Petri nets for qualitative modelling of biological networks

Claudine Chaouiya

Technologies Avancées pour le Génome et la Clinique

TAGC ERM 206

Université de la Méditerranée

chaouiya@univmed.fr

Petri net modelling

✓Mathematical and graphical formalism

✓Representation of concurrency/parallelism

✓Strong mathematical foundations

Properties

- Structural → P-invariants (conservative components)
  → T-invariants (repetitive components)
- Dynamical → liveness
  → boundness
  → reachability

- conservation
- flux modes
- stable states / equilibrium
- limited concentrations
- paths in the dynamics

Tools

- Analytical approaches → state equations, algebraic equations, graph analysis...
- Model checking
- Simulation

a variety of analysis tools and simulation shells available

Extensions Stochastic PN, Coloured nets, Hybrid nets...

Biological networks ?

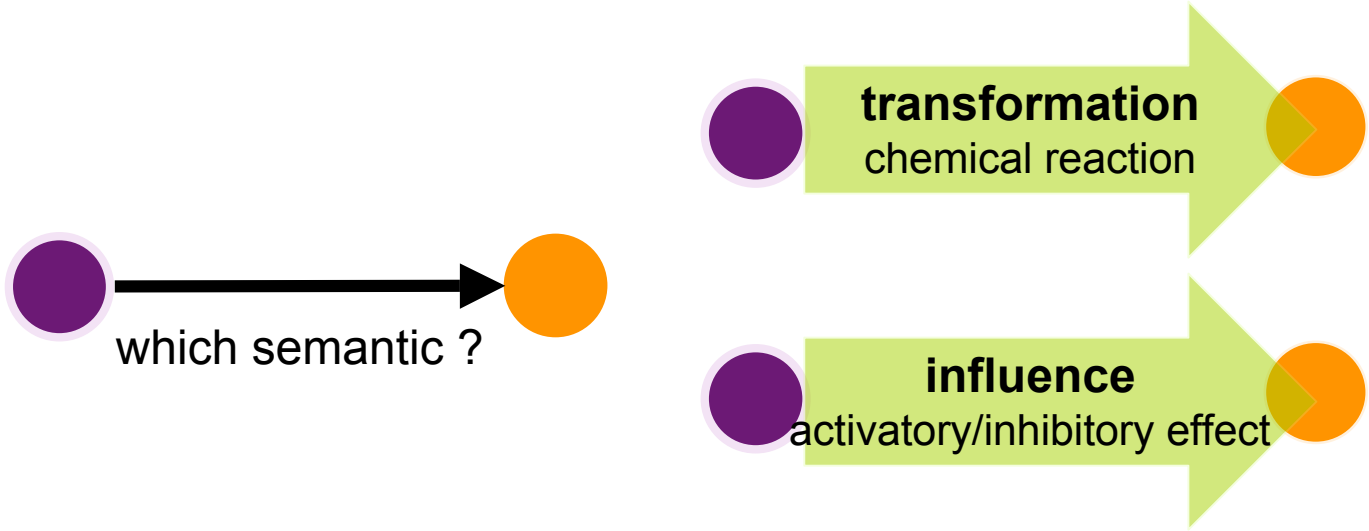
Different **abstraction levels** depending on

- The biological questions
- The nature and quality of available data

Molecular level: **biochemical network**

Gene cross-regulation level: **genetic network**

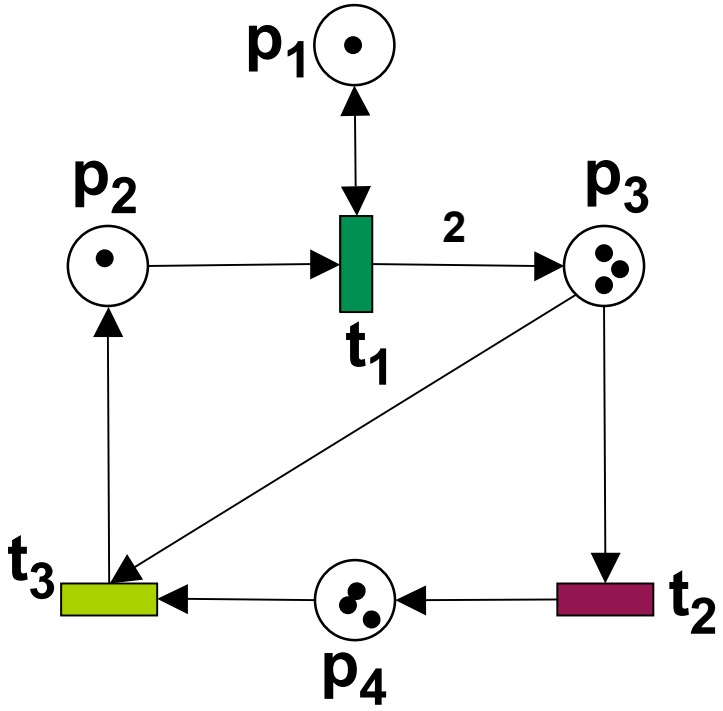
Tissue level: **inter-cellular network**



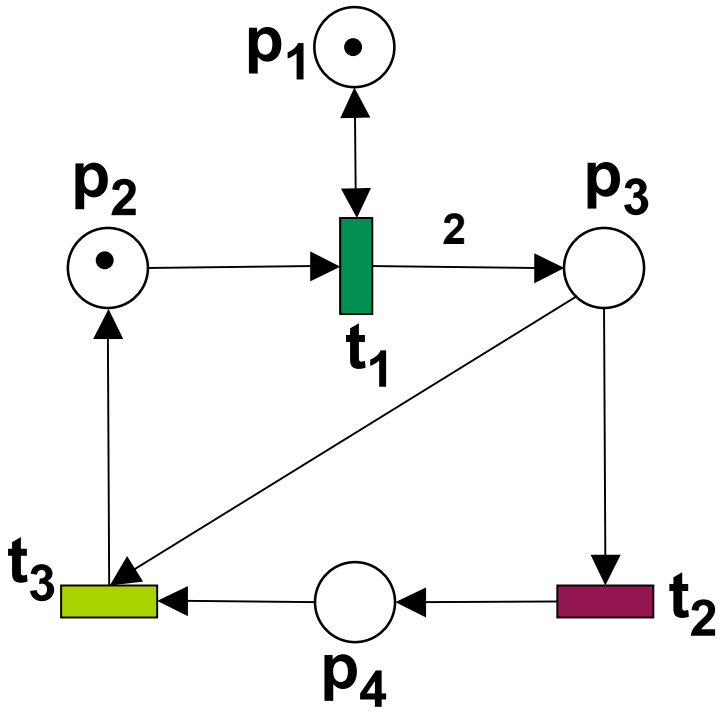
Outline

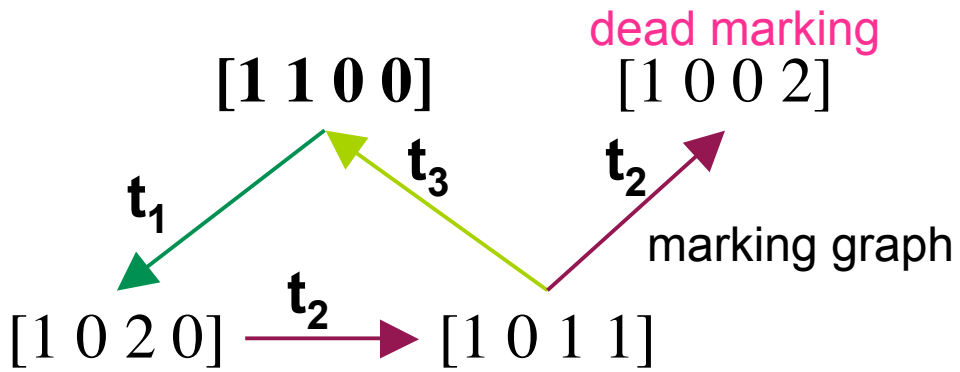
- Petri net basics
- Standard PN modelling of biochemical networks
- Standard PN modelling of *logical* regulatory networks
- Towards an integrated qualitative modelling of regulated metabolic networks

Petri net basics



Petri net basics





initial marking

$$M_0 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

PxT→ IN

$$Pre = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

TxP→ IN

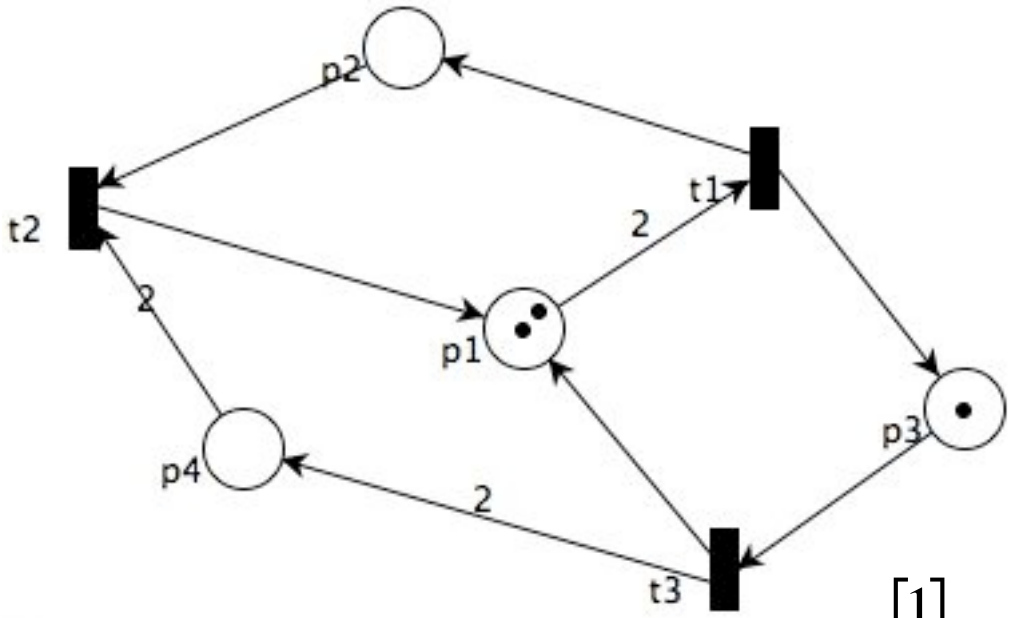
$$Post = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$C = Post^{\mathrm{T}} - Pre,$$

$$M_1 = M_0 + C \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix},$$

$$M_0[t_1>M_1 .$$

Petri net basics - invariants



$$C = \begin{bmatrix} -1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & -2 & 2 \end{bmatrix},$$

$$\text{T} - \text{invariant} : Cy = 0, \quad y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

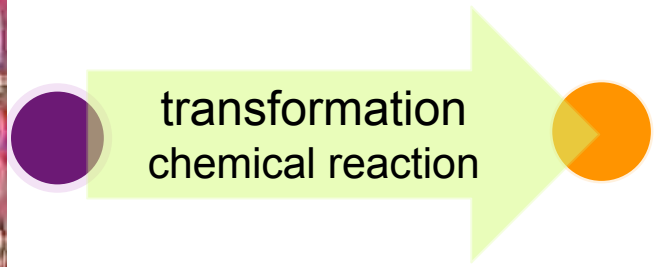
$$\text{P} - \text{invariants} : C^T x = 0, \quad x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 2 \\ 0 \\ 4 \\ 1 \end{bmatrix}$$

Outline

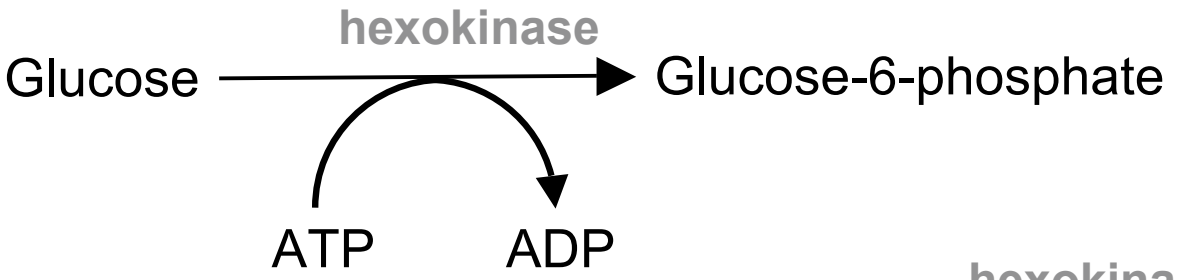
- Petri net basics
- **PN modelling of biochemical networks**
- PN modelling of *logical* regulatory networks
- Towards an integrated qualitative modelling of regulated metabolic networks

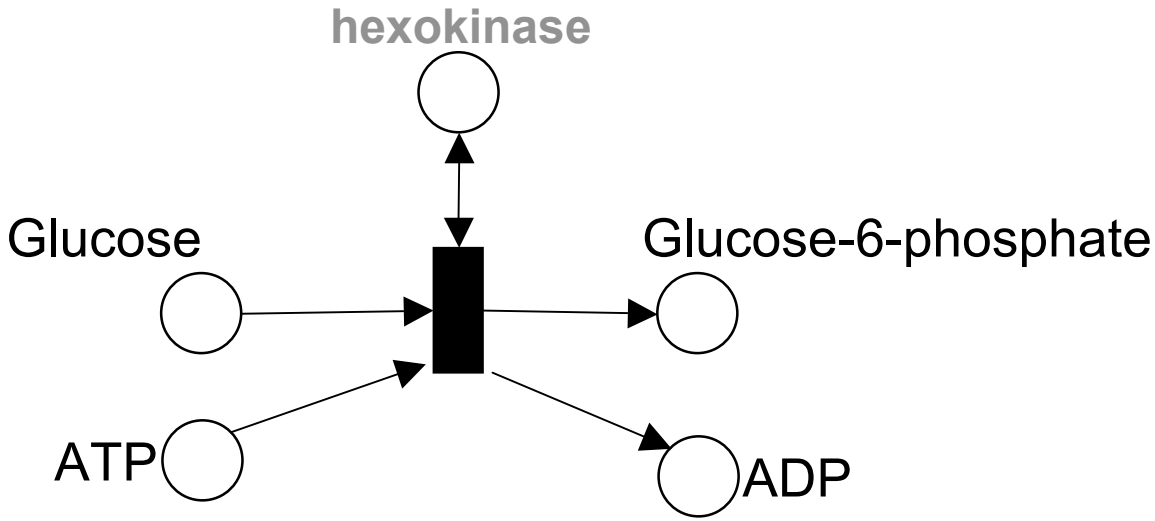
Modelling of biochemical networks

Qualitative PN modelling of metabolic networks



- **places:** reactants, products, enzymes ...
- **transitions:** reactions, catalysis ...
- **weighted arcs:** stoichiometry





Modelling of biochemical networks

PN based stoichiometric analysis (structural analysis)

- stoichiometry *i.e.* the reactants involved + molar ratios of consumption and production
- reaction directionality (reversible / irreversible)
- catalysing enzymes
- dynamics neglected

m, number of metabolites
q, number of reactions
N, (*q* x *m*) stoichiometric matrix (N_{ij} = stoichiometric coef. of i in reaction j)

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N}.\mathbf{r}(t) = 0,$$

in the PN representation C, the incidence matrix is the stoichiometric matrix

$\mathbf{c}$ $(m \text{ x} 1)$ vector of current metabolite concentrations,

$\mathbf{r}(t)$ $(q \text{ x } 1)$ a flux distribution.

Modelling of biochemical networks

PN based stoichiometric analysis (structural analysis)

(Pseudo)-steady state assumption, high turnover of substances (when compared to regulatory events), on longer time scales, metabolite concentrations and reaction rates are supposed constant:

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N}.\mathbf{r}(t) = 0$$

Modelling of biochemical networks

PN based stoichiometric analysis (structural analysis)

(Pseudo)-steady state assumption, high turnover of substances (when compared to regulatory events), on longer time scales, metabolite concentrations and reaction rates are supposed constant:

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N}.\mathbf{r}(t) = 0$$

Conservation relations: weighted sums of metabolite concentrations which remain constant in the system

$$\mathbf{N}^{\mathsf{T}}.\mathbf{y} = 0$$

correspond to P-invariants in PNs

Modelling of biochemical networks

PN based stoichiometric analysis (structural analysis)

(Pseudo)-steady state assumption, high turnover of substances (when compared to regulatory events), on longer time scales, metabolite concentrations and reaction rates are supposed constant:

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N}.\mathbf{r}(t) = 0$$

Conservation relations: weighted sums of metabolite concentrations which remain constant in the system

$$\mathbf{N}^{\mathsf{T}}.\mathbf{y} = 0$$

correspond to P-invariants in PNs

Elementary flux modes (EFMs), correspond to non decomposable steady state flux distributions using a minimal set of reactions (Schuster *et al.*,1999)

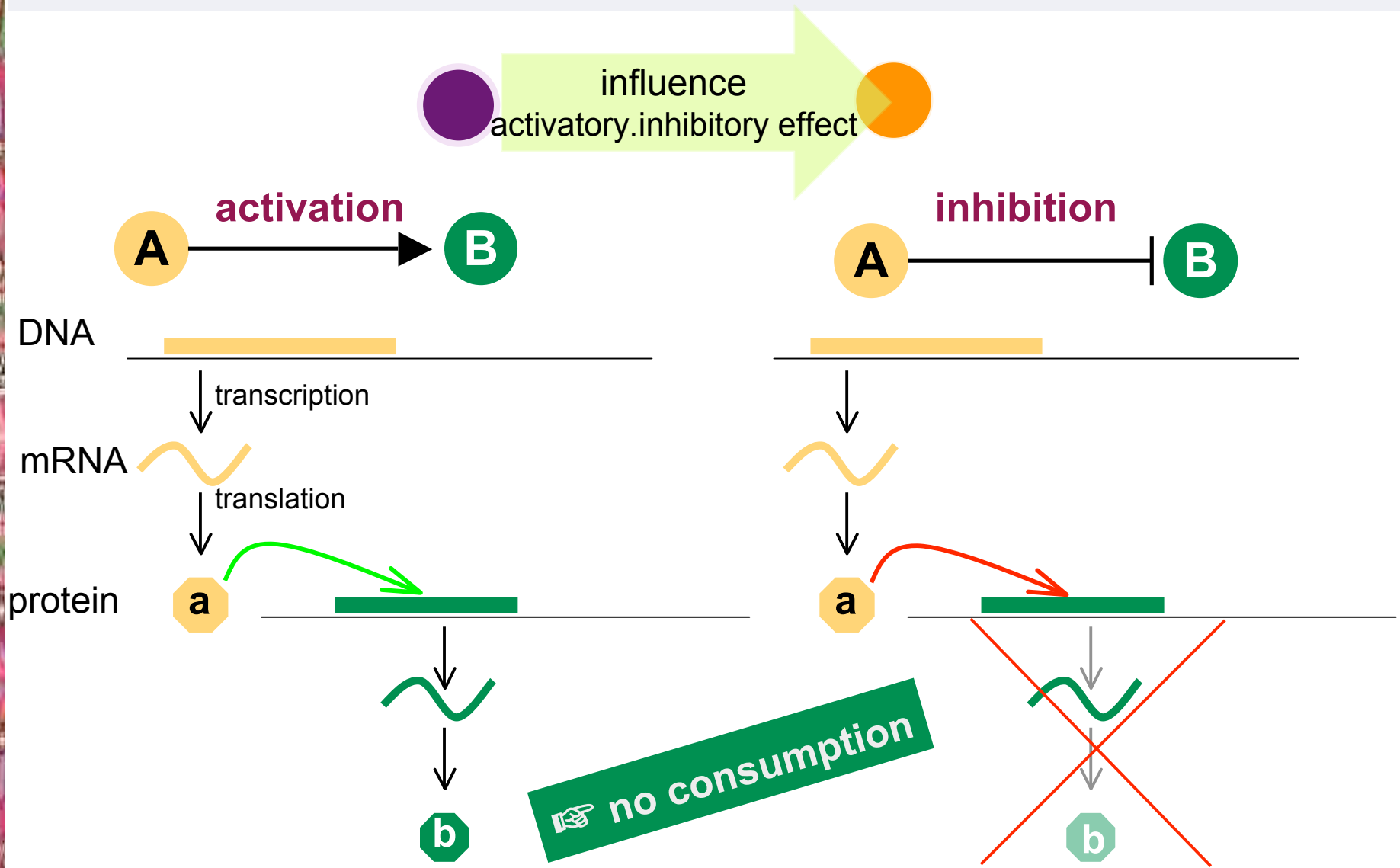
$$\mathbf{N}.\mathbf{r} = 0\,, \quad r_i \geq 0, \text{ if reaction } i \text{ irreversible.}$$

related to T-invariants in PNs

Outline

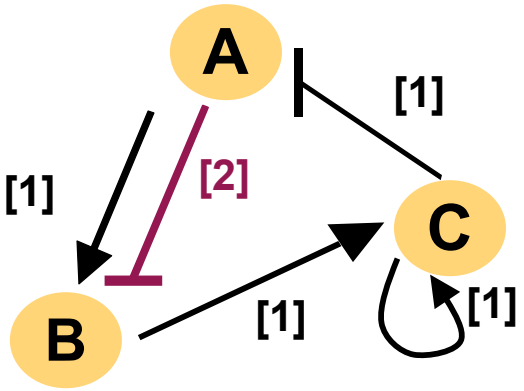
- Petri net basics
- PN modelling of biological networks, short overview
- Qualitative PN modelling of metabolic networks
- **PN modelling of *logical* regulatory networks**
- Towards an integrated qualitative modelling of regulated metabolic networks

Genetic regulatory networks, a schematic view



"quantity" of protein → level of expression of the corresponding gene

Logical approach
Dynamical modelling of gene networks



- ✓ A graph describes the interactions between genes or regulatory products
- ✓ Discrete levels of expression associated to each gene (logical variables)
- ✓ Levels associated to each interaction

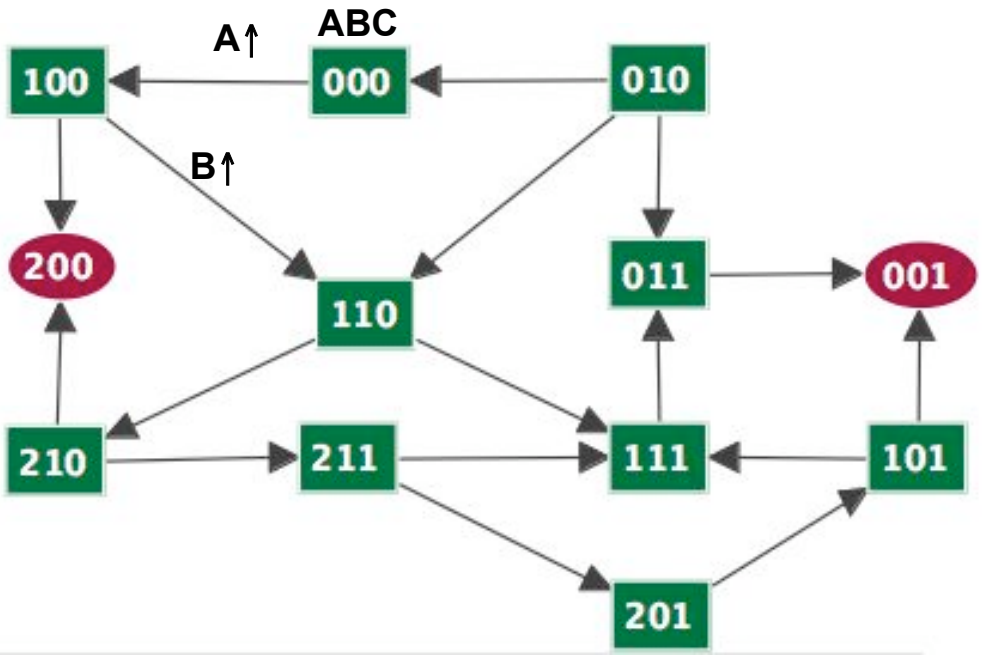
- ✓ ***Logical* parameters**

effects of combinations of incoming interactions

$K_B(\varnothing)=0$

$K_B(\{A,1\})=1$

$K_B(\{A,2\})=0$



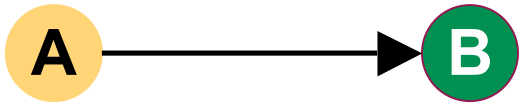
PN modelling of *logical* regulatory networks

Multi-valued Regulatory Petri Nets

Genetic regulatory networks described in terms of logical models (multi-level discretisation)

- ➔ two complementary places for each gene
- ➔ two transitions for each logical parameter (effect of interactions on a given gene)

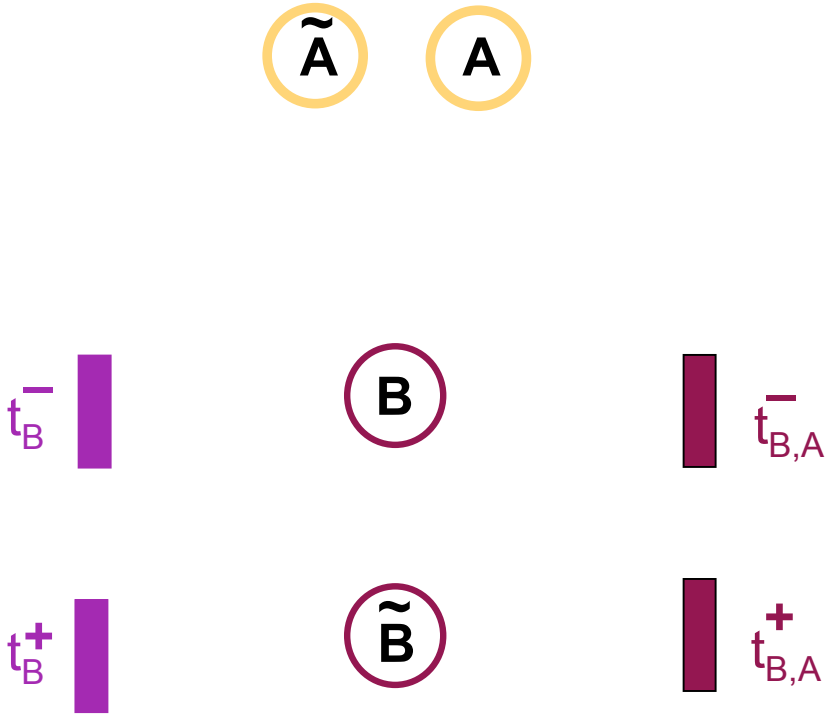
Example



$max_A=1$ $max_B=3$

$K_B(A)=2$ $K_B(\varnothing)=1$

M(A)+M(Ã)=1
M(B)+M(B̃)=3



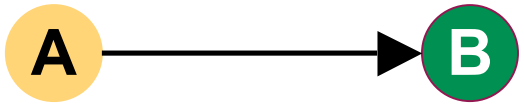
PN modelling of *logical* regulatory networks

Multi-valued Regulatory Petri Nets

Genetic regulatory networks described in terms of logical models (multi-level discretisation)

- ➔ two complementary places for each gene
- ➔ two transitions for each logical parameter (effect of interactions on a given gene)

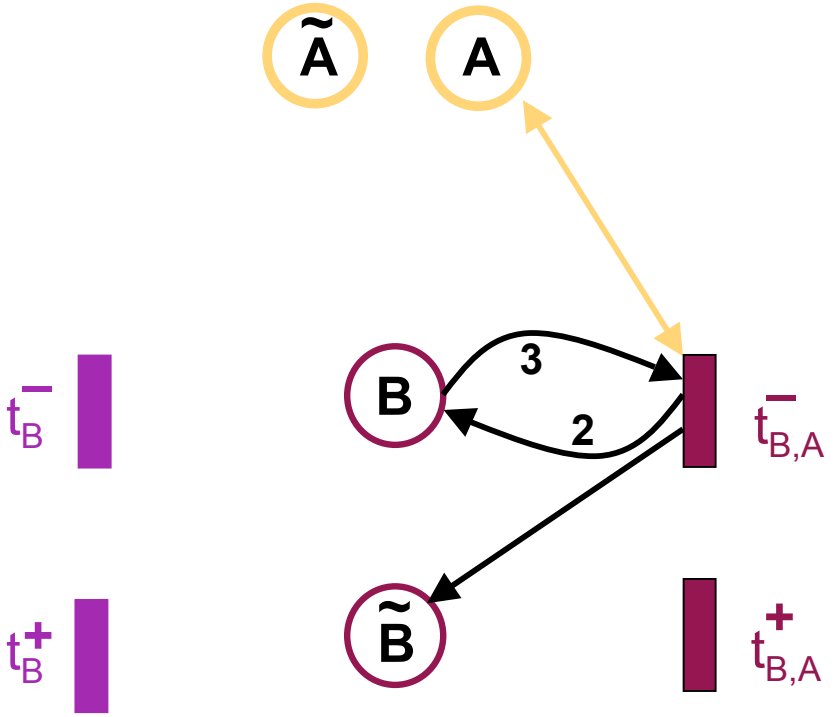
Example



$max_A=1$ $max_B=3$

$K_B(A)=2$ $K_B(\varnothing)=1$

M(A)+M(Ã)=1
M(B)+M(B̃)=3



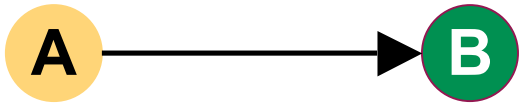
PN modelling of *logical* regulatory networks

Multi-valued Regulatory Petri Nets

Genetic regulatory networks described in terms of logical models (multi-level discretisation)

- ➔ two complementary places for each gene
- ➔ two transitions for each logical parameter (effect of interactions on a given gene)

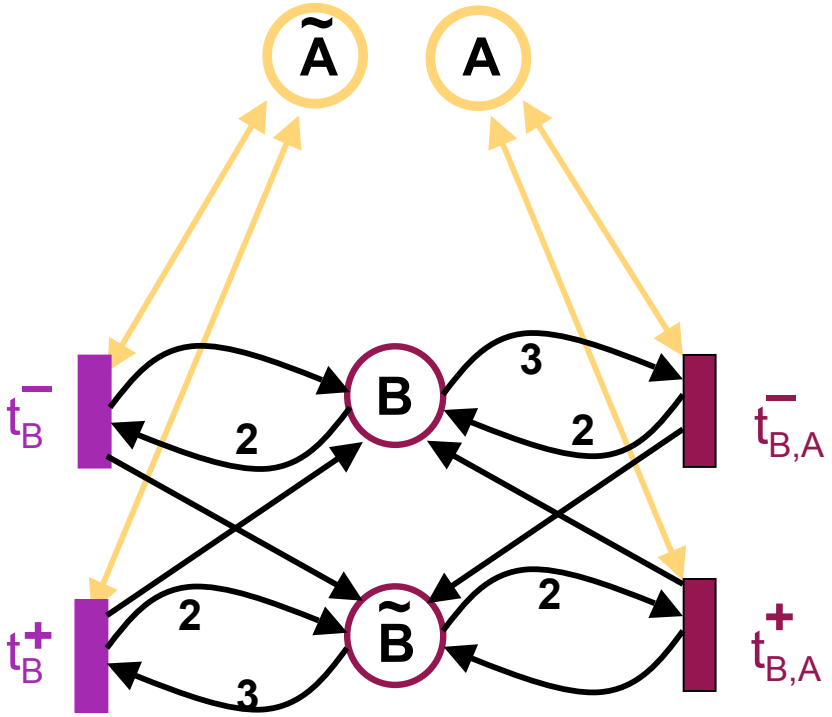
Example

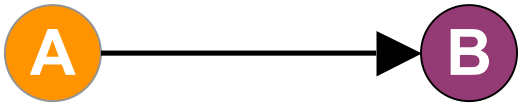


$max_A=1$ $max_B=3$

$K_B(A)=2$ $K_B(\varnothing)=1$

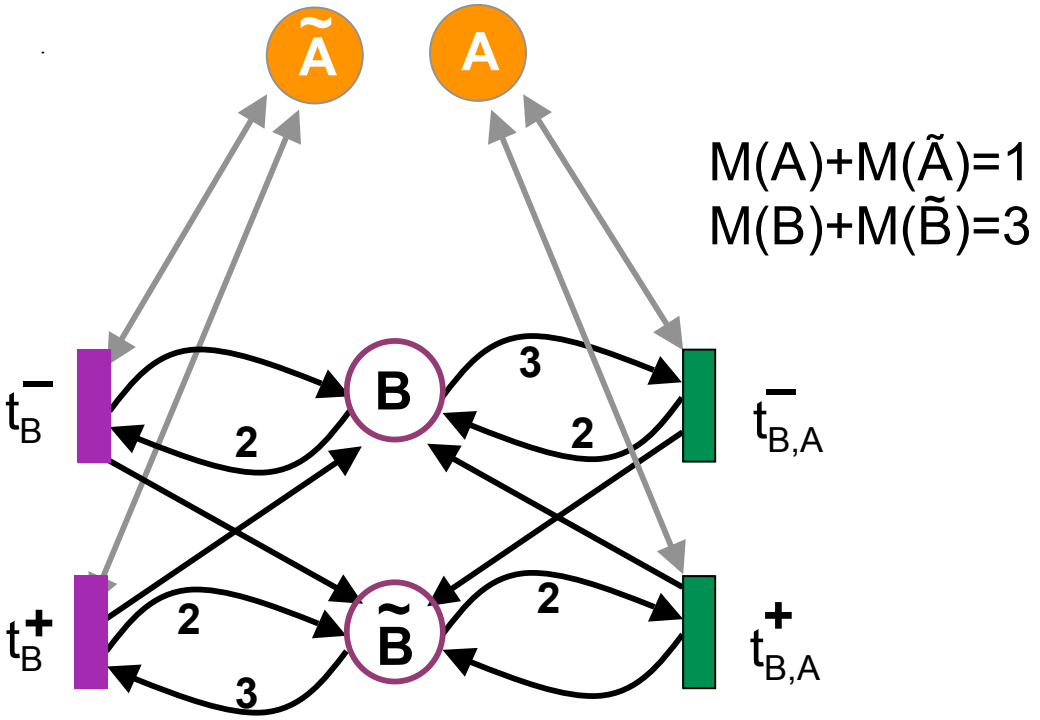
M(A)+M(Ã)=1
M(B)+M(B̃)=3





$max_A=1$ $max_B=3$

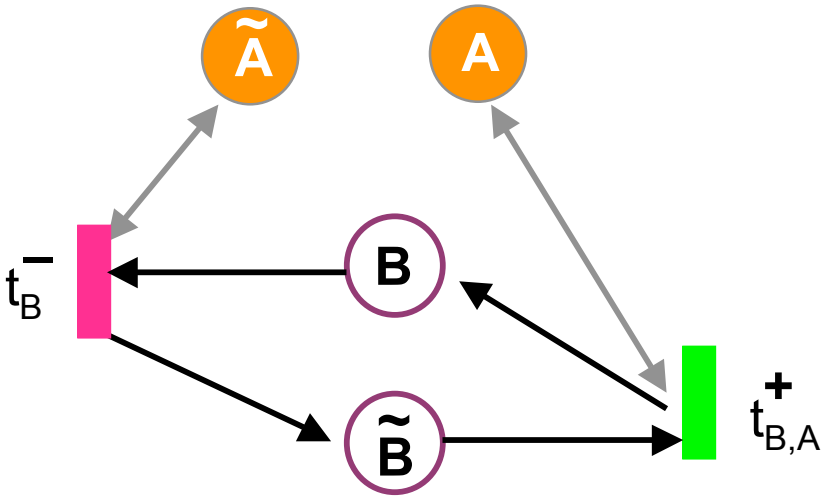
$K_B(A)=2$ $K_B(\varnothing)=1$



$M(A)+M(\tilde{A})=1$
$M(B)+M(\tilde{B})=3$

$K_B(A)=3$ $K_B(\varnothing)=0$

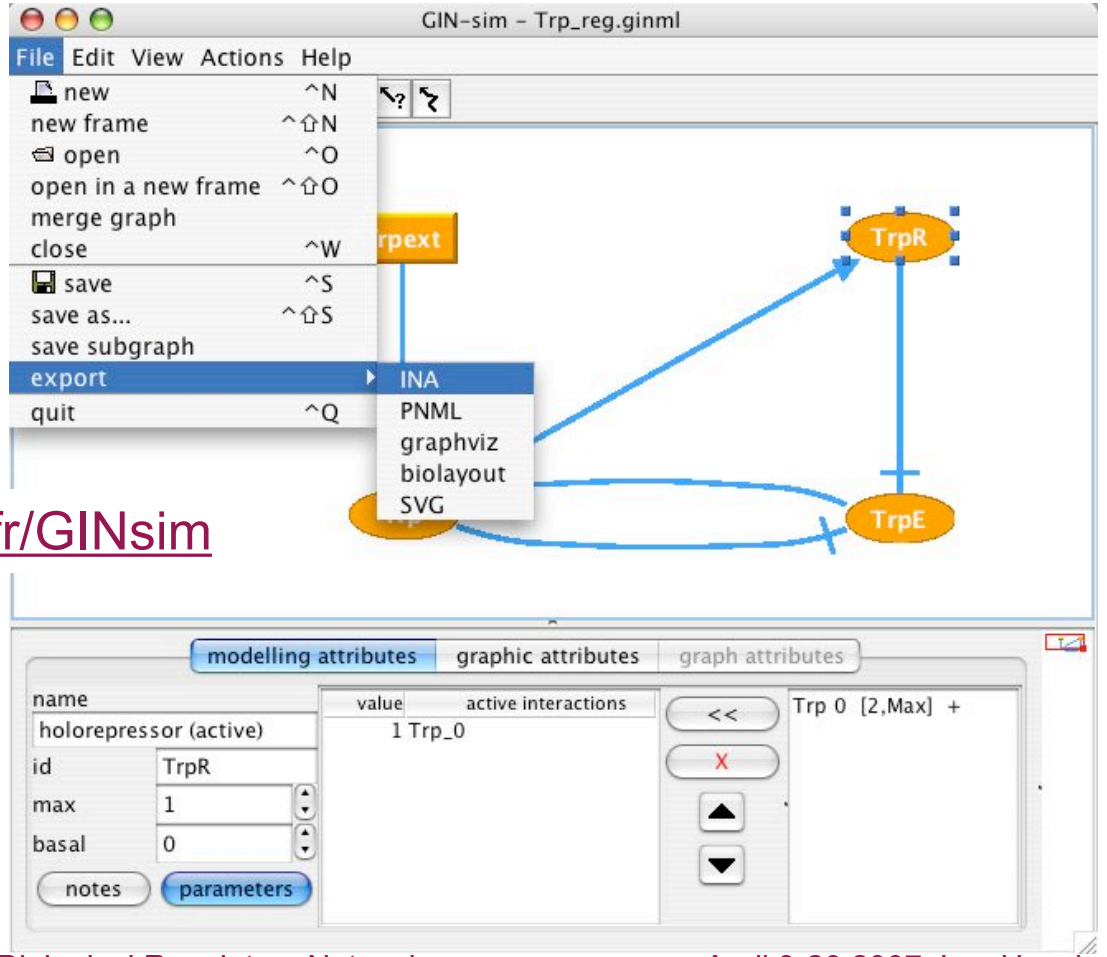
"extremal"
parameter values



PN modelling of *logical* regulatory networks

Multi-valued Regulatory Petri Nets

GINsim, a dedicated software for the definition, simulation, analysis of logical regulatory graphs
provides export facilities to INA format and PNML



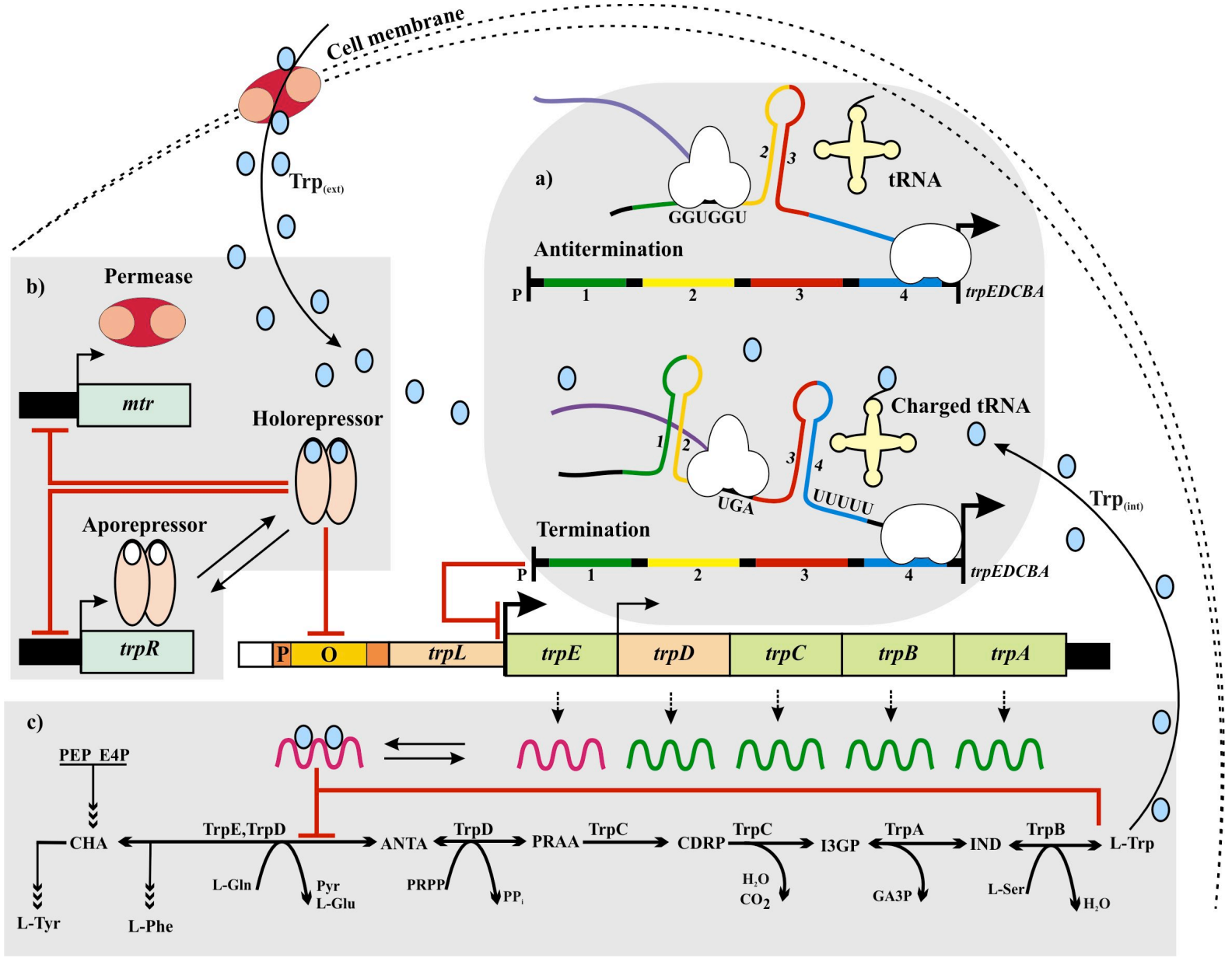
http://gin.univ-mrs.fr/GINsim

Outline

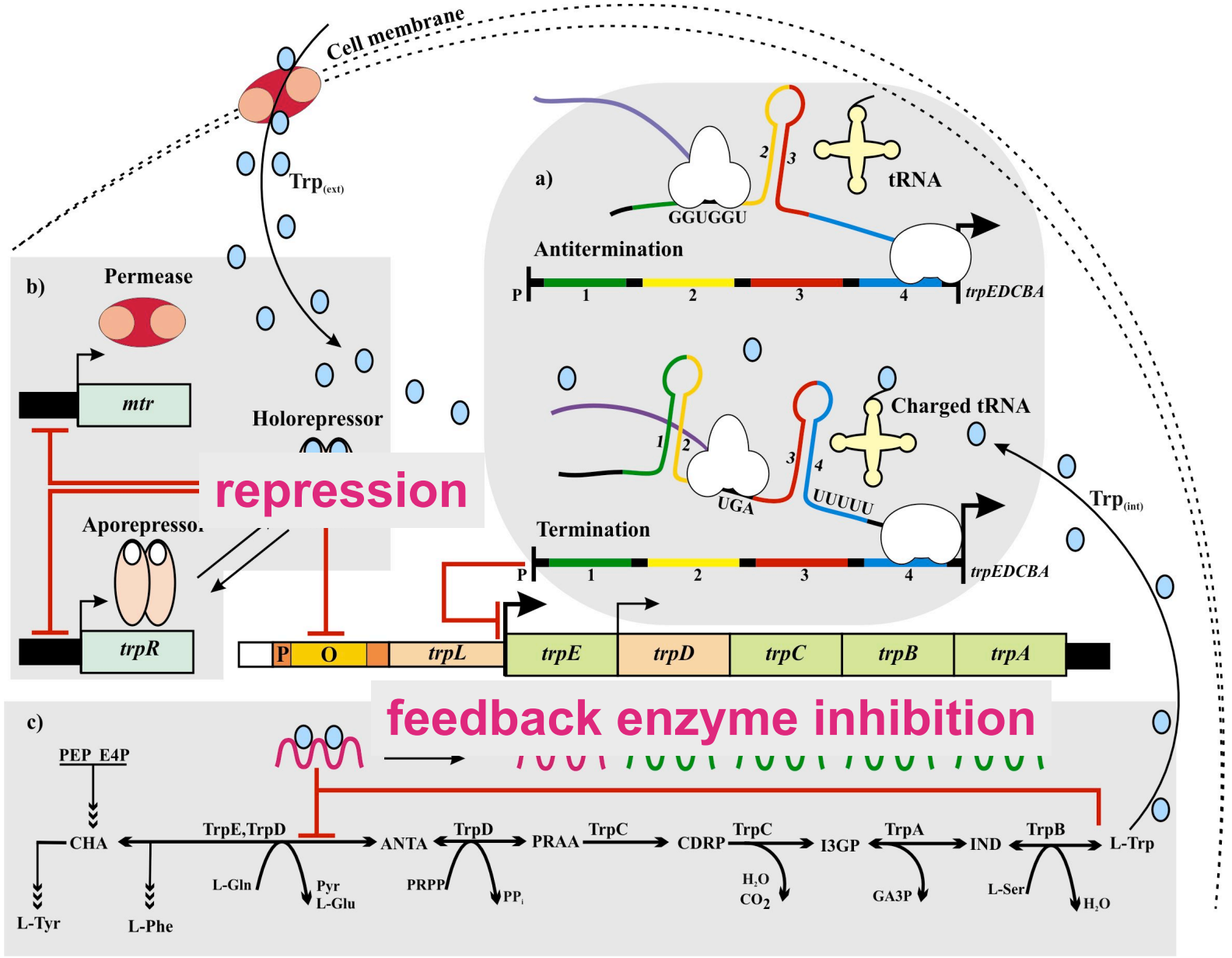
- Petri net basics
- PN modelling of biological networks, short overview
- Qualitative PN modelling of metabolic networks
- PN modelling of *logical* regulatory networks
- **Towards an integrated qualitative modelling of regulated metabolic networks**

E.Simão, E.Remy, D.Thieffry, C.Chaouiya (2005). *Qualitative Modelling of Regulated Metabolic Pathways: Application to the Tryptophan Biosynthesis in E.Coli.* Bioinformatics 21: ii190-196.

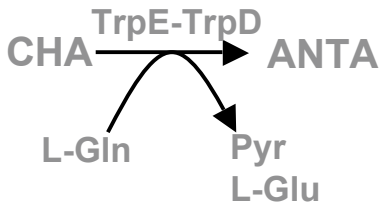
The Tryptophan Biosynthesis in *E. coli*



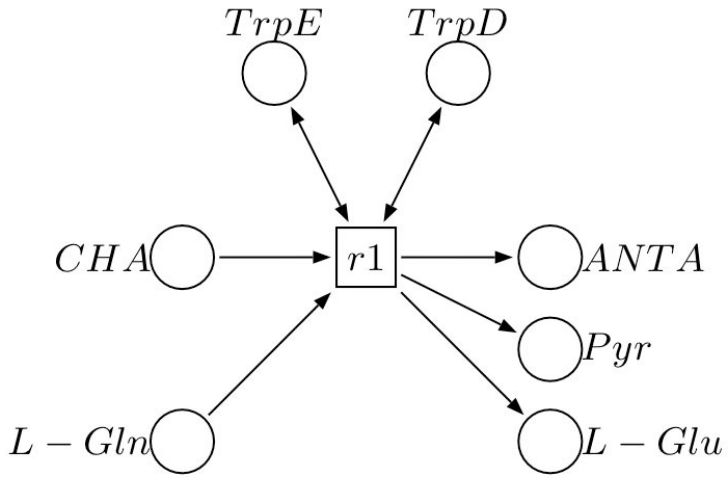
The Tryptophan Biosynthesis in *E. coli*



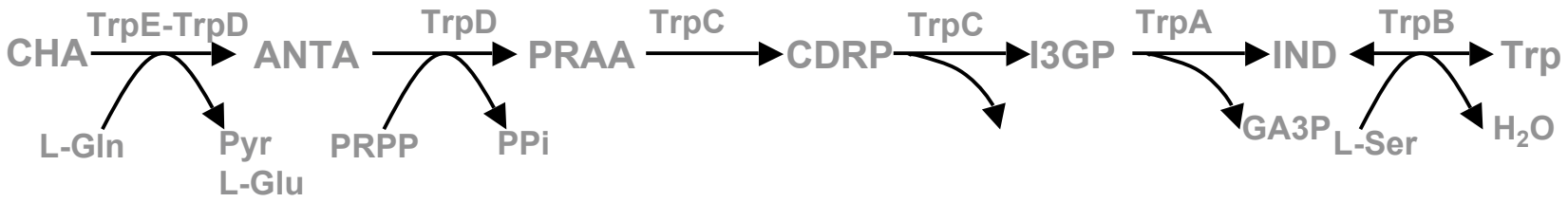
PN model of the metabolic pathway

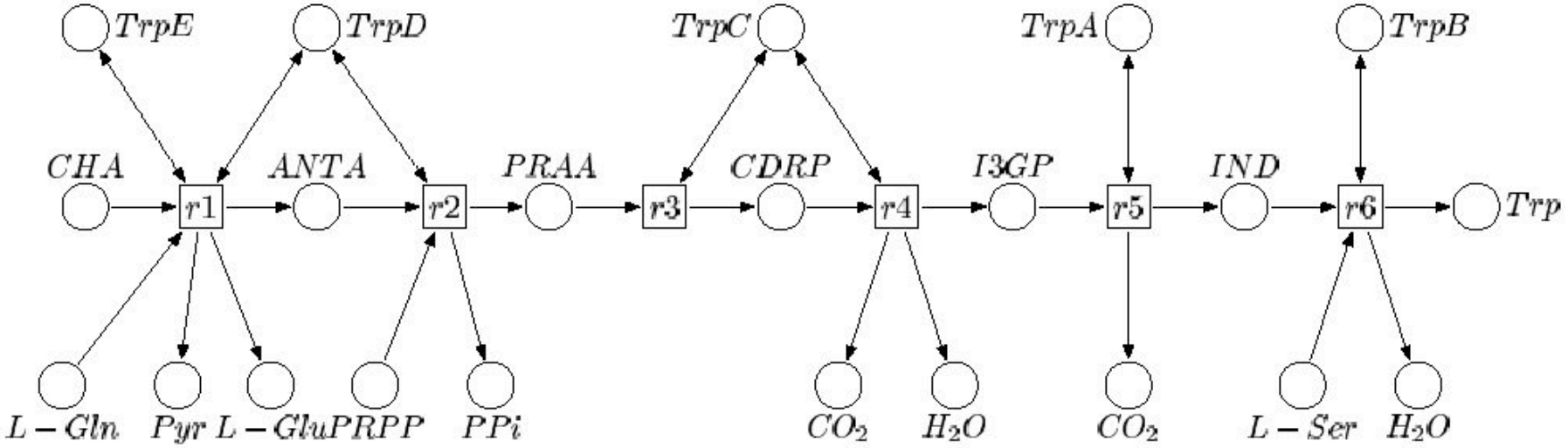


- metabolites / substrates / products / enzymes ... ➔ places
- reactions / catalysis ... ➔ transitions
- stoichiometry ➔ arcs and associated weights



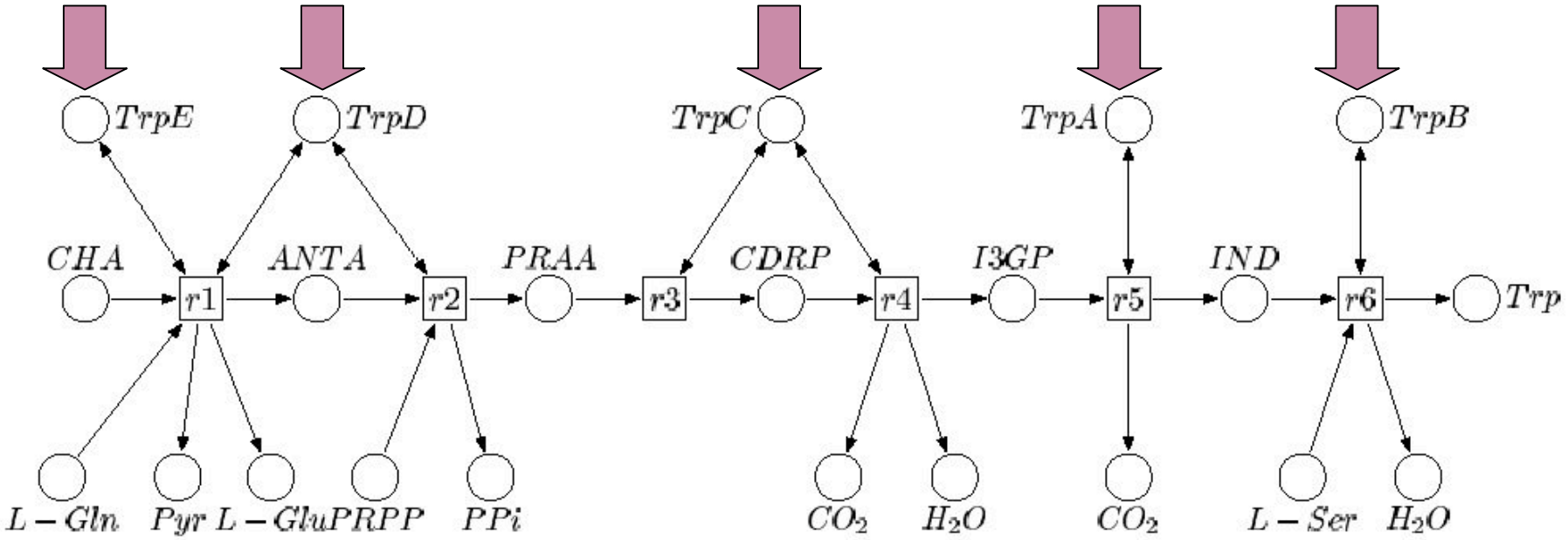
PN model of the metabolic pathway





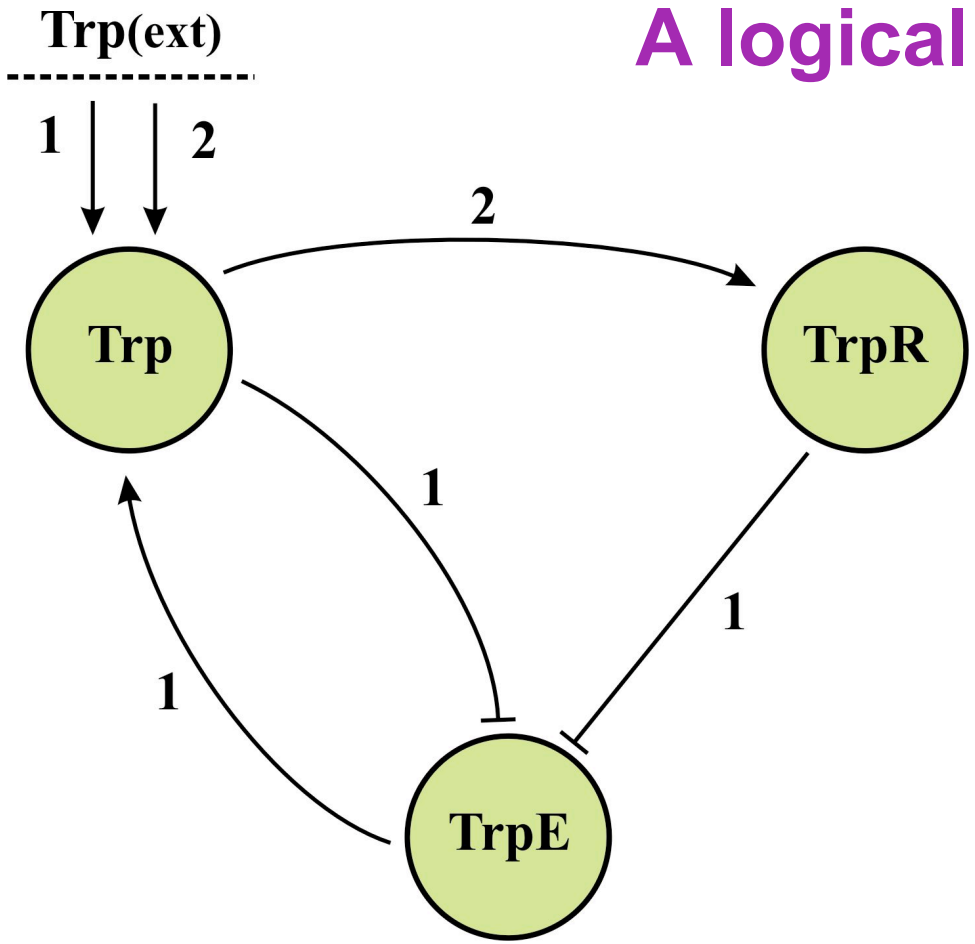
PN modelling of the regulation

REGULATION ???



The Tryptophan biosynthesis regulation

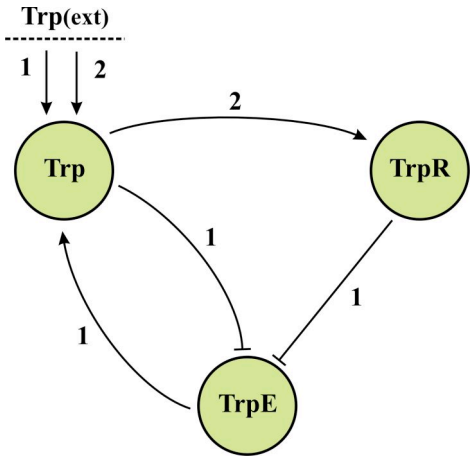
A logical model

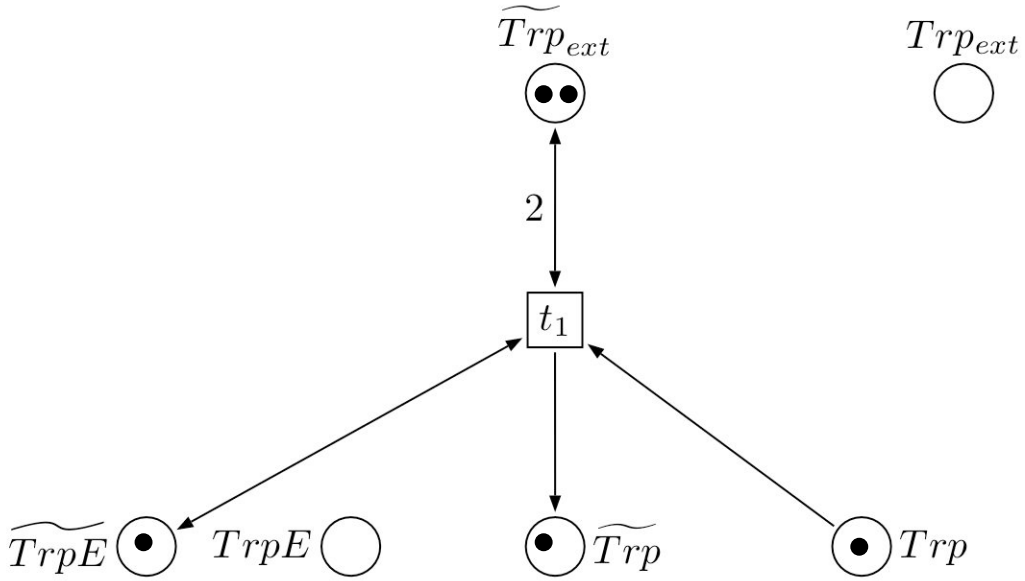


A logical model

The Tryptophan biosynthesis regulation

PN representation



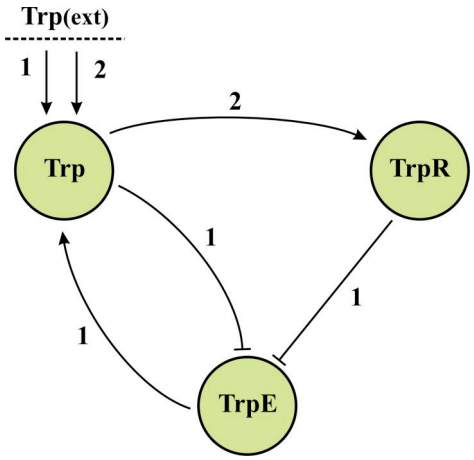


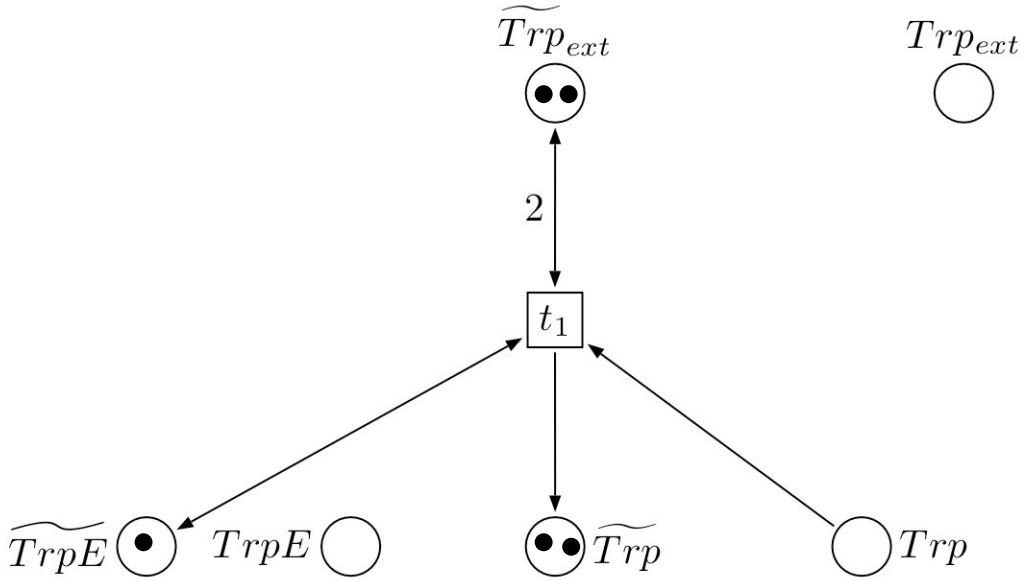
when Trp_{ext}=0 and TrpE=0, Trp→ 0

- gene or regulatory product ... ➔ 2 complementary places
- combination of incoming interactions ➔1 or 2 transitions

The Tryptophan biosynthesis regulation

PN representation



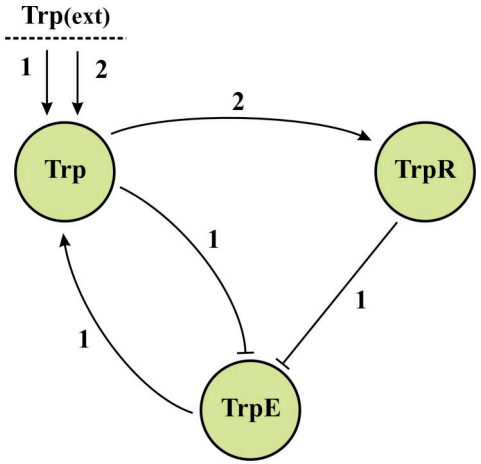


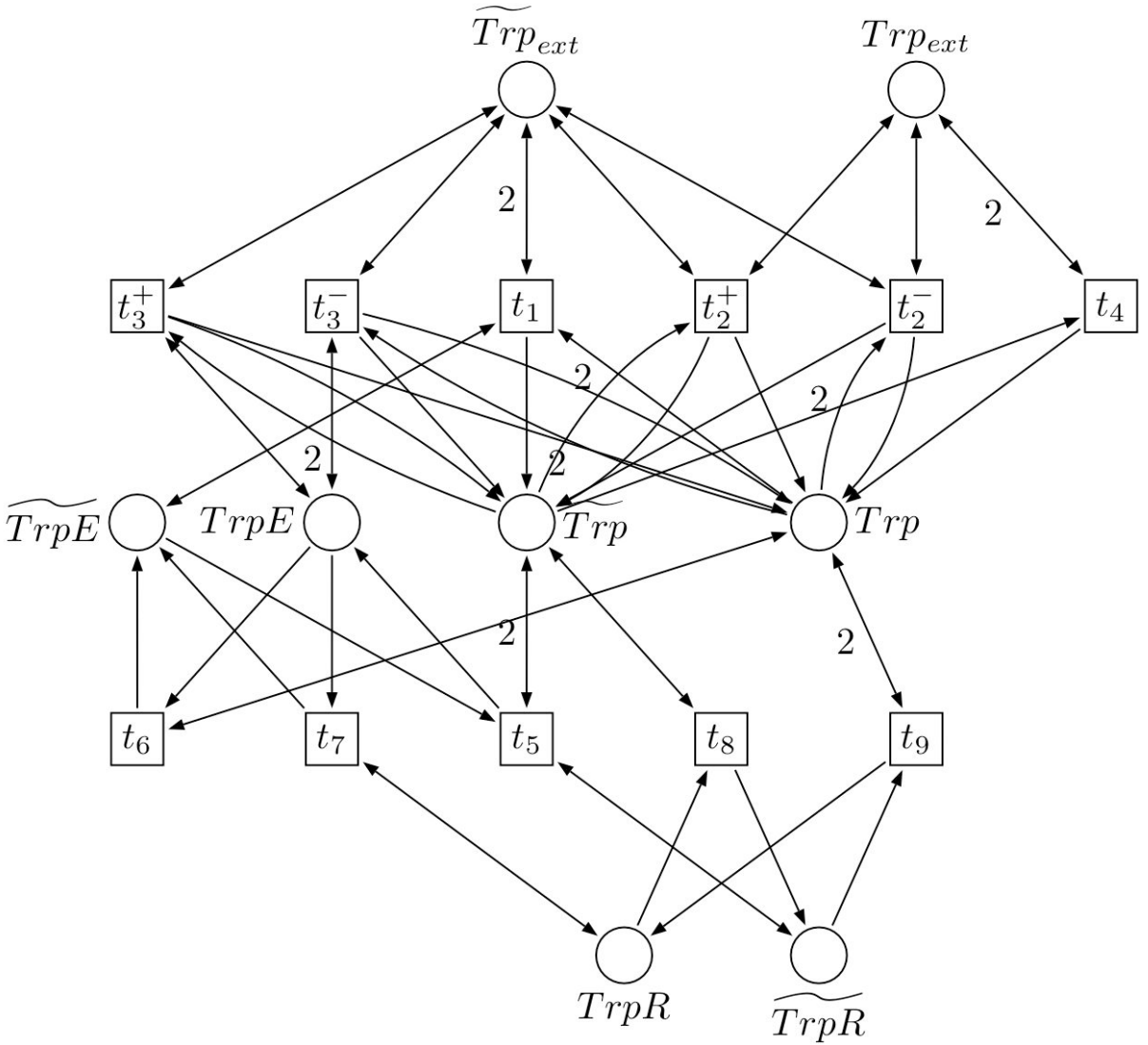
when Trp_{ext}=0 and TrpE=0, Trp→ 0

- gene or regulatory product ... ➔ 2 complementary places
- combination of incoming interactions ➔1 or 2 transitions

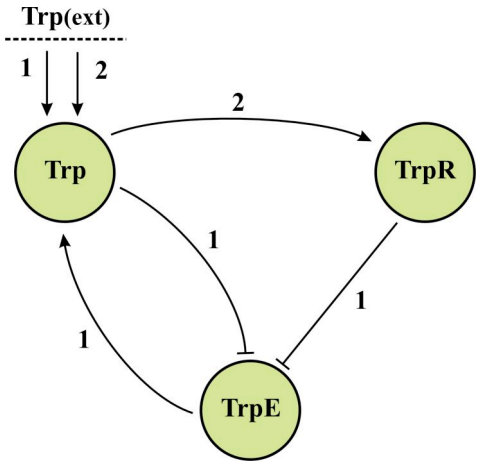
The Tryptophan biosynthesis regulation

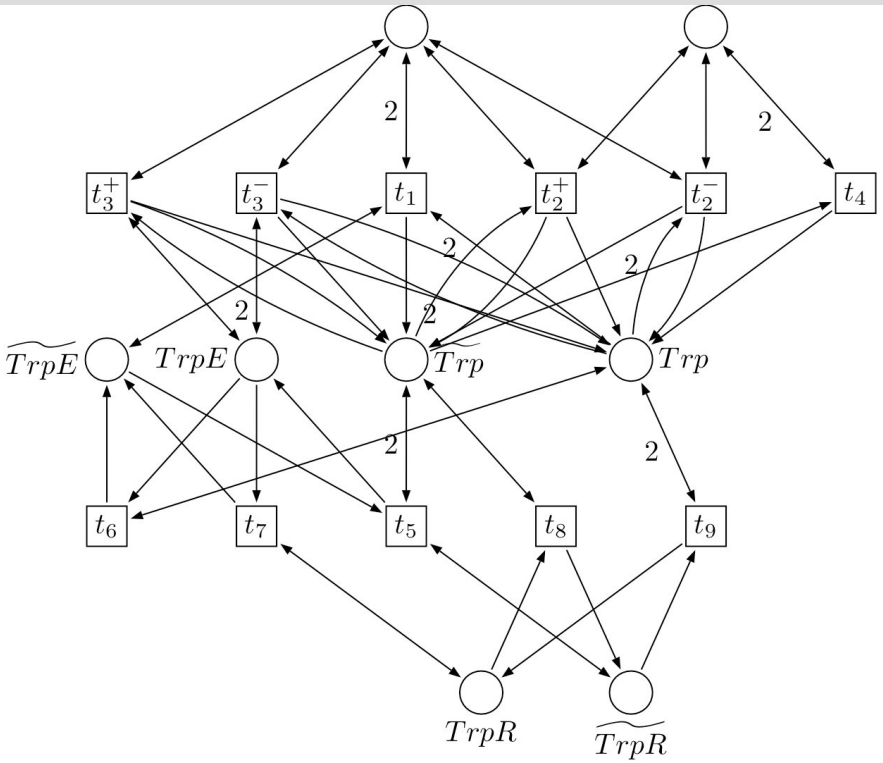
PN representation

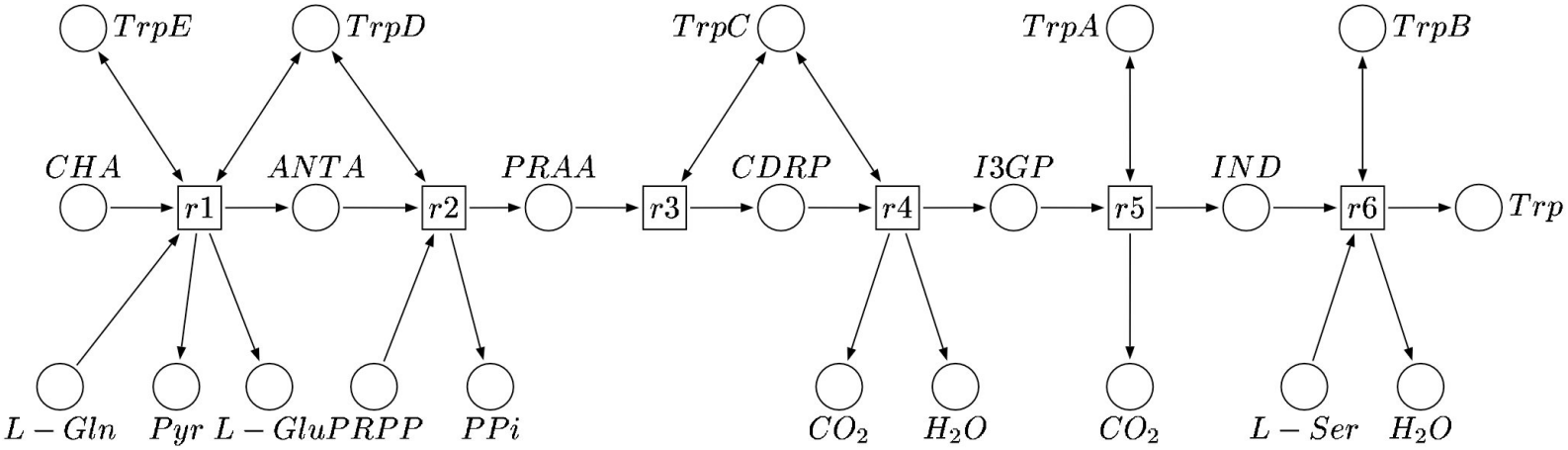




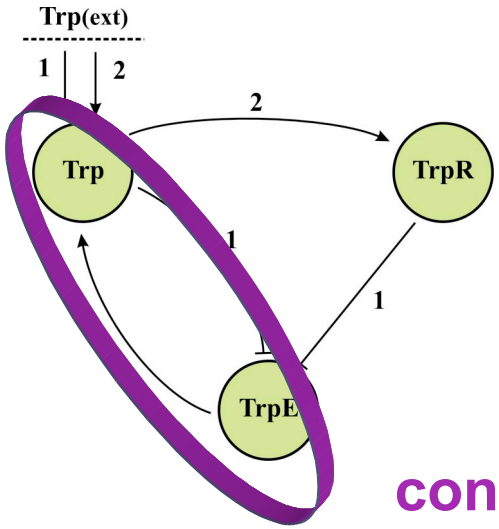
Integrated Petri net modelling

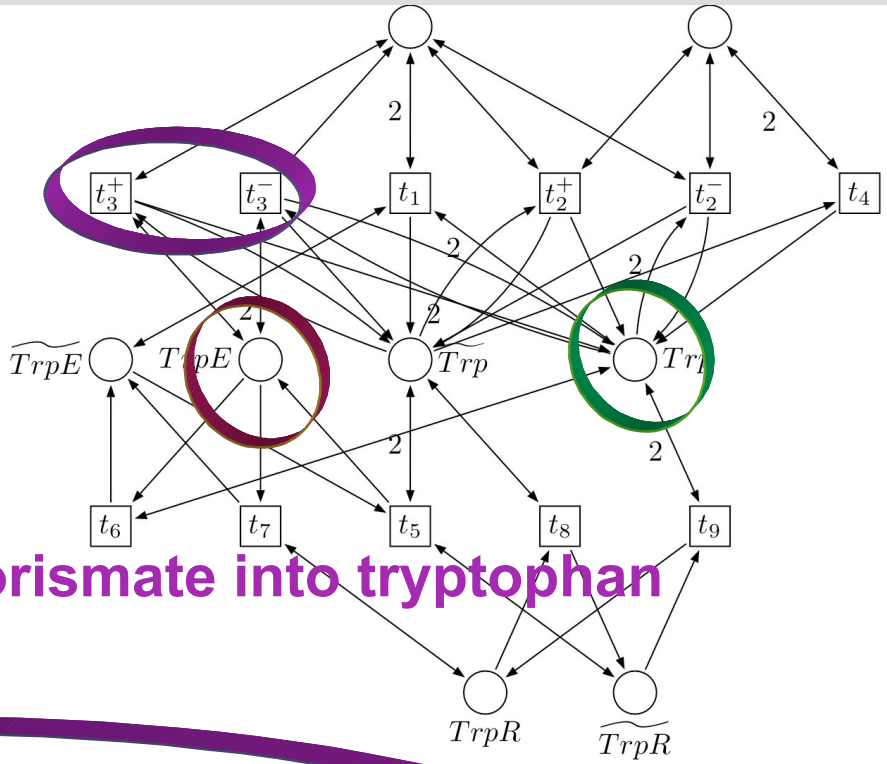




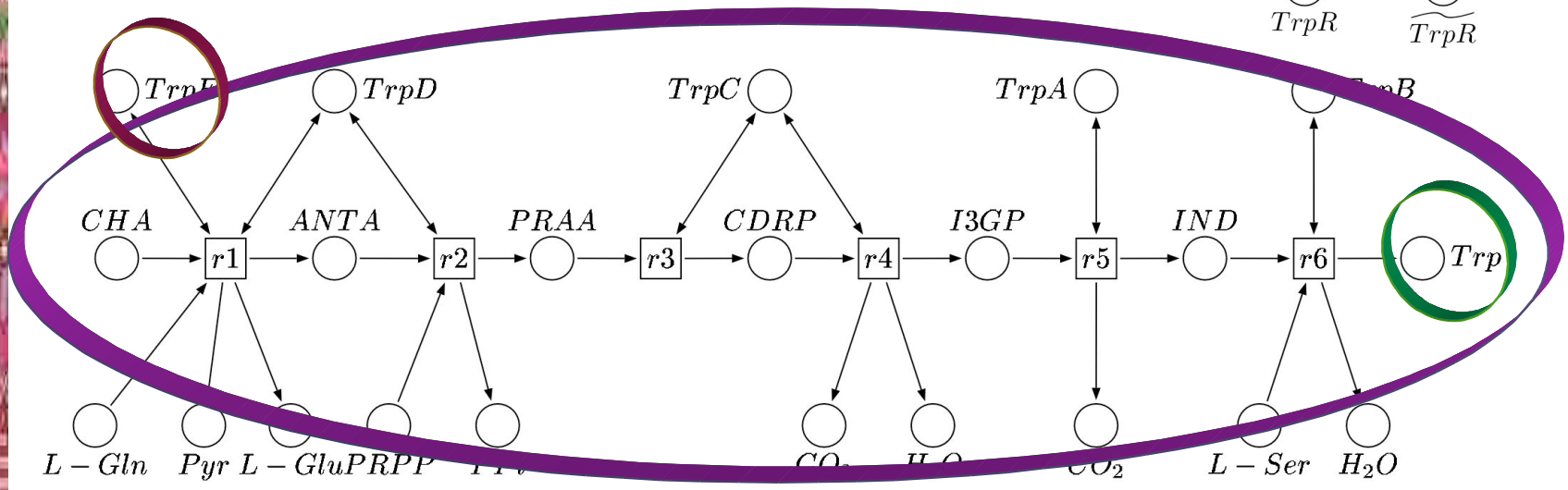


Integrated Petri net modelling

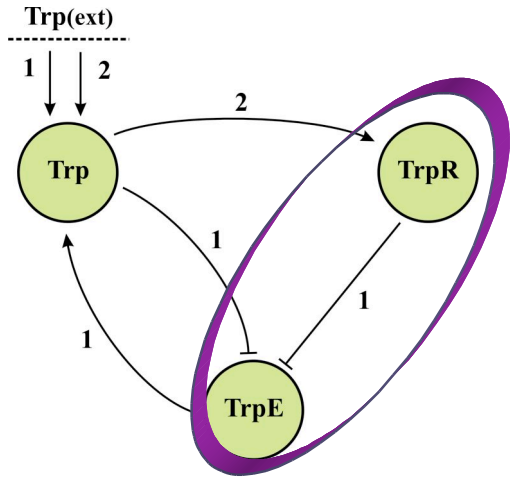


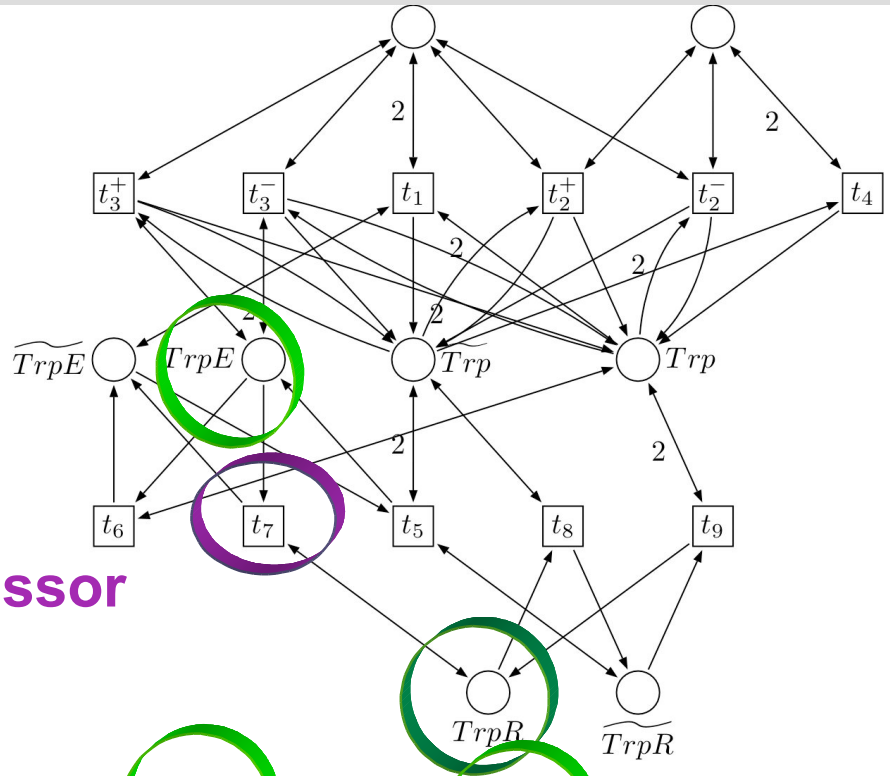


conversion of chorismate into tryptophan

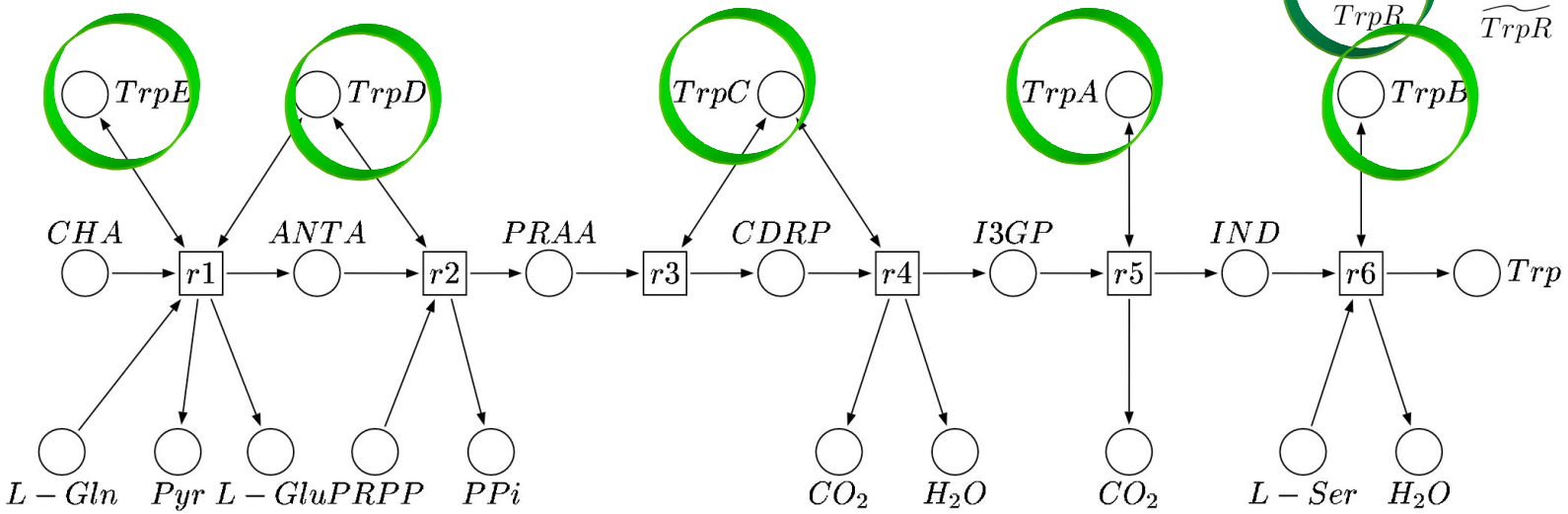


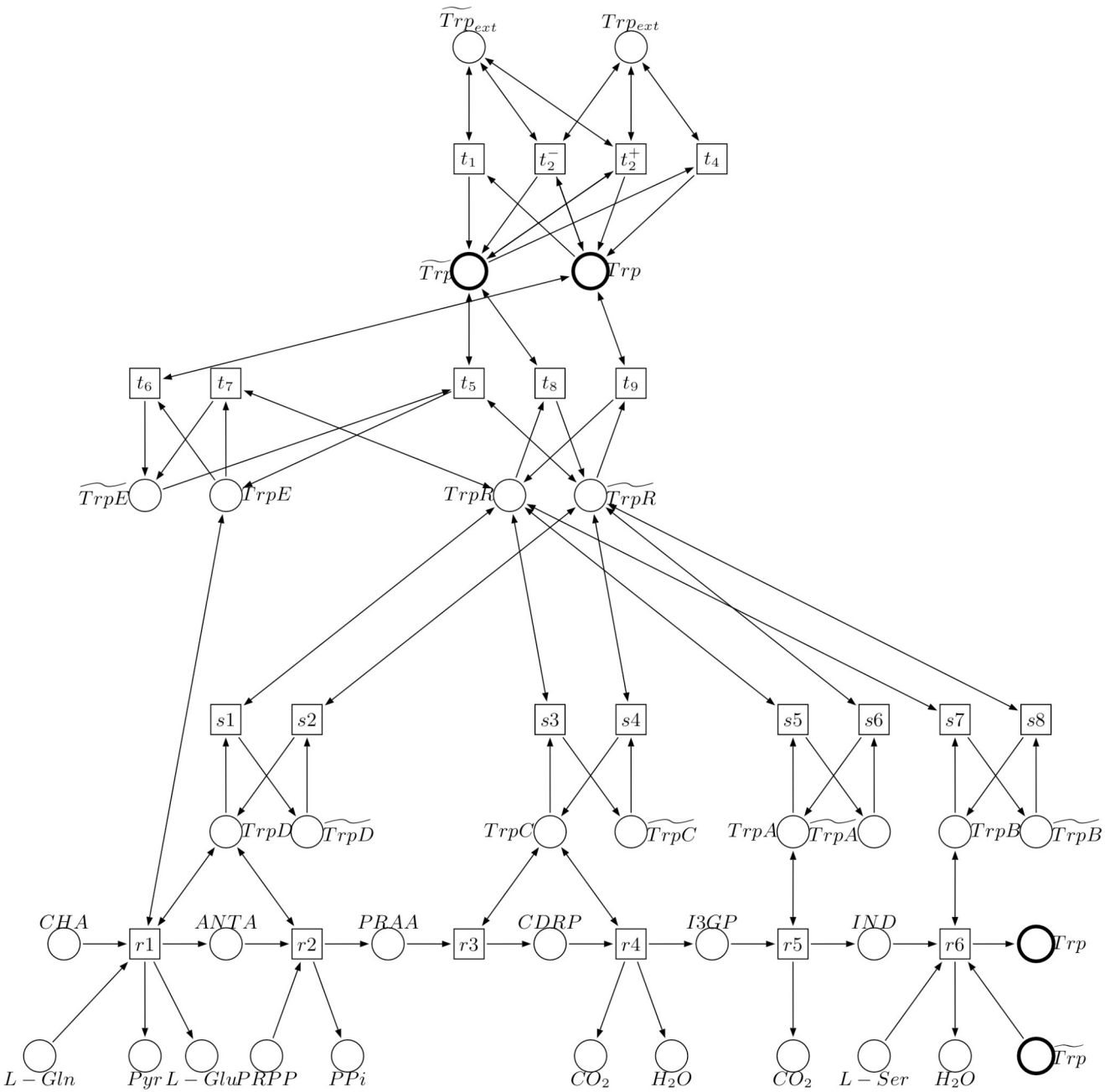
Integrated Petri net modelling



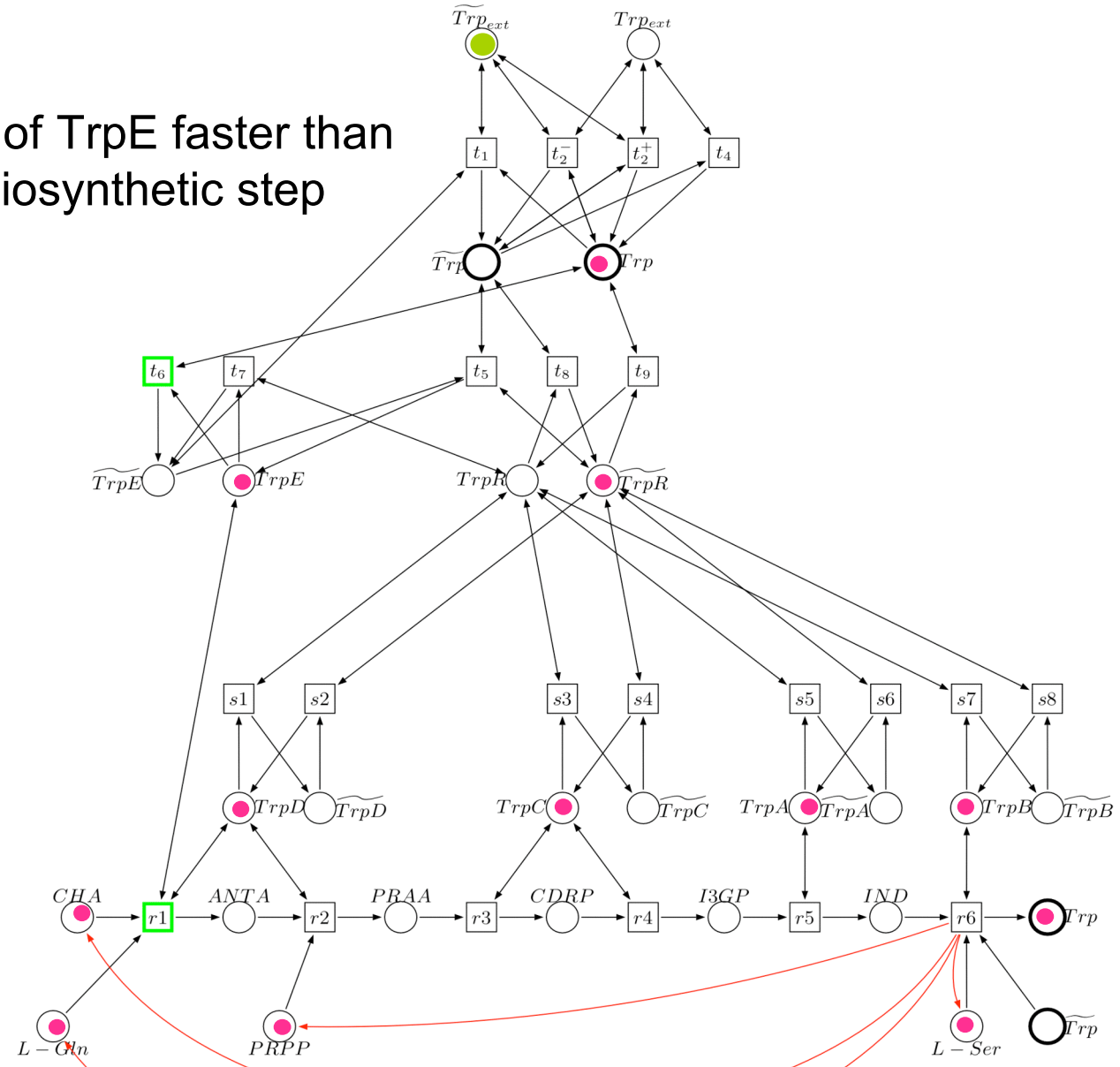


inhibitory effect of the holorepressor onto trpEDCBA

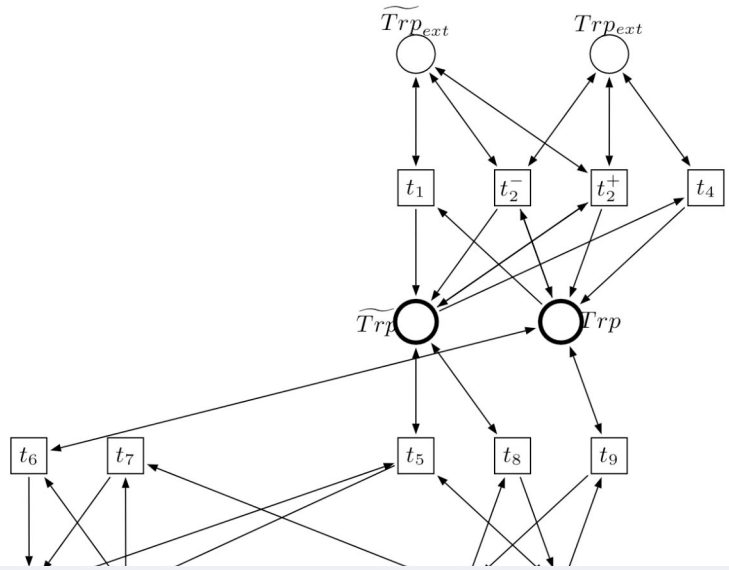



$\widetilde{Trp}_{ext}$
Trp_{ext}
t_1
t_2^-
t_2^+
t_4
$\widetilde{Trp}$
Trp
t_6
t_7
t_5
t_8
t_9
$\widetilde{TrpE}$
$TrpE$
$TrpR$
$\widetilde{TrpR}$
$s1$
$s2$
$s3$
$s4$
$s5$
$s6$
$s7$
$s8$
$TrpD$
$\widetilde{TrpD}$
$TrpC$
$\widetilde{TrpC}$
$TrpA$
$\widetilde{TrpA}$
$TrpB$
$\widetilde{TrpB}$
CHA
$r1$
$ANTA$
$r2$
$PRAA$
$r3$
$CDRP$
$r4$
$I3GP$
$r5$
IND
$r6$
Trp
$L-Gln$
Pyr
$L-Glu$
$PRPP$
PPi
CO_2
H_2O
CO_2
$L-Ser$
H_2O
$\widetilde{Trp}$

inhibition of TrpE faster than
the 1rst biosynthetic step
-> t6

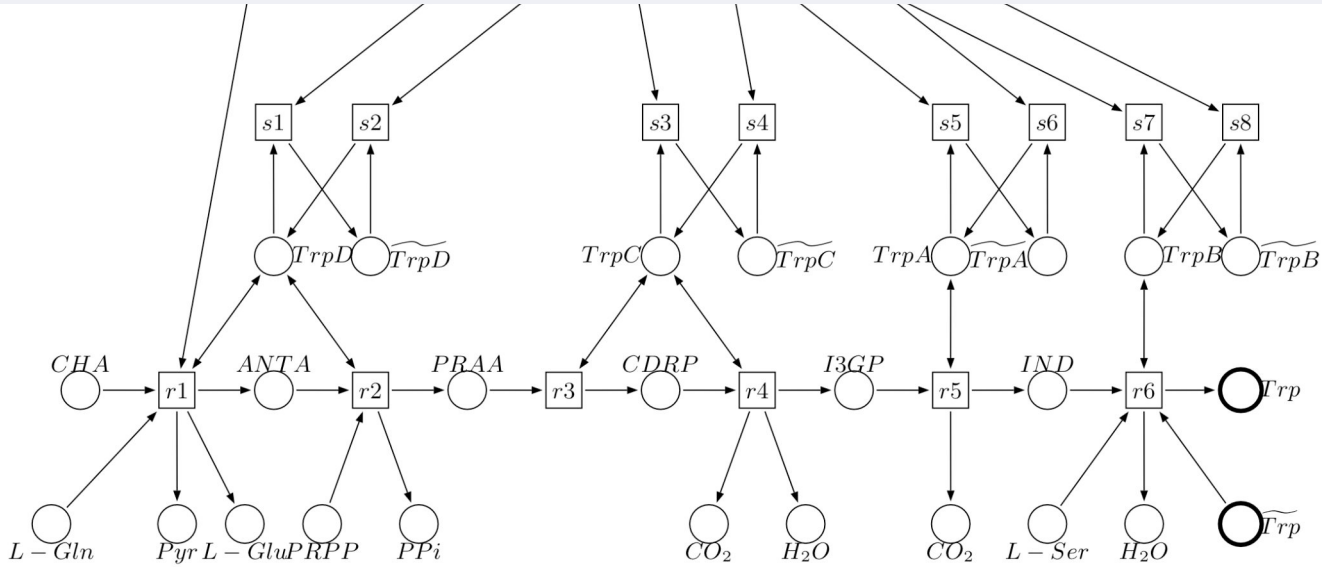


input compounds are not a restrictive resource



intermediate and full models (INA files) available at

http://gin.univ-mrs.fr



Integrated PN modelling - Analysis

Initial state: **no external tryptophan**, all input compounds present, all enzymes active, no internal tryptophan, no holorepressor

One cyclic attractor denoting **homeostatic levels** of **internal tryptophan** and **TrpE activity**

	CHA	ANTA	PRAA	CRDP	I3GP	IND	TrpE	TrpD	TrpC	TrpB	TrpA	Trp	TrpR	Trp_{ext}
	1	0	0	0	0	0	1	1	1	1	1	0	0	0
r1	**0**	**1**	0	0	0	0	1	1	1	1	1	0	0	0
r2	0	**0**	**1**	0	0	0	1	1	1	1	1	0	0	0
r3	0	0	**0**	**1**	0	0	1	1	1	1	1	0	0	0
r4	0	0	0	**0**	**1**	0	1	1	1	1	1	0	0	0
r5	0	0	0	0	**0**	**1**	1	1	1	1	1	0	0	0
r6	**1**	0	0	0	0	0	1	1	1	1	1	**1**	0	0
t6	1	0	0	0	0	0	**0**	1	1	1	1	1	0	0
t1	1	0	0	0	0	0	0	1	1	1	1	**0**	0	0

t5

Integrated PN modelling - Analysis

low external tryptophan, all input compounds present, all enzymes active, no internal tryptophan, no holorepressor

a **unique reachable dead marking** with a **moderate level** of internal **tryptophan** ; **repressor** and **TrpE inactive**

CHA	ANTA	PRAA	CRDP	I3GP	IND	TrpE	TrpD	TrpC	TrpB	TrpA	Trp	TrpR	Trp_{ext}
1	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**1**
---	---	---	---	---	---	---	---	---	---	---	---	---	---
1	0	0	0	0	0	**0**	1	1	1	1	**1**	0	1

Integrated PN modelling - Analysis

high external tryptophan, all input compounds present, all enzymes active, no internal tryptophan, no holorepressor

six reachable dead markings with a **high level** of **internal tryptophan**, the **six enzymes inactive**, the **repressor active**

CHA	ANTA	PRAA	CRDP	I3GP	IND	TrpE	TrpD	TrpC	TrpB	TrpA	Trp	TrpR	Trp_{ext}
1	0	0	0	0	0	1	1	1	1	1	0	0	2
---	---	---	---	---	---	---	---	---	---	---	---	---	---
1	0	0	0	0	0	0	0	0	0	0	2	1	2
0	1	0	0	0	0	0	0	0	0	0	2	1	2
0	0	1	0	0	0	0	0	0	0	0	2	1	2
0	0	0	1	0	0	0	0	0	0	0	2	1	2
0	0	0	0	1	0	0	0	0	0	0	2	1	2
0	0	0	0	0	1	0	0	0	0	0	2	1	2

Modelling of biological networks, Coloured PNs

- Tokens are distinguishable by means of "colour" sets
- Guards are associated to transitions
- Functions associated to arcs

→ compacted models (a CPN can be unfolded to a low level PN)

CPN modelling of metabolic pathways

Discrimination of alternative metabolic paths

K. Voss, M. Heiner, and I. Koch (2003) *In Silico Biol*, 3(3):367–87.

CPN modelling of regulatory networks (*logical* formalism)

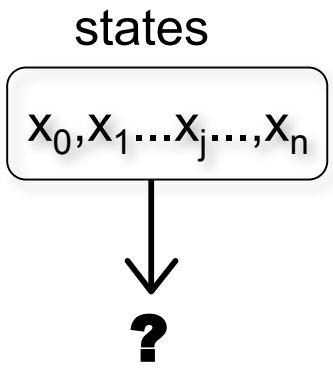
verify the coherence of the system under various hypotheses.

J.-P. Comet, H. Klaudel, S. Liauzu (2005) ICATPN - LNCS 3536, 208-227.

CPN modelling of *logical regulatory* networks

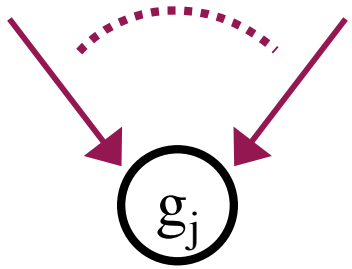
C.Chaouiya, E. Remy, D.Thieffry (2006) Qualitative Petri net modelling of genetic networks. TCSB VI, 95-112.

Recall that in the logical formalism, the evolution of the system is directed by logical parameters: in a given state, and for a given gene g_j select the relevant parameter



$\mathcal{I}_j(x)$

set of interactions operating on g_j in state x



if $K_j(\mathcal{I}_j(x)) \neq x_j$, gene g_j receives a **call for updating**

- if $K_j(\mathcal{I}_j(x)) > x_j$ then g_j is called to increase
- if $K_j(\mathcal{I}_j(x)) < x_j$ then g_j is called to decrease

CPN modelling of *logical regulatory* networks

Given a regulatory graph $R = (G, I, K)$, *an initial state* $\mathbf{x}_0$, *the corresponding Coloured Regulatory Petri Net*, $C(R) = (\Sigma, P, T, A, C, G, E, x_0)$, is defined *by:*

- $P = \{g_1, \ldots, g_n\}$ *the set of places,* $T = \{T_1, \ldots, T_n\}$ *the set of transitions.*
- Σ *the finite set of* ***colour sets****:* $\Sigma = \{ [0, \mathrm{Max}_i], i = 1, \ldots n\}$.
- C *the color function:* $C : P \rightarrow \Sigma,\ C(gi) = [0, \mathrm{Max}_i]$.
- $A \subseteq (P \times T \cup T \times P)$ *the set of arcs with*

  $\forall T_i \in T, \forall g_j \in Reg(i), (g_j, T_i) \in A$ *and* $(T_i, g_j) \in A, (g_i, T_i) \in A, (T_i, g_i) \in A$.
- E *the arc expression function defined as follows:* $\forall Ti \in T$,

  $\forall g_j \in {}^\bullet T_i \setminus \{g_i\}, \quad E(g_j, T_i) = E(T_i, g_j) = x_j,\ x_j \in C(g_j),$ (with ${}^\bullet T_i = Reg(i) \cup \{g_i\}$)

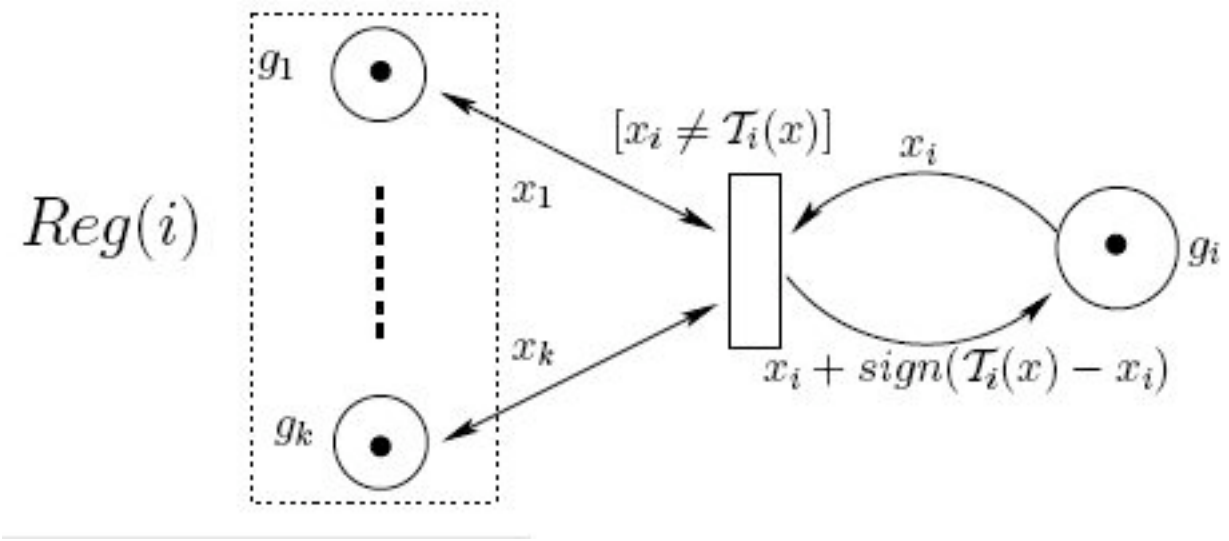
  $E(g_i, T_i) = x_i,\ x_i \in C(g_i),$

  $E(T_i, g_i) = x_i + sign(T_i(\mathbf{x}) - x_i)\ ,\ \mathbf{x} \in \prod_{gk \in P} C(g_k)$.
- $G = \{G_1, \ldots, G_n\}$ *is the set of guards;*

  for all transition T_i *a Boolean function,* G_i *is defined as follows:*

  $\forall \mathbf{x} \in \prod_{gk \in P} C(g_k)\ ,\ G_i(\mathbf{x}) = [\ T_i(\mathbf{x}) = x_i\ ]$.
- *The initial marking* $\mathbf{x}_0$ = *assigns to each place* g_i *one token with the required value in* $C(g_i)$

CPN modelling of *logical regulatory* networks



Efficient representation of $\mathcal{T}_i(x)$

Modelling of biological networks, Stochastic PNs

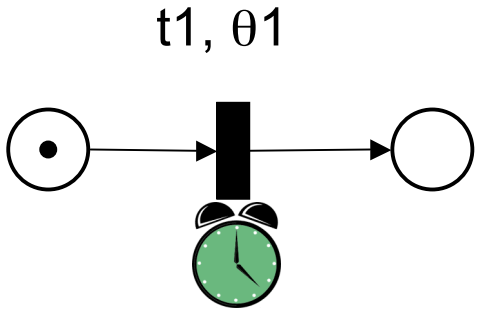
SPN modelling of stochastic molecular interactions (Gillespie's algorithm)

R. Srivastava, MS Peterson and WE Bentley (2001) Biotechnol Bioeng. Oct 5;75(1):120-9.

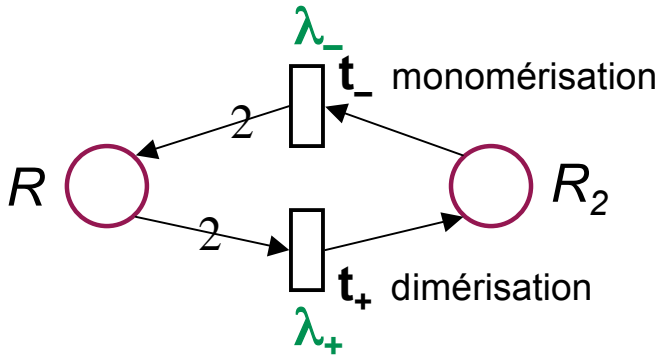
- Uncertainty attached to the data
- Environmental noise

→ Intrinsic noise (*i.e.* low molecular concentrations)

Stochastic time-delay associated to each transition (exponential distribution, may depend on the marking)



Example : $2R \leftrightarrow R_2$



reaction involving a unique reactant: $\lambda_- = k_-\ M(R_2)$

constant monomerisation rate

reaction involving two reactants: $\lambda_+ = c_+ M(R)(M(R)-1)$

constant dimerisation rate, $c_+ = k_+/V.N_A$

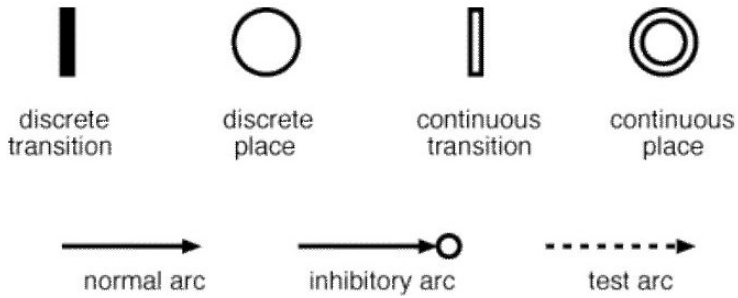
Modelling of biological networks, Hybrid PNs

HPN modelling of gene regulated metabolic networks

M.Chen and R.Hofestädt (2003), In Silico Biology 3, 0029

molecular concentration = continuous rather than discrete

discrete places (with tokens)
discrete transitions (with delays)
continuous places (with marks $\in IR^+$)
continuous transitions (with speeds

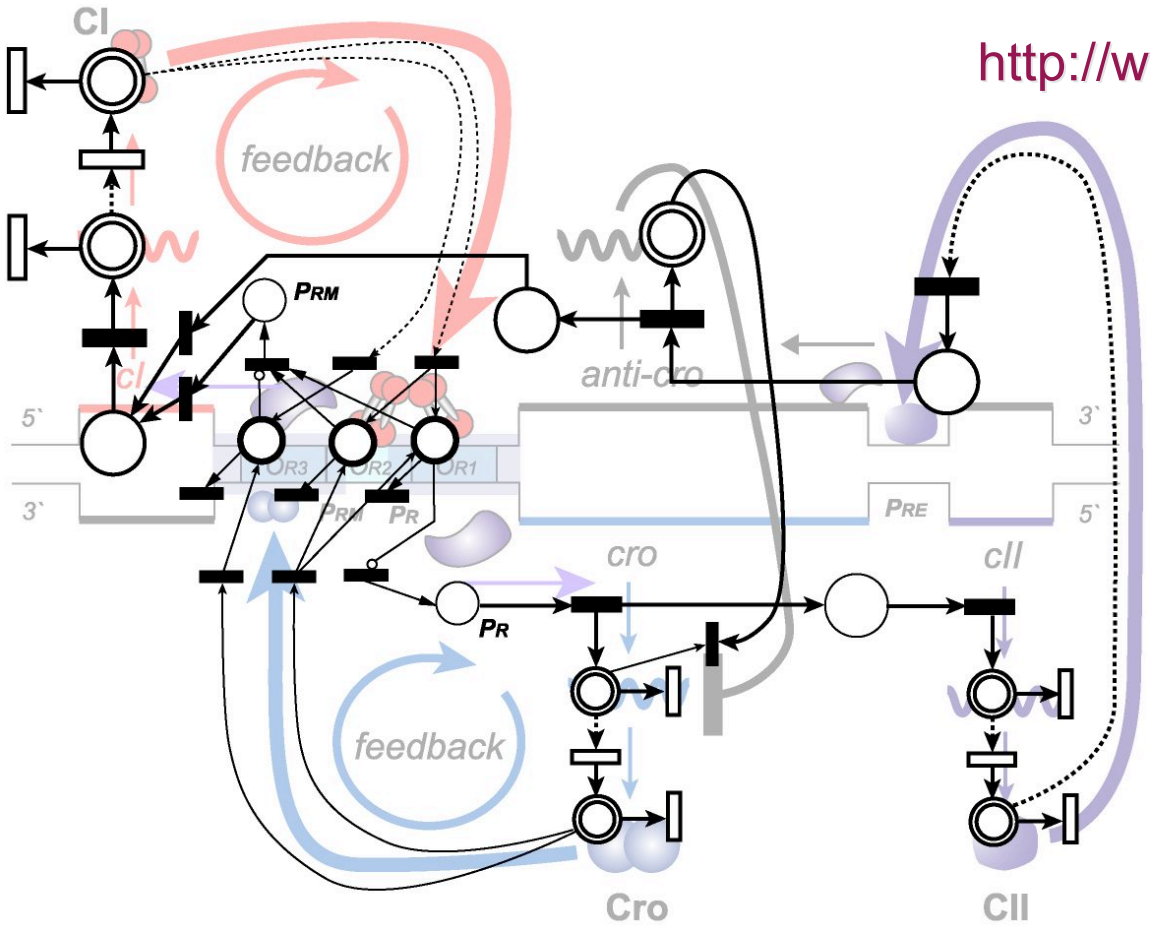


Modelling of biological networks, Hybrid PNs

HPN modelling of gene regulated metabolic networks

Lambda phage genetic switch feedback mechanism

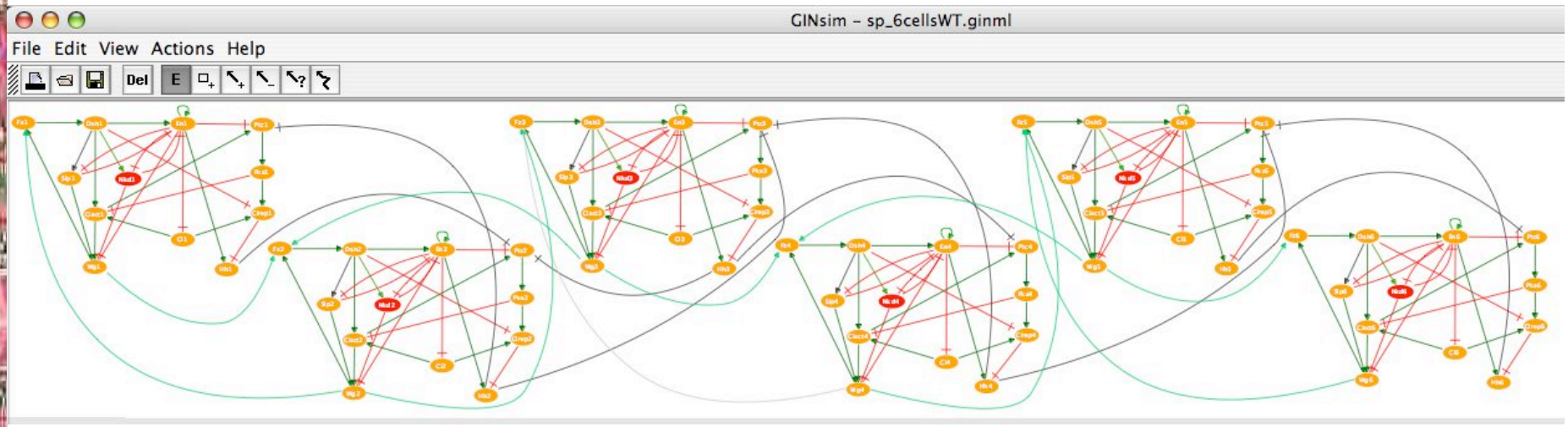
A. Doi, H. Matsuno, S. Miyano (2000) Currents in Computational Molecular Biology, 26-27.



http://www.genomicobject.net

Conclusions

- Segment polarity *logical* model (72 components)



Full state transition graph (considering all possible initial conditions) has **3 018 225 388 942 786 560** nodes

WT reachability analysis: two main multicellular outcomes

using the PN translation, generation of the marking graph

250 000 states (partial reduced marking graph, **stubborn reduction**)

→ 2 dead markings

Use of priorities + depth limitation

Conclusions

- Wide variety of PN based modelling for biological networks

  graphical representation, suitability to represent concurrency, well founded mathematical theory, available tools for analysis / simulation

  pure qualitative to sophisticated hybrid models

  structural analysis to pure simulation

  model checking

- Step by step modelling:

  integration of different levels of abstraction through the different PN extensions

  facing the problem of composition: defining PN buiding blocks

References

✓T. Murata (1989) Petri Nets: Properties, Analysis and Applications. Proceedings of the IEEE, Vol.77, No.4.

✓S. Hardy and P. Robillard (2004) Modeling and simulation of molecular biology using Petri nets: modeling goals of various approaches. J. of Bioinformatics and Computational Biology, 2(4):619-37.

✓I. Koch, S. Schuster and M. Heiner (2003) Simulation and analysis of metabolic networks by time dependent Petri nets. In Silico Biology 3-31.

✓C. Chaouiya and E. Remy and D. Thieffry (2006) Qualitative Petri Net Modelling of Genetic Networks. TCSB VI, 4220:95-112.

✓ H. Matsuno *et al.* (2003) Biopathways representation and simulation on hybrid functional Petri net. In Silico Biol. 3-32.

✓Srivastava R, Peterson MS, Bentley WE (2001) Stochastic kinetic analysis of the Escherichia coli stress circuit using sigma(32)-targeted antisense. Biotechnol Bioeng. Oct 5;75(1):120-9.