Characterization of Sphingolipid-Based Stress Responses in the Yeast *Saccharomyces cerevisiae* through Reverse Engineering

A Dissertation
Presented to
The Academic Faculty

by

Po-Wei Chen

In Partial Fulfillment
of the Requirements for the Degree
Doctorate Philosophy in Bioengineering in the
George W. Woodruff School of Mechanical Engineering

Georgia Institute of Technology
August 2015

Characterization of Sphingolipid-Based Stress Responses in the Yeast *Saccharomyces cerevisiae* through Reverse Engineering

Approved by:

Dr. Eberhard O. Voit, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Edward A. Botchwey
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Yusuf A. Hannun
Stony Brook Cancer Center
Stony Brook University

Dr. Melissa L. Kemp
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Mark P. Styczynski
Department of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Date Approved: July 16, 2015

To My Family

ACKNOWLEDGEMENTS

Pursuing a Ph.D. degree in a foreign country was probably the most important decision of my life so far. So, firstly I would like to thank my parents, who always conveyed a positive attitude toward this dream and who encouraged me to bravely face the challenging world. Their support and understanding have been the most precious gifts during this academic journey.

This dissertation could not have been completed without the help of many people. Dr. Eberhard Voit, my supportive Ph.D. advisor, never ceases to inspire me with astounding ideas and, more importantly, always guided me back to the right trail when I was lost. Our discussions throughout these years are the fundamental building blocks of this work. Also, I would like to sincerely thank my committee members, who always generously offered instruction and opinions on my research. Specifically, I want to thank Dr. Melissa Kemp and Dr. Mark Styczynski for their insightful input and encouragement to ensure that my models were realistic and useful. I also like to thank Dr. Edward Botchwey for his helpful advice regarding sphingolipids and it applications. Last not least, I sincerely appreciate the opportunity to gain practical experience in Dr. Yusuf Hannun's lipidomics lab, where I benefitted from his enormous experience with sphingolipid pathway systems and also from his data on sphingolipids, which were the basis of my research. I have always been feeling fortunate to have such a strong and supportive Ph.D. Thesis Committee.

My lab members, with a variety of backgrounds in Biochemistry, Electrical Engineering, Mechanical Engineering, Mathematics, and Bioinformatics, have always been willing to provide their professional advice and to share their keen observations regarding my research. I would particularly like to thank Dr. Yun Lee and Dr. Luis Fonseca, who are coauthors of my first paper (Chapter 2) and my following papers (Chapters 3 to 5), respectively. Dr. Lee led me into the field of Systems Biology and kindly offered solid training and coding skills in Matlab. Dr. Fonseca, as a Biochemist, provided many key suggestions that allowed us to close the gap between experimentalists and modelers.

Table of Contents

LIST OF TABLES

Page

LIST OF FIGURES

Page

SUMMARY

Sphingolipids regulate numerous cell functions through the activation of specific signaling cascades. In mammalian cells, these lipid-mediated processes include proliferation, differentiation, apoptosis, and cell trafficking, while lower eukaryotes like yeast use them as crucial signaling molecules in cell cycle control and stress responses, as well as a host of other phenomena. In most cases, proper signaling requires specific sphingolipids to assume altered concentrations or different sphingolipid species to be present in fine-tuned proportions, and detailed knowledge of particular alterations in sphingolipid profiles is therefore a prerequisite for understanding the control of fundamental cellular behaviors. Although the metabolism of sphingolipids has been investigated for several decades, such a detailed mechanistic and systemic understanding of the pathway system of sphingolipid biosynthesis and utilization is still lacking, due to the complicated topology of many the reaction steps in the system and to multiple means of regulation at the genomic, proteomic, and metabolic levels. As a consequence, it is still impossible to predict with reliability how cells will react and adapt to external stresses and which specific roles sphingolipids play in such stress responses.

Mathematical models of sphingolipid metabolism have been developed over the past decade to assist in investigations of the pathway structure and its regulatory mechanisms. Among these, the modeling framework of Biological Systems Theory (BST) has proven very beneficial in capturing the nonlinear behavior of the pathway and its functionality. In this thesis, I am using BST to develop computational approaches that shed light on the regulatory mechanisms with which baker's yeast responds to two types of stresses, namely heat and hydroxyurea (HU). The first of these represents a natural stress, while the second constitutes an artificial stress, which the cells had

never encountered before. This latter stress if of interest, as HU has been identified as a drug to treat various diseases, including sickle cell anemia and certain cancers.

Strong external stresses typically mandate the transition of a cellular system from its normal steady state to a different state, which may be transient or constitute a new steady state. In the first project of this work (Chapter 2), I perform a theoretical state transition analysis, which in some sense encompasses all possible strategies for reaching a required new steady state. This analysis makes use of the uniquely beneficial features of S-systems within the BST framework. Based on this theoretical foundation, I propose a sophisticated piecewise optimization strategy that appropriately captures the sphingolipid dynamics in yeast under 30 minutes of heat stress (Chapter 3). The large-scale simulations within this analysis reveal novel cellular response strategies. In particular, they demonstrate that the cells trigger a switch from an initial phase of sphingolipid biosynthesis to a later phase of sphingolipid retrieval from cell membranes.

Recent research has emphasized the importance of distinct variants of the same ceramide molecules, which differ in their fatty acyl CoA chain lengths and have specific signaling functions. To address the roles of these distinct variants, I propose in Chapter 4 a new model specifically for the biosynthesis and utilization of the different ceramide variants. This much more detailed model uses the previous model (Chapter 3) as boundary conditions and permits a detailed analysis of the dynamics of all ceramide variants in response to heat stress. The simulations with this model reveal interesting patterns of ceramide dynamics that are different for variants with long fatty acyl groups and with very long fatty acyl groups.

Stresses lasting only 30 minutes more or less exclude regulation via gene expression. Addressing this issue in Chapter 5, I analyze long-term, 20-hour exposure of yeast cells to hydroxyurea. The custom-tailored optimization approach I developed for this purpose, followed by a mass flow analysis, permits the novel identification and characterization of very subtle regulatory mechanisms of ceramide biosynthesis that are based on the cells' distinction between ceramides with saturated or unsaturated fatty acyl groups.

Taken together, this dissertation work reveals novel, and often rather subtle control mechanisms with which yeast cells coordinate complex responses to external stresses. Beyond the analysis of sphingolipids, this work demonstrates how innovative techniques of dynamic modeling and optimization can assist in the extraction of detailed information from modern metabolomics data.

CHAPTER 1

INTRODUCTION

1.1 Cellular Stress Response

Cells and organisms are regularly exposed to small fluctuations in their environments and have developed effective mechanisms of tolerance. Stronger perturbations lead to stresses, which are not as easily tolerated and require the cells to mount well-coordinated, multi-scale responses. These stress responses are very intriguing, because they offer superb windows into the complex strategies and mechanisms with which cells manage to live and thrive in a changing world. Understanding these responses in detail, however, is very difficult because of the complex coordination of intracellular processes, which include protein folding/unfolding, metabolic regulation, signaling transduction and gene expression. Furthermore, different environmental changes may trigger target genes through different functional paths. For instance, increased temperature causes changes in protein folding or unfolding, which alters enzyme activities immediately and further triggers signaling intermediates, while toxic agents bind to cell membrane acceptors to transduce the stress signals into the cells. Integrating these factors into an in-depth understanding of cellular stress responses is a true challenge.

Tremendous efforts over many decades have been devoted to understanding stress responses, and these have changed with the arrival of new and improved methods. Currently available experimental measurements of stresses and their consequences span a wide array of strategies, including assessments of correlations between the expression and prevalence of genes, enzymes, and metabolites. Furthermore, many attempts have been made to apply cellular responses beneficially to pharmaceutical product development, the maximization of biofuels production, and a variety of other targets. Nonetheless, a detailed, systemic

and mechanistic understanding of specific stress responses is by and large still missing, due to the complexity of cells and also due to the overall paucity of precise and sufficiently comprehensive data obtained at various different time scales. In this dissertation, I focus on a systemic understanding of stress responses in the baker's yeast *Saccharomyces cerevisiae*. A particular focus is the role of important molecules from the family of sphingolipids. The main strategy of this work is the development of modeling techniques that allow me to extract regulatory information from limited metabolic time series measurements.

1.2 Roles of Sphingolipids in Stress Responses of *S. cerevisiae*

Decades of research on sphingolipids have documented the enormous importance of this class of lipids in mediating a variety of critical cell functions. Sphingolipids exist in eukaryotic cells, where they serve not only as constituents of membranes but also as second messengers in different signaling transduction pathways. These can trigger specific changes in gene expression in organisms like baker's yeast and aid the control of cell proliferation, differentiation, cell trafficking and apoptosis in mammals [1-4]. Different sphingolipids often mediate overlapping but distinct cell functions, and it is frequently the balance between different sphingolipid species that evokes a critical response. In particular, the balance among ceramide, sphingosine, and sphingosine-1-phosphate is critical for regulating stress responses, programmed cell death, cell proliferation, differentiation, and cancer survival [5].

Baker's yeast (*Saccharomyces cerevisiae*) is a commonly accepted model for sphingolipid research because its metabolism is similar to that in mammalian cells and it is much easier to handle in the laboratory. Testing yeast sphingolipids has therefore become an important research focus for identifying regulatory patterns in their prevalence upon stresses. Heat and hydroxyurea treatment were chosen for my research as representatives of natural and artificial stressors, respectively. Heat is a particularly useful stressor for microorganisms as it is easily applied and measured, and because cells and organisms have regularly

experienced changes in temperature throughout evolution and developed very effective defenses. Indeed, it is well known that yeast reacts to modest heat stress with responses at several levels of its biological organization [6-8]. Numerous genes are up- or down-regulated within a few minutes, heat shock proteins are mobilized, transcription factors relocate between the cytosol and nucleus, the protective disaccharide trehalose begins to accumulate to high concentrations, and the metabolic profile of sphingolipids undergoes drastic changes. All these changes commence essentially immediately after a sufficient shift in temperature and may last for an hour or more. Some of these alterations, in turn, are known to serve as signals effecting secondary responses, for instance, by activating transcription factors and stress elements that trigger the expression of genes associated with heat stress.

Hydroxyurea is a chemotherapy drug that slows down or even halts DNA synthesis. The signaling roles of a specific sphingolipid species, ceramide, in mediating hydroxyurea stress responses have recently been identified. Although signaling roles of sphingolipids in mediating heat stress and hydroxyurea stress responses have been studied and understood to some degree, detailed metabolic and cellular response strategies of sphingolipids in these stress conditions have not yet been clearly identified, mainly because of the lack of a systemic understanding of the complex collective cellular responses upon stresses. In response to this situation, the main goal of this dissertation is to gain fundamental and systemic insights into yeast sphingolipid stress responses at the metabolic level. It is hoped that these insight will contribute to a detailed understanding of how critical sphingolipid species are altered upon environmental changes and how they, in turn, serve as signal transducers for further response actions.

1.2.1 Sphingolipid metabolism in yeast

Multiple environmental stresses, including heat or exposure to toxic agents, induce perturbations in sphingolipids concentrations, which secondarily result in the activation of distinct signaling pathways. Therefore, it is important to understand how different sphingolipid species are synthesized, transformed, or degraded, and how these processes are regulated upon stress.

Sphingolipids can be generated via two paths, namely *de novo* biosynthesis and the recycling of inositol phosphoceramide (IPC). *De novo* biosynthesis of sphingolipids is initiated by the condensation of serine and palmitoyl-CoA, a reaction which is catalyzed by serine palmitoyltransferase (SPT) (Figure 1.1). The product, 3-keto-dihydrosphingosine (3KDHS) is quickly reduced by KDHS reductase to dihydrosphingosine (DHS). DHS is the main source of so-called "ceramide backbone" compounds. It can be converted into different dihydroceramides (DHC), due to multiple variants of fatty acyl CoAs, which can serve as substrates for ceramide synthase. The reverse reaction, from DHC to DHS, is catalyzed by dihydroceramidase. DHC and DHS are key branch points in the sphingolipid biosynthesis pathway, because hydroxylases can irreversibly convert these compounds into phytoceramide (PHC) and phytosphingosine (PHS), respectively. PHC and PHS may undergo reversible reactions catalyzed by ceramide synthase (PHS to PHC) and phyotoceramidase (PHC to PHS); the forward reaction (PHS to PHC) requires one from among several different fatty acyl CoAs as substrate.

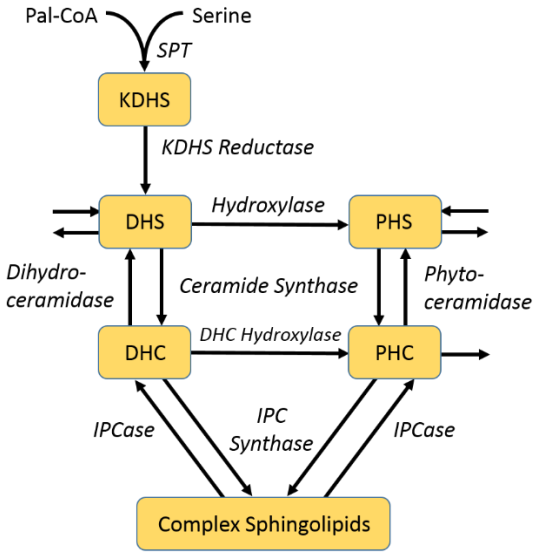


Figure 1.1. Greatly simplified diagram of ceramide metabolism. *Key metabolites are shown in boxes; while enzymes are represented in italics. See Text for abbreviations.*

De novo biosynthesis is not the only option for making sphingolipids available when needed. Equivalent to sphingomyelin in mammalian cells, yeast inositol phosphorylceramides (IPCs), which are also called complex sphingolipids, refer to a class of ceramides with one or two inositol groups attached. They can be formed from DHC and PHC substrates via catalysis by the enzyme IPC synthase. IPC can be irreversibly converted into mannose inositol phosphorylceramide (MIPC), which can furthermore form mannose di-inositol phosphorylceramide (MIP_2C). Importantly, IPC, MIPC and MIP_2C all can serve as sources of DHC and PHC, through catalysis by the enzyme IPCase. Thus, degradation of IPC compounds serves as the second path of ceramide production. Detailed diagrams of sphingolipid metabolism are shown in Chapters III and IV.

Investigating dynamic stress responses of a complicated pathway like sphingolipid metabolism can be very difficult for our unaided mind because key regulatory mechanisms act at multiple scales of the system and the inter-conversions of materials are dynamic and context dependent. Assistance can be found in mathematical and computational modeling, which has gradually been gaining sufficient power and efficacy to shed light on complex biomedical systems.

1.3 Mathematical modeling and reverse engineering

Mathematical modeling uses mathematical concepts and techniques to describe phenomena in the real world. While modeling has been used and applied in physics and engineering for many years, its effective application to larger biomedical systems is relatively new. For instance, simulations of models can reliably predict what might happen to engineered systems and machines and help engineers to design robust strategies that avoid dangerous or undesirable situations. The space shuttle program is a successful example of this type of simulation analysis. By designing physics-based models and putting mechanical and environmental parameters into these models it has been possible to predict precisely how much power the space shuttle needs, how much wind shear it can tolerate, what the perfect reentry angle is, and so on.

These predictions are accurate enough to send an engineered machine successful into the orbit. While simulations of extremely complicated systems work in many engineering cases, one might ask why one cannot make similarly reliable predictions for biomedical systems, cells, or organisms. Why can't we simply borrow the mature techniques and experiences from physics and engineering to gain an understanding of complex cellular machines?

One main challenge of modeling cellular systems, and a significant difference between modeling engineered and biomedical systems, is that the building blocks of cells are not as well known as engineering objects. While engineers build cars or airplanes from basic components in a bottom-up approach, biomedical engineers are forced to work in the reverse, top-down manner. Looking at a complex system, they must dissect often ill-characterized links between genes, enzymes or metabolites from measurements of high-level expressions, activities or concentrations. Converting these measurements into properly scaled numbers and imputing parameter values for mathematical models requires specifically designed algorithms that are hoped to reveal the unknown machineries that govern the observed behaviors of biomedical systems.

In this dissertation, I use a specific modeling framework for this reverse engineering task, namely, Biochemical Systems Theory (BST), which is based on favorably structured ordinary differential equations (ODEs). Based on power-law approximation, BST uses two main formulations. One is called a Generalized Mass Action (GMA) system and the other is called an S-system model. The base techniques for designing models in each formulation can be found in the following examples. Detailed descriptions can be found in Chapter 2 for S-system models and in Chapters III and IV for GMA systems.

1.3.1 GMA model

Building a GMA system of a metabolic pathway is intuitive and straightforward. First, one identifies the dependent variables, independent variables and reactions. For the illustration in Figure 1.2, X_1 to X_4 are metabolites and X_5 to X_9 are enzymes with constant amounts and activities. The metabolites X_3 and X_4 are defined as constant inputs, which means that they are independent variables, along with X_5

to X_9. Therefore, the model has two dependent and seven independent variables. Reactions describe how metabolites are converted, and the magnitudes of these conversions, referred to as fluxes, are the target of the model.

The dynamics of any dependent variable can at first be described as the sum of influxes minus the effluxes. Taking the first dependent variable in Figure 1.2 as an example, one obtains $\frac{dX_1}{dt} = Flux_1 + Flux_2 - Flux_3$. Following this logic for all dependent variables, we can easily build a flux based model.

Each flux can be described in many ways. In BST, the power-law function has been adopted to approximate any flux in the system. Using Figure 1.2 again as an example, metabolite X_4 is converted to X_1, and the process is catalyzed by enzyme X_6 and regulated by metabolite X_2. In a GMA system, this $Flux_2$ is formulated as $\gamma_{1,2}X_2^{f_{1,2,3}}X_4^{f_{1,4,4}}X_6^{f_{1,6,5}}$, where $\gamma_{1,2}$ is the rate constant and the indexed exponents $f_{i,j,k}$ are kinetic orders that capture the strength of each variable on the flux. (Please see Figure 1.2 for details of the numbering scheme.)

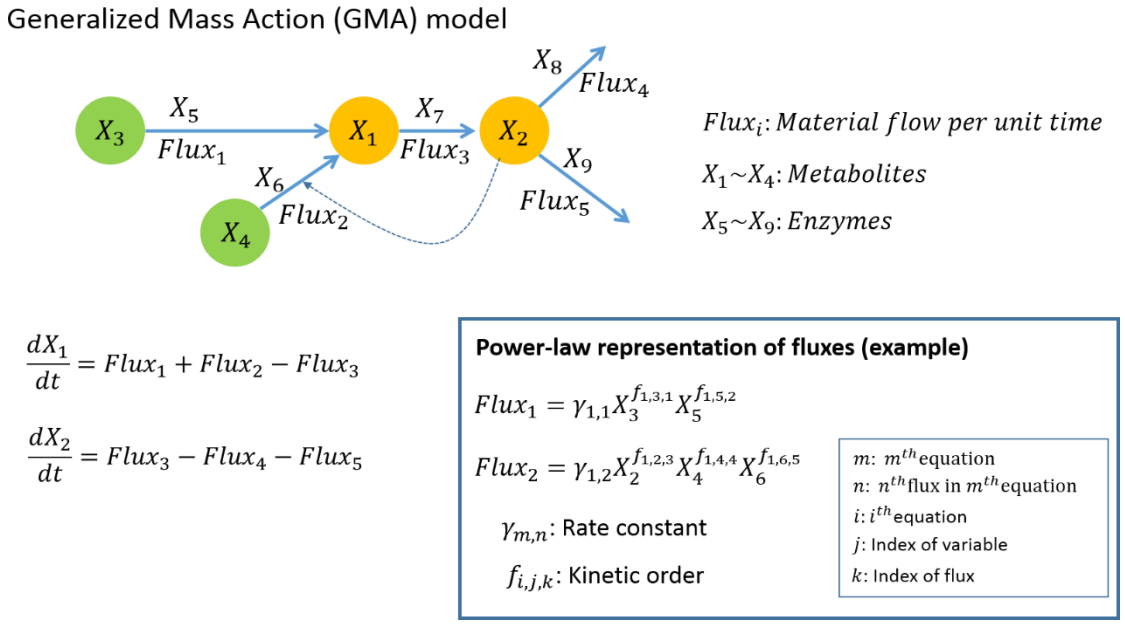


Figure 1.2. Example of a GMA model.

1.3.2 S-system model

Technically speaking, an S-system model can be seen as a power-law approximation of a GMA model. The rationale for this approximation is that the resulting S-system structure is mathematically simpler and permits certain analyses that are difficult for GMA systems. (Please refer to Chapter 2 for details.)

The key concept is that the rate of change of any dependent variable can be described as only one (aggregated) input term and one (aggregated) output term. These input and output term contain power-law representations of the contributions of variables associated with all influxes and all effluxes, respectively. As an example, consider Figure 1.3, which is the same metabolic pathway system as in Figure 1.2. We first consider dependent variable, X_1. To construct the S-system equation, one collects all dependent and independent variables associated with any of the inputs: X_2, X_3, X_4, X_5 and X_6. The input term of X_1 therefore can be formulated as $\alpha_1 X_2^{g_{1,2}} X_3^{g_{1,3}} X_4^{g_{1,4}} X_5^{g_{1,5}} X_6^{g_{1,6}}$, and following the same rule, the output term of X_1 can be formulated as $\beta_1 X_1^{h_{1,1}} X_7^{h_{1,7}}$, where α_i, β_i are rate constants and $g_{i,j}$, $h_{i,j}$ are kinetic orders. Note that the degradation term of X_1 is identical to the production term for X_2, which imposes constraints on the parameter values. Here, it is identical to *Flux*$_3$ in the GMA formulation.

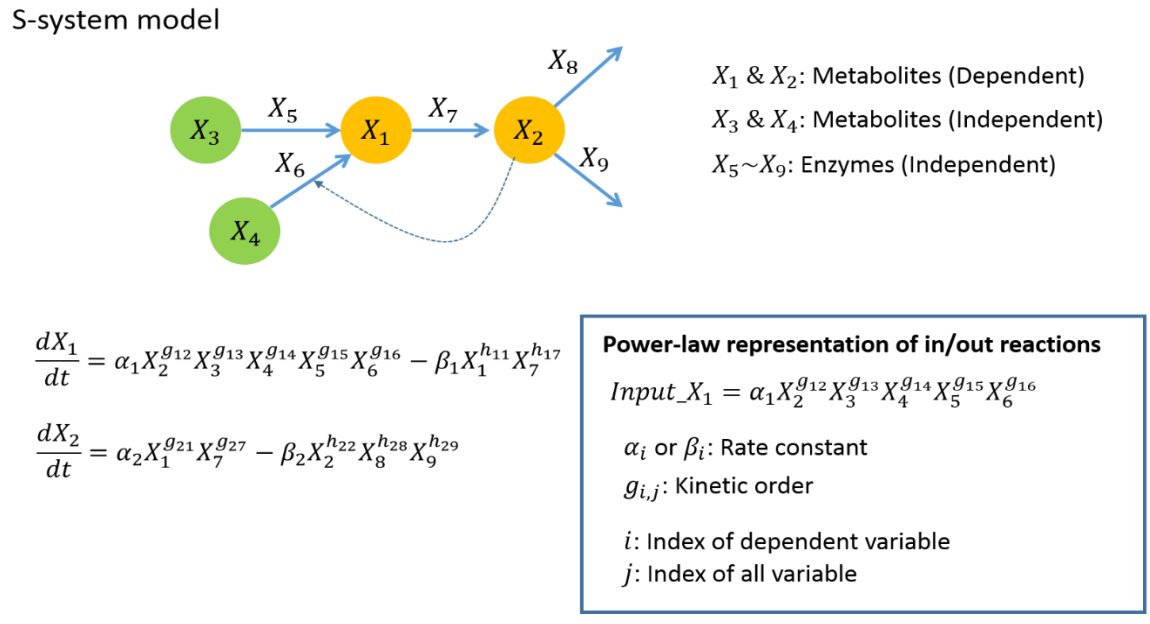


Figure 1.3. Example of an S-system model. Compare this representation with the GMA system in Figure 1.2.

1.4 Generic research strategy

Sphingolipid metabolism is altered upon stresses in order to "shift" the metabolic state of sphingolipid species such that it initiates signaling transduction. Such shifts are mainly achieved by changing the activities of many enzymes simultaneously. In order to understand these state transition mechanisms, I establish in this dissertation work a theoretical and computational framework for the assessment of stress responses, using S-system and GMA models. Chapter 2 contains a mathematically rigorous, theoretical analysis of system shifts to new steady states. Although the transitions in this case are exclusively between steady states, this fundamental research provides a solid platform for the further development of algorithms that will be used for lipidomics data in the form of actually measured wet-lab time series of concentrations.

Sphingolipid stress responses are dynamically and collectively coordinated. An interesting aspect of the collective cellular responses is the fact that they occur at distinct time scales. Some are effective immediately, while others require involvement of the entire sequence of gene expression, transcription, translation and protein modification before the end result takes effect. In Chapters III and IV, I focus on the particular dynamic roles of sphingolipids and ceramides, respectively, in the heat stress response of the baker's yeast *Saccharomyces cerevisiae*. Specifically, I investigate how the cell establishes the observed alterations in sphingolipid and ceramide profiles within a 30-minute time period of heat stress. It is clear that these altered metabolite profiles are the result of changes in the activities of some or all enzymes of sphingolipid metabolism. I will demonstrate that critical changes in activity can be inferred with novel computational approaches that use measured times series of different sphingolipid concentrations, combined with customized optimization strategies and the corresponding dynamic model that our laboratory has been developing and fine-tuning over the past decade [9-11].

Ceramide mediates cellular responses to the chemotherapy drug hydroxyurea (HU). Therefore, maintaining proper ceramide levels is a crucial task for cells when exposed to HU. In Chapter 5, I investigate

how cells manage to regulate levels of multiple ceramide species under 20 hours of HU stress. I use for this analysis a newly developed procedure that combines concepts of the optimization approach proposed in Chapter 3 with techniques of mass flow analysis.

CHAPTER 2

ANALYSIS OF OPERATING PRINCIPLES WITH S-SYSTEM MODELS [1]

2.1 Biological Design and Operating Principles

Biological design principles refer to structural or regulatory features of biological systems that are observed more often than expected. They are thought to have survived evolution, thereby making them apparently superior to hypothesized alternative structures that *a priori* might seem equally reasonable and valid [12,13]. The typical question in the investigation of design principles is: What is the advantage of a particular structural or regulatory feature over an otherwise equivalent design that lacks this feature?

Design principles are identified and investigated through comparisons with reference cases. In static network analysis, a candidate structure is declared a *motif* [14-17] if it is found significantly more often than in random graphs, as they were originally proposed over fifty years ago by Erdös and Rényi [18]. Within Biochemical Systems Theory (BST; [12,19-21]), the approach of choice has been the *Method of Controlled Mathematical Comparisons* (MCMC), which compares the observed dynamical system with a reference system that differs only in the one feature of interest [13,22]. MCMC originally focused on algebraic analyses, but was subsequently augmented with computational and statistical methods [12,23-26]. Dynamic biological systems that were successfully analyzed with respect to design principles include pathway topologies [12,23,24,27], immune cascades [23], gene regulatory circuits [28-30], signaling

[1] **Most of the material presented in this chapter was published in the article "Analysis of operating principles with S-system models", Mathematical Biosciences, 231(1), pp49-60, 2011.**

systems [31], and riboswitches [32]. A recent collection of papers on design principles can be found in a special issue on biological design principles in *Mathematical Biosciences* 231(1), 2011.

While design principles have become a fashionable topic of investigation in recent years, their dynamic counterparts, operating principles, have received only a small fraction of the attention. Operating principles address questions of the dynamics of a response as we observe or hypothesize it, in comparison to other, *a priori* equally valid alternatives [33]. The typical question asked in this case is: What is the advantage of utilizing or altering a particular process instead of an alternative process? As an example for this type of a question, the task was posed to optimize the product yield of a pathway by selecting and altering a small number of genes or enzymes. The results, which were not easy to predict without a quantitative analysis, demonstrated that the locations and magnitudes of optimal manipulations depend significantly on the regulatory signals in the pathway [33].

In a different context, the following question arose: If a system is forced to move to a new steady state, and if this state may be achieved either by drastically changing a few independent variables or by slightly changing many independent variables, which strategy is preferable? Alvarez and colleagues [34] analyzed this question heuristically for changes in yeast metabolism during the diauxic shift and came to the conclusion that many genes in the living yeast cell were changed to a modest degree. Yet a different aspect of operating principles was investigated in the response of yeast cells exposed to heat stress [25,35-38]. In this case, the lead questions were: Which genes are actually up-regulated in expression and by how much? What are the metabolic consequences of this up-regulation? Are there alternative up-regulation scenarios that might perform better? Can we find objective criteria explaining the emergence of the observed strategy?

One might ask whether operating principles are truly different from design principles, because the possible space of dynamic responses is clearly constrained, if not determined, by the physical and regulatory structure of a system. While design and operation are clearly coupled to some degree, their distinction seems both reasonable and necessary, because a cell or organism could theoretically respond to the same demand in different ways, even within exactly the same structural confines, as the diauxic shift study [34]

demonstrates. Furthermore, cells can be exposed to drastically different demands, which require appropriate responses within the same structural design. A good example is the generation of energy in blue-green algae, which may occur autotrophically per photosynthesis or heterotrophically per consumption of carbohydrates. It has been shown that the numerical distribution of metabolic fluxes, and thus the operation of the system, shifts dramatically between these two modes [39]. In a different example, it was shown that plant cells use the same pathway system, but with distinctly different, dynamically changing flux distributions, to produce woody materials during their development or in different transgenic strains [40,41].

As in the case of design principles, it is impossible to study operating principles in exhaustive generality. The analysis described here therefore focuses exclusively on one pertinent special case, namely, where a biological system must shift from its normal steady state to a new steady state, a response that is typical in the face of persistent changes in a cell's or organism's environment. While the two steady states will be at the center of the analysis, features of transients will also be discussed to some degree. In first approximation it may even be possible to consider slow-changing, longer-term trends as a series of different "almost-steady-states" [42].

Most analyses of design principles in the past had the benefit of clear reference systems that were topologically very similar to the system of interest. For instance, a system with feedback was compared to a system without this particular feedback. In the case of operating principles, it is not always *a priori* clear what the alternatives are. For instance, we cannot simply compare up-regulation of one process against unaltered operation, because the two would lead to different transients and steady states. Instead, the approach of a new steady state will almost always require alterations in larger sets of independent variables. Thus, the first important step in the analysis of operating principles is an exhaustive exploration of the *admissible set of operating strategies*. Once this set is characterized, the true discovery of *operating principles* consists of the selection of the one strategy that is superior to all others under the chosen criteria of optimality.

2.2 Methods and Theoretical Results

Canonical models, and in particular S-systems within Biochemical Systems Theory [12,19], are especially well suited for analyzing operating principles. As in the case of design principles, the primary reasons are twofold. First, these systems have a fixed structure, in which each component has a well-defined meaning and where system features are mapped onto parameters in a one-to-one fashion [12,21]. Secondly, S-systems permit a linear representation of their steady states within the language of linear algebra, upon a logarithmic transformation of all variables [43].

As described many times, S-systems always have the format

$$\dot{X}_i = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}} \ . \qquad (2.1)$$

Here, the X_i, $i = 1, \ldots, n$ are dependent variables, which may change under the action of the system, while X_i, $i = n+1, \ldots n+m$ are independent variables, which may affect the action of the system but are themselves not affected by the system. The parameters α_i and β_i are non-negative rate constants, and g_{ij} and h_{ij} are real-valued kinetic orders. The literature on these systems is quite rich (*e.g.*, see references in [21]).

The generic situation to be addressed here concerns a biological system, represented by S-system equations, that needs to respond to a changed environmental demand by assuming a new steady state. It is not difficult to imagine that this task usually has many solutions and that distinctly different settings of independent variables may lead to the same steady state with respect to the dependent variables. This multiplicity of possibilities is due to the fact that most systems contain many more processes than variables. Because these processes are usually under the control of independent variables, different choices of independent variables correspond to distinct solution strategies.

The non-trivial steady state of an S-system model can be formulated in matrix notation as

$$\mathbf{A}_D \cdot \mathbf{y}_D + \mathbf{A}_I \cdot \mathbf{y}_I = \mathbf{b} \qquad (2.2)$$

[44], where $\mathbf{y}_D$ denotes the vector of the logarithms of the dependent variables at steady state, $\mathbf{y}_I$ is the corresponding vector of independent variables, the elements of the matrices $\mathbf{A}_D$ and $\mathbf{A}_I$ are $a_{ij} = g_{ij} - h_{ij}$ for all i and j, separated into dependent and independent variables, and $b_i = \log(\beta_i/\alpha_i)$ for $i = 1, \ldots, n$.

In a typical analysis, all parameter values are known and one computes the non-trivial steady state, which may then be used for other diagnostics like stability, sensitivity, and gain analysis [12,21,43]. This steady state can be expressed explicitly as

$$\mathbf{y}_D = \mathbf{S} \cdot \mathbf{b} + \mathbf{L} \cdot \mathbf{y}_I \tag{2.3}$$

[44], where $\mathbf{S} = \mathbf{A}_D^{-1}$ and $\mathbf{L} = -\mathbf{A}_D^{-1}\mathbf{A}_I$ are the so-called sensitivity and logarithmic gain matrices, respectively.

For our purposes here, we must turn the task around. We assume that the system has to switch from some initial state to a target steady state that is mandated by new environmental demands. We furthermore suppose that we know the numerical values of the dependent variables at this target steady state. The question thus becomes how the independent variables should be changed to achieve this state (*cf.* [42,45]). Again using stress as an example, we might observe an altered metabolic steady state and ask which enzymes would have to be altered in activity to reach the stress state.

For ease of representation, we rewrite Eq. (2.3) as

$$-\mathbf{A}_D^{-1}\mathbf{A}_I\mathbf{y}_I = \mathbf{y}_D - \mathbf{A}_D^{-1}\mathbf{b} \tag{2.4}$$

and define

$$\mathbf{y}'_D = \mathbf{y}_D - \mathbf{A}_D^{-1}\mathbf{b}\,, \tag{2.5}$$

which yields the simplified representation

$$\mathbf{L}\mathbf{y}_I = \mathbf{y}'_D\,. \tag{2.6}$$

For the special case where $m = n$ and $\mathbf{L}$ has full rank, we can invert the system of equations and express the vector of independent variables as a unique linear function of the dependent variables in logarithmic space; namely we obtain

$$\mathbf{y}_I = \mathbf{L}^{-1}\mathbf{y}'_D . \qquad (2.7)$$

Expressed in words, we can demand numerical values for the dependent variables of a particular target steady state, and Eq. (2.7) determines how the independent variables have to be set to reach this state. If the new state is stable, and if the system starts within its basin of attraction, one may actually reach this state by starting the system at the original steady state and resetting the independent variables according to Eq. (2.7). Of course, we do not know how much time the dynamic system will require to come sufficiently close to the target.

For cases where $m < n$, the matrix $\mathbf{L}$ is "tall," which reflects an over-determined system that generally permits no solution. Nevertheless, for practical purposes we can compute a least squares solution, which minimizes the deviation from the target state and is given as the regression equation

$$\mathbf{y}_{I\_LS} = \mathbf{L}^{+}\mathbf{y}'_D . \qquad (2.8)$$

In the most pertinent case, the number of independent variables is larger than the number of dependent variables ($m > n$). This relationship is not always true, but it usually holds, because most systems contain more processes than pools. The matrix equation (2.6) now can no longer be inverted directly, and if the rank of $\mathbf{L}$ is r, the solution consists of an $m - r$ dimensional space. Even though an inversion is not directly possible, the solution space may be characterized with methods of linear algebra, where the starting point is the pseudo-inverse [46]. Specifically, the solution space, which consists of every admissible $\mathbf{y}_I$, can be spanned through the following steps. First, find a particular solution $\mathbf{y}_{I\_PS}$. Then use $\mathbf{y}_{I\_PS}$ and the span of the null space of $\mathbf{L}$ to describe the entire solution space as

$$\mathbf{y}_{I\_PS} = \mathbf{L}^{+}\mathbf{y}'_D \qquad (2.9)$$

$$\mathbf{y}_I = \mathbf{y}_{I\_PS} + \mathbf{B}\cdot\boldsymbol{\lambda} \qquad (2.10)$$

where $\boldsymbol{\lambda}$ is any given real-valued $(m–n)$-dimensional vector, rank($\mathbf{L}$) = n, and $\mathbf{B}$ is a matrix in which each column is a basis vector. Together, the set of these vectors constitutes a basis of the null space of $\mathbf{L}$.

2.2.1 Illustration Examples

The first example is a cascaded system (Figure 2.1), where the numbers of precursors and state variables are the same ($n = m = 4$) and the system has a unique solution. The cascade could describe the expression of a formerly inactive gene X_5, which becomes activated (X_1) and is subsequently transcribed; X_6 could model nucleotides that are assembled into mRNA (X_2); X_7 could represent amino acids, which are assembled into an enzyme (X_3), which subsequently catalyzed the conversion of a metabolic substrate X_8 into a product X_4. The final product could indirectly lead to the reduction of gene expression. The generic S-system representation of the model is

$$\begin{aligned}
\dot{X}_1 &= \alpha_1 X_5^{g_{15}} - \beta_1 X_1^{h_{11}} X_4^{h_{14}} \\
\dot{X}_2 &= \alpha_2 X_1^{g_{21}} X_6^{g_{26}} - \beta_2 X_2^{h_{22}} \\
\dot{X}_3 &= \alpha_3 X_3^{g_{32}} X_7^{g_{37}} - \beta_3 X_3^{h_{33}} \\
\dot{X}_4 &= \alpha_4 X_3^{g_{43}} X_8^{g_{48}} - \beta_4 X_4^{h_{44}}
\end{aligned} \tag{2.11}$$

Without loss of generality in this and the later illustration examples, all rate constants α_i and β_i were arbitrarily set to 1 and the independent variables were initially defined as 1.2. By this definition we know that $\mathbf{y}'_D = \mathbf{y}_D$ because $\mathbf{b} = \mathbf{0}$. The values of the kinetic order parameters in this and other systems are given in Table 2.1.

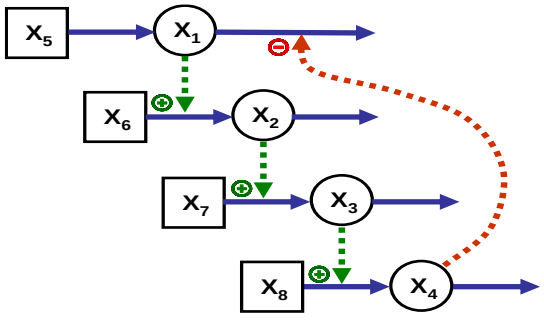


Figure 2.1: A cascaded system with as many dependent (circles) as independent (squares) variables.

Table 2.1: Numerical Values of Kinetic Order Parameters of the Illustration Examples

Cascade 1		**Linear Pathway**		**Cascade 2**		**Branched Pathway**			
Figure 2.1		**Figure 2.2**		**Figure 2.4**		**Figure 2.7**			
g_{15}	0.24	g_{15}	0.24	g_{15}	0.24	g_{15}	0.24	h_{11}	0.2
g_{21}	0.4	g_{21}	0.4	g_{21}	0.4	g_{21}	0.4	h_{13}	0.1
g_{26}	0.3	g_{24}	0.3	g_{26}	0.3	g_{23}	0.1	h_{16}	0.16
g_{32}	0.5	g_{32}	0.5	g_{32}	0.5	g_{26}	0.3	$h_{1,11}$	0.1
g_{37}	0.3	g_{36}	0.3	g_{37}	0.3	$g_{2,11}$	0.3	h_{22}	1
g_{43}	0.1	g_{43}	0.1	g_{43}	1	g_{32}	0.5	h_{27}	0.33
g_{48}	0.2	g_{47}	0.2	h_{11}	0.2	g_{39}	0.3	h_{29}	0.1
h_{11}	0.2	h_{11}	0.2	h_{14}	0.16	g_{42}	0.1	h_{33}	0.4
h_{14}	0.16	h_{14}	0.16	h_{22}	1	g_{47}	0.2	$h_{3,10}$	0.05
h_{22}	1	h_{22}	1	h_{33}	0.4			$h_{3,11}$	0.1
h_{33}	0.4	h_{26}	0.33	h_{44}	0.9			h_{44}	0.2
h_{44}	0.2	h_{33}	0.4					h_{48}	0.25
		h_{37}	0.1						
		h_{44}	0.2						
		h_{48}	0.25						

The second example of a simple linear pathway with feedback and an exogenous demand for product (Figure 2.2) is again characterized by $n = m = 4$. Because of precursor-product relationships, effluxes and influxes associated with subsequent pools are constrained, but these constraints have no real bearing on the

characterization of a set of independent variables that moves the system to the target steady state. The generic S-system model is

$$\begin{aligned}
\dot{X}_1 &= \alpha_1 X_5^{g_{15}} - \beta_1 X_1^{h_{11}} X_4^{h_{14}} \\
\dot{X}_2 &= \beta_1 X_1^{h_{11}} X_4^{h_{14}} - \beta_2 X_2^{h_{22}} X_6^{h_{26}} \\
\dot{X}_3 &= \beta_2 X_2^{h_{22}} X_6^{h_{26}} - \beta_3 X_3^{h_{33}} X_7^{h_{37}} \\
\dot{X}_4 &= \beta_3 X_3^{h_{33}} X_7^{h_{37}} - \beta_4 X_4^{h_{44}} X_8^{h_{48}}
\end{aligned} \qquad (2.12)$$

Again, all rate constants α_i and β_i were arbitrarily set to 1 and the independent variables to 1.2. The values of the kinetic orders are given in Table 2.1.

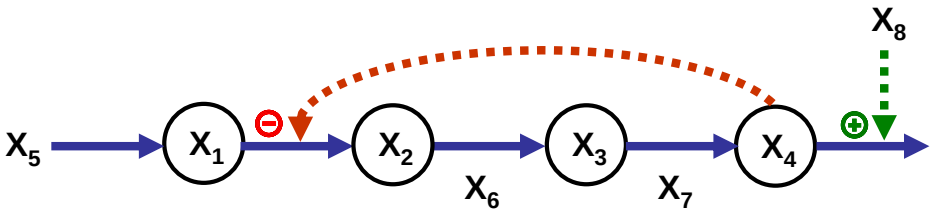


Figure 2.2: Linear pathway with feedback and an exogenous demand for product. The task of moving the system to a new steady state has a unique solution.

For our illustration, we start both systems arbitrarily at (1, 1, 1, 1) and let them reach their nominal steady states. As a response to an altered environmental demand we assume that all variables in the cascade and the linear pathway must move to a target value of 2. Thus, after the initial steady state has been established, we evoke the unique solutions for the cascaded and the linear systems and reset the independent variables as $\mathbf{X}_I$ = [2.8284 4.0 0.7937 1.4142]^T and $\mathbf{X}_I$ = [2.8284 0.2607 0.7579 1.5583]^T, respectively. The systems indeed respond by moving to the desired target states (Figure 2.3). The inverse solution does not convey anything about the transients.

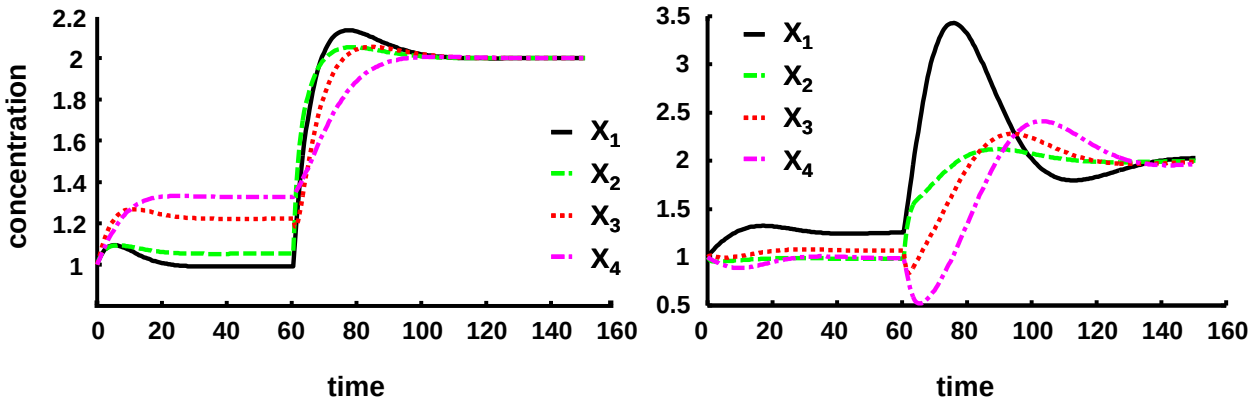


Figure 2.3: Resetting the independent variables to the computed vector moves the cascaded (left) and linear (right) pathway systems from the initial state to the1 desired target.

The third and fourth examples are cascaded and linear pathways with fewer independent than dependent variables (Figure 2.4). S-systems models were constructed according to well-documented guidelines, and the values of the kinetic orders for the cascaded system are given in Table 2.1. It could seem that these scenarios are rather unrealistic, but they do occur in cases like the ones shown here as well as in cases of strongly connected pathways where not all genes or enzymes are accessible to manipulations. If it is infeasible or impossible to alter some of the independent variables, *m* is in effect decreased and may become lower than *n*.

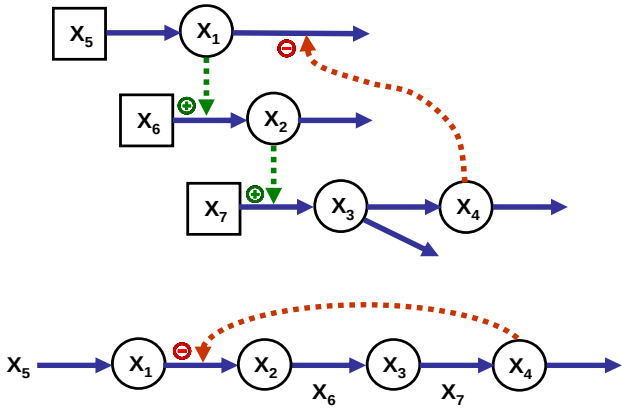


Figure 2.4: Over-determined cascaded and linear pathway systems with n = 4, m = 3. In the example of a linear pathway, the reaction between X_1 and X_2 may not be accessible to alterations.

This "unsolvable" situation may be addressed in different ways. First, instead of searching for an exact solution, one may solve the corresponding regression problem and find a set of independent variables that moves the system to a steady state that is close to the target state (Figure 2.5). In this case, the solution vector is $\mathbf{X}_I = [5.3889 \;\; 9 \;\; 0.6395]^T$, and we see that both X_3 and X_4 are not quite on target.

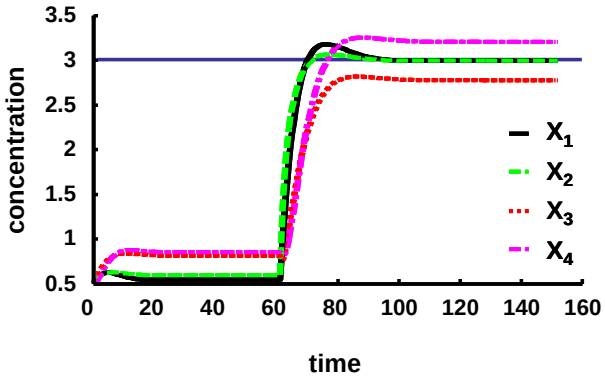


Figure 2.5: Least squares solution for the over-determined cascaded system in Figure 2.4.

As a variation on this theme, closeness to the target state may be defined differently for each dependent variable, through the use of appropriate weights. This strategy allows that some "important" dependent variables can be selected to come as closely as possible to their target values, while others are possibly not. Finally, one may ignore some of the dependent variables, whose specific values are not considered as important as those of other variables, and restrict the optimization to a subset of important dependent variables, thereby in effect reducing n. Examples for less important variables might be intermediates in linear pathways,

To be specific, suppose it is most important that variable X_4 of the cascaded pathway attain the target value, while other variables are of secondary importance. The original task can be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} \mathrm{L}_{11} & \mathrm{L}_{12} & \mathrm{L}_{13} \\ \mathrm{L}_{21} & \mathrm{L}_{22} & \mathrm{L}_{23} \\ \mathrm{L}_{31} & \mathrm{L}_{32} & \mathrm{L}_{33} \\ \mathrm{L}_{41} & \mathrm{L}_{42} & \mathrm{L}_{43} \end{bmatrix} \begin{bmatrix} y_5 \\ y_6 \\ y_7 \end{bmatrix} \quad \text{where} \quad y_{\mathrm{D}} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}, \ \mathbf{L} = \begin{bmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \\ L_{41} & L_{42} & L_{43} \end{bmatrix}, \text{ and } \quad y_{\mathrm{I}} = \begin{bmatrix} y_5 \\ y_6 \\ y_7 \end{bmatrix}. \quad (2.13)$$

To enforce that X_4 moves to the target, presumably at the cost of other variables, we separate the equation for X_4 in Eq. (2.13) from the rest, which yields

$$y_4 = \begin{bmatrix} \mathrm{L}_{41} & \mathrm{L}_{42} & \mathrm{L}_{43} \end{bmatrix} \begin{bmatrix} y_5 \\ y_6 \\ y_7 \end{bmatrix}. \quad (2.14)$$

Using the notation $\mathbf{L}_{123} = \begin{bmatrix} \mathrm{L}_{11} & \mathrm{L}_{12} & \mathrm{L}_{13} \\ \mathrm{L}_{21} & \mathrm{L}_{22} & \mathrm{L}_{23} \\ \mathrm{L}_{31} & \mathrm{L}_{32} & \mathrm{L}_{33} \end{bmatrix}$ and $\mathbf{L}_4 = [\mathrm{L}_{41} \quad \mathrm{L}_{42} \quad \mathrm{L}_{43}]$, the particular solution of $\mathbf{y}_I$ based on this separated equation is now given as

$$\mathbf{y}_I = \mathbf{y}_{I\_PS} + \mathbf{B}_4 \cdot \boldsymbol{\lambda}, \quad (2.15)$$

where

$$\mathbf{y}_{I\_PS} = \mathbf{L}_4^{+} y_4, \quad (2.16)$$

$\mathbf{B}_4$ is a 3×2 matrix where each column is a basis vector of the null space of $\mathbf{L}_4$, and $\boldsymbol{\lambda}$ is any real-valued 2-dimensional vector. Having enforced that the fourth variable will reach the target value, we still have options for the remaining independent variables. Namely, the equation

$$\begin{aligned} \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^T &= \mathbf{L}_{123} \mathbf{y}_I \\ &= \mathbf{L}_{123} (\mathbf{y}_{I\_PS} + \mathbf{B}_4 \cdot \boldsymbol{\lambda}) \end{aligned} \quad (2.17)$$

allows us to define criteria such as a least-squares error for the remaining variables, which correspond to different choices for $\boldsymbol{\lambda}$. For instance, we can use the pseudo-inverse to define

$$\boldsymbol{\lambda} = (\mathbf{L}_{123} \cdot \mathbf{B}_4)^{+} \left(\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^T - \mathbf{L}_{123} \cdot \mathbf{y}_{I\_PS} \right), \quad (2.18)$$

which yields the solution as

$$\mathbf{y}_I = \mathbf{y}_{I\_PS} + \mathbf{B}_4 \cdot (\mathbf{L}_{123} \cdot \mathbf{B}_4)^{+} \left(\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^T - \mathbf{L}_{123} \cdot \mathbf{y}_{I\_PS} \right). \tag{2.19}$$

The result of this operation is shown in Figure 2.6. In comparison with Figure 2.5, X_4 now reaches the target value 3 exactly, while the remaining variables approach the value 3 approximately. In particular, the improvement in X_4 is "paid for" with decreased performance of X_3. The solution vector of independent variables in this case is $\mathbf{X}_I = [5.1962 \;\; 9 \;\; 0.5989]^T$. If X_3 is most important in the same system, the solution vector is $\mathbf{X}_I = [5.6367 \;\; 9 \;\; 0.6934]^T$ and X_4 overshoots the target (plot not shown).

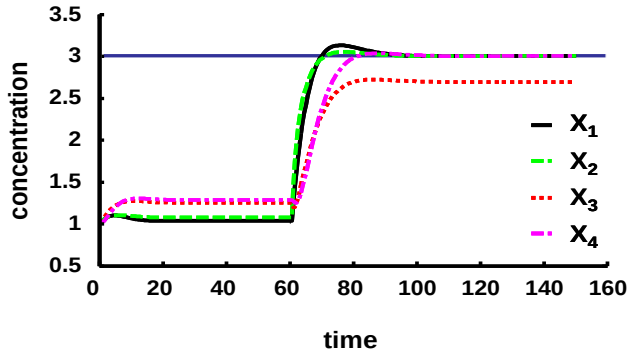


Figure 2.6: Solution for the over-determined cascaded system in Figure 2.4, where X_4 is forced to reach the target state 3.

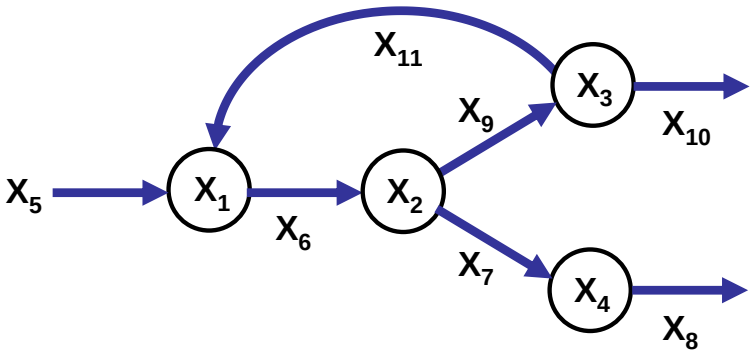


Figure 2.7: Branched pathway with a substrate cycle. The system contains four dependent variables (circles) and seven independent variables, which model the system input (X_5) and catalyzing enzymes (X_6, …, X_{11}). The system is representative of the most prevalent situation where n < m.

The most pertinent case is $n < m$. A representative example is the pathway shown in Figure 2.7, which has four dependent and seven independent variables. The S-system was constructed according to usual guidelines (see Table 2.1 for kinetic orders), and as before, the initial state was set to (1, 1, 1, 1) and all independent variables were initially set to 1.2. Furthermore, we assume that all target values for the dependent variables under the changed environmental demand are 2. The solution now consists of a space that can be expressed by a particular solution plus a linear span of a basis of the null space of $\mathbf{L} = -\mathbf{A}_D^{-1}\mathbf{A}_I$. The particular solution is computed as

$$\mathbf{y}_D = \mathbf{L}\mathbf{y}_I \tag{2.20}$$

$$\mathbf{y}_{I\_PS} = \mathbf{L}^{+}\mathbf{y}_D \tag{2.21}$$

and any feasible solution can be characterized by the particular solution plus an arbitrary vector in the null space of **L**:

$$\mathbf{y}_I = \mathbf{B}\cdot\boldsymbol{\lambda} + \mathbf{y}_{I\_PS} \tag{2.22}$$

where $\boldsymbol{\lambda}$ may be any 3-dimensional real-valued vector and **B** is a matrix in which each column is a basis vector of the null space of **L**.

Choosing any $\mathbf{y}_I$ inside this solution space is guaranteed to lead to the target steady state. Two admissible solutions in Cartesian space are $\mathbf{X}_I$ = [4.1059 0.9618 1.0410 0.7826 1.4962 2.8819 3.9459]^T and $\mathbf{X}_I$ = [3.1563 1.1943 0.6079 0.5089 0.9002 0.9657 1.4845]^T. These and other solutions within the admissible space move the system to the target steady state of (2, 2, 2, 2) as expected, but the transient behaviors of these systems are different, and it is not *a priori* clear how to manipulate them. Interestingly, it is possible to alter any solution to some degree in a targeted fashion by controlling the basis vectors of the three-dimensional null space of **L**. In the given numerical case, the basis vectors are

$$\begin{aligned}
\mathbf{B}_1 &= [0.3907\ \ 0.4237\ \ 0.5878\ \ 0.4702\ \ 0.1110\ \ 0.1462\ \ 0.2599]^T \\
\mathbf{B}_2 &= [-0.0443\ \ -0.0438\ \ -0.1182\ \ -0.0946\ \ 0.1504\ \ 0.9743\ \ -0.0361]^T \\
\mathbf{B}_3 &= [-0.0835\ \ -0.5963\ \ 0.0684\ \ 0.0547\ \ 0.2467\ \ -0.0272\ \ 0.7538]^T
\end{aligned} \tag{2.23}$$

These basis vectors can be computed directly in Matlab with the *Null* command, which applies singular value decomposition to obtain an orthogonal basis set.

Altering any of these basis vectors has different effects. For instance, increasing $\mathbf{B}_1$ by a positive factor causes all responses to speed up (Figure 2.8), while increasing $\mathbf{B}_2$ has the opposite effect (not shown). Increasing $\mathbf{B}_3$ causes X_3 and X_4 to accelerate while slowing down X_1 and X_2 (not shown). Thus, the transient behavior can be controlled to some degree through the basis vectors. However, the effects of such manipulations are difficult to predict, and it is more straightforward to use direct optimization methods as we will discuss them later.

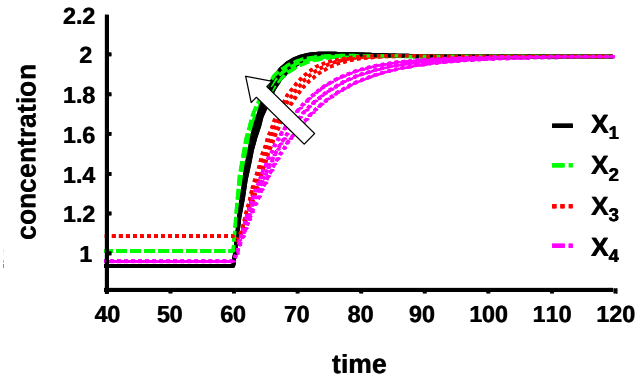


Figure 2.8: Manipulation of the basis vectors permits modest changes in transient speed. Here, increasing B_1 causes all transients to accelerate.

2.2.2 Functional Effectiveness

A key component of the Method of Controlled Mathematical Comparisons is the prior establishment of objective criteria of functional effectiveness [22,23]. In other words, before the comparison of two system structures is performed and interpreted, one formulates metrics according to which either the system of interest or some alternative is deemed superior. Typical criteria are stability, robustness, a short response time to stimuli, adequate responsiveness to external demands, and maybe a transient response profile that does not deviate too far from the nominal profile.

Operating principles have not yet been analyzed often enough to permit a listing of "typical" criteria of functional effectiveness, and judging by the case of design principles, they should be expected to change from one application to another. Among likely, generic criteria one might find similar measures of superiority as for design principles, because the different operating strategies discussed here lead to steady states that are characterized by the same sets of numerical values for the dependent variables, but different values for some of the independent variables. Thus, criteria might often include local stability, modest gains and sensitivities, and tolerance of the steady state to perturbations. Also as in the case of design principles, one might prefer fast response times and bounded transients. Another typical criterion in superior designs is a minimal accumulation of intermediates. Here, this criterion is automatically satisfied when the target profile of steady-state values is set, but if no target values for intermediates are specified, it may indeed serve as a criterion.

In addition to these rather evident criteria, different operating strategies are distinguishable in other respects. For instance, one might consider the metabolic burden, which is associated with the total mass of all protein (*cf.* discussion in [20]). Clearly, different operating strategies are likely to involve different amounts of enzymes, thereby affecting the total protein content of the cell. The metabolic burden can be an important issue of cellular protein economy because it was shown for the case of recombinant bacteria that the growth rate decreased monotonically with increasing numbers of introduced plasmid copies (*e.g.*, [47-49]). An increased metabolic burden can also be a disadvantage because it puts additional stress on the cell due to higher levels of transcription and translation. If minimal metabolic burden is indeed a pertinent criterion of functional effectiveness, the totality of changes in independent variables should be kept as small as possible.

A different criterion is the total number of changed independent variables. It is to be expected that most adaptations to a changed environmental demand will be implemented through the altered expression of key genes. If so, the more genes are involved in a response, the more transcription factors need to be mobilized, which must be expected to incur a physiological cost.

In comparing viable operating alternatives, one should expect that it is easier to up- or down-regulate some genes or enzyme activities than others. In fact, it might not be practically feasible to change some enzyme activities at all. If so, the corresponding independent variables in the model are off limits in the selection of any viable operating strategies. Other processes might be accessible to manipulations but limited in the degree of alteration.

In addition to these more or less generic criteria, some systems might require very specific constraints. As an example, we recently analyzed the biosynthesis of lignin in plant cell walls, where it is known that the monolignols, of which the lignin polymer is ultimately composed, have to be produced in certain ratios for an optimal cell wall composition [40]. These ratios could constitute a set of criteria of functional effectiveness that is germane to this very special case.

2.2.3 Optimal Operating Strategies

The computation of the pseudo-inverse in the S-system steady-state equations, along with the characterization of the null space, results in the space of all possible solutions, in which resetting the independent variables leads to a desired steady state in terms of the dependent variables. While it is mathematically and practically satisfying to have a concise representation of this solution space, one will wonder whether some admissible solutions are "better" than others. In other words, the question becomes how specific solutions within this space should be selected according to the chosen criteria of functional effectiveness. Clearly, the answer lies in some optimization procedure, which, interestingly, does not require an explicit characterization of the solution space *per se*.

The strategy is the following. We introduce a vector $\mathbf{d}$ that represents the change in the vector of independent variables such that the system reaches the target steady state $\mathbf{y}_D^T$, which is assumed to be known. With these definitions, we can write the task as

$$\mathbf{y}_D^T = \mathbf{A}_D^{-1}\mathbf{b} - \mathbf{A}_D^{-1}\mathbf{A}_I(\mathbf{y}_I + \mathbf{d}) . \qquad (2.24)$$

This expression can be reformulated as a linear constraint on $\mathbf{d}$. Namely, we can write

$$\mathbf{A}_D^{-1}\mathbf{A}_I\mathbf{d} = \mathbf{A}_D^{-1}(\mathbf{b} - \mathbf{A}_I\mathbf{y}_I) - \mathbf{y}_D^T . \tag{2.25}$$

Let

$$z_i = \begin{cases} 1 & \text{if the catalytic step coded by } d_i \text{ is induced to reach the steady state} \\ 0 & \text{otherwise} \end{cases} \tag{2.26}$$

If all z_i are set to 1, the optimization task allows every independent variable to change as long as the linear constraints are satisfied, but the identification of specific solutions still depends on the dimension and rank of $\mathbf{A}_D^{-1}\mathbf{A}_I$ as well as the chosen criteria of functional effectiveness. One of the most commonly used criteria for finding a particular solution is a minimized total squared error *E*, which in this case can be written as

$$E = \left\| \mathbf{A}_D^{-1}(\mathbf{b} - \mathbf{A}_I\mathbf{y}_I) - \mathbf{y}_D^T - \mathbf{A}_D^{-1}\mathbf{A}_I\mathbf{d} \right\|^2 , \tag{2.27}$$

where $\|\cdot\|^2$ is the 2-norm. The minimization of *E* results in a target steady state where the collective deviation in independent variables from their nominal values is minimized.

By contrast, if we are looking for a so far unknown subset of indicators $z_i = 0$ that satisfy some criteria of functional effectiveness, the task becomes a Mixed Integer Linear Programming (MILP) problem that allows us to find a minimum set of independent variables whose alteration is necessary for reaching the target steady state. This MILP has the form

$$\begin{aligned} &\min \sum_{i=1}^{m} z_i \text{ subject to} \\ &\mathbf{A}_D^{-1}\mathbf{A}_I\mathbf{d} = \mathbf{A}_D^{-1}(\mathbf{b} - \mathbf{A}_I\mathbf{y}_I) - \mathbf{y}_D^T \\ &d_i \geq 0 \\ &d_i \leq z_i D \text{ (D is an arbitrarily large positive number)} \\ &\mathrm{z}_i : 0/1 \\ &\forall i = 1,...,m \end{aligned} \tag{2.28}$$

The CPLEX solver in AMPL can be used to solve this type of MILP.

Similar to optimization tasks in the field of biotechnology, where the typical objective is the maximization of a metabolite pool or flux, it is here also possible to account for constraints on

concentrations and fluxes [50-53], as well as more complex limitations such as metabolic burden [20] or the feasibility of parameter regions that correspond to admissible physiological states [37].

2.3 Case Study

As an illustrative case study, we consider the response of yeast cells to heat stress. The first indications of such a response are observable within minutes of the initiation of heat stress: transcription factors are mobilized and translocated [54], and numerous genes respond with strong changes in expression [55-57]. At the proteomic level, heat shock proteins emerge in high numbers [58-60]. At the metabolic level, a significantly altered profile of sphingolipids guides the expression of some key genes [61], and, most important for the following illustration, the protective disaccharide trehalose is produced in huge amounts [62,63].

Several modeling studies have investigated the dynamics of trehalose upon heat shock in recent years [25,35,37,38,64], which allows us to keep the discussion of background information to a minimum. In a nutshell, material is siphoned off glycolysis at the level of glucose 6-phosphate and channeled toward the production of glucose 1-phosphate, UDPG, glycogen, trehalose 6-phospate and trehalose, which accumulates in large quantities. The enzyme trehalase splits trehalose into two glucose molecules and thereby completes the trehalose cycle (see Figure 2.9). Because the present study is focused on methodological advances rather than new biological insights, we take the S-system model of the trehalose cycle in [25] at face value and analyze alternative operating strategies.

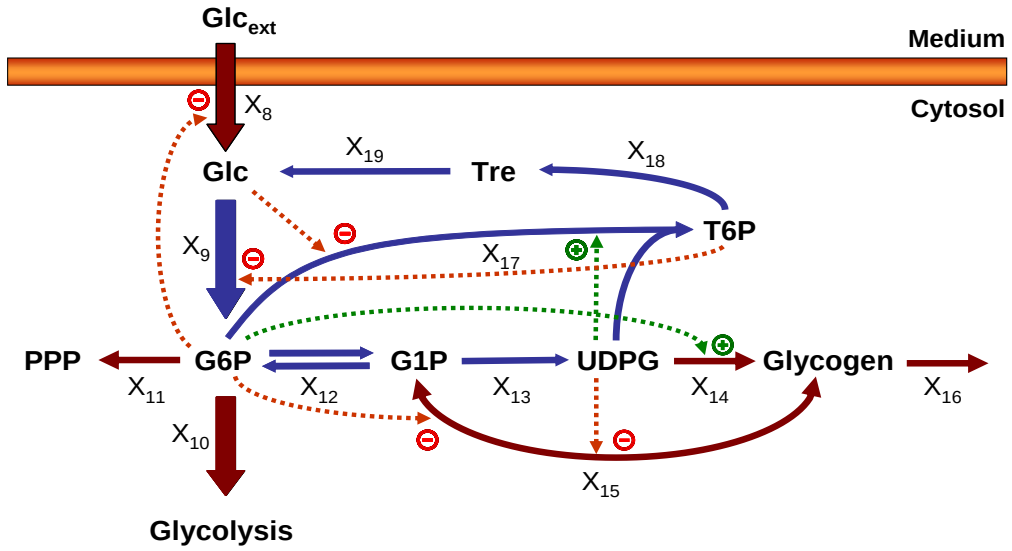


Figure 2.9: Diagram of the trehalose cycle (solid blue arrows) in yeast. Solid brown arrows represent other pertinent reactions. The main glycolytic flux is presented with heavy arrows. Red dotted arrows with associated minus signs indicate inhibition, while green dotted arrows associated with plus signs indicate activation. Abbreviations: Glc_{ext}: external glucose; Glc: internal glucose (X_1); G6P: glucose 6-phosphate (X_2); G1P: glucose 1-phosphate (X_3); UDGP: uridine diphosphate glucose (X_4); glycogen (X_5); T6P: trehalose 6-phosphate (X_6); Tre: trehalose (X_7); PPP: pentose phosphate pathway. X_8, …, X_{19} represent independent variables (see Table 2.3).

The S-system equations describing the system were taken directly from [25]. They are

Glucose: $$\dot{X}_1 = 31.912 X_0^{0.968} X_2^{-0.194} X_7^{0.00968} X_8^{0.968} X_{19}^{0.0323} - 89.935 X_1^{0.75} X_6^{-0.4} X_9$$

G6P: $$\dot{X}_2 = 142.72 X_1^{0.517} X_2^{-0.179} X_3^{0.183} X_6^{-0.276} X_9^{0.689} X_{12r}^{0.311}$$

$$- 30.120 X_1^{-0.00333} X_2^{0.575} X_3^{-0.17} X_4^{0.00333} X_{10}^{0.5111} X_{11}^{0.0667} X_{12f}^{0.411} X_{17}^{0.0111}$$

G1P: $$\dot{X}_3 = 7.8819 X_2^{0.394} X_3^{-0.392} X_4^{-0.010} X_5^{0.0128} X_{12f}^{0.949} X_{15r}^{0.0513}$$

$$- 76.434 X_2^{-0.412} X_3^{0.593} X_{12r}^{0.718} X_{13}^{0.180} X_{15f}^{0.103}$$

UDPG: $$\dot{X}_4 = 11.070 X_3^{0.5} X_{13} - 3.4556 X_1^{-0.0429} X_2^{0.214} X_4^{0.386} X_{14}^{0.857} X_{17}^{0.143}$$

Glycogen: $$\dot{X}_5 = 11.060 X_2^{0.040} X_3^{0.320} X_4^{0.160} X_{14}^{0.600} X_{15f}^{0.400}$$

$$- 4.9290 X_2^{-0.04} X_4^{-0.04} X_5^{0.25} X_{15r}^{0.200} X_{16}^{0.800}$$

T6P: $$\dot{X}_6 = 0.19424 X_1^{-0.300} X_2^{0.300} X_4^{0.300} X_{17} - 1.0939 X_6^{0.200} X_{18}$$

Trehalose: $$\dot{X}_7 = 1.0939 X_6^{0.200} X_{18} - 1.2288 X_7^{0.300} X_{19} \qquad (2.29)$$

Of primary interest here is the response of yeast to heat stress, which affects most of the reactions steps in the pathway. According to literature studies (cited in [25]), the alterations among the dependent and independent variables under heat stress are distinctly different, with some variables and steps changing substantially and other not as much or not at all (Tables 2.2 and 2.3).

Table 2.2: Dependent variables of the canonical model (Eq. 2.29) of the trehalose cycle. Steady-state values under optimal temperature conditions were collected from the literature [25]; heat-stress values (scaled by optimal steady-state values) computed with the S-system model upon changes in independent variables as shown in Table 2.3.

Metabolite	Variable Name	Steady-State Concentration [mM] under Optimal Temperature Conditions (from the Literature)	Computed Fold Change in Steady-State Concentration during Heat Stress (Scaled by Normal Steady State)
Glucose	X_1	0.03	1.46
Glucose 6-Phosphate	X_2	1	5.54
Glucose 1-Phosphate	X_3	0.1	3.99
Uridine Diphosphate Glucose	X_4	0.7	2.69
Glycogen	X_5	1	55.8
Trehalose 6-Phosphate	X_6	0.02	4.28
Trehalose	X_7	0.05	103

Table 2.3: Enzyme activities and transport steps in the canonical model (Eq. 2.29) and their fold-increases in response to heat stress

Catalytic or Transport Step	Variable Name	Heat-induced Fold-Increase in Activity Used in the Model
Glucose Transport into Cell	X_8	8
Hexokinase / Glucokinase	X_9	8
Phosphofructokinase	X_{10}	1
G6P Dehydrogenase	X_{11}	6
Phosphoglucomutase	X_{12}	16
UDPG pyrophosphorylase	X_{13}	16
Glycogen synthase	X_{14}	16
Glycogen phosphorylase	X_{15}	50
Glycogen use	X_{16}	16
α, α -T6P synthase	X_{17}	12
α, α -T6P phosphatase	X_{18}	18
Trehalase	X_{19}	6

In this case, n = 7 and m = 12, which indicates quite a bit of flexibility among different solutions. Application of the pseudo-inverse method reveals the space of all admissible solutions, and a possible solution is $\mathbf{X}_I = [5.4169\ 5.2877\ 0.9723\ 2.3770\ 2.3004\ 3.0197\ 2.8855\ 2.8574\ 2.1650\ 2.9739\ 4.4620\ 1.4873]^T$. Specifically, $\mathbf{X}_I$ is computed using the pseudo-inverse of $\mathbf{L}$ and the original basis of the null space of $\mathbf{L}$, which was obtained through singular value decomposition in MatLab, and $\boldsymbol{\lambda} = [1, 1\ 1\ 1\ 1]^T$. As to be expected, this vector of independent variables moves the system to the target steady state. However, the solution is much slower than the observed solution (Figure 2.10).

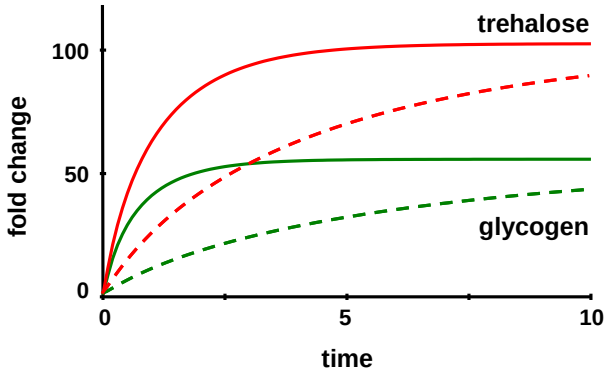


Figure 2.10: A possible solution within the space characterized by the pseudo-inverse method (dashed), in comparison with the nominal solution discussed in [25]. While both solutions reach the same steady state, the transient of the solution computed here is comparatively slow (see text for details).

The solution space obtained with the pseudo-inverse method is 5-dimensional, and a basis is

$\mathbf{B}_1 = [0.1635\ \ 0.1582\ \ 0.2144\ \ -0.1562\ \ 0.0986\ \ -0.1328\ \ -0.1543\ \ 0.9026\ \ 0.1099\ \ -0.0034\ \ -0.0034\ \ -0.0034]^T$

$\mathbf{B}_2 = [-0.2260\ \ -0.2188\ \ -0.3985\ \ 0.2716\ \ 0.3480\ \ 0.3708\ \ 0.4330\ \ 0.2639\ \ 0.3907\ \ -0.0017\ \ -0.0017\ \ -0.0017]^T$

$\mathbf{B}_3 = [-0.2260\ \ -0.2087\ \ -0.1967\ \ -0.7547\ \ 0.0363\ \ 0.0466\ \ 0.0021\ \ 0.0002\ \ 0.0016\ \ 0.3128\ \ 0.3128\ \ 0.3128]^T$

$\mathbf{B}_4 = [0.3576\ \ 0.3568\ \ 0.4235\ \ 0.1207\ \ 0.1771\ \ 0.2723\ \ 0.2627\ \ -0.1569\ \ 0.1578\ \ 0.3301\ \ 0.3301\ \ 0.3301]^T$

$\mathbf{B}_5 = [-0.1451\ \ -0.1291\ \ -0.2276\ \ 0.5556\ \ -0.1360\ \ -0.2193\ \ -0.3147\ \ 0.1558\ \ -0.1971\ \ 0.3527\ \ 0.3527\ 0.3527]^T$

As in the illustrative example, it is to some degree possible to affect the transient speed by manipulating the basis vectors. Tuning $\mathbf{B}_1$ or $\mathbf{B}_2$ causes the glycogen concentration to speed up but has almost no effect on trehalose or the other variables. Increasing $\mathbf{B}_3$ accelerates trehalose and no other variables, increasing $\mathbf{B}_4$ speeds up both trehalose and glycogen, while increasing $\mathbf{B}_5$ speeds up trehalose but slows down glycogen production (Figure 2.11).

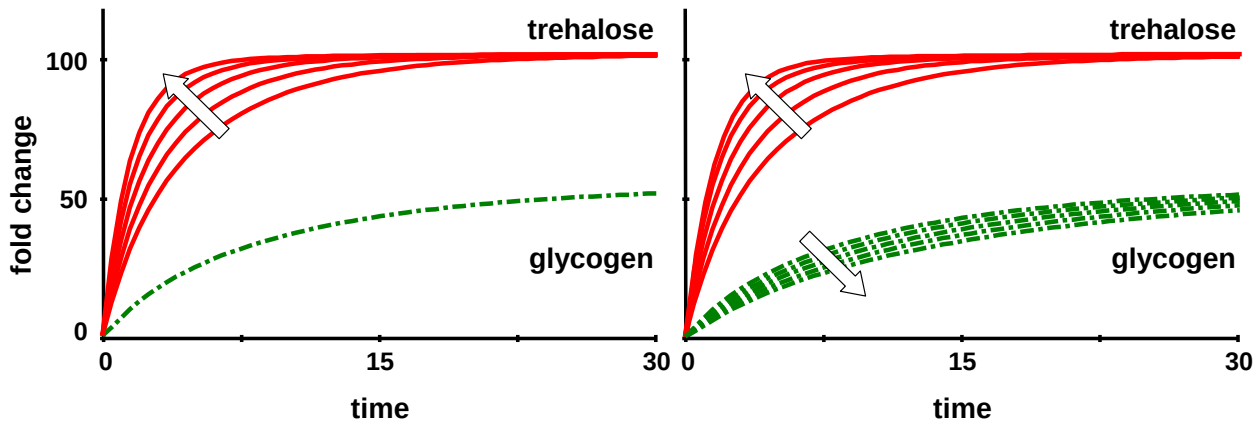


Figure 2.11: The solutions obtained with the pseudo-inverse method can be manipulated by modifying the basis vectors. In the left panel, basis vector B_3 was multiplied with factors 1, …, 5 (in direction of the arrow); this action did not affect the glycogen profile. In the right panel, basis vector B_5 was multiplied with factors 1, …, 5, in direction of the arrows. All solutions eventually reach the same target steady state.

In contrast to exploring the entire solution space, the direct optimization method allows us to select criteria of functional effectiveness *a priori* and to optimize the solution toward these criteria under the constraint that the target steady state is reached. As the first example, suppose the overriding criterion is to alter the independent variables as little as possible in magnitude. Least squares optimization toward this criterion yields a solution that of course reaches the target steady state and also exhibits only modest variations in enzyme activities (Table 2.4; column 3).

As a second example, we mandate to keep the number of altered independent variables to a minimum. MILP optimization reveals that this minimum number is 7, and the steady state is reached upon quite strong alterations in this minimum set (Table 2.4; column 4).

There are unlimited combinations on this theme, depending on the choice of criteria of functional effectiveness. For instance, we may consider a more complex scenario, which accommodates the following criteria. First, suppose that the phosphofructokinase step (X_{10}) cannot be altered. This situation was actually observed in yeast (*cf.* [35]), and rationale for this restriction was presented based on different types of analyses [37,38]. The restriction is easy to implement by fixing $X_{10} = 1$. Secondly, 20 transporters are involved in glucose uptake and fine-tuned for different glucose concentrations in the medium, which may

mean that the glucose transport step cannot be altered from what is observed. Thus, we enforce $X_8 = 8$, which corresponds to the observed level. Third, the rates of glycogen and trehalose production (V_5^+ and V_7^+) should be sufficiently large to achieve a timely response to elevated temperatures, which we implement by setting these flux rates to those actually observed and not permitting them to be altered. If subscript *e* indicates the nominal steady-state values under heat stress condition (as discussed in [25]), we obtain

$$
\begin{aligned}
& V_5^+ = V_{5e}^+ \\
& \Rightarrow \ln \alpha_5 + 0.04 y_{2e} + 0.32 y_{3e} + 0.16 y_{4e} + 0.6 y_{14} + 0.4 y_{15} = \\
& \quad \ln \alpha_5 + 0.04 y_{2e} + 0.32 y_{3e} + 0.16 y_{4e} + 0.6 y_{14e} + 0.4 y_{15e} \\
& \Rightarrow 0.6 y_{14} + 0.4 y_{15} = 0.6 y_{14e} + 0.4 y_{15e}
\end{aligned}
$$

$$
\begin{aligned}
& V_7^+ = V_{7e}^+ \\
& \Rightarrow \ln \alpha_7 + 0.2 y_{6e} + y_{18} = \ln \alpha_7 + y_{6e} + y_{18e} \\
& \Rightarrow y_{18} = y_{18e}
\end{aligned}
$$

These conditions impose further constraints on the system and are easily formulated in the MILP. The size of the solution space (number of free variables minus the number of linearly independent equality constraints) is now drastically reduced from 5 to 1. Within the constrained system, we can again identify the minimal set or least squares solution (Table 2.4, Columns 5 and 6) or could use some other criterion of function effectiveness.

Both results are interesting. First, the constrained least-squares solution turns out to be very similar to the nominal solution. Second, the minimum-set solution shows drastically different values than the nominal solution and identifies glycogen phosphorylase as the most dispensable reaction step. In an entirely different study [62], this same step was also identified as only modestly relevant for the trehalose response.

Table 2.4: Different implementations of computed heat stress responses, which all lead to exactly the same target steady state.

Catalytic or Transport Step	Nominal	Least Squares	Minimum Set	Minimum Set (X_8, X_{10}, V_5^+, and V_7^+ fixed)	Least Squares (X_8, X_{10}, V_5^+, and V_7^+ fixed)
Glucose transport	8	2.0096	4.6155	8 (fixed)	8 (fixed)
Hexokinase/Glucokinase	8	1.9440	4.4334	8	8
Phosphofructokinase	1	0.3577	1	1 (fixed)	1 (fixed)
G6P dehydrogenase	6	0.8745	1	1.6377	6.2371
Phosphoglucomutase	16	0.8435	1.1046	38.0406	15.5916
UDPG pyrophosphorylase	16	1.1110	1.5932	149.5541	14.9673
Glycogen synthase	16	1.0616	1.4620	217.1534	14.8016
Glycogen phosphorylase	50	1.0512	1	1	56.1937
Glycogen use	16	0.7965	1	42.5464	15.5396
α, α-T6P synthase	12	1.0942	2	12	12
α, α-T6P phosphatase	18	1.6413	3	18	18
Trehalase	6	0.5471	1	6	6

Table 2.4 seems to indicate that much "cheaper" solutions than the nominal solution can be found, which raises the question of why the nominal solution employs alterations in independent variables that are so much more dramatic than the least squares or minimum set solutions. At least one answer can be found in the response time: although all solutions reach exactly the same steady state, the nominal solution is more than ten times faster than the least squares and minimum-set solutions (Figure 2.12; note different time scales).

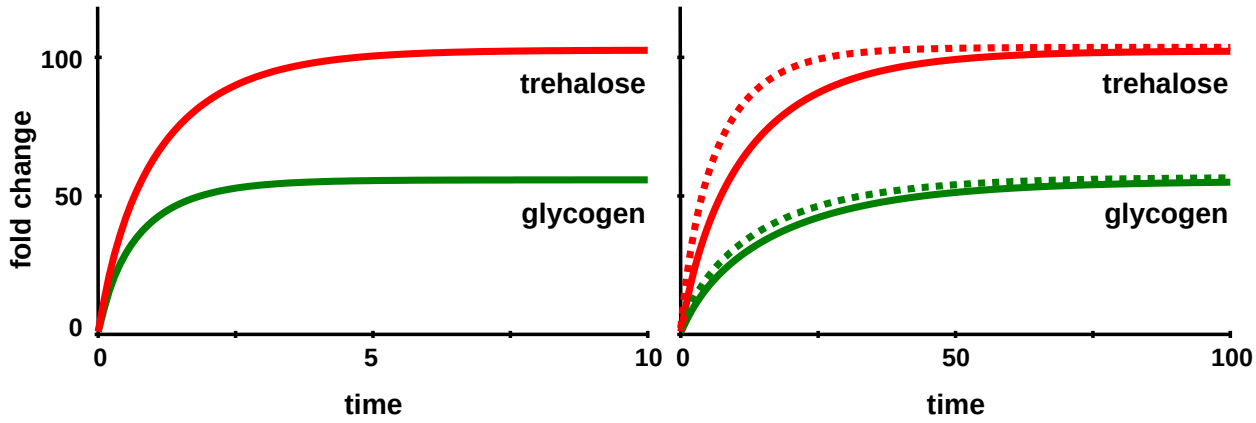


Figure 2.12: All solutions reach the exact same steady state and the transients have similar shapes, but the timing is quite different. While glycogen and trehalose in the nominal solution come close to their steady state values within about 5 time minutes (left panel), the same levels take ten or more times as long in the least squares (right panel; solid lines) and minimum set (right panel; dotted lines) solutions (note different time scales). Other variables respond on a more similar time scale (not shown).

The issue of drastically different transient speeds begs the question of whether and how the least squares and minimum set solutions could be accelerated. The most direct way of accomplishing acceleration arises if every flux contains its own independent variable. For instance, if every flux is governed by an enzyme which enters the flux with a kinetic order of 1, then multiplication by the same factor $\varphi > 1$ will speed up the dynamics of the entire system by φ. This advance does not come for free though, because the cost of the solution with respect to the chosen criterion increases and the result may no longer be optimal. For instance, the metabolic burden, which can be defined as the sum of independent variables, increases φ-fold. If the independent variables have different kinetic orders or appear in several equations, a systemic speed-up may still be possible. Specifically, one has to solve the equations

$$\prod_{j=n+1}^{n+m} \tilde{X}_j^{g_{ij}} = \varphi \prod_{j=n+1}^{n+m} X_j^{g_{ij}}$$

$$\prod_{j=n+1}^{n+m} \tilde{X}_j^{h_{ij}} = \varphi \prod_{j=n+1}^{n+m} X_j^{h_{ij}}$$

for all $i = 1, \ldots, n$. In the trehalose case, these conditions result in a set of 14 linear equations with 12 unknowns, which has no algebraic solution. Nevertheless, one can obtain a solution in a least squares sense, which indeed leads to an acceleration of the transients and approximately reaches the target steady state. The required changes in independent variables are presented in Table 5, where $\varphi_{LS} = 11.19$ and $\varphi_{MS} = 6.29$ are the acceleration factors for the least squares and the minimum set solutions, respectively. These factors are computed based on the settling time τ, which here is the amount of time needed for trehalose to reach and stay within 95% of its nominal heat shock value. While the resulting trehalose profiles are essentially the same as in the nominal scenario, the glycogen trends are still slower (Figure 2.13). Interestingly, the steps directly associated with the dynamics of trehalose are very similar to the nominal solution, and the glycogen phosphorylase step is again much lower (Table 2.5). Also of note is that glucose transport and the hexokinase step are almost the same.

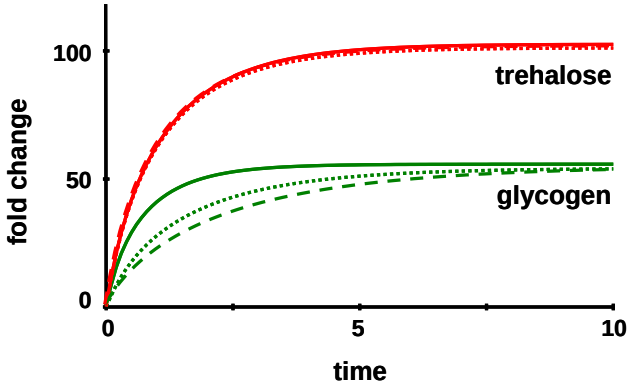


Figure 2.13: Acceleration of the least squares (dotted lines) and minimum set (dashed lines) solutions in the trehalose example leads to similar solutions as the nominal case (solid lines), but the accelerated solutions reach the target only approximately. The trehalose trend is now essentially indistinguishable from that in the nominal solution, while the rise in glycogen is slower. Time courses of the other variables essentially match those of the nominal case.

Table 2.5: Accelerated least squares and minimum set solutions for the trehalose cycle.

Catalytic or Transport Step	Nominal	Least Squares (accelerated)	Minimum Set (accelerated)
Glucose transport	8	22.4731	28.8238
Hexokinase/Glucokinase	8	21.7616	27.7065
Phosphofructokinase	1	5.2507	7.9468
G6P dehydrogenase	6	1.2416	1
Phosphoglucomutase	16	9.4252	6.8941
UDPG pyrophosphorylase	16	12.4311	9.9536
Glycogen synthase	16	11.8831	9.1365
Glycogen phosphorylase	50	11.7559	6.2449
Glycogen use	16	8.9158	6.2494
***α, α*-T6P synthase**	12	12.2458	12.4968
***α, α*-T6P phosphatase**	18	18.3691	18.7455
Trehalase	6	6.1230	6.2485

Distinctly different solutions to speeding up the transients could possibly be reached in two ways. First, the cell could initiate a fast transient toward a steady state with more extreme values than needed, and in a second phase relax these values toward the true target state. This strategy is expected to incur overshoots before the true target steady state is reached [42]. Second, it is possible to compute settings in independent variables that reach states that are not steady states. These computations require methods of nonlinear control theory, which were demonstrated for S-systems elsewhere [65].

2.3 Discussion

Deciphering how nature solves problems has been the dream of scientists for a long time. Consequently, enormous effort has been devoted to shining light on operating procedures in nature, dissecting systems, and identifying and characterizing processes that cells employ to solve specific problems. Given the seemingly unlimited variability and complexity of tasks that need to be addressed, a comprehensive understanding of operating procedures, let alone operating strategies or even operating principles, will not be gained in the foreseeable future. Nonetheless, the overwhelming magnitude of the challenge does not suggest that we give up, but that even small advances might be beneficial on our long journey.

We have shown in this article that a small sub-class of cellular tasks can be addressed quite efficiently with mathematical and computational tools. Namely, we propose methods for investigating the situation where a biological system is forced to move to a new steady state, which we assume to be known. For example, in the heat stress scenario discussed here, the cell must accumulate sufficient amounts of trehalose and possibly glycogen, while internal glucose and trehalose 6-phosphate need to be carefully controlled, because they cause adverse effects in high concentrations [63,66,67]. Thus, some pools in a pathway need to be altered substantially, while others must remain more or less at their nominal level. We show here that such tasks can be formulated rigorously in the language of linear algebra and constrained optimization.

The analysis yields two main results. First, it defines the entire solution space of the problem, and second, it allows a direct system optimization toward given criteria of functional effectiveness. The elegance of these solutions is primarily due to the special structure of S-system models, whose steady states are characterized by systems of linear equations. With the exception of Lotka-Volterra [68-70] and lin-log models [71,72], whose steady states are also governed by linear equations, it seems very difficult to obtain similarly general results with *ad hoc* models, such as pathway systems that are represented with Michaelis-Menten rate laws and their generalizations.

Interestingly, Generalized Mass Action (GMA) representations within BST [21,73], as well as other model structures, may permit numerical solutions under favorable conditions, although these solutions are not as general as in the case of S-systems. Namely, consider the important special case where each flux representation contains at most one independent variable, which enters the flux in a linear fashion, as it is typical for most enzymes. If all parameter values and the target steady state are known, all terms in the steady-state equations either become linear functions of one independent variable, or they do not contain an independent variable at all. Furthermore, outside the independent variables, all other components of each term combine to a single numerical value, so that the entire system of steady-state equations is linear in the independent variables. As in the cases shown here, this system may have a unique solution or be over- or underdetermined, and it can be analyzed in each case with methods of linear algebra and optimization. The condition of linearity with respect to independent variables can actually be further relaxed, for instance, to the requirement that the same independent variable, if it appears in different terms, always has the same kinetic order.

The tasks and solutions proposed here are reminiscent of optimization problems that have been analyzed in the field for two decades [20,37,50-53]. However, the two efforts represent different aspects of targeted alterations in pathways. In the typical optimization tasks in biotechnology or metabolic engineering, a metabolite pool or flux is to be maximized, while other features of the steady state profile are rather irrelevant as long as they remain within general physiological constraints. As a consequence, the task typically has a clearly defined, single optimal solution, although some cases permit alternative optima with the same value of the objective function. In the analysis here, the primary requirement is that the system must reach a specified steady-state profile. This task often admits an entire solution space, within which the system must operate. Within this space, questions of superiority of one solution over another with respect to selected criteria can be explored. Functional effectiveness is not usually considered in biotechnological optimization, but in the case analyzed here provides the metric for comparing alternative strategies and declaring one solution superior to another.

An unresolved issue is the definition of criteria for functional effectiveness, which are not necessarily known *a priori*. Is it advantageous to up-regulate just a few genes substantially, or is it better to up-regulate many genes by a small amount? We do not yet have answers to such questions, but we have taken a first step by asking these questions and by suggesting that it might be advisable to observe how nature solves tasks in order for us to develop ideas for what types of operating strategies might be candidates for optimality. Moreover, the work presented here suggests tools for comparing different solutions with objectivity and for declaring superiority of different alternatives once criteria are established.

CHAPTER 3

COORDINATION OF RAPID SPHINGOLIPID RESPONSES TO HEAT STRESS IN YEAST [2]

When yeast cells are exposed to heat, trehalose is produced in high quantities and sphingolipids respond rapidly and distinctly. The biochemical reactions governing this response form a complicated and dynamic metabolic system that regulates the concentrations of the key target sphingolipids, as well as their precursors and break-down products, in a highly coordinated and efficient manner. Changes in sphingolipids subsequently affect the expression of genes and, among a variety of response, the production of heat shock proteins (Figure 3.1).

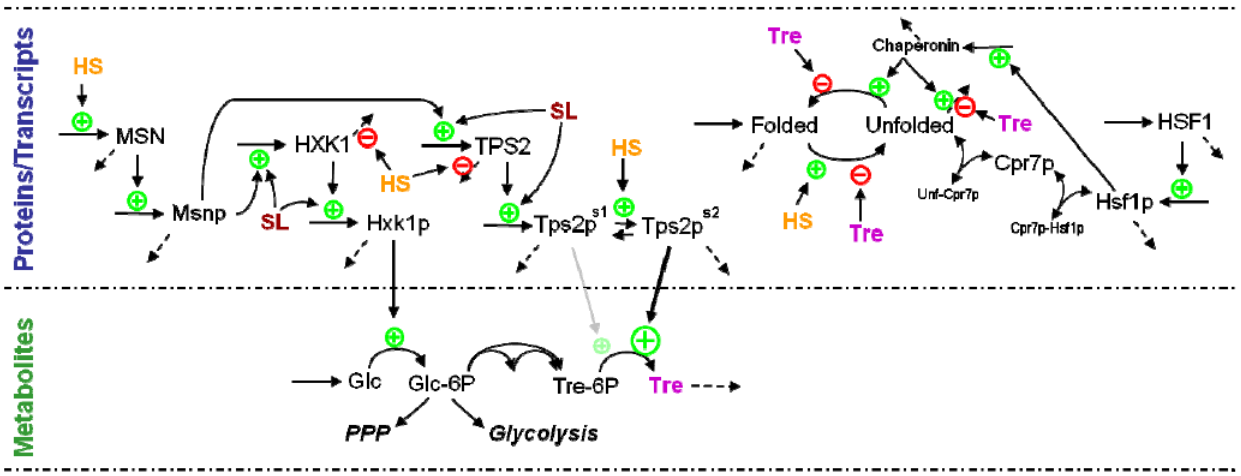


Figure 3.1. Multi-scale regulatory model of the heat stress response (from [74]). HS denotes location of direct action by heat stress; SL denote sphingolipid, while Tre refers to trehalose.

[2] Most of the material presented in this chapter was published in the articles "Canonical modeling of the multi-scale regulation of the heat stress response in yeast. *Metabolites* 2(1), 221-241, 2012" and "Coordination of rapid sphingolipid responses to heat stress in yeast", *PLOS Computational Biology*, 9(5), 2013.

In this chapter, I build upon the state-transition analyses of S-system models in Chapter 2 and expand the concepts toward a more sophisticated modeling approach that effectively addresses the dynamic heat stress responses of sphingolipids.

3.1 Introduction: Heat Stress Responses and the Roles of Sphingolipids

Any substantial increase in temperature has a direct effect on the macromolecules in a cell. Among them, proteins and lipids are most strongly affected. Nucleic acids can denature upon exposure to heat, but this process requires much higher temperatures of about 75°C – 100°C [75], which are outside the realm of tolerable heat stress.

Heat affects proteins in three ways. First, high temperature can modulate their synthesis from gene expression. In this context, Castells-Roca and colleagues investigated transcription rates and the stability of various mRNAs in *S. cerevisiae* following a temperature shift from 25°C to 37°C, and concluded that both were affected [76]. Second, processes of protein inactivation are temperature dependent. And third, heat can change a protein's folding state, which in turn may affect its function, as well as its removal by the proteasome. In particular, if the protein is an enzyme, its activity is influenced directly by its ambient temperature, according to an empirical relationship commonly called the Arrhenius effect or the Q_{10} effect.

Lipids are major constituents of membranes, and although the effects of heat are not completely understood, it appears that changes in temperature have an impact on membrane stiffness and fluidity [77]. Jenkins and coworkers [78] were among the first to connect sphingolipids to heat stress responses in yeast, demonstrating that these lipids play several particularly important roles (see also [79-83]). They subdivided the heat stress response into two phases. During the first phase, the cell needs to gain thermotolerance, which is at least partially accomplished with an accumulation of trehalose and the induction of heat shock proteins. Furthermore, the cell arrests its cell cycle in G0/G1, and this arrest lasts for approximately one

hour, during which time there is no growth. Once thermotolerance is achieved, the cell culture starts growing again in the second phase of the response, even if the temperature is still elevated.

The first response phase is directly associated with two distinct features of sphingolipids. First, the structural characteristics of complex sphingolipids, together with sterols, contribute to the physical organization of specific membrane microdomains within membranes, called lipid-rafts. These rafts are known to be associated with membrane fluidity, protein compartmentalization, and protein sorting and trafficking through membranes (*e.g.*, [84-86]). As core components of rafts, sphingolipids are thus directly involved in organizational structures with potential signaling functions, and alterations in these functions are effective at a short time scale [87].

The second role of sphingolipids in the early heat stress response is their capacity to serve as bioactive signaling molecules. This signaling function influences the regulation of the cell cycle response, nutrient uptake, and the synthesis of proteins, which can have important secondary effects, especially if heat shock proteins are not available to serve as protectors of other proteins [88,89]. Indeed, the groups of Ferguson-Yankey and Meier demonstrated that sphingolipid synthesis is required for an efficient initiation of translation, especially during heat stress [90,91]. Specifically, the translation rate is increased if sphingoid bases are synthesized and accumulate. Jenkins and collaborators [92] and Dickson and co-workers [79] showed that ceramides and simple sphingolipids, such as dihydrosphingosine and phytosphingosine (DHS and PHS), accumulate during heat stress in yeast. It appears that the short-term signaling role of sphingolipids is biphasic. In the first phase, sphingoid bases are required to regulate translation of heat shock mRNAs, a process that depends strongly on Pkh kinase, but not on Ypk kinases, which act downstream of Pkh. The second phase consists of a general increase in translation, which is dependent on the function of heat shock proteins. Without these heat shock proteins, the cell would run a severely elevated risk of protein aggregation or misfolding [91].

Sphingolipids also play roles over a longer time horizon. It has been known for a while that DHS induces the expression of a *STRE-LacZ* reporter gene, suggesting that the global stress element *STRE* can be activated by sphingolipid signals [79]. In particular, genes associated with the important trehalose stress

response contain multiple copies of *STRE*. Knock-outs or overexpression of genes coding for the synthesis of dihydrosphingosine-1-phosphate (DHS-1P) show changes that resemble thermotolerant and heat sensitive yeast phenotypes, indicating that DHS-1P is an important regulator of heat stress [93]. Phytosphingosine-1-phosphate is involved with the regulation of genes required for mitochondrial respiration [94]. More generally, modulations in any of the sphingolipid enzymes cause ripple effects that change the concentrations of many sphingolipids and, possibly, the expression of a variety of genes. Futerman and Hannun [95] summarized the long-term signaling mechanisms of simple sphingolipids including sphingosine-1-phosphate, sphingosine, ceramide and ceramide-1-phosphate in yeast.

Taken together it is evident that sphingolipids exert important roles within the coordinated heat-stress responses of a cell, and that these roles are pertinent over short and long time horizons. However, it is so far unclear how the cell is able to establish an appropriate sphingolipid profile very quickly in response to heat stress. To answer this question, we propose a computational analysis based on observed heat stress time courses and a dynamic model of sphingolipid biosynthesis and degradation that allows us to investigate the dynamic profiles of critical enzymes involved in the sphingolipid pathway.

3.2 Results

If changes in enzyme activities, for instance in response to heat, could be measured directly, the altered values could readily be entered into computational model equations [10,11], and solving the equations would show the time trajectories of all pertinent metabolites. Our task here requires the opposite task, which is much more complicated. Namely, we ask: can we infer from the metabolite time courses which enzymes have to be altered dynamically, and by how much, in order for the model to generate the observed time-dependent metabolic profile? The optimization-simulation strategy proposed for this analysis, as detailed in the *Methods* section, answers this question and reveals for the first time how the activities of key sphingolipid enzymes are adjusted by the cell during the heat stress response.

Specifically, we performed over 4,400 Monte-Carlo optimizations with random instantiations and selected from among these the 2,004 best models, based on the sum of squared errors (SSEs). To test the validity of these results, we also used the Akaike Information Criterion (AICc) for model selection and found that models selected based on SSEs and AICc were highly similar. Over 99% of the models identified through SSE were also identified with AICc; Appendix 3 contains further information on this comparative analysis. The thus selected models yielded dynamic trends for each sphingolipid enzyme during the heat stress response. These sets of individual trajectories reveal interesting insights. Namely, the trajectories collectively form tight, time-dependent activity ranges for those enzymes that control the influx to, and efflux from, the core of the sphingolipid biosynthetic pathway system. In other words, these enzymes always exhibit essentially the same dynamic activity patterns, independent of the randomly initialized start values. Most of the enzymes at the periphery of the pathway system, by contrast, exhibit widely varying activity profiles that are thus not identifiable from the available metabolic time series data. These results are described and discussed in detail in the following sections.

As a first validation of the collective results, we calculated the average of each computationally inferred enzyme activity at each time point and entered it into the pathway model (see *Methods* section) to check whether we were able to recoup the observed sphingolipid dynamics. The reconstructed sphingolipid dynamics indeed matches the original data quite well (Figure 3.2). This good match is by no means *a priori* guaranteed, because it is known that averages of parameter values from different good data fits do not necessarily correspond to good data fits themselves [96]. The averaged model was subsequently used for further interpretations of our results.

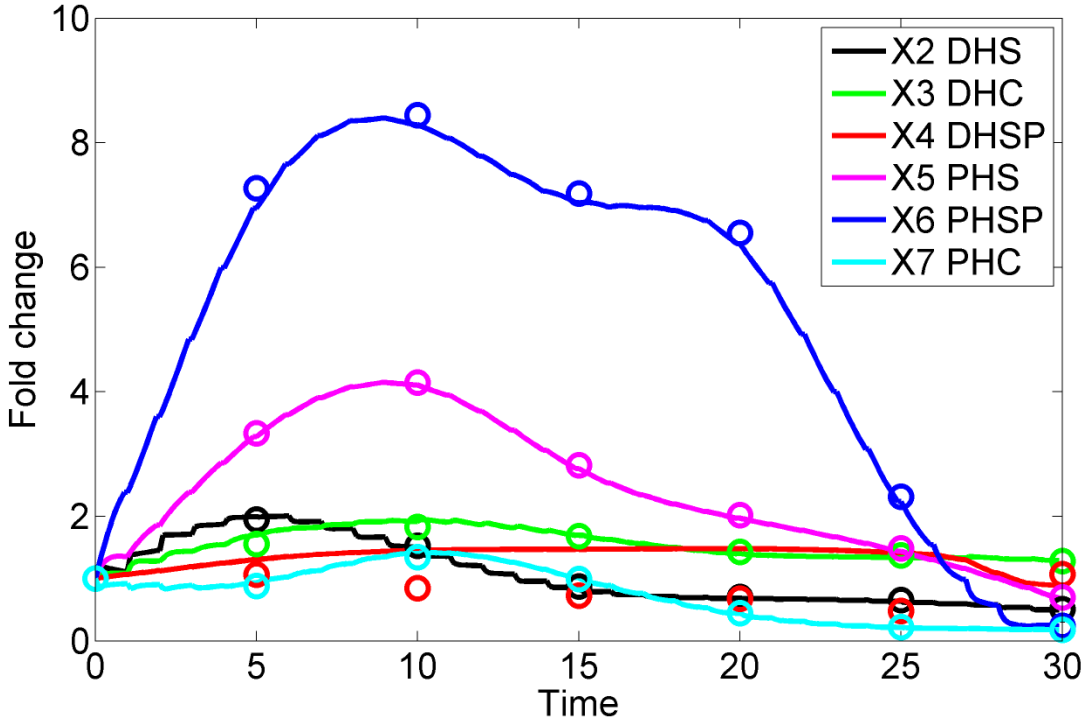


Figure 3.2. Data fit of the model with inferred enzyme activities. *Using averaged trends in enzyme activities leads to simulated metabolic profiles (lines) that reflect the observations (symbols; averaged from two experiments) quite well. The lines are segmented, because the model is solved with enzyme activities that are constant from each time point to the next, when they are dynamically reset. The fold changes in DHS-P do not seem to be modeled very well. The reason is that the absolute concentration of this metabolite is very small (Figure 3.11) and any fold change becomes vastly amplified.*

As a second, independent validation experiment, we explored changes in the concentrations of the complex sphingolipids IPC, MIPC, and $M(IP)_2C$ with the computationally inferred enzyme activities after a shift in temperature. In contrast to the profiles of simple sphingolipids (Figure 3.2), these trend lines are essentially flat, indicating that the complex sphingolipids do not change much during the heat stress response (Figure 3.3). This finding is directly consistent with experimental data [78] that were not used in our optimization.

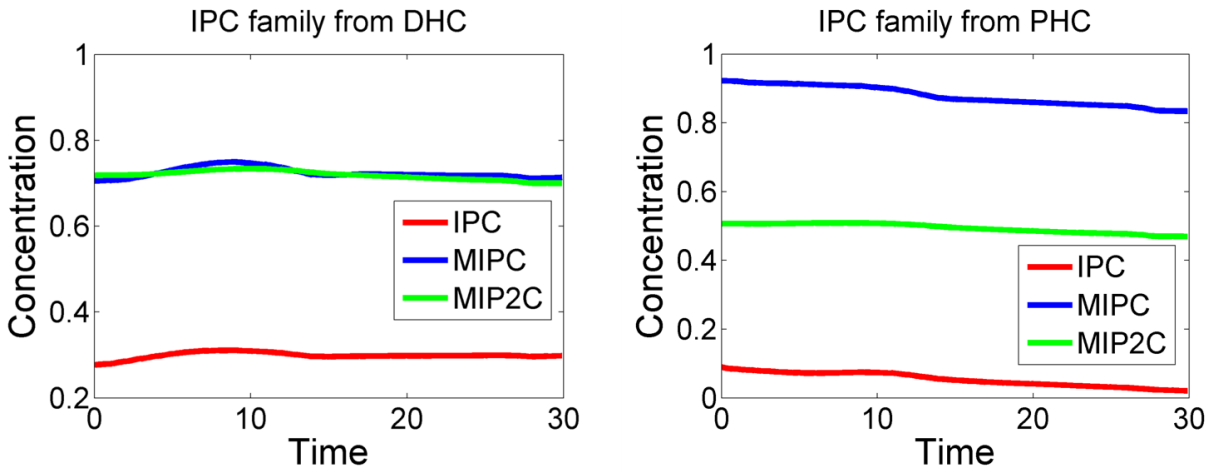


Figure 3.3. Concentration trends in complex sphingolipids. *While the sphingoid bases and ceramides exhibit strong responses to heat stress (Figure 3.2), the complex sphingolipids IPC, MIPC, and M(IP)2C remain almost constant. The left and right panels show levels of members of the IPC family, derived from dihydroceramide and phytoceramide, respectively.*

As a final validation approach, and quasi as a negative control, we fixed those key enzymes that were inferred to have tight activity ranges (X_{34}, X_{36}, X_{41}, X_{43}, X_{50}, X_{54}, X_{57} and X_{59}, see Figures. 3.4 and 3.5) at their nominal steady-state values and optimized all other enzyme activity profiles with the same methods as before. The resulting fit (Figure A3.7) is not good and much inferior to that in Figure 3.2; further details regarding this negative control are given in Appendix 3.

More interesting than these overall validation results are the trends in the individual enzyme activities (Figs. 3.4 – 3.9). Each panel in each of these figures shows grey lines, which are often so dense that they seem to form shaded areas. Each line is one of 2,004 simulated trend lines and represents the computationally inferred activity of the given enzyme at time points 1, …, 30, given a random initialization at $t = 0$. The red line in each panel shows the mean of the trend lines, while the dotted blue lines enclose 95% of the grey trend lines. The collective results from these panels are visualized in a different manner in Figure 3.10, where they are superimposed on the sphingolipid pathway system.

The first enzyme of interest, serine palmitoyltransferase (SPT; X_{57}) is the key bottleneck through which all *de novo* biosynthesis must pass (see red zone in Figure 3.10). The results show that the computationally inferred solution has SPT activity increasing briefly and then converging essentially to

zero within a few minutes (Figure 3.4). This pattern is seen in essentially all 2004 simulations with random initial settings (see *Methods* Section). The representation of fold changes seems most intuitive. However, the same results are also presented in Appendix 3 on a $\log_2$ scale, which stretches reduced activity levels. Changes in the subsequent, very fast step (3KDHS reductase; X_{27}) are less defined. A possible explanation is that the substrate of this reaction is toxic [97] and therefore never present in large concentrations, so that the capacity of the enzyme is not limiting. As a consequence, this enzyme activity does not contribute much to the error function that is to be minimized.

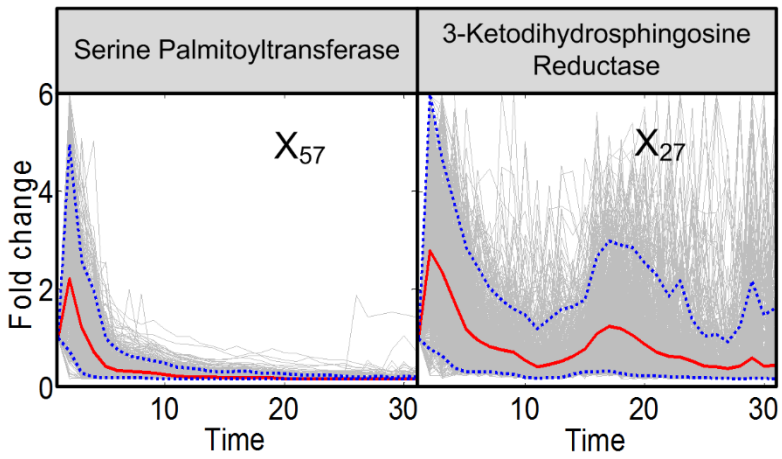


Figure 3.4. Trends in activities of enzyme at the entry point of sphingolipid biosynthesis. *Serine palmitoyltransferase and 3-KDHS reductase are enzymes responsible for the production and degradation of 3-KDHS, which is the key initial metabolite of sphingolipid biosynthesis. The enzymes are located in the red zone of Figure 3.10. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results.*

Similarly well defined as SPT are enzymes that catalyze the redistribution of material within the core of sphingolipid metabolism as well as the steps of sphingolipid removal (blue zone in Figure 3.10). These enzyme activities again rise quickly but approach a very small value shortly after (Figure 3.4). The very long chain fatty acid synthase and elongase (ELO1p; X_{59}) is responsible for the delivery of fatty acid-CoA to the sphingolipid system, while sphingosine-phosphate lyase (X_{50}) and GPI remodelase (X_{43}) are the only true exit routes out of central sphingolipid metabolism. The remaining enzymes in this group redistribute material within the pathway. Ceramide synthase (X_{34}) shows the same pattern as X_{59}, X_{50}, and

X_{43}, while sphingoid base kinase (X_{36}), sphingoid-1-phosphate phosphatase (X_{41}), and 4-hydroxylase (X_{54}) exhibit the same initial phase, but begin to rise more or less strongly after about 25 to 28 minutes of heat stress. These late increases in activity apparently indicate the first consequence of heat-induced gene expression. Among these enzymes, sphingoid-1-phosphate phosphatase (X_{41}) shows the strongest peak at 28 to 30 minutes by far. This enzyme is known to be a key regulator of sphingolipid metabolism and, in particular, of stress responses [81]. It plays an important role in regulating the crucial balance between ceramide and phosphorylated sphingoid base levels and thereby modulates later stress responses.

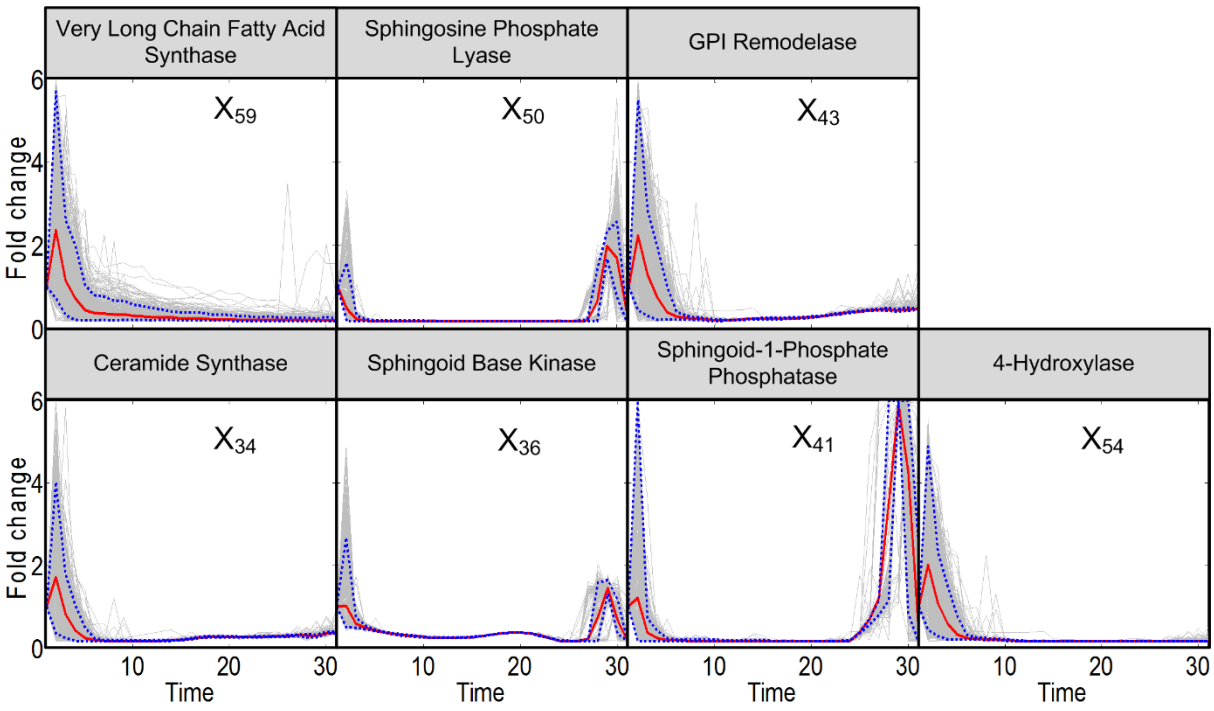


Figure 3.5. Trends in activities of enzymes in the core region of sphingolipid metabolism. *After an initial spike, all enzyme activities in this region are reduced to almost nil. The enzymes are located in the blue zone of Figure 3.10. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results.*

The two alkaline ceramidases exhibit rather different patterns. As with the previous enzymes, the activity of dihydroceramide alkaline ceramidase (dihydro-CDase; X_{29}), which converts dihydroceramide into dihydrosphingosine, decreases to almost zero, but much later and in a less defined manner. By contrast,

the activity of phytoceramide alkaline ceramidase (Phyto-CDase; X_{53}) shows tight trends consisting of three peaks, before returning to normalcy after about 30 minutes (Figure 3.6). These differences indicate that there is no "symmetry" between the function of dihydro- and phyto-forms of sphingolipids.

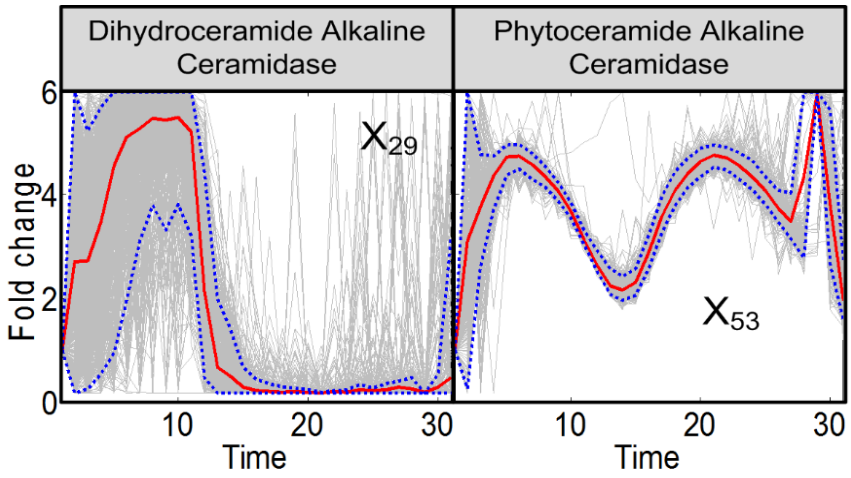


Figure 3.6. Trends in activities of the two alkaline ceramidases. *Dihydroceramide alkaline ceramidase and phytoceramide alkaline ceramidase, which convert the ceramide form into sphingosines, exhibit distinct activity patterns. The enzymes are shown with light blue circles in Figure 3.10. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results.*

The activity patterns of enzymes associated with complex sphingolipids are different; they are shown in Figure 3.7 (green zone in Figure 3.10). They all indicate a sustained level of hyper-activity for about 20 minutes, before becoming very low between about 20 and 28 minutes. These enzymes are inositol phosphorylceramide synthase (IPC synthase; X_{33}), mannosyl inositol phosphoceramide synthase (MIPC synthase; X_{35}), and mannosyl di-inositol phosphorylceramide synthase ($M(IP)_2C$ synthase; X_{55}), as well as inositol phosphosphingolipid phospholipase C (IPCase; ISC1 X_{51}), which returns IPC, MIPC and $M(IP)_2C$ to the dihydroceramide (DHC) and phytoceramide (PHC) pools.

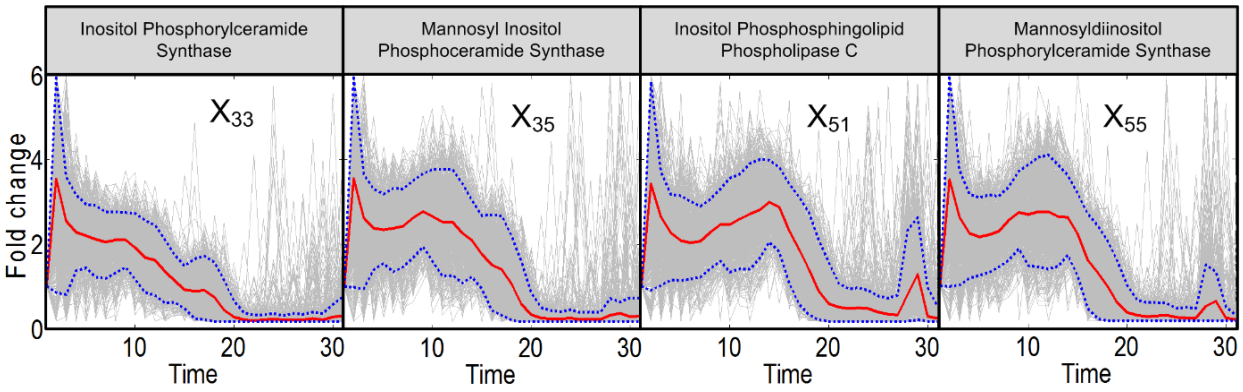


Figure 3.7. Trends in activities of enzymes associated with complex sphingolipids. *Enzymes interconverting complex sphingolipids are at first hyper-active, but tend to lose most activity at some point between 20 and 30 minutes. The enzymes are located in the green zone of Figure 3.10. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results.*

The remaining enzyme activities are not identifiable with our analysis. Some appear to be essentially unchanged throughout the measurement period of 30 minutes, during which the temperature remains elevated. Examples are fatty acid synthase (X_{52}), acetyl-coenzyme A carboxylase (X_{60}), and synthase (X_{63}) (Figure 3.8; yellow zone in Figure 3.10). Other enzyme activity patterns (X_{26}, X_{39}, X_{42}, X_{44}, X_{46}, X_{40}, X_{45}, X_{49}, X_{31}, X_{32}, X_{38}, and X_{56}) exhibit larger degrees of variation (Figure 8; pink and tan zones in Figure 9). On average, each pattern exhibits an individual Q_{10} effect, and subsequently stays more or less constant, decreases somewhat, or continues to increase slightly, but the trends are not clear. One reason for the large variability in these trends may be that the available metabolite data are not informative enough. It is also to be expected that the different processes catalyzed by these enzymes allow for a large degree of redundancy. For instance, serine is not only used in the SPT reaction, but also for the production of phosphoserine and in the serine hydroxymethyl transferase reaction, so that computationally inferred excesses in one reaction may be compensated numerically by a lower activity of one of the other two. Finally, as we discussed elsewhere [10], it is possible that these processes are not as well modeled as those at the core of sphingolipid biosynthesis, because they also participate in other pathway systems, such as phsopholipid or ergosterol metabolism.

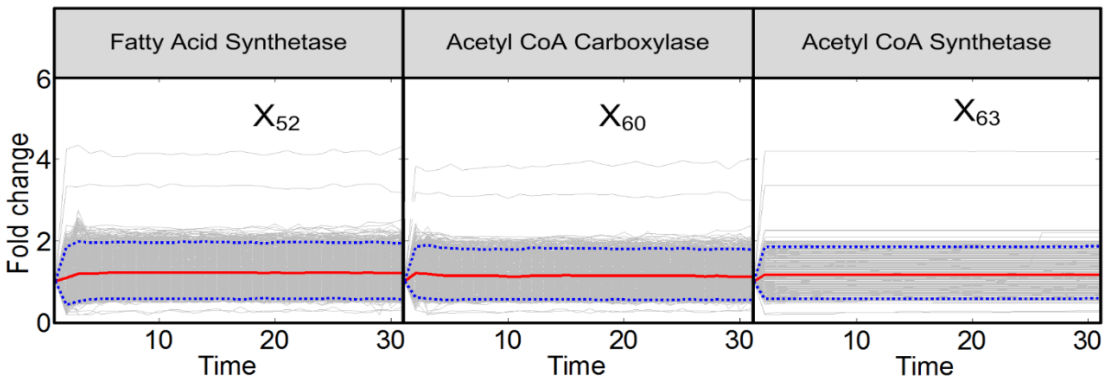


Figure 3.8. Trends in activities of enzymes associated with fatty acid CoA. *The enzymes shown here are responsible for CoA enlongation. The enzymes are located in the yellow zone of Figure 3.10. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results.*

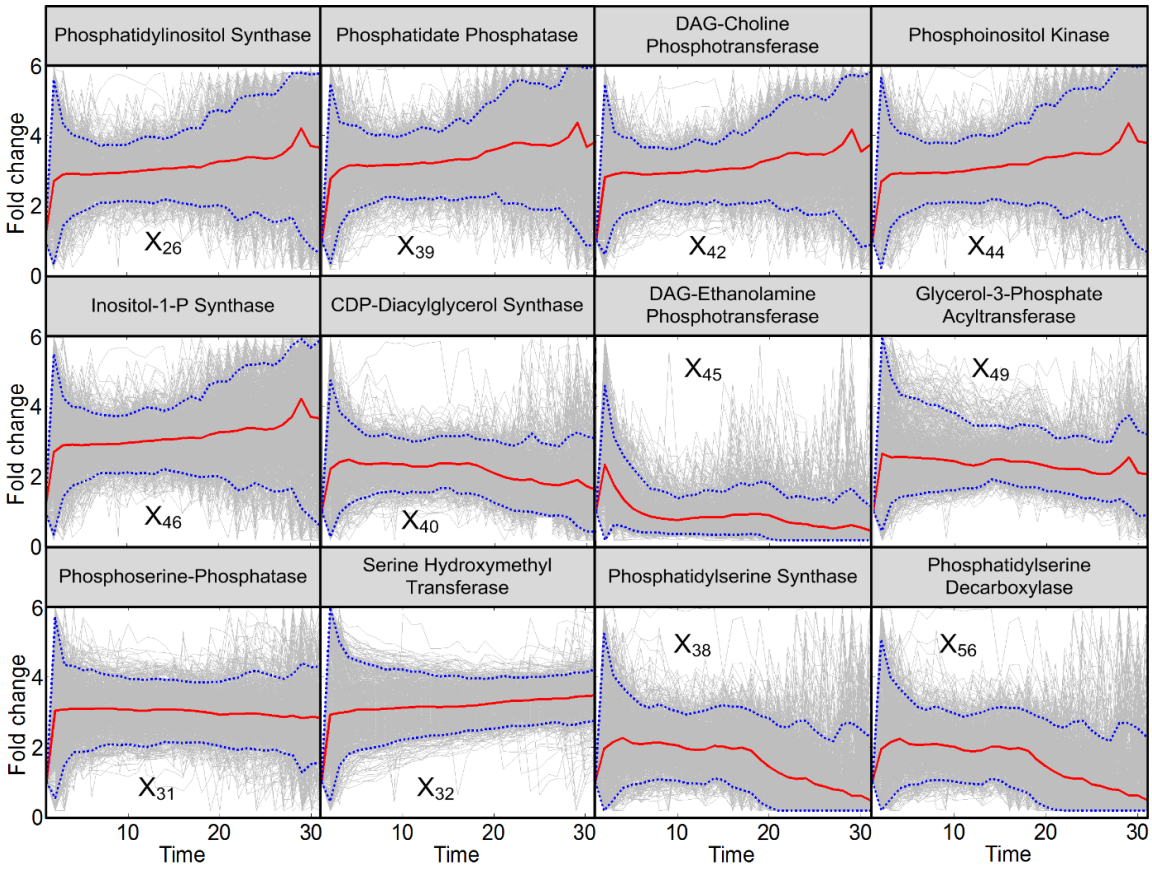


Figure 3.9. Trends in the remaining enzyme activities. *Activities of enzymes at the periphery of the pathway system are not identifiable, mainly due to insufficient information and the fact that these enzymes are also involved in other pathways. Enzymes in the two upper panels are related to the phospholipid metabolism and enzymes in the lower panel are related to serine metabolism. The enzymes are located in the tan and pink zones of Figure 3.10. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are averages, and dotted blue lines enclose 95% of the results.*

The computationally inferred patterns in enzyme activities are collectively depicted as colored zones in Figure 3.10.

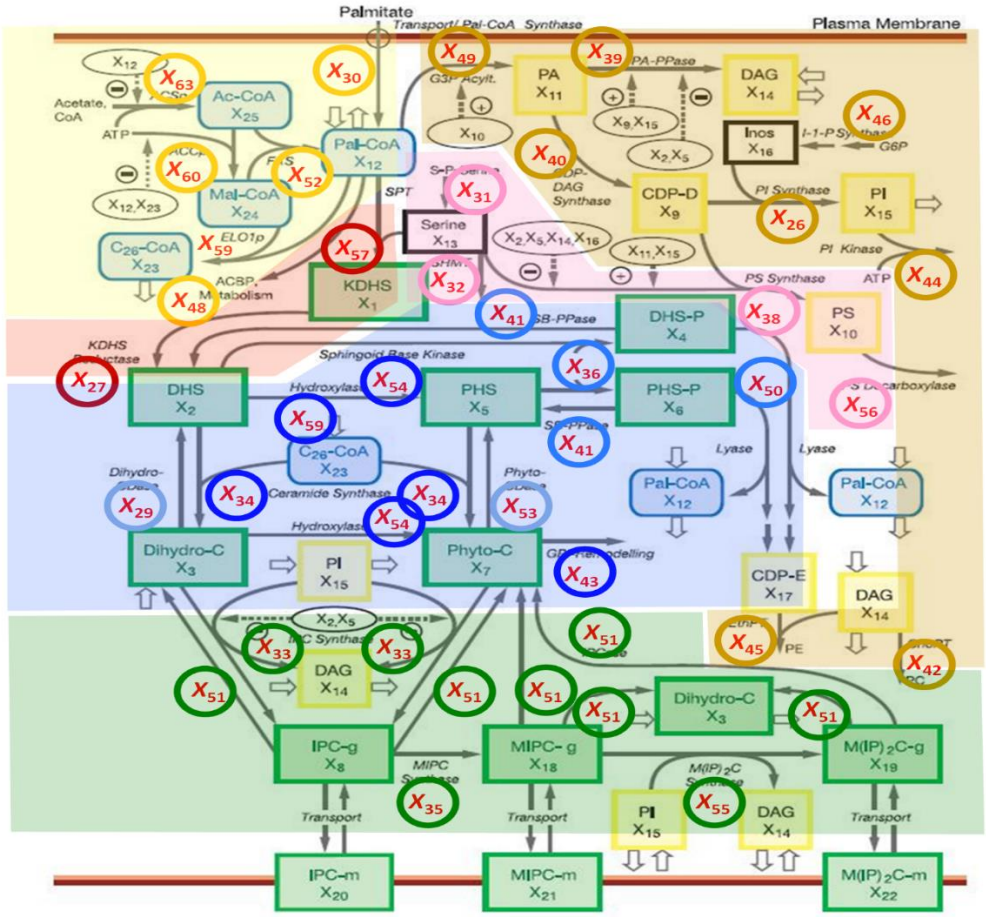


Figure 3.10. Zones of similar changes in enzyme activities. *The zones correspond to enzyme profiles in Figures. 3.4 (red), 3.5 and 3.6 (blue), 3.7 (green), 3.8 (yellow) and 3.9 (tan and pink) respectively. Abbreviations are: Green boxes (sphingolipid metabolism): KDHS (3 Ketodihydrosphingosine), DHS (Dihydrosphingosine), DHS-P (Dihydrosphingosine 1-phosphate), PHS (Phytosphingosine), PHS-P (Phytosphingosine 1-phosphate), DHC (Dihydroceramide), PHC (Phytoceramide), IPC-g (Inositol phosphorylceramide), MIPC-g (Mannosylinositol phosphorylceramide), $M(IP)_2C$-g (Mannosyldiinositol phosphorylceramide), IPC-m (Plasma membrane inositol phosphorylceramide), MIPC-m (Plasma membrane mannosylinositol phosphorylceramide), $M(IP)_2C$-m (Plasma membrane mannosyldiinositol phosphorylceramide). Yellow boxes (phospholipid metabolism): DAG (Sn-1,2-diacylglycerol), CDP-D (Cytidine diphosphate DAG), PS (Phosphatidylserine), PA (Phosphatidic acid), PI (Phosphatidylinositol), CDP-E (Cytidine diphosphate ethanolamine). Blue boxes (fatty-acid metabolism): Pal-CoA (Palmitoyl-Coenzyme), C26-CoA (Very long chain fatty acid), Mal-CoA (Malonyl coenzyme), Ac-CoA (Acetyl coenzyme). The base diagram was adapted from Alvarez-Vasquez et al., Nature 433(7024): 425-430, 2005.*

Most interesting are the red and blue zones, which control the influx to, redistribution within, and efflux out of the core of sphingolipid biosynthesis. The green zone contains the complex sphingolipds, which provide material for activities in the blue zone, even though their concentrations do not change much throughout the thirty minutes of heat stress (see Figure 3.3). The yellow, pink, and tan zones at the periphery contain fatty-acid CoAs, serine compounds and phospholipids, respectively. These are necessary for sphingolipid biosynthesis, but also for other pathways. Due to their multiple roles, they are presumably not modeled comprehensively, and their enzyme activities are not identifiable with the data and methods used here.

The trend lines, as well as their averages, collectively suggest that the sphingolipid heat stress response is achieved with quite moderate changes in many enzymes rather than very extensive changes in just a few key enzymes. This result is consistent with earlier studies in the context of the diauxic shift, which implied that cells probably satisfy altered metabolic demands with many small, rather than a few large, adjustments [9,98]. While it is impossible to identify the true advantage of this strategy unambiguously, the avoidance of large changes in any of the system components might be expected to reduce the risk of undesired side effects in neighboring pathways.

All trends in enzyme activities follow distinct patterns, which are the results of a balance among three forces induced by the shift in temperature from 30°C to 39°C: first, an essentially immediate increase in activity to a level of up to about four times the baseline, according to the enzyme's (typically unknown) Q_{10} value, which quantifies the Arrhenius effect (see Table A3.3 in Appendix 3); second, a possibly diminished activity due to partial protein unfolding and/or an altered half-life of the corresponding protein and/or mRNA; and third, changes in enzyme activity due to regulation and/or gene expression. These forces may be active to different degrees in overlapping time windows.

The three forces lead to different activity patterns. Most striking is the set of enzymes controlling the influxes and effluxes associated with the core of sphingolipid biosynthesis. Their pattern of heat

responses consists of enzyme activities that first exhibit a Q_{10} effect, which is subsequently counteracted by deactivation mechanisms that could be due to changes in RNA amounts, changes in half-lives or degradation rates of proteins or mRNAs, post-translational modifications, or heat induced gene depression [99]. Thus, after a few minutes, these enzyme activities essentially disappear.

3.2.1 Overall Heat Stress Response Strategy

Without any computational analysis, the measured data directly show which sphingolipids are apparently needed under heat stress at different points in time. Measured as absolute quantities, PHS increases by far the most in concentration, whereas PHS-P increases most relative to its baseline value. Interestingly, both adjustments are much stronger than in the corresponding dihydro-forms. For instance, the concentration of DHS-P remains very low throughout the observation period of 30 minutes (Figure 3.11). DHS reaches its modest peak earlier than PHS and PHS-P, whereas PHC reaches its peak later. It is difficult to discern the rationale for this timing and the differences in peak heights.

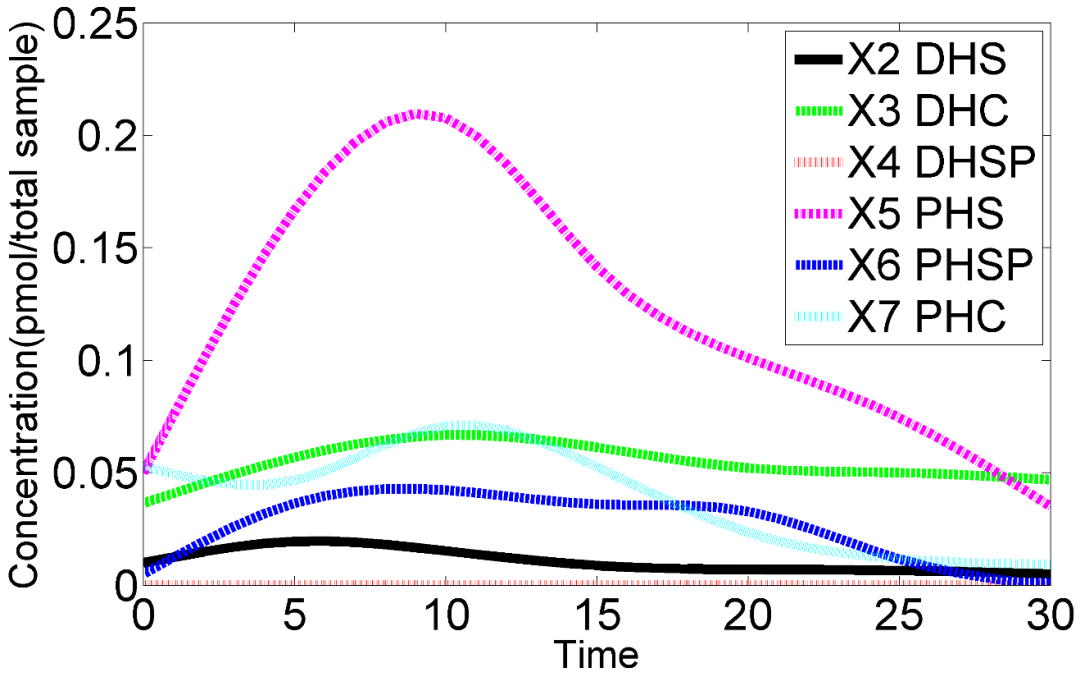


Figure 3.11. Smoothed time series data. *Absolute changes in six key sphingolipid metabolites in response to a temperature shift from 30C to 39C at time 0. The raw data were smoothed with a standard spline technique. See also Figure 3.2.*

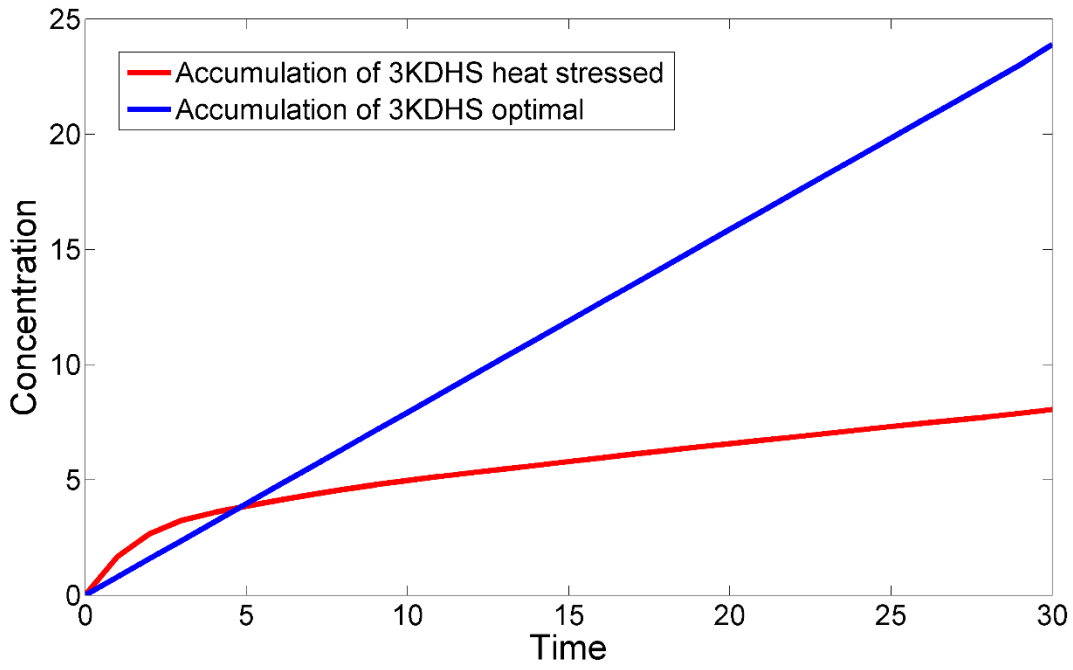


Figure 3.12. Smoothed time series data. *Absolute changes in six key sphingolipid metabolites in response to a temperature shift from 30°C to 39°C at time 0. The raw data were smoothed with a standard spline technique. See also Figure 3.1.*

What the computational analysis shown here suggests is how these observed adjustments are implemented by the cell. Initially, *de novo* biosynthesis increases quickly, but only for the first three or four minutes. The model actually allows us to quantify and compare the total amount of biosynthesis under optimal and heat stress conditions. Namely, we can record in the dynamic simulation the total production of 3-KDHS, while computationally omitting its degradation (Figure 3.12). Under optimal conditions, and with a constant influx of palmitate and serine, this accumulation is linear (blue line), and considering consumption as well, the concentration of 3KDHS is constant (results not shown). By contrast, under heat stress, the accumulation is faster for the first few minutes (red line), but it is increasingly reduced subsequently. Considering consumption as well, the concentration of 3KDHS decreases (results not shown).

In the next five to ten minutes, the patterns diverge strikingly. Probably most intriguing, both the input to, and the exit from, central sphingolipid metabolism are almost completely shut down. During this

time period, the cell not only counteracts the unavoidable Q_{10} effect in SPT, but down-regulates this enzyme to a mere residual amount, as shown in top left panel of Figure 3.4. Similarly, the exit routes through the lyase and remodelase steps lose activity about 5 minutes into the heat stress (Figure 3.5). The second step of *de novo* biosynthesis, KDHS reductase, is less dramatically affected (right panel in Figure 3.4), but deprived of substrate. This substrate deprivation appears to be safer than enzyme down-regulation, as 3KDHS is toxic [100] and any accumulation could be dangerous.

The computational deductions imply that *de novo* sphingolipid biosynthesis appears to be up-regulated only for the first few minutes [78]. To establish the needed changes in sphingolipid profile under heat stress, the cell appears to absorb and process residual substrate as vigorously as possible, but subsequently seems to count on the much more reliable use of existing complex sphingolipids for the generation of signaling molecules such as PHS, PHS-P and, to a lesser degree, DHS and DHC, and on a subsequent redistribution among the simple sphingolipid pools. This conclusion is based on the inferred reduction in biosynthesis after about five minutes, the shutting off of the lyase and remodelase steps, as well as three additional observations. First, IPCase (Figure 3.7) is strongly upregulated in a sustained manner for about 15 minutes. Second, the hydroxylase, which converts DHC into PHC and DHS into PHS, loses almost all activity throughout the measured time period (Figure 3.5). Third, processes leading to the synthesis of complex sphingolipids, including IPC synthase and the synthesis of PI and DAG, are down-regulated after about 15 minutes (Figure 3.7), thereby slowing down the genesis of new complex sphingolipids from simple sphingolipids. Several of the enzymes associated with complex sphingolipids begin to become active again about 28 minutes into the heat stress, which may be a consequence of changes in gene expression.

After 30 minutes, the six measured sphingolipid concentrations essentially return to their baseline levels. In stark contrast, the enzyme system has not returned to its original state, and several enzymes still exhibit an activity that is quite distinct from the profile under optimal temperature conditions. Thus, the cell, which is still under heat stress, is regaining a close resemblance of normalcy with respect to its metabolites, but this state is achieved with a significantly different flux and enzyme profile.

3.3 Discussion

In this part of the project, we have proposed a computational approach to analyze heat stress response strategies in yeast. Specifically, we have inferred how cells adjust their enzyme activities within sphingolipid metabolism, which has been demonstrated in numerous earlier reports as a heat sensitive signaling system. Using experimental measurements of metabolite concentrations following a shift in temperature, combined with a detailed dynamical model, we computationally inferred adjustments in enzyme activities that appear to be both sufficient and necessary for mounting the observed metabolic response. Rather than computing a single solution to the inverse task, we computed a comprehensive ensemble of over 4400 independent solutions and selected from among them the best 2004 solutions, based on SSE and AICc metrics. These 2004 solutions led to very similar trends in the activities of key enzymes, although not of enzymes at the periphery of the pathway system.

The computed results suggest, first, that the response to heat is not achieved by drastic changes in a few "key" enzymes, but that numerous enzymes are involved. Second, the dynamic alterations in activities differ substantially in both, magnitude and timing, as well as in the general shape of the enzyme activity trends throughout the observed 30-minute time window following the initiation of heat stress. The main surprise in our results is the deduction that the changes in sphingolipid profile are apparently not achieved by sustained increases in *de novo* biosynthesis but through a brief initial spike, followed by the retrieval of simple sphingolipids from membrane-associated complex sphingolipids, as well as a complicated redistribution scheme among the different ceramide and sphingosine forms. While this strategy was not expected, its seems to have merit, because the cell cannot be sure that new resources are quickly available for *de novo* synthesis of sphingolipids, while complex sphingolipids such as IPC, MIPC and $M(IP)_2C$ are integral components of membranes and therefore always available, with the possible exception of the most deprived situations. Thus, it seems that the cell sacrifices some of its membrane structures and recreates

them once the stress situation is under control. This sacrifice, however, is not very substantial, as the concentrations of complex sphingolipids change very little during the heat stress response (Figure 2). These results are consistent with experimental finding of Jenkins *et al.* [78], who studied different roles of sphingolipids during the heat stress response. Using isotope labeling, they showed that sphingoid bases and ceramides increase early on via *de novo* synthesis, but that IPC, MIPC and $M(IP)_2C$ remain essentially constant over a period of more than one hour. Wells *et al.* [83] also studied the formation of ceramide in response to heat stress and, using labeled phosphosphingolipids, and concluded that ceramide formation from IPC, MIPC, and $M(IP)_2C$ through the IPCase reaction was unlikely. However, the concentration profiles these authors observed were very different from those obtained by Cowart *et al.* [94], which we used here. In particular, under Wells' 39°C treatment, ceramide remained elevated at a level five times its baseline throughout the two-hour measurement period. Outside the fact that these authors studied a temperature shift from 24°C to 39°C, the differences in concentration profiles to those used here (Figure 1; Cowart *et al.* [94]) remain unexplained.

Although the computational results were obtained without any particular assumptions, some uncertainties are associated with the fact that many of the intermediate sphingolipids had not been measured and that the mathematical approach may not have revealed the one truly optimal solution. For instance, all results are obtained from large-scale simulations with a dynamical model that has been validated to some degree but could certainly be improved. Given the present data, it is unlikely that further simulations of the same type as shown here would lead to different results. However, if other metabolite concentrations could be measured, or if it were possible to determine some internal metabolic fluxes independently of the metabolite concentrations, the degree of reliability of our results would greatly increase.

The study presented here elucidates a systemic strategy with which the cell establishes the observed sphingolipid profile, but it does not address the specific roles of the various sphingolipids in the heat stress response. Interestingly, some of the simple sphingolipids that are known to have signaling roles do not change all that much, while others do. In particular, DHS, which activates the stress element STRE in the

expression of stress related genes, maximally rises to only about twice its normal level, about 5 minutes into the heat stress. Apparently, this increase is sufficient. By contrast, PHS-P, which was recently identified as an important gene regulator, rises to a level that corresponds to almost 10 times its baseline level and exhibits a sustained response that lasts over 20 minutes. PHS rises to a four-fold level. No direct signaling role is known, and it may just be that this compound is needed as a precursor of PHS-P.

The experiments generating the data used here exposed the cells to persistent heat stress. At the end of the 30-minute observation period, all six key sphingolipids have essentially returned to their normal levels, except for DHC, which still seems to be very slightly elevated. By contrast, many of the enzyme activities are not "back to normal." Expressed differently, the cell manages to mount a strong transient response, which is known to lead to longer-term genomic responses. Subsequently, within a total of just 30 minutes, it is able to adjust its catalytic machinery to the persistent heat conditions in such a manner that the fluxes exhibit a distinctly different activity pattern which, nevertheless, re-establishes a favorable metabolic state that is remarkably close to that under optimal conditions.

Our focus on sphingolipids sheds light on just one aspect of the well-coordinated, complex responses with which yeast adjusts to a new environmental condition. Nonetheless, this particular aspect is of special interest, as the roles of sphingolipids and their biosynthetic pathways have been preserved throughout evolution, from yeast to humans, where they are involved in numerous differentiation and disease processes (*e.g.*, [101-104]).

3.4 Methods

3.4.1 Data

The data, previously obtained in one of our labs, were described in the literature (see Supplements of [94]). They consist of duplicate 30-minute time courses of six key sphingolipids, collected following a step increase in temperature from 30°C to 39°C. Specifically, changes in metabolite concentrations were measured at baseline (t = 0; normal temperature) and at 5, 10, 15, 20, 25, and 30 minutes of heat stress. We used these measurements, averaged the duplicates, and then applied a smoothing spline technique to interpolate the trend of each time course so that concentration values at 31 time points (0, 1, …, 30 minutes) became available for each sphingolipid. The smoothed transients are shown as absolute concentrations in Figure 3.13 (see also Figure 3.2 for fold changes, which shows the smoothed data as symbols, along with a model fit based on averaged enzyme activities). For our computational analysis we used relative changes in each sphingolipid with respect to the baseline steady state before heat stress, which we directly obtained from the time series measurements, and scaled these with steady-state values, which were described in earlier work [11], to obtain actual concentrations.

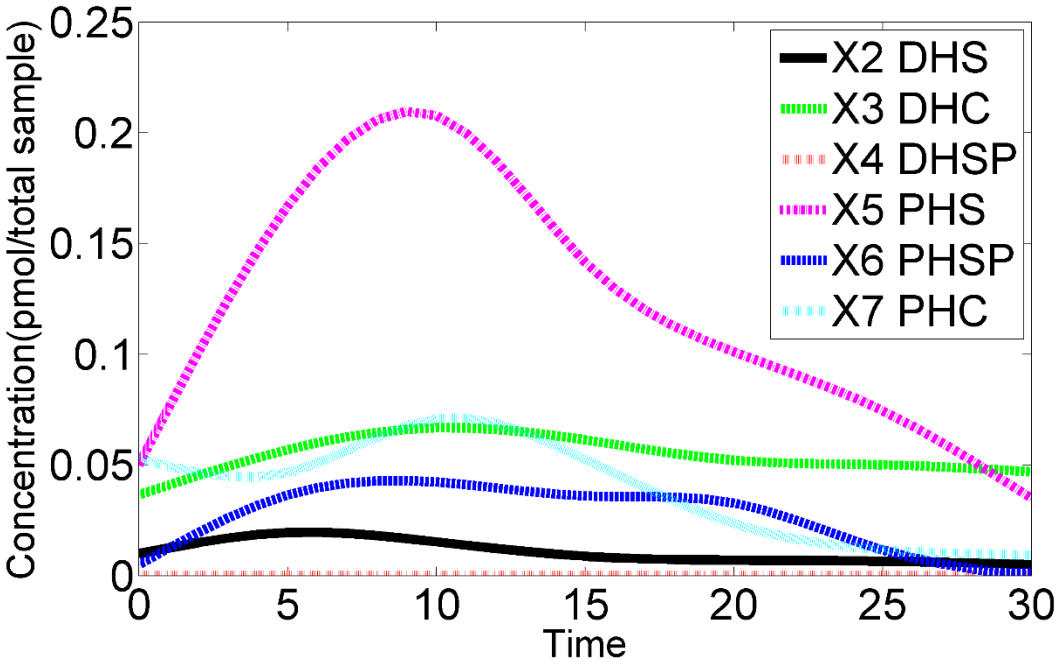


Figure 3.13. Changes in 3KDHS production under optimal and heat stress conditions. *The model allows the computation of 3KDHS accumulation without 3KDHS degradation. The accumulation patterns are distinctly different under optimal (blue) and heat stress (red) conditions.*

3.4.2 Mathematical Model

The biosynthesis, metabolic conversions, and degradation of sphingolipids constitute a complex, highly regulated pathway system (Figure 3.10) that exceeds intuitive capabilities and suggests computational modeling for quantitative systemic analyses. Over the past decade, we have developed a series of such models using a General Mass Action (GMA) formulation within the modeling framework of Biochemical Systems Theory (BST) [9-11,105]. Because these models have been described in detail elsewhere, we can keep their description here to a minimum.

The simple and complex sphingolipids, as well as other pertinent metabolites, are represented in the model as dependent variables, each of which satisfies an ordinary differential equation (ODE). Each ODE contains representations of the processes that produce or degrade this metabolite. According to the tenets of BST, each process is represented as a product of power-law functions, which consists of a rate

constant and of every variable directly affecting this process, raised to an exponent, called a kinetic order. Variable names and equations are presented in Appendix 3.

As an example for how to design a system equation, consider the dependent variable X_2, which represents dihydrosphingosine (DHS). This metabolite is generated from three possible sources. First, KDHS reductase (X_{27}) catalyzes the reduction of 3-keto-dihydrosphingosine (KDHS; X_1). The formulation of this process consists of a rate constant γ_{21} , which is multiplied by X_1 , raised to the kinetic order $f_{2,1,1}$, and by X_{27} , raised to the kinetic order $f_{2,27,2}$. Thus, the reduction process is modeled as $\gamma_{21} X_1^{f_{2,1,1}} X_{27}^{f_{2,27,2}}$. Second, DHS can be produced from dihydrosphingosine-1-phosphate (DHS-P; X_4), a process catalyzed by sphingoid 1-phosphate phosphatase (X_{41}). In analogy to the first process, this step is represented with its own rate constant, as well as the substrate and enzyme, which are both raised to appropriate kinetic orders. Third, dihydroceramide alkaline ceramidase (X_{29}) converts dihydroceramide (DHC; X_3) into DHS, and this process is formulated in an analogous manner. DHS is subject to three possible metabolic fates, namely through the ceramide synthase reaction toward DHC, through the 4-hydroxylase reaction toward phytosphingosine (PHS), and through the sphingoid base kinase reaction toward DHS-P. Taken together, the ODE equation describing the dynamics of DHS contains three influx terms and three efflux terms as shown in Eq. (3.1).

$$\begin{aligned}\frac{dX_2}{dt} &= \gamma_{21} X_1^{f_{2,1,1}} X_{27}^{f_{2,27,2}} + \gamma_{22} X_3^{f_{2,3,3}} X_{29}^{f_{2,29,4}} + \gamma_{23} X_4^{f_{2,4,5}} X_{41}^{f_{2,41,6}} \\ &- \gamma_{24} X_2^{f_{2,2,7}} X_{23}^{f_{2,23,8}} X_{34}^{f_{2,34,9}} - \gamma_{25} X_2^{f_{2,2,10}} X_{28}^{f_{2,28,11}} X_{36}^{f_{2,36,12}} - \gamma_{26} X_2^{f_{2,2,13}} X_{54}^{f_{2,54,14}}\end{aligned} \tag{3.1}$$

All differential equations for dependent variables are formulated in this manner. Values for all parameters were determined from the literature [11,106]. The complete model consists of 25 ordinary differential equations, including those representing the six key sphingolipids of interest here, namely

dihydrosphingosine, dihydroceramide, dihydrosphingosine 1-phosphate, phytosphingosine, phytosphingosine 1-phosphate and phytoceramide. The model furthermore contains 41 independent variables, which represent enzyme activities and metabolites such as ATP, palmitate, acetate and phosphoserine, which were assumed to be constant or considered unaffected by the dynamics of the pathway system. The model was rigorously tested and validated against data not used for model construction [10]. It was also recently combined with a model of the sterol pathway, which has relevance for the composition of membrane rafts [105]. An SBML version of the model can be found as in Appendix 3.

3.4.3 Piecewise Optimization Approach

As stated at the beginning of the *Results* section, it is our task to infer from the measured metabolite time courses which enzymes have to be altered dynamically, and by how much, in order for the model to generate the observed time-dependent metabolic profile? Mathematically, this inverse problem is underdetermined and furthermore complicated by the fact that the pathway is described by a system of nonlinear differential equations, as discussed before. If we were only concerned with a baseline steady state and the move of the system to a new steady state appropriate for heat stress conditions, we could use methods of linear algebra and pseudo-inverses, as we have demonstrated elsewhere [98]. However, here we are interested in the entire trajectories between stimulus (*i.e.*, the beginning of heat stress) and the cell's metabolic adjustments over 30 minutes.

We solved this dynamic inverse problem with an iterative, piecewise optimization approach. Specifically, we estimated optimal enzymatic profiles by minimizing the distance between the smoothed sphingolipid data and the simulation results at each time point, with 1-minute time intervals, from 0 to 30 minutes. At each time point, the optimization engine searched for the best set of enzyme activities, which were modeled as independent variables. To satisfy the specified objective function, we algorithmically minimized the distances between the six observed sphingolipid concentrations and the solutions produced by each trial set of independent variables. We executed this strategy 4144 times, using different random

values for initial settings. We then selected the 2004 best models based on residual errors (SSEs). In order to test the performance of this metric, we also selected models based on the Akaike criterion (AICc), and both criteria produced very similar results. Please see Appendix 3 for a detailed comparison of results using these two criteria. Subsequently, scanning all solutions throughout the 30-minute time period yielded dynamic alteration profiles in all enzymes as well as corresponding metabolite profiles that were consistent with the observed profiles throughout the experimental time period. Further details of this procedure are presented in the Appendix 3.

Each optimization run produced a dynamic enzymatic profile throughout the time period from 0 to 30 minutes. Due to the randomization of initial values and to the fact that the system is underdetermined, the solutions from different runs were different. Thus, instead of searching for a single unique solution, we studied an entire large ensemble of solutions and asked whether the solutions would reveal consistent trends of enzymatic profiles with in the potentially large solution space. Indeed, the overall result of this strategy was a set of surprisingly tight ranges for the key enzymes of sphingolipid biosynthesis

CHAPTER 4

DYNAMICS OF THE HEAT STRESS RESPONSE OF CERAMIDES WITH DIFFERENT FATTY-ACYL CHAIN LENGTHS IN BAKER'S YEAST [3]

4.1 Introduction

Recent experimental studies have demonstrated that not only the main sphingolipid metabolites in yeast differentially respond to heat stress, as it was discussed in Chapter 3, but that even variants of the same key sphingolipids, which differ in the lengths of their lipid backbones and fatty acyl head groups, can play distinct roles in cell signaling [107,108]. These alternate signaling roles of sphingolipids have been receiving increasing attention from academia and the biomedical industry. They also lead to the question of how exactly these different sphingolipid species are synthesized and how cells regulate and control the concentration of each sphingolipid variant.

Again, an experimental investigation of the control strategies used by the cells seems difficult, and we therefore develop computational methods here that allow us to shed light on these strategies. Using the results from our previous analysis (Chapter 3) quasi as boundary conditions, we narrow our focus on the smaller system of ceramides and design a much more detailed model that accounts for different ceramide species. As in earlier studies, we use Biochemical Systems Theory (BST) as the modeling framework and design a combinatorial modeling approach that includes data preprocessing, dynamic flux estimation, and

[3] **Most of the material presented in this chapter was published in the article "Dynamics of the heat stress response of ceramides with different fatty-acyl chain lengths in baker's yeast", *PLOS Computational Biology*, accepted on 6/2/2015.**

a modified multiple shooting optimization method. This custom-tailored method allows us to infer changes in enzymatic activities that lead to observed responses of the various ceramide species following heat stress.

4.1.1 A Brief Review of Pertinent Details of Sphingolipid Metabolism

Ceramides form a class of sphingolipids with a lipid backbone and a fatty acyl head group. Distinct variants of ceramides result from different backbones and fatty acyl CoAs. For example, C16-dihydroceramide (C16-DHC) is a ceramide with a DHS backbone and a C16 fatty acyl CoA (palmitoyl CoA) head group. Recent research has revealed that DHC species with long fatty acyl chain lengths, such as C18-DHC, as opposed to very long chain based DHC species, such as C26-DHC, have distinct signaling roles [108,109]. To understand and explain these subtle differences, it is necessary to characterize the metabolic mechanisms of ceramide biosynthesis and degradation.

Ceramide can be generated via two paths, namely *de novo* biosynthesis and conversion of inositol phosphoceramide (IPC). *De novo* biosynthesis of sphingolipids is initiated by the condensation of serine and palmitoyl CoA, a reaction which is catalyzed by serine palmitoyltransferase (SPT) (Figure 4.1). The product, 3-keto-dihydrosphingosine (3KDHS) is quickly reduced by KDHS reductase to dihydrosphingosine (DHS). DHS is a main source of ceramide backbone compounds. It can be converted into different dihydroceramides (DHC), due to multiple options for fatty acyl CoAs, which can serve as substrates for ceramide synthase [110,111]. The reverse reaction, from DHC to DHS, is catalyzed by dihydroceramidase [112]. DHC and DHS are key branch points in the sphingolipid biosynthesis pathway, because hydroxylases [113]can irreversibly convert these compounds into phytoceramide (PHC) and phytosphingosine (PHS), respectively. PHC and PHS may undergo reversible reactions catalyzed by ceramide synthase (PHS to PHC) and phyotoceramidase (PHC to PHS); the forward reaction (PHS -> PHC) requires one from among several different fatty acyl CoAs as substrate[114].

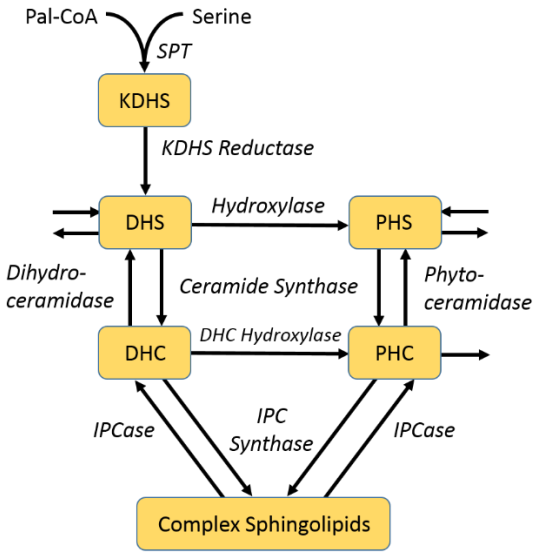


Figure 4.1. Greatly simplified diagram of ceramide metabolism. *Key metabolites are shown in boxes; while enzymes are represented in italics. See Text for abbreviations.*

De novo biosynthesis is not the only option for making sphingolipids available when needed. Equivalent to sphingomyelin in mammalian cells, yeast complex sphingolipids, including inositol phosphorylceramides (IPCs), refer to a class of ceramides with one or two inositol groups attached. They can be formed from DHC and PHC substrates via catalysis by the enzyme IPC synthase. IPC can be irreversibly converted into mannose inositol phosphorylceramide (MIPC), which can furthermore form mannose di-inositol phosphorylceramide (MIP_2C). Importantly, IPC, MIPC and MIP_2C all can serve as sources of DHC and PHC, through catalysis by the enzyme IPCase (Isc1). Thus, utilization of IPC compounds serves as the second path of ceramide production. [115]

The reactions described in the previous paragraphs form a complex metabolic network, in which different enzyme systems exert subtle control over the proper concentration profile of the various sphingolipid species. Focusing on the ceramide species, namely, DHC and PHC, we can describe the biosynthesis and utilization of DHC and PHC as a cooperation among five enzyme systems consisting of ceramide synthases, ceramidases, IPC synthases, IPCases, and DHC hydroxylases. Ceramide synthases and

IPCases mediate the formation of DHC and PHC, while ceramidases and IPC synthases use DHC and PHC as substrates. The hydroxylation reaction, catalyzed by DHC hydroxylases, converts DHC into PHC.

Our previous results [116] in Chapter 3 indicated that the enzymatic activities within the sphingolipid pathway can be grouped by their responses to heat stress, and thereby suggested that the various sphingolipid species are components of highly coordinated enzyme modules. An important detail we discovered is that the *de novo* biosynthesis of sphingolipids is activated immediately upon heat stress, but sustained only for a few minutes; this phase is followed by the activation of IPC utilization, which leads to the accumulation of sphingolipids that are retrieved from complex sphingolipids in the membranes. In this earlier, larger model, a single variable represented the concentration of all dihydroceramides (DHCs), irrespective of their fatty-acyl chain lengths, and a second variable represented the collective concentration of all phyotoceramides (PHCs). The DHC and PHC levels are controlled by five enzyme classes, namely ceramide synthase, dihydroceramidase and phyotoceramidase, IPC synthase, IPCase (Isc1), and dihydroceramide hydroxylase. A simplified representation is shown in Figure 4.1; for details of the sphingolipid pathway at large, see [10,11,116]. Ceramide synthase and DHC hydroxylase showed a similar pattern of activity changes, with a strong initial uptake of material followed by a weak profile toward the end. By contrast, dihydro- and phytoceramidase showed entirely different patterns. While dihydroceramidase exhibited very strong activation in the first 10 minutes, which then receded, phytoceramidase exhibited less pronounced peaks of activation. IPC synthase and IPCase had similar activities, again beginning with a sharp spike, which decreased much slower than for ceramide synthase.

While these inferred activity patterns shed some light on how cells regulate their sphingolipid or ceramide contents, they are too coarse to account for changes in the regulatory patterns that govern distinct ceramide species. The present study fills this gap.

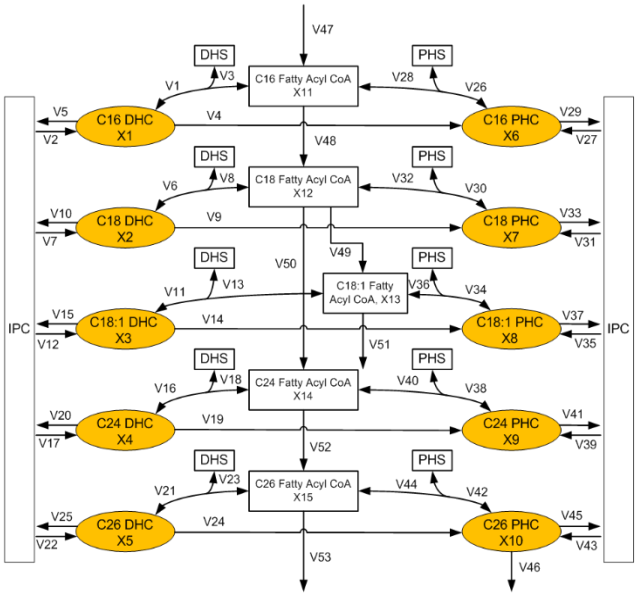
4.2 Results

The enzyme activities inferred computationally in our previous study [116] and Chapter 3 demonstrate how yeast resets its ceramide concentrations under heat stress. The alterations in activities are globally coordinated and grouped, and show distinct regulatory patterns across different enzymes and even ceramides with different fatty acid chains. In the following, we first briefly summarize pertinent findings from this previous research, which focused on heat-stress induced enzyme responses within the broader sphingolipid pathway system, and then zero in on details of the responses among enzymes directly associated with ceramides of different fatty acyl chain lengths.

4.2.1 Alterations in Enzymatic Activities within the Ceramide Pathway System

As mentioned above, the dynamics of ceramide species is governed by five groups of enzymes, which are shown in Figure 4.1 along with SPT and KDHS, which catalyze prior reactions steps. Our overall strategy for zeroing in on this sub-pathway is similar to our previous approach, although the detailed account here requires several methodological modifications, which are detailed in the section *Materials and Methods*. As an illustrative example, consider dihydroceramidase. In our previous study, this enzyme catalyzed a single reaction step from DHC to DHS. Now, dihydroceramidase is involved in five reactions, from C14 and C16 DHC to DHS, C18 DHC to DHS, and so on, to the conversion of C26 and C26:1 DHC into DHS. The relationships between fluxes in the previous and in the present model are shown in the right panel of Figure 4.2. References regarding the enzymatic reactions can be found in the section *Pertinent Details of Sphingolipid Metabolism.* The refined ceramide model discussed here shares many reactions with the previous sphingolipid model, but is expanded toward details regarding the availability of specific substrates.

The reactions of this detailed subsystem are presented in the left panel of Figure 4.2. The left side of the diagram shows the different DHC species, categorized by chain lengths, and the right side the corresponding PHC species. The processes in the center are associated with elongation and desaturation. We designed a dynamic pathway model of this system with methods of Biochemical Systems Theory [117-120], which have been widely documented in the literature.



Flux from [13]	Constrained fluxes
$V_{DHS->DHC}$	V1 + V6 + V11 + V16 +V21
$V_{PHS->PHC}$	V26 + V30 + V34 + V38 + V42
$V_{DHC->DHS}$	V3 + V8 + V13 + V18 + V23
$V_{PHC->PHS}$	V28 + V32 + V36 + V40 + V44
$V_{DHC->IPC}$	V5 + V10 + V15 + V20 + V25
$V_{PHC->IPC}$	V29 + V33 + V37 + V41 + V45
$V_{IPC->DHC}$	V2 + V7 + V12 + V17 + V22
$V_{IPC->PHC}$	V27 + V31 + V35 + V39 + V43
$V_{DHC->PHC}$	V4 + V9 + V14 + V19 + V24
$V_{PHC\_Remodel}$	V46
$V_{C16\_Fatty\_Acyl\_CoA}$	V47

Figure 4.2. Structure of the proposed model of ceramide dynamics. *Left panel: Detailed pathway of ceramide biosynthesis and utilization. The pathway consists of three major subsystems: synthesis and utilization of DHC (left column), synthesis and utilization of PHC (right column), and fatty acid elongation and desaturation (center column). V1, V6, V11, V16, V21, V26, V30, V34, V38 and V42 represent fluxes catalyzed by ceramide synthase. V3, V8, V13, V18, V23, and V28, V32, V36, V40 and V44 represent fluxes catalyzed by dihydroceramidase and phytoceramidase, respectively. Reactions exiting the system to the left or right, namely V5, V10, V15, V20, V25, V29, V33, V37, V41 and V45, represent fluxes catalyzed by IPC synthase, while reactions entering the system from the left or right, V2, V7, V12, V17, V22, V27, V31, V35, V39 and V43, represent fluxes catalyzed by IPCase (Isc1). V4, V9, V14, V19, V24 represent DHC hydroxylase. The vertically shown reactions represent reactions catalyzed by other enzymes, such as remodelase (V46), fatty acid elongases (V47, V48, V50, V51, V52, and V53), and desaturase (V49). The boxes marked "IPC" summarily account for the complex sphingolipids IPC, MIPC, and $M(IP)_2C$ and their interconversions. This simplification seems reasonable as all three, IPC, MIPC and $M(IP)_2C$, can serve as sources for the production of DHC and PHC. Right panel: The dynamic flux estimation is partly based on some fluxes whose magnitudes we took from our previous model [116]. The table summarizes these fluxes and indicates on the right how they constrain fluxes in our present model.*

Using the computational approach described in the *Methods* Section, we set out to infer detailed changes in enzymatic activities from time series measurements that had been generated specifically for the purpose of better understanding the heat stress response [107]. The data demonstrate that during the 30-minute heat stress response, different ceramide species respond in dramatically different ways, with some accumulating, others being reduced over time, and yet others maintaining a relatively stable concentration. We used these data to parameterize the model, while allowing the enzyme activity levels to change every five minutes. Custom-tailored methods, reminiscent of multiple-shooting optimization, were used for this purpose.

As a first diagnostics of the parameterization, we tested the goodness of fit of the model with time varying enzyme activities for the smoothed, interpolated ceramide data; the results are given in Figure 4.3, while the raw data are shown in Figure 4.10. Specifically, every three minutes the enzyme activities in the model were allowed to change, and this procedure yielded a reasonably good fit. The results are not entirely smooth due to the abrupt changes in enzyme activities between modeling windows. Even shorter windows were not considered necessary, considering the magnitude of noise in the data and our intent merely to characterize trends in enzyme activities, whereas longer windows (5 or 10 minutes) did not lead to satisfactory fits. As a consequence, the comparatively minor changes in sphingolipids do not seem to affect their levels much. Therefore, much smaller fluctuations in these major metabolites during stress environment should be expected in order for cells to minimize side effects in other pathways.

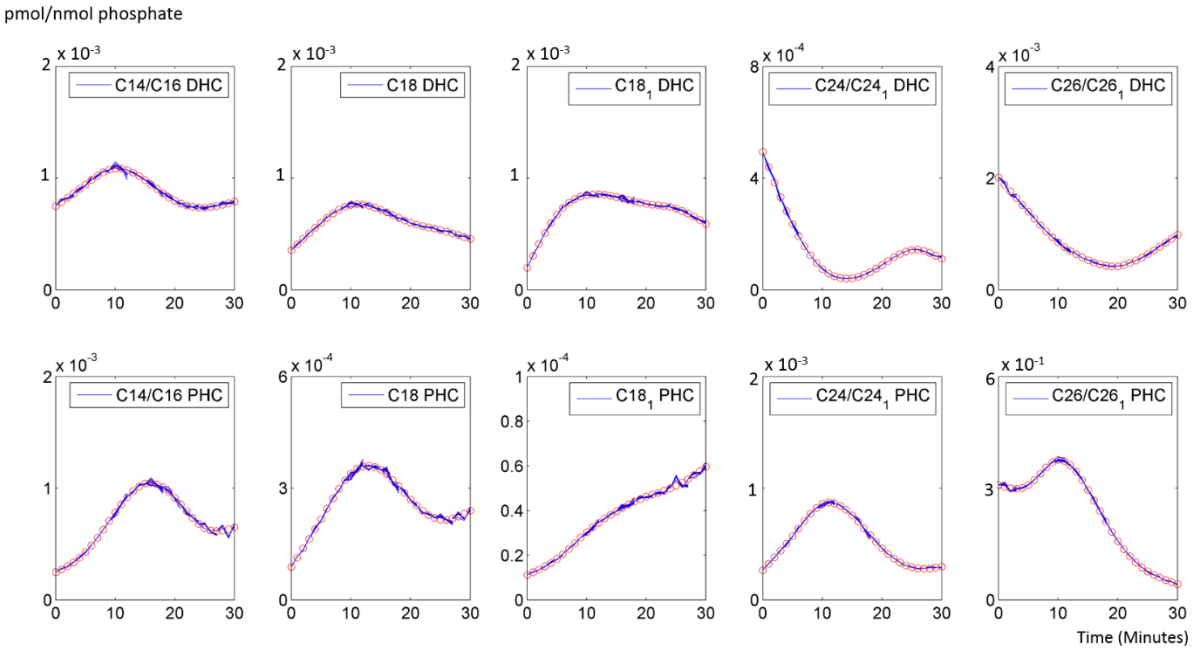


Figure 4.3. Smoothed and interpolated concentrations of ceramide species. *The concentrations were reconstructed with the model, using inferred enzymatic activities; for raw data, see Figure 4.10. The X-axis represents the 30 minutes of heat stress, while the Y-axis represents the concentrations of pertinent ceramides. In each graph, red circles represent the smoothed and interpolated trends in experimental data, and the segmented blue lines represent the corresponding ceramide concentrations computed with the dynamic model using optimized enzymatic activities. The discontinuities are due to the optimization method, which was gleaned from multiple shooting methods. See Methods Section and Text for further details.*

The model with these optimized parameter values fits the smoothed, interpolated metabolite data quite well, and large-scale simulations (see *4.4 Materials and Methods*) confirmed that the model settings are robust to moderate perturbations. Our computational inferences of enzyme activities, derived from these metabolic measurements, reveal the dynamic patterns that generate the various ceramide species with different carbon chain lengths.

One should note that the shapes, rather than the magnitudes, of the inferred fold changes are more indicative of actual changes. The reason is that, due to insufficient experimental information, our method only infers the product of a rate constant and the corresponding enzyme activity, but cannot assign

numerical values to the two factors separately. This situation is analogous to determining V_{max} in a Michaelis-Menten reaction, rather than k_{cat} and the total enzyme concentration.

1. Ceramide synthase

Ceramide synthases are involved in ten model reactions that are associated with the synthesis of C14/C16, C18, C18:1, C24/C24:1 and C26/C26:1 DHC and PHC. Their dynamically changing activities are shown in Figure 4.4. All ceramide synthases are activated immediately when heat stress commences, which is presumably due to the Arrhenius effect, but this spike in activity only lasts for two or three minutes. This observation is directly consistent with our previously published results[116], as well as other experimental findings [83,92]. Furthermore, the dynamic pattern of ceramide synthases show a trend based on different fatty acyl CoAs that are used as substrates. While the activities of ceramide synthases that use C14/C16 more or less return to the baseline after a few minutes, it appears that the activities exhibit a modest second peak between 15 and 20 minutes, which however, could be an artifact. In strong contrast, the activity for C26 undershoots after the initial spike and only recovers toward the end of the heat stress experiment. It is known that there is only one ceramide synthase associated with three genes (LAC1, LAG1 and LIP1) in yeast, which shows remarkably different substrate preferences [111,121,122].

The activities of ceramide synthase toward the synthesis of PHCs show different trends (shaded green in Figure 4.4). For substrates between C14 and 24, the activities gradually increase for about six to eight minutes and subsequently return slowly to their original values, which they reach after about 25 minutes. Interestingly, the activity for the very long chain substrates C26 and C26:1 is very similar for the synthesis of C26 DHC: it spikes to about 4- or 5-fold and immediately drops back to the baseline.

Note: The ceramide synthase activities we have estimated are actually products of ceramide synthase activities multiply by concentrations of corresponding fatty acyl CoAs. Since encountering difficulties finding measurements of fatty acyl CoAs in yeast under 30 minutes (short period) heat stress, we assumed concentrations of fatty acyl CoAs are close to their steady state in this short 30 minutes period so that these products directly refer to the ceramide synthase activities.

If fatty acyl CoAs concentration show a step change during heat stress, then the baseline activities of ceramide synthase would be changed, but the dynamic trends should be the same. If the fatty acyl CoAs exhibit dynamical changes in 30 minutes heat stress, then we would have to compensate the trend of ceramide synthase with the trend in fatty acyl CoA concentrations, which for now exceeds the capability of the current modeling approach.

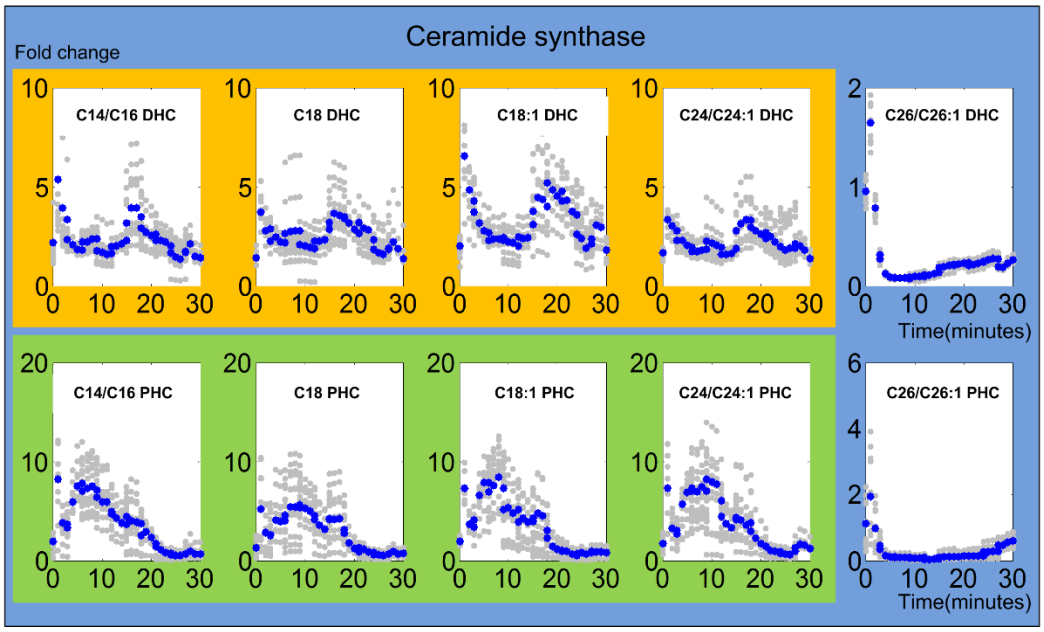


Figure 4.4. Trends in ceramide synthase activities. *The upper panel shows ceramide synthase activities toward DHC, while the lower panel shows ceramide synthase activities toward PHC. The X-axis represents the 30 minutes of heat stress and the Y-axis represents fold changes in activities. In each plot, the blue and gray dots represent averaged and individual enzyme activities, respectively. The yellow and green shading marks similar enzyme activity patterns involved in DHC (top panel) and PHC (bottom panel) synthesis, respectively.*

2. Ceramidase

Ceramidase catalyzes the reaction from ceramide to its sphingosine backbone. In our model, dihydroceramidase and phyotoceramidase are involved in five reactions that convert dihydroceramide species to dihydrosphingosines and phytoceramide species to phytosphingosines, respectively. Like the ceramide synthases, dihydroceramidase and phytoceramidase are activated immediately upon heat stress,

as it is shown in Figure 4.5. This result indicates the important fact that heat stress not only triggers the production of sphingolipids but also their degradation, and that these two events occur almost simultaneously.

The enzyme activities of ceramidases can be grouped according to the DHC or PHC substrates they use. For dihydroceramidase, reactions involving C14/C16, C18, and C18:1 DHC as substrates show similar enzymatic activity patterns, which immediately rise, but by about time 12 stop being active (shaded yellow in Figure 4.5). The reactions involving C24/C24:1 and C26/C26:1 DHC initially seem to change more slowly and peak slightly later than the enzymes for shorter-length substrates. For phytoceramidase, similar trends are observed for all chain lengths except C24 (shaded green in Figure 4.5), with peaks at about 5 minutes and thus a little later than for DHC. The pattern with regard to C24 PHC is not very distinctive, but seems to peak late into the heat stress. Note again that the absolute magnitudes of fold changes are to be considered with caution.

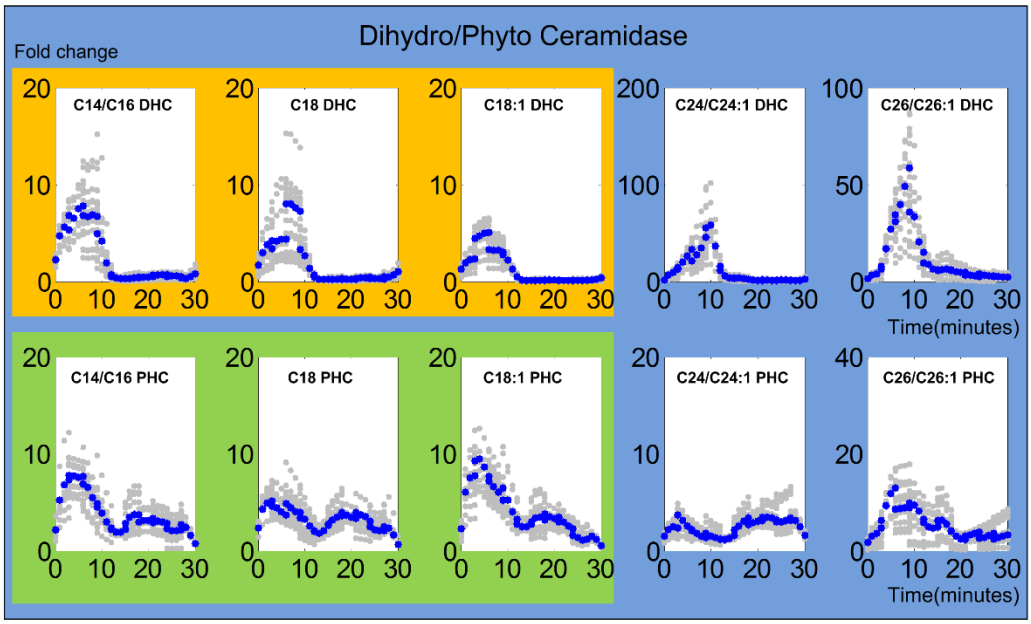


Figure 4.5. Trends in ceramidase activities. *The upper and lower panels show activity trends of dihydroceramidase and phytoceramidase, respectively. The X-axis represents the 30 minutes of the heat stress experiment, while the Y-axis represents the fold changes in activities. In each plot, the blue and gray dots represent averaged and individual enzyme activities, respectively. The yellow and green shading marks similar enzyme activity patterns involved in DHC (top panel) and PHC (bottom panel) conversion, respectively.*

3. IPC synthase

IPC synthase controls a different pathway of ceramide utilization, which results in the synthesis of complex sphingolipids that are members of the IPC family. For long chain (C14 – C18) DHC, an initial activation of IPC synthases is again followed by low activity profiles during the last 10 to 15 minutes (Figure 4.6). The same is observed for C14 – C24 PHC. By contrast, activities associated with very long C24 – C26 DHC and C26 PHC peak between 10 and 15 minutes. The activities for DHC return to low values within the experimental time period, while the activity with respect to C26 PHC is still increased after 30 minutes.

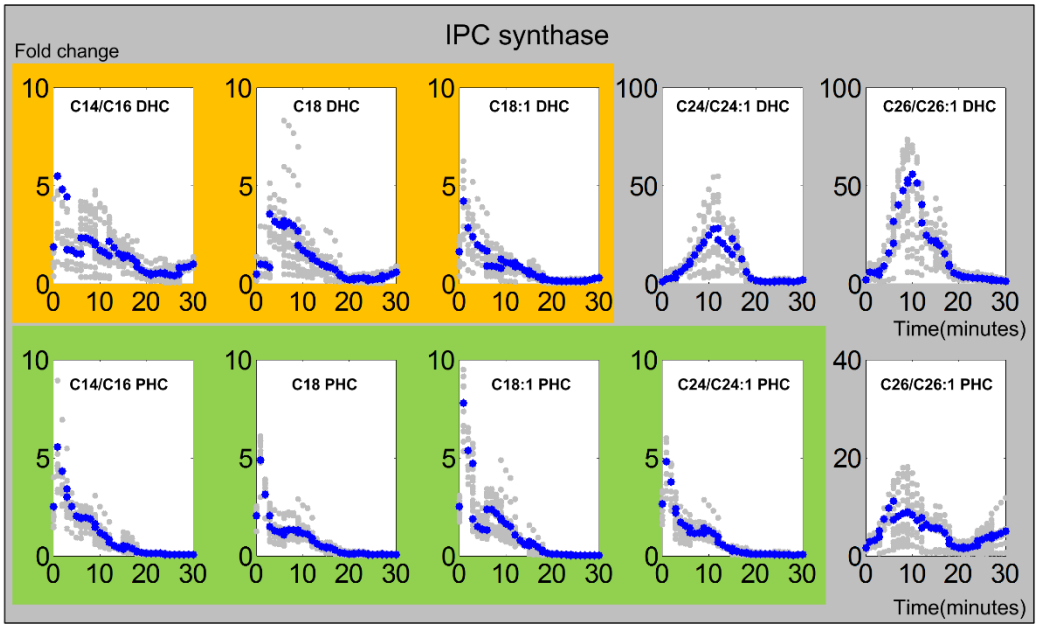


Figure 4.6. Trends in IPC synthase activities. *The upper and lower panels show IPC synthase activities from DHC to IPC and from PHC to IPC, respectively. The X-axis represents the 30 minutes of the heat stress experiment, and the Y-axis represents the fold change in activities. In each plot, the blue and gray dots represent averaged and individual enzyme activities, respectively. The yellow and green shading marks similar enzyme activity patterns involved in DHC (top panel) and PHC (bottom panel), respectively.*

4. ISC1/IPCase

ISC1p breaks IPC into ceramides and thus constitutes another important process affecting the production of DHC and PHC in yeast. Figure 4.7 presents the inferred Isc1 activities. In all cases except for C26, the enzyme activities gradually increase upon heat stress and peak at about 10 minutes, before they return to the baseline. By contrast, the production of the very long chain (C26/C26:1) DHC and PHC are very low throughout the experiment.

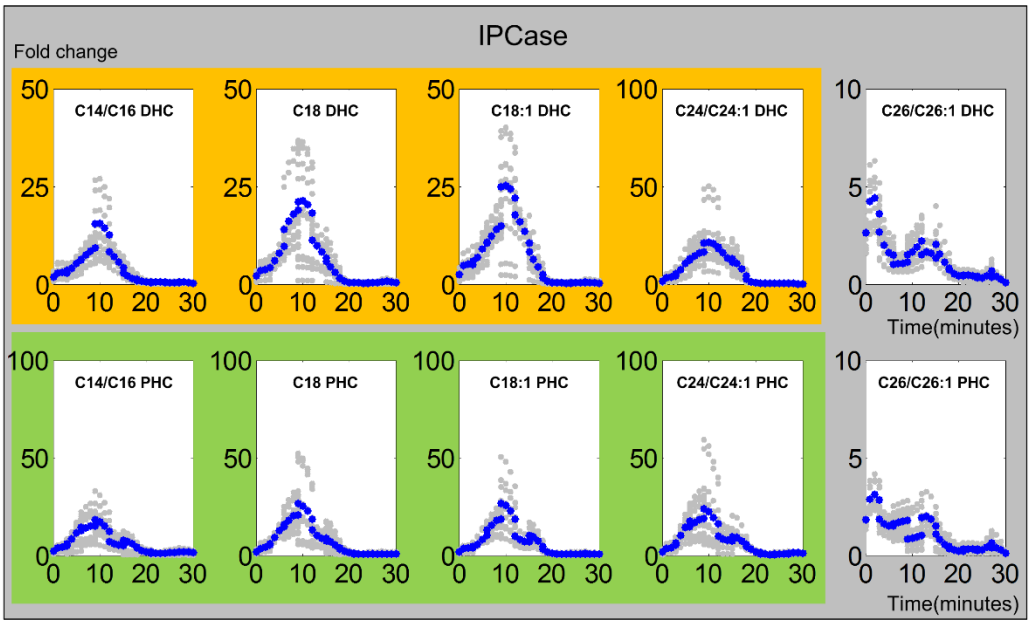


Figure 4.7. Trends in IPCase activities. *The upper and lower panels show IPCase activities from IPC to DHC and from IPC to PHC respectively. The X-axis represents the 30 minutes of the heat stress experiment, and the Y-axis represents the fold change in activities. In each plot, the blue and gray dots represent averaged and individual enzyme activities, respectively. The yellow and green shading marks similar enzyme activity patterns involved in DHC (top panel) and PHC (bottom panel) side, respectively.*

5. Dihydroceramide hydroxylase

DHC hydroxylase catalyzes the reaction from DHC to PHC. In our model, five reactions, associated with different substrate chain lengths, are catalyzed by this enzyme system. The inferred enzymatic activities are shown in Figure 4.8. Except for the initial burst that was also observed for other enzymes, reactions with long chain fatty acids show similar and rather low enzymatic activities. By contrast, reactions

involving very long chain fatty acid based DHC exhibit rather different dynamics. For C24/C24:1 DHC, the DHC hydroxylase is activated much later, and more strongly, than most other reaction steps, while for the C26/C26:1 substrate, the activity exhibits an immediate peak, as well as a second, low peak after about 15 to 20 minutes.

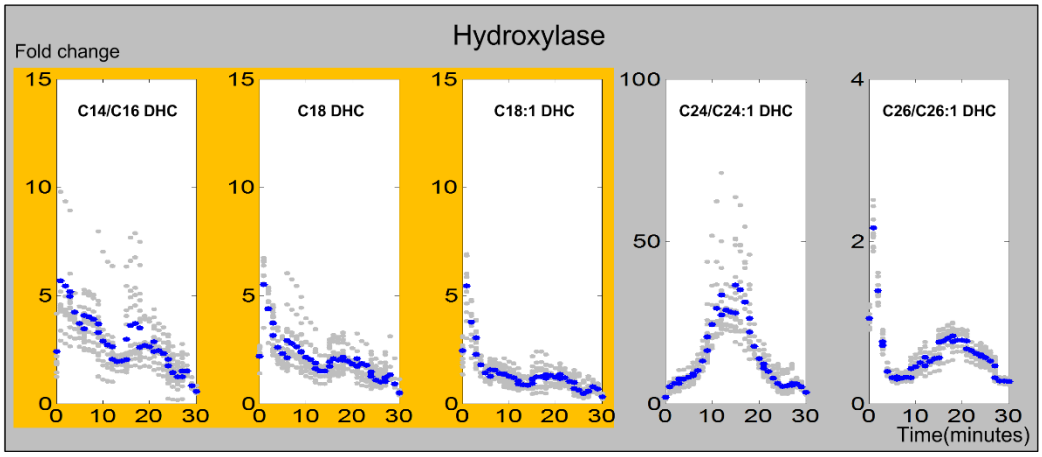


Figure 4.8. Trends in DHC hydroxylase activities. *The X-axis represents the 30 minutes of the heat stress experiment and Y-axis represents fold-changes in activities. In each plot, the blue and gray dots represent averaged and individual enzyme activities, respectively. The yellow shading marks similar DHC hydroxylase activities from DHC toward PHC.*

6. Other Enzymes

Fatty acid elongation is crucial for the backbone of ceramide biosynthesis, because it supplies the necessary fatty acyl CoAs of different lengths, from C14 to C26. As no time series measurements are available for the fatty acyl CoAs under heat stress, we assumed that these fatty acyl CoAs only show limited activation or degradation around their steady state. The corresponding enzyme activities in this elongation processes and also other accessory enzyme activities (remodelase, desaturase) are presented in Figure 4.9.

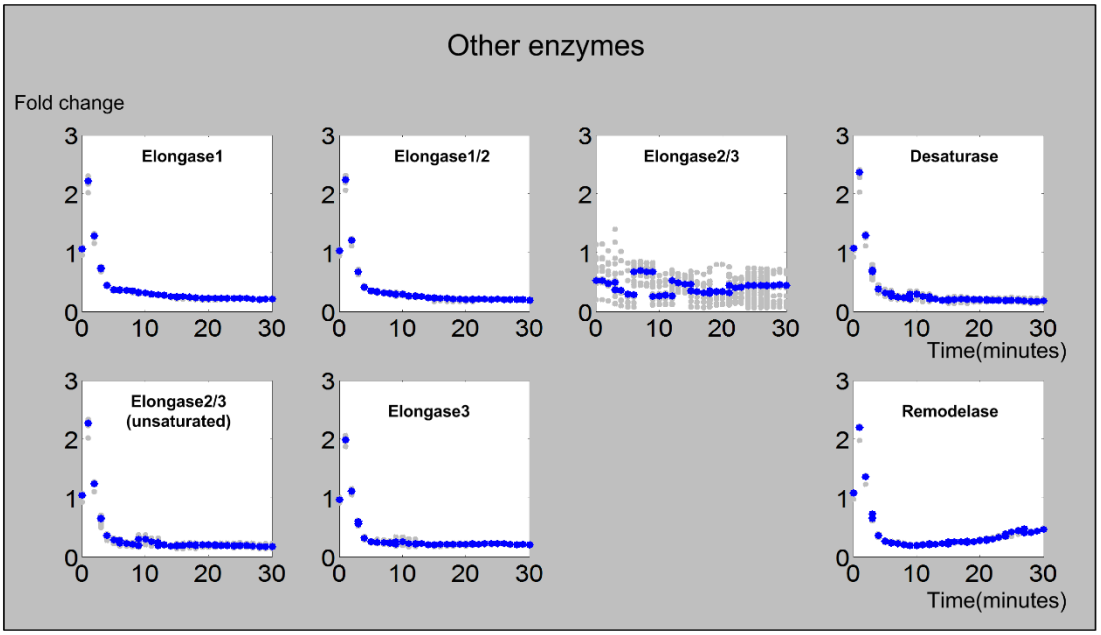


Figure 4.9. Activities of fatty acid elongases, desaturase, and remodelase. *The X-axis represents the 30 minutes of the heat stress experiment and the Y-axis represents fold-changes in activities. In each plot, the blue and gray dots represent averaged and individual enzyme activities, respectively.*

4.3 Discussion

Recent research has indicated that yeast ceramides of different types and with different fatty-acyl chain lengths can exert distinct signaling functions [108]. It is therefore important to investigate how yeast cells manage to make the different ceramide variants available when needed. Our results shed light on this question. The computational inferences of dynamic enzymatic activities in ceramide biosynthesis are in agreement with previous insights into the control of sphingolipids under heat stress, but have a much higher resolution. The earlier results had suggested that essentially all enzymes in the pathway are involved in the adjustment of sphingolipid metabolism under heat stress, rather than just a few key enzymes, as one might have expected. The same observation holds again for the results obtained here, where we zero in on the much smaller but very important ceramide subsystem. Our simulations suggest that, for each ceramide

species, the accumulation or degradation is governed by all associated enzymes. The results also render it evident that the ceramide heat stress response is not a haphazard panic reaction by the cell, but a highly cooperative adaptation whose implementation is shared among all associated enzymes.

The experimental metabolic time series data and our customized optimization method allowed us to infer trends in all five major enzymes: ceramide synthase, ceramidase, IPC synthase, IPCase, and DHC hydroxylase. Interestingly, these enzyme systems exhibit distinct dynamic activity patterns that strongly depend on the carbon chain lengths of their fatty acids.

The initial activation of ceramide synthase confirms experimental and computational results indicating that heat stress induces the *de novo* biosynthesis of ceramides. In particular, our earlier results had shown that the production of dihydrosphingosine (DHS), the precursor of DHC, is immediately and strongly triggered by heat stress. The results here indicate that some of this newly synthesized DHS is directly channeled into long chain DHC. After just a few minutes, the activity of ceramide synthase for C14-C24 returns to the former baseline under optimal temperature conditions. Especially for C18:1 and C24, a second peak is detected after about 20 minutes, which coincides with the time when long chain DHC ceramidases cease to be active. The synthesis of the corresponding PHC increases more slowly and peaks about six to eight minutes into the heat stress. This short delay between DHC and PHC synthesis may be explained with the fact that phytosphingosine, the substrate of PHC synthase, must first be produced from DHS. In contrast to these patterns, the activities for very long chain substrates (C26 DHC and PHC) only flare up very briefly, but strongly, and then remain rather low. Although this activity trend is short-lived, a comparison with our earlier results suggests that this pattern actually dominates the overall trend in ceramide synthase. Indeed, this suggestion aligns well with the fact that C26 PHC is by far the most prevalent ceramide variant under normal conditions (Figure 4.3).

By catalyzing the utilization of complex sphingolipids, IPCase is the second source for ceramides. In contrast to *de novo* biosynthesis, these processes initially increase much more slowly and exhibit a strong and long lasting activity peak between about 5 and 20 minutes (Figure 4.7). Subsequently, their activities essentially cease. Interestingly, the trends are very similar for DHC and PHC and for chain lengths up to

24, which may suggest that the same enzyme could catalyze all reactions for both DHC and PHC of chain lengths up to 24. The dynamics for C26 DHC and PHC is distinctly different.

The DHC hydroxylase reaction facilitates an internal redistribution between DHCs and PHCs. It is activated instantly for long chain DHCs, with a subsequent decrease in activity, whereas C24 activity occurs mostly between 10 and 20 minutes (Figure 4.8). The activities with respect to C26/C26:1 DHC show a mixed pattern with a very small magnitude.

The utilization of ceramides follows two routes, namely toward the sphingosine backbone via ceramidase, and toward IPC via IPC synthase. The long chain DHC ceramidases arguably exhibit the most striking pattern (Figure 4.5). Their activities rise immediately with heat stress and are sustained for about ten minutes, after which the activities drop quickly and cease altogether. As these activities are much higher than for the corresponding ceramide synthases, it appears that the cells are preferably channeling long chain material to DHS. The corresponding activities for PHC are not as clear-cut. They also increase, but not as quickly or strongly, and return to their baseline over the entire 30-minute time period. This difference between DHC and PHC substrates may again be explained with the fact that phytosphingosine and PHC must first be synthesized from dihydrosphingosine and DHC, respectively. Intriguingly, the activities for very long DHC substrates and for C26 PHC increase more gradually and peak between 5 and 10 minutes. As C26 PHC is the most prevalent ceramide species, this pattern dominates the overall trend in ceramidase.

Finally, IPC synthase incorporates ceramides into complex sphingolipids. Here, the activities for long chain DHC and PHC substrates rise instantly, and the activities essentially cease after about 20 minutes (Figure 4.5). By contrast, the highest use for very long chain DHC and PHC occurs at about 10 minutes of heat stress. As in other cases, these trends are in line with the overall trends we observed in our previous analysis.

Taken together, it appears that the immediate response to heat stress is the *de novo* synthesis of long chain DHC, its conversion into the corresponding PHC, and the return of some material to DHS. Also, some long chain DHCs and PHCs are incorporated into complex sphingolipids. Between 5 and 10 minutes of heat stress, long and very long chain PHCs are generated. Around 10 minutes into heat stress, complex

sphingolipids are used to generate DHC and PHC of all lengths. Also at this time, C24/C24:1 DHC is converted into PHC. Most activities are back to normal at the end of the 30-minute period, which is consistent with our earlier findings [116].

The molecular and cell-physiological reasons for the differences in activity patterns towards substrates of different lengths are unknown. The most straightforward hypotheses might be that the differences are due to:

(1) the existence of specific enzymes or isozymes for different substrates;

(2) different affinities of the same enzymes to substrates with different N-acyl chain lengths;

(3) compartmentalization of substrates and/or enzymes, which would allow the same enzyme to be regulated differently in its action on distinct substrates.

In mammalian cells, at least five ceramide synthases were identified [123], and they perform different but overlapping functions with respect to different fatty acyl CoAs. In yeast, LIP1, LAC1 and LAG1 are known as genes coding for subunits of ceramide synthase, but otherwise not much is known about the reactions associated with different ceramide species. Our computational inferences indicate that reactions using substrates with different fatty acyl chain lengths are grouped, often into long chain and very long chain classes, which could suggest that ceramide associated enzymes are regulated in a "fatty acyl chain length specific" manner. However, there is so far no evidence that different species of the enzymes in question, which seems to suggest that explanation (3) offers the most likely hypothesis. Further experimental research will be needed to support or refute this conclusion.

Beyond the particular application to ceramide metabolism, the computational modeling and inference methods in this work demonstrate how metabolite profiles, obtained as time series data, may be used to decipher *in vivo* strategies with which cells organize their responses to environmental stimuli.

4.4 Materials and Methods

4.4.1 Experimental Data

Ceramide time series data for the present study were obtained *de novo*, as described in [94]. *S. cerevisiae* cells were cultured overnight at an optimal temperature of 30°. In duplicate experiments, the cells were moved to a water bath that was kept 39° C, which causes heat stress in yeast, and sampled every 5 minutes between 0 and 30 minutes. The samples were analyzed with High Performance Liquid Chromatography and Mass Spectrometry (HPLC-MS) to yield heat stress time series data of the following dihydro- and phyto-ceramide species: C14, C16, C18, C18:1, C20, C22, C24, C26; here the numbers refer to fatty acyl chain lengths and ":1" refers to an unsaturated fatty acid with one double bond. Considering that some datasets were missing or measurements fell below the detection limit, we used C14/C16, C18, C18:1, C24/C24:1 and C26/C26:1 DHC/PHC to construct our mathematical model.

As in our previous study [116], we found it beneficial to interpolate the data in a smooth fashion. It seems reasonable to assume that the heat stress response is a continuous phenomenon, and that a minimally biased spline technique would reflect the true dynamics in acceptable approximation. The raw and smoothed, interpolated data are exhibited in Figure 4.10. One should note that the Y-axes in Figure 4.10 are quite different, which indicates a large variation in prevalence of the various ceramide species.

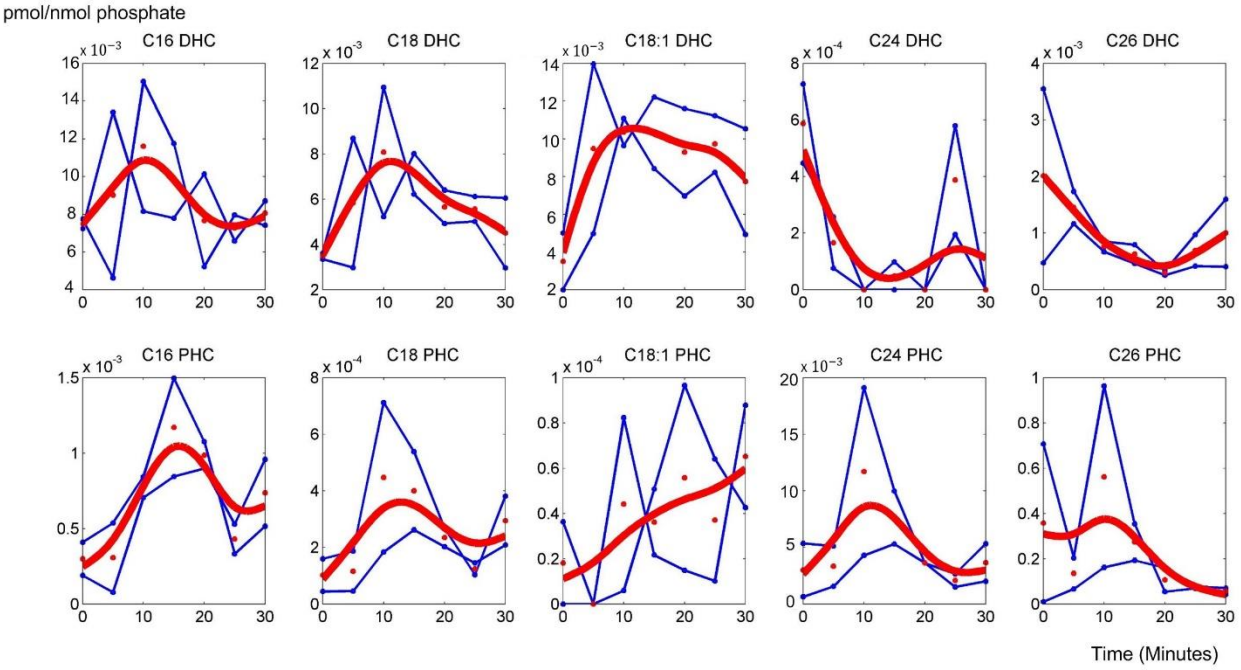


Figure 4.10. Raw duplicate time series concentrations of ceramide species (connected blue symbols) and their means (red symbols). The red curves show the smoothing spline interpolations of each dataset.

4.4.2 Previous Work on Enzyme Activity Inferences

Yeast responds to heat stress within minutes. Among the different aspects to this response, the concentration profile of sphingolipids starts to change in less than two minutes. Due to the complexity of the pathway, these changes in synthesis and degradation, which result from activity changes in enzymes, can hardly be inferred with intuition alone. To shed light on the response strategy, we recently proposed a customized computational approach to infer the dynamic changes in enzymatic activities from sphingolipid time series data [116]. These data consisted of time course measurements that were taken under heat stress conditions every 5 minutes until the end of a 30-minute interval and contained concentration measurements of dihydrosphingosine (DHS), dihydrosphingosine 1-phosphate (DHS1P), phytosphingosine (PHS), phytosphingosine 1-phosphate (PHS1P), dihydroceramide (DHC) and phytoceramide (PHC).

In order to infer the heat-induced changes in enzymatic activities, we used a Generalized Mass Action (GMA) model, in which every process was represented as a product of a rate constant and of all variables that directly affected the process, raised to a power [117]. The specific model was adopted from our earlier work [10,11] and included 31 metabolites as dependent variables and 64 enzymes or cofactors as independent variables. With this model as base structure, we developed a piecewise optimization approach.

First, the time series measurements were interpolated by smoothing splines and then re-sampled to produce time series values for every minute during the experimental time period. These values were entered into the GMA model. For each time interval, we formulated and solved the optimization problem of finding a set of enzymatic activities to establish the lipid profiles in that specific time frame. Given 31 time points, we thus found 30 corresponding sets of enzymatic activities that generated the observed dynamic metabolic profiles. An additional randomization scheme allowed us to infer solution spaces and confidence bands rather than point estimates. Several validation studies confirmed the results. The trajectories of the computed enzymatic activities revealed interesting regulatory mechanisms of sphingolipid metabolism, as described in the introduction.

4.4.3 Methodological Alterations for Explaining the Dynamics of Ceramide Variants

The basic concepts of the modeling method were taken from our previous study [116]. As before, we smoothed the heat stress time series data of the six DHC and PHC species with different chain lengths, considered one-minute time intervals, and developed a customized, piecewise optimization approach that allowed us to infer changes in enzyme activities in a step-by-step manner. Also as in earlier studies, the system was represented as a GMA model. The pathway system under investigation is depicted in Figure 2. It consists of three major parts: synthesis and utilization of DHC, synthesis and utilization of PHC, and fatty acid elongation.

In comparison to the previous inference of all sphingolipid enzyme activities, the subsystem we consider here is relatively small. However, by accounting for the different chain-length variants, the system

is much more detailed and leads to surprising complexity. In particular, the fatty acid elongation process becomes critical, whereas it was modeled only coarsely in the earlier analysis. It is our goal here to tease out the details of fatty acid elongation and the synthesis and degradation of variant ceramide species with different fatty acyl groups.

All responses outside the subsystem addressed here are expected to be the same as in the larger system, within normal biological variability, which allows us to use the previous model as a large set of dynamic boundary constraints that govern the sphingolipid system at large. For instance, fluxes entering the ceramide subsystem are directly taken from the large sphingolipid model. Similarly, the heat stress concentration of palmitoyl fatty acyl CoA can be directly imported from the sphingolipid model. The table in Figure 4.2 summarizes constraints on fluxes in the present model, imposed by the prior model.

The resulting ceramide subsystem has 53 fluxes and 15 dependent variables (Figure 4.2). X_1 to X_5, X_6 to X_{10} , and X_{11} to X_{15} represent the corresponding species of C16-DHC to C26-DHC, C16-PHC to C26-PHC, and C16-fatty acyl CoA to C26-fatty acyl CoA, respectively.

Details of the two core components of our approach, namely the estimation of dynamic fluxes and of enzyme activities, are shown in the flowchart of Figure 4.11. The first task, as shown in the upper panel of Figure 4.11, consists of checking the mass balances within the system and to construct the stoichiometric matrix that describes the production and degradation rates of the dependent metabolites. Because the system has considerably more fluxes than metabolites, we are faced with a highly underdetermined system. We are dealing with this situation by solving the system in 30 pieces, starting from the initial steady state to time 1, from time 1 to time 2, all the way to the end of the heat stress experiment (30^{th} minute). The following describes in more detail a customized optimization strategy with which we determine the flux distribution in each time point.

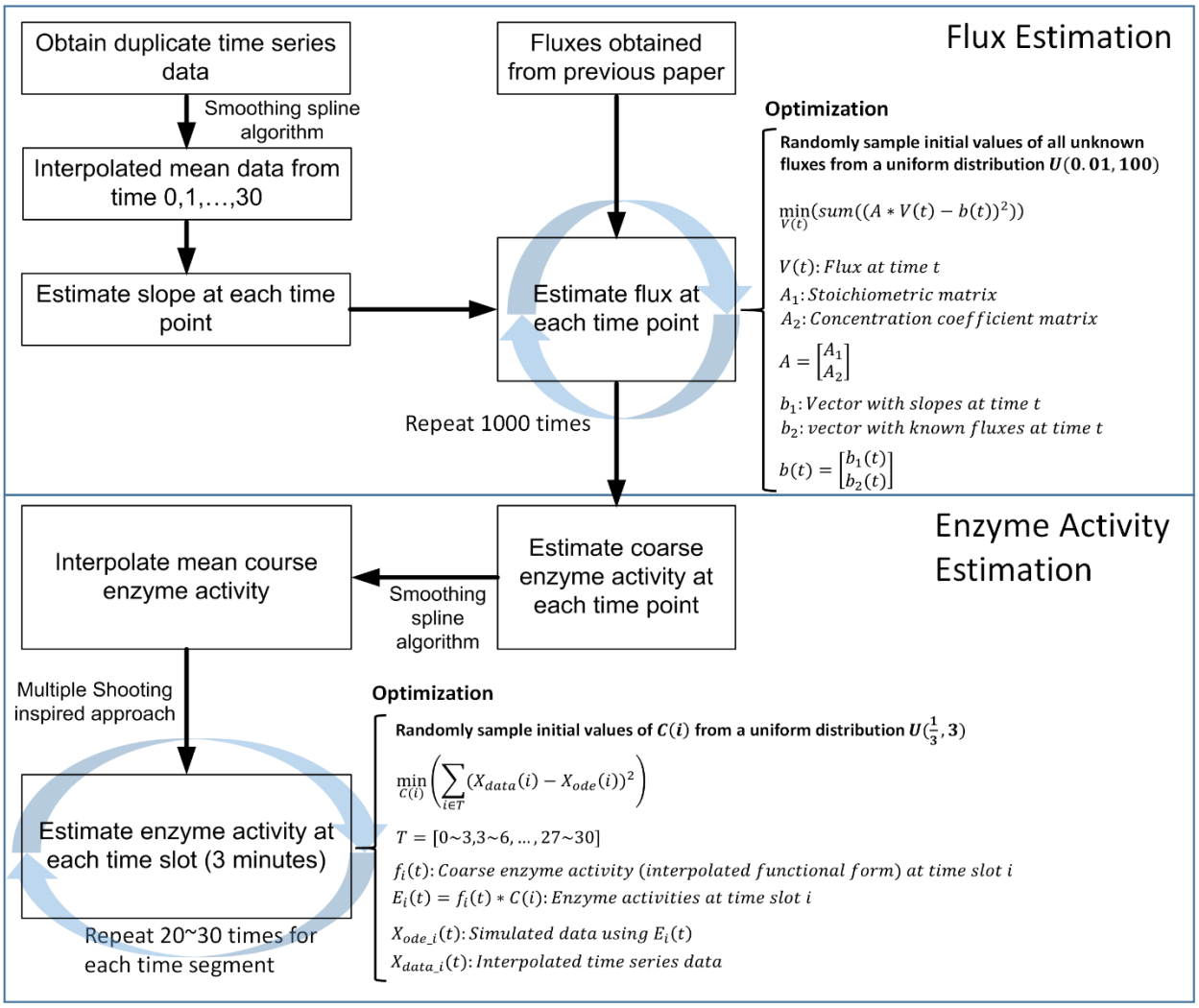


Figure 4.11. Details of the procedures of flux estimation and enzyme activity estimation.

Each flux is determined such that the model as a whole matches the observed data within a sufficiently small range of noise. Furthermore, all subsets of these fluxes, for instance, those representing the DHC hydroxylases, must collectively be consistent with the known fluxes of our previous model (see table in Figure 4.2). Also, all fluxes should change relatively smoothly from one time interval to the next. Finally, all fluxes are subject to upper and lower bounds. These tasks are formulated as a constrained nonlinear optimization program. Specifically, at each time point, we use two types of constraints. First, we constrain the system by ensuring that the slopes of the dependent variables (X) at a given time point t are

sufficiently close to $X(t) - X(t-1)$, which we accomplish by minimizing the sum of squared errors between these differences and the corresponding slopes of these X variables. Second, we constrain the system by ensuring that the fluxes entering the system from the outside are collectively equivalent to those of the former model (*cf.* Table in Figure 4.2). These two constraints can be formulated as a combined, single objective function. To achieve robustness of the solution, the system is solved repeatedly by assigning for all unknown fluxes initial values that are drawn randomly from the uniform distribution U(0.01, 100). The results suggested that 1,000 simulations return sufficiently diverse ensembles.

One could surmise that the rather strong variability in the time series data (Figure 4.10) would unduly affect the flux estimation. To test this hypothesis, we estimated fluxes based on interpolated concentration data that were perturbed in either direction by a random factor sampled from U(1/1.5, 1.5). We compared these estimated fluxes with those obtained from noise free interpolated data. The flux distributions show very similar patterns (Figure A4.1 and A4.2); detailed procedures are provided in the *Appendix 4*.

As an example, the resulting fluxes of reactions catalyzed by ceramide synthase are given in Figure 12; other fluxes are shown in the *Appendix* 4 (Figure A4.1). Once the flux distributions are computed for each time point, we examine the histogram of each flux in each time point to ensure that the solutions are well constrained; details are presented in the *Appendix 4* (Figure A4.3). The analysis revealed that most of the flux distributions at any given time point were rather tightly bell shaped, suggesting the use of the mean value of each flux at each time point as an appropriate, time-dependent estimate (Figure 4.13). We also redid the analysis with medians, but the results were essentially the same.

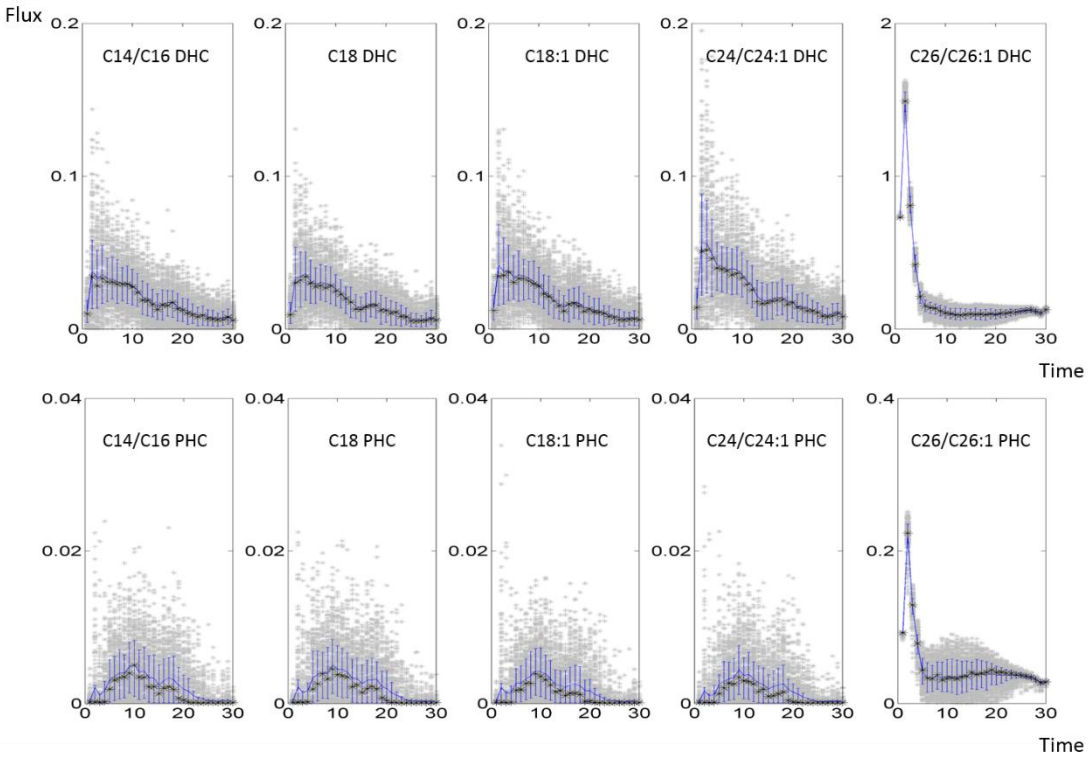


Figure 4.12. Dynamic flux distributions of 10 reactions catalyzed by ceramide synthase. *Each grey dot represents one simulation. The blue line indicates the means, the black asterisks show the medians, and blue bars represent 20th and 80th percentiles (blue bars).*

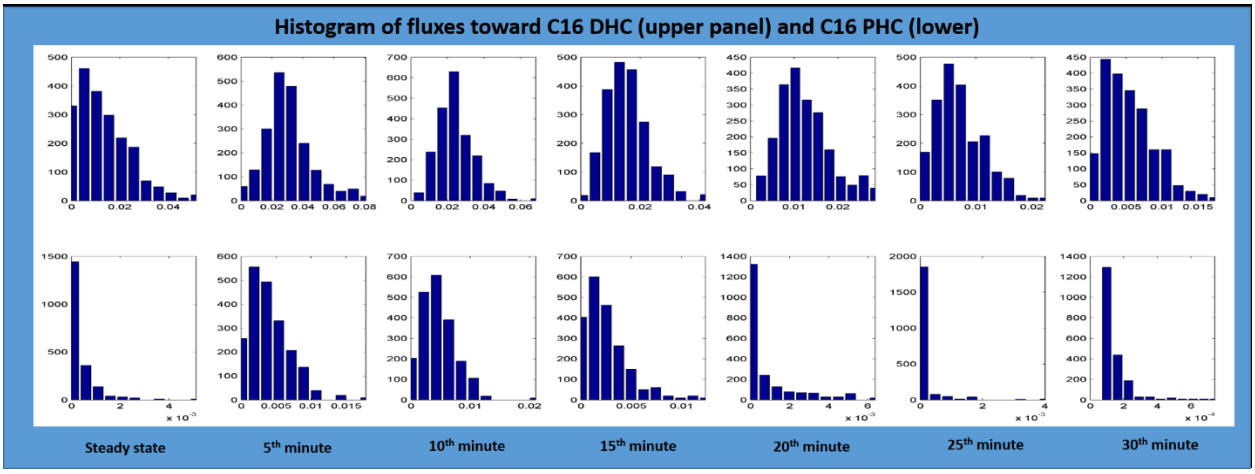


Figure 4.13. Histograms of ceramide synthase fluxes from C16 DHS to C16 DHC (upper panel) and from C16 PHS to C16 PHC (lower panel). *Time 0 represents normal steady-state temperature conditions at the beginning of the heat stress experiment, which lasts for 30 minutes.*

The method of Dynamic Flux Estimation [124] allows us to obtain important hints for how material flows within the system. The resulting flux estimates are also important for constructing a dynamic model of ceramide synthesis and degradation upon heat stress. For later purposes, we need to identify how substrates, enzyme, modulators and kinetic parameters contribute to the magnitude of a flux. Because the actual enzyme amounts and rate constants are not known, we assume, as it is commonly done, that enzyme activities enter a flux representation in a linear manner and that substrates contribute with a power of 1, which corresponds to a mass-action formulation. Given these assumptions we obtain coarse estimates of the product of the rate constant and the enzyme activity. This product is similar to a V_{max} value, which by definition consists of the product of k_{cat} and the total enzyme concentration. Some of these estimates are shown in Figure 14 as functions of time. (All estimates are presented in Figure A4.3 of the *Appendix*)

Based on these time dependent estimates of flux magnitudes, we employed smoothing splines with proper degrees to obtain smooth time trends in enzyme activities (please refer to lower panel in Figure 4.11). The functional representations of enzyme activities were entered into the ODE model in order to obtain trends in the different ceramide species. Under ideal conditions, these trends should match the observed concentration profiles. By using the residual errors between the results of the described modeling strategy and the data, we created an optimization strategy that iteratively refined the trends in enzyme activities. This strategy was gleaned from the method of Multiple Shooting, which is a well-documented methodology for fitting dynamical data in boundary value problems [125,126]. While most optimization methods that correspond to a single shooting strategy aim at searching for one parameter set to fit the observed dynamic trajectory in its entirety, multiple shooting splits the time series into successive time frames and initially searches for independent parameter sets that match the data one frame at a time. In many cases of complex dynamic systems, this type of multiple shooting has demonstrated a better performance than single shooting. Here, we address a separate initial value problem for each time frame, and the last data point in each time frame is not defined as a condition for the subsequent frame.

We applied the multiple-shooting inspired strategy to fit the various ceramide time series data. However, instead of searching for entirely new parameter sets for each time frame, the task here is simpler,

because the algorithm searches merely for slight adjustments of the coarse functional forms of enzyme activities that we had previously derived for each time frame to fit the ceramide data. Specifically, we associated unknown coefficients C to the enzyme activities in the ODEs. For example, from the 3rd to the 6th minute of heat stress, the functional representation of the enzyme activity, $f_{3-6minute}(t)$, was replaced with $C_{3-6minute} * f_{3-6minute}(t)$. With all coefficients set equal to one, the ODEs are unchanged. However, by using the coefficients as free parameters, we are now capable of adjusting the system dynamics in each time frame. Thus, we subdivided the 30-minute time frame of the heat stress experiment into 3-minute intervals and fit the data separately in each interval. Furthermore, to minimize bias, we executed the search algorithm with many random initial settings for each coefficient, each enzyme, and each timeframe so that we obtained ensembles of solutions within a larger solution space. The slightly adjusted enzyme activities were then collected for biological inference.

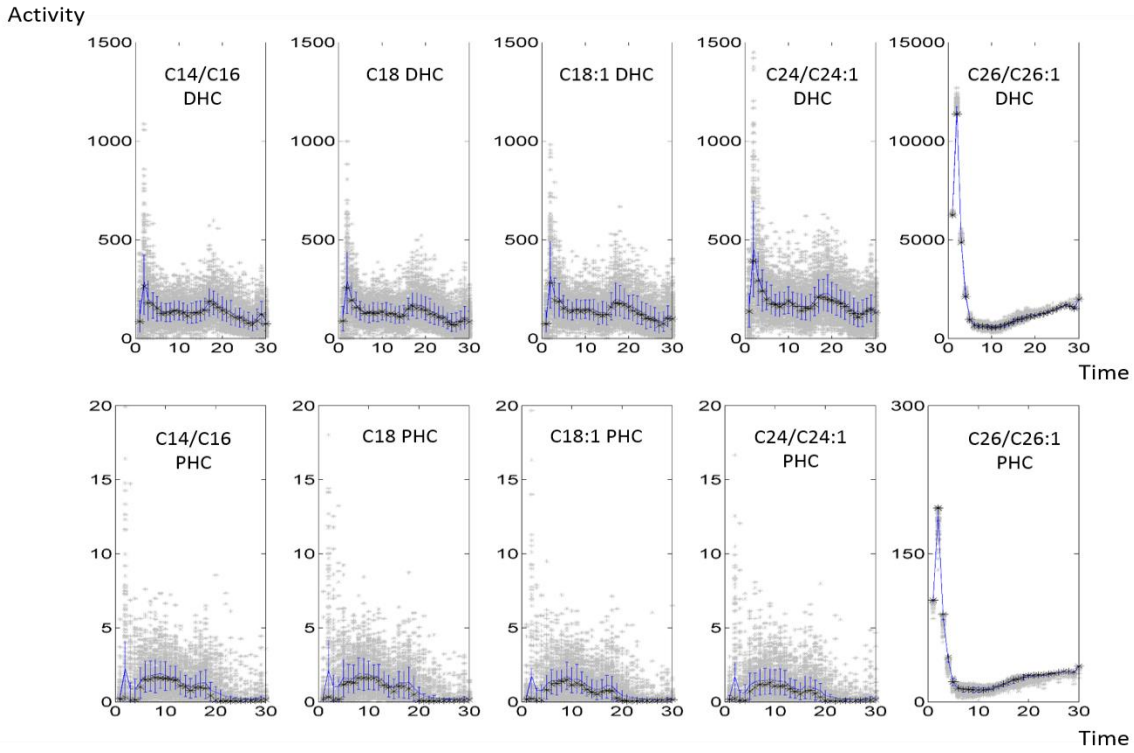


Figure 4.14. Rough estimation of ceramide synthase activities. *Each grey dot represents one simulation. The blue line indicates the means, the black asterisks show the medians, and blue bars represent 20th and 80th percentiles (blue bars)*

CHAPTER 5

ANALYSIS OF THE INVOLVEMENT OF DIFFERENT CERAMIDE VARIANTS IN THE RESPONSE TO HYDROXYUREA STRESS IN BAKER'S YEAST

5.1 Introduction

Mammalian sphingolipids actively participate in triggering crucial cell functions as second messengers. Within the sphingolipid family, ceramides have been investigated for several decades. More recently, they have been receiving increased attention because they are directly involved in signal transduction that ultimately leads to cell proliferation, differentiation, and apoptosis [107,109,127]. While the relationships between ceramide species, signaling cascades and gene expressions are gradually becoming clearer, it is still unknown how the concentrations of the various ceramide species are altered by cells in response to stresses. Two reasons for this lack in understanding are that sphingolipid metabolism constitutes a complex, highly regulated pathway system and that consistent metabolic and enzymatic time series measurements are difficult to obtain.

In contrast to mammalian cells, yeast cells are much more easily investigated, and because they are also eukaryotic, they have become valuable model systems for mammalian stress responses. In particular, the metabolism of sphingolipids is highly conserved and quite similar between mammalian cells and yeast, and it is feasible to measure stress responses more or less directly and with a time resolution of a few minutes. However, these measurements alone do not reveal the strategies that cells evoke to respond to stresses. Thus, using similar methods as in Chapters III and IV, I am proposing here a computational approach that uses experimental time coarse data of different yeast ceramide variants and is designed to

reveal and characterize patterns of metabolic regulation within the ceramide pathway under hydroxyurea stress.

The reasons for choosing hydroxyurea stress are the following. Hydroxyurea inhibits the enzyme ribonucleotide reductase and thereby decreases or even stalls DNA synthesis. Due to this property, hydroxyurea has been used increasingly as a treatment option for a variety of diseases including HIV infection and AIDS, sickle cell anemia, and myeloproliferative neoplasms. In spite of the growing interest in hydroxyurea, details of the mechanisms of action are not fully understood, and it is therefore not surprising that the investigation of cellular responses to hydroxyurea has become a highly interesting topic in biomedical research.

Using baker's yeast as a model organism, the roles of ceramides in signal transduction under hydroxyurea exposure are beginning to become clearer. For instance, the tolerance of yeast cells to hydroxyurea exposure decreases in knockouts of each one of several genes, including isc1 (IPCase) and sur4 (fatty acid elongation), in double gene knockouts such as lag1 and lac1 (ceramide synthase), as well as in strains overexpressing ydc1 (phytoceramidase) [108]. These findings strongly imply the involvement of ceramide in mediating hydroxyurea stress. Further research has suggested a signaling cascade, which starts with an increased concentration of C18:1 phytoceramide (C18:1 PHC, phytoceramide with a C18:1 fatty acyl CoA), which triggers the activation of a sub-domain of Cdc55/PP2A (a regulatory subunit for protein phosphatase 2A) that leads to a decrease in Swe1 (a mitosis inhibitor protein kinase) level, causes dephosphorylation of Clb2-Cdc28 (which activates Cdc28p to promote the cell cycle transition from G2 to M), and ultimately activates the G2/M checkpoint.

While it has thus been suggested that yeast C18:1 PHC mediates the hydroxyurea stress response via a multi-step signaling cascade, it is not clear which mechanisms alter the level of C18:1 PHC in the first place. These mechanisms are not easy to decipher with intuition alone, because the biosynthesis of C18:1 PHC is embedded in the complicated pathway system of sphingolipid biosynthesis (Figure 5.1). This system, to which I will refer in the following sections as *ceramide metabolism*, actually consists of parallel pathways that generate and utilize saturated and unsaturated ceramides with different fatty acyl chain

lengths. When formulated in sufficient detail to account for all different ceramide variants, this pathway is so complicated that targeted alterations of specific ceramide molecules can hardly be predicted without a targeted computational analysis (Figure 5.2). Nevertheless, it is important to shed light on the details of the involved metabolic processes because very specific ceramide variants appear to be the first responders to stresses, including hydroxyurea exposure and heat.

In the following, I will describe computational approaches that are based on a series of comprehensive sphingolipid pathway models that our team has been developing over the past decade [10,11,116]. In contrast to all previous studies, I focus here on the detailed cellular responses to hydroxyurea stress.

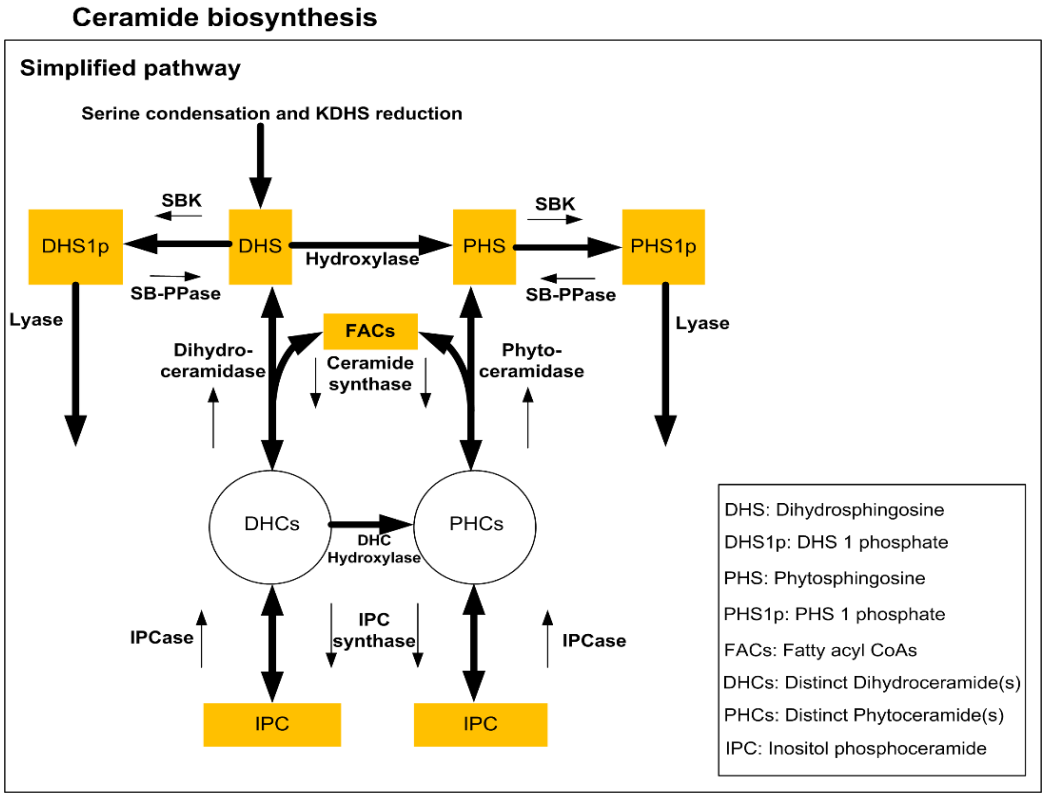


Figure 5.1. Ceramide biosynthesis pathway in simplified representation.

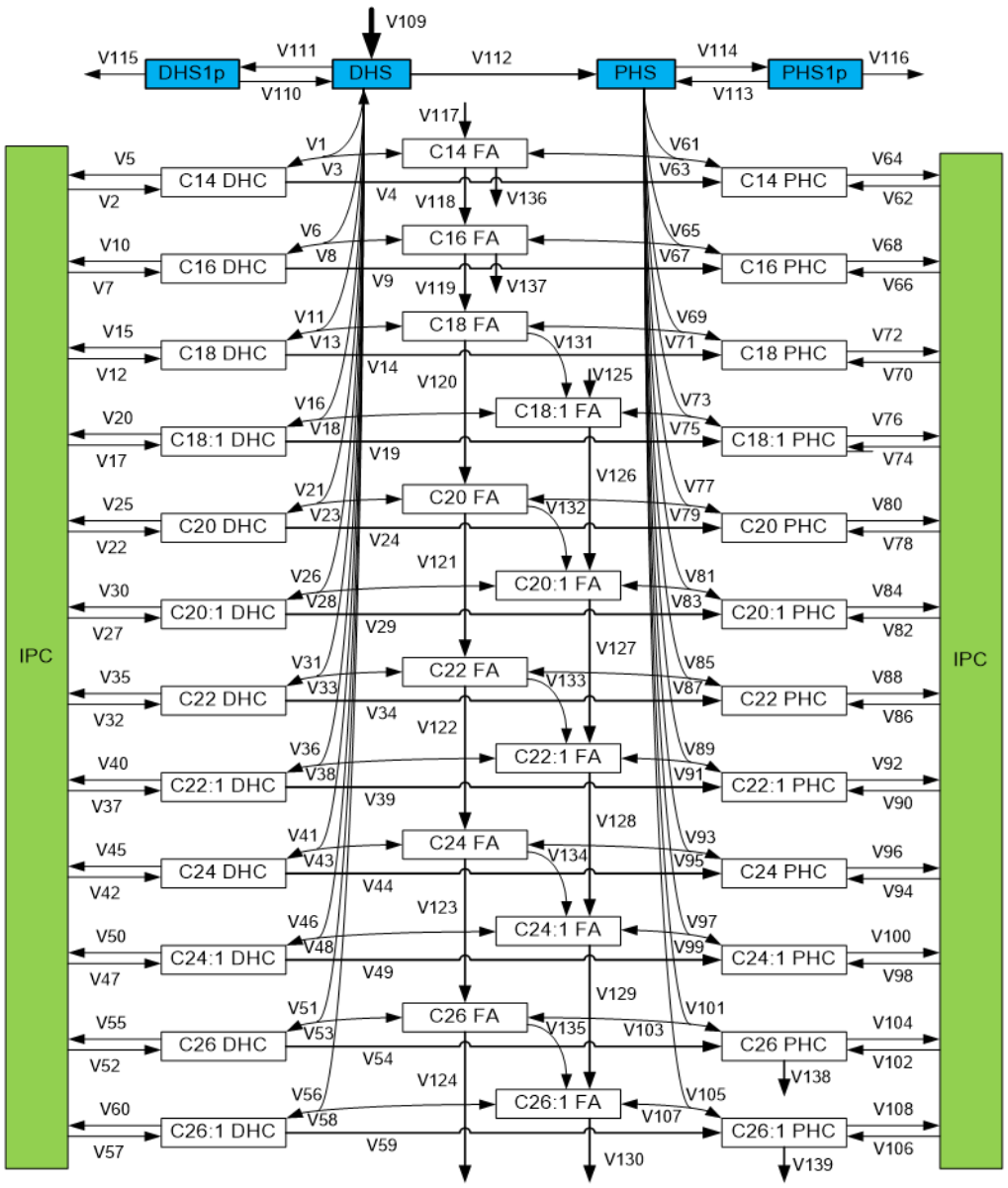


Figure 5.2. Ceramide metabolism consists of numerous parallel and cross reactions that lead to the production of distinct saturated and unsaturated dihydro- (DHC) and phyto- (PHC) ceramide variants with different fatty acyl (FA) chain lengths.

5.1.1 Ceramide Metabolism

Ceramide is a crucially important species within the group of sphingolipids. It consists of a long chain base (LCB) backbone and a fatty acyl group. Long chain bases usually contain 18 carbon bases, but

may also have C20 bases. The most prevalent LCBs in mammalian cells are sphingosine, dihydrosphingosine (DHS) and phytosphingosine (PHS), whereas yeast only produces DHS and PHS. Most ceramide variants possess fatty acyl groups ranging from C12 to C26. Since we only consider C18 LCBs in this research, the distinct ceramide species can be named in accordance with their fatty acyl groups. For instance, C16 dihydroceramide (C16 DHC) denotes the ceramide N-palmitoyldihydrosphingosine or Cer(d18:0/16:0).

In yeast, ceramides can be generated through *de novo* biosynthesis and through the degradation of complex sphingolipids (Figure 5.1). Sphingolipid biosynthesis begins with the condensation of serine and palmitoyl CoA, which is catalyzed by the enzyme serine palmitoyltransferase (SPT). This first step produces 3-keto-dihydrosphingosine (3KDHS), which is quickly reduced to DHS by the enzyme KDHS reductase. DHS is the key LCB and serves as a source for the other LCB, PHS, as well as for DHC. These reactions are reversible. Conversions between DHS and PHS are catalyzed by sphingoid base kinase (DHS to PHS) and sphingoid base phosphatase (PHS to DHS), respectively. DHS can also be converted to different variants of DHCs depending on the type of fatty acyl CoA the enzyme ceramide synthase uses. The reverse reaction from DHC to DHS is catalyzed by dihydroceramidase. PHS can be converted to PHC followed by the similar mechanism as between DHS and DHC, but the reverse reaction is catalyzed by the specific enzyme phytoceramidase. The final reaction of the pathway system is DHC hydroxylase, which catalyzes the reaction from DHC and PHC.

The degradation of complex sphingolipids constitutes an alternative path toward ceramide species. Complex sphingolipids include inositol phosphoceramide (IPC), mannose inositol phosphorylceramide (MIPC) and mannose di-inositol phosphorylceramide (MIP_2C); all of these can serve as sources of ceramide species. IPC can be converted reversibly into MIPC, and MIPC can be converted reversibly into MIP_2C. The degradation of complex sphingolipids is catalyzed by the enzyme isc1 (IPCase), which is the yeast homologue to the mammalian neutral sphingomyelinases. The reverse reactions from DHC/PHC to IPC are catalyzed by the enzyme IPC synthase.

In summary, the ceramide levels are controlled by five enzymes, namely ceramide synthase, dihydroceramidase and phyotoceramidase, IPC synthase, IPCase (Isc1), and DHC hydroxylase. These reactions, combined with the array of different ceramide variants, form a complicated metabolic pathway system (Figure 5.2 and *Methods* section) that renders an intuitive understanding of the dynamic responses of hydroxyurea very challenging. To shed light on this system, we have developed a computational strategy that is based on an ordinary differential equation (ODE) model and uses experimental time series data as input (for details see *Methods*). This approach will assist us in the dissection of the regulatory strategies with which ceramide metabolism responds to hydroxyurea exposure.

5.2 Results and Discussion

The concentrations of ceramides depend on the fluxes that enter or leave each pool and contribute or utilize mass. The magnitudes of these fluxes are mainly determined by two main factors: the level of substrates and the activity of the catalyzing enzyme. Within the framework of mass action kinetics, as well as other frameworks, the latter is often subsumed into the rate constant of the reaction. Using the techniques described in the *Methods* section, we have inferred these rate constants for all enzymatic steps involved. Summaries of these inferences are provided below. They are sorted by the five key enzymes of ceramide metabolism, namely, ceramide synthase, dihydroceramidase and phytoceramidase, IPC synthase, IPCase (Isc1), and DHC hydroxylase. Furthermore, because substrate concentrations are known from interpolation of the measured data, we are able to compute flux magnitudes. Thus, in the second half of this *Results* section, we describe insights from a detailed mass flow analysis that is based on these estimated flux magnitudes and offers novel insights into the mass flow patterns of different groups of ceramides.

5.2.1 Enzyme Activities

Activity patterns of the key enzymes ceramide synthase, dihydroceramidase or phytoceramidase, IPC synthase, IPCase, DHC hydroxylase and other auxiliary enzymes were inferred and analyzed for the

20-hour experimental time period. Many factors can potentially influence these activities, including protein quantity, structure (folding/unfolding), metabolic regulation, and post-translational modifications. While it is impossible to characterize these secondary effects based on the currently available data, our computational approach reveals interesting regulatory strategies in several aspects, although it focuses on overall enzyme activities and rate constants.

The main results consist of two types. Time series estimates are directly calculated from the optimization approach (see *Methods*). They are shown as gray dots, which represent individual simulation results, and blue lines, which depict simulation averages, as well as red linear regression lines. These regression lines immediately display the dynamic trends in enzyme activities during hydroxyurea exposure.

As an additional visualization, color-coded maps are computed from the dynamic time series. Specifically, the trends are simplified by two linear regressions, the first for the time period [0, 3], and the second for the time period [3, 20]; the cut at 3 hours reflects the measurement time point in the data. The trend for the first time window is shown in the left box at the top of each display associated with a reaction, while the trend for the second time window is shown in the right box on top. The wider box below these two displays the trend over the entire time period [0, 20]. These simplified representations immediately visualize the similarities and differences of dynamic enzyme activities within pools of reactions that involve the same specific enzyme.

1. Ceramide synthase

Ceramide synthase shows different patterns of activities when using DHS or PHS as substrate. In detail, the activities of ceramide synthase using DHS as a substrate exhibit a gradually decreasing pattern for fatty acyl chain lengths up to C24 DHC, while the corresponding reactions using PHS as substrate show increasing trends (except for C24:1 PHC). This finding can be further analyzed by means of the color-code maps in Figure 5.3. Interestingly, the reaction rates of these groups show opposite directions over the entire 20-hour experiment. Specifically, all color boxes at the bottom of each reaction for DHC are light blue, indicating a slight overall decrease, while they are pink for PHC, which indicates a slight overall increase.

In addition to these overall trends, one finds differences in activities during the [0, 3] and [3, 20] hour windows.

The different patterns of ceramide synthase activities may be interpreted in two ways. First, the biosynthesis of sphingolipids is proceeding at a moderate pace, since most of the activities of ceramide synthase are within a 2-fold range. At the same time, the slight but important trends of gradually decreasing or increasing ceramide synthase activities in DHS versus PHS may be due to competition between DHS and PHS toward ceramide synthase.

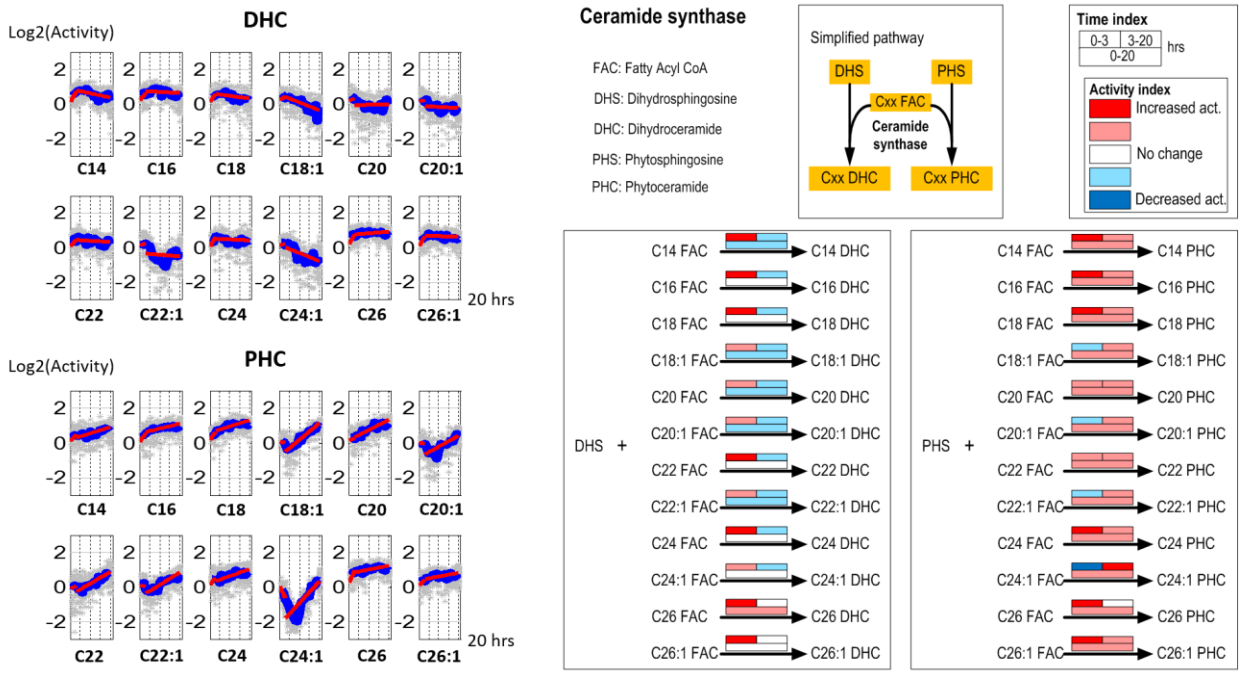


Figure 5.3. Ceramide synthase activities. Left panel: Time series estimates of ceramide synthase activities. Gray dots represent individual simulation and blue lines are averaged activities. Each subplot contains two segments of red lines indicating linear regression over 0-3 hours and 3-20 hours intervals. Right panel: Color-coded map. Each array of boxes associated with a reaction step indicates average slopes of the dynamic trends shown in the left panel for three time intervals: 0-3 (top left box), 3-20 (top right box) and 0-20 hours (wide bottom box).

2. Ceramidase

Ceramidases (dihydroceramidase and phytoceramidase) and IPC synthase catalyze the effluxes of ceramides; expressed differently, up-regulation of these enzymes decreases the ceramide concentrations. The cellular strategy of increasing dihydroceramidase activity to regulate the amount of DHC under heat stress has been documented. Our simulation results suggest a similar strategy under hydroxyurea stress (Figure 5.4). In a very consistent pattern, the activities of dihydroceramidase gradually increase while trends of phytoceramidase drop. These trends eventually result in a relatively low DHC level and a high PHC concentration.

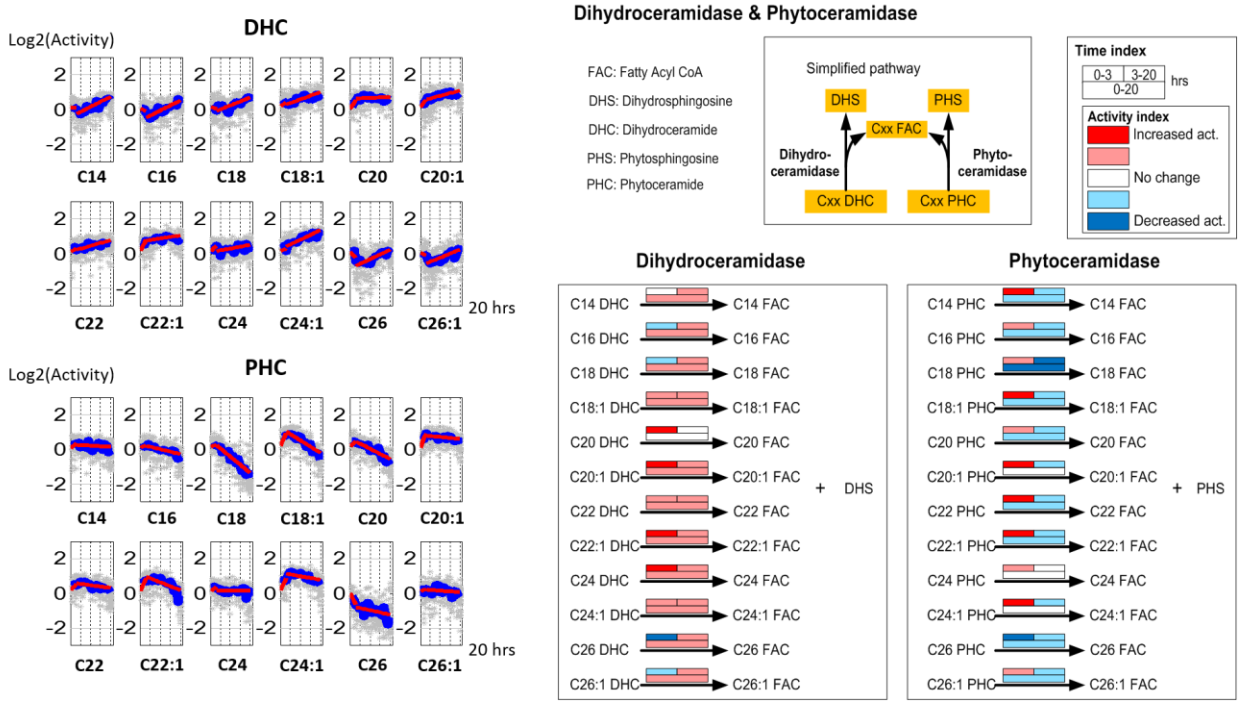


Figure 5.4. Dihydroceramidase and phytoceramidase activities. Left panel: Time series estimates of dihydroceramidase and phytoceramidase activities. Gray dots represent individual simulation and blue lines are averaged activities. Each subplot contains two segments of red lines indicating linear regression over 0-3 hours and 3-20 hours intervals. Right panel: Color-code map. Each array of boxes associated with a reaction step indicates average slopes of the dynamic trends shown in the left panel for three time intervals: 0-3 (top left box), 3-20 (top right box) and 0-20 hours (wide bottom box).

3. IPC synthase

The activities of IPC synthase, which use DHC or PHC as substrate, show similar patterns as dihydroceramidase and phytoceramidase, respectively (Figure 5.5). This finding could suggest that IPC synthase might be a regulator for balancing DHC and PHC concentrations. However, the existence of only one IPC synthase in yeast limits this potential regulatory role. In fact, the color-code representation in the right panel of Figure 5.5 exhibits lower decreases in activities for almost all ceramide species compared with ceramidases. The distinct patterns of time-dependent activities might result from substrate competition toward IPC synthase or from compartmentalization.

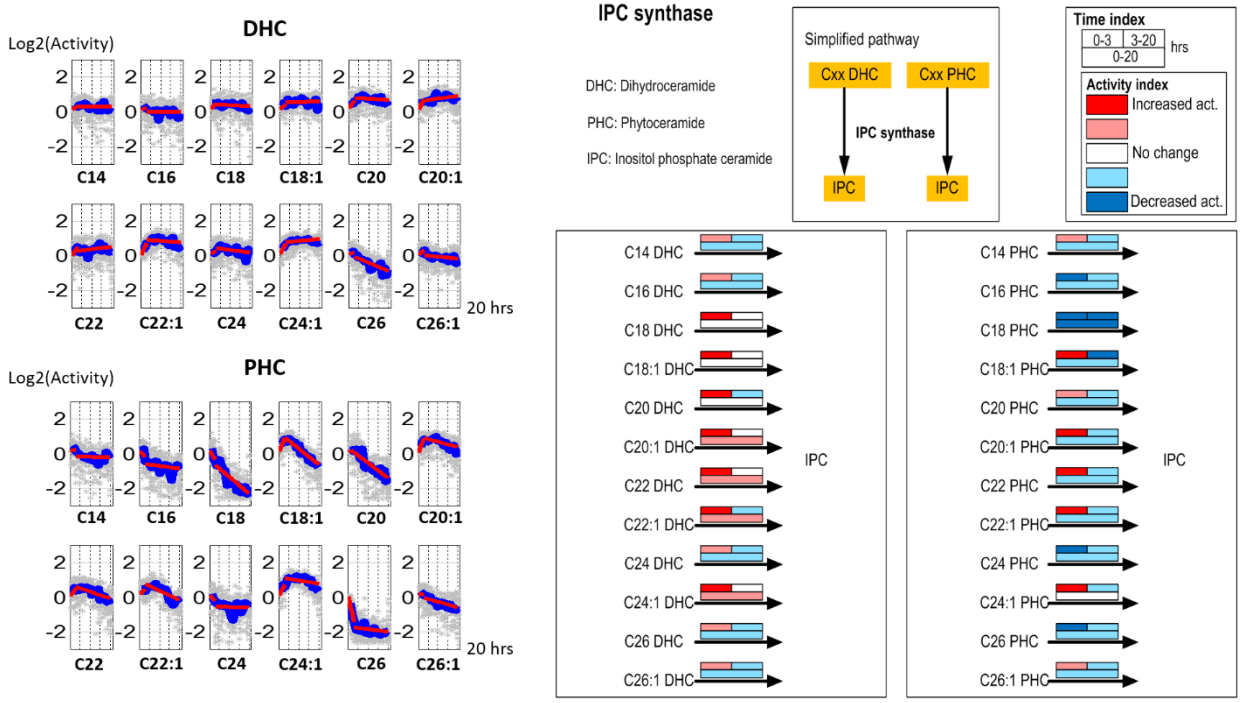


Figure 5.5. IPC synthase activities. Left panel: Time series estimates of IPC synthase activities. Gray dots represent individual simulation and blue lines are averaged activities. Each subplot contains two segments of red lines indicating linear regression over 0-3 hours and 3-20 hours intervals. Right panel: Color-code map. Each array of boxes associated with a reaction step indicates average slopes of the dynamic trends shown in the left panel for three time intervals: 0-3 (top left box), 3-20 (top right box) and 0-20 hours (wide bottom box).

Although the changes in activities of IPCase and the ceramidases look somewhat similar, the overall trends, as given in the color maps, are distinct. The strong contrast in ceramidase activity for DHC and PHC may suggest the existence of two enzyme variants (dihydroceramidase and phytoceramidase) while the similarity of IPC synthase activities suggests that there is only one IPC synthase enzyme in yeast.

4. IPCase (Isc1)

Under hydroxyurea stress, up-regulation of IPCase, or Isc1, has been reported in the literature. In our simulation, IPCase activities in most of the reactions match this finding over the 20-hour period, although the initial activity decreases in several cases. These cases, C22:1 DHC, C24:1 DHC, C18:1 PHC and C24:1 PHC, may be explainable with substrate competition toward IPCase, since IPC forms different substrates from IPC, MIPC and MIP_2C, and each of these contains distinct fatty acyl groups. However, the transient lower activities in the first 5-7 hours are compensated by later activation, which results in increased long term activities (Figure 5.6).

C18:1 PHC has been identified as a key signaling intermediate under hydroxyurea stress in yeast. Our simulation, instead of identifying a single key enzyme (such as IPCase) that would cause this increase, suggests a highly coordinated strategy that: (1) slightly increases ceramide synthase activity using PHS and C18:1 fatty acyl CoA as substrates (Figure 5.3); (2) slightly decreases phytoceramidase and IPC synthase activity using C18:1 PHC as substrate (Figure 5.4 and 5.5); and (3) subtly modulates IPCase activity (toward C18:1 PHC) to achieve the overall C18:1 PHC level (Figure 5.6). This postulated strategy suggests an energetically and metabolically effective way to control metabolites. A similar inference was made for changes in sphingolipid metabolism during the diauxic shift, where many enzyme activities were altered rather slightly, rather than a few enzymes that were altered much.

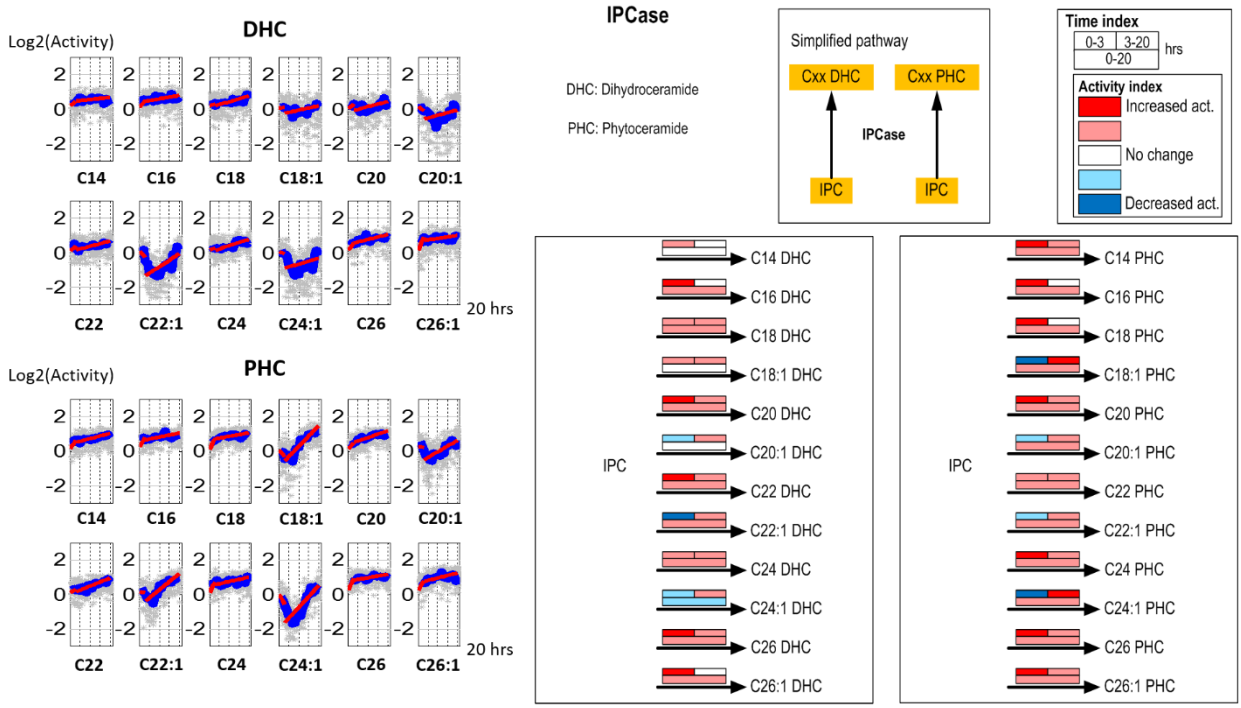


Figure 5.6. IPCase (Isc1) activities. Left panel: Time series estimates of IPCase (Isc1) activities. Gray dots represent individual simulation and blue lines are averaged activities. Each subplot contains two segments of red lines indicating linear regression over 0-3 hours and 3-20 hours intervals. Right panel: Color-code map. Each array of boxes associated with a reaction step indicates average slopes of the dynamic trends shown in the left panel for three time intervals: 0-3 (top left box), 3-20 (top right box) and 0-20 hours (wide bottom box).

5. DHC Hydroxylase

Simulations of DHC hydroxylase reveal no real trends, especially when the entire 3-20 hour experiment is considered (Figure 5.7). The activities do change somewhat for different DHC variants, but these slight changes may be compensatory.

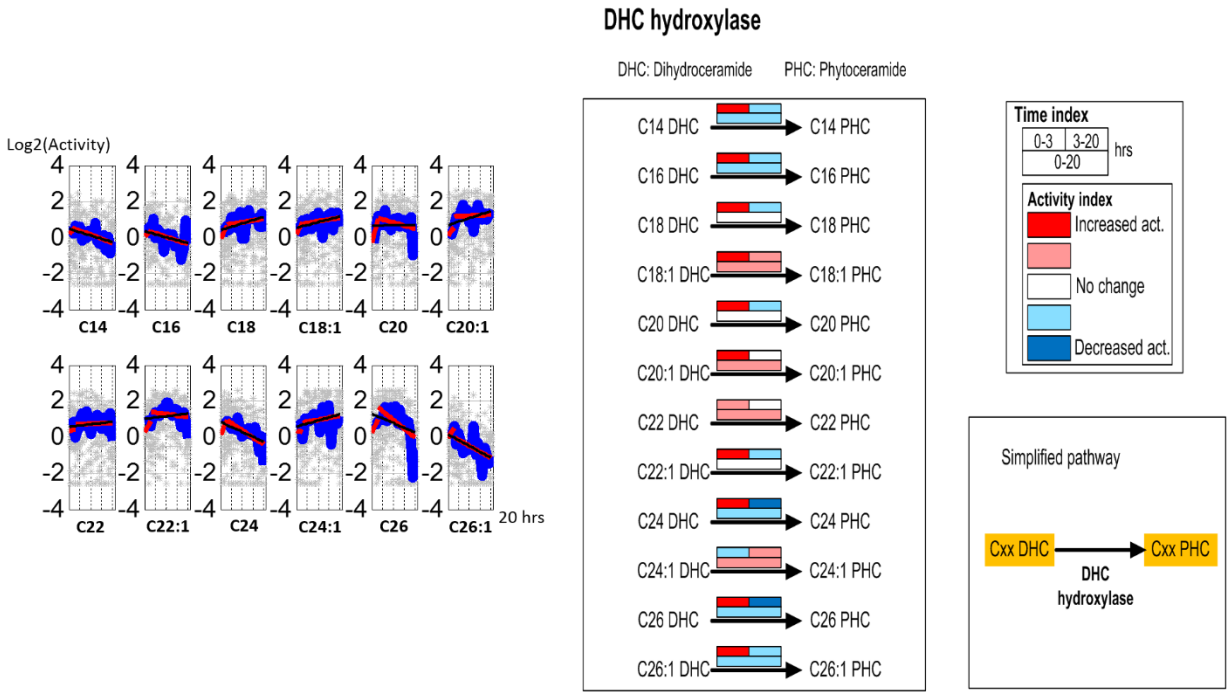


Figure 5.7. DHC hydroxylase activities. Left panel: Time series estimates of DHC hydroxylase activities. Gray dots represent individual simulation and blue lines are averaged activities. Each subplot contains two segments of red lines indicating linear regression over 0-3 hours and 3-20 hours intervals. Right panel: Color-code map. Each array of boxes associated with a reaction step indicates average slopes of the dynamic trends shown in the left panel for three time intervals: 0-3 (top left box), 3-20 (top right box) and 0-20 hours (wide bottom box).

6. Other enzymes

Enzymes associated with sphingolipid biosynthesis, fatty acid elongases and desaturase, are not as important as the enzymes listed above for regulating ceramide concentrations. However, these enzymes do affect the balances between auxiliary metabolites, such as DHS, PHS and fatty acyl CoAs, which are crucial for maintaining the appropriate ceramide concentrations. Therefore, it is beneficial to discuss the activities of these enzymes as well.

Sphingolipid biosynthesis and utilization

DHS biosynthesis (which here combines palmitate uptake, serine condensation and KDHS desaturation) remains constantly upregulated throughout the 20-hour experiment. This observation suggests

a persistent uptake of materials from the medium. The remaining enzymes, SBK (sphingoid base kinase), SB-PPase (sphingoid base phosphatase), lyases (which remove DHS-p and PHS-p from the system) and hydroxylase (which converts DHS into PHS), along with ceramide synthase and ceramidases, cooperatively maintain DHS, PHS, DHS-p and PHS-p concentrations at appropriate levels. (Figure 5.8).

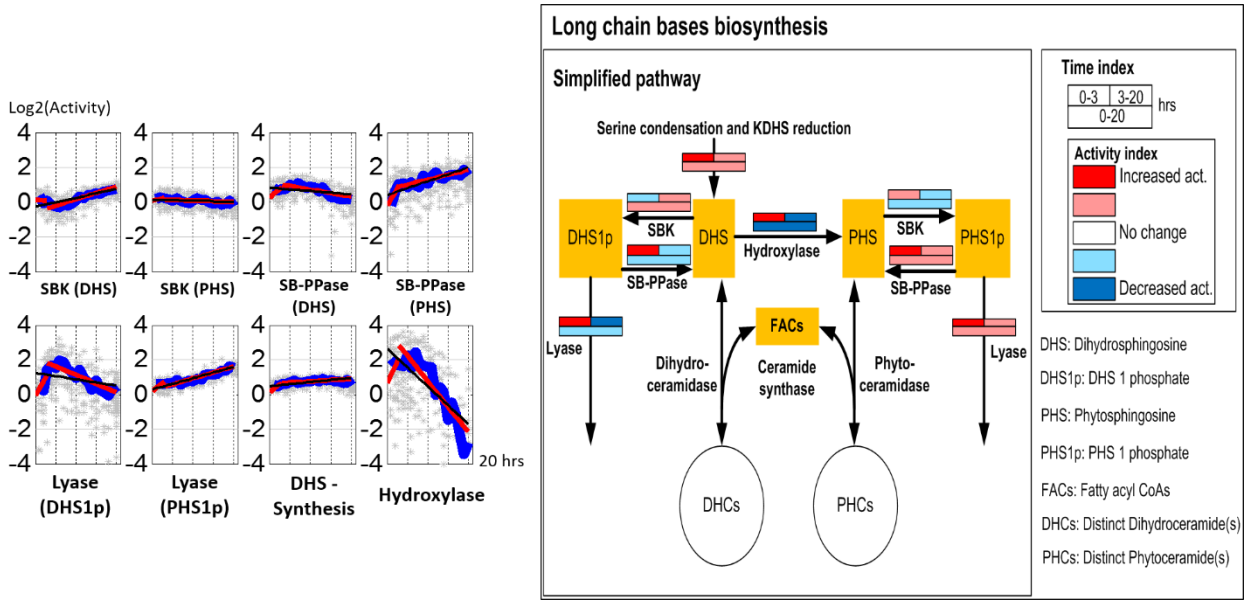


Figure 5.8. Enzyme activities of sphingolipid biosynthesis. Left panel: Time series activities of enzymes catalyzing sphingolipid biosynthesis. Gray dots represent individual simulation and blue lines are averaged activities. Each subplot contains two segments of red lines indicating linear regression over 0-3 hours and 3-20 hours intervals. Right panel: Color-code map. Each array of boxes associated with a reaction step indicates average slopes of the dynamic trends shown in the left panel for three time intervals: 0-3 (top left box), 3-20 (top right box) and 0-20 hours (wide bottom box).

Elongases and desaturase

Yeast expresses three elongases, Elo1, Elo2 and Elo3, which catalyze elongation reactions from C12 to C18, C14 to C22, and C18 to C26 fatty acid, respectively. Furthermore, desaturase (Ole1) catalyzes reactions from saturated to unsaturated fatty acyl CoAs. Time series estimates of these enzymes all show unremarkable patterns throughout the 20-hour experiment, even though the desaturase active rises slightly with time (Figure 5.9). The reason for their constancy might be that fatty acid elongation is an important

cellular event, and it appears that the demand for fatty acids of various chain lengths does not change much over time.

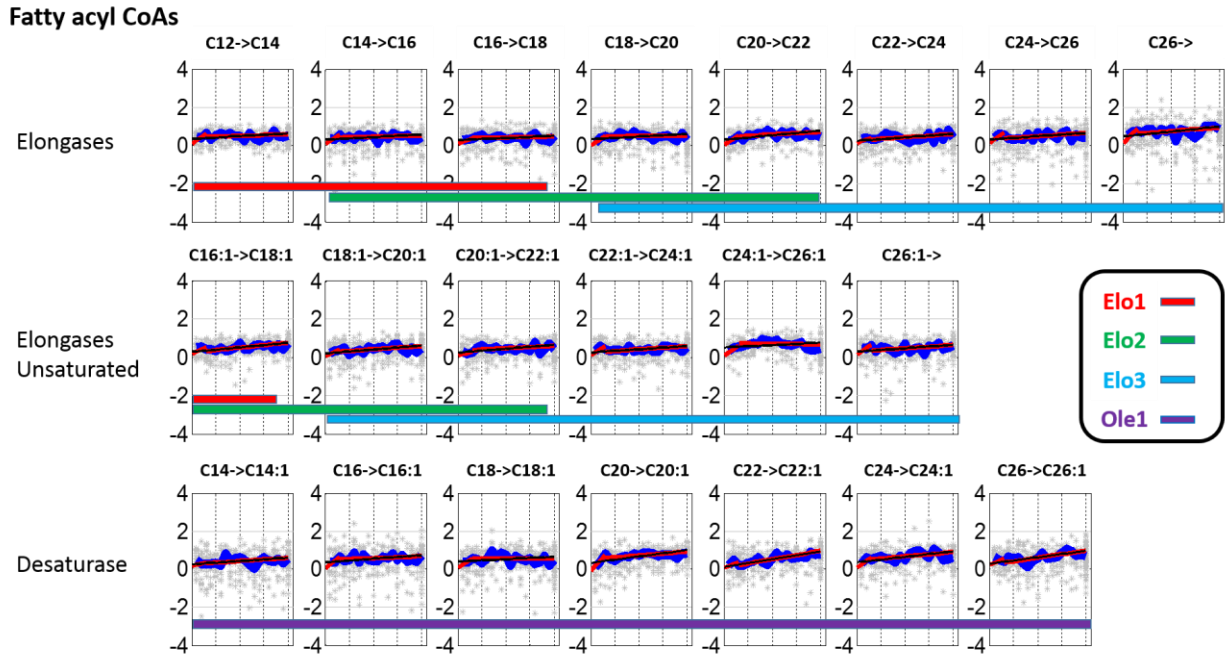


Figure 5.9. Time series estimates of elongases and desaturase.

5.2.2 Mass Flow Analysis

The flux into or out of a metabolite pool is determined by the concentrations of substrates and the rate constant including the enzyme activity. Thus, in order to assess changes in flux magnitudes, it is not sufficient to study changes in rate constants alone. We considered this fact in our approach and conducted a mass flow analysis that estimates materials entering and leaving pools in the system from estimated rate constants (details in *Methods*) and measured or interpolated substrate concentrations.

Once individual dynamic fluxes are computed, it is possible to assess the mass flow throughout the entire system during the 20 hours of hydroxyurea stress (see *Methods* for details). The mass flow around ceramides, including DHCs, PHCs and fatty acyl CoAs, suggests how materials are synthesized, degraded

or recycled during the 0-3 hour and 3-20 hour time periods, and throughout the entire 20 hours of hydroxyurea stress.

The flux distribution in Figure 5.10 indicates that sphingolipid biosynthesis is always active. DHCs and PHCs are maintained through two main fluxes. One is the net flux of the balance between ceramide synthase on the one hand and dihydroceramidase or phytoceramidase on the other. This net flux is always high, particularly during the [3, 20] hour time window. Interestingly, the net fluxes point from DHS to DHC, but from PHC to PHS. The other flux is controlled by DHC hydroxylase. It is similarly high and carries much more mass than the small positive net flow from IPC to PHCs in the first three hours, which is actually reversed during the second phase of the experiment. Thus, it is interesting to note that material flows in a consistent manner through the system from DHS to DHC to PHC to PHS to PHS-p, from where lyase removes material from ceramide metabolism. Expressed differently, there is no drastic change of this flux pattern throughout the 20 hours of stress, and there is no change in the direction of mass flow.

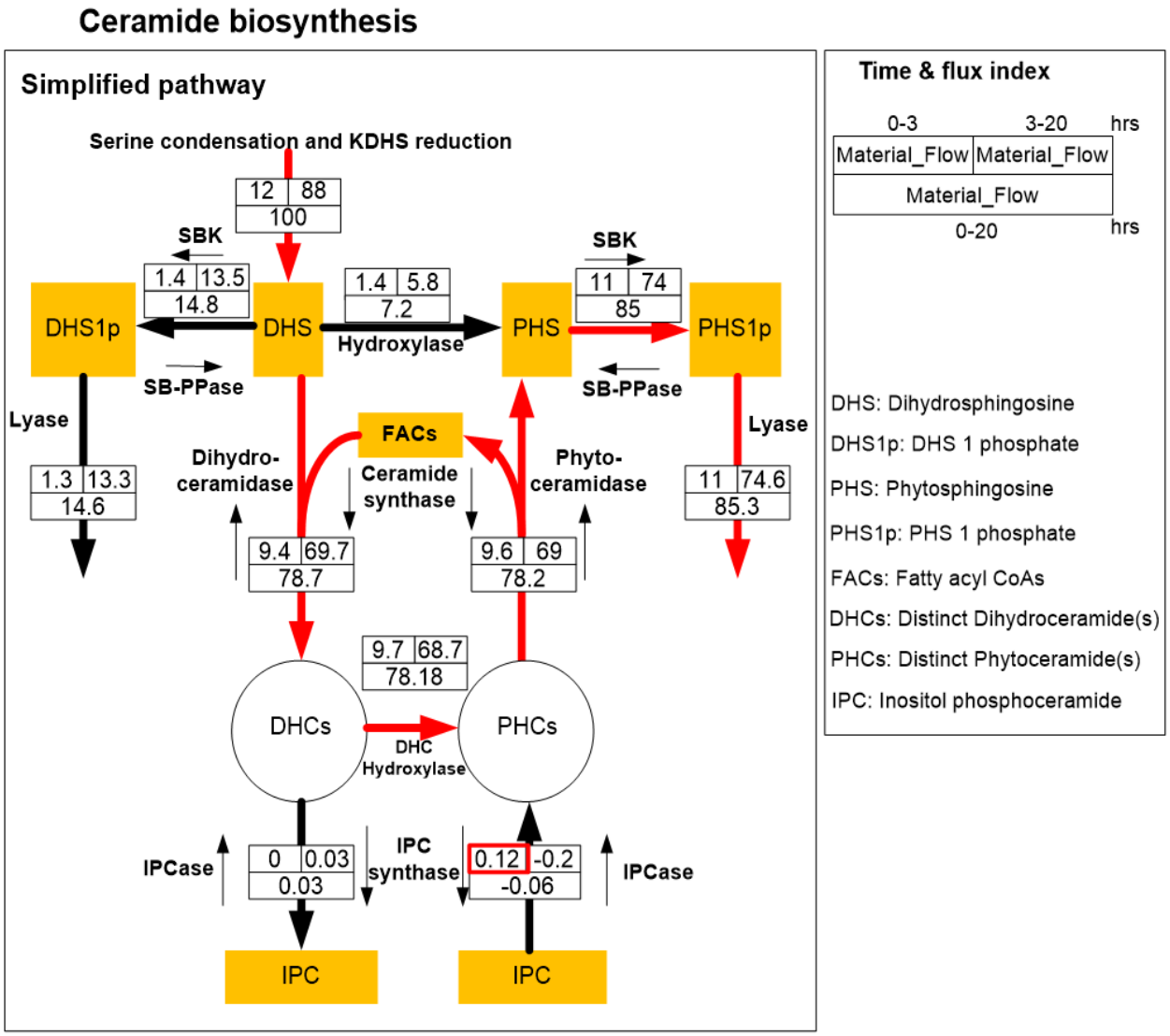


Figure 5.10. Mass flow through the pathway of ceramide biosynthesis (with different DHCs, PHCs and FAs merged into single pools). Red arrows indicate the main flow of material throughout 20 hours of hydroxyurea exposure.

This result raises the question of how the cells manage these fluxes in such a manner that all ceramide species achieve their target levels. Specifically, one must ask: if there is no evidence indicating any flux reversal, how is it possible that distinct saturated or unsaturated fatty acyl CoAs can be channeled toward the accumulation of a key species like C18:1 PHC?

To assess this question, we analyzed the mass flow for saturated and unsaturated fatty acyl CoAs separately. Intriguingly, the flow patterns now are distinctly different. Figure 5.11 shows the mass flow of the pathway for saturated and unsaturated fatty acyl CoAs in the left and right panels, respectively.

As a first result, one notices that ceramides with saturated fatty acyl group are mainly provided through the degradation of complex sphingolipids, whereas ceramides with unsaturated fatty acyl group are obtained from biosynthesis via DHCs.

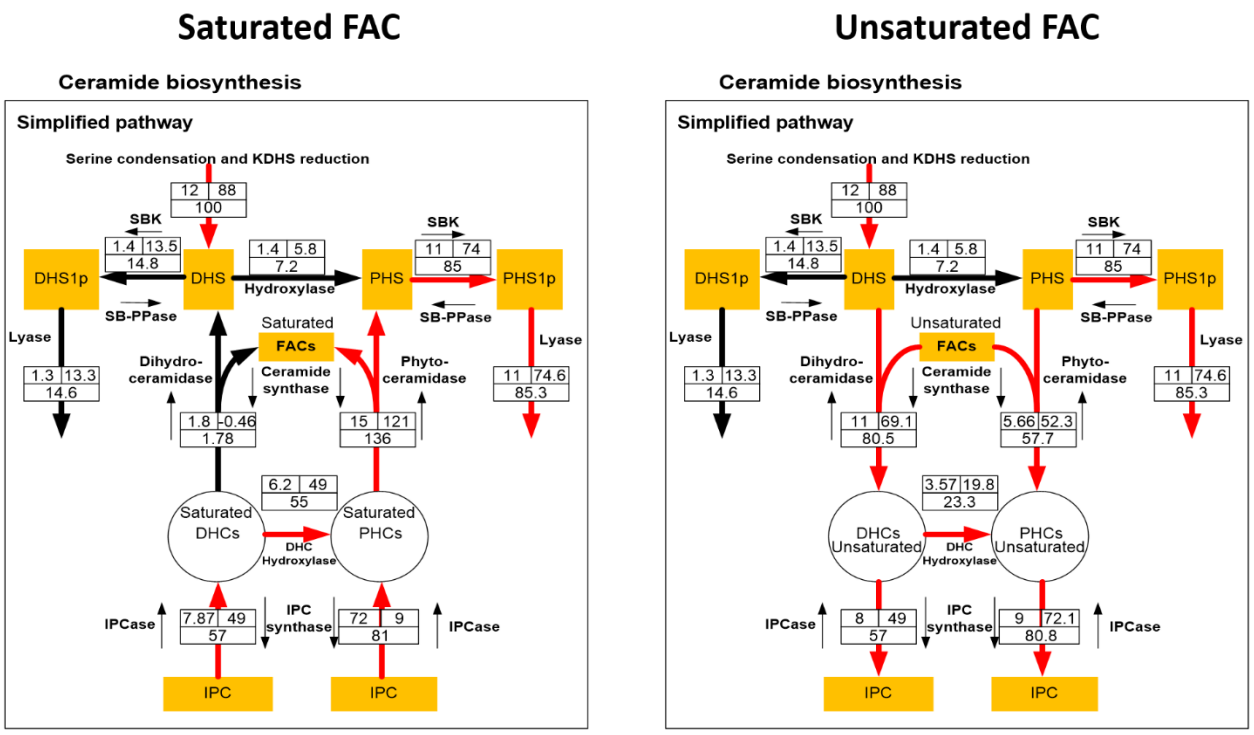


Figure 5.11. Mass flow analysis of the ceramide pathway separated for saturated and unsaturated fatty acyl groups. Red arrows indicate the main flow of mass in each system.

This result in turn suggests that increased Isc1 (IPCase) activity during 20 hours of hydroxyurea treatment affects saturated and unsaturated ceramides differently, which might explain findings regarding Isc1 in the literature. Such a differential effect does not seem unreasonable as saturated and unsaturated ceramides have distinct molecular shapes. According to our analysis, hydroxyurea increases the net flow from IPC to those ceramides containing saturated fatty acyl CoAs, and the PHC pool is being recycled via

PHS, PHS-p, and the lyase reaction. At the same time, Isc1 apparently prevents unsaturated ceramides from being channeled toward IPC. As a consequence, knocking down Isc1 indirectly permits the conversion of PHCs to IPC. PHCs, including C18:1 PHC, eventually become depleted, and cells are no longer tolerant to hydroxyurea and die.

Figure 5.10 can also be interpreted from the perspective of fatty acyl CoAs (FACs). The left panel demonstrates a strong net flux into saturated FACs, while the right panel suggests a strong net efflux of unsaturated FACs. A possible interpretation is the following: under hydroxyurea exposure, saturated FACs are used for the conversion of saturated DHCs into PHCs and can be converted into unsaturated FACs through the desaturase reaction, which exhibits increased activity. Unsaturated FACs can then be used to provide corresponding DHCs and PHCs.

These simulation results should be independently validated, but if they are correct, they provide novel insights into the regulatory mechanisms that govern—or at least contribute to—the control of distinct ceramide species. This regulation task is complicated by the fact that the system contains many metabolites but only a few enzymes, so that alterations in enzyme activities are quite limited when subtle changes of specific metabolites are needed. For example, increasing ceramide synthase activity elevates the rates of 12 reactions toward synthesizing distinct ceramide species in our model and does not by itself allow a change in just one or a few reaction products.

Our simulations suggest that cells manage to achieve hydroxyurea tolerance through the well-coordinated, differential usage of saturated and unsaturated fatty acyl groups. In other words, the substrate affinity of enzymes toward saturated or unsaturated fatty acyl groups seems to be distinct and appears to constitute an additional mode of regulation, which ultimately permits the fine-tuning of a desired ceramide profile.

While the analysis identifies Isc1 as an enzyme that differentiates substrates by saturation state, it is most likely not the only such enzyme. In fact, it seems that ceramide synthase, dihydro- and phyto-ceramidase, IPC synthase, DHC hydroxylase all help establish and maintain the appropriate ceramide profile.

5.3 Methods

5.3.1 Ceramide Time Coarse

Fold changes in ceramide concentrations, relative to the steady state at time 0, were measured in duplicate after 3 and 20 hours of exposure to hydroxyurea. Mean values of the duplicate measurements were interpolated with a smoothing spline function, and this interpolation constituted the baseline levels for the dynamic trends of ceramides. The interpolated mean trends do not reflect any variability in the raw data. Therefore, we allowed for 10% upper and lower boundaries relative to the interpolated baseline trends. The results are represented in Figure 5.11. The dataset we used to fit with our models, was randomly sampled from this 20% range around the interpolated mean values.

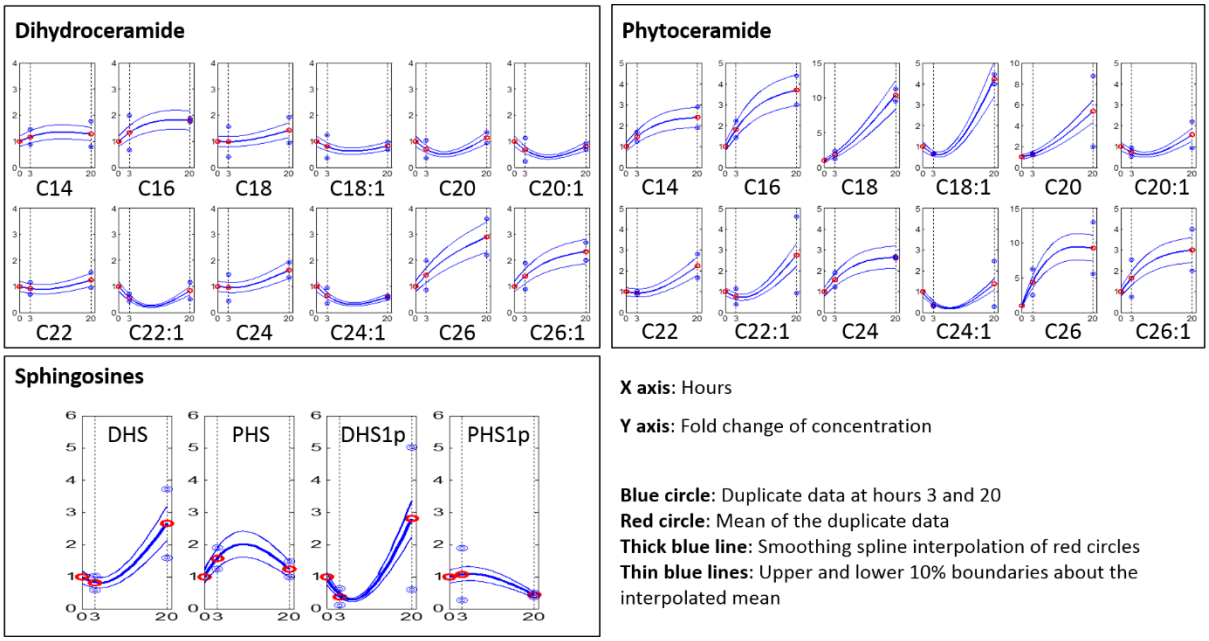


Figure 5.12. Duplicate time coarse data, mean data, interpolation of mean data and 10% variability tolerance bands are marked as blue circles, red circles, a thick blue line, and a thin blue line in each plot.

5.3.2 Mathematical Model

The ceramide pathway model considered here contains 138 reactions (fluxes) and 42 dependent variables (Figure 5.2). The pathway system accounts for many parallel reactions, for which kinetic parameters are not known. Furthermore, absolute ceramide concentrations are not available. To minimize the consequences of this paucity of information, we chose a mass action model to infer coarse trends in enzyme activities.

As an example for the design of this mass action model, consider flux V_1 in Figure 5.2, which represents the reaction from DHS + C16 Fatty acyl CoA to C16 DHC. The mass action model for this reaction is defined as $V_1 = \gamma_1 * [DHS] * [C16FAC]$, where [DHS] and [C16 FAC] represent the concentrations (or fold changes) of DHS and C16 Fatty acyl CoA (relative to their steady states), respectively, and γ_1 denotes the rate constant. Each rate constant is considered here to include the corresponding enzyme activity. This merging of factors is similar to the combination of k_{cat} and the total enzyme concentration in the Michaelis-Menten formalism. In the mass action model resulting from these settings, the steady state flux distribution (and the rate constants) can be estimated and tested quite rapidly.

5.3.3 Piecewise Optimization Approach & Linear Interpolation

Steady-state fluxes and steady-state enzyme activities were calculated with a constrained optimization approach. Using a resampling scheme for "data" from the range surrounding the first data point, averaged corresponding enzyme activities at the steady state were computed. They were tested by entering them into the mass action model and ensuring that they satisfied the steady state. If so, they were used as baseline activities.

Piecewise optimization, similar to the methods in our previous paper [116] and Chapter 4, was used to estimate dynamic trends in enzyme activities with a 1-hour resolution. The main technical difference was that, instead of fitting mean values of duplicate data, we created variability of the "data" by introducing 10% upper and lower bounds to the mean time courses, as discussed above. While this introduction of artificial variability increased the difficulties of fitting data, it constitutes a significant gain in confidence regarding the estimated enzyme activities. A flowchart of the process is presented in Figure 5.13.

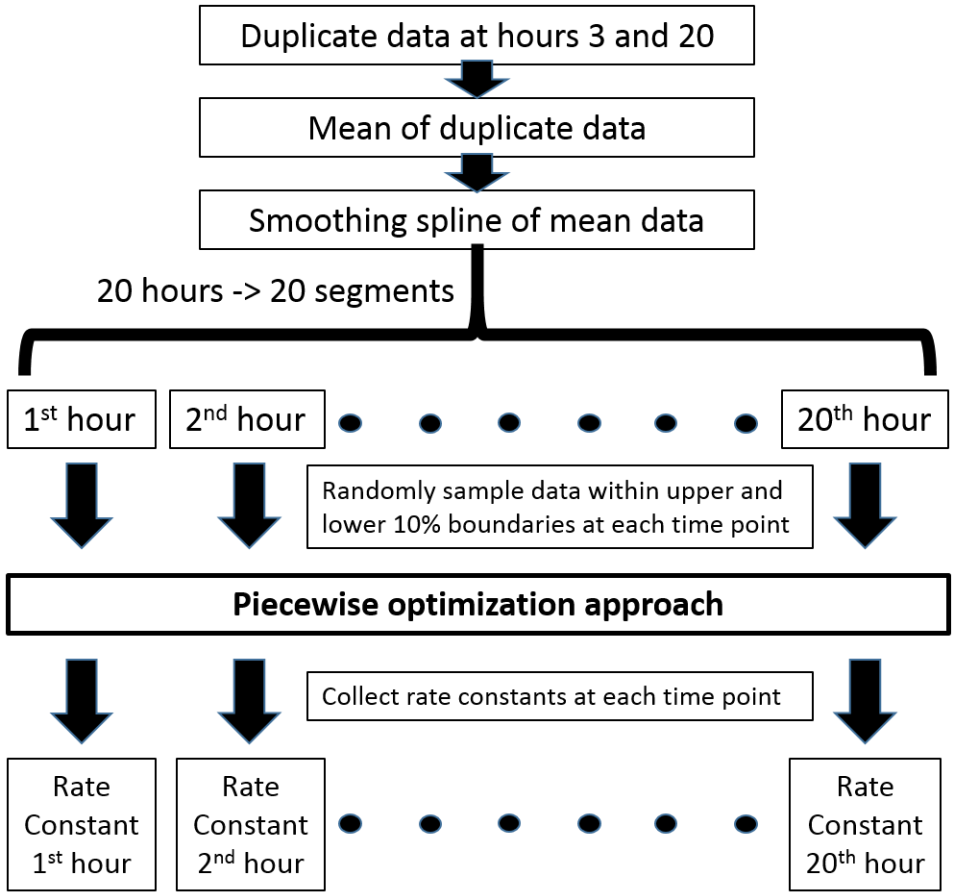


Figure 5.13. Flowchart of the piecewise optimization approach

The results of the piecewise optimization throughout the 20 hours of hydroxyurea exposure were stored for further analysis. It turned out that many dynamic trends in enzyme activities were almost linear or piecewise

linear from 0-3 hours and from 3-20 hours, which may be due to the fact that only three measurement time points are available data. Nonetheless, we utilized this observation and computed linear regression lines from 0-3 hours, 3-20 hours and also 0-20 hours of the estimated enzyme activities and recorded the slopes of these functions as indications of trends.

5.3.4 Dynamic Flux Re-estimation & Mass Flow Analysis

The piecewise optimization returned rate constants in one-hour intervals, but did not directly provide estimations of fluxes, which we needed for the mass flow analysis. Thus, fluxes were estimated as follows. Rate constants were estimated in a stepwise manner for each 1-hour interval. However, the fluxes within each 1-hour time interval are not constant because the substrate concentrations change. To obtain representative flux estimates for each 1-hour interval, we split each interval into 100 segments and calculated 100 sequential flux values per hour. Estimates of fluxes in each hourly interval were then represented as averages of these 100 flux segments.

For the mass flow analysis, we further averaged these fluxes from 0-3 hours, 3-20 hours and 0-20 hours. Mass flow distributions for these time intervals were then estimated through appropriate weighting: [average of fluxes from 0-3 hours]*[3 hours], [average of fluxes from 3-20 hours]*[17 hours] and [average of fluxes from 0-20 hours]*[20 hours]. These mass flows were normalized and displayed in Figure 5.9 and 5.10 of the 5.2 *Results* section.

CHAPTER 6

CONCLUSION

Biological Insights Regarding Heat Stress Responses, Gained by Reverse Engineering

Sphingolipid concentrations, even if measured in dense time series, are not by themselves sufficient to reveal the complex system of regulatory mechanisms that coordinate responses to external stresses. In addition, computational analyses are required that create a quantitative, functional context for these data. As demonstrated in Chapters III, IV and V, my computational approaches have the capability of taking limited metabolic time series data and converting them first into key enzymatic activities and subsequently into a functional explanation of stress responses. Thus, through simulations of very many combinations of activities, and restrictive subsequent filtering of results, it is possible to can gain a robust understanding of how cells process stresses.

Sphingolipids are key responders under a variety of stresses, and the regulatory strategies emerging from my computational analysis suggest a short period of sphingolipid biosynthesis followed by a phase of degradation of complex sphingolipids (Chapter 3). The cellular strategy furthermore seems to involve a short period of ceramide production (Chapter 4), which is well coordinated with distinct rates for long chain and very long chain fatty acyl CoAs groups under heat stress and a long-term production of ceramides (Chapter 5) that distinguishes between saturated and unsaturated fatty acyl CoAs. Under both types of stresses, sphingolipids and ceramides are maintained or controlled by: (1) alterations in the activities of important enzymes; (2) a fine-tuned balance between these enzyme activities and the metabolic fluxes associated with these enzymes; and (3) distinct substrates affinities toward these enzymes. The proposed modeling approaches clearly reveal these features of the response system from metabolic data, which are much easier to measure than fluxes or trends in enzyme activities over time. Beyond sphingolipids, the

approaches I developed are expected to be significant advances for a host of other reverse engineering tasks associated with dynamic biomedical systems.

Contributions to the Field of Modeling

In this work, I have developed several computational approaches. Chapter 2 describes generic, and in a sense comprehensive, strategies for cells or organisms to shift from a normal steady state to a new steady state that is adapted to altered environmental conditions. While demonstrated with the trehalose heat stress response in yeast, the methods described in this chapter are very general. They were fully executed for S-system models, which have the almost unique advantage of permitting straightforward algebraic analyses of steady states. It was also discussed that other systems, such as GMA systems, might be amenable to a similar analysis. One should furthermore expect that Lotka-Volterra models can be analyzed in an analogous fashion. Future work should work out the details of these types of analyses and moreover explore to what degree generic stoichiometric flux system might be analyzed with respect to transitions between steady states.

In Chapter 3, a novel piecewise optimization approach was developed and applied to the role of sphingolipids in heat stress responses; an expanded version of this strategy was used in Chapter 5. This type of stepwise optimization has not been reported in this form in the literature. While applied to specific heat stress responses, the proposed methods appear to be more or less directly applicable to a wide variety of reverse-engineering inference tasks addressing systems that change successively over time and where a global optimized inference throughout an extended time frame is not feasible.

Also in Chapter 4, a novel algorithm was proposed that bridge the methodology of dynamic flux estimation and the imputation of dynamic enzyme activity trends. This algorithm was successfully tested with a detailed analysis of ceramides under heat stress. The experiences gained from these modeling techniques will be beneficial for many types of future ensemble modeling and for the reverse engineering of metabolic—and possibly general biomedical—systems. Taken together, it is clear that the dynamic

responses of biomedical systems are complicated, and that purely computational inferences of their internal, governing mechanisms from global output data are even more challenging. Nonetheless, this work demonstrates that well-designed computational models and custom-tailored techniques can successfully address the inference of complex phenomena like stress responses. While the proposed techniques were demonstrated within the contexts of heat stress and sphingolipids, their applicability appears to be much broader.

Appendix A

Accessary Details of the computational method of Chapter 3

Model Equations

The model equations (Eq. A3.1) were adapted from Alvarez *et al.* [10,11]. One exception was the use of a single representation of IPCase rather than two in the original papers. Furthermore, we expanded the model from the original IPC mechanism, consisting of reversible reactions between the IPC family and ceramides, to a refined mechanism that distinguished between dihydroceramide and phytoceramide forms. To achieve this new mechanism, we included six more variables (X_{8b}, $X_{18b} - X_{22b}$) into the system to represent the IPC family as it reacts with phytoceramide, while leaving the original IPC variables (X_8, $X_{18} - X_{22}$) to represent the IPC family reacting with dihydroceramide. The corresponding equations are given below; please refer to Table A3.1 for definitions of variables.

$$\frac{dX_1}{dt} = \gamma_{11} X_{12}^{f_{1,12,1}} X_{13}^{f_{1,13,2}} X_{57}^{f_{1,57,3}} - \gamma_{12} X_1^{f_{1,1,4}} X_{27}^{f_{1,27,5}}$$

$$\frac{dX_2}{dt} = \gamma_{21} X_1^{f_{2,1,1}} X_{27}^{f_{2,27,2}} + \gamma_{22} X_3^{f_{2,3,3}} X_{29}^{f_{2,29,4}} + \gamma_{23} X_4^{f_{2,4,5}} X_{41}^{f_{2,41,6}} - \gamma_{24} X_2^{f_{2,2,7}} X_{23}^{f_{2,23,8}} X_{34}^{f_{2,34,9}} - \gamma_{25} X_2^{f_{2,2,10}} X_{28}^{f_{2,28,11}} X_{36}^{f_{2,36,12}} - \gamma_{26} X_2^{f_{2,2,13}} X_{54}^{f_{2,54,14}}$$

$$\frac{dX_3}{dt} = \gamma_{31} X_2^{f_{3,2,1}} X_{23}^{f_{3,23,2}} X_{34}^{f_{3,34,3}} + \gamma_{32} X_8^{f_{3,8,4}} X_{51}^{f_{3,51,5}} + \gamma_{33} X_{18}^{f_{3,18,6}} X_{51}^{f_{3,51,7}} + \gamma_{34} X_{19}^{f_{3,19,8}} X_{51}^{f_{3,51,9}} - \gamma_{35} X_3^{f_{3,3,10}} X_{29}^{f_{3,29,11}} - \gamma_{36} X_3^{f_{3,3,12}} X_{54}^{f_{3,54,13}} - \gamma_{37} X_2^{f_{3,2,14}} X_3^{f_{3,3,15}} X_5^{f_{3,5,16}} X_{15}^{f_{3,15,17}} X_{33}^{f_{3,33,18}}$$

$$\frac{dX_4}{dt} = \gamma_{41} X_2^{f_{4,2,1}} X_{28}^{f_{4,28,2}} X_{36}^{f_{4,36,3}} - \gamma_{42} X_4^{f_{4,4,4}} X_{41}^{f_{4,41,5}} - \gamma_{43} X_4^{f_{4,4,6}} X_{50}^{f_{4,50,7}}$$

$$\frac{dX_5}{dt} = \gamma_{51} X_2^{f_{5,2,1}} X_{54}^{f_{5,54,2}} + \gamma_{52} X_6^{f_{5,6,3}} X_{41}^{f_{5,41,4}} + \gamma_{53} X_7^{f_{5,7,5}} X_{53}^{f_{5,53,6}} - \gamma_{54} X_5^{f_{5,5,7}} X_{23}^{f_{5,23,8}} X_{34}^{f_{5,34,9}} - \gamma_{55} X_5^{f_{5,5,10}} X_{28}^{f_{5,28,11}} X_{36}^{f_{5,36,12}}$$

$$\frac{dX_6}{dt} = \gamma_{41} X_2^{f_{4,2,1}} X_{28}^{f_{4,28,2}} X_{36}^{f_{4,36,3}} - \gamma_{42} X_4^{f_{4,4,4}} X_{41}^{f_{4,41,5}} - \gamma_{43} X_4^{f_{4,4,6}} X_{50}^{f_{4,50,7}}$$

$$\frac{dX_7}{dt} = \gamma_{71}X_2^{f_{7,2,1}}X_{23}^{f_{7,23,2}}X_{34}^{f_{7,34,3}} + \gamma_{72}X_{8b}^{f_{7,8b,4}}X_{51}^{f_{7,51,5}} + \gamma_{73}X_3^{f_{7,3,6}}X_{54}^{f_{7,54,7}} + \gamma_{74}X_{18b}^{f_{7,18b,8}}X_{51}^{f_{7,51,9}}$$
$$+ \gamma_{75}X_{19b}^{f_{7,19b,10}}X_{51}^{f_{7,51,11}} - \gamma_{76}X_7^{f_{7,7,12}}X_{53}^{f_{7,53,13}} - \gamma_{77}X_2^{f_{7,2,14}}X_5^{f_{7,5,15}}X_7^{f_{7,7,16}}X_{15}^{f_{7,15,17}}X_{33}^{f_{7,33,18}}$$
$$- \gamma_{78}X_7^{f_{7,7,19}}X_{43}^{f_{7,43,20}}$$

$$\frac{dX_8}{dt} = \gamma_{82}X_2^{f_{8,2,6}}X_3^{f_{8,3,7}}X_5^{f_{8,5,8}}X_{15}^{f_{8,15,9}}X_{33}^{f_{8,33,10}} + \gamma_{83}X_{20}^{f_{8,20,11}} - \gamma_{85}X_8^{f_{8,8,14}}X_{35}^{f_{8,35,15}} - \gamma_{86}X_8^{f_{8,8,16}}X_{51}^{f_{8,51,17}}$$
$$- \gamma_{87}X_8^{f_{8,8,18}}$$

$$\frac{dX_9}{dt} = \gamma_{91}X_{11}^{f_{9,11,1}}X_{40}^{f_{9,40,2}} - \gamma_{92}X_2^{f_{9,2,3}}X_5^{f_{9,5,4}}X_9^{f_{9,9,5}}X_{11}^{f_{9,11,6}}X_{13}^{f_{9,13,7}}X_{14}^{f_{9,14,8}}X_{15}^{f_{9,15,9}}X_{16}^{f_{9,16,10}}X_{38}^{f_{9,38,11}}$$
$$- \gamma_{93}X_9^{f_{9,9,12}}X_{16}^{f_{9,16,13}}X_{26}^{f_{9,26,14}}$$

$$\frac{dX_{10}}{dt} = \gamma_{10,1}X_2^{f_{10,2,1}}X_5^{f_{10,5,2}}X_9^{f_{10,9,3}}X_{11}^{f_{10,11,4}}X_{13}^{f_{10,13,5}}X_{14}^{f_{10,14,6}}X_{15}^{f_{10,15,7}}X_{16}^{f_{10,16,8}}X_{38}^{f_{10,38,9}} - \gamma_{10,2}X_{10}^{f_{10,10,10}}X_{56}^{f_{10,56,11}}$$

$$\frac{dX_{11}}{dt} = \gamma_{11,1}X_{10}^{f_{11,10,1}}X_{12}^{f_{11,12,2}}X_{49}^{f_{11,49,3}} - \gamma_{11,2}X_2^{f_{11,2,4}}X_5^{f_{11,5,5}}X_9^{f_{11,9,6}}X_{11}^{f_{11,11,7}}X_{15}^{f_{11,15,8}}X_{39}^{f_{11,39,9}}$$
$$- \gamma_{11,3}X_{11}^{f_{11,11,10}}X_{40}^{f_{11,40,11}}$$

$$\frac{dX_{12}}{dt} = \gamma_{12,1}X_{30}^{f_{12,30,1}}X_{58}^{f_{12,58,2}} + \gamma_{12,2}X_{24}^{f_{12,24,3}}X_{25}^{f_{12,25,4}}X_{52}^{f_{12,52,5}} + \gamma_{12,3}X_4^{f_{12,4,6}}X_{50}^{f_{12,50,7}} + \gamma_{12,4}X_6^{f_{12,6,8}}X_{50}^{f_{12,50,9}}$$
$$- \gamma_{12,5}X_{12}^{f_{12,12,10}}X_{13}^{f_{12,13,11}}X_{57}^{f_{12,57,12}} - \gamma_{12,6}X_{10}^{f_{12,10,13}}X_{12}^{f_{12,12,14}}X_{49}^{f_{12,49,15}}$$
$$- \gamma_{12,7}X_{12}^{f_{12,12,16}}X_{48}^{f_{12,48,17}} - \gamma_{12,8}X_{12}^{f_{12,12,18}}X_{24}^{f_{12,24,19}}X_{59}^{f_{12,59,20}}$$

$$\frac{dX_{13}}{dt} = \gamma_{13,1}X_{31}^{f_{13,31,1}}X_{37}^{f_{13,37,2}} + \gamma_{13,2}X_{65}^{f_{13,65,3}}X_{66}^{f_{13,66,4}}$$
$$- \gamma_{13,3}X_2^{f_{13,2,5}}X_5^{f_{13,5,6}}X_9^{f_{13,9,7}}X_{11}^{f_{13,5,8}}X_{13}^{f_{13,13,9}}X_{14}^{f_{13,14,10}}X_{15}^{f_{13,15,11}}X_{16}^{f_{13,16,12}}X_{38}^{f_{13,38,13}}$$
$$- \gamma_{13,4}X_{12}^{f_{13,12,14}}X_{13}^{f_{13,13,15}}X_{57}^{f_{13,57,16}} - \gamma_{13,5}X_{13}^{f_{13,13,17}}X_{32}^{f_{13,32,18}}$$

$$\frac{dX_{14}}{dt} = \gamma_{14,1}X_2^{f_{14,2,1}}X_5^{f_{14,5,2}}X_9^{f_{14,9,3}}X_{11}^{f_{14,11,4}}X_{15}^{f_{14,15,5}}X_{39}^{f_{14,39,6}} + \gamma_{14,2}X_2^{f_{14,2,7}}X_5^{f_{14,5,8}}X_7^{f_{14,7,9}}X_{15}^{f_{14,15,10}}X_{33}^{f_{14,33,11}}$$
$$+ \gamma_{14,3}X_2^{f_{14,2,12}}X_3^{f_{14,3,13}}X_5^{f_{14,5,14}}X_{15}^{f_{14,15,15}}X_{33}^{f_{14,33,16}} + \gamma_{14,4}X_{15}^{f_{14,15,17}}X_{18}^{f_{14,18,18}}X_{55}^{f_{14,55,19}}$$
$$- \gamma_{14,5}X_{14}^{f_{14,14,20}}X_{17}^{f_{14,17,21}}X_{45}^{f_{14,45,22}} - \gamma_{14,6}X_{14}^{f_{14,14,23}}X_{42}^{f_{14,42,24}}$$

$$\frac{dX_{15}}{dt} = \gamma_{15,1}X_9^{f_{15,9,1}}X_{16}^{f_{15,16,2}}X_{26}^{f_{15,26,3}} - \gamma_{15,2}X_2^{f_{15,2,4}}X_5^{f_{15,5,5}}X_7^{f_{15,7,6}}X_{15}^{f_{15,15,7}}X_{33}^{f_{15,33,8}}$$
$$- \gamma_{15,3}X_2^{f_{15,2,9}}X_3^{f_{15,3,10}}X_5^{f_{15,5,11}}X_{15}^{f_{15,15,12}}X_{33}^{f_{15,33,13}} - \gamma_{15,4}X_{15}^{f_{15,15,14}}X_{28}^{f_{15,28,15}}X_{44}^{f_{15,44,16}}$$
$$- \gamma_{15,5}X_{15}^{f_{15,15,17}}X_{18}^{f_{15,18,18}}X_{55}^{f_{15,55,19}}$$

$$\frac{dX_{16}}{dt} = \gamma_{16,1}X_{46}^{f_{16,46,1}}X_{47}^{f_{16,47,2}} - \gamma_{16,2}X_9^{f_{16,9,3}}X_{16}^{f_{16,16,4}}X_{26}^{f_{16,26,5}}$$

$$\frac{dX_{17}}{dt} = \gamma_{17,1} X_4^{f_{17,4,1}} X_{50}^{f_{17,50,2}} + \gamma_{17,2} X_6^{f_{17,6,3}} X_{50}^{f_{17,50,4}} - \gamma_{17,3} X_{14}^{f_{17,14,5}} X_{17}^{f_{17,17,6}} X_{45}^{f_{17,45,7}}$$

$$\frac{dX_{18}}{dt} = \gamma_{18,1} X_8^{f_{18,8,1}} X_{35}^{f_{18,35,2}} + \gamma_{18,2} X_{21}^{f_{18,21,3}} - \gamma_{18,3} X_{18}^{f_{18,18,4}} X_{51}^{f_{18,51,5}} - \gamma_{18,5} X_{15}^{f_{18,15,8}} X_{18}^{f_{18,18,9}} X_{55}^{f_{18,55,10}} - \gamma_{18,6} X_{18}^{f_{18,18,11}}$$

$$\frac{dX_{19}}{dt} = \gamma_{19,1} X_{15}^{f_{19,15,1}} X_{18}^{f_{19,18,2}} X_{55}^{f_{19,55,3}} + \gamma_{19,2} X_{22}^{f_{19,22,4}} - \gamma_{19,3} X_{19}^{f_{19,19,5}} X_{51}^{f_{19,51,6}} - \gamma_{19,5} X_{19}^{f_{19,19,9}}$$

$$\frac{dX_{20}}{dt} = \gamma_{20,1} X_8^{f_{20,8,1}} - \gamma_{20,2} X_{20}^{f_{20,20,2}}$$

$$\frac{dX_{21}}{dt} = \gamma_{21,1} X_{18}^{f_{21,18,1}} - \gamma_{21,2} X_{21}^{f_{21,21,2}}$$

$$\frac{dX_{22}}{dt} = \gamma_{22,1} X_{19}^{f_{22,19,1}} - \gamma_{22,2} X_{22}^{f_{22,22,2}}$$

$$\frac{dX_{23}}{dt} = \gamma_{23,1} X_{12}^{f_{23,12,1}} X_{24}^{f_{23,24,2}} X_{59}^{f_{23,59,3}} - \gamma_{23,2} X_2^{f_{23,2,4}} X_{23}^{f_{23,23,5}} X_{34}^{f_{23,34,6}} - \gamma_{23,3} X_5^{f_{23,5,7}} X_{23}^{f_{23,23,8}} X_{34}^{f_{23,34,9}}$$

$$\frac{dX_{24}}{dt} = \gamma_{24,1} X_{12}^{f_{24,12,1}} X_{23}^{f_{24,23,2}} X_{25}^{f_{24,25,3}} X_{28}^{f_{24,28,4}} X_{60}^{f_{24,60,5}} - \gamma_{24,2} X_{12}^{f_{24,12,6}} X_{24}^{f_{24,24,7}} X_{59}^{f_{24,59,8}} - \gamma_{24,3} X_{24}^{f_{24,24,9}} X_{25}^{f_{24,25,10}} X_{52}^{f_{24,52,11}}$$

$$\frac{dX_{25}}{dt} = \gamma_{25,1} X_{12}^{f_{25,12,1}} X_{28}^{f_{25,28,2}} X_{61}^{f_{25,61,3}} X_{62}^{f_{25,62,4}} X_{63}^{f_{25,63,5}} - \gamma_{25,2} X_{12}^{f_{25,12,6}} X_{23}^{f_{25,23,7}} X_{25}^{f_{25,25,8}} X_{28}^{f_{25,28,9}} X_{60}^{f_{25,60,10}} - \gamma_{25,3} X_{24}^{f_{25,24,11}} X_{25}^{f_{25,25,12}} X_{52}^{f_{25,52,13}}$$

$$\frac{dX_{8b}}{dt} = \gamma_{26,1} X_2^{f_{26,2,1}} X_5^{f_{26,5,2}} X_7^{f_{26,7,3}} X_{15}^{f_{26,15,4}} X_{33}^{f_{26,33,5}} + \gamma_{26,2} X_{20b}^{f_{26,20b,6}} - \gamma_{26,3} X_{8b}^{f_{26,8b,7}} X_{35}^{f_{26,35,8}} - \gamma_{26,4} X_{8b}^{f_{26,8b,9}} X_{51}^{f_{26,51,10}} - \gamma_{26,5} X_{8b}^{f_{26,8b,11}}$$

$$\frac{dX_{18b}}{dt} = \gamma_{27,1} X_{8b}^{f_{27,8b,1}} X_{35}^{f_{27,35,2}} + \gamma_{27,2} X_{21b}^{f_{27,21b,3}} - \gamma_{18,3} X_{18b}^{f_{27,18b,4}} X_{51}^{f_{27,51,5}} - \gamma_{18,4} X_{15}^{f_{27,15,6}} X_{18b}^{f_{27,18b,7}} X_{55}^{f_{27,55,8}} - \gamma_{18,5} X_{18b}^{f_{27,18b,9}}$$

$$\frac{dX_{19b}}{dt} = \gamma_{28,1} X_{15}^{f_{28,15,1}} X_{18b}^{f_{28,18b,2}} X_{55}^{f_{28,55,3}} + \gamma_{28,2} X_{22b}^{f_{28,22b,4}} - \gamma_{28,3} X_{19b}^{f_{28,19b,5}} X_{51}^{f_{28,51,6}} - \gamma_{28,4} X_{19b}^{f_{28,19b,7}}$$

$$\frac{dX_{20b}}{dt} = \gamma_{29,1} X_{8b}^{f_{29,8b,1}} - \gamma_{29,2} X_{20b}^{f_{29,20b,2}}$$

$$\frac{dX_{21b}}{dt} = \gamma_{30,1} X_{18b}^{f_{30,18b,1}} - \gamma_{30,2} X_{21b}^{f_{30,21b,2}}$$

$$\frac{dX_{22b}}{dt} = \gamma_{31,1} X_{19b}^{f_{31,19b,1}} - \gamma_{31,2} X_{22b}^{f_{31,22b,2}} \quad \text{(A3.1)}$$

Table A3.1: Metabolites, Enzymes, Abbreviations, and Variable Names

Metabolites and their Representations in the Computational Analysis			
X_1	3-Keto-Dihydrosphingosine (KDHS)	X_{17}	Cytidine Diphosphate-Ethanolamine (CDP-Eth)
X_2	Dihydrosphingosine (DHS)	X_{18}, X_{18b}	Mannosylinositol Phosphorylceramide (MIPC-g) from DHC or PHC, respectively
X_3	Dihydroceramide (Dihydro-C)	X_{19}, X_{19b}	Mannosyldiinositol Phosphorylceramide ($M(IP)_2C$-g) from DHC or PHC, respectively
X_4	Dihydrosphingosine-1-P (DHS-P)	X_{20}, X_{20b}	Plasma Membrane Inositol Phosphorylceramide (IPC-m) from DHC or PHC, respectively
X_5	Phytosphingosine (PHS)	X_{21}, X_{21b}	Plasma Membrane Mannosylinositol Phosphorylceramide (MIPC-m) from DHC or PHC, respectively
X_6	Phytosphingosine-1-P (PHS-P)	X_{22}, X_{22b}	Plasma Membrane Mannosyldiinositol Phosphorylceramide ($M(IP)_2C$-m) from DHC or PHC, respectively
X_7	Phytoceramide (Phyto-C)	X_{23}	Very Long Chain Fatty Acid (C_{26}-CoA)
X_8, X_{8b}	Inositol Phosphorylceramide (IPC-g) from DHC or PHC, respectively	X_{24}	Malonyl-CoA (Mal-CoA)
X_9	CDP-Diacylglycerol (CDP-DAG)	X_{25}	Acetyl-CoA (Ac-CoA)
X_{10}	Phosphatidylserine (PS)	X_{28}	Adenosime-5'-Triphosphate (ATP)
X_{11}	Phosphatidic Acid (PA)	X_{37}	3-Phosphoserine (3-P-Serine)
X_{12}	Palmitoyl-CoA (Pal-CoA)	X_{47}	Glucose-6-P (G6P)
X_{13}	Serine	X_{58}	Palmitate
X_{14}	Sn-1,2-Diacylglycerol (DAG)	X_{61}	CoA
X_{15}	Phosphatidylinositol (PI)	X_{62}	Acetate
X_{16}	Inositol (I)		

Table A3.1 Continued

Enzymes and their Representations in the Computational Analysis			
X_{26}	Phosphatidylinositol Synthase (PI Synthase)	X_{45}	DG-Ethanolamine Phosphotransferase (EthPT)
X_{27}	3-Ketodihydrosphingosine Reductase (KDHS Reductase)	X_{46}	Inositol-1-P Synthase (I-1-P-Synth.)
X_{29}	Dihydroceramide Alkaline Ceramidase (Dihydro-CDase)	X_{48}	Acyl-CoA-Binding Protein (ACBP)
X_{30}	Palmitoyl Transport & Palmitoyl-CoA Synthase (Transp./Palmitoyl CoA Synthase)	X_{49}	Glycerol-3-Phosphate Acyltransferase (G3P Acyltransferase)
X_{31}	Phosphoserine-Phosphatase (P-Serine-PPase)	X_{50}	Sphingosine-Phosphate Lyase (Lyase)
X_{32}	Serine Hydroxymethyl Transferase (SHMT)	X_{51}	Inositol Phosphosphingolipid Phospholipase C (IPCase)
X_{33}	Inositol Phosphorylceramide Synthase (IPC Synthase)	X_{52}	Fatty Acid Synthetase (FAS)
X_{34}	Ceramide Synthase (Cer Synthase)	X_{53}	Phytoceramide Alkaline Ceramidase (Phyto-CDase)
X_{35}	Mannosyl Inositol Phosphoceramide Synthase (MIPC Synthase)	X_{54}	4-Hydroxylase (Hydroxylase; SYR2p-SUR2p)
X_{36}	Sphingoid Base Kinase	X_{55}	Mannosyldiinositol Phosphorylceramide Synthase ($M(IP)_2C$ Synthase)
X_{38}	Phosphatidylserine Synthase (PS Synthase)	X_{56}	Phosphatidylserine Decarboxylase (PS Decarboxylase)
X_{39}	Phosphatidate Phosphatase (PA-PPase)	X_{57}	Serine Palmitoyltransferase (SPT)
X_{40}	CDP-Diacylglycerol Synthase (CDP-DAG Synthase)	X_{59}	Very Long Chain Fatty Acid Synthase / Elongase (ELO1p)
X_{41}	Sphingoid-1-phosphate Phosphatase (SB-PPase)	X_{60}	Acetyl-Coenzyme A Carboxylase (ACCp)
X_{42}	DG-Choline Phosphotransferase (ChoPT)	X_{63}	Acetyl-Coenzyme A Synthetase (ACSp)
X_{43}	GPI Remodelase (Remodeling)	X_{65}	Not yet identified
X_{44}	Phosphoinositol Kinase (PI Kinase)	X_{66}	Not yet identified

Optimization Procedure

The optimization strategy of our approach can be broken down into several steps:

1. Fix rate constants and kinetic orders in the GMA model, as reported in[10].
2. Set upper and lower bounds for changes in enzyme activities.
3. Acquire smoothed heat stress metabolite measurements at time point *t*, where $t = 1 \ldots 30$.
4. At time point *t*, minimize the norm between the smoothed data and the simulation results of the GMA model with appropriate weights, as indicated in Eq. (A2), that give each metabolite equal importance. For the case $t = 1$, assign as initial guess the normal baseline value. For all other time points, assign the initial guess of each independent variable (enzyme) to the corresponding value in the previous solution. Thus, at each time point execute the following minimization:

$$\min_{XI(i,t),i=26,\ldots,66} \left(\sum_{i=2}^{7} \left(\frac{1}{XD_0(i)}\right)\left(XD_{Data}(i,t) - XD_{GMA}(i,t)\right)^2 + \sum_{i=1,8,..,25} 0.001\left(\frac{1}{XD_0(i)}\right)\left(XD_0(i) - XD_{GMA}(i,t)\right)^2\right) \quad \text{(A2)}$$

$XI(i,t)$: Independent variables indexed by $i = 26,\ldots 66$ at time t

$XD_0(i)$: Steady state of i^{th} dependent variables

$XD_{Data}(i,t)$: Heat stress data of i^{th} dependent variable at time t

$XD_{GMA}(i,t)$: Model calculation of i^{th} dependent variable at time t

5. Check the GMA simulation results for each iteration. If the GMA model produces negative or imaginary values for any of the dependent variables, then randomize the initial guess of the independent variables and return to step 4. Continue with step 5 until the GMA simulation produces reasonable (positive) values.
6. Collect the solutions of independent variables (enzymes) for the given time point. If $t = 30$, terminate.
7. Execute this iteration many (4144) times.

We randomly sampled the initial states (at the one-minute time point after the beginning of heat stress) of enzymatic activities in four-fold ranges (two-fold up and down) with respect to the normal steady-state values of enzyme activities. For example, the steady state for ceramide synthase is 1.65e-5 μM/min/mg, so that the initial guess for this enzyme was sampled from the range [0.825e-5, 3.3e-5]. This four-fold range only refers to the initial point. It is to be considered rather wide, because only one minute earlier the system had resided at its nominal steady state (under optimal temperature conditions). It would therefore seem

unrealistic to allow for, say, 10 or 100-fold changes. Once the system was initialized, the enzyme activities were allowed to vary much further from their optimal activity levels, namely within a 12-fold range.

This setting of 12-fold activity ranges was inspired by experimental data showing that yeast seems to respond to stress by changing many enzyme activities moderately, rather than changing a few key enzymes very strongly. At least this strategy was observed in the sphingolipid response to the diauxic shift in yeast [9]. In addition to this heuristic rationale, extensive preliminary testing suggested upper and lower bounds for all enzyme activities of about 6 times and 1/6 times the baseline levels. Modest variations in these bounds (to 10 and 1/10) were not influential, whereas bounds selected too small (2 and 1/2) did not allow enough flexibility and led to inferior minimization results, while bounds selected much larger (100 and 1/100) created solutions that appeared to be unrealistic.

For the minimization we used the Active Set Algorithm implemented in Matlab. This method is a local optimization algorithm that is based on Lagrange multipliers. While it is mathematically possible that a local algorithm would miss other acceptable solutions, it seems in our case biologically reasonable to search for enzyme profiles in moderately wide neighborhoods of their normal activity states, which can be expected to correspond to the basin of attraction of the local algorithm. We preferred a local algorithm over one of the evolutionary algorithms, because the latter, while excellent for global searches, sometimes have problems identifying good solutions within moderate ranges.

Log_2 Representation of Trends in Enzyme Activities

In order to provide greater resolution for reduced enzyme activities, Figures A3.1 to A3.6 show all simulation results with a $\log_2$ scale. These figures correspond to Figures. 3.4 to 3.9 in the main text.

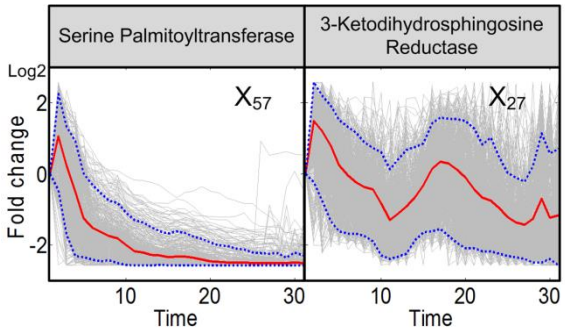


FIGURE A3.1. Trends in activities of enzyme at the entry point of sphingolipid biosynthesis. Serine palmitoyltransferase and 3-KDHS reductase are enzymes responsible for the production and degradation of 3-KDHS, which is the key initial metabolite of sphingolipid biosynthesis. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results. The figure corresponds to Figure 3.4 of the main text.

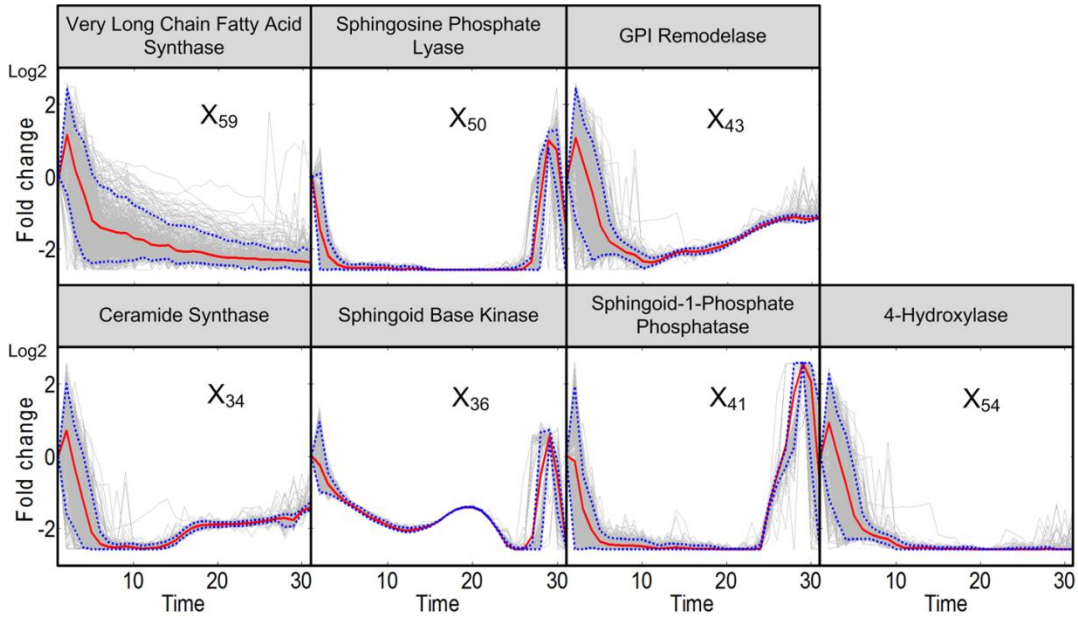


FIGURE A3.2. Trends in activities of enzymes in the core region of sphingolipid metabolism. After an initial spike, all enzyme activities in this region are reduced to almost nil. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results. The figure corresponds to Figure 3.5 of the main text.

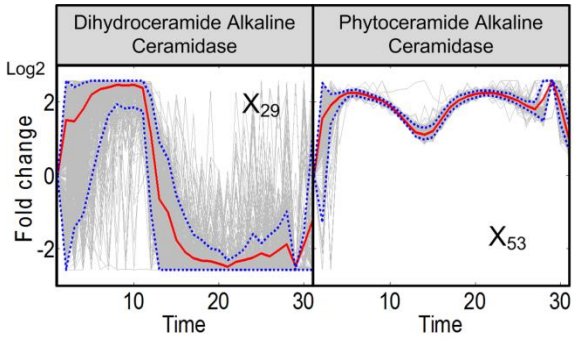


FIGURE A3.3. Trends in activities of the two alkaline ceramidases. Dihydroceramide alkaline ceramidase and phytoceramide alkaline ceramidase, which convert the ceramide form into sphingosines, exhibit distinct activity patterns. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results. The figure corresponds to Figure 3.6 of the main text.

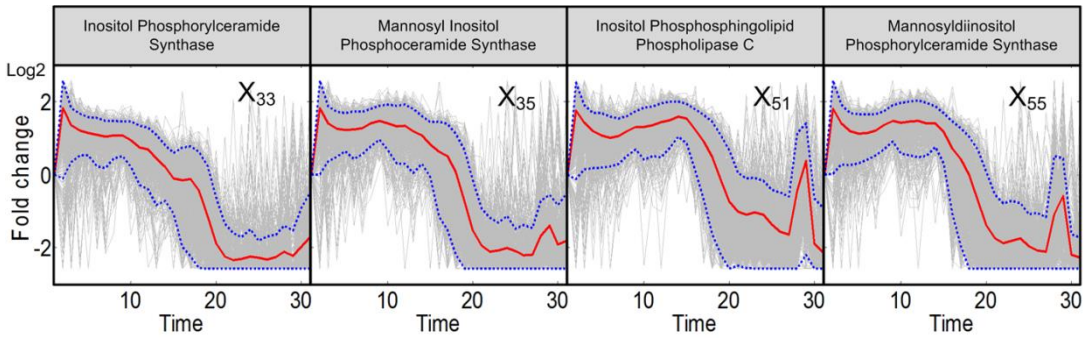


FIGURE A3.4. Trends in activities of enzymes associated with complex sphingolipids. Enzymes interconverting complex sphingolipids are at first hyper-active, but tend to lose most activity between 20 and 30 minutes. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results. The figure corresponds to Figure 3.7 of the main text.

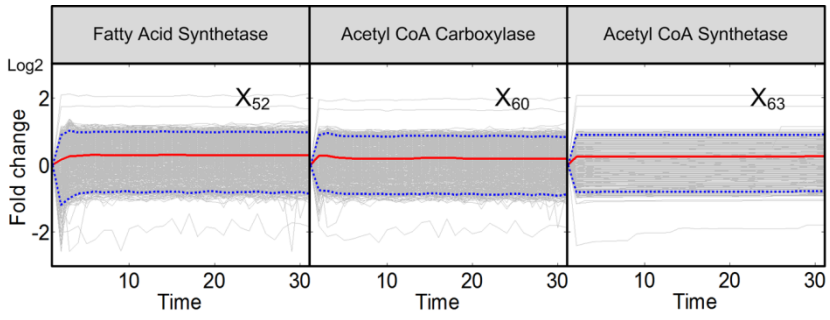


FIGURE A3.5. Trends in activities of enzymes associated with fatty acid CoA. The enzymes shown here are responsible for CoA enlongation. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are ensemble averages, and dotted blue lines enclose 95% of the results. The figure corresponds to Figure 3.8 of the main text.

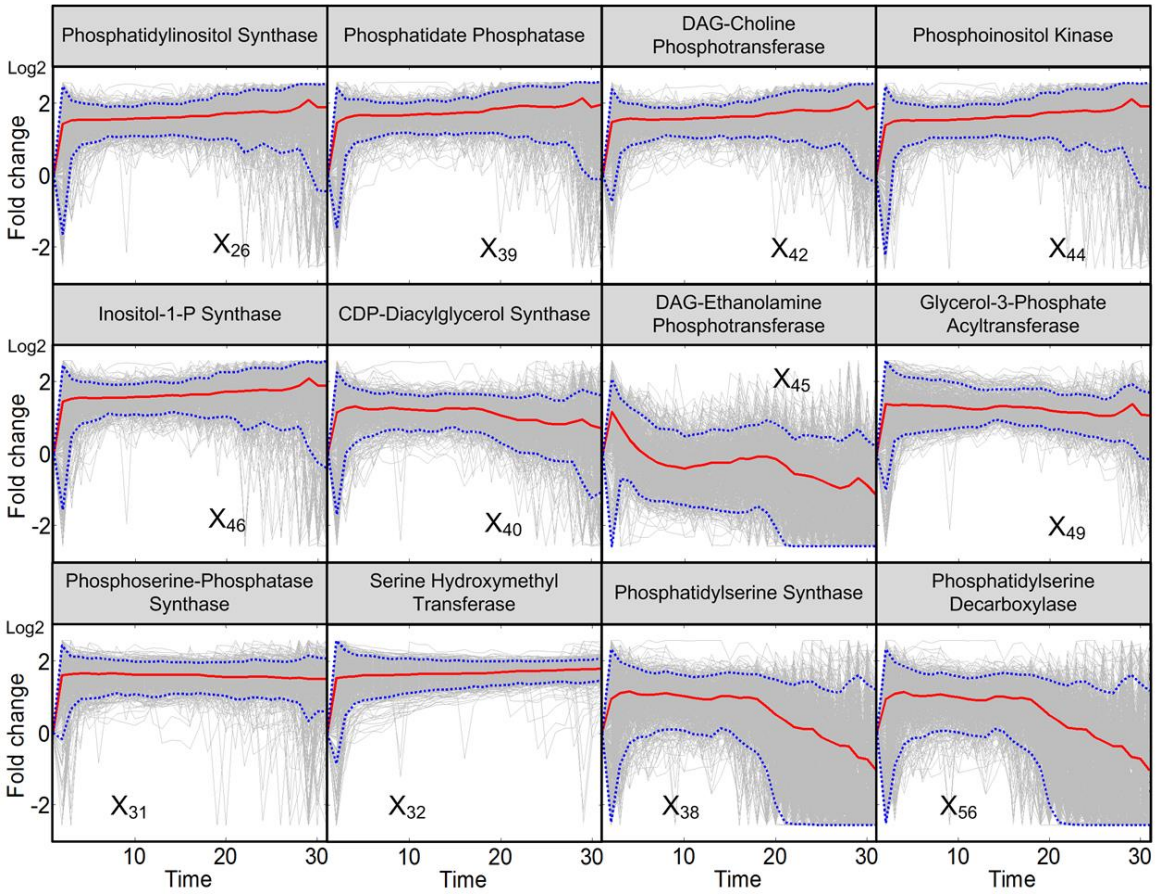


FIGURE A3.6. Trends in the remaining enzyme activities. Activities of enzymes at the periphery of the pathway system are not identifiable, mainly due to insufficient information and the fact that these enzymes are also involved in other pathways. Enzymes in two upper panels are related to the phospholipid metabolism and enzymes in the lower panel are related to serine metabolism. Grey lines are results of 2,000 individual iterations in the large-scale simulation. Red lines are averages, and dotted blue lines enclose 95% of the results. The figure corresponds to Figure 3.9 of the main text.

Table of Trends in Enzyme Activities

Figure 3.10 of the main text summarizes the results in earlier figures in a visual manner. Table A3.2 shows a different representation of the same results.

Table A3.2: Summary of Identifiable Dynamic Changes in Enzyme Activities in Response to Heat Stress. All enzyme activities initially climb to different degrees, presumably due to the Arrhenius effect. Subsequently, the trends are strikingly different. Numbers represent approximate fold changes in activities, while arrows and colors indicate the direction of change.

Enzyme	Time / Var.	0-5		5-10	10-15		15-20	20-25	25-30	
3-keto-dihydrosphingosine reductase	X_{27}	↑ 2.8		↓ 1.1	↑ 1.2		↓ 1	↓ 0.5	↓ 0.4	
Dihydroceramide aklaline ceramidase	X_{29}	↑ 5.5			↔ 0.3				↑ 0.5	
Inositol phosphorylceramide synthase	X_{33}	↑ 3.5		↓ 1			↓0.3		↑ 0.4	
Ceramide synthase	X_{34}	0-2.5	2.5-5	↔ 0.3			↑ 0.4			
		↑ 1.8	↓ 0.3							
Mannosyl inositol phosphoceramide synthase	X_{35}	0-2.5	2.5-5	↑ 2.8	↓ 0.3			↔ 0.3		
		↑ 3.5	↓ 2.5							
Sphingoid base kinase	X_{36}	0-2.5	2.5-5	↔ 0.3					25-28	28-30
		↑ 1.1	↓ 0.4						↑ 1.8	↓ 0.3
Sphingoid 1 phosphate phosphatase	X_{41}	0-2.5	2.5-5	↔ 0.3					25-29	29-30
		↑ 1.2	↓ 0.3						↑ 6	↓ 0.5
GPI remodelase	X_{43}	0-2.5	2.5-5	↓ 0.3	↔ 0.3			↑ 0.4	↑ 0.5	
		↑ 2.2	↓ 0.4							
Sphingosine phosphate lyase	X_{50}	↓ 0.2		↔ 0.3					25-28	28-30
									↑ 2	↓ 0.5
Inositol phosphosphingolipid phospholipase C	X_{51}	0-2	2-5	↑ 2.5	↑ 3		↓ 0.5	↓ 0.3	25-28	28-30
		↑ 3.4	↓ 2.1						↑ 1	↓ 0.3
Phytoceramide alkaline ceramidase	X_{53}	↑ 4.5		↓ 3.7	↓ 2.2		↑ 4.5	↓ 3.7	25-28	28-30
									↑ 6	↓ 2
4 hydroxylase	X_{54}	0-2	2-5	↔ 0.3						
		↑ 2	↓ 0.3							
Mannosyldiinositol phosphorylceramide synthase	X_{55}	0-2	2-5	↑ 2.7	10-13	13-15	↓ 0.5	↓ 0.3	25-28	28-30
		↑ 3.5	↓ 2.2		↓ 2.6	↓ 1.7			↑ 0.6	↓ 0.3
Serine palmitoyltransferase	X_{57}	0-2	2-5	↓ 0.3	↔ 0.3					
		↑ 2.2	↓ 0.5							
Very long chain fatty acid synthase	X_{59}	0-2	2-5	↓ 0.3	↔ 0.3					
		↑ 2.3	↓ 0.5							

Estimation of Q_{10} values for Enzymes of the Sphingolipid Pathway

The "initial jump" of enzyme activities allows us to estimate Q_{10} values for the different enzymes. These values are computed from the definition

$$Q_{10} = \left[\frac{R_1}{R_0}\right]^{\frac{10}{T_1 - T_0}}.$$

Here, R_0 is the base line enzyme activity under optimal conditions (30°C) and R_1 is the enzyme activity immediately following the temperature change to (39°C); the difference between the two is $T_1 - T_0$. The computed Q_{10} values are summarized as in Table A3.3. Values were only estimated for identifiable enzymes.

Table A3.3: Estimated Q_{10} Values, Based on the Initial Increases in Enzyme Activities.

Enzyme	**Variable**	**Q_{10}**
3-keto-dihydrosphingosine reductase	X_{27}	3.1394
Dihydroceramide aklaline ceramidase	X_{29}	*3.0150*
Inositol phosphorylceramide synthase	X_{33}	*4.0227*
Ceramide synthase	X_{34}	*1.9215*
Mannosyl inositol phosphoceramide synthase	X_{35}	*4.0227*
Sphingoid base kinase	X_{36}	*1.1117*
Sphingoid 1 phosphate phosphatase	X_{41}	*1.2246*
GPI remodelase	X_{43}	*2.4014*
Sphingosine phosphate lyase	X_{50}	*0.1673*
Inositol phosphosphingolipid phospholipase C	X_{51}	*3.8952*
Phytoceramide alkaline ceramidase	X_{53}	*3.5153*
Hydroxylase	X_{54}	*2.1601*
Mannosyldiinositol phosphorylceramide synthase	X_{55}	*4.0227*
Serine palmitoyltransferase	X_{57}	*2.4014*
Very long chain fatty acid synthase	X_{59}	*2.5230*

Assessment of Simulation Results with a "Negative Control"

The third assessment mentioned, but not detailed in the text, is a "negative control." We fixed eight key enzymes (X_{34}, X_{36}, X_{41}, X_{43}, X_{50}, X_{54}, X_{57} and X_{59}) at their optimal steady-state values and then performed the same optimization as before. The thus inferred enzyme activities of this "constrained model" do not produce good fits to the sphingolipid heat stress data (Figure A3.7). Specifically, the sum of squared errors (SSE) for Figure 3.2 is 5.79×10^{-4}, whereas it is 2.79×10^{-2} for Figure A3.7. In order to assess the quality of fit further, it is useful to display the residual errors of the individual optimizations with the constrained model in comparison to those obtained with the model in the text. Figure A3.8 clearly demonstrates that the residual errors of the 2,004 original simulations are much lower than those of 200 constrained simulations. In this representation, the X-axis shows the index of each individual simulation and the Y-axis shows the corresponding SSE. Figure A3.9 shows distributions of the SSEs in the two scenarios. These assessments clearly show that the SSEs for the constrained model are much larger than the corresponding SSEs for the original model in the body of the paper and therefore suggest that the key enzymes in the sphingolipid system must respond to heat stress in a coordinated manner.

One should note that SSEs of individual simulations are smaller than those of the averaged model fit. As mentioned in the text, parameterization with averaged values does not necessarily lead to good fits. In the present case, the averaged model in Figure 3.2 of the text is visually not all that different from the best individual simulation results (threshold: SSE < 1.25×10^{-5}) displayed in Figure A3.2 (upper panel). The similarity of these fits is shown in Figure A3.10. The higher SSE of the averaged fit is presumably due to the fact that the plots show fold changes and X6 (PHS-P) has a low steady-state of 0.005, while X3 (DHC) has a substantially higher level (0.0366), but does not change all that much in actual value.

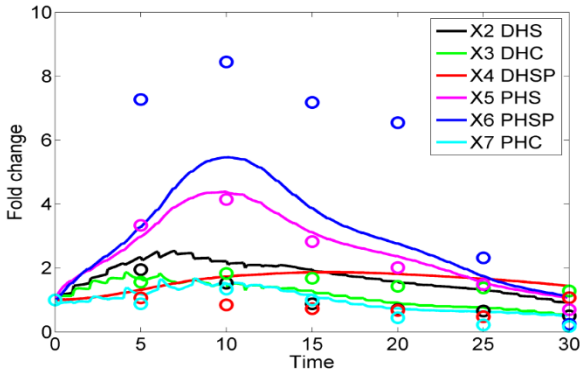


Figure A3.7: When the key enzymes are locked into their normal activity values and all other enzyme activities are allowed to be optimized, the fit of the best model to the experimental data is not very good.

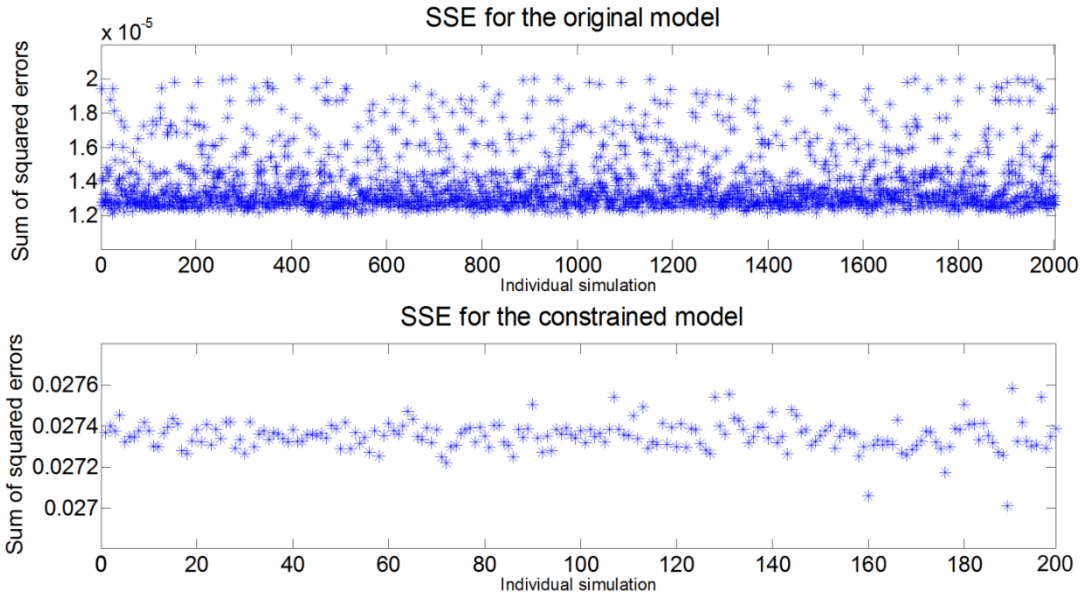


Figure A3.8. Sums of squared errors for individual optimizations. Upper panel: SSEs for 2,000 simulations with the original model. Lower panel: SSEs for 200 simulations with the constrained model. The X-axis shows the index of each individual simulation, while the Y-axis shows the corresponding sum of squared errors (SSE); note different scales.

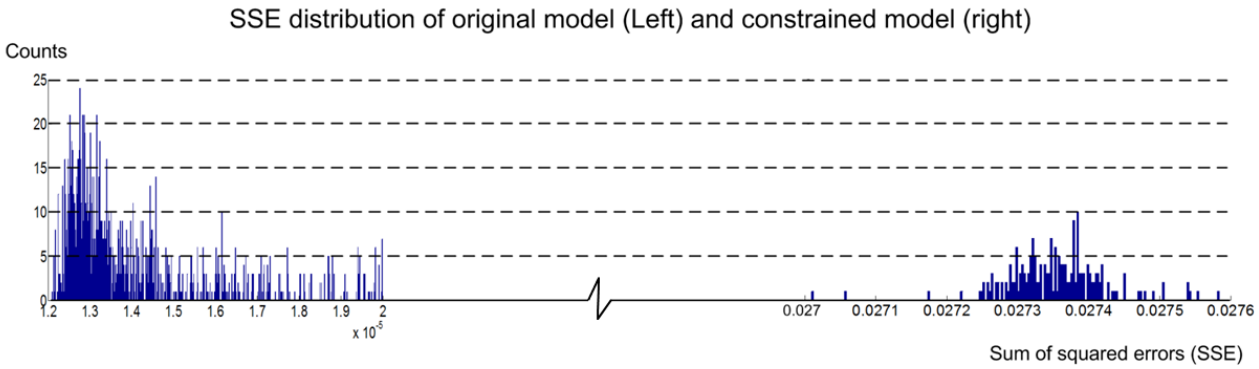


Figure A3.9. Distributions of sums of squared errors for individual simulations. The distribution on the left contains SSEs for the model in which all enzymes are allowed to change. The distribution on the right contains the corresponding SSE values for the constrained model.

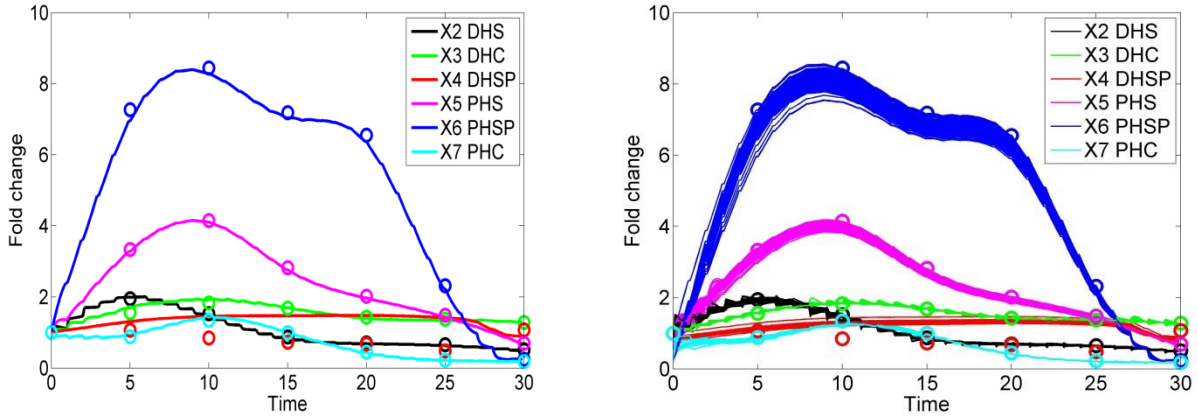


Figure A3.10. Comparison of data fits. Left panel: Data fitted with the unconstrained averaged model (identical to Figure 3.2 of the text). Right panel: 179 data fits with individual model simulations that resulted in SSE < 1.25×10^{-5} (cf. Figure A3.8).

Akaike Information Criterion

The Akaike Information Criterion (AIC) provides a measure of goodness-of-fit for multi-model inferences. We used this criterion to assess the quality of our model. While we are only using one model, the AIC value result may be helpful for future comparisons with alternative models. The formulation of AIC was adapted from [128]. AIC is based on the Kullback-Leibler (K-L) information measure $I(f, g)$ given below:

$$I(f,g) = \int f(x)\, log(\frac{f(x)}{g(x|\theta)})dx$$

Here, $f(x)$ refers to the true representation the model is supposed to capture and $g(x|\theta)$ is the model. $I(f, g)$ can be written in terms of expectations as

$$I(f,g) = \int f(x)\, log\big(f(x)\big)dx - \int f(x)\, log\big(g(x|\theta)\big)dx$$
$$= E_f\big[log\big(f(x)\big)\big] - E_f\big[log\big(g(x|\theta)\big)\big]$$

Here, the first term refers to the true representation, which however is unknown and therefore replaced with an unknown constant $C = E_f\big[log\big(f(x)\big)\big]$. The second term requires the estimation of the expectation of g given data y, $E_y E_x\left[log\left(g\big(x\big|\hat{\theta}(y)\big)\right)\right]$, where $\hat{\theta}(y)$ is the maximum likelihood estimator with respect to y.

Akaike ([129,130]) found that the maximum likelihood estimator for $E_y E_x\left[log\left(g\big(x\big|\hat{\theta}(y)\big)\right)\right]$ is biased, but that there is a relationship between the bias (called K) and the K-L information, namely:

$$log\left(L\left(\hat{\theta}(y)\right)\right) - K = C - E_{\hat{\theta}}[I(f,\hat{g})]$$

Furthermore, the bias K is equal to the number of estimated parameters. Akaike thus introduced AIC as

$$AIC = -2log\left(L\big(\hat{\theta}\big|y\big)\right) + 2K$$

For least square estimation, this definition reduces to

$$AIC = nlog(\sigma^2) + 2K$$

Where n is the number of data, σ^2: $\frac{\sum_{i=1}^{n}(\varepsilon_i)^2}{n}$ and ε_i is the sum of squared errors for each sample. To extend the applicability of AIC further, Akaike introduced AIC_c to deal with small samples, which are defined here as $\frac{sample\ size\ n}{\#\ paramter\ K} < 40$. The result is:

$AIC_c = AIC + \frac{2K(K+1)}{n-K-1}$.

We obtained the AIC and AICc values for our 4414 models by using the definitions given above. The parameter *n* represents the number ofdata, which in our case is $n = 30 * 6 = 180$. σ^2 is the mean sum of squared errors (SSE) for all simulations. As we have 30 points to fit in each simulation, σ^2 can be computed as

$$\sigma^2 = \frac{\sum_{i=1}^{n}(\sum_{j=1}^{30}(X(j) - Data(j))^2)}{n}$$

Here, *X* is the vector of six sphingolipids and *Data* represents the vector with the corresponding smoothed data.

In the text, we shown results of 2004 models selected from a much larger pool of initial models (4144 models), based on the sum of squared errors (SSEs). Computing the averaged model from these 2004 models, we further provided the validation for the averaged model by comparing the observed heat stress data with the heat stress profiles computed with this model. This comparison, yielding a good fit, supported the use of the averaged model for making inference regarding enzyme activities.

Now we provided the comparison between the SSEs criterion and AICc criterion in model selection. We have computed AICc values for all 4144 models and binned them in a histogram in Figure A3.11. The histogram very nicely shows that 2018 models (left-most column) are superior to the others. Comparing

this set of models with the set identified by SSEs (2004 models), we found that 98.66% (1991) of the 2018 models simultaneously satisfited both criteria, AICc and SSE. Now we can use either 2004 (SSE), or 2018 (AICc) good fit models to make the same inference of enzyme activities.

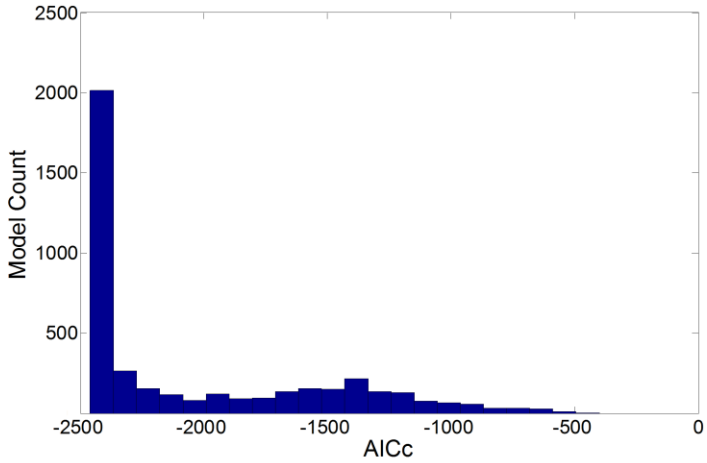


Figure A3.11: The histogram of 4144 models clearly indicates that the 2018 models in the left-most column are superior to all other parameterizations. 98.66% (1991) of the models from this column were also identified by SSEs, thereby demonstrating very strong consistency between the two measures of quality.

Appendix B

Further details of the computational methods in Chapter 4

Model equations

The model equations below correspond to the pathway in Figure 4.2 of the main article. $X_1 \sim X_5$ represent C14/C16, C18, C18:1, C24 and C26 DHC, respectively, while $X_6 \sim X_{10}$ represent C14/C16, C18, C18:1, C24 and C26 PHC, respectively. $X_{11} \sim X_{15}$ represent the corresponding C14/C16, C18, C18:1, C24 and C26 fatty acyl CoA variants. DHS and PHS were coded as X_{16} and X_{17}. Each V_i denotes a flux, which is modeled according to the Generalized Mass Action (GMA) framework within Biochemical Systems Theory[117]. For example, $V1 = \gamma_{1,1} X_{16}^{f_{1,1,16}} X_{11}^{f_{1,1,11}}$. In this formulation, $\gamma_{1,1}$ is a rate constant, and $f_{1,1,16}$ and $f_{1,1,11}$ are kinetic orders. The rate constants are at first unknown and must be specified from the data, while the kinetic orders were obtained from the previously published paper or assumed to have a value of 1, if they were unknown. X_{16} and X_{17} (DHS and PHS) are independent functions obtained from polynomial fitting of the data from published paper [94].

$$\frac{dX_1}{dt} = V1 + V2 - V3 - V4 - V5$$

$$\frac{dX_2}{dt} = V6 + V7 - V8 - V9 - V10$$

$$\frac{dX_3}{dt} = V11 + V12 - V13 - V14 - V15$$

$$\frac{dX_4}{dt} = V16 + V17 - V18 - V19 - V20$$

$$\frac{dX_5}{dt} = V21 + V22 - V23 - V24 - V25$$

$$\frac{dX_6}{dt} = V4 + V26 + V27 - V28 - V29$$

$$\frac{dX_7}{dt} = V9 + V30 + V31 - V32 - V33$$

$$\frac{dX_8}{dt} = V14 + V34 + V35 - V36 - V37$$

$$\frac{dX_9}{dt} = V19 + V38 + V39 - V40 - V41$$

$$\frac{dX_{10}}{dt} = V24 + V42 + V43 - V44 - V45 - V46$$

$$\frac{dX_{11}}{dt} = V47 - V1 - V26 - V48$$

$$\frac{dX_{12}}{dt} = V48 - V6 - V30 - V49 - V50$$

$$\frac{dX_{13}}{dt} = V49 - V11 - V34 - V51$$

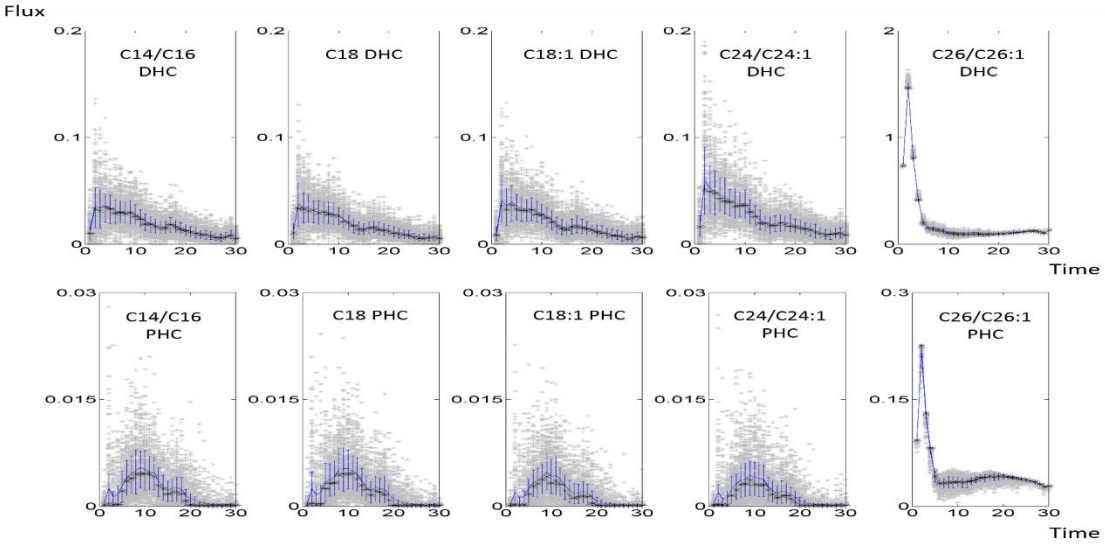
$$\frac{dX_{14}}{dt} = V50 + V51 - V16 - V38 - V52$$

$$\frac{dX_{15}}{dt} = V52 - V21 - V42 - V53$$

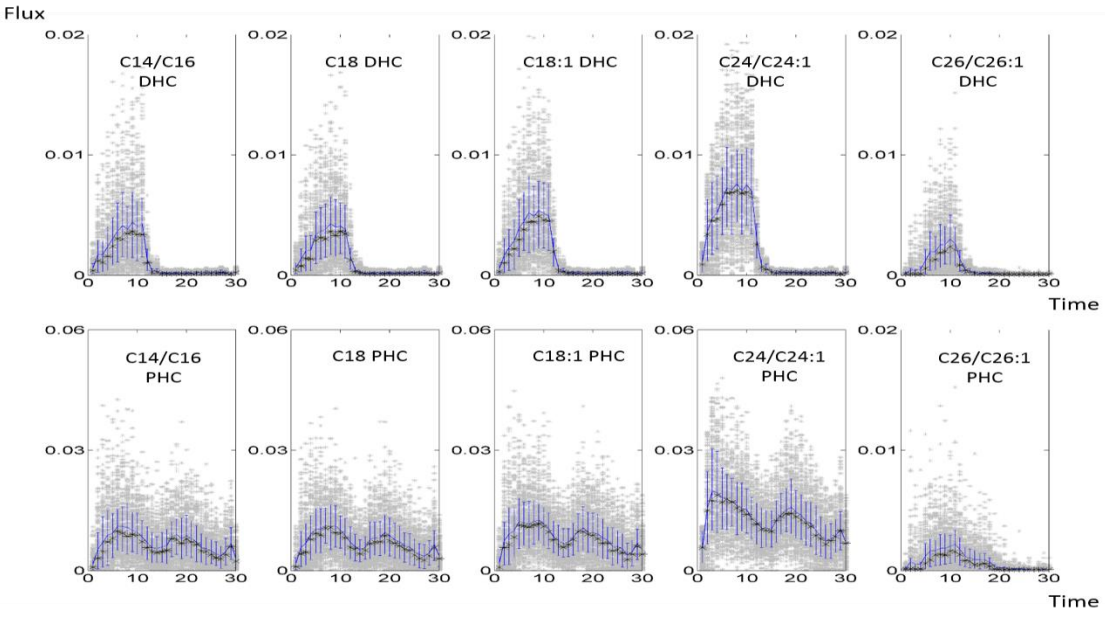
Flux distributions

The distributions of fluxes contributing to the dynamics of the key sphingolipid species addressed in this paper are given in the main text. Figure A4.1 provides the remaining flux distributions. Each set of results was collected from 2,000 simulations. Gray dots, blue lines, blue bars and black asterisks represent single simulation results, mean values, 20th and 80th percentiles and median values, respectively.

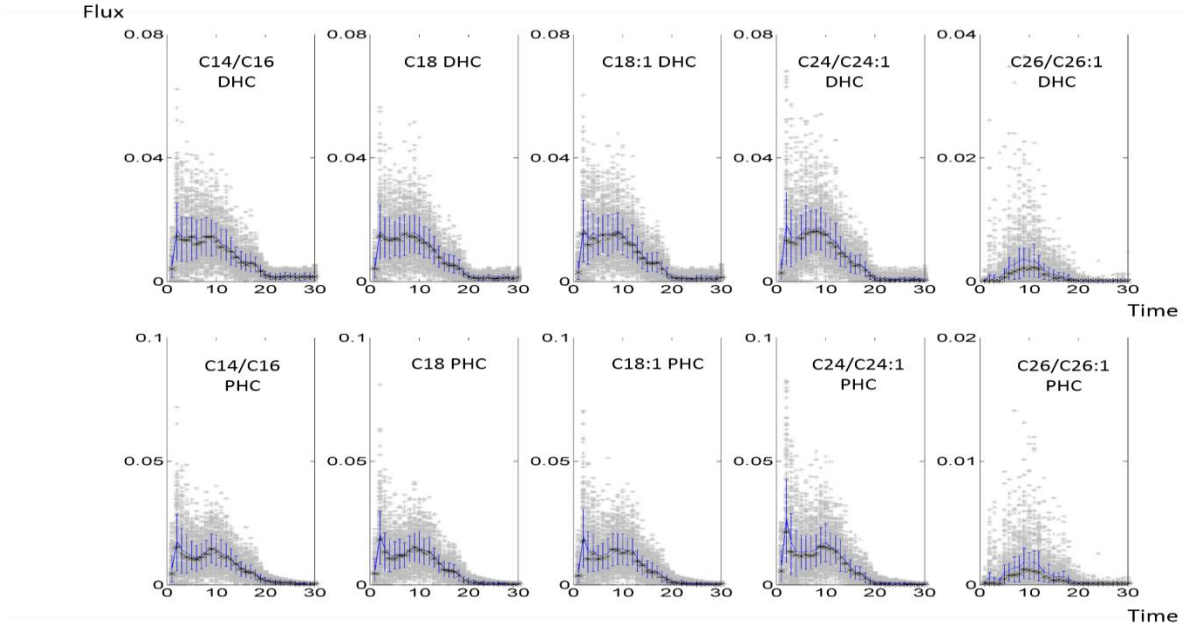
Ceramide synthase



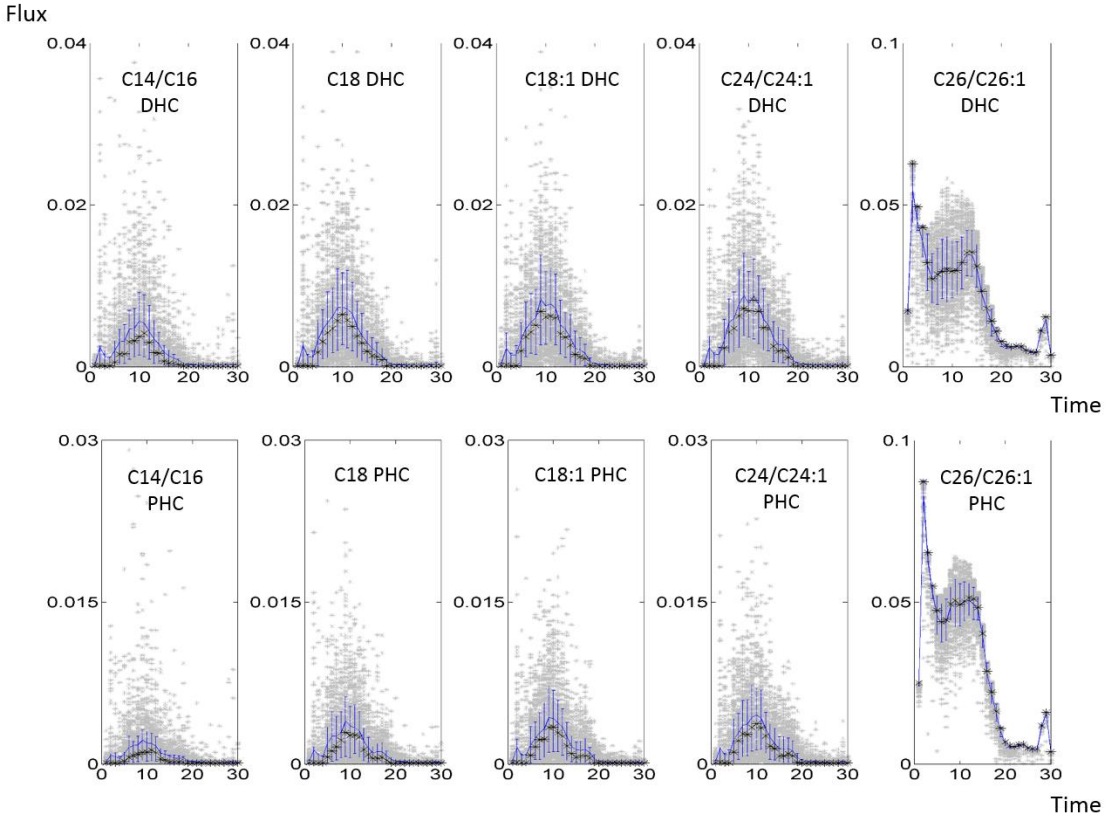
Dihydroceramidase & Phytoceramidase



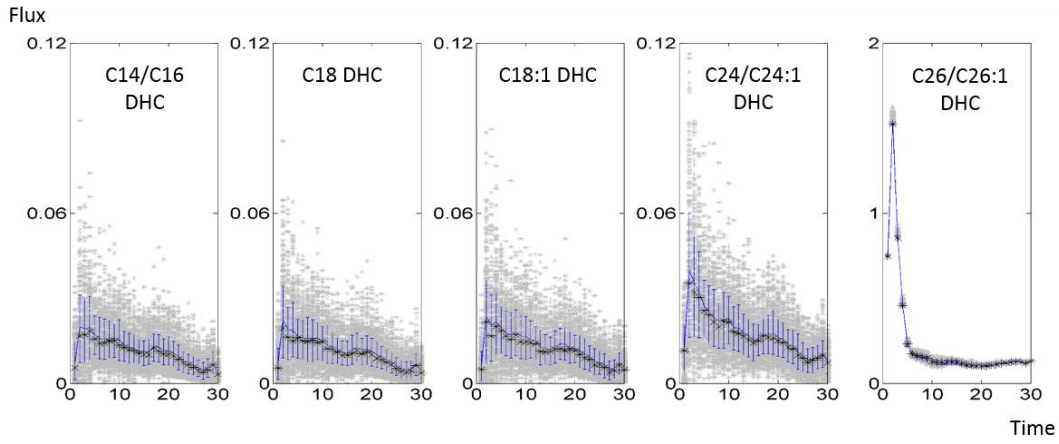
IPC synthase



IPCase



Hydroxylase



Elongases 1,2,3, Remodelase, and Desaturase

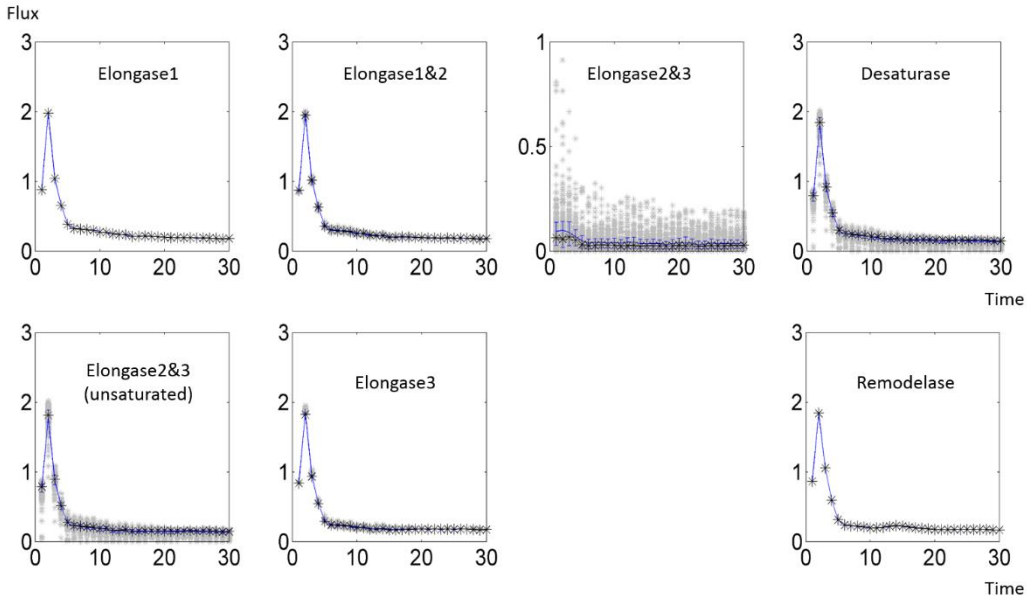
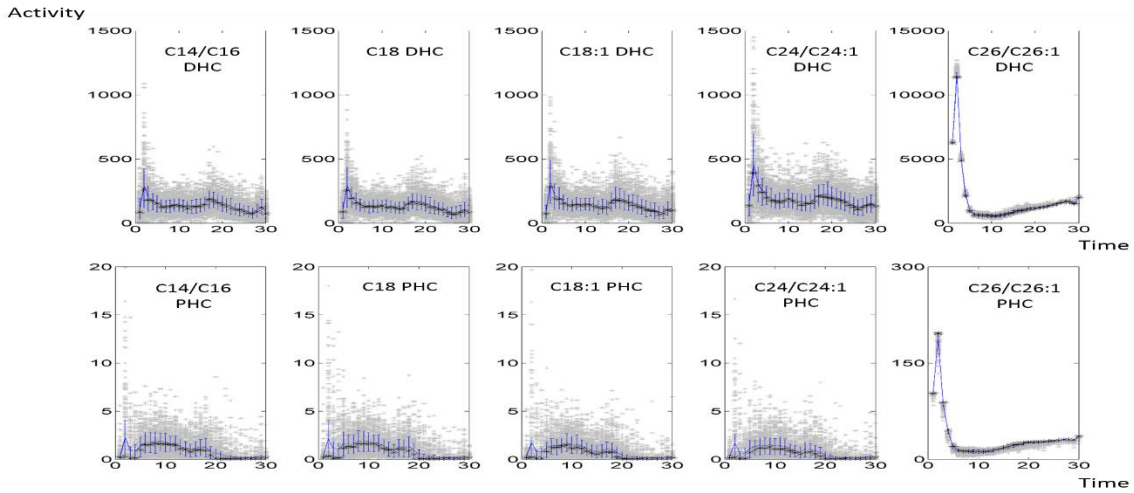


Figure A4.1: Distributions of fluxes not shown in the main text. Compare with Figs. 4.4-4.9 in Chapter 4.

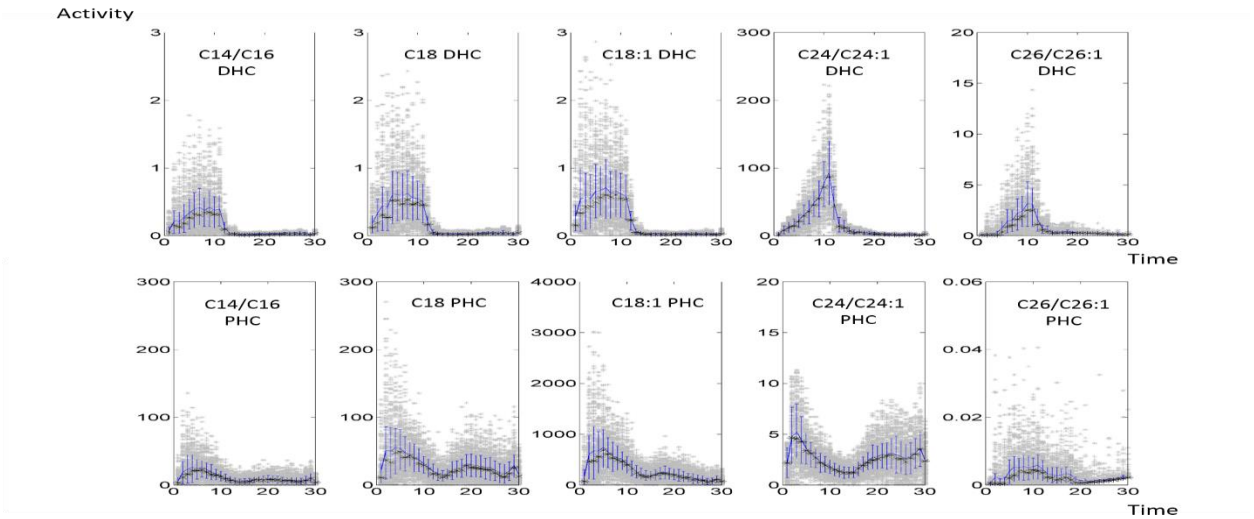
Magnitudes of Enzyme Activities

It was already mentioned in the main text that the actual enzyme amounts and rate constants are not known. As it is typical, we assume that enzyme activities enter a flux representation in a linear manner. With this assumption we obtain coarse estimates of the product of the rate constant and the enzyme activity. This product corresponds to a V_{max} value, which by definition consists of the product of k_{cat} and the total enzyme concentration. Coarse estimates of the ceramide synthase activities are presented in the main text; all other enzyme activities are shown here. Gray dots (which merge into a grey region in each plot) show the individual simulation results; blue lines indicate the means; vertical blue bars correspond to the 20th and 80th percentiles; black asterisks are medians. It is clear that the means and medians are quite similar.

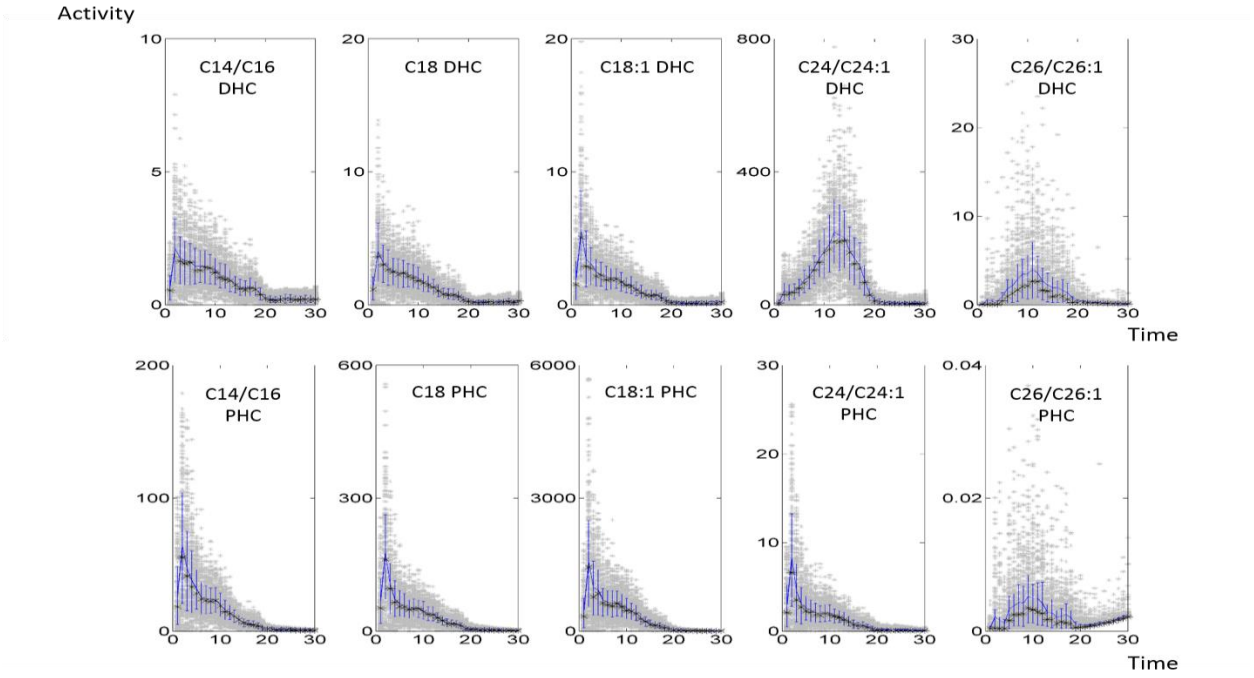
Ceramide Synthase



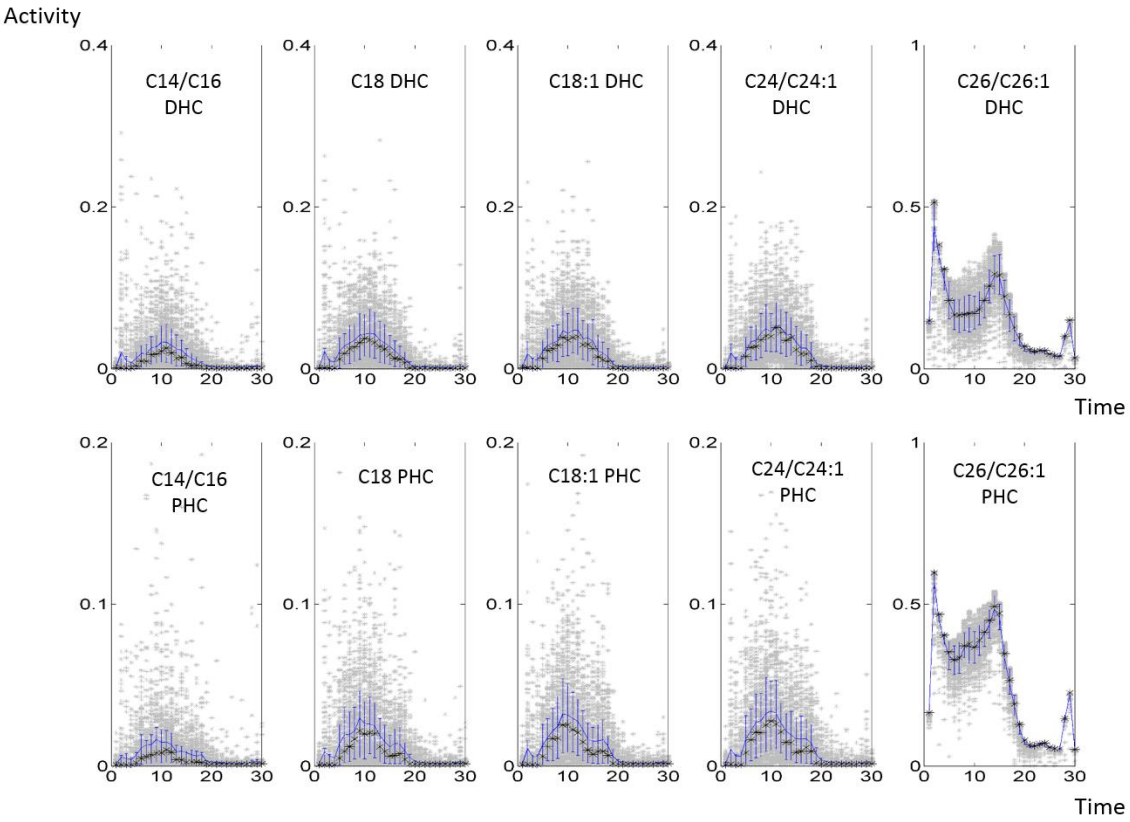
Dihydroceramidase & Phytoceramidase



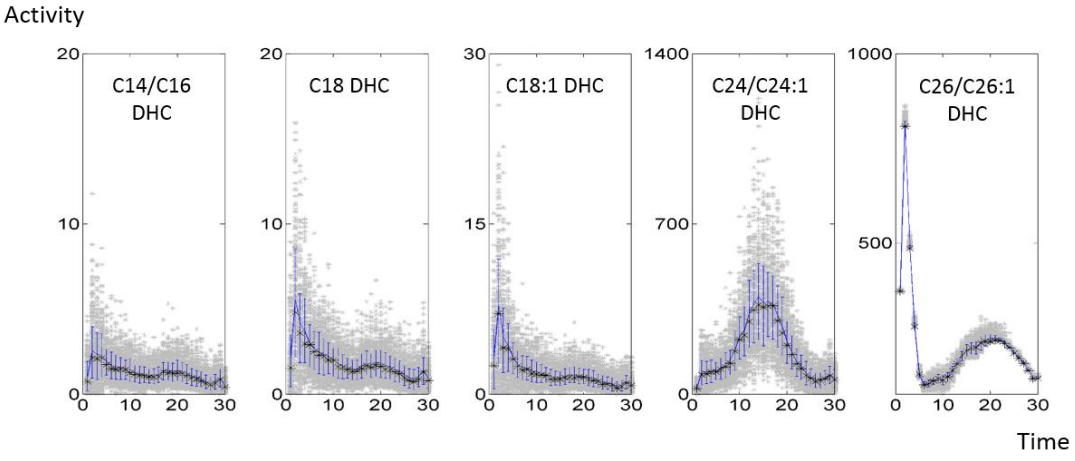
IPC synthase



IPCase



Hydroxylase



Elongases 1,2,3, Remodelase, and Desaturase

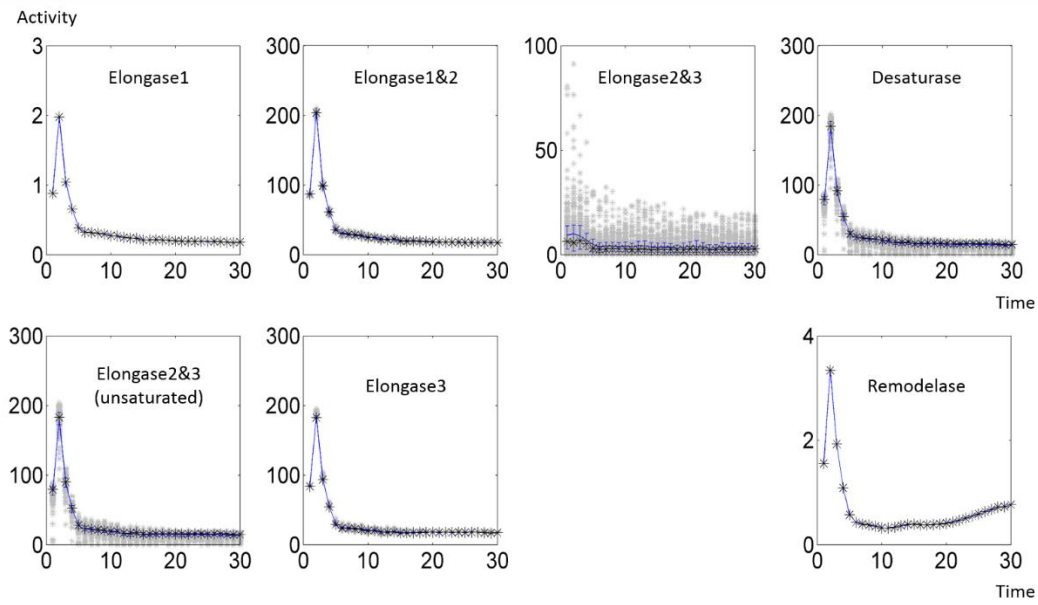
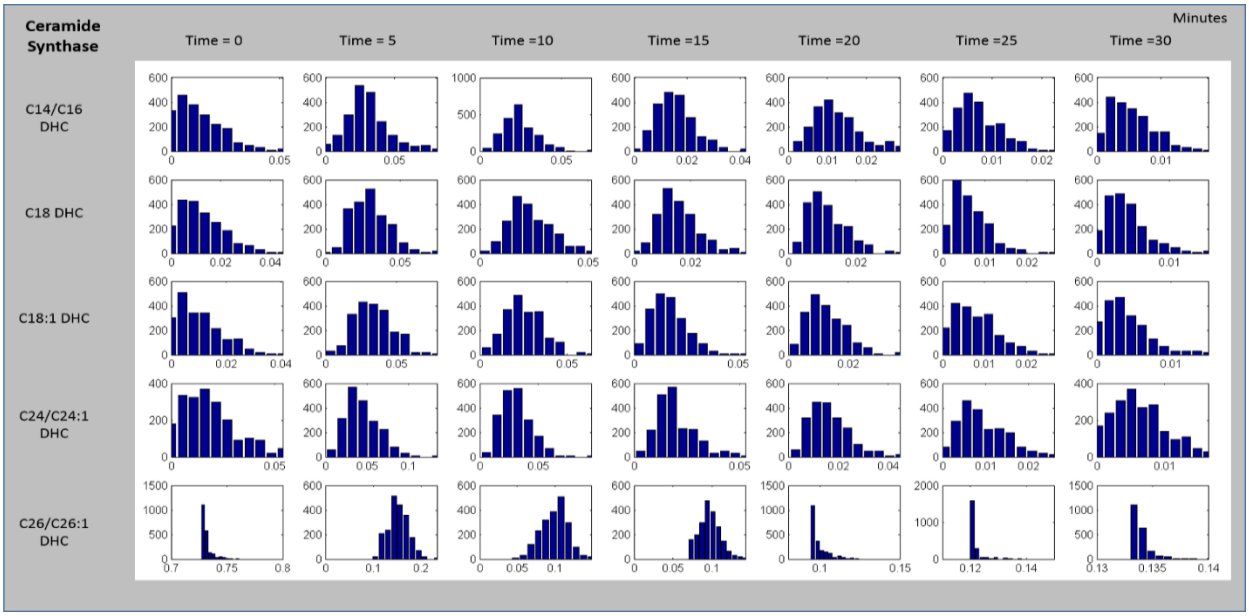


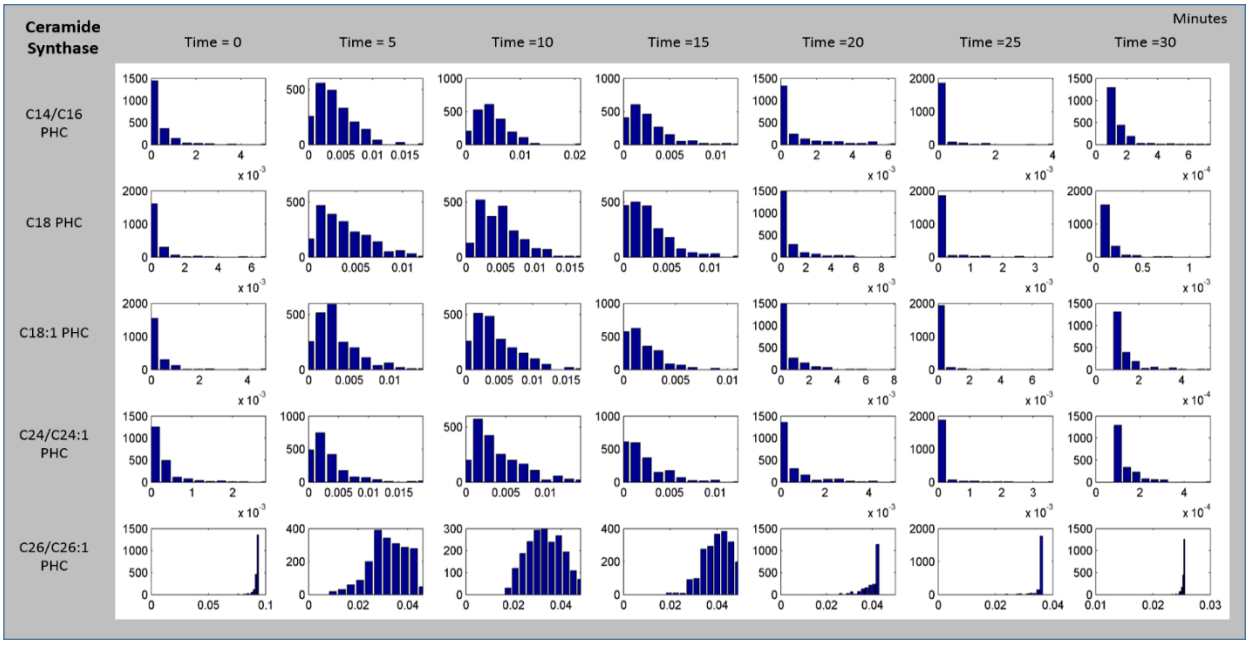
Figure A4.2: Simulation results for enzyme activities over time, as well as means (blue), medians (black asterisks), and 20th and 80th percentiles (blue bars). *Compare with Figure 4.11 in the Text.*

Histograms

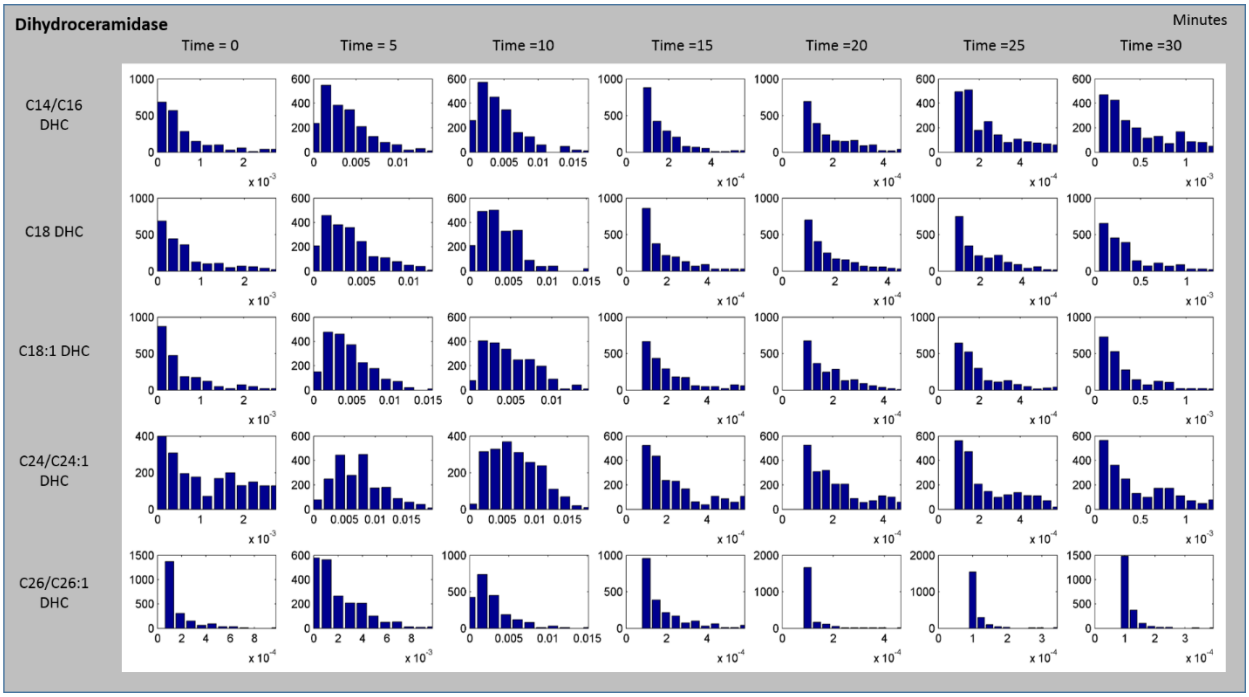
All histograms of each flux, from time point 0 to 30, were examined in order to ensure that the estimated fluxes were unimodal and well constrained in terms of a small variance. The main text showed the histograms of fluxes from DHS to C16 DHC and PHS to C16 PHC at the 0th, 5th, 10th, 15th, 20th, 25th and 30th minutes of heat stress. The histograms of all 53 fluxes are given here.

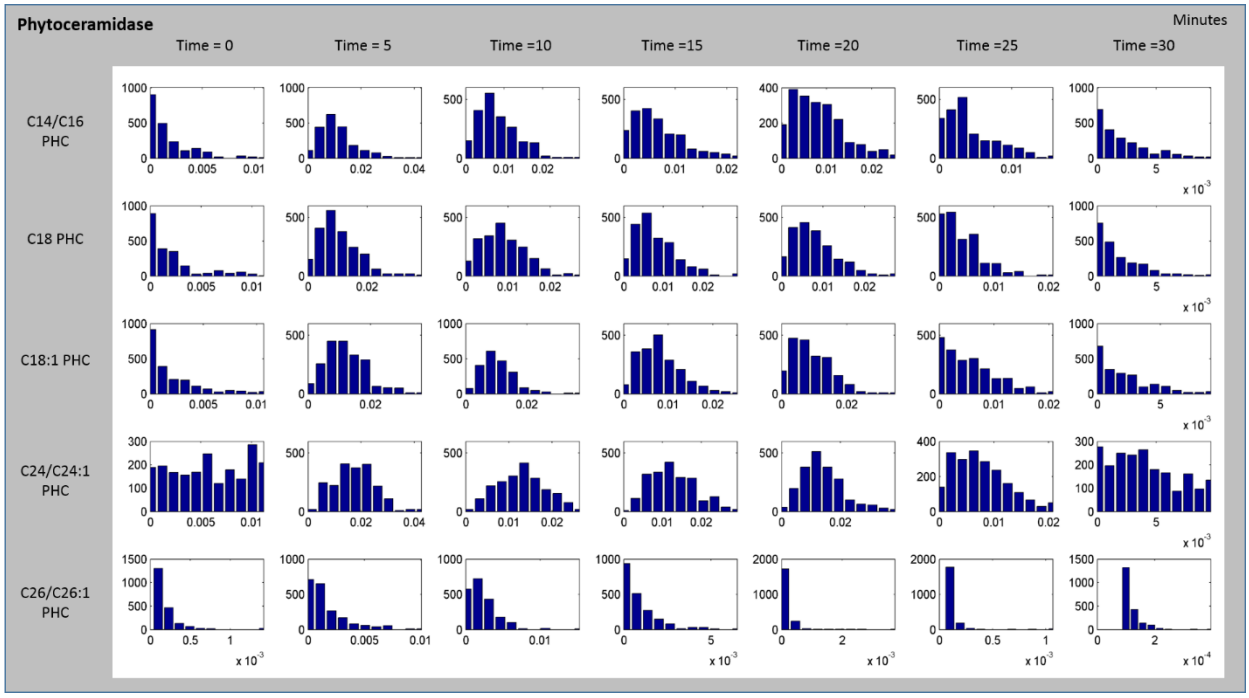
Ceramide synthase



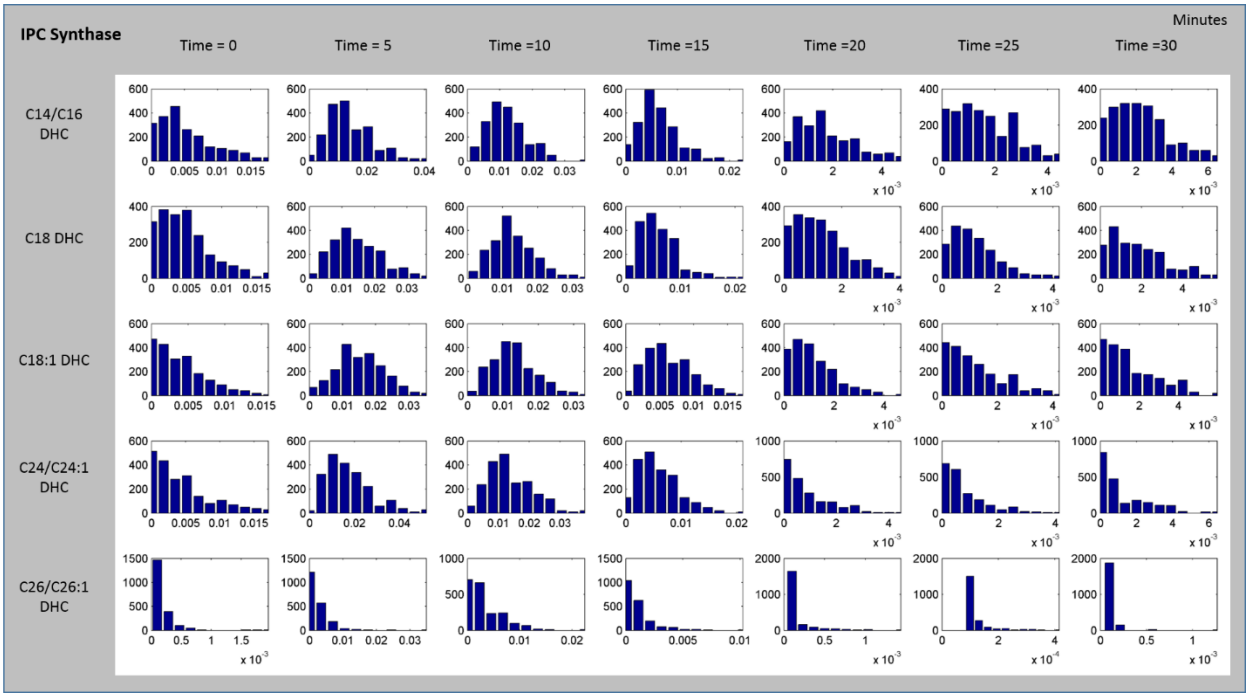


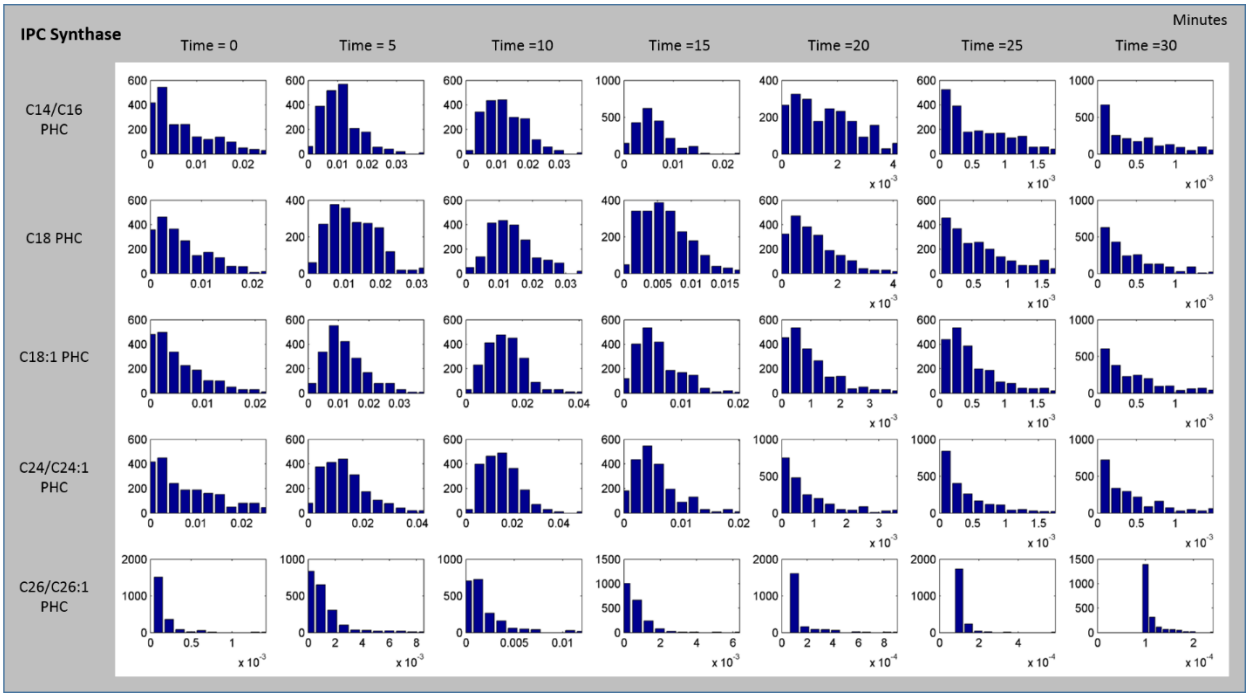
Dihydroceramidase & Phytoceramidase



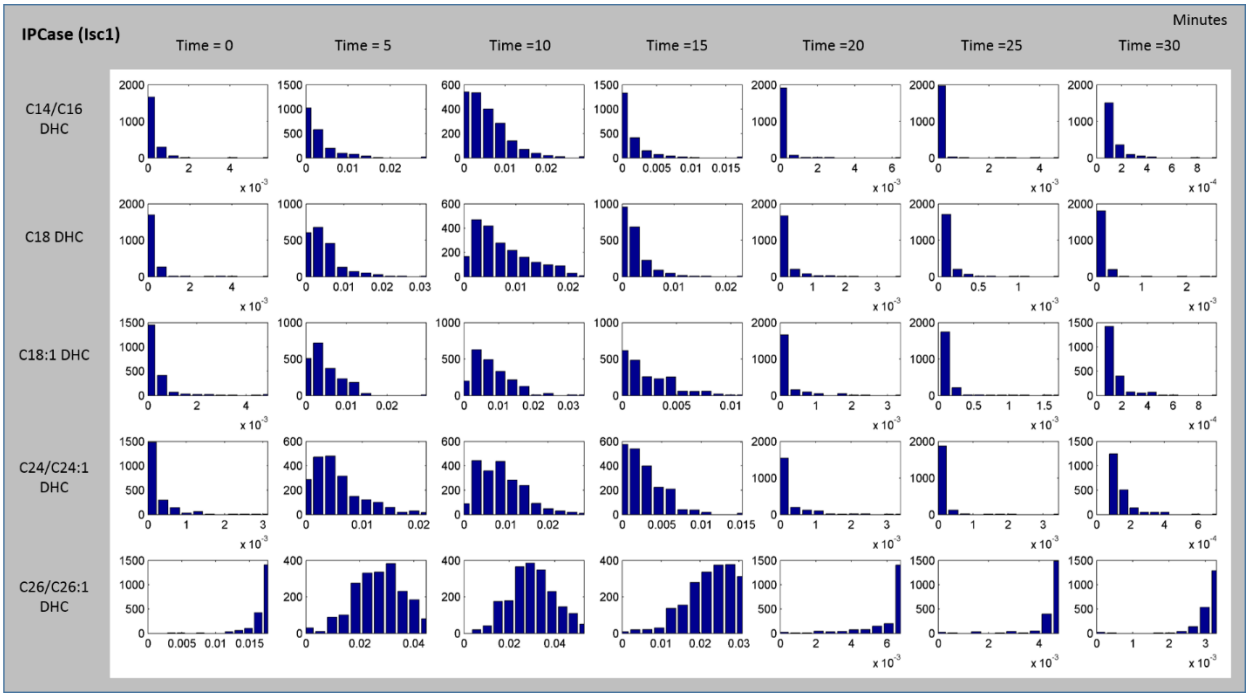


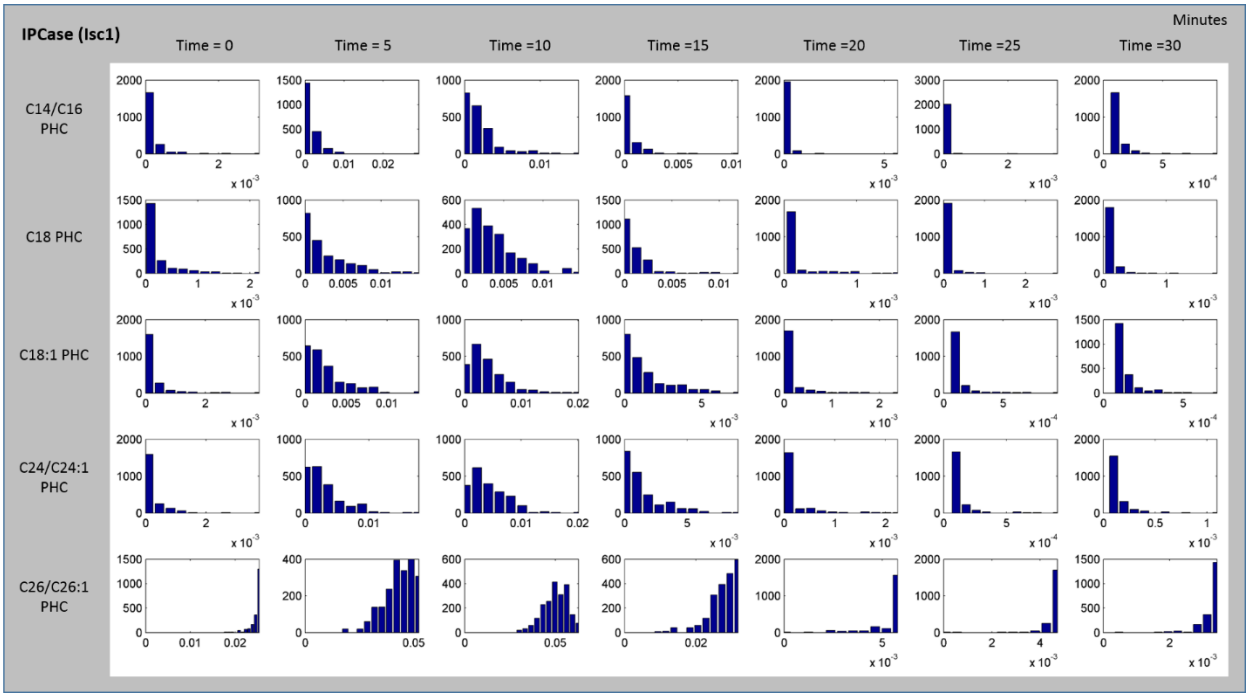
IPC synthase



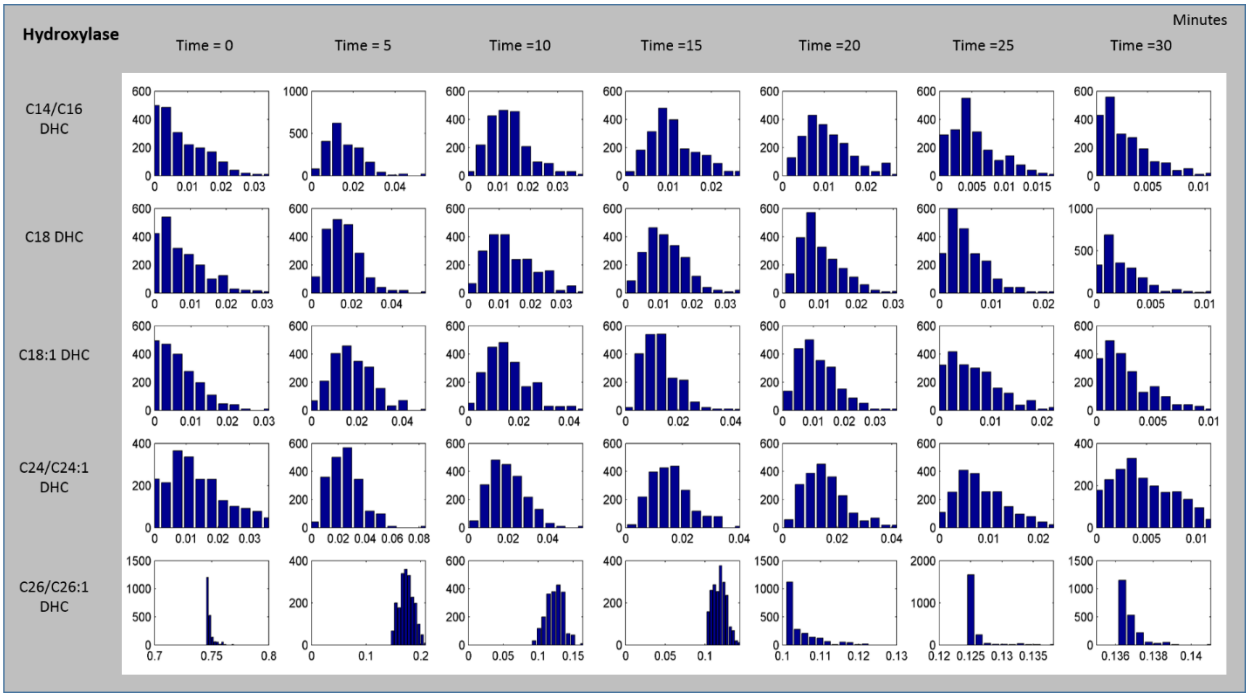


IPCase (Isc1)


IPCase (Isc1)
Minutes
Time = 0
Time = 5
Time =10
Time =15
Time =20
Time =25
Time =30
C14/C16 DHC
C18 DHC
C18:1 DHC
C24/C24:1 DHC
C26/C26:1 DHC


IPCase (Isc1)
Minutes
Time = 0
Time = 5
Time =10
Time =15
Time =20
Time =25
Time =30
C14/C16 PHC
C18 PHC
C18:1 PHC
C24/C24:1 PHC
C26/C26:1 PHC

DHC Hydroxylase



Elongases 1,2,3, Remodelase, and Desaturase

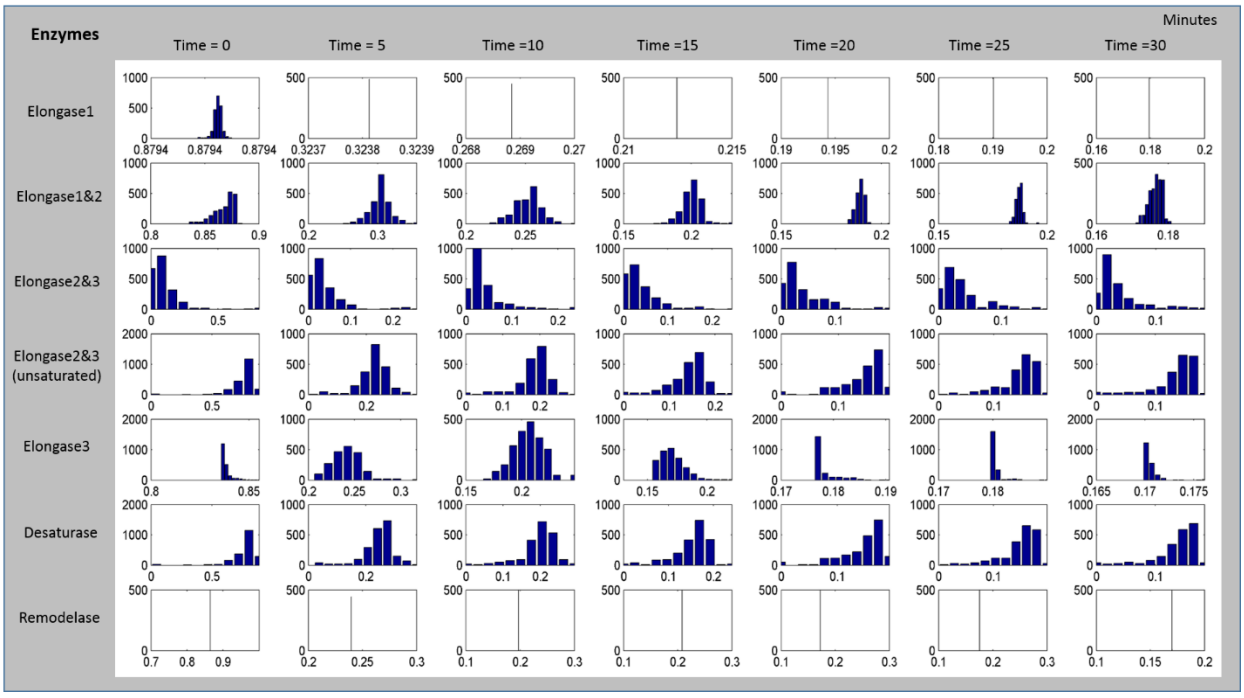


Figure A4.3: Histograms of fluxes at all measured time points. *Compare with Figure 4.12 in the Text.*

Impact of data variability on estimated distributions of dynamic fluxes

Results shown in the main text are calculated based on interpolation of mean of the duplicate raw data (Figure 4.10). In order to assess the impact of data variability, we performed the following step-wise analysis:

A. Adopt the original, smoothed and interpolated mean data as the baseline concentration.
B. Sample a uniformly distributed random variable from the range [1/1.5, 1.5].
C. Perturb the original dataset from time point 0 to 30 by the drawn random number and repeat for 1000 simulations.
D. For each set of randomly perturbed data, redo the simulation in the "Methods" Section to obtain dynamic fluxes.
E. Compare the fluxes based on perturbed data with the originally estimated fluxes.

The perturbed data are presented in Figure A4.4, and the perturbed flux distributions are shown in Figure A4.5. Gray dots, blue lines, blue bars and black asterisks represent single simulations, mean values, 20th and 80th percentiles, and median values, respectively. The flux distributions obtained from perturbed data are similar to the originally estimated fluxes (used in the main text). Increased variances are observed, but they do not influence the further estimation of enzyme activities.

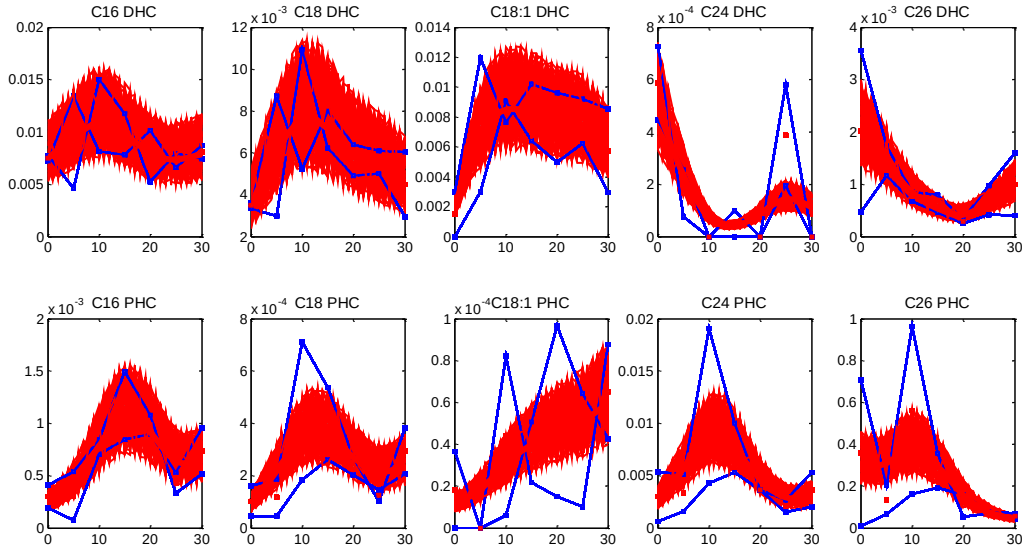
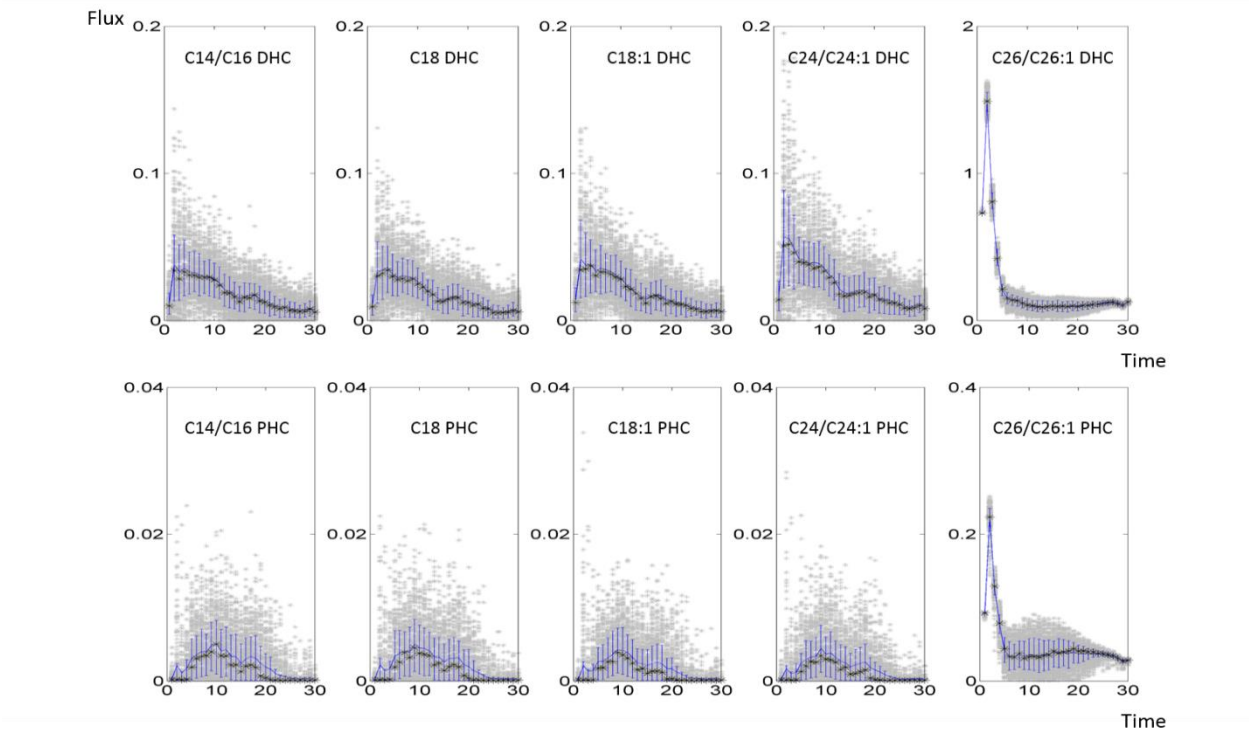
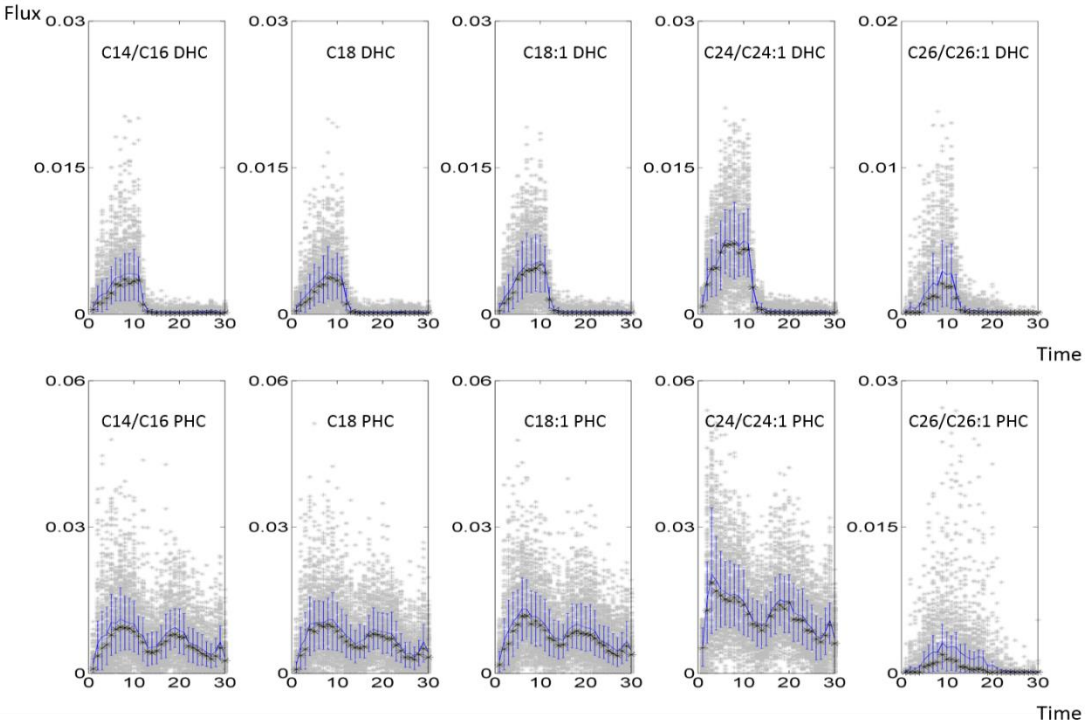


Figure A4.4: Perturbed data set based on 1.5 fold variation of the interpolated mean data.

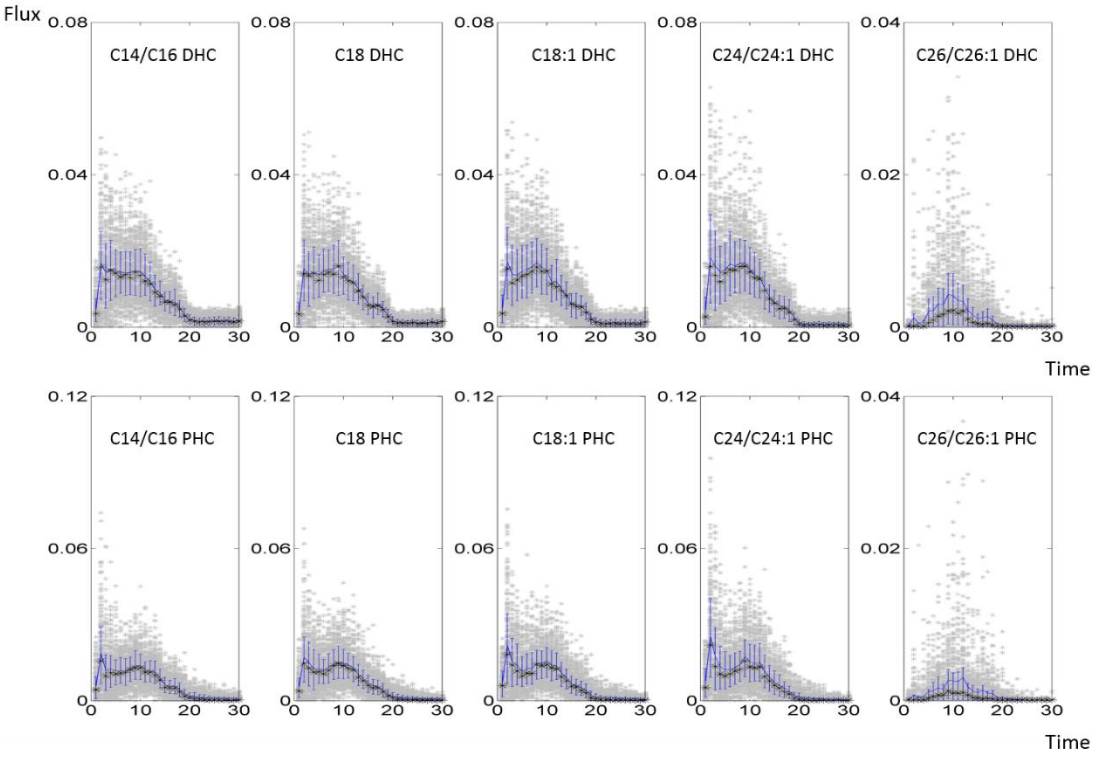
Ceramide Synthase



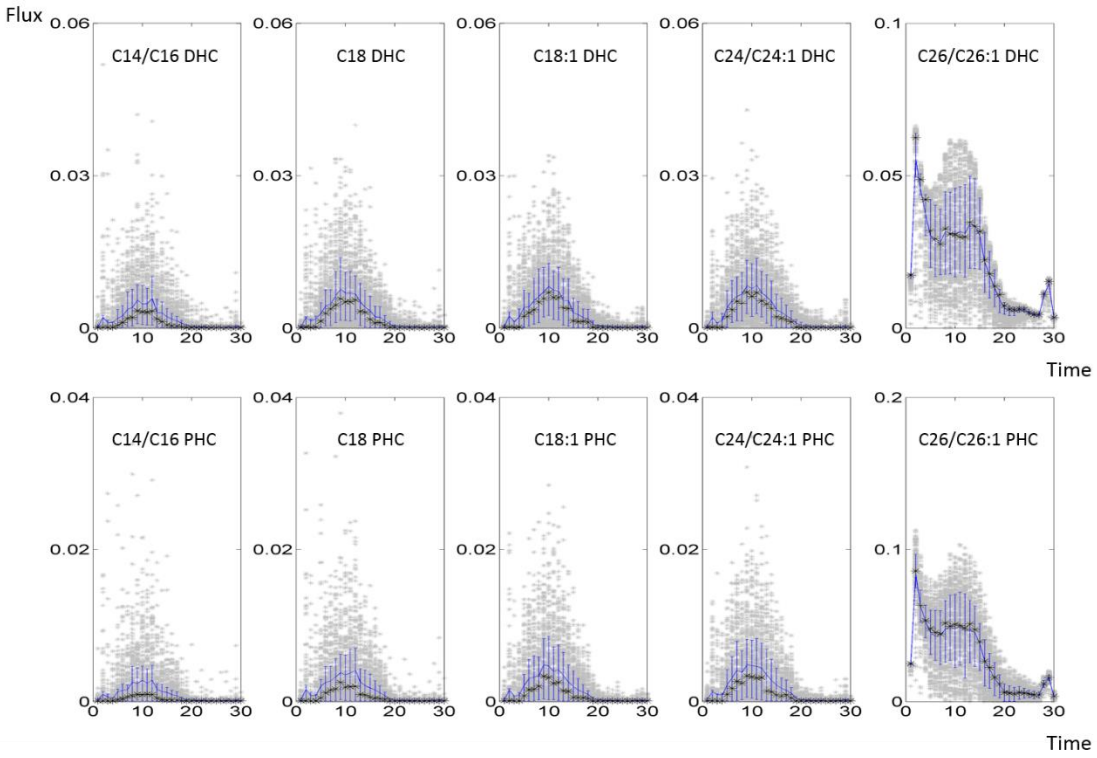
Dihydroceramidase and Phytoceramidase



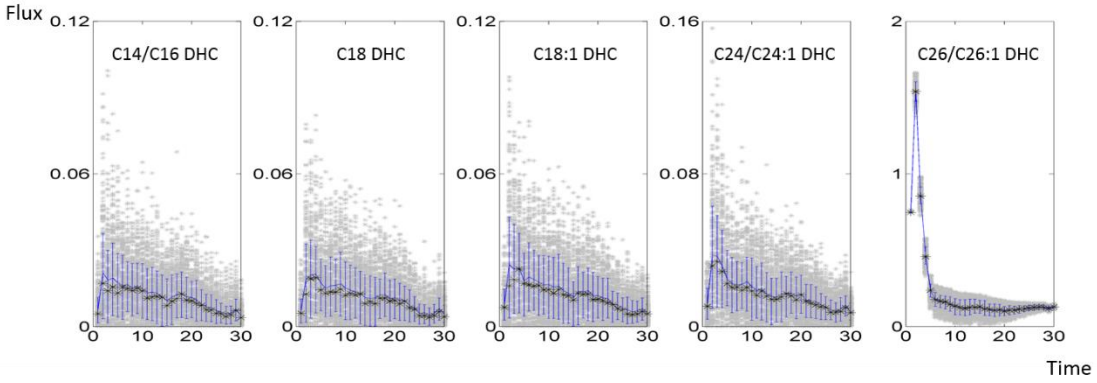
IPC Synthase



IPCase



Hydroxylase



Elongases 1,2,3, Remodelase, and Desaturase

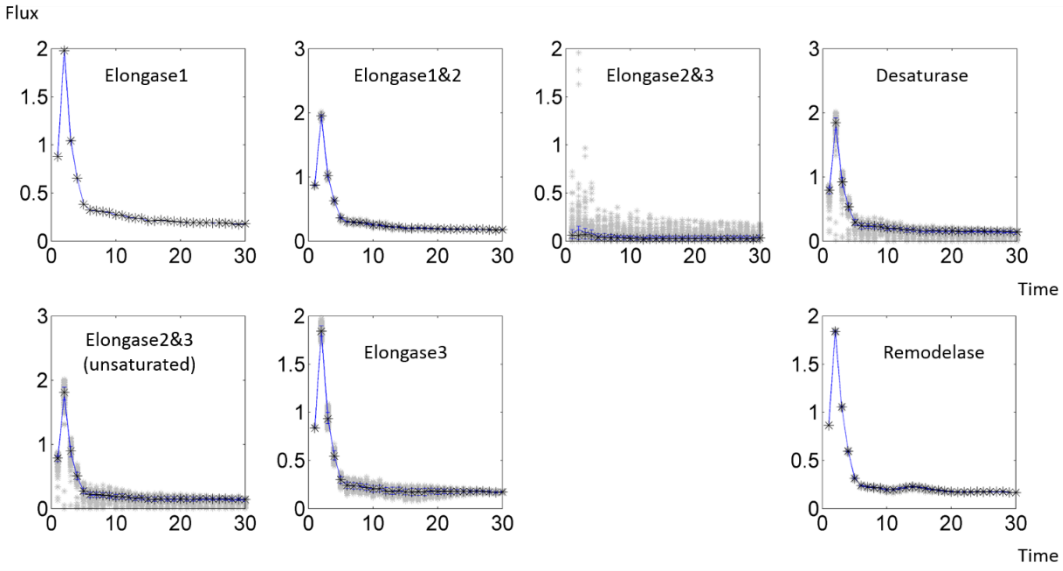


Figure A4.5: Flux distributions calculated from perturbed data set.

References

1. Gault CR, Obeid LM, Hannun YA (2010) An overview of sphingolipid metabolism: from synthesis to breakdown. Advances in Experimental Medicine and Biology 688: 1-23.
2. Hannun YA, Obeid LM (2008) Principles of bioactive lipid signalling: lessons from sphingolipids. Nature reviews Molecular cell biology 9: 139-150.
3. Lahiri S, Futerman AH (2007) The metabolism and function of sphingolipids and glycosphingolipids. Cell Mol Life Sci 64: 2270-2284.
4. Cowart LA, Okamoto Y, Pinto FR, Gandy JL, Almeida JS, et al. (2003) Roles for Sphingolipid Biosynthesis in Mediation of Specific Programs of the Heat Stress Response Determined through Gene Expression Profiling. Journal of Biological Chemistry 278: 30328-30338.
5. Pyne NJ, Pyne S (2010) Sphingosine 1-phosphate and cancer. Nat Rev Cancer 10: 489-503.
6. Hohmann S, Mager WH (2003) Heat Stress Responses: Springer.
7. Morimoto R (1993) Cells in stress: transcriptional activation of heat shock genes. Science 259: 1409-1410.
8. Slater MR, Craig EA (1987) Transcriptional regulation of an hsp70 heat shock gene in the yeast *Saccharomyces cerevisiae*. Molecular and Cellular Biology 7: 1906-1916.
9. Alvarez-Vasquez F, Sims K, Voit E, Hannun Y (2007) Coordination of the dynamics of yeast sphingolipid metabolism during the diauxic shift. Theoretical Biology and Medical Modelling 4: 42.
10. Alvarez-Vasquez F, Sims KJ, Cowart LA, Okamoto Y, Voit EO, et al. (2005) Simulation and validation of modelled sphingolipid metabolism in *Saccharomyces cerevisiae*. Nature 433: 425-430.
11. Alvarez-Vasquez F, Sims KJ, Hannun YA, Voit EO (2004) Integration of kinetic information on yeast sphingolipid metabolism in dynamical pathway models. Journal of Theoretical Biology 226: 265-291.
12. Savageau MA (1976) Biochemical systems analysis : a study of function and design in molecular biology. Reading, Mass.: Addison-Wesley Pub. Co. Advanced Book Program. xvii, 379 p.
13. Savageau MA (1985) A theory of alternative designs for biochemical control systems. Biomed Biochim Acta 44: 875-880.
14. Alon U (2006) An Introduction to Systems Biology: Design Principles of Biological Circuits. Boca Raton, FL: Chapman & Hall/CRC.
15. Barabási A-L, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. Nature Reviews Genetics 5: 101-113.
16. Lipshtat A, Purushothaman SP, Iyengar R, Ma'ayan A (2008) Functions of bifans in context of multiple regulatory motifs in signaling networks. Biophys J 94: 2566-2579.
17. Ma'ayan A, Cecchi GA, Wagner J, Rao AR, Iyengar R, et al. (2008) Ordered cyclic motifs contribute to dynamic stability in biological and engineered networks. Proc Natl Acad Sci USA 105: 19235-19249.
18. Erdős P, Rényi A (1959) On random graphs. Publicationes Mathematicae 6: 290–297.
19. Savageau MA (1969) Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. J Theor Biol 25: 365-369.
20. Torres NV, Voit EO (2002) Pathway Analysis and Optimization in Metabolic Engineering. New York: Cambridge University Press. xiv, 305 p.

21. Voit EO (2000) Computational Analysis of Biochemical Systems. A Practical Guide for Biochemists and Molecular Biologists. Cambridge, UK: Cambridge University Press. xii + 530 pp. p.
22. Irvine DH (1991) The method of controlled mathematical comparison. In: Voit EO, editor. Canonical Nonlinear Modeling. New York: Van Nostrand Reinhold. pp. 90-109.
23. Irvine DH, Savageau MA (1985) Network regulation of the immune response: Alternative control points for suppressor modulation of effector lymphocytes. J Immunol 134: 2100-2116.
24. Alves R, Savageau MA (2000) Effect of overall feedback inhibition in unbranched biosynthetic pathways. Biophys J 79: 2290-2304.
25. Voit EO (2003) Biochemical and genomic regulation of the trehalose cycle in yeast: review of observations and canonical model analysis. J Theor Biol 223: 55-78.
26. Schwacke JH, Voit EO (2004) Improved methods for the mathematically controlled comparison of biochemical systems. Theor Biol Med Model 1: 1.
27. Savageau MA, Coelho PM, Fasani RA, Tolla DA, Salvador A (2009) Phenotypes and tolerances in the design space of biochemical systems. Proceedings of the National Academy of Sciences of the United States of America 106: 6435-6440.
28. Hlavacek WS, Savageau MA (1996) Rules for coupled expression of regulator and effector genes in inducible circuits. J Mol Biol 255: 121-139.
29. Savageau MA (1998) Demand theory of gene regulation. I. Quantitative development of the theory. Genetics 149: 1665-1676.
30. Savageau MA (2001) Design principles for elementary gene circuits: Elements, methods, and examples. Chaos 11: 142-159.
31. Igoshin OA, Alves R, Savageau MA (2008) Hysteretic and graded responses in bacterial two-component signal transduction. Mol Microbiol 68.
32. Beisel CL, Smolke CD (2009) Design principles for riboswitch function. PLoS computational biology 5: e1000363.
33. Voit EO (2003) Design principles and operating principles: the yin and yang of optimal functioning. Math Biosci 182: 81-92.
34. Alvarez-Vasquez F, Sims KJ, Voit EO, Hannun YA (2007) Coordination of the dynamics of yeast sphingolipid metabolism during the diauxic shift. Theor Biol Med Model 4: 42.
35. Voit EO, Radivoyevitch T (2000) Biochemical systems analysis of genome-wide expression data. Bioinformatics 16: 1023-1037.
36. Alves R, Vilaprinyo E, Hernández-Bermejo B, Sorribas A (2008) Mathematical formalisms based on approximated kinetic representations for modeling genetic and metabolic pathways. Biotechnology and Genetic Engineering Reviews 25: 1-40.
37. Guillén-Gosálbez G, Sorribas A (2009) Identifying quantitative operation principles in metabolic pathways: a systematic method for searching feasible enzyme activity patterns leading to cellular adaptive responses. BMC bioinformatics 10.
38. Vilaprinyo E, Alves R, Sorribas A (2006) Use of physiological constraints to identify quantitative design principles for gene expression in yeast adaptation to heat shock. BMC bioinformatics 7: 184.
39. Navarro E, Montagud A, Fernández de Córdoba P, Urchueguía JF (2009) Metabolic flux analysis of the hydrogen production potential in *Synechocystis sp.* PCC6803. Int J Hydrogren Energy 34: 8828-8838.
40. Lee Y, Chen F, Dixon RA, Voit EO (2011) Integrative analysis of transgenic alfalfa (*Medicago sativa L.*) suggests new postulates for monolignol biosynthesis. PLoS Comput Biol. 7(5):e1002047.
41. Lee Y, Voit EO (2010) Mathematical modeling of monolignol biosynthesis in *Populus xylem*. Math Biosci. 231(1): 49-60.

42. Voit EO, Alvarez-Vasquez F, Hannun YA (2009) Computational Analysis of Sphingolipid Pathway Systems. In: Chalfant C, Del Poeta M, editors. Sphingolipids as Signaling and Regulatory Molecules. Austin, TX: Landes Bioscience.
43. Savageau MA (1969) Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation. J Theor Biol 25: 370-379.
44. Savageau MA (1972) The behavior of intact biochemical control systems. Curr Topics Cell Regulation 6: 63-129.
45. Voit EO (2009) A systems-theoretical framework for health and disease: inflammation and preconditioning from an abstract modeling point of view. Math Biosci 217: 11-18.
46. Penrose R (1955) A generalized inverse for matrices. Proceedings of the Cambridge Philosophical Society 51: 406-413.
47. Bentley WE, Mirjalili N, Andersen DC, Davis RH, Kompala DS (1990) Plasmid-encoded protein: The principal factor in the "metabolic burden" associated with recombinant bacteria. Biotech Bioeng 35: 668-681.
48. Snoep JL, Yomano LP, Westerhoff HV, Ingram LO (1995) Protein burden in *Zymomonas mobilis*: negative flux and growth control due to overproduction of glycolytic enzymes. Microbiology 141: 2329-2337.
49. Smits HP, Hauf J, Müller S, T.J. H, Zimmermann FK, et al. (2000) Simultaneous overepression of enzymes of the lower part of glycolysis can enhance the fermentative capacity of *Saccharomyces cerevisiae*. Yeast 16: 1325-1334.
50. Polisetty PK, Gatzke EP, Voit EO (2008) Yield optimization of regulated metabolic systems using deterministic branch-and-reduce methods. Biotechnology and bioengineering 99: 1154-1169.
51. Voit EO (1992) Optimization in integrated biochemical systems. Biotech Bioeng 40: 572-582.
52. Torres NV, Voit EO, Glez-Alcón C, Rodríguez F (1997) An indirect optimization method for biochemical systems. Description of method and application to ethanol, glycerol and carbohydrate production in *Saccharomyces cerevisiae*. Biotechn Bioeng 55: 758-772.
53. Hatzimanikatis V, Floudas CA, Bailey JE (1996) Analysis and design of metabolic reaction networks via mixed-integer linear optimization. AIChE Journal 42: 1277-1292.
54. Görner W, Durchschlag E, Martinez-Pastor MT, Estruch F, Ammerer G, et al. (1998) Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. Genes Dev 12: 586-597.
55. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, et al. (2000) Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11: 4241-4257.
56. Postmus J, Canelas AB, Bouwman J, Bakker BM, van Gulik W, et al. (2008) Quantitative analysis of the high temperature-induced glycolytic flux increase in *Saccharomyces cerevisiae* reveals dominant metabolic regulation. J Biol Chem 283: 23524-23532.
57. Ye Y, Zhu Y, Pan L, Li L, Wang X, et al. (2009) Gaining insight into the response logic of *Saccharomyces cerevisiae* to heat shock by combining expression profiles with metabolic pathways. Biochem Biophys Res Commun 385: 357-362.
58. Sanchez Y, Taulien J, Borkovich KA, Lindquist S (1992) Hsp104 is required for tolerance to many forms of stress. EMBO J 11: 2357-2364.
59. Davidson JF, Whyte B, Bissinger PH, Schiestl RH (1996) Oxidative stress is involved in heat-induced cell death in *Saccharomyces cerevisiae*. PNAS 93: 5116-5121.
60. Estruch F (2000) Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. FEMS Microbiol Rev 24: 469-486.
61. Cowart LA, Shotwell M, Worley ML, Richards AJ, Montefusco DJ, et al. (2010) Revealing a signaling role of phytosphingosine-1-phosphate in yeast. Molecular systems biology 6: 349.

62. Fonseca LL, Sánchez C, Santos H, Voit EO (2010) Complex coordination of multi-scale cellular responses to environmental stress. Submitted.
63. Hottiger T, Schmutz P, Wiemken A (1987) Heat-induced accumulation and futile cycling of trehalose in *Saccharomyces cerevisiae*. J Bacteriol 169: 5518-5522.
64. Aranda JS, Salgado E, Taillandier P (2004) Trehalose accumulation in *Saccharomyces cerevisiae* cells: experimental data and structured modeling. Biochemical Engineering Journal 17: 119-140.
65. Ervadi-Radhakrishnan A, Voit EO (2005) Controllability of non-linear biochemical systems. Math Biosci 196: 99-123.
66. Entian K.D., Fröhlich KU, Mecke D (1984) Regulation of enzymes and isoenzymes of carbohydrate metabolism in the yeast *Saccharomyces cerevisiae*. Biochim Biophys Acta 799: 181-186.
67. Thevelein JM, Hohmann S (1995) Trehalose synthase: guard to the gate of glycolysis in yeast?. Trends Biochem Sci 20: 3-10.
68. Lotka A (1924) Elements of Physical Biology. Baltimore: Williams and Wilkins; reprinted as 'Elements of Mathematical Biology'. Dover, New York, 1956.
69. Peschel M, Mende W (1986) The Predator-Prey Model: Do we Live in a Volterra World? Berlin: Akademie-Verlag.
70. Volterra V (1926) Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. Mem R Accad dei Lincei 2: 31-113.
71. Hatzimanikatis V, Bailey JE (1996) MCA has more to say. J Theor Biol 182: 233-242.
72. Wu L, Wang W, van Winden WA, van Gulik WM, Heijnen JJ (2004) A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. Eur J Biochem 271: 3348-3359.
73. Shiraishi F, Savageau MA (1992) The tricarboxylic-acid cycle in *Dictyostelium discoideum*. 1. Formulation of alternative kinetic representations. J Biol Chem 267: 22912-22918.
74. Fonseca LL, Chen P-W, Voit EO (2012) Canonical Modeling of the Multi-Scale Regulation of the Heat Stress Response in Yeast. Metabolites 2: 221-241.
75. Ussery DW (2001) DNA Denaturation; Brenner S, H. Miller J, editors. New York: Academic Press.
76. Castells-Roca L, García-Martínez J, Moreno J, Herrero E, Bellí G, et al. (2011) Heat shock response in yeast involves changes in both transcription rates and mRNA stabilities. PLoS ONE 6: e17272.
77. Murata N, Los DA (1997) Membrane fluidity and temperature perception. Plant Physiology 115: 875-879.
78. Jenkins GM, Richards A, Wahl T, Mao C, Obeid L, et al. (1997) Involvement of yeast sphingolipids in the heat stress response of *Saccharomyces cerevisiae*. Journal of Biological Chemistry 272: 32566-32572.
79. Dickson RC, Nagiec EE, Skrzypek M, Tillman P, Wells GB, et al. (1997) Sphingolipids are potential heat stress signals in *Saccharomyces*. Journal of Biological Chemistry 272: 30196-30200.
80. Hannun YA, Obeid LM (2008) Principles of bioactive lipid signalling: lessons from sphingolipids. Nat Rev Mol Cell Biol 9: 139-150.
81. Mandala SM, Thornton R, Tu Z, Kurtz MB, Nickels J, et al. (1998) Sphingoid base 1-phosphate phosphatase: A key regulator of sphingolipid metabolism and stress response. Proceedings of the National Academy of Sciences 95: 150-155.
82. Dickson RC (2008) Thematic review series: sphingolipids. New insights into sphingolipid metabolism and function in budding yeast. J Lipid Res 49: 909-921.
83. Wells GB, Dickson RC, Lester RL (1998) Heat-induced elevation of ceramide in *Saccharomyces cerevisiae* via *de novo* synthesis. Journal of Biological Chemistry 273: 7235-7243.
84. Bagnat M, Keränen S, Shevchenko A, Shevchenko A, Simons K (2000) Lipid rafts function in biosynthetic delivery of proteins to the cell surface in yeast. Proceedings of the National Academy of Sciences 97: 3254-3259.

85. Grimm MOW, Tschäpe JA, Grimm HS, Zinser EG, Hartmann T (2006) Altered membrane fluidity and lipid raft composition in presenilin-deficient cells. Acta Neurologica Scandinavica 114: 27-32.
86. Lingwood D, Simons K (2010) Lipid rafts as a membrane-organizing principle. Science 327: 46-50.
87. Futerman AH, Schuldiner M (2010) Lipids: The plasma membrane code. Nat Chem Biol 6: 487-488.
88. Friant S, Meier KD, Riezman H (2003) Increased ubiquitin-dependent degradation can replace the essential requirement for heat shock protein induction. EMBO J 22: 3783-3791.
89. Chung N, Jenkins G, Hannun YA, Heitman J, Obeid LM (2000) Sphingolipids signal heat stress-induced ubiquitin-dependent proteolysis. Journal of Biological Chemistry 275: 17229-17232.
90. Ferguson-Yankey SR, Skrzypek MS, Lester RL, Dickson RC (2002) Mutant analysis reveals complex regulation of sphingolipid long chain base phosphates and long chain bases during heat stress in yeast. Yeast 19: 573-586.
91. Meier KD, Deloche O, Kajiwara K, Funato K, Riezman H (2006) Sphingoid base is required for translation initiation during heat stress in *Saccharomyces cerevisiae*. Molecular Biology of the Cell 17: 1164-1175.
92. Jenkins GM, Cowart LA, Signorelli P, Pettus BJ, Chalfant CE, et al. (2002) Acute activation of *de novo* sphingolipid biosynthesis upon heat shock causes an accumulation of ceramide and subsequent dephosphorylation of SR proteins. Journal of Biological Chemistry 277: 42572-42578.
93. Mao C, Saba JD, Obeid LM (1999) The dihydrosphingosine-1-phosphate phosphatases of *Saccharomyces cerevisiae* are important regulators of cell proliferation and heat stress responses. Biochem J 342: 667-675.
94. Cowart LA, Shotwell M, Worley ML, Richards AJ, Montefusco DJ, et al. (2010) Revealing a signaling role of phytosphingosine-1-phosphate in yeast. Mol Syst Biol 6: 349.
95. Futerman AH, Hannun YA (2004) The complex life of simple sphingolipids. EMBO Rep 5: 777-782.
96. Voit EO (2012) A First Course in Systems Biology: Garland Science.
97. Merrill AH (2002) *De novo* sphingolipid biosynthesis: A necessary, but dangerous, pathway. Journal of Biological Chemistry 277: 25843-25846.
98. Lee Y, Chen P-W, Voit EO (2011) Analysis of operating principles with S-system models. Mathematical Biosciences 231: 49-60.
99. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, et al. (2000) Genomic expression programs in the response of yeast cells to environmental changes. Molecular Biology of the Cell 11: 4241-4257.
100. Linn SC, Kim HS, Keane EM, Andras LM, Wang E, et al. (2001) Regulation of *de novo* sphingolipid biosynthesis and the toxic consequences of its disruption. Biochemical Society Transactions 29: 831-835.
101. Dickson RC (1998) Sphingolipid functions in *Saccharomyces cerevisiae*: Comparison to mammals. Annual Review of Biochemistry 67: 27-48.
102. Futerman AH, Riezman H (2005) The ins and outs of sphingolipid synthesis. Trends in Cell Biology 15: 312-318.
103. Kolter T, Sandhoff K (2006) Sphingolipid metabolism diseases. Biochimica et Biophysica Acta (BBA) - Biomembranes 1758: 2057-2079.
104. Obeid LM, Linardic CM, Karolak LA, Hannun YA (1993) Programmed cell death induced by ceramide. Science 259: 1769-1771.
105. Alvarez-Vasquez F, Riezman H, Hannun YA, Voit EO (2011) Mathematical modeling and validation of the ergosterol pathway in *Saccharomyces cerevisiae*. PLoS ONE 6: e28344.
106. Sims KJ, Spassieva SD, Voit EO, Obeid LM (2004) Yeast sphingolipid metabolism: clues and connections. Biochemistry and Cell Biology 82: 45-61.

107. Montefusco DJ, Chen L, Matmati N, Lu S, Newcomb B, et al. (2013) Distinct Signaling Roles of Ceramide Species in Yeast Revealed Through Systematic Perturbation and Systems Biology Analyses. Sci Signal 6: rs14-.
108. Matmati N, Metelli A, Tripathi K, Yan S, Mohanty BK, et al. (2013) Identification of C18:1-Phytoceramide as the Candidate Lipid Mediator for Hydroxyurea Resistance in Yeast. Journal of Biological Chemistry 288: 17272-17284.
109. Montefusco DJ, Matmati N, Hannun YA (2014) The yeast sphingolipid signaling landscape. Chemistry and Physics of Lipids 177: 26-40.
110. Stiban J, Tidhar R, Futerman A (2010) Ceramide Synthases: Roles in Cell Physiology and Signaling. In: Chalfant C, Poeta M, editors. Sphingolipids as Signaling and Regulatory Molecules: Springer New York. pp. 60-71.
111. Schorling S, Vallée B, Barz WP, Riezman H, Oesterhelt D (2001) Lag1p and Lac1p Are Essential for the Acyl-CoA–dependent Ceramide Synthase Reaction in Saccharomyces cerevisae. Molecular Biology of the Cell 12: 3417-3427.
112. Mao C, Xu R, Bielawska A, Szulc ZM, Obeid LM (2000) Cloning and Characterization of a Saccharomyces cerevisiae Alkaline Ceramidase with Specificity for Dihydroceramide. Journal of Biological Chemistry 275: 31369-31378.
113. Haak D, Gable K, Beeler T, Dunn T (1997) Hydroxylation of Saccharomyces cerevisiae Ceramides Requires Sur2p and Scs7p. Journal of Biological Chemistry 272: 29704-29710.
114. Dickson RC, Lester RL (2002) Sphingolipid functions in Saccharomyces cerevisiae. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids 1583: 13-25.
115. Sawai H, Okamoto Y, Luberto C, Mao C, Bielawska A, et al. (2000) Identification of ISC1 (YER019w) as Inositol Phosphosphingolipid Phospholipase C inSaccharomyces cerevisiae. Journal of Biological Chemistry 275: 39793-39798.
116. Chen P-W, Fonseca LL, Hannun YA, Voit EO (2013) Coordination of Rapid Sphingolipid Responses to Heat Stress in Yeast. PLoS Comput Biol 9: e1003078.
117. Savageau MA (1969) Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. Journal of Theoretical Biology 25: 365-369.
118. Voit EO (2013) Biochemical Systems Theory: A Review. ISRN Biomathematics 2013: 53.
119. Torres NV, Voit. EO (2002) Pathway Analysis and Optimization in Metabolic Engineering: Cambridge University Press.
120. Voit EO (2000) Computational Analysis of Biochemical Systems. A Practical Guide for Biochemists and Molecular Biologists. Cambridge, United Kingdom: Cambridge University Press.
121. Guillas I, Kirchman PA, Chuard R, Pfefferli M, Jiang JC, et al. (2001) C26-CoA-dependent ceramide synthesis of Saccharomyces cerevisiae is operated by Lag1p and Lac1p. 2655-2665 p.
122. Luttgeharm KD, Cahoon EB, Markham JE (2015) A mass spectrometry-based method for the assay of ceramide synthase substrate specificity. Analytical Biochemistry http://dx.doi.org/10.1016/j.ab.2015.02.016.
123. Levy M, Futerman AH (2010) Mammalian ceramide synthases. IUBMB Life 62: 347-356.
124. Goel G, Chou I-C, Voit EO (2008) System estimation from metabolic time-series data. Bioinformatics 24: 2505-2511.
125. Ramsay JO, Hooker G, Campbell D, Cao J (2007) Parameter estimation for differential equations: a generalized smoothing approach. Journal of the Royal Statistical Society: Series B 69: 741-796.
126. Bock HG (1983) Recent advances in parameter identification techniques for O.D.E. Progress in Scientific Computing 2: 95-121.
127. Maceyka M, Spiegel S (2014) Sphingolipid metabolites in inflammatory disease. Nature 510: 58-67.
128. Burnham KP, Anderson DR (2002) Model Selection and Multi-Model Inference. New York: Springer-Verlag.

129. Akaike H (1973) Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Budapest FC, editors. Second International Symposium on Information Theory. pp. 267-281.
130. Akaike H (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control: AC-19:716-723.

VITA

Po-Wei Chen

Chen was born in Taipei, Taiwan. He attended public schools in Taipei and received a B.A. in Bio-industrial Mechatronics Engineering from National Taiwan University, Taipei, Taiwan in 2006. He served the mandatory military service in Taiwan from 2006 to 2007 as a Second Lieutenant, and joined Dr. Jui-Jen Chou's lab in National Taiwan University as a research assistant from 2007 to 2009 before coming to Georgia Tech to pursue a doctorate in Bioengineering in Dr. Eberhard Voit's lab. When he is not working on his research, Mr. Chen enjoys playing guitar, writing songs and taking photos.