Possibilities of Petri Net Theory to validate metabolic pathways

Ina Koch
Technical University of Applied Sciences Berlin
ina.koch@tfh-berlin.de
Monika Heiner
Brandenburg University of Technology at Cottbus
monika.heiner@informsatik.tu-cottbus.de

Bertinoro, 14th June 2004

Outline

- Introduction
- Petri net basics
- Analysis possibilities
- Sucrose-to-starch breakdown in the potato tuber
- Simulation of the net
- Conclusions

Introduction

Metabolic Control Analysis - MCA

Metabolic system: connected unit, steady state

MCA is based on solution of systems of differential equations

- MCA Kacser & Burns, *Symp.Soc.Exp.Bio.* (1973)
  Heinrich & Rapoport, *Eur.J.Biochem.* (1974)
- Biochemical systems theory

  Savageau, *J.Theor.Biol.* (1969)
- Flux oriented theory

  Crabtree & Newsholme, *Biochem.J.* (1987)

GEPASI Mendes, *Comp.Appl.Biosci.* (1993)

Introduction

Graph-Theory

- Hybrid graphs — Kohn & Letzkus, *J.Theor.Biol.* (1983)
- Bond graphs — Lefèvre & Barreto, *J.Franklin Inst.* (1985)
- Net-thermodynamics — Mikulecky, *Am.J.Physiol.* (1993)
  Weighted linear graphs — Goldstein & Shevelev, *J.Theor.Biol.* (1985)
  Goldstein & Selivanov, *Biomed.Biochim.Acta* (1990)
- Meta-nets (with gene expression systems)
  Kohn & Lemieux, *J.Theor.Biol.* (1991)
- Bipartite graphs — Zeigarnik & Temkin, *Kin.Catalysis* (1994)
- KING (KINetic Graphs) — Zeigarnik, *Kin.Catalysis* (1994)

Introduction

- Why is a model validation (check model consistency) useful?

- Before starting a quantitative analysis it should be sure that the model is valid.
- If the systems become larger with many interactions and regulations it could not be done manually any more.

- How model validation could be performed?
  By qualitative analysis

Basic dynamic properties: liveness, reversibility, boundedness, dead states, deadlocks, traps,

Basic structure properties: invariants, robustness, alternative pathways,

Petri net theory provides algorithms and tools to answer these questions.

etri net basics

Petri nets (PhD thesis of Carl Adam Petri 1962)
abstract models of information and control data flows, which allow to describe systems and processes at different abstraction levels and in a unique language
- developed for systems with causal concurrent processes

Applications: business processes, computer communication, automata theory, operating systems, software dependability

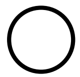
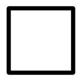
Biological networks: metabolic networks, signal transduction pathways, gene regulatory networks

Reddy, Mavrovouniotis, Liebman, *Proc. ISMB* (1993), *Comp.Mol.Med.* (1996)
Hofestädt, *J.Syst.Anal., Modell., Sim.* (1994), Hofestädt & Thelen, *In silico* Biol (1998)
Matsuno et al., *Proc.PSB* (2000), *In silico Biol.* (2003), *Proc.IACATPN* (2003) ,
Voss, Heiner, Koch, *BioSystems* (2004), Heiner, Koch, Will, *Proc.Comp.Methods Syst.Biol.* (2003)
Heiner, Voss, Koch, *In Silico Biology* (2003)

Metabolic Petri Net - MPN

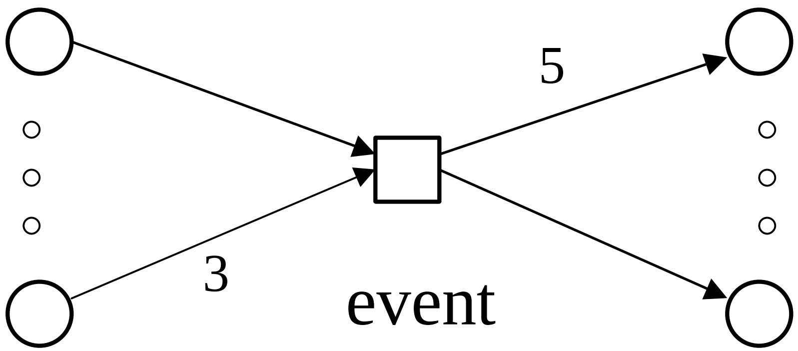
etri net basics

Petri nets: directed, labelled, bipartite graphs

Nodes: (vertiecs)	places	transitions
		
	passive elements	active elements
	conditions	events
	states	actions
	chemical compounds	chemical reactions
	metabolites	conversions of metabolites catalysed by enzymes

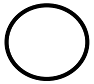
etri net basics

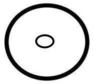
Arcs: pre-conditions post-conditions
(edges)

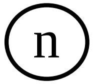


etri net basics

Tokens: movable objects in discrete units,
e.g. units of substances (mole)

 condition is not fulfilled

 condition is (one time) fulfilled



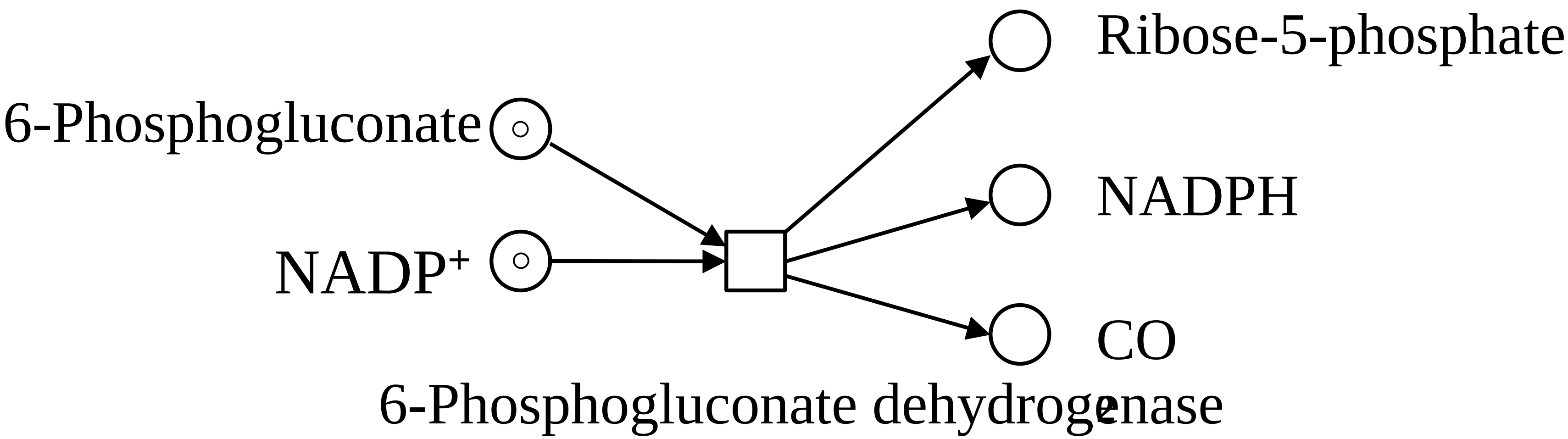
condition is n times fulfilled

Marking: system state, token distribution, initial marking

Token flow: occurring of an event (firing of a transition)

etri net basics

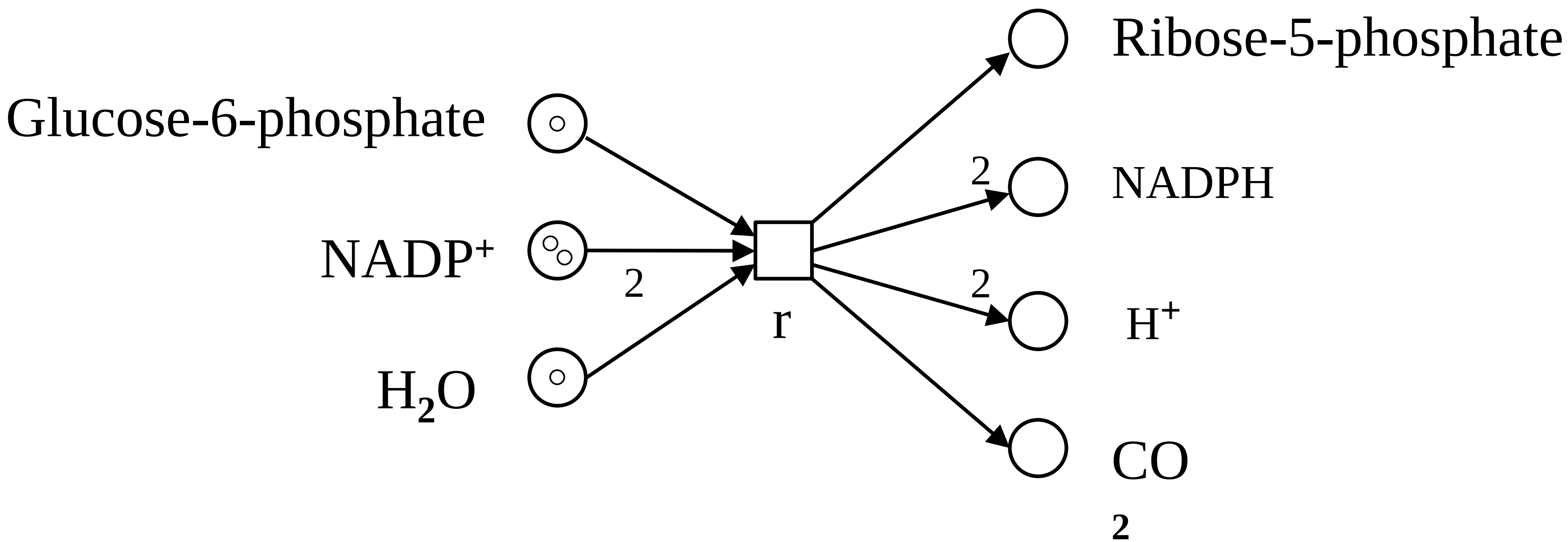
Example: Pentose Phosphate Pathway - one reaction



$6PG + NADP^{+} \rightarrow R5P + NADPH + CO_2$

etri net basics

Example: Pentose Phosphate Pathway - sum reaction



$$G6P + 2\ NADP^{+} + H_2O \rightarrow R5P + 2\ NADPH + 2\ H^{+} + CO_2$$

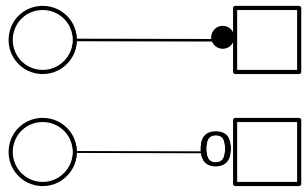
etri net basics

Special places:

input:	substrates	(source, e.g. sucrose)
output:	products	(sink, e.g. starch)

Special arcs: reading arcs

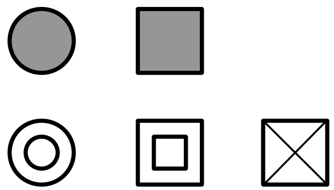
inhibitor arcs



Additional places & transitions:

logical

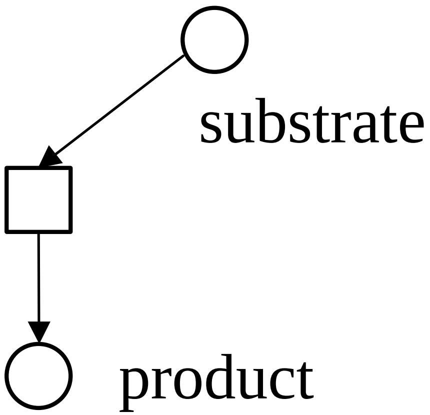
hierarchical



etri net basics

Transitions in MPNs:

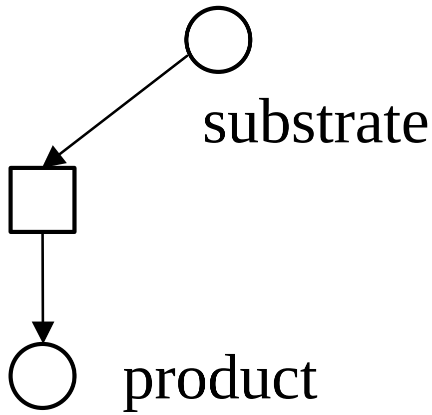
Reaction:



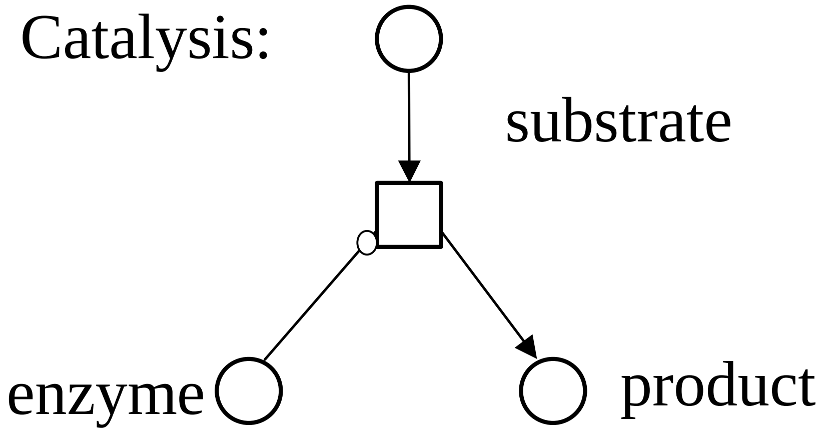
etri net basics

Transitions in MPNs:

Reaction:

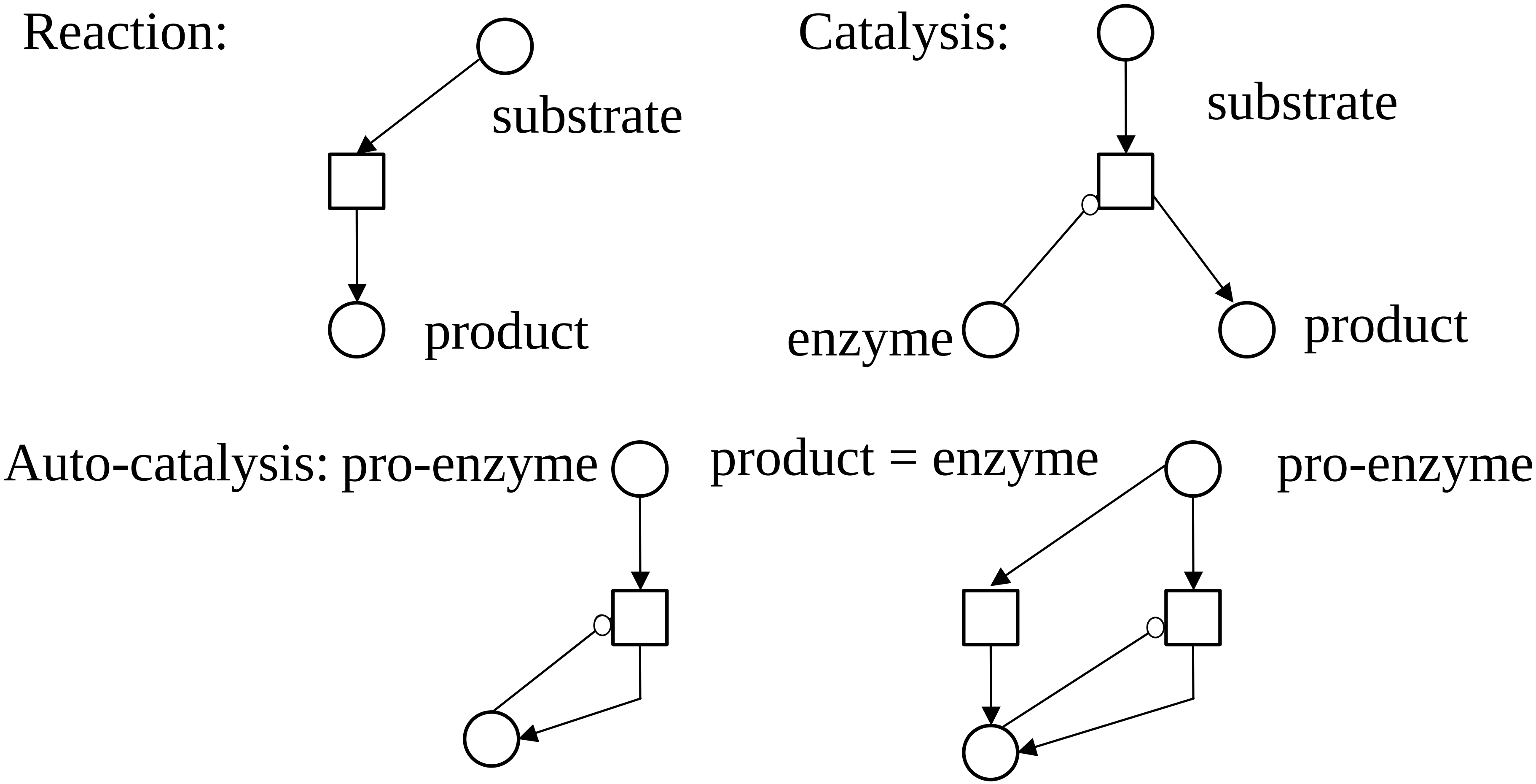


Catalysis:



etri net basics

Transitions in MPNs:



odel validation

(1) Dynamical (behavioural) properties

(2) Reachability analysis

(3) Structural analysis

(4) Invariant analysis

ynamic (behavioural) properties

liveness and reversibility

- a net is live, if all its transitions are live in the initial marking
- a net is reversible, if the initial marking can be reached from each possible state
- How often can a transition fire? (0-times, n-times, ∞ times)
- infinite systems behaviour, search for dead transitions
- prediction of system deadlocks

ynamic (behavioural) properties

boundedness

- a net is bounded, if there exists a positive integer number k, which represents a maximal number of tokens on each place in all states

- What is the maximal token number for a place?
  (0, 1, k, ∞) boundedness (k-bounded)

- for bounded nets special algorithms exist

eachability analysis

How many and which system states could be reached ? (0, 1, k, ∞)

- the reachability graph represents all possible states

- computational problem for large and dense biological networks

- for unbounded networks: computation of the coverability graph

- Is a certain system state again and again reachable?

progressiveness

- Is a certain system state never reachable? safety

tructural analysis

- aims at discovering net structures to derive conclusions on dynamic properties

Elementary properties:

ordinary: the multiplicity of every arc is equal one

homogeneous: for any place all outgoing edges have the same multiplicity

pure: there is no transition, for which a pre-place is also a post-place (loop-free)

conservative: for each place the sum of input arc weights is equal to the sum of output arc weights – a conservative net is bounded

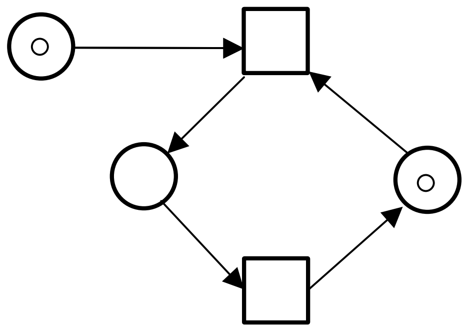
static conflict-free: there are no transitions with a common pre-place

connected, strongly connected: in graph-theoretical sense

tructural analysis

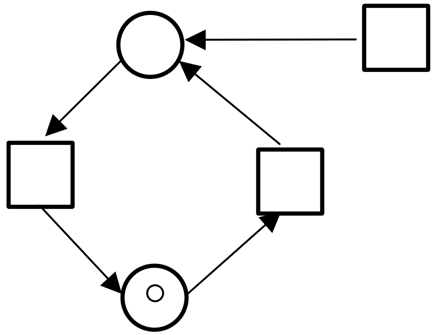
structural deadlock:

a set of places that delivers its tokens until a state is reached, where the place set is empty and there is no possibility to get a new token



trap:

the opposite situation that tokens cannot be removed from a place set



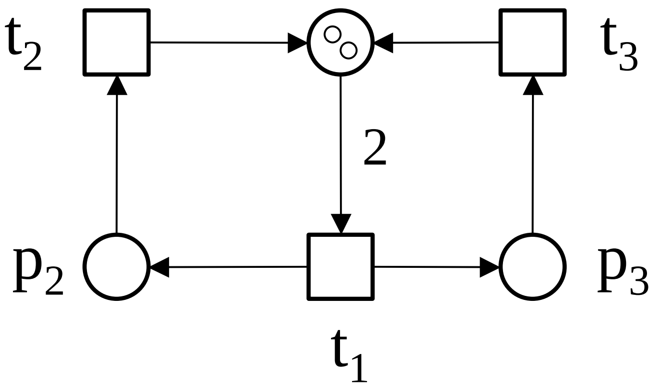
nvariant analysis

- properties, which are conserved during the working of the system
- independent of the initial marking
- only the net structure is relevant for their calculation

Are there invariant structures, which are independent from firing of the system?

Place-invariants (P-invariants)
Transition-invariants (T-invariants)

nvariant analysis



incidence matrix C = P x T

$$C = \begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \end{array} \begin{array}{c} \begin{array}{ccc} t_1 & t_2 & t_3 \end{array} \\ \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \end{array}$$

place (P-) invariant:

$$x\,C = 0$$

$$\begin{aligned} -2x_1 + x_2 + x_3 &= 0 \\ x_1 - x_2 \quad\quad &= 0 \\ x_1 \quad\quad - x_3 &= 0 \end{aligned}$$

transition (T-) invariant:

$$C\,y = 0$$

$$\begin{aligned} -2y_1 + y_2 + y_3 &= 0 \\ y_1 - y_2 \quad\quad &= 0 \\ y_1 \quad\quad - y_3 &= 0 \end{aligned}$$

nvariant analysis

Minimal semi-positive solutions are of interest with

- all components of the solution vector are ≥ 0
- basis of the semi-positive solution space such that none solution is contained in another solution, Lautenbach (1973)

The calculation

- of all integer solutions is in P
- of all semi-positive solutions is in P
- of all semi-positive integer solutions is NP-complete, Schrijver (1999)

Extreme Pathways, Schilling et al. (2000)

- minimal basis of semi-positive integer solutions (Hilbert-base)
- subset of T-invariants – biological interpretation?

nvariant analysis

Interpretation

P-invariants

- set of places, whose weighted sum of tokens is constant
- covered by P-invariants: sufficient condition for boundedness
- set of metabolites, whose total net concentrations remain unchanged

  ADP, ATP
  $NADP^+$, NADPH

T -invariants

- set of transitions, whose firing reproduces a given marking
- covered by T-invariants: necessary condition for liveness
- minimal set of enzymes which could operate at steady state
- indicate the presence of cyclic firing sequences

Elementary modes

Schuster, Hilgetag, Schuster (1993)

ucrose-to-starch-pathway in potato tuber

- rich in carbohydrates and energy
- a natural source of folate
- full of vitamin C
- low in calories
- good source of niacin, vitamin B6, iodine, thiamine, and minerals
- no cholesterol
- completely fat free

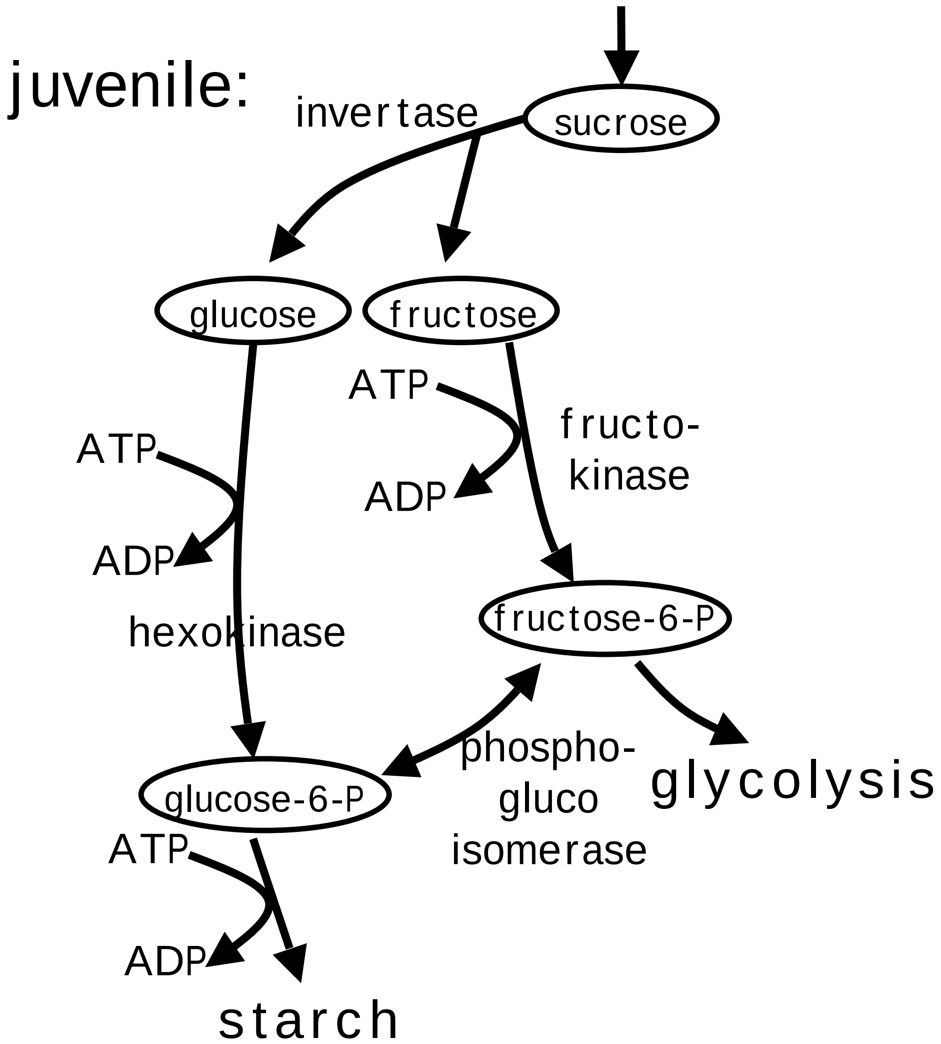
Fernie, Willmitzer, Trethewey (2002)

Research interest: increasing the starch content

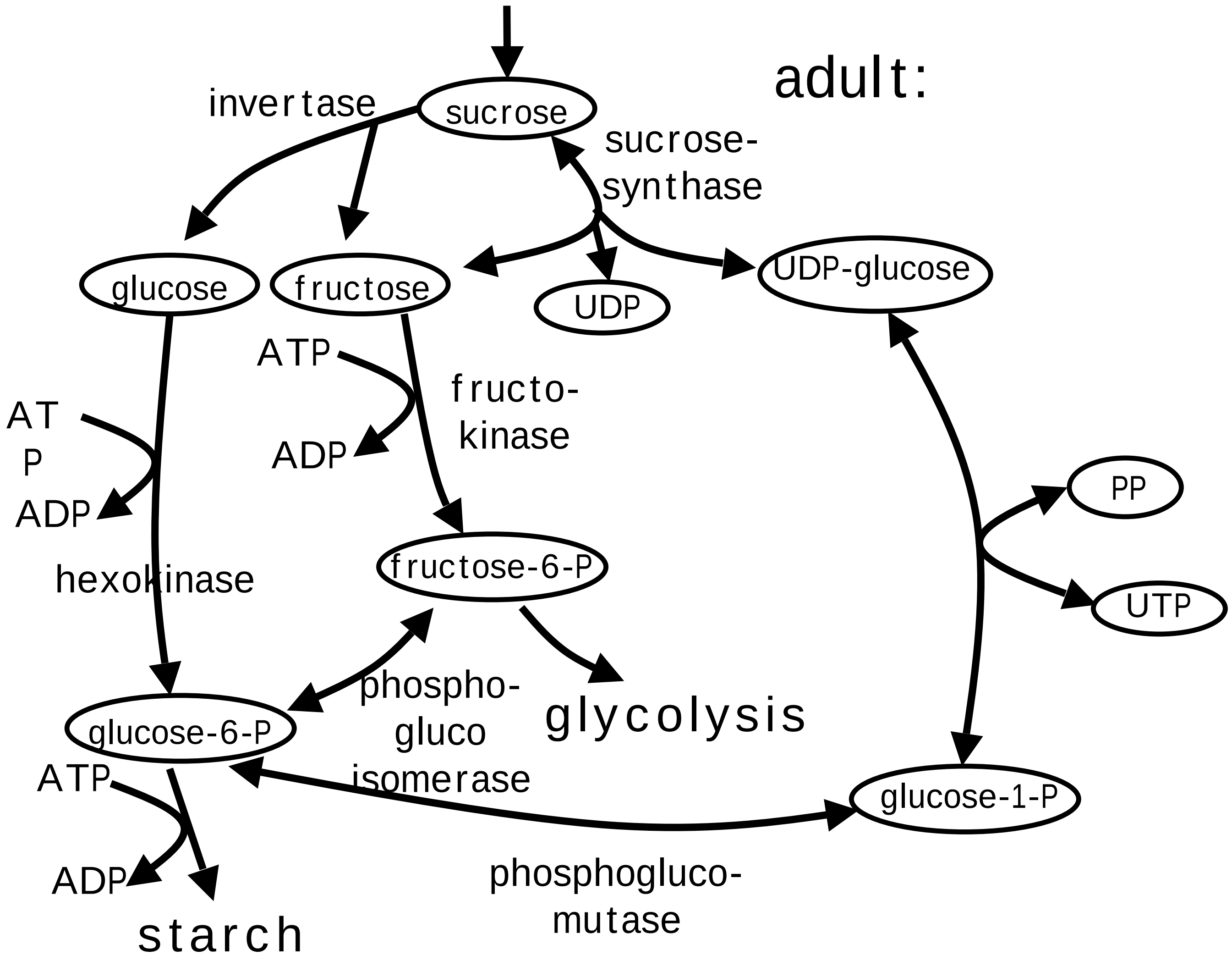
Co-operations: Max Planck Institute for Molecular Plant Physiology, Golm

Brandenburg University of Technology at Cottbus

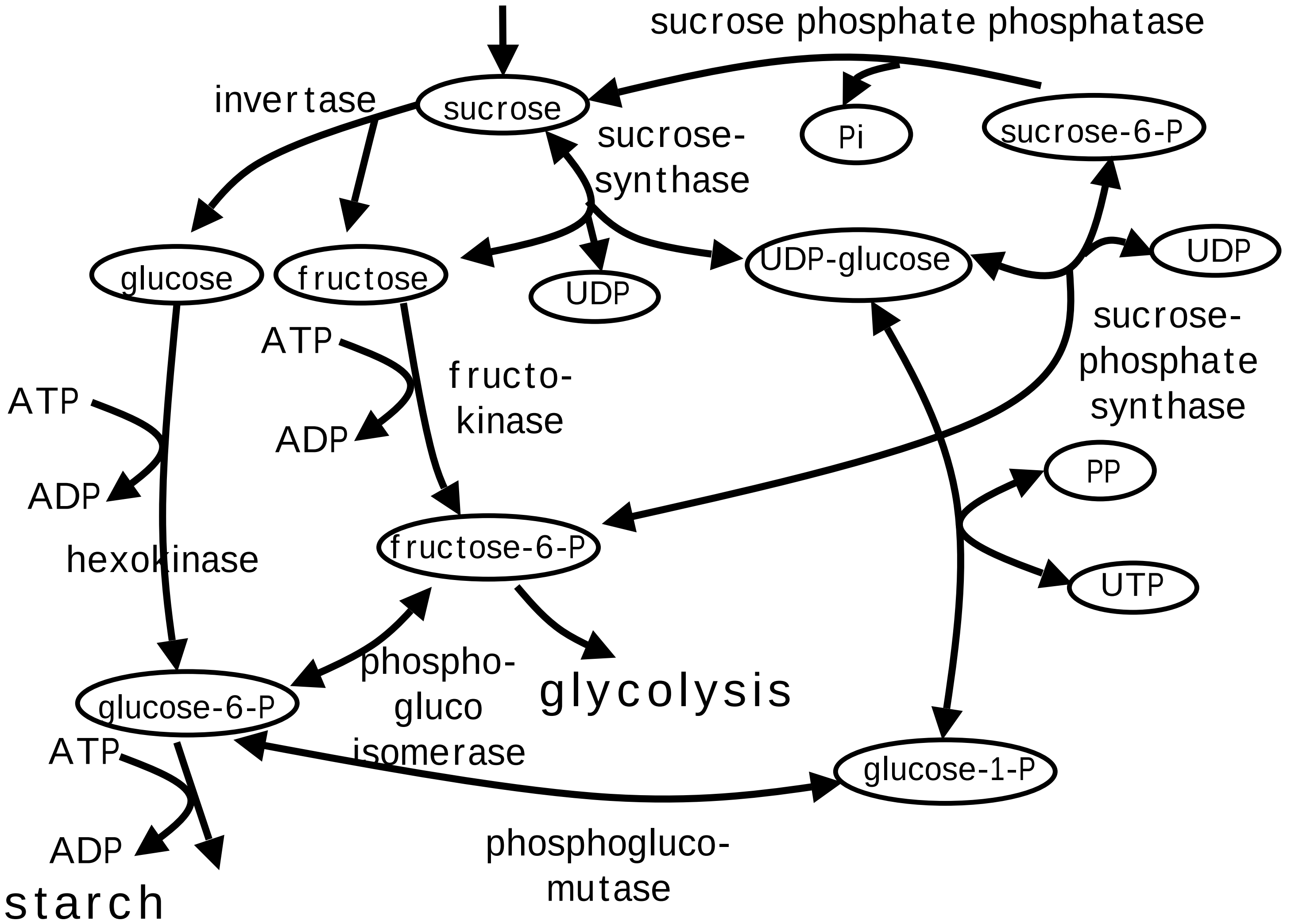
ucrose-to-starch-pathway in potato tuber



ucrose-to-starch-pathway in potato tuber

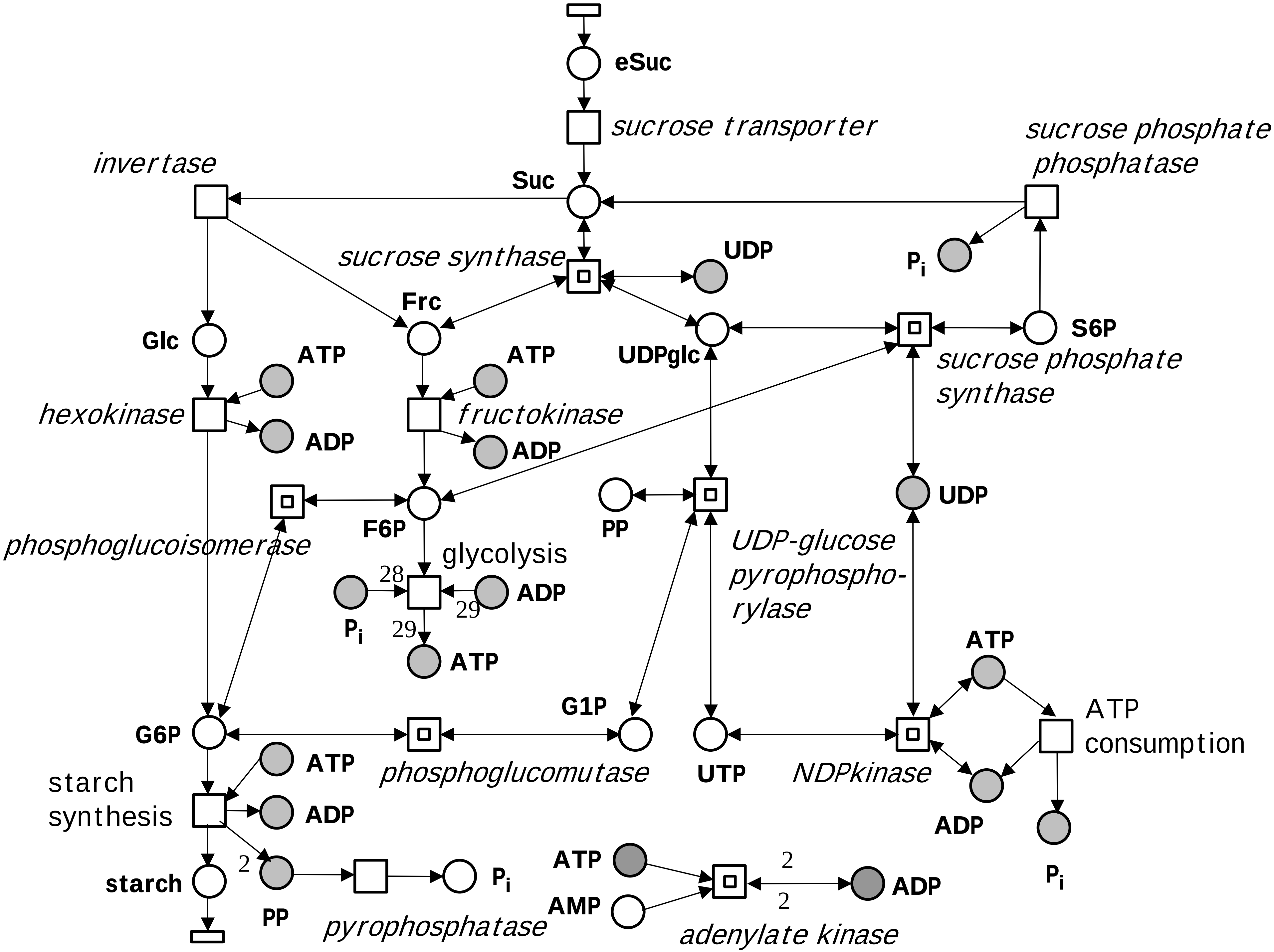


ucrose-to-starch-pathway in potato tuber



ucrose-to-starch-pathway in potato tuber

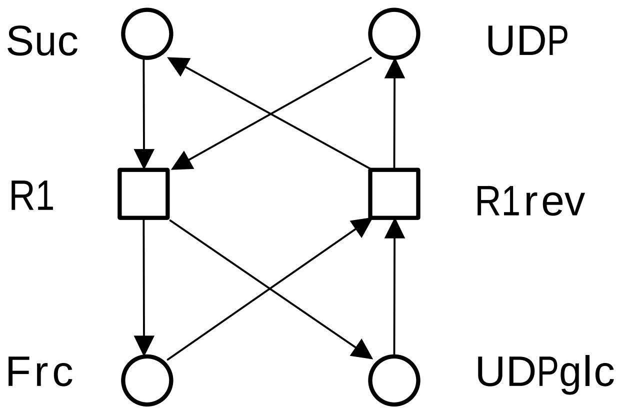
sucrose synthase:	Suc + UDP $\leftrightarrow$ UDPglc + Frc
UDP-glucose Pyrophosphorylase:	UDPglc + PP $\leftrightarrow$ G1P + UTP
phosphoglucomutase:	G6P $\leftrightarrow$ G1P
fructokinase:	Frc + ATP $\rightarrow$ F6P + ADP
phosophoglucoisomerase:	G6P $\leftrightarrow$ F6P
hexokinase:	Glc + ATP $\rightarrow$ G6P + ADP
invertase:	Suc $\rightarrow$ Glc + Frc
sucrose phosphate synthase:	F6P + UDPglc $\leftrightarrow$ S6P + UDP
sucrose phosphate phosphatase:	S6P $\rightarrow$ Suc + P_i
glycolysis (b):	F6P + 29 ADP + 28 P_i $\rightarrow$ 29 ATP
NDPkinase:	UDP + ATP $\leftrightarrow$ UTP + ADP
sucrose transporter:	eSuc $\rightarrow$ Suc
ATP consumption (b):	ATP $\rightarrow$ ADP + P_i
starch synthesis:	G6P + ATP $\rightarrow$ 2P_i + ADP + starch
adenylate kinase:	ATP + AMP $\leftrightarrow$ 2ADP
pyrophosphatase:	PP $\rightarrow$ 2 P_i



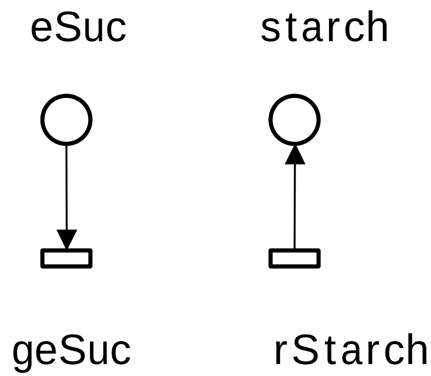
eSuc
sucrose transporter
sucrose phosphate phosphatase
invertase
Suc
sucrose synthase
UDP
P
i
Frc
S6P
Glc
ATP
ATP
UDPglc
sucrose phosphate synthase
hexokinase
fructokinase
ADP
ADP
UDP
PP
F6P
phosphoglucoisomerase
UDP-glucose pyrophospho-rylase
glycolysis
28
ADP
29
P
i
29
ATP
ATP
G1P
ATP consumption
G6P
phosphoglucomutase
UTP
NDPkinase
ATP
ADP
starch synthesis
ADP
ADP
P
i
2
starch
PP
pyrophosphatase
P
i
ATP
AMP
2
2
ADP
adenylate kinase

ucrose-to-starch-pathway in potato tuber

A hierarchical node:



Interface to the environment



Tools:

Editing: Ped Heiner BTU Cottbus

Animation: PedVisor http://www.informatik.tu-cottbus.de/~wwwdssz/

Qualitative analysis: INA Starke HU Berlin http://www.informatik.hu-berlin.de/~starke/ina.html

Qualitative analysis using INA

Elementary properties

The net is not statically conflict-free.
The net has transitions without pre-place.
The net is not covered by semipositive P-invariants.
The net is not structurally bounded.
The net is not safe.
The net has transitions without post-place.
The net is not ordinary.

The net is pure.
The net is not strongly connected.
The net is not bounded.
The net is not live and safe.
Transition 18.geSuc has no pre-place.
Transition 21.rStarch has no post-place.
The net is not conservative.

At least the following transitions are live: 0.SucTrans, 1.Inv, 18.geSuc,
At least the following places are simultaneously unbounded: 0.Suc, 1.eSuc, 2.Glc, 3.Frc,

The net is marked.
The net is not homogenous.
The net has no nonempty clean trap.
The net has no places without post-transition.
The net is connected.

The net is not marked with exactly one toke
The net has not a non-blocking multiplicity
The net has no places without pre-transition
Maximal in/out-degree: 6

ORD	HOM	NBM	PUR	CSV	SCF	CON	SC	Ft0	tF0	Fp0	pF0	MG	SM	FC	EFC	ES
N	N	N	Y	N	N	Y	N	Y	Y	N	N	N	N	N	N	N

Qualitative analysis using INA

Structural properties

DTP	CPI	CTI	B	SB	REV	DSt	BSt	DTr	DCF	L	LV	L&S
?	N	Y	N	N	?	?	?	?	?	?	?	N

- liveness could not be decided because the net is unbounded and the reachability graph cannot be calculated
- the coverability graph has more than 4 million states

smaller bounded version: more than 10^{10} states of the reachability graph

P-Invariant analysis

The net is not covered by P-invariants.

Following P-invariants were calculated:

1. UDPglc, UTP, UDP
2. ATP, AMP, ADP
3. G6P, F6P, G1P, UTP, ATP(2), ADP, S6P, P_i, PP(2)

T-Invariant analysis

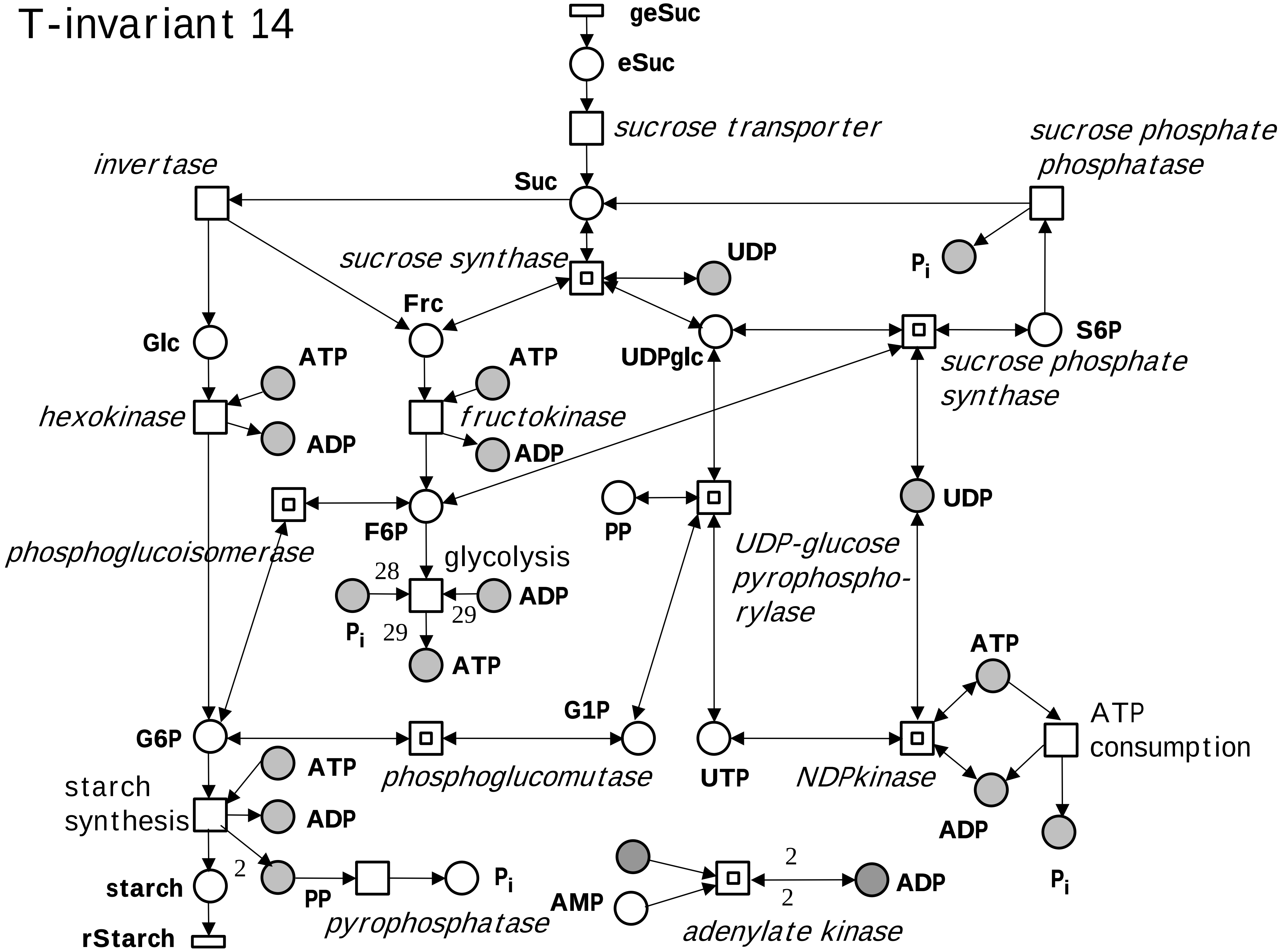
The net is covered by 19 T-invariants

7 trivials:
1. SPS, SPS_rev,
2. UGPase, UGPASE_rev,
3. SuSy_SuSy_rev,
4. PGM, PGM_rev,
5. NDPkin, NDPkin_rev,
6. AdK, AdK_rev,
7. PGI, PGI_rev

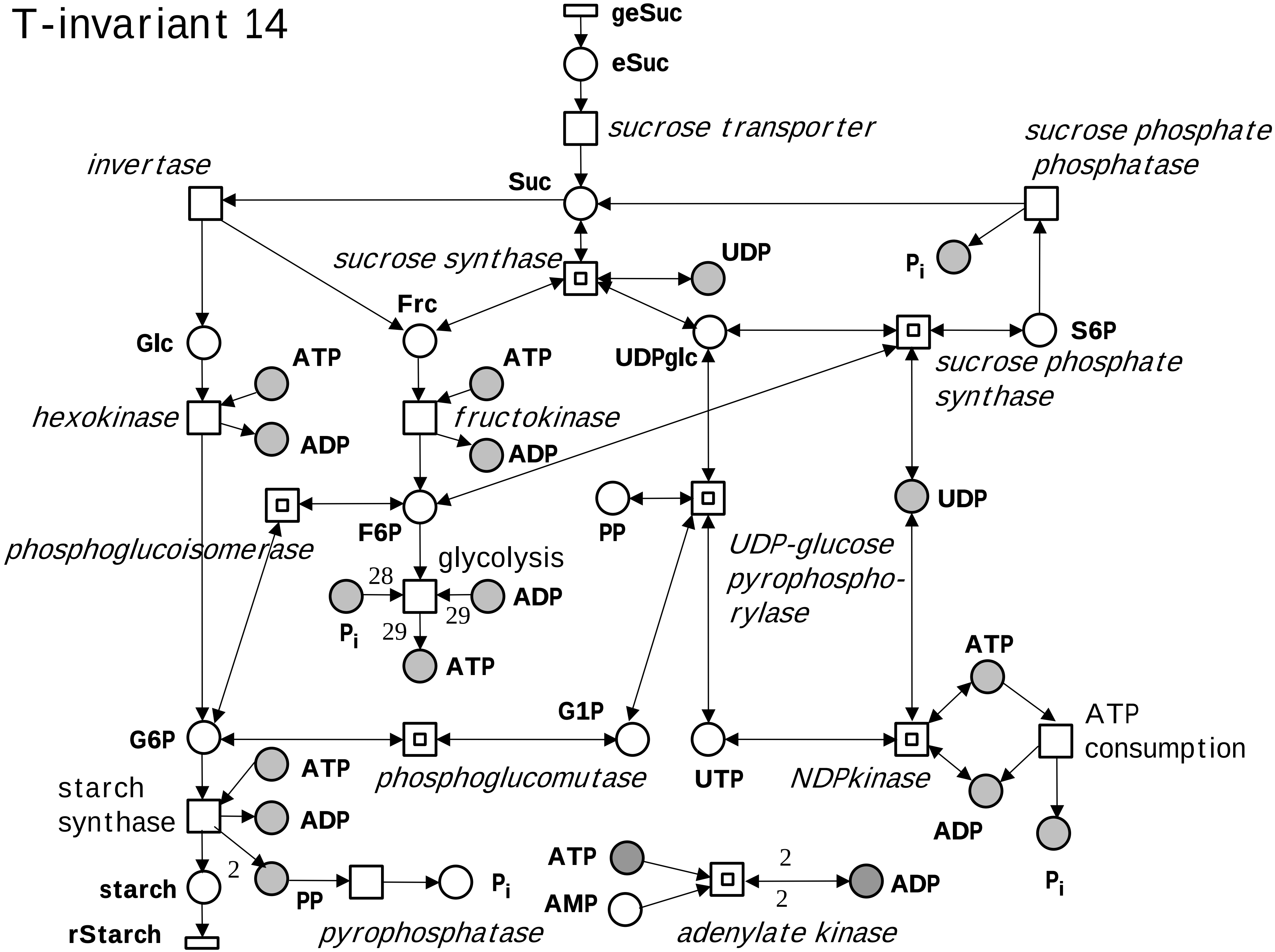
Example:

```
8 |   0.sucrose       :  1
  |   1.invertase     :  1
  |   4.R5            :  1
  |   9.hexokinase    :  1
  |  10.fructokinase  :  1
  |  18.geSuc         :  1
  |  19.glycolysis    :  2
  |  20.ATP           : 56
```

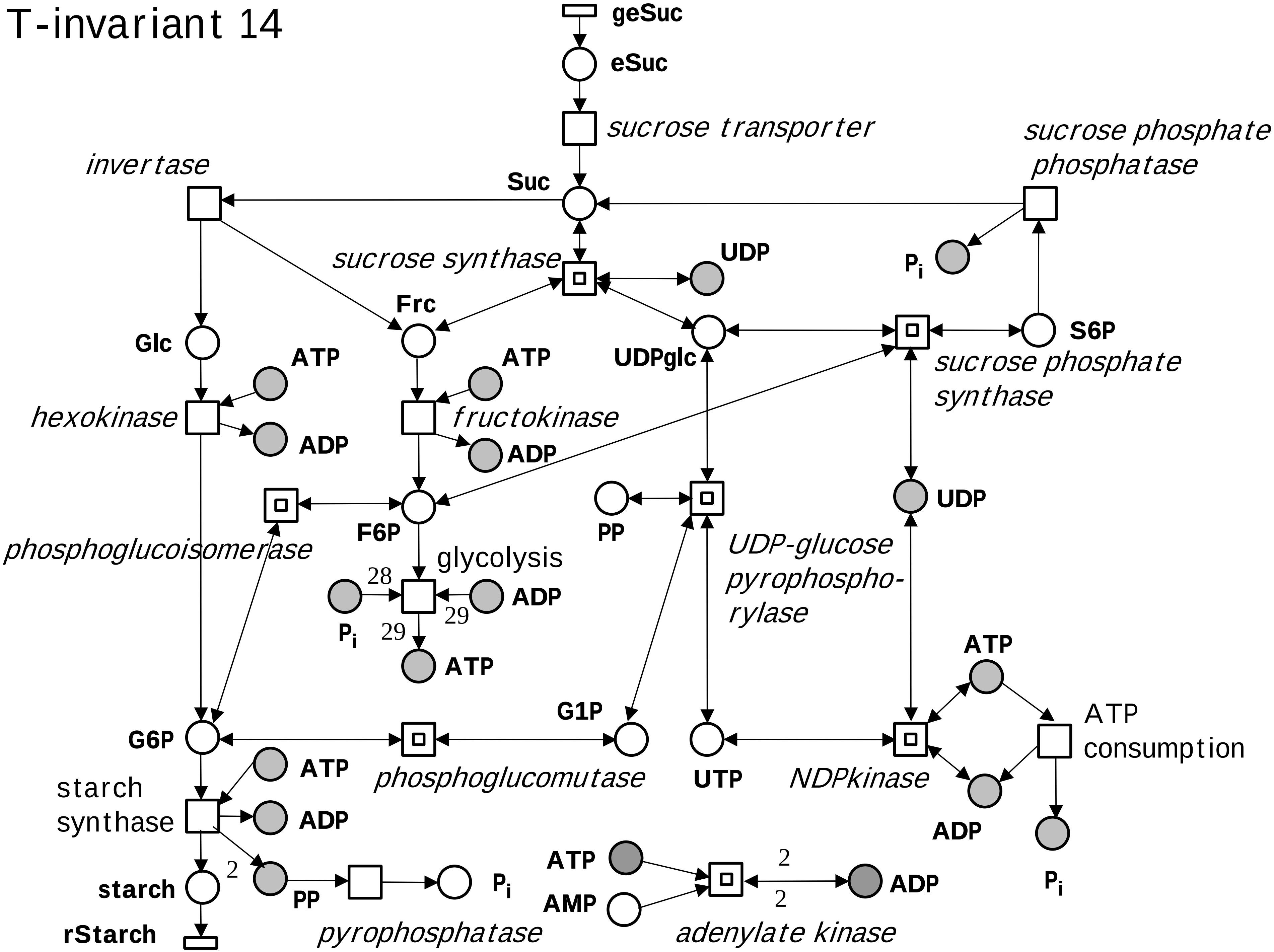
T-invariant 14



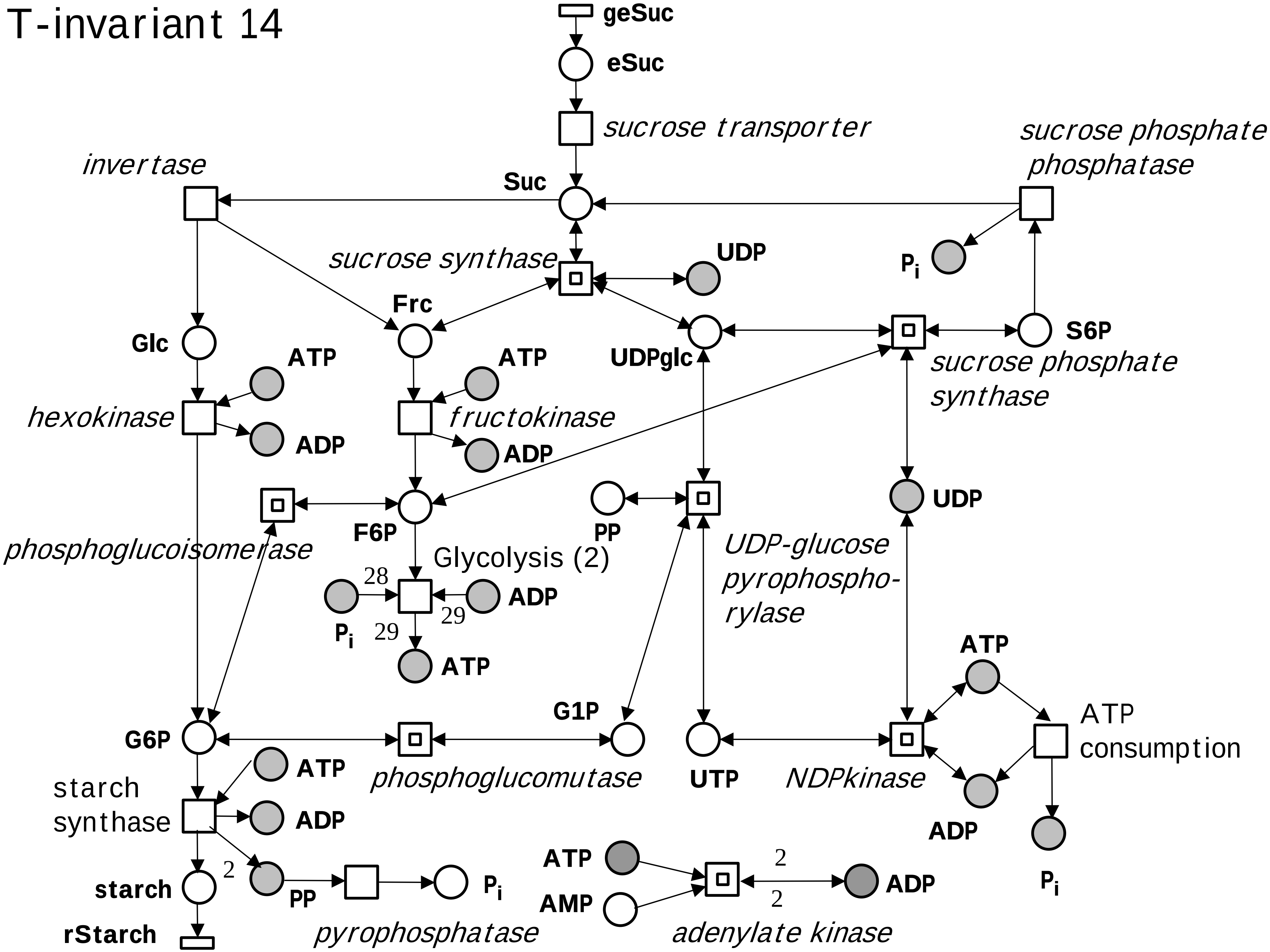
T-invariant 14



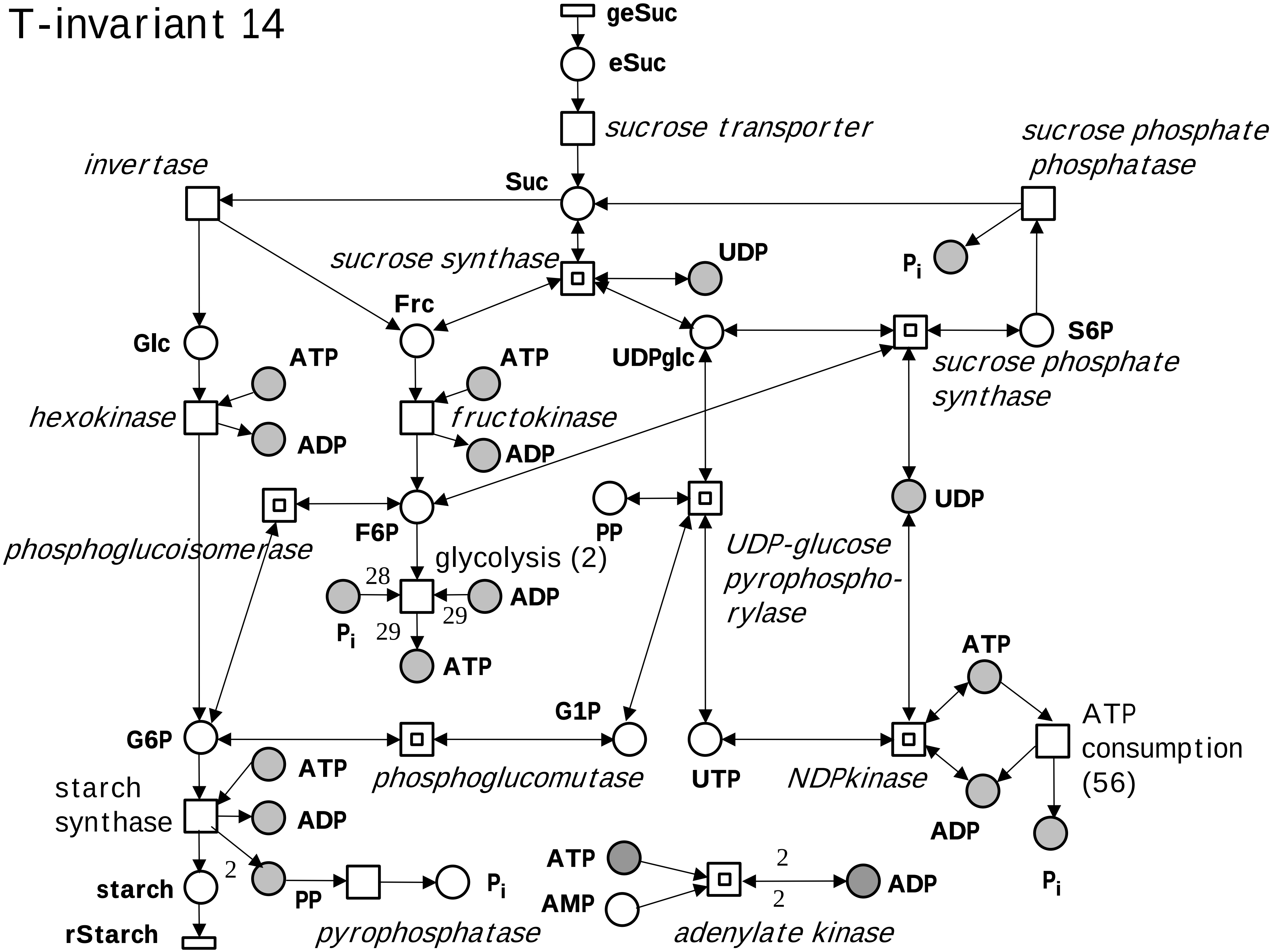
T-invariant 14



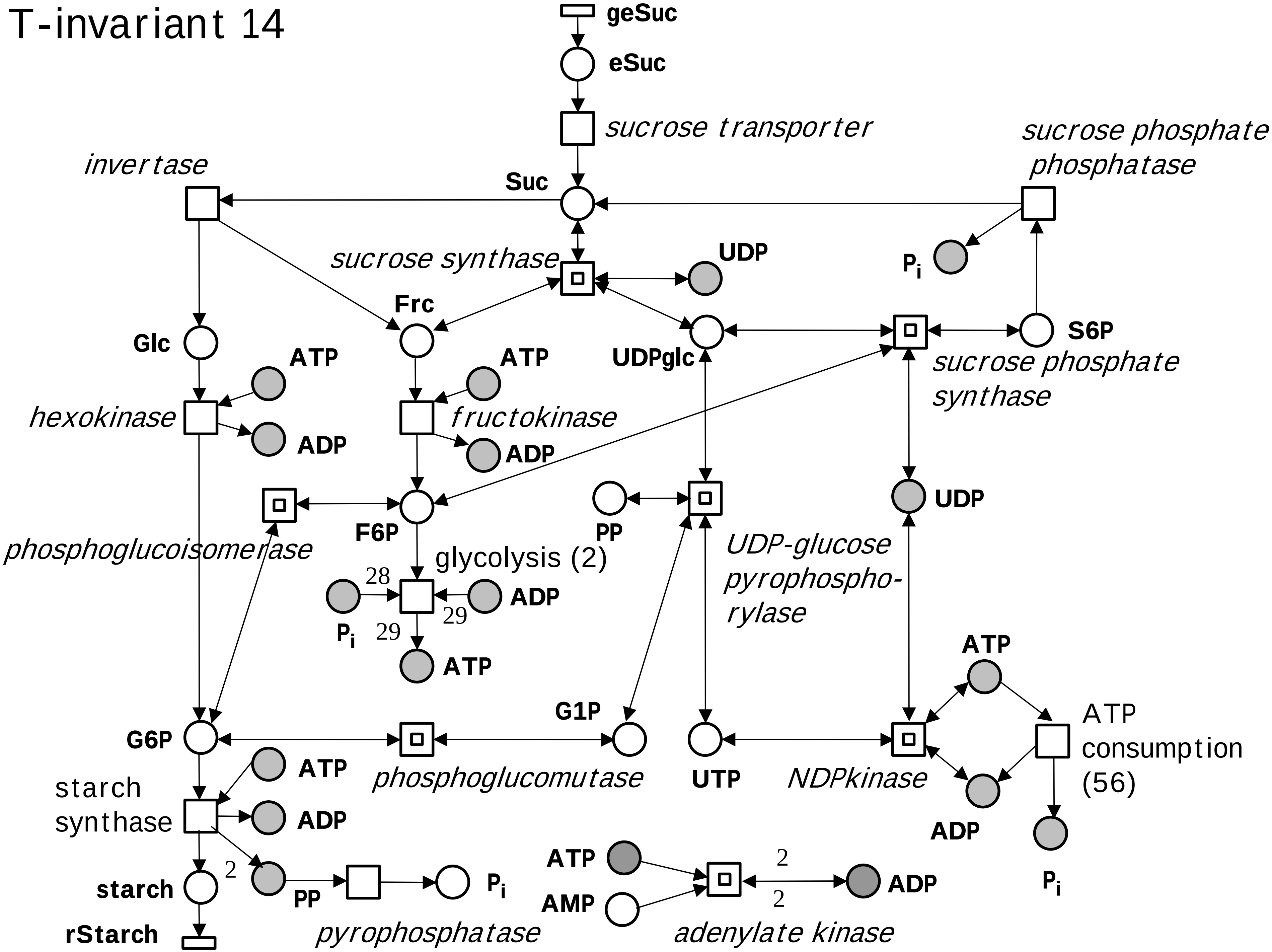
T-invariant 14



T-invariant 14



T-invariant 14



T-Invariant analysis

Invariant number	sucrose cleavage		hexoses go into		ATP	ATP used for cycling		
	SuSy	Inv	Glyc	StaSy	ATP cons	Inv SuSy_rev	Inv SPS, SPP	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			

Robustness

Robustness: sensitivity of the system against parameter
(fragility) changes (altered enzyme activity, mutations)
(Voit, 2000)

Stelling et al., Nature (2002): linear correlation between robustness and the number of elementary modes (T-invariants)

Our suggestion: - enzyme distribution over T-invariants
- number of alternative paths

Potato net: - fructokinase occurs in all T-invariants
- there is no enzyme that occurs in only one T-invariant

Conclusions & Outlook

Petri nets provide

(1) a unique description of biological networks

(2) methods for qualitative analysis to check models by the calculation of system properties.

(3) The complexity of biological systems make it necessary to extend Petri net methods.

(4) Automatic interpretation of T-invariants is necessary.

Projects

- Glycolysis-pentose phosphate pathway in erythrocytes

  Voss, Heiner, Koch, *BioSystems* in press (2004)

- Apoptosis Heiner, Koch, Will, *Proc.Comp.Methods Syst.Biol.* (2003)

  Heiner, Voss, Koch, *In Silico Biology* (2003)

Heiner & Koch, 25th International Conference on Application and Theory of Petri Nets, 21th - 25th June, Bologna, Italy (2004)

Ongoing projects:

1. The whole E.coli pathway Nina Kramer
2. The whole potato tuber pathway Nina Kramer

3. Detailed glycolysis with coloured Petri nets in human
   Thomas Runge, BTU Cottbus
4. G1/S - phase in mammalian cells Thomas Kaunath
   (tumour cell lines, Duchenne muscle dystrophy)

Thanks!

Monika Heiner (Brandenburg University of Technology at Cottbus)

Björn Junker (Max Planck Institute for Molecular Plant Physiology Golm)

Grazie per la vostra attenzione!