Towards integrated models of regulatory networks: the MetaGenoReg project

Hidde de Jong and Daniel Kahn

INRIA Grenoble - Rhône-Alpes

Hidde.de-Jong@inria.fr

http://ibis.inrialpes.fr

LBBE, Université de Lyon

Daniel.Kahn@univ-lyon1.fr

http://lbbe.univ-lyon1.fr

LBBE
BIOMETRIE ET BIOLOGIE EVOLUTIVE

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
INRIA
centre de recherche GRENOBLE - RHÔNE-ALPES
Lyon 1
INRA
Institut National de la Recherche Agronomique

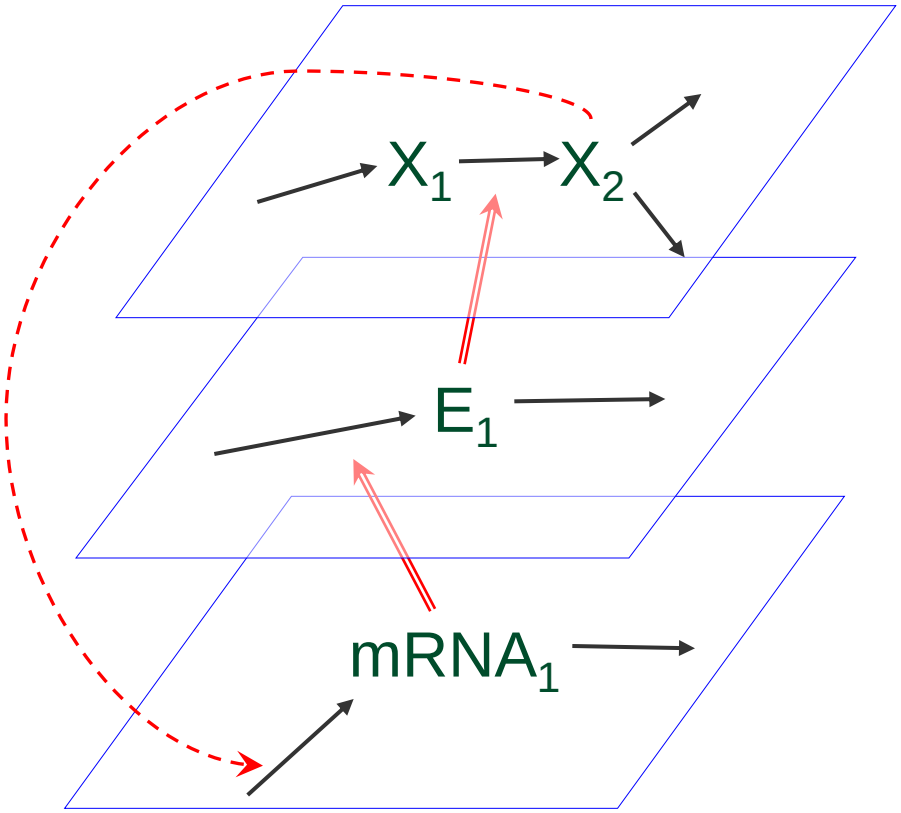
Overview

1. General question of biological regulation:
   the MetaGenoReg project
2. Analysis of metabolic coupling in gene regulatory networks
3. Metabolic model
4. Integrating gene and metabolic models

1. General question of biological regulation

- Cellular regulation involves several levels, including:
  - Gene regulatory networks
  - Metabolic regulation
- These levels interact:
  - Gene expression impacts metabolism through changes in enzyme concentrations
  - Conversely metabolism influences gene expression
- What is the rationale articulating both types of regulation?
  - Are they interchangeable ?
  - How much are they constrained?
  - What is the relative importance of gene and metabolic regulation?

‘Hierarchical’ analysis



module 1
metabolism

module 2
protein level

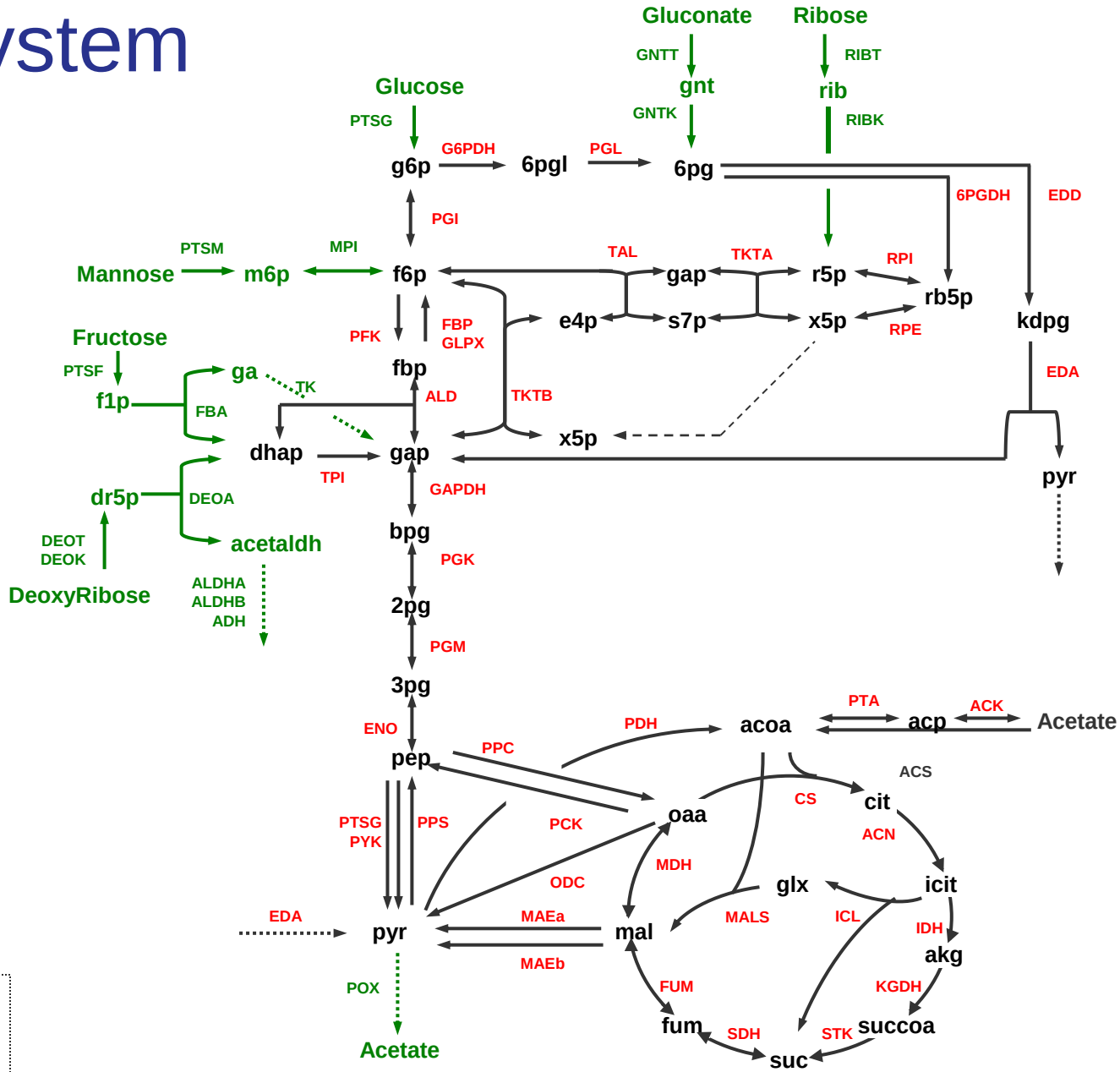
module 3
RNA level

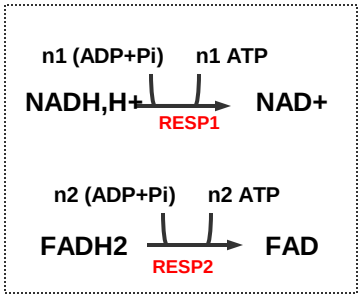
MetaGenoReg project outline

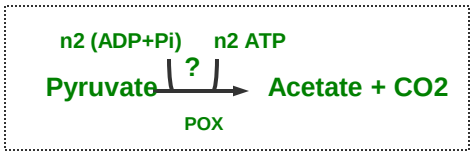
- Modelling combined metabolic and gene regulation
  - Reduce and simplify in order to understand the system's behaviour
  - Develop a method for joint modelling combining different approximations suited to both types of regulation
  - Measure their respective contribution
- Analyse the model's strengths and weaknesses from a systemic point of view
- Understand the biological rationale underlying the distribution of regulation between metabolism and gene expression

Biological system

E. coli carbon metabolism

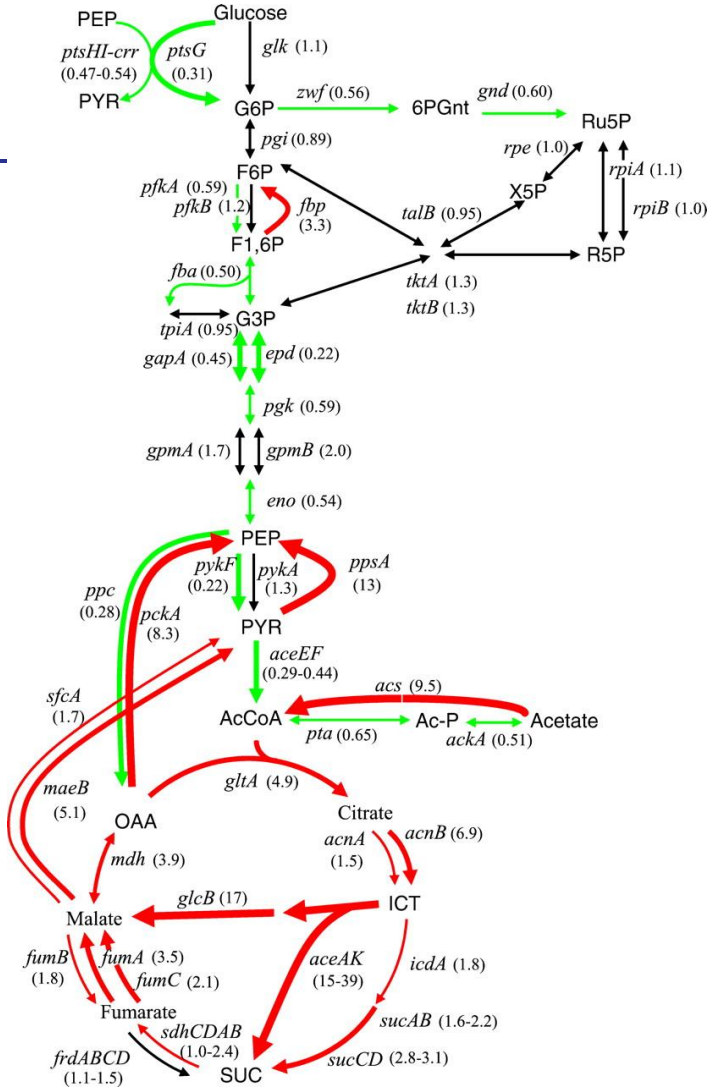






Glucose-acetate diauxie

- Well-characterised transition in *E. coli*
- Involves major changes
  - at the metabolic level: gluconeogenesis *vs.* glycolysis
  - at the gene expression level
- Strong interaction between metabolic and gene expression levels



Oh *et al.* (2002) *J Biol Chem.* 277:13175-83.

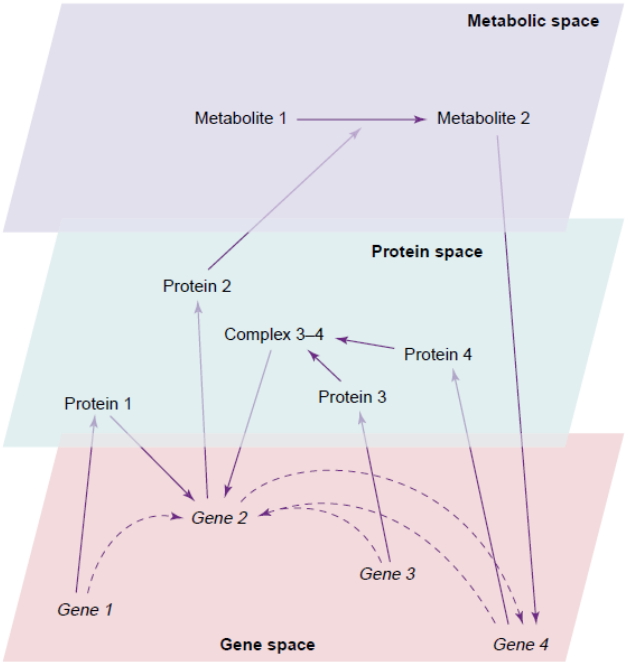
2. Analysis of metabolic coupling

- **Gene regulatory networks** control changes in gene expression in response to environmental perturbations
- They 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 include **direct** interactions (transcription regulation), but also **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?

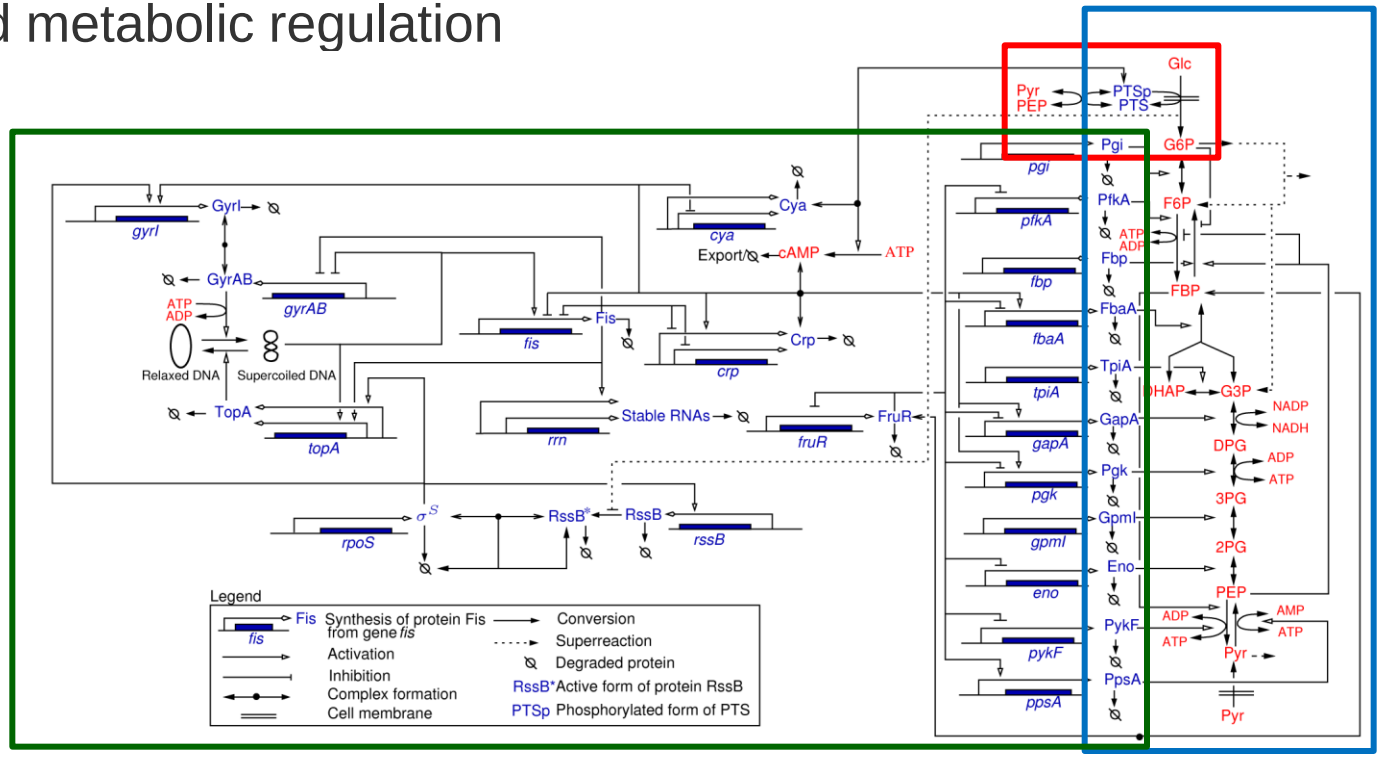
  How can we **derive indirect interactions** from underlying system of biochemical reactions?

  How do indirect interactions **influence system dynamics**?
- 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** can this 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?

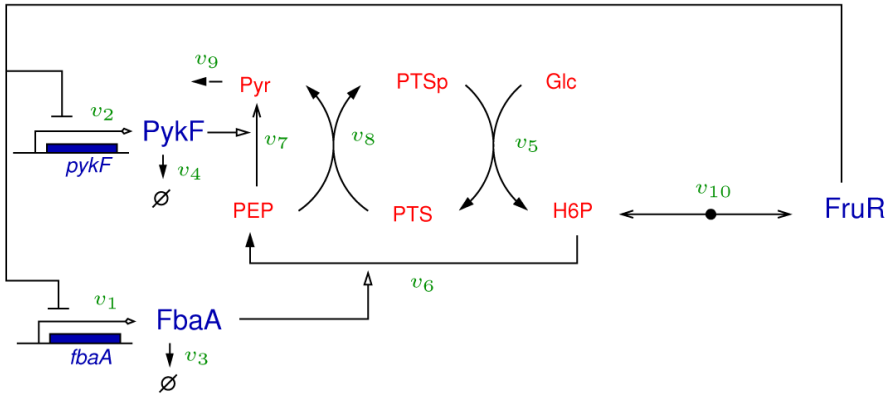
- Approach based on reduction of stoichiometric model of system of biochemical reactions, making following weak assumptions:
  - Distinct time-scale hierarchies between metabolism and gene expression: model reduction using **quasi-steady-state approximation**
  - Stability of fast subsystem: use of **control**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

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 $v : \mathbb{R}^n_+ \to \mathbb{R}^q$
  - 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

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

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**
- 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

- 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**

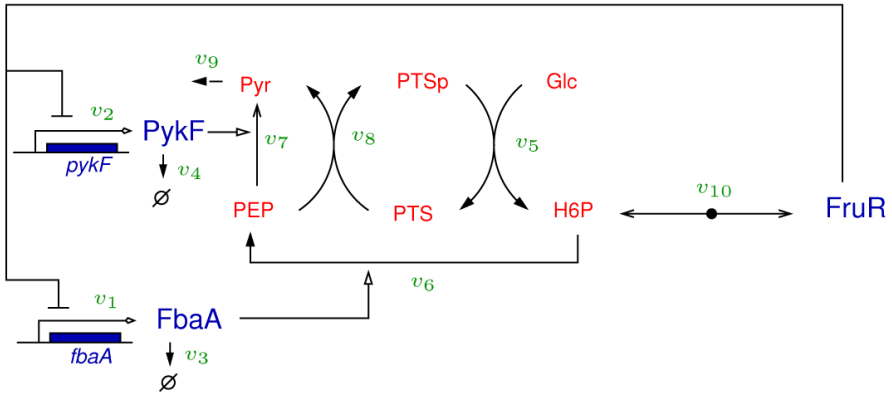
  Slow variables are typically **total protein concentrations**, fast variables **metabolites and biochemical complexes**

Kinetic models and time-scale hierarchy

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

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

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



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

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

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**

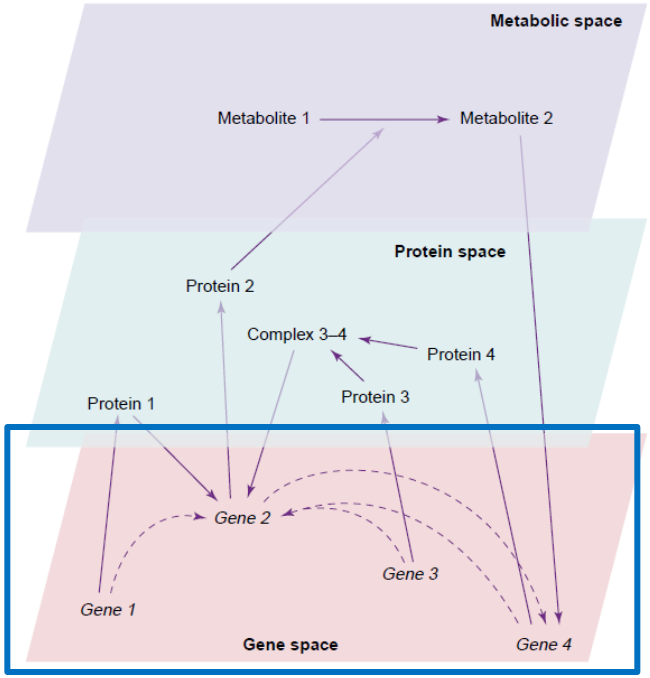
Analysis of metabolic coupling

- **Gene regulatory networks** control changes in gene expression in response to environmental perturbations
- They 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 include **direct** interactions (transcription regulation), but also **indirect** interactions (mediated by metabolism)

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



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 QSSA equation $N^f \, v^f(x^s, x^f) = 0$ 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)**

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

Concentration control coefficients

- **Concentration control coefficients** characterize the steady-state response of fast variables (metabolite concentrations) to changes in slow variables (enzyme concentrations)

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

Jacobian matrix and regulatory structure

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

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

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

$$M = N^f \boxed{\partial v^f(x^f, \bar{x^s})/\partial x^f}$$

Elasticity coefficients

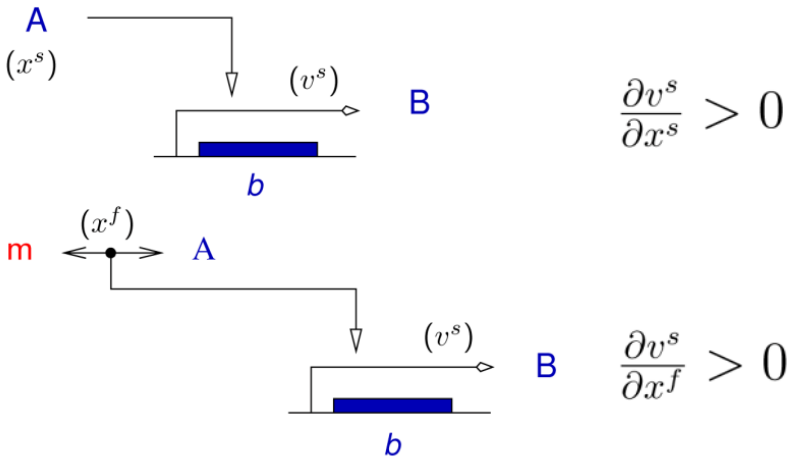
- System-level response expressed in terms of **elasticity coefficients**, which quantify the response of reaction rates to changes in variables

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

Determination of interaction signs

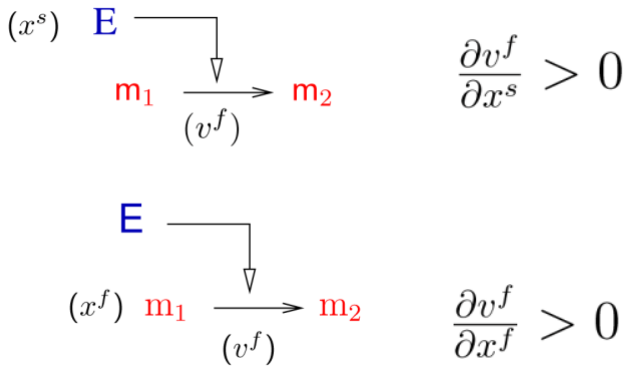
- Can we derive **signs for regulatory interactions** (elements of Jacobian matrix) without quantitative knowledge on rate laws and parameter values?

  Rate laws are generally monotone functions in variables, so signs of elasticities are known



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

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



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

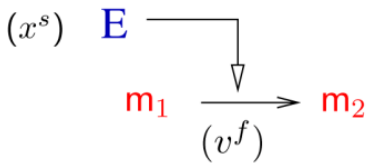
$$\frac{\partial v^f}{\partial x^f} > 0$$

Determination of interaction signs

- Can we derive **signs for regulatory interactions** (elements of Jacobian matrix) without quantitative knowledge on rate laws and parameter values?

  Rate laws are generally monotone functions in variables, so signs of elasticities are known

- 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

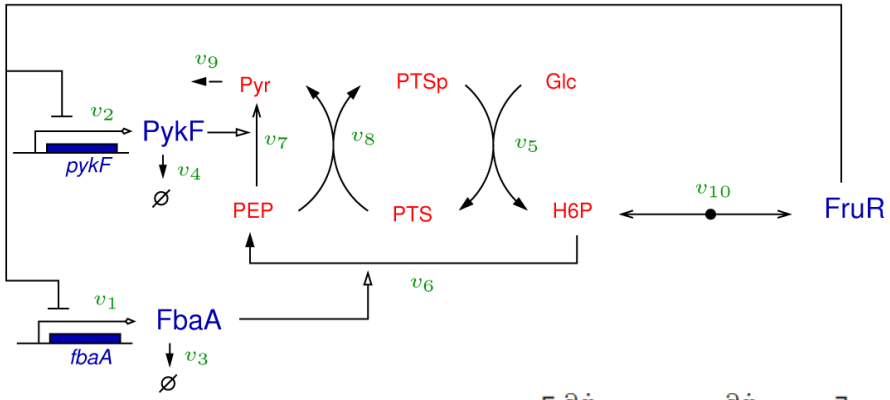
$$\mathcal{J} = \frac{\partial \dot{x}^s}{\partial x^s} = N^s \frac{\partial v^s(x^s, x^f)}{\partial x^s} + N^s \frac{\partial v^s(x^s, x^f)}{\partial x^f} \frac{\partial g(x^s)}{\partial x^s}$$

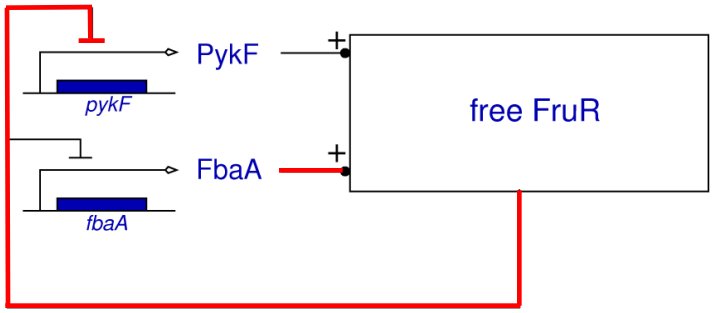
Symbolic Math Toolbox in Matlab

- 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



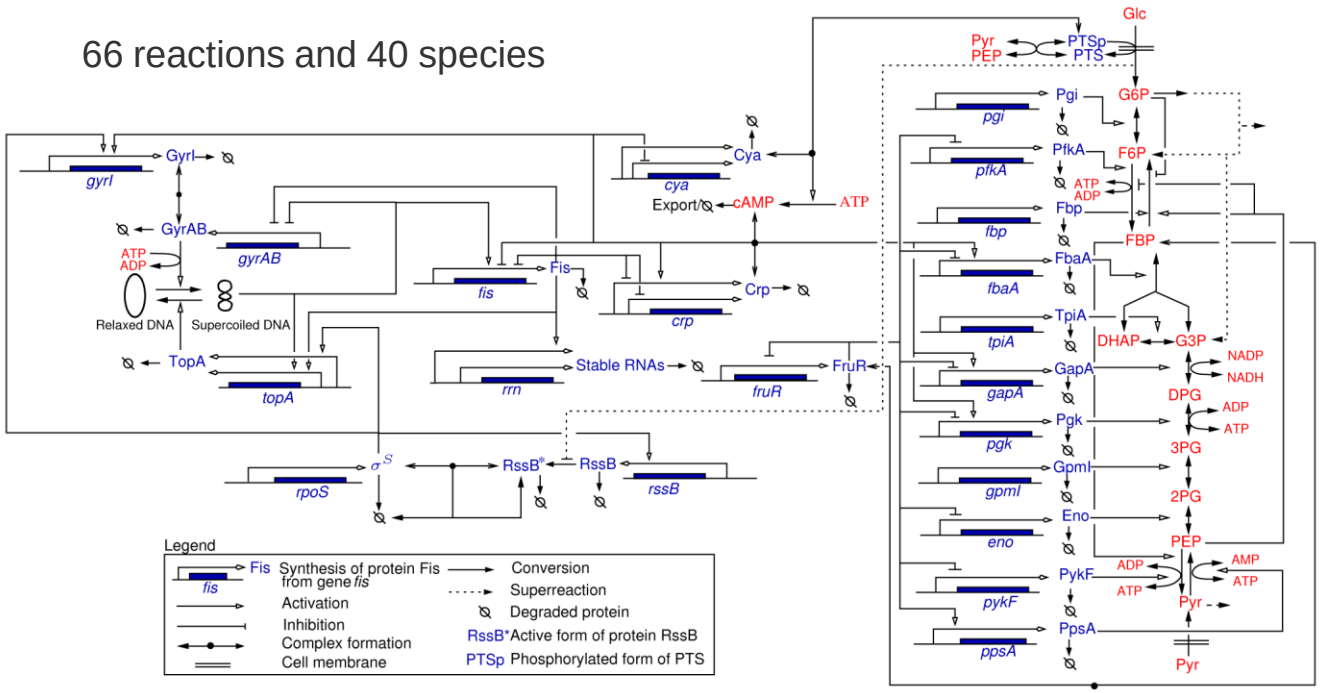


$$\mathcal{J} = \begin{bmatrix} \frac{\partial \dot{x}_{FbaA}}{\partial x_{FbaA}} & \frac{\partial \dot{x}_{FbaA}}{\partial x_{PykF}} \\ \frac{\partial \dot{x}_{PykF}}{\partial x_{FbaA}} & \frac{\partial \dot{x}_{PykF}}{\partial x_{PykF}} \end{bmatrix} \quad \text{and} \quad \text{sign}(\mathcal{J})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

Application to *E. coli* carbon assimilation

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

  Glycolysis/gluconeogenesis (growth on glucose/pyruvate)

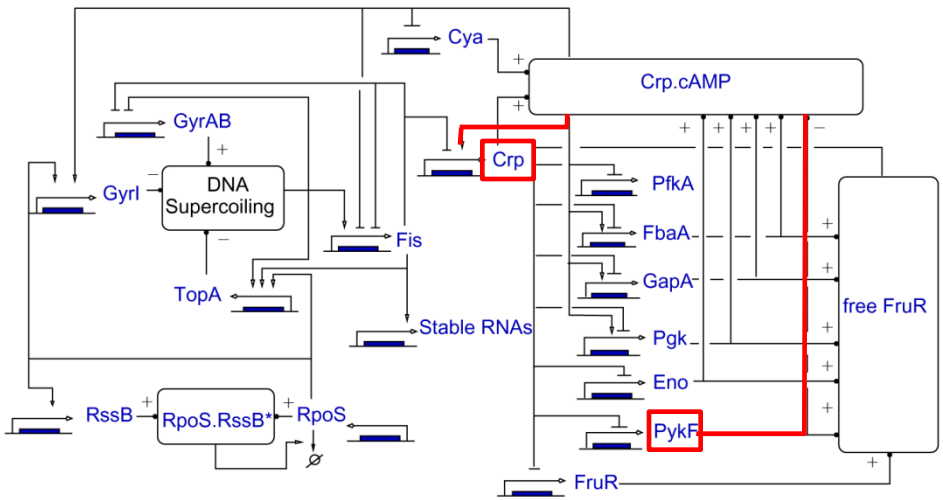


Application to *E. coli* carbon assimilation

- 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

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

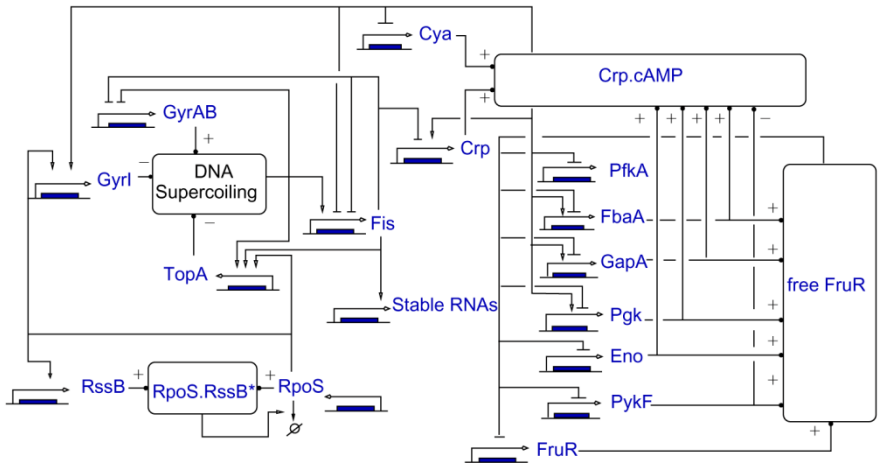
		Regulators															
		PfkA	FbaA	GapA	Pgk	Eno	PykF	Cya	Crp	Fis	GyrAB	GyrI	TopA	RpoS	RssB	stable RNAs	FruR
Genes	*pfkA*	0	-	-	-	-	-	0	0	0	0	0	0	0	0	0	-
	fbaA	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
	gapA	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
	pgk	0	-/+	-/+	-/+	-/+	-	+	+	0	0	0	0	0	0	0	-
	eno	0	-	-	-	-	-	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

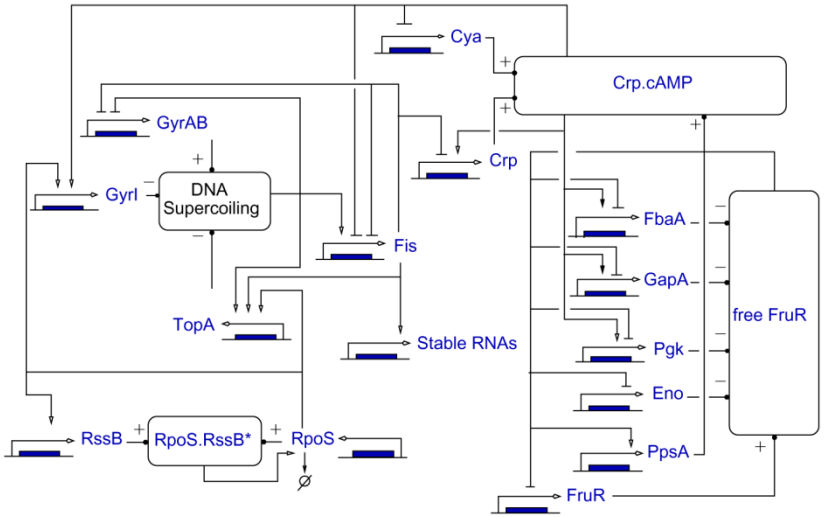
  Puccia and Levins (1985), *Qualitative Modeling of Complex Systems*, Harvard University Press

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 and concentration control



Network under glycolytic conditions



Network under gluconeogenic conditions

- Dynamic modification of feedback structure in response to environmental perturbations

Metabolic coupling and network dynamics

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

  How do indirect interactions **influence system dynamics**?

- **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-linear (PL) DE** models

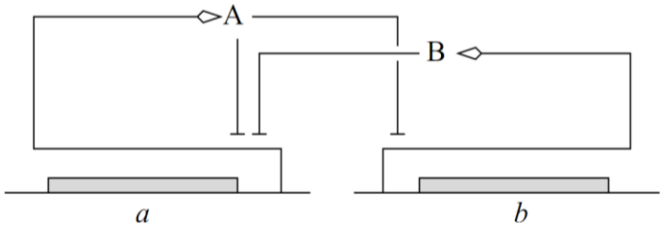
  PL models describe dynamics of gene regulatory networks by means of approximate, switch-like response functions

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

- Relation with discrete, logical models of gene regulation

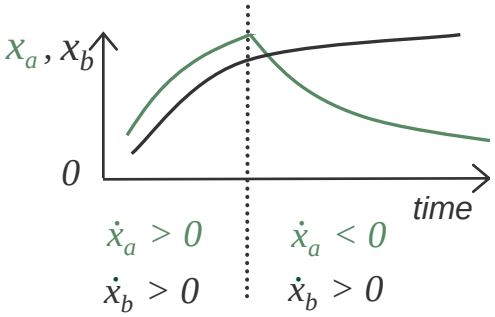
  Thomas and d'Ari (1990), *Biological Feedback,* CRC Press

Qualitative analysis of PL models

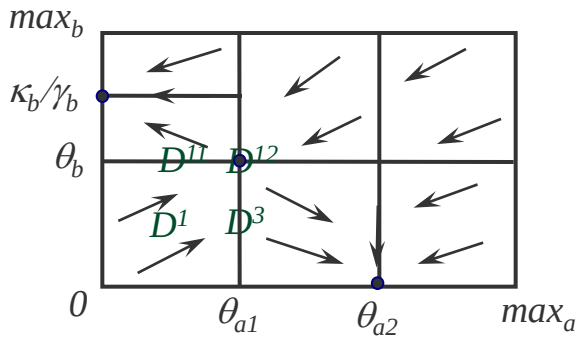


$$\dot{x}_a = \kappa_a \, s^-(x_a, \theta_{a2}) \, s^-(x_b, \theta_b) - \gamma_a \, x_a$$
$$\dot{x}_b = \kappa_b \, s^-(x_a, \theta_{a1}) - \gamma_b \, x_b$$

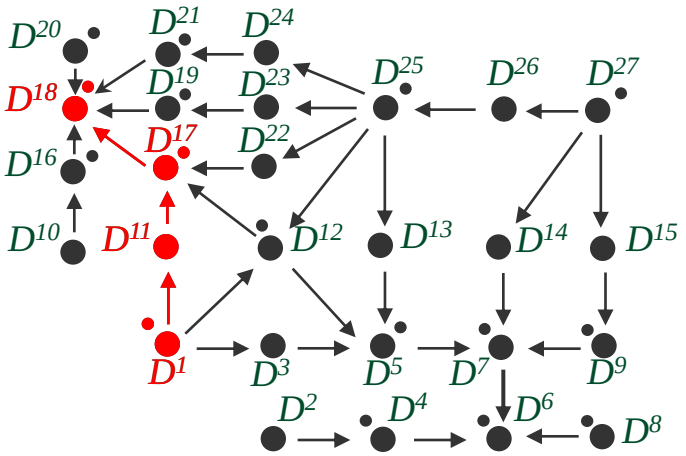
PL models using step functions



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

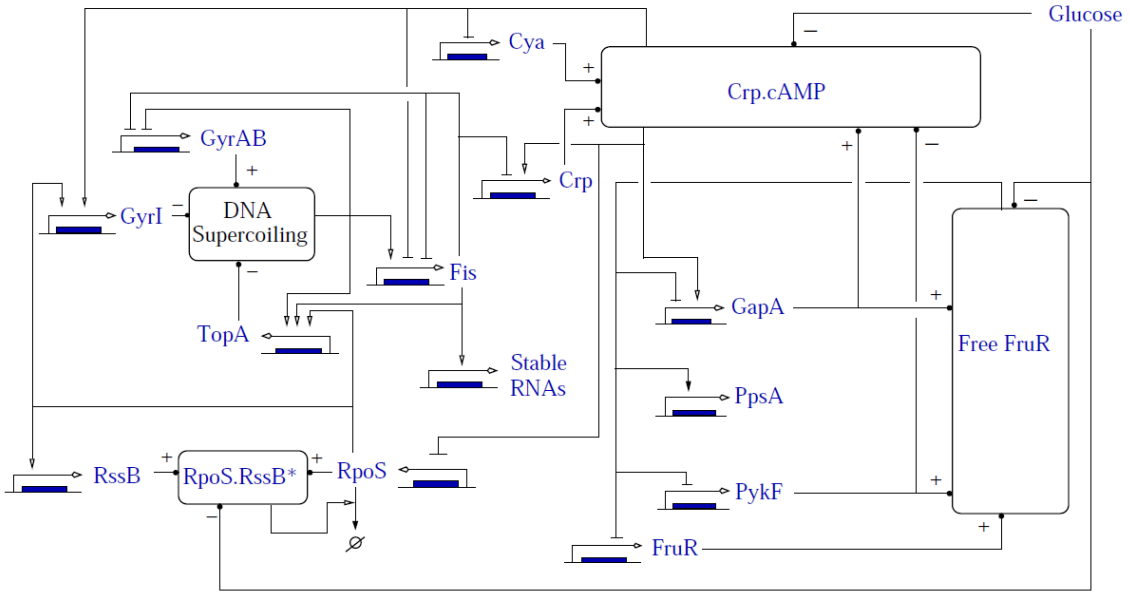
Predictions of qualitative dynamics, robust for large variations in parameter values

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

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

Formulation of PL models

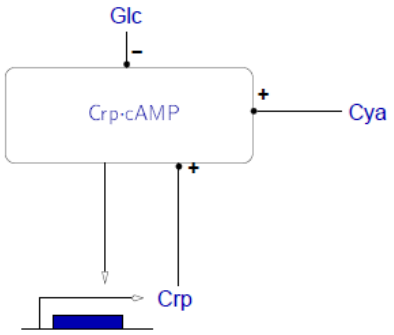
- Can PL models account for adaptation of gene expression in *E. coli* following glucose-acetate diauxie?
- Translation of network diagram into PL models

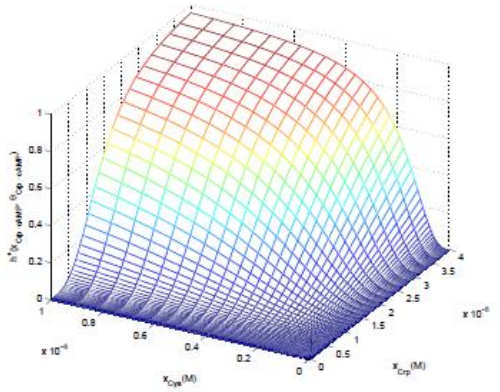


Baldazzi *et al.* (2012) *J. Theor. Biol.*, 295:100-115

Formulation of PL models

- Can PL models account for adaptation of gene expression in *E. coli* following glucose-acetate diauxie?
- Translation of network diagram into PL 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.* (2012) *J. Theor. Biol.*, 295:100-115

Dynamic analysis of metabolic coupling

- Can PL models account for adaptations of gene expression in *E. coli* following glucose-acetate diauxie?
- Comparison of model predictions with published data sets

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

- Indirect interactions induced by metabolic coupling are **essential for reproducing gene expression dynamics**

	crp	*fis*	*rpoS*	*fruR*	*gapA*	*ppsA*	*pykF*	Reference vs model
Experimental data	?	-	+	?	-	+	-	[29]
	-	-	+	+		+	-	[34]
						+	-	[35]
Model predictions	+	-	+	0	-	+	-	$\mathcal{M}_{neo}$ vs $\mathcal{M}_{glyco}$
	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.* (2012) *J. Theor. Biol.*, 295:100-115

Conclusions on metabolic coupling

- **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 to account for measured changes in gene expression**

Towards quantitative models?

- Above approach leads to models that view metabolism as intermediary between gene regulatory interactions
- However, metabolism is not explicitly modeled

  PL models aggregate and approximate complex rate functions in reduced model

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

- Moreover, models provide qualitative instead of quantitative picture of dynamics

  Qualitative models help provide intuitive idea of global system dynamics, but for some questions quantitative precision is required

Towards quantitative models?

- Another approach explicitly models metabolism and gene expression, followed by integration of two parts

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

- Approach based on suitable approximations of g
  - Approximations should provide good phenomenological description of metabolic rate laws
  - Minimal number of parameters to facilitate identification of parameter values from experimental data

3. Metabolic model

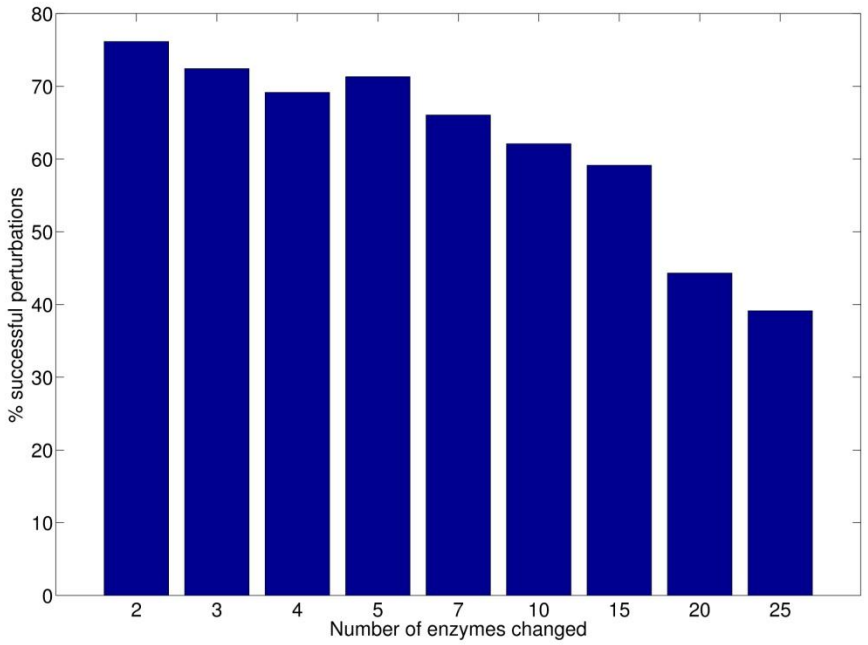
- Toy model entirely specified with ODEs
- 'Experimental' object used to test the quality of various reductions and approximations by comparison of simplified models with complete ODE model
- A suitable approximation would ideally allow us to calculate $g(x^s)$ analytically

Which approximations?

- Various types of linearisation of metabolic effects
- Compare reduced / approximated models with complete ODE-specified model

Assessing approximations of metabolism

- Randomly change enzyme concentrations in a 25-fold range on benchmark model (Matteo Brilli)
- Test steady-state obtention



Approximation 1, from MCT

$$\mathbf{v} = diag\ \boldsymbol{f}(\mathbf{x}) \cdot \mathbf{e}$$

$$d\mathbf{X}/d\mathbf{e} = \Gamma \cdot d\mathbf{v}/d\mathbf{e}$$

$$\Delta \ln \mathbf{X} \sim (diag\ \mathbf{X}_0)^{-1} \cdot \Gamma \cdot diag\ \boldsymbol{f}(\mathbf{X}_0) \cdot \Delta \mathbf{e}$$

Linearization around steady-state using control coefficients

Approximation 2, linlog

Linearization of kinetic laws:

$$\mathbf{v}(\mathbf{x}) \sim \mathit{diag}\ \mathbf{e} \cdot (\mathbf{A} + \mathbf{B} \cdot \ln \mathbf{x})$$

Steady-state implies:

$$\mathbf{N} \cdot \mathbf{v}(\mathbf{X}) = \mathbf{0}$$

$$\ln \mathbf{X} \sim - (\mathbf{N} \cdot \mathit{diag}\ \mathbf{e} \cdot \mathbf{B})^{-1} \cdot \mathbf{N} \cdot \mathit{diag}\ \mathbf{e} \cdot \mathbf{A}$$

Approximation 3, hyperbolic

Suggested from earlier work by Kacser:

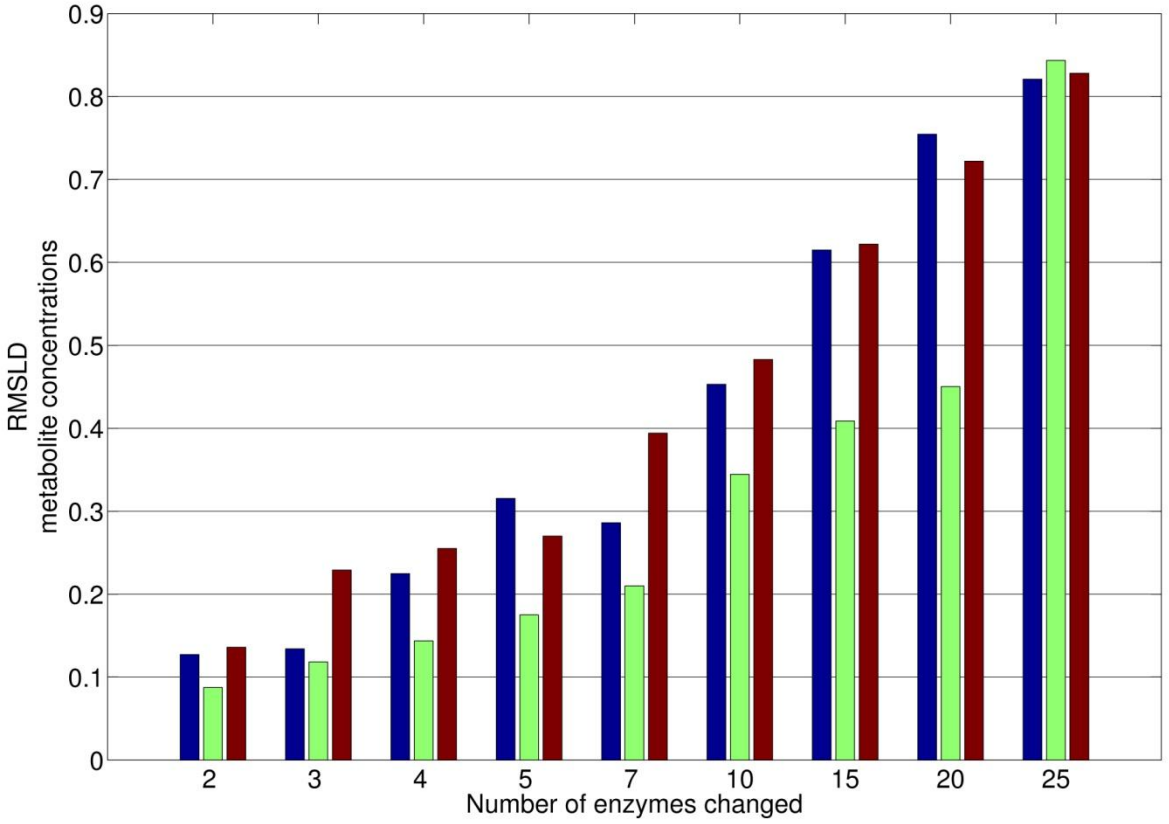
$$\Delta(1/\mathbf{X}) \sim (diag\ \mathbf{X}_0)^{-1} \cdot \mathbf{C}^X \cdot diag\ \mathbf{e}_0 \cdot \Delta(1/\mathbf{e})$$

$$\Delta(1/\mathbf{J}) \sim (diag\ \mathbf{J}_0)^{-1} \cdot \mathbf{C}^J \cdot diag\ \mathbf{e}_0 \cdot \Delta(1/\mathbf{e})$$

Linearization around steady-state using control coefficients

Metabolite estimates

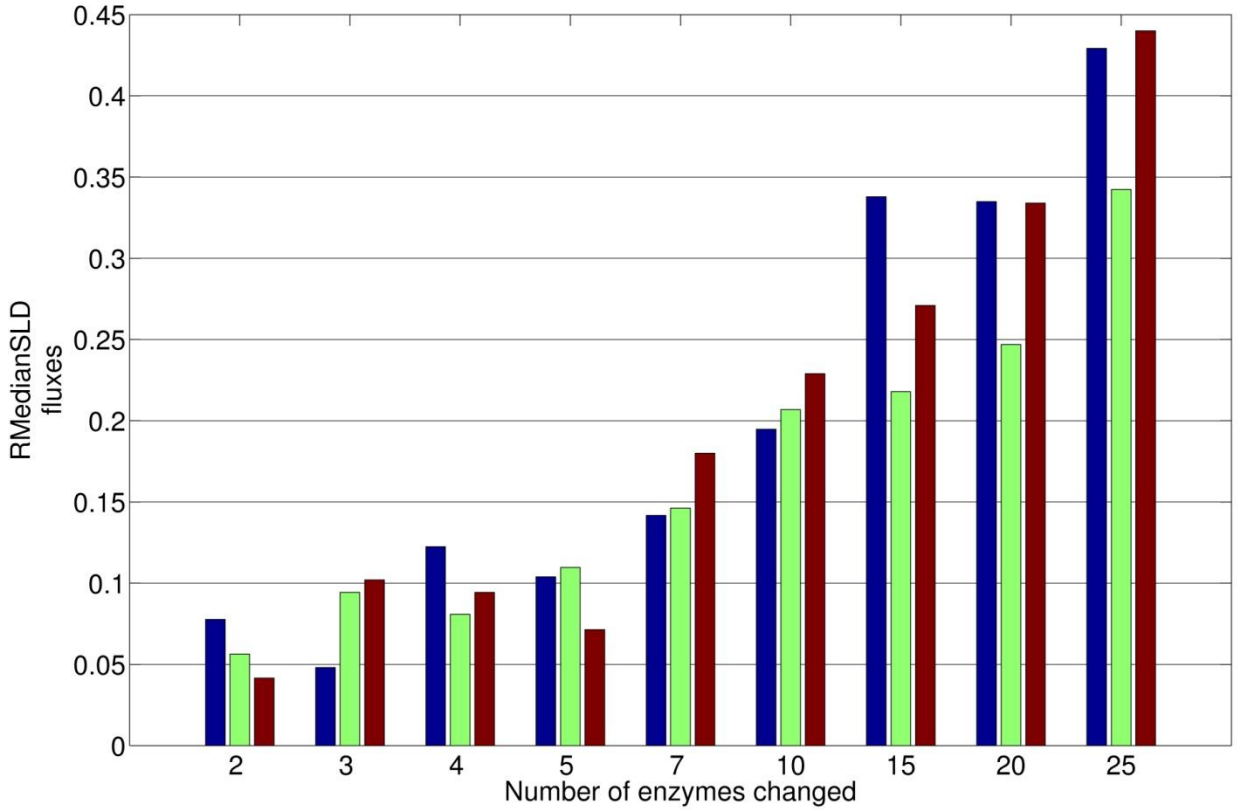
Root mean square Log deviation



Approximation MCA | Approximation linlog | Approximation hyperbolic

Flux estimates

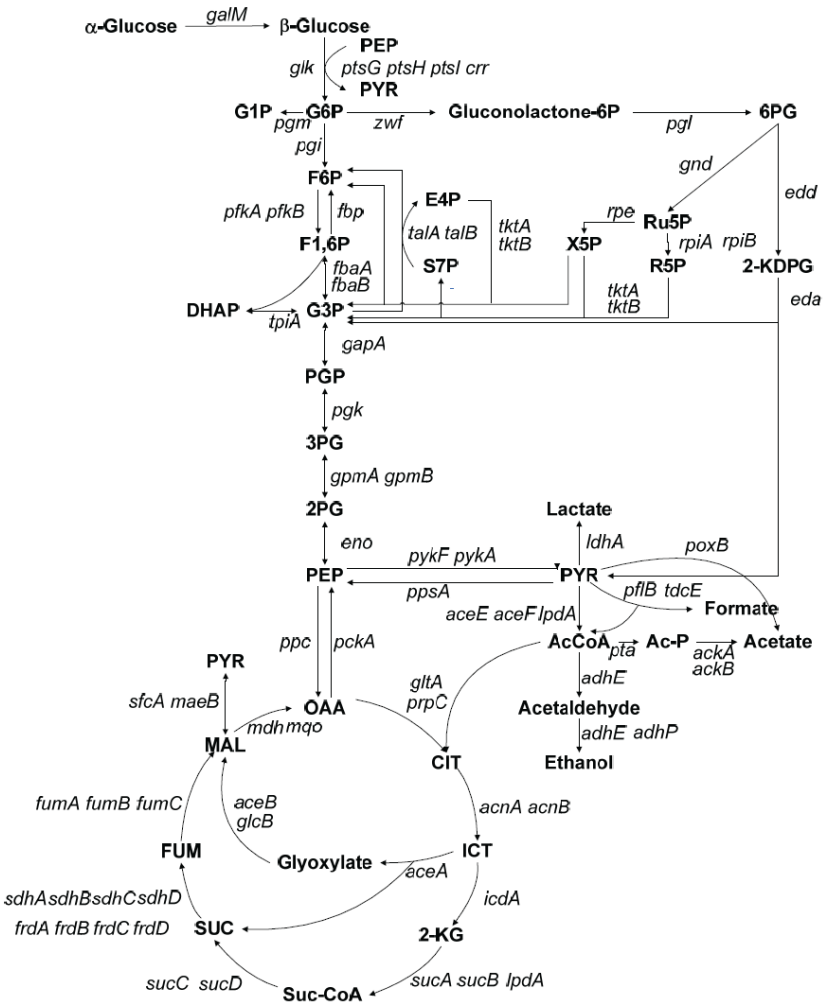
Median flux absolute Log deviation



Approximation MCA Approximation linlog Approximation hyperbolic

Model of *E. coli* carbon metabolism

- Simplified model
  - 32 reactions
  - 17 metabolites
- Linlog approximation
  $\mathbf{J} \sim \mathit{diag}\ \mathbf{e} \cdot (\mathbf{A} + \mathbf{B} \cdot \ln \mathbf{X})$
- → Independent linear regression possible for each reaction if sufficient data available :
  - Fluxes
  - Enzyme expression
  - Metabolite concentrations



Issues with metabolic model identification

- Difficulties to obtain high quality complete datasets
  - Fluxes, metabolite and enzyme concentrations with sufficient numbers of distinct observations
- Missing data can be handled efficiently
  - EM or maximum likelihood methods

    Berthoumieux *et al.* (2011) *Bioinformatics* 27:i186-i195
- Identifiability issues arise when there is insufficient variability or dependencies between metabolite concentrations because of
  - Reactions close to equilibrium
  - Steady-state constraints
  - Homeostasis

    Usefulness of dynamic non steady-state measurements (difficult to obtain)

Working around identifiability issues

- Use Principal Component Analysis on ln**X**
  - Reduce metabolite data by Singular Value Decomposition

$$\ln \mathbf{X} - \overline{\ln \mathbf{X}} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

  - Determine effective dimension of ln**X** from singular values σ_i, neglecting σ_i^2 smaller than experimental variance
  - Reduce metabolite data and reformulate identification accordingly

$$\mathbf{Y} = \mathbf{U}_r^T \left(\ln \mathbf{X} - \overline{\ln \mathbf{X}} \right)$$

  - Estimate parameters $\mathbf{B}_r$ for the reduced model such that

$$\mathbf{J} / \mathbf{e} - \overline{\mathbf{J} / \mathbf{e}} = \mathbf{B}_r \, \mathbf{Y}$$

  - One among an infinite number of choices for full parameters is

$$\mathbf{B} = \mathbf{B}_r \, \mathbf{U}_r^T$$

4. Integrating gene and metabolic models

- Identify separately the fast component (metabolic) and the slow component (gene expression)
- Use the resulting analytical model of metabolic steady-states as a ‘plugin’ function in the gene network model

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

- Critical issue: identification methods and, especially, quality and quantity of experimental data

Experimental data on metabolism

- Quantification of extra-cellular metabolites by means of nuclear magnetic resonance (NMR) spectroscopy

Experimental data on metabolism

- Quantification of intra-cellular metabolites by means of mass spectrometry

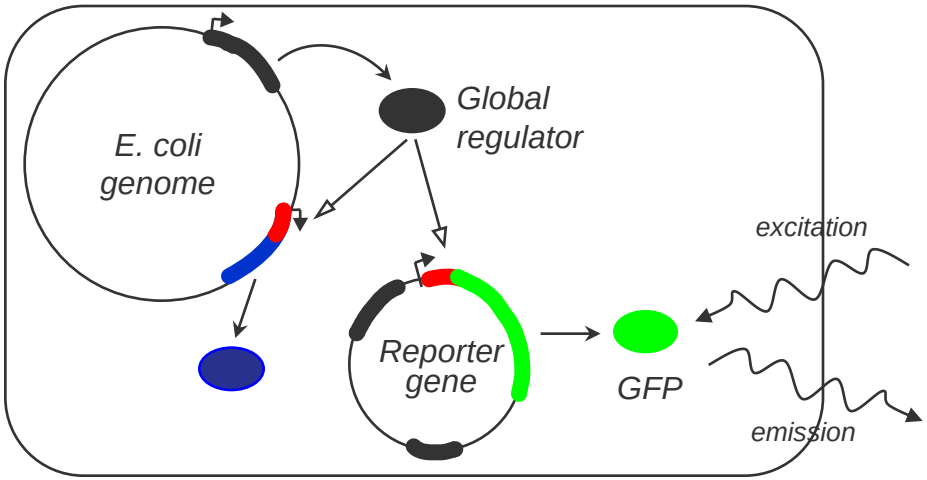
Isotope Dilution Mass Spectrometry

Experimental data on metabolism

- Quantification of intra-cellular metabolites by means of mass spectrometry

Experimental data on gene expression

- Quantification of gene expression by means of fluorescent and luminescent reporter genes
  - Expression of reporter gene is proportional to expression of target gene



Experimental data on gene expression

Prospect:
Roles of metabolic and gene regulation

- Identify parameters of the reduced system from data
- Study the metabolic response in the model when gene regulation is abolished
- Evaluate (quantify) the contribution of gene regulation to the metabolic response
- Conversely calculate the contribution of metabolic effects to gene regulation
- Understand the biological rationale underlying the distribution of regulation between metabolism and gene expression