BIOINFORMATICS ORIGINAL PAPER

Vol. 21 no. 7 2005, pages 1219–1226
doi:10.1093/bioinformatics/bti145

*Systems biology*

# Application of Petri net theory for modelling and validation of the sucrose breakdown pathway in the potato tuber

Ina Koch[1,*], Björn H. Junker[2] and Monika Heiner[3]

[1]Department of Bioinformatics, Technical University of Applied Sciences Berlin, Seestrasse 64, Berlin, 13347, Germany, [2] Institute of Plant Genetics and Crop Plant Research, Corrensstrasse 3, Gatersleben, 06466, Germany and [3]Computer Science Department, Brandenburg University of Technology at Cottbus, PO 10 13 44, Cottbus, 03013, Germany

Received on June 7, 2004; revised and accepted on November 9, 2004
Advance Access publication November 16, 2004

**ABSTRACT**

**Motivation:** Because of the complexity of metabolic networks and their regulation, formal modelling is a useful method to improve the understanding of these systems. An essential step in network modelling is to validate the network model. Petri net theory provides algorithms and methods, which can be applied directly to metabolic network modelling and analysis in order to validate the model. The metabolism between sucrose and starch in the potato tuber is of great research interest. Even if the metabolism is one of the best studied in sink organs, it is not yet fully understood.

**Results:** We provide an approach for model validation of metabolic networks using Petri net theory, which we demonstrate for the sucrose breakdown pathway in the potato tuber. We start with hierarchical modelling of the metabolic network as a Petri net and continue with the analysis of qualitative properties of the network. The results characterize the net structure and give insights into the complex net behaviour.

**Availability:** Free availability of the Petri net editor PED, the animator PedVisor via http://www-dssz.informatik.tu-cottbus.de/∼wwwdssz, and the analysis tool Integrated Net Analyser (INA) via http://www.informatik.hu-berlin.de/∼starke/ina.html

**Contact:** Ina.Koch@tfh-berlin.de

## INTRODUCTION

Systems Biology becomes more and more important with the drastically growing amount of not only sequence and structure data, but also of data on interacting proteins and gene expression, provided by the rapidly developing high-throughput technologies. Metabolic networks are a central paradigm in biology. The interactions of metabolites and enzymes characterize these networks and are responsible for the system function.

While kinetic models based on differential equations are able to simulate networks quantitatively, there are clear limitations to such models, so long as the data are incomplete. Even if the stoichiometric equations of each single reaction of a network are described in the literature, the detailed kinetic laws and constants are often incomplete or entirely missing. Generally, there are two approaches to overcome this problem: either the kinetic parameters of the taxonomically nearest neighbour species are used, for which they are available, or they are estimated. The less reliable information is available for a network of interest, the more the simulation results are questionable. Especially for large and/or dense network models, qualitative analysis can check the model for consistency and biologically meaningful behaviour by giving insights into the basic system behaviour.

The metabolic network investigated in this paper concerns the main carbon metabolism in *Solanum tuberosum* (potato) tubers. The conversion of sucrose through hexose phosphates to starch is the major flux in potato tuber metabolism. All enzymatic reactions of this metabolism are well characterized. Many transgenic experiments were performed highlighting the flexible and complex nature inherent to plant metabolism (ap Rees and Morell, 1990; Fernie *et al*., 2002b; Geigenberger, 2003).

In order to analyse this metabolism qualitatively we use methods originating from Petri net theory. Petri nets have been founded by Carl Adam Petri in his dissertation in 1962 to describe and simulate networks with causal concurrent processes. Since this time, many theorems and algorithms have been developed and implemented to analyse such systems (Murata, 1989; Starke, 1990). Applications of this theory comprise the analysis of technical systems, administrative systems and others. First applications of Petri net theory to biological systems were published about ten years ago (Reddy *et al*., 1993). Meanwhile, metabolic networks, signal transduction pathways and gene-regulatory networks have been modelled and analysed successfully using Petri net methods (see, e.g. Hofestädt, 1994; Reddy *et al*., 1996; Hofestädt and Thelen, 1998; Küffner *et al*., 2000; Tsavachidou and Liebman, 2002; Voss *et al*., 2003; Matsuno *et al*., 2003; Heiner *et al*., 2004). However, the vast majority of these papers deal with quantitative approaches.

Other existing methods for structural analysis of biological networks, like elementary modes (Schuster *et al*., 1993) or flux balance analysis (Edwards and Palsson, 2000), derive steady-state-related dynamic properties. They fail, generally, in the analysis of other important structural and more general dynamic properties of the given network.

*To whom correspondence should be addressed.

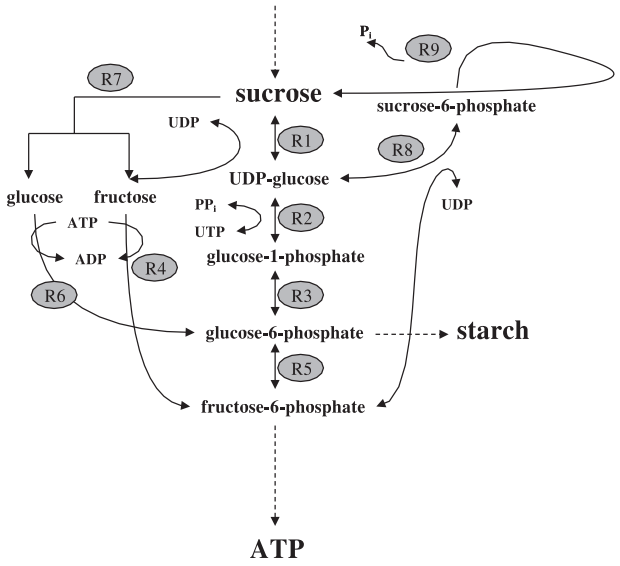



**Fig. 1.** The sucrose breakdown metabolism. For reaction numbers see text and Table 1.

In this paper, we explain the qualitative modelling of the carbon metabolism in the potato tuber as a Petri net and the validation process in detail. After a short informal introduction to Petri nets, we discuss model validation techniques for the sucrose breakdown metabolism using sound Petri net theory. Basic structural properties as well as dynamic behaviour will be discussed.

## METHODS

### The sucrose breakdown metabolism in potato tuber

The outline of the biological model is drawn in Figure 1. Sucrose delivered to the tuber can be cleaved by invertase in the apoplast or by either invertase or sucrose synthase in the cytosol. Invertase dominates during the early stages of tuber initiation, whereas sucrose synthase activities are high in the developed tuber (Appeldoorn *et al*., 1999).

The products of sucrose cleavage by invertase (R7), glucose and fructose, enter the invertase pathway by concerted action of fructokinase (R4) and hexokinase (R6) (Veramendi *et al*., 1999, 2002). The products of sucrose cleavage by sucrose synthase (R1), fructose and UDP-glucose, enter the sucrose synthase pathway by joint action of fructokinase (R4) and UDP-glucose pyrophosphorylase (R2). The formed hexoses are in equilibrium through the action of cytosolic isoforms of phosphoglucose isomerase (R5) and phosphoglucomutase (R3) (Fernie *et al*., 2002a). It is proven that also the process of sucrose (re)synthesis via two different pathways is possible within this tissue. One pathway is the reverse action of sucrose synthase (R1). The other one consists of the reactions catalysed by sucrose phosphate synthase (R8) and sucrose phosphate phosphatase (R9) (Geigenberger and Stitt, 1991).

The structure of a metabolic network is defined by its stoichiometric chemical equations. The reaction equations given in Table 1 characterize the sucrose breakdown metabolism. The reaction numbers according to Figure 1 are given in the first column, and the enzymes catalysing the reactions in the second column. Abbreviations for enzymes, metabolites and summarized reactions are shown in Table 2.

We extend the biological model in Figure 1 by reactions given implicitly, and by those, which represent the interface of the system to the environment. Thus, R12 and R14 realize the input of sucrose (the source) and the output of starch (the sink), respectively. R11 and R13 are reactions summarizing energy consuming reactions, while R15 and R16 represent the known

**Table 1.** The stoichiometric reaction system of the metabolism in Figure 1

| | | |
|---|---|---|
| R1 | SuSy | $Suc + UDP \leftrightarrows UDPglc + Frc$ |
| R2 | UGPase | $UDPglc + PP \leftrightarrows G1P + UTP$ |
| R3 | PGM | $G6P \leftrightarrows G1P$ |
| R4 | FK | $Frc + ATP \rightarrow F6P + ADP$ |
| R5 | PGI | $G6P \leftrightarrows F6P$ |
| R6 | HK | $Glc + ATP \rightarrow G6P + ADP$ |
| R7 | Inv | $Suc \rightarrow Glc + Frc$ |
| R8 | SPS | $F6P + UDPglc \leftrightarrows S6P + UDP$ |
| R9 | SPP | $S6P \rightarrow Suc + P_i$ |
| R10 | Glyc(b) | $F6P + 29\ ADP + 28\ P_i \rightarrow 29\ ATP$ |
| R11 | NDPkin | $UDP + ATP \leftrightarrows UTP + ADP$ |
| R12 | SucTrans | $eSuc \rightarrow Suc$ |
| R13 | ATPcons(b) | $ATP \rightarrow ADP + P_i$ |
| R14 | StaSy(b) | $G6P + ATP \rightarrow starch + ADP + PP$ |
| R15 | AdK | $ATP + AMP \leftrightarrows 2\ ADP$ |
| R16 | PPase | $PP \rightarrow 2P_i$ |

**Table 2.** Abbreviations used for enzymes, metabolites, summarized reactions and environment reactions

| | |
|---|---|
| Enzymes | |
| AdK | Adenylate kinase |
| FK | Fructokinase |
| HK | Hexokinase |
| Inv | Invertase |
| NDPkin | NDPkinase |
| PGI | Phosphoglucose isomerase |
| PGM | Phosphoglucomutase |
| PPase | Pyrophosphatase |
| SPP | Sucrose phosphate phosphatase |
| SPS | Sucrose phosphate synthase |
| SucTrans | Sucrose transporter |
| SuSy | Sucrose synthase |
| UGPase | UDP-glucose pyrophosphorylase |
| Summarized reactions | |
| ATPcons(b) | ATP consumption |
| Glyc(b) | Glycolysis |
| StaSy(b) | Starch synthesis |
| Environment (artificial) reactions | |
| geSuc | Generate eSuc |
| rStarch | Remove starch |
| Metabolites | |
| ADP | Adenosine diphosphate |
| AMP | Adenosine monophosphate |
| ATP | Adenosine triphosphate |
| eSuc | External sucrose |
| F6P | Fructose-6-phosphate |
| Frc | Fructose |
| G1P | Glucose-1-phosphate |
| G6P | Glucose-6-phosphate |
| Glc | Glucose |
| $P_i$ | Phosphate ion |
| PP | Pyrophosphate |
| S6P | Sucrose-6-phosphate |
| starch | Starch |
| Suc | Sucrose |
| UDP | Uridine diphosphate |
| UDPglc | Uridine diphosphate glucose |
| UTP | Uridine triphosphate |

AMP–ADP–ATP conversion, which is not specific for this metabolism. In Glyc(b) (R10), ATPcons(b) (R13) and StaSy(b) (R14), the symbol 'b' stands for bulk reactions indicating such a reaction is an abstract sum reaction of many other reactions, which are not included in our model, because these reactions represent complex systems. Glycolysis (R10) was modelled to yield only 29 instead of 36 ATP, because ∼20% of fructose-6-phosphate is utilized in the pentose phosphate cycle (ap Rees and Beevers, 1960).

## Modelling

A Petri net is a directed, bipartite and labelled graph. A Metabolic Petri net, called MPN, is a place/transition net, where the places (P) represent biological compounds (metabolites) and the transitions (T) represent the biochemical reactions between metabolites, which are catalysed by a certain enzyme that gives its name to the transition. Places can contain movable objects, the so-called tokens, which represent quantitative units (e.g. 1 mole) of a chemical compound. Arcs are allowed only between places and transitions and vice versa, but never between places or between transitions. The arc labels (multiplicities, weights) correspond to the stoichiometric numbers in the reaction equations.

Fluxes are modelled as token flows through the firing of transitions. A transition is enabled to fire, if the token amount of all pre-places satisfies the token amount required by the labels of the arcs connecting these pre-places with the considered transition. When the reaction takes place (the transition fires), the multiplicity of a given arc indicates the token amount, which has to flow through the arc from a pre-place to the transition or from the transition to a post-place, respectively.

The distribution of tokens over the places represents the marking of the net and characterizes a certain state of the system, i.e. distribution of compounds. The initial marking is the marking of the start state for any animation or marking-dependent analysis. For more detailed explanations see Heiner *et al*., 2004.

## Model validation

Model validation aims at checking the constructed model for system inconsistencies and deriving statements on structural and dynamic system properties, which reflect correctly the behaviour of the system in reality. In the following we introduce those model properties in terms of Petri net terminology, which are expected to be meaningful for metabolic network analysis. Afterwards, they are applied to our case study. Reading the text below, please bear in mind that transitions correspond to reactions, places to metabolites and marking/system states to token/substance distributions.

*Dynamic properties* The three orthogonal dynamic properties of a Petri net are liveness, reversibility and boundedness. An important expected property of a MPN is the infinite net behaviour in the sense that no part of the network will ever stop working, if a sufficient amount of input substrates enters the net. An unintended interruption of the substance fluxes is likely to indicate a modelling error. In order to validate the net we check it for dead transitions. A transition is dead at a given system state, if no state can be reached anymore, where the transition is enabled. A transition is live at a given state, if no state can be reached, where the transition is dead. A net is live, if all its transitions are live in the initial marking. This includes that there is no dead state reachable, where no transition at all is enabled, i.e. no unexpected total interruption of fluxes can occur. A net is reversible, if the initial system state can be reached again from each reachable state. This excludes obviously irreversible system behaviour like unexpected burning out or accumulation of certain compounds.

Let us consider boundedness, which is essential for any steady state network analysis. A net is bounded, if there exists a positive integer number, $k$, which represents an upper bound for the number of tokens on each place in all states of the net. If the net is bounded in every initial marking, it is said to be structurally bounded. If a net is bounded, the number of reachable states is finite.

Let us turn to analysis techniques to decide dynamic properties. Basically, three different ways of net analyses can be distinguished, first, the structural analysis, second, the invariant analysis and third, the reachability analysis.

## Structural analysis

Structural analysis aims at discovering certain net structures that allow conclusions on dynamic properties. It starts with the calculation of elementary properties, as e.g. if the net is ordinary, homogenous, conservative, pure, static conflict-free and connected or strongly connected in graph-theoretical sense. The first three properties give information on stoichiometric relations of biochemical networks. A net is ordinary, if all arc weights, i.e. the stoichiometric numbers in the chemical reaction equations, are equal to one. A net is homogenous, if, for any place, all arcs starting at this place have the same weights. These two properties hold typically in signal transduction pathways. A net is conservative, if for each transition the sum of input arc weights is equal to the sum of output arc weights. Therefore, all reactions work in a token-preserving way. If a net is conservative, it is bounded. A net is pure, if no transition exists, for which a pre-place is also a post-place. MPNs are pure as long as the reaction enzymes are not modelled explicitly. A net is static conflict-free, if there are no transitions sharing a pre-place. A MPN is usually not free of static conflicts, because compounds may be used by several reactions. Static conflicts may result in dynamic conflicts, establishing alternative system behaviour, if the tokens in shared places are not sufficient for the concurrent firing of all their enabled post-transitions. A network representing a biochemical metabolism should be connected, but must not be strongly connected.

Other interesting structural properties are structural deadlocks and traps. A set of places, which cannot get a new token as soon as it has lost all its tokens, is called a structural deadlock. In this case, every transition, which puts tokens into the place set, has also a pre-place in this set. Therefore, an insufficient marking of this place set makes all its post-transitions to dead ones. In MPNs, structural deadlocks indicate sets of compounds requiring a sufficient initial amount of tokens to make the net live. In biological context it means that the modelled metabolic network cannot run successfully without some metabolites being available at certain concentrations. The opposite situation that not all tokens can be removed from a set of places, is called a trap. This occurs, if every transition, which subtracts tokens from this place set, has also a post-place in this set, and thus returns always at least one token to this set. In biochemical systems, traps can represent sets of compounds for metabolite storage.

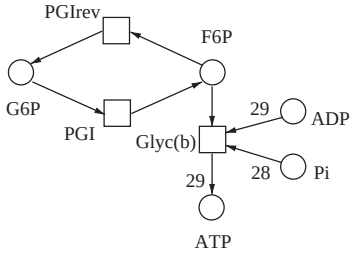
## Invariant analysis

The invariant analysis was introduced by Lautenbach (1973) and comprises the structural computation of dynamic net properties. Usually, an invariant is a property, which is valid in all states while the system is working. A transition-invariant (T-invariant) is a (multi-) set of transitions, the firing of which reproduces a given system state. In biological context, a T-invariant represents a set of chemical reactions reproducing a given distribution of chemical compounds, e.g. the steady state, and all T-invariants are likely to occur at the same time.

A T-invariant corresponds to a possible pathway through the network, which is defined by the net representation of the T-invariant, i.e. a subnet of the whole network consisting of the transitions belonging to the T-invariant, all their pre- and post-places and all arcs in between. If there is a T-invariant, which covers all transitions, then all reactions may contribute to some pathway.

Please note, we do not use the graph-theoretical term path for the interpretation of T-invariants, but the term pathway. A path in graph-theoretical sense is defined as a sequence of consecutive arcs through a given graph. In contrast, a pathway can exhibit branches describing alternatives and concurrencies in the corresponding subnet representation. A pathway can degenerate to a path.

A place-invariant (P-invariant) is a set of places, where the weighted sum of tokens remains constant while the net is working. P-invariants are structural deadlocks as well as traps and represent conservation relationships. Hence, if there is a P-invariant, which covers all places, then the




| C | PGI | PGIrev | Glyc(b) |
|---|---|---|---|
| G6P | –1 | 1 | 0 |
| F6P | 1 | –1 | –1 |
| ADP | 0 | 0 | –29 |
| ATP | 0 | 0 | 29 |
| Pi | 0 | 0 | –28 |

**Fig. 2.** The incidence matrix for a subnet of our Petri net model consisting of three transitions and five places.

net is bounded. In biological models, groups of certain metabolites, such as adenylates (AMP+ADT+ATP) or uridylates (UMP+UDP+UTP) often form P-invariants.

To compute net invariants we need the incidence matrix of the net, which corresponds to the stoichiometric matrix of the reaction system. The incidence matrix $C = P \times T$ of a place/transition net is defined as an integer matrix, where the places are listed as rows and the transitions as columns. The matrix elements contain the change of the token amount on a place, when the corresponding transition fires (Fig. 2). Please note, a non-pure net structure is not uniquely defined by its incidence matrix. T-invariants (P-invariants) are non-trivial, non-negative integer solutions of the linear equation system $C \cdot y = 0$ ($x \cdot C = 0$), where $C$ is the incidence matrix and $y$ ($x$) the solution vector. Therefore, the results do not depend on the initial marking. Owing to the meaning of a token flow, only those integer solutions are of interest, where all numbers of a solution vector are positive integers. These generally infinite sets of positive integer solutions can be obtained by calculating the uniquely defined basis of the solution space.

Such a basis consists of the so-called minimal invariants, which are pairwise linearly independent from each other, and the greatest common divisor of all elements of a solution vector is one. Minimal T-invariants correspond basically to the elementary modes. Each T-invariant, i.e. each possible pathway through the network, can be calculated from the minimal invariants by using only addition, multiplication with a non-negative integer and division by the greatest common divisor of all components.

In the following we consider only minimal invariants, which are therefore called invariants for short.

### Reachability analysis

The reachability analysis is based on the construction of the reachability graph or just the set of reachable states of the net. The result depends generally on the initial marking. The nodes of a reachability graph represent all possible states of the net. Thus, it can be finite only for bounded nets. The arcs connecting nodes of a reachability graph are labelled by single firing transitions and give altogether a graph representation of all possible single-step firing sequences. Consequently, concurrent behaviour is described by enumerating all interleaving firing sequences.

Realizable T-invariants, where the current marking actually allows the firing of the transitions of the T-invariant, correspond to cycles in the reachability graph, possibly to a cycle of length one (a firing transition without marking change). Realizable T-invariants with inherent concurrency correspond to several cycles in the reachability graph, one for each interleaving sequence. Based on the reachability graph, many detailed dynamic properties can be decided, e.g. if the reachability graph is strongly connected, then the net is reversible. Summarily, there exist various algorithms for reachability analysis, using different principles and data structures.

Models of biochemical systems are often unbounded in the first place. To decide unboundedness, an always finite reduced version of the reachability graph exists, the so-called coverability graph. It is generally possible to transform an unbounded system into a bounded one, still reflecting the steady state behaviour, by adding artificial nodes and arcs, which remove the output substances and convert them into input substances in correspondence with the total equation of the network. But often, even if the system is bounded, the large and complex biochemical networks exhibit too many states to be represented explicitly, or even symbolically.

*Upscalability* To date there exists no method in systems biology to compute all reachable states in really complex networks. The complexity of any algorithm to construct explicitly the reachability or coverability graph is known to be not primitive recursive. Even more sophisticated representation techniques, like symbolic state space representations, are of exponential complexity in the worst case. Hence, for smaller networks, especially for signal transduction pathways, all system states might be computable. In our example, like in other metabolic networks, the high stoichiometric numbers (e.g. R10) cause an excessive amount of reachable system states. This underlines the necessity of alternative analysis techniques as discussed in the two former subsections, which have in common the total avoidance of any state space construction. While most of the structural analysis techniques may be decided very efficiently, the computation of structural deadlocks, traps and minimal invariants is of exponential complexity. Concerning the latter problem, there are promising approaches to exploit modularization or incremental algorithms, which are still under evaluation.

### Software tools

The design of the model has been done using the graphical Petri net editor PED (Tiedemann, 1997). Model animation is supported by the simulator PedVisor (Menzel, 1996). The analysis of structural and dynamic properties is done by INA, The Integrated Net Analyser (Starke and Roch, 1999, http://www.informatik.hu-berlin.de/~starke/ina.html).

## RESULTS AND DISCUSSION

### The Petri net model of the sucrose breakdown

The Petri net model of the sucrose breakdown consists of 17 places and 25 transitions (Fig. 3). We model the reversible reactions as forward and backward reactions abstracted by so-called hierarchical transitions, which are introduced to improve the readability. They are drawn as two concentric squares (Fig. 4). Both forward and backward transitions are considered in the qualitative analysis. However, in the quantitative analysis, forward and backward reactions exhibit generally different reaction rates. Here, the reaction with the higher reaction rate is defined as forward reaction. The model contains seven hierarchical transitions: SuSy, SPS, PGI, PGM, NDPkin, UGPase and AdK.

The transitions ATPcons(b), StaSy(b) and Glyc(b) describe processes consisting of many subprocesses, which are not modelled and analysed here in detail. Therefore, they are not drawn as hierarchical transitions. Substrates as ATP and ADP participate in several reactions, such that representing them by only one place would yield many crossing arcs, resulting in an unreadable picture. Therefore, such places (UDP, $P_i$, PP, ATP and ADP) are modelled as logical places (grey-filled places), which are drawn several times, but refer always to the same place, i.e. compound.

The environment is modelled by an artificial transition without pre-places (geSuc) for the input of external sucrose into the system, and by another artificial transition without post-places (rStarch) for removing starch from the system. According to the firing rule, transitions without pre-places are always enabled, so they can always fire. Thus, in the model it is assumed that sucrose is always available, and that produced starch can be consumed by or deposited in the cell. Because nets of this type are not bounded, they are not covered by P-invariants.

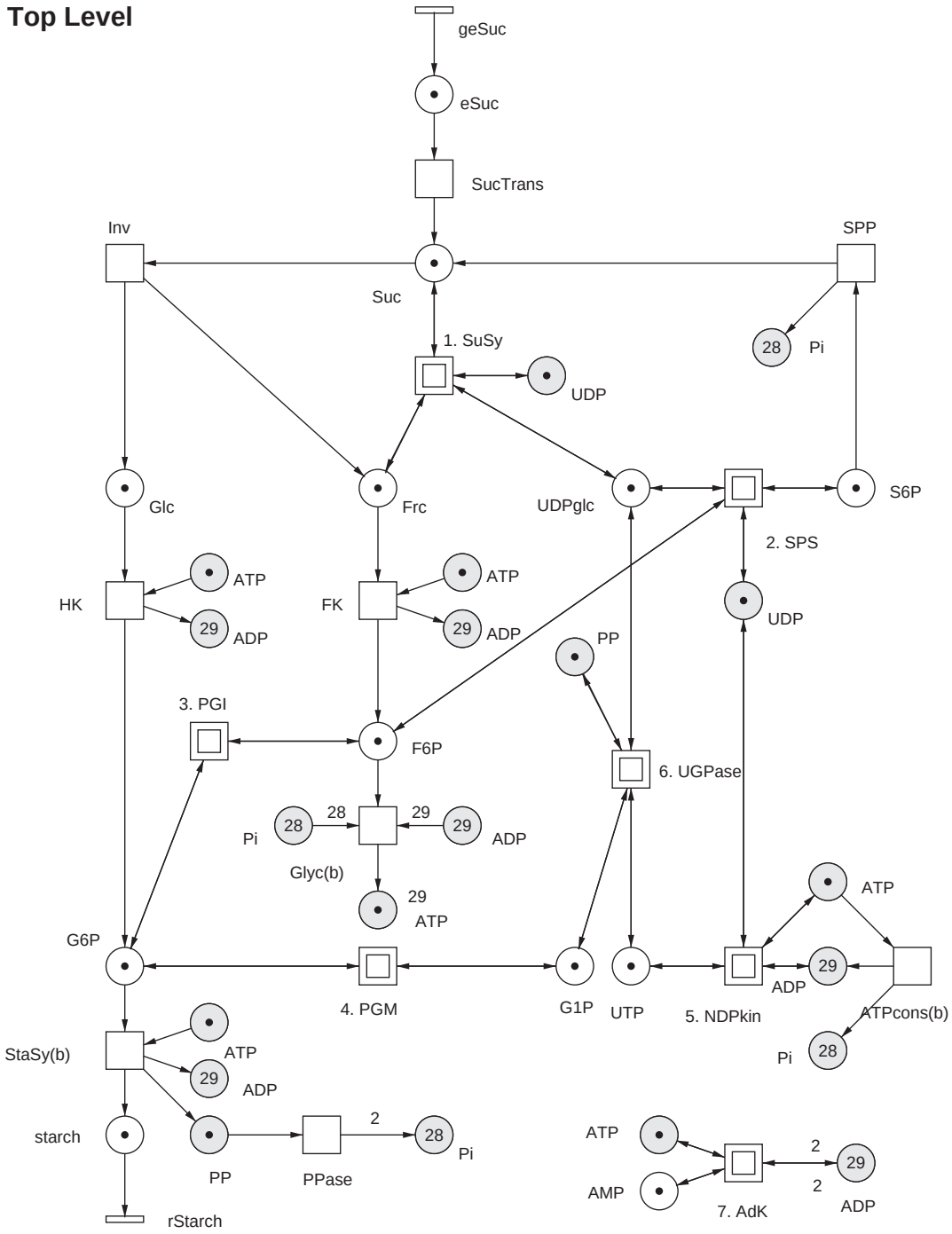



**Fig. 3.** The hierarchical Petri net model of the sucrose-to-starch metabolism in the potato tuber. The given marking reflects a state where each transition is enabled.

## Structural and reachability analysis

Our unbounded net in Figure 3 is not ordinary, not homogenous and not conservative. It is pure, but not static conflict free, and it is connected, but not strongly connected. Besides the P-invariants discussed below the net does not contain any structural deadlocks or traps.

Since we consider an unbounded model, we cannot compute the reachability graph. In addition, the computation of the coverability graph for this relatively small network is not successful, because of the huge amount of possible states (here, more than four millions). Even when constructing a smaller and bounded net version by summarizing the hexoses into one place, the set of reachable states

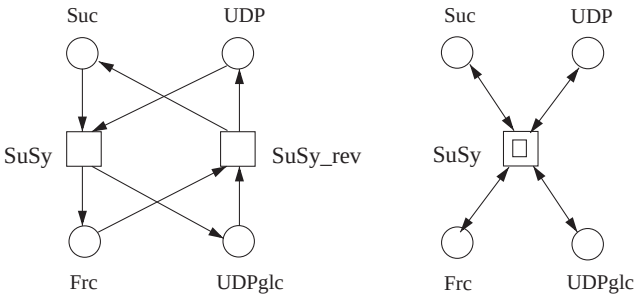



**Fig. 4.** The underlying structure of a hierarchical transition. Both forward and backward reactions are considered. For reversible reactions the backward reactions are indicated explicitly by an attached '_rev'.

consists of more than $10^{10}$ states (computed by a prototype of a symbolic analysis tool). Based on this reachability set, we can decide liveness and reversibility for the bounded model.

*P-invariant analysis*: The net contains, but is not covered by the following P-invariants:

(1) *Invariant A*: UDPglc, UTP, UDP.

(2) *Invariant B*: ATP, AMP, ADP.

(3) *Invariant C*: G6P, F6P, G1P, UTP, ATP(2), ADP, S6P, $P_i$, PP(2).

Invariants A and B comprise the metabolites containing uridine or adenosine residues, respectively. As the synthesis and degradation of nucleotides is not included in the model, the sum of these metabolites is not variable, which is likely to reflect the actual biological system behaviour. The turnover in utilization and regeneration of adenylates as energy donors is much higher than its biosynthesis and degradation. Invariant C represents the set of all compounds, which provide directly or indirectly a phosphate group. If the phosphate group is transferred from one compound to another, the sum of the phosphorylated metabolites is unchanged. If no phosphate is taken up or secreted by a cell, the sum of phosphate groups in all metabolites, including inorganic phosphate itself, will not change.

*T-invariant analysis*: The net is covered by T-invariants. For our model we obtain 19 T-invariants, 7 of them are trivial ones. They consist of only two transitions and indicate the reversible reactions.

*Trivial invariants*: (1) SPS, SPS_rev; (2) UGPase, UGPase_rev; (3) SuSy, SuSy_rev; (4) PGM, PGM_rev; (5) NDPkin, NDPkin_rev; (6) AdK, AdK_rev; and (7) PGI, PGI_rev.

Each of the 12 non-trivial T-invariants represent a combination of different sub-pathways. The incoming sucrose can be either cleaved by invertase or by sucrose synthase. The resulting hexose phosphates can be metabolized in glycolysis, which is the case in all non-trivial T-invariants, or can be incorporated into starch. The ATP produced by glycolysis can be utilized for ATP consumption by cellular processes other than starch synthesis, or be wasted by one of the three futile cycles. These are sucrose cleavage by invertase and sucrose (re)synthesis by (1) SPS/SPP or (2) reverse sucrose synthase, and (3) cleavage by sucrose synthase and sucrose (re)synthesis by SPS/SPP.

The non-trivial T-invariants are listed below. The firing counters given in parentheses indicate how often a reaction must fire to obtain this invariant, i.e. they give the relative firing rates in the steady state. Reactions without parentheses take place once.

## Invariants with sucrose cleavage by sucrose synthase

*Invariant 8*: geSuc, SucTrans, SuSy(29), UGPase, PGM_rev, FK(29), Glyc(b), StaSy(b), rStarch, SPS(28), SPP(28), NDPkin_rev.

*Invariant 9*: geSuc, SucTrans, SuSy, UGPase, PGM_rev, FK, Glyc(b), StaSy(b), rStarch, ATPcons(b)(28), NDPkin_rev.

*Invariant 10*: geSuc(15), SucTrans(15), SuSy(15), PGI_rev(14), UGPase(15), PGM_rev(15), FK(15), Glyc(b), StaSy(b)(29), rStarch(29), NDPkin_rev(15), PPase(14).

Sucrose is always cleaved by sucrose synthase. One hexose goes into glycolysis, and the other hexoses go into starch synthesis. In invariant 8 the produced ATP is utilized in sucrose cycling through SuSy, SPS and SPP, whereas in invariant 9 the produced ATP is consumed by other cellular processes. In invariant 10 all the other hexoses are incorporated into starch (firing counter 29). This invariant represents the main flux in the potato tuber.

## Invariants with sucrose cleavage by invertase

*Invariant 11*: geSuc, SucTrans, Inv(14), UGPase_rev(13), PGM(13), HK(14), FK, Glyc(b), StaSy(b), rStarch, SuSy_rev(13), NDPkin(13), PPase(14).

*Invariant 12*: geSuc(3), SucTrans(3), Inv(29), UGPase_rev(26), PGM(26), HK(29), FK(29), Glyc(b)(3), StaSy(b)(3), rStarch(3), SPS(26), SPP(26), NDPkin(26), PPase(29).

*Invariant 13*: geSuc, SucTrans, Inv, HK, FK(27), Glyc(b), StaSy(b), rStarch, SuSy(26), SPS(26), SPP(26), PPase.

*Invariant 14*: geSuc, SucTrans, Inv, HK, FK, Glyc(b), StaSy(b), rStarch, ATPcons(b)(26), PPase.

*Invariant 15*: geSuc(15), SucTrans(15), Inv(15), HK(15), FK(15), PGI_rev(13), Glyc(b)(2), StaSy(b)(28), rStarch(28), PPase(28).

*Invariant 16*: geSuc, SucTrans, Inv(29), HK(29), FK, PGI, UGPase_rev(28), PGM(28), Glyc(b)(2), SuSy_rev(28), NDPkin(28), PPase(28).

*Invariant 17*: geSuc(3), SucTrans(3), Inv(59), HK(59), FK(59), UGPase_rev(56), PGM(56), PGI(3), Glyc(b)(6), SPS(56), SPP(56), NDPkin(56), PPase(56).

*Invariant 18*: geSuc, SucTrans, Inv, HK, FK(57), PGI, Glyc(b)(2), SuSy(56), SPS(56), SPP(56).

*Invariant 19*: geSuc, SucTrans, Inv, HK, FK, PGI, Glyc(b)(2), ATPcons(b)(56).

Sucrose is always cleaved by invertase. The resulting hexoses can go into glycolysis and starch synthesis (invariants 11–15). In contrast to invariants 11–14, where the majority of produced hexoses are incorporated in a special type of ATP consumption, in invariant 15 two hexoses go into glycolysis and all the other hexoses are incorporated into starch (firing counter 28). Invariant 15 represents a bypass of invariant 10, the main flux, with sucrose cleavage using invertase instead of sucrose synthase. The remaining sub-pathways (sucrose to starch) in invariants 10 and 15 are qualitatively the same. In invariants 16–19 all resulting hexoses go into glycolysis.

Additionally, the invariants differ by their ATP utilization. Whereas ATP is consumed by other cellular processes (invariants 14 and 19), it can also be consumed by sucrose cycling through invertase and the reverse sucrose synthase reaction (invariants 11 and 16), through SuSy, SPS and SPP (invariants 13 and 18), and through invertase, SPS and SPP (invariants 12 and 17).

An interesting observation is the fact that the reaction catalyzed by AdK and SPS_rev do not occur in a non-trivial T-invariant. That

| invariant no. | Suc cleavage Susy | Suc cleavage Inv | hexoses go into Glyc | hexoses go into StaSy | ATP cons | ATP used for cycling Inv Susy_rev | ATP used for cycling Inv SPS, SPP | ATP used for cycling SuSy SPS, SPP |
|---|---|---|---|---|---|---|---|---|
| 8 | X | | X | X | | | | X |
| 9 | X | | X | X | X | | | |
| 10 | X | | X | X | | | | |
| 11 | | X | X | X | | X | | |
| 12 | | X | X | X | | | X | |
| 13 | | X | X | X | | | | X |
| 14 | | X | X | X | X | | | |
| 15 | | X | X | X | | | | |
| 16 | | X | X | | | X | | |
| 17 | | X | X | | | | X | |
| 18 | | X | X | | | | | X |
| 19 | | X | X | | X | | | |

**Fig. 5.** The distribution of the sub-pathways over the T-invariants.

means that removing AdK and/or SPS_rev would not influence the system behaviour. This is an example for new insights into the network behaviour by Petri net modelling and analysis, which would not have been obtained by contemplating or reasoning.

### Sub-pathways analysis

Let us consider the following biologically motivated sub-pathways: two possibilities of sucrose cleavage, the two ways for hexoses and the energy usage by five different ways. The combination of sub-pathways results in 20 theoretically possible pathways. Figure 5 gives an overview of the sub-pathways and their occurrence in the invariants. All 12 non-trivial T-invariants represent a combination of these sub-pathways. It is noticeable that the two possibilities for hexose utilization (starch synthesis and glycolysis) can both be present together in one invariant (invariants 8–15), while starch synthesis cannot occur without glycolysis, because no cellular process can take place without the availability of energy in the form of ATP or other similar cofactors.

Instead of the 20 combinatory possible invariants we get 12 non-trivial T-invariants, which means that 8 are missing. Two of these are the combinations of SuSy, starch synthesis and sucrose cycling, via invertase and the reverse sucrose synthase reaction, or via invertase and SPS/SPP. Five other missing invariants are the combination of SuSy with glycolysis, but without starch synthesis, together with one of the five possible sub-pathways for ATP utilization. The last missing invariant would describe an ATP production, but without starch synthesis and any way of ATP utilization, which makes no sense.

Apparently, the incoming sucrose that is needed to generate ATP in glycolysis, can only be cleaved by invertase or sucrose synthase, but not by both at the same time. This is also the case, if more than one sucrose after cleavage is used in glycolysis (invariants 12 and 17). The cleavage of sucrose by invertase and sucrose synthase can only be present together in one invariant, if sucrose cycling is involved (invariants 13 and 18).

Summarizing the discussion above it can be said that all T-invariants describe processes with biologically meaningful interpretation as well as all missing combinatorial possibilities of sub-pathways do not result in T-invariants, because they make no sense in the biological context.

## CONCLUSIONS

This paper describes the application of Petri net theory to model and validate the sucrose breakdown metabolism in the potato tuber. Petri nets represent not only a unique formalism to depict metabolic (and also other biochemical) networks, but also provide techniques, which can be applied for qualitative validation of the model before starting a quantitative analysis. An especially valuable technique is the calculation of T-invariants as well as P-invariants of a system. T-invariants reflect the main processes, i.e. pathways, taking place in the metabolic system in reality, while P-invariants reflect substance preservations. The presented detailed discussion of occurring T-invariants explains the net behaviour as possible combinations of sub-pathways, which correctly mirrors experimentally known results. Altogether, we give a validated model of sucrose breakdown, which provably corresponds to the current knowledge. Thus, it can serve as a sound basis for further investigations, such as kinetic analyses.

In the future, the existing net should be extended by other central metabolic processes (e.g. glycolysis, respiration and amino acid metabolism), first of all to get deeper insights into the whole metabolism in the potato tuber, and, as a side effect, to scrutinize Petri net methods for more complex networks.

## ACKNOWLEDGEMENTS

We would like to thank Thomas Runge for fruitful discussions, the anonymous referees for their valuable comments (5). This work is partly supported by the Federal German Ministry of Education and Research (BMBF), BCB project 0312705D.

## REFERENCES

ap Rees,T. and Beevers,H. (1960) Pathways of glucose dissimilation in carrot slices. *Plant Physiol.*, **35**, 830–838.

ap Rees,T. and Morell,S. (1990) Carbohydrate metabolism in developing potatoes. *Am. Potato J.*, **67**, 835–847.

Appeldoorn,N., de Bruijn,S., Koot-Gronsveld,E., Visser,R., Vreugdenhil,D. and van der Plas,L. (1999) Developmental changes in enzymes involved in the conversion of hexose phosphates and its subsequent metabolites during early tuberization of potato. *Plant Cell Environ.*, **22**, 1085–1096.

Edwards,J. and Palsson,B. (2000) Metabolic flux balance analysis and the *in silico* analysis of *Escherichia coli* K-12 gene deletions. *BMC Bioinformatics*, **1**(1) 1, Epub July 27 2000.

Fernie,A., Tauberger,E., Lytovchenko,A., Roessner,U., Willmitzer,L. and Trethewey,R. (2002a) Antisense repression of cytosylic phosphoglucomutase in potato (*Solanum tuberosum*) results in a severe growth retardation, reduction, in tuber number and altered carbon metabolism. *Planta*, **214**, 510–520.

Fernie,A., Willmitzer,L. and Trethewey,R. (2002b) Sucrose to starch: a transition in molecular plant physiology. *Trends Plant Sci.*, **7**, 35–41.

Geigenberger,P. (2003) Regulation of sucrose to starch conversion in growing potato tubers. *J. Exp. Bot.*, **54**, 457–465.

Geigenberger,P. and Stitt,M. (1991) Sucrose synthase catalyses a readily reversible reaction *in vivo* in developing potato tubers and other plant systems. *Planta*, **189**, 329.

Heiner,M., Koch,I. and Will,J. (2004) Model validation of biological pathways using Petri nets—demonstrated for apoptosis. *BioSystems*, **75** (Special issue), 15–28.

Hofestädt,R. (1994) A Petri net application of metabolic processes. *J. Syst. Anal. Modell. Simul.*, **16**, 113–122.

Hofestädt,R. and Thelen,S. (1998) Quantitative modelling of biochemical networks. *In Silico Biol.*, **1**, 39–53.

Küffner,R., Zimmer,R. and Lengauer,T. (2000) Pathway analysis in metabolic databases via differential metabolic display (dmd). *Bioinformatics*, **16**, 825–836.

Lautenbach,K. (1973) Exact liveness conditions of a Petri net class (in German). *GMD Report 82*, Bonn, Germany.

Matsuno,H., Fujita,S., Doi,A., Nagasaki,M. and Miyano,S. (2003) Towards pathway modelling and simulation. In Caradella,J. and Reioig,D. (eds) *Proceedings of the ICATPN 2004*, Bologna, Italy, June 2004, LNCS **2679**, pp. 3–22.

Menzel,T. (1996) Design and implementation of a Petri net tool kit framework integrating animation and simulation. *Major Individual Project (in German)*, BTU Cottbus, Department of CS.

Murata,T. (1989) Petri nets: properties, analysis and applications. In *Proc. IEEE*, **77**, 541–580.

Reddy,V., Mavrovouniotis,M. and Liebman,M. (1993) Petri net representation in metabolic pathways. In Hunter,H., Sears,D. and Shavlik,J. (eds) *Proceedings of the First International Conference on Intelligent Systems for Molecular Biology*, Bethesda, MD, USA, July 1993, AAAI Press, pp. 328–336.

Reddy,V., Mavrovouniotis,M. and Liebman,M. (1996) Qualitative analysis of biochemical reaction systems. *Comput. Biol. Med.*, **26**, 9–24.

Schuster,S., Hilgetag,C. and Schuster,R. (1993) Determining elementary modes of functioning in biochemical reaction networks at steady state. In Fritsh,R., Luitz,R.G., Behaira,M. and Behaira,M. (eds) *Proceedings of the Second Gauss Symposium*, Munich, Germany, August 1993, Walte de Gruyter Inc., pp. 101–114.

Starke,P. (1990) *Analysis of Petri Net Models (in German)*. B.G. Teubner, Stuttgart.

Starke,P. and Roch,S. (1999) *INA—The Intergrated Net Analyzer*. Humboldt University Berlin.

Tiedemann,R. (1997) PED—hierarchical Petri net editor (in German). *Internal Technical Report*, Department of CS, BTU Cottbus.

Tsavachidou,D. and Liebman,M. (2002) Modelling and simulation of pathways in the menopause. *JAMIA*, **9**, 461–871.

Veramendi,J., Fernie,A., Leisse,A., Willmitzer,L. and Trethewey,R. (2002) Potato hexokinase 2 complements transgenic *Arabidopsis* plants deficient in hexose 1 but does not play a key role in tuber carbon metabolism. *Plant Mol. Biol.*, **49**, 451–501.

Veramendi,J., Roessner,U., Renz,A., Willmitzer,L. and Trethewey,R. (1999) Antisense repression of hexokinase 1 leads to an overaccumulation of starch in leaves of transgenic potato plants but not to significant changes in tuber carbohydrate metabolism. *Plant Physiol.*, **121**, 123–133.

Voss,K., Heiner,M. and Koch,I. (2003) Steady state analysis of metabolic pathways using Petri nets. *In Silico Biol.*, **3**, 367–387.