REVERSE ENGINEERING HOMEOSTASIS IN MOLECULAR BIOLOGICAL SYSTEMS

A Thesis
Presented to
The Academic Faculty

by

Chang Feng Quo

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in
Bioengineering

Wallace H Coulter Department of Biomedical Engineering
Georgia Institute of Technology
August 2013

REVERSE ENGINEERING HOMEOSTASIS IN MOLECULAR BIOLOGICAL SYSTEMS

Approved by:

May D Wang, Ph.D., Advisor
Wallace H Coulter Department of Biomedical Engineering
Georgia Institute of Technology

Melissa L Kemp, Ph.D.
Wallace H Coulter Department of Biomedical Engineering
Georgia Institute of Technology

Alfred H Merrill, Jr., Ph.D.
School of Biology
Georgia Institute of Technology

Leonid Bunimovich, Ph.D.
School of Mathematics
Georgia Institute of Technology

Mark R Prausnitz, Ph.D.
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Date Approved: March 13, 2013

To Yuan,

and

my daughter, Chi-Chi

ACKNOWLEDGEMENTS

First, I want to thank

Dr. May D Wang, my advisor, for faith in me to explore this topic freely and operate with minimal constraint;

Dr. Al Merrill, for substantive and philosophical discussions, as well as personal accounts, that inspired me and helped me persevere, and

Dr. Melissa Kemp, Dr. Bunimovich, and Dr. Prausnitz, my committee members, for specific reviews and suggestions, that helped this work get some acceptance in the community.

In addition, this dissertation would not be possible without the following people:

- Elaine Wang, Chris Haynes, Jeremy Allegood — experimental data acquisition,
- Richard Moffitt, Neil Shah — parameter estimation,
- Chanchala Kaddi — technical discussions.

Finally, this work is supported by grants from the Lipid Metabolites and Pathways Strategy consortium (LIPIDMAPS GM069338), the Parker H. Petit Institute for Bioengineering and Bioscience, Johnson & Johnson, Bio Imaging Mass Spectrometry Initiative at Georgia Tech, National Institutes of Health (BRP R01CA108468, CCNE U54CA119338), Georgia Cancer Coalition (Distinguished Cancer Scholar Award to my advisor, Dr. May D Wang), and Microsoft Research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation is an initial study of how modern engineering control may be applied to reverse engineer homeostasis in metabolic pathways using high-throughput biological data. This attempt to reconcile differences between engineering control and biological homeostasis from an interdisciplinary perspective is motivated not only by the observation that robust behavior in metabolic pathways resembles stabilized dynamics in controlled systems, but also by the challenges forewarned in achieving a true meeting of minds between engineers and biologists.

To do this, a comparator model is developed and applied to model the effect of single-gene (SPT) overexpression on C16:0 sphingolipid *de novo* biosynthesis *in vitro*, specifically to simulate and predict potential homeostatic pathway interactions between the sphingolipid metabolites. Sphingolipid *de novo* biosynthesis is highly regulated because its pathway intermediates are highly bioactive. Alterations in sphingolipid synthesis, storage, and metabolism are implicated in human diseases. In addition, when variation in structure is considered, sphingolipids are one of the most diverse and complex families of biomolecules. To complete the modeling paradigm, wild type cells are defined as the reference that exhibits the "desired" pathway dynamics that the treated cells approach.

Key model results show that the proposed modern engineering control approach using a comparator to reverse engineer homeostasis in metabolic systems is: (a) effective in capturing observed pathway dynamics from experimental data, with no significant difference in precision from existing models, (b) robust to potential errors in estimating state-space parameters as a result of sparse data, (c) generalizable to model other metabolic systems, as demonstrated by testing on a separate independent dataset, and (d) biologically relevant in terms of predicting steady-state feedback as a result of homeostasis that is verified in literature and with additional independent data from drug dosage experiments.

CHAPTER I

INTRODUCTION: ROBUSTNESS IN METABOLIC SYSTEMS

Recent advancements in high-throughput biotechnology enable, and in fact necessitate, matching developments in the field of biology. Specifically, studies in genomics, proteomics, and metabolomics require comparable methods for data mining in order to handle vast volumes of biological data that can otherwise prove unwieldy. Furthermore, recognizing the opportunity to use this wealth of data to "understand biology at the system level" [77], '-omic' technology has also led to the emergence of "systems biology", which is focused on the abstraction of data, or modeling [124].

As an engineer-in-training, it is my intuition that "control" is integral to a system-level understanding of biology. More precisely, the premise of control in biology, in particular systems biology, is that robust behaviors in biological systems resembles stabilized dynamics in engineering systems under control. Consequently, control theory is essential to, and should complement, the study of systems biology.

A number of opportunities and challenges may emerge from systems biology for control theory, which can be broadly classified in four categories [146]. They are: (a) the role of control and signal processing techniques in bioinstrumentation, (b) the use of existing techniques in well-developed areas of control theory to analyze problems of interest to biologists, (c) the abstraction of new ideas for control engineering from biology, and (d) the formulation of new theoretical problems in control theory.

This dissertation is focused on the second category of problems. In particular, using lipidomic data on a metabolic pathway that is reported to be highly regulated in nature, a comparator is developed and applied to model the differences between treated and wild type cells, under the assumption that treated cells approach the same steady-state as the wild type in terms of intracellular lipid amounts. Thus, in this case, the effect of the applied comparator is to predict pathway feedback that may be responsible to maintain pathway

stability under a specific treatment condition.

1.1 Mechanisms of metabolic robustness: emergence and design

Biological robustness refers to the persistence of living organisms to maintain certain traits or behaviors under changing, and often times unfavorable, conditions [78, 94]. Specifically, phenotypic robustness refers to the ability of living organisms to adapt, reproduce, and evolve to changes in the environment over time as a result of genetic mutation, while component robustness refers to the persistence of metabolic function in living organisms under dynamic and relatively short-lived conditions in the micro-environment. The distinction between these sub-definitions occurs in terms of scale in space and time, where phenotypic robustness spans the levels of cell to population over multiple generation times while component robustness is generally limited to the cell level within a single generation. Component robustness, or robustness in biochemical systems, is the focus of this section.

Robustness in biochemical systems is an increasingly popular research topic for two reasons. First, it is a ubiquitous property, i.e., it is readily observed, e.g., in glycolysis [53], energy metabolism [32], protein synthesis [15], mitochondrial apoptosis [62], and chemotaxis [13, 176, 63]. Second, with sufficient understanding of the underlying mechanisms, it may be manipulated to serve the prevailing needs of society, e.g., for disease treatment [10, 79, 108, 154], in particular drug development [66, 68, 80, 101, 100]. For the latter reason, the public literature on this topic to date comprises mostly studies on representative mechanisms of robustness such as hierarchy [151], modularity and redundancy [97, 114], stochasticity [130], connectivity (and topology) [94], and more recently, feedback (and feedforward) control [56, 172]. Yet, insofar as such mechanisms are well-documented, there is comparatively little question (and even fewer answers) on the conception of robustness; furthermore, on how underlying assumptions on conception may bias the choice of one apparent mechanism, or few combined, over others to explain and perhaps in future to enable the robustness property in biochemical systems. Although not often discussed explicitly, a review of the public literature suggests that robustness in biochemical systems may be attributed to one of two notions: emergence and design.

In the context of systems theory, 'emergence' refers to the 'coming out' of more complex behaviors from simpler parts, especially where such 'higher' properties cannot be reasonably deduced from its 'lower' component interactions (see Bedau [18] for a more rigorous, and philosophical, treatment as well as the difference between strong and weak emergence). In other words, emergence is 'bottom-up'. In many instances, emergence is synonymous with synergy, where the oft-encountered phrase "the whole is greater than the sum of its parts" is an acceptable though simplified description of the phenomena. From the aforementioned list, mechanisms of robustness in biochemical systems that are implied to arise as a result of emergence include hierarchy, modularity and redundancy, stochasticity, and connectivity and topology.

On the other hand, 'design' refers to the specification of a particular construct, or set of constructs, to accomplish specific objectives under certain constraints [136]. In a well-designed system, various components are integrated to perform a well-defined function. In other words, design is 'top-down'. The apparently seamless combination of parts to serve a single purpose in a well-designed system has sparked a lively discussion of irreducible complexity throughout the years [19, 113], which is also implicated in the debate on evolution versus intelligent design. The point being that robustness of biochemical systems by 'design' implies, or in some perspectives necessitates, the existence of a designer (an argument better left for the interested reader to explore separately) or an almost incredible confluence of events. To reconcile this rather awkward implication with mainstream support for the theory of evolution, it is worthwhile to consider that perhaps 'welding points' in robust biochemical systems may have 'eroded' over time such that prior steps in evolution are not detectable, which makes robustness by design the best explanation. As a recently proposed mechanism of robustness, feedback (and feedforward) control may be attributed to design.

Here, the question of whether the property of robustness in biochemical systems is conceived from emergence or design is a basis to discuss the strength and implications of a decision to pursue one resultant mechanism, or few combined, over others to explain the phenomenon, not a mirror of the debate on evolution versus intelligent design. Although not often acknowledged, the starting point and subsequent path of research on robustness

in biochemical systems is naturally affected by the assumption, indeed acceptance, of either emergence or design as its basis.

What follows in this section is a survey of the robustness property in biochemical systems in terms of the aforementioned mechanisms, i.e., hierarchy, modularity and redundancy, stochasticity, connectivity and topology, and feedback (and feedforward) control; their assumed basis in emergence or design, and theoretical tradeoffs in implementing these mechanisms leading to the intriguing 'robust yet fragile' feature of such systems [5, 41, 83, 82]. To facilitate discussion, robustness of biochemical systems in this context shall refer to the ability of a cell signaling system to function reliably despite changes in effective molecular concentrations, i.e., to produce a desired output response if and only if there is an appropriate input stimulus. Finally, based on knowledge of these representative mechanisms, the future of research in biochemical robustness and its applications is discussed.

1.1.1 Hierarchy

Hierarchy refers to the stratification of component molecules into different biochemical levels of organization, e.g., genes (nucleotides), enzymes (proteins), and metabolites (lipids and carbohydrates) [157]. Specifically, order is enforced in the hierarchy where 'higher-order' molecules govern the action of 'lower-order' molecules, e.g., genes code for enzymes that act in turn on metabolites. Furthermore, signals from higher-order molecules are amplified in lower-order levels, e.g., a single strand of messenger RNA can be translated repeatedly to produce multiple enzymes. This cascade effect in cell signaling systems also implies that it is easier for signal transmission to occur in one direction from higher-order to lower-order molecules than the opposite.

Based on these features of stratification, ordering, and cascade effect, hierarchy as a mechanism ensures system robustness in two ways. First, by effectively compartmentalizing the functional molecules, weaknesses and failures are contained within particular levels when they occur. Second, by enforcing order and directional bias, variations in amounts of lower-order molecules, which occur most frequently, are not as easily propagated to higher-order molecules. In addition, because of the amplification cascade, hierarchy can also aid in

minimizing the cost of cell signaling where energy needed for transcription is higher than for translation. The bow-tie architecture observed in cell signaling systems illustrates these features of hierarchy as a mechanism of robustness in biochemical systems [80], where the knot in the middle represents higher-order molecules while the ends of the bow represents lower-order molecules.

At the same time, there are tradeoffs in implementing hierarchy as a mechanism of system robustness. First, in demarcating clearly stratified levels, there is an equivalent need to develop protocols to communicate between the corresponding levels [33]. For example, transcription and translation are protocols to facilitate communication between DNA, RNA, and enzymes. It follows that cell signaling, which traverses various biochemical levels, relies heavily on the efficacy of these protocols. Thus, while necessary, such protocols are points of weakness and fragility in the system. Second, because higher-order molecules govern lower-order molecules, additional fragility and risk is introduced to the system when perturbations can occur at the level of higher-order molecules, e.g., in virus infections where foreign DNA is introduced into the host cell and mistakenly replicated. Perturbations at this level are also more readily propagated and amplified because of the cascade effect. Third, also because of the cascade effect, sensitivity and signaling speed are lower because of an effective directional bias against signaling from lower-order to higher-order molecules. For instance in sensing, i.e., where signals are initiated by lower-order molecules, a higher stimulus threshold is required to generate an output response, e.g., in terms of sustained receptor activation over time or simultaneous activation of multiple receptors.

Hierarchy as a mechanism for robustness due to emergence may be supported by the endosymbiotic theory [84]. On this basis, it follows that to establish hierarchy as the mechanism of robustness for a given biochemical system it is critical to (1) identify and establish order between groups of functionally dissimilar or at least non-interchangeable molecules, and (2) discern the molecular protocols for communication between levels.

1.1.2 Modularity and redundancy

Modularity refers to the autonomy of component molecules, usually in self-contained functional groups, to perform specific actions that contribute towards the overall system objective. The trivial definition is that an individual enzymatic reaction is a module in a series of reactions that accomplishes a specific metabolic function. Modules in a given system may also be ordered, i.e., hierarchical, but not necessarily so. Modularity alone does not guarantee robustness and must be complemented by redundancy [114, 166], i.e., many modules exist to perform the same functions in a system. Thus, taken together, these features of modularity and redundancy ensure robustness in biochemical systems.

Modularity and redundancy are similar to hierarchy but key differences exist. First, modularity resembles hierarchy in terms of autonomy but modules are not necessarily hierarchical. The feature of being 'self-contained' ensures that failed modules are functionally removed from the system and do not interfere with overall system performance. At the same time, redundancy ensures that there are other modules that accomplish the same function to fill the void. Second, redundancy differs from the cascade effect in that the extent of redundancy is the same for all modules whereas in the cascade effect, duplication occurs only for some components, which are also largely lower-order molecules. As a result, redundancy does not also facilitate signal amplification unlike the cascade effect.

There are also two main issues in implementing modularity and redundancy as a mechanism for robustness in biochemical systems. First, in maintaining redundancy, it is costly because resources must be dedicated to create more than one component to perform the same function. Second, to ensure modularity also implies separating defined modules, which may difficult to do in the crowded molecular environment of the cytoplasm. As a result, the fragility of this mechanism for robustness is that biochemically similar molecules from one functional module may interfere or compete with molecules in another, leading to unintended system performance. These issues may be resolved where if modularity can be defined and implemented efficiently, then modules that perform functions common to several systems may also be shared. The sharing of such common modules may then balance

the additional costs required to maintain redundancy. In these considerations, functionality is the basis on which modularity and redundancy as a mechanism for robustness in biochemical systems may be attributed to emergence [97]. Thus, the challenge in research on modularity and redundancy as a mechanism for robustness in biochemical systems is to clearly define appropriate functional units while allowing for the possibility of duplication.

1.1.3 Stochasticity

Stochasticity in biochemical systems refers to the probabilistic nature of biochemical reactions as a result of the unpredictable motion of molecules [130, 131]. As a mechanism for robustness, stochasticity depends on maintaining threshold concentrations of molecules so that rates of reaction in a biochemical system remain effective to accomplish system function despite low odds of collision for individual molecules. At the same time, the stochastic nature of biochemical reactions also minimizes the probability of 'rogue' reactions as a result of perturbation by biochemically similar but functionally different molecules, i.e., by decreasing sensitivity.

Because of its probabilistic nature, stochasticity as a mechanism for robustness in biochemical systems suggests some readily observable features. First, such systems must be relatively simple. The low odds of reaction implies that theoretically, there is a limit on the number of reactions for a biochemical system to be robust via stochasticity. Second, to increase the odds of reaction in an already crowded cytoplasm, requisite molecules may be sequestered to maintain optimal concentrations. This may manifest in terms of increased number of vesicles in the cytoplasm or extensive use of existing organelles for compartmentalization. Third, the 'induced fit' theory of enzymatic action suggests that enzyme specificity may be relaxed for reactions in such systems to ensure overall system function. So, enzymes involved in such robust systems by stochasticity may be less specific in substrate recognition compared to other enzymes. Last, cofactor binding or positive cooperativity in enzymatic action may also increase the odds of reaction, which is another empirical feature.

From these features, it follows that stochasticity as a mechanism for robustness in biochemical systems can be attributed to emergence from component enzymatic reactions. Interestingly, the probabilistic nature of biochemical reactions also confers a degree of fragility as a result of a few 'stiff' combinations of parameters [36], which may require other considerations to overcome, e.g., by including other structural elements [9].

1.1.4 Connectivity and topology

Connectivity is a concept of graph theory that refers in general to the minimum number of nodes or edges to be removed before a network is disconnected. A more robust graph, or network, requires more network elements to be removed before it is disconnected. As a measure of robustness in graph theory, connectivity is also a mechanism of robustness in biochemical systems because of the nature of linkage in biochemical systems, e.g., gene, metabolic, and protein-protein interactions, specifically in terms of topology [6, 94, 175]. In particular, the scale-free topology of biochemical systems is directly linked to robustness in biochemical systems [4, 14, 153].

At its core, connectivity in biochemical systems is measured in terms of the frequency of nodes that are associated with specific number of edges, i.e., degree distribution. In random networks, the degree distribution resembles a Poisson distribution that peaks at the average degree. This means that because the edges are placed randomly in a random network, most nodes have similar numbers of edges. However, in real networks, e.g., biochemical networks, the degree distribution resembles a power-law distribution that deviates significantly from the Poisson distribution. This means that a small number of nodes are connected to many nodes, i.e., hubs, while a large number of nodes are connected to fewer nodes.

Assuming an equal chance of failure for individual nodes, robustness in biochemical systems is ensured as a result of scale-free topology because in this topology there are significantly fewer hubs compared to other less connected nodes. Thus, the basis of robustness in biochemical systems is emergence from the nature of linkage, or topology, between component molecules. It follows that the effect of connectivity to ensure robustness is appreciable only in large systems or networks where the ratio of hubs to nodes is non-trivial, e.g., the

World Wide Web, the Internet, and cells [2]. At the same time, while scale-free topology confers high error tolerance, i.e., robustness to high failure rates of nodes in general but is vulnerable to attacks, i.e., when a hub breaks down or is removed [3, 70]. It remains to be seen if other topologies, i.e., connectivity in terms of different degree distributions, may also facilitate robustness in biochemical systems.

1.1.5 Feedback and feedforward control

Feedback control, also known as 'closed-loop' control, refers to the use of current output to regulate future output [17, 173]. To do this, sensor units sample the system output at the present time, which is relayed to actuator units to guide or regulate system output at the next time, thus 'closing the loop'. Although well developed in control theory, the concept of feedback control entered mainstream research only recently in the context of robustness in biochemical systems [71, 146, 168]. This may be in part because of difficulties in reconciling its basis by (intelligent) design with support for the theory of evolution; the corollary is that there has been more attention on mechanisms of robustness based on emergence than on design as seen in the literature. Since then the concept of feedback control has been readily developed in the context of biochemical systems, e.g., in terms of concentration and buffering, [138, 143], homeostasis [37], and gene regulation [43, 142].

The feature of feedback control in biochemical systems is that sensor and actuator units are used to sample and incorporate the current output into the system dynamics to guide future output. This system architecture forms a closed loop between current and future outputs, which ensures robustness by guiding the output to converge to the desired state. Specifically, based on the difference between current output and the desired state from the sensor unit, the actuator unit regulates the system to generate future output that is closer to the desired state. This cycle terminates when the current output and desired state are the same. Thus, the effectiveness, and corresponding fragility, of feedback control as a mechanism for robustness depends on the design of sensor and actuator units [76].

There are various different ways to implement feedback control in biochemical systems,

e.g., proportional, integral, and derivative (PID) feedback in energy metabolism [32], integral feedback in bacterial chemotaxis [7, 177], and multiple feedback loops in E. coli tryptophan regulation [22]. Consequently, it can be difficult to determine the exact form that is responsible for robustness in a given biochemical system. This issue arises because, from the perspective of design, the various forms of feedback control are sufficient but not necessary to ensure robustness. Nonetheless, in the research of feedback control as a mechanism of robustness in biochemical systems, this challenge can still be overcome by the process of elimination, i.e., to first assume a specific form of feedback control and then test for particular dynamic features that are associated with it [103].

In addition to feedback control, feedforward control is also a form of control that ensures robustness in biochemical systems. In feedforward control, the input stimulus is not only received by the immediate system module, e.g., receptor, but is also relayed ahead to one or more subsequent modules. In other words, the input stimulus triggers a reaction in more than one module in the system. Such a mechanism is robust to ensure system output in response to input stimulus in case of intermediate module failure, amplifies the effects of the input stimuli, and also increases the response time of the system, e.g., in the use of a transcriptional factor to trigger the bacterial heat shock response [44].

1.1.6 Themes

Emergence is 'bottom-up' while design is 'top-down'. In other words, robustness in biochemical systems by emergence or design underscores two contrasting approaches to research on this topic, in particular to model biochemical systems based on modularity and robustness [27, 150]. The bottom-up approach is appealing because model interactions, namely enzymatic reactions, can be specified from traditional biological knowledge. However, in this approach it is also necessary to identify a large number of kinetic parameters. In most cases, it is possible to estimate but not measure these parameters because of data issues. Even so, parameter estimation is not without its challenges. The top-down approach requires only that a performance objective for the system be specified, namely robustness, so that accuracy of kinetic parameters is not as critical. However, at the present time,

the concept of biochemical robustness by design is difficult to reconcile with mainstream support for emergence. So, an interim solution to spur progress on this topic may be to combine top-down and bottom-up approaches so that problems of parameter estimation may be alleviated without the loss of biological knowledge [159].

At the same time, models of biochemical robustness by design may also play a role to elucidate molecular mechanisms in robust biochemical systems. The logical premise of this application is that because robustness is observed in a system, then an appropriate representation (model) of that system is also robust [115, 150]. For example, feedback control may be simulated by comparing the response of a robust system under perturbation to a baseline (wild type) response [125]. Then, the resulting feedback may be analyzed to yield insight into the contribution of, or interplay between, component molecules to ensure robustness in the biochemical system. Such an application is an example that illustrates how top-down approaches may be combined with bottom-up approaches to model robustness in biochemical systems.

A persistent theme in research on biochemical robustness is the development of more rigorous definitions of robustness in the context of biochemical systems, e.g., in terms of graphical representations of structure in biochemical systems [55], model reduction [128], and mathematical formulation towards a formal theory [81]. Although robustness is a ubiquitous property, current studies on this topic tend to be reported in different, often qualitative, terms [89, 95]. As a result, it is not easy to compare study findings to uncover deep, structural commonalities between robustness in different systems. Thus, the development of a standard, possibly analytical, expression for biological robustness is a key challenge, which will surely contribute to the discovery of fundamental principles for robustness.

Finally, robustness is a common, arguably definitive, property of biochemical systems that is manifest and achieved via different mechanisms. In practice it is generally the result of a combination of these mechanisms. For example, in the context of cell signaling, such combinations ensure robustness where some mechanisms may favor initial signal activation, e.g., modularity and redundancy, while others may favor continued signal transmission, e.g., hierarchy. Hence, research on these various mechanisms can be rallied to a common

objective, i.e., to elucidate the origin(s), execution [132], and rationale [30, 44] of robustness in biochemical systems. However, in both theory and practice, the process to integrate different mechanisms in a biochemical system to achieve robustness is not yet clear and also raises other intriguing questions, e.g., on the controllability of biochemical networks [23, 87, 92, 96]. Thus, the question of how different mechanisms are, or can be, integrated to ensure robustness in biochemical systems will surely be a focus for continuing research.

1.2 Metabolic system dynamics and control

1.2.1 Reaction kinetics

The dynamics of metabolic systems (or pathways) is primarily defined in terms of biochemical, or enzymatic, reaction kinetics. For example, generalized mass action (GMA) and flux balance analysis (FBA) are common approaches to quantify enzymatic reaction kinetics based on the underlying principle of conservation of mass. In engineering terms, these methods provide the equations of motion to describe metabolic system dynamics in terms of enzymatic reactions. They are written in the time domain, and take the form of ordinary differential equations (ODEs):

$$\dot{\vec{x}} = f(\vec{x}(t), t) \tag{1}$$

where $\vec{x}$ represents the variables of interest generally in terms of metabolite amount (concentration) or flux, and f is a function of metabolite amounts over time t, which describes the rate of change of these quantities based on an understanding of the enzymatic reactions between said pathway metabolites.

1.2.1.1 Generalized mass action

In generalized mass action (GMA), the metabolic system dynamics is described in terms of metabolite amount (concentration). The rate of change of metabolites is quantified based on the *law of mass action* to describe and predict enzymatic reactions in solution. Precisely, the law of mass action states that the rate of an elementary reaction, i.e., a reaction that proceeds in one step, is proportional to the product of the amounts (or concentrations) of participating molecules [58, 99]. More complex reactions that occur in multiple steps may be written as a series of elementary reactions.

Thus, consider the (elementary) reaction:

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C} \tag{2}$$

where substrates A, B react (irreversibly) to form product C, and define k as the (positive, forward) reaction rate constant. Then, the rate of change of these molecules are

written as:

$$\frac{d\mathrm{A}}{dt} = \frac{d\mathrm{B}}{dt} = -\frac{d\mathrm{C}}{dt} = -k\mathrm{AB} \tag{3}$$

Although the law of mass action was developed to describe molecules in solution, it may also be generalized to describe interactions between large numbers of individuals, e.g., wildlife populations in ecosystems. In theory, GMA is a simple and intuitive model that approximates aggregate interactions between individuals in large populations. In other words, the (kinetic) probability of interaction is proportional to the number of individuals. However, in practice, the proportionality constant is not easily obtained. In this case, this means that it is experimentally difficult to measure, or even estimate, the required enzymatic reaction rate constants.

1.2.1.2 Flux balance analysis

In flux balance analysis (FBA), metabolic system dynamics is described in terms of steady-state flux between metabolites under the assumption that particular objectives, e.g., homeostasis, are optimized [74, 119, 129, 162]. This approach combines metabolite amount (concentration), rate of reaction, and stoichiometry to represent metabolic flux. Thus, FBA is a constraint-based approach that seeks to describe the rates at which metabolites (variables) are exchanged between various states (compartments), i.e., flux, in a given system.

Thus, consider the reaction:

$$\mathrm{A} + \mathrm{B} \leftrightarrow \mathrm{C} \tag{4}$$

where substrates A, B react (reversibly) to form product C, and define $v_{forward}$, $v_{reverse}$ as the forward and reverse fluxes. Then, under the steady-state assumption, the rate of change of the balanced fluxes are written as:

$$\frac{d\vec{x}}{dt} = \mathcal{S} \cdot \vec{v} = \vec{0} \tag{5}$$

where

$$\vec{x} = \begin{pmatrix} \text{A} \\ \text{B} \\ \text{C} \end{pmatrix}, \qquad \mathcal{S} = \begin{bmatrix} -1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \vec{v} = \begin{pmatrix} v_{forward} \\ v_{reverse} \end{pmatrix}$$

$\vec{x}$ is a vector of molecular amounts (or concentrations), and $\mathcal{S}$ is the (balanced) stoichiometric matrix of the system that describes $\vec{v}$, a vector of (forward and reverse) fluxes that are unknown quantities of interest. Furthermore, fluxes may also be specified as internal, as shown in this example, or external to a system. FBA has been reported in models of biochemical processes in a variety of organisms that includes *E. coli* [42], yeast, and plants [57]. In theory, even for dense networks, fluxes in systems at equilibrium may be balanced to give linear equations. However, in practice, it is difficult to verify that flux is a constant quantity in metabolic systems.

1.2.2 Pathway regulation

In existing approaches to model metabolic systems, control is not usually differentiated from dynamics. Existing approaches describe only metabolic system dynamics, i.e., motion, in terms of parameters that must be specified, either by measurement or estimation. Examples of parameters in such approaches include enzymatic reaction rate constants (mass action) [58, 99], stoichiometric coefficients and flux rates (flux balance) [162], control coefficients and elasticities (metabolic control analysis) [46, 65, 73], and kinetic orders (biochemical systems theory) [134, 135]. Because metabolic system dynamics are described only in terms of such parameters, the problem of modeling metabolic systems can be rephrased into a question of parameter estimation, or system identification in engineering terms.

Importantly, using such approaches, it is implied that certain properties of metabolic system dynamics, in particular robustness, can be fully explained with the associated parameters. However, to do so using -omic data, the number of parameters in recently published models of metabolic systems tends to be significantly larger than the number of variables. This leads to the problem of data over-fitting. Furthermore, despite the numbers, these parameters contain proportionately little information about the metabolic system dynamics, e.g., in terms of parameter sensitivity [8, 59].

1.2.2.1 Differentiating control from dynamics

Especially in complex systems, control, i.e., how system dynamics is regulated, is essential to achieve system stability. Furthermore, it is necessary in reverse engineering stable metabolic systems that components of control be differentiated from components of dynamics and motion.

To illustrate the point, consider self-righting objects for instance (Figure 1). Precisely, the object geometry, a round bottom, is the basis for dynamics and motion, i.e., rotation about a point of stable equilibrium that represents the point of minimum gravitational potential energy. To ensure self-righting behavior, such objects usually contain ballast so that the center of mass is lower than its geometry suggests (assuming uniform density). This ensures that any tilting about the fixed point raises the center of mass and is followed by a subsequent return to equilibrium.

Thus, for a self-righting object, the ballast provides the necessary control mechanism to regulate the object dynamics and motion due to its geometry. If not for the recent construction of a self-righting object of uniform density, it may be easy to overlook the role of ballast as a control mechanism in such self-righting objects [39, 161]. [1]

1.2.2.2 Control theory in biology

Cybernetics is the study of structural complexity in animal and machine that enables communication and control [169], and is closely related to control theory. In biology, this approach was first applied to study organ systems in physiology, e.g., in circulation [60, 121], immunology [85, 155], the central [23, 28] and peripheral [49, 75] nervous systems, and even bone remodeling [51]. Recently, this approach has been applied to study metabolic systems that include biochemical pathways in *E. coli*, human hepatocytes and erythrocytes [20], as well as in adult rat cardiomyocytes and human skeletal muscle [61]. Nonetheless, while cybernetics emerged as the science of effective organization within systems, control theory was developed to influence and guide the dynamics of complex systems.

In general, control theory deals with the design of particular controllers to influence and

[1] In addition, see [40, 160] for how such geometry is found naturally in the shell of the Indian Star Tortoise.

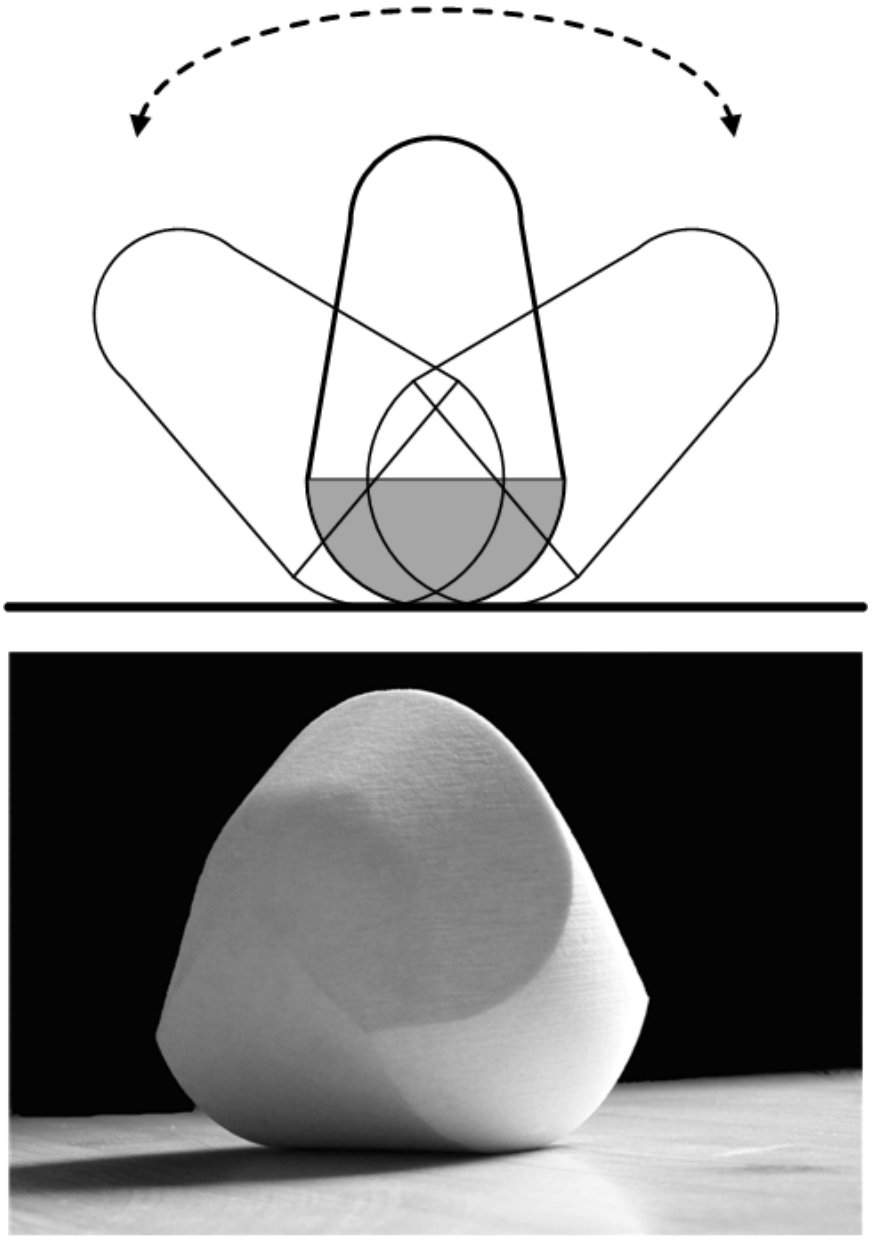


Figure 1: Stability of motion in self-righting objects: *(top)* ballast regulates motion in objects of non-uniform density; *(bottom)* geometry is solely responsible for self-righting in an object of uniform density, the Gömböc. (Gömböc image taken from http://www.thestar.com/article/269792)

guide the dynamics of complex systems. In biology, some examples include the design of optimal control strategies to manage wildlife populations [133, 170], or to maintain desirable pest populations in agricultural systems [24, 25, 26]. Recently, control theory has also been applied at the level of metabolic systems, e.g., to regulate the amount of cells in a bioreactor using robust sliding mode control [48], or to regulate cell function in *Escherichia coli* (E. coli) by constructing a bi-stable protein switch based on the switching properties of *lambda* phage [64]. In such applications for controller design, the key constraint is the difficulty involved in implementing the requisite actuators to effect control. In other words, the control of metabolic systems is limited by the extent to which biochemical mechanisms in these complex systems can be manipulated.

1.3 Reverse engineering homeostasis

Homeostatic pathways resemble engineering systems in that both types of systems are robust to disturbance within limit, i.e., system stability is maintained under small perturbations. Based on such observations, recent opinion express cautious optimism to reverse engineer the complexity and regulation of metabolic pathways by applying control theory in the context of systems biology [33, 156]. In particular for metabolic systems, the robustness of some cellular functions may be described in terms of control theory [91, 151]. For even more practical applications, it is envisioned that a theory of biological robustness may be developed to help understand the robustness of metabolic drug responses so as to improve drug therapy [80, 81]. Thus, and for various other reasons, reverse engineering robustness of metabolic systems, as a result of homeostasis, has attracted interest in the community not only to understand how homeostatic pathways may have come to be, but also with the potential of influencing and guiding the dynamics of deviated pathways.

1.3.1 Homeostasis as control

Homeostasis is the process of control, i.e., the regulation of pathway dynamics, to maintain a stable condition in metabolic systems. While homeostasis is highly complex, homeostatic mechanisms can be, and have been, simplified in terms of control theory. To illustrate the point, two examples of homeostasis in synaptic signaling and glycolysis are discussed in terms of open- and closed-loop control.

1.3.1.1 Open-loop control in synaptic signaling

Action potentials are propagated by the regulated flow of ions – calcium (Ca^{2+}), sodium (Na^{+}) and potassium (K^{+}) – into and out of cells via voltage- and receptor-gated channels on the cell membrane (Figure 2). Across gaps between cells, or synapses, propagation of action potentials is mediated by neurotransmitter molecules, e.g., acetylcholine, dopamine, norepinephrine, and glutamate [67, 122, 147].

Starting at rest, the post-synaptic membrane assumes a resting potential. An action potential arrives at the pre-synaptic membrane and depolarizes the membrane to open

voltage-gated Ca^{2+} channels. Ca^{2+} influx into the cell triggers exocytosis of small synaptic vesicles that contain neurotransmitters. These molecules couple with ion channels, e.g., Na^{+} channels, and other receptors that activate secondary messengers in the post-synaptic membrane. Ligand-receptor binding at the post-synaptic membrane depolarizes the membrane to generate an action potential. As the action potential at the post-synaptic membrane travels away from the synapse, its resting potential is restored.

To regulate synaptic signaling, neurotransmitter receptors are also present on the pre-synaptic membrane that may either inhibit or enhance exocytosis of synaptic vesicles. At the same time, released neurotransmitters may be (a) (re-)taken up by transport proteins on the pre-synaptic terminal membrane, e.g., dopamine, norepinephrine, and glutamate; (b) degraded, e.g., acetylcholine, or (c) taken up by neighboring glial cells, e.g., glutamate. Synaptic vesicles are recycled at the post-synaptic membrane by endocytosis.

Membrane potentials at the pre- and post-synaptic cells can be measured in terms of voltage. In this system, pre-synaptic membrane potential is the input and post-synaptic membrane potential is the output. The post-synaptic membrane potential depends on neurotransmitters, released at the pre-synaptic membrane, which diffuse across the synapse to bind to receptors on the post-synaptic membrane. The post-synaptic membrane potential does not affect the pre-synaptic membrane potential. Thus, synaptic signaling is regulated based on open-loop control, where the output does not inform the input.

1.3.1.2 Closed-loop control in glycolysis

Glycolysis is the primary pathway that breaks down glucose to synthesize adenosine triphosphate (ATP), the energy currency of the cell. The enzyme phosphofructokinase (PFK) is a key regulator of this pathway (Figure 3). In particular, PFK catalyzes the first committed step in glycolysis, i.e., the irreversible conversion of fructose-6-phosphate (F6P) into fructose-1,6,-bisphosphate (FBP) [21, 120].

At high levels of ATP, i.e., high energy levels in the cell, ATP allosterically inhibits the enzyme PFK by lowering the binding affinity of PFK for its substrate F6P at the catalytic site. Where energy (in the form of ATP) is spent, i.e., at low energy levels in the cell, ATP is

converted to adenosine monophosphate (AMP). AMP reverses the inhibition of ATP on the enzyme PFK. In other words, AMP enhances the activity of the enzyme PFK to increase glycolysis, so as to increase the production of ATP.

Amounts of intracellular F6P and ATP/AMP can be measured. In this system, F6P is the input and ATP/AMP are the output. The production of ATP/AMP depends on glycolysis, where PFK catalyzes the first committed step of the pathway. Critically, PFK activity is moderated by ATP/AMP. Thus, glycolysis is regulated based on closed-loop control, where the output informs the input.

1.3.1.3 A theory of stability

To aid the study of robustness in metabolic systems, a theory of stability is undoubtedly helpful to further reduce the complexity of homeostasis in terms of underlying principles. In particular, the Lyapunov theory of stability [90, 102] is especially useful because it provides: (a) an analytical definition of stability, and (b) a method of controller design that ensures system stability and consequent robustness. In this dissertation, this method of controller design, also known as the direct method of Lyapunov, is used to develop and apply a comparator model to infer potential feedback in a highly regulated metabolic pathway.

For the time-invariant case, the Lyapunov theory of stability is as follows:

Lyapunov stability theorem. *Let* $\vec{x} = \vec{0}$ *be an equlibrium point for* $\dot{\vec{x}}(t) = f(\vec{x}(t))$, *where* $\vec{x}(0) = \vec{0}$, $x \in \Re^n$, f *Lipschitz continuous in* $\mathcal{D} \subset \Re^n$, *and* $\mathcal{D}$ *contains the origin. Suppose that* $V(\vec{x}(t)) \in \mathbf{C}^1$ *positive definite in* $\mathcal{D}$ *such that* $\dot{V}(\vec{x}(t)) \leq 0$, $\forall \vec{x}(t) \in \mathcal{D}$, *then the equilibrium point is stable. If* $\dot{V}(\vec{x}(t)) < 0$, $\forall \vec{x}(t) \in \mathcal{D} \backslash \vec{0}$, *then the equilibrium point is asymptotically stable.*

A major outcome of the Lyapunov theory of stability is the provision of *sufficient* conditions to determine the stability of the origin, e.g., (0, 0) in 2-dimensional space, of complex systems. By specifying a candidate function, i.e., the candidate stability function (or *Lyapunov* function), that satisfies these conditions, a controller may be designed that ensures system stability without the need to solve accompanying differential equations that describe the system dynamics. This is known as the Lyapunov direct method, or the second method

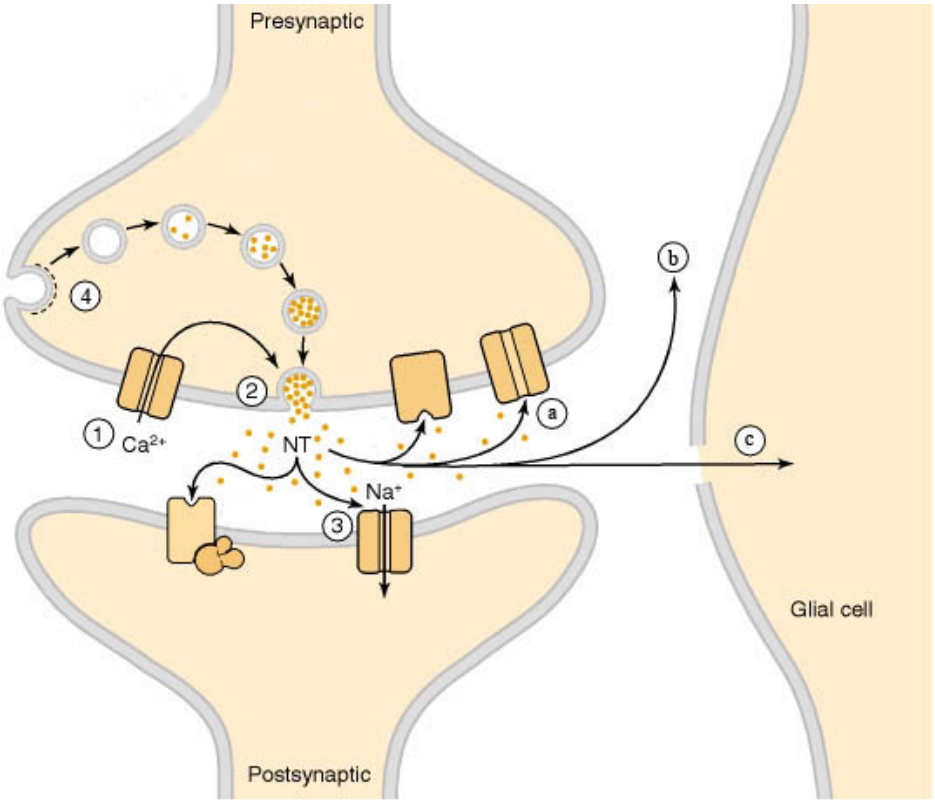


Figure 2: Synaptic transmission: an example of homeostasis in biochemical processes using open-loop control - [1] Ca^{2+} enters to trigger [2] exocytosis and release neurotransmitters (NT) that [3] bind to ion channels, e.g., Na^{+}, and other receptors in the post-synaptic terminal, [4] synaptic vesicles are recycled by endocytosis; released neurotransmitters may be [a] (re-)taken up by transport proteins, [b] degraded, [c] taken up by neighboring glial cells. (Image modified from source: Holz and Fisher [67] ©1999 American Society for Neurochemistry, National Center for Biotechnology Information (NCBI) Bookshelf)

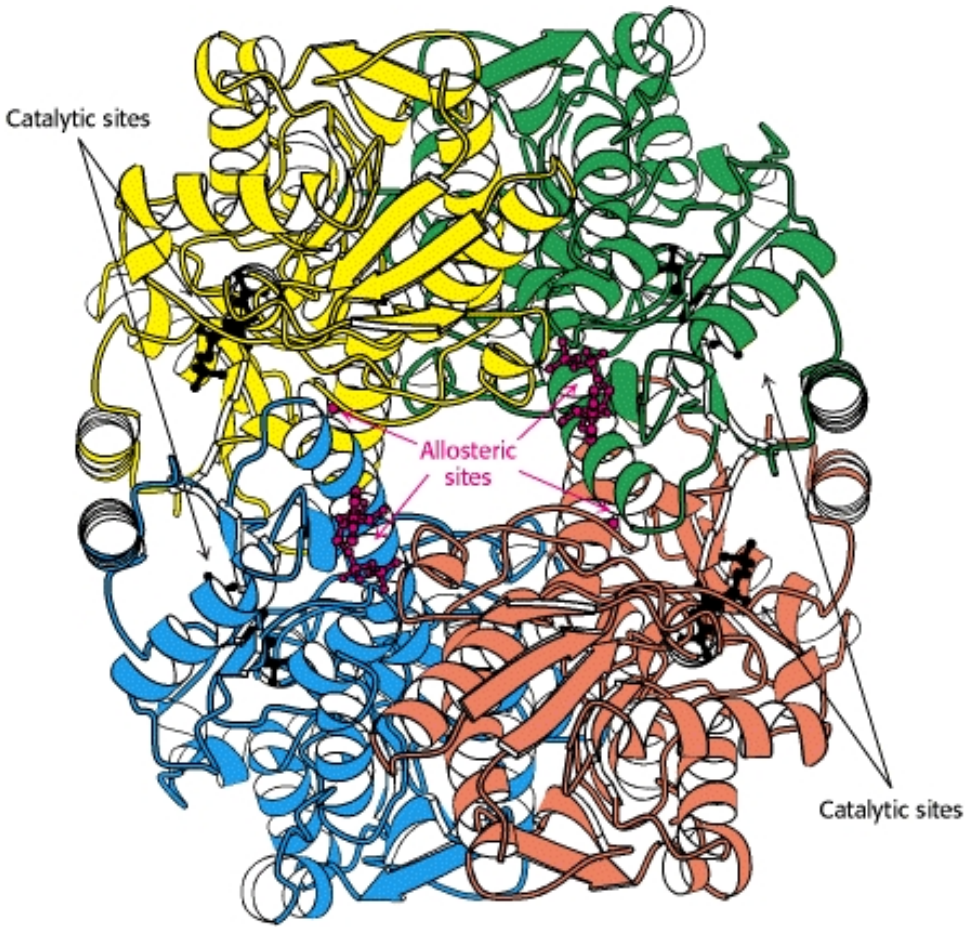


Figure 3: Structure of phosphofructokinase (PFK): this allosteric enzyme is a tetramer of four identical subunits and is the key regulator of glycolysis. PFK catalyzes the conversion of fructose-6-phosphate (F6P) to fructose-1,6,-bisphosphate (FBP). Adenosine triphosphate (ATP) inhibits, while adenosine monophosphate (AMP) enhances, PFK enzymatic activity by binding at the allosteric sites. (Image source: Berg *et al* [21] ©2002 W. H. Freeman and Company, National Center for Biotechnology Information (NCBI) Bookshelf)

of Lyapunov. For specific cases, the definition of candidate Lyapunov functions is left to the user.

The Lyapunov direct method is commonly used to design stabilizing controllers in dynamic systems. To do this, the control objective can be posed as a problem of stability of motion, where the origin represents desired (stable) steady-state dynamics. Then, a controller or steering vector $\vec{u}(t)$ may be specified with respect to the plant, so as to guide the system dynamics towards the origin. The key is to determine the dynamics of the controller based on the choice of an appropriate candidate Lyapunov function. In these cases, where controllers need not be unique, the choice of one controller over another may also be informed by other practical considerations, e.g., the cost of implementation.

In terms of regulation in metabolic systems, homeostasis may be interpreted as a steering vector in a complex system. At the same time, because the Lyapunov theory of stability provides only *sufficient* but not *necessary* conditions, the uniqueness of potential mechanisms that may be responsible for homeostasis in metabolic pathways requires additional experimental evidence to verify. When pathway redundancy in metabolic systems is also considered, it is likely that multiple modes of control, i.e., different enzymatic reactions, may exist to serve the same purpose.

1.3.2 Classical and modern control approaches

Classical control approaches deal mainly with single-input/single-output (SISO) systems in the frequency domain. The transfer function, i.e., function of the (scalar) system output in terms of the input, is computed from observations of frequency responses. Then based on the transfer function, system stability and performance is analyzed in terms of the roots of the characteristic equation, or 'poles' and 'zeros', using graphical methods. Feedback as a result of classical control is specified in terms of the system output, i.e., *closing the loop* based on open-loop behavior. Common modes of feedback in classical control are described in terms of proportional, integral, and derivative (PID) control.

In general, current models of homeostasis in metabolic systems are developed in terms of classical control and deal only with relatively simple cell behaviors that are well-studied

in the literature. This is because the key issue in applying control theory to model homeostasis in metabolic systems is the difficulty in observing pathway dynamics with sufficient temporal resolution, i.e., to measure metabolite amounts quickly enough. Classical control approaches, developed to work with SISO systems, involve less variables and parameters that require less data for abstraction.

In addition, because homeostatic mechanisms that may be responsible for control in metabolic systems are difficult to verify experimentally, current studies on homeostasis in metabolic systems are also focused on cases where extensive knowledge of the relevant biology is available *a priori*. Furthermore, in many cases, public literature is the only feasible resource to support any reasonable speculation on potential homeostatic mechanisms.

For example, bacterial chemotaxis in *Escherichia coli* is an instance of relatively simple homeostatic behavior that may be described in terms of classical control. Exact adaptation in bacterial chemotaxis is observed to be robust and not affected by changes in protein levels [7, 16]. Furthermore, such behavior may be described as the result of integral control that ensures convergence to steady state without error [177]. Using a classical control approach, details of the relevant transfer function can also be illustrated using dynamic input/output measurements, which provides insight into possible mechanisms that enable the robust behavior [141].

A second example deals with the regulation of heat shock response, also in *Escherichia coli*, which may be simplified as a feedback model for analysis using a number of potential feedback designs to study the costs and benefits of mounting the response with specific homeostatic mechanisms [44]. These examples illustrate classical control approaches to model relatively simple homeostatic behavior that involves analyzing single loops in a closed system.

Other examples of classical control models of homeostasis involve analyzing multiple loops in closed systems, e.g., tryptophan regulation and energy metabolism. The tryptophan system in *Escherichia coli* may be considered as three processes in series, i.e., transcription, translation, and synthesis, that involve multiple feedback loops [22, 163]. Energy metabolism in terms of glucose and ATP is also complex behavior that may be described in

terms of closed-loop regulation by various subsystems [32]. Moving on from single-variable analysis, a recent study of the adaptive response in membrane channels is based on using two variables, instead of one, to describe the dynamics of a 3-state model of membrane channel kinetics [50].

Thus, given single-variable data, classical control approaches are adequate to analyze and model simplified SISO models of metabolic systems. However, for metabolic systems, it is still difficult to interpret system dynamics and feedback in the frequency domain. Where high-throughput -omic technology is available such that there is sufficient data to support multi-input/multi-output (MIMO) models of metabolic systems, modern control approaches may prove to be more suitable to reverse engineer homeostasis in metabolic systems.

Given multi-variable data, modern control approaches were developed to handle MIMO systems, i.e., in terms of the dynamics of multiple internal states of a system (or *state-space*). To do this, the state-space is represented in terms of the rate of change of said states in the time-domain. Then, system stability may be analyzed in terms of stability theorems, usually based on optimizing specific stability functions (as in the case of the Lyapunov theory of stability). As a result of this approach, system feedback can be specified in terms of the state-space, which may prove useful to identify key controllable states. Common modes of feedback in modern control include adaptive, optimal, and robust control. A detailed comparison between classical and modern control approaches is shown in Table 1.

Thus, depending on the research objective, a modern control approach may prove to be more suitable than a classical control approach to reverse engineer homeostasis in metabolic systems using high-throughput -omic data, particularly in terms of:

- **handling many variables simultaneously** - modern control approaches handle vector systems using a state-space description but classical control approaches handle scalar systems,

- **representing time domain dynamics** - time rate of change of pathway dynamics is intuitive to biologists but transfer function 'poles' and 'zeros' in the frequency domain do not correspond directly to biological variables,

- **interpreting system feedback** - state-space feedback subject only to considerations for system stability in modern control are open to interpretation but PID feedback effectively constrains how pathway metabolites may interact in a homeostatic system.

Table 1: Comparison of classical vs. modern control approaches

	Classical control	Modern control
Data	Single variable	High-throughput (omic)
Variables	Input/output (scalar)	Internal states (vector)
Systems	Single-input/single-output (SISO)	Multi-input/multi-output (MIMO)
Representation	Frequency domain Transfer functions Frequency signals, e.g., impulse/step, are not intuitive and usually difficult to implement/manipulate in biology	Time domain State-space models Dynamics based on first principles, e.g., biochemical kinetics, are intuitive and appeal to current understanding
Stability	Roots of the characteristic equation, i.e., 'poles' and 'zeros' of the transfer function	Lyapunov stability theory, i.e., generalized concept of energy
Methods	Graphical analysis in the complex plane, e.g., root-locus	Numerical/computational analysis, e.g., matrix algebra
Feedback	Proportional, integral, derivative (PID) control, in terms of system output	Adaptive, robust, optimal control, in terms of internal states of a system

1.4 Thesis statement

In this present age of -omic biotechnology, modern control approaches are useful to model complex behaviors in biological systems, such as homeostasis in metabolic pathways, using high-throughput data. Although they can be useful in other ways, classical control approaches undermine the vast volumes, and wealth, of -omic data that are increasingly commonplace in biology. Thus, using -omic data, a modern control approach is suitable to reverse engineer homeostasis in metabolic systems.

In the rest of this dissertation, from an engineering and a modern control perspective, I describe the development and application of a comparator to reverse engineer homeostasis in a highly regulated metabolic pathway using mass spectrometry lipidomic time-series data. The metabolic pathway in question is the C16:0 sphingolipid *de novo* biosynthesis pathway, studied in a human embryonic kidney (HEK) cell line, where the effect of single-gene overexpression on homeostasis in sphingolipid *de novo* biosynthesis is of interest. Precisely, in these single-gene overexpressed (or treated) cells, the gene that codes for serine palmitoyltransferase (SPT), an enzyme in the *de novo* biosynthesis pathway, is overexpressed. The treated cells are assumed to approach the same steady-state dynamics as the wild type over time. Consequently, the outcome of the comparator is to: (a) model the differences between the treated and wild type cells, and (b) predict feedback in treated cells as a result of homeostasis.

In Chapter 2, I review the relevant sphingolipid biology and biomedical significance of sphingolipid *de novo* biosynthesis; and present experimental materials and methods as well as the development and application of the comparator model in terms of concept, assumptions, mathematical representation, and numerical implementation. In Chapter 3, I report the results of the comparator model in terms of *in silico* simulation; verify and interpret these results from a biological perspective; and demonstrate the generality of the comparator model with respect to additional data (on C26:0 sphingolipids). In Chapter 4, I discuss the biomedical applications and limitations of the proposed comparator model, and suggest what will be needed in the experimental data for the proposed model to be applied more fruitfully in the fields of biology and biomedical engineering.

From a broader perspective, it is also the goal of this thesis research to contribute to "the use of existing techniques in well-developed areas of control theory to analyze problems of interest to biologists" [146]. I strive to achieve this objective by proposing and developing the *first* use and application of a comparator model as a viable, albeit presently crude, tool for analysis in traditional case/control studies in biology.

CHAPTER II

MATERIALS AND METHODS: SPHINGOLIPID DE NOVO BIOSYNTHESIS

Using a modern control approach, a comparator model is developed to simulate and predict regulatory feedback in a case study metabolic pathway. Specifically, the comparator model is applied to a case study in the effect of single-gene overexpression on sphingolipid *de novo* biosynthesis in a human embryonic kidney (HEK) cell line. From the perspective of interdisciplinary research in particular, the comparator model appeals to the nature of conventional experiment design in the field of biology in terms of traditional case/control studies.

2.1 Sphingolipid biology

Sphingolipids are involved in key eukaryotic cell functions such as membrane structure recognition, signal transduction, intracellular regulation, and cell-cell interaction [109]. The general structure of a sphingolipid comprises a sphingoid base backbone, e.g., sphinganine, that is modified by addition of long-chain fatty acids, double bonds in the sphingoid base (to make sphingosine), and polar headgroups (Figure 4). When variation in these components is considered, sphingolipids are one of the most complex families of biomolecules [45, 123].

Long-chain fatty acids and their roles in sphingolipid biosynthesis, metabolism and function are of specific interest, e.g., in ceramide biosynthesis [88], maintenance of fatty acid levels [31], bilayer membrane composition in terms of lipid rafts [107, 52] and sterol affinity [117], mitochondrial permeability [116], and apoptotic signaling [139]. The metabolism of these sphingolipids is also studied in various cells, e.g., yeast [38], plants [12], mouse skin [104], and human lung endothelium [106].

De novo biosynthesis of sphingolipids with varying fatty acid chain lengths, e.g., C16:0

and C26:0, occurs via similar pathways. In particular, they share a common pool of palmitate and sphingoid bases. However, these chain-length variants affect cell functions differently. This suggests that biosynthesis of such chain-length variants is also regulated differently.

Alterations in sphingolipid biosynthesis, storage, and metabolism are implicated in human diseases [72]. For example, in synthesis, infantile-onset symptomatic epilepsy is described as a genetic defect in ganglioside GM3 synthase [144]; in storage, Tay-Sachs, Sandhoff, and Gaucher disease are characterized by glycosphingolipid accumulation [86]; in metabolism, ceramide functions as a tumor suppressor in various cancer cells [118]. Regulation of sphingolipid levels through *de novo* bio-synthesis is critical to maintain key cell functions, failing which leads to a variety of human diseases. By studying the dynamics and regulation of sphingolipid *de novo* biosynthesis, it may be possible to learn more about onset and development of these sphingolipid diseases.

2.1.1 Sphingolipid *de novo* biosynthesis

Sphingolipid *de novo* biosynthesis is a "necessary, but dangerous, pathway" [109]. *De novo* biosynthesis is necessary because sphingoid bases present in food are mostly degraded in the intestine [164]. Regulation of sphingolipid amounts through *de novo* biosynthesis is important because pathway intermediates such as ceramide and sphingomyelin are highly bioactive.

Three of the initial enzymes in the sphingolipid *de novo* biosynthesis pathway (Figure 5) are thought to be particularly important:

- *serine palmitoyltransferase* (SPT), which catalyzes the initial step of sphingolipid biosynthesis by condensation of L-serine and palmitoyl-coA (Pal-CoA),
- *(dihydro)ceramide synthase*, which adds the amide-linked fatty acid to form dihydroceramide, and
- *dihydroceramide desaturase*, which adds the 4,5-trans-double bond of the sphingoid base backbone to form ceramide.

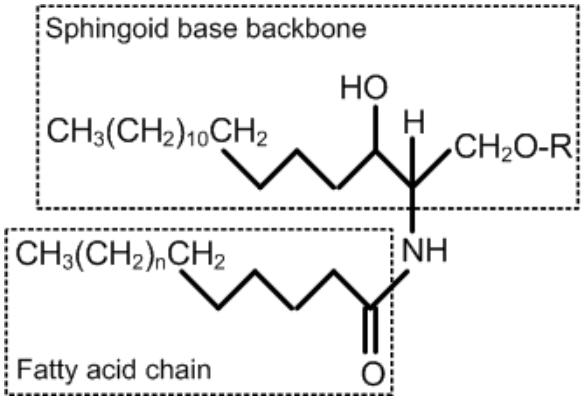


Figure 4: General molecular structure of sphingolipids with sphingoid base backbone, variable-length fatty acid chains, and headgroups (R).

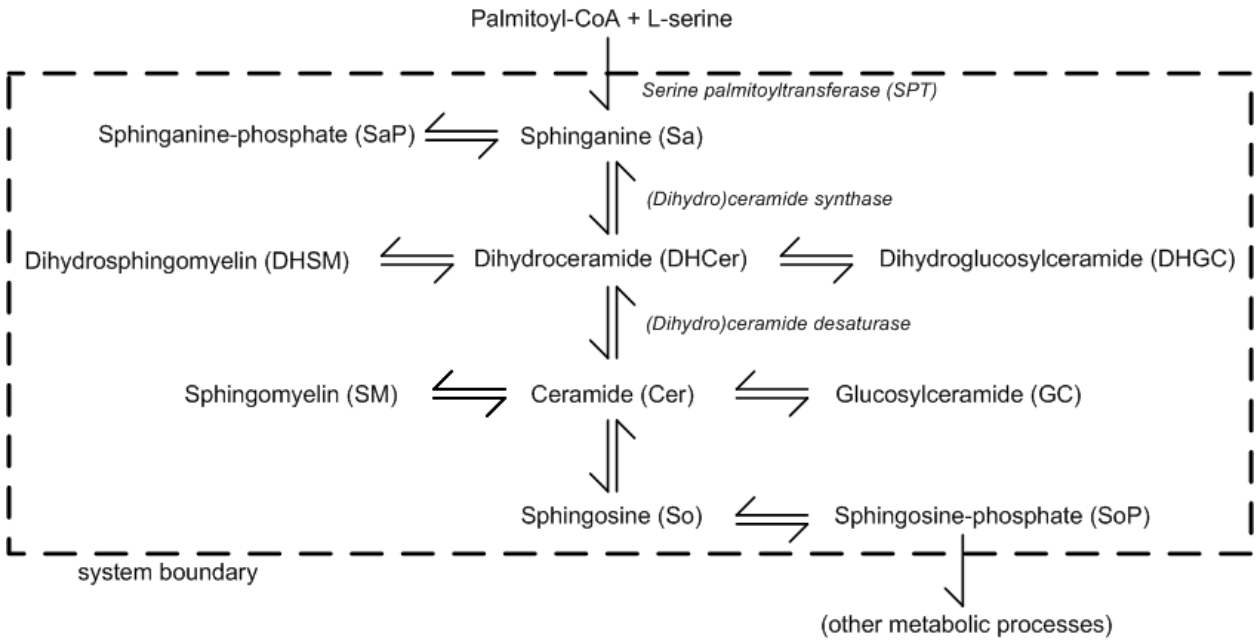


Figure 5: Sphingolipid *de novo* biosynthesis pathway, with key enzymes in *italics*; uni-/bi-directional arrows indicate irreversible/reversible reactions respectively; pathway intermediates, e.g., ceramide, sphingomyelin, are highly bioactive.

2.1.1.1 Sphingolipid-omics

The increasing wealth of quantitative lipidomic data makes the study of sphingolipid biology from a systems perspective promising and challenging at the same time [112]. The diversity and complexity of sphingolipids require researchers to possess a range of tools to work with these compounds. For example, to study dynamic changes in lipid amounts, computational methods that couple mass spectrometry with statistical algorithms to analyze the vast number of lipid species from cellular extracts are required [47]. In addition, the Lipidomics Gateway (www.lipidmaps.org) is a comprehensive website from the LIPID MAPS consortium, the leading group for lipidomics research, which contains standards, data, and tools for researchers interested in lipid biology.

2.1.2 Mass spectrometry lipidomics

Sphingolipids, i.e., sphingoid bases and (dihydro)N-acyl species, are extracted from samples of cultured human embryonic kidney (HEK) cells and quantified by liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) as described previously [110, 152]. For this case study, measurements from four technical replicates are taken at seven time points in uniform hourly intervals, i.e., at 0, 1, 2, 3, 4, 5, and 6 hours. The community standard is described by the LIPID MAPS consortium, which requires three technical replicates taken at eight time points, i.e., at 0, 0.5, 1, 2, 4, 8, 12, and 24 hours.

The data acquisition protocol for this case study is comparable to the community standard for time-series studies in terms of both (a) number of time points sampled, and (b) number of technical replicates. In addition, compared to the community standard, the data for this case study is also sampled more frequently at the onset of the experiment.

To quantify sphingoid bases labeled with [U-13C]-palmitate, additional multiple reaction monitoring (MRM) pairs corresponding to [M+16] precursor ions and [M+16] product ions are used. To quantify (dihydro)N-acyl sphingolipids labeled with a single [U-13C]-palmitate, MRM pairs corresponding to [M+16] precursor ions and both [M+0] and [M+16] product ions are used, providing discrimination of labeling on the sphingoid base and N-acyl moieties. These [M+16] isotopomers are designated BASE and FA respectively. To quantify

(dihydro)N-acyl sphingolipids labeled with two [U-13C]-palmitate molecules, MRM pairs corresponding to [M+32] precursor ions and [M+16] product ions are used; the [M+32] isotopomers are designated DUAL. Unlabeled sphingolipids are designated 12C. The peak areas of unlabeled and labeled sphingolipid isotopomers are integrated, converted to picomoles using the peak area of the internal standard, and normalized to the mg of protein in the extracted sample.

2.2 Comparator model

From engineering control, a *comparator* takes in 2 inputs: one each from the *reference* and the *plant* systems (Figure 6). In general, the reference system exhibits, and establishes, the desired response to an *input*. Under different scenarios, the input could be either known, e.g., a tracking signal, or unknown, e.g., a disturbance to the system. For the purpose of modeling, regardless of whether it is known or unknown, the input is common to both the reference and plant systems. In other words, the reference system is also a *model* for the plant system. Thus, the reference system is fully known while the plant system is unknown (fully or partially) and is therefore the system of interest.

The two inputs to the comparator are the outputs from the reference and plant systems. At each time, the difference between the two given inputs is calculated and used to compute an output by the comparator, which is feedback to the plant system at the next time. This process proceeds iteratively until the plant dynamics converges with the reference dynamics, i.e., when the difference between the outputs from the reference and plant systems is zero (Figure 4).

Thus, the key to convergence between the reference and plant systems lies in how the comparator output is: (a) computed, and (b) implemented, i.e., as actuated feedback to the plant system. As discussed before in Section 1.3.1, the Lyapunov theory of stability provides a direct method to specify how to accomplish the former based on a user-defined stability function.

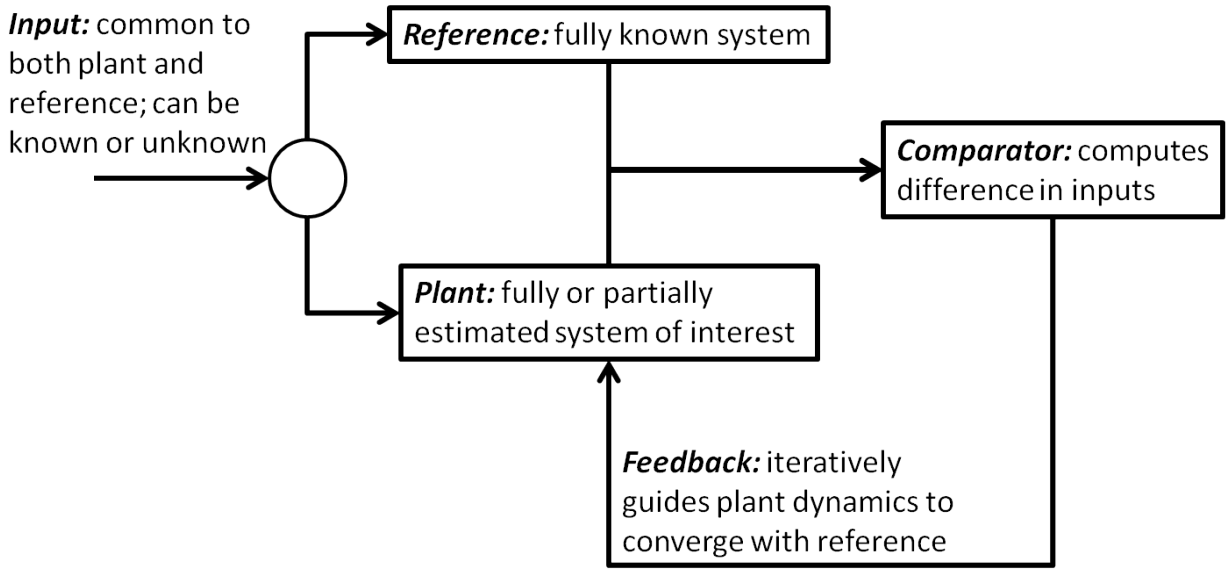


Figure 6: Block diagram of comparator model: the *comparator* determines the difference in input from the *reference* and *plant* systems, which is feedback to the plant. This process proceeds iteratively such that the plant dynamics converge to the reference dynamics, i.e., the difference between the plant and reference outputs are zero. Both the reference and plant systems received a common *input*.

2.2.1 Assumptions

The key assumptions are as follows:

2.2.1.1 Treatment effect on system stability

In this case study, the wild type dynamics is chosen as the *reference* system, which the treated cells, i.e., SPT overexpressed, follow to the same stable steady state. More precisely, pathway dynamics in treated cells is assumed to converge to the wild type over time, i.e., the difference between the two pathways, in terms of measurable metabolite amounts, goes to zero.

Thus, the underlying assumption here is that the effect of the treatment condition, i.e., single-gene (SPT) overexpression, is assumed to be large enough such that pathway dynamics in treated cells can be differentiated from the wild type; however, at the same time, it is small enough such that the treated cells do not converge to a second steady state that is different from the wild type. In other words, homeostasis remains relevant under these experimental conditions. This assumption is supported in theory by the pathway topology [171] and in practice from experimental data.

2.2.1.2 Determination of pathway differences

The comparator is a quantitative model of how deviations from the "desired" (stable) state (or response), represented by the wild type in this case, may be handled under homeostasis. The issue is that at any time, while the "desired" state may be observed in wild type cells, only the perturbed state is observed in the treated cells. In other words, without communicating with the wild type in real-time, it is implicitly assumed that the treated cells have knowledge of the "desired" state *a priori*.

Thus, the underlying assumption here is that a "desired" state is already prescribed within the cells, treated or wild type. Such a "desired" state could possibly be genetically predetermined so that if the cell encounters perturbation within certain homeostatic limits, such deviations from the "desired" state could be regulated to restore the system to stability.

2.2.2 Relevance to case/control studies in biology

In this case study, a comparator model is developed and applied to model the effect of single-gene (SPT) overexpression on sphingolipid *de novo* biosynthesis in human embryonic kidney (HEK) cells in terms of pathway regulation under homeostasis. In particular, as a result of homeostasis, treated cells are assumed to regulate sphingolipid *de novo* biosynthesis such that levels of sphingolipid metabolites in these cells approach the wild type over time. Stable isotope 13C-labeled palmitate is added to both treated and wild type cells in excess relative to intracellular sphingolipid metabolites to facilitate tracking the levels of these sphingolipid metabolites by mass spectrometry.

In other words, treated cells are represented as the *plant* system and the wild type as the *reference* system. Sphingolipid precursor, extracellular palmitate is the *input* that is common to both reference and plant systems. Then, by analyzing the dynamics of feedback gain from the *comparator* as a predictive model of homeostasis, interactions between sphingolipid metabolites in the system may be responsible to regulate pathway dynamics (Figure 7).

Thus, the comparator model appeals to, and indeed complements, traditional experiment design in the field of biology in that it provides a ready research paradigm for the analysis of homeostatic mechanisms using conventional case/control experiments, which correspond intuitively to the engineering language of plant/reference systems. Effectively, the proposed comparator model demonstrates that complexities of homeostasis can be reduced by borrowing well-established concepts from control theory to describe equally well-documented phenomena in biology. Consequently, so long as the experimental treatment does not result in responses that are beyond the reference system, the comparator model together with an appropriate theory of stability may be an additional analysis tool that is useful to reverse engineer homeostasis and contribute to research in biochemical robustness.

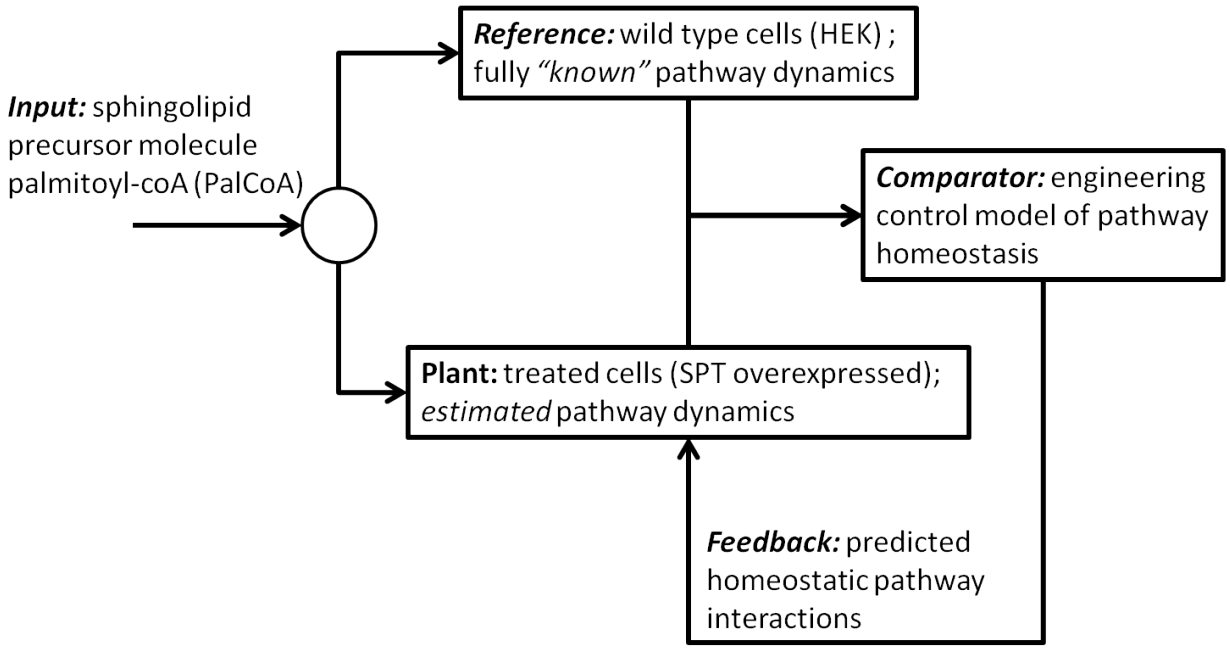


Figure 7: Proposed application of comparator model in a biological case/control study: in this case study on sphingolipid *de novo* biosynthesis, a comparator model from modern engineering control is applied to reverse engineer pathway homeostasis in treated cells (*plant*), i.e., single-gene (SPT) overexpressing cells; the pathway dynamics in the wild type are fully known (*reference*); to both treated and wild type cells, stable isotope 13C-labeled sphingolipid precursor palmitate (*input*) is added in excess relative to intracellular sphingolipid levels, which also faciliate mass spectrometry lipidomics; the key model result is predicted steady-state *feedback* to the treated cells that suggests potential homeostatic pathway interactions as a result of experimental treatment.

2.3 Mathematical representation

2.3.1 State-space model

There are 10 sphingolipid metabolites of interest in the *de novo* biosynthesis pathway based on the available data. They are:

1. sphinganine (Sa)
2. sphinganine phosphate (SaP)
3. dihydroceramide (DHCer)
4. dihydroglucosylceramide (DHGC)
5. dihydrosphingomyelin (DHSM)
6. ceramide (Cer)
7. glucosylceramide (GC)
8. sphingomyelin (SM)
9. sphingosine (So)
10. sphingosine phosphate (SoP)

As discussed in Section 2.1.2, extracellular palmitate, a precursor molecule that is labeled using stable isotope 13C, is added to facilitate the measurement of these sphingolipid metabolites using mass spectrometry.

Thus, these sphingolipid metabolites in the *de novo* biosynthesis pathway represent the internal states of the pathway system while the sphingolipid precursor molecule, palmitate, is the input that is common to both wild type and treated cells. Then, the state-space $\vec{x}(t)$ can be defined as the set of intracellular sphingolipid metabolite amounts where

$$\vec{x}(t) = [Sa(t), SaP(t),$$
$$DHCer(t), DHGC(t), DHSM(t),$$
$$Cer(t), GC(t), SM(t),$$
$$So(t), SoP(t)]^T \tag{6}$$

The dynamics of this system can be described in a simplified form using generalized mass action (GMA) [58, 99] and written as a system of linear ordinary differential equations (ODEs) as follows:

$$\dot{Sa}(t) = k_{in,Sa}PalCoA + k_{SaP,Sa}SaP + k_{DHCer,Sa}DHCer$$
$$- (k_{Sa,SaP} + k_{Sa,DHCer})Sa \tag{7}$$
$$\dot{SaP}(t) = k_{Sa,SaP}Sa - k_{SaP,Sa}SaP \tag{8}$$
$$\dot{DHCer}(t) = k_{Sa,DHCer}Sa + k_{DHGC,DHCer}DHGC$$
$$+ k_{DHSM,DHCer}DHSM + k_{Cer,DHCer}Cer$$
$$- (k_{DHCer,Sa} + k_{DHCer,DHGC} + k_{DHCer,DHSM} + k_{DHCer,Cer})DHCer \tag{9}$$
$$\dot{DHGC}(t) = k_{DHCer,DHGC}DHCer - k_{DHGC,DHCer}DHGC \tag{10}$$
$$\dot{DHSM}(t) = k_{DHCer,DHSM}DHCer - k_{DHSM,DHCer}DHSM \tag{11}$$
$$\dot{Cer}(t) = k_{DHCer,Cer}DHCer + k_{GC,Cer}GC + k_{SM,Cer}SM + k_{So,Cer}So$$
$$- (k_{Cer,DHCer} + k_{Cer,GC} + k_{Cer,SM} + k_{Cer,So})Cer \tag{12}$$
$$\dot{GC}(t) = k_{Cer,GC}Cer - k_{GC,Cer}GC \tag{13}$$
$$\dot{SM}(t) = k_{Cer,SM}Cer - k_{SM,Cer}SM \tag{14}$$
$$\dot{So}(t) = k_{Cer,So}Cer + k_{SoP,So}SoP - (k_{So,Cer} + k_{So,SoP})So \tag{15}$$
$$\dot{SoP}(t) = k_{So,SoP}So - (k_{SoP,So} + k_{SoP,out})SoP \tag{16}$$

where $k_{a,b}$ indicates the reaction rate constant for the enzymatic reaction going from a to b. In addition, because extracellular palmitate is added in excess relative to intracellular

amounts of sphingolipid metabolites, the amount of PalCoA as the system input is assumed to be constant and in excess for the length of the experiment *in vitro* and corresponding simulation *in silico*. On a related note, the Michaelis-Menten model of enzyme reaction kinetics [165] is not applied here because of the difficulties in estimating the requisite parameters [11, 34].

2.3.2 Comparator model

The comparator model is derived from the Lyapunov theory of stability based on scalar-input to a multi-variable system. As follows in this section, variables and parameters that deal with the *reference* system are denoted by subscript $_{ref}$.

2.3.2.1 Plant and reference systems

The state-space model for the *reference* system, i.e., wild type cells with (assumed) fully known processes, is:

$$\dot{\vec{x}}_{ref}(t) = \mathbf{A}_{ref}\vec{x}_{ref}(t) + \vec{b}_{ref}r(t) \tag{17}$$

for the internal states (sphingolipid metabolites) $\vec{x}_{ref}(t)$ of the *reference* system (wild type cells), where dynamics between these states (enzymatic reactions between sphingolipid metabolites) can be described in terms of linear ordinary differential equations (ODEs) with parameters (reaction rate constants $k_{a,b}$) contained in the matrix $\mathbf{A}_{ref}$.

The scalar system *input* (sphingolipid precursor palmitate, which is converted to intracellular palmitoyl-coA PalCoA) $r(t)$ can be scaled and targeted at various internal states by the actuation vector $\vec{b}_{ref}$; in this case, it is involved only in the first enzymatic reaction of the pathway.

The state-space model for the *plant* system, i.e., treated cells with fully or partially unknown processes that are of interest, is:

$$\dot{\vec{x}}(t) = \mathbf{A}\vec{x}(t) + \vec{b}u(t) \tag{18}$$

for the internal states (sphingolipid metabolites) $\vec{x}(t)$ of the *plant* system (treated cells),

where dynamics between these states (enzymatic reactions between sphingolipid metabolites) are described in terms of linear ODEs with parameters (reaction rate constants $k_{a,b}$) contained in the matrix $\mathbf{A}$.

The scalar *feedback* (from PalCoA and other system sphingolipid metabolites) $u(t)$ regulates the dynamics of these various states in the plant through the actuation vector $\vec{b}$. More precisely, the feedback $u(t)$ is the sum of two components such that:

$$u(t) = u_x(t) + u_r(t) = \vec{h}_x(t)^T \vec{x}(t) + h_r(t)r(t) \tag{19}$$

where $u_x(t)$, $u_r(t)$ are specific to the internal states and system input respectively and are modulated by feedback gain $\vec{h}_x(t)$, $h_r(t)$.

2.3.2.2 Derivation from Lyapunov stability

Let $\vec{e}(t)$ be the difference between *plant* and *reference* systems:

$$\vec{e}(t) \equiv \vec{x}(t) - \vec{x}_{ref}(t) \tag{20}$$

such that

$$\begin{aligned}
\vec{e}(t) &= \vec{x}(t) - \vec{x}_{ref}(t) \\
&\Rightarrow \\
\dot{\vec{e}}(t) &= \dot{\vec{x}}(t) - \dot{\vec{x}}_{ref}(t) \\
&= [\mathbf{A}\vec{x}(t) + \vec{b}u(t)] - [\mathbf{A}_{ref}\vec{x}_{ref}(t) + \vec{b}_{ref}r(t)] \\
&= [\mathbf{A}\vec{x}(t) + \vec{b}(\vec{h}_x(t)^T\vec{x}(t) + h_r(t)r(t))] - [\mathbf{A}_{ref}\vec{x}_{ref}(t) + \vec{b}_{ref}r(t)] \\
&= [\mathbf{A} + \vec{b}\vec{h}_x(t)^T]\vec{x}(t) - \mathbf{A}_{ref}\vec{x}_{ref}(t) + [\vec{b}h_r(t) - \vec{b}_{ref}]r(t) \\
&= [\mathbf{A} + \vec{b}\vec{h}_x^{*T} - \vec{b}\vec{h}_x^{*T} + \vec{b}\vec{h}_x(t)^T]\vec{x}(t) - \mathbf{A}_{ref}\vec{x}_{ref}(t) + [\vec{b}h_r(t) - \vec{b}h_r^* + \vec{b}h_r^* - \vec{b}_{ref}]r(t) \\
&= \mathbf{A}_{ref}[\vec{x}(t) - \vec{x}_m(t)] + \vec{b}[\vec{h}_x(t)^T - \vec{h}_x^{*T}]\vec{x}(t) + \vec{b}[h_r(t) - h_r^*]r(t) \\
&= \mathbf{A}_{ref}\vec{e}(t) + \vec{b}[\Delta\vec{h}_x(t)^T\vec{x}(t) + \Delta h_r(t)r(t)]
\end{aligned} \tag{21}$$

where

$$
\begin{aligned}
\mathbf{A} + \vec{b}\vec{h}_x^{*T} &= \mathbf{A}_m \\
\vec{b}h_r^* &= \vec{b}_m \\
\Delta\vec{h}_x(t) &= \vec{h}_x(t) - \vec{h}_x^* && \Rightarrow \Delta\dot{\vec{h}}_x(t) = \dot{\vec{h}}_x(t) \\
\Delta h_r(t) &= h_r(t) - h_r^* && \Rightarrow \Delta\dot{h}_r(t) = \dot{h}_r(t)
\end{aligned}
$$

and $\vec{h}_x^*$, h_r^* are the ideal steady-state feedback gains.

Then choose a candidate stability function, i.e., positive definite $V(\cdot)$:

$$
V(\vec{e}(t), \Delta\vec{h}_x(t), \Delta h_r(t)) = \vec{e}(t)^T\mathbf{P}\vec{e}(t) + \Delta\vec{h}_x(t)^T\Gamma_x^{-1}\Delta\vec{h}_x(t) - \gamma_r^{-1}\Delta h_r(t)^2 \tag{22}
$$

and take its derivative with respect to time as follows:

$$
\begin{aligned}
\dot{V}(t) &= \vec{e}(t)^T[\mathbf{A}_m^T\mathbf{P} + \mathbf{P}\mathbf{A}_m]\vec{e}(t) + 2\vec{e}(t)\mathbf{P}\vec{b}[\Delta\vec{h}_x(t)^T\vec{x}(t) + \Delta h_r(t)r(t)] \\
&\quad + 2\Delta\vec{h}_x(t)^T\Gamma_x^{-1}\Delta\dot{\vec{h}}_x(t) \\
&\quad + 2\Delta h_r(t)\gamma_r^{-1}\Delta\dot{h}_r(t) \\
&= \vec{e}(t)^T\mathbf{Q}\vec{e}(t) \\
&\quad + 2\Delta\vec{h}_x(t)^T[\vec{x}(t)\vec{e}(t)^T\mathbf{P}\vec{b} + \Gamma_x^{-1}\Delta\dot{\vec{h}}_x(t)] \\
&\quad + 2\Delta h_r(t)[r(t)\vec{e}(t)^T\mathbf{P}\vec{b} + \gamma_r^{-1}\Delta\dot{h}_r(t)]
\end{aligned} \tag{23}
$$

Thus, to ensure $\dot{V}(t)$ is negative semi-definite that guarantees asymptotic stability, i.e.,

$$
\dot{V}(t) = -\vec{e}(t)^T\mathbf{Q}\vec{e}(t) \leq 0 \tag{24}
$$

$\dot{\vec{h}}_x(t)$ and $\dot{h}_r(t)$, which are the rates of change of feedback gain, must be:

$$
\dot{\vec{h}}_x(t) = -\Gamma_x\vec{x}(t)\vec{e}(t)^T\mathbf{P}\vec{b} \tag{25}
$$

$$
\dot{h}_r(t) = -\gamma_r r(t)\vec{e}(t)^T\mathbf{P}\vec{b} \tag{26}
$$

where Γ_x, γ_x are also user-defined. Then, $\mathbf{P}$ is positive definite symmetric and satisfies the Lyapunov equation:

$$\mathbf{A}_m^T\mathbf{P} + \mathbf{P}\mathbf{A}_m + \mathbf{Q} = \vec{0} \tag{27}$$

where $\mathbf{Q}$ is also positive definite symmetric. $\mathbf{P}$ can be solved by specifying $\mathbf{Q}$ in addition to $\mathbf{A}_{ref}$ from equation (1). Finally, the initial conditions $\vec{h}_x(0)$, $h_r(0)$ may be chosen by matching:

$$\mathbf{A} + \vec{b}\vec{h}_x(0)^T = \mathbf{A}_{ref} \tag{28}$$

$$\vec{b}h_r(0) = \vec{b}_{ref} \tag{29}$$

to approximate the ideal steady-state feedback gains $\vec{h}_x^*$, h_r^*.

2.4 Implementation

2.4.1 Parameter estimation

2.4.1.1 State-space parameters

Reaction rate constants $k_{a,b}$ in the state-space model were estimated using a genetic algorithm to optimize a modified least-squares error between *in silico* pathway dynamics and corresponding experimental data [140]. Only C16:0 DUAL species data from both treated and wild type cell data were used. While additional information may be inferred from each of the labeled/unlabeled combinations of BASE, FA, or 12C species, they are not used here for simplification.

Inputs to the estimation routine include measured amounts of all sphingolipid metabolites determined from 4 technical replicates. Only 5 of 7 data points, from 0 to 4 hours, were used as training data for parameter estimation; the last 2 data points were used to verify estimation results. Furthermore, robustness of *in silico* model results, i.e., dynamics of feedback gain, is also tested by sensitivity analysis of the model results to perturbations in these reaction rate constants, where $k_{perturbed} = k_{original} * 10^{-1}, 10^{1}$.

2.4.1.2 Comparator parameters

The system input is extracellular palmitate that is in excess amounts relative to the pathway sphingolipid metabolites. Extracellular palmitate is converted to intracellular palmitoyl-coA (PalCoA). Subsequently, $r(t)$ represents the rate at which PalCoA is converted to sphinganine (Sa), i.e., $k_{in,Sa}PalCoA$, the first enzymatic reaction of the sphingolipid *de novo* biosynthesis pathway. Thus, $r(t)t$ is a boundary condition that is a scalar constant, numerical value 100.

The reference system actuation vector $\vec{b}_{ref}$ supplies only the common system input $r(t)$. The plant system actuation vector $\vec{b}$ supplies the control input in terms of two components: state feedback $u_x(t)$, and system input $u_r(t)$. Numerical values for $\vec{b}$ are specified through iteration, subject to non-negative model results, so as to minimize the least-squares error between *in silico* model results and *in vitro* experimental data.

Γ_x, γ_r, and $\mathbf{Q}$ are parameters of the candidate stability function, subject to conditions of

the Lyapunov theory of stability. These parameters affect the rate at which the plant system approaches (stable) steady state, and in practice, i.e., for controller design, may be tuned to achieve specific performance objectives. Here, as a first attempt to model homeostasis in a highly regulated metabolic pathway, these parameters are selected non-specifically.

Γ_x and γ_r correspond to the internal states of the pathway system $x(t)$ and the system input $r(t)$ respectively. Starting from an initial value of the identity matrix $\mathbf{I}$ for Γ_x and unity for γ_r, these parameters are tuned iteratively, by one order of magnitude to 0.1, to alleviate the stiffness problem for numerical integration [126].

$\mathbf{Q}$ is specified as the identity matrix $\mathbf{I}$. Then, based on the Lyapunov equation, $\mathbf{A}_{ref}^T\mathbf{P} + \mathbf{P}\mathbf{A}_{ref} + \mathbf{Q} = \vec{0}$, $\mathbf{P}$ can be solved in MATLAB using the built-in function `lyap()` once $\mathbf{Q}$ and $\mathbf{A}_{ref}$ are also specified.

$h_r(0)$ and $\vec{h}_x(0)$ represent initial conditions for feedback gains that correspond to the system input $r(t)$ and state-space $x(t)$. The numerical values are specified as 1 for $h_r(0)$ and $\vec{0}$ for $\vec{h}_x(0)$. Based on the candidate stability function subject to the Lyapunov theory of stability, the rate of change of these feedback gains is a function of the state-space difference $\vec{e}(t)$ between the reference and plant systems. In theory, initial conditions for the feedback gains affect only the *rate* of convergence; however in practice, if these initial conditions are poorly chosen, then convergence may not be achieved at all.

2.4.2 Modeling error

The modeling error is measured in terms of the root-sum-square (RSS) quantity between *in silico* model results and *in vitro* experimental data, which is defined as:

$$RSS = \sqrt{\sum_{t=t_0}^{t_f} (x_{insilico,t} - x_{invitro,t})^2} \tag{30}$$

where $x_{insilico,t}$ represents *in silico* model results at time t and $x_{invitro,t}$ represents *in vitro* experimental data at time t, where $t = 1, 2, 3, 4, 5, 6$.

2.4.3 Numerical details

The aforementioned mathematical representation is implemented in MATLAB using the built-in function `ode15s()` as follows:

- $\mathbf{A}, \mathbf{A}_{ref}, \Gamma_x, \mathbf{P}, \mathbf{Q} \in \Re^{10x10}$
- $\vec{x}(t), \vec{x}_{ref}(t), \vec{e}(t), \vec{h}_x(t), \vec{b}, \vec{b}_{ref} \in \Re^{10}$
- $u(t), r(t), h_r(t), \gamma_r \in \Re$

The proposed comparator model is first applied to data on the C16:0 sphingolipid *de novo* biosynthesis pathway. Then, to verify the generality of the proposed model, it is subsequently applied to data on the C26:0 sphingolipid pathway. While the two sphingolipid pathways are similar in terms of *de novo* biosynthesis, they are independent systems because:

- different enzymes are responsible to catalyze enzymatic reactions in each pathway system – a *family* of (dihydro)ceramide synthases (CerS1, CerS2, , CerS6) are responsible for dihydroceramide (DHCer) synthesis from sphinganine (Sa); in particular, CerS3 and CerS4 produce DHCer with very long chain fatty acids that are greater than or equal to 20 carbon atoms long,
- C16:0 and C26:0 sphingolipid metabolites are quantified using mass spectrometry independently,
- parameters of the state-space models, i.e., reaction rate constants, are estimated independently.

In reality, the underlying biological picture is likely to be much more complex. For instance, there may be cross-signaling events between the two pathways simply because sphingolipid metabolites are localized in the same regions on cellular membranes. Nonetheless in this case study, because such cross-signaling events in *de novo* biosynthesis between sphingolipid metabolites that are significantly different in fatty acid chain length is not yet well documented, modeling such events is not yet an issue.

2.4.3.1 C16:0 pathway

Parameters of the C16:0 pathway state-space model, i.e., reaction rate constants $k_{a,b}$, are available in Appendix A. All other parameters are defined as follows:

$$
\begin{aligned}
\texttt{input:} \quad & r(t) = 100 (= k_{in,Sa} PalCoA) \\
\texttt{actuation:} \quad & \vec{b}_{ref} = [1, 0, \ldots, 0]^T, \\
& \vec{b} = [10, -0.1, -10, 10, 10, -0.1, -0.1, -0.1, -0.1, 10]^T \\
\texttt{adaptation:} \quad & \mathbf{Q} = \mathbf{I}, \qquad \Gamma_x = 0.1\mathbf{I}, \qquad \gamma_r = 0.1 \\
\texttt{initial conditions:} \quad & h_r(0) = 1, \\
& \vec{h}_x(0) = [0, \ldots, 0]^T
\end{aligned}
$$

2.4.3.2 C26:0 pathway

Parameters of the C26:0 pathway state-space model, i.e., reaction rate constants $k_{a,b}$, are available in Appendix B. All other parameters are defined as follows:

$$
\begin{aligned}
\texttt{input:} \quad & r(t) = 100 (= k_{in,Sa} PalCoA) \\
\texttt{actuation:} \quad & \vec{b}_{ref} = [1, 0, \ldots, 0]^T, \\
& \vec{b} = [10, -0.1, -10, 10, 10, -0.1, -0.1, -0.1, -0.1, 10]^T \\
\texttt{adaptation:} \quad & \mathbf{Q} = \mathbf{I}, \qquad \Gamma_x = 0.1\mathbf{I}, \qquad \gamma_r = 0.1 \\
\texttt{initial conditions:} \quad & h_r(0) = 1, \\
& \vec{h}_x(0) = [0, \ldots, 0]^T
\end{aligned}
$$

CHAPTER III

RESULTS: PATHWAY DYNAMICS AND FEEDBACK

In this chapter, results of the proposed comparator model to reverse engineer homeostasis in a highly regulated metabolic pathway are reported as follows:

- to test (a) the efficacy of the proposed comparator model in capturing observed pathway dynamics, and (b) the robustness of predicted steady-state pathway feedback in response to simulated errors in state-space parameter estimation,
- to interpret and verify predicted steady-state pathway feedback in the context of homeostasis in C16:0 sphingolipid *de novo* biosynthesis, by comparison to literature as well as independent data on measured changes in sphingolipid metabolite levels in response to 4HPR dosage in both treated and wild type cells, and
- to demonstrate the generality of the proposed modeling approach, by applying the comparator model to a different metabolic system using independent data on C26:0 sphingolipid metabolites.

Here, to facilitate discussion of the model results, the 10 sphingolipid metabolites of interest in the *de novo* biosynthesis pathway can be grouped as follows:

- Sa* - Sa and SaP,
- So* - So and SoP,
- DH* - DHCer, DHGC, and DHSM,
- Cer* - Cer, GC, and SM.

where Sa*, So* are sphingoid bases, DH* are dihydrosphingolipids, and Cer* are sphingolipids.

3.1 Model efficacy and robustness

3.1.1 Pathway dynamics

In this subsection, the model results are reported in terms of: (a) state-space parameter fold change between treated and wild type cells, (b) root-sum-square (RSS) error between *in silico* model results and *in vitro* experimental data in wild type cells, i.e., *reference* system, (b) RSS error between *in silico* model results and *in vitro* experimental data in treated cells, i.e., *plant* system, *without* comparator, and (d) RSS error between *in silico* model results and *in vitro* experimental data in treated cells *with* comparator.

3.1.1.1 State-space parameter fold change between treated and wild type cells

Each state-space parameter represents the reaction rate constant of a specific enzymatic reaction in the pathway. Appendix A lists the estimated state-space parameters for the C16:0 sphingolipid biosynthesis pathway. As such, parameter fold change between comparable systems is often used as a preliminary indicator of enzymatic reactions that may proceed at noticeably different rates in one system than the other, i.e., in treated cells compared to the wild type, which may be biologically interesting.

Fold change in reaction rate constants between treated and wild type cells is reported in Table 2. Reaction rate constants with observed log_{10} fold change greater than magnitude "1" may be considered to be "biologically significant". In addition, based on these values, Figure 10 shows the log_{10} fold change of these reaction rate constants between treated and wild type cells. Fold change is "significant" in 9 of the 18 reaction rate constants in all.

The 9 reaction rate constants with potentially "biologically significant" fold change between treated and wild type cells are:

fold change $< 10^{-2}$ $k_{Sa,DHCer}$, $k_{GC,Cer}$,

$10^{-2} <$ **fold change** $< 10^{-1}$ $k_{DHCer,Sa}$, $k_{DHCer,DHGC}$, $k_{DHCer,Cer}$, $k_{Cer,So}$, $k_{So,Cer}$, and

$10^{1} <$ **fold change** $< 10^{2}$ $k_{Sa,SaP}$, $k_{DHGC,DHCer}$.

Table 2: Parameter fold change: C16:0 reaction rate constants

$k_{SaP,Sa}$	-0.1811
$k_{DHCer,Sa}$	-1.2927
$k_{Sa,SaP}$	1.3986
$k_{Sa,DHCer}$	-2.8536
$k_{DHGC,DHCer}$	1.3260
$k_{DHSM,DHCer}$	-0.1277
$k_{Cer,DHCer}$	0.5080
$k_{DHCer,DHGC}$	-1.5468
$k_{DHCer,DHSM}$	0.4930
$k_{DHCer,Cer}$	-1.1099
$k_{GC,Cer}$	-3.7469
$k_{SM,Cer}$	0.1851
$k_{So,Cer}$	-1.6749
$k_{Cer,GC}$	0.0072
$k_{Cer,SM}$	-0.6839
$k_{Cer,So}$	-1.2724
$k_{SoP,So}$	-0.1414
$k_{So,SoP}$	1.4213

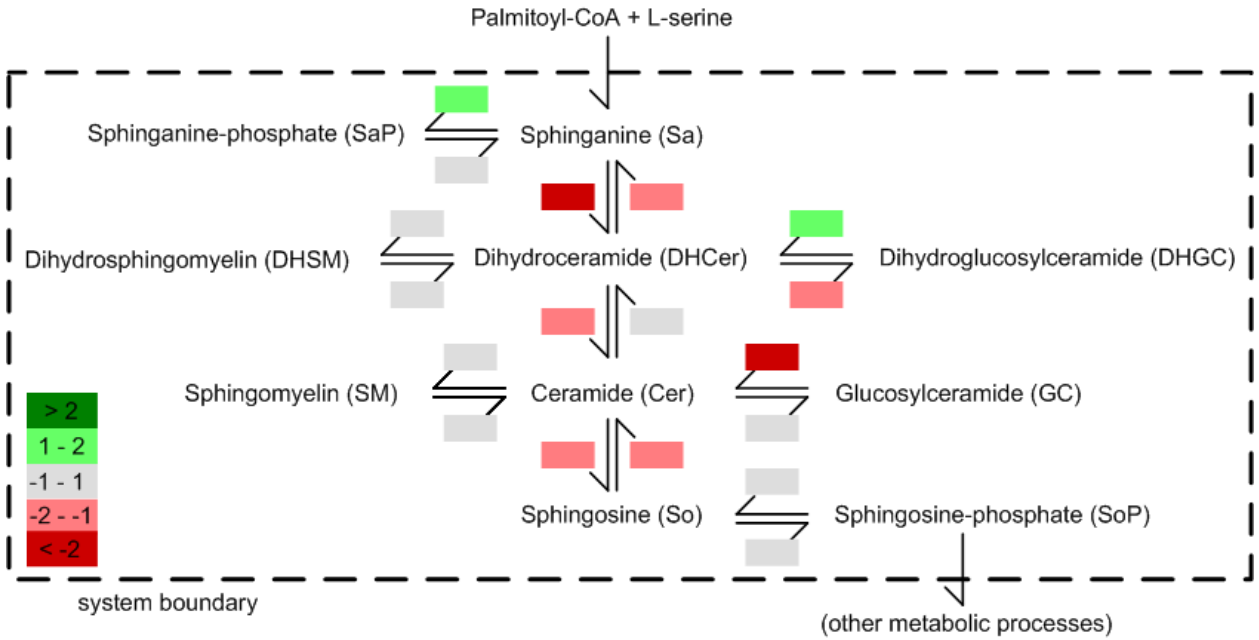


Figure 8: Parameter fold change between treated cells (*plant*) over wild type (*reference*): colors indicate range of fold change in terms of logarithmic ratio log_{10} of estimated reaction rate constants of C16:0 sphingolipid *de novo* biosynthesis pathway.

*3.1.1.2 Wild type cells (*reference *system)*

In the following series of time-series graphs of model results, the scatterplots represent median value and range of 4 technical replicates (where available) of *in vitro* experimental data in discrete time ($t = 0, 1, 2, ..., 6$) and broken lines represent *in silico* model results in terms of sphingolipid metabolite amounts.

Figure 8 shows the experimental data and model results for pathway dynamics in wild type cells in terms of sphingolipid metabolite amounts using the generalized mass action (GMA) state-space model. Experimental data was not available for Sa*.

For So*, the median data suggests that So increases quickly at first until time $t = 2h$ and approaches a steady state around $t = 4h$. At the same time, measurements for So are less consistent compared to other molecules and have a much larger range. The median values for SoP over time shows that SoP remains low throughout. This observation is supported by its range of measurements, which is consistently small except for at time $t = 6h$.

For DH*, technical replicates were not available and the measurements fluctuate significantly throughout the course of experiment. As such, it is not possible to estimate apparent trends in how the amounts of these molecules change over time.

For Cer*, experimental data is highly consistent across all replicates as seen from the small range of measurements. Thus, the data shows that Cer increases quickly at first till $t = 1h$ and approaches a steady state around $t = 4h$; GC increases very slightly throughout $t = 0h$ to $t = 6h$, and SM increases more quickly before $t = 3h$ compared to after $t = 4h$.

Comparing the model results with experimental data qualitatively, the state-space model of pathway dynamics based on GMA is a good fit to the experimental data for So*, Cer, and GC (4 metabolite species), where they fall clearly and consistently within the range of measurements for these metabolites. For SM, the model result reflects similar trends, i.e., an increase over time but at a much larger rate *in silico* than is measured *in vitro*. However for DH*, the model results are a poor fit to the experimental data, although even to the trained eye, it may not be straightforward to discern an apparent trend in pathway dynamics for these sphingolipid metabolites given the limited data.

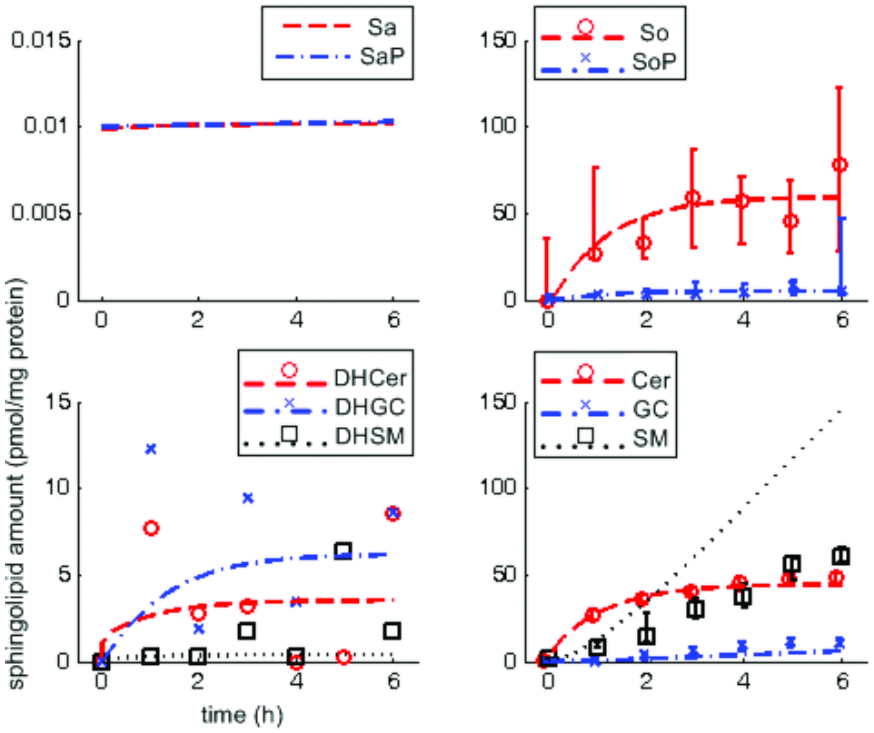


Figure 9: Experimental data and model results for C16:0 sphingolipid metabolites in wild type cells (*reference* system): abscissa - time (hours), ordinate – experimentally normalized sphingolipid metabolite amounts (pmol/mg protein); scatterplots represent *in vitro* experimental data (median and range indicated where available), broken lines represent *in silico* model results.

*3.1.1.3 Treated cells (*plant *system) without comparator*

Figure 9 shows the experimental data and model results of pathway dynamics in treated cells in terms of sphingolipid metabolite amounts using the generalized mass action (GMA) state-space model *without* the comparator.

For Sa*, the median data suggests that Sa increases very quickly at first until time $t = 2h$, then decreases almost immediately towards a steady state at $t = 3h$. However, the rate of the initial increase may be overestimated given that the measurements for Sa at $t = 1, 2h$ are spread over a larger range compared to measurements at subsequent times. The data for SaP is more consistent throughout the course of the experiment, which shows that SaP increases only slightly over time.

For So*, the data suggests that So increases continually throughout the course of experiment, although the range of measurements for So is much larger at times $t = 5, 6h$ compared to the preceding time. The data for SoP is similar to that for SaP and is consistent throughout, which shows that SoP also increases only slightly over time.

Technical replicates were not available for DH*. However, the available data does suggest gradual changes over time. DHCer appears to increase quickly from $t = 1h$ to $t = 2h$, then decreases more gradually from $t = 3h$ to $t = 5h$. DHGC appears to remain unchanged throughout the course of experiment. DHSM appears to increase quickly from $t = 1h$ to $t = 2h$ before approaching a steady state subsequently.

For Cer*, the experimental data is highly consistent across all replicates for GC and SM, as can be seen from the small range of measurements, but not for Cer where the range of measurements for Cer is much larger at times $t = 2, 3h$ compared to at other times. The data suggests that Cer increases quickly from $t = 1h$ to $t = 2h$, then decreases gradually from $t = 2h$ to $t = 5h$. DHGC increases only very slightly till $t = 6h$. SM increases uniformly and at a slightly higher rate over the same period of time.

Comparing the model results with experimental data qualitatively, *in silico* sphingolipid metabolite amounts based on the GMA state-space model *without* comparator are a good fit to the data for 5 metabolite species (SaP, SoP, DHGC, GC, and SM), where they fall clearly and consistently within the range of measurements for these metabolites. However,

for 4 metabolite species (Sa, So, DHCer, DHSM, and Cer), the model results consistently underestimate the experimental data. Furthermore, for 2 species (DHCer and Cer), the model results do not reflect changes in the experimental data sufficiently, in particular where rapid increases in the measured data from $t = 2h$ to $t = 3h$ followed by more gradual decreases from $t = 1h$ to $t = 5h$ are not captured in the state-space model.

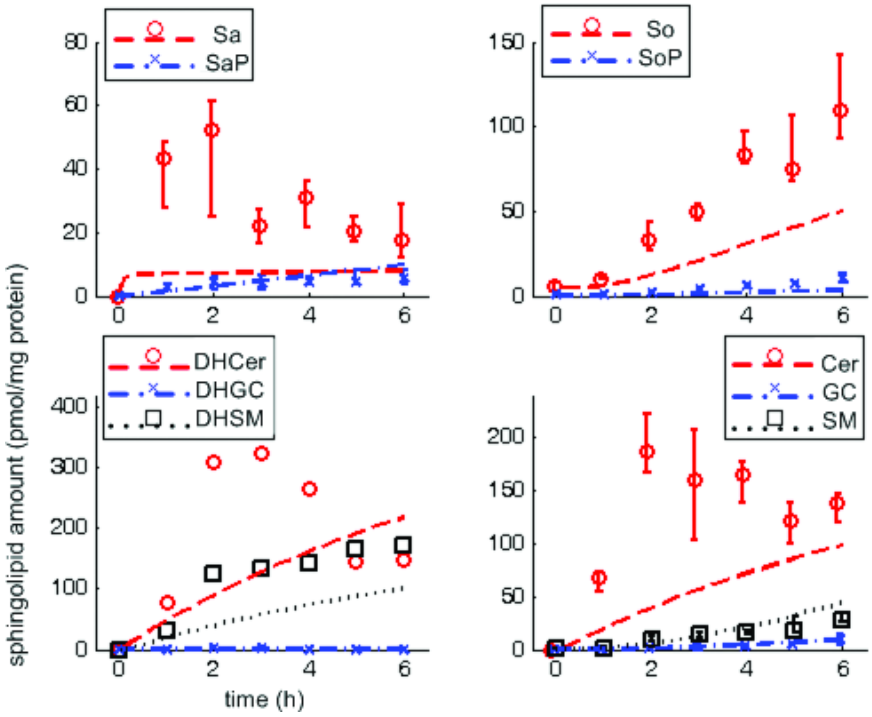


Figure 10: Experimental data and model results for C16:0 sphingolipid metabolites in treated cells (*plant* system) *without* comparator: abscissa - time (hours), ordinate – experimentally normalized sphingolipid metabolite amounts (pmol/mg protein); scatterplots represent *in vitro* experimental data (median and range indicated where available), broken lines represent *in silico* model results.

3.1.1.4 Root-sum-square (RSS) error in wild type and treated cells state-space model

Compared to model results for the wild type (Figure 8), model results for the treated cells (Figure 9) are not as good a fit to the experimental data. In particular, the model results for the treated cells are observed to have the poorest during intervals where significant changes are recorded in the measurement data. In terms of root-sum-square (RSS) error, the goodness-of-fit of the GMA state-space model of both wild type and treated cells *without* comparator is quantified and reported in Table 3. The RSS error is also equivalent to the L2-norm between model results and experimental data taken at discrete times, i.e., where $t = 0, 1, 2, ..., 6$. Taken together, these observations show that although the same algorithm was used for parameter estimation, i.e., a modified genetic algorithm described in Section 2.4.1, model results based on these estimated state-space parameters still differ in accuracy of matching experimental data for different datasets.

Table 3: State-space error: C16:0 wild type and treated cells *without* comparator

	Wild type	Treated cells ↯
C16:0 sphingolipids		
Sa	0.0002	61.4536
SaP	0.0004	5.5469
DHCer	6.9245	327.3739
DHGC	16.8494	3.0832
DHSM	6.0871	171.2229
Cer	4.7858	213.9073
GC	9.1085	41.2140
SM	88.0422	19.6755
So	20.2075	99.8462
SoP	2.4847	9.4587

Accuracy of state-space GMA model in terms of root-sum-square (RSS) error of *in silico* model results compared to *in vitro* experimental data in wild type and ↯ treated cells *without* comparator.

*3.1.1.5 Treated cells (*plant *system) with comparator*

Figure 10 shows the experimental data and model results of pathway dynamics in treated cells in terms of sphingolipid metabolite amounts using the generalized mass action (GMA) state-space model *with* the comparator. In particular, the model results are a good fit to the experimental data for 4 metabolite species (So, SoP, DHGC, and DHSM). The model results underestimate the measurement data for 1 metabolite species (Sa), and overestimate the data for 3 metabolite species (SaP, GC, and SM). For 2 metabolite species (DHCer and Cer), the model results do not capture the transient trend of the sphingolipid pathway dynamics from time $t = 1h$ to $t = 5h$.

In addition, to illustrate the effect of the comparator model to simulate state-space pathway dynamics in the treated cells, Table 5 shows the quantitative accuracy of the model results to experimental data in terms of comparison between the treated cells *with* and *without* comparator. It is noteworthy that the quantitative comparison shows that model accuracy in simulating state-space pathway dynamics is comparable in both the existing approach using GMA *without* comparator as in the proposed approach *with* comparator.

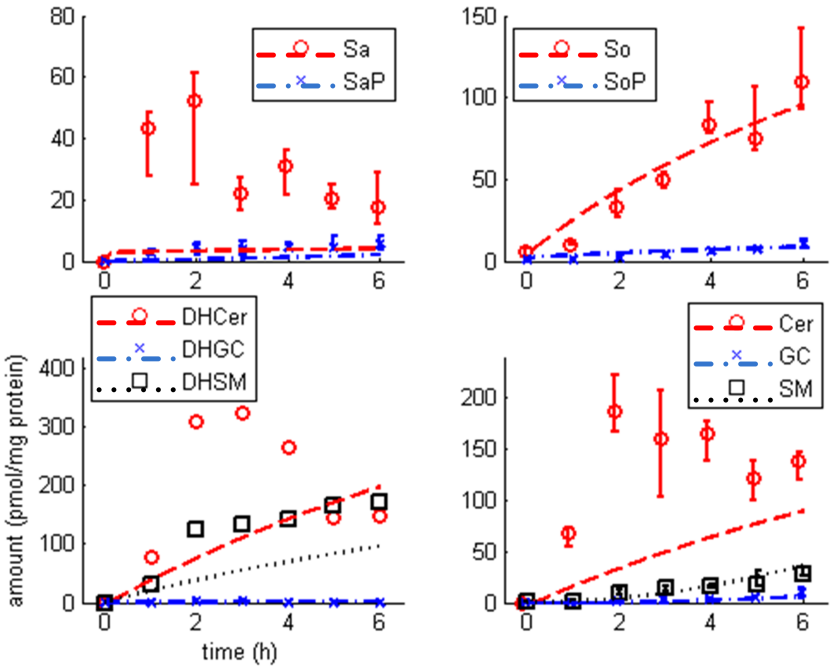


Figure 11: Experimental data and model results for C16:0 sphingolipid metabolites in treated cells (*plant* system) *with* comparator: abscissa - time (hours), ordinate – experimentally normalized sphingolipid metabolite amounts (pmol/mg protein); scatterplots represent *in vitro* experimental data (median and range indicated where available), broken lines represent *in silico* model results.

Table 4: State-space error: C16:0 treated cells *with* and *without* comparator

	Without comparator	*With* comparator
Treated cells only		
Sa	61.4536	70.1442
SaP	5.7469	8.2915
DHCer	327.3739	346.0480
DHGC	3.0832	2.9421
DHSM	171.2229	176.7805
Cer	213.9073	227.8527
GC	41.2140	46.6592
SM	19.6755	11.2155
So	99.8462	29.9782
SoP	9.4587	4.8419

3.1.2 Steady-state feedback

Figures 12 and 13 shows the key model outcome, i.e., predicted steady-state pathway feedback in terms of input and aggregate state feedback as well as individual sphingolipid metabolite feedback based on the proposed comparator model of homeostasis in C16:0 sphingolipid *de novo* biosynthesis in response to single-gene (SPT) overexpression in HEK cells.

Figure 12 shows the predicted steady-state pathway feedback in treated cells in terms of 2 components: (a) system input and sphingolipid precursor, palmitoyl-CoA (PalCoA), $u_r(t)$, and (b) aggregate state feedback from pathway sphingolipid metabolites $u_x(t)$. At steady-state, the magnitude of the feedback suggests the contribution of each component to homeostasis, i.e., a larger magnitude indicate greater contribution, and sign (+/-) indicates either enhancement (positive feedback) or inhibition (negative feedback) to the pathway. The range indicated at discrete time, i.e., $t = 1, 2, ..., 6$, is the result of sensitivity analysis of these model results in response to potential errors in state-space parameter estimation (as discussed in Section 2.4.1).

From the onset, i.e., at time $t = 0h$, $u_r(t)$ oscillates for a very brief period of time but with large peak-to-peak amplitude, then quickly decreases from the zero initial condition to settle at steady-state (approximately -20). On the other hand, throughout the course of simulation, $u_x(t)$ increases steadily from the zero initial condition to settle at steady-state (approximately +30). At steady-state, the magnitude of feedback in each of these components is similar but in different directions, i.e., the different signs suggest that input feedback $u_r(t)$ is inhibited while aggregate state feedback $u_x(t)$ is enhanced as a result of homeostasis.

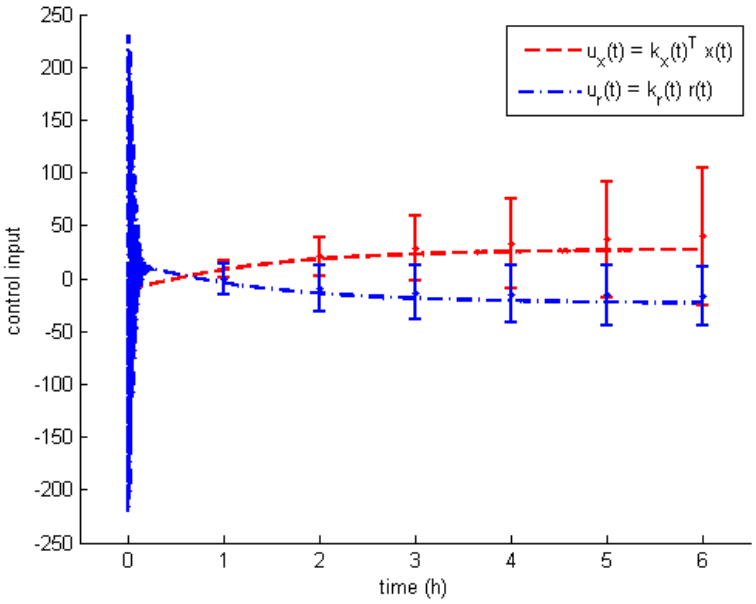


Figure 12: Predicted steady-state feedback for C16:0 sphingolipid metabolites pathway in treated cells in terms of: (a) system input and sphingolipid precursor, palmitoyl-coA (Pal-CoA), $u_r(t)$ (blue), and (b) aggregate state feedback from pathway sphingolipid metabolites $u_x(t)$ (red) - x-axis: time (hours), y-axis: magnitude (dimensionless), error bars: standard deviation from state-space parameter sensitivity analysis

Figure 13 shows the predicted steady-state feedback gain from individual pathway sphingolipid metabolites $h_x(t)$. These feedback gain are dimensionless; however, at steady-state, the magnitude and sign of these feedback values may be interpreted to suggest differential activity in pathway homeostasis. More precisely, the larger the magnitude, the greater the feedback; furthermore, the sign (+/-) of the gain values indicate either enhancement (positive feedback) or inhibition (negative feedback) to the pathway. In addition, the range indicated at discrete time, i.e., $t = 1, 2, ..., 6$, is the result of sensitivity analysis of these model results in response to potential errors in state-space parameter estimation (as discussed in Section 2.4.1).

Sa, SaP, So, and SoP are sphingoid bases; DHCer, DHGC, DHSM, Cer, GC, and SM are sphingolipids. Thus, from Figure 13:

(top left) Feedback gain for sphinganine (Sa) decreases quickly from the zero initial condition to settle at steady-state (approximately -1); there is no noticeable change in feedback gain for sphinganine-phosphate (SaP).

(top right) There is no noticeable change in feedback gain for sphingosine (So); however, feedback gain decreases quickly from the zero initial condition to settle at steady-state (approximately -1) for sphingosine-phosphate (SoP).

(bottom left) Feedback gain for dyhydroceramide (DHCer) increases quickly from the zero initial condition to settle at steady-state (approximately $+1$); on the other hand, it decreases quickly from the zero initial condition to settle at steady-state (approximately -1) for dihydroglucosylceramide (DHGC) and dihydrosphingomyelin (DHSM).

(bottom right) Feedback gain for ceramide (Cer) decreases gradually from the zero initial condition and approaches -0.07 at time $t = 6h$; there is no noticeable change from the zero initial condition for glucosylceramide (GC) and sphingomyelin (SM).

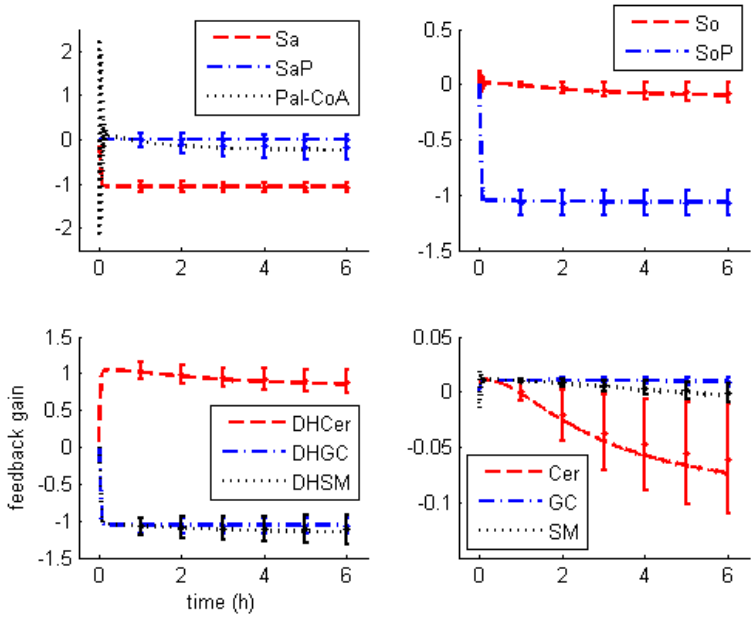


Figure 13: Predicted steady-state feedback for individual C16:0 sphingolipid metabolites in treated cells - x-axis: time (hours), y-axis: magnitude (dimensionless), error bars: standard deviation from state-space parameter sensitivity analysis

3.2 Biological interpretation and verification

3.2.1 Pathway input and aggregate state feedback

From Figure 12, these model results suggest that in response to single-gene (SPT) overexpression in treated cells, and when sphingolipid precursor PalCoA is added to the cells in excess amounts relative to intracellular sphingolipid metabolite levels, the metabolic effect of the pathway input (PalCoA) at steady-state is to inhibit sphingolipid *de novo* biosynthesis. On the other hand, the collective metabolic effect of pathway sphingolipid metabolites at steady-state is to enhance *de novo* biosynthesis. This observation suggests that, in this case of serine pamitolytransferase (SPT) upregulation (as a result of single-gene overexpression) where SPT is the enzyme that catalyzes for the initial pathway reaction of condensation between PalCoA and L-serine to form sphinganine, the excess input and reaction substrate PalCoA also inhibits the pathway dynamics.

This implication is intriguing in that in terms of controller design, such a regulatory strategy is also very specifically tuned to the type of system input. From a biochemical perspective, given that there may be only a limited number of admissible pathway inputs, such a regulatory strategy may be a viable option to maintain pathway homeostasis. For instance, there could be molecular mechanisms to slow the rate of extracellular palmitate uptake [54, 98, 105], reduce the rate of conversion from palmitate to PalCoA, or even increase the rate of other enzymatic reactions that require PalCoA as a substrate. Subsequently, given that PalCoA is an integral component of sphingolipid metabolites at least in terms of molecular structure, the implication that regulation of the sphingolipid *de novo* biosynthesis pathway may be highly dependent on intracellular PalCoA levels should not be too surprising.

The possibility that levels of intracellular PalCoA may be moderated as a result of SPT overexpression could also be interpreted as a result of experimental procedure, i.e., where only data on dual-chain isotope-labeled sphingolipid metabolites was used to develop this model (see Section 2.4.1 on parameter estimation). It could be that in reality, regardless of isotope-labeling, sphingolipid *de novo* biosynthesis and turnover is increased in SPT overexpressed cells, which leads to increased palmitate that is subsequently converted back

to PalCoA so that it is available to re-enter the pathway. Subsequently, because only data on dual-labeled sphingolipids is used for parameter estimation, it may appear that the amount of 13C-labeled PalCoA is moderated relative to the levels of all intracellular PalCoA, labeled or otherwise. This scenario may be tested experimentally by improving the tracking of various pools of labeled and unlabeled palmitate and PalCoA using mass spectrometry.

3.2.2 Individual sphingolipid metabolite feedback

Steady-state feedback gain in individual sphingolipid metabolites may be interpreted as follow: relative to other sphingolipid metabolites such as So and Cer, steady-state feedback gain from Sa, SoP, DHCer, DHGC, and DHSM is significantly different from the zero initial condition. This suggests that these sphingolipid metabolites may contribute more actively in pathway homeostasis. Furthermore, these particular sphingolipid metabolites also contribute to cell signaling. In particular, Sa and SoP are also cell signaling molecules that are reported to activate growth stimulation: Sa reverses growth inhibition [145] while SoP activates growth stimulation [148, 149]. In addition, So and Cer are also reported to activate growth inhibition [35, 111, 174].

Taken together in the context of pathway homeostasis, the similarities and differences in steady-state feedback gain for these sphingolipid signaling metabolites, which activate opposing responses, suggests that the experimental treatment, i.e., SPT overexpression at these levels, leads to growth stimulation in HEK cells [167]. Clinically, SPT overexpression is also implicated in cancer cell metastasis in human tumors [29].

Thus, in general, model results and interpretation from Figures 12 and 13 are consistent with current knowledge of sphingolipid biology (summarized in Table 5). The robustness of these model results, i.e., predicted steady-state feedback, in response to potential errors in parameter estimation underscores the utility and promise the proposed comparator model to simulate and predict homeostatic mechanisms in a highly regulated metabolic pathway. Furthermore, these model results and present interpretations raise rather new and perhaps interesting questions with regard to more specific aspects of homeostasis in sphingolipid *de novo* biosynthesis pathway that were not previously studied.

Table 5: Sphingolipid metabolite signaling: reported effects on cell growth

Sphingolipid metabolite	Reported effects on cell growth	Predicted steady-state feedback
Sphinganine (Sa)	Growth stimulation [145]	Strong inhibition
Sphinganine phosphate (SaP)		No change
Sphingosine (So)	Growth inhibition [35, 111, 174]	Weak inhibition
Sphingosine phosphate (SoP)	Growth stimulation [148, 149]	Strong inhibition
Dihydroceramide (DHCer)	Not available	Strong enhancement
Dihydroglucosylceramide (DHGC)	Not available	Strong inhibition
Dihydrosphingomyelin (DHSM)	Not available	Strong inhibition
Ceramide (Cer)	Growth inhibition [35, 111, 174]	Weak inhibition
Glucosylceramide (GC)	Not available	Weak enhancement
Sphingomyelin (SM)	Not available	Weak enhancement

3.2.3 Additional testing with $1\mu M$ 4HPR

4HPR (*4-hydroxyphenylretinamide*) is a synthetic retinoid that is under clinical evaluation as a therapeutic agent in a variety of cancers. It is reported to restrain tumor growth by inducing apoptosis. However, the required effective dosage to achieve toxicity differs significantly in SPT overexpressing cells and wild type HEK cells. In particular, a higher dose of 4HPR is required in SPT overexpressing cells compared to wild type HEK cells to achieve observable, or effective, cell toxicity (Figure 14).

In particular, although the mode of action, i.e., molecular mechanisms, of 4HPR has been linked to ceramide metabolism (Figure 15), its systemic effects are not fully understood. With respect to the three key enzymes involved in de novo sphingolipid synthesis, 4HPR is reported to enhance the activity of (a) SPT, i.e., serine palmitoyl-transferase that catalyzes the reaction of serine and palmitoyl-coA to form sphinganine (Sa), (b) (DH)CerS, i.e., (dihydro)ceramide synthase that catalyzes the formation of (dihydro)ceramide (DHCer) from sphinganine (Sa), and inhibits the activity of (c) DES, i.e., desaturase that catalyzes the formation of ceramide (Cer) from (dihydro)ceramide (DHCer). Subsequently, from a modeling perspective, the effect of 4HPR on sphingolipid *de novo* biosynthesis can be interpreted as a change in state-space parameters because the reaction rate constants are altered. In particular, 4HPR treatment is similar to SPT overexpression to the extent that SPT is overexpressed in both treatments.

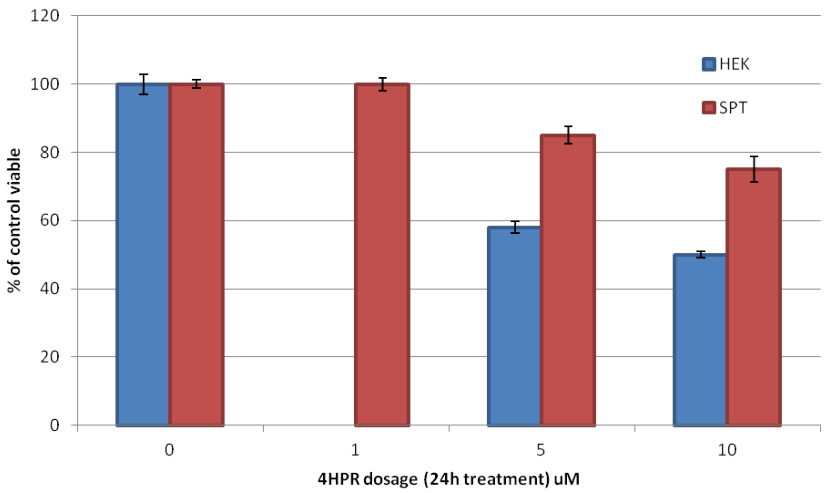


Figure 14: Effect of 4HPR on SPT and HEK cell viability – larger dose of 4HPR is required to achieve effective toxicity in SPT cells compared to wild type HEK cells: y-axis –number of viable cells (% of cells without 4HPR, experimental data), x-axis – 4HPR dosage

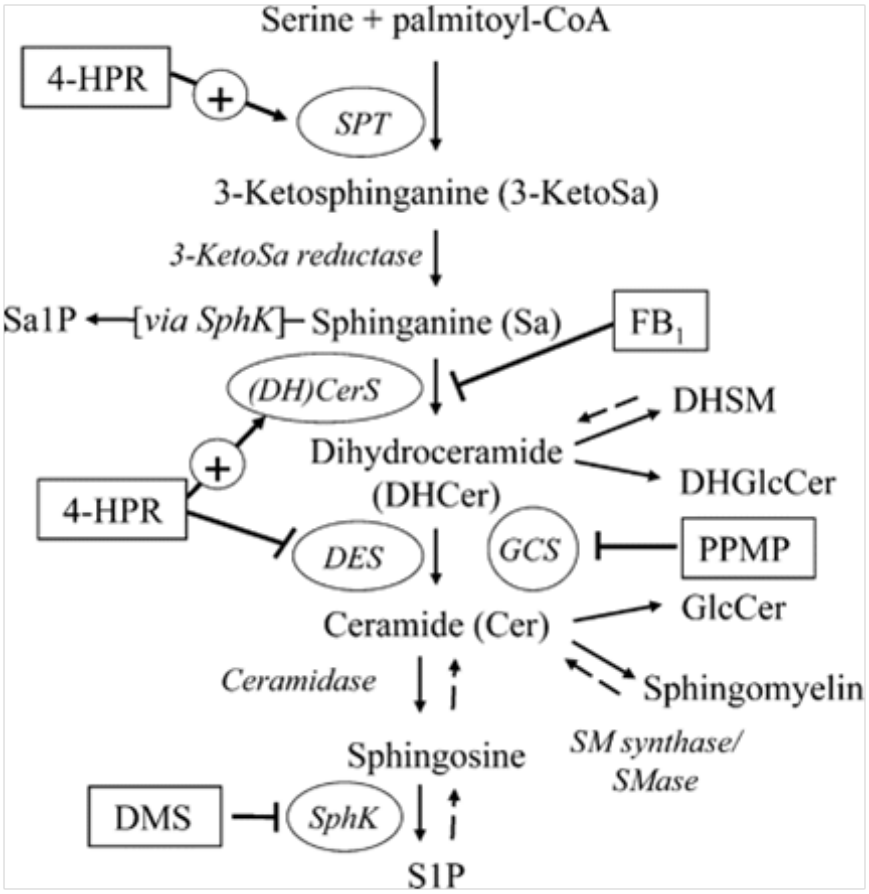


Figure 15: Action of 4HPR on sphingolipid metabolism: 4HPR enhances the activity of enzymes SPT, (DH)CerS, and inhibits DES

As a result of 4HPR treatment, from the known mechanisms of 4HPR (Figure 15), intracellular levels of Sa and DHCer are expected to increase while Cer levels are expected to decrease in HEK cells. Experimental data in Figure 16 illustrates the measured changes in the ratio of intracellular sphingolipid metabolite levels DH*:Cer* at time $t = 24h$ after 4HPR treatment, which is as expected in HEK cells based on the current understanding of 4HPR mechanisms on sphingolipid metabolism (Figure 15). However, the same expected changes in sphingolipid metabolite levels in SPT cells are not observed.

Because 4HPR treatment is similar to SPT overexpression to the extent that SPT activity is enhanced in both treatments, model results in terms of predicted steady-state feedback is verified in part by comparing Figures 13 and 16, where the predicted steady-state feedback (Figure 13) is quantitatively consistent with the observed changes in ratios of sphingolipid metabolite levels at time $t = 24h$ based on the available experimental data (Figure 16). In other words, quantitative agreement of the predicted steady-state feedback with additional independent experimental data suggests that the model predictions can be used to identify, without prior knowledge, which pathway metabolites may be responsible for differences in SPT vs. HEK cell response to 4HPR treatment, i.e., which metabolites may be controllable or are more involved in pathway homeostasis.

From a second perspective (that is not necessarily exclusive from the first), the predicted steady-state feedback specific to modeling homeostasis as a result of SPT overexpression could also contribute to a more comprehensive study to understand the metabolic effects of 4HPR by deconstructing its known mechanisms independently. In this case, modeling the effect of SPT overexpression, and its predicted steady-state feedback from available data, is the first step in testing and modeling individual mechanisms of 4HPR separately before these effects are combined. Consequently, quantitative agreement between the model results, i.e., predicted steady-state feedback, with independent 4HPR response data should lend confidence that the proposed approach using a comparator model is viable, given that the proposed approach and comparator model is built only on experimental data from SPT overexpression that represents only one of three reported mechanisms of 4HPR. Subsequently, future work may apply the proposed approach similarly but on other data

from (DH)CerS overexpression and DES underexpression separately in order to complete the 4HPR response model.

Thus, the proposed approach using a comparator model to reverse engineer homeostasis in metabolic pathways is useful in providing (a) a numerical approach to predict quantitative differences in sphingolipid metabolite levels observed between SPT and HEK cells as a result of specific experimental treatments, and (b) an analytical paradigm, in terms of homeostasis and control, to try and answer *why* these differences are observed.

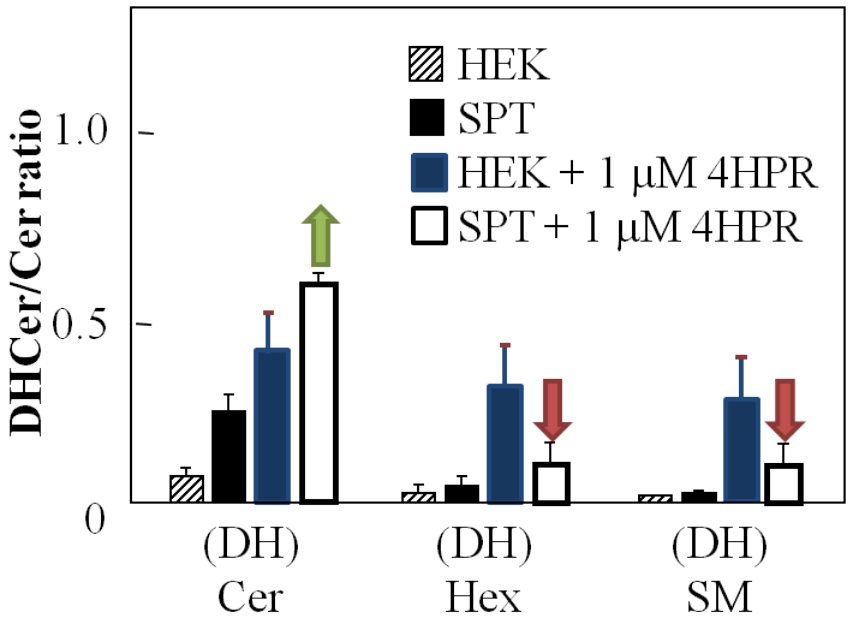


Figure 16: Effect of $1\mu M$ 4HPR on DHCer/Cer levels (experimental data): DHCer:Cer ratios are increased in HEK cells, which is consistent with expectation of increased DHCer and decreased Cer levels; changes in SPT ratios are consistent with predicted steady-state feedback (Figure 13)

3.3 Generality of proposed approach

3.3.1 Application to C26:0 sphingolipid metabolites

To demonstrate the generality of the proposed approach, the comparator model is applied to a second, independent dataset of C26:0 sphingolipid metabolites. Figures 17 and 18 show the experimental data and model results, in terms of sphingolipid metabolite amounts, for wild type and treated cells based on a similar GMA state-space model of pathway dynamics. For the wild type (Figure 17), no data were available for Sa*, and technical replicates were not available for DH*. For treated cells (Figure 18), technical replicates were not available for DH*.

C16:0 and C26:0 sphingnolipid metabolites have the same pathway topology for *de novo* biosynthesis, which also draws on a common pool of sphingolipid precursor PalCoA as the initial pathway reaction substrate. However, experimental data shows that the pathway dynamics of C26:0 sphingolipid metabolites is very different from the C16:0 sphingolipid metabolites. In both wild type and treated cells, similar amounts of sphingoid bases, So*, are observed for both the C16:0 and C26:0 sphingolipid pathways, but slightly less dihydrosphingolipids, DH*, and significantly less sphingolipids, Cer*, are detected in the C26:0 pathway compared to the C16:0 pathway. Because homeostasis is driven primarily by changes in system dynamics, differences in pathway dynamics between both systems from experimental data should be captured in the model results as well.

In Figure 17 for wild type cells, for So*, experimental data show that the amount of So increases gradually from $t = 0h$ and settles to steady-state at $t = 4h$. The amount of SoP remains low over the course of experiment. For DH*, noticeable changes are observed in DHCer over the course of experiment but without any obvious discernible trend, while amounts of DHGC and DHSM remain low throughout. For Cer*, the experimental data show that no dual-labeled Cer, GC, or SM were detected. The model results do not fit well for most metabolites, where they tend to overestimate the experimental data.

In Figure 18 for treated cells, for Sa*, the experimental data shows that Sa increases quickly from $t = 0h$ to $t = 2h$, then decreases to settle at steady-state by $t = 6h$. SaP increases gradually over the course of experiment. For So*, both So and SoP increase steadily

throughout the course of experiment, but So increases at a much faster rate. For DH*, the experimental data show that no dual-labeled DHCer, DHGC, or DHSM were detected. For Cer*, Cer and SM increase slightly after $t = 2h$ and $t = 5h$ correspondingly while no GC is detected. Qualitatively, the model results are a good fit with the experimental data for 5 metabolite species (SaP, SoP, DHGC, DHSM, and GC), fair for 4 metabolite species (Sa, So, Cer, and SM), and poor for 1 metabolite species (DHCer).

Figure 19 shows the experimental data and model results for C26:0 sphingolipid metabolites in treated cells with the comparator model. Qualitatively, the model results are a good fit for 5 metabolite species (SaP, SoP, DHGC, DHSM, and GC), fair for 3 metabolite species (Sa, So, and Cer), and poor for 2 metabolite species (DHCer and SM). Table 6 shows a quantitative measure of the accuracy of these model results in matching experimental data at discrete time in terms of RSS error. With regard to the GMA state-space model of pathway dynamics (top 2 rows), the model results are a better fit to the experimental data for the wild type than treated cells. Comparing the model results for treated cells with and without the comparator model (bottom 2 rows), goodness-of-fit differs most significantly for DH*, with an increase in RSS error of 19.4% with the comparator model, while RSS error is decreased for the other groups of sphingolipid metabolites. Overall, RSS error increases by 18.9% with the comparator model, where the increase in RSS error for DH* is most likely the primary factor for the difference.

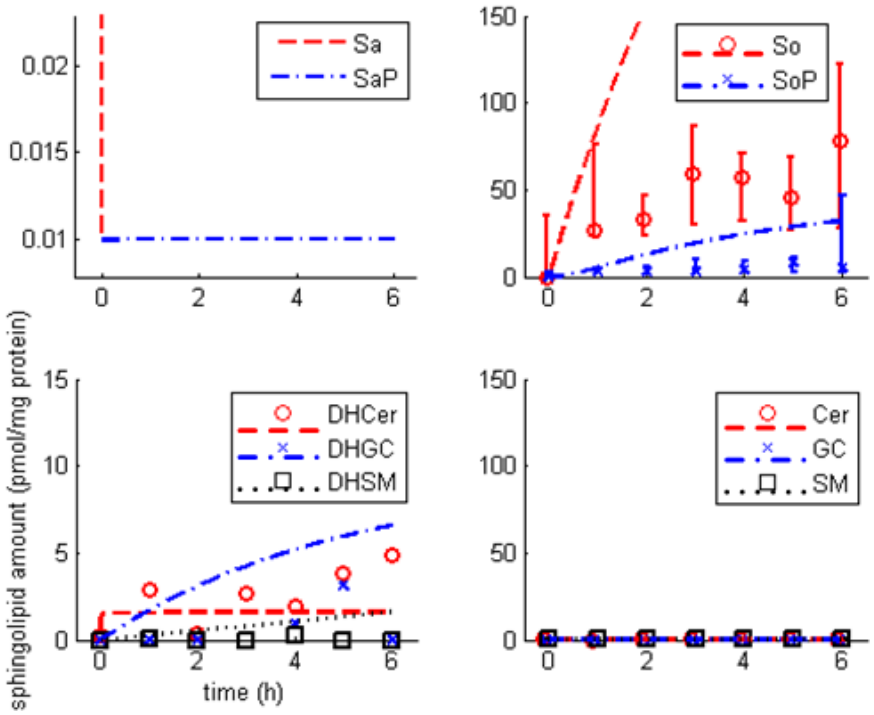


Figure 17: Experimental data and model results for C26:0 sphingolipid metabolites in wild type cells (*reference* system): abscissa - time (hours), ordinate – experimentally normalized sphingolipid metabolite amounts (pmol/mg protein); scatterplots represent *in vitro* experimental data (median and range indicated where available), broken lines represent *in silico* model results.

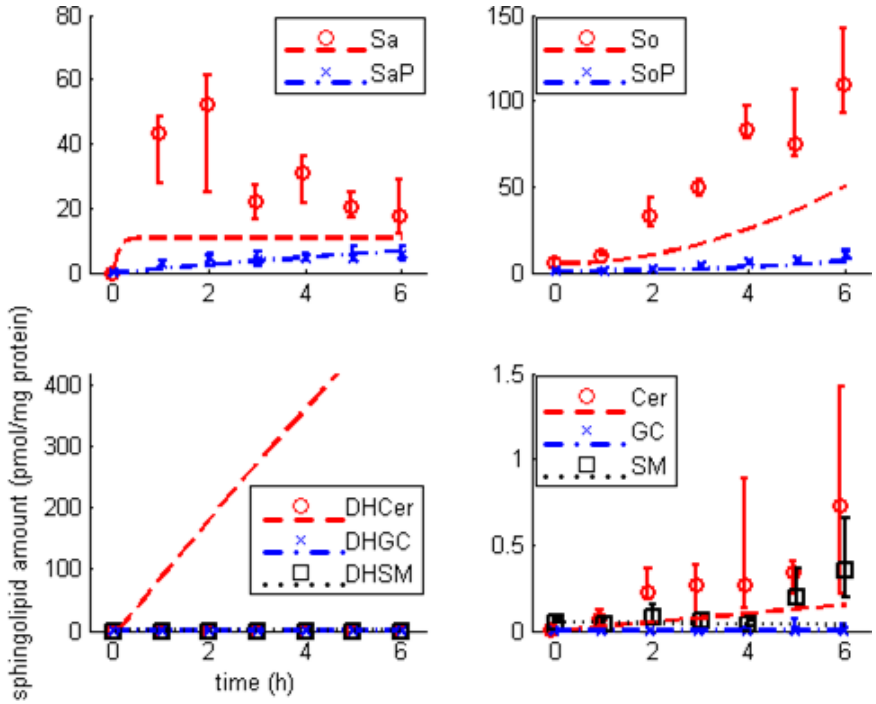


Figure 18: Experimental data and model results for C26:0 sphingolipid metabolites in treated cells (*plant* system) *without* comparator: abscissa - time (hours), ordinate – experimentally normalized sphingolipid metabolite amounts (pmol/mg protein); scatterplots represent *in vitro* experimental data (median and range indicated where available), broken lines represent *in silico* model results.

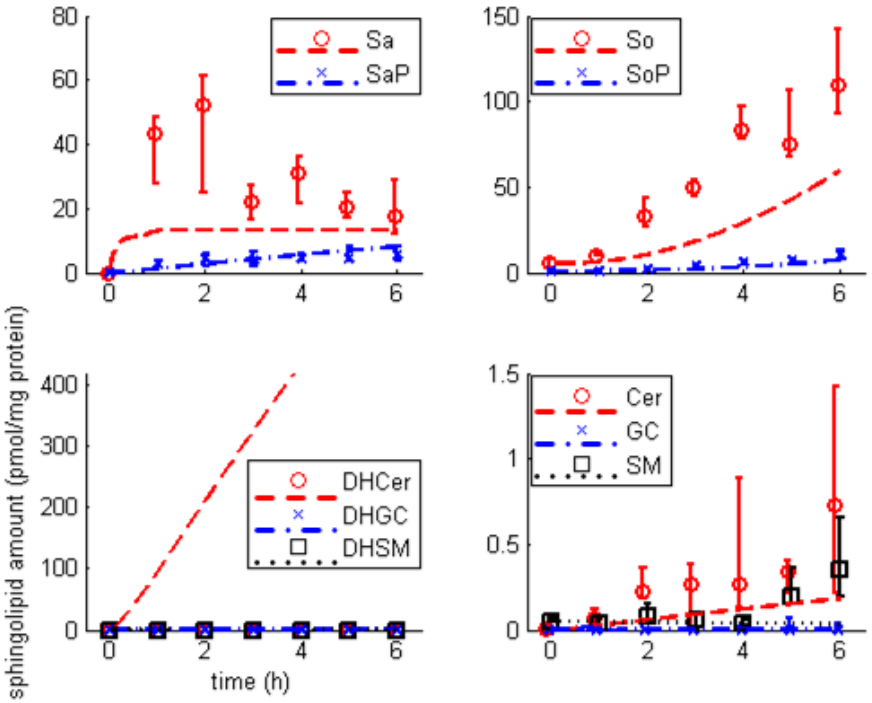


Figure 19: Experimental data and model results for C26:0 sphingolipid metabolites in treated cells (*plant* system) *with* comparator: abscissa - time (hours), ordinate – experimentally normalized sphingolipid metabolite amounts (pmol/mg protein); scatterplots represent *in vitro* experimental data (median and range indicated where available), broken lines represent *in silico* model results.

Table 6: State-space error: C26:0 sphingolipid metabolites

		Sa*	DH*	Cer*	So*	all
C26:0						
Wild type		n/a	13.78	0.29	458.23	458.43
Treated cells	*without* comparator	54.10	860.87	0.86	104.56	868.88
♯	*with* comparator	48.88 *(-9.6)*	1027.90 *(+19.4)*	0.82 *(-4.7)*	95.07 *(-9.1)*	1033.40 *(+18.9)*

♯ percentage change over RSS error in treated cells *without* comparator

3.3.1.1 Steady-state feedback

Figures 20 and 21 shows the key model outcome, i.e., predicted steady-state pathway feedback in terms of input and aggregate state feedback as well as individual sphingolipid metabolite feedback based on the proposed comparator model of homeostasis in C26:0 sphingolipid *de novo* biosynthesis in response to single-gene (SPT) overexpression in HEK cells. Because the C26:0 sphingolipid data and biology is different from the C16:0 metabolites, it is expected that the model results, i.e., predicted steady-state feedback, are different from the latter case.

Figure 20 shows the predicted steady-state pathway feedback in treated cells in terms of 2 components: (a) system input and sphingolipid precursor, palmitoyl-CoA (PalCoA), $u_r(t)$, and (b) aggregate state feedback from pathway sphingolipid metabolites $u_x(t)$. At steady-state, the magnitude of the feedback suggests the contribution of each component to homeostasis, i.e., a larger magnitude indicate greater contribution, and sign (+/-) indicates either enhancement (positive feedback) or inhibition (negative feedback) to the pathway.

Both components exhibit a brief period of transient oscillation from $t = 0h$ to $t = 1h$ with similar frequency. However, the peak-to-peak amplitude of oscillation is significantly larger for $u_r(t)$ at ~ 270 compared to $u_x(t)$ at ~ 140. In addition, $u_r(t)$ settles to steady-state at $\sim +130$ while $u_x(t)$ settles to steady-state at ~ 0. Compared to C16:0 sphingolipids, the dynamics of feedback control in C26:0 sphingolipids are clearly different in terms of period and amplitude of transient oscillations. At the same time, these feedback dynamics are consistent with designing a control policy that utilizes the least resources for pathway regulation, which is similar to the case for C16:0 sphingolipids.

Figure 21 shows the predicted steady-state feedback gain from individual pathway sphingolipid metabolites $h_x(t)$. These feedback gain are dimensionless; however, at steady-state, the magnitude and sign of these feedback values may be interpreted to suggest differential activity in pathway homeostasis. More precisely, the larger the magnitude, the greater the feedback; furthermore, the sign (+/-) of the gain values indicate either enhancement (positive feedback) or inhibition (negative feedback) to the pathway.

The steady-state feedback gain is non-zero for 5 metabolite species (PalCoA, So, DHGC,

DHSM, and GC), but is unchanged from the zero initial condition for the other sphingolipid metabolites. Thus, Figures 19 and 20 show that the predicted steady-state feedback gain for C26:0 sphingolipid metabolites are different from the C16:0 sphingolipid metabolites, which suggests a different set of homeostatic interactions may be responsible for pathway regulation in this system. On the other hand, predicted steady-state feedback due to system input palmitate $u_r(t)$ for both C16:0 and C26:0 sphingolipid metabolites are similar in that PalCoA is suggested to have a self-regulatory role, where excess input PalCoA appears also to inhibit the dynamics of sphingolipid *de novo* biosynthesis.

From a modeling perspective, the significance of these model results is that the proposed approach, using a comparator model to reverse engineer homeostasis in a highly regulated metabolic pathway, can be applied to other datasets in general. Because of a lack of prior knowledge, the biological interpretation and verification of these model results are not discussed in detail here. Furthermore, the model results are consistent with the understanding that homeostasis is driven primarily by pathway dynamics, as opposed to pathway topology, which can be derived directly from time-series experimental data.

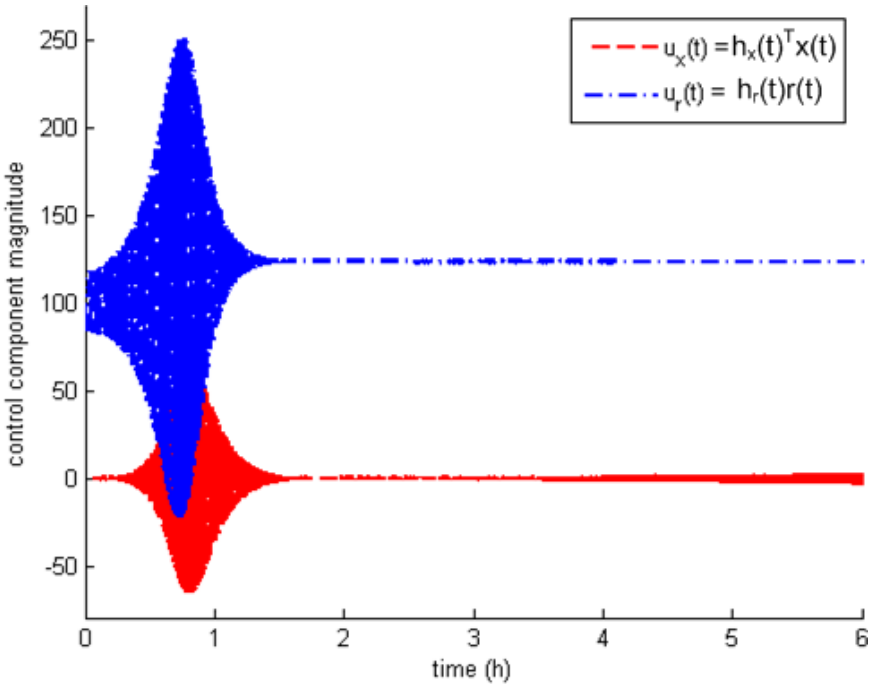


Figure 20: Predicted steady-state feedback for C26:0 sphingolipid metabolites pathway in treated cells in terms of: (a) system input and sphingolipid precursor, palmitoyl-coA (Pal-CoA), $u_r(t)$ (blue), and (b) aggregate state feedback from pathway sphingolipid metabolites $u_x(t)$ (red) - x-axis: time (hours), y-axis: magnitude (dimensionless)

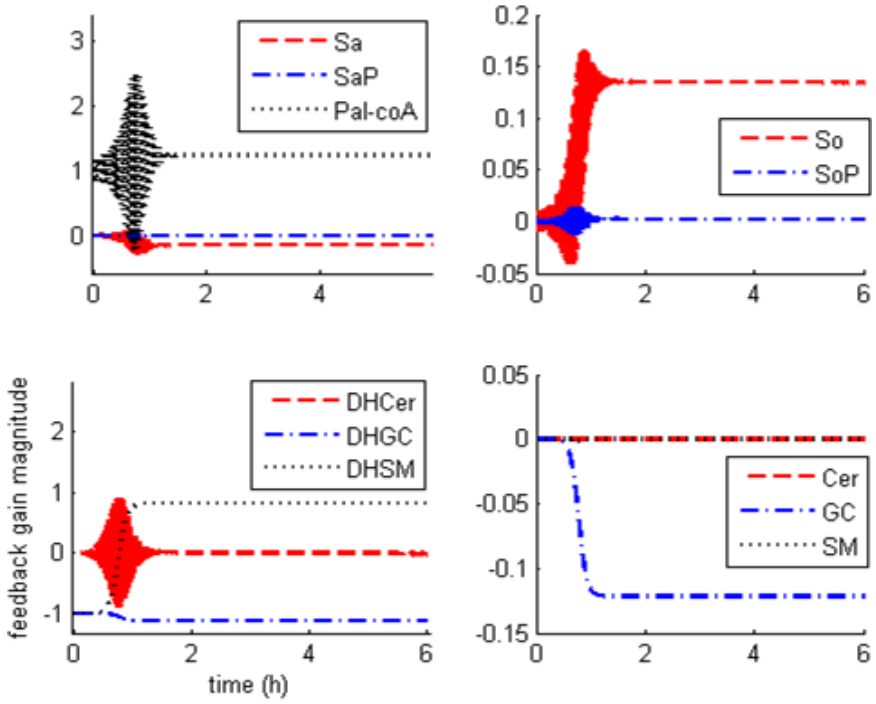


Figure 21: Predicted steady-state feedback for individual C26:0 sphingolipid metabolites in treated cells - x-axis: time (hours), y-axis: magnitude (dimensionless)

3.4 Summary

In summary, *in silico* results of the proposed approach using a comparator model to reverse engineer homeostasis in a highly regulated metabolic pathway have shown:

- (a) the efficacy of the proposed comparator model in capturing observed pathway dynamics, and (b) robustness of predicted steady-state pathway feedback in response to simulated errors in state-space parameter estimation,
- the validity of predicted steady-state pathway feedback in the context of homeostasis in C16:0 sphingolipid *de novo* biosynthesis, by comparison to literature as well as independent data on measured changes in sphingolipid metabolite levels in response to 4HPR dosage in both treated and wild type cells, and
- the generality of the proposed modeling approach, by applying the comparator model to a different metabolic system using independent data on C26:0 sphingolipid metabolites.

Using C16:0 sphingolipid *de novo* biosynthesis as a case study, the proposed approach to reverse engineer pathway homeostasis using a comparator model resulted in the prediction of steady-state pathway feedback, which is qualitatively and quantitatively verified with literature as well as additional independent experimental data. Furthermore, the comparator model is also applied to a separate independent dataset on C26:0 sphingolipids, where differences in model results are also consistent with differences in experimental data between the 2 datasets. As an initial study in "the use of existing techniques in well-developed areas of control theory to analyze problems of interest to biologists" [146], the outcomes of this case study on C16:0 sphingolipid metabolites have also been published in [125, 127].

CHAPTER IV

FUTURE WORK

Using C16:0 sphingolipid *de novo* biosynthesis as a case study, the concept, development, application, and verification of the proposed approach in using a comparator model to reverse engineer homeostasis in a highly regulated metabolic pathway have been discussed in the previous chapters. So far, the model results, i.e., predicted steady-state feedback to capture homeostatic pathway regulation, are encouraging and should warrant at least some further investigation in this area. This dissertation is an initial study of what will hopefully become a longer program in cellular control. In this chapter, I discuss limitations of the proposed comparator model and existing experimental data well as potential biomedical applications of the proposed approach.

4.1 Model and data limitations

4.1.1 Modeling nonlinear pathway dynamics

Using generalized mass action (GMA) to simplify system dynamics in this case study, a linear system of first-order ODEs is used to describe what could very possibly be nonlinear pathway dynamics. To account for nonlinear pathway dynamics, additional modeling parameters in terms of the plant and comparator dynamics could be written as follows:

$$\dot{\vec{x}}(t) = \mathbf{A}\vec{x}(t) + \vec{b}[u(t) + \vec{\alpha}(t)^T \Phi(\vec{x}(t))] \tag{31}$$

$$u(t) = \vec{h}_x(t)^T \vec{x}(t) + h_r(t) r(t) - \hat{\vec{\alpha}}(t)^T \Phi(\vec{x}(t)) \tag{32}$$

$$\dot{\hat{\vec{\alpha}}}(t) = \Gamma_\alpha \Phi(\vec{x}(t)) \vec{e}(t)^T \mathbf{P} \vec{b} \tag{33}$$

where $\Phi(\vec{x}(t))$ is a set of predefined nonlinear parameters, $\vec{\alpha}(t)$ is the true contribution of these parameters to the underlying pathway dynamics, and $\hat{\vec{\alpha}}(t)$ approximates the true contribution of these terms.

Particular nonlinear terms could be derived from the experimental data [137], although such approaches for parameter estimation in dynamic nonlinear systems do not rely on any prior knowledge of the system. Instead, model 'building blocks' that are generic mathematical functions that are not specific to the pathway biochemistry are tested for goodness-of-fit to the experimental data. As a result, such approaches to estimating nonlinear parameters may not provide any biological basis to support the parameters. Nonetheless, although it is beyond the scope of this dissertation, such a "model-free", or "top-down", approach to parameter estimation may still be useful to uncover radically new descriptions of the dynamics of metabolic pathways in the long term by essentially rewriting the current laws of biochemistry.

At the same time, to avoid data over-fitting, i.e., where model parameters may be added indiscriminately for the sole purpose of reducing error between model results and experimental data, measures of goodness-of-fit such as the Akaike information criterion [1, 69] may be useful to quantify the tradeoffs between precision and complexity of the model. For any proposed model to be of interest and utility to the field of biology at present, mathematical terms to model pathway dynamics must be informed and justified by an understanding of the underlying pathway biochemistry and dynamics.

4.1.2 Exploring control models: additional data requirements

The proposed comparator model is only a basic component of standard, as well as more advanced, controlled systems. In this case, its function is simple: to compute the difference between the *plant* and *reference* system outputs, which is feedback to the plant system so as to guide the plant to converge to the reference. However, in more advanced controllers, the comparator output is examined more carefully in significantly greater detail such that how the primary issue of interest lies in how the comparator output is, or may be, dealt with in terms of *feedback*. Consequently, the objective in engineering control design is to specify the appropriate feedback to achieve the desired system response.

Still, improved biological data acquisition is required in order to better meet, and indeed embrace, the theoretical and practical challenges of "[using] existing techniques in

well-developed areas of control theory to analyze problems of interest to biologists" [146]. Representative modes of (feedback) control are summarized in Table 7, where key features of these various modes of control also suggest what experimental data may be useful to explore and integrate, in the broad sense, the proposed approach using control theory to reverse engineer homeostasis in biochemical systems more fruitfully. In other words, additional experimental data is needed to properly calibrate such models of homeostasis based on control theory. Regardless, to achieve any real success, modelers surely need to work more closely with experimentalists.

In particular, for example, because the key feature of adaptive control is that the plant system is time-varying, i.e., where state-space parameters change with time, this suggests one of two possible demands on experimental data: *either* the pathway must be sampled more frequently, i.e., to gather time-series data, *or* over a contiguous period of experiment, the plant system should be subject to a sequence of different experimental treatments and such that each treatment affects the pathway dynamics differently. In the former case, the contribution of more time-series data to modeling is to enable detection (and derivation) of changes in reaction rate constants in real-time, i.e., to improve state-space parameter estimation with significantly improved temporal resolution. In the latter case, changes in the plant system as a result of a sequence of different experimental treatments also requires, and results in, different parameters to describe pathway dynamics following each treatment. In both cases, the effect of these changes in experimental procedure or data acquisition is that the key feature of adaptive control can be satisfied, i.e., a time-varying system where the state-space parameters change with time can be properly defined.

For robust control, apart from a fully known system, a range of inputs is required for the estimation of controller parameters in order to establish limits of input perturbation where it can be claimed that pathway homeostasis may indeed be described as such. Of the representative control modes summarized in Table 7, the data requirements for robust control may be the easier to meet, where the key factor to be varied could simply be the magnitude of pathway input, e.g., precursor molecule for the metabolic pathway of interest. Even so, depending on the resolution of the measurement instrumentation, sufficiently

different levels of pathway input that still lead to the same steady-state behavior may still be challenging to discern.

Another example: to apply optimal control as a model of homeostasis, where the desired system performance as well as system dynamics must be fully specified, substantial *a priori* knowledge of the metabolic pathway or system of interest will be useful, if not absolutely necessary, to define these key features. Optimal control implies direction, either internal or external, where system performance is driven towards achieving some known objective. Under heavily-regulated conditions for externally-driven objectives, e.g., ensuring pre-defined levels of particular metabolites for specific bioengineering projects, optimal control could be a viable model of regulation (but not homeostasis) [48, 64]. On the other hand, it is challenging to define "desired" system performance for homeostasis in metabolic pathways without controversy, given that biological variance all but ensures a range of acceptable (and healthy) behaviors in the wild type. To do so will require conclusive evidence that deviation of plant system response necessarily leads to "destructive" outcomes that are detrimental to cell survival. Still, in such cases, data on cancerous tumor cells could provide some support to modeling homeostasis in metabolic pathways based on optimal control.

Finally, for proportional, integral, and derivative (PID) control to be an effective model of homeostasis in metabolic pathways, a fully known system and coherent set of inputs are required. At the same time however, because it can be directly applied to scalar representations of pathway systems, PID control may be advantageous over other control modes from a modeling perspective in that even single-variable measurements, still sufficiently resolved in time, may suffice to support model development. For instance, PID control (specifically integral control) is reported to be an adequate model of regulation in bacterial chemotaxis [177, 176]. Nonetheless, it should also be noted that where some terms of PID control may seem more abstract than others, e.g., how integral control is defined in bacterial chemotaxis, *a priori* knowledge of the metabolic system goes a long way to illustrate how specific forms of PID control may be implemented via some of the known underlying molecular mechanisms.

Table 7: Representative control modes

Mode	Key feature(s)	Prior knowledge	Result
Adaptive	Time-varying system	Not required	Real-time control response Guaranteed system stability
Robust	Fully known system and known range of inputs	Some required, to determine limit of change in parameters	Pre-determined control response, i.e., does not react to unexpected disturbance Guaranteed system stability
Optimal	Fully known system and desired performance objectives	Fully required, to specify objective function for optimization	Pre-determined control response, i.e., does not respond to unexpected disturbance Guaranteed system stability
PID	Fully known system and inputs	Some required, to tune control response	Pre-determined control response, i.e., does not respond to unexpected disturbance System stability may not be guaranteed

4.2 Biomedical applications

4.2.1 Studying molecular mechanisms

The predicted steady-state feedback indicates which specific metabolites may be involved in potential homeostatic pathway interactions, but does not suggest how these homeostatic interactions may be accomplished *in vitro* in terms of observable molecular mechanisms, e.g., in terms of intracellular location, transport, or biochemistry of enzymatic reactions. More precisely, for instance in this case study, a particular molecular mechanism that may be of interest to biologists is how treated cells, i.e., SPT overexpressing cells, may sense that increased amounts of sphingolipid precursor PalCoA are present to determine if the balance of sphingolipid *de novo* biosynthesis and turnover is indeed upset. Could this be accomplished at the molecular level via the rate at which extracellular palmitate is taken up by the cells and converted to PalCoA?

Specific to this case study, because intracellular location is critical to sphingolipid metabolism, molecular imaging techniques, e.g., using molecular beacons, may be the immediate methods of choice to first determine if particular sphingolipid metabolites are indeed co-localized so that subsequent enzymatic reactions could possibly occur and be detected. Ultimately, unraveling such mechanisms at the molecular levels will go a long way to lend substantial experimental support to the use of control theory to reverse engineer homeostasis in metabolic pathways.

4.2.2 Impacting drug discovery and development

By identifying specific pathway metabolites that may be involved in homeostatic pathway interactions based on predicted steady-state feedback, model results from the proposed approach may also contribute to the process of drug discovery and development, e.g., in terms of identifying key pathway metabolites that may be targeted biochemically. Drug therapy is a common treatment for chronic disease where homeostasis is disrupted but remains viable [80]. However, drug-induced metabolic activity often conflict with homeostatic activities, leading to undesirable side effects. This problem affects drug discovery and development pipelines where drug candidates often fail in development because of latent toxicity.

The drug discovery and development process is a pipeline that involves different phases. In the discovery phase, disease targets and drug candidates are identified. *In silico* methods currently used for target identification focus on algorithms for pattern matching sequences and motifs, functional annotation, and mining expression datasets [93]. These methods tend to identify singular targets, where less emphasis is placed on assessing the targets in the context of metabolic pathways. Drug candidates are subsequently designed and evaluated for potency towards these targets, commonly to knock out disease-specific pathways. In the development phase, drug candidates are tested for toxicity in animal and human trials. In current pipelines, the separation of drug discovery and development into distinct phases has led to more failures than successes because "with discovery now driven primarily by chemistry and high-throughput screening, the biological effects and, in particular, the toxicity of new compounds is largely not appreciated until a compound enters development" [158]. This results in a very low turnover rate of drug candidates from concept to clinical deployment.

The prevalent strategy to improve the current situation is to try and increase the number of drug candidates discovered such that the number of drugs that are safe for market is not affected significantly. However, this does not improve the turnover rate and critically, it does not alleviate the burden of costs involved in the process. The recent trend of corporate mergers and acquisitions in the pharmaceutical industry reflects the growing financial strain on drug discovery and development as a result. An alternative solution is to consider toxicity earlier upstream in the discovery phase instead of doing so only in the development phase.

Thus, the proposed approach using a comparator model to reverse engineer homeostasis in metabolic pathways provides an analytical framework to identify and interpret aberrant metabolic effects in terms of interference with the dynamics of homeostasis in healthy pathways. In particular, in pathways affected by disease or drug toxicity, the clinical symptoms can be viewed as the result of sub-par pathway feedback that leads the metabolic system to undesirable, and ultimately unstable, points of equilibrium. It follows that a viable strategy for drug intervention in such aberrant pathways may be to try and complement or leverage the action of these inferior "homeostatic" feedback, instead of trying to knock out the

pathway entirely. Consequently for drug discovery and development, drug toxicity may be addressed by working with, not against, abnormal homeostatic activity in the context of diseased pathways.

4.2.3 Analyzing case/control studies

Finally, the proposed approach using a comparator model to reverse engineer homeostasis in metabolic pathways is a knowledge-based method, where the pathway dynamics of wild type cells is defined as a reference response for the treated cells to approach. In fact, this *plant-reference* paradigm is motivated by conventional experimental design in the field of biology where effects of experimental treatment are commonly studied using stable cell lines in comparison with the wild type, i.e., traditional case/control studies. Naturally, in such studies, there is only a single difference as a result of experimental treatment, e.g., single-gene mutation, between the treated cells and wild type.

Where dynamics of the metabolic pathway(s) of interest do not evolve beyond the wild type dynamics, such a plant-reference framework is justified where treated cells can be assumed to follow the wild type behavior. Consequently, together with its demonstrated generality in this dissertation, the proposed approach using a comparator model from engineering control to reverse engineer homeostasis in metabolic pathways can be a useful research tool to complement the analysis of data from traditional case/control studies in the field of biology.

Furthermore, taking a broader perspective on this subject, in applied research clinical data from individual patients in times of health may be similarly used as a *reference* to shortlist options for molecular intervention, e.g., drug therapy, in times of disease. Data collected from the patient in times of health may be used to build and calibrate a *reference* system, while data from the same patient in times of disease may be used to estimate a *plant* system. Thereafter, the clinical goal is guide the dynamics of the plant system to return to the dynamics of the reference system. In other words, each patient may (rightfully) provides her own "healthy" data as a reference response, which addresses the problem of biological variance in populations. Thus, if achieved, such a personalized approach to medicine stands

in stark contrast to the use of population statistics to define standards of health for the individual patient.

Even remaining within the limits of population (public) health, a reference system could similarly be built and calibrated using data from healthy subjects to test novel *intervention* treatments (as discussed in the previous Section 4.2.2). In other words, intervention therapy could be predicted *in silico* in models of plant systems based on data from diseased patients. Ultimately, such an application is based on the clinical need to unravel the figurative knot of biochemical interactions as a result of increasingly common drug cocktail therapy to manage "lifestyle diseases", because in many cases it is no longer sufficient to merely predict the initial or immediate patient response to particular drug therapies. As a result of such chronic diseases, an increasing number of patients are already subject to specific dosing regimens which may interfere with additional prescribed treatments.

APPENDIX A

ESTIMATED RATE CONSTANTS: C16:0 PATHWAY

Table 8: Estimated reaction rate constants: C16:0 sphingolipids, wild type

$k_{SaP,Sa}$	0.0043
$k_{DHCer,Sa}$	1.4431
$k_{Sa,SaP}$	0.0091
$k_{Sa,DHCer}$	1.0362e+04
$k_{DHGC,DHCer}$	1.6826
$k_{DHSM,DHCer}$	11.4186
$k_{Cer,DHCer}$	4.0843
$k_{DHCer,DHGC}$	2.9521
$k_{DHCer,DHSM}$	1.2753
$k_{DHCer,Cer}$	79.9867
$k_{GC,Cer}$	0.0345
$k_{SM,Cer}$	0.0025
$k_{So,Cer}$	4.1425
$k_{Cer,GC}$	0.0297
$k_{Cer,SM}$	0.0654
$k_{Cer,So}$	7.0930
$k_{SoP,So}$	9.1800
$k_{So,SoP}$	1.9684

Table 9: Estimated reaction rate constants: C16:0 sphingolipids, SPT overexpressed

$k_{SaP,Sa}$	0.0028
$k_{DHCer,Sa}$	0.0736
$k_{Sa,SaP}$	0.2271
$k_{Sa,DHCer}$	14.5170
$k_{DHGC,DHCer}$	35.6403
$k_{DHSM,DHCer}$	8.5091
$k_{Cer,DHCer}$	13.1570
$k_{DHCer,DHGC}$	0.0838
$k_{DHCer,DHSM}$	3.9688
$k_{DHCer,Cer}$	6.2111
$k_{GC,Cer}$	6.1811e-06
$k_{SM,Cer}$	0.0038
$k_{So,Cer}$	0.0876
$k_{Cer,GC}$	0.0302
$k_{Cer,SM}$	0.1353
$k_{Cer,So}$	0.3788
$k_{SoP,So}$	13.9785
$k_{So,SoP}$	1.4213

APPENDIX B

ESTIMATED RATE CONSTANTS: C26:0 PATHWAY

Table 10: Estimated reaction rate constants: C26:0 sphingolipids, wild type

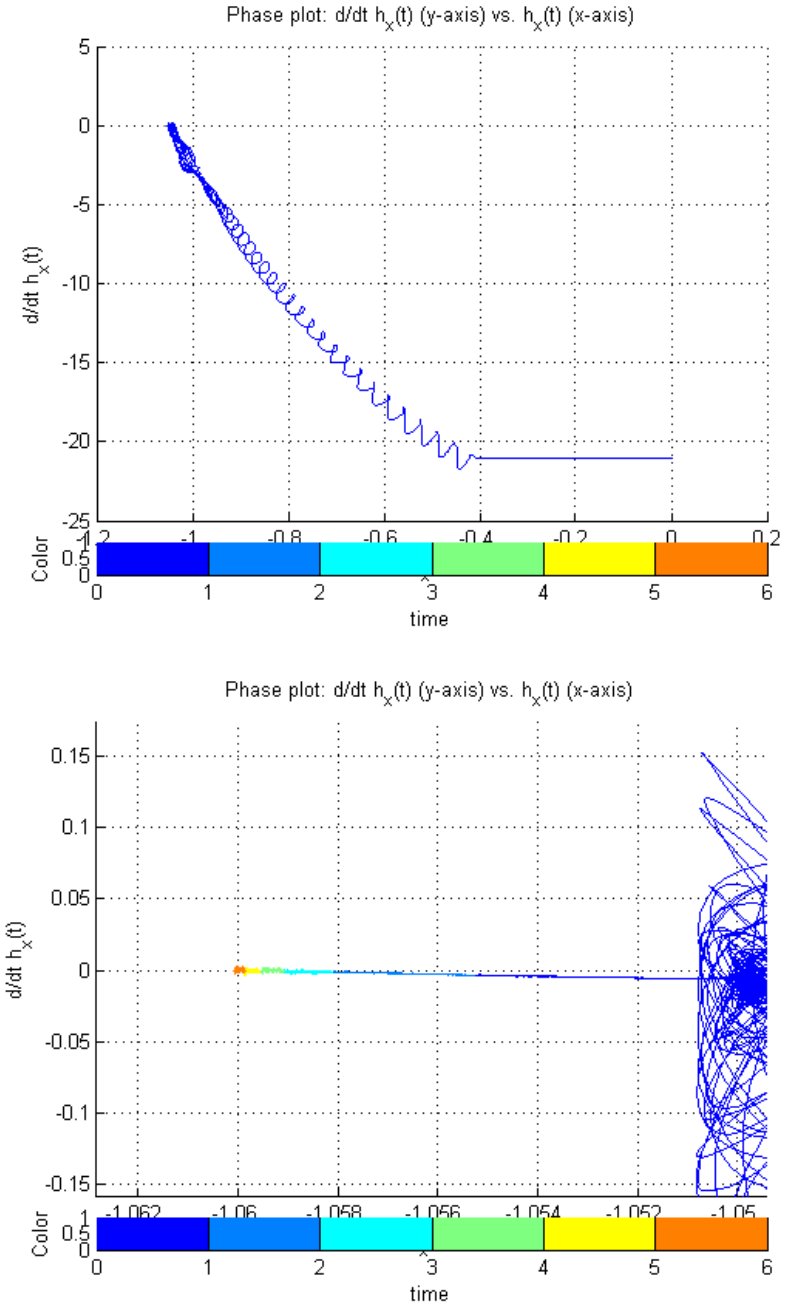
$k_{SaP,Sa}$	5.6743e-05
$k_{DHCer,Sa}$	2116.6
$k_{Sa,SaP}$	1.0271e-04
$k_{Sa,DHCer}$	1.1929e-04
$k_{DHGC,DHCer}$	0.2069
$k_{DHSM,DHCer}$	0.0047
$k_{Cer,DHCer}$	5.4449
$k_{DHCer,DHGC}$	1.2146
$k_{DHCer,DHSM}$	0.1711
$k_{DHCer,Cer}$	62.8240
$k_{GC,Cer}$	0.0853
$k_{SM,Cer}$	0.2108
$k_{So,Cer}$	0.2823
$k_{Cer,GC}$	0.0164
$k_{Cer,SM}$	0.0015
$k_{Cer,So}$	1125.4
$k_{SoP,So}$	0.0598
$k_{So,SoP}$	0.2289

Table 11: Estimated reaction rate constants: C26:0 sphingolipids, SPT overexpressed

$k_{SaP,Sa}$	0.05015
$k_{DHCer,Sa}$	9.1878
$k_{Sa,SaP}$	0.1218
$k_{Sa,DHCer}$	1.7913e-06
$k_{DHGC,DHCer}$	0.2244
$k_{DHSM,DHCer}$	4.5129e-08
$k_{Cer,DHCer}$	173.11
$k_{DHCer,DHGC}$	1.3972e-03
$k_{DHCer,DHSM}$	3.065e-05
$k_{DHCer,Cer}$	0.08087
$k_{GC,Cer}$	0.2582
$k_{SM,Cer}$	0.05618
$k_{So,Cer}$	2.1452e-06
$k_{Cer,GC}$	5.2146e-03
$k_{Cer,SM}$	1.6167e-04
$k_{Cer,So}$	112.16
$k_{SoP,So}$	2.7933e-06
$k_{So,SoP}$	0.04899

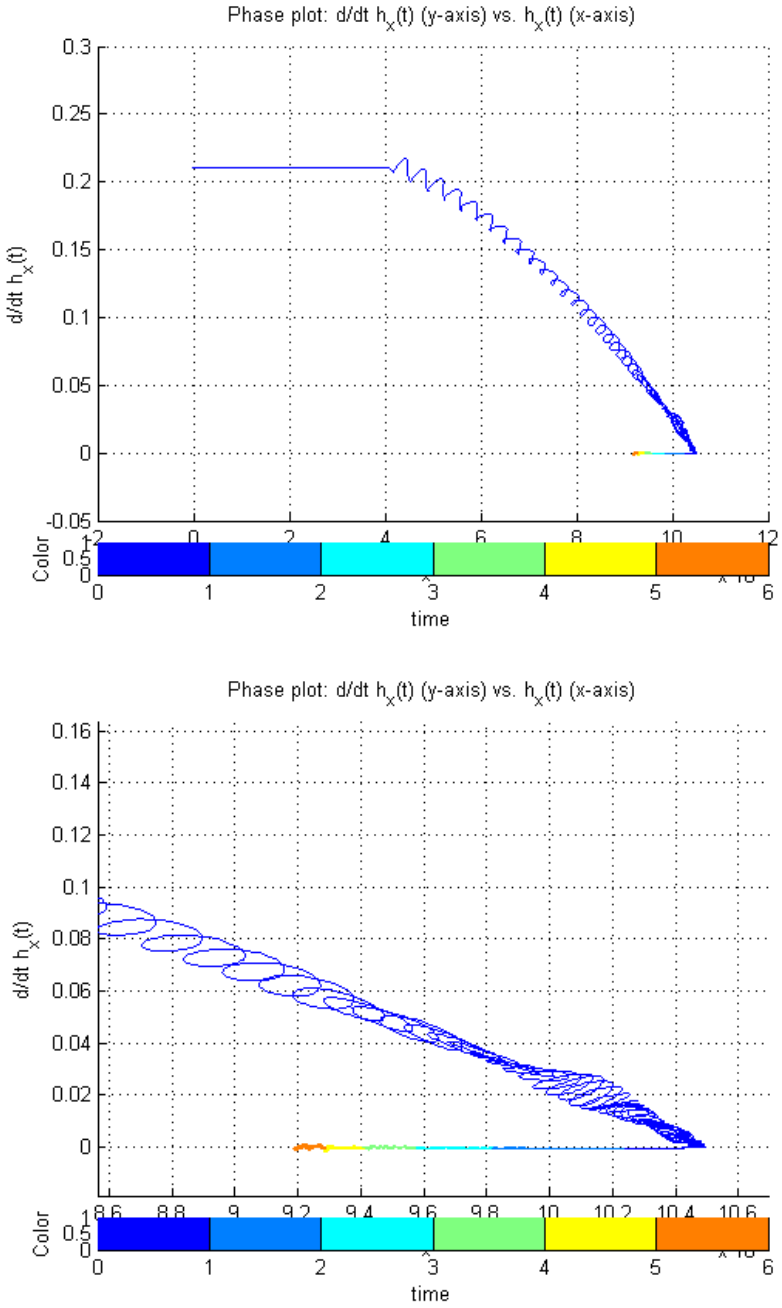
APPENDIX C

PHASE PLOTS: PREDICTED PATHWAY FEEDBACK



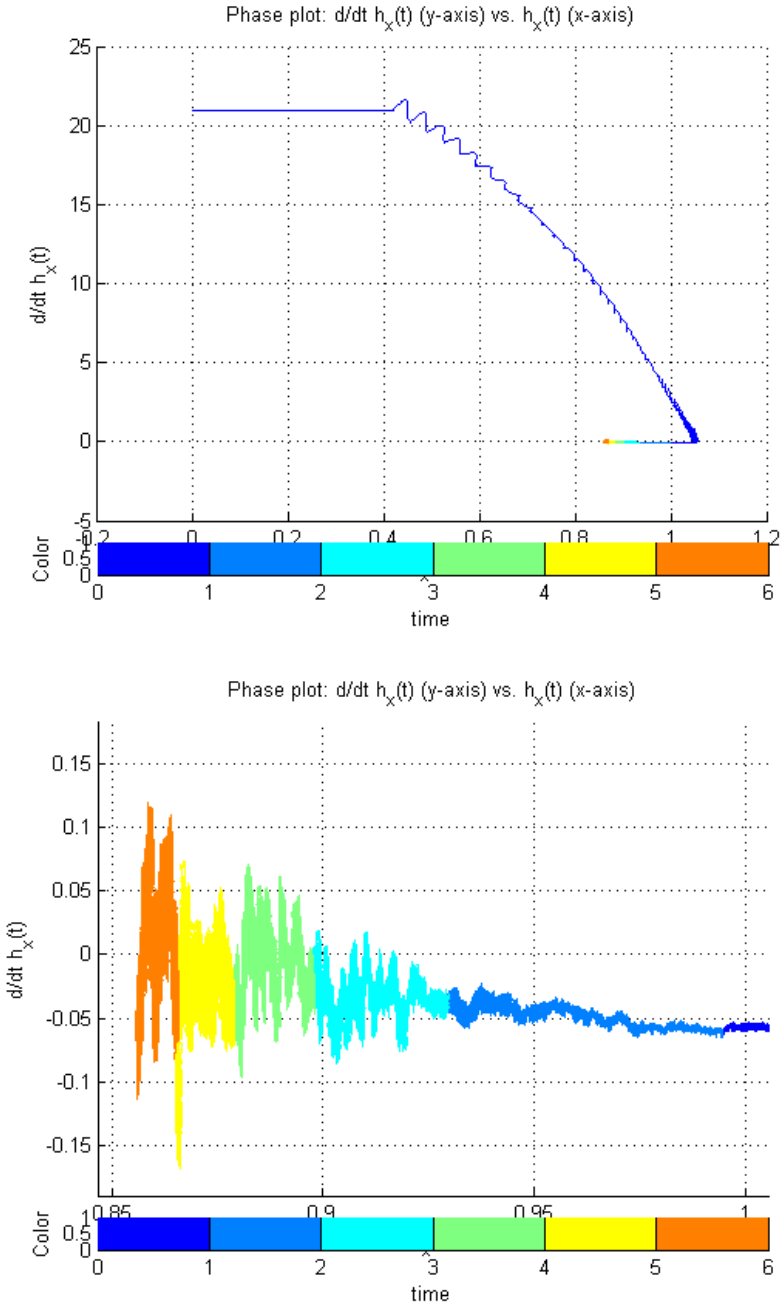
(zoomed in, from top)

Figure 22: Feedback gain phase plot: C16:0 Sphinganine (Sa) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



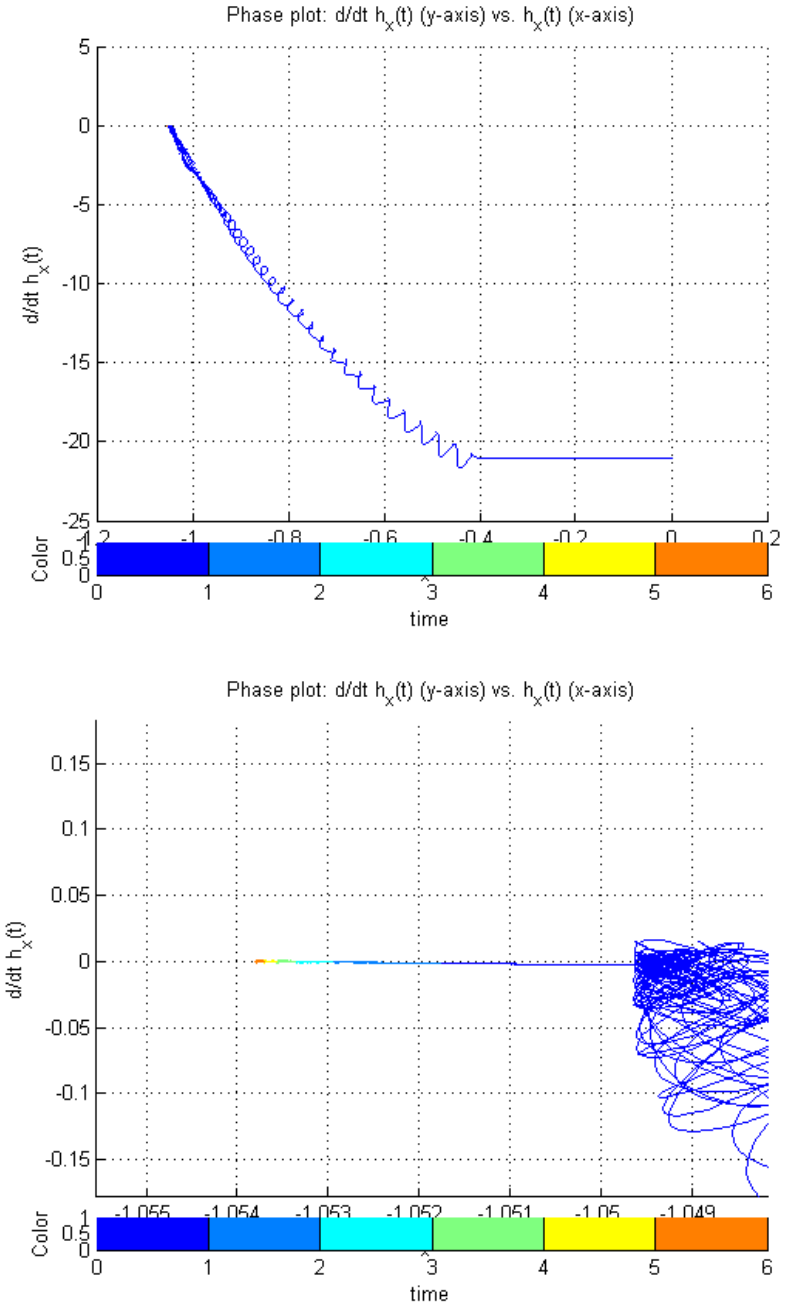
(zoomed in, from top)

Figure 23: Feedback gain phase plot: C16:0 Sphinganine-phosphate (SaP) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



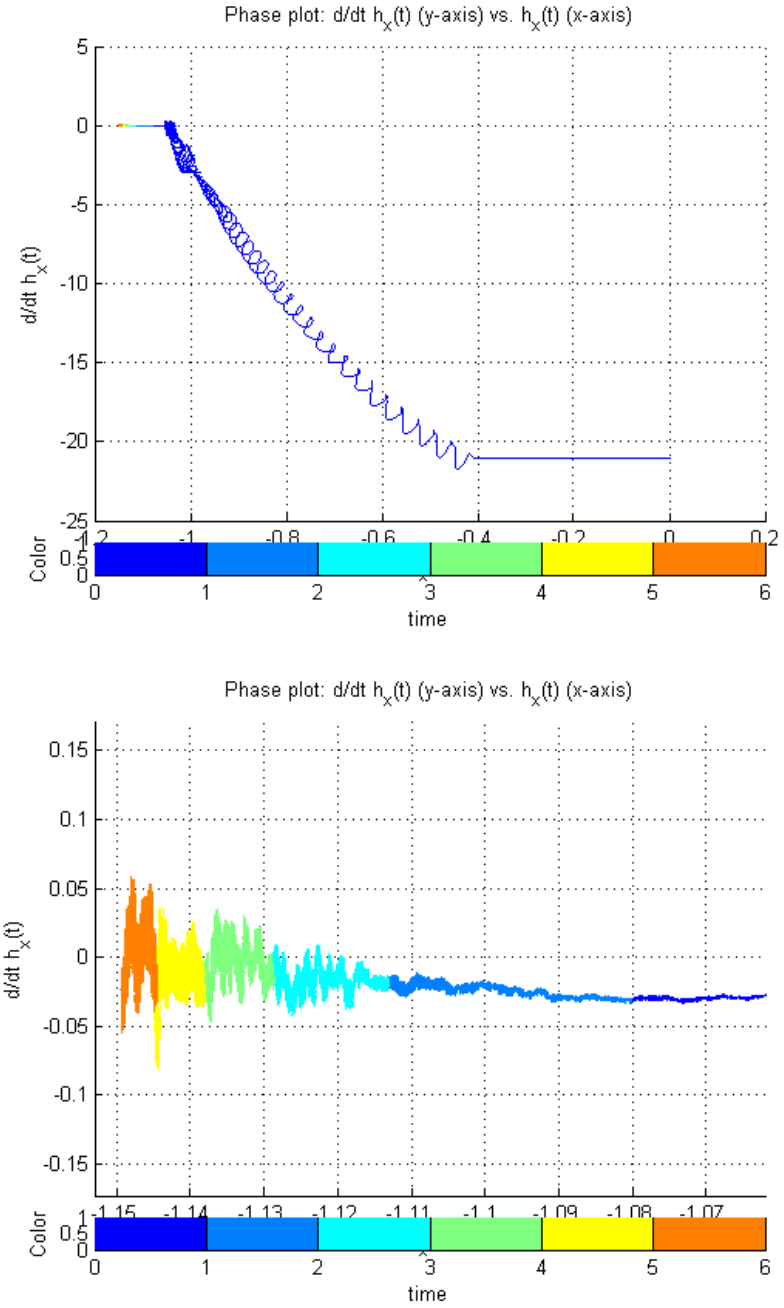
(zoomed in, from top)

Figure 24: Feedback gain phase plot: C16:0 Dihydroceramide (DHCer) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



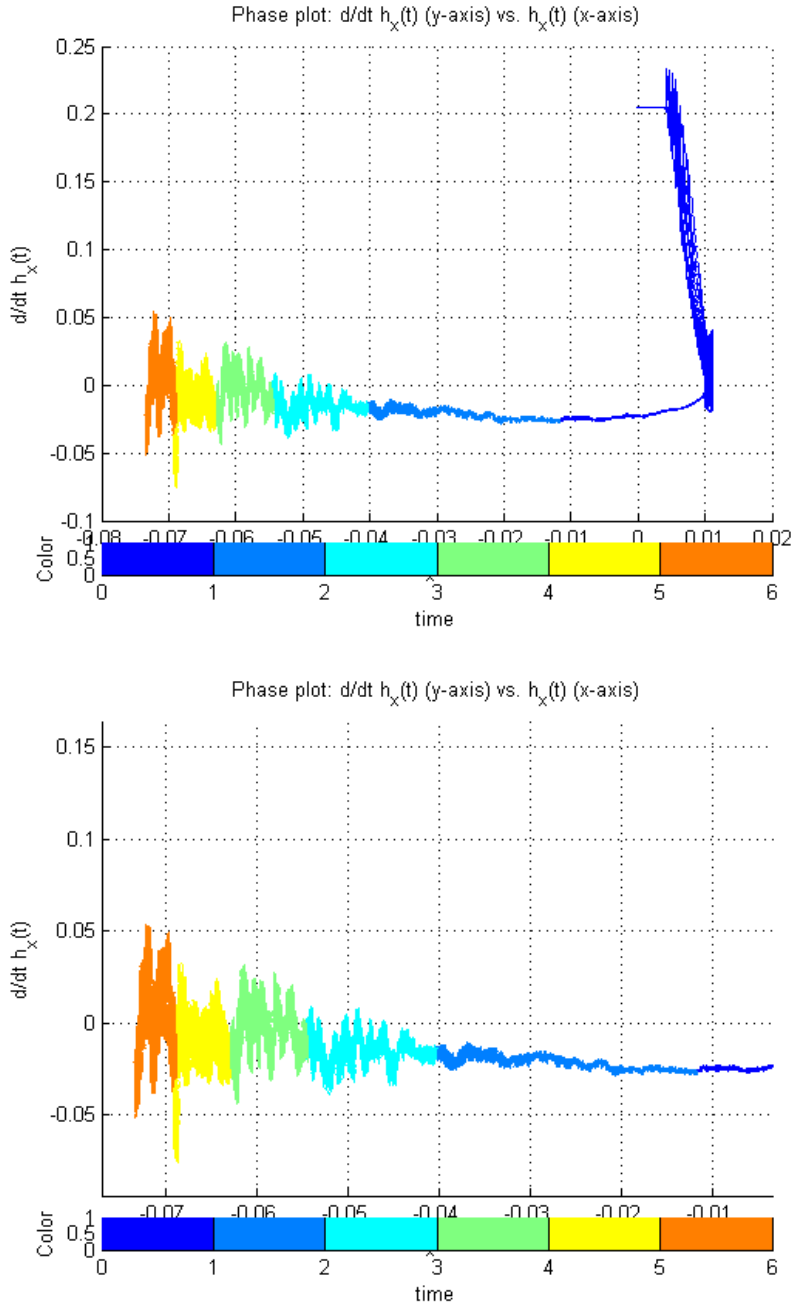
(zoomed in, from top)

Figure 25: Feedback gain phase plot: C16:0 Dihydroglucosylceramide (DHGC) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



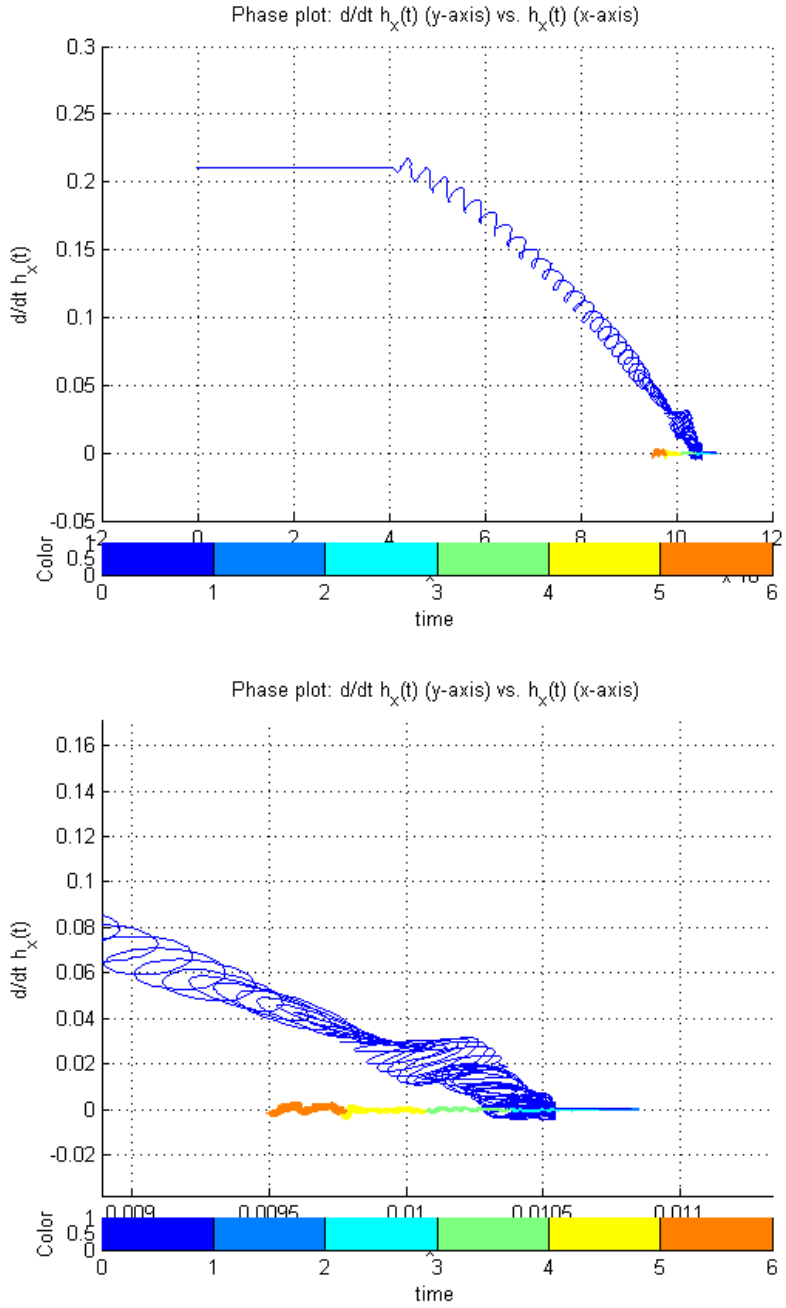
(zoomed in, from top)

Figure 26: Feedback gain phase plot: C16:0 Dihydrosphingomyelin (DHSM) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



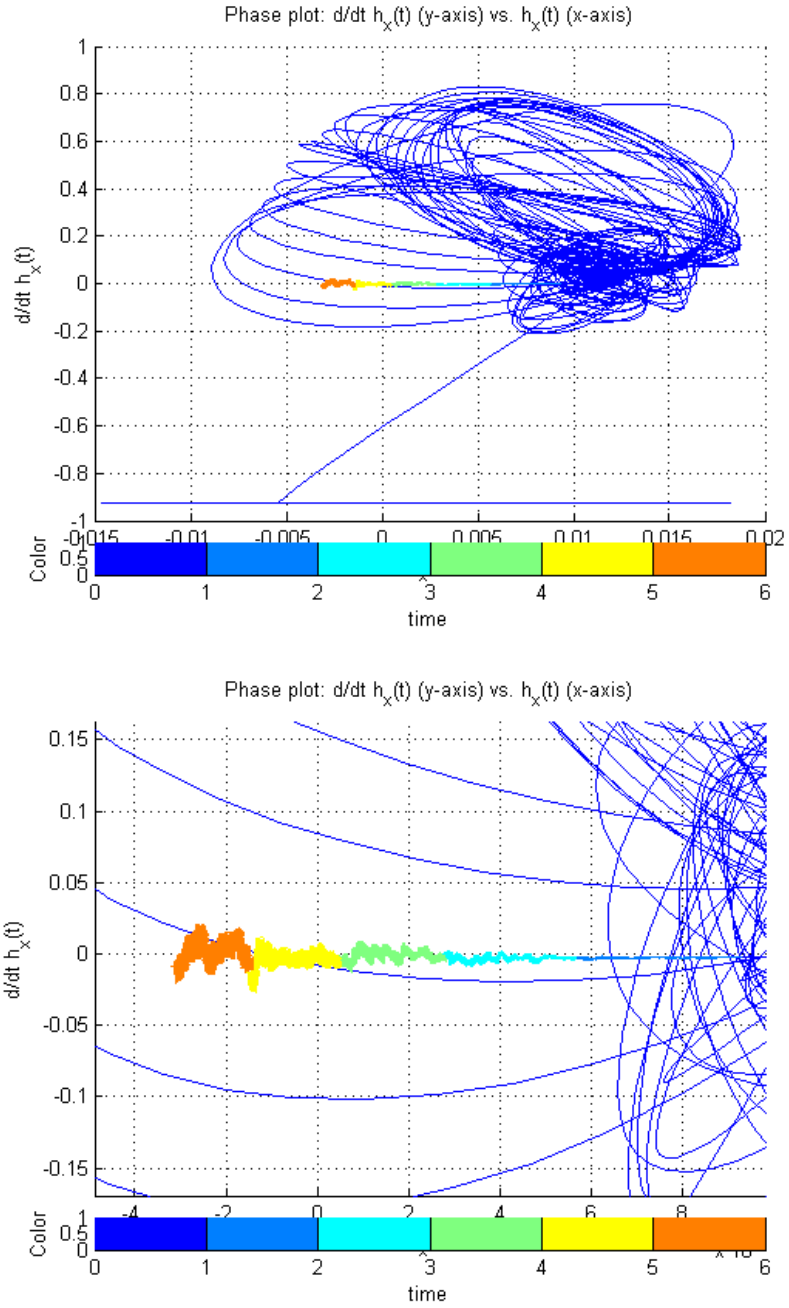
(zoomed in, from top)

Figure 27: Feedback gain phase plot: C16:0 Ceramide (Cer) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



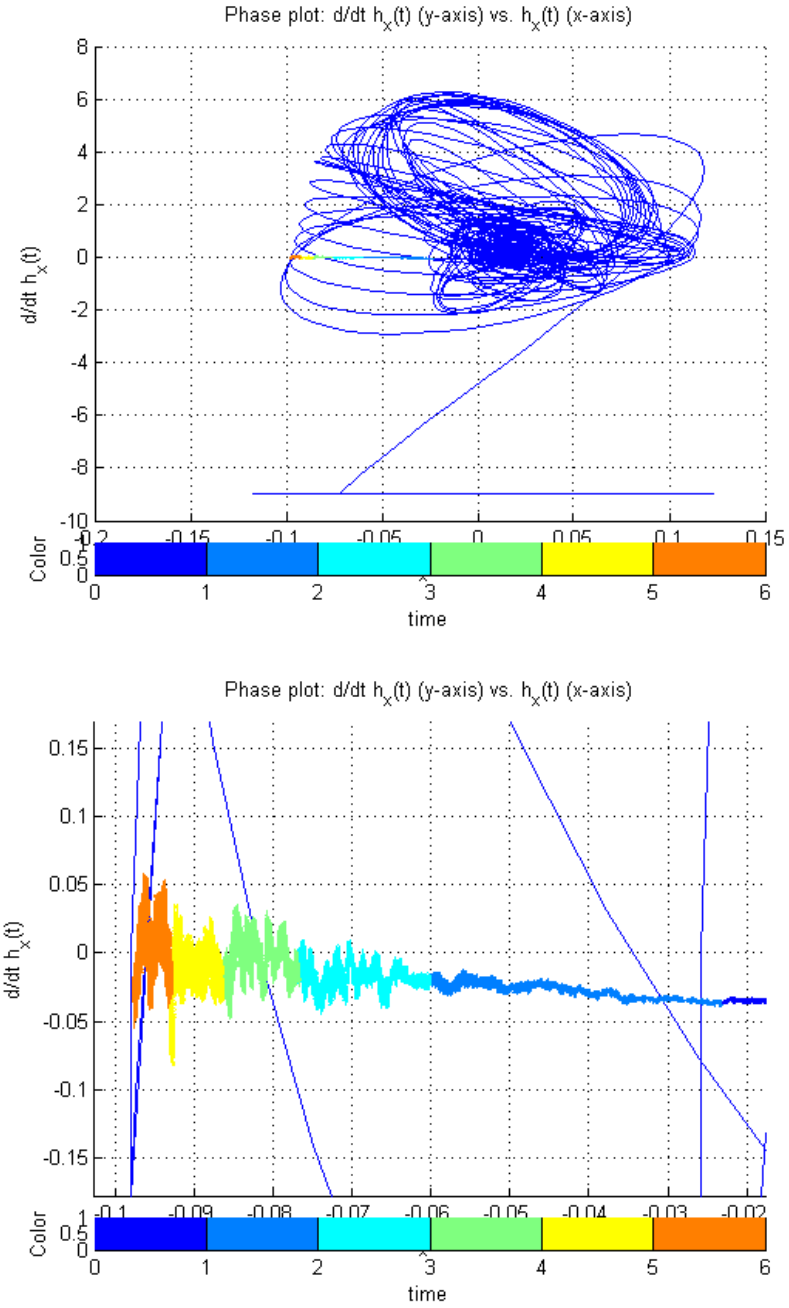
(zoomed in, from top)

Figure 28: Feedback gain phase plot: C16:0 Glucosylceramide (GC) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



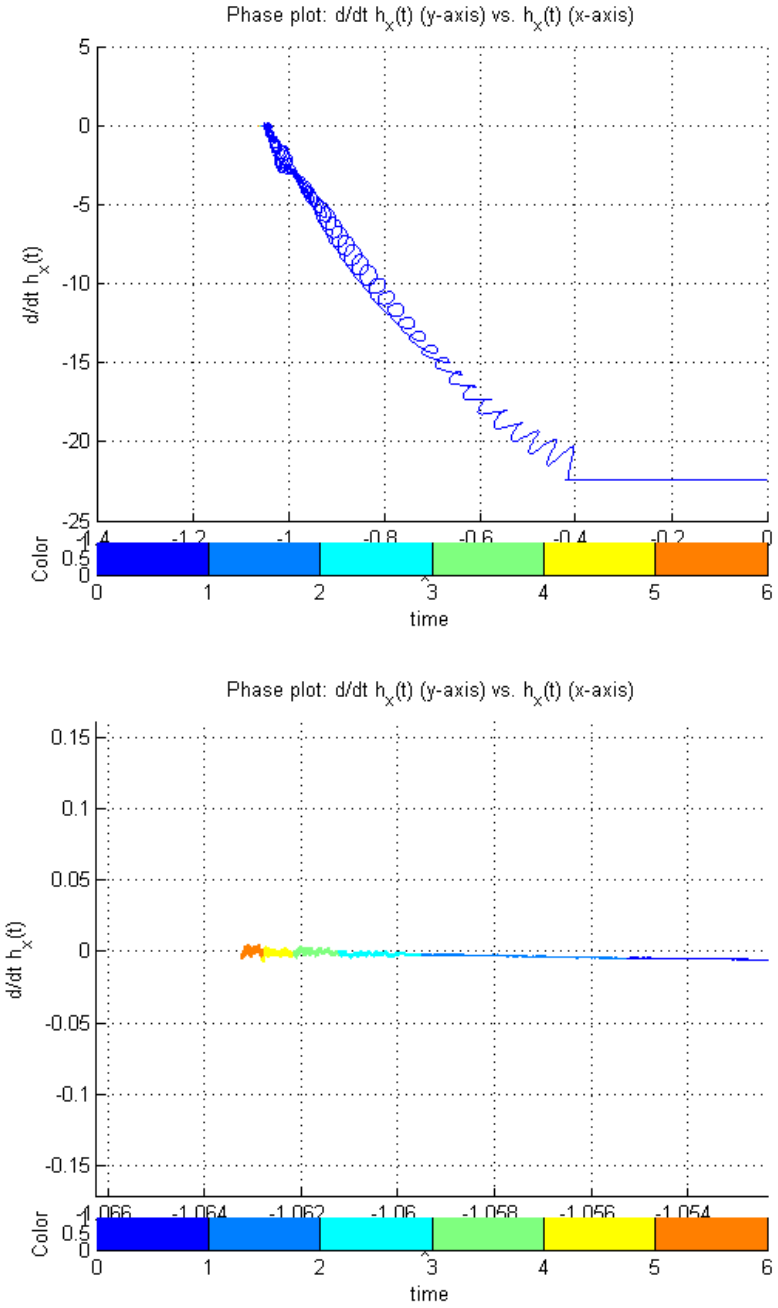
(zoomed in, from top)

Figure 29: Feedback gain phase plot: C16:0 Sphingomyelin (SM) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



(zoomed in, from top)

Figure 30: Feedback gain phase plot: C16:0 Sphingosine (So) - (axes: y, rate of change d/dt; x, gain value; *color*, time)



(zoomed in, from top)

Figure 31: Feedback gain phase plot: C16:0 Sphingosine-phosphate (SoP) - (axes: y, rate of change d/dt; x, gain value; *color*, time)

REFERENCES

[1] Akaike, H., "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[2] Albert, R. and Barabasi, A.-L., "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, pp. 47–97, 2002.

[3] Albert, R., Jeong, H., and Barabasi, A.-L., "Error and attack tolerance of complex networks," *Nature*, vol. 406, pp. 378–382, 2000.

[4] Aldana, M. and CLuzel, P., "A natural class of robust networks," *Proc Natl Acad Sci USA*, vol. 100, no. 15, pp. B710–B714, 2003.

[5] Alderson, D. L. and Doyle, J. C., "Contrasting views of complexity and their implications for network-centric infrastructures," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 40, no. 4, pp. 839–853, 2010.

[6] Almaas, E., "Biological impacts and context of network theory," *The Journal of Experimental Biology*, vol. 210, pp. 1548–1558, 2007.

[7] Alon, U., Surette, M. G., Barkai, N., and Leibler, S., "Robustness in bacterial chemotaxis," *Nature*, vol. 397, pp. 168–171, 1999.

[8] Alvarez-Vasquez, F., Sims, K. J., Cowart, A., Okamoto, Y., Voit, E. O., and Hannun, Y. A., "Simulation and validation of modeled sphingolipid metabolism in Saccharomyces cerevisiae," *Nature*, vol. 433, pp. 425–430, 2005.

[9] Ao, P., "Metabolic network modeling: including stochastic effects," *Computers and Chemical Engineering*, vol. 29, pp. 2297–2303, 2005.

[10] Araujo, R. P., Petricoin, E. F., and Liotta, L. A., "Mathematical modeling of the cancer cell's control circuitry: paving the way to individualized therapeutic strategies," *Current Signal Transduction Therapy*, vol. 2, pp. 145–155, 2007.

[11] Atkins, G. L. and Ninmo, I. A., "Review: current trends in the estimation of Michaelis-Menten parameters.," *Analytical Biochemistry*, vol. 104, no. 1, pp. 1–9, 1980.

[12] Bach, L. and Faure, J.-D., "Role of very-long-chain fatty acids in plant development, when chain length does matter," *Comptes Rendus Biologies*, vol. 333, no. 4, pp. 361–370, 2010.

[13] Baker, M. D., Wolanin, P. M., and Stock, J. B., "Systems biology of bacterial chemotaxis," *Current Opinion in Microbiology*, vol. 9, pp. 187–192, 2006.

[14] Balaji, S., Lakshminarayan, M. I., Aravind, L., and Babu, M. M., "Uncovering a hidden distributed architecture behind scale-free transcriptional regulatory networks," *Journal of Molecular Biology*, vol. 360, pp. 204–212, 2006.

[15] Bar, N. S., "Analysis of protein synthesis dynamic model in eukaryotic cells: input control," *Mathematical Biosciences*, vol. 219, pp. 84–91, 2009.

[16] Barkai, N. and Leibler, S., "Robustness in simple biochemical networks," *Nature*, vol. 387, pp. 913–917, 1997.

[17] Bechhoefer, J., "Feedback for physicists: a tutorial essay on control," *Review of Modern Physics*, vol. 77, pp. 783–836, 2005.

[18] Bedau, M. A., "Weak emergence," in *Philosophical Perspectives: Mind, Causation, and World* (Tomberlin, J., ed.), vol. 11, pp. 375–399, Malden MA: Blackwell, 1997.

[19] Behe, M. J., "The challenge of irreducible complexity," *Natural History*, vol. 111, no. 3, p. 74, 2002.

[20] Behre, J., Wilhelm, T., von Kamp, A., Ruppin, E., and Schuster, S., "Structural robustness of metabolic networks with respect to multiple knockouts," *Journal of Theoretical Biology*, vol. 252, pp. 433–441, 2008.

[21] Berg, J. M., Tymoczko, J. L., and Stryer, L., *Biochemistry.* New York, USA: W. H. Freeman, fifth ed., 2002.

[22] Bhartiya, S., Chaudhary, N., Venkatesh, K. V., Doyle, F. J., and III., "Multiple feedback loop deign in the tryptophan regulatory network of Escherichia coli suggests a paradigm for robust regulation of processes in series," *Journal of The Royal Society Interface*, vol. 3, pp. 383–391, 2006.

[23] Bianchi, A. L. and Gestreau, C., "The brainstem respiratory network: an overview of a half century of research," *Respiratory Physiology & Neurobiology*, vol. 168, pp. 4–12, 2009.

[24] Bor, Y. J., "Optimal pest management and economic threshold," *Agricultural Systems*, vol. 49, pp. 113–133, 1995.

[25] Bor, Y. J., "Some evidence for the existence of dynamic economic thresholds," *Agricultural Systems*, vol. 53, pp. 143–160, 1997.

[26] Bor, Y. J., "Uncertain control of dynamic economic threshold in pest management," *Agricultural Systems*, vol. 78, pp. 105–118, 2003.

[27] Bruggeman, F. J. and Westerhoff, H. V., "Approaches to biosimulation of cellular processes," *Journal of Biological Physics*, vol. 32, pp. 273–288, 2006.

[28] Cariani, P., "Symbols and dynamics in the brain," *Biosystems*, vol. 60, no. 1, 2001.

[29] Carton, J., J.Uhlinger, D., Batheja, A., Derian, C., Ho, G., Argenteri, D., and D'Andrea, M., "Enhanced serine palmitoyltransferase expression in proliferating fibroblasts, transformed cell lines, and human tumors," *Journal of Histochemistry & Cytochemistry*, vol. 51, no. 6, pp. 715–726, 2003.

[30] Celani, A. and Vergassola, M., "Bacterial strategies of chemotaxis response," *Proc Natl Acad Sci USA*, vol. 107, no. 4, pp. 1391–1396, 2010.

[31] Choi, J.-Y. and Martin, C. E., "The Saccharomyces cerevisiae FAT1 gene encodes an acyl-CoA synthetase that is required for maintenance of very long chain fatty acid levels," *Journal of Biological Chemistry*, vol. 274, no. 8, pp. 4671–4683, 1999.

[32] Cloutier, M. and Wellstead, P., "The control systems structures of energy metabolism," *Journal of The Royal Society Interface*, vol. 7, pp. 651–665, 2010.

[33] Csete, M. E. and Doyle, J. C., "Reverse engineering of biological complexity," *Science*, vol. 295, pp. 1664–1669, 2002.

[34] Currie, D. J., "Estimating Michaelis-Menten parameters: bias, variance and experimental design," *Biometrics*, vol. 38, no. 4, pp. 907–919, 1982.

[35] Cuvillier, O., "Sphingosine in apoptosis signaling," *BBA - Molecular and Cell Biology of Lipids*, vol. 1585, no. 2, 2002.

[36] Daniels, B. C., Chen, Y.-J., Sethna, J. P., Gutenkunst, R. N., and Myers, C. R., "Sloppiness, robustness, and evolvability in systems biology," *Current Opinion in Biotechnology*, vol. 19, no. 4, pp. 389–395, 2008.

[37] Davis, G. W., "Homeostatic control of neural activity: from phenomenology to molecular design," *Annual Reviews in Neuroscience*, vol. 39, pp. 307–323, 2006.

[38] Dickson, R. C., "New insights into sphingolipid metabolism and function in budding yeast," *Journal of Lipid Research*, vol. 49, no. 5, pp. 909–921, 2008.

[39] Domokos, G., "My lunch with Arnold," *Mathematical Intelligencer*, vol. 28, no. 4, pp. 31–33, 2006.

[40] Domokos, G. and Várkonyi, P. L., "Geometry and self-righting of turtles," *Proceedings of the Royal Society B*, vol. 275, no. 1630, pp. 11–17, 2008.

[41] Doyle, J. C., Alderson, D., Li, L., Low, S., Roughan, M., Shalunov, S., Tanaka, R., and Willinger, W., "The "robust yet fragile" nature of the internet," *Proc Natl Acad Sci USA*, vol. 102, no. 41, pp. 14497–14502, 2005.

[42] Edwards, J., Ibarra, R., and Palsson, B., "In silico predictions of escherichia coli metabolic capabilities are consistent with experimental data," *Nature Biotechnology*, vol. 19, pp. 125–130, 2001.

[43] El-Samad, H. and Khammash, M., "Modeling and analysic of gene regulatory network using feedback control theory," *International Journal of Systems Science*, vol. 41, no. 1, pp. 17–33, 2010.

[44] El-Samad, H., Kurata, H., Doyle, J. C., Gross, C. A., and Khammash, M., "Surviving heat shock: control strategies for robustness and performance," *Proceedings of the National Academy of Science USA*, vol. 102, no. 8, pp. 2736–2741, 2005.

[45] Fahy, E., Subramaniam, S., Brown, H. A., Glass, C. K., Merrill, A. H., Jr., Murphy, R. C., Raetz, C. R. H., Russell, D. W., Seyama, Y., Shaw, W., Shimizu, T., Spener, F., van Meer, G., VanNieuwenhze, M. S., White, S. H., Witztum, J. L., and Dennis, E. A., "A comprehensive classification system for lipids," *Journal of Lipid Research*, vol. 46, pp. 839–862, 2005.

[46] Fell, D., "Metabolic control analysis: a survey of its theoretical and experimental development," *Biochemical Journal*, vol. 286, pp. 313–330, 1992.

[47] Forrester, J. S., Milne, S. B., Ivanova, P. T., and Brown, H. A., "Computational lipidomics: a multiplexed analysis of dynamic changes in membrane lipid composition during signal transduction," *Molecular Pharmacology*, vol. 65, no. 4, pp. 813–821, 2004.

[48] Fossas, E., Ros, R. M., and Fabregat, J., "Sliding mode control in a bioreactor model," *Journal of Mathematical Chemistry*, vol. 30, no. 2, pp. 203–218, 2001.

[49] French, A. S., "The systems analysis approach to mechanosensory coding," *Biological Cybernetics*, vol. 100, no. 6, pp. 417–426, 2009.

[50] Friedlander, T. and Brenner, N., "Adaptive response by state-dependent inactivation," *Proceedings of the National Academy of Sciences USA*, vol. 106, no. 52, pp. 22558–22563, 2009.

[51] Frost, H., "Cybernetic aspects of bone modeling and remodeling, with special reference to osteoporosis and whole-bone strength," *American Journal of Human Biology*, vol. 13, pp. 235–248, 2001.

[52] Garner, A. E., Smith, D. A., and Hooper, N. M., "Sphingomyelin chain length influences the distribution of GPI-anchored proteins in rafts in supported lipid bilayers," *Molecular Membrane Biology*, vol. 24, no. 3, pp. 233–242, 2007.

[53] Gehrmann, E., Glaber, C., Jin, Y., Sendhoff, B., Drossel, B., and Hamacher, K., "Robustness of glycolysis in yeast to internal and external noise," *Physical Review E*, vol. 84, p. 201913, 2011.

[54] Glatz, J., Luiken, J., van Nieuwenhoven, F., and can der Vusse, G., "Molecular mechanism of cellular uptake and translocation of fatty acids," *Prostaglandins, Leukotrienes and Essential Fatty Acids*, vol. 57, no. 1, pp. 3–9, 1997.

[55] Goldstein, B. N., Ermakov, G., Centelles, J. J., Westerhoff, H. V., and Cascante, M., "What makes biochemical networks tick? a graphical tool for the identification of oscillophores," *European Journal of Bioechemistry*, vol. 271, pp. 3877–3887, 2004.

[56] Goulian, M., "Robust control in bacterial regulatory circuits," *Current Opinion in Microbiology*, vol. 7, pp. 198–202, 2004.

[57] Grafahrend-Belau, E., Schreiber, F., Koschutzki, D., and Junker, B. H., "Flux balance analysis of barley seeds: a computational approach to study systemic properties of central metabolism," *Plant Physiology*, vol. 149, pp. 585–598, 2009.

[58] Gulberg, C. M. and Waage, P., "Studies concerning affinity (translated)," *Journal of Chemical Education*, vol. 63, no. 12, pp. 1044–1047, 1986.

[59] Gupta, S., Maurya, M. R., Merrill, A. H., Jr., Glass, C. K., and Subramaniam, S., "Integration of lipidomics and transcriptomics data towards a systems biology model of sphingolipid metabolism," *BMC Systems Biology*, vol. 5, p. 26, 2011.

[60] Guyton, A. C., Coleman, T. G., and Granger, H. J., "Circulation: overall regulation," *Annual Reviews in Physiology*, vol. 34, pp. 13–44, 1972.

[61] Guzun, R. and Saks, V., "Application of the principles of systems biology and Wiener's cybernetics for analysis of regulation of energy fluxes in muscle cells in vivo," *International Journal of Molecular Science*, vol. 11, no. 3, pp. 982–1019, 2010.

[62] H. J. Huber, H. D., Wenus, J., Kilbride, S. M., and Prehn, J. H., "Mathematical modeling of the mitochondrial apoptosis pathway," *Biochemica et Biophysics Acta*, vol. 1813, pp. 608–615, 2011.

[63] Hansen, C. H., Endres, R. G., and Wingreen, N. S., "Chemotaxis in escherichia coli: a molecular model for robust precise adaptation," *PLoS Computational Biology*, vol. 4, no. 1, pp. 14–27, 2008.

[64] Hasty, J., Isaacs, F., Dolnik, M., McMillen, D., and Collins, J. J., "Designer gene networks: towards fundamental cellular control," *Chaos*, vol. 11, no. 1, pp. 207–220, 2001.

[65] Heinrich, R. and Rapoport, T. A., "A linear steady-state treatment of enzymatic chains. General properties, control and effector strength.," *European Journal of Biochemistry*, vol. 42, pp. 89–95, 1974.

[66] Hellerstein, M. K., "Exploiting complexity and robustness of network architecture for drug discovery," *The Journal of Pharmacology and Experimental Therapeutics*, vol. 325, pp. 1–9, 2008.

[67] Holz, R. W. and Fisher, S. K., "Synaptic transmission and cellular signaling: an overview," in *Basic neurochemistry: molecular, cellular and medical aspects* (Siegel, G. J., Agranoff, B. W., Albers, R. W., Fisher, S. K., and Uhler, M. D., eds.), part 2, Philadelphia: Lippincott-Raven, sixth ed., 1999.

[68] Hopkins, A., "Network pharmacology: the next paradigm in drug discovery," *Nature Chemical Biology*, vol. 4, no. 11, pp. 682–690, 2008.

[69] Hurvich, C. M., Simonoff, J. S., and Tsai, C.-L., "Smoothing parameter selection in nonparametric regression using an improved akaike information criterion," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 60, no. 2, pp. 271–293, 1998.

[70] Jing, Z., Lin, T., Hong, Y., Luo, J.-H., Cao, Z.-W., and Li, Y.-X., "The effects of degree correlations on network topologies and robustness," *Chinese Physics*, vol. 16, no. 12, pp. 3571–3581, 2007.

[71] Joslyn, C., "The semiotics of control and modeling relations in complex systems," *BioSystems*, vol. 60, pp. 131–148, 2001.

[72] Kacher, Y. and Futerman, A. H., "Genetic diseases of sphingolipid metabolism: pathological mechanisms and therapeutic options," *FEBS Letters*, vol. 580, no. 23, pp. 5510–5517, 2006.

[73] Kacser, H. and Burns, J. A., "The control of flux," *Symposia of the Society for Experimental Biology*, vol. 27, pp. 65–104, 1973.

[74] Kauffman, K. J., Prakash, P., and Edwards, J. S., "Advances in flux balance analysis," *Current Opinion in Biotechnology*, vol. 14, no. 5, pp. 491–496, 2003.

[75] Kiehn, O., "Locomotor circuits in the mammalian spinal cord," *Annual Review of Neuroscience*, vol. 29, pp. 279–306, 2006.

[76] Kim, J.-S., Valeyev, N. V., Postlethwaite, I., P. Heslop-Harrison, and, K.-H. C., and Bates, D. G., "Analysis and extension of a biochemical network model using robust control theory," *International Journal of Robust and Nonlinear Control*, vol. 20, pp. 1017–1026, 2010.

[77] Kitano, H., "Systems biology: a brief overview," *Science*, vol. 295, pp. 1662–1664, 2002.

[78] Kitano, H., "Biological robustness," *Nature Reviews Genetics*, vol. 5, pp. 826–837, 2004.

[79] Kitano, H., "Cancer as a robust system," *Nature Reviews Genetics*, vol. 4, pp. 227–235, 2004.

[80] Kitano, H., "A robustness-based approach to systems-oriented drug design," *Nature Reviews Drug Discovery*, vol. 6, no. 3, pp. 202–210, 2007.

[81] Kitano, H., "Towards a theory of biological robustness," *Molecular Systems Biology*, vol. 3, p. 137, 2007.

[82] Kitano, H., "Violations of robustness trade-offs," *Molecular Systems Biology*, vol. 6, p. 384, 2010.

[83] Kitano, H. and Oda, K., "Robustness trade-offs and host-microbial symbiosis in the immune system," *Molecular Systems Biology*, p. 2006.0022, 2006.

[84] Kitano, H. and Oda, K., "Self-extending symbiosis: a mechanism for increasing robustness through evolution," *Biological Theory*, vol. 1, no. 1, pp. 61–66, 2006.

[85] Kochel, B., "Control of the first-line human defence system: an autocatalytic model," *Kybernetes*, vol. 28, no. 4, pp. 430–440, 1999.

[86] Kolter, T. and Sandhoff, K., "Sphingolipid metabolism diseases," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1758, no. 12, pp. 2057–2079, 2006.

[87] Kwon, Y.-K. and Cho, K., "Analysis of feedback loops and robustness in network evolution based on boolean models," *BMC Bioinformatics*, vol. 8, p. 430, 2007.

[88] Lahiri, S., Park, H., Laviad, E. L., Lu, X., Bittman, R., and Futerman, A. H., "Ceramide synthesis is modulated by the sphingosine analog FTY720 via a mixture of uncompetitive and noncompetitive inhibition in an acyl-CoA chain length-dependent manner," *Journal of Biological Chemistry*, vol. 284, no. 24, pp. 16090–16098, 2009.

[89] Larhlimi, A., Blachon, S., Selbig, J., and Nikoloski, Z., "Robustness of metabolic networks: a review of existing definitions," *BioSystems*, vol. 106, pp. 1–8, 2011.

[90] LaSalle, J. P., "Recent advances in Liapunov stability theorem," *SIAM Review*, vol. 6, no. 1, pp. 1–11, 1964.

[91] Lauffenburger, D. A., "Cell signaling pathways as control modules: complexity for simplicity?," *Proceedings of the National Academy of Sciences of the USA*, vol. 97, no. 10, pp. 5031–5033, 2000.

[92] Layek, R., Datta, A., Pal, R., and Dougherty, E. R., "Adaptive intervention in probabilistic boolean networks," *Bioinformatics*, vol. 25, no. 16, pp. 2042–2048, 2009.

[93] León, D. and Markel, S., eds., *In silico technologies in drug target identification and validation*. Boca Raton FL, USA: Tayler & Francis Group, 2006.

[94] Lesne, A., "Complex networks: from graph theory to biology," *Letters in Mathematical Physics*, vol. 78, pp. 235–262, 2006.

[95] Lesne, A., "Robustness: confronting lessons from physics and biology," *Biological Reviews*, vol. 83, pp. 509–532, 2008.

[96] Lombardi, A. and Hornquist, M., "Controllability analysis of networks," *Physical Review E*, vol. 75, p. 056110, 2007.

[97] Lorenz, D. M., Jeng, A., and Deem, M. V., "The emergence of modularity in biology," *Physics of Life Reviews*, vol. 8, pp. 129–160, 2011.

[98] Luiken, J., van Nieuwenhoven, F., America, G., van der Vusse, G., and Glatz, J., "Uptake and metabolism of palmitate by isolated cardiac myocytes, from adult rats: involvement of sarcolemmal proteins," *Journal of Lipid Research*, vol. 38, pp. 745–758, 1997.

[99] Lund, E. W., "Guldberg and Waage and the law of mass action," *Journal of Chemical Education*, vol. 65, no. 10, pp. 548–550, 1965.

[100] Luni, C. and Doyle, D. J., "Robust multi-drug therapy design and application to insulin resistance in type 2 diabetes," *International Journal of Robust and Nonlinear Control*, vol. 21, pp. 1730–1741, 2011.

[101] Luni, C., Shoemaker, J. E., Sanft, K. R., Petzold, L. R., and Doyle, F. J., "Confidence from uncertainty - a multi-target drug screening method from robust control theory," *BMC Systems Biology*, vol. 4, p. 161, 2010.

[102] Lyapunov, A. M., *General problem of the stability of motion*. London, England: Taylor & Francis Ltd, 1992. Translated by A. T. Fuller.

[103] Ma, W., Trusina, A., El-Samad, H., Lim, W. A., and Tang, C., "Defining network topologies that can achieve biochemical adaptation," *Cell*, vol. 138, no. 4, pp. 760–773, 2009.

[104] Madison, K. C., Swartzendruber, D. C., Wertz, P. W., and Downing, D. T., "Sphingolipid metabolism in organotypic mouse keratinocyte cultures," *Journal of Investigative Dermatology*, vol. 95, no. 6, pp. 657–664, 1990.

[105] McArthur, M., Atshaves, B., Frolov, A., Foxworth, W., Kier, A., and Schroeder, F., "Cellular uptake and intracellular trafficking of long chain fatty acids," *Journal of Lipid Research*, vol. 40, pp. 1371–1383, 1999.

[106] Medler, T. R., Petrusca, D. N., Lee, P. J., Hubbard, W. C., Berdyshev, E. V., Skirball1, J., Kamocki, K., Schuchman, E., Tuder, R. M., and Petrache, I., "Apoptotic sphingolipid signaling by ceramides in lung endothelial cells," *American Journal of Respiratory Cell and Molecular Biology*, vol. 38, no. 6, pp. 639–646, 2008.

[107] Megha, Sawatzki, P., Kolter, T., Bittman, R., and London, E., "Effect of ceramide N-acyl chain and polar headgroup structure on the properties of ordered lipid domains (lipid rafts)," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1768, no. 9, pp. 2205–2212, 2007.

[108] Menolascina, F., Bellomo, D., Maiwald, T., Bevilacqua, V., Ciminelli, C., Paradiso, A., and Tommasi, S., "Developing optimal input design strategies in cancer systems biology with applications to microfluidic device engineering," *BMC Bioinformatics*, vol. 10, p. S4, 2009.

[109] Merrill, A. H., "De novo sphingolipid biosynthesis: a necessary, but dangerous, pathway," *Journal of Biological Chemistry*, vol. 277, no. 29, pp. 25843–25846, 2002.

[110] Merrill, A. H., Jr., Sullards, M. C., Allegood, J. C., Kelly, S., and Wang, E., "Sphingolipidomics: high-throughput, structure-specific, and quantitative analysis of sphingolipids by liquid chromatography tandem mass spectrometry," *Methods*, vol. 36, no. 2, pp. 207–224, 2005.

[111] Merrill, A. H., Jr., Sullards, M., Wang, E., Voss, K., and Riley, R., "Sphingolipid metabolism: roles in signal transduction and disruption by fumonisins," *Environmental Health Perspectives*, vol. 109, pp. 283–289, 2001.

[112] Merrill, A. H., Jr., Wang, M. D., Park, M., and Sullards, M. C., "(Glyco)sphingolipidology: an amazing challenge and opportunity for systems biology," *Trends in Biochemical Sciences*, vol. 32, no. 10, pp. 457–468, 2007.

[113] Miller, K. R., "The flaw in the mouse trap," *Natural History*, vol. 111, no. 3, p. 75, 2002.

[114] Min, Y., Jin, X., Chen, M., Pan, Z., Ge, Y., and Chang, J., "Pathway knockout and redundancy in metabolic networks," *Journal of Theoretical Biology*, vol. 270, pp. 63–69, 2011.

[115] Morohashi, M., Winn, A. W., Borisuk, M. T., Bolouri, H., Doyle, J., and Kitano, H., "Robustness as a measure of plausibility in models of biochemical networks," *Journal of Theoretical Biology*, vol. 216, pp. 19–30, 2002.

[116] Novgorodov, S. A., Gudz, T. I., and Obeid, L. M., "Long-chain ceramide is a potent inhibitor of the mitochondrial permeability transition pore," *Journal of Biological Chemistry*, vol. 283, no. 36, pp. 24707–24717, 2008.

[117] Nyholm, T. K. M., Grandell, P.-M., Westerlund, B., and Slotte, J. P., "Sterol affinity for bilayer membranes is affected by their ceramide content and the ceramide chain length," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1798, no. 5, pp. 1008–1013, 2010.

[118] Ogretmen, B. and Hannun, Y. A., "Biologically active sphingolipids in cancer pathogenesis and treatment," *Nature Reviews Cancer*, vol. 4, no. 8, pp. 604–616, 2004.

[119] Orth, J. D., Thiele, I., and Palsson, B. O., "What is flux balance analysis?," *Nature Biotechnology*, vol. 28, pp. 245–248, 2010.

[120] Paricharttanakul, N. M., Ye, S., Menefee, A. L., Javid-Majd, F., Sacchettini, J. C., and Reinhart, G. D., "Kinetic and structural characterization of phosphofructokinase from Lactobacillus bulgaricus," *Biochemistry*, vol. 44, no. 46, pp. 15280–15286, 2005.

[121] Pickering, W. D., Nikiforuk, P. N., and Merriman, J. E., "Analogue computer model of the human cardiovascular control system," *Medical & Biological Engineering*, vol. 7, no. 4, pp. 401–410, 1969.

[122] Pollard, T. D., Earnshaw, W. C., and Lippincott-Schwartz, J., *Cell biology*. Philadelphia, PA, USA: Saunders Elsevier, second ed., 2008.

[123] Pruett, S. T., Bushnev, A., Hagedorn, K., Adiga, M., Haynes, C. A., Sullards, M. C., Liotta, D. C., and Merrill, A. H., "Biodiversity of sphingoid bases ("sphingosines") and related amino alcohols," *Journal of Lipid Research*, vol. 49, no. 8, pp. 1621–1639, 2008.

[124] Quo, C. F., Kaddi, C., Phan, J., Amin, Z., Xu, M., Wang, M. D., and Alterovitz, G., "Reverse engineering biomolecular systems using high-throughput omic data: challenges, progress, and opportunities," *Briefings in Bioinformatics*, vol. 13, pp. 430–445, 2012.

[125] Quo, C. F., Moffitt, R. A., Merrill, A. H., Jr., and Wang, M. D., "Adaptive control model reveals systematic feedback and key molecules in metabolic pathway regulation," *Journal of Computational Biology*, vol. 18, no. 2, pp. 169–182, 2011.

[126] Quo, C. F. and Wang, M. D., "Quantitative analysis of numerical solvers for oscillatory biomolecular system models," *BMC Bioinformatics*, vol. 9, p. S17, 2008.

[127] Quo, C. F. and Wang, M. D., "Biological interpretation of model-reference adaptive control in a mass action kinetics metabolic pathway model," in *IEEE International Conference on Bioinformatics and Biomedicine*, pp. 265–268, Nov. 2011.

[128] Radulescu, O., Gorban, A. N., Zinovyev, A., and Lilienbaum, A., "Robust simplifications of multiscale biochemical networks," *BMC Systems Biology*, vol. 2, p. 86, 2008.

[129] Raman, K. and Chandra, N., "Flux balance analysis of biological systems: applications and challenges," *Briefings in Bioinformatics*, vol. 10, no. 4, pp. 435–449, 2009.

[130] Rao, C. V., Wolf, D. M., and Arkin, A. P., "Control, exploitation and tolerance of intracellular noise," *Nature*, vol. 420, pp. 231–238, 2002.

[131] Rocco, A., "Stochastic control of metabolic pathways," *Physical Biology*, vol. 6, p. 016002, 2009.

[132] Ross, J., "From the determination of complex reaction mechanisms to systems biology," *Annual Reviews in Biochemistry*, vol. 77, pp. 479–494, 2008.

[133] Runge, M. C. and Johnson, F. A., "The importance of functional form in optimal control solutions of problems in population dynamics," *Ecology*, vol. 83, no. 5, pp. 1357–1371, 2002.

[134] Savageau, M., "Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions," *Journal of Theoretical Biology*, vol. 25, no. 3, pp. 365–369, 1969.

[135] Savageau, M., "Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation," *Journal of Theoretical Biology*, vol. 25, no. 3, pp. 370–379, 1969.

[136] Savageau, M. A., "Design principles for elementary gene circuits," *Chaos*, vol. 11, p. 142, 2001.

[137] Schmidt, M. and Lipson, H., "Distilling free-form natural laws from experimental data," *Science*, vol. 324, no. 5923, pp. 81–85, 2009.

[138] Schmitt, B. M., "The concept of "buffering" in systems and control theory: from metaphor to math," *ChemBioChem*, vol. 5, pp. 1384–1392, 2004.

[139] Shabbits, J. A. and Mayer, L. D., "Intracellular delivery of ceramide lipids via liposomes enhances apoptosis in vitro," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1612, no. 1, pp. 98–106, 2003.

[140] Shah, N. A., Moffitt, R. A., and Wang, M. D., "Modified genetic algorithm for parameter selection of compartmental models," in *Proc. 29th Annual International Conference of the IEEE EMBS*, (Cite Internationale, Lyon, France), pp. 143–146, Aug. 2007.

[141] Shimizu, T. S., Tu, Y., and Berg, H. C., "A modular gradient-sensing network for chemotaxis in *Escherichia coli* revealed by responses to time-varying stimuli," *Molecular Systems Biology*, vol. 6, p. 382, 2010.

[142] Shin, Y.-J. and Bleris, L., "Linear control theory for gene network modeling," *PLoS ONE*, vol. 5, no. 9, p. e12785, 2010.

[143] Shinar, G. and Feinberg, M., "Design principles for robust biochemical reaction networks: what works, what cannot work, and what might almost work," *Mathematical Biosciences*, vol. 231, pp. 39–48, 2011.

[144] Simpson, M. A., Cross, H., Proukakis, C., Priestman, D. A., Neville, D. C. A., Reinkensmeier, G., Wang, H., Wiznitzer, M., Gurtz, K., Verganelaki, A., Pryde, A., Patton, M. A., Dwek, R. A., Butters, T. D.,

Platt, F. M., and Crosby, A. H., "Infantile-onset symptomatic epilepsy syndrome caused by a homozygous loss-of-function mutation of GM3 synthase," *Nature Genetics*, vol. 36, pp. 1225–1229, 2004.

[145] Solomon, J., Sharma, K., Wei, L., Fuijita, T., and Shi, Y., "A novel role for sphingolipid intermediates in activation-induced cell death in T cells," *Cell Death and Differentiation*, vol. 10, pp. 193–202, 2003.

[146] Sontag, E. D., "Some new directions in control theory inspired by systems biology," *IEE Systems Biology*, vol. 1, pp. 9–18, 2004.

[147] Soreq, H. and Seidman, S., "Acetylcholinesterase – new roles for an old actor," *Nature Reviews Neuroscience*, vol. 2, pp. 294–302, 2001.

[148] Spiegel, S. and Milstien, S., "Sphingosine 1-phosphate, a key cell signaling molecule," *The Journal of Biological Chemistry*, vol. 277, pp. 25851–25854, 2002.

[149] Spiegel, S. and Milstien, S., "Sphingosine-1-phosphate: An enigmatic signaling lipid," *Nature Reviews Molecular Cell Biology*, vol. 4, pp. 397–407, 2003.

[150] Stelling, J., "Mathematical models in microbial systems biology," *Current Opinion in Microbiology*, vol. 7, no. 5, pp. 513–518, 2004.

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

[152] Sullards, M. C., Allegood, J. C., Kelly, S., Wang, E., Haynes, C. A., Park, H., Chen, Y., and Merrill, A. H., "Structure-specific, quantitative methods for analysis of sphingolipids by liquid chromatography-tandem mass spectrometry: "inside-out" sphingolipidomics.," *Methods in Enzymology*, vol. 432, pp. 83–115, 2007.

[153] Sweetlove, L. I. and Fernie, A. R., "Regulation of metabolic networks: understanding metabolic complexity in the systems biology era," *New Phytologist*, vol. 168, no. 1, pp. 9–24, 2005.

[154] Szalay, M. S., Kovacs, I. A., Korcsmaros, T., Bode, C., and Csermely, P., "Stress-induced rearrangements of cellular networks: consequences for protection and drug design," *FEBS Letters*, vol. 581, pp. 3675–3680, 2007.

[155] Tauber, A. I., "The immune system and its ecology," *Philosophy of Science*, vol. 75, no. 2, pp. 224–245, 2008.

[156] Tomlin, C. J. and Axelrod, J. D., "Understanding biology by reverse engineering the control.," *Proceedings of the National Academy of Sciences of the USA*, vol. 102, pp. 4219–4220, 2005.

[157] Tsankov, A. M., Brown, C. R., Yu, M. C., Win, M. Z., Silver, P. A., and Casolari, J. M., "Communication between levels of transcriptional control improves robustness and adaptivity," *Molecular Systems Biology*, vol. 2, p. 65, 2006.

[158] Ulrich, R. and Friend, S. H., "Toxicogenomics and drug discovery: will new technologies help us produce better drugs?," *Nature Reviews Drug Discovery*, vol. 1, pp. 84–88, 2002.

[159] van Riel, N. A. W., "Dynamic modeling and analysis of biochemical networks: mechanism-based models and model-based experiments," *Briefings in Bioinformatics*, vol. 7, no. 4, pp. 364–374, 2006.

[160] Várkonyi, P. L. and Domokos, G., "Mono-monostatic bodies: the answer to Arnold's question," *Mathematical Intelligencer*, vol. 28, no. 4, pp. 34–38, 2006.

[161] Várkonyi, P. L. and Domokos, G., "Static equilibria of rigid bodies: dice, pebbles, and the Poincare-Hopf Theorem," *Journal of Nonlinear Science*, vol. 16, pp. 255–281, 2006.

[162] Varma, A. and Palsson, B. O., "Metabolic flux balancing: basic concepts, scientific and practical use," *Nature Biotechnology*, vol. 12, pp. 994–998, 1994.

[163] Venkatesh, K. V., Bhartiya, S., and Rubela, A., "Multiple feedback loops are key to a robust dynamic performance of tryptophan regulation in Escherichia coli," *FEBS Letters*, vol. 563, pp. 234–240, 2004.

[164] Vesper, H., Schmelz, E.-M., Nikolova-Karakashian, M. N., Dillehay, D. L., Lynch, D. V., Merrill, A. H., and Jr., "Sphingolipids in food and the emerging importance of sphingolipids to nutrition.," *Journal of Nutrition*, vol. 129, pp. 1239–1250, 1999.

[165] Voet, D. and Voet, J. G., *Biochemistry*. USA: John Wiley & Sons, Inc., third ed., 2004.

[166] Wang, Z. and Zhang, J., "Abundant indispensable redundancies in cellular metabolic networks," *Genome Biology and Evolution*, vol. 1, pp. 23–33, 2009.

[167] Wei, J., *Study of serine palmitoyltransferase and de novo synthesis of sphingolipids*. PhD dissertation, Georgia Institute of Technology, School of Biology, 2009.

[168] Wellstead, P., Bullinger, E., Kalamatianos, D., Mason, O., and Verwoerd, M., "The role of control and system theory in systems biology," *Annual Reviews in Control*, vol. 32, pp. 33–47, 2008.

[169] Wiener, N., *Cybernetics: or control and communication in the animal and the machine*. Cambridge MA, USA: MIT Press, second ed., 1965.

[170] Williams, B. K., "Optimal management of non-Markovian biological populations," *Ecological Modelling*, vol. 200, pp. 234–242, 2007.

[171] Wolf, D. M. and Arkin, A. P., "Motifs, modules, and games in bacteria," *Current Opinion in Microbiology*, vol. 6, pp. 125–134, 2003.

[172] Wolkenhauer, O., Sreenath, S. N., Wellstead, P., Ullah, M., and Cho, K.-H., "A systems- and signal-oriented approach to intravellular dynamics," *Biochemical Society Transactions*, vol. 33, pp. 507–515, 2005.

[173] Wolkenhauer, O., Ullah, M., Wellstead, P., and Cho, K.-H., "The dynamic systems approach to control and regulation of intracellular networks," *FEBS Letters*, vol. 579, pp. 1846–1853, 2005.

[174] Woodcock, J., "Sphingosine and ceramide signaling in apoptosis," *IUBMB Life*, vol. 58, no. 8, pp. 462–466, 2006.

[175] Wunderlich, Z. and Mimy, L. A., "Using the topology of metabolic networks to predict viability of mutant strains," *Biophysical Journal*, vol. 91, no. 6, pp. 2304–2311, 2006.

[176] Yi, T.-M., Andreas, B. W., and Iglesias, P. A., "Control analysis of bacterial chemotaxis signaling," *Methods in Enzymology*, vol. 422, pp. 123–140, 2007.

[177] Yi, T.-M., Huang, Y., Simon, M. I., and Doyle, J., "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proceedings of the National Academy of Sciences of the USA*, vol. 97, no. 9, pp. 4649–4653, 2000.

VITA

Chang Feng Quo ("Quo") was born, and spent his earlier years, in Singapore. He completed his early education at the Catholic High School, Raffles Institution, and Hwa Chong Junior; and thereafter enlisted in the Singapore Armed Forces, where he served with pride and excellence as a member of the 56th Singapore Armor Brigade. After completing his tour of duty, Quo enrolled at Georgia Tech for college and stayed for graduate school. Quo has a lovely daughter, Chi-Chi, to whom this thesis is dedicated.