Metabolic Coupling in Gene Regulatory Networks in Bacteria

Hidde de Jong

INRIA Grenoble - Rhône-Alpes
Hidde.de-Jong@inria.fr http://ibis.inrialpes.fr

UNIVERSITE JOSEPH FOURIER SCIENCES TECHNOLOGIE MEDECINE

cnrs

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

INRIA

centre de recherche GRENOBLE - RHÔNE-ALPES

Overview

1. Gene regulatory networks and metabolic coupling
2. Derivation of interactions induced by metabolic coupling
3. Analysis of network controlling genes involved in carbon assimilation in *E. coli*
4. Metabolic coupling and network dynamics
5. Conclusions

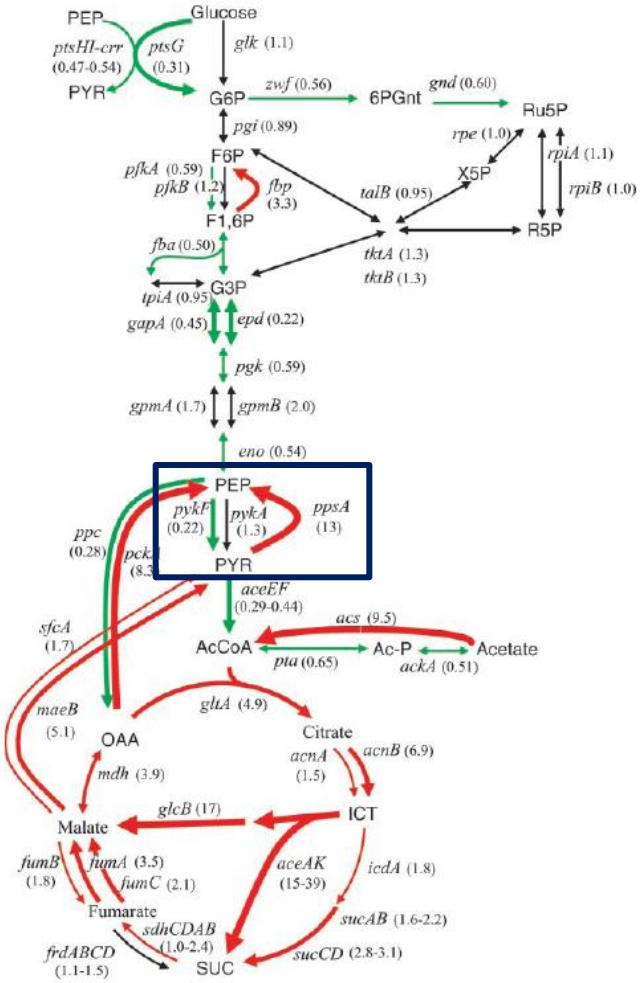
Gene regulatory networks

- The adaptation of bacteria to changes in their environment involves adjustment of gene expression levels

  Differences in expression of enzymes in central metabolism of *E. coli* during growth on glucose or acetate

  Oh *et al.* (2002), *J. Biol. Chem.*, 277(15):13175–83

- **Gene regulatory networks** control changes in expression levels in response to environmental perturbations



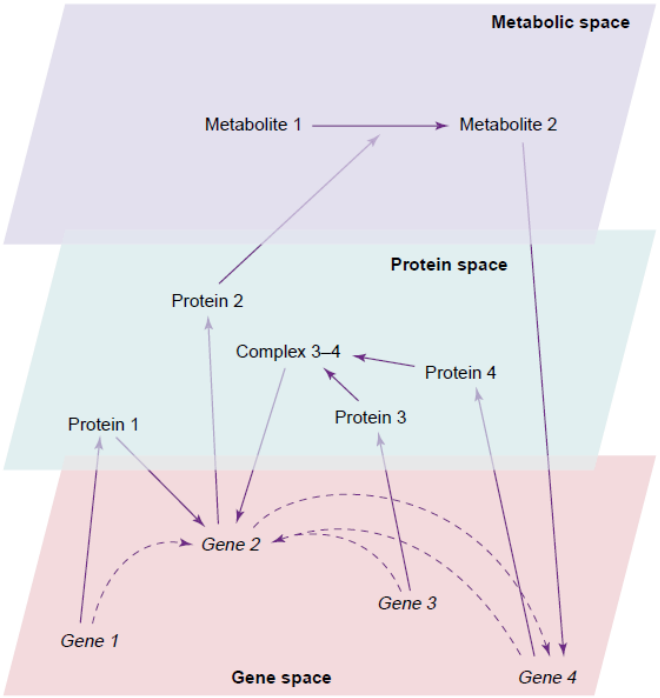
Gene regulatory networks

- **Gene regulatory networks** consist of genes, gene products (RNAs, proteins), and the regulatory effect of the latter on the expression of other genes

  Bolouri (2008), *Computational Modeling of Gene Regulatory Networks*, Imperial College Press

- Gene regulatory networks cannot be reduced to **direct** interactions (transcription regulation), but also include **indirect** interactions (mediated by metabolism)

  Brazhnik *et al.* (2002), *Trends Biotechnol.*, 20(11):467-72



Problem statement

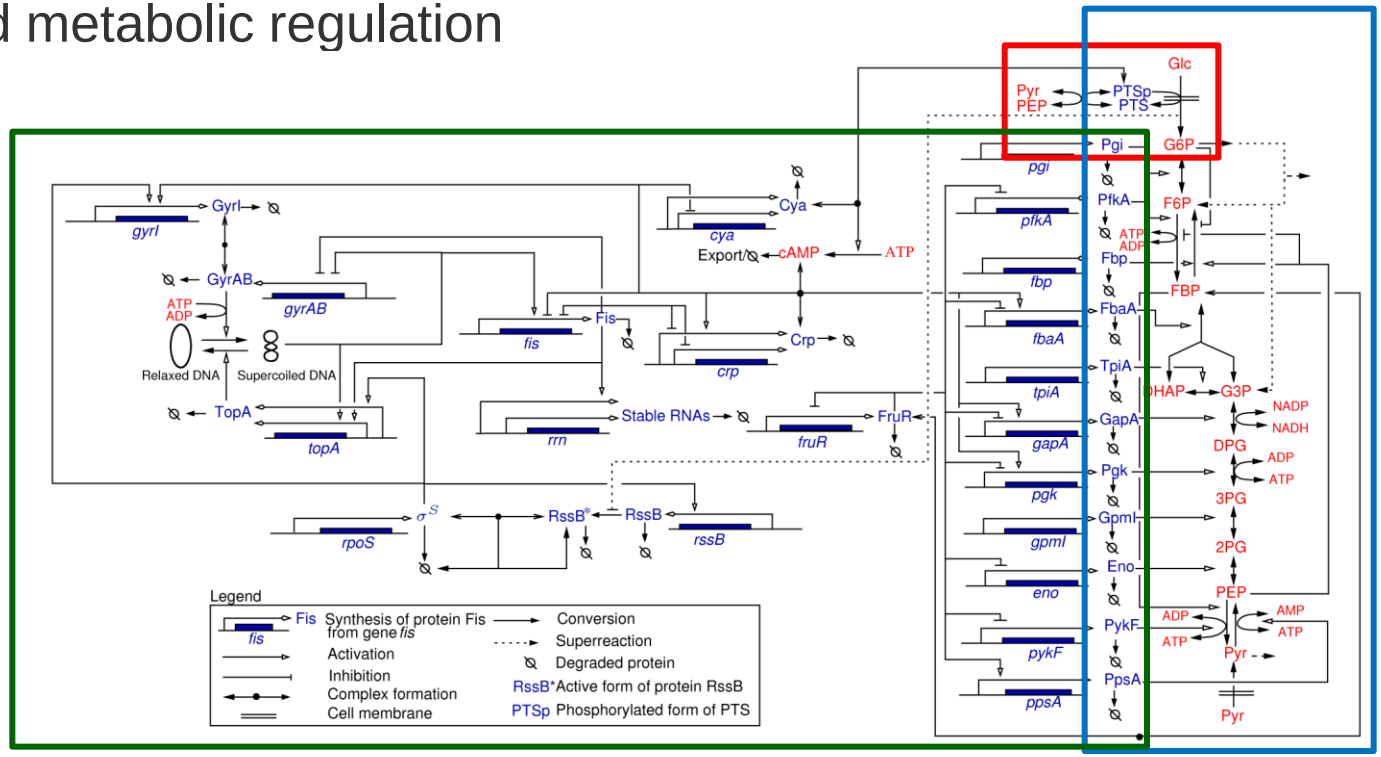
- Occurrence of indirect regulatory interactions between enzymes and genes: **metabolic coupling**
- By which method can we **analyze metabolic coupling** in gene regulatory networks in a principled way?

  How can we derive indirect interactions from underlying system of biochemical reactions?
- Practical constraints
  - Large systems (many species, many reactions)
  - Lack of information on specific reaction mechanisms
  - Lack of parameter values, lack of data to estimate parameter values

Problem statement

- Which **new insights** does this method give us into the functioning of the carbon assimilation network in *E. coli*?

  Upper part of glycolysis and gluconeogenesis pathways and their genetic and metabolic regulation



Outline of approach

- By which method can we **analyze metabolic coupling** in gene regulatory networks in a principled way?

  How can we derive indirect interactions from underlying system of biochemical reactions?

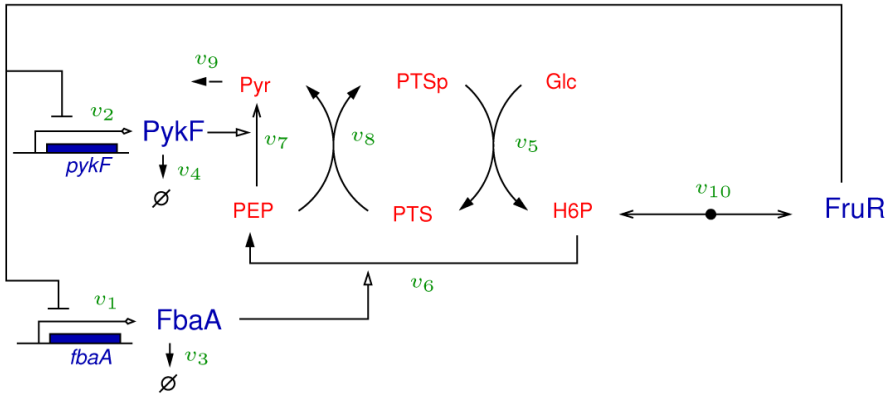
- Approach based on reduction of stoichiometric model of system of biochemical reactions, making following weak assumptions:
  - Distinct time-scale hierarchies between metabolism and gene expression: model reduction using **quasi-steady-state approximation**
  - Stability of fast subsystem: use of **control coefficients** from metabolic control theory

Baldazzi *et al.* (2010), *PLoS Comput. Biol.*, 6(6):e1000812

Kinetic models and time-scale hierarchy

- Kinetic model of form $\dot{x} = N\, v(x)$
  - Concentration variables $x \in \mathbb{R}^n_+$
  - Reaction rates $v : \mathbb{R}^n_+ \rightarrow \mathbb{R}^q$
  - Stoichiometry matrix $N \in \mathbb{Z}^{n \times q}$

Heinrich and Schuster (1996),
The Regulation of Cellular Systems, Chapman & Hall



Simplified model of glycolysis pathway, with metabolic and genetic regulation

$$\begin{aligned}\dot{x}_{PEP} = \; & 2 \cdot v_6(x_{H6P}, x_{PEP}, x_{FbaA}) \\ & -1 \cdot v_7(x_{Pyr}, x_{PEP}, x_{PykF}) \\ & -1 \cdot v_8(x_{PEP}, x_{Pyr}, x_{PTSp})\end{aligned}$$

Kinetic models and time-scale hierarchy

- Kinetic model of form $\dot{x} = N\,v(x)$
  - Concentration variables $x \in \mathbb{R}^n_+$
  - Reaction rates $v : \mathbb{R}^n_+ \to \mathbb{R}^q$
  - Stoichiometry matrix $N \in \mathbb{Z}^{n \times q}$

- Time-scale hierarchy motivates distinction between **fast** reaction rates $v^f \in \mathbb{R}^{q-p}$ and **slow** reaction rates $v^s \in \mathbb{R}^p$ such that

$$v = [v^s \; v^f]'$$

Typically, enzymatic and complex formation reactions are fast, protein synthesis and degradation are slow

Kinetic models and time-scale hierarchy

- Separation of fast and slow reactions motivates a linear transformation $T \in \mathbb{Z}^n \times \mathbb{Z}^n$ of the variables

$$\begin{bmatrix} x^s \\ x^f \end{bmatrix} = T\,x \quad \text{such that} \quad \begin{bmatrix} N^s & 0 \\ N^{s'} & N^f \end{bmatrix} = T\,N$$

- We call $x^s \in \mathbb{R}_+^m$ **slow variables** and $x^f \in \mathbb{R}_+^{n-m}$ **fast variables**

  Slow variables are typically **total protein concentrations**, fast variables **metabolites and biochemical complexes**

Kinetic models and time-scale hierarchy

- Separation of fast and slow reactions motivates a linear transformation $T \in \mathbb{Z}^n \times \mathbb{Z}^n$ of the variables

$$\begin{bmatrix} x^s \\ x^f \end{bmatrix} = T\,x \quad \text{such that} \quad \begin{bmatrix} N^s & 0 \\ N^{s'} & N^f \end{bmatrix} = T\,N$$

- We call $x^s \in \mathbb{R}_+^m$ **slow variables** and $x^f \in \mathbb{R}_+^{n-m}$ **fast variables**
- Separation of fast and slow variables allows $\dot{x} = N\,v(x)$ to be rewritten as coupled slow and fast subsystems

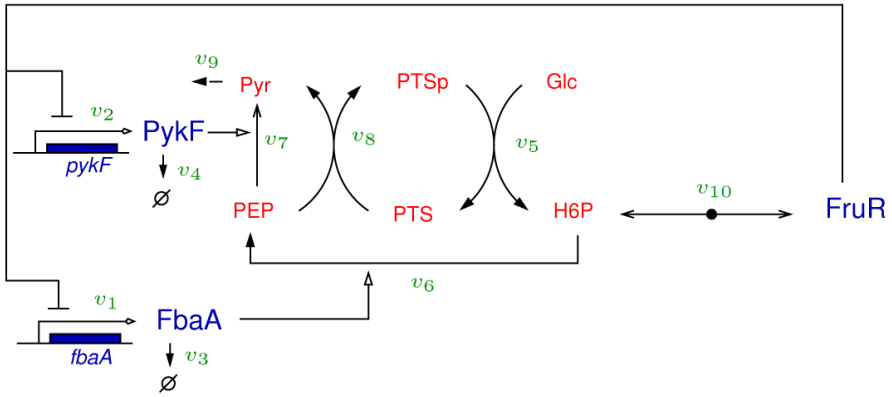
$$\dot{x}^s = N^s\,v^s(x^s, x^f)$$
$$\dot{x}^f = N^{s'}\,v^s(x^s, x^f) + N^f\,v^f(x^s, x^f) \approx N^f\,v^f(x^s, x^f)$$

Kinetic models and time-scale hierarchy

- Reduction of simplified kinetic model of glycolysis using time-scale separation

$$\dot{x}^s = N^s\, v^s(x^s, x^f)$$

$$\dot{x}^f = N^{s'}\, v^s(x^s, x^f) + N^f\, v^f(x^s, x^f) \approx N^f\, v^f(x^s, x^f)$$



$$\begin{bmatrix} \dot{x}_{FbaA} \\ \dot{x}_{PykF} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1(x_{FruR\cdot free}) \\ v_2(x_{FruR\cdot free}) \\ v_3(x_{FbaA}) \\ v_4(x_{PykF}) \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_{H6P} \\ \dot{x}_{PEP} \\ \dot{x}_{Pyr} \\ \dot{x}_{PTSp} \\ \dot{x}_{FruR\cdot free} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 2 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_5(x_{Glc}, x_{PTSp}) \\ v_6(x_{H6P}, x_{PEP}, x_{FbaA}) \\ v_7(x_{Pyr}, x_{PEP}, x_{PykF}) \\ v_8(x_{PEP}, x_{Pyr}, x_{PTSp}) \\ v_9(x_{Pyr}) \\ v_{10}(x_{H6P}, x_{FruR\cdot free}) \end{bmatrix}$$

Model reduction using time-scale hierarchy

- Separation of fast and slow variables allows original model to be rewritten as coupled slow and fast subsystems

$$\dot{x}^s = N^s \, v^s(x^s, x^f)$$

$$\dot{x}^f = N^{s'} \, v^s(x^s, x^f) + N^f \, v^f(x^s, x^f) \approx N^f \, v^f(x^s, x^f)$$

- Under **quasi-steady-state approximation (QSSA)**, fast variables are assumed to instantly adapt to slow dynamics

$$\dot{x}^f = 0 \;\Rightarrow\; N^f \, v^f(x^s, x^f) = 0$$

Mathematical basis for QSSA is given by Tikhonov's theorem

Heinrich and Schuster (1996), *The Regulation of Cellular Systems*, Chapman & Hall

Khalil (2001), *Nonlinear Systems*, Prentice Hall, 3rd ed.

Model reduction using time-scale hierarchy

- QSSA implicitly relates steady-state value of fast variables to slow variables

$$x^f = g(x^s),\ g\ :\ \mathbb{R}_+^m \rightarrow \mathbb{R}_+^{n-m}$$

- This gives **reduced model on the slow time-scale**

$$\dot{x}^s = N^s\, v^s(x^s, g(x^s))$$

Reduced model describes direct and indirect interactions between slow variables (total protein concentrations)

Mathematical representation of effective **gene regulatory network**

- But
  - Generally function g is not easy to obtain due to nonlinearities
  - Function g depends on unknown parameter values

Jacobian matrix and regulatory structure

- **Derivation of interaction structure** between slow variables by computation of **Jacobian matrix**

$$\mathcal{J} = \frac{\partial \dot{x}^s}{\partial x^s} = \underbrace{N^s \frac{\partial v^s(x^s, x^f)}{\partial x^s}}_{\text{Direct regulation by transcription factors}} + \underbrace{N^s \frac{\partial v^s(x^s, x^f)}{\partial x^f} \frac{\partial g(x^s)}{\partial x^s}}_{\text{Indirect regulation through metabolic coupling}}$$

- Implicit differentiation of $N^f \, v^f(x^s, x^f) = 0$ yields

$$\frac{\partial g(x^s)}{\partial x^s} = -M^{-1} \, N^f \, \frac{\partial v^f(x^s, x^f)}{\partial x^s}$$

where $M = N^f \, \partial v^f(x^f, \bar{x^s}) / \partial x^f$ is Jacobian matrix of fast system

Jacobian matrix and regulatory structure

- Relation between obtained expression for Jacobian matrix and **Metabolic Control Analysis (MCA)**

$$\frac{\partial g(x^s)}{\partial x^s} = \underbrace{-M^{-1}\, N^f}_{\text{Concentration control coefficients}} \frac{\partial v^f(x^s, x^f)}{\partial x^s}$$

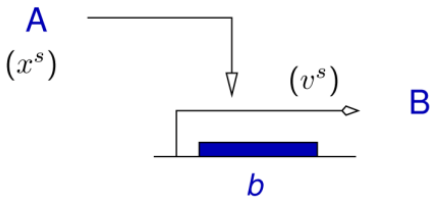
- **Concentration control coefficients** characterize the steady-state response of metabolic subsystem to changes in slow variables (enzyme concentrations)
- Concentration control coefficients are expressed in terms of **elasticity coefficients**, which quantify the changes in reaction rates to perturbations in slow variables

Heinrich and Schuster (1996), *The Regulation of Cellular Systems*, Chapman & Hall

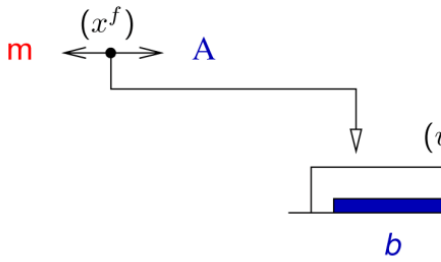
Determination of interaction signs

- Can we derive **signs for regulatory interactions** (elements of Jacobian matrix), without knowledge on rate laws and parameter values?
- Idea: exploit link with MCA, notably that **signs of elasticities** are known

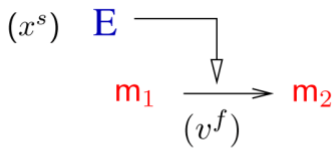
  Rate laws are generally monotone functions in variables



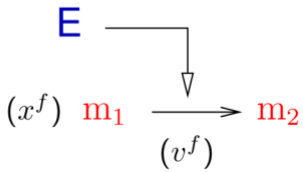
$$\frac{\partial v^s}{\partial x^s} > 0$$



$$\frac{\partial v^s}{\partial x^f} > 0$$



$$\frac{\partial v^f}{\partial x^s} > 0$$



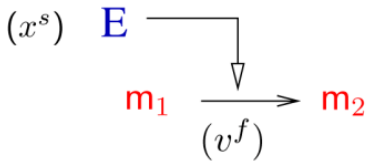
$$\frac{\partial v^f}{\partial x^f} > 0$$

Determination of interaction signs

- Can we derive **signs for regulatory interactions** (elements of Jacobian matrix), without knowledge on rate laws and parameter values?
- Idea: exploit link with MCA, notably that **signs of elasticities** are known

  Rate laws are generally monotone functions in variables

- But
  - Reversible reactions: signs of $\partial v^f(x^s, x^f)/\partial x^s$ change with flux direction



$$\frac{\partial v^f}{\partial x^s} > 0$$

  - Therefore, derive signs of regulatory interaction for given flux directions

Determination of interaction signs

- Resolution of signs of (large) algebraic expressions defining interaction signs by means of computer algebra tools

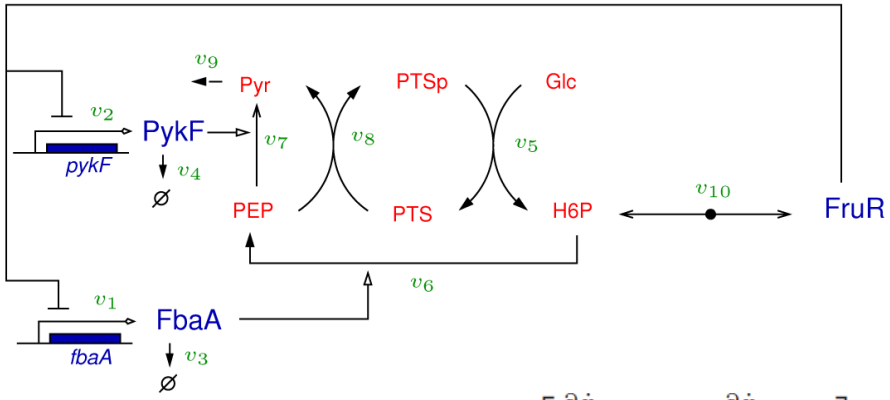
$$\mathcal{J} = \frac{\partial \dot{x}^s}{\partial x^s} = N^s \frac{\partial v^s(x^s, x^f)}{\partial x^s} + N^s \frac{\partial v^s(x^s, x^f)}{\partial x^f} \frac{\partial g(x^s)}{\partial x^s}$$

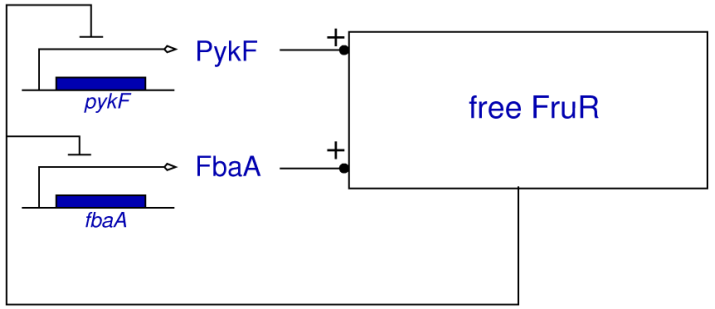
Symbolic Math Toolbox in Matlab

- Use of additional constraints in sign resolution
  - **Stability assumption for fast system**: necessary condition for stability is that coefficients of characteristic polynomial $det(M - \lambda I) = 0$ have same sign
  - **Experimental determination** of some of the signs of concentration control coefficients in $\frac{\partial g(x^s)}{\partial x^s}$ (if available)

Determination of interaction signs

- Derivation of interaction signs from simplified kinetic model of glycolysis
  - Enzymes influence expression of metabolic genes through metabolism (metabolic coupling)
  - Intuitive explation of metabolic coupling in this simple example



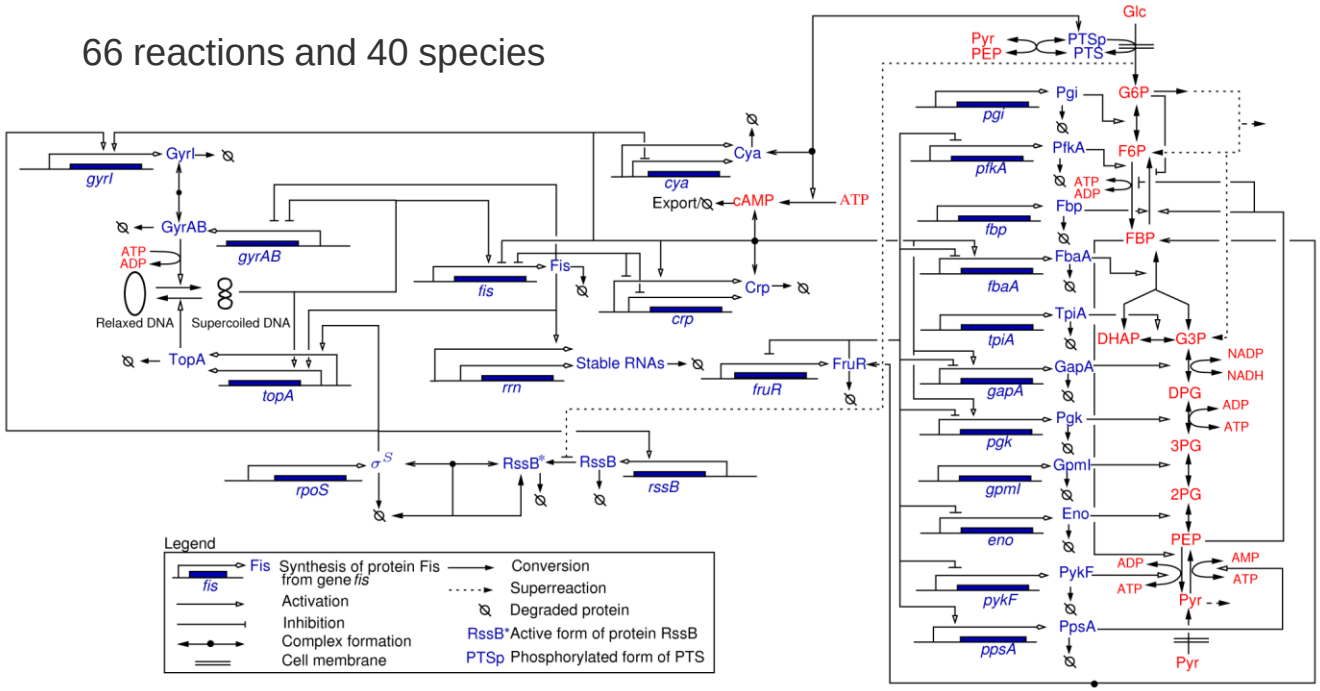


$$\mathcal{J} = \begin{bmatrix} \frac{\partial \dot{x}_{FbaA}}{\partial x_{FbaA}} & \frac{\partial \dot{x}_{FbaA}}{\partial x_{PykF}} \\ \frac{\partial \dot{x}_{PykF}}{\partial x_{FbaA}} & \frac{\partial \dot{x}_{PykF}}{\partial x_{PykF}} \end{bmatrix} \text{ and } \operatorname{sign}(\mathcal{J}) = \begin{bmatrix} - & - \\ - & - \end{bmatrix}.$$

Application to *E. coli* carbon assimilation

- Development of model of carbon assimilation network, analysis under following conditions:

  Glycolysis/gluconeogenesis (growth on glucose/pyruvate)

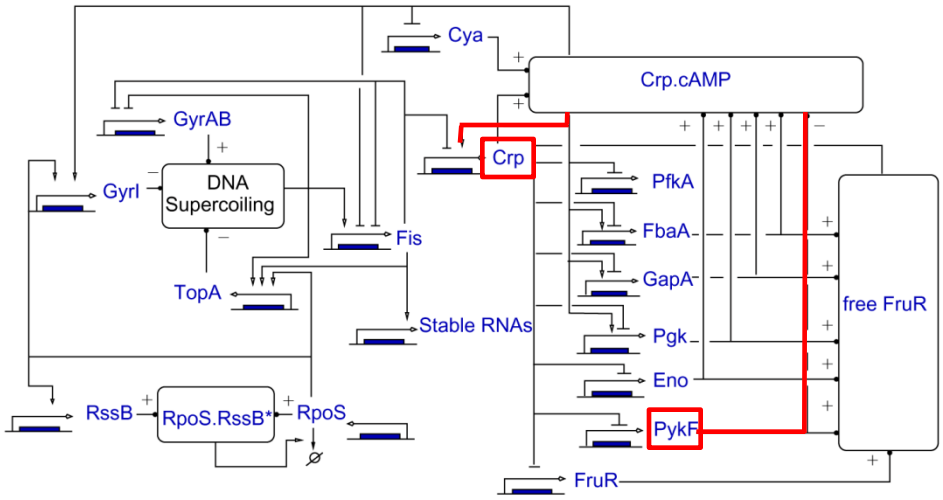


Application to *E. coli* carbon assimilation

- Development of model of carbon assimilation network, analysis under following conditions:

  Glycolysis/gluconeogenesis (growth on glucose/pyruvate)

$$\mathcal{J} = \frac{\partial \dot{x}^s}{\partial x^s}$$



Regulators

	PfkA	FbaA	GapA	Pgk	Eno	PykF	Cya	Crp	Fis	GyrAB	Gyrl	TopA	RpoS	RssB	stable RNAs	FruR
pfkA	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-
fbaA	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
gapA	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
pgk	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
eno	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-
pykF	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-
cya	0	-	-	-	-	+	-	-	0	0	0	0	0	0	0	0
crp	0	+	+	+	+	-	+	+	-	0	0	0	0	0	0	0
fis	0	0	0	0	0	0	-	-	-	+	-	-	0	0	0	0
gyrAB	0	0	0	0	0	0	0	0	-	-	+	+	0	0	0	0
gyrI	0	0	0	0	0	0	+	+	0	0	0	0	+	0	0	0
topA	0	0	0	0	0	0	0	0	+	+	-	-	+	0	0	0
rpoS	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0
rssB	0	0	0	0	0	0	0	0	0	0	0	0	+	0	0	0
rrn	0	0	0	0	0	0	0	0	+	0	0	0	0	0	0	0
fruR	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-

Glycolysis with allosteric effects

- Few fast variables couple metabolism to gene expression

Network is densely connected

- Contrary to what is often maintained, gene regulatory network is found to be **densely connected**
- Strong connectivity arises from metabolic coupling
  - $\mathcal{M}^0$: transcriptional network consisting of direct interactions only
  - $\mathcal{M}^2_{glyco}$: gene regulatory network in glycolytic growth conditions including direct and indirect interactions

	$\mathcal{M}^0$	$\mathcal{M}^1_{glyco}$	$\mathcal{M}^2_{glyco}$	$\mathcal{M}^1_{neo}$	$\mathcal{M}^2_{neo}$
Number of feedback loops	4	2388	9246	24	2257
Maximal loop length	2	12	12	6	12
Average connectivity	1.4	4.7	5.2	2.8	4.4

- Experimental evidence for indirect interactions in perturbation experiments (deletion mutants, enzyme overexpression)

Siddiquee *et al.* (2004), *FEMS Microbiol. Lett.,* 235:25–33
Baptist *et al.*, submitted

Network is largely sign-determined

- Derived gene regulatory network for carbon assimilation in *E. coli* is largely **sign-determined**

  Signs of interactions do not depend on explicit specification of kinetic rate laws or parameter values, but are structural property of system

Glycolysis with allosteric effects

		Regulators															
		PfkA	FbaA	GapA	Pgk	Eno	PykF	Cya	Crp	Fis	GyrAB	GyrI	TopA	RpoS	RssB	stable RNAs	FruR
Genes	*pfkA*	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-
	fbaA	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
	gapA	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
	pgk	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
	eno	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-
	pykF	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-
	cya	0	-	-	-	-	+	-	-	0	0	0	0	0	0	0	0
	crp	0	+	+	+	+	-	+	+	-	0	0	0	0	0	0	0
	fis	0	0	0	0	0	0	-	-	-	+	-	-	0	0	0	0
	gyrAB	0	0	0	0	0	0	0	0	-	-	+	+	0	0	0	0
	gyrI	0	0	0	0	0	0	+	+	0	0	0	0	+	0	0	0
	topA	0	0	0	0	0	0	0	0	+	+	-	-	+	0	0	0
	rpoS	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0
	rssB	0	0	0	0	0	0	0	0	0	0	0	0	+	0	0	0
	rrn	0	0	0	0	0	0	0	0	+	0	0	0	0	0	0	0
	fruR	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-

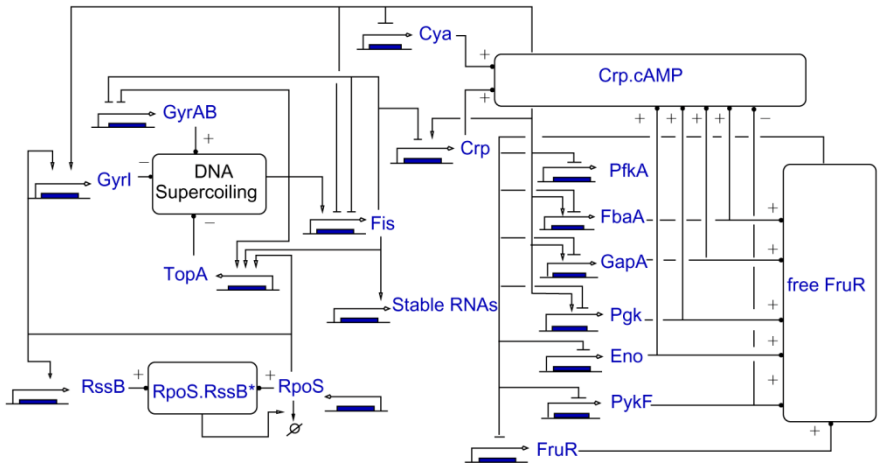
- Sign-determinedness not expected on basis of work in ecology

  Sufficient conditions for sign-determinedness can be formulated using expression for $\mathcal{J}$

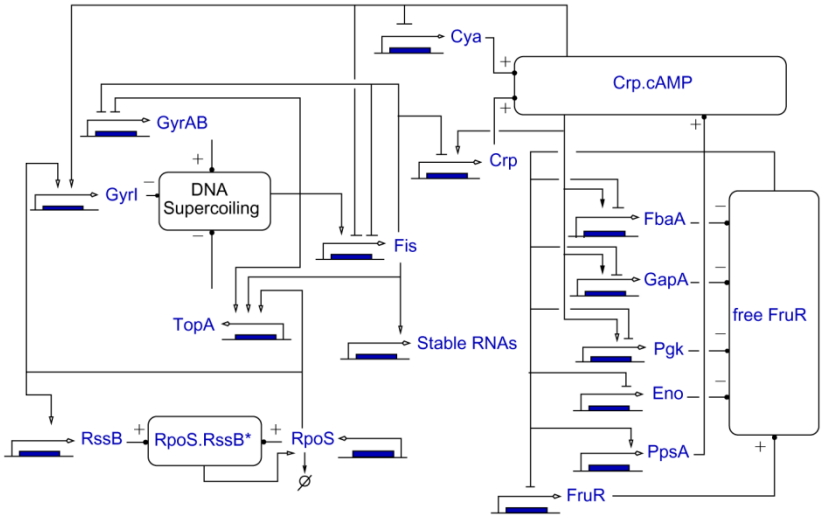
Baldazzi *et al.* (2010), *PLoS Comput. Biol.*, 6(6):e1000812

Interaction signs change with fluxes

- Radical changes in environment may **invert signs of indirect interactions**, because they change direction of metabolic fluxes and thus signs of elasticities



Network under glycolytic conditions



Network under gluconeogenic conditions

- Dynamic modification of feedback structure in response to environmental perturbations

Metabolic coupling and network dynamics

- Metabolic coupling changes network structure, but how does it affect **network dynamics**?
- **First approach**: reduce integrated network to gene regulatory network with metabolic coupling

$$\dot{x}^s = N^s \, v^s(x^s, g(x^s))$$

  - Description of effective network structure on time-scale of gene expression
  - Use of standard (qualitative or quantitative) models for describing direct and indirect interactions between genes

Qualitative modeling of network dynamics

- **Qualitative** models capture in simple manner complex dynamic of large regulatory networks without quantitative data

  Interesting in their own right, or first step towards fully quantitative modeling

- Approach based on description of network dynamics by means of **piecewise-affine (PA) DE** models

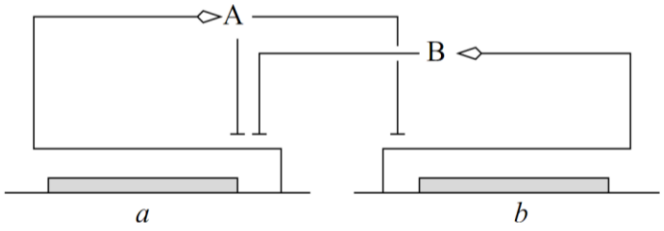
  PA models describe dynamics of gene regulatory networks by means of approximate, switch-like response functions

  Glass and Kauffman (1973), *J. Theor. Biol.*, 39(1):103-29

- Relation with discrete, logical models of gene regulation

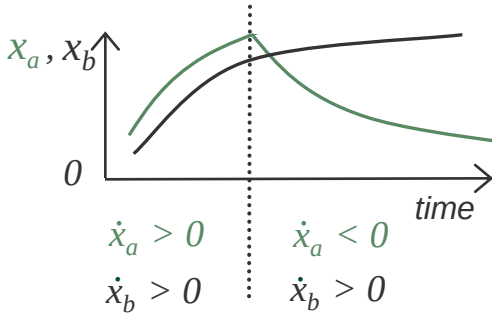
  Thomas and d'Ari (1990), *Biological Feedback,* CRC Press

Qualitative analysis of PA models

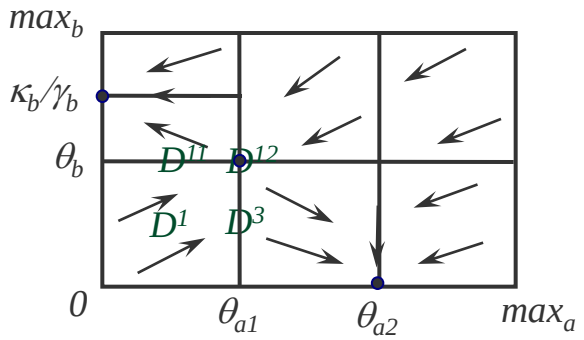


$$\dot{x}_a = \kappa_a \, s^-(x_a, \theta_{a2}) \, s^-(x_b, \theta_b) - \gamma_a \, x_a$$
$$\dot{x}_b = \kappa_b \, s^-(x_a, \theta_{a1}) - \gamma_b \, x_b$$

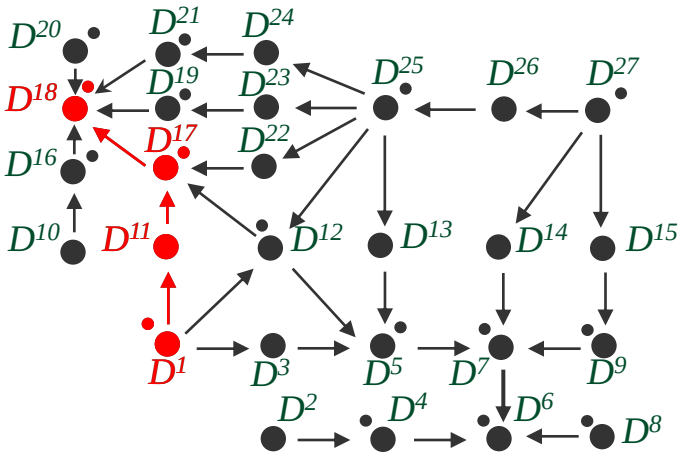
PA models using step functions



Model-checking for verification of system properties



Models easy to analyze, using inequalities



D^1: $\dot{x}_a > 0$, $\dot{x}_b > 0$

D^{17}: $\dot{x}_a < 0$, $\dot{x}_b > 0$

D^{18}: $\dot{x}_a = 0$, $\dot{x}_b = 0$

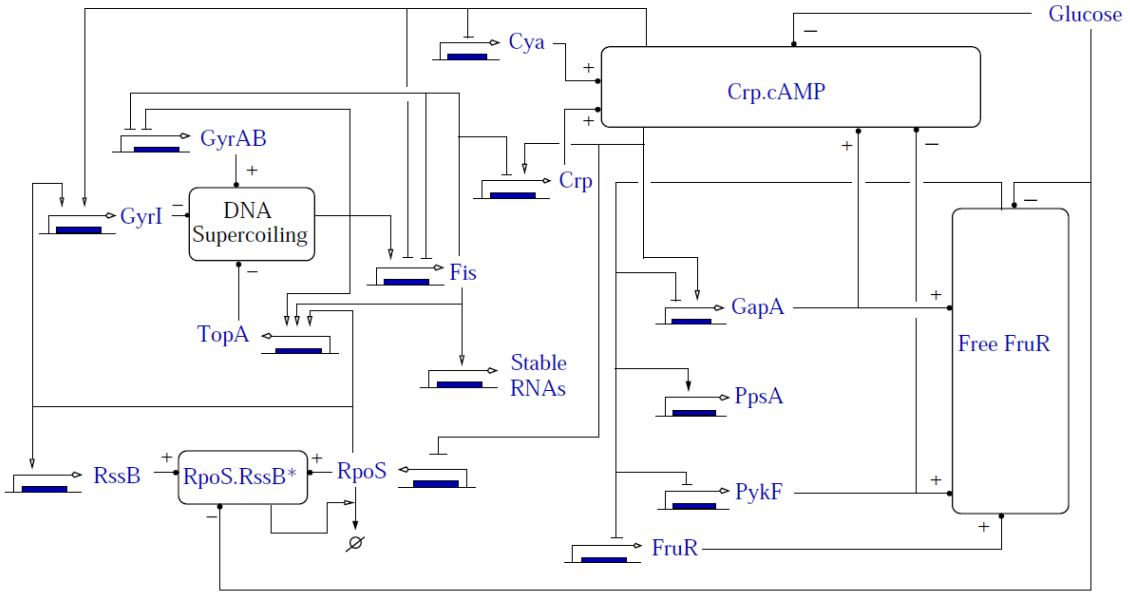
Predictions of qualitative dynamics, robust for large variations in parameter values

de Jong et al. (2004), Bull. Math. Biol., 66(2):301-40

Batt et al. (2005), Bioinformatics, 21(supp. 1): i19-i28

Formulation of PA models

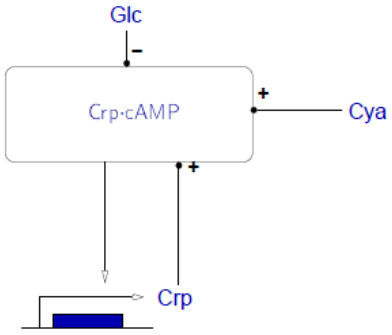
- Can PA models account for adaptations of gene expression in *E. coli* when bacteria following glucose-acetate diauxie?
- Translation of network diagram into PA models

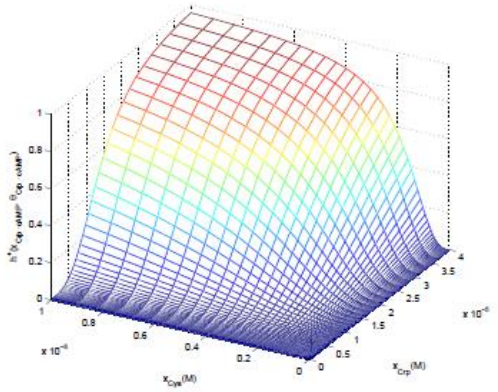


Baldazzi *et al.*, submitted

Formulation of PA models

- Can PA models account for adaptations of gene expression in *E. coli* when bacteria following glucose-acetate diauxie?
- Translation of network diagram into PA models
  - Straightforward for direct interactions…
  - … but also possible for indirect interactions





$$v_1(x_{Crp\cdot cAMP}) = \kappa_{crp}\, h^+(x_{Crp\cdot cAMP}, \theta_{Crp\cdot cAMP}, n_1)$$

$$x_{Crp\cdot cAMP} = g(x_{Crp}, x_{Cya}, u_{Glc}) = \frac{h^-(u_{Glc}, \theta_{Glc}, n_2)\, x_{Cya}}{h^-(u_{Glc}, \theta_{Glc}, n_2)\, x_{Cya} + K}\, x_{Crp}$$

$$v_1(x_{Crp}, x_{Cya}, u_{Glc}) = \kappa_{crp}\, h^-(u_{Glc}, \theta_{Glc}, n_2)\, h^+(x_{Crp}, \theta_{Crp}, n_3)\, h^+(x_{Cya}, \theta_{Cya}, n_4)$$

Baldazzi *et al.*, submitted

Dynamic analysis of metabolic coupling

- Can PA models account for adaptations of gene expression in *E. coli* when bacteria following glucose-acetate diauxie?
- Comparison of model predictions with published data sets: indirect interactions induced by metabolic coupling are **essential for reproducing gene expression dynamics**

  Steady-state mRNA concentration levels and initial transcriptional response of metabolic and regulatory genes

	crp	*fis*	*rpoS*	*fruR*	*gapA*	*ppsA*	*pykF*	Reference vs model
Experimental data	?	-	+	?	-	+	-	[29]
	-	-	+	+		+	-	[34]
						+	-	[35]
Model predictions	+	-	+	0	-	+	-	$\mathcal{M}_{neo}$ vs $\mathcal{M}_{gluco}$
	0	0	+	0	-/0	+/0	-/0	$\mathcal{M}^0_{neo/\text{Crp·cAMP}}$ vs $\mathcal{M}^0_{glyco/\text{Crp·cAMP}}$
	+	-	+	0	+	0	0	$\mathcal{M}^0_{neo/\text{free FruR}}$ vs $\mathcal{M}^0_{glyco/\text{free FruR}}$
	0	0	0	0	0	0	0	$\mathcal{M}^0$

Baldazzi *et al.*, submitted

Metabolic coupling and network dynamics

- Metabolic coupling changes network structure, but how does it affect **network dynamics**?
- **Second approach**: explicit modeling of metabolism using kinetic rate laws

$$N^f \, v^f(x^s, x^f) = 0 \;\Rightarrow\; x^f = g(x^s)$$
$$\dot{x}^s = N^s \, v^s(x^s, x^f)$$

  - Excellent examples available in literature

Kotte *et al*. (2010), *Mol. Syst. Biol*., 6: 355

Bettenbrock (2005), *J. Biol. Chem.*, 281(5):2578-84

  - But … rate laws are nonlinear, so no analytic expression for g, and ...
  - Obtaining reliable parameter values from data is currently bottleneck

Metabolic coupling and network dynamics

- Metabolic coupling changes network structure, but how does it affect **network dynamics**?
- **Modified second approach**: explicit modeling of metabolism using **approximate** kinetic rate laws

$$N^f\, v^f(x^s, x^f) = 0 \;\Rightarrow\; x^f = g(x^s)$$
$$\dot{x}^s = N^s\, v^s(x^s, x^f)$$

  - Approximate models that provide good phenomenological description of enzymatic rate laws: **linlog kinetics**
  - Estimation of parameter values in presence of noisy and missing data: **expectation-maximization (EM)** algorithm
  - Some preliminary results…

Berthoumieux *et al.* (2011), *Bioinformatics*, in press

Linlog models

- Linlog models approximate classical enzymatic rate laws:

$$v(x,u,e) = \mathrm{diag}(e) \cdot \left(a + B^x \cdot \ln(x) + B^u \cdot \ln(u)\right)$$

  - Internal and external metabolite concentrations $x \in \mathbb{R}_+^n$, $u \in \mathbb{R}_+^p$
  - Enzyme concentrations $e \in \mathbb{R}_+^m$
  - Parameters $a \in \mathbb{R}^m$, $B^x \in \mathbb{R}^{m \times n}$ and $B^u \in \mathbb{R}^{m \times p}$

Heijnen (2005), *Biotechnol. Bioeng.*, 91(5):534-45

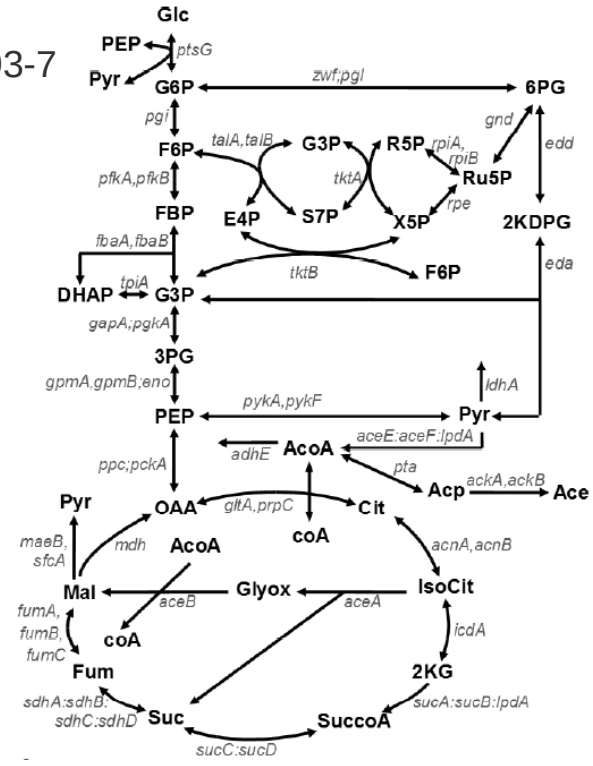
- Linlog models have several advantages for our purpose:
  - Analytical solution of g
  - Parameter estimation reduced to linear regression problem
  - Parameters have interpretation in terms of elasticity coefficients

Parameter estimation in linlog models

- High-throughput data sets are becoming available that allow estimation of parameters in linlog models

  Parallel measurement of enzyme and metabolite concentrations, and metabolic fluxes Ishii *et al.* (2007), *Science*, 316(5284):593-7

- Estimation of parameters in linlog models from experimental data
  - Technical problems: missing data, non-identifiability issues, …
  - EM approach for estimation of parameter values, tailored to linlog models



Berthoumieux et al. (2011), *Bioinformatics*, in press

Application to *E. coli* central metabolism

- Evaluation of results by comparing estimated and known signs of elasticities
  - Distinction between non-identifiable, non-significant, correctly and wrongly estimated elasticity signs
  - Discrepancies due to missing values, noise, reactions near equilibirum, and …

	Metabolite / Enzyme	Glc	PEP	G6P	Pyr	F6P	FBP	DHAP	3PG	AcoA coA	6PG	Ru5P	R5P	S7P	2KG	Suc	Fum	Mal	ATP ADP	Cit	NADPH NADP	NADH NAD	FAD	Ace	# complete datapoints
1	PtsG	0.29	-0.88	0.79	1.88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14
2	Pgi	0	0	-0.33	0	0.23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27
3	PfkA,PfkB	0	-0.13	0	0	-0.11	-0.23	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	18
4	FbaA,FbaB	0	0	0	0	0	-0.29	0.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
5	TpiA	0	0	0	0	0	0	-0.07	-0.22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
6	GapA;Pgk	0	0	0	0	0	-0.18	0	-0.05	0	0	0	0	0	0	0	0	0	0.32	0	0	-0.17	0	0	5
7	GpmA,GpmB;Eno	0	0.28	0	0	0	0	0	-0.12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24
8	PykA,PykF	0	0.19	0	0.43	0	-0.14	0	0	0	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	11
9	AceE:AceF:LpdA	0	0	0	0.05	0	0	0	0	-0.05	0	0	0	0	0	0	0	0	0	0	0	0.27	0	0	6
10	Zwf;Pgl	0	0	-0.1	0	0	0	0	0	0	-0.21	0	0	0	0	0	0	0	0	0	-1.38	0	0	0	2
11	Gnd	0	0	0	0	0	0	0	0	0	0.33	-0.08	0	0	0	0	0	0	0	0	-0.76	0	0	0	2*
12	Rpe	0	0	0	0	0	0	0	0	0	0	0.48	0	-0.39	0	0	0	0	0	0	0	0	0	0	28
13	RpiA,RpiB	0	0	-0.64	0	0	0	0	0	0	0	0.26	-0.1	0	0	0	0	0	0	0	0	0	0	0	18
14	TktA	0	0	0	0	0	0	0	0	0	0	0	0.01	-0.1	0	0	0	0	0	0	0	0	0	0	18
15	TalA,TalB	0	0	0	0	-0.25	0	0	0	0	0	0	0	0.29	0	0	0	0	0	0	0	0	0	0	27
16	TktB	0	0	0	0	-0.58	0	0	0	0	0	0.35	0	0	0	0	0	0	0	0	0	0	0	0	27
17	GltA,PrpC	0	0	0	0	0	0	0	0	0.22	0	0	0	0	0.01	0	0	0	0	0.49	0	-0.02	0	0	1*
18	AcnA;AcnB	0	0	0	0	0	0	0	0	0	0	0	0	0	0.99	0	0	0	0	0.55	0	0	0	0	5
19	IcdA	0	0	0	0	0	0	0	0	0.07	0	0	0	0	0.002	0	0	0	0	0	0.001	0	0	0	4
20	SucA:SucB:LpdA;SucC:SucD	0	0	0	0	0	0	0	0	0	0	0	0	0	1.26	0.3	0	0	0	0	0	-0.48	0	0	2*
21	SdhA:SdhB:SdhC:SdhD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.08	-0.44	0	0	0	0	0	0.4	0	21
22	FumA,FumB,FumC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.46	0.1	0	0	0	0	0	0	25
23	Mdh	0	0.15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.01	0	0.31	0	-0.12	0	0	3*
24	Ppc;PckA	0	0.29	0	0	0	-0.13	0	0	0	0	0	0	0	0	0	0	0.1	-1.15	0.21	0	0	0	0	9
25	MaeB,SfcA	0	0	0	0.08	0	0	0	0	-0.2	0	0	0	0	0	0	0	0.26	0	0	-0.003	0.04	0	0	2*
26	AceA;AceB	0	0	0	0	0	0	0	0	-0.11	0	0	0	0	0	-0.01	0	-0.02	0	0	0	0	0	0	25
27	μ	0	0.16	-0.11	0.06	-0.04	0	0	0.11	0.1	0	0	0.09	0	-0.06	0	0	0	-0.44	0	0.05	-0.01	0	0	0*
28	Edd;Eda	0	0	0	2.47	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	5
29	Pta;AckA,AckB	0	0	0	0.0015	0	0	0	0	-0.33	0	0	0	0	0	0	0	0	-1.42	0	-0.26	-0.83	0	2.4	2*
30	LdhA	0	0	0	2.67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.11	0	0	6
31	AdhE	0	0	0	0	0	0	0	0	0.91	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28
	% Missing Data	3	17	0	48	7	34	59	10	3	72	3	38	3	59	3	14	14	0	62	79	79	17	17	

Berthoumieux et al. (2011), *Bioinformatics*, in press

Conclusions

- **Metabolic coupling** gives rise to indirect interactions between enzymes and genes in gene regulatory networks

  Systematic derivation of effective structure of gene regulatory network on time-scale of gene expression

- Metabolic coupling leads to **densely-connected** networks with **robust** and **flexible** structure
  - Robust to changes kinetic properties (results not dependent on parameter values and rate laws)
  - Flexible rewiring of network structure following radical changes in environment (changes in flux directions)

- Including metabolic coupling in dynamic models is **essential for reproducing gene expression data**

Contributors and sponsors

Valentina Baldazzi, INRA Avignon
Matteo Brilli, INRIA Grenoble-Rhône-Alpes
Sara Berthoumieux, INRIA Grenoble-Rhône-Alpes
Eugenio Cinquemani, INRIA Grenoble-Rhône-Alpes
Hidde de Jong, INRIA Grenoble-Rhône-Alpes
Johannes Geiselmann, Université Joseph Fourier, Grenoble
Daniel Kahn, INRA, CNRS, Université Claude Bernard, Lyon
Yves Markowicz, Université Joseph Fourier, Grenoble
Delphine Ropers, INRIA Grenoble-Rhône-Alpes

Courtesy Guillaume Baptist (2008)

European Commission,
FP6, NEST program

Agence Nationale de la
Recherche, BioSys program