Abstract

Cell fractionation and immunohistochemical studies in the last 40 years have revealed the extensive compartmentation of plant metabolism. In recent years, new protein mass spectrometry and fluorescent-protein tagging technologies have accelerated the flow of information, especially for Arabidopsis thaliana, but the intracellular locations of the majority of proteins in the plant proteome are still not known. Prediction programs that search for targeting information within protein sequences can be applied to whole proteomes, but predictions from different programs often do not agree with each other or, indeed, with experimentally determined results. The compartmentation of most pathways of primary metabolism is generally covered in plant physiology textbooks, so the focus here is mainly on newly discovered metabolic pathways in plants or pathways that have recently been revised. Ultimately, all of the pathways of plant metabolism are interconnected, and a major challenge facing plant biochemists is to understand the regulation and control of metabolic networks. One of the best-characterized networks links sucrose synthesis in the cytosol with photosynthetic CO₂ fixation and starch synthesis in the chloroplasts. One of the key features of this network is how the transport of pathway intermediates and signal metabolites across the chloroplast envelope conveys information between the two compartments, influencing the regulation of several enzymes to coordinate fluxes through the different pathways. It is widely accepted that chloroplasts and mitochondria originated from prokaryotic endosymbionts, and that new transporters and regulatory networks evolved to integrate metabolism in these organelles with the rest of the cell. Curiously, the present-day locations of many metabolic pathways within the cell often do not reflect their evolutionary origin, and there is evidence of extensive shuffling of enzymes and whole pathways between compartments during the evolution of plants.

Key words: C₄ photosynthesis, coenzyme, compartmentation, gluconeogenesis, glycolysis, isoprenoid, starch, sucrose, vitamin.

Introduction

One of the characteristic features of eukaryotic cells is the compartmentation of metabolism and other cellular functions in different parts of the cell. Despite this division of labour, metabolism in every compartment depends, to some extent, on other parts of the cell for supplies of energy (ATP) or metabolic precursors, and all compartments rely on the nucleus and cytosolic ribosomes to provide most, if not all, of their enzymes and other proteins. Therefore, to understand how a eukaryotic cell works, it is necessary to know not only how metabolism and other processes are compartmented within the cell, but also how they are all linked together and controlled. Understanding how plant cells work is particularly challenging because of the presence of additional compartments—plastids, cell walls, and vacuoles (Fig. 1)—that are not found in all eukaryotic cells, and by their greater diversity of metabolic pathways. Not surprisingly, there are many unique aspects of metabolic compartmentation in plant cells (ap Rees, 1987).

In this review, the techniques, both old and new, that are used to investigate compartmentation of plant metabolism are described, and several examples of compartmented pathways are discussed. Information about the intracellular compartmentation of the major pathways of primary metabolism in plants can be found in any reputable plant physiology textbook. Therefore, the focus here is on some newly discovered pathways in plant metabolism, with a few other examples of metabolic pathways that have recently been revised. Once the location of a pathway has been established, the next challenge is to understand how...
different parts of the pathway are co-ordinately regulated. We are only just beginning to understand such regulatory networks and a brief overview of one of the few well-characterized examples—the control of photosynthetic carbon metabolism in leaves—will be given. Reconstruction of the evolutionary history of this pathway has revealed some surprising insights into the way plant metabolism became so highly compartmented.

Investigating compartmentation in plant metabolism

Cell fractionation

The traditional way to investigate metabolic compartmentation is by cell fractionation, which is usually a targeted approach that asks the question: ‘In which compartment is enzyme X located?’ To answer this question, cells or protoplasts are gently disrupted in an osmotically balanced medium, and the cell lysate is then fractionated by differential centrifugation and/or centrifugation on sucrose or Percoll density gradients. Each fraction is then assayed for activities of enzyme X and marker enzymes for each compartment, with co-localization across different fractions indicating co-localization within the cell. For enzymes that co-localize with organelle marker enzymes, further experiments on highly purified organelle preparations are carried out to see if the enzyme is really located within the organelle, or just stuck to the outside. These can include latency measurements, in which activity is compared in intact and lysed preparations of the organelle, and protease protection assays, where resistance to proteolytic degradation until the organelle is broken indicates that the enzyme is located inside. Such assays are also used to determine sub-organelar compartmentation in multi-compartment organelles such as mitochondria and chloroplasts. Cell fractionation studies rely on accurate measurements of the target and marker enzymes, and careful accounting of all the fractions is needed to check that all of the activity in the initial cell lysate is recovered after fractionation. This is particularly important where an enzyme is found in more than one compartment and the isoforms from different compartments show differential stability. A limitation of classical cell fractionation techniques is that a cytosolic location for an enzyme can only be inferred in a negative way, i.e. if it does not co-localize with any known organelle or membrane compartment, and cytosol-enriched fractions are almost invariably contaminated by proteins from broken organelles.

Cell fractionation studies have been carried out mostly on a few experimentally tractable systems, such as spinach leaves, castor bean seedlings, and cell suspension cultures, and information obtained from these model systems is generally assumed to be valid for comparable tissues from other species. However, such extrapolation can sometimes lead us astray, with the enzyme ADP-glucose pyrophosphorylase providing a good cautionary example. This enzyme synthesizes the substrate (ADP-glucose) for starch synthesis, which occurs only in plastids. In leaves and most non-photosynthetic organs, ADP-glucose pyrophosphorylase is similarly restricted to plastids—it has even been used as a plastidial marker enzyme in some studies. However, it has recently been shown that in the developing endosperm of cereals (e.g. barley, wheat, maize, and rice) and wild grass species the majority of the ADP-glucose pyrophosphorylase activity is in fact cytosolic, and most of the ADP-glucose needed for starch synthesis is produced in the cytosol and imported by the amyloplasts via a specific transporter in the amyloplast envelope (Denyer et al., 1996; Thorbjørnsen et al., 1996; Beckles et al., 2001; Sikka et al., 2001; Tetlow et al., 2003). The generally accepted plastidial location for ADP-glucose synthesis in other species and organs has recently been challenged by the proposal that ADP-glucose is synthesized in the cytosol, not by ADP-glucose pyrophosphorylase, but by sucrose synthase (Baroja-Fernandez et al., 2004). However, it is difficult to reconcile this proposal with the large body of biochemical and genetic evidence that plastidial ADP-glucose pyrophosphorylase does indeed supply the vast majority of ADP-glucose for starch synthesis in most species and organs (Neuhaus et al., 2005). Therefore, although sucrose synthase can undoubtedly synthesize ADP-glucose in vitro, the weight of evidence at present indicates that it makes little, if any, direct contribution to ADP-glucose synthesis in vivo.
In addition to locating enzymes, cell fractionation can also be used to investigate the intracellular compartmentation of metabolites. However, this is technically more challenging than enzyme localization because the turnover times of many metabolites are much shorter than the time needed for the fractionation. One way to overcome this problem is to remove water from the tissue to arrest metabolic activity, and carry out the fractionation under non-aqueous conditions. In a typical experiment, the tissue is first rapidly frozen in liquid nitrogen, lyophilized to remove all water, homogenized, and then centrifuged on gradients of organic solvents, such as n-heptane and carbon tetrachloride (Gerhardt and Heldt, 1987). The gradient is divided into fractions, which are then assayed for marker enzymes and the metabolites of interest. The separation of organelles by non-aqueous fractionation is much less effective than with aqueous systems, but it is still possible to partially resolve the cytosolic, plastidial, and vacuolar compartments of leaf tissue (Gerhardt and Heldt, 1987; Stitt et al., 1989) and potato tubers (Farré et al., 2001).

One important result from non-aqueous fractionation experiments was the finding of a significant pool of inorganic pyrophosphate (Pi) in the cytosol (Weiner et al., 1987). This pool of Pi allows UDP-glucose pyrophosphorylase to catabolize the UDP-glucose produced by cleavage of sucrose via sucrose synthase, and thus conserve much of the energy in the glycosidic bond of sucrose (Gerhardt et al., 1984). Cytosolic Pi can also contribute to energization of the tonoplast via the H+-pumping pyrophosphatase (Rea and Poole, 1993), and Pi can be transported into chloroplasts to replenish Pi pools in the stroma (Lunn and Douce, 1993). Non-aqueous fractionation of potato tubers was also instrumental in revealing the redox regulation of ADP-glucose pyrophosphorylase and starch synthesis (Tiessen et al., 2002).

**Immunohistochemistry**

Another technique that has been used widely to investigate compartmentation is immunohistochemistry. Specific antibodies are used to immuno-decorate tissue sections, and then visualized by light or electron microscopy after labelling of the antibodies with fluorescent, heavy-atom, or enzymatic markers (Tobin and Yamaya, 2001). Immunohistochemical techniques can be used for proteins that do not have enzymatic activity (e.g. transporters) or proteins of unknown function, and have also been used to locate non-protein targets such as cell wall polysaccharides and amino acids (Walker et al., 2001). Immunohistochemical methods also have the advantage that cytosolic proteins can be distinguished from those that are associated with intracellular structures that are too fragile or too low in abundance to isolate by simple cell fractionation techniques, such as the endoplasmic reticulum and Golgi bodies. However, a major limitation of this technique is the need for highly specific antibodies for each target of interest, because even weak cross-reaction with a particularly abundant proteins (e.g. Rubisco in leaves) can lead to mistaken conclusions. Furthermore, antibodies do not usually distinguish between active and inactive forms of their target protein, such as unprocessed pre-proteins or partially degraded proteins. Therefore, detection of an enzyme protein in a particular compartment does not necessarily indicate the presence of activity in that location.

**Proteome analysis by mass spectroscopy**

Traditional cell fractionation studies and immunohistochemistry are low-throughput, targeted approaches for investigating metabolic compartmentation. In the last few years, new technologies for identifying proteins by mass spectrometry have been combined with classical cell fractionation techniques to catalogue the proteomes of various plant cell compartments, particularly in the model species Arabidopsis thaliana. Such proteomic surveys have not only confirmed the locations of proteins identified in previous studies, but more importantly revealed the locations of numerous other proteins, including many whose functions are still unknown. Indeed, this has been one of the driving forces for such studies, as knowing the intracellular location of an ‘unknown’ protein can be a very useful clue to identifying its function.

To date, most proteomic studies in plants have focused on Arabidopsis thaliana, because the availability of the genome sequence makes protein identification relatively straightforward. A list of proteomic studies that have been undertaken in this species is given in Table 1. Proteomic analyses have also been carried out on the inner and outer envelope membranes of spinach chloroplasts (Ferro et al., 2002), bundle sheath and mesophyll cell chloroplasts from maize leaves (Majera et al., 2005), and rice etioplasts (Zychlinski et al., 2005). The results from such proteomic studies of plant organelles are available from several databases: PPDB, http://ppdb.tc.cornell.edu (van Wijk, 2004); SUBA, http://www.suba.bcs.uwa.edu.au (Heazlewood et al., 2005); plprot, http://www.plprot.ethz.ch (Kleffmann et al., 2006). Analysis of the chloroplast proteome confirmed that many of the enzymes involved in biosynthesis of fatty acids, lipids, amino acids, nucleotides, hormones, alkaloids, and isoprenoids are found in these organelles, in addition to the Calvin cycle enzymes and proteins belonging to the light-harvesting apparatus and photosynthetic electron transport chain (van Wijk, 2004).

In mitochondria, about 20% of the proteins identified have completely unknown functions (Heazlewood et al., 2004; Millar et al., 2005). Although many of the others have been assigned to general protein classes such as protein kinases and phosphatases, their specific functions are also unknown. This large number of poorly characterized proteins suggests that there are many metabolic pathways and regulatory processes still waiting to be discovered in these organelles.
Table 1. Proteomic studies of organelles and sub-organelar compartments in Arabidopsis thaliana

<table>
<thead>
<tr>
<th>Organelle</th>
<th>Tissue</th>
<th>Proteins identified</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chloroplasts</td>
<td>Rosette leaves</td>
<td>690</td>
<td>Kleffmann et al., 2004</td>
</tr>
<tr>
<td>Chloroplasts</td>
<td>Rosette leaves</td>
<td>426</td>
<td>Baginsky et al., 2005</td>
</tr>
<tr>
<td>Chloroplasts (thylakoid lumen)</td>
<td>Rosette leaves</td>
<td>81</td>
<td>Peltier et al., 2002</td>
</tr>
<tr>
<td>Chloroplasts (thylakoid lumen)</td>
<td>Rosette leaves</td>
<td>36</td>
<td>Schubert et al., 2002</td>
</tr>
<tr>
<td>Chloroplasts (thylakoid membrane)</td>
<td>Rosette leaves</td>
<td>154</td>
<td>Friso et al., 2004</td>
</tr>
<tr>
<td>Chloroplasts (envelope)</td>
<td>Rosette leaves</td>
<td>106</td>
<td>Ferro et al., 2003</td>
</tr>
<tr>
<td>Chloroplasts (envelope)</td>
<td>Rosette leaves</td>
<td>392</td>
<td>Froehlich et al., 2003</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>Stem, leaves, and cell culture</td>
<td>52</td>
<td>Kruft et al., 2001</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>Cell culture</td>
<td>91</td>
<td>Millar et al., 2001</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>Cell culture</td>
<td>114</td>
<td>Brugi`ere et al., 2004</td>
</tr>
<tr>
<td>Mitochondria (protein complexes)</td>
<td>Cell culture</td>
<td>14</td>
<td>Werhahn and Braun, 2002</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>Cell culture</td>
<td>416</td>
<td>Hazlewood et al., 2004</td>
</tr>
<tr>
<td>Peroxisomes</td>
<td>Greening cotyledons</td>
<td>29</td>
<td>Fukao et al., 2002</td>
</tr>
<tr>
<td>Peroxisomes</td>
<td>Etiolated cotyledons</td>
<td>19</td>
<td>Fukao et al., 2003</td>
</tr>
<tr>
<td>Vacuoles</td>
<td>Rosette leaves</td>
<td>402</td>
<td>Carter et al., 2004</td>
</tr>
<tr>
<td>Vacuoles (tonoplast)</td>
<td>Cell culture</td>
<td>163</td>
<td>Shimakawa et al., 2004</td>
</tr>
<tr>
<td>Vacuoles (tonoplast)</td>
<td>Cell culture</td>
<td>70</td>
<td>Szponarski et al., 2004</td>
</tr>
<tr>
<td>Nucleus</td>
<td>Rosette leaves</td>
<td>158</td>
<td>Bae et al., 2003</td>
</tr>
<tr>
<td>Nucleus</td>
<td>Cell culture</td>
<td>36</td>
<td>Ca´tilkowski et al., 2003</td>
</tr>
<tr>
<td>Nucleus (nucleolus)</td>
<td>Cell culture</td>
<td>217</td>
<td>Pendle et al., 2005</td>
</tr>
<tr>
<td>Plasmalemma</td>
<td>Cell culture</td>
<td>97</td>
<td>Marmagne et al., 2004</td>
</tr>
</tbody>
</table>

Proteomic studies have greatly increased our knowledge about compartmentation of enzymes and other proteins in plant cells, but current technologies do have some limitations. Whatever the approach used to identify proteins, the reliability of the results is absolutely dependent on the purity of the original organelle preparation, and the level of contamination by proteins from other compartments often sets the limit of sensitivity. Mass spectrometric techniques can also be limited by their dynamic range, with high abundance proteins swamping the analysis with redundant information, making it more difficult to find peptide signals from rare proteins. As with immunohistochemical techniques, mass spectroscopic methods do not distinguish between active and inactive proteins, so metabolic activities can only be inferred from such data. Analysis of very hydrophobic proteins is still technically challenging, and the difficulties of preparing highly purified membrane preparations have held back proteomic analyses of the plasmalemma, endoplasmic reticulum, and Golgi body. Another limitation is that proteins can still only be assigned to the cytosol in a negative way, by not being present in other compartments. New technologies that might offer solutions to some of these problems are iso- tope tagging (Dunkley et al., 2004) and laser micro-dissection and capture (Schad et al., 2005).

Fluorescent protein tagging and other in vivo imaging techniques

Plant tissues generally contain more than one cell type. For example, it has been estimated that mesophyll, epidermal, and vascular cells occupy 58%, 3%, and 1%, respectively, of the volume of spinach leaves (Winter et al., 1994), and 51%, 27%, and 4% of the volume of wheat leaves (Bowsher and Tobin, 2001), with most of the remainder being air spaces. As a result of this heterogeneity, any organelle preparation is likely to come from a mixture of cell types; for example, a mitochondrial preparation from leaves could include organelles from both photosynthetic and non-photosynthetic cells. These might have very different metabolic functions and protein compositions; for example, the photorespiratory glycine decarboxylase complex would be a major component only in the mitochondria from photosynthetic cells (Tobin et al., 1989). In the worst case scenario, preferential extraction of organelles from a minor cell type could give a very misleading impression of the metabolism in the tissue as a whole, so results obtained from cell fractionation studies on heterogeneous plant tissues must be interpreted with caution. Attempts to avoid this issue by using homogeneous cell suspension cultures still face the question of how to relate the data to whole plants.

An alternative approach to tackle this problem is to express fluorescently tagged versions of the protein of interest in the plant. These can be visualized by laser confocal microscopy in vivo to show the intracellular location of the protein, and also its distribution between different cell types and organs if the native promoter and other regulatory elements are used to drive expression of the tagged protein. In the beginning, this technique was used to locate just one or two proteins of special interest, but several larger-scale investigations have now been carried out (Cutler et al., 2000; Tian et al., 2004; Koroleva et al., 2005). These were essentially pilot studies for high-throughput approaches, using transient or stable expression systems to test different strategies for making the fusion protein constructs, and several different fluorescent proteins (e.g. GFP, YFP, and CFP). Some studies have used constitutive and/or strong promoters to drive expression of the fluorescently tagged proteins, but some authors have
argued that it is preferable to use native promoters and regulatory elements, because ectopically expressed proteins may not be properly targeted. For example, a protein that is normally expressed only in photosynthetic cells and located in the chloroplasts could be effectively trapped if expressed in the cytosol of non-photosynthetic cells with few or no plastids. Similarly, it is argued that expression of a protein at an abnormally high level could also lead to artefactual compartmentation by overloading intracellular trafficking mechanisms, or disrupting protein complexes whose integrity is needed for proper targeting. Nevertheless, at least one study (Tian et al., 2004) has found that addition of CaMV 35S enhancer sequences led to higher expression of the tagged proteins and stronger fluorescence signals, but did not change the intracelluar location of the proteins.

A very important consideration in such studies is where to attach the fluorescent tag to the target protein. Clearly the N-terminus of the protein should be avoided, as plastidial and mitochondrial transit peptides, and endomembrane signal peptides, are all located in this region. However, the C-terminus can also contain important targeting information, such as endoplasmic reticulum retention signals. As a default strategy, Tian et al. (2004) inserted the YFP tag about 10 amino acids upstream of the C-terminus of the protein, flanked by six- or nine-amino acid Gly- or Ala-rich peptides, with the aim of maximizing the integrity of the target protein and minimizing the chance of disrupting targeting regions. Using this strategy, every single one of the peroxisomonal, tonoplast, plasmalemma, cell wall, plasmodesmatal, cytoskeletal, nuclear, plastidial, and cytosolic proteins tested was located in the expected compartment, demonstrating that it would be feasible to use this approach to locate proteins in most compartments. For proteins of particular interest, a useful test for the reliability of the results is to check whether the fluorescently tagged fusion protein can complement a loss-of-function mutant. A good example is the YFP-tagged MEX1 maltose transporter, which was found to be targeted to the chloroplast envelope and was able to restore the wild-type phenotype to the mex1 mutant (Niittyla et al., 2004).

There are several other imaging techniques that can be used to study metabolism in vivo, although these are used more to detect metabolites than proteins. Nuclear magnetic resonance (NMR) spectroscopy has been used to determine the intracellular distribution of inorganic orthophosphate, and can detect changes in the intracellular pH and energy charge of living cells (Gout et al., 1992). NMR has also been used to measure rates of sucrose or water transport in phloem and xylem vessels, and for network flux analysis (Köckenberger, 2001; Ratcliffe and Shachar-Hill, 2006). A major disadvantage of in vivo NMR is its rather poor sensitivity, as it is only able to detect metabolites that are present in millimolar concentrations, and this has so far limited its usefulness for studies of plant metabolism. A more sensitive technique for in vivo imaging is positron emission tomography (PET), which detects short-lived isotopes ($^{11}$C or $^{15}$F), that have been used to label molecules of interest, for example, sugars and polyphenols (Gester et al., 2005). Unfortunately, the necessary equipment for PET is very expensive, and so far this technique has found little application outside the medical field.

Computational approaches

Proteomic and fluorescent protein-labelling studies have revealed the locations of over 4000 proteins in Arabidopsis thaliana, but the whole proteome of this species is estimated to contain another 12 000–24 000 proteins whose intracellular locations are unknown (Heazlewood et al., 2004). To tackle this huge problem, various computational approaches have been developed to predict intracellular location from the primary protein sequence. Some widely used prediction programs are: TargetP, http://www.cbs.dtu.dk/services/TargetP/ (Emanuelsson et al., 2000); Predotar, http://www.inra.fr/predotar/ (Small et al., 2004); iPSPORT, http://hc.ims.u-tokyo.ac.jp/iPSORT/ (Bannai et al., 2002); SubLoc, http://www.bioinfo.tsinghua.edu.cn/SubLoc/ (Hua and Sun, 2001). Most of these programs aim to identify N-terminal transit or signal peptide sequences that target proteins to plastids, mitochondria, or the endomembrane system (endoplasmic reticulum, plasmalemma, vacuoles) and apoplast. To assess the specificity and sensitivity of various prediction programs, several comparisons of predicted and experimentally determined sets of organelle proteins have been made. In general, the programs seem to be more successful at predicting plastidial proteins than mitochondrial or endomembrane/apoplastic proteins, although even the most favourable assessments found that at least 15% of plastidial proteins were mis-identified (Kleffmann et al., 2006). Combining results from several prediction programs was found to improve specificity, i.e. there were fewer false-positive predictions, but there was a trade-off in sensitivity, with a greater number of chloroplast proteins being wrongly rejected as non-plastidial (Richly and Leister, 2004). A similar finding was made for the mitochondrial prediction programs, which individually fail to recognize more than half of the proteins known to be located in the mitochondria (Heazlewood et al., 2004).

There are several reasons why prediction programs can give both false-positive and false-negative results. One trivial explanation is that the N-terminal sequence of the protein has been incorrectly identified. The primary sequences of transit and signal peptides are often poorly conserved between species, and the open reading frame (ORF) prediction programs used for genome annotation can misidentify a conserved downstream region as the start of the ORF if no full-length cDNA sequence is available. False negatives can also arise if the protein in question is imported into an organelle by an unconventional route (Miras et al., 2002; Villarejo et al., 2005), although this is probably quite rare. Perhaps one of the main reasons for the
rather mixed success of the first generation of prediction programs is that they were trained on relatively small data-sets, usually fewer than 100 proteins of known location. Presumably, the availability of much larger training sets will eventually allow the reliability of transit and signal peptide prediction programs to be improved, but at present they can only give a rough guide to a protein’s intracellular location. Post-translational modifications, such as prenylation or addition of glycosyl phosphatidylinositol, can also play a major role in protein targeting, especially to membranes (e.g. plasmalemma and Golgi bodies). Prediction programs have been developed to identify which proteins have potential sites for such modifications, and are therefore likely to be targeted to these compartments (Borner et al., 2002; Maurer-Stroh and Eisenhaber, 2005), but in most cases the results have yet to be confirmed experimentally.

New pathways in plant metabolism, and some old ones revisited

The compartmentation of plant metabolism has been painstakingly investigated over the last 40 years using the techniques described above. Although the locations of most of the pathways of primary metabolism appear to be well established, recent findings have revealed new details, or even shown that maps of some pathways need to be revised, providing both energy and carbon skeletons for other metabolic pathways in these two compartments (ap Rees, 1987). A more recent finding was that many of the glycolytic enzymes in the cytosol are closely associated with the outer mitochondrial membrane (Giege et al., 2003). It has been suggested that this microcompartmentation allows the product of glycolysis, pyruvate, to be directly channelled into the mitochondria to fuel the Krebs cycle. In the chloroplasts, several Calvin cycle enzymes appear to form multi-enzyme complexes that are also associated with membranes, in this instance the thylakoid membranes where NADPH and ATP are synthesized, and this association might also permit channelling of metabolites (Winkel, 2004). In yeast and animals, several glycolytic enzymes act as metabolic sensors and transcription factors, and are found in the nucleus as well as the cytosol (Kim and Dang, 2005). It seems probable that some plant enzymes also have such dual metabolic and regulatory functions.

Gluconeogenesis

One of the classical examples of a highly compartmented pathway in plants is the conversion of lipids to sugars in germinating oilseeds via β-oxidation and the glyoxylate cycle. This pathway involves reactions in the peroxisomes (glyoxysomes), cytosol, and mitochondria (ap Rees, 1987). Although the locations of the enzymes in the pathway have been known for many years as a result of a series of studies on germinating castor bean seedlings (Beevers, 1980), questions remained about the ways in which carbon is transported between the three compartments and the exact role of the peroxisomal malate dehydrogenase. Various schemes for the pathway have been proposed by different authors, some of which show the peroxisomal malate dehydrogenase (MDH) oxidizing malate to oxaloacetate, even though Mettler and Beevers (1980) pointed out that the high NADH:NAD⁺ ratio in the peroxisome, arising from β-oxidation, would be unfavourable for the MDH reaction to proceed in the oxidative direction. Recent analysis of knockout mutants that lack the peroxisomal MDH has shown that the reaction proceeds in the other direction, converting oxaloacetate to malate, and re-oxidizing the NADH produced by β-oxidation of fatty acids (Baker et al., 2006; I Pracharoenwattana, JE Cornah, SM Smith, personal communication). Other genetic studies on peroxisomal citrate synthase mutants have shown that this enzyme is essential for β-oxidation of fatty acids, and that citrate is exported from the peroxisomes for conversion to isocitrate in the cytosol or oxidation via the citric acid cycle in the mitochondria, unlike in yeast where the peroxisomes export acetyl carnitine (Pracharoenwattana et al., 2005).

Also related to fatty acid–sugar interconversions was the surprising discovery of a role for Rubisco in the synthesis of lipids in developing oilseeds (Schwender et al., 2004). Imported sucrose is a major source of carbon for storage-product synthesis in most developing seeds. In oilseeds, it was generally thought that sucrose was metabolized via glycolysis to acetyl CoA, the substrate for fatty acid synthesis. However, enzyme activity measurements and stable-isotope labelling studies in oilseed rape (Brassica napus) led to the proposal of an alternative pathway, in
which hexose-phosphates derived from sucrose are metabolized via the non-oxidative steps of the pentose-phosphate pathway in conjunction with Rubisco. The proposed pathway would allow synthesis of 20% more acetyl CoA than the glycolytic route, but with a 40% reduction in the amount of carbon lost as CO₂, (Schwender et al., 2004).

**C₄ photosynthesis**

Another classical example of compartmentation in plant metabolism is C₄ photosynthesis. The C₄ cycle maintains a high concentration of CO₂ at the site of fixation by Rubisco in the bundle sheath cells, and so helps to prevent photorespiration in C₄ plants (Hatch, 1987). One of the key features of this pathway is that it is compartmented between two different cell types—mesophyll and bundle sheath cells—as well as between chloroplasts, cytosol, and mitochondria. CO₂ is initially fixed in the mesophyll cell cytosol by PEP carboxylase, forming the C₄-acid oxaloacetate, which is converted to malate and/or aspartate and transported via the plasmodesmata into the bundle sheath cells. There, the malate/aspartate is decarboxylated in the chloroplasts (NADP-malic enzyme species), mitochondria (NAD-malic enzyme species), or cytosol and mitochondria (PEP carboxykinase species), releasing CO₂ to be refixed by Rubisco. The residual C₃ molecule (pyruvate or PEP) is transported back to the mesophyll cells, where pyruvate is converted back to PEP, the initial CO₂ acceptor, by pyruvate, Pi dikinase (PPDK) in the chloroplasts. Other pathways are also compartmented between mesophyll and bundle sheath cells in C₄ plants; for example, sucrose is synthesized preferentially in the mesophyll cells, whereas starch synthesis is usually restricted to the bundle sheath cells (Lunn and Furmank, 1997). A recent proteomic study of mesophyll and bundle sheath cell chloroplasts from maize leaves suggests that synthesis of lipids, tetrapyrroles, and isoprenoids is also preferentially located in the mesophyll cells, as is nitrogen assimilation, whereas sulphur assimilation takes place preferentially in the bundle sheath cells (Majeran et al., 2005).

For many years it was believed that C₄ photosynthesis in terrestrial plants was inseparably linked to the possession of Kranz anatomy, where the bundle sheath cells are surrounded by a characteristic wreath of mesophyll cells. However, Voznesenskaya et al. (2001) discovered an exception to this rule in the plant *Borszczowia aralocaspica*, from the family Chenopodiaceae. This plant, and another species from the Chenopodiaceae, *Biernertia cycloptera*, both lack Kranz anatomy, but otherwise show all the characteristics of C₄ photosynthesis (Voznesenskaya et al., 2001, 2002). The photosynthetic chlorenchyma cells of *Borszczowia aralocaspica* have spatially separated photosynthetic enzymes, and possess two types of chloroplasts and other organelles arranged in separate parts of the cell (Voznesenskaya et al., 2001). The chlorenchyma cells of *Biernertia cycloptera* have an unusual central cytoplasmic compartment containing mitochondria and granal chloroplasts (Voznesenskaya et al., 2002). This is connected by cytoplasmic channels through the vacuole to the peripheral cytoplasm, which appears to lack mitochondria and has only grana-deficient chloroplasts. Although quite different, the intracellular compartmentation in both of these species allows spatial separation of the carboxylation and decarboxylation phases of the C₄ cycle within the same cell, showing that compartmentation of C₄ cycle enzymes between two cell types is not absolutely essential for C₄ photosynthesis. Nevertheless, these two interesting species represent rare exceptions to the general rule that terrestrial C₄ plants have Kranz anatomy.

Phylogenetic studies of the occurrence of C₄ photosynthesis have shown that this pathway evolved independently at least 31 times, within 18 separate plant families. In evolutionary terms, the appearance of C₄ photosynthesis in all of these families seems to have occurred over a short period of time, when atmospheric CO₂ levels were much lower than today and would favour C₄ over C₃ photosynthesis. At first sight, it seems quite surprising that such a complex pathway should have apparently arisen so readily in so many different plant lineages, but some new observations suggest that several features of C₄ plants were already present in their C₃ ancestors. Hibberd and Quick (2002) found that photosynthetic stems and petioles of some C₃ plants contain high activities of the C₄-acid decarboxylating enzymes, and that these could form part of a C₄ cycle. They proposed that C₄ acids (e.g. malate and aspartate) are synthesized in other organs, such as roots, and then transported via the phloem and xylem to the stems where they are decarboxylated to supply CO₂ for photosynthesis. In this way, cells in the stem that do not have direct access to atmospheric CO₂ via stomata can still contribute to net carbon fixation. The presence of such a C₄ cycle in C₃ plants suggests that it is not such a giant leap from C₃ to C₄ photosynthesis, and therefore the phylogetic evolution of C₄ photosynthesis is less surprising (Hibberd and Quick, 2002).

**The plastidial pathway of isoprenoid synthesis**

It is not often that a major new metabolic pathway in plants is discovered, but a recent example is the 1-deoxy-D-xylulose-5-phosphate (DOXP) pathway of isoprenoid biosynthesis. This pathway is located in the plastids, and is responsible for the synthesis of carotenoids, phytol, plastquinone-9, and isoprene, as well as mono- and diterpenes (Lichtenthaler, 1999; Eisenreich et al., 2001; Tobin and Bowsher, 2005). Before the discovery of this pathway, it was thought that all isoprenoids, including photosynthetic pigments and electron transport chain components, were synthesized via the acetate/mevalonate pathway, which is located in the cytosol. An obvious advantage of having the DOXP pathway in the chloroplasts is that the photosynthesis-related isoprenoids can be
Vitamin and coenzyme synthesis

The biosynthetic pathways of several vitamins and coenzymes represent one of the final frontiers of plant primary metabolism. Fluxes through many of these pathways can be vanishingly small, making it difficult to measure either enzymatic activities or pathway intermediates. However, the discovery and analysis of auxotrophic mutants of Arabidopsis thaliana with lesions in the pathways of vitamin biosynthesis, together with new information from studies in bacteria, is slowly uncovering these pathways. A good example is the synthesis of biotin (vitamin H), the coenzyme of acetyl CoA carboxylase, and several other carboxylases, decarboxylases, and transcarboxylases. The last enzyme in the pathway of biotin biosynthesis, biotin synthase, catalyses the conversion of dethiobiotin to biotin, and this enzyme was found to be located in the mitochondrial matrix (Baldet et al., 1997), whereas the first committed step in the pathway, the conversion of pimeloyl CoA to 7-keto-8-aminopelargonic acid (KAPA) by KAPA synthase, takes place in the cytosol (Pinon et al., 2005). The enzymes catalysing the two intermediate steps in the pathway, DAPA aminotransferase and dethiobiotin synthase, have not yet been characterized in plants, but Arabidopsis thaliana proteins showing similarity to the bacterial enzymes are predicted to be located in the cytosol and mitochondria, respectively (Pinon et al., 2005).

The L-galactose pathway of L-ascorbic acid (vitamin C) synthesis is similarly divided between the cytosol and mitochondria, with the terminal enzyme, L-galactono-1,4-lactone dehydrogenase, being located in the inner mitochondrial membrane, in association with complex I of the mitochondrial electron transport chain (Smirnoff et al., 2001). The biosynthesis of pantothenate (vitamin B₅) is also split between cytosol and mitochondria, but in the reverse orientation, starting in the mitochondria (ketopantoate hydroxymethyltransferase) and ending in the cytosol (pantothenate synthase) (Coxon et al., 2005). In the thiamine (vitamin B₁) biosynthetic pathway, the mRNA from the THI1 gene encoding the thiazole biosynthetic enzyme contains two alternative AUG translation initiation codons (Chabregas et al., 2003). The first of these appears to be the preferred start codon and gives rise to a chloroplast-targeted protein, while the second AUG produces a slightly smaller protein that is targeted to the mitochondria. This suggests that the thiazole moiety of thiamine is synthesized predominantly in the chloroplasts, but can also be made in the mitochondria. It is not yet known to what extent the rest of the thiamine biosynthetic pathway is duplicated between these two organelles or other compartments. The biosynthesis of folic acid (vitamin B₉) is split between three compartments: cytosol, chloroplasts, and mitochondria (Saht et al., 2005). Folic acid is composed of three moieties: a pterin ring, p-aminobenzoic acid (p-ABA), and a polyglutamate chain (Glu₁₋₈). Dihydropterin and p-ABA are synthesized in the cytosol and chloroplasts, respectively, and then transported into the mitochondria where they are condensed with up to eight glutamate molecules to produce folic acid.

By contrast to these multi-compartmented pathways, the biosynthesis of the vitamin B₂ complex (pyridoxal, pyridoxine, pyridoxamine, and their derivatives) appears to be located entirely within the cytosol (Tambasco-Studart et al., 2005), and riboflavin (vitamin B₂) synthesis is probably restricted to the plastids (Sandoval and Roje, 2005). Except for the last two steps of haem biosynthesis, which are duplicated in the mitochondria, the synthesis of tetrpyrrole coenzymes (chlorophyll, haem, sirohaem, and phytochromobilin) is also restricted to the plastids (Moulin and Smith, 2005). Lipoic acid, the coenzyme of several oxidative decarboxylases in the mitochondria, is made in the mitochondria (Yasuno and Wada, 1998; Gueguen et al., 2000).

It is interesting to note that not all coenzymes are synthesized where their partner enzymes are located; for example, most of the biotin synthesized in the mitochondria probably ends up bound to the plastidial and cytosolic isoforms of acetyl CoA carboxylase. Presumably some of these coenzymes must be transported between different intracellular compartments, and perhaps between cells. One of the sucrose-H⁺ symporters in Arabidopsis thaliana (AtSuc5) can also transport biotin (Ludwig et al., 2000), and a folate transporter has been identified in the chloroplast envelope (Bedhomme et al., 2005). Apart from these examples, little is known about vitamin and coenzyme transporters in plants at present.

The apparent importance of the mitochondria in so many vitamin biosynthetic pathways is somewhat unexpected, given the dominance of the plastids in many other biosynthetic pathways. Some of these vitamin biosynthetic pathways could have been directly inherited from the bacterial endosymbiont that gave rise to the mitochondria, whereas others might have been transferred from the cyanobacterial ancestor of the chloroplasts. The dual targeting of the thiazole biosynthetic enzyme to both chloroplasts and mitochondria suggests one possible mechanism by which enzymes, and indeed whole pathways, could be shuffled between these organelles.
Communication between compartments

One of the consequences of metabolic compartmentation is the need for a whole host of transporters to move metabolites around the cell. These are usually integral membrane proteins that are less amenable to experimental investigation than soluble proteins, so only a few transporters have been thoroughly characterized in plants. Multi-compartmented pathways also need regulatory mechanisms to balance the fluxes through the separate branches of the pathway. A good example is photosynthetic carbon metabolism, where the rate of sucrose synthesis in the cytosol is co-ordinated with the rate of CO₂ fixation and starch synthesis in the chloroplasts. These pathways are linked by the triose-phosphate translocator in the chloroplast envelope, which imports Pi from the cytosol in exchange for triose-phosphates. The transported metabolites also act as signals that convey information about metabolic fluxes in the originating compartment; triose-phosphates exported from the chloroplast influence the level of the signal metabolite Fru2,6P₂ in the cytosol, which regulates the cytosolic fructose-1,6-bisphosphatase, and Pi released by sucrose synthesis in the cytosol is needed in the chloroplast for photophosphorylation (MacRae and Lunn, 2006). Pi is an allosteric inhibitor of the plastidial ADP-glucose pyrophosphorylase, and so the stromal concentration of Pi also influences the rate of starch synthesis. These and other mechanisms, including post-translational regulation of several enzymes, co-ordinate sucrose synthesis with the rate of CO₂ fixation, preventing excessive withdrawal of triose-phosphates from the Calvin cycle, but maintaining an adequate supply of Pi in the chloroplast for photophosphorylation.

There is emerging evidence that trehalose 6-phosphate (Tre6P), the intermediate of trehalose synthesis, could also be involved in communication between cytosol and chloroplast. Tre6P appears to be synthesized in the cytosol in plant cells, and the amount of this metabolite is strongly correlated with the level of sucrose (Lunn et al., 2006). It has been found that supplying Tre6P to isolated chloroplasts promotes the redox activation of ADP-glucose pyrophosphorylase, although the mechanism of this effect is not yet understood (Kolbe et al., 2005). From these observations it has been proposed that accumulation of sucrose in the leaf leads to a rise in the level of Tre6P in the cytosol, which in some way leads to activation of ADP-glucose pyrophosphorylase in the chloroplasts, and so diverts photosynthate away from sucrose and into starch while maintaining a steady rate of CO₂ fixation. Thus, one of the functions of Tre6P in plants could be to act as a messenger service between the cytosol and the chloroplasts.

Phylogenetic analysis of some of the enzymes of sucrose and starch synthesis has revealed some interesting clues to the evolutionary history of their compartmentation. Most of the enzymes in these pathways appear to have a cyanobacterial origin, presumably stemming from the endosymbiosis of the chloroplast ancestor. Therefore, it seems likely that both sucrose and starch would have been made in the chloroplasts of the earliest photosynthetic eukaryotes. During the evolution of plants, most (90–95%) of the ancestral cyanobacterial genes were lost from the chloroplast genome. Many were transferred to the nucleus, and in Arabidopsis thaliana these account for about 18% of the protein-coding genes in the nuclear genome (Martin et al., 2002). Despite their origin, the majority of these genes encode proteins that end up being located outside the plastids. These include the key enzymes of sucrose synthesis—sucrose-phosphate synthase and sucrose-phosphatase—as well as neutral/alkaline invertase, which hydrolyses sucrose in the cytosol (Lunn, 2002; Vargas et al., 2003). Interestingly, the acid invertases in the vacuole and apoplast appear to have a different evolutionary origin (Vargas et al., 2003).

Although the key enzymes of starch synthesis (ADP-glucose pyrophosphorylase, starch synthases, isoamylases, starch branching enzymes) are all encoded by nuclear genes, the proteins themselves are located in the plastids. Surprisingly, phylogenetic analysis showed that not all of these plastidial enzymes have a cyanobacterial origin. The starch branching enzymes are more closely related to enzymes from non-photosynthetic eukaryotes, and have entirely replaced any cyanobacterial form of the enzyme that might once have been present (Patron and Keeling, 2005). Thus, the pathway of starch synthesis appears to have been reconstituted in plastids by combining eukaryotic and cyanobacterial elements. Curiously, in red algae, whose plastids are derived from the same endosymbiotic event that gave rise to chloroplasts, the pathway for synthesis of floridean starch appears to have been reconstituted in a similar way, but in the cytosol, not the plastids (Patron and Keeling, 2005). These pathways, and the adaptation of pre-existing elements to form the C₄ photosynthetic pathway, both indicate that there has been extensive shuffling of genes and proteins between compartments during the evolution of plants.

Concluding comments

Identifying pathways and mapping their intracellular locations are the essential first steps for understanding plant metabolism. Genome sequencing has led to the discovery of new metabolic pathways in plants, and new technologies such as protein mass spectrometry and fluorescent protein tagging are providing fresh insights into the compartmentation of plant metabolism. However, despite the rapid advances of recent years, there are still some major gaps in our knowledge of primary metabolism in plants. For example, the biosynthetic pathways for several coenzymes have not yet been fully resolved, and there is only limited understanding of the synthesis, transport, and assembly of
many cell wall components. Knowledge of plant secondary metabolism is even more fragmentary. Collectively, plants synthesize a huge range of secondary metabolites (e.g. isoprenoids, alkaloids, and phenylpropanoids), but for many of these only the early steps in the biosynthetic pathways are known.

It is not only in biosynthetic plant metabolism where there are gaps in our knowledge. Plants often grow under nutrient-limited conditions, and recycling of nutrients from turnover of cell constituents (e.g. nucleotides, amino acids, membrane lipids, and photosynthetic pigments) allows them to reallocate limited resources to maximize their chances of survival or their ability to reproduce, for example, by degrading photosynthesis-related proteins in the leaves to supply nitrogen to developing seeds. Such reallocation can also have an important influence on the yield of crop plants, yet few of the catabolic and salvage pathways involved in nutrient recycling in plants have been fully characterized, and some are still a matter of guesswork based on information from other organisms.

Metabolite transport is yet another aspect of metabolic compartmentation where our knowledge is incomplete. Although the transport of metabolites can be inferred from the locations of the enzymes in the pathway, many of the proteins involved in membrane transport and intracellular trafficking have not yet been identified. In conclusion, it appears that there is still some way to go before there is a complete map of plant metabolism and, without this, efforts to identify regulatory networks or to understand plant metabolism as an integrated whole are hampered.

Acknowledgements

I thank Steve Smith (University of Western Australia) for providing preprints of unpublished papers and other useful background information. I also thank Mark Stitt and Alisdair Fernie (MPIMP, Golm) for helpful discussions and suggestions.

References


