Kinetic modelling of plant metabolic pathways

Johann M. Rohwer*
Triple-J Group for Molecular Cell Physiology, Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland, 7602 Stellenbosch, South Africa
* To whom correspondence should be addressed. E-mail: jr@sun.ac.za

Received 7 November 2011; Revised 14 February 2012; Accepted 20 February 2012

Abstract
This paper provides a review of kinetic modelling of plant metabolic pathways as a tool for analysing their control and regulation. An overview of different modelling strategies is presented, starting with those approaches that only require a knowledge of the network stoichiometry; these are referred to as structural. Flux-balance analysis, metabolic flux analysis using isotope labelling, and elementary mode analysis are briefly mentioned as three representative examples. The main focus of this paper, however, is a discussion of kinetic modelling, which requires, in addition to the stoichiometry, a knowledge of the kinetic properties of the constituent pathway enzymes. The different types of kinetic modelling analysis, namely time-course simulation, steady-state analysis, and metabolic control analysis, are explained in some detail. An overview is presented of strategies for obtaining model parameters, as well as software tools available for simulation of such models. The kinetic modelling approach is exemplified with discussion of three models from the general plant physiology literature. With the aid of kinetic modelling it is possible to perform a control analysis of a plant metabolic system, to identify potential targets for biotechnological manipulation, as well as to ascertain the regulatory importance of different enzymes (including isoforms of the same enzyme) in a pathway. Finally, a framework is presented for extending metabolic models to the whole-plant scale by linking biochemical reactions with diffusion and advective flow through the phloem. Future challenges include explicit modelling of subcellular compartments, as well as the integration of kinetic models on the different levels of the cellular and organizational hierarchy.

Key words: Differential equations, enzyme kinetics, flux-balance analysis, metabolic control analysis, metabolic flux analysis, parameterization, steady state, stoichiometry, time-course.

Introduction
The previous century brought huge advances in the fields of biochemistry and molecular biology; much of the cellular map has been elucidated, and the components of life were separated, identified, and their properties characterized. As such, this has been an era of structure and function, resulting in a build-up of a huge archive of knowledge about the components of life. However, most of this information is hugely reductionist in nature and, in the current post-genomic era the challenge has shifted from component identification and characterization towards assembling this disparate information into an integrated view at the molecular, cellular, and organismal level. Since all the genomes, transcriptomes, proteomes, and metabolomes by themselves constitute static information that provides little insight into the dynamic functioning of the living cell, this has led to the emergence and rapid growth of the field of systems biology (e.g. Kitano, 2002), which aims to integrate the information into a ‘systems’ view through a combination of interdisciplinary approaches. Importantly, this problem cannot, and will not, be solved by experimental analysis alone, and ‘omics’ experiments by themselves are not systems biology. Rather, an integrated view attempts to explain the behaviour of the system in terms of the properties of its components.

The rapid development of bioinformatics during the last two decades has facilitated the data analysis that is required for high-throughput experimentation at the genome scale. Most of the computational effort has in the past decade
been directed towards DNA and protein sequence analyses. However, it is becoming widely appreciated that the computational approach to understanding the dynamic aspects of cell processes is essential for understanding them and should therefore form part and parcel of the bioinformatic effort (Kitano, 2005). In addition to the computational systems biology endeavour, this requires additional experimentation, as the analysis has to be extended to the dynamic level, making time a central variable. In fact, it has been suggested that the integration of genetics, molecular biology, and cell biology into the interdisciplinary field of systems biology will enable biology to advance to the next level (Kitano, 2002; Westerhoff and Palsson, 2004).

Classically, when building a kinetic model of a pathway, a common strategy has been to reduce the system as far as possible while still maintaining its pertinent characteristics. As a result, such models have typically addressed questions relating to a part of a real system. Nevertheless, these models have played an important role, both in hypothesis testing about the (sub)system and in manipulating it, for example for industrial applications. Such models are, however, unsuitable for system-wide ‘omics’ data sets, which can quantify a substantial part of the mRNA transcript, protein, or metabolite levels in a particular cell. Such system-wide experimental analyses have become more and more commonplace in recent years, and will require larger and more detailed kinetic models to encapsulate their behaviour. The ultimate goal would be for such models to be replicas of the real system, as is the stated aim of several research projects, notably the Silicon Cell project (http://www.siliconcell.net), but see also E-cell (http://www.e-cell.org) and Physiome (http://www.physiome.org).

This review is organized as follows: the following section focuses on those models requiring only knowledge of the network stoichiometry. The next then introduces kinetic models, their requirements, and the software tools available for their simulation. Metabolic control analysis is also presented as a useful framework for analysing the regulation and control of such systems. The penultimate section shows how kinetic models of plant metabolism can be used to obtain a more complete understanding of the dynamic functioning of plant metabolic pathways, including the contribution of the various steps to the overall control and regulation as well as the identification of potential targets for experimental manipulation in order to achieve a desired function. Finally, concluding remarks and future perspectives are presented.

**Network stoichiometry and structural models**

Whatever the level at which a metabolic network is modelled, this always requires knowledge of the network stoichiometry. The stoichiometry defines the number and nature of all the reactants and products participating in a specific biochemical reaction, and thus describes how the various reactions of the network are connected (Hofmeyr, 2001; Palsson, 2006). To illustrate this, consider the simple reaction scheme of a hypothetical metabolic pathway in Fig. 1.

Similar to the pathway in Hofmeyr (2001), the reaction network has been chosen to be minimal, yet to display the most common features of metabolic pathways, namely a branch point and a moiety-conserved cycle. The latter is characterized by a set of species (S₂ and S₃ in the example of Fig. 1), the sum of whose concentrations remains constant. The metabolites in the cycle typically contain a ‘moiety’ that is conserved amongst them, and the existence of the moiety conservation places additional constraints on the system as the various forms of the moiety need to be balanced and regenerated. In the network of Fig. 1, this means that the fluxes through reactions 1 and 2 must be equal at steady state (see below). Typical examples of conserved moieties in metabolic networks include ATP+ADP (or ATP+ADP+AMP in the presence of an active adenylate kinase), as well as NADH+NAD⁺.

The connections between the reactants and the reactions of the network are described by stoichiometric coefficients, which, for each reaction, state how many molecules of each reactant or product participate in the reaction (for reactants, the stoichiometric coefficient is negative as they are consumed; for products, it is positive). The stoichiometry of the complete system can be conveniently summarized in matrix form (note that the so-called stoichiometric matrix N only refers to the internal metabolites of the network; the rows refer to metabolites and the columns to reactions):

$$
N = \begin{pmatrix}
S_1 & 1 & 2 & 3 & 4 & 5 \\
S_2 & 1 & 0 & -1 & -1 & 0 \\
S_3 & -1 & 1 & 0 & 0 & 0 \\
S_4 & 1 & -1 & 0 & 0 & 0 \\
S_5 & 0 & 0 & 0 & 1 & -1
\end{pmatrix}
$$

(1)

It is sometimes beneficial to include the external metabolites in N as well. This will introduce unbalanced metabolites into the network (i.e. metabolites that are either only consumed or only produced), but in the case of the initial pathway substrate(s) and final pathway product(s) this can be useful to calculate net stoichiometries, for example of elementary modes (see ‘Elementary mode analysis’).

If the species concentrations are now collated in a vector $s=[s_1, ..., s_5]^T$ and the reaction rates in a vector $v=[v_1, ..., v_5]^T$,
then the following set of ordinary differential equations (ODEs) encapsulates the kinetic model of the system,

$$\frac{ds}{dt} = Nv,$$  \hspace{1cm} (2)

with $N$ as in Equation 1. Equation 2 governs the time-dependent flow of matter through the network and forms the basis of all stoichiometric and kinetic models of metabolic pathways (Heinrich and Schuster, 1996; Fell, 1996; Hofmeyr, 2001). If the kinetic properties of the enzymes are unknown, these models are referred to as structural; in this case, Equation 2 is set equal to zero and the vector $v$ is treated as a set of variables, for which the equation is solved, generally only yielding information about flux values or flux ratios and not about concentrations. A number of examples are discussed below. In contrast, when kinetic information is available, the elements of $v$ are treated as functions of the species concentrations in $s$, allowing the time-dependent evolution of the system in terms of both fluxes and metabolite concentrations to be calculated. Such kinetic models are discussed separately later in this review. Note that every kinetic model also implicitly defines a structural model (as $N$ is known), and hence kinetic models can also be subject to structural analysis.

Three types of structural analysis that can be applied to models will now be presented in more detail: flux-balance analysis, metabolic flux analysis, and elementary mode analysis. In general, all structural models are based on the stoichiometric matrix and calculate the solution space of the system of ODEs within steady-state constraints:

$$\frac{ds}{dt} = Nv = 0$$ \hspace{1cm} (3)

Since the main focus of this review is kinetic modelling, only the main principles of these structural modelling approaches will be summarized, giving a few examples and referring the reader to further review literature.

Flux-balance analysis

Flux-balance analysis was pioneered in the laboratory of Palsson (Schilling et al., 1999, 2000), and aims to solve Equation 3 in order to calculate intracellular fluxes from measurement of a limited number of input and output fluxes (such as substrate consumption or product clearance) from the network. Since the overall stoichiometry of cellular processes is generally well known, such models are becoming available on a genome-wide scale, initially for microorganisms (e.g. Edwards and Palsson, 2000), but more recently also for plants (Poolman et al., 2009; de Oliveira Dal’Molin et al., 2010a, b).

The principle of flux-balance analysis relies on experimentally quantifying a set of metabolic fluxes in the network; normally, these are easily determinable fluxes such as substrate consumption or product excretion. These fluxes are then entered into Equation 3, and the equation is solved for the remaining unknown fluxes (Schilling et al., 1999). However, the system is typically under-determined and there is a significant number of degrees of freedom, as there are generally more internal metabolites than quantifiable boundary fluxes. By way of a simple illustration, consider the network in Fig. 1. The steady-state assumption in Equation 3 leads to a number of flux relationships. Using the symbol $J$ to denote steady-state flux (i.e. the reaction rate at steady state), two examples of such flux relationships are $J_3 = J_5$ (as they occur in the same branch), and $J_1 = J_3 + J_4$ (the total flux entering and leaving a branch point must be the same). As a result, not all the fluxes are independent, but a subset (the dependent ones) can be expressed as linear combinations of the independent fluxes. The complete set of flux relationships is obtained from null-space analysis of $N$ (Hofmeyr, 2001) and is given by

$$\mathbf{J} = \mathbf{KJ}_1 \text{ or specifically } \begin{pmatrix} J_1 \\ J_2 \\ J_3 \\ J_4 \\ J_5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} J_1 \\ J_3 \end{pmatrix}$$ \hspace{1cm} (4)

Note that here $J_1$ and $J_3$ have been chosen as independent fluxes, but $J_1$ and $J_5$ or $J_3$ and $J_5$ could equally as well have been chosen. Assuming now that, for example, only $J_1$ is experimentally determinable, this means that $J_3$, $J_4$, and $J_5$ cannot be uniquely determined as there is an internal degree of freedom. In flux-balance analysis, Equation 3 therefore has to be re-formulated as a linear programming problem by introducing an objective function for which the system is assumed to have been optimized, for example maximizing growth rate or the production of a particular metabolite, or minimizing nutrient uptake (Schilling et al., 1999; Reed and Palsson, 2003).

One of the disadvantages of flux-balance analysis is that it is unable to predict quantitatively to what extent certain enzyme activities should be altered to achieve a desired effect, such as an increase in the specific production rate of a product. The reason for this is that these models do not contain kinetic information and are thus only valid for the particular steady-state condition for which they were constructed. This limitation has been partially overcome through hybrid models, which typically incorporate additional constraints other than the network stoichiometry, such as enzyme capacity, thermodynamic information associated with reaction directionality, and biochemical loops (Price et al., 2003).

While the original flux-balance models were restricted to microorganisms due to the relative ease with which their exchange fluxes with the environment can be measured, a number of plant flux-balance models have recently appeared in the literature. These include, amongst others, genome-scale models for Arabidopsis (Poolman et al., 2009; de Oliveira Dal’Molin et al., 2010b; Williams et al., 2010), C4 plants (de Oliveira Dal’Molin et al., 2010a), Brassica (Hay and Schwender, 2011), and maize (Saha et al., 2011). One of the challenges of modelling higher organisms is to
account for subcellular compartmentation in a biologically realistic way (Lunn, 2007), as the data are often sparse. Flux-balance modelling of plant metabolism has recently been thoroughly reviewed, and the reader is referred to Sweetlove and Ratcliffe (2011) for further details.

Two useful extensions to flux-balance analysis have been the incorporation of enzyme and metabolite concentration data (Hoppe et al., 2007; De et al., 2008). This is important as it allows the use of proteomics and metabolomics data sets to build and refine flux-balance models; these data sets are now becoming widely available. For enzymes, constraints incorporating weighting coefficients corresponding to measured enzyme concentrations are formulated and incorporated into the model (De et al., 2008). In the case of metabolites, thermodynamic realizability is incorporated as a constraint to ensure that reaction directionality is always consistent with measured metabolite concentrations and the Gibbs free energy for the reaction (Hoppe et al., 2007). Both these approaches constrain the solution space of Equation 3 by providing additional information.

Importantly, any flux-balance model relies on an accurate reconstruction of the underlying metabolic network. The annotated genome sequence is often used as a starting point, but requires significant amounts of curation and error correction (for a review, see Fell et al., 2010). Other tools include the integration of metabolic databases (Radrich et al., 2010) and the metabolic scope algorithm (Ebenhöh et al., 2004; Handorf et al., 2005), which calculates what metabolites can in principle be produced from a given starting compound or compounds by a specific network, and thus identifies possible gaps or errors in the network.

**Metabolic flux analysis**

The aim of metabolic flux analysis is the same as that of flux-balance analysis, namely to obtain a global picture of intracellular fluxes by solving Equation 3. However, the approach is different: cells are supplemented with substrate that has been specifically labelled (most commonly with $^{13}$C) at specific positions. The cellular metabolism is then allowed to proceed until metabolic as well as isotopic steady state has been reached. The different isotopomers (i.e. molecules with the same chemical structure but differing in isotopic composition) of the various metabolites are quantified by analytical techniques such as nuclear magnetic resonance (NMR) or gas chromatography–mass spectrometry (GC-MS). The flux distributions are then derived from the isotopomer distributions by calculation and numerical analysis.

The approach has been pioneered by the groups of Wiechert (Wiechert and de Graaf, 1997; Wiechert et al., 1997; Wiechert, 2001) and Sauer (Dauner and Sauer, 2000; Fischer and Sauer, 2003), again focusing on microorganisms. With recent advances in analytical techniques as well as computational power, it has become possible to determine quantitatively the majority of the fluxes in central metabolism, also termed the ‘fluxome’ (Wittmann, 2007). Network flux analysis has, however, not been limited to microorganisms and has been applied to plant cells since the early 2000s; Roscher et al. (2000) provide an overview of strategies that can be employed to apply isotope labelling to the analysis of fluxes in plant metabolism. A number of reviews of plant metabolic flux analysis have appeared in the literature (Morgan and Rhodes, 2002; Kruger et al., 2003; Liboureul and Shachar-Hill, 2008; Kruger and Ratcliffe, 2009); the reader is referred to these papers for further details.

An extension to metabolic flux analysis needs to be mentioned, though, namely the modification of the approach to handle non-steady-state conditions (Baxter et al., 2007; Wahl et al., 2008). The advantage of this is that the experimental system does not have to be at steady state, a condition that is sometimes difficult to achieve, especially isotopic steady state. In turn, non-steady-state conditions require some computational modifications in the data analysis, and they suffer from the limitation that multiple pools of the same metabolite that get labelled at different rates cannot be accounted for, and the method does not work for carbon skeleton rearrangements (Baxter et al., 2007). However, when an experiment is carried out under transient metabolic conditions, $^{13}$C labelling has the potential to estimate kinetic parameters in addition to fluxes (Wahl et al., 2008).

As an illustrative example, metabolic flux analysis in combination with kinetic analysis has exposed a number of different flux distributions in the tricarboxylic acid cycle in plants (reviewed by Sweetlove et al., 2010). Many of these steady-state flux distributions are non-cyclic, demonstrating further functions of the pathway in addition to the cyclic oxidative mode that is normally taught in undergraduate biochemistry. Labelling studies have confirmed that these alternative fluxes actually operate in vivo, and have elucidated conditions under which flux distribution switches can occur (Sweetlove et al., 2010).

**Elementary mode analysis**

A third method of structural modelling is elementary mode analysis. An elementary flux mode is defined as a non-decomposable set of reactions that can carry a steady-state flux in a metabolic system (Schuster and Hilgetag, 1994; Schuster et al., 2000). The criterion of non-decomposability is important, as it ensures that an elementary flux mode cannot be written as a linear combination of two or more simpler elementary modes. In a metabolic network, any observed steady-state flux distribution can be written as a linear combination of different elementary flux modes, although this linear combination is not necessarily unique. However, the full set of elementary modes is uniquely defined for every metabolic system. A mathematical algorithm for detecting the elementary modes in a network has been developed (Schuster and Hilgetag, 1994), extended to the analysis of large networks (Klamt et al., 2005), and implemented computationally in the METATOOL program (Pfeiffer et al., 1999; von Kamp and Schuster, 2006).
To illustrate the concept, consider the network in Fig. 1. Assuming reactions 1, 2, and 4 to be reversible, and reactions 3 and 5 to be irreversible, the network has two elementary modes:

\[ J_1 - J_2 - J_3 \]
\[ J_1 - J_2 - J_4 - J_5 \]

The first mode represents flux through the upper branch, while the second one represents flux through the lower branch. The moiety-conserved cycle carries a flux in both elementary modes.

Elementary mode analysis has the following applications.

(i) It will show up all the substrate cycles or loops (sometimes also termed ‘futile cycles’) in the network. Such cycles can be driven (e.g. by the hydrolysis of ATP) or undriven. A driven cycle may appear to represent inefficiency, but this cannot be stated generally without further physiological analysis (see also the discussion on sucrose metabolism later in this review). Internal undriven cycles, in contrast, while readily identified from stoichiometric analysis, are not thought to be of biological significance as they carry no net steady-state flux.

(ii) All possible metabolic routes from a given pathway substrate to a metabolic end-product are enumerated.

(iii) Every elementary mode has an overall stoichiometry and therefore provides a yield of product molecules formed per substrate molecule utilized. Since the overall set of elementary modes is finite, it is possible to determine which metabolic route produces a particular product with the highest molar yield, and to distinguish higher yielding from lower yielding pathways, if these exist.

(iv) The definition of metabolic pathways in biochemical textbooks is sometimes ambiguous, with the start and end points not always uniquely defined. Here, elementary flux modes, and the related concept of extreme pathways, can be used to define a ‘metabolic pathway’ in a mathematically rigorous way (for a review, see Papin et al., 2003).

By way of example, elementary mode analysis was used to study central carbon metabolism and sucrose accumulation in sugarcane (Rohwer and Botha, 2001). As the metabolic network of sucrose synthesis and breakdown is highly branched, elementary modes were calculated to delineate all possible pathway routes. There are 14 elementary modes in total, which fall into three categories: those flux modes leading from extracellular glucose or fructose to vacuolar sucrose accumulation, those leading from extracellular glucose or fructose to glycolysis and respiration, and, finally, a set of five futile cycles (Rohwer and Botha, 2001). Futile cycling of sucrose (i.e. a continuous cycle of sucrose synthesis and breakdown) is energetically wasteful due to the consumption of ATP in the hexose kinase reactions. The breakdown and re-synthesis of sucrose had been established experimentally (Komor, 1994; Whittaker and Botha, 1997; Zhu et al., 1997), and it was postulated that this could negatively affect sucrose accumulation, or, conversely, that decreasing the level of futile cycling should lead to an increase in sucrose accumulation. This formed the basis for the construction of a detailed kinetic model of the pathways involved (Rohwer and Botha, 2001), and is discussed further later in this review.

Structural models: concluding remarks

Structural models have the advantage that they do not require kinetic information; however, as discussed above, this also limits their applicability. In contrast, dynamic models also contain the kinetic properties of each step of the reaction network in addition to the structure of the system and the stoichiometric constraints. This allows kinetic models to predict the actual values of individual fluxes and not merely their ratios, as well as intermediate concentrations. Since the main focus of this review is kinetic modelling, the next section is devoted to summarizing the general principles of this approach, as well as of the related framework of metabolic control analysis, which determines the control structure of a pathway and can be used as a tool for interrogating kinetic models.

Kinetic modelling of metabolism

The simulation of plant metabolic pathways with enzyme kinetic models (Schallau and Junker, 2010), as well as general systems approaches for modelling pathways (Pfau et al., 2011), have recently been reviewed. Rather than repeating this detail, this section will summarize the most important concepts and attempt to provide an introduction to the topic to the non-expert reader. Enzyme kinetics, the defining feature of these models, is first discussed. The types of analyses that can be performed with such models, as well as strategies for their construction and validation, are subsequently surveyed, followed by a brief listing of available modelling software and a summary of the type of information that can be gleaned from such models. An overview of metabolic control analysis concludes this section.

Enzyme kinetics

The stoichiometry of metabolic networks as a component of all metabolic models has been discussed above. If such a model is to simulate dynamical behaviour, it will have additionally to include information on how each of the constituent reaction rates depends on the concentrations of substrates, products, or effectors that interact with it. This is the domain of enzyme kinetics, which describes the above dependence with a mathematical rate law for each reaction.

An enzyme kinetic rate law is a function of one or more reactant concentrations and parameters that specifies how fast the reaction will proceed. To calculate the dynamics of the set of ODEs in Equation 2, such a rate law is needed for every enzyme-catalysed reaction step. An example is the
reversible Michaelis–Menten equation, which, for step 4 in the pathway of Fig. 1, reads

\[ v_4 = \frac{V_f}{K_{eq}} \left( \frac{S_1 - S_4}{K_m} \right) \]

where \( S_1 \) and \( S_4 \) are the concentrations of \( S_1 \) and \( S_4 \), \( V_f \) is the limiting rate (maximal activity) of enzyme 4 in the forward direction, \( K_{eq} \) is the equilibrium constant of the reaction, and \( K_1 \) and \( K_4 \) are the Michaelis constants for \( S_1 \) and \( S_4 \), respectively.

Enzyme kinetics is a vast field and it is impossible to cover all the aspects in detail here. Cornish-Bowden (2004) is a good introductory textbook, Segel (1993) is an encyclopaedic reference. It should be noted, though, that the development of computational systems biology has shifted the emphasis from detailed enzyme mechanistic studies towards an enzyme kinetics for modelling. This means that rather than providing information about the essential detail of a reaction (including kinetic and thermodynamic properties) to describe its behaviour (in terms of substrate, product, and effector dependence) adequately in models. Concomitantly, there has been a shift to the use of generic rate equations (Liebermeister and Klipp, 2006; Rohwer et al., 2006, 2007), which can describe the overall behaviour of enzyme-catalysed reactions as accurately as detailed mechanistic equations, but require fewer parameters. For example, for reaction 2 in Fig. 1 (\( S_3 + X_7 \rightarrow S_2 + X_8 \)), such an equation for a non-cooperative enzyme exhibiting normal hyperbolic Michaelis–Menten–type kinetics would be given by

\[ v = \frac{V_f}{K_{eq} \cdot K_i} \left( \frac{s \cdot x}{K_m + s \cdot x} \right) \]

with \( V_f \) and \( K_{eq} \) as defined previously, and \( K_i \) the half-saturation constant of reactant \( i \). In contrast, detailed mechanistic equations for the same reaction would differ for different enzyme mechanisms (ordered versus random versus ping-pong bi–bi), and moreover contain additional parameters such as inhibition constants for the various substrates and products (for details, refer to Segel, 1993; Cornish-Bowden, 2004).

In the case of more complicated mechanisms, appropriate terms in the generic rate equations can account for allostery and cooperativity (Rohwer et al., 2007). Moreover, such generic rate equations allow one to differentiate easily between thermodynamic and kinetic control of reaction rate, near-, and far-from-equilibrium reactions, and the contribution of rate capacity, mass action, and enzyme binding to the overall regulation of enzyme activity (Rohwer and Hofmeyr, 2010).

If transcription and translation are explicitly included in a dynamic model, then enzyme concentrations become variables and their differential equations are added to the system of equations (the limiting rate term \( V_f \) is actually the product of an enzyme concentration and a catalytic rate constant). Covalent modification of enzymes, for example phosphorylation, can be accounted for by treating each phosphorylated enzyme species separately as a variable. Genetic regulation and signal transduction networks can also be accounted for in this way. Hofmeyr and Westerhoff (2001) give a generalized framework for modelling such multilevel hierarchical networks and performing a control analysis on them (see ‘Metabolic control analysis’).

Types of analysis

Once a kinetic model has been assembled, the most common analysis types are time-course simulations and solving for the steady state.

Time-course simulations

Due to the highly non-linear nature of enzyme kinetic rate laws, it is generally impossible to integrate systems of ODEs as in Equation 2 analytically, and the solutions are typically obtained by numerical integration. Figure 2 provides a typical result when this approach is applied to the example pathway of Fig. 1. Note how the trajectories for both the concentrations and reaction rates reach a plateau where the variables no longer change with time; at this point the steady state has been reached.

Steady-state analysis

Metabolic networks, being open systems, frequently evolve to a (quasi) steady state. As shown by the latter quarter of the time-course in Fig. 2, such a steady state is characterized by species concentrations that remain constant with time, but at finite, constant reaction rates in the system. Each of the metabolites is produced at the same rate at which it is consumed. Note that at this point Equation 3 is satisfied (i.e. \( \text{d}N_v = 0 \)), and structural and kinetic analyses are thus fundamentally interconnected. A certain physiological condition can be characterized by a particular steady state, and physiological changes will result in a transient state as the system relaxes from the original to a new steady state.

Computationally, the steady state of a metabolic model can be obtained by numerically solving Equation 3 (as for the integration, analytical solutions can seldom be obtained due to the non-linear nature of the equations). A number of algorithms are available for solving such systems of equations, and the computational effort involved is considerably less than numerical integration of Equation 2. Consequently, time-course simulations are frequently skipped when the primary interest lies in steady-state data (see ‘Modelling sucrose metabolism in sugarcane’ for an example). In addition, the framework of metabolic control analysis (see below) is built on the assumption of a steady state.

Strategies for model building

This section briefly describes the process of assembling and validating a kinetic model of a metabolic pathway.
A kinetic model of a reaction network requires five types of data:

**Stoichiometry**

The stoichiometry of the network (N in Equation 2) must, of course, be known. In general, this is not a problem as metabolism is well charted and genome-scale metabolic networks have recently become available for an increasing number of organisms (Palsson, 2006), including plants (Poolman et al., 2009; de Oliveira Dal’Molin et al., 2010a, b). This has been discussed above.

**Enzyme kinetic data**

These include the enzyme kinetic rate laws and their associated Michaelis, inhibition, activation, and other constants. Previously, collecting these data from the biochemical literature was an arduous task, but a number of searchable and curated internet databases such as BRENDA (http://www.brenda-enzymes.org; Schomburg et al., 2004) and SABIO-RK (http://sabio.villa-bosch.de; Krebs et al., 2007) greatly facilitate this process. In addition, when kinetic data for a particular enzyme from a certain organism are unavailable, data from a closely related species can often be substituted.

Ideally, enzyme kinetic parameters should be determined under conditions that resemble the situation in vivo. These are most often not the conditions under which enzyme kinetic assays have historically been performed, since enzymes tended to be investigated for their kinetic mechanism under conditions for optimal activity in terms of pH, ionic strength, etc. To facilitate standardization and to integrate data from different laboratories into systems biology models, van Eunen et al. (2010) developed a set of standard conditions for measuring enzyme activities and performing kinetic assays in yeast. This was achieved in a multilaboratory collaboration by determining the cytosolic pH and concentrations of calcium, sodium, potassium, phosphorus, sulphur, and magnesium, and using these and literature data to formulate an experimentally feasible in vivo-like assay buffer. It is to be expected that similar in vivo-like media will be developed for other organisms. In addition, under the auspices of the Beilstein Institut in Germany, the STRENDA Commission (http://www.beilstein-institut.de/en/projects/strenda) has been set up to develop standardized guidelines for reporting enzyme data and to set up a searchable database of these data for easy retrieval (Apweiler et al., 2010).

**Thermodynamic data**

All enzyme-catalysed reactions are reversible in principle, and the degree of reversibility has to be adequately captured in kinetic models. Cornish-Bowden and Cárdenas (2001) investigated the effect of irreversible steps in computer models and concluded that these can have large effects on model behaviour if there is no way for downstream metabolites in a pathway to ‘communicate’ with upstream enzymes. Therefore, thermodynamic information should be included in kinetic models wherever possible. Fortunately, huge advances have been made in terms of collating equilibrium constant data (or the related standard Gibbs free energy of reaction) in the Thermodynamics of Enzyme-Catalyzed Reactions database (http://xpdb.nist.gov/enzyme_thermodynamics; Goldberg et al., 2004) and calculation of standard free energy changes for biochemical reactions from first principles (Alberty, 2006). An alternative approach is the group contribution method (Jankowski et al., 2008), which uses the Gibbs free energy contribution values of 74 distinct molecular substructures and 11 interaction
factors to estimate standard Gibbs free energies of lesser studied biochemical reactions.

Maximal enzyme activity data

The limiting rate enters an enzyme kinetic rate law either directly or as a product of enzyme concentration and $k_{\text{cat}}$, the catalytic rate constant. Since gene expression (and, as a consequence, the enzyme level) depends on the physiological conditions, these data have to be measured directly from the biological sample for which the model is to be built, and are seldom transferable from one laboratory to the next. Moreover, enzyme activity levels are tissue and organism specific. Recently, approaches have been developed to measure multiple enzyme activities on a large scale with a robotics platform (Gibon et al., 2004a), and a good correlation has been observed ($R^2=0.592$) between protein levels estimated from experimentally determined enzyme activities and published $k_{\text{cat}}$ values on the one hand, and quantitative proteomics using the emPAI protein abundance index on the other hand (Piques et al., 2009). Overall, it is now experimentally possible to obtain these data on a system-wide level.

Other model parameters

These include the concentrations of the ‘terminal’ source and sink metabolites, which are entered as fixed parameters into the model, the sums of the moiety-conserved species (e.g. NAD$^+$/NADH or ATP+ADP), and the initial concentrations of all the model variables (the latter are particularly pertinent if a time simulation is performed but irrelevant for steady-state analysis). These data can be obtained from metabolomics experiments or directed assays and, importantly, are also becoming available at the level of subcellular compartments (Krueger et al., 2011).

Once these data are available, the model is assembled and the system of ODEs is integrated (to obtain a time-course) or solved to calculate the steady state. In order to assess the quality of the model, it has to be validated. First, it is essential to check the model for internal consistency, for example by ensuring that no reactions can carry a steady-state flux when no external metabolites are present. In a subsequent step, the model output is typically compared with independent data which have not been used in the construction of the model. For example, such data may include measurements of fluxes or metabolite concentrations of the intact pathway in the biological sample under study. Validation data are thus directly comparable with model outputs, and are typically ‘system’ data, in contrast to the data for model construction which are usually mechanistic properties of the isolated components. This strategy forms the basis of the so-called ‘silicon cell’ approach (Snoep et al., 2006).

Sometimes the bottom-up approach to model building, as outlined above, is not feasible because insufficient data are available to determine all the model parameters. In such cases, model parameters may be ‘fitted’ in a top-down approach by adjusting them iteratively in an optimization routine until the model output (fluxes or concentrations) matches some experimentally determined value (see also the benzenoid network model described in ‘Top-down model assembly’ below). This strategy has its pitfalls, since the parameters—now being the output from a global optimization routine—are no longer mechanistically derived from enzyme properties. The problem is exacerbated when empirical rate equations are used, the parameters of which have no mechanistic interpretation. Also, the distinction between model construction and model validation is blurred, since the same type of data is used for both. Further, these models can be more limited in their applicability, since they may fail to describe physiological conditions removed from those that were used in their construction. Therefore, validation is especially important in these cases; this can be achieved, for example, by using data from a different physiological state. Note, however, that the distinction between model construction and validation is not always rigorously maintained in the modelling literature.

The topic of parameter estimation and experimental design in systems biology is a complex field, which there is not room to explore fully here. However, for further information the reader is referred to an excellent recent minireview series, specifically dedicated to this theme (Ashyraliyev et al., 2009; Bruggeman, 2009; Cedersund and Roll, 2009; Kreutz and Timmer, 2009).

Software for kinetic modelling

The large variety of problems that have been addressed with metabolic modelling has led to the development of an equally impressive list of simulation software. Different problems require different tools, yet there are certain elements of a model description (e.g. the stoichiometry or the kinetic equations) that are common to most, if not all, of the models. To facilitate interchange of models between simulation tools, and to provide a standard, the Systems Biology Markup Language (SBML, see http://sbml.org), an XML dialect for defining systems biology models, was developed by the international systems biology community (Hucka et al., 2003). As of November 2011, the SBML website lists 230 software programs that implement and support this standard. In theory, a model created in one software tool can then be exported as SBML and imported in another software. SBML is being actively developed and, while it has gone a long way towards providing standards and interoperability, in practice there are still some difficulties in exchanging models between programs, not least due to the fact that different programs support different analysis types.

Generically, any system of coupled ODEs such as in Equation 2 can of course be solved with standard numerical tools such as Mathematica (http://www.wolfram.com) or Matlab (http://www.mathworks.com/products/matlab), or their open-source alternatives Octave (http://www.gnu.org/software/octave) and SciPy (http://www.scipy.org). Specific add-ons or toolboxes are available that simplify the analysis
of SBML models in Mathematica (MathSBML) or Matlab (SBMLToolbox). There are, however, also a number of programs available that are dedicated to modelling such systems. In terms of these, a broad distinction can be made between console- or command-line-based tools on the one hand, and graphical user interface (GUI)-based tools on the other. While GUI tools tend to be easier to learn for the novice, console-based tools tend to be more versatile in terms of scripting complex analysis tasks. The overall functionality of many of these tools is equivalent, and the choice is often based on personal preference. A popular program from the GUI category is COPASI (Hoops et al., 2006); console-based programs include Jarnac, part of the Systems Biology Workbench (Sauro et al., 2003), and PySCeS (Olivier et al., 2005), which was developed in the author’s group.

Applications of models

So what is the use of a metabolic kinetic model and what types of questions can be answered with it? Once a kinetic model for a specific pathway has been developed and validated (see above), it can be used to calculate and predict pathway behaviour under conditions that have not yet been tested experimentally. While there are limitations in this approach (any model is a simplification of the real system and as such will never perfectly match the behaviour of the system it approximates), this can nevertheless be useful, as it could suggest new experiments that would test the model predictions. Simulations can be performed much more rapidly than experiments in the laboratory, especially if genetic manipulation would be involved in creating mutant strains with altered gene expression of specific pathway enzymes. Hence, when looking to manipulate a pathway with a specific outcome in mind, a model may explore different strategies and evaluate their efficacy, thus suggesting a directed experimental approach. This may be more effective than purely intuitive or random experimental manipulation. An example of this is provided by the kinetic modelling of sucrose futile cycling in sugarcane (Rohwer and Botha, 2001), which is discussed in greater detail later in this review.

In addition, a kinetic model can be used to explore the control structure of a pathway with metabolic control analysis. This can identify key control points that are most important for determining the value of specific fluxes or intermediate concentrations in the pathway. The framework of metabolic control analysis is introduced in the next section.

Metabolic control analysis

One of the advantages of kinetic modelling is that ‘What if?’ questions can be answered, for example, by changing parameters to values outside the range that were considered in an experiment. To be useful, a model should offer accurate predictions of pathway behaviour or be able to encompass a variety of experimental observations or conditions with a single model description. To assess model performance, methods to interrogate models are necessary, and an important such method is metabolic control analysis (MCA). The steady state of a system of coupled reactions has already been introduced above. When the steady-state assumption holds, the distribution of control amongst all the reactions in the system can be quantified using MCA. This framework was independently developed by Kacser and Burns (1973) and Heinrich and Rapoport (1974), and is in essence a sensitivity analysis that quantifies the dependence of metabolic variables (fluxes and metabolite concentrations) on the activities of the individual reactions in the system. These sensitivities are quantified with so-called control coefficients and elasticity coefficients.

The coefficients of MCA

An elasticity coefficient (or elasticity for short) is a local property of a particular enzyme-catalysed reaction and quantifies how the rate of this reaction depends on the concentration of substrate, product, or any external effector that affects its rate. It is equivalent to an apparent kinetic reaction order (reviewed in Kacser et al., 1995; Fell, 1996; Heinrich and Schuster, 1996). By way of example, in Fig. 1 the elasticity of reaction 4 to S1 is defined as

$$e_{v4, s1} = \left( \frac{\partial \ln v_4}{\partial \ln s_1} \right)_{s_2 ... s_5}$$  \hspace{1cm} (7)

where $v_4$ is the rate of reaction 4 and $s_1$ is the concentration of $S_1$. The subscript $s_2 ... s_5$ signifies that the other species are kept constant at their prevailing steady-state values.

In contrast to elasticities, control coefficients describe the change in a systemic variable as a result of a perturbation of a local reaction rate. While an elasticity is a property of an individual enzyme or reaction, a control coefficient is a function of the entire system and its value depends on all the system components. A control coefficient is defined as the fractional (i.e. relative) change in a system variable (steady-state flux or metabolite concentration) divided by the fractional change in an enzyme activity (Kacser et al., 1995; Fell, 1996; Heinrich and Schuster, 1996). Again with reference to Fig. 1, the flux control coefficient of step 2 on $J_5$ would be given by

$$C_{v5, s2} = \left( \frac{\partial \ln J_5}{\partial \ln v_2} \right)_{ss}$$  \hspace{1cm} (8)

where $J_5$ is the steady-state flux through reaction 5 and $v_2$ is the local activity of enzymatic step 2 that is perturbed. The subscript $ss$ signifies that the entire system is allowed to relax to a new steady state after the perturbation in $v_2$. Concentration control coefficients are defined similarly by replacing the flux ($J$) with a steady-state concentration ($s$) in Equation 8.

Summation theorems

There are two key corollaries to the framework of MCA. The first of these states that systemic properties (i.e. control coefficients) can be calculated from local properties (i.e.
elasticities) and vice versa (for the mathematical details, consult Fell, 1996). In this sense, MCA is a true ‘systems biological’ framework, since it quantifies the dependence of systemic behaviour on all of the components in the system.

The second corollary concerns the summation theorems which state that the control coefficients of all reactions on a particular flux add up to 1, and for a particular metabolite they add up to 0. The flux control summation theorem has dispelled the dogma of the so-called ‘rate-limiting step’ (which would imply that a single step has a flux control coefficient of 1 while those of all the other steps are 0). While such a scenario is possible in principle, the framework of MCA shows that flux control can be, and has been experimentally shown to be (see, for example, Groen et al., 1982), shared by steps in a pathway. Nevertheless, amongst non-specialists there is still widespread ignorance of this fundamental network property (e.g. Ishikawa et al., 2011).

Determining control coefficients

As has been thoroughly reviewed in Heinrich and Schuster (1996) and Fell (1996), control coefficients of metabolic pathways can be determined in various ways. In a directed experimental approach, the cellular level of an enzyme can be altered by molecular biological means, for example by cloning its gene behind a synthetic inducible promoter, or through techniques such as RNA interference. Alternatively, the system can be titrated with an inhibitor specific for a particular enzymatic step. By concomitantly measuring the steady-state variables (fluxes or metabolite concentrations), the control coefficient can be calculated directly from Equation 8.

An alternative method involves the determination of elasticities in the metabolic system. These can either be determined directly in enzyme assays, or calculated from differentiation of the enzymatic rate law (if known) and substitution of measured substrate and product concentrations. The control coefficients can then be calculated from the numerical elasticity values.

Finally, control coefficients can be calculated from models. If a properly validated kinetic model (see ‘Strategies for model building’) is available for a pathway, this enables direct computation of the control coefficients. The limitation of this method is that the control coefficients are of course those of the model, and the accuracy of the values obtained depends on how well the model describes the actual experimental system. This approach was used to determine the control of substrate cycling in sugarcane as outlined later.

This section has provided an introduction to the concepts of kinetic modelling and metabolic control analysis. The next section will briefly review some plant kinetic models from the literature.

Studying plant physiology with kinetic models

Various aspects of plant physiology have been analysed extensively with kinetic models, and the field has a history going back for more than two decades. Detailed surveys of these models, the pathways they are addressing, and the techniques used can be found in Morgan and Rhodes (2002) and Rios-Estepa and Lange (2007). A recent review (Arnold and Nikoloski, 2011) was devoted to a quantitative comparison and ranking of all the kinetic models that have been published on the Calvin–Benson cycle. This information will not be repeated here; instead, the purpose of this section is 2-fold. First, a summary of recent plant kinetic models that have been published since the last comprehensive review by Rios-Estepa and Lange (2007) is presented in Table 1. Importantly, the scope of these models is not limited to primary metabolism, but extends to secondary metabolism, the metabolism of xenobiotics, as well as gene-regulatory networks.

Secondly, three examples of plant kinetic models (one from the author’s own work) are discussed in greater detail to illustrate specific aspects of the modelling techniques and approaches.

Top-down model assembly

As outlined above, one of the problems associated with building kinetic models is that often the kinetic data for the constituent enzymes are not available. In such cases, kinetic parameters can be obtained from a ‘top-down’ approach by fitting the model parameters to experimental data from the intact system, for example fluxes and metabolite concentrations.

This ‘top-down’ method was followed by Colón et al. (2010) to develop a model of the benzenoid network in petunia flowers. The model comprises a network of 31 biochemical reactions describing the conversion of phenylalanine to various benzenoid volatiles. The experimental data used for fitting were metabolite pool sizes and labelling patterns obtained by feeding three different concentrations of deuterated phenylalanine to the flowers. Independent validation data were provided by transgenic flowers in which one of the pathway enzymes, BPBT, was down-regulated using RNA interference. The model was subsequently subject to MCA, showing that the enzyme phenylacetalddehyde synthase exerted the bulk of the control on the phenylacetalddehyde branch of the network. However, in other branches flux control was widely distributed. The significance of the results in this study is 2-fold: first, construction and assembly of the model enabled the authors to perform MCA with the model. This would not have been experimentally feasible for every step in the pathway. Secondly, the MCA was able to identify key flux-controlling steps and thus suggested possible future metabolic engineering strategies for perturbing secondary plant metabolism by targeting steps with large flux control coefficients, for example with a view to boosting secondary metabolite production.

Bottom-up model assembly

In contrast to the previous example, the ‘bottom-up’ approach entails assembling all the available data on the
isolated pathway components (such as enzyme kinetic parameters, etc., see above) into a model, and then appraising how well this model replicates the behaviour of the entire system. Curien et al. (2009) followed this strategy to build a detailed kinetic model of aspartate biosynthesis in Arabidopsis thaliana. The authors performed in vitro kinetic measurements on the constituent enzymes of the pathway, and then compiled and collated these into a detailed kinetic model, which could reproduce in vivo experimental measurements. The strength of this approach is that the authors could demonstrate that in vivo behaviour of the pathway can actually be explained in terms of independently collected in vitro data. The modelling results are significant for their explanatory power in identifying and clarifying the role of allosteric interactions, of which there are a great number in this pathway. The model identified some of these allosteric feedbacks whose function is not to couple supply and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case), but rather to ensure that fluxes in competing pathways and demand for an intermediate (as is commonly the case)

Finally, the model identified threonine as a potential high-level regulator as its concentration was the most sensitive variable in the system.

Modelling sucrose metabolism in sugarcane

In the remainder of this section, a brief overview of the author’s own work on the kinetic modelling of sugarcane will be provided. As explained earlier in this review, elementary mode analysis identified a number of substrate cycles in sugarcane metabolism; this concurrent sucrose breakdown and re-synthesis has also been demonstrated experimentally (Komor, 1994; Whittaker and Botha, 1997; Zhu et al., 1997). This prompted the construction of a kinetic model of the substrate cycling process, together with sucrose accumulation in the vacuole. The model, constructed with the ‘bottom-up’ approach, could replicate independent experimental flux and metabolite concentration validation data without resorting to parameter fitting (in no case was the discrepancy greater than a factor of two; Rohwer and Botha, 2001). Based on MCA, the control of each reaction on the substrate cycling of sucrose (defined as the control of the flux ratio between sucrose breakdown and sucrose accumulation in the vacuole) could be quantified. The five reactions with the numerically largest control coefficients were the uptake of fructose and glucose into the cytoplasm (–0.86 and –0.90, respectively), the transport of sucrose into the vacuole (–0.51), phosphorylation of glucose

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Comment</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essential oil biosynthesis</td>
<td>Monoterpenoid biosynthesis in peppermint; identified controlling enzymes for essential oil composition</td>
<td>Rios-Estepa et al. (2008, 2010)</td>
</tr>
<tr>
<td>Benzenoid network</td>
<td>Discussed in main text</td>
<td>Colon et al. (2010)</td>
</tr>
<tr>
<td>Aspartate metabolism</td>
<td>Discussed in main text</td>
<td>Curien et al. (2009)</td>
</tr>
<tr>
<td>Flavonoid pathway</td>
<td>Minimal kinetic model of temperature compensation and regulation of the pathway in Arabidopsis; temperature effects modelled with Arrhenius equation</td>
<td>Olsen et al. (2009)</td>
</tr>
<tr>
<td>Glucosinolate metabolism</td>
<td>Kinetic model of multifunctional enzymes in glucosinolate biosynthetic pathway incorporating measured enzymatic properties; model predicts different glucosinolate chain length profiles</td>
<td>Knoke et al. (2009)</td>
</tr>
<tr>
<td>Fenclorim metabolism</td>
<td>Modelled metabolism of the herbicide safener fenclorim (which enhances gene expression of detoxifying enzymes such as glutathione transferases) in Arabidopsis; parameters estimated by fitting to time-courses</td>
<td>Liu et al. (2009)</td>
</tr>
<tr>
<td>Photosystem II</td>
<td>Chlorophyll a fluorescence transients in pea leaves modelled with eight models with different reaction schemes but comprising the same electron carriers</td>
<td>Lazár and Jablonský (2009)</td>
</tr>
<tr>
<td>Plastoquinone kinetics</td>
<td>Three-state-variable model of plastoquinone-related electron transport kinetics in photosynthesis, describing delayed fluorescence; model validated with data from healthy and drought-stressed soybean plants</td>
<td>Guo and Tan (2009)</td>
</tr>
<tr>
<td>Hydrogen production</td>
<td>Bioreactor model of transition from oxygenic growth to anoxic H2 production in Chlamydomonas reinhardtii; dependence on light and sulphur availability</td>
<td>Fouchard et al. (2009)</td>
</tr>
<tr>
<td>Sucrose metabolism</td>
<td>Discussed in main text</td>
<td>Uys et al. (2007)</td>
</tr>
<tr>
<td>Circadian clock</td>
<td>Three-feedback-loop model of the plant clock gene-regulatory network in Arabidopsis incorporating data on experimentally established feedbacks between regulatory proteins and genes; the model predicts coupled morning and evening oscillators</td>
<td>Locke et al. (2006)</td>
</tr>
<tr>
<td>RuBisCO activation kinetics</td>
<td>Detailed kinetic model of the elementary reaction steps in the mechanism of RuBisCO from spinach leaves; rate constants of carbamylation, activation, carboxylation, and inhibition determined by fitting to multiple experimental data sets</td>
<td>McNevin et al. (2006)</td>
</tr>
</tbody>
</table>
by hexokinase (1.09), and breakdown of sucrose by neutral invertase (0.71). Rohwer and Botha (2001) found that a decrease in substrate cycling is predicted to translate into increased sucrose accumulation, and on the basis of these results predicted that overexpression of the plasma membrane glucose and fructose transporters, as well as of the vacuolar sucrose importer, and attenuation of neutral invertase would be the most promising biotechnological targets for reducing this futile cycling of sucrose and increasing sucrose accumulation. By way of experimental validation, Rossouw et al. (2007) demonstrated an increase in sucrose accumulation in sugarcane suspension cells by decreasing neutral invertase activity through RNA interference, albeit at the expense of reduced respiration and growth. Thus, while the increase in sucrose accumulation was correctly predicted by the model, its scope did not extend far enough to predict the effect of changes in neutral invertase activity on respiration and growth, perhaps suggesting that the cycling of sucrose is not ‘futile’ after all but fulfils another—as yet unidentified—function. More recently, these results were repeated in transgenic sugarcane plants, with Rossouw et al. (2010) demonstrating increased sucrose and decreased hexose levels, as well as reduced substrate cycling in the culm tissues of transgenics with reduced neutral invertase activity.

### Growth modelling of a sugarcane stalk in segments

The model of Rohwer and Botha (2001) is specific to medium-mature tissue of a single internode. However, the expression of metabolic enzymes changes significantly with growth, stem elongation, and internodal maturation (Botha et al., 1996); these effects are most pronounced during the first 10 internodes of the sugarcane stalk, and higher numbered internodes can be regarded as mature. This prompted the extension of the original sugarcane model to include more detail about culm biochemistry. Specifically, the isoforms of sucrose synthase and fructokinase were made explicit, carbon partitioning to respiration and fibre formation were treated separately, and partitioning to respiration occurred via either phosphofructokinase (PFK) or pyrophosphate-dependent PFK, on to aldolase and lower glycolysis (Uys et al., 2007). The stoichiometry of this extended network is shown in Fig. 3.

Furthermore, a consistent set of enzyme maximal activity data (Botha et al., 1996) measured in different internodes of the same plant allowed this model to simulate internodes 3–10 of a sugarcane stalk merely by substituting the appropriate internode-specific enzyme activities as model parameters. The stalk was thus modelled ‘in segments’ (Uys et al., 2007).

The salient features of this model are that as the internodes mature, there is a decreased carbon partitioning to glycolysis and respiration, as well as to fibre formation, in agreement with experimental data (Fig. 4). As discussed in Uys et al. (2007), the discrepancy between model and experiment for the carbon partitioning to glycolysis and respiration may perhaps be explained by the presence of an additional sink reaction for sucrose in the experimental system. The model only has the three carbon sink fluxes shown, so their flux partitioning values always add up to 100%, while this is not the case for the experimental data. Somewhat surprisingly, cytosolic sucrose concentration changes in the model remained relatively constant, as did the flux of sucrose accumulation into the vacuole, which led Uys et al. (2007) to conclude from these modelling results that accumulation of large amounts of sucrose in the vacuole is required to explain the observed increases in overall sucrose level with culm tissue maturation.

### Continuous growth modelling and phloem flow

The model described above distinguished between internodes purely by substituting internode-specific maximal activity data for the pathway enzymes. This is of course a simplification, and in reality the internodes are linked via the phloem—in other words, there is a flow of mass (mainly sucrose) between internodes that needs to be accounted for. In addition, the model did not simulate explicitly any subcellular compartments other than the cytosol.

To address these shortcomings, a method was developed for incorporating phloem flow into the sugarcane model. Phloem translocation has been the topic of a comprehensive modelling study by Thompson and Holbrook (2003a). In essence, an osmotic potential is established through an unequal distribution of solute in the phloem sieve tubes; the resulting pressure (turgor) gradient can then push fluids up or down. However, it is possible for solute to move faster than the actual fluid medium through so-called pressure–concentration waves (Thompson and Holbrook, 2003b, 2004; Thompson, 2005). The review by Thompson (2006) provides a good introduction to the mechanics of phloem flow. The studies by Höltä et al. (2006, 2009), Minchin et al. (1993), Henton et al. (2002), Minchin and Lacointe (2005), and Lacointe and Minchin (2008) have also modelled fluid flow in plants.

To couple kinetic models with models of fluid flow in plants, it was necessary to consider the vacuole, apoplast, symplast, and phloem explicitly as compartments. To model the distribution of sucrose throughout the culm, the stalk was considered as a one-dimensional object, and time-dependent metabolite concentration changes along this dimension were simulated with the following partial differential equation (Uys, 2009; L. Uys et al., unpublished results):

$$
\frac{\partial s}{\partial t} + \frac{\partial}{\partial z} \left( \bar{u} s \right) + \frac{\partial}{\partial z} \left( D_i \frac{\partial s}{\partial z} \right) = v_{rx} + v_{ir}
$$

which is illustrated here for only one metabolite S (with concentration s). Similar equations were constructed for all metabolites in all compartments. Time is indicated by t, z is the distance along the z-axis (length of stalk), s̅ is the phase average velocity of fluid flow for the particular compartment, $D_i$ is the diffusion coefficient of S, $v_{rx}$ is the sum of all metabolic reactions in which S participates, and $v_{ir}$ is the sum of all intercompartmental transport processes.
Fig. 3. Metabolic reactions of the extended model of sucrose accumulation and futile cycling in sugarcane parenchymal tissue (Uys et al., 2007). Metabolites are indicated by a small circle, enzymes by a numbered grey box. Isozymes are grouped by a surrounding box and numbered a, b, and c where applicable.

Fig. 4. Validation of the sugarcane growth model (Uys et al., 2007) with flux partitioning data from sugarcane tissue disks supplemented with glucose. Experimental data, indicated by symbols, are from Bindon and Botha (2002). Model data are indicated by lines. Flux ratios are defined as follows with numbering as in Fig. 3: vacuolar sucrose, $J_{11}/J_2$; fibre, $J_{13}/J_2$; glycolysis and respiration, $J_{14}/J_2$. Reprinted from Phytochemistry, 68, Uys L, Botha FC, Hofmeyr JHS, Rohwer JM. Kinetic model of sucrose accumulation in maturing sugarcane culm tissue. 2375–2392, 2007, with permission from Elsevier.

involving S. From left to right, the terms in Equation 9 simply state that the concentrations change with time, species may undergo advection (i.e. transport by a fluid—in this case the phloem sap—due to the fluid’s bulk motion in a particular direction), species may diffuse, and lastly they engage in metabolic reactions or move between compartments (cross-membrane transport). Equation 9 is termed an advection–diffusion–reaction system.

Equation 9 was computationally implemented in a model comprising five node–internode pairs using a finite volume method and analysed with the FiPy software (http://www.ctcms.nist.gov/fipy; Guyer et al., 2009). The pathway was a simplification and the model aimed to capture the essential features of the system but was not based in detail on experimentally determined data. Such ‘core’ or ‘toy’ models can often yield biologically relevant results, as in this case (Fig. 5), in spite of the model parameterization with default values. For example, the concentration of sucrose in the phloem was highest at the nodes due to phloem loading from the attached leaves, but then spread out in both directions to
create ‘saw-tooth’ profiles. Also, the ‘filling-up’ behaviour of the stalk could be observed; vacuolar sucrose was accumulated at first preferentially in the more mature internodes, but gradually the younger internodes also progressively accumulated sucrose (Uys, 2009).

This section reviewed some of the progress that has been made in understanding plant physiology and metabolism through mathematical modelling. One of the stumbling blocks in constructing accurate models is the shortage of enzyme kinetic and model verification data. Recent advances that may overcome these challenges, as well as possible future research directions, are discussed below.

The future of kinetic modelling in the ‘omics’ era

This review has dealt with the principles of modelling of metabolic pathways, starting out from structural approaches that require only stoichiometric information, and leading on to kinetic modelling, which requires enzyme kinetic data in addition to the stoichiometry. A few applications of each approach were discussed in detail. Looking into the future, the following areas will pose new challenges and have the potential of providing new solutions.

First, models need to be integrated on different levels of the cellular hierarchy (e.g. transcription, translation, signalling, and metabolism). Merely simulating the metabolic level on its own (or the signalling level, for that matter) has its shortcomings, as the other levels can potentially contribute to the control and regulation of cellular function. In this regard, computational systems biology can provide an invaluable tool (e.g. Wiechert, 2002), as it may point to cases where the hierarchical response is important for completing the picture. Building a kinetic model that realistically addresses the response of a cellular system to some perturbation would have to take account of the different levels, thus requiring the large-scale simultaneous measurement of specific RNA, protein, and metabolite levels, not only for one state, but in a time-course that reflects cellular dynamics. Klipp et al. (2005) pioneered this approach with a model of the yeast osmotic stress response at the metabolic, signalling, and gene expression level. Ultimately, for multicellular organisms this analysis will have to be extended to different levels of the organizational hierarchy (molecule, biochemical pathway, organelle, cell, tissue, organism, and population).

Secondly, new methods have to be developed for integrating and extracting information from high-throughput transcriptomics, proteomics, enzyme activity, and metabolomics data. Such data sets are now becoming more readily available in plants. For example, mRNA transcripts, enzyme activities, and metabolite levels have been measured quantitatively in the same samples in Arabidopsis over the diurnal cycle (Gibon et al., 2004b, 2006), and Sulpice et al. (2010) extended the same type of analysis to comparisons across 129 Arabidopsis accessions.

Thirdly, the huge data sets provided by transcriptomics, proteomics, and metabolomics experiments should also lay the foundation for new methods of MCA. For example, Camacho et al. (2005) used the framework of MCA in combination with kinetic modelling to investigate the origin of correlations in metabolomics data.

Finally, the combination of kinetic modelling with phloem flow in an advection–diffusion–reaction framework as outlined earlier paves the way for creating integrated models of whole-plant physiological function. The framework is not limited to sugarcane and can be applied to other plants as well, and has the potential also to be applied to xylem flow to obtain a more complete picture of water and solute homoeostasis.

In conclusion, the tenet of the work presented in this review is that a thorough understanding of cellular systems

![Fig. 5. Time-course of sucrose concentrations in the phloem and vacule in a toy model implementation of the advection–diffusion–reaction framework in Equation 9 (Uys, 2009). The y-axis shows the length of the stalk from 0 (young tissue) to 1 (older tissue). Time is shown on the x-axis. The stalk length was subdivided into 50 finite volume elements comprising five node–internode pairs, so that each node was modelled by a single volume element and the corresponding internode by nine volume elements (Uys, 2009). Sucrose concentrations are colour-coded according to the colour bar on the right-hand side of each figure. Units are arbitrary since the model was parameterized with default values.](image-url)
and their physiology requires quantitative experimentation linked to mathematical analysis and numerical modelling within a rigorous theoretical framework. Such understanding through predictive models will enhance our ability to manipulate organisms in a directed, targeted way. This approach holds a great future for biology in general, and for plant physiology in particular.

Acknowledgements

The author acknowledges financial support from the South African National Research Foundation (NRF). Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and therefore the NRF does not accept any liability in regard thereto.

References


lead to different simulations of the chlorophyll a fluorescence transients. *Journal of Theoretical Biology* 257, 260–269.


Rossouw D, Kossmann J, Botha F, Groenewald J. 2010. Reduced neutral invertase activity in the culm tissues of transgenic sugarcane plants results in a decrease in respiration and sucrose cycling and an increase in the sucrose to hexose ratio. *Functional Plant Biology* 37, 22–31.


