Modeling and analysis of flux distribution and bioproduct formation in *Synechocystis* sp. PCC 6803 using a new genome-scale metabolic reconstruction

Chintan J. Joshi[a], Christie A.M. Peebles[a,b], Ashok Prasad[a,b,*]

[a] Chemical and Biological Engineering, Colorado State University, Fort Collins, CO 80523, United States
[b] School of Biomedical Engineering, Colorado State University, Fort Collins, CO 80523, United States

A R T I C L E   I N F O

A B S T R A C T

Cyanobacteria are prokaryotes capable of performing oxygenic photosynthesis. This makes them an attractive candidate for genetic engineering to produce commercially important chemicals. However, optimally harnessing this potential requires an understanding of metabolic regulation in cyanobacteria under photoautotrophic conditions. Here we present an updated genome-scale metabolic network reconstruction (iSynCJ816) of *Synechocystis* sp. PCC 6803. This updated model, containing 816 genes and 1045 reactions, builds upon previously published models. New features include an unconstrained photo-respiratory reaction mechanism and a mechanism to account for changes in energy absorption from light at different wavelengths. We used Flux Balance Analysis (FBA) to calculate the flux distribution within iSynCJ816 and compared in silico predictions with values obtained by previous in vivo metabolic flux analyses in *Synechocystis* sp. PCC 6803. A qualitative growth comparison of 167 gene-deletion mutants with experimental studies resulted in an accuracy rate between 70 and 79%. We used the model to estimate both the maximum theoretical yield of each metabolite and the feasibility of engineering *Synechocystis* to increase CO_2 fixation. We found that it is theoretically possible to increase CO_2 fixation by up to 35% from wild-type levels. We also carried out a dynamic flux balance analysis of fluxes throughout a light-dark cycle and obtained results that qualitatively matched experimental observations.

1. Introduction

Cyanobacteria are the only oxygenic prokaryotes capable of converting abundantly available carbon dioxide and sunlight into chemical energy via photosynthesis. As primary producers in aquatic environments, they play an important role in CO_2 assimilation and oxygen evolution. Geological and geochemical research has indicated that oxygen evolution by photosynthetic cyanobacteria is primarily responsible for the presence of significant levels of molecular oxygen in Earth's atmosphere, a process that began approximately 3 billion years ago [1]. Today, cyanobacteria are primary photosynthesizers, accounting for nearly 30% of Earth's photosynthetic productivity [2]. Their photosynthetic capabilities, along with other useful properties such as fast generation times relative to plants, small genomes that are amenable to transformation and synthesis of large numbers of useful metabolites have made them an attractive platform for genetic modifications to produce commercially important chemicals such as biofuels [3], pharmaceuticals [4], and nutraceuticals [5].

Among the gamut of cyanobacterial strains, *Synechocystis* sp. PCC 6803 is a model organism that has been extensively studied [6] with the aim of developing it as a chassis for metabolic engineering purposes. However, rational strategies for metabolic engineering require tools to help us understand the complexity of the metabolic network, to account for the metabolic trade-offs and bottlenecks. Genome-scale metabolic reconstructions, along with simulation techniques like FBA, provide a set of valuable tools to rationally engineer microorganisms like *Synechocystis* to meet commercial and societal needs. Because genome-scale reconstructions integrate available information into a systems framework, models ideally build upon previous work and address new questions in each iteration. Once a comprehensive reconstruction is obtained, various sets of methods developed by computational systems biologists are applied to validate and make predictions. One such well-known computational framework is FBA, a constraint-based modeling method used to calculate the flow of carbon through various metabolites to optimize reaction flux at steady state [7]. In the past, FBA has been utilized to predict flux distribution through wild-type [8–12] and mutant strains [13–15], epistatic interactions [16–20], futile cycles [21], and gene essentiality analysis [22–24].

Since the first cyanobacterial genome-scale metabolic reconstruction was published in 2005 [25], there have been nine more reconstructions [12,25–33]. They have progressed from being qualitatively useful to quantitatively useful models. Each subsequent reconstruction was more detailed than the previous one due to the availability of better data; the most recently published models included 677 genes [33]. More recently, a combined model was built by using a computational tool to aggregate three recent models which yielded a network with 778 reactions [34]. However, this combined model does not appear to have included a gene-reaction association table and has not been subjected to manual curation or validation with experimental data. This is important since some of the recently published models have been built independently of each other and contain inconsistencies among the models.

In this paper, we present a new genome-scale metabolic reconstruction of *Synechocystis* sp. PCC 6803. This reconstruction reconciles inconsistencies in the gene annotation within previously published reconstructions by mining databases such as KEGG [35–37] and Cyanobase [38,39]. Molecular mechanisms of the photosynthetic network around the thylakoid membrane have been included to facilitate a better understanding of respiratory and photosynthetic interactions. It should be noted that these molecular mechanisms have been included in only the most recent reconstruction [40]. Our work greatly improves upon earlier reconstructions by including thermodynamic analysis of > 500 reactions and by testing for thermodynamically infeasible loops (TILs). The inclusion of thermodynamic information leads us to remove or constrain reactions that participated in TILs.

To validate the reconstruction, we applied constraints and environment-specific conditions to simulate flux distributions under autotrophic (growth on carbon dioxide and light) and heterotrophic (growth on glucose) growth. We compared the fluxes generated by our model with experimental fluxes that were previously obtained via metabolic flux analysis (MFA) under autotrophic and heterotrophic conditions [41,42]. To further validate the genome-scale reconstruction, we compared the *in silico* predicted growth/no-growth phenotypes of 167 metabolic gene knockouts with experimental results obtained from online databases and a detailed literature search. Our predictions showed better accuracy than previously published models.

We used FBA to simulate intracellular fluxes and demand fluxes of each metabolite (also called bioproduct production flux) under autotrophic conditions within the model to determine which metabolites could theoretically be produced with high productivities. Analysis of the data revealed that, by suitably maximizing bioproduct production flux, carbon fixation can be improved by up to 35%, and oxygen production can be improved by up to 12%. Interestingly principal component analysis of the flux distribution showed that flux distributions clustered together based on photosynthetic state (i.e. PSI/PSII ratio) and pentose phosphate pathway activity.

One of the main challenges in metabolic modeling of cyanobacteria is the intrinsic variability in metabolism due to the sinusoidal light-dark cycle of a typical day. FBA is a static method that assumes that metabolism is in steady state. There is significant interest to develop a dynamic version of FBA, called Dynamic Flux Balance Analysis (DFBA). DFBA applications and implementations have increased since the first formulation was released in 1994 [43]. DFBA has been applied to metabolic networks of *E. coli* [44], *S. cerevisiae* [45], *L. lactis* [46], *C. reinhardtii* [47], *H. sapiens* (red blood cells) [48], and *S. stipites* [45]. Most of the DFBA methods fall into three broad classes. (i) The first is a dynamic optimization approach [44] in which an objective function is optimized over the entire time trajectory. This is a non-linear programming problem that typically requires a significant amount of kinetic information, which is currently unavailable for most organisms. (ii) The second approach, which is perhaps the most popular, is a piecewise static optimization approach [44]. Here time is discretized into small steps, and the objective function is optimized at each time step. Consistency with the light-dark cycle is maintained by changing the external conditions at each time step by hand. However, this method implicitly assumes that the organism is unaware of the light-dark cycle and the metabolic network only responds to the current environment. As a result, this method cannot predict glycogen accumulation or other dynamically changing aspects of metabolism. (iii) The third approach involves embedding the FBA model into a kinetic model of the external fluxes. This has been called the direct approach (DA-DFBA) [13,49]. The kinetic equations concern processes like metabolite uptake and secretion, which are experimentally accessible. This method requires less information than the fully dynamic method, and at the same time is more predictive than the static method [47]. We therefore implemented DA-DFBA to simulate fluxes during the light-dark cycle using the iSynCJ816 model.

Previous attempts to simulate fluxes using a dynamic method include the recently published model iHK677 [33], in which the authors designed a biomass objective function that changes over time, reflecting the changing expression of metabolic genes over the light-dark cycle. However, this is not very different from the static approach in that the changes in biomass composition are hard constraints that have been imposed externally. A better treatment of dynamics should be able to predict the changes in composition itself.

The DA-DFBA optimization scheme we used is based on a lexicographic optimization scheme to guarantee the uniqueness of the flux distribution [47]. This scheme is equivalent to the assumption that the metabolic network of the organism has many hierarchically organized objectives, which is more plausible than the assumption of just one objective.

2. Materials and methods

2.1. Model reconstruction and enhancement

The initial draft of the *Synechocystis* sp. PCC 6803 metabolic network was extracted from online databases. The genomic and pathway-related information was extracted using MATLAB codes from online databases including KEGG [35–37], Cyanobase [38,39], METACYC [50], and an annotated genome sequence [51]. We performed a homology search with the BLASTp algorithm on NCBI (online) [52] on the extracted sequences present in Supplementary Table 1 against other genomes. We used an accepted identity cut-off of 44% and an accepted E-value of 10^{-50} for this homology search. The information concerning enzymes was extracted from BRENDA [53] and KEGG. The information specific to chemical species (metabolites) was taken from ChEBI [54] and PubChem [55]. The reaction-based information about photosynthetic machinery was adapted from previously published studies [56–59] and cyanobacteria textbooks [60,61]. We also utilized previously published metabolic reconstructions of *Synechocystis* sp. PCC 6803 [31–33] and enhanced the biochemical and genomic information of various pathways such as photorespiration, serine synthesis, electron transfer within and between light harvesting proteins, fatty acid synthesis, chlorophyll synthesis, amino acid metabolism, and purine and pyrimidine metabolism. A literature search also resulted in the addition of a newly discovered light-independent serine synthesis pathway [62]. In this way, the draft network was subjected to iterative manual gap-filling, consistent with the protocol described by Thiele and Palsson [63]. This model building process was used to generate the list of genes, reactions, and metabolites. Any further processing and analysis of the reconstruction and pathways included has been discussed in the Results and discussion section.

The metabolic reconstruction was made by assimilating all the above information in an SBML (.xml) file [Supplementary File 1]. The reconstruction was converted to a mathematical model in MATLAB 2013b and readable/writeable with SBML toolbox [64] and COBRA toolbox [65,66].

2.2. Modeling of important photosynthetic reactions

2.2.1. Photosystem II (PSII)

Inclusion of thylakoid membrane as a separate compartment enabled us to include mechanisms associated with the photosystem II complex (PSIIa, PSIIb, PSIIc, & OEC) [Supplementary Fig. 1B]. The following mechanisms were included: (i) electrons from the S-cycle utilizing oxygen evolving complex (OEC) are transferred to the P680 reaction center (OEC a–d), (ii) the uncharged reaction center is protonated by one photon, and an electron is transferred to the PSII-bound plastoquinone (Q_A) on the stromal side (PSIIa), and (iii) in two consecutive reactions, electrons from Q_A ($2 Q_A \rightarrow 2 Q_A^-$) were transferred to an unbound Q_B ($Q_B \rightarrow Q_B^{2-}$) (PSIIb-c). This Q_B^{2-} gets protonated to PQH_2 to diffuse through the thylakoid membrane. Protonation and deprotonation of P680 reaction center join water-splitting to PSIIa, and hence creates a flux mode connecting PSII reactions and oxygen evolution [61].

2.2.2. Cytochrome b_6/f (CBFC)

The Q-cycle mechanism of operation of Cytochrome b_6/f [Supplementary Fig. 1D] was implemented. This involves the following processes: (i) PQH_2 diffusing through thylakoid membrane, and regenerated at PSII, reduces the Rieske protein ([^{2}Fe-^{2}S] cluster) while releasing a proton in the lumen and a semiquinone on the lumen side of the complex (CBFCua); (ii) the reduced form of the Rieske protein reduces plastocyanin (PC) in the lumen through Cytochrome-f (CBFCub); (iii) the semiquinone on the lumen side transfers an electron to heme b_l (low state), while semiquinone is converted to quinone on the stromal side (CBFCuc-d), (iv) the heme b_l then transfers an electron to heme b_h (high state) and converts the quinone to a semiquinone on the stromal side (CBFCue), and (v) the heme b_l transfers an electron to heme b_h on the stromal side while converting the semiquinone on the stromal side to fully reduced quinone ($PQH2$) using protons from the stroma (CBFCuf). The 1st half of the Q-cycle utilizes the steps (i–iv), while only the 2nd half of this cycle utilizes the steps (i–iii, v) to regenerate PQ and PQH_2, which then diffuse within the membrane. The cytochrome $b6/f$ complex has been modeled in the thylakoid membrane (CBFCua-uf) as well as in a cytosolic membrane (CBFCpa-pf). This process overall mediates the transfer of 4 cytosolic protons to lumen side [61].

2.2.3. Photosystem I (PSI)

The PSI machinery [Supplementary Fig. 1C] is modeled as follows: (i) upon absorption of a photon, the P700 reaction center is protonated, and subsequently reduces ferredoxin through a series of electron transfers; (ii) the protonated P700 reaction center in the PSI complex is reduced by plastocyanin in the thylakoid (PSIb). The plastocyanin would be replaced by ferrocytochrome under the alternate cytochrome b_6/f electron transfer (PSI_2) process [61].

2.2.4. Ferredoxin (NADP$^+$) oxidoreductase

The ferredoxin reduced at PSI is then used to regenerate NADP$^+$ to NADPH (FNOR) to be utilized by all other parts of the cellular metabolic processes.

2.2.5. Alternate electron flow pathways

Other alternate electron transfer processes have also been modeled as follows: (i) Ferredoxin (PQ) reductase catalyzing reduction of quinone (PQ $\rightarrow$ PQH$_2$) using ferredoxin via heme; (ii) NAD(P)H dehydrogenase complexes, (iii) the Mehler reaction (MEHLER), (iv) the cytochrome c oxidases (CYTBD), and (v) the quinone (PQ) oxidase (CYO1b). These alternate electron flow pathways have been modeled in both the cytosol and the thylakoid membranes [61].

2.3. Thermodynamics – calculation and adjustment of $\Delta_r G'^o$ to $\Delta_r G'^m$, $\Delta_r G'^m_{min}$, and $\Delta_r G'^m_{max}$

The thermodynamic analysis of the metabolic reconstruction was performed according to the group contribution method developed by Jankowski and others [67]. The $\Delta_r G'^o$ was calculated using $\Delta_f G'$ and $U_{f,est}$ extracted from the previously published model of *E. coli* (iAF1260) [9]. $\Delta_r G'^o$ refers to the reference state of 1 M, which was adjusted to the reference state of 1 mM ($\Delta_r G'^m$) using Eqs. (1)–(2).

$$\Delta_r G'^o = \sum_{i=1}^{m} n_i \Delta_f G'^o_i \tag{1}$$

$$\Delta_r G'^m = \sum_{i=1}^{m} n_i \Delta_f G'^o_i + RT \, ln\left(\prod_{i=1}^{m} x_i^{n_i}\right) \tag{2}$$

where n_i is the stoichiometry of the metabolite i (which is negative if the metabolite is a reactant and positive if the metabolite is a product), R is the gas constant in kcal K^{-1} mol^{-1}, T is the temperature in K, and x_i is the metabolite activity as a proxy for metabolite concentration. Metabolite activity was set to 1 mM for all metabolites except H$^+$, H$_2$O, H$_2$, and O$_2$. The reference concentrations of H$_2$ and O$_2$ were set to the saturation concentration of these metabolites in H$_2$O at 1 atm and 298.15 K. We also take into account the proton gradient for all reactions involving H$^+$ transport across the biological membrane, which can be given by Eqs. (3)–(5).

$$\Delta_r G'^m = \Delta_r G'^m_{(transport)} + \Delta_r G'^m_{(intracellular)} \tag{3}$$

$$\Delta_r G'^m_{est(transport)} = \Delta_{\Delta pH} G \tag{4}$$

$$\Delta_{\Delta pH} G = -2.3 \, hRT \, (\Delta pH) \tag{5}$$

where ΔpH is the pH difference across the membrane, h is the number of protons transported across the membrane, and $\Delta_r G'^m_{(intracellular)}$ is the metabolic component of the transport reaction. *Synechocystis* sp. PCC 6803 has been known to grow best between pH levels of 7–8.5. We assume that the ΔpH across the cytoplasmic membrane is 0.5. We do not take into account the pH difference across carboxysomes. All the $\Delta_r G'^o$ and $\Delta_r G'^m$ used are provided in the supplementary information [Supplementary Table 2].

We estimated variations resulting from two possible sources: (i) variation due to large changes in activities of metabolites, and (ii) variation from uncertainty due to the Gibbs free energy of metabolite formation of metabolites. The variation arising from both of these sources can be given by the following Eqs. (6)–(8) [67].

$$U_{r,est} = \sqrt{\sum_{i=1}^{m} n_i^2 U_{f,i}^2} \tag{6}$$

$$\Delta_r G'^m_{max} = \sum_{i=1}^{m} n_i \Delta_f G'^o_i + RT \, ln\left(\prod_{i=1}^{products} x_{max}^{n_i}\right) + RT \, ln\left(\prod_{i=1}^{reactants} x_{min}^{n_i}\right) + U_{r,est} \tag{7}$$

$$\Delta_r G'^m_{min} = \sum_{i=1}^{m} n_i \Delta_f G'^o_i + RT \, ln\left(\prod_{i=1}^{products} x_{min}^{n_i}\right) + RT \, ln\left(\prod_{i=1}^{reactants} x_{max}^{n_i}\right) - U_{r,est} \tag{8}$$

where; x_{min} is minimal metabolite activity assumed to be 0.00001 M, and x_{max} is maximal metabolite activity assumed to be 0.02 M. As we did for the calculation of $\Delta_r G'^m$, the physiological ranges for dissolved gases are lower than that of other metabolites, so we set x_{min} of H$_2$, O$_2$, and CO$_2$ as 10^{-8} M, and x_{max} for these dissolved gases was set to their saturation concentration in water at 1 atm and 298.15 K, which is 0.000034 M, 0.000055 M, and 0.0014 M, respectively. Eqs. (6)–(8) provide upper and lower limits for the thermodynamic estimates.

The reaction reversibility information was initially set from that used in a previously published model of *Synechocystis* sp. PCC 6803 [33]. Reactions absent in the previously published model were assumed to be reversible. We then used the above calculated $\Delta_r G'_{min}$ and $\Delta_r G'_{max}$

to constrain directionality: (i) if the $\Delta_r G'_{max}$ was less than zero, the reaction was strictly irreversible; (ii) if the $\Delta_r G'_{min}$ was greater than zero, the reaction was thermodynamically infeasible in the forward direction; and (iii) if $\Delta_r G'_{max}$ was greater than zero and $\Delta_r G'_{min}$ was less than zero, the reaction was considered to be reversible. To detect the presence of thermodynamically infeasible loops, we coupled our analyses with flux variability analysis (Materials and methods, Section 2.6), checked whether flux ranges were consistent with the range of $\Delta_r G'^m$, and modified the directionality accordingly.

The reference state for metabolite concentrations, on which our analyses is based ($\Delta_r G'^m$) was chosen to be 1 mM as mentioned above. However, intracellular metabolite concentrations are known to be significantly different [68]. To keep this in consideration, we calculated the Gibbs free energy change of formation ($\Delta_r G'_{min}$, $\Delta_r G'_{max}$) for all reactions in the 0.00001–0.02 M concentration range. The calculated range of the free energy of formation was used to determine if the reaction was bi-directional or uni-directional. If the entire range including the uncertainty was positive or negative, the reaction was constrained to the backward or forward direction, respectively.

2.4. Biomass composition

Biomass composition for some of the previously published *Synechocystis* sp. PCC 6803 metabolic network computational studies [31] have been extracted from other experimental studies [25]. However, the deviations from the actual stoichiometric weightage of metabolites which contribute towards growth rate, among previously published reconstructions, only vary in the order of 10^{-6}. Hence, we chose to adopt the biomass objective equation from one such study, which had 114 biomass components [69].

2.5. Light composition

To account for optical variation of photosystem I (PSI) and photosystem II (PSII) during the light-dark cycle, we resolved sunlight flux into a fraction of photons interacting with PSI and PSII. The details of how this was implemented as well as a separate model file in which this is implemented are provided in Supplementary Text 1, and a supplementary model file respectively.

2.6. Flux balance analysis

FBA is a mathematical framework used to calculate the flow of metabolites through the metabolic network at steady state [70]. A stoichiometric matrix S, of size M by N, corresponding to metabolic reactions (N) and metabolites (M) can be calculated based on the metabolic model. Using this matrix, a system of differential equations can be written for the rate of change of concentrations of each metabolite. Under the assumption of steady state, reaction rates normalized to biomass production, a.k.a. fluxes given by this system of differential equations become a system of linear equations which is set to zero, given by Eq. (9a).

$$\sum_{j=1}^{N} S_{ij} \cdot v_j = 0 \ (OR) \ S \cdot v = 0 \tag{9a}$$

Here, ν is a vector of reaction flux and S_{ij} represents the stoichiometric coefficient for ith metabolite in jth reaction. To find the fluxes, an objective function is chosen that is hypothesized to be optimized by the organism, such as its growth rate. This makes it a linear programming problem (LPP) that can be solved by standard techniques by imposing additional constraints, discussed below, in addition to Eq. (9a). The objective function most commonly used for such models is an equation describing the growth rate of the organism. Growth rate reactions are described as given in Eq. (9b),

$$Max \sum_{j=1}^{N} c_j \cdot v_j \rightarrow \mu \ (OR) \ Max \ C^T \cdot v \tag{9b}$$

In the above equation, c_j and v_j refer to the weight in the final biomass and the flux of the product of the jth reaction, respectively, and μ refers to the growth rate of the organism. Constraints can be applied to secretion/uptake fluxes of metabolites, or applying upper (α_j) and lower (β_j) limits to intracellular reaction fluxes (v_j), given by Eq. (9c),

$$\alpha_j \leq v_j \leq \beta_j \tag{9c}$$

Reversible reactions can take either negative or positive value of fluxes, while irreversible values were constrained to take only positive values. Further, if any reactions were turned off, inactivated or deleted, the flux through the reaction was set to zero, as described by Eq. (10),

$$v_j = 0 \tag{10}$$

The mathematical definition of the problem formed by equations above (9) is called Linear Programming (LP). The LP problem can be constrained by experimentally measured uptake or secretion fluxes of various exchange reactions for nutrients and/or products. Constraints of the thermodynamic nature can be implemented by assigning reversibility of the reactions by calculating net change in Gibb's free energy. Variations to this problem are possible by using different choices of objective functions. The most common objective function used is microbial growth rate, which is also the one used here. A two-stage optimization problem can be formed by nesting one objective function into another (See Section 3.7). The linear programming problem was implemented using COBRA Toolbox with Gurobi 4.6.1 on MATLAB R2011b [66].

2.7. Flux variability analysis (FVA)

Genome-scale metabolic networks are usually under-determined systems that have many different solutions. Solutions are usually chosen based on optimizing one or more objective functions. Even so, it is very common to find multiple solutions that yield the same value of the objective function under the same set of constraints. To investigate this flux solution space, approaches called alternate optima analyses have been developed in previous studies [10,71]. Here, we implemented one of the basic types of alternate optima analyses called FVA.

In FVA we formulate a two-stage LP problem as follows. We start by determining the optimal value (Z_{obj}) of the linear objective function using Eqs. (9a)–(10) given above. Now, we use this value to constrain the original objective function (for which the optimal value was determined) and perform N (number of variables in the problem; for metabolic networks, the number of reactions) iterations, where a new objective function is assigned in each iteration. This new objective function is the flux through one of the reactions, which is separately maximized (Eq. (11a)) and minimized (Eq. (11b)), while keeping the rate of growth at its optimal value. After 2 N iterations of maximizing and minimizing each reaction flux, this procedure generates a range of flux values for each reaction where the exact same optimal value of the original objective function can be achieved. The set of ranges, therefore, define the boundaries of the optimal solution space.

$$
\begin{aligned}
&Max \quad v_j \\
&s.\,t. \quad S \cdot v = 0 \\
&\quad C^T \cdot v = Z_{obj} \\
&\qquad \alpha_j \leq v_j \leq \beta_j \ \text{for} \ j = 1\ldots n
\end{aligned}
\tag{11a}
$$

$$
\begin{aligned}
&Min \quad v_j \\
&s.\,t. \quad S \cdot v = 0 \\
&\quad C^T \cdot v = Z_{obj} \\
&\qquad \alpha_j \leq v_j \leq \beta_j \ \text{for} \ j = 1\ldots n
\end{aligned}
\tag{11b}
$$

2.8. Dynamic flux balance analysis

Here, we used DFBAlab with COBRA Toolbox on MATLAB R2014b, which implements the DA-DFBA described previously [47,72]. Uptake and secretion kinetics were set up by writing a mass balance for extracellular/nutrient chemical species which can, generally, be described by Eqs. (12a) and (12b).

$$\frac{d\mathbf{z}(t)}{dt} = \mathbf{f}(t, \mathbf{z}(t), \mathbf{v}(t, \mathbf{z}(t))) \tag{12a}$$

$$\boldsymbol{z}(t_0) = \boldsymbol{z}_0 \tag{12b}$$

Here, **f** represents the uptake kinetics as a function of time (t), metabolite concentration (**z**), and the LP flux solution (**v**) obtained from Eqs. (9a)–(9c); and **v** also contains the exchange fluxes. Our next step is to modify the metabolic network such that all reactions can only carry positive flux and define the LP problem in context of this metabolic network. Therefore, Eqs. (9a)–(9c) takes a time variant form given by Eqs. (13a)–(13c).

$$U(t, \mathbf{z}(t)) = \underset{\mathbf{v} \in \mathbb{R}^N}{\arg\min} \ \mathbf{c}^T \mathbf{v} \tag{13a}$$

$$\text{s. t. } \mathbf{S}.\ \mathbf{v} = \mathbf{b}(t, \mathbf{z}(t)) \tag{13b}$$

$$\mathbf{v} \geq 0 \tag{13c}$$

Here, U (t, **z** (t)) is a subset of R^N and is called *the solution set* for the right-hand side **b**, an element of R^M; and **b** represents the accumulation or utilization of chemical species. U (t, **z** (t)) is called *the solution set* because it contains all values of flux vector, **v**, where the objective function attains the optimum value. In Eq. (13a), *arg* min refers to all the points where the function, $\mathbf{c}^T\mathbf{v}$, is minimized. Therefore, it should be noted that U may contain alternate solutions, if any, which correspond to the optimum value of the objective function. It should also be noted that Eq. (13) has been written in standard LP form and maximizing $\mathbf{c}^T\mathbf{v}$ is equivalent to minimizing $-\mathbf{c}^T\mathbf{v}$. We have modified the weights (**c**) in Eqs. (13a)–(13c) as negative of the weights (**c**) in Eqs. (9a)–(9c). We refer to the Eqs. (12a) − (13c) as a dynamic system with a LP embedded.

2.9. Lexicographic optimization

FBA solutions are usually non-unique, leading to an indeterminacy problem for dynamic analysis, since the flux distribution at each time is sensitive to the flux distribution at previous times. It may be possible to reduce or resolve alternate solutions by formulating a set of hierarchical optimization problems [47,73] described by Eqs. (14a)–(15d).

$$g_1(V) = \underset{\mathbf{v} \in \mathbb{R}^N}{\min} \ \mathbf{c}^T \mathbf{v} \tag{14a}$$

$$\text{s. t. } \mathbf{S}.\ \mathbf{v} = \mathbf{b} \tag{14b}$$

$$\mathbf{v} \geq 0 \tag{14c}$$

$$g_2(V) = \underset{\mathbf{v} \in \mathbb{R}^N}{\min} \ \mathbf{c_2}^T \mathbf{v} \tag{15a}$$

$$\text{s. t. } \mathbf{S}.\ \mathbf{v} = \mathbf{b} \tag{15b}$$

$$\mathbf{c}^T \mathbf{v} = \mathbf{c}^T \mathbf{u} \tag{15c}$$

$$\mathbf{v} \geq 0 \tag{15d}$$

Here, $\mathbf{g}: S \rightarrow R^2$ and c_2 is a secondary optimization describing a secretion flux, uptake flux, and/or reaction flux like that of ethanol, oxygen, and/or ATP synthase. This is called Lexicographic optimization and is implemented as follows. First, a list of objective functions is made, in order of priority. The objective function with the highest priority is optimized first. Then the objective function with the next-highest priority is optimized with the optimum value of the first objective function as a constraint. The process is continued until the last objective function has been optimized.

Thus at each optimization step, an optimal flux vector is picked out from within the optima of the previous set. The number of solutions rapidly decline and a unique optimum is obtained with only a few steps. The priority ordering affects the final solution chosen. Thus, DA-DFBA involves additional hypotheses about metabolic objectives. These additional hypotheses are subordinate to the main one, for example maximization of glycogen production could be a subordinate objective to growth maximization. The validity of a hierarchy of lexicographic optimization can only be determined through comparisons of predicted fluxes as a function of time with experimentally observed fluxes over time. Though lexicographic optimization does involve a more complex hypothesis than simple FBA, it represents a reasonable biological hypothesis that organisms do not merely have one objective that governs metabolic fluxes but a hierarchy of objectives. We present and discuss predictions of the model utilizing three different schemes of lexicographic optimization coupled with DFBA. The parameters and set up of DFBA equations have been discussed in Supplementary Text 1.

2.10. Growth conditions and single gene deletions

The autotrophic growth of *Synechocystis* sp. PCC 6803 was simulated by setting the uptake rates of all the sources of carbon except carbon dioxide (CO_2) to zero. The CO_2 uptake was set to 3.7 mmol/gDW/h [25]. The light uptake rates were first determined by the growth condition being carbon limited or light limited. For our simulations at the given CO_2 uptake flux, the minimum amount of light required for maximum growth was calculated by dual maximization; i.e. calculating maximum growth rate (leaving light uptake flux unconstrained) using maximization of growth rate as objective function and then minimizing the light uptake (constraining growth rate equation to the value calculated in previous step) using minimization of light uptake as an objective function. The light above this minimum value is referred to as extra light. To simulate heterotrophic growth, we shut off the reaction for light absorption and set the lower bounds of glucose uptake to 0.85 mmol/gDW/h, as per previously published studies [42]. The carbon dioxide (CO_2) exchange with bulk was set to secretion by setting the lower bound to zero.

To simulate gene deletions, methods utilizing either FBA or Minimization of Metabolic Adjustment (MOMA) have been developed previously [15,74,75]. However, a study has shown that for *E. coli* K-12, the in silico FBA predictions of mutant and wild-type flux distributions are more accurate [74]. Hence, for our single gene deletion analysis, we used FBA to simulate gene deletions. To implement it, we apply a constraint given by Eq. (10) to all the reactions that are associated with the single gene deletion being simulated.

3. Results and discussion

3.1. Improvements in network reconstruction

A genome-scale metabolic reconstruction of *Synechocystis* (iSynCJ816) was developed based on existing genomic and biochemical information using previously established protocols [63] as detailed in the Materials and methods section. The initial reconstruction contains 816 genes [Fig. 1A], 925 metabolites [Fig. 1B], and 1060 reactions (metabolic, transport, GPR and non-GPR based reactions) distributed over 56 subsystems and 7 compartments (extracellular, cytoplasm, periplasmic space, periplasmic membrane, thylakoid lumen, thylakoid membrane, and carboxysome) [Fig. 1C]. This is larger than the recently reported reconstruction in [33] by 142 genes and even larger than the combination model constructed by Mohammadi et al. [34]. The initial step in the reconstruction was to carry out a comparative analysis of previously built reconstructions [31–33] that resulted in a net difference of 103 genes and 256 reactions that were present in at least one of the reconstructions and not others.

We found that there were many inconsistencies in the gene-reaction association among the two previously published models, iHK677 [33]
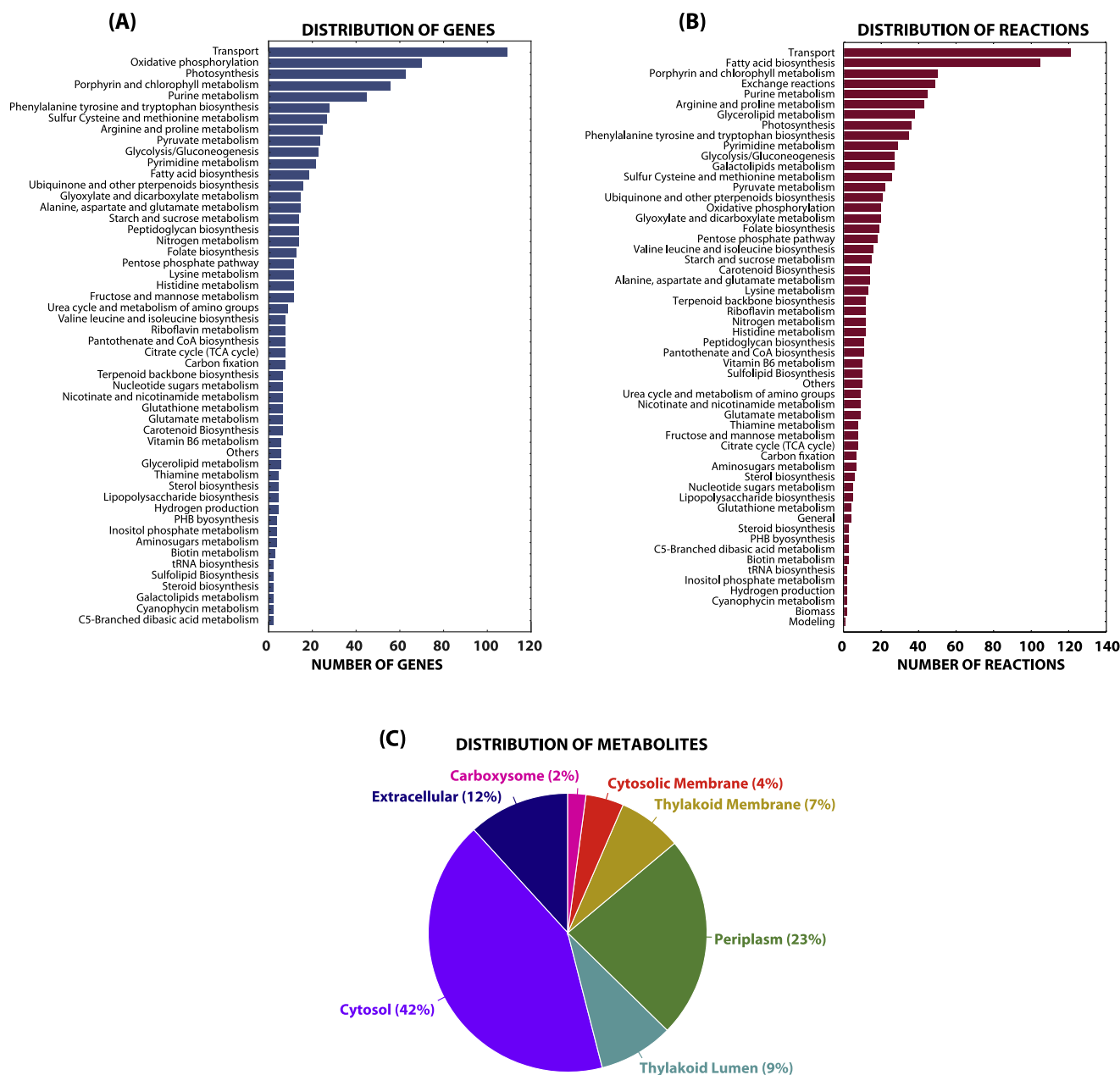
Fig. 1. Properties of the metabolic network reconstruction of *Synechocystis* sp. PCC6803. (A) Distribution of genes among various subsystems, (B) Distribution of metabolites among various compartments within the reconstruction, and (C) Distribution of reactions among various subsystems.

and iJN678 [31]. In all, we discovered and fixed 30 instances of such inconsistencies. Examples of such inconsistencies include sll1392 (Fe2 + transporter) in iJN678 that was correctly listed as slr1392 in iHK678, and sll1510 (G3P acyltransferase) in iJN678 which was slr1510 in iHK677. Several missing EC numbers were also added to facilitate construction of enzymatic information from within the model.

After manual checking, the 103 additional reactions were added to the model. We then searched for additional reactions using online databases as explained in the Materials and methods section and added a number of new reactions as well as the associated genes. The new reactions that were added, which were not present in previous models, belong to tRNA biosynthesis, amino acid and amino-sugar metabolism, sugar metabolism, nucleotide metabolism, nitrogen metabolism, terpenoid metabolism, fatty acid metabolism, photosynthesis, biosynthesis of secondary metabolism, and ion metabolism.

The photosynthetic electron transfer machinery was enhanced by the inclusion of the electron transfer mechanism through photosystems

I and II (PSI and PSII) as well as the electron transfer mechanism within cytochrome *b*6/*f*. To systemically study the photosynthetic process, we include the thylakoid and cytosolic membranes as separate compartments. The photosynthetic alternate electron flow pathways [56,76–78] were modeled taking into account the separate thylakoid and cytosolic membrane compartments. This implies that plastoquinone (PQ) involved in the electron transfer from *Psi*I to cytochrome *b*6/*f* is spatially limited to the thylakoid membrane, while the plastocyanin (PC) and ferricytochrome involved in electron transfer from cytochrome *b*6/*f* to PSI are spatially limited to the thylakoid lumen. For details on how the modeling of photosynthetic reactions was implemented, please refer to Materials and methods Section 2.2.

Our reconstruction, iSynCJ816, improves upon the previous reconstruction by the inclusion of multistep reaction cascades such as those involving photosystems I & II, glycine cleavage system, pyruvate dehydrogenases, cytochrome *b*6/*f*, etc., and expands the description of electron carrier promiscuity by the inclusion of 15 peroxidases and

NADH dehydrogenase catalyzed reactions. The distribution of reactions among various electron carriers is illustrated in supplementary information [Supplementary Fig. 2].

Previous studies [33] have reported the absence of the glyoxylate shunt, which produces malate from isocitrate via glyoxylate. We therefore have not included the glyoxylate shunt in our model. As shown by previously published results [79], a fully functional TCA cycle exists in *Synechocystis* sp. PCC 6803, which has been included in the reconstruction. The reconstruction also includes an additional carboxysome compartment, where three main sets of reactions of carbon-fixing occur: (i) carbonate dehydrogenase: conversion of HCO_3^- to CO_2, (ii) ribulose 1,5-bisphosphate carboxylase/oxygenase: CO_2 fixation and formation of 2-phosphoglycolate via oxygenase activity, and (iii) transport of by-products back to cytosol.

Once all the desired information about genes, reactions, and metabolites in each compartment was assembled, we performed charge balancing based on the charge on each metabolite as informed by various databases. After charge balancing, we also analyzed the distribution of the major electron carriers, ATP, protons, and H_2O [Supplementary Fig. 2].

A gap analysis was performed using the GapFind [66] tool in the COBRA toolbox on MATLAB, which finds blocked metabolites and assesses network connectivity. We found that there are 89 intracellular downstream gaps and 47 intracellular root gaps for a total of 136 blocked metabolites [Supplementary Table 3]. Gaps represent gaps in our knowledge, and we expect that these gaps will be filled with time and research. None of the gaps were in essential metabolites so they did not affect the simulations of the model. We have a larger number of gaps than previous models, which may have arisen because of incorporation of more genes, reactions, and compartments.

3.2. Thermodynamic analysis corrects reaction directionality and identifies unfavorable cycles

We used thermodynamic analysis to analyze reaction directionality for almost half the reactions in the model. The standard Gibbs free energy change of formation, $\Delta_f G'^o$, and reaction, $\Delta_r G'^o$, were estimated for compounds which occurred in both iAF1260 (*E. coli* model) [9] and our model. A total of 416 unique metabolites (~45%), common to both iAF1260 and iSynCJ816, are considered. Using this information, we were able to calculate $\Delta_r G'^o$ for 504 reactions (~49%) present in iSynCJ816 [Supplementary Table 2]. All $\Delta_r G'^o$ values were calculated using the contribution method implemented previously [67]. The conversion from $\Delta_r G'^o$ to $\Delta_r G'^m$ is shown in Materials and methods, Section 2.3.

In Fig. 2, we show a comparison of $\Delta_r G'^o$, $\Delta_r G'^m$, and the range $\Delta_r G'_{min}$ - $\Delta_r G'_{max}$. As discussed in the Materials and methods section,
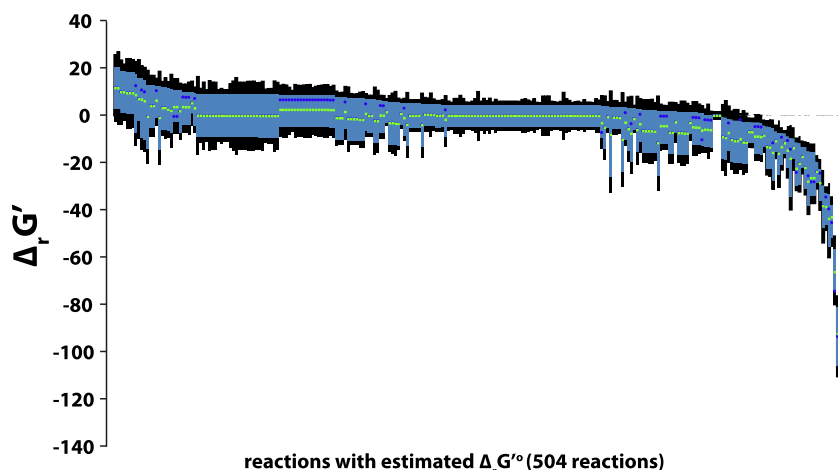
within a metabolite concentration range of [10−5, 0.02] M, reactions can be classified as either forward feasible, backward feasible or reversible depending upon the range $\Delta_r G'_{min}$–$\Delta_r G'_{max}$.

Among the reactions previously considered reversible, we found 16 reactions that had the entire thermodynamic range ($\Delta_r G'_{min}$, $\Delta_r G'_{max}$), including uncertainty ($U_{r,est}$), fall in the negative range. Hence, these reactions were restricted to the forward direction [Supplementary Table 4]. Among the reactions previously considered irreversible, we found 4 reactions in which the range ($\Delta_r G'_{min}$, $\Delta_r G'_{max}$) including uncertainty ($U_{r,est}$) was positive. Hence, the directionalities for these reactions were changed. The list of these reactions can be found in Supplementary Table 4. These changes did not affect growth rates in autotrophic or heterotrophic growth.

To ensure consistency, we also compared the free energy values with the sign of the reaction fluxes to ensure that fluxes were predicted in the correct direction for reversible reactions. For reversible reactions, if the range, [$\Delta_r G'_{min}$, $\Delta_r G'_{max}$] is entirely in the negative region, the calculated flux must be positive. In contrast, if the above range is entirely in the positive region, the calculated flux must be negative. We changed the directionality of reactions for which the product of flux and $\Delta_r G'$ resulted in a positive value.

We next identified thermodynamically infeasible loops (TILs) within the model using flux variability analysis as described previously [10]. A cycle was considered a TIL if flux through one of the reactions in the cycle achieved maximum possible flux by any given reaction. A reaction can achieve highest possible flux (1000 mmol/gDW/h) if it forms a loop or cycle with some other set of reactions that reverses this reaction [Supplementary Fig. 3]. After making thermodynamic corrections to the model, we identified a total of 13 unbounded reactions under heterotrophic growth conditions. A few examples that we found most interesting are discussed here. One TIL concerns the reversible reaction catalyzed by nucleoside-diphosphate kinase: ATP-GDP (NDPK1, sll1852), which regenerates GDP to GTP by transferring the phosphate from ATP. This reaction forms a TIL when forced in the backward direction in combination with the pyruvate kinases PYK and PYK3, which utilize ATP and GTP respectively [Fig. 3A]. PYK3 was also identified as a thermodynamically unfavorable (range of $\Delta_r G'$ was entirely positive) in the direction that utilizes GTP. Another similar TIL identified consisted of the reactions catalyzed by the enzymes NDPK3-PYK-PYK4. This TIL was based on the interaction between pyruvate phosphorylation (PYK4), the ATP-cytidine diphosphate (CDP) activity of nucleoside-diphosphate kinase (NDPK3), and ATP-utilizing pyruvate kinase (PYK) [Fig. 3B]. This TIL was resolved by removing pyruvate kinase reactions that utilize GTP and cytidine triphosphate (CTP). We chose to remove these reactions because thermodynamic analysis predicted that these reactions were thermodynamically unfavorable when utilizing GDP or CDP as substrates. Another TIL was found in the reactions catalyzed by

Fig. 2. Thermodynamic properties of the reactions for which $\Delta_r G'$ was calculated. The range of possible $\Delta_r G'$ values (kcal/mol) for the reactions in descending order. $\Delta_r G^o$ (magenta), $\Delta_r G^m$ (green), the range of $\Delta_r G^m$ (without uncertainty, blue), and range of $\Delta_r G^m$ (with uncertainty, black). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
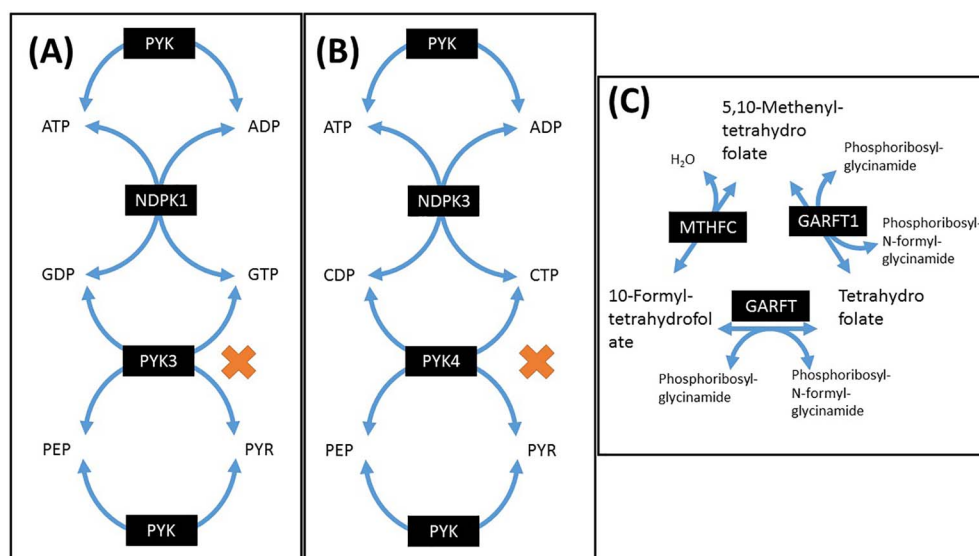
Fig. 3. Examples of thermodynamically infeasible cycles or futile cycles identified by our analysis. (A) Pyruvate kinase/Nucleoside diphosphate kinase (GDP utilizing)/Pyruvate kinase (GTP utilizing), (B) Pyruvate kinase/Nucleoside diphosphate kinase (CDP utilizing)/Pyruvate kinase (CTP utilizing), and (C) Interconversion of 5,10-Methenyltetrahydrofolate/10-tetrahydrofolate/tetrahydrofolate.
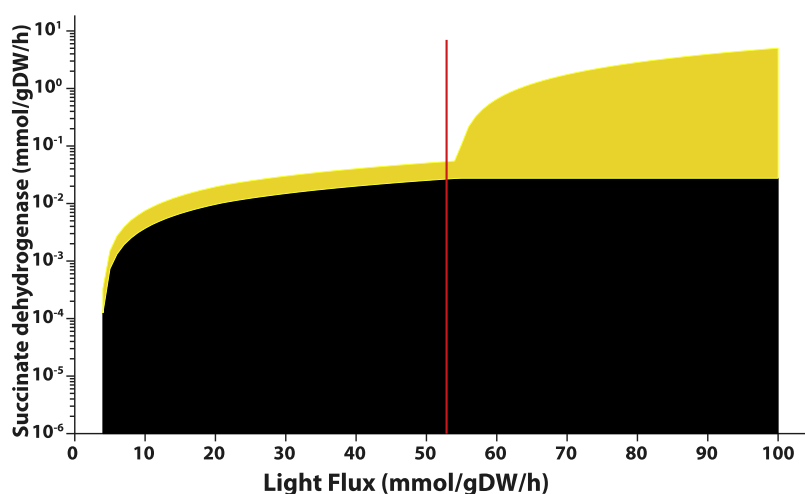

Fig. 4. Flux variability of succinate dehydrogenase under various light uptake conditions. The vertical red line indicates the light uptake under wild-type simulations, black region indicates the succinate dehydrogenase flux necessary for growth, and yellow region indicates electron transfer flux through succinate dehydrogenase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

phosphoribosylglycinamide formyltransferase, which can utilize either 10-formyltetrahydrofolate (GARFT) or 5, 10-methenyltetrahydrofolate (GARFT1) [Fig. 3C] as a substrate. To resolve the thermodynamically unfavorable cycle involving these two reactions, we removed GARFT1 to disrupt the thermodynamically unfavorable loop. Our final reconstruction after taking thermodynamic feasibility into account contained 816 genes and 1045 reactions.

3.3. Electron transfer in thylakoid membrane

The core electron transfer processes in the thylakoid membrane are well understood [61]. Light is absorbed by phycobilisomes, and light energy is transferred to PSI and PSII that acquire a charge due to the photon energy. Electrons originating within PSII reaction center flow through Q_A and Q_B to PQ located in the thylakoid lumen. Electrons accepted by PQ are then contributed to PSI through interactions with cytochrome b_6/f. The electrons at the PSI reaction center then reduce ferredoxin, which in turn interacts with ferredoxin (NADPH) reductase to generate NADPH. The charged PSII returns to its original state via the S-cycle of the oxygen-evolving complex. The charged PSI returns to the original state via oxidation of plastocyanin (reduced state). However, photosynthetic and respiratory electron transfer processes include a number of other reactions that are less well understood, and model simulations may provide some insight into their possible physiological

role. Below we discuss some of the key observations regarding the electron transfer machinery based on simulations of the model.

We performed FBA simulations as described and characterized various respiratory pathways according to maximum flux achievable at maximum growth rate [Supplementary Table 5]. We found that under carbon-limited (CL) conditions the achievable flux of succinate dehydrogenase (SUCD) was lowest among all the electron processing machinery. We also found that the flux of cytochrome oxidase (CYO1b2_syn) and superoxide dismutase (SOR) was highest and second highest, respectively, among all the electron processing machinery. However, experimental studies suggest a higher flux through SUCD than suggested by our simulations [56]. To investigate the role that SUCD may be playing, we fixed SUCD flux in the thylakoid compartment to that reported by experiment and found that this led to a reduction in growth rate of the organism [Supplementary Fig. 4] at optimal light levels. The cellular growth objective under these conditions was compromised since flux was pulled from 2-oxoglutarate, which is an important precursor for the production of amino acids that are in turn crucial for growth. Alternatively, SUCD could be playing an important role during high light conditions. To test this hypothesis, we performed flux variability simulations that confirmed that, for the same growth rate values under high light conditions, a major portion of the flux through SUCD arises due to its role in electron transfer [Fig. 4].

We find that in the absence of electron transfer through cytoplasmic

membrane a non-zero minimal flux is required through CO_2-transporting NADPH dehydrogenase (NDH1_3u) and SUCD. Further, our simulations suggest that the cytoplasmic membrane may be undergoing significant electron transfer in the absence of NDH1_3u. The absence of NADPH dehydrogenase in the thylakoid (NDH1_3u) pushes the electron transfer through the cytoplasmic version of NADPH dehydrogenase (NDH1_4pp). When both the cytoplasmic and the thylakoid variants of the enzyme are present, they both could be active. We did not find any experimental evidence that suggests that it is one or the other. Therefore, the higher than predicted flux through SUCD may arise due to these modes of operation of photosynthesis and oxidative pathways, which are present in both thylakoid and cytoplasmic membrane. It is also known that more than one electron transfer and oxidative pathway may be operational under any given environmental condition [40].

Previously published experimental studies have reported higher superoxide activities under CL conditions as compared with light limited (LL) conditions [80], which does not agree with what is seen in our simulations. However, a previous study in *Synechococcus* [81] suggested that cytochrome oxidase may provide an alternative pathway for electron transfer. Given high superoxide production rate, it can be argued that higher localized oxygen concentration will generate higher superoxide dismutase to cytochrome oxidase activity under CL conditions than under optimal (or LL) conditions. However, spatial concentration differences at this scale cannot yet be accounted for by an FBA model.

3.4. RuBisCO oxygenase and light-independent serine production

Recently, a light-independent serine production pathway was characterized in *Synechocystis* [62]. Previously published metabolic reconstructions have shown that in presence of photorespiration and absence of light-independent serine production pathway, all of the serine is produced via the photorespiratory pathway [69]. However, it is known that both serine production pathways are active during photosynthetic growth [62]. It is also known that RuBisCO oxygenase (RBCh) flux (which forms the photorespiratory precursor: 2-phosphoglycolate) forms 3%–5% of RuBisCO carboxylase (RBPC) flux [33,82]. Our simulations in presence of the newly identified light-independent serine production pathway predict that all of the serine is produced via the light-independent pathway and that photorespiration becomes dispensable. The reason for this is that photorespiration leads to a loss of carbon dioxide, which is the only source of carbon in autotrophic conditions, and thus lowers biomass. In presence of light-independent serine production pathway, the experimentally observed non-zero flux through photorespiration can be achieved via two mechanisms: (i) directly constraining RuBisCO oxygenase flux, resulting in the catabolism of 2-phosphoglycolate, or (ii) inclusion of a growth requirement on 2-phosphoglycolate catabolism. Previous metabolic models have dealt with this issue by constraining RuBisCO oxygenase flux to 5% [31] or 3% [33]. However, this appears to be an ad-hoc fix and there has been little discussion on the balance of photorespiratory and light-independent serine production pathway in the literature.

Experimental evidence suggests that photorespiration via RuBisCO oxygenase is an indispensable process under atmospheric conditions [83–87]. However, because there is no growth-associated demand (except possibly L-serine) for photorespiration in the presence of the light-independent serine production pathway, the optimal FBA prediction for flux through photorespiration will be zero. Given that (i) photorespiration is indispensable [87], (ii) biomass-related metabolites made by photorespiration can be made by another pathway [62], and (iii) non-zero photorespiratory flux has been observed in previous experiments [82], it is very likely that photorespiration may play roles beyond meeting metabolic requirements of biomass growth. Previous experiments suggest that photorespiration may play many roles [28,69,83–85,87]. Therefore, the overall metabolic role of photorespiration cannot be captured adequately if photorespiration is kept

manually fixed to a certain (3% or 5%) value under all environmental and genetic perturbations. It is important to leave photorespiration relatively unconstrained and allow the model to generate predictions on photorespiratory flux.

To generate a growth demand for photorespiratory flux, we hypothesized that a certain level of serine production occurs through light-independent serine production pathway while the remaining serine must be produced via photorespiration. Firstly, we assume a basal RBCh activity as 3% of total RuBisCO activity. Then, using this as a constraint, we calculated the viable flux of the light-independent serine pathway and the fraction of net serine production flux required through this pathway under optimal conditions. We found that 3% of total RuBisCO flux is directed towards its oxygenase flux, and light-independent serine production was calculated to be 41.72% of the value it would have, if it were meeting all serine requirements of the organism. The scheme devised here hypothesizes that both photorespiratory flux (3% of RuBisCO carboxylase flux) as well as flux through the light-independent serine production pathway is required to meet growth associated serine demand under light limited conditions (laboratory conditions). Our scheme is a plausible metabolic hypothesis about the role that photorespiration may be playing in the organism. This hypothesis can be experimentally tested through measurements of flux through the light-independent serine production pathway under different conditions. By setting an upper limit to light-independent serine production pathway, we were able to leave photorespiration unconstrained and capture effects on photorespiratory flux under conditions when there is a demand flux associated with a given metabolite (see Section 3.5).

We also simulated different scenarios that could result in flux through both the pathways. We constrained the flux through the light-independent pathway and calculated the flux through photorespiration. We constructed two possible scenarios for serine production: (i) total serine production through both the pathways combined increases with increase in flux through light-independent pathway [Fig. 5A]; and (ii) serine production via photorespiratory pathway decreases with increased flux through light-independent pathway [Fig. 5B]. The former scenario is associated with an increase in flux through NADP reduction (to facilitate hydroxypyruvate reductase, HPYRR1i_syn) and NADH oxidation (to facilitate phosphoglycerate dehydrogenase, 3PGDH, and glycolate dehydrogenase, GLYCTO_syn). This scenario creates an intracellular loop given by hydroxypyruvate reductase (NADH-utilizing) and 3-phosphoglycerate dehydrogenase (NADH-utilizing) [gray box, Fig. 5C]. Therefore, our simulations predict that an increase in pools of NAD (oxidized state) and NADPH (reduced state) may result in increased photorespiratory activity and light-independent serine production simultaneously without resulting in a trade-off between these two pathways. This would be very interesting to test experimentally in future work.

3.5. Model predicts theoretical increases in metabolic loads and carbon fixation

We evaluated the potential for *Synechocystis* sp. PCC 6803 to produce a variety of bioproducts. We took a generalized approach by assuming that each node within the metabolic network can produce a hypothetical bioproduct. We implemented a simple reaction scheme that directs flux from the metabolite at each node into a sink reaction. The flux of this reaction for bioproduct production was constrained to an arbitrary value (0.1 mmol/gDW/h) and other intracellular fluxes were calculated. Specifically, we analyzed HCO_3 uptake, RuBisCO carboxylase, RuBisCO oxygenase, and oxygen production.

Under this constant metabolic load, our simulations indicate that only 426 unique metabolic loads could be simulated with the same light uptake as the wild-type light uptake. In all cases tested, the results generated under constant light and light-independent serine production indicate that it is theoretically possible to increase inorganic carbon (Ci)
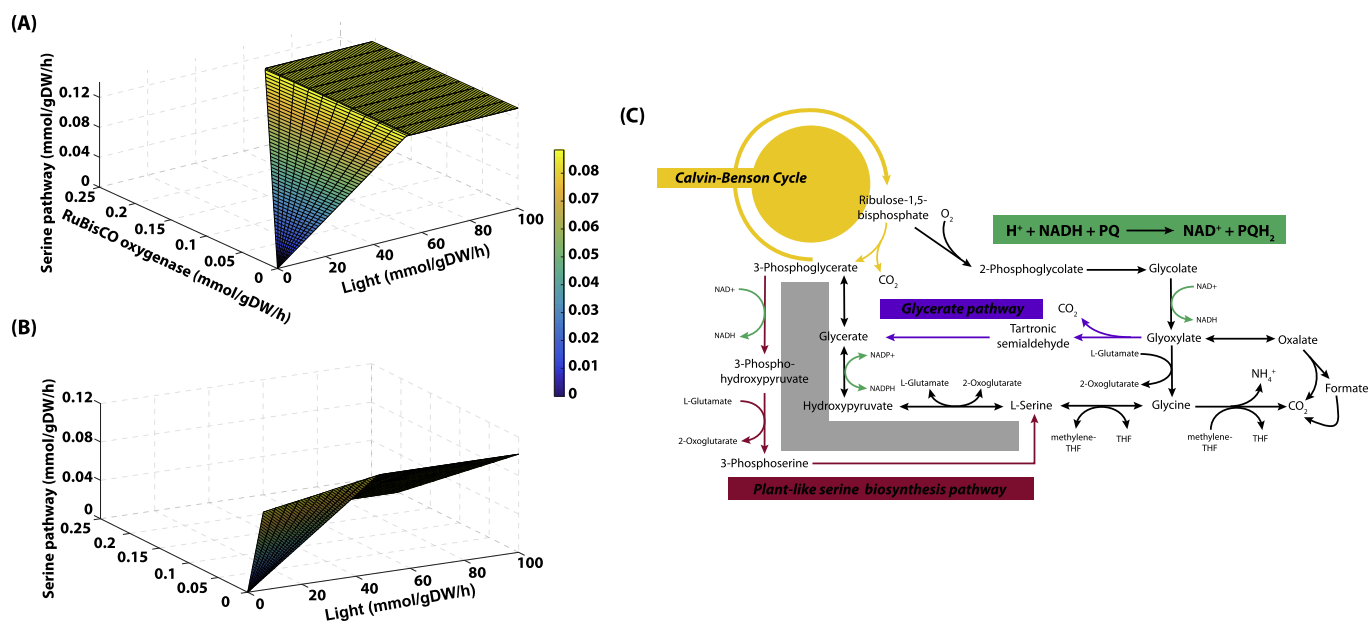
(A)

(B)

(C)

Fig. 5. Serine production via a light-independent pathway and photorespiratory pathway. (A) Flux through the serine pathway is shown as a function of light and Rubisco oxygenase flux, in presence of NADH/PQ dehydrogenase. Light-independent serine production pathway and photorespiration may both contribute to increasing serine production; (B) Flux through the serine pathway is shown as a function of light and RuBisCO oxygenase flux, in absence of NADH/PQ dehydrogenase. The increase in light-independent serine production pathway results in a decrease in photorespiration, while maintaining constant serine production; (C) Reaction scheme explaining the futile cycle that exists between the two pathways which results in increasing serine production as flux through light-independent pathway increases.
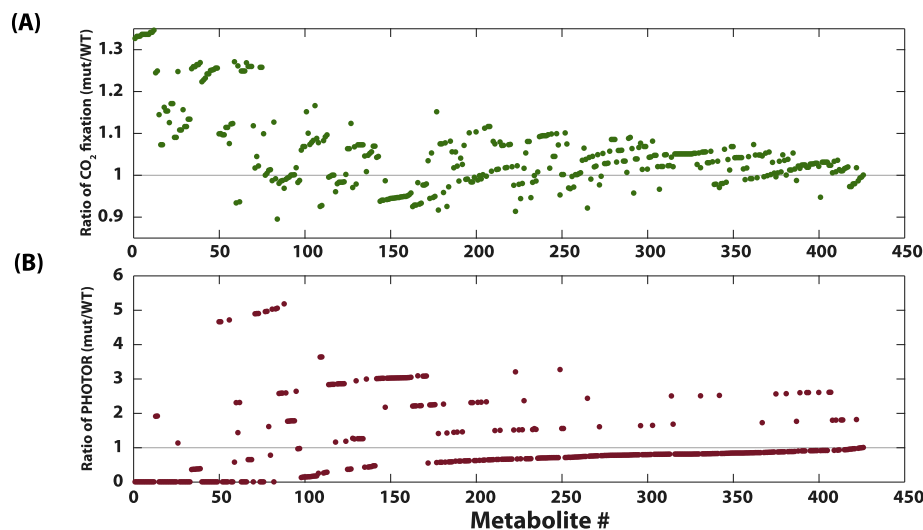
(A)

(B)

Fig. 6. Secretion of various metabolites may result in increased (A) CO_2 fixation or decrease in (B) photorespiration. Several metabolites show significant increase in CO_2 fixation and a decrease in photorespiration. Increased CO_2 remediation may be an additional criteria for picking target metabolites for commercial production. The y-axis represents the ratio between a strain that secretes a given metabolite and the wild-type value of either CO_2 fixation (A) or PHOTOR (B). The x-axis represents the metabolite number being secreted.

uptake [Fig. 6A], increase CO_2 fixation, and reduce photorespiratory flux [Fig. 6B] by the appropriate choice of bioproduct. We also quantified the effects of the choice of metabolic load on the rate of photosynthesis, defined as the ratio of the light uptake per carbon atom fixed. We found that 76 metabolic loads resulted in a higher photosynthetic rate than the wild-type rate.

Another possible application of these results concerns using *Synechocystis* sp. PCC 6803 for CO_2 remediation. CO_2 fixation reactions are faster than intracellular reactions contributing to growth, creating a bottleneck that prevents greater CO_2 fixation [88]. It has been argued that this bottleneck can be relieved by introducing a metabolic load, a high flux pathway, which transforms a metabolite into a molecule that does not contribute to the growth and can be secreted from the cell [89]. By inserting a metabolic load (production of bioproduct), our simulations of 426 metabolic loads show that it may be possible to separately process up to ~23% more inorganic carbon, produce about ~12% more oxygen, increase CO_2 fixation by ~35% [Fig. 6A], and

reduce photorespiration by ~87% [Fig. 6B] while producing bioproducts from different metabolites. We found 55 cases of zero photorespiration in our simulations. However, it has been argued that photorespiration may be a necessary process [84,87]. Our simulations suggest that a two-pronged approach, which involves applying a metabolic load (for industrially important molecules) and improving CO_2 fixation, may be a valuable strategy for CO_2 remediation.

3.6. Predicted metabolite yield and the trade-off with growth

We calculated the predicted theoretical yield for all the metabolites in the model after removing the nutrient metabolites, protons, inorganic phosphates (pi, ppi, and pppi), and extracellular metabolites. To obtain the theoretical yield for the remaining 829 metabolites, we maximized flux through that metabolite while allowing the growth-rate to fall to 0. However, theoretical yield calculations do not provide information on the growth-yield trade-off that is significant in practical

applications. We therefore performed flux maximization for each metabolite for 0%, 30%, 50%, 70%, and 100% of the wild-type growth rate. We found that 265 metabolites could not be secreted under any conditions. This leaves only 554 metabolites to analyze. For all growth shut-off conditions, the metabolite that achieves the maximum secretion (q_P) was molecular hydrogen (H_2) [Supplementary Table 6]. The plot of growth rate verses secretion (q_P) [Supplementary Fig. 5] shows a decreasing trend with a positive x-intercept and y-intercept.

Of the 265 metabolites that could not be secreted, 136 are the blocked metabolites discussed in Section 3.1. The remaining 129 metabolites cannot be secreted because of constraints placed by the network. Some examples of these metabolites are the individual S-states in the oxygen-evolving complex (S0, S1, S2, S3, and S4), which are modeled as individual metabolites. For the 554 remaining metabolites, we analyzed the flux distribution that maximizes product secretion under 50% reduction of the growth rate to obtain insight into the metabolic load on the network. Since high-dimensional flux vectors are difficult to analyze directly, we performed a Principal Component Analysis (PCA) and projected down to the space of the first three principal components that together capture 71% of the variation. In the space of the first two principal components, the flux distributions appeared spread out along two specific lines, one almost parallel to the x-axis with a small positive slope, and the other with a sharp negative slope (Fig. 7A). To analyze the importance of these lines, we examined the weights of the most important fluxes in each principal component, and the directions associated with increasing those fluxes. If a specific flux has a weight of *a1* in the expression for the first principal component and *a2* for the second, we plot the point (*a1,a2*) and connect it to the origin by a straight line (blue lines in Fig. 7). To interpret the direction described by these blue lines, it may be useful to consider a hypothetical flux distribution with only one non-zero flux of the ferredoxin oxidoreductase (FNOR) reaction. The projection of this flux vector on the principal component space in Fig. 7A will fall somewhere on the blue line corresponding to FNOR weights. Similarly, if only this flux increases, other things remaining constant, the data point would move outward parallel to this blue line. Using these directions as a guide to the eye, most of the variation along the first principal component appear to be variations in photon transfer reactions, photon absorption and electron transfer by PSII (PSIIa & PSIIb) and PSI, and in cytochrome *b6f* associated electron transfer reactions (CBFCu (a–c) & CBFCu (d–f)) with smaller contributions from the other fluxes shown. Most of the variation in the second principal component appears to be dominated by differences in PS1 absorption, cytochrome *b6f* reactions, FNOR, and a NAD(P)H dehydrogenase reaction (NDH1(1u)). We also plotted the flux distributions in the space of principal component 2 and principal component 3 [Fig. 7B]. Interestingly, principal component 3 separated the flux vectors into two distinct populations. Analysis of weights showed that principal component 3 was dominated by five reactions catalyzed by fructose 1,6-bisphosphate aldolase (FBA), transaldolase (TALA), sedoheptulose 1,7-bisphosphate phosphatase (SBP), sedoheptulose 1,7-bisphosphate D-glyceraldehyde-3-phosphate-lyase (FBA3), and fructose 1,7-bisphosphate phosphatase (FBP). These reactions are divided between reactions of the pentose phosphate pathway (TALA, SBP, FBA3) and glycolysis/gluconeogenesis (FBA, FBP). In Fig. 7B, the upper cluster consists of metabolites whose production causes an increase in flux through the pentose phosphate pathway (PPP) and leads to fructose 6-phosphate (F6P) production using transaldolase via sedoheptulose 7-phosphate. Metabolites that are part of this cluster include ribose phosphates, cofactors, nucleotides, terpenoids, phospholipids, oxaloacetate, and succinate. The lower cluster consists of metabolites that cause an increase in flux through glycolysis and led to F6P production using fructose bisphosphate phosphatase. Metabolites which are part of this cluster include nucleotide sugars, xylulose 5-phosphate, 3-phosphoglycerate, pyruvate, amino acids, lactate, ethanol, citrate, 2-oxoglutarate, acetyl-CoA, malate, sugars, and glycogen.
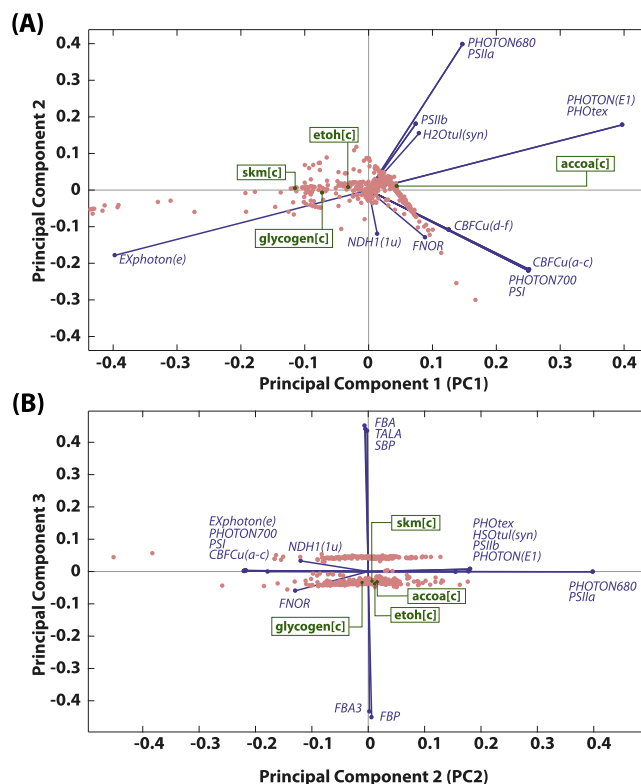

Fig. 7. Principal component analysis on flux distribution data obtained from simulation of secretion of a metabolite at 50% growth rate trade-off. (A) Flux distributions projected down to the space of the first two principal components (PC1 & PC2). Each dot corresponds to a flux distribution corresponding to a specific metabolite being secreted. The blue lines represent directions along which the flux corresponding to the indicated reaction is increasing. Only those reactions with large weight in the principal components were chosen (> 25% of the maximum variation within a principal component in either of the principal components). The abbreviations stand for the following: CBFCu (a–c) & CBFCu(d–f): Cytochrome b6f related electron transfer reaction; FNOR: Ferredoxin oxidoreductase related electron transfer; H2Otul(syn): water exchange from cytosol; NDH1_1u: NAD(P)H dehydrogenase reaction; PHOTON680 & PHOTON700: hypothetical reaction that converts a photon to a photon specific for PSII or PSI respectively; PHOTON (E1), PHOtex & ExPHOTON(e): hypothetical photon "transport" and "exchange" reactions; PSI: PSI reactions utilizing plastocyanin; PSIIa: photon absorption by PSII; PSIIb: PSII electron transfer reaction. (B) Flux distributions projected down to the space of the second and third principal components (PC2 & PC3). Comparison of PC2 and PC3 reveals the reactions which captured the most variation in flux distribution to either glycolysis or TCA cycle, and pentose phosphate pathway. The flux distribution corresponding to each metabolite being secreted is represented by a red dot. The reactions corresponding to the blue lines that are not mentioned in panel (A) are: FBA: Fructose 1,6-bisphosphate aldolase reaction; FBA3: Sedoheptulose 1,7-bisphosphate D-glyceraldehyde-3-phosphate-lyase, belongs to pentose phosphate pathway interconversions; FBP: fructose 1,7-bisphosphate phosphatase, belongs to gluconeogenesis; SBP: sedoheptulose 1,7-bisphosphate phosphatase, belongs to pentose phosphate pathway interconversions; TALA: Transaldolase, belongs to pentose phosphate pathway interconversions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We also analyzed the flux distribution of metabolites that may play a major role in the production of commercially important chemicals, including acetyl-CoA, shikimate, and ethanol. We chose these metabolites for their usefulness as commercially important molecules. Acetyl-CoA serves as a precursor for many important class of molecules such terpenoids, fatty acids, lipids, polyhydroxybutarate (PHB), tetracycline, and amino acids. Ethanol has a wide range of industrial uses, such as its use as an additive to gasoline used by cars and trucks. Shikimate is an important precursor for antibiotic synthesis. We also examined a fourth metabolite, glycogen, due to its importance in the night cycle. Of the two clusters separated by principal component 3, all four metabolites fall into the cluster in which fructose 6-phosphate is

produced via gluconeogenesis [Fig. 7B]. The FVA for acetyl-CoA production at a 50% growth reduction yielded 62 reactions for which flux varied by > 0.1 unit. These reactions belong to (i) pentose phosphate pathway such as transketolase 2, F6P and Xu5P phosphoketolases, and (ii) electron transfer machinery such as NADH dehydrogenase (active HCO3 transporters), NADPH dehydrogenase (utilizing 4 protons). We found that oxidative reactions (cytochrome oxidase), decarboxylation reactions (malic enzyme), and cytoplasmic electron transport (cytochrome *b6f*) were turned on. As a result, water and NADPH requirements were high, leading to increased water splitting/*PsiI* ratio. We observed similar trends for the other three metabolites, ethanol, shikimate and glycogen. Thus, principal component analysis of the flux distribution yields insights on the mode of operation of the metabolic network under metabolite demand, and flux distributions that cluster together show similar types of metabolic load.

3.7. Features of the autotrophic flux distribution

Autotrophic growth in the model is simulated by applying constraints to light uptake and carbon dioxide [Supplementary Fig. 6]. Other constraints on specific reactions are described in the Materials and methods. Given the flux constraint, an optimal solution was obtained using COBRA toolbox on MATLAB. Here, we leave the light uptake unconstrained such that maximum growth rate is obtained. Then, we calculated minimum light uptake corresponding to maximum growth rate. RuBisCO oxygenase activity was left unconstrained, but light-independent serine production was constrained such that ~3% of oxygenase was observed under low light conditions or optimal light conditions.

The simulations predicted some interesting flux distributions depending on whether the growth was simulated using bicarbonate (HCO_3^-) or carbon dioxide (CO_2) as carbon sources [Supplementary Table 7]. We used FVA at 100% growth to compare the flux distribution for these two conditions (CO_2 and HCO_3^-). Interestingly, we found that the range of fluxes was wider for simulations with CO_2 as compared to HCO_3^-. However, simulations with HCO_3^- resulted in the overall range being closer to the experimentally determined solution. The three exceptions where the experimentally observed flux value did not fall within calculated flux range for HCO_3^- growth simulation but did for growth with CO_2 are the reactions catalyzed by (i) G3P dehydrogenase (GAPDi (nadp)), (ii) triose phosphate isomerase (TPI), and (iii) pyruvate kinase (PYK). It should also be noted that optimal light calculated for HCO_3^- simulations was higher than that of CO_2 simulations. We also observed an increased proton influx for HCO_3^- growth simulations suggesting that ΔpH under these conditions is more positive than that for CO_2 growth simulations. The differences in growth condition seem to impact central carbon metabolism (CCM) flux distributions significantly. However, there is a lack of experimental studies that compare these two autotrophic growth conditions. These differences may be hard to test experimentally because of the predominant form of carbon present intracellularly at a pH of 7.5 to 9.5 is HCO_3^- and would have to chemically control for CO_2 to HCO_3^- conversion occurring inside the cell.

In both growth simulations, we noticed discrepancies with experimental results at three locations within CCM: (i) flux through oxidative pentose phosphate pathway (~0.004% of carbon) was much lower than observed experimentally (~13.5% of carbon), (ii) flux calculated (~1% of carbon) through phosphoglucoisomerase also contradict experiments (~19% of carbon), and (iii) acetyl-CoA (AcCoA) production from pyruvate was much lower (~0.02% of carbon) than that observed experimentally (35% of carbon). Interestingly, the last two discrepancies arise due to a single pathway which consists of two reactions involved in the production of acetyl-CoA [Supplementary Fig. 7]: (i) acetyl phosphate production from fructose 6-phosphate via xylulose 5-phosphate phosphoketolase (XU5PPK), and (ii) acetyl-CoA production from acetyl phosphate via phosphotransacetylase (PTAr). The first

discrepancy only occurs when optimal light uptake is simulated. Interestingly, we find that by increasing the light uptake by only an additional ~7% of the optimal value calculated, the flux through oxidative pentose phosphate pathway increases to the experimentally observed value. Thus, flux through this pathway is highly sensitive to light conditions and to excess electrons created by higher than optimal light conditions.

As expected, flux through TCA cycle is not cyclic, in either growth condition (CO_2 or HCO_3). Only 2/3 of the TCA cycle is active. About 1% of total carbon flows through the reaction cascade consisting succinate to oxaloacetate conversion, and about 2% of total carbon flows through the reaction cascade from oxaloacetate to 2-oxoglutarate (AKG) production. We do not see any flux through GABA shunt or succinate semialdehyde. As discussed previously, oxaloacetate not directed towards citric acid cycle was involved in the production of amino acids and nucleotides.

Finally, our model correctly predicts the directionality among the flux branching points but differs in the prediction of AcCoA, which is produced by F6PPK in our model rather than PDH. It must be noted that since many of these pathways were not identified in previous reconstructions, the experimental ^{13}C-based flux measurements do not contain values for these reactions. These values are estimated by us to have significantly high activity during light-dependent growth.

3.8. Dynamic flux balance analysis

To perform DFBA, we used the algorithm developed previously called DFBAlab [47]. Implementation of DFBAlab requires various user-defined parameters such as kinetic parameters of substrate consumption, light and dark phase time spans, initial conditions, and lexicographic scheme, i.e. a list of objective functions in order of priority. Here, we have analyzed various lexicographic schemes used to analyze growth dynamics of *Synechocystis* sp. PCC 6803. We employed 3 possible schemes based on the experimental observations (described in Table 1) with parameters as given in Table 2. Lexicographic schemes need to be specified separately for the light phase and the dark phase of growth. Our first scheme assumes that the primary objective of the organism is maximization of biomass production in the light phase and maximization of energy maintenance in the dark phase. However, this scheme was unable to reproduce experimental observations. The second scheme tests the hypothesis that glycogen production and glycogen utilization are the primary objectives of the organism in the two phases. This scheme too failed to reproduce experimental observations [Supplementary Figs. 9 & 10 and Supplementary Text 1]. We then hypothesized that the true objective function for a photosynthetic organism is probably a combination of both biomass maximization and glycogen accumulation in the light phase, while energy maintenance is likely to be the most important objective for the dark phase. We therefore constructed a new objective function that was a weighted convex combination of biomass growth and glycogen flux. This scheme was based on experimental observations that (i) glycogen levels increased during light phase and decreased during dark phase, but the decomposition of glycogen within the metabolic network does not contribute significantly to growth and carbon dioxide evolution takes place [90,91]; and (ii) ATP levels increased during light phase and decreased during dark phase.

As shown in Fig. 8, scheme 3 does qualitatively capture the dynamics of carbon dioxide, oxygen, glycogen, ATP, and growth during the light-dark cycle. Our analysis therefore indicates that photosynthetic organisms have metabolic objectives that are more complex than those of heterotrophic microbes and are composed of a combination of growth and storage requirements. Unfortunately, time-course metabolic flux measurements do not yet exist that would allow us to further explore these conclusions from our simulations.

Table 1
Priority list order for the lexicographic LP in scheme 1, scheme 2, and scheme 3.

Priority	Light Phase	Weights	Dark Phase	Weights
Scheme 1				
1	Maximize biomass production	1	Maximize ATP maintenance	1
2	Maximize ATP synthase (ATPase)	1	Minimize ATP synthase (ATPase)	1
3	Minimize ATP maintenance	1	Maximize biomass production	1
4	Maximize glycogen production	1	Maximize glycogen consumption	1
5	Maximize oxygen production	1	Maximize oxygen consumption	1
6	Maximize CO_2 consumption	1	Maximize CO_2 production	1
Scheme 2				
1	Maximize glycogen production	1	Maximize glycogen utilization	1
2	Maximize biomass production	1	Maximize biomass production	1
3	Maximize ATP synthase	1	Minimize ATP synthase	1
4	Minimize ATP maintenance	1	Maximize ATP maintenance	1
5	Maximize oxygen production	1	Maximize oxygen consumption	1
6	Maximize CO_2 consumption	1	Maximize CO_2 production	1
Scheme 3				
1	Maximize glycogen production	0.96	Maximize ATP maintenance	1
1	Maximize biomass production	0.04	Minimize ATP synthase (ATPase)	1
2	Maximize ATP synthase	1	Maximize biomass production	1
3	Minimize ATP maintenance	1	Maximize glycogen consumption	1
4	Maximize oxygen production	1	Maximize oxygen consumption	1
5	Maximize CO_2 consumption	1	Maximize CO_2 production	1

Table 2
Initial concentrations and parameters.

Variable/ Parameter	Name	Value	Units
ν_{CO_2},max	Maximum specific CO_2 uptake rate	0.249	mmol/gDW/h
K_{CO_2}	MM constant for CO_2 uptake	0.034	mmol/L
MT_{CO_2}	Mass transfer coefficient for CO_2	0.58	h^{-1}
K_{H,CO_2}	Henry's constant of CO_2	0.035	mol/L/atm
s^{CO_2}	concentration of CO_2 in atmosphere	0.035%	
$\nu_{max,glycogen}$	Maximum specific glycogen uptake rate	0.105	mmol/gDW/h
$K_{m,glycogen}$	MM constant for glycogen uptake	0.027	mmol/L
ν_{O_2},max	Maximum specific O_2 uptake rate	0.383	mmol/gDW/h
K_{O_2}	MM constant for O_2 uptake	0.135	mmol/L
MT_{O_2}	Mass transfer coefficient for O_2	0.6	h^{-1}
K_{H,O_2}	Henry's constant of CO_2	0.0013	mol/L/atm
s^{O_2}	concentration of O_2 in atmosphere	21%	
L	Depth of reactor	0.1	m
K_{e1}	Extinction coefficient constant 1	0.32	m^{-1}
K_{e2}	Extinction coefficient constant 2	0.03	m^{-1} (g/$m^3)^{-1}$
c_0	Initial CO_2 concentration	0.61	mmol/L
o_0	Initial O_2 concentration	0.125	mmol/L
g_o	Initial glycogen concentration	0.566	mmol/L
y_o	Initial biomass concentration	0.153	g/L
V	Volume of the reactor	140	L
F_{in}	Incoming flow rate	1.001	L/h
F_{out}	Outgoing flow rate	1.001	L/h

3.9. Heterotrophic flux distribution

Heterotrophic growth was simulated by allowing uptake of glucose in absence of light and maximizing the growth rate [Supplementary Fig. 8]. The glucose uptake was set to the experimentally determined value 0.85 mmol/gDW/h [42]. FVA simulations of reactions in central carbon metabolism were calculated and compared to experimentally determined fluxes [Supplementary Table 8] [42]. The main features of this comparison are discussed in the Supplementary Text 1.

Overall, we find that our model qualitatively predicts heterotrophic flux distributions, and in most cases there is quantitative agreement.
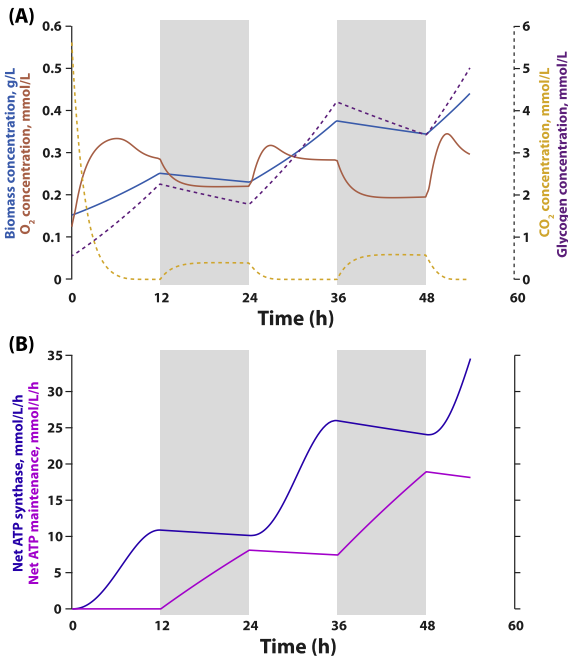

Fig. 8. Dynamic flux balance analysis using scheme 3. (A) Concentration of Biomass (purple), CO_2 (orange), O_2 (red), and glycogen (purple). (B) ATP synthase (blue) and ATP maintenance fluxes (red) over the simulated culture duration (LDLDL, 12:12:12:12:6). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Recent advances have been made in the identification of novel genes and reactions, such as plant-like serine production pathway [62] and a functional cyclic TCA cycle [79] in *Synechocystis* sp. PCC 6803. However, we find that in the light of these recent advances, repeating 13C metabolic flux analysis to experimentally measure fluxes within glycolysis and pentose phosphate pathway under heterotrophic growth conditions may lead to better understanding of cyanobacterial metabolism under dark growth conditions.
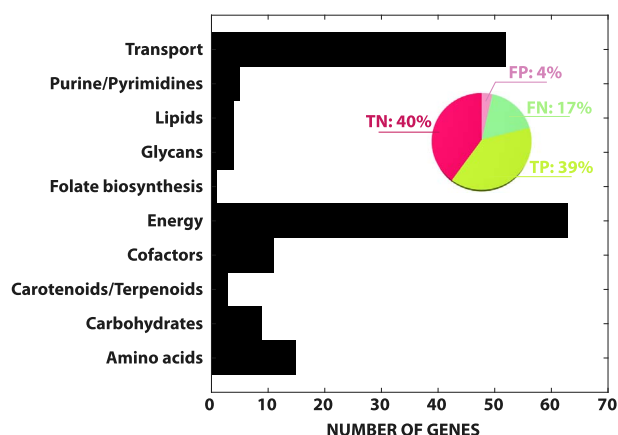
Fig. 9. Gene deletion analysis. The distribution of genes, for which gene deletion was compared with experimental results, among various different subsystems. The inset shows true positive (TP), false positive (FP), true negative (TN), and false negative (FN) percentages.

3.10. Single gene deletion analysis (experiment vs. simulation under autotrophic conditions)

Out of the 10 metabolic reconstructions published in the past, only two [31,32] addressed the performance of the model with respect to experimental data on single gene deletions. Here, we compared single gene deletions with experimental studies. Using information in Cyanobase [38,39] supplemented by a literature search, we located experimental evidence on 167 single gene deletions (details in Supplementary Table 9) and compared the observed phenotype with predictions of our model. When experimental results were ambiguous, such as very low growth for example, we assumed both possibilities and report our results below as a range. Comparison between the data on 167 gene deletions we collected and the model predictions under the respective gene deletions resulted in 39% true positive (growth predicted and observed), 31–40% true negative (growth neither predicted, nor observed), 4% false positive (growth predicted, but not observed), and 17–27% false negative (growth not predicted, but observed) results [Fig. 9] [Supplementary Table 9]. Therefore, out of 167 genes, 70–79% resulted in matches and 30–21% resulted in mismatches.

The mismatches that were predicted as essential belong mainly to photosynthesis (12 genes), peptidoglycan biosynthesis (4 genes), ion transporters (2 genes), ubiquinone and pterpenoid biosynthesis (2 genes). For example, sll1653 and sll2010 which make transcripts of demethylphylloquinone methyltransferase and UDP-*N*-acetylmuramoyl-L-alanyl-D-glutamate synthetase, respectively, are among this category of mismatches. The reason that deletion of these genes is predicted to be lethal in the model is that the products of the reaction catalyzed by these enzymes, which do not have an isozyme, are directly involved in the biomass growth objective. The mathematical formulation of the growth rate objective is rigid and determines the optimal flux distribution. However, in reality, the organism can probably accommodate a different biomass composition. For example, non-essential biomass components like terpenoids and peptidoglycans may be present in different amounts or may be substitutable by some other biomass components, and therefore show non-lethality. The rigidity of the growth rate objective function prevents growth being predicted in such instances by making all components of the biomass be equally essential. Hence, all the single deletions in non-isozymic or single pathway genes involved in the synthesis of a non-essential biomass component will also lead up to lethality.

A global gene essentiality analysis was also carried out by deleting every gene in the model, one at a time. The numbers of essential genes predicted under autotrophic, heterotrophic, and mixotrophic conditions were 389, 307, and 306, respectively. A list of 305 (37.2%) essential

genes appeared to be common to all the growth conditions. Consistent with previous studies [31,32], our model continues to show a significantly low genetic robustness (~ 47%) of *Synechocystis* under autotrophic conditions. Addition of newly added genes and pathways led to improvement in prediction of certain transporters. For example, a single gene (slr1145) deletion responsible for sodium transport through the cell which was predicted to be lethal in previous models, but is non-lethal in our model due to improvements in gene association (an isozymic multimer, sll1102 & sll1103 & sll1104) and addition of a literature-curated mechanism [92].

4. Conclusion

Cyanobacteria have garnered much interest as a resource for harnessing naturally available sunlight and carbon dioxide to produce commercially important chemicals. Here, we present a revised and updated genome-scale metabolic reconstruction of a cyanobacterium, *Synechocystis* sp. PCC6 803. This model makes improvements to the previously published models by including thermodynamics, incorporating a physiological mechanism for generating photorespiratory flux, and increases the level of detail of the molecular mechanisms involved in photosynthesis (PS) and oxidative phosphorylation (OXPHOS). Thermodynamic calculations allow assessment of reaction directionality and identification of thermodynamically unfavorable cycles.

FVA also allowed us to identify the trade-offs between two primary serine biosynthesis mechanisms in *Synechocystis* sp. PCC 6803: the photorespiratory pathway and serine biosynthesis. We propose that additional metabolic roles of photorespiration may make the process indispensible under normal conditions. Our hypothesis is that these metabolic roles may be emerging from dependency on metabolites involved in 2-phosphoglycolate metabolism. In this paper we applied a constraint-based scheme that assumes an enzymatic bottleneck in the light-independent serine production pathway. However, other possible schemes could be designed such as growth dependence on 2-phosphoglycolate (by making it part of the biomass growth equation) or implementing a demand flux for 2-phosphoglycolate. We hope that our work will spur more modeling and experiments to explore these hypotheses further.

Our analysis is an example of how systems biology can be useful in designing strategies to improve CO_2 remediation, and producing bioproducts, simultaneously. We also found a difference in water uptake and light absorption when growth was simulated using CO_2 and bicarbonate from the water. However, the differences in growth using CO_2 and HCO_3 have yet to be shown experimentally. We presented a global analysis of metabolite production from the organism and showed that principal component analysis could help in providing insights on the origins of metabolic load when over-producing any metabolite.

We conducted DFBA using a novel method called the direct approach. We studied three optimization schemes and found that to match qualitative observations of metabolite kinetics during the light-dark cycle, we require an objective that consists of both growth maximization and glycogen maximization. This objective function yielded a better match of the simulation results with the data. Lexicographic priority schemes may in fact be better than single-objective FBA for predicting the flux distribution of photosynthetic and even non-photosynthetic organisms under dynamic and static conditions.

We also simulated gene deletions and found that though growth under autotrophic conditions continues to exhibit low genetic robustness the percentage of lethal gene deletions were lower than that observed in previous models. Therefore, these results suggest the metabolic models of *Synechocystis* sp. PCC 6803 are still in its infancy, and as more data becomes available, the genetic robustness of future models may increase. However, we were able to compare 167 gene deletions in the model to Cyanobase and literature surveys, and were able to predict 77% accurately using iSynCJ816.

Though the expanded and updated model gives better agreement of flux distributions and mutant growth rates under autotrophic growth with experiments, there still exist many gaps in our understanding of the experimentally determined flux distributions. These include a balance between serine biosynthesis and photorespiration and the high flux variability in PS and OXPHOS metabolic network. Finally, most previous analyses and experimental data have been largely limited to constant light conditions. However during changing light conditions, the metabolic composition changes resulting in high variability of internal fluxes. Therefore, additional 'omics data needs to be coupled with such stoichiometric models to better explain flux distributions of photosynthetic microorganisms.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.algal.2017.09.013.

Acknowledgements

Author contributions

AP, CAMP and CJJ designed research, CJJ and AP carried out the research, and CJJ, AP and CAMP wrote the paper.

No conflicts, informed consent, human or animal rights applicable.

References

[1] J.J. Brocks, G.A. Logan, R. Buick, R.E. Summons, Archean molecular fossils and the early rise of eukaryotes, Science 285 (1999) 1033–1036, http://dx.doi.org/10.1126/science.285.5430.1033.

[2] B.D. Rae, B.M. Long, L.F. Whitehead, B. Förster, M.R. Badger, G.D. Price, Cyanobacterial carboxysomes: microcompartments that facilitate CO_2 fixation, J. Mol. Microbiol. Biotechnol. 23 (2013) 300–307, http://dx.doi.org/10.1159/000351342.

[3] N.E. Nozzi, J.W.K. Oliver, S. Atsumi, Cyanobacteria as a platform for biofuel production, Front. Bioeng. Biotechnol. Lipid (2013) 1–6, http://dx.doi.org/10.3389/fbioe.2013.00007.

[4] S. Vijayakumar, M. Menakha, Pharmaceutical applications of cyanobacteria-a review, JACME 5 (2015) 15–23, http://dx.doi.org/10.1016/j.jacme.2015.02.004.

[5] K. Gademann, Out in the green: biologically active metabolites produced by cyanobacteria, Chimia (Aarau) 65 (2011) 416–419, http://dx.doi.org/10.2533/chimia.2011.416.

[6] Y. Yu, L. You, D. Liu, W. Hollinshead, Y.J. Tang, F. Zhang, Development of *Synechocystis* sp. PCC 6803 as a phototrophic cell factory, Mar. Drugs 11 (2013) 2894–2916, http://dx.doi.org/10.3390/md11082894.

[7] J.D. Orth, I. Thiele, B.Ø. Palsson, What is flux balance analysis? Nat. Biotechnol. 28 (2010) 245–248, http://dx.doi.org/10.1038/nbt.1614.

[8] Y.K. Oh, B.O. Palsson, S.M. Park, C.H. Schilling, R. Mahadevan, Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data, J. Biol. Chem. 282 (2007) 28791–28799, http://dx.doi.org/10.1074/jbc.M703759200.

[9] A.M. Feist, C.S. Henry, J.L. Reed, M. Krummenacker, A.R. Joyce, P.D. Karp, L.J. Broadbelt, V. Hatzimanikatis, B.Ø. Palsson, A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information, Mol. Syst. Biol. 3 (2007) 121, http://dx.doi.org/10.1038/msb4100155.

[10] R. Mahadevan, C.H. Schilling, The effects of alternate optimal solutions in constraint-based genome-scale metabolic models, Metab. Eng. 5 (2003) 264–276, http://dx.doi.org/10.1016/j.ymben.2003.09.002.

[11] Y.C. Liao, T.W. Huang, F.C. Chen, P. Charusanti, J.S.J. Hong, H.Y. Chang, S.F. Tsai, B.O. Palsson, C.A. Hsiung, An experimentally validated genome-scale metabolic reconstruction of *Klebsiella pneumoniae* MGH 78578, iYL1228, J. Bacteriol. 193 (2011) 1710–1717, http://dx.doi.org/10.1128/JB.01218-10.

[12] A. Montagud, E. Navarro, P. Fernández de Córdoba, J.F. Urchueguía, K.R. Patil, Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium, BMC Syst. Biol. 4 (2010) 156, http://dx.doi.org/10.1186/1752-0509-4-156.

[13] J.L. Hjersted, M.A. Henson, Steady-state and dynamic flux balance analysis of ethanol production by *Saccharomyces cerevisiae*, IET Syst. Biol. 3 (2009) 167–179, http://dx.doi.org/10.1049/iet-syb.2008.0103.

[14] J.S. Edwards, B.O. Palsson, Metabolic flux balance analysis and the in silico analysis of *Escherichia coli* K-12 gene deletions, BMC Bioinf. 1 (2000) 1, http://dx.doi.org/10.1186/1471-2105-1-1.

[15] D. Segrè, D. Vitkup, G.M. Church, Analysis of optimality in natural and perturbed metabolic networks, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 15112–15117, http://dx.doi.org/10.1073/pnas.232349399.

[16] C.J. Joshi, A. Prasad, Epistatic interactions among metabolic genes depend upon environmental conditions, Mol. BioSyst. 10 (2014) 2578–2589, http://dx.doi.org/10.1039/c4mb00181h.

[17] D. Segrè, A. Deluna, G.M. Church, R. Kishony, Modular epistasis in yeast metabolism, Nat. Genet. 37 (2005) 77–83, http://dx.doi.org/10.1038/ng1489.

[18] E.S. Snitkin, D. Segrè, Epistatic interaction maps relative to multiple metabolic phenotypes, PLoS Genet. 7 (2011) e1001294, , http://dx.doi.org/10.1371/journal.pgen.1001294.

[19] D. Segrè, C.J. Marx, Introduction to focus issue: genetic interactions, Chaos 20 (2010) 26101, http://dx.doi.org/10.1063/1.3456057.

[20] P.C. Phillips, Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems, Nat. Rev. Genet. 9 (2008) 855–867, http://dx.doi.org/10.1038/nrg2452.

[21] L.F. de Figueiredo, T.I. Gossmann, M. Ziegler, S. Schuster, Pathways analysis of NAD + metabolism, Biochem. J. 439 (2011) 341–348, http://dx.doi.org/10.1042/BJ20110320.

[22] I. Rocha, J. Förster, J. Nielsen, Microbial Gene Essentiality - Protocols and Bioinformatics, Humana Press, 2008, http://dx.doi.org/10.1007/978-1-59745-321-9.

[23] P.F. Suthers, A. Zomorrodi, C.D. Maranas, Genome-scale gene/reaction essentiality and synthetic lethality analysis, Mol. Syst. Biol. 5 (2009) 301, http://dx.doi.org/10.1038/msb.2009.56.

[24] N. Jamshidi, B.Ø. Palsson, Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ661 and proposing alternative drug targets, BMC Syst. Biol. 1 (2007) 26, http://dx.doi.org/10.1186/1752-0509-1-26.

[25] A.A. Shastri, J.A. Morgan, Flux balance analysis of photoautotrophic metabolism, Biotechnol. Prog. 21 (2005) 1617–1626, http://dx.doi.org/10.1021/bp050246d.

[26] S.J. Hong, C.G. Lee, Evaluation of central metabolism based on a genomic database of *Synechocystis* PCC6803, Biotechnol. Bioprocess Eng. 12 (2007) 165–173, http://dx.doi.org/10.1007/BF03028644.

[27] P. Fu, Genome-scale modeling of *Synechocystis* sp. PCC 6803 and prediction of pathway insertion, J. Chem. Technol. Biotechnol. 84 (2009) 473–483, http://dx.doi.org/10.1002/jctb.2065.

[28] H. Knoop, Y. Zilliges, W. Lockau, R. Steuer, The metabolic network of *Synechocystis* sp. PCC 6803: systemic properties of autotrophic growth, Plant Physiol. 154 (2010) 410–422, http://dx.doi.org/10.1104/pp.110.157198.

[29] A. Montagud, A. Zelezniak, E. Navarro, P.F. de Córdoba, J.F. Urchueguía, K.R. Patil, Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium *Synechocystis* sp. PCC6803, Biotechnol. J. 6 (2011) 330–342, http://dx.doi.org/10.1002/biot.201000109.

[30] K. Yoshikawa, Y. Kojima, T. Nakajima, C. Furusawa, T. Hirasawa, H. Shimizu, Reconstruction and verification of a genome-scale metabolic model for *Synechocystis* sp. PCC6803, Appl. Microbiol. Biotechnol. 92 (2011) 347–358, http://dx.doi.org/10.1007/s00253-011-3559-x.

[31] J. Nogales, S. Gudmundsson, E.M. Knight, B.O. Palsson, I. Thiele, Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis, Proc. Natl. Acad. Sci. 109 (2012) 2678–2683, http://dx.doi.org/10.1073/pnas.1117907109.

[32] R. Saha, A.T. Verseput, B.M. Berla, T.J. Mueller, H.B. Pakrasi, C.D. Maranas, Reconstruction and comparison of the metabolic potential of Cyanobacteria *Cyanothece* sp. ATCC 51142 and *Synechocystis* sp. PCC 6803, PLoS One 7 (2012) e48285, , http://dx.doi.org/10.1371/journal.pone.0048285.

[33] H. Knoop, M. Gründel, Y. Zilliges, R. Lehmann, S. Hoffmann, W. Lockau, R. Steuer, Flux balance analysis of cyanobacterial metabolism: the metabolic network of *Synechocystis* sp. PCC 6803, PLoS Comput. Biol. 9 (2013) e1003081, , http://dx.doi.org/10.1371/journal.pcbi.1003081.

[34] R. Mohammadi, J. Fallah-Mehrabadi, G. Bidkhori, J. Zahiri, M. Javad Niroomand, A. Masoudi-Nejad, A systems biology approach to reconcile metabolic network models with application to *Synechocystis* sp. PCC 6803 for biofuel production, Mol. BioSyst. 12 (2016) 2552–2561, http://dx.doi.org/10.1039/c6mb00119j.

[35] M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, Data, information, knowledge and principle: back to metabolism in KEGG, Nucleic Acids Res. 42 (2014) D199–D205, http://dx.doi.org/10.1093/nar/gkt1076.

[36] M. Kanehisa, The KEGG database, Novartis Found. Symp. 247 (2002) 244–252 91-101-103, 119–128 https://doi.org/10.1038/nbt991.

[37] M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, M. Tanabe, KEGG for integration and interpretation of large-scale molecular data sets, Nucleic Acids Res. 40 (2012) gkr988, , http://dx.doi.org/10.1093/nar/gkr988.

[38] M. Nakao, S. Okamoto, M. Kohara, T. Fujishiro, T. Fujisawa, S. Sato, S. Tabata, T. Kaneko, Y. Nakamura, CyanoBase: the cyanobacteria genome database update 2010, Nucleic Acids Res. 38 (2009) D379–D381, http://dx.doi.org/10.1093/nar/gkp915.

[39] T. Fujisawa, S. Okamoto, T. Katayama, M. Nakao, H. Yoshimura, H. Kajiya-Kanegae, S. Yamamoto, C. Yano, Y. Yanaka, H. Maita, T. Kaneko, S. Tabata, Y. Nakamura, CyanoBase and RhizoBase: databases of manually curated annotations for cyanobacterial and rhizobial genomes, Nucleic Acids Res. 42 (2014) D666-70-D670 https://doi.org/10.1093/nar/gkt1145.

[40] W.F.J. Vermaas, Photosynthesis and Respiration in Cyanobacteria, (2001), http://dx.doi.org/10.1038/npg.els.0001670.

[41] J.D. Young, A.A. Shastri, G. Stephanopoulos, J.A. Morgan, Mapping photoautotrophic metabolism with isotopically nonstationary ^{13}C flux analysis, Metab. Eng. 13 (2011) 656–665, http://dx.doi.org/10.1016/j.ymben.2011.08.002.

[42] C. Yang, Q. Hua, K. Shimizu, Metabolic flux analysis in *Synechocystis* using isotope distribution from ^{13}C-labeled glucose, Metab. Eng. 4 (2002) 202–216, http://dx.doi.org/10.1006/mben.2002.0226.

[43] A. Varma, B.O. Palsson, Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110, Appl. Environ. Microbiol. 60 (1994) 3724–3731 (doi:PMC201879).

[44] R. Mahadevan, J.S. Edwards, F.J. Doyle, Dynamic flux balance analysis of diauxic growth in *Escherichia coli*, Biophys. J. 83 (2002) 1331–1340, http://dx.doi.org/10.1016/S0006-3495(02)73903-9.

[45] T.J. Hanly, M.A. Henson, Dynamic metabolic modeling of a microaerobic yeast co-culture: predicting and optimizing ethanol production from glucose/xylose mixtures, Biotechnol. Biofuels. 6 (2013) 44, http://dx.doi.org/10.1186/1754-6834-6-44.

[46] G.M. Oddone, D.A. Mills, D.E. Block, A dynamic, genome-scale flux model of Lactococcus lactis to increase specific recombinant protein expression, Metab. Eng. 11 (2009) 367–381, http://dx.doi.org/10.1016/j.ymben.2009.07.007.

[47] J.A. Gomez, K. Höffner, P.I. Barton, DFBAlab: a fast and reliable MATLAB code for dynamic flux balance analysis, BMC Bioinf. 15 (2014) 409, http://dx.doi.org/10.1186/s12859-014-0409-8.

[48] N. Jamshidi, J.S. Edwards, T. Fahland, G.M. Church, B.O. Palsson, Dynamic simulation of the human red blood cell metabolic network, Bioinformatics 17 (2001) 286–287 http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=11294796&retmode=ref&cmd=prlinks.

[49] J.L. Hjersted, M.A. Henson, Optimization of fed-batch *Saccharomyces cerevisiae* fermentation using dynamic flux balance models, Biotechnol. Prog. 22 (2006) 1239–1248, http://dx.doi.org/10.1021/bp060059v.

[50] R. Caspi, T. Altman, R. Billington, K. Dreher, H. Foerster, C.A. Fulcher, T.A. Holland, I.M. Keseler, A. Kothari, A. Kubo, M. Krummenacker, M. Latendresse, L.A. Mueller, Q. Ong, S. Paley, P. Subhraveti, D.S. Weaver, D. Weerasinghe, P. Zhang, P.D. Karp, The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases, Nucleic Acids Res. 42 (2014) D459-71-D471 https://doi.org/10.1093/nar/gkt1103.

[51] T. Kaneko, S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, S. Tabata, Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions, DNA Res. 3 (1996) 109–136, http://dx.doi.org/10.1093/dnares/3.3.109.

[52] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, T.L. Madden, NCBI BLAST: a better web interface, Nucleic Acids Res. 36 (2008) W5–W9, http://dx.doi.org/10.1093/nar/gkn201.

[53] M. Scheer, A. Grote, A. Chang, I. Schomburg, C. Munaretto, M. Rother, C. Söhngen, M. Stelzer, J. Thiele, D. Schomburg, BRENDA, the enzyme information system in 2011, Nucleic Acids Res. 39 (2011) D670–6, http://dx.doi.org/10.1093/nar/gkq1089.

[54] K. Degtyarenko, P. De Matos, M. Ennis, J. Hastings, M. Zbinden, A. Mcnaught, R. Alcántara, M. Darsow, M. Guedj, M. Ashburner, ChEBI: A database and ontology for chemical entities of biological interest, Nucleic Acids Res. 36 (2008), http://dx.doi.org/10.1093/nar/gkm791.

[55] Y. Wang, J. Xiao, T.O. Suzek, J. Zhang, J. Wang, S.H. Bryant, PubChem: a public information system for analyzing bioactivities of small molecules, Nucleic Acids Res. 37 (2009), http://dx.doi.org/10.1093/nar/gkp456.

[56] J.W. Cooley, W.F.J. Vermaas, Succinate dehydrogenase and other respiratory pathways in thylakoid membranes of *Synechocystis* sp. PCC 6803: capacity comparisons and physiological function, J. Bacteriol. 183 (2001) 4251–4258, http://dx.doi.org/10.1128/JB.183.14.4251-4258.2001.

[57] A. Latifi, M. Ruiz, C.C. Zhang, Oxidative stress in cyanobacteria, FEMS Microbiol. Rev. 33 (2009) 258–278, http://dx.doi.org/10.1111/j.1574-6976.2008.00134.x.

[58] W. Ma, T. Ogawa, Y. Shen, H. Mi, Changes in cyclic and respiratory electron transport by the movement of phycobilisomes in the cyanobacterium *Synechocystis* sp. strain PCC 6803, Biochim. Biophys. Acta Bioenerg. 1767 (2007) 742–749, http://dx.doi.org/10.1016/j.bbabio.2007.01.017.

[59] J. Barber, Photosystem II: its function, structure, and implications for artificial photosynthesis, Biochem. Biokhimiia. 79 (2014) 185–196, http://dx.doi.org/10.1134/S0006297914030031.

[60] H.-W. Heldt, B. Piechulla, Plant Biochem. (2011) 163–191, http://dx.doi.org/10.1016/B978-0-12-384986-1.00006-5.

[61] P.J. Lea, R.C. Leegood, Plant Biochemistry and Molecular Biology, John Wiley, 1999.

[62] F. Klemke, A. Baier, H. Knoop, R. Kern, J. Jablonsky, G. Beyer, T. Volkmer, R. Steuer, W. Lockau, M. Hagemann, Identification of the light-independent phosphoserine pathway as additional source for serine in the cyanobacterium *Synechocystis* sp. PCC 6803, Microbiology (2015), http://dx.doi.org/10.1099/mic.0.000055.

[63] I. Thiele, B.Ø. Palsson, A protocol for generating a high-quality genome-scale metabolic reconstruction, Nat. Protoc. 5 (2010) 93–121, http://dx.doi.org/10.1038/nprot.2009.203.

[64] S.M. Keating, B.J. Bornstein, A. Finney, M. Hucka, SBMLToolbox: an SBML toolbox for MATLAB users, Bioinformatics 22 (2006) 1275–1277, http://dx.doi.org/10.1093/bioinformatics/btl111.

[65] J. Schellenberger, R. Que, R.M.T. Fleming, I. Thiele, J.D. Orth, A.M. Feist, D.C. Zielinski, A. Bordbar, N.E. Lewis, S. Rahmanian, J. Kang, D.R. Hyduke, B.Ø. Palsson, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0, Nat. Protoc. 6 (2011) 1290–1307, http://dx.doi.org/10.1038/nprot.2011.308.

[66] D. Hyduke, D. Hyduke, J. Schellenberger, R. Que, R. Fleming, I. Thiele, J. Orth, A. Feist, D. Zielinski, A. Bordbar, N. Lewis, S. Rahmanian, J. Kang, B. Palsson, COBRA Toolbox 2.0, Protoc. Exch. (2011) 1–35, http://dx.doi.org/10.1038/protex.2011.234.

[67] M.D. Jankowski, C.S. Henry, L.J. Broadbelt, V. Hatzimanikatis, Group contribution method for thermodynamic analysis of complex metabolic networks, Biophys. J. 95 (2008) 1487–1499, http://dx.doi.org/10.1529/biophysj.107.124784.

[68] B.D. Bennett, E.H. Kimball, M. Gao, R. Osterhout, S.J. van Dien, J.D. Rabinowitz, Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*, Nat. Chem. Biol. 5 (2009) 593–599, http://dx.doi.org/10.1038/nchembio.186.

[69] J. Nogales, S. Gudmundsson, E.M. Knight, B.O. Palsson, I. Thiele, Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 2678–2683, http://dx.doi.org/10.1073/pnas.1117907109.

[70] A.M. Feist, B.O. Palsson, The biomass objective function, Curr. Opin. Microbiol. 13 (2010) 344–349, http://dx.doi.org/10.1016/j.mib.2010.03.003.

[71] J.L. Reed, B. Palsson, Genome-scale in silico models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states, Genome Res. 14 (2004) 1797–1805, http://dx.doi.org/10.1101/gr.2546004.

[72] K. Höffner, S.M. Harwood, P.I. Barton, A reliable simulator for dynamic flux balance analysis, Biotechnol. Bioeng. 110 (2013) 792–802, http://dx.doi.org/10.1002/bit.24748.

[73] S.M. Harwood, K. Höffner, P.I. Barton, Efficient solution of ordinary differential equations with a parametric lexicographic linear program embedded, Numer. Math. 133 (2016) 623–653, http://dx.doi.org/10.1007/s00211-015-0760-3.

[74] S.S. Fong, B.Ø. Palsson, Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes, Nat. Genet. 36 (2004) 1056–1058, http://dx.doi.org/10.1038/ng1432.

[75] Z.-X. Xu, Constrain-based analysis of gene deletion on the metabolic flux redistribution of *Saccharomyces cerevisiae*, J. Biomed. Sci. Eng. 1 (2008) 121–126, http://dx.doi.org/10.4236/jbise.2008.12020.

[76] N. Battchikova, M. Eisenhut, E.M. Aro, Cyanobacterial NDH-1 complexes: novel insights and remaining puzzles, Biochim. Biophys. Acta Bioenerg. 1807 (2011) 935–944, http://dx.doi.org/10.1016/j.bbabio.2010.10.017.

[77] M. Matsuo, T. Endo, K. Asada, Properties of the respiratory NAD(P)H dehydrogenase isolated from the cyanobacterium *Synechocystis* PCC6803, Plant Cell Physiol. 39 (1998) 263–267.

[78] C.A. Howitt, P.K. Udall, W.F. Vermaas, Type 2 NADH dehydrogenases in the cyanobacterium *Synechocystis* sp. strain PCC 6803 are involved in regulation rather than respiration, J. Bacteriol. 181 (1999) 3994–4003 (doi:0021-9193/99/$04.00 + 0).

[79] S. Zhang, D.A. Bryant, The tricarboxylic acid cycle in Cyanobacteria, Science 334 (2011) 1551–1553 (80-), https://doi.org/10.1126/science.1210858.

[80] M.R. Badger, S. von Caemmerer, S. Ruuska, H. Nakano, Electron flow to oxygen in higher plants and algae: rates and control of direct photoreduction (Mehler reaction) and rubisco oxygenase, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 355 (2000) 1433–1446, http://dx.doi.org/10.1098/rstb.2000.0704.

[81] S. Bailey, A. Melis, K.R.M. Mackey, P. Cardol, G. Finazzi, G. van Dijken, G.M. Berg, K. Arrigo, J. Shrager, A. Grossman, Alternative photosynthetic electron flow to oxygen in marine *Synechococcus*, Biochim. Biophys. Acta Bioenerg. 1777 (2008) 269–276, http://dx.doi.org/10.1016/j.bbabio.2008.01.002.

[82] J. Huege, J. Goetze, D. Schwarz, H. Bauwe, M. Hagemann, J. Kopka, Modulation of the major paths of carbon in photorespiratory mutants of *Synechocystis*, PLoS One 6 (2011), http://dx.doi.org/10.1371/journal.pone.0016278.

[83] Y. Allahverdiyeva, M. Ermakova, M. Eisenhut, P. Zhang, P. Richaud, M. Hagemann, L. Cournac, E.M. Aro, Interplay between flavodiiron proteins and photorespiration in *Synechocystis* sp. PCC 6803, J. Biol. Chem. 286 (2011) 24007–24014, http://dx.doi.org/10.1074/jbc.M111.223289.

[84] H. Bauwe, M. Hagemann, R. Kern, S. Timm, Photorespiration has a dual origin and manifold links to central metabolism, Curr. Opin. Plant Biol. 15 (2012) 269–275, http://dx.doi.org/10.1016/j.pbi.2012.01.008.

[85] M. Hagemann, A.P. Weber, M. Eisenhut, Photorespiration: origins and metabolic integration in interacting compartments, J. Exp. Bot. 67 (2016) 2915–2918, http://dx.doi.org/10.1093/jxb/erw178.

[86] C. Hackenberg, A. Engelhardt, H.C.P. Matthijs, F. Wittink, H. Bauwe, A. Kaplan, M. Hagemann, Photorespiratory 2-phosphoglycolate metabolism and photoreduction of O₂ cooperate in high-light acclimation of *Synechocystis* sp. strain PCC 6803, Planta 230 (2009) 625–637, http://dx.doi.org/10.1007/s00425-009-0972-9.

[87] M. Eisenhut, W. Ruth, M. Haimovich, H. Bauwe, A. Kaplan, M. Hagemann, The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 17199–17204, http://dx.doi.org/10.1073/pnas.0807043105.

[88] H. Knoop, R. Steuer, A computational analysis of stoichiometric constraints and trade-offs in Cyanobacterial biofuel production, Front. Bioeng. Biotechnol. 3 (2015) 47, http://dx.doi.org/10.3389/fbioe.2015.00047.

[89] J.W.K. Oliver, S. Atsumi, A carbon sink pathway increases carbon productivity in cyanobacteria, Metab. Eng. 29 (2015) 106–112, http://dx.doi.org/10.1016/j.ymben.2015.03.006.

[90] R. Saha, D. Liu, A. Hoynes-O'Connor, M. Liberton, J. Yu, M. Bhattacharyya-Pakrasi, A. Balassy, F. Zhang, T.S. Moon, C.D. Maranas, H.B. Pakrasi, Diurnal regulation of cellular processes in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803: insights from transcriptomic, fluxomic, and physiological analyses, MBio 7 (2016) e00464-16, , http://dx.doi.org/10.1128/mBio.00464-16.

[91] S.B. Gaudana, S. Krishnakumar, S. Alagesan, M.G. Digmurti, G.A. Viswanathan, M. Chetty, P.P. Wangikar, Rhythmic and sustained oscillations in metabolism and gene expression of *Cyanothece* sp. ATCC 51142 under constant light, Front. Microbiol. 4 (2013), http://dx.doi.org/10.3389/fmicb.2013.00374.

[92] M.J. Quintero, M.L. Montesinos, A. Herrero, E. Flores, Identification of genes encoding amino acid permeases by inactivation of selected ORFs from the *Synechocystis* genomic sequence, Genome Res. 11 (2001) 2034–2040, http://dx.doi.org/10.1101/gr.196301.