol. Rhythms*, vol. 27, no. 1, pp. 79–90, 2012.

[78] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, "Stochastic gene expression in a single cell," *Sci. Signal.*, vol. 297, no. 5584, p. 1183, 2002.

[79] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in Escherichia coli," *Nature*, vol. 403, no. 6767, pp. 339–342, 2000.

[80] Y. Cao, H. Li, and L. Petzold, "Efficient formulation of the stochastic simulation algorithm for chemically reacting systems," *J. Chem. Phys.*, vol. 121, no. 9, pp. 4059–4067, 2004.

[81] D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems," *J. Chem. Phys.*, vol. 115, no. 4, pp. 1716–1733, 2001.

[82] Y. Cao and L. Petzold, "Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems," *J. Comput. Phys.*, vol. 212, no. 1, pp. 6–24, 2006.

[83] A. Slepoy, A. P. Thompson, and S. J. Plimpton, "A constant-time kinetic Monte Carlo algorithm for simulation of large biochemical reaction networks," *J. Chem. Phys.*, vol. 128, no. 20, 2008.

[84] C. F. Lopez, J. L. Muhlich, J. A. Bachman, and P. K. Sorger, "Programming biological models in Python using PySB," *Mol. Syst. Biol.*, vol. 9, no. 646, p. 646, 2013.

[85] Dale E Seborg, D. A. Mellichamp, T. F. Edgar, and F. J. Doyle III, *Process Dynamics and Control*. Wiley, 2010.

[86] J. M. Bower and H. Bolouri, *Computational Modeling of Genetic and Biochemical Networks*. MIT Press, 2004.

[87] K. A. Fujita, Y. Toyoshima, S. Uda, Y.-i. Ozaki, H. Kubota, and S. Kuroda, "Decoupling of receptor and downstream signals in the Akt pathway by its low-pass filter characteristics," *Sci. Signal.*, vol. 3, no. 132, p. ra56, 2010.

[88] P. Hersen, M. N. McClean, L. Mahadevan, and S. Ramanathan, "Signal processing by the HOG MAP kinase pathway," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 20, pp. 7165–7170, 2008.

[89] D. Gonze, S. Bernard, and C. Waltermann, "Spontaneous synchronization of coupled circadian oscillators," *Biophys. J.*, vol. 89, no. 1, pp. 120–129, 2005.

[90] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C (2nd ed.): the art of scientific computing*, 1992, vol. 29, no. 4.

[91] A. N. van den Pol and F. E. Dudek, "Cellular communication in the circadian clock, the suprachiasmatic nucleus," *Neuroscience*, vol. 56, no. 4, pp. 793–811, 1993.

Appendix

S1 Example for Constructing and Simulating a GillesPy Model

```
import scipy as sp
import numpy as np
import matplotlib.pyplot as plt

import sys
sys.path.append('../')

import gillespy

class parameter_changing_model(gillespy.Model):
    """
    This toy example shows how we can simply simulate the same model for
    multiple parameter sets. Our model consists of the following reactions:

          0 -> S1
    S1 + S1 -> S2
         S2 -> 0

    So S1 being produced, dimerizing, and being degraded.
    """

    def __init__(self, parameter_values=None):

        # Initialize the model.
        gillespy.Model.__init__(self, name="simple1")

        # Parameters
        k1 = gillespy.Parameter(name='k1', expression=parameter_values[0])
        k2 = gillespy.Parameter(name='k2', expression=parameter_values[1])
        k3 = gillespy.Parameter(name='k3', expression=parameter_values[2])
        self.add_parameter([k1, k2, k3])

        # Species
        S1 = gillespy.Species(name='S1', initial_value=100)
        S2 = gillespy.Species(name='S2', initial_value=0)
        self.add_species([S1, S2])

        # Reactions
        rxn1 = gillespy.Reaction(
                name = 'S1 production',
                reactants = {},
                products = {S1:1},
                rate = k1 )

        rxn2 = gillespy.Reaction(
```

```
                name = 'dimer formation',
                reactants = {S1:2},
                products = {S2:1},
                rate = k2)

        rxn3 = gillespy.Reaction(
                name = 'dimer degradation',
                reactants = {S2:1},
                products = {},
                rate = k3)

        self.add_reaction([rxn1, rxn2, rxn3])


if __name__ == '__main__':

    # Here, we create the model objects. We have two different parameter
    # sets:

    set1 = [100, 0.1, 0.1]
    set2 = [100, 0.001, 0.1]

    # For set #2, dimers (S2) form much less readily.

    set1_model = parameter_changing_model(parameter_values = set1)
    set2_model = parameter_changing_model(parameter_values = set2)

    num_trajectories = 100

    # Let's simulate for both parameter sets, and compare the results
    set1_trajectories = gillespy.StochKitSolver.run(set1_model,
            number_of_trajectories = num_trajectories)
    set2_trajectories = gillespy.StochKitSolver.run(set2_model,
            number_of_trajectories = num_trajectories)

    # PLOTTING RESULTS
    # here, we will plot all trajectories with the mean overlaid
    from matplotlib import gridspec

    gs = gridspec.GridSpec(1,2)
    alp = 0.1 # alpha value

    # extract time values
    time = np.array(set1_trajectories[0][:,0])

    # Plot for parameter set #1
    ax0 = plt.subplot(gs[0,0])

    set1_S1 = np.array([set1_trajectories[i][:,1] for i in xrange(num_trajectories)]).T
    set1_S2 = np.array([set2_trajectories[i][:,2] for i in xrange(num_trajectories)]).T


    #plot individual trajectories
    ax0.plot(time, set1_S1, 'r', alpha = alp)
    ax0.plot(time, set1_S2, 'b', alpha = alp)

    #plot mean
    ax0.plot(time, set1_S1.mean(1), 'k--', label = "Mean S1")
    ax0.plot(time, set1_S2.mean(1), 'k:', label = "Mean S2")

    ax0.legend()
    ax0.set_xlabel('Time')
    ax0.set_ylabel('Species Count')
    ax0.set_title('Parameter Set 1')
```

```
# Plot for parameter set #2
ax1 = plt.subplot(gs[0,1])

set2_S1 = np.array([set2_trajectories[i][:,1] for i in xrange(num_trajectories)]).T
set2_S2 = np.array([set2_trajectories[i][:,2] for i in xrange(num_trajectories)]).T


#plot individual trajectories
ax1.plot(time, set2_S1, 'r', alpha = alp)
ax1.plot(time, set2_S2, 'b', alpha = alp)

#plot mean
ax1.plot(time, set2_S1.mean(1), 'k--', label = "Mean S1")
ax1.plot(time, set2_S2.mean(1), 'k:', label = "Mean S2")

ax1.legend()
ax1.set_xlabel('Time')
ax1.set_title('Parameter Set 2')



plt.tight_layout()
plt.show()
```

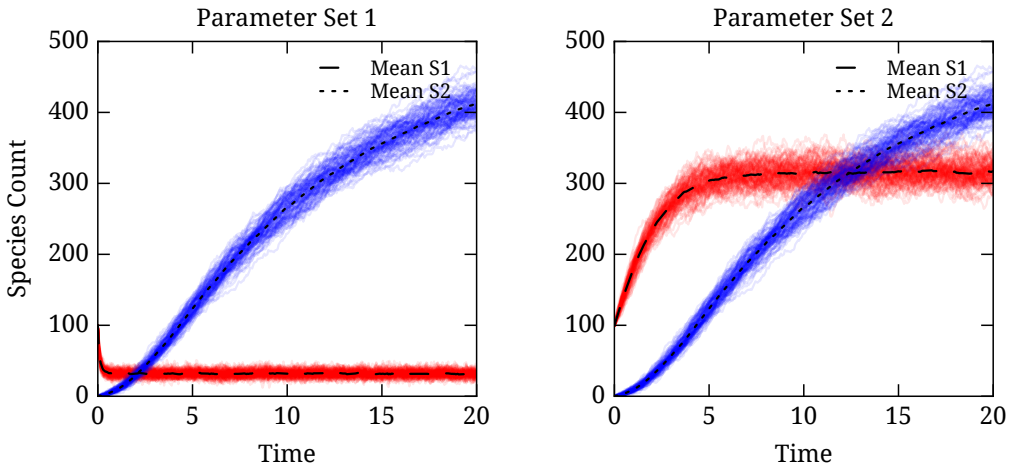

Figure S1: Gillespy simulation example. Direct output from the above code, showing simulation of two ensembles of 100 trajectories, each with differing parameterization.