Monitoring Primary Carbon Metabolism in Plants using Heavy Isotope Labelling and Mass Spectrometry

$^{13}CO_2$ Labelling, Detection and Estimation in Intact Plants

Pernilla Lindén
Faculty of Forest Sciences
Department of Forest Genetics and Plant Physiology
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2015

Acta Universitatis agriculturae Sueciae
2015:58

Cover: Word cloud of this thesis
(Designed by Pernilla Lindén using https://tagul.com/)

ISSN 1652-6880
ISBN (print version) 978-91-576-8314-4
ISBN (electronic version) 978-91-576-8315-1

Print: Arkitektkopia, Umeå 2015

Monitoring Primary Carbon Metabolism in Plants using Heavy Isotopes and Mass Spectrometry

Abstract

This thesis covers the possibilities and limitations of studying primary metabolism in intact plants, with special focus on heavy isotope labelling and mass spectrometry methodology. In paper I, a series of *Arabidopsis thaliana* mutants lacking one or both genes of mitochondrial malate dehydrogenase (mMDH) were characterised. We found that mMDH has a complex respiration-controlling role. In paper II, we continued the study of the single mutant, *mmdh1*. We developed a method using $^{13}CO_2$ to label whole plants and analyse their metabolic profiles by gas and liquid chromatography coupled to mass spectrometry (GC- and LC-MS). The results show that *mmdh1* had a metabolic phenotype that revealed an altered flux through photorespiration, especially under low CO_2 conditions. Combining incorporation data with metabolite pool size deepened our understanding regarding the role of mMDH in photorespiration, respiration, and cellular redox balance. The practical and theoretical aspects of ^{13}C-labelling of plants learned from this study were used for designing the experimental setup of paper III, a labelling study of developing wood in hybrid aspen. The labelling strategy, time resolution, and sampling had to be adapted to suit woody plants.

Two months old trees were labelled with $^{13}CO_2$ in a 4 h burst and then sampling of source leaves, stem phloem, and developing wood continued over 24 hours. Since sucrose is the main carbon transporter the analysis was concentrated around this metabolite. We found previously unrecognised temporal patterns in wood biosynthesis and an indication that the diurnal cycle serves as a cue in the regulation of carbon allocation in developing wood.

To study systems with complex structures such as plants, it is informative to study specific cell types rather than whole plants or whole organs. In paper IV, a method was developed for analysis of metabolite profiles in cell type specific cells. Isolated protoplast from Arabidopsis roots were sorted by fluorescence-activated cell sorting (FACS) and their metabolite profiles were analysed by GC-MS. The method is fast, robust and reliable for analysis of cell type specific metabolic profiling and a beginning to meet the call for new metabolite analysis techniques with a high spatial resolution.

Keywords: Plant Primary Carbon Metabolism, Mass Spectrometry, Heavy Isotope Labelling, $^{13}CO_2$, Photorespiration, Metabolomics, Carbon Allocation, Cell Sorting, Multivariate Statistical Analysis

Author's address: Pernilla Lindén, SLU, Department of Forest Genetics and Plant Physiology, 901 83 Umeå, Sweden
E-mail: Pernilla.Linden@slu.se

Dedication

To You who helped me reach this far. I am forever grateful.

What you get by achieving your goals is not as important as what you become by achieving your goals.

Henry David Thorau

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Tomaz T*, Bagard M*, Pracharoenwattana I, **Lindén P**, Lee CP, Carroll AJ, Stroher E, Smith SM, Gardestrom P, Millar AH (2010). Mitochondrial Malate Dehydrogenase Lowers Leaf Respiration and Alters Photorespiration and Plant Growth in Arabidopsis. *Plant Physiology* 154: 1143-1157

II **Lindén P**, Keech O, Stenlund H, Gardeström P, Moritz T. Reduced mitochondrial malate dehydrogenase activity has a strong effect on photorespiratory metabolism as revealed by ^{13}C-labelling (manuscript).

III Mahboubi A*, **Lindén P*,** Hedenström M, Moritz T, Niittylä T (2015). Carbon-13 tracking after $^{13}CO_2$ supply revealed diurnal patterns of wood formation in aspen. *Plant Physiology* in press.

IV Petersson S*, **Lindén P***, Moritz T, Ljung K. Cell-type specific metabolic profiling of *Arabidopsis thaliana* protoplasts as a tool for plant systems biology (submitted).

*To be considered as joint first authors.

Other papers by the author not included in this thesis:

Chawade A, **Lindén P**, Brautigam M, Jonsson R, Jonsson A, Moritz T, Olsson O (2012). Development of a model system to identify differences in spring and winter oat. *PLoS One*, 7(1), pp. e29792-e29792

The contribution of Pernilla Lindén to the papers included in this thesis was as follows:

I Performed the initial plant physiology work, for example established growth phenotype and germination rate under different growth conditions. Also involved in the mitochondria isolation and enzyme assays and oxygen consumption assay. Assisted with MS-analysis and had a minor part in manuscript writing.

II Planned and executed all experiments. Developed LC- and GC-MS methods. Active involvement in the design and construction of the labelling chamber. Performed data processing, metabolite identification, labelling estimations, and multivariate statistics. The majority of the manuscript including all figures and tables.

III Assisted with planning and evaluation of the labelling experiments. Developed the LC-MS method for labelled sugar phosphates and phenylalanine. Planned and ran the metabolite extraction and multivariate analysis (both LC- and GC-MS), data processing, labelling estimations, and statistics. Wrote part of materials and methods, took an active part in revision of the manuscript and made several figures.

IV Planned and executed metabolite analysis (GC-MS), data processing and metabolite identification, multivariate statistics. Took active part in planning and writing of manuscript. Generated most of the figures.

Abbreviations

2-PG	2-phosphoglycolate
3-PGA	3-phosphoglycerate
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
CBC	Calvin-Benson cycle
CE	Capillary Electrophoresis
F6P	Fructose-6-phosphate
FACS	Fluorescence-activated cell sorting
$FADH_2$	Flavin adenine dinucleotide (hydroquinone form)
FBA	Flux balance analysis
G1P	Glucose-1-phosphate
G6P	Glucose-6-phosphate
GAP	Glyceraldehyde-3-phosphate
GC	Gas chromatography
GFP	Green fluorescent protein
GLDH	Glutamate dehydrogenase
GS/GOGAT	Glutamine synthetase/ glutamine oxoglutarate aminotransferase
HP	Hydroxypyruvate
INST-MFA	Non-stationary metabolic flux analysis
LC	Liquid chromatography
LCM	Laser capture micro-dissection
LTQ	Linear ion trap
MDH	Malate dehydrogenase
MFA	Metabolic flux analysis
mMDH	Mitochondrial malate dehydrogenase
MRM	Multiple reaction monitoring
MS	Mass spectrometry

MSTFA	N-Methyl-N-(trimethylsilyl) trifluoroacetamide
NAD^+	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)
NAD-ME	NAD-malic enzyme
NADPH	Nicotinamide adenine dinucleotide phosphate
NDP-sugars	Nucleotide sugars
NMR	Nuclear magnetic resonance
OAA	Oxaloacetate
OPLS-DA	Orthogonal projection to latent structures-discriminant analysis
PCA	Principal component analysis
PEP	Phosphoenolpyruvate
P_i	Inorganic phosphate
PLS	Partial least squares
PLS-DA	PLS-discriminant analysis
PPP	Pentose phosphate pathway
Q	Quadrupole
QqQ	Triple quadrupole
RNA	Ribonucleic acid
RSD	Relative standard deviation
Ru5P	Ribulose-5-phosphate
Rubisco	Ribulose bisphosphate carboxylase/oxygenase
RubP	Ribulose-1,5-bisphosphate
TCA-cycle	Tricarboxylic acid cycle
TMS	Trimethylsilyl
TOF	Time of flight
UDP-glucose	Uridine diphosphate glucose

1 Introduction

This thesis summarises the work of my PhD project during the last four and a half years. My motivation for doing plant research is my curiosity and the hope of making a difference. The future is unwritten, but most certainly the world population will continue to increase for many years to come and so will global climate change. Thus, we will need to address these issues and I consider plant biology and biotechnology to have high potential in helping us to do so. We have powerful and precise plant molecular tools to our aid which have been thoroughly tested and proven safe. It saddens me that Europe refuses to embrace the possibilities that these tools present, effectively blocking the possibility of developing clean, sustainable, and yet efficient agriculture for the future.

This is one thesis among many from the Swedish University of Agricultural Sciences (SLU) published this year. Although, it might be a small contribution, seeing to the science community as a whole, it is not the accomplishment of one single scientist that will make a difference but rather what we can achieve together.

1.1 Plant Metabolism in Brief

We all learned in school that plants use light, water, and carbon dioxide from the air to produce sugar and oxygen. What actually takes place during this process, known as photosynthesis, and which metabolic pathways that are connected to it, is more complex and less common knowledge. I will give a brief introduction to the metabolic pathways that are most relevant for this thesis, namely, photosynthesis, carbon assimilation, photorespiration, and respiration.

Plants and other autotrophs have the unique ability to convert physical and chemical energy sources into carbohydrates in the absence of organic substrates. In other words, plants can convert atmospheric CO_2 into a form of carbon (-CHOH-) that can be used by the cell. This is the essence of photosynthesis. Photosynthesis is divided into two different stages. The first stage is the light dependent reaction where light energy is harvested in special reaction centres in the chloroplast. Water is split to donate electrons to produce high energy compounds, adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH), while O_2 is formed as a side product. The second stage of photosynthesis, carbon assimilation, is driven by the energy harvested from the light reactions and starts with the Calvin-Benson cycle which takes place in the chloroplast (figure1). In the first step of the Calvin-Benson cycle, CO_2 and water are enzymatically combined with the acceptor molecule ribulose-1,5-bisphosphate (RubP), generating two molecules of 3-phosphoglycerate (3-PGA). 3-PGA is in turn reduced to glyceraldehyde-3-phosphate (GAP), using photochemically produced energy and reducing equivalents in the form of ATP and NADPH. The regeneration of RubP completes the cycle (Taiz & Zeiger, 2002). Thus, the Calvin-Benson cycle can be divided into three phases; carboxylation, reduction and regeneration. The triose phosphates produced by the Calvin-Benson cycle exit the chloroplast and are further metabolised in the cytosol to sucrose. In the light, hexose phosphates derived from the Calvin-Benson cycle are used to synthesise starch in the chloroplast, where it is stored to be metabolised during the dark period. The primary carbon transporter in higher plants is sucrose. Sucrose is loaded from the source tissues into the phloem, transported via the phloem to sink cells/tissues where it is imported and metabolised. Many of these steps are symplastic transport, *i.e.* sucrose moves passively from cell-to-cell through the plasmodesmata. However, there are steps where sucrose is released into the apoplast and its cell-to-cell transport requires active transport (Sauer, 2007).

The enzyme ribulose bisphosphate carboxylase/oxygenase (Rubisco) catalyses the first reaction in the Calvin-Benson cycle where CO_2 is merged with RuBP, yielding two 3-PGA molecules. Rubisco is the most abundant enzyme in leaves, representing about 40 % of the soluble proteins (Taiz & Zeiger, 2002). As the name suggests, Rubisco has affinity for both CO_2 and O_2. Although CO_2 is highly favoured, the concentration of atmospheric O_2 is about 21 %, to be compared with a CO_2 concentration of about 0.04 %. At these concentrations Rubisco has approximately three times higher affinity for CO_2 than for O_2.

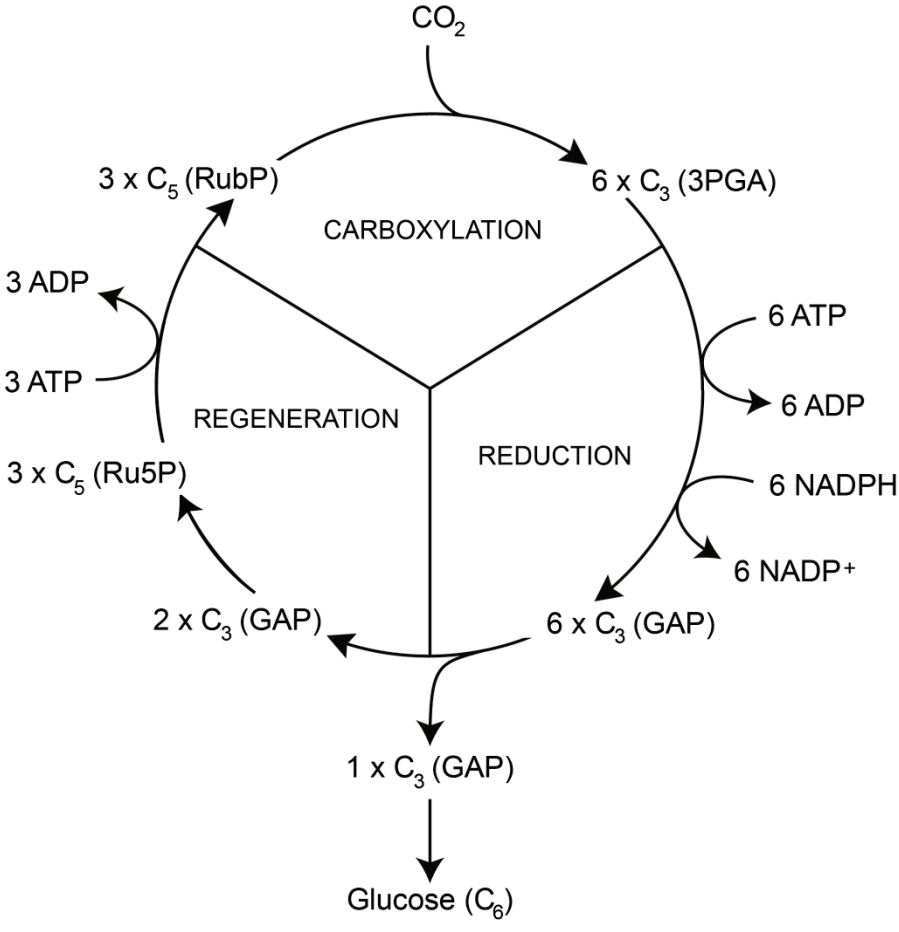


Figure 1. The Calvin-Benson cycle can be divided into three phases; 1) Carboxylation- where CO_2 is fixed to an acceptor molecule; 2) Reduction- where photochemically derived energy is used to form carbohydrates; 3) Regeneration- where the CO_2 acceptor molecule RubP is reformed. Abbreviations: 3PGA; 3-phosphoglycerate, GAP; glyceraldehyde-3-phosphate, Ru5P; ribulose-5-phosphate, RubP; ribulose-1,5-bisphosphate.

The oxygenation reaction is the beginning of photorespiration, one of the major pathways in plant metabolism. In oxygenation Rubisco produces one 3-PGA molecule and one 2-phosphoglycolate (2-PG) molecule. 2-PG is toxic and inhibits different enzymes involved in central carbon metabolism. The major function of photorespiration is thought to be to recycle and detoxify 2-PG to produce 3-PGA. Photorespiration is highly complex and involves four cell compartments; the chloroplast, the peroxisome, the mitochondria, and the cytosol. In the chloroplast 2-PG is hydrolysed to glycolate which is moved into the peroxisome. There it is oxidised to glyoxylate, which in turn is converted to

glycine and transported to the mitochondria. Here two glycine molecules are merged into one serine molecule, in the process NAD^+ is reduced to NADH and ammonia and CO_2 are also produced. The produced ammonia is used in the chloroplast in a later step of the pathway. Serine is transported back to the peroxisome where it donates its amino group to glyoxylate, forming hydroxypyruvate (HP). HP is reduced to glycerate and NADH is oxidised to NAD^+ in the process. Glycerate is transported to the chloroplast, where the final step in photorespiration takes place; glycerate is phosphorylated to 3-PGA which then re-enters the Calvin-Benson cycle (Florian *et al.*, 2013).

The described reactions are the carbon cycle of photorespiration but there is also a parallel nitrogen cycle (figure 2). In the photorespiratory nitrogen cycle the ammonium released from glycine decarboxylation in mitochondria is re-assimilated in chloroplasts by the glutamine synthetase/ glutamine oxoglutarate aminotransferase (GS/GOGAT) system. First, the ammonium is fixed by GS when glutamate is converted to glutamine in an ATP-dependent reaction. Then one glutamine and one α-ketoglutarate form two glutamate catalysed by GOGAT using reduced ferredoxin as reductant. Glutamate is then used in the peroxisome to transaminate glyoxylate to glycine. This is sufficient to form half of the glycine in the photorespiratory cycle. The amino group for the other half of glyoxylate to glycine conversion is supplied from serine when it is converted to hydroxypyruvate. Thus, the photorespiratory nitrogen cycle also involves reactions in chloroplasts, peroxisomes, and mitochondria linking photorespiration closely to nitrogen assimilation through the GS/GOGAT system which is common for the two processes (Maurino & Peterhansel, 2010).

Respiration is the metabolic process that supplies the cell with energy and carbon precursors for biosynthesis by the breakdown of carbon compounds. The plant cell can use many different carbon sources as substrates, *e.g.* sucrose, hexose and triose phosphates (from photosynthesis and starch degradation), fructans and other sugars, lipids, organic acids, and occasionally even proteins. The overall reaction of respiration can be viewed as the reversal of photosynthesis, completely oxidising sucrose to CO_2, while O_2, as the electron acceptor, is reduced to water.

The large amount of energy released from respiration is distributed over a series of reactions. These reactions can be organised into four major metabolic processes; glycolysis, the tricarboxylic acid (TCA) cycle, the pentose phosphate pathway (PPP), and oxidative phosphorylation. Glycolysis starts with a sugar that is oxidised in steps to produce pyruvate. The process yields small amounts of energy as ATP and reducing power as NADH. It mainly takes place in the cytosol, although most of the enzymes of the pathway are found in the chloroplast too. The oxidative PPP also takes place in the cytosol.

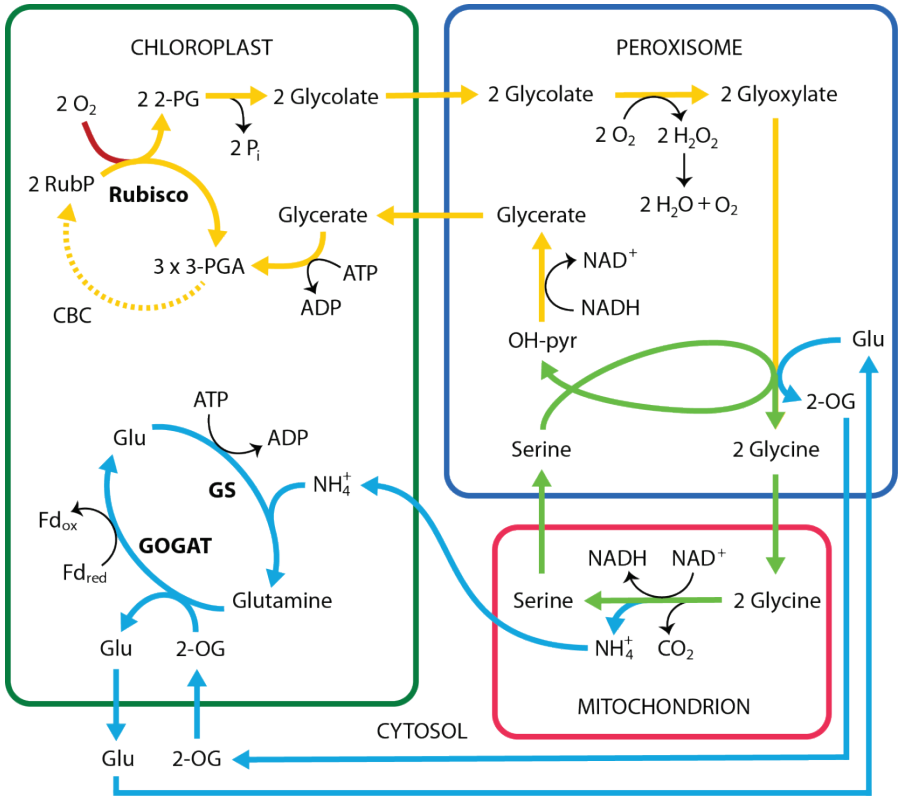


Figure 2. The carbon and nitrogen cycles of photorespiration starting with Rubisco binding oxygen to RubP, forming 2-PG. The Carbon cycle is indicated with yellow arrows and the nitrogen cycle with blue arrows. Green arrows indicate where the cycles coincide. Abbreviations: CBC; Calvin-Benson cycle, 2-PG; 2-phosphoglycolate, 3-PGA; 3-phosphoglycerate, Glu; Glutamate, RubP; ribulose-1,5-bisphosphate; OH-pyr; hydroxypyruvate, 2-OG; α-ketoglutarate, GC; Glutamine synthetase, GOGAT; Glutamine oxoglutarate aminotransferase.

In this process glucose-6-phosphate (G6P) is oxidised to ribulose-5-phosphate (Ru5P), releasing one molecule of CO_2 and two molecules of NADPH. Ru5P is in turn converted into other forms of C_5 sugars. In the TCA-cycle, which takes place in the mitochondria, pyruvate is oxidised to CO_2, in the process releasing large amounts of reducing power. For every pyruvate molecule oxidised, four NADH, one $FADH_2$ and one ATP molecule are released. The total yield from one pyruvate molecule is 15 ATP molecules. The mitochondrial respiration has an additional role in supplying carbon precursors for biosynthetic reactions, for example, α-ketoglutarate from the TCA-cycle is used for nitrogen assimilation. The final step of the oxidation of pyruvate is called oxidative phosphorylation

and involves the electron transport chain and the synthesis of ATP from ADP and inorganic phosphate (P_i) by ATP synthase. The electron transport chain takes place in the inner mitochondrial membrane and transfers electrons from NADH and $FADH_2$ to oxygen. The process releases substantial amounts of energy used to synthesise ATP (Taiz & Zeiger, 2002).

1.2 Metabolomics

A metabolite is a low-molecular-weight substance in an organism. Metabolism is the chemical reactions in the organism where enzymes aid in turning one metabolite into another. Metabolite reactions are organised into metabolic pathways. A metabolome is the complete set of metabolites in an organism (Oliver *et al.*, 1998). Metabolomics is the characterisation of the metabolome. While genomes can be sequenced for an organism, it is currently impossible to analyse a complete metabolome. This is due to the greater molecular complexity of metabolites compared to genes, which can be described by four letter codes or proteins that are defined by 20 amino acids (Fiehn, 2002). Even estimates of the number of metabolites in a well-studied model such as Arabidopsis are uncertain, and ranges from 5000- 25 000 (Trethewey, 2004; von Roepenack-Lahaye *et al.*, 2004; Huppe & Turpin, 1994). Looking at the plant kingdom in general the estimates become even more inexact, ranging from 50 000 to 200 000 (Fiehn, 2002; Pichersky & Gang, 2000). To analyse metabolites one must look at their elemental composition, structure (*e.g.* functional groups and polarity) and chemical properties. This comes with several analytical challenges. First, the chemical properties of metabolites are diverse, which puts high demands on the extraction in order to accommodate the variation. Second, the concentration differences between metabolites can be huge, demanding analysis techniques where the low abundant metabolites are not lost among the high abundant ones. Third, the spatial distribution of the metabolites *e.g.* specific organs, cell types, and sub cellular compartmentation. Fourth, the temporal distribution of the metabolites varies with the diurnal rhythm, season and lifespan of an organism. Fifth, genomic information cannot be used to aid in metabolite analysis (Hegeman, 2010).

To cope with this complexity there are several approaches of metabolite analysis and the definitions might vary between studies. In this work I use the definitions as described by Goodacre *et al.* (2004).

Metabolomics- analysis of all (or rather as many as possible) metabolites in a biological system.

Metabolic fingerprinting- analysis used for rapid screening and classification where identification and quantification of individual metabolites are not necessary.

Metabolic profiling- analysis of a set of metabolites that define or describe metabolic patterns for a group of metabolically or analytically related metabolites.

Metabolite target analysis- analysis of a limited number of metabolites or a compound class or metabolic pathway.

The different strategies can be divided in untargeted and targeted approaches. In untargeted metabolite analysis the study includes both identified and unknown metabolites. Metabolic fingerprinting and metabolomics are examples of untargeted approaches. If, on the other hand, the aim is to study known metabolites, this is referred to as targeted analysis. In this case there must be spectral reference information available, or it must be generated by running standards before the metabolites can be identified and quantified (Carter *et al.*, 2013; Kueger *et al.*, 2012). Metabolite profiling and metabolite target analysis are both targeted approaches.

The purpose with metabolomics research is to answer biological questions and thereby learn about the function of organisms and biological systems. Metabolite analysis can be applied to a range of studies, for example, as a diagnostic tool to investigate a plants metabolic response to a herbicide, or to compare genetically modified and conventional crop plants (to describe the response of plants to biotic or abiotic stresses) (Cook *et al.*, 2004), or to decipher gene function and understand metabolic regulation (Schauer & Fernie, 2006). Also, the combination of metabolite profiling and marker-assisted selection gives a higher understanding and possibility of influencing the chemical composition of crops (Fernie *et al.*, 2006).

1.3 Metabolite Analysis Techniques

There are many different techniques available for metabolite analysis. It is generally considered that several techniques used in combination are needed to get the best coverage of the metabolome. Common examples are; gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), capillary electrophoresis-mass spectrometry (CE-MS), nuclear magnetic resonance (NMR) spectroscopy and, less common, Fourier

transform-infrared (FT-IR) spectroscopy (Hiller K. *et al.*, 2011; t'Kindt *et al.*, 2009). NMR and FT-IR spectroscopy have lower sensitivity and selectivity compared to MS-based techniques. However, it is possible to separate samples on the basis of differences in their metabolite composition. Thus, they are good techniques for metabolite fingerprinting (Suzuki *et al.*, 2010). NMR has the advantages of being robust, reliable, and requiring little sample preparation. Another advantage is that it is non-destructive, which means that the sample can be analysed further. NMR provides both structural and quantitative information, which can be very useful for identification of unknown metabolites. Two-dimensional NMR (2D-NMR) offers a better selectivity than 1D NMR, but the sensitivity and the sample throughput (several hours per sample) are still low. 2D-NMR has been successfully used for analysis of ^{13}C-labelled biological samples (Giraudeau *et al.*, 2011).

A mass spectrometer consists of an ion source, a mass analyser, a detector and a computer system (figure 3). A sample is introduced by the sample inlet. A separation technique such as GC, LC or CE can be used. The sample can be directly injected or a direct probe technique can be used, like MALDI-TOF. For some ionisation techniques, such as ESI and APCI, the ionisation is performed under atmospheric pressure. In the ion source the molecules are ionised by being given a positive or negative charge. The ions continue to the mass analyser where the ions are separated according to m/z, after which they hit the detector that sends a translated signal to the computer system.

There are a range of options regarding MS techniques, including quadrupoles and ion traps which offer good sensitivity but are limited in their resolving power, or higher mass resolution instruments such as time-of-flight (TOF) or Orbitrap. Different mass spectrometers can be coupled together *e.g.* triple quadrupoles (QqQ), quadrupole-TOF (Q-TOF) or ion trap-Orbitrap (LTQ-Orbitrap) (Aharoni & Brandizzi, 2012). MS techniques have on the other hand, in combination with the right sample preparation and chromatographic separation, high sensitivity, specificity and a good dynamic range (Lu *et al.*, 2008). In this thesis LC-QqQ-MS and GC-TOF-MS techniques were combined for metabolite analysis. See Box 1 for a glossary of MS terms.

It is important to keep in mind that the reliability and suitability of sample preparation, data acquisition, data processing and data analysis are essential for correct biological interpretation in metabolomics studies. Technical variation can arise from many sources in a metabolomics study, from sampling, sample storage, extraction, derivatization, analysis and detection (Koek *et al.*, 2011). Therefore, it is important to take steps during analysis to be able to separate interesting biological variation from unwanted technical variation. One important first step is to rapidly quench metabolism at sampling. This can be

done by instantly freezing samples in liquid nitrogen, for example, by freeze clamping of an individual leaf (Hasunuma *et al.*, 2010), immersing the sample in liquid nitrogen (t'Kindt *et al.*, 2009; Huege *et al.*, 2007) or pouring liquid nitrogen over the whole plant rosette (Heise *et al.*, 2014; Szecowka *et al.*, 2013; Arrivault *et al.*, 2009). Samples should be stored at -80°C after freezing. It is important to keep the low temperature until proper additives are added during sample extraction, effectively stopping any further enzyme activity. Adding known concentrations of labelled standards (hereafter referred to as internal standards) to the extraction solution makes it possible to separate technical variation deriving from extraction and derivatization (for GC-MS) from biological variation. By adding an additional standard prior to analysis instrumental variation can be separated from other technical variation. The internal standards should preferably cover the chemical properties of the compounds to be analysed. For a GC-MS analysis, for example, it would be advised to have internal standards which cover amino acids, organic acids, sugars and fatty acids.

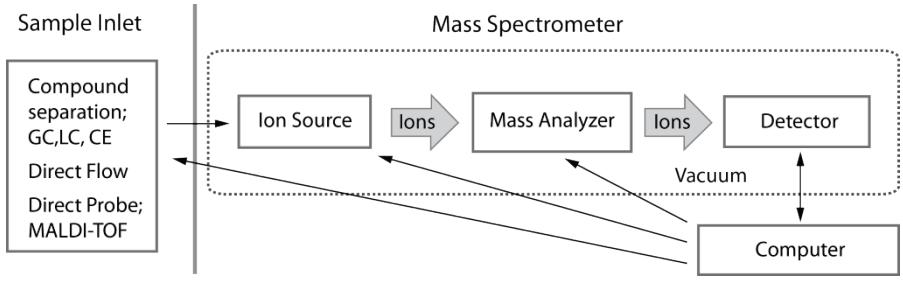


Figure 3. A mass spectrometer consists of an ion source, where the compounds are ionised; a mass analyser that separates ions with different mass-to-charge ratio (m/z); a detector that detects the ions and converts that to an electronic signal; and finally, a computer system that monitors and controls the different parts of the mass spectrometer. The area encircled with a dashed line indicates which part of the mass spectrometer that is operating under vacuum.

Box 1. Glossary of mass spectrometry terms. Many words in the glossary are extracted or adopted from the Dictionary of Mass Spectrometry by Mallet and Down 2010, published by WILEY.

Abundance- describes the relative occurrence of an ion.

Accuracy- the accuracy of a mass measurement is a measure of how close the obtained value is to the true value.

Atmospheric pressure chemical ionisation (APCI)- a gas phase CI process where the solvent acts as the CI reagent gas to ionise the sample.

Capillary electrophoresis (CE)- a liquid separation technique well suited for the separation of polar and charged compounds, since compounds are separated on the basis of their charge-to-mass ratio. Separation is performed in a buffer solution held in a length of silica capillary tubing. This configuration leads to very narrow peak widths and the dimensions of the tubing permit efficient cooling when high voltages are applied. Can be coupled to MS with EI, and requires a fast scanning mass analyser.

Capillary electrophoresis- mass spectrometry (CE-MS)- an advantage with this method is the ability to detect ionic metabolites such as sugar phosphates, nucleotides, amino acids and organic acids. A limitation is the poor concentration sensitivity due to the small sample volume (nanoliters). Coupled to MS this technique provides selective detection and structural information of metabolites.

Charge number (z)- the number of positive or negative charges on the ion.

Chemical ionisation (CI)- ionisation by the addition of a charged species to a neutral molecule or by removal of a proton. The production ion is an even-electron ion with low internal energy and hence, undergoes little fragmentation. Used in combination with GC-MS.

Chromatography- a collection of methods for the separation of mixtures of compounds.

Detector- a device that detects the ions produced in the mass spectrometer and produce a measurable signal, generally an electronic signal.

Dwell time- used in MRM to describe the time taken to analyse a particular transition.

Dynamic range (detector)- a measure of the range, smallest to largest, of the ions that can be measured at one time.

Dynamic MRM (DMRM)- the DMRM method monitors the analytes only around the expected retention time, this decreases the number of concurrent MRM transitions, and significantly extends the dwell time. Thus, DMRM can monitor more MRM transition in a single run.

Electron ionisation (EI)- the process of ionisation whereby an energetic electron passes sufficiently close to an atom to interact with and displace one of its electrons, leaving a radical cation. This is the most common ionisation for GC-MS.

Electrospray ionisation (ESI)- a solution of the analyte is sprayed from a fine needle held at high voltage. The solvent evaporates and a series of fine charged droplets are formed. As they diminish in size, the coulombic forces between ions of similar charge on the surface exceeds the surface tension of the liquid and spawn ever smaller droplets, leading eventually to a solvent free ion. ESI is ideal for coupling to LC-MS. The ions formed are introduced into the high vacuum of

the mass spectrometer through one or more small apertures or through a length of capillary tubing, assisted by efficient vacuum pumping.

Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR)- an ion storage device in which ions are trapped by crossed magnetic and electric fields and follow circular trajectories perpendicular to the magnetic field. They are excited to wider orbits by a radiofrequency pulse and the decay of their image currents is detected as a time domain signal, later converted to *m/z* values by fourier transform. Ion detection does not result in loss of the ions, so they can be stored for long periods in the ICR cell. This technique has a very high resolving power which means that ions with very small *m/z* differences can be separated.

Fragmentation pattern- the characteristic distribution of fragment ions in the mass spectrum of a particular compound, from which the compound might be identified.

Gas chromatography (GC)- separation technique of compounds in their gaseous state can be achieved by passing them in a stream of heated inert gas (carrier gas) through a long column containing an immobilised stationary phase for which different components will have different affinities. The sample is injected on to the top of the column which is placed in an oven with carefully controlled temperature and the gaseous mixture is passed through the column. It is common to use a temperature gradient to heat the column. Separated components elute from the end of the column and are passed to a detector.

Gas chromatography- mass spectrometry (GC-MS)- direct coupling of the output from a GC to the source of a mass spectrometer. Helium can be used as the carrier gas with a capillary column. EI or CI are the most common forms of ionisation.

Hard ionisation- for example EI where the target compound is significantly fragmented during ionisation.

High performance liquid chromatography (HPLC)- a type of column chromatography which separates compounds in liquid mixtures. The liquid is added to a column containing a packed stationary phase through which a moving phase of solvent is passed. The differential effect of retention by the stationary phase and elution by the moving solvent phase separate the compounds as they pass down the column. HPLC evolved from LC in the 70s when smaller stationary phase particles were found to give better separation but required higher pressure to force the solvents through. This led to the name high pressure LC which was gradually replaced by high performance LC. HPLC is often referred to as just LC.

Imaging mass spectrometry- in a MALDI-TOF instrument the incident laser desorption and ionising beam can be rastered over a two dimensional surface. Diagnostic ions specific to target molecules are detected and their abundance is converted into an image reflecting their intensities within the analysed material.

Liquid chromatography-mass spectrometry (LC-MS)- analysis of the eluting liquid from a LC column by direct linking to a mass spectrometer ion source. Ion sources that operate at atmospheric pressure are ideally suited for LC systems. The current range of sources can ionise nearly all classes of organic molecules.

m/z- the dimensionless quantity used in MS to describe the detected ion, where *m* is the mass of the ion in unified atomic mass units (u) and *z* is the number of charges on the ion, positive or negative. The term is used to label the x-axis in a mass spectrum. *m/z* is often referred to as the mass-to-charge ratio.

Matrix-assisted laser desorption-ionisation (MALDI)- introducing the analyte molecule in very low concentrations into a crystalline matrix produces a sample from which efficient ionisation can be obtained by the action of photons from a laser. The matrix is chosen for its ability to absorb the wavelength of the laser emission and to co-crystallise with the analyte. The initial action of the laser is to desorb high-energy particles from the matrix and these, in a secondary action, ionise the analyte by protonation and deprotonation.

Multiple reaction monitoring (MRM)- increasingly high degrees of specificity can be achieved in the detection and analysis of mixtures by using tandem MS in which the two analysers are set to transmit only pre-defined pairs of precursor and product ions.

Precursor ion- an ion that reacts to form a **product ion**. Also referred to as parent ion and daughter ion respectively.

Quadropole- a device containing four precisely parallel conductors onto which a DC and radiofrequency voltages can be applied to create a quadrupolar field, which can act to focus, store or analyse ions.

Resolving power- a measure of the ability of a mass spectrometer to separate two ions of different but defined *m/z* value. For two overlapping singly charged peaks m1 and m2 of equal height, the resolving power is defined as m1/Δm, where m1 is the *m/z* value of one ion and Δm is the mass difference between m1 and m2 such that the two peaks are resolved with a defined interpeak valley.

Selected ion monitoring (SIM)- target compound analysis, often in combination with an on-line chromatographic separation in which the analyser transmits one or more ions in a sequence. The selected ions are characteristic of the target analyte. This can increase the sensitivity by avoiding the time spent on scanning areas of the mass range where no ions of interest will be found.

Signal-to-noise ratio- the ratio of the intensity of an ion signal to that of the noise. It is used to determine the lower limit of detection and limit of quantitation and to calculate the precision and accuracy with which the analyte can be quantified.

Soft ionisation- ionisation in which a minimum of excess energy is imparted to the newly formed ion, resulting in little fragmentation. ESI is an example of soft ionisation.

Tandem mass spectrometry (MS-MS)- a technique that involves at least two stages. In the most common method a first analyser is used to isolate a precursor ion, which then undergoes spontaneously or by activation a fragmentation to yield product ions and neutral fragments. A second spectrometer analyses the product ions.

Time-of-flight mass spectrometer (TOF)- an instrument in which the ions are separated solely on the basis of their velocities without the application of any external force. Ions are accelerated down an evacuated flight tube at high potential, their velocities and flight times depending on the mass of the ion, so that lighter ions reach the detector before heavier ones. The *m/z* value is

related to the flight time, length of the flight tube and the accelerating voltage. In a linear TOF the ions travel down the flight tube and are detected at the opposite end of the source. Resolving power can be improved by use of a reflectron, or a series of reflectrons, which reflect ions back down the flight tube.

Triple quadrupole mass spectrometer (QqQ)- a tandem mass spectrometer consisting of three consecutive sets of quadrupoles. The first and the third acts as transmission quadrupoles and the second acts as a collision cell. Ions are selected in the first quadrupole, fragmented in the second, and the product ions are analysed in the third.

References Box 1: Rao *et al.* (2014); Mallet and Down (2010); Ramautar *et al.* (2009)

1.3.1 Liquid Chromatography Coupled to Mass Spectrometry

One specific LC-MS method, LC-QqQ-MS, will be described further here since this was the technique that was used in papers II and III. LC-QqQ-MS is a good analysis technique for metabolites with high molecular mass, *e.g.* secondary metabolites, co-factors, and sugar phosphates since it offers excellent sensitivity, reproducibility, and a broad dynamic range (Weckwerth, 2011). This type of instrument is commonly operated in multiple reaction monitoring mode (MRM), though full scan and other scanning methods such as selective ion monitoring (SIM) are also possible. During MRM mode the collision energy and product ion mass-to-charge ratio are pre-optimised for each analyte of interest in order to give the best signal. Technically this means that the first quadrupole selects for the parent ion while the second works as a collision cell to fragment the parent ion, the third quadrupole then isolates a pre-defined product ion. Each MRM scan takes time (dwell time) which limits the number of compounds that can be analysed (Lu *et al.*, 2008). A general rule of thumb is that the chromatographic peak width divided by the cycle time of the MRM scans determines the data points available for peak detection. The LC-QqQ-MS instrument used in this work has a dwell time of 50 ms and typical peak width using an amide column of about 6 s, giving 8-10 scans per peak. If the peaks of interest are well separated chromatographically, the chromatogram can be divided into segments and only the metabolites within a specific time window are scanned for.

1.3.2 Gas Chromatography Coupled to Mass Spectrometry

GC-MS is still a favourite among metabolomics researchers, catching central metabolites such as amino acids, organic acids, free fatty acids etc. (Weckwerth, 2011). It is very suitable for the task since it combines high separation efficiency with versatile, selective and sensitive detection. Also, the

fragmentation patterns obtained are highly reproducible and hence can be collected in libraries and used to identify and classify detected metabolites. The GC-MS mass spectra libraries are a big advantage compared to LC-MS analysis where such libraries are still rare when it comes to identification of metabolites in a high throughput manner. Even so, the identification of unknown metabolites remains a bottleneck. The limitation of GC-MS is that only non-polar, thermostable, and volatile compounds can be analysed (Goodacre *et al.*, 2004). Non-volatile compounds need to be derivatized prior to GC-MS analysis; making them less polar, volatile, and thermally stable. Commonly samples are derivatized with an oximation reagent followed by silylation. In our study we used methoxyamine for oximation and N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) for silylation. The oximation inhibits reducing sugars (*e.g.* glucose and fructose) from forming rings and limits them to open-chain structures, which typically results in two peaks per reducing sugar. The two peaks are a result of the reducing sugars limited rotation along the C=N bond. Another positive effect of oximation is that it protects against chemical equilibrium between an α-ketone or an aldehyde and an enol (keto–enol tautomerism) (Fiehn *et al.*, 2000). During silylation a reactive hydrogen from the analyte (-OH, -NH, -SH) is replaced by a trimethylsilyl group ($-Si(CH_3)_3$). Alcohols, carboxylic acids, amine, thiols and phosphates can be silylated. MSTFA has been reported to be the best silylation alternative for plant metabolites (Fiehn *et al.*, 2000).

1.4 Isotope Labelling

1.4.1 An Old Technique with New Possibilities

The use of isotopes for analysis and description of metabolic pathways is a well-established biochemical technique. It has been applied to great extent throughout the last century for identifying major pathways in plant primary metabolism. The general principle of the technique is to use a labelled precursor metabolite and to trace the isotope signature through the intermediate metabolite pools which connect the precursor and end product of the investigated metabolic pathway (Fernie *et al.*, 2005). The most classical example of a labelling study in plant science is probably the elucidation of the Calvin-Benson cycle, using $^{14}CO_2$-labelling in the green alga *Chlorella*. Calvin started to adopt ^{14}C as a principal tool in his work since it was cheap and available in large amounts due to the nuclear reactors that were constructed in the mid-1940s. They worked out a procedure using ion exchange columns and paper-chromatography to detect the labelled compounds after exposure to radioactive CO_2. It took an additional ten years to identify the compounds. In

1962, after twenty years of work, the complete pathway was finally published in Science (Calvin, 1962). However, Calvin and co-workers was not the first to publish results from ^{14}C-labelling experiments. Radioactive G1P and G6P were used in the mid-50s in several glucose metabolism studies (Gibbs & Beevers, 1955; Beevers & Gibbs, 1954; Bloom & Stetten Jr, 1953). In the early 70s $^{14}CO_2$ was used to study glycolate formation through ascorbate (Plaut & Gibbs, 1970) and there are several examples of radioactive sugars used to study different metabolic pathways later that century (Borowitz *et al.*, 1977; Ap Rees *et al.*, 1976). Even though ^{14}C tracers were the most common, in 1970 a paper was published describing the advantages with ^{13}C tracers (Flaumenhaft *et al.*, 1970).

Indeed this is the direction that labelling experiments have taken. Advantages with ^{13}C are, for example, its compatibility with MS and NMR techniques, that it enables positional labelling information, and is not radioactive. ^{13}C-labelled sucrose and glucose have been used extensively for studying metabolic flux in cell cultures by feeding tracer until steady-state is reached (Williams *et al.*, 2008; Allen *et al.*, 2007; Baxter *et al.*, 2007; Matsuda *et al.*, 2007; Sauer, 2004; Wittmann & Heinzle, 2002; Wiechert, 2001). See Box 2 for a glossary of flux analysis terms.

Fluxes cannot be measured directly like transcript, protein and metabolite abundance, but are estimated from other data sources. This puts high demands on experimental design, instrumental precision and the power of the computational models (Heise *et al.*, 2014). Advances in high-throughput technologies, such as mass spectrometry techniques, with increasingly better sensitivity and precision, opens up new possibilities for flux analysis. There are other ways of analysing fluxes besides using isotope labelling, but they will not be covered here. The most used isotope-based methods for steady state are flux balance analysis (FBA), and metabolic flux analysis (MFA). Attempts to study plant metabolism in cell suspension systems where the respective tissue has been taken from its *in-vivo* environment are to some extent artificial: typically the connections to transport tissues, *e.g.* phloem and xylem are interrupted and internal gas concentrations are changed. Hence, it is desirable to measure carbon flux *in planta* (Fernie *et al.*, 2005). Plant metabolic networks are more complex than those of other organisms, due to being sessile, ectothermic, autotrophic, their vast chemical diversity and the plant cells highly compartmentalised organisation, which could be the reason for the low success rate of the attempts made of metabolic engineering of plants (Allen *et al.*, 2009). Single gene modifications have proved to rarely result in the desired change in composition, yield or growth. This can be explained by the function of individual proteins or pathways being dependent on the state of the larger

metabolic network. Hence, there is a need for new and clever methods for monitoring plant metabolism and metabolite flux.

The most recent development in the field of flux estimations is quantitative flux estimation. The key components in this type of labelling studies are a predefined metabolic model of the system, labelling patterns of metabolites from experimental data and a computational application for flux estimation based on optimisation of the fit between the experimentally determined labelling data and the prediction model (Heise *et al.*, 2014).

Online resources for flux analysis are, for example, OpenFLUX (openflux.sourceforge.net/) and 13CFLUX2 (13cflux.net/13cflux2/)(Fernie & Morgan, 2013), a list of additional software tools can be found at the OMIC tools homepage (http://omictools.com/13c-fluxomics).

Box 2. Glossary of flux analysis terms.

Dynamic labelling- metabolic fluxes are obtained by analysing the time-course for the redistribution of the label.

Flux balance analysis (FBA)- analysis of the steady-state fluxes in a metabolic network based on stoichiometry and the measured input and output fluxes.

Isotope- an atom is defined by its number of protons and electrons, but the number of neutrons can vary over a small range, these are called isotopes. Some isotopes are unstable and many are radioactive.

Isotopically non-stationary MFA (INST-MFA)- quantification of intracellular metabolic fluxes based on computational analysis of dynamic isotope labelling routes.

Isotopic steady state- the point reached in a labelling experiment when the labelling of each metabolite in a network is constant.

Isotopomer- One of a set of molecules with the same chemical and isotopic composition, but differing in the positional arrangement of the isotopes. The term is also used to describe molecules with the same chemical structure but differ in isotopic composition.

Mass isotopomer- a set of isotopomers with the same mass. A compound with *n* carbon atoms has *n + 1* mass isotopomers.

Metabolic flux- the flow of material through a step or pathway.

Metabolic flux analysis (MFA)- steady-state analysis that relies on the isotopic labelling of individual components and models the distribution of labelled molecules.

Metabolic pathway- a route linking two or more metabolites in a network.

Metabolic steady-state- a state in which all the fluxes in a network are invariant, so that the concentrations of the metabolic intermediates are constant.

Network flux analysis- procedure for generating a map showing multiple fluxes in a metabolic network by extracting fluxes and kinetic parameters by fitting experimental data to the model of the network.

Steady-state labelling- fluxes are determined by measuring the redistribution of the label after the system has reached an isotopic steady state.

References Box 2: Fernie and Morgan (2013); Young *et al.* (2011); Liu *et al.* (2010); Ratcliffe and Shachar-Hill (2006).

1.4.2 The Challenge of Labelling Whole Plants

Labelling experiments with whole plants are preferably performed with $^{13}CO_2$ since CO_2 is their only carbon source and it is experimentally very easy to introduce. There are several challenges with flux estimations by labelling of whole plants. First, it is not trivial to find an optimal way to introduce the ^{13}C tracer without disturbing the system. Second, there are complications in data evaluation related to the dilution of the tracer by internal pools of unlabelled substrates (Fernie et al., 2005). Third, the assimilation of $^{13}CO_2$ in autotrophs leads to an uninformative isotopic steady-state where all metabolites are fully labelled irrespective of fluxes and pool sizes (Heise *et al.*, 2014). This means that steady-state assumption methods such as FBA and MFA cannot be used (Young *et al.*, 2008). Hence, for successful determination of metabolic flux in whole plants a different strategy is needed where a time series is collected and the enrichment of label is monitored over time (Kolling et al., 2013). There are studies where individual leaves were labelled by clamping a leaf, introducing the tracer only to that particular leaf (Kolling et al., 2013; Hasunuma et al., 2010). This strategy is perilous since labelled metabolites might be exported to other parts of the plant as well as unlabelled compounds being transported into the clamped leaf. This would result in errors when calculating and interpreting the ^{13}C-incorporation. Another possibility is to measure the unlabelled metabolite fraction and use existing data of metabolite pool sizes to calculate intracellular fluxes (Heise et al., 2014). There are few plant studies using $^{13}CO_2$ reported in the literature, reflecting the fact that the labelling kinetics of non-steady state experiments is very challenging, especially due to the vast number of isotopomers (Szecowka et al., 2013; Arrivault et al., 2009; Huege et al., 2007).

2 Objectives

The aim of the work presented in this thesis was to find new and better ways to study plant metabolism. Although the biological questions are the driving force of the research, some focus has also been put on the technical aspects and development of new ways to probe plant metabolism using heavy isotopes and mass spectrometry.

Paper I describes the impact on plant metabolism when an enzyme in the TCA-cycle is knocked out and introduces *mmdh1*, a single knockout mutant lacking a growth phenotype. In **paper II** the metabolic profile of *mmdh1* is further investigated by a labelling method using $^{13}CO_2$ in combination with mass spectrometry techniques. The labelling method is adapted in **paper III** to investigate carbon allocation in developing wood of hybrid aspen. The last study, **paper IV**, describes a method for measuring metabolite profiles in specific cell-types, which is an interesting direction for future studies of plant metabolism.

3 Methodological Considerations

3.1 Model Organisms

There are many model systems available in research. Mouse, rat and fruit fly are examples of common model systems used in medical science. In plant science one of the most used model systems is *Arabidopsis thaliana*, which has the common name Thale cress. For research in woody plants, species within the *Populus* genus are often used. The attributes of a good model system, no matter if it is a mouse, fly or plant, is that they are easy and fairly cheap to house (*e.g.* have non-specific living requirements, are small in size and that they have a short life-cycle with many off-springs), their genome is sequenced and that they are easy to transform. Furthermore, an established model system comes with a wealth of data and previous knowledge. In this study, Arabidopsis is predominantly used but in paper III, the model tree hybrid aspen (*Populus tremula x Populus tremuloides*) was used to investigate the carbon allocation in developing wood.

3.2 Design and Construction of Labelling Chamber

Introducing ^{13}C to plant leaves is not in itself complicated. The initial labelling experiments for this thesis were performed in 50 ml falcon tubes, with two severed leaves from Arabidopsis wild type (wt) plants. With a 60 ml syringe ^{13}C-labelled carbon dioxide was injected into the tubes and then they were kept under a light source (300- 500 $\mu mol\ m^{-2}\ s^{-1}$) for 15 or 60 min after which they were instantly frozen in liquid nitrogen. This very simple method was used to detect 32 metabolites involved in primary carbon metabolism, of which 16 were labelled. The set-up had many flaws, but the material obtained was adequate enough for method development of the mass spectrometry methods. However, to label whole plants in such a way that the obtained data is

adequate for accurate ^{13}C-incorporation estimations requires a more advanced labelling set-up. Seven functions were identified as necessary for a well-functioning labelling system, namely the possibility to;

1. label whole plants
2. work with the plants within the chamber without disturbing the inside environment
3. switch between gases ($^{12}CO_2$ or $^{13}CO_2$)
4. monitoring and logging of environmental conditions during experiments (time, temperature, humidity, $^{12}CO_2$, $^{13}CO_2$)
5. change the gas composition (*e.g.* concentration of CO_2)
6. rapid sampling and freezing
7. simultaneous treatment of multiple genotypes and adequate biological replicates

There are very few, if any, commercially available labelling chambers or they have serious limitations (*e.g.* single plant treatment or no possibility of regulating the gas mix or concentration). To avoid these limitations a labelling chamber was constructed that fulfilled the requirements.

A prototype was built by engineering students as a project course. The identified traits were given in the project specification. The resulting chamber was evaluated with good results showing clear ^{13}C-incorporation in the metabolic profiles of the plants. A larger experiment with four wild type and four *mmdh1* plants were performed in high and low CO_2 concentrations. The results showed no separation between genotypes in high CO_2 but a clear separation over time in low CO_2 (figure 4). Although these results were encouraging several flaws and limitations with the chamber were identified. For example; no automatic regulation of CO_2 concentration, slow and tedious sampling and problems with shading, too crowded space for the plants, no control or monitoring of temperature and humidity. These problems were taken in consideration for the construction of a second version of the chamber. This labelling chamber full-filled all the above mentioned criteria (further described in paper II). Figure 5 A shows a schematic over the 1st labelling chamber. The 2nd labelling chamber is presented as a photo in figure 5B and as a schematic overview in figure 5C. For the labelling experiment of hybrid aspen in paper III, due to the larger size of two month old hybrid aspen trees, a tent was used for introduction of the ^{13}C tracer.

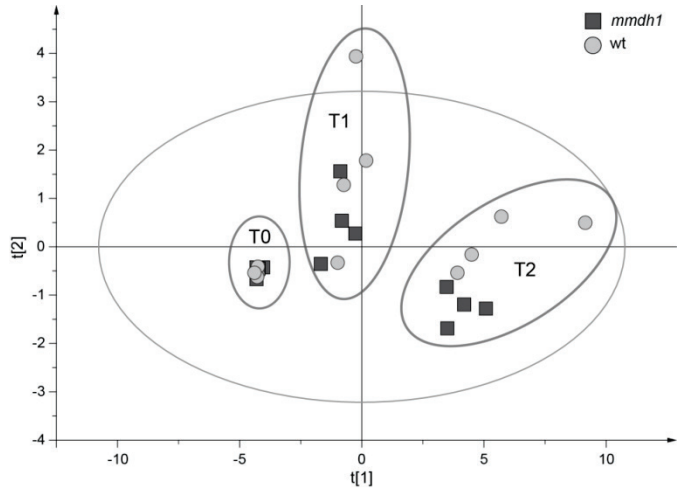


Figure 4. The first prototype of the labelling chamber showed a separation of wt (light grey circles) and *mmdh1* (dark grey squares) over time in low CO_2 concentrations and of genotypes at T2. To clarify the time point separation they were encircled. Time points: T0; control, T1; 30 min treatment and T2; 120 min treatment.

3.3 Special Considerations Regarding Labelled Metabolite Detection

3.3.1 Increasing Complexity in Metabolite Detection

The difference between detection of unlabelled and labelled metabolites is that a mass spectrum from a labelled metabolite changes depending on the degree of labelling and hence cannot be used for identification. Thus, to acquire metabolite quantifications for a labelled metabolite using GC-MS a targeted analysis approach must be used. A predefined list was put together with all possible isotopomers (*m/z* values) and a retention time window for each metabolite. The number of isotopomers needed to correctly calculate metabolite abundance is *(n+1)*, where *n* is the number of carbon atoms in the metabolite. Great time and care was put into creating a list of just under forty metabolites involved in primary carbon metabolism. The challenge was to find the fragment of each metabolite that contained the most back bone carbons in combination with a good detection level and without interference from other compounds. However, the first step was to analyse unlabelled plant material since this made it possible to use spectral libraries for metabolite identification and define the retention time window prior to fragment investigation. Next the

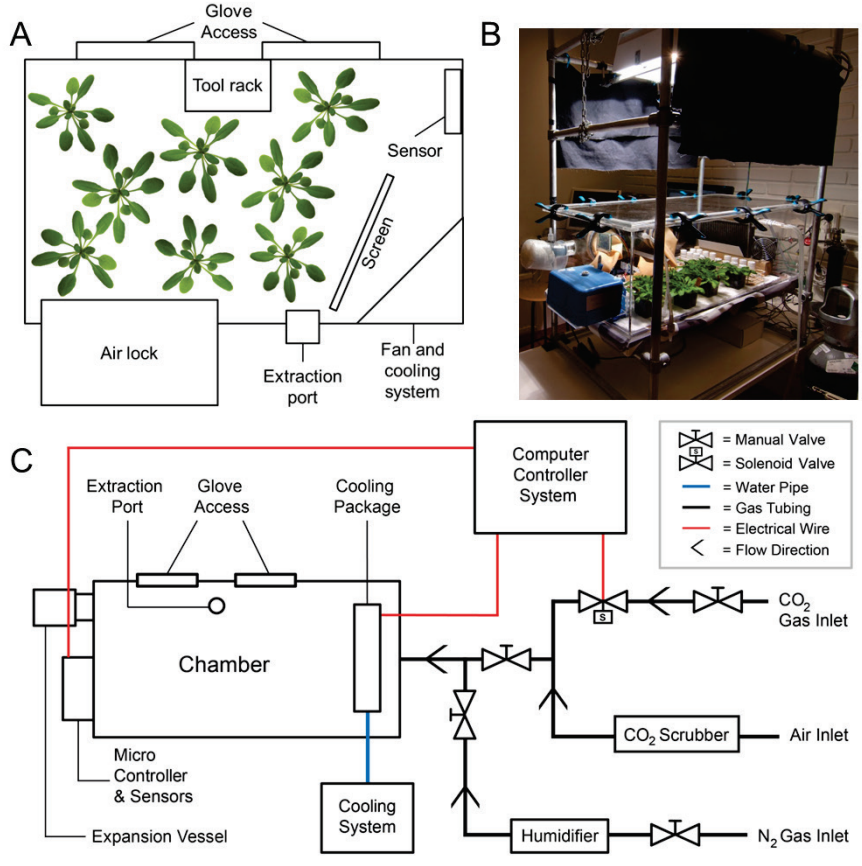


Figure 5. A) Schematics of the first prototype of the labelling chamber. B) Photograph of the 2nd version of the labelling chamber. C) Schematics of the 2nd and more advanced labelling chamber. The 2nd chamber was used for labelling Arabidopsis plants in paper II.

^{13}C-labelled leaf material was used to evaluate the quality of the selected fragments. The labelled data was investigated thoroughly and some fragments were replaced with others due to interference from other compounds. Large molecules, such as phenylalanine, which contain nine carbons, demand ten after each other following mass channels without interference to determine the full incorporation correctly. Therefore, some interference was to be expected. For the metabolites were it proved hard to find a fragment without interference the isotopomers before the interference were calculated (commonly the first two or three).

3.3.2 The Dilution Effect

Another issue with detection is that the low abundant metabolites became even harder to detect when they become labelled. This is due to the dilution effect that occurs during labelling. For example, if a four carbon metabolite has the exact same abundance in two samples where one sample is unlabelled and the other ^{13}C-labelled, the abundance of the unlabelled sample is determined by m_0 but for the labelled sample five isotopomers must be quantified, m_0- m_4, in order to determine its abundance. Hence, the intensity of m_0 in the unlabelled sample will be distributed over five isotopomers in the labelled sample, making the detection level for this compound lower than that of its unlabelled counterpart.

3.3.3 Experimental and Analytical Limitations

The dilution effect proved to be a problem for analysis of sugar phosphates in illuminated Arabidopsis leaves. Regardless of extensive method development using a LC-QqQ-MS/MS instrument with excellent sensitivity in MRM mode, many of the intermediates in the Calvin-Benson cycle were simply too low abundant to be detected in labelled leaf samples. Emphasis was put on a simple sample preparation for GC and LC samples together (figure 6). Thus, focus was on optimising the analysis in other ways than sample preparation. Different LC columns were evaluated and a thorough instrument optimization was performed in an effort to detect low abundant metabolites such as sugar phosphates. The specificity and sensitivity of MRM are great advantages compared to other techniques but there are limitations in the number of MRM transitions that can be scanned per injection. This is a liability when working with labelled material since *n+1* transitions are needed per labelled compound. For GC-MS analysis all channels between 50-800 *m/z* were scanned, which made it possible to detect all 191 ion channels for the 37 analysed metabolites in just one sample injection. Using LC-MS and MRM this corresponds to 12 injections per sample (assuming 16 MRM transitions per injection). Hence, the number of labelled metabolites analysed by MRM was limited to a few, but for these the sensitivity and precision was high. Using the LC-MS method three hexose phosphates (glucose-6-phosphate, fructose-6-phosphate and glucose-1-phosphate) and UDP-glucose were detected and the ^{13}C-incorporation was estimated.

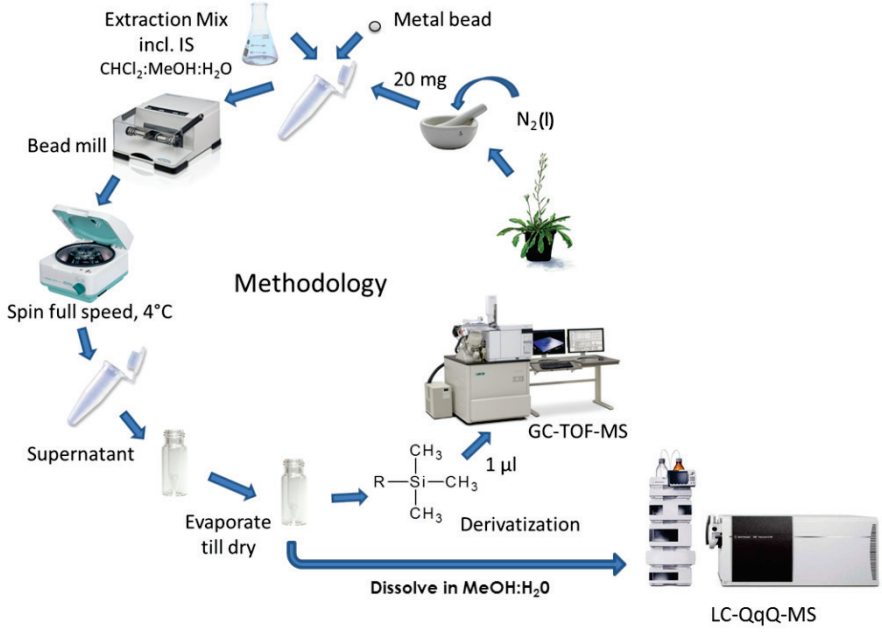


Figure 6. A common sample preparation methodology was used for both GC-and LC-MS samples; from harvest of the plant material, through the extraction process until the end of the extraction after the samples were dried down. At this step the samples were separated so the ones destined for LC-MS analysis were dissolved in 50:50 methanol and water, while the samples destined for GC-MS analysis were derivatized by methoxyamine and trimethylsilyl.

3.4 Data Processing

3.4.1 Mass Spectrometry Data

Before the acquired MS data can be used for any statistical or biological evaluation it needs to be processed. This is to make sure that the same corresponding peak in every sample is assigned the same identity, meanwhile handling peak shift and peak overlap. The main challenges in data processing are the sheer number of peaks per sample, unbiased data processing, peak alignment, and obtaining a unique entry per metabolite. There are three ways to do this; target analysis, peak picking, and deconvolution. For target analysis a predefined list with *m/z* values and retention time windows is used to quantify and identify known metabolites in a sample. The precision is generally higher than for peak picking and deconvolution and it only generates one entity per metabolite and sample. The disadvantage is the time consuming work of putting together the list of *m/z* and retention time windows for every

metabolite, another drawback is that no other peaks will be analysed and therefore many interesting peaks might be missed.

Peak picking is used to automatically extract compounds from raw data guided by peak parameters *e.g.* centroid, area, height, peak-width-at-half-maximum, signal-to-noise ratio and asymmetric peak shape. Peak picking is an explorative approach which includes all peak-like structures in the data, this result in a large amount of non-study specific structures being captured. There is a string of software available dealing with peak picking for GC, LC, CE-MS and NMR data. Examples of free software for peak picking are listed in Table 1 (Saito & Matsuda, 2010).

Table 1. Examples of open source software for peak picking of NMR, GC-MS and LC-MS data.

Software	**Type of Data**	**Webpage**
SpinAssign	NMR	http://prime.psc.riken.jp/?action=nmr_search
AMDIS	GC-MS	http://www.amdis.net/
MET-IDEA	GC-MS	http://www.noble.org/PlantBio/sumner/met-idea/
Metabolite Detector	GC-MS	http://metabolitedetector.tu-bs.de/
MZmine	LC-MS	http://mzmine.sourceforge.net/
XCMS	LC-MS	http://masspec.scripps.edu/xcms/download.php
MetAlign	GC-MS and LC-MS	http://www.wageningenur.nl/nl/Expertises-Dienstverlening/Onderzoeksinstituten/rikilt/show/MetAlign.htm

The third option, deconvolution, is a mathematical method that uses differences in the mass spectral information between overlapping metabolites in order to separate them. Simply put, deconvolution resolves unresolved peaks and transforms the raw data into peak tables with integrated peak areas for every metabolite in every sample plus a list of their mass spectra. These spectra can be used further for metabolite identification. The method can be automated and applied to a whole set of samples simultaneously, making this alternative superior to target analysis and peak picking. However, the perfect algorithm for deconvolution and peak integration has yet to be written and available programs have problems resolving how many metabolites there are in a cluster of peaks and also the variability of the mass spectra. In a perfect case the spectra would be the same for a metabolite between multiple samples, but in reality differences occur, complicating the process. Still, deconvolution is the most promising method because it can handle large data sets, be automated, performs automatic peak alignment, and give one area per metabolite and

sample. For the work presented in this thesis both targeted and deconvolution strategies were used.

3.4.2 Labelled Metabolite Data

It is difficult to automate processing of labelled data since there are so many isotopomers involved for every metabolite. To ensure the quality of the analysis it is important to screen through the calculated isotopomers of every sample and look for interferences. This must be done for every analysis as the interferences might differ between experiments. Also, it cannot be emphasised enough that without a good signal-to-noise level there might be errors introduced to the ^{13}C-estimations. To speed up the GC data analysis and keep the manual work to a minimum an in-house MATLAB script, Frag_calc (short for "fragment calculator"), was created. The input in Frag_calc was a list containing; a unique name, ion channel and retention time window for every metabolite to be analysed. The script calculated both abundance of each isotopomer and also the total abundance of all isotopomers for a metabolite summed up. To aid quality control of the quantifications a visualization plot for each metabolite was created in a designated project folder. The visualization plots presented the abundance with relative standard deviations for the different isotopomers for two genotypes so they could easily be compared (figure 7).

3.5 Incorporation Calculations and Multivariate Statistics

3.5.1 Calculation of ^{13}C-incorporation

Once lists of peak intensities for all the isotopomers were attained for the GC data, the next step was to determine and subtract the natural occurrence of ^{13}C (approx. 1.1 % of all carbons (Smith, 1972)) and the contribution from the TMS-groups (the natural isotopes of silicone; ^{28}Si 92.23 %, ^{29}Si 4.67 %, and ^{30}Si 3.1 % (De Laeter et al., 2003)). The isotope contribution from the TMS-groups and natural abundance of ^{13}C will be the same independent of whether a metabolite is labelled or not. This relation was used to calculate the actual ^{13}C-incorporation by sequential isotope compensation from a reference (unlabelled) spectra, deducting isotope contribution from both silica and natural abundance of ^{13}C (Figure 8). The same strategy was applied to the LC data, though only for natural abundance of ^{13}C.

In addition to Frag_calc, as was mentioned earlier, another in-house script, 13C_est (short for "^{13}C-incorporation estimation"), was developed to manage the incorporation calculations. The 13C_est script corrected for contribution from natural abundance of ^{13}C and TMS-groups and gave an output as

estimated ^{13}C-incorporation in percent. The script generated a text-file with the results and, just like Frag_calc, a visualization plot for each metabolite. The plot showed ^{13}C-incorporation as a column graph for each time point and genotype with relative standard deviation (RSD) values for every isotope in the reference spectra. The RSD values were of great value for the evaluation of the data quality. The reference (unlabelled) isotopes was coloured in dark blue and ^{13}C estimations in red. The subtracted contribution from the natural abundance of ^{13}C and the TMS-groups were shown in different colours (figure 9).

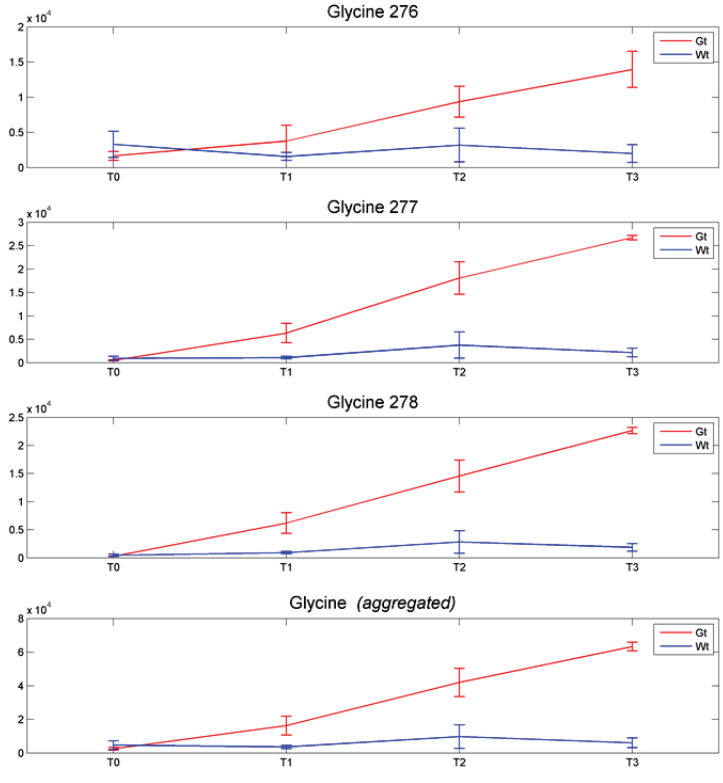
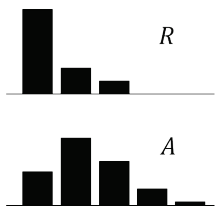


Figure 7. Visualization plot from the Frag_calc script for glycine in low CO_2 treatment. Wild type is blue and *mmdh1* red. Glycine is calculated from three *m/z*; 276-278, which are all presented. The bottom plot is the sum of all calculated *m/z*. Error bars are standard deviation calculated from four biological replicates.



$R = {}^{12}C\ reference\ MS\ (mass\ spectrum)$

$A = {}^{13}C\ incorporated\ MS$

$S_i = estimated\ MS\ for\ each\ degree\ (k)\ of\ {}^{13}C\ incorporation$

$k = number\ of\ incorporated\ {}^{13}C\ atoms$

$n = number\ of\ isotopes\ (n \leq k + 1)$

$for\ i = 0{:}k$

$$S_i(1{:}n) = R(1{:}n) \times \frac{A(i+1)}{R(1)}$$

$$S_i(S_i > A([1{:}n] + i)) = \mathrm{A}([1{:}n] + i)$$

$$\mathrm{A}([1{:}n] + i) = A([1{:}n] + i) - S_i(1{:}n)$$

end

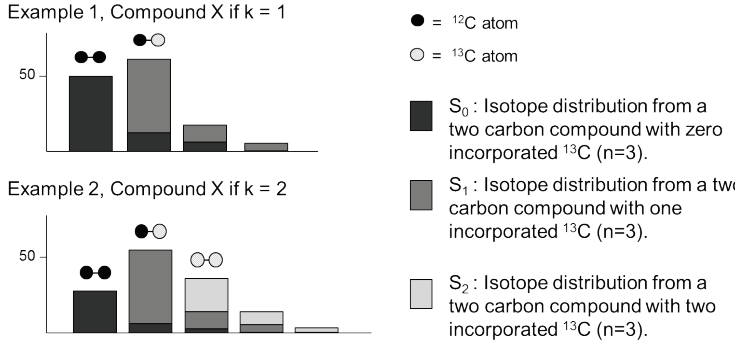


Figure 8. The principle behind the ^{13}C-incorporation calculations. TMS-groups and natural ^{13}C will be the same independently whether a metabolite is labelled or not. This was used to calculate the true ^{13}C incorporation by sequential isotope compensation from a reference (unlabelled) spectra, deducting isotope contribution from both silica and natural abundance of ^{13}C. The same strategy was applied to the LC data, but only for the natural ^{13}C abundance.

3.5.2 Multivariate Statistics for Biological Data

First the metabolite data is corrected for natural abundance of ^{13}C and TMS-groups, and normalized to fresh weight and internal standards. The next step is to use statistical analysis and/or mathematical models to interpret the biological information. The traditional approach would be to use a univariate method such as the student's t-test, but this is not a preferred strategy for this type of data since it assumes that there are more observations (samples) than variables (metabolites). Improper handling of covariance structures will lead to interpretational problems, *i.e.* type I and type II errors (false positives and false negatives, respectively). Also, information regarding relations between

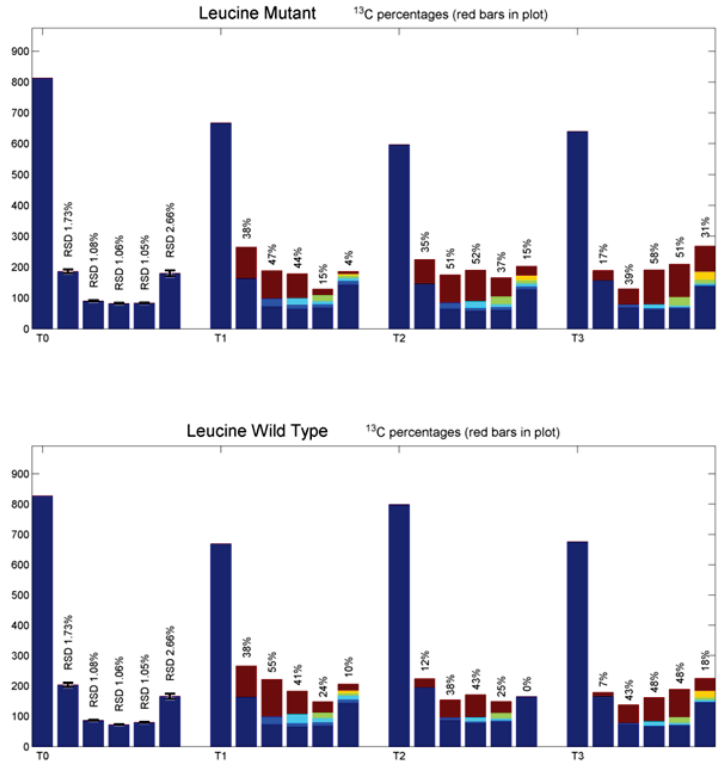


Figure 9. Visualization plot from the 13C_est script showing leucine in high CO_2 treatment. The isotope pattern from T0 was used as a reference spectrum to be able to eliminate unwanted contribution from natural ^{13}C and TMS-groups by sequential isotope compensation. Pure ^{13}C incorporation is shown in dark red and the percentage for each isotope is written above the column. These were summed up and presented in a separate text-file. For the unlabelled reference spectra, T0, error bars are RSD which is also written above the individual isotopomer.

variables will be overlooked. Other traditional statistical methods such as multiple linear regression are not preferred either since also here independent variables are assumed and this type of methods have problems with noisy data.

The better alternative for complex biological data such as metabolite profiles is to use chemometric methods. This type of multivariate analysis that is based on projection methods gives a good visualisation of the data and helps its interpretation. Chemometric methods convert the multi-dimensional data table into a low- dimensional model plane, usually consisting of one to five dimensions (Trygg *et al.*, 2006a). Principal component analysis (PCA) (Jackson, 1991) and partial least squares (PLS) (Trygg & Wold, 2002) are two common methods that can handle incomplete, noisy and collinear data. The

data needs to be organised in a two-dimensional table, an X matrix, where the samples are organised in the first column and the metabolite intensities in the horizontal direction. One should keep in mind that biological data is noisy in its nature, often with large biological variation between replicates.

There are yet other methods for statistical analysis and visualisation of multivariate data *e.g.* neural networks or self-organising maps, but these lacks the possibility to visualise how the variables influence the observations which is possible using chemometrics. I will give a brief overview of the chemometric methods used in this thesis.

3.5.3 Projection Methods for Multivariate Data

In short the principle of projection methods is transforming a matrix X into a swarm of data points in a multi-dimensional space. If matrix X has N rows (samples) and K columns (variables), each row (individual sample) can be represented as a point in a K-dimensional space. A samples position in this space is given by its coordinates (variables *e.g.* metabolic profile). Together the N observations will result in a swarm of points in the K-dimensional space. Points that lay close to each other are more biologically similar than points that are far away from each other. Projection methods find model hyperplanes of much lower dimensionality that approximates the matrix X (the swarm of points). Figure 10 gives an overview of the principle of the method. PCA is probably the most commonly used multivariate projection method. It extracts and visualises the systematic variation in a data matrix X. This is done by the two first principal components defining a plane (a window) into the K-dimensional space. By projecting the points in matrix X onto this two dimensional sub-space it is possible to visualise all samples. The coordinates of the samples are called scores T, and they are weighted averages of all variables in the matrix X (metabolites). The visualisation of the samples (scores T) are called a score plot. In the score plot it is possible to see how the samples relate to each other, for example clusters, trends, and outliers. To investigate what variables lie behind these relations in the score plot there is a corresponding plot, the loading plot P, which describes the influence (weight) of the X-variables (metabolites) in the model (Trygg *et al.*, 2006a). This is a powerful tool for understanding the underlying patterns of the data and is an excellent starting point to get an overview of a large data set (X matrix).

When additional information about the samples exist partial least squares projection to latent structures (PLS) method can be used.

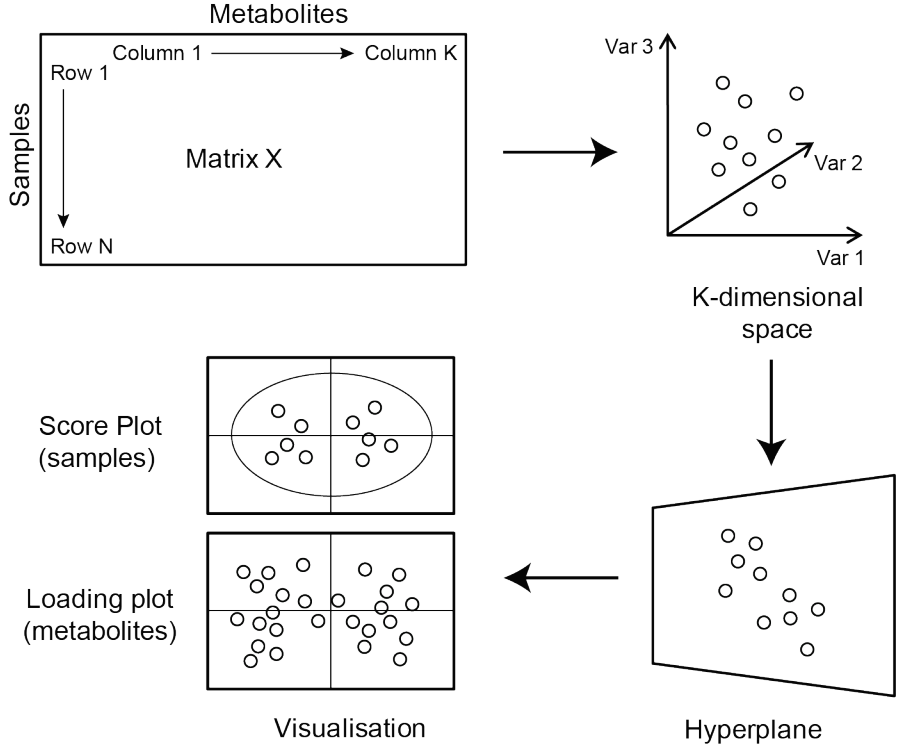


Figure 10. If matrix X has N rows (observations/samples) and K columns (variables/metabolites), each row (individual sample) can be represented as a point in a K-dimensional space. An observations position in this space is given by its coordinates (variables *e.g.* metabolic profile). Together the N observations will result in a swarm of points in the K-dimensional space. Points that lay close to each other are more biologically similar than points that are far away from each other. Projection methods find model hyperplanes of much lower dimensionality that approximate the X matrix (the swarm of points) and makes it easier to visualise.

The information, such as genotype or treatment is collected in the Y-matrix, which has the same number of rows as X. Each row in Y describes a group belonging for a sample where "1" means that the sample belongs to the group and "0" means that is does not. It is a well-used method if one is looking for a quantitative (or qualitative) relationship between the descriptor matrix X and a response matrix Y. The PLS method uses the additional information in the Y matrix to focus the model plane to the Y related variation in X. This means that the model will focus on the separation between genotypes or treatments rather than providing a general overview of all variation in X, as is done by a PCA model. Another possibility with the PLS method is that is can be used to predict the Y values (genotype or treatment) of new samples. When Y is qualitative, *e.g.* chlorophyll content, it is called PLS-discriminant analysis

(PLS-DA), to separate it from when Y is quantitative, *e.g.* genotype or treatment (Trygg *et al.*, 2006a).

Unfortunately, in many studies where the PCA method is the only tool applied the risk is that interesting information might be lost. There are more advanced multivariate methods that can take the analysis further. If the matrix X contains two known classes (*e.g.* mutant and wild type) orthogonal projection to latent structures-discriminant analysis (OPLS-DA) can shed light on class separation and simplify the interpretation. It can also explain some of the interclass variation. If there are no defined classes in the data set OPLS is a good option. This method aims to separate the systematic variation in the matrix X into two parts, one that is linearly related to Y and one that is unrelated (orthogonal) to Y. This division of X makes it possible to interpret the model but also to use the model to predict the classification of new samples. An OPLS model visualises two components, the Y-predictive and the Y-orthogonal (Trygg *et al.*, 2006b).

4 Results and Discussion

4.1 Characterisation of Mitochondrial Malate Dehydrogenase Mutants in Arabidopsis (Papers I & II)

In paper I we investigated the role of mitochondrial malate dehydrogenase (mMDH) in plant metabolism. mMDH is a TCA-cycle enzyme catalysing the NAD^+-dependent reversible reaction of malate to oxaloacetate (OAA). Its activity is also important for redox control of the mitochondrial matrix, through which it may participate in regulation of TCA cycle turnover (Nunes-Nesi *et al.*, 2005). Furthermore, its activity is closely linked to malate/OAA shuttling across the mitochondrial inner membrane. This process may be especially important for shuttling reductants between mitochondria and peroxisomes during photorespiration (Peterhansel *et al.*, 2010). There are eight genes encoding isoforms of NAD-dependent MDH in Arabidopsis, two are targeted to the peroxisomes, one to the chloroplast, three that have no apparent target and are assumed to be in the cytosol and two which are targeted to the mitochondria. The mitochondrial isoforms will be referred to as mMDH1 and mMDH2, where mMDH1 is the major isoform with nine times more transcript compared to mMDH2. Homozygous T-DNA insertion lines were isolated for two independent *mmdh1* (*mmdh1-1* and *mmdh1-2*) lines and one for *mmdh2*. Crossing the three lines provided two double mutants (*mmdh1-1mmdh2* and *mmdh1-2mmdh2*). Genotyping and biochemical analysis of isolated mitochondria from the double mutants confirmed that there was no mMDH activity. The *mmdh1-2mmdh2* and its constituent genotypes were used in further work. The double mutant will henceforth be referred to as *mmdh1mmdh2* and the single mutant of mMDH1, *mmdh1-2*, as *mmdh1*. If nothing else is stated the presented results were obtained in, and refer to, mature leaves. Plant growth was unaffected in the single mutants while it was distinctly reduced in the double mutants, also germination and seed set was affected. The growth phenotypes were observed in both long and short day-

growth conditions (figure 11). In total leaf extracts mMDH activity for long-day grown plants was decreased with 30 % in *mmdh1* and more than 40 % in the double mutant, while the activity was no different from wild type in *mmdh2*. This shows how important the mMDH activity is to the total NAD-MDH activity. Isolated mitochondria of wild type plants were used to analyse total mitochondrial protein extracts by 2D-gel separation. This revealed the known major mMDH1 spot and the much weaker mMDH2 spot. The mMDH1 spot was absent in *mmdh1* and *mmdh1mmdh2*. When comparing the mMDH activity of isolated mitochondria the double mutant did not show any activity, *mmdh1* had about 60 % less activity than wild type and *mmdh2* ~30 % less activity.

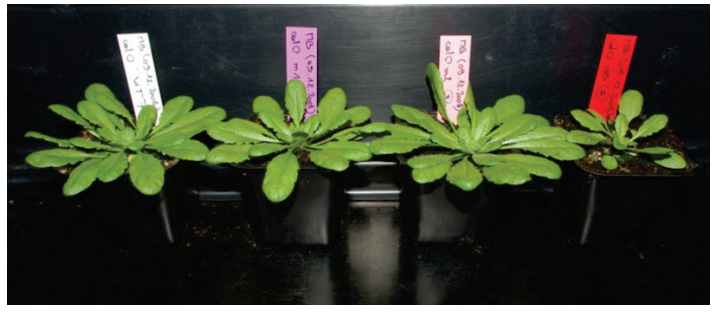


Figure 11. Growth phenotypes of five week old Arabidopsis plants grown in short day conditions. From left; wild type, *mmdh1*, *mmdh2*, *mmdh1mmdh2*.

Our study shows that mMDH can be removed without affecting plant vitality for the single knockouts while the double knockout had a major growth deficit. Hence, the loss of some mMDH can be compensated by other reactions. The double mutant for example could drive the malate-dependent respiratory flux through the electron transport chain at maximal speed using just the NAD-malic enzyme (NAD-ME). We believe the TCA-cycle is maintained by moving the NADH production to the cytosol through transport of malate out into the cytosol where cytosolic mMDH converts it to OAA. However, we found that other NAD-linked reactions also decreased in the double mutant. This would be consistent with mMDH operating in the NADH oxidation direction where OAA is reduced to malate in the light. When grown outdoors the double knockout could not survive and *mmdh1* had a 60 % decrease in seed production compared to wild type plants. This indicated that the compensatory mechanisms that set in when mMDH is reduced are too slow when the external fluctuations change too rapidly.

Light inhibits respiration by limiting the turnover of the TCA-cycle. In the light the mitochondrial redox state are increased due to an excess of NADH

which will contribute to inhibition of key steps in the TCA cycle, particularly steps catalysed by the pyruvate dehydrogenase complex and isocitrate dehydrogenase (Bykova et al., 2005). This would lead to a partial cycle in the light, and the outlined mechanism is strongly supported by several investigations in which several approaches have been applied (Hurry et al., 2005; Tcherkez et al., 2009; Sweetlove et al., 2010). Wild type, *mmdh1* and *mmdh2* showed a 40 % decrease in respiration in the light and the double mutant showed a similar inhibition but from a higher level in the dark (figure 4 Paper I). This indicates that the double knockout has a much higher respiration than the wild type and single mutants, both in darkened and illuminated leaves but that the overall change in respiration between dark and light is generally the same between all investigated genotypes. This implies that mMDH to some degree seems to have a limiting effect on the respiratory rate in leaves. A possible explanation for this could be that OAA is imported from the cytosol and converted to malate by mMDH. The competition with the electron transport chain for NADH and also the TCA-cycle through citrate synthase for OAA would result in limiting respiration. This limiting effect on respiration appears to be of high importance for the net CO_2 assimilation and biomass production considering the growth effects observed for *mmdh1mmdh2*.

Another interesting finding was regarding a plausible role of mMDH in photorespiration. The reactions of photorespiration that take place in mitochondria is the oxidation of glycine to serine, releasing one CO_2 and one ammonium ion while NAD^+ is reduced to NADH. The NADH is needed in the peroxisome, converting hydroxypyruvate into glycerate to balance the pathway. The double mutant exhibited a break in its CO_2 response curve at low CO_2 concentrations and elevated glycine levels (figure 8A in paper I). This suggests that mMDH has some role in photorespiration, probably through a limitation on the conversion rate of glycine to serine.

In conclusion, we found that mMDH has an effect on plant growth and photorespiration and that it decrease leaf respiration. The findings all seem to be connected to the cell trying to maintain the cellular redox balance.

In paper II we developed an analysis method using heavy isotopes and mass spectrometry to study primary carbon metabolism in mature leaves of *mmdh1* and wild type plants. Knowing that *mmdh1* lacked a growth phenotype but had a metabolic phenotype we wanted to investigate further how the decrease in mMDH would affect metabolism in a wider metabolic perspective.

A labelling chamber as described in chapter 3.2 was constructed, in which sets of eight plants were treated simultaneously. Metabolites in the primary carbon metabolism were analysed by LC- and GC-MS, optimised for ^{13}C-

labelled compounds. Since the data processing and calculations becomes more challenging when ^{13}C is involved, customised in-house scripts were developed to aid the data processing and ^{13}C-enrichment estimations. The metabolic profile and ^{13}C enrichment of both genotypes were analysed over a time course, under both high and low $^{13}CO_2$ conditions. The differences between *mmdh1* and wild type were expected to be small under high CO_2 conditions and that differences would appear under low CO_2. Multivariate statistics were used to investigate the metabolic profiles of the samples.

Indeed, the responses of wild type and *mmdh1* under high CO_2 treatment were found to be very similar. Under the low CO_2 treatment both the abundance and labelling of some metabolites were significantly different between the two genotypes. Interestingly, many of the observed differences were closely associated with photorespiration. In the proposed mechanism for export of reductants generated by mitochondrial glycine oxidation, NADH is reoxidised to NAD^+ by conversion of OAA to malate. The malate can then be exported to produce OAA and NADH outside the mitochondria. Finally, the OAA formed can be transported back to the mitochondria in exchange for malate. The net effect of this malate/OAA exchange is export of reductants from the mitochondrial matrix (see figure 5A in paper II). In accordance with paper I our data clearly indicated that in *mmdh1* limitation in OAA to malate conversion directly influenced the photorespiratory conversion of glycine to serine, catalysed by the glycine decarboxylase complex. Presumably this was caused by an increased NADH/NAD^+ ratio in the mitochondrial matrix. Similar increases in redox state have been observed following reductions in the capacity of glycine decarboxylation by antisense RNA (Bykova *et al.*, 2005). These results also suggest that *mmdh1* has limited capacity to shuttle the produced reductants from the mitochondrial matrix out to the peroxisomes. Furthermore, we saw a decrease in glutamate and glutamine pools together with an increase in α-ketoglutarate. This may be directly related to the reduced rates of ammonium production from mitochondrial glycine oxidation, which would limit its re-fixation via glutamine synthetase/ glutamine oxoglutarate aminotransferase (GS/GOGAT). As mentioned earlier an increased mitochondrial redox state can inhibit specific steps in the TCA-cycle and thereby limits its turnover, resulting in a partial TCA-cycle operating in the light.

Another interesting result was the accumulation of branched-chain amino acids in the *mmdh1* mutant under low CO_2. Valine, leucine and isoleucine were not measured in paper I, but a decrease in enzymes involved in their oxidation was observed. However, the ^{13}C-incorporation in these metabolites was not

correspondingly high, suggesting that their accumulation was due more to transfers between pools than *de novo* biosynthesis (see figure 5B in paper II).

These results are in accordance with paper I but the study in paper II also gave several other interesting findings regarding the role of mMDH on plant metabolism. For example, succinate proved to have a high level of ^{13}C-incorporation in *mmdh1* compared to wild type at every time point, under both high and low CO_2. Other TCA-cycle intermediates such as citrate and α-ketoglutarate did not show similar labelling patterns and the succinate pool was also much higher in *mmdh1* than in wild type plants. A possible explanation for this could be that the malate/OAA exchange in *mmdh1* is impaired, limiting the export of OAA from the mitochondria. Cytosolic PEP carboxylase might then fix $H^{13}CO_3^-$ to form OAA, which is reduced to malate by cytosolic MDH before it is transported back into the mitochondria and converted to succinate via fumarate at the cost of $FADH_2$. This labelled succinate then accumulate under these conditions. Such reactions could relieve some of the limitations imposed on glycine oxidation by a high matrix redox state.

Glutamine, glutamate and α-ketoglutarate showed a low level of ^{13}C-incorporation. The absence of label in these three amino acids could be due to the fact that they share the same carbon backbone and probably exchange carbon skeletons among their existing pools rather than being newly synthesized and thus labelled. On the other hand, we also found malate, fumarate and succinate to have a low labelling rate, but this is more likely due to the size and location of these pools in plant cells. These organic acids are stored in the vacuole in vast pools with little metabolic activity. Small, rapidly turned-over pools in other parts of the cell would therefore be masked by these huge vacuole pools.

^{13}C-labelling added another dimension to the analysis performed in paper I by providing information on metabolite dynamics. We found that hexoses, hexose phosphates, and sucrose rapidly became highly labelled. Two other sugars trehalose and maltose had similar pool sizes. However, the ^{13}C-incorporation revealed that maltose was highly labelled, similar to the other sugars, while trehalose remained poorly labelled throughout the ^{13}C treatment. This indicates that maltose is somehow a direct product of primary carbon incorporation. This is consistent with observations in earlier studies (Szecowka *et al.*, 2013; Schilling, 1982; Linden *et al.*, 1975).

In conclusion, the major finding in paper I was that mMDH has a role in photorespiration, biomass production, and regulation of respiration. The effect of the loss of mMDH resulted in redox imbalances to which the cell's metabolism had to adjust. The results in paper II highlight the striking

metabolic flexibility of plants. Even with a heavily reduced mMDH capacity the *mmdh1* mutant showed no visible growth phenotype. However, a new labelling method made it possible to analyse both abundance and estimate ^{13}C-incorporation of metabolites in primary carbon metabolism. The results revealed extensive metabolic changes in *mmdh1*, especially under limited CO_2 conditions.

4.2 Tracing ^{13}C in Plants Gives New Insights to Carbon Allocation and Transport (Paper III)

Having developed the labelling method in paper II, we aimed to broaden our scope to see how we could use the method for other research applications. In collaboration with colleagues working on wood formation in woody plants we got the opportunity to use the method for estimating ^{13}C-incorporation in source leaves, stem phloem, and developing wood in hybrid aspen.

Based on previous studies, the hypothesis was that the diurnal cycle has an effect on carbon allocation and wood formation in trees. Findings from previous studies, suggesting a diurnal effect on wood cell wall biosynthesis, has been mostly anatomical (Hosoo *et al.*, 2002; Yoshida *et al.*, 2000; Bobák & Nečesaný, 1967). The theory was further supported by findings of diurnal changes in transcript levels in the developing wood of eucalyptus trees (Solomon *et al.*, 2010). We developed a labelling system to investigate the patterns of carbon allocation in wood over a diurnal cycle. Two month old hybrid aspen were fed with $^{13}CO_2$ for four hours. Through feeding experiments it was established that most of the ^{13}C-tracer reached the cell wall within the first 24 hours. Hence, we focused on investigating ^{13}C transport and allocation in wood precursor metabolites and the main cell wall polymers (hemicellulose, cellulose, and lignin) over one diurnal cycle.

The developing wood of many tree species, including aspen, uses sucrose as its primary carbon source. Sucrose is synthesised in the leaves by photosynthetic CO_2 assimilation and is then transported via the stem phloem to the cambium, along the ray cells for final import to the wood forming cells. The phloem loading from the leaf cell is passive in *Populus* spp. while it has recently been shown that sucrose export from the ray cells to the developing wood requires active import (Mahboubi *et al.*, 2013). The imported sucrose is then metabolised to provide precursors for wood cell wall biosynthesis and energy for secondary growth. The metabolism of sucrose is initialised by splitting sucrose into glucose and fructose. These hexoses are quickly metabolised further to glucose- and fructose-6-phosphate (G6P and F6P). G6P

is a precursor to nucleotide sugars (NDP-sugars) which in turn forms hemicellulose. Sucrose can also be split into UDP-glucose and fructose. UDP-glucose is a precursor for cellulose and can also metabolise into NDP-sugars and from there on to hemicellulose. F6P can be metabolised further into phenylalanine which via the monolignol biosynthesis pathway produce hydroxycinnamyl alcohol monomers (figure 12). In poplar, lignin predominantly consists of the monolignols guaiacyl (G) and syringyl (S) (Mellerowicz *et al.*, 2001).

In order to investigate the influence of the diurnal cycle on wood biosynthesis, we performed two separate labelling experiments, one in the afternoon (experiment 1) and one in the morning (experiment 2) (figure 13). In both experiments we sampled six time points over 24 hours. In experiment 1, three tissue types were harvested; source leaf, stem phloem, and developing wood. In experiment 2, only developing wood was harvested. Samples were analysed by mass spectrometry (GC- and LC-MS) and NMR. We focused our analysis on sugars and metabolites closely linked to cell wall biosynthesis.

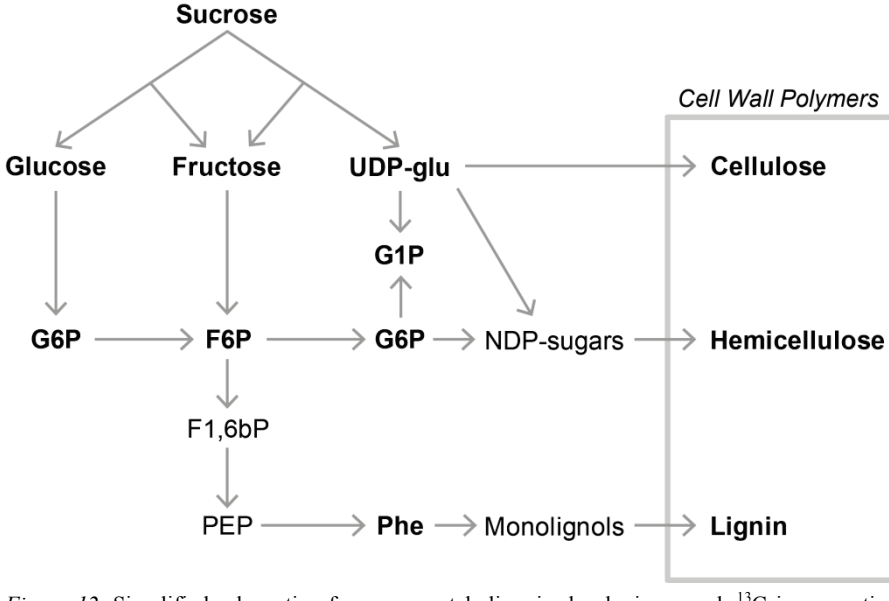


Figure 12. Simplified schematic of sucrose metabolism in developing wood. ^{13}C incorporation was determined for metabolites indicated in bold and also for the main cell wall polymers hemicellulose, cellulose, and lignin. Abbreviations: F6P; fructose-6-phosphate, G6P; glucose-6-phosphate, G1P; glucose-1-phorphate, UDP-glu; UDP-glucose, Phe; phenylalanine, PEP; phosphoenolpyruvate, F1,6bP; frucsote-1,6-bisphosphate.

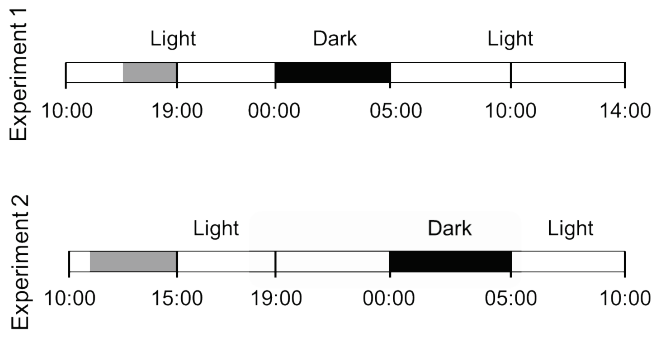


Figure 13. Time course of experiment 1 and experiment 2 for two months old hybrid aspen. The grey zone represents the 4 h $^{13}CO_2$ burst. Light period is indicated in white and dark in black. Samples were collected at six time points.

Source leaves and stem phloem had similar temporal labelling pattern for sucrose, peaking at the end of the 4 h $^{13}CO_2$ feeding. This similarity in labelling is in agreement with the passive phloem loading mechanism in popular trees which suggests that leaf and phloem sucrose pools are directly connected via plasmodesmata, implying a symplasmic phloem loading. In developing wood the ^{13}C-incorporation in sucrose did not peak until after nine hours after the introduction of the tracer. This delay could be caused by transport steps along the radial transport route from phloem to developing wood (figure 14A). Initially sucrose would follow a symplasmic route along the ray cells, from the ray cell an active import is required into the developing wood from the apoplasm. The apoplasmic transport in poplar wood has been estimated to be 8-80 times slower than the symplasmic transport (Sauter & Kloth, 1986). Taken together the apoplasmic and active transport steps in the radial phloem to developing wood provide a plausible explanation for the delay in ^{13}C- incorporation seen in sucrose in the developing wood compared to stem phloem. It has been shown that sucrose is to a large extent metabolised by invertases to fructose and glucose in the developing wood. Transgenic trees with only 4 % sucrose synthase activity, the other sucrose cleaving enzyme, had no obvious growth defects, which suggests that invertase can compensate for the reduced activity of sucrose synthase (Gerber *et al.*, 2014). Even so, when looking at the ^{13}C-incorporation in glucose and fructose it was much lower than what we observed for sucrose. Sucrose had 28 % ^{13}C-incorporation nine hours after the introduction of $^{13}CO_2$ while the corresponding incorporation for glucose was 3 % and 2 % for fructose. A possible explanation for this discrepancy could be the fast turnover of the glucose and fructose pools into hexose phosphates. Indeed the labelling patterns we obtained for the hexose phosphates (26 %) and UDP-glucose (32 %) in developing wood were very similar to what we found for sucrose (see figure 3 in paper III). To get an idea of the pool size of the sugars in developing wood we quantified UDP-

glucose levels and found that they were highest at the 10 o'clock time point with 2.2 µmol/g FW and lowest just before the dark period (midnight) with 1.4 µmol/g FW.

The ^{13}C-incorporation of phenylalanine, the precursor for lignin biosynthesis, peaked after nine hours and was estimated to 30 % in developing wood in a time course experiment referred to as experiment 2. This was comparable with sucrose, the reason for this is most likely that phenylalanine in developing wood derives from imported sucrose.

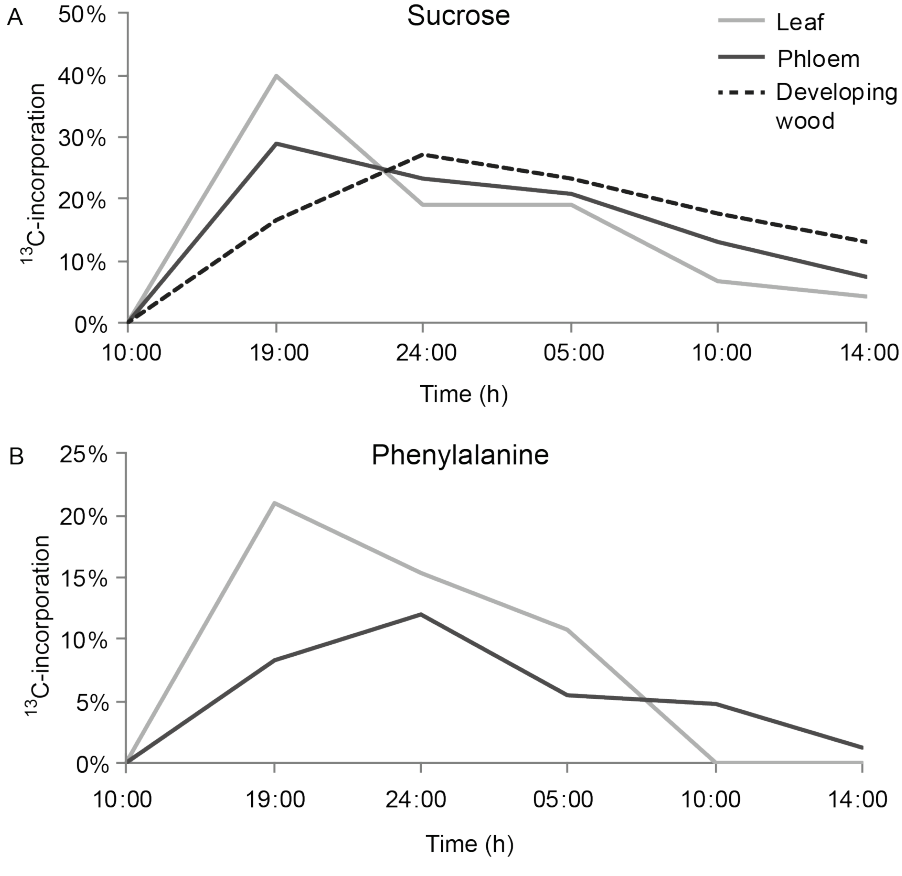


Figure 14. ^{13}C-incorpration data from experiment 1. A) incorporation in sucrose for source leaf, stem phloem, and developing wood. There is a four hour delay in the peak of incorporation in developing wood compared to source leaf and stem phloem. B) Incorporation in phenylalanine for source leaf and stem phloem. In opposite to sucrose phenylalanine had a delay before the labelled carbon reaches the stem phloem. Incorporation in developing wood could not be determined for this metabolite.

We quantified phenylalanine in developing wood over the time series and found that, similarly as for UDP-glucose, the highest levels were observed at ten o'clock in the morning with 225 nmol/g FW. The lowest level however was at the end of the dark period with 171 nmol/g FW. For experiment 1 it was possible to compare ^{13}C-incorporation in phenylalanine between tissue types. In source leaves the ^{13}C-incorporation peaked right after the end of the 4 h labelling burst and then continued dropping, while in phloem it was delayed an additional four hours and peaked just before the dark period (figure 14B). Unfortunately the ^{13}C-incorporation for developing wood in phenylalanine could not be determined due to a disturbance from another compound (4-hydroxy-benzoate) at one time point.

We used 2D-NMR to analyse the ^{13}C enrichment in cell wall polymers for both experiment 1 and experiment 2. In the first experiment, where the tracer was fed in the afternoon, there was a clear enrichment after four hours. The time course suggested that the dark period possibly might have an effect on the ^{13}C deposition in the cell walls. This was further supported in the second time course where the tracer was fed in the morning. A PCA of each experiment shows a clear separation between lignin peaks and cell wall carbohydrate peaks in the loadings plot and a possible correlation to the dark period in the score plot for experiment 1 and a stronger correlation in experiment 2 (see figure 6 in paper III). A possible reason for the more pronounced correlation in the experiment where the tracer was fed in the morning (experiment 2) might be that the tree is given more time to metabolise the $^{13}CO_2$ and hence for the label to reach the cell wall polymers.

In summary, the $^{13}CO_2$ labelling system that we used to investigate the carbon allocation in developing wood proved to be very functional. We could study carbon flux from source leaves to sink tissue in a woody plant. In our data we found evidence that the diurnal cycle possibly has an effect on wood cell wall biosynthesis.

4.3 The Beginning of Future Analysis of Plant Metabolites (Paper IV)

In Paper IV we developed a technique using fluorescence-activated cell sorting (FACS) in combination with mass spectrometry for metabolite profiling of specific cell types in Arabidopsis roots.

A plant consists of many different cell types that are separated in specialised tissues and plant organs. Plant organs are complicated structures and consist of many physiologically and morphologically different cell types. Each cell type has their own set of transcript, protein, and metabolite profiles

which gives them a specific role and function in the tissue where they are located and enables the integration with the plant as a whole (Wang *et al.*, 2012). To study systems with complex structures such as plants, it is informative to study specific cell types rather than whole plants or whole organs, as is done in most omics techniques. Undoubtedly information is lost that could give important insight into the function of individual cell types. There are examples of plant cell cultures of specific cell types (Endo *et al.*, 2009) but in such approaches the cells have been taken out of their natural environment and lost the interaction with neighbouring cells and tissues. Hence, it is difficult to interpret the biology in cultured cells and separate it from treatment effects. To be able to make a comprehensive and reliable study of the molecular biology, development and physiology of multicellular organisms, it is crucial to develop adequate methods. The behaviour of individual cell types examined must be in correspondence with their behaviour in the native environment within the organism (Wang *et al.*, 2012).

There has been a rapid development within the field of mass spectrometry, with new instruments being released with increasingly better sensitivity and resolution. This opens up for analysing much smaller sample concentrations and volumes than was previously possible (Andrews *et al.*, 2011; Makarov *et al.*, 2009; Marshall & Hendrickson, 2008). It is now possible to analyse the metabolite composition from very small amounts of plant tissue or even individual cells (Misra *et al.*, 2014; Rubakhin *et al.*, 2013; Heinemann & Zenobi, 2011).

To best use the benefits obtained from these sensitive analyses the sampling techniques must also be improved. From a historical perspective two main strategies have been used for isolation of specific cell types, dissection of a desired cell type or isolation of a cell type based on unique properties. More modern methods for dissection or cell sorting are laser micro-dissection (LCM) and FACS, which are based on mechanical and optical/mechanical separation techniques. LCM has been used for investigation of the Arabidopsis shoot meristem (Torti *et al.*, 2012) and also in combination with GC-TOF-MS for metabolic profiling of vascular bundles in Arabidopsis (Schad *et al.*, 2005). Drawbacks with LCM is that it is labour intense, time consuming, not suitable for collection of very small, complex three dimensional structures and that the cell yield is low (Wang *et al.*, 2012). Flow cytometry, or FACS, has emerged as a powerful technique for isolating specific cell types from plant tissues (Dudley *et al.*, 2010; Rochat *et al.*, 2010). Due to the presence of the cell wall, plants are more complicated to work with compared to cells without a cell wall. Hence, enzymatic degradation of the cell wall is required to isolate free protoplasts from the analysed tissue (Dolezel *et al.*, 2007). The technique

works by differentiating fluorescent cells in a liquid stream illuminated by a laser beam. The fluorescent light from the cells gives rise to electrical signals. The stream is broken into droplets downstream of the laser and the electrical signals are used to give charges to the drops. The charged drops pass between two charged plates and are deflected to fraction specific vessels. The technique can analyse thousands to millions of cells in a short time (Rubakhin *et al.*, 2013; Bonner *et al.*, 1972)

Transgenic lines of Arabidopsis expressing green fluorescent protein (GFP) in specific cell types have been used successfully for protoplasting and cell sorting by FACS. Examples of such studies are the isolation of specific cell types in roots (Albinsky *et al.*, 2010) and shoot apices (Yadav *et al.*, 2009). The technique has mainly been used for transcript profiling and cell type specific analysis of proteins (Petricka *et al.*, 2012; Rochat *et al.*, 2010; Allen *et al.*, 2003). It has also been combined with metabolite analysis (Moussaieff *et al.*, 2013; Pěnčík *et al.*, 2013; Petersson *et al.*, 2009). Moussaieff *et al.* (2013) pointed out that a problem with combining sorted cells with metabolite analysis is the high concentration of phosphates and other compounds from the sorting buffer. We found that this could be solved by replacing the FACS sheath buffer with a 0.7 % NaCl solution. The sodium chloride worked equally well as a sorting buffer for the protoplasts and did not disturb the metabolite analysis as the FACS buffer did. To reduce the stress on the root protoplasts we used a larger nozzle aperture and as a result a lower sheath gas pressure and slower sorting rate. Thus the protoplasts were suspended in a larger volume of buffer and needed to be concentrated before further analysis. The most efficient way turned out to be spinning down the protoplasts after sorting and discarding the buffer solution. This removed most of the NaCl solution and metabolite extraction could be performed with minimal losses, resulting in high sensitivity and good chromatographic separation in the GC-MS analysis (an overview of the method is found in figure 1, paper IV).

To validate the method we divided sorted protoplasts from wild type Arabidopsis into four samples that were then sorted after each other with 30 min delay. Thus sorting started directly after isolation or after 40 min, 60 min or 120 min after protoplast isolation. To verify the robustness of the method the sorting was executed on ten different days over a three month period. To minimise the biological variation in metabolite levels the protoplasts were sorted at the same time each day. Collected fractions contained one million cells and the metabolic profiles obtained from the GC-MS analysis was analysed by multivariate statistics.

Possible effects of cell sorting time and sampling day were investigated by PCA (figure 1B in paper IV). We found no trends related to either sorting time

or sampling day. As a complement to the PCA, an OPLS-DA analysis was performed, grouping the samples after sorting time and sorting day, but no significant model could be generated. Thus, no systematic variation in the data related to the sorting time or the day of sampling could be found. With this verification in combination with previous controls described in Petersson et al. (2009) we concluded that the method was robust, reproducible and could be used for cell-type specific analysis of metabolites in Arabidopsis.

For further method development we used a transgenic Arabidopsis line with GFP expression in the cortical and endodermal cell files of the root tip (line J0571) (Haseloff, 1999). This line has a stable and distinct GFP expression and has been used in several studies of root development (Albinsky *et al.*, 2010; Sozzani *et al.*, 2010; Allen *et al.*, 2009; Petersson *et al.*, 2009; Ubeda-Tomas *et al.*, 2008). Samples were sorted by FACS, collecting around one million GFP+ protoplasts and 4-5 million GFP- protoplasts. Protoplasts were spun down and the pellet was frozen in liquid nitrogen and stored at -80°C. The results showed, in accordance with the wild type method validation, that the experiments were stable between days regarding both isolation efficiency and the ratio of GFP+/GFP- cells isolated. Furthermore, we wanted to investigate possible differences between root tissue and sorted protoplasts. Again the J0571 line was used and analysed by GC-MS. The total ion chromatogram (TIC) showed that the metabolite concentration of root tissue and isolated protoplasts were in the same range (figure 3A, paper IV). We found that roots and sorted cells separated very well, probably due to the lack of cell walls in the isolated protoplasts. A PCA revealed that the separation between roots and sorted cells was mainly due to the higher abundance of plant sterols and long-chain carbon compounds in the protoplasts while the roots were found to have more sugars and amino acids. The largest differences between the sample types were the change in carbohydrate profiles. A lot of unidentified carbon chains were found in the protoplasts but they could not be identified. It is likely that they originate from the plasma membrane, but it could also be due to losses during sample extraction.

To compare GFP+ protoplasts with GFP- protoplasts the root samples were excluded from the model and OPLS-DA was used to analyse the difference between the two groups. GFP+ and GFP- samples could be separated and loadings showed that asparagine, α-ketoglutarate, glutamine, glutarate, malate, citrate and phosphoric acid correlated with the GFP+ protoplasts. Only an unspecified disaccharide and glycerate showed a clear correlation to GFP- samples.

Compared to previously published methods (Moussaieff *et al.*, 2013; Albinsky *et al.*, 2010), our method was more straightforward due to being

faster, simpler and more gentle to the protoplasts without losing accuracy, robustness or sensitivity. We found that different cell types had different metabolic profiles and that the concentration of some metabolites was different between these cell types. The cells sorted by this method could also be used for other analyses such as protein and transcript analysis.

There is a need for new methods with a high spatial resolution for metabolite analyses (Sweetlove *et al.*, 2014; Kruger *et al.*, 2012). It is very challenging to do metabolomics or metabolite profiling on isolated cell populations, but the development of new mass spectrometers opens up for new possibilities (Oikawa & Saito, 2012). It is important to keep in mind that the cells are subject to a great measure of stress during both protoplast isolation and possibly also during FACS analysis which might affect the metabolic profiles of the cells. With the method in paper IV we have shown that it is possible to use FACS for protoplast isolation and cell sorting of specific cell types and then analyse their metabolic profiles by GC-TOF-MS. Our analysis was untargeted but it could also be used for targeted analysis by LC-MS. Since the GFP expression does not affect the metabolite content it is also possible to combine this method with the labelling method used in paper II for cell type specific flux analysis. This is an interesting direction for further development of techniques to study metabolic flux in plants.

5 Conclusions and Future Perspectives

In paper I, we characterised a series of mutants lacking one or both genes of mitochondrial malate dehydrogenase. We found that mMDH has a complex respiration-controlling role. We continued to study one of the single mutants, *mmdh1*, in paper II**.** We developed an analysis method using ^{13}C-labelling to investigate primary carbon metabolism in mature plant leaves. We found that *mmdh1* had a metabolic phenotype that revealed an altered flux though photorespiration, especially under low CO_2 conditions. Thus, by looking at the changes in metabolism from a more global perspective we could deepen our understanding regarding the role of mMDH in photorespiration, respiration and cellular redox balance.

Emphasis was put on constructing a labelling chamber and developing MS methods to detect metabolites in primary carbon metabolism, but also how to calculate and estimate the ^{13}C-incorporation from the acquired data. However, we did not reach as far as to the integration of our labelling estimations with computer models of plant metabolism. This would be the logical next step to develop the method further. Isotopically non-stationary metabolic flux analysis (INST-MFA) has previously been used to estimate fluxes in photoautotrophic bacteria (Young *et al.*, 2011) and in a recent paper this method was adopted to measure fluxes in Arabidopsis leaves (Ma *et al.*, 2014). Ma *et al*. 2010 is the first publication where dynamic ^{13}C flux analysis has been successfully applied to map carbon fluxes in intact plants. An interesting aspect of this analysis is that it provides better flux resolution and does not require direct pool size measurements. However the method does not address the challenge of subcellular compartmentation.

Subcellular compartmentation gives plants more metabolic flexibility, specialisation and regulation, but it also complicates metabolic analyses and modelling since it makes the structure of the metabolic network more complex. The presence of large storage compartments such as the vacuoles and the

metabolically active plastids makes the task even harder compared with other organisms (Allen *et al.*, 2009). The method we present measures the metabolic turnover in leaves but it cannot distinguish between the subcellular localisation of metabolite pools. Heise *et al.* (2014) presents a protocol for estimating flux of photosynthetic carbon metabolism in whole plants where they use non-aqueous fractionation to address the issue of subcellular compartments. The protocol is estimated to take 8-10 weeks to perform which must be considered as too labour intense and time consuming to be an option for a high throughput analysis. Even so, the Heise study is an important and significant progress in the field of dynamic flux estimations in intact plants.

The practical and theoretical aspects of ^{13}C-labelling of plants which were learned from paper II**,** were used as a base for the experimental setup in paper III. In paper II the plants were sampled during continuous labelling and thus the ^{13}C-incoproration continually increased during the time series, and only leaves were harvested. In paper III, we aimed to follow the ^{13}C-incorporation from source leaf, via the stem phloem to its final destination in the developing wood. Due to the considerably larger size of two month old hybrid aspen trees compared to Arabidopsis plants, a tent was used as enclosure during the $^{13}CO_2$ treatment rather than the labelling chamber. This was practically possible since the trees were not sampled during labelling but before and after a 4 h $^{13}CO_2$ burst. From each of the ^{13}C treated trees leaf, phloem and developing wood tissue was sampled at six time points. Since carbon transport mainly takes place in the form of sucrose the metabolite analysis was concentrated to sucrose, hexoses and hexose phosphates and also phenylalanine, which is the precursor of lignin. In total eight metabolites were analysed by either GC-MS or a targeted LC-MS analysis. In developing wood UDP-glucose and phenylalanine were quantified, while for sucrose, fructose, glucose and the sugar phosphates the relative abundance was determined. We found previously unrecognised temporal patterns in wood biosynthesis and an indication that the diurnal cycle served as a cue in the regulation of carbon allocation in developing wood. The labelling system proved to be a viable way for studying source to sink carbon fluxes in trees. Paper III focus on the incorporation patterns of ^{13}C in developing wood but the study will continue by investigating transport between tissues and incorporation patterns of metabolites from source to sink.

Another promising way to study transport from source leaf to sink tissue is to combine the labelling chamber in paper II with an induced senescence system, such as individually darkened leaves (IDL), where leaves on a plant are covered so that little or no light can reach the leaf (Brouwer *et al.*, 2012;

Keech *et al.*, 2007). $^{13}CO_2$ labelling of whole plants with IDLs opens up for the possibility to investigate the import of newly synthesised metabolites from the leaves in light to the covered leaves. In such experiment the labelling should preferably be performed like in paper III, with a burst of $^{13}CO_2$ rather than continuous labelling, as this makes it possible to determine the time of transport between tissues.

Several recent reviews have described the increasing need for new metabolite analysis techniques with higher spatial resolution (Sweetlove *et al.*, 2014; Kruger *et al.*, 2012; Wang *et al.*, 2012). In Paper IV we present a method for fast and robust analysis of metabolic profiles in cell type specific cells. To study cell fate specification and cell type specific response to the environment cell type specific resolution is a requirement. FACS is a promising method for cell type specific metabolome studies due to the substantially higher cell yields compared to alternative methods. FACS in combination with high sensitivity MS is especially promising (Wang *et al.*, 2012). For a method to be useful for biological research it has to be able to cover a wide range of metabolites or quantify a few metabolites in a targeted manner with high precision. It must have a sufficient throughput to allow meaningful statistical analyses and be robust and reproducible. Our method fulfils these requirements and can in its current state be used to address a range of biological questions. Of special interest in this context is the possibility to combine the cell type specific metabolite analysis from paper IV with the labelling method from paper II. Many of the metabolites analysed in paper IV exhibited high enough abundance for there to be a viable chance of detecting labelled isotopomers. This would open up for not just metabolic profiling in specific cell types but also the determination of carbon fluxes between different tissue types.

References

Aharoni, A. & Brandizzi, F. (2012). High-Resolution Measurements in Plant Biology. *Plant Journal,* 70(1), pp. 1-4.

Albinsky, D., Sawada, Y., Kuwahara, A., Nagano, M., Hirai, A., Saito, K. & Hirai, M.Y. (2010). Widely targeted metabolomics and coexpression analysis as tools to identify genes involved in the side-chain elongation steps of aliphatic glucosinolate biosynthesis. *Amino acids,* 39(4), pp. 1067-1075.

Allen, D.K., Libourel, I.G.L. & Shachar-Hill, Y. (2009). Metabolic flux analysis in plants: coping with complexity. *Plant Cell and Environment,* 32(9), pp. 1241-1257.

Allen, D.K., Shachar-Hill, Y. & Ohlrogge, J.B. (2007). Compartment-specific labeling information in 13C metabolic flux analysis of plants. *Phytochemistry,* 68(16-18), pp. 2197-2210.

Allen, J., Davey, H.M., Broadhurst, D., Heald, J.K., Rowland, J.J., Oliver, S.G. & Kell, D.B. (2003). High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nature biotechnology,* 21(6), pp. 692-696.

Andrews, G.L., Simons, B.L., Young, J.B., Hawkridge, A.M. & Muddiman, D.C. (2011). Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (TripleTOF 5600). *Analytical chemistry,* 83(13), pp. 5442-5446.

Ap Rees, T., Fuller, W. & Wright, B. (1976). Pathways of carbohydrate oxidation during thermogenesis by the spadix of *Arum maculatum*. *Biochimica et Biophysica Acta (BBA)-General Subjects,* 437(1), pp. 22-35.

Arrivault, S., Guenther, M., Ivakov, A., Feil, R., Vosloh, D., Van Dongen, J.T., Sulpice, R. & Stitt, M. (2009). Use of reverse-phase liquid chromatography, linked to tandem mass spectrometry, to profile the Calvin cycle and other metabolic intermediates in Arabidopsis rosettes at different carbon dioxide concentrations. *The Plant Journal,* 59(5), pp. 826-839.

Baxter, C.J., Liu, J.L., Fernie, A.R. & Sweetlove, L.J. (2007). Determination of metabolic fluxes in a non-steady-state system. *Phytochemistry,* 68(16-18), pp. 2313-2319.

Beevers, H. & Gibbs, M. (1954). Participation of the Oxidative Pathway in Yeast Respiration. *Nature,* 173(4405), pp. 640-641.
Bloom, B. & Stetten Jr, D. (1953). PATHWAYS OF GLUCOSE CATABOLISM[1]. *Journal of the American Chemical Society,* 75(21), pp. 5446-5446.
Bobák, M. & Nečesaný, V. (1967). Changes in the formation of the lignified cell wall within a twenty-four hour period. *Biologia Plantarum,* 9(3), pp. 195-201.
Bonner, W., Hulett, H., Sweet, R. & Herzenberg, L. (1972). Fluorescence activated cell sorting. *Review of Scientific Instruments,* 43(3), pp. 404-409.
Borowitz, M.J., Stein, R.B. & Blum, J.J. (1977). Quantitative analysis of the change of metabolite fluxes along the pentose phosphate and glycolytic pathways in Tetrahymena in response to carbohydrates. *Journal of Biological Chemistry,* 252(5), pp. 1589-1605.
Brouwer, B., Ziolkowska, A., Bagard, M., Keech, O. & Gardestrom, P. (2012). The impact of light intensity on shade-induced leaf senescence. *Plant Cell and Environment,* 35(6), pp. 1084-1098.
Bykova, N., Keerberg, O., Pärnik, T., Bauwe, H. & Gardeström, P. (2005). Interaction between photorespiration and respiration in transgenic potato plants with antisense reduction in glycine decarboxylase. *Planta,* 222(1), pp. 130-140.
Calvin, M. (1962). PATH OF CARBON IN PHOTOSYNTHESIS. *Science,* 135(3507), pp. 879-&.
Carter, A.D., Bonyadi, R. & Gifford, M.L. (2013). The use of fluorescence-activated cell sorting in studying plant development and environmental responses. *International Journal of Developmental Biology,* 57(6-8), pp. 545-552.
Cook, D., Fowler, S., Fiehn, O. & Thomashow, M.F. (2004). A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America,* 101(42), pp. 15243-15248.
Dolezel, J., Greilhuber, J. & Suda, J. (2007). *Flow Cytometry with Plant Cells: Analysis of Genes, Chromosomes and Genomes*: John Wiley & Sons.
Dudley, E., Yousef, M., Wang, Y. & Griffiths, W. (2010). Targeted metabolomics and mass spectrometry. *Adv Protein Chem Struct Biol,* 80, pp. 45-83.
Endo, S., Pesquet, E., Yamaguchi, M., Tashiro, G., Sato, M., Toyooka, K., Nishikubo, N., Udagawa-Motose, M., Kubo, M. & Fukuda, H. (2009). Identifying new components participating in the secondary cell wall formation of vessel elements in Zinnia and Arabidopsis. *The Plant Cell Online,* 21(4), pp. 1155-1165.
Fernie, A. & Morgan, J. (2013). Analysis of metabolic flux using dynamic labelling and metabolic modelling. *Plant, cell & environment,* 36(9), pp. 1738-1750.
Fernie, A.R., Geigenberger, P. & Stitt, M. (2005). Flux an important, but neglected, component of functional glenomics. *Current Opinion in Plant Biology,* 8(2), pp. 174-182.

Fernie, A.R., Tadmor, Y. & Zamir, D. (2006). Natural genetic variation for improving crop quality. *Current Opinion in Plant Biology,* 9(2), pp. 196-202.

Fiehn, O. (2002). Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology,* 48(1-2), pp. 155-171.

Fiehn, O., Kopka, J., Trethewey, R.N. & Willmitzer, L. (2000). Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Analytical Chemistry,* 72(15), pp. 3573-3580.

Flaumenhaft, E., Uphaus, R. & Kats, J. (1970). Isotope biology of ^{13}C extensive incorporation of highly enriched ^{13}C in the Alga Chlorella vulgaris. *Biochimica et Biophysica Acta (BBA)-General Subjects,* 215(3), pp. 421-429.

Florian, A., Araújo, W.L. & Fernie, A.R. (2013). New insights into photorespiration obtained from metabolomics. *Plant Biology,* 15(4), pp. 656-666.

Gerber, L., Zhang, B., Roach, M., Rende, U., Gorzsás, A., Kumar, M., Burgert, I., Niittylä, T. & Sundberg, B. (2014). Deficient sucrose synthase activity in developing wood does not specifically affect cellulose biosynthesis, but causes an overall decrease in cell wall polymers. *New Phytologist,* 203(4), pp. 1220-1230.

Gibbs, M. & Beevers, H. (1955). Glucose Dissimilation in the Higher Plant. Effect of Age of Tissue. *Plant Physiology,* 30(4), pp. 343-347.

Giraudeau, P., Massou, S., Robin, Y., Cahoreau, E., Portais, J.-C. & Akoka, S. (2011). Ultrafast Quantitative 2D NMR: An Efficient Tool for the Measurement of Specific Isotopic Enrichments in Complex Biological Mixtures. *Analytical chemistry,* 83(8), pp. 3112-3119.

Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G. & Kell, D.B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology,* 22(5), pp. 245-252.

Haseloff, J. (1999). GFP variants for multispectral imaging of living cells. *Methods in Cell Biology,* 58, pp. 139-151.

Hasunuma, T., Harada, K., Miyazawa, S.I., Kondo, A., Fukusaki, E. & Miyake, C. (2010). Metabolic turnover analysis by a combination of in vivo C-13-labelling from (CO2)-C-13 and metabolic profiling with CE-MS/MS reveals rate-limiting steps of the C-3 photosynthetic pathway in Nicotiana tabacum leaves. *Journal of Experimental Botany,* 61(4), pp. 1041-1051.

Hegeman, A.D. (2010). Plant metabolomics-meeting the analytical challenges of comprehensive metabolite analysis. *Briefings in Functional Genomics,* 9(2), pp. 139-148.

Heinemann, M. & Zenobi, R. (2011). Single cell metabolomics. *Current Opinion in Biotechnology,* 22(1), pp. 26-31.

Heise, R., Arrivault, S., Szecowka, M., Tohge, T., Nunes-Nesi, A., Stitt, M., Nikoloski, Z. & Fernie, A.R. (2014). Flux profiling of photosynthetic carbon metabolism in intact plants. *Nature Protocols,* 9(8).

Hiller K., Metallo C. & G., S. (2011). Elucidation of Cellular Metabolism Via Metabolomics and Stable-Isotope Assisted Metabolomics. *Current Pharmaceutical Biotechnology,* 12(7), pp. 1075-1086.

Hosoo, Y., Yoshida, M., Imai, T. & Okuyama, T. (2002). Diurnal difference in the amount of immunogold-labeled glucomannans detected with field emission scanning electron microscopy at the innermost surface of developing secondary walls of differentiating conifer tracheids. *Planta,* 215(6), pp. 1006-1012.

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. *Phytochemistry,* 68(16-18), pp. 2258-2272.

Huppe, H.C. & Turpin, D.H. (1994). Integration of Carbon and Nitrogen Metabolism in Plant and Algal Cells. *Annual Review of Plant Physiology and Plant Molecular Biology,* 45(1), pp. 577-607.

Jackson, J. (1991). A user's guide to principal components.

Keech, O., Pesquet, E., Ahad, A., Askne, A., Nordvall, D., Vodnala, S.M., Tuominen, H., Hurry, V., Dizengremel, P. & Gardestroem, P. (2007). The different fates of mitochondria and chloroplasts during dark-induced senescence in Arabidopsis leaves. *Plant, cell & environment,* 30(12), pp. 1523-1534.

Koek, M.M., Jellema, R.H., van der Greef, J., Tas, A.C. & Hankemeier, T. (2011). Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics,* 7(3), pp. 307-328.

Kruger, N.J., Masakapalli, S.K. & Ratcliffe, R.G. (2012). Strategies for investigating the plant metabolic network with steady-state metabolic flux analysis: lessons from an Arabidopsis cell culture and other systems. *Journal of Experimental Botany,* 63(6), pp. 2309-2323.

Kueger, S., Steinhauser, D., Willmitzer, L. & Giavalisco, P. (2012). High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *The Plant Journal,* 70(1), pp. 39-50.

Linden, J., Schilling, N., Brackenhofer, H. & Kandler, O. (1975). Asymmetric Labelling of Maltose during Photosynthesis in $^{14}CO_2$. *Zeitschrift für Pflanzenphysiologie,* 76(2), pp. 176-181.

Liu, J.L., Grieson, C.S., Webb, A.A.R. & Hussey, P.J. (2010). Modelling dynamic plant cells. *Current Opinion in Plant Biology,* 13(6), pp. 744-749.

Lu, W., Bennett, B.D. & Rabinowitz, J.D. (2008). Analytical strategies for LC-MS-based targeted metabolomics. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences,* 871(2), pp. 236-242.

Ma, F., Jazmin, L.J., Young, J.D. & Allen, D.K. (2014). Isotopically nonstationary ^{13}C flux analysis of changes in *Arabidopsis thaliana* leaf metabolism due to high light acclimation. *Proceedings of the National Academy of Sciences of the United States of America,* 111(47), pp. 16967-16972.

Mahboubi, A., Ratke, C., Gorzsás, A., Kumar, M., Mellerowicz, E.J. & Niittylä, T. (2013). Aspen SUCROSE TRANSPORTER3 Allocates Carbon into Wood Fibers. *Plant Physiology,* 163(4), pp. 1729-1740.

Makarov, A., Denisov, E. & Lange, O. (2009). Performance Evaluation of a High-field Orbitrap Mass Analyzer. *Journal of the American Society for Mass Spectrometry,* 20(8), pp. 1391-1396.

Mallet, A. & Down, S. (2010). *Dictionary of mass spectrometry*: John Wiley & Sons.

Marshall, A.G. & Hendrickson, C.L. (2008). High-resolution mass spectrometers. *Annu. Rev. Anal. Chem.,* 1, pp. 579-599.

Matsuda, F., Wakasa, K. & Miyagawa, H. (2007). Metabolic flux analysis in plants using dynamic labeling technique: Application to tryptophan biosynthesis in cultured rice cells. *Phytochemistry,* 68(16-18), pp. 2290-2301.

Maurino, V.G. & Peterhansel, C. (2010). Photorespiration: current status and approaches for metabolic engineering. *Current Opinion in Plant Biology,* 13(3), pp. 248-255.

Mellerowicz, E., Baucher, M., Sundberg, B. & Boerjan, W. (2001). Unravelling cell wall formation in the woody dicot stem. In: Carpita, N.C., Campbell, M. & Tierney, M. (eds) *Plant Cell Walls* Springer Netherlands, pp. 239-274. Available from: http://dx.doi.org/10.1007/978-94-010-0668-2_15.

Misra, B.B., Assmann, S.M. & Chen, S. (2014). Plant single-cell and single-cell-type metabolomics. *Trends in plant science*.

Moussaieff, A., Rogachev, I., Brodsky, L., Malitsky, S., Toal, T.W., Belcher, H., Yativ, M., Brady, S.M., Benfey, P.N. & Aharoni, A. (2013). High-resolution metabolic mapping of cell types in plant roots. *Proceedings of the National Academy of Sciences,* 110(13), pp. E1232-E1241.

Nunes-Nesi, A., Carrari, F., Lytovchenko, A., Smith, A.M.O., Loureiro, M.E., Ratcliffe, R.G., Sweetlove, L.J. & Fernie, A.R. (2005). Enhanced photosynthetic performance and growth as a consequence of decreasing mitochondrial malate dehydrogenase activity in transgenic tomato plants. *Plant Physiology,* 137(2), pp. 611-622.

Oikawa, A. & Saito, K. (2012). Metabolite analyses of single cells. *The Plant Journal,* 70(1), pp. 30-38.

Oliver, S.G., Winson, M.K., Kell, D.B. & Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology,* 16(9), pp. 373-378.

Pěnčík, A., Simonovik, B., Petersson, S.V., Henyková, E., Simon, S., Greenham, K., Zhang, Y., Kowalczyk, M., Estelle, M., Zažímalová, E., Novák, O., Sandberg, G. & Ljung, K. (2013). Regulation of Auxin Homeostasis and Gradients in Arabidopsis Roots through the Formation of the Indole-3-Acetic Acid Catabolite 2-Oxindole-3-Acetic Acid. *The Plant Cell Online,* 25(10), pp. 3858-3870.

Peterhansel, C., Horst, I., Niessen, M., Blume, C., Kebeish, R., Kürkcüoglu, S. & Kreuzaler, F. (2010). Photorespiration. *The Arabidopsis Book*, p. e0130.

Petersson, S.V., Johansson, A.I., Kowalczyk, M., Makoveychuk, A., Wang, J.Y., Moritz, T., Grebe, M., Benfey, P.N., Sandberg, G. & Ljung, K. (2009). An Auxin Gradient and Maximum in the Arabidopsis Root Apex Shown

by High-Resolution Cell-Specific Analysis of IAA Distribution and Synthesis. *The Plant Cell Online,* 21(6), pp. 1659-1668.
Petricka, J.J., Schauer, M.A., Megraw, M., Breakfield, N.W., Thompson, J.W., Georgiev, S., Soderblom, E.J., Ohler, U., Moseley, M.A., Grossniklaus, U. & Benfey, P.N. (2012). The protein expression landscape of the Arabidopsis root. *Proceedings of the National Academy of Sciences,* 109(18), pp. 6811-6818.
Pichersky, E. & Gang, D.R. (2000). Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends in Plant Science,* 5(10), pp. 439-445.
Plaut, Z. & Gibbs, M. (1970). Glycolate formation in intact spinach chloroplasts. *Plant physiology,* 45(4), pp. 470-474.
Ramautar, R., Somsen, G.W. & de Jong, G.J. (2009). CE-MS in metabolomics. *Electrophoresis,* 30(1), pp. 276-291.
Rao, Z., Qin, H., Wei, Y., Zhou, Y., Zhang, G., Zhang, F., Shao, Y., Huang, J. & Wu, X.a. (2014). Development of a dynamic multiple reaction monitoring method for determination of digoxin and six active components of *Ginkgo biloba* leaf extract in rat plasma. *Journal of Chromatography B,* 959, pp. 27-35.
Ratcliffe, R.G. & Shachar-Hill, Y. (2006). Measuring multiple fluxes through plant metabolic networks. *The Plant Journal,* 45(4), pp. 490-511.
Rochat, L., Péchy-Tarr, M., Baehler, E., Maurhofer, M. & Keel, C. (2010). Combination of fluorescent reporters for simultaneous monitoring of root colonization and antifungal gene expression by a biocontrol pseudomonad on cereals with flow cytometry. *Molecular plant-microbe interactions,* 23(7), pp. 949-961.
Rubakhin, S.S., Lanni, E.J. & Sweedler, J.V. (2013). Progress toward single cell metabolomics. *Current Opinion in Biotechnology,* 24(1), pp. 95-104.
Saito, K. & Matsuda, F. (2010). Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annual Review of Plant Biology,* 61(1), pp. 463-489.
Sauer, N. (2007). Molecular physiology of higher plant sucrose transporters. *FEBS Letters,* 581(12), pp. 2309-2317.
Sauer, U. (2004). High-throughput phenomics: experimental methods for mapping fluxomes. *Curr Opin Biotechnol,* 15(1), pp. 58 - 63.
Sauter, J.J. & Kloth, S. (1986). Plasmodesmatal frequency and radial translocation rates in ray cells of poplar (*Populus x canadensis* Moench 'robusta'). *Planta,* 168(3), pp. 377-380.
Schad, M., Mungur, R., Fiehn, O. & Kehr, J. (2005). Metabolic profiling of laser microdissected vascular bundles of *Arabidopsis thaliana*. *Plant Methods,* 1.
Schauer, N. & Fernie, A.R. (2006). Plant metabolomics: towards biological function and mechanism. *Trends in Plant Science,* 11(10), pp. 508-516.
Schilling, N. (1982). Characterization of maltose biosynthesis from α-d-glucose-1-phosphate in *Spinacia oleracea.* L. *Planta,* 154(1), pp. 87-93.

Solomon, O.L., Berger, D.K. & Myburg, A.A. (2010). Diurnal and circadian patterns of gene expression in the developing xylem of Eucalyptus trees. *South African Journal of Botany,* 76(3), pp. 425-439.

Sozzani, R., Cui, H., Moreno-Risueno, M.A., Busch, W., Van Norman, J.M., Vernoux, T., Brady, S.M., Dewitte, W., Murray, J.A.H. & Benfey, P.N. (2010). Spatiotemporal regulation of cell-cycle genes by SHORTROOT links patterning and growth. *Nature,* 466(7302), pp. 128-132.

Suzuki, M., Kusano, M., Takahashi, H., Nakamura, Y., Hayashi, N., Kobayashi, M., Ichikawa, T., Matsui, M., Hirochika, H. & Saito, K. (2010). Rice-Arabidopsis FOX line screening with FT-NIR-based fingerprinting for GC-TOF/MS-based metabolite profiling. *Metabolomics,* 6(1), pp. 137-145.

Sweetlove, L.J., Obata, T. & Fernie, A.R. (2014). Systems analysis of metabolic phenotypes: what have we learnt? *Trends in Plant Science,* 19(4), pp. 222-230.

Szecowka, M., Heise, R., Tohge, T., Nunes-Nesi, A., Vosloh, D., Huege, J., Feil, R., Lunn, J., Nikoloski, Z., Stitt, M., Fernie, A.R. & Arrivault, S. (2013). Metabolic Fluxes in an Illuminated Arabidopsis Rosette. *The Plant Cell Online*.

t'Kindt, R., Morreel, K., Deforce, D., Boerjan, W. & Van Bocxlaer, J. (2009). Joint GC-MS and LC-MS platforms for comprehensive plant metabolomics: Repeatability and sample pre-treatment. *Journal of Chromatography B,* 877(29), pp. 3572-3580.

Taiz, L. & Zeiger, E. (2002). Plant physiology. Third Edition. *Plant physiology. Third Edition.* Sunderland, Massachusetts, Sinauer associates, Inc., Publishers, pp. i-xxvi, 1-690.

Torti, S., Fornara, F., Vincent, C., Andrés, F., Nordström, K., Göbel, U., Knoll, D., Schoof, H. & Coupland, G. (2012). Analysis of the Arabidopsis Shoot Meristem Transcriptome during Floral Transition Identifies Distinct Regulatory Patterns and a Leucine-Rich Repeat Protein That Promotes Flowering. *The Plant Cell Online,* 24(2), pp. 444-462.

Trethewey, R.N. (2004). Metabolite profiling as an aid to metabolic engineering in plants. *Current Opinion in Plant Biology,* 7(2), pp. 196-201.

Trygg, J., Gullberg, J., Johansson, A., Jonsson, P. & Moritz, T. (2006a). Chemometrics in metabolomics—an introduction. In: *Plant metabolomics* Springer, pp. 117-128.

Trygg, J., Holmes, E. & Lundstedt, T. (2006b). Chemometrics in Metabonomics. *Journal of Proteome Research,* 6(2), pp. 469-479.

Trygg, J. & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics,* 16(3), pp. 119-128.

Ubeda-Tomas, S., Swarup, R., Coates, J., Swarup, K., Laplaze, L., Beemster, G.T.S., Hedden, P., Bhalerao, R. & Bennett, M.J. (2008). Root growth in Arabidopsis requires gibberellin/DELLA signalling in the endodermis. *Nat Cell Biol,* 10(5), pp. 625-628.

von Roepenack-Lahaye, E., Degenkolb, T., Zerjeski, M., Franz, M., Roth, U., Wessjohann, L., Schmidt, J., Scheel, D. & Clemens, S. (2004). Profiling of Arabidopsis secondary metabolites by capillary liquid chromatography

coupled to electrospray ionization quadrupole time-of-flight mass spectrometry. *Plant Physiol,* 134(2), pp. 548-59.
Wang, D., Mills, E.S. & Deal, R.B. (2012). Technologies for systems-level analysis of specific cell types in plants. *Plant Science,* 197(0), pp. 21-29.
Weckwerth, W. (2011). Green systems biology - From single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *Journal of Proteomics,* 75(1), pp. 284-305.
Wiechert, W. (2001). 13C metabolic flux analysis. *Metab Eng,* 3(3), pp. 195 - 206.
Williams, T.C.R., Miguet, L., Masakapalli, S.K., Kruger, N.J., Sweetlove, L.J. & Ratcliffe, R.G. (2008). Metabolic Network Fluxes in Heterotrophic Arabidopsis Cells: Stability of the Flux Distribution under Different Oxygenation Conditions. *Plant Physiol.,* 148(2), pp. 704-718.
Wittmann, C. & Heinzle, E. (2002). Genealogy profiling through strain improvement by using metabolic network analysis: metabolic flux genealogy of several generations of lysine-producing corynebacteria. *Appl Environ Microbiol,* 68(12), pp. 5843 - 5859.
Yadav, R.K., Girke, T., Pasala, S., Xie, M. & Reddy, G.V. (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proceedings of the National Academy of Sciences,* 106(12), pp. 4941-4946.
Yoshida, M., Hosoo, Y. & Okuyama, T. (2000). Periodicity as a Factor in the Generation of Isotropic Compressive Growth Stress Between Microfibrils in Cell Wall Formation during a Twenty-Four Hour Period. *Holzforschung.* p. 469.
Young, J.D., Shastri, A.A., Stephanopoulos, G. & Morgan, J.A. (2011). Mapping photoautotrophic metabolism with isotopically nonstationary C-13 flux analysis. *Metabolic Engineering,* 13(6), pp. 656-665.
Young, J.D., Walther, J.L., Antoniewicz, M.R., Yoo, H. & Stephanopoulos, G. (2008). An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis. *Biotechnology and Bioengineering,* 99(3), pp. 686-699.

Acknowledgements

I want to thank my supervisors Thomas and Per for taking me on as a PhD student. I thought I had an idea about what doing a PhD meant before I started, I didn't. It takes one to know one. I am privileged to have had such experienced and enthusiastic supervisors to guide me.

I want to give my special thanks to Olivier. You may not be my supervisor on paper, but we both know that's exactly what you have been. You have taught me a lot about the world of science, guided and helped me. Besides work I will also remember your advice regarding the proper (French) way to cut a cheese and were to look for chanterelles and Karl-Johan mushrooms.

Åsa and Ulrika, your opinions and suggestions regarding my project have been very valuable to me. I congratulate myself for recruiting you to my reference group.

Karin, Sara, Amir and Totte, thanks for good and fruitful collaborations. I have very much enjoyed working with you and learning about your fields of expertise.

Sara, you get an extra hug for being such great office mate. When spending so much time there why not have fun and gossip a little?

Hasse and Annika, really, what had I done without you? What I know about multivariate statistics and mass spectrometry, I've learned from you guys.

Maria A and Linus- you make a reliable safety net. Maria, how many times have we calculated all kinds of concentrations together? Linus, you have been

a great office neighbour over the years and a very valuable computer technician when I have written my thesis from everywhere but my own desk.

Kate and Jenny, you are great in so many ways. You've been there for me both workwise and privately. It's been so much fun!
Jana, you are quite new in the office so you have not seen much of me. Even so, you helped me with the last corrections before printing. Thank you!

Krister, you probably spent a couple of month or so teaching me data processing and metabolite identification in the beginning of my PhD. Now I have taught others, hopefully with the same patience that you showed me. Inga-Britt, the lab would stop without you. Thank you for all the help and for being such good mentor. Jonas and Anders, you have been a great support and encouraged me to go on when I despaired over the LC method. Anything is possible as long as you stay positive, right?

Special thanks to past and present members of the Gardeström group and Moritz/Ljung groups for good scientific discussions and fika during group meetings. Ogonna, you make the best banana cake ever! Bas, Agnieszka, Gunilla, Daria and Simon, thanks for listening to my technical details regarding mass spectrometry and ^{13}C detection strategies even though it isn't your cup of tea.

Jocke and David S, thanks for being such excellent computer support, it's been needed more often than not since all computers apparently hate me.

Inga-Lis and Gun-Britt, thanks for your patience. I have probably asked you the same questions regarding payments and what forms to fill in and how on earth to manage the SLU internal webpage at least a hundred times. Sorry, but I still don't get the travelling system.

Big thanks to the students in the two DBT-courses, Mintlab och Heavymetlab, for engaging themselves in the design and construction of the labelling chambers.

Maria K, Anna K and Hanna J, we had many good lunches, dinners and movies together, good times! Hanna, extra thanks to you for taking care of Brutus the cat over so many holidays. I'm sorry he ate your plants every time…

Michael and Matilda, Carro and Jakob, my Tomtebo family, what would I have done without you? You have been there for me in all weathers, I am lucky to have such good friends so close by! Michael, special thanks for helping out with the language correction of my thesis.

Anna J, Andreas, Ann, Jonas, Lena and Lisa, thanks for all the joy, travels, bubbles and good times that we have shared!

Ellis, you end up between colleagues, friends and family because you are all of those- my sister in crimes!

Anna and Amanda with families, well you are my family. I cherish your advice and support. Special thanks to you Amanda for our great travels together.

Mum and Dad, thank you for always being there for me. Throughout my studies you have listened and supported me even though you never quite did grasp what the heck I have been doing all these years. I love you to bits.

David, you have heard and seen more than anyone of my transition into a PhD. You have been there for me and kept me above the surface. You mean the world to me.

Last, dear grandmothers Aina and Marianne, grandfather Kurt and uncle BOM, you were with me when I started my pursuit for a PhD, but you never got to see me finish. You live forever in my heart.