AN ABSTRACT OF THE THESIS OF

Robert M. Elder for the degree of Honors Baccalaureate of Science in Chemical Engineering presented on June 2, 2009. Title: Flux Balance Analysis of *Synechocystis* PCC 6803.

Abstract approved: _____

In order to improve photosynthetic hydrogen production, initial work has begun on a flux balance analysis of Synechocystis sp. strain PCC 6803, a model H2-producing organism. The mathematical model developed includes reactions relevant to photoautotrophic growth: glycolysis, the tricarboxylic acid cycle, the Calvin cycle, the photosynthetic electron transport chain, the cyanobacterial carbon concentrating mechanism (CCM), and ancillary reactions. The base model is adapted to simulate four situations: WT Synechocystis in low and high carbon conditions; M55, a Synechocystis mutant lacking Type I NAD(P)H dehydrogenase, a key component of the CCM; and Mallomonas papillosa, an organism completely lacking a CCM. To explore their metabolic capabilities, these models are subjected to different stresses: increased energy drain with fixed solar input and increased solar input with fixed biomass production. These experiments reveal that energy overflow metabolism is closely related to the ratio of Photosystems I and II, and that this ratio is strongly affected by mass and energy conservation constraints. Conceivably, energy overflow could be controlled to increase hydrogen production.

Key Words: Flux Balance Analysis, Synechocystis, metabolic engineering, cyanoabacteria

Corresponding e-mail address: rmelder@gmail.com

Flux Balance Analysis of *Synechocystis* PCC 6803

by

Robert M. Elder

A PROJECT

submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Chemical Engineering (Honors Scholar)

Presented June 2, 2009
Commencement June 2009

Honors Baccalaureate of Science in Chemical Engineering project of Robert M. Elder presented on June 2, 2009

APPROVED:

Mentor, representing Biological and Ecological Engineering

Committee Member, representing Biological and Ecological Engineering

---

Committee Member, representing Biological and Ecological Engineering

---

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

Robert M. Elder, Author

TABLE OF CONTENTS

LIST OF FIGURES

Flux Balance Analysis of *Synechocystis* PCC 6803

Introduction

In light of growing global energy demands, the inevitable depletion of fossil fuel reserves, and the apparent onset of climate change, the production of sustainable fuels has become increasingly attractive. One possible source for such a fuel is photosynthetic biological hydrogen production. Although many organisms are capable of converting solar energy into H_2, production rates are extremely low. These organisms must be engineered to produce hydrogen at commercial scales, but genetic engineering requires a thorough understanding of an organism's metabolism. Therefore, the aim of the current work is to elucidate the metabolic capabilities of a model hydrogen-producing organism, *Synechocystis* sp. strain PCC 6803. Initial work has focused on modeling the carbon concentrating mechanisms (CCMs) of *Synechocystis* since these are key factors in the energetics of this organism.

Background

Photosynthesis

Photosynthesis occurs in two phases. First, the reactions forming the photosynthetic electron transport chain (ETC) produce chemical energy (ATP) and reductive potential (NADPH). Second, the solar energy captured in NADPH and ATP is converted to other useful forms.

Photosynthetic electron transport is facilitated by two major enzyme complexes: photosystems I and II (PSI and PSII, respectively). These enzymes are aided in light capture by so-called light harvesting complexes (LHCs), different versions of which are expressed in response to different conditions.[1] The photosynthetic ETC is shown in Figure 1.

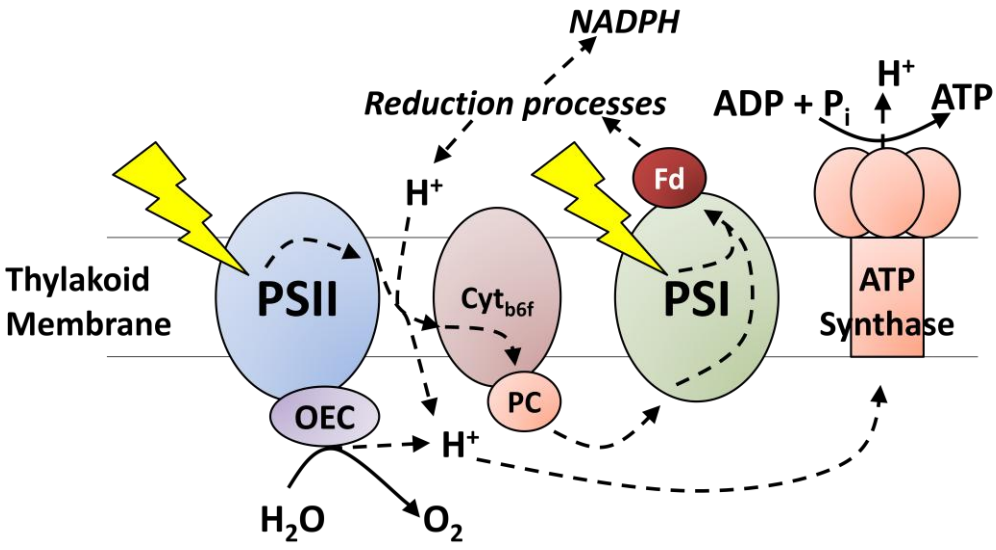

Figure 1: Photosynthetic electron transport chain

Photons captured by the antenna system, a complex array of LHCs, feed energy into PSII, which catalyses the first step of the photosynthetic ETC: the formation of a chlorophyll radical cation, the strongest oxidant found in biology.[1] Several components cause this chlorophyll to reduce water and pass electrons down the chain through multiple enzyme complexes, including PSI, to ferredoxin, which is used to generate NADPH, an electron carrying coenzyme used in many other biochemical reactions. PSI must also absorb photons to reduce ferredoxin. Simultaneously, the ETC and PSII release protons into the thylakoid membrane, creating an H^+ concentration gradient. ATP synthase uses this gradient to produce ATP.

The flow of electrons from PSII through PSI to ferredoxin is known as *linear* electron transport; this process produces a relatively high ratio of NADPH to ATP. When necessary, organisms can regulate this ratio by using *cyclic* electron transport, which moves electrons cyclically around PSI using ferredoxin in conjunction with Type I NAD(P)H dehydrogenase (NDH-1), increases the H^+ gradient, and generates additional ATP but less NADPH.[1] Type II NAD(P)H dehydrogenase (NDH-2) can also recycle e^- into the PQ pool but does not contribute to the H^+ gradient necessary for ATP production. The flux of e^- through the linear and cyclic pathways is

tightly regulated to produce the proper amounts of ATP and NADPH, allowing the organism greater flexibility. However, cyclic electron transport can limit the rate of biological hydrogen production by consumption of protons and reduced ferredoxin.[2] Another pathway, termed *pseudocyclic* e^- transport, prevents excess reduction of the PQ pool by allowing the organism to bleed off excess electrons to O_2.[3]

In the second phase, photosynthetic reduction, the solar energy captured in NADPH and ATP is converted to another useful form. This process, known as the Calvin Cycle, is catalyzed by the well-known enzyme ribulose bisphosphate carboxylase/oxygenase (RuBisCO), among others. As the cycle begins, ribulose-1,5-bisphosphate (RuBP) is combined with CO_2 by RuBisCO and split into two molecules of 3-phosphoglycerate (3-PGA). Some of these are regenerated into RuBP in a series of reactions; others are used to create higher carbohydrates, starches, and other useful metabolites.

Starches and carbohydrates can be fed into glycolysis to produce pyruvate, which is fed into the tricarboxylic acid cycle to generate a variety of other useful metabolites.[5] The TCA cycle also produces NADH necessary to feed oxidative phosphorylation. Thus, energy captured from sunlight and stored in various chemical forms can be used when needed, such as when light conditions are poor or it is dark.

Despite RuBisCO's important role, the enzyme is extraordinarily slow and inefficient: it catalyzes only a few reactions per second[4], which is meager compared to the tens of thousands catalyzed by many enzymes; and, as the name suggests, it catalyzes a competing oxygenation reaction. In this reaction, the beginning of what is known as photorespiration, RuBisCO reacts with oxygen to produce one molecule each of G3P and phosphoglycolate, the latter of which is converted to G3P through a long series of steps that result in the uptake of O_2 and evolution of CO_2.[5]

Although RuBisCO has a much higher specificity for CO_2 than for O_2[1], the concentration of O_2 at normal conditions is 25 times higher than CO_2.[5] Significant competition occurs between the carboxylation and oxygenation reactions.

Cyanobacterial Carbon Concentrating Mechanism

Photosynthetic organisms have evolved a variety of methods to circumvent RuBisCO's dual nature. Cyanobacteria have evolved a multi-component system collectively known as the carbon concentration mechanism (CCM). The cyanobacterial CCM acts to increase the concentration of carbon dioxide in the immediate vicinity of RuBisCO, which decreases the likelihood of unwanted photorespiration.

The system is highly effective: although RuBisCO has a low affinity for CO_2 with a Michaelis constant K_M higher than 150 μM, cyanobacteria can achieve maximum enzyme activity with external CO_2 concentrations lower than 15 μM.[6]

The CCM consists of two major components: uptake mechanisms for carbon species (CO_2 and HCO_3^-) and a specialized compartment to localize CO_2 near RuBisCO. Certain components of the CCM are constitutively expressed; others are expressed only in response to low levels of environmental carbon.

Active uptake mechanisms are required by cyanobacteria to achieve rates of uptake far in excess of diffusion alone. Five distinct mechanisms have been identified in *Synechocystis* PCC 6803, our model species: three HCO_3^- active transporters and two CO_2 uptake systems. The former group consists of SbtA and BicA, Na^+-dependent transporters, and BCT1, an ATPase, all of which are inducible under conditions of carbon limitation. The latter group consists of specialized NDH-1 complexes: NDH-1_4 is constitutively expressed while NDH-1_3 is inducible under carbon limitation. These are referred to as uptake systems rather than active transporters

because they rely on passive diffusion of CO_2 into the cell, which they convert to HCO_3^-.[7] The CCM mechanisms are shown in Figure 2.

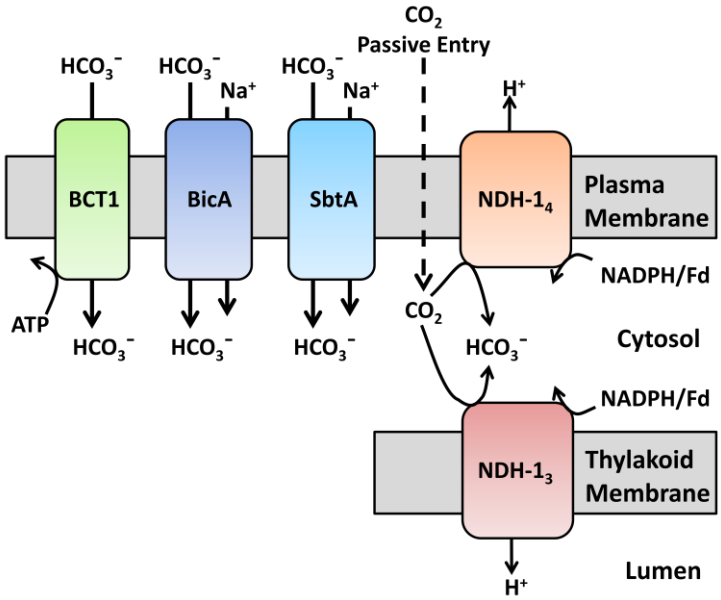

Figure 2: The carbon concentrating mechanism in *Synechocystis*

HCO_3^-, being an ionic species, is about 1000 times less permeable to lipid membranes than CO_2; therefore, it is the preferred species for carbon concentration elevation, and levels of HCO_3^- are maintained by the absence of carbonic anhydrase in the cytosol. The uptake mechanisms can concentrate carbon to levels 1000 times the extracellular concentration.[8]

A specialized compartment for CO_2 localization is realized in the carboxysome, a proteinaceous, RuBisCO-packed, polyhedral structure located within the thylakoid membrane.[7] Each cyanobacterium contains a large number of carboxysomes, which contain essentially all of their RuBisCO. HCO_3^- is transported across the protein shell, where localized carbonic anhydrases convert it to CO_2, which is used by RuBisCO in photosynthetic reduction.

Carboxysomes are supplemented with mechanisms to reduce leakage and recycle leaked CO_2. The proteinaceous carboxysome coat is related to the first function[9]; the carbon uptake systems, NDH-1_4 and NDH-1_3, fulfill the second function by converting leaked CO_2 into HCO_3^-.[10]

These complex adaptations, while effective, require a substantial investment of energy on the part of *Synechocystis*. In maintaining the strong gradient of carbon, active transporters require energy in the form of ATP or a Na^+ gradient, and maintaining the numerous enzymes and structures involved requires additional investments.

However, *Synechocystis* RuBisCO enzymes have evolved to operate more efficiently in an environment of concentrated CO_2. Although they have relatively lower affinities for substrates CO_2 and O_2, they have a much higher turnover rate, which allows cyanobacteria to produce fewer RuBisCO enzymes. Considering that RuBisCO constitutes 30% of the total proteins in most leaves[1], this likely saves considerable metabolic energy.[7]

Additionally, the CCM confers added flexibility on cyanobacteria. Under conditions of excess carbon, only the basal, constitutive mechanism, NDH-1_4, is active. Carbon limitation causes the expression BCT1, SbtA, BicA, and NDH-1_3, which drastically increase the ability to take up carbon. These mechanisms allow for efficient response to environmental conditions, such as temperature, pH, moisture, light, and oxygen level.[11]

It is, therefore, evident that the energy investment in the carbon concentrating mechanism is outweighed by the performance increase it bestows.

Biosolar Hydrogen Production

Biological hydrogen, the product of various hydrogenases, is ubiquitous throughout prokaryotes and eukaryotes, albeit in small quantities.[12] H_2 is produced by reducing elemental hydrogen with electrons from ferredoxin.[13]

In cyanobacteria, such as *Synechocystis* PCC6803, these hydrogenases contain [Ni-Fe] catalytic centers that are extremely sensitive to inactivation by O_2. This is one of the major barriers to biosolar hydrogen production. Reducing intracellular O_2 content, and thereby

increasing H_2 production, can be accomplished with natural mechanisms, such as consumption by respiration, chemical reduction via PSI, and reversible inactivation of PSII O_2 evolution.

Of the few H_2-producing organisms that have been closely studied, most were selected for ease of cultivation or as model systems. Even among this small subset, O_2 sensitivity and H_2 production differ by two orders of magnitude. For example, the hydrogenase of green algae *C. reinhardtii* is up to 100 times more active than those of cyanobacteria but also extremely sensitive to O_2.[1]

It is clear that improvements could be made by exploring the metabolic capabilities of numerous species.

The Basis for Mathematical Modeling

Quantitative modeling of the biochemical pathways that sustain life is a monumental task: the dynamics of these intricate systems, including kinetic and regulatory aspects, are difficult to measure and are not known with good certainty.

Flux Balance Analysis

One method, flux balance analysis (FBA), avoids the difficulties of dynamic models. This approach relies on a steady-state analysis using the stoichiometry of a given system, which is combined with experimentally attainable information like gross consumption and production rates. Although FBA seems to lack the sophistication required to analyze complex metabolic networks, the simple mass conservation constraint of stoichiometry defines an organism's widest limits, which may be further narrowed by kinetics and regulatory actions as these become available.[14]

Flux balance analysis is defined by the steady-state equation

$$S \cdot v = b,$$

where **S** is the stoichiometric matrix of the metabolic network, **v** is the vector of reaction fluxes, and **b** is a vector of metabolite production rates.[14] **b** is taken to be zero when cellular exchange fluxes are included in the model, because the conservation of mass is completely satisfied. Consider the simple reaction network shown in Figure 3.

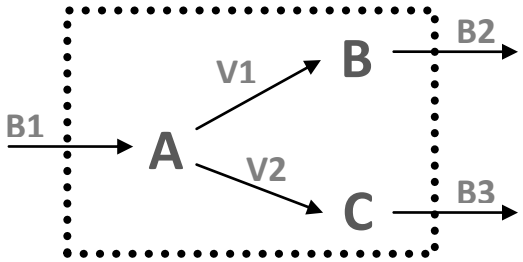

Figure 3: An example of a simple metabolic network

The stoichiometric matrix for this system is shown in Figure 4.

$$S = \begin{array}{c|ccccc} & \boldsymbol{V1} & \boldsymbol{V2} & \boldsymbol{B1} & \boldsymbol{B2} & \boldsymbol{B3} \\ \boldsymbol{A} & -1 & -1 & 1 & 0 & 0 \\ \boldsymbol{B} & 1 & 0 & 0 & -1 & 0 \\ \boldsymbol{C} & 0 & 1 & 0 & 0 & -1 \end{array}$$

Figure 4: Stoichiometric matrix of simple metabolic network shown in Figure 3

This example involves only three metabolites and five fluxes. Clearly, even modeling glycolysis and the tricarboxylic acid cycle, which are basic to any metabolic model, will create much more extensive matrices.

Typically, the number of metabolites is less than the number of reactions, creating an underdetermined system with a region containing infinite feasible solutions.[16] The feasible space is narrowed by applying physiologically relevant constraints to each flux: for example, reversible reactions may be considered to have negative lower bounds, or a reaction can be eliminated if it is known to be inactive under certain conditions. A particular solution may be found by choosing an objective function, such as production of biomass or a valuable metabolite, and maximizing the value using linear programming.[14]

Linear programming (LP) is a method for optimizing a particular linear objective function subject to constraints. In this application, the constraints are the conservation of mass and upper and lower boundaries on fluxes. The LP problem is formulated as:

Maximize $Z = c \cdot v$

Subject to $S \cdot v = 0$

$v_{i,min} \leq v_i \leq v_{i,max}$

The objective function **Z** is defined by **c**, a vector with arbitrary coefficients for the flux to be maximized and zero for all other fluxes.[16] **c** can also have multiple non-zero coefficients, in which case **Z** is a linear combination of fluxes. Again, consider the simple metabolic network above, with constraints included as shown in Figure 5.

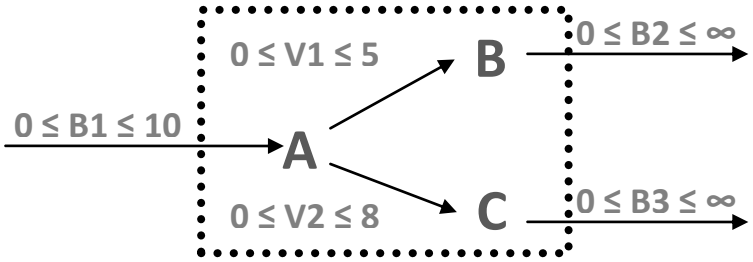

Figure 5: An example of a simple metabolic network with constraints

The solution space for the system and constraints in Figure 5, along with examples of objective functions and optimal solutions, is shown in Figure 6. It can be shown that optimal solutions are always on or between "corner points," which are defined by the intersection of multiple constraints. Essentially, linear programming (Simplex method) traverses these corner points to find the optimum. If the objective function falls on a line between two corner points, the LP problem has multiple optimal solutions: that is, different flux distributions can lead to the same optimum value of the objective function. In models of greater complexity, the optimal solution may fall on a hyperplane between numerous corner points.

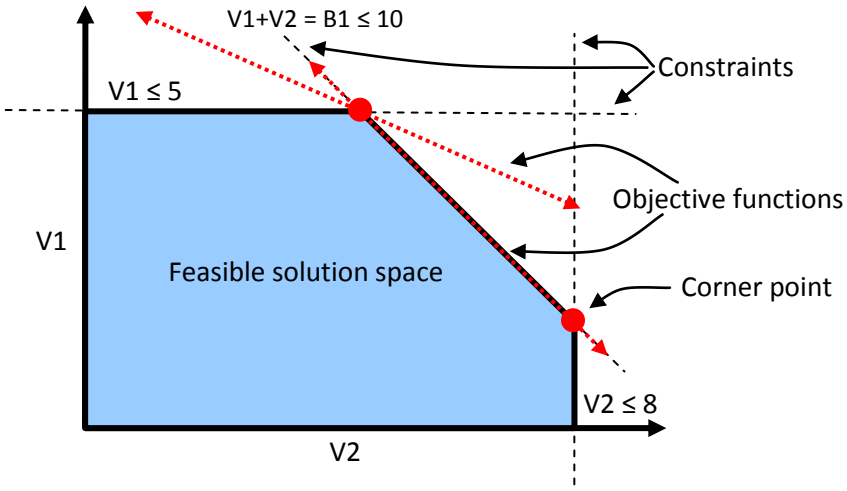

Figure 6: Solution space for a simple metabolic system

The great complexity of metabolic models often leads to multiple optima. The number of optima is related to network flexibility; consequently, as model organisms are given more energy, the number of optima tends to increase. Alternate optima have been interpreted in a variety of ways: for example, as the effect of exogenous factors, like temperature and stress, or as phenotypic heterogeneity within a population.[15]

Whatever the interpretation, a large number of optima makes it difficult to choose a representative flux distribution. One method, which assumes that an organism tends toward efficiency by minimizing total metabolic flux, is to choose the flux distribution with the lowest Euclidean norm. This approach was used when the models considered here showed multiple optima.

With the advent of rapid genome sequencing, existing biochemical knowledge can be synthesized into stoichiometric models approaching the genome-scale. The stoichiometry of many biochemical reactions is already well known.[16] Fully sequenced genomes often identify the specific gene sequence responsible for a particular reaction. Combining this information with FBA creates a framework for understanding whole cell metabolism by assisting with data

interpretation, predicting phenotypic responses to environmental conditions, and guiding for genetic engineering.

Random Sampling of Metabolic Models

The feasible space defined by constraints applied in flux balance analysis can also be studied using random sampling and statistical description.[17]

This approach has been used in a variety of ways. Simulating a variety of experiments and using Monte Carlo sampling to simulate measurement noise has allowed the selection of more informative experiments.[18] Certain reactions, termed a "high flux backbone," are those shown to be necessary for growth under different media conditions; the number of reactions in the backbone has been shown to reflect the resiliency of *H. pylori*, *E. coli*, and *S. cerevisiae*.[19,20] A model of the human cardiac mitochondria network described a known condition in diabetic patients. Decreased pyruvate dehydrogenase flux, which was believed to stem from unknown regulatory actions in diabetics, was shown to be a fundamental mass conservation limit.[21]

Random sampling has typically been accomplished using Monte Carlo methods. Specifically, the Markov Chain Monte Carlo "hit-and-run" sampler, which produces a valid solution during each iteration, is used with an improvement known as artificial centering to improve sampling behavior.[16] Collectively, these techniques are known as artificial centering hit-and-run (ACHR) and have been used for essentially all publications in the field.[17,22]

Experimental Goal

Although *Synechocystis* is capable of producing H_2, the natural production rate and efficiency are extraordinarily low. Additionally, the biochemical mechanisms of hydrogen formation are not well understood, which makes optimization challenging. The metabolic capabilities of cyanobacteria must be explored; it is posited that this can be accomplished using a flux balance analysis approach.

Materials and Methods

Model Overview

The reaction network of *Synechocystis* PCC6803 was adapted from previous models.[23,24] The final stoichiometric network covers the functions of photoautotrophic growth: glycolysis, the TCA cycle, the Calvin Cycle, the photosynthetic ETC, the cyanobacterial CCM, transport fluxes, pH regulation mechanisms for the thylakoid lumen and the cytosol, and maintenance reactions. Dark-only reactions such as the glyoxylate shunt or fermentative pathways are not included. The biomass production equation, involving many metabolites, was taken from a previous study.[24] The set of reactions is available in the Appendix.

Specific Model Formulations

WT *Synechocystis* was modeled under high and low carbon conditions. High carbon conditions are associated with the basal form of CCM: NDH-1_4 and SbtA/BicA are active. Since SbtA and BicA catalyze the same reaction, they are expressed as a single reaction. Low carbon conditions fully activate the CCM: in addition to the high carbon components, NDH-1_3 and BCT1 are also active.

In addition to WT Synechocystis, we also model a mutant (M55) and *Mallomonas papillosa*, an organism that lacks a CCM. M55 is an *ndhB* deletion mutant with impaired NDH function. M55 lacks all NDH-1 enzymes, but has active BCT1 and SbtA/BicA. In addition to the CCM, the *ndhB* deletion also impairs the ability of M55 to use cyclic e^- transport. The chemical interconversion of CO_2 and HCO_3^- takes on an important role in the M55 mutant; this reaction was considered kinetically irrelevant in the CCM-containing models. The CCM-lacking mutant does not possess BCT1, SbtA/BicA, NDH-1_3, or NDH-1_4, but retains NDH-1_2, which is important for cyclic e^- transport. Additionally, this model lacks a carboxysome, which is missing in *M. papillosa*.

Experiments

Two experiments were used to explore the metabolic capabilities of the modeled organisms. In the first experiment, random sampling is used to test the effects of energy drain and photosystem ratio. In the second experiment, alternate optima are enumerated at increased light levels in order to better understand the impact of light intensity on photosystem stoichiometry.

In the first experiment, each of the models was randomly sampled at different levels of ATP maintenance energy and ratio of PSI:PSII photon activity, with either ammonia (NH_3) or nitrate (NO_3) as a nitrogen source. Each set of random samples was averaged to produce a single representative data point for each set of conditions.

ATP maintenance energy is an umbrella ATP sink that accounts for various cellular maintenance activities, such as DNA repair. Literature indicates a value of 18.61 mmol ATP/g DCW produced; the value was varied from 0 to 50.[24] In a real organism, the ratio of PSI:PSII photon activity would vary with the effectiveness of the light harvesting complexes associated with each enzyme and with the amount of each enzyme; this ratio is varied from 1:8 to 8:1.

In the second experiment, the effect of increasing excess light level on each of these models is explored. First, the biomass production and carbon intake are set to constant amounts, and the FBA model is optimized for minimum photon usage. This results in a feasible model requiring some minimum photon input. Then, keeping biomass production and carbon intake constant, the light input is increased over the required amount, and all alternate optima are enumerated. From this set is chosen the most efficient optimum, which has the lowest Euclidean norm, as discussed above. ATP maintenance energy is fixed at 18.61 mmol ATP/g DCW.

Results and Discussion

The results of the first experiment, testing the effect of ATP maintenance energy and PSI:PSII ratio, are shown in Figure 7. Only the results for trials using NH_3 are shown because NO_3 trials did not exhibit qualitative behavioral differences, aside from a minor reduction in biomass production. This reduction in biomass production is expected: when growing on nitrate, *Synechocystis* converts NO_3 to NH_3 using NADPH. The range of values at each maintenance energy is caused by different ratios of PSI:PSII.

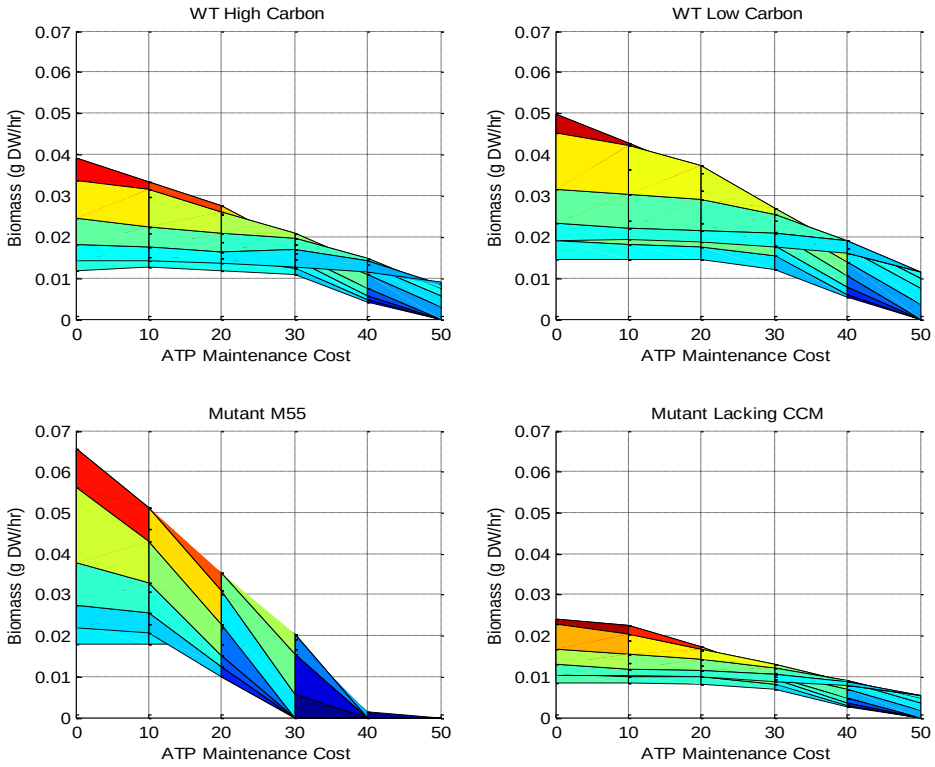

Figure 7: Results of random sampling of four models

As expected, biomass production decreases with increasing maintenance cost.

Surprisingly, the metabolically-deficient mutant M55 produces more biomass than either high- or low-carbon WT. This can be attributed to the lower investment in the energy-intensive CCM reactions; however, the growth rate of M55 is limited by CO_2 diffusion, which is

much slower than the uptake rate of the CCM. Therefore, for a given amount of solar energy, M55 will produce more biomass at substantially lower rate.

Attributing the higher biomass production of M55 to lower investment in CCM reactions seems inconsistent with the behavior visible in the WT model. By this rationale, the low carbon WT trial, with fully activated CCM, should produce less biomass than the high carbon WT trial, and *M. papillosa* should produce far more biomass than the high carbon WT trial. However, the opposite trend is observed.

This is attributed to the energy inefficiency of cyclic e^- transport: M55, lacking the ability to use this pathway, is able to use solar energy more efficiently. The other variants, with an active e^- recycling pathway, are less efficient. However, this inefficient pathway also increases versatility: additional control over ATP:NADPH production allows these organisms to survive even with high maintenance demands. From Figure 7, it is clear that M55 is less adaptable.

Figure 8 shows the location of the maximum biomass production rate for each ATP maintenance cost. Again, these results are only for NH_3 trials because the NO_3 trials were only slightly different.

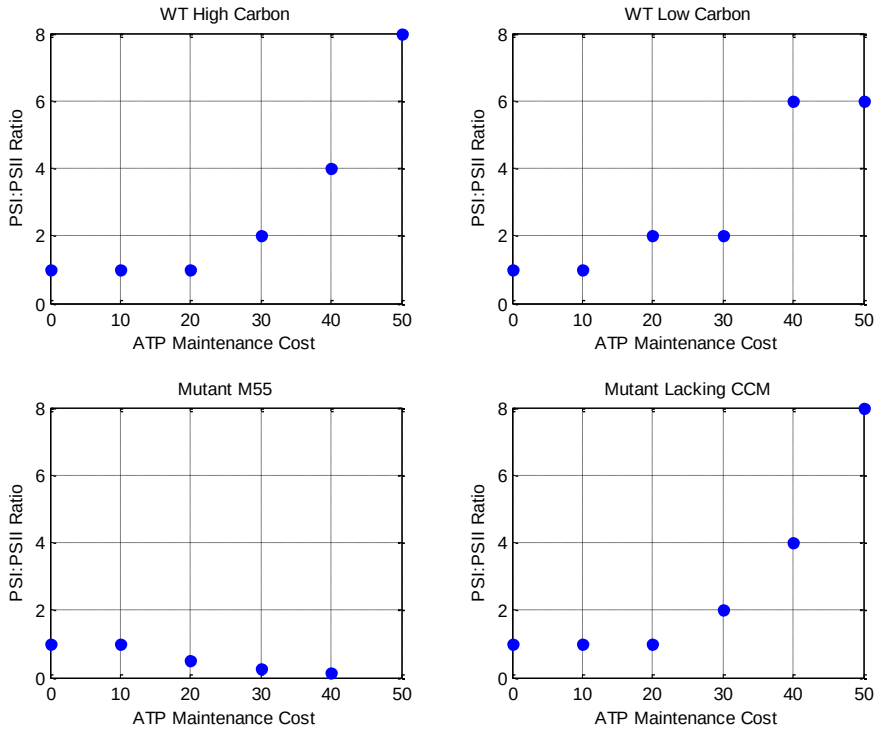

Figure 8: PSI:PSII ratio that maximizes biomass production at each maintenance cost

The ratio maximizing biomass production increases with ATP maintenance cost for all variants except for M55. This behavior is driven by the need to produce additional ATP: cyclic e^- transport must be used to shuttle electrons around PSI, generating ATP. However, this process cannot occur without sufficient PSI photon absorption: PSI requires solar input to energize electrons, allowing them to reduce ferredoxin and continue cyclic e^- transport. At lower PSI:PSII ratios, the e^- not accommodated by PSI overflow into an e^- sink reaction, which is typically the reduction of oxygen using thylakoid protons by Cyt_{553}, which carries electrons from Cyt_{b6f}. This reaction simultaneously wastes e^- and lowers the H^+ gradient necessary for ATP production. Therefore, higher ratios support additional cyclic e^- transport and higher biomass production rates.

M55, lacking NDH-1, is incapable of cyclic e^- transport. In this case, linear e^- transport must be used to meet the increased energy demand, so higher amounts of PSII are required.

Therefore, as shown in Figure 8, low PSI:PSII ratios enable the highest biomass production rate. M55 is also susceptible to e^- overflow.

Conceivably, energy overflow metabolism could be used to produce hydrogen rather than water if the necessary mechanisms were inserted.

The effect of increasing light level on the low carbon and M55 models is shown in Figure 9. Results are only shown for the WT Low Carbon and Mutant M55 models because the WT High Carbon and *M. papillosa* resemble WT Low Carbon.

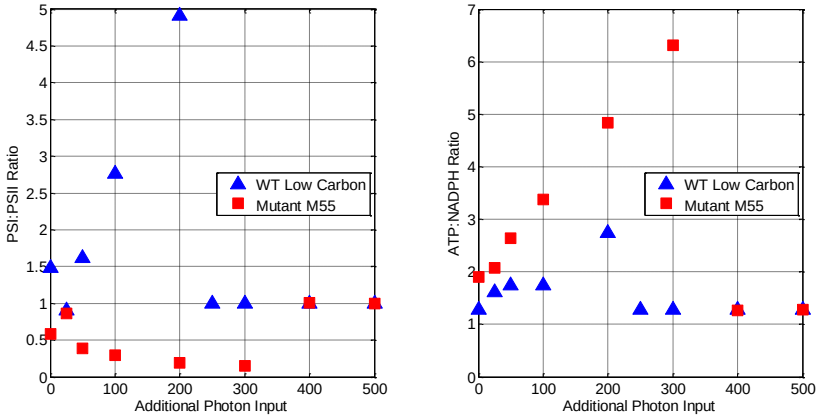

Figure 9: Effect of increasing excess light on PSI:PSII ratio and ATP:NADPH production ratio

This experiment is designed to explore the energy overflow metabolism in each of the models. Growth is limited by carbon availability, which is fixed, so any light energy in excess of the amount required for growth must be dissipated.

Both models use similar mechanisms to accomplish this goal. As light increases, a futile cycle forms between pyruvate kinase and phosphenolpyruvate synthase, which results in the consumption of ADP and the production of AMP. Combined with the action of adenylate kinase, which uses this AMP and an ATP to produce two ADP, the net result of this cycle is the production of ADP. This is necessary because ATP production, which requires ADP, increases with light level.

The most striking feature in Figure 9 is the sudden transition to a PSI:PSII ratio of 1 and ATP:NADPH ratio of 1.28. An ATP:NADPH ratio of 1.28 (9:7) is the predicted value for linear e^- transport, which is defined by a 1:1 ratio of PSI and PSII: therefore, both models transition to entirely linear e^- transport at high light levels.[3] A similar transition has been observed experimentally.[25] The high light state is characterized by excess NADPH being consumed via pseudocyclic e^- transport and a complex futile reaction network. This network uses phosphofructokinase, fructose bisphosphate aldolase, transaldolase, sedoheptulose-1,7 bisphosphate-GAP lyase, and sedoheptulose-1,7-bisphosphate 1-phosphohydrolase to shuffle sugars and dissipate ATP. Additionally, BCT-1 is used when possible to dissipate extra ATP when transporting HCO_3^- into the cell.

Excess energy could be dissipated more simply than the process above: the pyruvate-phosphoenolpyruvate futile cycle could use all additional ATP and pseudocyclic e^- transport could use all additional NADPH. Doubtless, distributions such as these are present in the numerous alternate optima generated in this experiment. However, distributions with a few high flux reactions are strongly disfavored by choosing the minimum Euclidean norm as the representative optimum; conversely, distributions with flux spread out evenly are favored. This seems reasonable from a biological standpoint: an organism is unlikely to strongly up-regulate a few reactions when a minor adaptation of many reactions will accomplish the same task, especially in the short-term. Additionally, pathways proceeding linearly through multiple intermediates are disfavored by the minimum Euclidean norm method.

Low light states, those prior to the transition, are characterized by increasing cyclic e^- transport, which is indicated by an increasing PSI:PSII ratio. No particular pattern is evident in the e^- sinks used to dissipate energy.

It is evident from Figure 9 that WT Low Carbon is better able to transition to the high light state at a lower light level than Mutant M55. The same is true of both WT High Carbon and the *M. papillosa* (data not shown). In most cases, increasing the light level increases the total metabolic flux; however, when the transition to the high light state occurs in M55, the Euclidean norm actually decreases. This suggests that M55 must overcome some critical threshold to access the high light state.

The nature of this threshold is unclear. The reaction missing from only Mutant M55 is NDH-1_2, which is largely unused by the other models. M55 displays several unique behaviors prior to transitioning to the high light state: it uses the pyruvate-phosphoenolpyruvate futile cycle before the transition, whereas the other models are able to limit their ATP production without this pathway; it uses linear e^- transport almost exclusively; and it produces a fixed amount of NADPH, with all excess energy from PSII being funneled through Cyt_{553} to an e^- sink reaction. The Cyt_{553} e^- sink, which is not used by the other models, also dissipates the thylakoid proton concentration gradient. These behaviors suggest that M55 has less control over its energy production than do the other models but do not elucidate the nature of the threshold.

Modulation of photosystem stoichiometry in response to light condition has been observed previously in *Synechocystis*.[25,26] This response improves photosynthetic efficiency in different light conditions and protects cells from photodamage under high light conditions.[25,27] In particular, high light conditions have been shown to decrease PSI:PSII to a value near 1.[25] Although this transition must be affected by regulatory mechanisms,[26] it appears that mass and energy conservation alone have major impacts.

BIBLIOGRAPHY

[1] O. Kruse, J. Rupprecht, J. H. Mussgnug, G. C. Dismukes, and B. Hankamer, *Photochem. Photobiol. Sci*., 2005, **4**, 957-969.

[2] L. Cournac, G. Guedeney, G. Peltier and P. M. *Vignais, J. Bacteriol.*, 2004, **186**, 1737-1746.

[3] J. Allen, *Trends in Plant Science*, 2003, **8**, 15-19.

[4] C. C. Mann, *Science*, 1999, **283**, 314–316.

[5] J. Berg, J. Tymoczko, and L. Stryer, *Biochemistry*. 6th ed, 2007. New York: W. H. Freeman.

[6] M. Badger and T. Andrews, *Progress In Photosynthesis Research*, proceedings of the VIIth International Congress on Photosynthesis, 1986, 601-609. Dordrecht: Martinus Nijhoff Publishers.

[7] G. Price, M. Badger, F. Woodger, and B. Long, *J. Experimental Botany*, 2008, **59**, 1441-1461.

[8] F. Woodger, M. Badger, and G. Price, *Plant Physiology*, 2005, **139**, 1959-1969.

[9] S. Heinhorst, E. Williams, F. Cai, C. Murin, J. Shively, and G. Cannon, *J. Bacteriol.*, 2006, **188**, 8087-8094.

[10] S. Maeda, M. Badger, G. Price, *Molecular Microbiology*, 2002, **43**, 425-435.

[11] M. Badger, G. Price, B. Long, and F. Woodger, *J. Experimental Botany*, 2006, **57**, 249-265.

[12] V. Boichenko and P. Hoffmann, *Photosynthetica*, 1994, **30**, 527-552.

[13] V. Boichenko, E. Greenbaum and M. Seibert, *Photoconversion of solar energy* in *Molecular to Global Photosynthesis*, 2004, 397-452. London: Imperial College Press.

[14] A. Varma and B. Palsson, *J. theor. Biol.*, 1993, **165**, 477-502.

[15] R. Sharan, Lecture on *Analysis of Biological Networks: Constraint-Based Modeling of Metabolic Networks*, 2006.

[16] I. Rocha, J. Förster, and J. Nielsen, *Methods in Molecular Biology*, 2008, **416**, 409-431.

[17] J. Schellenberger and B. Palsson, *J. Bio. Chem.*, 2009, **284**, 5457-5461.

[18] J. Savinell and B. Palsson, *J. theor. Biol.*, 1992, **155**, 201-213 and 215-242.

[19] E. Almaas, B. Kovacs, T. Vicsek, Z. Oltvai, and A. Barabasi, *Nature*, 2004, **427**, 839-843.

[20] E. Almaas, Z. Oltvai, and A. Barabasi, *PLoS Comput. Biol.*, 2005, **1**, e68.

[21] I. Thiele, et al., *J. Biol. Chem.*, 2005, **280**, 11683-11695.

[22] D. Kaufman and R. Smith, *Operations Research*, 1998, Jan-Feb.

[23] S. Hong and C. Lee, *Biotech. and Bioprocess Engineering*, 2007, **12**, 165-173.

[24] A. Shastri and J. Morgan, *Biotech. Progress*, 2005, **21**, 1617-1626.

[25] H. Ozaki, M. Ikeuchi, T. Ogawa, H. Fukuzawa, and K. Sonoike, *Plant Cell Physiol*., 2007, **48**, 451-458.

[26] K. Sonoike, Y. Hihara, and M. Ikeuchi, *Plant Cell Physiol*., 2001, **42**, 379-384.

[27] W. Chow, A. Melis, and J. Anderson, *Proc. Natl. Acad. Sci.*, 1990, **87**, 7502-7506.

APPENDIX

Reaction/Enzyme Name	Reactants		Products
Glucose-6-Phosphate Isomerase	G6P	↔	F6P
Transhydrogenase	NADPH + NAD + Hint	↔	NADH + NADP + Hthyl
Biomass Production	1.191 G6P + 49.06 NADPH + 2.82 NAD + 1.197 PYR + 3.7270 AcCoA + 0.133 GAP + 1.205 3PG + 1.002 PEP + 0.715 E4P + 2.039 OAA + 1.233 αKG + 0.16 SUCCoA + 53.35 ATP	→	2.82 NADH + 49.06 NADP + 0.103 Ac + 0.683 FUM + 1.017 CO_2 + 0.16 SUC + 3.887 CoA + 53.35 ADP
Fructose Bisphosphate Aldolase	GAP + DHAP	↔	F16P
Phosphofructokinase	F6P + ATP	→	Hint + F16P + ADP
Fructose Bisphosphatase	H_2O + F16P	→	F6P
Triose Phosphate Isomerase	DHAP	↔	GAP
GAP Dehydrogenase	NADH + Hint + PDG	↔	NAD + GAP
NADP-Dependent GAP Dehydrogenase	NADPH + Hint + PDG	→	NADP + GAP
Pyruvate Kinase	Hint + PEP + ADP	→	PYR + ATP
Phosphenolpyruvate Synthase	H_2O + PYR + ATP	→	PEP + AMP
Ribulose Bisphosphate Carboxylase (RuBisCO)	H_2O + CO_2c + R15P	→	2 3PG
D-Ribulose-5-Phosphate 1-Phosphotransferase	ATP + RU5P	→	ADP
Ribulose-5-Phosphate 3-Epimerase	X5P	↔	RU5P
Ribose-5-Phosphate Isomerase	R5P	↔	RU5P
Ribose-5-Phosphate Isomerase	RU5P	→	R5P
Transketolase	GAP + S7P	↔	X5P
GAP E4P Combination	GAP + E4P	→	S7P
Transaldolase	GAP + S7P	↔	F6P + E4P
Transketolase2	F6P + GAP	↔	X5P + E4P
Pyruvate Dehydrogenase	NAD + PYR + CoA	→	NADH + CO_2 + AcCoA
Citrate Synthase	AcCoA + OAA	↔	Hint + CIT + CoA
Aconitase	CIT	↔	ICIT
Isocitrate Dehydrogenase	NADP + ICIT	↔	NADPH + CO_2 + αKG
Malate Dehydrogenase	NAD + MAL	↔	NADH + Hint + OAA
Phosphotransacetylase Acetate Kinase	AcCoA + ADP	↔	Ac + CoA + ATP
Phosphoenolpyruvate Carboxylase	H_2O + HCO_3^- + PEP	→	OAA
Succinyl-CoA Synthetase	SUC + CoA + ATP	↔	SUCCoA + ADP
Fumerase	MAL	↔	H_2O + FUM
Phosphoglycerate Kinase	PDG + ADP	↔	3PG + ATP
Phosphoglycerate Mutase	2PG	↔	3PG
Enolase	H_2O + PEP	↔	2PG
NADPH Hydrogen	NADPH + Hint	↔	NADP + H_2

Reaction/Enzyme Name	Reactants		Products
Electron Chain 1	2 Hint + PQ + 2 Fd	→	PQH2 + 2 Fd
Electron Chain 2 (PSII)	4 Photon + 2 H_2O	→	4 Hthyl + O2 + 4 e0
Electron Chain 3	4 e0	→	4 e1
Electron Chain 4	8 Hint + 2 PQ + 4 e1	→	4 Hthyl + 2 PQH2
Electron Chain 5 Cytochrome b6 f Complex	2 PQH2	→	4 Hthyl + 2 PQ + 4 e2
Electron Chain 6	4 e2 + 4 PC	→	4 PC
Electron Chain 6 Reverse	4 PC	→	4 e2 + 4 PC
Electron Chain 7	4 e3 + 4 PC	→	4 PC
Electron Chain 7 Reverse	4 PC	→	4 e3 + 4 PC
Electron Chain 8 (PSI)	4 Photon + 4 e3	→	4 e4
Electron Chain 9	4 e4 + 4 Fd	→	4 Fd
Ferrodoxin NADP Oxidoreductase	NADP + Hint + 2 Fd	→	NADPH + 2 Fd
NADPH Dehydrogenase 1_3	NADPH + 3 Hint + 2 H_2O + 2 CO_2 + PQ	→	NADP + 4 Hthyl + 2 HCO_3^- + PQH2
NADPH Dehydrogenase 1_4	NADPH + 3 Hint + H_2O + 2 CO_2 + 0.50 O2	→	NADP + 2 HCO_3^- + 4 Hex
NADPH Dehydrogenase 1_2	NADPH + 5 Hint + PQ	→	NADP + 4 Hthyl + PQH2
Bicarbonate Transporter (BCT-1)	ATP + HCO_3^-ex	→	HCO_3^- + ADP
Carbonic Anhydrase	Hint + HCO_3^-	→	H_2O + CO_2c
CO_2 Carboxysome Diffusion	CO_2c	→	CO_2
Bicarbonate Diffusion (SbtA/BicA)	HCO_3^-ex	→	HCO_3^-
Bicarbonate CO_2 Equilibria	H_2O + CO_2	↔	Hint + HCO_3^-
ATP Synthase	4.670 Hthyl + ADP	→	4.670 Hint + ATP
Maintenance Reaction	H_2O + ATP	→	Hint + ADP
Electron Sink 1	0.50 O2 + PQH2	→	H_2O + PQ
Electron Chain 10	4 e2 + 4 Cyt553	→	4 Cyt553
Electron Chain 10 Reverse	4 Cyt553	→	4 e2 + 4 Cyt553
Electron Chain 11	4 e3 + 4 Cyt553	→	4 Cyt553
Electron Chain 11 Reverse	4 Cyt553	→	4 e3 + 4 Cyt553
Electron Sink 2	2 Hthyl + 0.50 O2 + 2 Cyt553	→	H_2O + 2 Cyt553
Adenylate Kinase	2 ADP	↔	ATP + AMP
Electron Sink 4 (pseudocyclic)	NADPH + Hint + 0.50 O2	→	NADP + H_2O
NADPH Dehydrogenase 2	NADH + Hint + PQ	→	NAD + PQH2
α-Ketoglutarate Dehydrogenase	NADPH + SUCCoA	→	NADP + CO_2 + αKG + CoA
Sedoheptulose-1,7-Bisphosphate GAP-Lyase	S17BP	↔	DHAP + E4P
Sedoheptulose-1,7-Bisphosphate 1-Phosphohydrolase	S17BP	↔	H_2O + S7P