UNIVERSITY OF CALIFORNIA, SANTA BARBARA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

Rudiyanto Gunawan, Ph.D.
Postdoctoral Research Fellow
Process Control Laboratory

Department of Chemical Engineering
Santa Barbara, CA 93106-5080
(805) 893-3120
(805) 893-4731 fax
gunawan@engineering.ucsb.edu

November 14, 2005

Professor Yves Brun
Department of Biology
Indiana University
Jordan Hall 142
1001 East Third Street
Bloomington, IN 47405-7005

Dear Professor Brun:

I am writing in regard to the advertised tenure-track faculty position in systems biology at Indiana University. I am a postdoctoral fellow in the Department of Chemical Engineering at the University of California Santa Barbara, working with Professor Francis J. Doyle III. My postdoctoral research belongs to the field of systems biology, with a particular interest in the elucidation of molecular bases for robustness properties in cells through model identification, systems analysis, and design of cellular networks.

I am interested in establishing a world-class research program in the area of systems biology that concentrates on the discovery of design principles underlying cellular behavior, such as robustness, and their physiological manifestations in diseases and infections. The success of the research will contribute to the understanding of disease development and the identification of possible drug targets for diseases and infections. Specifically, the research will focus on interacting cellular systems within a single cell (such as circadian rhythm, cell cycle, and metabolism) and in a cell population (for example, quorum sensing in bacterial populations). I will initiate and maintain close collaborations with experimental experts in each respective research area to ensure the significance of the research from both experimental and theoretical perspectives.

Please find enclosed my vitae, list of references, research summary, research and teaching interests, and several research articles representing my work. Thank you for your consideration. If you wish for more details or further information, please contact me.

Sincerely,

Rudiyanto Gunawan

RESEARCH SUMMARY

Rudiyanto Gunawan
Department of Chemical Engineering
University of California Santa Barbara
Santa Barbara, CA

The success of the human genome project has ushered in a new era that emphasizes a systemic or integrated approach to ascertain the cellular behavior arising from complex cellular networks. Scientists are now embarking on a quest to elucidate the organization and control of cellular networks that underlie the phenotypic behavior of a cell; these are the so-called "omics" such as genomics, transcriptomics, proteomics, metabolomics. Fueled by recent advances in molecular biology providing high-throughput and in-depth data of gene and protein interactions, it is increasingly clear that cell behaviors arise from complex interactions among the genes and proteins through crossover and cascade regulations and signal transductions, and thus can only be explained through a system-level understanding of these interactions. This is the goal of systems biology, which involves application of systems theoretic approaches and integration of experimental and computational research [1].

Control systems approaches have been instrumental in systems biology. Several concepts from control engineering, in particular robustness, have been used to define many characteristics of cellular behavior. Robustness refers to the ability of a cell to maintain its functions (phenotype) under intrinsic and extraneous uncertainties [2]. In biological systems, uncertainties can arise from the inherent stochastic nature of gene expression (intrinsic) [3] or from variations in the nutrients and signals concentration (extraneous). There exists a consensus in the literature that the complexity in the cellular network arises from the regulation and control required to achieve such robust behaviors [4]. Also, one salient feature of high robustness is the existence of fragility points in the cellular network to which a small perturbation can lead to catastrophic consequences [5]. The understanding of robustness and fragility trade-offs in biology can help elucidate disease development in healthy cells and identify possible drug targets in diseased cells [6, 7]. My research focus is to ascertain the underlying design principles of robustness in biology through model development and analysis.

1 Reverse-engineering of Cellular Networks

A system-level understanding of the functioning behavior of a cell requires an accurate representation of the underlying complex networks of gene and protein interactions. Advances in molecular biology have provided a glimpse of such complexity through diverse measurements of cellular activities. In systems biology, the goal of network inference or reverse engineering is to reconstruct the complex network of regulatory interactions from available measurements. Here, the reverse engineering effort faces two daunting problems: network size and complexity, and incomplete and inaccurate measurements. In addition,

complete knowledge of a cellular network entails the identification of not only the network architecture (topology) but also its dynamics. Indeed, implicit in the term "regulation" is the importance of dynamics of these interactions. Network inference from experiments has been extensively investigated in the field of engineering, which is known as system identification. In addition, many concepts in engineering, such as modularity, robustness, and optimality, have been observed in many biological systems. For these reasons, engineering approaches have been instrumental in the network inference of biological systems.

In practice, the reverse-engineering of a gene network should involve a careful design of experiment, using prior knowledge of the system to obtain the most informative measurements. Thus, this process should be iterative, in which the result from each trial is used to better design the next experiment. Figure 1 shows one realization of this iterative process, which includes four key steps: experiment, parameter estimation, model validation and optimal experiment design [8, 9]. Here, the parameter estimation from partial measurements is decomposed into two parts: state estimator and parameter identification. The state estimator involves an extension of the dynamic flux balance analysis [10] to obtain the fluxes (reaction rates) and concentrations that minimize the rate of production and accumulation of intermediates, called the State Regulator Problem (SRP) [11]. The availability of the full concentration and reaction rate estimates decouples the identification of kinetic parameters with respect to each reaction, which significantly reduces the complexity of the parameter estimation problem. In addition, the parameter estimation follows a Bayesian formulation for ease of additional data incorporation [12]. Model validity/invalidity is judged using several different criteria such as model prediction error, confidence (uncertainty) region of parameters, as well as existing biological information of the system such as robustness to certain external/internal disturbances. If further experiments are necessary to improve the model, model-based experiment design gives the optimal experimental conditions that maximize the amount of information for the model identification. Here, a measure of informativeness of data, such as the Fisher information matrix, can help formulate the optimal experiment design into an optimization problem. Application to the model development of caspase-activated apoptosis highlighted the importance of experimental design step in the reverse engineering of cellular systems [8].

2 System Analysis of Cellular Networks

The size and complexity of cellular networks prevent the deduction of robustness and fragility based solely on intuition. Systems analysis can help unravel this complexity. One such method is sensitivity analysis, in which linear sensitivities quantify how much the system behavior changes as the parameters are varied [13]. In cellular networks, high sensitivities point to the fragile links in the network on which cellular behavior strongly depends. Sensitivity analysis traditionally applies to continuous systems, such as differential equations, as these models are the most common representation of engineered systems. However, the characteristics of cellular processes, such as reactions involving low concen-

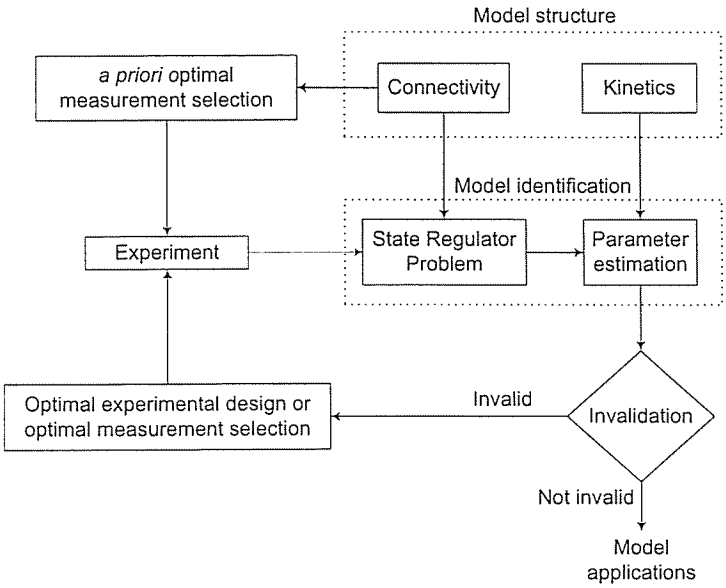

Figure 1: Iterative reverse-engineering of biological networks.

tration of molecules (nanomolar), limit the application of classical sensitivity analysis and thus require the development of new methodologies for analysis.

One focus in my research is the development of non-traditional sensitivity analysis to investigate common systems in biology, in particular discrete stochastic systems and oscillatory behavior. In [14], sensitivity analysis for discrete stochastic systems was developed in which the system dynamics are represented by the chemical master equation [15]. Here, the chemical reactions occur as discrete events due to the low copy number of species involved. Such behavior accurately describes many cellular processes such as gene expression [3]. Application of traditional sensitivity analysis to continuum representations of these systems can give incorrect results, in particular for systems with multiple steady states such as a common gene switch [14]. The results also indicate that the design of gene network in synthetic biology needs to address the stochastic behavior explicitly to capture *in vivo* dynamics.

Another common behavior of biology is oscillatory dynamics, for example circadian rhythm and cell cycle. In circadian rhythm, past sensitivity analyses have mainly focused on the period and amplitude of the oscillations [16, 17, 18]. Unfortunately, there exists very little work on the sensitivity analysis of the phase response, which underlies the synchronization of circadian rhythm to environmental cues such as light. Here, I developed the phase response analysis for oscillatory systems based on the concept of isochrons. An isochron of a limit cycle contains all points that evolve to the same phase (phase level sets). Different measures of phase sensitivity analysis can be derived from this approach including

the phase response curve, a common measure of transient phase response in chronobiology [19]. The application to *Drosophila melanogaster* (fruit fly) circadian model generated classifications in the gene regulation with respect to the period and phase modulations. The mRNA transcriptions were found to preferentially regulate phase response, while the nuclear translocations mostly affected the period modulation. Also, photic entrainment was found to modulate both period and phase responses, in agreement with experimental evidence in the literature [20].

References

[1] H. Kitano. Systems biology: a brief overview. *Science*, 295:1662–1664, 2002.

[2] J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle III, and J. Doyle. Robustness of cellular functions. *Cell*, 118(6):675–685, Sep 2004.

[3] H. H. McAdams and A. Arkin. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet.*, 15(2):65–69, Feb 1999.

[4] D. A. Lauffenburger. Cell signaling pathways as control modules: complexity for simplicity? *Proc. Natl. Acad. Sci. USA*, 97(10):5031–5033, May 2000.

[5] M. Csete and J. Doyle. Reverse engineering of biological complexity. *Science*, 295:1664–1669, 2002.

[6] M. Csete and J. Doyle. Bow ties, metabolism and disease. *Trends Biotechnol.*, 22(9):446–450, Sep 2004.

[7] H. Kitano. Cancer robustness: tumour tactics. *Nature*, 426(6963):125, Nov 2003.

[8] K. G. Gadkar, R. Gunawan, and F. J. Doyle III. Iterative approach to model identification of biological networks. *BMC Bioinformatics*, 6:155–174, Jun 2005.

[9] R. Gunawan, K. G. Gadkar, and F. J. Doyle III. Methods to identify cellular architecture and dynamics from experimental data. In Z. Szallasi, V. Periwal, and J. Stelling, editors, *System Modeling in Cellular Biology*. MIT Press, 2005.

[10] R. Mahadevan, J. S. Edwards, and F. J. Doyle III. Dynamic flux balance analysis of diauxic growth in *E. coli*. *Biophys. J.*, 83:1331–1340, 2002.

[11] K. G. Gadkar, J. Varner, and F. J. Doyle III. Model identification of signal transduction networks from data using a state regulator problem. *IEE Sys. Bio.*, 2:17–30, 2005.

[12] R. Gunawan, M. Y. L. Jung, E. G. Seebauer, and R. D. Braatz. Maximum a posteriori estimation of transient enhanced diffusion energetics. *AIChE J.*, 49:2114–2123, 2003.

[13] A. Varma, M. Morbidelli, and H. Wu. *Parametric Sensitivity in Chemical Systems.* Oxford University Press, New York, NY, 1999.

[14] R. Gunawan, Y. Cao, L. Petzold, and F. J. Doyle III. Sensitivity analysis of discrete stochastic systems. *Biophys. J.*, 88:2530–2540, 2005.

[15] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.

[16] J. Stelling, E. D. Gilles, and F. J. Doyle. Robustness properties of circadian clock architectures. *Proc. Natl. Acad. Sci. USA*, 101(36):13210–13215, Sep 2004.

[17] B. Ingalls. Autonomously oscillating biochemical systems: parametric sensitivity of extrema and period. *IEE Sys. Bio.*, 1:62–70, 2004.

[18] D. E. Zak, J. Stelling, and F. J. Doyle III. Sensitivity analysis of oscillatory (bio)chemical systems. *Comp. Chem. Eng.*, 29:663–673, 2005.

[19] C. H. Johnson. Forty years of PRCs–what have we learned? *Chronobiol. Int.*, 16(6):711–743, Nov 1999.

[20] R. Gunawan and F. J. Doyle III. Isochron-based phase response analysis of circadian rhythms. *Biophys. J.*, 2005. submitted.

RESEARCH PLAN

Rudiyanto Gunawan
Department of Chemical Engineering
University of California Santa Barbara
Santa Barbara, CA

1 Scope and Significance of Research

The success of genome projects has ushered in a new era that emphasizes a systemic or integrated approach, known as *systems biology*, to explain the overt cellular behavior based on genes and proteins, and their complex interconnections [1]. In systems biology, the functionality of a gene or protein is inferred from its (dynamical) interactions with other cellular components. The complexity of cellular networks is akin to many man-made systems, for example airplanes, internet networks, and integrated circuits, which result from one part design and one part evolution in technology (such as combining old modules with new parts). Cellular networks evolve in a similar manner through gene mutations and transfers under natural selection. Other similarities between cellular and well-engineered systems are their modularity and hierarchy that allow high evolvability through inter- and intra-operability among components in the networks, *e.g.*, gene transfers within genomes and between organisms [2, 3]. Therefore, applications of methodologies from fields such as engineering, computer science, and statistics, to problems in systems biology have been and will be instrumental.

Systems biology lies at the interface of biology and systems theory including control systems engineering. Control systems approaches have been instrumental in this relatively new discipline, for example in the elucidation of chemotaxis adaptive mechanism [4], in the identification of control motifs in regulatory networks [5], and in the unraveling of design principles in circadian rhythm architecture [6]. Other than the familiar use of the word "systems" as a designation for physical systems, the term here also refers to the study of physical systems through modeling, formulation of mathematical descriptions, analysis, and design [7]. Several concepts from control engineering, such as *robustness*, have diffused into systems biology to define many characteristics of cellular behavior [2, 8]. Robustness describes the ability of a system to maintain the desired performance/behavior under intrinsic and extraneous uncertainties [9]. Conversely, fragility represents the high sensitivity of system behavior to a particular perturbation. In biological systems, the system uncertainties can arise from the inherent stochastic nature of gene expression (intrinsic) [10] or variations in the extracellular nutrient or signal concentration (extraneous). In fact, there appears to be an intimate link between complexity and robustness in cellular functions [11]. Tools from systems engineering, such as sensitivity, stability, and robustness analysis, have been applied to quantify the robustness property of biological systems [6, 12]. The understanding of the robustness and fragility properties in cellular behaviors have multiple applications, such as

the elucidation of the evolutionary design principles of cellular networks, the understanding of disease development, and the identification of possible drug targets [2, 13].

This proposal outlines the research topics in the area of systems biology that I will undertake, which encompass different aspects of systems approaches in biology including system identification, analysis, and control. The research focuses on the implications of stochastic noise and system interactions on cellular behavior, such as robustness and fragility, using mathematical models of cellular processes and control system theory (sensitivity analysis, robustness analysis, multiscale system analysis). The overarching goal is to elucidate the causality in diseases and infections for better drug design and treatment regimen with specific targets in the diseased cells and minimal complications to healthy cells. New methodologies in systems theory will be developed to address problems that are common in biology, such as stochastic and oscillatory behavior. The research applications include modeling and treatment of infections caused by quorum-sensing bacterial populations, design of robust genetic switches, modeling and analysis of coupled oscillatory biological systems (circadian rhythm, cell cycle, and metabolism), and model identification of stochastic and oscillatory biological systems.

2 Modeling and Control of Bacterial Populations

In their natural environment, cells do not live independently, but rather they interact with each other. In a bacterial colony, the survival of a single cell should become secondary to that of the whole population. For example, cell-to-cell signalling, *quorum sensing*, is used to maintain critical population density to overwhelm host defenses and to synchronize specific gene expression for the population to survive in a certain environmental niche, such as biofilm formation and bacterial virulence [14, 15, 16]. In gram-negative bacteria, quorum sensing (QS) is achieved by a small diffusible signal molecule N-acylhomoserine lactones (AHLs) and a transcription factor with AHL-dependent activity. The same QS architecture was used in creating *E. coli* with a built-in population control [17]. Figure 1 illustrates a typical quorum sensing mechanism in gram-negative bacteria. Since the first discovery of QS in *V. fisheri* bioluminescence regulation [18], many more gram-negative bacteria were identified to employ QS, of which several can cause infections in humans and animals, including *P. aeruginosa*, *Yersinia* spp., and *Aeromonas* spp. [15]. In particular, *P. aeruginosa* commonly causes acute infections in immunocompromised patients and respiratory tract infections in cystic fibrosis patients [19]. Due to its importance, the *P. aeruginosa* QS system has been completely characterized [20] and single-cell dynamical models of its virulence have been created [21, 22, 23].

The goals of the research are to establish an accurate dynamical representation of QS population, to understand the implications of QS on the sensitivity and robustness properties of a single cell and the colony, and finally to identify the treatment options and procedures for infection caused by QS bacteria. The focus model in this research is *P. aeruginosa* virulence. The dynamics of the bacterial population will be represented

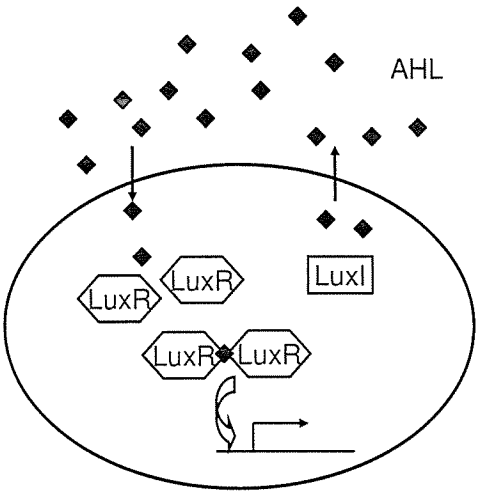

Figure 1: A representative quorum sensing mechanism in gram-negative bacteria. LuxI protein promotes the production of AHL signal molecules which are subsequently released out of the cell. As the bacterial colony proliferates, the external AHL accumulates and eventually activates LuxR-type transcription factors, which control many target genes such as those for biofilm formation and virulence.

using a population balance equation [24], modeling the spatial dependence and the different characteristics of the cells (level of transcription factors, protein, AHLs). Simulations of such model will require the coupling of numerical algorithms for a hyperbolic integro-partial differential equation [24, 25, 26] and computational fluid dynamics (such as finite volume methods [27]). The modeling approach will also consider the inherent stochastic behavior of cellular processes using a chemical master equation (CME) [28].

2.1 *Robustness Analysis of Biological Systems*

The size and complexity of cellular networks make the underlying mechanism for cellular behavior difficult to ascertain. Systems analysis can help to unravel this complexity. One such method is sensitivity analysis [29], in which linear sensitivities quantify how much the system behavior changes as the parameters are varied. Apart from its traditional use in analysis of differential equations, the sensitivity analysis has also been developed for biologically relevant systems such as stochastic [30] and oscillatory systems [31, 32, 33]. In cellular networks, high sensitivities point to the interconnections in the system on which cellular behavior strongly depends. Perturbations on these interconnections can potentially lead to a large disruption in the network behavior, *i.e.*, the network is fragile to the uncertainty in these pathways.

Here, the system analysis will focus on the robustness-fragility consequences of quorum sensing. A novel approach will be developed to characterize the global robustness

property based on the distance to bifurcation. The bifurcation distance characterizes the smallest perturbation in parameter set that leads to a change in the behavior (for example, from stable steady state to oscillations or chaos). Such distance is inversely correlated with the robustness of the system behavior. A related approach using the stability margin of a linearized system has previously been utilized as a robustness measure [12], which still represents a local analysis due to the linearization. The proposed robustness measure will be of general applicability for use in other biological systems. The system analysis of the QS population model will allow the investigation of robust and fragile points in the biofilm formation and bacterial virulence behavior, which may give potential targets for drug development.

2.2 *Treatment of Quorum-Sensing Bacteria Infection*

Finally, using the QS population model, the design of a bacterial infection treatment can be posed as a control problem to maximize drug efficiency while minimizing the side effects and/or the potential development of drug resistance. The quorum sensing molecule has become a new target in treating infection by these bacteria [34]. Different drugs for such infection, including AHL antagonists, common antibiotics, and possible new drug targets based on the above system analysis, will be modeled as external input or perturbation to the population. The control problem will then optimize the treatment regimen and determine the maximum efficacy of each drug, under the consideration of host response, drug toxicity, and possible emergence of resistance (especially, in antibiotic treatment).

3 Design of Genetic Switches

The understanding of design principles in biology can benefit from research efforts in the creation of small functional gene networks. Advances in recombinant DNA technology have allowed scientists to construct synthetic gene networks with specific functions such as a repressilator [35] and a gene switch [36]. These techniques set the foundation for building plug-and-play gene modules with predictable performance, which will make up a list of standardized parts [37]. From these parts, one can construct a functional module that can perform a specific task. Finally, the design effort in synthetic biology will become decoupled from the fabrication, analogous to the manufacture of an integrated chip. Existing methodologies for designing the gene modules take different approaches, such as combinatorial synthesis [38], design-then-mutate [39], and *in-silico* evolution [40]. However, none of these approaches considers the stochastic nature of cellular processes explicitly. The inherent stochastic noise can induce distinguishing behaviors that are not observable in continuum models [30, 41]. Thus, these plug-and-play gene modules should be designed to robustly perform under the stochastic noise *in-vivo* [42].

Gene switches represent a common building block in biological functions, for example in lactose utilization pathways in *E. coli* [43], λ-phage lysis-lysogeny decision [44], and mitogen-activated protein (MAP) kinase cascades [45]. Although the network structures

that give rise to the switch-like behavior are multiple, one recurring characteristic in most of these systems is multistability. These dynamics occurs when the system possesses more than one possible stable steady state, depending on the initial conditions. The most telling behavior of multistability is the hysteretic response around the switching point. From an engineering design perspective, such response is undesirable and should be minimized in the development of gene switches for synthetic biology. Although the proposed research will focus on a bistable switch, the approach will be generalizable to the design of other gene modules.

The quality of a gene switch is defined by its (i) fidelity and (ii) robustness. Several switch fidelity measures, such as sensitivity at the switch point (analogous to the Hill coefficient) and size of the hysteresis region, can be defined. In a system with large stochastic noise and/or two adjacent steady states, the cell dynamics can produce a flip-flop behavior between "ON" and "OFF" [41]. In a novel application, the switch design will optimize the dynamics of a gene switch to produce specific flip-flop frequencies. By combining this novel design with a downstream signal transduction acting as a bandwith filter, the resulting network can produce signal amplification only for very specific inputs that induce the switch to flip-flop at a frequency within the bandwith. Classical control approaches, in particular filter design, will contribute in the creation/selection of downstream signal transduction networks.

Bifurcation analysis has been used to design a bistable gene switch based on a deterministic model [46]. The bifurcation diagram maps the dynamical behavior of the system as a function of the parameters [47]. The engineering of the switch then reduces to finding the parameter combination that gives the desired switch response [43]. This approach however does not consider the stochastic effects in gene expressions that affect the network. The proposed methodology is to formulate the switch design in the framework of nonlinear programming. The engineering of the gene switch will become a constrained optimization problem. The stochastic effects will be directly coupled with the design using the stochastic simulation algorithm [28]. The objective function will also incorporate a measure of robustness to the stochastic effects to ensure a good switch performance under noise. The optimization framework presents a formal approach in engineering gene modules that also allows explicit consideration of cellular dynamics, in particular stochastic behavior, to ensure robust performance and cell viability. This formal approach will expedite the engineering of different gene modules to achieve cell-by-design goal in synthetic biology.

4 Oscillatory Systems in Biology

Oscillatory behavior represents one common dynamics in biological systems and controls many key cellular functions, such as cell cycle and circadian rhythm. Cell cycle regulates a sequence of checkpoints by which a cell grows and divides into two daughter cells. On the other hand, circadian rhythm regulates daily activities of most organisms on earth by mimicking the 24-hour earth's periodicity. Another system, though lesser known to oscil-

late, is the cellular metabolism, which produces reductive-oxidative cycles. Because of their importance, each of these cycles has been the subject of intensive experimental and computational research, but only as individual systems. However, recent evidence gives support to the existence of interactions among the systems [48, 49, 50]. For example, there exist clinical evidence supporting the relationship between circadian rhythm and the efficacy and toxicity of anti-tumor drugs (see [48] and references therein). The complete understanding of organism-level regulation by circadian rhythm, metabolic and mitotic cycles necessitates an integrative study of these systems as a whole [51]. Disturbances to the regulation of any of these cycles have many important physiological impacts from cancer to sleep disorders.

Advances in molecular biology have made detailed mechanistic models of these systems possible. The cell division cycle involves the regulation of several key proteins, most notably the cyclins and cyclin-dependent kinases (CDKs) [52, 53], with checkpoints between the cell cycle phases (G1-S-G2-M). Models exist that describe the regulation of cell cycle in eukaryotes (see for example [54, 55, 56]). On the other hand, the key genes of circadian rhythm in numerous organisms, from neurospora to *Drosophila* to mammals, have been identified [57]. For example, the mammalian circadian genes include *Per*, *Cry*, *Clock* and *Bmal1* [58]. There currently exist three models of varying detail describing the mammalian circadian rhythm based on the knowledge of its molecular biology [59, 60]. Finally, the cellular metabolism is the most studied system among the three for which public databases of models exist (for example, KEGG Pathway database at http://www.genome.jp/kegg/pathway.html, MetaCyc database at http://metacyc.org/).

The modeling effort in this research concerns the integration of existing models for each system based on evidence of interconnections in the literature. Several integrated models will be developed based on varying details of the individual model representation and their interconnections. The integrated model will be validated against experiments in the literature [49, 50]. The complete model will allow the investigation of co-regulations among the systems and responses to external inputs, as illustrated in Figure 2. The analysis of each individual system and the integrated model using the approach developed in Section 2.1, will show how system interconnections affect the robustness properties. In addition, the propagation of disturbances in a single system (such as those caused by diseases) to the others can be quantified using numerical simulations, validated to known physiological manifestations (*e.g.*, cancer), and utilized in hypothesis generation.

5 Design of Optimal Experiments for Stochastic and Oscillatory Systems

One key aspect of systems biology is the identification of biological networks from experimental data, which is known as reverse-engineering or network inference problem. The goals of reverse engineering in systems biology are multiple: (i) understanding of cellular function, (ii) hypothesis generation, and (iii) design of experiment. Here, the challenges stem from the high complexity of cellular networks and the lack of "quality" experimental data (see [61] and references therein). The high complexity arises mainly from the large number

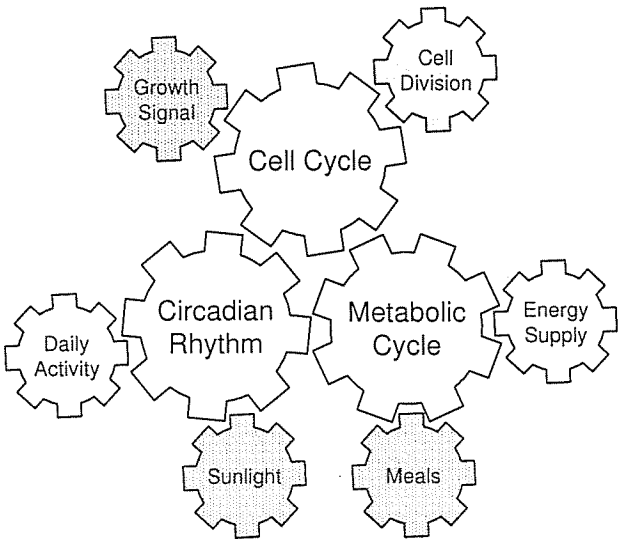

Figure 2: Integration of oscillatory systems: circadian rhythm, metabolic and mitotic cycles. Each cycle controls important functions in organisms, such as cell division, daily activity cycle (sleep-wake) and energy supply. The cycles also take external inputs, for example growth factors, light and meals.

of variables and their interconnections in a typical biological network (curse of dimensionality). The data quality issue relates to the identifiability problem in control theory [62], which concerns the effective amount of information in experimental data for the network identification. These two issues are coupled. The complexity of a network determines the appropriate model structures that are suitable for numerical simulations [63, 64, 65, 66]. The model structure then determines the type of experimental data necessary for model identification [63, 67]. Finally, the information contained in experimental data is limited by measurement error (noise) and conditions (*e.g.*, steady state measurements are of limited use in the identification of kinetic parameters).

An efficient reverse-engineering should involve an iterative process in which the result from each trial is used to iteratively improve the model identification [1, 61]. Figure 3 presents a diagram of such an iterative process. A similar procedure combining the dynamic flux balance analysis [68] and a Bayesian parameter estimation [69] have been developed and applied to the model identification of caspase-activated apoptosis [70]. The results of this study, as well as other identifiability analysis, suggest that the experiment design strongly determine the identifiability of a biological network [71].

The Fisher information matrix (FIM), arising from the field of information theory, can serve as a measure of information content in given noisy experimental data for the identification of system parameters [70, 72]. Experiments can then be designed to optimize an objective function based on the Fisher information matrix, such as the determinant of

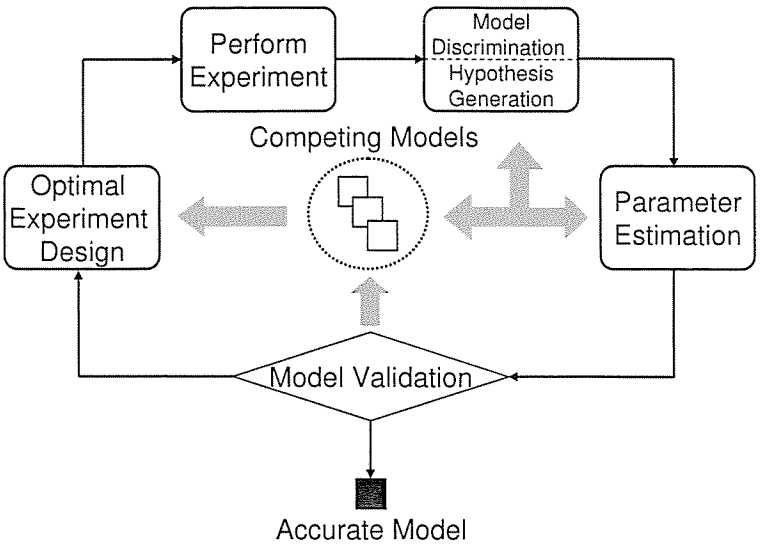

Figure 3: Iterative reverse-engineering of biological networks. Existing models of the network are used to design the next experiment to have the most informative data for parameter estimation and/or model discrimination. Model validation step uses independent experimental data to (in)validate the model predictions. The iterative procedure stops when a model of a desired predictive passes the model validation.

this matrix (as a representation of the volume of information). The optimization variables typically consist of parameterizations of the experimental protocol, but can also include the selection of key variables to be measured and the measurement frequency. Finally, the limitations in the experimental conditions and apparatus are represented as constraints in the optimization. The design of optimal experiment can be formulated as a constrained (mixed-integer) nonlinear programming for which efficient numerical optimizations exist (such as BARON [73] and MINOPT [74]).

The research focus in this area will be in the model-based optimal experiment design for biologically relevant systems, in particular stochastic and oscillatory systems. The optimization will utilize available knowledge of the system, represented by a model, to maximize the data informativeness in the subsequent experiment of an iterative model identification. These systems commonly arise in the modeling of biological systems, for example phage λ-infection in *E. coli* [44], circadian rhythms [75, 76], and cell cycle [77]. Biological relevance of stochastic system comes from the nature of gene transcription and translation which involve reactions with low copy number of molecules. At such low concentration, the processes occur as discrete events which continuum modeling can not fully represent [30, 41, 44]. Results from sensitivity analysis [30] give support to the heed for an explicit accounting of stochastic effects in the design of optimal experiments. In addition, the inclusion of stochastic effects will allow the use of current distributive measurements of cells, such as flow cytometry, to their full capability.

Similarly, oscillatory systems are commonly encountered in biology, as shown in the previous section. The dynamic attributes of these systems, such as period and phase, necessitate the development of specialized analysis [31, 32, 33]. Further, in circadian rhythm, the most common experimental data are in the form of actographs or actograms [78], which only contain relative phase information. Therefore, the design of experiments for oscillatory systems needs to adopt a different methodology that is suited for the dynamics as well as the available measurements. The proposed design of experiments will give an efficient and accurate model identification of stochastic and oscillatory systems in biology, which also matches the experimental procedures particular in each system.

References

[1] H. Kitano. Systems biology: a brief overview. *Science*, 295:1662–1664, 2002.

[2] M. Csete and J. Doyle. Bow ties, metabolism and disease. *Trends Biotechnol.*, 22:446–450, 2004.

[3] H. H. McAdams, B. Srinivasan, and A. P. Arkin. The evolution of genetic regulatory systems in bacteria. *Nature Rev.*, 5:169–178, 2004.

[4] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis. *Proc. Natl. Acad. Sci. USA*, 97:4649–4653, 2000.

[5] C. Rao and A. P. Arkin. Control motifs for intracellular regulatory networks. *Annu. Rev. Biomed. Eng.*, 3:391–419, 2001.

[6] J. Stelling, E. D. Gilles, and F. J. Doyle. Robustness properties of circadian clock architectures. *Proc. Natl. Acad. Sci. USA*, 101:13210–13215, 2004.

[7] C.-T. Chen. *Linear System Theory and Design.* Oxford University Press, New York, NY, 1999.

[8] H. Kitano. Biological robustness. *Nature Rev. Genet.*, 5:826–837, 2004.

[9] J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle III, and J. Doyle. Robustness of cellular functions. *Cell*, 118:675–685, 2004.

[10] H. H. McAdams and A. Arkin. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet.*, 15:65–69, 1999.

[11] D. A. Lauffenburger. Cell signaling pathways as control modules: complexity for simplicity? *Proc. Natl. Acad. Sci. USA*, 97:5031–5033, 2000.

[12] H. Schmidt and E. W. Jacobsen. Linear systems approach to analysis of complex dynamic behaviours in biochemical networks. *IEE Sys. Bio.*, 1:149–158, 2004.

[13] H. Kitano. Cancer robustness: tumour tactics. *Nature*, 426:125, 2003.

[14] C. Fuqua, M. R. Parsek, and E. P. Greenberg. Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. *Annu Rev Genet*, 35:439–468, 2001.

[15] P. Williams, M. Camara, A. Hardman, S. Swift, D. Milton, V. J. Hope, K. Winzer, B. Middleton, D. I. Pritchard, and B. W. Bycroft. Quorum sensing and the population-dependent control of virulence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 355:667–680, 2000.

[16] L.-H. Zhang and Y.-H. Dong. Quorum sensing and signal interference: diverse implications. *Molec. Microbiol.*, 53:1563–1571, 2004.

[17] L. You, R. S. Cox, R. Weiss, and F. H. Arnold. Programmed population control by cell-cell communication and regulated killing. *Nature*, 428:868–871, 2004.

[18] K. H. Nealson, T. Platt, and J. W. Hastings. Cellular control of the synthesis and activity of the bacterial luminescent system. *J. Bacteriol.*, 104:313–322, 1970.

[19] G. P. Bodey, R. Bolivar, V. Fainstein, and L. Jadeja. Infections caused by Pseudomonas aeruginosa. *Rev. Infect. Dis.*, 5:279–313, 1983.

[20] E. C. Pesci, J. P. Pearson, P. C. Seed, and B. H. Iglewski. Regulation of las and rhl quorum sensing in Pseudomonas aeruginosa. *J. Bacteriol.*, 179:3127–3132, 1997.

[21] D. L. Chopp, M. J. Kirisits, B. Moran, and M. R. Parsek. A mathematical model of quorum sensing in a growing bacterial biofilm. *J. Ind. Microbiol. Biotechnol.*, 29:339–346, 2002.

[22] A. U. Viretta and M. Fussenegger. Modeling the quorum sensing regulatory network of human-pathogenic Pseudomonas aeruginosa. *Biotech. Prog.*, 20:670–678, 2004.

[23] J. P. Ward, J. R. King, A. J. Koerber, P. Williams, J. M. Croft, and R. E. Sockett. Mathematical modelling of quorum sensing in bacteria. *IMA J. Math. Appl. Med. Biol.*, 18:263–292, 2001.

[24] D. Ramkrishna. *Population Balances: Theory and Applications fo Particulate Systems in Engineering.* Academic Press, San Diego, CA, 2000.

[25] C. D. Immanuel and F. J. Doyle III. Computationally efficient solution of population balance models incorporating nucleation, growth and coagulation: application to emulsion polymerization. *Chem. Eng. Sci.*, 58:3681–3698, 2003.

[26] R. Gunawan, I. Fusman, and R. D. Braatz. High resolution algorithms for multidimensional population balance equations. *AIChE J.*, 50:2738–2749, 2004.

[27] T. Barth and M. Ohlberger. Finite volume methods: foundations and analysis. In E. Stein, R. de Borst, and T. J. R. Hughes, editors, *Encyclopedia of Computational Mechanics*. John Wiley & Sons, Ltd., 2004.

[28] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.

[29] A. Varma, M. Morbidelli, and H. Wu. *Parametric Sensitivity in Chemical Systems*. Oxford University Press, New York, NY, 1999.

[30] R. Gunawan, Y. Cao, L. Petzold, and F. J. Doyle III. Sensitivity analysis of discrete stochastic systems. *Biophys. J.*, 88:2530–2540, 2005.

[31] D. E. Zak, J. Stelling, and F. J. Doyle III. Sensitivity analysis of oscillatory (bio)chemical systems. *Comp. Chem. Eng.*, 29:663–673, 2005.

[32] R. Gunawan and F. J. Doyle III. Phase sensitivity analysis of a circadian gene network. In *Proc. of 44th IEEE Conf. on Dec. and Control*, 2005.

[33] B. Ingalls. Autonomously oscillating biochemical systems: parametric sensitivity of extrema and period. *IEE Sys. Bio.*, 1:62–70, 2004.

[34] H. Suga and K. M. Smith. Molecular mechanisms of bacterial quorum sensing as a new drug target. *Curr. Opin. Chem. Biol.*, 7:586–591, 2003.

[35] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403:335–338, 2000.

[36] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in *escherichia coli*. *Nature*, 403:339–342, 2000.

[37] BioBricks. Registry of Standard Biological Parts. http://parts.mit.edu.

[38] C. C. Guet, M. B. Elowitz, W. Hsing, and S. Leibler. Combinatorial synthesis of genetic networks. *Science*, 296:1466–1470, 2002.

[39] Y. Yokobayashi, R. Weiss, and F. H. Arnold. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. USA*, 99:16587–16591, 2002.

[40] P. François and V. Hakim. Design of genetic networks with specified functions by evolution *in silico*. *Proc. Natl. Acad. Sci. USA*, 101:580–585, 2004.

[41] M. Samoilov, S. Plyasunov, and A. P. Arkin. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. USA*, 102:2310–2315, 2005.

[42] E. Alm and A. P. Arkin. Biological networks. *Curr. Op. Struct. Biol.*, 13:193–202, 2003.

[43] E. M. Ozbudak, M. Thattai, H. N. Lim, B. J. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilization network of *Escherichia coli*. *Nature*, 427:737–740, 2004.

[44] A. P. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in Phage lamda-infected *Escherichia coli* cells. *Genet.*, 149:1633–1648, 1998.

[45] J. E. Ferrell. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *TIBS*, 21:460–466, 1996.

[46] J. L. Cherry and F. R. Adler. How to make a biological switch. *J. theor. Biol.*, 203:117–133, 2000.

[47] S. H. Strogatz. *Nonlinear Dynamics and Chaos*. Perseus Books, Cambridge, MA, 1994.

[48] T. G. Granda and F. Lvi. Tumor-based rhythms of anticancer efficacy in experimental models. *Chronobiol Int*, 19:21–41, 2002.

[49] J. Rutter, M. Reick, and S. L. McKnight. Metabolism and the control of circadian rhythms. *Annu. Rev. Biochem.*, 71:307–331, 2002.

[50] P. L. Lowrey and J. S. Takahashi. Mammalian circadian biology: elucidating genome-wide levels of temporal organization. *Annu. Rev. Genomics. Hum. Genet.*, 5:407–441, 2004.

[51] M. U. Gillette and T. J. Sejnowski. Physiology. Biological clocks coordinately keep life on time. *Science*, 309:1196–1198, 2005.

[52] K. W. Kohn. Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol. Biol. Cell.*, 10:2703–2734, 1999.

[53] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100:57–70, 2000.

[54] V. Hatzimanikatis, K. H. Lee, and J. E. Bailey. A mathematical description of regulation of the G1-S transition of the mammalian cell cycle. *Biotechnol. Bioeng.*, 65:631–637, 1999.

[55] J. J. Tyson and B. Novak. Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions. *J. theor. Biol.*, 210:249–263, 2001.

[56] B. Novk and J. J. Tyson. A model for restriction point control of the mammalian cell cycle. *J. theor. Biol.*, 230:563–579, 2004.

[57] J. C. Dunlap. Molecular biology of circadian pacemaker systems. In J. C. Dunlap, J. J. Loros, and P. J. DeCoursey, editors, *Chronobiology: Biological Timekeeping*. Sinauer Associates, Inc., 2004.

[58] S. M. Reppert and D. R. Weaver. Molecular analysis of mammalian circadian rhythms. *Annu. Rev. Physiol.*, 63:647–676, 2001.

[59] D. B. Forger and C. S. Peskin. A detailed predictive model of the mammalian circadian clock. *Proc. Natl. Acad. Sci. USA*, 100:14806–14811, 2003.

[60] J.-C. Leloup and A. Goldbeter. Toward a detailed computational model for the mammalian circadian clock. *Proc. Natl. Acad. Sci. USA*, 100:7051–7056, 2003.

[61] R. Gunawan, K. G. Gadkar, and F. J. Doyle III. Methods to identify cellular architecture and dynamics from experimental data. In Z. Szallasi, V. Periwal, and J. Stelling, editors, *System Modeling in Cellular Biology*. MIT Press, 2005. in press.

[62] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, Englewood Cliffs, NJ, 1999.

[63] P. D'haeseleer, S. Liang, and R. Somogyi. Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics*, 16:707–726, 2000.

[64] T. Ideker and D. Lauffenburger. Building with a scaffold: emerging strategies for high- to low-level cellular modeling. *Trends Biotech.*, 21:255–262, 2003.

[65] J. Stelling. Mathematical models in microbial systems biology. *Curr. Op. Microbiol.*, 7:513–518, 2004.

[66] A.-L. Barabási and Z. N. Oltvai. Network biology: understanding the cell's functional organization. *Nature Rev. Genet.*, 5:101–113, 2004.

[67] D. W. Selinger, M. A. Wright, and G. M. Church. On the complete determination of biological systems. *Trends Biotech.*, 21:251–254, 2003.

[68] R. Mahadevan, J. S. Edwards, and F. J. Doyle III. Dynamic flux balance analysis of diauxic growth in *E. coli*. *Biophys. J.*, 83:1331–1340, 2002.

[69] R. Gunawan, M. Y. L. Jung, E. G. Seebauer, and R. D. Braatz. Maximum a posteriori estimation of transient enhanced diffusion energetics. *AIChE J.*, 49:2114–2123, 2003.

[70] K. G. Gadkar, R. Gunawan, and F. J. Doyle III. Iterative approach to model identification of biological networks. *BMC Bioinformatics*, 6:155–174, 2005.

[71] D. E. Zak, G. E. Gonye, J. S. Schwaber, and F. J. Doyle III. Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: insights from an identifiability analysis of an in silico network. *Genome Res.*, 13:2396–2405, 2003.

[72] R. Gunawan, D. L. Ma, M. Fujiwara, and R. D. Braatz. Identification of kinetic parameters in multidimensional crystallization processes. *Int. J. Modern Phys. B*, 16:367–374, 2002.

[73] M. Tawarmalani and N. V. Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Math. Programming*, 99:563–591, 2004.

[74] C. A. Schweiger and C. A. Floudas. Optimization framework for the synthesis of chemical reactor networks. *Ind. Eng. Chem. Res.*, 38:744–766, 1999.

[75] D. Gonze, J. Halloy, J.-C. Leloup, and A. Goldbeter. Stochastic models for circadian rhythms: effect of molecular noise on periodic and chaotic behavior. *C. R. Biologies*, 326:189–203, 2003.

[76] Daniel B Forger and Charles S Peskin. Stochastic simulation of the mammalian circadian clock. *Proc. Natl. Acad. Sci. USA*, 102:321–324, 2005.

[77] J. J. Tyson, C. I. Hong, C. D. Thron, and B. Novak. Simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM. *Biophys. J.*, 77:2411–2417, 1999.

[78] J. C. Dunlap, J. J. Loros, and P. J. DeCoursey, editors. *Chronobiology: Biological Timekeeping*. Sinauer Associates, Inc., 2004.

TEACHING INTERESTS

Rudiyanto Gunawan
Department of Chemical Engineering
University of California Santa Barbara
Santa Barbara, CA

I found my teaching experience at the University of Illinois to be enjoyable and rewarding. I was given the opportunity to lecture the "Open-ended Experimental Design" course (ChE469B) in the Fall of 2000, which was offered to master-level graduate students. The course included both lecture and laboratory sections to teach how to efficiently bring processes from design to production. The materials covered included data analysis, process modeling and numerical simulations, model identification, design of experiments, and finally process optimization. The course ended with capstone design projects that were carried out in the laboratory, for example the control of crystal size distribution. At the University of California Santa Barbara, I prepared and presented the tutorial for Bio-SPICE (a open-source software that includes a suite of toolboxes for biological modeling, simulation and analysis) as part of the systems biology course (ChE154) offered to undergraduates (junior and senior) and graduate students. From my teaching experience, I learned that relevant open-ended problems and case studies can serve as a powerful and effective teaching tool.

In the following, I outline the courses that I am interested and suited to teach with a brief explanation of the course content and the textbooks to use.

1 Systems Biology

The exponential growth of biological data and the complexity cellular systems necessitate the use of a systems perspective to fully understand how a cell accomplishes its myriad functions. Systems biology emerged to answer this challenge by combining approaches from science and engineering such as statistics, computer science, and systems engineering.

The systems biology course will teach students how to bring tools from mathematics, statistics, and engineering to study biological systems. This course will explore different aspects of systems biology, including system modeling and analysis, data mining and network inference, and experiment design in biology. In addition, the course will introduce various numerical approaches (simulations, optimizations), which are required to solve problems in systems biology. I will utilize case-studies taken from the literature, such as bacterial chemotaxis [1], λ-phage [2], circadian rhythm [3], and apoptosis [4], in which the students will apply tools from the course and present the results as a capstone project. The course will also incorporate hands-on experience in using available software for systems biology, such as Bio-SPICE (http://www.biospice.org).

As systems biology is still in its infancy, there does not yet exist a textbook that will sufficiently cover all the aspects of the course. Instead, the course will be based on lecture notes and selected readings from the literature, such as the upcoming book *System Modeling in Cellular Biology* [5]. This course will be offered to undergraduates with junior and senior standings and graduate students.

2 Metabolic Engineering

Fueled by advances in genomics and proteomics, the understanding of cellular metabolism enables the modification and optimization of cellular metabolic pathways, and even the introduction of new pathways, to produce desired biochemical products. The aim of this course is to introduce the basic principles and analytical methods in cellular metabolic engineering. The course topics include an overview of cellular metabolism, model representation and reverse engineering of metabolic networks, numerical optimization methods, metabolic system analyses (metabolic control analysis and metabolic flux balance analysis), and finally optimal design of metabolic networks. The course content will be based on the textbook by Stephanopoulos *et al.* [6] with selected chapters from Voit [7] and Heinrich and Schuster [8]. The capstone project involves a critical reading of recent literature on the general subject of metabolic engineering and a presentation by each student. The course will be open to advanced undergraduate students and graduate students.

3 Model-Based Experiment Design

Modeling and experimental efforts form an integral approach in a hypothesis-driven research in systems biology [9]. Mathematical models serve multiple uses in this approach including hypothesis generation and testing, system analysis, and experiment design. The aim of this course is to introduce different techniques in model-based experiment design. The topics in the course include model development, simulation, and validation, system analysis, data analysis, parameter identification, and optimal experiment design. Finally, the course will include a capstone project in which the students apply the model-based approach in wet-lab or *in-silico* experiments. The course materials are based on a textbook on system identification by Ljung [10] and selected chapters from other textbooks by Beck and Arnold [11] and Box *et al.* [12]. This course is appropriate for undergraduates with senior standings and graduate students.

4 Other Courses

Aside from the courses listed above, I am willing to teach any undergraduate or graduate courses with emphasis on applied mathematics, dynamical systems analysis, control, and optimization, and computational approach in life science and engineering.

References

[1] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis. *Proc. Natl. Acad. Sci. USA*, 97:4649–4653, 2000.

[2] A. P. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in Phage lamda-infected *Escherichia coli* cells. *Genet.*, 149:1633–1648, 1998.

[3] J.-C. Leloup and A. Goldbeter. Toward a detailed computational model for the mammalian circadian clock. *Proc. Natl. Acad. Sci. USA*, 100:7051–7056, 2003.

[4] M. Fussenegger, J. E. Bailey, and J. Varner. A mathematical model of caspase function in apoptosis. *Nature Biotech.*, 18:768–774, 2000.

[5] Z. Szallasi, V. Periwal, and J. Stelling, editors. *System Modeling in Cellular Biology.* MIT Press, 2005. in press.

[6] G. N. Stephanopoulos, A. A. Aristidou, and J. Nielsen. *Metabolic Engineering: Principles and Methodologies.* Elsevier Science, 1998.

[7] E. O. Voit. *Computational Analysis of Biochemical Systems.* Cambridge University Press, 2000.

[8] R. Heinrich and S. Schuster. *The Regulation of Cellular Systems.* Chapman & Hall, 1996.

[9] H. Kitano. Systems biology: a brief overview. *Science*, 295:1662–1664, 2002.

[10] L. Ljung. *System Identification: Theory for the User.* Prentice Hall, Englewood Cliffs, NJ, 1999.

[11] J. V. Beck and K. J. Arnold. *Parameter Estimation in Engineering and Science.* John Wiley & Sons, Inc., 1977.

[12] G. E. P. Box, W. G. Hunter, and J. S. Hunter. *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.* John Wiley & Sons, 1978.