Running head: The metabolic network of *Synechocystis* sp. PCC 6803

Correspondence:
Ralf Steuer
Humboldt University Berlin
Institute for Theoretical Biology (ITB)
Invalidenstrasse 43,
10115 Berlin, Germany
Phone: +49 (0)30 - 2093 9112
Fax: +49 (0)30 - 2093 8801
ralf.steuer@manchester.ac.uk

Journal Research Area: Systems Biology, Molecular Biology, and Gene Regulation
The metabolic network of *Synechocystis* sp. PCC 6803: Systemic properties of autotrophic growth

Henning Knoop¹, Yvonne Zilliges², Wolfgang Lockau², and Ralf Steuer¹,³

¹Humboldt University Berlin, Institute for Theoretical Biology (ITB), Invalidenstrasse 43, 10115 Berlin, Germany
²Humboldt University Berlin, Institute for Biology, Chausseestrasse 117, 10115 Berlin, Germany
³Manchester Interdisciplinary Biocentre, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, UK
Financial source: The work is funded by BMBF program FORSYS-Partner (Förderkennzeichen: 0315274B).
ABSTRACT

Unicellular cyanobacteria have attracted growing attention as potential host organisms for the production of valuable organic products and provide an ideal model to understand oxygenic photosynthesis and phototrophic metabolism. To obtain insight into the functional properties of phototrophic growth, we present a detailed reconstruction of the primary metabolic network of the autotrophic prokaryote *Synechocystis* sp. PCC 6803. The reconstruction is based on multiple data sources and extensive manual curation, and significantly extends currently available repositories of cyanobacterial metabolism. A systematic functional analysis, utilizing the framework of flux-balance analysis, allows to predict essential metabolic pathways and reactions and allows to identify inconsistencies in the current annotation. As a counterintuitive result, our computational model indicates that photorespiration is beneficial to achieve optimal growth rates. The reconstruction process highlights several obstacles currently encountered in the context of large-scale reconstructions of metabolic networks.
INTRODUCTION

Cyanobacteria are among the evolutionarily oldest organisms and are the only known prokaryotes capable of plant-like oxygenic photosynthesis. As primary producers in aquatic environments, they play an important role in global CO$_2$ assimilation and O$_2$ recycling. Recently, cyanobacteria have also attracted growing attention for economic purposes, including drug discovery and as prolific producers of natural products (Sielaff, 2006; Tan, 2007). In particular their ability to directly convert atmospheric CO$_2$ into biomass and organic compounds, driven by sunlight, offers considerable potential as a novel and renewable resource for bioenergetics (Deng and Coleman, 1999; Mascarelli, 2009; Atsumi et al., 2009; Lindberg et al., 2010).

Among the diverse cyanobacterial strains, _Synechocystis_ sp. PCC 6803 is one of the most extensively studied model organisms for the analysis of photosynthetic processes. With a rich compendium of genomic, biochemical and physiological data available, _Synechocystis_ sp. PCC 6803 therefore offers an ideal starting point to obtain insights into the systemic properties of phototrophic metabolism. The prerequisite of such a systemic description is a detailed reconstruction of the metabolic network of the organism, that is, a reconstruction of the comprehensive set of enzyme-catalyzed reactions required to support cellular growth and maintenance. Once a metabolic reconstruction is available, the vast array of methods developed by computational systems biology over the past decades allows us to dissect the functioning and interplay of possible metabolic routes and biochemical interconversions. In this respect, constraint-based modeling, most notably flux-balance analysis (FBA), has become a quasi-standard in the field. FBA is increasingly utilized to elucidate and characterize large-scale network properties, to direct discovery of novel or alternative pathways, to guide metabolic engineering, as well as for the conceptualization of high-throughput data (Oberhardt et al., 2009; Steuer and Junker, 2009). As one of its prime advantages, constraint-based modeling does not require knowledge of the kinetic parameters of individual metabolic reactions, making it applicable to large-scale, up to genome-scale, metabolic networks. Specifically, FBA allows the prediction of optimal steady-state fluxes that maximize a given objective function, usually the synthesis of biomass or biomass...
precursors required for growth. Though certainly not without pitfalls, the predictions of FBA have proven to reasonably reflect the modes of cellular operation, with manifold applications ranging from microorganisms to algae and plant metabolism (Varma and Palsson, 1994; Shastri and Morgan 2005; Feist et al., 2009; Oberhardt et al., 2009; Grafahrend-Belau et al., 2009; Boyle and Morgan, 2009, Poolman et al., 2009).

The aim of our contribution is to provide a high-quality stoichiometric reconstruction of the primary metabolic network of *Synechocystis* sp. PCC 6803 and a characterization of its functional properties during phototrophic growth. We identify inconsistencies in the current gene annotation and suggest several genes with putative novel or different enzymatic activity for further experimental validation. Our work significantly improves upon earlier reconstructions, which were either restricted to the primary carbon metabolism (Shastri and Morgan 2005; Hong and Lee, 2007; Navarro et al., 2009) or incomplete with respect to several metabolic routes (Fu, 2008; Kun et al., 2008). During the reconstruction process, we identified various obstacles that currently hamper the large-scale construction of metabolic models. First, despite the rather detailed annotation of the genome sequence of *Synechocystis* sp. PCC 6803, reflecting the high level of curation for this model organism, our analysis reveals a number of glaring gaps even within primary metabolic pathways. For several essential metabolic precursors no feasible and validated synthesis routes are known. Second, automated reconstruction and comparison of different resources is significantly compromised by heterogeneous naming conventions used in different databases, making a cross-database comparison a laborious and predominantly manual process. To facilitate model exchange and re-use, as well as to provide a resource for further studies, we therefore place specific emphasis on a consistent and standardized nomenclature. Our reconstruction complies with the minimal information requested for the annotation of biochemical models (MIRIAM) and is represented using the systems biology markup language (SBML), allowing to uniquely identify network components and enabling automated reasoning and model comparison (Le Novère et al., 2003; Hucka et al., 2003).

The computational model is subsequently utilized to explore and characterize possible flux states of *Synechocystis* sp. PCC 6803. Making use of the framework of FBA, we identify essential reaction sets and optimal synthesis routes for growth and maintenance, as well as the necessary re-organization of metabolic fluxes in the
transition from storage-utilizing night metabolism to phototrophic growth. Remarkably, our results indicate that the seemingly wasteful oxygenation of ribulose-1,5-bisphosphate (RuBP) is required to achieve an optimal flux state. This unintuitive result is discussed in the light of ongoing work to elucidate and understand photorespiratory metabolism (Eisenhut et al., 2006, 2008).

RESULTS

A metabolic model of *Synechocystis* sp. PCC 6803

We aim at a comprehensive reconstruction of the primary metabolic pathways of *Synechocystis* sp. PCC 6803. The scope of our reconstruction encompasses the main metabolic pathways, including photosynthetic light reactions, the Calvin cycle, glycolysis, the (incomplete) TCA cycle, the pentose-phosphate pathway, oxidative phosphorylation, photorespiration, amino-acid synthesis, as well as supply of nucleotides, lipid precursors, and cofactors. Our starting point is the annotated genome sequence of *Synechocystis* sp. PCC 6803 as obtained from the Cyanobase website (Nakamura et al., 1998; Nakao et al., 2010).

Following earlier strategies (Feist et al., 2009; Oberhardt et al., 2009), the reconstruction process comprises four steps: First, the annotated genome, together with pathway repositories such as the Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.kegg.com/), provide an initial draft network of the organism. Second, gaps and inconsistencies within the draft network are identified and the network is manually curated based on the available biochemical literature. Third, the set of reactions is converted into a *mathematical model*, using consistent annotations as well as established exchange formats, such as SBML. The conversion into a computational model entails the addition of several pseudo-reactions, such as ATP utilization for cellular maintenance or biomass formation, which do not correspond to genuine enzymatic reactions but reflect overall cellular processes. Finally, the model is evaluated using constraint-based analysis and other methods of computational systems biology. In particular, predictions about optimal flux states are validated using available phenotypic data. Details of the reconstruction process are given in Materials and Methods.
Our reconstruction results in a metabolic network of 380 reactions and 291 metabolic compounds. The set of reactions consists of 360 enzymatic reactions, including about 38 isoreactions, 4 spontaneous uncatalyzed reactions, 5 diffusion and transport reactions, as well as 11 auxiliary reactions to account for light absorption, photophosphorylation, biomass formation and cellular maintenance. All processes are assumed to take place in a single compartment, with the exception of transport reactions. The reconstruction is focused on a functional description of phototrophic growth and utilization of storage compounds in the absence of light. Uptake and utilization of organic molecules is not considered. All enzymatic reactions annotated for *Synechocystis* sp. PCC 6803 in KEGG and other repositories but not included in the core reconstruction are provided as a separate file. An overview of the model is shown in Fig. 1.

**Objectives for growth and maintenance**

During the reconstruction process, the initial draft model obtained from the annotated genome sequence was iteratively refined using extensive manual curation. Our primary objective was to create a metabolically functional model, capable to support cellular maintenance and growth, that is, a model that is capable to generate all relevant biomass components. In general, the biomass objective function (BOF) -- an equation that specifies all relevant biomass precursors in their appropriate molar ratios -- is difficult to obtain experimentally. An estimation of the macromolecular biomass composition for *Synechocystis* sp. PCC 6803 under different growth conditions was given by Shastri and Morgan (2007) and afterwards utilized in several studies. However, as the size of our reconstruction is considerably larger, the inclusion of additional components was necessary. To this end, we modified and extended the BOF given by Shastri and Morgan, based on information on the macromolecular composition reported for the cyanobacterium *Arthrospira (Spirulina) platensis* (Cogne et al., 2003). The BOF is given in Table 1 and detailed in Materials and Methods.

Knowledge of the BOF allows us to systematically check for available synthesis routes of biomass components. We performed iterative gap-filling, based on multiple pathway repositories, such as KEGG (http://www.kegg.com/), BioCyc (http://www.biocyc.com/),...
BRENDA (http://www.brenda-enzymes.org/), as well as curation of primary biochemical literature and identification of putative enzymes via BLAST search. As expected, the initial draft network did not support synthesis of all required intermediates. In particular, no complete synthesis routes for the amino acids glycine, serine, cysteine, methionine, asparagine, and histidine were annotated. In these cases, we proceeded according to the following scheme: First, we aimed to identify the missing steps to complete the synthesis pathways utilizing information from other cyanobacteria. For example, the enzyme for a methylenetetrahydrofolate reductase (EC 1.5.1.20) was not annotated for *Synechocystis*, but is annotated for *Cyanothece* sp. PCC 8801, which enables the identification of a candidate gene (*slr2141*) for this enzyme in *Synechocystis* sp. PCC 6803. Second, if no satisfactory results could be obtained by direct comparison, we searched the primary literature for a description of the respective pathways, starting with cyanobacterial metabolism and then broadening the search to plants (mainly *Arabidopsis thaliana*) and finally across all kingdoms. Our strategy resulted in complete synthesis routes for all amino acids, except asparagine and methionine. For the latter, the synthesis steps from *Microcystis aeruginosa* (EC 2.3.1.31, EC 2.5.1.49) were adopted to allow for a metabolically functional network.

Similar, the scheme described above was applied to test and complete the pathways for *de novo* synthesis of cofactors, including Coenzyme A (CoA), nicotinamide adenine dinucleotides (NAD/NADH), nicotinamide adenine dinucleotide phosphates (NADP/NADPH), flavin adenine dinucleotides (FAD/FADH), and tetrahydrofolate. Again, the initial draft network had to be supplemented with additional reactions. A complete list is provided in the Materials and Methods section and as Supplementary Information. For three reactions no reasonable candidate genes could be identified, namely UDP-glucose phosphorylase (EC 2.7.7.9), aminodeoxychorismate synthase (EC 2.6.1.85), and hydrolase of 5-amino-6-(5'-phosphoribitylamino)uracil (EC 3.1.3.-), although each step is necessary to allow for the synthesis of biomass precursors. Finally, it was verified that the model is capable to replenish each intracellular metabolite upon dilution by growth.

Main resources during the reconstruction process were the Cyanobase website (Nakamura et al., 1998), KEGG (Kanehisa and Goto, 2000), Meta- and BioCyc (Caspi et al., 2010), BRENDA (Chang et al., 2009), the CyanoMutants website (Nakamura et al., 1999), and to a lesser extend Uniprot (Uniprot Consortium, 2010). The final network is
provided as an MIRIAM-compliant SBML file (Supplementary File 1), and spreadsheet (Supplementary File 2).

**Functional properties of phototrophic growth**

Based on the reconstructed network, the functional properties of phototrophic growth are evaluated using flux-balance analysis (FBA). Specifically, the rate of change of the vector of metabolite concentrations $\mathbf{S}$ can be described by a differential mass-balance equation of the form

$$\frac{d\mathbf{S}}{dt} = \mathbf{N} \cdot \mathbf{v}$$

where $\mathbf{N}$ denotes the stoichiometric matrix and $\mathbf{v}$ the vector of reaction fluxes. Using the steady-state hypothesis for the vector of metabolite concentrations, any *stationary* flux distribution $\mathbf{v}^0$ has to satisfy the steady-state constraint $\mathbf{N} \mathbf{v}^0 = \mathbf{0}$. In addition to the constraint imposed by the mass-balance equation, FBA postulates that the distribution of fluxes is organized such that a given objective function, usually the synthesis of precursors required for cellular growth, is maximized. Mathematically, this assumption translates into a linear optimization problem for the stationary flux vector $\mathbf{v}^0$, such that

$$\text{BOF}[\mathbf{v}^0] \rightarrow \text{maximal}$$

under the conditions

$$\mathbf{N} \cdot \mathbf{v}^0 = \mathbf{0}$$

and

$$\mathbf{v}^\text{min}_i \leq \mathbf{v}^0_i \leq \mathbf{v}^\text{max}_i$$

for all metabolic fluxes $\mathbf{v}^0_i$ ($i=1,\ldots,m$), where $\mathbf{v}^\text{min}_i$ and $\mathbf{v}^\text{max}_i$ denote optional upper and lower bounds for each stationary flux value. The constraint optimization problem can be solved using standard techniques of linear programming (Klamt et al., 2007; Becker et al., 2007). The solution to the optimization problem results in a particular flux distribution that achieves a maximal growth yield, given predefined external conditions and
exchange fluxes. We note that, owing to redundancy and flexibility within the network, the optimal solution is usually not unique.

Once the constraint optimization problem is specified, FBA allows for a functional interrogation of the metabolic network with respect to (i) its ability to synthesize all metabolic precursors at their observed molar ratios, (ii) essential reactions and prediction of gene deletion studies, (iii) optimal routes for maximal growth, (iv) flux variability analysis, as well as (v) qualitative reorganization of metabolic routes under different growth conditions.

In the following, we are mainly interested in phototrophic growth under conditions of either limiting light or limiting CO₂ availability. Within the network, inorganic carbon is taken up from the environment as bicarbonate HCO₃⁻ by an active (ABC) transport (Omata et al., 1999). The intensity of light input is represented by flux of absorbed photons and translocation of H⁺ through the membrane, partially coupled to regeneration of NADPH. Photophosphorylation and the equations for the respiratory chain are adopted from Shastri and Morgan (2005). Figure 2 shows the predicted maximal growth rate as a function of input fluxes. Maximal growth is linearly dependent on either absorbed light or net CO₂ availability, with an offset for small light intensities. In addition to the biomass function, we assume a basal ATP demand of 0.27 mmol gDW⁻¹ h⁻¹ for cellular maintenance, see Materials and Methods.

In the following, all simulations are performed at a specified light intensity, whereas the model is allowed to adjust HCO₃⁻ uptake as necessary for optimal growth yield. In this way, we circumvent the necessity to utilize a two-step optimization strategy, as compared to Shastri and Morgan (2005). Likewise, we do not make use of auxiliary optimization criteria, such minimal total flux, to forestall variability within optimal flux distributions. Instead, the variability in the optimal solutions is described explicitly where appropriate. The constraint optimization problem usually results in relative flux values. Typical solutions can be compared with observed growth rates. For example, van Liere and Walsby (1982) report a typical carbon uptake of 0-13 μg C gDW⁻¹ h⁻¹, which translates into a flux of approximately up to 1.1 mmol CO₂ gDW⁻¹ h⁻¹. If we assume an uptake of 0.22 mmol CO₂ gDW⁻¹ h⁻¹, we obtain a growth rate of about 0.041 h⁻¹, well within the range of values observed for Synechocystis under many conditions. The photosynthetic quotient, the ratio of moles of O₂ released per mole of CO₂ fixed, is
approximately 1.43. Overall, all quantitative results are in good agreement with values obtained from previous models (Shastri and Morgan, 2005; Fu, 2008), with exceptions mainly due to the different network stoichiometry.

A predicted optimal flux distribution of central metabolism consistent with maximal growth is shown in Fig. 3. Flux values are reported relative to RuBisCO activity (100%). The solution is not unique, but considerable variability is observed in the utilization of the transaldolase (talB, slr1793, EC 2.2.1.2), the fructose-1,6-bisphosphate aldolases (cbbA and fda, slr0018 and slr0943, EC 4.1.2.13) and the fructose 1,6-bisphosphatase II/sedoheptulose 1,7-bisphosphatase (fbpl and glpX, slr2094 and slr0952, EC 3.1.3.11 and EC 3.1.3.37), corresponding to alternative optimal routes within the network (Poolman et al., 2003).

Common to all solutions is that the highest (absolute) flux is allocated to the glyceraldehyde-3-phosphate dehydrogenase (gap2, slr1342, EC 1.2.1.12). The reaction is a key reaction of the Calvin-Benson cycle and essential for growth. The corresponding mutant is not viable under phototrophic conditions (Koksharova et al., 1998). Likewise common to all solutions is a dominance of ribulose-phosphate 3-epimerase (cfxE, slr0807, EC 5.1.3.1) over the ribose-5-phosphate isomerase A (rpiA, slr0194, EC 5.3.1.6) with more than twice the flux allocated for optimal growth. Both reactions are essential within the network. At the level of glycerate-3-P (PG3) the flux is split towards biosynthesis and regeneration of the Calvin-Benson cycle, with a ratio of approximately 1:8. The optimal ratio is significantly higher than the textbook ratio of 1:5 since precursors for biosynthesis are withdrawn during regeneration, such as G6P for synthesis of glycogen and nucleotide sugars and ribose-5-P (R5P) for synthesis of folate, purines and histidine.

Our model includes photorespiration, which competitively inhibits the photosynthetic carboxylation of ribulose-1,5-bis-P (RuBP) by RuBisCO. Interestingly, the predicted optimal flux pattern allocates approximately 4-5 % of the flux through RuBisCO to the allegedly disadvantageous oxygenation reaction, thereby producing glyoxylate as an essential precursor for glycine, serine and cysteine formation. For the enzymes associated with the TCA cycle, the optimal flux distribution reflects the utilization of an incomplete cycle present in Synechocystis sp. PCC 6803. In particular, fumarate is mainly synthesized as a byproduct of arginine and nucleotide biosynthesis, and
channeled back to oxaloacetate for reuse in the TCA cycle and to serve as a precursor for the synthesis of diverse amino acids. The distribution shown in Fig. 3 is characteristic for photoautotrophic growth. We observed no qualitative changes with respect to different light intensities or different CO₂ availability.

**Gene deletion studies and metabolic robustness**

The validity of any computational model can be examined using phenotypic data and information on metabolic knockouts. In particular, one of the prime accomplishments of FBA is its ability to predict the viability or the expected (relative) decrease in maximal growth rate upon deletion of one or more genes. In this respect, we can assess the essentiality of each gene based on two criteria: The first requirement is the persistent ability of the network to produce all precursors required for biomass formation, as specified by the BOF, even in the absence of a specific gene. Second is the requirement to regenerate all metabolites and cofactors at a specified rate in the absence of the respective gene. Although usually only the first requirement is considered in the literature, both criteria are essential for viability and persistent growth. From an algorithmic point of view, the essentiality of each gene is evaluated by removing its corresponding reactions from the network and re-evaluating the constraint optimization problem for maximal biomass production. Genes are classified as essential if removal of the respective reactions prohibits the formation of biomass. Non-essential genes are further classified according to whether the maximal growth rate is reduced or remains unaltered. Within our network, 126 (of 337) genes are classified as essential for biomass formation, the removal of 26 genes results in a decreased maximal growth yield, whereas the removal of the remaining 185 genes has no effect on the maximal growth yield. The latter set includes a set of 38 isoenzymes. Replacing biomass formation by the ability to regenerate all metabolites results in 177 essential genes, whereas 160 genes are classified as non-essential, including 31 genes that result in a less efficient regeneration of metabolites and cofactors. A comprehensive list for essential genes is provided as Supplementary Information.

In total, 51 entries in CyanoMutants corresponding to mutational inactivation of distinct genes are directly related to our network reconstruction. Of these, 39 genes encode
non-essential genes within the network, of which 35 can be confirmed by a successful knock-out reported in CyanoMutants. The discrepant 4 genes correspond to cysteine desulfurase (slr0077), the NADH dehydrogenase type I subunit 4L (slr0522), a polyphosphate kinase (slr0290), and a pyrophosphatase (slr1622). The latter two reactions are easily compensated *in silico* by pairs of reactions acting together in a cycle - a solution that might not be feasible *in vivo* under the conditions tested. Essential genes are more difficult to reconcile, as our network does not account for possible roles in regulatory processes or toxicity and other adverse effects upon possible accumulation of metabolites. Of the 51 entries from CyanoMutants, 12 genes are classified as essential within our model, but only for 8 of those the construction of a fully-segregated mutant failed. The inactivation of the remaining 4 genes each resulted in a viable genotype. A more detailed comparison, including a summary of discrepant findings, is given in the Materials and Methods.

**From storage utilization to phototrophic growth**

In the absence of light, *Synechocystis* is able to temporarily utilize various storage compounds, such as glycogen and to a lesser extend cyanophycin and poly-beta-hydroxybutyrate (PHB), to provide energy and precursors for cellular maintenance. In this respect, our *in silico* model allows to simulate the necessary reorganization of metabolism during the transition from glycogen-utilizing night metabolism to phototrophic growth. Figure 4 shows the metabolic transitions on a two-dimensional plane with the flux of selected reactions color-coded as a function of absorbed photons (x-axis) and glycogen utilization (y-axis). Within each subplot, the transition from night metabolism (glycogen utilization) to phototrophic growth corresponds to a path from the upper left to the lower right corner. Along such a path, we observe several characteristic transitions in the re-organization of metabolic flux.

Under conditions of storage-utilization, corresponding to a point on the y-axis in Fig. 4, we assume glycogen as the only resource for cellular maintenance. In this case, the flux distribution is characterized by O₂ uptake and absence of RuBisCO-dependent carboxylation or oxygenation. Instead we observe flux through the TCA cycle utilizing a bypass via the succinate-semialdehyde dehydrogenase (slr0370, EC 1.2.1.16) to
circumvent the 2-oxoglutarate dehydrogenase not present in *Synechocystis* sp. PCC 6803. With increasing light intensity and concomitant decrease in glycogen mobilization, flux through the TCA cycle decreases. Instead, the photosynthetic light reactions start to generate the energy (ATP) requirements for cellular maintenance (fixed in the simulation at 0.27 mmol gDW\(^{-1}\) h\(^{-1}\)), but no CO\(_2\) fixation takes place yet. We note that the observed *in silico* flux distribution during mixotrophic growth agrees with experimental results from Howitt and Vermaas (1998), who did not observe any effect on mixotrophic and phototrophic growth by the deletion of two oxidases.

Only for a light intensity above a certain threshold, flux through the Calvin-Benson cycle is initiated, resulting in an increasing rate of RuBisCO-dependent carboxylation and oxygenation. Remarkably, oxygenation of ribulose-1,5-bisphosphate (RuBP) occurs slightly before the start of CO\(_2\) fixation, even though within the *in silico* optimization a complete suppression of photorespiration is possible - if favorable for maximal ATP production or biomass formation. The occurrence of a non-zero photorespiration in the constraint-optimization problem points to a beneficial role of the seemingly wasteful side reaction as a supply of precursors for glycine formation. In particular, during heterotrophic growth, glycine is formed as a result of proline degradation, which ceases in the presence of photorespiration. Consequently, the *in silico* model points to an ambivalent interpretation of the role of photorespiration during phototrophic growth, which is discussed in more detail below. We note that all results are contingent on the reconstructed network structure and the specified exchange (glycogen and light) fluxes imposed on the constraint optimization problem, and may therefore not necessarily replicate the *in vivo* reality. Nonetheless, the *in silico* simulation highlights the potential of cellular metabolism to adapt to changes in environmental conditions in complex, surprisingly subtle and sometimes non-intuitive ways.

**Photorespiration revisited**

One of the striking results of the *in silico* simulation is the occurrence of a non-zero rate of photorespiration during photoautotrophic growth. In *Synechocystis* sp. PCC 6803, like in most other photosynthetic organisms, the fixation of inorganic carbon is catalyzed by the bifunctional enzyme RuBisCO (Ribulose-1,5-bisphosphate carboxylase/oxygenase),
with the carboxylation of ribulose-1,5-bisphosphate (RuBP) yielding two molecules of 3-phosphoglycerate (PG3). However, carboxylation competes with RuBP oxygenation, a process that generates one molecule of PG3 and one molecule of the two carbon compound 2-phosphoglycolate (2PG). Photorespiration, the photosynthetic oxygenation of RuBP, is often described as one of the most wasteful processes on earth and is widely considered to be mainly an evolutionary relic - tracing back to the evolution of the enzyme in a low-O2 atmosphere.

In cyanobacteria the role and relevance of photorespiration is still only partially understood. Cyanobacteria have evolved dedicated CO2-concentration mechanisms (CCM) that significantly repress the oxygenase activity of RuBisCO and vastly improve photosynthetic performance (Badger and Price, 2002). However, the presence of CCM does not fully eradicate RuBP oxygenation, and Eisenhut et al. (2008) were recently able to show that an active photorespiratory 2PG metabolism is essential for cyanobacterial growth in the present O2 containing atmosphere. In particular, a non-negligible rate of photorespiratory 2PG generation necessitates quick conversion to glyoxylate to avoid toxification and inhibition of the Calvin-Benson cycle.

In this respect, our metabolic reconstruction allows to augment the current discussion on photorespiration from a network perspective. As shown above, the constraint optimization problem assigns a nonzero flux to the oxygenation of RuBP under conditions of photosynthetic growth, suggesting a non-trivial role of photorespiration to achieve a maximal rate of biomass formation. Indeed, an intermediate of photorespiration, glyoxylate, plays an important role in cyanobacterial metabolism and can serve as a precursor for glycine synthesis. More specifically, within our reconstruction, the biosynthesis of the amino acids glycine, serine and cysteine exclusively depends on the allocation of glyoxylate. Although it was suggested that cyanobacteria are capable of synthesizing serine directly from 3-phosphoglycerate via 3-phosphohydroxy pyruvate and phosphoserine (Colman and Norman, 1997), no candidate genes for such a conversion are currently known for Synechocystis. The assumption that glyoxylate is an essential precursor for serine synthesis is further supported by the fact that Eisenhut et al. (2006) were unable to obtain a completely segregated mutant with an inactivated serine hydroxymethyltransferase (SHMT, sll1931). The SHMT catalyzes interconversion of serine and glycine and would be
dispensable if an alternate synthesis pathway for serine exists. Correspondingly, our simulations suggest that during phototrophic growth the photorespiratory 2PG metabolism provides the required flux towards glyoxylate to allow for the formation of metabolic precursors and biomass. The computationally allocated rate of photorespiration is approximately 4.2% of the total RuBisCO activity, a value that is in good agreement with the 4-5% photorespiration assumed for *Synechocystis* [M. Hagemann, personal communication]. Prohibiting photorespiratory activity *in silico* results in a slightly decreased predicted maximal growth rate of 94% of the original value.

Consequently, in the context of our network reconstruction, photorespiration turns out as the most efficient way to allocate glyoxylate during phototrophic growth. However, this superior efficiency does not hold for non-phototrophic (dark) metabolism, as already observed in Fig. 4. In this case, the (rather low) demand for glyoxylate is met by degradation of proline, in a series of five enzymatic steps. Despite its slightly lower carbon efficiency, the synthesis of proline and subsequent conversion to glyoxylate yields additional units of NADH and ATP, as compared to photorespiration, making the pathway preferable in heterotrophic conditions when optimizing for total biomass formation. We emphasize that this difference highlights the fact that a criterion like metabolic efficiency is usually not a property of individual pathways, but must be interpreted in the context of the metabolic status and the demands of the remaining network.

Our computational results allow us to formulate two different scenarios: First, the beneficial role of photorespiration may be merely an artefact of an incomplete network topology. In particular, our results depend on the absence of alternative pathways for the formation of glycine and serine, most importantly the direct conversion of 3-phosphoglycerate to serine, via 3-P hydroxypyruvate and phosphoserine. Although no candidate genes are currently known, the pathway may be active in *Synechocystis* sp. PCC 6803. In this case, photorespiration would be completely suppressed in both, phototrophic and heterotrophic, conditions when optimizing for maximal growth yield (data not shown). Alternatively, the second scenario implies that the current annotation of the *Synechocystis* genome, augmented by our network reconstruction, does provide a correct picture of glycine metabolism. In this case, the metabolism of *Synechocystis* is
adapted to make use of the inevitable supply of the side product 2-phosphoglycolate under photoautotrophic conditions. In particular, there is no requirement for the organism to establish or maintain alternative and possibly more efficient pathways for glycine and serine formation. We note that this view does not contradict earlier findings on gene essentiality and growth in a high CO₂-atmosphere. Within our simulation photorespiration is not essential and in the absence of photorespiration all essential precursors can still be synthesized, albeit with slightly lower efficiency. In the context of our study, it is not possible to conclusively resolve which of the two scenarios mirrors the \textit{in vivo} reality. Nonetheless, we emphasize that only a thorough reconstruction of cyanobacterial metabolism, to the best of one's current knowledge, allows us to formulate and probe alternative hypotheses on the efficiency and optimality of reactions.

DISCUSSION AND PERSPECTIVES

Elucidation of the structure and function of metabolic networks is increasingly assisted by computational approaches that allow for a systemic view on possible interconversion routes. A prerequisite for such an analysis is the comprehensive reconstruction of the set of enzymatic reactions taking place in a single cell or organisms. Here, we have presented a manually curated reconstruction of the metabolic network of \textit{Synechocystis} sp. PCC 6803, an important unicellular model organism for phototrophic metabolism. Network reconstruction is usually based on the annotated genome sequence, in combination with various reaction and pathway repositories. However, most databases and repositories primarily focus on the 'parts lists' of enzymatic reactions, that is, they do not consider the resulting network as a functional entity. Conversely, network reconstruction specifically aims to take into account the multiple requirements that distinguish a list of reactions from a functioning network. Such requirements include the ability of the network to synthesize all essential intermediates and precursors for biomass production, to allow for energy (ATP) generation for cellular maintenance, to maintain the redox-balance, as well as to be capable to regenerate metabolites and cofactors upon dilution by growth. Failure to fulfill one or more of these requirements usually points to errors or gaps in the network stoichiometry and necessitates to revise the reconstructed network accordingly. In this sense, network reconstruction is a pre-
stage to metabolic modeling, and recognizes that cellular metabolism operates as a highly integrated network. Despite the great achievements over the last decade, resulting in more than a dozen reconstructions for several organisms, the systematic reconstruction of genome-scale metabolic networks is still in its infancy. While already considerable literature exists on algorithms for automated gap-filling, these algorithms are usually exemplified only on highly domesticated examples, such as recovery of a previously known complete network upon artificial deletion of individual reactions. As yet, high-quality reconstructions are therefore predominantly based on extensive manual curation utilizing information from primary biochemical literature. A persistent problem are also the diverse naming conventions used in different databases and existing reconstructions, as well as the sometimes inconsistent or erroneous annotation of enzymes, often only based on sequence comparison. As pointed out previously (Stitt et al., 2010), a great challenge for the future is the development and enforcement of publishing standards and naming conventions. Other problems are rooted in the interspecific and intraspecific variation and are less straightforward to overcome. For example, only recently it was shown than even within such a basic and exhaustively analyzed pathway as central carbon catabolism in *Escherichia coli*, previously unreported reactions occur (Nakahigashi et al., 2009). In this respect, we do not expect our reconstruction to be a faultless replicate of the *in vivo* reality, rather the purpose is to start and foster a discussion about possible inconsistencies and shortcomings of our current understanding of cyanobacterial metabolism. In particular, a high-quality network reconstruction, even if erroneous or incomplete in parts, offers tremendous possibilities to probe and understand the functioning of metabolism.

Most current reconstructions are evaluated using flux-balance analysis, a constraint-based method that allows to predict and evaluate optimal flux patterns with respect to a given objective function. In our case, the application of FBA led to several conclusions about (i) the typical optimal reaction routes during phototrophic growth, (ii) essential genes and synthesis pathways of photosynthetic metabolism as well as (iii) the necessary re-organization of metabolic flux during a transition from heterotrophic (glycogen-based) to phototrophic growth. The latter highlights the interconnectedness of cellular pathways as a system of supply and demand. We observe several characteristic
switches that emphasize that the efficiency of cellular reaction pathways is not an intrinsic property of the respective pathways, but depends on the current status and metabolic demand of the entire metabolism. The most prominent example for such a complex interdependency is the allocation of nonzero flux to the oxygenation reaction of RuBP. While for low light intensities the demand for glycine is met by alternative pathways, such as utilization of proline, the RuBisCO oxygenase reaction becomes more efficient with increasing light intensity and finally provides all glyoxylate required for cellular growth. This transition is not an effect of the efficiency of the pathway per se -- but rather reflects differences in the 'sales value' of cellular commodities under different growth conditions. We note that the fact that the seemingly wasteful side reaction of RuBisCO might be used in a constructive way is in good agreement with the known trade-off between specificity and overall turnover rate of RuBisCO (Tcherkez et al., 2006). The hypothesis resulting from our analysis might also explain the remarkably low specificity and a high turnover rates observed in some cyanobacteria (Witzel et al., 2010).

Similar consideration with respect to the beneficial role of seemingly wasteful side reactions might also hold for other organisms and tissues. For example, in a recent computational study Poolman et al. (2009) reported a nonzero RuBisCO oxygenase reaction as a result of a constraint optimization problem in heterotrophic Arabidopsis cells. While the scenario discussed therein is probably unlikely to occur in vivo, as also acknowledged by the authors (Poolman et al. 2009; Stitt et al., 2010), the prediction of flux through previously unrecognized reactions as a result of network optimization is not unusual (Schwender et al., 2004). In this sense, a thorough network reconstruction allows to evaluate the inevitable functional consequences, given the current annotation and knowledge of enzymatic interconversions, and thus opens the possibility to specifically design experiments to distinguish between alternate hypotheses.
MATERIALS AND METHODS

Abbreviations and nomenclature

2PG (2-Phosphoglycolate), ABC (ATP-binding cassette), AcCoA (Acetyl coenzyme A), BOF (Biomass objective function), CoA (Coenzyme A), Cytb6f (Cytochrome b6f complex), DHAP (Dihydroxyacetone phosphate), DW (dry weight), E4P (D-Erythrose 4-phosphate), F6P (D-Fructose 6-phosphate), FBA (Flux balance analysis), FBP (D-Fructose 1,6-bisphosphate), FNR (Ferredoxin-NADP+ reductase), G1P (D-Glucose 1-phosphate), PG2 (D-Glycerate 2-phosphate), PG3 (D-Glycerate 3-phosphate), G6P (D-Glucose 6-phosphate), GAP (D-Glyceraldehyde 3-phosphate), MTHF (5,10-Methylenetetrahydrofolate), PHB (Poly-3-hydroxybutyrate), PYR (Pyruvate), R5P (D-Ribose 5-phosphate), Ru5P (D-Ribulose-5-phosphat), RuBisCO (D-Ribulose 1,5-bisphosphate carboxylase/oxygenase), RuBP (D-Ribulose 1,5-bisphosphate), S7P (D-Sedoheptulose 7-phosphate), SBP (D-Sedoheptulose 1,7-bisphosphate), THF (Tetrahydrofolate), X5P (D-Xylulose 5-phosphate).

Metabolic network reconstruction

Network reconstruction was performed as outlined in the main text. Starting point of the reconstruction is the cyanobacteria genome database cyanobase (http://genome.kazusa.or.jp/cyanobase) containing the annotated nucleotide sequence of the genome of Synechocystis sp. PCC 6803 (Nakamura, 1998; Nakao, 2009). From the gene identifiers an initial list of enzymes and their associated reactions was assembled using the KEGG database (Kanehisa and Goto, 2000). Subsequently, the network was subjected to iterative gap-filling. Incomplete pathways were found for the biosynthesis of glycine, histidine, methionine and asparagine. To identify incomplete pathways, it was first ascertained whether the respective product can be produced under phototrophic conditions. Second, if no production was possible, it was established which set of reactions achieves production of this product in other organisms (other cyanobacterial species, and then broadening the search to plants). If applicable, candidate genes were identified in the Synechocystis genome using BLAST (http://blast.kazusa.or.jp/blast_search/cyanobase/genes). A list of all newly annotated...
candidate genes is provided as Supplemental Data 3, along with details of the comparison and the respective sequence identities and E-values.

With respect to glyoxylate metabolism, we adopted the pathway suggested by Eisenhut et al. (2008). In particular, glyoxylate that is synthesized during photorespiration serves as a precursor for the amino acids glycine, serine and cysteine. The substrate specificity for the transaminase (EC 2.6.1.-) coded by gene sll1559 remains unclear. Three scenarios are possible (Liepman et al., 2001): The gene can code for a serine-glyoxylate aminotransferase (EC 2.6.1.45), for a serine-pyruvate aminotransferase (EC 2.6.1.51) or for an alanine-glyoxylate aminotransferase (EC 2.6.1.44), or any combination thereof. Again, the effects on the results of FBA are negligible, as there is no difference in the overall cofactor utilization for these reactions. Therefore only the reaction for a serine-glyoxylate aminotransferase was included in the network. Despite extensive literature research, several open questions and incomplete pathways remained. It was not possible to identify an enzyme for the amination of aspartate to asparagine and no confirmed synthesis route for the amino acid methionine could be found. As methionine is an essential part of the biomass objective function, we implemented a pathway via O-acetyl-homoserine. This pathway is postulated to exist in Microcystis aeruginosa NIES-843, which is the nearest relative of Synechocystis sp. PCC 6803 according to 16SrDNA analysis (NIES-Collection, Japan), but no candidate gene can be identified in Synechocystis sp. PCC 6803.

The reconstruction process continued with the identification of biosynthesis pathways for all compounds of the BOF, including the synthesis of purine and pyrimidine nucleotides, nucleotide sugars and precursors of the lipid components of the biomass. Subsequently, the de novo biosynthesis routes of cofactors, namely NAD(P), CoA, glutathione, FAD and tetrahydrofolate was assessed. Transport mechanisms for the uptake of extracellular sulfate, hydrogen carbonate and nitrate were added as well as for an exchange of oxygen and carbon dioxide, respectively. We assumed nitrate as the sole nitrogen source, uptake of urea or ammonia was not considered. Carbon uptake was limited to a single transport reaction. The reactions of photophosphorylation and the respiratory chain were adapted from Shastri and Morgan (2004) and Hong and Lee (2007). Finally, a catalase (sll1987) and a thioredoxin-disulfide reductase (sll0600) were added.
An persistent problem in network reconstruction is that for many enzymes the electron acceptors and donors are unknown (see for example the proline dehydrogenase EC 1.5.99.8 or the glycolate dehydrogenase EC 1.1.99.14). In this case it was assumed that NADH acts as the final electron donor and NAD as the final acceptor, respectively. Similar, for many enzymes the cofactor specificity with respect to NAD(H) and NADP(H) is unclear. In this case, the respective enzyme was assumed to catalyze both reactions. This choice has only marginal consequences on the results of FBA, since the network includes a NAD/NADP transhydrogenase (slr1239, slr1434), which allows a direct conversion between NADH and NADPH and vice versa. We note that FBA does not straightforwardly allow to include quantitative differences in enzyme specificity but is primarily focussed on stoichiometric information.

The reconstructed network is provided as Supplemental Data 1 (SBML) and 2 (MS-Excel). A complete list of reactions added during the reconstruction process is provided as Supplemental Data 3. Supplemental Data 6 provides a list of additional enzymes associated with Synechocystis sp. PCC 6803 metabolism avaible from KEGG, but not included in the current model. The final network model is compliant with MIRIAM standards (Le Novère N et al., 2005) and contains links to CHEBI nomenclature (Degtyarenko K et al., 2008), a standard endorsed within recent publications (Herrgård et al., 2008). For appropriate tools to visualize and interrogate the network, see for example the software guide on www.sbml.org.

The Biomass Objective Function (BOF)

FBA requires the definition of a biomass objective function that specifies all cellular components in their appropriate molar ratios. A first BOF was provided by Shastri and Morgan (2004) and also used by Fu (2008). However, the respective BOF was restricted mainly to precursors for amino acid compositions as constituents of the biomass. A considerably more elaborate function was given by Cogne et al. (2003) for a metabolic model of Arthrospira platensis, a blue-green photolithoautotroph cyanobacterium. The function is composed of five classes of macromolecules, namely carbohydrates, proteins, lipids, RNA and DNA and the pigment chlorophyll. We utilized a modified version of the BOF of Cogne et al. (2003). Specifically, we neglect the flux to chlorophyll,
corresponding to only 0.16% of total biomass. Additionally, the lipid demands were simplified to lipid precursors only and the formation of carbohydrates was slightly simplified. The resulting BOF is given in Table 1.

**Reconciling existing models**

Several previous reconstructions of *Synechocystis* sp. PCC 6803 are available. To our knowledge, the first reconstruction was provided by Shastri and Morgan (2005) and subsequently extended by Hong and Lee (2007). Both reconstructions are restricted to the primary metabolism, including glycolysis, the TCA cycle, and the Calvin-Benson cycle. The model of Shastri and Morgan encompasses 93 metabolic reactions (with reversible reactions counted twice), including 23 lumped reactions for the formation of amino acids and nucleotides.

A considerably larger genome-scale model was recently published by Fu (2008), encompassing about 831 metabolic reactions. However, the biomass objective function was taken over from Shastri and Morgan, therefore the additional reactions cannot assume any functional role. Rather surprisingly, all previous reconstructions include a bacterial-like glyoxylate shunt, which was apparently suggested by Yang et al. (2002). We find no evidence for this assertion and assume the inclusion of the glyoxylate shunt in previous reconstructions was mainly motivated by the omission of a C2-Cycle.

**Flux-balance analysis**

Flux-balance analysis was based on available software tools and algorithms. All optimization studies were computed using the COBRA toolbox v. 1.3.1 (Becker et al., 2007) (available on http://gcrg.ucsd.edu/Downloads/Cobra_Toolbox), as well CellNetAnalyzer v.9.2 (Klamt et al., 2007) (available on http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html). The toolboxes were run under the MATLAB 7.6.0 environment, including the MATLAB optimization toolbox as a Linear Program Solver. For the COBRA toolbox a custom SBML file had to be generated (provided as Supplemental Data 7). See also (Oberhardt et al., 2009) and references therein for introductory tutorials on FBA.
For the constraint-optimization exchange fluxes only the light input was pre-assigned, whereas nitrate, sulfate and carbon (as hydrogen carbonate) uptake was allowed to adopt as necessary. In this way we avoid the two-step optimization problem employed by Shastri and Morgan (2005) who first optimize for biomass formation using a fixed carbon uptake and free light flux, followed by a second optimization for minimal light flux, given the previously obtained biomass formation as an additional constraint. During our simulations, the uptake of hydrogen carbonate during phototrophic growth was constrained between 0 and 0.45 mmol HCO$_3^-$ gDW$^{-1}$ h$^{-1}$, which is related to a maximum growth rate of 0.09h$^{-1}$. Limitation of growth by sulfate or nitrate was not considered. In addition to the BOF, we assumed a basal ATP demand (maintenance) for each FBA solution. As no reliable experimental data were available, the value was arbitrarily set to 10% of the maximal photosynthetic ATP generation used in Fig. 2, corresponding to a value of 0.27 mmol gDW$^{-1}$ h$^{-1}$. We note that the basal ATP demand, contrary to the situation in heterotrophic organisms, has no influence on the FBA solution. Likewise, the FBA optimization problem does not take dilution of metabolites, including cofactors, into account.

**Comparision of in silico knock out mutants and in in vivo measurements**

A crucial test for network reconstruction and FBA is the ability to correctly predict lethal gene knockouts. In our case, as described in the main text, we observe a significant overlap between *in silico* mutants and *in vivo* experiments. Several of the *in silico* predicted lethal genotypes can be confirmed by published data and information from the CyanoMutants repository (Nakamura et al., 1999). For example, the unavailability of a mutant lacking the *folB* gene (*slr*1626), coding for the dihydronopterin aldolase, an important enzyme involved in the folate biosynthesis (see CyanoMutants), indicates that there is only one possible pathway for folate biosynthesis. Correspondingly, the *in silico* gene deletion (*folB*) results in incapability of the model to produce the essential cofactor tetrahydrofolate.

Other examples include Quintero et al. (2000), who generated mutants showing an impaired arginine synthesis as a results of a knock-out of gene *slr*0902, coding for a ornithine carbamoyltransferase (ArgF). Also, an insertion in the gene (*slr*0661) coding a
pyrroline-5-carboxylate reductase (ProC) resulted in an impaired proline synthesis. Both mutants had to be exogenously supplied by a supplementation of the growth media with either arginine or proline, respectively. Autotrophic growth of either mutant could not be achieved. The phenotype of both mutants can be confirmed in silico. Under conditions of photoautotrophic growth, the genes coding for ArgF and ProC are classified as essential in the model but are dispensable if arginine and proline are supplied in silico, respectively.

Nonetheless, discrepant findings were observed for 4 genes for which a fully-segregated mutants were reported, despite being essential in the model. On closer examination, 2 of these 4 genes correspond to the hydrogen carbonate transporter Cmp (slr0040, slr0041). In silico essentiality is due to Cmp being the only carbon transport mechanism present in our reconstruction. However, in Synechocystis sp. PCC 6803 more than one carbon uptake mechanism exists. A further gene (slr0847) encodes CoaD, a phosphopantetheine adenylyltransferase (EC 2.7.7.3), located in biosynthesis pathway of Coenzyme A. Within our model, the enzyme catalyzes an essential step, the conversion from pantetheine 4’-phosphate to 3’-dephospho-CoA. Consequently a mutant lacking this gene would be unable to synthesise CoA. The fact that it was possible to construct a viable mutant implies that the mutant was either not fully segregated or there exists an unknown alternative pathway for the synthesis of CoA in Synechocystis. The last discrepant gene (sll0934) in annotated as CcmA, a carboxysome formation protein, as well as a 3-deoxy-7-phosphoheptulonate synthase, which catalyses the initial step of the aromatic amino acid synthesis. The latter is essential within our model. A fully-segregated mutant therefore suggests the existence of alternative steps within the formation of aromatic amino acids.
Supplemental Material

The following supporting information is supplied by the online version of this article.

Supplemental Data 1. Synechocystis Network File, including IDs for metabolites to the CHEBI and KEGG database and UniProt IDs for genes, SBML format

Supplemental Data 2. List of reactions, associated enzymes and genes implemented in the network with reference to literature, Excel file

Supplemental Data 3. Newly annotated genes for Synechocystis, Excel file

Supplemental Data 4. Comparison of gene knock out predictions in the model with in vivo data available from the CyanoMutants database, Excel file

Supplemental Data 5. Results of gene deletion study of all genes in the Synechocystis network, Excel file

Supplemental Data 6. List of additional enzymes associated to the Synechocystis metabolism available from the KEGG database, Excel file

Supplemental Data 7. Synechocystis Network compatible for usage with COBRA toolbox v. 1.3.1, SBML format

Supplemental Data 8. Instruction and parameter settings to recalculate the results in Fig. 2 - 4 using the COBRA toolbox v. 1.3.1, PDF file
Acknowledgements

We thank N. Swainston for advice on SBML and metabolite nomenclature, Eva Grafahrend-Belau for advice on the biomass objective function, as well as M. Hagemann and M. Eisenhut for advice on photorespiratory metabolism. We further thank all members of the FORSYS consortium "Systems Biology of cyanobacterial biofuel production" for discussion and support.
LITERATURE CITED


MICROBIAL CULTURE COLLECTION
National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki


Rasmusson AG, Geisler DA, Moller IM (2007) The multiplicity of dehydrogenases in the electron transport chain of plant mitochondria, Mitochondrion, Volume 8, Issue 1, Unique aspects of plant mitochondria: 47-60


Figure 1. An overview of metabolic pathways and reactions included in the metabolic reconstruction of *Synechocystis* sp. PCC 6803. The model consists of 380 reactions, including the primary metabolic pathways, transport reactions, light absorption, photophosphorylation, photosynthetic NADP⁺ reduction, biomass formation and cellular maintenance.

Figure 2. The predicted maximal growth rate as a function of light intensity and total CO₂ uptake.

Figure 3. A predicted flux distribution to achieve maximal phototrophic growth yield (BOF). All flux values are given relative to the synthesis of RuBP (100%). The simulation was performed for a given light intensity, while carbon uptake was allowed to adjust as necessary to achieve optimal yield. Intervals indicate variability in possible flux values with equivalent yield. CO₂ released by internal reactions re-fixed by RuBisCO. For abbreviations and nomenclature see Materials and Methods.

Figure 4. The transition from storage-utilizing night metabolism to phototrophic growth. Shown is the flux through selected reactions (color-coded), as a function of absorbed photons (x-axis) and glycogen utilization (y-axis), see colorbar for the respective ranges. (a) The biomass objective function (BOF); (b) flux through RuBisCO carboxylase; (c) flux through RuBisCO oxygenase; (d) net O₂ uptake; (e) flux through the TCA bypass; and (f) proline degradation. The *in silico* simulation assumes a constant ATP requirement of 0.27 mmol gDW⁻¹ h⁻¹ for cellular maintenance. All flux values are reported in [mmol gDW⁻¹ h⁻¹], except for BOF. Net O₂ uptake attains negative values for respiratory metabolism.

Figure 5. The key photorespiration pathways in *Synechocystis* sp. PCC 6803, taking into account the results of Eisenhut et al. (2006, 2008). Oxygenation of RuBP generates one molecule of PG3 and one molecule of 2PG. The latter is converted into glyoxylate and either recycled into the Calvin-Benson cycle or utilized as a precursor for glycine synthesis.
Table I. The Biomass objective function (BOF), adopted and modified from Cogne et al. (2003). All quantities are given in mmol, except biomass.

<table>
<thead>
<tr>
<th>Components</th>
<th>Reaction equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
<td>$1 \text{ ATP} + 2 \text{ GTP} + 0.058 \text{ L-Isoleucine} + 0.094 \text{ L-Leucine} + 0.036 \text{ L-Lysine} + 0.018 \text{ L-Methionine}$ $+ 0.039 \text{ L-Phenylalanine} + 0.054 \text{ L-Threonine} + 0.009 \text{ L-Tryptophan} + 0.076 \text{ L-Valine} + 0.053 \text{ L-Arginine} + 0.014 \text{ L-Histidine} + 0.101 \text{ L-Alanine} + 0.098 \text{ L-Aspartate} + 0.132 \text{ L-Glutamate} + 0.007 \text{ L-Cysteine} + 0.086 \text{ Glycine} + 0.038 \text{ L-Proline} + 0.054 \text{ L-Serine} + 0.033 \text{ L-Tyrosine} + 2 \text{ H}_{2}\text{O} \Rightarrow 2 \text{ GDP} + \text{1 AMP} + \text{Pyrophosphate} + 2 \text{ Orthophosphate} + \text{1 Protein}$</td>
</tr>
<tr>
<td>DNA</td>
<td>$0.222 \text{ dC} + 0.279 \text{ dT} + 0.222 \text{ dG} + 0.279 \text{ dA} \Rightarrow \text{1 Diphosphate} + \text{1 DNA}$</td>
</tr>
<tr>
<td>RNA</td>
<td>$0.322 \text{ GTP} + 0.262 \text{ ATP} + 0.2 \text{ CTP} + 0.216 \text{ UTP} \Rightarrow \text{1 Diphosphate} + \text{1 RNA}$</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>$0.1379 \text{ UDP-Glucose} + 0.8143 \text{ dTDP-Rhamnose} + 0.0478 \text{ Glycogen} \Rightarrow 0.8143 \text{ dTDP} + 0.1379 \text{ UDP} + \text{1 Carbohydrates}$</td>
</tr>
<tr>
<td>Lipids</td>
<td>$1 \text{ sn-Glycerol 3-phosphate} + 0.7153 \text{ UDP-D-galactose} + 0.1807 \text{ UDP-glucose} + 1.470 \text{ Acetyl-CoA} + 0.7350 \text{ H}_{2}\text{O} + 1.807 \text{ Hydrogen-sulfide} \Rightarrow 1.470 \text{ CoA} + 0.8960 \text{ UDP} + 0.7350 \text{ Orthophosphate} + \text{1 Lipids}$</td>
</tr>
<tr>
<td>Total</td>
<td>$0.8404 \text{ Protein} + 0.0039 \text{ DNA} + 0.013 \text{ RNA} + 0.123 \text{ Carbohydrates} + 0.0182 \text{ Lipids} \Rightarrow 1 \text{ g gDW}^{-1} \text{ h}^{-1} \text{ Biomass}$</td>
</tr>
</tbody>
</table>