Post-translational regulation of C₄-related proteins and the regulation of C₄ photosynthesis in leaves of NADP-malic enzyme monocot grasses.

By:

Emmanuel González Escobar

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

The University of Sheffield
Faculty of Science
Department of Animal and Plant Sciences

November, 2017
Abstract

C₄ photosynthesis is a complex trait that involves the efficient movement of carbon between mesophyll and bundle sheath cells, but gaps remain in our understanding on how protein phosphorylation regulates the relative flux through malate or aspartate in NADP-malic enzyme subtypes. Using mass spectrometry, 20 phosphosites in NADP-malic enzyme (ME), 9 phosphosites in aspartate aminotransferase (AAT) and 12 phosphosites in alanine aminotransferase were identified in the C₄ plant model Setaria viridis. In silico analysis shows that NADP-ME phosphorylation at Ser-343 may regulate NADP⁺ binding.

NADP-ME from illuminated leaves of Zea mays exhibited a higher affinity for L-malate at pH 8.4. High in vitro activity at pH 7.4 in illuminated leaves of Sorghum bicolor, and no change in affinity for L-malate at pH 7.4 or pH 8.4 reveals dependence on enzyme activity immediately after the onset of light. It is speculated that the same protein may be differently regulated in each grass species. The activation of AAT by C₄-acids also differed among the NADP-ME subtypes. Light activation in the presence of L-alanine and L-malate was only evident in Zea mays, whereas L-malate and L-alanine inhibited the activity in Sorghum bicolor and L-malate inhibited activity in Setaria viridis. In vitro activity assays also demonstrated that the affinity for L-aspartate was higher in darkened conditions and this activity may not be associated with C₄ photosynthesis.

The work presented in this study shows that there are distinct interspecies differences in the properties of NADP-ME and AAT. These observations are discussed in terms of a regulatory mechanism that controls the relative flux through malate and aspartate in a dual-decarboxylation system of Zea mays, but not in Sorghum bicolor. Together, these findings add to our understanding of the regulation of C₄ photosynthesis, thereby advancing the wider objective of engineering the C₄ trait into existing C₃ crops.
Acknowledgements

I would like to thank Prof Richard Leegood for giving me the opportunity to conduct this research at The University of Sheffield. I am appreciative of his feedback and support throughout the PhD degree and the writing of this thesis. I am particularly grateful of his patience and constructive criticism, particularly when developing new ideas. I am also appreciative of Dr Karen Bailey, who trained me in immunoblotting. In addition, I would like to thank Prof Carl Smythe, who kept me motivated and invited me to present my work to his research group. I am also grateful of Dr Chatchawal Phansopa for his help in the laboratory, guidance when setting up experiments and reminding me that the PhD degree was not all about the experiments but also about the experience, the people met and the ideas that come from it. I want to thank Dr Richard Beniston and Dr Adelina E. Acosta Martin for their support and training in mass spectrometry. I also appreciate the L3 and L4 project students that helped at various stages of this project. Overall, this research was possible with the continuous support from these individuals, funding by the European Commission (FP7) and discussion meetings held with the 3to4 consortium members.
Declaration

I confirm that this work is all my own, except where clearly indicated. I have referenced all my sources properly.
# Abbreviations

List of biological abbreviations and nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>Ammonium bicarbonate</td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulfate solution</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>BPB</td>
<td>Bromophenol blue</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CHAPS</td>
<td>3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate</td>
</tr>
<tr>
<td>CID</td>
<td>Collision induced dissociation</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>IAA</td>
<td>Iodoacetamide</td>
</tr>
<tr>
<td>IEF</td>
<td>Isoelectric focusing</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IPG</td>
<td>Immobilized pH gradient</td>
</tr>
<tr>
<td>MOAC</td>
<td>Metal oxide affinity chromatography</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>NADH</td>
<td>Nicotinamide adenine dinucleotide (reduced form)</td>
</tr>
<tr>
<td>NADPH</td>
<td>Nicotinamide adenine dinucleotide phosphate (reduced form)</td>
</tr>
<tr>
<td>NCBInr</td>
<td>National Center For Biotechnology Information non-redundant</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PPFD</td>
<td>Photosynthetic photon flux density</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts per million</td>
</tr>
<tr>
<td>rpkm</td>
<td>Reads per kilobase million</td>
</tr>
<tr>
<td>PSM</td>
<td>Peptide spectrum match</td>
</tr>
</tbody>
</table>
SDS     Sodium dodecyl sulfate
TEMED   Tetramethylethylenediamine
TFA     Trifluoroacetic acid
TGS     Tris-glycine-SDS
TiO₂    Titanium dioxide
Tris-Base Tris(hydroxymethyl) aminomethane
UniProtKB Universal Protein Knowledgebase
UV      Ultra-violet

**Enzymes**

AAT      Aspartate aminotransferase
AlaAT    Alanine aminotransferase
GDC      Glycine decarboxylase complex
LDH      Lactate dehydrogenase
MDH      Malate dehydrogenase
ME       Malic enzyme
PEPC     Phosphoenolpyruvate carboxylase
PEPCK    Phosphoenolpyruvate carboxykinase (ATP)
PGM      Phosphoglycerate mutase
Rubisco  Ribulose-1,5-bisphosphate carboxylase/oxygenase

**Metabolites**

2-OG     2-oxoglutarate
Ala      Alanine
Asp      Aspartate
Glu      Glutamate
Mal      Malate
OAA      Oxaloacetate
PGA      Phosphoglycerate
Pyr      Pyruvate
RuBP     Ribulose-1,5-bisphosphate
# List of Figures

## Chapter 1 – Introduction.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Overview of photosynthesis in higher plants.</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Photosynthetic carbon reduction cycle in all photosynthetic organisms.</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Photorespiratory pathway recycles 2-phosphoglycollate between chloroplasts, peroxisomes and mitochondria, eventually releasing CO₂.</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>Theoretical energy balance of a C₃ leaf in high atmospheric CO₂ and increase in air temperature.</td>
<td>10</td>
</tr>
<tr>
<td>1.5</td>
<td>Overview of the differences in the mechanism for carbon fixation in C₃ and C₄ plants.</td>
<td>11</td>
</tr>
<tr>
<td>1.6</td>
<td>Transduction of PAR in C₃ and C₄ plants in atmospheric CO₂.</td>
<td>12</td>
</tr>
<tr>
<td>1.7</td>
<td>Comparison between C₃ and C₄ anatomy in higher plants.</td>
<td>16</td>
</tr>
<tr>
<td>1.8</td>
<td>The phylogenetic distribution of 47 angiosperm clades with C₄ photosynthesis.</td>
<td>17</td>
</tr>
<tr>
<td>1.9</td>
<td>The path to C₄ photosynthesis.</td>
<td>18</td>
</tr>
<tr>
<td>1.10</td>
<td>Arrangement of mesophyll and bundle sheath cells in Kranz anatomy of C₄ plants.</td>
<td>19</td>
</tr>
<tr>
<td>1.11</td>
<td>Schematic showing interveinal distance comparison between C₃ and C₄ leaf cross-sections.</td>
<td>20</td>
</tr>
<tr>
<td>1.12</td>
<td>GDC compartmentalisation and the photorespiratory CO₂ pump in C₃–C₄ intermediate species.</td>
<td>22</td>
</tr>
<tr>
<td>1.13</td>
<td>C₄ leaf cross-section of <em>Zea mays</em> showing compartmentation of Rubisco and PEPC.</td>
<td>24</td>
</tr>
<tr>
<td>1.14</td>
<td>Three biochemical variations of the C₄ pathway using the NADP-malic enzyme, NAD-malic enzyme or PEPCK decarboxylase.</td>
<td>31</td>
</tr>
<tr>
<td>1.15</td>
<td>3-PGA/triose-phosphate shuttle operating between bundle sheath and mesophyll cells.</td>
<td>32</td>
</tr>
<tr>
<td>1.16</td>
<td>Modelled relationship between the rate of photosynthesis and intracellular concentration of CO₂ in leaves of C₃ and C₄ plants.</td>
<td>34</td>
</tr>
<tr>
<td>1.17</td>
<td>Overview of the global distribution of higher plants dominated by either C₃ or C₄ grasses.</td>
<td>36</td>
</tr>
</tbody>
</table>

## Chapter 2 – Methods.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>General workflow for the extraction of leaf proteins, sample processing and analysis by mass spectrometry.</td>
<td>40</td>
</tr>
<tr>
<td>2.2</td>
<td>2.5-week old <em>Setaria viridis</em> photographed before the harvesting of leaves.</td>
<td>41</td>
</tr>
<tr>
<td>2.3</td>
<td>Classical method for the extraction of proteins from plant leaves using liquid nitrogen.</td>
<td>42</td>
</tr>
<tr>
<td>2.4</td>
<td>General workflow for sample processing using an in-gel tryptic digestion approach, prior to analysis by mass spectrometry.</td>
<td>43</td>
</tr>
<tr>
<td>2.5</td>
<td>Workflow schematic illustrating the basic flow of analyte through TiO₂ phosphopeptide enrichment columns.</td>
<td>45</td>
</tr>
<tr>
<td>2.6</td>
<td>Proteome Discoverer workflow used for identifying novel phosphorylation sites.</td>
<td>47</td>
</tr>
<tr>
<td>2.7</td>
<td>Stained protein gel containing four replicates of <em>Setaria viridis</em> and <em>Sorghum bicolor</em> protein extracts from 7.5 h darkened leaves.</td>
<td>50</td>
</tr>
<tr>
<td>2.8</td>
<td><em>Setaria viridis</em> sample preparation workflow for protein isoform analysis.</td>
<td>52</td>
</tr>
</tbody>
</table>
2.9 Schematic showing the workflow for protein assays using *Setaria viridis*, *Sorghum bicolor* and *Zea mays* leaf proteins.

2.10 2.5-week old *Setaria viridis* and *Sorghum bicolor* leaves photographed before harvests.

2.11 The BSA standard curve that was used to determine protein concentration in leaf homogenates.

2.12 Example of a 96-well microtitre plate assay design used for determining Michaelis-Menten kinetics.

2.13 NADH and NADPH standard curves for determining enzyme reaction rates.

2.14 Reaction mechanism showing the carboxylation of phosphoenolpyruvate to oxaloacetate catalysed by PEPC in C4 photosynthesis.

2.15 Reaction mechanism showing the decarboxylation of L-malate by NADP-malic enzyme in bundle sheath chloroplasts of C4 plants.

2.16 Reaction mechanism for the aspartate and alanine aminotransferase assays.

## Chapter 3 – Identification of novel light-dependent phosphorylation sites of C4-related enzymes.

3.1 Schematic showing the role of PTMs on the functional complexity of the proteome.

3.2 *In vivo* reversible phosphorylation mechanism of serine, threonine and tyrosine residues.

3.3 Irreversible β-carboxylation of PEP to oxaloacetate by PEPC.

3.4 Regulation of C4-specific PEPC activity by *in vivo* serine phosphorylation.

3.5 Regeneration of PEP from pyruvate by PPDK in C4 plants.

3.6 Light activation mechanism of the C4-specific PPDK by reversible phosphorylation.

3.7 ATP-dependent decarboxylation of oxaloacetate by PEPCK in C4 plants.

3.8 *Setaria viridis* whole leaf proteins resolved by 1D SDS-PAGE.

3.9 Fluorescence stained gel showing phosphorylated proteins in *Setaria viridis*.

3.10 7 cm stained gels showing *Setaria viridis* whole leaf proteins resolved by 2D SDS-PAGE.

3.11 Protein schematic showing the similarities between PEPC P04711, K3XV32 and K3YPN6.

3.12A Phosphorylation at Ser-11 occurred after illumination.

3.12B Phosphorylation at Ser-752 occurred in darkened leaves.

3.13A Phosphorylation at Ser-180 occurred in darkened and illuminated leaves of *Setaria viridis*.

3.13B Phosphorylation at Ser-702 occurred in darkened and illuminated leaves of *Setaria viridis*.

3.14 Phosphorylation of PPDK in leaves of *Setaria viridis* compared to *Zea mays* PPDK.

3.15 PPDK phosphorylation at Thr-462 in *Setaria viridis*.

3.16 Phosphorylation at Thr-593 occurred strictly in darkened leaves of *Setaria viridis*.

3.17 NADP-malic enzyme isoforms identified in *Setaria viridis* leaf extracts compared to NADP-malic enzyme from *Zea mays*.

3.18 Peptide spectrum showing novel NADP-malic enzyme phosphorylation at Ser-429 in *Setaria viridis*.

3.19 Peptide spectrum showing alanine aminotransferase phosphorylations at Thr-167 and Ser-189.
The phosphorylation of alanine aminotransferase in *Setaria viridis* occurred predominantly after illumination.

Peptide spectra showing the phosphorylation at Ser-109 on phosphoglycerate mutase in *Setaria viridis*.

Peptide spectra showing the phosphorylations at Ser-520, Thr-527 and Thr-530 on phosphoglycerate mutase in *Setaria viridis*.

Schematic showing the alignment of enolase K3XWW9 and K3Z681 from *Setaria viridis*.

PEPC MS/MS spectra after TiO$_2$ phosphopeptide enrichment and no phosphopeptide enrichment.

Model relationship between light-intensity and rates of photosynthesis in C$_4$ plants.

Phosphorylation of PEPC from *Setaria viridis*.

Alignment of PEPCs from C$_4$ and C$_3$ plants.

*Zea mays* PEPC annotated with phosphorylation identified in *Setaria viridis*.

Schematic showing the active site of PEPC in higher plants.

Alignment of PPDKs from C$_4$ and C$_3$ plants.

Structure of PPDK from *Flaveria trinervia* reveals the position of phosphorylation sites identified in *Setaria viridis*.

Alignment of NAD- and NADP-malic enzymes across plant (C$_4$ and C$_3$), mammalian and avian origins.

Position of phosphorylation sites identified in *Setaria viridis*.

Alignment of aspartate aminotransferases from C$_4$ and C$_3$ plants.

Alignment of alanine aminotransferases from C$_4$ and C$_3$ plants.

Alignment of enolases (ENL) protein sequences.

Alignment of phosphoglycerate mutase protein sequences.

**Chapter 4 – The kinetics of C$_4$-acid decarboxylation by NADP-malic enzyme in response to pH and illumination in C$_4$ grasses.**

Oxidative decarboxylation of L-malate via NADP-malic enzyme.

The activity of PEPC in *Setaria viridis* corresponds to shifts in $K_M$ during the photoperiod.

NADP-malic enzyme $K_M$ for L-malate was highest during the dark to light transition in *Setaria viridis*.

The pH optimum of NADP-malic enzyme activity in *Sorghum bicolor* shifted after illumination.

NADP-malic enzyme from illuminated *Setaria viridis* leaves remained active across a broader pH range.

There was no change in the pH optimum of NADP-malic enzyme activity in *Zea mays* leaves.

Dark and light NADP-malic enzyme activity in leaves of *Sorghum bicolor, Setaria viridis* and *Zea mays* at the optimum pH.

No difference in NADP-malic enzyme activity from darkened and illuminated leaves of *Setaria viridis* in the presence of Mg$^{2+}$.

Activity of NADP-malic enzyme in response to Mg$^{2+}$ may be pH- and light-dependent in *Sorghum bicolor*.
4.10 The differences in light and dark activities of Sorghum bicolor NADP-malic enzyme in response to pH linked to shifts in $K_M$.

4.11 Link between pH and light/dark activity of NADP-malic enzyme in Setaria viridis.

4.12 NADP-malic enzyme in Zea mays maintained high substrate affinity in response to pH.

4.13 Schematic showing two distinct activities of NADP-malic enzyme from darkened and illuminated leaves of Sorghum bicolor.

4.14 Schematic showing NADP-malic enzyme activity from darkened and illuminated leaves of Setaria viridis.

4.15 Schematic showing NADP-malic enzyme activity from darkened and illuminated leaves of Zea mays.

4.16 Relative expression of NADP-malic enzyme and PEPCK in C₄ grasses.

Chapter 5 – The regulation of transamination and its involvement in controlling carbon trafficking in C₄ photosynthesis.

5.1 Kranz anatomy in C₄ plants.

5.2 C₄ pathway showing a theoretical dual-decarboxylation system using NADP-malic enzyme, PEPCK and NAD-malic enzyme.

5.3 Interconversion of aspartate and oxaloacetate by aspartate aminotransferase.

5.4 Interconversion of aspartate and oxaloacetate by aspartate aminotransferase.

5.5 Activity of aspartate aminotransferase in Sorghum bicolor, Setaria viridis and Zea mays in normalised leaf lysates.

5.6 Aspartate aminotransferase affinity for L-aspartate and 2-oxoglutarate declines in response to light in Setaria viridis.

5.7 Light-induced changes to aspartate aminotransferase affinity for L-Asp and 2-oxoglutarate in Sorghum bicolor.

5.8 Aspartate aminotransferase affinity for L-aspartate declines after the onset of light in Zea mays.

5.9 Dark and light Setaria viridis aspartate aminotransferase responds better to dimethyl-2-oxoglutarate.

5.10 3-PGA, l-alanine and l-malate affect aspartate aminotransferase activity differently in three C₄ species.

5.11 Aspartate aminotransferase activity in response to PEP and pyruvate.

5.12 Alanine aminotransferase reaction mechanism towards the formation of malate.

5.13 Alanine aminotransferase activity in illuminated Zea mays, Sorghum bicolor and Setaria viridis leaf lysates.

5.14 Alanine aminotransferase activity in Setaria viridis leaf lysates is higher than in Zea mays and Sorghum bicolor.

5.15 Interchange between aspartate and oxaloacetate by aspartate aminotransferase.

5.16 Relative expression of aspartate aminotransferase and alanine aminotransferase.

5.17 Simple schematic depicting a mechanism that could regulate the relative flux through malate or aspartate in a dual-decarboxylation system in Zea mays.

Chapter 6 – General Discussion.

6.1 C₄ plants grouped by their primary decarboxylase.

6.2 Interconversion of 3-PGA and PEP via phosphoglycerate mutase and enolase in a NADP-malic enzyme system.
# List of Tables

<table>
<thead>
<tr>
<th>Chapter 2 – Methods</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 MS/MS search engines used during the study.</td>
<td>48</td>
</tr>
<tr>
<td>2.2 Online protein sequence databases used for MS/MS analyses.</td>
<td>49</td>
</tr>
<tr>
<td>2.3 Programme for 7 cm IEF.</td>
<td>54</td>
</tr>
<tr>
<td>2.4 Programme for 24 cm IEF.</td>
<td>54</td>
</tr>
<tr>
<td>2.5 PEPC mastermix.</td>
<td>61</td>
</tr>
<tr>
<td>2.6 NADP-malic enzyme mastermix.</td>
<td>62</td>
</tr>
<tr>
<td>2.7 Aminotransferase mastermix.</td>
<td>63</td>
</tr>
<tr>
<td>2.8 Summary of enzyme assays.</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3 – Identification of novel light-dependent phosphorylation sites of C₄-related enzymes.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Reversible post-translational modifications that regulate the activity of plant proteins.</td>
<td>68</td>
</tr>
<tr>
<td>3.2 Summary of known post-translational modifications of plant proteins involved in carbon fixation.</td>
<td>79</td>
</tr>
<tr>
<td>3.3 Summary of proteins involved in C₄ photosynthesis identified by MS/MS.</td>
<td>86</td>
</tr>
<tr>
<td>3.4A Novel phosphorylation sites identified in 7.5 h darkened <em>Setaria viridis</em> leaves.</td>
<td>90</td>
</tr>
<tr>
<td>3.4B Novel phosphorylation sites identified in 15.5 h illuminated <em>Setaria viridis</em> leaves.</td>
<td>90</td>
</tr>
<tr>
<td>3.5 Protein isoforms identified in 7.5 h darkened <em>Setaria viridis</em> leaves.</td>
<td>98</td>
</tr>
<tr>
<td>3.6 Phosphopeptide validation for 7.5 h darkened <em>Setaria viridis</em> leaves.</td>
<td>115</td>
</tr>
<tr>
<td>3.7 Phosphopeptide validation for 7.5 h darkened <em>Sorghum bicolor</em> leaves.</td>
<td>115</td>
</tr>
<tr>
<td>3.8 Phosphorylation sites identified in 7.5 h darkened and 4 h illuminated leaves of <em>Megathyrsus maximus</em>.</td>
<td>116</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 4 – The kinetics of C₄-acid decarboxylation by NADP-malic enzyme in response to pH and illumination in C₄ grasses.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Changes to NADP-malic enzyme affinity for L-malate in response to illumination.</td>
<td>163</td>
</tr>
<tr>
<td>4.2 Comparing the $K_M$ for L-malate from previous studies.</td>
<td>175</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 5 – The regulation of transamination and its involvement in controlling carbon trafficking in C₄ photosynthesis.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Aspartate aminotransferase affinity for primary and secondary substrates after 15.5 h illumination.</td>
<td>187</td>
</tr>
<tr>
<td>5.2 Aspartate aminotransferase affinity for 2-oxoglutarate and dimethyl-2-oxoglutarate.</td>
<td>201</td>
</tr>
</tbody>
</table>
Table of Contents

Abstract ........................................................................................................................................ i
Acknowledgements .................................................................................................................. ii
Declaration .................................................................................................................................. iii
Abbreviations ........................................................................................................................... iv
List of Figures ............................................................................................................................. vi
List of Tables .............................................................................................................................. x
Chapter 1 ...................................................................................................................................... 1
  1.1 Food for thought .................................................................................................................. 1
    1.1.1 Agricultural sustainability, energy harvesting and photosynthesis ...................... 1
  1.2 The inefficiency of C₃ photosynthesis ............................................................................... 3
    1.2.1 The problem with Rubisco and photorespiration .............................................. 3
    1.2.2 C₃ photosynthesis will be limited by future climatic events ......................... 7
    1.2.3 Increasing crop productivity ............................................................................. 11
  1.3 The origins of C₄ photosynthesis ......................................................................................... 13
    1.3.1 Natural bypass of photorespiration ................................................................ 13
    1.3.2 The recurrent emergence of C₄ photosynthesis .............................................. 15
    1.3.3 Specialisation of mesophyll and bundle sheath cells .................................. 19
    1.3.4 Photorespiratory CO₂ pump in C₃–C₄ intermediates ................................ 21
    1.3.5 Establishing a fully functional C₄ cycle ......................................................... 23
    1.3.6 Integration of C₃ and C₄ cycles ..................................................................... 25
    1.3.7 Fine-tuning C₄-specific enzyme activity ....................................................... 28
  1.4 What does it mean to be C₄? .............................................................................................. 29
    1.4.1 Variations of the C₄ pathway ........................................................................... 29
    1.4.2 Increased capacity of metabolite movement ............................................... 33
    1.4.3 Physiology and performance of C₄ plants ..................................................... 34
  1.5 Aim of thesis and thesis structure ..................................................................................... 38
Chapter 2 – Methods .................................................................................................................. 40
  2.1 Proteomics and Mass Spectrometry .................................................................................. 40
    2.1.1 Growth of plant material .................................................................................. 40
    2.1.2 Extraction of leaf proteins for mass spectrometric analysis .......................... 42
    2.1.3 In-gel tryptic digestion ..................................................................................... 42
    2.1.4 C18 column clean-up ....................................................................................... 44
2.1.5 Phosphopeptide enrichment by titanium dioxide ........................................ 44
2.1.6 Mass spectrometry analysis ........................................................................ 46
2.1.7 General mass spectrometry protocol .......................................................... 46
2.1.8 Immunoblotting .......................................................................................... 46
2.1.9 Data acquisition and analyses ..................................................................... 48
2.1.10 Phosphopeptide validation ....................................................................... 50
2.2 Protein Isoform and Phosphoproteome Analysis ............................................. 52
  2.2.1 Sample preparation...................................................................................... 53
  2.2.2 Two-dimensional SDS-PAGE ................................................................. 53
  2.2.3 ProQ Diamond phosphoprotein staining ................................................... 55
  2.2.4 Coomassie staining ................................................................................... 55
2.3 Determining Enzyme Kinetics ........................................................................ 56
  2.3.1 Preparation of leaf homogenates for enzyme measurement ....................... 57
  2.3.2 Protein quantification .............................................................................. 58
  2.3.3 Measurement of enzymes ......................................................................... 59
  2.3.4 Phosphoenolpyruvate carboxylase assay .................................................. 61
  2.3.5 NADP-malic enzyme assay ...................................................................... 62
  2.3.6 Aspartate and alanine aminotransferase assay .......................................... 63
  2.3.7 Determining Michaelis-Menten kinetics .................................................... 65
Chapter 3 – Identification of novel light-dependent phosphorylation sites of C_4-related enzymes .......................................................................................................................... 66
  3.1 Introduction .................................................................................................... 66
    3.1.1 The role and diversity of post-translational modifications in plants .......... 67
    3.1.2 Phosphoenolpyruvate carboxylase .......................................................... 71
    3.1.3 Pyruvate, phosphate dikinase ................................................................... 74
    3.1.4 Phosphoenolpyruvate carboxykinase ...................................................... 77
    3.1.5 Identifying novel phosphorylation sites using LC-MS/MS ....................... 81
  3.2 Results – Novel phosphorylation sites of C_4-related enzymes ...................... 84
    3.2.1 Identifying key C_4 photosynthesis proteins in *Setaria viridis* .......... 84
    3.2.2 Phosphoenolpyruvate carboxylase phosphorylation .............................. 91
    3.2.3 Pyruvate, phosphate dikinase phosphorylation ....................................... 99
    3.2.4 NADP-dependent malic enzyme phosphorylation .................................. 103
    3.2.5 Alanine and aspartate aminotransferase phosphorylation ...................... 106
    3.2.6 Phosphoglycerate mutase and enolase phosphorylation ....................... 110
3.2.7 Phosphopeptide validation ................................................................. 114
3.3 Discussion .......................................................................................... 117
  3.3.1 Overcoming the challenges of phosphoproteomics ...................... 117
  3.3.2 Phosphorylation of phosphoenolpyruvate carboxylase .................. 122
  3.3.3 Phosphorylation of pyruvate, phosphate dikinase ......................... 127
  3.3.4 Phosphorylation of NADP-malic enzyme ....................................... 130
  3.3.5 Phosphorylation of aspartate and alanine aminotransferase .......... 134
  3.3.6 Phosphorylation of enolase and phosphoglycerate mutase .......... 137
  3.3.7 Quantitative proteomics and future study ...................................... 140
  3.3.8 Conclusion .................................................................................... 142

Chapter 4 – The kinetics of C₄-acid decarboxylation by NADP-malic enzyme in response to pH and illumination in C₄ grasses ................................................................. 143
  4.1 Introduction ....................................................................................... 143
    4.1.1 Regulation of NADP-malic enzyme in plants .............................. 143
    4.1.2 Regulation by pH and illumination .......................................... 145
    4.1.3 Identifying species specific NADP-malic enzyme properties .... 146
  4.2 Results ............................................................................................. 147
    4.2.1 Kinetics of phosphoenolpyruvate carboxylase and NADP-malic enzyme .................................................. 147
    4.2.2 NADP-malic enzyme is regulated by pH ................................... 149
    4.2.3 Optimal NADP-malic enzyme activity at low a Mg²⁺ concentration .... 153
    4.2.4 Michaelis-Menten kinetics of NADP-malic from three C₄ grasses ... 155
  4.3 Discussion ......................................................................................... 164
    4.3.1 Low pH optimum ensures NADP-malic enzyme activity at the onset of light ......................................................................................................................... 164
    4.3.2 PEPCK dependence may influence NADP-malic enzyme activity ... 171
    4.3.3 New insights on NADP-malic enzyme regulation in C₄ plants ...... 174
    4.3.4 Conclusion .................................................................................. 175

Chapter 5 – The regulation of transamination and its involvement in controlling carbon trafficking in C₄ photosynthesis ...................................................................................... 176
  5.1 Introduction ....................................................................................... 176
    5.1.1 The dual-decarboxylation system of C₄ photosynthesis .............. 176
    5.1.2 Regulation of aspartate and alanine aminotransferase ............... 179
  5.2 Results ............................................................................................. 181
5.2.1 The role and regulation of aspartate aminotransferase in C₄ photosynthesis .................................................................................................................. 181
5.2.2 The effect of metabolites on the activity of aspartate aminotransferase ..... 190
5.2.3 Specific activity of alanine aminotransferase ........................................... 195
5.3 Discussion ...................................................................................................... 198
  5.3.1 Aspartate aminotransferase affinity aspartate is sensitive to dark–light transitions .............................................................................................................. 198
  5.3.2 C₄ metabolites may regulate the activity of aspartate aminotransferase .... 202
  5.3.3 Alanine aminotransferase may be differently regulated in C₄ species ...... 207
  5.3.4 Conclusion .................................................................................................. 208
Chapter 6 – General Discussion ........................................................................ 209
  6.1 Regulating the relative flux through malate and aspartate ......................... 209
  6.2 Communication between C₃ and C₄ cycles and the formation of PEP .......... 216
  6.3 Conclusion .................................................................................................... 219
References .......................................................................................................... 221
Appendix A ......................................................................................................... 253
Appendix B ......................................................................................................... 258
Chapter 1

1.1 Food for thought

1.1.1 Agricultural sustainability, energy harvesting and photosynthesis

Agricultural sustainability refers to the agricultural capacity of producing enough food to sustain the population at any given time, while reducing environmental impacts (Hansen, 1996; Mitchell and Sheehy, 2006; Pretty, 2008; Godfray et al., 2010). The concept has received international attention in recent years due to a growing food crisis. As the global population rises, crop production must increase by at least 50% in the next 15 years to meet the minimum requirements to feed the population (Hibberd et al., 2008; Peterhänzel and Offermann, 2012). Today, crop production faces extreme weather patterns brought on by climate change, such as high temperatures and increased rainfall variability (Rosenzweig et al., 2001; McMichael et al., 2006; Gornall et al., 2010). It is evident that increased yields must be achieved using less land, which is declining due to urbanisation (Ramakrishnan, 2001; Zhu et al., 2010) and less water, due to an increase in water demand from cities (Pimentel et al., 1997; Postel, 2000; Pimentel et al., 2004).

The global population has nearly doubled in the last 40 years (Tilman, 1998). During the Green Revolution (1960–2000), crop productivity was increased using improved breeding strategies (rather than genetic manipulation), modernised farming techniques and application of more fertilisers and improved pesticides (Mitchell and Sheehy, 2006), which also resulted in reduced global food prices (Tilman et al., 2002; Evenson and Gollin, 2003). By the beginning of the 21st century, crop productivity rose 208% for wheat (Triticum aestivum), 109% for rice (Oryza sativa) and 157% for maize (Zea mays) (Pingali, 2012). However, at the rate that the current population is growing, projected yields are not enough to sustain developing countries. Furthermore, acquisition of farmland for bioenergy crop production (i.e. biofuels) will reduce the accessibility of land for agricultural use (Tokgoz et al., 2012; Popp et al., 2014; Vasile et al., 2016). The future of agriculture relies on the improvement of farmland management, enhanced use of technology as well as genetically enhanced crops (Tilman et al., 2002).
A consensus in the scientific community has called for a second Green Revolution, with hopes to create genetically enhanced crop plants. These should withstand the adverse effects of climate change, grow with fewer additives such as nitrogenous fertilisers and efficiently use photosynthetically active radiation (PAR) to assimilate CO\textsubscript{2} into carbohydrates during photosynthesis (Zhu et al., 2010; Covshoff and Hibberd, 2012; Leegood, 2013).

**Fig. 1.1. Overview of photosynthesis in higher plants.** In higher plants photosynthesis takes place inside chloroplasts. Captured energy is used to form ATP and NADPH. In the stroma of the chloroplast, ATP and NADPH are utilised by enzymes in the photosynthetic carbon reduction (PCR) cycle to fix CO\textsubscript{2} into carbohydrates (CH\textsubscript{2}O). Increasing the efficiency of photosynthesis in major food crops is the most feasible long-term solution that could mitigate the food crisis (Mitchell and Sheehy, 2006; Zhu et al., 2010; Covshoff and Hibberd, 2012). Photosynthesis has two phases: the light harvesting and the photosynthetic carbon reduction (PCR) cycle (Figure 1.1). During the day, solar energy is intercepted by chlorophyll-containing, light-harvesting complexes embedded in chloroplast thylakoid membranes, resulting in the splitting of water and release of O\textsubscript{2} (Candau et al., 1976; Kühlbrandt and Wang, 1991; McEvoy et al., 2005). Harvested energy is then used for the formation of ATP and reducing equivalents, which are used in the PCR cycle for the formation of carbohydrates.
1.2 The inefficiency of C₃ photosynthesis

1.2.1 The problem with Rubisco and photorespiration

All photosynthetic species use ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) for carbon fixation. In C₃ plants, carbon fixation is limited by the inefficiency of Rubisco, which has a low turnover rate and competing specificity for CO₂ and O₂ binding (Portis and Parry, 2007; Raines, 2011). During photosynthesis, Rubisco fixes atmospheric CO₂ through the carboxylation of ribulose-1,5-bisphosphate (RuBP) forming two molecules of 3-phosphoglycerate (3-PGA) (Sage, 1999). Rubisco also catalyses the oxygenation of RuBP, in an unfavourable and unavoidable side reaction with O₂, forming 3-PGA and 2-phosphoglycolate, the latter which cannot be utilised in the PCR cycle (Figure 1.2) and may be toxic in high concentrations (Ogren, 1984; Andrews and Lorimer, 1987; Maurino and Peterhänsel, 2010). To prevent the accumulation of 2-phosphoglycolate, the compound is removed from the chloroplast and metabolised into organic compounds, using energy and 25% of assimilated carbon (Björkman et al., 1969; Ehleringer and Björkman, 1977; Bauwe et al., 2010), in the process of photorespiration (Figure 1.3). Most importantly, the carbon present on 2-phosphoglycolate must be recycled back to the PCR cycle as efficiently as possible.

Although Rubisco has a higher affinity for CO₂ (K_M 9.7 µM), than O₂ (K_M 244 µM) (Cousins et al., 2010), the CO₂ concentration inside the chloroplast is about 1000-fold lower than that of O₂ (Ehleringer et al., 1991). Furthermore, in warm climates, the oxygenase activity of Rubisco increases relative to the carboxylase activity, both because the relative solubility of dissolved O₂ is greater than that of dissolved CO₂, thus lowering the ratio of dissolved CO₂ to O₂ in the chloroplast and because the kinetic properties of Rubisco alter in response to increasing temperature, thereby increasing its affinity for O₂ (Ku and Edwards, 1977; Chen and Spreitzer, 1992; Tcherkez et al., 2006; Leegood, 2013; Buchanan et al., 2015). In addition, the diffusion of CO₂ to the site of Rubisco is reduced when stomata close to prevent the loss of water through transpiration in high temperatures (Sage, 2004; Schulze et al., 2013). The loss of CO₂ near the site of Rubisco can reduce photosynthetic efficiency from 20–40% and increases in unfavourable
conditions such as warm temperatures (above 30 °C) and arid climates (Jordan and Ogren, 1984; Sharkey, 1988; Ehleringer et al., 1991; Gowik and Westhoff, 2011).

**Fig. 1.2. Photosynthetic carbon reduction cycle in all photosynthetic organisms.** Enzymes: 1, Rubisco; 2, phosphoglycerate kinase; 3, glyceraldehyde 3-phosphate (G3P) dehydrogenase, 4, triose phosphate isomerase; 5, phosphoribulokinase; 6, fructose bisphosphatase; 7, transketolase; 8, fructose bisphosphate aldolase; 9, sedoheptulose-1,7-bisphosphatase. Metabolites: RuBP, ribulose 1, 5-P; Ru5P, ribulose 5-P; R5P, ribose 5-P; 1,3-BPG, glycerate 1,3 bisphosphate; *3PGA, glycerate 3-P; *G3P, glyceraldehyde 3-P; *DHAP, dihydroxyacetone-P (triose-P); Xyl, Xylulose; Ery, erythrose; Sed, sedoheptulose, *Diffuses from stroma to cytosol.

Photorespiration is a complex detoxification and recycling process that takes place in the chloroplasts, peroxisomes, mitochondria and cytosol of higher plants (Figure 1.3) (Bauwe et al., 2010). Although it poses a problem for plants that solely rely on Rubisco for carbon fixation, photorespiration was not an issue when Rubisco first evolved around 3.5 billion years ago, due to an O2-free atmosphere (Leegood, 2013), but as photosynthetic species populated the globe, O2 became more readily available (Blankenship, 2010). This, however, was not a strong selective pressure to promote the
acquisition of a new carboxylase for carbon assimilation (Gowik and Westhoff, 2011). In fact, the bispecific nature of Rubisco evolved first and was subsequently maintained in all photosynthetic organisms, albeit preconditioned to operate in a high CO₂ atmosphere (Rachmilevitch et al., 2004; Andersson, 2008). Consequently, it is likely that photorespiration co-evolved with oxygenic photosynthesis in cyanobacteria to recycle carbon and to prevent the accumulation of glycollate from the unavoidable oxygenase activity of Rubisco, protect against photoinhibition by supplying reducing equivalents to other metabolic pathways or produce essential amino acids, such as serine and glycine (Kozaki and Takeba, 1996; Rachmilevitch et al., 2004; Bauwe et al., 2010; Maurino and Peterhänsel, 2010; Peterhänsel et al., 2012; Hagemann and Bauwe, 2016).

The reversible oxidation of glycine via the glycine decarboxylase complex (GDC), formed of glycine decarboxylase and serine hydroxymethyltransferase (SHMT) in the mitochondria (Figure 1.3, reaction 5), is an essential step in photorespiration (Walker and Oliver, 1986; Bauwe et al., 2010). Firstly, glycine decarboxylation generates ammonia, which is necessarily and efficiently refixed through the interconversion of glutamate and glutamine during nitrate assimilation (Linka and Weber, 2005). Secondly, 3-PGA can be regenerated from serine via pyruvate and glycerate, but the phosphorylation of glycerate to form 3-PGA is ATP-dependent (Figure 1.3, reaction 8) (Weber and von Caemmerer, 2010). Finally, despite that CO₂ is eventually produced through the carboxylation of glycine in the mitochondria, the uptake efficiency of CO₂ back into the chloroplasts and its subsequent use in the carboxylation reaction is not significant to recover from the losses of photorespiration (Schulze et al., 2013). Although loss of CO₂ through the photorespiratory pathway can have detrimental effects on C₃ plants by reducing nitrogen-use and water-use efficiency and limiting the regeneration of phosphate and RuBP, especially in warm and dry climates, photorespiration is important in all photosynthetic organisms (Rachmilevitch et al., 2004; Sage and Kubien, 2007; Bloom et al., 2010; Maurino and Peterhänsel, 2010; Bauwe et al., 2012). In cyanobacteria, glycollate is metabolised using a plant-like mechanism (Bauwe et al., 2010) or using a bacterial glycerate pathway, whereby glycollate is converted into glyoxylate, generating CO₂, but not ammonia, which does not need to be refixed through the ATP-dependent conversion of glutamate to glutamine (Figure 1.3, reaction 9) (Eisenhut et al., 2008; Hagemann and Bauwe, 2016).
Fig. 1.3. Photorespiratory pathway recycles 2-phosphoglycollate between chloroplasts (green), peroxisomes (blue) and mitochondria (orange), eventually releasing CO₂. Enzymes: 1, Rubisco; 2, phosphoglycollate phosphatase; 3, glycollate oxidase; 4, glutamate-glyoxylate aminotransferase; 5, glycine decarboxylase and serine hydroxymethyltransferase; 6, serine-glyoxylate aminotransferase; 7, pyruvate reductase; 8, glyceraldehyde kinase; 9, glutamine synthetase; 10, ferredoxin-dependent glutamate synthase. Transporters: A, glycerate-glycollate translocator; B, malate-glutamate/2-oxoglutarate translocator; C, amino acid translocator. Metabolites: RuBP, ribulose-1,5-bisphosphate; Ru5P, ribulose 5-phosphate; 3PGA, 3-phosphoglycerate; G3P, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate. Pᵢ, phosphate. Adapted from Bauwe et al. (2010).
Cyanobacteria mutants with the partial suppression of glycollate metabolism were found to have growth defects, while mutants completely lacking glycollate metabolism exhibited a lethal phenotype, suggesting that the removal of glycollate via photorespiration is essential (Eisenhut et al., 2008). Moreover, *Zea mays* (C₄) mutants, grown in ambient CO₂, lacking glycollate oxidase activity did not exhibit a viable phenotype, despite that the rate of photorespiration is substantially lower than in C₃ plants (Zelitch, 1968; Zelitch et al., 2009). Photorespiration is interwoven with other plant metabolic pathways of both C₃ and C₄ plants, and regardless of the amount of glycollate produced, and any disruption can have serious effects on the phenotype of photosynthetic species. The photorespiratory pathway may also play an intrinsic role in regulating the spatial availability of CO₂ and ammonia and mitigate the inhibition of photosynthesis by an excess of O₂ or phosphoglycollate (Chollet and Ogren, 1972; Ku and Edwards, 1977; Linka and Weber, 2005; Maurino and Peterhänsel, 2010; Weber and von Caemmerer, 2010).

1.2.2 C₃ photosynthesis will be limited by future climatic events

Since the evolution of oxygenic photosynthesis and increase in the atmospheric O₂, the rate at which Rubisco interacts with O₂ has increased, promoting photorespiration, which lowers photosynthetic efficiency and reduces carbon gain by at least 25% (Björkman et al., 1969; Ehleringer et al., 1991). Despite the competing carboxylase and oxygenase reactions, the reaction mechanism of Rubisco, with regards to structural differences and stabilisation during transition states, is optimised to differentiate between CO₂ and O₂ binding and maximise the rate of turnover (Lorimer et al., 1993; Tcherkez et al., 2006). However, Galmés et al. (2014) argues that the kinetic properties of Rubisco could still be improved such that its activity is optimised under future CO₂ concentrations. The efficiency of C₃ photosynthesis is not solely limited by the capacities of Rubisco. Transgenic plants with suppressed expression of Rubisco exhibited lower carbon fixation efficiencies when exposed to high light and temperature (Stitt and Schulze, 1994), but not if grown in high CO₂ (Masle et al., 1993). These studies also show that factors such as increasing temperature and CO₂, which are within the context of climate change, may influence how C₃ plants will respond in future climates.
At the current atmospheric concentration of CO\textsubscript{2} and O\textsubscript{2}, both carboxylation and oxygenation reactions occur at high rates (Leegood, 2013). In the next 50 years, the concentration of CO\textsubscript{2} is projected to double (from 360 to 600–1000 µmol mol\textsuperscript{−1}), which may increase the carboxylase activity of Rubisco by at least 70% (Rachmilevitch \textit{et al.}, 2004; Buchanan \textit{et al.}, 2015). However, increasing CO\textsubscript{2} availability in the atmosphere does not imply that the efficiency of photosynthesis will increase in C\textsubscript{3} plants. Several studies have shown that plants grown in elevated CO\textsubscript{2} suppress photorespiration and photosynthetic capacity increases by approximately 30% in the first instance (Björkman \textit{et al.}, 1969; Woodward, 2002). However, photosynthesis is eventually limited because of the capacity of the PCR cycle to regenerate enough RuBP to maintain high carboxylase activity (Griffin and Seemann, 1996; Sage and Kubien, 2007). Moreover, plants acclimate to elevated CO\textsubscript{2} after prolonged exposure, due to the accumulation of carbohydrates, which desensitise CO\textsubscript{2} utilisation by Rubisco or other enzymes of the PCR cycle (Moore \textit{et al.}, 1999; Rachmilevitch \textit{et al.}, 2004; Buchanan \textit{et al.}, 2015). In addition, prolonged exposure to high CO\textsubscript{2} can result in reduction of nitrogen-use efficiency, which could impair biomass production and photosynthesis (Smart \textit{et al.}, 1998; Bloom \textit{et al.}, 2002; Searles and Bloom, 2003).

Photosynthesis in plants can also be affected by temperature and it is one of many factors that can influence photosynthetic enzymes, plant development and leaf physiology (Berry and Bjorkman, 1980; Jordan and Ogren, 1984; Chen and Spreitzer, 1992; Leegood and Edwards, 1996; Bernacchi \textit{et al.}, 2002; Sage and Kubien, 2007). Moreover, the life-cycle of plants grown for grain production rely on seasonal changes in weather and temperature (Hatfield \textit{et al.}, 2011; Hatfield and Prueger, 2015). One of the suggestions to increase crop productivity is to extend the growing season, however, it is impractical because each crop species will have its optimum germination, maturation and harvest period, which is dependent on seasonal weather patterns (Covshoff and Hibberd, 2012). Certain crops, such as spinach (\textit{Spinacia oleracea}), prefer cooler, winter temperatures (Hatfield and Prueger, 2015). Extreme weather patterns or fluctuation in minimum and maximum daily temperatures were found to perturb developmental timings of annual crops, like tomato (\textit{Solanum lycopersicum}), and affected biomass production and seed viability (Ghosh \textit{et al.}, 2000; Sato \textit{et al.}, 2000; Pressman \textit{et al.}, 2002; Sato, 2006). Some fruits, such as apple (\textit{Malus domestica}) and grapes (\textit{Vitis vinifera}) are perennial plants, which require a certain number of hours of exposure to cooler temperatures. If
temperatures continue to increase, pollen production of perennial crops will be impaired, thus disrupting fruit production (Hatfield and Prueger, 2015). Fujii and Kennedy (1985) showed that the rate of photosynthesis in apple trees was dependent on seasonal changes, particularly during fruit development, while the rates of dark respiration and photorespiration were unaffected. An increase in temperature may lower the photosynthetic capacity of these plants during specific seasons, as the combined rate of photorespiration and dark respiration may surpass the rate of net carbon assimilation. Some C₃ vegetative crops like the soybean plant (Glycine max) are sensitive to illumination and leaf temperature (Hatfield and Prueger, 2015). Studies have shown that leaf shading increases the rate of photosynthesis and promotes gas exchange by decreasing stomatal resistance, but the rate of photorespiration increases proportionally (Mondal et al., 1978; Peet and Kramer, 1980). These effects could be more severe at higher temperatures and atmospheric CO₂ concentrations, and production yields may significantly suffer (Hatfield and Prueger, 2015). It is predicted that an increase in global temperatures will cause lower crop yields somewhere between 2.5% and 10% (Hatfield et al., 2011; Lobell et al., 2011) and these scenarios will be more prevalent if CO₂ emissions remain high (Schlenker and Roberts, 2009).

Increasing air temperature can also affect the rate of transpiration and energy balance of a leaf (Figure 1.4). High temperatures cause the exponential increase of atmospheric water vapour demand, which increases the rate at which water is transpired by the leaf until the rate of transpiration becomes limited by the rate of water absorption into the plant from the soil (Hatfield and Prueger, 2015). Increasing temperature can also limit photosynthesis, even when other conditions are at their optimum, by causing the premature closure of stomata in the daytime, thus increasing leaf temperature by at least 1 °C and decreasing the rate of gas exchange between the outside and inside of the leaf (Kimball et al., 1993; Bernacchi et al., 2002; Hatfield et al., 2011; Schulze et al., 2013; Hatfield and Prueger, 2015). Furthermore, the rate of evapotranspiration will drastically increase in extreme high temperatures, thus reducing the amount of water that the plant needs to absorb from the soil in order to meet the atmospheric evaporative demands (Hatfield and Prueger, 2011). In addition, altering the hydraulic conductance in plants may induce changes to the mechanisms that regulate the diurnal opening and closing of stomata, therefore limiting the rate of photosynthesis (Mathur et al., 2014; Bellasio et al., 2017).
Fig. 1.4. Theoretical energy balance of a C₃ leaf in high atmospheric CO₂ and increase in air temperature. Black arrows indicate the direction of stress response in the leaf. Red arrows (pointing up) indicate an increase and blue arrows (pointing down) indicate a decrease. Dashed arrows show predicted events. Adapted from Hatfield and Prueger (2011).

The thermal optimum of C₃ plants is between 10–25 °C, depending on plant species (Hatfield and Prueger, 2011). In low atmospheric CO₂ and at higher temperatures, photosynthesis becomes limited as the activation state of Rubisco declines with increasing temperature (Crafts-Brandner and Salvucci, 2000; Sage et al., 2002). Also, C₃ photosynthesis is limited by temperature only when atmospheric CO₂ concentration is high, because of a decline in the capacity to regenerate phosphate (Sage and Kubien, 2007). The thermal optima of C₃ plants can alter in response to increasing CO₂ concentration, thus increasing growth of certain plants in higher temperatures (Mawson and Cummins, 1989; Long, 1991; Ziska and Bunce, 1997; Atkin et al., 2006). However, even if elevated CO₂ may increase the growth in certain C₃ crops, it may not be a substantial amount to increase the grain production (Amthor, 2001). C₃ crop plant interactions with CO₂, temperature and water availability can have serious implications for crop production. It is also evident that increasing CO₂ concentration does not improve C₃ photosynthesis efficiency, and even though photorespiration may be suppressed by lowering the ratio of oxygenase to carboxylase activity. In summary, the inefficiency of C₃ photosynthesis poses a serious threat to C₃ crops grown for grain production if global temperatures and atmospheric CO₂ concentrations continue to increase. These effects will be more severe in drought-stricken regions.
1.2.3 Increasing crop productivity

Approximately 95% of land plants, including staple crops such as rice and wheat, use C₃ photosynthesis for carbon fixation (Sage, 2004). In these plants, CO₂ enters mesophyll cells and is directly fixed by Rubisco. However, maize is one of 4,500 grass species that use the alternative, and more efficient C₄ pathway (Sage, 1999). In the C₄-dicarboxylate cycle of C₄ photosynthesis, CO₂ is initially fixed by phosphoenolpyruvate carboxylase (PEPC) in mesophyll cells, forming a C₄-acid that diffuses to and is decarboxylated in bundle sheath cells, concentrating CO₂ near Rubisco (Figure 1.5). C₄ crops are anatomically and biochemically superior to C₃ crops, and are capable of maintaining high rates of photosynthesis even in hot and dry climates (Berry and Bjorkman, 1980; Pearcy and Ehleringer, 1984; Sage, 1999; Lara and Andreo, 2011).

![Diagram of C₃ and C₄ photosynthesis mechanisms](image)

**Fig. 1.5. Overview of the differences in the mechanism for carbon fixation in C₃ and C₄ plants.** In C₃ plants (left), CO₂ is fixed by Rubisco in mesophyll cells. C₄ plants (right) operate a carbon concentrating mechanism between mesophyll and bundle sheath cells. PEPC, phosphoenolpyruvate carboxylase.

With a global food crisis potentially developing by the end of this century, a great deal of attention has been given to the improvement of crop plants by converting to C₄ photosynthesis. Efforts to increase the photosynthetic potential of C₃ plants began after the realisation that plants operating the carbon concentrating mechanism of C₄ photosynthesis had a higher radiation use efficiency (RUE) than C₃ plants (Long *et al.*, 2006; Reynolds *et al.*, 2010; Zhu *et al.*, 2010; Covshoff and Hibberd, 2012).
Fig. 1.6. Transduction of PAR in C$_3$ and C$_4$ plants in atmospheric CO$_2$ (400 ppm). Arrows show the cellular process where energy is being spent. The size of the arrow shows the relative difference in the amount of energy lost. Shading from orange to yellow shows the direction of energy transduction from harvested sunlight to plant biomass (given as a percentage of the theoretical maximal energy conversion efficiency). The higher the percentage, the more energy the plant has invested in biomass production. Figure adapted from Zhu et al. (2010).

Theoretical models predict that C$_4$-augmented crop plants could potentially perform 50% more efficiently than non-C$_4$ plants, which translates to increased crop production efficiency (Covshoff and Hibberd, 2012). Efficient crop production is defined as the ratio of energy output for carbohydrate biosynthesis to solar energy input (Monteith and Moss, 1977). Approximately 70% of solar energy is lost to the environment before it is intercepted by plants; the remainder is used for carbohydrate biosynthesis and respiration, or dissipated as heat (Zhu et al., 2010). These models have also shown that C$_4$ plants, which have suppressed rates of photorespiration, can invest more energy into biomass production than C$_3$ plants (Figure 1.6). Therefore, converting C$_3$ crops to C$_4$ can potentially reduce the total amount of energy lost through photorespiration and increase photosynthetic efficiency by 30% (Zhu et al., 2010). However, until this can be implemented, the factors that promote photosynthetic efficiency in C$_4$ plants must be understood. This Chapter will review past and ongoing C$_4$ photosynthesis research, outline important evolutionary adaptations and identify the key traits and characteristics of the C$_4$ phenotype.
1.3 The origins of C₄ photosynthesis

1.3.1 Natural bypass of photorespiration

Since Rubisco is the only enzyme involved in CO₂ fixation in the vast majority of photosynthetic organisms, major chemical alternations might not have been selected for during evolution (Sage, 2004; Gowik and Westhoff, 2011). The interaction with O₂ is an inevitable and unavoidable consequence in the reaction mechanism that occurs when the 2,3-enediol intermediate form of RuBP reacts with O₂ rather than CO₂, because of a higher free energy activation with O₂ (Chen and Spreitzer, 1992; Tcherkez et al., 2006). Despite that different forms of Rubisco have evolved to respond better in low CO₂ to O₂ environments and increase the relative specificity for CO₂, some plants have found an alternative strategy for suppressing the oxygenase reaction by compartmentalising Rubisco with physical barriers, thus creating a subcellular environment where the ratio of CO₂ to O₂ is higher (Badger and Andrews, 1987; Sage, 1999). Additionally, some C₃ plants can restrict GDC activity to the bundle sheath, by the selective loss of GDC in the mesophyll, allowing photorespired CO₂ to be recycled more efficiently (Monson, 1999). Anatomical variation and spatial availability of CO₂ in C₃ plants can also increase photosynthetic capacity. In several photosynthetic species, including cyanobacteria, aquatic and land plants, this can be achieved using compartments that concentrate CO₂ at the site of Rubisco (Leegood, 2013). In microalgae, such as *Chlamydomonas reinhardtii*, CO₂ is concentrated in specialised compartments (pyrenoids) that maintain the intracellular CO₂ concentration high to reach maximal carboxylase activity (Badger et al., 1998; Mackinder et al., 2017). This is similar to the strategy adopted by cyanobacteria, which operate a carbon concentrating mechanism in carboxysomes (Badger et al., 2002). Changes to chloroplast positions can also affect the diffusion of CO₂ through the mesophyll of terrestrial C₃ plants (Tholen et al., 2008). Recently, Busch et al. (2013) showed that rice and wheat can restrict the efflux of photorespired CO₂ in photosynthesising cells using air spaces as physical barriers for diffusion. Although Busch et al. (2013) argue that this adaptation could have contributed to reducing the wastage of CO₂ produced during photorespiration, the delivery of atmospheric CO₂ into the site of Rubisco can be restricted by such structural barriers, which cause the CO₂ partial pressure in the stroma to be several-fold lower than that of the atmosphere (Evans
et al., 2009; Tosens et al., 2012). In rice, the entry of CO₂ may already be maximised as chloroplasts cover about 95% of the periphery of mesophyll cells, which increases the diffusion of CO₂ during high rates of photosynthesis (von Caemmerer and Evans, 1991; Tholen et al., 2008; Evans et al., 2009; Sage and Sage, 2009; Busch et al., 2013). Rice has also considerably high variation in the number of chloroplasts in bundle sheath cells, some of which are deficient of chloroplasts (Leegood, 2008; Sheehy et al., 2008). Theoretical models have also demonstrated that the diffusion efficiency and assimilation of CO₂ in tomato leaves correlates to the organisation and structure of mesophyll cells (Berghuijs et al., 2015). However, anatomical variations, such as the difference in photosynthetically competent mesophyll and bundle sheath cells in C₃ plants, make it difficult to assess the factors that contribute to the transport and utilisation CO₂ within C₃ plant cells (Berghuijs et al., 2016).

Before the elucidation of the C₄ photosynthetic pathway, Heinricher (1884) and Haberlandt (1904) noted that certain plant species of tropical origin had photosynthetically active chloroplasts in bundle sheath cells. In addition, chloroplasts found in bundle sheath cells of Zea mays were larger than those typically found in mesophyll cells of C₃ plants (Kiesselbach, 1916). Haberlandt (1904) originally suggested that the green parenchyma sheath cells in plants with wreath-like arrangement of mesophyll and bundle sheath cells (Kranz anatomy) might have specialised function. The accumulation of starch in bundle sheath cells (Rhoades and Carvalho, 1944), together with evidence that showed normal grana stacking in mesophyll chloroplasts but not bundle sheath chloroplasts in Zea mays (Hodge et al., 1955), further substantiated the biochemical specialisation of the bundle sheath. In light of these findings, Kortschak et al. (1965) demonstrated that radioactively labelled carbon in sugarcane was initially incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon. These observations were consistent with those demonstrated by Karpilov (1960), which showed that radioactively labelled carbon was incorporated into malate and aspartate, but not 3-PGA, whereas 3-PGA in soybean plants accounted for 80% of the radioactively labelled carbon.
photosynthesis, with 3-PGA produced much later. Based on these observations, Hatch and Slack (1966) elucidated the C₄ dicarboxylic acid pathway of photosynthesis, termed C₄ photosynthesis.

1.3.2 The recurrent emergence of C₄ photosynthesis

The C₄ pathway is the most effective trait that evolved in certain plants to overcome photorespiration (Sage, 1999). It has evolved independently over 60 times in the last 35 million years, through gradual anatomical and biochemical adaptations that allow CO₂ to be concentrated at the site of Rubisco (Sage, 2004). This is usually achieved using concentric tubes of vascular tissue tightly packaged by specialised bundle sheath and mesophyll cells in Kranz anatomy (Figure 1.7D). In C₄ plants, CO₂ is initially fixed into a four-carbon organic compound, usually malate or aspartate (Hatch, 1987), in mesophyll cells by PEPC, transported to and decarboxylated in bundle sheath cells forming CO₂ and a three-carbon acid that is subsequently transported back to mesophyll cells completing the cycle (Figure 1.7C). Unlike C₃ leaves, Rubisco is localised in the bundle sheath, where CO₂ is concentrated (Kanai and Edwards, 1999). As a result, the ratio of CO₂ to O₂ in the bundle sheath increases, allowing the carboxylase reaction of Rubisco to operate at near CO₂ saturation (Hatch, 1987; Sage, 1999; von Caemmerer and Furbank, 1999). Although the dual-cell configuration of Kranz anatomy is an important characteristic for the assembly of C₄ photosynthesis and compartmentalisation of Rubisco, it is not a requirement for all C₄ organisms (Badger et al., 1998; Edwards et al., 2004; Leegood, 2013; Berry et al., 2016). For instance, in a few single-cell chenopods such as Bieneritia sinuspersici, C₄ photosynthesis operates by compartmentalising Rubisco within a specialised region in the cell, concentrating CO₂ in equal proportions as plants with fully established Kranz anatomy (Sage, 2004; Rosnow et al., 2014).
**Fig. 1.7. Comparison between C₃ (A, B) and C₄ (C, D) anatomy in higher plants.** Leave cross-sections show the arrangement of mesophyll and bundle sheath cells in C₃ rice (B) and C₄ Sorghum (D) plants. ATP is needed to regenerate PEP via PPDK (C). Transverse leaf sections acquired from RT Furbank and RF Sage.

C₄ photosynthesis evolved from the gradual specialisation of ancestral C₃ plant characteristics, which were predisposed to C₄ photosynthesis (Hibberd and Quick, 2002; Sage, 2004; Christin and Osborne, 2013). Furthermore, all C₄-related enzymes derived from C₃ isoforms, which are abundant in photosynthetic cells (Hibberd and Quick, 2002), albeit serve different cellular functions and are differently regulated (Sage, 2004). The acquisition of C₄-related traits has been extensively studied in several plant taxa (Figure 1.8), mainly to understand how the various biochemical variants of the C₄ pathway were assembled multiple times from separate origins. While it is generally accepted that the
The evolution of C₄ photosynthesis was triggered, or at least favoured, by climatic events (i.e. decrease in atmospheric CO₂), Monson (2003) argues that its recurrent emergence was dependent on the genetic variability of C₃ lineages, mainly through genomic preconditioning such as gene duplications and neofunctionalisation.

From there, the C₄ pathway was assembled as an extension of the classical C₃ pathway, through two main phases (Figure 1.9): anatomical and biochemical specialisations (Monson et al., 1986; Sage, 1999; Sage, 2004; Sage et al., 2011). Comparative analyses in Alloteropsis semialata populations show that the convergent evolution of C₄ photosynthesis involved the optimisation of intermediary C₄ characteristics from C₄-like common ancestors, while others obtained C₄ features independently (Christin et al., 2012; Dunning et al., 2017). Furthermore, Gowik and Westhoff (2011) and Williams et
al. (2013) suggest that every transition into a C₄-like condition conferred certain advantages, and in some cases, certain species would not have necessarily continued to evolve C₄ photosynthesis, while others may still be evolving the C₄ pathway (Ludwig, 2012). Genomic and phylogenetic analyses predict that there are 4500 species in the grass family (Poaceae) that perform C₄ photosynthesis (equivalent to 60% of C₄ species), followed by 1500 species of sedges (Cyperaceae) and 1200 species of dicots, producing about a fourth of global O₂ (Sage, 2004; Edwards et al., 2010; Sage et al., 2012).

Fig. 1.9. The path to C₄ photosynthesis. Summary of the evolution of C₄ photosynthesis (in the direction of arrows). Example of C₃–C₄ intermediate species are shown to the right of each transition point. Figure adapted from Monson (1999), Monson and Rawsthorne (2000) and Gowik and Westhoff (2011).
1.3.3 Specialisation of mesophyll and bundle sheath cells

Trapping CO$_2$ at the site of Rubisco was accomplished by the development of a wreath-like arrangement of bundle sheath cells, spaced closely apart from neighbouring mesophyll cells and vasculature tissue. As a result, the intracellular CO$_2$ concentration in bundle sheath cells is about 10-fold higher than in mesophyll cells (Figure 1.10) (von Caemmerer and Furbank, 1999). Vein spacing in C$_4$ plants also facilitates the rapid diffusion of metabolites during photosynthesis (Osmond, 1971; Hatch and Osmond, 1976; Sage, 2004). Furthermore, C$_4$ plants typically have two mesophyll cells between veins, whereas C$_3$ plants can have more than five mesophyll cells between each vein (Figure 1.11) (Leegood and Walker, 1999; Ogle, 2003; Langdale, 2011).

![Fig. 1.10. Arrangement of mesophyll and bundle sheath cells in Kranz anatomy of C$_4$ plants.](image)

Kranz anatomy in a C$_4$ plant, such as *Zea mays*, is formed by specialised mesophyll and bundle sheath cells, which partition between carboxylation and decarboxylation reactions, respectively. CO$_2$ produced from the decarboxylation of the C$_4$-acid is concentrated in bundle sheath cells. Transverse leaf section provided by RC Leegood.
Fig. 1.11. Schematic showing interveinal distance comparison between C$_3$ (top) and C$_4$ (bottom) leaf cross-sections. Shorter interveinal distances in leaves of C$_4$ plants allow for the efficient movement of metabolites between mesophyll (M) and bundle sheath (BS) cells. Adapted from Langdale (2011).

The efficient transport of metabolites between mesophyll and bundle sheath cells is accomplished by exclusively using plasmodesmata, which are highly abundant in C$_4$ leaves (Evert et al., 1977; Furbank, 2016). Enhancement of bundle sheath organelles was also necessary with the specialisation of the bundle sheath in leaves of C$_4$ plants. Unlike C$_3$ leaves, C$_4$ photosynthesis is partitioned between mesophyll and bundle sheath cells, and only bundle sheath cells have chloroplasts that are capable of operating the PCR cycle because of the cell-specific expression of Rubisco (Majeran and van Wijk, 2009). Moreover, the increased photosynthetic capacity of bundle sheath cells is likely to have been co-dependent with the increased number and size of chloroplasts during the development of Kranz anatomy (Sage, 2004). This is supported by Flaveria and Moricandia C$_3$–C$_4$ intermediates, which do not fully express Kranz anatomy, but have increased number of organelles in bundle sheath cells (Brown and Hattersley, 1989). Brown and Hattersley (1989) suggest that increased number of organelles in bundle sheath cells in some C$_3$–C$_4$ intermediates was enough to compensate for reduced
photosynthetic capacity due to bundle sheath cells resembling that of C₃ plants, which have reduced photosynthetic capacity and few chloroplasts (Leegood, 2008).

Photosystem activity in chloroplast thylakoid membranes is specific in C₄ plants, though varies in certain species due to the mechanism used for decarboxylation (Laetsch, 1968; Pfündel et al., 1996; Kanai and Edwards, 1999). Photosystem activity also varies between mesophyll and bundle sheath cells in C₄ plants (Meierhoff and Westhoff, 1993; Pfündel et al., 1996; Edwards et al., 2001). For instance, Sorghum bicolor has bundle sheath cells with reduced grana stacking and lower photosystem II activity, but normal photosystem I activity, while mesophyll chloroplasts have comparable photosystem II and photosystem I activities with normal grana, resembling that of C₃ plants (Woo et al., 1970; Edwards et al., 2001; Majeran and van Wijk, 2009). By reducing photosystem II activity in bundle sheath cells, the intracellular O₂ concentration probably declines, thus preventing the unfavourable utilisation of O₂ by Rubisco (Chollet and Ogren, 1972; Chapman et al., 1980). Also, enhancement of mitochondria along with peroxisomes and chloroplasts in precursors of C₄ species, would have been a requirement before the establishment of cell-specific enzyme expression and activity, and indicative of the decarboxylation mechanism that is evolving, though exceptions are known to exist (Hatch, 1978; Brown and Hattersley, 1989; Meierhoff and Westhoff, 1993).

1.3.4 Photorespiratory CO₂ pump in C₃–C₄ intermediates

The specialisation of bundle sheath chloroplasts was an important phase for the development of the photorespiratory CO₂ pump, since all dual-cell C₄ species use bundle sheath cells for the decarboxylation of the C₄-acid (Brown and Hattersley, 1989; Sage, 2004). The evolutionary trajectory of C₄ evolution has been explored in C₃–C₄ intermediate species, which possess some Kranz-like leaf characteristics and reduced rates of photorespiration, though C₄-specific enzymes are not regulated or achieve similar activities as in C₄-like or C₄ plants (Ku et al., 1991). Furthermore, the photosynthetic CO₂ compensation point of Flaveria C₃–C₄ intermediates is between that of C₃ and C₄ plants, even under ambient O₂ concentration, suggesting a reduced rate of photorespiration (Ku et al., 1983). One major development in C₃–C₄ intermediate species, which enhanced photosynthetic capacity by reducing photorespiration, was the
restriction of GDC to bundle sheath cells (Sage, 2004; Bauwe et al., 2010; Sage et al., 2012). This restricts glycine decarboxylation in mitochondria of bundle sheath cells (Figure 1.12), which release photorespired CO₂ at the site of Rubisco (Schulze et al., 2013). Therefore, compartmentation of GDC increases the capacity to recycle photorespired CO₂ and reduces the loss of carbon through photorespiration or leakage from the cell periphery (Rawsthorne et al., 1988; Bauwe et al., 2010; Gowik and Westhoff, 2011).

![Fig. 1.12. GDC compartmentalisation and the photorespiratory CO₂ pump in C₃–C₄ intermediate species.](image)

Immunolocalisation experiments conducted in the C₃–C₄ intermediate species *Moricandia arvensis* showed that the P-protein subunit (encoded by *GLDP*) of GDC was solely expressed in bundle sheath cells, which resulted in the precursors of GDC in mesophyll cells to lose activity (Rawsthorne et al., 1988; Morgan et al., 1993). Loss of GDC activity in mesophyll cells is also evident in *Flaveria* C₃–C₄ intermediate species,
and was likely acquired by the gradual suppression of GLDP and GDC activity in mesophyll cells (Ku et al., 1991; Schulze et al., 2013), rather than rapid suppression as previously suggested (Sage, 2004; Sage et al., 2012). Furthermore, the gradual transition of GDC activity from mesophyll to bundle sheath cells would have been advantageous to Kranz-like C3 plants with C3-like chloroplasts and mitochondria in bundle sheath cells, which would not have the capacity to perform all the required GDC activity (Bauwe, 2011; Muhaidat et al., 2011; Schulze et al., 2013). Therefore, certain lineages of C3 species with Kranz-like anatomy or C3–C4 intermediates, in the early trajectory of C4 evolution, which had the capacity to operate the glycine shuttle between mesophyll and bundle sheath cells would have been positively selected during evolution, whereas those with low capacity to remove phosphoglycollate would have probably not survived (Sage, 2004).

1.3.5 Establishing a fully functional C4 cycle

C3–C4 intermediates typically show a distinct progression into C4 characteristics, but do not have optimised enzyme activity to achieve similar photosynthetic capacity of C4 plants (Engelmann et al., 2003). Similar to C3 plants, Flaveria C3–C4 intermediates have high concentration of PEPC and NADP-malic enzyme in photosynthetic cells, but the activities of these enzymes are not comparable to those found in C4 plants (Ku et al., 1991). The regulatory properties of PEPC evolved at a much later stage (Svensson et al., 2003; Gowik et al., 2004) and the significance of increased PEPC activity, with regards to C4 evolution, is more evident when comparing the changes in activity between C3 and C3–C4 intermediate species. For instance, the activity of PEPC from C3–C4 intermediates Flaveria linearis and Flaveria ramosissima is five and seven times higher than that of C3 relatives, respectively (Monson and Moore, 1989). In addition, the activity of PEPC between C4 Flaveria species can be up to 40 times higher than C3 variants within the genus (Monson and Rawsthorne, 2000; Svensson et al., 2003). Enhancement of PEPC activity in mesophyll cells was necessary to re-fix leaked CO2 from the bundle sheath, which would have readily diffused along a concentration gradient through the chloroplast membrane (Monson, 1999; von Caemmerer, 2000; Gowik and Westhoff, 2011).
While *Flaveria* C₃–C₄ intermediates had elevated PEPC activity at this evolutionary stage, the C₄ pump would not have been fully established, firstly because the regeneration of PEP was not catalysed by the C₄-specific pyruvate, phosphate dikinase (PPDK), which was not acquired at the same time as PEPC activity increased in mesophyll cells (Sage, 2004). This is supported by C₃–C₄ *Flaveria* intermediates with C₃-characteristics that operated a pathway for the interconversion of 3-PGA and PEP (Monson and Moore, 1989). Secondly, CO₂ may not have been initially fixed by PEPC, since Rubisco activity was still prominent in mesophyll cells (Gowik and Westhoff, 2011). Thirdly, Rubisco acquired bundle sheath-specific expression much later in *Flaveria* species, possibly because of the glycine shuttle operating between mesophyll and bundle sheath cells in C₃–C₄ intermediates that efficiently recycled photorespired CO₂ (Monson and Rawsthorne, 2000). Carbon isotope analysis in *Flaveria* intermediates showed that a large quantity of assimilated carbon was transported through the glycine shuttle rather than the C₄ pump, suggesting that integration of C₃ and C₄ cycles would only be achieved if the activities of PEPC and Rubisco were equal and localised (Monson *et al.*, 1988).

![Image](image-url)  
**Fig. 1.13.** C₄ leaf cross-section of *Zea mays* showing compartmentmentation of Rubisco (A) and PEPC (B). Rubisco is localised in bundle sheath cells and PEPC is localised in mesophyll cells of C₄ plants. Transverse leaf sections provided by RC Leegood.
The spatial separation of Rubisco to bundle sheath cells (Figure 1.13A) and PEPC to mesophyll cells (Figure 1.13B) was the final requirement for the complete integration of C$_3$ and C$_4$ cycles (Sage, 2004). Differential gene expression patterns in Sorghum bicolor show the restriction of PEPC, PPDK, carbonic anhydrase and photosystem II proteins in mesophyll cells, while Rubisco activase, NADP-malic enzyme and other enzymes of the PCR cycle are exclusively expressed in bundle sheath cells (Wyrich et al., 1998; Sage, 2004). In the Flaveria genus, the expression of Rubisco in mesophyll cells decreased along the evolutionary gradient of C$_4$ photosynthesis, with certain C$_4$-like species, like Flaveria brownii, exhibiting the preferential expression of Rubisco in bundle sheath cells, and Flaveria variants with C$_3$-characteristics and Nicotiana tabacum (C$_3$), expressing Rubisco in both mesophyll and bundle sheath cells (Bauwe, 1984; Reed and Chollet, 1985).

Once these restrictions were set, Rubisco would not compete with PEPC for carbon and the activity of PEPC would have been high enough to match the activity of Rubisco, creating a gradient of metabolites that diffused between C$_3$ and C$_4$ cycles and meet competing demands for ATP and CO$_2$ in mesophyll and bundle sheath cells, respectively (Monson, 1999; von Caemmerer and Furbank, 1999; von Caemmerer and Furbank, 2003; Gowik and Westhoff, 2011). Additionally, enhancement of mesophyll-specific carbonic anhydrase activity was necessary to generate enough bicarbonate for PEPC to avoid limiting the rate of photosynthesis (Hatch and Burnell, 1990). The reduction of carbonic anhydrase activity in bundle sheath cells was also an important step in C$_4$ evolution (Sage, 2004; Ludwig, 2012). As the intracellular concentration of CO$_2$ increases in bundle sheath cells, carbonic anhydrase would catalyse the conversion of CO$_2$ to bicarbonate, thus reducing the amount of substrate available for Rubisco (Burnell and Hatch, 1988). In C$_4$ species, carbonic anhydrase is expressed in large quantities in the cytosol of mesophyll cells, though the gene was acquired from a chloroplastic form (Ludwig, 2012), but lost its transit peptide due to a mutation to the sequence that encoded its chloroplastic expression (Tanz et al., 2009).
While C₃ plants contain high amounts of decarboxylating enzymes (NADP- and NAD-malic enzyme), these were expressed in vascular tissues, and not in photosynthetic cells and had to acquire cell-specific expression patterns (Hibberd and Quick, 2002; Sage, 2004). Both PEPC and NADP-malic enzyme in C₄ Flaveria species acquired C₄ characteristics through the duplication of an ancestral C₃ gene followed by neofunctionalisation, which led to cell-specific gene expression (Hermans and Westhoff, 1990; Marshall et al., 1996; Monson, 1999). Further establishment of the C₄ syndrome was dependent on substantial changes in gene expression patterns carried forward from ancestor C₃ species (Nelson and Dengler, 1992; Ku et al., 1996), but only needed slight changes to regulatory elements to achieve C₄-specific function (Gowik and Westhoff, 2011). For instance, PEPC expression in C₄ plants is controlled by a transcriptional cis-regulatory element (mesophyll expression module1, MEM1), which encodes for the mesophyll-specific expression of the ppcA gene of PEPC in Flaveria trinervia (Gowik et al., 2004). The MEM1 sequence is also found in the orthologous C₃ Flaveria pringlei gene, but it does not encode PEPC expression in mesophyll cells (Akyildiz et al., 2007). It was shown that two discrete changes in the MEM1 sequence were substantial to generate the mesophyll-specific expression of PEPC in Flaveria trinervia (Akyildiz et al., 2007). The cell-specific expression of the Rubisco found in C₄ plants is another well-documented example of subtle changes to gene expression. While most C₄-specific genes are regulated by transcriptional elements (Sheen, 1999), certain genes, like those in Flaveria bidentis encoding for bundle sheath-specific expression of the small subunit of Rubisco, are controlled by both transcriptional and posttranscriptional regulators (Patel et al., 2006). Furthermore, despite cell-specific expression, the expression patterns of Rubisco genes in C₄ plants are very similar to those found in their C₃ relatives (Berry et al., 2016). In both plant systems Rubisco genes are absent in non-photosynthetic tissues and ubiquitously expressed in photosynthetic tissues, where they are regulated by developmental signals, illumination or abiotic stress (Berry et al., 2013; Berry et al., 2016).

The eight large and small subunits of Rubisco are encoded by rbcL and rbcS genes, respectively (Patel et al., 2006; Berry et al., 2016). Despite that in the early stages of C₄ evolution these genes were expressed in both mesophyll and bundle sheath cells, Rubisco mRNAs were shown to only accumulate in the latter (Berry et al., 2013). Furthermore, Berry et al. (2016) suggested that nucleic acid binding proteins, such as RLSB, could be
affecting the regulation of cell-specific Rubisco genes. Recently, highly conserved RLSB proteins, from C₃ ancestral *Flaveria* plants, were shown to accumulate in bundle sheath cells and bind *rbcL* mRNAs, in a mechanism that describes the regulation of cell-specific Rubisco expression in C₄ plants (Yerramsetty *et al.*, 2017). Changes to the sequence of RLSB was observed to occur gradually through the progression of C₃ to C₄-like to C₄ *Flaveria* plants, which suggest that RLSB co-evolved with C₄ photosynthesis, albeit the functional significance of these changes through the transition to C₄-characteristics is unclear (Yerramsetty *et al.*, 2017).

The light-dependent expression of the small subunit of Rubisco in bundle sheath cells is similar to the mechanism previously described for the cell-specific expression of NAD-malic enzyme in the C₄ plant *Amaranthus hypochondriacus*. In *Amaranthus hypochondriacus*, the functional subunit (α subunit) of NAD-malic enzyme is preferentially expressed in bundle sheath cells (Long *et al.*, 1994). During the early stages of leaf development, and induced by light, the mRNAs corresponding to the α subunit accumulate in bundle sheath cells in amaranth (Long and Berry, 1996). However, the NAD-malic enzyme bundle sheath-specific expression pattern in amaranth occurs earlier than that of Rubisco (Long and Berry, 1996). During the early stages of leaf development, mRNA products of the small and large subunits of Rubisco initially accumulate in mesophyll and bundle sheath cells, resembling the expression pattern of C₃ plants, and bundle sheath cell specificity is acquired later through metabolism cues (Wang *et al.*, 1992; Wang *et al.*, 1993). Furthermore, acquisition of a C₄-specific Rubisco might have occurred with fewer changes to regulation patterns, since it was acquired from a photosynthetic counterpart, whereas NAD-malic enzyme, PEPC and PPDK, were recruited from separate, non-photosynthetic origins (Hatch, 1987; Furbank and Taylor, 1995). This also suggests that the regulatory patterns for the three latter enzymes were recruited around the same time as C₄-specific function evolved and therefore share similar regulatory elements (Long and Berry, 1996; Gowik and Westhoff, 2011).
1.3.7 Fine-tuning C₄-specific enzyme activity

Once the C₄ cycle was fully established, the enzymes that were recruited from ancestral C₃ counterparts to serve a C₄-specific function were optimised for C₄ photosynthesis (Sage, 2004). In order to maintain the metabolite gradient that operates the C₄ cycle, the rate at which PEPC fixes CO₂ in mesophyll cells must be equivalent to the rate of diffusion into bundle sheath cells, utilisation by Rubisco and regeneration of PEP (Leegood and Walker, 1999). In contrast, the rate of transport is no higher than one-sixth of the steady-state CO₂ assimilation in C₃ leaves (Edwards and Walker, 1983). The kinetic properties of C₄-related enzymes were also optimised during this phase. For instance, the C₄-specific PEPC affinity for PEP is lower than that of its C₃ counterpart, while the affinity for bicarbonate is higher in the C₄-form (Chollet et al., 1996; Gowik and Westhoff, 2011), and achieves higher maximum reaction rates than the C₃-form because PEP is available in larger quantities (Ku et al., 1996). Also, the C₃ PEPC form is sensitive to inhibition by malate, but this is reduced in the C₄-form, so that it can remain active when large quantities of malate are produced during C₄ photosynthesis (Bläsin et al., 2000; Svensson et al., 2003; Jacobs et al., 2008). Although the C₄-specific PEPC has evolved at least eight times through separate origins, all C₄ PEPCs share a large degree of similarity in Flaveria (Christin et al., 2007; Gowik and Westhoff, 2011). Moreover, it is likely that these features were the last to evolve, based on the absence of these regulatory mechanisms in PEPC from Flaveria C₃–C₄ intermediates (Sage, 2004). Similarly, the C₄-specific Rubisco has a lower affinity for CO₂, relative to C₃ counterparts, but can achieve a higher turnover rate because CO₂ is more readily available (Seemann et al., 1984). Similar to the regulatory features of the C₄ PEPC, Rubisco from Flaveria intermediates do not exhibit C₄-specific kinetic properties, but more closely resemble those of the C₃-form (Wessinger et al., 1989), suggesting that these features also evolved in the final phases of C₄ evolution (Sage, 2004). NADP-malic enzyme also underwent slight changes to its kinetic properties when it was recruited for C₄-specific function. Unlike the C₃-form, NADP-malic enzyme in C₄ species exhibits a higher maximum activity and has a higher affinity for malate (Drincovich et al., 2001). Also, unlike the C₄-specific PEPC, NADP-malic enzyme in Flaveria C₃–C₄ intermediates exhibited changes to affinity values for malate, somewhere between the
values of C₃ and C₄ variants (Sage, 2004), suggesting that some intermediates evolved these characteristics earlier than the regulatory properties PEPC or Rubisco evolved.

1.4 What does it mean to be C₄?

1.4.1 Variations of the C₄ pathway

During the evolution of the C₄ pathway, three decarboxylases were recruited (NAD(P)-malic enzyme and phosphoenolpyruvate carboxykinase [PEPCK]), creating three distinct biochemical variations of the CO₂ concentrating mechanism of C₄ photosynthesis (Hatch, 1987). In the NADP-malic enzyme subtype (Figure 1.14A), malate is formed in mesophyll cell chloroplasts, whereas in NAD-malic enzyme (Figure 1.14B) and PEPCK subtypes (Figure 1.14C), both malate and aspartate are formed (Wang et al., 2011). PEPCK operates as a secondary decarboxylase in several NADP- and NAD-malic enzyme subtypes, and does not appear to operate on its own (Walker et al., 1997; Wingler et al., 1999; Pick et al., 2011; Bräutigam et al., 2014). For instance, in NADP-malic enzyme subtypes, aspartate contributes to roughly 10–15% of assimilated carbon pool (Hatch and Mau, 1973; Hatch, 1987; Arrivault et al., 2016).

The coordination of C₃ and C₄ pathways depends on the cooperate activity of carboxylation and decarboxylation reactions (Bailey et al., 2007) and coordinate interaction of the C₄ pathway and mitochondrial metabolism (Leegood, 1985; Stitt and Heldt, 1985; Leegood and Walker, 1999). In PEPCK subtypes, for example, mitochondrial respiration generates the ATP required for the operation of PEPCK in bundle sheath cells (Hatch et al., 1988). Moreover, NAD-malic enzyme decarboxylation activity in the mitochondria does not compete with PEPCK decarboxylation (Figure 1.13C), because the rate at which ATP is generated for PEPCK activity is limiting (Carnal et al., 1993) and both decarboxylases contribute to the delivery of CO₂ at the site of Rubisco (Bräutigam et al., 2014). In PEPCK type C₄ plants, PEP is regenerated in bundle sheath cells from the decarboxylation of oxaloacetate requiring one molecule of ATP, whereas in NADP- and NAD-malic enzyme subtypes, PEP is regenerated via PPDK in mesophyll cells using two molecules of ATP per molecule of CO₂ assimilated.
Theoretical models predict that the energetic costs of PEPCK subtypes are slightly lower than in NADP-malic enzyme subtypes, requiring less ATP per CO₂ molecule assimilated, although require one more molecule of ATP than C₃ plants (Kanai and Edwards, 1999; Bräutigam et al., 2014).

Fig. 1.14.
Fig. 1.14. Three biochemical variations of the C₄ pathway using the (A) NADP-malic enzyme, (B) NAD-malic enzyme or (C) PEPCK decarboxylase. C₄ photosynthesis is characterised by the primary decarboxylase (dark blue) operating in bundle sheath cells. Enzymes (light blue): CA, carbonic anhydrase; PEPC, phosphoenolpyruvate carboxylase; AAT, aspartate aminotransferase; MDH, malate dehydrogenase; ME, malic enzyme; PEPCK, PEP carboxykinase; AlaAT, alanine aminotransferase. Metabolites (purple): OAA, oxaloacetate; Mal, malate; Asp, aspartate; Ala, alanine; Glu, glutamate; 2-OG, 2-oxoglutarate. Transporters: 1, DiT1; 2, BASS2/NHD1; 3, PPT. PCR, photosynthetic carbon reduction cycle. Dashed lines show diffusion routes. *AlaAT. Adapted from Wang et al. (2011).

Aside from changes to the kinetic properties of Rubisco in C₄ plants, all other enzymes of the PCR cycle present in bundle sheath cells of C₄ plants are unchanged from their C₃ counterparts (Ashton et al., 1990; Leegood and Walker, 1999). In C₃ leaves, triose phosphates that are generated in the PCR cycle in mesophyll cells are exported from the chloroplast, via a triose phosphate translocator, to be used in the cytosol for sucrose synthesis (Leegood and Walker, 1999). Since C₄ plants lack the ability to operate the PCR cycle in mesophyll cells due to the absence of Rubisco, phosphate, which is released from sucrose synthesis, is transported into bundle sheath chloroplasts and used for the regeneration of RuBP (Leegood and Walker, 1999). Triose phosphates accumulate in large quantities in C₃–C₄ intermediates Flaveria floridana and Moricandia arvensis, and the concentration is about 20-fold higher in C₄ plants than in C₃ plants (Leegood and von
Caemmerer, 1994). Movement of carbon through the 3-PGA/triose-phosphate shuttle (Figure 1.15) is another important feature in all C₄ plants and is likely to have evolved in C₃–C₄ intermediates to alleviate the burden on the PCR cycle in early developmental stages of C₄-competent bundle sheath chloroplasts (Leegood, 2013). In this shuttle, the concentration gradient of about 10 mM triose phosphate drives diffusion from mesophyll to bundle sheath cells, and in exchange return 3-PGA driven by a concentration gradient of about 9 mM (Stitt and Heldt, 1985).

**Fig. 1.15.** 3-PGA/triose-phosphate shuttle operating between bundle sheath and mesophyll cells. The conversion of 3-PGA to triose phosphate requires three enzymatic reactions catalysed by phosphoglycerate kinase, glyceraldehyde-3-phosphate dehydrogenase and triose phosphate isomerase (see Figure 1.2).
1.4.2 Increased capacity of metabolite movement

The operation of the C₄ cycle depends on the efficient and rapid movement of metabolites through diffusion driven transport (Leegood, 1985; Leegood and von Caemmerer, 1989; Leegood and Walker, 1999). In C₄ plants, the capacity of these transport mechanisms are enhanced to move large quantities of metabolites between mesophyll and bundle sheath cells (Hatch and Osmond, 1976). In contrast to C₃ plants, which require one transport step to produce triose phosphates, C₄ plants require at least 30 steps (Weber and von Caemmerer, 2010). In NADP-malic enzyme subtypes, large quantities of oxaloacetate diffuse from the cytosol to chloroplasts of mesophyll cells via a translocator with high specificity for oxaloacetate but not malate (Hatch et al., 1984). Another non-specific translocator with competing specificity for malate and oxaloacetate operates in mesophyll chloroplasts of Zea mays leaves, but is not efficient to operate under high load of C₄ photosynthesis when malate is in higher content than oxaloacetate (Day and Hatch, 1981; Hatch, 1987). During C₄ photosynthesis, a concentration gradient of about 18 mM malate drives diffusion from mesophyll cells to bundle sheath cells (Stitt and Heldt, 1985). In a recent study, however, Arrivault et al. (2016) noted that previous estimations were biased, since metabolite concentrations were determined from whole leaf tissue, instead of isolated mesophyll or bundle sheath cells.

These data also showed that malate, in Zea mays, contributes to the diffusion driven transport, and is about twofold more abundant in mesophyll than bundle sheath cells, resulting in a concentration gradient of 6 mM and not 18 mM as previously reported (Arrivault et al., 2016). Furthermore, in NADP-malic enzyme subtypes, aspartate contributes to about 10% of the carbon pool and moves from mesophyll to bundle sheath cells in an estimated gradient of 0.96 mM (Kanai and Edwards, 1999; Arrivault et al., 2016). In PEPCK subtypes, like Megathyrsus maximus, malate is imported into mitochondria in bundle sheath cells via a mitochondrial dicarboxylate carrier (DIC), but the transport of pyruvate from the mitochondrion to the cytosol has not been elucidated (Bräutigam et al., 2011). Also, despite that the transport of pyruvate from bundle sheath to mesophyll cells via BASS2 (BILE ACID:SODIUM SYMPORTER FAMILY PROTEIN 2) and NDH1 (sodium:proton antiporter) has been elucidated in Flaveria bidentis (NADP-malic enzyme) and Cleome gynandra (NAD-malic enzyme), it has yet to be described in Zea mays (Furumoto et al., 2011; Leegood, 2013).
1.4.3 Physiology and performance of C₄ plants

C₄ plants are characterised for their capacity to maintain high rates of photosynthesis in low CO₂ concentrations (Figure 1.16). This is because CO₂ can be concentrated around Rubisco, allowing Rubisco to operate under high CO₂ concentrations, thus suppressing the rate of photorespiration and inhibition by O₂ (Sage et al., 2012). To allow such a response, membrane permeability has been increased to allow the rapid diffusion of metabolites, without compromising on the leakage of CO₂ or bicarbonate (Hatch, 1987).

![Fig. 1.16. Modelled relationship between the rate of photosynthesis (PSₐ) and intracellular concentration of CO₂ in leaves of C₃ (blue) and C₄ plants (red). The initial slope of CO₂ assimilation in C₄ plants is greater than that of C₃ plants at a low intracellular CO₂ concentration, demonstrating the ability of C₄ plants to perform better in low CO₂. Redrawn from Pearcy and Ehleringer (1984).](image)

The proportion of CO₂ leaked from C₄ leaves is minimised by the capacity of PEPC to rapidly assimilate CO₂ into the C₄-acid, along with the expression of Kranz anatomy (Hatch and Osmond, 1976; Hatch, 1987; Cousins et al., 2007). However, the rate of photosynthesis in C₄ plants, at low CO₂, is limited by the rate at which bicarbonate is generated by carbonic anhydrase and delivered to PEPC (Sage and Kubien, 2007; Boyd
et al., 2015). Under such conditions, PEPC activity is independent of temperature unless the temperature falls below 10 °C, which increases the rate of PPDK inactivation (Shirahashi et al., 1978; Hatch, 1979; Long, 1983; Ohta et al., 2004) and the activity of PEPC becomes limited by the capacity of PPDK to regenerate PEP. In contrast and despite limited by low atmospheric CO₂, C₃ plants can achieve a higher maximal CO₂ assimilation rate because Rubisco is in higher content in C₃ leaves than in C₄ leaves (Ku et al., 1996). However, nitrogen-use efficiency is increased in C₄ plants because less Rubisco is required to operate the PCR cycle (Brown, 1978; Schmitt and Edwards, 1981; Sage et al., 1987).

C₄ grasses evolved under selective pressures such as increasing temperature, high light and declining CO₂ (Sage, 2004). A major characteristic of C₄ plants, especially those occupying open and arid grasslands, is to regulate the amount of water lost through transpiration by improving leaf hydraulics (Kocaçinar and Sage, 2004; Sack and Holbrook, 2006). As a result, stomata in C₄ plants can remain closed for longer during the daytime, when the light intensity is presumably higher, thus increasing water-use efficiency, which is at least twice as efficient as in C₃ plants (Hatch, 1987; Hatfield and Prueger, 2011). This is especially important in regions with higher temperatures and variable rainfall (Long, 1999; Gowik and Westhoff, 2011). There is also a distinct relationship between rate of assimilation and stomatal regulation. For instance, increased CO₂ fixation in C₄ grasses reduces stomatal conductance and minimises water loss (Osborne and Sack, 2012). A recent hydro-mechanical model of stomatal conductance of C₄ photosynthesis suggests that quicker responses in the mechanisms controlling stomatal opening and closure might have given C₄ plants the competitive advantage in achieving and maintaining high rates of CO₂ assimilation relative to C₃ plants (Bellasio et al., 2017). An improved hydraulic design is likely to have co-evolved with the CO₂ concentrating mechanism of C₄ photosynthesis and would have served a significant role for the emergence of C₄ plants in arid zones (Osborne and Sack, 2012; Bellasio et al., 2017).
Fig. 1.17. **Overview of the global distribution of higher plants dominated by either C₃ (yellow) or C₄ grasses (orange).** Forests (green) and cropland (red) are also indicated. Figure does not consider the seasonal emergence of C₃ and C₄ plants in some grasslands. Black dots show the regions of C₄ grasslands whose geological history are best described. Figure and annotation obtained from Edwards et al. (2010).

In addition, the CO₂ compensation point of C₄ plants does not increase as much as it does in C₃ plants in response to high temperatures (Sage and Kubien, 2007). Therefore, C₄ leaves are less sensitive to leaf temperature, and better adapted for higher temperatures, somewhere in the 25 to 30 °C range ( Ehleringer and Björkman, 1977). As a result, the landscape of grasses that populate the globe has drastically changed since the evolution of C₄ photosynthesis and it is believed that the C₄ syndrome was so successful that neighbouring ancestor C₃ grasses, mainly of tropical origin, were outcompeted as the atmospheric concentration of CO₂ decreased and temperatures increased (Sage, 1999; Edwards et al., 2010; Sage et al., 2012). C₄ grasses dominate in warm, tropical or subtropical environments (Figure 1.17), favouring higher daytime temperatures and lower altitudes, though need plenty of water and like their C₃ relatives are not adapted to drought-stress (Sage, 2004; Edwards et al., 2010). Although the C₄ pathway was selected as an advantageous trait for increasing carbon gain through higher photosynthetic efficiency (Monson, 2003), higher energetic costs of C₄-related reactions, such as the ATP-dependent regeneration of PEP via PPDK or ATP-dependent decarboxylation by PEPCK, make C₄ plants less competitive in regions where C₃ photosynthesis is favoured (Bauwe et al., 2010; Sage et al., 2012; Lundgren et al., 2016). It is likely, however, that the evolution of C₄ photosynthesis in grasslands allowed the broadening, rather than replacement, of existing ecological niches, thus allowing new C₄ populations to occupy
larger spaces, which gradually increased the ecological gap between C\textsubscript{3} and C\textsubscript{4} grasses (Edwards \textit{et al.}, 2010; Lundgren \textit{et al.}, 2015). The capacity of C\textsubscript{4} plants to operate at their optimum in regions with varying environmental conditions or extremes, such as temperature, water availability or soil salinity, which would otherwise directly promote photorespiration and dark respiration in C\textsubscript{3} plants (Brooks and Farquhar, 1985; Sharkey, 1988), have greatly contributed to C\textsubscript{4} evolution and the distribution of C\textsubscript{4} plants (Sage, 2004; Edwards \textit{et al.}, 2010). Moreover, the diversity of these environmental conditions could promote evolution of C\textsubscript{4} photosynthesis in some taxa, but not others (Lundgren \textit{et al.}, 2015). This is supported by the existence of several intermediate species, including \textit{Flaveria} intermediates (Powell, 1978), in arid and saline habitats, which are zones where C\textsubscript{4} species originated from (Sage, 2004). However, there is still some doubt as to how much photorespiration is reduced in C\textsubscript{4} plants, and whether the three distinct subtypes of C\textsubscript{4} photosynthesis are truly representative or rather an oversimplification of the C\textsubscript{4} syndrome. We will understand the factors that contribute to the efficiency of carbon movement and assimilation in C\textsubscript{4} plants by investigating how the enzymes involved in C\textsubscript{4} photosynthesis are regulated (Furbank, 2011; Bellasio and Griffiths, 2014; Wang \textit{et al.}, 2014; Arrivault \textit{et al.}, 2016).
1.5 Aim of thesis and thesis structure

This research project was part of an international consortium which attempts to engineer characteristics of C₄ photosynthesis into existing C₃ crops in hopes to improve photosynthetic performance and crop yields (Leegood, 2013). However, the post-translational regulatory mechanisms of C₄-related proteins and how that regulation translates to the efficiency of carbon assimilation through a diurnal cycle remains unclear. Therefore, it is vital to identify the factors that contribute to the regulation C₄-related proteins prior to engineering artificial protein networks into existing C₃ crops, such that the efficiency of carbon fixation, water-use efficiency and nitrogen-use efficiency are achieved. Post-translational modifications (PTMs) play a crucial role in plant biology (Friso and van Wijk, 2015) and their extensive regulatory mechanisms remain largely undiscovered in C₄ plants. Characterising unknown PTMs of C₄-related enzymes can help to understand the regulation of C₄ photosynthesis, prior to incorporation into C₃ plants. With this information, artificial regulatory mechanisms can be implemented in C₄ crop plant transformants, thus maintaining the expected photosynthetic efficiency (Komatsu et al., 2013; Furbank, 2016). Furthermore, exploiting these regulatory mechanisms can help maintain biochemical stability and increase abiotic stress tolerance of transformant crop plants grown in harsh nutrient-deficient environments (Hashiguchi and Komatsu, 2016).

The primary aim of this study was to identify light-dependent PTMs of key C₄-related proteins in leaves of *Setaria viridis* (NADP-malic enzyme subtype). Protein phosphorylation plays an important role in regulating the diurnal activity of PEPC, PEPCK and PPDK and it is hypothesised that several other C₄ enzymes, including NADP-dependent malic enzyme, aspartate aminotransferase and alanine aminotransferase undergo light-dependent phosphorylation. It is expected that not every phosphorylation site that is identified is under regulation or contributes to enzyme activity. However, identifying novel phosphorylation sites might provide some indication to undiscovered phosphoregulatory mechanisms that induce distinct changes to diurnal enzyme activity.
The secondary aim of this project was to ascertain whether NADP-malic enzyme, aspartate aminotransferase and alanine aminotransferase have similar kinetic properties across three C₄ NADP-malic enzyme monocot grasses, as well as explore the dual-decarboxylase system in NADP-malic enzyme subtypes and determine if there are distinct variations in the properties of aspartate and alanine aminotransferase in *Setaria viridis*, *Sorghum bicolor* and *Zea mays*. In the dual-decarboxylase system, C₄-acid decarboxylation is partitioned between NADP-malic enzyme and PEPCK or NAD-malic enzyme (Wang *et al.*, 2014). However, whether these pathways are differently regulated in closely related C₄ grasses is unclear. The C₄ grasses used in this study were chosen because they use NADP-malic enzyme as the primary decarboxylase, though have varying amounts of PEPCK activity (Hatch, 1987; Bräutigam *et al.*, 2014). Furthermore, *Sorghum bicolor* and *Zea mays* share a common ancestor whereas *Setaria viridis* evolved C₄ photosynthesis through a separate origin (Christin *et al.*, 2009). While it is accepted that NADP-malic enzyme activity is pH-dependent across C₄ species (Johnson and Hatch, 1970), it is hypothesised that each of the C₄ grasses in this study regulate NADP-malic enzyme differently and dependence on pH varies in dark and light conditions. Finally, it is hypothesised that there are interspecies differences in the kinetic properties of NADP-malic enzyme, aspartate aminotransferase and alanine aminotransferase.

The novel findings of this project are presented in Chapter 3, Chapter 4 and Chapter 5 of this thesis. A summary of our current understanding of the regulation of C₄-related proteins is presented in Chapter 3, followed by novel light-dependent phosphorylation sites identified in *Setaria viridis* (supplementary data in Appendix A). In Chapter 4, the light and dark *in vitro* activities and the corresponding Michaelis-Menten constants ($K_M$) of NADP-malic enzyme from *Setaria viridis*, *Sorghum bicolor* and *Zea mays* are presented and discussed. In Chapter 5, novel findings will show that the activity of aspartate aminotransferase is sensitive to dark and light changes and only aspartate aminotransferase in *Zea mays* is activated by L-malate and L-alanine. In Chapter 5 and 6, the implications of these findings will be discussed, with emphasis towards a putative, regulatory mechanism, which controls the flux of carbon, formation of PEP and efficiency of carbon assimilation in C₄ photosynthesis and sugar homeostasis during illuminated and darkened conditions.
Chapter 2 – Methods

2.1 Proteomics and Mass Spectrometry

2.1.1 Growth of plant material

Setaria viridis seeds accession A10 (harvested on April 2013) were received from Dr Asaph Cousins’ Lab (Washington State University, Pullman, WA, USA) in early 2014. Seed dormancy was overcome by cold stratification (Brutnell et al., 2010) or incubation at 45 °C prior to sowing (Rizal et al., 2013).

Fig. 2.1. General workflow for the extraction of leaf proteins, sample processing and analysis by mass spectrometry. Mass spectrometry method optimised for in-gel digestion of plant proteins. Samples were analysed using electrospray ionisation LC-MS/MS with ultra-sensitive Orbital Trap mass spectrometer.
Plants were grown for seed production and seeds were harvested from one-month old plants. To increase seed viability upon collection from bristles, seeds were dried at 40 °C for five days then stored at room temperature. Ten-month old *Setaria viridis* seeds were culled and sown into large trays containing nutrient rich M3 compost (ICL Levington, Ipswich, UK), supplemented with nitrogen (144 mg L$^{-1}$), phosphorus (73 mg L$^{-1}$) and potassium (239 mg L$^{-1}$) and covered with a thin layer of compost, following the planting method described by Jiang *et al.* (2013). Trays were watered with distilled water and propagator lids were fixed on potting trays. Lids were covered with a sheet of black cloth for no more than 36 h to allow for dark germination. Plants were grown in an environment controlled growth chamber using a 16 h photoperiod (28 °C/26 °C light/dark), maintaining 350 µmol m$^{-2}$ s$^{-1}$ photosynthetic photon flux density (PPFD), 70% relative humidity and CO$_2$ not exceeding 600 ppm. A week after germination, seedlings were transplanted into individual 7 cm pots and grown until the 6th leaf was fully emerged, equivalent to 2.5 weeks after germination (Figure 2.2).

![Mature leaf ready to harvest](image)

**Fig. 2.2.** 2.5-week old *Setaria viridis* photographed before the harvesting of leaves. During the harvest, cuts were made approximately half way from the tip and base of the leaf, indicated by the dashed line. One half-leaf section was collected per plant. 20 half-leaves were used for each time-point.
2.1.2 Extraction of leaf proteins for mass spectrometric analysis

All mass spectrometry protocols were performed under sterile conditions using proteomic-grade reagents and MilliQ water (18.2 MΩ cm⁻¹). To assess changes in protein post-translational modifications during the dark–light transition, 20 half-leaf sections from different 2.5-week old Setaria viridis plants (Figure 2.2) were harvested and flash-frozen in liquid nitrogen at 4 time-points: 2 h and 7.5 h into the dark period, 4 h and 15.5 h into the photoperiod. Frozen leaf tissue was ground to a fine powder in liquid nitrogen using a pestle and mortar, homogenised in 200 mM Bicine-KOH, pH 9.8 and 50 mM dithiothreitol (DTT) containing 1× ethylenediaminetetraacetic acid (EDTA)-free protease and PhosSTOP phosphatase inhibitors (Roche, Mannheim, Germany) (Figure 2.3). Crude extracts were cleared by centrifugation in Eppendorf tubes at 16,800 × g. Protein concentrations were determined spectrophotometrically at 595 nm following Bradford (1976) using bovine serum albumin (BSA) standards. Crude extracts were solubilised and boiled in sodium dodecyl sulfate (SDS) solubilisation buffer containing 100 mM Tris-HCl pH 6.8, 200 mM DTT, 20% (v/v) glycerol, 4% (w/v) SDS, 0.2% (w/v) bromophenol blue (BPB).

![Fig. 2.3. Classical method for the extraction of proteins from plant leaves using liquid nitrogen](image)

Proteins were extracted by grinding leaf tissue in liquid nitrogen using a pestle and mortar and homogenising in Bicine-KOH/DTT extraction buffer. Plant lysate was collected in Eppendorf tubes and cleared by centrifugation.

2.1.3 In-gel tryptic digestion

To analyse proteins by mass spectrometry, protein extracts were resolved by one-dimensional (1D) SDS polyacrylamide gel electrophoresis (PAGE) using 4–20% gradient pre-cast SDS-gels (Bio-Rad, Hertfordshire, UK). Gels were run at 180 V for 40
min in a Bio-Rad electrophoresis gel tank containing 1× Tris-Glycine-SDS (TGS) running buffer. After electrophoresis, gels were stained with Expedeo InstantBlue™ (Expedeo, Harston, UK). Bands were excised, placed into Eppendorf Lo-Bind microtubes (Figure 2.4) and destained at 37 °C in a solution containing 200 mM ammonium bicarbonate (ABC) and 40% acetonitrile (ACN). Gel pieces were dried down at 45 °C in a vacuum concentrator (Eppendorf, Stevenage, UK), and then reduced for 1 h at 56 °C with 200 µL 10 mM DTT and 50 mM ABC. Proteins were then alkylated in the dark at room temperature in a solution containing 55 mM iodoacetamide (IAA) and 50 mM ABC. Gel pieces were briefly washed with 50 mM ABC and 50% ACN then dried down. Tryptic digestion reactions were performed at 37 °C using 0.4 µg of trypsin from porcine pancreas (Sigma-Aldrich, Saint Louis, USA) in 70 µL containing 1 mM HCl, 40 mM ABC and 9% ACN. The following day, peptides were transferred to collection Lo-Bind microtubes and gel pieces were treated with 100% ACN and 5% formic acid to protonate and extract remaining peptides. Extracted peptides were slowly dried at 30 °C to remove interfering organic solvents, and stored at –20 °C until use.

Fig. 2.4. General workflow for sample processing using an in-gel tryptic digestion approach, prior to analysis by mass spectrometry. First, proteins were resolved on 4–20% or 4–12% pre-cast SDS Bis-Tris gel and stained with Coomassie InstantBlue™. Proteins in the SDS matrix were washed, reduced and alkylated to linearise peptides. Finally, proteins were treated with trypsin. Scissor symbols indicate trypsin cleavage sites at amino acids (orange circles) lysine (K) and arginine (R) within the protein (represented by the blue line). Peptides were de-salted using C18 spin tips to remove detergents and organic contaminants prior to analysis by mass spectrometry.
2.1.4 C18 column clean-up

To increase peptide recovery and spectrum resolution, dried down peptides were desalted using Pierce C18 spin columns (ThermoFisher Scientific, Waltham, USA). Spin columns were activated with 50% methanol and washed with 0.5% trifluoroacetic acid (TFA) and 5% ACN (equilibration buffer). Peptides were reconstituted in 50 µL equilibration buffer, applied to the C18 resin and centrifuged at 1000 × g, following manufacturer’s protocol (Pierce Biotechnology, Rockford, USA). Peptides were eluted in 70% ACN containing 0.1% formic acid.

2.1.5 Phosphopeptide enrichment by titanium dioxide

To enrich for phosphopeptides by titanium dioxide (TiO$_2$) metal oxide affinity chromatography (MOAC), dried down peptides were reconstituted in 150 µL 26% lactic acid, 0.4% TFA and 80% ACN (Buffer A) and applied to the surface of the activated TiO$_2$ resin (Figure 2.5) and centrifuged at 1000 × g. The flow-through was reapplied to the resin to maximise phosphopeptide binding. TiO$_2$ columns were washed with Buffer A at 3000 × g, and peptides were eluted in 1.5% ammonium hydroxide and 5% pyrrolidine. Samples were desalted using a graphite clean-up kit.
Fig. 2.5. Workflow schematic illustrating the basic flow of analyte through TiO$_2$ phosphopeptide enrichment columns. Four Eppendorf tubes per time-point, containing dried down peptides, were removed from the –20 °C freezer and equilibrated to room temperature. Peptides were reconstituted in 150 µL lactic acid and run through activated TiO$_2$ resins (blue cones) by centrifugation at low speed. Phosphopeptides were eluted into a new Eppendorf tube with 1.5% ammonium hydroxide and 5% pyrrolidine.
2.1.6 Mass spectrometry analysis

Mass spectrometry analyses, using a high-performance liquid chromatography (UHPLC) MS/MS Orbital Trap Elite mass spectrometer, were performed by Dr Richard G. Beniston (Senior Scientific Officer, biOMICS, The University of Sheffield) from February 2014 to November 2015 and Dr Adelina E. Acosta Martin (Facility Manager, biOMICS, The University of Sheffield) from April 2016 to June 2017. The protocols for in-gel protein digests, instrument calibration, sample injection and data analysis were issued and performed as a service by biOMICS and optimised through communication with Dr R. G. Beniston and Dr A. E. Acosta Martin.

2.1.7 General mass spectrometry protocol

Dried down peptides were reconstituted in 0.1% formic acid and 2% ACN and injected into a Dionex Ultimate 3000 UHPLC using a PepMap100 C18 trap column at a constant rate of 10 µL min⁻¹. The separation phase was performed over a 71 min gradient with increasing ACN concentrations from 2.4% to 72%, using a 15 cm PepMap100 C18 analytical column (2 µm particle size, 100 Å pore size, 75 μm I.D) (ThermoFisher Scientific, Waltham, USA) at a rate of 250 nL min⁻¹, at 35 °C. Nanospray ionization was performed at 2.0 kV, with the ion transfer capillary at 250 °C and S-lens setting of 60%. MS1 spectra ranging from 350–2000 m/z were acquired at a resolving power of 60,000. Following MS1 analysis, the top ten most abundant precursors were selected for MS2 analysis using collision induced dissociation (CID).

2.1.8 Immunoblotting

Immunoblots were performed using PEPC and NADP-malic enzyme antibodies raised from C₄ sequences, as described in Lundgren et al. (2016). Polypeptides were visualised with their appropriate Immunoglobulin G (IgG) secondary antibody grown in rabbit (polyclonal) or mouse (monoclonal) (Sigma-Aldrich, St. Louis, USA), in conjunction
with an enhanced chemiluminescence (ECL) blotting kit and Hyperfilm ECL (GE Healthcare, Buckinghamshire, UK).

Fig. 2.6. Proteome Discoverer workflow used for identifying novel phosphorylation sites. Resulting spectrum files were searched against spectral libraries using MASCOT and SEQUEST and validated using a target decoy database. Phosphorylations were determined using the PhosphoRS algorithm in Proteome Discoverer.
Table 2.1. **MS/MS search engines used during the study.** The main MS/MS search engines that were available during the study. All search engines score protein matches as a measure of how similar experimental spectra are to theoretical peptide spectrum files.

<table>
<thead>
<tr>
<th>Search Engine</th>
<th>Developer</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASCOT</td>
<td>Matrix Science</td>
<td><a href="http://www.matrixscience.com">http://www.matrixscience.com</a></td>
</tr>
<tr>
<td>SEQUEST</td>
<td>ThermoFisher Scientific</td>
<td><a href="http://www.thermo.com/com/cda/product/detail/0,1055,22209,00.html">http://www.thermo.com/com/cda/product/detail/0,1055,22209,00.html</a></td>
</tr>
<tr>
<td>MaxQuant</td>
<td>Max Planck Institute</td>
<td><a href="http://www.biochem.mpg.de/5111795/maxquant">http://www.biochem.mpg.de/5111795/maxquant</a></td>
</tr>
</tbody>
</table>

2.1.9 Data acquisition and analysis

The resulting peptide spectra were searched against target databases (Table 2.1), or theoretical spectral libraries from the National Center for Biotechnology Information non-redundant (NCBIInr) and the Universal Protein Knowledgebase (UniProtKB) protein sequence databases using MASCOT or SEQUEST. Peptide spectrum match (PSM) validation was performed using decoy amino acid sequences assembled from the target database during each analysis. Search parameters set as follows: digestion with trypsin with a maximum of two missed cleavages, MS1=5 ppm, MS2=0.2 Da, carbamidomethylation of cysteine (57.02 Da) as a static modification and serine (Ser, S) and threonine (Thr, T) phosphorylation (79.97 Da, HPO₃; 97.99 Da, H₃PO₄) and methionine oxidation (15.99 Da) as variable modifications. Phosphorylations were determined using PhosphoRSv3.1. PhosphoRS site probabilities were set to a minimum of 0.75 to compensate for poor PSM scores and due to the high degree of automation of PTM prediction algorithms (Zhao and Jensen, 2009). Low scoring phosphorylation sites were only considered if the phosphorylation was repeatedly assigned in subsequent experiments. For statistical confidence, accepted spectra were searched against decoy amino acid sequences. The false discovery rate (FDR) was defined at two stringencies (1% and 5%), requiring a minimum of two peptides per protein match, filtered above 95% confidence. Candidate phosphopeptide significance was determined at $P \leq 0.05$ ($E$-value) in MASCOT and the cross-correlation value (Xcorr) in SEQUEST, where $Xcorr \geq 2.15$ is significant.
Table 2.2. Online protein sequence databases used for MS/MS analyses. To increase the number of peptide spectrum matches, observed peptide spectra were searched against several plant protein databases. Uncharacterised proteins were subsequently searched against green plant protein sequences using BLAST.

<table>
<thead>
<tr>
<th>Target Database</th>
<th>Online Directory</th>
<th>FASTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCBInr</td>
<td>ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nr.gz</td>
<td>All Green Plants Viridiplantae</td>
</tr>
<tr>
<td>UniProtKB/Swiss-Prot</td>
<td>ftp.uniprot.org/pub/databases/uniprot/uniref/uniref100</td>
<td>Setaria italica Zea mays Sorghum bicolor</td>
</tr>
<tr>
<td>UniProtKB/TrEMBL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Protein matches, which remained uncharacterised due to unreviewed protein accessions in UniProtKB, were searched against green plant protein sequences using the Basic Local Alignment Search Tool (BLAST). For consistency and due to uncharacterised matches from protein sequence repositories, along with the use of protein sequences from related species, differences in protein amino acid sequences of inferred proteins of multiple accession numbers and multiple protein isoforms, phosphorylation sites were annotated by the position of the modified residue in the phosphopeptide and not the amino acid number in the protein sequence, unless specifically indicated. UniProt accession numbers are indicated in the text (Table 3.3).

Full length protein sequences and structural information were obtained from the UniProtKB database and annotated using FASTAnnotate. Alignments were performed using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). Protein schematics were drawn to scale using Adobe Illustrator and annotated with experimental data obtained (Appendix A). Peptide spectrum annotation was performed using pFind Studio (Fu et al., 2004; Li et al., 2005; Wang et al., 2007). Protein structure information was acquired from the Protein Data Bank (https://www.rcsb.org/) and proteins were modelled using PyMol (https://pymol.org). Additional protein information was acquired from BRENDA (http://www.brenda-enzymes.org/). Monoisotopic masses were calculated using the Peptide Mass Calculator (http://www.peptidesynthetics.co.uk).
2.1.10 Phosphopeptide validation

Novel phosphorylation sites were validated using a modified in-gel tryptic digestion. Half-leaves from 2.5-week old *Setaria viridis* and *Sorghum bicolor* plants were harvested at 7.5 h into dark and *Megathyrsus maximus* leaves were harvested at 7.5 h into the dark and 4 h into the light. Leaves were ground to a fine powder in liquid nitrogen, then homogenised in 200 Bicine-KOH, pH 9.8 and 40 mM DTT containing EDTA-free protease and phosphatase inhibitors. Lysate was centrifuged at 4 °C and aliquots pipetted into 0.5 mL Eppendorf tubes. Protein concentrations were determined following Bradford (1976).

![Stained protein gel](image)

**Fig. 2.7** Stained protein gel containing four replicates of *Setaria viridis* (lane 1–4, V1–V4) and *Sorghum bicolor* (lane 6–9, G1–G4) protein extracts from 7.5 h darkened leaves. 50 µg of protein was loaded into each lane of a 4–12%, 1.5 mm Bis-Tris SDS gel (NuPAGE), adding 15 µL of protein ladder into lane 5. Lane 10 was left empty. Gel was stained with InstantBlue™ stain for 15 min. Two sections (A and B) of the gel were excised per lane, per replicate, along the dashed lines. Dark band indicated by the black arrow below 98 kDa in Section A across eight samples is a combination of PEPC (109 kDa native) and PPDK (108 kDa native).
Protein extracts were solubilised in 4× lithium dodecyl sulfate (LDS) buffer and reduced using 10 mM DTT at 70 °C for 15 min then alkylated with 55 mM IAA for 15 min with dark incubation. 150 µg of protein was resolved using a 4–12% NuPAGE Bis-Tris 1.5 mm gel and subsequently stained with Coomassie dye, InstantBlue™ stain for 15 min. After staining, gel pieces were excised from two regions of the gel, between 110 kDa and 50 kDa, as indicated on Figure 2.7. Gel pieces were destained at room temperature and incubated overnight with 0.15 µg of trypsin at 37 °C. The following day, peptides were extracted in ACN and 5% formic acid by gentle agitation at room temperature, dried down using a SpeedVac at 45 °C and stored at −20 °C before use.
2.2 Protein Isoform and Phosphoproteome Analysis

**Fig. 2.8.** *Setaria viridis* sample preparation workflow for protein isoform analysis. *Setaria viridis* plants were grown in a controlled environment and harvested at four time-points by flash-freezing in liquid nitrogen. Proteins were extracted using a urea-thioura based buffer. Gels were either stained for whole protein or phosphoprotein, and the former subjected to analysis by mass spectrometry.
2.2.1 Sample preparation

To analyse protein isoforms by 2-dimensional (2D) SDS-PAGE, 350 mg of plant tissue from four specific time-points (2 h and 7.5 h into dark; 4 h and 15.5 h into light) was homogenised in a concentrated lysis buffer containing 8.2 M urea, 2.3 M thiourea, 35 mM Tris-base, 4.7% 3-[3-cholamidopropyl]dimethylammonio]-1-propanesulfonate (CHAPS), 1 μM pepstatin, 1× nuclease mix, EDTA-free protease and PhosSTOP phosphatase inhibitors (Roche, Mannheim, Germany). Crude extracts were transferred into Eppendorf tubes, incubated on ice for 30 min with agitation every ten min for the complete removal of nucleic acids, and centrifuged at 4 °C at 16,800 × g until liquid fraction appeared clear. The supernatant was collected, and proteins were quantified using the 2D Quant Assay Kit (GE Healthcare, Buckinghamshire).

2.2.2 Two-dimensional SDS-PAGE

Protein extracts were subjected to analysis by two-dimensional (2D) gel electrophoresis on 7 and 24 cm resolving gels using a 29:1 acrylamide:bisacrylamide solution containing 1.5 M Tris-HCl pH 8.8, 10% SDS, 0.5% (v/v) ammonium persulfate solution (APS) and 0.05% (v/v) tetramethylethylenediamine (TEMED). Proteins in native conditions were solubilised in rehydration buffer (lysis buffer, BPB and 1.6% (v/v) DeStreak) containing 1.6% (v/v) immobilized pH gradient (IPG) buffer (pH 4–7) and 20 mM DTT, vortexed and centrifuged at 13,300 × g for 5 min. 80 or 450 μg of protein for 7 or 24 cm 10% SDS gels, respectively, was fixed onto a pH 4–7 Immobiline DryStrip Gel (referred hereafter as IPG strip) (GE Healthcare, Buckinghamshire, UK) by passive rehydration. Isoelectric focusing (IEF) for 7 cm IPG strips was performed using a 12.5 h protocol (Table 2.3) at 300–2000 volts (V), 50 μA and 24 cm IPG strips were focused using a protocol at 750–10,000 V over the course of 15 h (Table 2.4).
Table 2.3. Programme for 7 cm IEF. Small format protocol used for isoelectric focusing of *Setaria viridis* proteins using 7 cm 10% SDS gels.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Voltage (V)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gradual voltage increase</td>
<td>300V</td>
<td>1h30</td>
</tr>
<tr>
<td>2</td>
<td>Gradual voltage increase</td>
<td>500V</td>
<td>3h</td>
</tr>
<tr>
<td>3</td>
<td>Voltage hold</td>
<td>500V</td>
<td>0h30</td>
</tr>
<tr>
<td>4</td>
<td>Gradual voltage increase</td>
<td>2000V</td>
<td>1h</td>
</tr>
<tr>
<td>5</td>
<td>Voltage hold</td>
<td>2000V</td>
<td>6h30</td>
</tr>
</tbody>
</table>

Table 2.4. Programme for 24 cm IEF. Large format protocol used for isoelectric focusing of *Setaria viridis* proteins using 24 cm 10% SDS gels. 24 cm gels were used for subsequent spot analysis and protein identification using tandem mass spectrometry.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Voltage (V)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gradual voltage increase</td>
<td>750</td>
<td>1h</td>
</tr>
<tr>
<td>2</td>
<td>Gradual voltage increase</td>
<td>2000</td>
<td>5h</td>
</tr>
<tr>
<td>3</td>
<td>Voltage hold</td>
<td>2000</td>
<td>2h</td>
</tr>
<tr>
<td>4</td>
<td>Gradual voltage increase</td>
<td>8000</td>
<td>2h</td>
</tr>
<tr>
<td>5</td>
<td>Voltage hold</td>
<td>8000</td>
<td>3h</td>
</tr>
<tr>
<td>6</td>
<td>Gradual voltage increase</td>
<td>10000</td>
<td>0h30</td>
</tr>
<tr>
<td>7</td>
<td>Voltage hold</td>
<td>10000</td>
<td>1h30</td>
</tr>
</tbody>
</table>

After IEF, IPG strips were immediately stored in plastic containers at –20 °C. Before the second dimension, IPG strips were immersed in 1× TGS containing 1% (w/v) DTT for 15 min then alkylated with 2.5% (w/v) IAA. Strips were then positioned on 7 or 24 cm SDS gels and stabilised using 1% (w/v) agarose. 7 cm gels were run at 25 V for 15 min, then at 200 V until bands migrated to the bottom of the gel. 24 cm gels were run at 600 V for 6.25 h with a maximum current of 10 μA for the first 3 h, then 40 μA per gel, while maintaining 20 °C.
2.2.3 ProQ Diamond phosphoprotein staining

For phosphoprotein profiling, 7 cm SDS-PAGE gels were fixed in 50% (v/v) methanol and 10% (v/v) acetic acid, then left to stain in the dark on a rocker for 2 h in 65 mL 3-fold diluted ProQ Diamond phosphoprotein gel stain (ThermoFisher Scientific, Wetham, USA), following Agrawal and Thelen (2009). After staining, gels were destained in 20% ACN and 5% sodium acetate, pH 4.0. ProQ Diamond stained gels were visualised with a FLA-5100 (FUJIFILM Life Science, Stamford, USA) using a 532 nm laser at 740 V intensity. Phosphoprotein signal was tuned against a PeppermintStick™ phosphoprotein standard (ThermoFisher Scientific, Weltham, USA).

2.2.4 Coomassie staining

7 and 24 cm 2D gels were stained using Coomassie Simply Blue (ThermoFisher Scientific, Wetham, USA) with gentle agitation at 4 °C. Stained gels were washed twice with ultra-pure water and imaged using the FLA-5100 at 500 V with 635 nm laser excitation.
2.3 Determining Enzyme Kinetics

Fig. 2.9. Schematic showing the workflow for protein assays using Setaria viridis, Sorghum bicolor and Zea mays leaf proteins. Plants were grown in four sets of 20 replicates in a controlled chamber under ambient CO$_2$ (400 ppm), long days at 28 °C and 900 µmol m$^{-2}$ s$^{-1}$ PPFD. Leaf tips of 2.5-week old plants were harvested by flash-freezing in liquid nitrogen. Proteins were extracted, quantified and normalised prior to measurement.
2.3.1 Preparation of leaf homogenates for enzyme measurement

To determine the enzymatic properties of PEPC, NADP-malic enzyme, aspartate aminotransferase and alanine aminotransferase in response to the dark to light transition, two 16 mm × 8 mm sections (Figure 2.10) were cut no more than 2 cm from each leaf tip of mature Setaria viridis and Sorghum bicolor plants at four time-points (2 h and 7.5 h into dark; 4 h and 15.5 h into light) and Zea mays at two time-points (7.5 h into dark and 15.5 h into light), placed into a microtube and flash-frozen in liquid nitrogen. Leaf proteins were extracted from the leaf section closer to the leaf tip, as detailed in Chapter 2, Methods Part I, using 1 mL of 200 mM Bicine-KOH, pH 9.8, including protease inhibitors. Crude extracts were cleared by centrifugation in Eppendorf tubes at 16,800 × g at 4 °C, then 250 µL aliquots were stored at −80 °C.

![Figure 2.10](image)

**Fig. 2.10.** 2.5-week old *Setaria viridis* (I) and *Sorghum bicolor* (II) leaves photographed before harvests. Proteins for enzyme assays were extracted from leaf section A. Chlorophyll content was measured in the adjacent leaf section (B). The 16 mm × 8 mm leaf sections were accurately cut using a metal template.
2.3.2 Protein quantification

Protein concentrations were determined for leaf homogenates spectrophotometrically at 595 nm using a BSA standard curve (Figure 2.11) following Bradford (1976).

![Standard Curve for BSA](image)

**Fig. 2.11. The BSA standard curve that was used to determine protein concentration in leaf homogenates.** Absorbance of bovine serum albumin standards (0–2000 µg), in solution, was determined spectrophotometrically. Linear regression analysis was performed, and the equation was used to calculate total protein concentration in leaf lysates.

Prior to each assay, crude extracts were normalised against the lowest protein concentration using equation 1,

\[
\left( \frac{A_n}{A_x} - 1 \right) \times U_i = U_F
\]

(1)

where \(A_n\) is the absorbance of a sample, \(A_x\) is the absorbance of the least concentrated sample, \(U_i\) is the initial crude extract volume (mL) and \(U_F\) is the volume (mL) of buffer to add each extract to normalise them against the least concentrated extract.
2.3.3 Measurement of enzymes

Rapid-assays are summarised in Table 2.8. Each enzyme assay consisted of six biological replicates across ten substrate concentrations (5000 range fold), each including three technical replicates on a 96-well microtitre plate (Figure 2.12). Reactions were initiated by the addition of 5 µL of leaf lysate (150 µL final volume) and monitored over 15 cycles (20 min) or until reaction reached saturation. For each assay, no leaf protein and no substrate controls were used to subtract non-enzyme mediated reaction background signals. Reaction rate given in terms of the amount of NADP (Figure 2.13A) or NADPH (Figure 2.13B) formed, per amount of protein used to initiate the reaction, per unit time.

Fig. 2.12. Example of a 96-well microtitre plate assay design used for determining Michaelis-Menten kinetics. Each plate contained three biological replicates (A, B and C), no substrate (=) and designated controls (×), across ten substrate concentrations to determine $K_M$ constants. Reactions were initiated by adding 5 µL of normalised protein extract to 145 µL of mastermix, giving a final volume of 150 µL and monitored spectrophotometrically over time in a continuous assay at 25 °C.
Fig. 2.13. NADH (A) and NADPH (B) standard curves for determining enzyme reaction rates. The assay was performed using serial dilutions of a freshly made NADH or NADPH stock solution. Linear regression analysis was performed, and the goodness of fit was determined. NADH curves were used for PEPC, PEPCK and aminotransferase assays. NADPH used for the NADP-malic enzyme assays.
2.3.4 Phosphoenolpyruvate carboxylase assay

Fig. 2.14. Reaction mechanism showing the carboxylation of phosphoenolpyruvate (PEP) to oxaloacetate (OAA) catalysed by PEPC in C₄ photosynthesis. Highlighted region in blue box indicates the direction of the *in vitro* reaction, monitoring the reduction of OAA to malate via MDH. The number of carbons is indicated below each metabolite. CA, carbonic anhydrase; PPDK, pyruvate phosphate dikinase; Pᵢ, phosphate.

Phosphoenolpyruvate carboxylase (PEPC; EC 4.1.1.31) was assayed at 340 nm, 25 °C, following the reduction of oxaloacetate (OAA) by nicotinamide adenine dinucleotide (NADH) (Figure 2.14). 150 μL reaction mixtures contained 97.32 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid-potassium hydroxide (HEPES-KOH), pH 7.5, 10 mM NaHCO₃, 5 mM glucose-6-phosphate (G6P), 5 mM MgCl₂, 0.2 mM NADH and 6 units of malate dehydrogenase (MDH) (Table 2.5). Absorbance signals of reaction mixtures containing no leaf protein were used to correct for non-enzymatic oxidation of NADH.

**Table 2.5. PEPC mastermix.**

<table>
<thead>
<tr>
<th>Stock Concentration</th>
<th>Reagent</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 M</td>
<td>HEPES-KOH, pH 7.5</td>
<td>97.32 mM</td>
</tr>
<tr>
<td>1 M</td>
<td>NaHCO₃</td>
<td>10 mM</td>
</tr>
<tr>
<td>1 M</td>
<td>G6P</td>
<td>5 mM</td>
</tr>
<tr>
<td>0.5 M</td>
<td>MgCl₂</td>
<td>5 mM</td>
</tr>
<tr>
<td>0.25 M</td>
<td>NADH</td>
<td>0.2 mM</td>
</tr>
<tr>
<td>1200 U mg⁻¹ (5 mg mL⁻¹)</td>
<td>MDH</td>
<td>6 U mL⁻¹</td>
</tr>
</tbody>
</table>
2.3.5 NADP-malic enzyme assay

**Fig. 2.15. Reaction mechanism showing the decarboxylation of L-malate by NADP-malic enzyme (ME) in bundle sheath chloroplasts of C₄ plants.** The *in vitro* assay monitors the formation of NADPH (in red) following the decarboxylation of L-malate by the NADP-dependent malic enzyme. In bundle sheath chloroplasts of C₄ plants, malate is oxidised to pyruvate (pyr), forming NADPH and CO₂. CO₂ then enters the PCR cycle. The number of carbons is indicated below each metabolite. RuBP, ribulose-1,5-bisphosphate; 3-PGA, 3-phosphoglycerate.

NADP-dependent malic enzyme (NADP-ME; EC 1.1.1.40) was spectrophotometrically assayed in the forward direction (decarboxylating), at 340 nm at 25 °C, following the reduction of NADP⁺ (Figure 2.15). The reaction mixture contained 98.90 mM Tris-HCl, pH 8.0, 5 mM MgCl₂ and 0.5 mM nicotinamide adenine dinucleotide phosphate (NADP⁺) (Table 2.6). The pH and concentration of MgCl₂, required to achieve optimal rates of reaction, were determined for NADP-malic enzyme from darkened and illuminated leaves.

**Table 2.6. NADP-malic enzyme mastermix.**

<table>
<thead>
<tr>
<th>Stock Concentration</th>
<th>Reagent</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 M</td>
<td>Tris-HCl, pH 8.0</td>
<td>98.90 mM</td>
</tr>
<tr>
<td>0.5 M</td>
<td>MgCl₂</td>
<td>5 mM</td>
</tr>
<tr>
<td>0.5 M</td>
<td>NADP⁺</td>
<td>0.5 mM</td>
</tr>
</tbody>
</table>
2.3.6 Aspartate and alanine aminotransferase assay

Fig. 2.16. Reaction mechanism for the aspartate (Asp) and alanine (Ala) aminotransferase assays. Aminotransferases were assayed by coupling to malate dehydrogenase (MDH) or lactate dehydrogenase (LDH). 2-OG, 2-oxoglutarate; OAA, oxaloacetate.

Aspartate aminotransferase (EC 2.6.1.1) was assayed spectrophotometrically at 340 nm by coupling to MDH (Table 2.7), following the oxidation of NADH to NAD⁺ (Figure 2.16) at pH 8.0. The range of L-aspartate and 2-oxoglutarate concentrations used in this assay was 0-25000 µM and 0–2000 µM, respectively. The activity of aspartate aminotransferase was assayed in the presence of L-malate, L-alanine, 3-PGA, PEP, pyruvate and dimethyl-2-oxoglutarate, an analogue of 2-oxoglutarate.

Alanine aminotransferase (EC 2.6.1.2) was assayed in the presence of L-alanine towards the formation of lactate at pH 7.5, by coupling to LDH. L-Alanine and L-aspartate were used as the negative control for aspartate and alanine aminotransferase, respectively.

Table 2.7. Aminotransferase mastermix. Mastermix used for aspartate and alanine aminotransferase assays. † Indicates alanine aminotransferase reagents.

<table>
<thead>
<tr>
<th>Stock Concentration</th>
<th>Reagent</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 M</td>
<td>Tris-HCl, pH 8.0 or 7.5†</td>
<td>50 mM</td>
</tr>
<tr>
<td>50 mg mL⁻¹</td>
<td>Pyridoxal phosphate</td>
<td>10 µg mL⁻¹</td>
</tr>
<tr>
<td>0.5 M</td>
<td>EDTA</td>
<td>2 mM</td>
</tr>
<tr>
<td>0.5 M</td>
<td>NADH</td>
<td>0.2 mM</td>
</tr>
<tr>
<td>6000 U mg⁻¹ (1.0 mg mL⁻¹)</td>
<td>MDH</td>
<td>4 U mL⁻¹</td>
</tr>
<tr>
<td>1000 U mg⁻¹ (9.3 mg mL⁻¹)</td>
<td>LDH†</td>
<td>4 U mL⁻¹</td>
</tr>
</tbody>
</table>
Table 2.8. Summary of enzyme assays.

<table>
<thead>
<tr>
<th>Assay</th>
<th>Wavelength</th>
<th>Spectroscopy</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEPC</td>
<td>340</td>
<td>Measures the concentration of NADH. Abs should <em>decrease</em> over time due to oxidation to NAD(^+).</td>
</tr>
<tr>
<td>PEPCK</td>
<td>340</td>
<td>Measures the concentration of NADH. Abs should <em>decrease</em> over time due to oxidation to NAD(^+).</td>
</tr>
<tr>
<td>NADP-ME</td>
<td>340</td>
<td>Measures the concentration of NADPH. Abs should <em>increase</em> over time due to NADP(^+) reduction to NADPH.</td>
</tr>
<tr>
<td>AspAT</td>
<td>340</td>
<td>Measures the concentration of NADH. Abs should <em>decrease</em> over time due to oxidation to NAD(^+).</td>
</tr>
<tr>
<td>AlaAT</td>
<td>340</td>
<td>Measures the concentration of NADH. Abs should <em>decrease</em> over time due to oxidation to NAD(^+).</td>
</tr>
<tr>
<td>Bradford</td>
<td>595</td>
<td>Measures the concentration of total protein present in the crude extract (Beer-Lambert Law).</td>
</tr>
<tr>
<td>Chlorophyll</td>
<td>652 (Sample) 750 (Blank)</td>
<td>Measures the concentration of chlorophyll (a) (mg L(^{-1})) in 80% ethanol or 80% chilled acetone (Appendix B).</td>
</tr>
</tbody>
</table>
2.3.7 Determining Michaelis-Menten kinetics

Raw data was analysed in Microsoft Excel and enzyme reaction rates were graphed using Prism 7 (GraphPad Software) and the Michaelis-Menten constants ($K_M$) were determined using a curve fitting algorithm in GraphPad, using the Michaelis-Menten model (2),

$$V_0 = V_{\text{max}} \cdot \frac{[S]}{[S] + K_M} \quad (2)$$

where $V_0$ is equal to the velocity of the reaction, $V_{\text{max}}$ is the maximum reaction rate, $[S]$ is the substrate concentration and $K_M$ is the Michaelis-Menten kinetics constant. Initial velocity is given as an arbitrary unit of absorbance (Abs) change over time (min).

**NB.** $V_{\text{max}}$ cannot be reported because enzyme assays were conducted using enzyme in crude leaf lysates and not purified protein. Also, for simplicity, $K_M$ will be referred to as affinity. However, it should be noted that $K_M$ and affinity are not interchangeable due to limitations of the Michaelis-Menten model. To properly discuss substrate affinity for an enzyme, the structural properties and binding interactions must be known.

Affinity fold change (i.e. increase or decrease) was calculated as a ratio of two $K_M$ values using equation (3),

$$\text{Fold change} = \frac{K_M^2}{K_M^1} - 1 \quad (3)$$

where $K_M^2$ is the new value (in response to a variable, for example light) and $K_M^1$ is the original value. An increase in $K_M$ corresponds to a decrease in affinity (positive fold change) and decrease in $K_M$ corresponds to an increase in affinity (negative fold change).
Chapter 3 – Identification of novel light-dependent phosphorylation sites of C₄-related enzymes

3.1 Introduction

C₄ photosynthesis has evolved independently in over 60 plant lineages from the classical C₃ pathway, involving the gradual transition to Kranz anatomy and recruitment of C₃ genes for C₄ function (Sage et al., 2012). The genetic mechanisms that underpinned the regulation of mesophyll- or bundle sheath-cell specific gene expression evolved in parallel across multiple C₄ origins (Sinha and Kellogg, 1996; Brown et al., 2011; Williams et al., 2013). C₄ plant lineages have also shared mechanisms for determining C₄-specific enzyme function through changes to amino acid sequences, though the degree of parallelism depends on the selective pressures acting on specific residues. For instance, the acquisition of the C₄-specific PEPC, through the evolution from the ancestral non-C₄ PEPC, occurred repeatedly across several independent C₄ lineages (Christin et al., 2007). Furthermore, Christin et al. (2007), also showed that particular PEPC codon mutations, at positions 517, 577, 579, 761 and 780 for serine and threonine amino acids, were under positive selection in independent C₄ lineages and may be determinants of C₄ characteristics. These amino acid substitutions, namely the A780S transition in Zea mays, have been shown to be contenders associated with the regulatory properties and kinetics of C₄-specific PEPCs (Dong et al., 1998; Bläsing et al., 2000; Svensson et al., 2003). The amino acids that evolved in parallel across independent lineages may be indications of the adaptations that were advantageous in determining C₄ function (Williams et al., 2013), but the significance of these residue mutations, with regards to the phosphoregulatory mechanisms underlying the kinetic properties of C₄-related enzymes, have not been studied in detail.

In plants, serine and threonine residues are frequently phosphorylated (Friso and van Wijk, 2015) and are prominent sites of regulation for C₄-related enzymes like PEPC, PPDK and PEPCK. These mechanisms of regulation are often complex, and may be dependent on illumination or circadian controllers (Jiao and Chollet, 1988; Jiao et al., 1991; Wilkins, 1992; Nimmo, 1998; Hartwell et al., 1999). The light-dependent phosphorylation of PEPC is well documented at the invariant N-terminal serine across
grass species (Jiao and Chollet, 1988; Jiao and Chollet, 1991), however despite evidence of the selection pressures of serine and threonine residues in C₄ lineages presented by Bläsing et al. (2000) and Christin et al. (2007), little is known about the regulatory characteristics that span across the C₄-specific PEPC sequence. Furthermore, the phosphoregulatory properties of other essential C₄-related proteins that evolved C₄-specific function, like aspartate and alanine aminotransferase (Hatch and Mau, 1973) or NADP-malic enzyme (Christin et al., 2009; Saigo et al., 2013), have not been described in C₄ plants.

3.1.1 The role and diversity of post-translational modifications in plants

Post-translational modifications (PTMs) play a pivotal role in the regulation of proteins and contribute to the functionality of the proteome (Figure 3.1). The diversity of PTMs, together with their reversible, dynamic nature, contribute to protein function, which is essential for the regulation of metabolic pathways, protein-protein interactions or activation of signal transduction pathways (Karve and Cheema, 2011). However, these tightly regulated, dynamic mechanisms are often overshadowed by spontaneous enzymatic reactions induced by reactive species such as free radicals or redox potentials, which non-specifically modify amino acids and result in irreversible changes to protein structure and function (Friso and van Wijk, 2015).

![Fig. 3.1. Schematic showing the role of PTMs on the functional complexity of the proteome. Expressed genes undergo splicing mechanisms creating polyadenylated gene transcripts (coloured lines, –AAA). Proteins shown as multi-coloured swirls.](image-url)
Table 3.1. **Reversible post-translational modifications that regulate the activity of plant proteins.** Common PTMs found in plant cells, structural characteristics and their general role in biology. Binding interactions between the amino acid and the PTM are highlighted in orange. Adapted from Friso and van Wijk (2015).

<table>
<thead>
<tr>
<th>PTM observed in vivo</th>
<th>Modifiable amino acid</th>
<th>Monoisotopic mass (Da)</th>
<th>General role in biology</th>
<th>Structure</th>
<th>Binding notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylation</td>
<td>Cysteine, Lysine</td>
<td>42.01056</td>
<td>Protein localisation</td>
<td><img src="image" alt="Acetyl group" /></td>
<td>An acetyl group (CH$_3$–CO) binds to the N-terminal (NH$_3^+$) of an amino acid.</td>
</tr>
<tr>
<td>Methylation</td>
<td>Lysine, Cysteine, Arginine, Glutamic acid</td>
<td>14.01565</td>
<td>Transcriptional regulation</td>
<td><img src="image" alt="Methyl group" /></td>
<td>A methyl group (CH$_3$) reacts with side-chains or C-terminals of target amino acids.</td>
</tr>
<tr>
<td>Oxidation</td>
<td>Methionine</td>
<td>15.99491</td>
<td>Protein stability and interactions</td>
<td><img src="image" alt="O-S bond" /></td>
<td>An oxygen atom binds to the sulphur atom in methionine.</td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>Serine, Threonine, Tyrosine, Histidine (rare)</td>
<td>79.96633</td>
<td>Signal transduction</td>
<td><img src="image" alt="Phosphate group" /></td>
<td>A phosphate group (PO$_4^{3-}$) reacts with a side chain hydroxyl group (–OH) of a serine, threonine or tyrosine residue (Figure 3.2).</td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>Lysine</td>
<td>114.04292</td>
<td>Development</td>
<td><img src="image" alt="Ubiquitin structure" /></td>
<td>Ubiquitinations are identified by a characteristic diglycine substituent conjugated with a lysine residue. Ubiquitin protein binds to a lysine residue forming an isopeptide bond between lysine and a glycine residue on ubiquitin.</td>
</tr>
</tbody>
</table>
In plants, proteins may be phosphorylated, ubiquitinated, acetylated or methylated (Table 3.1, 3.2) by enzymatic reactions mediated by regulatory enzymes (Friso and van Wijk, 2015). In the event of phosphorylation (Figure 3.2) protein kinases transfer the γ-phosphate of ATP onto tyrosine, threonine or serine residues (Hunter, 2007). Histidine and aspartate residues can also undergo phosphorylation, but at low frequencies (Friso and van Wijk, 2015).

![Fig. 3.2. In vivo reversible phosphorylation mechanism of serine, threonine and tyrosine residues.](image)

Protein kinases and phosphorylases regulate the reversible phosphorylation of serine, threonine and tyrosine residues (only functional groups shown), whereby phosphate groups are linked by phosphoester bonds (bottom) via an oxygen atom (in orange), resulting in an 80 Da mass shift in the target protein (top). Arrows show the direction of phosphorylation by protein kinases and dephosphorylation by phosphorylases.
Reversible modifications, such as phosphorylation, typically regulate protein activity through the diurnal cycle, but PTMs in plants like acetylation, play another crucial role in determining protein localisation for organelle-specific reactions. Carbon fixation in plants depends on cytosolic protein activity as well as protein activity in subcellular compartments such as chloroplasts, mitochondria and peroxisomes (Hodges et al., 2013). However, the PTM diversity of plastidic proteins and how they regulate plant metabolism is unclear (Lehtimäki et al., 2015). Additionally, proteins may undergo non-specific modifications, which may affect other PTMs on the same protein or nearby proteins. A single PTM can serve as a signal promoting the binding of regulator proteins that mediate subsequent specific or non-specific modifications (Hunter, 2007). Direct competition between PTMs can result in the blocking of amino acid side chains preventing site-specific modifications. For instance, cysteine residues are essential for plant development, immunity, pathogen defence, protein stability and enzyme activity (Kim et al., 2015) but are prone to non-specific PTMs typically by redox regulation (Michelet et al., 2013), due to a highly reactive thiol group in the side-chain. Under specific conditions, such as oxidative stress, cysteine residues are susceptible to spontaneous, non-enzymatic modifications, which block sites of regulation (Marino and Gladyshev, 2012; Chung et al., 2013; Bhattacharjee et al., 2015; Friso and van Wijk, 2015).

Enzymes involved in carbon fixation, such as Rubisco, PEPCK and NADP-malate dehydrogenase (MDH), as well as enzymes in the PCR cycle, are regulated by redox mechanisms (Table 3.2) (Ashton and Hatch, 1983; Drincovich and Andreo, 1994; Raines et al., 2000; Schürmann and Buchanan, 2008; Michelet et al., 2013; Gütle et al., 2016). These regulatory mechanisms are widespread in plants and other biological systems. For instance, in Mycobacterium tuberculosis, PEPCK preferentially catalyses the conversion of oxaloacetate to PEP, but under hypoxic and growth-limiting conditions, the anaplerotic reaction (in the reverse direction) towards the formation of oxaloacetate is favoured (Machová et al., 2014). These changes are controlled by the reduction of disulphide bridges forming between cysteine residues, either by maintaining reducing intracellular environments, or interaction with reducing agents, such as thioredoxin (Machová et al., 2014). It was shown that the formation of disulphide bridges in PEPCK resulted in the loss of enzyme activity (Carlson et al., 1978; Krautwurst et al., 1995), but such modifications can have more intrinsic effects on tertiary protein structures. It was recently shown that the interaction between Cys-391 and Cys-397, via a disulphide
bridge, induced changes to the structure and function of PEPCK from *Mycobacterium tuberculosis*, whereas the reduced form of these residues stabilised PEPCK and influenced the anaplerotic function (Machová et al., 2017).

In C₄ plants, NADP-MDH is activated in illuminated and reducing conditions. In *Zea mays*, NADP-MDH is inactive in darkened leaves, but when transferred from darkness to illumination, activity can be restored (Johnson and Hatch, 1970). In addition, when NADP-MDH from darkened leaves was subjected to 5 mM DTT, enzymatic activity was fully restored, suggesting that the reduced form of the enzyme was essential for catalysis (Johnson and Hatch, 1970). Furthermore, NADP-MDH was inactivated when extracted in the absence of thiol-reducing agents such as DTT or thioredoxin (Jacquot et al., 1981; Ashton and Hatch, 1983). The light activation of NADP-MDH is regulated by the reduction of a disulphide bridge occurring between Cys-10 and Cys-15 within the N-terminal of NADP-MDH (Decottignies et al., 1988).

### 3.1.2 Phosphoenolpyruvate carboxylase

PEPC (EC 4.1.1.31) plays a pivotal role in C₄ photosynthesis, catalysing the irreversible β-carboxylation of PEP to oxaloacetate (OAA) by utilising bicarbonate (HCO₃⁻) (Figure 3.3) in mesophyll cells of C₄ plants (Chollet et al., 1996; Cousins et al., 2007) and is by far one of the best characterised enzymes of the C₄ pathway. The active form of PEPC consists of four identical 109 kDa subunits (Hatch, 1978; Matsumura et al., 2002) and *in vivo* catalysis is dependent on a divalent cation, usually Mg²⁺, though Mn²⁺ and Co²⁺ can be replaced *in vitro* (O'Leary et al., 1981).

![Fig. 3.3. Irreversible β-carboxylation of PEP to OAA by PEPC.](image)

In C₄ plants, PEPC is activated by glucose-6-phosphate and allosterically inhibited by malate and aspartate (Huber and Edwards, 1975; Andreo et al., 1987). Inhibition by malate is less pronounced when assayed at pH 8.0, rather than in physiological pH 7.3 and sensitivity to inhibition by malate is reduced with increasing concentration of Mg²⁺.
(Hatch, 1978; Wedding et al., 1990; Echevarría et al., 1994; Duff and Chollet, 1995). Sensitivity to feedback inhibition by malate or activation by glucose-6-phosphate is regulated by light-dependent reversible phosphorylation (Jiao and Chollet, 1991; Chollet et al., 1996). The underlying mechanism was first described by Nimmo et al. (1984), demonstrating that the PEPC from Crassulacean acid metabolism (CAM) species was more sensitive to inhibition by malate during illumination than in the dark period, when PEPC is active. Furthermore, when the phosphorylated night form was purified and dephosphorylated in vitro, the sensitivity to malate inhibition increased (Nimmo et al., 1986).

Using the $^{32}$P-phosphorylation system, the C$_4$-form of PEPC from Zea mays was shown to undergo phosphorylation in vitro (Budde and Chollet, 1986). Consistent with the CAM-form, the phosphorylation of the C$_4$-specific PEPC was determined to occur predominantly at a single serine residue within the N-terminal of PEPC from Zea mays (Ser-15) and Sorghum bicolor (Ser-8) during illumination, when PEPC from C$_4$ plants is active (Jiao and Chollet, 1988; Jiao and Chollet, 1990; Jiao et al., 1991). The relationship between phosphorylation and sensitivity to malate inhibition was substantiated in phosphomimetic mutants of the Sorghum bicolor PEPC expressed in Escherichia coli, which showed that substituting the Ser-8 with aspartate (S8D) resulted in reduced inhibition by malate (Wang et al., 1992), whereas substitution to cysteine (S8C) showed no reduced sensitivity (Duff et al., 1993). Also, while the phosphorylation of the C$_4$-specific PEPC decreases sensitivity to inhibition by malate during illumination and causes about a two-fold increase in $V_{\text{max}}$, it does not alter the its affinity for bicarbonate or PEP (Vidal and Chollet, 1997). The phosphorylation of PEPC is essential to maintain high nocturnal carboxylase activity in CAM plants (Boxall et al., 2017), while in Flaveria bidentis (C$_4$), phosphorylation of PEPC is not essential to maintain high CO$_2$ assimilation rates during illumination (Furumoto et al., 2007).
Fig. 3.4. Regulation of C₄-specific PEPC activity by in vivo serine phosphorylation.
Light-dependent PEPC phosphorylation (serine in red) is regulated by PEPC kinase (PK), which is more active in the light. PEPC kinase is regulated by cytosolic pH and Ca²⁺. Dephosphorylation (serine in blue) is mediated by protein phosphatase 2A (PP). Phosphorylated PEPC can maintain higher maximal activity due to a decrease in sensitivity to inhibition by malate. Figure adapted from Chollet et al. (1996).

PEPC is maximally phosphorylated within one hour of illumination by PEPC kinase (Vidal and Chollet, 1997; Bailey et al., 2007) and dephosphorylated by a type 2A protein phosphatase shortly before the dark period in C₄ plants (Carter et al., 1990; Vidal and Chollet, 1997; Dong et al., 2001). Another study showed that the phosphorylation of PEPC in Zea mays occurred before the onset of light and was dephosphorylated before the dark period (Ueno et al., 2000), suggesting that to some degree, C₄ photosynthesis is controlled by circadian mechanisms (Wang et al., 2011). In fact, the activity of PEPC is largely dependent on the light-dependent activation of PEPC kinase (Figure 3.4) rather than changes of protein phosphatase 2A activity (Echevarría et al., 1990; Carter et al., 1991; McNaughton et al., 1991). Moreover, two protein kinases can phosphorylate PEPC at the N-terminal serine residue, but only the Ca²⁺-dependent protein kinase shows light dependency (Jiao and Chollet, 1988; McNaughton et al., 1991; Li and Chollet, 1993; Wang and Chollet, 1993). In the C₄-form, PEPC kinase is regulated by metabolic factors, such as protein turnover, photosynthesis and pH (Jiao et al., 1991; Jiao and Chollet, 1992; Li and Chollet, 1993), whereas activation of the CAM-form PEPC kinase is controlled by a circadian oscillator, rather than light to dark transitions (Duff et al., 1996; Giglioli-
Guivarc'h *et al.*, 1996; Vidal and Chollet, 1997). At the protein level, the activity of PEPC kinase is also regulated by ubiquitination mediated degradation (Agetsuma *et al.*, 2005). In relation to PEPC, the PEPC kinase is also preferentially expressed in mesophyll cells in *Zea mays*, suggesting the existence of additional regulatory controls at the gene level (Li *et al.*, 2010).

Recent studies have suggested that the activation mechanism of the C₄-form PEPC is dependent on cross-talk between phosphorylation and lysine ubiquitination. Lysine ubiquitination has received attention in recent years, because, like phosphorylation, it is reversible, widespread in eukaryotic cells, occurs at relatively higher frequencies and mediated by a myriad of regulator enzymes (Hunter, 2007; Huber, 2011; Friso and van Wijk, 2015). Monoubiquitination was first described by Uhrig *et al.* (2008), demonstrating that the non-photosynthetic PEPC was modified at Lys-628 in germinating castor oil seeds. Ruiz-Ballesta *et al.* (2014) demonstrated that PEPC monoubiquitination occurred at Lys-624 in developing *Sorghum bicolor* seeds, and proposed a novel network of regulatory mechanisms contributing to PEPC activity. In addition to this, recent evidence suggests that the activation cascade of PEPC depends on several modifications, other than phosphorylation. This is supported in maturing phosphate-deficient roots of *Hakea prostrata*, where PEPC deubiquitination was followed by phosphorylation, promoting enzyme activity *in vivo* (Shane *et al.*, 2013).

### 3.1.3 Pyruvate, phosphate dikinase

PPDK (EC 2.7.9.1) catalyses the ATP-dependent regeneration of PEP from pyruvate in mesophyll chloroplasts of C₄ plants (Figure 3.5). Similar to PEPC, the active form of PPDK is a homotetramer made up of four 94 kDa subunits (Hatch, 1978).

\[
\text{Pyruvate} + \text{ATP} + P_i \xrightleftharpoons{\text{Mg}^{2+}} \text{PEP} + \text{AMP} + \text{PP}_i
\]

*Fig. 3.5. Regeneration of PEP from pyruvate by PPDK in C₄ plants.*

The reaction mechanism of the C₄-specific PPDK involves two steps. The reaction begins when PPDK binds two ATP and phosphate to form a phosphohistidine intermediate,
AMP and diphosphate, and is subsequently followed by a ping-pong mechanism for the addition of phosphate to pyruvate to form PEP (Hatch, 1978; Roeske et al., 1988). PPDK is inhibited by PEP, AMP and PPi, but these products are readily utilised by high activities of PEPC, adenylate kinase and pyrophosphatase in mesophyll cells, thus the C4-specific PPDK reaction is favoured towards the formation of PEP (Hatch and Slack, 1968; Hatch, 1978). Mg2+ is also essential for catalysis and the increasing concentration of Mg2+ in the stroma during illumination correlates to the light activation mechanism of PPDK (Leegood and Walker, 1999).

An early study demonstrated that PPDK was less active in darkened leaves than illuminated leaves of *Amaranthus palmeri* (Slack, 1968). This study also showed that the *in vitro* activity of PPDK declined after transferring light-grown sorghum and sugarcane plants to a period of darkness. In addition to this, PPDK is activated *in vitro* with treatment by a stromal regulatory protein and phosphate and deactivated by a regulatory protein and ADP (Budde et al., 1985). To determine the activation properties of PPDK from *Zea mays*, PPDK was assayed *in vitro* using radioactively labelled phosphate. In this approach, it was demonstrated that PPDK underwent phosphorylation at a threonine residue by an ADP-dependent regulatory protein, resulting in the deactivation of the enzyme (Ashton and Hatch, 1983). Furthermore, the activity of PPDK depends on the degree of phosphorylation. It was shown that one 94 kDa subunit from the inactive form of PPDK contained two-fold higher content of phosphate than one subunit from the active form (Budde et al., 1985).

In *Zea mays*, PPDK is inactivated in the dark by phosphorylation at the active site (Thr-456) and re-activated by dephosphorylation (Roeske et al., 1988). Both the phosphorylation and dephosphorylation of PPDK are uniquely regulated by a single bifunctional regulatory protein (PPDK-RP) (Figure 3.6) (Chastain et al., 1997; Chastain et al., 2000). Although PPDK-RP is active in mesophyll cell chloroplasts where PPDK is localised, transcripts of PPDK-RP were predominantly found in bundle sheath cells of *Zea mays* (Li et al., 2010). Also, despite its essential role in the regulation of PPDK, PPDK-RP appears to be in low abundance in mesophyll chloroplasts (Wang et al., 2011). However, being localised in the stroma of the chloroplast, PPDK-RP is regulated by light-dependent changes in ADP, which increases after the light to dark transition, resulting in higher PPDK-RP activity and PPDK phosphorylation in darkened leaves (Chastain and Chollet, 2003). Conversely, the phosphatase reaction of PPDK-RP is
favoured in illuminated conditions, when ADP is limiting, resulting in fewer copies of
PPDK being phosphorylated.

Fig. 3.6. Light activation mechanism of the C₄-specific PPDK by reversible
phosphorylation. Phosphorylation (in red) and dephosphorylation (in blue) is regulated
by a bifunctional ADP-dependent regulator protein (PR). PPDK is inactivated by
phosphorylation in the dark and activated by dephosphorylation. Figure adapted from
Chastain et al. (1997).

The catalytic effect of phosphothreonine-456 of PPDK from *Zea mays* was determined
by direct mutagenesis in *Escherichia coli*, using serine (also a target for
phosphorylation), valine and aspartate substitutions (Chastain et al., 1997). In this study,
it was shown that a threonine to serine substitution at position 456 (T456S) did not hinder
the phosphorylation of PPDK nor affected catalysis, but resulted in less effective enzyme
inactivation, whereas the substitution to aspartate (T456D), mimicking a non-labile
phosphorylation, resulted in complete enzyme inactivation (Chastain et al., 1997).
Chastain et al. (1997) also showed that the dephosphorylated form (T456V) remained
active. Further analysis also showed that a second phosphomimetic mutant (T456E)
abolishes catalytic activity, while substitutions to tyrosine (T456Y, another target for
phosphorylation) and phenylalanine (T456F) substantially decreases activity (Chastain
et al., 2000).

In a study to determine the tolerance of low temperatures on the photosynthetic activity
in *Zea mays* showed that the accumulation of PPDK in mesophyll chloroplasts decreased
by 50% when plants were grown in low temperature, though transcript amounts were
unchanged, suggesting that low temperature had little effect on gene expression (Naidu
et al., 2003). While there are no other known regulatory sites on the C₄-specific PPDK,
Wang et al. (2011) speculates that PPDK may be controlled by more complex
mechanisms, thus explaining the difference between gene expression and accumulation
of PPDK in mesophyll chloroplasts. In fact, it has been recently demonstrated that the PPDK regulatory mechanism at Thr-527 in Zea mays is strictly controlled by light intensity rather than by the dark to light transition (Chen et al., 2014). It is important, however, that inactivation by decreasing light intensity had been previously noted (Burnell et al., 1986), but had not been attributed to a single residue. In the recent study, it is suggested that increasing the expression of PPDK in mesophyll cells does not imply an increase enzymatic activity and there could be other regulatory controls or regulatory sites that are still unaccounted for (Chen et al., 2014).

### 3.1.4 Phosphoenolpyruvate carboxykinase

Compared to PEPC and PPDK, the light-dependent phosphorylation of PEPCK (4.1.1.49) is the most recent to be elucidated in C₄ plants, yet other C₄-related proteins have not been shown to be regulated by phosphorylation (Table 3.2). In PEPCK- and NAD-malic enzyme subtypes and certain NADP-malic enzyme C₄ plants, such as Zea mays (Walker et al., 1997), PEPCK catalyses the cytosolic ATP-dependent decarboxylation of oxaloacetate (OAA) in bundle sheath cells, forming CO₂ and PEP (Figure 3.7) (Hatch, 1978).

\[
\text{OAA} + \text{ATP} \xrightleftharpoons{\text{Mn}^{2+}} \text{PEP} + \text{ADP} + \text{CO}_2
\]

**Fig. 3.7. ATP-dependent decarboxylation of OAA by PEPCK in C₄ plants.**

Initial observations noted that the activity of PEPCK and accumulation in bundle sheath cells changed very little between light and dark conditions (Walker et al., 2002) and regulating the activity PEPCK would be essential to prevent depletion of ATP or oxaloacetate during the light to dark transition of C₄ photosynthesis (Carnal et al., 1993). The light-dependent phosphorylation of PEPCK was first described in cucumber cotyledons by feeding seedlings with radioactively labelled phosphate (Walker and Leegood, 1995). In this study, the native form of PEPCK was purified and shown to undergo phosphorylation by treatment with PEPC kinase from Zea mays. Also, the larger and native 74 kDa protein was phosphorylated, whereas the truncated form (62 kDa) was not phosphorylated, suggesting that the phosphorylation site was in the N-terminal
extension that had been proteolysed (Walker and Leegood, 1995). Walker and Leegood (1995) also showed that upon the removal of ATP in the reaction assay, the phosphorylation of PEPCK was reversible by treatment with protein phosphatase 2A.

In C₄ plants, such as the PEPCK-type *Megathyrsus maximus* (previously *Panicum maximum*), PEPCK is phosphorylated in darkened leaves and dephosphorylated in illuminated leaves, in a regulatory mechanism that leads to PEPCK activation in illuminated leaves (Walker and Leegood, 1996; Walker et al., 2002). The role of phosphorylation in this mechanism was studied by Bailey *et al.* (2007), with emphasis towards the coordination between carboxylation by PEPC and decarboxylation by PEPCK in darkened and illuminated leaves of *Megathyrsus maximus*. In this study, the authors showed that PEPCK activation at full sunlight correlated to a decrease in its phosphorylation state, whereas the phosphorylation state of PEPC during illumination correlated with the light-induced activation of PEPC (Bailey *et al.*, 2007). Recently, the phoshoregulatory site of PEPCK from *Zea mays* was determined by mass spectrometry. In *Zea mays*, PEPCK was shown to undergo phosphorylation at Ser-55, Thr-58, Thr-59 and Thr-120, though phosphorylation at Ser-55 showed a higher degree of light-dependency (Chao *et al.*, 2014). Furthermore, the study showed that in *Zea mays* seedlings, Ser-55 was phosphorylated in illuminated leaves and dephosphorylated in darkened leaves, which is the opposite of what had previously been reported (Walker and Leegood, 1995; Walker and Leegood, 1996; Bailey *et al.*, 2007). Differing from previous findings, the activity of PEPCK was shown to be lower in illuminated leaves and higher in darkened leaves (Chao *et al.*, 2014).
Table 3.2. Summary of known post-translational modifications of plant proteins involved in carbon fixation. For each protein, the type of PTM is indicated (p, phosphorylation; ub, ubiquitination) and relevant references are cited.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Biological process</th>
<th>Post-translational regulation</th>
<th>Modifications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphoenolpyruvate carboxylase EC 4.1.1.31</td>
<td>C₄ photosynthesis</td>
<td>Light-dependent reversible phosphorylation decreases sensitivity to malate inhibition Mono-ubiquitination cross-talk with phosphorylation</td>
<td>pSer-15 ub-Lys-624</td>
<td>(Ruiz-Ballesta et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Shane et al., 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Uhrig et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Jiao and Chollet, 1991)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Huber and Edwards, 1975)</td>
</tr>
<tr>
<td>Phosphoenolpyruvate carboxylase kinase EC 2.7.11.1</td>
<td>C₄ photosynthesis</td>
<td>Ubiquitination mediated degradation</td>
<td></td>
<td>(Agetsuma et al., 2005)</td>
</tr>
<tr>
<td>Glycerate kinase EC 2.7.1.31</td>
<td>Photorespiration</td>
<td>Thioredoxin regulated disulphide bridge (S-S)</td>
<td></td>
<td>(Bartsch et al., 2010)</td>
</tr>
<tr>
<td>Pyruvate, phosphatase dikinase EC 2.7.9.1</td>
<td>C₄ photosynthesis</td>
<td>Light-intensity mediated reversible phosphorylation by PPDK-RP kinase-phosphatase</td>
<td>pThr-527 pSer-528</td>
<td>(Chen et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Chastain et al., 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Roeske et al., 1988)</td>
</tr>
<tr>
<td>NADP-malic enzyme EC 1.1.1.40</td>
<td>C₄ photosynthesis</td>
<td>None documented in C₄ plants</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>NADP-malate dehydrogenase EC 1.1.1.82</td>
<td>C₄ photosynthesis</td>
<td>Thioredoxin regulated disulphide bridge (S-S)</td>
<td>67 ↔ 72</td>
<td>(Decottignies et al., 1988)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Leegood and Walker, 1983)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Jacquot et al., 1981)</td>
</tr>
<tr>
<td>Phosphoenolpyruvate carboxykinase EC 4.1.1.49</td>
<td>C₄ photosynthesis</td>
<td>Light-dependent reversible phosphorylation</td>
<td>pSer-55</td>
<td>(Chao et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Bailey et al., 2007)</td>
</tr>
<tr>
<td>Protein</td>
<td>Biological process</td>
<td>Post-translational regulation</td>
<td>Modifications</td>
<td>References</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>---------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>Carbonic anhydrase EC 4.2.1.1</td>
<td>C₄ photosynthesis</td>
<td>None documented in the C₄ photosynthesis protein</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Alanine aminotransferase</td>
<td>C₄ photosynthesis</td>
<td>None documented in the C₄ photosynthesis protein</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Aspartate aminotransferase EC 2.6.1.1</td>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Pyruvate dehydrogenase EC 1.2.4.1</td>
<td>Respiration</td>
<td>Reversible phosphorylation</td>
<td>(Tovar-Méndez et al., 2003) (Thelen et al., 2000)</td>
<td></td>
</tr>
<tr>
<td>Sucrose phosphate synthase EC 2.4.1.14</td>
<td>Sucrose metabolism</td>
<td>Reversible light-regulated phosphorylation in maize</td>
<td>pSer-162</td>
<td>(Huber and Huber, 1996)</td>
</tr>
<tr>
<td>Rubisco EC 4.1.1.39</td>
<td>C₃ photosynthesis</td>
<td>Disulphide bridge Phosphorylation, methylation, acetylation</td>
<td>Cys-247</td>
<td>(Friso and van Wijk, 2015) (Hodges et al., 2013) (Houtz et al., 2008)</td>
</tr>
<tr>
<td>Transketolase EC 2.2.1.1</td>
<td>Pentose-phosphate pathway</td>
<td>Phosphorylation</td>
<td>Ser-428</td>
<td>(Rocha et al., 2014)</td>
</tr>
</tbody>
</table>
3.1.5 Identifying novel phosphorylation sites using LC-MS/MS

The adaptation of proteomic pipelines for large-scale studies has grown in recent years through the technological advances of high-throughput mass spectrometry. Proteomics, along with genomics and metabolomics, has become a powerful tool for identifying proteins and characterising protein function and is essential for elucidating the factors that contribute to the complexity of the proteome (Porubleva and Chitnis, 2000; Park, 2004). Using high resolution mass spectrometry, not only can proteins be inferred from peptide spectra, but post-translational modifications (PTMs) can be accurately identified to amino acid specificity. The analysis of complex mixtures is most frequently achieved by coupling mass spectrometry to high performance liquid chromatography (Pitt, 2009), as in the case of the Orbital Trap mass spectrometer (Aebersold and Mann, 2003; Hu et al., 2005; Perry et al., 2008), in which ions can be separated on a liquid elution gradient prior to injection into the instrument. However, success in proteomic studies, such as phosphoproteomics, depends on several factors that may outweigh the technological improvements and optimisations during analysis. Phosphoproteomic studies are hindered by experimental challenges such as sample preparation, which may affect phosphopeptide recovery and yield computational challenges, like poorly annotated protein databases (Ma, 2010). In plants, these challenges are more pronounced due to incomplete genome sequences and uncharacterised protein entries. Furthermore, proteins from plant samples may be underrepresented due to incomplete protein extraction from various specialised cell types or organelles (Jorrín et al., 2007), presence of nucleic acids and non-protein contaminants that affect subsequent analysis (Haynes and Roberts, 2007), high concentration of in-soluble membrane proteins, which are lost during detergent-free extractions using non-volatile buffers and high content of hydrophobic molecules that may block analytical columns (Abdallah et al., 2012; Bagag et al., 2013).

The success in proteomic studies relies on increased resolution, sensitivity and accuracy of analytical protocols. Early proteomic experiments were designed to understand cellular function at the protein level by mapping proteins using 2D protein gels, which were limited to small-scale experiments due to the increased difficulty in identifying multiple proteins in more complex mixtures (O'Farrell, 1975; Mann et al., 2001; Graves and Haystead, 2002). This was overcome by the development of mass spectrometry technology, which had the necessary sensitivity and resolution to accurately identify
proteins in complex mixtures. The study of proteins by mass spectrometry became capable after the development of ionisation strategies, which transfer the protein analyte from the liquid to the gas phase, maintaining ions in their charged state. There are two commonly used methods for ionisation: 1) matrix-assisted laser desorption/ionization (MALDI) and 2) electrospray ionisation (ESI); the latter ionisation source is used in OrbiTrap-based mass spectrometers (Aebersold and Mann, 2003; Hu et al., 2005). During the first stage of analysis (MS1), ions (peptides) are accelerated through positively and negatively charged plates, separated and selected by their mass-to-charge ratio (m/z) using an electromagnetic field. In tandem mass spectrometry, ions are isolated in a collision chamber and fragmented with an inert gas during collision-induced dissociation (CID) (Marcotte, 2007; Dephoure et al., 2013), which breaks down peptides into smaller fragments, most commonly at the peptide bond, resulting in b-ions (N-terminal fragments) and y-ions (C-terminal fragments), which are analysed, during the second stage of analysis (MS2) (Aebersold and Mann, 2003).

For rapid protein identification, the resulting peptide spectra are searched against theoretical mass spectra, which are generated from the computational digestion of protein sequence databases (Zhang et al., 2013). For phosphopeptide determination, observed spectra are searched against theoretical peptide spectra generated assuming all possible serine, threonine or tyrosine phosphorylation events (Dephoure et al., 2013). Phosphorylated peptides bearing a phosphate group (80 Da), unlike dephosphorylated peptides, behave differently during CID fragmentation. The behaviour of the labile phosphoester bond during fragmentation can affect phosphorylation site assignment due to peptide spectra generated from multiple fragment ion products, often from the partial neutral loss of phosphoric acid (H₃PO₄; 98 Da), through the gas-phase β-elimination of phosphoserine or phosphothreonine (Schweppe et al., 2003; Thingholm et al., 2009; Solari et al., 2015). The neutral loss of phosphoric acid from phosphoserine or phosphothreonine is a favourable event during CID, unlike phosphotyrosine which is resistant to gas-phase β-elimination due to stabilisation of β-protons in the benzene ring (Figure 3.2) (Mann et al., 2002; Schweppe et al., 2003). Moreover, the neutral loss of water (18 Da), along with the loss of phosphate competes with mass spectra for the loss of phosphoric acid, which, under specific conditions, can hinder phosphorylation site assignment (Cui et al., 2014). Although these complex chemical interactions pose challenges for neutral loss scanning methods and subsequent mass spectrum analysis, the
versatility of new proteomic approaches, as well as tailored prediction algorithms and mass spectra visualisation tools have enhanced the accuracy of PTM identification by mass spectrometry (Kwon et al., 2006; Audagnotto and Dal Peraro, 2017).

To understand the significance of convergent C₄-related enzyme evolution and identify the determinants for C₄-specific enzyme activity, the phosphorylation of C₄ proteins in three NADP-malic enzyme C₄ subtypes was studied using high-resolution tandem mass spectrometry. Recent large-scale phosphoproteomic studies in C₃ plant models have demonstrated that several metabolic proteins undergo phosphorylation, regardless of their subcellular compartmentation, indicating that phosphoregulatory mechanisms are ubiquitous in C₃ plant organelles and in the cytosol (Nakagami et al., 2010; Hodges et al., 2013; Liu et al., 2014; Lv et al., 2014; Ye et al., 2016). The enzymes that are involved in carbon fixation and photorespiration in C₄ plants operate between two specialised cell types (Hatch, 1987; Kanai and Edwards, 1999), in the cytosol, as well as in chloroplasts, mitochondria and peroxisomes (Hatch and Osmond, 1976; Hodges et al., 2013), and may be, like their C₃ counterparts, subjected to PTMs. However, the diversity of these modifications and how they regulate C₄-specific protein activity across subcellular compartments as well as how they contribute to the efficiency of C₄ photosynthesis is unclear. High-resolution mass spectrometry has been previously used in targeted studies to describe the light-dependent phosphorylation sites on PEPCK (Chao et al., 2014) or light-intensity dependent phosphorylation sites on PPDK in Zea mays (Chen et al., 2014). However, there have not been any large-scale phosphoproteomic experiments in C₄ grasses, so the underlying regulatory properties of C₄-related proteins involved in C₄ photosynthesis are unaccounted for. The aim of this investigation was to determine novel phosphorylation sites on key C₄-related proteins from darkened and illuminated leaves of Setaria viridis, as well as in Sorghum bicolor and Megathyrsus maximus. In this Chapter, novel light-dependent phosphorylation sites identified on C₄-related proteins are presented and discussed. This was the first attempt to analyse the phosphoproteome in Setaria viridis, with hopes to elucidate novel phosphoregulatory mechanisms and further understand the role of phosphorylation and how changes of protein activity, through a diurnal cycle, are regulated. The comprehensive list of novel phosphorylation sites identified are listed in Appendix A.
3.2 Results – Novel phosphorylation sites of C₄-related enzymes

3.2.1 Identifying key C₄ photosynthesis proteins in *Setaria viridis*

To determine novel phosphorylation sites in the C₄ model *Setaria viridis*, proteins were extracted from 2.5-week old leaves 30 min before the onset of light and 4 h and 15.5 h into the photoperiod. Leaf extracts were subjected to 1D SDS-PAGE and stained with Coomassie Blue protein dye (Figure 3.8). Two bands above 100 kDa were excised to identify previously described phosphorylation of PEPC (phosphoserine at position 15, residues 12–20, HHSIDAQLR) and PPDK (phosphothreonine at position 527, residues 524–535, GGMTSHAAVVAR) in *Zea mays* (Roeske *et al.*, 1988; Jiao *et al.*, 1991). The tolerances for subsequent MS/MS spectra validations were established against the aforementioned phosphopeptide controls. Twenty-three proteins involved in carbon fixation were identified, of which seven were specific to C₄ photosynthesis (Table 3.3). 1,702 spectra were matched from 155 observed peptides of PEPC at 7.5 h into the dark compared to 1,669 peptide spectrum matches (PSMs) from 131 observed peptides of PPDK. The number of PEPC PSMs increased to 1,983 at 15.5 h into the light and amino acid sequence coverage increased by 16%. Similarly, PPDK sequence coverage increased by 41% between 7.5 h into the dark and 15.5 h into the light, resulting in 2,358 PSMs.
Fig. 3.8. *Setaria viridis* whole leaf proteins resolved by 1D SDS-PAGE. The 4–12% SDS gel was stained with Coomassie Blue protein dye. Lane 1 and lane 2 contained protein and phosphoprotein standards, respectively. 10 µg of protein was loaded into each lane, across four replicates (A–D) at three time-points: 7.5 h into the dark (H1A–H1D, lane 3–6), 4 h into the light (H2A–H2D, lane 7–10) and 15.5 h into the light (H3A–H3D, lane 11–14). Lane number 15 did not contain any protein. The gel was imaged using a FLA-5100 with 635 nm laser excitation. PEPC and PPDK are indicated at 109 kDa and 102 kDa, respectively.

Unlike PEPC and PPDK, the recovery of NADP-malic enzyme in *Setaria viridis* was less consistent, but protein sequence coverage increased by 58% after the dark to light transition. From the total 143 theoretical peptides, assuming a maximum of two missed cleavages filtered above 500 Daltons, 48 peptides were identified in NADP-malic enzyme from the 7.5 h into dark extracts; this increased by 13% after plants were illuminated for four hours, and then again by 2% towards the end of the photoperiod. The number of PSMs increased 4-fold from 228 to 1,179 after 15.5 h illumination; however, this large difference was due to the increased resolution of observed peptide spectra rather than an increased number of total observed peptides.
### Table 3.3. Summary of proteins involved in C4 photosynthesis identified by MS/MS.

Accession numbers, protein mass (kDa) and the calculated isoelectric point (pl) obtained from the UniProtKB protein database.

<table>
<thead>
<tr>
<th>Protein name</th>
<th>UniProt Accession</th>
<th>Mass (kDa)</th>
<th>Calc. pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine aminotransferase</td>
<td>K4A868</td>
<td>59.3</td>
<td>8.09</td>
</tr>
<tr>
<td></td>
<td>K3ZSX0</td>
<td>53.2</td>
<td>7.64</td>
</tr>
<tr>
<td></td>
<td>K3YSB2</td>
<td>50.2</td>
<td>8.63</td>
</tr>
<tr>
<td></td>
<td>K3XHJ0</td>
<td>50.2</td>
<td>8.68</td>
</tr>
<tr>
<td>Aspartate aminotransferase</td>
<td>K4AG31</td>
<td>16.1</td>
<td>5.01</td>
</tr>
<tr>
<td></td>
<td>K3Z6L1</td>
<td>44.6</td>
<td>6.28</td>
</tr>
<tr>
<td></td>
<td>K3YSM6</td>
<td>47.6</td>
<td>6.84</td>
</tr>
<tr>
<td></td>
<td>K3XHZ4</td>
<td>47.3</td>
<td>8.22</td>
</tr>
<tr>
<td>Enolase</td>
<td>K3XWW9</td>
<td>47.9</td>
<td>5.31</td>
</tr>
<tr>
<td></td>
<td>K3Z681</td>
<td>47.9</td>
<td>5.81</td>
</tr>
<tr>
<td></td>
<td>P42895</td>
<td>48.1</td>
<td>5.97</td>
</tr>
<tr>
<td></td>
<td>P26301</td>
<td>48.0</td>
<td>5.33</td>
</tr>
<tr>
<td>NADP-dependent malate dehydrogenase</td>
<td>K3YHB4</td>
<td>52.8</td>
<td>6.68</td>
</tr>
<tr>
<td></td>
<td>K3ZT87</td>
<td>49.3</td>
<td>6.35</td>
</tr>
<tr>
<td></td>
<td>K4ACE3</td>
<td>35.5</td>
<td>6.09</td>
</tr>
<tr>
<td>NAD-dependent malate dehydrogenase</td>
<td>K3ZU36</td>
<td>40.8</td>
<td>7.59</td>
</tr>
<tr>
<td></td>
<td>K3Z7W1</td>
<td>34.3</td>
<td>7.39</td>
</tr>
<tr>
<td></td>
<td>K3XJN7</td>
<td>35.5</td>
<td>7.85</td>
</tr>
<tr>
<td></td>
<td>K3Z7Q4</td>
<td>35.5</td>
<td>8.10</td>
</tr>
<tr>
<td>NADP-dependent malic enzyme</td>
<td>K3XG11</td>
<td>63.8</td>
<td>7.15</td>
</tr>
<tr>
<td></td>
<td>K3XFH4</td>
<td>65.1</td>
<td>6.16</td>
</tr>
<tr>
<td></td>
<td>K3ZRI5</td>
<td>70.0</td>
<td>6.77</td>
</tr>
<tr>
<td>Phosphoenolpyruvate carboxylase</td>
<td>K3XV32</td>
<td>109.9</td>
<td>6.34</td>
</tr>
<tr>
<td></td>
<td>K3YPN6</td>
<td>109.8</td>
<td>6.20</td>
</tr>
<tr>
<td>Pyruvate, phosphate dikinase</td>
<td>K3Z3Q6</td>
<td>95.7</td>
<td>5.12</td>
</tr>
</tbody>
</table>

Smaller sized or less abundant proteins in the leaf at the time of harvesting showed a lower number of PSMs and below 50% sequence coverage on average, suggesting that analyte recovery may have been overwhelmed by large quantities of more abundant ions (peptides) from PEPC and PPDK. PEP carboxykinase (PEPCK) was not identified in any of the *Setaria viridis* leaf lysates, but 38% of the total sequence was observed in *Megathyrsus maximus* at 7.5 h into the dark and 4 h into the light (Table 3.7) using a similar experimental approach. To evaluate the subset proteins that undergo in vivo phosphorylation, gels were stained with ProQ Diamond phosphoprotein gel stain. The
light-dependent phosphorylation of PEPC was evident in illuminated leaf lysates, whereas no banding was observed in darkened leaves (Figure 3.9). Conversely, the phosphorylation of PPDK was more prominent at 7.5 h into dark. The phosphorylation pattern of PEPC and PPDK at 4 h and 15.5 h into the light was less consistent across the biological replicates, and there was no distinguishable pattern of light-dependency between the light time-points. Phosphoprotein profiling using 2D gels (Figure 3.10) concurred with Figure 3.9 and showed that PEPC phosphorylation increased from dark to light, while PPDK phosphorylation remained constant throughout the 16h light and 8h dark cycle.

![Phosphoprotein Profiling](image)

**Fig. 3.9. Fluorescence stained gel showing phosphorylated proteins in *Setaria viridis*.** 10 µg of protein from 7.5 h into the dark, 4 h and 15.5 h into the light were resolved using a 12% SDS gel. After electrophoresis, the gel was stained with ProQ Diamond phosphoprotein fluorescent gel stain. 8 µL of PeppermintStick™ phosphoprotein standard was loaded into lane 2 (lane 1 is not shown). Phosphorylation of PEPC and PPDK is indicated. The gel was visualised with 532 nm laser excitation.
Fig. 3.10. 7 cm stained gels showing *Setaria viridis* whole leaf proteins resolved by 2D SDS-PAGE. 80 μg of protein was fixed onto a pH 4–7 IPG strip by passive rehydration, followed by IEF. After IEF, proteins were resolved on 10% SDS gels and stained with ProQ Diamond (top) or Coomassie InstantBlue™ (bottom). PEPC and PPDK are indicated by the red and white rectangles, respectively.
Using tandem mass spectrometry, PEPC was confirmed to undergo phosphorylation at Ser-11 (Figure 3.12A, residues 9–17, HHSIDAQLR), strictly after leaves were illuminated. However, during validation experiments, a triply charged phosphopeptide ion (residues 2–17, ASKFVE-KHHSIDAQLR, m/z 604.99) was observed four times in 7.5 h into the dark protein extracts (Table 3.7). The phosphorylation of PEPC at 7.5 h into the dark was not evident from the phosphoprotein stained gel (Figure 3.10) and presumably remained present in the dark period as an intermediate PEPC isoform during the light to dark transition. While the phosphorylation at Ser-11 in PEPC generally exhibited a strict light-dependent regulatory mechanism, PPDK phosphorylation at Thr-462 (Figure 3.15, residues 459–470, GGMTSHAAVVAR) was identified in every leaf sample mixture, irrespective of the harvesting time-point in Setaria viridis and Sorghum bicolor leaf lysates. Additionally, the phosphorylation at Ser-463 was detected four times, although in low confidence, at 7.5 h into dark in validation experiments using a 1% FDR (Table 3.6). The phosphorylation of Ser-463 had not been identified in Setaria viridis protein extracts. In total, 69 putative phosphorylation sites from key C₄ photosynthesis enzymes in Setaria viridis were identified by LC-MS/MS (Appendix A): 12 sites in alanine aminotransferase, 9 sites in aspartate aminotransferase, 13 sites in malate dehydrogenase, 20 sites in NADP-malic enzyme, seven sites in PEPC including the identified phosphorylation at Ser-15 in Zea mays, and eight sites in PPDK, including the phosphorylation identified at Thr-527 in Zea mays.
Table 3.4. Novel phosphopeptides identified in *Setaria viridis* leaf lysates using phosphopeptide enrichment. Phosphorylation sites identified in 7.5 h darkened (A) and 15.5 h illuminated (B) *Setaria viridis* by TiO$_2$ enrichment. The resulting spectra were searched against UniProtKB *Setaria italica* protein sequences using SEQUEST. The percentage of the protein sequence that was experimentally observed is given for each protein match. Phosphorylation site localisation probabilities ($P$) are indicated in parentheses to the right of the modified amino acid ($P > 0.75$ is significant). Statistical confidence is given by Xcorr, where Xcorr $\geq 2.15$ is significant.

Table 3.4A. Novel phosphorylation sites identified in 7.5 h darkened *Setaria viridis* leaves.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Coverage (%)</th>
<th>Phosphopeptide</th>
<th>Position in peptide</th>
<th>MH$^+$ (Da)</th>
<th>Monoisotopic mass (Da)</th>
<th>Xcorr</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPDK</td>
<td>K3Z464</td>
<td>25.27</td>
<td>NDNDLTAS(0.99)DLKELVAQYK GGMT(0.97)SHAAVVAR</td>
<td>Ser-8 Thr-4</td>
<td>2116.98</td>
<td>2035.99</td>
<td>3.99</td>
</tr>
</tbody>
</table>

Table 3.4B. Novel phosphorylation sites identified in 15.5 h illuminated *Setaria viridis* leaves.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Coverage (%)</th>
<th>Phosphopeptide</th>
<th>Position in peptide</th>
<th>MH$^+$ (Da)</th>
<th>Monoisotopic mass (Da)</th>
<th>Xcorr</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGM</td>
<td>K3XF0</td>
<td>8.69</td>
<td>AHGTAVGLPSDDDMGS(1)EVGH NALGAGR</td>
<td>Ser-17$^|$</td>
<td>2801.18</td>
<td>2704.21</td>
<td>3.48</td>
</tr>
<tr>
<td>PEPC</td>
<td>K3XV32</td>
<td>38.90</td>
<td>RGDFADEGFATATES(0.95)DIEET LKR</td>
<td>Ser-14</td>
<td>2537.12</td>
<td>2456.13</td>
<td>4.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LLAPGKVS(1)EDDKLVEYDALLI FTAATLEHGMPVS(1)PKPEWR</td>
<td>Ser-8$^|$ Ser-15$^|$</td>
<td>2652.36</td>
<td>2571.37</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RVDT(1)ALKNIGIDER</td>
<td>Thr-4</td>
<td>1679.85</td>
<td>1598.86</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NQTVDLVFTAHPTQS(0.99)VRR HHS(1)IDAQLR</td>
<td>Ser-15 Ser-3$^|$</td>
<td>2149.05</td>
<td>2068.07</td>
<td>1.50</td>
</tr>
</tbody>
</table>
3.2.2 Phosphoenolpyruvate carboxylase phosphorylation

Seven phosphorylation sites (Figure 3.11) were identified in *Setaria viridis* PEPC (K3XV32), including the phosphoserine described in *Zea mays* PEPC (P04711). PEPC-2 (K3YPN6) was also identified in *Setaria viridis*, but despite conserved active site residues, the extent of similarity between PEPC-2 and the C₄ PEPC was not investigated. Furthermore, the N-terminal phosphopeptide was not detected in PEPC-2. Based on the phosphoproteomic data and information obtained from the UniProtKB database, candidate protein K3XV32 was selected as the true counterpart to *Zea mays'* C₄ PEPC primarily because of 1) the fully conserved N-terminal phosphopeptide which bears the phosphorylated regulatory site described in *Zea mays*, 2) higher sequence similarity to *Zea mays* PEPC and 3) more substantial peptide recovery, suggesting that PEPC is relatively more abundant in leaf lysates than PEPC-2 (Table 3.5).

PEPC sequence information was obtained from the UniProtKB database, aligned using Clustal Omega and illustrated to scale as Figure 3.11. Sequence similarity between C₄ PEPC in *Zea mays* (P04711) and PEPC in *Setaria viridis* (K3XV32) is 84% covering 814 identical positions, which is close to the 83% similarity between PEPC and PEPC-2, whereas similarity between PEPC-2 and P04711 is slightly lower at 80%, based on 781 identical positions. All phosphorylated residues, except for Ser-378 in K3XV32, are conserved in the *Zea mays* C₄ PEPC. PEPC in *Setaria viridis* was phosphorylated at Ser-11 (Figure 3.12A, residues 9–17, HHSIDAQLR) after the onset of light and remained phosphorylated until the end of the photoperiod (Table 3.4B). While Ser-12 (PEPC-2, residues 9–18, HQSIDAQLR) corresponds to Ser-11 in PEPC, it was not phosphorylated. The phosphorylation of Ser-25 of PEPC occurred strictly after the onset of light and corresponds to Ser-26 in PEPC-2, which was phosphorylated at the same time-point.
Fig. 3.11. Protein schematic showing the similarities between PEPC P04711, K3XV32 and K3YPN6. Proteins sequences are represented by rectangles and the number of amino acids in the sequence is indicated at the far right of each protein. Light or dark phosphorylation sites are highlighted in yellow or blue boxes, respectively. Phosphorylation sites present in darkened and illuminated leaves are shown in half-yellow-half-blue boxes. The regulatory site at Ser-15 in Zea mays, corresponding to Ser-11 in K3XV32, is indicated by a red dot above the modified residue. PEPC structural data for Zea mays obtained from Matsumura et al. (2002).
Fig. 3.12A. Phosphorylation at Ser-11 occurred after illumination. PEPC peptide spectra showing the ion fragmentation pattern of residues 9–17 (above spectrum) along its mass-to-charge (m/z) ratio.
Fig. 3.12B. Phosphorylation at Ser-752 occurred in darkened leaves. PEPC peptide spectra showing the ion fragmentation pattern of residues 748–757 (above spectrum) along its mass-to-charge (m/z) ratio. Charge of the peptide is indicated by the blue number above the peptide sequence. The detected b- and y-ions (including loss of H$_2$O, 18 Da and loss of (PO$_4$)$_3^-$, 80 Da) are indicated in green and yellow, respectively. Spectrum quality is determined by the total number of b- and y-ions observed along the peptide sequence and the size of peaks (relative intensity, %).
The phosphorylations at Ser-25, Ser-378 and Thr-379 were present in illuminated conditions, but absent in darkened leaves. Conversely, the phosphorylation at Ser-752 (Figure 3.12B, residues 748–757, LNIGSRPAKR) solely occurred in darkened leaves and was not identified in illuminated leaves. As shown in Figure 3.9, the phosphorylations at Ser-180 (Figure 3.13A, residues 180–189, SLLQKHARIR) and Ser-702 (Figure 3.13B, residues 688–708, FTAATLEHGMHPVSPKPEWR) occurred in darkened and illuminated leaves and showed no light-dependent regulation. Additionally, the light-dependent phosphorylations of Thr-658 and Thr-820 occurred in PEPC-2, but the corresponding residues in PEPC were not phosphorylated.
Fig. 3.13A. Phosphorylation at Ser-180 occurred in darkened and illuminated leaves of *Setaria viridis*. PEPC peptide spectra showing the ion fragmentation pattern of residues 180–189 (above spectrum) along its mass-to-charge (m/z) ratio. Protonated precursor ion in red (ca 1200 m/z).
Fig. 3.13B. Phosphorylation at Ser-702 occurred in darkened and illuminated leaves of *Setaria viridis*. PEPC peptide spectra showing the ion fragmentation pattern of residues 688–708 (above spectrum) along its mass-to-charge \((m/z)\) ratio. Charge of the peptide is indicated by the blue number above the peptide sequence. The detected b- and y-ions (including loss of H₂O, 18 Da and loss of \((\text{PO}_4)^{3-}\), 80 Da) are indicated in green and yellow, respectively. Putative phosphorylation sites highlighted in red.
Table 3.5. Protein isoforms identified in 7.5 h darkened *Setaria viridis* leaves. Information acquired from the validation of phosphopeptides in *Setaria viridis* at 7.5 h into dark from four biological replicates (R1-R4) using MaxQuant. Unique peptides are unique to the protein group indicated by the protein accession, whereas all other peptides, which exist in related proteins, are labelled as razor peptides. Sequence coverage is given as a percentage of the total number of amino acids present in the sum of the experimentally observed unique peptides and razor peptides against the total number of amino acids in the protein sequence. The intensity-based absolute quantification value (iBAQ) is a mathematical approximation used to compare the relative amount of one protein to another within the same complex mixture.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Number of razor peptides</th>
<th>Number of unique peptides</th>
<th>Mass (kDa)</th>
<th>Coverage (%)</th>
<th>iBAQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>AlaAT</td>
<td>K3ZSX0</td>
<td>R1 0 R2 0 R3 0 R4 0</td>
<td>R1 4 R2 3 R3 2 R4 2</td>
<td>53.2</td>
<td>20.5</td>
<td>366910</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K4A8B1</td>
<td>R1 0 R2 0 R3 0 R4 0</td>
<td>R1 1 R2 0 R3 0 R4 0</td>
<td>58.2</td>
<td>3.0</td>
<td>3011</td>
</tr>
<tr>
<td>Enolase 1</td>
<td>K3XWW9</td>
<td>R1 7 R2 7 R3 5 R4 5</td>
<td>R1 9 R2 10 R3 9 R4 6</td>
<td>47.9</td>
<td>45.3</td>
<td>1202000</td>
</tr>
<tr>
<td>Enolase 2</td>
<td>K3Z681</td>
<td>R1 0 R2 0 R3 0 R4 0</td>
<td>R1 7 R2 7 R3 7 R4 4</td>
<td>47.9</td>
<td>38.2</td>
<td>2003400</td>
</tr>
<tr>
<td>NADP-malic enzyme</td>
<td>K3XG11</td>
<td>R1 3 R2 3 R3 3 R4 3</td>
<td>R1 38 R2 39 R3 34 R4 38</td>
<td>63.9</td>
<td>56.7</td>
<td>2208700</td>
</tr>
<tr>
<td>NADP-malic enzyme</td>
<td>K3XFW4</td>
<td>R1 0 R2 1 R3 1 R4 0</td>
<td>R1 3 R2 4 R3 2 R4 2</td>
<td>65.1</td>
<td>13.4</td>
<td>412370</td>
</tr>
<tr>
<td>NADP-malic enzyme</td>
<td>K3ZRI5</td>
<td>R1 0 R2 0 R3 0 R4 0</td>
<td>R1 19 R2 13 R3 16 R4 22</td>
<td>69.5</td>
<td>49.9</td>
<td>3637300</td>
</tr>
<tr>
<td>PEPC</td>
<td>K3XV32</td>
<td>R1 14 R2 16 R3 18 R4 21</td>
<td>R1 120 R2 12 R3 133 R4 141</td>
<td>109.98</td>
<td>75.2</td>
<td>5812700</td>
</tr>
<tr>
<td>PEPC</td>
<td>K3YPN6</td>
<td>R1 3 R2 1 R3 2 R4 3</td>
<td>R1 3 R2 5 R3 6 R4 8</td>
<td>109.82</td>
<td>27.6</td>
<td>790010</td>
</tr>
</tbody>
</table>
3.2.3 Pyruvate, phosphate dikinase phosphorylation

Eight phosphorylation sites were identified in *Setaria viridis* pyruvate, phosphate dikinase (PPDK, K3Z3Q6), including the previously described phosphopeptide at position Thr-527 in *Zea mays* (Roeske et al., 1988), which corresponds to Thr-462 in *Setaria viridis* (Figure 3.14). The extent of sequence similarity between K3Z3Q6 and P11155 is 83% covering 786 identical amino acid positions, including all identified phosphorylated residues. The phosphorylation at Thr-462 (Figure 3.15, residues 459–470, GGM\_SHAAVVAR) occurred at 7.5 h into the dark and 4 h and 15.5 h into the light, suggesting that PPDK phosphorylation at Thr-462 may be under more extensive regulation rather than being simply induced by the dark to light transition. Based on the ratio between the phosphopeptide and its unmodified counterpart, as a relative quantitative measure, the phosphorylated phosphopeptide was observed in higher amounts in 7.5 h darkened leaves than in illuminated leaves. Aside from Thr-462, phosphorylations at Ser-104, Ser-394 and Thr-593 (Figure 3.16, residues 584–597, AVRQIMAP\_TELRL) occurred at 7.5 h into dark, whereas the phosphorylations at Thr-557, Thr-704, Thr-721 and Thr-753, occurred strictly after the onset of light. Additionally, the phosphorylations at Thr-753 and Thr-462 were identified in the phosphopeptide enrichment experiments.
Proteins sequences (rectangles) are annotated based on information obtained from UniProtKB. Light and dark phosphorylation sites identified in K3Z3Q6 are highlighted in yellow or blue boxes, respectively (light-independent phosphosites are highlighted in blue and yellow). Described phosphorylation sites are indicated by a red dot above the modified residue. Purple dots above residues indicate metal and substrate binding sites.
Fig. 3.15. PPDK phosphorylation at Thr-462 in *Setaria viridis*. PPDK peptide spectrum showing the ion fragmentation pattern of residues 459–470 (above spectrum) along its mass-to-charge (m/z) ratio. Charge of the peptide is indicated by the blue number above the peptide sequence. The detected b- and y-ions (including loss of H$_2$O, 18 Da and loss of (PO$_4$)$_3^-$, 80 Da) are indicated in green and yellow, respectively. Putative phosphorylation sites highlighted in red. Protonated precursor ion indicated in red.
Fig. 3.16. Phosphorylation at Thr-593 occurred strictly in darkened leaves of *Setaria viridis*. PPDK peptide spectrum showing the ion fragmentation pattern of residues 584–597 (above spectrum) along its mass-to-charge (m/z) ratio. Charge of the peptide is indicated by the blue number above the peptide sequence. The detected b- and y-ions (including loss of H$_2$O, 18 Da and loss of (PO$_4$)$_3^-$, 80 Da) are indicated in green and yellow, respectively. Putative phosphorylation sites highlighted in red.
3.2.4 NADP-dependent malic enzyme phosphorylation

Four NADP-dependent malic enzyme proteins were identified in *Setaria viridis*: K3XFH6, K3XFW4, K3XG11 and K3ZRI5 (Table 3.3). Compared to the primary NADP-malic enzyme in *Zea mays* (P16243), K3ZRI5 is 36% identical, while the protein sequence similarity between P16243 and K3XFH6, K3XFW4 and K3XG11 is higher at 83%, 70% and 75%, respectively. In addition, NADP-malic enzyme K3XG11 was approximately 60-fold and 535-fold more abundant in darkened leaf lysates than K3ZRI5 and K3XFW4, respectively (Table 3.5). Using a standard MS approach without enriching for phosphopeptides, 21 phosphorylation sites in NADP-malic enzyme were identified. Five putative phosphorylation sites were identified in K3XFH6 during the light time-points, of which three were significant (Xcorr > 2.15). The phosphorylation at Ser-429 (Figure 3.18, residues 417–430, VWLVDKGLIVSSR) corresponded to the phosphorylation at Ser-380 in K3XFW4 and Ser-363 in K3XG11 (Figure 3.17). Although, the phosphorylations at Ser-373, Ser-379 and Ser-380 in K3XFW4 were identified in separate experiments, the phosphopeptides and phosphorylation sites were not significant. Furthermore, the triply phosphorylated peptide was not detected in the light samples.
Fig. 3.17. NADP-malic enzyme isoforms identified in *Setaria viridis* leaf extracts compared to NADP-malic enzyme from *Zea mays* (top). Protein sequences (rectangles) were obtained from UniProtKB. Light or dark occurring phosphorylation sites are highlighted in yellow or blue, respectively. Phosphosites present in light and dark highlighted in blue and yellow. Proteins are aligned at the metal binding sites.
Fig. 3.18. Peptide spectrum showing novel NADP-malic enzyme phosphorylation at Ser-429 in *Setaria viridis*. NADP-malic enzyme peptide spectrum showing the ion fragmentation pattern of residues 417–430 (above spectrum) along the mass-to-charge (*m/z*) ratio. Charge of the peptide is indicated by the blue number above the peptide sequence. The detected b- and y-ions (including loss of H$_2$O, 18 Da and loss of (PO$_4$)$_3^-$, 80 Da) are indicated in green and yellow, respectively. Putative phosphorylation sites highlighted in red.
3.2.5 Alanine and aspartate aminotransferase phosphorylation

Seven phosphopeptides were identified in alanine aminotransferase, of which six were significant (Xcorr > 2.15). In total, 12 phosphorylation sites were identified in alanine aminotransferase (Figure 3.20). Reoccurring phosphorylations occurred at Thr-167 and Ser-189 (Figure 3.19, residues 166–190, ATGAYSHSQIKGLRDAIAAGIASR) in 7.5 h into dark and 4 h after the onset of light and remained phosphorylated until 30 min before the dark period. The phosphorylations at Thr-57, Ser-62, Thr-203, Ser-207, Thr-288, Ser-290, Ser-322 and Ser-326 occurred in illuminated leaves, whereas the phosphorylations at Thr-167, Ser-189 and Ser-206 occurred in darkened and illuminated leaves. Only one putative phosphorylation site, at Ser-285, was present in darkened leaves. Conversely, aspartate aminotransferase was phosphorylated at more residues during the dark period. Seven phosphopeptides, of the nine identified in aspartate aminotransferase, were significant (Xcorr > 2.15). The phosphorylations at Ser-10, Ser-11, Ser-99, Ser-266, Ser-273, Ser-341 and Ser-391, were only found in darkened leaves, whereas the phosphorylations of Ser-300, Thr-394 and Thr-397 were only present after illumination (Figure 3.21).
Fig. 3.19. Peptide spectrum showing alanine aminotransferase phosphorylations at Thr-167 and Ser-189. Alanine aminotransferase peptide spectrum showing the ion fragmentation pattern of residues 166–190 (above spectrum) along its mass-to-charge (m/z) ratio. Charge of the peptide is indicated by the blue number above the peptide sequence. The detected b- and y-ions (including loss of H2O, 18 Da and loss of (PO4)3−, 80 Da) are indicated in green and yellow, respectively. Putative phosphorylation sites highlighted in red. * N-acetylation.
**Fig. 3.20. The phosphorylation of alanine aminotransferase in *Setaria viridis occurred predominantly after illumination***. Protein sequences (rectangles) were obtained from the UniProtKB database and aligned using Clustal Omega. Key amino acids are indicated by the coloured lines and residue number. Novel sites of phosphorylation are indicated in red. Light or dark occurring phosphorylation sites are highlighted in yellow or blue, respectively. Phosphorylations occurring in light and dark are highlighted in yellow and blue boxes.
Fig. 3.21. The phosphorylation of aspartate aminotransferase in *Setaria viridis* occurred predominantly in darkened leaves. Protein sequences (rectangles) were obtained from the UniProtKB database and aligned using Clustal Omega. Key amino acids are indicated by the coloured lines and residue number. Novel sites of phosphorylation are indicated in red. Light or dark occurring phosphorylation sites are highlighted in yellow or blue, respectively.
3.2.6 Phosphoglycerate mutase and enolase phosphorylation

Three phosphopeptides were identified in phosphoglycerate mutase (K3XFX0), one of which was detected in the phosphopeptide enrichment experiments (Table 3.4B). The phosphorylation at Ser-109 (Figure 3.22A, residues 93–120, AHGTAVGLPSDDDMGN-SEVGHNALGAGR) occurred in 4 h and 15.5 h illuminated leaves of *Setaria viridis*, and during phosphopeptide validation experiments, the phosphorylation was detected in three replicates of 7.5 h darkened *Setaria viridis* and *Sorghum bicolor* leaves (Table 3.6 and Table 3.7). The phosphopeptide was not detected in darkened leaves in previous experiments. Additionally, the phosphorylations at Ser-520, Thr-527 (Figure 3.22B, residues 520–549, SGGIQILTSHLQPVPVAIGGPGLHPGVKFR) occurred at 7.5 h into the dark period and 15.5 h into the photoperiod, but were absent 4 h after illumination. Either Ser-528 or Thr-530 could be phosphorylated in 7.5 h darkened and 15.5 h illuminated leaves, but the identified sites were not significant ($P < 0.75$). However, based on a 240.9 Da mass shift between the identified phosphopeptide and the unmodified counterpart, the peptide (residues 520–549) could be triply phosphorylated in darkened and illuminated leaves.

Two enolase proteins with 90% sequence similarity (K3XWW9 and K3Z681), represented as Figure 3.23, were identified in *Setaria viridis*. The 7.5 h into dark and 4 h into light phosphorylation at Ser-121 (K3XWW9, residues 112–132, LGANAILAVSLAVCKAGAMVK) corresponded to the phosphorylation at Ser-120 (K3Z681, residues 112–133, LGANAILAVSLAVCKAGASIKK) in an almost identical phosphopeptide. The phosphorylation of Ser-129 was only present in protein K3Z681, while the phosphorylations at Thr-334 and Thr-337 only occurred in K3XWW9.
Fig. 3.22A. Peptide spectra showing the phosphorylation at Ser-109 (residues 93–120) on phosphoglycerate mutase in *Setaria viridis*. 
Fig. 3.22B. Peptide spectra showing the phosphorylations at Ser-520, Thr-527 and Thr-530 on phosphoglycerate mutase in *Setaria viridis*. Phosphoglycerate mutase peptide spectra showing the ion fragmentation pattern of residues 520–549 (above spectrum) along its mass-to-charge (m/z) ratio. Charge of the peptide is indicated by the blue number above the peptide sequence. The detected b- and y-ions (including loss of H₂O, 18 Da and loss of (PO₄)³⁻, 80 Da) are indicated in green and yellow, respectively. Putative phosphorylation sites highlighted in red.
Fig. 3.23. Schematic showing the alignment of enolase K3XWW9 and K3Z681 from *Setaria viridis*. Protein sequences are represented by the rectangles. Putatively identified phosphorylation sites occurring in illuminated leaves or darkened leaves are highlighted in yellow or blue, respectively. Phosphorylation sites occurring in illuminated and darkened leaves are highlighted in blue and yellow.
3.2.7 Phosphopeptide validation

To validate the phosphopeptides that may be under light-dependent regulation, proteins between 110 kDa and 50 kDa from dark-grown *Setaria viridis* plants were analysed using a modified mass spectrometry protocol. Gradient lengths were optimised for the analyte and the resulting peptide spectra were searched against *Setaria italica* protein sequences using MaxQuant. Stringent tolerances were set for the classification of class I phosphopeptides and the FDR was set at 1%. Additional validations were carried out in two other C₄ species: *Sorghum bicolor* (Table 3.7) and *Megathyrsus maximus* (Table 3.8). The degree of reproducibility of putative phosphorylation sites was determined by the total number of times each phosphorylation site was observed across four biological replicates. Eight phosphorylation sites were validated in 7.5 h darkened *Setaria viridis* leaves (Table 3.7). Phosphoglycerate mutase was phosphorylated at Ser-17 in *Megathyrsus maximus* (Table 3.8), Ser-81 in *Sorghum bicolor* (searched against *Zea mays* protein sequences) and Ser-109 in *Setaria viridis* (searched against *Setaria italica* protein sequences). This phosphorylation was detected in several experiments (Table 3.4B; Table 3.6; Table 3.7; Table 3.8).
Table 3.6. Phosphopeptide validation for 7.5 h darkened *Setaria viridis* leaves. Class I phosphopeptides are given by localisation probabilities ($P \geq 0.75$ is significant) and score difference values (score diff. $\geq 5$ is significant) across four replicates (R1–R4). Candidate phosphorylation sites are underlined. Phosphopeptides identified in *Sorghum bicolor* are in Table 3.7.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Phosphopeptide</th>
<th>Residue</th>
<th>Times observed</th>
<th>Localisation probability</th>
<th>Score difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R1</td>
<td>R2</td>
</tr>
<tr>
<td>PGM</td>
<td>K3XF0</td>
<td>AHGTAVGLPSDDDMGNS**(1)**EVGHNALGAGR</td>
<td>S109</td>
<td>3 of 4</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>PEPC</td>
<td>K3XV32</td>
<td>ASKPVEKHHS**(1)**IDAQLR</td>
<td>S11</td>
<td>4 of 4</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>PPDK</td>
<td>K3Z3Q6</td>
<td>GGMT**(0.78)**HAAVVAR</td>
<td>S463</td>
<td>4 of 4</td>
<td>0.74</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 3.7. Phosphopeptide validation for 7.5 h darkened *Sorghum bicolor* leaves.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Phosphopeptide</th>
<th>Residue</th>
<th>Times observed</th>
<th>Localisation probability</th>
<th>Score difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R1</td>
<td>R2</td>
</tr>
<tr>
<td>PGM</td>
<td>P30792</td>
<td>AHGTAVGLPSDDDMGNS**(1)**EVGHNALGAGR</td>
<td>S81</td>
<td>3 of 4</td>
<td>0.99</td>
<td>-</td>
</tr>
<tr>
<td>PEPC</td>
<td>P04711</td>
<td>FTAATLEHGMPPVS**(0.97)**PK</td>
<td>S708</td>
<td>1 of 4</td>
<td>0.98</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F<strong>S</strong>(0.68)SWMGGDRDGNPR</td>
<td>S286</td>
<td>1 of 4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 3.8. Phosphorylation sites identified in 7.5 h darkened and 4 h illuminated leaves of *Megathyrsus maximus*. Novel phosphorylation sites identified in *Megathyrsus maximus* (Guinea grass). Experimental peptide spectra were searched against UniProtKB and NCBI *Zea mays* and *Megathyrsus maximus* protein sequences using MASCOT and phosphorylations were determined using PhosphoRS. Protein sequence coverage (out of 100%) is indicated for each protein. Phosphorylation sites are annotated by the position of the modified residue in the phosphopeptide sequence. Phosphorylation site localisation probabilities (*P*) are indicated in parentheses to the right of the modified amino acid, where *P* ≥ 0.75 is significant. The calculated mass of the protonated precursor ion (MH⁺) and the monoisotopic mass (Da) of the unmodified peptide are shown. Statistical confidence is given by the *E*-value, where an *E*-value ≤ 0.05 is significant. Reoccurring phosphorylation sites are indicated by $^§$.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Coverage (%)</th>
<th>Phosphopeptide</th>
<th>Position in peptide</th>
<th>MH⁺ (Da)</th>
<th>Monoisotopic mass (Da)</th>
<th><em>E</em>-value</th>
<th>Time point</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGM</td>
<td>P30792</td>
<td>24.33</td>
<td>AHGTAVGLPSDDDMGNS(1)EVGHNALGAGR</td>
<td>Ser-17</td>
<td>2801.20</td>
<td>2704.21</td>
<td>0.00389</td>
<td>TP3</td>
</tr>
<tr>
<td>PEPCK</td>
<td>Q9SLZ0</td>
<td>20.87</td>
<td>S(1)APTTPIK</td>
<td>Ser-1</td>
<td>894.43</td>
<td>813.45</td>
<td>0.0248</td>
<td>TP2</td>
</tr>
<tr>
<td>PEPCK</td>
<td>Q9SLZ0</td>
<td>20.87</td>
<td>S(1)APTTPIK</td>
<td>Ser-1</td>
<td>894.43</td>
<td>813.45</td>
<td>0.0382</td>
<td>TP3</td>
</tr>
<tr>
<td>PEPCK</td>
<td>33329363</td>
<td>70.72</td>
<td>S(0.98)APT(0.51)TPIKDGAATSFAAISEEDR</td>
<td>Ser-1, Thr-4§</td>
<td>2566.11</td>
<td>2405.16</td>
<td>0.00486</td>
<td>TP3</td>
</tr>
<tr>
<td>PEPCK</td>
<td>33329363</td>
<td>70.72</td>
<td>S(0.03)APT(0.49)T(0.49)PIKDGAATSTFAAALSEEER</td>
<td>Thr-4 or Thr-5</td>
<td>2500.17</td>
<td>2419.18</td>
<td>0.0433</td>
<td>TP3</td>
</tr>
<tr>
<td>PEPC</td>
<td>P04711</td>
<td>15.26</td>
<td>HHS(1)IDAQLR</td>
<td>Ser-3§</td>
<td>1156.53</td>
<td>1075.54</td>
<td>0.00324</td>
<td>TP3</td>
</tr>
<tr>
<td>PPDK</td>
<td>P11155</td>
<td>23.31</td>
<td>GGMT(1)SHAAVVAR</td>
<td>Thr-4§</td>
<td>1236.56</td>
<td>1155.57</td>
<td>0.00367</td>
<td>TP2</td>
</tr>
<tr>
<td>PPDK</td>
<td>P11155</td>
<td>24.24</td>
<td>GGMT(0.5)S(0.5)HAAVVAR</td>
<td>Thr-4, Ser-5</td>
<td>1236.56</td>
<td>1155.57</td>
<td>0.0546</td>
<td>TP3</td>
</tr>
</tbody>
</table>
3.3 Discussion

3.3.1 Overcoming the challenges of phosphoproteomics

Phosphorylated proteins account for approximately 0.1% of the proteome and phosphorylated peptides are difficult to detect because of their lower abundance relative to unmodified peptides (Seo and Lee, 2004; Tichy et al., 2011). To increase the recovery of phosphopeptides in *Setaria viridis*, protein extracts were first subjected to phosphopeptide enrichment using TiO$_2$. TiO$_2$ enrichment was specifically chosen due to its low-cost, high specificity for phosphorylation sites and tolerance towards a range of detergents and buffers readily used during protein extractions (Tichy et al., 2011). However, phosphopeptide enrichment in *Setaria viridis* was not successful, possibly due to low protein concentrations in plant cell lysates, which led to sub-optimal phosphopeptide binding. Future approaches may need pre-fractionation using strong cation exchange chromatography and multi-step enrichment techniques to gradually reduce proteome complexity (Zhao and Jensen, 2009; Fíla and Honys, 2012). Low quality spectra and high background noise indicate that there may have been contamination from unspecific binding during the enrichment (Figure 3.24). While TiO$_2$ is highly selective towards phosphopeptides, acidic residues or negatively charged contaminants may have bound to the resin instead, co-eluting with phosphopeptides. Phosphopeptide signal suppression was evident with most of the phosphorylation profiles in *Setaria viridis*, including the PEPC N-terminal phosphopeptide (Figure 3.24). The majority of phosphopeptides identified by TiO$_2$ phosphopeptide enrichment were singly phosphorylated (Table 3.4, 3.6), while containing several serine and threonine residues that were not mistaken as putative phosphorylation sites, which is likely to occur when site-determining ions are absent (Taus et al., 2011). The presented data suggests that phosphorylation sites were correctly assigned despite loss of phosphopeptide resolution. The use of phosphopeptide enrichment, however, for this analytical approach was not efficient mainly due to low peptide recovery, resulting in approximately 80–90% of protein sequences being absent.
Fig. 3.24. PEPC MS/MS spectra after A) TiO₂ phosphopeptide enrichment and B) no phosphopeptide enrichment. Phosphopeptide (residues 9–17, HH₅IDAQLR, m/z 578.77, z = 2+) shows increased background signals after phosphopeptide enrichment (A) when compared to the phosphopeptide (residues 9–17, HH₅IDAQLR, m/z 386.18, z = 3+) from non-enriched samples (B).
Insufficient fragment ion information from mass spectra hindered phosphorylation site assignment in *Setaria viridis*. While the observation was more notable after phosphopeptide enrichment, less abundant ions in the analyte showed diminished peptide resolution across all experiments, which not only affected PTM assignment but also made it difficult to properly infer proteins or distinguish protein isoforms. PTM assignment was also hindered by a low abundance of phosphopeptides in sample mixtures. It is not known at which stage the loss of phosphopeptides occurred, but addition of protease and phosphatase inhibitors in subsequent protein extracts contributed to increased phosphopeptide recovery. In phosphopeptide enrichments, strong binding affinities between multiply phosphorylated peptides and the enrichment resin may have made it difficult to elute several phosphopeptides. Phosphopeptide elution can be increased in future studies by optimising the pH of the peptide elution buffer, since phosphate groups contribute to the solution charge, affecting the elution efficiencies of phosphorylated peptides (Villen and Gygi, 2008). While subsequent analyses did not include a phosphopeptide enrichment step, it remains clear that to counter the effects of sub-stoichiometric phosphorylation, plant cell lysates should undergo several steps to reduce sample complexity and reduce dephosphorylated peptides, which can suppress signals from the modified counterparts.

Singly and multiply phosphorylated peptides in the mixture can also affect the overall elution of phosphopeptides. In these experiments, proteins were digested with trypsin to guarantee the presence of either lysine or arginine, generating peptides of basic nature, which aid the elution of singly phosphorylated peptides after their dephosphorylated counterparts, reducing the loss of phosphopeptides during the early stages of the flow gradient (Steen *et al.*, 2006). The effects are reversed in the presence of multiply phosphorylated peptides, and may be lost due to the predominantly negative charge. In multiply phosphorylated peptides, the rate of peptide hydrolysis by trypsin may decrease if phosphorylation sites are in proximity to the cleavage site (Figure 3.25), increasing peptide lengths, thereby increasing the number basic residues, which may counter the effects of multiply phosphorylated peptides (Schlosser *et al.*, 2001; Steen *et al.*, 2006). Phosphopeptides could also be lost due to lower ionisation efficiencies, but in a similar approach using ESI-LC-MS/MS, phosphopeptides showed better ionisation and detection efficiencies than dephosphorylated peptides (Steen *et al.*, 2006).
**Fig. 3.25. Schematic showing trypsin inaction due to a phosphorylation site along a peptide.** A) Tryptic digestion of a phosphopeptide (underlined) showing cleavage sites (red lines) at lysine (K) and arginine (R). The phosphorylation site is indicated by P. B) Hypothetical protein peptide digestion affected by phosphorylation next to cleavage site (K). Red lines crossing peptide sequence represent a successful cleavage. Negative charge (‒); positive charge (+).

The light intensity for these experiments was maintained at 350 µmol m$^{-2}$ s$^{-1}$, which was approximately six-fold lower than the intensity of full sunlight (Peri *et al.*, 2009). Although this may have had minimal effect on proteins that undergo light-dependent phosphorylation, enzymes that are regulated by more complex light-intensity mechanisms, like PPDK (Chen *et al.*, 2014), may have been less phosphorylated at key regulatory sites after the dark to light transition or visa-versa. Therefore, these phosphorylations may have been more difficult to detect because of their low abundance in the analyte. In addition to this, given that the rate of photosynthesis and activity of key enzymes could be affected by lower light intensities (Figure 3.26), subsequent enzyme kinetic experiments described in Chapter 4 and Chapter 5 were conducted using plants grown at 900 µmol m$^{-2}$ s$^{-1}$. 
Fig. 3.26. Model relationship between irradiance and rates of photosynthesis in C\textsubscript{4} plants. Low rates of photosynthesis are observed when light-intensity is low. Rates of photosynthesis increase with increasing light intensities. Schematic is adapted from experimental data collected by Milner and Hiesey (1964) and Bräutigam et al. (2014).

Aside from experimental difficulties for PTM assignment in plants, incomplete or poorly annotated databases affect protein identification (van Wijk, 2001; Abdallah et al., 2012).

In UniProtKB, there are currently 267 unreviewed protein entries and 17 reviewed entries for Setaria viridis, compared to 40,779 and 27 entries, respectively, in the Setaria italica database. Choosing Setaria italica protein sequences as the target database for this study was essential for the identification of proteins from MS/MS peptide spectra. Although peptide spectra from Setaria viridis could be searched against the well-annotated and larger Zea mays protein sequences, Setaria italica is more comparable due to high genome similarity (Bennetzen et al., 2012). Initial experimental data was searched using MASCOT and NCBI green plant protein sequences. While the advantage of using a larger database and filtering peptides using a narrow mass window may assist in increasing protein identification, database search times were substantially longer and did not yield significant results, as the probability of peptide spectra mismatches increased (Hsieh et al., 2010). There were similar issues with the Sorghum bicolor database in UniProtKB and resulting MS/MS peptide spectra were searched against the Zea mays protein database, which aided the identification of Sorghum bicolor proteins and subsequent phosphorylation site assignment.
3.3.2 Phosphorylation of phosphoenolpyruvate carboxylase

In *Zea mays* and *Sorghum bicolor*, the regulatory phosphorylation of PEPC occurs at Ser-15 and Ser-8 respectively (Jiao and Chollet, 1990; Jiao et al., 1991; Chollet et al., 1996; Vidal and Chollet, 1997). However, regulatory phosphorylation of PEPC at additional serine or threonine residues has not been described. Using a non-targeted approach in *Setaria viridis*, seven novel phosphorylation sites were identified (Figure 3.27). PEPC phosphorylation in *Setaria viridis* increased after illumination (Figure 3.10), but this may be due to increased content of PEPC through the photoperiod, since PEPC content in leaves can vary in response light intensity (Slack et al., 1969; Hatch, 1987). Also, increased PEPC content was only discernible from 2D gels using 80 µg of total protein (Figure 3.10) and not 1D gels using 10 µg of total protein (Figure 3.9) and phosphorylated PEPCs in darkened leaves may be in very low content to be detected by phosphoprotein staining.

Fig. 3.27. Phosphorylation of PEPC from *Setaria viridis*. Cleavage sites at lysine (K) and arginine (R) are highlighted in red. Phosphorylation sites indicated by asterisk. Active sites are indicated. Sequence annotated with FASTAnnotate.
The phosphorylation at Ser-11 (residues 9–17, **HHSIDAQLR**) in *Setaria viridis*, which corresponds to Ser-15 in *Zea mays*, occurred after illumination (Jiao and Chollet, 1988; Bailey *et al.*, 2007). The phosphorylation of Ser-11 in darkened leaves was only detected in a larger peptide (residues 2–17, **ASKPVEKHHSIDAQLR**), suggesting that a missed cleavage at Lys-8 could be affecting the retention of this larger phosphorylated peptide, as depicted in Figure 3.25. Furthermore, the identification of phosphorylation sites at Ser-25, Ser-180, Ser-378, Thr-379, and Ser-752 could be affected by the proximity of cleavage sites (Figure 3.27). In fact, all these phosphorylation sites were detected in phosphopeptides with two missed cleavages and it is possible that small phosphopeptides were lost during analysis.

**Fig. 3.28. Alignment of PEPCs from C₄ and C₃ plants.** PEPC and PEPC-2 (housekeeping enzyme) are aligned and the phosphorylation sites identified in *Setaria viridis* are shown. Phosphorylations identified in illuminated leaves highlighted yellow and phosphorylations found in both light and dark highlighted in blue and yellow. Similar residues highlighted green.
Ser-11 is conserved in major C₄ and C₃ plants (Figure 3.28) and it may have similar regulatory properties in *Setaria viridis* as it does in closely related C₄ plants. In addition, Ser-11 may be a substrate for PEPC kinase. The non-photosynthetic PEPC in C₃ plants undergoes regulatory phosphorylation *in vitro* and *in vivo* by the light- and Ca²⁺-dependent PEPC kinase at the invariant N-terminal serine (Wang and Chollet, 1993; Duff and Chollet, 1995; Li *et al.*, 1996; Zhang and Chollet, 1997), which suggests that these regulatory mechanisms are conserved in plant species, but whether these regulatory mechanisms extend to other highly conserved residues like Ser-25 or Thr-379 is unknown. The position of conserved positively charged lysine and arginine residues around putative phosphorylation sites could also act as recognition binding sites for PEPC kinase (Ueno *et al.*, 1997).

*Fig. 3.29. Zea mays PEPC annotated with phosphorylation identified in Setaria viridis.* Annotated residues (red markings) correspond to the phosphorylated sites identified in *Setaria viridis*. C1 indicates the region important for catalysis and PEP binding. Cyan lines and markers indicate active sites. Green line indicates the phosphorylated Ser-378 residue in *Setaria viridis*, which corresponds to Asn-384 in *Zea mays*. Blue tubes are alpha helices and yellow arrows show beta sheets. Protein structure taken from Matsumura *et al.* (2002).
Using the PEPC structure from *Zea mays*, the positions of the identified phosphorylated residues were determined (Figure 3.29). Using this model, residues Asn-384, Thr-385 and Ser-708 (corresponding to Ser-378, Thr-379 and Ser-702 in *Setaria viridis*), are exposed and could be readily phosphorylated by kinases, whereas residues Ser-185 and Ser-758 (corresponding to Ser-180 and Ser-752 in *Setaria viridis*) are less accessible. In *Arabidopsis thaliana*, phosphorylation of the C₃-form PEPC occurs at Ser-704 (de la Fuente van Bentem *et al.*, 2008), which corresponds to Ser-702 in *Setaria viridis*, but this residue is not known to be phosphorylated in C₄ species, and could be a determinant for C₄-specific PEPC function as previously found for Ser-774 in *Flaveria* species (Bläsing *et al.*, 2000; Jacobs *et al.*, 2008).

Fig. 3.30. Schematic model showing the active site of PEPC in higher plants. The phosphate and carboxylate of PEP interacts with positively charged lysine (Lys) and arginine (Arg) residues. Active site in *Setaria viridis* at His-172. Figure taken from Andreo *et al.* (1987).
More importantly, Ser-752 is conserved in all PEPCs and is near the active site at His-177 in Zea mays or His-172 in Setaria viridis, suggesting phosphorylation could have direct interaction with the carboxylation reaction (Kai et al., 2003). Furthermore, the cavity between His-177 and Arg-647, shown as C1 in Figure 3.29, allows for movement of PEP during catalysis (Andreo et al., 1987; Matsumura et al., 2002). It is therefore possible that the negative charge from a nearby phosphorylated residue (Ser-752) may counteract the positive charge of Arg-647 or Lys-606 (Andreo et al., 1987; Scheeff et al., 2009). Moreover, the phosphorylation of Ser-752 was only identified in darkened leaves. Phosphorylation at Ser-752 could influence the way that PEP interacts with the catalytic site. A phosphate interacting with the positively charged lysine and arginine residues at the catalytic site might also alter the shape of the catalytic pocket and hinder the entrance of Mg$^{2+}$ and PEP into the active site or disrupt the interaction with histidine and cysteine residues, as shown in Figure 3.30. This, together with the identification solely in the dark period, suggests that phosphorylation at Ser-752 might correlate with PEPC inactivation in darkened leaves and regulate PEP binding during the dark period. This region of PEPC may also contribute to C$_4$-specific function and contain regulatory properties (Bläsing et al., 2000).

Additionally, the side chain of Ser-185 in Zea mays (Ser-180 in Setaria viridis) may interact with the side chains of Arg-182, Arg-183, Arg-231 and Arg-372 during the allosteric binding of glucose-6-phosphate (Bläsing et al., 2000; Matsumura et al., 2002; Kai et al., 2003). If Ser-185 (Zea mays numbering) is phosphorylated, the negative charge of the phosphate group could repel the phosphate from glucose-6-phosphate, thus affecting its binding to PEPC. Ser-180 is also conserved in all C$_3$ and C$_4$ PEPCs and in Setaria viridis, Ser-180 was phosphorylated in darkened and illuminated leaves. It is therefore possible that phosphorylation at Ser-180 is a necessary structural component to regulate activator binding, rather than prevent or promote binding at one specific time-point.
3.3.3 Phosphorylation of pyruvate, phosphate dikinase

Unlike PEPC, the regulatory phosphorylation of PPDK at Thr-462 in *Setaria viridis* occurred in both darkened and illuminated leaves, as evident from phosphoprotein stained gels (Figure 3.9 and Figure 3.10). In the phosphoprotein stained 1D gel (Figure 3.9), the PPDK band intensity from darkened leaves appeared to be greater than PPDK from illuminated leaves. This was confirmed by relative quantitation of Thr-462, which showed that the phosphopeptide was in higher amounts in darkened leaves. Unlike PEPC, there is little indication whether the content of PPDK varied between darkened and illuminated conditions, but 2D gels (Figure 3.10) suggest that there was a higher PPDK content per 80 µg of total protein in darkened leaves than in illuminated leaves. The phosphorylation of Thr-462 in darkened leaves is consistent with the regulatory site previously identified in *Zea mays* (Ashton and Hatch, 1983; Chastain *et al.*, 2000; Chastain and Chollet, 2003; Chen *et al.*, 2014). Moreover, the higher degree of phosphorylation at Thr-462 in darkened leaves is consistent with the inactivation mechanism by phosphorylation (Budde *et al.*, 1985).

The phosphorylation of Ser-463 was also identified in darkened leaves of *Setaria viridis* (Figure 3.6), but with a lower localisation probability (\(P = 0.78\)) when compared to the phosphorylation at Thr-462 (\(P = 0.99\)). Additionally, the phosphorylation at Ser-463 was also identified in *Megathyrsus maximus* (\(P = 0.50\)) in illuminated leaves. However, despite Ser-463 being less accessible for phosphorylation, it is possible that Ser-463 in *Setaria viridis*, like in leaves of *Zea mays*, might be a substrate for PPDK-RP (Chen *et al.*, 2014). In a previous study, the phosphoserine was identified in both active and inactive preparations of PPDK, while the phosphothreonine was only present in the inactive PPDK and although Chen *et al.* (2014) identified Ser-528 as a substrate for PPDK-RP, the phosphorylation at Ser-463 in *Setaria viridis* may be the result of a non-enzymatic, nitrogen to oxygen phosphoryl migration (Plapinger and Wagner-Jauregg, 1953; Rathlev and Rosenberg, 1956; Budde *et al.*, 1985).
Fig. 3.31. Alignment of PPDKs from C₄ and C₃ plants. Phosphorylation sites identified in *Setaria viridis* are indicated. Phosphorylations identified in illuminated leaves highlighted yellow and those found in darkened leaves highlighted blue. Phosphorylations found in both light and dark highlighted in blue and yellow. Similar residues highlighted green. Substrate binding sites highlighted in purple. His-364 (*Setaria italica* numbering) is the active site. Green triangle represents phosphorylation of Ser-528 identified in *Zea mays*.

Thr-704, Thr-721 and Thr-753 are conserved across C₄ and C₃ plants (Figure 3.31), however there is no indication that these residues are phosphorylated. Moreover, the proximity of Thr-753 to the metal binding site (Glu-756) could affect binding mechanisms. Of the phosphorylation sites identified, Thr-462, Thr-557 and Thr-753 are close to the catalytic site of PPDK and may influence catalytic properties (Figure 3.32). Thr-593, Thr-721, Thr-704 and possibly Ser-394 are exposed to the surface of the protein and may be targets for phosphorylation. However, whether any of these putative sites are regulated cannot be determined from these data.
Fig. 3.32. Structure of PPDK from *Flaveria trinervia* reveals the position of phosphorylation sites (in red) identified in *Setaria viridis*. Green rectangle indicates the main catalytic domain of PPDK, including the position of the catalytic site at His-464 and Cys-842 (not labelled). Blue tubes are alpha helices and yellow arrows show beta sheets. PPDK structure taken from Minges *et al.* (2017).

In the crystal structure of the C₄ PPDK, the phosphorylation at Thr-753 may directly affect catalysis by interacting with the PEP binding site, which is indicated by the green rectangle on Figure 3.32. In addition, the phosphorylation at Thr-753 was only detected in illuminated leaves and might hinder PEP binding, thus limiting the reverse reaction towards the formation of pyruvate. This could regulate selectivity towards pyruvate binding and not PEP binding during illumination, when PPDK is needed to regenerate PEP. Furthermore, in *Zea mays*, binding of PEP during phosphate transfer results in a swivelling-domain motion at the C-terminal active site (Nakanishi *et al.*, 2005). During
this motion, PEP interacts with surrounding residues and may be easily be influenced by proximal negatively charged moieties (Minges et al., 2017). The phosphorylation at Thr-557 could also influence the binding interaction between PEP and PPDK, but Thr-557 is not as close to the catalytic site as Thr-753.

3.3.4 Phosphorylation of NADP-malic enzyme

The regulatory phosphorylation of NADP-malic enzyme has not been described in plants, but phosphorylation has been described in mice (Mus musculus) (Huttlin et al., 2010). In this study, 20 putative phosphorylation sites were identified. Although four isoforms were identified, protein K3XG11 was 53-times more abundant in darkened leaves than protein K3XFW4 and six-times more abundant than K3ZRI5 (Table 3.5). However, only two phosphorylation sites were identified on the most abundant protein isoform (Figure 3.17). The phosphorylation at Ser-373 had been previously identified in Zea mays using a similar proteomic approach (I. Abreu, personal communication). Ser-343 is conserved in three isoforms in Setaria viridis (Figure 3.33). Aside from NAD-malic enzyme from Zea mays and Arabidopsis thaliana, this serine residue is conserved in both NAD- and NADP-malic enzyme of C3 and C4 origin (Figure 3.33). Not all the phosphorylation sites were identified on every protein isoform detected in Setaria viridis leaf lysates (Figure 3.34). However, because of high sequence homology, it is possible that some phosphopeptides were inferred from different isoforms when in fact they were only from a single isoform. Based on sequence similarity to the main NADP-malic enzyme isoform from Zea mays and molecular masses, proteins K3XFH6 and K3XG11 might be the primary NADP-malic enzyme isoforms in Setaria viridis and correspond to the two major C4-specific isoforms previously identified in Zea mays (Maurino et al., 1996).
Fig. 3.33. Alignment of NAD- and NADP-malic enzymes across plant (C4 and C3), mammalian and avian origins. Putative phosphorylation sites identified in illuminated leaves highlighted yellow and those found in darkened leaves highlighted blue. Phosphorylations found in both light and dark highlighted in blue and yellow. Highly conserved residues highlighted green. Metal binding sites highlighted in red. Red triangle indicates phosphoserine-336 identified in Mus musculus and Ser-346 in Columbia livia (discussed below). Blue arrows indicate NADP+ interaction sites.
In mice (*Mus musculus*), NADP-malic enzyme is phosphorylated at Ser-336 (Ser-373 in *Setaria italica* numbering), though its phosphoregulatory properties have not been described (Huttlin *et al.*, 2010). In addition, studies have suggested that Asp-335 is selective towards NADP+ co-factor binding rather than NADP+, because the negative charge of aspartate repels the negative charge of 2′-phosphate of NADP+ (Wierenga *et al.*, 1986; Scrutton *et al.*, 1990; Yang *et al.*, 2002). Although it was shown that this interaction may not occur since Asp-335 interacts with Arg-354, the presence of phosphate groups from nearby phosphorylated residues may confer some disadvantages to NADP+ binding. Therefore, the phosphorylation at Ser-373 might conflict with NADP+ binding since the 2′-phosphate might interact with the phosphoserine-373 as has been shown in the cytosolic-form of NADP-malic enzyme from pigeon (*Columba livia*) (Yang *et al.*, 2002). The phosphorylation at Ser-373, which only occurred in darkened leaves, may be part of a mechanism to reduce the activity of NADP-malic enzyme during the dark period when CO2 fixation is not required. The phosphorylation of Ser-379 and Ser-380 on the same phosphopeptide from protein K3FW4 may cause additional charge differences and disrupt NADP+ binding (Figure 3.34). In addition to this, one protein isoform identified in *Setaria viridis* (K3ZRI5) and NAD-malic enzyme from *Zea mays* and *Arabidopsis thaliana* have an alanine residue instead of serine in position 373, which cannot undergo phosphorylation. This substitution might also indicate preference towards NAD+ binding in the NAD-malic enzyme from *Zea mays* and *Arabidopsis thaliana*. This assumption would also indicate that the identified malic enzyme in *Setaria viridis* (K3ZRI5) might be NAD- rather than NADP-dependent, but despite inferred to be NAD+-dependent on the UniProt protein database, there is no other evidence to verify this. In addition to Ser-346 (pigeon numbering), NADP+ interaction with the side-chain ammonium group of Lys-362 is essential for NADP-dependent enzymatic activity (Yang *et al.*, 2002). It was previously shown that the pigeon NADP-malic enzyme with a lysine to alanine substitution at position 362 (K362A) drastically decreased the affinity for NADP+ (Kuo *et al.*, 2000). All the proteins that do not have a serine at the first NADP+ binding site, indicated above Ser-373 by the blue arrow on Figure 3.33, have an alanine at the second NADP+ binding site. Since Lys-362 serves an important role in determining NADP+ selectivity in NADP-malic enzyme from pigeon (Yang *et al.*, 2002), plant proteins lacking the lysine residue at this position may have preference for NAD+ binding instead. Both residues at these positions may be determinants for NAD-specific malic enzyme activity.
Fig. 3.34. Position of phosphorylation sites identified in *Setaria viridis*. Cyan rectangle (C1) indicates the catalytic centre of NADP-malic enzyme. Blue tubes are alpha helices and yellow arrows show beta sheets. Phosphorylation sites (*Setaria italica* numbering) are highlighted red on the protein. The labelled residues are the phosphorylation sites identified on each protein isoform. Only Ser-429 (K3XFH6) and Ser-380 (K3XFW4) occur on both proteins. Protein model based on the pigeon cytosolic NADP-malic enzyme taken from Yang *et al.* (2002).

The position of Ser-373, Ser-379 and Ser-380, shown in Figure 3.34, further demonstrates show this cluster of phosphorylations could be affecting NADP⁺ binding during the dark period. The phosphorylation at Ser-379 and Ser-380 could potential cause a charge repulsion with NADP⁺ as it interacts with Ser-373 during catalysis. Although not shown in Figure 3.34, the phosphorylations at Ser-302, Thr-313 and Ser-360 might disrupt binding of the metal ion at the aspartate binding site (Figure 3.33). These
phosphorylation sites are the first indication of a phosphoregulatory mechanism of the C₄-specific NADP-malic enzyme from *Setaria viridis*. Recent research using a recombinant NADP-malic enzyme from *Zea mays*, suggests that NADP-malic enzyme is inactivated by phosphorylation in the dark period (V.G. Maurino, personal communication). In *Setaria viridis*, Ser-373, which corresponds to Ser-419 in *Zea mays*, is possibly the most likely target for regulatory phosphorylation because of its role in the NADP⁺ binding mechanism (Kuo *et al.*, 2000; Yang *et al.*, 2002) and previous identification in mice (Huttlin *et al.*, 2010). However, validation of this phosphorylation in *Setaria viridis* or identification in *Sorghum bicolor* has been unsuccessful. The recovery of the NADP-malic enzyme containing the phosphorylation at Ser-373 was very low, relative to the other isoforms identified (Table 3.5). However, the relative abundance coefficient (iBAQ) does not accurately quantify all isoforms expressed *in vivo* and is rather an approximation. Increasing the recovery of phosphopeptides of protein K3XFW4 may be the first step in future studies for validating the phosphorylation at Ser-373 in *Setaria viridis*.

3.3.5 Phosphorylation of aspartate and alanine aminotransferase

Three phosphorylation sites were solely identified in aspartate aminotransferase from illuminated leaves of *Setaria viridis* (Figure 3.35A), which could be regulating activity during the photoperiod. Aspartate aminotransferase is phosphorylated at Ser-49 and Ser-149 in rat (*Rattus norvegicus*) (Lundby *et al.*, 2012), though these serine residues are not completely conserved in plant species. *Setaria italica* aspartate aminotransferase (K3YSB2) could potentially be phosphorylated at Ser-73 (corresponding to Ser-149 in rat), but there is no evidence of this phosphorylation. Ser-300 and Ser-341 from protein K3YSB2 are highly conserved in both C₃ and C₄ species and could be potential targets for regulatory phosphorylation, while amino acid substitutions at these positions could indicate specific isoform activity (Figure 3.35A). In C₄ plants there are at least two major aspartate aminotransferase isoforms, compared to three that were identified in *Arabidopsis thaliana* (Numazawa *et al.*, 1989; Wilkie and Warren, 1998). Having identified four isoforms in *Setaria viridis*, it is difficult to ascertain which ones have C₄-specific function. In addition, there could be differences in the regulatory properties
between the aspartate aminotransferase isoform localised in the chloroplasts and mitochondria (Hatch and Mau, 1973; Meister et al., 1996), which cannot be ascertained using whole leaf protein extracts.

**Fig. 3.35A. Alignment of aspartate aminotransferases from C₄ and C₃ plants.** Putative phosphorylation sites identified in illuminated leaves highlighted yellow and those found in darkened leaves highlighted blue. Highly conserved residues highlighted green.

Conversely, alanine aminotransferase was predominantly phosphorylated after illumination Figure 3.35B). If the degree of phosphorylation (total number of phosphorylation sites identified) were important for enzyme activity, then it would suggest that the phosphorylation mechanisms are different for alanine and aspartate aminotransferase. While it appeared that Thr-167, Ser-189 and Ser-206 are not light-dependent, because they were present in both darkened and illuminated leaves, phosphorylation at these residues could contribute to enzymatic activity or stability. Unlike that phosphorylation sites identified in aspartate aminotransferase, those identified in alanine aminotransferase are highly conserved in *Zea mays* and C₃ species, except for Ser-326 (Figure 3.35B). Whether the cysteine residue in position 326 (K3ZSX0 numbering) contributes to additional regulatory mechanisms, such as redox regulation, is unknown. Despite catalysing the amino transfer in two similar reactions, protein sequences between the aspartate and alanine aminotransferase are very dissimilar, probably due to high specificity for either aspartate or alanine, respectively (Ashton et al., 1990; Duff et al., 2012). Moreover, the optimal activity of aspartate and alanine aminotransferase might directly be associated with upstream enzymes like PEPC.
or NADP-malic enzyme, respectively (Andrews et al., 1971; Hatch and Mau, 1973). In addition to this, alanine aminotransferase is not associated with chloroplasts (Meister et al., 1996) and the regulatory properties of the identified isoforms might be more similar as opposed to the aspartate aminotransferase isoforms, which may be differently compartmentalised.

**Fig. 3.35B. Alignment of alanine aminotransferases from C_4 and C_3 plants.** Putative phosphorylation sites identified in illuminated leaves highlighted yellow and those found in darkened leaves highlighted blue. Phosphorylations found in both light and dark highlighted in blue and yellow. Highly conserved residues highlighted green.

The activities of major aspartate and alanine aminotransferase isoforms in C_4 plants are higher in dark-grown plants, while minor isoforms differ very little in activity between darkened and illuminated conditions (Hatch and Mau, 1973). However, from the phosphorylation data, it is not possible to ascertain whether the changes to the degree of phosphorylation in darkened and illuminated conditions might be regulating the activity of aspartate or alanine aminotransferase. In an early study on the activity of aspartate and aminotransferase in C_4 plants, Hatch and Mau (1973) suggest that the aminotransferases evolved C_4-specific function, possibly after the C_4 cycle was fully established. Also, the activity of these enzymes is linked to decarboxylation by PEPCK in bundle sheath cells or NAD-malic enzyme in mitochondria in bundle sheath cells, and not with C_4 subtypes.
that predominantly use NADP-malic enzyme (Andrews et al., 1971; Huber and Edwards, 1975; Chapman and Hatch, 1981; Meister et al., 1996). Therefore, it is possible that aspartate aminotransferase activity evolved together with downstream PEPCK activity and the phosphoregulatory mechanisms that evolved with PEPCK could have been adapted to regulate the activity of aspartate aminotransferase (Gowik and Westhoff, 2011). PEPCK, unlike PEPC, is inactivated by the degree of phosphorylation in dark-grown plants (Bailey et al., 2007). Alternatively, regulatory elements could have evolved along with the regulatory phosphorylation of PEPC activation, as aspartate aminotransferase could have evolved C₄-specific activity in order to increase the options for carbon flux during high rates of carboxylation by PEPC (Bräutigam et al., 2014; Wang et al., 2014).

3.3.6 Phosphorylation of enolase and phosphoglycerate mutase

Although not directly associated with the C₄ pathway, the activity of phosphoglycerate mutase and enolase in leaves of C₄ plants, particularly in Zea mays, has been considered to contribute for the supply of 3-PGA and PEP (Huber and Edwards, 1975; Furbank and Leegood, 1984). However, the rate at which this interconversion occurs and how it may be regulated is unknown (Arrivault et al., 2016). In this study, novel light-dependent phosphorylation sites identified in enolase and phosphoglycerate mutase from Setaria viridis suggest that the formation of PEP through enolase and phosphoglycerate mutase is regulated. In yeast (Saccharomyces cerevisiae), enolase undergoes phosphorylation at Ser-10 (Pearlman et al., 2011), but this residue is not conserved in plants (Figure 3.36).
Enolase catalyses the interconversion of 2-PGA and PEP (Zhang et al., 1997) and its activity might provide an alternative route for the formation of PEP. The phosphorylation at Ser-129 occurred only in enolase-2 in Setaria viridis, though enolase-2 from Zea mays and Triticum aestivum could potentially be phosphorylated (Figure 3.36). This might be indicative of a specific mechanism regulating the activity of enolase-2 in Setaria viridis.

In addition, despite Thr-334 and Thr-337 being conserved in enolase-1 and enolase-2, only enolase-1 underwent phosphorylation in illuminated leaves in Setaria viridis. Because of the proximity to the binding site (Asp-329) and active site (Lys-354), phosphorylation at these two residues might hinder substrate binding by charge repulsion with the phosphate of 2-PG or PEP.
Fig. 3.37. Alignment of phosphoglycerate mutase protein sequences. Novel phosphorylation sites indicated in yellow (occurring in light) and yellow and blue (occurring in light and dark). Red triangle shows phosphoserine-74 in *Trypanosoma brucei*.

Phosphoglycerate mutase catalyses the interconversion of 2-PGA and 3-PGA and might serve as a direct link with the activity of enolase and formation of PEP (Huber and Edwards, 1975). While Ser-520 and Thr-527 were only identified in *Setaria viridis*, the phosphorylation at Ser-109 was detected in illuminated and darkened leaves of *Setaria viridis* (Table 3.4B, 3.6), *Sorghum bicolor* (Table 3.7) and *Megathyrsus maximus* (Table 3.8), suggesting that it serves an important role in the catalytic mechanism. In fact, this residue is important for catalysis and exists as a phosphointermediate in several non-plant species (Collet et al., 2001; Jedrzejas and Setlow, 2001; Potters et al., 2003; Rigden et al., 2003). Based on sequence homology, it is possible that the phosphorylation site identified in this study might be part of the reaction mechanism rather than a phosphoregulatory site.

<table>
<thead>
<tr>
<th>Phosphoglycerate Mutase</th>
<th>Setaria italica</th>
<th>Zea mays</th>
<th>Sorghum bicolor</th>
<th>Arabidopsis thaliana</th>
<th>Triticum aestivum</th>
<th>Oryza sativa</th>
<th>Escherichia coli</th>
<th>Trypanosoma brucei</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C₄</strong></td>
<td>DDDMGNSEVGHNLGAGRIFAQG</td>
<td>DDDMGNSEVGHNLGAGRIFAQG</td>
<td>DDDMGNSEVGHNLGAGQIYAQG</td>
<td>EDDMGNSEVGHNLGAGRIFAQG</td>
<td>DDDMGNSEVGHNLGAGRIFAQG</td>
<td>DDDMGNSEVGHNLGAGQIYAQG</td>
<td>-RQMNSEVGHVNLGAGRIVYQD</td>
<td>DADMGNSEVGHNLGAGRVLQG</td>
</tr>
<tr>
<td><strong>S109</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C₃</strong></td>
<td>Setaria italica</td>
<td>Zea mays</td>
<td>Sorghum bicolor</td>
<td>Arabidopsis thaliana</td>
<td>Triticum aestivum</td>
<td>Oryza sativa</td>
<td>Escherichia coli</td>
<td>Trypanosoma brucei</td>
</tr>
<tr>
<td></td>
<td>P LLDKGGIGIQLTLSHTLQP</td>
<td>P LLDKNGRLIQILTSHTLQP</td>
<td>P LRDKGKGVQTLTSHTLN</td>
<td>P ALDKEGKLQILTSHTLKP</td>
<td>P MLDKGSGIQLTLSHTLQP</td>
<td>P LRDKGNNVQPLTSHTLN</td>
<td>-GQAHTTAHTNL</td>
<td>P VRDAGEGNLMPLTSHTLAP</td>
</tr>
</tbody>
</table>
In *Trypanosoma brucei*, phosphoglycerate mutase undergoes phosphorylation at Ser-74 and because of the low rates of formation and disappearance during the reaction, the phosphoserine is produced as an intermediate of the phosphatase reaction when phosphate is removed from either 2-PGA or 3-PGA (Jedrzejas *et al.*, 2000; Collet *et al.*, 2001). Despite being distant to plants, it has been suggested that trypanosomes retain several plant-like genes, most of which are involved in central metabolism and at some point, in their evolutionary past, trypanosomes contained chloroplasts (Hannaert *et al.*, 2003; Martin and Borst, 2003). However, sequence similarity between plant and *Trypanosoma brucei* phosphoglycerate mutase (Figure 3.37) does not necessarily imply that the phosphoregulatory elements in the protozoa are retained in plant species. Due to the high reproducibility of Ser-109 across three of the investigated C₄ plants, the phosphorylation at Ser-109 might be a phosphointermediate or at least an important residue essential for catalysis. It would be interesting to investigate whether the replacement of serine to alanine would affect the catalytic properties of phosphoglycerate mutase in plants, as previously investigated in *Bacillus stearothermophilus*, which drastically reduced enzyme activity (Jedrzejas *et al.*, 2000; Jedrzejas and Setlow, 2001).

### 3.3.7 Quantitative proteomics and future study

In this study, several novel phosphorylation sites were identified on key C₄-related proteins, suggesting that multiple sites, in addition to those already identified, might regulate enzymatic activity. However, it is unlikely that all phosphorylation sites identified are under regulation. To ascertain the sites that are under regulation, the degree of phosphorylation and how it changes during dark to light transitions requires further investigation. This can be done by either quantifying the signal intensity of phosphoprotein stained gels or using quantitative proteomics. One method commonly used to quantify peptides by mass spectrometry is using stable isotope labelling by amino acids in cell culture (SILAC) in conjunction with immunoprecipitation (Zhang and Neubert, 2009; Matthes *et al.*, 2014), however, the incorporation of heavy amino acids into proteins has only recently been optimised for *Arabidopsis thaliana* seeds and may not be compatible with seeds of other plant species (Lewandowska *et al.*, 2013). An alternative to metabolic labelling is using isobaric tags for relative and absolute
quantitation (iTRAQ), whereby peptides are chemically tagged in vitro (Wiese et al., 2007), however is not as reproducible as SILAC (Zhang and Neubert, 2009).

Having identified putative phosphopeptides, selected reaction monitoring (SRM) is more applicable for quantifying phosphorylation than using metabolic labelling. In complex mixtures, particularly in untargeted approaches, the analysis of proteins of interest can be overwhelmed by more abundant proteins, resulting in poor consistency in fragmentation spectra and detection of the associated peptides of interest (Lange et al., 2008). In SRM, the peptide of interest can be selected by defining a search range given its corresponding m/z value, thus the elution of ions can be monitored in the run and interfering ions, which are out of the pre-defined monitoring range, can be filtered out (Lange et al., 2008). These analyses are performed using triple quadrupole mass spectrometers, which can monitor the precursor ion and fragment ion of a given peptide over time (Williamson et al., 2006), however, unlike tandem mass spectrometers, SRM-capable mass spectrometers utilise a non-scanning approach, which by not recording all fragmentation spectra detected, the targeted peptide can be preferentially observed (Lange et al., 2008). This allows for the selection of ions of interest in a complex mixture, which can then be accurately quantified. However, for SRM to be effective for PTM identification and quantification, the m/z value of the peptide must be known.

It is possible that there are additional phosphorylation sites that have not been identified in this study due to the loss of peptides of interest during analysis or low amount of phosphopeptides due to suboptimal concentration of proteins from the leaf lysate. One method to increase the concentration of phosphopeptides is by preparing peptides using a gel-free approach. However, the heterogeneity of the protein suspension and nature of globular proteins in their native three-dimensional state makes it difficult to guarantee that the trypsin-mediated proteolysis has been equally applied to every protein in the sample (Zhang et al., 2013). Also, despite trypsin is widely used, the dynamic mechanism of proteolytic digestion does not guarantee complete coverage of the protein sequence and peptides are readily lost during analysis (McLachlin and Chait, 2001), making it difficult to map all modifications on the protein of interest. To overcome this, intact proteins can be analysed, thereby reducing the loss of peptides of interest. However, fragmentation efficiencies of intact proteins are lower than that of peptides, compromising on the fragment ion spectra needed for subsequent mass spectrometry analysis and determination of site-specific modifications (Moradian et al., 2014).
addition to this, fractionating intact proteins by SDS-PAGE is hindered by reduced compatibility with non-volatile buffers required for the extraction from the polyacrylamide matrix (Garcia, 2010; Zhang et al., 2013). Intact proteins can also be enriched using phosphoprotein enrichment, however, the binding interaction between an intact, folded phosphoprotein and resin particle in the enrichment column is difficult to predict, thus making phosphoprotein enrichment less reliable than phosphopeptide enrichment (Fíla and Honys, 2012). Alternatively, larger polypeptides, instead of intact proteins, can be analysed using size-dependent peptide fractionation. In this method, proteins are partially digested to form large polypeptide fragments, which are then selectively fractionated by size (Garcia, 2010). However, this approach is not optimised for large-scale phosphoproteomic studies (Zhang et al., 2013). Despite being several methods to overcome the difficulties of phosphoproteomics, contaminant peptides from human proteins, co-migrating proteins or from the autolysis of trypsin can still pose challenges (Seo and Lee, 2004).

3.3.8 Conclusion

It was hypothesised that C4-related proteins in *Setaria viridis* undergo light-dependent phosphorylation similar to the mechanisms elucidated for PEPC and PPDK in *Zea mays*. The findings in this chapter verified that the PEPC and PPDK undergo light-dependent phosphorylation at the conserves residues as in *Zea mays*, and that the degree for phosphorylation of PPDK increased after the transition into darkness, as previously observed (Chao et al., 2014). In addition, several other C4-related enzymes such as NADP-malic enzyme, were phosphorylated, supporting the hypothesis that phosphoregulatory mechanisms might regulate the activity of other key C4-related proteins. While the exact effect of these modifications cannot be ascertained from this data alone, structural evidence has shown that the phosphorylation of NADP-malic enzyme at Ser-373, along with Ser-379 and Ser-380 in darkened leaves might regulate NADP⁺ co-factor binding when there is no requirement for NADP-malic enzyme activity. This region of NADP-malic enzyme in C4 plants might contain selective factors for either NAD⁺ or NADP⁺ co-factor binding, and therefore could be a prime region for post-translational regulation.
Chapter 4 – The kinetics of C₄-acid decarboxylation by NADP-malic enzyme in response to pH and illumination in C₄ grasses

4.1 Introduction

4.1.1 Regulation of NADP-malic enzyme in plants

NADP-malic enzyme (EC 1.1.1.40) is one of the three decarboxylases in C₄ plants, along with NAD-malic enzyme and PEPCK. It is widespread in higher plants, including Crassulacean acid metabolism (CAM) plants, and present in various tissues including leaves, seeds and roots (Edwards and Andreo, 1992; Drincovich et al., 2001). NADP-malic enzyme is found in bundle sheath chloroplasts of C₄ plants and catalyses the oxidative decarboxylation of L-malate (Figure 4.1), forming pyruvate and releasing CO₂ at the site of Rubisco (Slack et al., 1969; Hatch, 1987; Maurino et al., 1997; Drincovich et al., 2001).

$$\text{L-malate} + \text{NADP}^+ \leftrightarrow \text{pyruvate} + \text{NADPH} + \text{CO}_2$$

Fig. 4.1. Oxidative decarboxylation of L-malate via NADP-malic enzyme.

Aside from its essential role in C₄ photosynthesis, cytosolic, non-photosynthetic NADP-malic enzyme isoforms in C₄ plants are involved in cellular defence mechanisms, nitrogen assimilation and supplying reducing equivalents for use in central metabolism, much like their counterparts present in C₃ plants (Drincovich et al., 2001; Tausta et al., 2002; Maier et al., 2011; Badia et al., 2015; Ludwig, 2016). The C₄-specific function of NADP-malic enzyme, with regards to its subcellular compartmentalisation in bundle sheath chloroplasts and activation by illumination (Lai et al., 2002), was acquired from an ancestral, chloroplastic NADP-malic enzyme following a gene duplication event (Monson, 1999; Tausta et al., 2002). In the Flaveria species, which has C₃ and C₄ plant variants, NADP-malic enzyme belongs to a small gene family, which encodes its expression in the cytosol and chloroplasts (Rajeevan et al., 1991). Genes encoding the chloroplastic NADP-malic enzyme protein found in Flaveria trinervia (C₄) and Flaveria
tringlei (C₃) were shown to have high similarity and likely to have evolved from a common ancestor (Lipka et al., 1994; Marshall et al., 1996). The C₃-specific NADP-malic enzyme, the likely ancestor to the C₄-specific form, is also present in chloroplasts of C₃ plants, but is expressed in low amounts (Maurino et al., 1997; Drincovich et al., 1998).

Despite having multiple functions, photosynthetic NADP-malic enzyme isoforms in bundle sheath chloroplasts of C₄ plants evolved a specialised role in decarboxylase activity to suit the demands of malate metabolism in C₄ photosynthesis (Maurino et al., 1996; Maurino et al., 1997; Drincovich et al., 2001; Tsuchida et al., 2001). In Zea mays, NADP-malic enzyme isoforms found in the cytosol and chloroplasts show distinct kinetic properties and are regulated differently by pH and malate (Johnson and Hatch, 1970; Nishikido and Wada, 1974; Asami et al., 1979; Edwards and Andreo, 1992; Maurino et al., 1996; Tausta et al., 2002). When compared to C₃ relatives, NADP-malic enzyme in C₄ plants has a higher affinity for L-malate and operates at a higher pH optimum (Nishikido and Wada, 1974). Furthermore, in Zea mays, the activity of NADP-malic enzyme is dependent on pH, which alters its sensitivity to inhibition by L-malate (Asami et al., 1979). At pH 8.0, NADP-malic enzyme is in its active homotetramer form consisting of identical 62 to 68 kDa subunits, and has a high $V_{\text{max}}$ and a high affinity for NADP⁺ during illumination (Edwards and Andreo, 1992; Kanai and Edwards, 1999). The active form dissociates into homodimers at pH 7.0 (Iglesias and Andreo, 1990), which decreases its affinity for L-malate and NADP⁺ and becomes more sensitive to allosteric inhibition by high malate concentrations (Asami et al., 1979; Edwards and Andreo, 1992; Kanai and Edwards, 1999; Detarsio et al., 2007). While NADP-malic enzyme remains active in all oligomeric forms, it is most active at pH 8.0 (Edwards and Andreo, 1992). Holaday and Lowder (1989) suggest that the weak acidity of the malate binding site promotes catalysis in vivo at the high optimal pH, but Michaelis-Menten kinetics are lost at pH 9.0, where some degree of cooperative interaction is observed (Holaday and Lowder, 1989; Edwards and Andreo, 1992).
4.1.2 Regulation by pH and illumination

The relationship between pH and illumination may be part of a more complex mechanism for the regulation of NADP-malic enzyme-dependent decarboxylation in C_4 plants. The C_4-specific NADP-malic enzyme has a lower affinity for L-malate at an alkaline pH, albeit activity increases with increasing pH (Edwards and Andreo, 1992). In addition, the effect of pH on the activity of NADP-malic enzyme varies with the L-malate concentration (Johnson and Hatch, 1970). It is also evident that pH and Mg^{2+}, both of which are factors that regulate the activity of NADP-malic enzyme (Johnson and Hatch, 1970; Asami et al., 1979), increase in the chloroplast stroma when illuminated (Edwards and Andreo, 1992). During illumination, the pH in the thylakoid space decreases and increases in the stroma, due to the proton flux into the former (Neumann and Jagendorf, 1964; Heldt et al., 1973; Werdan et al., 1975). In darkened conditions, pH in the stroma decreases, which consequently inhibits CO_2 fixation, and is considerably a necessary mechanism to make sure carbohydrate oxidation occurs exclusively in illuminated conditions (Wer dan et al., 1975). This oscillation of pH in darkened and illuminated conditions may be playing a much greater role in the activation of NADP-malic enzyme in C_4 plants; and being chloroplastic (Slack et al., 1969), it is conceivable that the activity is regulated by light-dependent changes in the stromal pH (Asami et al., 1979; Edwards and Andreo, 1992).

Apart from its regulation by pH, the activity of NADP-malic enzyme is regulated by changes in thiol-disulphide bonds in Zea mays leaves (Drincovich and Andreo, 1994; Alvarez et al., 2012; Saigo et al., 2013). The \textit{in vitro} activity of NADP-malic enzyme from dark-grown plants can be stimulated in the presence of DTT (reducing agent), and under such conditions, the activity in darkened leaves is close to the activity of the light-form (Drincovich and Andreo, 1994). However, apart from the light-dependent regulation by pH or redox regulation, more extensive post-translational mechanisms that modulate enzyme activity, like those described for other C_4-related proteins, have not been described of NADP-malic enzyme in C_4 grasses. For instance, the phosphoregulatory properties of PEPCK and its involvement in C_4-acid decarboxylation have been extensively studied in C_4 plants (Walker and Leegood, 1996; Walker et al., 1997; Walker et al., 2002; Bailey et al., 2007; Chao et al., 2014). This poses an issue with regards to engineering C_4 photosynthesis into target C_3 crops, and there is growing
advocacy for PEPCK subtypes, with both PEPCK and NAD-malic enzyme activity, to be used as an engineering blueprint rather than plants which predominantly use NADP-malic enzyme (Bräutigam et al., 2014; Wang et al., 2014). Furthermore, PEPCK may be subjected to different regulatory mechanisms than NADP-malic enzyme, simply based on its location in the cytosol, rather than in plastids.

4.1.3 Identifying species specific NADP-malic enzyme properties

It is important to ascertain if the catalytic properties of NADP-malic enzyme from closely related C₄ plants are similar or if distinct differences in the molecular design of NADP-malic enzyme from each of the C₄ grasses may be contributing to species-specific C₄-function and regulation. Putatively identified light-dependent phosphorylation sites presented in Chapter 3 are the first evidence of the post-translational regulation of NADP-malic enzyme in Setaria viridis, but whether these sites confer any biological significance, is unknown. Furthermore, phosphorylation at the identified residues may alter the physiochemical properties of NADP-malic enzyme, which could possibly alter sensitivity to pH and induce changes in the catalytic properties for its reaction with L-malate. This investigation was aimed to understand the significance of chloroplastic NADP-malic enzyme activity and its regulation in response to illumination and pH, and aid our attempts in engineering the C₄-specific NADP-malic enzyme into target C₃ crops. This Chapter will show that the activity and kinetics of NADP-malic enzyme from three NADP-malic enzyme type C₄ grasses varies in response to pH, illumination and reducing conditions with DTT (reducing agent). Data presented in this study will also demonstrate that the NADP-malic enzyme regulatory mechanism by light activation applies differently to NADP-malic enzyme subtypes.
4.2 Results

4.2.1 Kinetics of phosphoenolpyruvate carboxylase and NADP-malic enzyme

The dark and light regulation of PEPC and NADP-malic enzyme was determined by comparing their in vitro activities. The activity of these enzymes was monitored in continuous assays for approximately 20 min using a 96-well plate format, consisting of 150 µL reaction volumes and a small volume of crude enzyme extract normalised against the lowest concentration. Testing the design of this approach, the in vitro activities of PEPC showed a decrease in affinity for PEP at 7.5 h into dark (Figure 4.2) consistent with the activation mechanism by light-dependent phosphorylation demonstrated by Bailey et al. (2007). The $K_M$ for PEP decreased approximately 0.45-fold after illumination and remained relatively low until the end of the photoperiod. There was no significant difference in $K_M$ between 4 h and 15.5 h into light and 2 h into dark time-points.

**Fig. 4.2.** The activity of PEPC in *Setaria viridis* corresponds to shifts in $K_M$ during the photoperiod. The kinetic properties of PEPC from darkened (2 and 7.5 h into dark period) and illuminated (4 and 15.5 h into light) leaves were determined by enzyme assays. Immunoblot indicating PEPC content in 2 µL of leaf lysate is shown to the right. Data are based on six biological replicates and error bars show the standard error (SE).
This approach was also used to assay the activity of NADP-malic enzyme in *Setaria viridis*. The $K_M$ for L-malate was higher at 7.5 h into the dark and 4 h into the light, when compared to 2 h into the dark and 15.5 h into the light (Figure 4.3). However, despite that the $K_M$ for L-malate increased by 1.3-fold from 2 h into the dark and 4 h into the light, NADP-malic enzyme maintained high activity. There was no significant difference in $K_M$ for L-malate in 15.5 h illuminated and 2 h darkened leaves. Considering light availability at these two transition points, it is possible that the light-dependent mechanism for enzyme activation occurs or is in operation at least 4 h after the onset of light.

**Fig. 4.3.** NADP-malic enzyme $K_M$ for L-malate was highest during the dark to light transition in *Setaria viridis*. The kinetic properties of NADP-malic enzyme from darkened (2 and 7.5 h into dark period) and illuminated (4 and 15.5 h into light) leaves were determined by enzyme assays. Immunoblot indicating NADP-malic enzyme content in 2 µL leaf lysate shown to the right. Data are based on six biological replicates and error bars show the standard error (SE).

The initial results suggest that malate metabolism by NADP-malic enzyme occurred predominantly towards the end of the photoperiod in *Setaria viridis*. These observations may warrant further investigation into the how NADP-malic enzyme responds to during the photoperiod and dark period in other C₄ grasses.
4.2.2 NADP-malic enzyme is regulated by pH

In the initial test assays the optimal conditions for assaying malic enzyme were obtained from Ashton et al. (1990). To further investigate the light-dependent regulation of NADP-malic enzyme, the optimal assay conditions in *Setaria viridis*, *Sorghum bicolor* and *Zea mays* were determined. NADP-malic enzyme from *Sorghum bicolor* displayed two distinct pH optima. One at pH 7.4 in illuminated leaves and another at pH 8.2 in darkened leaves (Figure 4.4). At a low pH, the activity of NADP-malic enzyme was highest in illuminated leaves, but lowest in darkened leaves. Conversely, at a high pH, the activity of NADP-malic enzyme was highest in darkened leaves and lowest in illuminated leaves. The differences in NADP-malic enzyme activity between 2 and 7.5 h darkened or 4 and 15.5 h illuminated leaves was not significant.

![Graph showing pH optima of NADP-malic enzyme activity](image)

**Fig. 4.4. The pH optimum of NADP-malic enzyme activity in *Sorghum bicolor* shifted after illumination.** The optimum activity of NADP-malic enzyme, in terms of NADPH produced over time, was determined using a pH range from 6.8 to 9.2. NADP-malic enzyme from *Sorghum bicolor* displayed two distinct pH optima in light and dark (indicated by dashed lines). pH optima were determined at four time-points: 2 and 7.5 h into dark (blue lines, black markers) and 4 and 15.5 h into the light (orange lines, clear markers). Curve fitting was performed in GraphPad Prism 7, considering the standard deviation error bars.
Unlike *Sorghum bicolor*, NADP-malic enzyme in *Setaria viridis* from darkened and illuminated leaves remained active across a broader pH range (Figure 4.5). Furthermore, NADP-malic enzyme activity in darkened leaves decreased after pH 8.2, whereas enzyme activity in illuminated leaves decreased after pH 9.2. Also, unlike *Sorghum bicolor*, the pH response of NADP-malic enzyme from *Setaria viridis* illuminated leaves was broader and peaked at pH 8.6, while the activity in darkened leaves peaked around pH 8.2, which was similar to the activity observed in darkened leaves of *Sorghum bicolor* (Figure 4.4).

![Graph showing pH activity of NADP-malic enzyme](image)

**Fig. 4.5. NADP-malic enzyme from illuminated *Setaria viridis* leaves remained active across a broader pH range.** The optimum activity of NADP-malic enzyme was determined using a pH range from 6.8 to 9.2. NADP-malic enzyme from *Setaria viridis* displayed higher tolerance to increasing pH, peaking at pH 8.2 and pH 8.6 (indicated by dashed lines), in darkened and illuminated leaves, respectively. The differences in optimum activity between light and dark were not significant. pH optima were determined at four time-points: 2 and 7.5 h into dark (blue lines) and 4 and 15.5 h into the light (orange lines). Curve fitting was performed in GraphPad Prism 7, considering the standard deviation error bars.

In *Zea mays*, the NADP-malic enzyme pH curve (Figure 4.6) was similar to that of *Sorghum bicolor* and like *Sorghum bicolor*, the activity of NADP-malic enzyme in *Zea mays* displayed a distinct peak at pH 8.2, but there was no difference in the pH optimum between 7.5 h darkened and 15.5 h illuminated leaves.
There was no change in the pH optimum of NADP-malic enzyme activity in *Zea mays* leaves. The optimum activity of NADP-malic enzyme was determined using a pH range from 6.8 to 9.2. Activity at pH 8.2 (dashed line). pH optima were determined at two time-points: 7.5 h into dark (black squares) and 15.5 h into the light (clear squares). Curve fitting was performed in GraphPad Prism 7, considering the standard deviation error bars.

NADP-malic enzyme activity in illuminated *Zea mays* leaves exhibited low activity at a lower pH, which was similar to the activity of NADP-malic enzyme from illuminated *Setaria viridis* leaves at pH 7.0 (Figure 4.5). Additionally, the activity of NADP-malic enzyme from illuminated leaves of *Zea mays* dropped substantially at pH 6.8. This was not observed in *Sorghum bicolor* or *Setaria viridis*. The activity of NADP-malic enzyme from darkened leaves in response to pH was similar in *Sorghum bicolor* (Figure 4.4) and *Zea mays* (Figure 4.6). In *Setaria viridis*, NADP-malic enzyme displayed a flatter peak at the pH optimum (Figure 4.5) and the maximum activity was approximately three times lower compared to *Sorghum bicolor* and *Zea mays* (Figure 4.7). The differences of NADP-malic enzyme activity across all three C₄ grasses, in response to increasing pH, were mostly observed in illuminated leaves.

In *Zea mays*, there was no significant difference in the NADP-malic enzyme activity between illuminated and darkened leaves. Similarly, the dark and light activities of NADP-malic enzyme in *Setaria viridis* were not significant around the pH optimum, but were distinguishable below and above the pH optimum. In *Sorghum bicolor*, NADP-malic enzyme from illuminated leaves remained active at a lower pH (Figure 4.4), unlike NADP-malic enzyme activity from *Setaria viridis* or *Zea mays*, which was much lower.
at pH 6.8. Additionally, the activity of NADP-malic enzyme in *Sorghum bicolor* in illuminated leaves, unlike that of *Setaria viridis* (Figure 4.5) was not restored with increasing pH, and declined about 0.5-fold lower than the optimum activity at pH 7.4.

![NADP-malic enzyme activity](image)

**Fig. 4.7. Dark and light NADP-malic enzyme activity in leaves of *Sorghum bicolor*, *Setaria viridis* and *Zea mays* at the optimum pH.** Activity of NADP-malic enzyme in leaf lysates normalised against the lowest concentration. Data are based on six biological replicates and the error bars show the standard deviation.
4.2.3 Optimal NADP-malic enzyme activity at low a Mg\textsuperscript{2+} concentration

The optimum activity of NADP-malic enzyme was determined in the presence of Mg\textsuperscript{2+}, its preferred divalent metal co-factor as previously reported by Johnson and Hatch (1970). For both *Setaria viridis* and *Sorghum bicolor*, there was no NADP-malic enzyme activity in the absence of Mg\textsuperscript{2+}. NADP-malic enzyme from *Setaria viridis* (Figure 4.8), was active in the presence of 1–15 mM at the optimum pH. Activity declined in the presence of 20 and 30 mM. In addition, there was no difference in Mg\textsuperscript{2+} dependence between darkened and illuminated leaves.

![Setaria viridis, pH 8.4](image)

**Fig. 4.8.** No difference in NADP-malic enzyme activity from darkened and illuminated leaves of *Setaria viridis* in the presence of Mg\textsuperscript{2+}. NADP-malic enzyme was assayed in the decarboxylation direction, towards the formation of pyruvate, in the presence of MgCl\textsubscript{2}(0–30 mM), at pH 8.4. Data are based on six biological replicates and error bars show the standard deviation.

At the light pH optimum (pH 7.4), the activity of NADP-malic enzyme from *Sorghum bicolor* (Figure 4.9) did not significantly change in the presence of 1 and 30 mM MgCl\textsubscript{2}. NADP-malic enzyme activity was less stable at the dark pH optimum (pH 8.2). Unlike *Setaria viridis*, the activity of *Sorghum bicolor* NADP-malic enzyme was light-dependent. In addition, the activity of NADP-malic enzyme from darkened leaves was lower than illuminated leaves in both optimal and suboptimal pH.
Fig. 4.9. Activity of NADP-malic enzyme in response to Mg$^{2+}$ may be pH- and light-dependent in *Sorghum bicolor*. NADP-malic enzyme assayed in the decarboxylation direction, towards the formation of pyruvate, in the presence of MgCl$_2$ (0–30 mM), at pH 7.4 (light optimum) and pH 8.2 (dark optimum). Data are based on six biological replicates and error bars show the standard deviation.

The optimal Mg$^{2+}$ concentration for *Setaria viridis* and *Sorghum bicolor* NADP-malic enzyme was determined at 2 mM. Dependence on Mg$^{2+}$ may be pH-dependent in *Sorghum bicolor*. The optimal concentration of Mg$^{2+}$ was not determined for *Zea mays*. 
4.2.4 Michaelis-Menten kinetics of NADP-malic from three C₄ grasses

To further analyse the effects of illumination and pH, the dark and light Michaelis-Menten kinetics of NADP-malic enzyme, at the identified pH and Mg²⁺ optima, were investigated. In addition, to test the effect of reducing conditions, the activity of NADP-malic enzyme was measured by adding 10 mM DTT (reducing agent) into the reaction assay. In *Sorghum bicolor*, the previous results showing that NADP-malic enzyme had two distinct pH optima under illuminated or darkened conditions were further substantiated by shifts in $K_M$ for L-malate under optimal and suboptimal conditions (Figure 4.10). There was no change in the $K_M$ for L-malate under optimal and suboptimal pH in 15.5 h illuminated leaves. At pH 7.4 and pH 8.2, NADP-malic enzyme $K_M$ for L-malate in darkened *Sorghum bicolor* leaves decreased in the presence of DTT. In darkened leaves, the $K_M$ for L-malate was nearly 1.8-fold higher than the $K_M$ in illuminated leaves in optimal pH and about 0.66-fold higher in the suboptimal pH (Figure 4.10), suggesting that the affinity for L-malate increases after the dark to light transition, regardless of pH. This observation concurred with previous findings in *Setaria viridis* (Figure 4.3), in which the $K_M$ for L-malate was about 1.3-fold higher at 7.5 h into dark and remained high after 4 h of illumination, until lowering after 15.5 h of illumination. However, the $K_M$ for L-malate decreased in response to suboptimal pH, regardless the light condition.
Fig. 4.10.

<table>
<thead>
<tr>
<th>pH 7.4 (K_M ± SE µM)</th>
<th>pH 7.4 + DTT (K_M ± SE µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark</td>
<td>Light</td>
</tr>
<tr>
<td>236.6 ± 22.39</td>
<td>154.9 ± 18.98</td>
</tr>
<tr>
<td>Light</td>
<td>Dark</td>
</tr>
<tr>
<td>209.9 ± 12.68</td>
<td>125.0 ± 6.17</td>
</tr>
<tr>
<td>Light</td>
<td></td>
</tr>
<tr>
<td>236.6 ± 22.39</td>
<td>154.9 ± 18.98</td>
</tr>
</tbody>
</table>
The differences in light and dark activities of *Sorghum bicolor* NADP-malic enzyme in response to pH linked to shifts in $K_M$. Michaelis-Menten curves for the NADP-malic reaction in darkened (black dots) and illuminated (clear dots) *Sorghum bicolor* leaves in response to increasing L-malate concentrations. $K_M$ was determined at two pH optima (pH 7.4 in green and pH 8.2 in blue) and in response to DTT (+) and no added DTT (–). Data are based on six biological replicates and error bars show the standard deviation (SD). SE, standard error. Units for initial velocity are Abs/min.

<table>
<thead>
<tr>
<th>pH 8.2 ($K_M \pm \text{SE } \mu\text{M})$</th>
<th>pH 8.2 + DTT ($K_M \pm \text{SE } \mu\text{M})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark 443.7 ± 56.18</td>
<td>Light 142.3 ± 29.07</td>
</tr>
<tr>
<td>Light 318.8 ± 14.82</td>
<td>Light 200.8 ± 20.67</td>
</tr>
</tbody>
</table>
The relationship between optimal and suboptimal pH and NADP-malic enzyme activity was not conclusive in *Setaria viridis*. At the optimal pH (pH 8.4), NADP-malic enzyme did not display Michaelis-Menten kinetics and the $K_M$ for L-malate was ambiguous without the addition of DTT (Figure 4.11). However, Michaelis-Menten kinetics were restored at the suboptimal (pH 7.4) in both darkened and illuminated leaves. In addition, at a higher pH, the activity of NADP-malic enzyme in *Setaria viridis* was higher, but the affinity for L-malate was low. Conversely, enzyme activity was low at a lower pH, but the affinity for L-malate was higher.
Fig. 4.11.

<table>
<thead>
<tr>
<th>pH 7.4 (K_M ± SE µM)</th>
<th>pH 7.4 + DTT (K_M ± SE µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark</td>
<td>Light</td>
</tr>
<tr>
<td>20.9 ± 4.10</td>
<td>35.7 ± 6.92</td>
</tr>
</tbody>
</table>
Fig. 4.11. Link between pH and light/dark activity of NADP-malic enzyme in *Setaria viridis*. Michaelis-Menten curves for the NADP-malic reaction in darkened (black dots) and illuminated (clear dots) *Setaria viridis* leaves in response to increasing L-malate concentrations. $K_M$ was determined at the optimal pH 8.4 (blue) and suboptimal pH 7.4 (green), in response to DTT (+) and no added DTT (−). Data are based on six biological replicates and error bars show the standard deviation (SD). SE, standard error. Units for initial velocity are Abs/min.

<table>
<thead>
<tr>
<th>pH 8.4 ($K_M$ ± SE μM)</th>
<th>pH 8.4 + DTT ($K_M$ ± SE μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark</td>
<td>Light</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>Ambiguous</td>
</tr>
</tbody>
</table>
7.5h Dark, pH 7.4

15.5h Light, pH 7.4

<table>
<thead>
<tr>
<th>pH 7.4 (K_M ± SE μM)</th>
<th>pH 7.4 + DTT (K_M ± SE μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark</td>
<td>Light</td>
</tr>
<tr>
<td>10.8 ± 4.42</td>
<td>9.1 ± 4.21</td>
</tr>
</tbody>
</table>

Fig. 4.12.
Fig. 4.12. NADP-malic enzyme in *Zea mays* maintained high substrate affinity in response to pH. Michaelis-Menten curves for the NADP-malic reaction in darkened (black dots) and illuminated (clear dots) *Zea mays* leaves in response to increasing L-malate concentrations. $K_M$ was determined at a single pH optimum (pH 8.2 in blue) and a suboptimal pH (pH 7.4 in green), in response to DTT (+) and no added DTT (−). Affinity for L-malate remained consistently high (low $K_M$) in darkened and illuminated leaves in high and low pH. Presence of DTT had a marginal effect on NADP-malic enzyme affinity for L-malate. Data are based on six biological replicates and error bars show the standard deviation (SD). SE, standard error. Units for initial velocity are Abs/min.
In *Zea mays*, NADP-malic enzyme maintained a higher substrate affinity across all time-points (Figure 4.12), unlike *Sorghum bicolor* or *Setaria viridis*. In the presence of DTT, there was a significant change in $K_M$ in darkened and illuminated leaves in suboptimal pH (pH 7.4). These differences were greater than those observed in *Sorghum bicolor* and *Setaria viridis* leaves in response to DTT. However, at the pH optimum (pH 8.2), a marginal increase in $K_M$ was observed, but was not significant. In *Zea mays*, *Sorghum bicolor* and *Setaria viridis*, the effect of DTT on the activity and $K_M$ for L-malate was not discernible in these assays. Furthermore, there were no distinct patterns in the change of $K_M$ at the optimal and suboptimal pH in response to DTT across the three C$_4$ grasses, and therefore the effect of DTT on the activity of NADP-malic enzyme was inconclusive.

**Table 4.1. Changes to NADP-malic enzyme affinity for L-malate in response to illumination.** Summary of the change in NADP-malic enzyme affinities for L-malate in response to the light transition, under optimal and suboptimal pH. Up-arrows indicate increase in affinity, while down-arrows indicate decrease in affinity. *Fold change calculated in the presence of DTT.*

<table>
<thead>
<tr>
<th></th>
<th>Optimal pH</th>
<th>Suboptimal pH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affinity (fold change)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Sorghum bicolor</em></td>
<td>0.65 ↑</td>
<td>0.40 ↑</td>
</tr>
<tr>
<td><em>Setaria viridis</em></td>
<td>0.40* ↑</td>
<td>0.71 ↓</td>
</tr>
<tr>
<td><em>Zea mays</em></td>
<td>0.12 ↓</td>
<td>0.16 ↑</td>
</tr>
</tbody>
</table>

It was evident that under optimal conditions, the NADP-malic enzyme affinity range for L-malate was very high in *Zea mays*, moderate in *Sorghum bicolor*, but much lower in *Setaria viridis*. In illuminated leaves, NADP-malic enzyme affinity for L-malate increased by almost 0.65-fold in *Sorghum bicolor*, 0.40-fold in *Setaria viridis*, but decreased about 0.12-fold in *Zea mays* after the dark to light transition at the pH optimum (Table 4.1). Furthermore, under optimal and suboptimal conditions, NADP-malic enzyme affinity for L-malate in *Zea mays* leaves remained higher than in *Sorghum bicolor* and *Setaria viridis*. However, since NADP-malic enzyme in *Setaria viridis* did not display Michaelis-Menten kinetics under optimal conditions and $K_M$ was ambiguous, the true relationship between affinity for L-malate and pH cannot be determined.
4.3 Discussion

4.3.1 Low pH optimum ensures NADP-malic enzyme activity at the onset of light

The C₄-specific NADP-malic enzyme is generally assayed at pH 8.0 (Ashton et al., 1990; Edwards and Andreo, 1992), which correlates to the physiological pH in the chloroplast stroma during illumination, when NADP-malic enzyme is active in C₄ plants (Neumann and Jagendorf, 1964; Heldt et al., 1973; Kanai and Edwards, 1999). In this study, NADP-malic enzyme activity was determined in low and high pH and varying concentrations of Mg²⁺. These results show that the activity of NADP-malic enzyme from darkened and illuminated leaves of Setaria viridis (Figure 4.5) and Zea mays (Figure 4.6) was at its optimum at a high pH (pH 8.0), which is consistent with previous findings (Johnson and Hatch, 1970; Nishikido and Wada, 1974; Edwards and Andreo, 1992). However, the activity of the enzyme from illuminated leaves of Sorghum bicolor was at its optimum at a lower pH (Figure 4.4), and activity declined as the pH increased. In fact, NADP-malic enzyme from Sorghum bicolor exhibited two distinct pH optima (Figure 4.13). These results indicate that there are interspecies differences in the regulation of NADP-malic enzyme activity in closely related C₄ grasses. Also, while in vitro enzyme assays have been performed using 5 mM Mg²⁺ (Ashton et al., 1990), maximal activity was achieved using 2 mM Mg²⁺ (Figure 4.8, 4.9). This correlates to the physiological concentration of Mg²⁺ in the chloroplast stroma, which rises from about 0.5 mM to 2 mM upon illumination (Portis and Heldt, 1976; Heldt, 1979; Ishijima et al., 2003). Apart from serving as a co-factor in the reaction, Mg²⁺ protects NADP-malic enzyme from becoming oxidised (Drincovich and Andreo, 1994). Whether 2 mM Mg²⁺ can grant the same protection as 10 mM used in previous experiments, is unclear. It was clear however, that the activity of NADP-malic enzyme decreased when using Mg²⁺ concentrations over 10 mM. This suggests that a high concentration of Mg²⁺ can lead to the blockage of substrate binding sites in the C₄-specific NADP-malic enzyme, and should be taken into consideration for future in vitro assays. This has been previously speculated to occur in the NADP-malic enzyme from Escherichia coli, whereby concentrations of 4 mM Mg²⁺ and above inhibited enzyme activity (Bologna et al., 2007). Loss of NADP-malic enzyme activity could have been due to the oxidised form of the enzyme present in the reaction assay. However, these effects were kept to a minimum, because NADP-malic enzyme
was extracted in reducing conditions with 5 mM DTT to prevent the formation of disulphide bonds.

NADP-malic enzyme from illuminated *Sorghum bicolor* leaves displayed maximal activity at pH 7.4, but there was no change in affinity for L-malate between pH 7.4 and pH 8.2. This suggests that NADP-malic enzyme from *Sorghum bicolor* is maximally active at the onset of light, when the pH in the stroma is close to pH 7.0 (Werdan et al., 1975). Optimal NADP-malic enzyme activity from dark-grown *Sorghum bicolor* plants was achieved at pH 8.2, which is close to the pH optimum generally used to assay NADP-malic enzyme activity (Ashton et al., 1990).

![Fig. 4.13. Simple model showing two distinct activities of NADP-malic enzyme from darkened (blue line) and illuminated (orange line) leaves of *Sorghum bicolor*. Activity of NADP-malic enzyme is shown as high and low, which is modelled from the changes in activity shown in Figure 4.4. Dark and light periods are shown in light blue and light yellow shaded regions, respectively. A simple model shows the changes in pH in the stroma during light transitions (bottom).](image-url)
A simple model of NADP-malic enzyme activity in *Sorghum bicolor*, Figure 4.13, shows that when the activity of NADP-malic enzyme from illuminated leaves is high, the activity of the NADP-malic enzyme from darkened leaves is low. These inverse changes in dark and light NADP-malic enzyme activity align with the diurnal change of the pH in the stroma, which rises after illumination (Heldt et al., 1973; Werdan et al., 1975). The pH in the stroma declines at the beginning of the dark period and inhibits carbon fixation (Werdan et al., 1975), and it is unlikely that NADP-malic enzyme maintains high *in vivo* activity at a high pH after the light to dark transition. The difference in the pH optimum between light-grown and dark-grown *Sorghum bicolor* plants suggests that NADP-malic enzyme exists as two separate isoforms during light transitions. Previous studies have shown that the photosynthetic NADP-malic enzyme in *Zea mays* and *Flaveria floridana* has a high pH optimum, while the non-photosynthetic form, which is present in both C₃ and C₄ species, has a low pH optimum (Drincovich et al., 1991; Marshall et al., 1996; Maurino et al., 1996; Casati et al., 1997; Casati et al., 1999). This suggests that the distinct differences in the pH optimum observed in these assays may have been due to non-photosynthetic NADP-malic enzyme activity in crude leaf extracts.

However, it is unclear whether the formation of NADPH at a low pH was due to activity from the non-photosynthetic form of NADP-malic enzyme in the crude leaf lysates. While this may be the case, it is uncertain why NADP-malic enzyme activity was relatively lower at the pH optimum of the C₄-form (Edwards and Andreo, 1992). Also, if NADP-dependent activity came from the non-photosynthetic isoform, then the two distinct pH optima would have been evident in assays conducted in *Setaria viridis* and *Zea mays*, which may contain appreciable amounts of non-photosynthetic NADP-malic enzyme (Drincovich et al., 2001), and thus could have contributed to the formation of NADPH. A model of NADP-malic enzyme activity from illuminated and darkened leaves of *Setaria viridis* (Figure 4.14) shows that NADP-malic enzyme from illuminated leaves had a broader pH dependency and it did not exhibit two distinct pH optima, like *Sorghum bicolor*. Instead, the activity of NADP-malic enzyme rose as pH increased after illumination, and the pH optimum was determined at pH 8.4, irrespective of the light condition, aligning to the pH optimum of NADP-malic enzyme in C₄ plants (Ashton et al., 1990) and light-dependent activity of NADP-malic enzyme *in vivo* (Asami et al., 1979; Edwards and Andreo, 1992).
Fig. 4.14. Simple model showing NADP-malic enzyme activity from darkened (blue line) and illuminated (orange line) leaves of *Setaria viridis*. Activity of NADP-malic enzyme is shown as high and low, which is modelled from the changes in activity shown in Figure 4.5. Dark and light periods are shown in light blue and light yellow shaded regions, respectively. A simple model shows the changes in pH in the stroma during light transitions (bottom).

When compared to the modelled activity of NADP-malic enzyme from darkened and illuminated leaves of *Zea mays* (Figure 4.15) it is unlikely that activity from the non-photosynthetic NADP-malic enzyme was observed. NADP-malic enzyme activity was lower at pH 7.4 in *Zea mays* and *Setaria viridis*, while high activity was observed in high pH, suggesting that NADP-dependent activity came from the C₄-specific malic enzyme, which has a high pH optimum. Had there been non-photosynthetic NADP-malic enzyme activity, such activity would have been distinctly higher at a lower pH, as observed in *Sorghum bicolor*. Another possibility is that *Zea mays* and *Setaria viridis* may have a relatively lower amount of the non-photosynthetic NADP-malic enzyme. However, this is unlikely since the non-photosynthetic form is the second major isoform of NADP-

**Fig. 4.15.** Simple model showing NADP-malic enzyme activity from darkened (blue line) and illuminated (orange line) leaves of *Zea mays*. Activity of NADP-malic enzyme is shown as high and low, which is modelled from the changes in activity shown in Figure 4.6. Dark and light periods are shown in light blue and light yellow shaded regions, respectively. A simple model shows the changes in pH in the stroma during light transitions (bottom).

Since the activity of NADP-malic enzyme was not as high at a low pH in neither *Setaria viridis* or *Zea mays*, there may be other factors regulating activity of NADP-malic enzyme in *Sorghum bicolor*. Light-dependent differences at the pH optimum may be due to post-translational regulation mechanisms occurring at least 30 min before the onset of light and 30 min before the dark period. The latter would correlate with the previous finding in *Setaria viridis*, in which NADP-malic enzyme displayed a higher affinity for L-malate 30 min before dark period, while the affinity for L-malate was about 1.3-fold lower 4 h after the onset of light (Figure 4.3, 4.14). Furthermore, in *Sorghum bicolor*, the affinity for L-malate was about 0.46-fold higher at pH 7.4 than at pH 8.2, 30 min before
the onset of light, which does not concur with previous findings that showed that the C₄-form has a lower affinity for L-malate at pH 7.0 when the tetramer dissociates (Asami et al., 1979; Detarsio et al., 2007). However, before the onset of light, a high affinity for L-malate might be required to achieve and maintain a high reaction rate since malate has not accumulated in the chloroplast. Therefore, the overlap in activity of NADP-malic enzyme between darkened and illuminated leaves may contribute to achieving a seamless transition in activity between light conditions, and may also correlate to the time it takes for metabolites to accumulate in C₄ plants. In this case, it is unlikely that NADP-malic enzyme in *Sorghum bicolor* is regulated by diurnal changes in the pH of the stroma, but pH serves more as fine-control of enzyme isoforms. An additional time-point, around 4.5 h into the dark period would ascertain whether this mechanism occurs solely at dark to light transitions.

NADP-malic enzyme may alternatively be regulated by redox mechanisms, which can exert more control over the activation and inactivation of chloroplastic enzymes, than through diurnal changes in the pH of the stroma (Jacquot et al., 1981; Leegood and Walker, 1983; Drincovich and Andreo, 1994; Schürmann and Buchanan, 2008). Addition of 10 mM DTT to lysates from illuminated leaves containing NADP-malic enzyme from *Sorghum bicolor* increased the affinity for L-malate when assayed at pH 7.4 (Figure 4.10), but had no effect on the activity of NADP-malic enzyme. Conversely, addition of DTT to illuminated leaf lysates increased the maximal activity of NADP-malic enzyme at pH 8.2 and decreased affinity for L-malate (Figure 4.10). This might indicate that at pH 8.2 the oxidised-form of the enzyme, from *Sorghum bicolor*, might have a higher sensitivity to allosteric inhibition (non-competitive), possibly by malate, than the reduced-form, which displayed a higher maximal activity. However, this does not align with the catalytic properties of NADP-malic enzyme, whereby sensitivity to allosteric inhibition by malate increases at a lower pH (Asami et al., 1979). A previous study has shown that NADP-malic enzyme from the intermediate C₃–C₄ species *Flaveria floridana* is not inhibited by malate at pH 7.0 (Casati et al., 1999), suggesting that the enzyme assayed in illuminated leaves of *Sorghum bicolor* may be distinctly different than the C₄-form found in leaves of *Zea mays*. The calculated pI for NADP-malic enzyme from these C₄ grasses ranges from 5.8 to 6.7 (isoform K3XG11 has pI 7.15) and under the range of pH used in these assays (pH 6.8 to 9.2), the enzymes should exist as net-negatively charged species. Any differences to their activities or affinities for L-malate
might be due to the action of the reducing agent (DTT) on cysteine-cysteine disulphide bonds, which confers stability to the protein. There are ten cysteine residues in *Sorghum bicolor*, nine residues in *Zea mays* and seven residues in *Setaria viridis*, which could theoretically be affected. However, since these assays were performed using enzyme in crude leaf extracts, kinetic constants like $V_{\text{max}}$, $K_{\text{cat}}$ or $K_i$ cannot be calculated, and the exact effects of DTT on the activity of NADP-malic enzyme cannot be determined.

The NADP-malic enzyme *in vivo* activity in illuminated leaves of *Sorghum bicolor* may increase because malate is at high concentrations (Leegood and von Caemmerer, 1989) and therefore, a high affinity for malate might not be required. Also, despite that the lower affinity for L-malate at pH 8.2 in illuminated leaves of *Sorghum bicolor*, the maximal activity of NADP-malic enzyme increased (Figure 4.10), which aligns with the *in vivo* NADP-malic enzyme activity, which is activated by a light-induced thiol-disulphide interchange (Drincovich and Andreo, 1994). The effect of DTT on the activity of NADP-malic enzyme is also consistent with that of other chloroplastic enzymes, which are regulated by light-dependent redox mechanisms (Jacquot et al., 1981; Ashton and Hatch, 1983; Leegood and Walker, 1983; Drincovich and Andreo, 1994; Raines et al., 2000; Schürmann and Buchanan, 2008; Gütle et al., 2016).

Addition of 10 mM DTT to darkened and illuminated leaf lysates from *Zea mays* lowered the affinity for L-malate by fourfold at pH 7.4. A lower affinity at pH 7.4 may affect the rate of malate decarboxylation at the onset of light, because malate has not accumulated in the stroma. Conversely, the changes in affinity for L-malate were negligible between the oxidised- and reduced-form of the enzyme at pH 8.2. This is significant as it ensures that NADP-malic enzyme reaches maximal activity, with no major changes to its affinity for L-malate at a high pH in the light, regardless whether the enzyme is reduced or not. This mechanism may grant some protection since the redox state of disulphide bonds of NADP-malic enzyme may depend greatly on cellular stimuli or can be blocked under certain conditions such as immune responses (Drincovich and Andreo, 1994; Michelet et al., 2013; Bhattacharjee et al., 2015; Friso and van Wijk, 2015).
4.3.2 PEPCK dependence may influence NADP-malic enzyme activity

NADP-malic enzyme from illuminated leaves of Setaria viridis had a broader pH-dependence and exhibited a higher affinity for L-malate at light transitions (Figure 4.14). This suggests that the *in vivo* NADP-malic enzyme activity might be required at the onset of light, in a mechanism similar to that of NADP-malic enzyme from Sorghum bicolor. Transcriptomic analyses performed in leaves of Sorghum bicolor, Setaria italica, which is genetically similar to Setaria viridis (Bennetzen et al., 2012) and Zea mays (Figure 4.16B and 4.16C) show the relative expression of NADP-malic enzyme and PEPCK in these C₄ grasses. Therefore, it is conceivable that the activity of NADP-malic enzyme might be needed at the onset of light in Setaria viridis and Sorghum bicolor, because these two grasses have lower expression of PEPCK, whereas Zea mays has an appreciable amount of PEPCK (Figure 4.16). In addition to this, since NADP-malic enzyme is localised in the chloroplast, the activity of NADP-malic enzyme would depend more on fine-control by pH than PEPCK (Edwards and Andreo, 1992). Despite being closely related, the pH-dependency of NADP-malic enzyme from Sorghum bicolor was drastically different than that of Zea mays. Unlike NADP-malic enzyme from illuminated leaves of Sorghum bicolor, the activity of NADP-malic enzyme from Zea mays was relatively lower at pH 7.4 (Figure 4.6, 4.15). This suggests that decarboxylation by NADP-malic enzyme is not required in the same degree in Zea mays at the onset of light, as it seems for NADP-malic enzyme from Sorghum bicolor and Setaria viridis. One possibility might be that at the onset of light, the C₄-transfer acid in Zea mays is preferentially decarboxylated by PEPCK, instead of NADP-malic enzyme. This might also indicate that NADP-malic enzyme from Zea mays is differently regulated during light transitions, than the enzyme from Sorghum bicolor and Setaria viridis. Furthermore, the flux through aspartate and decarboxylation by PEPCK might be more energetically favoured than the flux through malate after the onset of light, since PEP does not need to be regenerated via PPDK using an extra ATP (Wang *et al.*, 2014).

In addition, the conversion of oxaloacetate to malate in the chloroplast of mesophyll cells depends on the MDH, which must also undergo activation by light (Johnson and Hatch, 1970; Jacquot *et al.*, 1981; Ashton and Hatch, 1983; Leegood and Walker, 1983). It is unlikely that the supply of NADPH, which is needed for the conversion of oxaloacetate to malate by reduction, is limiting the rate of malate formation at the onset of light, since
large amounts of NADPH are produced within milliseconds of illumination or supplied from metabolite exchanges (Huber and Edwards, 1975; Furbank and Leegood, 1984; Stitt and Zhu, 2014). Instead, it is conceivable that decarboxylation by PEPCK might not be equally as accessible in *Sorghum bicolor* and *Setaria viridis*, simply because expression of PEPCK is lower in these two grasses (Figure 4.16C). This is also supported by the inability to detect PEPCK in leaves of *Sorghum bicolor* or *Setaria viridis* by mass spectrometry and suggests that the abundance of PEPCK is lower than that of NADP-malic enzyme, which was readily detected in leaves of *Sorghum bicolor* and *Setaria viridis*. Although there may be supplementary PEPCK activity, this shows how much lower the amount of PEPCK is relative to the primary decarboxylase. This is further supported by the inability to detect NADP-malic enzyme in *Megathyrsus maximus* using a similar approach, while PEPCK, its primary decarboxylase (Bräutigam et al., 2014), was readily detected. Comparatively, the expression of PEPCK in *Megathyrsus maximus* is about 88-fold higher than in *Setaria italica* and *Sorghum bicolor* and about 10-fold higher than in *Zea mays.*
Fig. 4.16. Relative expression of NADP-malic enzyme (B) and PEPCK (C) in C₄ grasses. Activity of NADP-malic enzyme also shown (A). Expression in reads per kilobase million (rpkm). Expression data obtained from PA Christin (unpublished).

In addition, compared to the maximal activity of Sorghum bicolor and Zea mays, the activity of Setaria viridis was substantially lower (Figure 4.16A), while transcriptomic evidence shows that NADP-malic enzyme is highly expressed in Setaria italica (Figure 4.16B). $K_M$ plots showed that NADP-malic enzyme at that pH optimum (pH 8.4) did not saturate with 1 mM L-malate (Figure 4.11). This suggests that the activity of NADP-malic enzyme from Setaria viridis might have been lower because the concentration of L-malate was limiting the rate of the in vitro reaction, rather than containing less NADP-malic enzyme than Sorghum bicolor or Zea mays. Alternatively, there could have been active inhibitors in the Setaria viridis leaf lysates that affected the activity of NADP-malic enzyme. A comparative assessment between the maximal activities of NADP-malic enzyme can be further ascertained by mixing Zea mays and Sorghum bicolor leaf
lysates with *Setaria viridis* leaf lysates and determine whether activity in the mixed assays is inhibited.

### 4.3.3 New insights on NADP-malic enzyme regulation in C$_4$ plants

These results indicate that the pH optimum of NADP-malic enzyme activity varies in C$_4$ grasses. However, the regulatory mechanisms that promote high activity of NADP-malic enzyme at a lower pH in *Sorghum bicolor*, but not *Setaria viridis* and *Zea mays* are unknown. One major limitation of this investigation is that NADP-malic enzyme was assayed in crude leaf extracts. Firstly, the exact concentration of NADP-malic enzyme cannot be determined because several isoforms of the enzyme are present in the leaf lysate and can, if not equally, contribute to NADP-dependent activity. In addition, kinetic coefficients like $V_{\text{max}}$ and $K_{\text{cat}}$ cannot be calculated and precise changes in NADP-malic enzyme activity in response to illuminate cannot be ascertained. Secondly, the formation of NADPH may have come from mitochondrial NAD-malic enzyme activity, because it can utilise NADP$^+$ as a co-factor and contribute to NADP-dependent activity (Edwards and Andreo, 1992).

The $K_M$ determined for L-malate in illuminated leaves of *Sorghum bicolor* was close to the $K_M$ previously reported in crude leaf extracts (Table 4.2). The results in this Chapter show that affinity for L-malate decreases during the dark period in *Sorghum bicolor*, which is consistent with the light-dependent activation mechanism. More importantly, these results indicate that the affinity for L-malate does not change in *Zea mays*, and is significantly lower than that found in purified NADP-malic enzyme (Drincovich et al., 2001). This further suggests that there were different isoforms of NADP-malic enzyme in the crude leaf lysate that may have contributed to the reaction. In this case, $K_M$ values reported may be the cumulative kinetic properties of all NADP-malic enzyme isoforms present in the leaf lysate. In addition to the two major isoforms of NADP-malic enzyme in C$_4$ grasses, evidence in the *Flaveria* species suggests that there are three or four NADP-malic genes in total, one of which encodes a cytosolic enzyme (Marshall et al., 1996; Lai et al., 2002). If the cytosolic enzyme were present in crude leaf lysates, then it is conceivable that its pH optimum is different than that of the chloroplastic form.
Table 4.2. Comparing the $K_M$ for $l$-malate from previous studies.

<table>
<thead>
<tr>
<th>Plant species</th>
<th>$K_M$ for $l$-malate (mM)</th>
<th>$K_M$ for $l$-malate (mM) (this study)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Zea mays</em></td>
<td>0.19</td>
<td>0.006 0.007</td>
<td>(Drincovich <em>et al</em>., 2001)</td>
</tr>
<tr>
<td><em>Sorghum bicolor</em></td>
<td>0.10</td>
<td>0.44 0.15</td>
<td>(Nishikido and Wada, 1974)</td>
</tr>
<tr>
<td><em>Setaria viridis</em></td>
<td>0.25</td>
<td>– –</td>
<td>(Nishikido and Wada, 1974)</td>
</tr>
</tbody>
</table>

The reason for using crude leaf extracts in this study was to maximise the recovery of NADP-malic enzyme in its phosphorylated state. It is likely that during extraction some enzymes may have been dephosphorylated, thereby losing the chemical modifications that may confer some biochemical change. In addition to this, it is not known to what degree NADP-malic enzyme remained phosphorylated during the reaction assays. Unless the phosphorylated form of NADP-malic enzyme is purified, it is not possible to ascertain whether the phosphorylation state is affecting enzymatic activity. The kinetic properties of NADP-malic enzyme determined in this study may represent both the phosphorylated and dephosphorylated forms of NADP-malic enzyme.

4.3.4 Conclusion

It was shown that the NADP-malic enzyme from *Sorghum bicolor* leaves was distinctly different from the enzyme in *Setaria viridis* and *Zea mays*, supporting the hypothesis that interspecies variation may account for different NADP-malic enzyme requirements. With regards to regulation, pH may modulate the activities of dark and light forms, thus ensuring a seamless transition of chloroplastic activity between light conditions. This aligns with the initial hypothesis and previous speculation (Asami *et al*., 1979; Edwards and Andreo, 1992). In addition to this, the affinity for $l$-malate decreased in the light in *Sorghum bicolor*, whereas there was no change in the affinity for $l$-malate in *Zea mays*. These responses might be particularly important in *Sorghum bicolor*, because of higher requirements for malate decarboxylation than *Zea mays*, which may operate a dual-decarboxylation system, whereby decarboxylation is partitioned between PEPCK and NADP-malic enzyme (Gutierrez *et al*., 1974; Wang *et al*., 2014).
5.1 Introduction

5.1.1 The dual-decarboxylation system of C₄ photosynthesis

C₄ photosynthesis is a highly complex, but efficient process, which involves the movement of carbon between specialised mesophyll and bundle sheath cells, which partition carboxylation and decarboxylation pathways and compartmentalises cell-specific enzymes such as PEPC and Rubisco (Leegood, 2002; von Caemmerer and Furbank, 2003). To achieve the efficient movement of organic carbon, mesophyll cells surround bundle sheath cells and vascular tissue (Figure 5.1), creating a densely packed, wreath-like arrangement of cells that establishes an intimate relationship between the two cell types.

**Fig. 5.1. Kranz anatomy in C₄ plants.** C₄ leave cross-section (top) shows structure of Kranz anatomy in Zea mays made up of concentric tubes of vascular tissue bundle sheath cells and mesophyll cells. Biochemical specialisation of these tissues indicated (bottom).
This arrangement also promotes the rapid diffusion of organic carbon along concentration gradients via plasmodesmata, ensuring the movement of intermediate metabolites that drive the C₄ cycle and contribute to the carbon pool (Osmond, 1971; Hatch and Osmond, 1976; Leegood, 1985; Stitt and Heldt, 1985; Stitt and Zhu, 2014). In their elucidation of C₄ photosynthesis, Hatch and Slack (1966) showed that radioactively labelled CO₂ in sugarcane was rapidly incorporated into oxaloacetate, malate and aspartate, but not initially into metabolites produced in the PCR cycle as previously shown (Benson, 1954; Bassham and Kirk, 1960; Bassham and Calvin, 1962). Further dissection of the C₄ trait revealed that there were distinct biochemical variations in the mechanism used for photosynthetic carbon fixation, thus creating NADP-malic enzyme, PEPCK and NAD-malic enzyme subtypes of the C₄ pathway (Hatch et al., 1975; Furbank, 2011; Furbank, 2016). In NADP-malic enzyme subtype grasses, malate is the predominant C₄-acid utilised during C₄ photosynthesis (Ku et al., 1996), though aspartate also contributes, if not equally, to the CO₂ pool in bundle sheath cells (Andrews et al., 1971; Hatch, 1971; Meister et al., 1996; Bräutigam et al., 2014; Arrivault et al., 2016; Furbank, 2016; Ludwig, 2016). Recent radioactive CO₂ labelling experiments conducted in Zea mays, a NADP-malic enzyme subtype, demonstrated that gradients of aspartate, as well as alanine and phosphoenolpyruvate (PEP), contribute to the transfer of carbon via the PEPCK decarboxylation pathway (Arrivault et al., 2016).

It is apparent that NADP-malic enzyme subtype C₄ plants operate a dual-decarboxylation system (Figure 5.2), whereby C₄-acid decarboxylation, through the conversion of oxaloacetate to aspartate by transamination or to malate by reduction, is divided between PEPCK and NADP-malic enzyme, respectively (Wingler et al., 1999; Furbank, 2011; Pick et al., 2011; Bellasio and Griffiths, 2014; Bräutigam et al., 2014; Furbank, 2016). Also, PEPCK activity in bundle sheath cells may be linked to NAD-malic enzyme decarboxylation in the mitochondria in NAD-malic enzyme type C₄ plants (Bräutigam et al., 2014). One proposed advantage of utilising multiple decarboxylase pathways is that they increase the flexibility of carbon trafficking mechanisms, thus creating pools of intermediate metabolites that are used as a source of carbon in the absence of CO₂ fixation or during intervals of low-light (Leegood and von Caemmerer, 1989; Stitt and Zhu, 2014).
C₄ pathway showing a theoretical dual-decarboxylation system using NADP-malic enzyme (6), PEPCK (7) and NAD-malic enzyme (9). Metabolites (purple/grey): PEP, phosphoenolpyruvate; OAA, oxaloacetate; Asp, aspartate; Glu, glutamate; 2-OG, 2-oxoglutarate; Mal, malate; Pyr, pyruvate; Ala, alanine. Enzymes: 1, carbonic anhydrase; 2, pyruvate, phosphate dikinase; 3, PEP carboxylase; 4, aspartate aminotransferase; 5, NADPH-malate dehydrogenase; 6, NADP-malic enzyme; 7, PEP carboxykinase; 8, NAD-malate dehydrogenase; 9, NAD-malic enzyme; 10, alanine aminotransferase. PCR, photosynthetic carbon reduction cycle. Blue arrows indicate movement of assimilated CO₂; red arrows show decarboxylation reactions. Adapted from Wang et al. (2014).

It is also plausible that the use of multiple decarboxylases was favoured during the evolution of C₄ photosynthesis, increasing the options for carbon flux, decreasing the concentration of malate required to operate the C₄ cycle, and alleviating the burden on a single transport mechanism (Furbank and Leegood, 1984; Bräutigam et al., 2014; Wang et al., 2014). In fact, several shuttle mechanisms are present in leaves of C₄ plants that aid in the efficiency of net carbon assimilation. In all C₄ plants, a large proportion of 3-phosphoglycerate (3-PGA) produced in bundle sheath cell chloroplasts is transported to mesophyll chloroplasts via the 3-PGA/triose-phosphate shuttle, where it is reduced to triose-phosphates, which are transported back to the bundle sheath, and subsequently used for the regeneration of ribulose-1,5-bisphosphate (RuBP) (Hatch, 1987). Similarly, the interconversion of 3-PGA and PEP is proposed to serve in the flux of carbon from bundle sheath to mesophyll cells and increase the amounts of C₄ metabolites, though its
regulation is not known (Furbank and Leegood, 1984; Leegood and von Caemmerer, 1989; Bräutigam et al., 2014; Arrivault et al., 2016). The regulation of carbon movement through the relative flux of aspartate and malate in plants operating a dual-decarboxylation system is not entirely understood. Furthermore, there has been some debate as to how much the secondary decarboxylation by PEPCK contributes to photosynthetic efficiency of C₄ plants and whether flux through aspartate is necessary for maintaining biochemical stability in plants operating a NADP-malic enzyme system (Bräutigam et al., 2014; Wang et al., 2014). Additionally, there is no evidence in C₄ grasses that the interconversion between aspartate and oxaloacetate via aspartate aminotransferase or pyruvate and alanine via alanine aminotransferase is regulated at the protein level. The formation of aspartate and alanine is linked to PEPCK and NAD-malic enzyme decarboxylation (Wang et al., 2014) (Figure 5.2, reactions 4 and 10) and alanine might control the flux of carbon through this decarboxylation pathway by regulating the activity aspartate aminotransferase. To address whether alanine can stimulate the activity of aspartate aminotransferase and understand the regulation of carbon movement, through shuttle mechanisms associated with decarboxylation by PEPCK and NADP-malic enzyme, the kinetics of aspartate and alanine aminotransferase were studied in three NADP-malic enzyme C₄ grasses: Setaria viridis, Sorghum bicolor and Zea mays.

5.1.2 Regulation of aspartate and alanine aminotransferase

Aspartate aminotransferase (EC 2.6.1.1) catalyses the interconversion of L-aspartate (Asp) and 2-oxoglutarate (2-OG) to L-glutamate (Glu) and oxaloacetate (OAA) (Figure 5.3), by coupling to pyridoxal 5-phosphate (PLP), via a ping-pong kinetic mechanism (Toney, 2014) in mesophyll and bundle sheath cells of C₄ plants typically using the NAD-malic enzyme or PEPCK decarboxylation pathway (Hatch, 1987). In the first reaction, aspartate aminotransferase coupled with PLP reacts with L-aspartate forming pyridoxamine 5-phosphate (PMP) enzyme and oxaloacetate, then aspartate
aminotransferase coupled with PMP reacts with 2-oxoglutarate forming L-glutamate and regenerates PLP (Toney, 2014). In NAD-malic enzyme subtypes, there are three aspartate aminotransferase isoforms, two of which are specific to C₄ photosynthesis and present either in the mesophyll cytosol or in mitochondria of bundle sheath cells (Kanai and Edwards, 1999). The third isoform does not contribute to C₄-specific activity and resembles the C₃-form localised in chloroplasts (Taniguchi and Sugiyama, 1990). In Panicum miliaceum (NAD-malic enzyme subtype) there are three major alanine aminotransferase isoforms (Kanai and Edwards, 1999). In plants alanine aminotransferase (EC 2.6.1.2) catalyses the interconversion between L-alanine and 2-oxoglutarate to L-glutamate and pyruvate (Figure 5.4) and besides its essential role in C₄ photosynthesis, it is involved in gluconeogenesis, glycolysis and amino acid metabolism (McAllister and Good, 2015).

\[
\text{Ala} + \text{2-OG} \rightleftharpoons \text{pyruvate} + \text{Glu}
\]

**Fig. 5.4. Interconversion of aspartate and oxaloacetate by aspartate aminotransferase.**

Aspartate and alanine aminotransferase serve a central role in C₄ photosynthesis (Andrews *et al.*, 1971; Hatch and Mau, 1973; Pick *et al.*, 2011; Toney, 2014; Wang *et al.*, 2014), however their underlying regulatory mechanisms and catalytic properties for the transamination of oxaloacetate and L-aspartate by aspartate aminotransferase or pyruvate and L-alanine by alanine aminotransferase, along with their involvement in the flux of carbon with regards to the partitioning control of PEPCK and NADP-malic enzyme decarboxylation, are yet to be elucidated in C₄ grasses. Large pools of metabolites are required to drive the C₄ cycle (Leegood and Furbank, 1984; Leegood and von Caemmerer, 1989; Stitt and Zhu, 2014), however the effect of feedback inhibition or activation by transfer acids such as malate and alanine on the activity of aspartate aminotransferase has not been explored in NADP-malic enzyme subtypes. Since aspartate aminotransferase acts as a bridge for the PEPCK pathway, activation by alanine or malate might be a requirement in C₄ grasses that can switch between NADP-malic enzyme and PEPCK decarboxylation. It is therefore hypothesised that C₄ grasses with the ability to use both decarboxylation pathways, such as Zea mays, might be conditioned to maintain high aspartate aminotransferase activity in the presence of malate and alanine. The results in this Chapter will show that the kinetic properties of aspartate
aminotransferase are sensitive to dark and light changes and that the response to effector metabolites may contribute to a much greater role in regulating the movement of carbon through the PEPCK and NADP-malic enzyme decarboxylation pathways. These novel findings indicate that the efficiency and flexibility of C₄ photosynthesis depends on the proper regulation of aspartate and alanine aminotransferase in NADP-malic enzyme subtypes.

5.2 Results

5.2.1 The role and regulation of aspartate aminotransferase in C₄ photosynthesis

The kinetic properties of aspartate aminotransferase were determined in the forward direction, following the conversion of L-aspartate to oxaloacetate and glutamate utilising 2-oxoglutarate (Wilkie and Warren, 1998; Toney, 2014). Novel phosphoproteomic data presented in Chapter 3 showed that aspartate aminotransferase isoforms undergo phosphorylation predominantly in darkened conditions. To further investigate the regulation of aspartate aminotransferase in response to illumination, changes in Michaelis-Menten kinetics were determined for its primary substrate, L-aspartate, and secondary substrate, 2-oxoglutarate, in Setaria viridis, Sorghum bicolor and Zea mays.

![Fig. 5.5. Activity of aspartate aminotransferase in Sorghum bicolor, Setaria viridis and Zea mays.](image)

The activity of aspartate aminotransferase was determined towards the formation of oxaloacetate using 2.5 mM L-aspartate and 2-oxoglutarate. Data are based on six biological replicates and the error bars show the standard deviation.
There was no difference in the activity of aspartate aminotransferase in the presence of 2.5 mM L-aspartate, in leaf lysates of *Sorghum bicolor*, *Setaria viridis* and *Zea mays* (Figure 5.5). In addition, aspartate aminotransferase was not active in the presence of L-alanine. To determine the concentration of L-aspartate and 2-oxoglutarate required to reach saturation, the activity of aspartate aminotransferase was assayed in response to increasing substrate concentrations. The initial activity of aspartate aminotransferase was high in *Setaria viridis* and *Sorghum bicolor* darkened and illuminated leaves, but low in *Zea mays*. In *Setaria viridis* (Figure 5.6), the aspartate aminotransferase $K_M$ for L-aspartate increased from $1624 \pm 233.8$ µM to $2334 \pm 392.8$ µM after the dark to light transition, equivalent to approximately a 0.44-fold decrease in affinity for L-aspartate. Similarly, the $K_M$ for 2-oxoglutarate increased 0.20-fold in response to illumination.
Fig. 5.6. Aspartate aminotransferase affinity for L-aspartate and 2-oxoglutarate declines in response to light in *Setaria viridis*. Michaelis-Menten curves for the aspartate aminotransferase reaction in darkened (black lines, black dots) and illuminated (yellow lines, clear dots) *Setaria viridis* leaves in response to increasing L-aspartate (Asp) or 2-oxoglutarate (2-OG) concentrations. $K_M$ was determined at pH 8.0. Data are based on six biological replicates and error bars show the standard deviation (SD). SE, standard error. Units for initial velocity are Abs/min.
The Michaelis-Menten curve of *Sorghum bicolor* (Figure 5.7) was similar to that of *Setaria viridis* (5.6). However, the change in the aspartate aminotransferase $K_M$ for L-aspartate was more substantial in response to the dark to light transition in *Sorghum bicolor* (Figure 5.7). In *Sorghum bicolor*, the $K_M$ for L-aspartate increased from $1782 \pm 256.2 \mu M$ to $5185 \pm 1378.0 \mu M$ after illumination, which was about twice as high as the $K_M$ for L-aspartate in *Setaria viridis*. Conversely, the $K_M$ for 2-oxoglutarate declined in response to light in *Sorghum bicolor* leaves. In *Setaria viridis* and *Zea mays*, the $K_M$ for 2-oxoglutarate increased after illumination.
Fig. 5.7. Light-induced changes to aspartate aminotransferase affinity for L-Asp and 2-oxoglutarate in *Sorghum bicolor*. Michaelis-Menten curves for the aspartate aminotransferase reaction in darkened (black lines, black dots) and illuminated (yellow lines, clear dots) *Sorghum bicolor* leaves in response to increasing L-aspartate or 2-oxoglutarate (2-OG) concentrations. $K_M$ was determined at pH 8.0. Data are based on six biological replicates and error bars show the standard deviation (SD). SE, standard error. Units for initial velocity are Abs/min.
Zea mays $K_M$(Asp)

$K_M$

- 154.3 ± 36.5 µM (Dark)
- 2473 ± 402.1 µM (Light)

Zea mays $K_M$(2-OG)

$K_M$

- 110.1 ± 24.9 µM (Dark)
- 142.8 ± 7.4 µM (Light)

<table>
<thead>
<tr>
<th>L-Aspartate ($K_M$ ± SE µM)</th>
<th>2-oxoglutarate ($K_M$ ± SE µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark 154.3 ± 36.5</td>
<td>2473 ± 402.1</td>
</tr>
<tr>
<td>Light 110.1 ± 24.9</td>
<td>142.8 ± 7.4</td>
</tr>
</tbody>
</table>

**Fig. 5.8.** Aspartate aminotransferase affinity for L-aspartate declines after the onset of light in Zea mays. Michaelis-Menten curves for the aspartate aminotransferase reaction in darkened (black dots) and illuminated (clear dots) Zea mays leaves in response to increasing L-aspartate or 2-oxoglutarate (2-OG) concentrations. Data are based on six biological replicates and error bars show the standard deviation (SD). SE, standard error. Units for initial velocity are Abs/min.
While the changes in $K_M$ for L-aspartate and 2-oxoglutarate were significantly different in *Setaria viridis* and *Sorghum bicolor*, aspartate aminotransferase displayed similar rates of activity. In contrast, the activity of aspartate aminotransferase in *Zea mays* decreased by approximately 0.7-fold in saturating conditions with 10 mM L-aspartate (Figure 5.8). Aspartate aminotransferase from darkened leaves reached saturation in the presence of at least 2.5 mM L-aspartate, while the aspartate aminotransferase in illuminated leaves did not saturate until 10 mM L-aspartate. Additionally, the $K_M$ for L-aspartate increased from $154.3 \pm 36.5 \, \mu M$ to $2473 \pm 402.1 \, \mu M$ after the dark to light transition in *Zea mays*, which was higher than the observed changes to the affinity for L-aspartate in *Setaria viridis* and *Sorghum bicolor* in response to light. The affinity for L-aspartate and 2-oxoglutarate decreased in the dark to light transition in the three C₄ grasses, but the dark and light affinity differences for L-aspartate were most evident in *Zea mays*, followed by *Sorghum bicolor* and *Setaria viridis* (Table 5.1).

Table 5.1. Aspartate aminotransferase affinity for primary and secondary substrates after 15.5 h illumination. Summary of the change in aspartate aminotransferase affinities for L-aspartate and 2-oxoglutarate in response to light. In 15.5 h illuminated leaves, the decrease in affinity for L-aspartate was highest in *Zea mays*, moderate in *Sorghum bicolor* and smallest in *Setaria viridis*. The affinity for 2-oxoglutarate increased in *Sorghum bicolor* after the transition to light, but decreased in both *Setaria viridis* and *Zea mays*. Up-arrows indicate increase in affinity (a negative fold change, decrease in $K_M$), while down-arrows indicate decrease in affinity (a positive fold change, increase in $K_M$).

<table>
<thead>
<tr>
<th></th>
<th>L-Aspartate</th>
<th>2-oxoglutarate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affinity (fold change)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Sorghum bicolor</em></td>
<td>1.9 ↓</td>
<td>0.51 ↑</td>
</tr>
<tr>
<td><em>Setaria viridis</em></td>
<td>0.44 ↓</td>
<td>0.20 ↓</td>
</tr>
<tr>
<td><em>Zea mays</em></td>
<td>15.0 ↓</td>
<td>0.30 ↓</td>
</tr>
</tbody>
</table>
Aspartate aminotransferase affinity for dimethyl-2-oxoglutarate, an analogue of 2-oxoglutarate produced during glucose metabolism (Odegaard et al., 2010), was also tested. The changes in $K_M$ for 2-oxoglutarate and dimethyl-2-oxoglutarate were larger in *Zea mays* followed equally by *Sorghum bicolor* and *Setaria viridis* (Figure 5.9). Aspartate aminotransferase activity was reduced across all three C$_4$ species, but was relatively higher in *Setaria viridis*. Moreover, the response to dimethyl-2-oxoglutarate was more consistent in darkened and illuminated *Setaria viridis* leaves, unlike *Sorghum bicolor* and *Zea mays*. The activity of aspartate aminotransferase in *Sorghum bicolor* and *Zea mays* declined in the presence of dimethyl-2-oxoglutarate and less consistent than in *Setaria viridis*. 
**Fig. 5.9.** Dark and light *Setaria viridis* aspartate aminotransferase responds better to dimethyl-2-oxoglutarate. Michaelis-Menten curves for the aspartate aminotransferase reaction in darkened (black dots) and illuminated (white dots) leaves in response to increasing dimethyl-2-oxoglutarate concentrations. Data are based on six biological replicates and error bars show the standard deviation (SD). SE, standard error. Units for initial velocity are Abs/min.

<table>
<thead>
<tr>
<th>Species</th>
<th>Dimethyl-2-oxoglutarate (K_M ± SE µM)</th>
<th>Dark</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Sorghum bicolor</em></td>
<td></td>
<td>10.8 ± 1.6</td>
<td>152.1 ± 32.1</td>
</tr>
<tr>
<td><em>Setaria viridis</em></td>
<td></td>
<td>83.4 ± 7.8</td>
<td>72.6 ± 7.6</td>
</tr>
<tr>
<td><em>Zea mays</em></td>
<td></td>
<td>36.9 ± 3.1</td>
<td>1049 ± 230</td>
</tr>
</tbody>
</table>

**Species**

- *Sorghum bicolor*
- *Setaria viridis*
- *Zea mays*
5.2.2 The effect of metabolites on the activity of aspartate aminotransferase

The sensitivity of aspartate aminotransferase activity was tested in response to metabolites produced during C₄ photosynthesis. In Setaria viridis, Sorghum bicolor and Zea mays, increasing concentrations of L-malate reduced activity in darkened and illuminated leaves (Figure 5.10). However, the magnitude of the inhibitory effect of L-malate was different across the three C₄ species. In Figure 5.10B, aspartate aminotransferase from Sorghum bicolor illuminated leaves showed the highest decrease in activity in the presence of 20 and 50 mM L-malate when compared to the controls (paired t-test: df=5, P < 0.001). This trend was closely followed by Setaria viridis (Figure 5.10A), in which the changes of enzyme activity between controls and 10, 20 and 50 mM L-malate was also significant (paired t-test: df=5, P < 0.001). Conversely, the activity in darkened and illuminated Zea mays leaves was the least affected by increasing L-malate concentrations (Figure 5.10C) when compared to aspartate aminotransferase activity in Setaria viridis and Sorghum bicolor. In darkened leaves of Sorghum bicolor, the inhibitory effect of L-malate was evident in the presence of 10 mM (Figure 5.10B), whereas the activity of aspartate aminotransferase in illuminated leaves was unaffected, suggesting that in the light, the enzyme was less responsive to increasing L-malate concentrations. Additionally, the activity of aspartate aminotransferase from darkened leaves was not further reduced in the presence of 20 or 50 mM L-malate, while activity decreased in illuminated leaves. This effect was also observed in Setaria viridis leaves (Figure 5.10A), however aspartate aminotransferase activity in Setaria viridis was lower in the presence of 10 mM L-malate. Although the differences in activity between control reaction assays and treatment with L-malate were significant, the changes in activity between the different concentrations of L-malate were not significant.
Fig. 5.10.
Fig. 5.10. 3-PGA, L-alanine and L-malate affect aspartate aminotransferase activity differently in three C₄ species. Activity bar graphs show the rate of aspartate aminotransferase activity from darkened (grey bars) and illuminated (white bars) *Setaria viridis* (A), *Sorghum bicolor* (B) and *Zea mays* (C) leaves in response of 10, 20 and 50 mM L-malate and L-alanine, and 1 mM 3-PGA. An inhibitory effect is observed by bars lowering relatively to controls, whereas an activation effect is observed by bars increasing relatively to controls. The changes to aspartate aminotransferase activity in the presence of L-malate and L-alanine were most significant in *Sorghum bicolor* and least significant in *Zea mays*. In *Setaria viridis*, activity lowered in response to L-malate, but not L-alanine. 1 mM 3-PGA had no effect on the activity of aspartate aminotransferase across all species. 3-PGA, phosphoglycerate. The significance between the average activity between control and metabolite concentrations were determined by paired t-tests, where * = P < 0.05, ** = P < 0.005, *** = P < 0.0005. Error bars show the standard deviation.

Unlike *Sorghum bicolor* and *Zea mays*, the activity of aspartate aminotransferase from darkened and illuminated *Setaria viridis* leaves was not inhibited by increasing concentrations of L-alanine (Figure 5.10A–C). However, despite that the activity of aspartate aminotransferase was lower in response to L-alanine in *Sorghum bicolor*, enzyme activity in *Zea mays* in response to L-alanine was enhanced across all concentrations (Figure 5.10C), though not significantly in darkened leaves.
Fig. 5.11.
Fig. 5.11. Aspartate aminotransferase activity in response to PEP and pyruvate. Activity bar graphs show the rate of aspartate aminotransferase activity from darkened (grey bars) and illuminated (white bars) *Setaria viridis* (A), *Sorghum bicolor* (B) and *Zea mays* (C) leaves in response of 5, 10 and 25 mM pyruvate and 10, 20 and 50 mM phosphoenolpyruvate (PEP). PEP reduces aspartate aminotransferase activity equally across C4 species, while pyruvate has no effect on activity. The significance between the average activity between control and metabolite concentrations were determined by paired t-tests, where * = P < 0.05, ** = P < 0.005, *** = P < 0.0005. Error bars show the standard deviation.

The activity of aspartate aminotransferase in *Setaria viridis*, *Sorghum bicolor* and *Zea mays* was differently affected by L-malate and L-alanine and it was evident that the sensitivity to L-malate and L-alanine of light and dark activities were most significant in *Sorghum bicolor* (Figure 5.10B). Activity in *Setaria viridis* was most affected by L-malate, but not by L-alanine (Figure 5.10A), whereas increasing concentrations of L-malate and L-alanine had the least inhibitory effects on aspartate aminotransferase activity in *Zea mays* (Figure 5.10C). There was no effect on the activity of 1 mM 3-PGA on aspartate aminotransferase across all C4 species.
Aspartate aminotransferase activity was unaffected by 5 mM pyruvate (Figure 5.11A–C). Changes to mean activities between controls and 25 mM pyruvate were most significant in darkened Setaria viridis leaves (paired t-test; df = 5, \(P < 0.001\)) as shown in Figure 5.11A, although, the standard deviation showed a large degree of variation in enzyme activity across the six biological replicates. Furthermore, the inhibitory effect of pyruvate was more evident in illuminated Sorghum bicolor leaves in the presence of 10 mM pyruvate (paired t-test; df = 5, \(P = 0.0037\)) (Figure 5.11B) than in the presence of 25 mM pyruvate (paired t-test; df = 5, \(P = 0.0359\)). The activity of aspartate aminotransferase in Zea mays leaves was the least affected by pyruvate (Figure 5.11C), although changes in activity in illuminated leaves in the presence of 10 mM and 25 mM were significant (paired t-test; df = 5, \(P < 0.05\)). Conversely, all C₄ plants were equally affected by increasing concentrations of PEP. Aspartate aminotransferase remained active in darkened and illuminated leaves in up to 50 mM PEP, but activity decreased 0.81-fold in Zea mays, 0.84-fold in Sorghum bicolor and 0.83-fold in Setaria viridis, when compared to controls (Figure 5.11A–C).

5.2.3 Specific activity of alanine aminotransferase

Alanine aminotransferase from Sorghum bicolor, Setaria viridis and Zea mays was assayed in vitro in the reverse direction towards the formation of pyruvate by coupling to lactate dehydrogenase (LDH). No enzyme activity was observed in reaction mixtures containing 20 mM L-alanine. Based on alanine aminotransferase assays in Arabidopsis thaliana (Miyashita et al., 2007), the concentration of L-alanine was increased to 25, 50 and 75 mM. However, after 20 min, there was no activity. In the negative control, there was no activity observed in the presence of L-aspartate.
Alanine aminotransferase (AlaAT) reaction mechanism towards the formation of malate. Metabolites in purple were added to the reaction mixture and metabolites in green were produced during the reaction. Assay measures the formation of malate through the oxidation of NADH (in red). Pyr, pyruvate; 2-OG, 2-oxoglutarate; OAA, oxaloacetate. Aspartate aminotransferase (AspAT) and malate dehydrogenase (MDH).

Alanine aminotransferase was also assayed in the forward direction towards the formation of malate, coupling to aspartate aminotransferase and malate dehydrogenase (MDH) (Figure 5.12). Activity was measured 10 min after priming the reaction assay with leaf protein. Preliminary measurements showed that the activity of alanine aminotransferase was highest in Setaria viridis, when compared to Sorghum bicolor and Zea mays (Figure 5.13).

Fig. 5.13. Alanine aminotransferase activity in illuminated Zea mays, Sorghum bicolor and Setaria viridis leaf lysates. Alanine aminotransferase was assayed in the forward direction towards the formation of alanine in leaf lysates normalised against the lowest total protein concentration. Preliminary data are based on one biological replicate from 15.5 h illuminated leaves.
The assay was repeated with four biological replicates, normalised against the lowest total protein concentration. The second assay showed that the activity of alanine aminotransferase was highest in *Setaria viridis* (Figure 5.14), consistent with preliminary data in Figure 5.13. There was no significant difference in the activity of alanine aminotransferase in *Zea mays* and *Sorghum bicolor*. Also, there was no change in activity between darkened and illuminated leaf protein.

![Alanine aminotransferase activity in Setaria viridis leaf lysates is higher than in Zea mays and Sorghum bicolor.](image)

**Fig. 5.14.** Alanine aminotransferase activity in *Setaria viridis* leaf lysates is higher than in *Zea mays* and *Sorghum bicolor*. Alanine aminotransferase was assayed in the forward direction towards the formation of alanine in leaf lysates normalised against the lowest total protein concentration. Data are based on four biological replicates and error bars show the standard deviation.
5.3 Discussion

5.3.1 Aspartate aminotransferase affinity aspartate is sensitive to dark–light transitions

C₄ plants that use NADP-malic enzyme as the primary decarboxylase may also use a supplementary decarboxylation pathway via PEPCK through the conversion of oxaloacetate to aspartate and then back to oxaloacetate (Pick et al., 2011; Bräutigam et al., 2014; Wang et al., 2014) (Figure 5.2). Previous studies have noted that the C₄-acid transfer via aspartate aminotransferase is essential in plants with low NADP-malic enzyme activity (Andrews et al., 1971; Hatch and Mau, 1973; Meister et al., 1996). However, the regulation of carbon flux through aspartate in C₄ grasses that predominantly use NADP-malic enzyme remains unclear. The results in this Chapter show that the catalytic properties of aspartate aminotransferase from three NADP-malic enzyme subtypes are sensitive to dark to light transitions and there may be interspecies differences, particularly with regards to regulation by transfer acids. In the three C₄ species studied, the affinity for L-aspartate declined in illuminated leaves, which suggests that the conversion of aspartate to oxaloacetate may be regulated in illuminated conditions, but could correlate to the initial increase of aspartate in illuminated leaves (Leegood and von Caemmerer, 1989). Increased affinity for L-aspartate in darkened conditions is consistent with previous findings that showed that major aspartate aminotransferase isoforms have increased activity in dark-grown plants (Hatch and Mau, 1973). In activity assays (Figure 5.5), there was no change in aspartate aminotransferase activity in darkened and illuminated leaves, which is not characteristic of major isoforms as previously shown (Hatch and Mau, 1973). It is possible that activity from minor isoforms was observed and could have masked changes in dark and light activity of major isoforms.

In *Megathyrsus maximus* (PEPCK subtype), the $K_M$ for L-aspartate is 2.3 mM (Numazawa et al., 1989), which is close to the $K_M$ for L-aspartate in illuminated leaves of *Setaria viridis* (2.3 mM) and *Zea may* (2.5 mM), but not *Sorghum bicolor* (5.2 mM), which may require relatively less PEPCK activity (Gutierrez et al., 1974). In addition, the $K_M$ for oxaloacetate in *Megathyrsus maximus* is 0.049 mM (Numazawa et al., 1989) and although not determined in this study, the $K_M$ for oxaloacetate may be lower than
that of L-aspartate in the NADP-malic enzyme subtypes. Large fluxes of oxaloacetate are generated by PEPC in the light, but oxaloacetate does not accumulate to high concentrations in the cell, because it is quickly converted to malate by reduction via MDH in NADP-malic enzyme subtypes (Hatch, 1987). In species like Zea mays that can use PEPCK as a supplementary decarboxylation pathway, a relatively higher affinity for oxaloacetate may be needed to maintain high rates of aspartate aminotransferase activity towards the formation of aspartate. In Sorghum bicolor, the $K_M$ for L-aspartate in illuminated leaves was substantially higher than the $K_M$ for L-aspartate in Setaria viridis and Zea mays. Unlike Zea mays, Sorghum bicolor almost exclusively uses NADP-malic enzyme (Gutierrez et al., 1974), whereas Zea mays can use both NADP-malic enzyme and PEPCK (Hatch, 1987). Thus, a relatively lower affinity for L-aspartate may correlate to both the lower activity of PEPCK in Sorghum bicolor leaves and relatively lower expression of PEPCK than in Zea mays (Figure 4.16).

Differences in the decarboxylation chemistry of these C₄ grasses may have also contributed to the decrease in affinity for 2-oxoglutarate in Sorghum bicolor, while in both Setaria viridis and Zea mays, the affinity for 2-oxoglutarate increased after the dark to light transition. Unlike oxaloacetate, the amount of aspartate available in the Zea mays leaf can vary in response to light intensity or dark to light transitions, and the concentration increases upon illumination and then rapidly declines, but remains between 25 to 60 mM in the whole cell (Hatch and Osmond, 1976; Leegood and Furbank, 1984; Hatch, 1987; Leegood and von Caemmerer, 1989). Therefore, it is unlikely that the activity of aspartate aminotransferase towards the formation of oxaloacetate will be limited by the amount of aspartate when the $K_M$ for aspartate is high, but if the $K_M$ for oxaloacetate were high, then the reverse reaction (towards the formation of aspartate) may be limited by the availability of oxaloacetate. Alternatively, if the affinity for L-aspartate were higher, then aspartate aminotransferase would readily catalyse the conversion of aspartate to oxaloacetate, decreasing the amount of substrate for PEPCK.
Fig. 5.15. Interchange between aspartate and oxaloacetate by aspartate aminotransferase (AAT).

In illuminated leaves, there was a 15-fold decrease in affinity for L-aspartate in Zea mays, which correlates to the increase in aspartate after the onset of light. This suggests that there may be additional need for aspartate aminotransferase activity towards the formation of oxaloacetate in darkened conditions. This activity might not be associated with C₄ photosynthesis, and instead aspartate aminotransferase in darkened leaves may be catalysing the formation of oxaloacetate from aspartate for cellular respiration or for biosynthesis. In fact, oxaloacetate serves a central role in metabolism (Buchanan et al., 2015) and requirements may differ between dark and light conditions (Figure 5.15). There may also be a link between the phosphorylation sites identified solely in darkened leaves of Setaria viridis and the relatively higher affinity for L-aspartate in darkened conditions. From these observations, it is probable that the activity of aspartate aminotransferase is partitioned between photosynthesis in illuminated conditions and other cellular processes in darkened conditions, somewhat similar to the regulation of photosynthetic PEPC in leaves and anaplerotic activity in non-photosynthetic tissues (O'Leary et al., 2011; Shane et al., 2013; Ruiz-Ballesta et al., 2014). These results also suggest that the regulation of aspartate aminotransferase may be an additional control for downstream PECK activity. This is supported by the fact that the amount of PECK protein and PECK activity in C₄ plants changes very little in darkened and illuminated conditions and PECK needs to be regulated to prevent depletion of ATP (Carnal et al., 1993; Walker et al., 2002). Although PECK is regulated by reversible phosphorylation, this occurs in darkened leaves of Megathyrsus maximus (Walker and Leegood, 1995; Walker et al., 1997; Bailey et al., 2007). In Zea mays, PECK appears to be phosphorylated in illuminated leaves (Chao et al., 2014) and its regulation might be distinctly different than in PECK subtypes. Therefore, the regulation of aspartate
aminotransferase in illuminated conditions may serve to regulate PEPCK activity by controlling the supply of oxaloacetate.

In addition to this, there may be some aspartate aminotransferase activity associated with chloroplasts or mitochondria (Hatch and Mau, 1973; Hatch, 1987; Meister et al., 1996). This may contribute to additional regulation by pH, as observed for NADP-malic enzyme. Moreover, the stability of aspartate aminotransferase and its interaction with competitive inhibitors can be influenced by pH (Bonsib et al., 1975). When assayed at pH 8.0 in the presence of dimethyl-2-oxoglutarate, an analogue of 2-oxoglutarate (Odegaard et al., 2010), aspartate aminotransferase still followed Michaelis-Menten kinetics. However, the affinity for dimethyl-2-oxoglutarate was higher than the affinity for 2-oxoglutarate in darkened leaves of *Sorghum bicolor, Setaria viridis* and *Zea mays* (Table 5.2), suggesting that in the absence of 2-oxoglutarate, dimethyl-2-oxoglutarate could be an effective replacement.

Table 5.2. Aspartate aminotransferase affinity for 2-oxoglutarate and dimethyl-2-oxoglutarate. Up-arrows indicate an increase in affinity against the original $K_M$ value for 2-oxoglutarate (a negative fold change, decrease in $K_M$), while down-arrows indicate a decrease in affinity (a positive fold change, increase in $K_M$).

<table>
<thead>
<tr>
<th></th>
<th>Dark</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affinity (fold change)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Sorghum bicolor</em></td>
<td>0.9↑</td>
<td>0.7↓</td>
</tr>
<tr>
<td><em>Setaria viridis</em></td>
<td>0.5↑</td>
<td>0.6↑</td>
</tr>
<tr>
<td><em>Zea mays</em></td>
<td>0.66↑</td>
<td>6.3↓</td>
</tr>
</tbody>
</table>

In the presence of 2 mM dimethyl-2-oxoglutarate, aspartate aminotransferase from illuminated *Zea mays* leaves did not saturate and in *Sorghum bicolor*, the activity of aspartate aminotransferase in the presence of dimethyl-2-oxoglutarate declined. This indicates that the properties of aspartate aminotransferase from illuminated leaves of *Zea mays* and *Sorghum bicolor* may be distinctly different than that of aspartate aminotransferase from *Setaria viridis*, which remained equally active in darkened and illuminate leaf lysates in the presence of dimethyl-2-oxoglutarate. This may modulate $C_4$-specific function after the onset of light. Furthermore, computer analyses have predicted that dimethyl-2-oxoglutarate is a weak inhibitor of aspartate aminotransferase.
(Bonsib et al., 1975), which is consistent with the minimal effect of dimethyl-2-oxoglutarate on the activity of the transaminase (Figure 5.9). However, these effects may be more pronounced in the presence of stronger dicarboxylic inhibitors and may conflict with native substrate binding and affect the role of aspartate aminotransferase in C₄ photosynthesis. To assess the stability of the enzyme, it would be useful to determine if the activity of aspartate aminotransferase and affinity for L-aspartate and 2-oxoglutarate changes in low and high pH.

5.3.2 C₄ metabolites may regulate the activity of aspartate aminotransferase

C₄ photosynthesis involves the efficient movement of carbon between mesophyll and bundle sheath cells by the transfer acids malate, aspartate, alanine, pyruvate and PEP (Hatch, 1987; Bräutigam et al., 2014). Previous studies have shown that these transfer acids are crucial for the coordination of carboxylation and decarboxylation pathways (Chapman and Hatch, 1981; Furbank and Leegood, 1984; Leegood, 1985; Stitt and Heldt, 1985; Leegood and von Caemmerer, 1989; Bräutigam et al., 2014; Arrivault et al., 2016). Results presented in this Chapter have shown that the activity of aspartate aminotransferase was activated or inhibited in the presence of C₄ transfer acids. This suggests that similar to the coordination of C₃ and C₄ cycles afforded by the exchange of transfer acids, there could be additional mechanisms that regulate decarboxylation by either NADP-malic enzyme or PEPCK/NAD-malic enzyme in C₄ species that operate the dual-decarboxylation system (Figure 5.2) (Chapman and Hatch, 1979).
Fig. 5.16. Relative expression of aspartate aminotransferase (AAT) and alanine aminotransferase (AlaAT). Expression in reads per kilobase million (rpkm). Transcriptomic data obtained from PA Christin (unpublished).

This is supported by the inhibition of aspartate aminotransferase activity from Sorghum bicolor in the presence of L-alanine (Figure 5.10A), while aspartate aminotransferase from Setaria viridis was not inhibited (Figure 5.10B) and in Zea mays, the transaminase was activated (Figure 5.10C). This suggests that aspartate aminotransferase from Sorghum bicolor might lack regulatory controls that would otherwise maintain high in vivo activity in an excess of alanine. This may also reflect the decarboxylation pathway in Sorghum bicolor, which unlike Setaria viridis or Zea mays, Sorghum bicolor uses NADP-malic enzyme and may not have high in vivo activities of aspartate aminotransferase or alanine aminotransferase (Andrews et al., 1971; Gutierrez et al., 1974). Conversely, aspartate aminotransferase from Zea mays contains more aspartate aminotransferase activity (Andrews et al., 1971), albeit has a lower expression of aspartate aminotransferase than Sorghum bicolor and Setaria viridis (Figure 5.16). More importantly, transcriptomic data shows that the expression of alanine aminotransferase in Zea mays is substantially higher than that of Setaria viridis and Sorghum bicolor. This may indicate that aspartate aminotransferase in Zea mays is adapted to maintain activity in a high concentration of L-alanine, because of the relative dependence on alanine aminotransferase activity. Moreover, increasing the concentration of L-alanine stimulated the activity of aspartate aminotransferase from illuminated Zea mays leaf lysates. This suggests that aspartate aminotransferase may be linked to alanine
aminotransferase activity, and is consistent with decarboxylation via PEPCK/NAD-malic enzyme pathways, which require both aspartate aminotransferase and alanine aminotransferase activity (Andrews et al., 1971; Bräutigam et al., 2014; Wang et al., 2014).

Aspartate aminotransferase from darkened leaves of Zea mays was not activated by L-alanine (Figure 5.10C), suggesting that the dark-form of the transaminase may not be preconditioned for C₄-specific activity. In the absence of L-alanine, the formation of oxaloacetate seemed to be favoured in darkened conditions (Figure 5.8), but as noted, may not be associated with C₄-specific activity. Also, since aspartate aminotransferase from Zea mays appeared to be stimulated by the C₄ metabolite solely in illuminated leaf lysates, this might correspond to an activation mechanism for C₄-specific function and could signal the formation of oxaloacetate in illuminated leaves and stimulate PEPCK activity. Conversely, aspartate aminotransferase from Setaria viridis was not affected by L-alanine, which may reflect that Setaria viridis depends more on the flux of L-alanine than Sorghum bicolor, but may not utilise a secondary decarboxylation pathway via PEPCK/NAD-malic enzyme as readily as Zea mays. This is partially supported by the relatively higher alanine aminotransferase activity in Setaria viridis (Figure 5.14) and higher alanine aminotransferase expression when compared to alanine aminotransferase from Sorghum bicolor, but substantially lower expression of alanine aminotransferase when compared to Zea mays (Figure 5.16).

In addition to this, L-malate stimulated the activity of aspartate aminotransferase in illuminated Zea mays leaf lysates, whereas in darkened and illuminated Setaria viridis leaf lysates, the activity of aspartate aminotransferase was inhibited by an excess of L-malate, which is consistent with previous findings (Chapman and Hatch, 1979). Furthermore, the activation of aspartate aminotransferase towards the formation of oxaloacetate in the presence of L-malate may link malate metabolism via NADP-malic enzyme and oxaloacetate decarboxylation via PEPCK (Figure 5.17).
Fig. 5.17. Simple schematic depicting a mechanism that could regulate the relative flux through malate or aspartate in a dual-decarboxylation system in *Zea mays*. Red arrows indicate increase rate of formation (in the indicated direction; metabolite in bold), blue arrows (dashed lines) indicate reduced rate of formation and orange arrows indicate PEPC inhibition by aspartate or malate. For simplicity, other feedback inhibition mechanisms are not shown. Figure based on experimental data shown in Figures 5.10 and 5.11 and data from Chapman and Hatch (1979). Asp, aspartate; Pyr, pyruvate; Mal, malate; OAA, oxaloacetate.

This link was previously shown when the formation of pyruvate via malate metabolism was stimulated in the presence of aspartate (Chapman and Hatch, 1979). This implies that NADP-malic enzyme activity towards the formation of pyruvate is favoured when aspartate is formed, but this may affect PEPCK activity, because the formation of oxaloacetate may not be favoured. Alternatively, if the formation of oxaloacetate is favoured and aspartate declines, then the formation of pyruvate via NADP-malic enzyme may decline, implying that malate would be more readily available since the rate at which is it decarboxylated may decrease. Conversely, if there were more aspartate and pyruvate, it would imply that the rate at which malate is decarboxylated increases. The link between aspartate, pyruvate and malate (Figure 5.17), as part of an extended aspartate–malate shuttle, is further supported by the decrease in the rate at which aspartate is converted to oxaloacetate in an excess of pyruvate (Figure 5.11C).
Therefore, in the same way that aspartate can induce the formation of pyruvate and stimulate the activity of NADP-malic enzyme (Müller et al., 2008), then malate might induce the formation of oxaloacetate and stimulate PEPCK activity. This activation effect, which was solely evident in aspartate aminotransferase from Zea mays, may reflect the fine control of PEPCK and NADP-malic enzyme decarboxylation, particularly when the activity of either decarboxylase is reduced or not favoured. In rats (Rattus norvegicus), aspartate aminotransferase activation by L-malate may contribute to the fine control and reversibility of the transaminase, and most importantly, may regulate the relative flux through aspartate or malate (McKenna et al., 2006). Whether the oxaloacetate formed in this mechanism contributes to malate metabolism or PEPCK decarboxylation is not clear. The inhibitory effect of L-malate observed in Sorghum bicolor and Setaria viridis may suggest the relatively lower dependence on PEPCK-dependent decarboxylation compared to Zea mays, as previously shown in transcriptomic analyses (Figure 4.16).

10 mM PEP had no effect on the activity of aspartate aminotransferase from illuminated leaves of Setaria viridis (Figure 5.11A) and Zea mays (Figure 5.11C), while 20 mM and 50 mM PEP affected the activity of aspartate aminotransferase from the three C₄ grasses equally. This suggests that the activity of aspartate aminotransferase may be linked to the activity of PEPC, but may not be as significant as control by L-malate or L-alanine, since the concentration of PEP used in the experiments was well above the physiological amounts (Leegood and von Caemmerer, 1989). The amount of PEP increases upon illumination, but is rapidly assimilated into oxaloacetate by PEPC during photosynthesis (Leegood and von Caemmerer, 1988; Leegood and von Caemmerer, 1989) and PEP can serve as an activator for PEPC activity (Tovar-Méndez and Muñoz-Clares, 2001). Aspartate can inhibit PEPC activity (Mareš et al., 1979), and therefore a inhibitory mechanism regulated by PEP or aspartate may link aspartate aminotransferase and PEPC activities, respectively, and might regulate the formation of oxaloacetate when the rate of PEP carboxylation is reduced or in darkened conditions, when photosynthesis is not occurring. This may lead to higher aspartate aminotransferase activity in darkened leaves and contribute to the fine control of PEPC activity during light transitions, in addition to regulation by phosphorylation or other C₄ metabolites (Budde and Chollet, 1986; Doncaster and Leegood, 1987; Hatch, 1987; Leegood and von Caemmerer, 1988).
5.3.3 Alanine aminotransferase may be differently regulated in C₄ species

Alanine aminotransferase activity measured towards the formation of pyruvate was negligible when assayed in the presence of 20, 50 and 75 mM L-alanine. Previous studies have shown that alanine aminotransferase has a high $K_M$ for L-alanine ($K_M$ 3.03 mM in *Zea mays* compared to $K_M$ 10.4 mM in *Arabidopsis thaliana*) and the reaction assays require a high concentration of the substrate (Miyashita *et al.*, 2007; Duff *et al.*, 2012; Kendziorek *et al.*, 2012; McAllister *et al.*, 2013). In *Zea mays*, the $K_M$ for pyruvate is lower than the $K_M$ for L-alanine (Duff *et al.*, 2012), which suggests some selectivity towards pyruvate binding. This may also be part of the mechanism to preferentially bind pyruvate in illuminated conditions, since the amount of pyruvate decreases with increasing light intensity, and while the amount of alanine declines after the onset of light, it remains between 30 to 70 mM in darkened or illuminated conditions (Leegood and Furbank, 1984; Leegood and von Caemmerer, 1989).

Preference for pyruvate might be significant for the formation alanine in C₄ plants, particularly in PEPCK/NAD-malic enzyme subtypes, where the balance of amino acids between mesophyll and bundle sheath cells must be kept by transferring alanine from bundle sheath to mesophyll cells after aspartate is transferred from mesophyll to bundle sheath cells (Hatch, 1987; Weber and Bräutigam, 2013). Furthermore, a concentration gradient for the movement of pyruvate from bundle sheath cells to mesophyll cells is not required (Arrivault *et al.*, 2016). In addition to this, theoretical models used to model the energy costs of C₄ photosynthesis have also shown that in PEPCK subtypes, the transport of alanine, presumably after the conversion from pyruvate via alanine aminotransferase, and PEP via PEPCK compensates the reduction of malate metabolism and pyruvate transport to regenerate PEP via PPDK (Wang *et al.*, 2014). This suggests that activity of alanine aminotransferase may be preferential for the pyruvate reaction, towards the formation of alanine, since a concentration gradient of alanine is required to operate the C₄ cycle (Arrivault *et al.*, 2016).

The relative activity of alanine aminotransferase in the three C₄ grasses did not correlate to transcriptomic analyses (Figure 5.16). High expression of alanine aminotransferase and aspartate aminotransferase expression in *Zea mays* correlates to aspartate–oxaloacetate and pyruvate–alanine interconversion pathways in species that use PEPCK
and NADP-malic enzyme (Hatch, 1987; Wang et al., 2014). However, changes to alanine aminotransferase activity as seen on Figure 5.14 may not correlate to requirements of PEPCK, but rather to the relative activity of NADP-malic enzyme. This is supported by the activity of alanine aminotransferase being lower in Zea mays than in Setaria viridis, when Zea mays contains relatively more PEPCK than either Sorghum bicolor or Setaria viridis (Gutierrez et al., 1974).

Considering these results, the alanine aminotransferase reaction towards the formation of alanine requires further optimisation and investigation in these C₄ species. One major problem was that the reaction took more than 20 min to have detectable changes in NADH oxidation, compared to only a few minutes to see a decline in the concentration of NADH in the aspartate aminotransferase reaction. Also, whether the reaction towards the formation of alanine is being masked by catalysis in the reverse direction, is unclear, but may be unlikely since activity was not previously observed in high concentrations of L-alanine. Improvements to the assay of the activity of alanine aminotransferase and determining its kinetic properties in darkened and illuminated leaves may provide some indication to its regulation and role in NADP-malic enzyme subtype C₄ grasses.

5.3.4 Conclusion

The results in this chapter showed that solely the aspartate aminotransferase from illuminated Zea mays leaf lysates was activated by an excess of L-malate and L-alanine, whereas aspartate aminotransferase activity in Sorghum bicolor was inhibited by both L-malate and L-alanine. In Setaria viridis, the inhibitory effect was only observed in the presence of L-malate. These results support the role of aspartate aminotransferase as a link between PEPCK and NADP-malic enzyme decarboxylation, as well as align with the hypothesis that different NADP-malic enzyme C₄ subtypes might have differing demands for aspartate aminotransferase activity and PEPCK-dependent decarboxylation. In addition, these effects were more predominant in illuminated leaf lysates and might correspond to a C₄-specific aspartate aminotransferase isoform in NADP-malic enzyme subtypes. Effector metabolites, rather than light conditions, might assert more control over the activation of the transaminase, which appeared to have a lower affinity for L-aspartate after transition into the light period.
Chapter 6 – General Discussion

6.1 Regulating the relative flux through malate and aspartate

In C₄ photosynthesis, assimilated carbon is delivered at the site of Rubisco via two main pathways through fluxes of malate or aspartate (Bellasio and Griffiths, 2014). Historically, C₄ plants exist as three distinct subtypes, NAD- and NADP-malic enzyme and PEPCK, and are classified by their primary decarboxylase (Hatch, 1987). It has become apparent, however, that certain NADP-malic enzyme C₄ grass species can partition between NADP-malic enzyme and PEPCK decarboxylation using malate and aspartate as C₄ transfer acids. Furthermore, in PEPCK subtypes, NAD-malic enzyme contributes to malate metabolism in mitochondria of bundle sheath (Hatch, 1987; Bräutigam et al., 2014). Recent evidence suggests that the decarboxylation pathways of C₄ grasses are flexible, and multiple decarboxylases may have been selected for during evolution (Furbank, 2011; Wang et al., 2014). Moreover, the relative amounts of PEPCK, NADP-malic enzyme or NAD-malic enzyme vary among subtypes. For instance, certain NADP-malic enzyme subtypes, such as Zea mays, have appreciable amounts of PEPCK activity, while others, such as Sorghum bicolor or Setaria viridis have extremely low PEPCK activity (Gutierrez et al., 1974). Similarly, certain NAD-malic enzyme species, such as Panicum decompositum contain substantial amounts of PEPCK activity, while other NAD-malic enzyme species within the Panicum genus do not (Gutierrez et al., 1974). Interestingly, the decarboxylation chemistry of the dicot NAD-malic enzyme C₄ plant, Cleome gynandra, can be developmentally regulated and mid-aged leaves are able to use both NAD-malic enzyme and PEPCK decarboxylase pathways, with enhanced aspartate aminotransferase and alanine aminotransferase activity in bundle sheath mitochondria (Sommer et al., 2012). However, the regulation of carbon flux through malate or aspartate has not been properly understood in NADP-malic enzyme monocot grasses.

The findings presented in this thesis show that NADP-malic enzyme and aspartate aminotransferase may be differently regulated in NADP-malic enzyme subtype grasses and that the regulation of aspartate aminotransferase by C₄ transfer acids may reflect the relative flux through malate or aspartate and coordination of NADP-malic enzyme and
PEPCK decarboxylation in *Zea mays*. However, these mechanisms may not be present in C₄ grasses with lower expression of PEPCK, such as *Setaria viridis* and *Sorghum bicolor*, which may be less dependent on the flux through aspartate (Gutierrez et al., 1974). These results describe three distinct regulatory properties of NADP-malic enzyme and aspartate aminotransferase that could control the flux of carbon between mesophyll and bundle sheath cells of NADP-malic enzyme subtypes. These results also show that the classification of C₄ plants as distinct subtypes may be an oversimplification and may not accurately reflect the regulation of carbon flux, as the interspecies differences in the catalytic properties of NADP-malic enzyme and activation of aspartate aminotransferase in response to C₄ metabolites have shown.

There is also doubt to how much the malate–pyruvate pathway for NADP-malic enzyme and the aspartate–alanine pathway for PEPCK/NAD-malic enzyme overlap. Plants using a dual-decarboxylation system may operate an oxaloacetate–aspartate pathway via aspartate aminotransferase and a pyruvate–alanine pathway via alanine aminotransferase. Although alanine aminotransferase is primarily used for converting pyruvate to alanine from the decarboxylation of malate via NAD-malic enzyme in mitochondria of NAD-malic enzyme subtypes (Bräutigam et al., 2014), alanine aminotransferase could catalyse the interconversion of pyruvate and alanine downstream of malate decarboxylation by NADP-malic enzyme. Enhanced alanine aminotransferase activity was observed in leaves of *Cleome gynandra* in mid-development, while older leaves had reduced alanine aminotransferase activity (Sommer et al., 2012).

While Leegood (1985), Stitt and Heldt (1985) and Arrivault et al. (2016) have identified the metabolites and their respective concentration gradients that contribute to the flux of carbon, this rate of diffusion must be balanced with the rate of metabolism (von Caemmerer and Furbank, 2003). Based on modelling analysis (Wang et al., 2014), it is likely that the conversion of oxaloacetate to aspartate in mesophyll cells reduces carbon trafficking through malate, thus alleviating chloroplast transport mechanisms. However, converting too much aspartate may reduce the ratio of oxaloacetate to malate in mesophyll cells, hindering the uptake of oxaloacetate into chloroplasts for reduction to malate (Leegood, 2002). Also, movement of carbon through aspartate may reduce pools of malate necessary for downstream metabolism and decarboxylation by NADP-malic enzyme. This indicates that the diurnal regulation of the relative flux through aspartate or malate might be essential in plants that operate a dual-decarboxylation system.
Contribution of aspartate, along with malate, to the carbon pool was the subject of speculation since the elucidation of the carbon concentrating mechanism of C₄ photosynthesis (Hatch and Slack, 1966; Slack \textit{et al.}, 1969). Recent evidence suggests that the utilisation of aspartate and malate in certain C₄ plants may improve assimilation efficiencies and reduce energy requirements (Wang \textit{et al.}, 2014) and such flexibility in decarboxylation chemistry can allow C₄ plants to adapt better to environmental changes such as fluctuating light intensities (Furbank, 2011; Stitt and Zhu, 2014).

There has also been concern on which decarboxylation mechanism is most efficient and there is growing advocacy for PEPCK subtypes, with both PEPCK and NAD-malic enzyme activity, to be used as an engineering template for engineering the C₄ trait into existing C₃ crops. Recent evidence has shown that the PEPCK decarboxylation pathway requires fewer ATP than the NADP-malic enzyme pathway because PEP does not need to be regenerated via PPDK (Wang \textit{et al.}, 2014). However, decarboxylation via PEPCK does not generate reducing equivalents and these must be generated via NAD-malic enzyme or enhanced photosystem II activity in bundle sheath cells (Bräutigam \textit{et al.}, 2014). Since NADP-malic enzyme species have reduced photosystem II activity, because NADP-malic enzyme can generate reducing equivalents, relying fully on PEPCK decarboxylation might not be selected for during the evolution of decarboxylation mechanisms (Furbank, 2011). However, photosystem II activity varies between C₄ species and certain NADP-malic enzyme species, such as \textit{Zea mays}, have higher photosystem II activity than other NADP-malic enzyme grasses (Hardt and Kok, 1978; Walker and Izawa, 1979). Enhancement of bundle sheath cell photosystem II activities in species such as \textit{Zea mays}, which use a dual-decarboxylation system may compensate the generation of reducing equivalents that are not formed via NADP-malic enzyme. Differences in the kinetics of NADP-malic enzyme and aspartate aminotransferase may correlate to the relative rates of O₂ evolution in bundle sheath cells of NADP-malic enzyme subtypes. For instance, in \textit{Sorghum bicolor}, the activity of NADP-malic enzyme might be required at the onset of light, whereas in \textit{Zea mays}, the activity of NADP-malic enzyme at the onset of light might not be equally necessary. This may correlate to the absence of photosystem II activity in bundle sheath cells of \textit{Sorghum bicolor} (Hatch, 1978; Meierhoff and Westhoff, 1993), whereas \textit{Zea mays} contains appreciable amount of photosystem II and PEPCK activity in bundle sheath cells. The degree of O₂ evolution in bundle sheath cells of C₄ plants, particularly of NADP-malic enzyme subtypes, may
be interwoven with the relative flux through aspartate, as previously speculated (Chapman and Hatch, 1981; Meister et al., 1996). Furthermore, if the formation of reducing equivalents in the bundle sheath is not regulated, then it can influence the amount of 3-PGA that is translocated to mesophyll cells, via the triose phosphate porter, for reduction to triose phosphates, as well as alter the amount of triose phosphates that are translocated back to bundle sheath cells and used to regenerate RuBP under high photosynthetic loads (Hatch, 1987; Furbank, 2011). An early attempt in transforming C₃ plants with copies of the Zea mays NADP-malic enzyme caused bleaching of leaves in rice and reduced photosynthetic capacity probably by photoinhibition due to an increase in the supply of NADPH (Tsuchida et al., 2001). Although overexpression of NADP-malic enzyme can be deteriorating for transgenic C₃ plants, reduced grana stacking in bundle sheath cells of C₄ plants protects against photoinhibition by preventing the accumulation of NADPH in chloroplasts. This study is an example of how overexpressing the C₄-specific NADP-malic enzyme can induce phenotypic changes. Therefore, the formation of reducing equivalents in C₄-augmented plants must be considered, such that engineered protein networks do not perturb other transport mechanisms, such as the 3-PGA/triose-phosphate shuttle via the triose phosphate porter, which is relatively less expressed in C₃ grasses (Weber and von Caemmerer, 2010; Bräutigam et al., 2011). Compared to C₃ species, C₃–C₄ intermediates, such as Flaveria floridana and Moricandia arvensis, and C₄ species have considerably higher amounts of 3-PGA and triose phosphates (Badger et al., 1984; Leegood and von Caemmerer, 1994).

In addition to this, since PEPCK is cytosolic, its regulation may be simpler than that of NADP-malic enzyme, which is chloroplastic and may depend on regulation by light-dependent changes in the stromal pH (Asami et al., 1979; Edwards and Andreo, 1992; Bräutigam et al., 2014). It is argued, however, that NADP-malic enzyme might be the simplest to engineer, despite having to consider the expression of additional transporters for oxaloacetate, malate, 3-PGA and triose phosphates (Bräutigam et al., 2008; Weber and von Caemmerer, 2010; Bräutigam et al., 2011; Leegood, 2013). There may also be some importance to the selection of decarboxylation by NADP-malic enzyme, which was acquired as the primary decarboxylase in over 50% of C₄ grasses (Sage et al., 2011). Also, the catalytic properties of NADP-malic enzyme may not only differ between C₄ subtypes, but also between species using the same primary decarboxylation pathway. For instance, NADP-malic enzyme from Sorghum bicolor exhibited two distinct pH optima,
one of which contributed to maximal activity at pH 7.4. This may ensure high in vivo activity at the onset of light when the pH of the stroma is low (Werdan et al., 1975). In *Setaria viridis*, also a NADP-malic enzyme subtype, NADP-malic enzyme activity in illuminated leaf lysates had a broader pH dependence, and despite that maximal activity was not observed at a low pH, like in *Sorghum bicolor*, the differences in activity between low and high pH were negligible (Figure 4.5). These changes were not as distinct as the changes in NADP-malic enzyme activity in *Zea mays* between pH 7.0 and pH 8.4 (Figure 4.6). Unlike *Sorghum bicolor* and *Setaria viridis*, *Zea mays* has appreciable amounts of PEPCK expression and activity (Walker et al., 1997; Furumoto et al., 1999; Wingler et al., 1999). These results also indicate that in NADP-malic enzyme subtypes, the activity and regulation of NADP-malic enzyme activity may be influenced by a gradient of PEPCK-dependent decarboxylation. The dependence on PEPCK decarboxylation is higher in *Zea mays* and lower in *Sorghum bicolor* (Gutierrez et al., 1974) and PEPCK activity in *Setaria viridis* may be somewhere between that of *Zea mays* and *Sorghum bicolor* (Figure 6.1) (Gutierrez et al., 1974).

**Fig. 6.1.** *C₄* plants grouped by their primary decarboxylase. *C₄* species are grouped by their primary decarboxylase. Positioning depends on the relative expression of NADP-malic enzyme (far right), NAD-malic enzyme (far left) and PEPCK (towards the centre of the cross). NADP-malic enzyme subtypes used in this study are indicated by the boxes. Figure from APM Weber (unpublished).
While the kinetics of NADP-malic enzyme may depend on PEPCK, the regulation and activity of aspartate aminotransferase may depend on diurnal PEPC activity and correlate to changes in the affinity of PEP between dark and light transitions. Figure 4.2 shows that the affinity for PEP was lowest in 7.5 h darkened leaves. This is consistent with the downregulation of PEPC in darkened conditions to prevent PEP carboxylation (Jiao and Chollet, 1988; Jiao and Chollet, 1991; Chollet et al., 1996; O’Leary et al., 2011). Additionally, the amounts of PEP may be higher in the absence of carbon fixation and PEP content may not decrease as rapidly when PEPC is not active (Leegood and von Caemmerer, 1989). Conversely, in darkened leaves, aspartate aminotransferase exhibited a higher affinity for aspartate and 2-oxoglutarate (Figure 5.6), albeit the activity in the dark period was lower than in illuminated leaves (Figure 5.10B). If the concentration of PEP were higher, possibly due to the reduced PEPC activity in darkened leaves, then the activity of aspartate aminotransferase might be downregulated as well. This is supported by the inhibition of aspartate aminotransferase activity in the presence of PEP (Figure 5.11). In addition to this, PEPC affinity for PEP increased after the dark to light transition, but aspartate aminotransferase affinity for aspartate and 2-oxoglutarate declined. This implies that when PEP is readily metabolised by PEPC, the activity of aspartate aminotransferase can be higher, and it may be less inhibited by PEP. In the absence of PEP, the activity of aspartate aminotransferase was higher in illuminated leaf lysates of *Setaria viridis* (Figure 5.10B).

In addition to this, the activity of alanine aminotransferase may depend on NADP-malic enzyme, rather than the relative PEPCK expression in NADP-malic enzyme subtypes. This is supported by the relative differences between NADP-malic enzyme activity and alanine aminotransferase activity in the C₄ grasses studied. Figure 4.7 showed that in *Setaria viridis* the activity of NADP-malic enzyme was substantially lower than that in *Sorghum bicolor* and *Zea mays*. However, the activity of alanine aminotransferase was relatively higher in *Setaria viridis*, compared to the activity in *Sorghum bicolor* and *Zea mays*. This observation is supported by the relative expression of alanine aminotransferase in Figure 5.16 and corresponds to the higher activity of alanine aminotransferase in C₄ species with reduced NADP-malic enzyme activity (Andrews et al., 1971; Hatch, 1987). It is important to note however, that the rate of NADP-malic enzyme activity was substantially higher than that of alanine aminotransferase, and this observation is not a direct comparison between changes in specific activities.
Novel phosphorylation data presented in Chapter 3 showed that PEPC, aspartate and alanine aminotransferase and NADP-malic enzyme undergo light-dependent phosphorylation (Appendix A). Along with the in vitro activity assays, it appeared that there were changes in the dark and light in vitro activities of these enzymes. However, whether the phosphorylation sites identified serve a biological function remains unclear. These can be further investigated using phosphomimetic mutants expressed in Escherichia coli, similar to the approach used to study the effect of phosphorylation of PPDK (Chastain et al., 1997; Chastain et al., 2000). Using site-directed mutagenesis of Setaria viridis C₄-related proteins, such as PEPC and NADP-malic enzyme, the phosphorylation sites identified in this study can be further validated. Recombinant proteins can be purified and assayed in vitro and assessed whether the phosphorylations affect enzymatic activity. Furthermore, kinetic constants, like $V_{\text{max}}$ and $K_{\text{cat}}$ can be calculated and compared to the in vitro activities of the enzymes from crude leaf lysates. While this approach may prove useful, it might not be truly representative of in vivo phosphorylation. For instance, phosphorylations are dynamic and the degree of phosphorylation of a protein may change in response to intracellular stimuli or environmental cues, such as varying light intensities (Chen et al., 2014; Friso and van Wijk, 2015). While phosphomimetic mutants can show whether a certain modification can alter enzyme activity, the genetically introduced modifications are static and cannot dynamically respond to stimulants in vitro. A supplementary approach would be to selectively purify phosphorylated proteins from the crude leaf lysate. Anti-phosphoprotein antibodies are commonly used for phosphoprotein identification by immunoblotting or mass spectrometry, but up to now their applications for immunoprecipitating phosphorylated proteins and determining physiological enzymatic activity are limited. This is partly due to reactions conducted in non-physiological conditions, which are likely to lead to loss of enzyme activity and because of low binding efficiencies of commercially available anti-phosphoprotein antibodies, namely the differences between anti-phosphotyrosine and anti-phosphoserine or anti-phosphothreonine, which consequently enrich phosphorylated tyrosine residues with higher efficiency than phosphoserines and phosphothreonines (Grønborg et al., 2002). One method to overcome this might be to develop gel activity assays, whereby proteins are kept and assayed in their native conditions (Covian et al., 2012). In addition, by selectively purifying proteins of interest using antibodies, proteins may be as close to their physiological phosphorylation state and the effects of phosphorylation on protein
activity can be determined. Also, certain issues such as ensuring phosphomimetic proteins are structurally comparable to the candidate wildtype proteins can be avoided.

6.2 Communication between C₃ and C₄ cycles and the formation of PEP

The efficient and rapid movement of C₄ metabolites depends on the coordinate regulation of C₃ and C₄ pathways (Furbank and Leegood, 1984; Leegood and Furbank, 1984). However, one interaction that has not been fully understood is the interchange of 3-PGA and PEP using phosphoglycerate mutase and enolase. This pathway was first elucidated by Huber and Edwards (1975), suggesting that PEP could be formed via phosphoglycerate mutase and enolase without using energy (Figure 6.2). This is significant because the two main pathways that form PEP in C₄ photosynthesis, via the reactions catalysed by PEPC or PPDK, require energy. However, it is not known how rapidly the interconversion between 3-PGA to PEP occurs or whether it is regulated (Arrivault et al., 2016), but an early study estimated that the interconversion between 3-PGA and PEP occurred at one-third the maximal rate of the reactions in the PCR cycle (Monson and Moore, 1989).

![Diagram of C₃ and C₄ cycles](image)

**Fig. 6.2. Interconversion of 3-PGA and PEP via phosphoglycerate mutase and enolase in a NADP-malic enzyme system.** The interconversion of 3-PGA and PEP in C₄ plants may link C₃ and C₄ cycles, thereby generating PEP in a sequential, non-energy requiring mechanism. Adapted from Huber and Edwards (1975).
Light-dependent phosphorylation sites identified in phosphoglycerate mutase and enolase are the first indication that this pathway is regulated in C₄ plants. However, since the interconversion of 3-PGA and PEP is not energy-dependent, like PEPCK or PPDK, its regulation might not be as critical as that of other key C₄ enzymes (M. Stitt, personal communication). The interconversion between 3-PGA and triose phosphates is crucial in C₄ plants and the 3-PGA that is translocated from bundle sheath to mesophyll cells needs to be reduced to triose phosphates and not necessarily used to form PEP. The conversion of 3-PGA to triose phosphates requires three steps and energy (Huber and Edwards, 1975) (Figure 1.2). First, 3-PGA is converted to 1,3 bisphosphoglycerate via phosphoglycerate kinase, which requires activation by ATP. Then, 1,3 bisphosphoglycerate is reduced to glyceraldehyde-3-phosphate via glyceraldehyde-3-phosphate dehydrogenase using NADPH. Finally, glyceraldehyde-3-phosphate is converted to dihydroxyacetone phosphate (triose phosphate) via triose phosphate isomerase (Anderson, 1971; Huber and Edwards, 1975).

Alternatively, the conversion of 3-PGA to PEP via phosphoglycerate mutase and enolase does not require energy, and therefore the flux of carbon through this pathway may offer less resistance than the energy-dependent reduction of 3-PGA to triose phosphates. Therefore, these two pathways could potentially compete for 3-PGA. The conversion between 3-PGA to triose phosphates via the reactions catalysed by phosphoglycerate kinase, glyceraldehyde-3-phosphate dehydrogenase and triose phosphate isomerase in mesophyll cells is crucial in C₄ plants and alleviates the burden of 3-PGA dependent O₂ evolution from bundle sheath cells especially in NADP-malic enzyme subtypes, which lack the capacity to generate reducing equivalents due to reduced photosystem II activity in bundle sheath cells (Chapman et al., 1980; Leegood, 2013). Triose phosphates that are formed in mesophyll cells must be transported back to bundle sheath chloroplasts and are essential to regenerate RuBP in the PCR cycle (Hatch, 1987).

Formation of PEP through the conversion from 3-PGA might be a way of increasing C₄ metabolites by feeding in from the PCR cycle (Leegood and von Caemmerer, 1994) and may be advantageous in C₃–C₄ intermediate species or C₄ species such as Sorghum bicolor and Setaria viridis, which solely rely on one decarboxylation pathway (Gutierrez et al., 1974). Furthermore, metabolites could effectively move between C₃ and C₄ cycles, thus contributing to the larger pool of metabolites in mesophyll or bundle sheath cells (Arrivault et al., 2016). Communication between C₃ and C₄ pathways along with
mechanisms allowing to switch between using fluxes malate or aspartate, can increase the flexibility of the C₄ phenotype and maintain photosynthetic performance during intervals of low-light or reduced rates of carbon fixation (Huber and Edwards, 1975; Furbank and Leegood, 1984; Leegood and von Caemmerer, 1989; Bellasio and Griffiths, 2014; Stitt and Zhu, 2014). In addition to this, the interconversion of 3-PGA and PEP might have evolved when the C₄ pump was first established and PEPC activity was enhanced, but there might not have been sufficient PPDK activity in mesophyll cells to regenerate PEP since PPDK did not acquire C₄-specific function until much later (Sage, 2004). It has been suggested that reduced rates of PEP formation might have limited the progression of C₄-evolution in C₃–C₄ intermediates (Peisker, 1986) and certain C₃–C₄ Flaveria intermediates might have used the interconversion of 3-PGA and PEP to increase the amount of C₄ metabolites needed to operate the CO₂ pump (Monson and Moore, 1989). The importance of this interconversion in Flaveria C₃–C₄ intermediates is further supported by the higher activities of phosphoglycerate mutase and enolase when compared to the activities in Flaveria cronquistii (C₃), Flaveria trinervia (C₄) (Monson and Moore, 1989), Zea mays and Spinacea oleracea (C₃) (Furbank and Leegood, 1984). Although the non-energy requiring interconversion between 3-PGA and PEP could have been selected for during the early stages of C₄-evolution, it is speculated that its importance declined after PPDK acquired C₄-specific function and could regenerate PEP at higher rates (Monson and Moore, 1989).

Both phosphoglycerate mutase and enolase are present in appreciable amounts in mesophyll cells of C₄ plants (Ku and Edwards, 1975) and activity might be strictly dependent on illumination, which would correlate to the increase in the amount of PEP and 3-PGA available in the whole leaf (Furbank and Leegood, 1984; Leegood and von Caemmerer, 1989). While this pathway may not be the primary route for PEP formation in C₄ plants, its reversibility, together with requiring no energy, can be the only way that carbon is moved between C₃ and C₄ cycles. Such communication between C₃ and C₄ cycles might be necessary in C₄ plants and intermediate species, since they generate higher amounts of 3-PGA and triose phosphates during photosynthesis than C₃ plants (Badger et al., 1984; Leegood and von Caemmerer, 1994; Arrivault et al., 2016). It would be interesting to determine if there are any changes in the kinetic properties of phosphoglycerate mutase and enolase in response to the dark to light transition in both the forward and reverse reactions. It might also be important to understand how the
phosphorylation of phosphoglycerate mutase and enolase might affect enzymatic activity. In addition to this, it might be beneficial to consider this pathway when engineering C4 photosynthesis into C3 crops, since it might have served an important role during the early stages of C4 photosynthesis and may have confer some biochemical stability for the CO2 pump (Monson and Moore, 1989). Moreover, this pathway might maintain high photosynthetic efficiency during light fluctuations by supplying metabolites when the rate of CO2 assimilation declines (Furbank and Leegood, 1984; Leegood and von Caemmerer, 1989; Stitt and Zhu, 2014).

With regards to C3–C4 intermediate species, there is some concern as to how PEP is regenerated in C3–C4 intermediates and C4 variants of Alloteropsis semialata populations. Transcriptomic analyses have shown that the relative expression of PPDK in the C4 variant of Alloteropsis semialata is about 0.5-fold lower than the expression in Setaria italica (unpublished), raising concern whether there is significant PPDK activity in mesophyll cells. Comparatively, PEPCK expression in Alloteropsis semialata is approximately 122-fold higher than the expression in Setaria italica, whereas the relative expression of NADP-malic enzyme in Setaria italica is approximately 12-fold higher than in Alloteropsis semialata (unpublished). Furthermore, evidence suggests that there is no substantial formation of pyruvate in Alloteropsis semialata via NADP-malic enzyme, suggesting the Alloteropsis semialata might be a true PEPCK subtype, and therefore PEP cannot be efficiently regenerated via PPDK (P.A. Christin, personal communication). It was also found that C4 Alloteropsis semialata have high PEP phosphatase activity and this might serve as an alternative pathway for the formation of PEP (P.A. Christin, personal communication).

6.3 Conclusion

In summary, it was shown that NADP-malic enzyme, aspartate aminotransferase and alanine aminotransferase undergo light-dependent phosphorylation. This study also showed that PEPC and PPDK were phosphorylated at the previously described residues in Zea mays and Sorghum bicolor, but there are additional phosphorylation sites that have not been described in other C4 plants. It is possible that some of these phosphorylated amino acids affect enzymatic activity. Aside from the phosphorylation of key C4-related
proteins, the light-dependent phosphorylation of phosphoglycerate mutase and enolase was shown. This might be important for regulating the flux of carbon through C₃ and C₄ cycles and if there are phosphoregulatory controls, it might be an indication to the importance of this pathway and formation of C₄ metabolites.

In addition, this study showed that the properties of NADP-malic enzyme and aspartate aminotransferase are species-specific. It is apparent that the flexibility of NADP-malic enzyme and aspartate or alanine aminotransferase activity might be important for maintaining high photosynthetic performance. Certain NADP-malic enzyme plants, like *Zea mays*, may have the ability to switch between NADP-malic enzyme and PEPCK decarboxylation. This is shown by the differences in catalytic properties of NADP-malic enzyme and aspartate aminotransferase between *Zea mays* and *Sorghum bicolor*. This study further supports the fact that the classification of C₄ plants as distinct subtypes might not be truly representative of the interspecies differences that exist.

With regards to engineering the C₄ trait into existing C₃ crops, efforts should consider engineering multiple decarboxylase pathways, such that the metabolic robustness of the carbon concentrating mechanism of C₄ photosynthesis is maintained and responds to environmental changes such as fluctuating light intensities (Wang et al., 2014). Furthermore, there may not be a single species that provides the best blueprint for engineering C₄ photosynthesis, as species-dependent changes in NADP-malic enzyme activity have been shown to exist. In addition to this, post-translational regulatory mechanisms must be validated such that C₄-related proteins expressed in target C₃ species maintain the expected photosynthetic capacity. This is particularly important since NADP-malic enzyme, aspartate aminotransferase and alanine aminotransferase exist in both C₃ and C₄ plants, but these may be regulated differently (Hibberd and Quick, 2002).
References


Portis AR and Heldt HW (1976) Light-dependent changes of the Mg$^{2+}$ concentration in the stroma in relation to the Mg$^{2+}$ dependency of CO$_2$ fixation in intact chloroplasts. *Biochim Biophys Acta* **449**: 434–446.


Appendix A

Putative phosphorylation sites identified in darkened and illuminated leaves of *Setaria viridis*.

Phosphorylation sites identified in darkened and illuminated leaves of *Setaria viridis* by tandem mass spectrometry. Peptide spectra were searched against NCBI and UniProtKB plant protein sequences using MASCOT (Table A) and SEQUEST (Table B). Phosphorylations were determined using PhosphoRS. Phosphorylation site localisation probabilities (P > 0.75 is significant) are indicated in parentheses to the right of the modified amino acid. Phosphopeptide significance is measured against an E-value (MASCOT) and Xcorr (SEQUEST), where E-value < 0.05 or Xcorr ≥ 2.15, is significant. *Reoccurring phosphorylation sites; **phosphorylations identified by phosphopeptide enrichment. Time-points (TP): TP2, 7.5 h into the dark (shaded); TP3, 4 h into the light; TP4, 15.5 h into the light.

Table A. Proteins identified using MASCOT.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Coverage (%)</th>
<th>Phosphopeptide</th>
<th>Position in peptide</th>
<th>MH⁺ (Da)</th>
<th>Monoisotopic mass (Da)</th>
<th>E-value</th>
<th>TP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEPC</td>
<td>526117397</td>
<td>56.43</td>
<td>HHS(1)IDAQLR</td>
<td>Ser-3**</td>
<td>1156.53</td>
<td>1075.54</td>
<td>0.25786</td>
<td>TP3</td>
</tr>
<tr>
<td>PEPC</td>
<td>P04711</td>
<td>18.35</td>
<td>HHS(1)IDAQLR</td>
<td>Ser-3*</td>
<td>1156.53</td>
<td>1075.54</td>
<td>0.00432</td>
<td>TP3</td>
</tr>
<tr>
<td>PPDK</td>
<td>514750542</td>
<td>60.53</td>
<td>GGMT(1)SHAAVVAR</td>
<td>Thr-4**</td>
<td>1252.55</td>
<td>1155.57</td>
<td>0.00011</td>
<td>TP2</td>
</tr>
<tr>
<td>PPDK</td>
<td>Q6AVA8</td>
<td>23.97</td>
<td>GGMT(1)SHAAVVAR</td>
<td>Thr-4*</td>
<td>1252.55</td>
<td>1155.57</td>
<td>2.47E-06</td>
<td>TP2</td>
</tr>
<tr>
<td>PPDK</td>
<td>514750542</td>
<td>60.53</td>
<td>GGMT(0.98)S(0.02)SHAAVVAR</td>
<td>Thr-4*</td>
<td>1236.56</td>
<td>1155.57</td>
<td>0.00005</td>
<td>TP2</td>
</tr>
<tr>
<td>PPDK</td>
<td>Q6AVA8</td>
<td>23.97</td>
<td>GGMT(0.98)S(0.02)SHAAVVAR</td>
<td>Thr-4*</td>
<td>1236.56</td>
<td>1155.57</td>
<td>7.93E-07</td>
<td>TP2</td>
</tr>
<tr>
<td>PPDK</td>
<td>514750542</td>
<td>62.55</td>
<td>GGMT(0.97)S(0.02)SHAAVVAR</td>
<td>Thr-4*</td>
<td>1236.56</td>
<td>1155.57</td>
<td>0.00865</td>
<td>TP3</td>
</tr>
<tr>
<td>PPDK</td>
<td>P11155</td>
<td>20.49</td>
<td>GGMT(1)SHAAVVAR</td>
<td>Thr-4*</td>
<td>1236.56</td>
<td>1155.57</td>
<td>0.00009</td>
<td>TP3</td>
</tr>
<tr>
<td>Protein</td>
<td>Accession</td>
<td>Coverage (%)</td>
<td>Phosphopeptide</td>
<td>Position in sequence</td>
<td>MH* (Da)</td>
<td>Monoisotopic mass (Da)</td>
<td>Xcorr</td>
<td>TP</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>--------------</td>
<td>------------------------</td>
<td>----------------------</td>
<td>----------</td>
<td>------------------------</td>
<td>-------</td>
<td>----</td>
</tr>
<tr>
<td>PGM</td>
<td>K3XFX0</td>
<td>51.79</td>
<td>S(0.92)GGIQILT(0.69)S(0.69)HT(0.69)QPVPVAIGPGPLHGV</td>
<td>S520*; T527*; S528 or T530</td>
<td>3373.65</td>
<td>3132.75</td>
<td>2.27</td>
<td>TP2</td>
</tr>
<tr>
<td>PGM</td>
<td>K3XFX0</td>
<td>84.16</td>
<td>AHGTAVGLPSDDMNS(1)EVGHNALGAGR</td>
<td>S109**</td>
<td>2785.21</td>
<td>2704.22</td>
<td>4.39</td>
<td>TP3</td>
</tr>
<tr>
<td>PGM</td>
<td>K3XFX0</td>
<td>79.05</td>
<td>AGKPLLDDS(0.96)GGIQILT(0.88)S(0.57)HT(0.57)QPVPVAIG</td>
<td>S109*</td>
<td>2801.20</td>
<td>2704.22</td>
<td>3.63</td>
<td>TP4</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K3ZSX0</td>
<td>72.80</td>
<td>LLEAT(0.5)GIS(0.5)T(0.5)VPGS(0.5)GFQKDEVHFLRKLFDMMFLPSREVQVLVS(0.91)FHT(53.9)VS(0.53)K</td>
<td>T424 or T428; S427 or S432 S285; T288 or S290* S206*</td>
<td>2760.31</td>
<td>2599.37</td>
<td>2.08</td>
<td>TP2</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K3ZSX0</td>
<td>79.50</td>
<td>VAS(0.88)IALS(0.24)PNVPGQI FMGVNPPPKPSIDIS(0.64)YLYK LEAT(0.5)GIS(0.5)T(0.5)VPGS(0.5)GFQKDEVHFLRVLFDMMFLPSREVQVLVS(0.75)VS(0.75)K</td>
<td>S322 or S326; S348 T424 or T428; S427 or S432 T288; S290*</td>
<td>4149.03</td>
<td>3972.08</td>
<td>2.81</td>
<td>TP3</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K3ZSX0</td>
<td>71.34</td>
<td>VAS(0.99)IALS(0.98)PNVPGQI FMGVNPPPKPSIDISLKYFAATAEARS(0.99)KGITVR IIFTNVPHALGQKPLT(0.96)EPR</td>
<td>S322; S326 S206* T57</td>
<td>4149.04</td>
<td>3972.08</td>
<td>2.38</td>
<td>TP4</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K4A868</td>
<td>64.44</td>
<td>AT(0.07)GAYSH(0.46)HS(0.46)QG AT(0.09)GAYSHSQGIKLRLDAIAAGIAS(1)R</td>
<td>T167*; S171 or T167*; S189*</td>
<td>1667.80</td>
<td>1544.80</td>
<td>1.62</td>
<td>TP2</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K4A868</td>
<td>66.48</td>
<td>AT(0.09)GAYSHSQGIKLRLDAIAAGIAS SSDRFLFLFAMAAS(0.93)VAVE</td>
<td>T167*; S189*</td>
<td>2673.26</td>
<td>2470.29</td>
<td>3.13</td>
<td>TP3</td>
</tr>
<tr>
<td>Protein</td>
<td>Accession</td>
<td>Coverage (%)</td>
<td>Phosphopeptide</td>
<td>Position in sequence</td>
<td>MH* (Da)</td>
<td>Monoisotopic mass (Da)</td>
<td>Xcorr</td>
<td>TP</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>--------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>----------</td>
<td>------------------------</td>
<td>-------</td>
<td>----</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K4A868</td>
<td>60.00</td>
<td>AT(0.92)GAYSHSQGIKGLRDAIAAGIAS(1)R DGFPANADDIFIT(1)DGAS(1)P GVHLMQQILIR</td>
<td>T167*; S189*</td>
<td>2673.25</td>
<td>2470.29</td>
<td>3.03</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2703; S207</td>
<td>3306.45</td>
<td>3113.52</td>
<td>3.08</td>
</tr>
<tr>
<td>AlaAT</td>
<td>K4A8B1</td>
<td>12.83</td>
<td>AT(0.93)GAYSHSQGIKGLRDAIAAGIAS(1)R</td>
<td>T167*; S189*</td>
<td>2673.25</td>
<td>2470.29</td>
<td>3.40</td>
<td>TP3</td>
</tr>
<tr>
<td>AAT</td>
<td>K3XHJ0</td>
<td>45.22</td>
<td>IISMRQQLFDALKS(0.86)R</td>
<td>S391</td>
<td>1901.96</td>
<td>1804.99</td>
<td>2.24</td>
<td>TP2</td>
</tr>
<tr>
<td>AAT</td>
<td>K3YSB2</td>
<td>76.20</td>
<td>KHMPFFDVAYQGFAS(0.48)GS(0.48)LDEDAFSVRLFVKMAS(0.53)T(0.53)AAFPVS(0.81)S(0.77)PAASVAARSKVHGGKSQLKRLRPMYS(0.98)NPFPFHG</td>
<td>S275 or S277</td>
<td>3488.62</td>
<td>3407.65</td>
<td>2.58</td>
<td>TP2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S3 or T4; S10; S11</td>
<td>2827.24</td>
<td>2586.32</td>
<td>2.48</td>
</tr>
<tr>
<td>AAT</td>
<td>K3YSB2</td>
<td>74.45</td>
<td>RGMEVFVAQS(0.98)YSKKNGLY</td>
<td>S300</td>
<td>2514.18</td>
<td>2433.20</td>
<td>2.53</td>
<td>TP3</td>
</tr>
<tr>
<td>AAT</td>
<td>K3YSM6</td>
<td>53.27</td>
<td>IAGNLNMYLPFGGS(1)IK</td>
<td>S99</td>
<td>1887.87</td>
<td>1806.89</td>
<td>2.31</td>
<td>TP2</td>
</tr>
<tr>
<td>AAT</td>
<td>K3Z6L1</td>
<td>43.61</td>
<td>LGLYAAFPKHFS(0.99)ACGKIQS(0.75)QYTSQGASSISQK</td>
<td>S266; S273</td>
<td>3486.62</td>
<td>3226.63</td>
<td>2.71</td>
<td>TP2</td>
</tr>
<tr>
<td>AAT</td>
<td>K3Z6L1</td>
<td>51.57</td>
<td>IS(0.55)YAAAMS(0.55)T(0.97)LQT(0.91)AMEIKIKEANMALRPPV</td>
<td>S387 or T393; T394; T397</td>
<td>3486.62</td>
<td>3203.70</td>
<td>2.23</td>
<td>TP4</td>
</tr>
<tr>
<td>AAT</td>
<td>K4AG31</td>
<td>68.75</td>
<td>LVVEAAAT(1)IS(1)VFQK</td>
<td>T22; S23</td>
<td>1493.67</td>
<td>1290.71</td>
<td>1.70</td>
<td>TP4</td>
</tr>
<tr>
<td>Endoase</td>
<td>K3XWW9</td>
<td>83.18</td>
<td>QKLGANAILVS(1)LAVCKAGAMV</td>
<td>S121*</td>
<td>2393.27</td>
<td>2255.28</td>
<td>2.83</td>
<td>TP2</td>
</tr>
<tr>
<td>Endoase</td>
<td>K3XWW9</td>
<td>80.49</td>
<td>LGANAILVS(1)LAVCKAGAMV</td>
<td>S121*</td>
<td>2153.11</td>
<td>1999.12</td>
<td>1.52</td>
<td>TP3</td>
</tr>
<tr>
<td>Endoase</td>
<td>K3XWW9</td>
<td>55.16</td>
<td>LTDEIQQQQQIVGDDLVLVT(0.9)9NPT(1)RVAK</td>
<td>T334; T337</td>
<td>2982.47</td>
<td>2821.51</td>
<td>1.00</td>
<td>TP4</td>
</tr>
<tr>
<td>Endoase</td>
<td>K3Z681</td>
<td>75.06</td>
<td>QLVLVPVFAPNVINGSHAGNKAMQEFMILTPT(0.45)GAS(0.45)LGANAILVS(1)LAVCKAGAS(1)N</td>
<td>T179 or S183</td>
<td>4507.29</td>
<td>4442.29</td>
<td>4.71</td>
<td>TP2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S120*; S129*</td>
<td>2357.21</td>
<td>2097.22</td>
<td>3.55</td>
</tr>
<tr>
<td>Endoase</td>
<td>K3Z681</td>
<td>44.49</td>
<td>LGANAILVS(1)LAVCKAGAS(1)N</td>
<td>S120*; S129*</td>
<td>2357.20</td>
<td>2097.22</td>
<td>3.25</td>
<td>TP3</td>
</tr>
<tr>
<td>Endoase</td>
<td>K3Z681</td>
<td>66.97</td>
<td>LGANAILVS(0.99)LAVCKAGA</td>
<td>S120*</td>
<td>2235.23</td>
<td>2097.22</td>
<td>2.08</td>
<td>TP3</td>
</tr>
</tbody>
</table>
Accession

K3XFH6

K3XFH6

K3XFW4

K3XFW4

K3XFW4

K3XG11

K3ZRI5

K3ZRI5

K3XV32

Protein

NADP-ME

NADP-ME

NADP-ME

NADP-ME

NADP-ME

NADP-ME

NADP-ME

NADP-ME

PEPC

88.69

42.86

37.40

78.36

69.83

69.83

38.14

90.77

Coverage
(%)
88.26

1301.72

S180*
T147 or S152
S752

RLVT(0.5)ELGKS(0.5)K
LNIGS(1)RPAKR

256

2742.46
2484.16

S435*; S442*
S702**

1191.64

1210.66

3348.68

1205.62

S11
S25; T33

2912.55

2925.36

T302; T313
S435*; S442*

1748.91

1638.86

S362 or S363
S357

2385.20

3448.77

2711.37

1110.65

1129.67

1220.74

2581.50
2387.17

3130.70

1082.62

2709.59

2722.41

1667.92

1557.88

2224.26

3325.78

2550.43

1699.99

2567.39

2802.30
1982.91

1814.03
4476.33

Monoisotopic
mass (Da)
1557.88
4476.33

1895.01
4816.29

MH+
(Da)
1638.86
4816.29

S530; T548

S354

S541; S547

S373*, S379*;
S380*

Position in
sequence
S429*
T263*; T282*;
S271 or T277
S429*
T263*; T282*;
S271 or T277
S360; S369

RLHLLLS(1)QHPREYVT(1)AEC
HRPVVLHK
VKPDVILGLS(1)AVGGLFS(1)K
FTAATLEHGMHPPVS(1)PKPEW
R
S(1)LLQKHARIR

AGLLGAVR
KVKPDVILGLS(1)AVGGLFS(1)
KEVLEALK
RS(100.0)LAPDRLR

IVVAGAGSAGIGVVNAAS(0.99)
R
KT(1)YRMFNDDVQGT(1)AGVAI

KIS(1)AHIAAAVAAKAYELGLA
T(1)R
VWLVDSKGLIVS
(0.5)S(0.5)R

GIAELIALEIS(0.9)R

GLIYPPFS(1)NIRKIS(1)AHIA
ANVAA
LLGGTLADHTFLFLGAGEAGT

VWLVDSKGLIVSS(0.99)R
LALYT(0.89)ALGGVRPS(0.6)AC
LPIT(0.6)IDVGT(0.89)NNEELL
QKVWLVDSKGLIVSS
(0.67)R
NDEFYIGLRQK
LALYT(0.71)ALGGVRPS(0.84)A
CLPIT(0.71)IDVGT(0.71)NNEE
SHLVFNDDIQGTAS
LLNDEFYIGLRQK (0.96)VVLA
GLLAS(0.96)LK
IWLVDS(1)KGLIVS(1)S(1)RK

Phosphopeptide

2.17

1.06

0.71

2.80
2.57

2.68

1.38

1.76

1.40

1.59

1.96

2.39

3.60

3.20

0.85

1.92

TP2

TP2

TP2

TP4
TP2

TP4

TP3

TP3

TP3

TP3

TP2

TP2

TP4

TP4

TP2

TP4

TP4
TP4

TP3

3.05
2.37
3.53

TP3

TP

2.07

Xcorr


<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Coverage (%)</th>
<th>Phosphopeptide</th>
<th>Position in sequence</th>
<th>MH* (Da)</th>
<th>Monoisotopic mass (Da)</th>
<th>Xcorr</th>
<th>TP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEPC</td>
<td>K3XV32</td>
<td>92.22</td>
<td>HHS(1)IDAQLR</td>
<td>S11**</td>
<td>1136.52</td>
<td>1075.54</td>
<td>1.28</td>
<td>TP3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FTAATLEHGMHFPVS(1)PKPEW</td>
<td>S702**</td>
<td>2468.16</td>
<td>2387.17</td>
<td>2.16</td>
<td>TP3</td>
</tr>
<tr>
<td>PEPC</td>
<td>K3XV32</td>
<td>94.81</td>
<td>HHS(1)IDAQLR</td>
<td>S11**</td>
<td>1136.83</td>
<td>1075.54</td>
<td>2.24</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LLA PGKVS(1)EDDK</td>
<td>S25**</td>
<td>1393.66</td>
<td>1270.67</td>
<td>2.31</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RS(1)LLQKHAR</td>
<td>S180*</td>
<td>1230.65</td>
<td>1107.65</td>
<td>2.08</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VVLGYRDKLYS(1)T(1)R</td>
<td>S378; T379</td>
<td>1828.88</td>
<td>1667.93</td>
<td>1.94</td>
<td>TP4</td>
</tr>
<tr>
<td>PEPC-2</td>
<td>K3YPN6</td>
<td>43.52</td>
<td>GGGPTHLA1LSQPDT(0.92)IHH</td>
<td>T658</td>
<td>2346.16</td>
<td>2223.17</td>
<td>2.41</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LLA PGKVS(1)EDDK</td>
<td>S26**</td>
<td>1393.66</td>
<td>1270.67</td>
<td>2.31</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VT(1)LDLLEMFVAKGDPGIAAV</td>
<td>S820</td>
<td>2561.26</td>
<td>2464.28</td>
<td>1.97</td>
<td>TP4</td>
</tr>
<tr>
<td>PPDK</td>
<td>K3Z3Q6</td>
<td>46.03</td>
<td>GGMT(0.98)SHAAVVAR</td>
<td>T462**</td>
<td>1236.56</td>
<td>1155.57</td>
<td>3.33</td>
<td>TP2</td>
</tr>
<tr>
<td></td>
<td>K3Z3Q6</td>
<td>87.07</td>
<td>GGMT(0.98)SHAAVVAR</td>
<td>T462**</td>
<td>1236.55</td>
<td>1155.57</td>
<td>3.31</td>
<td>TP2</td>
</tr>
<tr>
<td>PPDK</td>
<td>K3Z3Q6</td>
<td>58.96</td>
<td>GGMT(0.33)GAAVIS(0.46)MPGMDT(0.6)VNLGLNDEVAAGGLGAKS(0.33)GAVGAVGMPGMDT(0.6)VNLGLNDEVAAGGLGAKS</td>
<td>T462**</td>
<td>1236.55</td>
<td>1155.57</td>
<td>4.24</td>
<td>TP3</td>
</tr>
<tr>
<td></td>
<td>K3Z3Q6</td>
<td>43.88</td>
<td>GGMT(0.98)SHAAVVAR</td>
<td>T462**</td>
<td>1236.55</td>
<td>1155.57</td>
<td>3.08</td>
<td>TP3</td>
</tr>
<tr>
<td>PPDK</td>
<td>K3Z3Q6</td>
<td>188.55</td>
<td>GGMT(0.99)SHAAVVAR</td>
<td>T557</td>
<td>2005.99</td>
<td>1925.01</td>
<td>1.96</td>
<td>TP3</td>
</tr>
<tr>
<td></td>
<td>K3Z3Q6</td>
<td>91.16</td>
<td>GGMT(0.99)SHAAVVAR</td>
<td>T462**</td>
<td>1236.56</td>
<td>1155.57</td>
<td>2.71</td>
<td>TP3</td>
</tr>
<tr>
<td></td>
<td>K3Z3Q6</td>
<td></td>
<td>GGMT(0.99)SHAAVVAR</td>
<td>T462**</td>
<td>1236.57</td>
<td>1155.57</td>
<td>2.99</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td>K3Z3Q6</td>
<td></td>
<td>AIFEAAIAMT(0.88)NQGVQVFP</td>
<td>T462**</td>
<td>1236.56</td>
<td>1155.57</td>
<td>3.47</td>
<td>TP4</td>
</tr>
<tr>
<td></td>
<td>K3Z3Q6</td>
<td></td>
<td>GGMT(0.99)SHAAVVAR</td>
<td>T753**</td>
<td>2416.26</td>
<td>2335.29</td>
<td>3.19</td>
<td>TP4</td>
</tr>
</tbody>
</table>
Appendix B

Measuring chlorophyll content.

To normalise protein samples, chlorophyll content from harvested leaf sections was determined. Chlorophyll assays were performed following Arnon (1949) and Leegood (1993). 1.28 cm² leaf sections were placed in a test tube and submerged in 3 mL 80% (v/v) ethanol, and incubated in the dark at 70 °C for 35 min. Extra time was added to allow complete leaf bleaching. Chlorophyll content was determined spectrophotometrically at A₆₅₂. The instrument was blanked at A₇₅₀ to correct for sample turbidity (Ritchie, 2006). Chlorophyll content was determined in 80% ethanol was calculated using equation A,

\[
\text{Chl } a = A_{652} \times 27.8 \quad (A)
\]

To extract in 80% acetone, leaf tissue was ground to a fine powder in a liquid nitrogen chilled mortar and homogenised in 1 mL 80% chilled acetone. Plant lysate was transferred to an Eppendorf and centrifuged at high speed at 4 °C. Supernatant was collected into a light impermeable glass vial and placed on ice. The remaining pellet was further washed with acetone until completely bleached; the total volume per chlorophyll extract did not exceed 3.5 mL. Glass vials were removed from ice and let to equilibrate at room temperature before measuring chlorophyll. Chlorophyll a in 80% acetone was quantified spectrophotometrically at A₆₅₂ and content in mg L⁻¹ was determined using equation B,

\[
\text{Chl } a \text{ mg L}^{-1} = \frac{A_{652} \times 1000}{34.5} \quad (B)
\]

NB. Protein extracts were normalised against the lowest whole protein concentration.