ABSTRACT

Title of Document:	MAPPING METABOLIC FLUXES IN PLANT CELLS TO UNDERSTANT CARBON – NITROGEN INTERACTIONS AND NITROGEN STORAGE AND CYCLING
	Shilpa Nargund, Doctor of Philosophy, 2012
Directed By:	Ganesh Sriram, Assistant Professor Department of Chemical and Biomolecular Engineering

Plants provide commodities like food, fiber, fuel and chemicals. Understanding plant metabolism will help find genetic engineering targets that enhance production of these commodities. Interactions between the macronutrients – carbon (C) and nitrogen (N) determine growth and developmental functions in plants (Nunes-Nesi, Fernie, and Stitt 2010; Sakakibara, Takei, and Hirose 2006) and are regulated by complex mechanisms that need systems-level analyses. Metabolic fluxes, the rates of C flow in metabolic pathways, provide a system-wide view of metabolism and are quantified by steady state metabolic flux analysis (MFA) wherein isotopic tracers (^{13}C, ^{15}N) are fed to the cells and the resulting labeling patterns of biomass components are used to fit the fluxes. In this study we i) statistically designed isotope labeling experiments (ILEs) *in silico* to enhance accuracy of flux estimates through the pentose phosphate pathway (PPP) ii) conducted MFA on heterotrophic cell suspensions of *Arabidopsis thaliana* (Arabidopsis), a model plant, to investigate regulatory role of light in cell metabolism and iii) conducted MFA on cell

suspensions of poplar (*Populus tremula* x *Populus alba*; clone N 717-B4), a potential biofuel crop, to understand C-N interactions. *In silico* label design studies determined that accuracy of flux estimates in the PPP improves by ILEs with 1,2-^{13}C glucose and measuring labeling patterns of sugars, especially ribose. Metabolic fluxes, estimated by the designed ILEs on Arabidopsis cells, under continuous light or dark, showed negligible changes between treatments indicating that light does not regulate central carbon metabolism in heterotrophic Arabidopsis cells. The designed ILEs improved confidences of non-oxidative PPP flux estimates by 40-80% from previous studies (Masakapalli et al. 2009a). ILEs on poplar cell suspensions, grown in batch cultures, displayed unexpected back-mixing between unlabeled seed biomass and newly synthesized labeled biomass. Novel metabolic network models were developed that successfully account for observed back-mixing. ILEs on poplar cells, subjected to different C-N supply treatments to understand C-N interactions show significant differences in labeling patterns. Design of ILEs and subsequent improvement in flux estimates and the improvements in modeling metabolic networks are the novel contributions of this work.

MAPPING METABOLIC FLUXES IN PLANT CELLS TO UNDERSTAND CARBON – NITROGEN INTERACTIONS AND NITROGEN STORAGE AND CYCLING

By

Shilpa Nargund

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctorate of philosophy
2012

Advisory Committee:
Ganesh Sriram, Assistant Professor, Chemical and Biomolecular Engineering, Chair
Gary Coleman, Associate Professor, Plant Science and Landscape Architecture
Raymond Adomaitis, Professor, Chemical and Biomolecular Engineering
Srinivasa Raghavan, Professor, Chemical and Biomolecular Engineering
Heven Sze, Professor, Cell Biology and Molecular Genetics

Dedication

I dedicate my Ph.D. dissertation to Pushkar, the wind beneath my wings

I dedicate this work to my parents for their unrelenting faith, support and love

I dedicate this work to my friends Avani, Falguni, and Neil for being my best cheerleaders

Acknowledgements

I thank, first and foremost, Dr. Ganesh Sriram, my advisor, for his excellent guidance, encouragement, patience and creating an invigorating atmosphere for doing research.

I thank Dr. Gary Coleman for his support and guidance in my work on poplar. I wish to thank Dr. Vitali Tugarinov for lending his expertise on NMR data acquisition.

I thank my lab co-workers Ashish Misra, Xiaofeng Zhang, Ho-Man Yeung, Yuting Zheng, and present and past undergraduate researchers for their immense support and creating a lively atmosphere in our lab.

Finally I thank my committee members Dr. Srinivasa Raghavan, Dr. Heven Sze and Dr. Raymond Adomaitis for their careful critique of my research work.

Table of Contents

Dedication ii

Acknowledgements iii

Table of Contents iv

1. Chapter 1a: Introduction 1

Chapter 1b: Methods in metabolic flux analysis 3

1.1 Mathematical modeling of metabolism in plant cells 3

1.1.1 Cataloging reactions, their atom rearrangements and reversibilities 3

1.1.2 Modeling compartmentalization of reactions in subcellular organelles 3

1.1.3 Simplifying the model 4

1.1.4 Modeling the above features mathematically 4

1.1.5 Choosing free fluxes 6

1.2 Simulating an ILE in the metabolic network model 7

1.3 Statistical methods to estimate fluxes and their errors 10

1.4 Processing MS and NMR data 10

1.5 Software available for MFA 11

Outline of the work done in this study 12

2. Chapter 2: Designer labels for plant metabolism: Statistical design of isotope labeling experiments for improved quantification of flux in complex plant metabolic networks 14

Abstract 15

2.1. Introduction 16

2.2. Methods 25
2.2.1. Metabolic network models for the PPP and the GABA shunt 25
2.2.2. Simulations of ILEs by cumomer balancing 27
2.2.3. Isotopomer measurements simulated during ILE design 27
2.2.4. Statistical flux identifiability 29
2.3. Results and Discussion 31
2.3.1. Plant PPP fluxes are best identified with 100% 1,2-^{13}C Glc (with our set of isotopomer measurements) 31
2.3.2. *In silico* ILEs corroborate the superiority of 1,2-^{13}C Glc over 1-^{13}C Glc in identifying PPP fluxes 34
2.3.3. Is 1,2-^{13}C Glc always the best choice for the plant PPPs? 37
2.3.4. Superior performance of 1,2-^{13}C Glc is largely independent of PPP flux values 40
2.3.5. Labeling information contained in hexose and pentose sugars is critical in elucidating PPP compartmentation 42
2.3.6. MS outperforms NMR in identifying PPP fluxes 45
2.3.7. Performance of Glc labels in pairs and triads of ILEs in estimating PPP fluxes 46
2.3.8. Performance of labeled carbon sources for GABA shunt 51
2.4. Summary and outlook 52
Acknowledgments 55

3. Chapter 3: Isotope-assisted metabolic flux analysis on heterotrophic *Arabidopsis thaliana* under contrasting light treatments reveals negligible effect of light on central carbon metabolism 56

Abstract 57

3.1. Introduction 58

3.2. Results 61

3.2.1. Biomass accumulation and protein content differs between the two light treatments 61

3.2.2. Isotopic steady state attained between day 5 and 6 61

3.2.3. Serine and glycine labeling show that light-grown do not fix CO_2 photosynthetically but undergo serine hydroxymethyltransferase reaction 61

3.2.4. *In silico* design of isotope labeling experiments predicts that 1,2-^{13}C glucose estimates fluxes in pentose phosphate pathway the best 65

3.2.5. Metabolic network model 65

3.2.6. Difference in growth between the light treatments affects isotopomer abundances 71

3.2.7. Flux distributions in central carbon metabolism in the two light treatments are similar 71

3.2.8. Metabolic network model validation 74

3.3. Discussion 77

3.3.1. Central carbon metabolism fluxes minimally affected by light treatments because carbon signaling substitutes for light signaling 78

3.3.2. Limitations and criticisms 80

3.4. Conclusion 81
3.5. Materials and methods 82
3.5.1. Arabidopsis suspension cell cultures 82
3.5.2. Extraction and quantification of biomass components for estimation of biomass effluxes 82
3.5.2.1. Protein quantification 82
3.5.2.2. RNA quantification 83
3.5.2.3. Soluble metabolite quantification 83
3.5.2.4. Lipid quantification 83
3.5.2.5. Starch quantification 84
3.5.2.6. Cell wall quantification 84
3.5.3. Growth rates 84
3.5.4. Extraction and measurement of isotopomer abundances of biomass components by GC-MS 85
3.5.4.1. Amino acids and soluble metabolites 85
3.5.4.2. Glucose from soluble metabolites 86
3.5.4.3. Ribose from RNA 86
3.5.5. Extraction and measurement of isotopomer abundances of biomass components by NMR 86
3.6. Acknowledgements 87
4. Chapter 4: Elucidating carbon – nitrogen interactions in poplar suspension cells by metabolic flux analysis 88
4.1. Introduction 89

4.2. Results 90
4.2.1. Growth and nutrient uptake rates differ between the C-N supply treatments 91
4.2.2. The ^{13}C label is diluted by the initial seed biomass in batch cultures. 91
4.2.3. Seed biomass backmixes with the newly synthesized biomass 94
4.2.4. Modeling the backmixing of seed biomass with newly synthesized biomass 95
4.2.5. The backmixing model performs better than previously used models 97
4.2.6. ILEs carried out for three subculture cycles reduce the dilution of ^{13}C label by seed biomass 100
4.3. Discussion 101
4.4. Material and methods 102
4.4.1. Poplar suspension cell cultures 102
5. Chapter 5: Conclusions 103
Future directions 104
Bibliography 1066

1. Chapter 1a: Introduction

Plants are important sources of food, fiber, chemicals, pharmaceutical products and fuels. They are living chemical factories whose potential remains fairly untapped. There has been increasing interest in plants as biofuels and biomaterials. The desirable traits of a biofuel or biomaterial crop are i) high growth rate ii) low fertilizer utilization iii) high biomass per unit land area and iv) growth round the year. Not all traits are present in one plant species and it is desirable to genetically tailor potential plants so that they are more economical as biomaterial crops (Sticklen 2006). Understanding plant metabolism is key to the success of such endeavors.

Plant metabolism is regulated to a great extent by the availability of water, light, carbon and nitrogen (Nunes-Nesi, Fernie, and Stitt 2010). We are particularly interested in studying the carbon-nitrogen (C-N) interactions and nitrogen (N) storage and cycling in plant metabolism. C-N interactions are important because N is required for production of amino acids which are building blocks for proteins, nucleotides and several other metabolites all of which have C backbones. Thus uptake and assimilation of C and N occur mutually and drive the accumulation of biomass in plants. To improve crop biomass it is important to understand these interactions that are regulated at various levels. In this work, we study plant metabolism at the fluxomic level using the tool metabolic flux analysis.

Metabolic fluxes represent rates of carbon flow through metabolic pathways and are important indicators of cell physiology (Stephanopoulos and Stafford 2002a). While certain metabolic fluxes (e.g. fluxes leading to accumulation of biomass) can be

directly measured, others can only be measured by indirect techniques such as isotope-assisted metabolic flux analysis (isotope MFA). This technique involves conducting isotope labeling experiments (ILEs) wherein isotopically labeled (e.g. ^{13}C, ^{15}N, ^{17}O) substrates are fed to the plant cell culture or tissue and the isotopomer (*isotop*e is*omer*) abundances of intracellular metabolites such as proteins, lipids, sugars and nucleotides, are measured using nuclear magnetic resonance (NMR) or mass spectrometry (MS). For a given metabolite, the isotopomers formed depend on the atom rearrangements that occur during reactions involving the metabolite whereas the relative abundances of these isotopomers depend on the relative fluxes of the reactions. Since knowledge of atom rearrangements and reaction networks is fairly well documented, the fluxes can be estimated by fitting the measured isotopomer abundances (***Iexp***) to isotopomer abundances (***Isim)*** that are obtained by simulated the ILE on a mathematical model of the metabolic network. This chapter illustrates the steps involved in i) building a mathematical model of the metabolic network ii) simulating an ILE and iii) fitting ***Iexp*** to ***Isim***.

Chapter 1b: Methods in metabolic flux analysis

1.1 Mathematical modeling of metabolism in plant cells

A metabolic network is an interconnected web of several metabolic reactions. The following steps are required to build a model of the metabolic network

1.1.1 Cataloging reactions, their atom rearrangements and reversibilities

Freely available resources such as Kyoto Encyclopedia of Genes and Genomes (KEGG) (Masoudi-Nejad et al. 2008a), MetaCyc Encyclopedia of metabolic pathways (Zhang et al. 2005) and plant biochemistry textbooks provide information on reactions present in metabolic networks of particular organisms and their atom rearrangements. Reversibility of a reaction depends on the accompanying change in Gibbs free energy (ΔG). A negative ΔG implies that the reaction is irreversible. ΔG values can also be obtained from the above mentioned sources. The catalog should also include influx and efflux reactions such as uptake of nutrients and efflux of metabolites into sink components such as cell wall, protein, lipid, starch and nucleotides to model the flow of material in and out of the cell.

1.1.2 Modeling compartmentalization of reactions in subcellular organelles

Plant metabolic networks are unique in that they compartmentalize certain metabolic pathways within specific subcellular organelles such as the mitochondria, plastids, glyoxysome, etc with some degree of duplication between the organelles. For example, the glycolysis pathway and pentose phosphate pathway (PPP) are present

both in the cytosol and the plastid whereas the tricarboxylic acid cycle is primarily located in the mitochondrion. Compartmentalization information can be found in literature (Singh 1998a). Metabolite pools in duplicated pathways that are separated by organelle membranes may or may not be in isotopic equilibrium with each other depending on fluxes through the duplicated pathways and the rates of their transport across membranes. To model duplicated pathways in different compartments, include two pools of the common metabolites and transport reactions between the corresponding pools (see Sec. 4. for an example).

1.1.3 Simplifying the model

Each reaction is associated with a flux and of the thousands of reactions in plant metabolism; the model usually includes only those that participate in the primary carbon metabolism. This is because i) isotopomers of secondary metabolism are not easily measured ii) computational load increases with number of metabolites. Additionally, in steady state MFA, to simplify the model further, metabolites that do not undergo carbon atom rearrangements during sequential reactions are lumped together (Winden et al. 2001a).

1.1.4 Modeling the above features mathematically

Flux balance equations help mathematically model metabolic reactions and their stoichiometries. Flux balancing is based on the principle of conservation of mass. For a given metabolite,

$$v_{generation} - v_{consumption} = v_{accumulation}$$

Where $v_{generation}$: sum of fluxes that generate the metabolite,

$v_{consumption}$: sum of fluxes that consume the metabolite

$v_{accumulation}$: accumulation of the metabolite

At metabolic steady state there is no accumulation of intracellular metabolites in the network i.e.

$$v_{generation} - v_{consumption} = 0$$

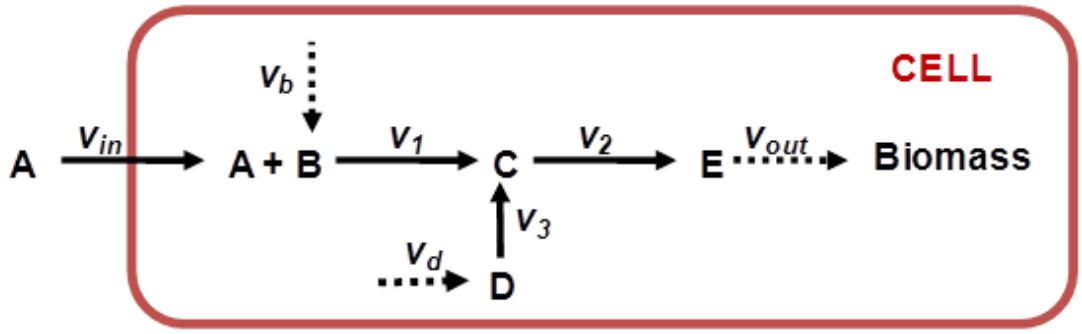


Fig. 1.1 Metabolic fluxes in a cell. Arrows represent fluxes and metabolites are represented by capital letters. The feed substrate A enters the cell via the flux v_{in} and E leaves to form biomass via the flux v_{out}. The dashed lines indicate that there are multiple reaction steps are required to generate these metabolites. The fluxes v_1, v_2 and v_3 are intracellular fluxes that can be estimated by fitting the measured data to a mathematical model of the metabolic network.

Flux balance equations for metabolites A, B, C, D and E in the example network (Fig. 1.1) are shown below

A: $v_{in} - v_1 = 0$;

B: $v_b - v_1 = 0$;

C: $v_1 + v_3 - v_2 = 0$

D: $v_d - v_3 = 0$

E: $v_2 - v_{out} = 0$

These equations can be represented in matrix format as follows

$$\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_{in} \\ v_b \\ v_d \\ v_1 \\ v_2 \\ v_3 \\ v_{out} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Further, the matrices are denoted as

$$\boldsymbol{S}.\boldsymbol{v} = 0$$

where $\boldsymbol{S}$ is the stoichiometric matrix and $\boldsymbol{v}$ is a vector containing the flux values. Measuring any 2 fluxes will satisfy the degree of freedom of this system of equations and thus enable us to calculate the other fluxes.

1.1.5 Choosing free fluxes

Oftentimes in larger metabolic networks, it is not possible to measure enough number of fluxes to satisfy the degree of freedom and hence to estimate these additional fluxes, isotopomer measurements (***Iexp***) are used. The fluxes estimated using ***Iexp*** are called "free fluxes". All other fluxes, called dependent fluxes, can be expressed as linear combinations of the free and measured fluxes.

The number of free fluxes *f* in a metabolic network with *m* metabolites and *n* net fluxes is given by $f = n - rank(\boldsymbol{S})$. Alternately $f = n - m - p$ where p is the number of measured fluxes. Several sets of free fluxes exist and they can be linearly transformed into each other. For small metabolic networks (like the example network) it is possible to figure out the free fluxes manually (by writing flux balances on paper) but the same is not true for large networks as the number of combinations increase. We can make use of the constraints placed on fluxes by the stoichiometric matrix to find a set of fluxes that are independent of each other. The stoichiometric matrix can be expressed as

$$\boldsymbol{S}_c.\boldsymbol{v}_c + \boldsymbol{S}_m.\boldsymbol{v}_m = 0,$$

where $\boldsymbol{v}_m$ and $\boldsymbol{v}_c$ are the measured and estimated fluxes respectively leading to

$$\boldsymbol{v_c = (S_c^{-1}).S_m.v_m}$$

One method of determining free fluxes in a metabolic network is by ensuring that the matrix $\boldsymbol{S_c}$ is invertible. There are other methods documented in literature for finding free fluxes as well (Quek et al. 2009).

1.2 Simulating an ILE in the metabolic network model

To estimate the free fluxes, the ***Iexp*** obtained by conducting an ILE is fitted to ***Isim*** which is obtained by simulating the ILE on the metabolic network model *in silico*. To simulate an ILE in the model, the paths of the isotopes through the reactions in the model need to be tracked. This can be done if the atom rearrangements during reactions are known. Isotopomer balances, similar to flux balances, help simulate ILEs. Fig. 1.2b illustrates isotopomer balances for the metabolite C in the example network in Fig. 1.1.

The isotopomers of the metabolites E and C are identical at steady state. Given guessed values of the free fluxes, the ***Isim*** of metabolite E can be computed. Thus the ***Iexp*** of E can be fitted to the ***Isim*** of E by varying the guess values of free flux within the stoichiometrically feasible flux space.

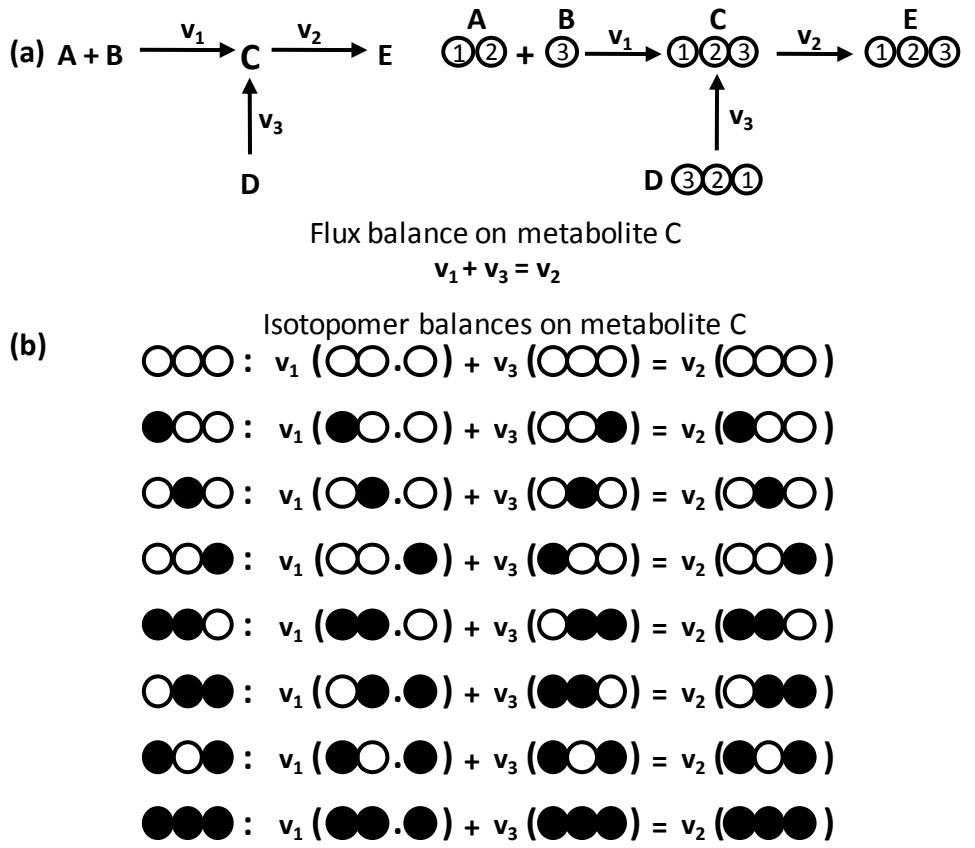


Fig. 1.2. Isotopomer balancing for simulating an ILE mathematically. The intracellular fluxes in the example network (Fig. 1) can be estimated by fitting experimentally measured isotopomer abundances to simulated abundances. The ILEs are simulated in the mathematical model by conducting isotopomer balancing. To balance the isotopomers of the metabolite C, the **(a)** flux balance equation and atom rearrangements of the reactions involving C are required. The number of carbon atoms in a metabolite are represented by the number of circles. The atom rearrangements that occur in the three reaction v_1, v_2 and v_3 are depicted with respect to the metabolite C. For instance the 1st carbon atom of C can come from the 1st carbon atom of A or 3rd carbon atom of D. **(b)** The isotopomers of the metabolite C are balanced in this example network. C is a three carbon atom metabolite and therefore has 2^3 isotopomers and 2^3 isotopomer balance equations. Empty circles represent ^{12}C whereas filled

circles represent ^{13}C. The isotopomer which is being balanced is listed at the left before its equation. It is clear that the isotopomer equations are nonlinear owing to the presence of the condensation reaction v_1.

A metabolite with n carbon atoms has 2^n isotopomers and correspondingly 2^n isotopomer balance equations. These equations can be non-linear if condensation reactions are involved. In Fig. 1.2b, the 1st term involving the flux v_1 renders the equations non-linear. The number of isotopomer balance equations can be reduced by lumping metabolites in the network and restricting the model to as few reactions as possible. To solve the non-linear equations analytically, several techniques such as cumomer, bondomer, elementary metabolite unit (EMU) and fluxomer balancing have been developed. The cumomer and EMU balancing techniques decompose the non-linear isotopomer equations to cascades of linear equations that can be solved analytically in sequence (Wiechert et al. 1997a; Antoniewicz, Kelleher, and Stephanopoulos 2007a). Both these techniques are equivalent in simulating ILEs and are most commonly used currently. The bondomer technique keeps track of intact and broken bonds to simulate the ILE using probability equations. This technique has primarily been used in ILEs that employ uniformly labeled feed substrates in order to avoid dealing with complex probability equations (van Winden, Heijnen, and Verheijen 2002; Sriram and Shanks 2004). Fluxomer balancing is the most recently developed technique and it accomplishes decoupling the non-linear terms by using a single variable for both the fluxes and isotopomers (Srour, Young, and Eldar 2011).

1.3 Statistical methods to estimate fluxes and their errors

Optimization algorithms are used to intelligently search the flux space iteratively and find the points that best explain the experimental observations i.e. the isotopomer abundances. The optimizations are repeated several times starting at different points in the flux space to ensure that the global optimum is reached. To ensure statistical significance, the errors in experimental measurements are projected onto flux estimates via techniques such as Monte Carlo simulations or other non-linear approaches (Antoniewicz, Kelleher, and Stephanopoulos 2006a; Möllney et al. 1999a).

1.4 Processing MS and NMR data

Raw MS and NMR data need to be processed in order to obtain isotopomer abundances. MS data has to be corrected for the presence of naturally abundant isotopes of all elements in the metabolites and the areas under the NMR peaks need to be determined to obtain isotopomer abundances. MS and NMR are capable of measuring only linear combinations of abundances of isotopomers of a metabolite (Fig. 1.3).

It is sometimes possible to obtain all 2^n isotopomer abundances for certain metabolites. For instance, three distinct fragments of the three carbon atom amino acid serine (when derivatized by TBDMS) can be measured by MS and from these all 8 isotopomers can be computed. These linear combinations of isotopomer abundances can easily be simulated during simulation of an ILE.

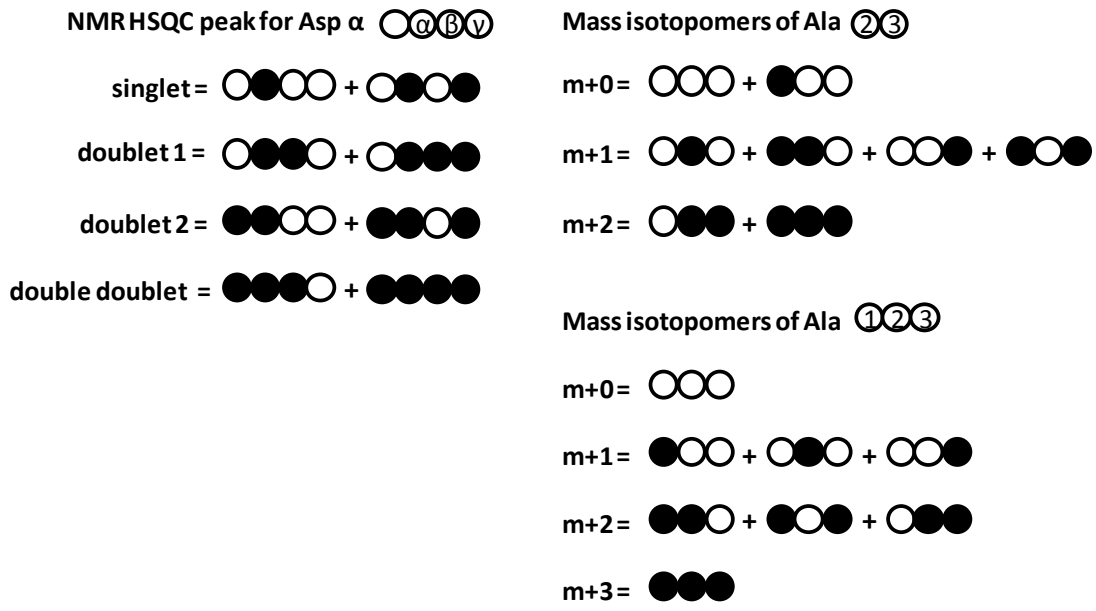


Fig. 1.3. Linear combinations of isotopomer abundances are measured by NMR and MS. Filled circles represent ^{13}C atoms and empty ^{12}C **(a)** The 2D-NMR peak corresponding to Asp α carbon atom can split into 4 peaklets, the relative areas under which represent the relative abundances of the linear combinations of isotopomers illustrated. The singlet is obtained when the α carbon atom is ^{13}C and its immediate neighbors are ^{12}C. Doublet 1 and doublet 2 are obtained when the α carbon atom and one of its immediate neighbors are ^{13}C. The double doublet is obtained when the α carbon and both its neighbors are ^{13}C. The isotope label on the γ position does not affect the peak splitting. **(b)** Mass spectrometry fragments metabolites and measures their masses. If m is the molecular weight of the fragment with n carbon atoms then its mass can range from m+0 to m+n depending on the isotope labeling of the carbon atoms. These mass isotopomers represent linear combinations of positional isotopomers.

1.5 Software available for MFA

It is imperative to automate parts or all of the above mathematical techniques using

programming languages to obtain accurate results in a reasonable timeframe. Various software have been developed over the years that can simulate ILEs on mathematical models of metabolic networks, estimate fluxes using optimization algorithms and find flux confidence intervals using statistical methods. Depending on the capabilities of the software, additional computational tools maybe required to determine free fluxes and process MS or NMR data. The various software viz. 13CFLUX2, FiatFlux, NMR2Flux, Metran, and OpenFLUX differ mainly in their ILE simulation technique, optimization algorithm and statistical method used for estimating flux standard deviations.

NMR2Flux, developed by Sriram et al. (Sriram et al. 2004a), is capable of simulating ILEs using cumomer or bondomer balancing. This is the software used in this work. NMR2Flux is in C language and is currently not available publically. The user needs to input the metabolic reactions and their atom rearrangements in a comma separated value (csv) file that can be read by the software. The user also needs to determine and specify the free fluxes. Corrected MS or NMR data is read from csv files. The software is capable of conducting sensitivity analysis thus allowing design of experiments. It uses the optimization algorithm simulated annealing to search the flux space and Monte Carlo simulations to compute standard deviations of fluxes.

Outline of the work done in this study

Chapter 2 describes the statistical design study conducted to identify the best isotopic labels and information rich measurements to accurately identify fluxes in two important plant pathways. This chapter is an accepted manuscript in the journal –

Molecular Biosystems.

Chapter 3 describes the MFA study on Arabidopsis cell suspensions subjected to two contrasting light treatments. This chapter is the draft of a manuscript that will be submitted for publication to the journal – *Plant physiology*.

Chapter 4 describes the MFA study on poplar cell suspensions grown under different carbon – nitrogen supply treatments. This chapter is an early draft of two manuscripts that will be submitted to the journals – *Metabolic Engineering* and *Plant Physiology*. This work was mainly done by myself, Dr. Ashish Misra and Mr. Xiaofeng Zhang and is part of a broader study on poplar via collaboration with Dr. Gary Coleman at the Plant science and landscape architecture department.

2. Chapter 2: Designer labels for plant metabolism: Statistical design of isotope labeling experiments for improved quantification of flux in complex plant metabolic networks

Shilpa Nargund[1] and Ganesh Sriram[1]*

[1]Department of Chemical and Biomolecular Engineering, University of Maryland, College Park, MD 20742, USA

Molecular Biosystems, Manuscript in press. DOI: 10.1039/C2MB25253H

Author contributions

SN and GS conceived this study and jointly made improvements to the computer program NMR2Flux+ to enable the necessary computation. SN performed the computation, wrote substantial portions of the manuscript and drafted a revised version after review. GS wrote and critically edited the manuscript. Both authors read and approved the final version of the manuscript prior to submission.

Abstract

Metabolic fluxes are powerful indicators of cell physiology and can be estimated by isotope-assisted metabolic flux analysis (MFA). The complexity of the compartmented metabolic networks of plants has constrained the application of isotope-assisted MFA to them, principally because of poor identifiability of fluxes from the measured isotope labeling patterns. However, flux identifiability can be significantly improved by a priori design of isotope labeling experiments (ILEs). This computational design involves evaluating the effect of different isotope label and isotopomer measurement combinations on flux identifiability, and thereby identifying optimal labels and measurements toward evaluating the fluxes of interest with the highest confidence. This article reports ILE designs for two major, compartmented plant metabolic pathways – the pentose phosphate pathway (PPP) and γ-aminobutyric (GABA) shunt. Together, these pathways represent common motifs in plant metabolism including duplication of pathways in different subcellular compartments, reversible reactions and cyclic carbon flow. To compare various ILE designs, we employed statistical A- and D- optimality criteria. Our computations showed that 1,2-^{13}C Glc is a powerful and robust label for the plant PPPs, given currently popular isotopomer measurement techniques (single quadrupole mass spectrometry [MS] and 2-D nuclear magnetic resonance [NMR]). Further analysis revealed that this label can estimate several PPP fluxes better than the popular label 1-^{13}C Glc. Furthermore, the concurrent measurement of the isotopomers of hexose and pentose moieties synthesized exclusively in the cytosol or the plastid compartments (measurable

through intracellular glucose or sucrose, starch, RNA ribose and histidine) considerably improves the identifiability of PPP fluxes in the individual compartments. Additionally, MS-derived isotopomer measurements outperform NMR-derived measurements in identifying PPP fluxes. The potency of 1,2-^{13}C Glc can be improved substantially by combining it with other labels (e.g. 3-^{13}C Glc, 1-^{13}C Glc and U-^{13}C Glc) in parallel ILEs. For the GABA shunt, we calculated that 100% 2-^{13}C Ala and 100% U-^{13}C Gln constitute the best labels. We anticipate that the ILE designs presented in this article can enhance the quality of flux estimates in these two complex plant pathways. In the future, these ILE designs can be further improved by leveraging recent analytical and computational developments in isotope-assisted MFA.

2.1. Introduction

Metabolic flux analysis (MFA), a powerful technique to quantify cellular physiology (Stephanopoulos and Stafford 2002b), involves the systemwide quantification of carbon traffic through cellular biochemical reactions. Metabolic fluxes represent a substantial portion of the "action" occurring in a cell or tissue (Stephanopoulos 2002). Therefore, they are as important as other indicators of phenotype such as transcript levels, protein levels and enzyme activities (L. J. Sweetlove, Last, and Fernie 2003). Metabolic flux maps are instrumental toward a comprehensive understanding of metabolism (L. Sweetlove, Fell, and Fernie 2008; L. J. Sweetlove, Last, and Fernie 2003; Schwender, Ohlrogge, and Shachar-Hill 2003a; Schwender, Ohlrogge, and Shachar-Hill 2004; Fernie, Geigenberger, and Stitt 2005; Schwender 2008). By

enabling visualization of carbon traffic in metabolic pathways (Sriram et al. 2004b; Sriram, Fulton, and Shanks 2007; Rontein et al. 2002; Schwender, Shachar-Hill, and Ohlrogge 2006; Libourel and Shachar-Hill 2008; Baxter et al. 2007; Alonso et al. 2007; Williams et al. 2008a; Allen, Libourel, and Shachar-Hill 2009; Allen, Ohlrogge, and Shachar-Hill 2009), these maps can suggest strategic metabolic engineering targets (Schwender 2008; Stephanopoulos 1999), identify unknown (Schwender et al. 1996; Schwender et al. 2004) or apparently futile metabolic pathways(Sriram et al. 2008) and potentially contribute toward building predictive models of metabolism (L. J. Sweetlove, Last, and Fernie 2003; Minorsky 2003; Ratcliffe and Shachar-Hill 2006; Rios-Estepa and Lange 2007).

Isotope-assisted MFA is a powerful method of quantifying fluxes, especially in sophisticated metabolic networks such as those of plants. In this method the biological system of interest is fed a designed mixture of labeled (e.g. ^{13}C) and unlabeled (e.g. ^{12}C) stable isotopes in an isotope labeling experiment (ILE). Fluxes are iteratively evaluated from the ensuing isotopic labeling patterns (such as isotope isomers [isotopomers]) of metabolites and biomass components by computational techniques such as metabolic network modeling, isotopomer balancing, and global optimization (Stephanopoulos and Stafford 2002b; Libourel and Shachar-Hill 2008; Wiechert 2001). This task is nontrivial because flux evaluation is a challenging parameter estimation problem in an extensive parameter space (Antoniewicz, Kelleher, and Stephanopoulos 2006b), wherein the fluxes are parameters that have to be estimated from isotopomer abundances and other measurements.

Flux identifiability, the confidence with which a flux can be estimated from the

information contained in isotope labeling patterns, is a valuable measure of the quality of information obtained from an ILE. Flux identifiability depends on properties of the ILE, and this dependence can be examined on two levels. On one level (structural identifiability), flux identifiability depends on **(a)** the layout or topology (stoichiometry and carbon atom rearrangements) of the metabolic network under investigation, **(b)** which labeled carbon sources are supplied and which atoms of these carbon sources are labeled and **(c)** which metabolites are analyzed and which isotopomers of these metabolites are measured. On another level (statistical identifiability), flux identifiability also depends on **(d)** the values of the fluxes in the network, **(e)** the relative proportions of the supplied labeled carbon sources and **(f)** the measurement errors of the labeling patterns(Möllney et al. 1999b). Whereas the metabolic network layout and flux values (**a** and **d**) are beyond the control of the investigator, all the other factors (**b**, **c**, **e**, **f**) can be chosen judiciously to enhance flux identifiability. However, doing so is a difficult task because the optimal choices are usually not obvious and have to be determined by sophisticated mathematical procedures that compare different ILE designs on the basis of their statistical quality. Insufficient flux identifiability is an acute problem in plant metabolic networks due to their complexity, which arises due to the duplication of pathways in multiple intracellular compartments with different fluxes in each compartment, the existence of many bypasses and cyclic pathways, myriad interconnections between metabolic subnetworks and incompletely known biochemistries.

This work is motivated by the success of previous investigations on microorganisms as well as *in silico* work on mammalian cell and plant embryo metabolism (Möllney

et al. 1999b; Libourel, Gehan, and Shachar-Hill 2007a; Araúzo-Bravo and Shimizu 2003; Metallo, Walther, and Stephanopoulos 2009; Crown, Ahn, and Antoniewicz 2012; Walther et al. 2012), which showed that the *a priori* design of ILEs can lead to a multifold increase in the information obtained from the ILEs (Williams et al. 2008a; Möllney et al. 1999b; Libourel, Gehan, and Shachar-Hill 2007a; Wiechert et al. 2001). In this article, we report the *a priori* design of ILEs for quantifying fluxes in two important plant metabolic pathways: the pentose phosphate pathway (PPP) and the γ-aminobutyric acid (GABA) shunt. The PPP (Fig. 2.1a) is a central pathway in plant metabolism and carries substantial carbon flux (Kruger and von Schaewen 2003a). Although the PPP wastes a sixth of the carbon that passes through it as CO_2 (Heldt 2004), it regenerates the reductant NADPH (which is necessary for synthesizing fatty acids and certain amino acids) and provides carbon skeletons for the synthesis of nucleotides, phenylpropanoids, lignin and some amino acids (Kruger and von Schaewen 2003a).

Furthermore, the complexity of the PPPs also raises several open questions: (i) whether the compartmentation of the PPPs is standard across plants, (ii) how the distribution of carbon traffic in the PPPs varies with environmental conditions such as light, temperature and nitrogen source availability, and (iii) how plants regulate carbon traffic through the PPPs at the level of gene expression. These questions necessitate the development of a tool to precisely quantify carbon traffic in the PPPs.

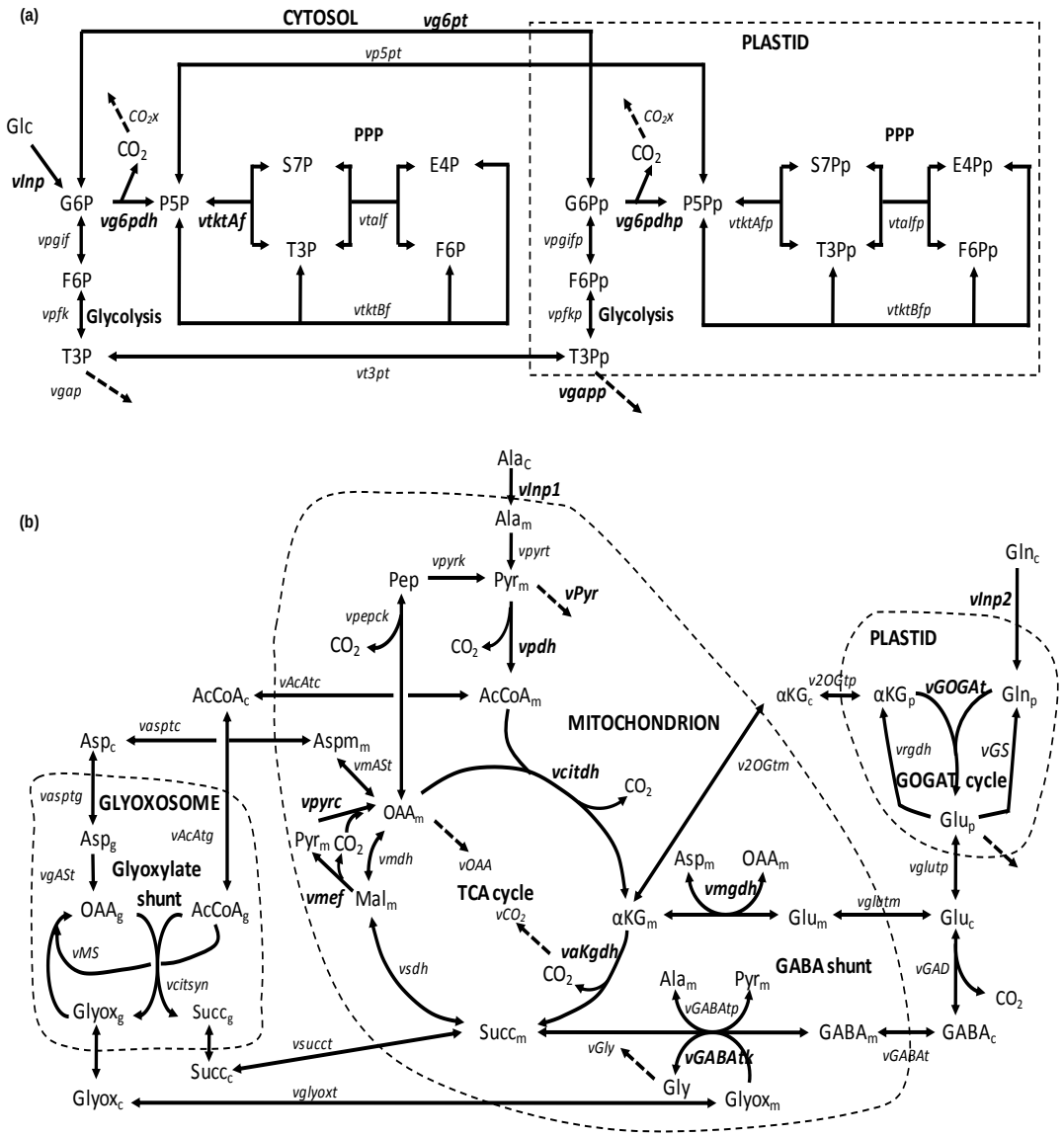


Fig. 2.1 (a) PPP metabolic model. Glycolysis and PPP duplicated in cytosol and plastid contain 13 metabolites and 32 fluxes (19 net fluxes and 13 reversibility extents). Flux balances allow dependent fluxes to be expressed as linear combinations of directly measurable (*vInp and vgapp*) and free fluxes (*vg6pdh, vtktAf, vg6pdhp* and *vg6pt).* Dashed arrows indicate metabolites leaving the system. The plastidic compartment is enclosed by a dashed box; suffix 'p' indicates metabolites in the plastidic compartment. Abbreviations: E4P, erythrose-4-phophate; F6P, fructose-6-phosphate; Glc, glucose; G6P, glucose-6-

phosphate; P5P, pentose phosphates; S7P, sedoheptulose-7-phosphate; T3P, triose phosphates. **(b) GABA shunt model.** The TCA cycle in the mitochondrion, the glyoxylate shunt in the glyoxysome, the GOGAT cycle in the plastid and the GABA shunt across the mitochondrion and the cytosol together comprise a network with 31 metabolites and 34 net fluxes. The measured fluxes are *vInp1, vInp2* and *vPyr;* the free fluxes *are vPdh, vCitdh, vaKgdh, vMgdh, vMef, vPyrc, vGABAtk* and *vGogat*. The intracellular compartments are separated from each by dash lined regions. Dashed arrows indicate metabolites leaving the system. Suffixes 'm', 'p', 'c' and 'g' indicate metabolites in the mitochondrial, plastidic, cytosolic and glyoxysome compartments respectively. Abbreviations: AcCoA, acetyl CoA; Ala, alanine; Asp, aspartate; GABA, γ-aminobutyric acid; Gln, glutamine; Gly, glycine; Glu, glutamate; Glyox, glyoxylate; αKG, alpha-ketoglutarate; Mal, malate; OAA, oxaloacetate; Pep, phosphoenolpyruvate; Pyr, pyruvate; Succ, succinate.

Although previous investigations have addressed these questions to an extent, there are gaps in knowledge and a clear picture of PPP flux distribution across compartments is lacking. In the pioneering MFA work of Dieuaide-Noubhani et al. (Dieuaide-Noubhani et al. 1995), the metabolic redistribution of 1-^{14}C and 2-^{14}C glucose (Glc) into sucrose, free Glc and starch in maize root tip cells suggested that the PPP was mostly active in the plastid. Subsequent work by Shachar-Hill and co-workers (Paula Alonso, Dale, and Shachar-Hill 2010; Schwender, Ohlrogge, and Shachar-Hill 2003a; Alonso, Val, and Shachar-Hill 2011) as well as Schwender and co-workers on *Brassica napus* embryos, maize embryos and maize endosperm used models in which the oxidative branch of PPP (from glucose-6-phosphate [G6P] to the pentose phosphates) was present in both the cytosol and the plastid, whereas the non-

oxidative branch (the rest of the PPP) was present only in the plastid. However, other evidence points to the possibility that the PPP operates in both compartments. For example, Krook et al. (Krook et al. 1998) fed 1-^{13}C Glc to carrot suspension cells and compared the labeling in sucrose (synthesized from cytosolic hexose phosphates (Usuda and Edwards 1980)) and starch (synthesized from plastidic hexose phosphates (Streb et al. 2009)). From this, they qualitatively inferred that the PPP is present in both the cytosol and the plastid. Sriram et al. (Sriram et al. 2004b), in their work on soybean embryos, observed substantially different isotopomer abundances in hydrolysis products of sugars associated with glycosylated protein (derived from cytosolic hexose phosphates) and starch, and fitted their labeling data to a model that contained the oxidative and non-oxidative PPPs in both the cytosol and the plastid. Allen and co-workers conducted ILEs on soybean embryos and analyzed the labeling patterns of amino acids belonging to the large and small subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase, which are synthesized in the plastid and cytosol respectively. This novel method revealed isotopic differences between amino acids originating in the cytosol and the plastid and also found evidence of flux through the PPP. Together, these results suggest that the allocation of flux to the PPPs in the different compartments may vary(Allen et al. 2012). In a recent, elaborate study, Masakapalli et al. (Masakapalli et al. 2009b) fed 1-^{13}C, 2-^{13}C and U-^{13}C Glc to *Arabidopsis thaliana* suspension cells and examined three metabolic models that could best account for the concomitant labeling patterns. These three models differed in the subcellular localization of the PPP – one model contained the oxidative and non-oxidative reactions of PPP only in the plastid, another model contained the

oxidative reactions in cytosol and plastid with the non-oxidative reactions restricted to the plastid and the third model contained the oxidative and non-oxidative reactions of PPP in both the plastid and the cytosol. Surprisingly, Masakapalli et al. observed that all three models explained the data equally well. This illustrates the difficulty in identifying PPP fluxes accurately in plant metabolic networks. Our study aims to address these issues by determining the best isotopic labels and critical isotopomer measurements that can help obtain additional information that will help quantify the PPP fluxes more accurately. Although elaborate and comprehensive MFA studies on PPP in plants exist (Sriram et al. 2004b; Sriram, Fulton, and Shanks 2007; Masakapalli et al. 2009b; Roscher, Kruger, and Ratcliffe 2000; Alonso, Val, and Shachar-Hill 2011; Allen, Ohlrogge, and Shachar-Hill 2009), there has been little focus on the design of ILEs involving non-trivial isotope labels, especially for the PPP. We anticipate that the PPP flux estimates of previous studies can be significantly improved by employing the isotope labeling strategies proposed in this article.

The GABA shunt (Fig. 2.1b) is a highly interconnected pathway that acts as a crosslink between carbon and nitrogen metabolism (Fait et al. 2008). This pathway involves the conversion of glutamate to succinate via the non-protein amino acid GABA instead of via the tricarboxylic acid (TCA) cycle; therefore, the GABA shunt is a bypass of the TCA cycle. Although GABA is known to play various crucial roles in plants (integration of carbon and nitrogen metabolism (Bouche and Fromm 2004), defense against insect attack (Bown, Hall, and MacGregor 2002) and pollen tube development (Palanivelu et al. 2003)) and animals (neurotransmitter (Fait et al. 2008;

Shelp, Bown, and Faure 2006; Shelp, Bown, and McLean 1999)) little is known about carbon flow through the GABA shunt (Fait et al. 2008) relative to that through the TCA cycle. Researchers have hypothesized that the GABA shunt is a metabolic highway that carries significant carbon flux during normal conditions and even greater flux when a plant faces stress (Fait et al. 2008). If this hypothesis is true, then the GABA shunt is one of the first major pathways taken by nitrogen after it enters primary metabolism. This article also reports the design of judicious combinations of labeled carbon sources fed in ILEs that will help test this hypothesis through MFA. Previous isotope-assisted MFA studies of plant metabolic networks (Sriram et al. 2004b; Sriram, Fulton, and Shanks 2007; Schwender, Shachar-Hill, and Ohlrogge 2006; Alonso et al. 2007; Allen, Ohlrogge, and Shachar-Hill 2009; Iyer et al. 2008) were designed to investigate most central carbon metabolic pathways but did not focus on the GABA shunt.

In the article we identify isotope labels and label combinations that improve the identifiabilities of important fluxes in PPP and GABA shunt pathways. We also identify the biomass components that contribute maximum labeling information toward flux identifiability. Additionally, we compare the usefulness of labeling information obtained from the two commonly used isotopomer measurement techniques – mass spectrometry (MS) and nuclear magnetic resonance (NMR).

2.2. Methods

2.2.1. Metabolic network models for the PPP and the GABA shunt

We modeled metabolic networks by using steady state flux balance equations of the form:

$$\mathbf{S.v} = \mathbf{0}$$

where **v** is a vector containing all fluxes and **S** is a stoichiometric matrix that represents metabolite balances in terms of the fluxes. An outcome of this relationship is that several fluxes in the network ("dependent" fluxes) are expressible as linear combinations of a smaller set of parameters, which includes: (i) a set of fluxes termed "free" fluxes(Wiechert and Graaf 1997), (ii) the few fluxes that are directly measureable (e.g. carbon source uptake), (iii) reversibility extents, relevant to pairs of reversible reactions and (iv) "scrambling extents" which, for pairs of reactions that have identical stoichiometries but different carbon atom rearrangements, are ratios indicating how the net flux is split across the two carbon atom rearrangements.

Our model of the PPP (Fig. 2.1a) is based on reaction stoichiometries and carbon atom rearrangements from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Masoudi-Nejad et al. 2008b) as well as previous studies of the PPP (Kruger and von Schaewen 2003a). This model comprises glycolysis and the PPP, each duplicated in the cytosol and the plastid compartments. The sole carbon source in the model is Glc (taken up through the flux *vInp*). Carbon exits the network either as triose phosphates from the plastid (flux *vgapp*) or as CO_2 (flux vco_2x). We lumped the three carbon atom-metabolites dihydroxyacetone phosphate, glyceraldehyde-3-phosphate,

phosphoenolpyruvate and pyruvate into a single metabolite named triose-3-phosphates (T3P) (Winden et al. 2001b); we also lumped the five-carbon atom metabolites ribose 5-phosphate, ribulose-5-phosphate and xylulose 5-phoshpate into a single metabolite named pentose-5-phosphates (P5P). The absence of carbon atom rearrangements between these metabolites and their relatively rapid equilibration (Winden et al. 2001b) justifies this lumping. We modeled the intercompartmental transport of metabolites, such as the reversible transport of G6P, P5P and T3P between the cytosol to the plastid, as bidirectional fluxes. Overall, the metabolic model contains 13 metabolites and 32 fluxes. Of these, two fluxes (*vInp* and *vgapp*) are measurable, and four (*vg6pdh, vtktAf, vg6pdhp, vg6pt*) are free fluxes. Additionally the model contains 13 reversibility extents. Our model of the GABA shunt (Fig. 2.1b) is also based on reaction stoichiometries and carbon atom rearrangements from KEGG and on previous studies on this pathway(Fait et al. 2008; Clark et al. 2009). The model comprises the TCA cycle in the mitochondrion, the glutamine (Gln)-α-ketoglutarate aminotransferase (GOGAT) cycle in the plastid, the glyoxylate shunt in the glyoxysome and the GABA shunt that spans the mitochondrion and the cytosol. We lumped the metabolites fumarate and malate into a single pool (Winden et al. 2001b) and modeled the intercompartmental transport of metabolites as bidirectional fluxes. The network has 31 metabolites and 65 fluxes of which three fluxes (*vInp1*, *vInp2* and *vPyr*) are measurable; eight (*vpdh, vcitdh, vpyrc, vmef, vGOGAt, vGABAtk, vmgdh, vaKgdh*) are free fluxes. Additionally the model contains 22 reversibility extents and one scrambling extent. The only carbon sources in the model are Ala and Gln (taken up through the fluxes *vInp1* and *vInp2*

respectively). The metabolites pyruvate, CO_2, plastidic glutamate, mitochondrial oxaloacetate and mitochondrial glycine exit the metabolic network.

2.2.2. Simulations of ILEs by cumomer balancing

We simulated the isotopomer abundances of metabolites that are measurable by MS and NMR by using cumomer balancing (Wiechert et al. 1999) with stoichiometrically feasible flux values. Cumomers (*cum*ulative isotop*omers*) are defined as sums of specific isotopomers. This transformation enables conversion of nonlinear isotopomer balance equations to cascades of linear cumomer balance equations. Cumomers can be easily transformed back into isotopomers(Wiechert et al. 1999). Cumomer balancing provides identical results as the more recent technique of elementary metabolite unit (EMU) balancing, but may require longer simulation times.

2.2.3. Isotopomer measurements simulated during ILE design

For the identifiability analysis we simulated a comprehensive list of isotopomers ($\boldsymbol{I_{sim}}$) of metabolites that are known to be measureable by MS and NMR (Table 2.1 and Table 2.2 list the corresponding metabolites). Usually researchers use either MS or NMR to measure isotopomers; however, since these techniques often provide complementary labeling information for a given metabolite (Christensen and Nielsen 1999), we simulated measurements from both techniques. To differentiate between the fluxes of pathways duplicated in the cytosol and plastid, our PPP model incorporated several metabolites that are known to be synthesized exclusively in one of these two compartments.

Analytical technique	Metabolites measured
MS	Ala_m, Glc, Gly_m, His_m, Phe_m, ribose, Ser_m, starch, Tyr_m, Val_m
1-D or 2-D NMR	Ala_n, Gly_n, His_n, LVAgc, LVAgp, LVArc, Phe_n, Ser_n, Tyr_n, Val_n

Table 2.1. Isotopomer measurements in I_{sim} in the PPP model. The isotopomer abundances of amino acids and carbohydrates whose metabolic precursors are known to be exclusively or predominantly synthesized in the cytosol and the plastid were included in $\boldsymbol{I_{sim}}$. Subscripts 'm' and 'n' indicate isotopomer measurements by MS and NMR respectively for the same metabolite. Abbreviations: Glc, glucose; LVAgc, levulinic acid obtained by hydrolysis of cytosolic glucose, LVAgp, levulinic acid obtained by hydrolysis of plastidic glucose; LVArc, levulinic acid obtained by hydrolysis of cytosolic (RNA) ribose.

Analytical technique	Metabolites measured
MS	Ala_m, Arg_m, Asp_m, Gly_m, Glu_m, Ile_m, Leu_m, Lys_m, Pro_m, Ser_m, Thr_m, Val_m
1-D or 2-D NMR	Ala_n, Arg_n, Asp_n, Glu_n, Ile_n, Leu_n, Lys_n, Met_n, Pro_n, Ser_n, Thr_n, Val_n

Table 2.2. Isotopomer measurements in I_{sim} in the GABA model. The isotopomer abundances of amino acids whose metabolic precursors are known to be synthesized in specific compartments were included in $\boldsymbol{I_{sim}}$. Subscripts 'm' and 'n' indicate isotopomer measurements by MS and NMR respectively for the same metabolite.

The subcellular compartmental origins of particular metabolites are well established (Singh 1998b) whereas those of others may be determined by finding the localizations of the enzymes that catalyze their formation reactions (Huang et al. 2010; Wu, Xiao, and Chou 2011; Chou and Shen 2010; Briesemeister, Rahnenführer, and Kohlbacher 2010; Emanuelsson et al. 2000; Chi and Nam 2012). Metabolites predominantly or exclusively synthesized in the cytosol include soluble Glc, ribose (from RNA) and Ala, whereas those with a plastidic origin include Val, His, Phe and Tyr, starch (Heldt 2004; Singh 1998b) as well as Gly and Ser. Our GABA shunt included the compartment-specific metabolites Ala, Gly, Ser Val, Ile, Pro, Thr, Asp, Glu, Lys and Arg.

2.2.4. Statistical flux identifiability

The mathematical techniques of quantifying identifiability of fluxes in a metabolic network have been established previously (Möllney et al. 1999b; Libourel, Gehan, and Shachar-Hill 2007a; Wiechert et al. 1997b) and are discussed briefly here. The covariance of fluxes with respect to noisy isotopomer measurements is an indicator of flux identifiability, and the diagonal elements of the covariance matrix (***Cov***) represent variances of the corresponding fluxes (Press et al. 1992). The premise of identifiability analysis is that the ***Cov*** can be computed without prior knowledge of the true flux values. *A priori* identifiability analysis thus necessitates use of guessed values of free fluxes required to compute the matrix ***Cov*** which is given by the inverse of the Hessian (***H***) of the chi-square function (χ^2) between $\boldsymbol{I}_{sim}$ and experimentally measured isotopomers ($\boldsymbol{I}_{meas}$**,** this term vanishes thus allowing *a priori*

analysis), (Press et al. 1992)

$$Cov(f,m) = [H(\chi^2)]^{-1}$$

$$H(\chi^2) = \frac{\partial^2 \chi^2}{\partial f_k \partial f_l},$$

where *k and l* are counters that go over all free fluxes. Comparing covariance matrices obtained under different ILE designs amounts to comparing their statistical flux identifiabilities. Previous studies have used scalar statistical criteria such as A- and D- optimality criteria to compare covariance matrices (Möllney et al. 1999b; Libourel, Gehan, and Shachar-Hill 2007a). The A-optimality criterion is defined as the trace of the covariance matrix and the D-optimality criterion is the determinant of the covariance matrix.

$$D_{crit} = \det(Cov)$$

$$A_{crit} = \frac{\mathrm{tr}(Cov)}{n},$$

where *n* is the number of rows or columns in ***Cov***.

A_{crit} signifies the arithmetic mean of the variances whereas $D_{crit}^{1/n}$ signifies the geometric mean of the variances. Since the criteria are proportional to flux variances, high identifiability corresponds to small A- and D- criteria. This work uses the A-criterion as a measure of identifiability since it has certain advantages over the D-criterion that were highlighted by Libourel et. al. (Libourel, Gehan, and Shachar-Hill 2007a). Briefly, the difference between the arithmetic mean (AM) and the geometric mean (GM) of the variances of fluxes is greater for the D-optimal ILE designs (Libourel, Gehan, and Shachar-Hill 2007a). This follows from the AM-GM inequality. This means that D-optimality criterion may lead to a needle shaped

confidence region of fluxes, i.e. a D-optimal ILE design may resolve all but one flux with acceptable confidence interval. In this work we verified this claim for 5 pairs of optimal A- and D- designs (data not shown).

Because the A-criterion is inversely proportional to flux identifiability, we use the term "information yield" (IY), the square root of inverse of A-criterion, as a metric to compare different ILE designs:

$$IY = \frac{1}{\sqrt{A_{crit}}}$$

We implemented all isotopomer simulations and IY calculations on our flux evaluation computer program NMR2Flux+ (Sriram et al. 2004b; Sriram et al. 2008). NMR2Flux+ uses cumomer balancing (Wiechert et al. 1999) (Sec. 2.2) to simulate isotopomer abundances from a given set of fluxes, and uses the global optimization algorithm simulated annealing (Pardalos and Romeijn 2002) to evaluate fluxes from a given set of isotopomer abundances.

2.3. Results and Discussion

2.3.1. Plant PPP fluxes are best identified with 100% 1,2-^{13}C Glc (with our set of isotopomer measurements)

The choice of an appropriately labeled carbon source is paramount in ILE design because it crucially determines both the amount of information obtainable from the experiment and the cost of the experiment. Therefore, our first objective was to determine which of the commercially available labels of glucose provides the maximal information toward identifying fluxes in the plant PPPs. Toward this we

computed IY for ILEs that employ each of the eight commercially available Glc labels and naturally abundant Glc, mixed in different proportions. This analysis (Fig. 2.2a) revealed that the PPP fluxes are best identified with 100% 1,2-^{13}C Glc (IY = 21.9 [arbitrary units]), followed by 100% 3-^{13}C Glc (IY = 19.6) and the popularly used(Williams et al. 2008a; Sriram et al. 2008; Masakapalli et al. 2009b) label 100% 1-^{13}C Glc (IY = 18.2).

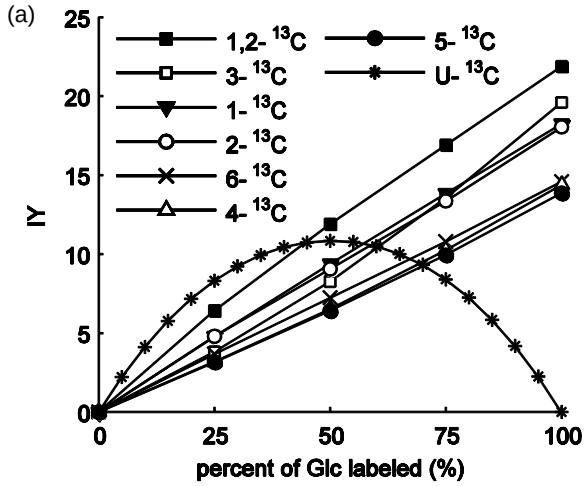


Fig. 2.2 100% 1,2-^{13}C Glc outperforms other commercially available Glc labels in estimating PPP fluxes. This plot depicts simulated IY against extents of labeling of Glc. The values on the horizontal axis indicate the percentage of the label in the supplied Glc; the rest of the supplied Glc is naturally abundant.

For all labels except U-^{13}C Glc, IY increases with the proportion of labeled Glc, implying that dilution of these labels with naturally abundant glucose reduces the information available from them. The exception, U-^{13}C glucose, is explained by the fact that 100% U-^{13}C Glc completely labels all carbon atoms of intracellular metabolites with ^{13}C, thus resulting in no differential distribution of label by different

pathways. Therefore, it is essential to dilute this label with naturally abundant glucose. Additionally, U-^{13}C Glc generates isotopomers that are suitable for measurement by NMR(Sriram, Fulton, and Shanks 2007; Szyperski 1998); therefore, it is not surprising that this label is often employed at proportion as low as 5%(Sriram, Fulton, and Shanks 2007) and 10%(Szyperski 1995).

Furthermore, we analyzed the performances of 55 Glc labels that, to our knowledge, are commercially unavailable except through custom synthesis (all isotopomers of Glc except those listed in. Nine of these 56 Glc labels performed better than 1,2-^{13}C Glc, of which the three best labels were 100% 3,4,5,6-^{13}C Glc (IY = 23.7), closely followed by 100% 1,2,4-^{13}C (IY = 23.3) and 3,5,6-^{13}C Glc (IY = 22.9) (Table 2.3).

Atom(s) of Glc labeled ^{13}C	**IY**
3,4,5,6	23.7
1,2,4	23.3
3,5,6	23.2
1,4,5,6	22.9
2,3	22.6
3,4,6	22.6
1,2,5	22.3
2,4,5,6	22.0
2,3,6	21.9
1,2,4,5,6	21.7

Table 2.3. Performances of commercially unavailable Glc labels for the plant PPP network. This list contains the top 10 best-performing Glc labels. All labels are at 100% of total Glc. Nine Glc labels perform better than 1,2-^{13}C Glc, which is the best performing amongst the commercially available Glc labels.

2.3.2. *In silico* ILEs corroborate the superiority of 1,2-^{13}C Glc over 1-^{13}C Glc in identifying PPP fluxes

The results presented above lead to the question: how does an IY value translate to actual flux identifiability – does the slightly higher IY of the best commercial label 1,2-^{13}C Glc (IY = 21.9) over the popularly used label 1-^{13}C Glc (IY = 18.24) imply that 1,2-^{13}C Glc is significantly better in identifying PPP fluxes? To answer this question, we performed *in silico* ILEs as follows. From arbitrarily chosen values of the four free fluxes in the PPP model, we simulated isotopomer abundances resulting from ILEs employing either 100% 1,2-^{13}C Glc or 100% 1-^{13}C Glc. We then treated these simulated isotopomer abundances as surrogate experimental measurements and allowed NMR2Flux+ to evaluate, by minimizing χ^2 through global optimization, a set of fluxes that best accounted for these surrogate measurements. Repeating this flux evaluation several (478) times from random initial points, we selected the evaluations that converged to χ^2 values less than 20 (corresponding to a confidence level of 99.96% for four degrees of freedom). This resulted in distributions for each flux (Figs. 2.3-2.4), which we compared to the initially chosen ("true") flux values from which we had simulated the isotopomer abundances. Interestingly, some fluxes had bimodal distributions, e.g. *vt3pt* (Fig. 2.3a), *vpfk, vpgifp, vpfkp, vg6pdhp and vg6pt* (not shown), whereas other fluxes had unimodal distributions, e.g. *vtktAf* (Fig. 2.4b). The bimodal distributions exhibited a major peak close to the true flux value and a minor peak far away from it.

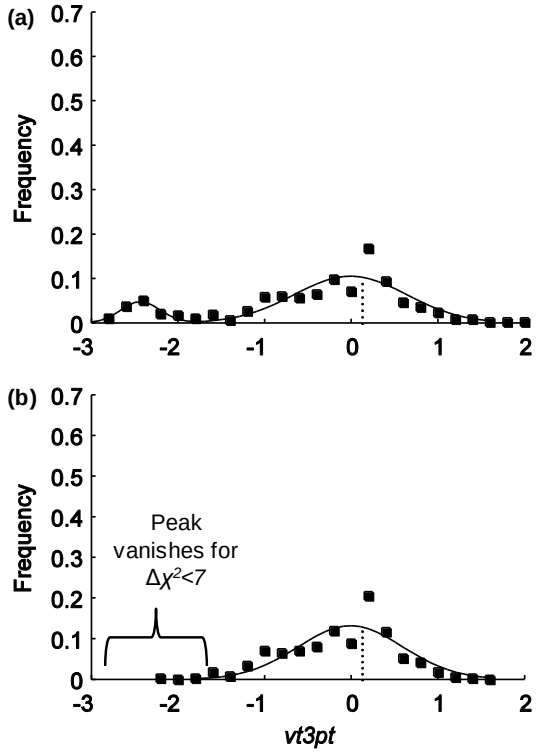


Fig. 2.3 Enforcing stricter criteria for the χ^2 goodness-of-fit function can improve flux estimates and eliminate bimodal distributions. We simulated isotopomer abundances for ILEs employing 100% 1,2-^{13}C Glc, treated these simulated isotopomer abundances as surrogate experimental measurements and then, by minimizing χ^2 through global optimization, evaluated a set of fluxes that best accounted for these surrogate measurements. We repeated this flux evaluation 478 times from random initial points to obtain a flux distribution. Enforcing a goodness-of-fit criterion of $\chi^2 = 20$ (corresponding to a confidence level of 99.96% for four degrees of freedom) gave bimodal distributions for some fluxes (the distribution of the flux *vt3pt* is shown here); enforcing the stricter criterion of $\chi^2 = 7$ eliminated the minor peak and retained the major peak around the "true" flux value from which we had originally simulated the isotopomer abundances. Therefore, enforcing a stricter goodness-of-fit fitting criterion can significantly improve flux estimates.

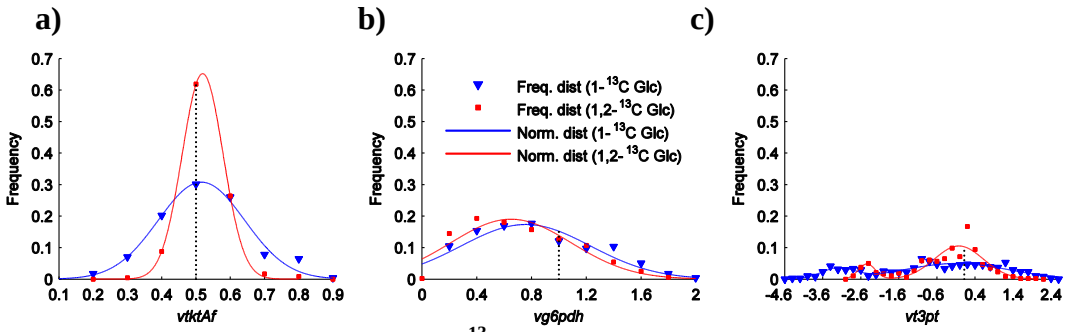


Fig. 2.4 *In silico* ILEs with 100% 1,2-^{13}C Glc estimate several fluxes better than 100% 1-^{13}C Glc. We obtained flux distributions as explained in the caption of Fig. 3 and the text. The flux **(a)** *vtktAf* is more identifiable by the *in silico* ILE with 100% 1,2-^{13}C Glc (red symbols and line) than with 100% 1-^{13}C Glc (blue symbols and line), as the distribution corresponding to 1,2-^{13}C Glc clusters closer to the true flux (dotted line). Both labels identify the flux **(b)** *vg6pdh* to nearly the same extent. The flux **(c)** *vt3pt* is identifiable only with 100% 1,2-^{13}C Glc. This illustrates the superiority of 1,2-^{13}C Glc over 1-^{13}C Glc in estimating fluxes in the compartmented plant PPPs.

For instance, the major peak in the distribution of the flux *vt3pt* represented 385 out of 478 (~81%) flux evaluations and was centered around the true flux value of 0.2, whereas the minor peak that represented the remaining 19% of the flux evaluations was centered away from the true flux value (Fig. 2.3a). However, the points on the minor peak corresponded to χ^2 values between 7 and 20; therefore, using a stringent cutoff of $\chi^2 < 7$ completely eliminated the minor peak (Fig. 2.3b). Applying a $\chi^2 < 7$ cutoff to eight other fluxes that initially showed a bimodal distribution eliminated their minor peaks and retained the major peaks centered around the true flux values. This suggests that accurate identification of PPP fluxes requires a stringent χ^2 cutoff value heuristically learned from a priori simulations. In the distributions discussed in

the rest of Sec. 3.2 (Fig. 2.4), we only consider flux sets corresponding to $\chi^2 < 7$. These flux distributions showed that while both 1,2-^{13}C Glc and 1-^{13}C Glc identified certain fluxes equally well (e.g. the cytosolic oxidative PPP flux *vg6pdh*), 1,2-^{13}C Glc identified certain fluxes significantly better than 1-^{13}C Glc (e.g. the non-oxidative PPP flux *vtktAf*) (Fig. 2.4a,b). Furthermore, 1,2-^{13}C Glc was able to reasonably identify certain fluxes that were not at all identified by 1-^{13}C Glc (e.g. the intercompartmental T3P transport flux *vt3pt*) (Fig. 2.4c). Apart from corroborating the identifiability results of Sec. 3.1, this outcome demonstrates that relatively small increases in IY could translate to significant differences in flux identifiability. Overall, 1,2-^{13}C Glc identified as many as 10 out of the 19 fluxes in the model very close to their "true" values, including glycolytic fluxes in the cytosol (*vpgif, vpfk*) and the plastid (*vpgifb*) and the oxidative PPP flux in the cytosol (*vg6pdh*). However, 1,2-^{13}C Glc does not identify all fluxes satisfactorily – certain fluxes, including glycolytic fluxes in the plastid (*vpfkp*), the oxidative PPP fluxes in the plastid (*vg6pdhp*) and the fluxes of the intercompartmental G6P and P5P transporters (*vg6pt, vp5pt*) were not well identified by this label. Therefore we examined if combining this label with other labels would increase flux identifiability (Sec. 3.7).

2.3.3. Is 1,2-^{13}C Glc always the best choice for the plant PPPs?

Two recent investigations that have focused on designing labels for mammalian PPPs serve as a benchmark for our work. In a study that focused on the mammalian PPP (unicompartmental model, reactions of the PPP assumed irreversible)(Crown and Antoniewicz 2012), Crown and Antoniewicz identified 100% 2,3,4,5,6-^{13}C Glc and

its complement 100% 1-^{13}C Glc as the best Glc labels for this pathway. In another study focused on mammalian primary metabolic pathways including the PPP (unicompartmental model, reversible reactions included)(Metallo, Walther, and Stephanopoulos 2009), Metallo et al. identified 1,2-^{13}C Glc as the best among 11 Glc labels for estimating PPP and glycolysis fluxes. Interestingly, our work converged on the result of Metallo et al. – we identified 1,2-^{13}C Glc as the best commercially available label and its complement 3,4,5,6-^{13}C Glc as the best commercially unavailable label for the plant PPPs. This similarity is despite several major differences between the metabolic network models and isotopomer measurements considered by Metallo et al. and our study – our PPP model contains compartmentalized duplicates of the PPP and glycolysis as well as intercompartmental transport reactions, a hallmark of plant cells. Additionally, we have considered a larger number and variety of readout metabolites, including essential amino acids not synthesized by mammalian cells, sugars and nucleic acids as well as two complementary isotopomer measurement techniques (NMR and MS). Metallo et al. rationalized the superiority of 1,2-^{13}C Glc by showing that if Glc were labeled at the C-2 atom, it can enrich a larger number of carbon atoms of PPP metabolites than if it were labeled at other carbon atoms (e.g. C-4). This is due the repeated breakage and re-formation the C-1–C-2, C-2–C-3 and C-3–C-4 bonds of Glc in the reversible reactions of the PPP; conversely, the bonds between C-4–C-5 and C-5–C-6 remain largely intact(Kruger and von Schaewen 2003a; Metallo, Walther, and Stephanopoulos 2009). To advance this line of reasoning, we examined how many isotopomer abundances change significantly (by > 0.01 units on a scale ranging from

0 to 1 units) when the two most important fluxes in the PPP network – the oxidative PPP fluxes in the cytosol (*vg6pdh*) and the plastid (*vg6pdhp*) – are each perturbed by 20%. This calculation showed that 100% 1,2-^{13}C Glc surpasses both 100% 1-^{13}C Glc and 30% U-^{13}C Glc in both the sum of isotopomer abundance changes (Fig. 2.5a) and the number of altered isotopomers (Fig. 2.5b).

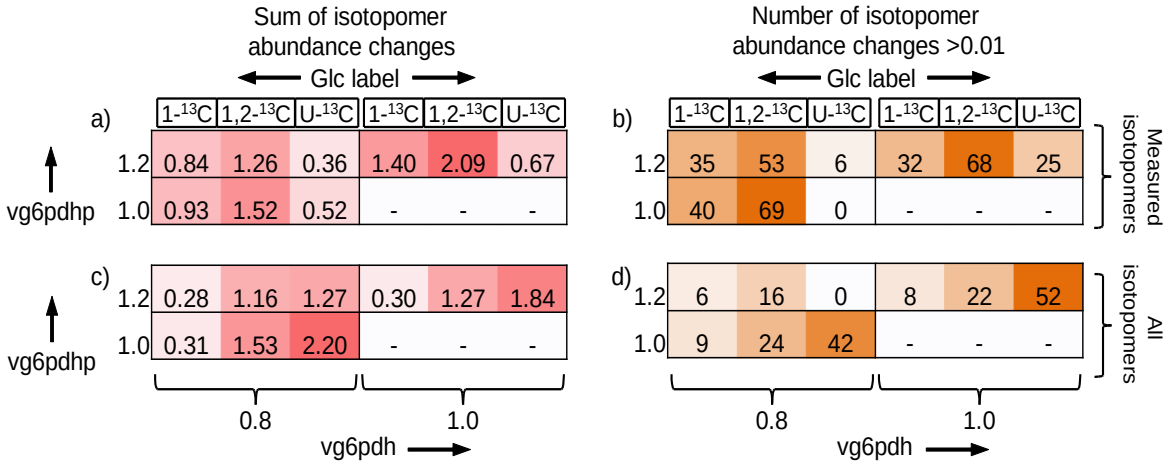


Fig. 2.5 MS- and NMR- derived isotopomer abundances are most sensitive to flux changes in the ILE with 100% 1,2-^{13}C Glc. To examine the sensitivities of isotopomers to flux changes, we simulated ILEs employing 100% 1-^{13}C, 100% 1,2-^{13}C and 30% U-^{13}C Glc, and specifically examined how the perturbation of the cytosolic oxidative PPP flux *vg6pdh* (1.0→ 0.8) and the plastidic oxidative PPP flux *vg6pdhp* (1.0→ 1.2) alters isotopomer abundances. **(a)** the sum of changes and **(b)** the number of changes > 0.01 in the MS- and NMR- derived subset of isotopomer abundances indicates that 100% 1,2-^{13}C Glc renders the isotopomer abundances most sensitive to the given flux changes. Interestingly, **(c)** the sum of changes and the **(d)** number of changes >0.01 in all 2^n isotopomers of all PPP metabolites indicates that 30% U-^{13}C Glc renders isotopomer abundances most sensitive to the given flux changes.

However, this result applies to the set of isotopomer MS- and NMR- derived isotopomer measurements considered in our model. Although this is a large set of isotopomers, MS and NMR can only measure a subset of all 2^n isotopomers of an *n*-carbon metabolite. Unexpectedly, we found that if all 2^n isotopomers of each biomass component could be measured (instead of our subset of MS- or NMR- derived measurements), 30% U-^{13}C Glc surpasses the other two labels in both the um of isotopomer abundance changes (Fig. 2.5c) and the number of altered isotopomers (Fig. 2.5d). Therefore, 1,2-^{13}C Glc is the best label for the plant PPPs with respect to the subset of isotopomers that can be measured by the currently poplar versions of MS (single quadrupole) and NMR (2-D [^{13}C, ^{1}H] HSQC(Iyer et al. 2008; Sriram et al. 2007) or [^{1}H, ^{1}H] TOCSY(Iyer et al. 2008)). However, the measurement of all possible isotopomers of all PPP metabolites may potentially result in a new experimental design. Quantifying the abundances of all isotopomers of the five-, six- and seven- carbon metabolites of the PPP may require significant advancement of the measurement techniques. However, recent developments such as tandem MS can measure all 2^n isotopomers of four-carbon metabolites such as Asp (Choi and Antoniewicz 2011; Choi, Grossbach, and Antoniewicz 2012). Therefore, it is reasonable to expect that further improvements may make it possible to measure a large fraction of the 2^n isotopomers of each PPP metabolite.

2.3.4. Superior performance of 1,2-^{13}C Glc is largely independent of PPP flux values

A central premise of statistical flux identifiability analysis is that the information

contained in an ILE is generally independent of the values of the fluxes. Therefore, the best labels determined by assuming one set of flux values should also be the best labels for any stoichiometrically feasible set of flux values. To examine if this was the case, we repeated the analysis described in Sec. 2.3.1 and Fig. 2.2 for 36 randomly chosen combinations of free flux values that spanned the stoichiometric range of the metabolic network. This analysis (Fig. 2.6) revealed that 100% 1,2-^{13}C Glc was the best-performing label in 32 of the 36 (89%) combinations of flux values and that 100% 3-^{13}C was the second best Glc label in 20 out of 36 (56%) combinations. The consistent performance of 1,2-^{13}C Glc over a range of feasible flux values validates the premise stated above and suggests that this label should provide significant information for most plant PPP networks.

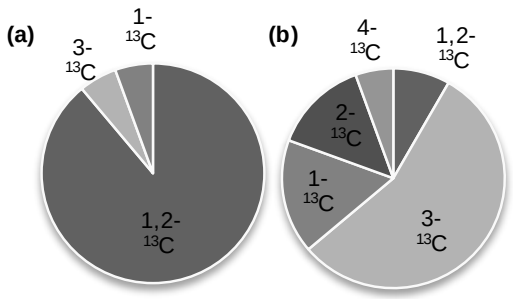


Fig. 2.6 Superior performance of 1,2-^{13}C Glc label is consistent across different sets of free flux values. Calculation of IY for 36 ILEs with different sets of free flux values shows that **(a)** 100% 1,2-^{13}C Glc performs the best for 89% of the free flux combinations and **(b)** 100% 3-^{13}C Glc performs second-best for 56% of free flux combinations, thus validating the premise that the identifiability analysis is fairly independent of free flux values. Some of the isotopomer abundances in $\boldsymbol{I_{sim}}$ were absent in these simulations; nonetheless, this does not affect the performance of the Glc labels(Libourel, Gehan, and Shachar-Hill 2007a).

2.3.5. Labeling information contained in hexose and pentose sugars is critical in elucidating PPP compartmentation

Traditionally, isotope-assisted MFA has involved the measurement of labeling in proteinogenic amino acids derived from acid hydrolysis of a cell pellet or a protein extract (Szyperski 1995; Szyperski 1998; Schmidt, Nielsen, and Villadsen 1999). This experimentally straightforward technique provides metabolic information from various parts of the central carbon metabolic network because the biosynthetic precursors of the amino acids are distributed throughout this network. However, this may be inadequate for the plant PPPs. The complex carbon rearrangements and intercompartmental transfers of sugars in the plant PPPs may not be reflected in the few amino acids that originate from PPP metabolites. The measurement of labeling in hexose and pentose sugars of the PPP may perhaps reveal more information, especially on compartmentation. For example, carbohydrates such as intracellular sucrose or glucose and sugars in glycosylated protein (mannose, glucosamine) reflect cytosolic hexose phosphates, whereas starch reflects plastidic G6P. Therefore, a comparative analysis of the isotopomers of these compounds can potentially reveal differences between cytosolic and plastidic G6P pools, as conceptualized by Roscher et al.(Roscher, Kruger, and Ratcliffe 2000) and experimentally demonstrated by Sriram et al.(Sriram et al. 2004b) Similarly, ribose in nucleic acids likely has a predominantly cytosolic origin; hence its isotope labeling may contrast with the pentose backbone of histidine, which has a plastidic origin. Consequently, comparing the isotopomers of ribose from nucleic acids and histidine from protein will reveal

differences in the P5P pools of the cytosol and the plastid.

However, the extraction, processing and analysis of many of the aforementioned compounds is laborious and not surprisingly, their concurrent measurement for isotope-assisted MFA is rare and has not been reported for plants. Therefore, it is worthwhile to computationally analyze the incremental benefit of measuring the labeling in hexose and pentose sugars. Toward this, we simulated IY for ILEs with varying proportions of 1,2-^{13}C Glc by sequentially including in $\boldsymbol{I_{sim}}$ labeling measurements from the following biomass components: (i) only proteinogenic amino acids, (ii) intracellular glucose, (iii) starch and (iv) RNA ribose. Clearly, each successive labeling measurement substantially increases IY, the highest incremental benefit being in the case of RNA ribose (Fig 2.8a). Further, to examine whether metabolic information from hexose and pentose sugars is more pronounced in MS or NMR measurements, we calculated IY for exclusively MS measurements (Fig. 2.7a) and exclusively NMR measurements (Fig. 2.7b). With NMR, one can measure hexose and pentose sugars through their five-carbon acid hydrolysis product levulinic acid (LVA)(Sriram et al. 2007): LVA resulting from hydrolysis of glycosylated protein is LVAgc, that resulting from starch hydrolysis is LVAgp and that resulting from ribose hydrolysis is LVArc. Again, each successive labeling measurement provides significant additional information with the highest incremental benefit being in the case of ribose-derived LVA (Fig. 2.7a,b).

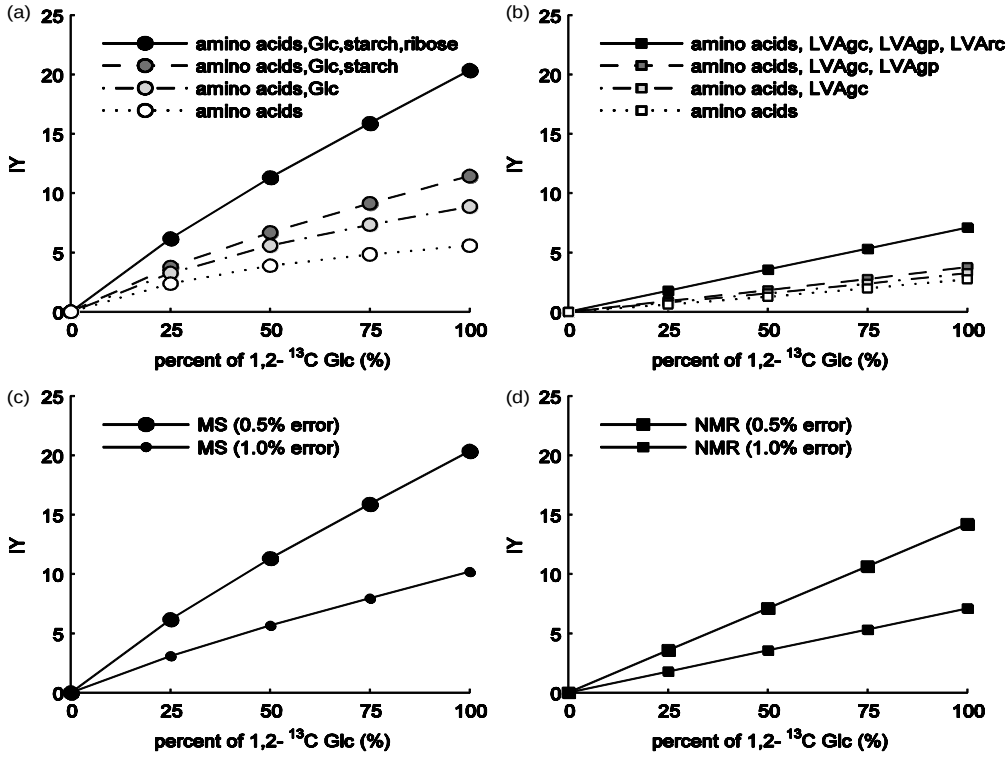


Fig. 2.7 Measurement of isotopomers of hexose and pentose metabolites with cytosolic and plastidic origins substantially improves identifiability. IYs of ILEs with different percents of 1,2-^{13}C Glc and including isotopomer measurements of different biomass components by **(a)** MS and **(b)** NMR show measuring that ribose isotopomers contributes substantially to improving identifiabilities of compartmented PPP fluxes. Additionally, for the same metabolites and measurement errors, the ability of the **(c)** MS to accurately identify the PPP fluxes is greater than that of the **(d)** NMR. Abbreviations: LVAgc, cytosolic glucose-derived levulinic acid, LVAgp, plastidic glucose (starch)-derived levulinic acid; LVArc, cytosolic ribose-derived levulinic acid.

The concurrent measurement of hexose and pentose phosphate pools from the cytosol and the plastid increases IY likely because it disentangles the effects of the oxidative

and the non-oxidative PPPs in the two compartments. In both the cytosol and the plastid, G6P loses its C-1 carbon atom to form P5P in the oxidative PPP. Therefore, reduced labeling of cytosolic or plastidic G6P in an ILE employing 1-^{13}C Glc or 1,2-^{13}C Glc can reveal the presence of the oxidative PPP in the corresponding compartment. However, the complex carbon rearrangements in the downstream non-oxidative PPP in both compartments can confuse this interpretation. For example, high fluxes through the cytosolic oxidative PPP and the plastidic non-oxidative PPP can superimpose over each other, making the compartmentation unidentifiable. However, many of the rearrangements in the non-oxidative PPP are captured in the pentose phosphates synthesized in the respective compartments. Therefore the addition of pentose phosphate isotopomer measurements from the cytosol and the plastid can provide information orthogonal to that contained in the hexose phosphates. This explains the superior performance of the combination of isotopomer measurements from intracellular glucose, starch, RNA and histidine.

2.3.6. MS outperforms NMR in identifying PPP fluxes

The simulations presented above also shed light on the relative efficacies of MS- and NMR- derived isotopomers in estimating PPP fluxes. MS and NMR measure different linear combinations of the isotopomers of a particular metabolite(Christensen and Nielsen 1999): MS measures the mass isotopomer abundances of metabolite fragments, whereas 1-D NMR measures positional ^{13}C enrichments and 2-D NMR measures populations of isotopomers containing different sequences of ^{13}C -^{13}C or ^{12}C-^{13}C bonds. As per the results in Fig. 2.7, MS measurements yield substantially

higher IY values than NMR measurements; therefore, MS is clearly better than NMR in measuring fluxes through the plant PPPs. It could be argued that this is due to the higher analytical sensitivity of MS – based on our previous experience and the literature(Sriram et al. 2004b; Christensen and Nielsen 1999), errors in MS-derived isotopomer abundances are typically less than 0.005 (or 0.5%), whereas errors in NMR-derived abundances are of the order of 0.01 (or 1%). To examine whether MS measurements were superior solely due to their higher precision, we compared the MS and NMR measurements of the same metabolites at the same precision level. Fig. 2.7c and 2.7d depict that even if MS- and NMR- derived isotopomers had identical precision; MS gives higher IY values than NMR. This implies that, for the given metabolites whose labeling was measured, MS performs better than NMR in identifying plant PPP fluxes because the types of isotopomers it measures are more sensitive to PPP fluxes, and not only because it is a more precise technique. In other words, the MS measurements confer higher structural identifiability due to their increased sensitivity to PPP fluxes as well as higher statistical identifiability due to their higher precision. The superiority of MS over NMR is specific to the PPPs; the relative strengths of these techniques may compare differently for a network with a different topology.

2.3.7. Performance of Glc labels in pairs and triads of ILEs in estimating PPP fluxes

We quantified the performance of ILEs that simultaneously employed pairs of commercially available Glc labels at different proportions (data not shown). This

analysis revealed that there was no merit in using mixtures of Glc labels, because in general, an ILE employing a mixture of two labels had a lower IY than ILEs that individually employed each label constituting the mixture. For instance, an ILE employing a mixture of 50% 1,2-^{13}C Glc and 50% 1-^{13}C Glc (IY = 17.5) had a lower IY than one employing 100% 1,2-^{13}C (IY = 21.9; Fig. 2.2a) or 100% 1-^{13}C Glc (IY = 18.2; Fig. 2.2a). This is because a mixture of two labels suffers from the dilution effect that occurs when a single label is diluted with a naturally abundant version of the carbon source (Fig. 2.2a). Isotopomers resulting from one of the two labels mask those of the other, thus diminishing the information obtained from the label mixture. Nevertheless, the deployment of two Glc labels in two or more parallel ILEs followed by flux evaluation from the combined measurements of both the ILEs should enhance flux identifiability. This will leverage the unique information offered by the two labels without allowing one to mask the other (Schwender, Shachar-Hill, and Ohlrogge 2006; Libourel, Gehan, and Shachar-Hill 2007a). To explore this possibility, we examined the performance of pairs and triads of commercially available Glc labels when used in parallel ILEs. All labels performed relatively better when paired with 100% 1,2-^{13}C Glc and relatively worse when paired with 50% U-^{13}C Glc performed (Fig. 2.8a). The pair {1,2-^{13}C, 3-^{13}C Glc} performed the best (Fig. 2.8a), the performance of each label increasing with increase in its proportion and reaching a maximum at 100% proportions of both labels (Fig. 2.8b). The IY of the combination {100% 1,2-^{13}C, 100% 3-^{13}C Glc} was 29.5 (Fig. 2.8a), significantly higher than that of just 100% 1,2-^{13}C Glc (IY = 21.9; Fig. 2.2a).

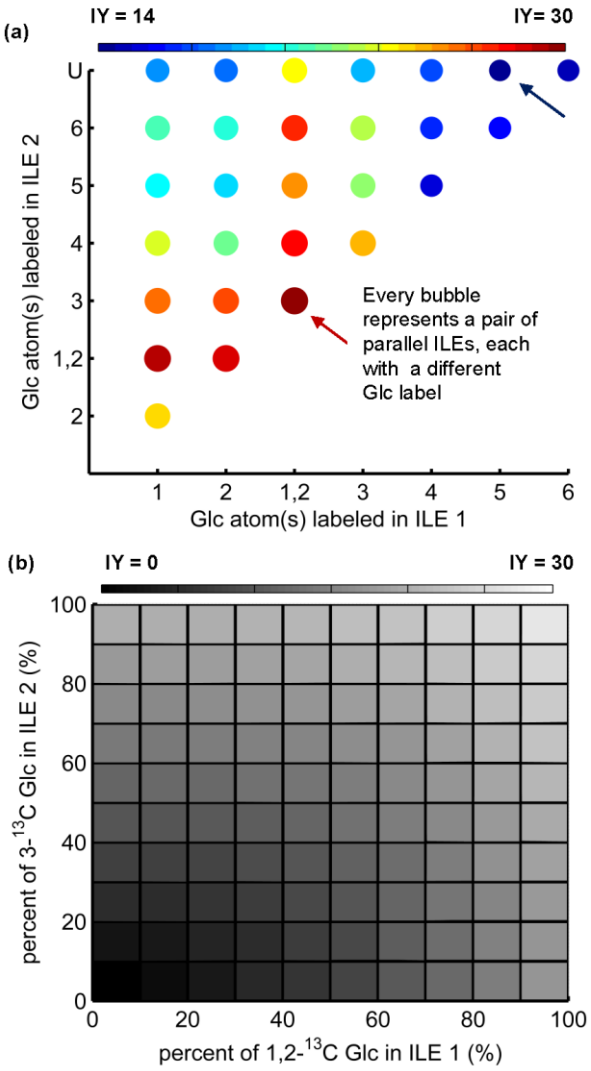


Fig. 2.8 Identifiabilities of PPP fluxes improve upon pooling isotopomer abundances from two parallel ILEs are pooled. (a) The axes list the 100% Glc labels (50% in case of U-^{13}C) in two parallel ILEs – ILE1 and ILE2. Each bubble represents a pair of parallel ILEs whose IY was obtained by pooling the isotopomer measurements from these ILEs. Both the sizes of the bubbles and their color (as indicated in the color bar) are independently proportional to IY. Pooling isotopomer measurements from ILEs with 100% 1,2-^{13}C Glc and 3-^{13}C Glc (IY = 29.5, highlighted by red arrow) is the most advantageous and better than the single ILE with 100% 1,2-^{13}C Glc (IY=21.9). **(b)** IY (lighter shades of gray correspond to

increased IY as shown in the color bar) of pooled isotopomer measurements from ILEs with 1,2-^{13}C and 3-^{13}C increases with increase in their proportions, i.e. dilution of any of the two labels with naturally abundant Glc is undesirable.

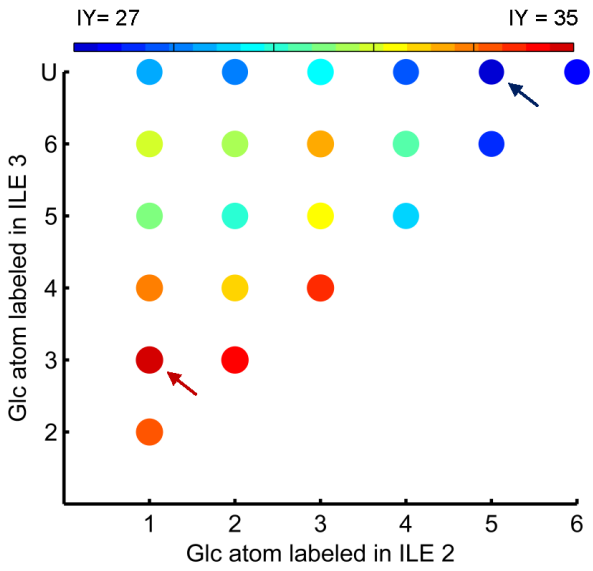


Fig. 2.9 Performances of Glc labels for triads of ILEs. This bubble plot is similar to the one in Fig. 8a. After locking 100% 1,2-^{13}C Glc as one label, we determined IYs for triads of ILEs employing this label and two other Glc labels. Pooling measurements from ILEs using 100% 1,2-^{13}C Glc, 100% 3-13C Glc and 100% 1-13C Glc respectively is the most advantageous, with IY = 34.8.

Identifying optimal triads of ILEs was a more difficult problem due to the large number of possible triads. To circumvent this difficulty, we locked 100% 1,2-^{13}C Glc as one of the labels in the triad (due to its superior performance established in Sec. 3.1-3.4). Then we explored two other labels that could be used with 1,2-^{13}C Glc in parallel experiments. The triad {100% 1,2-^{13}C Glc, 100% 1-^{13}C, 100% 3-^{13}C Glc}

performed the best with IY = 34.8 (Fig. 2.9), significantly higher than the best pair{1,2-^{13}C, 3-^{13}C Glc} (IY = 29.5; Fig. 2.8a) or single label 1,2-^{13}C (IY = 21.9; Fig. 2.1). We investigated the best labels or label combinations for the plant PPPs when exclusively MS-derived or exclusively NMR-derived isotopomer measurements are available (Table 2.4; all labels at 100% proportion).

Measured by	Best five single ILEs		Best five paired ILEs			Best five triads of ILEs			
	Label	IY	Label 1	Label 2	IY	Label 1	Label 2	Label 3	IY
MS	1,2	20.3	1,2	1	27.3	1,2	1	3	32.6
MS	1	17.7	1,2	3	27.0	1,2	1	2	32.5
MS	3	17.7	1,2	2	26.6	1,2	1	4	32.3
MS	2	16.8	1,2	4	26.4	1,2	3	2	32.0
MS	6	13.7	1,2	6	25.5	1,2	3	4	32.0
NMR	3	7.5	3	1,2	10.6	1,2	3	4	11.9
NMR	1,2	7.1	3	4	9.4	1,2	3	2	11.8
NMR	2	5.2	3	2	9.3	1,2	3	6	11.5
NMR	4	4.0	4	1,2	9.0	1,2	3	1	11.0
NMR	5	3.7	3	6	8.9	1,2	3	5	10.9

Table 2.4. Performance of Glc labels when I_{sim} comprised solely MS- or solely NMR-derived isotopomer measurements. This list contains five of the best performing Glc labels and their IYs for single, pairs and triads of ILEs. MS measurements have higher IYs when compared to corresponding NMR measurements.

1,2-^{13}C, 3-^{13}C and 1-^{13}C Glc consistently featured amongst the top labels, 1,2-^{13}C Glc being more prominent when MS measurements are available and 3-^{13}C Glc being more prominent when NMR measurements are available. This suggests that more or

less the same set of labels is optimal when either MS or NMR measurements are available, although the use of MS leads to substantially greater flux identifiability.

2.3.8. Performance of labeled carbon sources for GABA shunt

We chose Ala and Gln as the carbon sources for the GABA shunt network because they enter this network via completely different routes – Ala enters through the mitochondrial TCA cycle and Gln through the plastidic GOGAT cycle. Therefore, their differential labeling can potentially provide significant flux information for this network. We evaluated IY for commonly available Ala and Gln labels at different proportions. Table 2.5 lists the five best performing labels of Ala and Gln, of which the combination {100% 2-^{13}C Ala, 100% U-^{13}C Gln} has the highest IY (= 71.8) toward estimating GABA shunt fluxes.

Ala label	**Percent Ala**	**Gln label**	**Percent Gln**	**IY**
2	100	U	100	71.8
2	100	U	75	70.3
2	100	U	50	68.1
2	100	3	75	65.8
2	100	3	100	65.6

Table 2.5. Five best-performing Ala and Gln labels for GABA shunt network. An ILE with 100% 2-^{13}C Ala and 100% U-^{13}C Gln corresponds to the best IY.

The combination {100% 2-^{13}C Ala, 75% U-^{13}C Gln} has nearly the same IY (=70.3) 75% 3-^{13}C Gln and may serve as a good substitute especially as it offers high identifiability at less than 100% proportion, i.e. at reduced experimental cost.

The fluxes in the GABA network are more interlinked than those in the PPP network

due to the presence of many cyclic pathways. This makes it difficult to track the fates of different carbon atoms. Additionally, the fates of labeled carbon atoms originating from Ala and Gln cannot be distinguished from each other once they are assimilated into succinate ($Succ_m$). Thus speculating on the reasons for the relative merits of different Ala and Gln labels is not easy. Nevertheless, simulations showed that an ILE employing 100% 2-^{13}C Ala and 100% U-^{13}C Gln generated a greater number of isotopomers in the GABA network than other ILEs (data not shown), which partially explains the superior performance of this label combination.

2.4. Summary and outlook

This article explores in substantial detail the design of ILEs toward quantifying fluxes through two important, complex, compartmented plant metabolic pathways. We determined optimal combinations of commercially available Glc labels for the PPP as well as Ala and Gln labels for the GABA shunt. In particular, we established that given currently popular isotopomer measurement techniques (single quadrupole MS and 1-D or 2-D NMR), 1,2-^{13}C Glc is a powerful and robust label for the plant PPPs. We also calculated that its potency can substantially be improved by combining it with other labels (e.g. 3-^{13}C Glc, 1-^{13}C Glc and U-^{13}C Glc) in parallel ILEs. We showed that measuring the labeling patterns of hexose and pentose moieties synthesized exclusively in the cytosol or the plastid is important toward evaluating fluxes in the individual compartments. Specifically, the concurrent measurement of RNA ribose, intracellular glucose or sucrose and starch, although laborious, adds critically to the information obtained from the ILE. Additionally, we showed that MS

outperforms NMR in identifying fluxes in the PPPs. The label designs and measurements proposed in this study have not been simultaneously employed for plant MFA. End-users of this work should bear in mind that although our metabolic models are representative of typical scenarios in plant cells, our optimal ILE designs obtained are most directly applicable to the ILEs employing the carbon sources and isotopomer measurements that we have considered. Researchers employing other labeled carbon sources (e.g. sucrose instead of glucose for the PPP), significantly different isotopomer measurements (e.g. certain intracellular metabolites not considered by us) or techniques (e.g. tandem MS) should repeat our analyses with appropriate changes to the model. Furthermore, the choice of labeled substrates, usually the largest contributing factor to flux identifiability, is heavily influenced by the costs of the substrates. The results in our study are not cost-sensitive since changes in prices are dependent on factors beyond the control of investigators. Therefore, we have examined ILE designs involving both commercially available "catalog" labels and exotic, expensive labels that may be available only through custom synthesis. Ultimately, a balance of the aforementioned factors will enable an end-user to select an appropriate ILE. Currently, we are employing the label designs proposed in this study to investigate metabolic fluxes in *Arabidopsis thaliana* and poplar cell suspensions.

Designing ILEs is a rigorous computational process due to the variety of available label and measurement possibilities that need to be optimized. Nevertheless, it offers valuable insights toward performing an efficient ILE and ensures that maximum information if gained from the ILE. In the future, this work can be advanced by

making use of recently reported improvements in isotope MFA on the analytical and computational fronts. On the analytical front, it is necessary to expand the spectrum of intracellular metabolites whose labeling is analyzed as well as to use high resolution instruments such as liquid chromatography-MS (Kleijn et al. 2007). In this context, Antoniewicz and co-workers' tandem MS methodology (Choi and Antoniewicz 2011; Antoniewicz, Kelleher, and Stephanopoulos 2007b) and novel NMR methods are likely to enable the measurement of a much larger subset of isotopomers than is currently possible. On the computational front, it is essential to use optimization algorithms to efficiently probe the multidimensional space of all available isotope labels for a given ILE. Recently, Stephanopoulos and co-workers applied a genetic algorithm (Metallo, Walther, and Stephanopoulos 2009), while Palsson and co-workers applied Monte Carlo sampling toward this purpose (Schellenberger et al. 2012). Furthermore, Antoniewicz and co-workers have pioneered an EMU-based technique that rationally deduces the optimal labels for an ILE by tracing the number of different ways a product isotopomer can be synthesized from given substrate isotopomers. Current implementations of this technique (Crown and Antoniewicz 2012; Crown, Ahn, and Antoniewicz 2012) have focused on illustrative or irreversible networks; in the future this technique may provide significant insights on complex networks such as the ones explored in this article. Finally, methods that integrate other omics studies such as transcriptomics and proteomics with MFA can augment and enhance the flux information available from isotope MFA.

Acknowledgments

This work was partially funded by the National Science Foundation (award number IOS 0922650) as well as Department of Chemical and Biomolecular Engineering, University of Maryland and A. James Clark School of Engineering, University of Maryland (faculty startup grant to GS).

3. Chapter 3: Isotope-assisted metabolic flux analysis on heterotrophic *Arabidopsis thaliana* under contrasting light treatments reveals negligible effect of light on central carbon metabolism

Shilpa Nargund[1], Ho-Man Yeung[1], Vitali Tugarinov[2] and Ganesh Sriram[1]*

[1]Department of Chemical and Biomolecular Engineering, University of Maryland, College Park, MD 20742, USA

[2]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA

Author contributions

SN and GS conceived this study. SN performed the experiments and computations under the direction of GS. SN wrote the manuscript and GS edited it. VT developed the NMR acquisition methods. HY developed GC-MS techniques to measure glucose and ribose isotopomers.

Abstract

To gain a system-wide perspective on the effects of light on metabolism in heterotrophic plant cells, we performed isotope-assisted metabolic flux analysis (isotope MFA) on *Arabidopsis thaliana* (Arabidopsis) cell suspensions grown in continuous light or dark. Isotope MFA uses isotopic tracers to map carbon traffic (fluxes) through metabolic pathways and thus provides valuable information toward understanding cell physiology. We used statistical design methods to judiciously choose the isotopic tracers (in this study, differently ^{13}C-labeled isomers of glucose), and determined the isotopically labeled metabolite measurements that can provide the most information on primary metabolic fluxes. On the basis of these results, we performed three parallel tracer experiments that used [1-^{13}C], [1,2-^{13}C] and 30% [U-^{13}C] glucose on Arabidopsis cells grown under continuous light or dark. After attainment of isotopic and metabolic steady state, we analyzed the ^{13}C labeling patterns of several biomass components using both gas chromatography-mass spectrometry (GC-MS) and 2-D [^{13}C, ^{1}H] nuclear magnetic resonance (NMR). Together, the three isotope labeling analyses provided over 1300 distinct data points under each light treatment which was collectively fitted to a multicompartmental, central carbon metabolic network model. Isotope MFA revealed that light does not affect flux distributions at intracellular branch-points despite differences in biomass accumulation and protein content between the two light treatments. The similar metabolic fluxes between the two conditions indicate that light plays no role in regulating metabolism of heterotrophic Arabidopsis cells

3.1. Introduction

Light plays dual roles in plant metabolism – the provision of energy for CO_2 fixation and the regulation of a wide spectrum of metabolic. Consequently, light affects various facets of plant metabolism including carbon and nitrogen fixation, the accumulation of various metabolites as well as growth and development, at the systems level (Bläsing et al. 2005). Multiple metabolic pathways such as the Calvin cycle, glycolysis, TCA and phenylpropanoid biosynthesis pathway are regulated by light as shown by gene expression studies on *Arabidopsis thaliana* (Arabidopsis) leaves (Ma et al. 2001). Particularly, the Calvin cycle enzymes glyceraldehyde 3-phosphate dehydrogenase, fructose-1,6-bisphosphatase, sedoheptulose-1,7-bisposphatase and phophoribulokinase show increased activity in light (Bukhov 2004). Light also induces glutamine synthetase, an essential enzyme in nitrogen assimilation (Peterman and Goodman 1991) and represses glucose-6-phosphate dehydrogenase which catalyzes the first, NADPH-providing step of the oxidative pentose phosphate pathway (PPP) (Scheibe, Geissler, and Fickenscher 1989; Hutchings, Rawsthorne, and Emes 2005). Furthermore, the absence of light causes accumulation of vegetative storage proteins (Berger et al. 1995), whereas its presence causes accumulation of starch (Hendriks et al. 2003). Employing metabolic flux analysis to investigate the role of light in plant metabolism will enable a systems-level analysis and complement similar transcriptomic studies on Arabidopsis plants (K. Thum et al. 2008). Metabolic fluxes, the rates of carbon flow through the metabolic pathways, can be quantified by isotope-assisted metabolic flux analysis

(MFA). It involves conducting isotope labeling experiments wherein isotopically labeled (e.g. ^{13}C, ^{15}N, ^{17}O) substrates are fed to the plant cell culture or tissue and the isotopomer (*isotop*e i*somer*) abundances of biomass components such as proteins, lipids, sugars and nucleotides, are measured using nuclear magnetic resonance (NMR) or mass spectrometry (MS). The fluxes are estimated by fitting these isotopomer abundances to a model of the metabolic network.

Investigators have only recently begun exploring the metabolic flux landscape in Arabidopsis. In a pioneering metabolic flux analysis study on Arabidopsis cell suspensions, Williams et al. (2008). quantified fluxes in cell suspensions grown under two oxygenic conditions using [1-^{13}C] glucose to conclude that increased oxygen supply increases biomass accumulation but does not affect the flux distribution (Williams et al. 2008b). In 2009, Lonien et al., mapped the fluxes in wild type, and low seed oil phenotype mutant embryos – regulatory mutant *wrinkled1* and a double mutant in plastidic pyruvate kinase (*pkp$\beta_1$pkpα*). The flux through the pyruvate kinase reaction, which contributes precursors to lipid synthesis, surprisingly, was significantly lower in the *wrinkled1* mutant, than in *pkp$\beta_1$pkpα* despite having higher enzyme levels. They concluded that the single reaction step mutation in *pkp$\beta_1$pkpα* was compensated by other glycolytic reactions, but the transcription factor mutation in *wrinkled1* affects multiple reactions to effectively lower the flux through pyruvate kinase (Lonien and Schwender 2009a). To query the role of the duplicate PPPs in the cytosol and plastid of plant cells, Masakapalli et al., in 2010, performed MFA on Arabidopsis suspension cells with three parallel isotope labeling experiments (ILEs;

[1-^{13}C], [2-^{13}C] and 10% [U-^{13}C] glucose). Based on their flux estimates and biochemical evidences from other studies, they concluded that the cytosol contained only the oxidative steps of the PPP while the plastid contained both the oxidative and non-oxidative reactions of the PPP (Masakapalli et al. 2010). Williams et al., in 2010 verified that the genome scale metabolic network model developed by Poolman et al., ((Poolman et al. 2009) can predict the flux estimates obtained by MFA on Arabidopsis cell suspensions subjected to high temperature or hyperosmolarity (Williams et al. 2010).

Previously, researchers have studied the effect of light on gene and protein expressions in Arabidopsis plants (K. E. Thum et al. 2003; Phee et al. 2004). However, the effect of light on the fluxome of Arabidopsis plant cells is not known. Knowledge of fluxomic changes in heterotrophic plant cells due to light regulation will enhance our understanding of the role of light and may bridge the gaps between the transcriptomic and proteomic studies.

In this study, the metabolic fluxes in Arabidopsis cell suspensions, subjected to continuous light or dark, were quantified by robustly designed ILEs which employed [1-^{13}C], [1,2-^{13}C] and 30% [U-^{13}C] glucose. Over 1300 mass isotopomers (the largest data set measured for MFA on Arabidopsis) obtained under each light treatment were fitted to a multicompartmental model of primary carbon metabolic pathways in plant cells. Distinct biomass accumulation and protein contents observed under the two light treatments lead to negligible difference in flux distribution in the primary carbon metabolism of Arabidopsis cell suspensions.

3.2. Results

3.2.1. Biomass accumulation and protein content differs between the two light treatments

Dark grown cell suspensions accumulated lesser biomass (239±12 mg) than light grown cell suspensions (271±11mg) in light (two sample t-test, p-value = 0.0001;Fig. 3.1a) despite consuming the same amount of glucose (1.12±0.04 mmol glucose/day/mg biomass in dark and 1.13±0.05 mmol glucose/day/mg biomass in light (two sample t-test, p-value = 0.98)). The partitioning of carbon into protein also differed between the two light treatments (Fig.3.1b).

3.2.2. Isotopic steady state attained between day 5 and 6

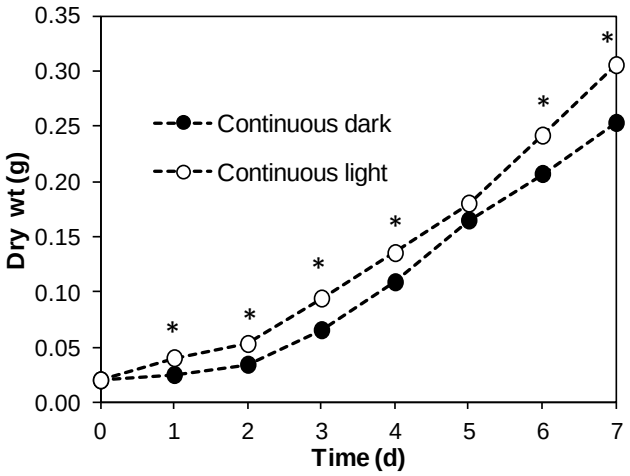
Proteinogenic amino acids from Arabidopsis cells grown on 50% U-^{13}C glucose for 4, 5 and 6 days after subculture show negligible difference in ^{13}C enrichments between day 5 and 6 (Fig. 3.2). This indicates that isotopic steady state is attained during this period. Therefore the MFA was carried out on day 6.

3.2.3. Serine and glycine labeling show that light-grown do not fix CO_2 photosynthetically but undergo serine hydroxymethyltransferase reaction

Photosynthetic fixation of CO_2, if present, affects the mass isotopomer distributions (MIDs) of metabolites and consequently the flux estimation. Therefore it is important to verify the occurrence of CO_2 fixation in plant cells. Photosynthetic fixation of CO_2 leaves an isotopomeric signature in the labeling of amino acids serine and glycine

during ILEs. Carbon from CO_2 is preferentially fixed to the 1st position of 3-phosphoglycerate and subsequently to the 1st position of serine and glycine.

(a)



(b)

Biomass component	Continuous light (% dry weight)	Continuous dark (% dry weight)
Cell wall	34.6 ± 2.9	47.6 ± 2.9
Protein	20.8 ± 1.5*	16.6 ± 0.4*
Lipid	4.1 ± 0.5	2.3 ± 1.0
Starch	1.9 ± 0.3	1.9 ± 0.5
Soluble metabolites	9.1 ± 0.7	5.3 ± 0.6

Fig. 3.1. Consistent growth characteristic differences between continuous light and dark treatments. (a) Difference in biomass accumulation between the two light treatments was observed reproducibly (n=3 under each treatment) and consistently over multiple subculture cycles (the errors bars are small and therefore hidden by the symbols). **(b)** The protein content was also different between the two light treatments. *Means are different with p-values < 0.05.

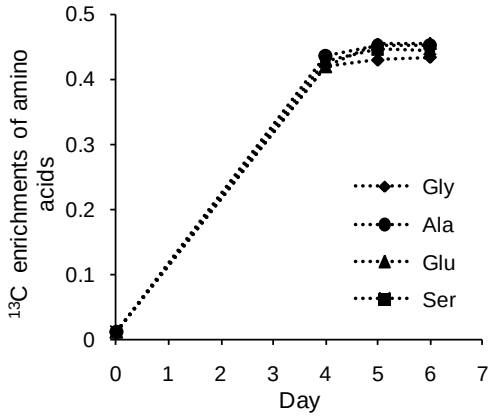


Fig. 3.2. Isotopic steady state. The constant ^{13}C enrichments of proteinogenic amino acids glycine, alanine, glutamate and serine between day 5 and 6 demonstrate attainment of isotopic steady state.

The labeling on the 1st C atom on glycine and serine extracted from light-grown suspension cells under 30% [U-^{13}C] glucose do not exhibit these signatures and are very similar to those extracted from dark grown cell suspensions indicating that CO_2 is not fixed in the light grown cell suspensions.

Interestingly, the 3rd carbon atom of serine is differently labeled in cells under light than dark (Fig. 3.3c) which may be due to the reversibility of serine hydroxymethyltransferase reaction (Ser $\leftrightarrow$ Gly + CO_2) which causes bond breakage and formation and is known to be induced by light (Turner et al. 1993; McClung et al. 2000)**.**

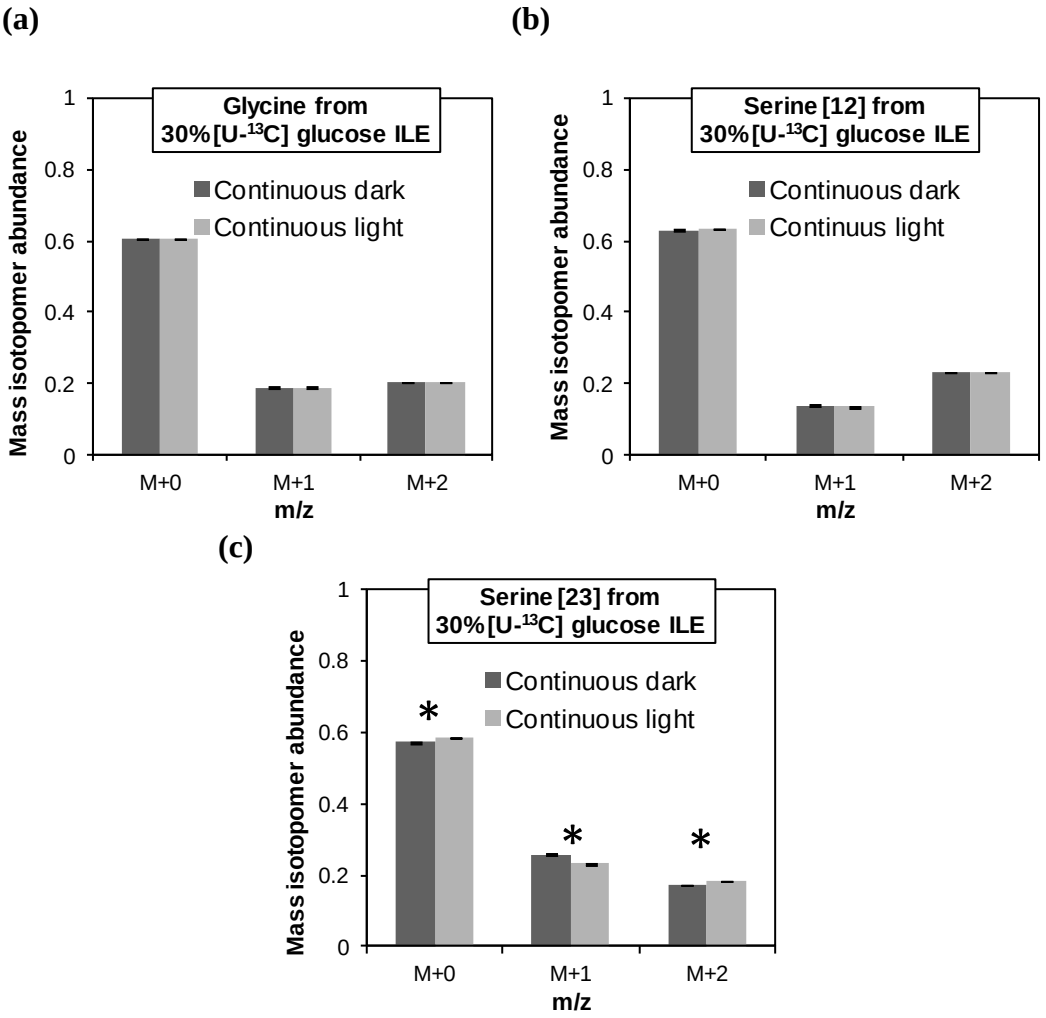


Fig. 3.3. Photosynthetic fixation of CO2 absent in light-grown Arabidopsis cells. The MIDs of **(a)** glycine and **(b)** fragment of serine containing only the 1st two C atoms (serine [12]) from the light-grown suspensions cells under 30% U-^{13}C glucose ILE are similar in both light and dark-grown cells. **(c)** the 3rd carbon atom of serine is differently labeled between the light treatments ***** Means are different with p-value < 0.05.

3.2.4. *In silico* design of isotope labeling experiments predicts that 1,2-^{13}C glucose estimates fluxes in pentose phosphate pathway the best

In a previous *in silico* design study (Nargund and Sriram 2012), we found that the label 100% [1,2-^{13}C] glucose performed the best in identifying plant PPP fluxes. We are particularly interested in the flux changes in the PPP since this pathway shares several of the light sensitive enzymes with the Calvin cycle pathway (Bukhov 2004). Thus three parallel ILEs were conducted with i) 100% [1,2-^{13}C] glucose – the label that best identifies PPP fluxes ii) 100% [1-^{13}C] glucose – a label popularly used to estimate PPP fluxes and iii) 30% [U-^{13}C] glucose – an economic label suitable for estimating tricarboxylic acid cycle (TCA) fluxes. The above combination of ILEs was computationally predicted to reveal more flux information than any single or paired ILEs as shown in our previous study (Nargund and Sriram 2012). The design study also predicted that the labeling information in hexose and pentose sugars, especially ribose, improved PPP flux estimates significantly. Accordingly, we measured the isotopomer abundances of ribose (from RNA) and intracellular glucose.

3.2.5. Metabolic network model

Construction of the metabolic network model is influenced by i) the organism being probed, ii) the range of readout metabolites measured and iii) their mass isotopomer distributions (MIDs). In this MFA study, the readout metabolites measured – proteinogenic amino acids, organic acids, ribose and intracellular glucose, allow estimation of fluxes in the glycolysis, PPP and TCA cycle. The carbon atom

rearrangements and reversibilities of the reactions in these pathways were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG), plant biochemistry textbooks and previous literature (Singh 1998a; Kruger and von Schaewen 2003b). Plant metabolic networks are complicated due to duplication of the glycolysis and PPP in the cytosolic and plastidic compartments. Generally, the compartmentalized duplicate pathways are included in the model if i) the duplicate pathways have distinct fluxes and ii) the duplicate metabolite pools are isotopically distinguishable. It is not possible to determine the former *a priori* but the latter can be verified by comparing the MIDs of corresponding readout metabolites from the two compartments. It may also happen that the duplicate pathways have different fluxes but the duplicate metabolite pools are in isotopic equilibrium due to rapid exchange by intercompartmental transport reactions in which case a single compartmental model should be used.

Table 3.1 lists the biosynthetic precursor(s) and compartment of origin of metabolites, as recorded in literature and as modeled in our metabolic network. Several corresponding readout metabolites that originate from the duplicate pools in the cytosol and plastid are not in isotopic equilibrium under both the light treatments and across the three ILEs (Fig. 3.4).

Glycerol originates from cytosolic triose phosphates (T3P) and has a distinct MID from that of serine which arguably originates from plastidic T3P; therefore two pools of T3P are included in the model (although serine participates in the serine hydroxymethyltransferase reaction (serine ↔ glycine) and may originate in the cytosol or mitovhondrion it is fairly similar to phenylalanine and tyrosine which are

also made from plastidic T3P and therefore a valid comparison with the cytosolic glycerol). Similarly, alanine, synthesized arguably from cytosolic pyruvate is distinguishable from valine that is synthesized from plastidic pyruvate and therefore two separate pools of pyruvate are included. The three carbon atom metabolites T3P and pyruvate can also rapidly exchange with each other to attain isotopic equilibrium.

Readout metabolite	Biosynthetic precursor	Compartment of origin according to literature	Reference	Compartment of origin in our model
Alanine	Pyr	Cytosol/mitochondrion	(Singh 1998b)	Cytosol
Glycine	T3P	Cytosol, mitochondrion, plastid		Plastid
Valine	Pyr	Plastid		Plastid
Leucine	Pyr, AcCoA	Plastid		Plastid
Isoleucien	OAA, Pyr	Plastid		Cytosol, plastid
Proline	α-KG	Cytosol	(Funck et al. 2008)	Cytosol
Serine	T3P	Cytosol, mitochondrion, plastid	(Singh 1998b)	Plastid
Threonine	OAA	Plastid	(Fait et al. 2008)	Cytosol
Pheylalanine	T3P, E4P	Plastid	(Singh 1998b; Heldt 2004)	Plastid
Aspartate	OAA	Plastid	(Fait et al. 2008)	Cytosol
Glutamate	α-KG	Mitochondria, plastid	(Singh 1998b)	Cytosol
Histidine	P5P	Plastid		Plastid
Tyrosine	T3P, E4P	Plastid		Plastid
Glycerol	T3P	Cytosol	(Saha et al. 2006; Ohlrogge & Jaworski 1997)	Cytosol
Malic acid	Malic acid	Mitochondria		Cytosol
Ribose	P5P	Cytosol		Cytosol
glucose	G6P	Cytosol		Cytosol

Table 3.1. Compartment of origin of readout metabolites. To estimate the fluxes in the duplicate pathways in separate compartments, labeling information of readout metabolites originating in these compartments is crucial. The compartments of readout metabolites as recorded in literature and as modeled in this work are listed below. Abbreviations: α-KG, α-ketoglutarate; AcCoA, acetyl coenzyme A; E4P, erythrose-4-phosphate; G6P, glucose-6-phosphate; OAA, oxaloacetic aicd; P5P, pentose-5-phosphate; Pyr, pyruvate; T3P, triose

phosphate.

We found that, in the cytosol, glycerol (from T3P) and alanine (from pyruvate) are isotopically different and in the plastid, serine (from T3P) and valine (from pyruvate) are distinguishable, thus warranting inclusion of separate pools.

Separate pools of pentose-5-phosphate (P5P) were included in the model despite the fact that comparison between the readout metabolites, histidine from the plastid and ribose from the cytosol, was inconclusive (the MIDs of histidine and ribose are not comparable because histidine contains an additional carbon atom that it acquires from C1 metabolism). Duplicate pools of E4P and S7P were also included to allow completion of the PPP pathways in both compartments.

Based on the above observations, we constructed a model that has both the cytosolic and the plastidic glycolysis and PPPs (Fig. 3.5). The metabolites glucose-6-phosphate (G6P), P5P, T3P and pyruvate can be transported across the intercompartmental membranes between the cytosolic and plastidic pools (Schwender, Ohlrogge, and Shachar-Hill 2003b). The cytosol and mitochondrion are treated as a single compartment due to lack of labeling data to distinguish between the two. The reactions of the γ-aminobutyric acid (GABA) shunt and the reactions of the TCA cycle that it bypasses undergo the same carbon atom rearrangements and therefore the fluxes through the two competing pathways cannot be distinguished with the labels used in this study. Indeed, the similarity of the MIDs of GABA and glutamate support this assumption (data not shown).

The metabolic network model has 23 metabolites and 48 net fluxes of which 17 were

measured and 7 independent fluxes. Additionally, there were 20 reversibility extents for the reactions that are reversible and 1 scrambling extent which indicates the extent to which the reaction catalyzed by succinate dehydrogenase (succinate → fumarate) undergoes one of the two possible carbon atom rearrangements.

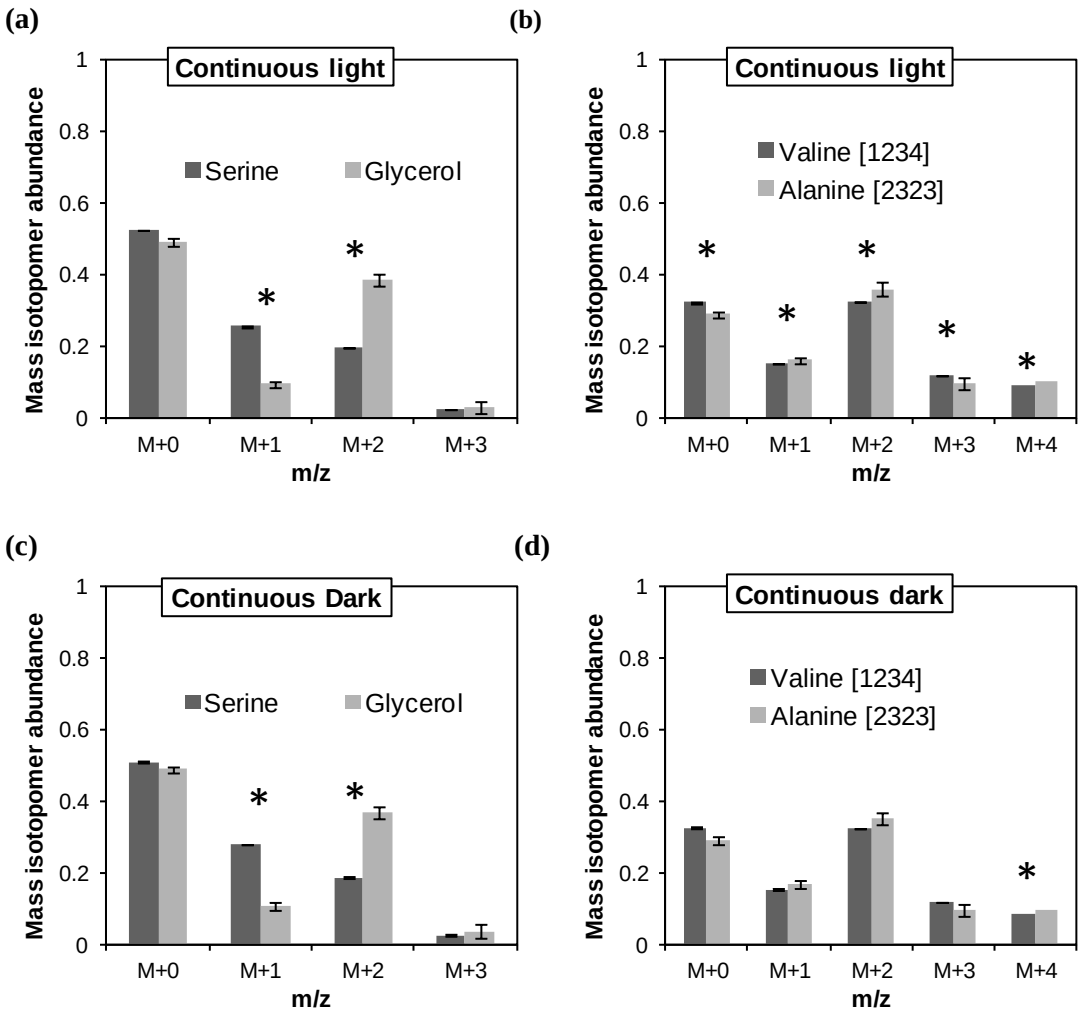


Fig. 3.4. Metabolite pools are not in isotopic equilibrium between the cytosolic and plastidic compartments. The MIDs of **(a,c)** serine and glycerol and **(b, d)** alanine and valine are distinguishable from each other under both the light treatments. The MIDs were obtained from the [1,2-^{13}C] glucose ILE. Similar trends were seen under the other two ILEs as well. *****

Means are different with p-values < 0.05

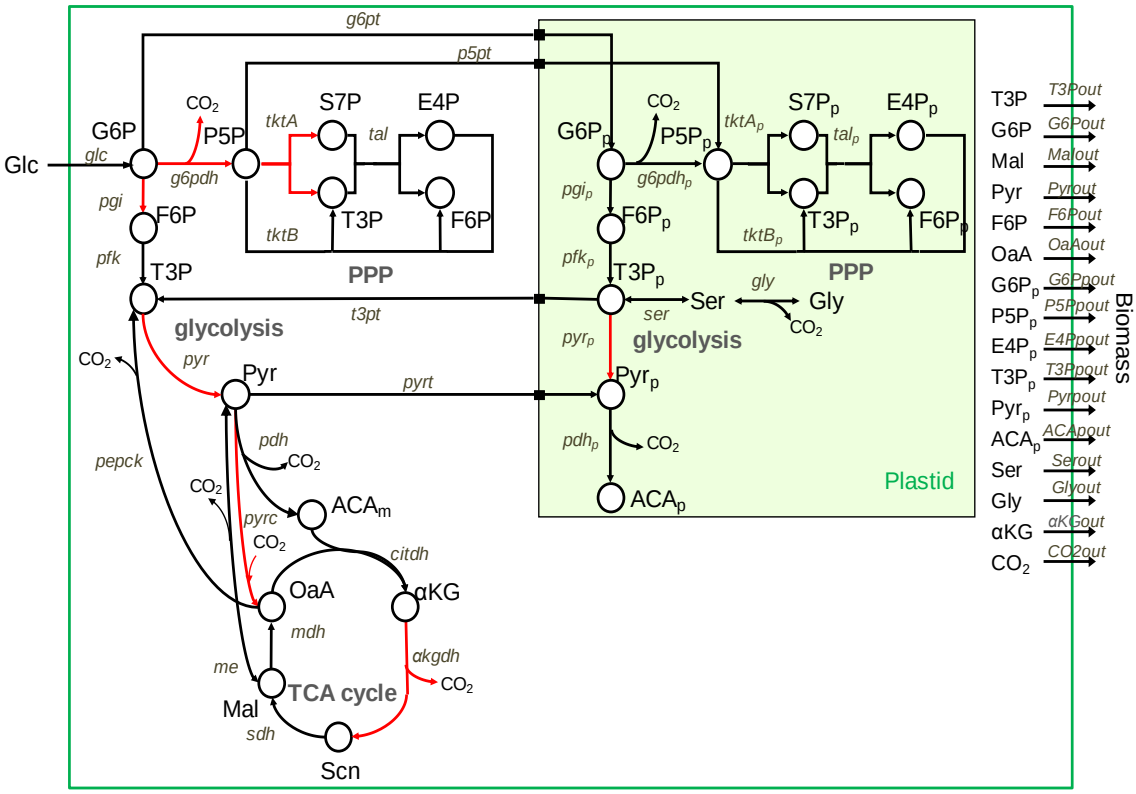


Fig. 3.5. Metabolic network model. The model consists of the glycolysis and PPP duplicated in cytosol and plastid and the TCA cycle in the cytosol. There are 23 metabolites and 69 fluxes (48 net fluxes and 20 reversibility extents). The free fluxes *g6pdh*, *tktA*, *pgi*, *pyr*, *pyrc*, *αkgdh* and pyr_p (marked in red) are fitted to the model whereas the other fluxes, which are linear combinations of the free fluxes, are calculated. The plastidic compartment is enclosed by a dashed box; suffix 'p' indicates metabolites in the plastidic compartment. Abbreviations: αKG, α-ketoglutarate; ACA, acetyl coenzyme A; E4P, erythrose-4-phophate; F6P, fructose-6-phosphate; Glc, glucose; G6P, glucose-6-phosphate; Gly, glycine; Mal, malic acid; OaA, oxaloacetic acid; P5P, pentose phosphates; Pyr, pyruvate; S7P, sedoheptulose-7-

phosphate; Scn, succinate; Ser, serine; T3P, triose phosphates.

3.2.6. Difference in growth between the light treatments affects isotopomer abundances

The ^{13}C enrichments of proteinogenic amino acids extracted from cells grown on 29% U-^{13}C glucose were significantly lower than the expected 29% (between 26-29%) (Fig. 3.6). Since the cells cannot distinguish between ^{12}C and ^{13}C isotopes (Kruger et al. 2007a) the dilution of the label can only be due to the presence of the seed biomass added during subculture. Every week during subculture, aliquots of the culture with known mass of cells, are transferred to new shake flasks with fresh media. This seed biomass takes up the fresh media to synthesize new biomass. During an ILE, the naturally labeled (1.13% ^{13}C) seed biomass is transferred to labeled media (containing 100% [1-^{13}C] or 100% [1,2-^{13}C] or 30% [U-^{13}C] glucose) and it contributes a substantial proportion (10-12%) to the total biomass. To account for the seed biomass, we adjusted the isotopomer abundances of all amino acids based on the extent of dilution estimated from the 30% [U-^{13}C] glucose ILE data (data not shown).

3.2.7. Flux distributions in central carbon metabolism in the two light treatments are similar

The MIDs of proteinogenic amino acids, soluble organic acids, ribose and glucose were measured by GC-MS whereas NMR was used to measure labeling only in the proteinogenic amino acids from the 30% [U-^{13}C] glucose ILE. On combining the data from the two biological replicates, under each light treatment, we obtained 458, 424 and 474 mass isotopomers from the [1-^{13}C], [1,2-^{13}C] and 30% [U-^{13}C] glucose ILEs

respectively.

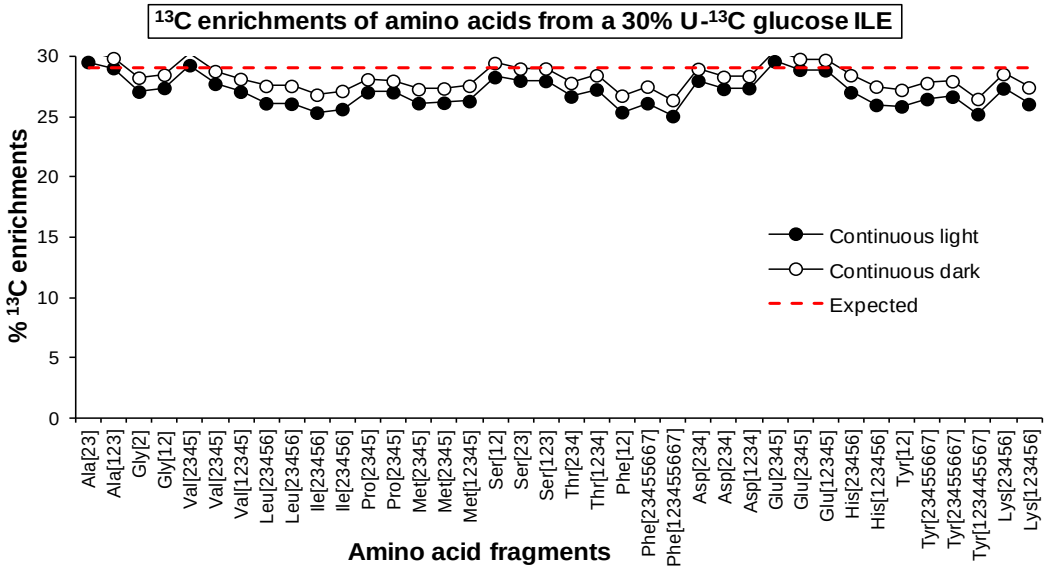


Fig. 3.6. Dilution of the ^{13}C label observed. The ^{13}C enrichments of amino acids extracted from cells grown on 30% U-^{13}C glucose are consistently lower than the expected 30%. This dilution is attributed to the presence of the initial unlabeled seed biomass.

The biological standard deviations of amino acid mass isotopomers were within 0.01-0.8% whereas those of soluble organic acids were between 0.02-3.5%. The amino acids measured by NMR did not have biological replicates and were assigned 5% errors uniformly. The biological averages of the mass isotopomers from all three ILEs were pooled to obtain 678 mass isotopomers that were fit to the model using the program NMR2Flux+. For given stoichiometrically feasible guess values of free fluxes, NMR2Flux+ simulates the three ILEs consecutively to calculate the chi-square function (χ^2) between corresponding simulated and measured MIDs of metabolites (listed in Table 3.1). This χ^2 function is minimized by the global

optimization algorithm - simulated annealing followed by the local optimization algorithm - Powell's method. The data points with biological standard deviations lower than 1% (most amino acid MIDs) led to very high χ^2 and therefore these errors were increased to 1% artificially.

Preliminary data fittings were used to i) ascertain that the global minimum was reached i.e. all fluxes had unimodal distributions and ii) refine the metabolic network model. Although all of the free fluxes had unimodal distributions, the free fluxes *tktf*, *pyrf*, *akgdh*, clustered far away from one of their constraints at higher χ^2 and therefore these constraints were tightened. The serine and glycine MIDs fit poorly and the inclusion of the serine hydroxymethyltransferase reaction improved their fits. The methionine MIDs were eliminated from the data set because they did not fit satisfactorily. We suspect that its MIDs were not accurately measured because it had very low abundance.

The fluxes estimated by using the model described above, explains the measured data satisfactorily (Fig. 3.7). Despite significant differences in growth rates and biomass compositions, the metabolic fluxes in the central carbon metabolism between the two light treatments remain unaltered (Fig. 3.8).

The poorly identified net fluxes (Table 3.2) were the anaplerotic fluxes (*me*, *pyrc* and *pepck*), through which 19±7% of the carbon enters the TCA cycle and the transport fluxes (*g6pt*, *t3pt*, *p5pt*). The reversibility extents of the reactions *tktA*, *tal*, *tktap*, *talp*, *pyrt*, *me* and *ser* and the scrambling extent of *sdh* were also not well identified, but reversibility and scrambling extents are usually not very identifiable (Wiechert et al. 1997a; Nargund and Sriram 2012).

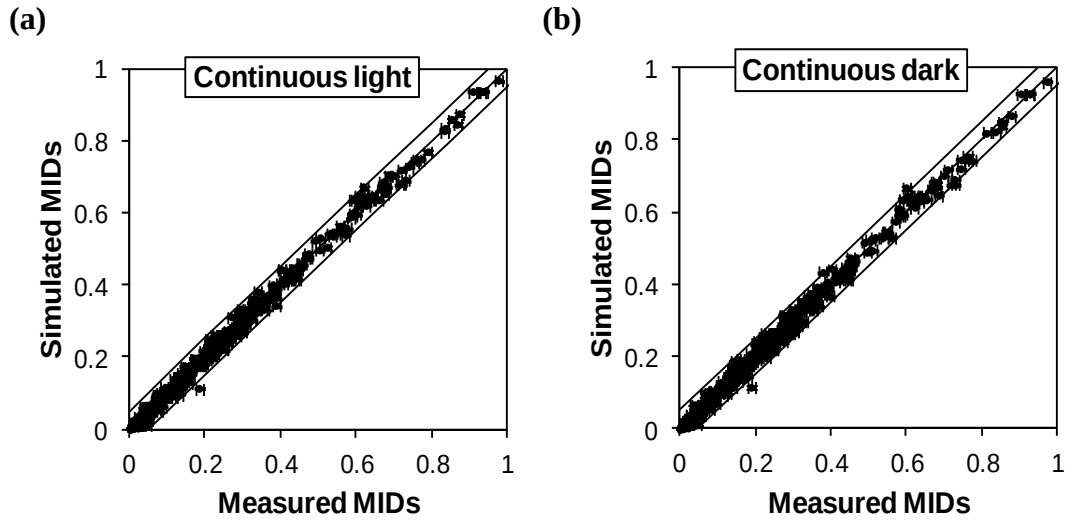


Fig. 3.7. Estimated fluxes account for measured MIDs. The measured and simulated MIDs when plotted against each other lie close to the diagonal line and within 5% error under both **(a)** continuous light and **(b)** continuous dark treatments.

Monte Carlo simulations, carried out ~400 times for each light treatment, were used to obtain confidence intervals of fluxes and most of the net fluxes are remarkably well identified (Table 3.2).

3.2.8. Metabolic network model validation

The assumption that two separate P5P pools are present in the cytosol and plastid, even though the isotopic equilibrium between histidine and ribose could not be ascertained, was validated posteriori. The MID of plastidic P5P was calculated from the MIDs of histidine (measured) and CO_2 (simulated using the flux estimates obtained). We found that the MIDs of the plastidic and cytosolic P5P (measured as

ribose) were distinct under [1-^{13}C] and [1,2-^{13}C] glucose ILE (data not shown) thus validating our model. Furthermore it shows that [1-^{13}C] and [1,2-^{13}C] glucose ILEs are more sensitive to PPP carbon rearrangements than the 30% [U-^{13}C] glucose ILE.

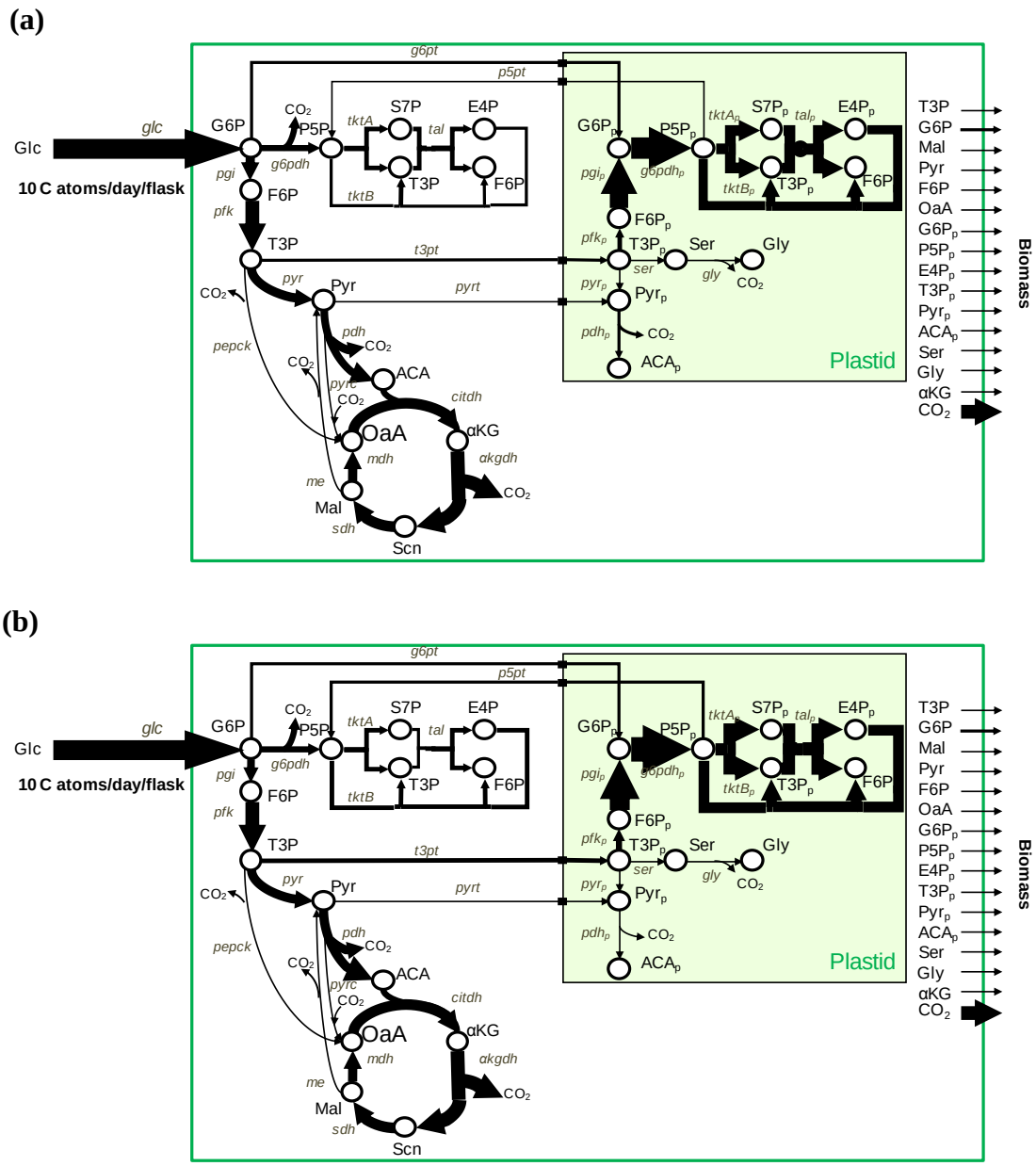


Fig. 3.8. Carbon redistribution is identical between the two light treatments. The fluxes relative to glucose uptake between both the **(a)** continuous light and **(b)** continuous dark

treatments are nearly identical.

Reaction name	Stoichiometry	Continuous dark				Continuous light			
		Net flux		Reversibility		Net flux		Reversibility	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Cytosolic glycolysis									
glci	→ G6P	0.804	0.007	irrev.		0.806	0.009	irrev.	
pgi	G6P → F6P	0.321	0.023	0.990	0.017	0.359	0.030	0.986	0.031
pfk	F6P → T3P + T3P	0.552	0.028	0.218	0.050	0.540	0.031	0.229	0.064
pyr	T3P → pyr	0.758	0.025	irrev.		0.749	0.017	irrev.	
Cytosolic PPP									
g6pdh	G6P → P5P + CO_2	0.216	0.016	irrev.		0.214	0.017	irrev.	
tktA	P5P + P5P → S7P + T3P	0.117	0.012	0.029	0.062	0.092	0.016	0.088	0.132
tal	S7P + T3P → F6P + E4P	0.117	0.012	0.019	0.054	0.092	0.016	0.054	0.090
tktB	P5P + E4P → F6P + T3P	0.117	0.012	0.625	0.033	0.092	0.016	0.680	0.045
Plastidic glycolysis									
pgip	G6Pp → F6Pp	-0.811	0.097	0.990	0.027	-0.738	0.095	0.972	0.113
pfkp	F6Pp → T3Pp + T3Pp	-0.269	0.040	0.240	0.052	-0.212	0.036	0.332	0.062
pyrp	T3Pp → pyrp	0.062	0.017	irrev.		0.094	0.018	irrev.	
pdhp	pyrp → ACAp + CO_2	0.175	0.001	irrev.		0.223	0.001	irrev.	
Plastidic PPP									
g6pdhp	G6Pp → P5Pp + CO_2	0.952	0.098	irrev.		0.858	0.093	irrev.	
tktAp	P5Pp + P5Pp → S7Pp + T3Pp	0.274	0.035	0.039	0.100	0.268	0.039	0.023	0.059
talp	S7Pp + T3Pp → F6Pp + E4Pp	0.274	0.035	0.016	0.049	0.268	0.039	0.013	0.042
tktBp	P5Pp + E4Pp → F6Pp + T3Pp	0.268	0.035	0.684	0.057	0.259	0.039	0.541	0.113
TCA cycle									
pdh	pyr → ACA + CO_2	0.687	0.029	irrev.		0.661	0.020	irrev.	
citdh	ACA + OaA → αKG + CO_2	0.687	0.029	irrev.		0.661	0.020	irrev.	
akgdh	αKG → Scn + CO_2	0.669	0.029	irrev.		0.630	0.020	irrev.	
sdh	Scn → Mal	0.669	0.029	0.980	0.034	0.630	0.020	0.972	0.039
scrambling	Scn → Mal			0.689	0.321			0.534	0.357
mdh	Mal → OaA	0.565	0.025	0.992	0.019	0.518	0.024	0.963	0.045
Anaplerotic reactions									
me	Mal → pyr + CO_2	0.104	0.036	0.276	0.240	0.113	0.028	0.221	0.194
pyrc	pyr + CO_2 → OaA	0.020	0.018			0.019	0.019	irrev	
pepck	OaA → T3P + CO_2	-0.110	0.039	0.576	0.101	-0.137	0.031	0.474	0.077

Table 3.2. Metabolic fluxes in Arabidopsis cells suspensions under continuous light or dark. The fluxes (mol/day/flask) estimated by fitting the biomass effluxes and MIDs of readout metabolites to the model (Fig. 3.8) are listed below.

Table. 3.2 is continued on next page.

Reaction name	Stoichiometry	Continuous dark				Continuous light			
		Net flux		Reversibility		Net flux		Reversibility	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Serine hydroxymethlytransferase reactions									
Ser	$T3P_p \rightarrow Ser$	0.007	0.000	0.373	0.229	0.010	0.000	0.357	0.230
Gly	$Ser \rightarrow Gly + CO_2$	0.005	0.000	0.726	0.078	0.006	0.000	0.674	0.084
Intercompartmental transport									
t3pt	$T3P_p \rightarrow T3P$	-0.350	0.072	0.438	0.127	-0.285	0.075	0.536	0.154
P5Pt	$P5P \rightarrow P5P_p$	-0.133	0.037	0.621	0.135	-0.061	0.048	0.875	0.083
pyrtp	$pyr \rightarrow pyr_p$	0.148	0.017	0.431	0.250	0.170	0.018	0.502	0.202
g6pt	$G6P \rightarrow G6P_p$	0.145	0.026	0.398	0.112	0.124	0.031	0.475	0.145
Biosynthetic effluxes									
T3Ppout	$T3P_p \rightarrow$	0.012	0.000			0.017	0.001		
T3Pout	$T3P \rightarrow$	0.002	0.000			0.001	0.000		
Serout	$Ser \rightarrow$	0.002	0.000			0.004	0.000		
glyout	$Gly \rightarrow$	0.005	0.000			0.006	0.000		
E4Ppout	$E4P_p \rightarrow$	0.006	0.000			0.009	0.000		
Pyrout	$pyr \rightarrow$	0.007	0.000			0.011	0.000		
Pyrpout	$pyr_p \rightarrow$	0.034	0.000			0.040	0.001		
ACApout	$ACA_p \rightarrow$	0.175	0.001			0.223	0.001		
OaAout	$OaA \rightarrow$	0.009	0.000			0.013	0.000		
αKGout	$\alpha KG \rightarrow$	0.018	0.001			0.030	0.003		
G6Pout	$G6P \rightarrow$	0.121	0.002			0.109	0.003		
G6Ppout	$G6P_p \rightarrow$	0.004	0.001			0.004	0.000		
P5Ppout	$P5P_p \rightarrow$	0.002	0.000			0.003	0.000		
Malout	$Mal \rightarrow$	0.003	0.000			0.003	0.000		
F6Pout	$F6P \rightarrow$	0.003	0.000			0.003	0.000		
CO2out	$CO_2 \rightarrow$	3.364	0.048			3.211	0.064		

Abbreviations: αKG, α-ketoglutarate; ACA, acetyl coenzyme A; E4P, erythrose-4-phophate; F6P, fructose-6-phosphate; Glc, glucose; G6P, glucose-6-phosphate; Gly, glycine; Mal, malic acid; OaA, oxaloacetic acid; P5P, pentose phosphates; Pyr, pyruvate; S7P, sedoheptulose-7-phosphate; Scn, succinate; Ser, serine; T3P, triose phosphates.

3.3. Discussion

Metabolic fluxes in heterotrophic Arabidopsis cell suspensions acclimated to continuous dark and light were quantified using MFA. The fluxes estimated in this work are very well identified in comparison to previous such studies (Williams et al.

2008b; Masakapalli et al. 2010) and this may be attributed to i) robustly designed ILEs employing [1-^{13}C], [1,2-^{13}C] and [U-^{13}C] glucose ii) fitting a large dataset (678 data points) that included mass isotopomers of critical readout metabolites such as ribose and iii) the high degree of agreement between the biological replicates. We found that the light treatments affect the biomass accumulation and protein production, but leave the carbon flux distribution in the primary carbon metabolism relatively unaltered.

3.3.1. Central carbon metabolism fluxes minimally affected by light treatments

Light is a major regulatory factor in plant cells. It regulates photosynthesis, development, flowering and growth (Eckardt 2007; Horvath 2009). How do light grown cells escape light's regulatory control? This can happen because the regulatory signals of light can be substituted by the regulatory signals of carbon (K. E. Thum et al. 2003; Krouk et al. 2009). The regulatory pathways of light and carbon are highly intertwined. Gutierrez and co-workers studied gene expressions in hydroponically grown Arabidopsis leaves and roots under combinations of stimuli by carbon, light and nitrogen. They found that external sugars and photosynthate sugars, both affect gene expressions similarly in roots and concluded that the heterotrophic roots sense light in the form of carbon (Krouk et al. 2009). Another study by Strasser et al., found that the arrested development of quintuple phytochrome mutants of Arabidopsis could be partially rescued in certain cases by introduction of sucrose in the growth media, which implies that carbon can substitute for light signaling (Strasser et al. 2010).

These evidences from previous literature suggest that, the similarity in the metabolic fluxes in cells between the two light treatments observed in this study may be because carbon signaling masks the signaling of light to render the metabolism of the cells identical under both treatments. Furthermore, in another study, Arellano and co-workers used Arabidopsis cell suspensions that harvested light energy, to study gene expressions when subjected to short periods of dark or high light (González-Pérez et al. 2011) and on mining their gene expression data (GEO accession #: GSE22671), we found that the expressions of the glycolytic and PPP genes do not differ between the two treatments (data not shown). We observe a similar response under the two treatments at the fluxomic level despite differences in the exposure times to dark and light.

Although counterintuitive, the differences in protein content can be attained with negligible changes in the central metabolic fluxes. Indeed, this has been observed in many MFA studies (Williams et al. 2008b). Because the precursors to the amino acids are acquired from several nodes of the primary metabolic pathways, significant changes in protein content can be achieved with insignificant changes in intracellular fluxes. For example, the protein content increases from 16.6% in the dark to 20.8% in light whereas the percentage of flux from plastidic pyruvate (a precursor to 4 amino acids) to protein only increases from 15% in the dark to 16% in the light. Therefore the growth and protein content differences between the two light treatments can be achieved with negligible changes in the intracellular metabolic fluxes.

3.3.2. Limitations and criticisms

The pooled data from multiple ILEs improves the accuracy of flux estimates (Libourel, Gehan, and Shachar-Hill 2007b; Nargund and Sriram 2012), but it may be difficult to fit data from multiple ILEs if the cells behavior differ. For instance, the glucose uptake rate of the cell suspension grown on [1,2-^{13}C] glucose was significantly lower than under [1-^{13}C] and [30% U-^{13}C] glucose ILEs. Consequently the constraints on the glucose uptake had to be relaxed to accommodate the difference and this may have increased the confidence intervals of the flux estimates.

The confidence intervals of flux estimates, when obtained by Monte Carlo simulations, are affected by the standard deviations of the isotopomer abundances. But since the χ^2 of fits are inversely proportional to standard deviations it is common practice to assign, on an ad hoc basis, higher standard deviations (1-5% in this study) than observed. The effect of the χ^2 values on the accuracy of flux estimation should perhaps be analyzed to enable better statistical treatment of the flux estimation problem.

The compartmental PPP flux estimates obtained in this study can be further improved by measuring the MIDs of additional metabolites from the cytosolic and plastidic compartments such as E4P, F6P and S7P. These metabolites can currently be measured by liquid chromatography – mass spectrometry (Huck et al. 2003) but their compartment of origin is difficult to determine. Use of extraction techniques that help obtain metabolite pools from specific subcellular compartments may solve the above problem.

3.4. Conclusion

This study advances the use of MFA in plant systems by implementing carefully designed ILEs on Arabidopsis cell suspensions under two light treatments. The improvements in flux estimates made in this study advocate designing ILEs for the specific pathways of interest and also measuring the MIDs of large number of varied metabolites. The growth and biomass composition differences between the two treatments suggest that light affects metabolism whereas the nearly identical metabolic flux maps suggest that light's role is limited only to providing energy via photophosphorylation. The regulatory signaling of light is ceded to that of carbon which is identical between the two light treatments and therefore leads to negligible changes in flux distributions. The fluxomic evidence of carbon signaling substituting light signaling in this study, corroborates previous gene expression studies (K. E. Thum et al. 2003; Krouk et al. 2009; Strasser et al. 2010; González-Pérez et al. 2011) to fortify our understanding of carbon and light signaling interactions in heterotrophic plant cells. The theme emerging from these systems level analyses is that light and carbon signaling are highly integrated, especially in heterotrophic plants cells. We believe that further systems level studies will shed more light on the regulatory networks of plant metabolism.

3.5. *Materials and methods*

3.5.1. Arabidopsis suspension cell cultures

Arabidopsis Columbia-0 suspension cell cultures were grown in 125ml Erlenmeyer flasks on an orbital shaker (New Brunswick 44R) at 125rpm under 24.5 °C and were subjected to continuous light or dark. The suspensions were subcultured every 7d by transferring 600mg of cells into 45ml of Murashige and Skoog media containing 3% (w/v) glucose, 1 mg/l 1-naphthaleneacetic acid and 0.05 mg/l kinetin. Three parallel ILEs with 100% 1-^{13}C, 100% 1,2-^{13}C and 30% U-^{13}C Glucose (Cambridge Isotopes) were conducted under both the treatments. The difference in the glucose concentration between the fresh and spent media helped estimate the glucose uptake rates.

3.5.2. Extraction and quantification of biomass components for estimation of biomass effluxes

Arabidopsis cells, harvested by vacuum filtration through glass microfiber filter paper, were immediately frozen in liquid nitrogen to arrest metabolism. The samples were then freeze-dried and stored at -80 °C until further analysis.

3.5.2.1. Protein quantification

20 mg of ground, freeze dried cells were contacted with 600 µl of phosphate buffer saline for 15 min on ice, centrifuged and the supernatant containing the protein collected. The above steps were repeated twice. The supernatants from each contact were pooled and the protein content quantified by a Bradford assay (Biorad, Hercules,

CA).

3.5.2.2. RNA quantification

RNA was extracted using the RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) and quantified spectrophotometrically by measuring the absorbance at the wavelength of 260 nm.

3.5.2.3. Soluble metabolite quantification

Soluble metabolites were extracted and quantified as described by Fiehn et al. (Fiehn et al. 2000). Briefly, 70 mg of ground, freeze dried cells were contacted with 1.4 ml methanol and 50 µl water, heated at 75 °C for 15 min, centrifuged and the supernatant containing the soluble metabolites collected. The above steps were repeated once. The supernatant was dried, redissolved in 100 µl solution of methoxyamine hydrochloride in pyridine (20mg/ml) and heated at 30 °C for 90 min. Next the soluble metabolites were derivatized by adding 100 µl N-(tert-butyldimethylsilyl)-N-methyl trifluoroacetamide (MTBSTFA) and heating at 70 °C for 1 h. Soluble metabolites were quantified by GC with Norleucine as an internal standard.

3.5.2.4. Lipid quantification

On extraction of soluble metabolites, the cells were dried, contacted with 1ml hexane, subjected to 40 °C for 30 min, centrifuged and supernatant collected. The above process was repeated twice. The supernatants were pooled, dried, redissolved in 1ml methanolic HCl (3N) and heated at 70 °C for 1h to produce fatty acid methyl esters (FAME). On cooling, the methanolic HCl was contacted with hexane three times. The

FAMEs in the hexane layer were quantified by GC with octanoic acid as the internal standard.

3.5.2.5. Starch quantification

Starch was extracted as described in (Foster, Martin, and Pauly 2010) .The cell biomass, stripped of lipids and soluble metabolites, was dried, redissolved in 1.5 ml of 0.1 M sodium acetate buffer (pH 5.0), heated at 70 °C for 20 min and cooled on ice. 35 µl of 0.01% sodium azide, 35 µl of amylase (50 µg/ml H_2O) and 18.7 units of pullulanase were added to the solution and it was incubated overnight at 37 °C to hydrolyze the starch to glucose. The supernatant was collected and the glucose concentration in it measured by a glucose analyzer (YSI 2700).

3.5.2.6. Cell wall quantification

The cell biomass remaining on extraction of all the above components contains cell wall and mineral ash predominantly.

3.5.3. Growth rates

The Arabidopsis cells were subcultured and 3 biological replicates were harvested every day for seven days by vacuum filtration through glass microfiber filter paper (Whatman). The cells were weighed before and after freeze drying overnight. The biomass effluxes were estimated using the biomass compositions and growth rates.

3.5.4. Extraction and measurement of isotopomer abundances of biomass components by GC-MS

A Varian 450-GC in line with a Varian 300 MS was used to measure mass isotopomer distributions of various biomass components. The GC had aVF-5ms column (30 m x 0.25 mm x 0.25 µm; Varian, Inc) with helium as the carrier gas. The MS was acquired in electron ionization mode with the ion source at 280 °C. The Varian MS workstation (version 6.9.3) software was used in tandem with the NIST mass spectral library (National Institute of Standards and Technology, Gaithersburg, MD) to identify and quantify the MS data. An in-house MATALB program was used to correct the mass isotopomer distributions (MIDs) for the naturally abundant isotopes of non metabolic C, H, N, O, S and Si.

3.5.4.1. Amino acids and soluble metabolites

20 mg of ground, freeze dried cells were vacuum hydrolyzed with 6N HCl for 5 h at 160 °C in hydrolysis tubes (Pierce Endogen, Rockford, IL). The hydrolysate was dried overnight in a RapidVap (Labconco, Kansas City, MO), reconstituted in 1 ml water, filtered, freeze dried and stored in -80 °C until further analysis. Soluble metabolites were extracted as mentioned above. To improve volatility in the GC, the amino acid and soluble metabolites were derivatized with 100 µl of N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide (MTBSTFA) in 100 µl dimethylformamide by heating at 70 °C for 1 h. A sample volume of 1 µl and split ratio 50 was injected into the GC column with the carrier gas flow rate at 1.0 ml/min. The oven temperature started at 150 °C for 2 min, ramped to 230 °C at 3 °C /min, to

240 °C at 2 °C/min and then to 275 °C at 10 °C/min where it was held constant for 6 min.

3.5.4.2. Glucose from soluble metabolites

The glucose in the soluble metabolites was extracted as described above. To improve volatility in the GC, the sample was dried and derivatized by heating with 50 µl hydroxylamine hydrochloride (20mg/ml in pyridine) at 90 °C for 1h and then on adding 100 µl propionic anhydride at 60 °C for 30 min. It was dried under a stream of N_2 and reconstituted in 100 µl ethyl acetate. A sample volume of 1 µl at split ratio 75 was injected into the GC column with the carrier gas flow rate at 0.9 ml/min. The oven temperature was held at 180 °C for 2 min and ramped up to 300 °C at 5 °C/min and held constant for 4 min.

3.5.4.3. Ribose from RNA

The RNA, extracted as described above, was hydrolyzed to ribose with 1 ml of 2N HCl at 100 °C for 2. To improve volatility in the GC, ribose was derivatized with the same procedure as glucose. A sample volume of 9 µl at split ratio 10 was injected into the GC column with the carrier gas flow rate at 1.5 ml/min. The oven temperature was held at 180 °C for 2 min and ramped up to 300 °C at 20 °C/min and held constant for 2 min.

3.5.5. Extraction and measurement of isotopomer abundances of biomass components by NMR

Protein was extracted from 60 mg of dried, ground cells as described above. Protein

extracted in phosphate buffer saline was dialyzed to eliminate the phosphate salts as they affect NMR acquisition. The protein was then freeze dried and reconstituted in 500 µl D_2O. The 2-D [^{13}C, ^{1}H] HSQC spectra was acquired with the following parameters at 25 °C: ^{13}C (F1) resonance frequency, 150 MHz; ^{1}H (F2) resonance frequency, 600 MHz; spectral width along ^{13}C (F1) dimension, 6039 Hz; spectral width along ^{1}H (F2) dimension, 8371 Hz; number of complex data points, 4096 (^{13}C) × 1024 (^{1}H); number of scans, 4. The software Bruker Topspin 2.1 was used to program the acquisition, view and phase the spectra. The software NMRviewJ (One Moon Scientific, Inc.; available free of charge at http://www.onemoonscientific.com) was used to extract the peaks and an in-house MATLAB package NMRisotopomer (to be made freely available at http://openwetware.org/wiki/Sriram_Lab) was used to quantify the areas under the peaks. The ^{13}C enrichments of amino acids needed for computing isotopomer abundances were calculated from their respective MIDs.

3.6. Acknowledgements

We thank Prof. Diane Bassham (Department of Genetics, Development, and Cell Biology, Iowa State University) for the Arabidopsis cell suspension cultures and Prof. Jacqueline V. Shanks (Department of Chemical and Biological Engineering, Iowa State University) for a gift of ^{13}C labeled glucose. This work was funded by the University of Maryland, Department of Chemical and Biomolecular Engineering faculty startup grant and the Minta Martin Foundation, University of Maryland.

4. Chapter 4: Elucidating carbon – nitrogen interactions in poplar suspension cells by metabolic flux analysis

Ashish Misra[1], Shilpa Nargund[1], Xiaofeng Zhang[1], Gary D. Coleman[2] and Ganesh Sriram[1]*

[1]Department of Chemical and Biomolecular Engineering, University of Maryland, College Park, MD 20742, USA

[2]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD 20742

1st draft of manuscripts to be submitted to the journal *Metabolic Engineering* and *Plant Physiology*

Author contributions

GS, GC, AM and SN conceived this study. GC developed the poplar cell suspension cultures. SN, AM and XZ performed the experiments and computations under the direction of GS. SN wrote the 1st draft and GS edited it.

4.1. Introduction

Plants produce a host of primary and secondary metabolites that are of prime economic interest. C and N metabolism are intimately intertwined and to a large extent dictate the fate of the nutrients entering plants (Sakakibara, Takei, and Hirose 2006). The availability of N affects not only photosynthesis (CO_2 fixation) but also photorespiration (undesirable O_2 fixation) and respiration which are important cellular functions in plants (Nunes-Nesi, Fernie, and Stitt 2010). The balance between C and N also affects the partitioning of biomass into sugars, starch and organic acids, flowering time and root architecture (Sakakibara, Takei, and Hirose 2006; Gutierrez et al. 2007). The interactions between C and N themselves respond to a multitude of factors: light and water availability, adenosine triphosphate (ATP), reducing factors (required for assimilation of inorganic N), and pH balance and consequently they are regulated by various signaling molecules such as nitrate, ammonium, sugars, amino acids (especially glutamine, glutamate and aspartate), and organic acids (Nunes-Nesi, Fernie, and Stitt 2010). Understanding the intricacies of C-N interactions will help us modulate plants for desired purposes.

Poplar is a potential cellulosic biofuel source and hence there is tremendous interest in understanding N cycling and C-N interaction in this tree. Woody perennial trees have evolved mechanisms to recycle N considerably from senescing leaves during fall and winter (Zhu and Coleman 2001; Cantón, Suárez, and Cánovas 2005). Uptake and assimilation of N needs energy in the form of ATP and reducing power (Nunes-Nesi, Fernie, and Stitt 2010). N recycling results in more efficient use of N saving the

tree precious resources. Studying the existing metabolic machinery that helps recycle N and genetically eliminating possible bottle necks in this process will help trees utilize N more efficiently and make them better biomaterial crops. Poplars are capable of recycling up to 80% of the N in leaves to perennial tissues during leaf senescence (Pregitzer et al. 1990). Poplars produce bark storage proteins (BSPs) which serve as nitrogen sinks when the N from senescing leaves is remobilized. Probing the intracellular fluxes in poplar cell suspensions under various N conditions will help understand the N cycling and C-N interactions.

4.2. Results

To study the C-N interactions in poplar, we subjected poplar cell suspensions to four different C-N supply treatments (Fig. 4.1) – optimal C, N (C++,N++++++), optimal C, low N (C++, N+), low C, optimal N (C+, N++++++), and low C, low N (C+, N+). ILEs were carried out on all these treatments.

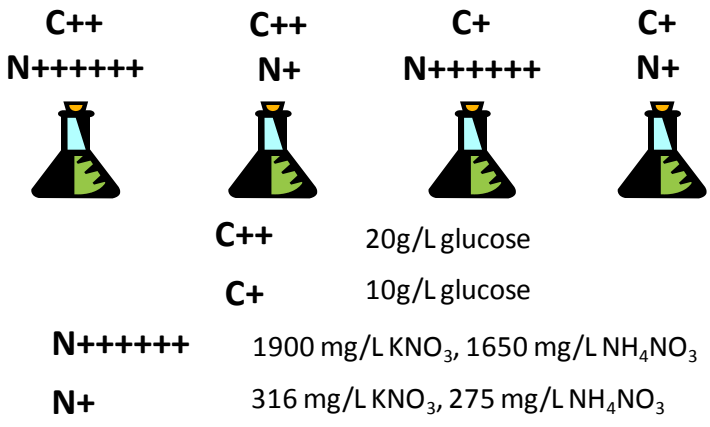


Fig. 4.1. Poplar cell suspensions acclimated to different C-N supply treatments. The different C-N supply treatments allow us to study the C-N interactions especially under C or

N deficient conditions. The treatments were chosen by trial and error.

4.2.1. Growth and nutrient uptake rates differ between the C-N supply treatments

The change in biomass and consumption of glucose and nitrogen sources were measured for 7 days after subculture. The cell suspensions under the different treatments show different growth characteristics. The cell suspensions under C++, N++++++ treatment show the highest growth rate whereas those under the N+ treatments grow slower (Fig. 4.2). The biomass yield/mg C uptake (Fig. 4.3) for all treatments is similar suggesting that irrespective of the status of C or N, the uptake of C is constant. Contrary to this, the biomass yield/mg N uptake is higher for the low N treatments (Fig. 4.4).

4.2.2. The ^{13}C label is diluted by the initial seed biomass in batch cultures

We conducted ILEs with 100% 1-^{13}C, and 30% U-^{13}C glucose on all four C-N treatments. The ^{13}C enrichments of proteinogenic amino acids extracted from cells grown on 27% U-^{13}C glucose were significantly lower than the expected 27% (between 22-28%) (Fig.4.5). This is a surprising result because the cells cannot distinguish between ^{12}C and ^{13}C isotopes (Kruger et al. 2007b). To explain this, we hypothesized that the dilution could be due to one of the following reasons – i) photosynthetic fixation of unlabeled CO_2 or ii) initial seed biomass. To test photosynthetic fixation of CO_2, the cells were grown on 100% U-^{13}C glucose under continuous light and dark.

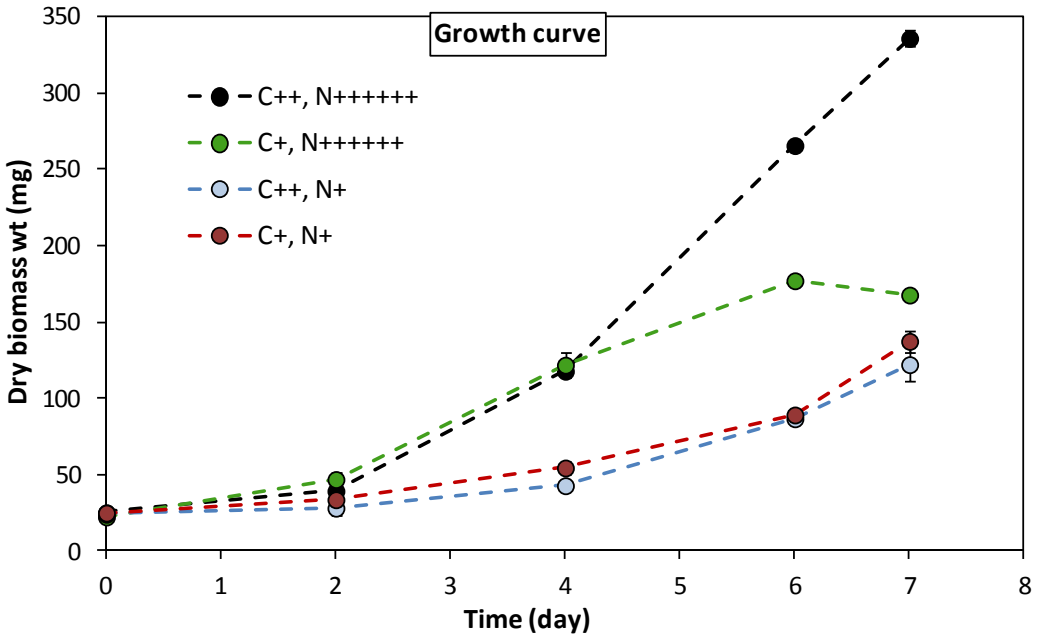


Fig. 4.2 Growth rates between cells under the four treatments are different.

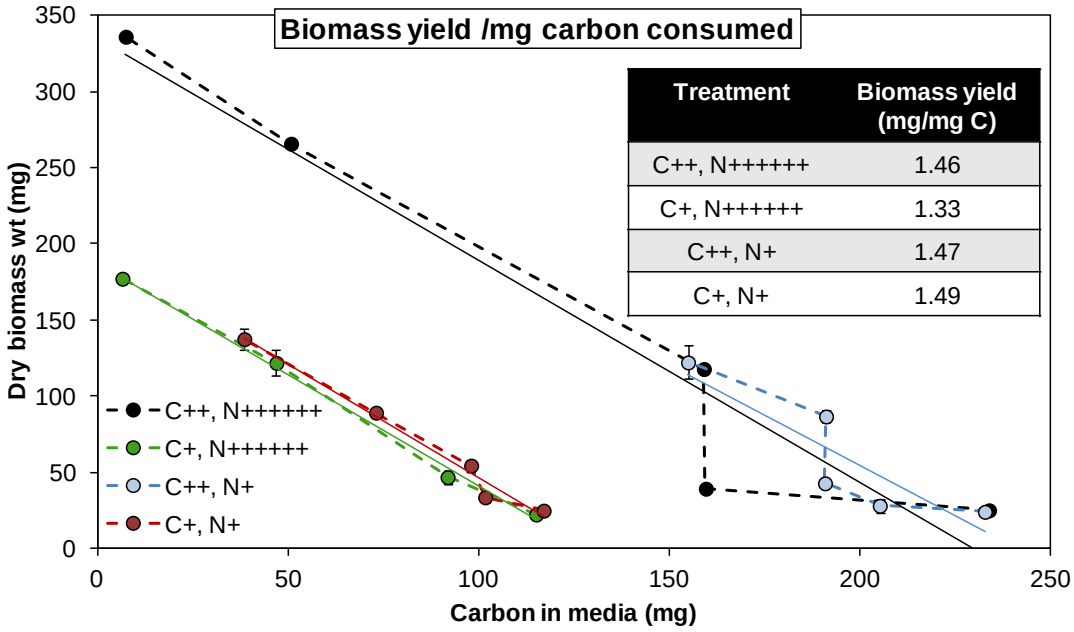


Fig. 4.3 Biomass yield per mg carbon consumed is similar across all C-N treatments.

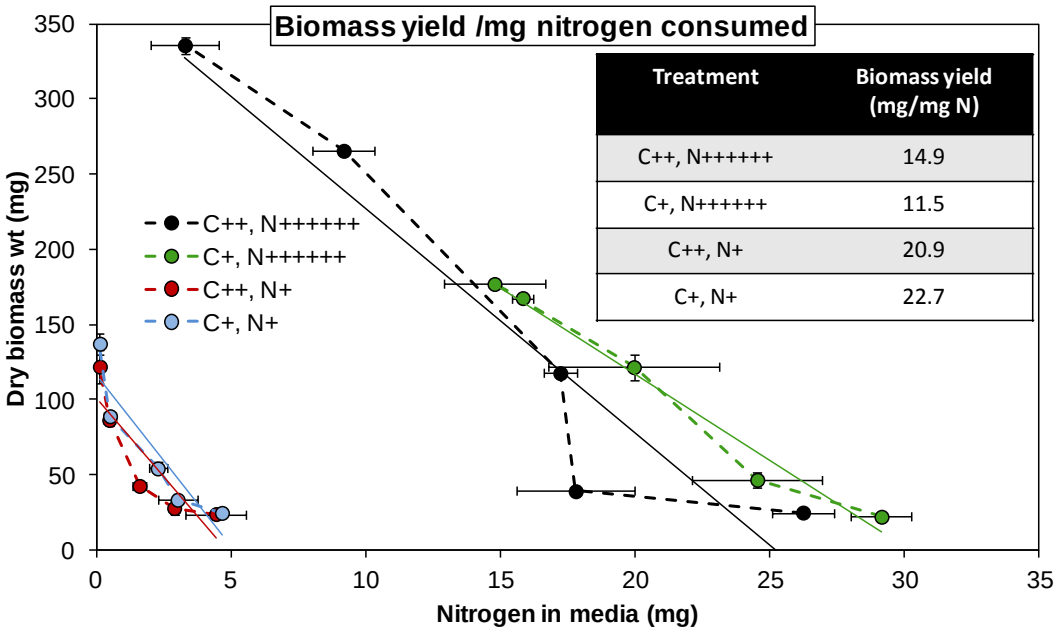


Fig. 4.4 Biomass yield per mg N consumed is higher in cells grown under low N.

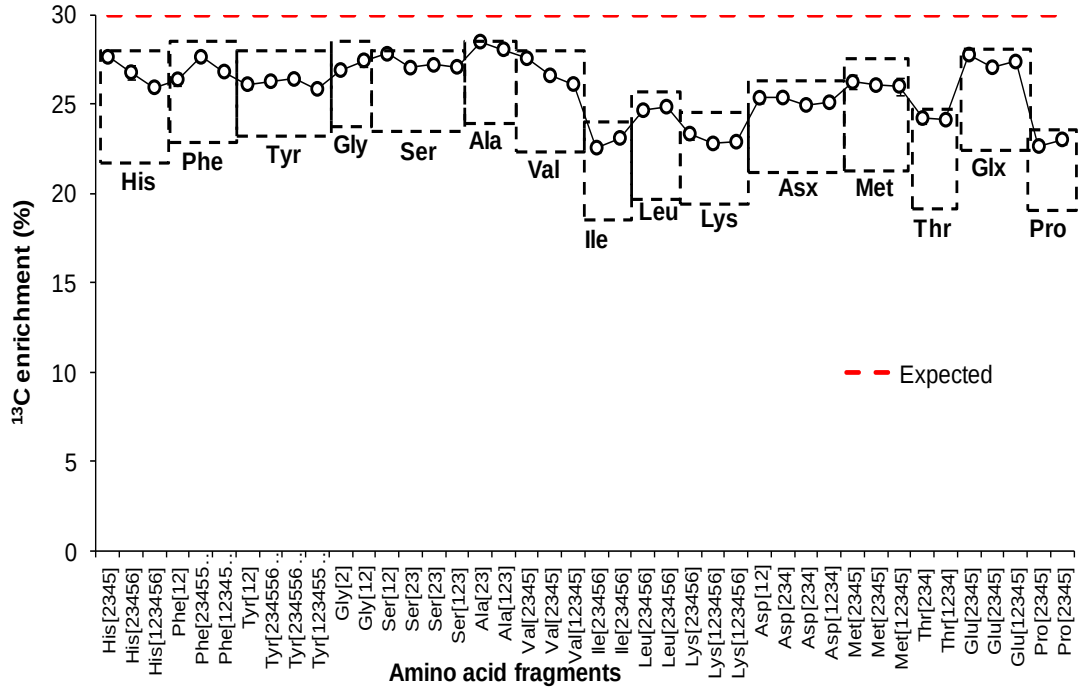


Fig. 4.5. ^{13}C enrichments of amino acids from cells grown on 30% U-^{13}C glucose are lower than the expected 30%. Dilution of the ^{13}C label observed is different between the various amino acids.

Absence of dilution of the ^{13}C label in the dark grown cell metabolites would implicate the role of photosynthetic CO_2 fixation. Similar extents of dilutions of ^{13}C label in amino acids were observed under both light and dark (Fig. 4.6) suggesting that photosynthetic fixation of CO_2 is not responsible for the dilution. Evidence of anaplerotic fixation of CO_2 (data not shown) was detected but this alone could not explain the extensive dilution observed.

The poplar cell suspensions are grown in batch cultures. Every week aliquots of the culture, with known mass of cells, are transferred to new shake flasks with fresh media. This seed biomass takes up the fresh media to synthesize new biomass. During an ILE, the naturally labeled (1.13% ^{13}C) seed biomass is transferred to labeled media (containing 30% U-^{13}C or 100% U-^{13}C or 100% 1-^{13}C glucose) and it contributes a substantial proportion (10-12%) to the total biomass. This can readily explain the extensive dilution of ^{13}C label that occurs in amino acids.

4.2.3. Seed biomass backmixes with the newly synthesized biomass

In a 100% U-^{13}C glucose ILE, we expect to obtain only unlabeled (from the seed biomass) and fully ^{13}C labeled (from the newly synthesized biomass) metabolites i.e. the mass spectra of metabolites should show only the smallest and the largest mass isotopomers. Contrary to this, several intermediate mass isotopomers were detected in all amino acid mass isotopomer distributions (MIDs) (Fig. 4.7). These can occur only if the seed biomass backmixes with the newly synthesized biomass to produce combinations of ^{12}C and ^{13}C.

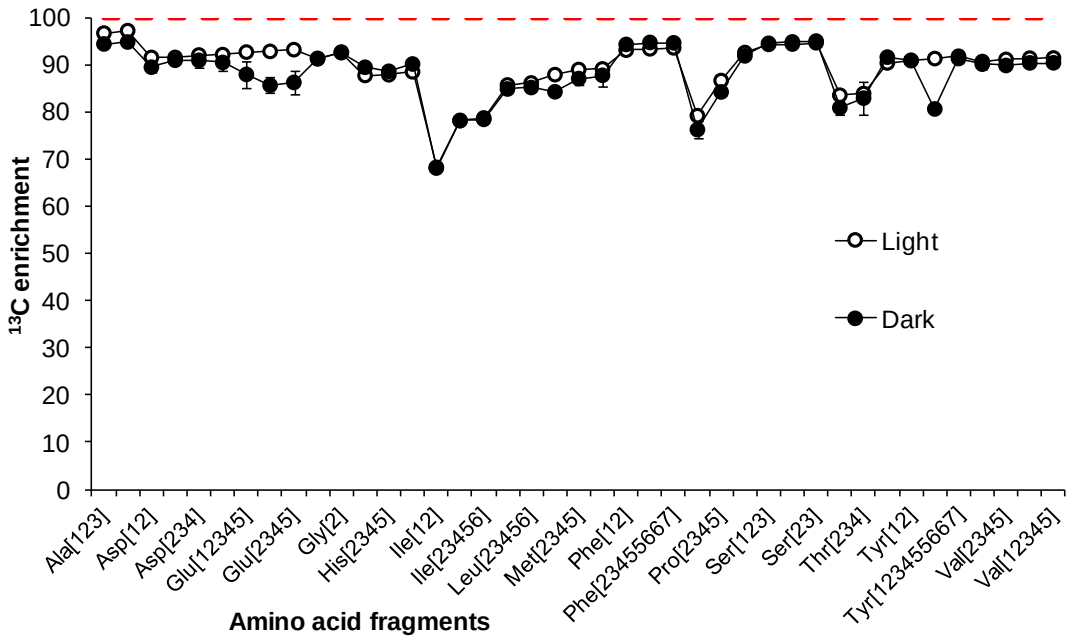


Fig. 4.6. The dilution of ^{13}C enrichments of amino acids is not due to photosynthetic fixation of unlabeled CO_2. ILEs with 100% U-^{13}C glucose were conducted in light and dark. Since photosynthesis can occur only in light, the similar dilution of ^{13}C label under both light treatments suggests that photosynthetic fixation of unlabeled CO_2 is not the cause of dilution.

Backmixing also explains the differences in the ^{13}C enrichments among amino acids.We hypothesize that the metabolites, that show signatures of backmixing, are biosynthesized from the labeled media as well as from the seed culture by degradation pathways. Since the metabolites undergo several cycles of biosynthesis and degradation, the ^{12}C and ^{13}C combinations are formed.

4.2.4. Modeling the backmixing of seed biomass with newly synthesized biomass

The metabolic network model consists of the glycolysis and the PPP both in the

cytosol and the plastidic compartments and the tricarboxylic acid cycle (TCA) in the cytosol. Glucose is the carbon source and the various biomass components leave the system.

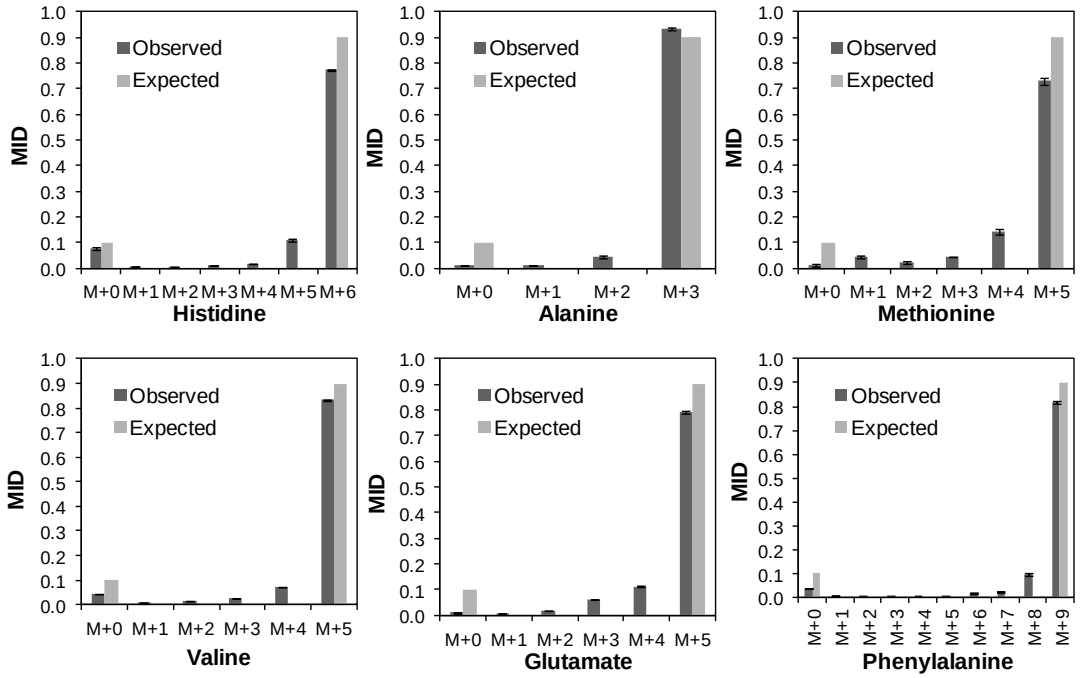


Fig. 4.7. Mass isotopomer distributions of metabolites from a 100% U-^{13}C glucose ILE. Presence of mass isotopomers between the lowest and highest mass isotopomers indicate that backmixing between the seed biomass and newly synthesized biomass occurs.

In this network, to depict the dilution the amino acids that show signatures of backmixing (the 100% U-^{13}C glucose ILE provides the most prominent signatures) enter the metabolic network as unlabeled metabolites (Fig. 4.8). The unlabeled amino acid influxes are constrained to 12% of their biomass effluxes (calculated using the biomass composition) since the seed biomass contributes up to 12% of the total biomass. Biosynthesis and degradation pathways model the backmixing by allowing

the unlabeled amino acids to mix with their biosynthetic labeled versions and degrade and reform to attain combinations of ^{12}C and ^{13}C. Except His, Leu, Ile, Lys, Pro and Thr all other amino acids showed signatures of backmixing. The amino acid influxes and the parameters that belong to the biosynthesis and degradation reactions are called backmixing parameters.

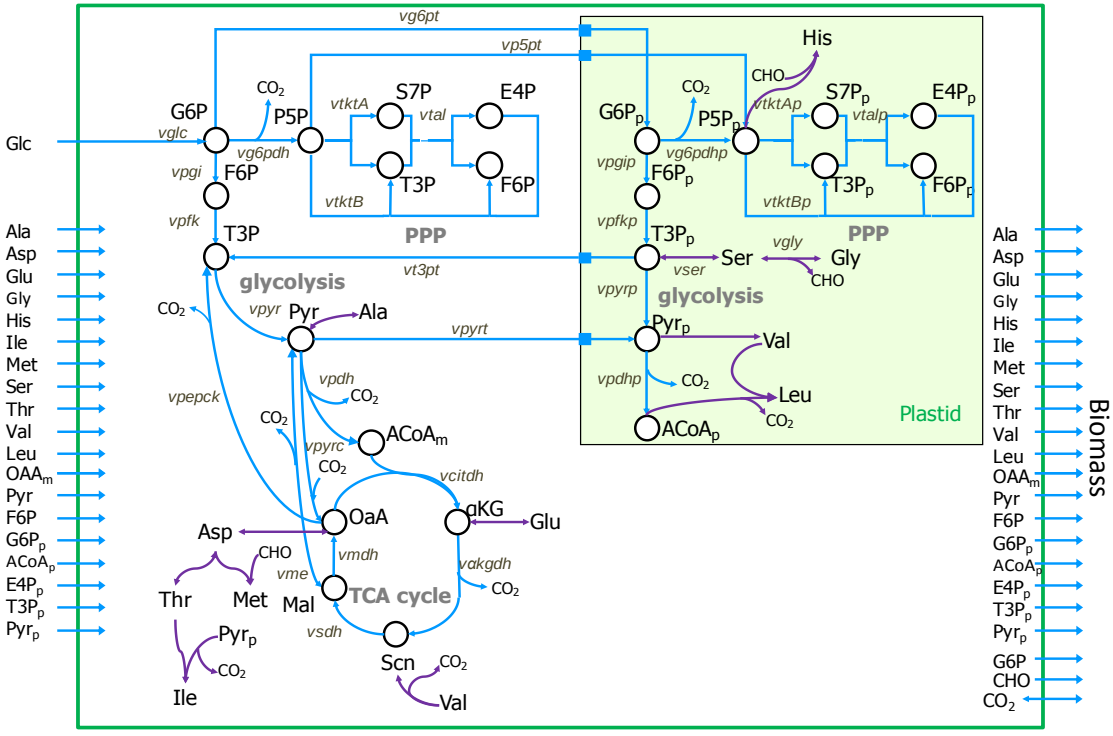


Fig. 4.8. Backmixing model. In this model the metabolites that show signatures of backmixing enter the metabolic networks along with labeled glucose to model the dilution. Additionally the biosynthetic and degradation pathways help model the backmixing by allowing the metabolites to be synthesized by labeled glucose and from their incoming streams.

4.2.5. The backmixing model performs better than previously used

models

Previously, investigators have observed similar dilution effects in cells suspensions/embryos due to the presence of seed biomass/embryo material (Lonien and Schwender 2009b) and have accounted for it by i) adjusting the labeling of the feed substrate (Glc labeling adjusted model) or ii) correcting the MIDs by a factor based on the seed biomass/total biomass (MIDs adjusted model). In this study, we compare the above models to the backmixing model described above by fitting the MIDs obtained from the 1-^{13}C, 30% U-^{13}C and 100% U-^{13}C ILEs simultaneously.

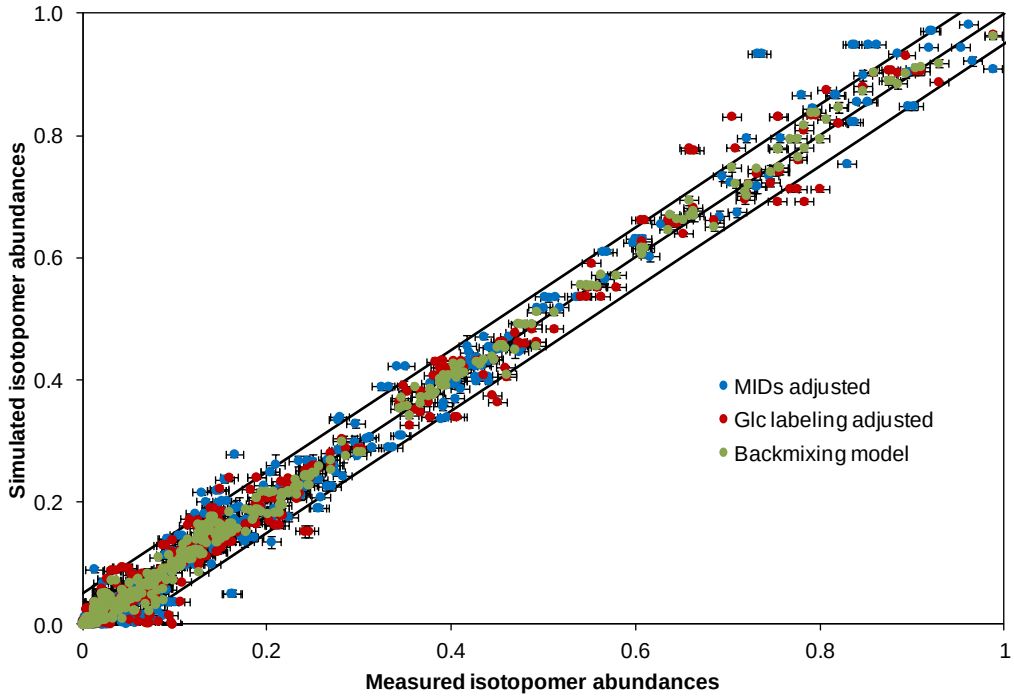


Fig. 4.9. The fluxes estimated by the backmixing model explain the measured MIDs the best. The isotopomer abundances simulated by the backmixing model match the measured isotopomer abundances better than the other two models and consistently lie within the 5% error boundaries.

The MIDs from the 1-^{13}C and 30% U-^{13}C glucose ILEs constrain the intracellular fluxes and the backmixing fluxes whereas those from the 100% U-^{13}C glucose ILE constrain only the backmixing fluxes. The backmixing model explains the isotopomer abundances to a greater extent than the other models. The measured data is matched best by the simulated data obtained from the backmixing model (Fig. 4.9).

A preliminary flux map produced using the backmixing model shows that the significant improvement of the fits is achieved with very little flux through the backmixing reactions (Fig. 4.10).

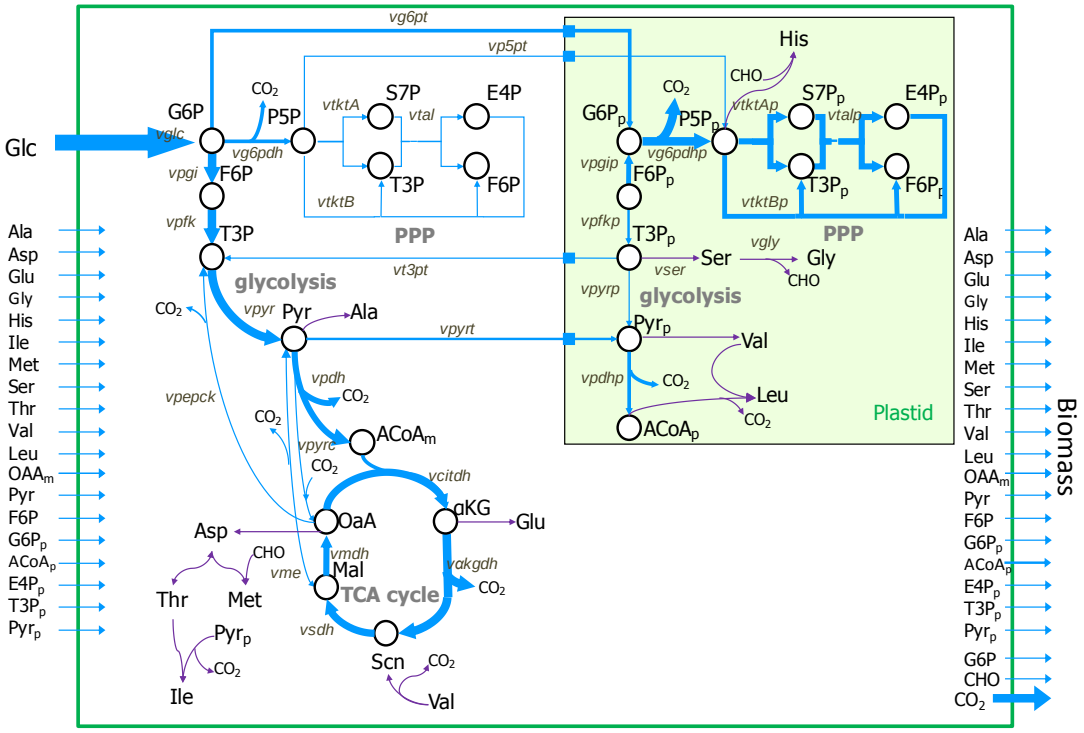


Fig. 4.10. Metabolic flux map created using flux estimates obtained from the backmixing model

4.2.6. ILEs carried out for three subculture cycles reduce the dilution of ^{13}C label by seed biomass

To eliminate the unlabeled seed biomass altogether, we conducted ILEs with 100% 1-^{13}C, 30% U-^{13}C and 100% U-^{13}C glucose for three subculture cycles instead of one. Despite, the longer labeling duration, the ^{13}C enrichments of metabolites, although higher, are not all enriched up to the expected values (Fig. 4.11). Therefore we will be using the backmixing model to fit the labeling data from all poplar cell suspension ILEs.

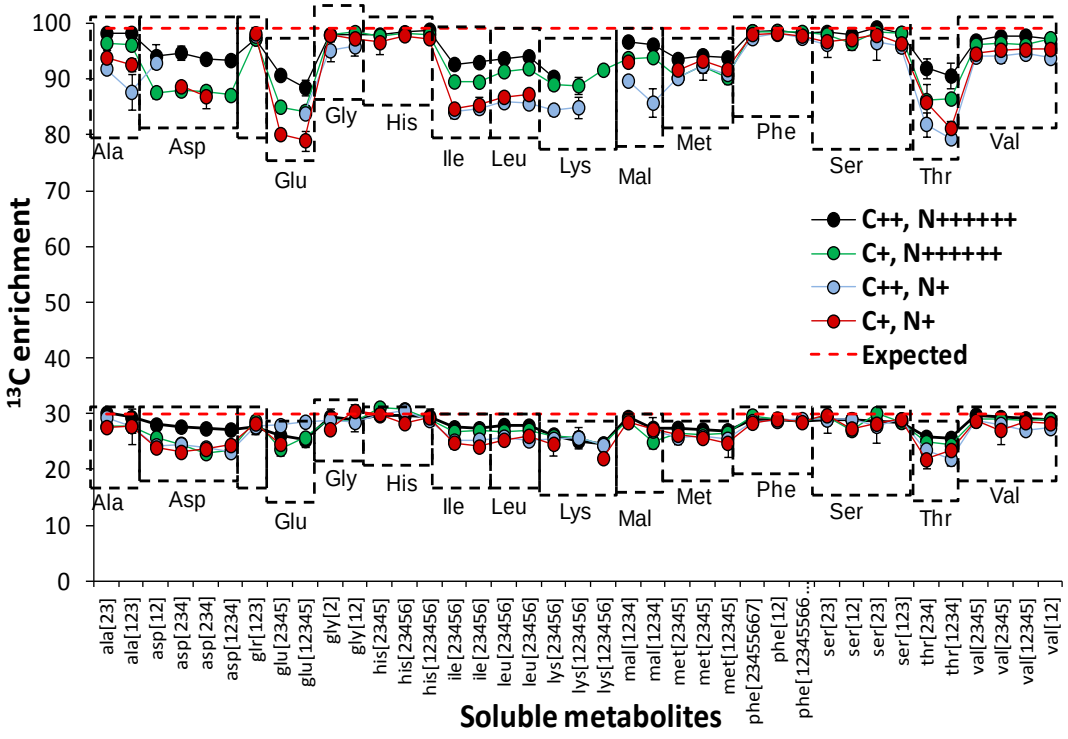


Fig. 4.11. Conducting ILEs for three subculture cycles does not eliminate the unlabeled seed biomass completely. The metabolites from cells grown on 30% U-^{13}C or 100% U-^{13}C glucose for three subculture cycles do not attain the expected ^{13}C enrichments suggesting that the unlabled seed biomass is not completely eliminated

4.3. Discussion

The growth and nutrient uptake characteristics between the C-N treatments indicate that the decrease in the C supply has negligible effect on metabolism whereas the decrease in N supply does. The cells grown under low N exhibit higher biomass accumulation per mg N consumed thus suggesting that the cells are making fewer N containing metabolites. Flux maps will help us understand the pathways regulation under different C-N treatments.

Flux estimates are obtained by carefully balancing the ^{12}C and ^{13}C atoms in the metabolites of the metabolic network. Therefore it is paramount to identify all sources of ^{12}C and ^{13}C for accurate book-keeping. This is especially important in plant systems because MFA is often carried out on batch cultures or embryos. The effect of label dilution by seed biomass material on flux estimates has not been examined before. We found that we have to account not only for the dilution sources but also the backmixing of the seed biomass with the newly synthesized biomass. The backmixing model proposed performs better than previous models used to explain dilution effects. The backmixing model was made primarily based on the MIDs of metabolites obtained under the 100% U-^{13}C Glc ILE. Therefore ILEs with 100% U-13C Glc should be conducted to determine the backmixing extents in the various metabolites. Significant improvements were achieved in the fits with only small fluxes through the backmixing reactions. This shows the potency of careful modeling of the metabolic network.

4.4. Material and methods

4.4.1. Poplar suspension cell cultures

Poplar (*Populus deltoides*) suspension cell cultures were grown in 125ml Erlenmeyer flasks on an orbital shaker at 125rpm under 21 °C. The suspensions were subcultured every 7d by transferring 600mg of cells into 30ml of Murashige and Skoog media containing 2% (w/v) glucose, 0.1% of 1mg/ml vitamins, 0.1% of 1mg/ml 2,4-Dicholorophenoxyacetic acid, 0.01% of 1mg/ml 1-naphthaleneacetic acid, 0.001% of 1mg/ml 6-Benzylaminopurine and 0.0008% BASTA at pH 5.7. Three parallel ILEs with 100% 1-^{13}C, 100% 1,2-^{13}C and 30% U-^{13}C Glucose (Cambridge Isotopes) were conducted under all C-N supply treatments.

The other methods used in this study are similar to those mentioned in Section 3.5.

5. Chapter 5: Conclusions

The fluxome is the phenotypic culmination of interactions between all other components such as the genome, transcriptome, proteome and metabolome. Fluxomics is an important 'omics' field that has not yet been fully exploited in plant studies. We have developed and engaged fluxomic tools, especially metabolic flux analysis (MFA), to probe metabolism in two plant cell suspensions – *Arabidopsis thaliana* (Arabidopsis), a model plant and poplar, a model tree and potential biofuel crop. Although flux maps have been generated for Arabidopsis prior to this study, we have generated the first flux map for poplar cell suspensions.

The isotope labeling experiment (ILE) design study carried out in this work (Chapter 2) successfully identified isotopic labels (1,2-^{13}C glucose) and information rich metabolites (ribose) that helped estimate the fluxes in the pentose phosphate pathways (PPP) in the Arabidopsis cell suspensions with much greater accuracy than reported in literature previously (Chapter 3). We showed that use of multiple parallel ILEs improves the flux estimates further and that in the case of PPP flux estimation, the measurement of isotopomers by mass spectrometry measurement outperforms that by nuclear magnetic resonance.

The flux maps obtained by carrying out MFA on Arabidopsis cell suspensions grown under continuous light and dark showed that the light and dark grown cells are regulated only by carbon signaling and not light signaling. This study provides fluxomic evidence that light and carbon signaling are highly intertwined and adds to other systems level studies that have found similar responses in plant cells under light

and carbon stimuli (Chapter 3).

MFA was conducted on poplar that is known to perform distinctive nitrogen storage and recycling. Poplar cell suspensions acclimated to different carbon – nitrogen (C-N) supply treatments displayed different growth rates and biomass yield/ mg nitrogen uptake but very similar biomass yield/ mg carbon uptake. This suggests that cells can direct nitrogen allocation with more flexibility than carbon allocation. Poplar cell suspensions, grown in batch cultures, demonstrated via ILEs, that the seed biomass dilutes ^{13}C label and affects the mass isotopomer distributions of metabolites to a great extent. This phenomenon was also observed in Arabidopsis cell suspensions but to a lesser extent. The ILEs carried out for three subculture cycles instead of one were unable to eliminate the seed biomass completely. The dilution and accompanying backmixing of the seed biomass and newly synthesized biomass was successfully modeled to obtain satisfactory fits of the mass isotopomer distributions of metabolites (Chapter 4).

The work done in this study advances the use of MFA in studying plant physiology by enhancing both computational and experimental approaches in flux estimation. The insights obtained from this work, at once showcase the robustness of the MFA technique and the challenges in improving it further for application to plant systems.

Future directions

Metabolic flux maps of poplar cell suspensions under the different C-N supply treatments will shed new light on the mechanisms by which nitrogen metabolism is tuned. Integration of information from the flux maps with information obtained by

proteomic and transcriptomic studies on poplar, that are being carried out in our collaborator, Dr. Gary Coleman's lab, can potentially help discover more details on the nitrogen storage and recycling mechanisms of poplars.

Bibliography

Allen, Doug K, Russell W Laclair, John B Ohlrogge, and Yair Shachar-hill. 2012. “Isotope Labelling of Rubisco Subunits Provides in Vivo Information on Subcellular Biosynthesis and Exchange of Amino Acids Between Compartments.” *Plant, Cell & Environment* (February 24). doi:10.1111/j.1365-3040.2012.02485.x. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-3040.2012.02485.x/abstract.

Allen, Doug K, Igor GL Libourel, and Yair Shachar-Hill. 2009. “Metabolic Flux Analysis in Plants: Coping with Complexity.” *Plant, Cell & Environment* 9999 (9999). doi:10.1111/j.1365-3040.2009.01992.x. http://dx.doi.org/10.1111/j.1365-3040.2009.01992.x.

Allen, Doug K., John B. Ohlrogge, and Yair Shachar-Hill. 2009. “The Role of Light in Soybean Seed Filling Metabolism.” *The Plant Journal* 58 (2): 220–234. doi:10.1111/j.1365-313X.2008.03771.x.

Alonso, Ana P, Dale L Val, and Yair Shachar-Hill. 2011. “Central Metabolic Fluxes in the Endosperm of Developing Maize Seeds and Their Implications for Metabolic Engineering.” *Metabolic Engineering* 13 (1): 96–107.

Alonso, Ana P., Fernando D. Goffman, John B. Ohlrogge, and Yair Shachar-Hill. 2007. “Carbon Conversion Efficiency and Central Metabolic Fluxes in Developing Sunflower (Helianthus Annuus L.) Embryos.” *The Plant Journal* 52 (2) (October): 296–308. doi:10.1111/j.1365-313X.2007.03235.x.

Antoniewicz, Maciek R., Joanne K. Kelleher, and Gregory Stephanopoulos. 2006a. “Determination of Confidence Intervals of Metabolic Fluxes Estimated from Stable Isotope Measurements.” *Metabolic Engineering* 8 (4) (July): 324–337. doi:10.1016/j.ymben.2006.01.004.

———. 2006b. “Determination of Confidence Intervals of Metabolic Fluxes Estimated from Stable Isotope Measurements.” *Metabolic Engineering* 8 (4) (July): 324–337. doi:10.1016/j.ymben.2006.01.004.

———. 2007a. “Elementary Metabolite Units (EMU): A Novel Framework for Modeling Isotopic Distributions.” *Metabolic Engineering* 9 (1) (January): 68–86. doi:10.1016/j.ymben.2006.09.001.

———. 2007b. “Accurate Assessment of Amino Acid Mass Isotopomer Distributions for Metabolic Flux Analysis.” *Analytical Chemistry* 79 (19) (October 1): 7554–7559. doi:10.1021/ac0708893.

Araúzo-Bravo, Marcos J, and Kazuyuki Shimizu. 2003. “An Improved Method for Statistical Analysis of Metabolic Flux Analysis Using Isotopomer Mapping Matrices with Analytical Expressions.” *Journal of Biotechnology* 105 (1-2) (October 9): 117–33.

Baxter, Charles J., Henning Redestig, Nicolas Schauer, Dirk Repsilber, Kiran R. Patil, Jens Nielsen, Joachim Selbig, Junli Liu, Alisdair R. Fernie, and Lee J. Sweetlove. 2007. “The Metabolic Response of Heterotrophic Arabidopsis Cells to Oxidative Stress.” *Plant Physiology* 143 (1) (January 1): 312–325. doi:10.1104/pp.106.090431.

Berger, Susanne, Erin Bell, Avi Sadka, and John E. Mullet. 1995. “Arabidopsis Thaliana Atvsp Is Homologous to Soybean VspA and VspB, Genes Encoding Vegetative Storage Protein Acid Phosphatases, and Is Regulated Similarly by Methyl Jasmonate, Wounding, Sugars, Light and Phosphate.” *Plant Molecular Biology* 27 (5) (March): 933–942. doi:10.1007/BF00037021.

Bläsing, Oliver E., Yves Gibon, Manuela Günther, Melanie Höhne, Rosa Morcuende, Daniel Osuna, Oliver Thimm, Björn Usadel, Wolf-Rüdiger Scheible, and Mark Stitt. 2005. “Sugars and Circadian Regulation Make Major Contributions to the Global

Regulation of Diurnal Gene Expression in Arabidopsis." *The Plant Cell Online* 17 (12) (December 1): 3257 –3281. doi:10.1105/tpc.105.035261.
Bouche, Nicolas, and Hillel Fromm. 2004. "GABA in Plants: Just a Metabolite?" *Trends in Plant Science* 9 (3) (March): 110–115. doi:10.1016/j.tplants.2004.01.006.
Bown, Alan W., Dawn E. Hall, and Kennaway B. MacGregor. 2002. "Insect Footsteps on Leaves Stimulate the Accumulation of 4-aminobutyrate and Can Be Visualized Through Increased Chlorophyll Fluorescence and Superoxide Production." *Plant Physiology* 129 (4) (August 1): 1430–1434. doi:10.1104/pp.006114.
Briesemeister, Sebastian, Jörg Rahnenführer, and Oliver Kohlbacher. 2010. "Going from Where to Why—interpretable Prediction of Protein Subcellular Localization." *Bioinformatics* 26 (9) (May 1): 1232–1238. doi:10.1093/bioinformatics/btq115.
Bukhov, N. G. 2004. "Dynamic Light Regulation of Photosynthesis (A Review)." *Russian Journal of Plant Physiology* 51 (6): 742–753. doi:10.1023/B:RUPP.0000047822.66925.bf.
Cantón, Francisco R., María Fernanda Suárez, and Francisco M. Cánovas. 2005. "Molecular Aspects of Nitrogen Mobilization and Recycling in Trees." *Photosynthesis Research* 83 (2) (February): 265–278. doi:10.1007/s11120-004-9366-9.
Chi, Sang-Mun, and Dougu Nam. 2012. "WegoLoc: Accurate Prediction of Protein Subcellular Localization Using Weighted Gene Ontology Terms." *Bioinformatics* (January 31). doi:10.1093/bioinformatics/bts062. http://bioinformatics.oxfordjournals.org/content/early/2012/01/31/bioinformatics.bts062.
Choi, Jungik, and Maciek R. Antoniewicz. 2011. "Tandem Mass Spectrometry: A Novel Approach for Metabolic Flux Analysis." *Metabolic Engineering* 13 (2) (March): 225–233. doi:10.1016/j.ymben.2010.11.006.
Choi, Jungik, Matthew T. Grossbach, and Maciek R. Antoniewicz. 2012. "Measuring Complete Isotopomer Distribution of Aspartate Using Gas Chromatography/Tandem Mass Spectrometry." *Analytical Chemistry* 84 (10) (May 15): 4628–4632. doi:10.1021/ac300611n.
Chou, Kuo-Chen, and Hong-Bin Shen. 2010. "Plant-mPLoc: A Top-Down Strategy to Augment the Power for Predicting Plant Protein Subcellular Localization." *PLoS ONE* 5 (6) (June 28): e11335. doi:10.1371/journal.pone.0011335.
Christensen, B., and J. Nielsen. 1999. "Isotopomer Analysis Using GC-MS." *Metabolic Engineering* 1 (4): 282–290.
Clark, Shawn M., Rosa Di Leo, Preetinder K. Dhanoa, Owen R. Van Cauwenberghe, Robert T. Mullen, and Barry J. Shelp. 2009. "Biochemical Characterization, Mitochondrial Localization, Expression, and Potential Functions for an Arabidopsis {gamma}-aminobutyrate Transaminase That Utilizes Both Pyruvate and Glyoxylate." *J. Exp. Bot.* 60 (6) (April 1): 1743–1757. doi:10.1093/jxb/erp044.
Crown, Scott B, Woo Suk Ahn, and Maciek R Antoniewicz. 2012. "Rational Design of 13C-labeling Experiments for Metabolic Flux Analysis in Mammalian Cells." *BMC Systems Biology* 6 (1) (May 16): 43. doi:10.1186/1752-0509-6-43.
Crown, Scott B., and Maciek R. Antoniewicz. 2012. "Selection of Tracers for 13C-Metabolic Flux Analysis Using Elementary Metabolite Units (EMU) Basis Vector Methodology." *Metabolic Engineering* 14 (2) (March): 150–161. doi:10.1016/j.ymben.2011.12.005.
Dieuaide-Noubhani, Martine, Gerard Raffard, Paul Canioni, Alain Pradet, and Philippe Raymond. 1995. "Quantification of Compartmented Metabolic Fluxes in Maize Root Tips Using Iotope Distribution from 13C- or 14C-labeled Glucose." *J. Biol. Chem.* 270 (22) (June 2): 13147–13159. doi:10.1074/jbc.270.22.13147.

Eckardt, Nancy A. 2007. “Light Regulation of Plant Development: HY5 Genomic Binding Sites.” *The Plant Cell* 19 (3) (March): 727–729. doi:10.1105/tpc.107.052233.

Emanuelsson, Olof, Henrik Nielsen, Søren Brunak, and Gunnar von Heijne. 2000. “Predicting Subcellular Localization of Proteins Based on Their N-terminal Amino Acid Sequence.” *Journal of Molecular Biology* 300 (4) (July 21): 1005–1016. doi:10.1006/jmbi.2000.3903.

Fait, Aaron, Hillel Fromm, Dirk Walter, Gad Galili, and Alisdair R. Fernie. 2008. “Highway or Byway: The Metabolic Role of the GABA Shunt in Plants.” *Trends in Plant Science* 13 (1) (January): 14–19. doi:10.1016/j.tplants.2007.10.005.

Feild, T S, D W Lee, and N M Holbrook. 2001. “Why Leaves Turn Red in Autumn. The Role of Anthocyanins in Senescing Leaves of Red-osier Dogwood.” *Plant Physiology* 127 (2) (October): 566–574.

Fernie, Alisdair R, Peter Geigenberger, and Mark Stitt. 2005. “Flux an Important, but Neglected, Component of Functional Genomics.” *Current Opinion in Plant Biology* 8 (2) (April): 174–182. doi:10.1016/j.pbi.2005.01.008.

Fiehn, Oliver, Joachim Kopka, Richard N. Trethewey, and Lothar Willmitzer. 2000. “Identification of Uncommon Plant Metabolites Based on Calculation of Elemental Compositions Using Gas Chromatography and Quadrupole Mass Spectrometry.” *Analytical Chemistry* 72 (15): 3573–3580. doi:10.1021/ac991142i.

Foster, Cliff E., Tina M. Martin, and Markus Pauly. 2010. “Comprehensive Compositional Analysis of Plant Cell Walls (Lignocellulosic Biomass) Part II: Carbohydrates.” *Journal of Visualized Experiments* (37) (March 12). doi:10.3791/1837. http://www.jove.com/video/1837/comprehensive-compositional-analysis-plant-cell-walls-lignocellulosic.

González-Pérez, Sergio, Jorge Gutiérrez, Francisco García-García, Daniel Osuna, Joaquín Dopazo, Óscar Lorenzo, José L Revuelta, and Juan B Arellano. 2011. “Early Transcriptional Defense Responses in Arabidopsis Cell Suspension Culture Under High-light Conditions.” *Plant Physiology* 156 (3) (July): 1439–1456. doi:10.1104/pp.111.177766.

Gutierrez, Rodrigo, Laurence Lejay, Alexis Dean, Francesca Chiaromonte, Dennis Shasha, and Gloria Coruzzi. 2007. “Qualitative Network Models and Genome-wide Expression Data Define Carbon/nitrogen-responsive Molecular Machines in Arabidopsis.” *Genome Biology* 8 (1): R7. doi:10.1186/gb-2007-8-1-r7.

Heldt, Hans-Walter. 2004. *Plant Biochemistry*. 3rd ed. Academic Press.

Hendriks, Janneke H.M., Anna Kolbe, Yves Gibon, Mark Stitt, and Peter Geigenberger. 2003. “ADP-Glucose Pyrophosphorylase Is Activated by Posttranslational Redox-Modification in Response to Light and to Sugars in Leaves of Arabidopsis and Other Plant Species.” *Plant Physiology* 133 (2) (October 1): 838 –849. doi:10.1104/pp.103.024513.

Horvath, David. 2009. “Common Mechanisms Regulate Flowering and Dormancy.” *Plant Science* 177 (6) (December): 523–531. doi:10.1016/j.plantsci.2009.09.002.

Huang, Tao, Xiao-He Shi, Ping Wang, Zhisong He, Kai-Yan Feng, LeLe Hu, Xiangyin Kong, Yi-Xue Li, Yu-Dong Cai, and Kuo-Chen Chou. 2010. “Analysis and Prediction of the Metabolic Stability of Proteins Based on Their Sequential Features, Subcellular Locations and Interaction Networks.” *PLoS ONE* 5 (6) (June 4): e10972. doi:10.1371/journal.pone.0010972.

Huck, Jojanneke H J, Eduard A Struys, Nanda M Verhoeven, Cornelis Jakobs, and Marjo S van der Knaap. 2003. “Profiling of Pentose Phosphate Pathway Intermediates in Blood Spots by Tandem Mass Spectrometry: Application to Transaldolase Deficiency.” *Clinical Chemistry* 49 (8) (August): 1375–1380.

Hutchings, David, Stephen Rawsthorne, and Michael J. Emes. 2005. “Fatty Acid Synthesis and the Oxidative Pentose Phosphate Pathway in Developing Embryos of Oilseed Rape (Brassica Napus L.).” *J. Exp. Bot.* 56 (412) (February 1): 577–585. doi:10.1093/jxb/eri046.

Iyer, Vidya V., Ganesh Sriram, D. Bruce Fulton, Ruilian Zhou, Mark E. Westgate, and Jacqueline V. Shanks. 2008. “Metabolic Flux Maps Comparing the Effect of Temperature on Protein and Oil Biosynthesis in Developing Soybean Cotyledons.” *Plant, Cell & Environment* 31 (4) (April): 506–517. doi:doi:10.1111/j.1365-3040.2008.01781.x.

Kleijn, Roelco J., Jan-Maarten A. Geertman, Beckley K. Nfor, Cor Ras, Dick Schipper, Jack T. Pronk, Joseph J. Heijnen, Antonius J.A. van Maris, and Wouter A. van Winden. 2007. “Metabolic Flux Analysis of a Glycerol-overproducing *Saccharomyces Cerevisiae* Strain Based on GC-MS, LC-MS and NMR-derived ^{13}C-labelling Data.” *FEMS Yeast Research* 7 (2): 216–231. doi:10.1111/j.1567-1364.2006.00180.x.

Krook, J, D Vreugdenhil, C Dijkema, and L van der Plas. 1998. “Sucrose and Starch Metabolism in Carrot (Daucus Carota L.) Cell Suspensions Analysed by 13-labelling: Indications for a Cytosol and a Plastid-localized Oxidative Pentose Phosphate Pathway.” *J. Exp. Bot.* 49 (329) (December 1): 1917–1924. doi:10.1093/jexbot/49.329.1917.

Krouk, Gabriel, Daniel Tranchina, Laurence Lejay, Alexis A. Cruikshank, Dennis Shasha, Gloria M. Coruzzi, and Rodrigo A. Gutiérrez. 2009. “A Systems Approach Uncovers Restrictions for Signal Interactions Regulating Genome-wide Responses to Nutritional Cues in Arabidopsis.” *PLoS Comput Biol* 5 (3) (March 20): e1000326. doi:10.1371/journal.pcbi.1000326.

Kruger, Nicholas J, and Antje von Schaewen. 2003a. “The Oxidative Pentose Phosphate Pathway: Structure and Organisation.” *Current Opinion in Plant Biology* 6 (3) (June): 236–246. doi:10.1016/S1369-5266(03)00039-6.

———. 2003b. “The Oxidative Pentose Phosphate Pathway: Structure and Organisation.” *Current Opinion in Plant Biology* 6 (3) (June): 236–246. doi:10.1016/S1369-5266(03)00039-6.

Kruger, Nicholas J., Joanna E. Huddleston, Pascaline Le Lay, Nicholas D. Brown, and R. George Ratcliffe. 2007a. “Network Flux Analysis: Impact of 13C-substrates on Metabolism in Arabidopsis Thaliana Cell Suspension Cultures.” *Phytochemistry* 68 (16-18): 2176–2188. doi:10.1016/j.phytochem.2007.03.033.

———. 2007b. “Network Flux Analysis: Impact of 13C-substrates on Metabolism in Arabidopsis Thaliana Cell Suspension Cultures.” *Phytochemistry* 68 (16-18): 2176–2188. doi:10.1016/j.phytochem.2007.03.033.

Libourel, Igor G. L., and Yair Shachar-Hill. 2008. “Metabolic Flux Analysis in Plants: From Intelligent Design to Rational Engineering.” *Annual Review of Plant Biology* 59: 625–650.

Libourel, Igor G.L., Jackson P. Gehan, and Yair Shachar-Hill. 2007a. “Design of Substrate Label for Steady State Flux Measurements in Plant Systems Using the Metabolic Network of Brassica Napus Embryos.” *Phytochemistry* 68 (16-18): 2211–2221. doi:10.1016/j.phytochem.2007.04.033.

———. 2007b. “Design of Substrate Label for Steady State Flux Measurements in Plant Systems Using the Metabolic Network of Brassica Napus Embryos.” *Phytochemistry* 68 (16-18): 2211–2221. doi:10.1016/j.phytochem.2007.04.033.

Lonien, Joachim, and Jorg Schwender. 2009a. “Analysis of Metabolic Flux Phenotypes for Two Arabidopsis Thaliana Mutants with Severe Impairment in Seed Storage Lipid Synthesis.” *Plant Physiol.* (September 15): pp.109.144121.

doi:10.1104/pp.109.144121.
———. 2009b. “Analysis of Metabolic Flux Phenotypes for Two Arabidopsis Thaliana Mutants with Severe Impairment in Seed Storage Lipid Synthesis.” *Plant Physiol.* (September 15): pp.109.144121. doi:10.1104/pp.109.144121.
Ma, Ligeng, Jinming Li, Lijia Qu, Janet Hager, Zhangliang Chen, Hongyu Zhao, and Xing Wang Deng. 2001. “Light Control of Arabidopsis Development Entails Coordinated Regulation of Genome Expression and Cellular Pathways.” *The Plant Cell Online* 13 (12) (December 1): 2589–2607. doi:10.1105/tpc.010229.
Masakapalli, Shyam K, Pascaline Le Lay, Joanna E Huddleston, Naomi L Pollock, Nicholas J Kruger, and R George Ratcliffe. 2010. “Subcellular Flux Analysis of Central Metabolism in a Heterotrophic Arabidopsis Cell Suspension Using Steady-state Stable Isotope Labeling.” *Plant Physiology* 152 (2) (February): 602–619. doi:10.1104/pp.109.151316.
Masakapalli, Shyam K., Pascaline Le Lay, Joanna E. Huddleston, Naomi L. Pollock, Nicholas J. Kruger, and R.George Ratcliffe. 2009a. “Subcellular Flux Analysis of Central Metabolism in a Heterotrophic Arabidopsis Thaliana Cell Suspension Using Steady-state Stable Isotope Labeling.” *Plant Physiology* (November 25): pp.109.151316. doi:10.1104/pp.109.151316.
———. 2009b. “Subcellular Flux Analysis of Central Metabolism in a Heterotrophic Arabidopsis Thaliana Cell Suspension Using Steady-state Stable Isotope Labeling.” *Plant Physiology* (November 25): pp.109.151316. doi:10.1104/pp.109.151316.
Masoudi-Nejad, Ali, Susumu Goto, Takashi R. Endo, and Minoru Kanehisa. 2008a. “KEGG Bioinformatics Resource for Plant Genomics Research.” In *Plant Bioinformatics*, ed. D. Edwards, 406:437–458. Methods in Molecular Biology. Humana Press. http://dx.doi.org/10.1007/978-1-59745-535-0_21.
———. 2008b. “KEGG Bioinformatics Resource for Plant Genomics Research.” In *Plant Bioinformatics*, ed. D. Edwards, 406:437–458. Methods in Molecular Biology. Humana Press. http://dx.doi.org/10.1007/978-1-59745-535-0_21.
McClung, C. Robertson, Meier Hsu, Janet E. Painter, Jennifer M. Gagne, Sharon D. Karlsberg, and Patrice A. Salomé. 2000. “Integrated Temporal Regulation of the Photorespiratory Pathway. Circadian Regulation of Two Arabidopsis Genes Encoding Serine Hydroxymethyltransferase.” *Plant Physiology* 123 (1) (May): 381–392.
Metallo, Christian M., Jason L. Walther, and Gregory Stephanopoulos. 2009. “Evaluation of 13C Isotopic Tracers for Metabolic Flux Analysis in Mammalian Cells.” *Journal of Biotechnology* 144 (3) (November): 167–174. doi:10.1016/j.jbiotec.2009.07.010.
Minorsky, Peter V. 2003. “Achieving the in Silico Plant. Systems Biology and the Future of Plant Biological Research.” *Plant Physiology* 132 (June 1): 404–409.
Möllney, Michael, Wolfgang Wiechert, Dirk Kownatzki, and Albert A. de Graaf. 1999a. “Bidirectional Reaction Steps in Metabolic Networks: IV. Optimal Design of Isotopomer Labeling Experiments.” *Biotechnology and Bioengineering* 66 (2): 86–103.
———. 1999b. “Bidirectional Reaction Steps in Metabolic Networks: IV. Optimal Design of Isotopomer Labeling Experiments.” *Biotechnology and Bioengineering* 66 (2): 86–103.
Nargund, Shilpa, and Ganesh Sriram. 2012. “Designer Labels for Plant Metabolism: Statistical Design of Isotope Labeling Experiments for Improved Quantification of Flux in Complex Plant Metabolic Networks.” *Molecular BioSystems* (October 9). doi:10.1039/C2MB25253H. http://pubs.rsc.org/en/content/articlelanding/2012/mb/c2mb25253h.

Nunes-Nesi, Adriano, Alisdair R. Fernie, and Mark Stitt. 2010. “Metabolic and Signaling Aspects Underpinning the Regulation of Plant Carbon Nitrogen Interactions.” *Molecular Plant* 3 (6) (November 1): 973 –996. doi:10.1093/mp/ssq049.

Palanivelu, Ravishankar, Laura Brass, Anna F Edlund, and Daphne Preuss. 2003. “Pollen Tube Growth and Guidance Is Regulated by POP2, an Arabidopsis Gene That Controls GABA Levels.” *Cell* 114 (1) (July 11): 47–59. doi:10.1016/S0092-8674(03)00479-3.

Pardalos, Panos M., and H. Edwin Romeijn. 2002. *Handbook of Global Optimization*. Springer.

Paula Alonso, Ana, Val L. Dale, and Yair Shachar-Hill. 2010. “Understanding Fatty Acid Synthesis in Developing Maize Embryos Using Metabolic Flux Analysis.” *Metabolic Engineering* 12 (5) (September): 488–497. doi:10.1016/j.ymben.2010.04.002.

Peterman, T K, and H M Goodman. 1991. “The Glutamine Synthetase Gene Family of Arabidopsis Thaliana: Light-regulation and Differential Expression in Leaves, Roots and Seeds.” *Molecular & General Genetics: MGG* 230 (1-2) (November): 145–154.

Phee, Bong-Kwan, Jin-Hwan Cho, Sebyul Park, Jin Hee Jung, Youn-Hyung Lee, Jong-Seong Jeon, Seong Hee Bhoo, and Tae-Ryong Hahn. 2004. “Proteomic Analysis of the Response of Arabidopsis Chloroplast Proteins to High Light Stress.” *PROTEOMICS* 4 (11): 3560–3568. doi:10.1002/pmic.200400982.

Poolman, Mark G., Laurent Miguet, Lee J. Sweetlove, and David A. Fell. 2009. “A Genome-Scale Metabolic Model of Arabidopsis and Some of Its Properties.” *Plant Physiology* 151 (3) (November 1): 1570–1581. doi:10.1104/pp.109.141267.

Pregitzer, Kurt S., Donald I. Dickmann, Ron Hendrick, and Phu V. Nguyen. 1990. “Whole-tree Carbon and Nitrogen Partitioning in Young Hybrid Poplars.” *Tree Physiology* 7 (1-2-3-4) (December 1): 79 –93. doi:10.1093/treephys/7.1-2-3-4.79.

Press, William H., Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. 2nd ed. Cambridge University Press.

Quek, Lake-Ee, Christoph Wittmann, Lars Nielsen, and Jens Kromer. 2009. “OpenFLUX: Efficient Modelling Software for 13C-based Metabolic Flux Analysis.” *Microbial Cell Factories* 8 (1): 25. doi:10.1186/1475-2859-8-25.

Ratcliffe, R.G., and Y. Shachar-Hill. 2006. “Measuring Multiple Fluxes Through Plant Metabolic Networks.” *The Plant Journal* 45 (4) (February): 490–511. doi:doi:10.1111/j.1365-313X.2005.02649.x.

Rios-Estepa, Rigoberto, and B. Markus Lange. 2007. “Experimental and Mathematical Approaches to Modeling Plant Metabolic Networks.” *Phytochemistry* 68 (16-18): 2351–2374. doi:10.1016/j.phytochem.2007.04.021.

Rontein, Denis, Martine Dieuaide-Noubhani, Erick J. Dufourc, Philippe Raymond, and Dominique Rolin. 2002. “The Metabolic Architecture of Plant Cells. Stability of Central Metabolism and Flexibility of Anabolic Pathways During the Growth Cycle of Tomato Cells.” *Journal of Biological Chemistry* 277 (46) (November 8): 43948–43960. doi:10.1074/jbc.M206366200.

Roscher, Albrecht, Nicholas J. Kruger, and R.George Ratcliffe. 2000. “Strategies for Metabolic Flux Analysis in Plants Using Isotope Labelling.” *Journal of Biotechnology* 77 (1) (January 28): 81–102. doi:10.1016/S0168-1656(99)00209-6.

Sakakibara, Hitoshi, Kentaro Takei, and Naoya Hirose. 2006. “Interactions Between Nitrogen and Cytokinin in the Regulation of Metabolism and Development.” *Trends in Plant Science* 11 (9) (September): 440–448. doi:10.1016/j.tplants.2006.07.004.

Scheibe, R, A Geissler, and K Fickenscher. 1989. “Chloroplast Glucose-6-phosphate Dehydrogenase: Km Shift Upon Light Modulation and Reduction.” *Archives of*

Biochemistry and Biophysics 274 (1) (October): 290–297.
Schellenberger, Jan, Daniel Zielinski, Wing Choi, Sunthosh Madireddi, Vasiliy Portnoy, David Scott, Jennifer Reed, Andrei Osterman, and Bernhard Palsson. 2012. "Predicting Outcomes of Steady-state 13C Isotope Tracing Experiments Using Monte Carlo Sampling." *BMC Systems Biology* 6 (1) (January 30): 9. doi:10.1186/1752-0509-6-9.
Schmidt, K., J. Nielsen, and J. Villadsen. 1999. "Quantitative Analysis of Metabolic Fluxes in Escherichia Coli, Using Two-dimensional NMR Spectroscopy and Complete Isotopomer Models." *Journal of Biotechnology* 71 (1-3): 175–189.
Schwender, Jorg. 2008. "Metabolic Flux Analysis as a Tool in Metabolic Engineering of Plants." *Current Opinion in Biotechnology* 19 (2) (April): 131–137. doi:10.1016/j.copbio.2008.02.006.
Schwender, Jorg, Fernando Goffman, John B. Ohlrogge, and Yair Shachar-Hill. 2004. "Rubisco Without the Calvin Cycle Improves the Carbon Efficiency of Developing Green Seeds." *Nature* 432 (7018) (December 9): 779–782. doi:10.1038/nature03145.
Schwender, Jorg, John B. Ohlrogge, and Yair Shachar-Hill. 2003a. "A Flux Model of Glycolysis and the Oxidative Pentosephosphate Pathway in Developing Brassica Napus Embryos." *Journal of Biological Chemistry* 278 (32) (August 8): 29442–29453. doi:10.1074/jbc.M303432200.
———. 2003b. "A Flux Model of Glycolysis and the Oxidative Pentosephosphate Pathway in Developing Brassica Napus Embryos." *Journal of Biological Chemistry* 278 (32) (August 8): 29442–29453. doi:10.1074/jbc.M303432200.
Schwender, Jorg, John Ohlrogge, and Yair Shachar-Hill. 2004. "Understanding Flux in Plant Metabolic Networks." *Current Opinion in Plant Biology* 7 (3) (June): 309–317. doi:10.1016/j.pbi.2004.03.016.
Schwender, Jorg, M. Seemann, H. Lichtenthaler, and M. Rohmer. 1996. "Biosynthesis of Isoprenoids (carotenoids, Sterols, Prenyl Side-chains of Chlorophylls and Plastoquinone) via a Novel Pyruvate/glyceraldehyde 3-phosphate Non-mevalonate Pathway in the Green Alga Scenedesmus Obliquus." *Biochemical Journal* 316 (May 15): 73–80.
Schwender, Jorg, Yair Shachar-Hill, and John B. Ohlrogge. 2006. "Mitochondrial Metabolism in Developing Embryos of Brassica Napus." *Journal of Biological Chemistry* 281 (45) (November 10): 34040–34047. doi:10.1074/jbc.M606266200.
Shelp, Barry J., Alan W. Bown, and Denis Faure. 2006. "Extracellular {gamma}-aminobutyrate Mediates Communication Between Plants and Other Organisms." *Plant Physiology* 142 (4) (December 1): 1350–1352. doi:10.1104/pp.106.088955.
Shelp, Barry J., Alan W. Bown, and Michael D. McLean. 1999. "Metabolism and Functions of Gamma-aminobutyric Acid." *Trends in Plant Science* 4 (11) (November 1): 446–452. doi:10.1016/S1360-1385(99)01486-7.
Singh, B. K. 1998a. *Plant Amino Acids (Books in Soils, Plants, & the Environment)*. 1st ed. CRC.
———. 1998b. *Plant Amino Acids (Books in Soils, Plants, & the Environment)*. 1st ed. CRC.
Sriram, Ganesh, D. Bruce Fulton, Vidya V. Iyer, Joan Marie Peterson, Ruilian Zhou, Mark E. Westgate, Martin H. Spalding, and Jacqueline V. Shanks. 2004a. "Quantification of Compartmented Metabolic Fluxes in Developing Soybean Embryos by Employing Biosynthetically Directed Fractional 13C Labeling, Two-dimensional [13C, 1H] Nuclear Magnetic Resonance, and Comprehensive Isotopomer Balancing." *Plant Physiology* 136 (2) (October 1): 3043–3057. doi:10.1104/pp.104.050625.
———. 2004b. "Quantification of Compartmented Metabolic Fluxes in Developing Soybean Embryos by Employing Biosynthetically Directed Fractional 13C Labeling, Two-

dimensional [13C, 1H] Nuclear Magnetic Resonance, and Comprehensive Isotopomer Balancing." *Plant Physiology* 136 (2) (October 1): 3043–3057. doi:10.1104/pp.104.050625.

Sriram, Ganesh, D. Bruce Fulton, and Jacqueline V. Shanks. 2007. "Flux Quantification in Central Carbon Metabolism of Catharanthus Roseus Hairy Roots by 13C Labeling and Comprehensive Bondomer Balancing." *Phytochemistry* 68 (16-18): 2243–2257. doi:10.1016/j.phytochem.2007.04.009.

Sriram, Ganesh, Vidya V. Iyer, D. Bruce Fulton, and Jacqueline V. Shanks. 2007. "Identification of Hexose Hydrolysis Products in Metabolic Flux Analytes: A Case Study of Levulinic Acid in Plant Protein Hydrolysate." *Metabolic Engineering* 9 (5-6): 442–451. doi:10.1016/j.ymben.2007.07.003.

Sriram, Ganesh, Lola Rahib, Jian-Sen He, Allison E. Campos, Lilly S. Parr, James C. Liao, and Katrina M. Dipple. 2008. "Global Metabolic Effects of Glycerol Kinase Overexpression in Rat Hepatoma Cells." *Molecular Genetics and Metabolism* 93 (2) (February): 145–159. doi:10.1016/j.ymgme.2007.09.008.

Sriram, Ganesh, and Jacqueline V. Shanks. 2004. "Improvements in Metabolic Flux Analysis Using Carbon Bond Labeling Experiments: Bondomer Balancing and Boolean Function Mapping." *Metabolic Engineering* 6 (2) (April): 116–132. doi:10.1016/j.ymben.2004.02.003.

Srour, Orr, Jamey D Young, and Yonina C Eldar. 2011. "Fluxomers: a New Approach for 13C Metabolic Flux Analysis." *BMC Systems Biology* 5 (1): 129. doi:10.1186/1752-0509-5-129.

Stephanopoulos, Gregory. 1999. "Metabolic Fluxes and Metabolic Engineering." *Metabolic Engineering* 1 (1) (January): 1–11.

———. 2002. "Metabolic Engineering: Perspective of a Chemical Engineer." *AIChE Journal* 48 (5): 920–926.

Stephanopoulos, Gregory, and D. E. Stafford. 2002a. "Metabolic Engineering: a New Frontier of Chemical Reaction Engineering." *Chemical Engineering Science* 57 (14): 2595–2602.

———. 2002b. "Metabolic Engineering: a New Frontier of Chemical Reaction Engineering." *Chemical Engineering Science* 57 (14): 2595–2602.

Sticklen, Mariam. 2006. "Plant Genetic Engineering to Improve Biomass Characteristics for Biofuels." *Current Opinion in Biotechnology* 17 (3) (June): 315–319. doi:10.1016/j.copbio.2006.05.003.

Strasser, Bárbara, Maximiliano Sánchez-Lamas, Marcelo J. Yanovsky, Jorge J. Casal, and Pablo D. Cerdán. 2010. "Arabidopsis Thaliana Life Without Phytochromes." *Proceedings of the National Academy of Sciences* (February 22). doi:10.1073/pnas.0910446107. http://www.pnas.org/content/early/2010/02/12/0910446107.

Streb, Sebastian, Barbara Egli, Simona Eicke, and Samuel C. Zeeman. 2009. "The Debate on the Pathway of Starch Synthesis: A Closer Look at Low-Starch Mutants Lacking Plastidial Phosphoglucomutase Supports the Chloroplast-Localized Pathway." *Plant Physiology* 151 (4) (December 1): 1769–1772. doi:10.1104/pp.109.144931.

Sweetlove, L, D Fell, and A Fernie. 2008. "Getting to Grips with the Plant Metabolic Network." *Biochemical Journal* 409 (January 1): 27–41.

Sweetlove, L. J., R. L. Last, and A. R. Fernie. 2003. "Predictive Metabolic Engineering: a Goal for Systems Biology." *Plant Physiology* 132 (2): 420–5.

Szyperski, Thomas. 1995. "Biosynthetically Directed Fractional 13C-labeling of Proteinogenic Amino Acids. An Efficient Analytical Tool to Investigate Intermediary Metabolism." *European Journal of Biochemistry* 232 (2): 433–448.

doi:10.1111/j.1432-1033.1995.tb20829.x.
———. 1998. “13C-NMR, MS and Metabolic Flux Balancing in Biotechnology Research.” *Quarterly Reviews of Biophysics* 31 (01): 41–106.
Thum, Karen E., Dennis E. Shasha, Laurence V. Lejay, and Gloria M. Coruzzi. 2003. “Light- and Carbon-Signaling Pathways. Modeling Circuits of Interactions.” *Plant Physiology* 132 (2) (June 1): 440–452. doi:10.1104/pp.103.022780.
Thum, Karen, Michael Shin, Rodrigo Gutiérrez, Indrani Mukherjee, Manpreet Katari, Damion Nero, Dennis Shasha, and Gloria Coruzzi. 2008. “An Integrated Genetic, Genomic and Systems Approach Defines Gene Networks Regulated by the Interaction of Light and Carbon Signaling Pathways in Arabidopsis.” *BMC Systems Biology* 2 (1) (April 4): 31. doi:10.1186/1752-0509-2-31.
Turner, S R, R Hellens, R Ireland, N Ellis, and S Rawsthorne. 1993. “The Organisation and Expression of the Genes Encoding the Mitochondrial Glycine Decarboxylase Complex and Serine Hydroxymethyltransferase in Pea (Pisum Sativum).” *Molecular & General Genetics: MGG* 236 (2-3) (January): 402–408.
Usuda, Hideaki, and Gerald E. Edwards. 1980. “Localization of Glycerate Kinase and Some Enzymes for Sucrose Synthesis in C3 and C4 Plants 1.” *Plant Physiology* 65 (5) (May): 1017–1022.
Walther, Jason L, Christian M Metallo, Jie Zhang, and Gregory Stephanopoulos. 2012. “Optimization of 13C Isotopic Tracers for Metabolic Flux Analysis in Mammalian Cells.” *Metabolic Engineering* 14 (2) (March): 162–171. doi:10.1016/j.ymben.2011.12.004.
Wiechert, Wolfgang. 2001. “13C Metabolic Flux Analysis.” *Metabolic Engineering* 3 (3) (July): 195–206. doi:10.1006/mben.2001.0187.
Wiechert, Wolfgang, and Albert A. de Graaf. 1997. “Bidirectional Reaction Steps in Metabolic Networks: I. Modeling and Simulation of Carbon Isotope Labeling Experiments.” *Biotechnology and Bioengineering* 55 (1): 101–117.
Wiechert, Wolfgang, Michael Möllney, Nichole Isermann, Michael Wurzel, and Albert A. de Graaf. 1999. “Bidirectional Reaction Steps in Metabolic Networks: III. Explicit Solution and Analysis of Isotopomer Labeling Systems.” *Biotechnology and Bioengineering* 66 (2): 69–85.
Wiechert, Wolfgang, Michael Mollney, Soren Petersen, and Albert A. de Graaf. 2001. “A Universal Framework for 13C Metabolic Flux Analysis.” *Metabolic Engineering* 3 (3) (July): 265–283. doi:10.1006/mben.2001.0188.
Wiechert, Wolfgang, Claudia Siefke, Albert A. de Graaf, and Achim Marx. 1997a. “Bidirectional Reaction Steps in Metabolic Networks: II. Flux Estimation and Statistical Analysis.” *Biotechnology and Bioengineering* 55 (1): 118–135.
———. 1997b. “Bidirectional Reaction Steps in Metabolic Networks: II. Flux Estimation and Statistical Analysis.” *Biotechnology and Bioengineering* 55 (1): 118–135.
Williams, Thomas C R, Mark G Poolman, Andrew J M Howden, Markus Schwarzlander, David A Fell, R George Ratcliffe, and Lee J Sweetlove. 2010. “A Genome-scale Metabolic Model Accurately Predicts Fluxes in Central Carbon Metabolism Under Stress Conditions.” *Plant Physiology* 154 (1) (September): 311–323. doi:10.1104/pp.110.158535.
Williams, Thomas C.R., Laurent Miguet, Shyam K. Masakapalli, Nicholas J. Kruger, Lee J. Sweetlove, and R. George Ratcliffe. 2008a. “Metabolic Network Fluxes in Heterotrophic Arabidopsis Cells: Stability of the Flux Distribution Under Different Oxygenation Conditions.” *Plant Physiology* 148 (2) (October 1): 704–718. doi:10.1104/pp.108.125195.
———. 2008b. “Metabolic Network Fluxes in Heterotrophic Arabidopsis Cells: Stability of

the Flux Distribution Under Different Oxygenation Conditions." *Plant Physiology* 148 (2) (October 1): 704–718. doi:10.1104/pp.108.125195.

Winden, Wouter A. van, Joseph J. Heijnen, Peter J. T. Verheijen, and Johan Grievink. 2001a. "A Priori Analysis of Metabolic Flux Identifiability from ^{<FONT SIZE='-1'>13</FONT>}C-labeling Data." *Biotechnology and Bioengineering* 74 (6): 505–516. doi:10.1002/bit.1142.

———. 2001b. "A Priori Analysis of Metabolic Flux Identifiability from 13C-labeling Data." *Biotechnology and Bioengineering* 74 (6): 505–516.

van Winden, Wouter A., Joseph J. Heijnen, and Peter J.T. Verheijen. 2002. "Cumulative Bondomers: A New Concept in Flux Analysis from 2D [13C, 1H] COSY NMR Data." *Biotechnology and Bioengineering* 80 (7): 731–745.

Wu, Zhi-Cheng, Xuan Xiao, and Kuo-Chen Chou. 2011. "iLoc-Plant: a Multi-label Classifier for Predicting the Subcellular Localization of Plant Proteins with Both Single and Multiple Sites." *Molecular bioSystems* 7 (12) (December): 3287–3297. doi:10.1039/c1mb05232b.

Zhang, Peifen, Hartmut Foerster, Christophe P. Tissier, Lukas Mueller, Suzanne Paley, Peter D. Karp, and Seung Y. Rhee. 2005. "MetaCyc and AraCyc. Metabolic Pathway Databases for Plant Research." *Plant Physiology* 138 (1) (May 1): 27–37. doi:10.1104/pp.105.060376.

Zhu, Baolong, and Gary D. Coleman. 2001. "Phytochrome-Mediated Photoperiod Perception, Shoot Growth, Glutamine, Calcium, and Protein Phosphorylation Influence the Activity of the Poplar Bark Storage Protein Gene Promoter (bspA)." *Plant Physiol.* 126 (1) (May 1): 342–351. doi:10.1104/pp.126.1.342.