UC San Diego
UC San Diego Previously Published Works

Title

Integration of lipidomics and transcriptomics data towards a systems biology model of sphingolipid metabolism

Permalink

https://escholarship.org/uc/item/2fd169v6

Journal

BMC Systems Biology, 5(1)

ISSN

1752-0509

Authors

Gupta, Shakti
Maurya, Mano R
Merrill Jr, Alfred H
et al.

Publication Date

2011-02-08

DOI

http://dx.doi.org/10.1186/1752-0509-5-26

Supplemental Material

https://escholarship.org/uc/item/2fd169v6#supplemental

Peer reviewed

Integration of lipidomics and transcriptomics data towards a systems biology model of sphingolipid metabolism

Shakti Gupta[1], Mano R Maurya[1], Alfred H Merrill Jr[2], Christopher K Glass[3], Shankar Subramaniam[1,3,4*]

Abstract

Background: Sphingolipids play important roles in cell structure and function as well as in the pathophysiology of many diseases. Many of the intermediates of sphingolipid biosynthesis are highly bioactive and sometimes have antagonistic activities, for example, ceramide promotes apoptosis whereas sphingosine-1-phosphate can inhibit apoptosis and induce cell growth; therefore, quantification of the metabolites and modeling of the sphingolipid network is imperative for an understanding of sphingolipid biology.

Results: In this direction, the LIPID MAPS Consortium is developing methods to quantitate the sphingolipid metabolites in mammalian cells and is investigating their application to studies of the activation of the RAW264.7 macrophage cell by a chemically defined endotoxin, Kdo$_2$-Lipid A. Herein, we describe a model for the C$_{16}$-branch of sphingolipid metabolism (i.e., for ceramides with palmitate as the N-acyl-linked fatty acid, which is selected because it is a major subspecies for all categories of complex sphingolipids in RAW264.7 cells) integrating lipidomics and transcriptomics data and using a two-step matrix-based approach to estimate the rate constants from experimental data. The rate constants obtained from the first step are further refined using generalized constrained nonlinear optimization. The resulting model fits the experimental data for all species. The robustness of the model is validated through parametric sensitivity analysis.

Conclusions: A quantitative model of the sphigolipid pathway is developed by integrating metabolomics and transcriptomics data with legacy knowledge. The model could be used to design experimental studies of how genetic and pharmacological perturbations alter the flux through this important lipid biosynthetic pathway.

Background

Sphingolipids (SL) are categorized as lipids with a sphingoid base backbone [1] that is often derivatized with an amide-linked fatty acid to make ceramides (Cer) and more structurally complex SL with diverse biological functions [2]. SL in essentially every subcategory, from the lipid backbones [3] to complex SL [4], are highly bioactive and play important roles in diseases [5,6]; hence, methods for "lipidomic" analysis of SL and SL metabolism are important for an in-depth understanding of these enigmatic compounds. In recent years, a number of large-scale experimental and bioinformatics projects have begun to address the complexity of the lipidome. Examples include the Lipid Metabolites and Pathways Strategy (LIPID MAPS) Consortium [7], The Lipid Library [8], CYBERLIPID CENTER [9] and Lipid-Bank [10]. In particular, LIPID MAPS has become a comprehensive resource for information on classification, structures and quantitative data on lipids and provides an opportunity for developing quantitative models of lipid synthesis and metabolism thus facilitating a mechanistic and systems-level understanding.

The de novo biosynthesis of SL begins with production of the sphingoid base, which utilizes serine and palmitoyl-coenzyme A (CoA) and various fatty acyl-CoAs to make N-acylsphinganines (dihydroceramides, DHCer) that are desaturated to Cer (N-acylsphingosines) and incorporated into more complex SL such as ceramide 1-phosphate (CerP), sphingomyelin (SM), glucosyl- and galactosyl-ceramide (GlcCer and GalCer) and more

* Correspondence: shankar@ucsd.edu
[1]Department of Bioengineering, University of California, San Diego, 9500 Gilman Dr., La Jolla CA 92093, USA
Full list of author information is available at the end of the article

complex glycosphingolipids [2,11]. Ceramide can also be synthesized by recycling of sphingosine from turnover of SL such as SM [11,12]; furthermore, sphingosine and sphinganine are phosphorylated to sphingosine 1-phosphate (S1P) and sphinganine 1-phosphate (DHSph1P) which are intermediates of sphingoid base degradation [13] and cell signaling molecules [14].

Due to the complexity of this pathway, and the paucity of data for its many metabolites, there are only a few models of SL metabolism available in the literature [15-18]. The LIPID MAPS Consortium [7] has quantified the global changes in lipid metabolites and genes in RAW 264.7 macrophage cells treated with Kdo_2-Lipid A (KLA). KLA is the active component of inflammatory lipopolysaccharide which functions as a selective agonist of Toll-like receptor 4 (TLR4) and mimics bacterial infection. The measurements are carried out over a 24-hour time period and the data is freely available via the LIPID MAPS website [7]. The goal of the work presented here is to develop a predictive kinetic model for SL metabolism using the lipidomics and transcriptomics data from the LIPID MAPS studies. This manuscript is organized as follows: we first briefly discuss the experimental data preprocessing and the methodology used for estimating the rate parameters, then we present the results of parameter estimation, followed by discussion and conclusions.

Methods

Network simplification

A detailed metabolic reaction network was developed using the information available in the literature and the KEGG pathways database [19] (Figure 1). The C_{16}-branch of Cer biosynthesis (i.e., the Cer and DHCer with palmitate as the N-acyl-linked fatty acid) was selected for developing the model because this is a major subspecies for all categories of complex SL in the RAW264.7 cells. VANTED software was used to draw the reaction network [20]. It is common in modeling studies for the network to contain several unmeasured nodes (e.g. metabolites and genes); in our pathway (Figure 1), quantities are known for all of the metabolites and genes except DHGalCer and GalCer (which are present in such small amounts that they are below the limit of detection until the cells are activated) and are expressed as leaf-nodes (last metabolite in each branch) in the network. One of the steps in our matrix-based fast algorithm for parameter-estimation requires experimental data on all metabolites except on the leaf-nodes/ metabolites in the network. A detailed procedure for simplifying the network, if the network contains unmeasured components, is described in our previous work [21]. The leaf-nodes were exempted from the model described in this paper because the reactions leading to

the unmeasured leaf metabolites were combined with the default degradation of their precursors.

Experimental data and preprocessing

The LIPID MAPS Consortium has measured all the major lipids in mouse macrophage RAW 264.7 cells grown in 10% serum and treated with KLA. All metabolites were measured in pmol/μg-DNA units. The time-courses of lipids and genes were measured under two conditions: (a) normal condition designated as control and (b) inflamed condition (stimulation by KLA). Time series comprising of 8 data points at 0, 0.5, 1, 2, 4, 8, 12 and 24 hr were measured with three biological replicates consisting of three technical replicates each. The three technical replicate experiments were performed on the same day with a single batch of cells. In addition, each time course was repeated three times on different days with different batches of cells (biological replicates). For kinetic modeling purposes, outlier points were detected by a t-test and were excluded at each time-point. The resulting data from all the replicates were averaged at each time-point. Data was processed for all metabolites under the treatment with KLA and control conditions.
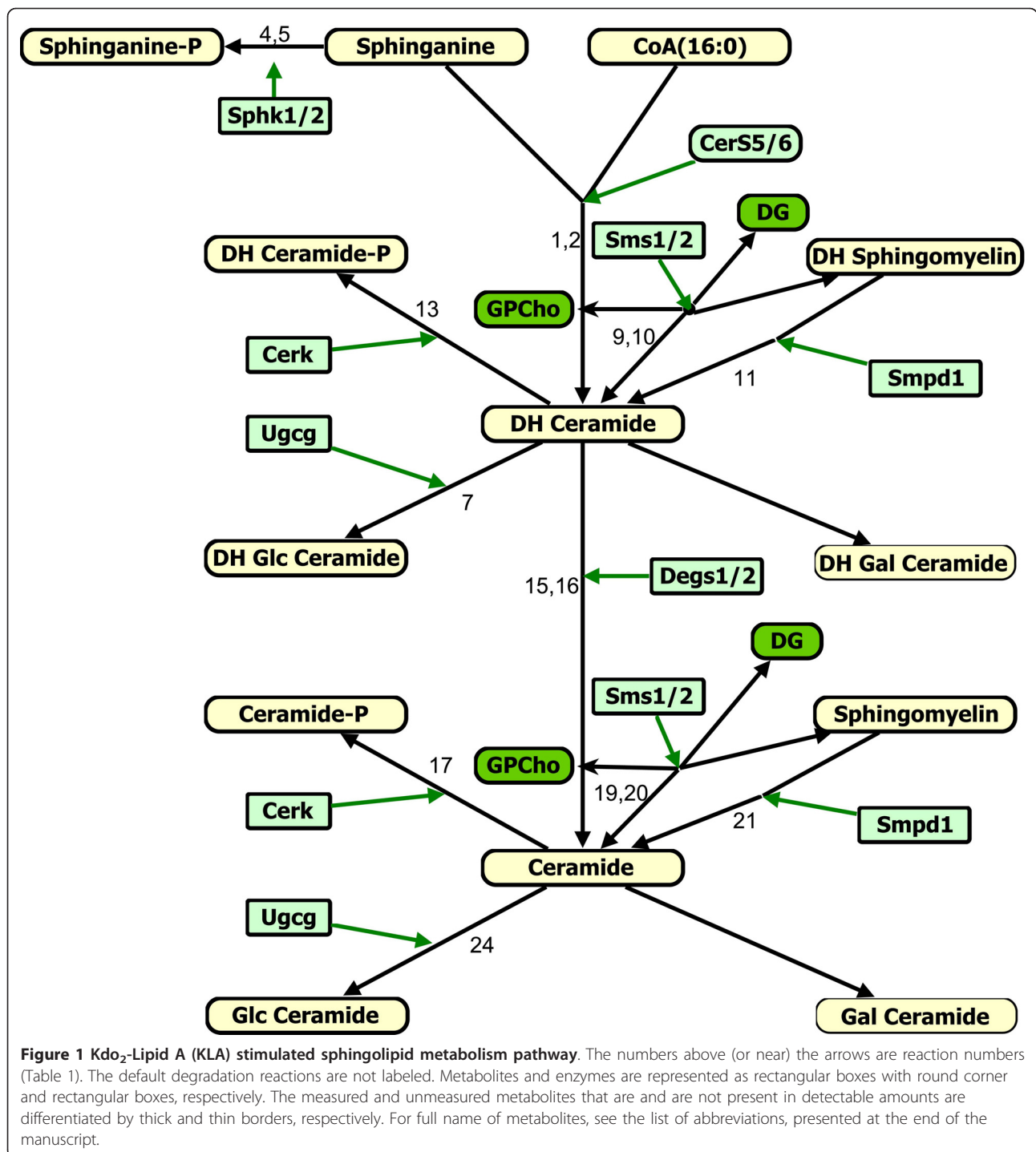
Serine palmitoyltransferase activity *in vitro* was determined using the assay conditions described previously [22] but with [U-^{13}C]palmitoyl-CoA as the labeled substrate. After extraction of the products, the amounts of [^{13}C]3-ketosphinganine were determined by liquid chromatography, electrospray-ionization tandem mass spectrometry [23]. The assay was conducted using 100 μg of protein obtained by sonication of RAW264.7 cells in the buffer used for the assay.

Development of a kinetic model and parameter estimation

We have developed a kinetic model of the SL metabolism. The procedure is similar to the one described in our previous work [21,24], but it is presented here through an example of reactions from the SL network for completeness. The reaction rates were described by linear or law of mass action kinetics with the assumption that for enzymatic reactions, the substrate concentrations are much smaller as compared to the corresponding Michaelis constant, K_m. For example, the following types of reaction schemes and rate expressions were used:

Reaction	Rate expressions
A → B	k * [A]
A + B → A.B	k * [A] * [B]
[Enz]A → B	k * [Enz] * [A]

The enzymes involved in SL metabolism can be regulated at multiple levels from mRNA expression to

Figure 1 Kdo$_2$-Lipid A (KLA) stimulated sphingolipid metabolism pathway. The numbers above (or near) the arrows are reaction numbers (Table 1). The default degradation reactions are not labeled. Metabolites and enzymes are represented as rectangular boxes with round corner and rectangular boxes, respectively. The measured and unmeasured metabolites that are and are not present in detectable amounts are differentiated by thick and thin borders, respectively. For full name of metabolites, see the list of abbreviations, presented at the end of the manuscript.

posttranslational modification. mRNA data on the genes involved in the pathway is available from microarray experiments (Additional file 1 Table S1). However, the corresponding proteomics data is not available. Hence, in our modeling approach, we have captured the effect of temporal changes in transcription and protein levels by utilizing the microarray data with a three hour time

delay as an input to the model. This assumes that the corresponding protein profile is similar to the mRNA time-course with the three hour delay, i.e., $p(t) = g(t-3)$ where $p(t)$ and $g(t)$ denote the level of protein and gene-mRNA, respectively, at time t which is in hrs. A three hour time-delay was chosen based on the general consensus on delay in the protein synthesis from mRNA

[25]. Here is an example of model formulation; the enzymes Sphk1 and Sphk2 catalyze the conversion of DHSph into DHSph1P (Figure 1). This effect was functionally captured through the reactions:

$$[Sphk1]\,DHSph \xrightarrow{k_{f4}} DHSph1P \quad \text{For the activity of Sphk1}$$

$$[Sphk2]\,DHSph \xrightarrow{k_{f5}} DHSph1P \quad \text{For the activity of Sphk2}$$
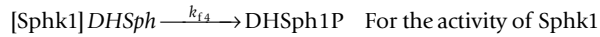
The effective rate of DHSph1P production was written as $(k_{f4} \times [Sphk1] + k_{f5} \times [Sphk2]) \times [DHSph]$. The flux expressions obtained from this scheme were linear in rate parameters and nonlinear in metabolite concentrations. The matrix-based approach to estimate the rate

constants is described below in terms of the reaction numbers labeled in Figure 1 and listed in Table 1. Eq. 1 describes the rate of change of [DHSph1P] and [C_{16} DHGlcCer].

$$\frac{d[DHSph1P]}{dt} = k_{f4}[DHSph][Sphk1] + k_{f5}[DHSph][Sphk2]$$
$$\qquad - k_{f6}[DHSph1P] \qquad\qquad\qquad (1)$$
$$\frac{d[C_{16}\ DHGlcCer]}{dt} = k_{f7}[C_{16}\ DHCer][Ugcg] - k_{f8}[C_{16}\ DHGlcCer]$$

where the rate constants k_i (i = f4, f5, f6, f7, f8) are as defined in Table 1.

If the metabolite concentrations were known and the rate parameters were unknown, then the ordinary

Table 1 The estimated values of the rate parameters in the model of sphingolipid metabolism

No.	Reactions	Parameter Names	Values
1	DHSph + CoA16 + CerS6 → C_{16} DHCer	k_{f1}	1.30E+01 ± 1.01E-02
2	DHSph + CoA16 → C_{16} DHCer	k_{f2}	5.62E-02 ± 1.30E-02
3	C_{16} DHCer →	k_{f3}	1.11E-03 ± 3.20E-03
4	DHSph + Sphk1 → DHSph1P	k_{f4}	1.49E-04 ± 1.11E-03
5	DHSph + Sphk2 → DHSph1P	k_{f5}	1.61E-02 ± 8.38E-04
6	DHSph1P →	k_{f6}	5.28E-01 ± 1.83E-02
7	C_{16} DHCer + Ugcg → C_{16} DHGlcCer	k_{f7}	1.81E-02 ± 2.71E-03
8	C_{16} DHGlcCer →	k_{f8}	2.99E-01 ± 3.25E-02
9	C_{16} DHCer + Sms1 + C_{16} GPCho ←→	k_{f9}	1.15E-01 ± 6.85E-03
	C_{16} DHSM + Sms1 + C_{16} DG	k_{b9}	1.7176 ± 5.47E-03
10	C_{16} DHCer + Sms2 + C_{16} GPCho ←→	k_{f10}	3.88E-01 ± 1.02E-02
	C_{16} DHSM + Sms2 + C_{16} DG	k_{b10}	8.63E-01 ± 1.23E-02
11	C_{16} DHSM + Smpd1 → C_{16} DHCer	k_{f11}	1.07E-01 ± 5.62E-03
12	C_{16} DHSM →	k_{f12}	2.49E-02 ± 3.01E-03
13	C_{16} DHCer + Cerk → C_{16} DHCerP	k_{f13}	5.00E-05 ± 7.20E-06
14	C_{16} DHCerP →	k_{f14}	8.25E-02 ± 6.57E-03
15	C_{16} DHCer + Degs1 → C_{16} Cer	k_{f15}	9.07E-03 ± 9.56E-03
16	C_{16} DHCer → C_{16} Cer	k_{f16}	8.03E-01 ± 1.39E-02
17	C_{16} Cer + Cerk → C_{16} CerP	k_{f17}	1.29E-04 ± 7.64E-06
18	C_{16} CerP →	k_{f18}	1.12E-01 ± 4.05E-03
19	C_{16} Cer + Sms1 + C_{16} GPCho ←→	k_{f19}	6.50E-02 ± 3.13E-02
	C_{16} SM + Sms1 + C_{16} DG	k_{b19}	4.72E-02 ± 8.67E-03
20	C_{16} Cer + Sms2 + C_{16} GPCho ←→	k_{f20}	4.48E-01 ± 2.50E-02
	C_{16} SM + Sms2 + C_{16} DG	k_{b20}	1.3343 ± 2.52E-02
21	C_{16} SM + Smpd1 → C_{16} Cer	k_{f21}	1.91E-05 ± 1.98E-02
22	C_{16} SM →	k_{f22}	9.79E-03 ± 1.28E-03
23	C_{16} Cer →	k_{f23}	0.00E+00 ± 2.55E-03
24	C_{16} Cer + Ugcg → C_{16} GlcCer	k_{f24}	2.39E-02 ± 3.05E-03
25	C_{16} GlcCer →	k_{f25}	4.58E-03 ± 3.43E-03

Microarray data was used for protein concentration.

X → means default degradation of the metabolite X.

Unit of first order reaction is 1/hr.

Unit of second order reaction is 1/hr when it involves protein as a modifier as we have used fold change data with respect to control for these variables.

Unit of third order reaction is µg DNA/(pmol * hr) when it involves protein as a modifier.

Unit of third order reaction is µg DNA/(Ratio Int * hr) when it involves DG and protein as modifiers.

differential equations (ODEs) can be rearranged in a matrix format as shown in Eq. 2.

$$
\underbrace{\begin{bmatrix} \dfrac{d[DHSph1P]}{dt} \\ \dfrac{d[C_{16}DHGlcCer]}{dt} \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} [DHSph][Sphk1] & 0 \\ [DHSph][Sphk2] & 0 \\ -[DHSph1P] & 0 \\ 0 & [C_{16}DHCer][Ugcg] \\ 0 & -[C_{16}DHGlcCer] \end{bmatrix}^{T}}_{\mathbf{X}} * \underbrace{\begin{bmatrix} k_{f4} \\ k_{f5} \\ k_{f6} \\ k_{f7} \\ k_{f8} \end{bmatrix}}_{\mathbf{b}} \quad (2)
$$

The coefficients in the matrix **X** are linear/nonlinear functions of metabolite and gene concentrations. All the equations used in the simulation are listed in Appendix A. **X** is completely defined. The left hand side of the equations (matrix **Y**) was computed using discretization and the experimental data (Eq. 3).

$$
\left. \frac{dx}{dt} \right|_{t=t_k} = \frac{x|_{t=t_k} - x|_{t=t_{k-1}}}{t_k - t_{k-1}} \quad (3)
$$

Eq. 2 contains known matrices **X** and **Y**, and the only unknown in this equation is the rate-constant vector **b**. The constrained least-squares approach (Matlab® [26] function *lsqlin*) was used to solve **b**. *lsqlin* optimized the solution with objective function (Eq. 4) with additional constraints that all parameter values have to be positive.

$$
\min \| \mathbf{Y} - \mathbf{Xb} \|^2 \text{ where } \mathbf{Y} \text{ and } \mathbf{Xb} \text{ are: } \left. \frac{dy}{dt} \right|_{exp} \text{ and } \left. \frac{dy}{dt} \right|_{pred} \quad (4)
$$

The estimated values of the parameters were further refined by using generalized constrained nonlinear optimization (Matlab® function *fmincon*) where the objective (Eq. 5) was to minimize the weighted fit-error between the experimental and predicted metabolite concentrations. The algorithm of *fmincon* does not require a matrix form. Thus, numerical integration was used (e. g. Matlab® function *ode23*) to simulate the system to circumvent the discretization error. The combined use of *lsqlin* and *fmincon* made the overall process computationally efficient. The objective function for use with *fmincon* was:

$$
\min_{K, X_0} \sum_{i=1}^{nsp} \left(\sum_{j=1}^{nt} (y_{i,j,exp} - y_{i,j,pred}(K, X_0))^2 \right) \quad (5)
$$

K : prameters (rate constants)
X_0: Initial conditions (species concentrations)

where, nt is the number of time-points and nsp is the number of species.

The initial concentrations of the metabolites were also optimized in a narrow range around the experimental values. When data on more than one condition was available, then all the data was used to compute the fit-error by simulating the model several times individually and minimizing the objective function collectively.

Estimation of uncertainty in the optimized parameters

The variation among the different technical and biological replicates of lipid and gene data should be accounted for in the estimated values of the parameters. Hence, an uncertainty analysis was performed on the parameters. Their standard error of mean (SEM) was computed as follows:

1. Compute the SEM in the lipid and gene data at each time point (Additional file 1 Table S1b).
2. Create a candidate data set for parameter estimation by generating nsp x m random matrix utilizing the normal distribution, scale it with the corresponding SEM and add the scaled-value matrix to the mean-value data on lipids and genes (Additional file 1 Table S1a).
3. Estimate the parameters using the candidate data set thus producing a one parameter-value set.
4. Repeat steps 2-3 k times to generate k parameter-value sets (k = 10 in our simulation).
5. Compute the SEM for each parameter across the k sets.

Results
Fit to experimental data

ODEs were generated for all metabolites in the network and effective rate constants were estimated for the simplified SL reaction network using the approach described in the "Methods" section. Table 1 lists the reactions and the corresponding estimated reaction-rate parameters included in the model. Lipid metabolism and signaling are complex processes and the mechanisms involved are only partially known. KLA treatment generated increases in most sphingolipids in RAW cells. The increase in sphinganine (DHSph), which doubled in 4 hours, followed by increases in downstream metabolites, including *N*-acyl-sphinganines (dihydroceramides, DHCer) after a lag of approximately 2 to 4 hours indicted the induction of *de novo* sphingolipid biosynthesis. To account the effect of KLA signaling, the time-courses of the concentrations of DHSph, CoA16, C_{16} DG, C_{16} GPCho and microarray data were used as input to the network (Additional file 1 Figure S1, Table S1). Linear interpolations of these time-courses were used in the integration. The decline in SL control experiment data suggested that the control experiments were not at steady state [7]. The assumption of steady-state was circumvented by including (fitting) data obtained in two experimental scenarios, namely, the treatment with KLA

and the corresponding control data set, during parameter-estimation. The microarray data was used to represent the corresponding protein concentration with 3 hr time delay. For microarray data, fold change with respect to the control was used in the simulation corresponding to the treatment experiment (Additional file 1 Table S1). A fold-value of 1 was used in the simulation corresponding to the control experiment. CerS5 and Degs2 did not show significant change with respect to the control in microarray data, thus the corresponding reactions (reaction 2 and 14) were written without these enzymes. The omission of these enzymes indicated their constant activity in these reactions. The simplified model is a reliable predictive model as evidenced by the good fit to experimental data for most metabolites (Figure 2). For C_{16} DHCer (top-left panel), the fit is good up to 8 hrs after which it deteriorates. One possible reason for this deterioration in the fit is the discrepancy between mRNA and protein levels for the enzymes CerS5/6.

There were two problems in the optimization of parameters: (1) the wide range of metabolite concentrations, and (2) irregular time intervals (longer intervals at later times). The concentration of metabolites varied between 0.01 pmol/μg DNA (for C_{16} DHCerP) and 700 pmol/μg DNA (for C_{16} SM). Due to the orders of magnitude difference in the metabolite concentrations, the fits were biased towards metabolites with high concentration and resulted in a poor fit for the metabolites with lower concentrations. To resolve this issue, the experimental values and predicted values for each metabolite were scaled/normalized by its maximum experimental value. Then, the sum of squares of the normalized fit-error on all metabolites was minimized. This scaling, essentially, normalized the maximum experimental concentrations to 1 for all metabolites and resulted in approximately equal weight for all metabolites. Further, the data measured at irregular time intervals (12 hr difference in last two measurements) also created a problem in optimization and
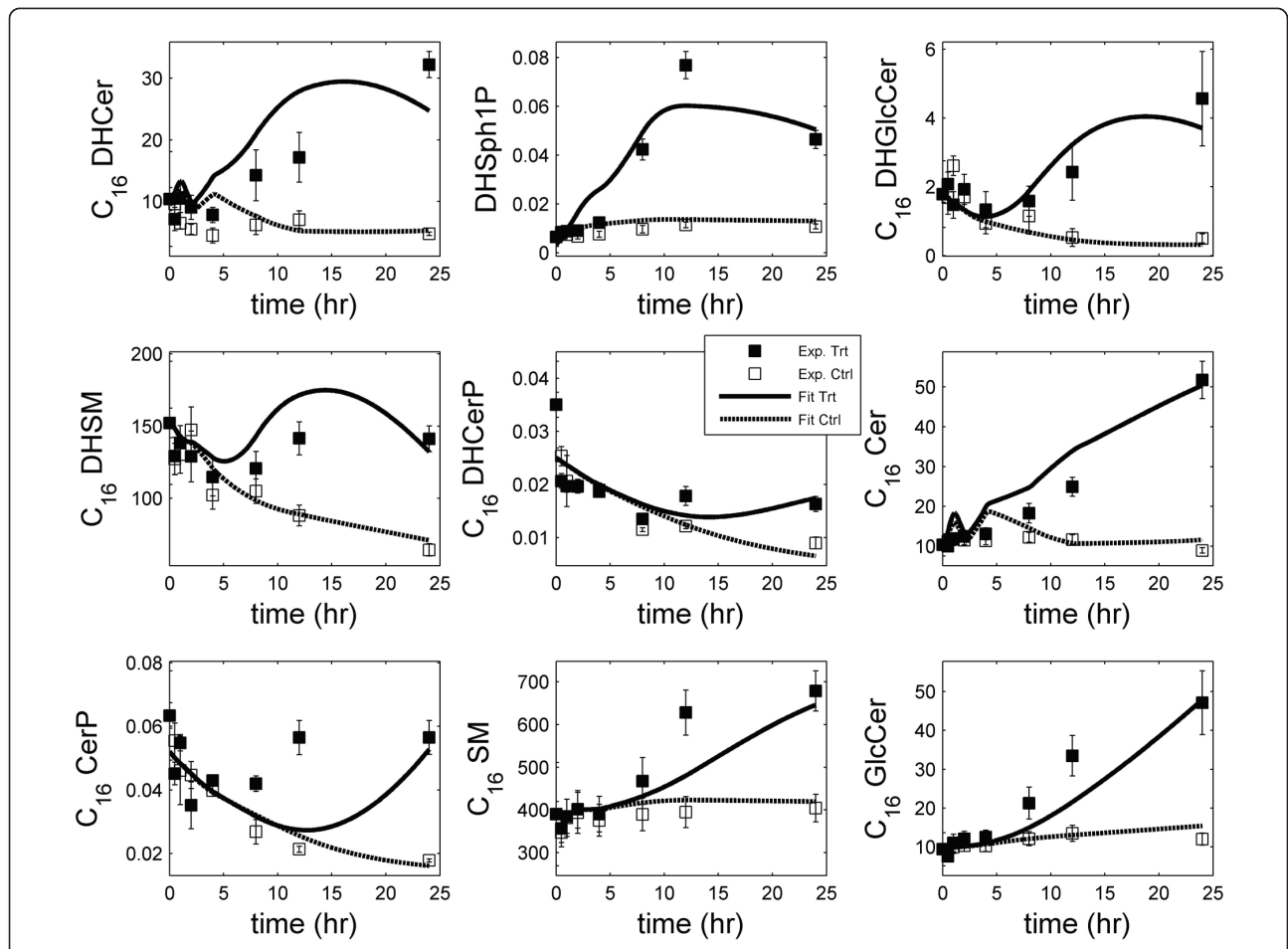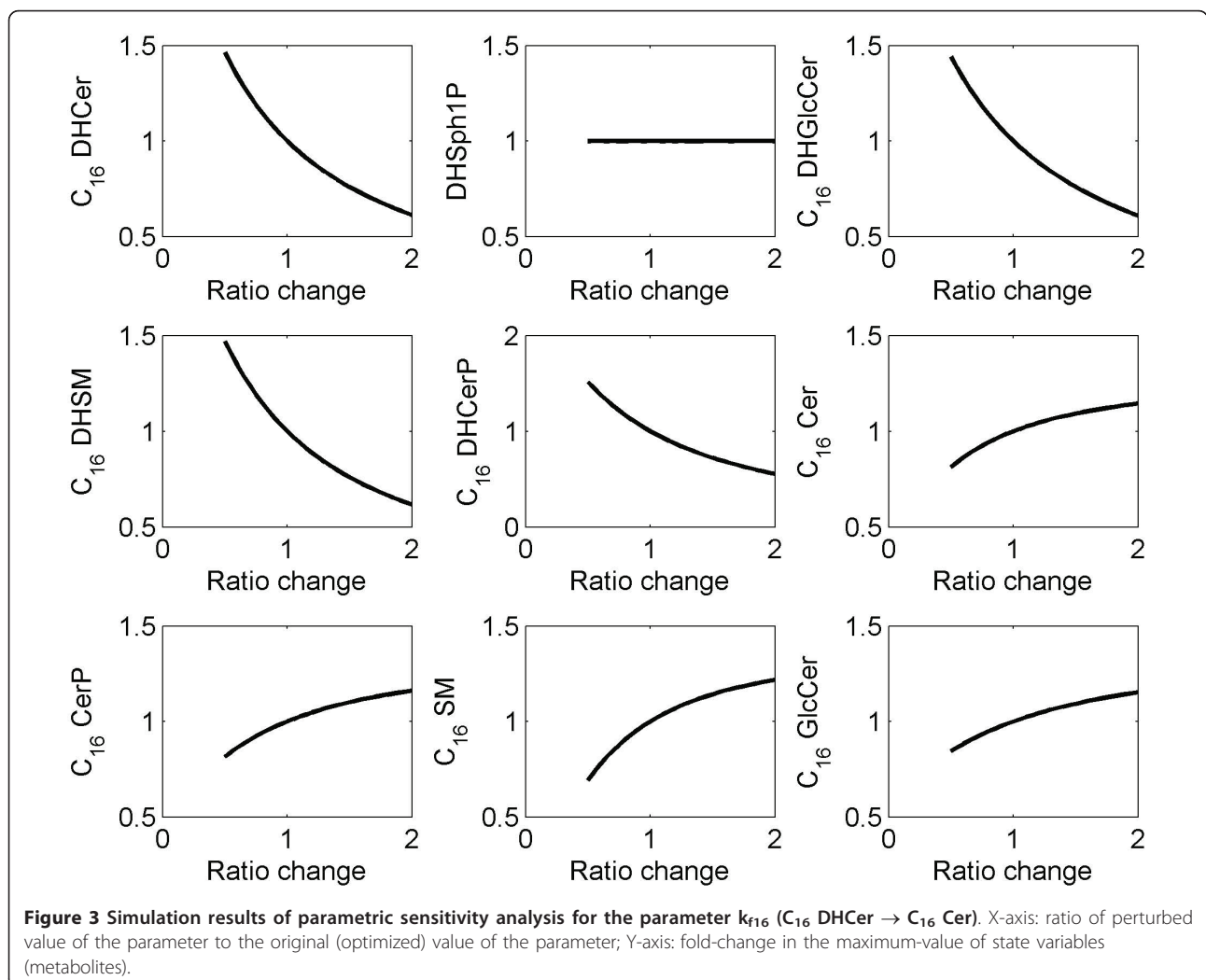

Figure 2 Simulation results of kinetic modeling of sphingolipids network: Fit of the predicted response (control and treatment with KLA) to the corresponding experimental data. In the legend, 'Ctrl' refers to control and 'Trt' refers to KLA treatment of RAW 264.7 cells. The error-bars shown on the experimental data are standard-error of mean.

led to relatively poor fit at later time points. From parameter-estimation view-point, the measurements should be made at equal intervals so that equal weight is assigned to the entire time-course. To account for this, the point-wise-error was scaled by the 1/4 power of the length of the corresponding time-interval, resulting in a higher weight for later time points. Consequently, the quality of fit for the later time-points was improved. For most time points, the difference between the predicted and experimental data was within the standard-error of mean (Figure 2). The good fit was obtained for all metabolites under both treatment and control conditions.

Parametric sensitivity and time-scale analysis

Parametric sensitivity analysis: Parametric sensitivity analysis was performed by varying all parameters (one at a time) by two-fold up and down from its original (optimized) value. The sensitivity of each metabolite was studied by plotting *the fold-change at its maximum concentration as compared to the maximum concentration corresponding to the original value of the parameter* vs. *the ratio-change in the value of the perturbed parameter* (Figure 3). The numerical values of the sensitivity, i.e. the slope of the plots at the optimized value of the parameter, are listed in Additional file 1 Table S2. For each parameter and each metabolite, monotonic increase, decrease or no change was observed depending upon the respective location of the parameter and the metabolite chosen in the network. For example, an increase in the parameter k_{f16} (C_{16} DHCer → C_{16} Cer) (Figure 3) produced a decrease in all upstream metabolites except DHSph1P (Sphinganine-1P). An increase is observed for Cer (Figure 3 sensitivity ~ 0.25) and its downstream products such as Cer-1P, GlcCer, Gal-Cer and SM. This is meaningful mechanistically because the increase in k_{f16} increases the flux of the reaction C_{16} DHCer → C_{16} Cer. If the level of C_{16} DHCer were not to change much, then one would expect almost proportional increase in C_{16} Cer (sensitivity ~ 1). However, this is not

Figure 3 Simulation results of parametric sensitivity analysis for the parameter k_{f16} (C_{16} DHCer → C_{16} Cer). X-axis: ratio of perturbed value of the parameter to the original (optimized) value of the parameter; Y-axis: fold-change in the maximum-value of state variables (metabolites).

true because the level of C_{16} DHCer is reduced (sensitivity ~ -0.64). This results in a net sensitivity value of less than 1. Similarly, an increase in k_{f16} results in decreased DHCer-1P, DHGlcCer, and DHSM levels. These metabolites are products of DHCer and hence the sensitivities are consistent with the structure of the biochemical reaction network (Figure 1). Small to moderate sensitivities (Additional file 1 Table S2) suggest that the biochemical system is robust with respect to parametric perturbations.

Uncertainty analysis on the parameters: The SEM of the parameters is calculated as described in the Methods section and the results are reported in Table 1. Overall, the parameters are well-behaved and the parameter-estimation procedure is reliable because the fractional SEM for the lipid and gene data and the fractional SEM for most of the parameters are of the same order (20-30%).

Time-scale analysis: Time-scale is an important intrinsic property of dynamical biological systems. While the time-scale of metabolites, at which they evolve, can be gleaned by analyzing several time-courses under different conditions, a more systemic picture can be obtained through eigenvalue and eigenvector analysis of the Jacobian matrix of ordinary differential equations at the steady-state conditions. In our computations, the steady-state was identified by simulating the system corresponding to the control condition (no stimulus) for a long time (t = 1000 hr). The Jacobian matrix was computed through numerical differentiation of the expressions on the right hand sides of the ODEs with respect to the state variables. The eigenvalues were split into three broad ranges. For each eigenvalue, the metabolites with substantial contribution to the corresponding eigenvector were identified. Depending upon the eigenvalues and metabolites significantly contributing to the corresponding eigenvectors, these metabolites have been divided into three categories as listed in Table 2. When a metabolite contributed significantly in two eigenvectors spanning in two different eigen value ranges, the metabolite was assigned to the smaller eigen value because the fast manifold only determines its initial transients and the slow manifold governs the later response leading to steady state. A comparison of Table 2 and Figure 2 shows that the time scale of the metabolite is dependent on its location (whether it is leaf node metabolite or intermediate metabolite) and its concentration. The medium time-scale metabolites (column 2 in Table 2) are the leaf node metabolites having low concentration (~ 0.01-1 pmole/μg of DNA); the slow time-scale metabolites (column 3 in Table 2) are the leaf node metabolite with high concentration (~ 10^2-10^3 pmole/μg of DNA). DHCer and Cer have fast time-scale because of their intermediate location in the network, moderate concentration (~10 pmole/μg of DNA) and high flux through these nodes in the production of SM.

Discussion

This study has used the large data sets from mass spectrometric measurements of the SL and the microarray data of the mRNAs of RAW264.7 cells generated by the LIPID MAPS Consortium to evaluate a model for SL metabolism.

Importance of including the transcriptomics data and the data on fatty acyl-CoA

In this study, we have included the microarray data for the lipid-related genes using a time-delay of three-hours to account for the time for mRNA translation and protein translocation. The use of mRNA data for kinetic modeling was motivated by good correlations between specific genes and its metabolic products in the sphingolipid pathway [27]. In general, protein levels follow the qualitative profile of mRNA with appropriate time-delays. However, we note the potential caveat that the protein levels for some proteins may not be even qualitatively similar to the transcriptional levels of their genes. The discrepancy of mismatch between the mRNA and protein profiles can arise due to several factors such as dependence of mRNA translation on ribosome binding site (RBS) sequences, post translational modification of the protein, protein translocation and its stability. This may introduce errors in the estimated values of the kinetic parameters.

To delineate the importance of the transcriptional data, we first developed the mathematical model without using the gene/protein data. The rate parameters were estimated using the lipidomics data alone. We visualized the fit to experimental data (equivalent of the plots shown in Figure 2; data not shown). A reasonable fit was obtained for all metabolites except C_{16}-DHCer. To resolve disagreement in the shape of DHCer, we identified the reactions in which these lipids were consumed or produced. We also analyzed the time-course of the mRNA levels for the genes related to these reactions. Most of the genes exhibited differential regulation at later time points (i.e., the ratio of data with KLA treatment to control being significantly different from 1). Among several differentially regulated genes, the prominent were CerS6 (reaction 1), Smpd1 (reactions 9 and

Table 2 Results of eigenvalue based time-scale analysis of the metabolites

Fast (~1 hr)	Medium (~10 hr)	Slow (~100 hr)
C_{16} DHCer	DHSph1P	C_{16} SM
C_{16} Cer	C_{16} DHGlcCer	C_{16} GlcCer
	C_{16} DHSM	
	C_{16} DHCerP	
	C_{16} CerP	

17), Degs1 (reaction 13) and Ugcg (reaction 20). The up-regulation and involvement of CerS6 in the production of DHCer suggested that the gene data must be included in the network to capture the DHCer dynamics. After adding the gene data with delay, we observed substantial improvement in the fit to the experimental data for DHCer (Figure 2).

The profile of palmitoyl-CoA16 also increases monotonically during 0-24 hrs (Additional file 1 Figure S1). Even though [DHSph] increases till 4 hr and comes back to control level at 24 hr, the influx to DHCer is maintained throughout 0-24 hr because of increase in palmitoyl-CoA and CerS6 at later time points. Due to this reason, DHCer shows a monotonically increasing profile during 0-24 hr. Thus, the profiles of CerS6 and palmitoyl-CoA are important to obtain a good fit on DHCer.

Rate parameters for the enzymes

We compared the combined rate constant for the enzymes CerS5/CerS6 (pmol/hr/µg-DNA) with an estimation of the maximum flux through the de novo biosynthetic pathway based on the activity of serine palmitoyltransferase assayed in vitro with optimal concentrations of substrates, which was 30 + 1 pmole/min/mg protein for control (unstimulated) RAW264.7 cells. To convert the estimated values to the same units, we used estimations (by measurement) for the relationships between DNA, cell number and protein amount of ~3 µg DNA/10^6 RAW264.7 cells, and 10^6 cells have ~0.25 mg protein. Thus, the computed value for C16-SL biosynthesis by CerS5/CerS6 is ~1 pmol/min/mg protein, which is about one order of magnitude lower than the theoretical maximum rate of sphinganine production by the cells. This might mean that the calculated value is low due to inaccuracies in some of the modeling approximations, including the use of linear kinetics instead of Michaelis-Menten kinetics; however, the differences might be real because serine palmitoyltransferase has usually not been found to be operating at V_{max} because its substrates are not saturating [28,29] and other CerS also contribute to the utilization of the sphinganine that is produced de novo. In addition, sphinganine and sphinganine 1-phosphate are elevated to some extent in RAW264.7 cells, which implies that CerSs are not trapping all of the sphinganine that is made.

Similarity in the rate parameters for same gene/enzyme involved in different reactions

Structurally, DHCer and Cer differ by one double bond (Figure 4). However, both are converted to their corresponding derivatives by four genes, namely Ugcg, Cerk and Sms1/2 (Figure 1). To check the effects of this double bond on the rate parameter, we compared the rate parameters for the above genes/enzymes in their reactions.
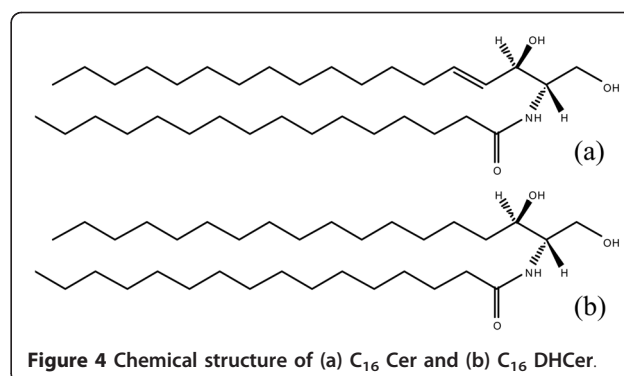
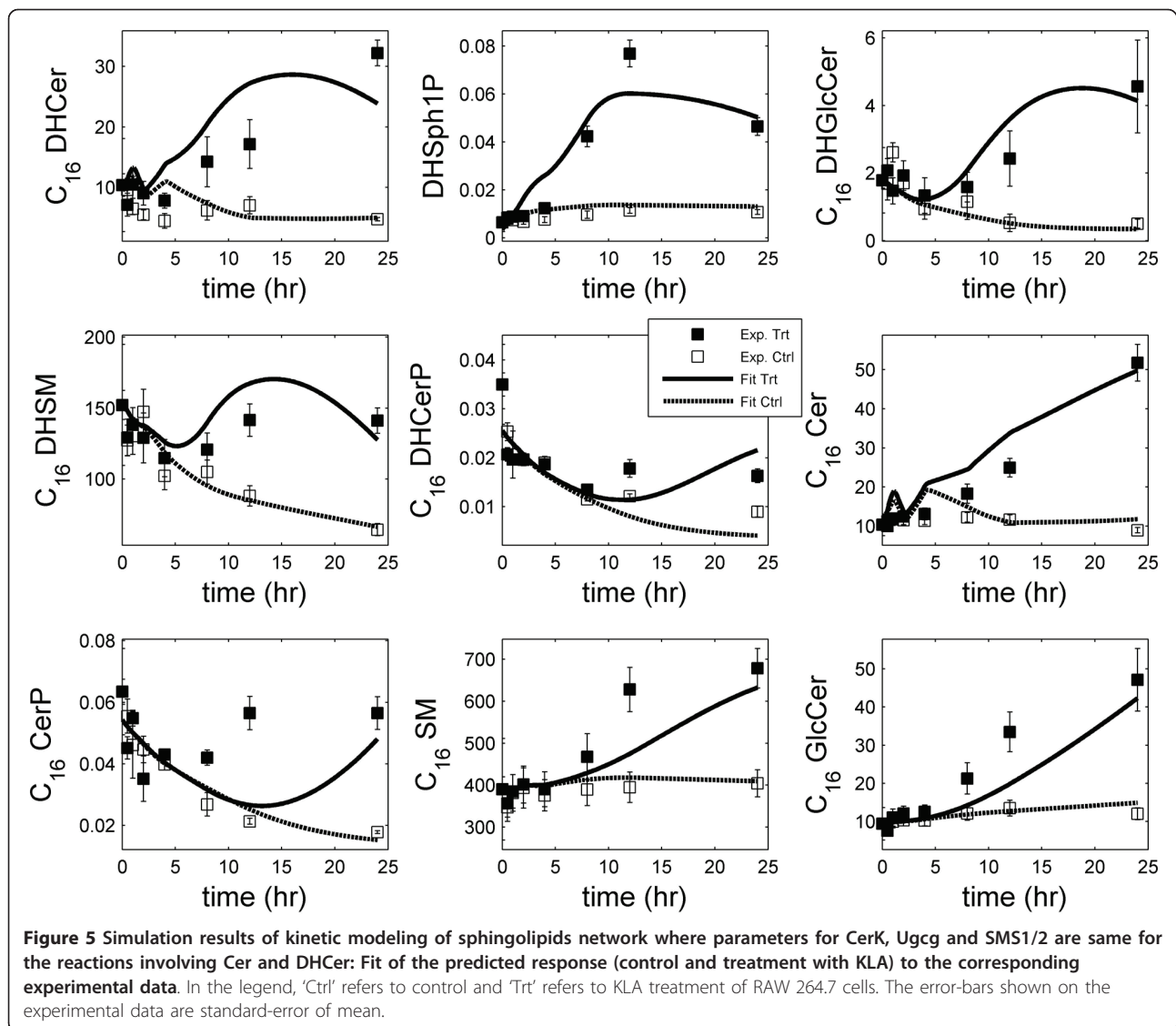**Figure 4 Chemical structure of (a) C$_{16}$ Cer and (b) C$_{16}$ DHCer**.

The rate parameter for SMS2 (k_{f10} and k_{f20}) and Ugcg (k_{f7} and k_{f24}) was found approximately similar for the both reactions involving DHCer and Cer. For Sms1 and Cerk, the rate parameters for the two reactions differed by a factor of 2 (Table 1). To further test whether we can find the common rate constants for these enzymes in their reactions, we carried out the optimization using same rate constant for these enzymes in their reactions. We were able to get reasonably good fit for all of the metabolites (Figure 5). In the modified optimization, the values obtained for these rate-parameters were between the corresponding values for the two reactions obtained in the original optimization (Additional file 1 Table S3). This result suggests that the affinities of these enzymes are similar for both the substrates Cer and DHCer.

Consideration of different chain-lengths

C$_{16}$ sphingolipids are used in this model because (1) these are present in higher amounts than the SL of larger chain-lengths, and (2) they showed significant differences between the treatment and control experiments in response to KLA. We can extend our model to include the metabolites with higher chain-lengths (e.g. C$_{18}$, C$_{20}$, C$_{22}$ and C$_{24}$). The reactions in the existing model (C$_{16}$-based) will serve as a template that can be instantiated for higher chain-lengths. To estimate the rate parameters for the entire model, the parameters for the C$_{16}$-based model can provide bounds for the rate parameters in the overall model for the similar types of reactions. For example, the rate constants for the reactions C$_n$ DHCer → C$_n$ DHGlcCer (where n = 18, 20, 22 and 24), can be constrained with a factor of 0.5 to 2 of the rate constant for the reaction C$_{16}$ DHCer → C$_{16}$ DHGlcCer in the model developed here.

Use of the model for in-silico perturbation experiments

The model can be used to perform knockdown (KD) or other perturbation experiments in-silico. Such computational studies provide useful insights into the behavior of the system which can add in designing the actual perturbation experiments. These simulations can assist in finding

Figure 5 Simulation results of kinetic modeling of sphingolipids network where parameters for CerK, Ugcg and SMS1/2 are same for the reactions involving Cer and DHCer: Fit of the predicted response (control and treatment with KLA) to the corresponding experimental data. In the legend, 'Ctrl' refers to control and 'Trt' refers to KLA treatment of RAW 264.7 cells. The error-bars shown on the experimental data are standard-error of mean.

the propagation of the effects of KD in the network and lead to a better design of experiments (e.g. when the measurements should be made). Simulation can also suggest whether a high level of KD is needed such that the differences between the KD and control scenarios would be statistically significant. The data generated from new perturbation experiments can be used to further refine the model. For the enzymes, the KD perturbations can be simulated by decreasing the corresponding rate parameter because amount of the (active) enzyme directly affects the rate parameters. The KD perturbation simulation results will be similar to the sensitivity analysis results on the corresponding rate parameter. For example, the effect of 50% KD of the enzyme for the reaction 16 (Figure 1 Degs2), can be predicted from the simulation for sensitivity analysis in which the value of the parameter k_{f16} is reduced by 50% (Figure 3).

The case of knockdown of Cerk (cermide kinase) is also interesting due to its direct effect on Ceramide-1-phosphate (CerP) which inhibits apoptosis and induces cell growth [30]. The result of sensitivity analysis for the corresponding rate parameter, k_{f17} (C_{16} Cer + Cerk $\rightarrow$ C_{16} CerP), is shown in Additional file 1 Figure S2. Decrease in k_{f17} results in a corresponding decrease in the CerP level. These changes have also been observed in recent experiments validating our parametric sensitivity analysis and the predictive ability of the model [31]. CerP has also been implicated in regulating the homeostasis of calcium [32] thereby affecting the activity of several signaling pathways. CerP and Cerk mediate the effect of cytokines to activate cytosolic phospholipase A2 (cPLA$_2$) and cyclooxygenase 2 (COX-2), resulting in increased production of prostaglandin E2 (PGE2), a mediator of inflammation [33]. Hence, it has been

hypothesized that Cerk could be a potential target for anti-inflammatory drugs [34,35].

Conclusion

Use of systems biology approaches is becoming more common in the study of lipids to elucidate their functions and roles in human health and diseases such as arthritis and cancer. Systems biology has already been recognized as an indispensable tool in pathway-based drug discovery. Here we have applied a matrix-based approach to develop a dynamic model of SL metabolism by integrating legacy information on the lipid pathways with novel experimental data. The metabolic pathway was reconstructed using information from the KEGG database and the existing literature. Based upon the network map reconstructed, we have developed an ordinary differential equations-based mathematical model. Parameter-estimation used a two-step approach. In the first step, a matrix-based approach provided an initial guess. The parameter-values were further refined in the second step. The resulting model fitted the experimental data well for all species and demonstrated that the integrated metabolic and signaling network and the experimental data are consistent with each other. The robustness of the model parameters was also validated through sensitivity analysis. Though we have used this two-step approach previously and applied it to eicosanoid lipid pathway, the major distinction lies in its application to the SL metabolic pathway and the integration of transcriptomic data with the metabolomic data along with legacy knowledge to develop the kinetic model. Previous computational models of sphingolipid metabolism were for non-mammalian systems in which only a few metabolites were measured as compared to the total number of metabolites in the reaction network. In comparison, our model is based on a large amount of experimental time-course data where the concentrations of most metabolites and mRNA levels of genes in the network are measured. This provides a more context-specific model for RAW cells in particular and for mammalian cell systems in general.

Appendix A

The flux expressions for the reactions shown in Figure 1 are as follows:

$v_1 = k_{f1}[DHSph][CoA16][CerS6]$
$v_2 = k_{f2}[DHSph][CoA16]$
$v_3 = k_{f3}[C_{16} DHCer]$
$v_4 = k_{f4}[DHSph][Sphk1]$
$v_5 = k_{f5}[DHSph][Sphk2]$
$v_6 = k_{f6}[DHSph1P]$
$v_7 = k_{f7}[C_{16} DHCer][Ugcg]$
$v_8 = k_{f8}[C_{16} DHGlcCer]$

$v_9 = k_{f9}[C_{16} DHCer][Sms1][C_{16} GPCho] - k_{b9}[C_{16} DHSM][Sms1][C_{16} DG]$
$v_{10} = k_{f10}[C_{16} DHCer][Sms2][C_{16} GPCho] - k_{b10}[C_{16} DHSM][Sms2][C_{16} DG]$
$v_{11} = k_{f11}[C_{16} DHSM][Smpd1]$
$v_{12} = k_{f12}[C_{16} DHSM]$
$v_{13} = k_{f13}[C_{16} DHCer][Cerk]$
$v_{14} = k_{f14}[C_{16} DHCerP]$
$v_{15} = k_{f15}[C_{16} DHCer][Degs1]$
$v_{16} = k_{f16}[C_{16} DHCer]$
$v_{17} = k_{f17}[C_{16} Cer][Cerk]$
$v_{18} = k_{f18}[C_{16} CerP]$
$v_{19} = k_{f19}[C_{16} Cer][Sms1][C_{16} GPCho] - k_{b19}[C_{16} SM][Sms1][C_{16} DG]$
$v_{20} = k_{f20}[C_{16} Cer][Sms2][C_{16} GPCho] - k_{b20}[C_{16} SM][Sms2][C_{16} DG]$
$v_{21} = k_{f21}[C_{16} SM][Smpd1]$
$v_{22} = k_{f22}[C_{16} SM]$
$v_{23} = k_{f23}[C_{16} Cer]$
$v_{24} = k_{f24}[C_{16} Cer][Ugcg]$
$v_{25} = k_{f25}[C_{16} GlcCer]$

The differential equations describing the rate of change of metabolite concentrations are:

$d[C_{16} DHCer]/dt = v_1 + v_2 - v_3 - v_7 - v_9 - v_{10} + v_{11} - v_{13} - v_{15} - v_{16}$
$d[DHSph1P]/dt = v_4 + v_5 - v_6$
$d[C_{16} DHGlcCer]/dt = v_7 - v_8$
$d[C_{16} DHSM]/dt = v_9 + v_{10} - v_{11} - v_{12}$
$d[C_{16} DHCerP]/dt = v_{13} - v_{14}$
$d[C_{16} Cer]/dt = v_{15} + v_{16} - v_{17} - v_{19} - v_{20} + v_{21} - v_{23} - v_{24}$
$d[C_{16} CerP]/dt = v_{17} - v_{18}$
$d[C_{16} SM]/dt = v_{19} + v_{20} - v_{21} - v_{22}$
$d[C_{16} GlcCer]/dt = v_{24} - v_{25}$

Additional material

Additional file 1: Supporting Material. This file is in PDF format and contains additional Figures S1-S2 and additional Tables S1-S3.

Abbreviations used

Cer: Ceramide; CerP: Ceramide-1-phosphate; Cerk: ceramide kinase; Degs1: degenerative spermatocyte homolog 1 (Drosophila); Degs2: degenerative spermatocyte homolog 2 (Drosophila), lipid desaturase; CerS5: ceramide synthase 5; CerS6: ceramide synthase 6; Smpd1: sphingomyelin phosphodiesterase 1, acid lysosomal; Sphk1: sphingosine kinase 1; Sphk2: sphingosine kinase 2; Sms1: sphingomyelin synthase 1; Sms2: sphingomyelin synthase 2; Ugcg: UDP-glucose ceramide glucosyltransferase; C16 Cer: N-(hexadecanoyl)- sphing-4-enine (C16 Ceramide); C16 CerP: N-(hexadecanoyl)-sphing-4-enine-1-phosphate (C16 Ceramide-1-phosphate); C16 GlcCer: N-(hexadecanoyl)-1-β- sphing-4-enine (C16 Glucosylceramide); C16 SM: N-(hexadecanoyl)- sphing-4-enine -1-phosphocholine (C16 Sphingomyelin); C16 DHCer: N-(hexadecanoyl)-sphinganine (C16 Dihydroceramide); C16 DHCerP: N-(hexadecanoyl)-sphinganine-1-phosphate; C16 DHGlcCer: N-(hexadecanoyl)-1-β-glucosyl-sphinganine; C16 DHSM: N-(hexadecanoyl)-sphinganine-1-phosphocholine; DHSph: Sphinganine; DHSph1P: Sphinganine-1-phosphate; DHGalCer: Dihydro Galactosylceramide; GalCer: Galactosylceramide;

Author details
[1]Department of Bioengineering, University of California, San Diego, 9500 Gilman Dr., La Jolla CA 92093, USA. [2]School of Biology & Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332-0230, USA. [3]Department of Cellular and Molecular Medicine, University of California, San Diego, 9500 Gilman Dr., La Jolla CA 92093, USA. [4]Department of Chemistry & Biochemistry, San Diego Supercomputer Center and Graduate Program in Bioinformatics, University of California, San Diego, 9500 Gilman Dr., La Jolla CA 92093, USA.

Authors' contributions
SG designed the simulations, wrote the computer program, analyzed the experimental data and the simulation results and wrote the manuscript. AHM assisted in refining the reaction network. MRM assisted in designing some of the simulations, wrote part of the computer program and contributed in writing the manuscript. AHM, CKG and SS assisted in revising the manuscript. Sphingolipid and enzyme activity measurements were carried out in the laboratory of AHM. Transcriptomics experiments were carried out in the laboratory of CKG. SS supervised the modeling study. All authors have read and approved the final manuscript.

References
1. Fahy E, Subramaniam S, Brown HA, Glass CK, Merrill AH, Murphy RC, Raetz CRH, Russell DW, Seyama Y, Shaw W, *et al*: **A comprehensive classification system for lipids.** *Journal of lipid research* 2005, **46(5)**:839-862.
2. Merrill AH, Wang MD, Park M, Sullards MC: **(Glyco)sphingolipidology: an amazing challenge and opportunity for systems biology.** *Trends Biochem Sci* 2007, **32(10)**:457-468.
3. Zheng W, Kollmeyer J, Symolon H, Momin A, Munter E, Wang E, Kelly S, Allegood JC, Liu Y, Peng Q, *et al*: **Ceramides and other bioactive sphingolipid backbones in health and disease: lipidomic analysis, metabolism and roles in membrane structure, dynamics, signaling and autophagy.** *Biochim Biophys Acta* 2006, **1758(12)**:1864-1884.
4. Lopez PH, Schnaar RL: **Gangliosides in cell recognition and membrane protein regulation.** *Curr Opin Struct Biol* 2009, **19(5)**:549-557.
5. Zeidan YH, Hannun YA: **Translational aspects of sphingolipid metabolism.** *Trends Mol Med* 2007, **13(8)**:327-336.
6. Wennekes T, van den Berg RJ, Boot RG, van der Marel GA, Overkleeft HS, Aerts JM: **Glycosphingolipids–nature, function, and pharmacological modulation.** *Angew Chem Int Ed Engl* 2009, **48(47)**:8848-8869.
7. Murphy RC, Fiedler J, Hevko J: **Analysis of nonvolatile lipids by mass spectrometry.** *Chem Rev* 2001, **101(2)**:479-526.
8. KEGG BRITE database. [http://www.genome.jp/kegg/brite.html].
9. SphinGOMAP pathways. [http://sphingolab.biology.gatech.edu].
10. Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, Kitano H, Kohlbacher O, Neuweger H, Schneider R, Tenenbaum D, *et al*: **Visualization of omics data for systems biology.** *Nat Methods* 2010, **7(3 Suppl)**:S56-68.
11. Bartke N, Hannun YA: **Bioactive sphingolipids: metabolism and function.** *Journal of lipid research* 2009, **50(Suppl)**:S91-96.
12. Kitatani K, Idkowiak-Baldys J, Hannun YA: **The sphingolipid salvage pathway in ceramide metabolism and signaling.** *Cell Signal* 2008, **20(6)**:1010-1018.
13. Fyrst H, Saba JD: **Sphingosine-1-phosphate lyase in development and disease: sphingolipid metabolism takes flight.** *Biochim Biophys Acta* 2008, **1781(9)**:448-458.
14. Maceyka M, Milstien S, Spiegel S: **Sphingosine-1-phosphate: the Swiss army knife of sphingolipid signaling.** *Journal of lipid research* 2009, **50(Suppl)**:S272-276.
15. Henning PA, Merrill AH, Wang MD: **Dynamic pathway modeling of sphingolipid metabolism.** *Conf Proc IEEE Eng Med Biol Soc* 2004, **4**:2913-2916.
16. Alvarez-Vasquez F, Sims KJ, Cowart LA, Okamoto Y, Voit EO, Hannun YA: **Simulation and validation of modelled sphingolipid metabolism in Saccharomyces cerevisiae.** *Nature* 2005, **433(7024)**:425-430.
17. Henning P, Moffitt R, Allegood J, Wang E, Merrill A, Wang M: **Computationally predicting rate constants in pathway models.** *Conf Proc IEEE Eng Med Biol Soc* 2005, **5**:5093-5096.
18. Garcia J, Shea J, Alvarez-Vasquez F, Qureshi A, Luberto C, Voit EO, Del Poeta M: **Mathematical modeling of pathogenicity of Cryptococcus neoformans.** *Mol Syst Biol* 2008, **4**:183.
19. Kyoto Encyclopedia of Genes and Genomes (KEGG). [http://www.genome.ad.jp/kegg/].
20. Junker BH, Klukas C, Schreiber F: **VANTED: A system for advanced data analysis and visualization in the context of biological networks.** *BMC Bioinformatics* 2006, **7**.
21. Gupta S, Maurya MR, Stephens DL, Dennis EA, Subramaniam S: **An Integrated Model of Eicosanoid Metabolism and Signaling Based on Lipidomics Flux Analysis.** *Biophysical Journal* 2009, **96(11)**:4542-4551.
22. Merrill AH Jr: **Characterization of serine palmitoyltransferase activity in Chinese hamster ovary cells.** *Biochim Biophys Acta* 1983, **754(3)**:284-291.
23. Sullards MC, Allegood JC, Kelly S, Wang E, Haynes CA, Park H, Chen Y, Merrill AH Jr: **Structure-specific, quantitative methods for analysis of sphingolipids by liquid chromatography-tandem mass spectrometry: "inside-out" sphingolipidomics.** *Methods Enzymol* 2007, **432**:83-115.
24. Maurya MR, Subramaniam S: **A kinetic model for calcium dynamics in RAW 264.7 cells: 1. Mechanisms, parameters, and subpopulational variability.** *Biophys J* 2007, **93(3)**:709-728.
25. Yoshikawa K, Kita Y, Kishimoto K, Shimizu T: **Profiling of eicosanoid production in the rat hippocampus during kainic acid-induced seizure: dual phase regulation and differential involvement of COX-1 and COX-2.** *The Journal of biological chemistry* 2006, **281(21)**:14663-14669.
26. The Mathworks, Inc.© 1994 - 2010. [http://www.mathworks.com/].
27. Dennis EA, Deems RA, Harkewicz R, Quehenberger O, Brown HA, Milne SB, Myers DS, Glass CK, Hardiman GT, Reichart D, *et al*: **A Mouse Macrophage Lipidome.** *J Biol Chem* 2010, **285(51)**:39976-39985.
28. Messmer TO, Wang E, Stevens VL, Merrill AH Jr: **Sphingolipid biosynthesis by rat liver cells: effects of serine, fatty acids and lipoproteins.** *J Nutr* 1989, **119(4)**:534-538.
29. Merrill AH, Wang E, Mullins RE: **Kinetics of long-chain (sphingoid) base biosynthesis in intact LM cells: effects of varying the extracellular concentrations of serine and fatty acid precursors of this pathway.** *Biochemistry* 1988, **27(1)**:340-345.
30. Gomez-Munoz A: **Ceramide 1-phosphate/ceramide, a switch between life and death.** *Biochim Biophys Acta* 2006, **1758(12)**:2049-2056.
31. Niwa S, Graf C, Bornancin F: **Ceramide kinase deficiency impairs microendothelial cell angiogenesis in vitro.** *Microvasc Res* 2009, **77(3)**:389-393.
32. Hinkovska-Galcheva V, VanWay SM, Shanley TP, Kunkel RG: **The role of sphingosine-1-phosphate and ceramide-1-phosphate in calcium homeostasis.** *Curr Opin Investig Drugs* 2008, **9(11)**:1192-1205.
33. Pettus BJ, Kitatani K, Chalfant CE, Taha TA, Kawamori T, Bielawski J, Obeid LM, Hannun YA: **The coordination of prostaglandin E2 production by sphingosine-1-phosphate and ceramide-1-phosphate.** *Mol Pharmacol* 2005, **68(2)**:330-335.
34. Saxena S, Banerjee M, Shirumalla RK, Ray A: **Ceramide kinase: a potential anti-inflammatory target?** *Curr Opin Investig Drugs* 2008, **9(5)**:455-462.
35. Lamour NF, Chalfant CE: **Ceramide kinase and the ceramide-1-phosphate/cPLA2alpha interaction as a therapeutic target.** *Curr Drug Targets* 2008, **9(8)**:674-682.