

RECONSTRUCTION OF A GENOME-SCALE METABOLIC MODEL FOR
AUXENOCHLORELLA PROTOTHECOIDES

by
Jacob Tamburro

Copyright by Jacob Tamburro, 2025

All Rights Reserved

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Quantitative Bioscience Engineering).

Golden, Colorado

Date _____

Signed: _____

Jacob Tamburro

Signed: _____

Dr. Nanette Boyle
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____

Dr. Nanette Boyle
QBE Program Director
Quantitative Biosciences and Engineering

ABSTRACT

Microalgae show strong potential as alternative energy platforms for biofuel production and high-value product synthesis, yet the complexity of photosynthetic metabolism has posed significant modeling and engineering challenges. Recent developments in genome-scale metabolic models (GEMs) have substantially advanced our understanding of microalgal biology by integrating omics data, improving light-harvesting simulations, and automating large portions of the reconstruction process. Building upon these methodological gains, a new GEM *iPro4643* was reconstructed for *Auxenochlorella protothecoides*, capturing the organism's metabolic versatility under autotrophic, mixotrophic, and heterotrophic conditions. This model highlights both shared pathways with other algal species and features such as rapid heterotrophic growth and energy storage dynamics that are especially relevant to industrial-scale applications. Additionally, an in-depth ^{13}C metabolic flux analysis (MFA) of heterotrophically grown *A. protothecoides* reveals that actual carbon flux is similar to that predicted by the model while diverging from model predictions optimized strictly for growth, pointing to an underestimation of flux toward starch. These findings underscore the need for objective functions and modeling frameworks that incorporate energy storage and regulatory constraints, particularly under fluctuating environmental conditions. By uniting insights from state-of-the-art GEM methodologies, a refined and experimentally validated metabolic model, and empirical MFA data, this body of work paves the way for more accurate predictive capabilities and targeted metabolic engineering strategies aimed at enhancing lipid productivity and other commercially valuable bioproducts in microalgae.

TABLE OF CONTENTS

ABSTRACT	iii
LIST OF FIGURES.....	vii
LIST OF TABLES	viii
LIST OF ABBREVIATIONS	ix
ACKNOWLEDGMENTS.....	xi
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 CHARTING THE STATE OF GEMS IN MICROALGAE: PROGRESS, CHALLENGES, AND INNOVATION.....	4
2.1 Abstract.....	4
2.2 Introduction.....	4
2.3 <i>Chlamydomonas reinhardtii</i> : A Keystone Species for Microalga GEM Reconstruction	7
2.4 Challenges and Limitations of Algal Genome-scale Metabolic Models.....	8
2.5 Automation of Model Reconstruction.....	10
2.6 Modeling Light Harvesting.....	11
2.7 Models With a Focus on Reactions Outside of Carbon Metabolism	12
2.8 Model Robustness.....	13
2.9 Integration of Additional Omics Data into Models.....	13
2.10 Dynamic Metabolic Modeling	14
2.11 Prospects for Future Microalgal Genome-scale Metabolic Models.....	15
2.12 Conclusion	16
CHAPTER 3 RECONSTRUCTION OF A GENOME-SCALE METABOLIC MODEL FOR <i>AUXENOCHLORELLA PROTOTHECOIDES</i>	18
3.1 Abstract.....	18
3.2 Introduction.....	19
3.3 Materials and Methods.....	20
3.3.1 Strains and growth conditions	20
3.3.2 Biomass Composition.....	20
3.3.3 Model Reconstruction.....	24

3.3.4	Flux Balance Analysis (FBA).....	24
3.4	Results	25
3.4.1	Network Reconstruction and Curation	25
3.4.2	Biomass Composition and biomass function.....	26
3.4.3	Experimentally measured model constraints.....	28
3.4.4	Flux distribution	29
3.4.5	Essential Gene Analysis	33
3.5	Discussion.....	33
3.5.1	Reconstruction of the genome-scale metabolic model	33
3.5.2	Biomass composition.....	34
3.5.3	Flux distribution	35
3.6	Conclusion	37
CHAPTER 4	¹³ C METABOLIC FLUX ANALYSIS OF HETEROTROPHICALLY GROWN AUXENOCHLORELLA PROTOTHECOIDES UTEX 250	38
4.1	Abstract.....	38
4.2	Introduction.....	39
4.3	Method	40
4.3.1	¹³ C Growth Conditions and Sampling	40
4.3.2	¹³ C Metabolic Flux Analysis	41
4.4	Results.....	42
4.4.1	¹³ C Metabolic Flux Analysis	42
4.5	Discussion.....	43
4.6	Conclusion	45
CHAPTER 5	CONCLUSION	46
5.1	Future work.....	47
REFERENCES	48
APPENDIX A	64

A.1	^{13}C MFA Fluxes.....	64
A.2	^{13}C Network.....	69

LIST OF FIGURES

Figure 2.1	Historical perspective on the generation of algal genome-scale metabolic models, organized by species and year of publication.....	7
Figure 2.2	Challenges for algae genome scale model construction include the need for automated algorithms for model construction specifically for algae, improved light modeling and integrating -omics data to improve predictability. Addressing these will enable the design of dynamic models that can better predict growth in dynamic conditions, such as day/night cycles.....	10
Figure 3.1	Compartmental data for reactions and metabolites in iApro4643 as well as transport reactions.	26
Figure 3.2	Macromolecule biomass composition of <i>A. protothecoides</i> under autotrophic (APM1-Glucose media), mixotrophic and heterotrophic (APM1 media) growth conditions	27
Figure 3.3	Total measured fatty acid composition of <i>A. protothecoides</i> as determined by FAMES measured on GC-MS under autotrophic, mixotrophic and heterotrophic conditions.....	27
Figure 3.4	In silico simulated autotrophic growth of <i>A. protothecoides</i> grown on atmospheric CO ₂ (APM1-glucose media) using experimental data to constrain the model. The predicted fluxes are normalized to a CO ₂ uptake of 100mmol/g·DW·hr (100mmol total carbon) for ease of comparison with the other trophic condition results.	30
Figure 3.5	In silico simulated mixotrophic growth of <i>A. protothecoides</i> grown on glucose (APM1 media) using experimental data to constrain the model. The predicted fluxes are normalized to a glucose uptake of 16.6mmol/g·DW·hr (100mmol total carbon) for ease of comparison with the other trophic condition results.	31
Figure 3.6	In silico simulated heterotrophic growth of <i>A. protothecoides</i> grown on glucose (APM1 media) using experimental data to constrain the model. The predicted fluxes are normalized to a glucose uptake of 16.6mmol/g·DW·hr (100mmol total carbon) for ease of comparison with the other trophic condition results.	32
Figure 4.1	Experimentally determined fluxes of <i>A. protothecoides</i> grown on glucose (APM1 media) using carbon enrichment data and INCA. The fluxes are normalized to a glucose uptake of 16.6mmol/gDW-hr (100mmol total carbon) for ease of comparison with the GEM results. The four compartments are labelled on the flux map.	43

LIST OF TABLES

Table 2.1	Algae species that currently have GEMs reconstructed for them as well as research and cell factory applications of each species.	5
Table 2.2	The table displays the genome-scale metabolic models (GEMs) currently published, organized by species and year of publication from oldest to newest.	15
Table 3.1	Experimentally measured mid-exponential growth rate (gDW·h), Glucose uptake (mmol/gDW·h) rate and ethanol excretion rates (mmol/gDW·h) for each trophic condition. Glucose uptake measurements were not taken on autotrophic cultures as there was no glucose in the media. Autotrophic and mixotrophic cultures were assessed for Ethanol content however the values were below the limit of detection on the Cedex bioanalyzer.	28
Table 3.2	The table shows the exact lower and upper bound constraints that were put on the metabolic models when running the pFBA.	29
Table A.1	¹³ C MFA experimentally determined fluxes	64

LIST OF ABBREVIATIONS

Genome-scale Metabolic Model.	GEM
Flux Balance Analysis.	FBA
Parsimonious Flux Balance Analysis.	pFBA
Metabolic Flux Analysis.	MFA
Carbon Dioxide.	CO ₂
Transient Metabolic Model.	TMM
Protein Constrained Flux Balance Analysis.	PC-FBA
Rapid Annotation of Photosynthetic Systems.	RAPS
Variability Flux Sampling.	VFS
Flux Variability Analysis.	FVA
Protein-constrained Metabolic Model.	PC-model
Optical Density.	OD
Sulfuric Acid.	H ₂ SO ₄
Sodium Hydroxide.	NaOH
Hydrochloric Acid.	HCl
N-(tertbutyldimethylsilyl)-N-methyl-triflouracetamide.	MTBSTFA
Fatty Acid Methyl Ester.	FAME
Deoxyribonucleic Acid.	DNA
Ribonucleic Acid.	RNA
Basic Local Alignment Search Tool.	BLAST
Energy-generating Cycle.	EGC
Metabolic Model Tests.	MEMOTE
Tricarboxylic Acid.	TCA
Glyceraldehyde-3-Phosphate.	G3P
Pentose Phosphate Pathway.	PPP
Nicotinamide Adenine Dinucleotide Phosphate.	NADPH
Methoxamine.	MOX
Isotopomer Network Compartmental Analysis.	INCA
Isocitrate.	ICT

ADP-Glucose Pyrophosphorylase.	AGP
UDP-Glucose Pyrophosphorylase.	UGP
Acetyl-CoA Carboxylase.	ACC
Diacylglycerol Acyltransferase.	DGAT
Isotopically Nonstationary Metabolic Flux Analysis.	INST-MFA
Major Lipid Droplet Protein 1.	MLDP1
Sedoheptulose-1,7-Bisphosphatase.	SBPase

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Boyle, for giving me the opportunity to engage in exciting research as well as for providing an environment where I can grow as a scientist. Working in your lab has broadened my scientific perspective and your mentorship has allowed me to grow both academically and personally.

I am also grateful to Dr. Posewitz, the head of my previous lab, for first introducing me to the world of research. Your encouragement and insight allowed me to discover my passion for science and I appreciate the foundation you provided for my continued growth.

Finally, I wish to thank Dr. Burch, the postdoctoral researcher I worked under in Dr. Posewitz's lab, for mentoring me during my formative research experiences. Your patience in teaching me new techniques and the trust that you showed in me from the beginning were integral to my continued passion for research.

To each of you, thank you for inspiring me to push boundaries, guiding me through challenges, and shaping my journey into research. Your collective mentorship has made this work possible and far more valuable than you.

CHAPTER 1

INTRODUCTION

Quality of life in modern society is, at present, inextricably linked with the availability of fossil fuels. Transportation and electricity generation are currently met primarily by burning fossil fuels such as coal, natural gas and petroleum. Additionally, a substantial number of products that we rely on are derived from petroleum, from most plastics to pharmaceuticals, road surfacing and cleaning products¹. However, fossil fuels are a finite resource with many sources expected to run out by 2100 at current consumption rates². In addition to the looming threat of resource depletion, the combustion of fossil fuels also release greenhouse gases that trap heat from the sun in the earth's atmosphere³.

Carbon dioxide (CO₂) is the greenhouse gas that is most widely reported on as it is the most abundantly produced greenhouse gas with 35.8 gigatons of CO₂ produced in 2023 alone⁴. A large portion of the CO₂ emitted is produced by burning fossil fuels and has shifted the composition of the atmosphere enough to cause a measurable increase in global average temperatures³. Despite the considerable harm that has already occurred due to anthropogenic climate change, work can still be done to reduce future impact by transitioning to more sustainable production methods.

One alternative to fossil fuels being explored is the increased use of cell factories. Cell factories are microbial cells that are optimized to increase the production of desired products. Using microorganisms to generate useful products is not a new concept which has been utilized by humans since the beginning of civilization in the production of fermented products such as beer and bread. Of the microorganisms which could be harnessed for use as fuel, microalgae present a particularly attractive platform. Microalgae boast a high photosynthetic efficiency, meaning they can convert sunlight, water, and carbon dioxide into biomass at rates exceeding those of many terrestrial plants⁵. Unlike conventional crops that require arable land, microalgae can be cultivated in bioreactors or ponds on non-arable land, thereby avoiding direct competition with food crops. Besides presenting an attractive platform for biofuel development⁶, microalgae have the potential to sustainably produce biopolymers and biopharmaceuticals⁷. Moreover, microalgae-based systems can be used for waste processing and bioremediation, capturing and converting waste streams into valuable byproducts.

At the core of using living organisms as cell factories is metabolic engineering. Metabolic engineering involves modifying and optimizing an organism's biochemical pathways to boost the

production of a specific substance. Drawing on biology, genetics, biochemistry, and systems engineering, scientists manipulate genes and regulatory circuits to redirect carbon and other nutrients, thereby increasing yields of the desired product⁸. This process serves as a linchpin for transforming naturally occurring microalgae and other microorganisms into more efficient, robust, and specialized cell factories.

Genome-scale metabolic models (GEMs) are comprehensive, stoichiometric networks derived from an organism's annotated genome. They detail all known metabolic reactions and map how substrates flow through these reactions to generate products. In practice, GEMs act as blueprints, helping metabolic engineers identify bottlenecks or alternate pathways for product synthesis. While GEMs are invaluable for generating hypotheses, they become far more powerful when augmented by real-world experimental data. Integrating measurements such as carbon uptake rates, metabolite excretion profiles, and biomass composition yields models that more accurately represent physiological states. For example, a biomass formula quantifies the building blocks (lipids, proteins, carbohydrates) essential for growth; incorporating this biomass equation narrows the model's solution space and aligns predicted metabolic fluxes with those observed in living cultures.

To employ GEMs in practice, researchers often use Flux Balance Analysis (FBA), a mathematical method that calculates the flow of metabolites through the network under steady-state conditions. Relying on linear optimization, FBA determines how cells allocate resources to maximize or minimize a chosen objective function, most commonly biomass production or a target metabolite. By systematically testing all possible flux distributions within the stoichiometric constraints of the GEM, FBA can predict how microalgae might shift their metabolism when engineered to overproduce lipids or pigments. Although maximizing growth is typically the default objective, scientists can redefine the objective to prioritize the production of a specific compound, such as a biofuel precursor or a high-value metabolite. By changing the objective function and comparing results trade-offs between growth and product yield can be explored.

However, traditional FBA assumes that cells may use any viable combination of metabolic pathways to meet the objective, which can complicate flux predictions. In practice, cells often favor simpler, less energetically demanding routes, which is an aspect not always captured by standard FBA. Parsimonious Flux Balance Analysis (pFBA) refines this approach by penalizing unnecessary fluxes, producing results more consistent with experimentally observed metabolic

behavior. By applying the principle of minimal metabolic effort, pFBA generates flux distributions that more accurately mirror real cellular processes.

To experimentally determine the metabolism of an organism *in vivo* researchers can utilize ^{13}C metabolic flux analysis (MFA). ^{13}C is an isotope of carbon that is different from the more abundant ^{12}C isotope of carbon with the ^{13}C isotope having an additional neutron increasing the atomic mass. Based on the difference in atomic mass, the presence of ^{12}C or ^{13}C carbon can be determined in molecules when measured with mass spectrometry (MS). MFA determines the flux of an organism in metabolic steady state at mid exponential growth by feeding an organism isotopically labelled carbon and tracking the distribution of ^{13}C through metabolites. To determine the distribution of ^{13}C in cellular metabolites, metabolites are extracted and measured by gas chromatography mass spectrometry (GC-MS). The resulting ^{13}C composition of metabolites is then used with a metabolic network and software such as isotopic network compartment analysis (INCA)⁹ that models fluxes through the network that best fit the experimentally determined ^{13}C carbon enrichments in the metabolites measured.

This thesis explores methods for determining the intricacies of metabolism in algae. This is done by first reviewing GEMs available for microalgae as well as the challenges associated with GEM reconstruction and methods that researchers have overcome these challenges, followed by chapters on the reconstruction of our GEM and ^{13}C MFA of *Auxenochlorella protothecoides*.

CHAPTER 2

CHARTING THE STATE OF GEMS IN MICROALGAE: PROGRESS, CHALLENGES, AND INNOVATION

Jacob Tamburro¹, Nanette R. Boyle^{1,2}

2.1 Abstract

Genome-scale metabolic models (GEMs) provide a systems-level framework for understanding and engineering microalgal metabolism. This review explores the evolution of GEMs in microalgae, highlighting advances in light modeling, automation, and multi-omics integration. Special emphasis is placed on *Chlamydomonas reinhardtii* as a model species. Limitations of current models, particularly for microalgae, are discussed, alongside promising developments in dynamic modeling and machine learning. Together, these innovations chart a path toward more predictive, adaptable GEMs that can accelerate biotechnological applications of microalgae in sustainable production systems.

2.2 Introduction

Microalgae have demonstrated significant potential for the sustainable production of biofuels and other valuable products. As cell factories, microalgae can be optimized for biofuel production⁶, wastewater processing¹⁰, and the creation of high-value bioproducts such as nutraceuticals and pharmaceuticals¹¹. Microalgae have also been shown to have a photosynthetic efficiency of 4.4%¹², considerably higher than the photosynthetic efficiency of terrestrial plants which is typically between 1-2%¹³. The advantage in photosynthetic efficiency for microalgae then translates to higher growth rates and annual yields compared to terrestrial plants¹⁴. Additionally, microalgae offer a sustainable method for producing a wide array of bio-based products (see Table 2.1), many of which can serve as sustainable alternatives to petroleum-based products¹⁵.

¹ Colorado School of Mines, Quantitative Bioscience Engineering

² Colorado School of Mines, Chemical and Biological Engineering

Table 2.1 Algae species that currently have GEMs reconstructed for them as well as research and cell factory applications of each species.

Species	Research and Cell Factory Applications
<i>Auxenochlorella protothecoides</i>	Biofuel ¹⁶ , nutraceutical and pharmaceutical production ¹⁷
<i>Chlamydomonas reinhardtii</i>	Model green microalgae ¹⁸ , biofuel ¹⁹ , nutraceuticals and food supplements production ²⁰
<i>Chlorella variabilis</i>	Wastewater remediation, Nutraceutical ²¹ and biofuel production ²²
<i>Chlorella vulgaris</i>	Nutraceutical ²³ and biofuel production ²⁴
<i>Chromochloris zofingiensis</i>	Nutraceutical ²⁵ and biofuel production ²⁶
<i>Dunaliella salina</i>	Wastewater remediation, high salinity tolerance ²⁷ and nutraceutical production ²⁸
<i>Emiliana huxleyi</i>	Broad salinity tolerance ²⁹ and nutraceutical production ³⁰
<i>Fragilariopsis cylindrus</i>	Extremophile ³¹ and nutraceutical production ^{32,33}
<i>Haematococcus pluvialis</i>	Nutraceutical ³⁴ and biofuel production ³⁵
<i>Isochrysis galbana</i>	Broad salinity tolerance ³⁶ , nutraceutical ³⁷ and biofuel production ³⁸
<i>Nannochloropsis gaditana</i>	Nutraceutical and biofuel production ³⁹
<i>Nannochloropsis salina</i>	Nutraceutical ⁴⁰ and biofuel production ⁴¹
<i>Phaeodactylum tricornerutum</i>	Model diatom, nutraceutical and biofuel production ⁴²
<i>Scenedesmus obliquus</i>	Wastewater remediation, nutraceutical ⁴³ and biofuel production ⁴⁴
<i>Schizochytrium limacinum</i>	Nutraceutical ⁴⁵ and biofuel production ⁴⁶
<i>Thalassiosira pseudonana</i>	First microalgae sequenced ⁴⁷ , wastewater remediation, nutraceutical ⁴⁸ and biofuel production ⁴⁹

Unfortunately, algae have not fully realized their potential as cellular factories due to a number of challenges associated with economical production at large scale⁵⁰. A main driver of the overall cost of production is the productivity of the algae (growth rate x production rate) which influences the choice of photobioreactors, separation and labor costs⁵¹⁻⁵³. Maximizing productivity can lead to lower downstream costs, and one tool that has been proven to be successful in rerouting carbon in metabolism is metabolic engineering, specifically the use of metabolic models, to predict and implement genetic changes that can improve overall productivity⁵⁴ and product specific productivity^{55,56}.

Computational tools provide powerful means to investigate the complexities of metabolism. Among the computational methods employed in metabolic engineering, genome-scale metabolic models (GEMs) stand out due to their relative ease of implementation and comprehensive, systems level approach. GEMs are *in silico* representations of an organism's metabolic capacity based on the organism's sequenced genome, enumerating all reactions and metabolites encoded within. Experimental data, such as carbon uptake and excretion, biomass composition and growth rate can be used to constrain the model⁵⁷. Dramatically decreasing costs for high quality genome sequencing has led to increased sequence data for GEM reconstruction⁵⁸, and advances in genome annotation have enabled more complete simulations of metabolic processes. GEMs can

be used to identify gene knockouts that lead to increased yield or productivity. They can also be used to evaluate changes in yield due to the incorporation of heterologous metabolic pathways, narrowing the potential mutants to be screened in the lab and drastically decreasing research and development investment⁵⁹. By representing the entire metabolic capacity of an organism, GEMs have also been used to identify genetic targets that are not easy to predict *a priori* as having an impact on the productivity of a specific product^{60,61}. The utilization of GEMs is not limited to screening genetic changes, GEMs can additionally be applied to understand how an organism will respond to environmental changes. These applications include media optimization and predictions on the most crucial nutrients for growth⁶². GEMs also can be utilized to rapidly provide predictions on the changes that varying growth conditions will have phenotypically⁶³. *In silico* studies provide an effective alternative to traditional experiments, enabling researchers to investigate the impact of thousands of genetic or environmental changes in a fraction of the time it takes to create and characterize in lab^{64,65}.

GEMs have been extensively employed to study metabolism across a wide range of organisms, with the majority of existing literature and models focused on heterotrophic systems such as bacteria and yeast. These models have proven highly effective for simulating metabolic fluxes, predicting genetic engineering targets, and optimizing growth conditions. However, applying GEMs to photoautotrophic organisms, particularly eukaryotic microalgae, presents a distinct set of challenges. These include the need to account for light-dependent metabolism, complex regulatory networks, and dynamic environmental interactions such as diel light cycles. In this review, we examine the specific difficulties encountered when constructing and utilizing GEMs for photoautotrophic microalgae, as well as the current limitations that hinder their broader adoption and predictive accuracy. A dedicated section explores the role of *Chlamydomonas reinhardtii*, which has emerged as a cornerstone species in algal systems biology and a model for developing and refining GEMs in microalgae. Finally, we highlight future directions in GEM research, including the integration of dynamic modeling, multi-omics data, and machine learning techniques, all of which are poised to significantly advance the utility of GEMs in both fundamental research and applied biotechnology.

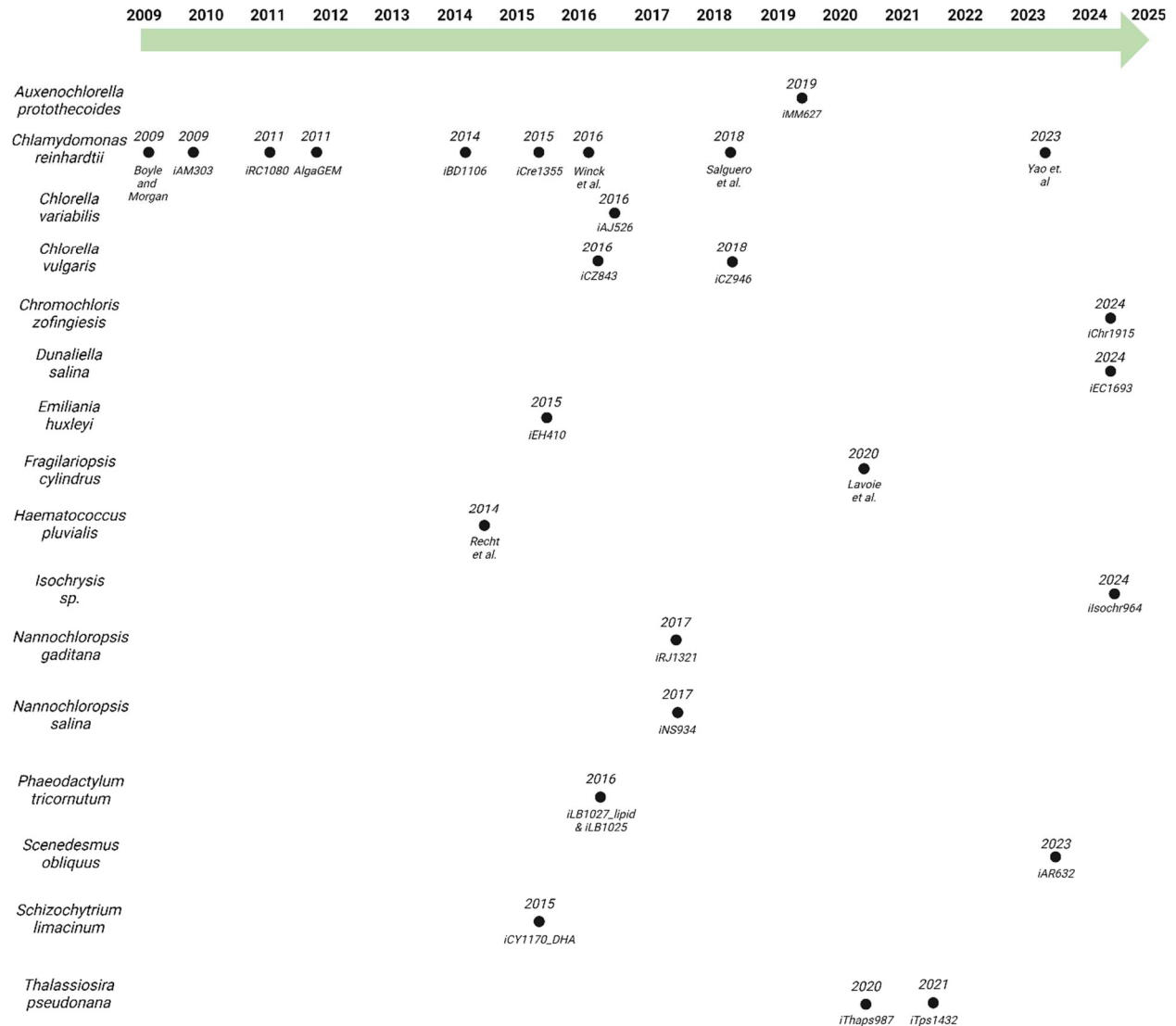


Figure 2.1 Historical perspective on the generation of algal genome-scale metabolic models, organized by species and year of publication.

2.3 *Chlamydomonas reinhardtii*: A Keystone Species for Microalga GEM Reconstruction

Chlamydomonas reinhardtii has received extensive attention in scientific research¹⁸, emerging as a pivotal organism for studying microalgae and aquatic photosynthetic systems⁶⁶. As a model green microalga, *C. reinhardtii* has served as the foundation for GEMs in algal species. The first GEM for *C. reinhardtii* was developed by Boyle and Morgan in 2009⁶⁷, marking the first GEM constructed for any algal species (see Figure 2.1). Another noteworthy GEM is *iCre1355*⁶⁸, which has served as a foundational platform for subsequent models. Derived like many of the currently available GEMs from the earlier *iRC1080*⁶⁹, *iCre1355*⁶⁸ incorporates updates based on

improvements made to the annotation of the genome, rectifying inaccuracies in gene-protein reactions. This improved model has been utilized to predict reaction kinetics⁷⁰ and growth under varying light conditions⁷¹. *iCre1355*⁶⁸ was also utilized in the development of the first diurnal metabolic model developed by Metcalf and Boyle⁷². This transient metabolic model (TMM) incorporated quantitative, time dependent transcriptomic data to constrain the availability of the associated gene products and metabolic reactions and more accurately predict growth in diurnal conditions. The most recent GEM for *C. reinhardtii*, developed by Yao et al.⁷³, merged the *iCre1355* and *iGR774* models, replacing the chloroplast reactions in *iCre1355*⁶⁸ with the more detailed *iGR774*⁷⁴ chloroplast specific model to create a more comprehensive framework. Yao et al. utilized the protein constrained flux balance (PC-FBA) demonstrating the first implementation of proteomic data into an microalgal GEM. With these advancements, *C. reinhardtii*'s GEMs continue to be at the forefront of advancing algal biotechnology, significantly contributing to the understanding of microalgal metabolism and algal GEM reconstructions.

2.4 Challenges and Limitations of Algal Genome-scale Metabolic Models

Genome-scale Metabolic Models (GEMs) are a powerful and rapidly advancing tool for understanding cellular metabolism, however, like any complex modeling approach, there are challenges that researchers continue to address to unlock their full potential. A challenge across all organisms is inaccurate or incomplete genome annotations, which leads to gaps that need to be manually filled. This issue is particularly pronounced in photoautotrophic organisms, as fewer well-annotated reference genomes are available for comparison. Many inaccuracies arise from homology-based annotations, which, while faster than manual curation, can assign functions without biochemical validation. Based on how automated annotation algorithms work, poor annotations can be carried through to new organisms. An additional challenge with annotation is that many metabolic pathways and reactions, particularly in non-model organisms, are still being discovered or refined, which can create gaps in the models that require extensive manual curation or assumptions to fill⁷⁵. Beyond annotation issues, GEMs also face limitations due to their reliance on stoichiometric reactions rather than reaction kinetics. By ignoring reaction kinetics, the entire metabolic network can be modeled; but it comes at a cost because the level of detail is greatly reduced. This modeling approach can be used as a first approximation and more detailed kinetic analysis can be performed on a much smaller and simple network. To enhance their accuracy, GEMs can integrate omics data such as transcriptomics and proteomics. This data

provides crucial insights into cellular states and responses. However, aligning diverse omics datasets with GEMs is another challenge, requiring sophisticated computational techniques. Fortunately, advancements in data integration and computational methods are allowing GEMs to incorporate omics data more effectively and enhance their predictive power⁷⁶. However, even with these advancements in annotation and omics integration, GEMs still face limitations due to key assumptions most notably the reliance on steady state conditions that pose unique challenges in photosynthetic organisms.

Adopting a steady-state assumption poses significant challenges for GEMs in photosynthetic microalgae, where complex diel fluctuations and regulatory mechanisms make strict steady-state models less representative of metabolic dynamics. While this assumption is important mathematically, converting a set of ordinary differential equations to a set of linear equations, it limits the application of GEMs to steady growth conditions. This is particularly pronounced in photosynthetic organisms due to typical growth in diel light conditions which results in substantial fluctuations in metabolism⁷⁷ due to the shift from day to night and vice versa. Photosynthesis also involves numerous regulatory mechanisms, such as photoprotection⁷⁸, photosynthetic quenching⁷⁹, and variations in photon flux⁸⁰, all of which are difficult to represent with static models. To account for these regulatory elements, additional proteomic and transcriptomic data is necessary; however, this data is commonly not available for microalgae. An extension of this problem is the use of a single objective function (most often maximum biomass). While this objective function matches the cellular objective for heterotrophic bacteria quite well⁸¹, this objective is especially problematic in photosynthetic organisms due to the decoupling of carbon and energy inputs and the time-dependent nature of cellular division in diel light. Additionally, the biomass function for algae is more complex and dynamic than those seen in heterotrophic organisms, as many can grow in autotrophic, mixotrophic, and heterotrophic states. With each trophic state requiring different biomass functions⁸² along with different conditions cells have to optimize for dependent on conditions. These can vary such as minimizing energy usage when light is not present or the formation of storage products in preparation for environmental changes. Additionally autotrophic and mixotrophic growth results in biomass composition being more dependent on the environment, changing with light intensity throughout the day under diel conditions⁸³. These diverse challenges demand more advanced, integrated GEM approaches for photosynthetic microalgae (see Figure 2.2), which the following sections will explore.

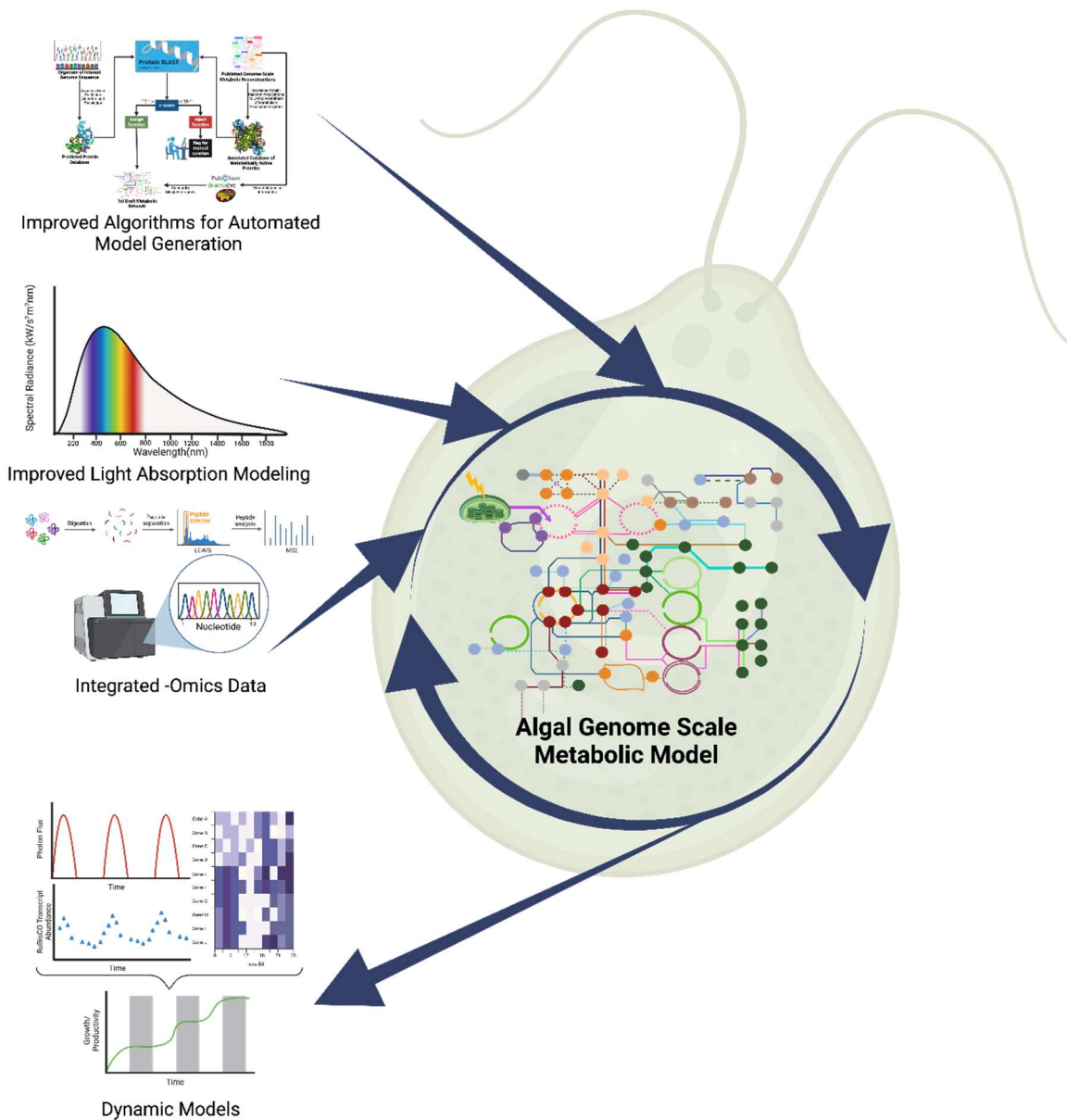


Figure 2.2 Challenges for algae genome scale model construction include the need for automated algorithms for model construction specifically for algae, improved light modeling and integrating -omics data to improve predictability. Addressing these will enable the design of dynamic models that can better predict growth in dynamic conditions, such as day/night cycles.

2.5 Automation of Model Reconstruction

The genome-scale metabolic model (GEM) *iChr1915*⁸⁴ for *C. zofigiensis* represent significant advancement in the automatic curation of photosynthetic metabolic networks.

*iChr1915*⁸⁴ utilized an algorithm called Rapid Annotation of Photosynthetic Systems (RAPS)⁸⁵ to automate much of the process. Other GEM automation tools exist such as model SEED⁸⁶ and CarveMe⁸⁷, however these automation tools are not tailored for use on algae. The model SEED⁸⁶ framework plantSEED⁸⁸ is, as its name would suggest, better suited for the reconstruction of plant GEMs as it carries over many highly conserved reactions in plants to avoid issues with gap filling. Including these conserved reactions in algal GEM reconstructions doesn't properly represent the diversity of microalgal metabolism⁸⁹ and variation from plant metabolism⁹⁰. CarveMe⁸⁷ additionally is primarily for the reconstruction of prokaryotes and bacterial communities with reactions pulled from the BiGG database⁹¹ excluding reactions unique to eukaryotic organisms. The use of RAPS⁸⁵ enabled the development of a high quality first draft network in only 20 minutes; the resulting model only required minimal manual curation. RAPS⁸⁵ facilitates the automated curation of GEMs for photosynthetic algae by leveraging manual curation efforts already invested in published models and using these to generate new models. This method addresses a key challenge in GEM development: the time-consuming nature of manual curation and annotation gaps. By using RAPS⁸⁵, researchers can streamline the initial stages of model development, allowing them to focus on gap-filling and other manual curation efforts that will lead to a high-quality network. This hybrid approach reduces the time-intensive nature of fully manual curation by automating the initial draft creation and filling metabolic gaps while still incorporating the precision of expert intervention where needed.

2.6 Modeling Light Harvesting

The first GEM for microalgae to account for different wavelengths of photons in its metabolic network was *iRC1080*⁶⁹, a model for *C. reinhardtii* allowing for variations in light conditions to influence the model. *iRC1080*⁶⁹ achieved this by defining spectral ranges associated with all the photon-utilizing reactions in the network connecting and allowing for 11 distinct light sources such as solar light as well as halogen and LED lights to be modeled. The metabolic network was also verified with over 90% of transcripts predicted by *iRC1080*⁶⁹ being found in experimental transcriptomic data. Additionally, *iRC1080*⁶⁹ accurately predicted photosynthetic efficiency to be 2%, matching experimental results. The coupling of light wavelengths with reactions marked a substantial improvement on previous models and allows for the optimization of light sources as well as elucidating the phenotypic results of varying light conditions. Similarly, the *C. variabilis* model *iAJ526*⁹² accounts for varying light conditions by simulating the effects of twelve different

light sources on growth rate and uptake rates. These light sources were like those modelled in iRC1080⁶⁹ representing light sources that have been utilized in algal growth, but had a greater focus on modeling different combinations of LED light and didn't include sunlight. Three of these light conditions were experimentally validated, confirming predictions made by the model that white light would provide the best growth followed by red/blue light then red light. iAJ526⁹² predicts higher growth rates than those observed experimentally under all light conditions with the authors attributing the differences to issues with the model's lack of growth kinetics and photoinhibition. These models advance GEM reconstruction in algae and other photosynthetic organisms by offering a more robust representation of the effects light intensity and composition have on metabolism.

Innovations in GEMs for microalgae have also addressed other limitations traditionally seen in GEMs, particularly those affecting photosynthetic organisms. For instance, the *T. pseudonana* model iTps1432⁶² incorporates the application of photon loss reactions to simulate photosynthetic quenching. By including these reactions, iTps1432⁶² offers valuable insight into photon loss reactions with particular interest coming from predictions around cyclic electron flow at low light intensities. At these lower light intensities, the model predicts that a significant portion of total electron flow is made up of cyclic electron flow supporting other findings highlighted in the paper that cyclic electron flow is important for ATP generation at low light⁹³. Cyclic electron flow is not only important for ATP generation and modeling light dynamics but has also been demonstrated to be important in lipid biosynthesis pathways in algae⁹⁴. This highlights the potential improvements adding light dynamics reaction within GEMs can provide. With this added insight these models can be better applied to determine targets for improving metabolic engineering outcomes under autotrophic conditions.

2.7 Models With a Focus on Reactions Outside of Carbon Metabolism

Models such as iMM627⁹⁵ for *A. protothecoides* and iRJI321⁹⁶ for *N. gaditana* incorporate predictions of hydrogen production, demonstrating how GEMs can be utilized for potential products beyond biomass and hydrocarbons. The iMM627⁹⁵ model integrates two objective functions maximizing both biomass and hydrogen production. This allows for the exploration of algae to generate green hydrogen which has garnered increased interest in recent years⁹⁷. By incorporating multiple objectives, the model can more completely utilize the GEMs metabolic network and better represent reactions outside of central carbon metabolism. The *N. salina* model

*iNS934*⁹⁸ provides a more detailed representation of nitrogen metabolism, capturing the intricate balance between carbon fixation and nitrogen assimilation, while also incorporating a variety of nitrogen sources. This allows *iNS934*⁹⁸ to integrate essential reactions not directly tied to carbon metabolic pathways, addressing gaps present in earlier models and offering more flexibility when optimizing media recipes.

2.8 Model Robustness

The Lavoie et al. *F. cylindrus* model⁹⁹ focuses on reaction robustness to help analyze how metabolic networks maintain stability under stress or environmental shifts. This robustness analysis, combines flux balance analysis (FBA) with minimization of metabolic adjustment (MOMA)¹⁰⁰ allows for better prediction of how networks respond to perturbations made by knock outs. This improves the GEMs' ability to simulate stress responses, addressing a significant aspect of how *F. cylindrus* survives well in its very dynamic environment¹⁰¹. The Recht et al. model¹⁰² for *H. pluvialis* further incorporates variability flux sampling (VFS), an additional step on the commonly used flux variability analysis (FVA)¹⁰³. VFS enables more accurate flux predictions and a deeper analysis of metabolic pathways as it not only predicts the range of possible fluxes, as is done in FVA, but also includes determinations about the probabilities of various fluxes. Incorporating VFS allows for better understanding of pathways that are activated as a stress response as demonstrated in the models focus on exploring the shift towards fatty acid synthesis under nitrogen starvation. By allows the model to incorporate a probabilistic approach into the stoichiometric model this aids in representing the dynamics reactions present under nitrogen stress conditions¹⁰⁴.

2.9 Integration of Additional Omics Data into Models

Models have additionally taken the novel approach of incorporating omics data to address limitations created by representing metabolism purely stoichiometrically. The Yao et al. model⁷³ for *C. reinhardtii* does this by incorporating RNA sequencing data to assume the proteome of the organism as well as enzyme data to create a protein-constrained metabolic model (PC-model). This allows for the model to better represent the dynamics that are lost in the conventional approach of representing metabolism only stoichiometrically. Another model, *iEH410*¹⁰⁵ for *Emiliana huxleyi*, introduces dynamic FBA, significantly improving the simulation of internal regulation of metabolic reactions by moving beyond static flux distributions and better reflecting

real-time cellular responses. Models that incorporate omics data have great potential in better representing the complex regulatory mechanisms present around metabolism¹⁰⁶ as well as applications under varying growth conditions¹⁰⁷.

2.10 Dynamic Metabolic Modeling

Transient Metabolic Models (TMMs) offers a promising avenue for future research utilizing the value of GEMs while offering a dynamic model. The first TMM for microalgae was developed by Metcalf and Boyle⁷² in *C. reinhardtii* to model growth in diel light. The model was based on experimental transcriptomics data based on growth in 12:12 hour day:night cycles as this data was used to constrain the availability of the associated enzymatic reactions based on gene expression data. Additionally, the TMM also decoupled the biomass objective functions from the standard static biomass equation allowing it to better simulate the cells adapting to the changing environmental conditions over a day. This is a substantial improvement on GEMs, addressing one of their key challenges: that they are generally static stoichiometric representations of metabolism. The dynamics of the TMM also allow for better targeting for metabolic engineering as these models better represents the fluctuations in metabolism over the course of a day rather than at a single point in the day. Specifically for microalgae as photosynthetic organisms, the ability of the Metcalf and Boyle model to simulate changes in metabolism over the course of a day allows for improvements in understanding metabolic dynamics under diel conditions.

Table 2.2 The table displays the genome-scale metabolic models (GEMs) currently published, organized by species and year of publication from oldest to newest.

Species	Model	Year	Reactions	Metabolites
<i>Auxenochlorella protothecoides</i>	<i>iMM627</i> ⁹⁵	2019	1,963	2,115
<i>Chlamydomonas reinhardtii</i>	Boyle and Morgan ⁶⁷	2009	484	458
	<i>iAM303</i> ¹⁰⁸	2009	259	267
	<i>iRC1080</i> ⁶⁹	2011	2,190	1,706
	<i>AlgaGEM</i> ¹⁰⁹	2011	1,725	1,862
	<i>iBD1106</i> ¹¹⁰	2014	2,445	1,959
	<i>iCre1355</i> ⁶⁸	2015	2,394	1,845
	Winck et al. ¹¹¹	2016	3,554	2,342
	Salguero et al. ¹¹²	2018	3,726	2,436
	Yao et al. ⁷³	2023	2,641	2,240
<i>Chlorella variabilis</i>	<i>iAJ526</i> ⁹²	2016	1,455	1,236
<i>Chlorella vulgaris</i>	<i>iCZ843</i> ⁶³	2016	2,294	1,770
	<i>iCZ946</i> ¹¹³	2018	2,294	1,770
<i>Chromochloris zofingiensis</i>	<i>iChr1915</i> ⁸⁴	2024	3,413	2,652
<i>Dunaliella salina</i>	<i>iEC1693</i> ¹¹⁴	2024	4,614	3,732
<i>Emiliana huxleyi</i>	<i>iEH410</i> ¹⁰⁵	2015	410	363
<i>Fragilariopsis cylindrus</i>	Lavoie et al. ⁹⁹	2020	2,144	1,707
<i>Haematococcus pluvialis</i>	Recht et al. ¹⁰²	2014	2,622	1,975
<i>Isochrysis sp.</i>	<i>iIsochr964</i> ¹¹⁵	2023	4,315	1,879
<i>Nannochloropsis gaditana</i>	<i>iRJ1321</i> ⁹⁶	2017	1,918	1,862
<i>Nannochloropsis salina</i>	<i>iNS934</i> ⁹⁸	2017	2,345	1,985
<i>Phaeodactylum tricornerutum</i>	<i>iLB1027_lipid</i> ⁶⁰	2016	4,456	2,172
	<i>iLB1025</i> ⁶⁰	2016	2,156	1,704
<i>Scenedesmus obliquus</i>	<i>iAR632</i> ¹¹⁶	2023	1,476	1,549
<i>Schizochytrium limacinum</i>	<i>iCY1170_DHA</i> ¹¹⁷	2015	1769	1659
<i>Thalassiosira pseudonana</i>	<i>iThaps987</i> ¹¹⁸	2020	2,477	2,456
	<i>iTps1432</i> ⁶²	2021	6,073	2,789

2.11 Prospects for Future Microalgal Genome-scale Metabolic Models

Future advances in GEM formulation will enable more sophisticated models that will be better suited to predicting the dynamic and complex metabolism of microalgae. While GEMs have traditionally relied on steady-state assumptions using FBA, incorporating regulatory constraints has successfully been demonstrated in the Yao et al. model and *iEH410*. While both these GEMs incorporated transcriptomics data, there are further advancements that can be made to the reconstruction of future GEMs incorporating multi-omics data. Tools such as GECKO 2.0¹¹⁹ allow for pipelines for the implementation of enzyme kinetic parameters and proteomic

data into GEMs which has already been utilized in multiple species of yeast, *E. coli* and *Homo sapiens*. By adding additional layers of omics data GEMs can address limitations that are presented in many of the currently available static stoichiometric models.

While data availability remains a constraint for many microalgal species, machine learning offers exciting opportunities. In particular, deep learning, which uses neural networks to perform multi-level predictions¹²⁰, has already improved genome annotations in bacterial metagenomes¹²¹. Applying similar approaches to microalgae could enable the reconstruction of GEMs for the vast number of algal species that remain unculturable¹²². This could not only deepen our understanding of these organisms but also help design more effective cultivation strategies.

Another promising avenue is the integration of GEMs with Transient Metabolic Models (TMMs), which simulate metabolic changes over time and under varying environmental conditions. While a TMM has been developed for *Chlamydomonas reinhardtii*⁷², other microalgae including those with GEMs (see Table 2.2) currently lack such dynamic models. Expanding TMMs to include additional species and conditions such as UV radiation, temperature fluctuations, and nutrient availability could dramatically enhance the applicability of GEMs in modeling real-world scenarios¹²³⁻¹²⁵.

Altogether, these innovations multi-omics integration, machine learning, and dynamic modeling represent the future of microalgal GEMs. They offer a more comprehensive understanding of algal metabolism, particularly under diel cycles and photosynthetic fluctuations, moving the field closer to realizing the full potential of microalgae in biotechnology and sustainability applications.

2.12 Conclusion

Microalgae hold immense potential for contributing to a sustainable future through their applications in biofuels, bioremediation, and the production of high-value products. The development of genome-scale metabolic models (GEMs) has emerged as a powerful tool in understanding the complex metabolic networks of these organisms, enabling researchers to optimize their metabolic pathways effectively. However, while GEMs have made significant strides, they are not without limitations. Issues related to incomplete genome annotations, static assumptions, and the integration of multi-omics data continue to pose challenges. To address these limitations and fully harness the capabilities of microalgae, there is a pressing need for the

creation of more GEMs across a diverse array of algal species. Expanding the repertoire of GEMs will enhance our understanding of algal metabolism and facilitate the development of tailored strategies for metabolic engineering. By addressing the existing challenges and improving GEM methodologies, we can pave the way for a more environmentally friendly future, ultimately contributing to a more sustainable and productive bioproduct landscape.

CHAPTER 3
RECONSTRUCTION OF A GENOME-SCALE METABOLIC MODEL FOR
AUXENOCHLORELLA PROTOTHECOIDES

Based on a paper in preparation for publication
Jacob Tamburro¹, Mark Vigliotti², Nanette R. Boyle^{1,2}

3.1 Abstract

Microalgae have recently gained prominence as platforms for biofuel production and the synthesis of high value bioproducts. Among these, *Auxenochlorella protothecoides* stands out for its significant lipid accumulation, rapid growth, and low metabolic engineering requirements, making it a prime candidate for renewable biofuel production. To fully exploit its metabolic potential, genome-scale metabolic models (GEMs) offer a powerful computational framework for predicting metabolic behavior and guiding in silico engineering strategies. The development of a genome-scale metabolic model for *A. protothecoides* is presented, with a focus on autotrophic, mixotrophic, and heterotrophic growth conditions. The model, called *iApro4643*, was reconstructed through a combination of automated methods (RAPS) and manual curation. Flux balance analysis (FBA) simulations were used to generate flux maps, incorporating experimental data to reflect metabolic shifts across different trophic conditions. FBA also identified 65 essential genes for photoautotrophic growth and 63 for mixotrophic growth, with notable differences in essential genes related to photosynthesis. Biomass composition analysis showed significant shifts, with starch accumulation under heterotrophic growth and a protein-dominant profile in autotrophic cultures. Fatty acid profiles also shifted, with autotrophic cultures predominantly producing linolenic acid (18:3) and mixotrophic and heterotrophic cultures favoring oleic acid (18:1), suggesting potential applications in biodiesel production. These findings demonstrate the metabolic flexibility of *A. protothecoides* and present a base for using the GEM to optimize metabolic processes for biotechnological applications.

¹ Colorado School of Mines, Quantitative Bioscience Engineering

² Colorado School of Mines, Chemical and Biological Engineering

3.2 Introduction

As global concerns over climate change intensify, there is increasing demand for sustainable, low-carbon alternatives to fossil fuels. One potential solution to reducing reliance on fossil fuels as well as offering a low-carbon solution are biofuels, sustainable fuels derived from biological material. Biofuels have been utilized as a fuel source since the introduction of the Model T in 1908 with ethanol as one intended fuel sources for the vehicle¹²⁶. Ethanol, derived largely from corn in the United States, while still in use today as a fuel blend has disadvantages as a future biofuel source such as requiring arable land. Microalgae provide a promising candidate for a future source for biofuels due to having higher growth rates¹⁴ and higher photosynthetic efficiency than terrestrial plants¹²⁷ as well as not requiring arable land. There is also interest in producing value added products from microalgae with particular interest in the production of carotenoids¹²⁸. Carotenoids are tetraterpenoid pigments produced in plants, algae and photosynthetic bacteria¹²⁹ which have commercial potential due to their nutritional and supplemental value¹³⁰. The ability to produce carotenoids adds value to photosynthetic organisms as a feedstock in the production of biofuels.

The microalgae *Auxenochlorella protothecoides* shows significant promise as source of sustainable biofuels due to its capacity for high lipid content, high growth rate and low metabolic engineering costs. *A. protothecoides* can accumulate up to 60% lipid content when subjected to nitrogen deprivation¹⁶. *A. protothecoides* also produces carotenoids such as lutein and zeaxanthin, demonstrating the potential for *A. protothecoides* outside of biofuel production¹³¹. Furthermore *A. protothecoides* has a doubling time of 11 hours when cultivated with an organic carbon source. *A. protothecoides* is additionally an attractive candidate for metabolic engineering because it can perform homologous recombination¹³². Homologous recombination is a genetic process that is commonly found in yeast that allows for precise, low-cost and rapid genetic engineering¹³³. The strain of *A. protothecoides* (UTEX250) used in this paper is a hybrid that is vegetative diploid¹³⁴ which has the benefit of preventing the loss of modifications due to recombination during meiosis and sexual reproduction. The combination of high lipid accumulation, high growth rate, and capacity for inexpensive genetic engineering enhances *A. protothecoides*' potential as a scalable commercially viable cell factory.

To optimize cell factory productivity, researchers have leveraged genome-scale metabolic models (GEMs)^{65,135}, computational stoichiometric representations of the metabolic reactions present in an organism. GEMs represent the metabolic networks of an organism by gene-protein

reactions based on the annotated genome of the organism to create constraint-based, stoichiometric models. The network is then solved by flux balance analysis (FBA) based on experimental constraints as well as an objective function. FBA is a computational method used to predict the flow of metabolites through a metabolic network under steady-state conditions, allowing for the estimation of growth rates and the identification of active metabolic pathways¹³⁶. A previously developed GEM has already been utilized to increase lutein production in another strain of *A. protothecoides*¹³⁷(sp. 0710). GEMs have additionally been used to optimize media composition resulting in a 61% increased yield of fatty acids in related algae *Chlorella vulgaris*¹³⁸. Unlike traditional *in vitro* methods, which can be time-intensive and costly, GEMs allow researchers to conduct rapid *in silico* experiments, assessing the potential impact of genetic or metabolic modifications limiting the need for physical testing.

Our GEM *iApro4643*, the most comprehensive model of *A. protothecoides* to date, consisting of 2820 metabolites and 3357 reactions *iApro4643* provides a powerful computational tool for the efficient exploration of strategies to enhance the productivity of *A. protothecoides*, aiding in the development of *A. protothecoides* as an industrially relevant cell factory. The method for automated reconstruction as well as experimentally determined constraints are outlined below.

3.3 Materials and Methods

3.3.1 Strains and growth conditions

Auxenochlorella protothecoides strain UTEX 250 was generously provided by the Merchant lab at UC Berkeley. Heterotrophic, mixotrophic and autotrophic cultures were grown at 25°C at 120 rpm in baffled 250mL Erlenmeyer flasks with the mixotrophic and autotrophic cultures being grown with 120 $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ and atmospheric CO₂. The cultures were grown in APM1 media as described by Duenas et al.¹³⁹ supplemented with 20 g/L glucose for both heterotrophic and mixotrophic growth conditions.

3.3.2 Biomass Composition

Biomass composition data for mixotrophic and heterotrophic cultures were determined using cell pellets from 3 biological replicate flasks. Cell pellets were obtained by centrifuging 200 μL , 1mL or 5mLs of culture in 2mL tubes or 15mL falcon tubes at an optical density (OD) of 7.0 during mid-exponential growth. The 1mL cell pellets were centrifuged at $10,000 \times g$ for 5 minutes while the 5mL cell pellets were centrifuged at $1,800 \times g$ for 10 minutes the 5mL cell

pellets were then transferred to 2mL tubes. For autotrophic cultures cell pellets were obtained from 3 biological replicate flasks by centrifuging 2mL or 20mLs of culture in 2mL tubes or 50mL falcon tubes at an optical density of 0.5. The 2mL cell pellets were centrifuged at $10,000 \times g$ for 5 minutes while the 20mL cell pellets were centrifuged at $1,800 \times g$ for 10 minutes the 20mL cell pellets were then transferred to 2mL tubes. The collected cell pellets were then stored in a freezer at -20°C . For each biomass method used for biomass composition 3 biological replicates were used with 3 technical replicates making a total of 9 samples used for each.

Dry weight was used to determine the total biomass for the biomass function with 20mLs of culture used for autotrophic cultures and 5mL used for mixotrophic and heterotrophic cultures used. Dry weight was determined by vacuum filtration of the representative amount of culture through pre-dried and pre-weighed $1.6 \mu\text{m}$ pore size glass microfiber filters. After filtering the filters with the cells on them were rinsed with APM1 media and dried overnight at 80°C before being weighed again the next day⁸⁴. Glucose uptake was determined on 1mL supernatant samples taken in 1.5mL tubes throughout mid-exponential growth leading up to harvesting the cell pellets. The spent media samples were run on a Roche Cedex bioanalyzer to determine glucose content in the supernatant. Ethanol content was also determined on the Roche Cedex bioanalyzer on the same supernatant samples. Total lipids were determined by the Folch method on 20mL (autotrophic) and 5mL (mixotrophic and heterotrophic) cell pellets. Cell pellets were first lysed by three rounds of bead beating with glass beads in the 2mL tubes for one minute using a Mini-Beadbeater-16. The lysed cell contents were then transferred to glass 20mL sample vials using four rounds of 1mL methanol being added to the 2mL tube then transferred with the lysed cell contents to the 20mL sample vial. 8mL of chloroform was then added to the 4mL of methanol and vortexed after which 3mL of water was added to achieve a 8:4:3 ratio of chloroform:methanol:water¹⁴⁰. The mixture was then vortexed again followed by centrifuging the vial at $300 \times g$ for 5 minutes to decrease the amount of time for the phases to separate. After the phases separate the bottom layer containing the chloroform and lipids was transferred to a pre-weighed open top vial and another 2mL of chloroform was added to the vial still containing the methanol and water. The mixture was again vortexed and centrifuged as described above with the bottom layer being transferred again after which the open top vial was allowed to evaporate. Once all the chloroform evaporated the vial containing the lipids was weighed again to determine the total lipid content.

Total carbohydrates were determined through an anthrone assay on 2mL (autotrophic) and 200µL (mixotrophic and heterotrophic) cell pellets. To conduct the anthrone assay both cell pellets were resuspended in 2mL of milli-Q water and 100µL of mixture was then transferred to a 1.5mL tube with 900µL 75% H₂SO₄ containing 2mg/mL anthrone on ice. A set of glucose standards were also prepared and added in parallel to the samples. The solution and standards were then vortexed and heated at 100°C for 15 minutes. After which the tubes were put onto ice and incubated until they fully cooled down. The samples and standards were then transferred to 10mm light path semi-micro cuvettes and measured at 578nm¹⁸. Total protein was determined by a Pierce BCA assay on 2mL (autotrophic) and 200µL (mixotrophic and heterotrophic) cell pellets. The cells were first lysed by three rounds of bead beating with glass beads in the 2mL tubes then resuspended in 1mL of 0.2M NaOH. The resuspended cell contents were then centrifuged at 10,000×g for 5 minutes and the supernatant was run according the Thermo Scientific Pierce BCA assay procedure¹⁴¹.

Chlorophyll content was determined on 2mL (autotrophic) and 200µL (mixotrophic and heterotrophic) cell pellets by acetone extraction. Cell pellets were lysed by bead beating using a Mini-Beadbeater-16 after being lysed the cells were put on ice with 1.5mL of acetone then being added to the lysed pellet. The acetone and cell contents were then centrifuged at 10,000×g for 5 minutes at 4°C the supernatant was then removed and put in quartz cuvettes with a light path length of 10mm. The spectrophotometer was then blanked with pure acetone at 645, 663 and 750nm then the samples were then measured at 645, 663 and 750nm¹⁴². The equation (Equation 1) was used to determine the chlorophyll content with 750nm absorbance included to determine if cell debris was present in the sample.

$$Total\ chlorophyll\ \left(\frac{\mu g}{mL}\right) = (20.2 \times A_{645}) + (8.02 \times A_{663}) \quad (3.1)$$

Amino acid composition was determined through first hydrolyzing 2mL (autotrophic) and 1mL (mixotrophic and heterotrophic) cell pellets in 700µL 6M HCl in a vacuum hydrolysis tube for 24 hours at 110°C. The hydrolyzed cell contents were then transferred to a 1.5mL tube and dried in a SpeedVac after which the dried samples were resuspended in milli-Q water and moved to .22µm spin-X centrifuge filter. The filters were centrifuged at 10,000×g for 1 minute removing debris from the samples then once again dried in a SpeedVac. The dried samples were then resuspended in 50µLs of pyridine and incubated at 40°C for 90 minutes while occasionally vortexing the sample. 70µLs of MTBSTFA (N-(tertbutyldimethylsilyl)-N-methyl-

trifluoroacetamide) was added to the samples to derivatize the amino acids to their tert-butyltrimethylsilyl derivatives. The samples were then incubated at 60°C for 30 minutes after which they incubated at room temperature for 90 minutes. The derivatized samples were then transferred to 2mL autosampler vial with 150µL insert and run on an Agilent 6890 GC with a 5973 mass spectrometer with a DB-1701 column^{84,143}.

Fatty acids were determined on 2mL (autotrophic) and 1mL (mixotrophic and heterotrophic) cell pellets through converting the fatty acids into their FAMES (fatty acid methyl esters). This was done by first resuspending cell pellets with 500µL of water and transferring the mixture to a 4mL glass vial. Next 1mL of saponification reagent (40g/L potassium hydroxide dissolved in methanol) was added to the vial then vortexed and incubated at 100°C for 90 minutes. The fatty acid esters were then methylated by adding 1.5mL of methylation reagent (0.7M hydrochloric acid in methanol) vortexing the mixture and incubating at 60°C overnight. The FAMES were then extracted by adding 1mL of hexane followed by vortexing the sample. 150µL of the sample was then transferred to a 2mL autosampler vial with 150µL insert and run on an Agilent 6890 GC with a 5973 mass spectrometer with a DB-WAX column¹⁴⁴.

Starch content was determined on 2mL (autotrophic) and 1mL (mixotrophic and heterotrophic) cell pellets. The first step was bead beating the pellets using glass beads in a Mini-Beadbeater-16 after being lysed the pigments were removed by adding 1mL of 95% ethanol followed by centrifuging at 10,000×g for 5 minutes. The supernatant was then removed and 200µL of 100mM sodium acetate (pH 4.5) was added, the sample was then moved to a 4mL glass culture vial. The sample was then autoclaved at 121°C for 30 minutes to convert the particulate starches to colloidal starches. After autoclaving the sample was transferred to a new 1.5mL tube then 400µL of 80 unit/mL solution of amyloglucosidase (Sigma Aldrich, S9144) was added and the sample was then vortexed and incubated at 55°C for 1 hour¹⁸. Once the incubation was complete 400µL of pure ethanol was added followed by another round of vortexing. The sample was then incubated at 15°C for 15 minutes and centrifuged at 10,000×g for 5 minutes. The supernatant was then transferred to another tube with was analyzed on the Roche Cedex bioanalyzer to determine glucose content.

DNA content was determined using a Qubit 3.0 fluorometer, determined on 2mL (autotrophic) and 200µL (mixotrophic and heterotrophic) cell pellets. The cell pellet was lysed using glass beads in a Mini-Beadbeater-16, the lysed cell pellet was then resuspended in 1mL of milli-Q water. The sample was then centrifuged at 10,000×g for 5 minutes after which the sample

was run using a Qubit 3.0 fluorometer following the procedure outlined. The measured DNA content was then used to determine the RNA content assuming that there was a 28-fold higher RNA content than DNA content¹⁴⁵. The ratio was determined in *C. reinhardtii* however being as both organisms are green microalgae the assumption is believed to be applicable.

3.3.3 Model Reconstruction

To reconstruct the genome-scale metabolic network for *A. protothecoides*, we utilized the RAPS (Rapid Annotated of Photosynthetic Systems) algorithm⁸⁵. The RAPS algorithm, implemented in Python, integrates several computational tools and modules, including COBRAPy¹⁴⁶, Biopython¹⁴⁷, LibSBML¹⁴⁸, Matplotlib¹⁴⁹, and Open Babel¹⁵⁰. The reconstruction process began with a draft network, which was generated by comparing the predicted protein sequences of *A. protothecoides* to those in existing donor models using BLAST. Proteins with a significant match (E-value < 10^{-20}) were automatically incorporated into the draft network, while sequences without a high-confidence match were flagged for manual curation.

After the initial draft model was automatically reconstructed, it underwent a brief manual curation. Reactions lacking genetic support, such as transport and maintenance reactions, were incorporated directly from existing databases to ensure modelling completeness. Then, a biomass reaction specific to *A. protothecoides* for each trophic growth condition was incorporated. To identify and address erroneous energy-generating cycles (EGCs) commonly found in automatically reconstructed metabolic networks, we followed the procedure outlined by Frtizemeier, et al.¹⁵¹

3.3.4 Flux Balance Analysis (FBA)

To perform flux balance analysis (FBA) on the reconstructed metabolic model of *A. protothecoides*, COBRAPy¹⁴⁶, a Python package for constraint-based modeling, was used. Here, parsimonious flux balance analysis (pFBA) was used for each simulation. To simulate different growth conditions, we defined specific environmental constraints by adjusting the uptake rates of key metabolites. Any metabolite present in the media was left unconstrained. Autotrophic growth in the light was modeled by permitting carbon dioxide uptake as the only carbon source. Light uptake reactions were incorporated following the light-modeling and photon flux approach developed by Chang, et al.⁶⁹ CO₂ uptake was minimized to determine the minimal amount of CO₂ required to achieve the measured growth rate. Mixotrophic growth in the light was simulated by allowing the uptake of glucose, oxygen, and carbon dioxide, reflecting the

simultaneous operation of autotrophic and heterotrophic metabolisms. CO₂ uptake was constrained to the same value as that in the autotrophic case. Glucose uptake was set to the experimentally measured value in both the mixotrophic and heterotrophic cases, and biomass production was maximized in both cases as well. A summary of constraints can be found in Table 3.2.

3.4 Results

3.4.1 Network Reconstruction and Curation

After brief manual curation of the automatically generated first draft model using the Rapid Annotated of Photosynthetic Systems (RAPS) algorithm, the final model consists of 2820 metabolites and 3357 reactions across 7 compartments. Of those 3357 reactions, 587 of them are transport reactions. Light reactions were incorporated to simulate photoautotrophic and mixotrophic growth. Biomass reactions corresponding to experimentally determined biomass compositions in each trophic growth condition (autotrophic, mixotrophic, and heterotrophic) were also introduced. Eyespot and flagella reactions that had acceptable E-scores were removed from the model since those organelles are not present in *Auxenochlorella*. After manual curation, the published model achieved a MEMOTE score¹⁵² of 85% with a near 100% score in stoichiometric, mass balance, charge balance, and unbounded flux consistency.

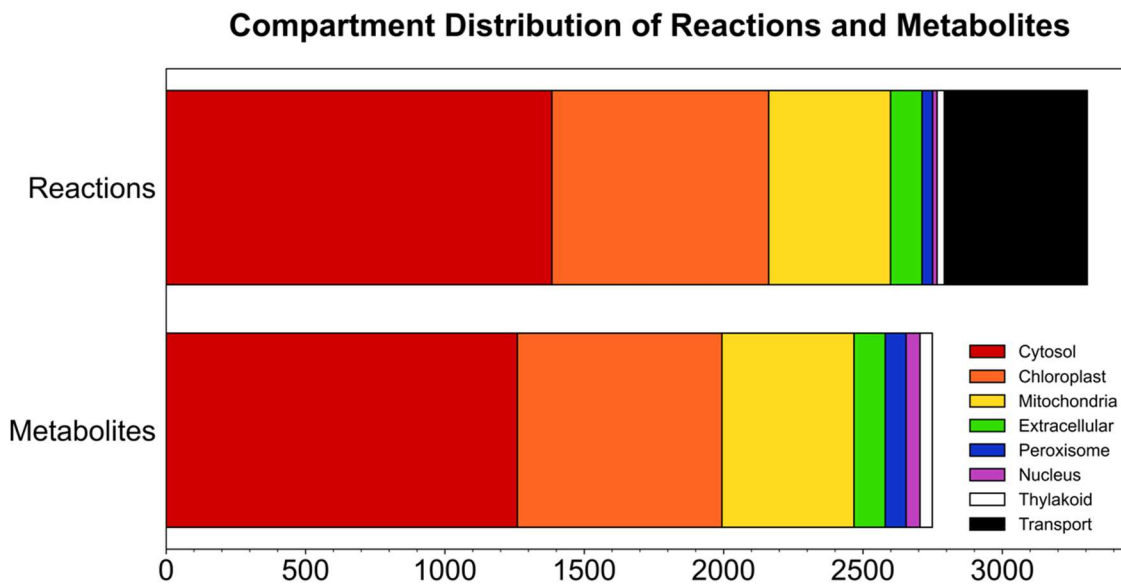


Figure 3.1 Compartmental data for reactions and metabolites in iApro4643 as well as transport reactions.

3.4.2 Biomass Composition and biomass function

The biomass composition of cells was determined at mid-exponential growth under autotrophic, mixotrophic and heterotrophic growth conditions. The experimentally derived biomass composition data along with assumed RNA content based on measured DNA was used to construct a comprehensive biomass formation equation for each growth condition. The results of the biomass equations are presented in Figure 3.2.

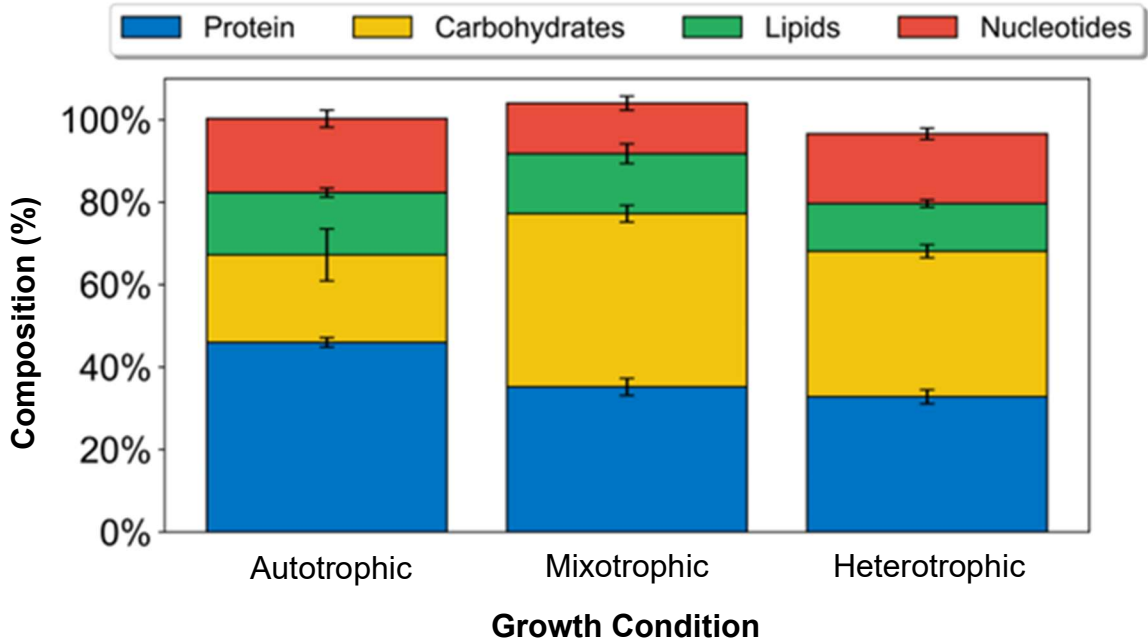


Figure 3.2 Macromolecule biomass composition of *A. protothecoides* under autotrophic (APM1-Glucose media), mixotrophic and heterotrophic (APM1 media) growth conditions

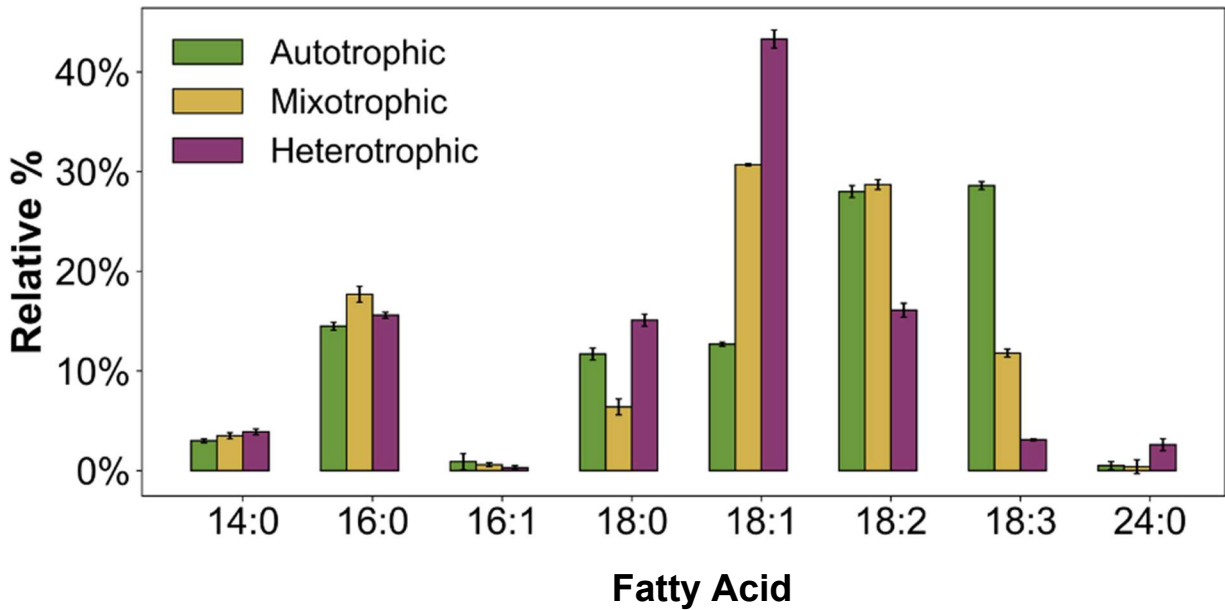


Figure 3.3 Total measured fatty acid composition of *A. protothecoides* as determined by FAMES measured on GC-MS under autotrophic, mixotrophic and heterotrophic conditions.

3.4.3 Experimentally measured model constraints

To have the pFBA simulations best represent in vivo growth conditions glucose uptake rates for mixotrophic and heterotrophic cultures were used to constrain the carbon uptake of the simulation. For autotrophic growth the simulation was constrained by the light availability that was measured experimentally. For each trophic growth condition experimentally determined growth rates were also included to constrain the simulations. The constraints used for the pFBA simulations are available in Table 3.2. The growth rates for mixotrophic and heterotrophic growth were unsurprisingly considerably higher than that seen with autotrophic growth utilizing atmospheric CO₂. Additionally, the growth rate and glucose uptake rate seen in heterotrophic growth were higher than those seen in mixotrophic growth demonstrated a degree of glucose uptake inhibition due to light.

Table 3.1 Experimentally measured mid-exponential growth rate (gDW·h), Glucose uptake (mmol/gDW·h) rate and ethanol excretion rates (mmol/gDW·h) for each trophic condition. Glucose uptake measurements were not taken on autotrophic cultures as there was no glucose in the media. Autotrophic and mixotrophic cultures were assessed for Ethanol content however the values were below the limit of detection on the Cedex bioanalyzer.

Model constraint	Autotrophic	Mixotrophic	Heterotrophic
Growth rate (gDW·h)	0.00225	0.035	0.062
Glucose uptake rate (mmol/gDW·h)	n/a	0.94	2.66
Ethanol excretion rate (mmol/gDW·h)	n/a	n/a	.0235

Table 3.2 The table shows the exact lower and upper bound constraints that were put on the metabolic models when running the pFBA.

Simulation	Growth Rate (gDW·h)	Glucose Exchange (mmol/gDW·h)	CO ₂ Exchange (mmol/gDW·h)	O ₂ Exchange (mmol/gDW·h)	Photon Exchange (mmol/gDW·h)
Minimize CO (Autotrophic)	0.002248, 0.002248	0, 0	-100, 0	0, 100	-100, 0
Maximize Biomass (Mixotrophic)	0, 100	-0.94, 0	-0.1049, 100	-100, 100	-3.85, 0
Maximize Biomass (Heterotrophic)	0, 100	-2.66, 0	0, 100	-100, 0	0, 0

3.4.4 Flux distribution

Simulations in our reconstructed model *ipro4643* were done under autotrophic, mixotrophic and heterotrophic conditions using pFBA. For autotrophic growth the pFBA's objective function was to maximize CO₂ usage while at a set growth rate and can be seen in Table 3.2. For mixotrophic and heterotrophic conditions the pFBA was run using the objective to maximize biomass productivity and is not therefore constrained by growth rate but instead uptake rate. The predicted fluxes are substantially different between the different growth conditions. In the autotrophic condition (Figure 3.4) most of the flux is as expected being committed to regenerating substrate for photosynthesis. There is also flux towards biomass synthesis that occurs in the cytosol. The simulation for the mixotrophic growth condition (Figure 3.5) does not resemble the autotrophic growth condition and has far more in common with the heterotrophic growth condition. In the mixotrophic condition (Figure 3.5) most of the predicted flux passes through glycolysis and eventually into the TCA cycle in the mitochondrial while utilizing the glyoxylate shunt. There is also considerable flux to pentose phosphate pathway for gluconeogenesis with essentially no uptake CO₂ for photosynthesis with the model also predicting starch synthesis. The heterotrophic flux predictions (Figure 3.6) focuses heavily on glycolysis in the cytosol and the TCA cycle in the mitochondria with a more limited focus on the generation biomass. The model also predicts that under the heterotrophic conditions that *A. protothecoides* will undergo gluconeogenesis through the formation of glyceraldehyde-3-phosphate (G3P) from pyruvate.

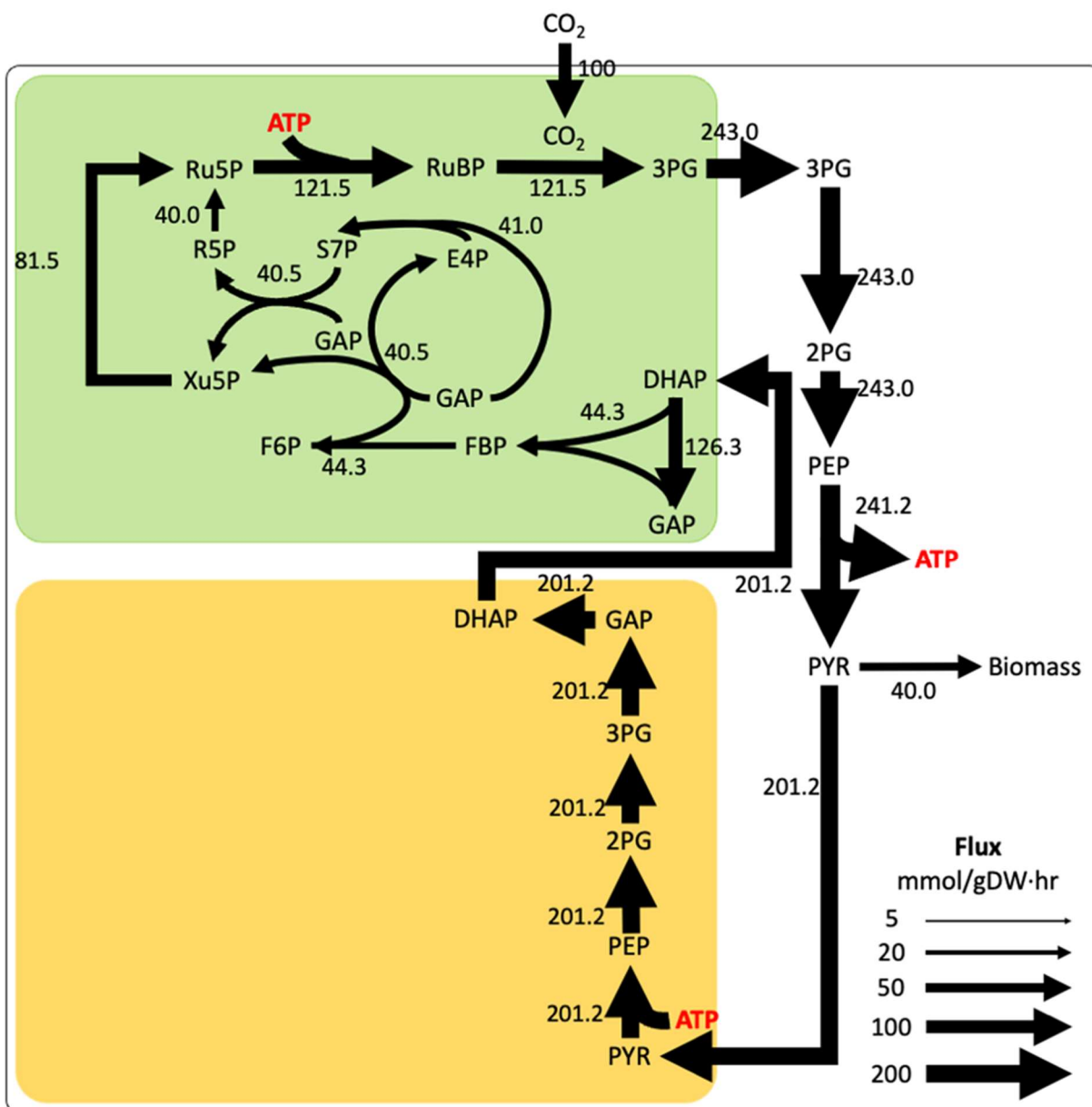


Figure 3.4 In silico simulated autotrophic growth of *A. protothecoides* grown on atmospheric CO₂ (APM1-glucose media) using experimental data to constrain the model. The predicted fluxes are normalized to a CO₂ uptake of 100mmol/g·DW·hr (100mmol total carbon) for ease of comparison with the other trophic condition results.

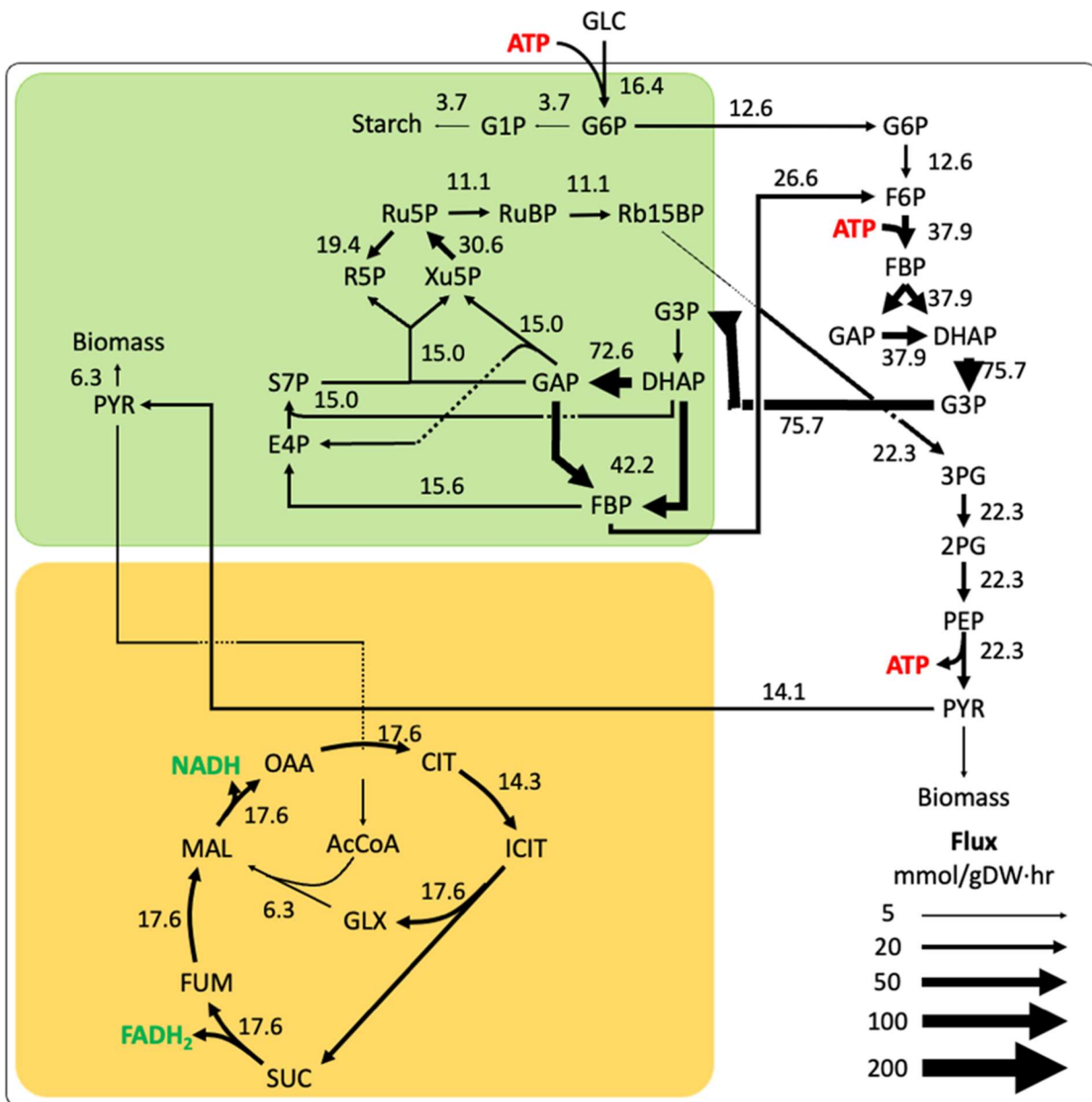


Figure 3.5 In silico simulated mixotrophic growth of *A. protothecoides* grown on glucose (APM1 media) using experimental data to constrain the model. The predicted fluxes are normalized to a glucose uptake of 16.6mmol/g·DW·hr (100mmol total carbon) for ease of comparison with the other trophic condition results.

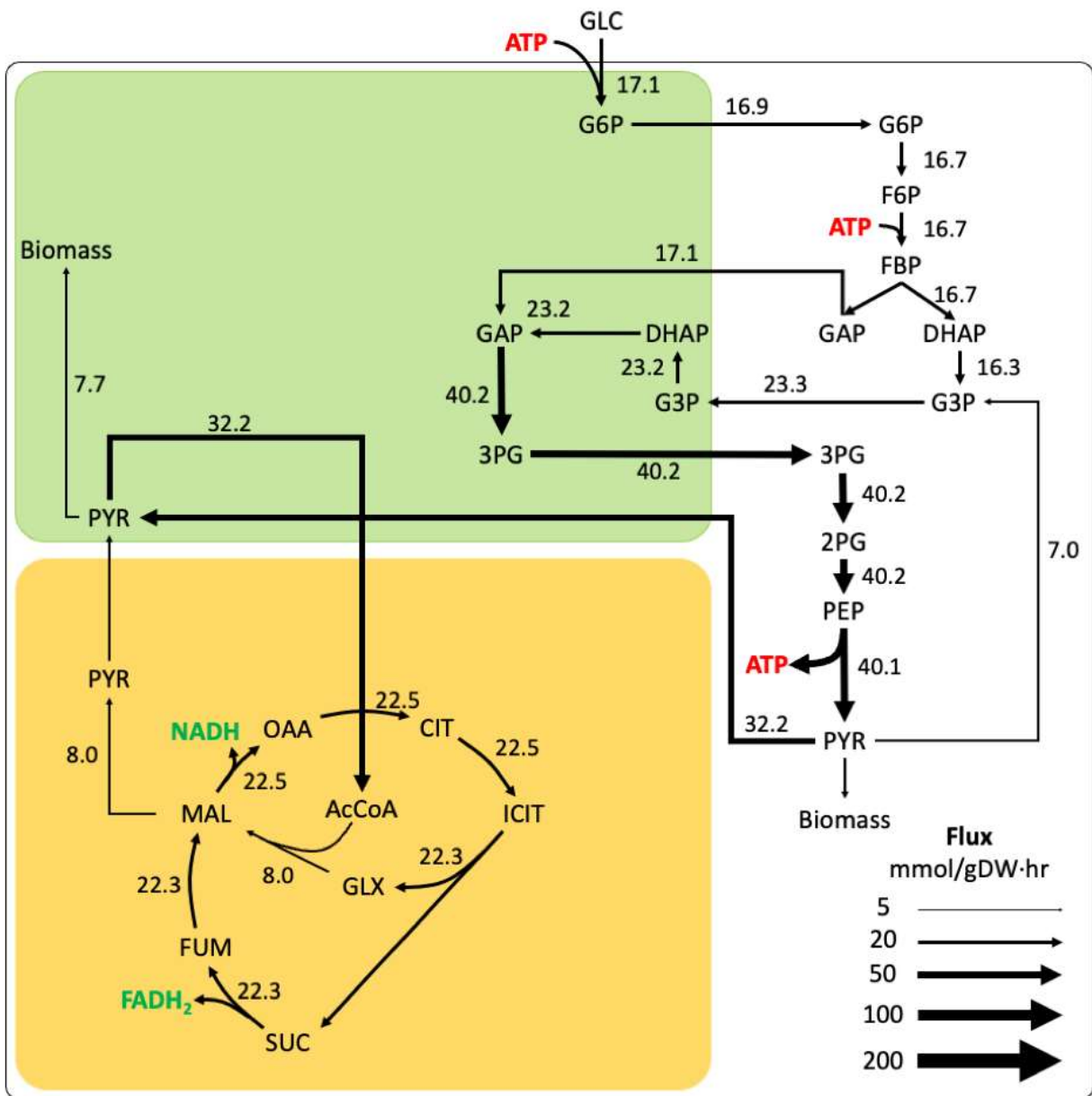


Figure 3.6 In silico simulated heterotrophic growth of *A. protothecoides* grown on glucose (APM1 media) using experimental data to constrain the model. The predicted fluxes are normalized to a glucose uptake of 16.6mmol/g·DW·hr (100mmol total carbon) for ease of comparison with the other trophic condition results.

3.4.5 Essential Gene Analysis

FBA was used to perform a gene essentiality analysis for both photoautotrophic and mixotrophic growth in continuous light with CO₂ uptake unconstrained in both conditions. The experimentally determined glucose uptake rate of 0.94 mmol/g·DW·hr was used for mixotrophic growth. Essential genes are those that do not allow the model to produce biomass when knocked out. There were 65 genes found to be essential for photoautotrophic growth and 63 for mixotrophic growth. With the size of the model the number of essential genes may appear small because metabolism is in many cases redundant; multiple genes can carry out the same function¹⁵³. The key difference between the essential genes under autotrophic and mixotrophic growth conditions are the genes encoding ribulose-1 5-bisphosphate carboxylase/oxygenase and phosphoribulokinase which are, as expected, essential in the photoautotrophic conditions. Both genes encode enzymes required to catalyze the reactions in photosynthesis that are light independent but regenerate substrate for the light dependent reactions to take place.

3.5 Discussion

3.5.1 Reconstruction of the genome-scale metabolic model

The development of the genome-scale metabolic model for *Auxenochlorella protothecoides* followed an iterative workflow that began with the automated generation of a draft model using RAPS⁸⁵. This algorithm employs BLASTp to compare annotated genes of the target organism against those from template organism models. For *A. protothecoides*, the template models used were those from the organisms *Chlamydomonas reinhardtii* (iCre1355)⁶⁸ and *Nannochloropsis gaditana* (iRJ1321)⁹⁸. The lack of standardization in reaction and metabolite annotation across these models posed challenges in ensuring that reactions and metabolites were correctly aligned between templates. The annotation for *A. protothecoides* introduced additional complexity due to the organism's polyploidy, resulting in multiple copies (alleles) of each gene. This increased the runtime of RAPS, as the algorithm had to handle a more extensive gene dataset. Despite these challenges, the draft model was successfully generated on a local machine (MacBook Pro, 2020, M1 chip, 16 GB RAM) within 24 hours.

Manual curation of the draft model was necessary to refine its accuracy and ensure biological plausibility. This included removing specific reactions from the *Nannochloropsis gaditana* model that did not align with *Chlamydomonas reinhardtii* reactions, particularly those that were overly

specific to the original model's biology. Additionally, reactions related to energy-generating cycles required curation to better reflect the metabolic characteristics of *A. protothecoides*. Some reactions involving the eyespot and flagella, which are absent in this organism, were also removed. Overall, the process highlights the iterative nature of development. Each refinement step, including addressing polyploidy-related complexities and ensuring alignment between template models, contributes to building a robust and accurate model for *A. protothecoides*.

3.5.2 Biomass composition

The growth rates for heterotrophic and mixotrophic growth on glucose were expectedly considerably higher than the autotrophic growth rate with atmospheric CO₂ (Table 3.1). This large disparity in growth rates has been demonstrated previously in other strains of *A. protothecoides*¹⁵⁴ which is in part due to *A. protothecoides* lacking a pyrenoid¹⁵⁵. Because of the substantially higher growth rates while utilizing an organic carbon source a great deal of the research in *A. protothecoides* has been focused on mixotrophic and heterotrophic growth¹⁵⁶⁻¹⁵⁸. *A. protothecoides* also shows a considerable difference in biomass composition when grown with glucose as an organic carbon source (Figure 3.2). When grown in nutrient replete conditions with 20g/L glucose there was not a substantial shift in the total lipid content present in the cells between autotrophic and mixotrophic cultures. This is consistent with others observation in *A. protothecoides* that show that for transition of starch to triglycerides to occur, continuous stress is required¹⁵⁹. There was however a substantial shift in biomass composition from protein in autotrophic cultures to carbohydrates making up the largest portion of the biomass. The shift was most pronounced in heterotrophic cultures where over 90% of total carbohydrates were starch. Mixotrophic cultures also increased total starch content but were still proportionally with the starch to total carbohydrate ratio seen in autotrophic cultures.

The proportion of starch to total carbohydrates is only one aspect of a considerable shift in metabolism due to the presence of light. The other data that demonstrates a considerable shift in metabolism is that growth rates and glucose uptake are both lower in mixotrophic cultures than the heterotrophic cultures. The only difference between the culturing conditions was that mixotrophic cultures were grown with light and heterotrophic cultures were grown in the dark. This inhibition in growth rate and carbon storage due to light has been demonstrated in other strains of *A. protothecoides*^{160,161}. It is hypothesized that the inhibition is a result of a significant

increases in glyceraldehyde-3-phosphate (G3P) diverting flux from glycolysis and lower glucose uptake¹⁶¹.

Another considerable change that is a result of different light conditions is that when grown heterotrophically almost all the chlorophyll in the cells is degraded. There is also a substantial reduction in chlorophyll between autotrophic cultures and mixotrophic cultures however mixotrophic cultures still maintain some chlorophyll content. This degreening process is well known in *A. protothecoides*, chlorophyll and light harvesting proteins are degraded when an organic carbon source is available¹⁶². Degreening may be responsible for the reduction in protein content in the glucose media cultures as even the mixotrophic culture while still retaining some chlorophyll has a much lower proportion than autotrophic cultures. The degreening process additionally offers benefits as it can lower harvesting costs as chlorophyll doesn't need to be separated from the lipids harvested¹⁶³.

The fatty acids profile in *A. protothecoides* when grown on glucose has a sizable difference dependent on trophic conditions Figure 3.3. The vast majority of the fatty acids across each condition were 16 and 18 carbon fatty acids ideal for biodiesel production¹⁶⁴. The most abundant fatty acid in autotrophic cultures was linolenic acid (18:3) while in mixotrophic and heterotrophic cultures the most abundant fatty acid was oleic acid (18:1). This aligns with fatty acid data from previous studies where increasing glucose in media resulted in the conversion of linolenic acid (18:3) to oleic acid (18:1)¹⁶⁵. The ability to shift the fatty acid profile towards monounsaturated fatty acids further demonstrates the potential that *A. protothecoides* has as a cell factory. This is because monounsaturated fatty acids have garnered particular interest for biodiesel production^{166,167}.

3.5.3 Flux distribution

pFBA simulations were conducted to predict flux distributions under autotrophic, mixotrophic, and heterotrophic conditions. Under autotrophic growth (Figure 3.4), the model projects a high flux through the Calvin–Benson cycle and its intermediates, aligning with the expectation that this cycle serves as the sole energy-generating pathway under strictly autotrophic conditions. This reliance on the Calvin–Benson cycle is further underscored by the necessity of maintaining substrate availability¹⁵³, especially given that *A. protothecoides* lacks a pyrenoid¹⁵⁵. However, the minimal flux predicted toward starch or other storage compounds conflicts with the established importance of energy reserves for photosynthetic organisms, which

must sustain metabolic activity during dark periods¹⁶⁸. This discrepancy most likely arises because the model's objective function, which maximizes CO₂ production, does not account for the organism's requirement to stockpile energy in anticipation of times when photosynthesis is not feasible. This highlights the nuance that is required in determining an objective function that captures diurnal cycles and storage demands.

Based on the model predictions for the mixotrophic condition (Figure 3.4), where most flux proceeds through glycolysis and into the mitochondrial TCA cycle while also engaging the glyoxylate shunt, it appears that the organism preferentially channels exogenous carbon through heterotrophic pathways despite the presence of light. This outcome aligns with observations in many algae that, when supplied with organic substrates, tend to reduce their reliance on photosynthesis and instead leverage glycolysis and downstream oxidative pathways for energy¹⁶⁹. This is supported by *A. protothecoides* removing essentially all its chlorophyll even in the presence of light when provided with an organic carbon source. The predicted use of the pentose phosphate pathway (PPP) for gluconeogenesis further suggests a need for NADPH and ribose-5-phosphate for biosynthesis, which could reflect either increased anabolic demands or a metabolic strategy for maintaining redox balance. This opens the possibility of future work to experimentally determine fluxes to confirm whether the model's predictions accurately capture *in vivo* flux partitioning or potentially overlook regulatory constraints.

Under strictly heterotrophic conditions (Figure 3.6), the model predicts a pronounced emphasis on cytosolic glycolysis and mitochondrial TCA cycle activity, with comparatively limited flux directed toward biomass generation. This outcome is consistent with an organism solely reliant on an organic carbon source, for which oxidative pathways are prioritized to meet immediate energy and metabolic demands¹⁷⁰. The model also indicates that *A. protothecoides* engages in gluconeogenesis from pyruvate to glyceraldehyde-3-phosphate (G3P), which could serve both anabolic and redox-balancing roles^{161,171}. These findings suggest that, in the absence of light, the cells rely heavily on high flux through catabolic routes to power growth while maintaining essential metabolic functions. However, as with the mixotrophic predictions, experimental validation of these heterotrophic fluxes would be valuable for identifying potential regulatory layers or kinetic factors not captured by the current model with experimental fluxes explored in the next chapter.

3.6 Conclusion

Auxenochlorella protothecoides holds a great deal of promise as a cell factory in the production of biofuels. UTEX 250 is a strain of particular interest due to ease of engineering¹¹ and its demonstrated rapid growth rate doubling in 11 hours during exponential growth. In this paper we present the GEM *iApro4643* representing the latest and most comprehensive model for *A. protothecoides* comprising 2820 metabolites and 3357 reactions distributed across seven compartments. The model was used to run pFBA simulations to determine fluxes under autotrophic, mixotrophic and heterotrophic conditions. *iApro4643* was constructed using RAPS²⁵ which aided greatly in removing much of the time consuming process of manual reconstruction. With computation tools such as *iApro4643* future metabolic engineering can be guided in a less time consuming and more productive manner.

CHAPTER 4

¹³C METABOLIC FLUX ANALYSIS OF HETEROTROPHICALLY GROWN AUXENOCHLORELLA PROTOTHECOIDES UTEX 250

Based on a paper in preparation for publication

Jacob Tamburro¹, Nanette R. Boyle^{1,2}

4.1 Abstract

Microalgae-based biofuels have emerged as a promising avenue for reducing reliance on fossil fuels, owing to the rapid growth, high photosynthetic efficiency, and minimal land requirements of microalgal systems. *Auxenochlorella protothecoides* (UTEX 250) is particularly appealing due to its robust lipid accumulation under nitrogen deprivation and its capacity for homologous recombination, which simplifies precise genetic engineering. To elucidate the organism's metabolic capabilities, a genome-scale metabolic model (GEM), *iApro4643*, was reconstructed to capture the behavior of *A. protothecoides* under various growth modes; however, accurate in silico predictions require robust experimental validation. In this study, we performed a comprehensive ¹³C metabolic flux analysis (MFA) under heterotrophic conditions using uniformly and partially labeled ¹³C glucose, tracking isotopic labeling in multiple metabolite pools fatty acids, amino acids, organic acids, and starch and cell wall sugars via gas chromatography mass spectrometry (GC–MS). These isotopic data were integrated in the INCA software suite to estimate intracellular fluxes, revealing that *A. protothecoides* directs most carbon through glycolysis into the mitochondrial tricarboxylic acid (TCA) cycle and then preferentially uses the glyoxylate shunt at the isocitrate branch point, while a substantial fraction of carbon is diverted toward starch synthesis, a phenomenon underestimated by the GEM. Our findings highlight the need for experimental validation to refine computational models and open new opportunities for engineering strategies, such as reducing starch biosynthesis and enhancing lipid formation, to improve the commercial viability of *A. protothecoides* for biofuel production.

¹ Colorado School of Mines, Quantitative Bioscience Engineering

² Colorado School of Mines, Chemical and Biological Engineering

4.2 Introduction

Increasing global awareness of climate change has amplified the search for sustainable energy sources that can mitigate the dependence on fossil fuels. Biofuels derived from renewable biomass offer a promising route for achieving lower carbon emissions and reduced environmental impact. Among the various forms of biomass, microalgae have garnered particular attention due to their rapid growth rates, minimal land requirements, and high photosynthetic efficiency¹²⁷. These characteristics not only make microalgae excellent candidates for biofuel production but also enable the simultaneous generation of valuable co-products such as carotenoids, which have nutritional relevance¹³⁰.

Auxenochlorella protothecoides (UTEX 250) is a compelling microalgal strain for biofuel applications, owing to its ability to accumulate high lipid content up to 60% under nitrogen deprivation¹²⁷ and its rapid growth rates under heterotrophic conditions. Moreover, *A. protothecoides* can be engineered at relatively low cost because it performs homologous recombination¹³². This attribute, more commonly associated with yeast, streamlines precise genetic modifications and thus accelerates the development of tailored production strains¹³³. *A. protothecoides* also synthesizes carotenoids such as lutein and zeaxanthin, enhancing its overall commercial viability as a feedstock for both biofuel and high-value products¹³¹.

To optimize these desirable traits, the genome-scale metabolic model (GEM) *iApro4643* was reconstructed. *iApro4643*, encompassing 2820 metabolites and 3357 reactions, to elucidate the metabolic networks of *A. protothecoides*. The metabolism of *A. protothecoides* was determined *in silico* by *iApro4643* under autotrophic, mixotrophic and heterotrophic growth conditions. While the fluxes determined by *iApro4643* offer valuable insight comparison with experimental flux values is important for refining and contextualizing the model's predictions.

¹³C metabolic flux analysis (MFA) offers a powerful means of experimentally determining *in vivo* flux distributions within an organism's metabolic network, complementing the insights gained from *in silico* models like *iApro4643*. To perform ¹³C MFA, cultures are fed isotopically labeled carbon sources often mixtures of ¹³C- and ¹²C-labeled substrates under conditions designed to maintain metabolic steady state. As the organism consumes these substrates, the heavier ¹³C atoms are incorporated into the organism's intracellular metabolites. These labeled metabolites are then extracted and analyzed using high-resolution instruments such as GC-MS, which can distinguish between isotopologues based on differences in mass. The resulting isotopic labeling patterns reflect the underlying fluxes through metabolic pathways, as specific branch

points and enzymatic reactions create distinct isotopomer distributions. By applying computational tools like isotopic network compartment analysis (INCA)⁹, researchers systematically reconcile the measured labeling data with a metabolic network model to estimate fluxes that best explain the observed isotopomer frequencies. This approach enables high-resolution mapping of carbon flow across central and secondary metabolic pathways, providing critical validation points for GEMs and identifying key nodes or bottlenecks for targeted metabolic engineering¹⁷².

In the present study, we describe the methods used to evaluate the model's predictions under heterotrophic growth conditions using uniformly and partially labeled ¹³C glucose. By integrating experimental labeling data with isotopic network compartment analysis (INCA)⁹, we demonstrate how these approaches facilitate accurate measurement of metabolic fluxes and critical intracellular pathways. This methodology elucidates the fate of carbon through key metabolic nodes, laying the groundwork for further refinement of GEM predictions and offering valuable insights into improving biofuel production in *A. protothecoides*.

4.3 Method

4.3.1 ¹³C Growth Conditions and Sampling

Cultures were grown heterotrophically in 20% uniformly ¹³C glucose, 20% 1,2 labelled ¹³C glucose and 60% unlabeled ¹²C glucose for the labeling experiment. The cultures were grown in seed flasks with the labeled glucose until reaching mid exponential growth and were used to seed the experimental flasks. The experimental and seed flasks were grown in the conditions referenced in the strain and growth conditions section of chapter 3. The experimental flasks had glucose uptake samples taken during exponential growth for glucose uptake rate and 1mL samples for ¹³C MFA collected at mid exponential (OD 7.0) were immediately quenched. Quenching was performed by freezing saline solution until it formed a slushy mixture at a 4 to 1 ratio of the samples in 15mL falcon tubes. Once the quenching solution is ready samples are added to the solution and centrifuged for 10 minutes at 1800×g and -1°C. The pellets and 500uL of solution were then transferred to 2mL tubes and centrifuged for 10 minutes at 8000×g at -1°C; after centrifuging the supernatant was removed and the pellets were stored at -20°C.

4.3.2 ¹³C Metabolic Flux Analysis

Isotope labelling data was collected on a GC-MS for fatty acids, amino acids, organic acids, starch sugars and cell wall sugars. The methods for the determination of fatty acids and amino acids were identical to those described in the biomass composition section, while the starch method was identical up until the glucose measurement step. For organic acid derivatization cell pellets were put through a freezing cycle of 30 seconds in liquid nitrogen followed by 30 seconds on ice with the process repeated 3 times to aid in extraction of metabolites. The cell pellets were then dried in a SpeedVac to remove any remaining water present after which 50uL of MOX reagent was added to dried cell pellet then vortexed and incubated at 40°C for 90 minutes. After the incubation 70uL of N-tert-Butyldimethylsilyl-N-methyltrifluoroacetamide + 1% Tertbutyldimethylchlorosilane was added and incubated at 70°C for 30 minutes. Samples were then incubated at room temperature overnight as this appears to improve peak strength and run on an Agilent 6890 GC with a 5973 mass spectrometer with a DB-1701 column¹⁷³.

For starch sugars the change that was made from the method described in the biomass composition section was that the supernatant sample was dried in a SpeedVac. The cell wall sugars were extracted by adding 50uL of 6M HCl to a cell pellet and incubated for 1 hour at 30°C followed by the addition of 250uL of water and incubation at 110°C for 1 hour. After the incubation the sample was cooled to room temperature and 50uL of 5M NaOH was added to neutralize the acid, the solution was filtered through a .22um spin-X centrifuge filter and dried in a SpeedVac. Both the cell wall and starch sugars were derivatized using the same method for determination on the GC-MS 50uL hydroxylamine hydrochloride 2% solution in pyridine was added to each dried pellet then vortexed and incubated at 90°C for 1 hour. After the incubation, the samples were then centrifuged at 10,000×g for 30 seconds and 100uL of propionic anhydride was added. Following the addition of propionic anhydride the sample was incubated at 60°C for 30 minutes then centrifuged at 10,000×g for 30 seconds with the supernatant being taken and run on an Agilent 6890 GC with a 5973 mass spectrometer with a DB-1701 column¹⁷⁴. The enrichments determined in the above methods were then integrated into INCA⁹ with an assumed error of 3%. To determine metabolic fluxes the network was constrained by the measured enrichments, glucose uptake rate and ethanol production rate.

4.4 Results

4.4.1 ^{13}C Metabolic Flux Analysis

^{13}C MFA was performed on heterotrophic *A. protothecoides* with glucose as the sole carbon source with the culture grown in the dark. Samples ^{13}C enrichment samples were taken at mid-exponential phase after which the samples were run on a GS-MS to determine relative abundance of ^{13}C carbon. The enrichment data was then processed with INCA⁹ to determine fluxes and a flux map was produced (Figure 4.1) from these experimentally determined fluxes. The fluxes determined by INCA primarily utilize glycolysis followed by the tricarboxylic acid (TCA) cycle in the mitochondria up until isocitrate (ICT) then favoring the glyoxylate shunt. There is also a considerable amount of flux to starch synthesis with approximately a third of flux leading to starch synthesis.

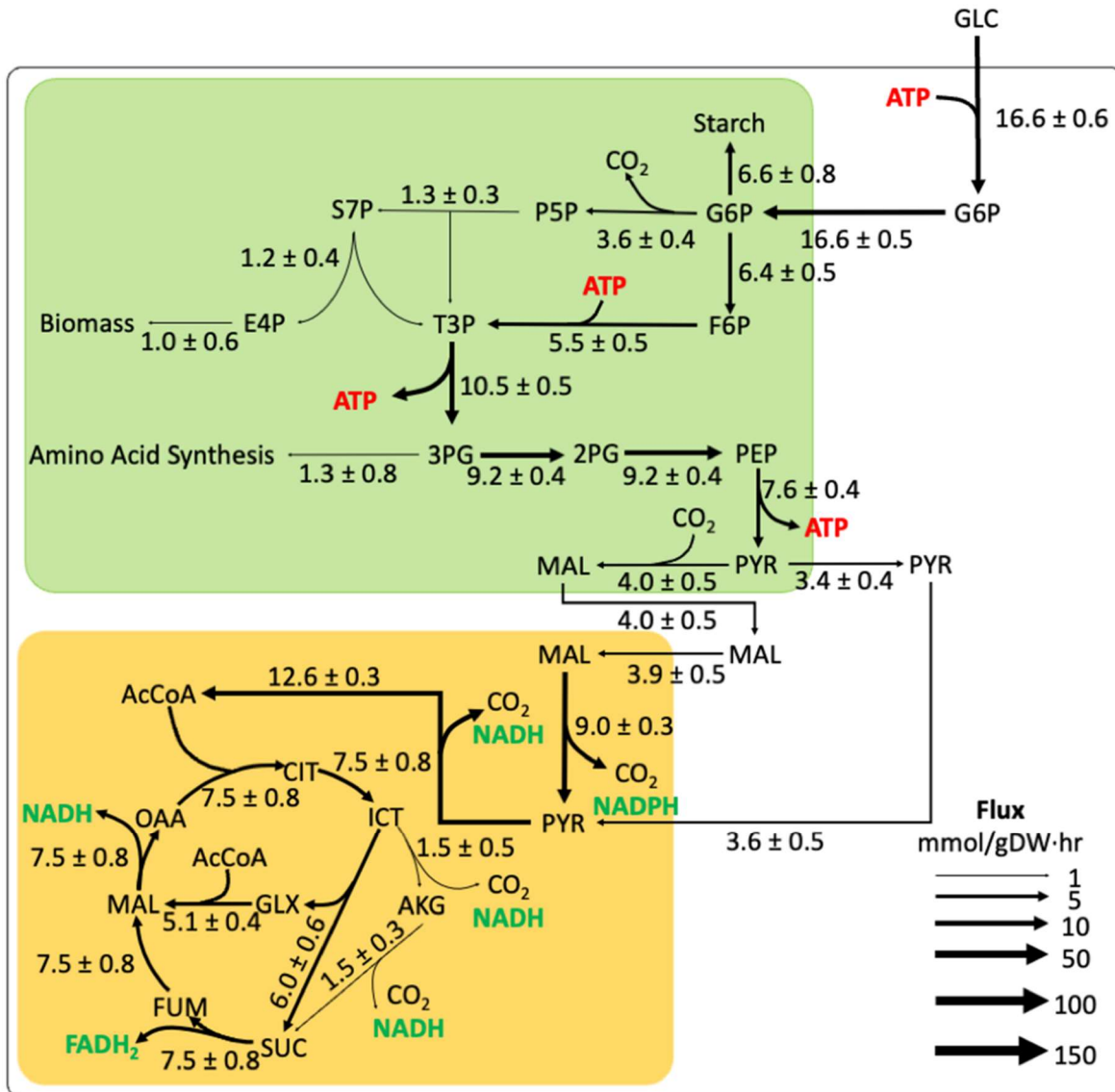


Figure 4.1 Experimentally determined fluxes of *A. protothecoides* grown on glucose (APM1 media) using carbon enrichment data and INCA. The fluxes are normalized to a glucose uptake of 16.6 mmol/gDW-hr (100 mmol total carbon) for ease of comparison with the GEM results. The four compartments are labelled on the flux map.

4.5 Discussion

The results of the ^{13}C metabolic flux analysis (MFA) confirm that *A. protothecoides* grown heterotrophically on glucose channels most of its carbon through glycolysis and into the mitochondrial (TCA) cycle until (ICT), where it favors the glyoxylate shunt. These findings are broadly consistent with the predictions made by the *iApro4643* genome-scale metabolic model (GEM), which likewise suggests that heterotrophic metabolism in *A. protothecoides* is dominated by mitochondrial TCA cycle activity. While the exact flux values did not align perfectly one-to-

one with the model's predictions, their proportional distributions were comparable. Such discrepancies between measured and modeled fluxes may stem from biological phenomena not fully accounted for in the GEM, such as potential enzyme degradation by reactive oxygen species¹⁷⁵ or unmodeled regulatory effects that can alter metabolic flows *in vivo*¹⁷⁶.

An additional difference in the experimentally derived fluxes and the model's output was the substantial flux to starch synthesis. The model, based on an objective function intent to maximize growth, did not predict major starch accumulation. Instead, it prioritized direct biomass precursors and energy production over storage compounds. In contrast, the *in vivo* observations revealed that roughly one-third of the carbon flux was directed toward starch, suggesting that *A. protothecoides* invests more heavily in carbon storage under these conditions than the model would anticipate. Many microorganisms uptake nutrients in excess of needs to compete with other microorganisms as well as to prepare for eventuality that nutrients become scarce, factors that the model does not consider¹⁷⁷. This discrepancy highlights the importance of validating computational predictions with experimental data, as purely *in silico* approaches may overlook physiological strategies employed by the organism particularly those related to storage or other non-optimal metabolic modes.

From an applied perspective, a high proportion of carbon routed to starch is less desirable if the goal is to maximize fatty acid (and thus biodiesel) production¹⁷⁸. Currently, nitrogen deprivation is often used to induce lipid accumulation, but this strategy comes at the expense of reduced growth rates^{159,179}. Identifying methods to shift *A. protothecoides* metabolism more directly toward fatty acid production without incurring a significant growth penalty is then a key challenge.

To steer *A. protothecoides* away from starch and toward lipid biosynthesis, several strategies may be considered. Gene knockouts or knockdowns of critical enzymes in the starch biosynthetic pathway such as ADP-glucose pyrophosphorylase (APG)¹⁸⁰ could diminish starch accumulation, redirecting carbon into lipid pathways¹⁸¹. Preventing starch synthesis by this method does however come with trade offs resulting in growth inhibition in some cases¹⁸¹. UDP-glucose pyrophosphorylase (UGP) is potentially the better target that has been demonstrated through the targeting of UGP1 and UGP2 to have no reduction in growth compared with the wild type while increasing lipid content¹⁸². Conversely, overexpression of enzymes involved in fatty acid synthesis, like acetyl-CoA carboxylase (ACC) or diacylglycerol acyltransferase (DGAT), have been demonstrated to bolster the capacity for lipid production¹⁸³⁻¹⁸⁵. The increased lipid content

also comes without the growth penalty observed in the knockout of AGP¹⁸³⁻¹⁸⁵. These engineering approaches can be integrated into an iterative cycle of GEM-guided design and experimental validation.

4.6 Conclusion

Overall, the present study provides critical experimental validation of the *A. protothecoides* genome-scale metabolic model *iApro4643* through ¹³C metabolic flux analysis, highlighting both congruencies and deviations between in silico predictions and in vivo measurements. The organism's pronounced routing of carbon toward starch, rather than exclusively into biomass precursors or lipid accumulation, emphasizes the physiological strategies it employs and underscores the limitations of relying solely on computational predictions that prioritize optimal growth. These observations have important implications for engineering *A. protothecoides* as a biofuel cell factory, indicating that enhancing lipid productivity may necessitate targeted genetic modifications to reduce starch synthesis or upregulate fatty acid biosynthetic pathways. Future iterations of the GEM could integrate this flux data, alongside regulatory and post-translational considerations, to yield more accurate simulations that guide metabolic engineering efforts. By bridging the gap between computational modeling and empirical flux measurements, researchers can develop more robust strategies to optimize *A. protothecoides* for sustainable biofuel production and further refine our understanding of the metabolic complexity underlying microalgal growth.

CHAPTER 5

CONCLUSION

The state current state of genome-scale metabolic models (GEMs) in microalgae has greatly improved by iteratively building off one another. GEMs have evolved from relatively simplistic frameworks to powerful predictive tools by incorporating photoautotrophic features, omics data, and dynamic simulations. Many challenges remain, such as bridging the gap between static assumptions and the complex diel cycles inherent in photosynthetic organisms however improvements continue to be made. The successful reconstruction of the *Auxenochlorella protothecoides* GEM *iApro4643*, which captures trophic flexibility (autotrophic, mixotrophic, and heterotrophic) as well as condition-specific biomass compositions demonstrates another step forward in algal GEM development. This model underscores how automating the reconstruction process with refined algorithms (e.g., RAPS⁸⁵) can drastically reduce manual curation time, enabling researchers to focus on critical processes like energy-generating cycles and compartment-specific refinements. The reconstruction also reveals the complexity of determining objective functions in autotrophic organisms, particularly regarding energy storage (e.g., starch) that can be critical for microalgal survival.

¹³C metabolic flux analysis (MFA) on heterotrophically grown *A. protothecoides* validates major in silico predictions especially the organism's reliance on glycolysis, the TCA cycle, and the glyoxylate shunt while highlighting discrepancies in starch flux compared to the model's output. The experimental data confirm that carbon metabolism in *A. protothecoides* can be diverted into more favorable storage products, representing a significant consideration for biofuel-oriented metabolic engineering strategies. Aligning the GEM's objective function with in vivo physiology including carbon storage, environmental fluctuations, and regulatory constraints will thus be essential for accurately predicting metabolic responses in broader operational contexts. Moving forward, deeper integration of multi-omics data, kinetic constraints, and dynamic modeling frameworks offers the greatest potential for producing predictive GEMs that can guide metabolic engineering in real time, accelerating the development of microalgae as viable cell factories for sustainable biofuel production and high value bioproduct synthesis.

5.1 Future work

^{13}C Isotopically nonstationary metabolic flux analysis (INST-MFA) represents an adaptation of traditional Metabolic Flux Analysis (MFA) that focuses on metabolism during a transient state rather than at steady state. In this approach, cells growing under autotrophic conditions are fed a pulse of ^{13}C -labeled substrate, and samples are rapidly quenched over time. The key is to capture how the labeled carbon flows through metabolic pathways before the system reaches a fully labeled equilibrium. This is particularly critical for autotrophic organisms, which fix a single carbon atoms at a time via rubisco. If experiments were conducted only at metabolic steady state, all carbon intermediates would eventually become uniformly labeled, masking the ^{13}C distributions that allow for the cell's carbon flux to be determined.

Applying ^{13}C INST-MFA in autotrophs not only reveals how carbon flows through key pathways in real time but also helps identify novel targets for metabolic engineering. A future avenue of research could be utilizing ^{13}C INST-MFA to determine the flux of *A. protothecoides* under autotrophic conditions. By comparing experimental flux measurements to predictions generated by computational models, researchers can uncover potential bottlenecks or alternative pathways that might otherwise remain hidden. Such information is especially valuable when exploring the metabolism of promising mutants to determine metabolic differences with the wildtype strain.

In this context, two notable mutants from our collaborators at Berkley merit further investigation using INST-MFA. The first is a MLDP1 knockout¹⁸⁶, in which the loss of the major lipid droplet protein 1 causes the formation of a single large lipid droplet rather than multiple smaller ones. When nitrogen is reintroduced to the medium, this mutant struggles to mobilize the stored lipid, implying a disruption in lipid turnover. The second mutant overexpresses sedoheptulose-1,7-bisphosphatase (SBPase), a Calvin cycle enzyme previously shown in *Chlamydomonas reinhardtii* to enhance photosynthetic activity at high CO_2 concentration¹⁸⁷. Comparing flux profiles between these two mutants and the wild-type strain would provide critical insights into how changes in lipid storage or enhanced carbon fixation impact overall metabolism, ultimately guiding future metabolic engineering strategies.

REFERENCES

- 1 Eneh, O. C. A review on petroleum: Source, uses, processing, products and the environment. *Journal of applied sciences* **11**, 2084-2091 (2011).
- 2 Höök, M. & Tang, X. Depletion of fossil fuels and anthropogenic climate change—A review. *Energy Policy* **52**, 797-809 (2013). <https://doi.org/10.1016/j.enpol.2012.10.046>
- 3 Filonchyk, M., Peterson, M. P., Zhang, L., Hurynovich, V. & He, Y. Greenhouse gases emissions and global climate change: Examining the influence of CO₂, CH₄, and N₂O. *Science of The Total Environment* **935**, 173359 (2024). <https://doi.org/10.1016/j.scitotenv.2024.173359>
- 4 Liu, Z., Deng, Z., Davis, S. J. & Ciais, P. Global carbon emissions in 2023. *Nature Reviews Earth & Environment* **5**, 253-254 (2024). <https://doi.org/10.1038/s43017-024-00532-2>
- 5 Thoré, E. S. J., Muylaert, K., Bertram, M. G. & Brodin, T. Microalgae. *Current Biology* **33**, R91-R95 (2023). <https://doi.org/10.1016/j.cub.2022.12.032>
- 6 Makareviciene, V. & Sendzikiene, E. Application of Microalgae Biomass for Biodiesel Fuel Production. *Energies* **15** (2022). <https://doi.org/10.3390/en15114178>
- 7 Goshtasbi, H. *et al.* Harnessing microalgae as sustainable cellular factories for biopharmaceutical production. *Algal Research* **74**, 103237 (2023). <https://doi.org/10.1016/j.algal.2023.103237>
- 8 Park, S. Y., Yang, D., Ha, S. H. & Lee, S. Y. Metabolic Engineering of Microorganisms for the Production of Natural Compounds. *Advanced Biosystems* **2**, 1700190 (2018). <https://doi.org/10.1002/adbi.201700190>
- 9 Young, J. D. INCA: a computational platform for isotopically non-stationary metabolic flux analysis. *Bioinformatics* **30**, 1333-1335 (2014). <https://doi.org/10.1093/bioinformatics/btu015>
- 10 Ahmed, S. F. *et al.* Progress and challenges of contaminate removal from wastewater using microalgae biomass. *Chemosphere* **286**, 131656 (2022). <https://doi.org/10.1016/j.chemosphere.2021.131656>
- 11 Abu-Ghosh, S., Dubinsky, Z., Verdelho, V. & Iluz, D. Unconventional high-value products from microalgae: A review. *Bioresour Technol* **329**, 124895 (2021). <https://doi.org/10.1016/j.biortech.2021.124895>
- 12 Huntley, M. E. & Redalje, D. G. CO₂ Mitigation and Renewable Oil from Photosynthetic Microbes: A New Appraisal. *Mitigation and Adaptation Strategies for Global Change* **12**, 573-608 (2007). <https://doi.org/10.1007/s11027-006-7304-1>
- 13 Vasudevan, P. T. & Briggs, M. Biodiesel production—current state of the art and challenges. *Journal of Industrial Microbiology and Biotechnology* **35**, 421-421 (2008). <https://doi.org/10.1007/s10295-008-0312-2>

- 14 Chung, I. K., Beardall, J., Mehta, S., Sahoo, D. & Stojkovic, S. Using marine macroalgae for carbon sequestration: a critical appraisal. *Journal of Applied Phycology* **23**, 877-886 (2011). <https://doi.org/10.1007/s10811-010-9604-9>
- 15 Khanra, A. *et al.* Green bioprocessing and applications of microalgae-derived biopolymers as a renewable feedstock: Circular bioeconomy approach. *Environmental Technology & Innovation* **28** (2022). <https://doi.org/10.1016/j.eti.2022.102872>
- 16 Patel, A., Matsakas, L., Rova, U. & Christakopoulos, P. Heterotrophic cultivation of *Auxenochlorella protothecoides* using forest biomass as a feedstock for sustainable biodiesel production. *Biotechnol Biofuels* **11**, 169 (2018). <https://doi.org/10.1186/s13068-018-1173-1>
- 17 Polat, E., Yavuztürk-Gül, B., Ünver, H. & Altınbaş, M. Biotechnological product potential of *Auxenochlorella protothecoides* including biologically active compounds (BACs) under nitrogen stress conditions. *World Journal of Microbiology and Biotechnology* **39**, 198 (2023). <https://doi.org/10.1007/s11274-023-03642-z>
- 18 Harris, E. H. CHLAMYDOMONAS A MODEL ORGANISM. *Annual Review of Plant Biology* **52**, 363-406 (2001). <https://doi.org/10.1146/annurev.arplant.52.1.363>
- 19 Pandey, S., Kumar, P., Dasgupta, S., Archana, G. & Bagchi, D. Gradient Strategy for Mixotrophic Cultivation of *Chlamydomonas reinhardtii*: Small Steps, a Large Impact on Biofuel Potential and Lipid Droplet Morphology. *BioEnergy Research* **16**, 163-176 (2023). <https://doi.org/10.1007/s12155-022-10454-w>
- 20 Masi, A. *et al.* *Chlamydomonas reinhardtii*: A Factory of Nutraceutical and Food Supplements for Human Health. *Molecules* **28** (2023).
- 21 Tran, D. T., Van Do, T. C., Nguyen, Q. T. & Le, T. G. Simultaneous removal of pollutants and high value biomaterials production by *Chlorella variabilis* TH03 from domestic wastewater. *Clean Technologies and Environmental Policy* **23**, 3-17 (2020). <https://doi.org/10.1007/s10098-020-01810-5>
- 22 Sati, H., Chokshi, K., Soundarya, R., Ghosh, A. & Mishra, S. Seaweed-based biostimulant improves photosynthesis and effectively enhances growth and biofuel potential of a green microalga *Chlorella variabilis*. *Aquaculture International* **29**, 963-975 (2021). <https://doi.org/10.1007/s10499-021-00667-9>
- 23 Bito, T., Okumura, E., Fujishima, M. & Watanabe, F. Potential of *Chlorella* as a Dietary Supplement to Promote Human Health. *Nutrients* **12** (2020). <https://doi.org/10.3390/nu12092524>
- 24 Moradi, P. & Saidi, M. Biodiesel production from *Chlorella Vulgaris* microalgal-derived oil via electrochemical and thermal processes. *Fuel Processing Technology* **228**, 107158 (2022). <https://doi.org/10.1016/j.fuproc.2021.107158>

- 25 Zhang, Y., Ye, Y., Bai, F. & Liu, J. The oleaginous astaxanthin-producing alga *Chromochloris zofingiensis*: potential from production to an emerging model for studying lipid metabolism and carotenogenesis. *Biotechnol Biofuels* **14**, 119 (2021). <https://doi.org/10.1186/s13068-021-01969-z>
- 26 Vitali, L. *et al.* Lipid content and fatty acid methyl ester profile by *Chromochloris zofingiensis* under chemical and metabolic stress. *Biomass Conversion and Biorefinery* (2023). <https://doi.org/10.1007/s13399-023-04153-5>
- 27 Hu, H. *et al.* Salinity controlling enhanced high-salinity pickle wastewater treatment coupling with high-value fatty acid production by *Dunaliella salina*. *Journal of Cleaner Production* **448**, 141732 (2024). <https://doi.org/10.1016/j.jclepro.2024.141732>
- 28 Xi, Y., Bian, J., Luo, G., Kong, F. & Chi, Z. Enhanced β -carotene production in *Dunaliella salina* under relative high flashing light. *Algal Research* **67**, 102857 (2022). <https://doi.org/10.1016/j.algal.2022.102857>
- 29 Sheward, R. M., Gebühr, C., Bollmann, J. & Herrle, J. O. Short-term response of *Emiliania huxleyi* growth and morphology to abrupt salinity stress. *Biogeosciences* **21**, 3121-3141 (2024). <https://doi.org/10.5194/bg-21-3121-2024>
- 30 Aveiro, S. S. *et al.* The Polar Lipidome of Cultured *Emiliania huxleyi*: A Source of Bioactive Lipids with Relevance for Biotechnological Applications. *Biomolecules* **10** (2020).
- 31 Bayer-Giraldi, M., Uhlig, C., John, U., Mock, T. & Valentin, K. Antifreeze proteins in polar sea ice diatoms: diversity and gene expression in the genus *Fragilariopsis*. *Environmental Microbiology* **12**, 1041-1052 (2010). <https://doi.org/10.1111/j.1462-2920.2009.02149.x>
- 32 Vaezi, R., Napier, J. A. & Sayanova, O. Identification and functional characterization of genes encoding omega-3 polyunsaturated fatty acid biosynthetic activities from unicellular microalgae. *Mar Drugs* **11**, 5116-5129 (2013). <https://doi.org/10.3390/md11125116>
- 33 Guerin, S., Raguenes, L., Croteau, D., Babin, M. & Lavaud, J. Potential for the Production of Carotenoids of Interest in the Polar Diatom *Fragilariopsis cylindrus*. *Mar Drugs* **20** (2022). <https://doi.org/10.3390/md20080491>
- 34 Mularczyk, M., Michalak, I. & Marycz, K. Astaxanthin and other Nutrients from *Haematococcus pluvialis*-Multifunctional Applications. *Mar Drugs* **18** (2020). <https://doi.org/10.3390/md18090459>
- 35 Hosseini, A., Jazini, M., Mahdih, M. & Karimi, K. Efficient superantioxidant and biofuel production from microalga *Haematococcus pluvialis* via a biorefinery approach. *Bioresour Technol* **306**, 123100 (2020). <https://doi.org/10.1016/j.biortech.2020.123100>
- 36 Alkhamis, Y. & Qin, J. G. Cultivation of *Isochrysis galbana* in Phototrophic, Heterotrophic, and Mixotrophic Conditions. *BioMed Research International* **2013**, 983465 (2013). <https://doi.org/10.1155/2013/983465>

- 37 Matos, J. *et al.* Bioprospection of *Isochrysis galbana* and its potential as a nutraceutical. *Food & Function* **10**, 7333-7342 (2019). <https://doi.org/10.1039/C9FO01364D>
- 38 Sánchez, Á., Maceiras, R., Cancela, Á. & Pérez, A. Culture aspects of *Isochrysis galbana* for biodiesel production. *Applied Energy* **101**, 192-197 (2013). <https://doi.org/10.1016/j.apenergy.2012.03.027>
- 39 Mitra, M., Patidar, S. K., George, B., Shah, F. & Mishra, S. A euryhaline *Nannochloropsis gaditana* with potential for nutraceutical (EPA) and biodiesel production. *Algal Research* **8**, 161-167 (2015). <https://doi.org/10.1016/j.algal.2015.02.006>
- 40 Koh, H. G. *et al.* Optimization and mechanism analysis of photosynthetic EPA production in *Nannochloropsis salina*: Evaluating the effect of temperature and nitrogen concentrations. *Plant Physiol Biochem* **211**, 108729 (2024). <https://doi.org/10.1016/j.plaphy.2024.108729>
- 41 Fakhry, E. M. & El Maghraby, D. M. Lipid accumulation in response to nitrogen limitation and variation of temperature in *Nannochloropsis salina*. *Bot Stud* **56**, 6 (2015). <https://doi.org/10.1186/s40529-015-0085-7>
- 42 Butler, T., Kapoore, R. V. & Vaidyanathan, S. *Phaeodactylum tricornutum*: A Diatom Cell Factory. *Trends in Biotechnology* **38**, 606-622 (2020). <https://doi.org/10.1016/j.tibtech.2019.12.023>
- 43 Qin, S., Liu, G.-X. & Hu, Z.-Y. The accumulation and metabolism of astaxanthin in *Scenedesmus obliquus* (Chlorophyceae). *Process Biochemistry* **43**, 795-802 (2008). <https://doi.org/10.1016/j.procbio.2008.03.010>
- 44 Yang, L. *et al.* Removal of ofloxacin with biofuel production by oleaginous microalgae *Scenedesmus obliquus*. *Bioresour Technol* **315**, 123738 (2020). <https://doi.org/10.1016/j.biortech.2020.123738>
- 45 Bouras, S., Katsoulas, N., Antoniadis, D. & Karapanagiotidis, I. T. Use of Biofuel Industry Wastes as Alternative Nutrient Sources for DHA-Yielding *Schizochytrium limacinum* Production. *Applied Sciences* **10** (2020).
- 46 Bi, Z., He, B. B. & McDonald, A. G. Biodiesel Production from Green Microalgae *Schizochytrium limacinum* via in Situ Transesterification. *Energy & Fuels* **29**, 5018-5027 (2015). <https://doi.org/10.1021/acs.energyfuels.5b00559>
- 47 Armbrust, E. V. *et al.* The Genome of the Diatom *Thalassiosira Pseudonana*: Ecology, Evolution, and Metabolism. *Science* **306**, 79-86 (2004). <https://doi.org/10.1126/science.1101156>
- 48 Peng, M. *et al.* Effects of light quality on the growth, productivity, fucoxanthin accumulation, and fatty acid composition of *Thalassiosira pseudonana*. *Journal of Applied Phycology* **36**, 1667-1678 (2024). <https://doi.org/10.1007/s10811-024-03245-7>

- 49 El-Sheekh, M. M., Galal, H. R., Mousa, A. S. H. H. & Farghl, A. A. M. Improving the biodiesel production in the marine diatom *Thalassiosira pseudonana* cultivated in nutrient deficiency and sewage water. *Environmental Science and Pollution Research* **31**, 63764-63776 (2024). <https://doi.org/10.1007/s11356-024-35409-w>
- 50 Bošnjaković, M. & Sinaga, N. The Perspective of Large-Scale Production of Algae Biodiesel. *Applied Sciences* **10** (2020).
- 51 Acién, F. G., Fernández, J. M., Magán, J. J. & Molina, E. Production cost of a real microalgae production plant and strategies to reduce it. *Biotechnology Advances* **30**, 1344-1353 (2012). <https://doi.org/10.1016/j.biotechadv.2012.02.005>
- 52 Awasthi, M. K. *et al.* Refining biomass residues for sustainable energy and bio-products: An assessment of technology, its importance, and strategic applications in circular bio-economy. *Renewable and Sustainable Energy Reviews* **127**, 109876 (2020). <https://doi.org/10.1016/j.rser.2020.109876>
- 53 Stichnothe, H., Storz, H., Meier, D., de Bari, I. & Thomas, S. in *Developing the Global Bioeconomy* (eds Patrick Lamers, Erin Searcy, J. Richard Hess, & Heinz Stichnothe) 11-40 (Academic Press, 2016).
- 54 Hu, J., Wang, D., Chen, H. & Wang, Q. Advances in Genetic Engineering in Improving Photosynthesis and Microalgal Productivity. *International Journal of Molecular Sciences* **24** (2023).
- 55 Yan, J., Kuang, Y., Gui, X., Han, X. & Yan, Y. Engineering a malic enzyme to enhance lipid accumulation in *Chlorella protothecoides* and direct production of biodiesel from the microalgal biomass. *Biomass and Bioenergy* **122**, 298-304 (2019). <https://doi.org/10.1016/j.biombioe.2019.01.046>
- 56 Song, I. *et al.* The generation of metabolic changes for the production of high-purity zeaxanthin mediated by CRISPR-Cas9 in *Chlamydomonas reinhardtii*. *Microbial Cell Factories* **19**, 220 (2020). <https://doi.org/10.1186/s12934-020-01480-4>
- 57 Bernstein, D. B., Sulheim, S., Almaas, E. & Segrè, D. Addressing uncertainty in genome-scale metabolic model reconstruction and analysis. *Genome Biology* **22**, 64 (2021). <https://doi.org/10.1186/s13059-021-02289-z>
- 58 Pareek, C. S., Smoczynski, R. & Tretyn, A. Sequencing technologies and genome sequencing. *J Appl Genet* **52**, 413-435 (2011). <https://doi.org/10.1007/s13353-011-0057-x>
- 59 Mekanik, M., Fotovat, R., Motamedian, E. & Jafarian, V. Improvement of Lutein Production in *Auxenochlorella protothecoides* Using Its Genome-Scale Metabolic Model and a System-Oriented Approach. *Appl Biochem Biotechnol* **195**, 889-904 (2023). <https://doi.org/10.1007/s12010-022-04186-y>

- 60 Levering, J. *et al.* Genome-Scale Model Reveals Metabolic Basis of Biomass Partitioning in a Model Diatom. *PLoS One* **11**, e0155038 (2016). <https://doi.org/10.1371/journal.pone.0155038>
- 61 Yang, J. E. *et al.* One-step fermentative production of aromatic polyesters from glucose by metabolically engineered *Escherichia coli* strains. *Nature Communications* **9**, 79 (2018). <https://doi.org/10.1038/s41467-017-02498-w>
- 62 van Tol, H. M. & Armbrust, E. V. Genome-scale metabolic model of the diatom *Thalassiosira pseudonana* highlights the importance of nitrogen and sulfur metabolism in redox balance. *PLoS One* **16**, e0241960 (2021). <https://doi.org/10.1371/journal.pone.0241960>
- 63 Zuniga, C. *et al.* Genome-Scale Metabolic Model for the Green Alga *Chlorella vulgaris* UTEX 395 Accurately Predicts Phenotypes under Autotrophic, Heterotrophic, and Mixotrophic Growth Conditions. *Plant Physiol* **172**, 589-602 (2016). <https://doi.org/10.1104/pp.16.00593>
- 64 Ofaim, S., Sulheim, S., Almaas, E., Sher, D. & Segre, D. Dynamic Allocation of Carbon Storage and Nutrient-Dependent Exudation in a Revised Genome-Scale Model of *Prochlorococcus*. *Front Genet* **12**, 586293 (2021). <https://doi.org/10.3389/fgene.2021.586293>
- 65 Nocon, J. *et al.* Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. *Metabolic Engineering* **24**, 129-138 (2014). <https://doi.org/10.1016/j.ymben.2014.05.011>
- 66 Calatrava, V. *et al.* *Chlamydomonas reinhardtii*, a Reference Organism to Study Algal-Microbial Interactions: Why Can't They Be Friends? *Plants (Basel)* **12** (2023). <https://doi.org/10.3390/plants12040788>
- 67 Boyle, N. R. & Morgan, J. A. Flux balance analysis of primary metabolism in *Chlamydomonas reinhardtii*. *BMC Systems Biology* **3**, 4 (2009). <https://doi.org/10.1186/1752-0509-3-4>
- 68 Imam, S. *et al.* A refined genome-scale reconstruction of *Chlamydomonas* metabolism provides a platform for systems-level analyses. *Plant J* **84**, 1239-1256 (2015). <https://doi.org/10.1111/tpj.13059>
- 69 Chang, R. L. *et al.* Metabolic network reconstruction of *Chlamydomonas* offers insight into light-driven algal metabolism. *Mol Syst Biol* **7**, 518 (2011). <https://doi.org/10.1038/msb.2011.52>
- 70 Arend, M. *et al.* Proteomics and constraint-based modelling reveal enzyme kinetic properties of *Chlamydomonas reinhardtii* on a genome scale. *Nature Communications* **14**, 4781 (2023). <https://doi.org/10.1038/s41467-023-40498-1>
- 71 Shene, C., Asenjo, J. A. & Chisti, Y. Metabolic modelling and simulation of the light and dark metabolism of *Chlamydomonas reinhardtii*. *The Plant Journal* **96**, 1076-1088 (2018). <https://doi.org/10.1111/tpj.14078>

- 72 Metcalf Alex, J. & Boyle Nanette, R. Rhythm of the Night (and Day): Predictive Metabolic Modeling of Diurnal Growth in *Chlamydomonas*. *mSystems* **7**, e00176-00122 (2022). <https://doi.org/10.1128/msystems.00176-22>
- 73 Yao, H., Dahal, S. & Yang, L. Novel context-specific genome-scale modelling explores the potential of triacylglycerol production by *Chlamydomonas reinhardtii*. *Microb Cell Fact* **22**, 13 (2023). <https://doi.org/10.1186/s12934-022-02004-y>
- 74 Bjerkelund Rokke, G., Hohmann-Marriott, M. F. & Almaas, E. An adjustable algal chloroplast plug-and-play model for genome-scale metabolic models. *PLoS One* **15**, e0229408 (2020). <https://doi.org/10.1371/journal.pone.0229408>
- 75 Karp, P. D., Weaver, D. & Latendresse, M. How accurate is automated gap filling of metabolic models? *BMC Systems Biology* **12**, 73 (2018). <https://doi.org/10.1186/s12918-018-0593-7>
- 76 Sen, P. & Orešič, M. Integrating Omics Data in Genome-Scale Metabolic Modeling: A Methodological Perspective for Precision Medicine. *Metabolites* **13** (2023).
- 77 Fisher, N. L. *et al.* Light-dependent metabolic shifts in the model diatom *Thalassiosira pseudonana*. *Algal Research* **74**, 103172 (2023). <https://doi.org/10.1016/j.algal.2023.103172>
- 78 Goss, R. & Jakob, T. Regulation and function of xanthophyll cycle-dependent photoprotection in algae. *Photosynthesis Research* **106**, 103-122 (2010). <https://doi.org/10.1007/s11120-010-9536-x>
- 79 Schubert, H., Andersson, M. & Snoeijs, P. Relationship between photosynthesis and non-photochemical quenching of chlorophyll fluorescence in two red algae with different carotenoid compositions. *Marine Biology* **149**, 1003-1013 (2006). <https://doi.org/10.1007/s00227-006-0265-9>
- 80 Schnurr, P. J., Espie, G. S. & Allen, G. D. The effect of photon flux density on algal biofilm growth and internal fatty acid concentrations. *Algal Research* **16**, 349-356 (2016). <https://doi.org/10.1016/j.algal.2016.04.001>
- 81 Orth, J. D. *et al.* A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Molecular Systems Biology* **7**, 535 (2011). <https://doi.org/10.1038/msb.2011.65>
- 82 Matos, Â. P., Cavanholi, M. G., Moecke, E. H. S. & Sant'Anna, E. S. Effects of different photoperiod and trophic conditions on biomass, protein and lipid production by the marine alga *Nannochloropsis gaditana* at optimal concentration of desalination concentrate. *Bioresource Technology* **224**, 490-497 (2017). <https://doi.org/10.1016/j.biortech.2016.11.004>
- 83 Jallet, D., Caballero, M. A., Gallina, A. A., Youngblood, M. & Peers, G. Photosynthetic physiology and biomass partitioning in the model diatom *Phaeodactylum tricornutum* grown in a sinusoidal light regime. *Algal Research* **18**, 51-60 (2016). <https://doi.org/10.1016/j.algal.2016.05.014>

- 84 Meagher, M. *et al.* Genome-scale metabolic model accurately predicts fermentation of glucose by *Chromochloris zofingiensis*. *Algal Research* **84**, 103805 (2024). <https://doi.org/10.1016/j.algal.2024.103805>
- 85 Metcalf, A. J., Nagygyor, A. & Boyle, N. R. Rapid Annotation of Photosynthetic Systems (RAPS): automated algorithm to generate genome-scale metabolic networks from algal genomes. *Algal Research* **50** (2020). <https://doi.org/10.1016/j.algal.2020.101967>
- 86 Devoid, S. *et al.* in *Systems Metabolic Engineering: Methods and Protocols* (ed Hal S. Alper) 17-45 (Humana Press, 2013).
- 87 Machado, D., Andrejev, S., Tramontano, M. & Patil, K. R. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Research* **46**, 7542-7553 (2018). <https://doi.org/10.1093/nar/gky537>
- 88 Seaver, S. M. D. *et al.* High-throughput comparison, functional annotation, and metabolic modeling of plant genomes using the PlantSEED resource. *Proceedings of the National Academy of Sciences* **111**, 9645-9650 (2014). <https://doi.org/10.1073/pnas.1401329111>
- 89 Catalanotti, C., Yang, W., Posewitz, M. C. & Grossman, A. R. Fermentation metabolism and its evolution in algae. *Frontiers in Plant Science* **4** (2013). <https://doi.org/10.3389/fpls.2013.00150>
- 90 Tamoi, M. & Shigeoka, S. Diversity of regulatory mechanisms of photosynthetic carbon metabolism in plants and algae. *Bioscience, Biotechnology, and Biochemistry* **79**, 870-876 (2015). <https://doi.org/10.1080/09168451.2015.1020754>
- 91 Schellenberger, J., Park, J. O., Conrad, T. M. & Palsson, B. Ø. BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinformatics* **11**, 213 (2010). <https://doi.org/10.1186/1471-2105-11-213>
- 92 Juneja, A., Chaplen, F. W. R. & Murthy, G. S. Genome scale metabolic reconstruction of *Chlorella variabilis* for exploring its metabolic potential for biofuels. *Bioresour Technol* **213**, 103-110 (2016). <https://doi.org/10.1016/j.biortech.2016.02.118>
- 93 Bailleul, B. *et al.* Energetic coupling between plastids and mitochondria drives CO₂ assimilation in diatoms. *Nature* **524**, 366-369 (2015). <https://doi.org/10.1038/nature14599>
- 94 Chen, H. *et al.* Ca²⁺-regulated cyclic electron flow supplies ATP for nitrogen starvation-induced lipid biosynthesis in green alga. *Scientific Reports* **5**, 15117 (2015). <https://doi.org/10.1038/srep15117>
- 95 Mekanik, M., Motamedian, E., Fotovat, R. & Jafarian, V. Reconstruction of a genome-scale metabolic model for *Auxenochlorella protothecoides* to study hydrogen production under anaerobiosis using multiple optimal solutions. *International Journal of Hydrogen Energy* **44**, 2580-2591 (2019). <https://doi.org/10.1016/j.ijhydene.2018.12.049>

- 96 Shah, A. R., Ahmad, A., Srivastava, S. & Jaffar Ali, B. M. Reconstruction and analysis of a genome-scale metabolic model of *Nannochloropsis gaditana*. *Algal Research* **26**, 354-364 (2017). <https://doi.org/10.1016/j.algal.2017.08.014>
- 97 Borges, P. T. *et al.* Photosynthetic green hydrogen: Advances, challenges, opportunities, and prospects. *International Journal of Hydrogen Energy* **49**, 433-458 (2024). <https://doi.org/10.1016/j.ijhydene.2023.09.075>
- 98 Loira, N. *et al.* Reconstruction of the microalga *Nannochloropsis salina* genome-scale metabolic model with applications to lipid production. *BMC Syst Biol* **11**, 66 (2017). <https://doi.org/10.1186/s12918-017-0441-1>
- 99 Lavoie, M. *et al.* Genome-Scale Metabolic Reconstruction and in Silico Perturbation Analysis of the Polar Diatom *Fragilariopsis cylindrus* Predicts High Metabolic Robustness. *Biology (Basel)* **9** (2020). <https://doi.org/10.3390/biology9020030>
- 100 Segrè, D., Vitkup, D. & Church, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences* **99**, 15112-15117 (2002). <https://doi.org/10.1073/pnas.232349399>
- 101 Yoshida, K., Seger, A., Kennedy, F., McMinn, A. & Suzuki, K. Freezing, Melting, and Light Stress on the Photophysiology of Ice Algae: Ex Situ Incubation of the Ice Algal diatom *Fragilariopsis cylindrus* (Bacillariophyceae) Using an Ice Tank. *Journal of Phycology* **56**, 1323-1338 (2020). <https://doi.org/10.1111/jpy.13036>
- 102 Recht, L. *et al.* Metabolite profiling and integrative modeling reveal metabolic constraints for carbon partitioning under nitrogen starvation in the green algae *Haematococcus pluvialis*. *J Biol Chem* **289**, 30387-30403 (2014). <https://doi.org/10.1074/jbc.M114.555144>
- 103 Gudmundsson, S. & Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinformatics* **11**, 489 (2010). <https://doi.org/10.1186/1471-2105-11-489>
- 104 Recht, L., Zarka, A. & Boussiba, S. Patterns of carbohydrate and fatty acid changes under nitrogen starvation in the microalgae *Haematococcus pluvialis* and *Nannochloropsis* sp. *Applied Microbiology and Biotechnology* **94**, 1495-1503 (2012). <https://doi.org/10.1007/s00253-012-3940-4>
- 105 Knies, D. *et al.* Modeling and Simulation of Optimal Resource Management during the Diurnal Cycle in *Emiliania huxleyi* by Genome-Scale Reconstruction and an Extended Flux Balance Analysis Approach. *Metabolites* **5**, 659-676 (2015). <https://doi.org/10.3390/metabo5040659>
- 106 Carthew, R. W. Gene Regulation and Cellular Metabolism: An Essential Partnership. *Trends in Genetics* **37**, 389-400 (2021). <https://doi.org/10.1016/j.tig.2020.09.018>
- 107 Gim, G. H., Ryu, J., Kim, M. J., Kim, P. I. & Kim, S. W. Effects of carbon source and light intensity on the growth and total lipid production of three microalgae under different culture

- conditions. *Journal of Industrial Microbiology and Biotechnology* **43**, 605-616 (2016). <https://doi.org/10.1007/s10295-016-1741-y>
- 108 Manichaikul, A. *et al.* Metabolic network analysis integrated with transcript verification for sequenced genomes. *Nature Methods* **6**, 589-592 (2009). <https://doi.org/10.1038/nmeth.1348>
- 109 Gomes de Oliveira Dal'Molin, C., Quek, L.-E., Palfreyman, R. W. & Nielsen, L. K. AlgaGEM – a genome-scale metabolic reconstruction of algae based on the *Chlamydomonas reinhardtii* genome. *BMC Genomics* **12**, S5 (2011). <https://doi.org/10.1186/1471-2164-12-S4-S5>
- 110 Chaiboonchoe, A. *et al.* Microalgal Metabolic Network Model Refinement through High-Throughput Functional Metabolic Profiling. *Frontiers in Bioengineering and Biotechnology* **2** (2014).
- 111 Winck, F. V. *et al.* Analysis of Sensitive CO₂ Pathways and Genes Related to Carbon Uptake and Accumulation in *Chlamydomonas reinhardtii* through Genomic Scale Modeling and Experimental Validation. *Front Plant Sci* **7**, 43 (2016). <https://doi.org/10.3389/fpls.2016.00043>
- 112 Mora Salguero, D. A. *et al.* Development of a *Chlamydomonas reinhardtii* metabolic network dynamic model to describe distinct phenotypes occurring at different CO₂ levels. *PeerJ* **6**, e5528 (2018). <https://doi.org/10.7717/peerj.5528>
- 113 Zuniga, C. *et al.* Predicting Dynamic Metabolic Demands in the Photosynthetic Eukaryote *Chlorella vulgaris*. *Plant Physiol* **176**, 450-462 (2018). <https://doi.org/10.1104/pp.17.00605>
- 114 Cunha, E., Sousa, V., Vicente, A., Geda, P. & Dias, O. Towards a genome-scale metabolic model of *Dunaliella salina*. *IFAC-PapersOnLine* **58**, 37-42 (2024). <https://doi.org/10.1016/j.ifacol.2024.10.007>
- 115 Sengupta, A. *et al.* A Novel Draft Genome-Scale Reconstruction Model of *Isochrysis* sp: Exploring Metabolic Pathways for Sustainable Aquaculture Innovations. *Microbiol. Biotechnol. Lett* **52**, 141-151 (2024). <https://doi.org/10.48022/mb.2309.09011>
- 116 Ray, A., Kundu, P. & Ghosh, A. Reconstruction of a Genome-Scale Metabolic Model of *Scenedesmus obliquus* and Its Application for Lipid Production under Three Trophic Modes. *ACS Synth Biol* **12**, 3463-3481 (2023). <https://doi.org/10.1021/acssynbio.3c00516>
- 117 Ye, C. *et al.* Reconstruction and analysis of the genome-scale metabolic model of *schizochytrium limacinum* SR21 for docosahexaenoic acid production. *BMC Genomics* **16**, 799 (2015). <https://doi.org/10.1186/s12864-015-2042-y>
- 118 Ahmad, A., Tiwari, A. & Srivastava, S. A Genome-Scale Metabolic Model of *Thalassiosira pseudonana* CCMP 1335 for a Systems-Level Understanding of Its Metabolism and Biotechnological Potential. *Microorganisms* **8** (2020). <https://doi.org/10.3390/microorganisms8091396>

- 119 Domenzain, I. *et al.* Reconstruction of a catalogue of genome-scale metabolic models with enzymatic constraints using GECKO 2.0. *Nature Communications* **13**, 3766 (2022). <https://doi.org/10.1038/s41467-022-31421-1>
- 120 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436-444 (2015). <https://doi.org/10.1038/nature14539>
- 121 Boer, M. D. *et al.* Improving genome-scale metabolic models of incomplete genomes with deep learning. *iScience* **27** (2024). <https://doi.org/10.1016/j.isci.2024.111349>
- 122 Sharma, N. K. & Rai, A. K. Biodiversity and biogeography of microalgae: progress and pitfalls. *Environmental Reviews* **19**, 1-15 (2010). <https://doi.org/10.1139/a10-020>
- 123 El-Sheekh, M. M., Alwaleed, E. A., Ibrahim, A. & Saber, H. Detrimental effect of UV-B radiation on growth, photosynthetic pigments, metabolites and ultrastructure of some cyanobacteria and freshwater chlorophyta. *International Journal of Radiation Biology* **97**, 265-275 (2021). <https://doi.org/10.1080/09553002.2021.1851060>
- 124 Al Jabri, H. *et al.* Cultivating Microalgae in Desert Conditions: Evaluation of the Effect of Light-Temperature Summer Conditions on the Growth and Metabolism of Nannochloropsis QU130. *Applied Sciences* **11** (2021).
- 125 Ikarán, Z., Suárez-Alvarez, S., Urreta, I. & Castañón, S. The effect of nitrogen limitation on the physiology and metabolism of chlorella vulgaris var L3. *Algal Research* **10**, 134-144 (2015). <https://doi.org/10.1016/j.algal.2015.04.023>
- 126 Solomon, B. D., Barnes, J. R. & Halvorsen, K. E. Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass and Bioenergy* **31**, 416-425 (2007). <https://doi.org/10.1016/j.biombioe.2007.01.023>
- 127 Chisti, Y. Constraints to commercialization of algal fuels. *Journal of Biotechnology* **167**, 201-214 (2013). <https://doi.org/10.1016/j.jbiotec.2013.07.020>
- 128 Ren, Y., Sun, H., Deng, J., Huang, J. & Chen, F. Carotenoid Production from Microalgae: Biosynthesis, Salinity Responses and Novel Biotechnologies. *Marine Drugs* **19** (2021).
- 129 Maoka, T. Carotenoids as natural functional pigments. *Journal of Natural Medicines* **74**, 1-16 (2020). <https://doi.org/10.1007/s11418-019-01364-x>
- 130 Kvangsakul, J. *et al.* Supplementation with the carotenoids lutein or zeaxanthin improves human visual performance. *Ophthalmic and Physiological Optics* **26**, 362-371 (2006). <https://doi.org/10.1111/j.1475-1313.2006.00387.x>
- 131 Xiao, Y. *et al.* Photosynthetic Accumulation of Lutein in Auxenochlorella protothecoides after Heterotrophic Growth. *Marine Drugs* **16** (2018).

- 132 Park, S.-H., Kyndt, J. A. & Brown, J. K. Comparison of *Auxenochlorella protothecoides* and *Chlorella* spp. Chloroplast Genomes: Evidence for Endosymbiosis and Horizontal Virus-like Gene Transfer. *Life* **12** (2022).
- 133 Capecchi, M. R. Altering the Genome by Homologous Recombination. *Science* **244**, 1288-1292 (1989). <https://doi.org/10.1126/science.2660260>
- 134 Craig, R. *et al.* in *Population Genetics Group 57 (PopGroup57)* 169 (2024).
- 135 He, M., Wen, J., Yin, Y. & Wang, P. Metabolic engineering of *Bacillus subtilis* based on genome-scale metabolic model to promote fengycin production. *3 Biotech* **11**, 448 (2021). <https://doi.org/10.1007/s13205-021-02990-7>
- 136 Orth, J. D., Thiele, I. & Palsson, B. O. What is flux balance analysis? *Nat Biotechnol* **28**, 245-248 (2010). <https://doi.org/10.1038/nbt.1614>
- 137 Mekanik, M., Fotovat, R., Motamedian, E. & Jafarian, V. Improvement of Lutein Production in *Auxenochlorella protothecoides* Using Its Genome-Scale Metabolic Model and a System-Oriented Approach. *Applied Biochemistry and Biotechnology* **195**, 889-904 (2023). <https://doi.org/10.1007/s12010-022-04186-y>
- 138 Li, C.-T. *et al.* Utilizing genome-scale models to optimize nutrient supply for sustained algal growth and lipid productivity. *npj Systems Biology and Applications* **5**, 33 (2019). <https://doi.org/10.1038/s41540-019-0110-7>
- 139 Duenas, M. A., Craig, R. J., Gallaher, S. D., Moseley, J. L. & Merchant, S. S. Leaky ribosomal scanning enables tunable translation of bicistronic ORFs in green algae. *bioRxiv*, 2024.2007.2024.605010 (2024). <https://doi.org/10.1101/2024.07.24.605010>
- 140 Iverson, S. J., Lang, S. L. C. & Cooper, M. H. Comparison of the bligh and dyer and folch methods for total lipid determination in a broad range of marine tissue. *Lipids* **36**, 1283-1287 (2001). <https://doi.org/10.1007/s11745-001-0843-0>
- 141 Scientific, T. Pierce BCA protein assay kit. *Pierce BCA* **449** (2013).
- 142 Palta, J. P. Leaf chlorophyll content. *Remote Sensing Reviews* **5**, 207-213 (1990). <https://doi.org/10.1080/02757259009532129>
- 143 Antoniewicz, M. R., Kelleher, J. K. & Stephanopoulos, G. Accurate Assessment of Amino Acid Mass Isotopomer Distributions for Metabolic Flux Analysis. *Analytical Chemistry* **79**, 7554-7559 (2007). <https://doi.org/10.1021/ac0708893>
- 144 Christie, W. W. Gas chromatography-mass spectrometry methods for structural analysis of fatty acids. *Lipids* **33**, 343-353 (1998). <https://doi.org/10.1007/s11745-998-0214-x>

- 145 Valle, O., Lien, T. & Knutsen, G. Fluorometric determination of DNA and RNA in Chlamydomonas using ethidium bromide. *Journal of Biochemical and Biophysical Methods* **4**, 271-277 (1981). [https://doi.org/10.1016/0165-022X\(81\)90067-1](https://doi.org/10.1016/0165-022X(81)90067-1)
- 146 Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRApy: COncstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology* **7**, 74 (2013). <https://doi.org/10.1186/1752-0509-7-74>
- 147 Cock, P. J. *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422-1423 (2009). <https://doi.org/10.1093/bioinformatics/btp163>
- 148 Bornstein, B. J., Keating, S. M., Jouraku, A. & Hucka, M. LibSBML: an API Library for SBML. *Bioinformatics* **24**, 880-881 (2008). <https://doi.org/10.1093/bioinformatics/btn051>
- 149 Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90-95 (2007). <https://doi.org/10.1109/MCSE.2007.55>
- 150 Banck, M., James, C., Morley, C., Vandermeersch, T. & Hutchison, G. Open Babel: An open chemical toolbox. *J. Cheminf* **3**, 33 (2011).
- 151 Fritzemeier, C. J., Hartleb, D., Szappanos, B., Papp, B. & Lercher, M. J. Erroneous energy-generating cycles in published genome scale metabolic networks: Identification and removal. *PLoS Computational Biology* **13**, e1005494 (2017). <https://doi.org/10.1371/journal.pcbi.1005494>
- 152 Lieven, C. *et al.* MEMOTE for standardized genome-scale metabolic model testing. *Nature Biotechnology* **38**, 272-276 (2020). <https://doi.org/10.1038/s41587-020-0446-y>
- 153 Sambamoorthy, G. & Raman, K. Understanding the evolution of functional redundancy in metabolic networks. *Bioinformatics* **34**, i981-i987 (2018). <https://doi.org/10.1093/bioinformatics/bty604>
- 154 Chen, H., Sosa, A. & Chen, F. Growth and Cell Size of Microalga *Auxenochlorella protothecoides* AS-1 under Different Trophic Modes. *Microorganisms* **12** (2024).
- 155 Darienko, T. & Pröschold, T. Genetic variability and taxonomic revision of the genus *Auxenochlorella* (Shihira et Krauss) Kalina et Puncocharova (Trebouxiophyceae, Chlorophyta). *Journal of Phycology* **51**, 394-400 (2015). <https://doi.org/doi.org/10.1111/jpy.12279>
- 156 Korozi, E. *et al.* Continuous Culture of *Auxenochlorella protothecoides* on Biodiesel Derived Glycerol under Mixotrophic and Heterotrophic Conditions: Growth Parameters and Biochemical Composition. *Microorganisms* **10** (2022).
- 157 Polat, E. & Altınbaş, M. Optimization of *Auxenochlorella protothecoides* lipid content using response surface methodology for biofuel production. *Biomass Conversion and Biorefinery* **12**, 2133-2147 (2022). <https://doi.org/10.1007/s13399-020-00798-8>

- 158 Çakır, Z. B., Yılmaz, H., Ertan, F., Tanriseven, A. & Özkan, M. Carrot pomace alone supports heterotrophic growth and lipid production of *Auxenochlorella protothecoides*. *Biomass Conversion and Biorefinery* **14**, 7315-7327 (2024). <https://doi.org/10.1007/s13399-022-02683-y>
- 159 Andeden, E. E., Ozturk, S. & Aslim, B. Effect of alkaline pH and nitrogen starvation on the triacylglycerol (TAG) content, growth, biochemical composition, and fatty acid profile of *Auxenochlorella protothecoides* KP7. *Journal of Applied Phycology* **33**, 211-225 (2021). <https://doi.org/10.1007/s10811-020-02311-0>
- 160 Markou, G. *et al.* Effects of Monochromatic Illumination with LEDs Lights on the Growth and Photosynthetic Performance of *Auxenochlorella protothecoides* in Photo- and Mixotrophic Conditions. *Plants* **10** (2021).
- 161 Xiao, Y. *et al.* Inhibition of glucose assimilation in *Auxenochlorella protothecoides* by light. *Biotechnology for Biofuels* **13**, 146 (2020). <https://doi.org/10.1186/s13068-020-01787-9>
- 162 Hörtensteiner, S., Chinner, J., Matile, P., Thomas, H. & Donnison, I. S. Chlorophyll breakdown in *Chlorella protothecoides*: characterization of degreening and cloning of degreening-related genes. *Plant Molecular Biology* **42**, 439-450 (2000). <https://doi.org/10.1023/A:1006380125438>
- 163 Rismani-Yazdi, H. *et al.* High-productivity lipid production using mixed trophic state cultivation of *Auxenochlorella* (*Chlorella*) *protothecoides*. *Bioprocess and Biosystems Engineering* **38**, 639-650 (2015). <https://doi.org/10.1007/s00449-014-1303-5>
- 164 Pydimalla, M., Husaini, S., Kadire, A. & Kumar Verma, R. Sustainable biodiesel: A comprehensive review on feedstock, production methods, applications, challenges and opportunities. *Materials Today: Proceedings* **92**, 458-464 (2023). <https://doi.org/10.1016/j.matpr.2023.03.593>
- 165 Krzemińska, I. & Oleszek, M. Glucose supplementation-induced changes in the *Auxenochlorella protothecoides* fatty acid composition suitable for biodiesel production. *Bioresource Technology* **218**, 1294-1297 (2016). <https://doi.org/10.1016/j.biortech.2016.07.104>
- 166 Stansell, G. R., Gray, V. M. & Sym, S. D. Microalgal fatty acid composition: implications for biodiesel quality. *Journal of Applied Phycology* **24**, 791-801 (2012). <https://doi.org/10.1007/s10811-011-9696-x>
- 167 Moser, B. R. Impact of fatty ester composition on low temperature properties of biodiesel-petroleum diesel blends. *Fuel* **115**, 500-506 (2014). <https://doi.org/10.1016/j.fuel.2013.07.075>
- 168 Stitt, M. & Zeeman, S. C. Starch turnover: pathways, regulation and role in growth. *Current Opinion in Plant Biology* **15**, 282-292 (2012). <https://doi.org/10.1016/j.pbi.2012.03.016>

- 169 Joun, J., Sirohi, R. & Sim, S. J. The effects of acetate and glucose on carbon fixation and carbon utilization in mixotrophy of *Haematococcus pluvialis*. *Bioresource Technology* **367**, 128218 (2023). <https://doi.org/10.1016/j.biortech.2022.128218>
- 170 Cao, Y. *et al.* Metabolomic exploration of the physiological regulatory mechanism of the growth and metabolism characteristics of *Chlorella vulgaris* under photoautotrophic, mixotrophic, and heterotrophic cultivation conditions. *Biomass and Bioenergy* **173**, 106775 (2023). <https://doi.org/10.1016/j.biombioe.2023.106775>
- 171 Ren, X., Chen, J., Deschênes, J.-S., Tremblay, R. & Jolicoeur, M. Glucose feeding recalibrates carbon flux distribution and favours lipid accumulation in *Chlorella protothecoides* through cell energetic management. *Algal Research* **14**, 83-91 (2016). <https://doi.org/10.1016/j.algal.2016.01.004>
- 172 Choi, K. R. *et al.* Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. *Trends in Biotechnology* **37**, 817-837 (2019). <https://doi.org/10.1016/j.tibtech.2019.01.003>
- 173 Young, J. D., Allen, D. K. & Morgan, J. A. in *Plant Metabolism: Methods and Protocols* (ed Ganesh Sriram) 85-108 (Humana Press, 2014).
- 174 McConnell, B. O. & Antoniewicz, M. R. Measuring the Composition and Stable-Isotope Labeling of Algal Biomass Carbohydrates via Gas Chromatography/Mass Spectrometry. *Analytical Chemistry* **88**, 4624-4628 (2016). <https://doi.org/10.1021/acs.analchem.6b00779>
- 175 Rezayian, M., Niknam, V. & Ebrahimzadeh, H. Oxidative damage and antioxidative system in algae. *Toxicology Reports* **6**, 1309-1313 (2019). <https://doi.org/10.1016/j.toxrep.2019.10.001>
- 176 Chung, C. H., Lin, D.-W., Eames, A. & Chandrasekaran, S. Next-Generation Genome-Scale Metabolic Modeling through Integration of Regulatory Mechanisms. *Metabolites* **11** (2021).
- 177 Ghoul, M. & Mitri, S. The Ecology and Evolution of Microbial Competition. *Trends in Microbiology* **24**, 833-845 (2016). <https://doi.org/10.1016/j.tim.2016.06.011>
- 178 Karatzos, S., van Dyk, J. S., McMillan, J. D. & Saddler, J. Drop-in biofuel production via conventional (lipid/fatty acid) and advanced (biomass) routes. Part I. *Biofuels, Bioproducts and Biorefining* **11**, 344-362 (2017). <https://doi.org/10.1002/bbb.1746>
- 179 Griffiths, M. J., van Hille, R. P. & Harrison, S. T. L. Lipid productivity, settling potential and fatty acid profile of 11 microalgal species grown under nitrogen replete and limited conditions. *Journal of Applied Phycology* **24**, 989-1001 (2012). <https://doi.org/10.1007/s10811-011-9723-y>
- 180 Figueroa, C. M., Asencion Diez, M. D., Ballicora, M. A. & Iglesias, A. A. Structure, function, and evolution of plant ADP-glucose pyrophosphorylase. *Plant Molecular Biology* **108**, 307-323 (2022). <https://doi.org/10.1007/s11103-021-01235-8>

- 181 Li, Y. *et al.* Chlamydomonas starchless mutant defective in ADP-glucose pyrophosphorylase hyper-accumulates triacylglycerol. *Metabolic Engineering* **12**, 387-391 (2010). <https://doi.org/10.1016/j.ymben.2010.02.002>
- 182 Zhang, R. *et al.* UDP-glucose pyrophosphorylase as a target for regulating carbon flux distribution and antioxidant capacity in Phaeodactylum tricornutum. *Communications Biology* **6**, 750 (2023). <https://doi.org/10.1038/s42003-023-05096-3>
- 183 Chen, D. *et al.* Overexpression of acetyl-CoA carboxylase increases fatty acid production in the green alga Chlamydomonas reinhardtii. *Biotechnology Letters* **41**, 1133-1145 (2019). <https://doi.org/10.1007/s10529-019-02715-0>
- 184 Zhang, Y., Pan, Y., Ding, W., Hu, H. & Liu, J. Lipid production is more than doubled by manipulating a diacylglycerol acyltransferase in algae. *GCB Bioenergy* **13**, 185-200 (2021). <https://doi.org/10.1111/gcbb.12771>
- 185 Li, D.-W. *et al.* Constitutive and Chloroplast Targeted Expression of Acetyl-CoA Carboxylase in Oleaginous Microalgae Elevates Fatty Acid Biosynthesis. *Marine Biotechnology* **20**, 566-572 (2018). <https://doi.org/10.1007/s10126-018-9841-5>
- 186 Craig, R. J. *et al.* Allodiploid hybridization, loss-of-heterozygosity and aneuploidy in the green alga Auxenochlorella, an emerging model for discovery research and bioengineering. *bioRxiv*, 2025.2002.2007.637104 (2025). <https://doi.org/10.1101/2025.02.07.637104>
- 187 Hammel, A. *et al.* Overexpression of Sedoheptulose-1,7-Bisphosphatase Enhances Photosynthesis in Chlamydomonas reinhardtii and Has No Effect on the Abundance of Other Calvin-Benson Cycle Enzymes. *Frontiers in Plant Science* **11** (2020). <https://doi.org/10.3389/fpls.2020.00868>

APPENDIX A

A.1 ¹³C MFA Fluxes

Table A.1 ¹³C MFA experimentally determined fluxes

Reaction	Equation	Flux	STD Error
R1	GLC.c -> G6P.c	2.5225	0.0935
R2 net	G6P.c <-> G6P.h	2.5225	0.0752
R3	G6P.c -> Ara.b + CO2.c	4.83E-06	0.0534
R4	G6P.c -> Gal.b	4.83E-06	0.0486
R5	G6P.h -> Sta.b	1	0.1279
R6 net	G6P.c <-> F6P.c	8.42E-07	0.042
R7 net	G6P.h <-> F6P.h	0.979	0.075
R8	G6P.h -> P5P.h + CO2.h	0.5435	0.0593
R9 net	P5P.h + P5P.h <-> S7P.h + T3P.h	0.1968	0.0478
R10 net	S7P.h + T3P.h <-> F6P.h + E4P.h	0.0099	0.0291
R11 net	P5P.h + E4P.h <-> F6P.h + T3P.h	-1.07E-07	0.0395
R12 net	F6P.c <-> T3P.c + T3P.c	-0.15	0.0704
R13 net	F6P.h <-> T3P.h + T3P.h	0.8389	0.0694
R14 net	T3P.c <-> T3P.h	-0.45	0.1969
R15 net	T3P.h <-> 3PG.h	1.6015	0.071
R16 net	3PG.h <-> 2PG.h	1.4015	0.067
R17 net	2PG.h <-> PEP.h	1.4015	0.067
R18	PEP.h -> PYR.h	1.158	0.0596
R19	PYR.h -> ACCoA.h + CO2.h	0.0235	0.0897
R20	PEP.h + CO2.h -> OAA.h	1.00E-07	0.0735

Table A.1 Continued

R21 net	OAA.h <-> MAL.h	1.00E-07	0.0735
R22 net	MAL.h <-> PYR.h + CO2.h	-0.6138	0.0749
R23	PEP.h + PEP.h + E4P.h -> Tyr.b + CO2.h	0.0217	0.0531
R24	PEP.h + PEP.h + E4P.h -> Phe.b + CO2.h	0.0251	0.0491
R25	PYR.h + PYR.h + ACCoA.h -> Leu.b + CO2.h + CO2.h	1.00E-07	0.0489
R26	3PG.h -> Ser.b	0.2	0.1198
R27	ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h -> Pal.b	1.00E-07	0.0941
R28	ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h -> Ster.b	1.00E-07	0.0497
R29	ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h -> Ole.b	1.00E-07	0.0439
R30	ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h + ACCoA.h -> Lino.b	1.00E-07	0.0461
R32 net	3PG.c <-> 2PG.c	-2.00E-07	0.0787
R33 net	2PG.c <-> PEP.c	-2.00E-07	0.0787
R34	PEP.c -> PYR.c	0.0205	0.0812
R35 net	PEP.c + CO2.c <-> OAA.c	-0.0205	0.0558
R36	PYR.c -> ACCoA.c + CO2.c	1.32E-06	0.0429
R37 net	OAA.c <-> MAL.c	-0.0205	0.1075
R38 net	MAL.c <-> MAL.h	-0.6138	0.0429
R39 net	MAL.c <-> PYR.c + CO2.c	-4.67E-08	0.0677

Table A.1 Continued

R40	PYR.c -> Ala.b	1.00E-07	5.43E-05
R42 net	PYR.c <-> PYR.h	-0.5207	0.0589
R43 net	PYR.c <-> PYR.m	0.5412	0.0707
R44	PYR.m -> LAC.m	1.13E-07	0.1522
R45 net	LAC.m <-> LAC.c	1.13E-07	0.1522
R46	LAC.c -> LAC.e	1.13E-07	0.1522
R47	SUCC.m -> SUCC.e	1.02E-07	0.161
R48	PYR.m -> ACCoA.m + CO2.m	1.9085	0.0435
R49	ACCoA.m + OAA.m -> CIT.m	1.1342	0.0549
R50 net	CIT.m <-> ICIT.m	1.1342	0.0549
R51	ICIT.m -> AKG.m + CO2.m	0.2283	0.0776
R52	AKG.m -> SUCC.m + CO2.m	0.2283	0.0488
R53	ICIT.m -> GLX.m + SUCC.m	0.9059	0.0928
R54	SUCC.m -> FUM.m	1.1342	0.1214
R55	FUM.m -> MAL.m	1.1342	0.1214
R56	ACCoA.m + GLX.m -> MAL.m	0.7743	0.0545
R57	MAL.m -> OAA.m	1.1342	0.0517
R58	MAL.m -> PYR.m + CO2.m	1.3675	0.052
R59 net	MAL.m <-> MAL.c	-0.5932	0.076
R60	AKG.m -> Glu.b	9.59E-08	0.0627
R61	GLX.m -> Gly.b	0.1317	0.1089
R62	PYR.m -> Ala.b	2.44E-04	0.0498
R63	PYR.m + PYR.m -> Val.b + CO2.m	1.00E-07	0.0493
R64 net	CO2.c <-> CO2.h	-1.76E-11	0.0709

Table A.1 Continued

R65 net	CO2.c <-> CO2.m	-3.7325	0.0626
R66	G6P.c -> G6P.b	4.83E-06	0.0532
R67	G6P.h -> G6P.b	4.82E-06	0.0487
R68	F6P.c -> F6P.b	0.15	0.0844
R69	F6P.h -> F6P.b	0.15	0.1155
R70	P5P.h -> P5P.b	0.15	0.1067
R71	E4P.h -> E4P.b	0.15	0.0933
R72	T3P.h -> T3P.b	6.15E-06	0.053
R73	T3P.c -> T3P.b	0.15	0.2481
R74	3PG.c -> 3PG.b	1.00E-07	0.0963
R75	3PG.h -> 3PG.b	2.57E-06	0.0457
R76	PEP.c -> PEP.b	9.97E-08	0.1258
R77	PEP.h -> PEP.b	0.15	0.163
R78	PYR.h -> PYR.b	1.55E-07	0.1294
R79	PYR.c -> PYR.b	4.61E-09	0.0488
R80	PYR.m -> PYR.b	1.00E-07	0.0493
R81	ACCoA.m -> ACCoA.b	9.99E-08	0.09
R82	ACCoA.h -> ACCoA.b	1.00E-07	0.0457
R83	ACCoA.c -> ACCoA.b	1.00E-07	0.067
R84	OAA.c -> OAA.b	1.00E-07	0.0456
R85	OAA.m -> OAA.b	9.99E-08	0.0748
R86	SUCC.m -> SUCC.b	1.00E-07	0.0458
R87	AKG.m -> AKG.b	9.97E-08	0.0457
R88	0*AKG.m -> AKG.s	0.0012	0.0476

Table A.1 Continued

R89	0*PYR.c -> PYR.s	3.65E-04	0.048
R90	0*PYR.m -> PYR.s	6.91E-04	0.0473
R91	0*PYR.h -> PYR.s	3.09E-04	0.0437
R92	0*MAL.h -> MAL.s	1.00E-07	0.0548
R93	0*MAL.m -> MAL.s	1.4103	0.0457
R94	0*MAL.c -> MAL.s	1.00E-07	0.0451
R95	0*CIT.m -> CIT.s	7.56E-04	0.0428
R96	0*LAC.m -> LAC.s	4.14E-04	0.0498
R97 net	T3P.h + E4P.h <-> S7P.h	-0.1868	0.0544
R98	CO2.c -> CO2.e	3.7531	0.1475
R99	3PG.c -> T3P.c	1.00E-07	0.0556
R100	OAA.c -> Asp.c	2.00E-07	0.1384
R101	Asp.c -> Asp.b	1.00E-07	0.1491
R102	Asp.c -> Thr.b	9.99E-08	0.0515
R103	P5P.h -> Ru15bp.h	1.00E-07	0.0286
R104	Ru15bp.h + CO2.h -> 3PG.h + 3PG.h	1.00E-07	0.0286
R105	ACCoA.h -> Ac.h	0.0235	0
R106	Ac.h -> Ac.h	8.95E-04	0.0472
R107	Ac.h -> EtOh.h	0.0235	0
R108	EtOh.h -> EtOh.e	0.0235	0
R109	ACCoA.c + OAA.c -> CIT.c	1.22E-06	0.0519
R110 net	CIT.c <-> ICIT.c	1.22E-06	0.0519
R111	ICIT.c -> AKG.c + CO2.m	1.22E-06	0.0519
R112 net	AKG.c <-> AKG.m	1.03E-06	0.0742

Table A.1 Continued

R113	AKG.c -> AKG.b	9.74E-08	0.1047
R114	AKG.c -> Glu.b	1.00E-07	0.0499

A.2 ¹³C Network

GLC.c (abcdef) -> G6P.c (abcdef)

G6P.c (abcdef) <-> G6P.h (abcdef)

G6P.c (abcdef) -> Ara.b (dfbce) + CO2.c (a)

G6P.c (abcdef) -> Gal.b (abcdef)

G6P.h (abcdef) <-> Sta.b (abcdef)

G6P.c (abcdef) <-> F6P.c (afbcde)

G6P.h (abcdef) <-> F6P.h (afbcde)

G6P.h (abcdef) -> P5P.h (bcdef) + CO2.h (a)

P5P.h (abcde) + P5P.h (fghij) <-> S7P.h (abfghij) + T3P.h (cde)

S7P.h (abcdefg) + T3P.h (hij) <-> F6P.h (abchij) + E4P.h (defg)

P5P.h (abcde) + E4P.h (fghi) <-> F6P.h (abfghi) + T3P.h (cde)

F6P.c (abcdef) <-> T3P.c (ebf) + T3P.c (dac)

F6P.h (abcdef) <-> T3P.h (ebf) + T3P.h (dac)

T3P.c (abc) <-> T3P.h (abc)

T3P.h (abc) <-> 3PG.h (bca)

3PG.h (abc) <-> 2PG.h (abc)

2PG.h (abc) <-> PEP.h (abc)

PEP.h (abc) -> PYR.h (abc)

PYR.h (abc) -> ACCoA.h (ab) + CO2.h (c)

PEP.h (abc) + CO2.h (d) -> OAA.h (abdc)

OAA.h (abcd) <-> MAL.h (abcd)

MAL.h (abcd) <-> PYR.h (abd) + CO2.h (c)

PEP.h (abc) + PEP.h (def) + E4P.h (ghij) -> Tyr.b (higfadjbc) + CO2.h (e)

PEP.h (abc) + PEP.h (def) + E4P.h (ghij) -> Phe.b (jgfhiadbc) + CO2.h (e)

PYR.h (abc) + PYR.h (def) + ACCoA.h (gh) -> Leu.b (adbech) + CO2.h (f) + CO2.h (g)

3PG.h (abc) -> Ser.b (abc)

ACCoA.h (ab) + ACCoA.h (cd) + ACCoA.h (ef) + ACCoA.h (gh) + ACCoA.h (ij) + ACCoA.h (kl) + ACCoA.h (mn) + ACCoA.h (op) -> Pal.b (abcdefghijklmnop)

ACCoA.h (ab) + ACCoA.h (cd) + ACCoA.h (ef) + ACCoA.h (gh) + ACCoA.h (ij) + ACCoA.h (kl) + ACCoA.h (mn) + ACCoA.h (op) + ACCoA.h (qr) -> Ster.b (abcdefghijklmnpqr)

ACCoA.h (ab) + ACCoA.h (cd) + ACCoA.h (ef) + ACCoA.h (gh) + ACCoA.h (ij) + ACCoA.h (kl) + ACCoA.h (mn) + ACCoA.h (op) + ACCoA.h (qr) -> Ole.b (abcdefghijklmnpqr)

ACCoA.h (ab) + ACCoA.h (cd) + ACCoA.h (ef) + ACCoA.h (gh) + ACCoA.h (ij) + ACCoA.h (kl) + ACCoA.h (mn) + ACCoA.h (op) + ACCoA.h (qr) -> Lino.b (abcdefghijklmnpqr)

3PG.h (abc) <-> 3PG.c (abc)

3PG.c (abc) <-> 2PG.c (abc)

2PG.c (abc) <-> PEP.c (abc)

PEP.c (abc) -> PYR.c (abc)

PEP.c (abc) + CO2.c (d) <-> OAA.c (abdc)

PYR.c (abc) -> ACCoA.c (ab) + CO2.c (c)

OAA.c (abcd) <-> MAL.c (abcd)

MAL.c (abcd) <-> MAL.h (abcd)

MAL.c (abcd) <-> PYR.c (abd) + CO2.c (c)

PYR.c (abc) -> Ala.b (abc)

2PG.h (abc) <-> 2PG.c (abc)

PYR.c (abc) <-> PYR.h (abc)

PYR.c (abc) <-> PYR.m (abc)

PYR.m (abc) -> LAC.m (abc)

LAC.m (abc) <-> LAC.c (abc)

LAC.c (abc) -> LAC.e (abc)

SUCC.m (abcd) -> SUCC.e (abcd)

PYR.m (abc) -> ACCoA.m (ab) + CO2.m (c)

ACCoA.m (ab) + OAA.m (cdef) -> CIT.m (caebfd)

CIT.m (abcdef) <-> ICIT.m (afcbed)

ICIT.m (abcdef) -> AKG.m (badcf) + CO2.m (e)

AKG.m (abcde) -> SUCC.m (badc) + CO2.m (e)

ICIT.m (abcdef) -> GLX.m (df) + SUCC.m (baec)

SUCC.m (abcd) -> FUM.m (abcd)

FUM.m (abcd) -> MAL.m (badc)

ACCoA.m (ab) + GLX.m (cd) -> MAL.m (acbd)

MAL.m (abcd) -> OAA.m (abcd)

MAL.m (abcd) -> PYR.m (abd) + CO2.m (c)

MAL.m (abcd) <-> MAL.c (abcd)
AKG.m (abcde) -> Glu.b (abcde)
GLX.m (ab) -> Gly.b (ab)
PYR.m (abc) -> Ala.b (abc)
PYR.m (abc) + PYR.m (def) -> Val.b (daebc) + CO2.m (f)
CO2.c (a) <-> CO2.h (a)
CO2.c (a) <-> CO2.m (a)
G6P.c (abcdef) -> G6P.b (abcdef)
G6P.h (abcdef) -> G6P.b (abcdef)
F6P.c (abcdef) -> F6P.b (abcdef)
F6P.h (abcdef) -> F6P.b (abcdef)
P5P.h (abcde) -> P5P.b (abcde)
E4P.h (abcd) -> E4P.b (abcd)
T3P.h (abc) -> T3P.b (abc)
T3P.c (abc) -> T3P.b (abc)
3PG.c (abc) -> 3PG.b (abc)
3PG.h (abc) -> 3PG.b (abc)
PEP.c (abc) -> PEP.b (abc)
PEP.h (abc) -> PEP.b (abc)
PYR.h (abc) -> PYR.b (abc)
PYR.c (abc) -> PYR.b (abc)
PYR.m (abc) -> PYR.b (abc)
ACCoA.m (ab) -> ACCoA.b (ab)
ACCoA.h (ab) -> ACCoA.b (ab)
ACCoA.c (ab) -> ACCoA.b (ab)
OAA.c (abcd) -> OAA.b (abcd)
OAA.m (abcd) -> OAA.b (abcd)
SUCC.m (abcd) -> SUCC.b (abcd)
AKG.m (abcde) -> AKG.b (abcde)
0*AKG.m (abcde) -> AKG.s (abcde)
0*PYR.c (abc) -> PYR.s (abc)
0*PYR.m (abc) -> PYR.s (abc)
0*PYR.h (abc) -> PYR.s (abc)

0*MAL.h (abcd) -> MAL.s (abcd)
0*MAL.m (abcd) -> MAL.s (abcd)
0*MAL.c (abcd) -> MAL.s (abcd)
0*CIT.m (abcdef) -> CIT.s (abcdef)
0*LAC.m (abc) -> LAC.s (abc)
T3P.h (abc) + E4P.h (defg) <-> S7P.h (becgafd)
CO2.c (a) -> CO2.e (a)
3PG.c (abc) -> T3P.c (cab)
OAA.c (abcd) -> Asp.c (abcd)
Asp.c (abcd) -> Asp.b (abcd)
Asp.c (abcd) -> Thr.b (cabd)
ACCoA.c (ab) -> Ac.c (ab)
Ac.c (ab) -> Ac.e (ab)
Ac.c (ab) -> EtOh.c (ab)
EtOh.c (ab) -> EtOh.e (ab)