The carbon assimilation network in *Escherichia coli* is densely connected and largely sign-determined by directions of metabolic fluxes

V. Baldazzi[1,4], D. Ropers[1], Y. Markowicz[1,2,3], D. Kahn[1,3,4,5], J. Geiselmann[1,2,3], H. de Jong[1]

PLoS Comput. Biol., 6(6):e1000812, 2010

1 INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE | INRIA | centre de recherche GRENOBLE - RHÔNE-ALPES

2 UNIVERSITE JOSEPH FOURIER SCIENCES TECHNOLOGIE MEDECINE

3 cnrs

4 INRA Institut National de la Recherche Agronomique

5 UJF

Lara, née le 26 juillet 2010

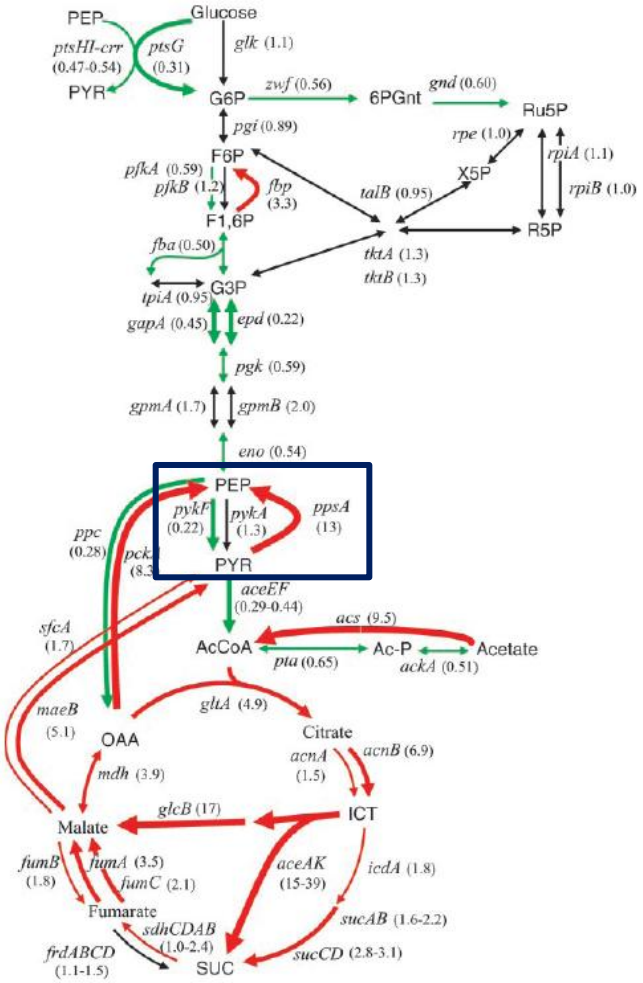
Gene regulatory networks

- 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



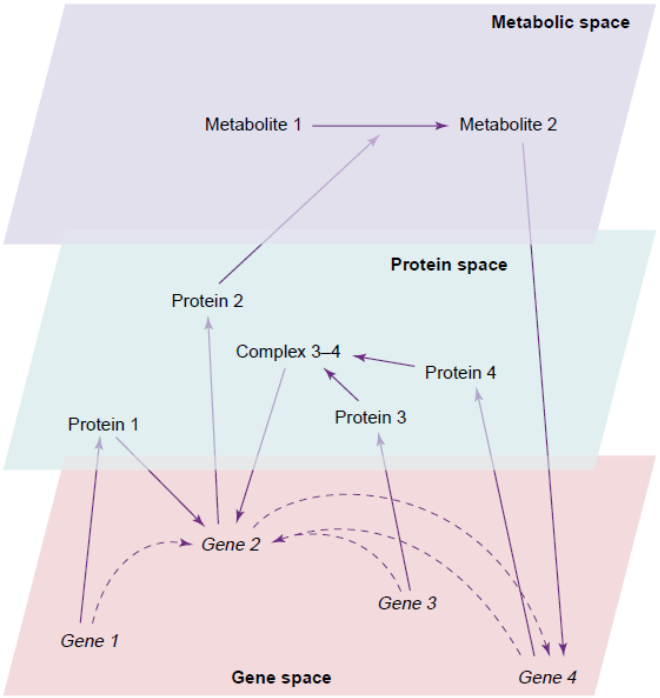
Gene regulatory networks

- **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

- By which method can we **derive the interaction structure** of gene regulatory networks in a principled way?

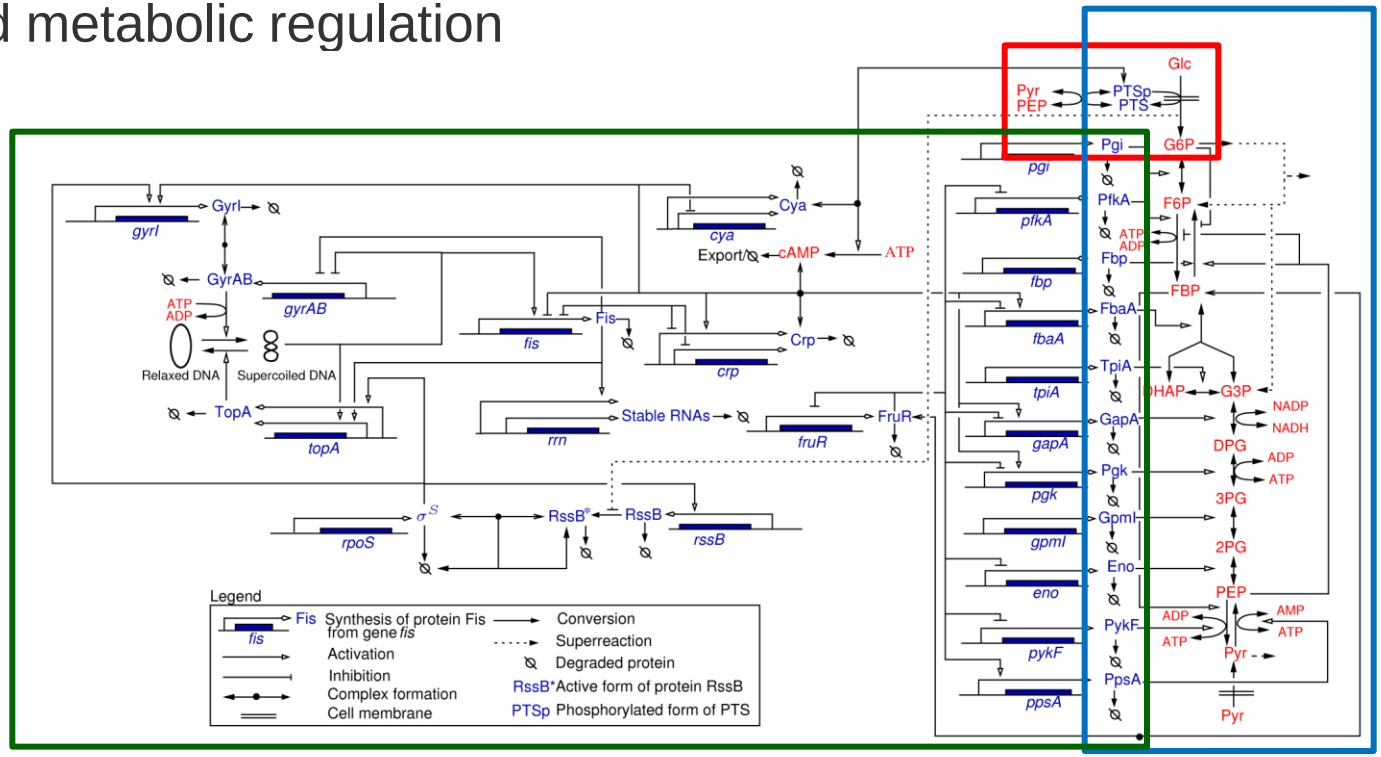
  How can we obtain indirect interactions from underlying system of biochemical reactions?

- 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 **derive the interaction structure** of gene regulatory networks in a principled way?

  How can we obtain 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

Model reduction using time-scale hierarchy

- Basic form of model $\dot{x} = N\, v(x)$
  - Concentration variables $x \in \mathbb{R}^n_+$
  - Reaction rates $v : \mathbb{R}^n_+ \to \mathbb{R}^q$
  - Stoichiometric 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

$$v = \begin{bmatrix} v^s & v^f \end{bmatrix}'$$

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

Model reduction using 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**, while $N^s \in \mathbb{Z}^m \times \mathbb{Z}^p$ and $N^{s'} \in \mathbb{Z}^{n-m} \times \mathbb{Z}^p$ are stoichiometric matrices for slow reactions and $N^f \in \mathbb{Z}^{n-m} \times \mathbb{Z}^{q-p}$ is stoichiometric matrix for fast reactions

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

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 dependencies between slow variables (total protein concentrations)

Mathematical representation of effective **gene regulatory network**

- Notice
  - 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**

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

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

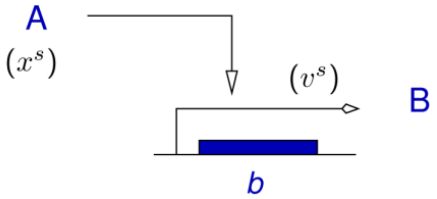
$$\frac{\partial g(x^s)}{\partial x^s} = \underbrace{-M^{-1} \, N^f}_{\text{Concentration control coefficients}} \frac{\partial v^f(x^s, x^f)}{\partial x^s}$$

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

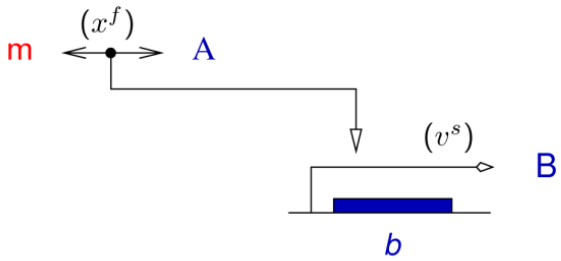
Determination of interaction signs

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

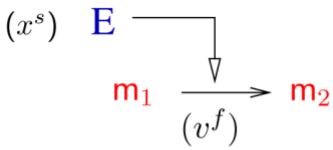
  Rate laws are generally monotone functions in variables



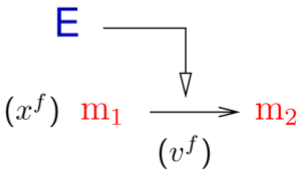
$$\frac{\partial v^s}{\partial x^s} > 0$$



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



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

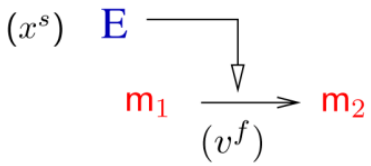


$$\frac{\partial v^f}{\partial x^f} > 0$$

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 fact that **signs of elasticities** are known

  Rate laws are generally monotone functions in variables
- Notice
  - Reversible reactions: signs of $\partial v^f(x^s, x^f)/\partial x^s$ change with flux direction



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

Determination of interaction signs

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

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

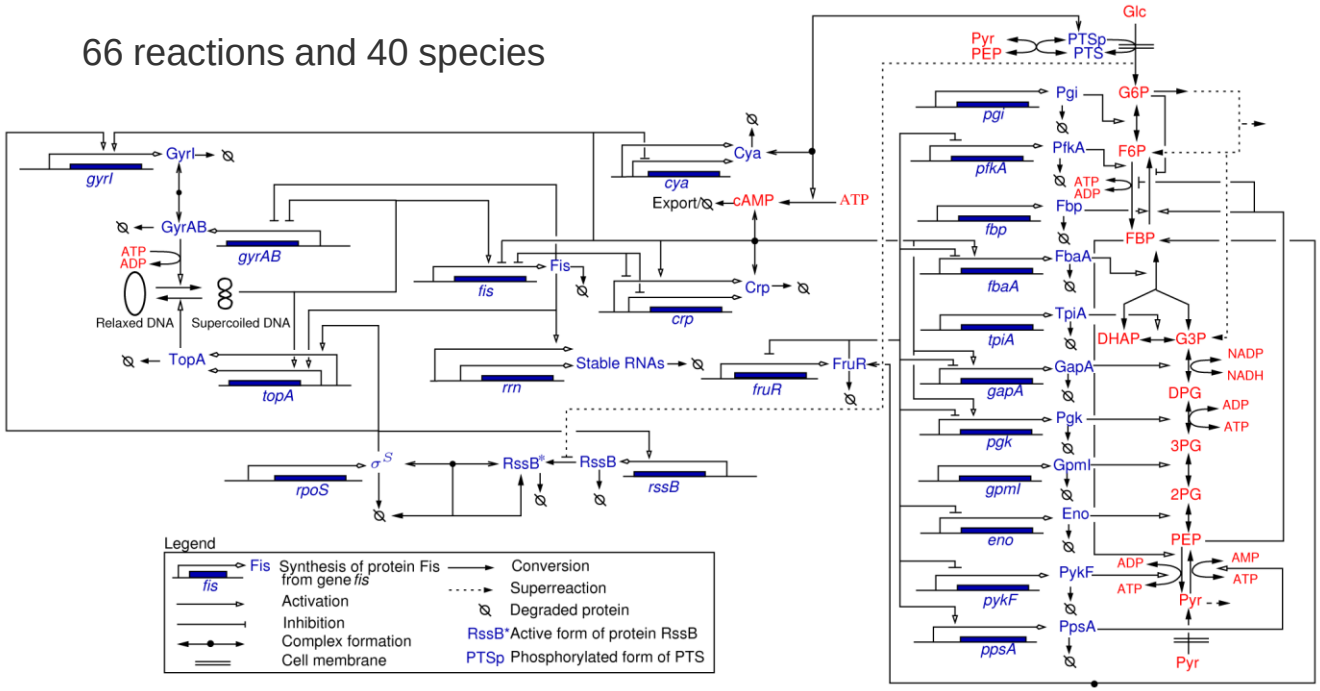
Symbolic Math Toolbox in Matlab

- 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)

Application to *E. coli* carbon assimilation

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

  Glycolysis/gluconeogenesis (growth on glucose/pyruvate)

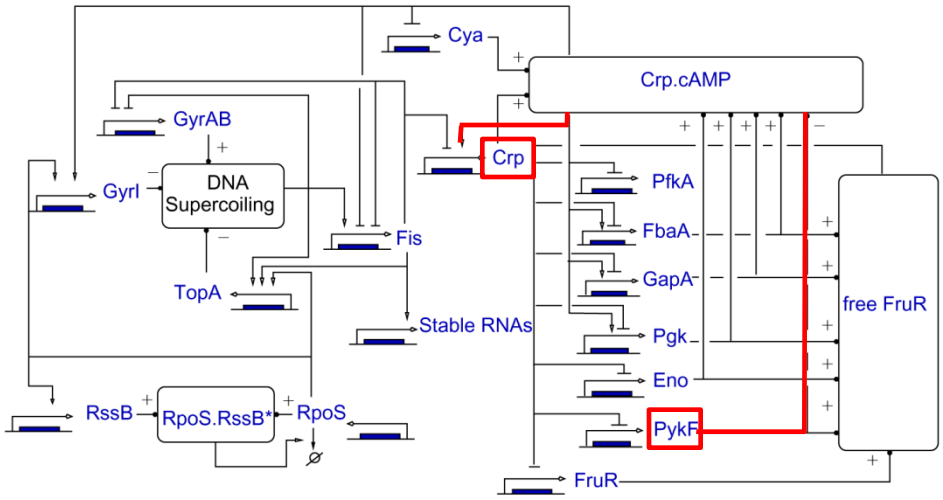


Application to *E. coli* carbon assimilation

- 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 indirect interactions mediated by metabolism
  - $\mathcal{M}^0$: 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

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	-

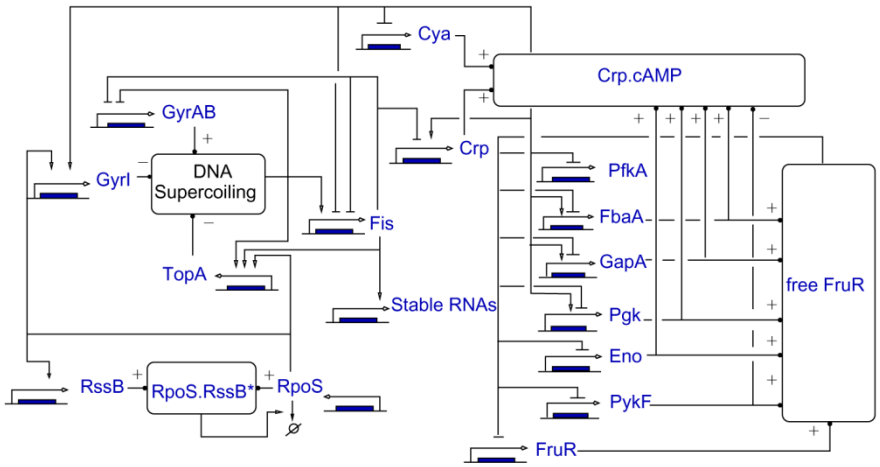
- Sign-determinedness not expected on basis of work in ecology

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

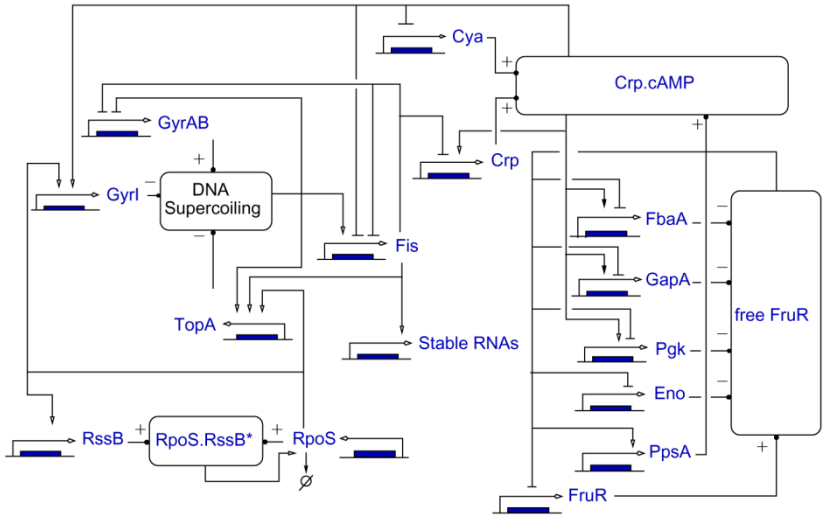
Baldazzi *et al.* (2010), *PLoS Comput. Biol.*, 6(6):e1000812

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

Conclusions

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

  Weak assumptions on time-scale hierarchies and stability

- Obtained network is at the same time **robust** and **flexible**
  - 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)

- Results on *E. coli* network raise several issues:
  - To which extent do observations carry over to other regulatory systems in bacteria and higher organisms?
  - How do indirect interactions affect dynamics of networks?

Contributors and sponsors

Grégory Batt, INRIA Rocquencourt
Valentina Baldazzi, INRA Avignon
Sara Berthoumieux, INRIA Grenoble-Rhône-Alpes
Bruno Besson, Genostar SA
Hidde de Jong, INRIA Grenoble-Rhône-Alpes
Estelle Dumas, INRIA Grenoble-Rhône-Alpes
Johannes Geiselmann, Université Joseph Fourier, Grenoble
Jean-Luc Gouzé, INRIA Sophia-Antipolis
Radu Mateescu, INRIA Grenoble-Rhône-Alpes
Pedro Monteiro, INRIA Grenoble-Rhône-Alpes/IST Lisbon
Michel Page, INRIA Grenoble-Rhône-Alpes/Université Pierre Mendès France, Grenoble
Corinne Pinel, Université Joseph Fourier, Grenoble
Caroline Ranquet, Université Joseph Fourier, Grenoble
Delphine Ropers, INRIA Grenoble-Rhône-Alpes

Courtesy Guillaume Baptist (2008)

ACI IMPBIO

Ministère de la Recherche, IMPBIO program

European Commission, FP6, NEST program

INRIA, AE program

Agence Nationale de la Recherche, BioSys program

Overview

1. Gene regulatory networks in bacteria
2. Reconstruction of gene regulatory networks from underlying biochemical reaction systems
3. Analysis of carbon assimilation network in *E. coli*
4. Conclusions