Chapter 27

Using a Multi-compartmental Metabolic Model to Predict Carbon Allocation in *Arabidopsis thaliana*

Maksim Zakhartsev

Abstract

The molecular mechanism of loading/unloading of sucrose into/from the phloem plays an important role in sucrose translocation among plant tissues. Perturbation of this mechanism results in growth phenotypes of a plant. In order to better understand the coupling of sucrose translocation with metabolic processes a multi-compartmental metabolic network of *Arabidopsis thaliana* was reconstructed and optimized with respect to biomass growth, both in light and in dark conditions. The model can be used to perform flux balance analysis of metabolic fluxes through the central carbon metabolism and catabolic and anabolic pathways. Balances and turnover of energy (ATP/ADP) and redox metabolites (NAD(P)H/NAD(P)) as well as proton concentrations in different compartments can be estimated. Importantly, the model can be used to quantify the translocation of sucrose from source to sink tissues through phloem in association with an integral balance of protons, which in turn is defined by the operational modes of the energy metabolism (light and dark conditions). This chapter describes how a multi-compartmental model to predict carbon allocation is constructed and used.

Key words Energy metabolism, Multi-compartment metabolic model, Central carbon metabolism, Sucrose metabolism, Sucrose transport, Flux balance analysis, Diurnal growth

1 Introduction

In growing vascular plants, sucrose is the most widespread sugar used to supply both carbon and energy, from "source" tissues (e.g., autotrophic mesophyll) to "sink" tissues (e.g., heterotrophic roots), where it is used to build up biomass [1] (Fig. 1). This process of redistribution of photosynthetically fixed carbon between plant tissues is called sucrose translocation. The phloem, consisting of sieve elements, companion cells, and parenchyma cells, plays an important role in this process. During photosynthesis in plastids of mesophyll cells, triose phosphates (GAP, DHAP) are synthesized and exported into the cytoplasm to be transformed into sucrose, primary metabolites, and finally biomass. During growth under light, one part of fixed carbon is exported and

Johannes Liesche (ed.), *Phloem: Methods and Protocols*, Methods in Molecular Biology, vol. 2014,
https://doi.org/10.1007/978-1-4939-9562-2_27,

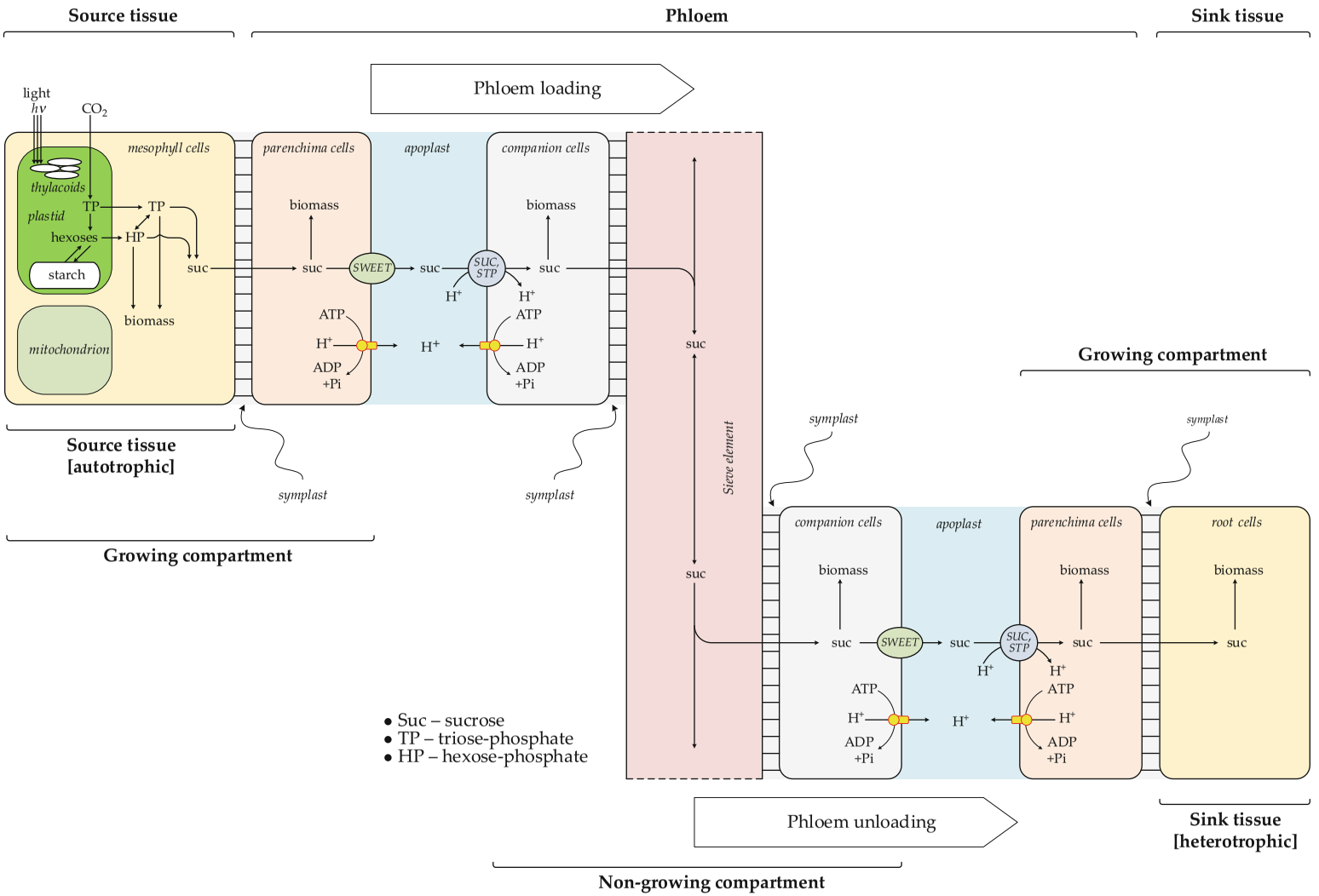

Fig. 1 Phloem-centric block diagram of sucrose translocation from source (autotrophic) to sink (heterotrophic) tissues. In this version, the unidirectional flux of sucrose is presented, although it can be inversed in dependence on the location of starch deposits, since sucrose concentration in phloem is a homeostatic variable. The diagram is based on the synthetic of views from [1, 11, 12, 17, 38]. From the modeling point of view, the cells joined via symplast were logically merged into growing and nongrowing compartments, which are separated by apoplast. *SWEET*—sucrose efflux transporters; *SUC*, *STP*—sucrose-proton symporters

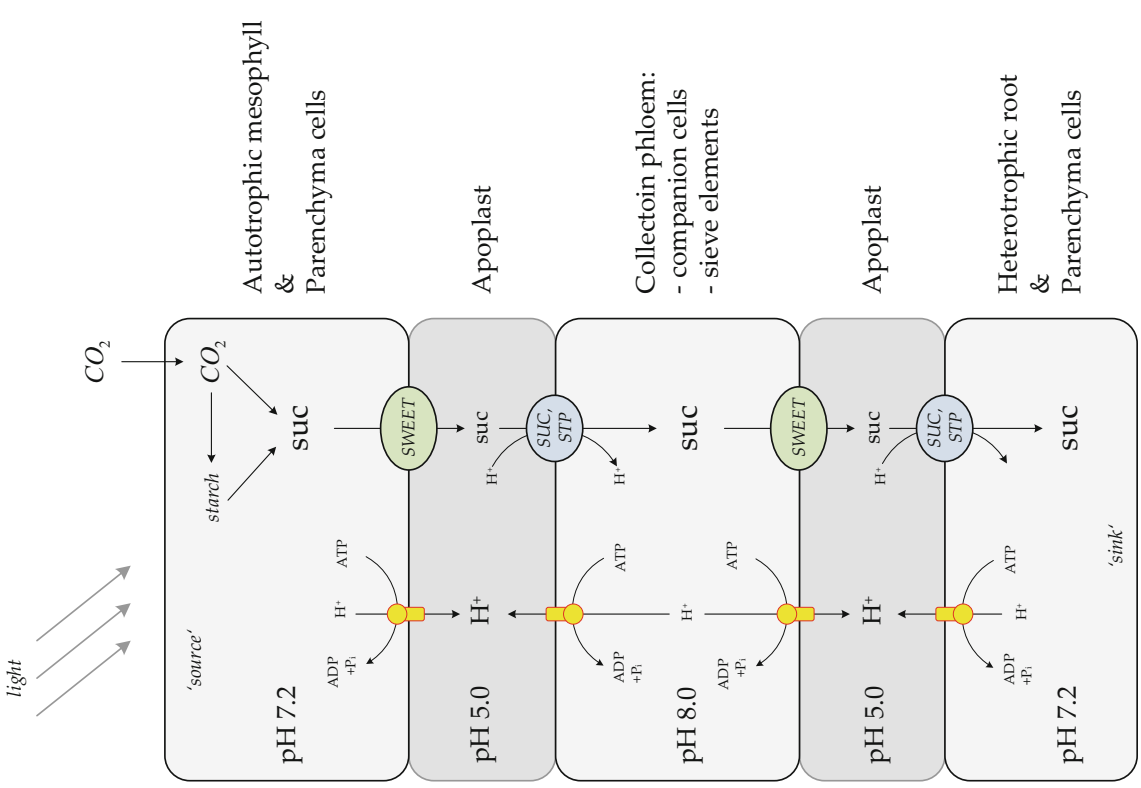

Fig. 1 (continued)

another part is stored in the form of starch, typically in plastids in the leaf [2, 3]. Starch (i.e., the carbon repository) is the primary carbon source for biomass formation during growth in the dark [2]. The diurnal dynamics of starch accumulation is generally well documented [2] and modeled [4]. Perturbation of these tightly regulated metabolic processes results in growth phenotypes of the plants. For example, disruption of the plant's ability to invest carbon into daytime storage of starch in the *pgm* mutant [5] results in higher cytosolic sucrose levels, higher respiration rates, retarded growth [6], low seed yield [7], and slow root growth at night [8]. The cause-effect relationship of sucrose translocation between organs and phenotypic traits of plants is an important area of current plant research [9].

Sucrose is translocated between organs through the phloem. The sucrose is loaded from the source tissues into the phloem sieve elements, which form a low-resistance pathway for mass flow of the sugar-rich phloem sap (Fig. 1). Sucrose is unloaded from the phloem in the sink tissues [10, 11] into the apoplast between the parenchyma and the companion cells (Fig. 1). The molecular mechanism of apoplastic sucrose transport involves several efflux/influx carriers [10–12]. There are three families of sugar transporters known in *Arabidopsis thaliana* of which at least some members specifically transport sucrose: *SWEET*, *SUC*, and *STP*, which differ in functional properties. The members of the *SWEET* family facilitate sucrose transport across membrane along its gradient [13], whereas the members of *SUC* and *STP* families transport sucrose across membranes as symport with protons [14–16] against the sucrose gradient, but along the proton gradient, which is generated by plasma membrane H^+-ATPase activity [11, 17].

All plant tissues simultaneously express members of all three families of sucrose transporters (*SWEET, SUC*, *STP*), but in different quantities [18–20]. This fact indicates that efflux and influx events through the apoplast (including into/from phloem) are coordinated to form a net flux of sucrose from any sources (i.e., photosynthetic cells, starch deposits in leaves/roots) to any tissue that requires sucrose for metabolic needs. Consequently, the plants in which the sucrose transporters are mutated or knocked out exhibit specific traits [21–23]. The sucrose concentration in phloem can be considered as the homeostatic variable connecting source production and sink demand. Since *Arabidopsis thaliana* stores starch in leaves the net flux of sucrose is directed from the leaf as the source tissue through the phloem to a root as the sink tissue [10, 11]. Therefore, it is valid to generalize the molecular mechanism of sucrose translocation among organs as presented in Fig. 1. Such generalized view of the molecular mechanisms takes into account only two chemical motive forces (sucrose and proton gradients) and the respective transporters that use them (*SWEET* efflux transporters and *SUC/STP* sucrose-proton symporters).

Intercellular diffusion of sucrose through plasmodesmata might be included in a future version if quantitative data under the corresponding experimental conditions becomes available.

The analysis of metabolic networks with a rational/quantitative approach is an efficient tool for engineering of plant systems [24, 25]. Model-based approaches are increasingly used also in plant biology to gain functional insights and even to predict metabolic processes [26, 27]. In general, development of a stoichiometric steady-state model of a metabolic network involves several stages: (1) collection of a priori knowledge and reconstruction of a network of metabolic reactions, and elimination of information gaps [28, 29]; (2) testing of the proposed metabolic network model to eliminate structural gaps by means of topological analysis of a stoichiometric matrix; (3) formulation of constraints and an objective function for the model; (4) network optimization using constraint-based analyses such as flux balance analysis (FBA); (5) a posteriori consistency check of the identified model by metabolic flux analysis (MFA); and (6) validation of the model with experimental datasets [30, 31]. This workflow follows an iterative approach, which improves the model with each iteration (Fig. 2). Based on this workflow successful metabolic models of *Arabidopsis thaliana* operating at different scales have been developed [24, 26, 29, 32–35], including multi-tissue genome-scale metabolic models [33, 36–38]. However, each of these metabolic models is dedicated to understanding a specific biological phenomenon, such as photosynthesis [39], the Calvin-Benson cycle [40], metabolic costs of amino acid and protein synthesis in light and dark conditions [41], stability of metabolic fluxes in central metabolism of *Arabidopsis thaliana* root cells in relation to oxygen levels [42, 43], or contribution of sucrose translocation to energy, redox, and proton balances in growing *Arabidopsis thaliana* [38]. The last model is genome scale and it describes biomass growth both in light and in dark conditions with corresponding formation and consumption of starch and sucrose. The model has taken into account major pathways of primary metabolism such as sugar metabolism, central carbon metabolisms, photosynthesis, photorespiration, energy and redox metabolism, proton turnover, nutrient transportation, sucrose translocation from source to sink tissues, and biomass growth in various compartments.

This model is described here in detail together with instructions on how to apply it. As a network model, it consists of compartments, compounds, and reactions. Reactions transform compounds into other compounds. The collection of all reactions constitutes the mathematical model. For each balanced compound, an ordinary differential equation is generated, which describes its mass balance including all processes of formation, degradation, and transport of the respective compound. Relevant terms for understanding this model and its parameters are described below before providing instructions on how to use it.

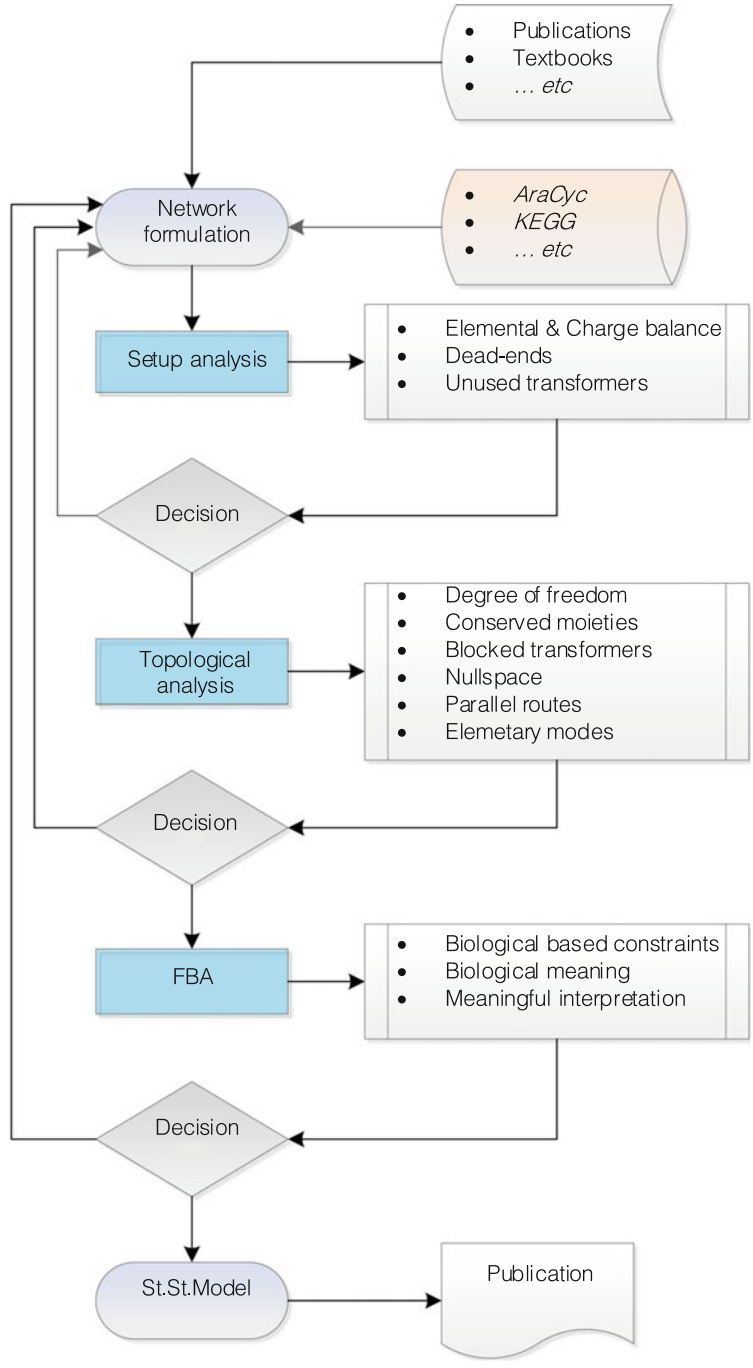

Fig. 2 The generalized workflow of development of steady-state (i.e., stoichiometric) models. The algorithm implies iterative development of the model throughout different stages

1.1 Model Assumptions and Biologically Based Constraints

Each stoichiometric metabolic model is formulated under specific biological assumptions that describe the conditions under which the model is valid. The assumptions are required in order to reduce the complexity of the model. With regard to the present model, this means that assumptions are made to focus on the mechanisms that mainly contribute to sucrose translocation.

1.2 Compartmentalization

Model compartmentalization is an important step in setting up of the model. The compartments serve as container for the network component's reactions and polymerizators (i.e., transformers). Transport steps are located at the compartment boundaries. The compartments are important when separating the functions of conserved moieties participating in energy and redox reactions, and also in cases of differential usage of the same metabolites (e.g., triosephosphates, protons). In the multi-compartmental model there are two types of compartments: (1) tissue and (2) subcellular. In this model [38], the principal molecular mechanism of sucrose translocation is the basis for a multi-compartmental metabolic model (Fig. 1).

1.3 Network Reconstruction

The structured metabolic network of *Arabidopsis thaliana* was reconstructed based on the decision to take into account only those metabolic processes, which contributed to better understanding of sugar metabolism, central carbon metabolism, energy and redox metabolisms, proton turnover, biomass growth, and sucrose translocation among tissues in *Arabidopsis thaliana*. The reconstructed network consists of **transformers** which are located in the compartments or at its boundary. The transformer holds attributes and behavior for describing of (1) enzymatic or nonenzymatic reaction of transformation from one or several substrates to one or several products, (2) polymerization, and (3) transport. A **transporter** is a transformer capable of transporting compounds across the compartment boundary. Some transformation steps can be concatenated in linear pathways (e.g., nucleotide or amino acid biosynthesis) and, therefore, create a **transformer subset**, which can be lumped to an overall transformer in order to reduce the model complexity without changing the degrees of freedom. All transformers in the model were manually transferred from the database, and then manually curated. The transformers operate with compounds, which hold attributes for chemical entity like a metabolite or protein.

1.4 Biomass Composition

The growth of biomass is an objective function ($\mathbf{Z}$) for the flux balance analysis (FBA; Eq. 1); therefore the composition of the biomass is a very important constraint, which defines output of the model. The biomass was assumed to be composed of the following major polymers: proteins, lipids, RNA/DNA, carbohydrates (i.e., starch, sucrose, cellulose, pectin), and ash [34, 42]. Polymerization

forms large molecules like proteins from small molecules, the so-called monomer building blocks (e.g., amino acids). In metabolic networks, polymers are formed by condensation reactions. Monomer composition for each macromolecular constituent should be a priori known from experimental data, literature, or databases.

1.5 Model Annotation

Each transformer, compound, and gene in the model has unique identifier, which allows for a unique identification of the network components with regard to the compartment. The annotation of each transformer step in the model has been made on the base of gene-protein-reaction association acquired from various databases (e.g., NCBI, KEGG, TAIR). Each biochemical reaction (classified with E.C. number) is referred to corresponding KEGG reaction (KeggID). Additionally, reactions are annotated with accession numbers of genes whose protein products perform this biochemical process (GeneID). In case of lumped reactions, all corresponding GeneIDs are associated with the lumped reaction. The annotation allows mapping transcriptomic and proteomic data [44] in order to validate the presence of reactions in the model.

1.6 Network Setup

The basic information on the stoichiometry of metabolic reactions can be collected from existing databases, textbooks, publications, etc. The detailed consistency check of the network stoichiometry is required in order to find and eliminate **gaps** in the network setup. The **stoichiometry** of biochemical reactions is checked on the basis of elemental and charge balances of the reactions, which relies on known elemental composition and charge (at specific pH) of metabolites. The consistency check allows identifying and revising/eliminating: (1) reactions with incorrect stoichiometry; (2) compounds which are missing/incorrect the elemental composition or charge; and (3) inconsistent metabolic reactions, polymerizations, and transport steps with unbalanced elements or charges. A check for **dead ends** (a compound that is linked to only one transformer) and **unused transformers** allows identifying and eliminating transformers and network regions which are not capable of operating at steady state. The detection of unused transformation steps is done by analyzing the base vectors of the null-space. A sufficient condition for an unused transformation step is that for an unused transformation step all corresponding entries of the base vectors are zero. A special care must be paid to the known directionality of the reaction/processes since it is an important constraint for the flux balance analysis. A transformation step is called **irreversible** if the direction of the net flux is fixed. A check for irreversible transformers allowed visualizing and approving those transformers, which are irreversible from the thermodynamic point of view.

1.7 Topological Analysis of the Network

The reconstructed network of reactions can be presented in terms of the matrix algebra using the stoichiometric matrix of the network (**S**) and the rate vector of metabolic fluxes (**v**) (Eq. 1). The stoichiometric matrix holds the stoichiometric factors of the individual transformers (reactions/transporters/polymerizators) for each metabolite node in the network. The topological analysis analyzes the topology of the **S** and it is a very powerful tool in optimizing the structure of the metabolic network. This analysis allows identifying false and/or unused network structures. The degree of freedom estimates the dimensions of the network, and further requirements to solve it with respect to experimentally derived measurements. The **total degree of freedom** denotes the number of fluxes which can be freely chosen to determine the system's steady state. The total degree of freedom equals (1) the number of base vectors of the null-space and (2) the number of fluxes minus the number of linearly independent balance equations. The **inner degree of freedom** denotes the number of fluxes which can be freely chosen to determine a steady state while all fluxes crossing the system boundary are set to zero. The inner degree of freedom thus equals the number of parallel routes/reaction cycles. The **outer degree of freedom** is defined by the total degrees of freedom minus inner degrees of freedom. Analysis of **conserved moiety** reveals sets of compounds, whose sum always remained the same even under dynamic conditions. The number of conserved moieties equaled the number of linearly dependent balance equations. Analysis of conserved moieties can reveal false-positive conserved moieties which had accidentally appeared due to the topology of the reconstructed network. Accidental false-positive conserved moieties must be eliminated and only biologically relevant conserved moieties (e.g., ATP/ADP, NAD(P)H/NAD(P), acetyl-CoA/CoA/succinyl-CoA) should remain in the network. Analysis of **blocked transformers** identifies regions of the network, which were not capable of operating at steady state due to the irreversibility constraints of the transformer equation. **Null-space** of **S** is a set of all solutions of the linear equation system $\mathbf{S} \times \mathbf{v}=0$. Computation of the base vector of the null-space reveals "functional" and "non-functional" regions of a metabolic network. Computation of **parallel routes and cycles** reveals a set of reactions/transporters/polymerizators which (1) are capable of maintaining a steady state, (2) could not be decomposed, and (3) in total did not consume or produce an external substrate/product. From condition (3) follows that the net reaction of parallel routes and reaction cycles is zero. In the complex reconstructed networks, there is a high chance for the accidental formation of parallel routes and cycles which increases internal degree of freedom. Parallel routes and cycles substantially increase the number of elementary flux modes without leading to new phenotypic behavior. Therefore, their number had to be minimized in order to use elementary

mode analysis for evaluation of maximal product yield. The set of all solutions of the linear equation system $\mathbf{S} \times \mathbf{v}=0$ are called elementary flux modes of **S**, if (1) the solutions obey reversibility criteria and (2) every nonzero element of **v** is essential for maintaining the steady state. Computation of **elementary modes** reveals a unique set of smallest sub-networks that allows a reconstructed network to function in steady state. Elementary mode analysis takes into account stoichiometry and thermodynamics when evaluating whether a particular metabolic route or (sub-)network is feasible and likely for a given set of proteins/enzymes.

1.8 Flux Balance Analysis (FBA)

The distribution of intracellular metabolic fluxes through the reconstructed metabolic network was investigated under pseudo-steady-state assumption ($\mathbf{S} \times \mathbf{v}=0$) using constraint-based flux balance analysis (FBA). In order to calculate metabolic flux distribution, maximization of biomass formation was formulated as the objective function (**Z**). The FBA is a method to calculate flux distribution in the metabolic network to find maximum yield of a specified compound by linear optimization of an objective function:

$$\begin{array}{cc} \text{maximize} & \mathbf{Z} = \mathbf{c}^{\mathrm{T}}\mathbf{v} \\ \text{subject to} & \mathbf{S} \times \mathbf{v} = \mathbf{0} \\ \text{and} & v_i^{\min} \le v_i \le v_i^{\max} \end{array} \quad (1)$$

where **Z** is the objective function, **c** is a vector of weighting factors, **v** is a vector of metabolic fluxes, v_i is the ith element of **v**, and $v_i^{\min}$ and $v_i^{\max}$ are the minimum and maximum constraints on v_i. The topology of the stoichiometric matrix (**S**), together with lower ($v_i^{\min}$) and upper ($v_i^{\min}$) boundaries for the vector of metabolic fluxes, in combination with thermodynamic properties of reactions (e.g., irreversibility), and steady-state condition ($\mathbf{S} \times \mathbf{v}=0$) provided with sufficient constraints to optimize objective function. The linear optimization is a procedure for locating the maximum or minimum of a linear function of variables that are subject to linear constraints.

1.9 Model Verification

The model verification is checking that the model definition (equations) matches the real world (experimental data). The designed metabolic model is subjected to FBA under two specific constraints: "light" and "dark" scenarios. Each scenario is the set of boundary conditions for FBA which resembles biological essence. For example, under the "light" conditions the photons, CO_2, water, and protons are consumed, whereas O_2, starch, and biomass are produced; under the "dark" conditions photons = 0, O_2, protons, and starch are consumed, whereas CO_2, water, and biomass are produced. Additional verification of the metabolic model can be done on the base of comparison of metabolite turnover ratios predicted

by the metabolic model with the empirical values. For example, such turnover ratios for the plants are photon/CO_2 [28], photon/NO_3^- [28], CO_2/NO_3^- [28], P/O in mitochondria [45], RQ [46], H^+_{in}/ATP in mitochondria [47, 48], ATP/NADH in cytoplasm [49], ATP/NADPH in plastid [46, 50], etc.

1.10 Model Availability

The stoichiometric model (exported in SBML, MATLAB formats), support datasets (model documentation, stoichiometric matrix, FBA constraints, and solutions for different conditions), metabolite database (e.g., elemental composition, charge, molecular weight, external database reference), transformer database (e.g., identifier, trivial name, stoichiometric equation, belonging to pathway, EC number of the reaction, external database reference), and gene database (gene ID in ATG format, catalyzed reaction, gene definition, gene locus, etc.) were uploaded to FAIRDOM (Data and Model Management service to the European Systems Biology community) [51] under ZucAt project (from German "zucker" and *Arabidopsis thaliana*) at https://fairdomhub.org/projects/37 and are available for download under DOI: 10.15490/seek.1.investigation.74.9.

2 Materials

2.1 Model Organism

1. The model organism is thale cress *Arabidopsis thaliana* (L.) Heynh [NCBI:txid3702].

2.2 Reference Genome

1. The reference genome of *Arabidopsis thaliana* (assembly TAIR10) [https://www.ncbi.nlm.nih.gov/genome/4].

2.3 Databases for Network Reconstruction

1. AraCyc [52] and Pathway Tool v19.0 [53] for the primary information on the metabolic reactions regarding stoichiometry, direction, and genes associated with particular biochemical reactions (classified with Enzyme Commission number).
2. TAIR [54, 55], KEGG [56] for verification of metabolic reactions.
3. ChEBI [57] for molecular structure, elemental composition, and charge state of metabolites.

2.4 Software

1. Insilico Discovery™ (Insilico Biotechnology AG, Stuttgart, Germany; www.insilico-biotechnology.com) for model reconstruction, consistency check of the network setup, topological analysis of the stoichiometric matrix, constrained flux balance analysis, mapping transcriptome and proteome data.
2. Matlab or software that can read SBML code, such as GNU Octave, can be used to run the model script.

3 Methods

3.1 Aims of Modeling

1. Formulate the question that you intend to answer with the aid of modeling (*see* **Note 1**).

3.2 Compiling the Databases

1. Collect the compound database from available sources (e.g., ChEBI); provide each compound with a unique identifier (CompoundID); make sure that the elemental composition and charge state of the compounds are available.
2. Collect the gene database from available sources (e.g., NCBI, KEGG, TAIR), and use gene accession numbers in ATG format as the unique identifier (GeneID).
3. Collect the transformer database from available sources (e.g., AraCyc, Pathway Tools); give to each transformer the unique identifier (TransformerID); make sure that the stoichiometric equation uses the corresponding CompoundID.
4. Perform gene-protein-reaction (GPR) association of the transformers with particular biochemical reactions (denoted with EC number) and genes (GeneID) whose protein products perform the biochemical reaction.
5. Annotate the transformers with KeggID (available from source database KEGG) in accordance with their EC number and metabolite specificity.
6. Make sure that the transformers that perform irreversible reactions are correctly denoted.

3.3 Network Formulation

1. Investigate the biological phenomenon in detail, and derive molecular mechanisms that accurately describe the biological processes.
2. Formulate biologically based assumptions and natural constraints under which the model is valid (*see* **Note 2**).
3. Design compartmentalization of the model [(1) top compartment, (2) tissues, and (3) subcellular] based on knowledge of biological processes taking place in the corresponding compartments (Fig. 3) under the derived assumptions (*see* **Note 3**).
4. Choose the metabolic pathways that mostly contribute to the modeled biological phenomenon and describe it in detail. This is the most challenging and creative step in the network reconstruction workflow since it requires detailed a priori knowledge about biochemistry and physiology of the studied organism (*see* **Note 4**).
5. Arrange transformers in corresponding compartments and organize them into pathways.

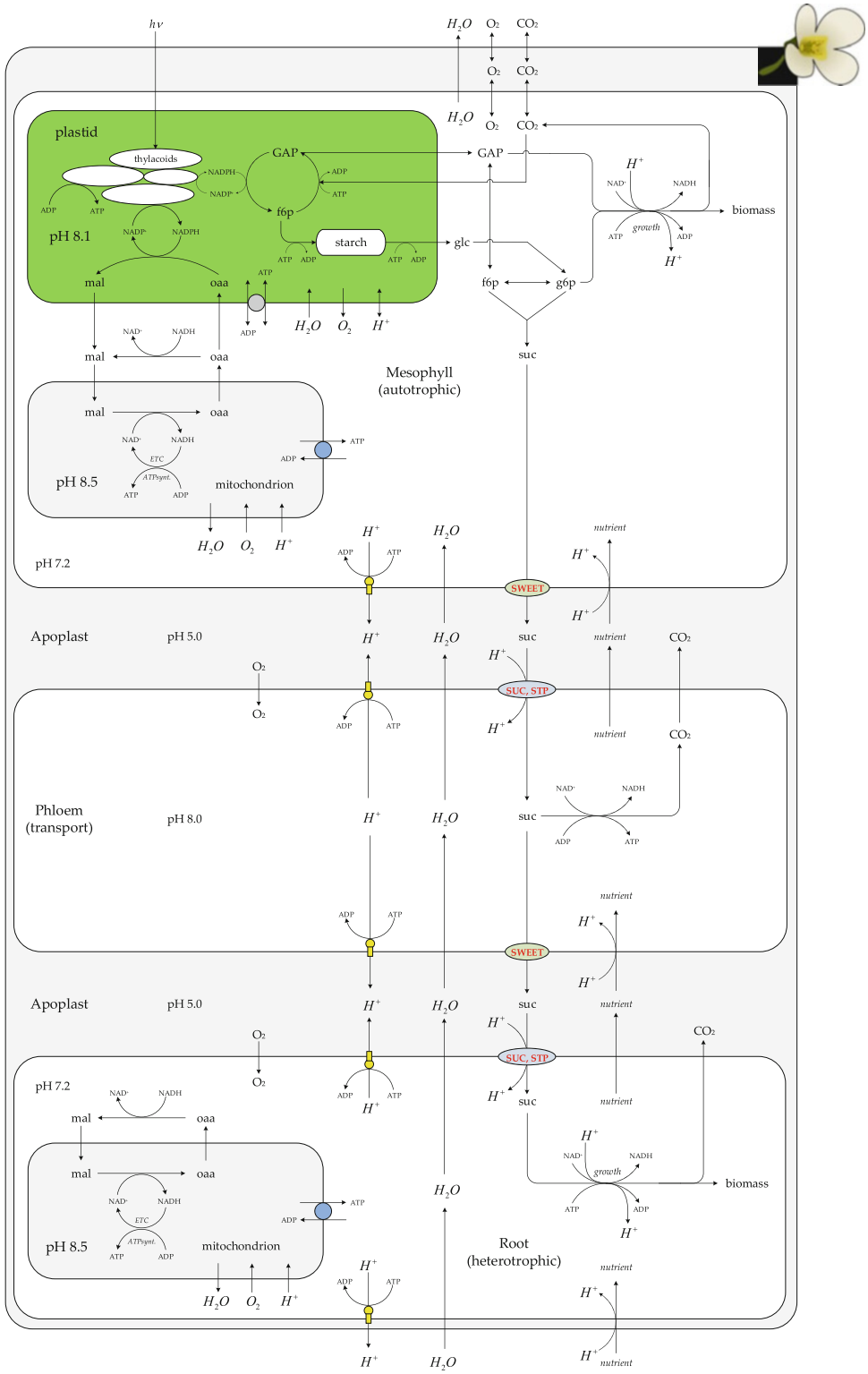

Fig. 3 Schematic circuit of the central carbon, energy, and redox metabolisms of *Arabidopsis thaliana* (adopted from [38]). The model consists of super-compartment "plant," which includes growing autotrophic sub-compartment "mesophyll," nongrowing transport sub-compartment "phloem," and growing heterotrophic sub-compartment "root." The inner space of the super-compartment "plant" was defined as of "apoplast." The "mesophyll" contains subcellular compartments "plastid" and "mitochondrion" while the

6. Choose or create and place the transporters on boundaries between corresponding compartments; make sure that transporters exchange all required compounds among compartments; make sure that symport/antiport properties of transporters are correctly denoted.
7. Formulate biomass composition in terms of w/w ratio among macromolecular constituents (i.e., polymers: proteins, DNA/RNA, carbohydrates) and fractional ratio among monomers (e.g., amino acids, nucleotides, sugars, fatty acids); pay attention to the side reactions (e.g., ATP hydrolysis as a result of protein polymerization) (*see* **Note 5**). The formed biomass is irreversibly exported from the model.

3.4 Checking Network Setup and Consistency

1. Check the elemental and charge balance for each transformer; in case of an inconsistency, revise the corresponding stoichiometric equation in the transformer database and update the transformer in the model (*see* **Note 6**).
2. Identify dead-end compounds; in case of inconsistency eliminate the dead ends and update the model (*see* **Note 7**).
3. Identify unused transformers; in case of inconsistency eliminate the unused transformers and update the model (*see* **Note 8**).
4. Identify irreversible transformers; this is a double-checking step to make sure that the irreversible transformers are present in the network and correctly allocated.

3.5 Topological Analysis

1. Get the degrees of freedom (total, inner, and outer) of the stoichiometric matrix; the lower the degrees of freedom, the better; reduction of parallel routs and cycles reduces the inner degree of freedom; revise the network formulation and update the model accordingly (*see* **Note 9**).
2. Identify conserved moieties; if characteristic conserved moieties cannot be identified, the missing relation might serve for debugging network models; or if nontypical conserved moieties are identified, the network setup must be revised to eliminate them; revise the network formulation and update the model accordingly (*see* **Note 10**).

Fig. 3 (continued) "root" compartment contains only "mitochondrion." Details of metabolic pathways were hidden in order to focus only on the specificity of the sucrose synthesis/translocation in relation of H^+-turnover, and nutrient and water transport between tissues. *hν*—light photon; *GAP*—glyceraldehyde 3-phosphate; *suc*—sucrose; *g6p*—glucose-6-phosphate; *f6p*—fructose-6-phosphate; *oaa*—oxaloacetate; *mal*—malate; *H*$^+$—proton; *ETC*—electron transport chain that performs oxidative phosphorylation; *growth*—collective set of reactions resulted in formation of biomass; *ATPsynt*—ATP synthase; *nutrient*—nutrients such as NO_3^-, HPO_4^{2-}, SO_4^{2-}; *SWEET*—sucrose efflux transporters; *SUC*, *STP*—sucrose-proton symporters

3. Identify blocked transformers; eliminate blocked transformers from the network setup by revising the stoichiometric equations of transformers, because blocked transformers indicate the presence of contradicting irreversible transformers; revise the network formulation and update the model accordingly (*see* **Note 11**).
4. Identify transformer subsets; transformers that compose the subset can be lumped into overall transformer; revise the network formulation and update the model (*see* **Note 12**). This step is optional.
5. Compute the null-space of the stoichiometric matrix. Using the result of the analysis, identify and eliminate nonfunctioning regions from the network setup; revise the network formulation and update the model accordingly (*see* **Note 13**).
6. Compute parallel routes and cycles in the network. Using the result of the analysis, try to eliminate the parallel routes and cycles from the network setup; revise the network formulation and update the model accordingly (*see* **Note 14**).
7. Compute the elementary modes in the network; for the complex networks, this step is computationally demanding (*see* **Note 15**); this analysis allows selecting sets of elementary modes which lead from specific input substrate (e.g., glucose) to specific output product (e.g., an amino acid, biomass) and compare the maximal specific yields among them; additionally it allows identifying the essential transformers (*see* **Note 16**); these analyses allow making decision on the optimization of the network setup; revise the network formulation and update the model accordingly.

3.6 Flux Balance Analysis (FBA)

1. Define the objective function (**Z**; Eq. 1) for FBA (*see* **Note 17**); in case of the model [38], it was "maximization of the biomass production" under different growth scenarios: light/dark.
2. Formulate constraints for flux balance analysis; the network stoichiometry and transformer attributes (e.g., irreversibility of reactions) provide a set of linear constraints under steady-state conditions ($\mathbf{S} \times \mathbf{v}=0$; Eq. 1); additional constraints defining uptake rate restrictions or split ratios can be formulated; by default, the minimum constraint for a network flux ($v_i^{\min}$) is zero; the maximum constraints ($v_i^{\max}$) can be set for desired flux value; the constraints can be applied to the following fluxes: external exchange reactions, transport, intracellular reactions, and polymerizations; some fluxes can be constrained to zero to inactivate the corresponding transformer; the constraints must be biologically meaningful (*see* **Note 18**) and coherent with the model assumptions (*see* **Note 2**).

3. Run linear optimization; unbounded linear optimizations must be avoided (*see* **Note 19**).
4. Interpret the results of FBA in the context of the biological process. The FBA reports the fluxes through each transformer, as well as the overall stoichiometric equation of the model.
5. Optionally, estimate the turnover of the node compounds by summarizing input/output fluxes.

3.7 Model Verification

1. Test if the transcriptome or proteome dataset matches the model's expression/protein levels under the corresponding experimental conditions. If the transformers are correctly annotated, than the transcriptome or proteome dataset can be overlaid on the model. For an example, *see* [44].
2. Test if metabolite turnover ratios predicted by FBA of the metabolic model match with empirical values. Characteristic turnover ratios for plants are photon/CO_2 [28], photon/NO_3^- [28], CO_2/NO_3^- [28], P/O in mitochondria [45], RQ [46], H^+_{in}/ATP in mitochondria [47, 48], ATP/NADH in cytoplasm [49], ATP/NADPH in plastid [46, 50], etc.

4 Notes

1. Aim of the model in [38] was to quantitatively estimate the translocation of sucrose between tissues through phloem of growing *Arabidopsis thaliana* in association with the integral balance of protons, which in turn is defined by operational modes of the energy metabolism.
2. The multi-compartmental model of *Arabidopsis thaliana* is formulated under the following biochemical assumptions:
   (a) The model describes steady-state growth of *Arabidopsis thaliana* biomass, which is the basic/general assumption/constraint for the stoichiometric model.
   (b) Sucrose is by far the most important carbohydrate that is exchanged among tissues in *Arabidopsis*.
   (c) The anatomy of the whole plant is reduced to only three compartments: autotrophic growing mesophyll, heterotrophic growing root, which are interconnected by non-growing transport phloem. The model is restricted to the mesophyll as always being the "source" tissue of sucrose both in light conditions (during photosynthesis) and in dark conditions (non-photosynthetic condition, carbon derived from starch). The root is always defined as the "sink" tissue for sucrose. The root takes up sucrose and uses it for biomass formation, but does not store or reload it into the phloem. Thus, the exchange of sucrose

is a unidirectional net flux from autotrophic mesophyll via phloem to heterotrophic root (Fig. 1).

(d) The phloem is the nongrowing compartment that only has a role in transport processes and delivers nutrients and water from root to mesophyll, and sucrose from mesophyll to root. In this sense, the compartment phloem also fulfils the role of xylem. To keep the model simple, no biomass-forming metabolic reactions in the phloem are included. Nevertheless, some metabolic pathways that feed H^+-ATPase activity and recycle ATP/ADP moiety are placed in the phloem compartment (Fig. 3).

(e) Starch is considered as a polymer-metabolite that can be formed and accumulated only in autotrophic compartments (mesophyll) during daytime growth and used as the only carbon and energy source during the dark growth phase. Starch is not degraded during light growth phase at all. Starch is not formed and accumulated in heterotrophic tissues (root).

(f) There are two pools of sucrose and starch: (1) available for metabolism and translocation between compartments and (2) the constituent of the biomass, i.e. not available for the metabolism.

(g) Exchange of amino acids and organic acids (glutamate, acetic, citric, malate acids) between tissues is not considered.

(h) Fatty acid biosynthesis is considered to be cytoplasmic to reduce complexity of redox balance in plastids.

(i) ATP/ADP is made exchangeable between subcellular compartments (cytosol, mitochondrion, plastid) through the adenine nucleotide translocator.

(j) NAD(P)H/NAD(P) cannot be exchanged between subcellular compartments. However, the reduced equivalent can be transferred between subcellular compartments through malate/oxaloacetate shuttle.

(k) Nutrient uptake (nitrate, orthophosphate, and sulfate) follows a proton-symport mechanism.

(l) Proton transport against the pH gradient is always active (ATP dependent), unlike transport along the pH gradients, which is passive.

(m) The only metabolite from the central carbon metabolism exchangeable between plastid and cytoplasm in the mesophyll is D-glyceraldehyde 3-phosphate (GAP). Other potentially exchangeable metabolites (i.e., other triose phosphates, hexose phosphates) were not considered in the model to reduce the complexity.

(n) Gas exchange between plant and environment is considered as being passive.

3. The cells that are symplasmically connected were logically merged into growing and nongrowing compartments, which are separated by the apoplast (Fig. 1). A super-compartment "plant" includes (1) the growing autotrophic sub-compartment "mesophyll" (mesophyll and parenchyma cells) with subcellular compartments "plastid" and "mitochondrion," (2) the growing heterotrophic sub-compartment "root" (root and parenchyma cells) with only one subcellular compartment "mitochondrion," and (3) nongrowing transport compartment "phloem" (sieve elements and companion cells) (Fig. 3). Under steady-state assumption on the day/night timescale, the phloem has been considered as nongrowing compartment in this model. The inner space of the super-compartment "plant" served as the extracellular compartment "apoplast" through which the exchange processes between tissue compartments were routed. Subcellular compartmentalization is required due to independent operation of energy (i.e., ATP/ADP), redox (NADH/NAD, NADPH/NADP), and other (e.g., AcCoA/CoA) conserved moieties. Subcellular compartment "plastid" is required to separate photosynthetic light reactions, sulfate, and nitrite reduction, Calvin-Benson cycle with associated triose phosphates, hexose phosphates, NADPH/NADP moiety, ATP synthesis, and starch synthesis. Subcellular compartment "mitochondrion" is required to separate TCA cycle and its associated NADH/NAD conversion activity and ATP synthesis via oxidative phosphorylation. Also, the subcellular compartmentalization is required in order to separate pools of similar metabolites between compartments with different metabolic specificities and to orchestrate their exchange through transport steps [34]. Since sucrose translocation depends on the formation of proton gradients, proton balancing in the source, phloem, and sink tissues is required for the quantitative assessment of the metabolic costs for sucrose translocation. Therefore, the main metabolic processes involved in the production/utilization of protons in various metabolic processes in the different compartments had to be included.
4. The metabolic reactions from the following pathways were used in the reconstructed network: 2-oxoglutarate decarboxylation to succinyl-CoA; adenosine nucleotide de novo biosynthesis; amino acid biosynthesis; aspartate degradation; Calvin-Benson cycle; chorismate biosynthesis; fatty acid biosynthesis; folate metabolism; glutamine biosynthesis; glycolysis; homoserine biosynthesis; inosine-5′-phosphate biosynthesis; maintenance; malate/oxaloacetate shuttle; nitrate reduction;

ornithine biosynthesis; oxidative phosphorylation; PRPP biosynthesis; pentose-phosphate pathway; phosphorus metabolism; photosynthesis light reactions; photorespiration; purine nucleotide de novo biosynthesis; pyrimidine ribonucleotide interconversions; pyruvate decarboxylation to acetyl-CoA; pyruvate fermentation; starch biosynthesis; starch degradation; sucrose biosynthesis; sucrose degradation; sulfate reduction; TCA cycle; UDP-glucose biosynthesis; UDP-glucuronate biosynthesis; and uridine-5′-phospahte biosynthesis.

5. The biomass of *Arabidopsis thaliana* is assumed to be composed of the following major polymers: proteins, lipids, RNA/DNA, carbohydrates (i.e., starch, sucrose, cellulose, pectin), and ash [34, 42]. The average length of a protein was accepted as 330 amino acids, based on the analysis of FASTA protein sequence. The amino acid composition of a protein was accepted from literature [34]. The lipid composition was assumed to be similar to yeast and considered as a polymer of 9.6% hexadecanoate, 20.6% *cis*-octadecenoate, 34.7% *cis*-hexadecenoate, 2.4% octadecanoate, 0.006% linoleate, and 32.4% glycerol. RNA composition was derived by reversed sequence analysis of protein amino acid sequences in FASTA format available from TAIR: 32% A, 26% U, 9% C, and 33% G. DNA composition was calculated from genome GC value (GC % = 36.0528); thus G = C = 18.0264% and A = T = 31.9736%. Overall, the biomass macromolecular composition was considered to be (w/w) 20% cellulose, 20% pectin, 10% starch, 12% sucrose, 25% proteins, 9% lipids, 1% RNA/DNA, and 3% ash. Anatomically, the mesophyll and phloem compartments together form a shoot. The mass ratio between shoot and root was accepted as 85/15 (w/w) based on own measurements of fresh weight of 21-day-old plants grown under long-day (16-h light) conditions in hydroponic culture. The biomass composition is a very important constraint for flux balance analysis and it defines the output from the model.
6. Always check the consistency of the network stoichiometry through elemental and charge balance of the transformer reaction before using advanced network analysis tools. Inconsistent stoichiometry can lead to wrong/incomplete results like wrong information on reaction cycles/parallel routes.
7. Remove dead ends either by removing the corresponding transformers or by extending the network through additional transformation steps if you are interested in obtaining networks free of unused transformation steps. Set the balance status of compounds to non-balanced or introduce additional transport steps to the external compartment.

8. Use the unused transformation steps method for identifying regions which are not capable of operating at steady state. Set the balance status of compounds to non-balanced or introduce additional transport steps to the external for reducing the set of unused transformers.
9. The inner degree of freedom corresponds with the number of solutions obtained by the parallel routes and reaction cycles method while the total degree of freedom equals the number of solutions (base vectors) computed by the null-space method. Reduction of the degrees of freedom reduces a number of possible steady states of the model.
10. There are characteristic conserved moieties like in metabolic networks, for example, amount of NAD + amount of NADH = const. Only biologically meaningful conserved moieties (e.g., NTP/NDP [(N)ucleotides], NAD(P)H/NAD(P), acetyl-CoA/CoA/succinyl-CoA, CoQ/CoQH2, phosphocreatine/creatin, THF/10-formyl-THF/5,10-methenyl-THF/5,10-methylene-THF) should remain in the network. All biologically nontypical conserved moieties must be eliminated from the network setup in order to fit the biological sense.
11. Use the blocked transformation steps method for identifying network regions which are not capable of operating at steady state due to the irreversibility constraints of the transformer equations.
12. Some transformers can be organized into linear pathways and, therefore, can be lumped to an overall transformer in order to reduce the model complexity without changing the degrees of freedom.
13. Computation of base vectors of the null-space immediately reveals "functioning" and "nonfunctioning" regions of a cellular network. When setting up cellular networks, null-space analysis assists for setting the balance status of compounds. Note that the null-space base vectors are not unique. This means that although the number of base vectors always corresponds with the degree of freedom, the base vectors of two test runs may therefore differ from each other. In addition, reversibility criteria of individual reaction steps are normally not fulfilled. For obtaining unique solutions of network functions compute elementary flux modes.
14. Parallel routes and reaction cycles increase the number of elementary flux modes substantially without leading to new phenotypic behavior. Remove parallel routes and reaction cycles if you are interested in maximal product yields rather than potential bypasses. Note that metabolic fluxes in parallel routes/reaction cycles cannot be observed by experimentally determined uptake and secretion rates. Therefore, another way

to get rid of the structural uncertainties due to presence of parallel routes and cycles is to use experimentally derived intracellular fluxes for any transformer that is part of the parallel route and cycle to constrain the problem (e.g., ATP hydrolysis for maintenance, ^{13}C-flux).

15. Computing elementary flux modes in large cellular systems with many degrees of freedom is a complex problem. Try to decrease the degrees of freedom by removing transport steps from/to compounds of minor interest and/or removing parallel routes/reaction cycles.
16. Essential transformation steps are transformers essential for maintaining a rate like the production rate of a biomass. That is, essential transformation steps cannot be replaced by other transformation steps, which means that if an essential step is deleted a certain function is unsustainable.
17. FBA can give a large number of mathematically acceptable solutions to the steady-state problem. However, solutions of biological interest are the ones which produce the desired metabolites in the correct proportion. The objective function defines the proportion of these metabolites. For instance, when modeling the growth of an organism the objective function is generally defined as biomass. Therefore, the experimentally derived stoichiometry of biomass constituents becomes important. In FBA, the objective function is subjected to either maximization or minimization depending on the goal.
18. Examples of FBA constraints [38] for different scenarios (add-on constraints). The set of constraints is coherent with the model assumptions (*see* **Note 2**). The objective function—maximization of biomass formation under light growth phase assumes that:
    (a) The biomass production rate is constrained (T.Biomass.ext <= 5000) (*see* **Note 19**).
    (b) A plant consumes CO_2 as the only carbon source (T.CO2.ext >= 0).
    (c) All formed starch becomes a part of the biomass (T.Starch.ext = 0).
    (d) There is no starch degradation under light and therefore starch kinase is inactive (SK = 0).
    (e) Photorespiration flux was fixed at 20% of flux through RuBisCo (0.25*RPC_plastide—RPC2_plastide = 0) at all tested conditions.
    (f) Additionally, the impact of different degrees of cyclic electron flow-through photosynthesis light reactions (particularly through ferredoxin-plastoquinone reductase, FQR) was estimated by means of add-on constraints:

- Either . . . cyclic electron flow is inactive (FQR = 0)
- Or . . . flux through ATP/ADP translocator between plastid and cytoplasm is inactive (T.ADP.plastid—T. ATP.plastid = 0), and thus plastid's ATP balance becomes self-sufficient
- Or . . . flux through FQR varies relative to the noncyclic electron flow through ferredoxin-NADP+-oxidoreductase (FNR) ((0 . . . 0.5)*FNR—FQR = 0)
- Or . . . flux through photorespiration pathway is zero (RPC2_plastide = 0; GLYK = 0), but cyclic electron flow through FQR is subjected to optimization

19. An unbounded problem can occur when constraints do not restrict objective function: for example, when you maximize your biomass production and your inputs are not restricted or when a combination of reactions allows unlimited biomass production, for example running futile cycles from reversible reactions in backward direction.

Acknowledgments

The author acknowledges Prof. Dr. Waltraud X. Schulze (Plant Systems Biology, University of Hohenheim, Stuttgart, Germany) for the collaboration in this research, which has been supported by a research grant of the German Research foundation (DFG). This publication has been supported by Research Council of Norway grant 248792, project of Digital Life Norway in which Norwegian University of Life Sciences (NMBU) is the node for Digital Production Biology.

References

1. Taiz L, Zeiger E (2010) Plant physiology, 5th edn. Sinauer Associates, Inc., Sunderland
2. Scialdone A, Mugford ST, Feike D, Skeffington A, Borrill P, Graf A, Smith AM, Howard M (2013) Arabidopsis plants perform arithmetic division to prevent starvation at night. elife 2:e00669. https://doi.org/10.7554/eLife.00669
3. Chapin FS, Schulze ED, Mooney HA (1990) The ecology and economics of storage in plants. Annu Rev Ecol Syst 21:423–447
4. Seaton DD, Ebenhöh O, Millar AJ, Pokhilko A (2014) Regulatory principles and experimental approaches to the circadian control of starch turnover. J R Soc Interface 11(91):20130979. https://doi.org/10.1098/rsif.2013.0979
5. Caspar T, Huber SC, Somerville C (1985) Alterations in growth, photosynthesis, and respiration in a starchless mutant of *Arabidopsis thaliana* (L.) deficient in chloroplast phosphoglucomutase activity. Plant Physiol 79:1–7
6. Schulze W, Stitt M, Schulze ED, Neuhaus HE, Fichtner K (1991) A quantification of the significance of assimilatory starch for growth of *Arabidopsis thaliana* L. Heynh. Plant Physiol 95:890–895
7. Schulze W, Schulze ED, Stader J, Heilmeier H, Stitt M, Mooney HA (1994) Growth and reproduction of *Arabidopsis thaliana* in relation to storage of starch and nitrate in the wild-type and in starch-deficient and nitrate-uptake-deficient mutants. Plant Cell Environ 17:795–809

8. Yazdanbakhsh N, Sulpice R, Graf A, Stitt M, Fisahn J (2001) Circadian control of root elongation and C partitioning in *Arabidopsis thaliana*. Plant Cell Environ 34:877–894
9. Milne RJ, Byrt CS, Patrick JW, Grof CP (2013) Are sucrose transporter expression profiles linked with patterns of biomass partitioning in Sorghum phenotypes? Front Plant Sci 4:223. https://doi.org/10.3389/fpls.2013.00223
10. Lemoine R, La Camera S, Atanassova R, Dédaldéchamp F, Allario T, Pourtau N, Bonnemain J-L, Laloi M, Coutos-Thévenot P, Maurousset L, Faucher M, Girousse C, Lemonnier P, Parrilla J, Durand M (2013) Source to sink transport and regulation by environmental factors. Front Plant Sci 4:272. https://doi.org/10.3389/fpls.2013.00272
11. Ludewig F, Flügge UI (2013) Role of metabolite transporters in source-sink carbon allocation. Front Plant Sci 4:231. https://doi.org/10.3389/fpls.2013.00231
12. Kühn C, Grof CPL (2010) Sucrose transporters of higher plants. Curr Opin Plant Biol 13 (3):287–297. https://doi.org/10.1016/j.pbi.2010.02.001
13. Chen L-Q (2014) SWEET sugar transporters for phloem transport and pathogen nutrition. New Phytol 201(4):1150–1155. https://doi.org/10.1111/nph.12445
14. Shiratake K (2007) Genetics of sucrose transporter in plants. In: Genes, genomes and genomics, Global Science Books, vol 1. Global Science Books, Ltd., Ikenobe, pp 73–80
15. Boorer KJ, Loo DDF, Frommer WB, Wright EM (1996) Transport mechanism of the cloned potato H+/sucrose cotransporter StSUT1. J Biol Chem 271:25139–25144
16. Gahrtz M, Stolz J, Sauer N (1994) A phloem-specific sucrose-H+ symporter from *Plantago major* L. supports the model of apoplastic phloem loading. Plant J 6:697–706
17. Flügge U-I, Häusler RE, Ludewig F, Gierth M (2011) The role of transporters in supplying energy to plant plastids. J Exp Bot 62 (7):2381–2392. https://doi.org/10.1093/jxb/erq361
18. Zimmermann P, Hirsch-Hoffman M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiol 136:2621–2632
19. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana* development. Nat Genet 37(5):501–506. https://doi.org/10.1038/ng1543
20. Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J 50(2):347–363. https://doi.org/10.1111/j.1365-313X.2007.03052.x
21. Gottwald JR, Krysan PJ, Young JC, Evert RF, Sussman MR (2000) Genetic evidence for the *in planta* role of phloem-specific plasma membrane sucrose transporters. Proc Natl Acad Sci U S A 97(25):13979–13984
22. Stadler R, Truernit E, Gahrtz M, Sauer N (1999) The AtSUC1 sucrose carrier may represent the osmotic driving force for anther dehiscence and pollen tube growth in Arabidopsis. Plant J 19:269–278
23. Chen LQ, Qu XQ, Hou BH, Sosso D, Osorio S, Fernie AR, Frommer WB (2012) Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. Science 35:207–211
24. Shachar-Hill Y (2013) Metabolic network flux analysis for engineering plant systems. Curr Opin Biotechnol 24(2):247–255. https://doi.org/10.1016/j.copbio.2013.01.004
25. Sweetlove LJ, Obata T, Fernie AR (2014) Systems analysis of metabolic phenotypes: what have we learnt? Trends Plant Sci 19 (4):222–230. https://doi.org/10.1016/j.tplants.2013.09.005
26. Baghalian K, Hajirezaei M-R, Schreiber F (2014) Plant metabolic modeling: achieving new insight into metabolism and metabolic engineering. Plant Cell Online 26 (10):3847–3866. https://doi.org/10.1105/tpc.114.130328
27. Poolman MG, Assmus HE, Fell DA (2004) Applications of metabolic modelling to plant metabolism. J Exp Bot 55(400):1177–1186. https://doi.org/10.1093/jxb/erh090
28. Arnold A, Nikoloski Z (2014) Bottom-up metabolic reconstruction of Arabidopsis and its application to determining the metabolic costs of enzyme production. Plant Physiol 165 (3):1380–1391. https://doi.org/10.1104/pp.114.235358
29. de Oliveira Dal'Molin CG, Quek L-E, Palfreyman RW, Brumbley SM, Nielsen LK (2010) AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. Plant Physiol 152(2):579–589. https://doi.org/10.1104/pp.109.148817
30. Vanrolleghem PA, Heijnen J (1995) Metabolic network modelling: improving predictions of microbial metabolism by maximal

incorporation of knowledge on biochemical reaction stoichiometry. In: Proceedings 9th Forum for Applied Biotechnology, Med. Fac. Landbouww. Univ. Gent, 60/4a, 1933–1940

31. Thiele I, Palsson BO (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc 5(1):93–121. https://doi.org/10.1038/nprot.2009.203
32. Stitt M, Sulpice R, Keurentjes J (2010) Metabolic networks: how to identify key components in the regulation of metabolism and growth. Plant Physiol 152(2):428–444. https://doi.org/10.1104/pp.109.150821
33. Mintz-Oron S, Meir S, Malitsky S, Ruppin E, Aharoni A, Shlomi T (2012) Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. Proc Natl Acad Sci 109(1):339–344. https://doi.org/10.1073/pnas.1100358109
34. Poolman MG, Miguet L, Sweetlove LJ, Fell DA (2009) A genome-scale metabolic model of *Arabidopsis* and some of its properties. Plant Physiol 151(3):1570–1581. https://doi.org/10.1104/pp.109.141267
35. Cheung CYM, Ratcliffe RG, Sweetlove LJ (2015) A method of accounting for enzyme costs in flux balance analysis reveals alternative pathways and metabolite stores in an illuminated Arabidopsis leaf. Plant Physiol 169 (3):1671–1682. https://doi.org/10.1104/pp.15.00880
36. Gomes De Oliveira Dal'molin C, Quek L-E, Saa PA, Nielsen LK (2015) A multi-tissue genome-scale metabolic modelling framework for the analysis of whole plant systems. Front Plant Sci 6:4. https://doi.org/10.3389/fpls.2015.00004
37. Grafahrend-Belau E, Junker A, Eschenröder A, Müller J, Schreiber F, Junker BH (2013) Multiscale metabolic modeling: dynamic flux balance analysis on a whole-plant scale. Plant Physiol 163(2):637–647. https://doi.org/10.1104/pp.113.224006
38. Zakhartsev M, Medvedeva I, Orlov Y, Akberdin I, Krebs O, Schulze WX (2016) Metabolic model of central carbon and energy metabolisms of growing *Arabidopsis thaliana* in relation to sucrose translocation. BMC Plant Biol 16(1):262. https://doi.org/10.1186/s12870-016-0868-3
39. Poolman MG, Fell DA, Thomas S (2000) Modelling photosynthesis and its control. J Exp Bot 51(suppl 1):319–328. https://doi.org/10.1093/jexbot/51.suppl_1.319
40. Arnold A, Nikoloski Z (2011) A quantitative comparison of Calvin–Benson cycle models. Trends Plant Sci 16(12):676–683. https://doi.org/10.1016/j.tplants.2011.09.004
41. Sajitz-Hermstein M, Nikoloski Z (2010) A novel approach for determining environment-specific protein costs: the case of *Arabidopsis thaliana*. Bioinformatics 26(18):i582–i588. https://doi.org/10.1093/bioinformatics/btq390
42. Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG (2008) Metabolic network fluxes in heterotrophic Arabidopsis cells: stability of the flux distribution under different oxygenation conditions. Plant Physiol 148(2):704–718. https://doi.org/10.1104/pp.108.125195
43. Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG (2010) Subcellular flux analysis of central metabolism in a heterotrophic *Arabidopsis* cell suspension using steady-state stable isotope labeling. Plant Physiol 152(2):602–619. https://doi.org/10.1104/pp.109.151316
44. Piques M, Schulze WX, Höhne M, Usadel B, Gibon Y, Rohwer J, Stitt M (2009) Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis. Mol Syst Biol 5(1). https://doi.org/10.1038/msb.2009.68
45. Lee CP, Gu Q, Xiong Y, Mitchell RA, Ernster L (1996) P/O ratios reassessed: mitochondrial P/O ratios consistently exceed 1.5 with succinate and 2.5 with NAD-linked substrates. FASEB J 10(2):345–350
46. Hoefnagel MHN, Atkin OK, Wiskich JT (1998) Interdependence between chloroplasts and mitochondria in the light and the dark. Biochim Biophys Acta 1366(3):235–255. https://doi.org/10.1016/S0005-2728(98)00126-1
47. Pänke O, Rumberg B (1997) Energy and entropy balance of ATP synthesis. Biochim Biophys Acta 1322(2–3):183–194. https://doi.org/10.1016/S0005-2728(97)00079-0
48. Pänke O, Rumberg B (1999) Kinetic modeling of rotary CF0F1-ATP synthase: storage of elastic energy during energy transduction. Biochim Biophys Acta 1412(2):118–128. https://doi.org/10.1016/S0005-2728(99)00059-6
49. Badger MR (1985) Photosynthetic oxygen exchange. Annu Rev Plant Physiol 36 (1):27–53. https://doi.org/10.1146/annurev.pp.36.060185.000331
50. Portis AR, McCarty RE (1976) Quantitative relationships between phosphorylation, electron flow, and internal hydrogen ion concentrations in spinach chloroplasts. J Biol Chem 251(6):1610–1617

51. Wolstencroft K, Owen S, Krebs O, Nguyen Q, Stanford NJ, Golebiewski M, Weidemann A, Bittkowski M, An L, Shockley D, Snoep JL, Mueller W, Goble C (2015) SEEK: a systems biology data and model management platform. BMC Syst Biol 9(1):1–12. https://doi.org/10.1186/s12918-015-0174-y
52. Zhang P, Foerster H, Tissier CP, Mueller L, Paley S, Karp PD, Rhee SY (2005) MetaCyc and AraCyc. Metabolic pathway databases for plant research. Plant Physiol 138(1):27–37. https://doi.org/10.1104/pp.105.060376
53. Karp PD, Paley SM, Krummenacker M, Latendresse M, Dale JM, Lee TJ, Kaipa P, Gilham F, Spaulding A, Popescu L, Altman T, Paulsen I, Keseler IM, Caspi R (2010) Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. Brief Bioinform 11(1):40–79. https://doi.org/10.1093/bib/bbp043
54. Huala E, Dickerman AW, Garcia-Hernandez-M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang M, Huang W, Mueller LA, Bhattacharyya D, Bhaya D, Sobral BW, Beavis W, Meinke DW, Town CD, Somerville C, Rhee SY (2001) The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res 29 (1):102–105
55. Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P, Huala E (2008) The Arabidopsis Information Resource (TAIR): gene structure and function annotation. Nucleic Acids Res 36:D1009–D1014
56. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014) Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 42 (D1):D199–D205. https://doi.org/10.1093/nar/gkt1076
57. Hastings J, de Matos P, Dekker A, Ennis M, Harsha B, Kale N, Muthukrishnan V, Owen G, Turner S, Williams M, Steinbeck C (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acids Res 41(D1): D456–D463. https://doi.org/10.1093/nar/gks1146