UNIVERSITY OF CALIFORNIA, SAN DIEGO

Kinetic Model Reconstruction of Phytoplankton Light-Dependent Reactions and Implementation Towards Membrane Constraints

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Bioengineering

by

Andrew Vasco Basilio

Committee in charge:

Professor Bernhard Palsson, Chair
Professor Christian Metallo
Professor Karsten Zengler

2017

The Thesis of Andrew Vasco Basilio is approved and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California, San Diego

2017

Dedication

To God first and foremost, for giving me guidance and direction, as well as the willingness to take risks and tackle challenges with enthusiasm.

To my entire family and all my friends, thank you for your continuous support and believing in me.

To Mom, thank you for the lessons you have taught me throughout my life and for being a role model in the field of engineering leadership. Your love has motivated me to keep pushing forward, to reach for the stars, to move quickly, and to not be afraid of failure.

To Dad, thank you for inspiring me to reach my highest potential. You're journey in obtaining an MS and PhD in Aerospace Engineering while working full-time as a project manager and being an excellent father has shown me that anything is possible. No obstacle is too difficult.

To my grandparents, thank you for always being by my side. I promise to keep progressing and make you all proud one day.

Epigraph

"You think that because you understand 'one' that you must therefore understand 'two' because one and one make two. But you forget that you must also understand 'and.'"

- Donella H. Meadows, Thinking in Systems: A Primer

Table of Contents

List of Figures

List of Tables

List of Abbreviations

MASS Toolbox: Mass-Action Stoichiometric Simulation Toolbox

PSII: Photosystem II

Cyt b6/f: cytochrome b6/f

PSI: Photosystem I

PFD: Photon Flux Density

PQ: Plastoquinone

PQH_2: Plastoquinol

PC: Plastocyanin

Fd: Ferredoxin

COBRApy: Constraint-Based Reconstruction and Analysis using Python

Acknowledgments

I would like to thank Professor Bernhard Palsson for his support as Chair of the committee. He has given me an opportunity to work in the lab, allowing me to dive into the field of systems biology and learn techniques and concepts associated with the field.

I would also like to thank Jared Broddrick for giving me the extensive guidance and feedback throughout the entire research process. His advice has been extremely valuable, especially when introducing me to the field of systems biology and explaining the nuances and intricacies of the biological mechanism within the light reactions of photosynthesis.

I would also like to thank Anand Sastry and Bin Du for their assistance whenever I needed help using Mathematica, installing MASS Toolbox onto my personal computer, and teaching me how to build a kinetic model.

I would also like the thank Dr. Laurence Yang for his guidance regarding the mathematics and concepts of systems biology and for giving me insight when I was learning about genome-scale modeling and COBRApy methods.

I would also like to thank Sarah, Erol, Edward, Erol, Billy, Hoang, and other lab members who have allowed me to discuss my ideas and helped me through the process.

I would also like to thank my friend Kimberly Apolonio for always being by my side throughout my entire final year of graduate school at UCSD and for motivating and empowering me throughout my research, the writing of this thesis, and beyond.

Abstract of Thesis

Kinetic Model Reconstruction of Phytoplankton Light-Dependent Reactions and Implementation Towards Membrane Constraints

by

Andrew Vasco Basilio

Master of Science in Bioengineering

University of California, San Diego, 2017

Professor Bernhard Palsson

Phytoplankton photosynthesis accounts for much of the oxygen present in the world, which can affect many different global systems. Additionally, understanding phytoplankton photosynthesis and other metabolic processes are important because of their potential to be engineered for human purposes. In this thesis, I constructed a computational kinetic model of the light-dependent photosynthetic reactions of phytoplankton using the MASS Toolbox, an in-house computational tool developed by the UCSD Systems Biology Research Group. This model was based upon the one constructed by Kroon and Thoms and my model simulations were validated against this model. The kinetic model was constructed based on the fact that the photosystem II : cytochrome b6/f : photosystem I ratio was 1:1:1. The objective, therefore, for the

second portion of the thesis was to determine thylakoid membrane crowding constraints and solve for the optimized ratio of photosystem II, cytochrome b6/f, and photosystem I. I calculated the crowding constraint equation using the model and incorporating parameters from literature, with hopes that it can be applied to the Phaeodactylum tricornutum genome-scale model in the future.

In the last portion of the thesis, actual physiological ratios of photosystem II, cytochrome b6/f, and photosystem I from the organism T. weissflogii were incorporated into the MASS Toolbox model of phytoplankton. I then simulated the model and observed the differences compared to the original model with the 1:1:1 protein ratio.

Chapter 1: Introduction

Organisms convert light energy to chemical energy in a process called photosynthesis. This phenomenon occurs in both marine and terrestrial environments, but aquatic photosynthesis contributes to approximately 45% of the total photosynthesis on earth[5]. Therefore, to study a topic that has global effects, it is imperative to conduct research on photosynthetic organisms within aquatic environments. While there is a vast number of different types of organisms within earth's oceans, ponds, and lakes, the focus of this thesis will specifically be on phytoplankton, which include diatoms.

Diatoms are organisms found throughout the world's oceans, accounting for approximately one-fifth of the oxygen produced on earth[1]. Because of their large impact upon the ecosystem, it is worth studying these organisms in depth. Diatoms are also amenable to a variety of genetic engineering methods, implicating their usefulness within the realm of biotechnology for socioeconomic efforts.

In addition to the ease of genetically engineering diatoms, Phaeodactylum tricornutum, a common diatom, is studied due to its simplicity in growth requirements and small, manageable genome[2]. Because of these characteristics, much work has been conducted to sequence, analyze, and reconstruct its genome along with associated reactions. Accurately describing its phenotype due to perturbations within the external environment through experimental procedures alone is extremely difficult, especially since Phaeodactylum tricornutum lives in an ecosystem that is highly dynamic[3]. Therefore, to assist with phenotypic characterization of this organism as a function of important environmental factors such as photon flux density, it is imperative to construct

a computational model to facilitate analysis of this organism. Manual curation of a computational model requires knowledge of the metabolic reactions within the organism, the connections among these reactions, and knowledge of various parameters such as the rate constants and initial conditions of each reaction. From here, systems biology methods can be applied to aid in discovery of new biological pathways or novel genetic engineering applications.

Studying the intricate interactions of every metabolite and reaction simultaneously is the main theme of systems biology. This approach is different from the experimental, reductionist methods used by molecular biologists throughout most of the 20th century. These methods are more specifically termed computational systems biology, a field that can be split into two main over-arching branches: knowledge-discovery through data-mining methods and simulation-based analysis[4]. The former consists of bioinformatics methods to make sense of complex data sets through pattern searching, while the latter focuses on creating predictions based upon established parameters, reaction properties, and reaction connectivity within an *in silico* environment[4]. Computational systems biology has greatly facilitated experimental efforts. For example, these methods can be used to quickly create hypotheses and predictions regarding effects of drugs, which can then be tested by experimentalists to create and discover new cures much quicker than a "trial-and-error" method.

Here, simulation-based tools were applied to phytoplankton in general as well as Phaeodactylum tricornutum. One aim focuses on manual curation of a kinetic model of phytoplankton in general while the other aim focuses on implementing membrane protein surface area constraints using mathematics which can eventually be applied to the

Phaeodactylum tricornutum genome-scale model to establish a more accurate model of flux.

Chapter 2: Manual Curation of a Computational Kinetic Model of Phytoplankton Photosynthetic Metabolism

2.1 Introduction: Overview of Photosynthetic Light Reactions

Photosynthesis can be divided into the light reactions and the dark reactions. The light reactions of photosynthesis involve capturing the energy of light in the form of a photon and are controlled by four main membrane proteins: photosystem II, cytochrome b6/f, photosystem I, and ATP synthase. This portion of photosynthesis only occurs when light is present[9]. Together these proteins direct electron movement originating from the energetic photon through the electron transport chain in photosynthetic organisms for the purpose of ATP and NADPH production.

For photosynthesis to begin, a photon of light needs to be absorbed and directed towards a reaction center so that this energy can eventually be used for charge separation, a step that is vital for the electron transport chain to proceed[5]. In diatoms, light energy is absorbed by chlorophyll pigment molecules and passed along towards a reaction center in a chain reaction procedure called exciton transfer[9]. For these energetic transfers to occur, it is necessary that molecular constituents of the entire complex are located close to each other. Throughout this process, over 90% of energy absorbed is focused onto the reaction center and partakes in photosynthesis[7,8]. Once the energy is transferred to the reaction center, the transport of an energetic electron through the electron transport chain begins, leading to a series of reactions.

The first major protein involved in diatom photosynthesis, as well as a majority of photosynthetic organisms, is Photosystem II (PSII). PSII structure is highly conserved

between distantly related species[10], consisting of multiple subunits that in total weigh about 700 KDa[11]. The two plastoquinone molecules, Q_a and Q_b, located within PSII play a major role in moving electrons through the complex. Q_a first receives an electron from pheophytin which in turn receives an electron from P680, a pigment found in the reaction center of PSII[9]. PQ, a free-floating plastoquinone molecule found in a pool within the lipid bilayer of the thylakoid membrane then attaches to PSII. Q_a then transfers this electron to the attached PQ molecule, which then becomes designated as Q_b. Q_b then becomes negatively charged, but only after it receives another electron, giving it a total negative charge of -2 so that the Q_b molecule gives electrons to plastoquinol (PQH_2), which is released from PSII into the lipid bilayer to travel to the cytochrome b6/f complex.

The cytochrome b6/f complex is the second main protein of the light reactions composed of two cytochrome b6/f molecules, a cytochrome f molecule, an iron-sulfur protein, and a subunit IV protein in which PQH_2 can dock after it leaves PSII. Cytochrome b6/f ultimately oxidizes the PQH_2 into PQ and undergoes a series of reactions, causing two H+ ions to be released into the lumen of the thylakoid[5]. Eventually, electrons are moved from the cytochrome b6/f complex to plastocyanin which in turn transfers its electrons to the P700 pigment in photosystem I (PSI)[13].

PSI activation by photons is similar to PSII in that a special type of chlorophyll at the reaction center is activated, which loses an electron to an electron acceptor. In the case of PSI, the P700 pigment at the reaction center loses an excited electron to A_0, another form of chlorophyll, which then passes the electron to phylloquinone (A_1), and ultimately to various Fe-S centers (F_x, F_{ab})[9]. These Fe-S centers then pass electrons to

ferredoxin (Fd) for the sake of making the intended product NADPH. The protons pumped into the thylakoid throughout this process form a proton gradient, providing the potential energy needed to form ATP from ADP and phosphate.

Although this process seems linear, it is actually an interconnected web of reactions running simultaneously, making it almost impossible to study the exact dynamics of the network using conventional molecular biology techniques. In these situations, it is highly advantageous to construct a computational kinetic model based upon first principles of chemistry so that a holistic view of network dynamics, flux, and pooling is revealed.

Constructing a kinetic model requires prerequisite knowledge. First, all metabolites and involved reactions must be known and entered into the model to the best of one's ability. A missing metabolite or an incorrect connection can drastically affect the results of the simulation. Next, all the rate constants need to be known as well as the initial conditions. If all these parameters are known for every reaction, a unique and specific solution is produced. The paper "From Electron to Biomass: A Mechanistic Model to Describe Phytoplankton Photosynthesis and Steady-State Growth Rates" by Bernd M. A. Kroon and Silke Thoms provided the metabolites, connections, rate constants, and equilibrium constants while in-house Mass-Action Stoichiometric Simulation (MASS) Toolbox methods were implemented to create a computational model based on these parameters.

The MASS Toolbox, a tool developed in the UCSD Systems Biology Research Group, is a computational method based on Mathematica software which is specifically catered towards construction and analysis of systems of chemical reactions. These

chemical reactions are interconnected in such a way that the products of one reaction can become the reactants of several other reactions, ultimately leading to a web-like connection throughout the entire system. Fascinatingly, all these reactions and their interconnections can be represented within a single mathematical entity: the stoichiometric matrix. The stoichiometric matrix houses the coefficients of each compound within a chemical reaction, in which each column represents a reaction in the system and each row represents a molecule that acts as a reactant or product depending on which column its element populates[14]. Furthermore, all of the different possible solutions of flux throughout the system, denoted as the feasible solution, can be determined by following some first-principles of physics: the conservation of mass, charge, and energy. Determination of this feasible set can be done by using the equation

$$S \bullet v = 0 \qquad (1.1)$$

where S is the stoichiometric matrix and v is the vector of fluxes. The product of these two matrices must equal zero in order to satisfy conservation of mass. The different possible values of the v matrix that satisfy this equation represent the feasible solution.

The MASS Toolbox interface is exactly like a Mathematica notebook. Once all of the reactions and parameters are inputted into the notebook, MASS Toolbox automatically forms the stoichiometric matrix of the entire system, allowing for analysis. In the first portion of this thesis, the photosynthetic network of the light reactions was constructed, verified, and validated against the publication by Kroon and Thoms that demonstrated their own computational network of the same system.

2.2 Results

2.2.1 Reconstruction Overview and Photosystem II Reconstruction using MASS Toolbox

The network was constructed by focusing on one protein complex at a time, starting with PSII. Reactions were constructed based on the schematic in the diagram below by Kroon and Thoms:

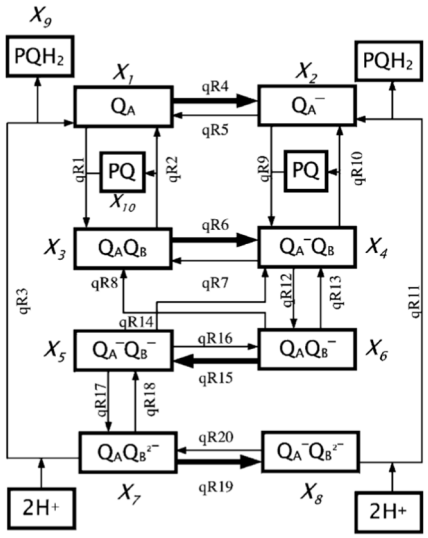

Figure 2.1: Network of PSII Electron Flow and Associated Reactions (Kroon and Thoms). Species encased in rectangles signify metabolites, thin arrows represent connections, and bold arrows represent steps where irradiance energy is inputted.

Construction of the PSII model was based on MASS Toolbox syntax shown in the figure below:

```
rxns = {

  (*Photosystem II reactions*)

  reaction["qr1", {metabolite["Qa"], metabolite["PQ"]},
   {transform[complex[metabolite["Qa"], metabolite["Qb"]]]}, {1, 1, 1}, False],
  reaction["qr2", {transform[complex[metabolite["Qa"], metabolite["Qb"]]]},
   {metabolite["PQ"], metabolite["Qa"]}, {1, 1, 1}, False],
```

Figure 2.2: MASS Toolbox Syntax for Inputting Reactions. The syntax shows the input required to form the qR1 and qR2 reactions only. The reaction name, metabolites, and associated stoichiometric coefficients are entered. Entering "False" means the reaction is non-reversible.

The output of the entered syntax can be checked by running the code and comparing the result to an actual metabolic model as shown below:

Out[3]//TableForm=

$$PQ + Qa \xrightarrow{qr1} Qa\_Qb$$

$$Qa\_Qb \xrightarrow{qr2} PQ + Qa$$

Figure 2.3: Output of MASS Toolbox Reaction Input for qR1 and qR2 for PSII

Next, rate constants of PSII were entered into the model. The rate of a chemical reaction is dependent on its rate constant as well as concentration of its products. According to "A Generalized Statement of the Law of Mass Action" by James K. Baird, if a chemical reaction can be written in the general form as:

$$v_1[1] + v_2[2] <=> v_3[3] + v_4[4] \tag{1.2}$$

then it's net reaction rate can be written as:

$$r = k_f c_1^{v_1} c_2^{v_2} - k_r c_3^{v_3} c_4^{v_4} \tag{1.3}$$

where v_1-v_4 represents the stoichiometric coefficient of the compound, [1]-[4] represents the unique compound, r represents the reaction rate, c represents the concentrations of the specific product or reactant, k_f represents the forward rate constant, and k_r represents the reverse rate constant[16]. Since most of the reactions in PSII, cyt b6/f, and PSI are modeled as forward reactions, the focus of the rate equation will be on the quantity related to the forward reaction. It shows that the rate of reaction is directly proportional to the concentration of the reactants, their stoichiometric coefficients, and the forward rate constant. Because the concentration of each reactant and flux of each reaction is dependent on time, the entire system is dynamic, forming a feedback loop where reaction rates change until flux and concentration steady-state solutions are produced. These steady-state solutions can then be analyzed in further research efforts.

A list of rate constants for most reactions involved in the network are shown in the table below:

Table 2.1: Rate Constants for Reactions in Kinetic Model (Kroon and Thoms). Rate constants in this table describe most reactions except for reactions qR4, qR6, qR15, and qR20 in PSII and reactions R1a, R1b, R1c, R1d in PSI. These rate constants depend on photon flux density, chlorophyll surface area, and various biophysical constants.

Rate constant	Unit	Value	Reactions	Process
bk1	$m\,s^{-1}$	5	bR1, bR4	Oxidation of b6/f by PC^+
bk2	$m\,s^{-1}$	10	bR2, bR5	Re-oxidation of PQH_2 by b6/f
bk3	$m\,s^{-1}$	0.99	bR3, bR6	Lumenal proton release by b6/f
bk4	$m\,s^{-1}$	1	bR7	Stromal proton binding by b6/f
bk5	$m\,s^{-1}$	5	bR8	Reduction of PQ at Q_i site of b6/f
bk6	$m\,s^{-1}$	2	bR9, bR12	Release of Fd from b6/f
kabE	$m\,s^{-1}$	1.733	qR17	Electron transfer between Q_A and Q_B
kabO	$m\,s^{-1}$	4.621	qR12	Electron transfer between Q_A and Q_B
kav	$m\,s^{-1}$	0.34	qR2, qR10	Release of PQ from PSII
kbackA	$m\,s^{-1}$	2.567×10^{-4}	qR5, qR7, qR16, qR20	Charge recombination in PSII
kbackB	$m\,s^{-1}$	2.773×10^{-5}	qR8, qR14	Charge recombination in PSII
kbaE	$m\,s^{-1}$	0.0347	qR18	Electron transfer between Q_A and Q_B
kbaO	$m\,s^{-1}$	0.2718	qR13	Electron transfer between Q_A and Q_B
kc	$m\,s^{-1}$	0.03	bR10, bR11	Docking of Fd^- to b6/f
k_{P700}	$m\,s^{-1}$	6.9	R4	PSI reduction of PC
kpqh2	$m\,s^{-1}$	2	qR3, qR11	Release of PQH_2 from PSII
k_{PSI}	$m\,s^{-1}$	0.2	R2a-d	Reduction of Fd by PSI
kterm	$m\,s^{-1}$	0.12	R3	Terminal Fd re-oxidation
kva	$m\,s^{-1}$	0.86	qR1, qR9	Docking of PQ to PSII

These rate constants are then entered in the MASS Toolbox interface as shown below:

```
In[3074]:= TableForm[
            param = {
              (*units are milliseconds^-1*)
              k["br1"] → 5,
              k["br4"] → 5,
              k["br2"] → 10,
              k["br5"] → 10,
              k["br3"] → 0.99,
              k["br6"] → 0.99,
              k["br7"] → 1,
              k["br8"] → 5,
              k["br9"] → 2,
              k["br12"] → 2,
              k["qr17"] → 1.733,
              k["qr12"] → 4.621,
              k["qr2"] → 0.34,
              k["qr10"] → 0.34,
              k["qr5"] → 0.0002567,
              k["qr7"] -> 0.0002567,
              k["qr16"] -> 0.0002567,
              k["qr20"] → 0.0002567,
              k["qr8"] → 0.00002773,
              k["qr14"] -> 0.00002773,
              k["qr18"] → 0.0347,
              k["qr13"] → 0.2718,
              k["br10"] → 0, (* If cyclic e- transfer present around PSI, then k["br10"]=0.03. If closed, then = 0 *)
              k["br11"] → 0, (* If cyclic e- transfer present around PSI, then k["br11"]=0.03. If closed, then = 0 *)
              k["R4"] → 6.9,
```

Figure 2.4: Rate Constants in MASS Toolbox Interface. Only a portion of the rate constants entered into MASS Toolbox are shown. The brackets contain the reaction name.

Upon examination of figure 2.1, it is important to note that the reversible reactions (qR5, qR7, qR16, and qR20) pertaining to Q_a oxidation and reduction are considered and incorporated into the model. While these reversible reactions follow the principle of mass action, the forward reactions follow a slightly different rule. In the forward reactions in PSII designated by qR4, qR6, qR16, and qR20, the rate law is dependent upon the photon flux density, number of chlorophyll surrounding PSII, the optical surface area capable of trapping photons, and various biophysical constants. These unique reactions are designated as thick arrows in Figure 2.1.

During reactions qR4, qR6, qR16, qR20, a photon hits a chlorophyll antenna molecule until the energy is transferred to the reaction center[17]. From there, the energy is transferred to a unique chlorophyll called P680 within the reaction center, which then transfers the energy to pheophytin and ultimately Q_a, a plastoquinone molecule[17]. This causes Q_a to become negatively charged as shown in Figure 2.1 in boxes X_2, X_4, X_5, X_8. In

reactions qR12 and qR17, the negative charge from Q_a is transferred onto another plastoquinone molecule called Q_b. It is important to note the Q_a and Q_b are similar molecules when isolated, but are structurally different when observed in vivo[5]. For an electron to be transferred from Q_a to Q_b, a free-floating plastoquinone designated PQ in the lipid bilayer must bind to a pocket on the protein containing the reaction center[5]. PQ is found in a pool within the lipid bilayer and once it binds to this pocket and obtains electrons from Q_a, it is called Q_b. This explains the forward reactions qR1 and qR9 and their associated reverse reactions qR2 and qR10, all of which incorporate PQ as a metabolite. As one traces the path of electron movement, Q_b actually obtains two reduced equivalents, ultimately forming Q_b^{2-} in a process call the two-electron gate[5]. Q_b^{2-} can potentially become reoxidized, hampering forward electron movement. Therefore, to prevent this reverse reaction from happening, two hydrogen ions become abstracted from the stroma and attach to Q_b^{2-} ultimately forming plastoquinol[5]. This molecule is abbreviated as PQH_2 and detaches itself from the binding pocket once the two hydrogens attach themselves[5]. This molecule then diffuses through the thylakoid lipid bilayer due to its hydrophobicity until it encounters the cytochrome b6/f complex, which will be discussed in further detail.

A brief overview of the electron transfer pathway and its construction has been presented, but it is imperative to dive into greater detail when discussing reactions qR4, qR6, qR15, and qR20. The rate laws of these reactions are not simply due to the rate constant and concentration of the reactants; many calculations needed to be performed using the MASS Toolbox to ensure that the rate laws for these reactions were properly implemented. Ultimately, the rate laws of these four equations in PSII are as follows[19]:

$$qR4 = Exc_{II}\Phi_{PSII,\,1} \tag{1.4}$$

$$qR6 = Exc_{II}\Phi_{PSII,\,3} \tag{1.5}$$

$$qR15 = Exc_{II}\Phi_{PSII,\,6} \tag{1.6}$$

$$qR19 = Exc_{II}\Phi_{PSII,\,7} \tag{1.7}$$

The right side of each of these equations can furthermore be broken down as follows[19]:

$$Exc_{II} = (\sigma_{PSII})(PFD)(cf) \tag{1.8}$$

$$\Phi_{PSII,\,i} = \frac{X_i}{1+J\sum_i X_i} \tag{1.9}$$

where PFD represents the photon flux density in units of (μmol of quanta)(m^{-2})(s^{-1}), cf represents the unitless conversion factor of 6.02 x 10^{14}, X_i represents the concentration of a specific open PSII unit when referring to figure 2.1, A is unitless biophysical parameter, and J is a unitless biophysical parameter. Determining the values of each of these variables will be explained in subsequent sections in greater detail.

First, in order to solve for Exc_{II}, the antenna sizes of each reaction center needs to be calculated. This can be done by determining the number of chlorophyll a and chlorophyll b surrounding each PSII protein. In the paper by Kroon and Thoms, they present different scenarios in which different numbers of chlorophyll a and b are presented[19]. These values will be very important for analysis and validation of the model and are shown in the following table:

Table 2.2: Various Antenna Scenarios (Kroon and Thoms). The figure above shows different scenarios of different numbers of chlorophyll that surround each reaction center of PSII and PSI. These values will be important for validation and verification of the constructed kinetic model.

	Number of proteins per PSII			chl *a/b* molecules per PSII			Number of proteins per PSI			chl *a/b* molecules per PSI		
Scenario	CP43 + 47	Lhcb3 to Lhcb6	Trimers of Lhcb1/b2	chl *a*	chl *b*	chl *a* + *b*	Core	Lhca1 to Lhca4	Trimers of Lhcb1/b2	chl *a*	chl *b*	chl *a* + *b*
1	1	4	4	261	158	419	90	8	2	254	85	339
2	1	4	3.5	240	143	383	90	8	1	212	54	266
3	1	4	3	219	127	346	90	8	0	170	23	193
4	1	4	2.5	198	112	310	90	7	0	160	20	180
5	1	4	2	177	96	273	90	6	0	150	17	167
6	1	4	1.5	156	81	237	90	5	0	140	14	154
7	1	4	1	135	65	200	90	4	0	130	11	141
8	1	4	0.5	114	50	164	90	3	0	120	9	129
9	1	4	0	93	34	127	90	2	0	110	6	116

Nine scenarios for PSII and PSI were chosen with varying amounts of peripheral antennae sizes as indicated by the numbers of chl *a*, chl *b* and chl *(a + b)*.

Knowing the number of antenna chlorophyll surrounding each reaction center allows for determination of the total optical cross-sectional area per protein, given that the optical cross sectional area of chlorophyll a is 8.6749 x 10^{-21} m^2 and of chlorophyll b is 9.1222 x 10^{-21} m^2 [19]. Therefore, the total cross-sectional area of the optical surface area can easily be determined as shown in the following equation[19]:

$$\sigma_{PSII} = \sigma_{chl\ a} N_{aII} + \sigma_{chl\ b} N_{bII} \tag{1.10}$$

in which σ_{PSII} represents the total optical surface area of PSII, $\sigma_{chl\ a}$ represents the optical cross-sectional surface area of a single chlorophyll a molecule, $\sigma_{chl\ b}$ represents the optical cross-sectional surface area of a single chlorophyll b molecules, N_{aII} represents the total number of chlorophyll a molecules surrounding a single PSII reaction center, and N_{bII} represents the total number of chlorophyll b molecules surrounding a single PSII reaction center. This equation was entered into the MASS Toolbox interface in the following syntax as shown in the figure below:

```
sigmachla = 8.6749 * (10^-21);(* units=m^2 per chlorophyll a *)
sigmachlb = 9.1222 * (10^-21);(* units=m^2 per chlorophyll b *)
NaII = 114(* Number of chlorophyll A per PSII. *);
NbII = 50(* Number of Chlorophyll B per PSII. *);
sigmaPSII = sigmachla * NaII + sigmachlb * NbII (*Entire optical surface area per PSII that can absorb photons*);
```

Figure 2.5: MASS Toolbox Input for Solving Optical Surface Area. This figure shows a small portion of the syntax used for calculating total optical surface area of PSII. The values shown for N_{aII} and N_{bII} use values from antenna scenario 8 as shown in table 2.2 as an example.

Next, $\Phi_{PSII,\,I}$ needs to be determined in order to solve for the rate law. It is important to keep in mind the important biophysical constants listed in the table below:

Table 2.3: List of Relevant Biophysical Constants (Kroon and Thoms). These values will be used for important calculations regarding the rate laws for reactions qR4, qR6, qR15, qR19.

Constant	Value	Unit	Definition
k_t^{ox}	$\frac{540}{N_{aII}+N_{bII}}$	$n\,s^{-1}$	Trapping constant for open PSII centers. Depends on the number of chl *a* and chl *b* molecules in PSII (Table 2)
k_2	2.3	$n\,s^{-1}$	Reduction of Q_A
k_{-1}^{ox}	0.3	$n\,s^{-1}$	Charge recombination in PSII units with open centers
k_d^{ox}	0.001	$n\,s^{-1}$	Non-radiative losses of the radical pair in open centers
k_t^{red}	$\frac{84.6}{N_{aII}+N_{bII}}$	$n\,s^{-1}$	Trapping constant for closed PSII centers. Depends on the number of chl *a* and chl *b* molecules in PSII (Table 2)
k_{-1}^{red}	0.34	$n\,s^{-1}$	Charge recombination in PSII units with closed centers
k_d^{red}	0.99	$n\,s^{-1}$	Non-radiative losses of the radical pair in closed centers
k_{UU}	2.6	$n\,s^{-1}$	Interunit exciton transfer
k_{loss}	0.45	$n\,s^{-1}$	Exciton losses
k_{rad}	0.084	$n\,s^{-1}$	Radiative antennae decay

These values are then plugged into the equations that represent the biophysical parameters A and J. These equations are as follows (Kroon and Thoms):

$$A = \frac{\alpha_p(k_{UU}+\beta+k_{loss})}{(\beta+k_{loss})(k_{UU}+\alpha_p+\alpha_d+k_{loss}} \tag{1.11}$$

and

$$J = \frac{k_{UU}(\alpha_p+\alpha_d-\beta)}{(\beta+k_{loss})(k_{UU}+\alpha_p+\alpha_d+k_{loss})} \tag{1.12}$$

The biophysical constants as well as their implementation in the calculation of various parameters were entered into the MASS Toolbox interface as shown in the following figure:

```
In[3096]:= (*Biophysical Rate Constants for Analysis*)

krad = 0.084 (*units: ns^-1*);
kloss = 0.45 (*units: ns^-1*);
kUU = 2.6 (*units: ns^-1*);
kdred = 0.99 (*units: ns^-1*);
kredinv = 0.34 (*units: ns^-1*);
ktred = 84.6 / (NaII + NbII) (*units: ns^-1*);
kdox = 0.001 (*units: ns^-1*);
koxinv = 0.3 (*units: ns^-1*);
ktwo = 2.3 (*units: ns^-1*);
ktox = 540 / (NaII + NbII) (*units: ns^-1*);
beta = (ktred * kdred) / (kredinv + kdred) (*units: ns^-1*);
Bparameter = krad / (beta + kloss) (*units: none*);
alphap = (ktox * ktwo) / (koxinv + ktwo + kdox) (*units: ns^-1*);
alphad = (ktox * kdox) / (koxinv + ktwo + kdox) (*units: ns^-1*);
Cparameter = (-krad * (alphap + alphad - beta)) / ((beta + kloss) * (kUU + alphap + alphad + kloss)) (*the parameter C*);
Aparameter = (alphap * (kUU + beta + kloss)) / ((beta + kloss) * (kUU + alphap + alphad + kloss)) (*the parameter A*);
Jparameter = (kUU * (alphap + alphad - beta)) / ((beta + kloss) * (kUU + alphap + alphad + kloss)) (*the parameter J*);
```

Figure 2.6: Biophysical Constants and Parameters. These values will be used to calculate the rate laws for reactions qR4, qR6, qR15, and qR19.

Knowing parameters A and J allows for calculation of equation (1.9), although its derivation is fairly difficult[19].

At this point, the expression solving for reaction rates for qR4, qR6, qR15, and qR19 are known, but simply plugging in values of concentrations of open Q_a centers is inadequate because the concentrations of these species changes over time; the system is dynamic and therefore needs to have this concept considered when coded into MASS Toolbox. The syntax used in MASS Toolbox considers the dynamics of the system and changing concentrations of open Q_a states is shown in the following syntax:

```
In[3081]:= setCustomRateLaws[model,
 {
  "qr4" →
   (sigmaPSII * parameter["PFD"] * cf * Aparameter *
     (metabolite["Qa"][t] /
       (1 + Jparameter * (metabolite["Qa"][t] + transform[complex[metabolite["Qa"], metabolite["Qb"]]][t] +
            transform[complex[metabolite["Qa"], metabolite["Qb-"]]][t] + transform[complex[metabolite["Qa"], metabolite["Qb2-"]]][t])))),

  "qr6" →
   (sigmaPSII * parameter["PFD"] * cf * Aparameter * (transform[complex[metabolite["Qa"], metabolite["Qb"]]])[t] /
      (1 + Jparameter * (metabolite["Qa"][t] + transform[complex[metabolite["Qa"], metabolite["Qb"]]][t] +
           transform[complex[metabolite["Qa"], metabolite["Qb-"]]][t] + transform[complex[metabolite["Qa"], metabolite["Qb2-"]]][t]))),

  "qr15" →
   (sigmaPSII * parameter["PFD"] * cf * Aparameter * (transform[complex[metabolite["Qa"], metabolite["Qb-"]]])[t] /
      (1 + Jparameter * (metabolite["Qa"][t] + transform[complex[metabolite["Qa"], metabolite["Qb"]]][t] +
           transform[complex[metabolite["Qa"], metabolite["Qb-"]]][t] + transform[complex[metabolite["Qa"], metabolite["Qb2-"]]][t]))),

  "qr19" →
   (sigmaPSII * parameter["PFD"] * cf * Aparameter * (transform[complex[metabolite["Qa"], metabolite["Qb2-"]]])[t] /
      (1 + Jparameter * (metabolite["Qa"][t] + transform[complex[metabolite["Qa"], metabolite["Qb"]]][t] +
           transform[complex[metabolite["Qa"], metabolite["Qb-"]]][t] + transform[complex[metabolite["Qa"], metabolite["Qb2-"]]][t]))) (*,
```

Figure 2.7: Rate Law Modification for qR4, qR6, qR15, and qR19 on MASS Toolbox Interface. Shown here is the syntax used to modify the rate laws so that time-dependent concentration changes in open Q_a centers plug themselves back into the reaction rate equations.

When using MASS Toolbox, the default rate law equation is based on properties of mass-action kinetics. However, the MASS Toolbox command "setCustomRateLaws" allows for modification of the rate law, whether the desired input is a simple constant or something more complex. The syntax shown in figure 2.10 represents the direct input of equations (1.4)-(1.7) into the model.

To ensure that the rate laws of qR4, qR6, qR15, and qR19 were modified when compared to the other standard reactions in PSII, the "model" command was run on MASS Toolbox to check the rate laws. The resulting output is shown in the following figure:

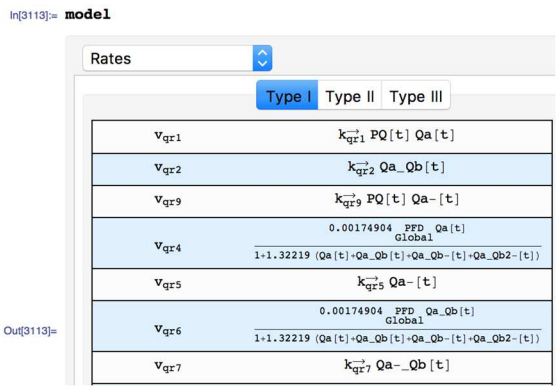

Figure 2.8: Comparison Between Modified Rate Laws and Default Rate Laws. Only a portion of PSII reactions are shown, but the difference in rate laws is noticeable after modification. Reactions qR4 and qR6 depend on parameters A and J in equations (1.11) and (1.12) as well as photon flux density.

2.2.2 Cytochrome b6/f Reconstruction using MASS Toolbox

Construction of the cytochrome b6/f (cyt b6/f) electron flow pathway kinetic model was completed after construction of the PSII kinetic model. The construction performed on MASS Toolbox was based on a figure from the publication by Kroon and Thoms shown in the following figure:

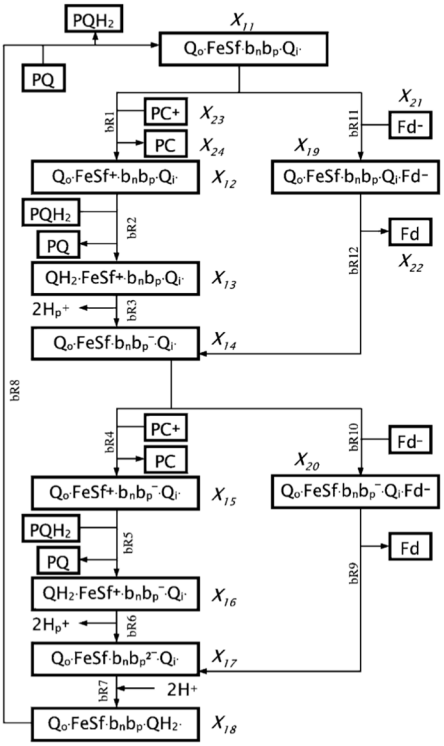

Figure 2.9: Network of Cytochrome b6/f Electron Flow and Associated Reactions (Kroon and Thoms). This complex is modeled in between PSII and PSI.

From this figure, it is simple to see how this complex is connected to PSII that precedes it. PQH_2, which is metabolite X_9 in PSII shown in figure 2.1, connects to cytochrome b6/f through reactions bR2, bR5, and bR8. Construction of this complex using the MASS Toolbox was similarly performed as in construction of PSII. The figure below shows a portion of the syntax used in the MASS Toolbox interface to construct the model:

```
(* Cytochrome b6/f electron flow reactions *)
reaction["br1",
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp"], metabolite["bn"], metabolite["Qi"]]],
  metabolite["PC+"]},
 {metabolite["PC"], transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f+"], metabolite["bp"], metabolite["bn"],
    metabolite["Qi"]]]}, {1, 1, 1, 1}, False],
reaction["br2",
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f+"], metabolite["bp"], metabolite["bn"], metabolite["Qi"]]],
  metabolite["PQH2"]},
 {metabolite["PQ"], transform[complex[metabolite["QH2"], metabolite["FeS"], metabolite["f+"], metabolite["bp"], metabolite["bn"],
    metabolite["Qi"]]]}, {1, 1, 1, 1}, False],
(*br3 BE CAREFUL WITH H+ INSERTION. INPUT INITIAL CONDITION*)
reaction["br3",
 {transform[complex[metabolite["QH2"], metabolite["FeS"], metabolite["f+"], metabolite["bp"], metabolite["bn"], metabolite["Qi"]]]},
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]]},
 {1, 1}, False],
reaction["br4",
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]],
  metabolite["PC+"]},
 {metabolite["PC"], transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f+"], metabolite["bp-"], metabolite["bn"],
    metabolite["Qi"]]]}, {1, 1, 1, 1}, False],
reaction["br5",
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f+"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]],
  metabolite["PQH2"]},
 {metabolite["PQ"], transform[complex[metabolite["QH2"], metabolite["FeS"], metabolite["f+"], metabolite["bp-"], metabolite["bn"],
    metabolite["Qi"]]]}, {1, 1, 1, 1}, False],
(*br6 BE CAREFUL WITH H+ INSERTION. INPUT INITIAL CONDITION*)
reaction["br6",
 {transform[complex[metabolite["QH2"], metabolite["FeS"], metabolite["f+"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]]}
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp2-"], metabolite["bn"], metabolite["Qi"]]]},
 {1, 1}, False],
(*br7 BE CAREFUL WITH H+ INSERTION. INPUT INITIAL CONDITION*)
reaction["br7",
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp2-"], metabolite["bn"], metabolite["Qi"]]]},
 {transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp"], metabolite["bn"], metabolite["QH2"]]]},
 {1, 1}, False],
```

Figure 2.10: Syntax for Inputting Cyt b6/f electron flow pathways into MASS Toolbox. Only a portion of the code is shown. Unlike several reactions in PSII, it is important to note that all reactions here follow the normal mass-action rules regarding reaction rates. No photon energy input is required.

Once the model was built, continuity and connection to PSII was verified. This was done by running the "model" command on the MASS Toolbox interface and selecting the "pathway map" option on the output. If all connections are correct, the output should give an image of a connected network. If not, the discontinuity will be apparent when the "model" command is run and the pathway map is checked. The following is a set of reactions in which connectivity is present. It will serve as a simple example, since the actual model displays a crowded network due to the large number of metabolites and reactions:

$$A + B \leftrightarrow C \tag{1.13}$$

$$C + D \leftrightarrow E \tag{1.14}$$

$$E + B \leftrightarrow G \tag{1.15}$$

The resulting pathway map is shown below:

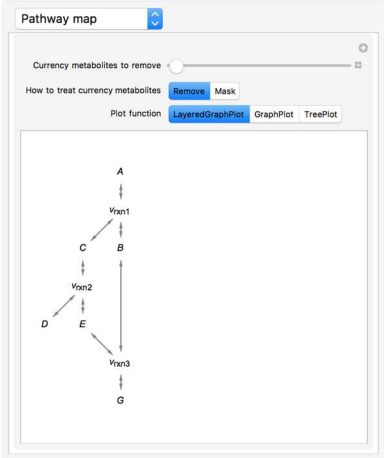

Figure 2.11: Pathway Map of Simple Set of Reactions Displaying Connectivity. All three reactions have at least one reactant or product that is present in another reaction, forming one complete network.

Shown below is a series of reactions in which connectivity is not present.

$$A + B \rightarrow C \tag{1.16}$$

$$C + D \rightarrow E \tag{1.17}$$

$$F + G \rightarrow H \tag{1.18}$$

The resulting pathway map from reactions (1.16) - (1.18) is therefore shown below:

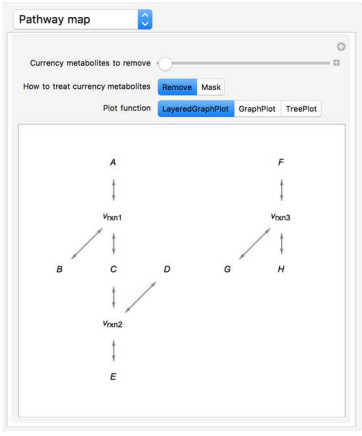

Figure 2.12: Pathway Map of Simple Set of Reactions Displaying No Connectivity. Reaction (1.18) is disconnected from reaction (1.16) and (1.17).

Using the methods described above, the connectivity between PSII and cyt b6/f was verified by ensuring no physical gaps existed in the pathway output.

In terms of the biochemistry, reactions bR9, bR10, bR11, and bR12 in cyt b6/f refer to cyclic electron flow between cyt b6/f and PSI. This cyclic electron flow plays a major role in determination of proton yield as a function of irradiance, which plays a significant role in validating the model and will be discussed later. All other reactions in cyt b6/f incorporate a linear electron flow pathway. Electrons are transferred throughout the entire cyt b6/f complex, designated as $Q_oFeSfb_nb_pQ_i$[19]. Q_o and Q_i are sites for quinone oxidation and reduction from when PQH_2 arrives from the PSII complex[19]. FeSf designates the Rieske iron-sulfur center, a protein that assists in electron transfer through the reduction or oxidation of iron within the iron-sulfur complex[9]. Reactions bR3 and bR6 show the movement of H+ ions into the stroma, and the flux of these reactions play a large role in the determination of H+ ions produced per electron that travels through the electron transport chain. The construction of the PSI complex was started after construction of the cyt b6/f complex was complete.

2.2.3 Photosystem I Reconstruction Using MASS Toolbox

PSI is the third protein complex involved in the electron transport chain. This complex is connected to the cyt b6/f complex through the metabolite plastocyanin (PC), which is located on the lumen side of the membrane of the thylakoid disc[20]. Plastocyanin becomes reduced and travels through the lumen side of the membrane until it contacts PSI. From there, plastocyanin loses the electron it was carrying and becomes oxidized by P700, a special type of chlorophyll in PSI[9]. P700 is a special type of chlorophyll with a reduction potential of -1.2 V, making it a very strong reducing agent[21]. Because of P700's strong

reduction potential, it donates an electron to another chlorophyll a molecule, designated A_o that acts as a primary acceptor, which then passes its electron to phylloquinone, designated A_1 and acts as a secondary acceptor[22]. The reduced phylloquinone then passes on its electron, reducing an iron-sulfur complex designated F_x[22]. Electrons are then transferred from F_x to the final iron-sulfur complexes within PSI designated F_{AB}, and finally to ferredoxin, the final electron carrier that reduces $NADP^+$ into NADPH, the final product of the light reactions[9]. It is important to realize that ferredoxin can be reduced during any time that an electron is present in the final iron-sulfur F_{AB} complex. Additionally, A_o can be reduced at any time, whether or not either A_1, F_x, or F_{AB} is reduced. The diagram displaying the connectivity by Kroon and Thoms is shown below:

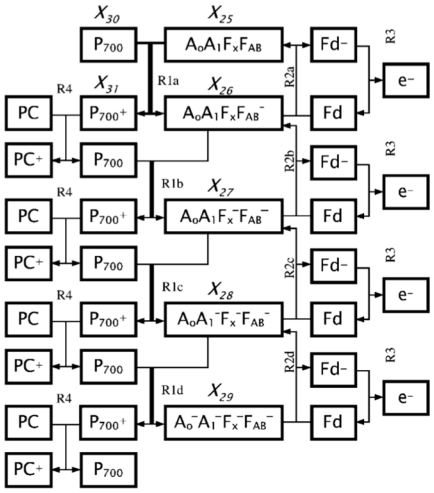

Figure 2.13: Network of Photosystem I Electron Flow and Associated Reactions (Kroon and Thoms). Electrons flow from through plastocyanin, P700, the phylloquinone and iron-sulfur complexes, and ferredoxin to reduce $NADP^+$ into NADPH.

The photosystem I electron flow network was constructed just like the photosystem II and cytochrome b6/f complex before it. Additionally, connectivity among all three of the

complexes was verified using the same method described in Figure 2.13. The syntax used to construct the photosystem I electron flow pathways is shown below:

```
(* Photosystem I electron flow reactions *)
reaction["R1a", {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx"], metabolite["Fab"]]], metabolite["P700"]},
 {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx"], metabolite["Fab-"]]], metabolite["P700+"]}, {1, 1, 1, 1}, False],
reaction["R1b", {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx"], metabolite["Fab-"]]], metabolite["P700"]},
 {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["P700+"]}, {1, 1, 1, 1}, False],
reaction["R1c", {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["P700"]},
 {transform[complex[metabolite["Ao"], metabolite["A1-"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["P700+"]}, {1, 1, 1, 1}, False],
reaction["R1d", {transform[complex[metabolite["Ao"], metabolite["A1-"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["P700"]},
 {transform[complex[metabolite["Ao-"], metabolite["A1-"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["P700+"]},
 {1, 1, 1, 1}, False],
reaction["R4", {metabolite["PC"], metabolite["P700+"]}, {metabolite["PC+"], metabolite["P700"]}, {1, 1, 1, 1}, False],
reaction["R2d", {transform[complex[metabolite["Ao-"], metabolite["A1-"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["Fd"]},
 {transform[complex[metabolite["Ao"], metabolite["A1-"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["Fd-"]}, {1, 1, 1, 1}, False],
reaction["R2c", {transform[complex[metabolite["Ao"], metabolite["A1-"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["Fd"]},
 {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["Fd-"]}, {1, 1, 1, 1}, False],
reaction["R2b", {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx-"], metabolite["Fab-"]]], metabolite["Fd"]},
 {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx"], metabolite["Fab-"]]], metabolite["Fd-"]}, {1, 1, 1, 1}, False],
reaction["R2a", {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx"], metabolite["Fab-"]]], metabolite["Fd"]},
 {transform[complex[metabolite["Ao"], metabolite["A1"], metabolite["Fx"], metabolite["Fab"]]], metabolite["Fd-"]}, {1, 1, 1, 1}, False],
reaction["R3", {metabolite["Fd-"]}, {metabolite["Fd"], metabolite["e-dark"]}, {1, 1, 1}, False]
```

Figure 2.14: Syntax for inputting Photosystem I electron flow pathways into MASS Toolbox.

Once all three complexes were constructed and the connections among them were verified, the next step was to validate the model.

Validation of the system was done by running simulations of the model and comparing the results to the publication “From Electron to Biomass: A Mechanistic Model to Describe Phytoplankton Photosynthesis and Steady-State Growth Rates” by Kroon and Thoms. Validation of the MASS Toolbox model was achieved when simulation outputs and parameter inputs were identical to that presented in the publication.

Simulation of a model requires several precursor steps. Once the model is constructed, all rate constants for every reaction and all initial concentrations of each metabolite must be provided into the model. This is because the entire network is dynamic, changing over time, and is therefore described using a set of ordinary differential equations. Rate constants by Kroon and Thoms are provided and these values were entered in the model as shown in in figure 2.4. Initial conditions provided by Kroon and Thoms is shown in the following table:

Table 2.4: List of Initial Conditions for Each Metabolite (Kroon and Thoms). X_n (n = 1, 2, … 31) describes the metabolites from the three complexes shown in Figures 2.1, 2.12, and 2.15. The values are unitless because they describe the number of actual molecules per PSII/Cyt b6f/PSI complex where the ratio of these complexes in the system is assumed to be 1:1:1[19].

Variable	Unit	Initial value	Definition
X_1	None	$\frac{x}{x+1}$	Fraction of PSIIs with oxidized Q_A and vacant Q_B site
X_2	None	0	Fraction of PSIIs with reduced Q_A and vacant Q_B site
X_3	None	$\frac{1}{x+1}$	Fraction of PSIIs with oxidized Q_A and oxidized Q_B
X_4	None	0	Fraction of PSIIs with reduced Q_A and oxidized Q_B
X_5	None	0	Fraction of PSIIs with reduced Q_A and reduced Q_B
X_6	None	0	Fraction of PSIIs with oxidized Q_A and reduced Q_B
X_7	None	0	Fraction of PSIIs with oxidized Q_A and doubly reduced Q_B
X_8	None	0	Fraction of PSIIs with reduced Q_A and doubly reduced Q_B
X_9	None	0	Number of plastoquinol per PSII
X_{10}	None	4	Number of plastoquinone per PSII
X_{11}	None	1	Fraction of cyt b6/f complexes in neutral state
X_{12}	None	0	Fraction of cyt b6/f complexes with oxidized cyt f
X_{13}	None	0	Fraction of cyt b6/f complexes with reduced Q_o site
X_{14}	none	0	Fraction of cyt b6/f complexes with one of the *b* hemes in reduced state
X_{15}	None	0	As X_{14} additionally with oxidized cyt f
X_{16}	None	0	As X_{15} additionally with reduced Q_o site
X_{17}	None	0	Fraction of cyt b6/f complexes with both *b* hemes in reduced state
X_{18}	None	0	Fraction of cyt b6/f complexes with doubly protonated Q_i site
X_{19}	None	0	As X_{11} with reduced ferredoxin at Q_i site
X_{20}	None	0	As X_{14} with reduced ferredoxin at Q_i site
X_{21}	None	0	Number of reduced ferredoxin per PSI
X_{22}	None	1	Number of neutral ferredoxin per PSI
X_{23}	None	0	Number of oxidized plastocyanin per PSI
X_{24}	None	1	Number of neutral plastocyanin per PSI
X_{25}	None	1	Fraction of PSIs with $A_oA_1F_xF_{AB}$ in neutral state
X_{26}	None	0	Fraction of PSIs with reduced F_{AB}
X_{27}	None	0	As X_{26} additionally with reduced F_x
X_{28}	None	0	As X_{27} additionally with reduced A_1
X_{29}	None	0	As X_{28} additionally with reduced A_o
X_{30}	None	1	Fraction of PSIs with P700 in neutral state
X_{31}	None	0	Fraction of PSIs with oxidized P700

These values were incorporated into the MASS Toolbox using the syntax shown below:

```
TableForm[
  ic = {
    Qa → 0.08994,
    Qa- → 0.,
    Qa_Qb → 0.910059,
    Qa-_Qb → 0.,
    Qa-_Qb- → 0.,
    Qa_Qb- → 0.,
    Qa_Qb2- → 0.,
    Qa-_Qb2- → 0.,
    PQH2 → 0.,
    PQ → 4. ,|
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp"], metabolite["bn"], metabolite["Qi"]]] → 1.,
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f+"], metabolite["bp"], metabolite["bn"], metabolite["Qi"]]] → 0.,
    transform[complex[metabolite["QH2"], metabolite["FeS"], metabolite["f+"], metabolite["bp"], metabolite["bn"], metabolite["Qi"]]] → 0.,
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]] → 0.,
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f+"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]] → 0.,
    transform[complex[metabolite["QH2"], metabolite["FeS"], metabolite["f+"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]] → 0.,
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp2-"], metabolite["bn"], metabolite["Qi"]]] → 0.,
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp"], metabolite["bn"], metabolite["QH2"]]] → 0.,
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp"], metabolite["bn"], metabolite["Qi"],
       metabolite["Fd-"]]] → 0.,
    transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"], metabolite["bp-"], metabolite["bn"], metabolite["Qi"],
       metabolite["Fd-"]]] → 0.,
    Fd- → 0.,
```

Figure 2.15: Syntax Used to Input Initial Conditions Into MASS Toolbox. Only a portion of the metabolites is shown, but the input is identical for the remaining metabolites that are not shown.

Complexes of multiple metabolites within close proximity are constructed using the "complex[]" command. The "transform[]" command combines every species of the complex and treats it as a single metabolite, which is important for rules of mass action.

The first simulation of four that was run to aid in validation of the entire model was to observe how rate of photosynthesis is affected as irradiance changes. Rate of photosynthesis, in this case, is defined as the number of electrons that flow through PSII per millisecond to oxidize the negatively charged ferredoxin ion so that it ultimately forms NADPH. This is equivalent to the rate constant of terminal ferredoxin re-oxidation multiplied by the concentration of negatively charged ferredoxin ion[19]. Irradiance is defined as the μmol of photons per m^2 per second. The specific output of this graph shown in the publication by Kroon and Thoms was based upon antenna scenario 6 shown in Table 2.2 in which the number of chlorophyll a and chlorophyll b per PSII complex is 156 and 81, respectively, and the number of chlorophyll a and chlorophyll b per PSI complex is 140 and 14, respectively. The rate constant was also changed among five different values: 0.15 ms^{-1}, 0.30 ms^{-1}, 1.0 ms^{-1}, 4 ms^{-1}, 150 ms^{-1}.

Antenna Scenario 6 was inputted into the model using the following syntax shown below:

```
sigmachla = 8.6749 * (10 ^ -21) ;(* units=m^2 per chlorophyll a *)
sigmachlb = 9.1222 * (10 ^ -21) ;(* units=m^2 per chlorophyll b *)
NaII = 156(* Scenario 6. Number of chlorophyll A per PSII. *);
NbII = 81(* Scenario 6. Number of Chlorophyll B per PSII. *);
sigmaPSII = sigmachla * NaII + sigmachlb * NbII (*Entire optical surface area per PSII that can absorb photons*);
cf = 6.02 * 10 ^ 14 (*conversion factor that changes units from micromole to quanta, and seconds to milliseconds*);
NaI = 140; (*Number of chlorophyll A per PSI. Scenario 6 from Table 2. To solve P/E curve in fig. 7*)
NbI = 14; (*Number of chlorophyll B per PSI. Scenario 6 from Table 2. To solve P/E curve in fig. 7*)
sigmaPSI = (sigmachla * NaI) + (sigmachlb * NbI);
PFDplugin = 1000;
```

Figure 2.16: Chlorophyll Antenna Scenario 6 Incorporated Into Model. The number of chlorophyll also affect the rate constants shown in figure 2.5.

The PFDplugin shown in the bottom of Figure 2.16 shows the value of the irradiance in units of μmol of photons per m^2 per second. This PFDplugin value influences the EXCI value that participates in the value of the rate expression for reactions R1a, R1b, R1c, R1d, which are all dependent on light energy, in PSI as shown in figure 2.8. The rate expressions for these reactions, as shown by Kroon and Thoms, is demonstrated in the following equations:

$$R1a = EXCI*\Phi_{PSI}*X_{30}*X_{25} \quad (1.19)$$

$$R1b = EXCI*\Phi_{PSI}*X_{30}*X_{26} \quad (1.20)$$

$$R1c = EXCI*\Phi_{PSI}*X_{30}*X_{27} \quad (1.21)$$

$$R1d = EXCI*\Phi_{PSI}*X_{30}*X_{28} \quad (1.22)$$

where $X_{25} – X_{30}$ are the concentrations of metabolites shown in figure 2.13. Therefore, for a PFDplugin value of 1000, the resulting EXCI value is 0.676099. EXCI can be determined by the equation[19]:

$$EXCI = sigmaPSI * PFDplugin * cf \quad (1.23)$$

The syntax used on MASS Toolbox interface to determine the value of EXCI with a PFDplugin value of 1000 is shown below:

```
In[465]:= EXCI = sigmaPSI * PFDplugin * cf
Out[465]= 0.676099
```

Figure 2.17: MASS Toolbox Syntax and Input Used to Calculate EXCI. This value is important for determining the rate expression for all reactions in PSI that depend on light input.

The quantum yield of Φ_{PSI} is equivalent to the value of 1 since it is at a maximum value and is independent of open antenna centers, unlike PSII[19]. Recall that the quantum yield

of PSII shown in equation 1.9, designated as Φ_{PSII} and incorporated into the rate equations of light-dependent reactions of PSII (reactions qR4, qR6, qR15, qR19), will always be a fractional value less than and dependent on open Q_a centers[19].

In addition to figure 2.4 that reveals the incorporation of rate constants of regular, light-independent reactions within the three complexes, and in addition to figure 2.7 that reveals the rate law modifications for light-dependent reactions in PSII, the figure below shows the syntax used on MASS Toolbox to modify the rate laws of reactions R1a – R1d:

```
k["R1a"] → EXCI,
k["R1b"] → EXCI,
k["R1c"] → EXCI,
k["R1d"] → EXCI ,
```

Figure 2.18: Syntax on MASS Toolbox to Implement Rate Laws for Light-Dependent Reactions of PSI. Each value is also multiplied by Φ_{PSI}, which equals 1, and therefore the value of EXCI alone is sufficient.

Once these parameters are inputted, they needed to be updated into the model so that the model could be simulated. The syntax for updating the parameters and simulating the solution is shown below:

```
In[423]:= updateParameters[model, {parameter["PFD"] → 1000}];
In[424]:= {concProfile, fluxProfile} = simulate[model, {t, 0, 10^8}];
```

Figure 2.19: Syntax to Update Parameters and Simulate the Solution. The PFD value of 1000 is used to modify the rate expression of PSI light-dependent reactions. The "simulate" command solves the concentration for all metabolites and fluxes for all reactions between time t = 0 ms until t = 10^8 ms.

Once the model is simulated, a concentration is provided for every metabolite and a flux value is provided for every reaction for every single time point. Once these values are determined, they can be used for a wide variety of calculations, which will be demonstrated in subsequent sections of this thesis.

2.2.4 Model Validation Part 1: Photosynthetic Rate as a Function of Irradiance and Ferredoxin Oxidation Rate Constant

The first step in validation of the model was to ensure that the effects of changing the ferredoxin reaction rate constant altered the photosynthetic rate as a function of irradiance in the same manner as portrayed in the publication by Kroon and Thoms. Photosynthetic rate can be described as the rate of reaction R3 in PSI shown in figure 2.15 in which ferredoxin ion becomes oxidized to give off an electron[19]. This is therefore equivalent to the product of the rate constant of terminal ferredoxin oxidation and the concentration of ferredoxin ion. The result is shown below:

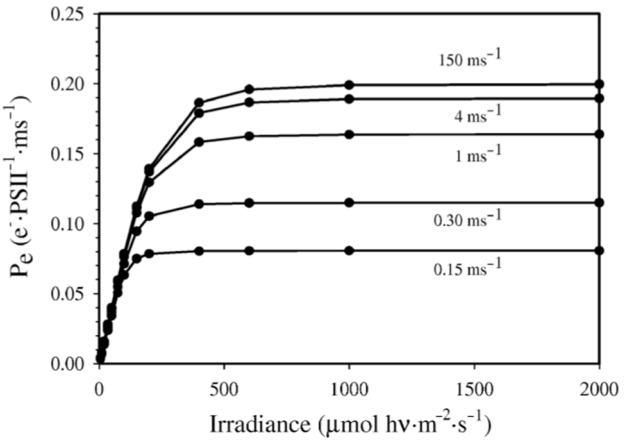

Figure 2.20: Photosynthetic Rate vs. Irradiance With Varying Ferredoxin Oxidation Rate Constants (Kroon and Thoms). Antenna Scenario 6 was used in which 156 chlorophyll a molecules and 81 chlorophyll b molecules are associated with PSII. 140 chlorophyll a molecules and 14 chlorophyll b molecules are in PSI.

Since the MASS Toolbox reconstructed model has stored values of concentrations for every metabolite at every time point, it is imperative to ensure that the system reaches steady-state in terms of concentration. This can be achieved by observing the concentration profiles and choosing a time point at which all metabolites maintain a

steady-state. The syntax used in MASS Toolbox to observe the concentration profile is shown in the figure below:

```
In[544]:= plotSimulation[concProfile,
            PlotFunction → LogLinearPlot, Legend → {Position → Left}]
```

Figure 2.21: MASS Toolbox Command for Observing Concentration Profile for Each Metabolite

The resulting graphical output from the syntax is shown below:

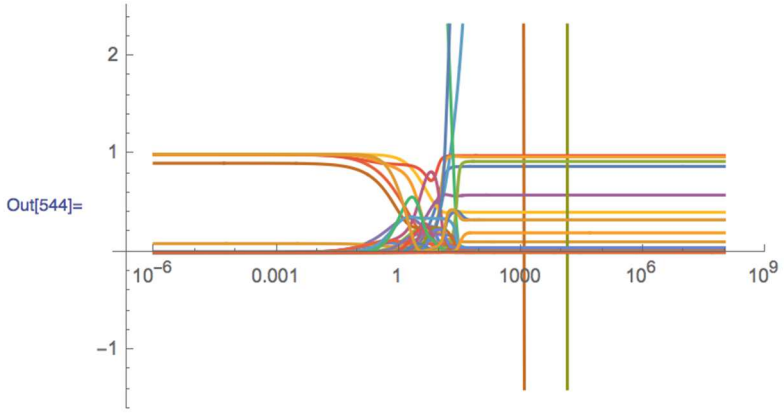

Figure 2.22: Concentration Profiles of Every Metabolite in the System. Each line represents a metabolite's concentration, where the x-axis is time in milliseconds and the y-axis is the number of molecules present per protein complex (PSII, Cyt b6/f, PSI). After approximately $10^{1.5}$ milliseconds, most metabolites have achieved a steady state flux.

Since almost all metabolites have hit a steady state concentration after $10^{1.5}$ milliseconds, it is ideal to pick a concentration at a time point beyond $10^{1.5}$ milliseconds after a simulation is run. The syntax used on MASS Toolbox to calculate photosynthetic rate as a function of irradiance is shown in the following figure:

```
(* VALIDATION of Photosynthetic Rate vs. Irradiance. Varying Fd oxidation rate const. *)
0.12 * (metabolite["Fd-"] /. concProfile /. t → 10 000)
0.0701909
```

Figure 2.23: MASS Toolbox Syntax to Calculate Photosynthetic Rate vs. Irradiance. The input shows the default rate constant of 0.12 multiplied by the concentration of ferredoxin ion at 10000 ms, a value chosen because it ensures that ferredoxin ion has reach steady state. This is equivalent to the photosynthetic rate. The output of 0.0701909 simply shows the photosynthetic rate for this specific rate constant value of 0.12.

This process is repeated for every irradiance point as shown in figure 2.20 calculated by Kroon and Thoms. The resulting graphical representation of the MASS Toolbox model portraying photosynthetic rate vs. irradiance with varying ferredoxin oxidation rate constant is shown below:

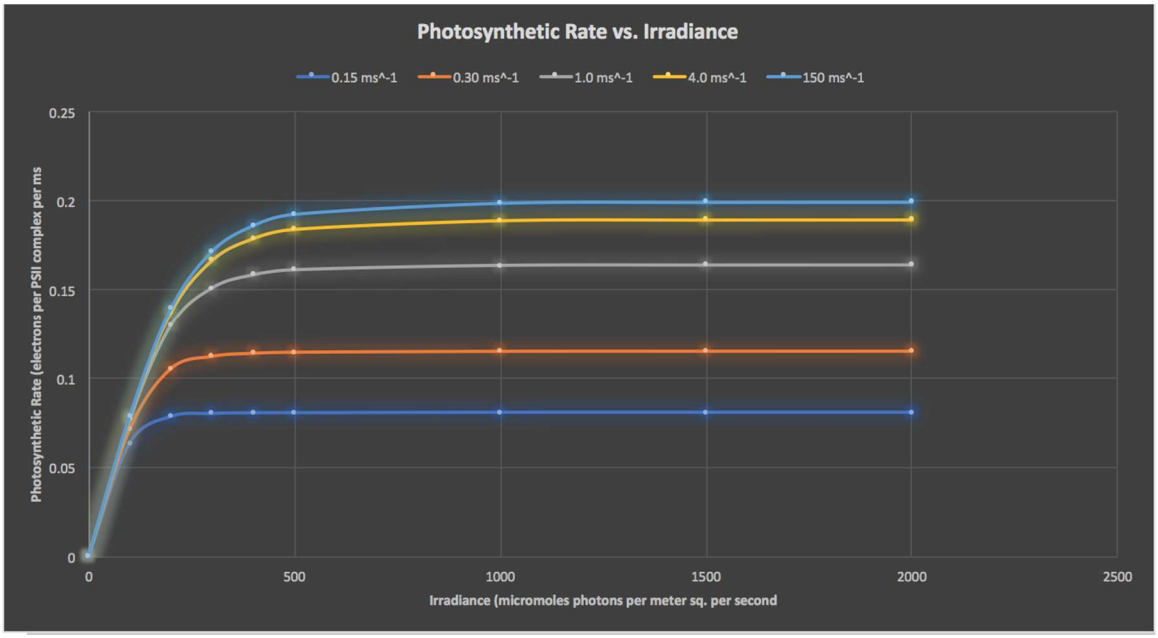

Figure 2.24: MASS Toolbox Output of Photosynthetic Rate vs. Irradiance with Varying Ferredoxin Oxidation Rate Constant. Each line describes a different ferredoxin oxidation rate constant.

The graph shown in figure 2.20 is identical to the graphs shown in figure 2.24, therefore providing the first line of validation evidence that the constructed MASS Toolbox model is identical to the model constructed by Kroon and Thoms.

2.2.5 Model Validation Part 2: Photosynthetic Rate as a Function of Irradiance, Antenna Size, and Presence/Absence of Cyclic Electron Flow

The next step towards complete validation of the MASS Toolbox model was to simulate the model and observe the effects of antenna size and the effects of the presence or absence of cyclic electron flow on the relationship between photosynthetic rate and irradiance. The validation method is similar to the method conducted in part 1 in which the effects of varying the ferredoxin oxidation rate constant were observed. However, there are several key differences.

The publication by Kroon and Thoms shows a graph of the relationship between photosynthetic rate and irradiance and the effects of varying antenna size as well as the presence and absence of cyclic electron flow on this relationship. This graph is shown in the following figure:

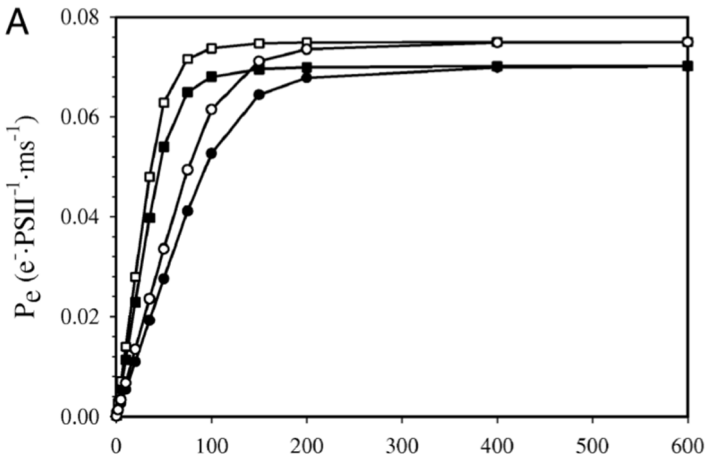

Figure 2.25: Photosynthetic Rate vs. Irradiance and Effects of Antenna Size and Cyclic Electron Flow (Kroon and Thoms). As described in the paper, square data points describe antenna scenario 2 and circular data points describe antenna scenario 8. Open data points represent the absence of cyclic electron flow from PSI to cyt b6/f, while closed data points represent the presence of cyclic electron flow.

For antenna scenario 2, the PSII complex has 240 chlorophyll a molecules and 143 chlorophyll b molecules. For antenna scenario 8, the PSII complex has 114 chlorophyll a molecules and 50 chlorophyll b molecules and the PSI complex has 120 chlorophyll a molecules and 9 chlorophyll b molecules. The MASS Toolbox model was constructed to easily implement changes in chlorophyll antenna sizes as well as propagation of chlorophyll number into relevant biophysical constants necessary for simulations of the model. The figure below shows how antenna scenario 2 was incorporated into MASS Toolbox:

```
sigmachla = 8.6749 * (10 ^ -21) ;(* units=m^2 per chlorophyll a *)
sigmachlb = 9.1222 * (10 ^ -21) ;(* units=m^2 per chlorophyll b *)
NaII = 240(* Scenario 2. Number of chlorophyll A per PSII. *);
NbII = 143(* Scenario 2. Number of Chlorophyll B per PSII. *);
sigmaPSII = (sigmachla * NaII) + (sigmachlb * NbII) ; (* Entire optical surface area per PSII that can absorb photons *)
cf = 6.02 * 10 ^ 14 ; (* conversion factor that changes units from micromole to quanta, and seconds to milliseconds *)
NaI = 212; (* Scenario 2. Number of chlorophyll A per PSI. *)
NbI = 54; (* Scenario 2. Number of chlorophyll B per PSI. 7*)
sigmaPSI = (sigmachla * NaI) + (sigmachlb * NbI); (* Entire optical surface area per PSI that can absorb photons *)
PFDplugin = 1000;
```

Figure 2.26: MASS Toolbox Syntax: Parameter Input – Chlorophyll a and Chlorophyll b Input for Scenario 2. The first two lines of syntax describe the surface area per chlorophyll a and b capable of absorbing proteins.

The figure below shows how antenna scenario 8 was incorporated into MASS Toolbox:

```
sigmachla = 8.6749 * (10 ^ -21) ;(* units=m^2 per chlorophyll a *)
sigmachlb = 9.1222 * (10 ^ -21) ;(* units=m^2 per chlorophyll b *)
NaII = 114(* Scenario 8. Number of chlorophyll A per PSII. *);
NbII = 50(* Scenario 8. Number of Chlorophyll B per PSII. *);
sigmaPSII = (sigmachla * NaII) + (sigmachlb * NbII) ; (* Entire optical surface area per PSII that can absorb photons *)
cf = 6.02 * 10 ^ 14 ; (* conversion factor that changes units from micromole to quanta, and seconds to milliseconds *)
NaI = 120;(* Scenario 8. Number of chlorophyll A per PSI. *)
NbI = 9; (* Scenario 8. Number of chlorophyll B per PSI. 7*)
sigmaPSI = (sigmachla * NaI) + (sigmachlb * NbI); (* Entire optical surface area per PSI that can absorb photons *)
PFDplugin = 1000;
```

Figure 2.27: MASS Toolbox Syntax: Parameter Input – Chlorophyll a and Chlorophyll b Input for Scenario 8.

Changing antenna number is a simple matter of adjusting the constants in the inputs designated above. These values then carry over onto the description of biophysical rate constants described in figure 2.8 and encoded on MASS Toolbox as shown on figure 2.9.

Next, it is important to be able to incorporate the presence and absence of cyclic electron flow between PSI and cyt b6/f within the MASS Toolbox model. Looking at the electron flow diagram of cyt b6/f in figure 2.12, there is a portion of the model where ferredoxin ion from PSI is connected to cyt b6/f. Ferredoxin ion can either be oxidized to transfer an electron to $NADP^+$ to form NADPH or recycled back to cyt b6/f complex and oxidized to ultimately pump more protons into the lumen. The reactions involved in this cyclic electron flow are bR11 and bR10. To ensure that these reactions are present in the model, their reactions rate constants are set to equal 0.03 ms^{-1}. The syntax used on MASS Toolbox is shown below:

```
k["br10"] → 0.03,
k["br11"] → 0.03,
```

Figure 2.28: MASS Toolbox Syntax to Ensure Presence of Cyclic Electron Flow. Nonzero rate constants ensure that some electrons from ferredoxin ion oxidation flow back to the cyt b6/f complex.

To ensure strictly linear electron flow, these values are simply switched to zero.

The model was simulated using the same methods as described previously. Values for photosynthetic rate were also calculated by multiplying 0.12 (default ferredoxin oxidation rate constant) by the concentration of ferredoxin ion. When calculated on the MASS Toolbox, the output was as follows:

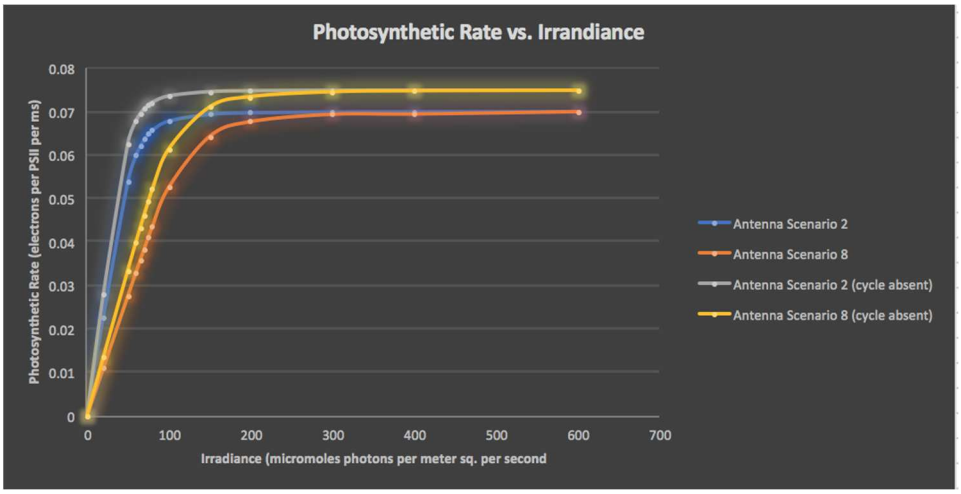

Figure 2.29: Photosynthetic Rate vs. Irradiance and Effects of Chlorophyll Antenna Size and Cyclic Electron Flow. This graph shows how irradiance affects photosynthetic rate. Antenna scenario 2 and 8 are compared as well as the influence of cyclic electron flow.

A quick comparison of figure 2.25 and figure 2.29 show that the results in both are identical, thus providing further support that the model constructed on MASS Toolbox is identical to the model constructed by Kroon and Thoms.

2.2.6 Model Validation Part 3: PSI Cyclic Rate as a Function of Irradiance and Antenna Size

To further validate the MASS Toolbox model, the effects of irradiance on the rate of cyclic electron flow through PSI was observed. Additionally, the effects of antenna size were observed on cyclic rate as well. The publication by Kroon and Thoms shows the output of their model: as irradiance increased, cyclic electron flow from PSI towards

cyt b6/f increased but levels out after a certain point. The output is shown in the figure below:

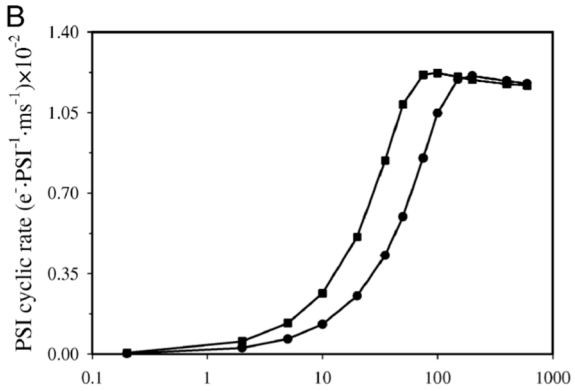

Figure 2.30: PSI Cyclic Rate vs. Irradiance (Kroon and Thoms). PSI cyclic rate is in units of electrons the flow through PSI per millisecond. Irradiance is in its usual units of micromoles of photons per m^2 per seconds.

Simulating the model on the MASS Toolbox was identical to the methods conducted when simulating the model to observe photosynthetic rate in the previous section. However, the equations used on MASS Toolbox needed to be modified to produce the output describing PSI cyclic rate. The syntax used on MASS Toolbox to calculate the PSI cyclic rate after inputting a specific photon flux density value is shown below:

```
(* VALIDATION of PSI Cyclic Rate *)

(0.03 * (transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"],
        metabolite["bp"], metabolite["bn"], metabolite["Qi"]]] /. concProfile /. t → 10 000) *
   (metabolite["Fd-"] /. concProfile /. t → 10 000)) +

 (0.03 * (transform[complex[metabolite["Qo"], metabolite["FeS"], metabolite["f"],
        metabolite["bp-"], metabolite["bn"], metabolite["Qi"]]] /. concProfile /. t → 10 000) *
   (metabolite["Fd-"] /. concProfile /. t → 10 000))
```

Figure 2.31: MASS Toolbox Syntax to Calculate PSI Cyclic Rate. The reaction rates of bR10 and bR11 are added to produce a value of the total number of electrons that pass through PSI per millisecond.

The syntax above is split into two visibly different portions, with the top portion referring to bR11 and the bottom portion referring to bR12. The constant 0.3 in front of both portions

of code refer to the rate constant which describes the "docking of Fd^- to b6/f" for both reactions[19]. The rest of the syntax shows time points that were chosen at 10,000 milliseconds to ensure that a steady state concentration was achieved for all metabolites involved in the mass-action rules that determine the rate of bR10 and bR11. When the MASS Toolbox model was simulated and the syntax from figure 2.31 was run across a wide spectrum of irradiance levels, the resulting output showed a graph that was similar to the output published by Kroon and Thoms:

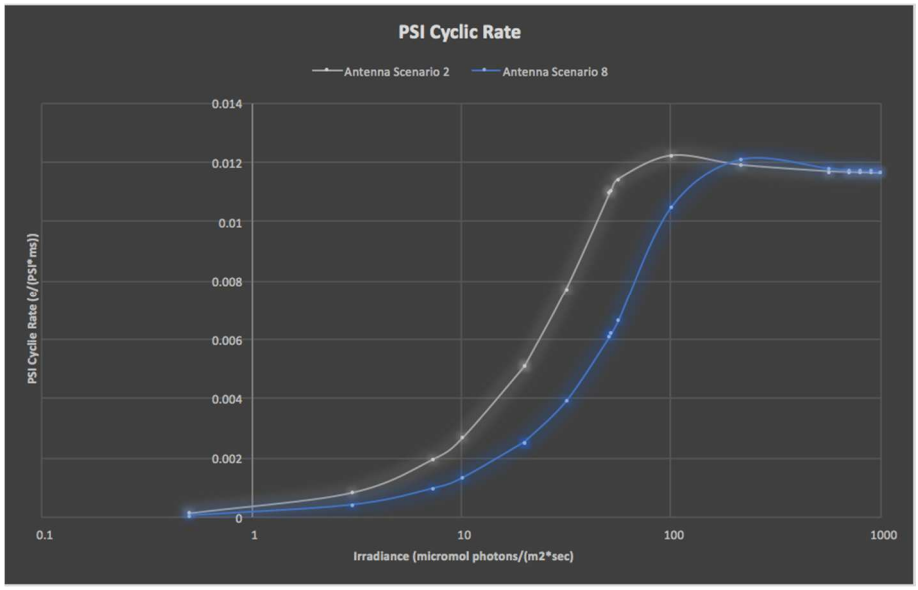

Figure 2.32: MASS Toolbox Model Output for PSI Cyclic Rate vs. Irradiance Levels: A quick glance of figure 2.30 shows that it is identical to figure 2.32, therefore providing additional evidence that the constructed MASS Toolbox model is identical in characteristics to the model constructed by Kroon and Thoms.

2.2.7 Model Validation Part 4: Proton Yield as a Function of Irradiance and Antenna Size

This was the final validation method conducted that demonstrated with confidence that the model constructed using the MASS Toolbox was identical to that of model constructed by Kroon and Thoms. The MASS Toolbox was further validated by observing

the effects of irradiance on the number of protons that is pumped into the lumen per electron that flows through the electron transport chain. For every electron that passes through the electron transport chain, a few protons are pumped into the lumen. One proton is pumped into the lumen for every electron that is produced by the electron transport chain, initially leading to a proton-to-electron ratio of 1:1[19]. Further observation of the cytochrome b6/f complex in figure 2.12 shows that reactions bR3 and bR6 add an additional two protons per electron that flows through this part of the complex[19]. Therefore, if reactions bR10 and bR11 are removed, causing solely linear electron flow due to lack of ferredoxin ion's ability to bind back to cytochrome b6/f from PSI, an electron can only flow through either bR3 or bR6 once, causing the maximum possible number of protons translocated towards the lumen per electron to be two. Taking account both the effects of water splitting and linear electron flow, the maximum ratio of protons pumped into the lumen per electron is therefore 3:1[19].

However, if cyclic electron flow is present, then some of the electrons that flow through reactions bR3 and bR6 will end up on reduced ferredoxin in PSI and flow back towards the cytochrome b6/f complex through reactions bR10 and bR11, thus allowing slightly more protons to be pumped into the lumen for every same electron. This causes the proton-to-electron ratio to increase slightly above 3. In the model constructed by Kroon and Thoms, the resulting proton-to-electron ratio resulting from varying degrees of irradiance is shown below:

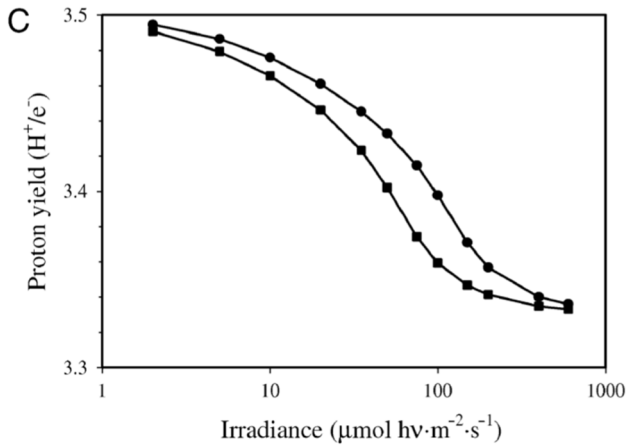

Figure 2.33: Proton-to-Electron Ratio as a Function of Irradiance (Kroon and Thoms). This figure is a result from the publication by Kroon and Thoms, showing how the number of protons pumped into the lumen decreases per electron as irradiance increases.

The MASS Toolbox model was simulated in the exact same fashion using different irradiance values inputted into the photon flux density entry syntax on the model. The syntax used to carry out the calculations required to solve for proton-to-electron ratio is shown below:

```
(* Validation of Proton-to-Electron Ratio *)

1 + ((2 * v["br3"] /. fluxProfile /. t → 10 000) + (2 * v["br6"] /. fluxProfile /. t → 10 000)) /
   ((v["br3"] /. fluxProfile /. t → 10 000) + (v["br6"] /. fluxProfile /. t → 10 000) -
     ((v["br11"] /. fluxProfile /. t → 10 000 ) + (v["br10"] /. fluxProfile /. t → 10 000)))
```

Figure 2.34: MASS Toolbox Syntax for Calculating Proton-to-Electron Ratio. All flux vales are taken at a time point of 10,000 ms to ensure that a steady-state condition was achieved.

The syntax was encoded in this way to account for the electrons coming from water splitting as well as the electrons coming from oxidized ferredoxin. The mathematical equation is:

$$1 + \frac{2*v[br3]+2*v[br6]}{v[br3]+v[br6]-[v[br11]+v[br10]]} \tag{1.24}$$

The 1 in the first quantity is used to represent the proton coming from the water-splitting reaction. While the numerator of the second quantity represents the protons going through the cytochrome b6/f electron flow pathways. The denominator of the second quantity represents the electrons flowing through the cytochrome b6/f minus the number of electrons that flow into the cytochrome b6/f pathway from PSI. This helps establish the ratio of protons produced per electron. This ratio ends up being more than two because the denominator diminishes to a value smaller than 1. Therefore equation 1.24 is modeled exactly by the syntax in figure 2.34. The resulting MASS Toolbox simulation and calculation using equation 1.24 shows the output below:

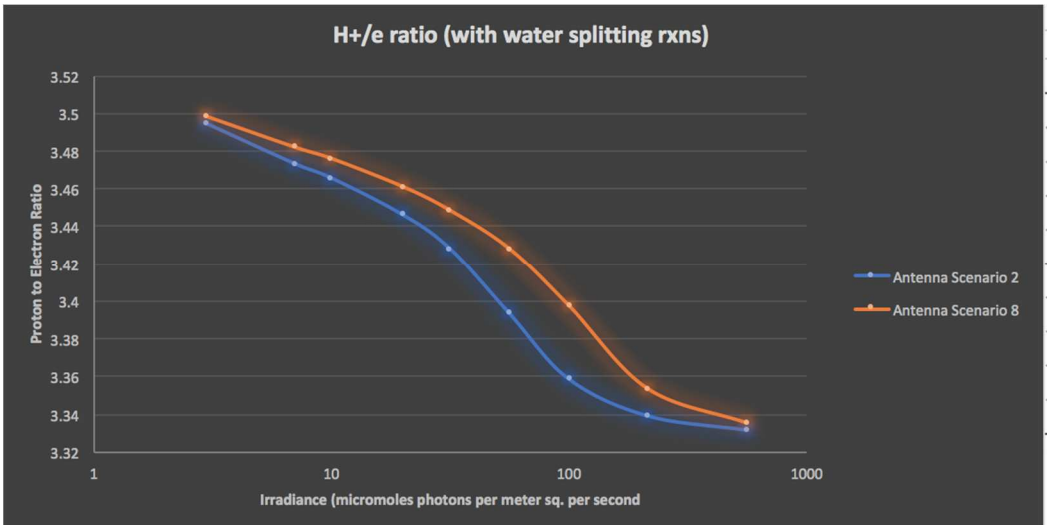

Figure 2.35: Proton-to-Electron Ratio as a Function of Irradiance. The two graphs represent two different antenna scenarios. This was the final validation step for the reconstructed MASS Toolbox model.

Since the graphical outputs shown in both figures 2.33 and 2.35 are identical, it further validates that reconstruction performed on MASS Toolbox is identical to the model by Kroon and Thoms.

2.3 Discussion

With the MASS Toolbox model validated multiple times against the published phytoplankton model constructed by Kroon and Thoms, it can now be used as a tool to help advance research in the field of photosynthesis, especially for phytoplankton. In addition to the simple user interface, the model is also useful in that it is modifiable. Depending on the user's goal, the model can be formulated and tailored to assist in research efforts for a variety of photosynthetic organisms. This is because of the similarities and conserved structures in the electron transport chain among a variety of photosynthetic organisms.

Perusing the construction of the MASS Toolbox model, it consists of metabolites in the correct stoichiometric ratios, reactions in the appropriate network connections, equilibrium constants, rate constants, initial conditions, and a variety of modifiable biophysical constants used to ultimately affect rate constants. Each of these subsections is arranged in an intuitive order so that changes can be implemented. For example, although photosynthetic machinery is well-conserved, antenna complexes surrounding the photosystem complexes and reaction centers may be composed of different sizes for different organisms. This is easily modifiable, just as it was demonstrated through the different antenna scenarios in the validation procedures shown in prior sections. These different antenna sizes then affect the optical cross sectional area and ultimately the rate constants for all reactions that absorb light within both photosystem complexes. Therefore, if an organism has a different number of chlorophyll a or chlorophyll b molecules, this can be implemented into the MASS Toolbox model. Some photosynthetic organisms such as Phaeodactylum tricornutum also have additional pigment types such as chlorophyll c while

other organisms have phycobilisomes, and these could also be incorporated into the model, perhaps through the increasing of optical surface area and minor modification to biophysical rate constants[19].

It is also important to note that this kinetic model was constructed with a photosystem II-to-cytochrome b6/f-to-photosystem I ratio of 1:1:1. Realistically, in some situations the ratios of these protein complexes may be different when expressed on an actual thylakoid membrane. Modifying the MASS Toolbox model to take these different ratios into consideration may be as simple as changing the initial conditions of certain metabolites or perhaps a complete change of rate constants needs to be described. More tests would need to be conducted and compared to experimental results to validate these changes.

Although the light reactions of photosynthesis are well-conserved, there may still be differences in parameters depending on the organism. The metabolic pathways among lower and higher photosynthetic organisms are similar, but concentrations of metabolites and structures of some enzymes may be different. If so, some equilibrium and rate constants will be different from the reconstructed MASS Toolbox model presented above. Fortunately, these values can be easily manipulated, assuming that organism-specific parameters can be determined.

Chapter 3: Using the Constructed Kinetic Model to Solve for Thylakoid Membrane Constraints

3.1 Introduction

With the completion of the kinetic model, it is important to attempt and utilize this tool to answer a biological question. Within molecular biology, constraints play a large role in the determination of a set of flux solutions. When studying and observing a macroscale environment, taking into consideration the size of an individual molecule or a collection of molecules is almost irrelevant. However, when dealing with single cells or organelles, the size of molecules and protein complexes becomes much more important. This portion of the research will focus on general molecular crowding constraints to later implement towards Phaeodactylum tricornutum, and outputs from the reconstructed MASS Toolbox kinetic model of phytoplankton will assist in these calculations.

The diatom Phaeodactylum tricornutum is a form of phytoplankton that impacts the global environment through the collective process of photosynthesis conducted by all organisms of this species within the ocean[23]. Therefore, studying its photosynthetic potential is of great importance. While the reconstructed phytoplankton MASS Toolbox kinetic model is useful and easily modifiable, it is not specifically tailored towards the organism Phaeodactylum tricornutum. Incorporation of kinetic elements of antenna specific to this organism such as chlorophyll c and fucoxanthin could potentially be incorporated into this model, however it is beyond the scope of this thesis. These can be incorporated by entering their surface area contribution to the reaction centers in the model. Even more important is the fact that the kinetic model described previously does not

correct parameters to account for molecular crowding of the thylakoid disc by the photosystem I, cytochrome b6/f, and photosystem II. This would require ingenious modification of the basic initial conditions and rate constants for each metabolic reaction. Instead of using the kinetic model to calculate the constraints of molecular crowding, a more advantageous alternative is to modify the genome-scale model of Phaeodactylum tricornutum, adding constraints to the model, and simulating the model using a variety of optimization methods.

The genome-scale model was constructed by Jared Broddrick, a PhD student in the UCSD Systems Biology Research Group. Genome-scale models contain all of the genes and proteins they encode as well as all of the network connections within the model, and historically, simulations of these models have successfully predicted an output 70-90% of the time[24]. Simulation of the genome-scale model of Phaeodactylum gives a set of plausible solutions and a single solution when using an optimizing function. However, it does not take into account maximum number of proteins that the thylakoid membrane can contain, which in turn can affect photosynthesis. Implementing these crowding constraints is therefore necessary. Using the MASS Toolbox kinetic model, it is possible to solve for parameters that can aid in implementing crowding constraints.

3.2 Methods

3.2.1 Overview of Governing Equations of Membrane Constraints

Determining the parameters of molecular crowding were done by following the methods presented in “Economics of Membrane Occupancy and Respiro-Fermentation” by Kai Zhuang, Goutham N. Vemuri, and Radhakrishnan Mahadevan. The researchers focused on E. coli and how crowding of glucose importer proteins and respiratory proteins

on the E. coli cell membrane affected respiratory potential as well as growth rate. They did this by implementing membrane constraints on the genome-scale model of E. coli constructed by Adam Feist of the UCSD Systems Biology Research Group. These membrane constraint parameters were determined through a series of equations relevant to first principles of chemistry. Because the concept of determining molecular crowding constraints of photosystem II, cytochrome b6/f, photosystem I, and ATP synthase within a single thylakoid membrane are similar in principle to determining the molecular crowding constraints of glucose transport proteins and respiratory proteins imbedded within the E. coli cell membrane, the techniques used in the paper were adopted towards thylakoid membrane constraints.

3.2.2 Determining the Governing Equations for Photosynthetic Membrane Constraints

The following mathematical methods are described by Zhuang, Vemura, and Mahadevan. It is important to initially determine how much membrane area is occupied per metabolic reaction pertaining to each protein complex. This is achieved by observing the relationship between a specific reaction's flux, the concentration of the enzyme of the reaction, and the turnover number of the enzyme itself, demonstrated in the equation below:

$$v_i = k_{cat,\,i}[E_i] \tag{3.1}$$

where v_i represents the flux and rate of reaction i, $k_{cat,\,i}$ represents the turnover rate of the enzyme, and $[E_i]$ represents the concentration of the enzyme[25]. Since the four photosynthetic protein structures are neither enzymes nor transporters, a parallel needed to be drawn between this equation and the photosynthetic machinery. Since each of these photosystems as well as the cytochrome b6/f complex involve multiple reactions within

each complex, the flux v_i for each complex was assumed to be equivalent to the slowest reaction within each complex at steady state. This was designated as the rate limiting step. The parallel to enzyme concentration in equation 3.1 was set to be the concentration of the reactant of the rate limiting step at steady state. Knowing these two values allows for calculation of the $k_{cat,\,i}$ value.

The $k_{cat,\,i}$ value is then used to assist in calculation of the membrane cost of a reaction, which incorporates membrane surface area and turnover rate[25]. This is shown in the equation below:

$$c_i = \frac{m_i}{k_{cat,i}} \tag{3.2}$$

C_i is defined as the membrane cost and m_i is defined as the surface area of membrane consumed per enzyme[25]. It is important to note that equation 3.2 is only true specifically for a flux of 1 mmol/gdw/hr. This value can easily be scaled depending on the flux going through the protein complex and will ultimately be important for establishing constraints.

Next, specific membrane area is calculated using the surface area to volume ratio of a single thylakoid disc and dividing it by the density of a single thylakoid disc. The specific membrane area, designated M_{cyt}, can be calculated as follows:

$$M_{cyt} = \frac{R_{S/V}}{D} \tag{3.3}$$

where $R_{S/V}$ designates surface area to volume ratio and D designates density[25]. This value is used to solve for the cytoplasmic membrane budget, designated B. The equation for B is described as the product of the specific membrane area and the fraction of membrane surface area that can accommodate proteins, thus describing the total amount of thylakoid surface area that can contain proteins[25]. Depending on the cellular structure and the

organism, this fraction can be drastically different. There still needs to be enough lipid to support the structure of the membrane, which is why the entire thylakoid membrane surface area cannot contain proteins. The equation representing the cytoplasmic membrane budget can be described as:

$$B = f * M_{cyt} \quad (3.4)$$

where f designates the fraction of membrane surface area that can accommodate proteins[25]. Equation 3.4 is the constraining value in the following equation where

$$\sum v_i C_i \leq B \quad (3.5)$$

3.2.3 Determining Parameter Values of Membrane Constraints for Phaeodactylum tricornutum

To determine the membrane constraints for photosystem II, cytochrome b6/f, and photosystem I, the specific parameters for each of these complexes needs to be implemented into the set of equations presented in the prior section. Solving for the turnover number $k_{cat,\,i}$ for each complex was determined by solving for the appropriate flux and reactant concentration values as revealed in equation 3.1. The overall rate of each reaction complex was determined to be the same rate as the slowest reaction in the complex, the rate-limiting reaction. The reaction flux v_i and reactant concentration of the rate-limiting reaction $[E_i]$ was determined by simulating the MASS Toolbox kinetic model, obtaining the flux value of the slowest feasible reaction to designate its rate as equivalent to the entire protein complex, and obtaining the concentration of the reactant species of this rate-limiting reaction. It is important to notice that the flux and concentration values fluctuate and oscillate during the initial stages of the simulation, so it is imperative to choose values that have achieved steady-state. This was done by plotting the results of the

simulation and choosing an appropriate time point at which all metabolite concentrations have achieved a constant concentration. The figure below shows a plot of the concentrations after simulation of the MASS Toolbox kinetic model:

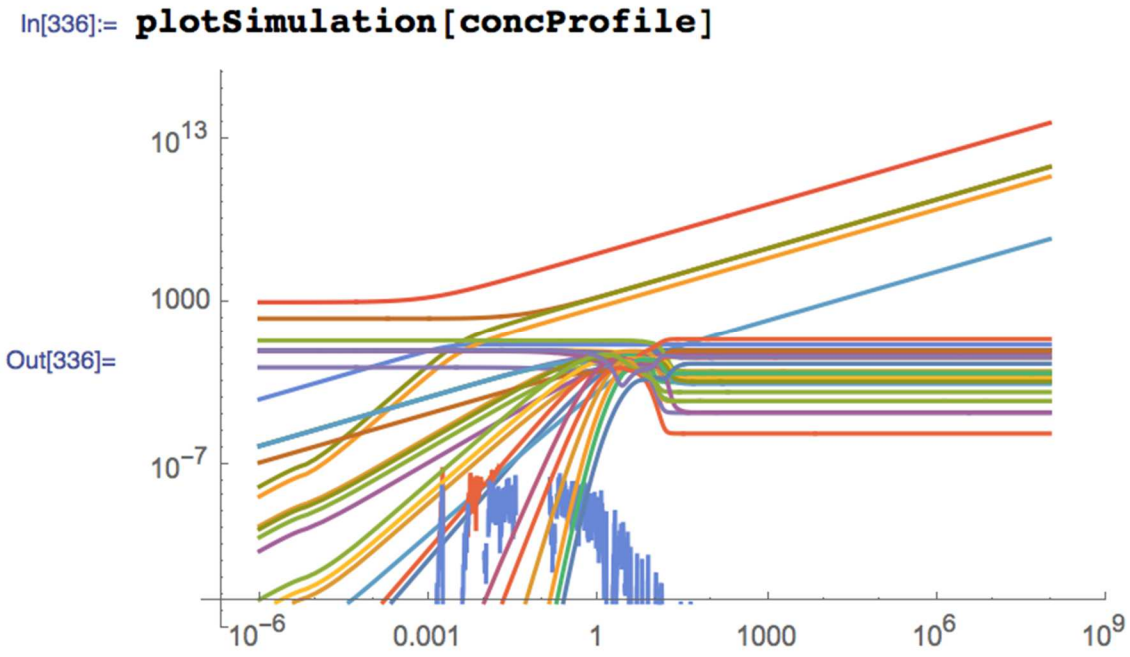

Figure 3.1: Plot of Concentration vs. Time For Each Metabolite in System. X-axis represents milliseconds, y-axis represents the number of molecules present per electron transport chain consisting of photosystem II, cytochrome b6/f, and photosystem I in a 1:1:1 ratio. Irradiance of 1000 micromoles per meter sq. per second and antenna scenario 2 was used in the specific example.

The legend for all profiles shown in figure 3.1 is shown in following figure:

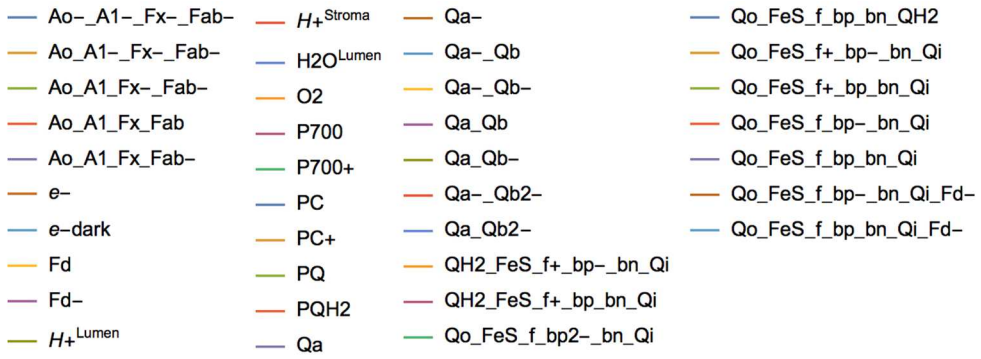

Figure 3.2: Legend of All Metabolites In the Light-Dependent Reactions

The MASS Toolbox syntax used to plot figure 3.1 is plotSimulation[concProfile]. In this simulation output with the specific parameters and conditions noted above in figure 3.1, the four ascending lines simply represent the buildup of protons in the lumen and stroma, oxygen produced from the water-splitting reaction, and electrons from the ferredoxin anion oxidation reaction. During physiological conditions within photosynthetic organisms, the concentrations of these four products do not keep increasing; these will eventually terminate or achieve a steady-state value. This is because the reconstructed kinetic model only contains the protein complexes responsible for the light reactions when it actually should be connected to another system called the dark reactions. In this system, the buildup of electrons from ferredoxin ion oxidation form NADPH which drains into the dark reactions. Also, the oxygen produced from the initial water-splitting reaction is eventually expelled from the organism as a waste product. In physiological conditions, pH homeostasis within the stroma and lumen is also maintained. Therefore, when observing the other metabolites, steady-state concentrations seem to be attained anywhere between 10^1 and 10^2 milliseconds.

However, there is one main anomaly when observing the plot of concentrations. The quickly oscillating profile at the bottom of figure 3.1 represents the metabolite $Q_AQ_B^{2-}$ within photosystem II, which is metabolite X_7 in figure 2.1. Although this value does not hit a steady-state concentration value like the other metabolites, its decreasing profile makes sense: refilling of the plastoquinone pool shown in reactions qr2 and qr10 in figure 2.1 takes a relatively long time because of the diffusion process that needs to occur. This time delay causes a buildup of $Q_A^-Q_B^{2-}$, causing this species to become preferentially oxidized as opposed to $Q_AQ_B^{2-}$ to form plastoquinol (PQH_2). Therefore, the concentration profile of $Q_AQ_B^{2-}$ should not impact the choice of where steady-state occurs for the system.

Next, the model was simulated using the same parameters and the flux profile was observed. Using the command plotSimulation[fluxProfile], the flux profiles for every reaction in the model was graphically displayed. This is shown below:

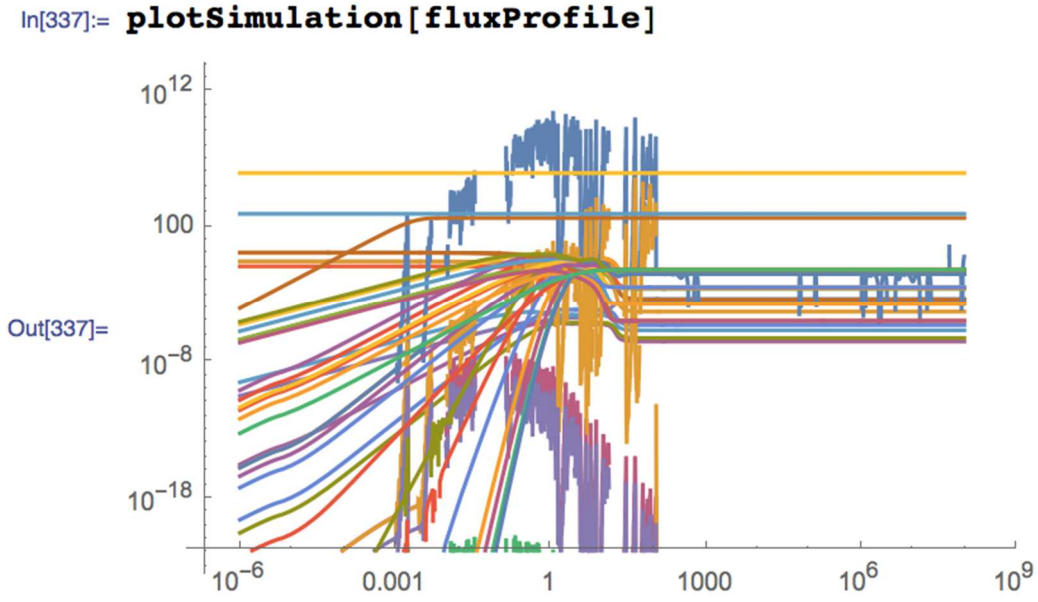

Figure 3.3: Plot of Flux vs. Time For Each Reaction in System. X-axis represents time in milliseconds, while the y-axis describes the number of molecules flowing through the reaction per electron transport chain per millisecond, in which the ratio of photosystem II to cytochrome b6/f to photosystem I is 1:1:1.

The legend for all flux profiles shown in figure 3.3 is shown in the following figure:

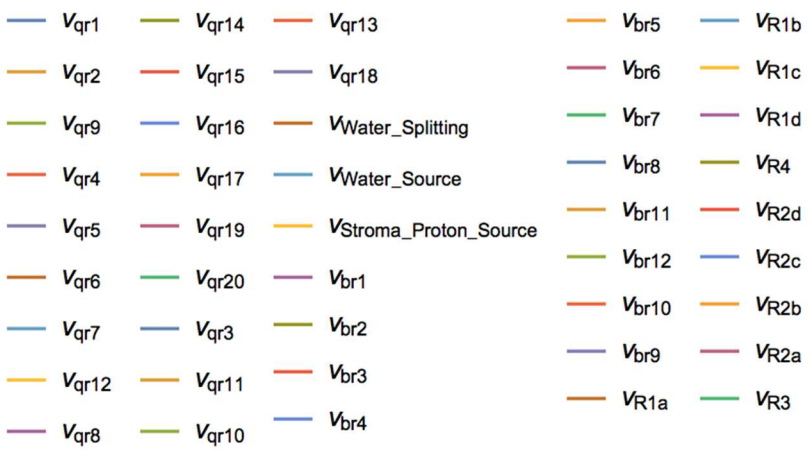

Figure 3.4: Legend of Fluxes for All Reactions of Light-Dependent Reactions.

Although rather crowded with multiple reaction profiles, figure 3.2 assists in efforts to determine after which time points the system has hit steady-state. However, it is much more beneficial to observe profiles that are associated solely with a specific protein complex.

To calculate the turnover number of the photosystem I complex as a whole, a value required to eventually solve for membrane constraints, the concentration and flux profiles shown in figures 3.1 and 3.3 were filtered to only show reactions and metabolites relevant to photosystem II. To display the metabolites of relevant to photosystem II, the following syntax was used:

```
In[341]:= plotSimulation[FilterRules[concProfile, {metabolite["Qa"], metabolite["Qa-"], metabolite["PQ"],
            transform[complex[metabolite["Qa"], metabolite["Qb"]]],
            transform[complex[metabolite["Qa-"], metabolite["Qb"]]],
            transform[complex[metabolite["Qa-"], metabolite["Qb-"]]],
            transform[complex[metabolite["Qa"], metabolite["Qb-"]]],
            transform[complex[metabolite["Qa"], metabolite["Qb2-"]]],
            transform[complex[metabolite["Qa-"], metabolite["Qb2-"]]],
            metabolite["PQH2"]}]]
```

Figure 3.5: MASS Toolbox Syntax to Only Plot Metabolites Associated With Photosystem II

The result of the above transformation created a concentration profile output that is shown below:

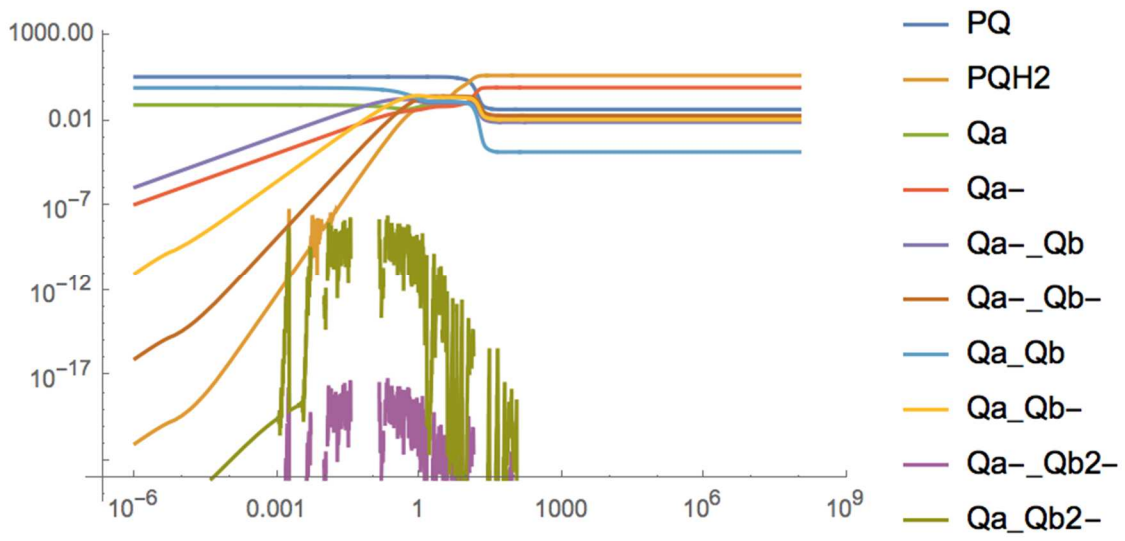

Figure 3.6: Concentration vs. Time of Metabolites for Photosystem II. X-axis represents time in milliseconds. Y-axis represents the number of molecules of that specific metabolite per electron transport chain. The metabolites at steady state are PQH_2, Qa^-, PQ, Qa^-Qb^-, Qa, QaQb, Qa^-Qb, $QaQb^-$. The two oscillating profiles represent $QaQb^{2-}$ and Qa^-Qb^{2-}.

To observe the flux values of all metabolites within photosystem II, a similar syntax as shown in figure 3.5 was applied to the flux profile. The result of solely photosystem II flux is shown below:

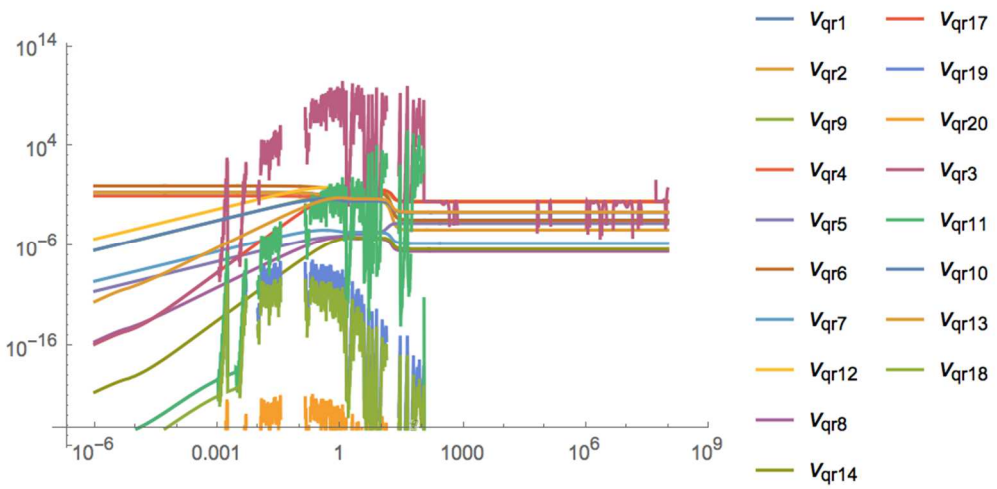

Figure 3.7: Flux vs. Time For Each Reaction of Photosystem II. Flux profiles oscillate but become relatively constant after approximately 100 milliseconds. X-axis represents time in milliseconds while the y-axis represents how many molecules flow through a specific reaction per millisecond.

Observing figures 3.6 and 3.7 show that an appropriate time point to calculate the concentration and flux values of any metabolite or reaction, respectively, to ensure that it is at steady state is any value after 100 milliseconds.

Exact values for calculating concentration of photosystem II metabolites was achieved using the following syntax:

```
In[380]:= tpoints = {10 000};

conc = FilterRules[concProfile, {metabolite["Qa"], metabolite["Qa-"], metabolite["PQ"],
     transform[complex[metabolite["Qa"], metabolite["Qb"]]],
     transform[complex[metabolite["Qa-"], metabolite["Qb"]]],
     transform[complex[metabolite["Qa-"], metabolite["Qb-"]]],
     transform[complex[metabolite["Qa"], metabolite["Qb-"]]],
     transform[complex[metabolite["Qa"], metabolite["Qb2-"]]],
     transform[complex[metabolite["Qa-"], metabolite["Qb2-"]]],
     metabolite["PQH2"]}] /. t → tpoints;

TableForm[conc[[All, 2]], TableHeadings → {conc[[All, 1]], tpoints}]
```

Out[382]//TableForm=

	10 000
PQ	0.049288
PQH2	4.81768
Qa	0.0123658
Qa-	0.944547
Qa-_Qb	0.00884722
Qa-_Qb-	0.0216387
Qa_Qb	0.000157055
Qa_Qb-	0.0124436
Qa-_Qb2-	-3.27718×10^{-69}
Qa_Qb2-	-2.10392×10^{-36}

Figure 3.8: MASS Toolbox Syntax to Calculate Steady-State Concentrations of Photosystem II Metabolites. The left column shows the metabolites of photosystem II while the right column shows the steady state concentrations per electron transport chain that contains a photosystem II to cytochrome b6/f to photosystem I ratio of 1:1:1.

Furthermore, exact values of the fluxes can be solved through the following syntax:

```
tpoints = {10 000};

fluxes = FilterRules[fluxProfile, {v["qr1"], v["qr2"], v["qr3"], v["qr4"], v["qr5"],
      v["qr6"], v["qr7"], v["qr8"], v["qr9"], v["qr10"], v["qr11"],
      v["qr12"], v["qr13"], v["qr14"], v["qr15"], v["qr16"], v["qr17"],
      v["qr18"], v["qr19"], v["qr20"]}] /. t → tpoints ;

In[478]:= TableForm[fluxes[[All, 2]], TableHeadings → {fluxes[[All, 1]], tpoints}]
```

Out[478]//TableForm=

	10 000
v_{qr1}	0.000524159
v_{qr2}	0.0000533988
v_{qr9}	0.0400372
v_{qr4}	0.0372716
v_{qr5}	0.000242465
v_{qr6}	0.000473377
v_{qr7}	2.27108×10^{-6}
v_{qr12}	0.040883
v_{qr8}	3.45062×10^{-7}
v_{qr14}	6.00041×10^{-7}
v_{qr15}	0.037506
v_{qr16}	5.55466×10^{-6}
v_{qr17}	0.0374999
v_{qr19}	-6.34137×10^{-36}
v_{qr20}	-8.41252×10^{-73}
v_{qr3}	−0.168302
v_{qr11}	-2.62157×10^{-34}
v_{qr10}	0.00300806
v_{qr13}	0.00338218
v_{qr18}	-7.3006×10^{-38}

Figure 3.9: MASS Toolbox Syntax to Calculate Steady-State Flux of Photosystem II Reactions. The left column represents the names of all reactions relevant to photosystem II while the right column represents the flux values at steady-state 10,000 milliseconds after simulation.

Observing these flux values helps determine which pathway among the 20 reactions is the main pathway in which most of the electrons flow. It is imperative to determine this main pathway in order to solve for membrane constraints, and this kinetic model allows for easy visualization of all fluxes. Figure 3.9 shows that the main flux flows through reactions qR4, qR9, qR12, qR15, and qR17. Since these reactions are in the main pathway, the slowest of these main reactions can be considered the rate-limiting step and therefore of identical value to the rate or flux of the entire photosystem II

complex. Looking at Table 2.1, out of the reactions in the main flux pathway of photosystem II, reaction qR17 has a value of 1.733 ms^{-1}, making this the rate value of the entire photosystem II complex as well.

The same method as above was applied towards the electron flow pathway of the cytochrome b6/f complex. The following flux values were for all reactions of the cytochrome b6/f complex are as follows:

```
In[487]:= TableForm[fluxes[[All, 2]], TableHeadings → {fluxes[[All, 1]], tpoints}]
Out[487]//TableForm=
                10 000
v_br1           0.0374999
v_br2           0.0374999
v_br3           0.0374999
v_br4           0.0374999
v_br5           0.0374999
v_br6           0.0374999
v_br7           0.0374999
v_br8           0.0374999
v_br11          0
v_br12          0.
v_br10          0
v_br9           0.
```

Figure 3.10: Steady State Flux Values of Cytochrome b6/f Reactions. The left column designates the reaction names while the right column describes the steady state flux values in electrons per millisecond after 10,000 milliseconds.

Therefore, the main electron flow pathway goes from reaction bR1 until bR8. Table 2.1 shows that reactions bR3, bR6, and bR7 all have approximately the same rate constant value and are therefore we can assume that the overall rate and k_{cat} value of the cytochrome b6/f complex is 1 millisecond.

The same method was then applied to photosystem I. The flux values for all reactions in photosystem I were calculated using the MASS Toolbox and are shown below:

```
In[492]:= TableForm[fluxes[[All, 2]], TableHeadings → {fluxes[[All, 1]], tpoints}]
Out[492]//TableForm=
```

	10 000
v_{R1a}	0.0000111911
v_{R1b}	0.000207124
v_{R1c}	0.00383341
v_{R1d}	0.0709481
v_{R4}	0.0749998
v_{R2d}	0.0709481
v_{R2c}	0.00383341
v_{R2b}	0.000207124
v_{R2a}	0.0000111911
v_{R3}	0.0749998

Figure 3.11: Steady State Flux Values of Photosystem I Reactions. The left column designates the reaction names while the right column designates the steady state fluxes in units of electrons per millisecond after 10,000 milliseconds.

The reactions that constitute the main pathway of flux can be determined by observing the pathways with highest flux. Reactions R1c, R1d, R2c, R2d, R3, and R4 carry the most amount of electron flux in this complex. Since R3 is a reaction that is separate from the complex, the next slowest reactions are both R2a and R2b with a rate constant value of 0.2 ms^{-1}. Therefore, the rate and k_{cat} value of the entire photosystem I complex is also 0.2 ms^{-1}.

These k_{cat} values can therefore be plugged into equation 3.2, helping to solve for the membrane cost for each reaction. The only remaining parameter that needs to be solved for this is the membrane occupancy for each protein complex.

For the cyanobacterium Synechococcus sp. the shape of the photosystem II protein complex in its dimeric form is that of an elliptical disk with a length of 15.1 nm +/- 1.5 nm and a width of 18.5 nm +/- 1.0 nm[26]. Calculating the area of an ellipse, this means that the amount of membrane area occupied by a single photosystem II protein complex is 279.35 nm^2. Although the organism of interest in this thesis is Phaeodactylum tricornutum, the core complex subunits of photosystem II are highly conserved through

evolution, regardless of the species; usually it is only the peripheral antenna that is responsible for catching energy of photons that is different among species[27]. Therefore, it is a safe assumption that the size and structure of photosystem II for Synechococcus sp. is nearly identical to Phaeodactylum tricornutum. This is therefore the m_i value for photosystem II in equation 3.2.

Calculating the membrane area occupied by cytochrome b6/f was slightly more involved than the calculation pertaining to photosystem II. The cytochrome b6/f complex is a 220 kDa dimer[28]. The given molecular weight is important because it allows for calculation of the solvent accessible surface area. Large molecular complexes usually fold into each other and thus only a portion of area is accessible towards the external environment, otherwise known as the solvent accessible surface area. Using the following equation:

$$A_s = 5.3M^{0.76} \tag{3.7}$$

where A_s denotes the solvent accessible surface area, and M denotes the molecular mass, the cytochrome b6/f solvent accessible surface area was therefore calculated to be 515.54895 nm^2 [29]. To simplify calculations, the cytochrome b6/f structure was assumed to be a basic sphere, meaning the surface accessible surface area of the cytochrome b6/f complex was assumed to be equivalent to the surface area of a sphere. If the surface area of the sphere is 515.54895 nm^2, the radius is therefore calculated to be 6.405 nm. The radius of the sphere is identical to the radius of the circular cross-sectional area of the sphere which passes through the cellular membrane. Therefore, the membrane area occupied by the protein complex is 128.8872 nm^2. This value is then plugged into equation 3.2 to calculate the membrane cost of the cytochrome b6f complex.

Calculation of the membrane occupancy area of photosystem I was conducted in a similar manner to the cytochrome b6/f complex. The volume of photosystem I is approximately 657 nm^3 [30]. Assuming that the shape of photosystem I is spherical, the cross-sectional area intersecting the membrane can be calculated from the volume of the sphere, leading to a membrane area occupancy value of 91.3681 nm^2. This value can then be used in equation 3.2 to calculate the membrane cost of photosystem I.

Next, specific membrane area, M_{cyt}, was calculated using values of the thylakoid surface area to volume ratio and density of a single thylakoid disk as shown in equation 3.3. Thylakoids are on average 300-600 nm[31]. For the sake of calculation of the surface area to volume ratio, the diameter of the thylakoid disk was assumed to be 450 nm. The thickness of the lipid bilayer of a thylakoid disk is approximately 4.0 nm and the stromal thickness is approximately 4.5 nm[31]. Assuming the shape of a thylakoid disk is a very thin cylinder, the height of this cylinder is then 12.5 nm. The surface area of a single thylakoid disk is 335757.7149 nm^2 while the volume is 1988039.101 nm^3, leading to a surface area to volume ratio of 0.16 nm^{-1}. Density of the thylakoid stroma is assumed to be the same as water: 1 gram per cm^3. Dividing the thylakoid value of the surface area to volume ratio by the density of water allowed for calculation of the specific membrane area.

To solve for the cytoplasmic membrane budget shown in equation 3.4, the specific membrane area is multiplied by the fraction of membrane area that can house proteins. 70% of the thylakoid membrane is occupied by proteins while 80% of grana thylakoid membranes are occupied by proteins[32]. Therefore, the cytoplasmic membrane budget was solved by multiplying the specific membrane area by 0.70.

The relative membrane costs for each protein complex assume a flux of 1 mmol/gdw/hr. Therefore, a variable coefficient is multiplied by each relative membrane cost, indicating the total membrane cost depending on the flux that flows through each complex, which is ultimately dependent on external factors such as irradiance and size of chlorophyll antenna. The summation of the product between the flux going through a protein complex and the membrane cost is constrained by the cytoplasmic membrane budget, ultimately providing a solution space of various fluxes.

3.3 Results

Table 3.1: Biophysical Values for Each Protein Complex

Protein Complex	**Membrane Surface Area per complex (m_i)**	**K_{cat} (units: hr^{-1})**	**Specific Membrane Area (C_i) (units: (nm^2/gdw)**
PSII	279.35 nm^2	$6.24*10^6$	$4.4776*10^{-5}$
Cytochrome b6/f	128.8872 nm^2	$3.6*10^6$	$3.5802*10^{-5}$
PSI	91.3681 nm^2	$7.2*10^5$	$1.269*10^{-4}$

Table 3.2: Biophysical Values for Thylakoid Membrane

Variable	Value
Surface area to volume ratio ($R_{S/V}$)	0.16 nm^{-1}
Specific membrane area of cytoplasmic membrane (M_{cyt})	$1.6*10^6$ cm^2/g
Fraction of cell membrane that can accommodate proteins (f)	0.70
Cytoplasmic Membrane Budget (B)	0.112 nm^2/g

Since the C_i value for each protein complex describes the thylakoid membrane area required to achieve a flux of 1 mmol/gdw/hr, the total thylakoid membrane area occupied by a certain protein complex can be determined from multiplying the actual flux by the C_i value. Therefore, the constraint equation is:

$$X_{PSII}\,(4.4776*10^{-5}) + X_{cyt\ b6f}\,(3.5802*10^{-5}) + X_{PSI}\,(1.269*10^{-4}) < 0.112 \qquad (3.8)$$

Where C_i in parenthesis is in units of nm^2/g, B on the right-hand side of the inequality is in units of nm^2/g and X_{PSII}, $X_{cyt\ b6f}$, and X_{PSI} represent the fluxes going through each protein complex in unitless, scalar form.

3.4 Discussion

Knowing the constraint equation is particularly helpful for determining what flux values through each protein complex, and ultimately, how many protein complexes are allowable on a single thylakoid membrane. Using mixed-integer linear programming methods, the maximum total amount of each protein complex and the maximum amount of flux occurring on a single thylakoid membrane can be calculated.

Constraints can be implemented towards the genome-scale model of Phaeodactylum tricornutum using constraint-based reconstruction and analysis methods using Constraint-based reconstruction and analysis using Python (COBRApy). Adding an imaginary, constraining metabolite to the genome-scale model can be used to force flux down a certain path because it adds another row to the S matrix containing all the metabolites and reactions in the system. The stoichiometric coefficients of this imaginary metabolite, which represent the specific membrane area per protein complex, are then multiplied by the fluxes flowing through each membrane complex. Summation of the product of each specific membrane areas for each protein complex and their flux is then constrained to be less than or equal to the cytoplasmic membrane budget.

Chapter 4: Addressing Different Photosystem II, Cytochrome b6/f, Photosystem I Ratios

4.1 Introduction

In chapter 3, membrane constraints were calculated in order to modify the Phaeodactylum tricornutum genome-scale model for future research efforts. The constraint equation can be incorporated into the genome-scale model using mixed-integer linear programming techniques which will find the optimized possible fluxes as long as the membrane surface areas are constrained to the allowable bounds. From this ratio of optimized fluxes, the ratio of each protein complex can be determined.

However, another approach is to incorporate established ratios from the literature into the MASS Toolbox kinetic model, observe the effects of irradiance on various outputs as shown in Chapter 2, and study the deviations from the original network which was built assuming that the ratio of Photosystem II: Cytochrome b6/f: Photosystem I is 1:1:1. The MASS Toolbox model allows for simple modification of various parameters to observe the effects of different ratios of protein complexes.

4.2 Methods

Modifying the ratios of protein complexes in the MASS Toolbox model was done by changing the values of the initial conditions, since the initial conditions shown on table 2.4 are for a default protein complex ratio of 1:1:1. In high light conditions, T. weissflogii has a photosystem II: cytochrome b6/f: photosystem I complex ratio of

4.4:1.9:1[33]. T. weissflogii is a diatom, although a different species from Phaeodactylum tricornutum, but this protein ratio was incorporated in the MASS Toolbox model.

Since the photosystem II to photosystem I ratio was 4.4 to 1, all base initial condition values of photosystem II were multiplied by 4.4. The original initial value of Q_A went from 0.08994 to 0.395736, Q_AQ_B went from 0.910059 to 4.0042596, PQ went from 4 to 17.6. The number of cytochrome b6f complexes went from 1 to 1.9. All chlorophyll antenna sizes of photosystem II reactions centers were also multiplied by 4.4, but photosystem I antenna sizes remained the same. The model was then simulated and outputs were compared to those shown in Chapter 2.

4.3 Results

The photosynthetic rate as a function of irradiance and observing the effects of cyclic electron flow presence and absence as well as the effects of antenna size are shown in the following figure:

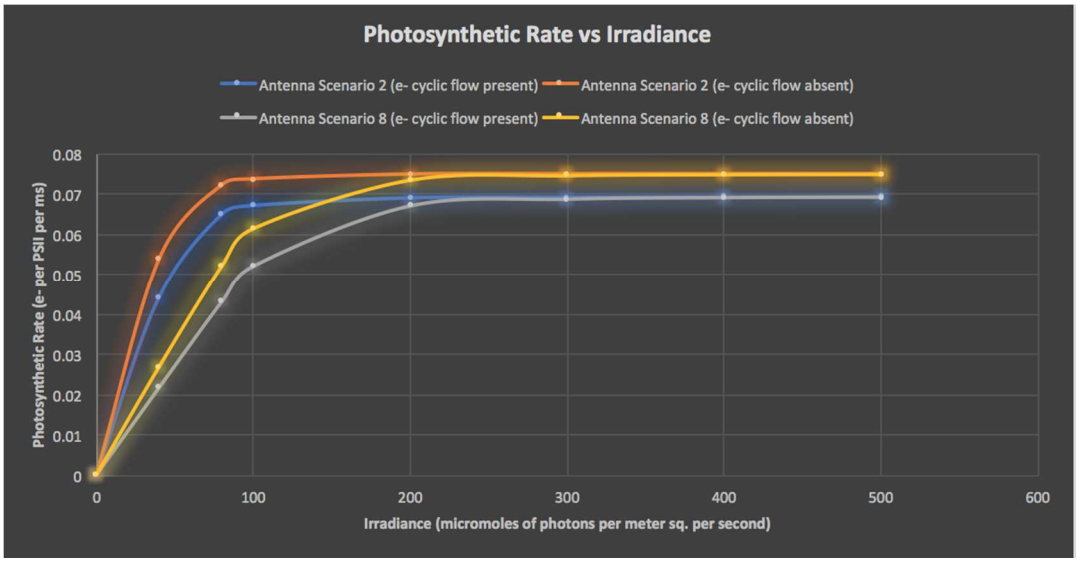

Figure 4.1: Photosynthetic Rate vs. Irradiance and Effects of Chlorophyll Antenna Size and Cyclic Electron Flow (Modified Protein Ratios). These values are compared to those in figure 2.29, in which the photosystem II to cytochrome b6/f to photosystem I protein ratio is 1:1:1

The figure below shows the output of the model with modified protein ratios in regards to photosystem I electron cyclic rate:

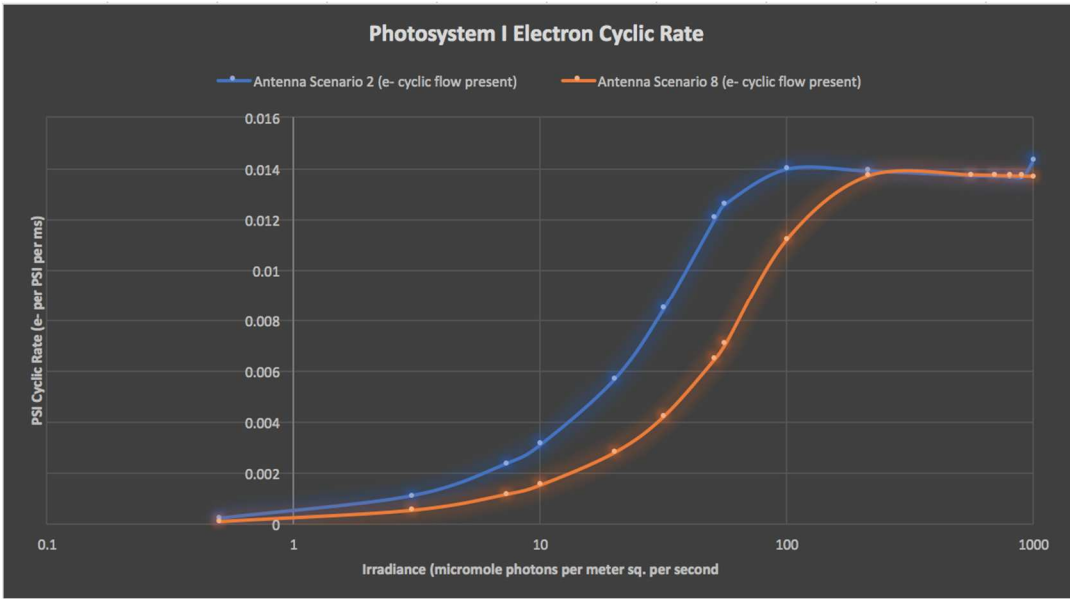

Figure 4.2: MASS Toolbox Model Output for PSI Cyclic Rate vs. Irradiance Levels (Modified Protein Ratios)

The figure below shows the proton yield as a function of irradiance and the effects of different chlorophyll antenna sizes using the modified protein ratio:

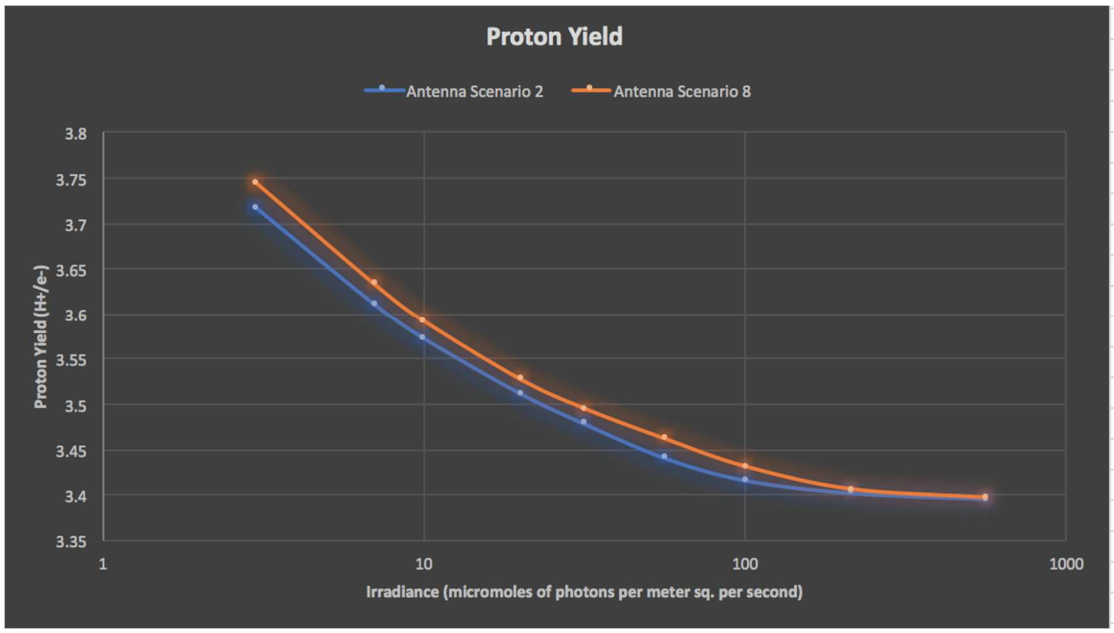

Figure 4.3: Proton-to-Electron Ratio as a Function of Irradiance (Modified Protein Ratio)

The following figure shows the photosynthetic rate as a function of irradiance and rate constant of ferredoxin oxidation:

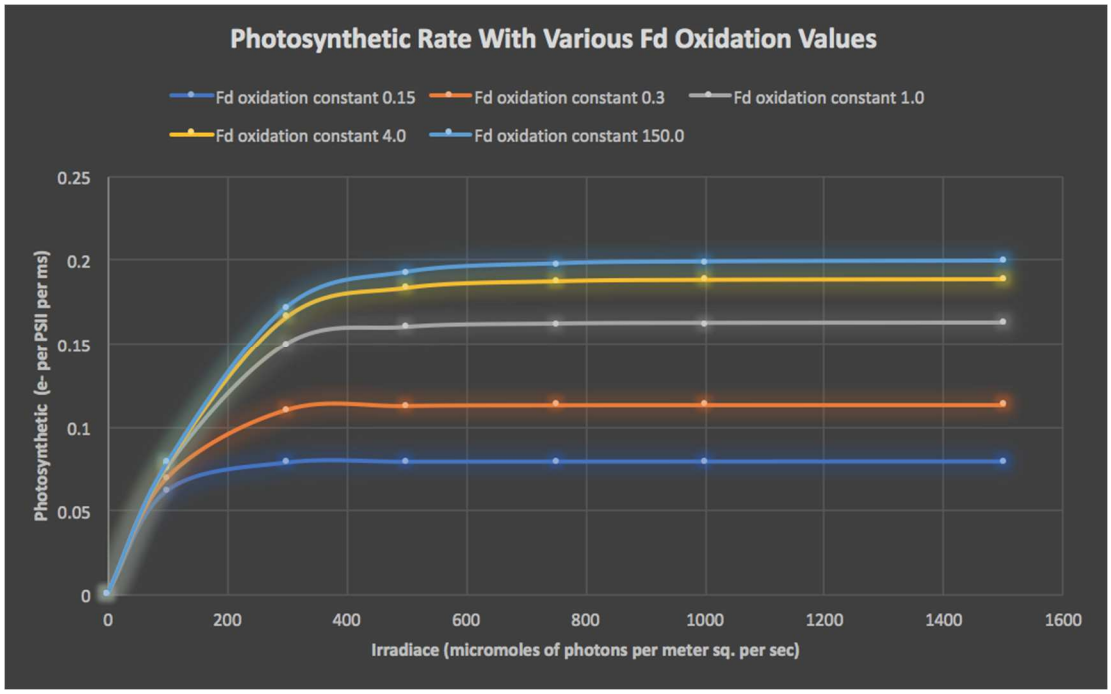

Figure 4.4: MASS Toolbox Output of Photosynthetic Rate vs. Irradiance with Varying Ferredoxin Oxidation Rate Constant (Modified Protein Ratios)

4.4 Discussion

Comparing the results from figure 2.25 and figure 4.1, the outputs values are almost identical. One possible reason is that even if there are more photosystem II complexes and cytochrome b6/f complex when compared to photosystem I, the photosystem I complex is already saturated with electrons flowing through it, therefore producing a bottleneck effect that prevents photosynthetic rate from increasing.

Comparing the results from figure 2.32 and figure 4.2 show that the profiles look very similar. However, closer examination of the outputs show that photosystem I cyclic rate plateaus at 0.014 e- per photosystem I per millisecond for the modified protein ratios while the photosystem I cyclic rate plateaus at approximately 0.0116 e- per photosystem I

per millisecond for the original protein ratios. One possible reason for the increased photosystem I cyclic rate with the modified protein ratios is that the slightly higher number of cytochrome b6f complexes relative to photosystem I mean that there is a greater number of metabolite X_{11}, otherwise known as the cytochrome b6/f complex itself, in the system. A higher number of X_{11} means that more electrons can flow down reaction bR11, indicative of cyclic electron flow, combine with more ferredoxin ion from photosystem I, and increase the photosystem I cyclic rate.

Comparing the results from figure 2.35 and figure 4.3 show that the profiles are drastically different. First, maximum values of proton yield per electron flowing into the system of modified protein ratios is higher than the original model. Secondly, the proton yield per electron in the modified protein ratios model shows that proton yield drops much faster as a function of irradiance as opposed to the original model. There could be multiple reasons for this deviation such as the increased number of cytochrome b6f complexes that impact the kinetics of the system. One plausible explanation is that physiological protein complex ratios have a higher proton to electron ratio at low irradiance levels, since this ultimately allows for more ATP produced. However, more realistically, to account for the different protein ratios, the biophysical rate values in figure 2.6 would most likely need to be modified. More research needs to be conducted in this specific portion.

Comparing the results from figure 2.24 to figure 4.4 show that the profiles are nearly identical for the system using both modified protein ratios and original protein ratios. The reason for this may be the same as why figures 2.25 and 4.1 are almost identical: regardless of the greater number of higher number of photosystem II and

cytochrome b6f, photosystem I ultimately acts as a bottleneck, limiting the number of photons that flow through PSII and limiting the maximum number of ferredoxin ions that can be oxidized to form NADPH.

Ideally, reconciling the use of membrane constraints in Chapter 3 to ultimately affect the Phaeodactylum tricornutum genome-scale model and determine the protein complex ratios that exist within the thylakoid membrane and then implement those ratios into the MASS Toolbox kinetic model to produce a solution space of fluxes to overlap in both models is the end objective. Although theoretically possible, there are many complex factors that need to be taken into account. Simply changing the initial conditions of various metabolites and optical surface area of photosystem reaction centers may be inadequate to describe protein ratios. Several of the biophysical constants shown in figure 2.6 were simply provided by Kroon and Thoms, but a further search into the literature has shown that the derivation of many of these values is computationally intensive. Therefore, more work would need to be conducted to observe how the protein complex ratios affect these values, which in turn would affect the overall dynamics of the system. Additionally, the MASS Toolbox kinetic model can be modified and tailored for specific use towards Phaeodactylum tricornutum by adding parameters that include chlorophyll c. Further experimental efforts can be conducted to validate the solution overlap between the Phaeodactylum tricornutum genome-scale model and kinetic model.

References

1. Daboussi, Fayza, Sophie Leduc, Alan Marechal, Gwendoline Dubois, Valeria Guyot, Christophe Perez-Michaut, Alberto Amato, Angela Falciatore, Alexandre Juillerat, Marine Beurdeley, Daniel F. Voytas, Laurent Cavarec, and Philippe Duchateau. "Genome Engineering Empowers the Diatom Phaeodactylum Tricornutum for Biotechnology." *Nature Communications* (2014): n. pag. Web.

2. Levering J, Broddrick J, Dupont CL, Peers G, Beeri K, Mayers J, et al. Genome-Scale Model Reveals Metabolic Basis of Biomass Partitioning in a Model Diatom. *PLoS ONE* 11(5): e0155038. doi:10.1371/journal.pone.0155038

3. Wagner, Heiko, Torsten Jakob, Johann Lavaud, and Christian Wilhelm. "Photosystem II Cycle Activity and Alternative Electron Transport in the Diatom Phaeodactylum Tricornutum under Dynamic Light Conditions and Nitrogen Limitation." *Photosynthesis Research* 128.2 (2016): 151-61. Web.

4. Kitano, H. "Systems Biology: A Brief Overview." *Science* 295.5560 (2002): 1662-664. Web.

5. Falkowski, Paul G., and John A. Raven. Aquatic Photosynthesis. 2nd ed. Princeton: Princeton UP, 2007. Print

6. Fotiadis, D., P. Qian, A. Philippsen, P.A. Bullough, A. Engel, and C. N. Hunter. "Structural Analysis of the Reaction Center Light-harvesting Complex I Photosynthetic Core Complex of Rhodospirillum Rubrum Using Atomic Force Microscopy." *Journal of Biological Chemistry* 279.3 (2003): 2063-068. Web.

7. Krause, G., and E. Weis. "Chlorophyll Fluorescence And Photosynthesis: The Basics." *Annual Review of Plant Physiology and Plant Molecular Biology* 42 (1991): 313-49. Web.

8. Björkman, Olle, and Barbara Demmig. "Photon Yield of O2 Evolution and Chlorophyll Fluorescence Characteristics at 77 K among Vascular Plants of Diverse Origins." *Planta* 170.4 (1987): 489-504. Web.

9. Nelson, David L., and Michael M. Cox. *Lehninger Principles of Biochemistry.* 5th ed. New York: W. H. Freeman, 2008. Print.

10. Satoh, Kimiyuki, and Thomas J. Wydrzynski. *Photosystem II: The Light-Driven Water:Plastoquinone Oxidoreductase.* New York: Springer, 2005. Web.

11. Suga, Michihiro, Fusamichi Akita, Kunio Hirata, Go Ueno, Hironori Murakami, Yoshiki Nakajima, Tetsuya Shimizu, Keitaro Yamashita, Masaki Yamamoto, Hideo Ago, and Jian-Ren Shen. "Native Structure of Photosystem II at 1.95 Å Resolution Viewed by Femtosecond X-ray Pulses." *Nature* 517 (2014): 99-103. Web.

12. Whitmarsh, John, and Govindjee. "Photosystem II." *Encyclopedia of Life Sciences* (2002). Web

13. Gross, Elizabeth L. "Plastocyanin: Structure and Function." *Photosynthesis Research* 37.2 (1993): 103-16. Web.

14. Palsson, Bernhard Ø. *Systems Biology Constraint-Based Reconstruction and Analysis*. Cambridge: Cambridge UP, 2015. Print.

15. Graige, M. S., G. Feher, and M. Y. Okamura. "Conformational Gating of the Electron Transfer Reaction $Q_A^{-\cdot}Q_B$ -> $Q_AQ_B^{-\cdot}$ in Bacterial Reaction Centers of *Rhodobacter sphaeroides* Determined by a Driving Force Assay." *Proceedings of the National Academy of Sciences* 95.20 (1998): 11679-1684. Web.

16. Baird, James K. "A Generalized Statement of the Law of Mass Action." *Journal of Chemical Education* 76.8 (1999): 1146. Web.

17. Dekker, Jan P., and Rienk Van Grondelle. "Primary Charge Separation in Photosystem II." *Photosynthesis Research* 63.3 (2000): 195-208. Web

18. Wijn, Rik De, and Hans J. Van Gorkom. "Kinetics of Electron Transfer from Q_A to Q_B in Photosystem II" *Biochemistry* 40.39 (2001): 11912-1922. Web.

19. Kroon, Bernd M. A., and Silke Thoms. "From Electron to Biomass: A Mechanistic Model Describe Phytoplankton Photosynthesis And Steady-State Growth Rates" *Journal of Phycology* 42.3 (2006): 593-609. Web.

20. Cramer, W.a., S.e. Martinez, P.n. Furbacher, D. Huang, and J.I. Smith. "The Cytochrome B6f Complex." *Current Opinion in Structural Biology* 4.4 (1994): 536-44. Web.

21. Webber, Andrew N., and Wolfgang Lubitz. "P700: The Primary Electron Donor of Photosystem I" *Biochimica Et Biophysica Acta (BBA) – Bioenergetics* 1507. 1-3 (2001): 61-79. Web.

22. Joliot, Pierre, and Anne Joliot. "In Vivo Analysis of the Electron Transfer within Photosystem I: Are the Two Phylloquinones Involved?" *Biochemistry* 38.34 (1999): 11130-1136. Web.

23. Guglielmi, Gerard, Johann Lavaud, Bernard Rousseau, Anne-Lise Etienne, Jean Houmard, and Alexander V. Ruban. “The Light-harvesting Antenna of the Diatom Phaeodactylum Tricornutum. Evidence for a Diadinoxanthin-binding Subcomplex.” *FEBS Journal* 272.17 (2005): 4339-348. Web.

24. Price, Nathan D., Jason A. Papin, Christophe H. Schilling, and Bernhard O. Palsson. “Genome-scale Microbial in Silico Models: The Constraints-based Approach.” *Trends in Biotechnology* 21.4 (2003): 162-69. Web.

25. Zhuang, K., G. N. Vemuri, and R. Mahadevan. “Economics of Membrane Occupancy and Respiro-fermentation.” *Molecular Systems Biology* 7.1 (2014): 500. Web.

26. Rögner, M., J.P. Dekker, E.j. Boekema, and H.T. Witt. “Size, Shape and Mass of the Oxygen-evolving Photosystem II Complex from the Thermophilic Cyanobacterium Synechococcus sp.” *FEBS Letters 219.1 (1987): 207-11. Web.*

27. Ballottari, Matteo, Julien Girardon, Luca Dall’osto, and Roberto Bassi. “Evolution and Functional Properties of Phytosystem II Light Harvesting Complexes in Eukaryotes.” *Biochimica Et Biophysica Acta (BBA) – Bioenergetics* 1817.1 (2012): 143-57. Web

28. Cramer, William. “Cramer Lab Project Page: Cytochrome B6f Complex.” *Structure-Function Studies of the Cytochrome B6f Complex.* Purdue University, 2010. Web. 20 May 2017.

29. Miller, Susan, Arthur M. Lesk, Joël Janin, and Cyrus Chothia. “The Accessible Surface Area and Stability of Oligomeric Proteins.” *Nature* 328 (1987): 834-36. Web.

30. Kitmitto, A., A. O. Mustafa, A. Holzenburg, and R. C. Ford. “Three-dimensional Structure of Higher Plant Photosystem I Determined by Electron Crystallography.” *Journal of Biological Chemistry* 273.45 (1998): 29592-9599. Web.

31. Kirchhoff, Helmut. “Diffusion of Molecules and Macromolecules in Thylakoid Membranes.” *Biochimica Et Biophysica Acta (BBA) – Bioenergetics* 1837.4 (2014): 495-502. Web.

32. Kirchhoff, Helmut. “Significance of Protein Crowding, Order and Mobility for Photosynthetic Membrane Functions.” *Biochemical Society Transactions* 36.5 (2008): 967-70. Web.

33. Strzepek, Robert F., and Paul J. Harrison. “Photosynthetic Architecture Differences in Coastal and Oceanic Diatoms.” *Nature* 431 (2004): 689-92. Web.