Network Flux Analysis of Central Metabolism in Plants

Shyam Kumar Masakapalli

Trinity term 2011

D.Phil Thesis

University of Oxford

Department of Plant Sciences

Exeter College

Supervisors: Professor R. George Ratcliffe & Dr. Nicholas J Kruger

Abstract

Network Flux Analysis of Central Metabolism in Plants

Shyam Kumar Masakapalli, Exeter College

Submitted for the degree of D.Phil., Trinity Term 2011

The aim of this thesis was to develop stable-isotope steady-state metabolic flux analysis (MFA) based on ^{13}C labeling to quantify intracellular fluxes of central carbon metabolism in plants. The experiments focus on the analysis of a heterotrophic cell suspension culture of *Arabidopsis thaliana* (L) Heynh. (ecotype Landsberg *erecta*). The first objective was to develop a robust methodology based on combining high quality steady-state stable labeling data, metabolic modeling and computational analysis. A comprehensive analysis of the factors that influence the outcome of MFA was undertaken and best practice established. This allowed a critical analysis of the subcellular compartmentation of carbohydrate oxidation in the cell culture. The second objective was to apply the methodology to nutritional perturbations of the cell suspension. A comparison of growth on different nitrogen sources revealed that transfer to an ammonium-free medium: (i) increased flux through the oxidative pentose phosphate pathway (oxPPP) by 10% relative to glucose utilisation; (ii) caused a substantial decrease in entry of carbon into the tricarboxylic acid cycle (TCA); and (iii) increased the carbon conversion efficiency from 55% to 69%. Although growth on nitrate alone might be expected to increase the demand for reductant, the cells responded by decreasing the assimilation of inorganic N. Cells were also grown in media containing different levels of inorganic phosphate (Pi). Comparison of the flux maps showed that decreasing Pi availability: (i) decreased flux through the oxPPP; (ii) increased the proportion of substrate fully oxidised by the TCA cycle; and (iii) decreased carbon conversion efficiency. These changes are consistent with redirection of metabolism away from biosynthesis towards cell maintenance as Pi is depleted. Although published genome-wide transcriptomic and metabolomic studies suggest that Pi starvation leads to the restructuring of carbon and nitrogen metabolism, the current analysis suggests that the impact on metabolic organisation is much less extreme.

Acknowledgements

I am indebted to Professor R. George Ratcliffe and Dr. Nick Kruger for their patience, careful supervision, encouragement and consistent unfailing enthusiasm throughout the years of my D.Phil. It has been both a privilege and a pleasure to have experienced the opportunity to be supervised by two leading international experts in this challenging field of flux analysis. I sincerely thank you both for being the sort of supervisors every student needs - astute, supportive, enthusiastic, and inspiring.

Thanks also to everyone at the Department of Plant Sciences that I've had the pleasure of sharing a lab with and learning from the last 4 years. I especially thank Dr. Thomas Williams for his advice, technical help with NMR and GC-MS and relevant discussions. I also thank Ms. Katherine Beard, Mr. Maurice Cheung, Ms. Sarah McCraw and Mr. Andy Hall for discussions of relevant topics, moral support and encouragement. I would like to express my appreciation to Mr. Pedro Bota for his help with maintenance of GC-MS and other lab activities.

I acknowledge the financial assistance provided by University of Oxford and Exeter College via Clarendon scholarship, an Overseas Research Student (ORS) award and Mr. Krishna Pathak (Exeter College) scholarship. Finally, I would like to thank my family and friends for their understanding, encouragement and constant support over the years. I dedicate this thesis to my beloved family.

Abbreviations

Fru-6-P	fructose 6-phosphate
AMM	atom mapping matrices
ANOVA	analysis of variance
C3-P	three-carbon phosphate ester pool
CCF	cell culture filtrate
CI	confidence intervals
CLE	carbon labeling experiments
EMA	elementary modes analysis
EMP	Embden-Meyerhof-Parnas pathway
EMU	elementary metabolite unit
EPA	extreme pathway analysis
FBA	flux balance analysis
FDA	fluorescein diacetate
GABA	γ-amino butyric acid
GC-MS	gas chromatography-mass spectroscopy
Glc-6-P	glucose 6-phosphate
ILE	isotope labeling experiments
IMM	isotopomer mapping matrices
LF	line-fitting
MC	Monte Carlo
MFA	metabolic flux analysis
MID	mass isotopomer distribution
MSCorr	mass correction software
MSTFA	N-methyl-N-trimethylsilyl trifluroacetamide
MTHF	Methylene tetra-hydrofoloate
N	nitrogen
NMR	nuclear magnetic resonance
oxPPP	oxidative pentose phosphate pathway
Pi	phosphate
PPP	pentose phosphate pathway
RE	relative errors
SIM	single ion monitoring
SNR	signal to noise ratio
SQP	sequential quadratic programming
TBDMS	N-methyl-N-t-butyl dimethylsilyl
t-BDMCS	t-butyl-dimethyl chlorosilane
TFA	trifluoro acetic acid
TIC	total ion current
TSP	trimethyl silyl propionate

Table of Contents

5.1B Mass isotopomer measurements of metabolites derived from *Arabidopsis* suspension grown on normal MS medium fed with either 99%[1-^{13}C]glucose, 99%[2-^{13}C]glucose or 20%[$^{13}C_6$]glucose

5.1C ^{13}C NMR measurements of metabolites derived from *Arabidopsis* suspension grown on N40 (nitrate only) medium fed with 99%[1-^{13}C]glucose or 99%[2-^{13}C]glucose

5.1D Mass isotopomer measurements of metabolites derived from *Arabidopsis* suspension grown on N40 medium (nitrate only) fed with either 99%[1-^{13}C]glucose, 99%[2-^{13}C]glucose or 20%[$^{13}C_6$]glucose

5.2 Metabolic Network of *Arabidopsis* cell suspension culture (Used in Chapters 5 and 6)

5.3A Comprehensive list of model reactions tested while model construction of *Arabidopsis thaliana* suspension culture.

5.3B List of Amino acid recycling reactions tested in the models

5.4 Non linear statistical analysis of free fluxes of MS and N40 models

6.1A ^{13}C NMR measurements of metabolites derived from *Arabidopsis* suspension grown on Pi- medium fed with either 99%[1-^{13}C]glucose , 99%[2-^{13}C]glucose or 20% [$^{13}C_6$]glucose

6.1B ^{13}C NMR measurements of metabolites derived from *Arabidopsis* suspension grown on MS medium fed with either 99%[1-^{13}C]glucose or 99%[2-^{13}C]glucose

6.1C ^{13}C NMR measurements of metabolites derived from *Arabidopsis* suspension grown on Pi+ medium fed with either 99%[1-^{13}C]glucose or 99%[2-^{13}C]glucose

6.1D Mass isotopomer measurements of metabolites derived from *Arabidopsis* suspension grown on different phosphate levels (Pi-, MS and Pi+) fed with either 99%[1-^{13}C]glucose, 99%[2-^{13}C]glucose or 20%[$^{13}C_6$]glucose

1 Introduction

1.1 Statement of aims

The major objective of this thesis is to develop stable-isotope steady-state metabolic flux analysis (MFA) based on ^{13}C labeling to quantify intracellular fluxes of central carbon metabolism in plants. The specific aims are:

- To establish *in silico* strategies using the software 13C-FLUX® for analysing the pathways of carbohydrate oxidation in plants (Chapter 3).
- To comprehensively analyse the factors that influence the outcome of MFA and establish best practices (Chapter 3). The major factors studied are methods for i) data handling of isotopomer measurements (obtained by GC-MS and NMR) and ii) strategies to account for pre-existing unlabelled pools of metabolites that are observed to be very common in labelling experiments on plant systems.
- To validate the robustness of the computational strategies developed by analysing existing data sets of heterotrophic Arabidopsis (*Arabidopsis thaliana*) cell suspension cultures with focus on compartmentation (Chapter 4)
- To define metabolic flux phenotypes of Arabidopsis cell suspension cultures under nutritional perturbations of inorganic nitrogen (Chapter 5) and phosphorus (Chapter 6).

1.2 Metabolic Flux- a crucial feature of cell function

Investigation of the fluxes (the rates of metabolite interconversion) through a metabolic network is crucial to elucidate the functional behavior of cells. Metabolic flux analysis (MFA) is a well established diagnostic tool used to describe and quantify metabolism at the systems

level (Wiechert and de Graff, 1996; Libourel and Shachar-Hill, 2008). The flux maps generated by MFA quantify multiple fluxes through various reactions and pathways and thus provide a functional description of the metabolic phenotype of the organism under study (Ratcliffe and Shachar-Hill, 2006). This system-wide measurement of fluxes, which has been termed the 'fluxome', complements the interaction between and within the genome, transcriptome, proteome and metabolome to define the physiological state of the model system as a whole (Sauer 2004; Sweetlove et al., 2003). Studies involving flux analysis of different phenotypes gives insight into metabolic control (Sweetlove and Fernie, 2005), predictive modeling (Ratcliffe and Shachar-Hill, 2001; Sweetlove et al., 2003), identification of novel metabolic routes (Boatright et al., 2004) and the rational design of metabolic engineering strategies (Ratcliffe and Shachar-Hill, 2006).

The key strategies for MFA is using ^{13}C carbon labeling experiments (CLE) along with integration of biological, analytical, mathematical and computational principles (Wiechert et al., 2001). The present chapter is focused on introducing two major aspects of this approach – (i) basic concepts and prerequisites of steady state MFA; and (ii) modeling design, one of the key aspects of ^{13}C based MFA in order to effectively understand the underlying basic concepts for efficient modeling.

1.3 ^{13}C MFA – a key diagnostic tool for flux analysis

MFA based on ^{13}C labeling experiments (known as ^{13}C MFA) on cells or tissues at their steady state followed by mathematical analysis of the labeling data is currently the principal method used to quantify intracellular fluxes throughout the metabolic network. In principle, ^{13}C MFA

deals with the computation of the unknown intracellular fluxes from the experimentally measured extracellular fluxes (such as uptake rate and biomass effluxes) and

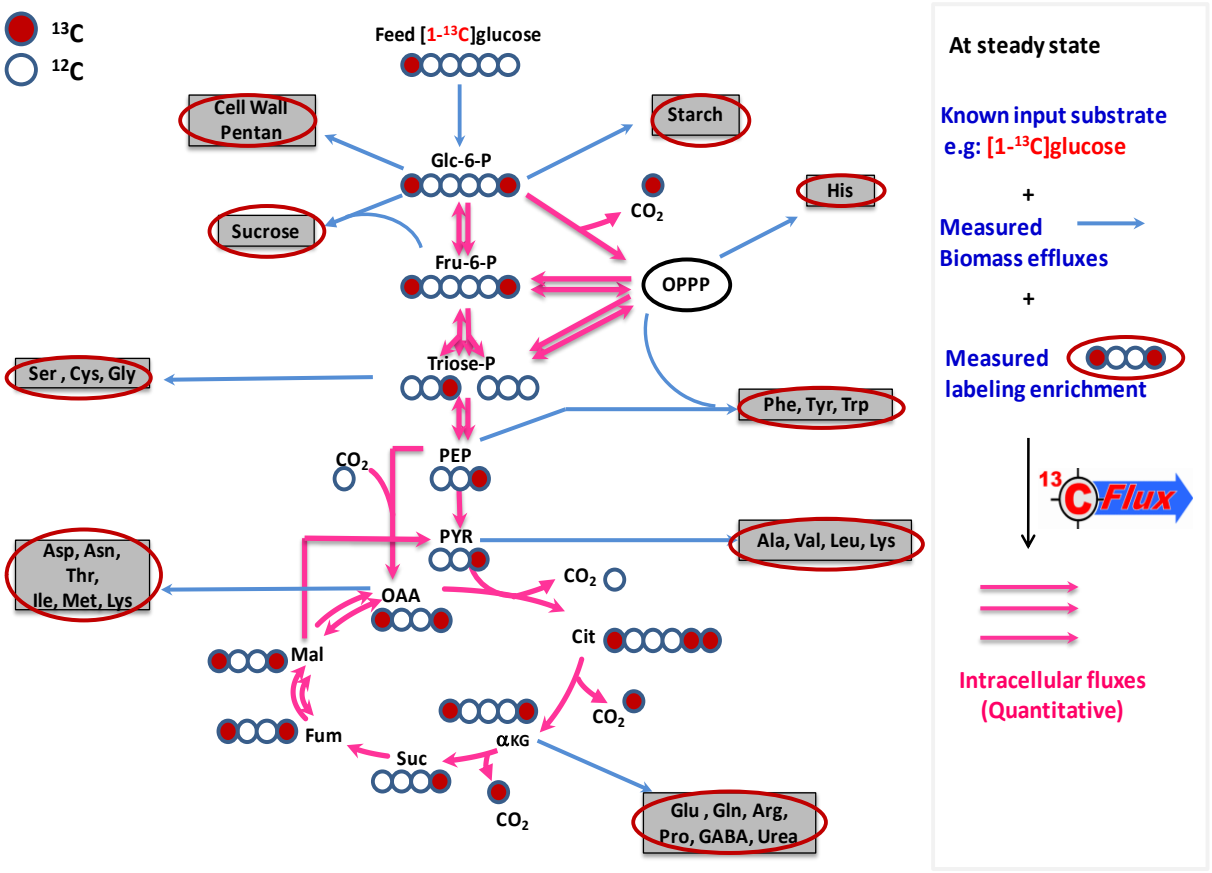


Figure 1.1: Strategy for steady state ^{13}C Metabolic Flux Analysis (^{13}C MFA) with a simplified cellular metabolic network. The biological material is supplied with a ^{13}C-labelled substrate (e.g., [1-^{13}C]glucose] until the system reaches isotopic steady state. The label is redistributed within the network as a result of fluxes i.e. activities through individual reactions. First the label reaches the central metabolites like sugars, organic acids, amino acids and then enters the biomass (output) metabolites like protein, cell wall, starch and sucrose. At isotopic steady state, the label enrichment of biomass metabolites reports on central metabolites and the fluxes. The extracellular fluxes (blue lines) are experimentally quantified and the label enrichment within each metabolite (in red) is measured by either NMR or GC-MS. These measurements provide constraints for quantifying intracellular fluxes (in pink) by mathematical modeling using custom software such as 13CFLUX. In a flux map the thickness of arrow is often used to depict the flux through the reaction. The label enrichment of metabolites and thickness of arrows in this figure are purely hypothetical and only used for illustrative purposes.

information on the fractional labeling of intracellular metabolites obtained by nuclear magnetic resonance (NMR) or mass spectroscopy (MS) techniques (see Fig 1.1). It has the capacity to accommodate reversible steps, cyclic fluxes and subcellular compartmentation (Masakapalli et al., 2010). The reliability of the flux maps obtained depends on the structure of the network, the labelled substrate used, the label measurements obtained, the flux values and the biological assumptions made (Wiechert et al., 2001; Roscher et al., 2000).

1.4 Basic assumptions and prerequisites of steady state ^{13}C MFA

The basic assumptions for applying steady-state ^{13}C MFA to determine fluxes through the network are three fold. The first is that there is no appreciable isotopic effect on metabolism either due to enzymatic preference between different labelled metabolites or other effects of the presence of label. Evidence in support of this assumption on metabolism has recently been obtained in higher plants in which [^{13}C]glucose labelling had no detectable effect on the intracellular fluxes (Kruger et al., 2007). The second is that no omitted reaction alters the distribution of label within metabolites in the network. Finally, there is a presumption that the carbon atom transitions (i.e. the relationship between carbon atoms in the substrate and the product) of individual reactions in the network are well defined (Wiechert and de Graff, 1997). However, unknown reactions and the occurrence of metabolite channeling could lead to uncertainties (van Winden et al., 2001, Williams et al., 2011).

The key prerequisites for steady-state ^{13}C MFA are: (i) the system is at metabolic and isotopic steady state; (ii) the metabolic network for the organism under study is precisely defined; (iii) rates of substrate consumption and biomass are available to provide adequate constraints on

the network; and (iv) isotopomer measurement data of network intermediates and products resulting from the metabolism of a defined labelled substrate are available.

1.4.1 Metabolic and isotopic steady state condition

Steady state MFA requires the system to reach metabolic and isotopic steady state; i.e., the amounts of internal metabolites and their fractional isotopic enrichment are constant over time. The metabolic steady state can be established by measurements of the rate of growth and biomass accumulation. The isotopic steady state can be obtained by analyzing the labeling enrichment of metabolites in a time-course. In a typical plant cell, the central network metabolites usually reach isotopic steady state within a few hours of labeling whereas polymers like protein, cell wall and starch can take several days (Troufflard et al., 2007).

1.4.2 Precise formulation of metabolic network

In practice, the biochemical reactions and their corresponding carbon transitions are obtained from standard biochemical texts, biochemical pathway database of model organisms (e.g., KEGG, AraCyc) and the scientific literature. Linear reactions (where there is no change in carbon position) are frequently considered as a single composite step, and rapidly exchanging metabolite pools are usually combined into a single pool. For a compartmentalized system (e.g., a plant cell) with duplicated reactions and parallel pathways in different subcellular organelles, localisation of various metabolites and amino acids has to be taken into account (Ratcliffe and Shachar-Hill, 2006). Also, the traditional reactions as conventionally depicted in biochemical schemes may need to be modified to more precisely define the redistribution of label arising from the reactions catalysed by component enzymes of the network (e.g., traditional reactions of the non-oxidative pentose phosphate pathway can be replaced by half-

reactions (Kleijn et al., 2005)). At present, networks of central carbon metabolism have been implemented successfully for ^{13}C MFA studies in microbes and plants (Rontein et al., 2002; Schwender and Ohlrogge., 2002; Schwender et al., 2003; Schwender et al., 2004; Schwender et al., 2006; Sriram et al., 2004; Spielbauer et al., 2006; Alonso et al., 2007; Kruger et al., 2007; Williams et al., 2008; Masakapalli et al., 2010).

1.4.3 Determination of substrate consumption and biomass production

The rates of substrate uptake and biomass accumulation provide important constraints in most steady-state MFA experiments. They supply the external boundaries of the system that aid in obtaining intracellular fluxes using the label patterns of network metabolites. The biomass measurements such as protein, lipid, cell wall components, starch, storage pools of free amino acids, organic acids, sugars and CO_2 release provide the output constraints for metabolic modeling. These can be measured using either standard biochemical methods (Williams et al., 2008;, Sriram et al., 2006) or by incubating the system with [U-^{14}C]substrate followed by measuring the amount of radioactivity in targeted end products of fractionated extracts (Masakapalli et al., 2010).

1.4.4 Labelled input substrate and isotopomer measurement data

In general, the choice of one or more labelled substrates depends on the physiological context of the system. Most studies on heterotrophic plant tissues are generally labeled with either [^{13}C]glucose (Williams et al., 2008; Masakapalli et al., 2010) or [^{13}C]sucrose (Sriram et al., 2007, Lonien and Schwender 2009). Also multiple labeled substrate such as labeled sugars and amino acids are successfully employed in plants for flux analysis (Schwender et al., 2006, Allen et al., 2007).

The labelling enrichment of metabolites and end-products is analysed by nuclear magnetic resonance (NMR) spectroscopy and/or mass spectrometry (MS) (Ratcliffe and Shachar-Hill, 2006; Allen and Ratcliffe, 2009) to quantify the relative abundances of positional isotopomers or mass isotopomers, respectively. Analysis of the labelling patterns is a routine application for both techniques, and ideally both approaches should be used to maximize the confidence in the deduced fluxes (Kleijn et al., 2007). The appropriate methods for data handling of isotopomer measurements obtained from NMR and MS are considered in Chapter 3 of this thesis.

1.5 Modeling design: *in silico* modeling and simulations of ^{13}C MFA experiments

Mathematical modeling and simulation are vital scientific tools for examining metabolism at the systems level. Over the past few years, several modeling methods and tools have been developed and implemented for flux analysis. Broadly the modeling methods can be categorized into elementary modes analysis (EMA, Schuster et al., 1999), extreme pathway analysis (EPA, Schilling et al., 2000), energy balance analysis (EBA, Beard et al., 2002)), flux balance analysis (FBA, Edwards et al., 2002) and isotope labelling experiments (ILE). Two approaches of ILE exist, dynamic and steady state, with the later approach being widely used for analysis of the highly integrated network of reactions of central carbon metabolism. For the branching linear network of secondary metabolism dynamic labeling may be more informative. For a description of these methods and case studies one can refer to recent reviews (Ratcliffe and Shachar-Hill, 2006, Libourel and Shachar-Hill, 2008, Rios-Estepa and Lange, 2007). FBA and ILE methods are of special interest due to their potential for defining intracellular fluxes and can be complementary (Fig 1.2, Libourel and Shachar-Hill, 2008).

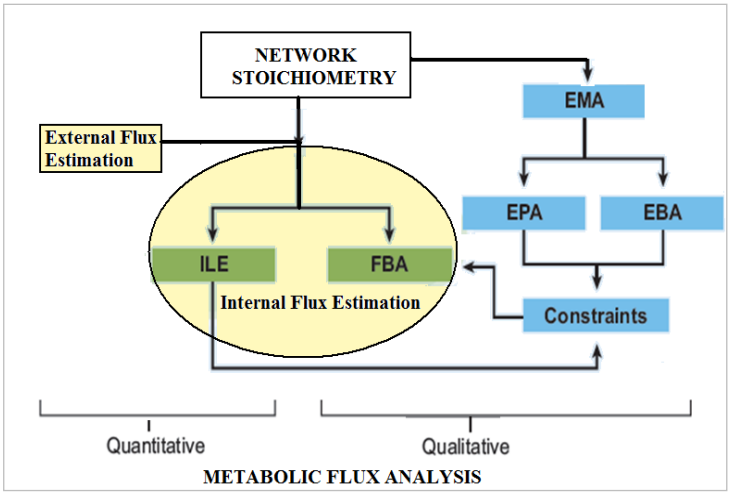


Figure 1.2: Overview of flux analysis modeling methods. ILE (^{13}C MFA) is used for quantitative flux analysis (Adapted from Libourel and Shachar-Hill, 2008) with slight modifications)

The internet resources and software packages for flux analysis are well documented (Ratcliffe and Shachar-Hill, 2006;, Rios-Estepa and Lange, 2007). Among these the widely used modeling tool for analysis of steady state labeling experiments (^{13}C MFA) is 13C-FLUX (Wiechert et al. 2001). This can be used for *in silico* simulation, experimental design, data evaluation and statistical analysis of ^{13}C labeling experiments. The tool has been used on microbial and plant systems and evolved over the years with regular changes and newer versions (13C-FLUX current version 20050329) available free for academic use (http://www.uni-siegen.de/fb11/simtec/software/ 13cflux/?lang=de). The ability of this software to simulate both NMR and MS data along with a range of optimisation and statistical routines makes it a preferred tool for ^{13}C MFA in plants (Ratcliffe and Shachar-Hill., 2006). The subsequent sections and Chapter 3 explore ^{13}C MFA and the utility of 13C-FLUX with emphasis on design of models and modeling strategies for generating reliable flux maps.

1.5.1 Stoichiometric modeling and development of ^{13}C MFA

All flux analysis methods follow the principles of stoichiometric modeling (Llaneras and Pico 2008) in which a biochemical network with m metabolites (M) and n fluxes (v) is defined in a stoichiometry matrix (S) of size m x n. A stoichiometry matrix is built from mass-balance

equations around the internal metabolites (M) and the reactions (v). Under a steady state assumption i.e. when the internal metabolite pool sizes and the fluxes are constant, the following expression is obtained which forms the linear algebraic mass balance equations (i.e., rate of production equals rate of removal):

$$dM/dt = S.v = 0$$

where, dM/dt is the change in metabolite concentration with change in time (constant at steady state), S is the stoichiometry matrix and v is the vector of all metabolic fluxes. The system contains a certain degree of freedom (n-rank(S)) which defines the number of free fluxes in the network from which all other dependent fluxes can be determined.

In most natural metabolic networks the FBA method is basically underdetermined and its solution depends on additional constraints and assumptions of objective functions (e.g. maximum growth, biomass accumulation) (Varma and Palsson, 1994; Libourel & Shachar-Hill 2008). The major limitations of this approach are its inability to determine fluxes through reversible reactions, parallel pathways and metabolic cycles (Wiechert et al., 2001, Tang et al., 2007). The shortcomings of this method and other pure stoichiometry-based methods (Vallino and Stephanopoulous, 1993; Varma and Palsson, 1994) led to the development of ^{13}C MFA based on carbon/isotope labeling experiments (Marx et al., 1996; Wiechert et al., 2001) which allows intracellular fluxes to be quantified. While FBA method has its limitations, it is now considered to be a useful complement to ^{13}C MFA (Sweetlove and Ratcliffe, 2011).

1.5.2 The concept of isotopomer networks is the central scheme of ^{13}C MFA modeling

1.5.2.1 *Isotopomer concept*

Isotopic isomers, or isotopomers, are chemically identical compounds that differ only in the distribution of isotope within the molecule, and these provide the most rigorous framework for interpretation of isotopic labelling studies. A metabolite with n carbon atoms has 2^n positional isotopomers which can be distinguished by binary suffix 0 and 1 (0 for absence and 1 for presence of isotope label) and n+1 mass isotopomers (positional isotopomers having equal mass). A metabolite (for example with two carbons) can be experimentally measured and computationally modeled for their isotopomer fractions (four isotopomers #00, #01, #10 & #11), positional enrichment (C1 & C2), or mass isotopomers (M0, M1, M2) (Figs. 1.3a, 1.3b). These measurements provide constraints for modeling in ^{13}C MFA (Wiechert 1999). The modeling tools generate isotopomer balance equations from the stoichiometric network via the carbon transition network which conserve the fate of each carbon flow.

Although the computational modeling of labelling experiments was initiated by generating atom mapping matrices (AMMs) (Zupke and Stephanopoulos, 1994), these were subsequently refined in the construction of isotopomer mapping matrices (IMMs) (Schmidt et al., 1997). IMMs have the advantage of holding label information and isotopomer mass balance expressions in a single matrix equation for a metabolite pool (Rios-Estepa and Lange, 2007). Many other approaches for modeling of isotope labelling systems have evolved over the years such as the cumomer concept, the bondomer concept and elementary metabolite unit (EMU)) or cumomers network decomposition methods. All of them provide the same information (i.e. relation of flux to isotopomers) but differ in their computational efficiency as discussed below.

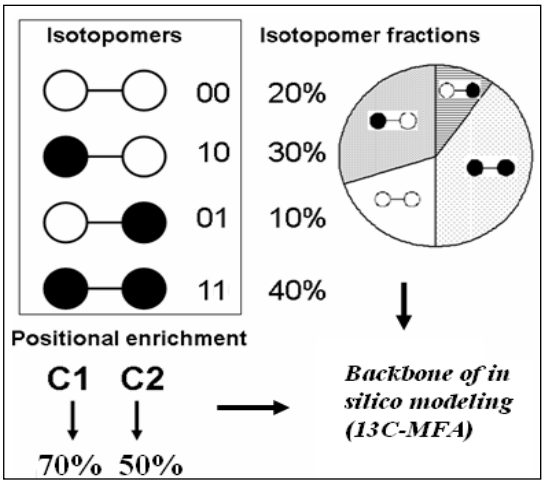


Figure 1.3a: Example showing isotopomers of a 2-carbon metabolite with their hypothetical fractional abundance (sum to 100%) and corresponding positional enrichments

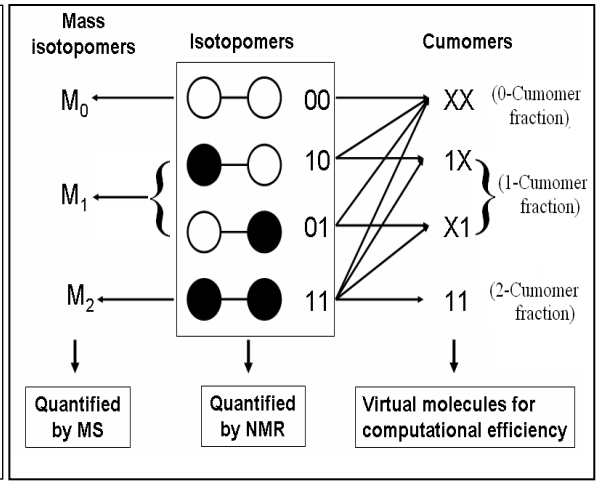


Figure 1.3b: Example showing relation between isotopomers, mass isotopomers and cumomers of a 2-carbon compound

Wiechert et al., 1999 proposed the "concept of cumomer redundancy" and introduced an approach for cumomer network analysis which is computationally effective. Cumomers are set of cumulated isotopomers specified by the subscript "X" and "1" (X-depicts an atom that is either labeled or unlabeled and 1 indicates a labelled atom). Isotopomers can be readily transformed to cumomers and vice versa (Fig 1.3b). The cumomers are referred as "virtual molecules" which can be dealt as n-cumomer (in which n is the number of specified labeled atoms) fractions that form a cascade of cumomer sub-networks during computational modelling. The 0-cumomer fraction (#XX), 1-cumomer fraction (#X1 & #1X) and 2-cumomer fraction (#11) can be deduced for any metabolite (Fig 1.3b). Cumomer network and cumomer balance equations are readily formed from the isotopomer network and isotopomer balance equations respectively. 3D visualisation of the cumomer networks by tools such as CumoVis (Droste et al., 2008) can help in isotope labelling experimental design and path tracing for carbons of interest. It should be noted that the software framework 13C-FLUX (Wiechert et

al., 2001) is based on the isotopomer and cumomer concept and is by far the most widely used analytical approach.

1.5.2.2 *Bondomer concept*

Sriram et al., 2004 used the concept of bondomer balancing for ^{13}C MFA modeling. Following ^{13}C labelling experiment, the carbon-carbon connectivity information is explored to derive bondomers of a metabolite which differ in the positions and number of C-C intact bonds (van Winden et al., 2002). The approach can be used only with universally labelled substrates (e.g [$^{13}C_6$] glucose) thereby limiting the scope of using positional labeled substrates (e.g [1-^{13}C]glucose) that may have higher sensitivity to deduce fluxes of interest.

1.5.2.3 *Improving the speed of analysis – Elementary metabolite units and cumomer network decomposition*

Recently, two new concepts have emerged which are known to reduce the computational time needed for modeling isotope labeling systems and as a consequence have potential advantages for high throughput ^{13}C MFA. First, Antoniewicz et al., (2007) proposed the concept of EMU (elementary metabolite units) for modeling multiple isotopic tracer experiments (^{13}C, ^{2}H, ^{18}O). EMUs are distinct subsets of the atoms of a compound (i.e., 2^n-1 EMUs for a compound with n atoms). For example, a two carbon compound X forms three EMUs (X_1 and X_2 of EMU size 1 and X_{12} of EMU size 2). The approach uses only a fraction of all possible EMUs (i.e. the minimum number of variables) thereby reducing the dimension of the isotope network. The algorithm decomposes the whole biochemical network into EMUs of different sizes (i.e., all molecular fractions with EMU size 1, EMU size 2 and so on) until the substrates are traced back, and finally considers the minimal set of EMU reactions for simulation model. Second, Wietzel et al., (2007) proposed an alternative approach based on "cumomer network

decomposition" which exploits graph theory concepts and algorithms. Here, the isotope labeling system defined by the cumomer network is further decomposed into smaller units. This approach is orthogonal to EMU modeling. However, these methods both provide the same information as that of the isotopomer and cumomers based models. Further, at present, the software tools for implementation of these approaches are not readily available. Hence, these methods will not be discussed further.

1.5.3 Flux definitions-net, exchange and reversibility

All fluxes in a network are categorized as net fluxes, exchange fluxes or as reversibilities. The traditional forward and backward fluxes are found to be inappropriate for the interpretation of bidirectional reactions. Hence, the terms net flux (V^{net}) and exchange flux (V^{xch}) were introduced (Wiechert and de Graaf, 1996) which can be derived from the forward and backward fluxes.

$$V^{net} = V_{forward} - V_{backward} \qquad V^{xch} = \min (V_{forward}, V_{backward})$$

V^{xch} has a possible range of 0 to infinity [0, ∞], making it unstable mathematically at large exchange fluxes. This is overcome by transforming the exchange flux range to a [0,1] scale either by the terms $\mathbf{V^{xch[0,1]}}$ **(**Wiechert and de Graaf, 1996) or reversibility(r) (Sriram et al., 2004). These are represented as

$$V^{xch[0\text{-}1]} = V^{xch} / (1+V^{xch})$$

$$\text{Reversibility(r)} = V_{backward}/V_{forward}$$

For a bidirectional reaction, $V^{xch[0\text{-}1]}$ or a reversibility flux of 0 represents irreversibility and 1 represents complete reversibility i.e at thermodynamic equilibrium.

1.5.4 Relation between flux and label measurements

There is a strong relationship between the flux and label redistribution in the network (Wiechert et al., 1997). The number and quality of label measurements define whether the model is over-determined or structurally undeterminable (van Winden et al., 2001; Iserman

and Wiechert, 2003). The model is over-determined when the number of independent measurements available is greater than the number of unique fluxes to be estimated. When the available measurements cannot resolve the fluxes of some pathways or reactions, these fluxes are referred to as being structurally undeterminable. However, this does not influence the other determinable parts of the network. Therefore precise experimental measurements are needed for reliable flux estimation (Wiechert et al., 2001, Roscher et al., 2000). It is possible to perform *a priori* flux identifiability analysis and substrate design to obtain crucial measurements needed for estimating fluxes of interest (Libourel et al., 2007). Moreover, in practice, over-determined models with significantly large number of measurements are still needed to estimate fluxes throughout the network.

1.5.5 Simulations of carbon labelling experiment

In principle, the input (uptake of substrate) and output (biomass effluxes) fluxes are considered to be unidirectional and these are normally constrained (C) by experimentally measured values. All the intracellular fluxes are estimated relative to the uptake flux (usually set to 100%). The exchange fluxes of thermodynamically irreversible reactions are constrained to be zero and the corresponding net fluxes are constrained to be positive. The stoichiometric equations of the network allow a number of free fluxes (F) from which all other dependent fluxes (D) can be solved. 13C-FLUX aids in the choice of free fluxes and considers user defined linear equality and inequality constraints between net and exchange fluxes (Wiechert et al., 2001). All these parameters are entered into a textual notation file (known as an ".ftbl" file) following the nomenclature suggested by Wiechert et al., (2001), which is processed by the 13C-FLUX routines. In Chapter 3 of this thesis, various *in silico* strategies needed for compartmented flux analysis in plants are further tested and presented. Further enhancements

including the Monte Carlo simulations of multiple models, non-linearised statistical analysis and other key consideration of data handling, are introduced in the appropriate sections. In Chapters 4, 5 and 6 the *in silico* strategies developed were applied to perform robust flux analysis in compartmented models to understand the metabolic phenotypes of *Arabidopsis* cell suspension cultures under nutritional perturbation of nitrogen sources and phosphate availability.

1.6 Choice of experimental system

In order to meet the objectives stated, a heterotrophic cell suspension culture of *Arabidopsis thaliana* (L) Heynh. (ecotype Landsberg *erecta*) was chosen as the experimental system. The *Arabidopsis* cell suspension is experimentally tractable, as it has been used successfully for MFA (Williams et al., 2008, Masakapalli et al., 2010), and it avoids the difficulties that might arise from multiple cell types in differentiated tissues. Moreover many of the general metabolic features of the chosen *Arabidopsis* culture are similar to the heterotrophic metabolism that occurs in roots (Lehmann et al., 2009). The cell suspension was chosen for the investigation to simplify the task of testing different models of the subcellular compartmentation of the oxidative pentose phosphate pathway (Chapter 4) and to establish metabolic phenotypes of compartmented systems under different nutritional regimes of nitrogen (Chapter 5) and phosphorus (Chapter 6).

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals

[1-^{14}C]glucose (specific activity, 2.00 GBq mmol^{-1}), [6-^{14}C]glucose (specific activity, 2.07 GBq mmol^{-1}), and [U-^{14}C]glucose (specific activity, 10.4 GBq mmol^{-1}) were purchased from GE Healthcare Life Sciences. [3,4-^{14}C]glucose (specific activity, 1.11 GBq mmol^{-1}) was purchased from American Radiolabeled Chemicals, and [2-^{14}C]glucose (specific activity, 1.67 GBq mmol^{-1}) was obtained from Perkin-Elmer. [2-^{13}C]glucose (99 atom %) was purchased from Omicron Biochemicals, and [1-^{13}C]glucose (99 atom %) and [$^{13}C_6$]glucose (99 atom %) were obtained from Isotec.. TBDMS (*N*-methyl-*N*-(*t*-butyldimethylsilyl) trifluoroacetamide + 1% *t*-butyl-dimethylchlorosilane, Regis Technologies Inc), MSTFA (*N*-methyl-*N*-(trimethylsilyl) trifluroacetamide (HiChrom)), Dioxane, Trimethyl silyl propionate (TSP). All enzymes were from Roche Diagnostics. General chemicals and chromatography resins were purchased from Sigma-Aldrich or Merck unless otherwise stated.

2.1.2 Software

13CFLUX, Agilent Chemstation, MSCorr, NUTS (Acorn NMR), Matlab, Metalign, Amdis

2.1.3 Analytical equipment

Scintillation counter, NMR, GC-MS (Agilent), Hansatech O_2 electrode, Microplate reader (ELx 808, Bio-tek instruments Inc).

2.2 Arabidopsis cell suspension culture

2.2.1 Maintenance of light grown Arabidopsis suspension

Cell suspensions of *Arabidopsis thaliana* (ecotype Landsberg *erecta*) (May and Leaver, 1993) were maintained on an orbital shaker at 150 rpm in 250 mL Erlenmeyer flasks sealed with a double layer of aluminium foil under a 16 h light, 8 h dark cycle at 22°C (Williams *et al.*, 2008). Every 7 d, 15 mL of cell suspension was subcultured into 100 mL of fresh Murashige and Skoog (MS) medium containing the micro- and macronutrients including vitamins (Duchefa) supplemented with 3% (w/v) glucose (final conc. 166 mM), 0.5 mg L^{-1} naphthaleneacetic acid (NAA), and 0.05 mg L^{-1} kinetin. The growth medium was prepared by mixing equal volumes of autoclaved 6% (w/v) glucose and 8.8 g L^{-1} MS, 1 mg L^{-1} NAA and 0.1 mg L^{-1} kinetin corrected to pH 5.8 with 0.2 M KOH prior to autoclaving.

2.2.2 Treatment with different inorganic phosphate (Pi) levels

2.2.2.1 *Growth of heterotrophic Arabidopsis cell suspension and harvesting*

Heterotrophic *Arabidopsis* cell suspensions were produced by subculturing 6 mL of a 7 d old light-grown cell suspension into 36 mL of fresh growth medium containing different inorganic phosphate (Pi) concentrations in 100 ml Erlenmeyer flasks followed by incubation in the dark at 22°C in an orbital shaker set at 150 rpm. Growth medium was made as above (section 2.2.1) except that the macronutrients were adjusted to contain three different levels of Pi: (a) 0 mM KH_2PO_4 (Pi-) for the phosphate starved condition; (b) 1.25 mM KH_2PO_4 (MS) for normal growth; and (c) 5 mM KH_2PO_4 (Pi+) for the phosphate abundant condition. (See Appendix 2.1 for complete composition of media). To establish phosphate starved (Pi-) cultures, 7 d old light grown cultures were allowed to settle, 60 ml of cell culture filtrate (CCF) was discarded from

the supernatant followed by transfer of 6 ml cells from the remaining homogeneous suspension into 36 ml of growth medium containing 0 mM KH_2PO_4. This was done to allow for the slow growth of cells under Pi- conditions. Cell growth was monitored until the cells reached stationary phase. Cells were harvested by vacuum filtration through a single paper filter (Whatman #4), transferred to a pre-weighed container, the fresh weight determined and the samples were immediately frozen in liquid N_2. Tissue was stored at -80^{o}C prior to lyophilisation. Samples of CCF were also collected and stored at -80^{o}C.

2.2.2.2 *Isotope feeding experiments to establish isotopic steady state*

To establish the isotopic and metabolic state conditions of Pi- cultures, the cells were grown with 20% [U-$^{13}C_6$]glucose for 5, 10, 15 and 20 d. MS and Pi+ cultures were also fed with 20% [$^{13}C_6$]glucose until day 5.5 when the cells are in their exponential phase of growth and reach isotopic steady state (Williams et al., 2008). As mentioned in the following sections, the cells were harvested, extracted and the soluble metabolites analysed by GC-MS to establish the isotopic steady state conditions.

2.2.2.3 *Steady state isotope feeding experiments for 13C MFA*

For the labelling experiments, cells were subcultured into Pi-, MS and Pi+ growth medium supplemented with the selected labelled glucose. Three different feeding strategies were used: 99% [1-^{13}C]glucose; 99% [2-^{13}C]glucose; and a mixture of 20% [$^{13}C_6$]glucose and 80% unlabeled glucose (Sigma). Cells were grown in the dark at 22°C until they reached isotopic steady state (5.5 d for MS and Pi+, 15 d for Pi-), harvested and stored.

2.2.3 Growth with different nitrogen sources

Arabidopsis cell suspensions were subcultured (as mentioned in 2.2.2.1) into growth media containing different inorganic nitrogen (N) levels: (a) normal MS growth medium, which contained 18.79 mM KNO_3 and 20.61 mM NH_4NO_3 (MS); (b) 20 mM KNO_3 (N20); (c) 40 mM KNO_3 (N40); and (d) 20 mM NH_4Cl plus 20 mM KCl (NH_4-20). The growth media containing glucose (3% w/v), micronutrients, macronutrients (with appropriate N levels), vitamins and hormones was prepared as described in section 2.2.1.

2.2.3.1 *Isotope feeding experiments to establish isotopic steady state*

Based on physiological characterisation, MS (containing 18.79 mM KNO_3 and 20.61 mM NH_4NO_3) and N40 (40 mM KNO_3) conditions were selected for further ^{13}C feeding experiments. To establish the isotopic and metabolic steady state conditions of Arabidopsis suspension cultures under MS and N40 conditions, the cells were grown with 20% [$^{13}C_6$]glucose for 5.5d when the cells are in their exponential phase of growth and would reach isotopic steady state (Williams et al., 2008). As mentioned in the following sections, the cells were harvested, extracted and the soluble metabolites analysed by GC-MS to establish the isotopic steady state conditions.

2.2.3.2 *Steady state isotope feeding experiments for 13C MFA*

For the labelling experiments, cells were subcultured into growth medium containing different N sources (MS or N40) and selected labelled glucose. Three different feeding strategies were used: 99% [1-^{13}C]glucose; 99% [2-^{13}C]glucose; and a mixture of 20% [$^{13}C_6$]glucose and 80% unlabeled glucose (Sigma). Cells were grown in the dark at 22°C until they reached isotopic steady state (5.5 d), harvested and stored until further extraction and analysis.

2.3 Physiological studies

2.3.1 Measuring growth and cell viability

Growth was measured from the fresh weight and dry weight (after lyophilisation) of the harvested cells. Cell viability was measured microscopically by dual staining using fluorescein diacetate (FDA) and Evans blue. The living cells metabolise FDA to form fluorescein (Excitation max-494nm, Emission max-518nm) which can be observed using a UV fluorescence microscopy (FITC band pass filter), whereas Evans blue accumulates in the dead cells as blue protoplasmic stain (Turner and Novacky, 1974). For the cell viability assay, 2 µl of FDA (1mg/ml) solution and 2 µl of Evans blue (10% W/V) were added to 1 ml of Arabidopsis suspension culture followed by incubation at room temperature for 3 min. 30 µl of cell suspension was then mounted on a slide and analyzed using microscopy (Bright field and UV fluorescence). Photographs were taken to count the % of living cells. Viability of the cell suspensions was assessed by determining the proportion of around 150 cells spread over three different locations on each slide that stained green. Growth and cell viability of Pi- cultures were monitored for 20 d, whereas the MS and Pi+ cultures were monitored for 8 days, when they reached stationary phase.

2.3.2 Phosphorus extraction and quantification

2.3.2.1 *Extraction of free inorganic phosphate from the cells and culture filtrate*

Pi was extracted as mentioned by Bozzo *et al.* (2006). 10 mg of lyophilised *Arabidopsis* cells were ground in 500 µl of 10% (w/v) perchloric acid and centrifuged at 13000 rpm for 10 min. Supernatants were collected in separate tubes and were neutralized with 5 M KOH/1 M triethanolamine. The neutralized extracts were centrifuged as above to discard the potassium

perchlorate pellet, and the supernatants were used for determination of soluble Pi. CCF was used directly for quantification of Pi levels in the medium during the growth period.

2.3.2.2 *Inorganic phosphate quantification assay*

Pi forms a phosphomolybdate complex with ammonium molybdate in the presence of Zn^{2+} at pH 5 after reduction with ascorbic acid. The complex produces a chromophore that shows maximum absorption near 850 nm. In the absence of phosphate, the chromophore was not produced. This property of the phosphomolybdate complex was utilized for the phosphate assay (Drueckes et al., 1995; Saheki et al., 1985).

Molybdate reagent was prepared by addition of ammonium molybdate (final conc. 10 mM) to zinc acetate (15 mM), adjusted to pH 5.0 with 36% HCl. This was stored at 4ºC in the dark. A 10% (w/v) ascorbic acid solution, pH 5.0 adjusted with 40%(w/v) NaOH, was prepared immediately before use. Standard solutions of phosphate (KH_2PO_4) were prepared to obtain a standard curve in the range of 3 – 150 nmol phosphate. For the assay, the reagent mixture was prepared by the addition of molybdate reagent (1 part) to the reducing reagent (4 parts). In a microtitre plate, 30 µl test samples containing Pi were incubated with 125 µl of reagent mixture at 37ºC for 1 h. The absorbance of the phosphomolybdate complex (bluish) was measured at 720 nm (the maximum wavelength available) in an automated microplate reader ELx 808 (Bio-tek instruments Inc).

2.3.3 Rate of oxygen consumption

Measurements of the rate of oxygen consumption were made using a Clark type oxygen electrode (Hansatech oxygraph). Electrodes were calibrated using air saturated water (Truesdale and Downing, 1954) to set the maximum level and a pinch of sodium hydrosulphite

was used to remove all oxygen from the chamber in order to set the minimum level. 1 ml of the *Arabidopsis* cell suspension was suspended in the electrode chamber and the rate of O_2 consumption monitored at 22°C. The rates were measured as nmol O_2 consumed per min. The data were converted into nano moles of O_2 consumed per min per gram fresh weight of cells. Each treatment was measured in triplicate.

2.3.4 Glucose concentration

An enzymatic assay was used to quantify glucose in the culture medium of cell suspension cultures. In a microtitre plate 5 µl of sample was mixed with 45 µl deionised H_2O and 200 µl of assay cocktail containing 150 mM Tris (pH 8.1), 5 mM $MgCl_2$, NAD (0.2 mg/ml stock) and ATP (0.66 mg/ml stock). The baseline absorbance (a) was measured at 340 nm. To start the reaction, 5 µl of 0.5 U hexokinase/ glucose-6-phosphate dehydrogenase (G6PDH, Roche diagnostics) mix was added and incubated at room temperature for 30 min. The final absorbance (b) was again recorded at 340 nm and the concentration of glucose in the sample calculated according to the difference in the final and initial absorbance ($A_{340} = b - a$). Samples were diluted with water as appropriate so that the assay was not saturated to ensure that the absorbance measurements fell within the linear range of operation of the spectrophotometer. Alternatively, glucose concentration in the medium was determined from ^1H NMR of CCF (Section 2.8.2).

2.3.5 Acid Phosphatase assay

Enzymatic assay of acid phosphatase (EC 3.1.3.2) in the CCF was performed according to the Sigma protocol (Bergmeyer et al., 1974) available from their website. Para-nitrophenyl

phosphate (pNPP), a chromogenic substrate for most phosphatases yields para-nitrophenol (pNP), an yellow soluble product that can be measured by a spectrophotometer at 410 nm.

In an eppendorf, 100 µl citrate buffer (90 mM, pH 4.8) and 100 µl pNPP (15.2 mM) were mixed by inversion and equilibrated to 37 °C in an incubator. 20 µl of CCF was added to the mixture and incubated for exactly 10 min at 37 °C. Then the reaction was stopped by adding 800 µl of NaOH (100 mM). The mixture was transferred to a cuvette (light path 1 cm) and the absorbance recorded at 410 nm. To obtain the absorbance of controls (or blanks) 20 µl of CCF was added after the addition of NaOH. Units of acid phosphatase per ml of CCF was calculated as follows:

Acid phosphatase units/ml of CCF = (A_{410} CCF - A_{410} Blank)* total vol of solution (in ml)/ time of assay (in mins)* extinction coefficient of pNP at 410 nm* volume of enzyme used (in ml)

2.4 Extraction procedures after stable isotope labeling

2.4.1 Extraction and fractionation of soluble metabolites

2.4.1.1 *Perchloric acid extraction*

Soluble metabolites were extracted from lyophilised cells using perchloric acid (Kruger et al., 2008). 1 ml of 3 M ice cold perchloric acid was added to 100 mg of lyophilised cells and ground in a pre-cooled mortar and pestle. The ground sample was transferred to a 15 ml falcon tube placed on ice. 5 ml of ice-cold 1 M perchloric acid was added, mixed by vortex to obtain a homogeneous suspension and allowed to stand on ice for 30 minutes. Samples were centrifuged at 3500 rpm for 30 minutes at 4^{o}C to pellet insoluble tissue and the supernatant

decanted and retained. The pellet was re-suspended with another 5 ml of ice-cold 1 M perchloric acid, centrifuged as above and the supernatant from this was combined with the previous step in a 50 ml falcon tube. The tissue pellet was retained and stored at -80^{o}C for extracting protein, starch and cell wall components later. The pooled perchloric extracts were then neutralised by adding about 6.5 ml of 2 M ice cold KOH, and the samples allowed to stand on ice for 30 minutes to allow insoluble potassium perchlorate ($KClO_4$) to precipitate. The pH of the solution was checked and adjusted to be within 5.0-7.0 with KOH/HCl and the samples centrifuged as described above. The resulting supernatant was collected and freeze dried for 3 days. After freeze drying, the samples were re-dissolved in 3 ml of deionised water. The pH of the samples were checked and adjusted to be between 4.0 and 6.0. Samples were centrifuged again as described above to remove any undissolved $KClO_4$ and the supernatant collected. 1 ml of the supernatant was stored for ion-exchange fractionation and subsequent GC-MS analysis. The remaining 2 ml of supernatant was freeze dried and stored at -80^{o}C for NMR analysis.

2.4.1.2 *Ion exchange chromatography*

Preparation and charging of ion-exchange columns (Dowex-50 and Dowex-1)

Anion exchange resin (Dowex-1, 1*8-200(Sigma)) was suspended three times in excess deionised water, allowed the resin to settle and the unsettled material was removed. The soaked resin was charged by first washing it three times in excess 1 M sodium acetate followed by washing repeatedly with 0.1 M acetic acid until the washings became acidic.

Cation exchange resin (Dowex-50, 50*8-200(Sigma)) was prepared by adding 2 volumes of 4 M HCl and incubating in a water bath at 100^{o}C for 1 h. The resin was then washed with deionised water in a Buchner funnel until the pH of the washings reached neutral. Following

this the resin was resuspended in 2 volumes of 10% (w/v) NaOH and heated in a water bath at 100^{o}C for 1 h. The resin was again washed with deionised water as described above, resuspended in 2 volumes of 2 M HCl and stored at 4 °C until needed. The resin was charged for use by washing three times with excess 2 M HCl followed by washing with deionised water until the pH of the washings is the same as pure deionised water.

Ion-exchange fractionation of the samples

Anionic and cationic exchange columns were prepared by pipetting 500 µl of appropriate charged resin slurry into 1000 µl plastic pipette tips stopped with glass beads , through which excess liquid is allowed to drain off. The cationic column was placed on top of the anionic column and placed inside a 5 ml scintillation tube. The pH of the samples to be fractionated was adjusted to between 3 and 6 using 0.2 M KOH or 0.1 M HCl. Sample (1 ml) was added to the cation exchange column and the liquid allowed to pass through to the anion exchange column. The columns were then washed with 2.5 ml deionised water and the liquid eluted into the scintillation vial was retained as the neutral fraction (principally contains sugars). The anionic and cation exchange columns were then separated and placed in 5 ml scintillation tubes. The cation exchange column was eluted with 2 ml of 1 M NH_4OH to obtain the basic fraction (mainly contains amino acids) and the anion exchange column was eluted with 2 ml of 2 M formic acid to obtain the acidic fraction (mainly contains organic acids). The separated neutral, acidic and basic fractions were dried in a vacuum evaporator and stored at -80 ^{o}C until further use for GC-MS and NMR analysis.

2.4.2 Extraction and hydrolysis of protein

2.4.2.1 *Extraction with urea/thiourea*

Protein was extracted from half of the insoluble pellet obtained after the perchloric acid extraction of the lyophilised cells (Williams et al., 2008). The pellets were transferred to a 15 ml falcon tube and 200 µl of 2 M KOH was added to neutralise any residual perchloric acid, followed by addition of 3 ml 6 M urea/2 M thiourea. The samples were carefully homogenised by a polytron for approximately 30 sec and the homogenate was split into two 2 ml eppendorf tubes and centrifuged at 13000 rpm for 30 min at 4°C. The supernatant was collected in fresh eppendorfs and again centrifuged as above to remove any cell debris. The supernatants containing the soluble protein were transferred to a 15 ml falcon tube and retained for protein precipitation.

2.4.2.2 *Precipitation*

Protein was precipitated by adding 100% trichloroacetic acid so that its final concentration was 15% (v/v). Samples were vortexed and stood on ice for 24 hours in a cold room. Precipitated protein was pelleted by centrifugation of the samples for 30 minutes at 4^{o}C and 3500 rpm. The supernatant was decanted carefully (hazardous waste). The pellets were washed twice with 4 ml of ice-cold acetone, centrifuged as above to remove acetone. The pellets were allowed to stand in a flow hood to evaporate any acetone left before hydrolysis.

2.4.2.3 *Standard hydrolysis*

Extracted protein pellets, present in 15 mL falcon tubes were resuspended in 2 mL of 6 M HCl. A hole was punched through the lid of the tube and fitted with a 100 µl plastic pipette tip to provide a vent and to prevent pressure building up during incubation. Tubes were incubated at

100^{o}C in a dri-block for 24 h. Samples were cooled to room temperature and the hydrolysates (containing amino acids) were transferred to eppendorf tubes and dried in a speed-vac overnight to remove HCl. The dried samples were redissolved in 1 ml deionised water and passed through a 0.22 μm syringe filter to remove particulates before adjusting the pH to between 3 and 6. For purification of amino acids from the protein hydrolysates, the samples were fractionated by cationic exchange column (section 2.4.1.2). The fraction obtained were dried again and stored at -80^{o}C until further use for GCMS and NMR analysis.

2.4.3 Starch digestion

Starch was extracted from the remaining half of the insoluble pellet obtained after the perchloric acid extraction of the lyophilised cells. The pellets were transferred to a 15 ml falcon tube and washed with water twice. Each sample was gelatinized by autoclaving the residue in 3 ml of 25 mM sodium acetate for 3 h at 121°C, 1.04 bars pressure followed by enzymic digestion with 20 U α-amylase (Sigma-Aldrich) and 5 U amyloglucosidase (Sigma-Aldrich) for 16 h at 37°C. The samples were centrifuged and the supernatant containing the glucose monomers was collected, freeze dried and stored at -80°C until further analysis by GC-MS and NMR.

2.4.4 Cell wall hydrolysis

De-starched insoluble material was hydrolysed in a screw-top tubes by addition of 1 ml of 2 M trifluoroacetic acid (TFA) at 120°C for 1 h. The tubes were cooled to room temperature and centrifuged at 13000 rpm for 2 min. The supernatant containing hydrolysed monosaccharides were transferred to a new tube and stored at -20°C until further use. Prior to GC-MS analysis, 10 μl of the supernatant was transferred to an eppendorf and vacuum dried to remove TFA.

The samples were dissolved in 20 µl methanol and again vaccum dried to remove any residual TFA. The dried samples containing the monosaccharides were analysed by GC-MS.

2.5 In vivo NMR

In vivo NMR spectra of *Arabidopsis thaliana* cell suspensions were recorded using a perfusion system designed for 10 mm NMR probes. Heterotrophic *Arabidopsis* cells were carefully packed into the NMR tube to about 6 cm in length (or ~3 g FW) above the Vyon® porous plastic filter so that the sample was centered within the detection coil of the probe. Two peristaltic pumps were used to circulate the growth medium through the reservoir and the packed cells at 3.33 ml/min. Oxygen gas was bubbled into the medium to ensure sufficient oxygenation of the cells. A safety water-pump powered vacuum line was used to prevent overflow of the system. To obtain spectra under hypoxic conditions, the pumps were stopped.

2.5.1 In vivo ^{13}C NMR

To record *in vivo* ^{13}C NMR spectra of 5 d heterotrophic Arabidopsis cultures, the tube containing the packed cells was placed inside the magnet with both peristaltic pumps running and oxygenation of the medium reservoir. A capillary containing 2.5 M 1,4-dioxane in $^{2}H_2O$ was placed coaxially in the NMR tube to act as an internal standard. About 10% $^{2}H_2O$ was added to the growth medium to give a total volume of 42 ml. The spectrometer was tuned, shimmed and locked before acquisition. *In vivo* ^{13}C NMR spectra of Arabidopsis suspension metabolising unlabelled glucose (~ 90 to 102 mM) were recorded at 150.844 MHz using a 10 mm broadband probe (P6753) for 2 h. To monitor the metabolism of cells fed with about 25% positionally labelled glucose, 10 ml of perfusion medium containing unlabelled glucose was replaced with 10 ml of 100 mM [1-^{13}C]glucose (prepared by dissolving in 9 ml water and 1 ml

D_2O). The spectrometer was immediately tuned and shimmed before acquisition of i*n vivo* ^{13}C NMR spectra for 24 h (96 spectra @ 128 scans/spectrum in an array). For Pi- cultures the spectrum were acquired for 48 h. The acquisition parameters employed were a relaxation delay of 6 s, a 90^o observe pulse of 24 μs, and an acquisition time of 1.016 s. Waltz 16 ^1H decoupling was used during the detection period. For processing the spectra, the arrays were zero filled, baseline corrected and a line-broadening (exponential multiplication) of 4 Hz was applied. After Fourier transformation, manual phase correction was done followed by linear regression to flatten the spectral baseline. Appropriate slices of the array were added to obtain the time-course kinetic data. Chemical shifts were measured relative to the 1,4-dioxane signal at 67.30 ppm, and peaks were detected and assigned on the basis of literature values and authentic standards. The in-vivo ^{13}C NMR data is not presented in this thesis.

2.5.2 In vivo ^{31}P NMR

To record *in vivo* ^{31}P spectra of 5 d heterotrophic Arabidopsis cultures, the tube containing the packed cells was placed inside the magnet with both peristaltic pumps running and oxygenation of the medium reservoir. A capillary containing 0.5 M methylene diphosphonate (MDP) in Tris buffer (pH 8.9) was placed coaxially in the NMR tube to act as an internal standard. About 10% ^2H_2O was added to the perfusion medium. The spectrometer was tuned, shimmed and locked before acquisition. 1D ^{31}P NMR spectra were recorded at 242.816 MHz using a 10 mm selective frequency ^{31}P probe (P6725) for 13.45 hrs (12 spectra @ 512 scans/spectrum). The acquisition parameters employed were a relaxation delay of 0.765 s, a 90^o observe pulse of 36 us, and an acquisition time of 1.235 s. No ^1H decoupling was used. For processing the spectra, the arrays were summed, zero filled twice, baseline corrected and line-broadening (exponential multiplication) of 8 Hz was applied. After Fourier transformation,

manual phase correction was done followed by linear regression to flatten the spectral baseline. Spectra were referenced to MDP at 0 ppm. The peaks detected were assigned with reference to literature and known standards.

2.6 Radiorespirometric analysis

This was done to study the release of $^{14}CO_2$ by *Arabidopsis* cell suspension cultures metabolizing positionally labeled glucose and gluconate**.** Stock solutions of positionally labelled [^{14}C]glucose and [^{14}C]gluconate (1.5 μCi in 750 μl for each substrate) were carefully prepared. The radioactivity of each [^{14}C] substrate stock was counted by taking 5 μl aliquot thrice as mentioned below (section 2.7). All the stocks were prepared a day before the ^{14}C feeding experiments, stored at -20°C overnight and thawed before use.

On day 4.5, 50 μl of [1-^{14}C]-, [2-^{14}C]-, [3,4-^{14}C]-, [6-^{14}C]glucose or [1-^{14}C]gluconate were pippetted into sterile 100 ml flasks. To each flask 5 mL of a homogeneous suspension of *Arabidopsis* cells along with their growth medium was added. Each flask was then immediately sealed with a rubber bung and aerated in the dark at 22°C on an orbital shaker set at 100 rpm. Released $^{14}CO_2$ was collected in 0.5 mL of 10% (w/v) KOH in a vial suspended in the flask (Harrison and Kruger, 2008). The vials containing KOH solution were collected and replaced with fresh vials at 2 h intervals for 12 h and again 24, 36, and 48 h after the beginning of the incubation. The radioactivity of $^{14}CO_2$ captured by each KOH vial was counted as described below (section 2.7). Ratios of $^{14}CO_2$ release from positionally labelled glucose were determined as described by Harrison and Kruger (2008). Experiments were conducted by taking four replicate flasks for each culture condition and ^{14}C feeding. The fresh weight of

cells transferred to each flask was measured from the remaining homogeneous suspension to verify the consistency of the amount of cultures used in the feeding experiment.

2.7 Radiolabelling for quantification of biosynthetic outputs

Fluxes to biosynthetic products were obtained from the redistribution of label following metabolism of [U-^{14}C]glucose by cell suspensions.

2.7.1 Incubation of Arabidopsis cell culture with [U-14C]glucose

Cells were grown with unlabeled glucose under the conditions used for the stable isotope labeling experiments. After 5 d (or 10 d for Pi- cells), 5 mL aliquots of the cell suspension were transferred to sealed 100 mL conical flasks and supplemented with 74 kBq [U-^{14}C]glucose (2 μCi per flask). After 24 h (5 d for Pi- cells), the cells were harvested by filtration and washed with 5 ml of MS medium. The cells and medium were immediately frozen in liquid nitrogen and later stored at -20°C. The production of $^{14}CO_2$ was determined in replicate flasks in which the incubation was terminated at the appropriate time by injection of 1 mL of 12 M formic acid to stop metabolism and to release dissolved CO_2 from the medium (Harrison and Kruger, 2008). Released $^{14}CO_2$ was collected in 0.5 mL (later 1 ml) of 10% (w/v) KOH in a vial suspended within each sealed culture flask.

2.7.2 Extraction of Arabidopsis cells after 14C labeling

Frozen cells were combined with 10 mL of 100% methanol at –20°C and frozen in liquid nitrogen. The sample was thawed on ice for 2 h and then centrifuged at 4,500 rpm for 20 min at 5°C. The supernatant was removed and stored on ice. The pellet was resuspended in 4 mL of cold methanol, and after storage on ice for 5 min, the suspension was centrifuged as above and

the supernatant was combined with the supernatant from the initial extraction. The insoluble cell residue was then extracted twice with 4 mL of 100% methanol for 5 min at 70°C. Samples were centrifuged to separate methanol-soluble material from the insoluble residue, and the supernatants were combined with that from the cold extractions. Finally, the insoluble residue was washed with two further 4 mL aliquots of 100% methanol, and the insoluble material was stored at –80°C.

2.7.3 Fractionation of methanolic cell extracts

Distilled water and chloroform were added to the methanolic fraction to generate a $CHCl_3$: CH_3OH: water mixture (4:4:2, v/v). The mixture was centrifuged at 3,500*g* for 10 min at 20°C to promote phase separation. The aqueous methanol phase was retained, and the chloroform phase (lipid) was rinsed with a further two washes of adjusted ratios of water and methanol before storage at –20°C and measurement of the total radioactivity. The aqueous methanol fractions from each wash were pooled with the initial aqueous methanol fraction, and this combined aqueous methanol fraction was then dried by rotary evaporation at 25°C and resuspended in 5 mL of water. Total radioactivity in this fraction was measured before further separation into basic, acidic, and neutral fractions by solid-phase extraction on Dowex ion-exchange columns (Kruger et al., 2007). The aqueous basic and acidic fractions consisting of amino acids and organic acids respectively were freeze dried, resuspended in 1 mL of water and total radioactivity measured. Each neutral fraction was freeze dried, resuspended in 2 mL water, total radioactivity measured and a fraction is separated into sucrose, glucose, and fructose using thin-layer chromatography (Scott and Kruger, 1995).

Total radioactivity in the methanol-insoluble material was determined by incubating a 100 µL aliquot of insoluble residue (from the 2 ml water suspension made before) overnight with 400 µL of Solvable tissue solubilizer (Perkin-Elmer) at 60°C followed by liquid scintillation counting. Remaining material was further analysed by autoclaving for 3 hours at 121°C, 1.04 bars pressure followed by a two step enzymatic digestion to obtain the starch, cell wall and protein components. First the insoluble material was incubated with 400 µl sodium acetate buffer containing 20 U α-amylase (Sigma-Aldrich) and 10 U amyloglucosidase (Sigma-Aldrich) for 16 h at 37°C, followed by addition of 400 µl Tris-cl (pH 7.8, 400 mM) containing 10 U protease, type XIV(Sigma) and further incubation for 16 h at 37°C. The digest was centrifuged at 13000g for 5 min and the supernatant was separated into fractions by ion-exchange chromatography (Malone et al., 2006). The resulting aqueous acidic and basic fractions represented starch, cell wall material, and protein, respectively. Radioactivity in the remaining undigested residue (cell wall) was determined by washing the residue to remove the supernatant including digested material and enzymes, followed by overnight incubation at 60 °C with tissue solubilizer (Kruger et al., 2007).

2.7.4 Extraction of ethanol, excreted organic acids and protein from the incubation medium after ^{14}C labeling

Radioactive ethanol (^{14}C) was recovered from 5 mL of cell culture filtrate by distillation after adding 2 mL of 100% ethanol to the sample to aid recovery. Distillation was repeated twice with further addition of 2 mL of 100% ethanol respectively. Recovered [^{14}C]ethanol was counted before and after evaporation. 1 ml of CCF was fractionated though anionic column to obtain the organic acids. Protein was precipitated from the remaining CCF by the addition of BSA (2 mg/ml) followed by 2% DOC (deoxycholate) and 15% v/v TCA (trichloroacetic acid).

Standing on ice overnight resulted in a clear protein pellet which was separated by centrifugation at 13000g for 15 min. The pellet was washed with 1 ml acetone and centrifuged at 13000g for 15 min. The acetone washing was repeated twice and the radioactivity on the protein pellet was measured.

2.7.5 Determination of radioactivity

Radioactivity was determined by liquid scintillation counting in a Beckman LS 6500 scintillation counter. An appropriate aliquot of each fraction was mixed with four volumes of scintillation fluid (Optiphase HiSafe) and total disintegrations per minute (dpm) counted. Each sample was counted for 5 min. In general the counting efficiency was observed to be typically greater than 90%. Prior to counting, each sample was placed in the dark at 4^{o}C overnight to reduce the effect of chemi-luminescence.

2.7.6 Separation of sugars in neutral fraction by TLC

Half of the neutral fractions obtained by ion-exchange fractionation of aqueous methanolic samples were lyophilyzed and redissolved in 50 µl of 50% ethanol. The samples (25 µl) were spotted on to a TLC plate (20 cm x 20 cm, Polygram ® CEL 300, Layer 0.1 mm cellulose MN300, Macherey-Nagel, Germany), dried and separated by solvent (formic acid: butanone: tert-butanol: H_2O in the ratio 1.5: 3: 4: 1.5 (Scott and Kruger 1995) in a chamber under the fumehood, until the solvent front was 2-3 cm from the top of the plate. The plates were removed from the tank, dried under the fumehood followed by rerunning for two more times in the solvent for better separation of sugars. The dried plates were carefully sealed with polyethylene cling film and exposed to high resolution phosphor imaging screens (Make: Fuji) placed in a radioactive cassette (McCormick et al., 2008). The exposure was done for two

different time intervals, 14 h and 24 h, after which the images on the phosphor screens were captured by Molecular imager Fx (Biorad). Bands of glucose, fructose and sucrose were clearly distinguished from the raw 1-D images and their corresponding intensities (count/mm^2) were measured using the software Quantity One (version 4.4.1, Biorad). The relative abundances of the sugars were normalised to the total volume of the neutral fraction to get the corresponding radioactive counts of each sugar.

2.7.7 Proportions of metabolites for biomass quantification

The GC-MS spectra of various fractions obtained from soluble extracts, protein hydrolysates and cell culture filtrates were used to measure the relative proportions of amino acids, organic acids and sugars. Appropriate standard curves of standard metabolites were obtained by GC-MS and then used to obtain the relative proportions of the metabolites (Appendix 2.2). The relative proportions of metabolites, their molar masses and their corresponding incorporation of ^{14}C into each metabolite was used to determine the metabolite network output fluxes as detailed in Chapter 5.

2.8 NMR spectroscopy of biogenic sample extracts

All spectra were recorded using a Varian Unity Inova 600 MHz spectrometer (Varian Inc., Palo Alto, California). The acquisition parameters were controlled by VNMRJ software (Varian). The spectrometer was manually tuned and locked on the ^2H_2O present within the sample by adjustment of Z0, power, gain and phase. Samples were auto shimmed using the routines "Z1Z2", "low non-spins", "All non-spins" and "All Zs" until adequate line width of the reference signal was obtained (< 3.3 Hz for dioxane in ^{13}C NMR and <1.5 Hz for TSP in ^1H NMR). Peaks were assigned on the basis of literature values and authentic standards.

2.8.1 One dimensional ^{13}C NMR

One-dimensional ^{1}H-decoupled ^{13}C NMR spectra were recorded at 20°C and 150.844 MHz using a 10-mm-diameter broadband probe (P6753 with tuning rod 2.4) after the freeze dried samples had been redissolved in 3.3 mL of $^{2}H_2O$ (10%) containing 10 mM EDTA, 25 mM 1,4-dioxane, and 10 mM KH_2PO_4/K_2HPO_4 buffer (pH 7.5). The acquisition parameters employed for the samples obtained from the 20% [$U^{-13}C_6$] glucose feeding experiments were a relaxation delay of 6 s, a 90^{o} observe pulse of 24 μs, an acquisition time of 1.016 s and Waltz 16 ^{1}H decoupling with nuclear Overhauser effect. All the other samples obtained from specifically labelled glucose ([1-^{13}C] or [2-^{13}C]) feeding experiments were subjected to a relaxation delay of 19 s, a 90^{o} observe pulse of 12.50 μs, an acquisition time of 1.015 s and Waltz 16 ^{1}H decoupling without nuclear Overhauser effect. Large numbers of scans were generally required to generate spectra of sufficient quality for quantitative analysis of the peak intensities, and spectra were accumulated in blocks (8 or 10 blocks @ 1024 scans/block) to check for unwanted spectral changes. For neutralised perchloric acid extracts (containing soluble metabolites), a total acquisition time of about 20 h or 46 h was used for the spectra of 6 s or 19 s relaxation delay respectively. Spectra of starch digests were acquired for 6 to 10 h. Spectra were also obtained from the anionic, cationic and neutral fractions of the soluble extracts (for 6 h to 12 h) for the detection of amino acids, organic acids and sugars respectively.

2.8.1.1 *Spectral processing*

Spectra were processed using NutsPro software (Acorn NMR, Professional version 20070315). All the blocks of each spectrum were summed, zero filled, baseline corrected and a line-broadening (exponential multiplication) of 1 Hz was applied. After Fourier transformation, phase correction (auto and manual) was done followed by linear regression to flatten the

spectral baseline. Chemical shifts were measured relative to the 1,4-dioxane signal at 67.30 ppm. Positional isotopomers were specified using 0 to denote ^{12}C, 1 to denote ^{13}C, and X to denote either ^{12}C or ^{13}C.

2.8.1.2 *Calculation of peak area and errors*

The relative abundances of positional isotopomers and the errors on them were measured as suggested by Williams et al., 2008). Spectral deconvolution (line-fitting) using the line-fitting (LF) subroutine and absolute peak heights calculation using the define peaks (DP) subroutine were carried out to obtain the peak areas and their relative errors as detailed with an example in Chapter 3 (NMR data handling for ^{13}C MFA).

2.8.2 One dimensional ^{1}H NMR

One-dimensional ^{1}H NMR spectra of CCF and perchloric acid extracts obtained from feeding experiments were recorded at 20°C and 600 MHz using a 5-mm diameter $^{1}H\{^{13}C/^{15}N\}$ triple resonance probe (P6173_3res). The acquisition parameters employed were a 90° pulse angle, a 10 ppm spectral width, relaxation delay of 0.02 s, a VNMRJ "Presat" delay of 1.98 s, and an acquisition time of 4 s. The presaturation delay was used for water suppression. The samples were added to a known amount of $^{2}H_2O$ containing Trimethyl silyl propionate (TSP, final conc 1 mM; final vol 800 µl), and the spectrum acquired for about 6 to 12 min. The ^{1}H FIDs obtained were processed by NUTS software. They were zero filled, baseline corrected and a line-broadening (exponential multiplication) of 1 Hz was applied. After Fourier transformation, phase correction (auto and manual) was done followed by linear regression to flatten the spectral baseline. Chemical shifts were measured relative to the TSP signal at 0 ppm. The peak abundances were obtained by manual integration.

2.8.3 One dimensional ^{14}N NMR

^{14}N NMR was used to quantify the nitrate and ammonium levels in the CCF or sample extracts. The broad band probe ^{14}N NMR spectra were recorded at 20°C and 280 MHz low pass filter using a 10-mm-diameter broadband probe (P6753 with capacitors 36D in channel 1 and "S" in channel 2). NaOCN (30 µl of 1 M) was used in a capillary as an internal standard. The parameters used for acquisition were 128 scans at 4 sec relaxation delay, 25 µs or 90 degree observe pulse. The spectral width was 429 to -30 ppm.

2.9 Gas chromatography-mass spectrometry

The GC-MS measurements were performed on an Agilent 79890 GC coupled to an Agilent 5975 quadrupole MS detector, electron impact ionisation (70 eV) equipped with either a Varian VF5-ms column (30 m, 10 m guard column, 0.25 mm inner diameter) or an Agilent HP 5-ms column (30 m, 0.25 mm inner diameter) at the facility in the Department of Plant Sciences, University of Oxford, UK. Samples (5– 50 µl) to be derivatized were dried in 1.5 ml eppendorf tubes using a speed-vac system to ensure complete removal of water. Authentic commercial standards were analysed concurrently by GC-MS, to obtain the elution profiles and to identify mass fragments of all metabolites.

2.9.1 Derivatization of amino acids and organic acids

The amino acids and organic acid extracts were derivatised by TBDMS (Antoniewicz et al., 2007) and/or TMS (Lisec et al., 2006). To obtain the TBDMS derivatives, the dried samples were first dissolved in 25 µl of pyridine and incubated at 37^{o}C, shaking at 900 rpm for 30 min. Then 35 µl of MtBSTFA + 1% t-BDMCS (*N*-methyl-*N*-(*t*-butyldimethylsilyl) trifluoroacetamide + 1% *t*-butyl-dimethylchlorosilane, Regis Technologies Inc*)* was added and

the mixture was incubated at 60^{o}C, shaking at 900 rpm for 30 min. To obtain the TMS derivatives, 43 µl of pyridine was added to each dried sample and incubated at 37^{o}C for 1 h at 900 rpm. Then 60 µl MSTFA (*N*-methyl-*N*-(trimethylsilyl)trifluroacetamide (HiChrom)) was added and the mixture was incubated at 37^{o}C for 30 min, shaking at 900 rpm. The derivatised samples were centrifuged for 10 min at 13000 rpm to pellet any insoluble material and the supernatant were transferred to 8 mm glass vials (Chromacol) and sealed with a septum cap.

2.9.2 Derivatisation of sugars (MeOX TMS method)

50 µl of 20 mg ml^{-1} methoxyamine hydrochloride (Sigma) in pyridine was added to the sample and incubated at 37^{o}C, shaking at 900 rpm for 2 h. Then 70 µl MSTFA was added and the mixture was incubated at 37^{o}C for 30 min at 900 rpm. Samples were centrifuged for 10 min at 13000 rpm to pellet any insoluble material and the supernatant was transferred to glass vials.

2.9.3 Optimised injection and temperature parameters

Before the samples were run, the mass spectrometer was calibrated using the Agilent autotune method. The injected sample volume was 1 µl, and this was injected in purged splitless mode. For separation of the TBDMS derivatised metabolites, the injection port temperature was 230^{o}C. The oven program was 120^{o}C for 5 min, then a 4^{o}C/min ramp to 270^{o}C, held for 3 min, then a 20^{o}C/min ramp to 320^{o}C and held for 1 min. The carrier gas flow was helium at 1.3 ml min^{-1}. For separation of the TMS derivatised amino acid and organic acid samples the oven temperature was initially held constant at 70^{o}C for 5 min then ramped up to 280^{o}C at 5^{o}C/min and held there for 3 min before being decreased to 70^{o}C at 120^{o}C/min. The carrier gas flow was helium at 0.6 ml min-1. For MeOX-TMS derivatised sugar samples the oven temperature was initially held constant at 70^{o}C for 5 min then ramped up to 350^{o}C at 5^{o}C/min. The

temperature was then immediately decreased to 330^{o}C and held there for 5 min before being decreased to 70^{o}C at 120^{o}C/min. The carrier gas flow was helium at 0.6 ml min^{-1}.

Mass spectra of all the samples were acquired for m/z 146-600 by scanning (at 4.38 or 3.19 scans s^{-1}) and/or single ion monitoring (SIM) mode (dwell time of 10 to 30 ms for each group of ions). A solvent delay of 10 min was set so that the mass spectrometer was turned on after elution of the bulk of the solvent from the column.

2.9.4 Metabolite identification and mass isotopomer data handling

Analysis of ^{13}C incorporation in each metabolite is crucial for obtaining the mass isotopomer abundances. The raw GC-MS spectra need to be handled carefully for accurate assessment of mass isotopomer distributions in metabolites (Antoniewicz et al., 2007). The detailed procedure of handling GC-MS data for ^{13}C MFA studies is mentioned in Chapter 3.

2.10 Metabolic modeling

Metabolic modeling was performed using 13C-FLUX (version 20050329; Wiechert et al., 2001) according to the protocol outlined in Chapter 3. After defining the metabolic network, the free fluxes were fitted to the biomass measurements and labeling data using the Donlp2 (P. Spelluci, Technische Universität Darmstadt, Germany) routine in 13C-FLUX. This sequential quadratic programming algorithm is superior to the CooolEvoAlpha evolutionary algorithm in 13C-FLUX, in terms of both computational efficiency and avoidance of local minima (Wiechert et al., 2001). The optimizer was used in conjunction with Monte Carlo simulations and bootstrap sampling of the isotopomer abundances based on the measured values and their deviations to generate solutions for the flux map. The bootstrap Monte Carlo sampling method

was invoked by the command: Donlp2 "model file name".ftbl -a 0.00001 –m "number of simulations" -bootstrap -logAllF > "output file name".txt&watch "grep Resi "output file name".txt|tail –20". Each simulation generated a set of free fluxes after minimizing the sum of squared standardized differences (the residuum) between the experimental measurements and their predicted values. The detailed instruction to use the Monte Carlo simulations along with other linux tool (Feasible infeasible residuum extractor) to collect the flux estimates is available in Appendix 2.4. Principal component analysis of the free fluxes obtained from multiple simulations was performed using SIMCA-P 11.5 (Umetrics) to identify the set of best fit feasible solutions with low values of the residuum. This set of outputs was used to generate a mean flux solution that was then used as a starting point for a further optimization run with Donlp2. EstimateStat, the statistical analysis component of 13C-FLUX and/or a non-linear statistical method (Antoniewicz et al., 2006) was used to compare the merits of different modeling schemes and to assign confidence intervals to the flux values (See Chapter 3 and Appendix 2.3 (custom code for non-linear statistical analysis)). Other specific modeling activities like adding proxy reactions to account for pre-existing pools (Chapter 3), MANOVA of Monte Carlo flux estimates obtained from simulations after Jackknife of the datasets (by removing one dataset at a time) are mentioned in Chapter 5.

3 *In silico* strategies, experimental considerations and data handling of isotopomers for ^{13}C-based network flux analysis

3.1 Introduction

Metabolic flux analysis based on ^{13}C labeling experiments (CLE) followed by mathematical analysis of the labeling data is currently the principal method used to quantify intracellular fluxes in the central metabolic network in heterotrophic cells and tissues. The strategy of CLE involves the integration of the biological, analytical, experimental, mathematical and computational principles necessary for MFA (Wiechert et al., 2001). Mathematical modeling, simulations, experimental design and data handling are vital scientific tools that need considerable attention to develop robust methods for understanding metabolism at the systems level. The specific aims of this chapter are (i) to establish and validate robust *in silico* strategies for ^{13}C-based network flux analysis using the software framework 13C-FLUX® (version 20050329; Wiechert et al., 2001). Extensive tests were conducted to validate several approaches for determining flux confidence intervals. The *in silico* modeling strategies and the utility of the software framework 13C-FLUX for analysing steady-state labelling experiments are discussed (see Section 3.2); and (ii) to comprehensively analyse the factors that influence the outcome of MFA and establish best practices. The major factors studied are methods for (a) data handling of isotopomer measurements (obtained by GC-MS and NMR; see Sections 3.3 and 3.5) and (b) strategies to account for pre-existing unlabelled pools (see Section 3.4) of metabolites that are observed to be common in labelling experiments in plant cells.

3.2 *In silico* strategies for ^{13}C-metabolic flux analysis

3.2.1 Example metabolic network

To illustrate the strategies developed, extensive tests were conducted on a simple example metabolic network (PPP model; Fig 3.1) that is supplied with the 13C-FLUX software package. The stoichiometry and atom transformations for the reactions, ^{13}C-substrate(s) fed and other parameters of the model network are presented in Appendix 3.1 in "FTBL" file format. We assumed the system to be at isotopic steady state and set the free fluxes in the network to the values shown in Fig 3.1. The net uptake rate, *Vupt* was constrained to 1. Simulation of the network (using the routine *Cumonet)* with the predefined fluxes and specifically labelled glucose substrate as isotopic tracer resulted in the isotopomer and cumomer abundances of the metabolites. Now the objective was to test how well we could estimate the expected fluxes and calculate the flux confidence intervals. Using 13CFLUX, which modeling strategies would be statistically robust? To be consistent with the commonly used analytical techniques, both NMR and GC-MS labelling measurements (Table 3.1) were selected to fully exploit the capacity of 13C-FLUX to handle both types of data.

3.2.2 Bootstrap Monte Carlo simulations can be used to explore the flux space

The extent to which labelling information in metabolites defines fluxes throughout the network was examined by obtaining flux solutions using the fractional abundances of the cumomers or mass isotopomers of the metabolites presented in Table 3.1. Random free fluxes were fitted to the labeling data using the *Donlp2* (P. Spelluci, Technische Universita¨t Darmstadt, Germany) routine in 13C-FLUX. This sequential quadratic programming algorithm is superior to the evolutionary algorithm (*CooolEvoAlpha*) in 13C-FLUX, in terms of both computational

efficiency and avoidance of local minima (Wiechert et al., 2001). The optimizer (*Donlp2*) was used in conjunction with Monte Carlo simulations with and without the bootstrap sampling of the isotopomer/cumomer abundances. Two important features of the Monte Carlo sampling method have been elucidated and a strategy has been developed for optimal analysis.

First, the Monte Carlo sampling method with and without the bootstrap sampling of the label measurements differed significantly, with bootstrap sampling the preferred method. Without the bootstrap sampling, each simulation generated a population of pseudo-measurement data points which were observed to differ significantly from the actual measurements **(**Fig. 3.2 A,B,C**)**. In contrast simulations with bootstrap sampling provided an effective method for statistical sampling of populations in which the pseudo-measurements were predominantly within the errors assigned to the actual measurements **(**Fig. 3.2 A,B,C**)**. Further analysis has pointed to the fact that without bootstrap the populations of pseudo measurements are not derived from the actual measurements, rather they come from the simulated measurements. This is undesirable as the simulated measurements may not represent the actual measurements if the model starts parameter fitting with poor starting points.

Multiple simulations with the pseudo-measurements propagate the errors in the measurements to the fluxes. The bootstrap Monte Carlo sampling method was invoked by the command: *Donlp2* "model file name".ftbl -a 0.00001 -m "number of simulations" -bootstrap -logAllF . "output file name".txt&watch "grep Resi "output file name".txt|tail 20". Each simulation generated a set of free fluxes after minimizing the sum of squared standardized differences (the residuum) between the representative measurements and their predicted values (Fig. 3.3). In practice thousands of Monte Carlo simulations with bootstrap sampling are required for optimal flux analysis.

Secondly, it was observed that the Monte Carlo simulations generated flux estimate solutions which fell into three broad categories (a) infeasible solutions which did not satisfy all the defined network constraints; (b) feasible solutions with high residua (local minima); and (c) feasible solutions with low residua (best fits) (Fig. 3.4). The current analysis on the simple PPP model did not generate any infeasible solutions, but they were observed while modeling the published complex metabolic networks of sunflower embryos (Alonso et al., 2009) and *Arabidopsis* cell suspension cultures (Masakapalli et al., 2010). The residua for the feasible solutions vary over a wide range, and a subset of the solutions corresponding to a cutoff residuum value (threshold) can be selected by visual inspection of the distribution of residua or by principal component analysis of the mean-centered, unit variance-scaled flux solutions (Masakapalli et al., 2010). The infeasible and local minima solutions, if any, were rejected. The best fit flux estimates were considered as reliable solutions for deriving individual fluxes and the confidence regions of the fluxes (Kruger and Ratcliffe., 2009) as discussed in the following sections.

Monte Carlo simulations with bootstrap sampling of the isotopomer abundances provided an effective method for exploring the flux space that was consistent with the observed redistribution of label. Stochastic sampling of the starting fluxes is routine in the modeling process (Schwender et al., 2006), and combining this with statistical sampling of the data points improves the fitting procedure (Kelly et al., 1990; Namba, 2004). Surprisingly, the availability of a subroutine (i.e using bootstrap Monte Carlo) to exploit this method in the commonly used 13C-FLUX software appears to have been overlooked.

3.2.3 Mean fluxes of Monte Carlo (MC) flux space aid in obtaining the global best fit solution

The bootstrap MC simulations of the PPP model (with 99% [1-^{13}C] glucose as substrate and the corresponding measurements (Table 3.1)) generated many feasible solutions with different flux estimates (Fig 3.5) and residua (Figs 3.5, 3.6). The estimates of four free fluxes (Vemp1_NET, Vppp2_XCH, Vppp4_XCH, Vppp6_XCH) were tightly distributed, whereas the two fluxes (Vppp3_XCH and Vppp5_XCH) were much more scattered. The number of Monte Carlo simulations needed to explore the flux space depends on the network size, the precision of the measurements and number of free fluxes. About hundred simulations were found to be sufficient for the PPP model which had six free fluxes.

The MC flux space of each flux was obtained by calculating the mean and the standard deviations of the selected flux estimates from the distribution. The mean solution could not be considered as the global best fit since it differed from the expected fluxes (Fig 3.7). However it was observed that these mean flux values provided an ideal starting point for re-optimising the fit to obtain the optimal flux values. Therefore, to reduce the numerous feasible solutions to a global best fit, the mean values from the MC flux space were used as starting free flux values for a final run of the optimizer with the actual measurement data set (Fig 3.8). This run was performed without bootstrap Monte Carlo sampling, and it resulted in a global best fit solution with the lowest residuum (Fig 3.6). The global best fit flux estimates agreed well (Fig 3.7) with the expected values. The confidence limits for the global best fit fluxes were obtained by various strategies as discussed in section 3.2.4.

Often, optimization with random starting free flux values may result in poor flux estimates. The above method exploits the bootstrap Monte Carlo simulations to define the flux space and then accurately find the optimal flux estimates.

3.2.4 Strategies for the determination of fluxes and their confidence intervals

3.2.4.1 *Linear approximation method (using EstimateStat)*

The measurement sensitivity matrix (dMeas/dFlux) shows the effect on the measurements due to change in the free fluxes, and it is used to compute a covariance matrix by linear approximation of the fluxes (Mollney et al., 1999; Wiechert, 2001). From this the confidence intervals of the flux estimates are obtained. In 13C-FLUX, the routine *EstimateStat* computes the confidence interval from the sensitivity and covariance matrices. In practice, net fluxes are usually well defined statistically, but the exchange fluxes (due to non linearity) can typically only be determined within an order of magnitude. From the current analysis of the PPP model, it is observed that the best fit flux estimates were similar to the expected values, that Vemp1_NET and certain exchange fluxes were well defined, but that the fluxes Vppp3_XCH and Vppp5_XCH were poorly determined (Table 3.2, *EstimateStat* result). It should be noted that the confidence limits of well defined exchange fluxes should ideally be within the limits of 0 and 1, but that was not the case here since the upper limits of Vppp2_XCH, Vppp4_XCH were more than 1, which is mainly attributed to the linear approximations. Can we overcome this limitation by other non-linear methods?

3.2.4.2 *Confidence intervals from flux estimates obtained from Monte-Carlo simulations*

Alternatively, Monte-Carlo simulations can be used to generate confidence limits in a non-linear statistical analysis (Press et al., 1992; Witzel et al., 2007; Joshi et al., 2006; Antoniewicz

et al., 2006). This method is known to quantify both the net and exchange fluxes more reliably (Weitzel et al., 2007). A large number of Monte Carlo simulations (with bootstrap) generated a distribution of flux values, which were first sorted in ascending order. The 68% confidence intervals were obtained by discarding 16% of the solutions from top and bottom of the distribution. The minimum (min) and maximum (max) range of fluxes obtained are presented in Table 3.2. and the results show the asymmetric confidence intervals for the exchange fluxes. It was observed that the confidence intervals decreased as the number of simulations increased but the procedure was computationally demanding.

3.2.4.3 *Non-linear analysis method for determining true confidence intervals (CI)*

Antoniewicz et al. (2006) developed an efficient algorithm for accurate determination of flux confidence intervals by a non-linear statistical technique. In this method the confidence intervals of each flux are deduced individually by determining the minimised sum of squared residuals as a function of the flux value. A threshold value for the change in residuum (or χ^2) corresponding to 68% (0.9889) and 95% (3.84) are used as cut off values to obtain the confidence limits of the fluxes. This feature is not present in the current version of 13C-FLUX, so a tool was developed that operates in conjunction with 13C-FLUX to deduce the confidence limits (Appendix 2.3 for code).

Starting with the global best fit flux estimates (Section 3.2.3), the flux of interest was constrained to a series of values, while all the other fluxes were determined by re-optimisation. This was repeated until the flux reached infinity or the residuum obtained was greater than the threshold set for the confidence limits (Fig 3.9). The residuum obtained was plotted against the constrained values for the flux, allowing the minimum and maximum confidence limits to be deduced. The confidence intervals (CI) obtained by this method are presented in Table 3.2, and

as expected they are similar to the values obtained with the Monte Carlo method (Antoniewicz et al., 2006).

3.2.5 Sub-grouping of the ^{13}C NMR measurements improves the fitting procedure

In principle, ^{13}C NMR measurements can be grouped and scaled separately for each carbon atom of a metabolite (Mollney et al., 1999). To validate the influence of grouping the data, flux analysis studies were conducted with the published ^{13}C NMR datasets of heterotrophic *Arabidopsis thaliana* cell suspension cultures (Masakapalli et al., 2010). Label measurements for a particular metabolite were initially grouped together for the fitting process, but subsequently it was established that better fits (lower residua) could be obtained by treating ^{13}C peaks with multiplet structures as separate subgroups. For example, the measurements of a four carbon compound xx1x, x11x & xx11 (carbon 3 and its satellites) are placed in a subgroup, similarly x1xx, 11xx & x11x (carbon 2 and its satellites) in another subgroup and so on. In one fit it was observed that the total residuum contributed by leucine measurements reduced by 20-fold on sub-grouping. Simulations in 13C-FLUX showed that the reduction in the residuum could be attributed to the inclusion of additional scaling factors, compensating for discrepancies in the relative intensities of the signals from different carbon atoms within a metabolite, and that the subgrouping procedure did not compromise the accuracy of flux estimation. Subgrouping multiplet signals avoids the need to establish the relative intensities of multiplets from the same metabolite, and it is used routinely when two-dimensional NMR methods, which often mask the true relative intensities, are used to analyze [$^{13}C_6$]glucose labeling experiments (Sriram et al., 2004).

Further *in silico* tests were carried on the PPP model to study the effect of sub-grouping on the fluxes and its potential for implementation. Three different cases of representation of additional cumomers of the metabolite, GA3P were tested (Table 3.3). In Case 1, the actual measurements were used. In case 2, the GA3P cumomers of carbon-2 (#x1x, #11x+#x11) and carbon-3 (#xx1, #x11) were hypothesized to be noisy (differ by 20% from expected values) and were handled as a single group. This was done to represent a condition where the actual intensities in different carbons of a metabolite were perturbed due to analytical differences (e.g differences due to relaxation delays in NMR). In Case 3, the GA3P cumomers from case 2 were sub-grouped based on carbon numbers.

Perturbed measurements influenced the flux estimates (Vppp2_XCH, Vppp3_XCH, Vppp4_XCH and Vppp5_XCH) and the expected global best fit solution could not be obtained (Table 3.4). This showed that systematic errors in the measurements could result in poor flux estimates and this needs to be overcome by appropriate data handling. The sub-grouping of the cumomer measurements did not compromise accurate estimation of the fluxes (Table 3.4) but reduced the errors between individual subgroups (Table 3.5). This was achieved due to the inclusion of the additional scaling factors for each carbon group. In conclusion, sub-grouping should be used for accurate scaling of measurements and hence the strategy was implemented wherever necessary in subsequent analyses.

3.2.6 Strategies for optimal substrate(s) design and detecting key measurements from the sensitivity matrix (dcum/dflux)

There is a strong relationship between the flux and label redistribution in the network and therefore precise experimental measurements are needed for reliable flux estimation (Roscher

et al., 2000; Wiechert et al., 2001). *A priori* identification of optimal substrate(s) and crucial measurements could save undue experimental efforts. The 13C-FLUX routine *Estimatestat* generates a large sensitivity matrix (dcum/dflux) which represents the sensitivity of every cumomer to the free fluxes. The accuracy of the sensitivity analysis depends on the free fluxes defined (which are unknown *a priori*) and the network stoichiometry, but it is independent of the actual measurements. A software tool (not shown) was developed which uses the outputs of *Cumonet* (to generate the cumomer list) and *Estimatestat* (to analyse the dcum/dflux sensitivity matrix) to derive the most significant cumomers that best define each individual flux. The analysis can be repeated with different input substrates (labeled at different positions) in order to detect optimal substrate(s) and the key measurements to define fluxes of interest.

The tool was tested on the PPP network in which the exchange fluxes Vppp3_XCH (Rul5P ⇔ Rib5P) and Vppp5_XCH (Xul5P + Rib5P ⇔ GA3P + Sed7P) were poorly defined. The dcum/dflux matrix analysis implied that the cumomer measurements of Rib5P (#1xxxx and #xxxx1), Sed7P (#1xxxxxx and #xx1xxxx) and many others were sensitive to the rates of Vppp3_XCH and Vppp5_XCH. Adding these sensitive cumomers of Rib5P and Sed7P to the model greatly improved the definition of the two exchange fluxes (Fig 3.10). Thus the sensitivity matrix (dcum/dflux) has the potential to improve the efficiency of flux analysis studies by identifying critical cumomer measurements for the flux solution.

3.2.7 Strategy of multi-model simulations improves the reliability of fluxes obtained

Selection of the labelled substrate plays a crucial role in ^{13}C MFA. It has been found that a labelled substrate may not be informative for monitoring a part of the network particularly well and that parallel experiments with other substrates are essential (Kruger et al. 2003, Schwender

et al., 2006). The parallel experiments with different substrates can be simulated together (known as multi-models) to increase the statistical quality of the flux estimates and also to determine the fluxes that are poorly defined with a single labelled substrate (Schwender et al., 2006, Alonso et al., 2007). Also optimal substrate design studies on a metabolic network of *Brassica napus* embryos showed improvement in the information content and flux determinability when multi-models with different substrates were used (Libourel et al., 2007).

In order to combine separate labeling experiments into a multimodel (e.g. a double model (Fig. 3.11)), all the reactions of the basic network model have to be repeated with different flux names and metabolites names (e.g. V1<flux name>, V2<flux name> for two different label experiments). The identical fluxes are constrained to be equal in the "EQUALITIES" section of the .ftbl file and the different input labelled substrates are defined (Schwender et al., 2006, Alonso et al., 2007). The optimisation routines fit all the measurements simultaneously to give optimal flux estimates which best fit the combination of the two experiments. However, the computational speed decreased in multimodels which is attributed to the large network and additional constraints.

To test the potential of the multi-model approach, single and double models of the PPP network were constructed. Flux analysis was conducted using the measurement datasets presented in Table 3.1 with either [1-^{13}C]glucose (1) and/or [2-^{13}C]glucose (2) as substrate. Various combinations of substrates (1,1; 1,2 and 2,2) were tested and the results are presented in Table 3.6. The flux estimation and the confidence regions of the fluxes improved in the double models (e.g. Vppp6_XCH, Fig 3.12), which is due to the additional sensitivity of the measurements. Within the double models, feeding with different substrates (Double[1][2]) was often better than the replicate models (Double[1][1] or Double[2][2]) as shown in Fig 3.12 for

Vppp6_XCH. Flux determinability depends on the sensitivity of the measurements and hence improves with additional measurements and multi-models provide additional constraints for optimal flux estimation.

3.2.8 Summary of modeling design for ^{13}C MFA using 13C-FLUX software and other tools developed

In summary, ^{13}C MFA using 13C-FLUX requires several iterative steps (Fig. 3.13, Fig. 3.14):

Step 1: Establish a model network (with biochemical reactions along with their carbon transitions).

Step 2: Provide the parameters (experimental and assumed) in textual format (FTBL file), including isotopomer measurements, label input, free and constrained fluxes, inequalities and network stoichiometry.

Step 3: Run the simulations using the *CumoNet* routine, which generates the matrices and vectors to simulate the metabolic activity of the network and to check the feasibility of the model. It also generates the simulated isotopomer and cumomer abundances for a defined set of free fluxes.

Step 4: Estimate fluxes through least square fitting of experimental and predicted measurements by a range of optimisation tools including an evolutionary algorithm (CooolEvoAlpha) and sequential quadratic programming algorithm (*Donlp2*, which includes the facility for Monte Carlo simulations with bootstrap sampling of isotopomer abundances). Ideally the optimization routine with bootstrap Monte Carlo simulations (Fig 3.3) is preferred to generate a correspondingly large set of flux solutions, each with a set of predicted labeling

measurements. Some of the solutions fail to satisfy constraints in the model and are identified as infeasible solutions; but the majority of the solutions are feasible, and these can be extracted from the 13C-FLUX output files with a customized software tool (Appendix 2.4). The procedure for determining the global best fit for the model and the optimal flux values is summarized in figure 3.8.

Step 5: Statistical evaluation of flux estimates (section 3.2.4, Fig. 3.8) using the output from *EstimateStat*, Monte Carlo simulations and non-linearised statistics. The ultimate result is a flux map of the network under steady state conditions.

3.3 GC-MS data handling for ^{13}C-based metabolic flux analysis

Mass isotopomer distribution (MID) of cellular metabolites obtained from stable-isotope labeling experiments provide information for metabolic flux analysis. The reliability of flux estimates depend on the accuracy and precision of the mass isotopomers that are measured by gas chromatography-mass spectrometry (GC-MS). Even small errors in the MID may contribute towards large errors in flux estimates (Antoniewicz et al., 2006). Therefore, the MID of all the metabolite fragments should be assessed for potential errors. This section focuses on a thorough consideration of appropriate methods for handling GC-MS data in order to accurately assess the abundance, minimize the potential errors over the mass isotopomers, and also validate the reliability of fragments. For the current study, representative GC-MS spectra of authentic standards and metabolites obtained from *Arabidopsis* suspension cultures fed with ^{13}C and unlabelled glucose were used.

Based on the analysis below, it was observed that considerable attention on spectral processing methods improves the accuracy and precision of the mass isotopomer abundances. The key considerations include

i) GC-MS profiling of metabolites

ii) Baseline correction of the spectra

iii) Choice of MID source, qualitative and quantitative assessment of the fragments

- *Extraction of MIDs (Individual scans and averaged-scans of GC peak)*
- *Correction of natural abundance and calculation of fractional enrichments*
- *Assessment of the quality of MID and validation of the fragments*
- *Concentration effect of metabolites on the MID*

- *Defining appropriate errors on mass isotopomer measurements*

3.3.1 GC-MS profiling of metabolites

Profiling of derivatised metabolites is the first step for their accurate identification which are typically based on elution times in the gas chromatogram (GC) and the fragmentation pattern they form in the mass spectroscopy (MS) due to ionization. The GC-MS spectra from authentic standards of amino acids (Appendix 2.2, Appendix 3.2), organic acids (Appendix 2.2) and few sugars derivatised with either TBDMS, TMS or MeOX-TMS were acquired to record the elution times and the fragmentation patterns (Appendix 3.2). The spectral peaks were visualised and processed using the software packages AMDIS (NIST, National Institute of Standards and Technology, Maryland) and Chemstation (Agilent). The identification and elution times of each derivatised metabolite was established by comparison of the mass isotopomer distribution (MID) of each peak with an extensive database of TMS-derivatized plant metabolites (Kopka et al., 2005) using the software MSSearch (NIST, Golm database) and authentic standard libraries of TBDMS-derivatised amino acids and organic acids generated in our laboratory and literature.

To illustrate the profiling of metabolites, GC-MS analysis of a TBDMS-derivatised alanine standard was considered (Fig 3.15). Alanine was derivatised to alanine-2TBDMS with one TBDMS group bonding to COO^- and another to NH^+ . This derivative eluted at 13.34 – 13.74 minutes during GC. The electron impact ionization during the subsequent MS analysis resulted in fragment ions due to loss of molecular groups from the carbon or non-carbon skeleton (Table 3.7). The resulting mass spectrum consisted predominantly of fragment ions with m/z values of 260 ($[M-57]^+$), 232 ($[M-85]^+$) and 158 ($[M-159]^+$). Minor fragments with m/z values

of 302 ($[M-15]^+$ and/or $f302^+$) and 317 ($[M-0]^+$, complete molecule) were also detected. These fragmentation patterns are typical for TBDMS-derivatised amino acids (Wahl et al., 2004, Wittmann et al., 2002). Each fragment retains specific carbon atoms of the metabolite. For example, the $[M-57]^+$ fragment retains all the carbons of alanine and the relative abundances of the mass isotopomer pools m+0, m+1, m+2 and m+3 are obtained from processing of the intensities of corresponding mass ranges (m/z 260-263).

3.3.2 Baseline correction of spectra

The raw GC-MS spectra need to be baseline corrected at first for accurate assessment of mass isotopomer distributions in metabolites. The MIDs without baseline correction are known to produce errors in the isotopomer estimates (Antoniewicz et al., 2007). Therefore, prior to analysis, each GC/MS spectra was baseline corrected in order to reduce the influence of background noise. The software MetAlign (Lommen 2009) with its default parameters was used for baseline correction.

3.3.3 Choice of MID source, qualitative and quantitative assessment of the fragments

Analysis of the extent of ^{13}C incorporation into each metabolite is crucial to obtain the mass isotopomer abundances needed for flux analysis studies. A metabolite peak in GC/MS is formed from successive mass scans of the sample eluting from the GC. However, different scans may be influenced by different levels of noise depending on the signal strength and the potential presence of overlapping fragments from other metabolites. Wahl et al. (2004) showed that for flux analysis studies mass isotopomers have to be validated across individual scans and the most appropriate scan (or scans) used in subsequent analysis. Other studies have argued that influences due to differential interactions of heavy atoms (^{13}C) versus normal atoms within

the GC column could result in isotopic fractionation and thereby influence the relative mass isotopomer abundances across an elution profile and for this reason have advocated using the average of all scans across a peak (Allen and Ratcliffe, 2009). To assess the relative merits of these contrasting strategies for obtaining MIDs, comparisons from individual scans or values averaged from scans across an elution profile was studied. The data obtained from GC-MS analysis of TBDMS derivatised standards and biogenic metabolites obtained from 20% [$^{13}C_6$]glucose fed *Arabidopsis* suspension culture were used.

3.3.3.1 ***Extraction of MIDs***

Mass isotopomer distributions (MID) are basically detected in the mass spectra that contain series of masses (m/z) and their corresponding ion counts (abundances). The GC peaks of each metabolite was obtained using the software Chemstation (Agilent). Data were extracted either manually or in batches of samples where overlaid chromatograms were highly reproducible, using a cutomised macro to automatically output the m/z and average ion counts of all the metabolites whose peak boundaries were pre-defined based on the elution times of the GC. The extracted MID of each metabolite was saved as a tab delimited file for further processing.

For illustration, the GC-MS spectra of a natural alanine standard and biogenically derived [U-^{13}C] alanine are presented (Figs 3.16A, 3.18A**).** For each sample, the individual scans extracted for further analysis are indicated by arrows. Representative MID (Figs 3.16B, 3.18B) extracted from individual scans or the averaged-scans show the fragment ions of alanine-2TBDMS with major peaks at m/z 158, 232, 260, and 302. Differences between the fragments of unlabelled and ^{13}C labeled samples can be clearly distinguished (Figs 3.16B, 3.18B).

3.3.3.2 *Correction of isotopomer profiles for natural abundance and calculation of fractional enrichments*

Correction of mass effects due to natural abundance of stable mass isotopes of common atoms (^{13}C, ^{2}H, ^{15}N, ^{17}O, ^{18}O, ^{29}Si, ^{30}Si, ^{33}S etc) is a prerequisite for determining the extent of labeling of metabolites from a supplied [^{13}C] substrate. The MIDs of each fragment ion obtained from individual scans and the averaged mass spectra were corrected for the presence of naturally occurring heavy isotopes attached to the carbon backbone of the derivative using mass correction software MSCorr (Wahl et al., 2004). In addition to mass correction, this tool estimates the measurement noise and standard deviation of each mass isotopomer by taking signal strength and other parameters into consideration. The identity and chemical formulae of fragment ions required for MSCorr are obtained from the fragmentation pattern of the derivatised metabolite deduced manually and reported in the literature (Leimer et al., 1977, Huege et al., 2007, Wittmann et al., 2002). From this information MSCorr is used to deduce the absolute and proportional abundances of mass isotopomers of all the metabolites along with their errors (Wahl et al., 2004).

3.3.3.3 *Assessment of the quality of MID and validation of the fragments*

The accuracy and precision of the fragment ions were evaluated by comparing the measured mass isotopomer proportions with the theoretically anticipated values. The fraction of ^{13}C in an unlabelled fragment is expected to be present at natural abundance (~1.13% per carbon) which can be calculated based on the number of carbons. For example, in a three carbon molecule, the expected fractions of m+0, m+1, m+2 and m+3 are 0.966, 0.034, 0.000 and 0.000, respectively (when expressed to a precision of three decimal places). The most abundant isotopomer (m+0), as a representative of a fragment can readily be used for accuracy

assessment. The precision of the measurements are estimated from the errors derived using MSCorr tool (Wahl et al., 2004) and are expressed as mol %.

Analysis of the standard alanine-2TBDMS fragments ions (Figs 3.16, 3.17) obtained from individual (scans 530 to 571) and the averaged-scans illustrate the common features encountered in the qualitative and quantitative assessment of GC-MS data. In particular, these data show:

- The mass isotopomer fractions of the complete derivatised alanine molecule $[M-0]^+$ obtained from all scans were found to deviate significantly from the distribution expected values for natural abundance. Moreover, the errors calculated for these measurements are large (typically greater than 50% of the estimate). For these reasons information obtained from the intact molecule $[M-0]^+$ was considered unreliable and thus rejected.
- Other fragment ions $[M-15]^+$, $[M-57]^+$, $[M-85]^+$, $[M-159]^+$ and $f302^+$ showed mass isotopomer distributions much closer to the expected values and were therefore further analyzed for their accuracy and precision (Fig 3.17).
- Although the mass proportions derived from individual and averaged scans of the fragment $[M-159]^+$ were consistent (within 0.02 - 0.36 % noise), it was also rejected due to poor accuracy determined by the deviation from the expected value (by about 4% to 6%).
- Individual scans from small peak intensities (e.g. scan 530 and 568) showed very high errors and deviated significantly from the expected values. These scans also differed from other scans that were extracted from higher peak intensities (e.g., scans 535 to 562). However, the small peak intensities had minimal influence on the accuracy of the mass proportions derived from averaged-scans of $[M-15]^+$, $[M-57]^+$, $[M-85]^+$ and $f302^+$ peaks.

- The mass proportions obtained from most of the individual scans and averaged-scans were consistent in the fragments $[M-57]^+$ and $[M-85]^+$ with related errors of < 0.002 % and < 0.012 % respectively for the averaged scans. These are the most reliable fragments of alanine. Also the average of all scans across a metabolite peak provides representative MIDs of the metabolite fragments.
- The fragments $[M-15]^+$ and $f302^+$ result in same mass m/z = 302 and can be influenced by each other. These two fragments result from the loss of a methyl group from the TBDMS derivative ($[M-15]^+$) or from alanine ($f302^+$). Therefore, $[M-15]^+$ contains all three C-skeleton atoms (m/z 302-305) whereas $f302^+$ has only two (m/z 302-304). The potential interference caused by these overlapping signals can confound interpretation of these fragment and needs to be thoroughly verified. The error associated with individual scans of fragments $[M-15]^+$ (m/z 302-305) and $f302^+$ (m/z 302-304) is high (about 3.1 mol% and 2.3 mol% respectively). However, the mass proportions of averaged scans agree with the expected values and are comparatively more precise, with errors of < 0.53mol%and < 0.46 mol% respectively. The reliability of these fragments depends on the relative contribution of the overlapping fragments. Although these two fragments seem to be valid in the current analysis, the influence of the signal overlap on the accuracy and precision could be much more pronounced in ^{13}C-labeled samples. Thus, it is important to examine the accuracy and errors of mass fractions in molecules possessing isotopomer distributions similar to those likely to be encountered in MFA labeling studies. For this reason, a similar analysis was performed on metabolites labeled biogenically by growing cells in an appropriate ^{13}C-substrate.

Figures 3.18 and 3.19 represents the mass proportions of fragment ions and their errors obtained from individual and the averaged-scans acquired from biogenically derived [U-^{13}C]alanine. Here, [^{13}C]alanine was obtained from heterotrophic *Arabidopsis* suspension grown on 20% [$^{13}C_6$]glucose. Based on steady state conditions the average ^{13}C in alanine was expected to be about 20%. It must be noted that the expected values in biogenic derived samples are at best estimates which can differ based on the metabolic system under study. In a 3-carbon fragment (e.g., $[M-57]^+$), with 20%[U-^{13}C]alanine the expected mass fractions are 0.2 (i.e. 20%) in m+0 and 0.8 (i.e. 80%) in m+3. Analysis of the GC-MS data obtained for alanine assessment for ^{13}C labeled samples are verified below.

- The mass proportions of $[M-0]^+$ were inaccurate and noisy, hence rejected. This is in agreement with the analysis from unlabelled alanine.
- The mass proportions obtained from averaged-scans were consistent with the individual scans for the rest of fragments implying that averaged-scans of a GC peak can be used as representative MID for this molecule.
- Quantitative analysis of mass proportions obtained from averaged scans for $[M-57]^+$ and $[M-85]^+$ agreed with the expected values and were accurate with an error of < 0.005 mol% and < 0.025 mol% respectively (Fig 3.19)
- Fragment $[M-159]^+$, was unreliable due to high error (0.68 mol% to 1.55 mol%), and was consistent with the previous analysis of unlabelled alanine
- The overlapping fragments $[M-15]^+$ and $f302^+$ at m/z 302 were analyzed in detail (Fig 3.20). The mass proportions from $[M-15]^+$ agreed with both the expected values and also with the [M-57]+ (same number of alanine carbon atoms). The errors were between 0.11 mol% and 0.18 mol% and hence the fragment can be accepted at this precision. The mass proportions of

f302^{+} varied from the expected values and also the errors were very high (15.9 mol% to 23.5 mol%), thereby rejecting the fragment. The analysis implies that [M-15]$^{+}$ is the major contributing fragment for the MIDs between m/z 302 to 305. As the [M-57]$^{+}$ is redundant to the [M-15]$^{+}$ and also more reliable due to low errors, it was preferred for flux analysis studies.

3.3.3.4 *Average ^{13}C can be used to validate the fragments*

In addition, it was observed that the average ^{13}C derived from mass isotopomer fractions of metabolite fragments can also be used reliably for validations, by comparing with the expected values (natural abundance of ≈1.13% ^{13}C per carbon in unlabelled metabolite fragments). Average ^{13}C per carbon is calculated by equation 3.1. The mass proportions of the fragments obtained from standards, biogenically derived unlabelled and ^{13}C-labeled metabolites were extensively tested to validate the fragments. The current study shows how average ^{13}C can be used to test the accuracy and estimate the precision of the fragments as illustrated below.

$$\text{Average}^{13}\text{C} = \left(\sum_{i=0-N} n_{m+i} \times A \right) \Big/ N \qquad \text{---> Equation 3.1}$$

n_{m+i} : number of ^{13}C containing carbons in mass isotopomer (m+i). n=0 for m+0, n=1 for m+1 and so on .... until i equals to total number of carbons (**N**) in the fragment

A : corresponding fractional abundance of the mass isotopomer

To illustrate this approach, Tables 3.8A and 3.8B present the average ^{13}C in each fragment ion of alanine-2TBDMS obtained from biogenically derived unlabelled and ~20%[U-^{13}C] alanine.

- The fragments [M-0]$^{+}$ and [M-159]$^{+}$ were rejected as the average ^{13}C was much higher than the expected value in all the samples.

- The average ^{13}C from the fragments $[M-15]^+$, $[M-57]^+$ and $[M-85]^+$ were in agreement with the expected values (< 1.13 atom% for unlabelled alanine) and were deemed reliable.
- The average ^{13}C of labelled fragments (from biogenic 20% [U-^{13}C]alanine) processed from individual and averaged-scans over the GC peaks were also analyzed (Fig 3.21). Fragment $f302^+$ was rejected based on the biogenic derived [U-^{13}C]alanine in which it deviated significantly from the expected value of 20%.
- The average ^{13}C of a few individual scans obtained from low peak intensities (e.g. 262, 264) deviated slightly from the majority of scans which agreed with the expected values. However, these did not influence the results from averaged-scans, which agreed well with the expected values for fragments $[M-15]^+$, $[M-57]^+$ and $[M-85]^+$. All these conclusions are consistent with the validation of alanine fragments from analysis of unlabelled standards that contained ^{13}C at natural abundance.
- Reliability on each mass isotopomer (measurement accuracy and precision) can be assigned based on the average ^{13}C measured in standards or unlabelled samples. For example the mass fractions of fragments $[M-15]^+$, $[M-57]^+$ and $[M-85]^+$ can be assigned to be less than the limits of measured ^{13}C values i.e. less than 1.04, 1.03 or 1.00 mol% respectively. Although not recommended, if we were to use the unreliable fragments $[M-0]^+$ and $[M-159]^+$, the errors assigned on the mass fractions should be huge which were measured (based on average ^{13}C) to be 74.13 mol% and 5.55 mol% respectively.

3.3.3.5 *Effect of metabolite concentration on the MID*

There could be significant correlation between the MID and concentrations of the metabolites. Metabolite concentrations may affect the accuracy of the mass fractions and are machine dependent (Antoniewicz et al., 2007). The influence on the accuracy needs to be accounted for

by assigning appropriate errors to each mass isotopomer that are measured on a particular GC-MS system. Hence, in the current study, the concentration effect of metabolites on the accuracy of measured MID acquired with Agilent 79890 GC/5975 were tested. The standard amino acid concentrations injected to the GC-MS were in the range of 5.2 to 166.6 ng (Appendix 2.2). Current analysis has shown that, the accuracy or precision of the mass fractions were not influenced by concentration effects as depicted below.

Figure 3.22 shows the analysis of GC-MS data of alanine and glycine obtained at different concentrations. The mass fractions of the fragment $[M-57]^+$ at different concentrations is presented. The mass fractions were consistent and accurate at concentrations over the range 20-166.6 ng/µl, whereas at low concentrations (5.2-10.4 ng/µl) they deviated slightly from the expected values. At low concentrations the errors were estimated to be < 1.2 mol% and 0.26 mol% for alanine and glycine respectively. This could be mainly due to the high signal-to-noise ratios at very low concentrations. At higher concentrations (20-166.6 ng/µl) the maximum noise for alanine and glycine were 0.19 mol% and 0.14 mol% respectively. No significant trends of concentration effects were detected. Nevertheless to ensure suitable signal-to-noise in GC-MS spectra, biological samples were routinely analysed over a range of dilutions and an appropriate concentration selected. This was done both to avoid overloading of the GC-column and also to circumvent reliance on very low signals of desired metabolites in subsequent flux analysis.

3.3.3.6 *Assignment of mass isotopomers measurement errors*

The errors associated with measurements of mass isotopomer fractions need to be defined to provide the appropriate weights to the values used in the subsequent flux-fitting procedure.

Based on the above analysis, three potential sources of error estimates derived for each fragment may be used:

i) Errors derived for each fragment of a test sample using the tool MSCorr. Signal strength is taken into consideration.

ii) Accuracy and precisions based on the average ^{13}C measured from standard and unlabelled biogenic-metabolites that mimic the experimental condition

iii) Errors due to concentration effects and instrumental noise of the GC-MS used (For most standard metabolites it is observed to be < 1 mol %)

The maximum of these three is used as a reliable estimate of error for flux analysis.

3.3.4 Conclusions related to GC-MS data handling

The major conclusions from the current analysis are

- MID from averaged scans provide accurate and reliable estimates of the extent of labelling of metabolite fragments.
- All metabolite fragments have to be validated for their accuracy and precision before using them for flux analysis.
- Appropriate errors on the mass isotopomers have to be assigned.
- Unlabelled standards can be used for validation of most of the fragments and can further be verified by using ^{13}C labelled metabolite (standards or biogenic derived). Care should be taken while verifying scans that could overlap.
- Average ^{13}C can also be used to validate the fragments with their accuracy and precision levels. Accordingly errors on mass fractions can be assigned.

3.4 Strategies to account for pre-existing unlabelled metabolite pools

Labelling information from protein, starch and cell wall components, is essential to deduce the fluxes in the compartmented network of plant metabolism, since these measurements report on the labelling of subcellular pools of intermediates that would otherwise be inaccessible to conventional analysis (Ratcliffe and Shachar-Hill, 2006). In steady state ^{13}C feeding experiments conducted on *Arabidopsis* cell suspension cultures, the mass isotopomer measurements of protein-derived amino acids, starch and cell wall components may be biased by inclusion of pre-existing unlabelled pools in the analysis (Chapter 5 and Chapter 6). Similar observations have been reported in earlier studies on *Arabidopsis* embryos (Lonien and Schwender 2009) and *Arabidopsis* cell suspension cultures (Williams et al., 2010). This bias can be attributed to the pools of metabolites that are contained within the initial cell inoculum used in the stable isotope feeding experiments which dilute the contribution of de-novo synthesized metabolites to the labeling of the respective pools. The pools that are most affected by this are usually free metabolites that are stored in vacuole, starch granules, cell wall components and storage or structural proteins in the cell. Also, the problem is likely to be more pronounced in tissues with very slow rate of metabolism. For example, in *Arabidopsis* cell suspensions grown under phosphorus starvation (Chapter 6), some of the soluble amino acid pools have substantial contributions from unlabelled pre-existing metabolites even after many days of ^{13}C feeding. This bias in the mass isotopomer fractions resulting from the influence of pre-existing unlabelled metabolite pools needs to be corrected in order to obtain reliable flux estimates.

After extensive tests on *Arabidopsis* models, appropriate experimental, mathematical and computational strategies were developed to correct the bias in mass isotopomer data of metabolites resulting from unlabelled pre-existing pools. These strategies are presented below, and their implementation illustrated by considering representative mass isotopomer data of metabolites obtained from steady state feeding experiments on heterotrophic *Arabidopsis* suspension cultures fed with 99% [1-^{13}C]-, 99% [2-^{13}C]- and 20% [$^{13}C_6$]glucose separately for 5 days. These examples focus on identifying and correcting for the effect of unlabelled pools in representative amino acids (alanine, glycine, serine and glutamate) that usually report on fluxes from different parts of the metabolic network.

3.4.1 Establishing the presence of unlabelled pools in ^{13}C MFA studies

The first major step is to ascertain the presence of unlabelled pools. This is done in two ways:

i. First, by comparing the mass isotopomer fractions of storage compounds with those of their precursors. As illustrated in Figure 3.23, comparison of labelling in free amino acids against those of the protein-derived amino acids (obtained from either 99% [1-^{13}C]-, 99% [2-^{13}C]- or 20% [$^{13}C_6$]glucose feeding experiments) has established significant differences in the mass fractions. The m+0 of the protein-derived amino acids were significantly higher than those of the corresponding free amino acids. This implies the presence of pre-existing unlabelled protein pools which bias the measurements of mass fractions.
ii. Secondly, by measuring the average abundance of ^{13}C (Section 3.3.3.4) in metabolites derived from feeding of uniformly labelled substrate (example 20% [$^{13}C_6$]glucose). The measure of average ^{13}C in each metabolite is used to determine the extent of labelling and also to examine the isotopic steady state conditions. At isotopic steady state, the average

^{13}C in most of the metabolites obtained from *Arabidopsis* cells supplied with 20% [$^{13}C_6$]glucose is expected to be ~20%. However, the average ^{13}C in all protein-derived amino acids were significantly lower (Fig 3.24) than that in the corresponding free amino acids which confirm the dilution effect of unlabelled pre-existing protein pools in *Arabidopsis* cells. The free amino acids (Fig 3.24) seem to be at near isotopic steady state as their average ^{13}C are close to the expected value of 20%. Further interrogation of individual metabolites is done on a case by case basis as discussed in subsequent sections.

3.4.2 Pre-existing pool correction methodology

Once the occurrence of unlabelled pools is ascertained, the next step is to correct for their contribution to the measurement. Two correction methods were tested:

i. Modeling adjustments: by adding proxy reactions in the model network
ii. Mathematical: by correcting the biased mass fractions of the metabolite pools

3.4.2.1 *Modeling adjustments to account for pre-existing unlabelled pools*

The contribution of pre-existing unlabelled pools to the mass isotopomer measurements of metabolites may be accommodated by incorporating additional proxy reactions into the network to account for the bias in the mass fractions (Lonien and Schwender 2009; Williams et al., 2010; Kruger et al., 2012). The pre-existing unlabelled pool of each metabolite was determined by a proxy uptake reaction ("set as a free or constrained flux") of the corresponding unlabelled metabolite ($Vupt_{AAunlab}$, Fig 3.25A) during the fitting process. The uptake reaction was compensated for by an additional efflux (VAA_{eff}+$Vupt_{AAunlab}$ set as "determined", Fig 3.25A) reaction so that the model was constrained to the actual measured net metabolite biosynthesis (VAA_{eff}, Fig 3.25A). This ensures that the introduction of the

unlabelled metabolite pool has no net bearing on the flux solutions of the network. Therefore for each metabolite, one additional free flux and a determined flux need to be incorporated, thereby increasing the total number of free fluxes. In a condition where the pre-existing protein turn-over influences the labeling of soluble amino acids, even more proxy reactions may be needed (Fig 3.25B) to account for the bias. Usually, steady-state MFA of central carbon metabolism in plants are over-determined i.e., the number of free fluxes is usually much less (30-60) than the number of isotopomer measurements (>1000) used in the fitting process (Kruger et al., 2012). Hence, additional free proxy reactions that can account for pre-existing pools of all protein-derived amino acids, sugars such as sucrose, cell wall components (hexoses and pentans) and starch can be accommodated in the model, albeit at the cost of increasing complexity.

Preliminary simulations established that it was possible to analyse a version of the *Arabidopsis* model network containing up to 101 free fluxes (data not shown). Adopting this method required some specific modifications to the fitting procedure in order to handle the large number of free fluxes. Specifically, the proxy free fluxes needed to be constrained during Monte-Carlo analysis of the model to avoid generating an unacceptably high proportion of infeasible flux solutions. Subsequently, in the final optimization to find the best fit, these were set as free fluxes. Additionally, in these models the proxy uptake flux value could not be set as zero; this limitation potentially compromises the utility of such models for analysis of labeling data from a metabolite that does not contain a contribution from pre-existing pools. The minimum proxy uptake flux permissible by the software is 0.0001, and this may influence the fitting of metabolites with very small biosynthetic fluxes. Therefore, it is desirable, wherever possible; to remove reactions carrying negligible fluxes and thereby minimize the total number

of free fluxes in the network. To reduce the reliance on proxy fluxes in the model, an alternative mathematical approach for correcting for the presence of pre-existing pools of metabolites is explained below.

3.4.2.2 *Mathematical pool correction of biased mass fractions*

Mathematical pool correction was used to correct the mass isotopomer fractions by removing the contribution of the pre-existing pools. Mathematical pool correction is based on four steps (see below in a box for representation)

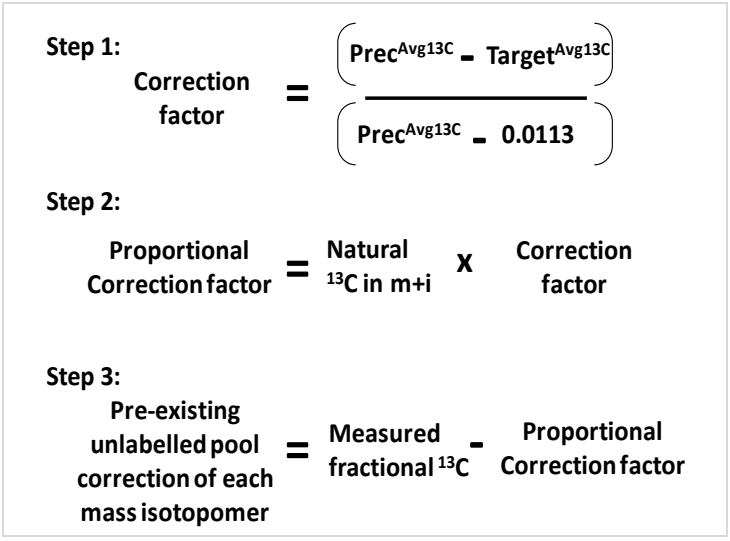
Step 1: **Calculate average ^{13}C of each metabolite fragment** (e.g. $[M-57]^+$ of amino acids) obtained from feeding of uniformly labeled substrate (example 20% [$^{13}C_6$]glucose).

Step 2: **Estimate a correction factor** based on the average ^{13}C measured from target metabolite fragment, the measured (or theoretical) isotopic abundance of the precursor metabolite pools and the natural abundance of ^{13}C (1.13%) in pre-existing pools as shown in the equation below. The average ^{13}C of precursors are derived from central metabolite(s) that are at near isotopic steady state and contribute carbons for the biosynthesis of the target metabolites. For example, the average ^{13}C of free soluble alanine (precursor) can be used to correct the protein-derived alanine (target metabolite from protein hydrolysate). In case the immediate precursor metabolite could not be detected or was itself biased due to the presence of stored pools, the average ^{13}C of other central metabolites that are retrobiosynthetically related can be used. For example, in the absence of reliable estimates of the labeling of α-ketogluratate, the average ^{13}C abundance of TCA cycle intermediates citrate and malate, taken together, may be used as estimates of the labeling of the precursor to correct measurements of the target metabolites free glutamate and protein-derived glutamate which may be biased by

pre-existing pools. Similarly, the average ^{13}C of the supplied respiratory substrate (20% [$^{13}C_6$]glucose) or of free alanine can be used to estimate correction factors for sucrose, starch and cell wall components. The selection of appropriate precursor and target metabolites is dependent on the network structure, metabolic phenotype and the measurements available, as shown in the case studies below (section 3.4.3).

Step 3: **Calculate the proportional correction factors of each mass isotopomer** of the target metabolite fragment by multiplying the respective natural abundance with correction factor. The proportional correction factors of the mass fractions represent the contribution of the pre-existing pool in each mass isotopomer that needs to be subtracted from the measurements.

Step 4: **Obtain the correct mass isotopomer fractions** of metabolite fragments by subtracting the contribution of pre-existing pools. The corrected mass isotopomers of each metabolite fragment are scaled and these represent the measurements of *labelled* metabolites (synthesized *de novo*) which can be used for ^{13}C MFA. Measurements obtained from feeding of positionally labelled substrates (e.g. 99% [1-^{13}C]-, 99%[2-^{13}C]glucose) can be corrected using the proportional correction factors derived from [$^{13}C_6$]glucose provided all the experiments are conducted in parallel on the same batch of cells.



Prec^{Avg13C} : Average ^{13}C of the precursor metabolite fragment

Target^{Avg13C} : Average ^{13}C of the target metabolite fragment

Natural ^{13}C in m+i : Expected natural abundance of mass isotopomer (i = 0,1,2…n for a molecule containing n carbon atoms)

In summary, this method relies on measuring the average ^{13}C in each metabolite obtained from feeding of uniformly labeled substrates and then correcting the labeling of products based on the average ^{13}C abundance in central metabolites or precursors that reach isotopic steady state. Most of the metabolites can be subjected to mathematical pool corrections to account for the bias in measurements of mass fractions arising from the contribution from pre-existing pools. However, it is important to note that not all metabolites in a network are amenable to such mathematical pool correction. In particular it is difficult to apply this approach to metabolites formed by reactions which take up atmospheric CO_2 and/or methylene-tetrahydrofolate (M-THF, 1-carbon) since it is not possible to accurately predict the average ^{13}C abundance in the proportion of such a metabolite formed during labeling with [$^{13}C_6$]glucose. Thus, pre-existing pool corrections in methionine, histidine, arginine, glycine and serine (that have active CO_2 or M-THF contributions) are best achieved by modeling adjustments with additional proxy

reactions as described above. Thus in practice it is likely that flux analysis will involve a combination of both pool correction methods to reduce the complexity of the network and improve the accuracy and precision of the flux determinations.

3.4.3 Case studies

The following case studies illustrate the implementation of the strategies outlined above for pre-existing unlabelled pool correction. Representative measurements of mass isotopomer fractions of soluble and protein-derived amino acid fragments [M-57^{+}] of alanine, glycine, serine and glutamate (Figs 3.23 and 3.24) obtained from steady state feeding experiments on heterotrophic *Arabidopsis* suspension cultures fed with 99% [1-^{13}C]-, 99% [2-^{13}C]- and 20% [$^{13}C_6$]glucose separately for 5 days are used.

Case1: Alanine

The raw measurements of soluble and protein-derived alanine differed and the pre-existing pools are detected in protein-derived alanine (Fig 3.23, Fig 3.26A). The average ^{13}C of the free soluble alanine (19.6 ± 0.34%) was in good agreement with the expected (~20%) implying this pool reached isotopic steady state (Fig 3.24). The average ^{13}C of soluble alanine was used to correct the protein-derived alanine mathematically. After correction there was no significant difference between the mass isotopomers of soluble and protein-derived alanine (Fig 3.26B). Corrected protein-derived alanine isotopomers can be used for flux analysis. This also validates the approach of correction.

Case2: Serine and Glycine

The raw measurements of soluble and protein-derived serine/glycine differed and the pre-existing pools are detected in protein-derived serine/glycine (Fig 3.23). The average ^{13}C of

soluble glycine (18.1 ± 0.26%) and serine (18.6 ± 0.08%) were slightly less than the theoretically expected value of 20% (assuming no isotopic dilution). This can only be explained by two assumptions a) that there are unlabelled pools of these free amino acids in the cells and/or b) that it is due to the metabolic property of the phenotype in which unlabelled or partially labelled 1-carbon molecules (like CO_2 and/or methylene tetra-hydrofolate (MTHF)) enter the metabolism thereby diluting the average ^{13}C label in central intermediates. The reversible reactions of glycine decarboxylation (glycine ↔ CO_2 + MTHF) and serine hydroxymethyl transferase (serine ↔ glycine + MTHF) allow the entry of CO_2 and MTHF and may dilute labelling in the free glycine and serine pools. The extent of dilution due to these 1-carbon molecules depends on the fluxes of the related reactions. In this instance, the second assumption seems more likely since most of the soluble metabolites in *Arabidopsis* suspension reach isotopic steady state within few hours (data not shown, based on ^{13}C *in vivo* NMR studies). Therefore, the approach of adding additional proxy reactions is preferred to account for the pre-existing pools in protein-derived serine and glycine (Fig 3.27). Also uptake of 1-carbon molecules needs to be simulated to avoid bias arising from the incorporation of unlabelled carbon from this source.

Case3: Glutamate

The raw measurements of soluble and protein-derived glutamate differed and pre-existing pools were detected in protein-derived glutamate (Fig 3.23, Fig 3.28A). The average ^{13}C of soluble glutamate (18.9 ± 0.04%) was observed to be slightly less than the expected value of 20% (assuming no isotopic dilution). This may be explained from the dilution effect associated with the introduction of unlabelled carbon from CO_2 via the reaction catalysed by PEP carboxylase (PEPc; PEP + CO_2 → OAA). The flux via this route may dilute the labeling of

TCA cycle intermediates and amino acids synthesized from them (Fig 3.29). As the organic acids in these *Arabidopsis* cells are at isotopic steady state their average ^{13}C can be used for correction of pre-existing pools in amino acids derived from TCA cycle. The average ^{13}C of citrate and malate taken together (measured to be 19.06%) is used to estimate the labeling of precursors to correct the label measurements of target metabolites i.e. free glutamate and protein-derived glutamate. After isotopic correction the difference between the mass isotopomers of soluble and protein-derived glutamate decreased significantly (Fig 3.28B).

3.5 ^{13}C-NMR data handling

The positional isotopomers from ^{13}C NMR provides with additional sensitivity in deriving accurate fluxes. For accurate assessment, the following experimental and analytical tests were conducted.

- Tests to establish appropriate relaxation delay: 6 seconds vs. 19 seconds
- How long to acquire the spectra for low signal to noise - 10 arrays vs. 8 arrays vs. 6 arrays vs. 4 arrays vs. 2 arrays vs. 1 array
- Line fitting for obtaining accurate isotopomer abundances - Satellite peaks are very common in ^{13}C spectra and deconvolution is an essential criteria
- Calculation of peak area and errors from ^{13}C NMR spectra

3.5.1 Tests to establish appropriate relaxation delay (comparison of 6s vs. 19s)

Comparison of the effect of relaxation delays of 6s (Fig 3.30 A) vs. 19s (Fig 3.30 B) on ^{13}C peak intensities of alanine carbons has shown significant differences. The carboxyl carbon of alanine (C-1) is underestimated at 6s relaxation delay (Figs 3.30, 3.31). It confirms that incomplete relaxation leads to poor measurements. Hence, all the samples obtained from

specifically labelled glucose ([1-^{13}C] or [2-^{13}C]) feeding experiments (Chapters 5 and 6) were subjected to a relaxation delay of 19 s. Alanine is chosen as a representative metabolite for analysis as its isotopomers covers the complete range of the spectrum from 17 ppm (C-3) to 176.6 ppm (C-1; carboxyl group) and is detectable in most of the soluble extracts obtained from *Arabidopsis* cultures grown on [1-^{13}C]glucose.

3.5.2 Effect of acquisition time on the measurements

Comparison of the alanine carbon peak areas and its errors showed that errors on measurements reduce with acquisition time (Fig 3.32). This was due to the reduction in noise. It is recommended to run for long durations. ^{13}C NMR of neutralised perchloric acid extracts (containing soluble metabolites) from Chapters 5 and 6, are subjected to a total acquisition time of about 46 h to obtain reliable measurements for flux analysis.

3.5.3 Calculation of peak area and errors from ^{13}C NMR spectra

The relative abundances of positional isotopomers and the errors on them were measured as presented below in a number of steps. Spectral deconvolution (line-fitting) using the line-fitting (LF) subroutine and absolute peak heights calculation using the define peaks (DP) subroutine were carried out to obtain the peak areas and their relative errors. During LF, resonance frequency, signal intensity, linewidth and the fractional contribution of a Lorentzian lineshape were varied by the software to minimise the difference between the real and simulated spectra (Fig 3.33). In the LF subroutine, peaks of interest were manually selected, and the fit optimised to obtain the absolute peak areas. The relative peak areas were calculated with reference to the area of dioxane peak. Absolute peak heights and chemical shifts were obtained by using the define peak (DP) subroutine along with the automatic peak picking

which uses the RMS noise factor and a predefined minimum height. The signal to noise ratio (SNR) of dioxane was measured which in turn is used to obtain the absolute RMS noise of dioxane and all the other peaks in the spectrum based on corresponding peak heights. The relative errors (RE) on each peak area (obtained by line fitting) was calculated by the custom formula $RE = 0.62(SNR)^{-0.66}$ (Williams et al., 2009).

The different steps used in error estimation of 13C NMR measurements is presented below with an example.

Step 1: Line fitting- deconvolution of peaks, obtaining peak area, peak height

Step 2: Obtain SNR of Dioxane (3.3 ml of PCA extracts from 100-150mg (dw) of plant tissue grown to isotopic steady state are tested for consistency in relation to 25 mM 1,4-dioxane used)

Step 3: Identify the metabolite peaks and assign the isotopomers

Step 4: Obtain absolute RMS noise

Step 5: Obtain SNR of each peak by the formulae- Peak height/RMS noise

Step 6: Calculate error $0.62(SNR)^{-0.66}$. Error estimation from signal to noise (Williams et al., 2008).

Step 7: Calculate Relative abundances and error of each peak. Relative area * error

Relative area is obtained in relation to Dioxane abudance (set to 100)

Step 8: Combine the same isotopomer satellites and errors SQRT(sd_sat1^2+sd_sat2^2)

Step 9: Additional noise for 13CFLUX analysis

Error ± 2.5% of absolute abundance (Based on the observation that most of the peaks gives a base error of at least 2.5% (Williams et al., 2008))

Calculation of peak area and errors demonstrated with an example:

One-dimensional ^{13}C NMR spectra were obtained from 3.3 ml of perchloric acid extracts of *Arabidopsis* cultures (100 mg dw) grown on 99%[1-^{13}C]glucose. Peaks of alanine and dioxane are used in this example.

Step 1: Line fitting

LINE	HZ	PPM	HEIGHT	REL_HT	WIDTH	AREA	REL_AREA
1	26613.86	176.452	3.89E+08	0.188	3.34	5.11E+09	5.11E+09
2	10150.61	67.299	1.4E+10	6.732	1.95	1.07E+11	1.07E+11
3	7776.38	51.558	1.07E+08	0.052	-5	2.11E+09	2.11E+09
4	7759.27	51.445	5.7E+08	0.275	4.61	1.03E+10	1.03E+10
5	7741.3	51.326	1.2E+08	0.058	-4.41	2.09E+09	2.09E+09
6	2588.54	17.162	2.11E+08	0.102	1.75	1.45E+09	1.45E+09
7	2572.69	17.057	1.69E+10	8.148	2.82	1.87E+11	1.87E+11
8	2553.55	16.93	1.89E+08	0.091	1.51	1.13E+09	1.13E+09

Step 2: Signal-to-noise of dioxane (25 mM) = 1636.9

Step 3 to Step 7: Peak identification

Step 3			**Step 4**	**Step 5**	**Step 6**	**Step 7**	
PPM	Metabolite and its isotopomers		Measured S/N of Dioxane = 1636.9 Absolute Height of Dioxane = 1.4E+10 Absolute RMS Noise = 8532534	S/N of each peak	Error 0.62x^-0.66	REL AREA (Dioxane set to 100)	Relative error
176.452	**Alanine**	#10x		45.599	0.050	4.769	0.238
67.299	**Dioxane**	Dioxane		1636.900	0.005	100.000	0.469
51.558	**Alanine**	#011+#110		12.552	0.117	1.968	0.230
51.445	**Alanine**	#010		66.782	0.039	9.635	0.373
51.326	**Alanine**	#011+#110		14.103	0.108	1.946	0.210
17.162	**Alanine**	#x11		24.718	0.075	1.356	0.101
17.057	**Alanine**	#x01		1981.16	0.004	174.761	0.723

				1			
16.93	**Alanine**	#x11		22.150	0.080	1.050	0.084

Step 8 and Step 9:

		Step 8		**Step 9 (Use for ^{13}C MFA)**	
Metabolite and its isotopomers		Rel. abundance	error	Rel. abundance	sd
Alanine	#10x	4.769	0.238	4.769	0.357
	#010	9.635	0.373	9.635	0.614
	#011+#110	3.914	0.312	3.914	0.409
	#x01	174.761	0.723	174.761	5.092
	#x11	2.405	0.132	2.405	0.192

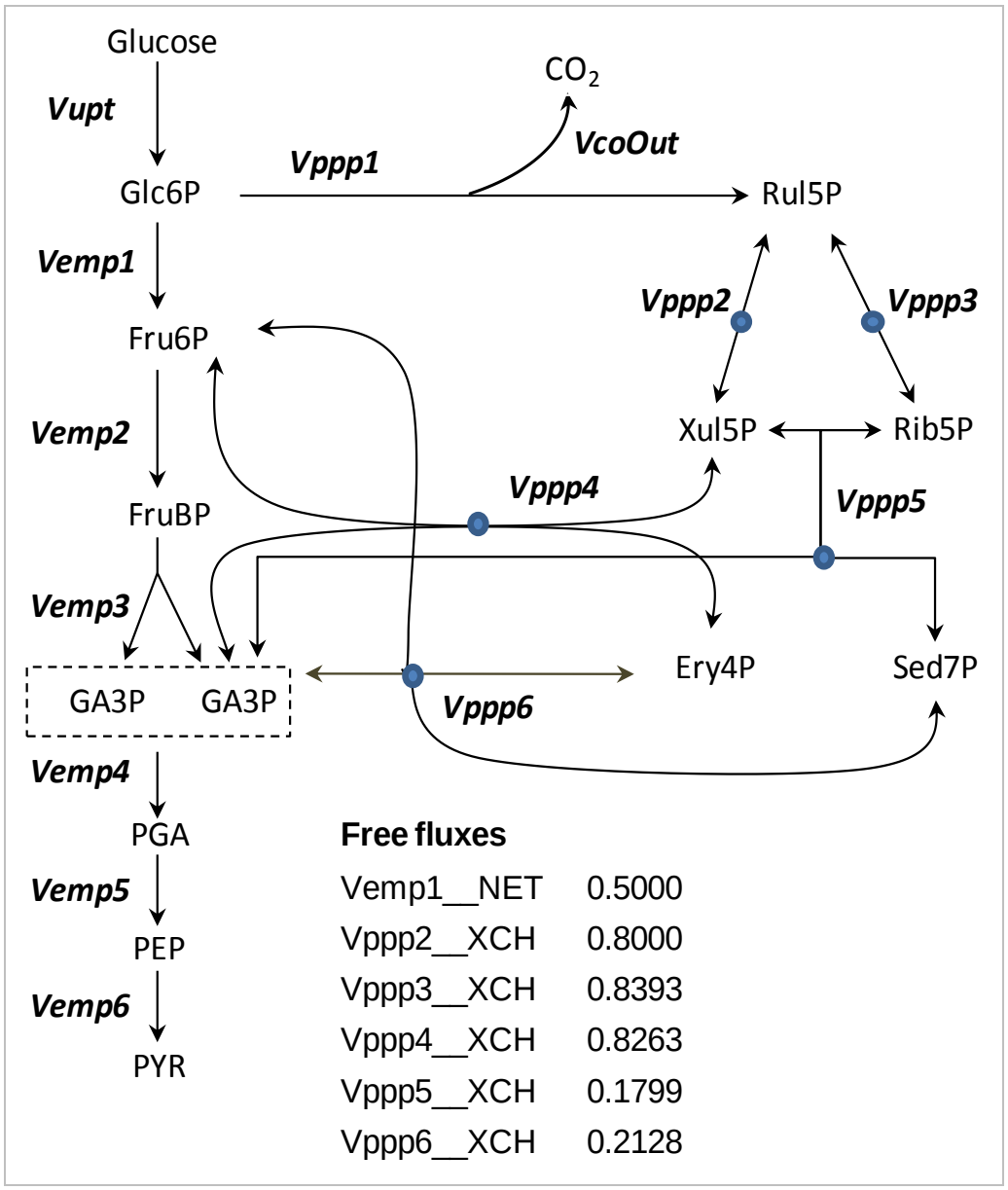


Figure 3.1 Example metabolic network (PPP model) with the Embden-Meyerhof-Parnas pathway (emp) and pentose phosphate pathway (ppp) reactions. The fluxes are represented in bold italics. Unidirectional arrows represent irreversible reactions with only net flux, whereas reversible reactions with both net and exchange fluxes are presented as arrows with dots. The glucose uptake rate is constrained to 1. The values of the six free fluxes (Net and XCH[0-1]) were set to the values shown.

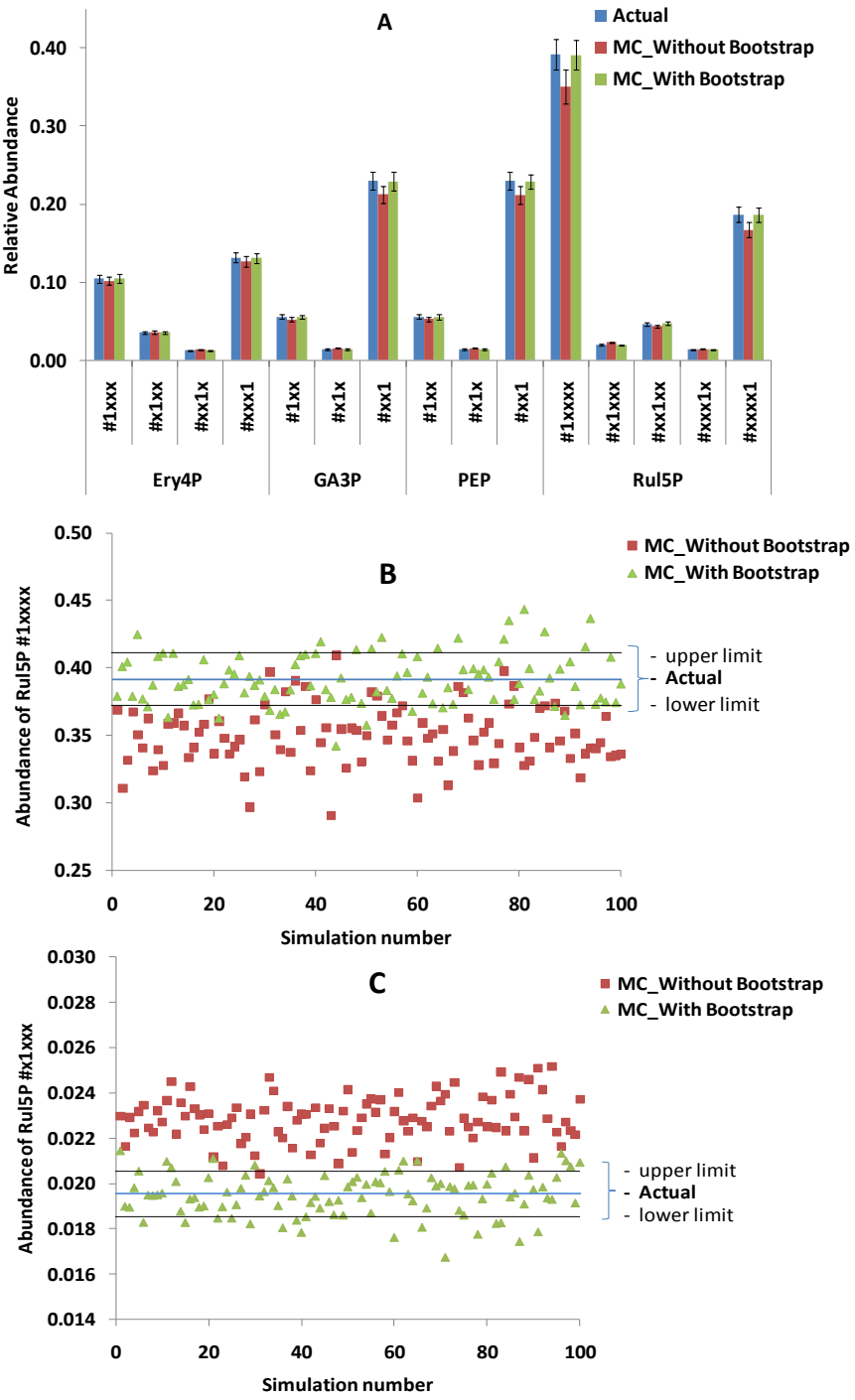


Figure 3.2 A Mean abundances of cumomers obtained from actual and 100 Monte Carlo (MC) simulations with and without bootstrap sampling. The distribution of Rul5P #1xxxx (**B)** and #x1xxx (**C**) pseudo-measurements representing the overestimated and underestimated data points respectively obtained without bootstrap sampling of actual measurements. The pseudo-measurements obtained with bootstrap sampling predominantly fell within the limits of the actual measurements.

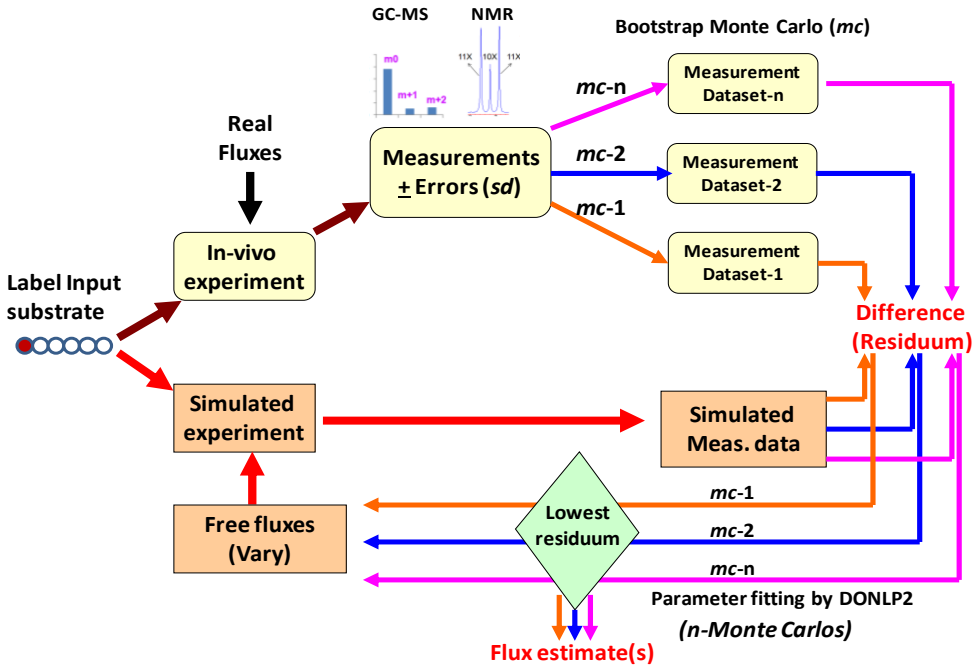


Figure 3.3 Optimisation procedure used by 13C-FLUX for iterative fitting of measured labelling data to the simulated data. Bootstrap Monte Carlo simulations generate populations of measurement datasets that provide an effective method for exploring the flux space. Each simulation starts with random free fluxes that vary iteratively to generate simulated measurements that improve the fit by reducing the residuum. The flux estimates that give the lowest residuum are considered to be optimal solutions.

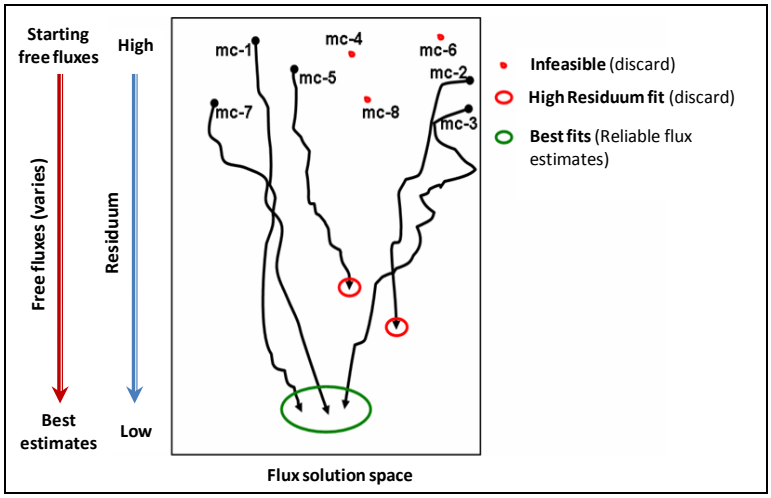


Figure 3.4 Illustration depicting the quality of the flux estimate solutions obtained by Monte Carlo simulations. The residuum decreases and free fluxes vary for each Monte Carlo (mc) simulation. The global best fits with low residuum values are reliable flux estimates from which the flux space can be defined. Solutions with high residuum values (local minima) and infeasible points are discarded.

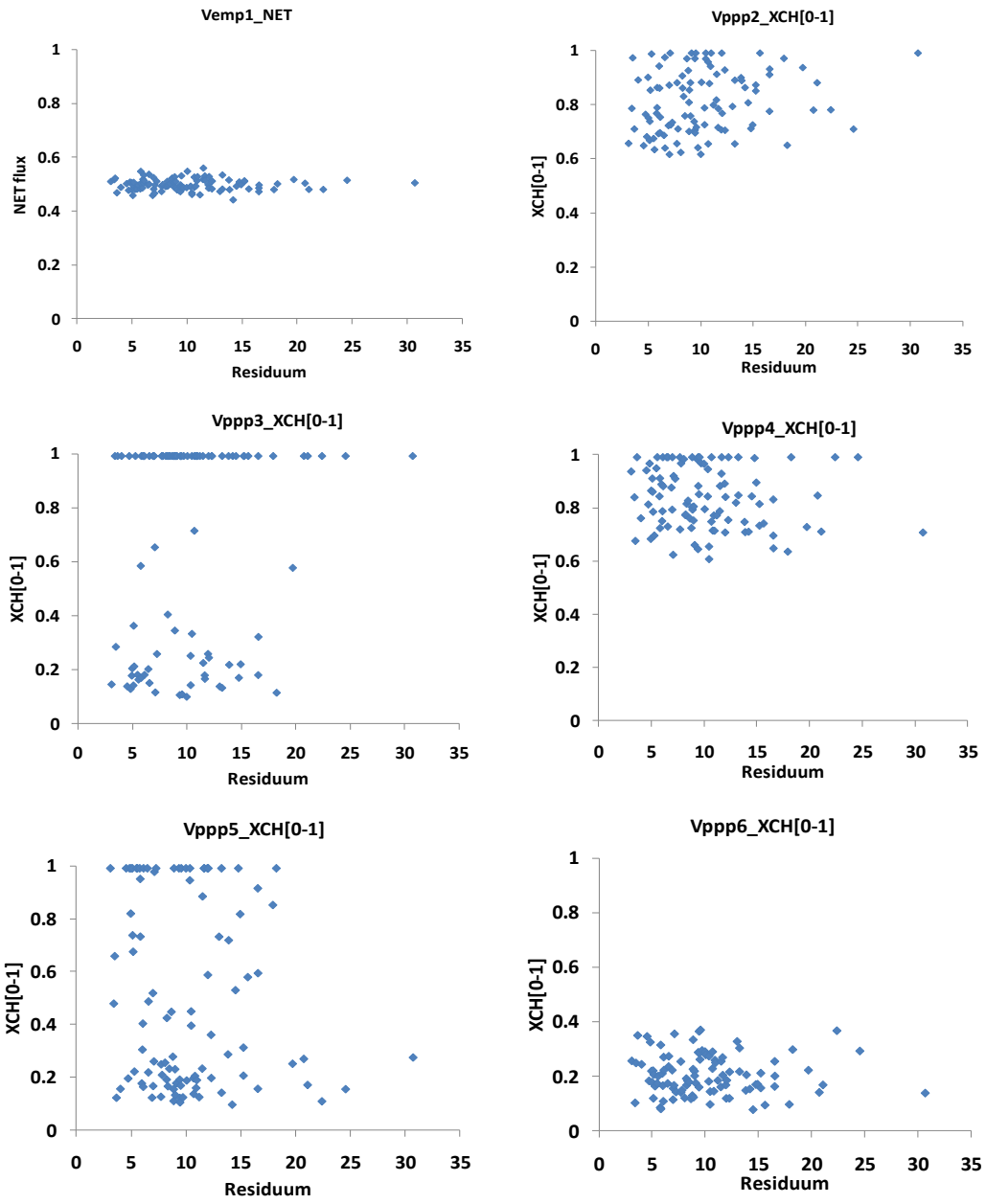


Figure 3.5 Monte Carlo flux estimates of all the free fluxes and the corresponding residua obtained by bootstrap sampling of measurements and iterative simulations. Vemp1_NET, Vppp2_XCH, Vppp4_XCH and Vppp6_XCH are tightly distributed whereas Vppp3_XCH and Vppp5_XCH were random. The global best fit solution is obtained from these estimates as mentioned in the text and Fig. 3.7

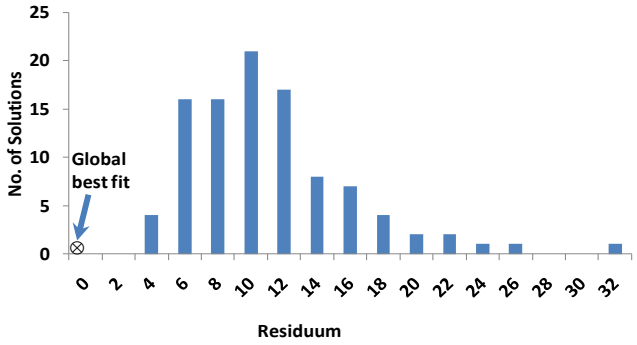


Figure 3.6 The residua and frequency of Monte Carlo solutions of the PPP model. The residua fell within the range 3 to 32. The global best fit with the lowest residuum of 0 is indicated by an arrow.

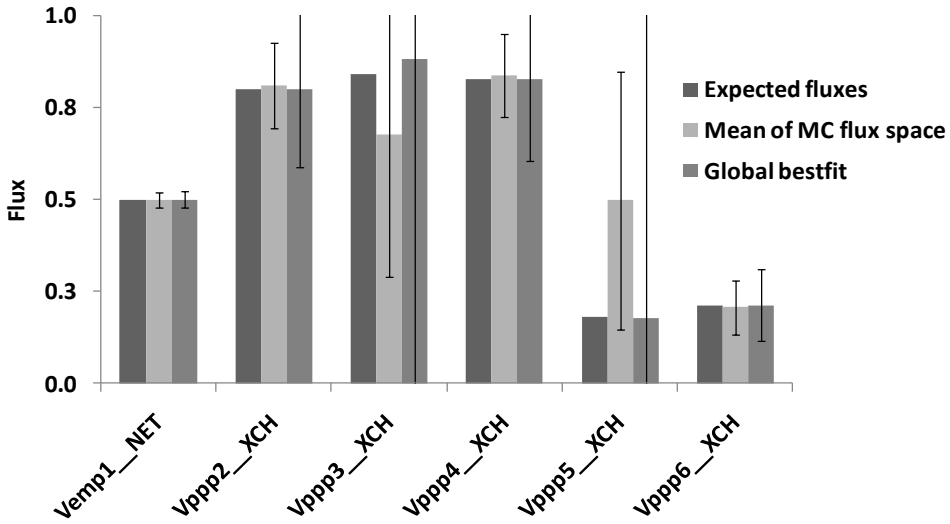


Figure 3.7 Fluxes of expected, Monte Carlo flux space and the global best fit solution. Monte Carlo flux space was obtained from 100 best fit flux solutions (mean ± sd). The global best fit was obtained by refitting the mean fluxes obtained from the flux space with the actual measurements and the error bars were obtained by the 13C-FLUX routine *EstimateStat* (section 3.2.4.1). The limits of error bars are set to be < 1 for XCH[0-1] fluxes.

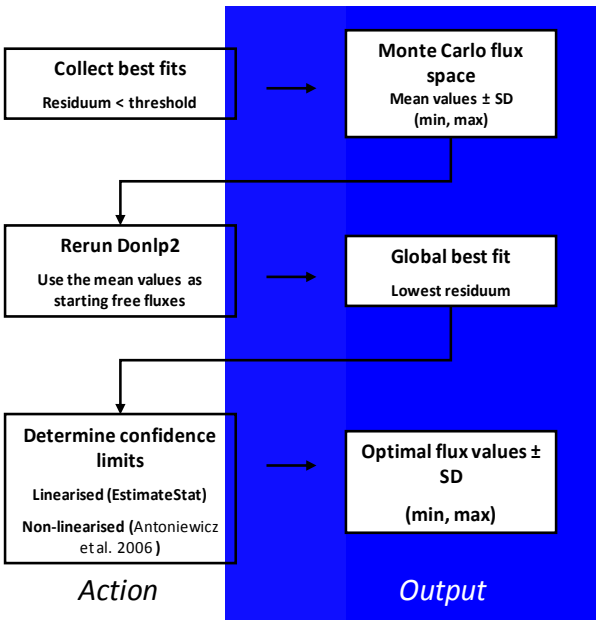


Figure 3.8 Procedure for determining the global best fit and statistically robust flux confidence intervals. The best fit flux estimates obtained by the bootstrap Monte Carlo method are collected, based on threshold cutoff of the residuum, and the Monte Carlo flux space is derived as mentioned in section 3.2.4. The mean values are used as the starting free fluxes for a final optimization which result in global best fit solution with the lowest residuum. The confidence limits are derived by linearised approximation or non-linearised statistics as discussed in section 3.2.4.3.

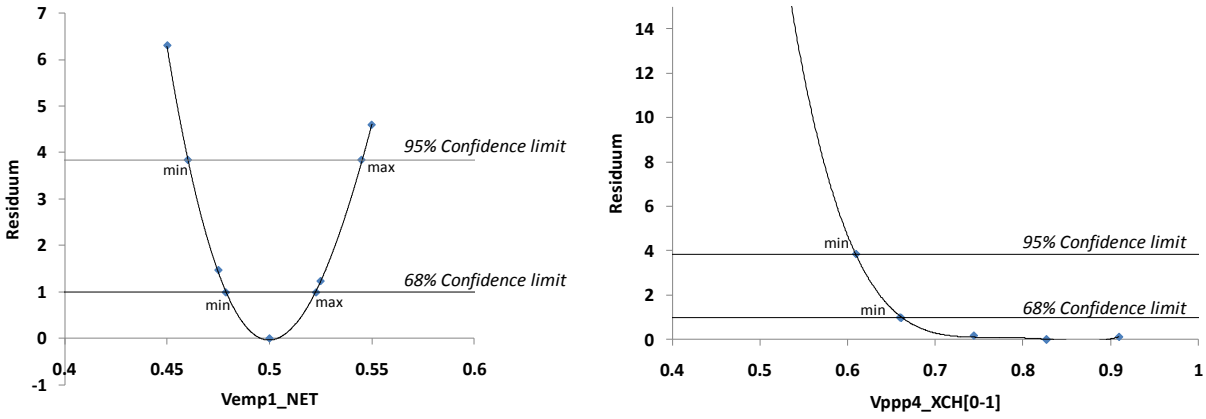


Figure 3.9 Determination of true flux confidence intervals (CI) by the non-linearised statistical method. The residuum is plotted against the flux values which were varied. The 68% and 95% confidence limits are presented as horizontal lines and these were used as a cutoff to deduce the minimum (min) and maximum (max) flux intervals. The non-symmetric nature of the flux Vppp4_XCH is clearly illustrated

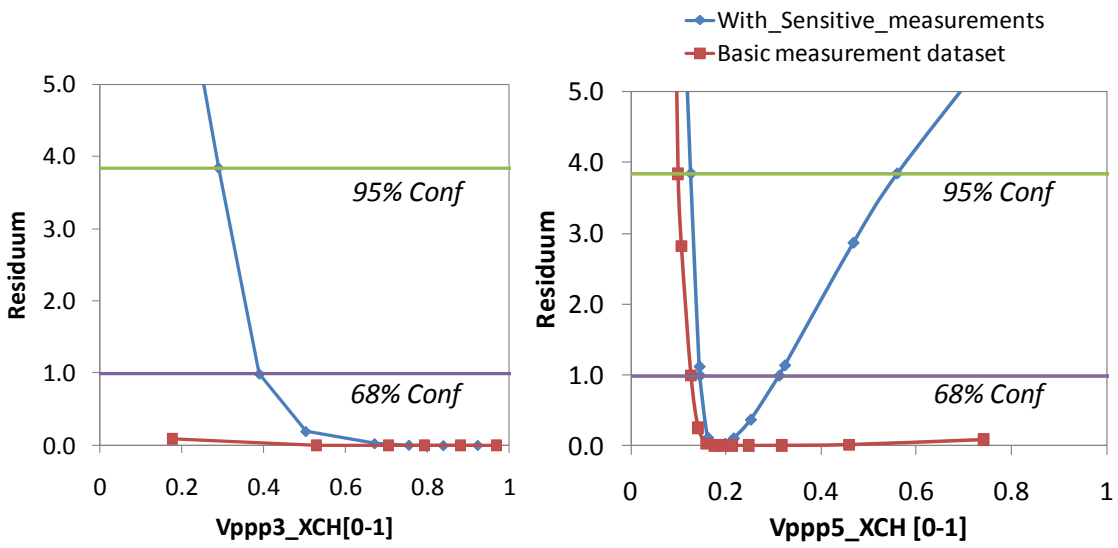


Figure 3.10 Effect of sensitive measurements on fluxes. Vppp3_XCH and Vppp5_XCH are poorly defined by the basic measurement dataset (Table 3.1). Flux analysis with the addition of sensitive measurements for Rib-5P (#1xxxx and #xxxx1) and Sed-7P (#1xxxxxx and #xx1xxxx) improved the definition of the exchange fluxes Vppp3_XCH and Vppp5_XCH. The flux confidence intervals (CI) were obtained by the non-linearised statistical method. The residuum is plotted against the flux values which were varied. The 68% and 95% confidence limits are presented as horizontal lines and used as cutoff to deduce the minimum (min) and maximum (max) flux intervals.

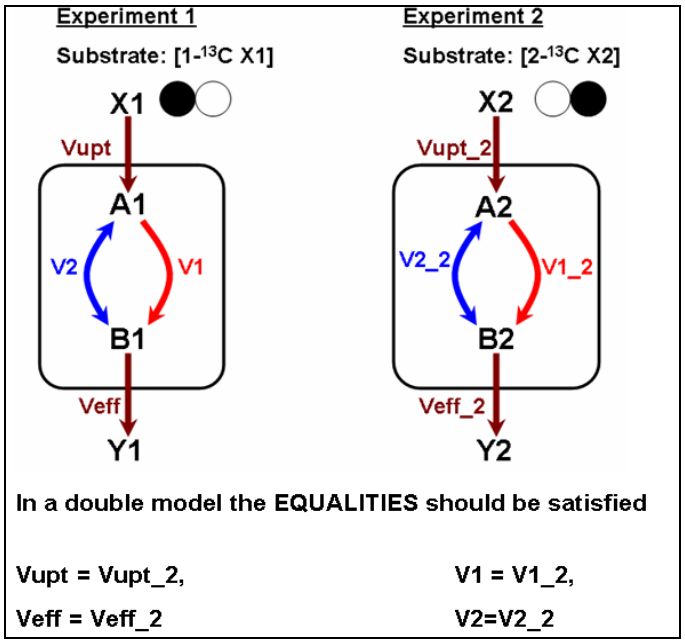


Figure 3.11 An example illustrating a double model derived from two separate experiments by flux EQUALITY constraints.

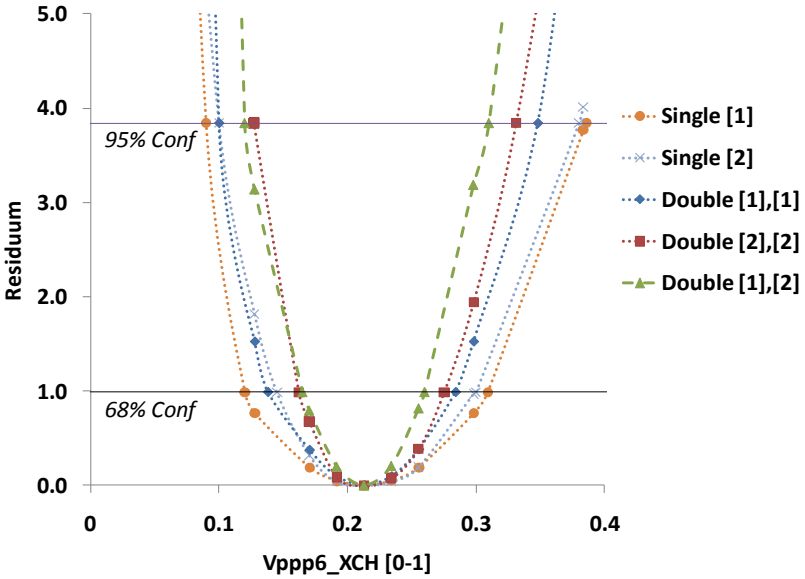


Figure 3.12 Effect of multi-model simulations on flux confidence intervals (CI) of Vppp6_XCH by the non-linearised statistical method. The single or double models were obtained by combination of 1-[^{13}C]glucose [1] and/or 2-[^{13}C]glucose [2] as substrate (s). The minimum and maximum flux intervals derived at 68% confidence intervals are presented in Table 3.6. The confidence on the flux was better in double models in comparison to the single. Within the double models, feeding with different substrates (Double[1][2]) was better than the replicate models (Double[1][1] or Double[2][2]) for Vppp6_XCH.

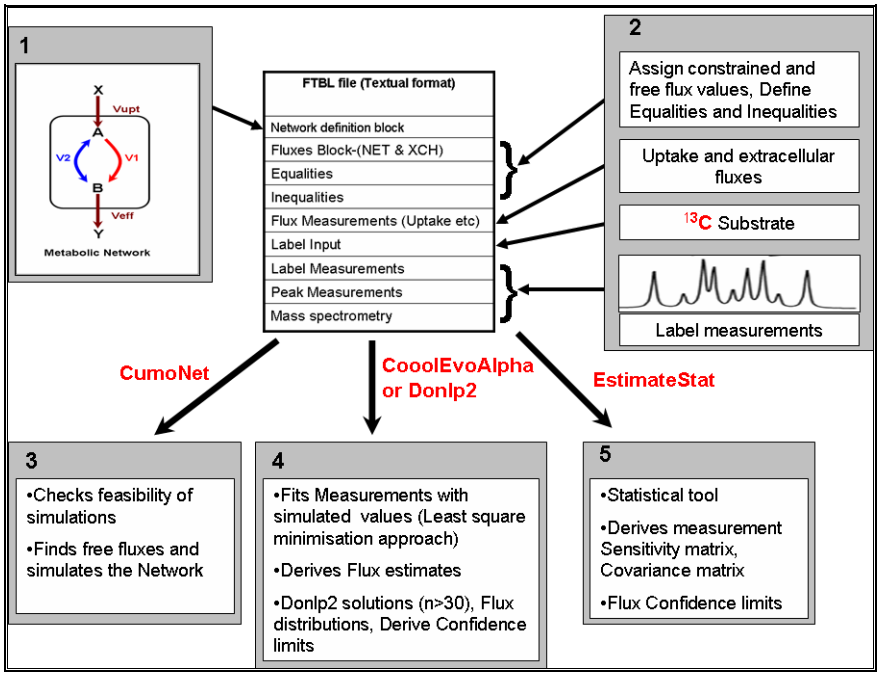


Figure 3.13 Outline of 13C-FLUX framework (Wiechert et al., 2001). The pre-requisites for flux analysis are fed into a textual notation file (FTBL file). The routines *CumoNet*, *CooolEvoAlpha*, *Donlp2* and *EstimateStat* are used iteratively to generate the flux estimates.

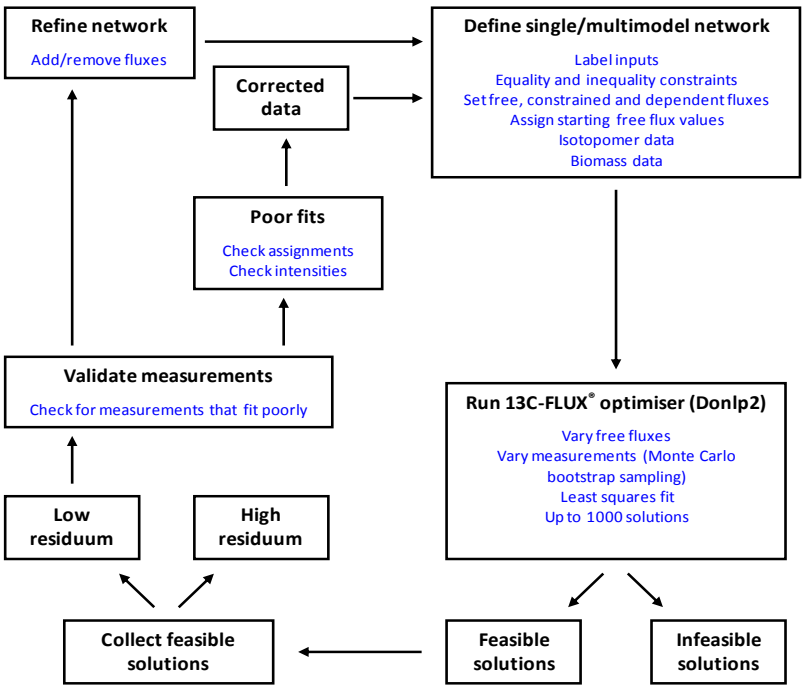


Figure 3.14 Iterative procedures for determining metabolic fluxes from steady-state stable isotope labeling experiments.

Table 3.1 NMR and GC-MS measurements selected for flux estimation with either [1-^{13}C]glucose or [2-^{13}C]glucose as substrate. These measurements were selected out of the 458 cumomers obtained from the *Cumonet* simulation of the network with predefined fluxes (Fig 3.1) and labeled substrate. The isotopomer/cumomer data was taken with random noise with a standard deviation of 5% for NMR measurements and 1 mol% for GC-MS measurements. Cumomers were defined by the presence (1) or absence (0) of label at specific positions. Positions in which the labelling of carbon atoms is undetermined are indicated by x.

		Substrate: 99%[1-^{13}C]glucose		Substrate: 99%[2-^{13}C]glucose		
Metabolite	Cumomer group	Value	Deviation	Value	Deviation	Cumomer
Ery4P	C1234	0.105	0.005	0.253	0.013	#1xxx
		0.036	0.002	0.069	0.003	#x1xx
		0.013	0.001	0.134	0.007	#xx1x
		0.132	0.007	0.123	0.006	#xxx1
GA3P	C123	0.056	0.003	0.117	0.006	#1xx
		0.014	0.001	0.235	0.012	#x1x
		0.230	0.012	0.215	0.011	#xx1
PEP	C123	0.056	0.003	0.117	0.006	#1xx
		0.014	0.001	0.235	0.012	#x1x
		0.230	0.012	0.215	0.011	#xx1
Rul5P	C12345	0.392	0.020	0.551	0.028	#1xxxx
		0.020	0.001	0.405	0.020	#x1xxx
		0.047	0.002	0.096	0.005	#xx1xx
		0.014	0.001	0.191	0.010	#xxx1x
		0.187	0.009	0.174	0.009	#xxxx1
	Mass isotopomers	Value	Deviation	Value	Deviation	
PEP	m0	0.726	0.010	0.545	0.010	
	m1	0.249	0.010	0.345	0.010	
	m2	0.025	0.010	0.108	0.010	
	m3	0.000	0.010	0.002	0.010	

Table 3.2 Comparison of methods for calculation of fluxes and their confidence intervals. Flux estimates for the PPP model with 99%[1-^{13}C]glucose uptake and the measurements defined in Table 3.1 were obtained by exploring the Monte Carlo flux space followed by determination of the global best fit (section 3.2.4). The 68% confidence limits over the best fit fluxes were calculated using linear approximations (EstimateStat), Bootstrap Monte Carlo (MC) solutions and the non-linear analysis method (Antoniewicz et al., 2006) for determining confidence intervals (CI).

	Expected	Best fit ± sd	MC 68%	CI 68%
		(EstimateStat)	(min, max)	Best fit (min, max)
Vemp1__NET	0.5	0.500 ± 0.022	(0.480, 0.518)	0.5 (0.479, 0.523)
Vppp2__XCH	0.8	0.800 ± 0.213	(0.686, 0.957)	0.8 (0.665, 0.99)
Vppp3__XCH	0.839	0.882 ± 50.088	(0.169, 0.99)	0.882 (0, 0.99)
Vppp4__XCH	0.826	0.827 ± 0.222	(0.711, 0.99)	0.827 (0.660, 0.99)
Vppp5__XCH	0.18	0.177 ± 3.508	(0.155, 0.99)	0.177 (0.12, 0.99)
Vppp6__XCH	0.213	0.213 ± 0.098	(0.130, 0.291)	0.213 (0.120, 0.305)

Table 3.3 Testing the impact of sub-grouping NMR measurements according to the carbon atoms in a metabolite. GA3P cumomers were selected for testing the hypothesis under the three cases mentioned below. The perturbed measurements fitted poorly as represented by the increased residuum, whereas subgrouping the perturbed measurements improved the fit. Sub-grouping did not influence the flux solutions.

Metabolite	Cumomer group	Abundance	sd	Cumomer
Case1: Actual measurements				
GA3P	C123	0.05563	0.00278	#1xx
		0.00092	0.00005	#11x
		0.01422	0.00071	#x1x
		0.00374	0.00019	#x11+#11x
		0.23015	0.01151	#xx1
		0.00282	0.00014	#x11
Case 2: With perturbed measurements - carbons 2 and 3 were hypothesized to be altered by about 20% from their true values				
GA3P	C123	0.05563	0.00278	#1xx
		0.00092	0.00005	#11x
		0.01707	0.00085	#x1x
		0.00449	0.00022	#x11+#11x
		0.27617	0.01381	#xx1
		0.00338	0.00017	#x11
Case 3: Subgrouping of Case 2 measurements				
GA3P	C1	0.05563	0.00278	#1xx
		0.00092	0.00005	#11x
	C2	0.01707	0.00085	#x1x
		0.00449	0.00022	#x11+#11x
	C3	0.27617	0.01381	#xx1
		0.00338	0.00017	#x11

Table 3.4 Effect of sub-grouping of GA3P on the flux estimates. Sub-grouping of the cumomer measurements did not compromise accurate estimation of the fluxes

	Expected	**Disrupted**	**Subgrouped**
Vemp1__NET	0.500 ± 0.018	0.502 ± 0.017	0.500 ± 0.018
Vppp2__XCH	0.796 ± 0.194	0.640 ± 0.135	0.798 ± 0.215
Vppp3__XCH	0.714 ± 29.959	0.152 ± 0.064	0.880 ± 51.848
Vppp4__XCH	0.829 ± 0.214	0.941 ± 0.269	0.828 ± 0.222
Vppp5__XCH	0.191 ± 3.436	0.990 ± 0.432	0.176 ± 3.623
Vppp6__XCH	0.213 ± 0.097	0.220 ± 0.086	0.213 ± 0.098

Table 3.5 Weighted sum of square residuals of GA3P cumomers derived from global best fit solutions under the three cases mentioned in Table 3.3. Sub-grouping reduced the errors between individual subgroups.

		Case1: *Actual*	***Case2:*** *Disrupted*	***Case3:*** *Subgrouped*
GA3P	$1xx	0.0001	4.4751	0.0002
GA3P	$11x	0.0004	0.4360	0.0003
GA3P	$x1x	0.0000	1.0849	0.0000
GA3P	$x11+$11x	0.0001	1.4353	0.0000
GA3P	$xx1	0.0001	0.0511	0.0000
GA3P	$x11	0.0003	0.3931	0.0000
Total Residuum		**0.001**	**7.876**	**0.0005**

Table 3.6 Flux estimates for the PPP model using single and double models fed with a combination of 1- [^{13}C]glucose [1] and/or 2-[^{13}C]glucose [2] as substrate (s). Best fit fluxes are represented with standard deviations obtained by *EstimateStat.* Monte Carlo flux space (MC) at 68% confidence limits (section 3.2.4.2), and Confidence interval (CI) at 68% by the method of Antoniewicz et al., 2006 (section 3.2.4.3) are presented as (minimum, maximum). "nd" implies non deteriminable flux.

Flux	Model	Expected flux	Bestfit (*EstimateStat*)	Monte Carlo flux space 68% (min, max)	CI 68% (min, max)
Vemp1_NET	Single [1]	0.5	0.50 ± 0.022	(0.480, 0.518)	(0.479, 0.523)
	Single [2]	0.5	0.50 ± 0.011	(0.488, 0.506)	(0.491, 0.509)
	Double [1],[1]	0.5	0.50 ± 0.015	(0.487, 0.514)	(0.484, 0.515)
	Double [2],[2]	0.5	0.50 ± 0.008	(0.492, 0.504)	(0.493, 0.506)
	Double [1],[2]	0.5	0.50 ± 0.008	(0.493, 0.506)	(0.492, 0.508)
Vppp2_XCH	Single [1]	0.8	0.80 ± 0.213	(0.686, 0.957)	(0.665, 0.99)
	Single [2]	0.8	0.80 ± 1.680	(0.666, 0.99)	(0.650, 0.99)
	Double [1],[1]	0.8	0.80 ± 0.151	(0.693, 0.919)	(0.687, 0.947)
	Double [2],[2]	0.8	0.800 ± 1.188	(0.670, 0.99)	(0.670, 0.99)
	Double [1],[2]	0.8	0.800 ± 0.153	(0.699, 0.949)	(0.685, 0.952)
Vppp3_XCH	Single [1]	0.839	0.882 ± 50.08	(0.169, 0.99)	nd
	Single [2]	0.839	0.851 ± 60.77	(0.136, 0.99)	nd
	Double [1],[1]	0.839	0.882 ± 35.41	(0.191, 0.99)	nd
	Double [2],[2]	0.839	0.851 ± 42.97	(0.139, 0.99)	nd
	Double [1],[2]	0.839	0.845 ± 30.07	(0.192, 0.99)	nd
Vppp4_XCH	Single [1]	0.826	0.827 ± 0.222	(0.711, 0.99)	(0.660, 0.99)
	Single [2]	0.826	0.826 ± 1.797	(0.602, 0.99)	(0.600, 0.99)
	Double [1],[1]	0.826	0.827 ± 0.157	(0.733, 0.99)	(0.70, 0.99)
	Double [2],[2]	0.826	0.826 ± 1.27	(0.122, 0.99)	(0.61, 0.99)
	Double [1],[2]	0.826	0.827 ± 0.148	(0.719, 0.953)	(0.690, 0.99)
Vppp5_XCH	Single [1]	0.180	0.177 ± 3.508	(0.155, 0.99)	(0.12, 0.99)
	Single [2]	0.18	0.179 ± 4.651	(0.143, 0.99)	(0.130, 0.99)
	Double [1],[1]	0.18	0.177 ± 2.481	(0.169, 0.868)	(0.120, 0.99)
	Double [2],[2]	0.18	0.179 ± 3.289	(0.169, 0.99)	(0.13,0.99)
	Double [1],[2]	0.18	0.179 ± 2.327	(0.172, 0.812)	(0.120, 0.99)
Vppp6_XCH	Single [1]	0.213	0.213 ± 0.098	(0.130,0.291)	(0.120, 0.305)
	Single [2]	0.213	0.213 ± 0.159	(0.137, 0.307)	(0.145, 0.300)
	Double [1],[1]	0.213	0.213 ± 0.069	(0.144, 0.262)	(0.138, 0.284)
	Double [2],[2]	0.213	0.213 ± 0.112	(0.174, 0.872)	(0.162, 0.275)
	Double [1],[2]	0.213	0.213 ± 0.048	(0.162, 0.256)	(0.165, 0.260)

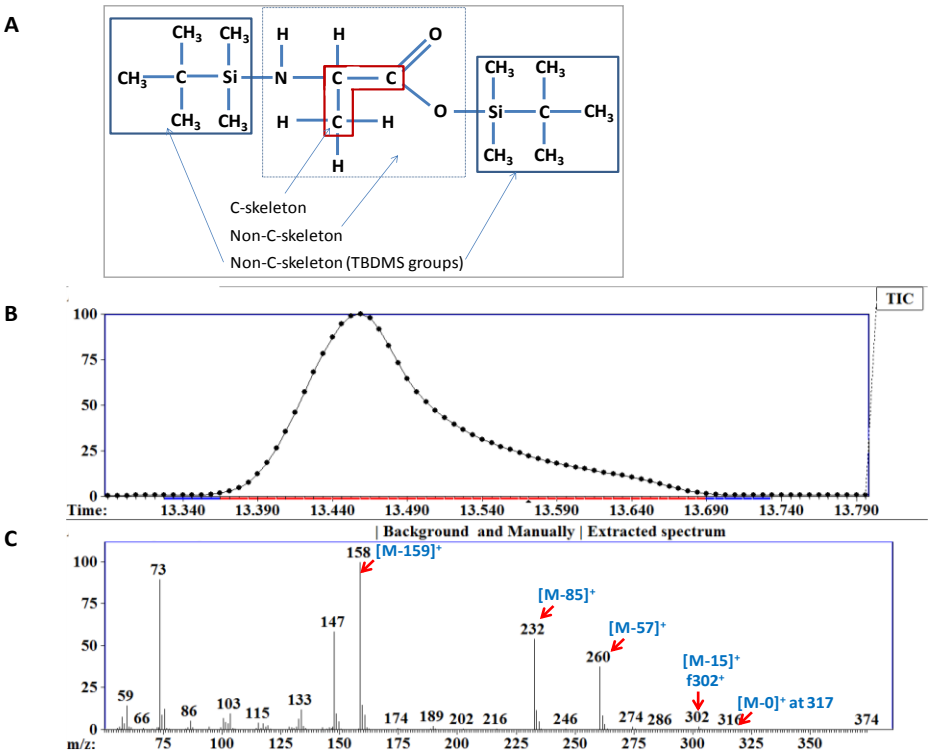


Figure 3.15 GC-MS spectral analysis of alanine derivatised with TBDMS. (A) Structure of alanine-2TBDMS ($C_{15}H_{35}O_2N_1Si_2$) indicating the carbon skeleton and non-carbon component of the parent metabolite and the components added by chemical derivitisation. (B) The total ion current (TIC) elution profile of alanine-2TBDMS during GC. (C) The mass spectrum of fragment ions generated from alanine-2TBDMS by ionization during MS. The structures of characteristic fragments indicated by arrows are described in Table 3.7.

Table 3.7 Fragmentation ions of alanine-2TBDMS formed by ionization during MS. The loss of structures from the complete molecule $[M-0]^+$ results in fragment ions with different masses (m/z).

Fragment	m/z (m+0)	Lost from C-skeleton	Lost from Non-C-skeleton	Remainder of derivative	Carbon retained from C-skeleton
$[M-0]^+$	317	-	-	$(C_6H_{15}Si)_2$	1,2,3
$[M-15]^+$	302	-	CH_3 from TBDMS	$C_5H_{12}SiC_6H_{15}Si$	1,2,3
$[M-57]^+$	260	-	$C(CH_3)_3$	$C_2H_6SiC_6H_{15}Si$	1,2,3
$[M-85]^+$	232	CO	$C(CH_3)_3$	$C_2H_6SiC_6H_{15}Si$	2,3
$[M-159]^+$	158	COO	$C_6H_{15}Si$	$C_6H_{15}Si$	2,3
$[f302]^+$	302	CH_3	-	$(C_6H_{15}Si)_2$	1,2

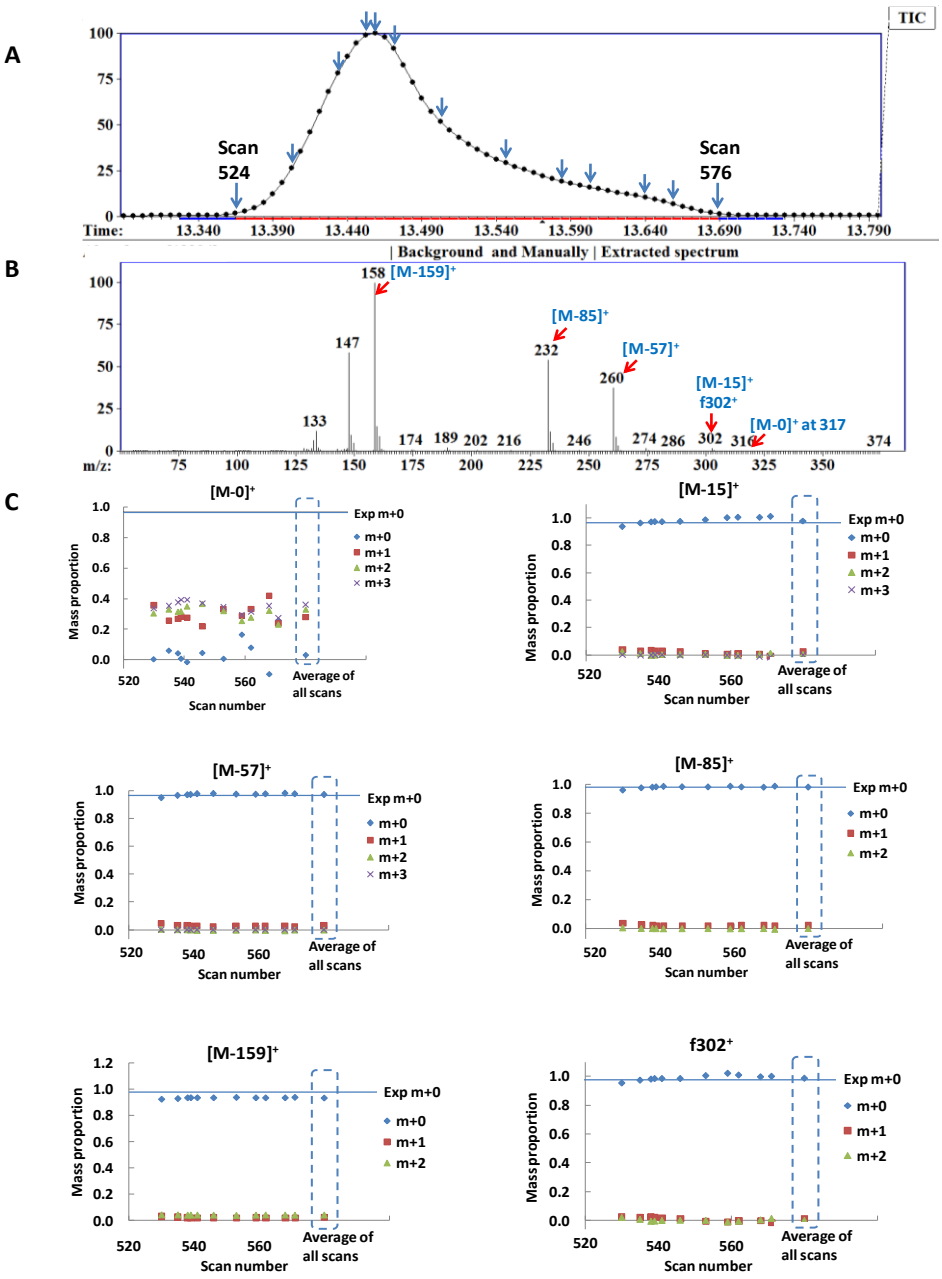


Figure 3.16 MID analysis of alanine-2TBDMS standard. (A) GC profile of derivatised alanine with arrows indicating scans from which MID were extracted for analysis. MID of average scan is obtained by extracting data from all scans between the peak boundaries (scan 524-576). (B) Representative MS profile of alanine-2TBDMS with its characteristic fragment ions (C) Relative mass isotopomer abundance of each fragment ion of alanine measured in selected individual scans and the average of all scans (dashed rectangle). The mass fractions are validated for accuracy by comparing with the expected values at natural abundance (presented as horizontal line for m+0). Note that mass effects due to natural abundance of mass isotopes of common atoms were corrected prior to calculations.

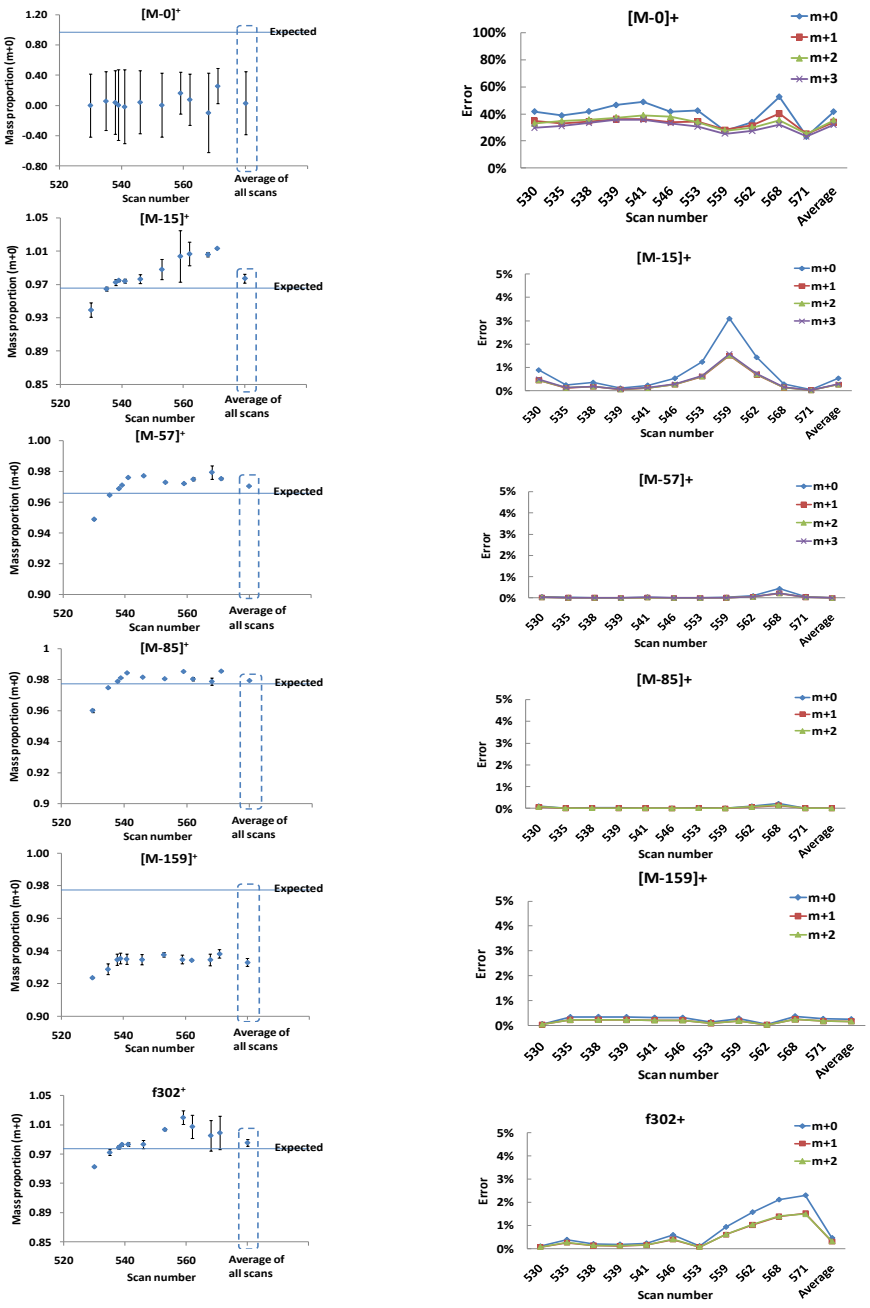


Figure 3.17 The accuracy and precision of isotopomer measurements of alanine-2TBDMS fragments analysed by GC-MS. The graphs relate to data presented in Fig 3.16. A. Relative abundance of mass isotopomer, m+0 in each scan across the metabolite elution from GC and the averaged-scans of the eluted peak. Accuracy is determined by comparing the measured and expected values of isotopic abundance. The precision levels of each observation are presented as standard error bars. B. Precision (mol% error) of all mass isotopomers in each fragment ion. Fragments with poor accuracy (eg., $[M\text{-}0]^+$) and/or precision ($f302^+$) are often rejected. See text for discussion.

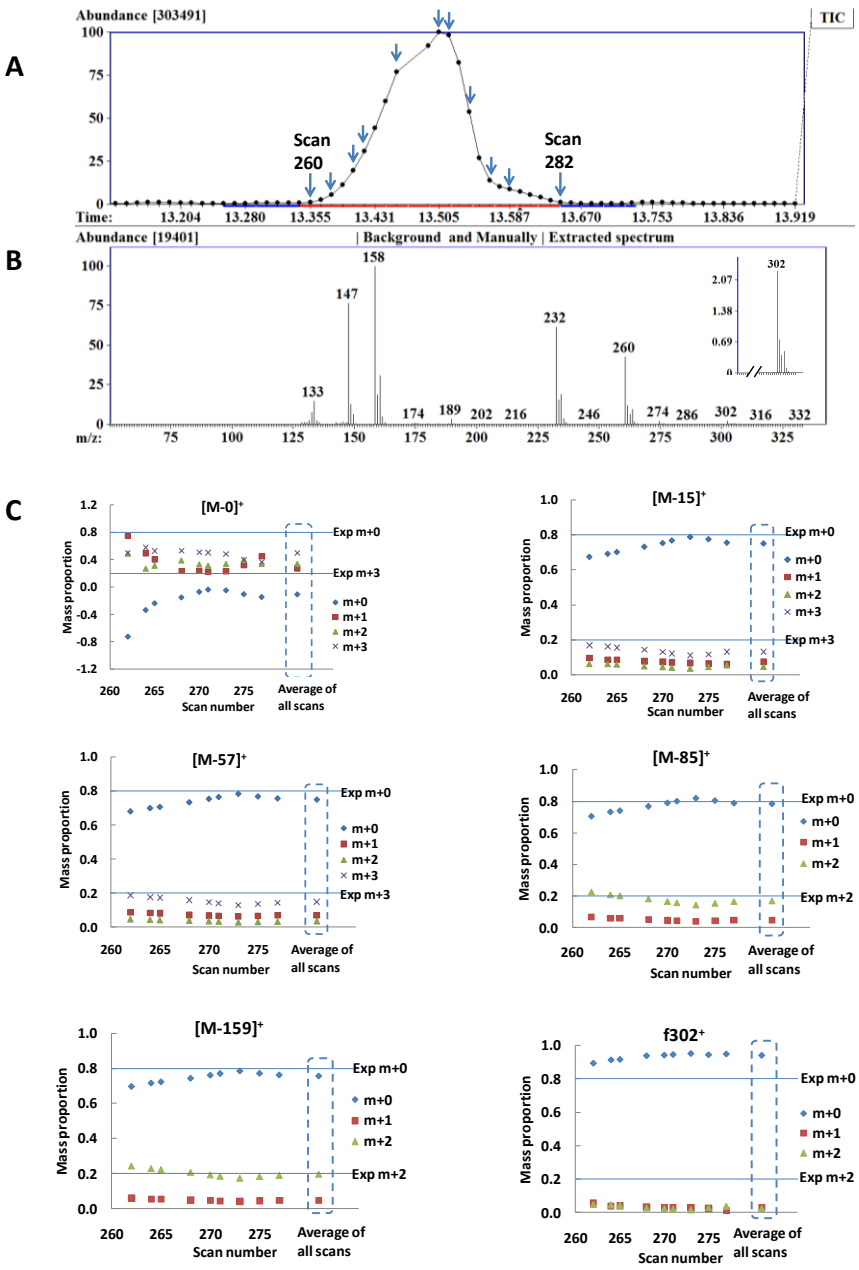


Figure 3.18: MID analysis of ~20% [U-^{13}C]alanine-2TBDMS obtained from Arabidopsis cells. (A) Gas chromatogram with arrows pointing to each scan from which MID was extracted for analysis. MID of average scan is extracted over entire peak boundaries. B) MID of [U-^{13}C] alanine-2TBDMS with its fragment ions C) Mass fractions of alanine carbons of each fragment ion measured across the selected individual scans and the average of all scans (dotted vertical rectangle). The mass fractions are validated for accuracy by comparing with the expected values of approximately 20% [U-^{13}C] alanine (presented as horizontal line). Note that mass effects due to natural abundance of mass isotopes of common atoms were corrected prior to calculations.

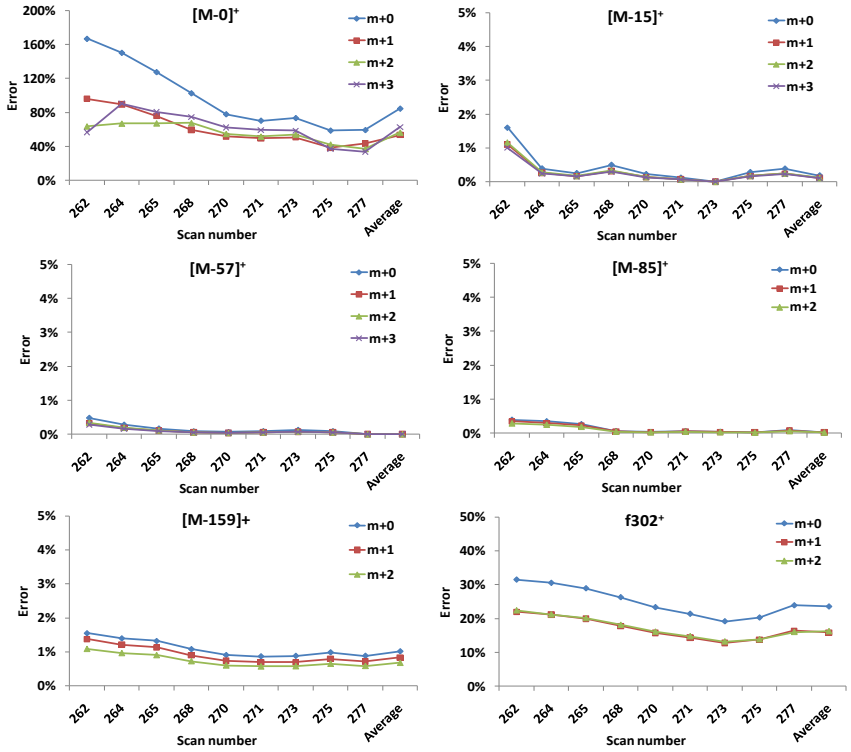


Figure 3.19: Precision (mol % error) of all the isotopomers fractions presented in Fig 3.18. Fragments with poor precision are either rejected or accepted with the measurement errors defined. See text for more discussion

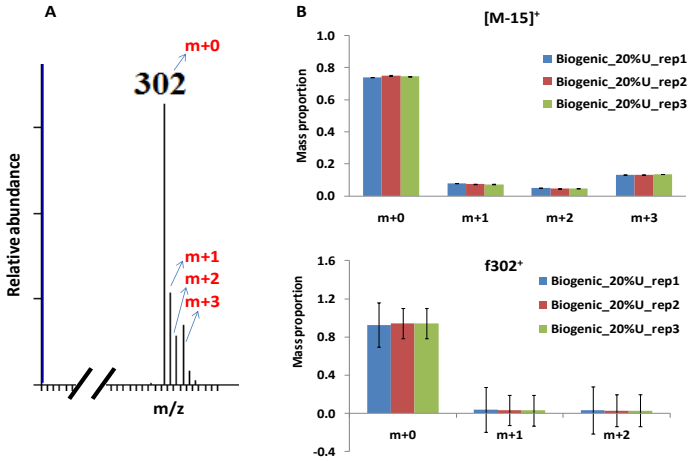


Figure 3.20: Analysis of the MID of the overlapping fragments $[M-15]^+$ and $f302^+$ from biogenic ≈20% [U-^{13}C] alanine-2TBDMS. The fragments overlap at m/z 302. $[M-15]^+$and $f302^+$ contains 3- and 2-carbon atoms of alanine, respectively, both with mass isotopomers m+0 at 302. The MID (A) was extracted and the mass fractions and associated errors (B) are calculated after mass correction (using MSCorr). Analysis is conducted from three separate replicate experiments. Fragment $f302^+$ is rejected whereas fragment $[M-15]^+$ accepted based on the accuracy and precision.

Table 3.8A Average ^{13}C of standard and biogenically derived unlabelled alanine-2TBDMS fragments. The fragments are validated based on their accuracy and precision levels. The expected average ^{13}C is ~1.13%.

Fragments	Mass (m/z) range	Alanine C-atoms	Percentage ^{13}C content (mean ± SE, (n=3)		Deviation from accuracy (mol%)	Validation
			Standard	Biogenic		
$[M-0]^+$	317-320	1,2,3	69.37 ± 1.5	74.13 ± 2.3	68.2-72.99	Reject
$[M-15]^+$	302-305	1,2,3	0.81 ± 0.11	1.04 ± 0.04	0.08-0.32	Accept
$[M-57]^+$	260-263	1,2,3	0.94 ± 0.05	1.03 ± 0.03	0.10-0.18	Accept
$[M-85]^+$	232-234	2,3	1.01 ± 0.03	1.00 ± 0.05	0.12-0.13	Accept
$[M-159]^+$	158-160	2,3	5.43 ± 0.10	5.55 ± 0.01	4.3-4.4	Reject
$[f302]^+$	302-304	1,2	0.77 ± 0.09	1.01 ± 0.09	0.11-0.36	Accept

Table 3.8B Average ^{13}C of biogenic derived ~20% [U-^{13}C] alanine-2TBDMS fragments. The fragments are validated based on their accuracy and precision levels. The expected average ^{13}C is about ~20%.

fragments	mass (m/z) range	alanine C-atoms	average ^{13}C ± precision (n=3)	deviation from accuracy (mol%)	validation
$[M-0]^+$	317-320	1,2,3	80.31 ± 2.80	60.31	Reject
$[M-15]^+$	302-305	1,2,3	18.92 ± 0.28	1.08	Accept
$[M-57]^+$	260-263	1,2,3	19.58 ± 0.34	0.42	Accept
$[M-85]^+$	232-234	2,3	19.66 ± 0.32	0.34	Accept
$[M-159]^+$	158-160	2,3	22.16 ± 0.19	2.16	Reject
$[f302]^+$	202-304	1,2	4.62 ± 0.64	15.38	Reject

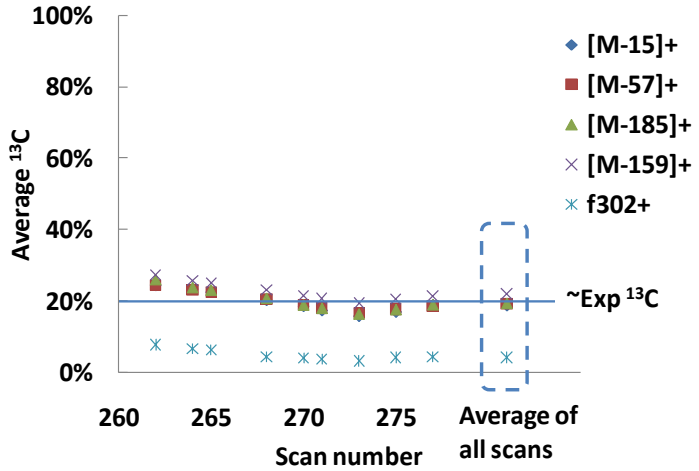


Figure 3.21 Average ^{13}C of biogenic ~20% [U-^{13}C]alanine-2TBDMS fragments analysed from MID of individual scans or averaged-scans (dotted rectangle). Averaged-scans are a representative of individual scans and is used for validations.

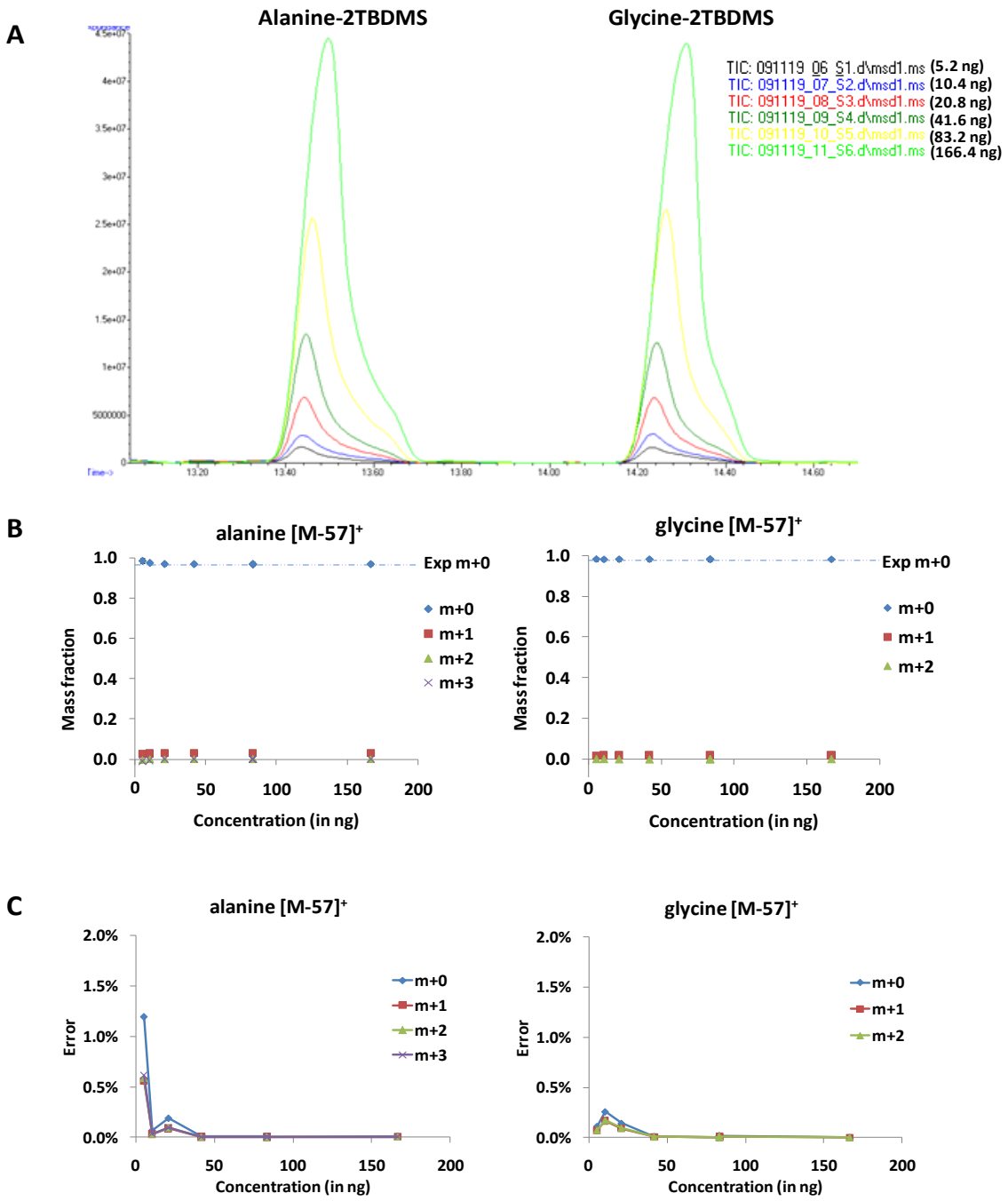


Figure 3.22 Concentration effect of metabolites on the MID. (A) GC of alanine-2TBDMS and glycine-2TBDMS at different concentrations (5.2 to 166.6 ng/µl) (B) Measured m*ass fractions* of alanine-$[M-57]^+$ and glycine-$[M-57]^+$. Accuracy of mass isotopomer m+0 in relation to expected abundance (dotted line) is presented. (C) Precision of mass isotopomers at each concentration (mol% error) shows errors < 1.2 mol%. At very low concentrations the errors were marginally higher and were better at higher concentrations. However, no significant trends of concentration effects were detected.

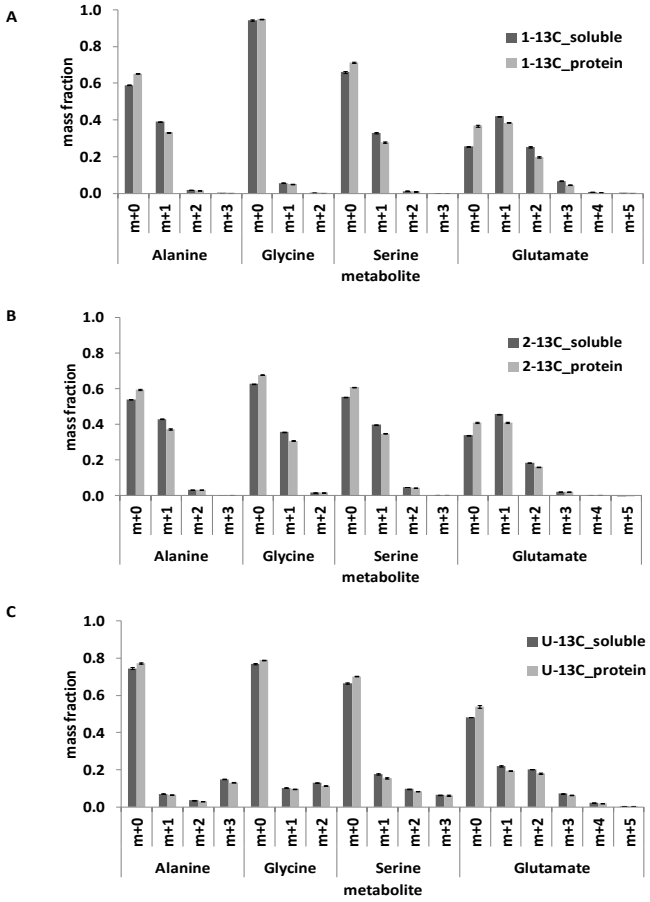


Figure 3.23 Mass isotopomer fractions of representative soluble and protein-derived amino acid fragments [M-57^{+}] obtained from A) 99% [1-^{13}C]glucose B) 99% [2-^{13}C]glucose and C) 20% [$^{13}C_6$]glucose fed *Arabidopsis* suspension. The mass isotopomer fractions significantly differ. The greater m+0 fraction in protein-derived amino acids implies the existence of pre-existing unlabelled protein pools. Values are the mean ± SE from analysis of 3 replicates.

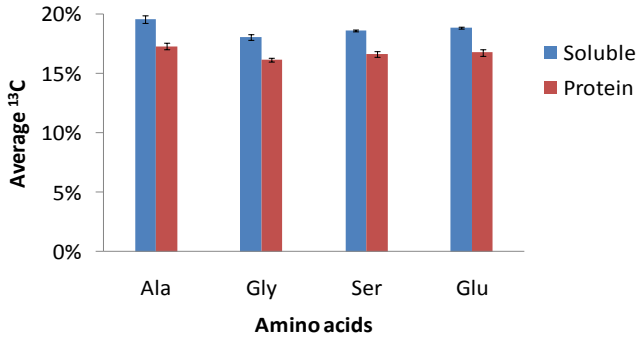


Figure 3.24 Average ^{13}C of representative soluble and protein-derived amino acid fragments [M-57^{+}] obtained from 20% [$^{13}C_6$]glucose fed *Arabidopsis* suspension. Average ^{13}C in protein-derived amino acids are less than the soluble implying the existence of pre-existing unlabelled protein pools that needs to be accounted for flux analysis studies. Soluble amino acids are at near isotopic steady state with an expected average of about 20% (see text).

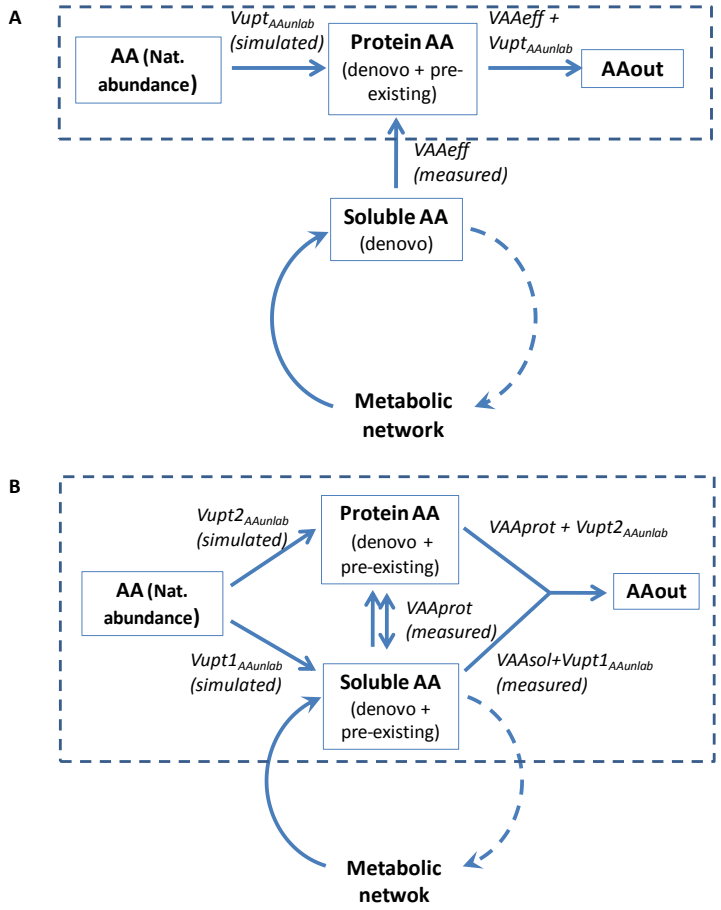


Figure 3.25 *In silico* modeling adjustments to account for pre-existing unlabelled metabolite pools. A) When protein-derived amino acids (ProteinAA) is a mixture of *de novo* synthesized and pre-existing unlabelled pool, additional proxy reactions are incorporated with the freedom to simulate uptake of unlabelled substrate (AA at natural abundance, $Vupt_{AAunlabelled}$). The biased mass fractions of protein are accounted for by simulating relative contributions of $Vupt_{AAunlabelled}$(proxy) in relation to metabolically synthesized (VAAeff). The uptake reaction is compensated by an additional efflux (VAA_{eff}+$Vupt_{AAunlab}$) so that there is no net influence on other fluxes. B) When both protein-derived and soluble amino acids (SolubleAA) are unlabelled and there is significant turnover, many proxy reactions as described above (in A) can be incorporated. Relative biosynthetic fluxes of soluble (VAAsol) and protein (VAAprot) would be needed.

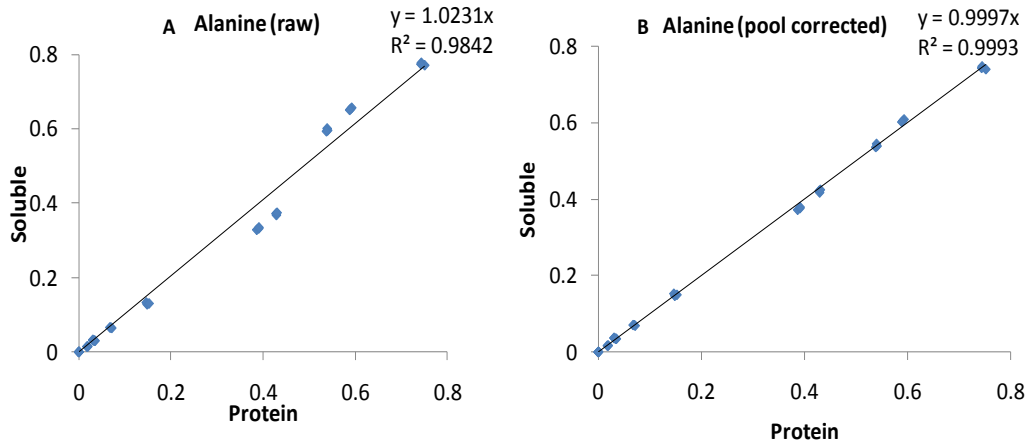


Figure 3.26 Linear regression of soluble vs. protein-derived alanine mass isotopomers measured (A) and after mathematical pool correction (B). All the mass isotopomers of alanine presented in Fig 1 are used in the analysis.

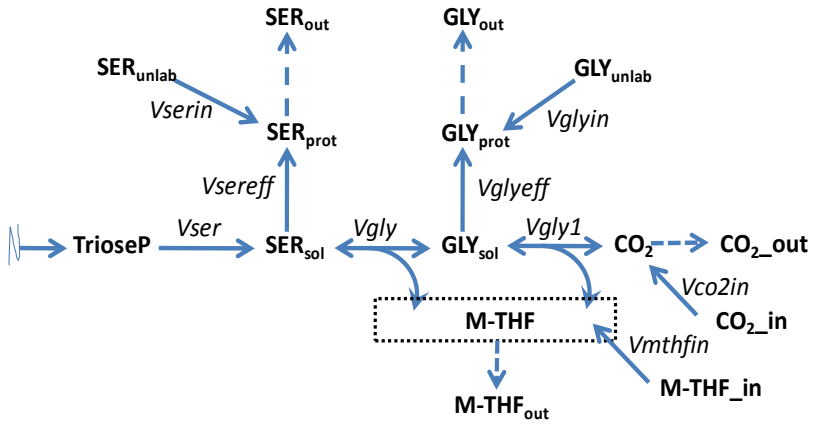


Figure 3.27 Case illustrating the use of modeling adjustments to account for pre-existing unlabelled pools in protein-derived glycine (GLY_{prot}) and serine (SER_{prot}). The proxy uptake reactions *Vserin* and *Vglyin* that take up unlabelled serine and glycine respectively simulates the proportion of pre-existing protein pools. The soluble pools are simulated by reversible reactions of glycine decarboylation (glycine ↔ CO_2 + MTHF) and serine hydroxymethyl transferase (serine ↔ glycine + MTHF) to allow the entry of CO_2 and MTHF into the pools.

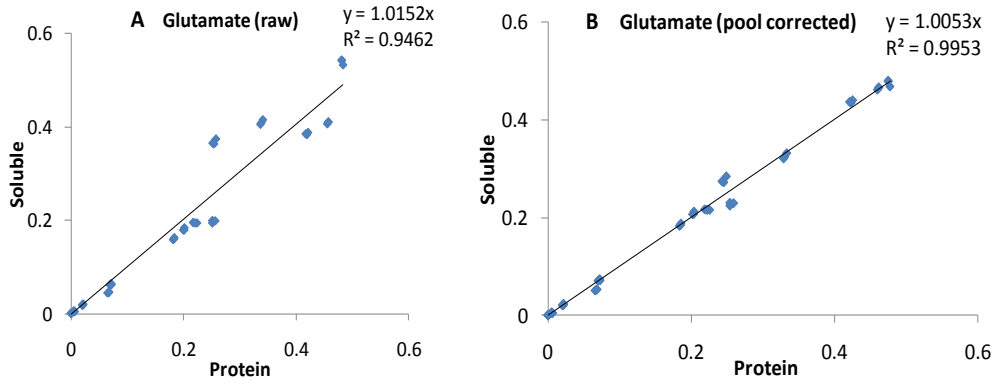


Figure 3.28. Linear regression of soluble vs. protein-derived glutamate mass isotopomers measured (A) and after pool correction mathematically (B). All the mass isotopomers of glutamate presented in Fig 1 are used in the analysis.

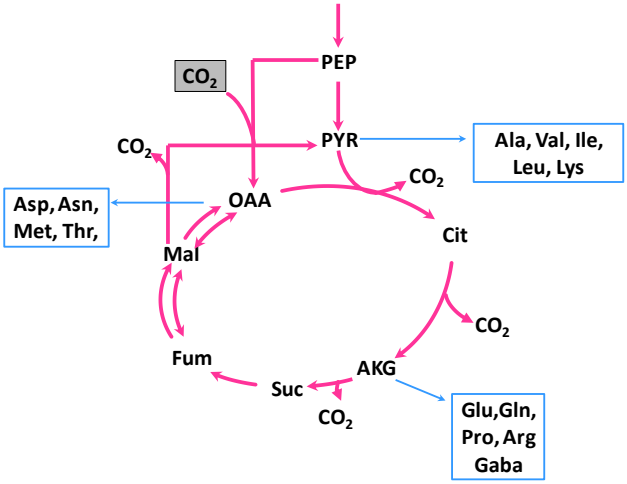


Figure 3.29. Route of dilution of TCA cycle intermediates and amino acids synthesized from them. CO_2 can enter the metabolic pathways via the reaction of PEP carboxylase (PEP + CO_2 → OAA) diluting the labeling of TCA cycle intermediates and amino acids synthesized from them. The average ^{13}C in TCA cycle intermediates (at isotopic steady state) are used to correct the pre-existing pools in protein-derived amino acids.

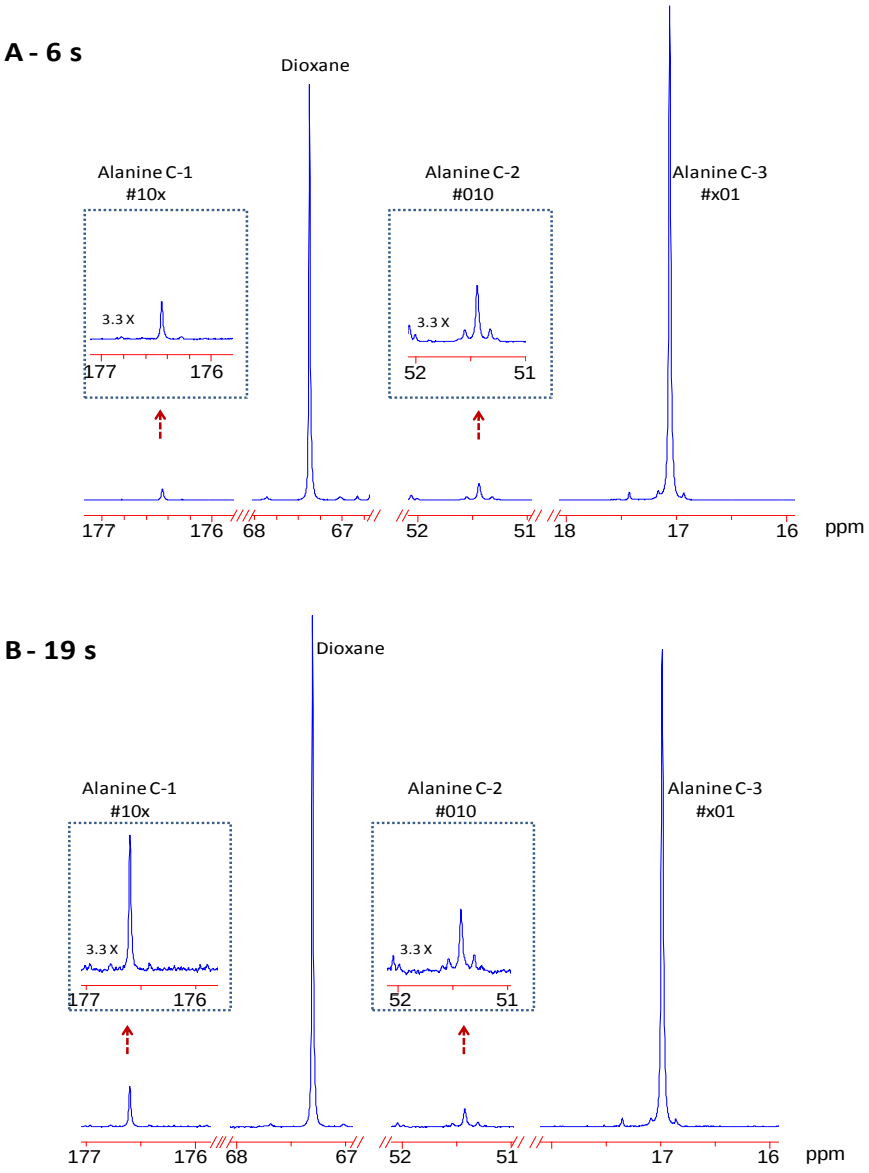


Figure 3.30 Comparison of relaxation delays of 6s (A) and 19s (B) on ^{13}C peak intensities of alanine carbons. Alanine is chosen as its isotopomers covers the complete range of the spectrum from 17 ppm (C-3) to 176.6 ppm (C-1; carboxyl group) and is detectable in most of the PCA extracts obtained from Arabidopsis cultures grown on [1-^{13}C]glucose. Spectra were acquired from the same sample using either 6 s (A) or 19 s (B) relaxation delay. Each carbon position of the isotopomers is defined by 1 (^{13}C), 0 (^{12}C) or x (^{13}C or ^{12}C). Dioxane was used as an internal standard. The carboxyl carbon of alanine (C-1) is underestimated when the spectra is not fully relaxed as can be qualitatively compared from the insets. The actual peak abundances are presented in Fig 3.31.

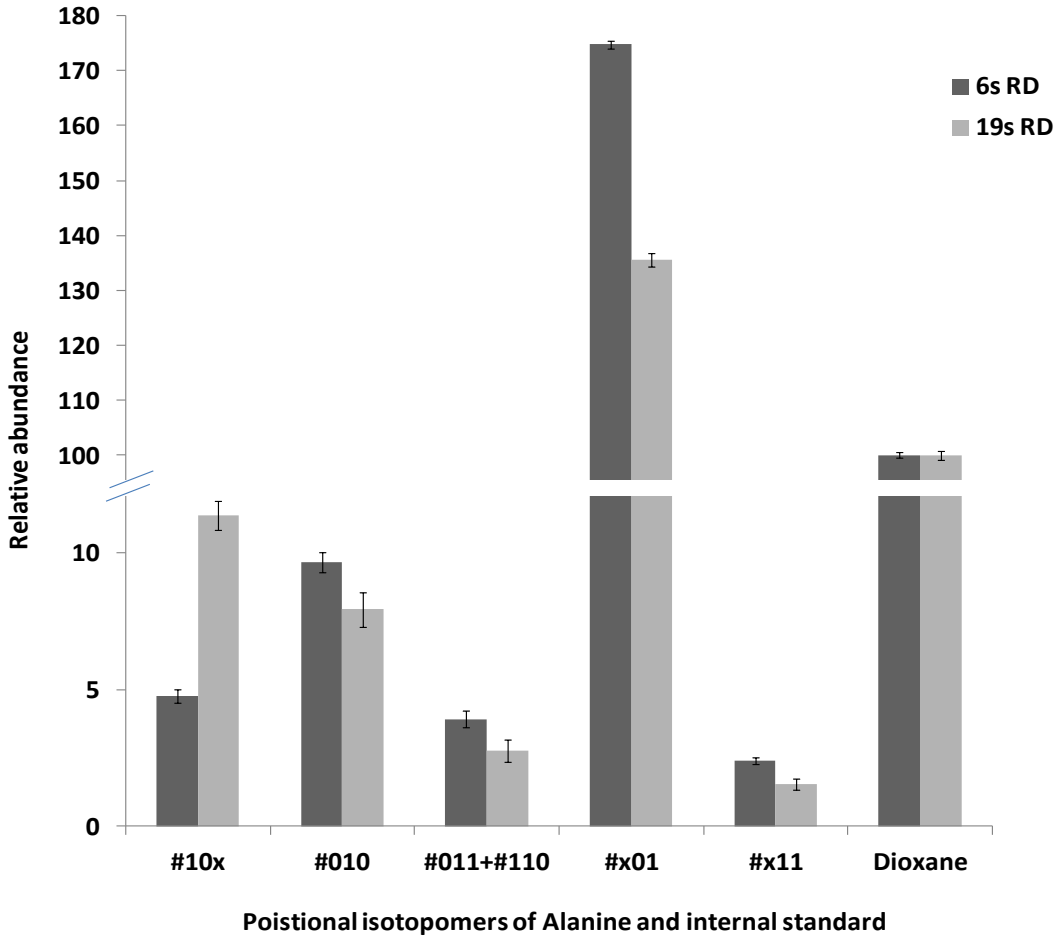


Figure 3.31 Effect of relaxation delay on the isotopomer abundances of a representative metabolite alanine obtained from ^{13}C spectra (Fig 3.30). The carboxyl carbon of alanine (C-1, #10x) is underestimated when the spectra is not fully relaxed at 6 s. Accordingly, spectra obtained using a 19 s relaxation delay were acquired for accurate flux analysis. Relative peak abundance is presented as the mean ± SD of measurements calculated as mentioned in section 3.5.3.

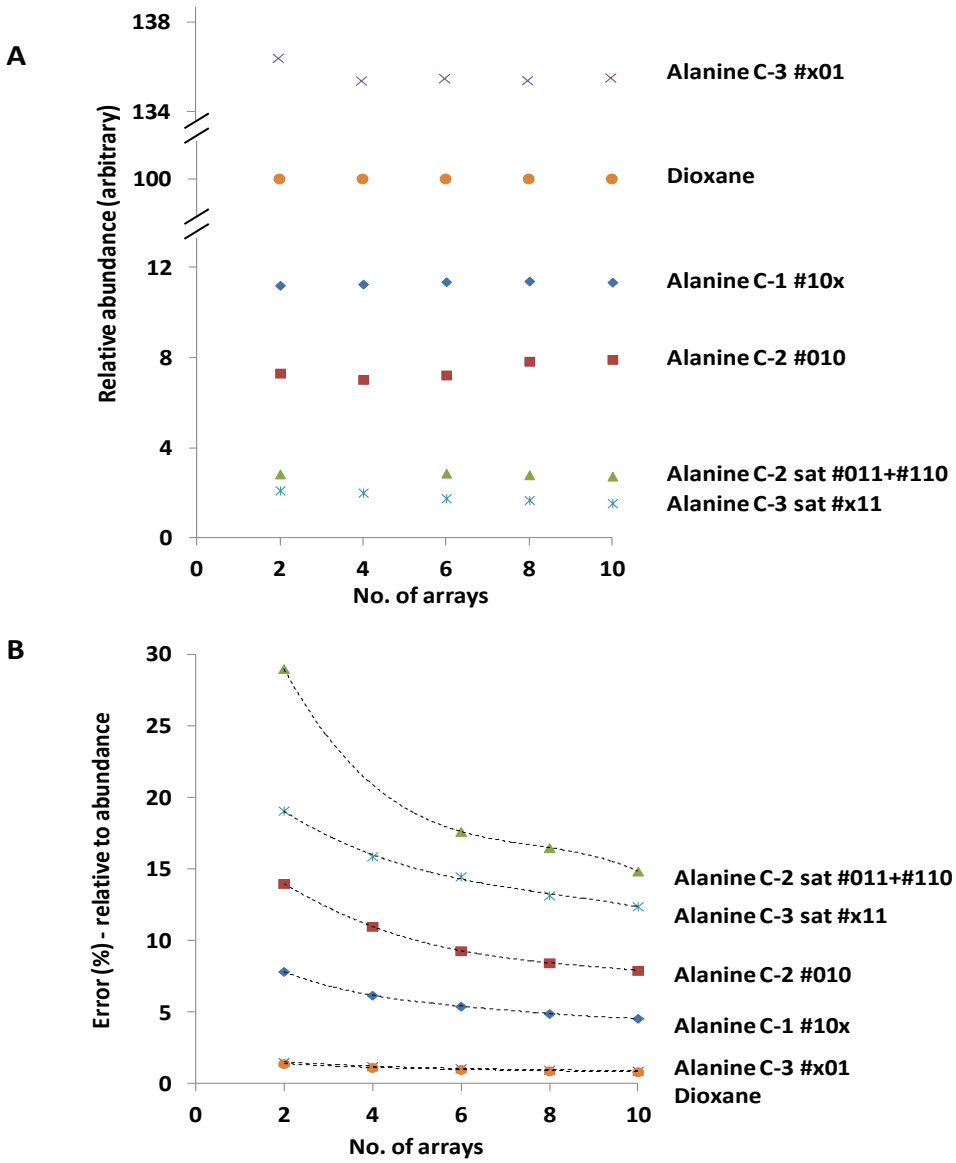


Figure 3.32 Effect of acquisition time on the relative abundances (A) and percent errors (B) of positional isotopomers of alanine obtained from one-dimensional ^{13}C-NMR. The reliability of the measurements improved with the number of arrays. The measurements are made from a spectra of representative perchloric acid extract containing soluble metabolites of *Arabidopsis* suspension culture grown on [1-^{13}C]glucose (Fig 3.30B). The spectra were accumulated for 46 h in 10 arrays (or blocks), with each array containing 1024 scans acquired using a relaxation delay of 19 s. The abundances and errors were obtained as described in Sections 3.5.3 (Line-fitting and error estimation). Each carbon position of the isotopomers is defined by 1 (^{13}C), 0 (^{12}C) or X (^{13}C or ^{12}C). No reliable signals were detected for satellite peak (sat) of Alanine C-2 at 4 arrays and hence these isotopomers are omitted from the comparison. Dioxane was used as an internal standard whose errors were estimated to be same as C-3 of alanine (#x01).

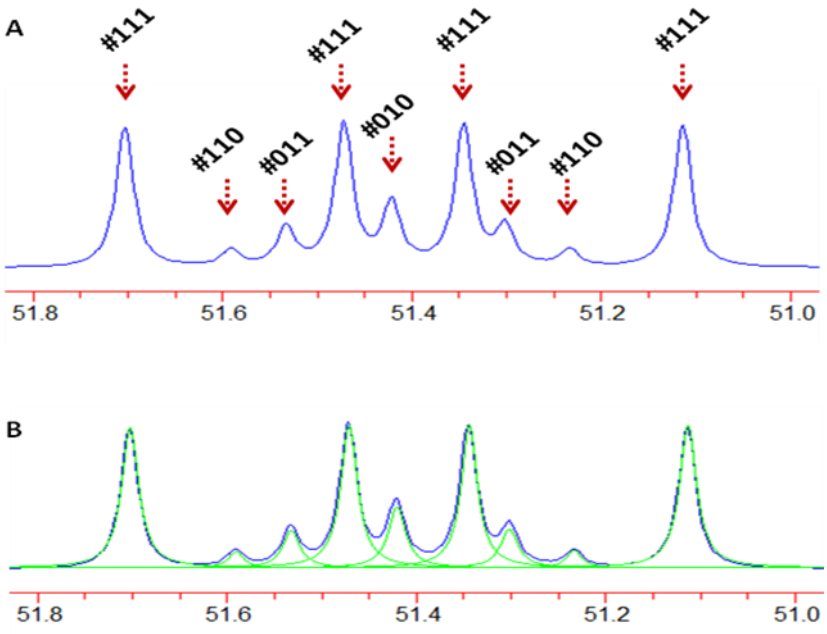


Figure 3.33 A) ^{13}C spectra of alanine C-2 isotopomers with its satellite peaks. B) Deconvoluted spectra by line-fitting (green) that was used for obtaining the relative intensities and errors of all the isotopomers. This is a representative spectra obtained from [U-^{13}C]glucose fed *Arabidopsis* culture.

4 Subcellular flux analysis of central metabolism in a heterotrophic *Arabidopsis thaliana* cell suspension using steady-state stable isotope labeling

4.1 Introduction

4.1.1 Objectives and data source for flux analysis

The major objective of this chapter was to validate the robustness of the computational strategies developed in Chapter 3 and to establish subcellular compartmented flux maps of heterotrophic *Arabidopsis thaliana* cell suspension with a focus on carbohydrate oxidation via glycolysis and oxidative pentose phosphate pathway (OxPPP) in the cystosol and plastid. This work has been published in Plant Physiology (Masakapalli et al., *2010*). The experimental datasets used in this chapter were obtained from steady state [^{13}C]glucose feeding experiments performed by my co-authors. My contribution to this chapter was the reanalysis of the existing ^{13}C NMR datasets, network formulations, mathematical modeling, flux analysis and write-up along with my supervisors (Prof R.G. Ratcliffe and Dr N.J. Kruger) whom I acknowledge.

4.1.2 Subcellular compartmentation in plants

Extensive subcellular compartmentation, with unique locations for many steps and pathways, as well as the duplication of other steps and pathways in different compartments, adds greatly to the structural complexity of the plant metabolic network (Lunn, 2007; Kruger and Ratcliffe, 2008). The need to consider discrete pools of metabolites in specific compartments, and the transporters that link them, complicates the quest for a detailed, predictive understanding of the regulation of plant metabolism; and as a result it remains difficult to manipulate flows of

material through the central metabolic network in a predictable way (Carrari et al., 2003; Kruger and Ratcliffe, 2008; Sweetlove et al., 2008).

The emergence of steady-state metabolic flux analysis (MFA) as a practicable systems biology tool for generating flux maps of the central metabolic pathways in plants offers new opportunities for analyzing plant metabolic phenotypes (Ratcliffe and Shachar-Hill, 2006; Schwender, 2008; Libourel and Shachar-Hill, 2008; Kruger and Ratcliffe, 2009). In this approach, substrates labeled with stable isotopes are introduced into the network, and fluxes are determined by measuring the labeling of the system after it has reached an isotopic and metabolic steady state. Subcellular compartmentation of steps and pathways can be incorporated into the model that describes the redistribution of the label, and flux maps of central carbon metabolism in plant cells typically aim to distinguish between fluxes in the cytosol, mitochondria and plastids. In principle compartmented fluxes can be deduced from the labeling patterns of metabolites that are synthesized in a single compartment (Ratcliffe and Shachar-Hill, 2005, 2006; Allen et al., 2007), and this property has been exploited with varying degrees of success in different studies.

Most recent flux maps shows there is less consensus in the description of the pathways of carbohydrate oxidation, with the subcellular compartmentation of glycolysis and the oxidative pentose phosphate pathway (PPP) being described at various levels of detail. For example, a recent model of soybean (*Glycine max*) cotyledons included complete pathways for glycolysis and the PPP in both cytosol and plastids (Iyer et al., 2008); whereas models for developing sunflower embryos (Alonso et al., 2007), maize root tips (Alonso et al., 2007) and an Arabidopsis cell culture (Williams et al., 2008) confined the oxidative and non-oxidative steps of the PPP to the plastid. The limited compartmentation of the pathways of carbohydrate

oxidation in many flux maps can be traced to an analysis of developing oilseed rape embryos (Schwender et al., 2003) which concluded that the labeling data could be satisfactorily explained without duplication of glycolysis and the PPP. Since some of the enzymes of these pathways are known to be present in both compartments, the model implied fast exchange of intermediates across the plastid envelope, giving rise to a situation in which the physically compartmented pathways are functionally uncompartmented (Schwender et al., 2003; Ratcliffe and Shachar-Hill, 2006).

Although the extent to which the steps of the PPP are duplicated in the cytosol and plastids has not been fully established, most of the biochemical evidence suggests that the enzymes for the oxidative steps at least are present in both compartments (Kruger and von Schaewen 2003). So modeling the PPP as fully duplicated in both cytosol and plastids, or alternatively as exclusively plastidic, ignores the most likely structure of the network. Models based on all three network structures have now been tested in an Arabidopsis cell culture. Building on an MFA analysis that had been based on steady-state data from a single labeled substrate (Williams et al., 2008), similar experiments were conducted with multiple substrates and the labeling datasets were simultaneously fitted to models in which the subcellular compartmentation of the PPP was varied. Acceptable fits to the data were found for all three models, emphasizing the importance of building the most appropriate level of subcellular compartmentation into the model at the outset. Moreover the flux maps showed that the data were consistent with a major contribution from the cytosol to the oxidative steps of the PPP.

4.2 Biosynthetic outputs

The Arabidopsis cell culture was maintained in the dark with glucose as the sole respiratory substrate, and the principal activities of the metabolic network, together with the associated biosynthetic outputs, were assessed by determining the redistribution of radiolabel following metabolism of [U-^{14}C]glucose (Table 4.1). Overall the pattern of labeling was comparable to that observed previously for heterotrophic cell cultures of Arabidopsis and is similar to that reported for a wide range of non-photosynthetic plant material (Kruger et al., 2007). However, extending the analysis to include components that are frequently ignored during similar fractionation procedures, such as the organic phase obtained during chloroform/methanol fractionation and the medium following removal of cells, allowed a more comprehensive analysis of the fate of metabolized glucose and measurement of the labeling of lipids and ethanol. The data show that almost 40% of the metabolized glucose was released as CO_2, while less than 25% of the carbon was converted into biopolymers (comprising starch, cell wall, protein and lipid), with the remainder accumulating mainly as soluble sugars, organic acids and amino acids. While this analysis defines the major output fluxes associated with growth, further information is required to determine the fluxes that generate the precursors needed for biomass production, and this can be provided by steady-state MFA.

4.2.1 Metabolic network

In MFA, the metabolic network is reduced to a model that can account for the observed redistribution of the label. Typically this leads to a model in which: (i) multiple steps between branch points in the network are represented as a single reversible or irreversible step; and (ii) subcellular pools of metabolites are treated as single pools in the absence of compartment-

specific information about their labeling. The scope of the model is strongly influenced by the choice of label precursor, and by the extent and quality of the measurements that define the redistribution of the label after the system has reached an isotopic and metabolic steady state. It has already been shown that statistically robust flux maps of central carbon metabolism can be obtained for heterotrophic Arabidopsis cell suspensions using [1-^{13}C]glucose as the label source (Williams et al., 2008), and here the aim was to establish whether improved definition of the subcellular compartmention of carbohydrate oxidation could be obtained by combining the results of experiments with [1-^{13}C]glucose, [2-^{13}C]glucose and [U-$^{13}C_6$]glucose as label precursors.

The existing model, in which the oxidative and non-oxidative reactions of the PPP were confined to the plastids (Fig. 4.1A), was compared with two alternatives: (i) a model in which the oxidative steps of the PPP were allowed to occur in both the cytosol and plastids, with the non-oxidative steps confined to the plastids (Fig. 4.1B); and (ii) a model in which the complete PPP was present in both compartments (Fig. 4.1C). Models were constructed in the format used by the steady-state MFA software 13C-FLUX and full details of the model in which the oxidative steps of the PPP were duplicated in the cytosol and plastids (Fig. 4.1B) are given in Appendix 4.1. In essence the model defines the relationship between the carbon atoms in substrates and products for each step in the network so that the underlying fluxes can be deduced from the redistribution of label in the steady-state labeling experiment. The models used here are directly related in most respects to the one used earlier (Williams et al., 2008) and they were constructed in the same way. Several features of the models, and the biochemical networks they represent, should be noted.

First, analysis of the subcellular distribution of flux depends on obtaining information about the steady-state labeling of metabolic intermediates in specific compartments. Here the labeling of the plastidic pools was deduced from the labeling of starch, and from amino acids that are known to be synthesized exclusively in the plastid (Cys, Gly, His, Ile, Leu, Lys, Phe, Ser, Trp, Tyr, and Val). Similarly the labeling of cytosolic intermediates was deduced from the labeling of sucrose and alanine (Miyashita et al., 2007); and arginine, γ-aminobutyrate (GABA), glutamate, glutamine and proline were assumed to be derived from mitochondrial 2-oxoglutarate. In contrast it was not possible to deduce information about the subcellular labeling of oxaloacetate and malate, and the synthesis of these metabolites was not assigned to a specific compartment.

Secondly, because the model describes the redistribution of label through the network, expected biochemical features of the network may be concealed or omitted. For example flux through the GABA shunt is combined with the flux through the TCA cycle between 2-oxoglutarate and fumarate because these pathways have identical labeling outcomes in the model. Similarly sucrose cycling is omitted because theoretical analysis has shown that in steady-state analysis this can only be correctly analyzed with information on the labeling of the cytosolic and vacuolar glucose pools (Kruger et al., 2007). The flux supported by the putative glucose-6-phosphatase (Alonso et al., 2005) is omitted for the same reason, because there is no easy way to measure the labeling of the cytosolic and vacuolar glucose pools, and ignoring the compartmentation of glucose is likely to produce very misleading flux values (Kruger et al., 2007). In contrast, discrete cytosolic and plastidic pools of glucose 6-phosphate were retained in the model because the glucosyl units of sucrose and starch differed appreciably in their

patterns of labeling, establishing that they were synthesized from isotopically distinct precursors (Fig 4.2).

Thirdly, the steps catalyzed by transketolase (TK) and transaldolase (TA) in the non-oxidative steps of the PPP are represented in terms of half-reactions to maximize the possibility of resolving cytosolic and plastidic contributions to the PPP fluxes (Kruger and von Schaewen, 2003; Kleijn et al., 2005, Selivanov et al., 2005, Masakapalli et al., 2010).

4.3 Data handling and mathematical modeling

Heterotrophic Arabidopsis cell cultures were incubated with [1-^{13}C]glucose, [2-^{13}C]glucose or 10% [U-$^{13}C_6$]glucose, the label redistribution at isotopic and metabolic steady state was quantified by ^{13}C NMR, and flux solutions, constrained by measurements of glucose consumption and biosynthetic outputs (Table 4.2), were calculated in 13C-FLUX using the protocol summarized in Chapter 3 (Figures 3.8 and 3.14). Label was routinely measured in soluble metabolites, including amino acids, citrate, fumarate, GABA, malate and sucrose, in the amino acids derived from protein hydrolysate, and in the glucose derived from starch. The best fit flux solutions presented here were based largely on sucrose, starch and protein hydrolysate labeling data, and full details of the measurements used in the fitting process are given in Appendix 4.2. Measurement errors were generally set at 5% for amino acid data, and 15% for other metabolites, but a few measurements were assigned larger deviations as specified in Appendix 4.2.

Label measurements for a particular metabolite were initially grouped together for the fitting process, but subsequently it was established that better fits (lower residua) could be obtained by treating ^{13}C peaks with multiplet structure as separate subgroups. Simulations in 13C-

FLUX showed that the reduction in the residuum could be attributed to the inclusion of additional scaling factors, compensating for discrepancies in the relative intensities of the signals from different carbon atoms within a metabolite, and that the subgrouping procedure did not compromise the accuracy of flux estimation. Subgrouping multiplet signals avoids the need to establish the relative intensities of multiplets from the same metabolite, and it is used routinely when two dimensional NMR methods, which often mask the true relative intensities, are used to analyze [U-$^{13}C_6$]glucose labeling experiments (Sriram et al., 2004).

Each of the three labeling experiments was performed twice, and various combinations of experiments were analyzed simultaneously in 13C-FLUX by setting up models in which the reaction network was replicated two, three or six times. Software limitations, arising from the large size of the network and the quantity of data, placed some restriction on this approach, but a three-substrate model, comprising a sub-network for each labeled precursor, i.e. [1-^{13}C]glucose, [2-^{13}C]glucose or 10% [U-$^{13}C_6$]glucose, and two datasets for each sub-network, worked well, and a single flux solution corresponding to the entire dataset could be generated by constraining equivalent fluxes in the three sub-networks to be equal in the model definition file. Ultimately this procedure allowed a more reliable definition of the fluxes in the network than could be obtained by analyzing different labeling experiments separately (see below).

Preliminary fits of the data to the model were used to refine the model and to identify errors in the labeling data. In this iterative process, 13C-FLUX was used to generate 50-100 solutions for a given model and dataset. Changes were then made to either the model or the dataset, for example the addition of a pathway or the correction of an incorrectly annotated data point, and the fitting process was repeated. Comparing the residua for the five best fits provided the basis for accepting or rejecting a change to the model or the dataset. For example the reaction

catalyzed by isocitrate dehydrogenase, and the mitochondrial uptake of pyruvate, were both defined as unidirectional because making them reversible had no effect on the fit. Similarly CO_2 uptake at natural abundance was removed from the model because it had no discernable effect on the fit. In contrast adding unidirectional plastidic pyruvate uptake improved the fit, as did the inclusion of separate cytosolic and plastidic pools for PEP and triose phosphate. However, the latter metabolites were combined to create a single pool of 3-carbon phosphate esters in each compartment since this removed indeterminable fluxes from the network and had no effect on the fit. The iteration process lead to the model in Figure 4.3, and the validated dataset in Appendix 4.2, and thus provided a secure foundation for testing the extent to which the method could reveal the subcellular compartmentation of carbohydrate oxidation (Fig. 4.1).

4.4 Best fit fluxes and statistical analysis

The six datasets, comprising over 700 positional isotopomer measurements and 29 biomass fluxes, were initially analyzed in a three-substrate model based on the flux network in Figure 4.3. Estimates for the free fluxes were obtained by running the optimization routine 1000 times, generating a correspondingly large set of flux solutions, each with a set of predicted labeling measurements. Some of these solutions failed to satisfy constraints in the model, and were identified as infeasible solutions by the software; but the majority were feasible solutions, and these were extracted from the 13C-FLUX output file with a customized linux script (Appendix 2.4). The residua for the feasible solutions varied over a wide range (Fig. 4.4A), and a subset of the solutions corresponding to residua below 3000 was selected for further analysis. This cut-off was justified both by visual inspection of the distribution of residua (Fig. 4.4A) and by principal components analysis of the mean-centered, unit variance scaled flux solutions (Fig. 4.4B). The one dimensional scores plot corresponding to the first principal

component confirmed the similarity between the selected feasible solutions and provided further justification for rejecting the solutions with residua above the cut-off.

Monte Carlo simulations with bootstrap sampling of isotopomer measurements provides an effective method for characterizing the size and shape of the flux space (Wiback et al., 2004; Joshi et al., 2006) and the selected solutions showed considerable variability in the definition of the free fluxes (Fig. 4.5). Some fluxes were tightly defined, for example many of those in the TCA cycle, while others were normally distributed over a wider range of values. To reduce the hundreds of selected feasible solutions to a global best fit, the mean values from the Monte Carlo flux space were used as starting free flux values for a final run of the optimizer with the validated dataset. This run was performed without bootstrap Monte Carlo sampling, and it resulted in a global best fit solution with the lowest residuum (Table 4.3). These procedures were repeated for the models in which the full PPP was either confined to the plastids (Fig. 4.1A) or present in both compartments (Fig. 4.1C), and the final best fit fluxes for the three reaction networks are given in Table 4.3.

Statistically reliable flux values were generated for all three models of the PPP (Table 4.3) with only minor differences in residuum. The flux distribution for the steps involving sugar phosphates differed substantially for the three models (Fig. 4.6), and it was noticeable that the bulk of the flux through the oxidative steps of the PPP switched to the cytosolic pathway when this option was included in the network structure. Thus while the flux through the oxidative steps was necessarily confined to the plastids in the absence of a cytosolic pathway (Fig 4.1A, 4.6A), the plastidic pathway made only a minor contribution when the cytosolic pathway was included in the model (Figs 4.1B,C, 4.6B,C). The degrees of freedom in the models increased with the number of fluxes, and as expected this led to a small improvement in the residuum

achieved in the fitting procedure but an increase in the standard deviations for the flux estimates. For example including the full PPP in both cytosol and plastids (Fig 4.1C) produced a marginal improvement in residuum, but it produced a flux map in which 28 of the 33 fluxes shared by all three models had the largest standard deviations.

The validity of the flux maps was tested by comparing the predicted contribution of the carbon atoms of metabolized glucose with the relative specific activity of $^{14}CO_2$ released from specifically labeled glucose (Fig 4.7). Ratios of $^{14}CO_2$ labeling from [1-^{14}C]-, [2-^{14}C]-, [3,4-^{14}C]- and [6-^{14}C]glucose are typically used to characterize carbohydrate oxidation (Harrison and Kruger, 2008), and the pattern of $^{14}CO_2$ release predicted by the global best-fit flux estimates for each model was not significantly different from the experimentally determined pattern (Fig 4.7), implying that each of the flux maps provides an adequate explanation of the observed data for the release of $^{14}CO_2$ from positionally labeled glucose.

Since the models are indistinguishable, it can be concluded that the subcellular compartmentation of the PPP cannot be determined unambiguously from the available labeling data alone. However on current biochemical evidence the most likely model is the one in which the oxidative steps of the PPP are duplicated in the cytosol and the plastids (Fig 4.1B, 4.6B; Kruger and von Schaewen, 2003). The global best-fit solution for this network is presented in detail in Table 4.4, and it provides good agreement between the observed labeling data and the isotopomer abundances predicted from the global best fit flux map (Fig 4.8A). There is also excellent correspondence between the measured, scaled and predicted relative fractional abundance of positional isotopomers for individual metabolites – as exemplified by the data on protein-derived aspartate/asparagine (Fig 4.8B). More than 80% of the total residuum was confined to 176 measurements (25% of total), and the flux map obtained after

removal of these values was essentially identical to that obtained from the complete data set suggesting that the poorly fitting measurements did not have a major influence on the flux solution (Fig 4.9). Furthermore, there was no consistency in the identities of the isotopomers that contributed most to the residua between duplicate feeding studies involving identically labeled substrates, implying that the discrepancies between the observed and predicted values were due to poor estimates of measurement error, rather than a failure of the model to account adequately for the pattern of label redistribution in a particular section of the network.

Finally, *in silico* analysis of the chosen network revealed that the reliability of the flux estimates improved by combining data from labeling studies using [1-^{13}C]-, [2-^{13}C] and [U-$^{13}C_6$]glucose – this was particularly noticeable for the oxidative steps of the PPP in the cytosol and the plastids - and it also emphasized the importance of using well defined labeling measurements (Fig 4.10).

4.5 Discussion

4.5.1 Quantifying the compartmented fluxes of central metabolism

The subcellular compartmentation of metabolism in eukaryotes complicates steady-state MFA (Roscher et al., 2000), and the problem is acute in plants because of the more extensive duplication of steps and pathways in the metabolic network (Kruger and Ratcliffe, 2008). Capturing the characteristic compartmentation of the pathways of carbohydrate oxidation in a realistic model of heterotrophic plant metabolism exemplifies the problem, and it would be particularly useful if this could be done by comparing models for networks compartmented to different extents, since this would allow an *in vivo* assessment of the functional compartmentation of the network. Since most steady-state MFA models are based on over-

determined datasets, reflecting the wealth of stable isotope labeling information that can be obtained from nuclear magnetic resonance (NMR) or mass spectrometry (MS) (Ratcliffe and Shachar-Hill, 2006), it is easy to add extra steps to the model to allow it to match biochemical reality more closely. However the relationship between fluxes and measurements is not always obvious, and even if the extra fluxes are potentially resolvable they may not be determinable with any great confidence.

In practice many steady-state MFA investigations of the PPP in plants use models in which the compartmentation of the pathway is pre-defined. For example a single plastidic pathway has been assumed in maize root tips (Dieuaide-Noubhani et al., 1995; Alonso et al., 2007), tomato cell cultures (Rontein et al., 2002) and Arabidopsis cell cultures (Williams et al., 2008); whereas complete duplication of the pathway in the cytosol and plastid has been assumed in the analysis of cultured soybean embryos (Sriram et al., 2004; Iyer et al., 2008). In some tissues, for example oilseed rape embryos (Schwender et al., 2003) and soybean embryos (Allen et al., 2009), the modeling task is made easier because the observed labeling patterns imply that intermediates in glycolysis and the PPP are in fast exchange between the cytosol and the plastids, ensuring that duplicate pathways in different physical locations function indistinguishably from a single uncompartmented pathway. However in most of the tissues that have been examined so far, including soybean embryos in some instances (Sriram et al., 2004), the hexose phosphate pools in the cytosol and the plastids are found to be labeled to different extents. Typically this evidence might be based on differences in the labeling patterns of the glucosyl moieties from sucrose and starch (Fig 4.2), and it is then necessary to consider the subcellular compartmentation of glycolysis and the oxidative PPP explicitly.

Inspection of the labeling patterns in products derived specifically from either the cytosolic or plastidic hexose phosphate pools may provide circumstantial evidence for the operation of the full PPP in both compartments (Krook et al., 1999), justifying the construction of a model with duplicate pathways. However, intuitive interpretation of the data is complicated by the multiple metabolic processes that contribute to the labeling of these metabolites (Kruger et al., 2003), and a more robust quantitative approach is needed to compare the results of fitting all the labeling data to models in which the pathways are duplicated to different extents. Here, steady-state MFA was used to compare three different models of the oxidative PPP (Fig 4.1) and the analysis indicates that several aspects of the protocol contribute to its effective implementation. First, it is necessary to work with over-determined models so that the statistical significance of the flux solutions can be properly evaluated. Secondly, for the PPP, it is important to define the carbon transition network in terms of the half-reactions for TA and TK to maximize the chance of being able to discriminate between parallel events in the cytosol and plastid. Thirdly, it is essential to use high quality data from experiments with several differently labeled substrates in order to deduce fluxes with sufficient confidence to allow a comparison of the models (Fig 4.10).

Analysis of three models for the oxidative PPP (Fig 4.1) emphasized the difficulty of determining the flux distribution with sufficient confidence to be able to discriminate between the models. There was little to distinguish a plastid-only model (Fig 4.1A) from the model in which the oxidative steps were duplicated in the cytosol (Fig 4.1B), even though the latter is more plausible on the basis of molecular evidence. The increase in the degrees of freedom in the second model (Fig 4.1B) produced only a marginal improvement in the residuum obtained for the global best fit, and the main observations that support the model are the non-random

allocation of flux between the oxidative steps in the two compartments (Fig 4.5), and the reliability of these flux estimates, as reflected in their confidence intervals (Table 4.3). Adding further degrees of freedom by duplicating the complete oxidative PPP in the two compartments (Fig 4.1C) led to a further, modest, improvement in the fit and a radically different flux distribution through the pathways of carbohydrate oxidation (Table 4.3), but it also reduced the confidence in the flux values, with most of the fluxes showing an increase in standard deviation. Thus this model provides a less reliable flux map and can be rejected on the basis that the fit to the available data is less convincing. Such a conclusion is always subject to the proviso that data might become available that would put further constraints on the assessment of the compartmentation of the pathway (Kruger and Ratcliffe, 2009), but currently the available data imply only a plastidic location for the non-oxidative steps of the PPP in Arabidopsis cells. A similar, but apparently less comprehensive, analysis of the same problem in sunflower embryos (Alonso et al., 2007a) found no improvement in the residuum, which is surprising given the extra degrees of freedom in the models with both cytosolic and plastidic pathways, and it was concluded that the data could be adequately described by model in which the oxidative PPP was confined to the plastid.

The difficulties encountered in analyzing the compartmentation of the oxidative PPP by steady-state MFA highlight some of the practical limitations of the approach. The objective is to capture a realistic metabolic phenotype, but in practice the flux map is a representation of the movement of the chosen label through the system. Resolving the contribution of chemically distinct pathways between two metabolites in the same compartment, or of duplicated steps and pathways in distinct compartments, is not always possible or even attempted. For example the steps from citrate to 2-oxoglutarate are routinely assigned to the

mitochondrion in steady-state MFA even though there may be parallel fluxes in the cytosol. Thus if a function of MFA is to guide metabolic engineering (Kruger and Ratcliffe, 2008; Libourel and Shachar-Hill, 2008) then current flux maps are unlikely to be useful in predicting the result of manipulating a compartment-specific enzyme activity. Similarly, even if MFA provides evidence that a physically compartmented pathway is functionally uncompartmented as a result of fast exchange of intermediates (Schwender et al., 2003; Allen et al., 2009), the resulting metabolic phenotype still conceals the associated compartmental demands for reducing/oxidizing power and energy. It can be concluded that it is essential to push steady-state MFA to its limits to obtain the most informative model for the flux distribution.

4.5.2 Metabolism in a heterotrophic Arabidopsis cell suspension

The network structure used to determine the biochemically preferred flux map (Fig 4.6) differs from the one used in the previous analysis of Arabidopsis cell suspension cultures (Williams et al., 2008) in two main ways. First, the models differ in their treatment of the lower section of glycolysis. In steady-state MFA of plant metabolism there is usually insufficient labeling information to support a model with separate cytosolic and plastidic pools of both glyceraldehyde 3-phosphate and phospho*enol*pyruvate. In the earlier model (Williams et al., 2008) each metabolite was assumed to be in a common uncompartmented pool, implying fast exchange across the plastid envelope; whereas here it was assumed that the two metabolites were at isotopic equilibrium in a three-carbon phosphate ester pool (C3-P; Fig 4.3) within each compartment. This assumption allowed the degree of compartmentation of the pools to be tested by not arbitrarily assuming fast exchange between the cytosol and plastid and it led to a marked alteration in the flux distribution, since all three flux maps showed a significant net flux from the plastid to the cytosol (Fig 4.6) at the level of C3-P and only a limited exchange

flux (Table 4.3). Secondly, the biochemically preferred model that emerged from the comparison of three different models for the oxidative PPP (Fig 4.1) allowed the oxidative steps to operate in both the cytosol and plastids. This compartmentation is supported by estimates of transcript abundance based on publically available microarray data for genes encoding isozymes of the PPP and associated plastid envelope translocators (Masakapalli et al., 2010).

A striking feature of the flux map generated from the biochemically preferred model is that the oxidative steps of the PPP are almost completely confined to the cytosol. Since most biosynthetic processes requiring NADPH are plastidic, this flux distribution implies extensive exchange of reducing power equivalents across the plastid envelope. However, this is unlikely to be problematic since previous studies on maize root tips suggest that the oxidative PPP in the cytosol and plastids can co-operate in providing NADPH for biosynthesis throughout the cell (Averill et al., 1998). Nevertheless, the dominance of the cytosolic pathway in catalyzing flux through the oxidative section of the PPP is unexpected, and it is unclear whether this is a general characteristic of the organization of carbohydrate metabolism in higher plants since compartmentation of the PPP is not included in most flux maps. The only other plant tissues for which information is available are rosy periwinkle hairy root cultures and developing soybean embryos, and in both of these only 25-50% of flux through the oxidative section of the PPP is cytosolic (Sriram et al., 2004, 2007; Iyer et al., 2008). However, in these systems total flux through the oxidative PPP is considerably higher than that observed in Arabidopsis cells, and when expressed as a proportion of the rate of sugar consumption the cytosolic fluxes in all three systems are similar (rosy periwinkle, 42.6%; soybean, 47.4%; Arabidopsis, 37.2%) implying that variation in compartmentation of PPP flux is due to fluctuations in plastidic

activities rather than to a decrease in the significance of the cytosolic flux. Moreover, irrespective of the exact proportion of the oxidative PPP flux occurring in the cytosol, two recent studies demonstrate the importance of the cytosolic activity. First, disrupting the expression of the two genes encoding cytosolic isozymes of glucose 6-phosphate dehydrogenase (*G6PD5* and *G6PD6*) in Arabidopsis through insertional inactivation alters seed lipid content (Wakao et al., 2008). Secondly, supplementation or replacement of cytosolic glucose 6-phosphate dehydrogenase with an alternative isoform improves resistance of tobacco leaves to *Phytophthora infestans*, and this is attributed to changes in the provision of NADPH and stimulation of the hypersentitive defense response (Scharte et al., 2009). Both observations highlight a requirement for the oxidative section of the PPP in the cytosol to support normal metabolic activity.

4.5.3 Energy balance in plant metabolic networks

The metabolic activity of the Arabidopsis cell culture appears to be broadly representative of a wide range of heterotrophic plant systems and one measure of this can be found in its efficiency in converting sugar to biomass components. The proportion of the [U-^{14}C]glucose metabolised to macromolecules and soluble compounds (Table 4.1) indicates a carbon use efficiency of 60%, in good agreement with the 65% calculated from a comparison of substrate uptake and estimated CO_2 output in the previous study of Arabidopsis cells grown under equivalent conditions (Williams et al., 2008). These estimates of carbon conversion efficiency (CCE) are within the range obtained for other heterotrophic plant systems, including a rosy periwinkle hairy root culture (24%; Sriram et al., 2007), maize root tips (42-47%; Alonso et al., 2007), developing sunflower embryos (50%; Alonso et al., 2007) and a tomato cell culture (52-68%; Rontein et al., 2002). Although considerably higher values have been reported for

developing embryos of oilseed rape (85-95%) and soybean (82-83%), in these systems there is appreciable reassimilation of respired CO_2 through Rubisco, and light makes significant contributions to the provision of ATP and/or reductant required for biosynthesis (Allen et al., 2009; Goffman et al., 2005). In contrast, the energy and reducing power required for biosynthesis to support growth in heterotrophic Arabidopsis cells must be generated exclusively by oxidation of glucose, the sole respiratory substrate, through the central pathways of carbon metabolism.

In contrast to the situation in developing soybean embryos (Iyer et al., 2008), inspection of the fluxes determined for the Arabidopsis cells establishes that the oxidative PPP is insufficient to meet the NADPH requirements for biosynthesis predicted by the network (Table 4.5). This observation is the result of two assumptions: (i) the assimilation of equal amounts of nitrate and ammonium (Table 4.5); and (ii) the NADPH-dependence of glutamate synthesis (Table 4.6). In this situation reductive biosynthesis will require other $NADP^+$-dependent processes, and in the network considered here the demand could be met by the operation of $NADP^+$-isocitrate dehydrogenase (Table 4.5).

In contrast, the production of NADH is more than enough for its utilization in reductive processes, including the demands for ATP synthesis. Indeed, if it is assumed that excess NADH (and $FADH_2$) are re-oxidised through the mitochondrial electron transport chain to produce ATP through oxidative phosphorylation (at stoichiometries of 2.3–2.5 and 1.4–1.5 ATP per NADH and $FADH_2$, respectively; Hinkle, 2005), then the maximum yield of ATP generated by the metabolic network is over seven times that required to support the flux to biosynthetic products. The actual yield of ATP will depend on the extent of engagement of the alternative oxidase and the activity of mitochondrial uncoupling proteins, but the comparison

suggests the metabolic network has a considerable over-capacity for ATP production. Of course, these calculations ignore the energy requirements of other metabolic processes such as substrate cycling, turnover of proteins and other macromolecules, as well as synthesis and turnover of RNA that are not included in the present network. Moreover, ATP is used in other cellular processes such as substrate uptake, transport of metabolites across intracellular membranes, and maintenance of trans-membrane ion and electrical gradients. Such activities can account for over half of the cellular ATP consumption in microorganisms (Nielsen et al., 2003), and the findings presented here suggest that an even greater proportion of the potential chemical energy released during metabolism in actively growing plant cells could be needed to meet these requirements. Recently, analysis of energy demands in a genome-scale stoichiometric model of Arabidopsis has led to a similar conclusion that only a relatively small proportion of the potential ATP generated by the cell is used for biosynthesis (Poolman et al., 2009).

This conclusion has implications for determination of flux through plant metabolic networks by flux balance analysis (FBA). This approach depends on the selection of an appropriate objective function to obtain an optimum solution within the potential flux space defined by the stoichiometry of the biochemical reaction network, plus other constraints, such as the metabolic outputs of the system (Kauffman et al., 2003). Typically, the optimisation criterion in microorganisms is maximisation of yield of biomass, ATP or a product of interest, or alternatively maximisation of flux, i.e. maximisation of the rate of production of one or more components required for growth (Schuster et al., 2008). A recent study on developing endosperm of barley used a refinement of this approach in which the objective function was maximisation of growth per unit flux – this is a combination of maximization of biomass yield,

and minimisation of overall flux (Grafahrend-Belau et al., 2009). The logic behind this optimisation criterion is that cellular metabolism has evolved to fulfil its functions with the most economic use of available resources (Holzhütter, 2004), but application of this criterion requires that the major energy-consuming processes have been adequately incorporated into the metabolic model. The demonstration that a large proportion of the potential biochemical energy released through respiration in heterotrophic plant cells is dissipated, or used by processes other than those directly associated with the synthesis of macromolecules and soluble components that accumulate during growth, means that ATP expenditure for such purposes must be accurately assessed if FBA is to generate meaningful flux maps of the network of central carbon metabolism in plants.

4.6 Conclusion

Steady-state MFA of multiple labeling experiments in an Arabidopsis cell suspension shows that the experimental observations can be explained by any one of three different models of the subcellular compartmentation of the oxidative PPP. Other biochemical evidence favors the model in which the oxidative steps are duplicated in the cytosol and plastids, and this leads to the conclusion that the bulk of the flux through the oxidative steps is cytosolic. This illustrates the point that interesting features of the metabolic phenotype can be missed with an inappropriate model and emphasizes the importance of basing the model on all the available biochemical evidence.

Table 4.1 *Metabolism of [U-^{14}C]glucose by cell suspension cultures of Arabidopsis*

Samples (5 ml, 0.618 ± 0.017 g fresh weight) from 6 d old cultures grown in glucose at natural abundance ^{13}C were supplemented with 185 kBq [U-^{14}C]glucose (yielding a specific activity of 308 ± 15 MBq mol^{-1}) and incubated for 24 h prior to extraction in methanol and subsequent metabolic fractionation. Total ^{14}C uptake is the sum of radioactivity in CO_2, ethanol and the chloroform soluble, aqueous methanol soluble and methanol insoluble fractions in each cell culture. Each value is the mean ± SE of the distribution of radioactivity in four cell cultures.

Metabolic fraction	Radioactivity in specified fraction (dpm)		
CO_2	1,076,468	±	35,772
Chloroform soluble	63,439	±	5,420
Neutral lipids	41,242	±	2,217
Free fatty acids	1,083	±	225
Neutral P-lipids	5,046	±	722
Acidic P-lipids	21,597	±	2,921
Methanol soluble	1,021,858	±	26,185
Organic acids/P-esters	109,533	±	553
Amino acids	386,305	±	13,201
Asp/Glu	31,938	±	2,563
Others	334,737	±	20,786
Sugars	389,986	±	17,377
Sucrose	251,596	±	16,562
Glucose	58,680	±	5,384
Fructose	37,060	±	2,298
Others	3,943	±	623
Methanol insoluble	591,272	±	53,215
Starch	72,424	±	11,971
Protein	116,647	±	10,699
Cell wall	103,763	±	13,417
Residue	77,010	±	21,924
Ethanol	18,500	±	2,115
Total ^{14}C uptake	2,771,536	±	67,703
Total ^{14}C recovered	2,361,265	±	25,419
Recovery (%)	84.2	±	2.1

Table 4.2 Summary of biosynthetic output constraints applied in flux determinations using 13C-FLUX

Metabolite (as defined in model)	Flux name (as defined in model)	Molar output flux (relative to *upt* = 1)
Sucrose – glucosyl unit	*suc(g)OUT*	0.0760 ± 0.0039
Sucrose – fructosyl unit	*suc(f)OUT*	0.0760 ± 0.0039
Starch – glycosyl unit	*starchOUT*	0.0422 ± 0.0070
Cell wall – hexosyl unit	*cellwall(h)OUT*	0.0790 ± 0.0112
Cell wall – pentosyl unit	*cellwall(p)OUT*	0.0316 ± 0.0045
Lipid – acyl component	*lipidOUT*	0.0657 ± 0.0045
Lipid – glycerol unit	*glycerolOUT*	0.0024 ± 0.0002
Ethanol	*ethanolOUT*	0.0202 ± 0.0023
Citrate	*citOUT*	0.0253 ± 0.0001
Malate	*malOUT*	0.0207 ± 0.0001
Succinate	*succOUT*	0.0104 ± 0.0001
Alanine	*alaOUT*	0.0359 ± 0.0014
Arginine	*argOUT*	0.0114 ± 0.0004
Aspartate/asparagine	*aspOUT*	0.0187 ± 0.0013
Cysteine	*cysOUT*	0.0055 ± 0.0002
Glutamate/glutamine	*gluOUT*	0.0246 ± 0.0017
Glycine	*glyOUT*	0.0532 ± 0.0021
Histidine	*hisOUT*	0.0051 ± 0.0002
Isoleucine	*ileOUT*	0.0132 ± 0.0005
Leucine	*leuOUT*	0.0223 ± 0.0009
Lysine	*lysOUT*	0.0164 ± 0.0006
Methionine	*metOUT*	0.0008 ± 0.00004
Phenylalanine + tyrosine	*pheOUT*	0.0133 ± 0.0004
Proline	*proOUT*	0.0162 ± 0.0006
Threonine	*thrOUT*	0.0179 ± 0.0007
Tryptophan (2C unit)	*trp(a)OUT*	0.0065 ± 0.0002
Tryptophan (2C unit)	*trp(b)OUT*	0.0065 ± 0.0002
Tryptophan (3C unit) + serine	*trp(c)OUT*	0.0307 ± 0.0010
Tryptophan (4C unit)	*trp(d)OUT*	0.0065 ± 0.0002
Valine	*valOUT*	0.0239 ± 0.0009

Table 4.3. *Influence of metabolic network structure on determination of flux estimates*

Fluxes were determined by fitting isotopomer abundances and biosynthetic output estimates to a family of models differing in the extent of pentose phosphate pathway (PPP) compartmentation (Fig 4.1). Each model contained a complete plastidic PPP pathway. Free fluxes varied during the fitting procedure are indicated in ***bold***. Net fluxes and normalized hyperbolic transformed exchange fluxes are the global best fit estimate ± SD as determined by EstimateStat and are calculated relative to the rate of glucose uptake (set to unity). A normalized exchange flux for which SD takes the estimated value outside the permissible [0,1] range is considered indeterminate (Ind).

Flux name	Flux (relative to glucose uptake)		
	No cytosolic PPP	Cytosolic oxidative PPP steps only	Complete cytosolic PPP
Net flux			
chex1	0.317 ± 0.062	0.191 ± 0.083	0.665 ± 0.074
chex2	0.241 ± 0.062	0.115 ± 0.083	0.365 ± 0.187
chex3	0.778 ± 0.016	0.786 ± 0.021	0.894 ± 0.040
phex1	0.049 ± 0.067	0.209 ± 0.101	-0.160 ± 0.090
phex2	0.305 ± 0.063	0.442 ± 0.088	0.228 ± 0.187
phex3	0.072 ± 0.016	0.075 ± 0.020	0.000 ± 0.040
cppp1	-	0.372 ± 0.028	0.266 ± 0.019
cppp2a	-	-	-0.224 ± 0.200
cppp2b	-	-	-0.112 ± 0.100
cppp2c	-	-	-0.112 ± 0.100
cppp3a	-	-	-0.112 ± 0.100
cppp3b	-	-	-0.112 ± 0.100
pppp1	0.405 ± 0.018	0.000 ± 0.039	0.000 ± 0.027
pppp2a	0.138 ± 0.006	0.127 ± 0.009	0.204 ± 0.102
pppp2b	0.118 ± 0.006	0.107 ± 0.009	0.184 ± 0.102
pppp2c	0.256 ± 0.012	0.234 ± 0.018	0.388 ± 0.204
pppp3a	0.138 ± 0.006	0.127 ± 0.009	0.204 ± 0.102
pppp3b	0.138 ± 0.006	0.127 ± 0.009	0.204 ± 0.102
tca1	0.650 ± 0.006	0.661 ± 0.009	0.696 ± 0.009
tca2	0.650 ± 0.006	0.661 ± 0.009	0.696 ± 0.009
tca3	0.624 ± 0.006	0.635 ± 0.009	0.670 ± 0.009
tca4	0.572 ± 0.006	0.583 ± 0.009	0.618 ± 0.009
tca5	0.573 ± 0.006	0.584 ± 0.009	0.619 ± 0.009
ana1	0.241 ± 0.004	0.241 ± 0.004	0.243 ± 0.005
ana2	0.039 ± 0.003	0.040 ± 0.004	0.039 ± 0.004
ana3	0.026 ± 0.001	0.026 ± 0.001	0.029 ± 0.003

cmex	0.610 ± 0.007	0.621 ± 0.009	0.657 ± 0.009
cpex	0.112 ± 0.015	0.109 ± 0.019	0.181 ± 0.037
gpt	0.496 ± 0.062	0.251 ± 0.088	-0.118 ± 0.082
tpt/ppt	0.540 ± 0.124	0.801 ± 0.171	0.523 ± 0.471
xpt	-	-	0.603 ± 0.309
Normalized exchange flux			
chex1	0.990 ± 0.102	0.990 ± 0.108	0.423 ± 0.066
chex2	0.682 ± 0.055	0.667 ± 0.053	0.803 ± 0.073
phex1	0.476 ± 0.040	0.475 ± 0.054	0.148 ± 0.103
cppp2a	-	-	0.990 ± 0.276
cppp2b	-	-	0.320 ± 0.082
cppp2c	-	-	Ind
cppp3a	-	-	Ind
cppp3b	-	-	0.000 ± 0.194
pppp2a	Ind	Ind	Ind
pppp2b	0.990 ± 0.141	0.990 ± 0.166	0.990 ± 0.588
pppp2c	0.990 ± 0.120	0.990 ± 0.115	0.990 ± 0.502
pppp3a	Ind	Ind	Ind
pppp3b	0.268 ± 0.039	0.299 ± 0.059	0.252 ± 0.123
tca5	0.719 ± 0.043	0.719 ± 0.043	0.730 ± 0.044
gpt	0.970 ± 0.051	0.695 ± 0.067	0.462 ± 0.096
tpt/ppt	0.227 ± 0.140	0.441 ± 0.115	0.000 ± 1.424
xpt	-	-	0.309 ± 0.362

Table 4.4 Metabolic fluxes in heterotrophic Arabidopsis cells

Fluxes were determined by fitting isotopomer abundances and biosynthetic output estimates to the model shown in Fig 4.3. Internal network fluxes were initially calculated relative to the rate of glucose uptake (set to unity) and then rescaled using the rate of glucose uptake determined from the metabolism of [U-^{14}C]glucose (Table 4.1). Net fluxes are the global best fit estimate ± SD as determined by EstimateStat. Exchange fluxes are the global best fit estimates (with lower and upper asymmetric 68% confidence intervals) obtained from the optimum normalized [0,1] exchange flux ± SD determined after hyperbolic transformation. Where the bounds of the normalized exchange flux lie outside the permissible [0,1] range, an exchange flux of 0 (≤ 0) or infinity (≥ 1) is given. Processes for which no exchange flux is given are considered irreversible and were constrained to zero in the fitting procedure.

Flux name	Metabolic process	Net flux	Exchange flux
		(μmol h^{-1} g fresh $weight^{-1}$)	
Metabolic inputs			
upt	Glucose uptake	10.11 ± 0.346	-
Hexose/triose phosphate metabolism			
chex1	Cytosolic phosphoglucose isomerase	1.93 ± 0.836	1000.89 (75.26, ∞)
chex2	Cytosolic Fru-6-P → triose-P/PEP interconversion	1.16 ± 0.836	20.28 (16.07, 26.08)
chex3	Cytosolic pyruvate kinase	7.95 ± 0.210	-
phex1	Plastidic phosphoglucose isomerase	2.11 ± 1.02	9.15 (7.34, 11.38)
phex2	Plastidic Fru-6-P → triose-P/PEP interconversion	4.47 ± 0.891	-
phex3	Plastidic pyruvate kinase	0.76 ± 0.204	-
cppp1	Cytosolic pentose phosphate pathway - oxidative steps	3.76 ± 0.280	-
pppp1	Plastidic pentose phosphate pathway - oxidative steps	0.00 ± 0.389	-
pppp2a	Plastidic transketolase - C7/C5 conversion	1.28 ± 0.089	9.78 (0 , ∞)
pppp2b	Plastidic transketolase - C6/C4 conversion	1.08 ± 0.089	1000.89 (47.17, ∞)
pppp2c	Plastidic transketolase - C5/C3 conversion	2.36 ± 0.178	1000.89 (70.46, ∞)
pppp3a	Plastidic transaldolase - C7/C4 conversion	1.28 ± 0.089	13.16 (0, ∞)
pppp3b	Plastidic transaldolase - C6/C3 conversion	1.28 ± 0.089	4.31 (3.19, 5.64)

Tricarboxylic acid cycle			
tca1	Mitochondrial pyruvate dehydrogenase	6.68 ± 0.089	-
tca2	Citrate synthase	6.68 ± 0.089	-
tca3	Aconitase/isocitrate dehydrogenase	6.42 ± 0.089	-
tca4	2-oxoglutarate → fumarate convertion	5.89 ± 0.089	-
tca5a+b	Fumarase	5.90 ± 0.089	51.86 (42.28, 64.90)
Anaplerotic fluxes			
ana1	PEP carboxylase	2.44 ± 0.040	-
ana2	Mitochodrial malic enzyme	0.41 ± 0.036	-
ana3	Plastidic malic enzyme	0.26 ± 0.015	-
Transporters/exchanges			
cmex	Pyruvate uptake by mitochondria	6.27 ± 0.089	-
cpex	Pyruvate uptake by plastids	1.11 ± 0.195	-
gpt	Plastidic Glc-6-P/Pi translocator	2.54 ± 0.888	23.01 (17.03, 32.36)
tpt/ppt	Plastidic triose-P/Pi & PEP/Pi translocators	8.09 ± 1.731	7.97 (4.90, 12.62)
Biosynthetic outputs			
CO_2OUT	Respired CO_2	23.47 ± 0.781	-
suc(g)OUT	Sucrose synthesis - glucosyl moiety	0.77 ± 0.039	-
suc(f)OUT	Sucrose synthesis - fructosyl moiety	0.77 ± 0.039	-
starchOUT	Starch production	0.43 ± 0.071	-
cellwall(h)OUT	Cell wall (hexosyl units)	0.80 ± 0.113	-
cellwall(p)OUT	Cell wall (pentosyl units)	0.32 ± 0.045	-
lipidOUT	Acyl units derived from acetylCoA	0.66 ± 0.045	-
glycerolOUT	Glycerol unit of lipids	0.02 ± 0.002	-
ethanolOUT	Ethanol accumulation in medium	0.20 ± 0.023	-
citOUT	Accumulation of citrate	0.26 ± 0.001	-

malOUT	Accumulation of malate	0.21 ± 0.001	-
succOUT	Accumulation of succinate	0.11 ± 0.001	-
p5pOUT	Synthesis of His and Trp from pentose-P	0.12 ± 0.003	-
pc3pOUT	Synthesis of Cys, Gly, Phe, Ser, Trp and Tyr from plastidic triose-P/PEP	1.17 ± 0.025	-
pe4pOUT	Synthesis of Phe, Trp and Tyr from plastidic Ery-4-P	0.20 ± 0.004	-
ppyrOUT	Synthesis of Ile, Leu, Lys and Val from plastidic pyruvate	1.23 ± 0.027	-
cpyrOUT	Synthesis of Ala from cytosolic pyruvate	0.36 ± 0.014	-
pacOUT	Synthesis of Leu from plastidic acetylCoA	0.23 ± 0.009	-
oaaOUT	Synthesis of Asp, Asn, Met and Thr from OAA	0.38 ± 0.015	-
αkgOUT	Synthesis of Arg, Glu, Gln and Pro from 2-oxoglutarate	0.53 ± 0.019	-

Table 4.5. *Coenzyme requirements for the biosynthetic outputs from the central metabolic network in heterotrophic Arabidopsis cells.*

The coenzyme requirements for biosynthesis were determined from the fluxes in Table 4.2 using the reaction stoichiometries in Table 4.6, and are presented as mean ± SE based on the estimates of label partitioning following metabolism of [U-^{14}C]glucose in Table 4.1. The coenzyme requirements for lipid assume synthesis of triacylglycerol containing only oleoyl (C18:0) side chains. Costs for amino acid production are based on the assimilation of equal amounts of NO_3^- and NH_4^+. The ATP requirement for protein production is calculated from the cost of synthesising proteins from the component amino acids, and is equivalent to 4 ATP per amino acid residue. Accumulating organic acids are withdrawn directly from intermediate pools in the network and incur no additional cost. Generation of coenzymes by the central metabolic network is determined from the *net* fluxes of the internal interconversions (i.e. excluding output steps) of the model defined in Figure 4.3 using conventional reaction stoichiometries and ignoring the possible contribution of inorganic pyrophosphate as a phosphoryl donor. NADH and NADPH yield in italics assume conversion of isocitrate to 2-oxoglutarate by $NADP^+$-dependent isocitrate dehydrogenase. ATP generation is attributed to substrate level phosphorylation, and $FADH_2$ production is assumed to be equivalent to 0.6x NADH. Values for coenzyme generation are the mean ± SE of estimates obtained from the best-fit flux solutions of 501 Monte Carlo simulations; values in parentheses are determined from the global best-fit flux solution of the network in which the oxidative steps of the PPP are located in both the cytosol and plastids as detailed in Table 4.4.

Product	Coenzyme requirement (molar flux relative to glucose uptake)		
	ATP	NADH	NADPH
Sucrose	0.076 ± 0.0039	-	-
Starch	0.042 ± 0.0070	-	-
Cell wall	0.079 ± 0.0112	-0.039 ± 0.0045	-
Lipid	0.066 ± 0.0045	0.068 ± 0.0045	0.066 ± 0.0045
Ethanol	-	0.020 ± 0.0023	-
Amino acids	0.726 ± 0.0090	0.144 ± 0.0028	1.207 ± 0.0141
Protein	0.268 ± 0.0285	-	-
Total requirement	1.257 ± 0.0332	0.193 ± 0.0073	1.273 ± 0.0148
Generated by the central metabolic network	0.999 ± 0.0011 (1.001)	3.964 ± 0.0024 *3.330 ± 0.0021* (4.055) *(3.420)*	0.771 ± 0.0022 *1.407 ± 0.0018* (0.770) *(1.405)*

Table 4.6 Coenzyme stoichiometries for generation of biosynthetic products from intermediates of the central metabolic network

Product	Precursor	Coenzyme requirement (mol mol^{-1} precursor)						Comment
		NH_4^+ as N source			NO_3^- as N source			
		ATP	NADH	NADPH	ATP	NADH	NADPH	
Starch	Hexose-P	1	0	0	1	0	0	
Cell wall hexosyl unit	Hexose-P	1	0	0	1	0	0	Ignores generation of NADH in uronic acid production
Cell wall pentosyl unit	Hexose-P	1	-2	0	1	-2	0	
Sucrose	2(Hexose-P)	1	0	0	1	0	0	
Lipid acyl chain	AcCoA	1	1	1	1	1	1	Cost for acyl chain elongation, assumes enoyl-ACP reductase uses NADH
Lipid glycerol unit	Triose-P	0	1	0	0	1	0	
Ethanol	Pyr	0	1	0	0	1	0	
Ala	Pyr	1	0	1	1	1	4	
Arg	αKG + 2(AcCoA)	9	-1	4	9	2	13	Assumes 2ATP needed for regeneration of AcCoA from acetate, and includes costs of amide group production

Asx	OAA	4	0	1	4	2	7	Assumes all present as Asn and includes cost of amide group production
Cys	3PGA + AcCoA + SO_4^{2-}	6	-1	5	6	0	8	Includes cost of sulphate reduction
Glx	αKG	2	0	1	2	2	7	Assumes all present as Gln and includes cost of amide group production
Gly	3PGA	1	0	1	1	1	4	
His	Rib-5-P + ATP	7	-3	2	7	-1	10	Includes costs for production of R5PP and ring N, and regeneration of ATP from aminoimidazole carboxamide ribonucleotide
Ile	OAA + Pyr	4	0	5	4	2	11	Includes cost of reassimilating NH_3
Leu	2(Pyr) + AcCoA	1	-1	2	1	0	5	
Lys	OAA + Pyr	4	0	4	4	2	10	Includes cost of ε-amino group formation and allowance for ATP not generated in conversion of SuccCoA to Succ
Met	OAA + SO_4^{2-}	10	-1	8	10	0	11	Includes costs of sulphate reduction, production of Ser and its conversion to Pyr
Phe	2(PEP) + E4P	2	0	2	2	1	5	
Pro	αKG	2	0	3	2	1	6	

Ser	3PGA	1	0	1	1	1	4	
Thr	OAA	3	0	3	3	1	6	
Trp	2(E4P) + PEP + R5PP + 3PGA	5	0	2	5	1	5	Includes cost for production of Ser, R5PP and ring N, and allowance for ATP not generated in formation of Pyr from 3PGA
Tyr	2(PEP) + E4P	2	-1	2	2	0	5	
Val	2(Pyr)	1	0	1	1	1	4	

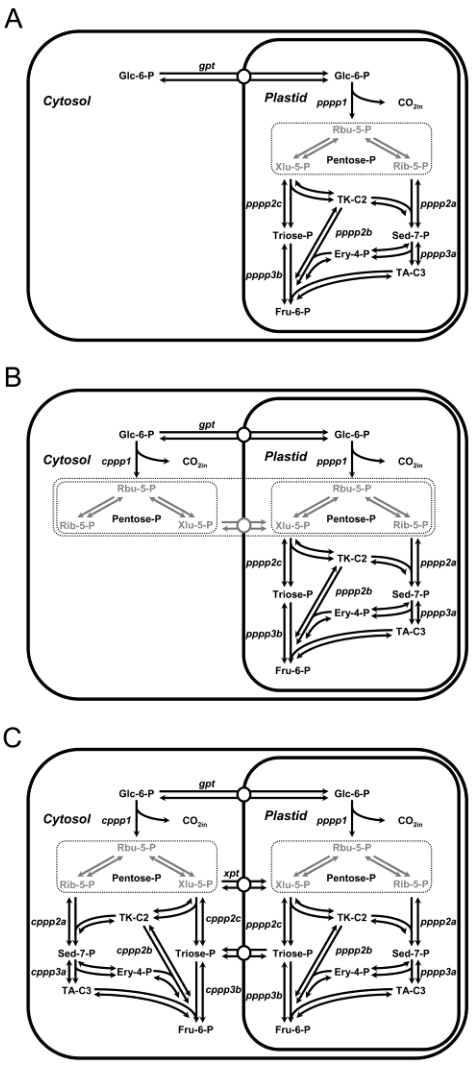


Figure 4.1. Alternative metabolic models for the subcellular compartmentation of the PPP. A, Oxidative and non-oxidative reactions of the PPP confined to the plastid; B, oxidative steps of the PPP in both the cytosol and plastid, with the non-oxidative steps in the plastid; and C, the complete pathway in both compartments. Net and exchange fluxes are shown with uni- and bi-directional arrows respectively, and the direction of a unidirectional arrow corresponds to a positive flux in the output of the model. Flux names are given in italics and the pentose phosphates are grouped into a single indistinguishable pool in all three models. Abbreviations: Ery-4-P, erythrose 4-phosphate; Fru-6-P, fructose 6-phosphate; Glu-6-P, glucose 6-phosphate; Rib-5-P, ribose 5-phosphate; Rbu-5-P, ribulose 5-phosphate; Sed-7-P, sedoheptulose 7-phosphate; TA-C3, the three-carbon substituted enzyme intermediate in the reaction catalyzed by transaldolase; TK-C2, the two-carbon substituted enzyme intermediate in the reaction catalyzed by transketolase; Xlu-5-P, xylulose 5-phosphate.

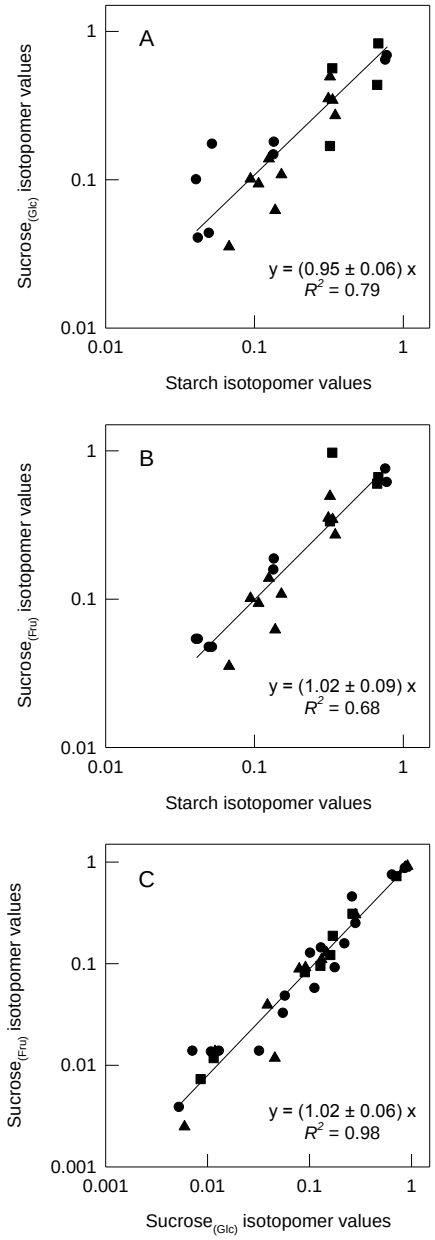


Figure 4.2 Comparison of relative fractional abundance of isotopomers in carbohydrates derived from cytosolic and plastidic hexose phosphate pools. Equivalent isotopomer groups of glycosyl units of starch and both glucosyl and fructosyl moieties of sucrose were compared after labelling cells with [1-^{13}C]glucose (●), [2-^{13}C]glucose (■) or [U-^{13}C]glucose (▲). The comparisons were: A, starch and the glucosyl moiety of sucrose; B, starch and the fructosyl moiety of sucrose; C, the glucosyl and fructosyl moieties of sucrose. For each hexosyl unit, the abundance of a specific isotopomer was normalized relative to the sum of signal intensities of the isotopomer groups used in the comparison. The line of best fit was determined by linear regression analysis and the gradient ± SE is reported along with the coefficient of determination (R^2).

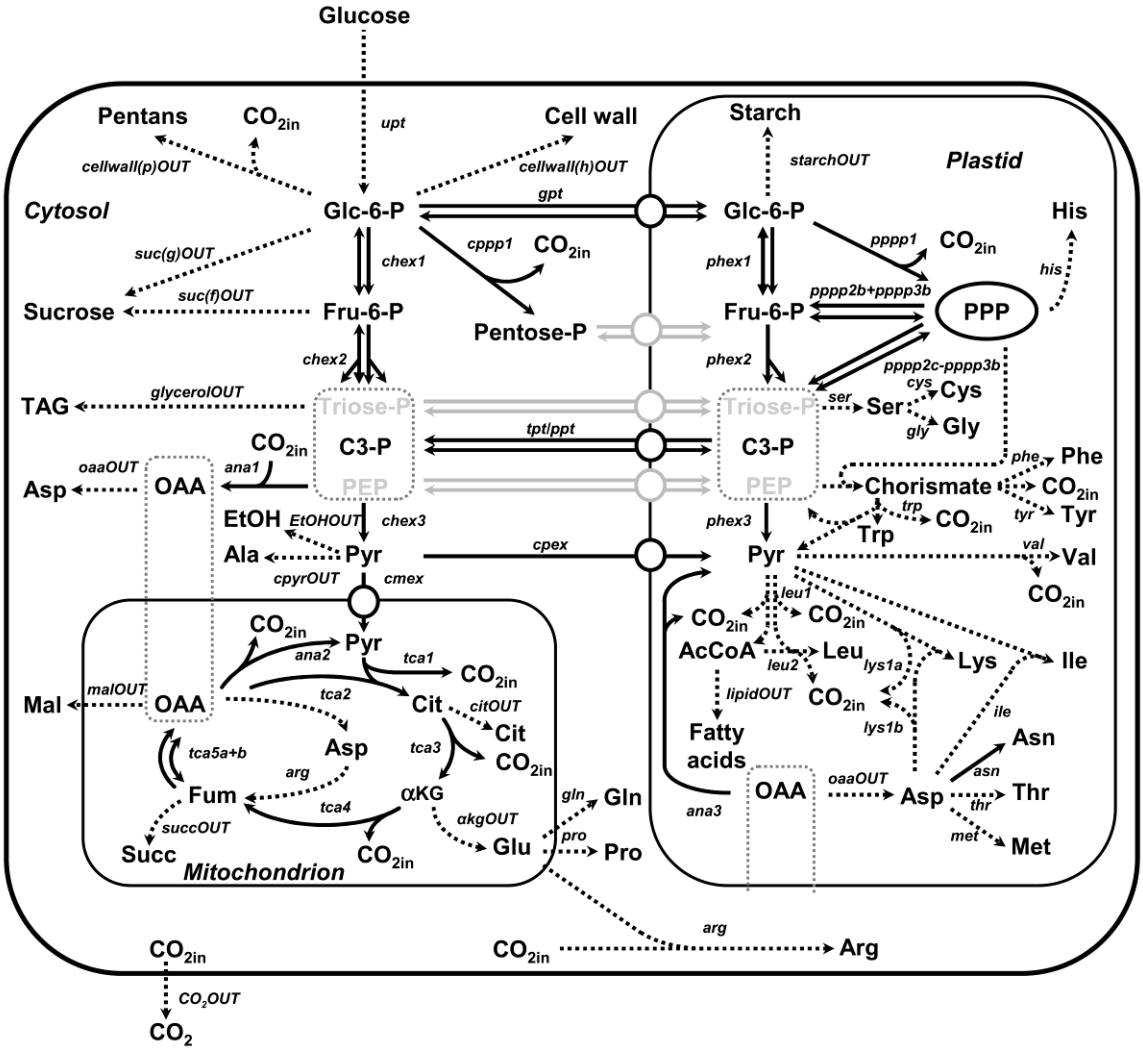


Figure 4.3 Metabolic model of central carbon metabolism with the oxidative steps of the plastidic oxidative PPP duplicated in the cytosol. Output fluxes used to constrain the model are shown with dotted black arrows; net and exchange fluxes deduced from the ^{13}C-label redistribution are shown with solid black uni- and bi-diectional arrows, respectively. The direction of a unidirectional arrow corresponds to a positive flux in the output of the model and the dotted grey boxes indicate that the enclosed metabolites were considered as a single pool. Standard abbreviations are used for the amino acids. Other abbreviations: C3-P, three-carbon phosphate ester pool; cit, citrate; Fru-6-P, fructose 6-phosphate; fum, fumarate; Glc-6-P, glucose 6-phosphate; αKG, 2-oxoglutarate; OAA, oxaloacetate.

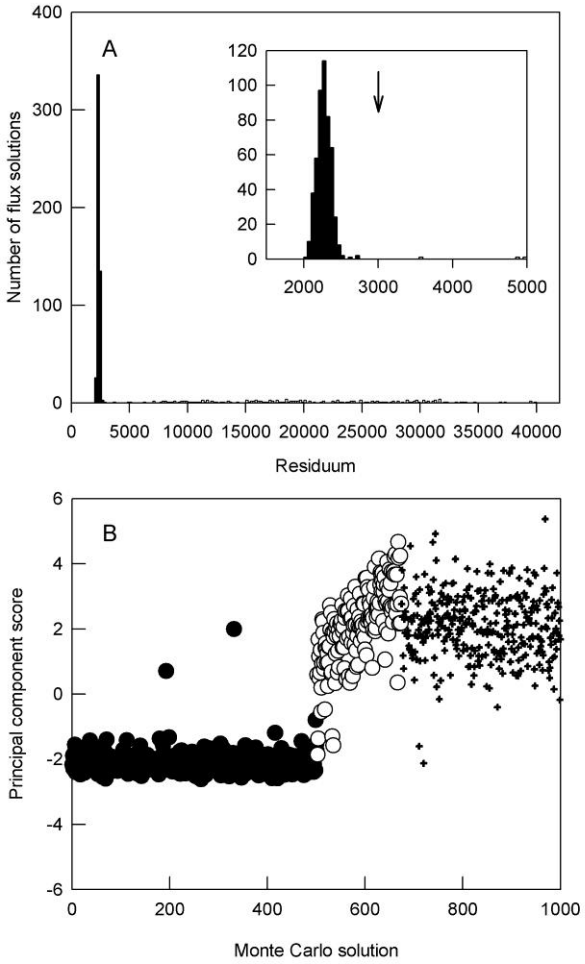


Figure 4.4 Establishing the threshold for the best fit flux solutions used to define the subset of Monte Carlo flux space for calculating the global best fit. A, Histogram showing the frequency distribution for the residua of the feasible solutions. The inset shows an expansion of the region covering the lowest residua with an arrow to indicate the threshold for inclusion in the subsequent analysis. B, Distribution of scores for the first component following principal component analysis of the flux solutions from 1000 Monte Carlo simulations. Feasible flux solutions with residua <3000 (●) form a distinct cluster that is well resolved from those with residua >3000 (○). The latter have a more scattered distribution which is similar to that obtained from random combinations of fluxes that produce infeasible flux maps (✚).

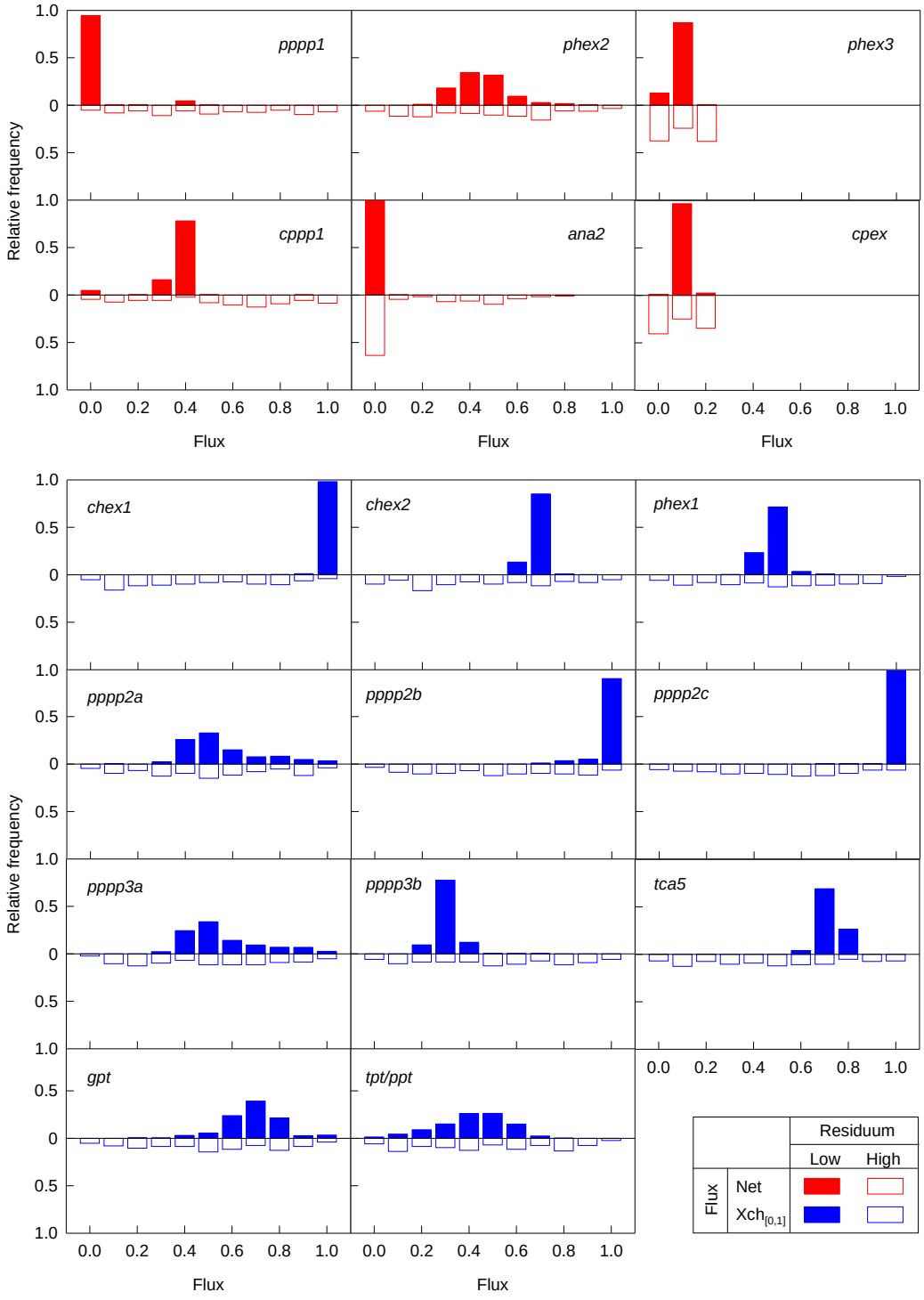


Figure 4.5. Distribution of estimates for free fluxes in the metabolic network from Monte Carlo simulations. $Xch_{[0,1]}$ is the normalized hyperbolic transformation of the exchange flux.

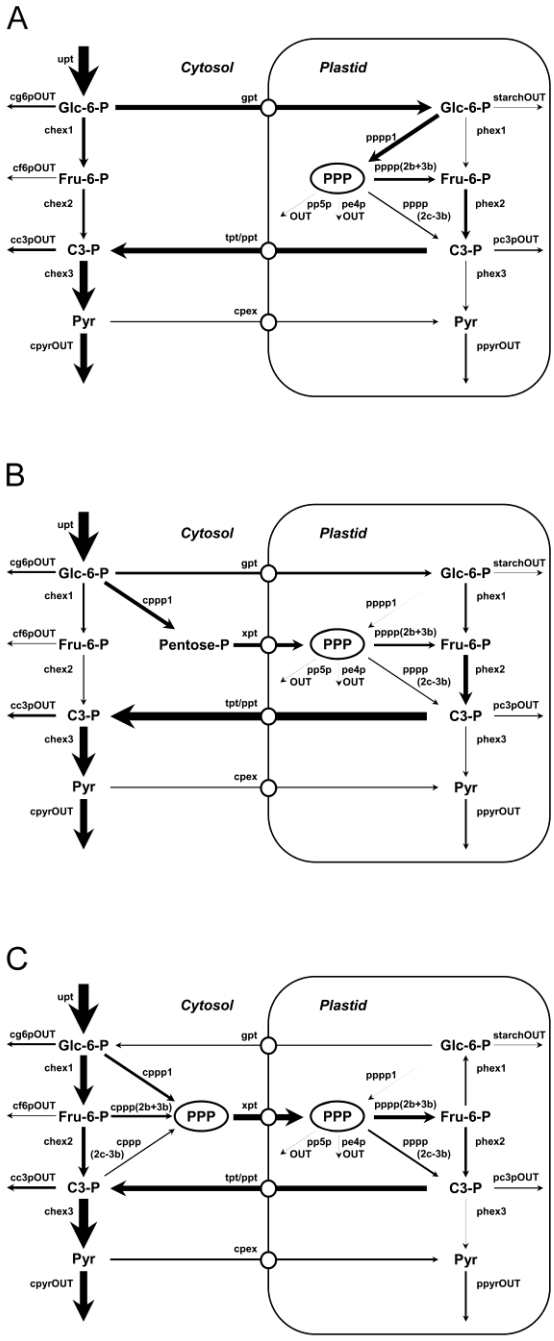


Figure 4.6 Flux maps corresponding to three models for the subcellular compartmentation of the oxidative PPP. A, Oxidative and non-oxidative reactions of the PPP confined to the plastid; B, oxidative steps of the PPP in both the cytosol and plastid, with the non-oxidative steps in the plastid; and C, the complete pathway in both compartments. The width of the arrows is proportional to the net molar flux for each step as listed in Table 4.3.

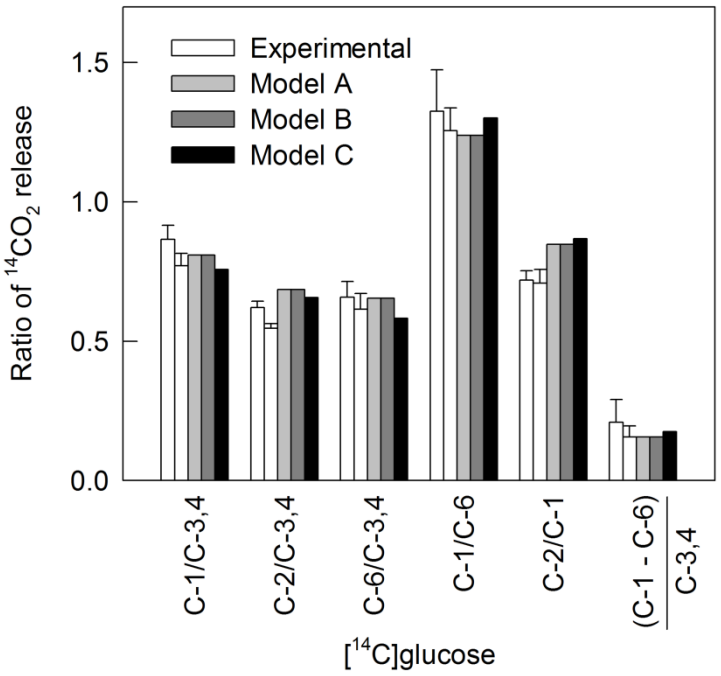


Figure 4.7 Validation of Arabidopsis flux map by radiorespirometric analysis. The ratios of the specific activity of $^{14}CO_2$ released from cell cultures supplied with [^{14}C]glucose specifically labeled in the carbon positions indicated was compared with those predicted from the flux maps. The latter were derived from the specific abundance of label in CO_2 determined by simulations in CumoNet using different glucose isotopomers as the input substrate. Experimental data are shown as the mean ± SD of ratios from four replicate measurements for each of two independent cell cultures (white bars). The predicted ratios were determined from the global best-fit solutions (see Table 4.3, Table 4.4) for the networks in which the PPP is confined to the plastid (Model A, light gray), the oxidative steps of the PPP are located in both the cytosol and plastids (Model B, dark gray), or the entire pathway is present in both compartments (Model C, black). There is no significant difference in the experimentally determined ratios for the two separately grown cultures as assessed by repeat measures ANOVA, and all the predicted ratios are within the 95% prediction intervals of the ratios determined experimentally.

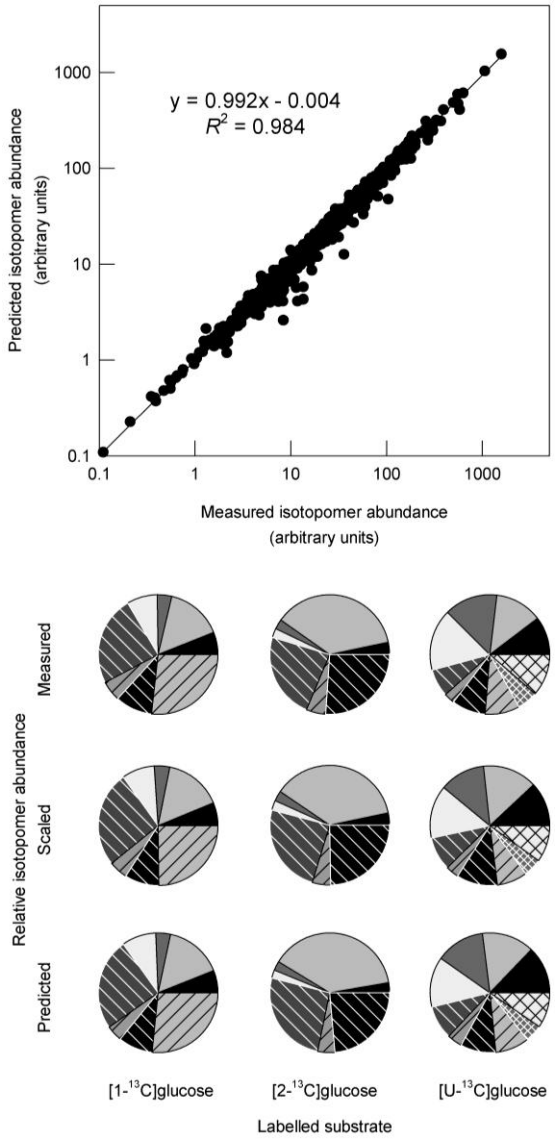


Figure 4.8 Assessment of goodness-of-fit of isotopomer data to the global best fit solution of the metabolic network defined in Figure 4.3. Data from cell cultures labeled separately with [1-^{13}C]glucose, [2-^{13}C]glucose and [U-$^{13}C_6$]glucose were used in these evaluations. A. Comparison of predicted and measured isotopomer abundances for all metabolites analyzed. B. Comparison of the predicted relative fractional abundance of isotopomers of aspartate with measured values, before and after scaling of subgroups of ^{13}C peaks with multiplet structure. Shading is conserved for individual isotopomers in the pie charts for a given labeled substrate.

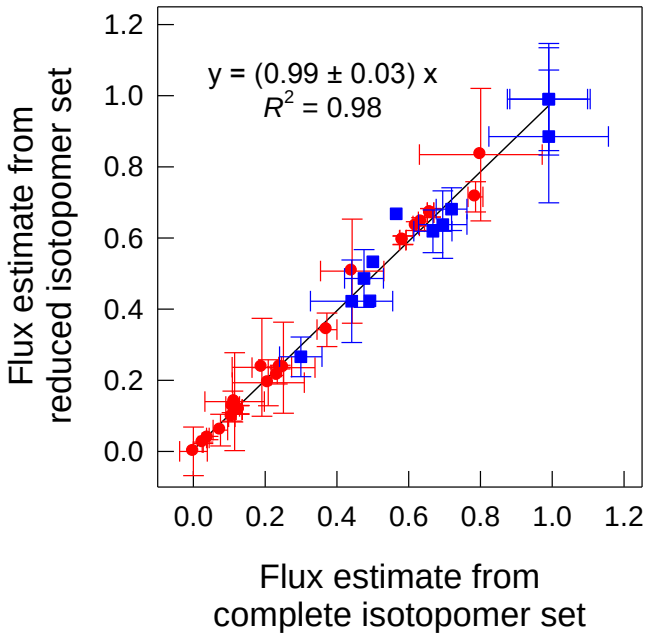


Figure 4.9 Assessment of robustness of global best-fit flux solution. Estimates of free fluxes derived from the complete isotopomer data set were compared with those obtained after removal of the 176 measurements that contributed 80% of the total residuum, together with the 49 measurements that were thereby orphaned. Error bars indicate the standard deviation of the best fit estimate for each net flux (●) or normalized exchange flux (■). The residuum of the flux solution obtained using the reduced set of isotopomer measurements was 245. Linear regression analysis demonstrates that the flux maps obtained from the complete and reduced data set are equivalent.

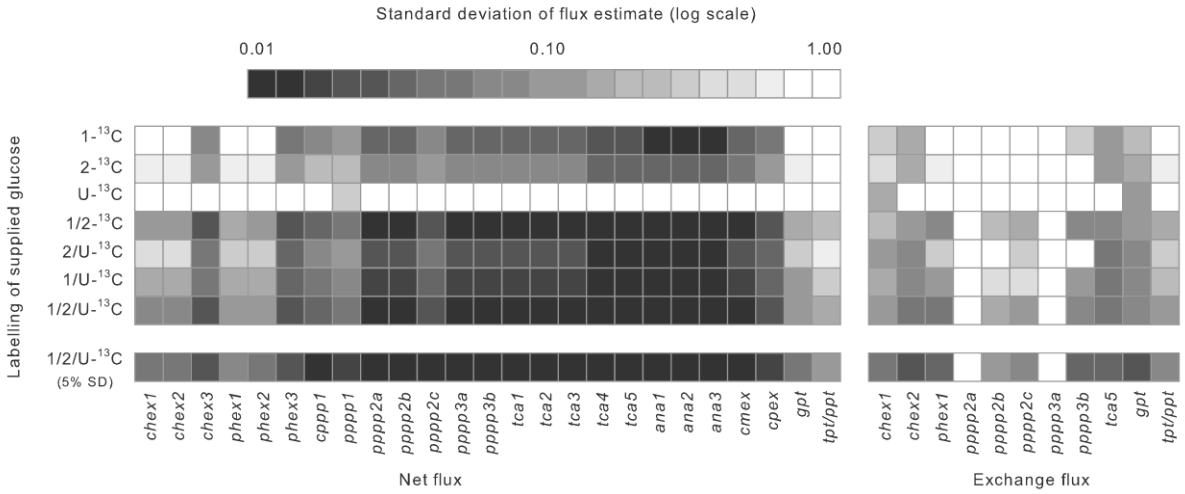


Figure 4.10 Statistical analysis of flux estimates obtained with different steady-state labelling strategies. The standard deviations of optimum flux estimates were derived from the global best fit solution of the metabolic network defined in Fig. 4.3 using the EstimateStat routine in 13C-FLUX. Isotopomer measurements obtained from analysis of extracts of cell cultures labeled separately with either [1-^{13}C]glucose (1-^{13}C), [2-^{13}C]glucose (2-^{13}C) or [U-$^{13}C_6$]glucose (U-^{13}C) were used independently or in combination. Fluxes were expressed relative to the rate of glucose consumption (set to unity) and standard deviations are coded from black (low; flux well defined) to white (high; flux poorly defined) as indicated by the logarithmic grayscale at the top of the figure. The bottom row contains an analysis of data combined from all three feeding strategies (1/2/U-^{13}C) in which the standard deviation of all isotopomer measurements was reduced to 5%.

5 Effect of nitrogen sources on fluxes through central metabolism in Arabidopsis cell cultures.

5.1 Introduction

The previous chapter on metabolic flux analysis (MFA) of a heterotrophic cell suspension of *Arabidopsis thaliana* grown on normal MS containing about 20 mM NH_4 and 40 mM NO_3 suggested that assimilation of inorganic nitrogen (N) might be constrained by the provision of reducing power (NADPH). Since growth on nitrate alone should increase the relative demand for reductant (Fig 5.1), it is possible to test the hypothesis that the supply of reducing power limits biosynthesis by comparing metabolic phenotypes of cells grown on different inorganic sources of N.

The major objective of this chapter is to assess this proposal by measuring the flux distribution in the central network of heterotrophic carbon metabolism using steady-state MFA, and then evaluating the contribution of the oxidative pentose phosphate pathway (oxPPP) and other steps to the supply of reductant. The specific aims were to examine the effect of an ammonium-free medium on cellular biomass production, flux distribution through central metabolism, carbon conversion efficiency and study the provision of reductant for biosynthesis.

Radio-respirometric studies using specifically-labelled [^{14}C]substrates (glucose and gluconate) established that adaptation to N is associated with a change in the flux distribution in central carbon metabolism in *Arabidopsis* cell suspension. Biomass output fluxes were

quantified from measurements of [U-^{14}C]glucose metabolism, and additional metabolic constraints needed for an over-determined system were provided by quantification of mass and positional isotopomers of metabolites obtained by GC-MS and NMR, respectively, after supplying the cells with [1-^{13}C]-, [2-^{13}C]- or [$^{13}C_6$]-glucose. Multiple fluxes across the central network were obtained by mathematical modeling using 13CFLUX® at metabolic and isotopic steady-state. Compartmented flux maps, which give the ultimate metabolic readouts of the phenotypes, were obtained for cells grown on different N sources.

5.1.1 Significance of studying the effect of N in plant metabolism

Nitrogen (N) is an essential macronutrient for maintenance of life in plants as a key constituent of amino acids, protein, nucleotides, chlorophyll and other cellular metabolites (Kraiser et al., 2011). Only a few plant species enter into symbiotic relationships with microbes (e.g., legume-rhizobia) that can fix atmospheric nitrogen. The majority of plants depend on the nitrogen resources available in the soil as an organic or inorganic form; with the later being the main source in temperate climate conditions (Dechorgnat et al., 2011). Also, crops are dependent on fertilizers rich in inorganic form of N - nitrate, ammonium or urea. Hence, a full understanding of regulation of N uptake and assimilation by plants with innovative systems biology and other omic approaches will lead to better crop productivity (Preface, Special issue on Nitrogen, Journal of Experimental Botany, 2011).

Transcriptome analysis shows that the availability of nitrogen for assimilation has a major impact on the gene expression of several central and secondary metabolic pathways in *Arabidopsis* seedlings (Scheible et al., 2004). Also, metabolomic profiling studies in plants has established that significant changes in metabolism can be observed by different N regimes

(Urbanczyk-Wochniak and Fernie, 2005) and that carbon and N metabolism are tightly coordinated (Gutierrez et al., 2007; Nunes-Nesi et al., 2010). This implies any perturbation in the nitrogen levels or sources may influence the metabolic activity of the central pathway reactions thereby altering the phenotype. In this chapter, I intend to understand the effect of inorganic N sources on the metabolic flux phenotypes of the *Arabidopsis* cell suspension.

5.2 Results

5.2.1 Growth of *Arabidopsis* cell suspension cultures under different N sources

In order to choose appropriate feeding conditions that could meet the objectives, experiments were conducted to evaluate the growth of *Arabidopsis* cell suspension cultures initially on four different levels of inorganic N sources. All the feeding experiments were conducted in flasks containing 42 ml of medium (Section 2.2.3). The concentrations of different N sources used in this study were

(a) MS (39.4 mM NO_3^- and 20.61 mM NH_4^+) normal;

(b) N20 (20 mM NO_3^-) only nitrate;

(c) N40 (40 mM NO_3^-) only nitrate; and

(d) NH_4-20 (20 mM NH_4) only ammonium

The fresh weight of N40 cultures was significantly higher than MS (Fig 5.2). After 5.5 days of growth the total fresh weight of MS cells was measured to be 8.024 ± 0.92 g/flask and that of N40 cultures was 11.026 ± 0.63 g/flask. Fresh weight of N20 was not significantly different from either MS or N40 cultures. Cultures fed with only ammonium (NH_4-20) survived for two

days. The death of these could be due to ammonium toxicity reported for higher plants (Britto & Kronzucker 2002).

5.2.2 The uptake and assimilation of nitrate and ammonia

The measure of the nitrate and ammonium proportions metabolized by the cells is necessary to estimate the net biosynthetic demand of reducing power. Also the availability of N sources throughout the growth period needs to be established to select ideal conditions for MFA studies. The uptake and assimilation of NO_3^- and/or NH_4^+ by the cells on different N sources was measured from ^{14}N NMR spectra.

The concentrations of NO_3^- and/or NH_4^+ taken up by the cells from day 0 to day 8 (Fig 5.3, Fig 5.4) were obtained from the cell culture filtrate (CCF). Day 0 corresponds to fresh media with known concentrations of N. It was established that the NO_3^- levels in the CCF from the MS treatment (Fig 5.3A) were not completely depleted, whereas the available NH_4^+ was taken up by the cells by day 6. The NO_3^- in the CCF of the N20 treatment (Fig 5.3B) could be detected until day 6. In contrast to N20, the N40 treatment (Fig 5.3C) had detectable NO_3^- levels in the CCF until day 8. Cells grown on only ammonium (NH4-20) did not survive beyond day 2, hence were not considered further.

The MS and N40 treatments were considered for further flux analysis studies for two reasons. First, both conditions have sufficient and similar levels of NO_3^- throughout the entire growth period (Figs 5.3A & 5.3C). Second, the presence (MS) and absence (N40) of NH_4^+ in the growth medium is hypothesized to influence the demand for reducing power thereby altering fluxes in the network for their maintenance and biosynthesis.

From the total uptake of the N sources, about equal proportions of NO_3^- and NH_4^+ were metabolized by normal Arabidopsis cell cultures (MS) as evidenced from measurements made at day 5.5 (Fig 5.5). This measurement was used for calculating the capacity of the cells to meet the demand for reductant and ATP.

5.2.3 Rate of glucose uptake

Depletion of substrate over the entire growth period can be observed using ^{1}H NMR of the culture medium (Allen et al., 2009). Aliquots of culture filtrate from MS and N40 cultures at days-0, 2, 4, 6 and 8 were subjected to ^{1}H NMR which clearly shows the depletion of glucose in MS and N40 cultures (Fig 5.6). The area of signals from proton of carbon 1 (C1α, C1β) were used to measure the concentrations of glucose (Fig 5.6C). The rates of glucose uptake by the cells in the MS (23.56 ± 2.16 mM/flask/day) and N40 (36.45 ± 6.93 mM/flask/day) conditions is obtained from the depletion phase between day 2 and day 6.

5.2.4 Validation of isotopic steady state condition and establishing the existence of unlabelled pools

As already discussed in previous chapters, steady state MFA requires the system to reach isotopic steady state. In a typical plant cell, the central free metabolites usually reach isotopic steady state within few hours of labelling, whereas polymers like protein, cell wall and starch take several days (Troufflard et al., 2007). This is also verified by in-vivo ^{13}C-NMR studies on the *Arabidopsis* cell suspension cultures (data not shown) which established that the central free metabolites reach isotopic steady state within 2-3 hours. Therefore, labelling analysis of soluble metabolites will be ideal for establishing isotopic steady state conditions. Also previous studies on the same cell lines of Arabidopsis suspension cultures on normal

medium established that the labelling in protein reaches isotopic steady state by 5.5 days (Williams et al., 2008).

In the current study, the isotopic steady state condition for MS and N40 cells was verified by monitoring the extent of average ^{13}C labeled in metabolites after feeding the cells with 20% [$^{13}C_6$]glucose (uniformly labeled) for 5.5 days. At isotopic steady state the average ^{13}C in the metabolites extracted from MS and N40 cells was expected to be about 20%. A two stage analysis was performed. First, the focus was to estimate the label in soluble metabolites – amino acids, organic acids and sugars - to establish the isotopic steady state condition. N40 cells contained very low or negligible amounts of soluble amino acids as no major signals were detected from the GC-MS profiles (Fig 5.7). This was also later confirmed from biomass estimations (section 5.2.6). Hence the current analysis was based on organic acids and sugars which were detected. Average ^{13}C of organic acids and sugars were measured to be about 20% (Fig 5.8, Table 5.1) which established the isotopic steady state condition of MS and N40 cells after 5.5 days of feeding with labeled substrate. However, it must be noted that the average ^{13}C of some of these metabolites is slightly less than 20% (marked by dotted horizontal line in Fig 5.8) which could be due to either the metabolic property of the cells to assimilate atmospheric CO_2 (unlabelled) via the PEP carboxylase reaction or to the presence of pre-existing stored pools that have very slow turnover rate.

Secondly, the focus was to measure the extent of label in protein, cell wall and starch to establish the presence of any pre-existing pools. The average ^{13}C of most of the metabolites obtained from hydrolysis of protein, starch and cell wall was considerably less than the expected value of 20% (Table 5.1). This can only be explained by the presence of unlabelled pre-existing pools of protein, cell wall and starch. Their existence is further verified in the

modeling exercise of ^{13}C MFA (section 5.2.8) and is accounted for as described in chapter 3 in detail.

5.2.5 ^{14}C radio-respirometric analysis indicates that central metabolism responds to the nitrogen source

Radio-respirometric analysis involves the measurement of $^{14}CO_2$ released from cells metabolizing radio-isotopically labeled substrates. The relative rates of $^{14}CO_2$ released from non-photosynthesising plant cells fed with different positionally labeled ^{14}C substrates provides a suitable approach for estimating flux through the central pathways of carbohydrate oxidation (Stitt and ap Rees 1980; Garlick et al., 2002; Harrison and Kruger, 2008). Previous radio-respirometric studies on *Arabidopsis* cell suspension have validated the approach of using an alkaline trap (KOH) to capture the released $^{14}CO_2$ without altering the metabolism (Kruger et al., 2007; Harrison and Kruger, 2008).

To study the effect of nitrogen sources on the carbohydrate oxidation reactions, MS and N40 *Arabidopsis* suspension cultures (4.5 days old, dark grown) were incubated with 3.7 kBq of [1-^{14}C]-, [2-^{14}C]-, [3,4-^{14}C]-, [6-^{14}C]glucose or 3.7 kBq [1-^{14}C]gluconate for 48 hrs. The released $^{14}CO_2$ was monitored at 2 h intervals for 12 h and again 24, 36, and 48 h after the beginning of the incubation. These specifically labeled respiratory substrates released different amounts of CO_2 based on the cellular metabolism of the MS and N40 phenotypes (Fig 5.9). Further the relative ratios of CO_2 reflected the metabolic differences in the carbohydrate oxidative reactions under MS and N40 conditions (Fig 5.10). This analysis was used as a quick basis for evaluating the effect of N sources on the patterns of carbon

metabolism prior to conducting further experiments to elucidate detailed metabolic flux phenotypes by ^{13}C MFA.

After 24 hours of incubation, these cultures are known to reach isotopic steady state (5.5 d old) as supported by previous experiments on the same cultures (Masakapalli et al., 2010). Hence making the comparisons at this time point (Fig 5.11) sheds light on the carbohydrate oxidation status under different N sources. Detailed analysis of the amount of the $^{14}CO_2$ released (Figs 5.9, 5.11A) in MS and N40 conditions followed an expected trend of C3,4 > C1 > C6 > C2 from specific positions in glucose. Also the ratios of $^{14}CO_2$ released (Figs 5.10, 5.11B) from different combinations of positionally labelled substrates imply oxidation of glucose via a combination of glycolysis, oxidative pentose phosphate pathway (oxPPP) followed by the TCA cycle, with recycling of triose-P to hexose-P and some pentan synthesis (Stitt and ap Rees 1980; Malone et al., 2006; Kruger et al., 2007; Harrison & Kruger 2008**)**. Significant differences in the ratios (Fig 5.11B) imply alterations in carbohydrate oxidation under different N conditions. The release of CO_2 in [1-^{14}C]gluconate fed culture indicates the relative activity of the oxPPP. The proportion of $^{14}CO_2$ released through the oxPPP was greater (Figs 5.9, 5.11A) in N40 than MS which could be due to an expected increase in demand for NADPH in N40 cells.

5.2.6 Biomass measurements

As already introduced before, the output fluxes that need to be measured for steady state flux analysis in plants are predominantly directed towards the synthesis of protein, cell wall, lipids, starch, organic acids, free amino acids and neutral sugars such as sucrose. Radiolabelling experiments can be used to quantify flux into biomass components by measuring the amount

of radioactivity in the end products (Alonso et al., 2007; Masakapalli et al., 2010). Previous biomass studies on the *Arabidopsis* cell suspension have established that under normal growth conditions, the uptake of label is linear over 24 h, and the proportion of radioactivity recovered in each of the major chemical fractions was the same after 12 and 24 h incubation (Masakapalli et al., 2010). This implies *de novo* synthesized metabolite pools would reach isotopic steady state within 24 hours.

5.2.6.1 ***Rate of Biomass accumulation (relative to glucose)***

To obtain the biosynthetic output constraints for flux analysis and also to investigate the effect of nitrogen sources on the biomass rates, MS and N40 *Arabidopsis* suspension cultures (4.5 days old, dark grown) were incubated with 74 kBq [U-^{14}C] glucose for 24 hrs. As described in detail (Chapter 2 , materials and methods), the cells and the culture filtrate were harvested and the radioactivity in various metabolic fractions that contain the free amino acids (cationic fraction), organic acids (anionic fraction), sugars (neutral), ethanol (from culture filtrate), cell wall, protein and starch (from insoluble residue), lipids (chloroform fraction) and un-metabolised glucose (from culture filtrate) were measured (Table 5.2A). Total $^{14}CO_2$ released by the cells was measured from parallel flasks by acidifying the medium with formic acid followed by trapping the CO_2 by KOH (Table 5.2A). The biomass proportions and the relative carbon atom% of soluble amino acids, organic acids and protein-amino acids were measured from GC-MS of sample fractions obtained from parallel unlabelled feeding experiments (Table 5.2b). As an advanced approach the extent of labelling in each metabolite (derived from [$^{13}C_6$]glucose labelling experiments, section 5.2.4) at steady state conditions were used to deduce the actual carbon atom abundances that were generated due to metabolism (Table 5.2B). This approach ensures there is no influence of pre-existing pools, a common feature in

plant cells; on the abundance of *de novo* synthesized biomass. The proportions of biosynthesized sugars in the neutral fraction were derived from the TLC separation of the neutral fraction (from ^{14}C samples) followed by phosphor screen image analysis (Fig 5.12). The biomass accumulation rates were calculated as proportional and molar output fluxes relative to the glucose uptake as described in Table 5.2C. All the biomass fluxes were used as constraints in the ^{13}C MFA of the MS and N40 models.

5.2.6.2 *Analysis of [U-^{14}C]glucose metabolism shows altered biomass production*

Label incorporation into end-products was affected by the nitrogen source. Relative to MS, cells in N40 showed (Figs 5.12, 5.13, Tables 5.2A, 5.2B, 5.2C):

- Enhanced accumulation of sucrose and organic acids;
- Decreased protein synthesis and very low levels of free amino acids;
- Increased carbon conversion efficiency – 69% versus 55% (MS).

While this analysis defines the major output fluxes associated with growth, further information is required to determine the fluxes that generate the precursors needed for biomass production, and this can be provided by steady-state MFA.

5.2.7 Label measurements

As discussed in previous chapters, ^{13}C-based steady state MFA depends on the additional constraints of label measurements in the form of isotopomers of key metabolites of the network. To summarize, there is a strong relation between the isotopomers and flux. The labelling in biosynthetically derived product metabolites can report on the labelling of their immediate precursors or central metabolites that in turn report metabolic fluxes in a retro-biosynthetic approach (Szyperski 1998). In plants, measuring fluxes reliably via parallel

pathways occurring in different compartments depends on the labelling of reporter molecules. For example labelling of starch and sucrose would report on the label of cytosolic and plastidic G6P respectively, which may have critical information to derive fluxes through the parallel glycolytic reactions (Allen et al., 2007; Masakapalli et al., 2010). Anther crucial parameter for optimal flux analysis is the choice of labeled substrate (e.g., [1-^{13}C]glucose) used in the feeding studies. Typically, in 13C MFA studies, different labeled substrates and the quality of isotopomer measurements provide different degrees of sensitivity for determining fluxes through a set of reactions (Libourel *et al.*, 2007; Masakapalli et al., 2010). Therefore, care must be taken to obtain high quality measurement datasets in order to deduce robust flux maps that could distinguish the phenotypes under different N sources.

5.2.7.1 ***Choice of labeled substrate***

Parallel feeding experiments with multiple substrates has been successfully applied to obtain statistically robust flux maps (Chapter 4; Schwender et al., 2006; Masakapalli et al., 2010). Based on Chapter 4 and the isotopic steady state conditions established (section 5.2.4), the *Arabidopsis* cell suspension cultures were fed with 99%[1-^{13}C]glucose (1-^{13}C), 99%[2-^{13}C]glucose (2-^{13}C) or 20%[$^{13}C_6$]glucose (U-^{13}C) independently for 5.5 days under different nitrogen sources (MS and N40 conditions). The cells were harvested and the labelling in free metabolites and other end products - protein, starch and cell wall were analysed as below. A summary of the feeding strategies and measurements made is presented in Table 5.3

5.2.7.2 ***Analysing the redistribution of label***

Label measurements were obtained from analysis of extracts of the cell cultures by ^{13}C-NMR and GC-MS (Ratcliffe and Shachar-Hill, 2006; Allen and Ratcliffe, 2009). Both these

analytical approaches were used to maximise the confidence in the deduced fluxes (Kleijn *et al.*, 2007; Kruger et al., 2012).

The positional isotopomers of the soluble metabolites obtained from 1-^{13}C, 2-^{13}C and U-^{13}C feeding experiments, were measured by ^{13}C-NMR (Figs 5.14, 5.15, 5.16). Positional labelling of carbons from glucose monomers of starch were obtained from NMR spectra of 1-^{13}C and 2-^{13}C samples (Fig 5.17). The abundances from the NMR peaks were obtained by line-fitting and the errors over the measurements were estimated as described in Chapter 3 (NMR data handling for 13C MFA).

The mass isotopomers of metabolites were obtained from GC-MS of individual sample fractions obtained from 1-^{13}C, 2-^{13}C and U-^{13}C feeding experiments. Cationic and anionic fractions containing the free amino acids (Fig 5.7) and organic acids (Fig 5.18) respectively were derivatised with TBDMS. The neutral fractions (Fig 5.19) containing the sugars were derivatised with MeOX TMS. The labelling of amino acids derived from protein (Fig 5.20; TBDMS derivatised) and sugars from cell wall and starch (Figs 5.21, 5.22, MeOX TMS derivatised) were also obtained. All the metabolites from each GC-MS spectra were carefully identified, the resulting mass fragments were validated for their accuracy and precision by an array of tools (Chapter 3). The relative mass isotopomers and their errors were obtained as detailed in Chapter 3. All the mass isotopomers were corrected for the effect of natural abundance before further analysis.

Qualitative analysis of the total ion chromatogram from GC-MS has indicated the inability of N40 cells to store significant amounts of free aminoacids (Fig 5.7). ^{13}C NMR of the perchloric acid extracts further confirmed the low abundance of soluble amino acids (Figs 5.14, 5.15,

5.16). N40 cells accumulated significant amounts of sucrose and organic acids as is evident from the ^{13}C NMR spectra (Figs 5.14, 5.15, 5.16). The NMR and GC-MS isotopomer measurements that are used for modeling are depicted in Appendices 5.1A (NMR data of MS), 5.1B (GC-MS data of MS), 5.1C (NMR data of N40) and 5.1D (GC-MS data of N40). The GC-MS data presented in this thesis were corrected for existence of any unlabelled pools by methods developed in Chapter 3 (GC-MS Data handling).

5.2.8 Metabolic modeling

Metabolic modeling was done in 13C-FLUX (version 20050329; Wiechert et al., 2001) using all the *in silico* strategies detailed in Chapter 3 to obtain robust flux maps of *Arabidopsis* suspension cultures grown under the two different N sources (MS and N40). The major prerequisites for steady state MFA - a model network of Arabidopsis, isotopomer measurements and biomass rates measured at isotopic steady state were fulfilled. The followings experimental data obtained under different N sources were used for constraining the parameter fitting of the metabolic models

i) the molar output fluxes (Relative to glucose uptake) towards biomass components derived from [U-^{14}C]glucose feeding experiments (Table 5.2C). The glucose uptake rate was constrained to 1.

ii) the positional isotopomers derived from ^{13}C-NMR and mass isotopomers derived from GC-MS analysis of metabolites from separate feeding experiments with 99%[1-^{13}C]glucose (1-^{13}C), 99%[2-^{13}C]glucose (2-^{13}C) or 20%[$^{13}C_6$]glucose (U-^{13}C) (Appendices 5.1A, 5.1B, 5.1C and 5.1D). The measurements were handled as discussed in chapter 3 for accurate assessment and precise error estimation.

5.2.8.1 *Metabolic model refinement*

The final metabolic network (FTBL file) with all the reaction stoichiometry and carbon transitions is presented in Appendix 5.2 (MODEL NETWORK). The model network was comprehensively tested for the activity of various reactions (Appendices 5.3A, 5.3B).The same central network of *Arabidopsis* model defined in Chapter 4 (Fig 4.3) was used with two major refinements.

First, additional inputs of external CO_2 and MTHF (methylene tetra-hydrofolate, a 1 carbon compound, defined as CX in the model) were included to account for any natural or partially labelled pools of these metabolites influencing the fitting of related metabolites. For example, the PEP carboxylase reaction (PEP + CO_2 → Malate) can take up atmospheric CO_2 and can contribute to the labelling in TCA intermediates or amino acids produced from them (Fig 3.29). Arginine biosynthesis is another process in which CO_2 is assimilated. Similarly, MTHF can enter metabolism via the reversible reactions of glycine decarboxylation (glycine ↔ CO_2 + MTHF) and serine hydroxymethyl transferase (serine ↔ glycine + MTHF). Other reactions involving MTHF include histidine and methionine biosynthesis. The extent of dilution of metabolites due to these 1-carbon molecules depends on the fluxes of the relevant reactions.

Secondly, proxy reactions were defined for modeling adjustments (see Chapter 3, section 3.4.2.1, Fig 3.25A) to account for unlabelled pools in methionine, histidine, glycine, serine and arginine. This was done on the basis of preliminary modeling attempts in which mass isotopomer, m+0 of these metabolites from protein data, poorly fit and implied existence of unlabelled pools. Also evidence of average ^{13}C labelling in these metabolites (Table 5.1) further confirmed it. Along with the preexisting unlabelled pools, these metabolites were also

influenced by additional uptake of 1-carbon molecules (CO_2 or MTHF). These proxy reactions along with free uptake reactions of CO_2 and CX were tested to fit these metabolites better without influencing other central fluxes. The modeling adjustments were done in such a way that there was no net influence of the proxy reactions on the central fluxes.

5.2.9 Steady-state ^{13}C MFA established that the flux distribution through central metabolism is dependent on the nitrogen source

5.2.9.1 *Fitting of experimental data to metabolic models*

To obtain the flux phenotype of *Arabidopsis* cell suspension cultures under each condition of N source (MS or N40), a triple model containing three sub-networks was used to simultaneously fit the experimental data from the three parallel feeding experiments (@3 replicates for each labelling substrates, a total of 9 datasets). This strategy of multi-model simulations (section 3.2.7) was implemented successfully in obtaining compartmented flux maps of *Arabidopsis* suspension cultures (Chapter 4) and was shown to increase the statistical quality of the flux estimates (Chapter 4, Schwender et al., 2006, Alonso et al., 2007). The parameter fitting of the models, bootstrap Monte Carlo simulations to determine the flux space, simulations for best fit flux estimates and statistical analysis were performed as detailed in chapter 3 (Section 3.2.8, Figs 3.13 and 3.14). Each triple-model of MS or N40 was constrained by their respective (i) biomass output fluxes and (ii) the mass and positional isotopomers of amino acids, organic acids and sugars extracted from cells supplied with [1-^{13}C]-, [2-^{13}C]- or 20% [$^{13}C_6$]glucose.

5.2.9.2 *The metabolic flux phenotypes*

In an additional approach, all the experimental datasets were checked for consistencies by removing one dataset at a time (Jack-knife method) followed by re-optimisation with 100 bootstrap Monte-Carlo simulations to determine the flux space. The same flux space is expected if all the datasets are consistent. Statistical tests of Jack-knife solutions with canonical clustering using MANOVA (multiple analysis of variance) in Matlab showed that all the datasets within a treatment (nine datasets of MS and nine datasets of N40) were consistent, and the flux solutions were different for MS and N40 cultures (Fig 5.23). As all the measurement datasets were established to be consistent, parameter fitting and optimizations were redone on triple-models with all the nine datasets. The flux space of each model was obtained from 500 bootstrap Monte-Carlo simulations. The mean of these flux estimates was used to refit the data to obtain the global bestfit solution.

Assessment of goodness-of-fit of isotopomer abundances to the model predicted abundances showed good agreement in both the conditions of MS (Fig 5.24) and N40 (Fig 5.24). The best fit models passed the Chi^2 criteria (Wiechert et al., 2001) implying that the flux estimates through the network can adequately explain the measurements made (Table 5.4). The best fit flux estimates (Table 5.5) and the flux maps (Figs 5.25, 5.26) defined the metabolic phenotype of *Arabidopsis* cell suspension under different nitrogen sources (MS and N40). The comparative fluxes are presented as histograms in Figures 5.27A and 5.27B (Net fluxes) and Figure 5.27C (XCH[0-1] fluxes). All fluxes were relative to glucose uptake of 1 and were subjected to direct comparisons. The confidence intervals over the free fluxes were also verified by non-linearised statistics method (Fig 5.28a for cytosolic and plastidic PPP fluxes. See Appendix 5.4 for all free fluxes).

Switching to ammonium-free medium (N40; nitrate only) has considerable effect on the fluxes through the network thereby altering the metabolic flux phenotype. The major differences in various pathways of the metabolic network are discussed

i) Oxidative Pentose phosphate pathway

In N40 cultures the net flux through the oxPPP (cppp1+ pppp1) was enhanced by about 9.69% in comparison to MS. This is established from the distribution of Monte Carlo flux estimates of combined ppp1 fluxes from the cytosol (cppp1) and platid (pppp1) (Fig 5.28b). The values for oxPPP obtained from the Monte Carlo flux determinations differed significantly from a normal distribution for N40, $D(85) = 0.100$, $p = 0.035$ and only narrowly failed to differ significantly from a normal distribution for MS, $D(209) = 0.060$, $p = 0.063$ using the Kolmogorov-Smirnov test (with Lilliefors significance correction), and both differed significantly from normality when assessed using the Shapiro-Wilk test, $p < 0.001$. Moreover, the variances of the flux estimates for the two culture conditions were significantly different using Levene's test, $F(1,292) = 30.636$, $p < 0.001$, and this was confirmed by the variance ratio, Hartley's $F_{Max} = 2635$, $p << 0.01$. Therefore the distributions were compared using the non-parametric Mann-Whitney test*. In this analysis oxPPP flux in N40 (Median = 0.317) was significantly greater than that in MS (Median = 0.289), $U = 1784$, $z = -10.74$, $p < 0.001$, $r = -0.63$, indicating a large effect.

The cells responded to nitrate by enhanced generation of reductant (NADPH) by increasing flux marginally via the oxPPP. This aspect of Nitrate-specific response of enhanced expression of oxPPP releated genes (G6PDH and PGDH) is supported by transcriptomic studies in *Arabidopsis* (Patterson et al., 2010, Scheible et al., 2004). The nitrate (NO_3^-) taken

up by the cells is reduced to nitrite (NO_2^-) and ammonium (NH_4^+) which is assimilated into amino acids. This process depends on NADH, ferridoxin and NADPH (Scheible et al., 2004) that are provided by cellular metabolism in response to Nitrogen.

ii) TCA cycle, carbon conversion efficiency and anaplerotic reactions

N40 cultures (Nitrate only) adapted by substantial decrease in flux (by 37.65%) through the tricarboxylic acid (TCA) cycle and decreased CO_2 efflux resulting in higher carbon-conversion efficiency (69% vs. 55%) in comparison to MS. The question of enhanced nitrate reductase (NR) activity under low CO_2 concentration is previously reported (Bloom et al., 2010*)* but this may need further investigation. The enhanced carbon-conversion efficiency is due to increase in flux towards sucrose biosynthesis in N40 (increase by 56 %). Sucrose is known to have role as a signal molecule which can mimic light for NR activity (Cheng et al., 1992).

Although there is less flux directed towards TCA cycle in N40 cells, the biomass measurements indicated enhanced accumulation of organic acids (Fig 5.13). It is known that the organic acids synthesized can act as counter anions to maintain the pH balance of the cell by replacing the OH^- ions that would be resulted while synthesis of amino acids from nitrate (Scheible et al., 2004). In fact the reduced flux in TCA cycle in N40 is due to reduced requirement of carbon backbone for biosynthesis of amino acids. Therefore, flux towards TCA cycle depends on the nature of the available nitrogen source in order to fulfill the precursor requirements for the biosynthesis of storage compounds. Enhanced organic acid but reduced amino acid and protein is in agreement with less requirement of PEP carboxylase (PEPc) activity in N40 whose flux was reduced by 26.8% in comparison to MS.

5.2.10 Capacity of the cells to meet NADPH demand

The capacity of the cells to meet the demand for reductant (NADPH) and other energy provisions via ATP and NADH was calculated from the flux maps (Chapter 5, Masakapalli et al., 2010). N40 cells have the capacity to meet the NADPH demand via the oxPPP. MS cells could not meet the increased demand for reductant using the oxPPP (Fig 5.29), and the flux map pointed to a substantial contribution from an NADP-dependent isocitrate dehydrogenase (tca3).

5.3 Summary

The ammonium-free medium containing NO_3 as sole source of N altered the cellular metabolic phenotype. Surprisingly, cells responded to the ammonium-free medium by decreasing the assimilation of inorganic N while not increasing their demand for reductant. Ammonium-free cells had the capacity to meet the NADPH demand via oxPPP flux alone, whereas MS cells depended on additional sources of reductant. The shift away from nitrogen assimilation towards storage of carbohydrate in cells on an ammonium-free medium decreased flux through the TCA cycle and increased carbon conversion efficiency.

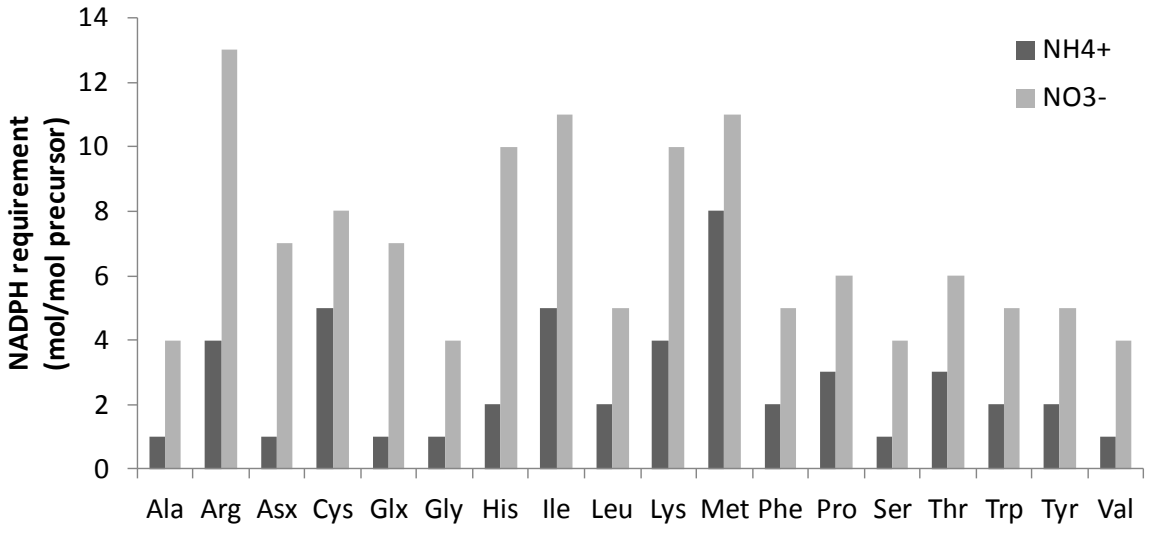


Figure 5.1 Biosynthetic demand of reducing power (NADPH) for each amino acid biosynthesis under different sources of inorganic nitrogen (NH_4^+ and NO_3^-)**.** Cost of assimilating NO_3^- is greater than NH_4^+. The values are obtained from coenzyme stoichiometries for generation of biosynthetic products from intermediates of the central metabolic network (Masakapalli et al., 2010)

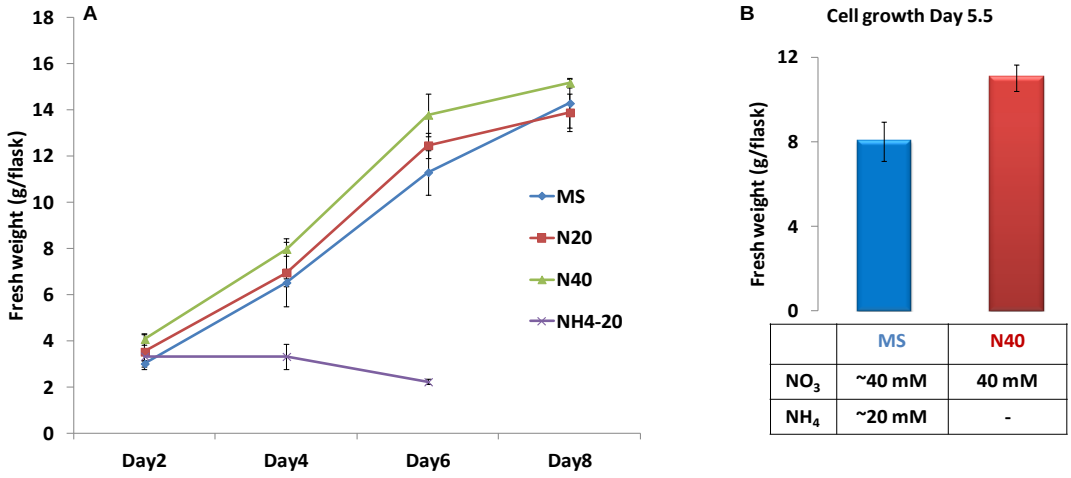


Figure 5.2 Effect of different nitrogen sources on fresh weight of cells. A) Cells were harvested every 2 days from flasks (~42 ml) of heterotrophic Arabidopsis suspension cultures and fresh weight measured. Ammonium NH4-20 treated cells died after day 2 B) Fresh weight of MS and N40 cultures at day 5.5. Values are average ± sd (n=4 for each time point).

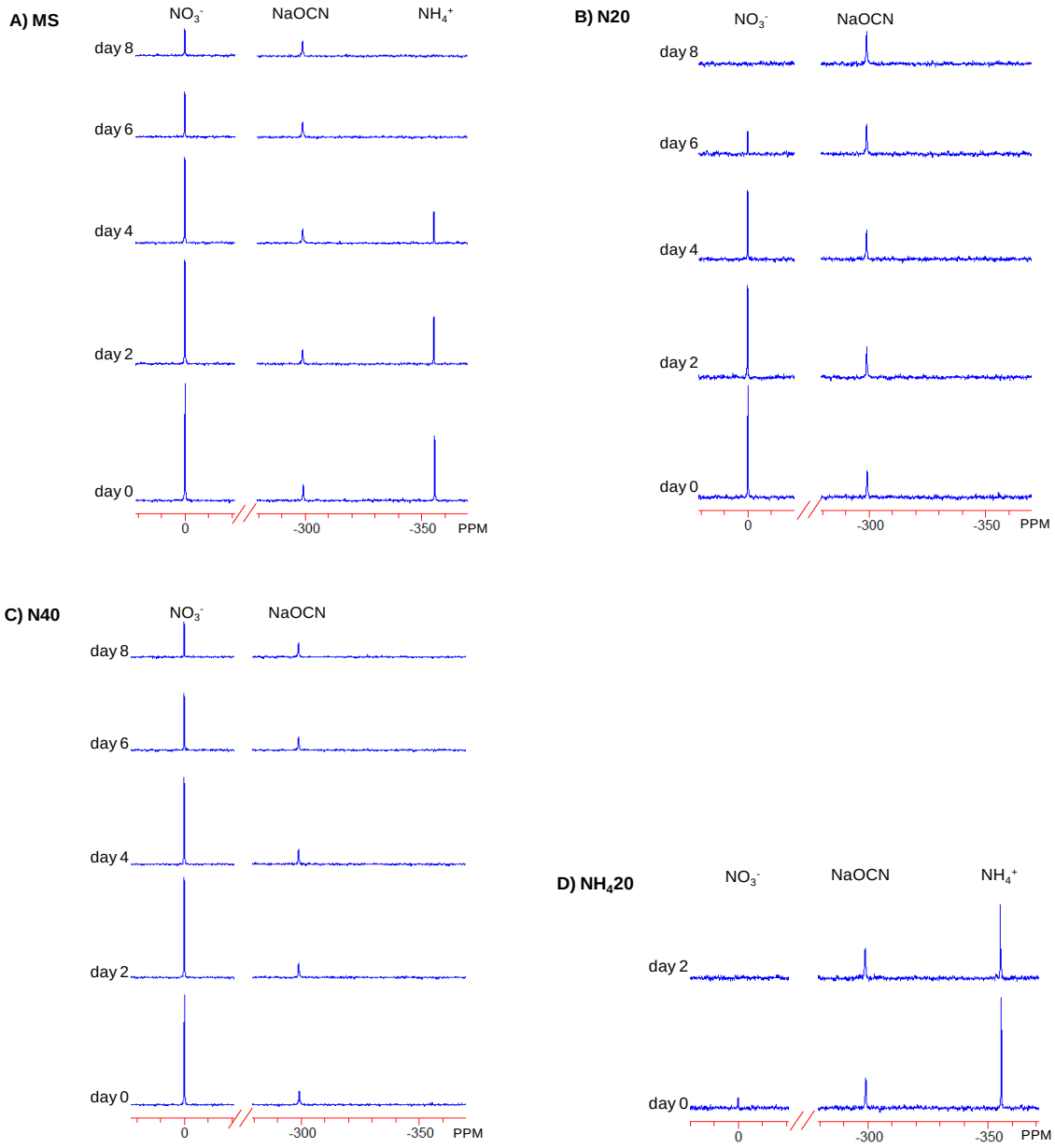


Figure 5.3 Assimilation of nitrate and ammonium. ^{14}N NMR spectra of cell culture filtrate obtained from *Arabidopsis* cultures grown under different N sources for different number of days (A,B,C,D). In each spectra NO_3^- peak is at 0 ppm and NH_4+ at -356 ppm. NaOCN is internal standard. The concentrations are presented in Fig 5.4

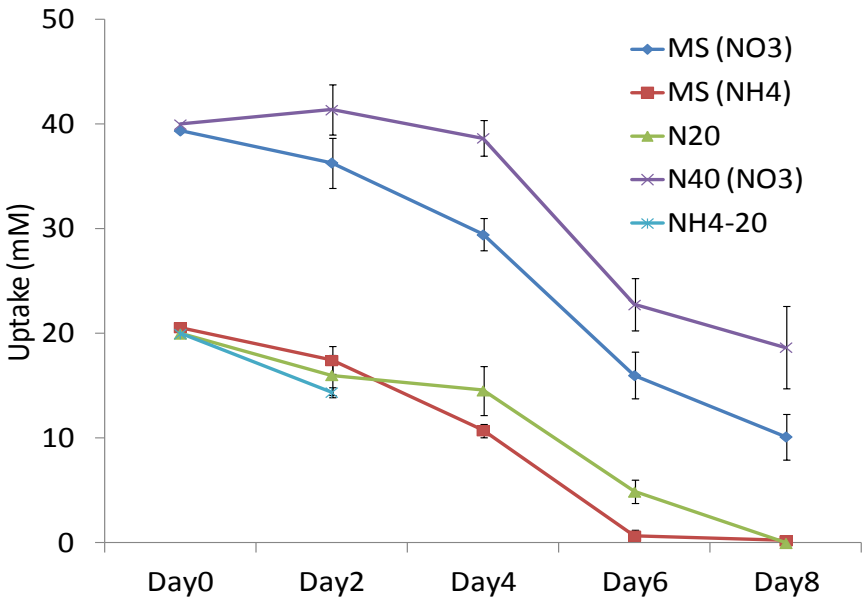


Figure 5.4 Concentration of NO_3^- and NH_4^+ in the culture filtrate of Arabidopsis cell suspension grown on different N sources. The concentrations were measured from signal intensities of ^{14}N-NMR (Fig 5.3).

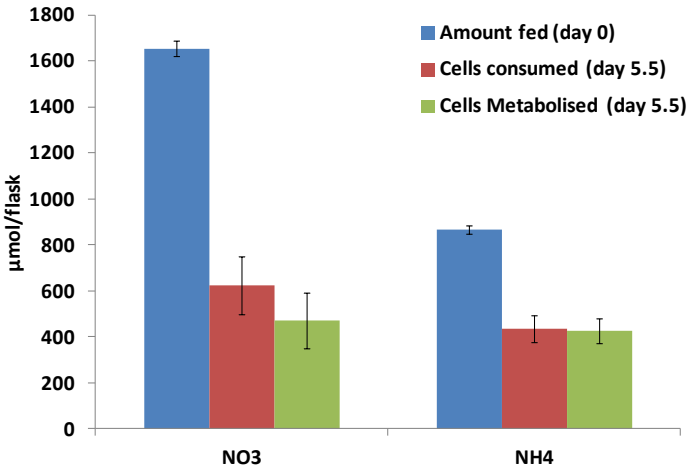


Figure 5.5 Uptake and metabolism of NO_3^- and NH_4^+ in normal Arabidopsis cell cultures (MS) measured at Day 5.5**.** The amount fed and consumed by the cells was obtained from the ^{14}N NMR spectra of culture filtrate harvested at Day 0 and Day 5.5 respectively. The amount of metabolized NO_3^- and NH_4^+ by the cells was obtained from ^{14}N NMR of PCA extracts. The cells assimilate about equal proportions of NO_3^- and NH_4^+

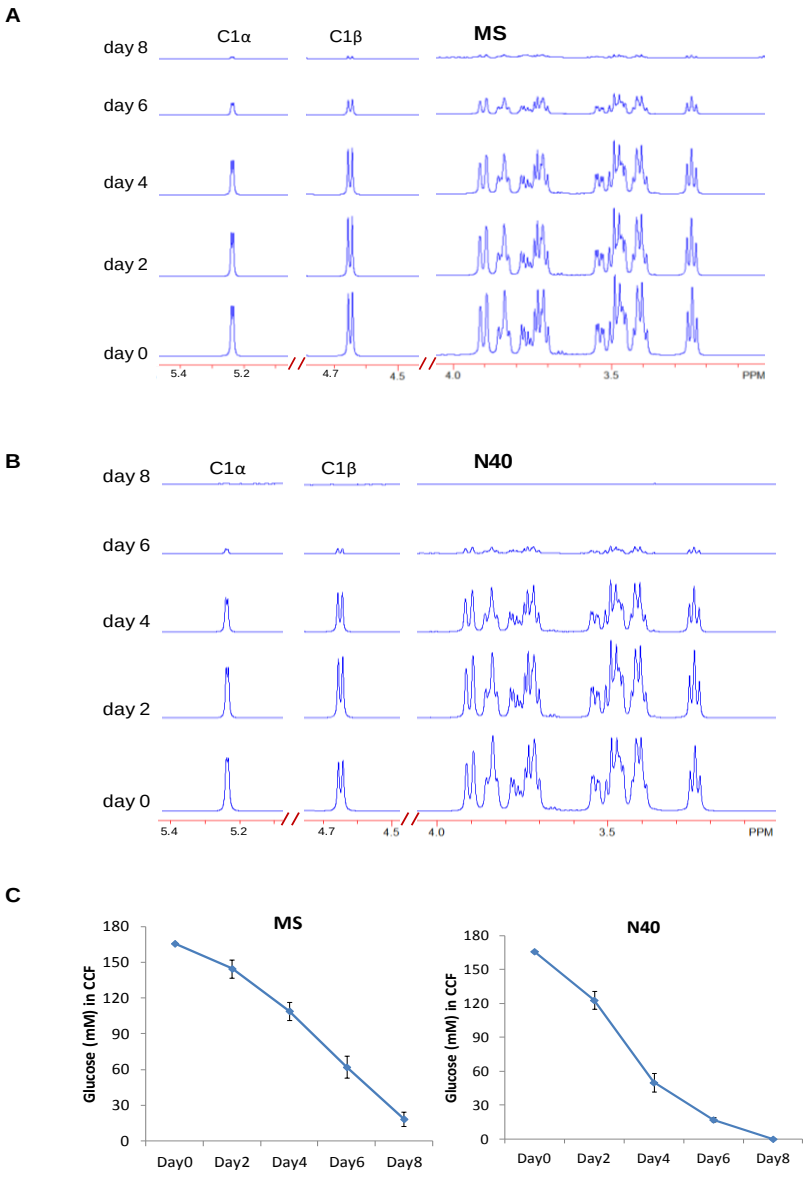


Figure 5.6 Rate of glucose uptake in MS and N40 cells was measured from its depletion detected from the signal intensities of ^{1}H NMR spectra (A, B) of the culture filtrate during the growth period. The area of signals from the proton of carbon 1 (C1α, C1β) was used to measure the concentrations of glucose (C) at days 0, 2, 4, 6 and 8. The initial concentration of glucose (3%) in the medium was 166 mM. The MS and N40 cells were grown in several flaks containing total volume of 42 ml medium. At each time point culture filtrate from 3 separate flasks was harvested and ^{1}H NMR acquired on a standard aliquot. The glucose concentration is presented in main text as mean ± sd from three replicates. TSP was used as an internal standard (signal at ppm 0, not shown) to scale different samples in ^{1}H NMR.

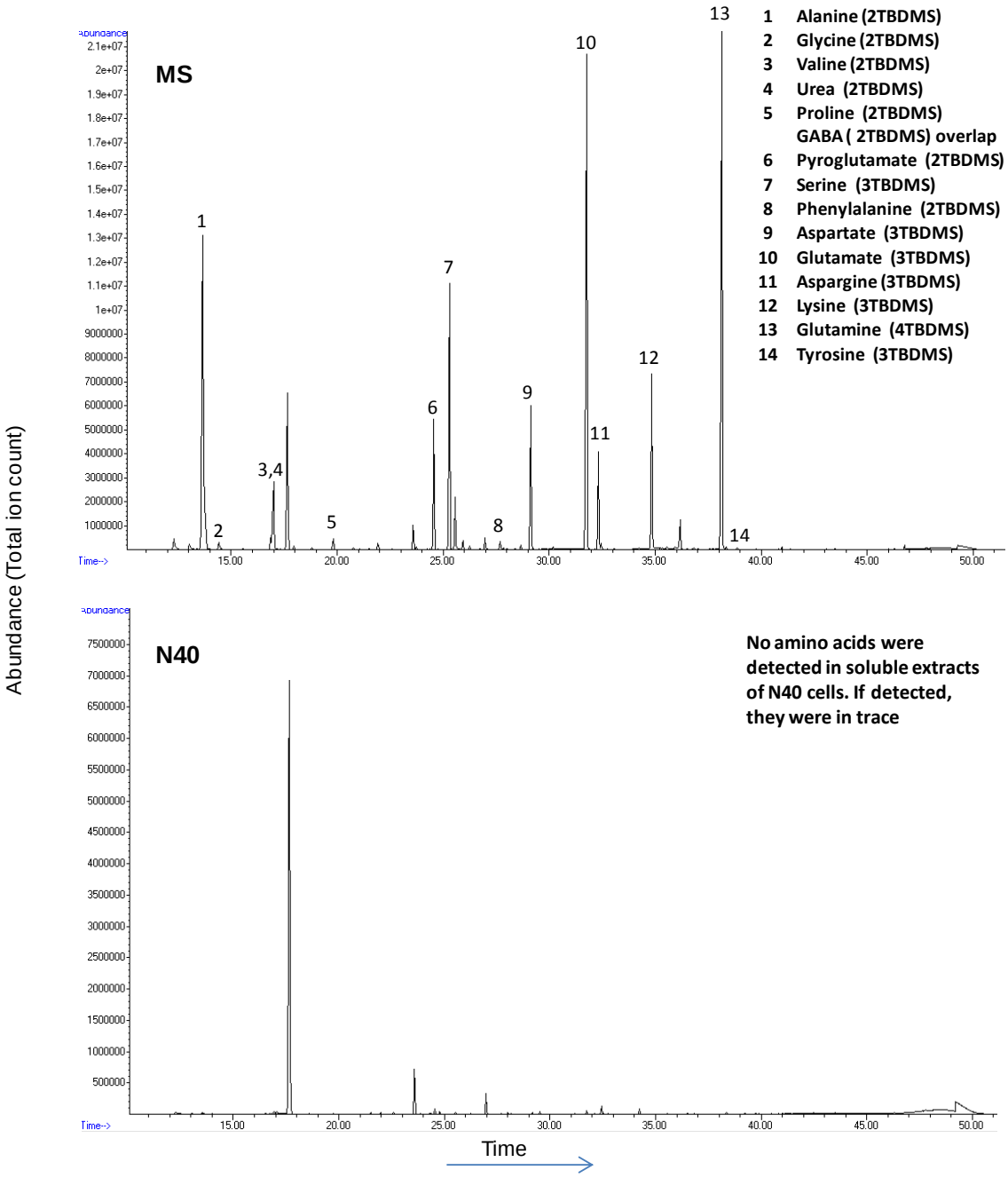


Figure 5.7 Total ion chromatograms of cationic fractions (amino acids) obtained from MS and N40 Arabidopsis cell suspensions fed with 20% [$^{13}C_6$] cells for 5.5 days. No amino acids were detected in the N40 conditions. The fractions were derivatised with TBDMS and the GC-MS spectra were acquired as described in Chapter 2.

Table 5.1 Average labelling of metabolites obtained from 20% [$^{13}C_6$]glucose fed MS and N40 cells of Arabidopsis cultures over 5.5 days. The average ^{13}C is calculated from the mass isotopomer fractions of metabolites that were measured from the GC-MS analysis of suitably derivatised metabolites. The data are presented as average ^{13}C (in %) with standard deviations obtained from 3 replicates for each condition. The extent of labelling in soluble metabolites from sugars and organic acids of MS and N40 establishes the near isotopic steady state condition (~20% average ^{13}C, see text, Fig 5.8). The extent of labelling in protein, cell wall and starch implies the presence of unlabelled preexisting pools which needs to be corrected for ^{13}C MFA (Chapter 3).

Metabolite Source	Derivatised molecule	Carbon atoms	Fragment Ion	Average ^{13}C (%) ± sd	
				MS	N40
Protein	Alanine-2TBDMS	1,2,3	M-57	17.28 ± 0.14	16.42 ± 0.37
	Glycine-2TBDMS	1,2	M-57	15.87 ± 0.11	15.30 ± 0.04
	Valine-2TBDMS	1,2,3,4,5	M-57	15.25 ± 0.57	14.13 ± 0.48
	Leucine-2TBDMS	1,2,3,4,5,6	M-57	13.51 ± 0.15	12.82 ± 0.07
	Isoleucine-2TBDMS	1,2,3,4,5,6	M-57	10.18 ± 0.10	9.20 ± 0.13
	Proline-2TBDMS	2,3,4,5	M-85	14.96 ± 0.36	13.69 ± 0.24
	Methionine-2TBDMS	1,2,3,4,5	M-57	14.48 ± 0.11	13.63 ± 0.34
	Serine-3TBDMS	1,2,3	M-57	16.70 ± 0.59	15.62 ± 0.26
	Threonine-3TBDMS	1,2,3,4	M-57	14.52 ± 0.31	13.31 ± 0.30
	Phenylalanine-2TBDMS	1,2,3,4,5,6,7,8,9	M-57	15.13 ± 0.22	14.06 ± 0.33
	Aspartate-3TBDMS	1,2,3,4	M-57	16.03 ± 0.51	14.65 ± 0.29
	Glutamate-3TBDMS	1,2,3,4,5	M-57	16.82 ± 0.63	16.06 ± 0.34
	Lysine-3TBDMS	1,2,3,4,5,6	M-0	14.64 ± 0.61	13.34 ± 0.35
		1,2,3,4,5,6	M-15	14.18 ± 1.17	12.56 ± 0.90
		1,2,3,4,5,6	M-57	13.70 ± 1.35	12.30 ± 1.37
	Arginine (-NH2)-3TBDMS	1,2,3,4,5,6	M-57	14.27 ± 0.41	12.92 ± 0.16
	Histidine-3TBDMS	1,2,3,4,5,6	M-58	15.11 ± 0.25	14.06 ± 0.23
	Tyrosine-3TBDMS	1,2,3,4,5,6,7,8,9	M-59	15.46 ± 0.11	14.23 ± 0.38
Neutral fraction of PCA extract (Sugars),	Fructose MeOX 5TMS	4,5,6	M-262	18.53 ± 0.25	18.77 ± 0.13
		1,2,3,4	M-205	18.59 ± 0.15	18.81 ± 0.09
	Fructose MeOX 5TMS	4,5,6	M-262	18.17 ± 0.28	18.25 ± 0.14
		1,2,3,4	M-205	17.46 ± 0.28	18.11 ± 0.35
	Glucose MeOX 5TMS	1,2	M-409	20.42 ± 0.29	20.20 ± 0.60
		3,4,5,6	M-250	19.75 ± 0.28	19.68 ± 0.33
	Glucose MeOX 5TMS	1,2	M-409	18.63 ± 0.45	18.66 ± 0.09
		3,4,5,6	M-250	19.13 ± 0.13	19.12 ± 0.06
	Sucrose 8TMS	1,2,3,4,5,6 (glucosyl)	M-90	19.03 ± 0.43	19.55 ± 0.10
	Glycerol 3TMS	1,2,3	M-15	17.50 ± 0.43	16.77 ± 0.87
Anionic fraction of PCA extract (Oraganic acids)	Succinate-2TBDMS	1,2,3,4	M-57	18.22 ± 0.53	18.73 ± 0.17
	Maleic acid-2TBDMS	1,2,3,4	M-57	19.04 ± 0.19	18.71 ± 0.01
	Malic acid-3TBDMS	1,2,3,4	M-57	18.65 ± 0.08	18.78 ± 0.13
	Citric acid-4TBDMS	1,2,3,4,5,6	M-57	18.67 ± 0.16	18.75 ± 0.07
	Pyroglutamate-2TBDMS	1,2,3,4,5	M-57	17.54 ± 0.77	10.23 ± 3.85
Cell wall	Arabinose MeOX 4 TMS	5	M-364	12.42 ± 0.22	11.60 ± 0.65
		3,4,5	M-160	13.78 ± 0.19	12.68 ± 0.83

	Xylose MeOX 4 TMS major signal	5	M-364	11.64 ± 0.19	13.92 ± 0.11
		3,4,5	M-160	13.24 ± 0.21	15.39 ± 0.18
	Xylose MeOX 4 TMS minor signal	3,4,5	M-160	12.86 ± 0.16	15.07 ± 0.10
	Hex1 Mannose	3,4,5,6	M-250	17.81 ± 0.13	14.73 ± 3.74
	Hex2 Galactose	3,4,5,6	M-250	17.85 ± 0.19	18.05 ± 0.02
	Hex3 Glucose1_wall	3,4,5,6	M-250	16.16 ± 0.17	16.74 ± 0.21
	Hex4 Glucose2_wall	3,4,5,6	M-250	17.93 ± 0.20	18.18 ± 0.09
	Pentan_wall	5	M-364	15.30 ± 0.11	15.02 ± 1.64
		3,4,5	M-160	16.66 ± 0.18	15.68 ± 0.13
Starch	Glucose MeOX 5TMS	1,2	M-409	13.57 ± 1.30	17.05 ± 0.09
		3,4,5,6	M-250	14.34 ± 1.25	17.45 ± 0.23
	Glucose MeOX 5TMS	3,4,5,6	M-250	14.20 ± 1.28	17.47 ± 0.27

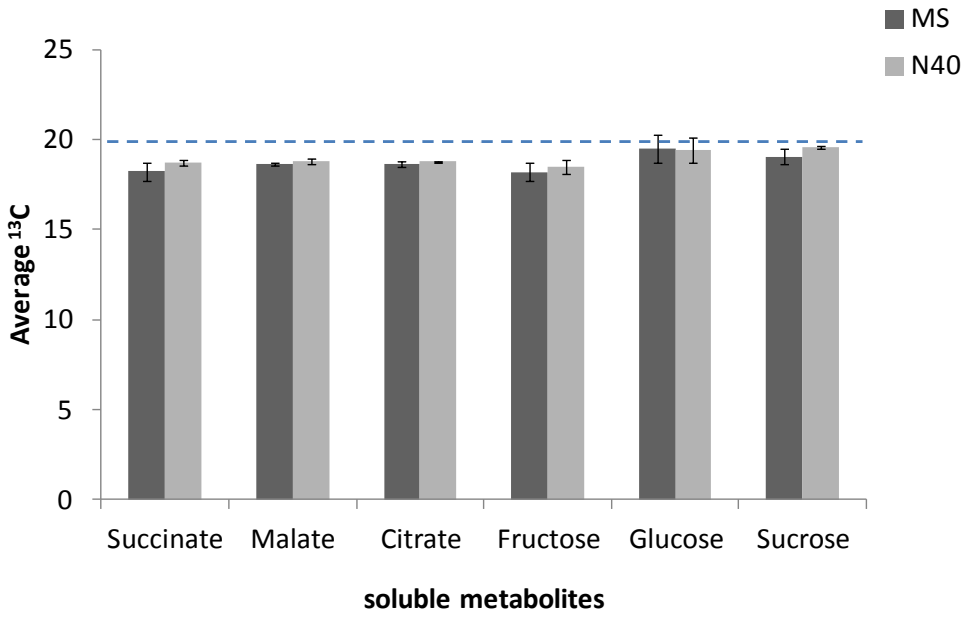


Figure 5.8 Extent of labelling of organic acids and sugars obtained from 20% [$^{13}C_6$] glucose fed MS and N40 cells of *Arabidopsis* cultures over 5.5 days. The average ^{13}C was measured from GC-MS analysis (Table 5.1). The average ^{13}C of these central metabolites in MS and N40 was about 20% implying they reached near isotopic steady state. Soluble amino acids were not present in N40 conditions hence not presented.

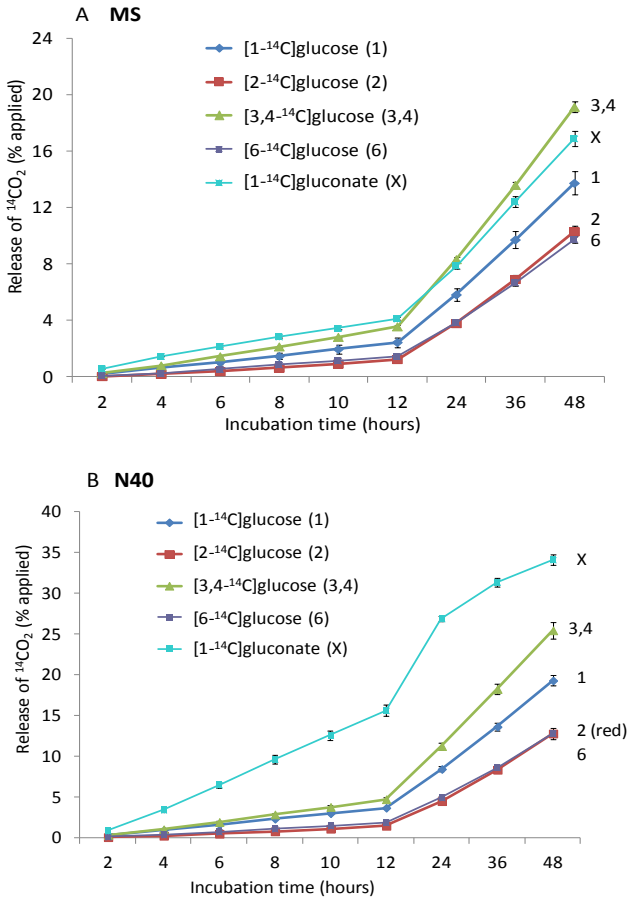


Figure 5.9 Oxidation of specifically labeled [^{14}C] substrates under different N sources (MS (A) and N40 (B)). Release of $^{14}CO_2$ was measured from *Arabidopsis* suspension cultures (4.5 day old in MS and N40 conditions) after incubating with 3.7 kBq of [1-^{14}C]-, [2-^{14}C]-, [3,4-^{14}C]-, [6-^{14}C]glucose or 3.7 kBq [1-^{14}C]gluconate until 48 hrs in separate flasks. Released $^{14}CO_2$ was collected by alkali traps at time points 2h, 4h, 6h, 8h, 12h, 24h, 36h and 48 h. Each value presented is the cumulative $^{14}CO_2$ release as the percentage of applied label ± SE from four replicate flasks. N40(B) shows very high gluconate (indicated as X) metabolism in comparison to MS.

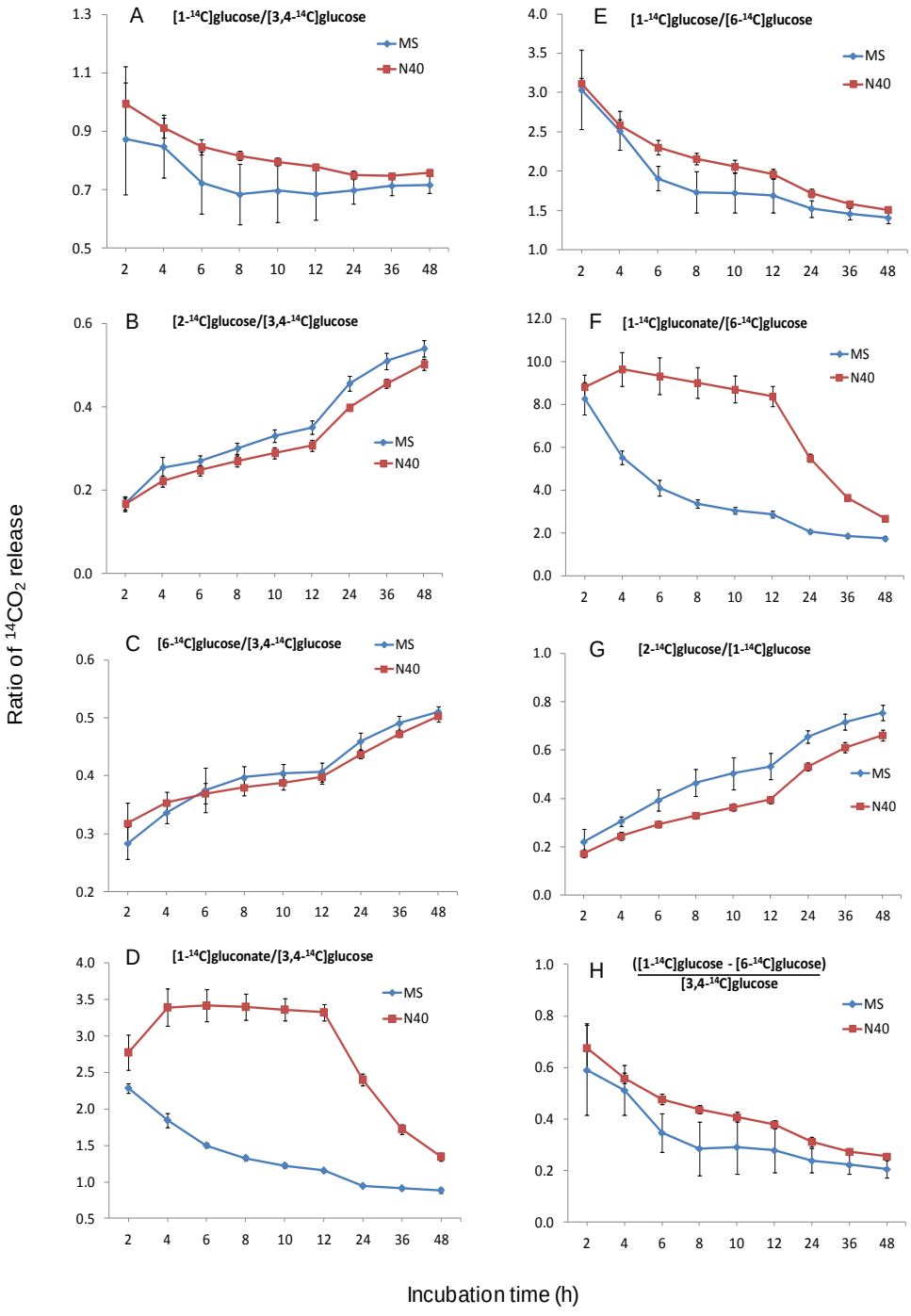


Figure 5.10 Comparison of the relative rates of oxidation of specific positions within labelled substrates by the *Arabidopsis* cell suspension.. Ratios were calculated from the cumulative CO_2 released (Fig 5.9) by different combinations of specifically labeled substrates. Significant differences in the patterns of carbohydrate oxidation were observed in Arabidopsis suspension cultures grown under different N conditions (MS and N40). Each point is the mean ± SE for ratios determined from four separate samples of cell culture.

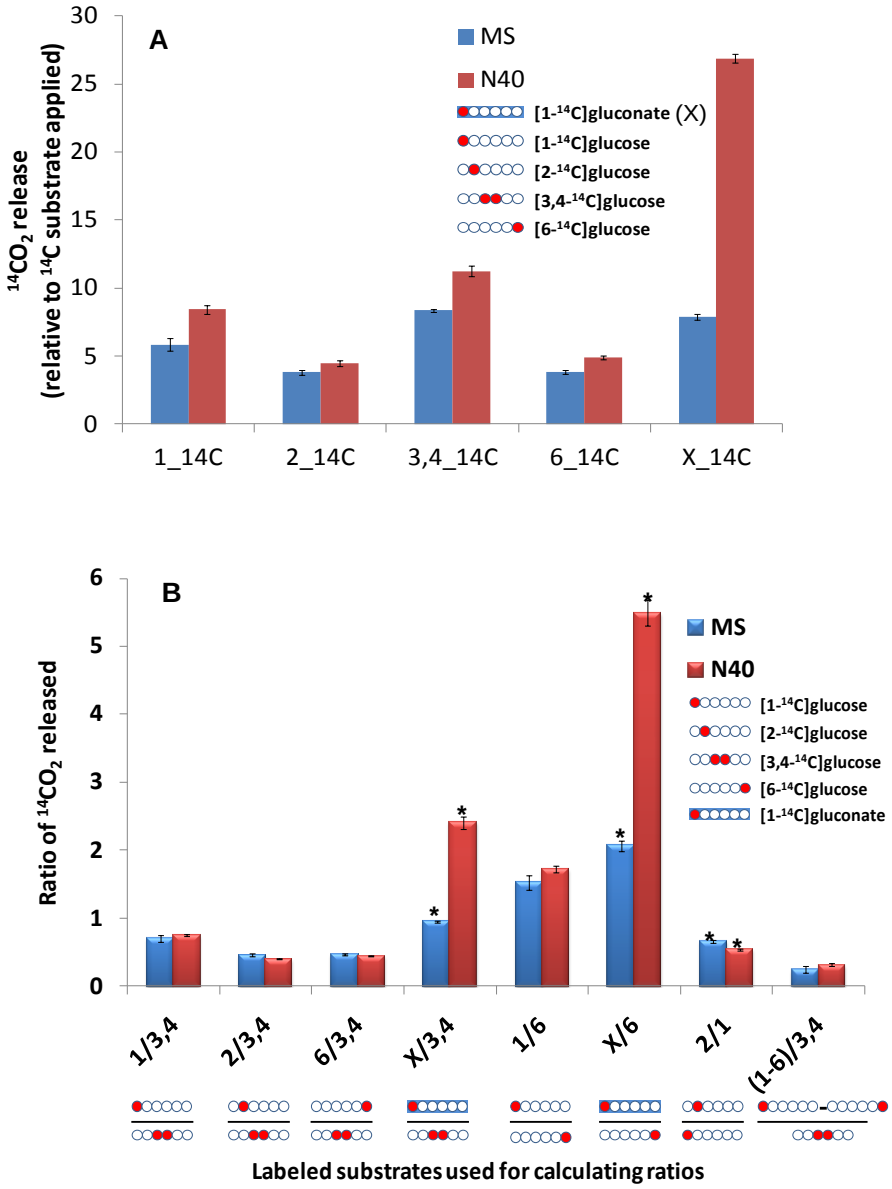


Figure 5.11 Oxidation of specifically labeled [^{14}C] substrates under different N sources after 24 hours of incubation. *Arabidopsis* suspension cultures were incubated with specifically labeled [^{14}C] substrates at 4.5 days (Figs 5.9 & 5.10). After 24 hours, they are known to reach isotopic steady state (5.5 d old) as supported by previous experiments on the same cultures (Masakapalli et al., 2010). The amount of the $^{14}CO_2$ released (A) in MS and N40 conditions followed an expected trend of C3,4 > C1 > C6 > C2 from specific positions in glucose. The ratios of $^{14}CO_2$ released (B) from different combinations of positionally labelled substrates were significantly different (*). Predominantly the differences were observed in the metabolism of gluconate (X). The values are means ± SE determined as in Fig 5.9 and 5.10.

Table 5.2A Radioactive distribution of each biomass fraction and individual metabolites. 4.5 days old *Arabidopsis* suspension cultures grown on MS and N40 conditions were incubated with 74kBq [U-^{14}C]glucose for 24 hours. The total radioactivity (dpm) measured from various biomass fractions are shaded. The radioactivity of individual metabolites from each fraction were estimated from their relative carbon atom % (Table 5.2B for amino acids and organic acids) and TLC (Fig 5.12) for sugars from methanol soluble fraction). The total ^{14}C metabolized by the cells and the recovery % are also presented. All values are mean radioactive counts ± sd derived from 4 separate flasks.

Biomass fractions	Biomass Metabolite	Radioactivity (dpm) in MS and N40 conditions (n=4)	
		MS	**N40**
$^{14}CO_2$ released (alkali trap)			
	$^{14}CO_2$	342026.9 ± 13942.6	419668.7 ± 23773.6
Chloroform extract			
	Lipids	15129.0 ± 2518.1	13184.7 ± 3471.7
Ethanol from cell culture filtrate			
	Ethanol	1076.3 ± 52.2	877.0 ± 31.5
Aqueous methanol extract (Alcohol Solubles)			
Cationic fraction (Amino acids)		87279.7 ± 4064.7	19500.2 ± 6976.1
	Ala	9647.8 ± 4957.4	1440.1 ± 690.0
	Arg	0.0 ± 0.0	0.0 ± 0.0
	Asn	0.0 ± 0.0	0.0 ± 0.0
	Asp	963.3 ± 167.5	36.1 ± 31.3
	Cys	0.0 ± 0.0	0.0 ± 0.0
	Glu	6097.3 ± 1352.8	656.1 ± 414.4
	Gln	39086.8 ± 19300.9	0.0 ± 0.0
	Gly	373.5 ± 265.4	504.6 ± 157.4
	His	0.0 ± 0.0	0.0 ± 0.0
	Iso	0.0 ± 0.0	0.0 ± 0.0
	Leu	77.5 ± 20.2	106.1 ± 55.9
	Lys	20144.7 ± 10601.7	0.0 ± 0.0
	Met	0.0 ± 0.0	0.0 ± 0.0
	Phe	186.1 ± 76.8	806.9 ± 716.0
	Pro	1062.3 ± 478.9	535.9 ± 148.4
	Ser	5065.7 ± 2280.5	594.3 ± 147.9
	Thr	0.0 ± 0.0	0.0 ± 0.0
	Trp	0.0 ± 0.0	0.0 ± 0.0
	Tyr	28.2 ± 8.1	17.1 ± 15.8
	Val	254.7 ± 86.9	109.3 ± 52.5
	Pyroglu	2652.8 ± 558.3	14507.1 ± 6608.1
	Gaba	1639.2 ± 2167.5	186.8 ± 162.3
Anionic fraction (Organic acids)		50361.1 ± 983.4	160687.2 ± 28031.0
	Citrate	21644.3 ± 2889.0	64943.0 ± 20066.9

	Isocitrate	1.2	±	1.3	64.2	±	49.8
	Malate	25267.5	±	2314.2	86324.4	±	9970.2
	Maleic acid	82.2	±	62.7	1189.1	±	589.9
	Pyroglutamate	237.8	±	87.3	44.0	±	7.2
	Succinate	3127.9	±	1289.1	8122.5	±	1674.4
Neutral fraction (sugars from TLC)		119864.8	±	9978.8	488024.3	±	7656.6
	Fructose	30511.7	±	2255.5	123040.4	±	1571.4
	Glucose	30511.7	±	2255.5	123040.4	±	1571.4
	Sucrose	58841.5	±	5469.3	241943.4	±	4625.5
Alcohol insoluble residue (AIR)							
Cationic fraction (Protein)		61366.0	±	3210.6	81870.0	±	398.7
	Ala	3904.4	±	312.7	5784.0	±	438.3
	Arg	2476.2	±	192.5	3713.1	±	831.3
	Asn	0.0	±	0.0	0.0	±	0.0
	Asp	3550.0	±	227.7	4993.2	±	238.6
	Cys	1627.6	±	130.8	1323.8	±	933.2
	Glu	147.3	±	11.4	228.0	±	18.6
	Gln	0.0	±	0.0	0.0	±	0.0
	Gly	2837.9	±	152.1	4420.2	±	49.0
	His	1096.9	±	54.9	1086.8	±	321.9
	Iso	1654.2	±	179.3	2384.6	±	389.4
	Leu	4715.9	±	449.3	6909.0	±	968.0
	Lys	8112.5	±	543.1	9713.3	±	1346.7
	Met	1627.6	±	130.8	1323.8	±	933.2
	Phe	4105.0	±	409.5	5983.6	±	618.7
	Pro	6703.9	±	691.6	7346.4	±	1100.7
	Ser	4734.2	±	95.4	6747.6	±	838.0
	Thr	2012.0	±	271.7	2918.1	±	682.6
	Trp	4105.0	±	409.5	5983.6	±	618.7
	Tyr	4402.0	±	309.6	5616.2	±	89.9
	Val	3553.5	±	237.1	5394.8	±	865.7
Anionic fraction + Cell wall	Cell wall	26852.7	±	6094.7	67698.8	±	3705.4
Neutral	Starch	56074.0	±	4704.4	102086.9	±	3991.6
Total ^{14}C metabolised and recovery							
Total 14C metabolised		760030.5	±	25866.1	1353597.8	±	34189.6
Cell culture filtrate		3586554.5	±	173877.2	2922371.1	±	104876.1
Total recovered		4346585.0	±	151202.5	4275968.9	±	114876.2
Applied radioactivity/flask (74kBq)		4392713.6	±	165545.8	4392713.6	±	165545.8
Recovery		98.95%	±	3.44%	97.34%	±	2.62%

Table 5.2B Composition of metabolites in the soluble (cationic and anionic fractions) and protein of *Arabidopsis* cell suspension culture (5.5 days old) in normal (MS) and ammonium-free (N40) media . The metabolites from each fraction were quantified using GC-MS standard curves (Appendix 2.2) and the relative carbon atom % (calculated as noted below this table) was used to estimate the radioactivity (Table 5.2A) in each metabolite from which relative carbon fluxes are calculated

A. Soluble cationic extract (amino acids fraction):

	Abundance (nanograms/ul of sample injected in GC-MS)								Extent of labelling from [$^{13}C_6$]glucose feeding expt(%)		Relative C atom %					
Metabolite	MS			N40			MW*	No of C	MS	N40**	MS			N40		
Ala	75.50	±	18.16	0.08	±	0.03	89.09	3	100.00	100.00	10.91	±	2.95	8.24	±	3.19
Arg	0.00	±	0.00	0.00	±	0.00	210.70	6	96.63	96.63	0.00	±	0.00	0.00	±	0.00
Asn	nd			nd			132.10	4	nd	nd	nd			nd		
Asp	11.07	±	4.04	0.00	±	0.00	133.10	4	91.22	91.22	1.10	±	0.08	0.21	±	0.12
Cys	nd			nd	±		121.20	3	nd	nd	nd			nd		
Glu	57.13	±	19.99	0.04	±	0.02	147.10	5	96.63	96.63	6.95	±	0.71	3.75	±	1.68
Gln	491.51	±	255.21	0.00	±	0.00	146.10	5	96.63	96.63	45.46	±	13.62	0.00	±	0.00
Gly	3.64	±	0.52	0.03	±	0.00	75.07	2	90.00	90.00	0.42	±	0.16	2.71	±	0.47
His	0.00	±	0.00	0.00	±	0.00	209.60	6	nd	nd	0.00	±	0.00	0.00	±	0.00
Iso	0.00	±	0.00	0.00	±	0.00	131.20	6	66.87	66.87	0.00	±	0.00	0.00	±	0.00
Leu	0.76	±	0.26	0.01	±	0.00	131.20	6	66.87	66.87	0.09	±	0.01	0.61	±	0.26
Lys	200.08	±	47.02	0.00	±	0.00	182.60	6	80.17	80.17	22.77	±	6.33	0.00	±	0.00
Met	0.00	±	0.00	0.00	±	0.00	149.20	5	91.22	91.22	0.00	±	0.00	0.00	±	0.00
Phe	1.21	±	0.34	0.04	±	0.02	165.20	9	78.73	78.73	0.21	±	0.04	4.29	±	1.99
Pro	7.41	±	1.98	0.02	±	0.00	115.10	5	90.11	90.11	1.20	±	0.28	2.82	±	0.26
Ser	51.08	±	13.65	0.04	±	0.01	105.10	3	94.91	94.91	5.74	±	1.34	3.27	±	0.72

Thr	0.00	±	0.00	0.00	±	0.00	119.10	4	91.22	91.22	0.00	±	0.00	0.00	±	0.00
Trp	nd			nd			204.20	11	nd	nd	nd			nd		
Tyr	0.22	±	0.07	0.00	±	0.00	181.20	9	78.73	78.73	0.03	±	0.00	0.09	±	0.05
Val	2.04	±	0.63	0.01	±	0.00	117.10	5	84.51	84.51	0.29	±	0.05	0.64	±	0.28
Pyroglu	22.04	±	7.79	0.60	±	0.12	129.11	5	96.18	96.18	3.03	±	0.29	72.23	±	8.58
Gaba	6.17	±	2.59	0.01	±	0.01	103.12	4	96.63	96.63	1.81	±	1.36	1.12	±	0.68
B. Soluble Anionic extract (organic acids fraction):																
Citrate	23.80	±	3.52	99.55	±	29.29	192.00	6	93.33	93.77	43.02	±	3.55	39.80	±	3.71
Glycolate	0.00	±	0.00	0.00	±	0.00	76.05	2	0.00	0.00	0.00	±	0.00	0.00	±	0.00
Isocitrate	0.00	±	0.00	0.11	±	0.06	192.12	6	93.33	93.77	0.00	±	0.00	0.04	±	0.02
Malate	28.80	±	2.72	132.14	±	24.27	134.09	4	93.24	93.90	50.14	±	2.24	54.32	±	4.08
Maleic_acid	0.09	±	0.03	2.09	±	0.85	138.10	4	95.22	93.55	0.16	±	0.07	0.72	±	0.18
Pyroglu	0.23	±	0.07	0.10	±	0.03	129.11	5	87.70	51.16	0.47	±	0.10	0.03	±	0.00
Succinate	3.23	±	0.92	11.75	±	3.92	118.09	4	91.09	93.63	6.20	±	1.48	5.10	±	0.59
Urea	0.00	±	0.00	0.00	±	0.00	60.06	1	0.00	0.00	0.00	±	0.00	0.00	±	0.00
C. Protein hydrolysate (amino acids):																
Ala	73.54	±	6.28	40.78	±	13.99	89.09	3	86.40	82.09	6.36	±	0.15	7.06	±	0.31
Arg	66.43	±	3.36	44.48	±	21.27	210.70	6	71.37	64.60	4.04	±	0.21	4.53	±	0.58
Asn	nd			nd			132.10	4	nd	nd	0.00	±	0.00	0.00	±	0.00
Asp	80.76	±	6.46	44.91	±	16.28	133.10	4	80.13	73.25	5.79	±	0.13	6.10	±	0.16
Cys	nd	±	nd	nd	±	nd	121.20	3	nd	nd	2.65	±	0.06	1.62	±	0.66
Glu	2.83	±	0.25	1.63	±	0.56	147.10	5	84.11	80.29	0.24	±	0.00	0.28	±	0.01
Gln	nd	±	nd	nd	±	nd	146.10	5	nd	nd	0.00	±	0.00	0.00	±	0.00

Gly	73.57	±	5.88	43.72	±	16.33	75.07	2	79.34	76.52	4.63	±	0.10	5.40	±	0.05
His	27.91	±	2.84	12.24	±	5.96	209.60	6	75.53	70.28	1.79	±	0.00	1.33	±	0.23
Iso	39.07	±	4.22	21.31	±	6.22	131.20	6	50.89	46.02	2.69	±	0.09	2.91	±	0.28
Leu	83.71	±	7.88	44.57	±	13.73	131.20	6	67.57	64.11	7.68	±	0.21	8.44	±	0.68
Lys	192.89	±	25.96	98.89	±	43.55	182.60	6	70.87	63.67	13.23	±	0.53	11.86	±	0.94
Met	36.78	±	3.22	14.79	±	8.15	149.20	5	72.40	68.17	2.65	±	0.06	1.62	±	0.66
Phe	54.63	±	5.24	30.04	±	9.77	165.20	9	75.67	70.30	6.68	±	0.20	7.31	±	0.44
Pro	114.56	±	17.37	51.79	±	20.51	115.10	5	74.80	68.45	10.92	±	0.55	8.98	±	0.79
Ser	109.20	±	10.55	59.71	±	22.65	105.10	3	83.52	78.08	7.72	±	0.14	8.24	±	0.57
Thr	45.58	±	6.54	24.89	±	9.11	119.10	4	72.61	66.53	3.30	±	0.34	3.56	±	0.47
Trp	nd	±	nd	nd	±	nd	204.20	11	nd	nd	6.68	±	0.20	7.31	±	0.44
Tyr	63.08	±	6.53	31.75	±	11.64	181.20	9	77.29	71.16	7.17	±	0.07	6.86	±	0.06
Val	59.91	±	5.56	33.51	±	9.94	117.10	5	76.27	70.63	5.79	±	0.08	6.59	±	0.61

Note:

***** MW is the molecular weight of the commercial samples used for standard curves in GC-MS.

****** Extent of average labelling in amino acids (cationic fraction) of N40 is assumed to be equal to MS, as they were not detectable. Anyway the influence of this on the relative fractions of metabolites in cationic extracts of N40 is negligible due to the very low radioactivity detected in them.

nd : not detected in GC-MS. Cys and Trp are lost during acid hydrolysis of protein. Gln and Asn gets converted to Glu and Asp respectively. While assigning the radioactivity - Trp and Cys are assumed to be equal proportions to Phe and Met respectively.

Relative carbon atom (%) calculation: The labeled carbon atom abundance of each metabolite is calculated by the equation (Abundance/MW)*No of Carbons*extent of label. The relative carbon atoms were obtained from the relative proportions of the labeled carbon atom abundance. The extent of label in each metabolite is used as the initial abundance measurement could be influenced by unlabelled pre-existing pools.

Table 5.2C: Fluxes to biosynthetic outputs presented as A) proportional carbon flux and (B) molar output flux relative to glucose uptake rate of 1. The proportional carbon flux is based on radioactive proportions presented in Table 5.2A. The molar output fluxes were calculated as noted below and were used as constraints in the parameter fitting of the models using 13CFLUX (section 5.2.8)

	A) proportion carbon flux from glucose					B) Molar output flux (Relative to Glucose uptake =1) [(Prop. carbon flux * 6)/no. of C atoms]				
Metabolite	MS		N40		No of C atoms	Output fluxes	MS		N40	
14CO2	0.450	± 0.003	0.310	± 0.012	1	CO_2[#]	2.700	± 0.019	1.859	± 0.070
Fructose	0.040	± 0.004	0.091	± 0.003	6	fruOUT	*		*	
Glucose	0.040	± 0.004	0.091	± 0.003	6	gluOUT	*		*	
Sucrose-glucosyl	0.039	± 0.005	0.089	± 0.004	6	suc(g)OUT	0.079	± 0.009	0.180	± 0.007
Sucrose-fructosyl	0.039	± 0.005	0.089	± 0.004	6	suc(f)OUT	0.079	± 0.009	0.180	± 0.007
Cellwall C6	0.029	± 0.006	0.041	± 0.003	6	cellwall(h)OUT	0.029	± 0.006	0.041	± 0.003
Cellwall C5	0.006	± 0.001	0.009	± 0.001	5	cellwall(p)OUT	0.008	± 0.001	0.011	± 0.001
Starch	0.074	± 0.004	0.075	± 0.002	6	starchOUT	0.074	± 0.004	0.075	± 0.002
Lipid-glycerol[@]	0.002	± 0.000	0.001	± 0.000	3	glycerolOUT	0.003	± 0.000	0.002	± 0.000
Lipid-FA[@]	0.018	± 0.002	0.009	± 0.002	2	lipidOUT	0.055	± 0.007	0.027	± 0.007
Ethanol	0.001	± 0.000	0.001	± 0.000	2	ethanolOUT	0.004	± 0.000	0.002	± 0.000
Citrate	0.029	± 0.005	0.048	± 0.014	6	citOUT	0.029	± 0.005	0.048	± 0.014
Malate	0.033	± 0.002	0.065	± 0.006	4	malOUT	0.050	± 0.003	0.097	± 0.009
Succinate	0.004	± 0.002	0.006	± 0.001	4	SuccOUT	0.006	± 0.002	0.009	± 0.002
Ala	0.018	± 0.006	0.005	± 0.000	3	alaOUT	0.035	± 0.013	0.011	± 0.000
Arg	0.003	± 0.000	0.003	± 0.001	6	argOUT	0.003	± 0.000	0.003	± 0.001
Asx	0.006	± 0.000	0.004	± 0.000	4	aspOUT	0.009	± 0.001	0.006	± 0.000
Cys	0.002	± 0.000	0.001	± 0.001	3	cysOUT	0.004	± 0.000	0.002	± 0.001
Glx	0.064	± 0.025	0.011	± 0.005	5	gluOUT	0.077	± 0.029	0.014	± 0.006
Gly	0.004	± 0.000	0.004	± 0.000	2	glyOUT	0.013	± 0.001	0.011	± 0.000

His	0.001	±	0.000	0.001	±	0.000	6	hisOUT	0.001	±	0.000	0.001	±	0.000
Iso	0.002	±	0.000	0.002	±	0.000	6	ileOUT	0.002	±	0.000	0.002	±	0.000
Leu	0.006	±	0.000	0.005	±	0.001	6	leuOUT	0.006	±	0.000	0.005	±	0.001
Lys	0.037	±	0.013	0.007	±	0.001	6	lysOUT	0.037	±	0.013	0.007	±	0.001
Met	0.002	±	0.000	0.001	±	0.001	5	metOUT	0.003	±	0.000	0.001	±	0.001
Phe	0.006	±	0.000	0.005	±	0.000	9	pheOUT	0.004	±	0.000	0.003	±	0.000
Pro	0.010	±	0.001	0.006	±	0.001	5	proOUT	0.012	±	0.001	0.007	±	0.001
Ser	0.013	±	0.003	0.005	±	0.001	3	serOUT	0.026	±	0.005	0.011	±	0.001
Thr	0.003	±	0.000	0.002	±	0.001	4	thrOUT	0.004	±	0.001	0.003	±	0.001
Trp	0.005	±	0.000	0.004	±	0.000	11	trpOUT	0.003	±	0.000	0.002	±	0.000
Tyr	0.006	±	0.000	0.004	±	0.000	9	tyrOUT	0.004	±	0.000	0.003	±	0.000
Val	0.005	±	0.000	0.004	±	0.001	5	valOUT	0.006	±	0.000	0.005	±	0.001
Gaba	0.002	±	0.003	0.000	±	0.000	4	gabaOUT	0.003	±	0.004	0.000	±	0.000

Note:

@ Assumed that the bulk of lipid is phospholipid with 35C in fatty acids per glycerol backbone. The total radioactivity in lipid was divided into glycerol (no of C = 3) and fatty acid (FA, no of C = 35) proportions based on the calculation; (Total radioactivity of lipids * No of C)/38.

CO_2 efflux was not used in the model as it was allowed to be free to be determined by the model. The relative fluxes through other metabolites serve as constraints while parameter fitting of the models in 13C-FLUX

* Used with sucrose output fluxes (suc(g)OUT or suc(f)OUT) as they are assumed to be produced from sucrose degradation. Also they along with sucrose moieties come out as output fluxes from the same precursor. The glucose and fructose were measured from TLC and the possible contamination of external glucose was corrected (see Fig 5.12).

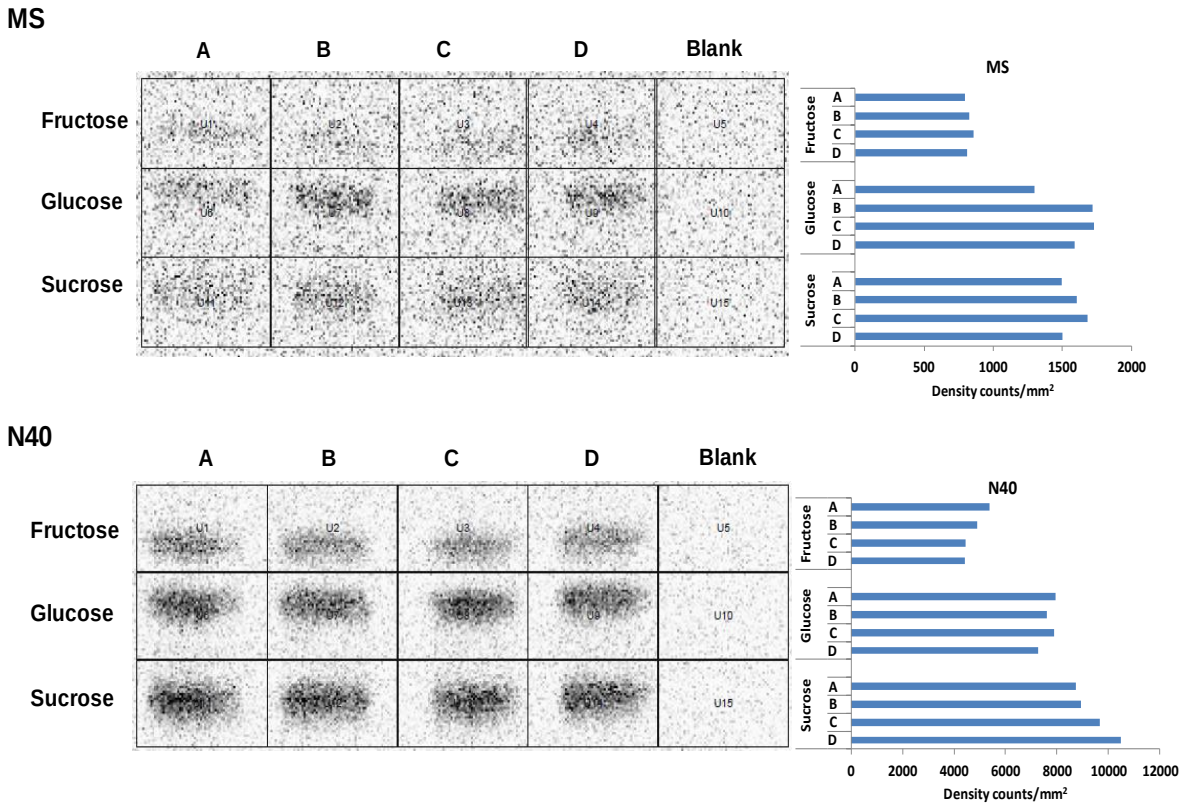


Figure 5.12 Phosphor-imaging of TLC plates showing bands of relative proportions of sugars from neutral fraction. Bands of glucose, fructose and sucrose were clearly distinguished from the raw 1-D images and their corresponding intensities (density count/mm^2) were measured using the software Quantity One (version 4.4.1, Biorad). The glucose band is due to the *de novo* synthesized plus un-metabolised glucose that was taken up by the cultures from the medium. *De novo* synthesised fructose and glucose were considered to be from sucrose degradation and to be equal, thereby allowing to correct for the contaminating glucose. The relative abundances of the sugars were normalised to the total radioactivity of the neutral fraction (see Table 5.2A for radioactivity of metabolized sugars). Table 5.2C presents the sucrose output flux (7.9% in MS and 18 % in N40) relative to glucose uptake. GC-MS of neutral fractions (Fig 5.19) from unlabelled samples showed that glucose, fructose and sucrose constitute > 99% with traces of metabolically generated myoinisitol and glycerol which are not considered in the current analysis.

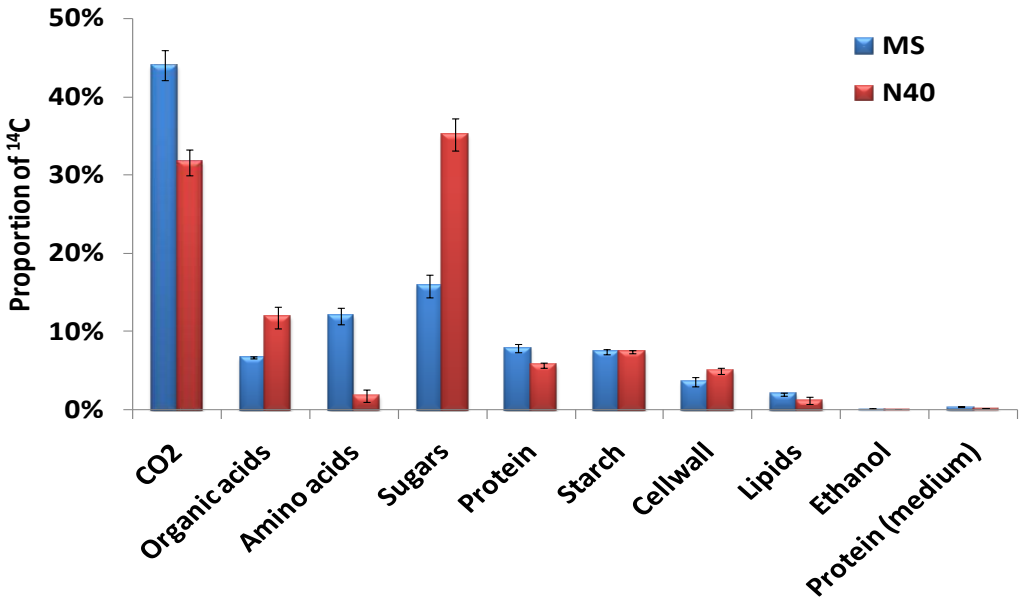


Figure 5.13 Proportion of ^{14}C metabolized into each fraction in *Arabidopsis* suspension cultures grown under different N sources (MS and N40). Significant differences were observed in CO_2 release, organic acids, amino acids, sugars, protein, cell wall and lipid. The values are obtained from Table 5.2A and are presented as mean proportions of radioactivity in relation to metabolized glucose ± sd derived from 4 separate flasks

Table 5.3 Summary of feeding strategies and measurements made with NMR and GC-MS. Representative spectra are presented in figures as mentioned.				
Treatment 1	MS	MS	MS	MS
Treatment 2	N40	N40	N40	N40
Choice of substrates	99%[1-^{13}C] glucose	99%[2-^{13}C] glucose	20%[$^{13}C_6$] glucose	Unlabelled glucose
No. of days grown	5.5	5.5	5.5	5.5
No. of replicates	3	3	3	3
^{13}C-NMR (Positional isotopomers)				
Solubles	Fig 5.12	Fig 5.13	Fig 5.14	-
Starch	Fig 5.15	Fig 5.15	-	-
GC-MS (Mass isotopomers)				
Amino acids	Fig 5.7			
Organic acids	Fig 5.16			
Sugars	Fig 5.17			
Protein	Fig 5.18			
Cell wall	Fig 5.19			
Starch	Fig 5.20			

[1-^{13}C]glucose fed – MS vs N40

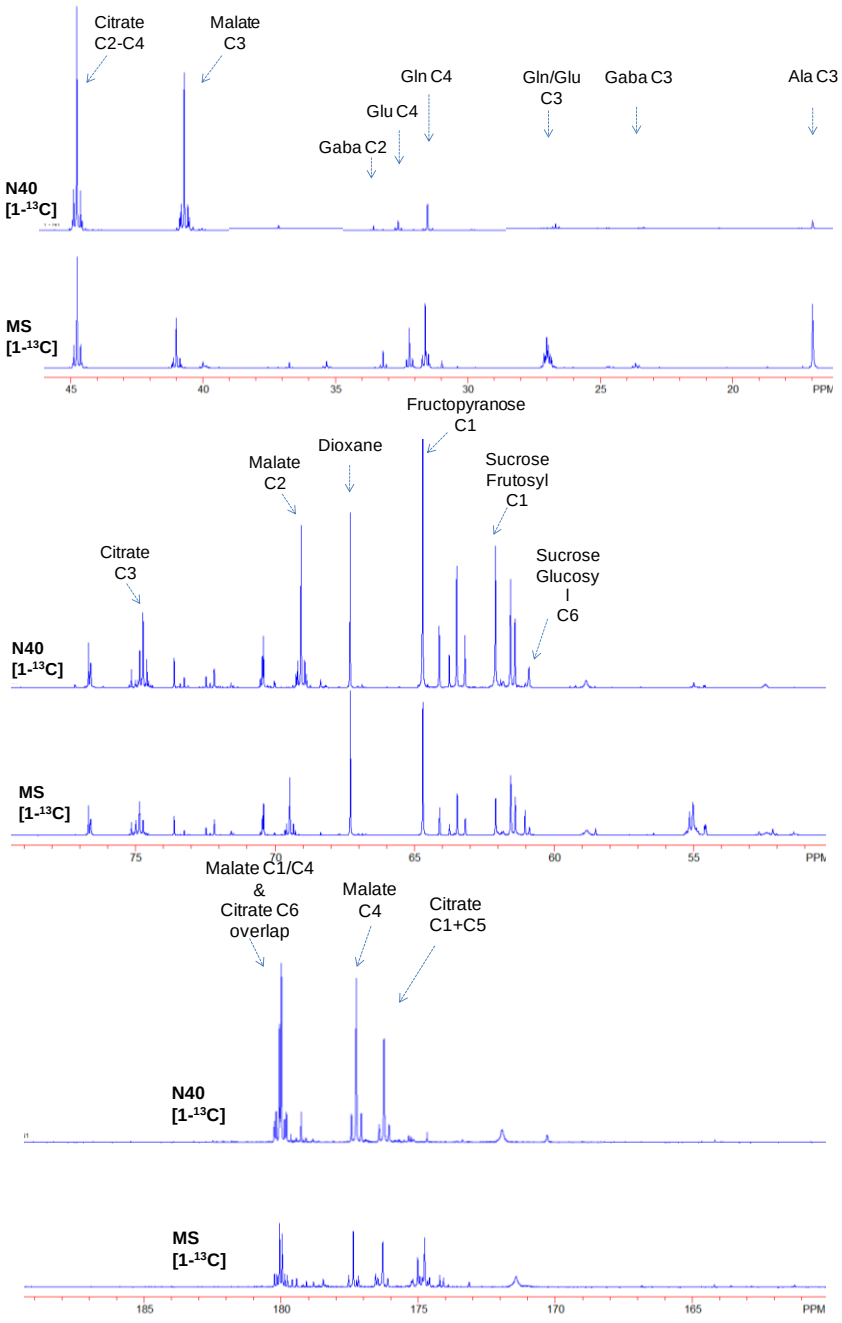


Figure 5.14 ^{13}C NMR of perchloric acid extracts obtained from *Arabidopsis* cell suspension cultures fed with 99% [1-^{13}C]glucose under normal (MS) and nitrate-only (N40) conditions. Three selected regions from the whole spectra are presented. The peak abundances and the errors were measured as detailed in Chapter 3.

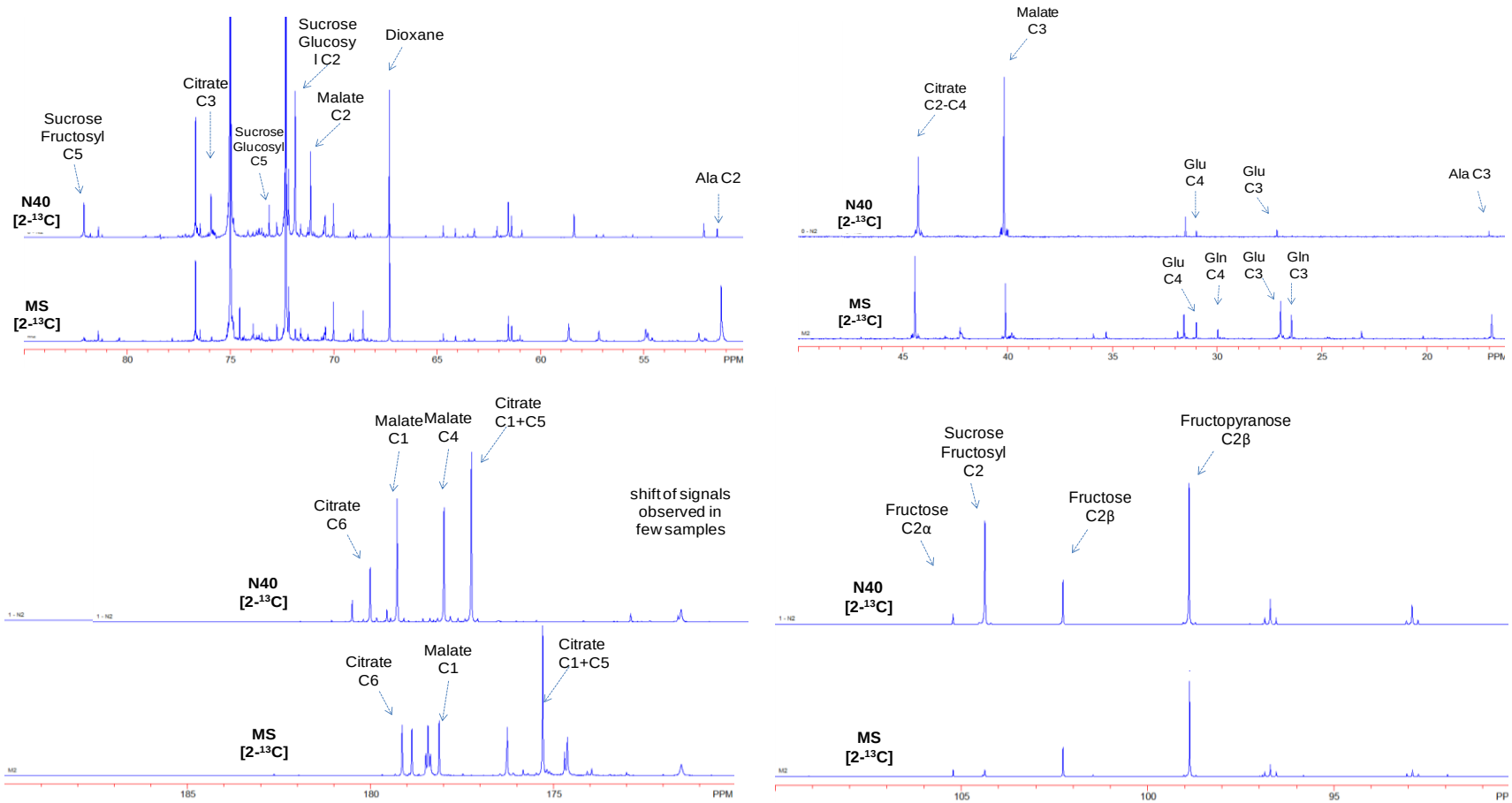


Figure 5.15 ^{13}C NMR of perchloric acid extracts obtained from *Arabidopsis* cell suspension cultures fed with 99% [2-^{13}C]glucose under normal (MS) and nitrate-only (N40) conditions. Four selected regions from the whole spectra are presented. The peak abundances and the errors are measured as detailed in Chapter 3.

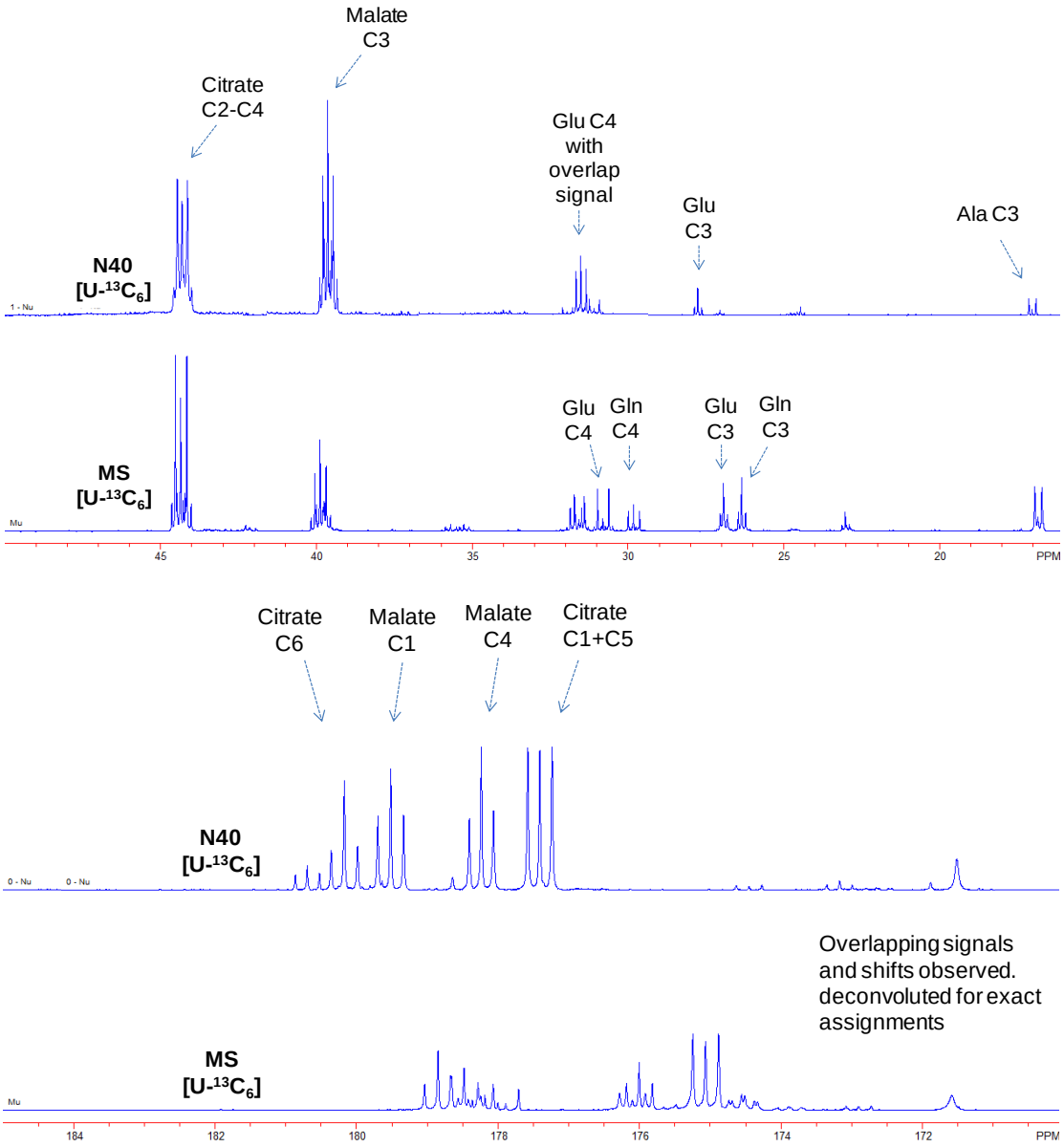


Figure 5.16 ^{13}C NMR of perchloric acid extracts obtained from *Arabidopsis* cell suspension cultures fed with 20% [$^{13}C_6$]glucose under normal (MS) and nitrate-only (N40) conditions. Two selected regions from the whole spectra are presented. The peak abundances and the errors are measured as detailed in Chapter 3.

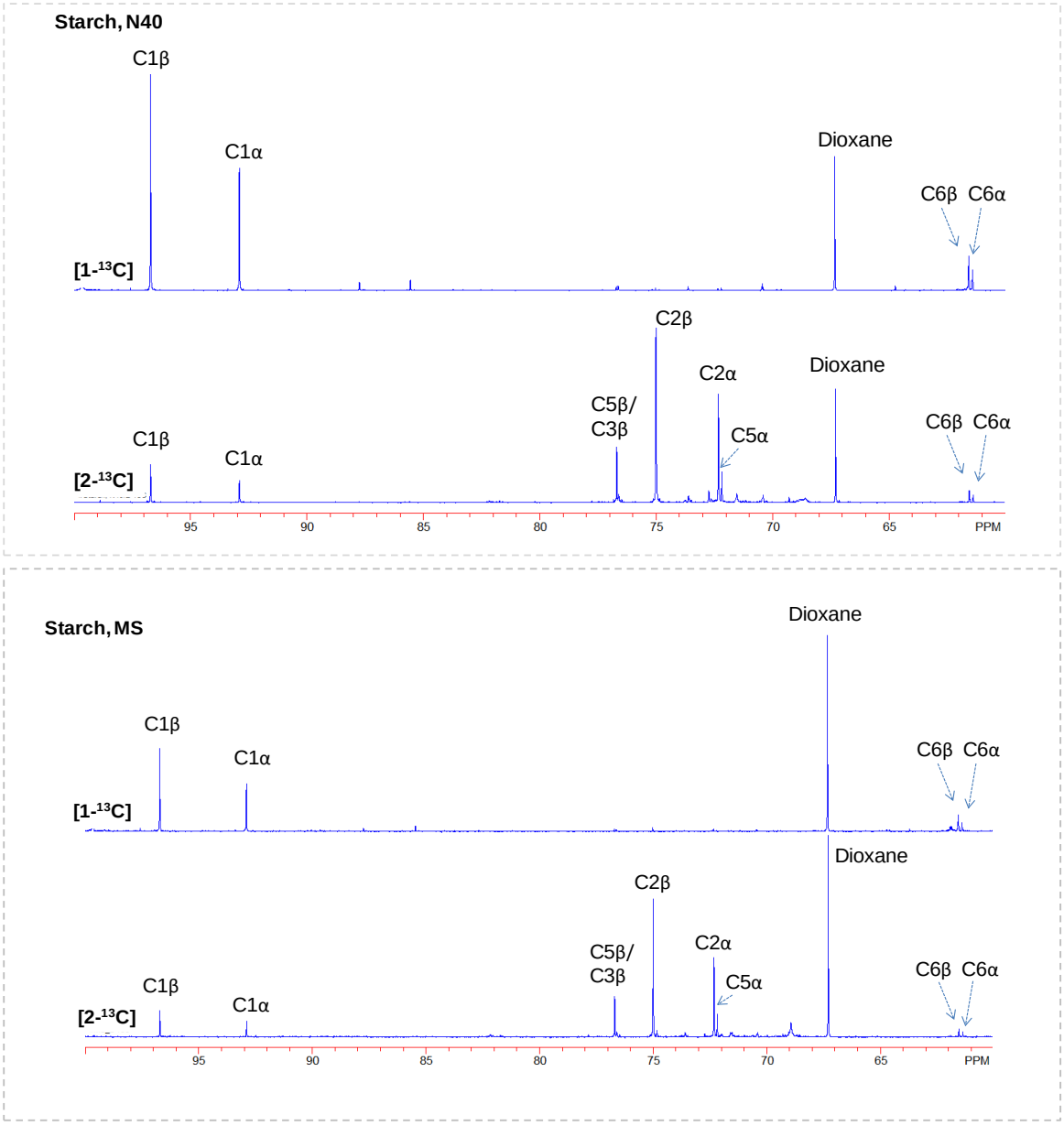


Figure 5.17 ^{13}C NMR of enzymatically digested starch from MS and N40 samples showing the fractional abundances of glucose carbons. The starch was obtained from [1-^{13}C]glucose and [2-^{13}C]glucose feeding experiments on *Arabidopsis* cell suspension cultures. The peak abundances and the errors are measured as detailed in Chapter 3.

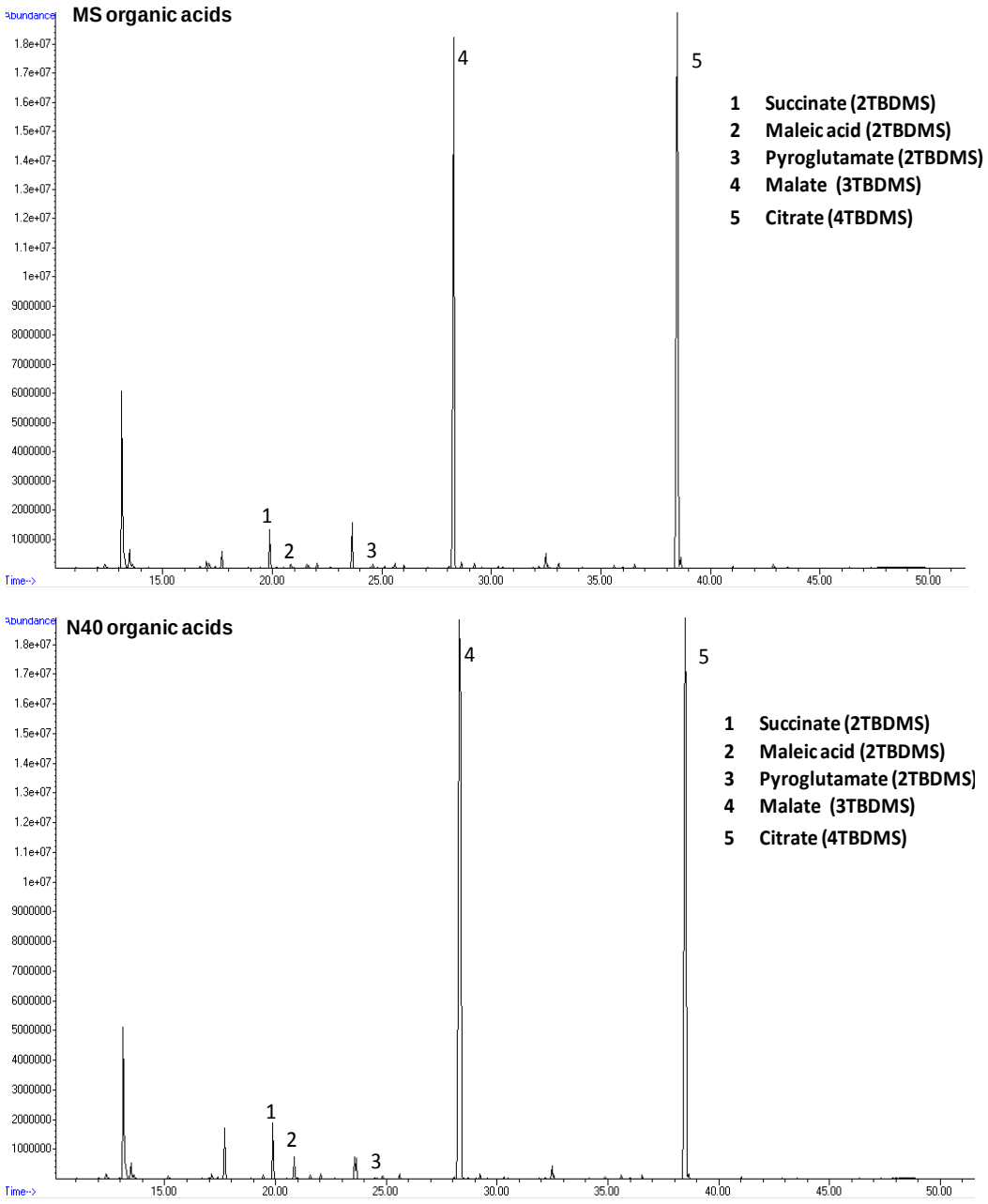


Figure 5.18 Representative total ion chromatograms (from GC-MS) of the anionic fraction containing organic acids obtained from *Arabidopsis* cell suspension cultures under normal (MS) and nitrate-only (N40) conditions. The samples were TBDMS derivatised. The mass isotopomer fragments from each metabolite were measured and handled as detailed in Chapter 3.

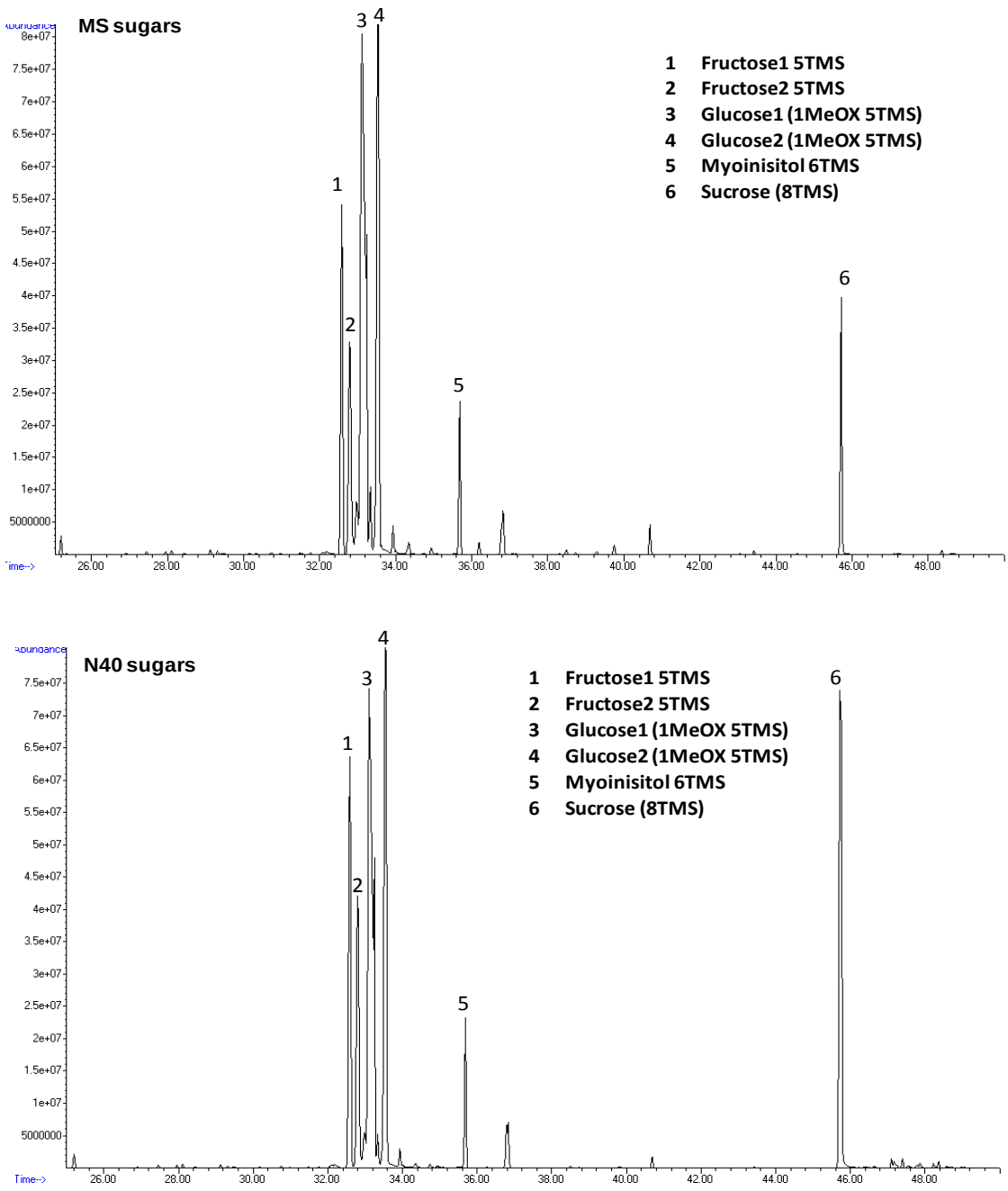


Figure 5.19 Representative total ion chromatograms (from GC-MS) of the neutral fraction containing sugars obtained from *Arabidopsis* cell suspension cultures under normal (MS) and nitrate-only (N40) conditions. The sugars were MeOX-TMS derivatised. The mass isotopomer fragments from each metabolite were measured and handled as detailed in Chapter 3.

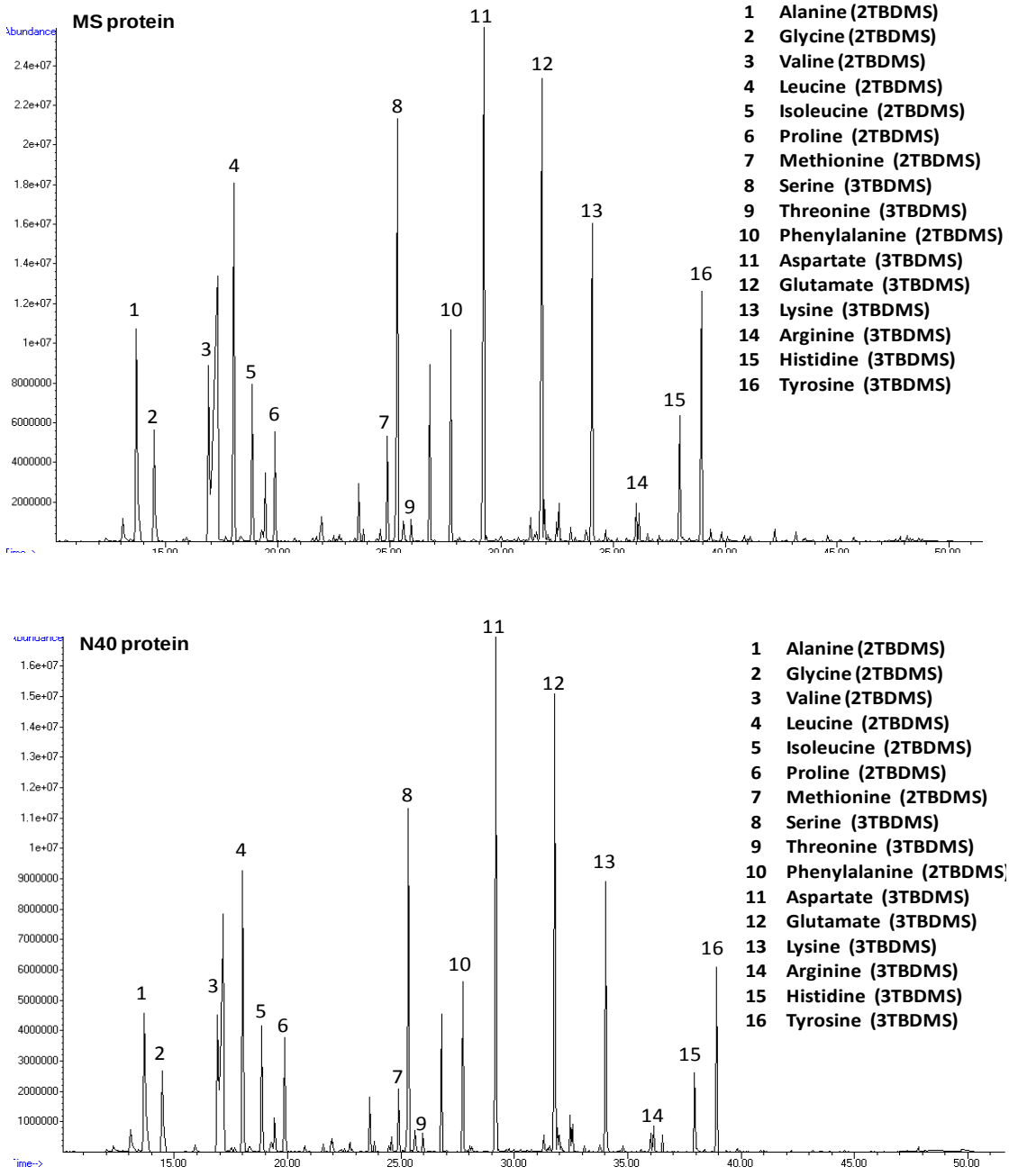


Figure 5.20 Representative total ion chromatograms (from GC-MS) of protein hydrolysate obtained from *Arabidopsis* cell suspension cultures under normal (MS) and nitrate-only (N40) conditions. The aminoacids were TBDMS derivatised. The mass isotopomer fragments from each metabolite were measured and handled as detailed in Chapter 3.

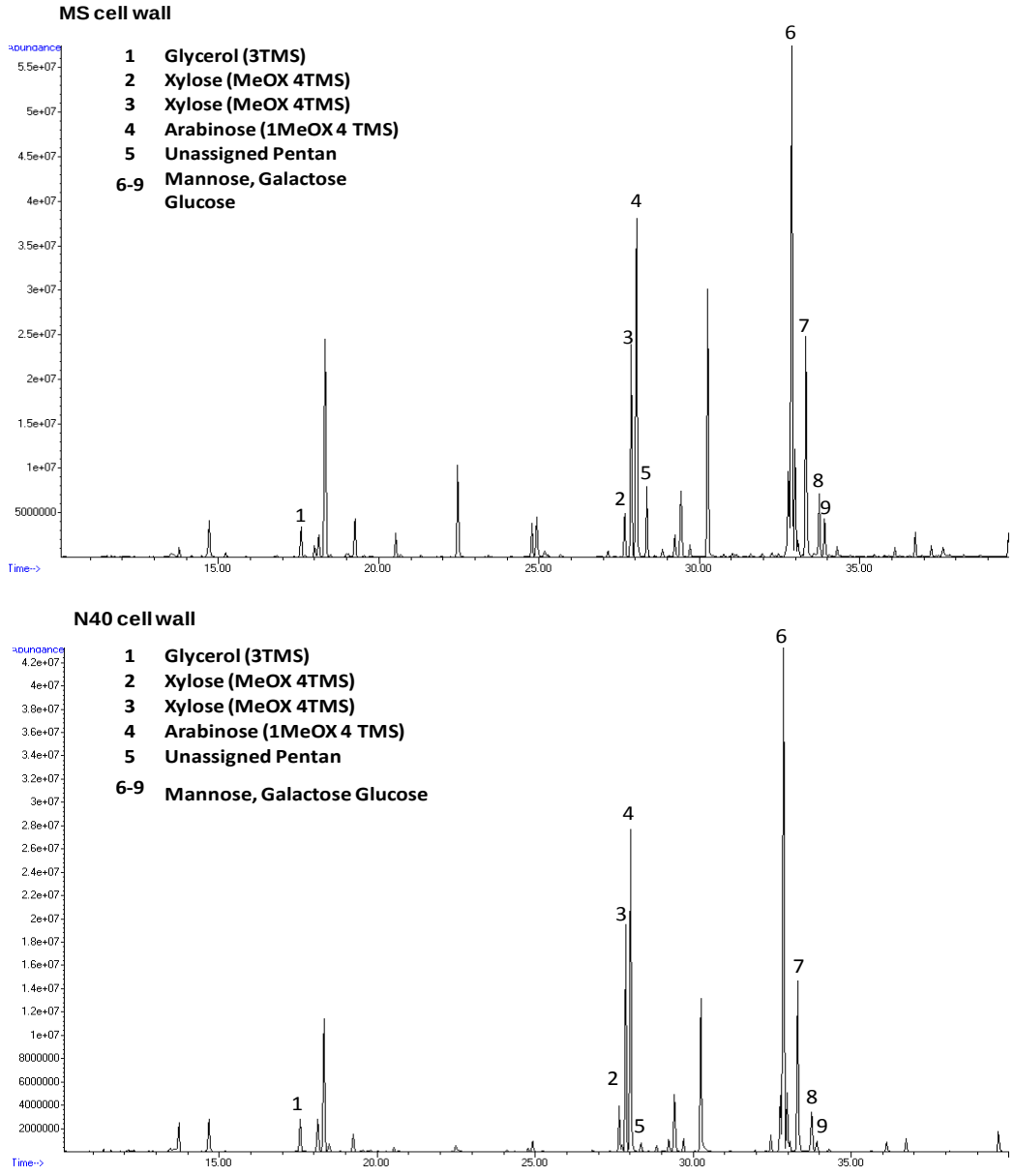


Figure 5.21 Representative total ion chromatograms (from GC-MS) of cell wall hydrolysate obtained from *Arabidopsis* cell suspension cultures under normal (MS) and nitrate-only (N40) conditions. The samples were MeOX-TMS derivatised. The mass isotopomer fragments from each metabolite were measured and handled as detailed in Chapter 3.

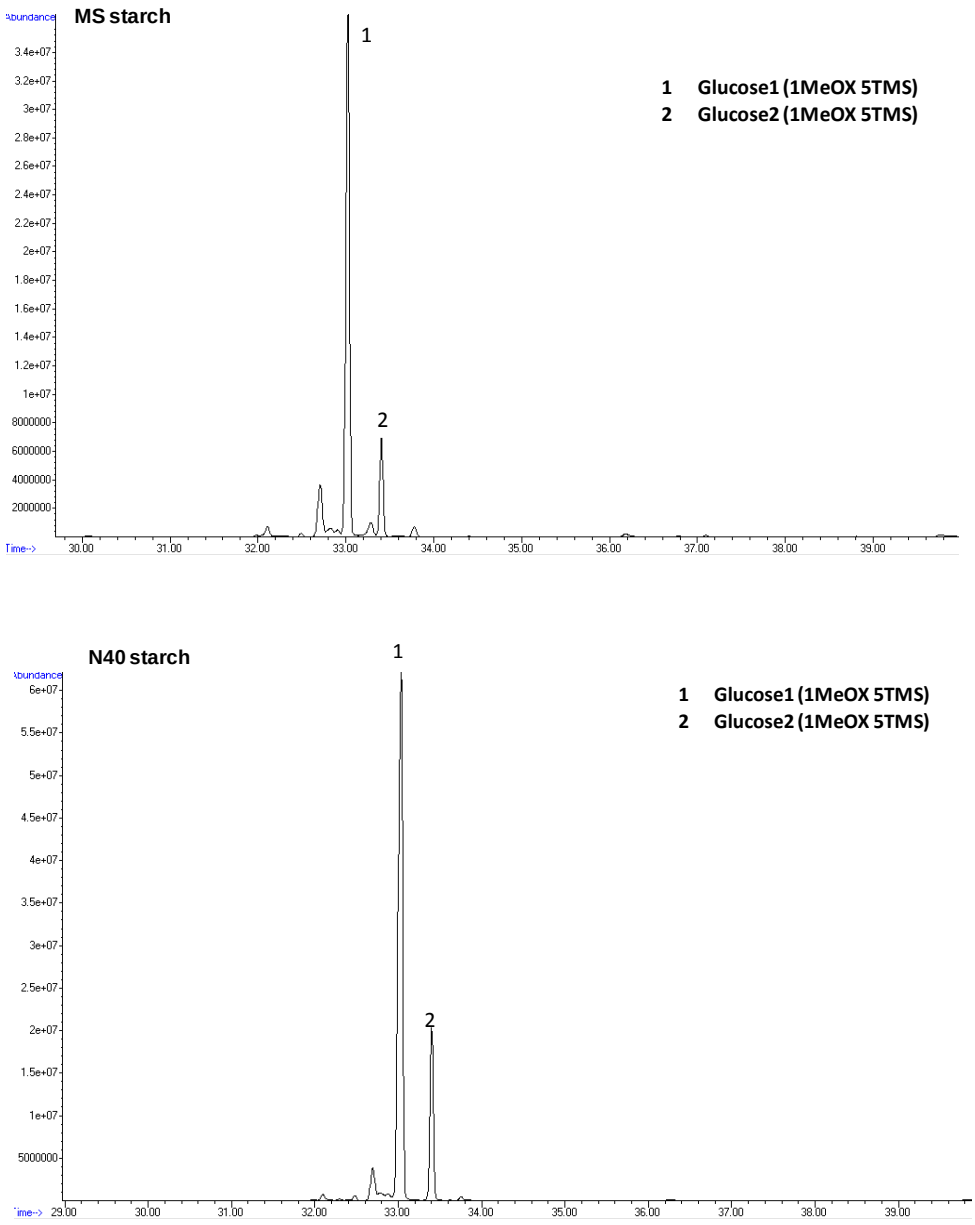


Figure 5.22 Representative total ion chromatograms (from GC-MS) of enzymatically digested starch containing glucose. Starch was extracted from *Arabidopsis* cell suspension cultures grown under normal (MS) and nitrate-only (N40) conditions. The samples were MeOX-TMS derivatised. The mass isotopomer fragments from glucose were measured and handled as detailed in Chapter 3.

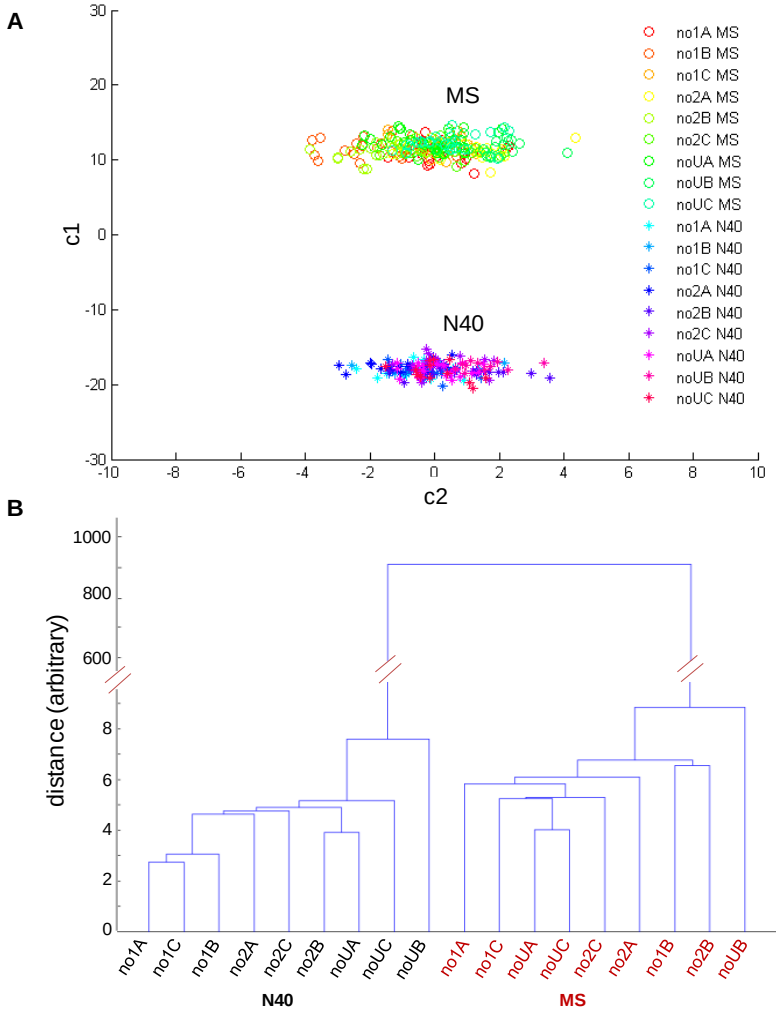


Figure 5.23 Consistency checking of the replicate datasets by investigating the flux estimates obtained by Jack-knife method. Each individual dataset is removed (one at a time) and Monte-Carlo flux solutions obtained were subjected to canonical clustering analysis using MANOVA in matlab. The c1 and c2 in (**A**) are the canonicals which that explain the maximum differences between the groups. All the solutions from either MS or N40 models cluster together implying the consistencies in datasets. Significant differences were observed between MS and N40. The tree below (B) points to the level of differences within and between the groups.

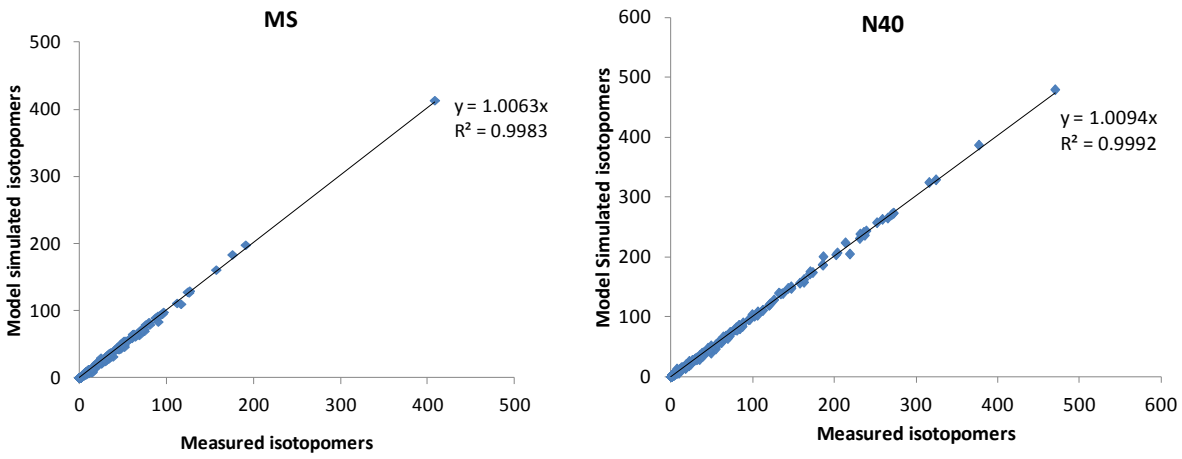


Figure 5.24 Measured Vs Predicted isotopomer abundances of MS and N40 best fit model. The linear regression ($R^2 > 0.99$) showed good agreement between the measured and model predicted values.

Table 5.4: Statistical performance of the best fit models of MS and N40		
	A) MS	**B) N40**
Type of network	Triple-model	Triple-model
Experimental datasets	99%[1-^{13}C]glucose(3 replicas) 99%[2-^{13}C]glucose(3 replicas) 20%[$^{13}C_6$]glucose(3 replicas)	99%[1-^{13}C]glucose(3 replicas) 99%[2-^{13}C]glucose(3 replicas) 20%[$^{13}C_6$]glucose(3 replicas)
Residuum of global best fit solution	1332.33	1778.76
No. of Free fluxes	30	30
No. of measurements (NMR + GCMS)	2390	2540
No of metabolite fragments (GC-MS)/sub-groups (NMR)	530	522
Degree of freedom	1830	1988
Chi^2 (at 95% significance level)	1930.63	2092.84
Validation	Passed (Residuum < Chi2)	Passed (Residuum < Chi2)

Table 5.5 The best fit flux estimates of *Arabidopsis* cell suspension culture grown under normal MS medium (~40 mM NO_3 and ~20 mM NH_4-) and N40 (40 mM NO_3) conditions. The 68% confidence limits (sd) of all the fluxes obtained by *EstimateStat* and the free fluxes further verified by non-linearised statistics (95% CI) are presented. Significant differences in the TCA cycle fluxes and net PPP flux (cppp1+pppp1) were observed in Arabidopsis flux phenotypes under different N sources. All fluxes are relative to glucose uptake of 1± 0.05. "nd" is non determinable flux.

		MS				**N40**			
NET fluxes	**Flux name**	**Best fit flux**	sd (68% linear approx)	95% CI of Free fluxes (Non-linear)		**Best fit flux**	sd (68% linear approx)	95% CI of Free fluxes (Non-linear)	
				min	max			min	max
Hexose/triose phosphate metabolism									
Glucose uptake (Constrained)	*upt*	1	0.05			1	0.05		
Cytosolic phosphoglucose isomerase	*chex1*	-0.161	0.041	-0.245	-0.09	-0.141	0.026	-0.19	-0.10
Cytosolic Fru-6-P → triose-P/PEP interconversion	*chex2*	-0.24	0.041			-0.322	0.026		
Cytosolic pyruvate kinase	*chex3*	0.925	0.042			0.606	0.018		
Plastidic phosphoglucose isomerase	*phex1*	0.679	0.038			0.517	0.025		
Plastidic Fru-6-P → triose-P/PEP interconversion	*phex2*	0.868	0.039			0.723	0.025		
Plastidic pyruvate kinase	*phex3*	0.084	0.038	0.055	0.086	0.054	0.013	0.036	0.056
Pentose phosphate pathway									
Cytosolic pentose phosphate pathway - oxidative steps	*cppp1*	0.078	0.05	0	0.164	0	0.032	0	0.024
Plastidic pentose phosphate pathway - oxidative steps	*pppp1*	0.215	0.063	0.149	0.315	0.317	0.026	0.284	0.326
Plastidic transketolase - C7/C5 conversion	*pppp2a*	0.1	0.005			0.107	0.003		
Plastidic transketolase - C6/C4 conversion	*pppp2b*	0.089	0.005			0.099	0.003		

Plastidic transketolase - C5/C3 conversion	*pppp2c*	0.189	0.01			0.206	0.006		
Plastidic transaldolase - C7/C4 conversion	*pppp3a*	0.1	0.005			0.107	0.003		
Plastidic transaldolase - C6/C3 conversion	*pppp3b*	0.1	0.005			0.107	0.003		
Transporters/exchanges									
Pyruvate uptake by mitochondria	*cmex*	0.885	0.012			0.591	0		
Pyruvate uptake by plastids	*cpex*	0	0.041	0	0.03	0.003	0.013	0	0.0205
Plastidic Glc-6-P/Pi translocator	*gpt*	0.968	0.071			0.909	0.042		
Plastidic triose-P/Pi & PEP/Pi translocators (Net direction is from Plastid to Cytosol)	*tpt*	1.681	0.093			1.45	0.053		
Tricarboxylic acid cycle									
Mitochondrial pyruvate dehydrogenase	*tca1*	0.948	0.005			0.591	0.003		
Citrate synthase	*tca2*	0.948	0.005			0.591	0.003		
Aconitase/isocitrate dehydrogenase	*tca3*	0.92	0.005			0.543	0.003		
2-oxoglutarate → Succinate	*tca4*	0.824	0.005			0.522	0.003		
Succinate → fumarate convertion	*tca5*	0.818	0.005			0.513	0.003		
Fumarase	*tca6a+b*	0.822	0.005			0.513	0.003		
Anaplerotic fluxes									
PEP carboxylase	*ana1*	0.272	0.015			0.199	0.011		
Mitochodrial malic enzyme	*ana2*	0.063	0.014	0.04	0.09	0	0.011	0	0.013
Plastidic malic enzyme	*ana3*	0.008	0.002			0.004	0.001		
XCH[0-1] fluxes		**MS**	sd			**N40**	sd		
Cytosolic phosphoglucose isomerase	*chex1*	0.916	0.014	0.887	0.945	0.973	0.016	0.945	0.99
Cytosolic Fru-6-P → triose-P/PEP interconversion	*chex2*	0.314	0.02	0.269	0.351	0.206	0.019	0.172	0.236
Plastidic phosphoglucose isomerase	*phex1*	0.701	0.064	0.581	0.829	0.65	0.032	0.596	0.707

Plastidic transketolase - C7/C5 conversion	*pppp2a*	0.32	nd	0.118	1	0.353	nd	nd	nd
Plastidic transketolase - C6/C4 conversion	*pppp2b*	0.177	0.015	0.148	0.212	0.344	0.018	0.31	0.382
Plastidic transketolase - C5/C3 conversion	*pppp2c*	0.307	0.016	0.262	0.341	0.274	0.014	0.248	0.301
Plastidic transaldolase - C7/C4 conversion	*pppp3a*	0.156	nd			0.325	nd	nd	nd
Plastidic transaldolase - C6/C3 conversion	*pppp3b*	0.096	0.03	0	0.21	0	0.012	0	0
Plastidic Glc-6-P/Pi translocator	*gpt*	0	0.073			0	0.076		
Plastidic triose-P/Pi & PEP/Pi translocators	*tpt*	0.146	0.283	0	0.53	0	0.29	0	0
Fumarase	*tca6*	0.966	0.06	0.85	0.99	0.99	0.107	0.934	0.99
PEP carboxylase (Exchange is for PEP carboxykinase)	*ana1*	0.015	0.014	0	0.042	0	0.012	0	0

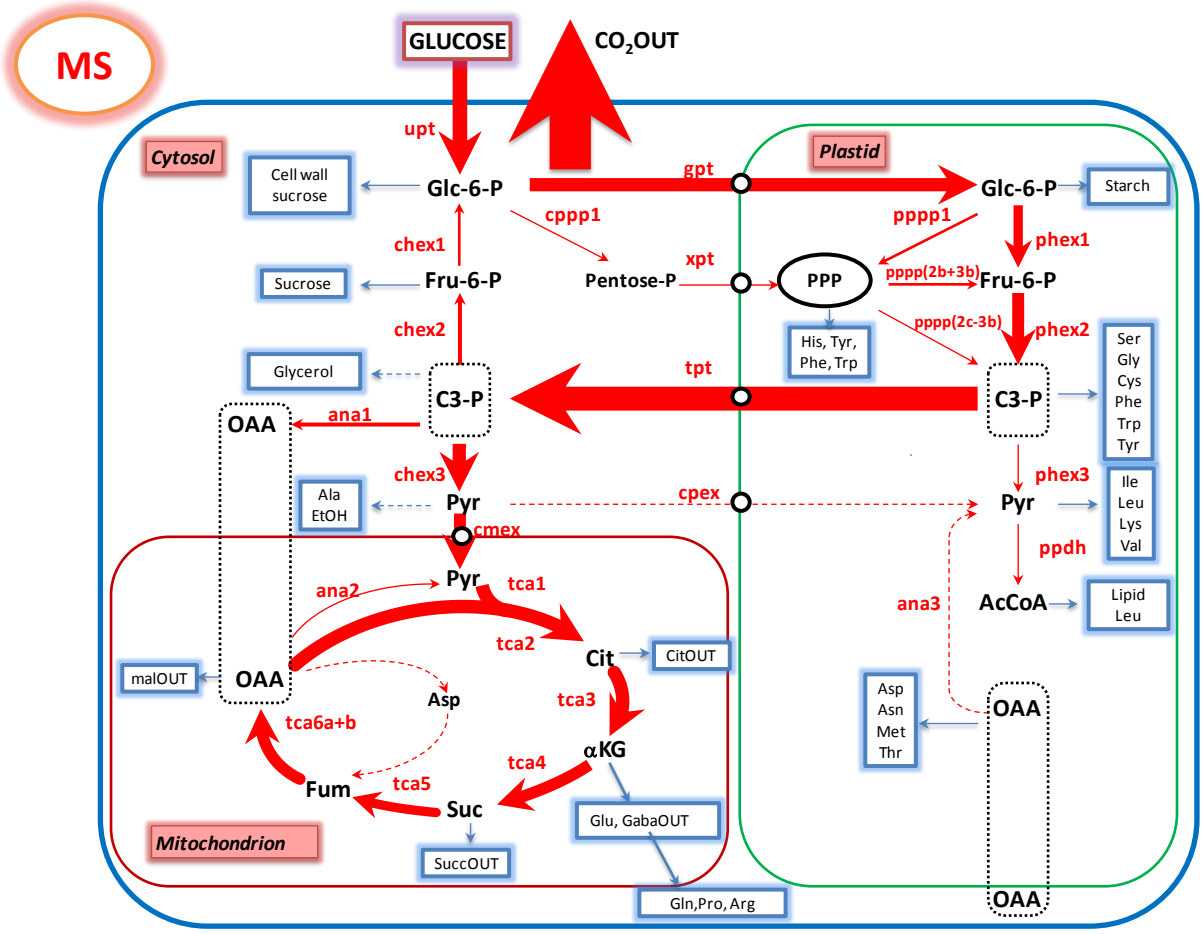


Figure 5.25 Flux map of *Arabidopsis* cell suspension culture grown under normal MS medium with ~40 mM NO_3 and ~20 mM NH_4. The best fit net fluxes presented here as the thickness of arrow are estimated from parameter fitting of triple model with labelling data from three separate feeding experiments of [1-^{13}C]glucose, [2-^{13}C]glucose and [$^{13}C_6$]glucose and biomass constraints derived from [U-^{14}C]glucose feeding. All fluxes are relative to glucose uptake. The dotted lines indicate < 5% of flux relative to glucose uptake. The net fluxes and the XCH[0-1] with their confidence limits are presented in Table 5.5 and Figure 5.27.

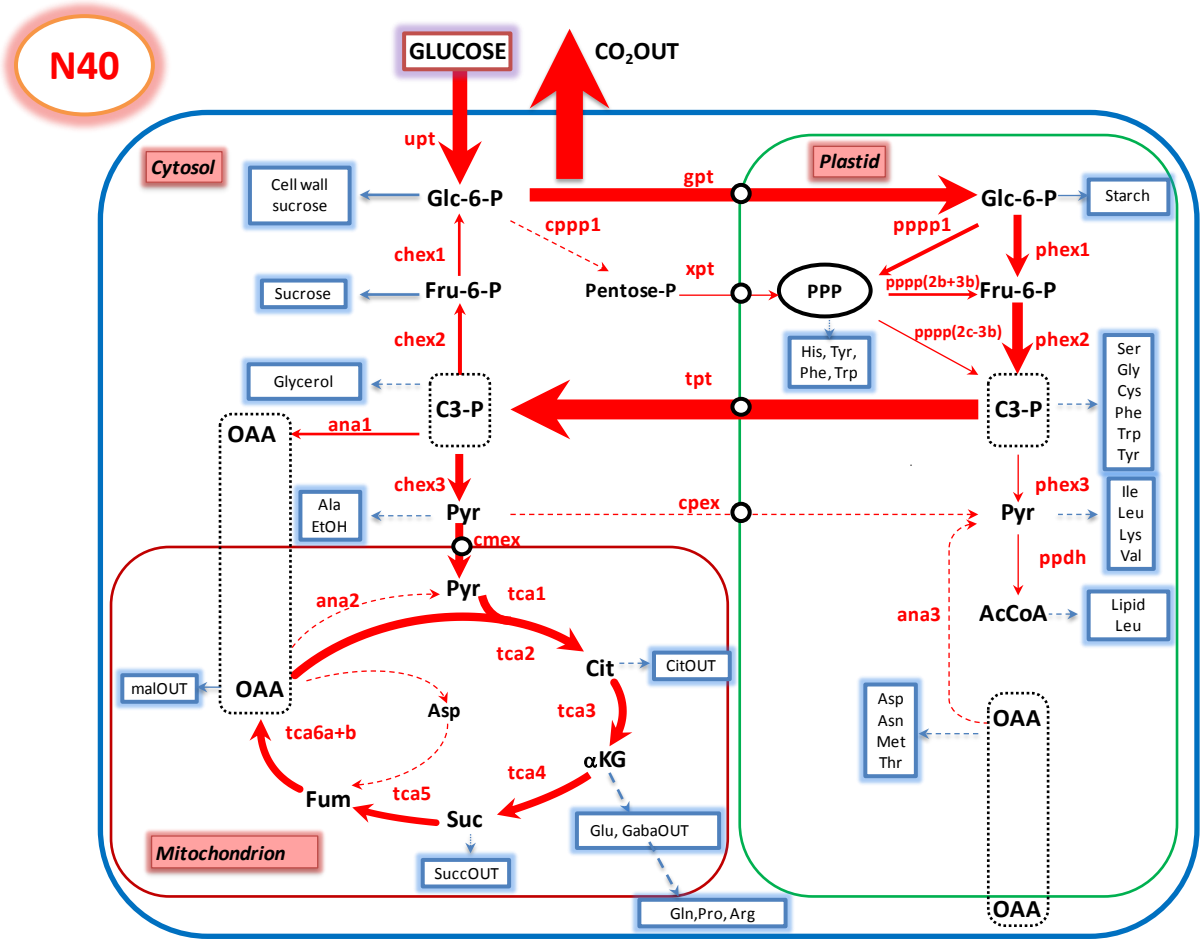


Figure 5.26 Flux map of *Arabidopsis* cell suspension culture grown under N40 medium (nitrate only with 40 mM NO_3). The best fit net fluxes presented here as the thickness of arrow are estimated from parameter fitting of triple model with labelling data from three separate feeding experiments of [1-^{13}C]glucose, [2-^{13}C]glucose and [$^{13}C_6$]glucose and biomass constraints derived from [U-^{14}C]glucose feeding. All fluxes are relative to glucose uptake. The dotted lines indicate < 5% of flux relative to glucose uptake. The net fluxes and the XCH[0-1] with their confidence limits are presented in Table 5.5 and Figure 5.27.

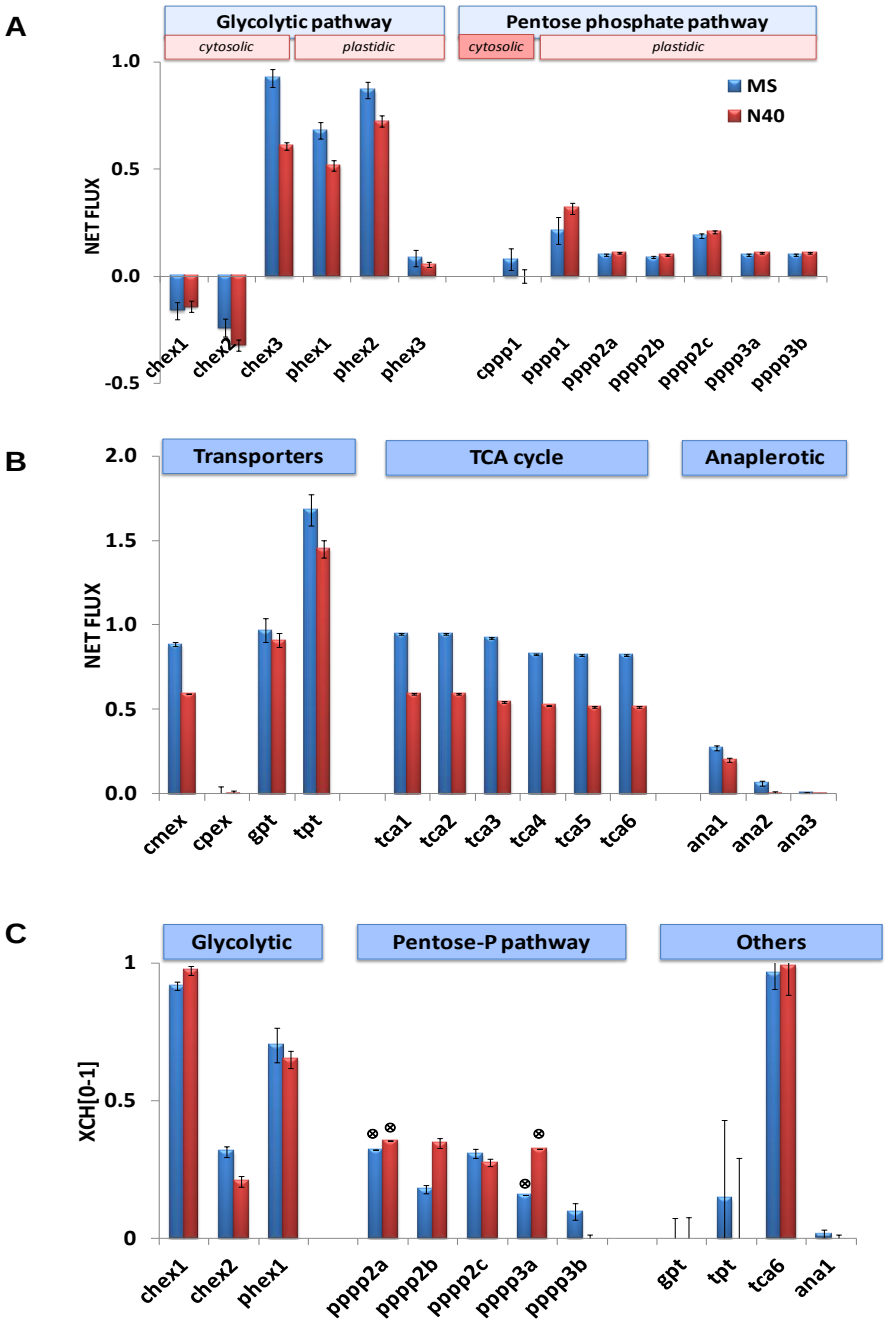
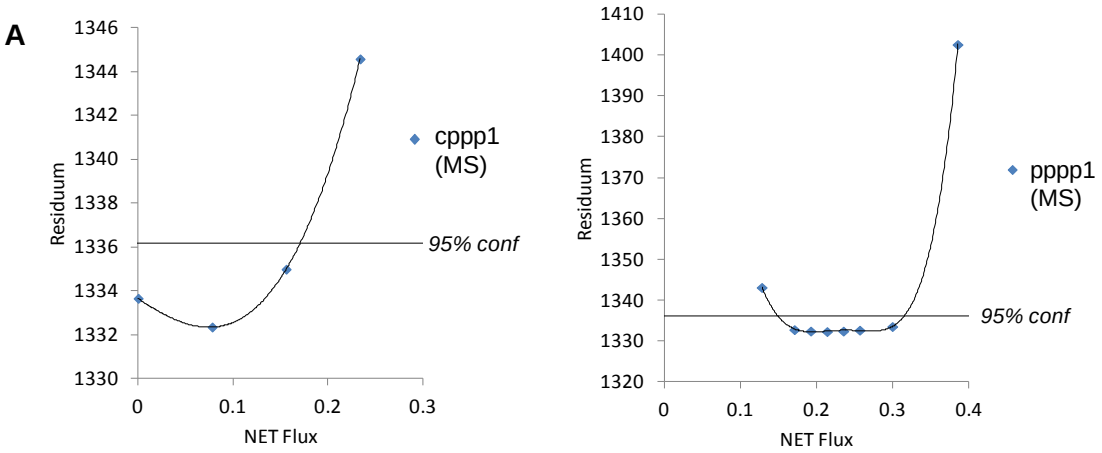
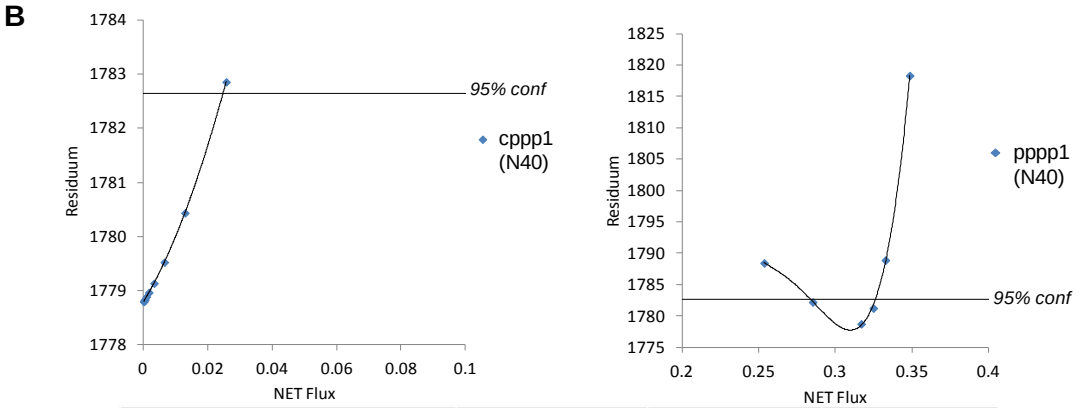


Figure 5.27 Comparison of the best fit flux estimates of MS and N40 cultures. Net fluxes through the central pathways are presented in A and B. XCH[0-1] fluxes are presented in C. Poorly defined fluxes (⊗) have no error bars. The 68% confidence limits on the fluxes are obtained by *EstimateStat* and were verified by non-linearised statistics. Significant differences were observed in Arabidopsis flux phenotypes under different N sources.



oxPPP in MS	optimal	min	max
cppp1 (cytosolic)	0.078	0	0.165
Pppp1 (plastidic)	0.214	0.149	0.316



oxPPP in N40	optimal	min	max
cppp1 (cytosolic)	0.0001	0	0.0247
Pppp1 (plastidic)	0.317	0.284	0.326

Figure 5.28a Non linear statistical analysis of confidence limits of oxPPP steps though the cytosol and plastid in MS and N40 models. The confidence limits of fluxes via individual compartments doesn't reflect the total estimates of oxPPP as they may be inversely related. Monte-Carlo flux estimates (Fig 5.26b) could be used for statistical comparison to establish difference in net oxPPP flux between the models if any.

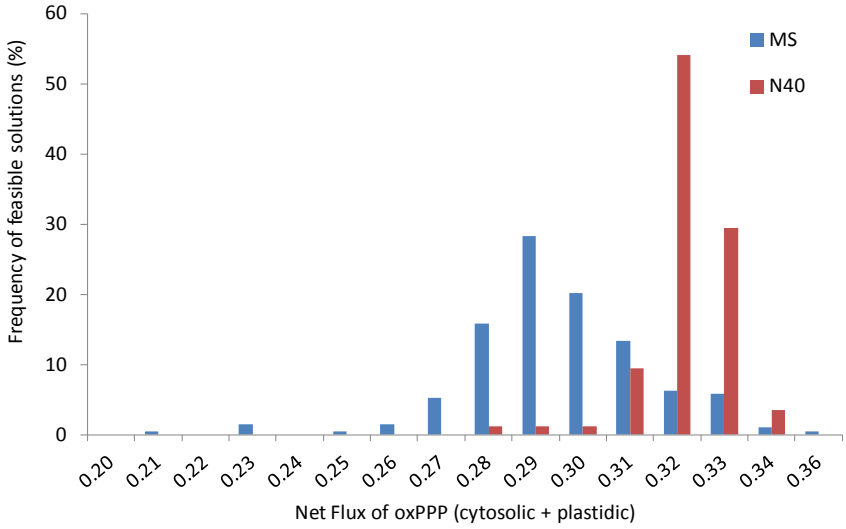
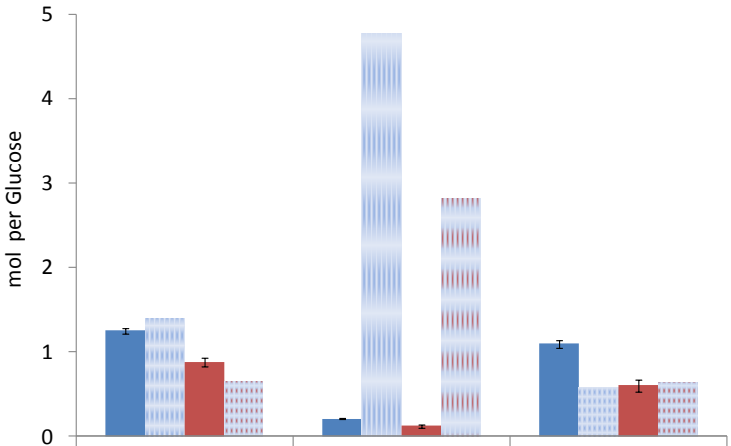


Figure 5.28b Distribution of Monte-Carlo flux estimates of total flux via oxidative pentose phosphate pathway (oxPPP-cytosolic + plastidic) from MS and N40 models. The distributions were significantly (statistical tests mentioned in text) different between the models which established enhanced oxPPP flux in N40 cells.



	ATP	NADH	NADPH
MS-demand	1.251	0.213	1.096
MS-generated	1.402	4.773	0.585
N40-demand	0.878	0.126	0.606
N40-generated	0.648	2.817	0.634

Figure 5.29. The capacity of the cells to meet the demand for reductant (NADPH), ATP and NADH. The net demand was calculated from the biomass data and the ability of the cells to meet the demand was calculated from the flux maps (Chapter 4, Masakapalli et al., 2010). N40 cells have the capacity to meet the NADPH demand via the oxPPP. MS cells could not meet the increased demand for reductant using the oxPPP, and the flux map (Fig 5.25) pointed to a substantial contribution from an NADP-dependent isocitrate dehydrogenase (tca3).

6 Metabolic flux phenotypes of *Arabidopsis* cells grown on different levels of inorganic phosphate

6.1 Introduction

6.1.1 Background

Previous chapters has established that ^{13}C based steady state MFA with an approach of multi-model simulations that can fit data from parallel feeding experiments (Chapter 3), is robust in determining compartmented fluxes in plant cells (Chapter 4) and can explain metabolic phenotypes under perturbed nutritional conditions (Chapter 5). However, all these studies were done on *Arabidopsis* cell suspension phenotypes that grew rapidly. In this chapter, the developed strategies were applied to slowly growing cells under severe nutritional stress. The question I wished to address was the extent to which the cells would alter the fluxes through the central metabolism in order to maintain the available resources. The literature indicated that phosphate starvation would result in plant phenotypes with a very slow growth rate (Plaxton and Tran, 2011), and this condition was selected in this study because of the global significance of phosphate availability for plant productivity. In fact a recent review on the metabolic response to phosphate starvation (Plaxton and Tran, 2011) has highlighted the need for a thorough investigation of the central metabolic network under these conditions.

6.1.2 Significance of phosphorus in metabolic context

Phosphorus (P) is an essential macronutrient required for plant growth and metabolism. Modern agriculture depends on application of phosphate (Pi) to support crop yields. There is an increasing anxiety about P supply for agriculture. According to several reports, the world

will reach peak phosphate production around the year 2034, after which the depletion starts (Cordell et al., 2009). The current estimates suggest global rock reserves are finite and likely to be exhausted within some hundreds of years (Dawson and Hilton 2011). This may depend on many factors including demand and supply. Alternative sources may become more relevant in the future. The world is looking for a shift from a Pi+ (abundant) world to Pi- (limited) world over an uncertain time frame. A thorough understanding of metabolic responses to phosphate starvation at cellular level could be useful for modifying plants for optimal Pi utilisation.

The projected global depletion of non-renewable inorganic phosphate (Pi) reserves has prompted researchers to undertake genome-wide transcriptomic (Morcuende et al., 2007) and metabolite profiling (Huang et al., 2008) studies to understand the metabolic response of plants to Pi limitation and develop strategies for optimal Pi utilisation. These studies suggest Pi starvation leads to the restructuring of carbon and nitrogen metabolism in plants, but do not establish the impact this has on the cellular metabolic activity *per se*.

6.1.3 Objectives

The specific aims of this chapter are:

- To establish the metabolic phenotypes of *Arabidopsis* cell suspension associated with P status by varying the supply of inorganic phosphate (Pi).

Pi- **(0 mM)** **MS** **(1.25 mM)** **Pi+** **(5 mM)**

- To quantify the central metabolic fluxes of heterotrophic carbon metabolism using steady-state metabolic flux analysis (MFA)
- To examine the effect of Pi availability on cellular biomass production and carbon use efficiency

Radio-respirometric studies using specifically-labelled [^{14}C]substrates (glucose and gluconate) established that adaptation to Pi-limitation is associated with a change in the flux distribution in central carbon metabolism in *Arabidopsis* cell suspension. Biomass output fluxes were quantified from measurements of [U-^{14}C]glucose metabolism, and additional metabolic constraints needed for an over-determined system were provided by the distribution of mass isotopomers of amino acids and other metabolites after supplying the cells with [1-^{13}C]-, [2-^{13}C]- or [$^{13}C_6$]glucose. Multiple fluxes across the central network were obtained by mathematical modeling at metabolic and isotopic steady-state. Flux maps, the ultimate metabolic readouts of the Arabidopsis cell suspension phenotypes at different levels of phosphorus nutrition were deduced.

6.2 Physiological characterisation of cell cultures under different Pi levels

Arabidopsis cultures fed on glucose in the dark showed significantly restricted growth (Fig 6.1A) under phosphorus limitation (Pi- [0 mM Pi]) when compared to normal (MS [1.25 mM Pi]) and high phosphorus conditions (Pi+ [5 mM Pi]). The cell cultures take up available inorganic phosphate (Fig 6.1B), glucose (Fig 6.1C) and oxygen (Fig 6.1D) from the medium, implying that they are metabolically active during the growth period. *In vivo* ^{31}P NMR of 5 d old Pi- cultures showed severe depletion of internal pools of phosphorylated metabolites in comparison to MS and Pi+ cultures (Fig 6.1E). The Pi- cells were viable until 20 days without any significant increase in biomass. This was confirmed by microscopically observing the cells after dual staining using fluorescein diacetate (FDA) and Evans blue (section 2.3.1). Feeding experiments with ^{13}C labeled glucose substrate until day 20 (section 6.4.1) and ^{14}C glucose uptake experiments until 15 days (section 6.5) further validate their viability as they uptake and metabolise the glucose.

6.3 Initial characterisation of carbohydrate oxidation at differing Pi levels - radiorespirometric analysis

The effect of Pi on sugar metabolism was examined by monitoring the release of $^{14}CO_2$ from positionally ^{14}C-labelled glucose and gluconate. Dark grown (day 5) Pi- (0 mM), MS (1.25mM) and Pi+ (5mM) suspension cultures were fed with [1-^{14}C]glucose, [2-^{14}C]glucose, [3,4-^{14}C]glucose, [6-^{14}C]glucose and [1-^{14}C]gluconate for 48 hours. $^{14}CO_2$ released was captured by alkaline trap at time intervals of 2hrs, 4hrs, 6hrs, 8hrs, 10hrs, 12hrs, 24hrs, 36hrs and 48hrs). Ratios and $^{14}CO_2$ released (relative to % label applied) were calculated from the scintillation data.

For all cultures (Fig 6.2), the trend of $^{14}CO_2$ release from C3,4 > C1 > C6 > C2 implies oxidation of glucose via a combination of glycolysis and the oxidative pentose phosphate pathway (oxPPP) followed by TCA cycle, with recycling of triose-P to hexose-P and some pentan synthesis (Harrison and Kruger, 2008). The pattern of ratios obtained after 24 h in Pi- cultures was significantly different from those grown in MS or Pi+ medium (Fig 6.3). This indicates that fluxes through the central pathways of carbohydrate oxidation are altered. $^{14}CO_2$ released from [1-^{14}C]gluconate (X) was less than [3,4-^{14}C]glucose (3,4) in Pi- cultures (Fig 6.2A) but generally greater in MS (Fig 6.2B) and Pi+ cultures (Fig 6.2C). Additionally, the 1/6 ratio and 2/1 ratio in Pi- cultures differed from those in the other two conditions (Fig 6.3). In combination, these differences implied that relative flux through the oxPPP is lower in Pi- cells than in those grown in MS or Pi+ conditions.

6.4 Validation of isotopic steady state condition

6.4.1 Phosphate-starved cultures take longer time to reach isotopic steady state.

Analysis of preliminary experiments of growing Pi- cultures on 20% [$^{13}C_6$] glucose for 5 and 10 days showed that a longer incubation was necessary to reach isotopic steady state. Mass isotopomer abundances of soluble amino acids and organic acids were obtained from GC-MS analysis of cationic and anionic fractions of perchloric acid extracts after 10, 15 and 20 d (Fig 6.4). Some metabolites reached an isotopic steady state as indicated by an average ^{13}C abundance of 20%, but others did not. In particular:

- Most soluble metabolites reached an isotopic steady state between day 15 and day 20 (Figs 6.4A, 6.4B).
- Lys, Leu, Ile, Val, Phe and Tyr were exceptions and took longer (Fig 6.4A).

- Interestingly malate reached isotopic equilibrium, whereas fumarate and succinate did not, implying significant activities of PEPcarboxylase and malate dehydrogenase, a characteristic feature of Pi limitation in plants.
- Amino acids obtained from protein hydrolysates did **not** reach an isotopic steady state (Figs 6.4 C, 6.4D). This is assumed to be due to the pre-existing pools of protein. For flux analysis these unlabelled pools have to be accounted as explained in Chapter 3.

6.4.2 Ideal conditions for steady state flux analysis under different Pi levels was established

As the normal MS and Pi+ cells grow faster, assumptions based on previous Chapters and the average labelling from [$^{13}C_6$]glucose feeding experiments indicated these cells reached isotopic steady state by day 5.5. The ideal conditions selected for further ^{13}C feeding experiments and flux analysis for Pi- was day 15, and for MS and Pi+ it was day 5.5

6.5 Biomass measurements

6.5.1 Analysis of [U-^{14}C]glucose metabolism shows that phosphorus stress alters biomass production

Biomass production by Arabidopsis cells grown with Pi at 5 mM (Pi+), 1.25 mM (MS) or 0 mM (Pi-) was assessed by supplying cultures with [U-^{14}C]glucose for either 24 h (MS (4.5 days to 5.5 days) and Pi+ (4.5 days to 5.5 days)) or 120 h (Pi- (from day 10 to day 15)). The biomass accumulation rates measured (Fig 6.5) show that

- Pi availability affected the redistribution of label into end products.

- Pi-limited cells increased excretion of organic acids and amino acids into the culture medium.
- The proportion of carbon converted to biomass (carbon use efficiency) in Pi- cells (24%) was less than that of cells grown in MS (57%) or Pi+ (66%).
- Pi-limitation decreased the proportion of C incorporated into amino acids and protein.

6.6 ^{13}C label measurements

Label measurements of free amino acids, organic acids, sugars, amino acids derived from protein, sugars derived from cell wall and starch, were obtained from extracts of Arabidopsis cell suspension cultures supplied with 99%[1-^{13}C]-, 99%[2-^{13}C]- or 20% [$^{13}C_6$]glucose that were grown to isotopic steady state for each condition of Pi. As already detailed in Chapter 3 and Chapter 5, mass isotopomers were obtained by GC-MS after TBDMS or TMS derivatisation. Positional isotopomers were quantified by ^{13}C-NMR. Spectral data analysis was done using: AMDIS (NIST), MSSearch (NIST), Chemstation (Agilent), NUTS (Acorn NMR) and infinite dimensional matrix calculus. The datasets of isotopomer measurements were handled as described in Chapter 3 for accurate assessment and precise error estimation. The final data used for metabolic modeling is presented in Appendices 6.1A, 6.1B, 6.1C and 6.1D.

6.7 Metabolic modeling

Metabolic modeling and statistical analysis were conducted using 13C-FLUX on the same model used in Chapter 5 (Appendix 5.2). The model refinements, fitting of experimental data to the metabolic models and the confidence intervals were verified by non-linear statistical

analysis as described in Chapters 3and 5. The followings experimental data obtained under different Pi levels was used for constraining the parameter fitting of the metabolic models

iii) the molar output fluxes (Relative to glucose uptake) towards biomass components derived from [U-^{14}C]glucose feeding experiments (Table 6.1). The glucose uptake rate was constrained to 1.

iv) the positional isotopomers derived from ^{13}C-NMR and mass isotopomers derived from GC-MS analysis of metabolites from separate feeding experiments with 99%[1-^{13}C]glucose (1-^{13}C), 99%[2-^{13}C]glucose (2-^{13}C) or 20%[$^{13}C_6$]glucose (U-^{13}C) (Appendices 6.1A, 6.1B, 6.1C and 6.1D). The measurements were handled as discussed in chapter 3.

6.7.1 Availability of Pi alters the flux distribution through central metabolism

The best fit flux estimates (Table 6.2) are presented in the form of flux maps (Figs 6.6, 6.7 and 6.8) and histograms (Fig 6.9). The flux estimates show that Pi-limitation:

- ❖ decreases flux through the predominantly cytosolic oxidative pentose phosphate pathway (cppp1);
- ❖ increases the anaplerotic flux via PEP carboxylase (ana1) and malic enzyme (ana2);
- ❖ increases relative TCA cycle fluxes and CO_2 efflux.
- ❖ also decreased the relative rate of inorganic nitrogen assimilation, a major sink for NADPH in these cells.

6.8 Conclusions

Although phosphorus starvation significantly reduced growth of heterotrophic Arabidopsis cultures, the cells remained metabolically active (and thus viable) for prolong periods. Radio-respirometric analysis established that adaptation to P-limitation was associated with a change in the flux distribution in central carbon metabolism in Arabidopsis cell cultures. Pi-limitation altered the cellular metabolic phenotype. Carbon conversion efficiency decreased in Pi-limited cultures. Excretion of organic and amino acids in Pi-limited cultures was accompanied by changes in fluxes through the TCA cycle and associated anaplerotic steps. Pi-limitation decreased flux through the oxidative pentose phosphate pathway, but increased the relative importance of the plastidic contribution. Pi-limitation also decreased the relative rate of inorganic nitrogen assimilation, a major sink for NADPH in these cells. Curiously, Pi limitation was associated with a shift in the relative importance of the cytosolic and plastidic contributions to the oxidative pentose phosphate pathway.

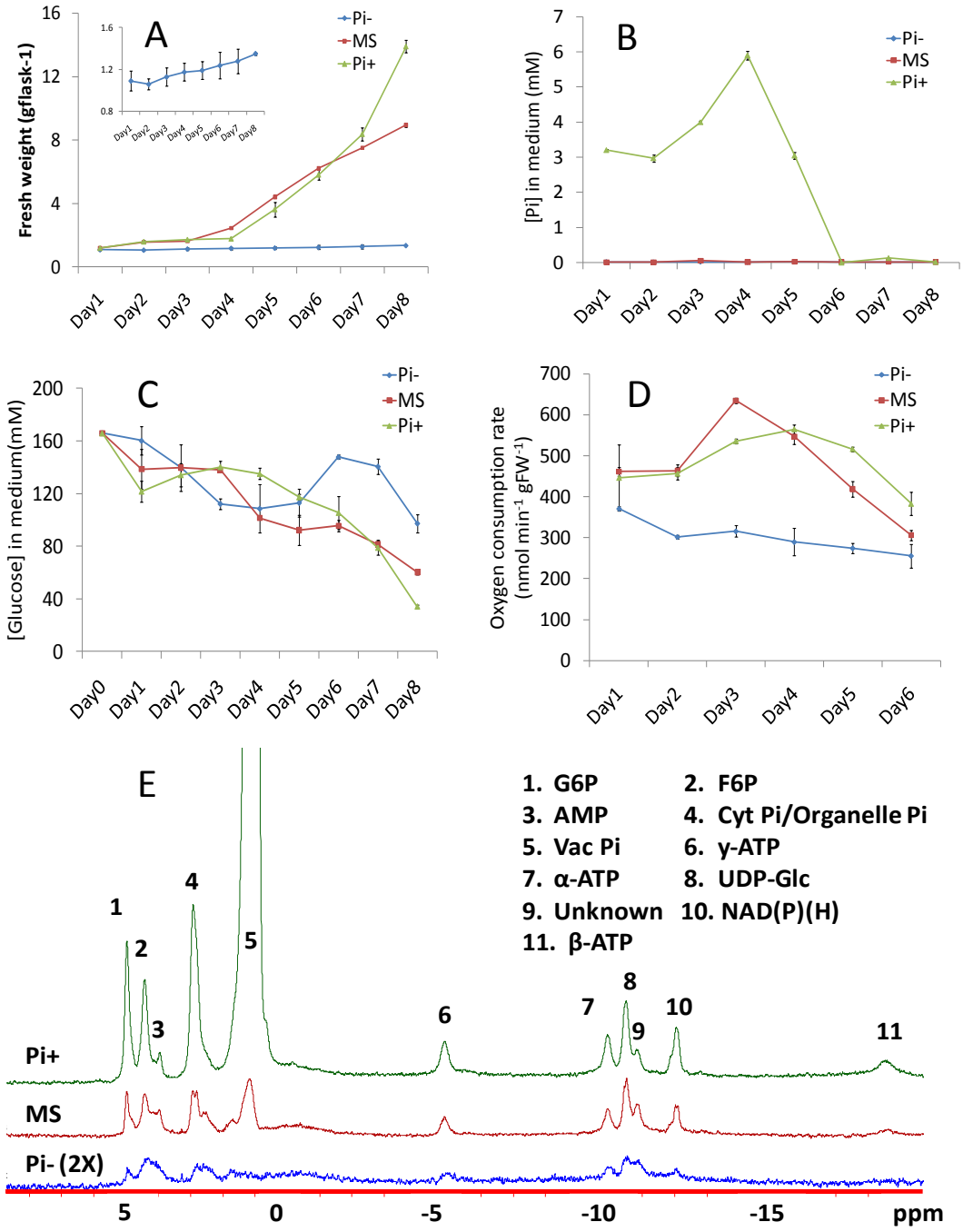


Figure 6.1 Effect of [Pi] on growth of Arabidopsis suspension cultures. A, fresh weight; B, Pi uptake; C, glucose consumption; D, oxygen consumption; E, profile of intracellular phosphorylated intermediates at day 5 determined by in vivo ^{31}P-NMR.

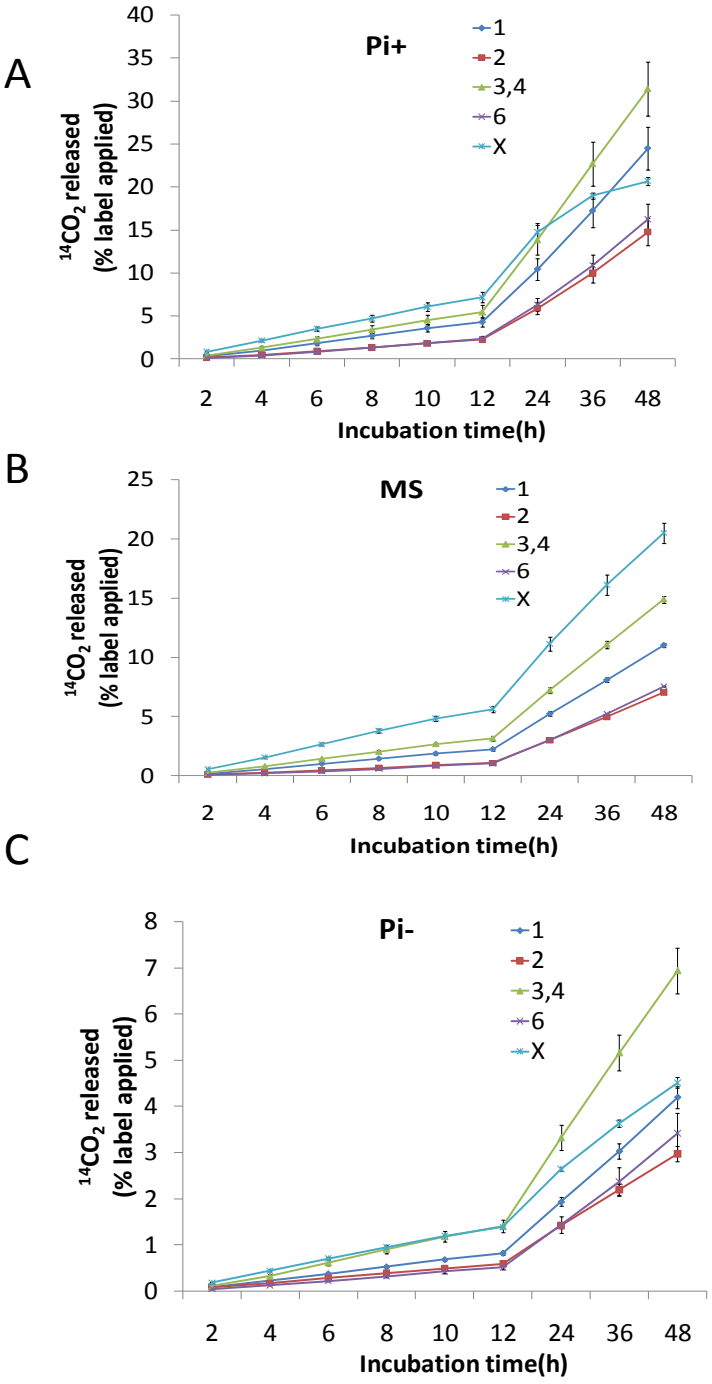


Figure 6.2 $^{14}CO_2$ release from positionally labeled ^{14}C-substrates. Dark-grown (day 5) suspension cultures were supplied with [1-^{14}C]glucose (1), [2^{14}C]glucose (2), [3,4-^{14}C]glucose (3,4), [6-^{14}C]glucose (6) and [1-^{14}C]gluconate (X) for 48 h and metabolized $^{14}CO_2$ was monitored over the time intervals indicated. Cultures were grown in: A, Pi+ medium; B, MS medium; C, Pi- medium.

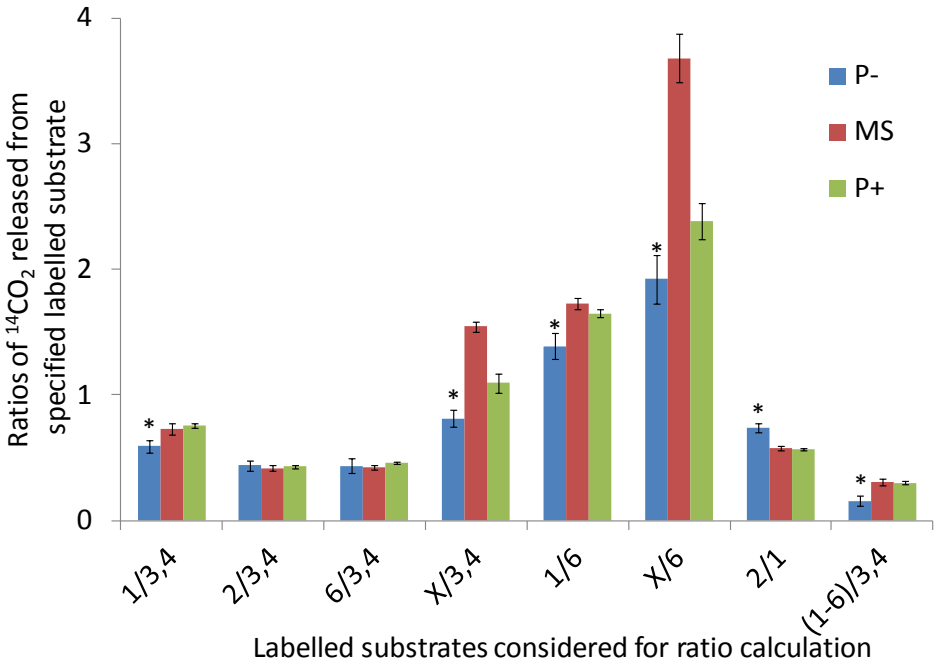


Figure 6.3 The ratios of $^{14}CO_2$ release from positionally labelled glucose from Dark-grown (day 5) *Arabidopsis* suspension cultures supplied with [1-^{14}C]glucose (1), [2^{14}C]glucose (2), [3,4-^{14}C]glucose (3,4), [6-^{14}C]glucose (6) and [1-^{14}C]gluconate (X) at 24 h time point. The ratios of $^{14}CO_2$ release from substrates labeled in various positions after 24 h incubation shows significant changes in patterns of carbohydrate oxidation. Several ratios changed significantly (*) when Pi availability was reduced implying that Pi-limitation alters the relative activities of the different pathways of carbohydrate oxidation

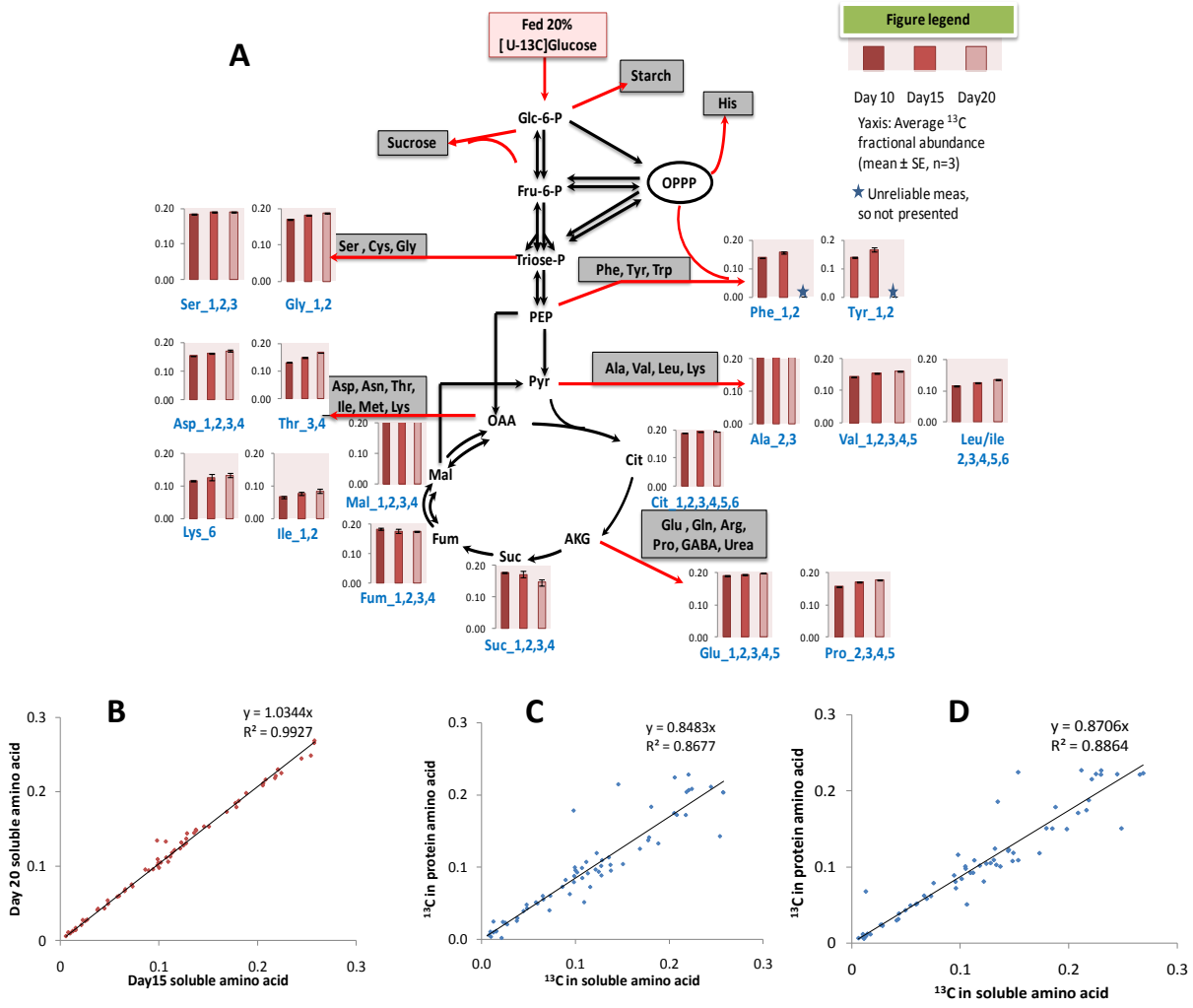


Fig 6.4: Isotopic labelling of metabolites in the central network of carbon metabolism in Pi- cultures. (A) depicts the average ^{13}C abundance in soluble metabolites of Pi starved Arabidopsis cultures fed with 20% [U-^{13}C]glucose for 10, 15 and 20 days. For each metabolite the carbon positions included in the analysed fragment are indicated in blue. Additional figures represent the comparisons of ^{13}C enrichment in (B) soluble amino acids at day 15 vs. day 20; and soluble vs. protein-derived amino acids at day 15 (C) and day 20 (D). The ^{13}C fractional abundance of soluble extracts was measured by GC-MS after derivatisation of the samples with MSTFA. Spectral data analysis was done using a combination of tools: AMDIS32(NIST), MSSearch(NIST), Chemstation (Agilent) and Mass Correction tool as explained in chapter 3

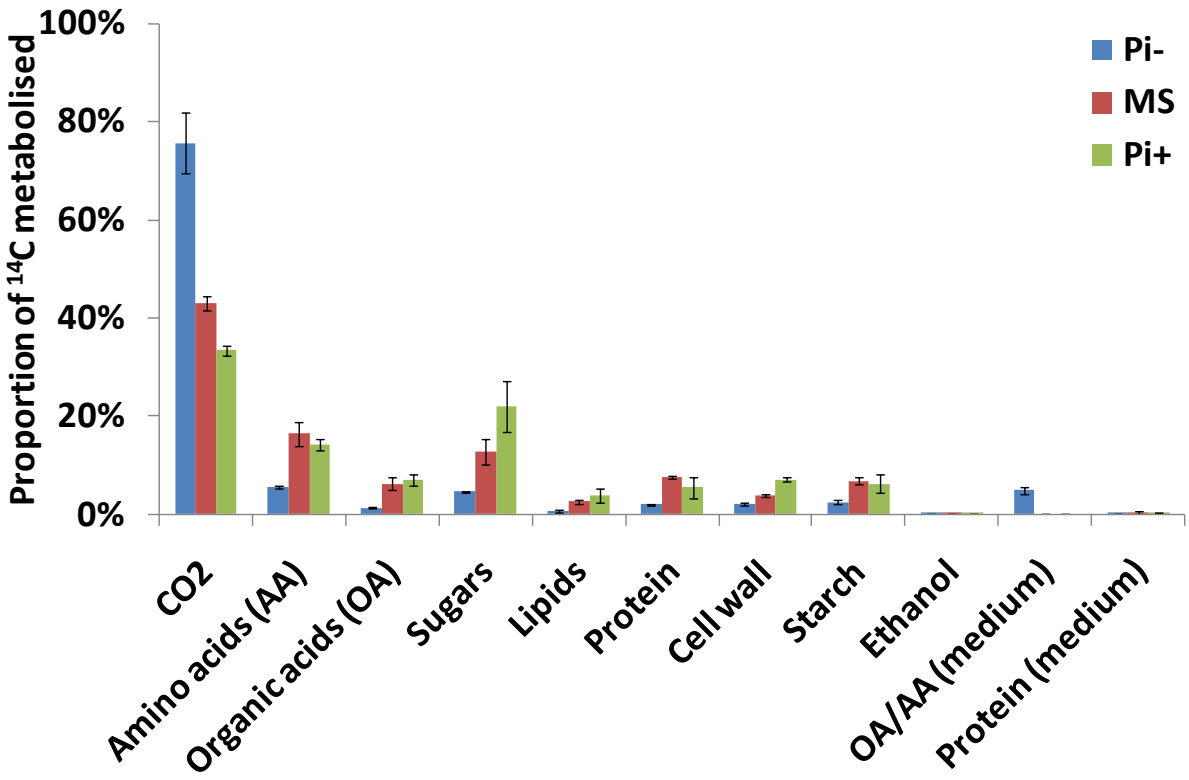


Figure 6.5 Proportion of ^{14}C metabolized into each fraction in *Arabidopsis* cell suspension culture grown under different Pi levels (Pi-, MS and Pi+). Significant differences were observed in CO_2 release and other biomass components. The values presented as mean proportions of radioactivity in relation to metabolized glucose ± sd derived from 4 separate flasks in each condition.

Table 6.1: Fluxes to biosynthetic outputs of *Arabidopsis* cell suspension culture grown under different Pi levels (Pi-, MS and Pi+). The values are presented as molar output fluxes relative to glucose uptake rate of 1. These were calculated based on radioactive proportions (Fig 6.5) and relative quantifications of individual metabolites as described in Chapter 5.

	Pi-	**MS**	**Pi+**		sd (**Pi-**)	sd (**MS**)	sd (**Pi+**)
$^{14}CO_2$	4.5410	2.5825	2.0191		0.0000	0.5201	0.3104
Sucrose-glucosyl	0.0230	0.0651	0.1097		0.0012	0.0219	0.0295
Sucrose-fructosyl	0.0230	0.0651	0.1097		0.0012	0.0219	0.0295
Cell wall C6	0.0178	0.0315	0.0581		0.0030	0.0029	0.0072
Cell wall C5	0.0047	0.0083	0.0153		0.0008	0.0008	0.0019
Starch	0.0256	0.0689	0.0621		0.0045	0.0151	0.0176
Lipid-glycerol	0.0011	0.0041	0.0061		0.0003	0.0011	0.0024
Lipid-FA	0.0187	0.0717	0.1061		0.0051	0.0185	0.0414
Ethanol	0.0031	0.0045	0.0055		0.0007	0.0017	0.0010
Citrate	0.0067	0.0262	0.0145		0.0012	0.0024	0.0056
Isocitrate	0.0008	0.0028	0.0003		0.0007	0.0013	0.0001
Malate	0.0138	0.0363	0.0343		0.0018	0.0133	0.0108
Oxoglutarate	0.0045	0.0000	0.0000		0.0005	0.0000	0.0000
Succinate	0.0023	0.0102	0.0476		0.0005	0.0053	0.0246
Urea	0.0291	0.0025	0.0006		0.0059	0.0016	0.0008
Glycolate	0.0003	0.0003	0.0001		0.0001	0.0000	0.0001
Ala	0.0167	0.0984	0.0784		0.0022	0.0093	0.0298
Arg	0.0013	0.0034	0.0026		0.0005	0.0006	0.0010
Asx	0.0045	0.0105	0.0068		0.0003	0.0014	0.0023
Cys	0.0007	0.0030	0.0019		0.0001	0.0007	0.0009
Glx	0.0637	0.0975	0.0790		0.0023	0.0032	0.0145
Gly	0.0076	0.0221	0.0192		0.0009	0.0025	0.0066
His	0.0006	0.0016	0.0019		0.0001	0.0003	0.0002
Iso	0.0004	0.0020	0.0020		0.0001	0.0007	0.0004
Leu	0.0014	0.0058	0.0052		0.0001	0.0013	0.0014
Lys	0.0026	0.0115	0.0077		0.0003	0.0027	0.0036
Met	0.0005	0.0018	0.0012		0.0000	0.0004	0.0005
Phe	0.0012	0.0059	0.0049		0.0001	0.0014	0.0005
Pro	0.0050	0.0145	0.0103		0.0009	0.0019	0.0040
Ser	0.0103	0.0438	0.0298		0.0007	0.0034	0.0075
Thr	0.0042	0.0089	0.0198		0.0009	0.0027	0.0031
Trp	0.0007	0.0025	0.0015		0.0001	0.0005	0.0007
Tyr	0.0056	0.0025	0.0019		0.0009	0.0005	0.0008
Val	0.0025	0.0040	0.0063		0.0002	0.0009	0.0007
Gaba	0.0229	0.0172	0.0136		0.0044	0.0039	0.0052

Table 6.2 The best fit flux estimates of *Arabidopsis* cell suspension culture grown under varying supply of inorganic phosphate (Pi): Pi- (Phosphate stress, 0 mM), MS (normal medium, 1.25 mM) and Pi+ (Phosphate abundant, 5 mM). The 68% confidence limits (sd) of all the fluxes were obtained by *EstimateStat*. The confidence limits of free fluxes were further verified by non-linearised statistics (95% CI) which were in good agreement with sd obtained by *EstimateStat* and hence not presented. Significant differences in the TCA cycle fluxes, anaplerotic fluxes and net PPP flux (cppp1+pppp1) were observed in. All fluxes are relative to glucose uptake of 1± 0.05. "nd" is non determinable flux.

NET FLUXES [0-1]		**Pi-**	**MS**	**Pi+**		**sd (Pi-)**	**sd (MS)**	sd **(Pi+)**
Hexose/triose phosphate metabolism								
Cytosolic phosphoglucose isomerase	*chex1*	-0.014	0.189	0.053		0.139	0.036	0.025
Cytosolic Fru-6-P → triose-P/PEP interconversion	*chex2*	-0.037	0.124	-0.057		0.139	0.036	0.025
Cytosolic pyruvate kinase	*chex3*	1.422	1.012	0.724		0.018	0.013	0.063
Plastidic phosphoglucose isomerase	*phex1*	0.825	0.352	0.415		0.141	0.031	0.023
Plastidic Fru-6-P → triose-P/PEP interconversion	*phex2*	0.899	0.536	0.601		0.140	0.034	0.024
Plastidic pyruvate kinase	*phex3*	0.000	0.046	0.137		0.006	0.010	0.059
Triose-P isomerase	*Vtpi*	-0.039	0.120	-0.063		0.139	0.036	0.025
Pentose p pathway								
Cytosolic pentose phosphate pathway - oxidative steps	*cppp1*	0.000	0.138	0.287		0.020	0.037	0.012
Plastidic pentose phosphate pathway - oxidative steps	*pppp1*	0.116	0.147	0.000		0.012	0.046	0.021
Plastidic transketolase - C7/C5 conversion	*pppp2a*	0.041	0.097	0.097		0.004	0.004	0.004
Plastidic transketolase - C6/C4 conversion	*pppp2b*	0.033	0.087	0.089		0.004	0.004	0.004
Plastidic transketolase - C5/C3 conversion	*pppp2c*	0.074	0.184	0.186		0.008	0.007	0.009
Plastidic transaldolase - C7/C4 conversion	*pppp3a*	0.041	0.097	0.097		0.004	0.004	0.004
Plastidic transaldolase - C6/C3 conversion	*pppp3b*	0.041	0.097	0.097		0.004	0.004	0.004
Transporters/exchanges								
Pyruvate uptake by mitochondria	*cmex*	1.371	0.852	0.640		0.016	0.007	0.015
Pyruvate uptake by plastids	*cpex*	0.031	0.057	0.000		0.006	0.011	0.061
Plastidic Glc-6-P/Pi translocator	*gpt*	0.967	0.568	0.477		0.142	0.059	0.025
Plastidic triose-P/Pi & PEP/Pi translocators	*tpt*	-1.800	-1.028	-1.090		0.279	0.071	0.071

Tricarboxylic acid cycle								
Mitochondrial pyruvate dehydrogenase	*tca1*	1.537	0.861	0.640		0.004	0.004	0.005
Citrate synthase	*tca2*	1.537	0.860	0.640		0.004	0.004	0.004
Aconitase/isocitrate dehydrogenase	*tca3*	1.529	0.831	0.626		0.004	0.004	0.004
2-oxoglutarate → Succinate	*tca4*	1.430	0.699	0.520		0.004	0.004	0.004
Succinate → fumarate conversion	*tca5*	1.427	0.689	0.473		0.004	0.004	0.004
Fumarase	*tca6a+b*	1.429	0.692	0.475		0.004	0.004	0.004
Anaplerotic fluxes								
PEP carboxylase	*ana1*	0.303	0.259	0.246		0.017	0.008	0.016
Mitochodrial malic enzyme	*ana2*	0.166	0.008	0.000		0.017	0.008	0.016
Plastidic malic enzyme	*ana3*	0.001	0.008	0.007		0.000	0.001	0.002
EXCHANGE FLUXES [0-1]		**Pi-**	**MS**	**Pi+**		**sd (Pi-)**	**sd (MS)**	sd **(Pi+)**
Cytosolic phosphoglucose isomerase	*chex1*	0.970	0.952	0.962		0.024	0.019	0.034
Cytosolic Fru-6-P → triose-P/PEP interconversion	*chex2*	0.395	0.417	0.242		0.102	0.006	0.022
Plastidic phosphoglucose isomerase	*phex1*	0.684	0.651	0.710		0.027	0.079	0.064
Plastidic transketolase - C7/C5 conversion	*pppp2a*	0.225	0.289	0.990		nd	nd	nd
Plastidic transketolase - C6/C4 conversion	*pppp2b*	0.636	0.132	0.000		0.024	0.012	0.012
Plastidic transketolase - C5/C3 conversion	*pppp2c*	0.316	0.254	0.212		0.025	0.012	0.009
Plastidic transaldolase - C7/C4 conversion	*pppp3a*	0.178	0.349	0.323		nd	nd	nd
Plastidic transaldolase - C6/C3 conversion	*pppp3b*	0.781	0.011	0.141		0.034	0.011	0.017
Plastidic Glc-6-P/Pi translocator	*gpt*	0.889	0.299	0.383		0.022	0.038	0.032
Plastidic triose-P/Pi & PEP/Pi translocators	*tpt*	0.000	0.065	0.000		nd	0.186	0.249
Fumarase	*tca6*	0.980	0.942	0.894		0.039	0.044	0.072
PEP carboxylase (for PEP carboxykinase)	*ana1*	0.077	0.026	0.011		0.014	0.004	0.008

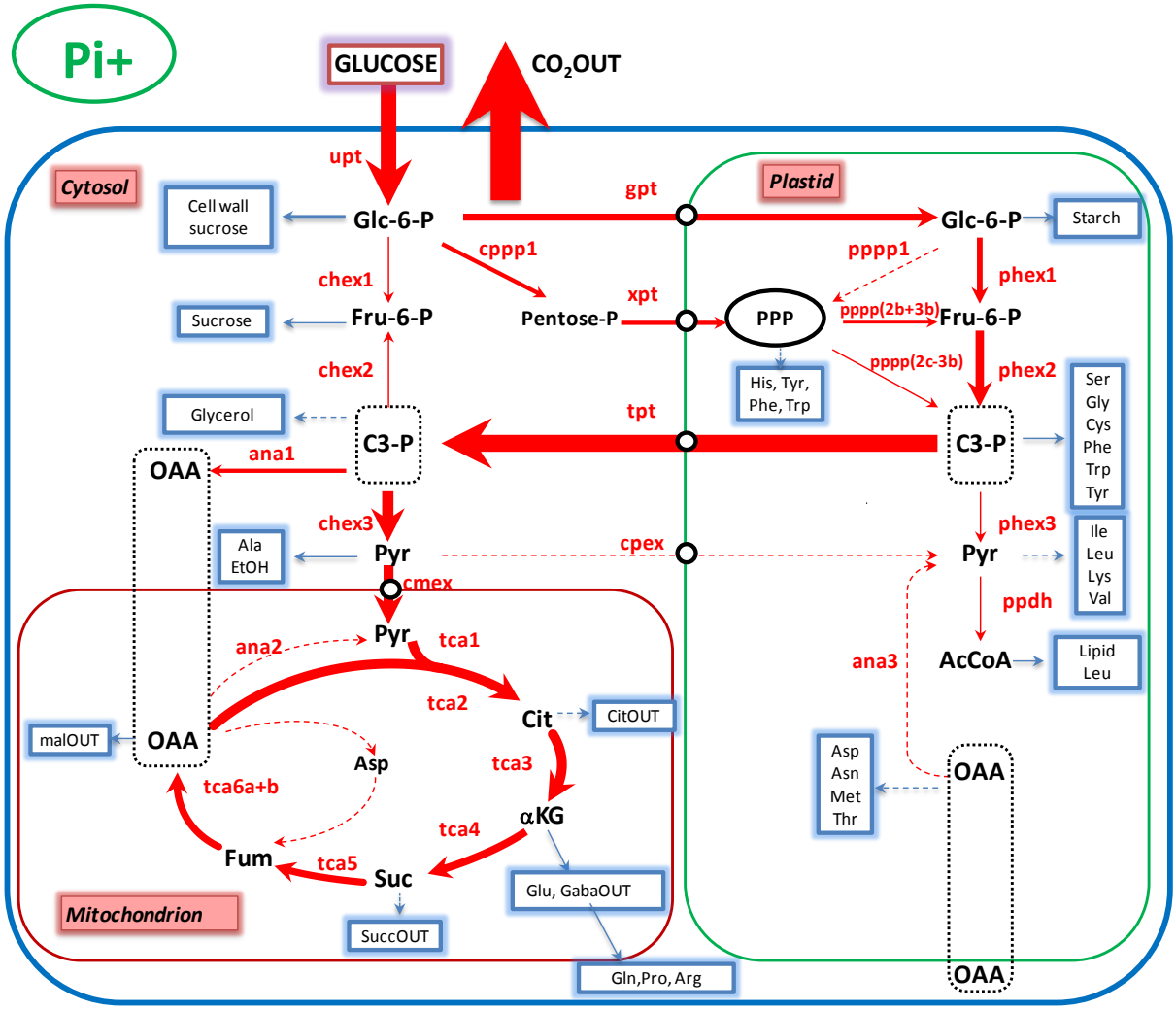


Figure 6.6 Flux map of *Arabidopsis* cell suspension culture grown under Pi+ (5 mM). The best fit net fluxes presented here as the thickness of arrow are estimated from parameter fitting of triple model with labelling data from three separate feeding experiments of [1-^{13}C]glucose, [2-^{13}C]glucose and [$^{13}C_6$]glucose and biomass constraints derived from [U-^{14}C]glucose feeding. All fluxes are relative to glucose uptake. The dotted lines indicate < 5% of flux relative to glucose uptake. The net fluxes and the XCH[0-1] with their confidence limits are presented in Figure 6.9

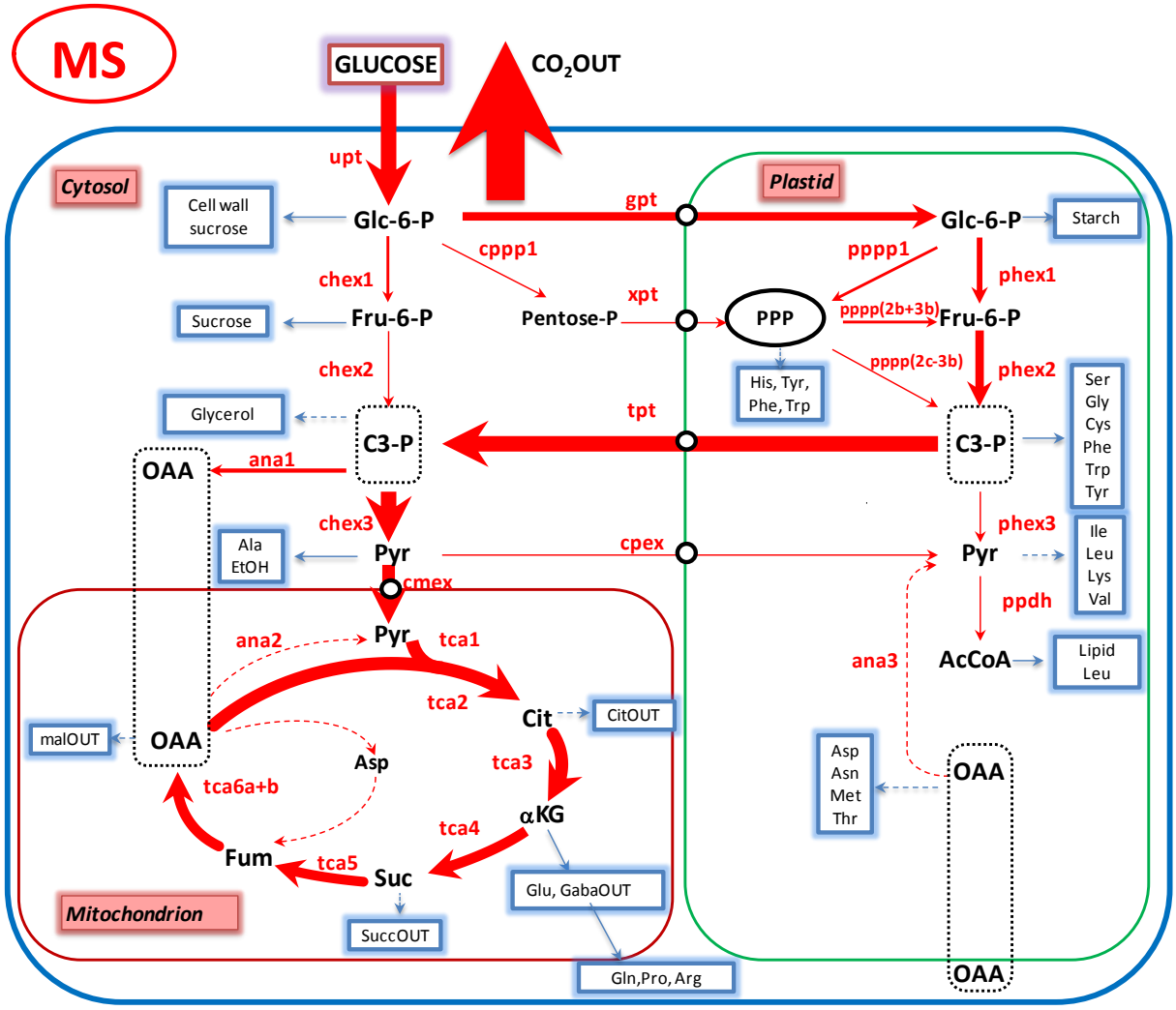


Figure 6.7 Flux map of *Arabidopsis* cell suspension culture grown under MS (1.25 mM phosphate). The best fit net fluxes presented here as the thickness of arrow are estimated from parameter fitting of triple model with labelling data from three separate feeding experiments of [1-^{13}C]glucose, [2-^{13}C]glucose and [$^{13}C_6$]glucose and biomass constraints derived from [U-^{14}C]glucose feeding. All fluxes are relative to glucose uptake. The dotted lines indicate < 5% of flux relative to glucose uptake. The net fluxes and the XCH[0-1] with their confidence limits are presented in Figure 6.9

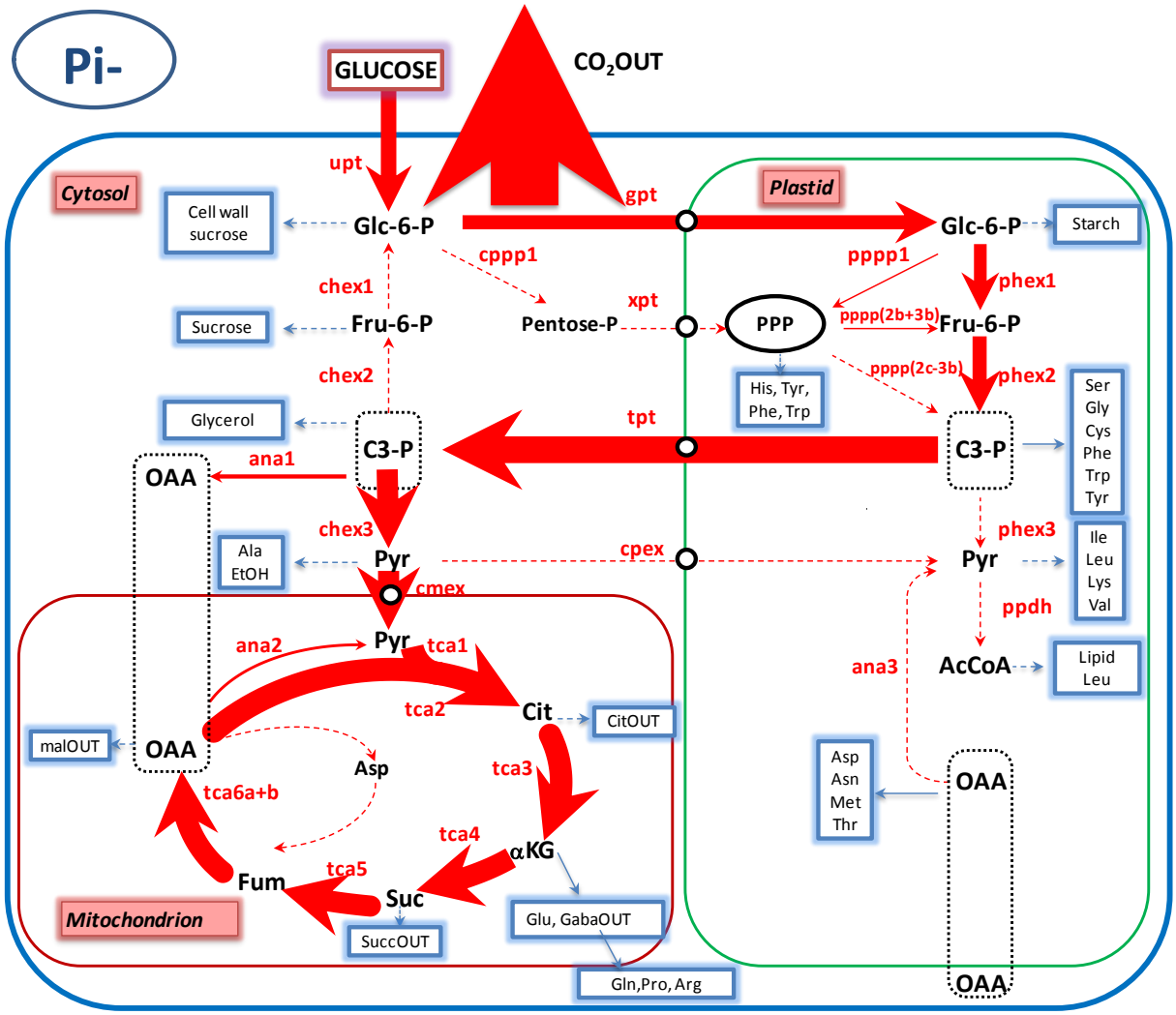


Figure 6.8 Flux map of *Arabidopsis* cell suspension culture grown under severe phosphate starvation (0 mM phosphate). The best fit net fluxes presented here as the thickness of arrow are estimated from parameter fitting of triple model with labelling data from three separate feeding experiments of [1-^{13}C]glucose, [2-^{13}C]glucose and [$^{13}C_6$]glucose and biomass constraints derived from [U-^{14}C]glucose feeding. All fluxes are relative to glucose uptake. The dotted lines indicate < 5% of flux relative to glucose uptake. The net fluxes and the XCH[0-1] with their confidence limits are presented in Figure 6.9

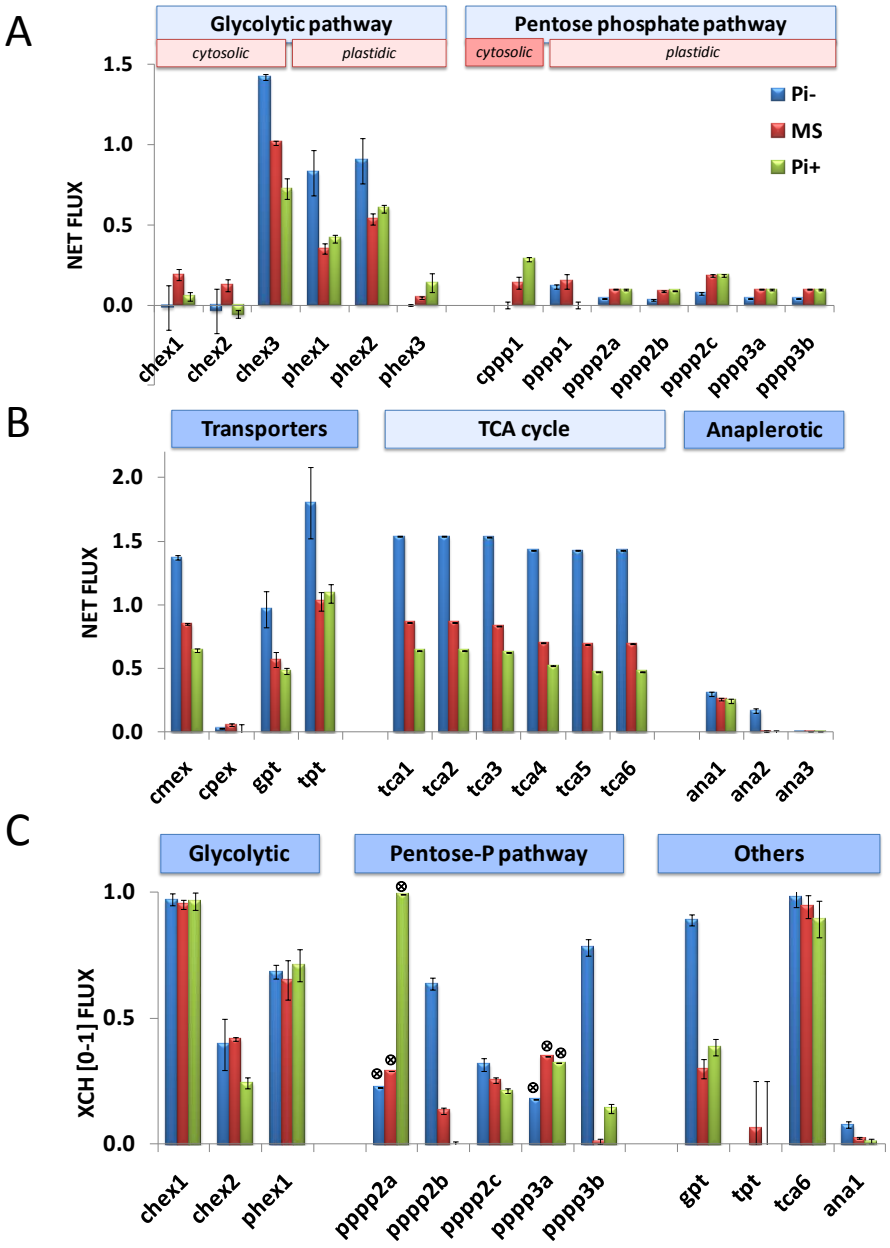


Figure 6.9 Comparison of the best fit flux estimates of Pi-, MS and Pi+ cultures. Net fluxes through the central pathways are presented in A and B. XCH[0-1] fluxes are presented in C. Poorly defined fluxes (⊗) have no error bars. The 68% confidence limits on the fluxes are obtained by *EstimateStat* and were verified by non-linearised statistics. Significant differences in the TCA cycle fluxes and oxPPP were observed in Arabidopsis flux phenotypes.

7 General discussion

7.1 Summary of results

In this work a robust methodology for steady-state metabolic flux analysis (MFA) based on ^{13}C labelling to quantify intracellular fluxes of central carbon metabolism in plants was developed. This involved combining high quality steady-state stable labelling data, metabolic modeling and computational analysis. The modeling strategies presented in Chapter 3 were validated and established for ^{13}C-based network flux analysis using the software framework 13C-FLUX® (version 20050329; Wiechert et al., 2001). In addition, the significance of obtaining high quality isotopomer measurements and accurate assessment of errors over the GC-MS and NMR data was discussed in Chapter 3. Results from Chapter 4 indicated that the *in silico* strategies and the approaches established here have the capacity to distinguish compartmented fluxes of carbohydrate oxidation through parallel reactions occurring in the cytosol and plastid in *Arabidopsis* cell suspension culture using high quality datasets.

Results from Chapters 5 showed that nutritional perturbations of the cell suspension with different sources of inorganic nitrogen (MS (nitrate + ammonia) and N40 (nitrate only)) alter the cellular metabolic phenotype. Growth (in terms of fresh weight) was higher in N40 than normal MS conditions. Growth on nitrate alone (N40) might be expected to increase demand for reductant substantially, but there was marginal increase in flux through the oxPPP (about 9.29%). Surprisingly, cells responded to the ammonium-free medium by decreasing the assimilation of inorganic N while not increasing their demand for reductant substantially. At the same time biomass measurements showed enhanced biosynthesis of sugars and organic acids. Ammonium-free cells (N40) had the capacity to meet the NADPH demand via oxPPP

flux alone, whereas MS cells depended on additional sources of reductant. The shift away from nitrogen assimilation towards storage of carbohydrate in cells on an ammonium-free medium decreased flux through the TCA cycle and increased carbon conversion efficiency. These results might also suggest that the cells adapt to the change in medium by increasing their size rather than dividing.

Results from Chapter 6 showed that severe nutritional stress of inorganic phosphate (Pi) influenced the metabolic phenotype. The flux maps showed that decreasing Pi availability decreased flux through the oxPPP, increased the proportion of substrate fully oxidised through the TCA cycle; and decreased carbon-conversion efficiency. Pi-limitation also decreased the relative rate of inorganic nitrogen assimilation, a major sink for NADPH in these cells. Curiously, Pi limitation was associated with a shift in the relative importance of the cytosolic and plastidic contributions to the oxPPP, i.e., flux maps of Pi- (0 mM) cells showed the oxPPP flux to be predominantly plastidic in contrast to Pi+ (5 mM) in which the oxPPP flux was predominantly cytosolic. The normal cells (MS (1.25 mM) showed contributions of the oxPPP in both the compartments. This result could have been due to enhanced demand of ribose sugars in the cytosol for cell division with increase in phosphate. Tests with appropriate mutants or knockouts for oxPPP reactions should be done to confirm the speculation.

In conclusion, robust methods for stable-isotope steady-state metabolic flux analysis (MFA) based on ^{13}C labelling were established to obtain compartmented flux maps in heterotrophic tissue and successfully extended to understand the effect of perturbations on central fluxes.

The remaining parts of this chapter deals with several common key issues related to flux analysis in plants that were encountered during this investigation. These include metabolic

modeling (section 7.2), revealing compartmented fluxes (section 7.3), isotopic steady state (section 7.4) and the stability of central metabolism to perturbation (section 7.5). The future scope of research in related areas is also discussed (section 7.6).

7.2 Metabolic modeling

Definition of the metabolic network, i*n silico* simulation, data evaluation and statistical analysis form the core components of metabolic modeling for flux analysis. For statistical robustness Monte-Carlo simulations can be used for flux analysis in plants (Schwender et al., 2006, Masakapalli et al., 2010). It has been observed that stochastic sampling of the starting fluxes is routine in the modeling process and combining this with statistical sampling of the measurement data points improves the fitting procedure. This can be done by bootstap-Monte Carlo simulations as detailed in Chapter 3. Monte Carlo simulations with bootstrap sampling of the isotopomer abundances provided an effective method for exploring the flux space that was consistent with the observed redistribution of label. Surprisingly, the availability of a subroutine to exploit this method in the commonly used 13C-FLUX software appears to have been overlooked. In the current work all the models (Chapter 4, 5 and 6) were subjected to bootstrap Monte-Carlo simulations to define the flux space.

Flux estimates obtained from Monte Carlo simulations can be used to generate confidence limits (Antoniewicz et al., 2006), but this is computationally demanding. Antoniewicz et al. (2006) developed an efficient algorithm for accurate determination of flux confidence intervals by a non-linear statistical technique. All the free flux estimates of models obtained in Chapter 5 and 6 were verified for the confidence limits by this method.

7.3 Revealing compartmented fluxes

The extent to which parallel fluxes can be distinguished depends on the labelling of metabolites that report on intermediates in specific compartments (Ratcliffe and Shachar-Hill, 2005, 2006; Allen et al., 2007). For example, starch is considered to report labelling of the plastidic glucose 6-phosphate pool, and sucrose or cell wall report on the cytosolic glucose 6-phosphate pool. The differences in labelling patterns in these metabolites can be indistinguishable due to rapid rate of exchange as observed in soyabean embryos(Allen et al., 2009), or distinguishable as in Arabidopsis suspension cultures (Williams et al. 2008; Masakapalli et al., 2010). The labelling information could be lost in single feeding experiment. Parallel feeding experiments would provide additional confirmation of the labelling differences of the same metabolite. The parallel experiments with different substrates can be simulated together (known as multi-models) to increase the statistical quality of the flux estimates and also to determine the fluxes that are poorly defined with a single labelled substrate (Schwender et al., 2006, Alonso et al., 2007). In this thesis all the flux maps were obtained from parallel feeding experiments followed by multi-model simulations in order to obtain robust compartmented flux maps.

7.4 Isotopic steady state

Steady state MFA requires the system to be at metabolic and isotopic steady state (Roscher et al., 2000). In a typical plant cell, the central free metabolites usually reach isotopic steady state within few hours of labelling whereas polymers like protein, cell wall and starch take several days (Troufflard et al., 2007). Therefore, labelling analysis of soluble metabolites will be ideal for establishing isotopic steady state conditions.

From the current studies conducted on *Arabidopsis* cell suspension cultures, the mass isotopomer measurements of proteinogenic amino acids, starch and cell wall components were biased by inclusion of pre-existing unlabelled pools in the analysis (Chapter 5 and Chapter 6). Similar observations have been reported in earlier studies on *Arabidopsis* embryos (Lonien and Schwender 2009) and *Arabidopsis* cell suspension cultures (Williams *et al.*, 2010). The problem is observed to be more pronounced in tissues with very slow rate of metabolism. For example, under phosphorus starvation (Chapter 6), even some of the soluble amino acid pools have substantial contributions from unlabelled pre-existing metabolites after many days of ^{13}C feeding. This bias in the mass isotopomer fractions resulting from the influence of pre-existing unlabelled metabolite pools needs to be corrected in order to obtain reliable flux estimates. The strategies developed (Chapter 3) for accounting for the bias were used here to obtain reliable flux maps (Chapter 5, 6). Hence, isotopic steady state criteria can be slightly relaxed when dealing with plants in which large pools of stored metabolites and macromolecules are commonly found.

7.5 Stability of central metabolism to perturbations

Rational metabolic engineering in plants is constrained by the lack of comprehensive understanding of the metabolic changes associated with perturbation(s) such as genetic manipulations, nutritional stress or source of substrate. Studies on central metabolic fluxes in plants highlighted two major aspects of the effect of perturbations: (a) the stability of the network in which only minor or no changes could be observed in relative fluxes; and (b) significant alterations in metabolism (Kruger and Ratcliffe, 2009). ^{13}C MFA on *Arabidopsis* cell suspension under elevated O_2 conditions (Williams et al., 2008) showed no alteration in relative fluxes through central carbon metabolism implying the stability of the network. In

contrast differences in fluxes were observed in studies on *Brassica napus* embryos grown under different N source. When the embryos were supplied with glutamine the direction of flux in one of the TCA cycle reactions was from 2-oxoglutarate to citrate (Schwender et al., 2006), but when grown on inorganic nitrogen (Junker et al., 2007) the direction of flux was as expected in a traditional TCA cycle i.e., from citrate to 2-oxoglutarate. Also, differences in fluxes in plant central metabolism due to perturbations like temperature (Iyer et al., 2008; Williams et al., 2010) and hyperosmotic stress (Williams et al., 2010) have been observed.

In the current studies, it is observed that *Arabidopsis* suspension culture grown on different N sources has altered fluxes phenotypes, although the net flux through the oxPPP was unaffected. Under phosphate starvation (Pi-), varied flux phenotypes were obtained. However, the fluxes were stable under enhanced phosphate nutrition (Pi+). The physiological conditions, availability of resources and the capacity of the cells to meet the energy requirements for maintenance seem to be the major factors in maintaining the stability of the central network fluxes in plants.

7.6 Future directions

All the studies in this thesis show that ^{13}C-based steady state MFA is a powerful systems biology tool that has the capacity to define the metabolic flux phenotypes under perturbed conditions. Recent studies show how constraints-based flux balance analysis (FBA) is able to complement MFA (Williams et al., 2010). Also non-stationary 13C MFA method for the analysis of labelling time-courses in central metabolism should provide temporal information of fluxes during the course of growth. This method has been successfully demonstrated in microbes (Antoniewicz et al., 2007, Wahl et al., 2008, Noack et al., 2011); and could be

attempted in plants. Also flux models needs to be tested for the possibility of metabolite channeling (Williams et al., 2011). MFA in combination with other tools will be significant in better understanding of metabolic flux phenotypes.

8 Literature Cited

Allen DK, Ohlrogge JB, Shachar-Hill Y (2009) The role of light in soybean seed filling metabolism. Plant J **58:** 220-234

Allen DK, Ratcliffe RG (2009) Quantification of isotope label. In: Schwender, J, ed. Plant Metabolic Networks. New York, USA: Springer, 105-149

Allen DK, Shachar-Hill Y, Ohlrogge JB (2007) Compartment-specific labeling information in ^{13}C metabolic flux analysis of plants. Phytochem **68:** 2197-2210

Alonso AP, Goffman FD, Ohlrogge JB, Shachar-Hill Y (2007) Carbon conversion efficiency and central metabolic fluxes in developing sunflower (*Helianthus annuus* L.) embryos. Plant J **52**: 296-308

Alonso AP, Raymond P, Rolin D, Dieuaide-Noubhani M (2007) Substrate cycles in the central metabolism of maize root tips under hypoxia. Phytochem **68**: 2222–2231

Alonso AP, Vigeolas H, Raymond P, Rolin D, Dieuaide-Noubhani M (2005) A new substrate cycle in plants. Evidence for a high glucose-phosphate-to-glucose turnover from in vivo steady-state and pulse-labeling experiments with [^{13}C]glucose and [^{14}C]glucose. Plant Physiol **138:** 2220-2232

Antoniewicz MR, Kelleher JK, Stephanopoulos G (2006) Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metab Eng **8**: 324-337

Antoniewicz MR, Kelleher JK, Stephanopoulos G (2007) Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. Anal. Chem **79** : 7554-7559

Antoniewicz MR, Kelleher JK, Stephanopoulos G (2007) Elementary metabolite units (EMU): A novel framework for modeling isotopic distributions. Metab Eng **9**: 68–86

Antoniewicz MR, Kraynie DR, Laffend LA, Gonzalez-Lergier J, Kelleher JK, Stephanopoulos G (2007) Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of E. coli producing 1,3-propanediol. Metab Eng 9:277

Averill RH, Bailey-Serres J, Kruger NJ (1998) Co-operation between cytosolic and plastidic oxidative pentose phosphate pathways revealed by 6-phosphogluconate-deficient genotypes of maize. Plant J **14**: 449-457

Beard DA, Liang S, Qian H (2002) Energy balance for analysis of complex metabolic networks. Biophy J 83:79–86

Bergmeyer HU, Gawehn K, Grassl M (1974) in *Methods of Enzymatic Analysis* (Bergmeyer, H.U. ed) Volume I, Second Edition, 457-458, Academic Press, Inc., New York, NY

Bloom AJ, Burger M, Rubio-Asensio JS, Cousins AB (2010) Carbon dioxide enrichment inhibits nitrate assimilation in wheat and Arabidopsis. Science **328:** 899-903

Boatright J, Negre F, Chen X, Kish CM, Wood B (2004) Understanding in vivo benzenoid metabolism in petunia petal tissue. Plant Physiol **135**: 1993–2011

Bozzo GG, Dunn EL, Plaxton WC (2006) Differential synthesis of phosphate-starvation inducible purple acid phosphatase isozymes in tomato (*Lycopersicon esculentum*) suspension cells and seedlings, Plant Cell Environ. **29:** 303–313.

Britto DT, Kronzucker HJ (2002) NH4+ toxicity in higher plants: A critical review. J. Plant Physiol. **159:** 567–584

Carrari F, Urbanczyk-Wochniak E, Willmitzer L, Fernie AR (2003) Engineering central metabolism in crop species: learning the system. Metab Eng **5:** 191-200

Cheng CL, Acedo GN, Cristinsin M, Conkling M (1992) Sucrose mimics the light induction of Arabidopsis nitrate reductase gene transcription. Proc. Natl. Acad. Sci. **89:** 1861-1864

Cordell D, Drangert JO, White S (2009) The story of phosphorus: Global food security and food for thought. Global Environmental Change **19:** 292–305

Cossins E, Beevers H (1963) Ethanol metabolism in plant tissues. Plant Physiol **38:** 375-380

Dawson CJ, Hilton J (2011) Fertiliser availability in a resource-limited world: Production and recycling of nitrogen and phosphorus. Food Policy **36:** S14–S22

Dechorgnat J, Nguyen CT, Armengaud P, Jossier M, Diatloff E, Filleur S, Daniel-Vedele1 F (2011) From the soil to the seeds: the long journey of nitrate in plants. J. Exp Bot. 62: 1349-1359

Dieuaide-Noubhani M, Raffard G, Canioni P, Pradet A, Raymond P (1995) Quantification of compartmented metabolic fluxes in maize root-tips using isotope distribution from C13-labeled or C14-labeled glucose. J Biol Chem **270:** 13147–13159.

Droste P, Weitzel M, Wiechert W (2008) Visual exploration of isotope labeling networks in 3D Bioprocess Biosyst Eng **31:** 227–239

Drueckes P, Schinzel R, Palm D (1995) Photometric Microtiter Assay of Inorganic Phosphate in the Presence of Acid-Labile Organic Phosphates. Anal Biochem **230:** 173–177

Edwards JS, Covert M, Palsson BO (2002) Metabolic modelling of microbes: The flux-balance approach. Environ Microbiol **4:** 133–140

Garlick AP, Moore C, Kruger NJ (2002) Monitoring flux through the oxidative pentose phosphate pathway using [1-14C]gluconate. Planta **216:** 265–272

Goffman FD, Alonso AP, Schwender J, Shachar-Hill Y, Ohlrogge JB (2005) Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. Plant Physiol **138**: 2269–2279

Grafahrend-Belau E, Schreiber F, Koschützki D, Junker BH (2009) Flux balance analysis of barley seeds: a computational approach to study systemic properties of central metabolism. Plant Physiol **149**: 585–598

Gutiérrez RA, Lejay LV, Dean A, Chiaromonte F, Shasha DE, Coruzzi MN (2007) Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in Arabidopsis. Genome Biology **8:** R7

Harrison PW, Kruger NJ (2008) Validation of the design of feeding experiments involving [^{14}C]substrates used to monitor metabolic flux in higher plants Phytochem **69:** 2920-2927

Hinkle PC (2005) P/O ratios of mitochondrial oxidative phosphorylation. Biochim Biophys Acta **1706**: 1–11

Holzhütter HG (2004) The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. Eur J Biochem **271**: 2905–2922

Huang CY, Roessner U, Eickmeier I, Genc Y, Callahan DL, Shirley N, Langridge P, Bacic A (2008) Metabolite Profiling Reveals Distinct Changes in Carbon and Nitrogen Metabolism in Phosphate-Deficient Barley Plants (*Hordeum vulgare* L.). Plant Cell Physiol **49**: 691-703

Huege J, Sulpice R, Gibon Y, Lisec J, Koehl K, Kopka J (2007) GC-EI-TOF-MS analysis of in vivo carbon-partitioning into soluble metabolite pools of higher plants by monitoring isotope dilution after 13CO2 labelling. Phytochem **68**: 2258–2272

Isermann N, Wiechert W (2003) Metabolic isotopomer labeling systems. Part II: structural flux identifiability analysis. Math. Biosci **183**: 175–214

Iyer VV, Sriram G, Fulton DB, Zhou R, Westgate ME, Shanks, JV (2008) Metabolic flux maps comparing the effect of temperature on protein and oil biosynthesis in developing soybean cotyledons. Plant Cell Environ **31:** 506-517

Joshi M, Seidel-Morgenstern A, Kremling A (2006) Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. Metab Eng **8**: 447-455

Junker BH, Lonien J, Heady LE, Rogers A, Schwender J (2007) Parallel determination of enzyme activities and *in vivo* fluxes in *Brassica napus* embryos grown on organic or inorganic nitrogen source. Phytochem **68:** 2232-2242

Kauffman KJ, Prakash P, Edwards JS (2003) Advances in flux balance analysis. Curr Opin Biotechnol **14**: 491–496

Kelly EB, McMillen DP, Buono MJ (1990) Constructing confidence intervals using the bootstrap: an application to a multi-product cost function. The Review of Economics and Statistics **72**: 339-344

Kleijn RJ, Geertman JMA, Nfor BK, Ras C, Schipper D, Pronk JT, Heijnen JJ, van Maris AJA, van Winden WA (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:** 216-231

Kleijn RJ, van Winden WA, van Gulik WM, Heijnen JJ (2005) Revisiting the ^{13}C-label distribution of the non-oxidative branch of the pentose phosphate pathway based upon kinetic and genetic evidence. FEBS J **272:** 4970-4982

Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmuller E, Dormann P, Weckwerth W, Gibon Y, Stitt M, Willmitzer L, Fernie AR, Steinhauser D (2005) GMD@CSB.DB: the Golm Metabolome Database. Bioinformatics **21:** 1635-1638

Kraiser T, Gras DE, Gutiérrez AG, González B, Gutiérrez RA (2011) A holistic view of nitrogen acquisition in plants *J. Exp. Bot.* **62**: 1455-1466

Krook J, Vreugdenhil D, Dijkema C, van der Plas LHW (1999) Sucrose and starch metabolism in carrot (*Daucus carota* L.) cell suspensions analysed by ^{13}C-labelling: indications for a cytosol and a plastid-localized oxidative pentose phosphate pathway. J Exp Bot **49:** 1917-1924

Kruger NJ, Huddleston JE, Le Lay P, Brown ND, Ratcliffe RG (2007) Network flux analysis: impact of 13C-substrates on metabolism in *Arabidopsis thaliana* cell suspension cultures. Phytochem **68**: 2176–2188

Kruger NJ, Le Lay P, Ratcliffe RG (2007). Vacuolar compartmentation complicates the steady-state analysis of glucose metabolism and forces reappraisal of sucrose cycling in plants. Phytochem **68**: 2189–2196

Kruger NJ, Ratcliffe RG (2007) Dynamic metabolic networks: going with the flow. Phytochem **68:** 2136-2138

Kruger NJ, Ratcliffe RG (2008). Metabolic organization in plants: a challenge for the metabolic engineer. In: Bohnert HJ, Nguyen H, Lewis NG, eds. Advances in Plant Biochemistry and Molecular Biology, Volume 1: Bioengineering and Molecular Biology of Plant Pathways. Amsterdam, The Netherlands: Elsevier, 1-27

Kruger NJ, Ratcliffe RG (2009) Insights into plant metabolic networks from steady-state metabolic flux analysis. Biochimie **91**: 697-702

Kruger NJ, Ratcliffe RG, Roscher A (2003) Quantitative approaches for analysing fluxes through plant metabolic networks using NMR and stable isotope labelling. Phytochem Reviews **2**: 17–30

Kruger NJ, Troncoso-Ponce MA, Ratcliffe RG (2008) ^{1}H-NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. Nat Protocols **3:** 1001-1012

Kruger NJ, von Schaewen A (2003) The oxidative pentose phosphate pathway: structure and organisation. Curr Opin Plant Biol **6:** 236-246

Kruger NJ, Masakapalli SK, Ratcliffe RG (2012) Strategies for investigating the plant metabolic network with steady-state metabolic flux analysis: lessons from an Arabidopsis cell culture and other systems. J Exp Bot **63**, in press.

Lehmann M, Schwarzländer M, Obata T, Sirikantaramas S, Burow M, Olsen CE, Tohge T, Fricker MD, Møller BL, Fernie AR (2009) The metabolic response of Arabidopsis roots to oxidative stress is distinct from that of heterotrophic cells in culture and highlights a complex relationship between the levels of transcripts, metabolites and flux. Mol Plant **2**: 390–406

Leimer KR, Rice RH, Gehrke CW (1977) Complete mass spectra of N-trifluoroacetyl-n-butyl esters of amino acids. J Chromatogr **141:** 121-144

Libourel IG, Gehan JP, Shachar-Hill Y (2007) Design of substrate label for steady state flux measurements in plant systems using the metabolic network of *Brassica napus* embryos. Phytochem **68**: 2211–2221

Libourel IG, Shachar-Hill Y (2008) Metabolic flux analysis in plants: from intelligent design to rational engineering. Annu Rev Plant Biol **59**: 625-650

Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protocols **1:** 387-396

Llaneras F, Picó J (2008) Stoichiometric modelling of cell metabolism. J Biosci Bioeng **105**: 1-11

Lommen A (2009) MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Anal Chem **81:** 3079–3086

Lonien J, Schwender J (2009) Analysis of metabolic flux phenotypes for two *Arabidopsis* mutants with severe impairment in seed storage lipid synthesis. Plant Physiol **151**: 1617-1634

Lunn JE (2007) Compartmentation in plant metabolism. J Exp Bot **58:** 35-47

Malone JG, Mittova V, Ratcliffe RG, Kruger NJ (2006) The response of carbohydrate metabolism in potato tubers to low temperature. Plant Cell Physiol **47:** 1309-1322

Marx A, de Graaf AA, Wiechert W, Eggeling L, Sahm H (1996) Determination of the fluxes in central metabolism of Corynebacterium glutamicum by NMR spectroscopy combined with metabolite balancing. Biotechnol. Bioeng **49:** 111-129

Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG (2010) Subcellular flux analysis of central metabolism in a heterotrophic *Arabidopsis thaliana* cell suspension using steady-state stable isotope labeling. Plant Physiol **152:** 602-619

May M, Leaver C (1993) Oxidative stimulation of glutathione synthesis in *Arabidopsis thaliana* suspension cultures. Plant Physiol **103:** 621–627

McCormick AJ, Cramer MD, Watt DA (2008) Regulation of photosynthesis by sugars in sugarcane leaves. Journal of Plant Physiol **165:** 1817–1829

Miyashita Y, Dolferus R, Ismond KP, Good AG (2007) Alanine aminotransferase catalyses the breakdown of alanine after hypoxia in *Arabidopsis thaliana*. Plant J **49:** 92–102

Mollney M, Wiechert W, Kownatzki D, de Graaf AA (1999) Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. Biotechnol. Bioeng **66:** 86–103

Morcuende R, Bari R, Gibon Y, Zheng W, Pant BD, Bläsing O, Usadel B, Czechowski T, Udvardi MK, Stitt M, Scheible WR (2007) Genome-wide reprogramming of metabolism and regulatory networks of *Arabidopsis* in response to phosphorus. Plant Cell Environ. **30:** 85–112

Namba A (2004) Simulation studies on bootstrap empirical likelihood tests. Communications in Statistics – Simulation and Computation **33:** 99-108

Nielsen J, Villadsen J, Lidén G (2003) Bioreaction Engineering Principles, Ed 2. Springer, New York, pp 124-142

Noack S, Nöh K, Moch M, Oldiges M, Wiechert W (2011) Stationary versus non-stationary ^{13}C-MFA: A comparison using a consistent dataset. J Biotech, **154:** 179–190

Nunes-Nesi A, Sulpice R, Gibon Y, Fernie AR (2008) The enigmatic contribution of mitochondrial function in photosynthesis. J Exp Bot **59:** 1675–1684

Patterson K, Cakmak T, Cooper A, Lager I, Rasmusson AG, Escobar MA (2010) Distinct signaling pathways and transcriptome response signatures differentiate ammonium- and nitrate-supplied plants. Plant Cell Environ **33:** 1486-1501

Plaxton WC, Tran HT (2011) Metabolic Adaptations of Phosphate-Starved Plants. Plant Physiol **156:** 1006-1015

Poolman MG, Miguet L, Sweetlove LJ, Fell DA (2009) A genome-scale metabolic model of *Arabidopsis thaliana* and some of its properties. Plant Physiol 151:1570-1581

Press WH, Teukolsky SA, Vetterling WT , Flannery BP (1992) Numerical Recipes in C: The Art of Scientific Computing, Ed 2. Cambridge University Press, Cambridge, UK

Ratcliffe RG ,Shachar-Hill Y (2006) Measuring multiple fluxes through plant metabolic networks. Plant J **45:** 490-511

Ratcliffe RG, Shachar-Hill Y (2001) Probing plant metabolism with NMR. Annu. Rev. Plant Physiol. Plant Mol. Biol **52:** 499–526

Ratcliffe RG, Shachar-Hill Y (2005) Revealing metabolic phenotypes in plants: inputs from NMR analysis. Biol Rev **80:** 27-43

Rios-Estepa R, Lange BM (2007) Experimental and mathematical approaches to modelling plant metabolic networks. Phytochem **68:** 2351–2374

Rontein D, Dieuaide-Noubhani M, Dufourc EJ, Raymond P, Rolin D (2002) The metabolic architecture of plant cells. Stability of central metabolism and flexibility of anabolic pathways during the growth cycle of tomato cells. J. Biol. Chem **277:** 43948–43960

Roscher A, Kruger NJ, Ratcliffe RG (2000) Strategies for metabolic flux analysis in plants using isotope labelling. J. Biotechnol. **77:** 81–102

Saheki S, Takeda A, Shimazu T (1985) Assay of inorganic phosphate in the Mild pH Range, suitable for measurement of glycogen phosphorylase activity. Anal Biochem **148:** 277-281

Sauer U (2004) High-throughput phenomics: experimental methods for mapping fluxomes. Curr. Opin. Biotechnol. **15:** 58–63

Scharte J, Schön H, Tjaden Z, Weis E, von Schaewen A (2009) Isoenzyme replacement of glucose-6-phosphate dehydrogenase in the cytosol improves stress tolerance in plants. Proc Natl Acad Sci **106:** 8061–8066

Scheible WR, Morcuende R, Czechowski T, Fritz C, Osuna D, Palacios-Rojas N, Schindelasch D, Thimm O, Udvardi MK, Stitt M (2004) Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen. Plant Physiol **136:** 2483-2499

Schilling CH, Letscher D, Palsson BØ (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. J. Theor. Biol **203:** 229–248

Schmidt K, Carlsen M, Nielsen J, Villadsen J (1997) Modelling isotopomer distribution in biochemical networks using isotopomers mapping matrices. Biotechnol. Bioeng. **55:** 831-840

Schuster S, Dandekar T, Fell DA (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. Trends Biotechnol **17:** 53–60

Schuster S, Pfeiffer T, Fell DA (2008) Is maximization of molar yield in metabolic networks favoured by evolution? J Theor Biol **252:** 497-504

Schwender J (2008) Metabolic flux analysis as a tool in metabolic engineering of plants. Curr Opin Biotechnol **19:** 131-137

Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y (2004) Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. Nature **432**: 779–782

Schwender J, Ohlrogge JB (2002) Probing in vivo metabolism by stable isotope labeling of storage lipids and proteins in developing Brassica napus embryos. Plant Physiol **130**: 347–361

Schwender J, Ohlrogge JB, Shachar-Hill Y (2003) A flux model of glycolysis and the oxidative pentosephosphate pathway in developing *Brassica napus* embryos. J Biol. Chem **278:** 29442-29453

Schwender J, Shachar-Hill Y, Ohlrogge JB (2006) Mitochondrial metabolism in developing embryos of *Brassica napus*. J. Biol. Chem **281**: 34040–34047

Scott P, Kruger NJ (1995) Influence of elevated fructose-2,6-bisphosphate levels on starch mobilization in transgenic tobacco leaves in the dark. Plant Physiol **108**: 1569-1577

Selivanov VA, Meshalkina LE, Solovjeva ON, Kuchel PW, Ramos-Montoya A, Kochetov GA, Lee PWN, Cascante M (2005) Rapid simulation and analysis of isotopomer distributions using constraints based on enzyme mechanisms: an example from HT29 cancer cells. Bioinformatics **21:** 3558-3564

Spielbauer G, Margl L, Hannah LC, Romisch W, Ettenhuber C, Bacher A, Gierl A, Eisenreich W, Genschel U (2006) Robustness of central carbohydrate metabolism in developing maize kernels. Phytochem **67**:1460–1475

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou R, Westgate ME, Spalding MH, Shanks JV (2004) Quantification of compartmented metabolic fluxes in developing soybean embryos by employing biosynthetically directed fractional (13)C labeling, two-dimensional [(13)C, (1)H] nuclear magnetic resonance, and comprehensive isotopomer balancing. Plant Physiol **136**: 3043-3057

Sriram G, Gonzalez-Rivera O, Shanks JV (2006) Determination of biomass composition of *Catharanthus roseus* hairy roots for metabolic flux analysis. Biotechnol Prog **22:** 1659–1663

Stitt M, ap Rees T (1980) Estimation of the activity of the oxidative pentose phosphate pathway in pea chloroplasts. Phytochem **19**: 1583–1585

Sweetlove LJ, Fell D, Fernie AR (2008) Getting to grips with the plant metabolic network. Biochem J **409:** 27-41

Sweetlove LJ, Fernie AR (2005) Regulation of metabolic networks: understanding metabolic complexity in the systems biology era. New Phytologist **168**: 9–24

Sweetlove LJ, Last RL, Fernie AR (2003) Predictive metabolic engineering: a goal for systems biology. Plant Physiol **132**: 420-425

Sweetlove LJ, Ratcliffe RG (2011). Flux-balance modelling of plant metabolism. Frontiers in Plant Science **2:** 38

Szyperski T (1998) ^{13}C-NMR, MS and metabolic flux balancing in biotechnology research. Quart. Rev. Biophys. **31:** 41–106

Tang Y, Pingitore F, Mukhopadhyay A, Phan R, Hazen TC, Keasling JD (2007) Pathway confirmation and flux analysis of central metabolic pathways in *Desulfovibrio vulgaris* hildenborough using gas chromatography-mass spectrometry and Fourier transform-ion cyclotron resonance mass spectrometry. J. Bacteriol **189:** 940–949

Troufflard S, Roscher A, Thomasset B, Barbotin JN, Rawsthorne S, Portais JC (2007) In vivo ^{13}C NMR determines metabolic fluxes and steady state in linseed embryos. Phytochem **68:** 2341-2350

Truesdale GA, Downing AL (1954) Solubility of oxygen in water. Nature **173:** 1236
Turner JG, Novacky A (1974) The quantitative relation between plant and bacterial cells involved in the hypersensitive reaction. Phytopathology **64:** 885 890

Urbanczyk-Wochniak E, Fernie AR (2005) Metabolic profiling reveals altered nitrogen nutrient regimes have diverse effects on the metabolism of hydroponically-grown tomato (*Solanum lycopersicum*) plants J. Exp. Bot. **56:** 309-321

Vallino JJ, Stephanopoulos G (1993) Metabolic flux distribution in Corynebacterium glutamicum during growth and lysine overproduction. Biotechnol. Bioeng. **41:** 633-646

van Winden WA, Heijnen JJ, Verheijen PJT (2002) Cumulative bondomers: A new concept in flux analysis from 2D [C-13,H-1] COSY NMR data. Biotechnol. Bioeng. **80:** 731–745

van Winden WA, Heijnen JJ, Verheijen PJT, Grievink J (2001) A priori analysis of metabolic flux identifiability from 13C-labeling data. Biotechnol. Bioeng. **74:** 505–516

Varma A, Palsson BO (1994) Metabolic flux balancing—Basic concepts, scientific and practical use. Bio-Technology **12:** 994–998

Wahl SA, Dauner M, Wiechert W (2004) New tools for mass isotopomer data evaluation in ^{13}C flux analysis: mass isotope correction, data consistency checking, and precursor relationships. Biotechnol Bioeng **85:** 259-268

Wahl SA, **Nöh K**, **Wiechert W** (2008) ^{13}C labeling experiments at metabolic nonstationary conditions: An exploratory study. *BMC Bioinformatics* **9:** 152

Wakao S, Andre C, Benning C (2008) Functional analyses of cytosolic glucose-6-phosphate dehydrogenases and their contribution to seed oil accumulation in Arabidopsis. Plant Physiol **146:** 277-288

Weitzel M, Wiechert W, Nöh K (2007) The topology of metabolic isotope labeling networks. BMC Bioinfo **8:** 315

Wiback SJ, Famili I, Greenberg HJ, Palsson BØ (2004) Monte Carlo sampling can be used to determine the size and shape of the steady-state flux space. J Theor Biol **228:** 437-447

Wiback SJ, Mahadevan R, Palsson BO (2004) Using metabolic flux data to further constrain the metabolic solution space and predict internal flux patterns: The *Escherichia coli* spectrum. Biotechnol. and Bioeng **86:** 317–331

Wiechert W (2001) ^{13}C Metabolic Flux Analysis. Metab. Eng.**3:** 195-206

Wiechert W, de Graaf AA (1996) In vivo stationary flux analysis by 13C labelling experiments. Adv Biochem Eng Biotechnol **54:** 109-154

Wiechert W, de Graaf AA (1997) Bidirectional reaction steps in metabolic networks.I. Modeling and simulation of carbon isotope labeling experiments. Biotechnol Bioeng. **55:** 101–117

Wiechert W, Mollney M, Isermann N, Wurzel W, de Graaf AA (1999) Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems. Biotechnol. Bioeng **66:** 69–85

Wiechert W, Mollney M, Petersen S, de Graaf AA (2001) A universal framework for ^{13}C metabolic flux analysis. Metab. Eng **3:** 265–283

Wiechert W, Siefke C, de Graaf AA, Marx A (1997) Bidirectional reaction steps in metabolic networks.II. Flux estimation and statistical analysis. Biotechnol. Bioeng **55:** 118–135

Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG (2008) Metabolic network fluxes in heterotrophic Arabidopsis cells: stability of the flux distribution under different oxygenation conditions. Plant Physiol **148:** 704-718

Williams TCR, Poolman MG, Howden AJ, Schwarzlander M, Fell DA, Ratcliffe RG, Sweetlove LJ (2010) A genome-scale metabolic model accurately predicts fluxes in central carbon metabolism under stress conditions. Plant Physiol **154:** 311-323

Williams TCR, Sweetlove LJ, Ratcliffe RG (2011) Capturing metabolite channeling in metabolic flux phenotypes. *Plant Physiol* **157:** 981-984

Wittmann C, Hans M, Heinzle E (2002) In vivo analysis of intracellular amino acid labeling by GC/MS. Anal. Biochem. **307:** 379-382

Zupke C, Stephanopoulos G (1994) Modeling of isotope distributions and intracellular fluxes in metabolic networks using atom mapping matrices. Biotechnol Prog **10:** 489-498