REVIEW PAPER

Strategies for investigating the plant metabolic network with steady-state metabolic flux analysis: lessons from an Arabidopsis cell culture and other systems

N. J. Kruger*, S. K. Masakapalli and R. G. Ratcliffe*

Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

* To whom correspondence should be addressed. E-mail: nick.kruger@plants.ox.ac.uk or george.ratcliffe@plants.ox.ac.uk

Received 2 September 2011; Revised 24 October 2011; Accepted 26 October 2011

Abstract

Steady-state 13C metabolic flux analysis (MFA) is currently the experimental method of choice for generating flux maps of the compartmented network of primary metabolism in heterotrophic and mixotrophic plant tissues. While statistically robust protocols for the application of steady-state MFA to plant tissues have been developed by several research groups, the implementation of the method is still far from routine. The effort required to produce a flux map is more than justified by the information that it contains about the metabolic phenotype of the system, but it remains the case that steady-state MFA is both analytically and computationally demanding. This article provides an overview of principles that underpin the implementation of steady-state MFA, focusing on the definition of the metabolic network responsible for redistribution of the label, experimental considerations relating to data collection, the modelling process that allows a set of metabolic fluxes to be deduced from the labelling data, and the interpretation of flux maps. The article draws on published studies of Arabidopsis cell cultures and other systems, including developing oilseeds, with the aim of providing practical guidance and strategies for handling the issues that arise when applying steady-state MFA to the complex metabolic networks encountered in plants.

Key words: Arabidopsis, cell culture, metabolic flux analysis, heterotrophic carbon metabolism, isotope labelling, mass spectrometry, metabolic modelling, network analysis, nuclear magnetic resonance, primary metabolism.

Introduction

The flow of material between two metabolites in a metabolic network constitutes a metabolic flux, and collections of fluxes between different points in the network can be assembled into a flux map. Since the purpose of metabolic networks is to support such fluxes, flux maps define a fundamental metabolic phenotype that reflects this function (Ratcliffe and Shachar-Hill, 2005). Unfortunately most of the fluxes supported by a typical network cannot be measured directly and they have to be inferred from isotopic labelling experiments, typically with 13C-labelled substrates, using the techniques of metabolic flux analysis (MFA; Roscher et al., 2000; Ratcliffe and Shachar-Hill, 2006). One of these techniques, steady-state 13C-MFA, has been used extensively to analyse the network of central carbon metabolism in plant tissues, and the development of this field over the last decade has been charted in review articles (Kruger et al., 2003; Schwender et al., 2004; Ratcliffe and Shachar-Hill, 2006; Kruger and Ratcliffe, 2007, 2009; Libourel and Shachar-Hill, 2008; Schwender, 2008, 2009; Allen et al., 2009a). In passing it may be noted that this approach greatly extends the scope for determining fluxes in the central metabolic network in comparison with the earlier use of 14C-labelling (ap Rees, 1980; Fernie et al., 2005).

In steady-state 13C-MFA, the cellular system is incubated with a 13C-labelled precursor and the redistribution of the label into metabolic intermediates and end-products is analysed when the system has reached an isotopic and metabolic steady state. This redistribution of label depends on the structure of the metabolic network and the fluxes it supports, and the fluxes can be deduced by fitting a model of the network to the labelling data (Fig. 1; Wiechert et al., 2001). Although informative, steady-state MFA is experimentally
and computationally non-trivial, and so this paper aims to provide some practical guidance on the implementation of the technique. Specifically the article addresses issues that commonly arise when applying the technique to the complex network of plant metabolism.

**The plant metabolic network**

**Definition of the network for steady-state MFA**

The complete metabolic network encompasses thousands of reactions that can be roughly divided into the highly connected pathways of central carbon metabolism, and the more specialized, and frequently linear, pathways of secondary metabolism. Steady-state $^{13}$C-MFA is primarily a tool for analysing the flux distribution in central metabolism, and the method hinges on finding a flux distribution through the network that explains the observed redistribution of the label into metabolic intermediates and end-products. So to proceed it is necessary to define the network of reactions that are considered to be responsible for the redistribution of the label.

In contrast to the genome-scale networks of plant metabolism that are currently being developed for flux balance analysis, and which typically include many hundreds of reactions (Sweetlove and Ratcliffe, 2011), the networks used for steady-state MFA are usually constructed manually on the basis of well established biochemical knowledge. Moreover these networks are considerably less detailed than conventional maps of the pathways of central metabolism or genome-scale models, because consecutive steps between branch points in the network necessarily carry the same
net flux and are therefore represented as a single reaction. Typically the network will include inputs from one or more substrates, a representation of the pathways of central metabolism, with each step defined as reversible or irreversible, and a series of outputs to carbon dioxide and biomass components, such as protein, lipid, and stored carbohydrate (Fig. 2). Each step in the network contributes to the redistribution of the $^{13}$C-label, and so the model that describes the network has to include information on the fate of individual carbon atoms (Supplementary Fig. S1 available at JXB online). This is provided in the carbon atom mapping matrix, which specifies the relationship between the carbon atoms in adjacent metabolites (Zupke and Stephanopoulos, 1994). Thus the network used for steady-state MFA is specifically constructed to reveal the metabolic fluxes that redistribute the label, rather than to model every biochemical reaction in the system.

**Subcellular compartmentation**

Subcellular compartmentation is a prominent feature of the plant metabolic network (Lunn, 2007; Kruger and Ratcliffe, 2008; Sweetlove *et al.*, 2008) and it affects the implementation of steady-state MFA in three ways.

First, it is necessary to assign steps and pathways to the correct compartment or compartments on the basis of current knowledge. This appears to be straightforward for the pathways of central metabolism, but there is a steady flow of reports of proteins and enzyme activities in unexpected compartments (e.g. fumarase; Pracharoenwattana *et al.*, 2010), and these pathways will need to be considered explicitly if they support significant fluxes in particular tissues.

Secondly, including subcellular compartmentation increases the analytical and computational challenge, because information that reports on the labelling of metabolite pools...
in specific compartments is required to deduce the fluxes in
those compartments (Ratcliffe and Shachar-Hill, 2005, 2006;
Allen et al., 2007). Much of this information is lost on
extraction and, although the techniques of subcellular
fractionation might help (Allen et al., 2009a; Kruger et al.,
2011), they are rarely incorporated into the extraction
procedures that are currently used in MFA. Accordingly,
the labelling pattern of compartmented intermediates is
usually deduced from the labelling of an end-product that is
synthesized in a particular compartment. For example,
starch is generally considered to be synthesized exclusively
from the plastidic glucose 6-phosphate pool (Streb et al.,
2009), and this allows the labelling of that pool to be deduced
from the labelling of the glucosyl moieties of starch.

Finally, some of the intermediates that are present in
multiple compartments are able to exchange between
compartments, which leads to the mixing of pools that
might be differentially labelled. In some instances, the rate
of exchange may be sufficiently rapid to eliminate any
intrinsic difference between the labelling of the two pools,
allowing the pools to be merged in the model with
a labelling pattern equivalent to that of the extracted
metabolite; while for other exchangeable metabolites there
may be evidence that the labelling pattern of the two pools
is different. For example, in Arabidopsis cell suspensions
the difference in the steady-state labelling of the glucosyl
moieties of sucrose and starch indicates that the cytosolic
and plastidic pools of glucose 6-phosphate have different
labelling patterns (Williams et al., 2008; Masakapalli et al.,
2010), whereas in Brassica napus (Schwender et al., 2003) and
soybean (Allen et al., 2009b) embryos there are no significant
differences in the labelling of end-products derived from the
cytosolic and plastidic glucose 6-phosphate pools, allowing
them to be treated as a single pool. In practice a typical
network model will merge the compartmented pools of some
metabolites, while maintaining others as separate pools.

Ignoring or oversimplifying the subcellular compartmenta-
tion of the network can have significant implications for flux
maps and the metabolic conclusions that can be drawn from
them. For example, glucose is generally expected to be
regenerated from the hexose phosphate pool in heterotro-
phic tissues, either by sucrose cycling (Dieuaide-Noubhani
et al., 1995) or by the action of a glucose-6-phosphatase
(Alonso et al., 2005). Surprisingly high rates of glucose
regeneration have been deduced from steady-state MFA for
several tissues, including maize root tips (Dieuaide-Noubhani
et al., 1995; Alonso et al., 2007c), cultured tomato cells
(Rontein et al., 2002), and the endosperm of developing maize
seeds (Alonso et al., 2011), implying that glucose regeneration
is a major sink for ATP. In all these studies it was assumed
that glucose regeneration occurred in the cytosol, but a
theoretical analysis has shown that this assumption has
a critical bearing on the outcome. In particular, if the vacuole
is involved in regenerating glucose, for example through
the action of a vacuolar invertase or glucose-6-phosphatase, then
estimates of the pathway flux from steady-state MFA are
likely to be unreliable in the absence of compartment-specific
information on the labelling of glucose in the cytosolic and
vacuolar pools (Kruger et al., 2007b). This information is not
readily available, and in this instance assuming that the
labelling is the same has the potential to overestimate sucrose
cycling and/or glucose-6-phosphatase activity. So, in the
absence of a demonstration that the labelling of the cytosolic
and vacuolar glucose pools is indeed the same, the network
has to be simplified to exclude the regeneration of glucose.

Indeterminable and poorly defined fluxes

The next point to appreciate is that the reliability of the
fluxes that can be deduced by fitting the network model to
the labelling data is strongly dependent on the available
measurements (Wiechert et al., 2001). The labelling of any
metabolite is influenced to different extents by the fluxes in
the network, and in turn some measurements are more useful
than others in determining a particular flux. Thus, in a typical
network, some fluxes will be well determined by the available
data, while others will be poorly defined. The latter group
may include some fluxes that are not constrained at all, and
are thus structurally indeterminable (Fig. 3). Indeterminable
fluxes typically arise when there are two routes between a pair
of metabolites, but no labelling measurements in the existing
data set to discriminate between them. In this situation, an in
silico analysis may be able to identify label measurements
that could solve the problem; however, if the solution is
impracticable (e.g. it may not be possible to acquire the
missing data), then the alternative is to simplify the network
to eliminate the indeterminable fluxes. For example, in many
systems, it is not possible to discriminate between the
conversion of 2-oxoglutarate to succinate via the tricarbox-
ylic acid (TCA) cycle or the GABA (γ-aminobutyric acid) shunt, and so the two alternatives have to be merged into
a single pathway. The flux through such a merged pathway is
often attributed to the TCA cycle, although strictly this is
unjustifiable without further information.

Iteration of the network structure

In practice the definition of the metabolic network that best
explains the experimental data is likely to be an iterative
process along the following lines:

- Set up the largest possible network based on the available
biochemical information on the relevant pathways and
their compartmentation.
- Identify and eliminate structurally indeterminable fluxes
by simplifying the network.
- Use the resulting network to deduce the best-fit flux map
from the labelling data set.
- Identify all well-defined zero fluxes and consider eliminat-
ing them.
- Identify all poorly defined fluxes and consider eliminating
them if they cannot be improved by further data collection.

This procedure will lead to the largest flux map in which
all the fluxes are well defined, and the underlying network
can then be fixed and used for comparative studies of the
system under different conditions. The rationale behind this
scheme is that a flux map is only useful if the fluxes are well defined, and the largest possible network that generates such fluxes is likely to provide the most informative metabolic phenotype.

Experimental considerations

Cells and tissues

Steady-state MFA is applicable to systems in a metabolic steady state that have been labelled to an isotopic steady state. In this state, the label enrichment of each carbon atom is constant, but to be informative the distribution of the label should be non-uniform, with different carbon atoms being labelled to different extents. These requirements constrain the choice of experimental system in two fundamental ways (Roscher et al., 2000; Ratcliffe and Shachar-Hill, 2006).

First, the metabolic state of the tissue of interest may be changing over time, for example in response to a stress or a diurnal cycle, and the rate of change may be too rapid for the system to attain an isotopic steady state that reflects the current metabolic steady state. Metabolites and end-products turn over on time scales that range from seconds to days (Ratcliffe and Shachar-Hill, 2006; Troufflard et al., 2007), and so it takes a similar length of time for the corresponding pools to reach isotopic equilibrium. Thus steady-state MFA is most readily applied to cell suspensions and excised tissues that can be cultured under constant conditions for extended periods, typically days, in the presence of a labelled substrate. If this is impracticable, then the alternative is to use non-steady-state methods to analyse labelling time-courses, but this approach is better suited to the analysis of secondary metabolism (Heinzle et al., 2007), and widely applicable methods for the analysis of central metabolism are only just becoming available (Nöhl and Wiechert, 2011).

Secondly, steady-state 13C-MFA exploits the principle that if different pathways to the same metabolite lead to different labelling patterns, then the relative fluxes through the pathways can be deduced from the observed labelling pattern. The approach fails if the labelling strategy leads to uniform labelling of every metabolite, and the most important situation in which this could occur is in photoautotrophic organisms, where the assimilation of labelled carbon dioxide as the sole carbon source would lead to an isotopic steady state in which all the carbon atoms in every metabolite and end-product would be equally labelled (Roscher et al., 2000; Shastri and Morgan, 2007). Thus steady-state MFA cannot be used to deduce fluxes from such an experiment, and as before the alternative is to analyse labelling time-courses (Shastri and Morgan, 2007; Hasunuma et al., 2010; Huege et al., 2011).

Isotopic and metabolic steady states

Steady-state MFA assumes that the system is in a metabolic and isotopic steady state. Evidence for the metabolic steady state can be obtained by measuring the rates of substrate utilization, end-product accumulation, and growth; while justification for the isotopic steady state can be obtained by analysing a labelling time-course. For example, it was found that the rates of glucose consumption and biomass accumulation in a heterotrophic Arabidopsis cell culture were the same at 4.5 d and 5.5 d after subculture; and it was also shown that the labelling of a range of low molecular weight metabolites, as well as the labelling of the amino acids derived from hydrolysed protein, reached a constant value after 5 d (Williams et al., 2008). It has also been established that growth on 13C-labelled glucose had no discernible effect on the growth of the cells and the flux distribution in the network, thus verifying an implicit assumption of steady-state MFA (Kruger et al., 2007a).

The requirement for an isotopic steady state is less restrictive than it appears because it only applies to the measurements that are used to deduce the fluxes through the

Fig. 3. Exploring flux determinability using a simple model network. Flux determinability in a network containing four free fluxes, R2net, R2xch, R4xch, and R5xch, was assessed for flux distributions corresponding to (A–D) R2net=0, (E–H) R2net=0.4, and (I–L) R2net=1 (Supplementary information at JXB online). The reliability of the flux estimates is dependent on the distribution of flux through the network, ranging from excellent determinability (I) to almost complete indeterminability (J), where the flux estimate shows an almost uniform distribution across the full range of permissible values. An exhaustive analysis of this and related models shows that flux indeterminability may result from the structure of the network, the flux distribution in the network, or insufficient/inappropriate label measurements.
network. In principle, this relaxes one of the constraints on the application of steady-state MFA, since it suggests that it might be possible to deduce the fluxes through central metabolism purely on the basis of measurements of the labelling of metabolites that turn over rapidly and thus reach isotopic steady state within minutes or hours. However, in practice, labelling information from protein, starch, and cell wall components is often essential for deducing the fluxes in the compartmentalized network of plant metabolism, since these measurements report on the labelling of subcellular pools of intermediates that would otherwise be inaccessible to conventional analysis. Thus, to date, all the published flux maps of the central metabolic network in plants have been based in part on labelling information derived from end-products that usually take days to reach isotopic equilibrium.

Potential strategies for avoiding the prolonged labelling time required for some of these end-products to reach isotopic equilibrium include the incomplete digestion of starch granules to reveal the labelling of recently synthesized starch (Keeling et al., 1988), starvation prior to labelling to deplete stores of starch and other carbohydrates (Dieuaide-Noubhani et al. 1995), and the use of inducible reporter proteins (Shaikh et al., 2009). Only the second of these options has been implemented so far in steady-state MFA of the central metabolic network in plants.

An alternative strategy, applicable when it is clear that the labelling of soluble amino acids is higher than of those derived from hydrolysed protein, is to assume that the extracted protein is a mixture of protein that has been labelled to isotopic steady state during the course of the experiment and pre-existing unlabelled protein (Lonien and Schwender, 2009; Williams et al., 2010). This assumption can be readily incorporated into the metabolic network (Fig. 4), allowing the abundance of the unlabelled pool to be determined during the fitting process. Note that the number of parameters is usually much less than the number of measurements in steady-state MFA, so the extra parameters associated with this approach to the problem of the protein failing to reach an isotopic steady state can be easily accommodated in the fitting procedure.

13C-labelling strategies

There are three steps in developing the labelling strategy:

- Selection of the substrate.
- Selection of the substrate isotopeomers to be used as the source of label.
- Experimental design.

The first step, the selection of the substrate(s), is usually obvious from the physiological context. Carbohydrates are the most commonly used substrates, reflecting their status as the primary form in which organic carbon is translocated in plants (Turgeon and Wolf, 2009) and the limitation of steady-state MFA to heterotrophic and mixotrophic metabolism. Thus, for most purposes, 13C-label can be supplied in the form of glucose. While sucrose would be the usual source of carbohydrate both in vivo and in many culture systems, sucrose is converted to hexose units before it enters metabolism and so there is no advantage from a computational perspective in using the more expensive 13C-labelled sucrose as a source of label. Occasionally there might be a compelling physiological reason for using sucrose, as found for Arabidopsis embryos grown in liquid culture (Lonien and Schwender, 2009), and other substrates can be used if there is evidence for their involvement in the normal metabolism of the tissue or cell type. For example, developing oilseed embryos are usually grown on a mixture of carbohydrates and amino acids (Allen et al., 2007), and typically steady-state MFA is then based on experiments with both labelled glucose and labelled glutamine (Allen et al., 2007; Alonso et al., 2010a).

The next step is to decide on the positional labelling of the substrate. For example, [1-13C]glucose, [2-13C]glucose, and [U-13C6]glucose are easily obtainable and, in principle, a flux map of the central metabolic network could be based on a steady-state labelling experiment with any of them. However, the three experiments would not provide
equivalent information about the flux distribution, because the structure of the network, the fluxes that it supports, and the extent of the labelling information all determine the reliability with which a particular labelled precursor can report on a particular flux. In silico analysis provides the surest way to optimize the labelling strategy (Libourel et al., 2007), and the usual conclusion is that it will be necessary to use more than one labelled substrate to achieve adequate coverage of the network (Schwender et al., 2004, 2006).

This conclusion is supported by a statistical analysis of the flux estimates obtained with different labelling strategies in a heterotrophic Arabidopsis cell culture (Masakapalli et al., 2010). Labelling data sets from [1-\(^{13}\)C]glucose, [2-\(^{13}\)C]glucose, and [U-\(^{13}\)C_6]glucose experiments were each able to define the net and exchange fluxes in the network, but statistical analysis showed that the flux estimates from the [U-\(^{13}\)C_6]glucose experiment were markedly worse than those obtained with the other two experiments. However, simultaneous analysis of pairs of data sets improved the flux estimates, sometimes substantially, and there was a further improvement when all three data sets were combined.

The remaining question relates to experimental design, and in particular the optimum way in which to conduct labelling experiments with multiple substrates. In the microbial field, it is common for the system to be incubated with a mixture of labelled substrates (Wiechert et al., 2001). However, in silico analysis suggests that it may often be better to run parallel experiments with each of the labelled substrates separately, and then to analyse the data from these experiments simultaneously, using a procedure that constrains the flux solutions for each experiment to be the same (Fig. 5A–D).

This approach has been used to produce well-defined flux maps for several systems, including oilseed rape embryos (Schwender et al., 2006), Arabidopsis cells (Masakapalli et al., 2010), and maize embryos (Alonso et al., 2010a).

**Fig. 5.** Factors influencing the reliability of flux estimates obtained by steady-state MFA. The impact of the labelling strategy, the precision of the label measurements, and the extent of replication were assessed using a simple model network (Supplementary information at JXB online). There are four free fluxes in the model identified as R_{2net}, R_{2xch}, R_{4xch}, and R_{5xch}. (A–D) Flux estimates based on mass isotopomer measurements from experiments using four different labelling strategies: (1) 50% [1-\(^{13}\)C]substrate; (2) a mixture containing 50% [1-\(^{13}\)C]substrate and 20% [\(^{13}\)C_3]substrate; (3) separate experiments with 50% [1-\(^{13}\)C]substrate and 20% [\(^{13}\)C_3]substrate, analysed simultaneously; and (4) 20% [\(^{13}\)C_3]substrate. Flux estimates are presented as a box plot representing the median and interquartile range, with the whiskers displaying the 90% confidence range and crosses identifying the 95% confidence limits. The dotted lines indicate the maximum exchange flux (\(v_{xch}=0.99\)) permitted in 13C-FLUX. The reliability of the flux estimates obtained by combining information from two different labelling experiments performed in parallel (3) exceeds that obtained with either substrate alone (1 and 4) or the mixture (2). (E–H) Flux estimates based on positional isotopomer measurements from 100% [1-\(^{13}\)C]substrate experiments. The median flux estimates are indicated by the solid white lines and the 95% confidence intervals are shown for data with a coefficient of variation of 10% (light grey), 5% (grey), 2% (dark grey), and 1% (black). Increasing measurement precision increases the reliability of the flux estimates, but the confidence intervals are not necessarily symmetrical; and increasing the number of replicate experiments is beneficial.

**Determination of biomass constraints**

Rates of biomass accumulation provide important constraints on the interpretation of the labelling data in most steady-state MFA experiments, because the flux distribution that is deduced from the redistribution of the input label has to be able to support the observed outputs from the network. These outputs, which might include some or all of
the following—protein, lipid, nucleic acids, storage carbohydrates, cell wall components, and low molecular weight metabolites that accumulate in the vacuole—can be determined in tissue extracts using standard methods. Sriram et al. (2006) provide a typical protocol, although in practice there are many variations, reflecting both researcher preferences and the specific requirements for particular cells and tissues. A sensitive alternative is to incubate the system with a [U-14C]substrate and then to analyse the redistribution of the label in a fractionated extract after the pools of metabolic intermediates have been labelled to isotopic steady state (Masakapalli et al., 2010). The high sensitivity of this method makes it likely to be particularly attractive in systems where the relative growth rate in the metabolic steady state is slow.

Analysing the redistribution of label

The labelling of metabolites and end-products at isotopic steady state is analysed by nuclear magnetic resonance (NMR) spectroscopy and/or mass spectrometry (MS) (Ratcliffe and Shachar-Hill, 2006; Allen and Ratcliffe, 2009). The analysis of 13C-labelling patterns is a routine application for both techniques, and ideally both approaches should be used to maximize the confidence in the deduced fluxes (Kleijn et al., 2007).

In outline, NMR has the advantage of minimal sample handling, and the NMR signals provide information on the labelling of specific carbon atoms. This leads to the quantification of positional isotopomers (i.e. molecules with the same chemical structure that differ in isotopic (13C) composition at particular carbon atoms), and as well as characterizing the redistribution of the label, this information can be readily used to identify the pathways that are involved in the metabolic redistribution of the label. Another feature of the NMR method is that the labelling of adjacent carbon atoms usually leads to characteristic line shapes that can be invaluable in tracking carbon–carbon bond formation and breakage.

MS methods, whether gas chromatography (GC) MS or liquid chromatography (LC) MS, quantify the labelling of metabolites in terms of mass isotopomers (i.e. molecules or fragments that differ in the number of labelled atoms). Mass isotopomer abundances can be measured with high precision, and the sensitivity of MS greatly exceeds that of NMR, permitting the analysis of smaller samples and less abundant metabolites. The two methods provide complementary information about the labelling of the metabolic network and, while either method can provide sufficient information for steady-state MFA, typical NMR and MS data sets will constrain individual fluxes in the flux map to different extents, emphasizing the desirability of using both NMR and MS in the analysis.

The key analytical objective in steady-state MFA is to collect sufficient labelling information to allow the problem of flux determination to be overdetermined. Thus in contrast to early applications to plants, where the number of measurements was just sufficient to allow the unknown fluxes to be determined from a set of algebraic relations (Dieuaide-Noubhani et al., 1995; Roscher et al., 2000; Fernie et al., 2001; Rontein et al., 2002), the aim is to collect a much larger data set that will allow a statistically robust fitting procedure. Fitting the model to an excess of measurements is akin to a regression analysis and leads to estimates of the reliability of the deduced fluxes; whereas the earlier approach leads to an exact solution for the unknown fluxes with no estimate of the confidence intervals.

When implementing this model-fitting approach, it should be noted that not all measurements are equal, because the extent to which an individual measurement constrains a particular flux depends on the network structure. This leads to the notion that it should be possible to identify a minimal data set that would still be sufficiently large to over-determine the fluxes by a substantial margin, but which would minimize the analytical effort required to obtain the measurements (Chang et al., 2008). However the optimum data set for a network depends on the fluxes that it supports, and since these are initially unknown it is difficult to make reliable predictions at the outset of an investigation. So, in practice, most investigations are not based on an in silico analysis of the optimum data set, and it is more usual to adopt an empirical approach in which the labelling of a range of metabolites and macromolecules is measured without specific reference to the extent to which any given isotopomer might be useful. Thus a typical data set for an Arabidopsis cell culture might include measurements of soluble metabolites, including amino acids, organic acids, and sucrose, amino acids derived from protein hydrolysates, and the glucose derived from hydrolysed starch (Williams et al., 2008, 2010; Masakapalli et al., 2010). Two important principles should guide the collection of the labelling data set.

First, high accuracy and precision are essential for deducing a useful flux map during the fitting process. Accuracy requires careful assignment of the NMR signals and MS fragments, and the rejection of any measurement where there is a possibility of overlap with other signals (Antoniewicz et al., 2007a). The fitting of the network model to the data is an iterative process, and this gives the opportunity to check the assignment and resolution of poorly fitting signals during the analysis. The precision of the measurements is also important, since this has a major bearing on the fitting process and the confidence limits on the deduced fluxes (Fig. 5E–H). Imprecise measurements reduce the stringency of the fitting process, but this will lead to poorly defined flux estimates with large standard deviations. On the other hand, if the precision of the measurements is overestimated, then it may be difficult to obtain a solution in the fitting process that passes a χ2 test. Thus a realistic assessment of the precision of the measurements is a prerequisite for deducing reliable flux values (Williams et al., 2008).

Secondly, as emphasized earlier, it is important to analyse metabolites and macromolecules that can report on the labelling patterns of intermediates within specific subcellular compartments (Ratcliffe and Shachar-Hill, 2005, 2006; Allen et al., 2007). This is the only routine way in which compartmental information can be obtained in steady-state MFA, and the availability of this information determines the extent to which the flux map will capture the subcellular
compartmentation of studies of all eukaryotic systems and it is one of the reasons why the application of steady-state MFA to such systems is less routine than in prokaryotes (Niklas et al., 2010).
by a principal component analysis of the flux solutions (Masakapalli et al., 2010). The mean of the selected solutions can then be used as a set of starting fluxes for a final fit of the model to the data, leading to the final best fit, which is the set of free flux estimates that comes closest to explaining the labelling data.

There are several ways in which the statistical significance of the final flux map can be evaluated. At one level, it is desirable for the model to pass the $\chi^2$ test, and most recently published flux maps do so, even though there was originally an expectation that the models used for steady-state MFA might satisfy this condition rather rarely (Wiechert et al., 2001). Assuming that the model does pass the test, then the next task is to assign confidence limits to the deduced fluxes. In steady-state applications of MFA to plants, an approximate method that is available as a subroutine (EstimateStat) in the commonly used 13C-FLUX software is often used to estimate standard deviations for the fluxes (e.g. Schwender et al., 2006; Williams et al., 2008; Alonso et al., 2010a). However, this method is strictly inapplicable to the non-linear models that describe the redistribution of label in steady-state MFA, and a more rigorous non-linear method for the determination confidence intervals has been proposed (Antoniewicz et al., 2006) and implemented in more recent software packages such as OpenFLUX and METRAN. The method is not currently supported in other MFA software, but it can be implemented manually in 13C-FLUX by constraining a specific flux to a series of values and repeatedly re-optimizing the fit of the model to the data. This procedure has been successfully applied in an analysis of developing Arabidopsis embryos (Lonien and Schwender, 2009) and in Arabidopsis cell cultures (Williams et al., 2011).

**Working with the flux map**

The biological applications of steady-state MFA are discussed at length in the review articles cited in the Introduction, and so this section focuses on several generic points that arise in the interpretation of flux maps.

The first point to consider is the validation of the flux map. Typically authors cite good agreement between the predicted and experimental labelling measurements, a successful $\chi^2$ test, and acceptable confidence limits for the flux estimates as evidence that the flux map provides a good description of the steady-state labelling experiment. Beyond this, it is possible to validate a model by making predictions about network properties from the flux map. These predictions could include the outcomes of positional radiorespirometric analysis (Masakapalli et al., 2010) or measurements of carbon conversion efficiency (e.g. Alonso et al., 2010a). It is also possible to check that extractable enzyme activities are sufficient to support the measured fluxes, although in a substantial study in which this was undertaken, it was found that most enzyme activities were in considerable excess and that changes in enzyme activity between two conditions were not consistent with the observed changes in metabolic flux (Junker et al., 2007).

More fundamentally, it may be useful to validate aspects of the network itself. Typically the network structure is defined at the outset of the investigation on the basis of the available information, but it is then quite common to add or remove steps during the iterative fitting of the model. For example, in the analysis of the heterotrophic Arabidopsis cell culture, the possibility of synthesizing glycine via threonine aldolase was tested and it was concluded that there was no evidence from the labelling data to justify inclusion of the reaction in the model (Masakapalli et al., 2010). On the other hand, improved fits to the data were obtained by including a step converting serine to pyruvate, but it carried a surprisingly high flux. It was therefore necessary to seek further evidence for the potential contribution of the pathway before including it in the model, and in fact a targeted labelling experiment showed that the direct conversion of serine to pyruvate carried a negligible flux (Fig. 6), showing that the improvement in the MFA flux map was an artefact of the fitting procedure.

This leads to an important practical question that is discussed infrequently in the steady-state MFA literature—the mechanism for choosing between alternative network structures that both provide an adequate description of the labelling data. In principle, as demonstrated in a simple bacterial network, it should be possible to use the $\chi^2$ criterion (Dauner et al., 2001), but in practice this may not be sufficiently reliable even in bacterial systems (Beste et al., 2011). The same problem can arise in plants. For example, in a steady-state MFA study of developing soybean embryos, it was observed that two distinct networks, one including the oxidative pentose phosphate pathway and no contribution from Rubisco and the Calvin cycle, and the other including Rubisco and the Calvin cycle but with no contribution from the oxidative steps in the oxidative pentose phosphate pathway, both provided an adequate explanation of the experimental data (Allen et al., 2009b). Ultimately the decision in favour of the flux map that included the Rubisco flux was based on the better agreement between the measured and predicted values of carbon dioxide production with this network. Similarly, the redistribution of label in a steady-state analysis of Arabidopsis cell cultures could be explained by any one of three models of the subcellular compartmentation of the oxidative pentose phosphate pathway, and it was concluded that the preferred model could only be identified on the basis of other biochemical evidence (Masakapalli et al., 2010).

These observations have an interesting corollary, which is that if a network that explains the labelling data is altered to include an extra pathway—perhaps at the suggestion of a reviewer—and if the modification does not improve the fit, then this in itself is not sufficient to reject the alternative model. In fact there are numerous instances in the literature of statements to the effect that this or that modification of the network did not improve the fit, followed by the conclusion that the original network can be preferred, but in the absence of further biochemical evidence there is really no justification for such a conclusion. Thus significant conclusions about alternative networks can only be drawn...
from steady-state MFA when one network provides a significantly better explanation of the labelling data than the alternative.

Finally, although the flux map is primarily a representation of the interconversion of primary metabolites, it can also provide quantitative insights into processes such as carbon conversion efficiency, ATP utilization, and the supply of reductant for biosynthesis. These processes have been the focus of numerous steady-state MFA investigations in plants, because the flux maps can be used to identify the relative importance of specific pathways to net carbon dioxide production, to compare the maximum potential ATP yield from the network with the demand for ATP implied by the fluxes to biomass, and to assess the relative importance of specific pathways for the supply of reductant. Thus the availability of the flux map allows the investigator to probe the specific contribution of individual steps to processes that would previously have been analysed at a more superficial level. In comparison, and in contrast to the focus of much of the bacterial work, there have been

---

**Fig. 6.** Validation of a feature of a flux map using targeted labelling. (A) Part of a flux map obtained from a steady-state MFA investigation of a heterotrophic *Arabidopsis* cell culture. The thickness of the arrows is proportional to the flux between the metabolites. Including a direct link between serine and pyruvate improved the fit, but the extra step carried an unexpectedly high flux. (B). A high net flux from serine to pyruvate would allow label to be transferred from [2-13C]glycine to alanine, generating a characteristic mixture of isotopomers ([2-13C] and [2, 3-13C]alanine) arising from the combined effect of glycine decarboxylase and serine hydroxymethyltransferase (Hartung and Ratcliffe, 2002). (C) 1H-decoupled 13C NMR spectra recorded at 150.9 MHz from extracts of *Arabidopsis* cells incubated with [2-13C]glycine. The serine signals confirm the expected extensive labelling of C2 and C3, but the alanine signals show only natural abundance signals with no satellite peaks, indicating a negligible flux through the putative pathway from serine to pyruvate. (SKM, unpublished results.)
surprisingly few applications of the method so far to mutant and transgenic lines (Fernie et al., 2001; Alonso et al., 2007b; Lonien and Schwender, 2009).

Concluding remarks

Although by no means routine, robust experimental and computational protocols now exist for the application of steady-state MFA to a range of heterotrophic and mixotrophic plant systems. The flux maps of primary metabolism that have emerged from such studies provide informative metabolic phenotypes that allow the contribution of individual steps and pathways to be assessed within the context of the whole network. This information is useful both in relation to gaining a deeper understanding of how the network supports the growth and development of plants, and for identifying targets for metabolic engineering.

Current methodological developments suggest that the scope and power of the approach may well be set to increase. For example, the power of some of the recently developed software for flux analysis, such as METRAN, OpenFLUX, and FIA, has yet to be harnessed and should increase the rate at which statistically reliable flux maps of plant metabolism can be generated. Similarly, the current development of tandem MS analysis (Choi and Antoniewicz, 2011), capillary electrophoresis MS/MS (Haumuma et al., 2010), and LC-MS methods (Luo et al., 2007; Arrivault et al., 2009; Alonso et al., 2010b) for MFA will facilitate the analysis of pathway intermediates, reducing the dependence of steady-state MFA on measurements of end-products that reach isotopic equilibrium very slowly. Improvements in subcellular fractionation (Krueger et al., 2011) should also lead to more effective and direct analysis of the redistribution of label at isotopic steady state. Current developments in the flux balance analysis (FBA) of genome-scale models of plant metabolism (Sweetlove and Ratcliffe, 2011) are also likely to impact on steady-state MFA, because recent studies have shown that the fluxes predicted by this method can be in good agreement with fluxes determined experimentally (Williams et al., 2010; Hay and Schwender, 2011). This suggests that it would be advantageous to use the in silico FBA method to identify circumstances in which a more time-consuming and demanding analysis by steady-state MFA would be justifiable.

Finally it should be noted that rapid developments are now taking place in the analysis of labelling time-courses (Young et al., 2007; Noack et al., 2010), and that this approach offers several advantages over steady-state MFA (Nöh and Wiechert, 2011). In particular, the duration of the experiment is much shorter, reducing the cost, because it is no longer necessary for the system to reach isotopic steady state; and the approach is especially useful for systems that cannot be maintained in a metabolic steady state for long periods. The analysis of labelling time-courses is also the only feasible approach for the analysis of photautotrophic systems labelled with carbon dioxide (Shastri and Morgan, 2007). Should the analysis of labelling time-courses become routine for determining flux maps of primary metabolism, and it should be noted that the approach presents many computational and analytical challenges that are only just being overcome, then it will eventually replace steady-state MFA for many purposes. Steady-state MFA will then be seen to have been just a necessary stepping stone on the way to a more complete understanding of the metabolic fluxes that underpin cell function.

Supplementary data

Supplementary data are available at JXB online.

Supplementary information. Metabolic flux analysis of a model network.

Figure S1. Outline of the model metabolic network.

Table S1. Fluxes used in analysis of the model network.

Table S2. Summary of fractional abundance of selected metabolite cumomer groups following metabolism of [1-13C]substrate.

Table S3. Summary of fractional abundance of metabolite mass isotopomers following metabolism of [13C]substrates

Acknowledgements

SKM acknowledges financial support from the University of Oxford (Clarendon Scholarship), Exeter College (Mr Krishna Pathak Scholarship), and a UK Overseas Research Student award.

References


from intelligent design to rational engineering.


