Qualitative biochemical pathway analysis using Petri nets

Ina Koch
Technical University of Applied Sciences Berlin
http://www.tfh-berlin.de/bi/

Monika Heiner
Brandenburg University of Technology Cottbus
http://www.informatik.tu-cottbus.de/~wwwdssz/

Federal Ministry of Education and Research

Marseille, February 25th 2004

Berlin Center for Genome Based Bioinformatics

Outline

- Introduction
- Petri net Basics
- Sucrose-to-Starch Pathway in potato tuber
- Model validation
- Summary & Outlook
- Simulation of the net

Introduction

Towards system-level understanding of biological systems

The roots:

- N. Wiener "Cybernetics or Control and Communication in the Animal and the Mach
  The MIT Press, *Cambridge* (1948) → Cybernetics, Biological cybernetics

- W.B. Cannon "The wisdom of the body", Norton, New York
  → Concept of Homeostasis

- L. van Bertalanffy "General System Theory" *Braziler*, *New York* (1968)
  → First general theory of the system

Description and analysis of biological systems

at the physiological level → at the molecular level

Introduction

I. System structure identification

regulatory relationships of genes, interactions of proteins, physical structure of organisms, (high-throughput DNA microarray, RT-PCR)

II. System behaviour analysis

sensitivity against external perturbations cell response to certain chemicals, estimation of side effects

III. System Control

How we can transform cancer cells to turn them into normal cells or cause apoptosis? Can we control the differentiation status of a specific cell into a stem cell and control it to differentiate into the desired cell type?

IV. System Design

with the aim of providing cures for diseases, design and growth organs form the patient’s own tissue , metabolic engineering for product optimisation

Introduction

I. System structure identification

regulatory relationships of genes, interactions of proteins, physical structure of organisms, (high-throughput DNA microarray, RT-PCR)

II. System behaviour analysis

sensitivity against external perturbations cell response to certain chemicals, estimation of side effects

III. System Control

How we can transform cancer cells to turn them into normal cells or cause apoptosis? Can we control the differentiation status of a specific cell into a stem cell and control it to differentiate into the desired cell type?

IV. System Design

with the aim of providing cures for diseases, design and growth organs form the patient’s own tissue, metabolic engineering for product optimisation

Introduction

Bioinformatics should provide

- databases for storing experimental data at different description levels
- editor software for editing biological networks - unique representation of networks
- data visualisation software to represent also large networks
- simulation software - metabolic pathways, signal transduction pathways, cell? , organism?
- system analysis techniques - qualitative analysis, quantitative analysis, stochastic analysis, model validation methods
- hypothesis generator and experiment planning advisor tools

Provided by Petri net theory

Introduction

Metabolic Control Analysis - MCA

Metabolic system: connected unit, steady state

- Homogenous distribution of metabolites over the enzymes
- rates of enzyme effect are proportional to the enzyme concentrations

MCA bases on solution of systems of differential equations

- MCA H. Kacser, J.A. Burns *Symp.Soc.Exp.Bio*. **27**: 65 (1973)
  R. Heinrich, T.A. Rapoport *Eur.J.Biochem*. **42**: 89, 97 (1974)
- Biochemical systems theory
  A.M. Savageau *J.Theor.Biol*. **25**: 365, 370 (1969)
- Flux oriented theory
  B. Crabtree, E.A. Newsholme *Biochem.J*. **247**: 113 (1987)

GEPASI P. Mendes *Comp.Appl.Biosci*. **9**:563 (1993)

Introduction

Graph-Theory

- Hybrid graphs — M.C. Kohn, W.J. Letzkus *J.Theor.Biol.* **100**: 293 (1983)
- Bond graphs — J. Lefèvre, J. Barreto *J.Franklin Inst.* **319**: 201 (1985)
- Net-thermodynamics — D. Mikulecky *Am.J.Physiol.* **245**: R1 (1993)
- Weighted linear graphs

  B.N. Goldstein, E.L. Shevelev *J.Theor.Biol.* **112**: 493 (1985)

  B.N. Goldstein, V.A. Selivanov *Biomed.Biochim.Acta* **49**: 645 (1990)
- Meta-nets (with gene expression systems)

  M.C. Kohn, D.R. Lemieux *J.Theor.Biol.* **150**: 3 (1991)
- Bipartite graphs — A.V. Zeigarnik, O.N. Temkin *Kin.Catalysis* **35**: 674 (1994)
- KING (KINetic Graphs) — A.V. Zeigarnik *Kin.Catalysis* **35**: 656 (1994)

Model Validation

- 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 anymore.

- How model validation could be performed?
  By qualitative analysis

  Basic structure properties: invariants, robustness, alternative pathways, knockout simulation

  Basic dynamic properties: dead states, deadlocks, traps, liveliness

↓

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

Petri 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

Metabolic/Biological Petri-Nets - MPN/BPN

Reddy et al. (1993, 1996), Matsuno et al. (2003,2003)

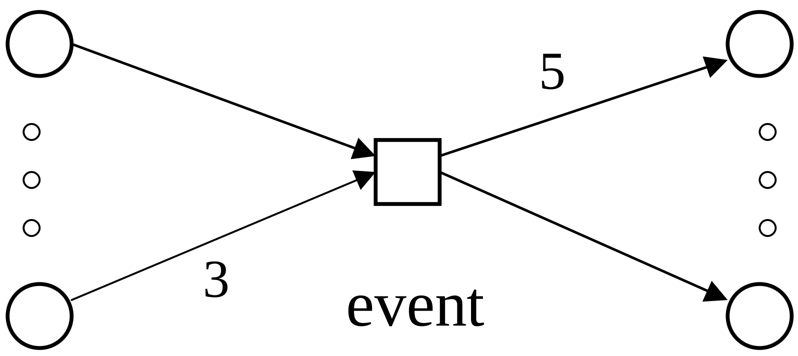
Petri net basics

Petri nets: Two-coloured, labelled, directed, bipartite graphs

Vertices: (nodes)	places	transitions
	○	□
	passive elements	active elements
	conditions	events
	states	actions
	chemical compounds	chemical reactions
	metabolites	conversions of metabolites catalysed by enzymes

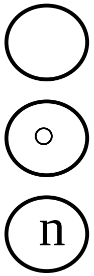
Petri net basics

Edges: (arcs) pre-conditions post-conditions



Petri net basics

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



condition is not fulfilled

condition is (one time) fulfilled

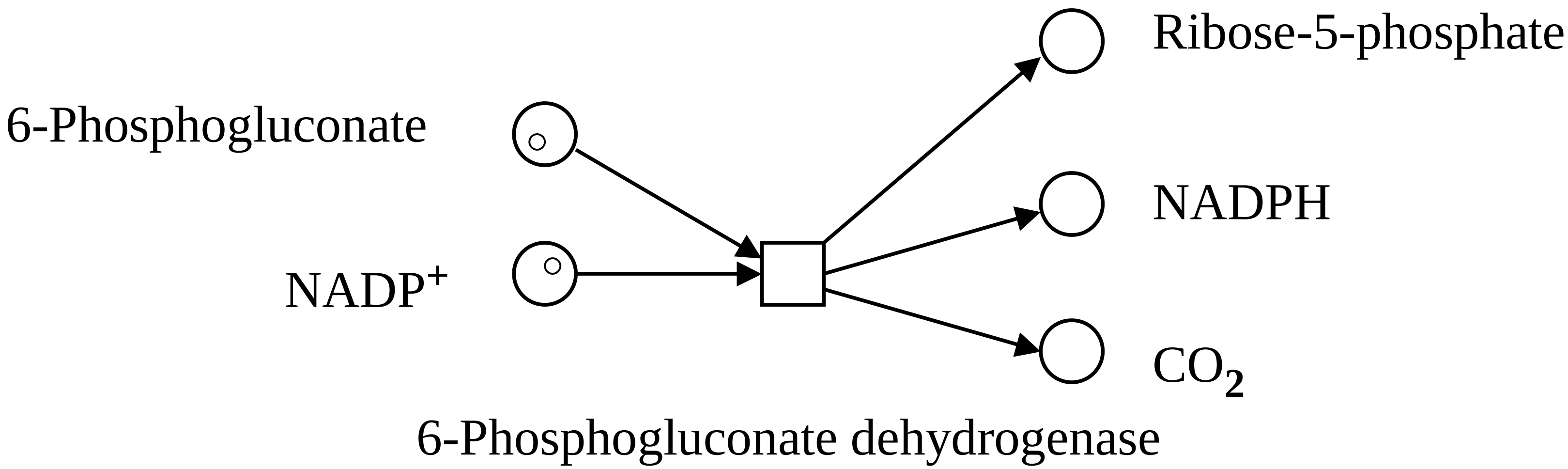
condition is n times fulfilled

Marking: system state, token distribution
initial distribution

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

Petri net basics

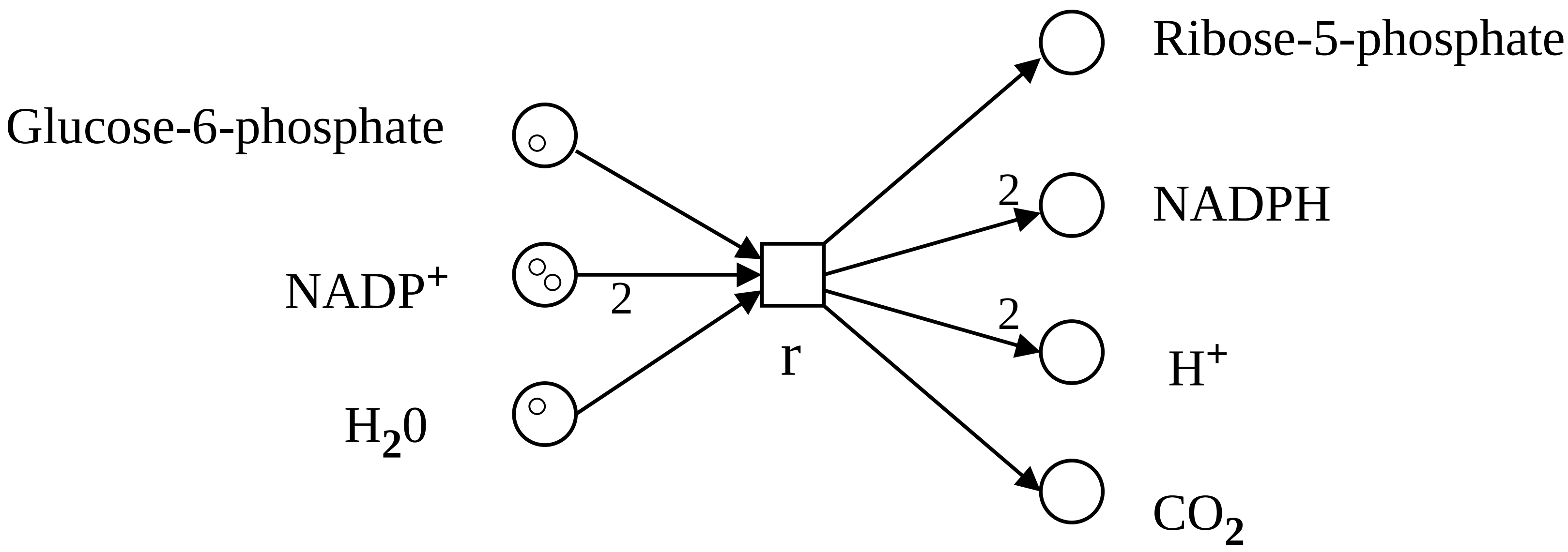
Example: Pentose Phosphate Pathway - one reaction



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

Petri net basics

Example: Pentose Phosphate Pathway - sum reaction



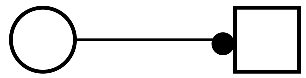
$$G6P + 2NADP^{+} + H_2O \rightarrow R5P + 2NADPH + 2H^{+} + CO_2$$

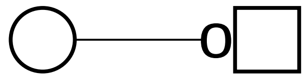
Petri net basics

Special places:

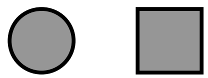
input:	substrates	(source, e.g. sucrose)
output:	products	(sink, e.g. starch)

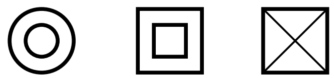
Special edges:

reading edges 

inhibitor edges 

Additional places & transitions:

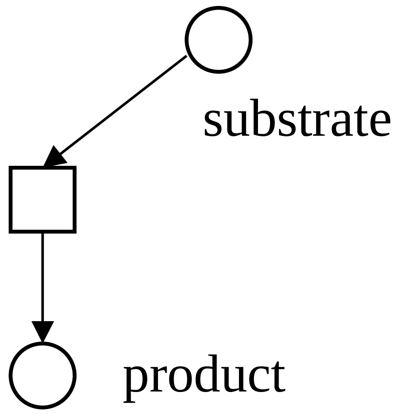
logical 

hierarchical 

Petri net basics

Transitions in MPNs:

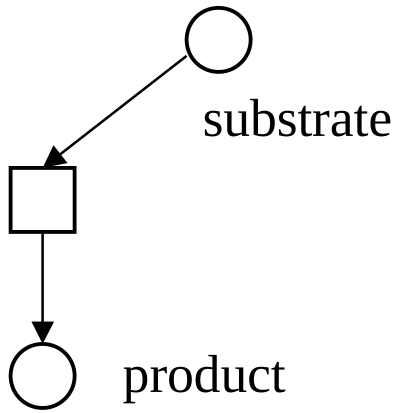
Reaction:



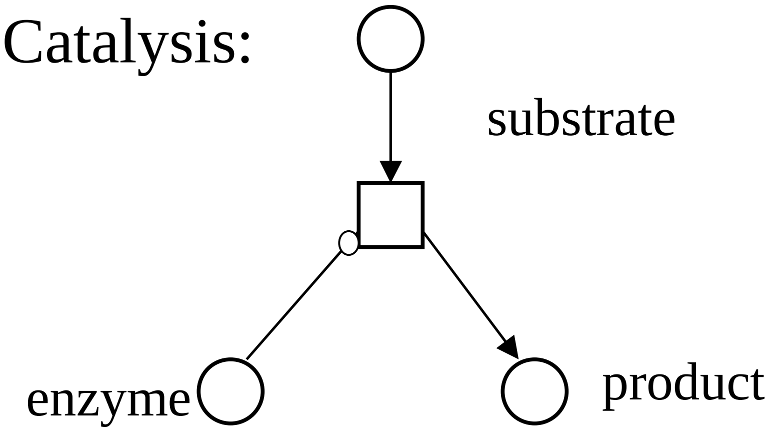
Petri net basics

Transitions in MPNs:

Reaction:



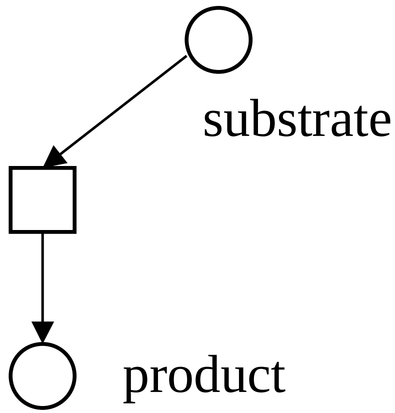
Catalysis:



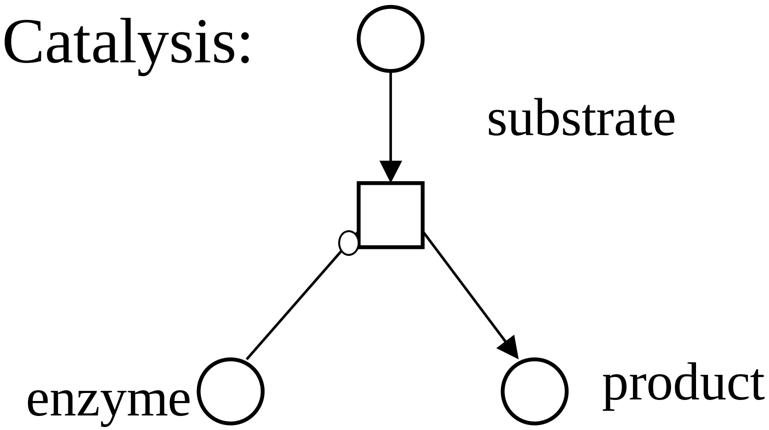
Petri net basics

Transitions in MPNs:

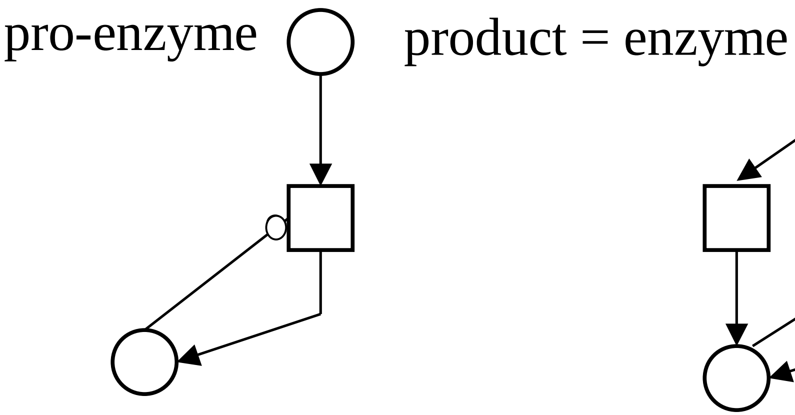
Reaction:

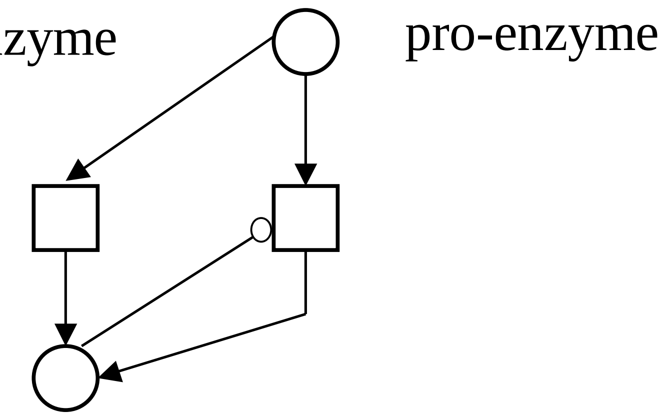


Catalysis:



Auto-catalysis:





Petri net basics

Questions of the qualitative analysis

Dynamical (behavioural) properties

- How often can a transition fire?(0-times, n-times, ∞ times) liveliness
- What is the maximal token number for a place? (0, 1, k, ∞) boundedness (k-bounded)
- Is a certain system state again and again reachable? progressiveness
- Is a certain system state never reachable? safety
- How many and which system states could be reached ? (0, 1, k, ∞) reachability analysis

Petri net basics

Questions of the qualitative analysis

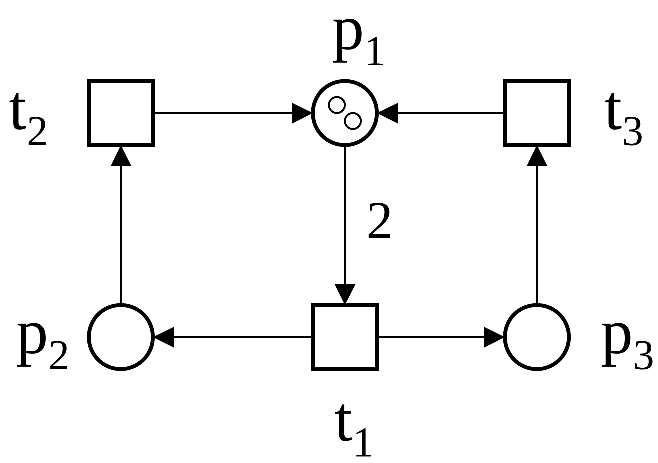
Static (structural) properties

- 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 or P- invariants/Transition-invariants or T-invariants

Petri net basics



incidence matrix

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

transition invariant: $C\,y = 0$

set of transitions, whose firing reproduces a given marking

$$-2y_1 + y_2 + y_3 = 0$$
$$y_1 - y_2 = 0$$
$$y_1 - y_3 = 0$$

Petri net basics

Minimal semi-positive T-invariants

- each net behaviour can be described by linear combination of these invariants,

  K. Lautenbach in *Advances in Petri Nets 1986 Part I, LNCS* ***254,*** Springer (1987)

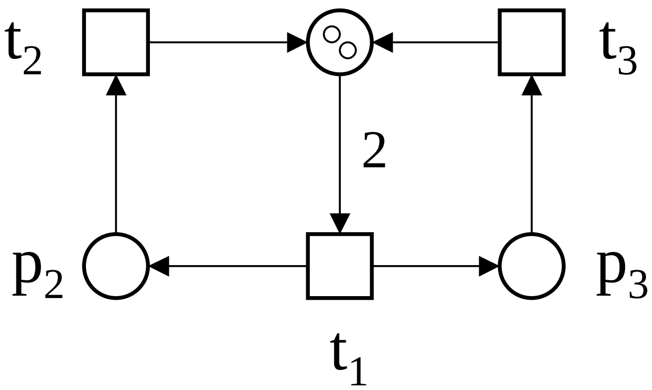
- covered by T-invariants: necessary condition for liveliness

Biological interpretation

- minimal set of enzymes which could operate at steady state
- set of reactions that can be in a state of continuous operation
- indicate the presence of cyclic firing sequences

  S. Schuster, C. Hilgetag, R. Schuster *Proc.Sec.Gauss Symp*. (1993) Elementary modes

Petri net basics



incidence matrix

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

place invariant: $x\,C = 0$

set of places, whose weighted sum of tokens is always constant

$$-2x_1 + x_2 + x_3 = 0$$

$$x_1 - x_2 = 0$$

$$x_1 - x_3 = 0$$

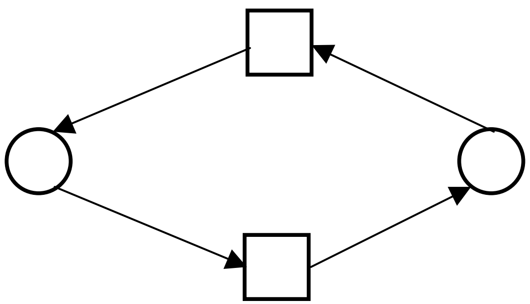
Petri net basics

Minimal semi-positive place invariants

- all possible P-invariants can be computed from the minimal set of semi-positive P-invariants by linear combination
- covered by P-invariants: sufficient condition for boundedness

Biological interpretation

- set of metabolites, whose total net concentration remains unchanged in the course of a reaction



ADP, ATP
$NADP^+$, NADPH

Sucrose-to-starch-pathway in potato tuber

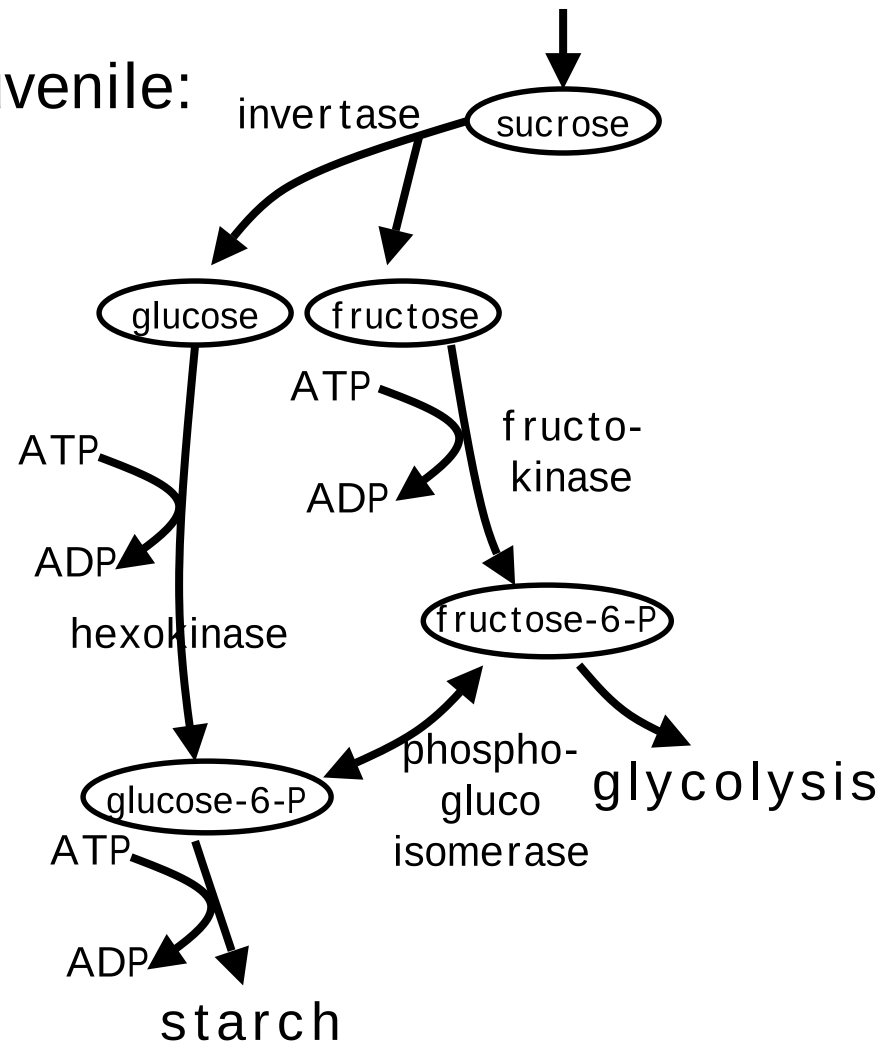
- rich in carbohydrates and energy
- a natural source of folate
- full of vitamin C
- low in calories
- good source of niacin, vitamin B6, iodine, thiamine, and minerals
- no cholesterol
- completely fat free

Research interest: increasing the starch content
full understanding of the pathway

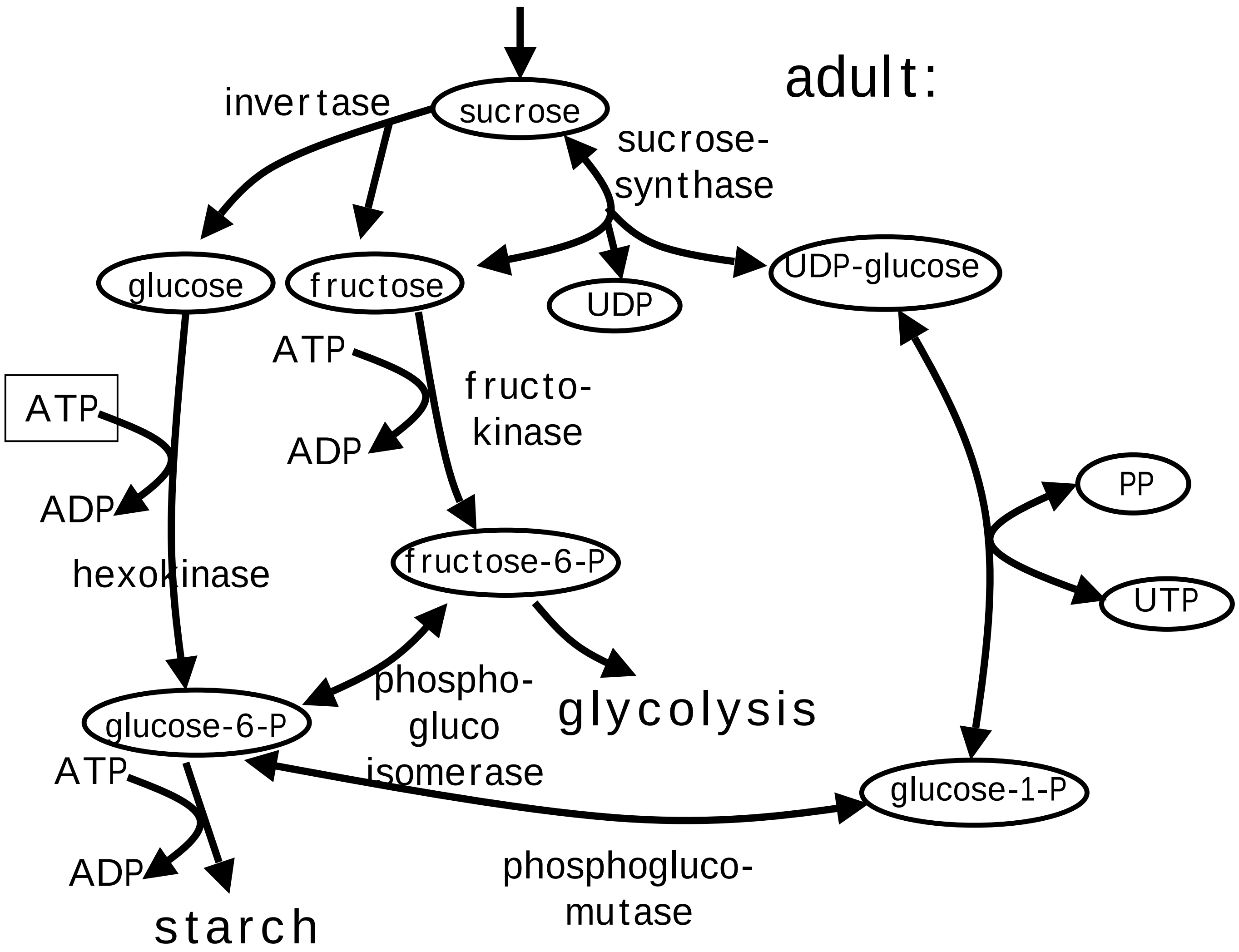
Co-operations: Max Planck Institute for Molecular Plant Physiology, Golm
Brandenburg University of Technology Cottbus

Sucrose-to-starch-pathway in potato tuber

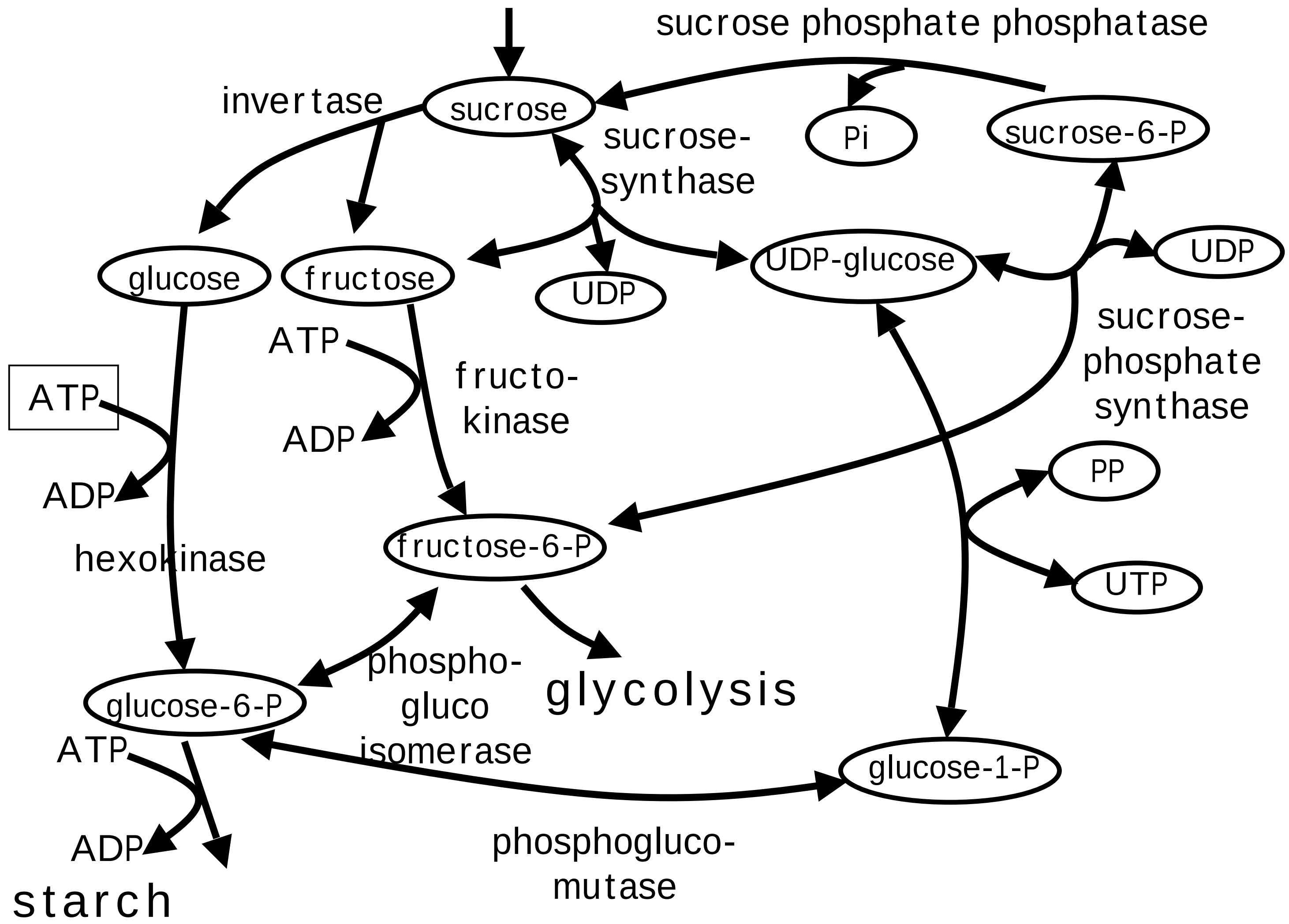
juvenile:



Sucrose-to-starch-pathway in potato tuber

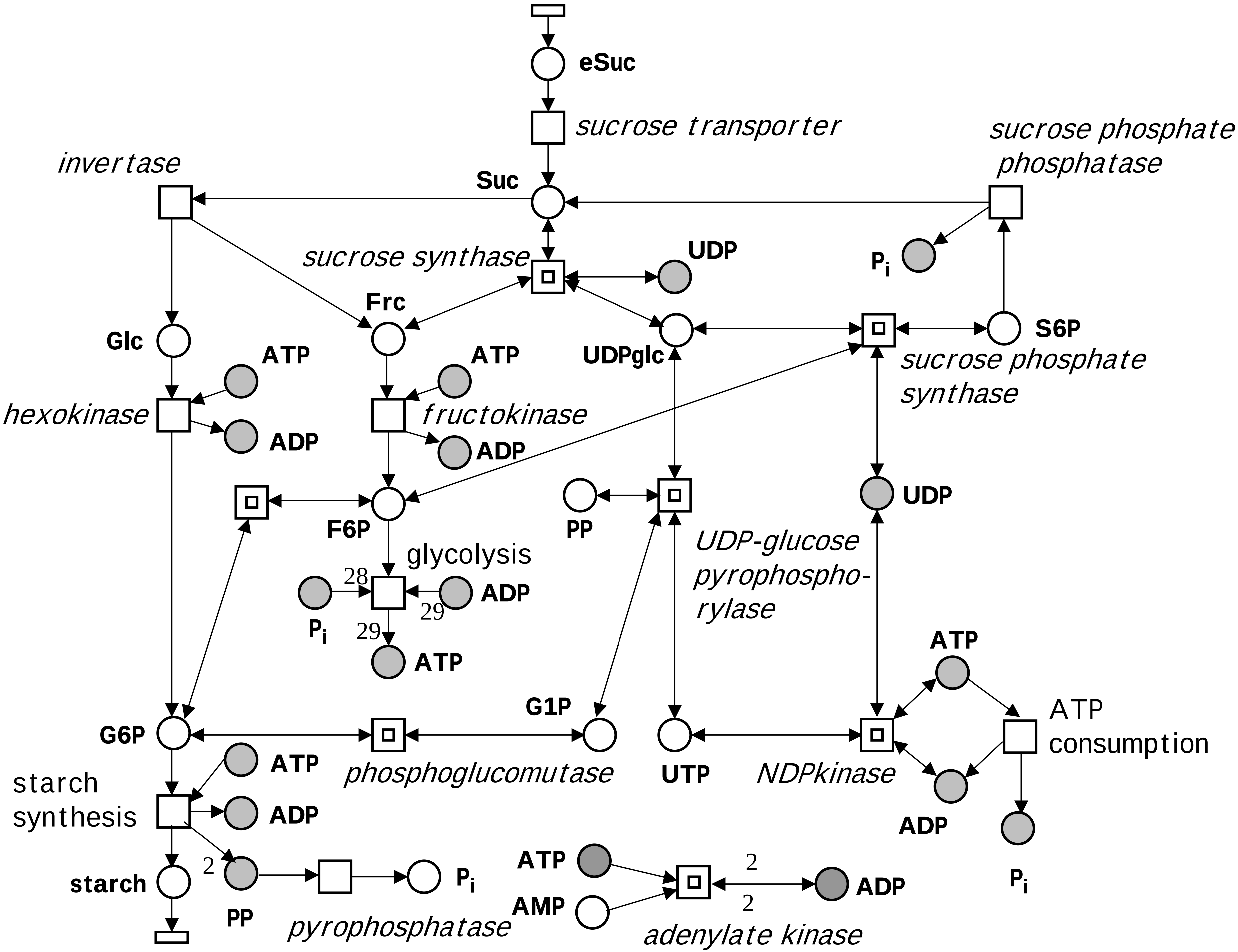


Sucrose-to-starch-pathway in potato tuber



Sucrose-to-starch-pathway in potato tuber

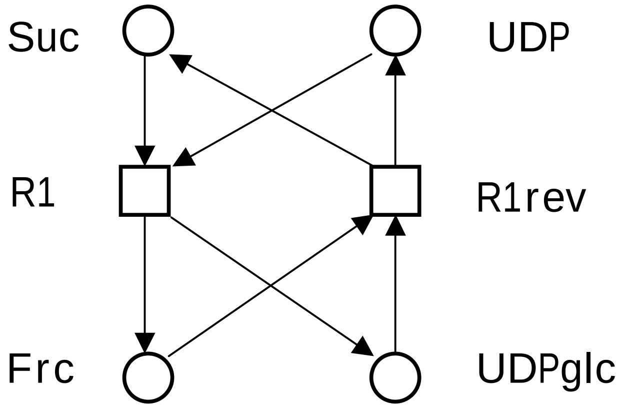
sucrose transporter:	eSuc $\rightarrow$ Suc
sucrose synthase:	Suc + UDP $\leftrightarrow$ UDPglc + Frc
invertase:	Suc $\rightarrow$ Glc + Frc
hexokinase:	Glc + ATP $\rightarrow$ G6P + ADP
fructokinase:	Frc + ATP $\rightarrow$ F6P + ADP
sucrose phosphate synthase:	F6P + UDPglc $\leftrightarrow$ S6P +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



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

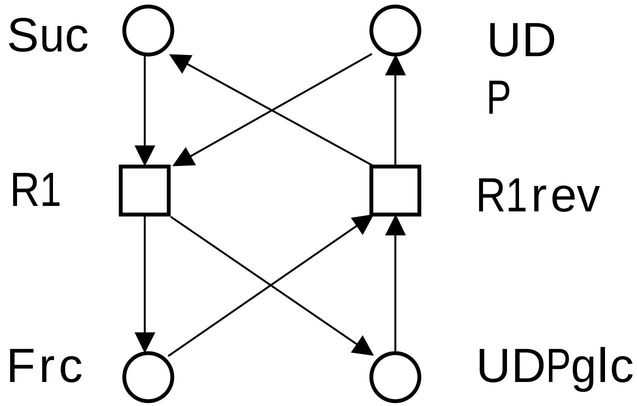
Sucrose-to-starch-pathway in potato tuber

A hierarchical node:

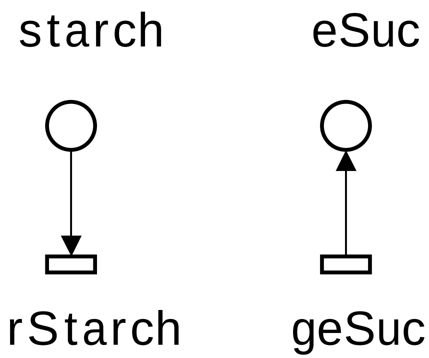


Sucrose-to-starch-pathway in potato tuber

A hierarchical node:



Interface to the environment



Qualitative analysis using INA

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

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

ORD: ordinary - the multiplicity of every edge is equal one

HOM: homogenous - for any place all outgoing edges have the same multiplicity

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

Qualitative analysis using INA

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

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

Ft0/tF0 - transitions without pre-places/post places
Fp0/pF0 - places without pre-transitions/post-transitions

CPI: covered by P-invariants - there is a P-invariant, which assigns a positive value to each place

CTI: covered by T-invariants - there is a T-invariant, which assigns a positive value to each transition

Qualitative analysis using INA

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

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

B: bounded - the number of tokens is bounded to a number k in any reachable marking

DSt: dead state - when no transition can fire any more

DTr: Deadlock - if the place is once empty no token will ever enter it

Trap - no tokens could leave this place

L: live - no state is reachable, in which a transition is dead

L&S: live & save - there is not more than one token on a place in any reachable marking

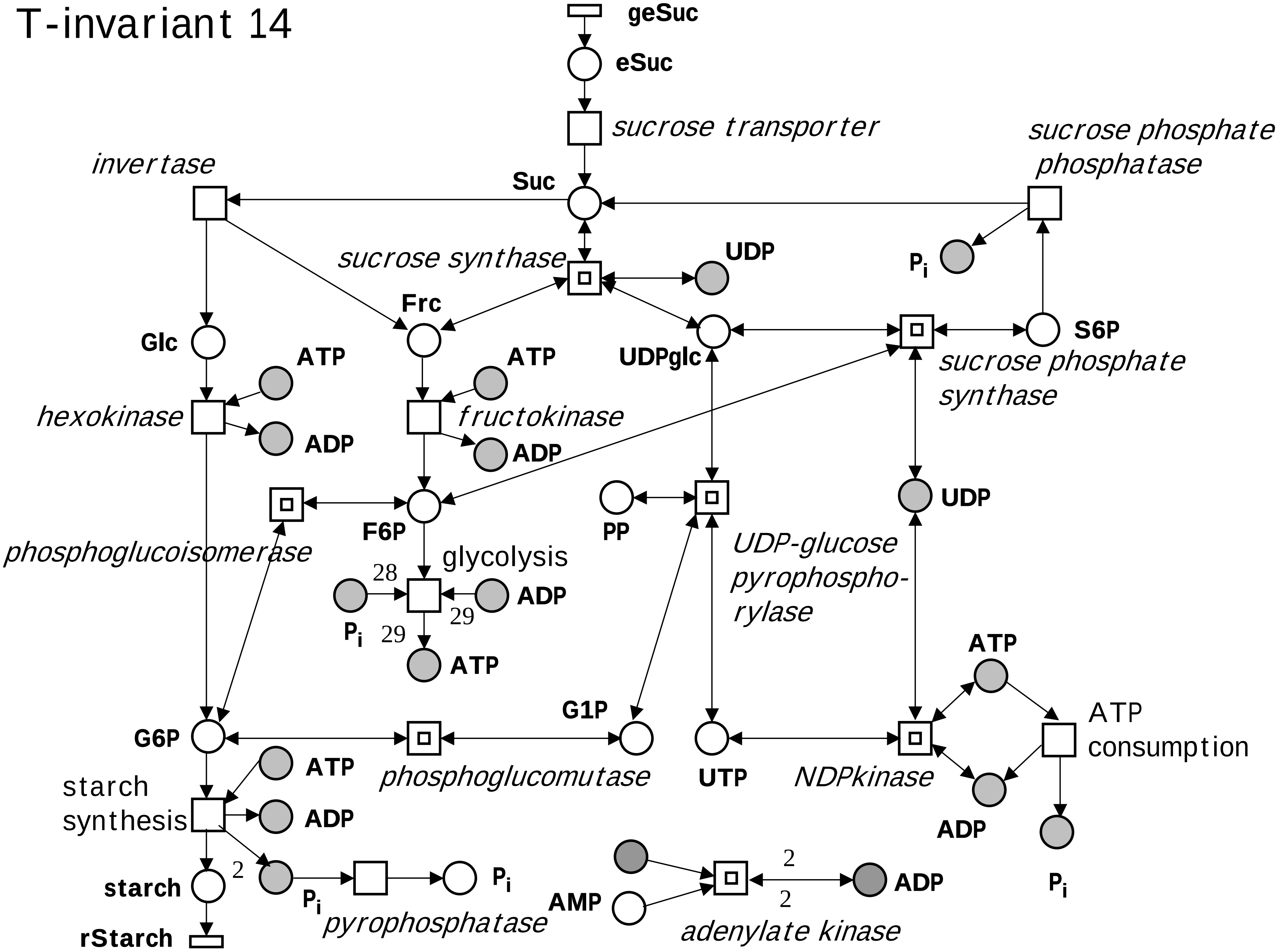
Qualitative analysis using INA

- the net is covered by 19 T-invariants
- 7 trivial T-invariants

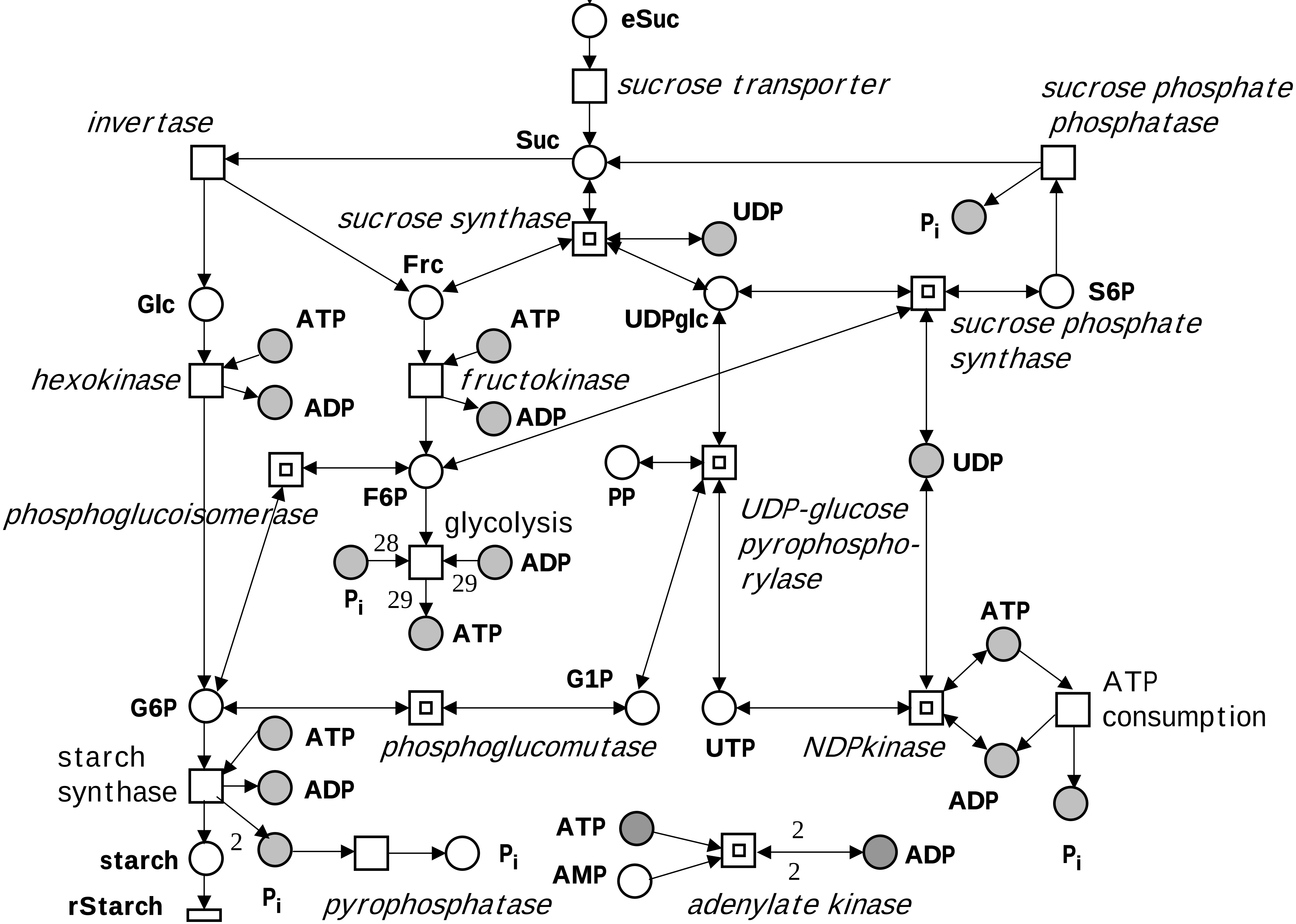
Example:

14	geSuc	:	1
	sucrose transporter	:	1
	invertase	:	1
	hexokinase	:	1
	fructokinase	:	1
	phosphoglucoisomerase_reverse	:	1
	glycolysis	:	1
	starch synthesis	:	1
	ATPconsumption	:	26
	pyrophosphatase	:	1

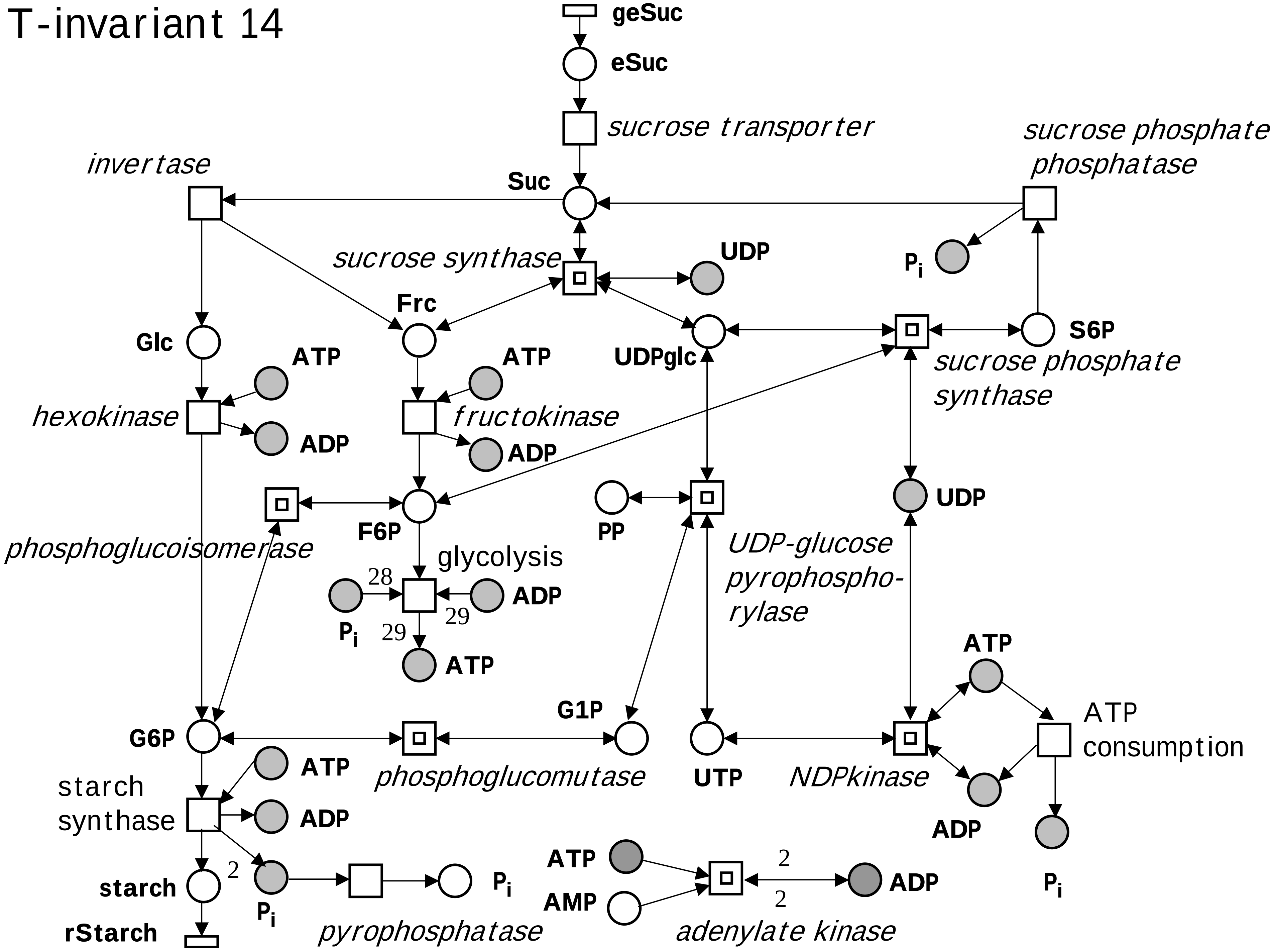
T-invariant 14



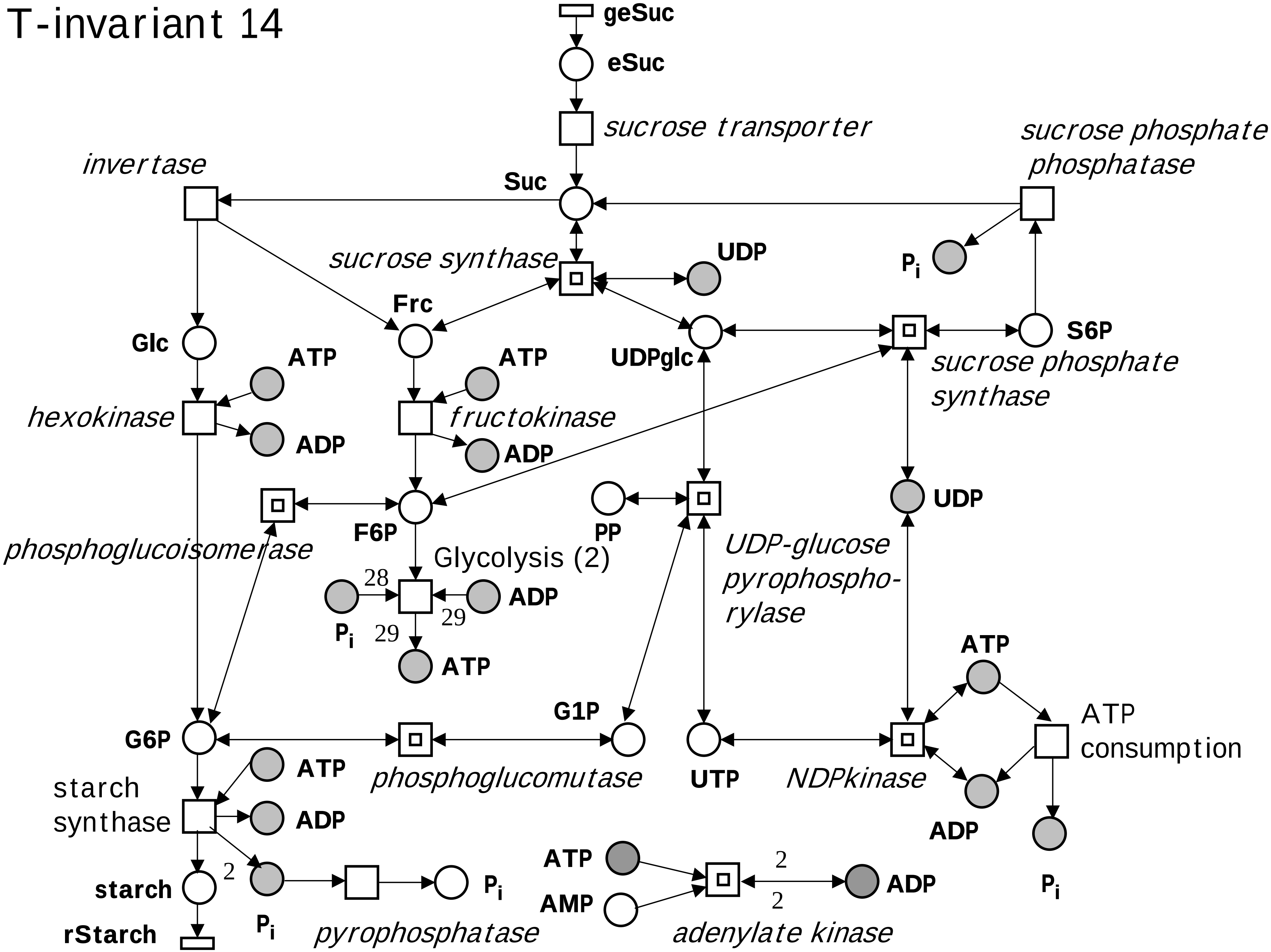
T-invariant 14



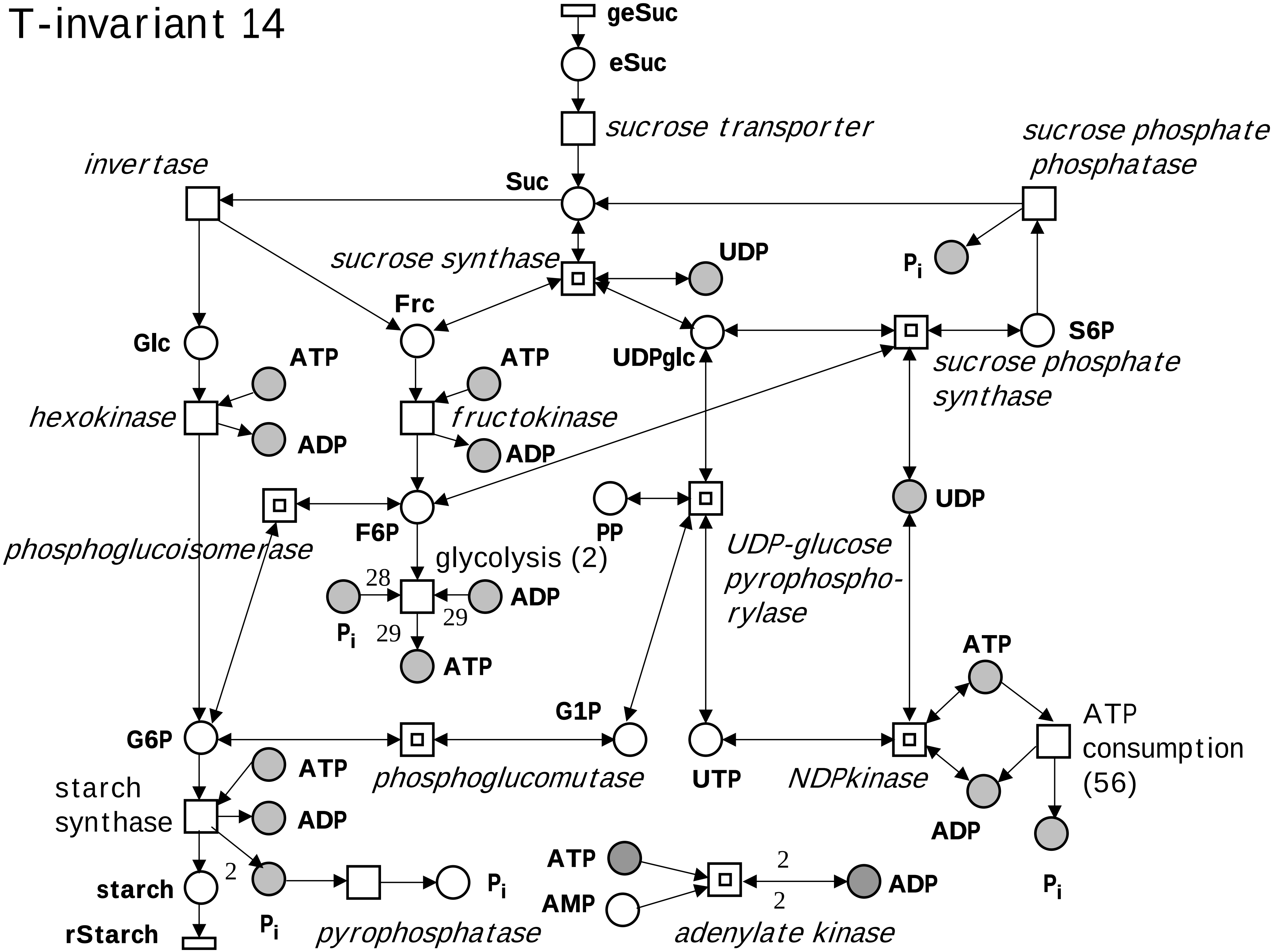
T-invariant 14



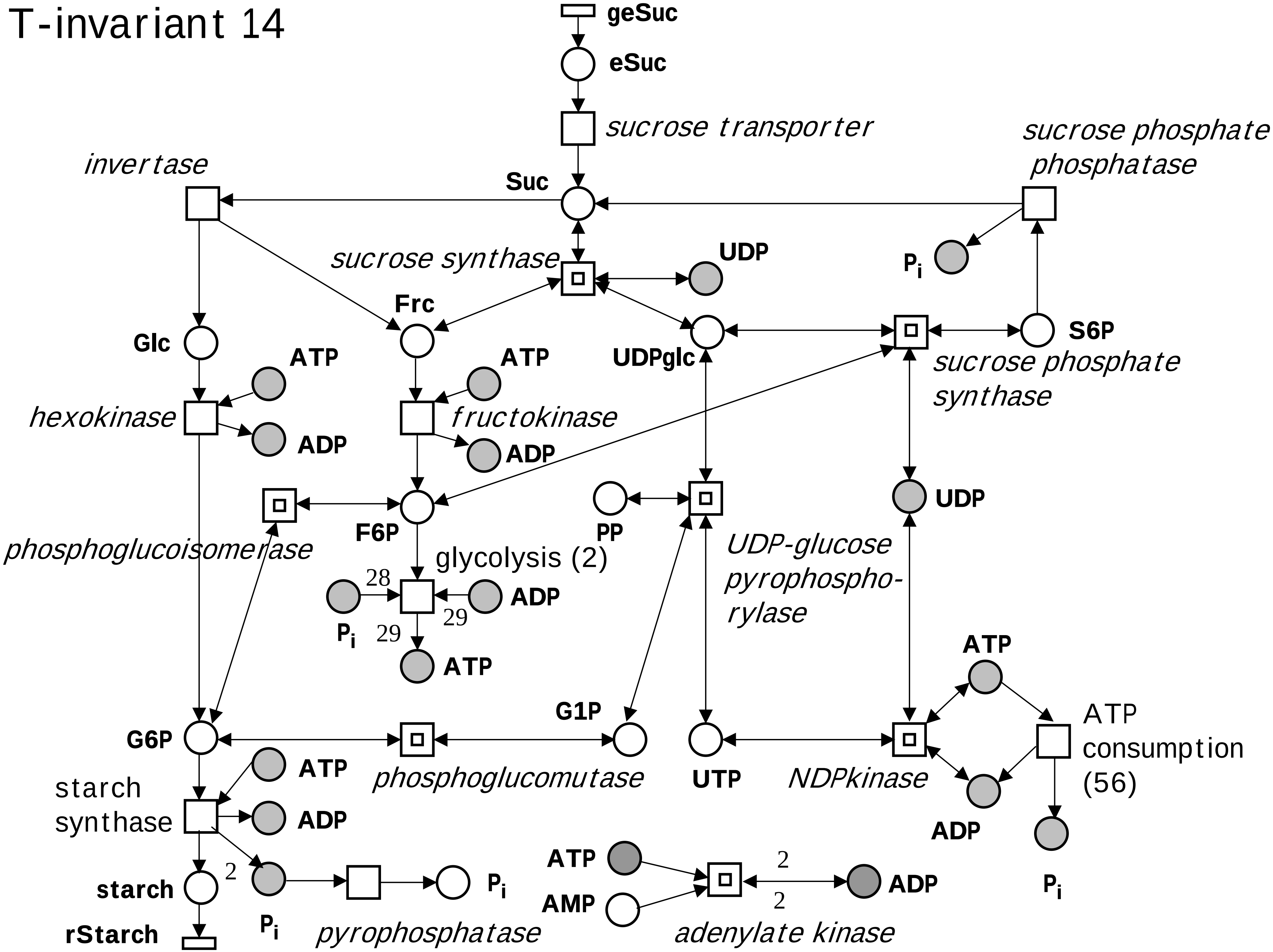
T-invariant 14



T-invariant 14



T-invariant 14



Qualitative analysis using INA

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

(Voit, Computational Analysis of Biochemical Systems, Cambridge University Press 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

Qualitative analysis using INA

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

Summary & Outlook

- Petri net basics
- Metabolic Petri nets: unique description of biological networks

- Qualitative analysis: model checking
  calculation of systems properties
- Modelling, simulation, analysis: sucrose-to-starch-pathway
  in potato tuber

- Used free available tools:
  Editing: `Ped` M. Heiner BTU Cottbus
  Simulation: `Pedframe` http://www.informatik.tu-cottbus.de/~wwwdssz/
  qualitative Analysis: `INA` P.H.Starke HU Berlin
  http://www.informatik.hu-berlin.de/~starke/ina.html

Summary & Outlook

- Glycolysis-pentose phosphate pathway in erythrocytes

  K.Voss, M.Heiner,I.Koch *BioSystems* in press (2004)
- Apoptosis M.Heiner, I.Koch, J. Will *Comp.Methods Syst.Biol.* LNCS **2602**:173 (2003)

  M.Heiner, K.Voss, I.Koch *In Silico Biology* 3: 031 (2003)

Ongoing projects:

1. The whole E.coli pathway Nina Kramer
2. The whole potato tuber pathway Björn Junker/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!

Qualitative analysis using INA