Journal of Theoretical Biology 295 (2012) 100–115

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Importance of metabolic coupling for the dynamics of gene expression following a diauxic shift in *Escherichia coli*

Valentina Baldazzi [a,*], Delphine Ropers [b], Johannes Geiselmann [b,c], Daniel Kahn [d], Hidde de Jong [b]

[a] INRA, Plantes et Systèmes de Culture Horticoles, Domaine St Paul, Site Agroparc, 84941 Avignon Cedex 9, France
[b] INRIA Grenoble – Rhône-Alpes, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, France
[c] Laboratoire Adaptation et Pathogénie des Microorganismes (CNRS UMR 5163), Université Joseph Fourier, Bâtiment Jean Roget, Faculté de Médecine-Pharmacie, La Tronche, France
[d] Laboratoire de Biométrie et Biologie Evolutive (CNRS UMR 5558), Université Lyon 1, INRA, Villeurbanne, France

ARTICLE INFO

Article history:
Received 24 May 2011
Received in revised form
7 November 2011
Accepted 8 November 2011
Available online 28 November 2011

Keywords:
Systems biology
Modeling and identification
Microbiology
Hybrid models

ABSTRACT

Gene regulatory networks consist of direct interactions, but also include indirect interactions mediated by metabolism. We investigate to which extent these indirect interactions arising from metabolic coupling influence the dynamics of the system. To this end, we build a qualitative model of the gene regulatory network controlling carbon assimilation in *Escherichia coli*, and use this model to study the changes in gene expression following a diauxic shift from glucose to acetate. In particular, we compare the relative variation in the steady-state concentrations of enzymes and transcription regulators during growth on glucose and acetate, as well as the dynamic response of gene expression to the exhaustion of glucose and the subsequent assimilation of acetate. We find significant differences between the dynamics of the system in the absence and presence of metabolic coupling. This shows that interactions arising from metabolic coupling cannot be ignored when studying the dynamics of gene regulatory networks.

1. Introduction

The reorganization of gene expression patterns in response to environmental changes is controlled by so-called gene regulatory networks, which ensure the coordinated expression of clusters of functionally related genes. The interactions in the network may be direct, as in the case of a gene coding for a transcription factor regulating the expression of another gene. Most of the time, however, regulatory interactions are indirect, *e.g.*, when a gene encodes an enzyme producing a transcriptional effector (Brazhnik et al., 2002). The latter interactions arise from the fact that gene regulation is embedded in a complex, multi-level system that tightly integrates gene expression with metabolism. We call the occurrence of indirect interactions between enzymes and genes, mediated by metabolism, *metabolic coupling*.

In previous work (Baldazzi et al., 2010), we showed how indirect interactions arising from metabolic coupling can be derived from a model of the underlying biochemical reaction network. We applied this approach to the carbon assimilation network in *Escherichia coli* investigating how the structural properties of the network are modified by the inclusion of metabolic interactions. Our results showed that the derived gene regulatory network is densely connected, contrary to what is usually assumed. Moreover, we found that the signs of the indirect interactions are largely fixed by the direction of metabolic fluxes, independently of specific parameter values and rate laws, and that a change in flux direction may invert the sign of indirect interactions. This leads to a feedback structure that is at the same time robust to changes in the kinetic properties of enzymes and that has the flexibility to accommodate radical changes in the environment.

It remains an open question, however, to which extent the indirect interactions induced by metabolic coupling influence the dynamics of the system. This is a key issue for understanding the relative contributions of the regulation of gene expression and metabolism during the adaptation of the cell to changes in its environment. Indirect interactions could be essential in shaping the response of the cell, giving it the required flexibility to adapt to external perturbations. However, one could also argue that indirect interactions only have a fine-tuning effect, simply refining the gene expression levels at which the system would stabilize without metabolic coupling.

In order to decide between these two hypotheses, and obtain a clearer view of the role of metabolic coupling in the adaptation of gene expression, we use the topology obtained in our previous work (Baldazzi et al., 2010) to build a dynamic model of the gene regulatory network controlling carbon assimilation in *E. coli*, and use this model to study the changes in gene expression following

* Corresponding author. Tel.: +33 4 32722447; fax: +33 4 32722432.
E-mail address: valentina.baldazzi@avignon.inra.fr (V. Baldazzi).

0022-5193/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2011.11.010

a diauxic shift from glucose to acetate. More specifically, we develop a qualitative model using piecewise-linear (PL) differential equations that allows us to encode the regulatory logic of the system in a simple way (Batt et al., 2008; Glass and Kauffman, 1973). Even though good quantitative models of carbon metabolism in *E. coli* exist in the literature (Bettenbrock et al., 2005; Chassagnole et al., 2002; Kotte et al., 2010), not much is known about precise mechanisms and parameter values for gene regulation. Moreover, qualitative models are an appropriate tool for analyzing if metabolic coupling can induce *major* changes in the gene expression dynamics, *i.e.* will not only have an effect on quantitative but also qualitative properties of the system dynamics.

We build several qualitative models, corresponding to a network topology including all, some, or none of the indirect interactions, respectively. The dynamical properties of the models are analyzed and compared with available experimental data. In particular, we compare the steady-state concentrations of enzymes and transcription regulators during growth on glucose and acetate, as well as the dynamic response of gene expression to the exhaustion of glucose and the subsequent assimilation of acetate. We find significant differences between the dynamics of the system in the absence and presence of metabolic coupling. This confirms that indirect interactions are essential for correctly reproducing the observed adaptation of gene expression to a change in carbon source. More specifically, indirect interactions involving the (in)activation of the regulators Crp and FruR by the flux-signaling metabolites cAMP and FBP, respectively, are key for the adjustment of gene expression levels when changing the carbon source from glucose to acetate. Our work thus underlines the importance of metabolic coupling in gene regulatory networks, and shows that such indirect interactions cannot be neglected when studying the adaptation of an organism to changes in its environment.

2. Methods

2.1. Metabolic coupling in gene regulatory networks

In Baldazzi et al. (2010) a mathematical method is proposed, based on a combination of time-scale approximations and sensitivity criteria from metabolic control analysis, to reconstruct the indirect interactions between genes from a model of the underlying network of biochemical reactions. In the following, we briefly summarize the basic principles of this approach.

We start by building a stoichiometry model of the network. The model takes the form of a system of ordinary differential equations (ODEs) and describes the rate of change of the concentrations of the different molecular species in the network

$$\dot{x} = N v(x), \quad x(0) = x_0 \tag{1}$$

where $x \in \mathbb{R}^n_+$ denotes the vector of concentrations and $v : \mathbb{R}^n_+ \to \mathbb{R}^q$ the vector of reaction rates. $N \in \mathbb{Z}^{n\times q}$ is a stoichiometry matrix. In the presence of conserved quantities, N is the reduced stoichiometry matrix (Reder, 1988). Contrary to kinetic models, the rate laws are not explicitly specified, but only the functional dependency of the reaction rates on specific molecular species are given. We do not develop the rate laws, because only the signs of the partial derivatives are used for reconstructing the indirect interactions (see below).

As a first step, the model can be simplified by making the *quasi-steady-state* (*QSS*) *approximation* (Heinrich and Schuster, 1996). Two different time-scales are distinguished, one corresponding to the slow processes (protein synthesis and degradation) and one to the fast processes (complex formation and enzymatic reactions). Based on time-scale separation, the original model can be rewritten as two distinct subsystems

$$\dot{x}^s = N^s v^s(x^s, x^f), \quad x^s(0) = x^s_0 \tag{2}$$

$$\dot{x}^f = N^f v^f(x^s, x^f), \quad x^f(0) = x^f_0 \tag{3}$$

where $x^s \in \mathbb{R}^m_+$ and $x^f \in \mathbb{R}^{n-m}_+$ are the vectors of slow and fast variables, respectively, and N^s, N^f and v^s, v^f the corresponding stoichiometry matrices and rate vectors. The slow variables describe total concentrations of the proteins, and the fast variables concentrations of metabolites and complexes of proteins and signaling molecules.

The QSS approximation makes the assumption that at the time-scale of the slow processes the fast part of the system is at steady state, instantly adapting to the dynamics of the slow variables, *i.e.* $N^f v^f(x^s, x^f) = 0$ (see Heinrich and Schuster, 1996; Khalil, 2001 for the conditions under which the QSS approximation is valid). This means that, after an initial transient, the dynamics of the fast system can be well approximated by an algebraic function of the slow variables, if such a function can be found: $x^f = g(x^s)$, $g : \mathbb{R}^m_+ \to \mathbb{R}^{n-m}_+$. The resulting reduced system, at the slow time-scale, takes the following form:

$$\dot{x}^s = N^s v^s(x^s, g(x^s)) \tag{4}$$

The interest of this model is that it represents the structure of direct and indirect interactions between the slow variables, that is, the structure of the gene regulatory network. Since we have not specified the rate laws, and in general these are complex nonlinear functions of the concentration variables, it is not possible to obtain a closed-form expression for the function g. We therefore follow another strategy to characterize the indirect interactions between the slow variables, by studying the Jacobian matrix $\mathcal{J} \in \mathbb{R}^m \times \mathbb{R}^m$ of the system in Eq. (4)

$$\mathcal{J} = \frac{\partial \dot{x}^s}{\partial x^s} = N^s \frac{\partial v^s(x^s, g(x^s))}{\partial x^s} + N^s \frac{\partial v^s(x^s, g(x^s))}{\partial x^f} \frac{\partial g(x^s)}{\partial x^s} \tag{5}$$

The Jacobian matrix includes the direct effect of each slow variable on the others (first term) and the indirect effect via coupling through the fast system (second term). It accounts for direct regulation of gene expression by transcription factors as well as indirect regulation through metabolic intermediates. Applying the implicit function theorem to the QSS equation, $\partial g(x^s)/\partial x^s$ can be further developed as

$$\frac{\partial g(x^s)}{\partial x^s} = -M^{-1} N^f \frac{\partial v^f(x^s, x^f)}{\partial x^s} \tag{6}$$

where $M = N^f \partial v^f(x^f, x^s)/\partial x^f$ is the Jacobian matrix of the fast system. Eq. (6) describes the response of the fast system around its steady state to changes in the slow variables. $-M^{-1}N^f$ is known as the matrix of (non-normalized) concentration control coefficients in metabolic control analysis (Heinrich and Schuster, 1996).

The symbolic computation of the elements of $\mathcal{J}$ gives rise to complex expressions in terms of elementary partial derivatives of the reaction rates with respect to the concentration variables, called (non-normalized) elasticities (Heinrich and Schuster, 1996). Even in the absence of a precise specification of kinetic rate laws and parameter values, most elasticities have a well-defined sign, given a convention on the positive flux direction. Together with constraints imposed by the stability of the steady state of the fast system, the signs of the elasticities are used to evaluate the signs of the elements of the Jacobian matrix. Notice that the sign of the elasticity of a reversible reaction with respect to its enzyme concentration can be positive or negative, depending on the flux direction. As a consequence, the signs of the

elements of $\mathcal{J}$ depend on the growth conditions, as the latter may influence the flux directions, and thus the signs of (some of) the elasticities. In particular, following a sudden change in the substrate availability (for instance, during a diauxie), some of the indirect interactions may invert their sign, thus modifying the regulation of their target genes. This is an important difference between direct and indirect interactions: whereas transcriptional interaction are fixed in sign and direction, interactions mediated by metabolism can change according to the environmental conditions, allowing for a partial rewiring of the network.

Fig. 2 illustrates the application of the method to a small example, describing the synthesis of cyclic AMP (cAMP) and its effect on gene expression in response to glucose depletion. Panels A and B show the reactions playing a role in this subsystem and the stoichiometry model obtained by separating the time-scales for the expression of the genes *fis*, *cya*, and *crp* on one hand, and for cAMP synthesis and the formation of the complex Crp · cAMP on the other. A graphical representation of the reduced network, displaying the regulatory interactions between the slow variables of the system, is shown in panel C. Some of the interactions are direct, such as those between Fis and its targets. However, many of the interactions are indirect, in particular the influence of Crp, Cya, and external glucose on the metabolic coupling species Crp · cAMP, which in turn controls the expression of the genes in the network. Notice that in determining the signs of the Jacobian elements (panel D), we excluded the direct effect of a slow variable on itself due to nonspecific protein degradation and growth dilution. The results for this example are intuitive and easy to derive, but this is no longer the case when large networks like that shown in Fig. 1 are studied. The expressions for the elements of the Jacobian matrix become complex and computer algebra tools may be necessary to resolve their signs (Baldazzi et al., 2010).

2.2. Including metabolic coupling in dynamic models of gene regulatory networks

Modeling the dynamics of a gene regulatory network amounts to specifying the right-hand side of Eq. (4). However, recall from Section 2.1 that we have not fully specified the rate laws. Analysis of Eq. (5), defining the Jacobian matrix $\mathcal{J}$ and thus the structure of the gene regulatory network, provides information that can be exploited for qualitatively modeling the network dynamics. By making the common assumption that gene regulatory interactions are monotone (at least over a large concentration range), and sigmoidal (de Jong, 2002; Polynikis et al., 2009), we show how approximate functions for v^s can be defined for the qualitative analysis of the network dynamics.

Consider first the case of direct interactions, and more specifically the case $v_i^s = v_i^s(x_j^s)$, for some reaction i and slow variable x_j^s. Typically, v_i^s represents a protein synthesis rate and x_j^s the (total) concentration of a transcription factor. Notice that the existence of a direct interaction implies, with Eq. (5), that $\partial v_i^s/\partial x_j^s$ is non-zero. *Monotonicity* of v_i^s puts the further restriction on $\partial v_i^s/\partial x_j^s$ that its sign is either positive or negative. More precisely, if the sign of the (i,j)th element of $\mathcal{J}$ is positive (negative), then the transcription factor is an activator (inhibitor). *Sigmoidality* of v_i^s means that $\partial^2 v_i^s/(\partial x_j^s)^2$ has a (single) zero. A simple function that is both monotone and sigmoidal is the Hill function (Plahte et al., 1998; Wittmann et al., 2009; Ironi and Panzeri, 2009)

$$v_i^s(x_j^s) = \kappa_i h^+(x_j^s) = \kappa_i \frac{(x_j^s)^n}{(x_j^s)^n + \theta_j^n} \quad \text{or } v_i^s(x_j^s) = \kappa_i h^-(x_j^s) = \kappa_i (1 - h^+(x_j^s)) \tag{7}$$

where θ_j is a so-called threshold concentration for x_j^s, and n a cooperativity constant that determines the steepness of the sigmoidal functions. When several transcription regulators

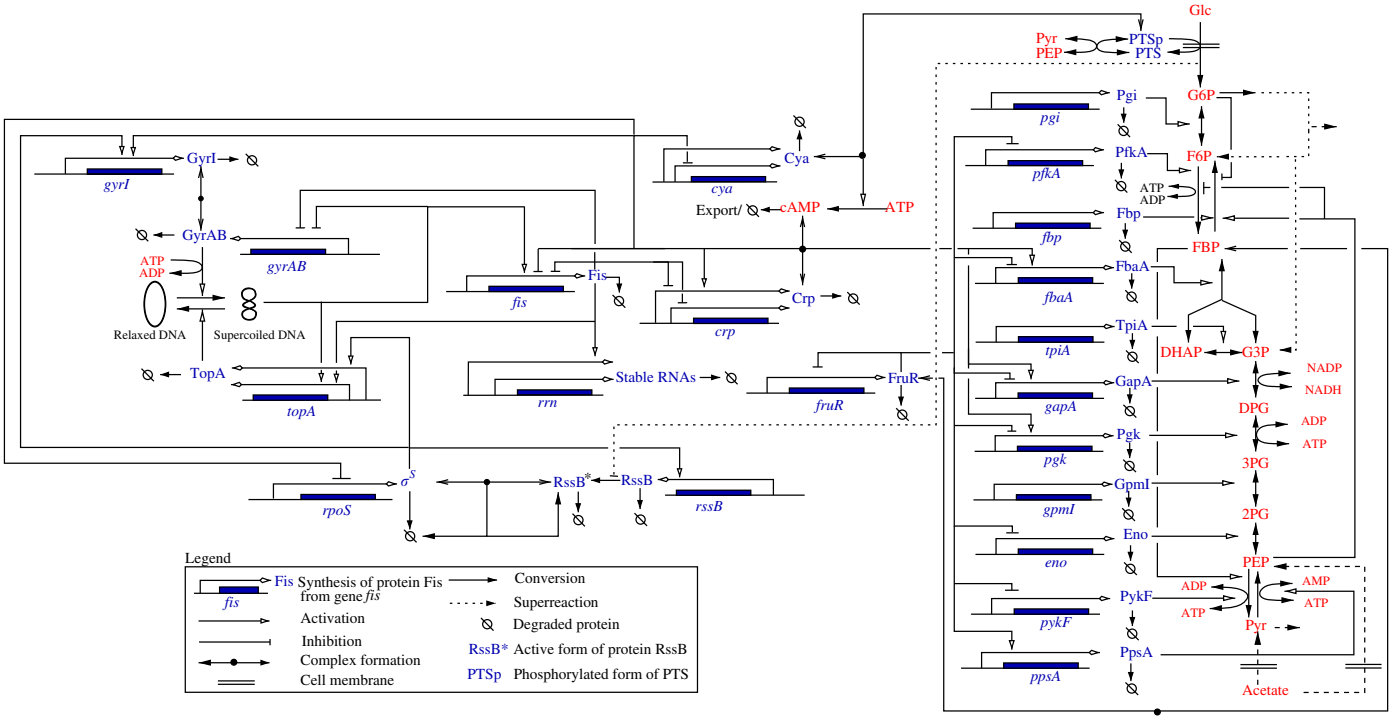


Fig. 1. Carbon assimilation network in *E. coli*, consisting of the glycolysis and gluconeogenesis pathways and their genetic and metabolic regulation (Baldazzi et al., 2010). Metabolites are indicated in red, enzymes and transcription factors in blue. The global regulators (left in the figure) control the expression of some the enzymes, thus affecting metabolism (right in the figure). Some signaling metabolites (cAMP and FBP) have an influence on the activity of specific transcription factors thus giving rise to an indirect control of gene expression through metabolic coupling. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

directly regulate the expression of a gene, that is, $v_i^s = v_i^s(\ldots,x_{j-1}^s,x_j^s,x_{j+1}^s,\ldots)$, the possibly combinatorial effect of these regulators can be defined by means of sums, subtractions, and products of Hill functions (Plahte et al., 1998; Wittmann et al., 2009).

Can metabolic coupling also be modeled by expressions of Hill functions? In order to investigate this, consider the case of an indirect interaction $v_i^s = v_i^s(x_k^f)$ with $x_k^f = g_k(x_j^s)$, for some fast variable x_k^f. For example, x_j^s might represent an enzyme concentration and x_k^f the concentration of the complex of a transcription factor and a signaling metabolite under the control of x_j^s. From Eq. (5) it follows that, in the case of an indirect interaction, $\partial v_i^s/\partial x_k^f$ and $\partial g_k(x_j^s)/\partial x_j^s$ are non-zero. The above-mentioned assumptions of monotonicity and sigmoidality further imply that $\partial v_i^s/\partial x_k^f$ has a definite (positive or negative) sign and that $\partial^2 v_i^s/(\partial x_k^f)^2$ crosses zero (once), respectively. By Eq. (5) and the chain rule, the partial derivatives characterizing the indirect effect of the slow variables can be expressed as follows:

$$\frac{\partial v_i^s}{\partial x_j^s} = \frac{\partial v_i^s}{\partial x_k^f}\frac{\partial g_k(x_j^s)}{\partial x_j^s} \tag{8}$$

$$\frac{\partial^2 v_i^s}{(\partial x_j^s)^2} = \frac{\partial^2 v_i^s}{(\partial x_k^f)^2}\left(\frac{\partial g_k(x_j^s)}{\partial x_j^s}\right)^2 + \frac{\partial v_i^s}{\partial x_k^f}\frac{\partial^2 g_k(x_j^s)}{(\partial x_j^s)^2} \tag{9}$$

From these expressions we can derive conditions for the monotonicity and sigmoidality of indirect interactions. For v_i^s to be monotone, $\partial g_k(x_j^s)/\partial x_j^s$ needs to have a definite sign, at least over a concentration range of interest. This means that the relevant concentration control coefficient of the fast system needs to be sign-determined. As to the sigmoidal dependence of v_i^s on x_j^s, the first term in Eq. (9) is guaranteed to cross zero, since $\partial^2 v_i^s/(\partial x_k^f)^2$ does and $(\partial g_k(x_j^s)/\partial x_j^s)^2$ is a positive multiplicative factor. In order for the entire expression to have a sigmoidal shape, the addition of the second term should preserve the zero-crossing. This occurs, for example, when over the concentration range of interest the fast variables are linear functions of the slow variables, in which case $\partial^2 g_k(x_j^s)/(\partial x_j^s)^2$ equals 0. The above argument can be generalized to the case that g is a multivariate function, *i.e.* $x_k^f = g_k(\ldots,x_{j-1}^s,x_j^s,x_{j+1}^s,\ldots)$.

In order to illustrate these conditions, we consider the subnetwork describing the activation of the transcription factor Crp, as a typical example of metabolic coupling. This includes, in particular, the control of the *crp* synthesis rate by Crp and Cya, through the effect of the latter on the concentration of Crp · cAMP when glucose is absent (Fig. 3). We develop a kinetic model of this subnetwork, using parameter values derived from the experimental literature (Ropers et al., 2011), separate the fast and slow variables as described in Fig. 2, and derive an analytical expression for the function g (Appendix A). As shown in Fig. 3A, g is an increasing function of x_{Crp} and x_{Cya}, which means that the concentration control coefficients $\partial g(x_{Crp},x_{Cya},u_{Glc})/\partial x_{Crp}$ and $\partial g(x_{Crp},x_{Cya},u_{Glc})/\partial x_{Cya}$ are both positive. Together with the fact that $v_1(x_{Crp\cdot cAMP})$ is a positive Hill function, this yields the monotonicity of $\partial v_1(x_{Crp\cdot cAMP})/\partial x_{Crp}$ and $\partial v_1(x_{Crp\cdot cAMP})/\partial x_{Cya}$. Is v_1 also a sigmoidal function of x_{Crp} and x_{Cya}? Notice that g is linear in x_{Crp}. As a consequence, following the discussion in the previous paragraph, $\partial v_1(x_{Crp\cdot cAMP})/\partial x_{Crp}$ is sigmoidal. Moreover, g is a saturating function of x_{Cya} with the Michaelian constant $K = 0.66\times10^{-8}$ (Appendix A). This means that in a large part of the concentration range over which x_{Cya} varies, the dependence of g on x_{Cya} is well-approximated by a linear function, and $\partial v_1(x_{Crp\cdot cAMP})/\partial x_{Cya}$ also takes a sigmoidal shape. The resulting monotonicity and sigmoidality of the two indirect interactions give rise to the combined response surface of Fig. 3B, which shows the dependence of the *crp* synthesis rate $v_1(x_{Crp\cdot cAMP})$ on the (total) concentrations of Crp and Cya (x_{Crp} and x_{Cya}). The response surface is conveniently modeled by a multiplication of Hill functions.[1]

The above analysis shows how sigmoidal interactions naturally arise in a typical example of metabolic coupling. While feasible for a small subnetwork, this can no longer be done for large networks like the one shown in Fig. 1, due to the difficulty of explicitly solving for g. In what follows we assume that just like direct interactions, indirect interactions are generally monotone and sigmoidal, and can be modeled by combinations of Hill functions. Numerical simulation studies have shown that, in comparison with classical kinetic models with fully specified rate laws, the approximations are generally able to preserve the (qualitative) dynamics of the network quite well (Ropers et al., 2011; Davidich and Bornholdt, 2008). In Discussion section we return to this issue.

2.3. Qualitative modeling of gene regulatory networks

In order to study the role and the relevance of the indirect interactions for the dynamics of the system, in particular its response to environmental changes, we adopt a qualitative modeling approach. More specifically, we use *piecewise-linear* (*PL*) *differential equations* to model regulatory interactions between genes and tools developed for the qualitative analysis of these models. The PL models, introduced by Glass and Kauffman (1973), are based on the use of step-function approximations of the Hill functions describing direct and indirect regulatory interactions (Section 2.2). Here we informally summarize the basic notions underlying the modeling of gene regulatory networks by means of PL models, while mathematical details can be found elsewhere (Batt et al., 2008; de Jong et al., 2004a).

PL models are directly obtained from the dynamical models discussed in the previous section by replacing the Hill functions in the rate laws in the right-hand side of Eq. (4) by step functions

$$s^+(x_j^s,\theta_j) = \begin{cases} 1 & \text{if } x_j^s > \theta_j \\ 0 & \text{if } x_j^s < \theta_j \end{cases}$$

$$s^-(x_j^s,\theta_j) = 1 - s^+(x_j^s,\theta_j)$$

In comparison with Hill functions, step functions discontinuously change their value at the threshold concentration. However, notice that this approximation preserves the switch-like and combinatorial character of gene regulation. On the formal level, the PL models are related to logical models of regulatory networks (see Fauré and Thieffry, 2009; Morris et al., 2010 for reviews on logical models and Snoussi, 1989; Wittmann et al., 2009 for a discussion of the relation between logical and PL models).

An advantage of the use of step functions is that they facilitate the analysis of the qualitative dynamics of the PL models, even in higher-dimensional systems (Kappler et al., 2003). The thresholds of the concentration variables define a rectangular partition of the state space, such that in every region not located on a threshold, the PL model reduces to an analytically solvable system of differential equations. Moreover, in every such region the derivatives (trends) of the concentration variables have a determinate sign, which is shown to be invariant for rather weak constraints on the parameter values. The definition of these constraints can generally be inferred from available data in the experimental

[1] Notice that if the parameter K had a much smaller value, the effect of Cya would be saturating over most of the concentration range and the indirect interaction would only involve Crp.

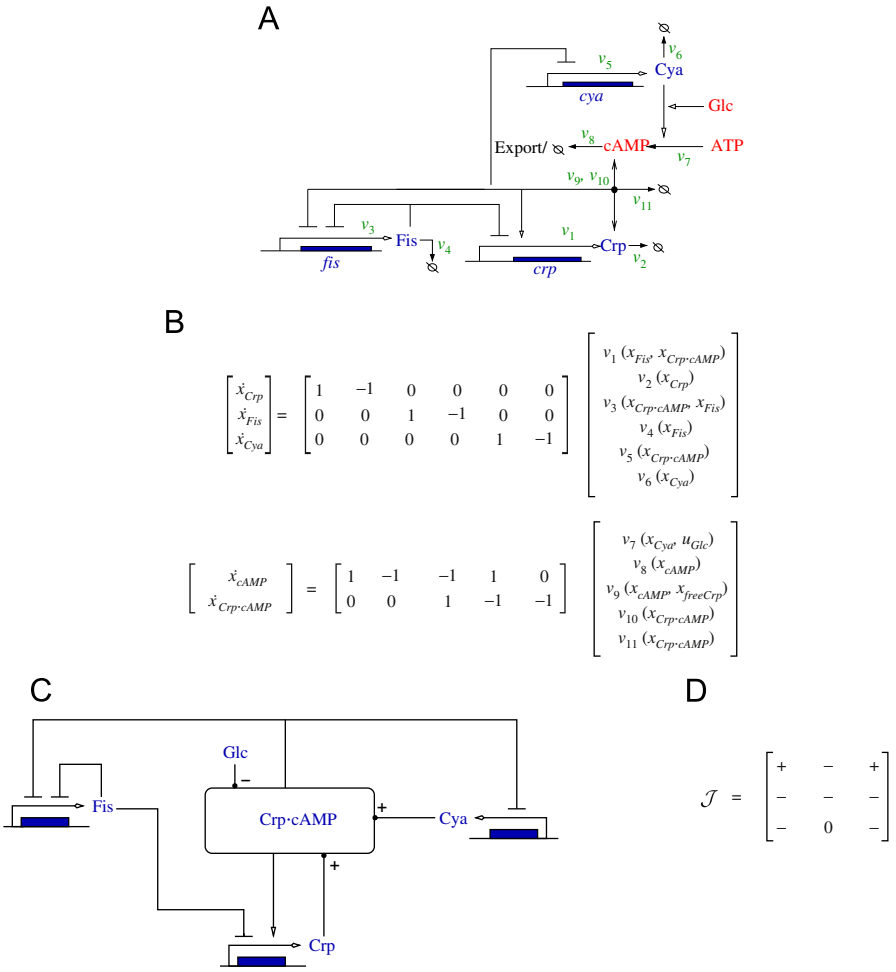


Fig. 2. Example of the derivation of the gene regulatory network involved in the activation of the transcription regulator Crp. A: Network diagram (see Fig. 1 for the legend). The reactions correspond to protein synthesis (v_1, v_3, v_5), protein degradation (v_2, v_4, v_6, v_{11}), enzymatic reactions (v_7), association and dissociation reactions (v_9, v_{10}), and export (v_8). The synthesis of the second messenger cAMP, and its regulation through the activation of adenylate cyclase Cya under conditions of glucose depletion, are summarized in a single reaction (v_7). Proteins are shown in red, metabolites in blue, and reactions in green. B: Stoichiometry model of the network, with an explicit separation of time-scales following the distinction between slow reactions (protein synthesis and degradation) and fast reactions (enzymatic reactions and complex formation). The slow variables are the total concentrations of Crp (x_{Crp}), Fis (x_{Fis}), and Cya (x_{Cya}), while the fast variables are the concentrations of cAMP (x_{cAMP}) and Crp · cAMP ($x_{Crp\cdot cAMP}$). The glucose concentration is an input variable (u_{Glc}). For simplicity, ATP and external cAMP are not explicitly modeled in reactions v_7 and v_8. C: Derived gene regulatory network showing direct and indirect interactions between genes and their products, obtained from the stoichiometry model in *B*. The concentrations of Crp, Cya, and glucose influence the concentration of the fast coupling species Crp · cAMP, which in turn affects the expression of the three genes in the network (*fis*, *cya*, and *crp*). In addition, the diagram shows direct interactions, such as the regulation of the expression of *fis* and *crp* by Fis. D: Jacobian matrix showing the sign of the derived regulatory interactions between the slow variables $[x_{Crp}, x_{Fis}, x_{Cya}]'$ of the system. Nonspecific protein degradation and growth dilution have been excluded from the computation of indirect regulatory interactions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

literature or by intuitive reasoning, even in the absence of quantitative information on parameter values.

The dynamics of PL models is conveniently described by a so-called *state transition graph*. The states in the graph correspond to the regions in the state space. The transitions between the states arise from solutions of the PL model that enter one region from another. Each transition thus corresponds to a discrete event, namely the crossing of a threshold by one or more concentration variables, possibly entailing a change in the derivative of these variables. The state transition graph can be inferred from the model using the above-mentioned constraints on parameter values rather than exact numerical values. Dynamic properties of the network, such as the stability of steady states, the existence of limit cycles, and the reachability of certain states of interest from a given initial state can be formulated in terms of the graph topology. Moreover, large graphs can be conveniently and efficiently analyzed using model-checking tools (Monteiro et al., 2008).

Genetic Network Analyzer (*GNA*) is a computer tool specifically developed for the analysis of gene regulatory networks by means of PL models (de Jong et al., 2003). It allows the user to define a regulatory network, build a model of this network, determine the steady states of the system, generate a state transition graph starting from an initial state (qualitative simulation of the network dynamics), and analyze the latter graph using model-checking tools. GNA has been used for the analysis of a number of bacterial regulatory networks (de Jong et al., 2004b; Usseglio Viretta and Fussenegger, 2004; Sepulchre et al., 2007; Ropers et al., 2006).

3. Results

3.1. Metabolic coupling in network controlling carbon assimilation in E. coli

Fig. 1 shows a part of the network involved in the assimilation of carbon sources in *E. coli*. It consists of the glycolytic and gluconeogenic pathways, the phosphotransferase system, the

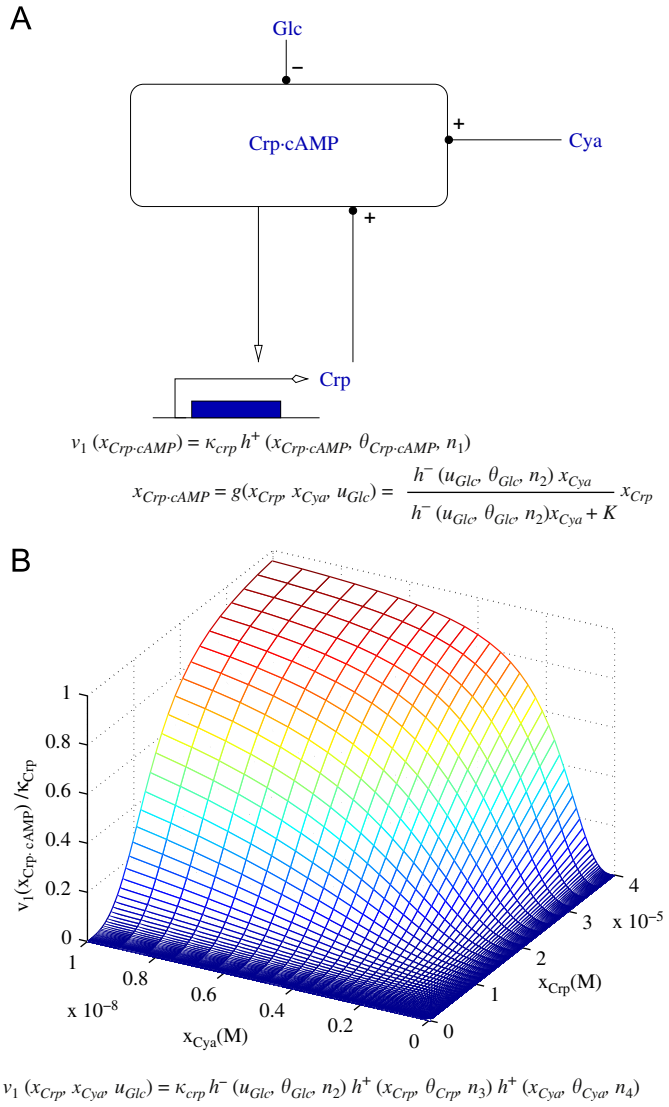


Fig. 3. Example of modeling an indirect interaction in the Crp activation network of Fig. 2. A: Regulation of *crp* expression by Crp, Cya, and glucose through metabolic coupling, with an explicit model of the indirect interaction. The expression for v_1, representing the synthesis rate of *crp*, is obtained by solving the QSS equation in a kinetic model of the Crp activation network (Appendix A). $\theta_{Crp\cdot cAMP}$ and θ_{Glc} denote threshold concentrations, κ_{crp} a synthesis parameter, n_1 and n_2 cooperativity constants, and K a (lumped) kinetic parameter. B: Plot of $v_1(x_{Crp\cdot cAMP})/\kappa_{crp}$ as a function of x_{Crp} and x_{Cya} when glucose is absent ($h^-(u_{Glc},\theta_{Glc},n_2) \approx 1$). The resulting sigmoidal surface can be approximated by fitting a multiplicative expression of Hill functions, where θ_{Crp} and θ_{Cya} are the threshold concentrations, and n_3 and n_4 the cooperativity constants.

genes coding for the enzymes, their key transcriptional regulators (Crp, FruR, and Fis), as well as other global regulators of transcription (RpoS, DNA supercoiling, etc.). We distinguish two different growth conditions: growth on the carbon-rich substrate glucose (glycolysis) and growth on the carbon-poor substrate acetate (gluconeogenesis). Since glucose and acetate are utilized by distinct metabolic pathways, the flux distributions during glycolysis and gluconeogenesis are significantly different. A diauxic shift from glucose to acetate notably leads to the inversion of glycolytic fluxes (Oh et al., 2002).

The network of biochemical reactions in Fig. 1 can be transformed into a network of direct and indirect regulatory interactions following the method outlined in Section 2.1. To this end, we use a stoichiometry model of the network describing the dependence of the reaction rates on the substrates, products, enzymes, transcription factors, and other molecular species (Baldazzi et al., 2010). In the case of reversible reactions, the signs of the elasticities, and therefore the signs of the indirect interactions, depend on the directions of the metabolic fluxes. We therefore consider the cases of glycolysis and gluconeogenesis separately.

Fig. 4 shows the resulting network for glycolysis (panel *A*) and gluconeogenesis (panel *B*), respectively called $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$. The boxes represent so-called *coupling species*, which channel indirect influences on gene expression through metabolism. For instance, the concentration of Crp · cAMP is modulated by Cya, Crp, and the glycolytic enzymes. Via this route, the glycolytic enzymes indirectly regulate targets of Crp · cAMP. Apart from Crp · cAMP, the network includes three other coupling species: free FruR, RpoS · RssB, and DNA supercoiling. A positive or negative sign indicates the effect of slow variables on the concentration of each coupling species. Notice that, in the case of the *E. coli* network, these signs have been determined based solely on local kinetic information (signs of elasticities) and stability assumptions, without fully specifying the rate laws. This is a (somewhat surprising) property of the *E. coli* network, which may not hold in general though (see Baldazzi et al., 2010 for a discussion on sign-determinedness and its theoretical conditions).

The comparison of the two networks in Fig. 4 shows that some indirect interactions appear in one growth condition and disappear in the other, such as the control of the concentration of free FruR, and its targets, by PykF and PpsA. Moreover, the same interaction may have an opposite sign depending on growth conditions. For instance the effect of GapA on the concentration of free FruR is positive during growth on glucose, and negative when acetate is used as carbon source. Metabolic coupling thus allows the structure of regulatory interactions to be rewired by changes in environmental conditions, potentially enhancing the adaptive capacity of the system (Baldazzi et al., 2010).

The networks in Fig. 4 have been simplified by leaving out some of the glycolytic enzymes. We notably omit enzymes whose expression does not significantly change between growth on glucose and acetate (Oh et al., 2002), enzymes that are constitutively expressed, and enzymes that do not regulate any other gene in the network. Moreover, we keep only a single representative of each group of enzymes with common regulators. This leaves us with GapA, PpsA, and PykF, whose expression patterns are known to characterize the two modes of carbon catabolism considered here (Oh et al., 2002). Notice that these simplifications of the network tend to reduce the number of indirect interactions, and thus their influence on the network dynamics. If an effect of the indirect interactions on the qualitative dynamics is visible in the simplified networks, this will *a fortiori* be the case in the extended networks as well.

In order to precisely assess the effect of the indirect interactions on the network dynamics, we additionally define a number of reference networks in which all indirect interactions mediated by the coupling species Crp · cAMP and free FruR are systematically removed. The reference networks are called $\mathcal{M}^0_{glyco/\text{Crp·cAMP}}$, $\mathcal{M}^0_{neo/\text{Crp·cAMP}}$, $\mathcal{M}^0_{glyco/\text{free FruR}}$, and $\mathcal{M}^0_{neo/\text{free FruR}}$. As an extreme case, a network consisting of transcriptional regulatory interactions only is also considered. This last network is simply called $\mathcal{M}^0$, and is the same for glycolysis and gluconeogenesis. The topology of $\mathcal{M}^0$ networks is shown in Fig. 5 (glycolysis) and in Section S2 of the Supplemental Information (SI) of this paper (gluconeogenesis).

When omitting indirect interactions, we make an exception for the effect of transcription regulators, even when their activity is controlled by a metabolic effector (*e.g.*, cAMP for Crp). Although not direct in the sense of Eq. (5), including these interactions in the reference networks has the advantage of complying with the usual notion of transcriptional regulatory networks employed in curated databases (Keseler et al., 2011). As above, the approach is

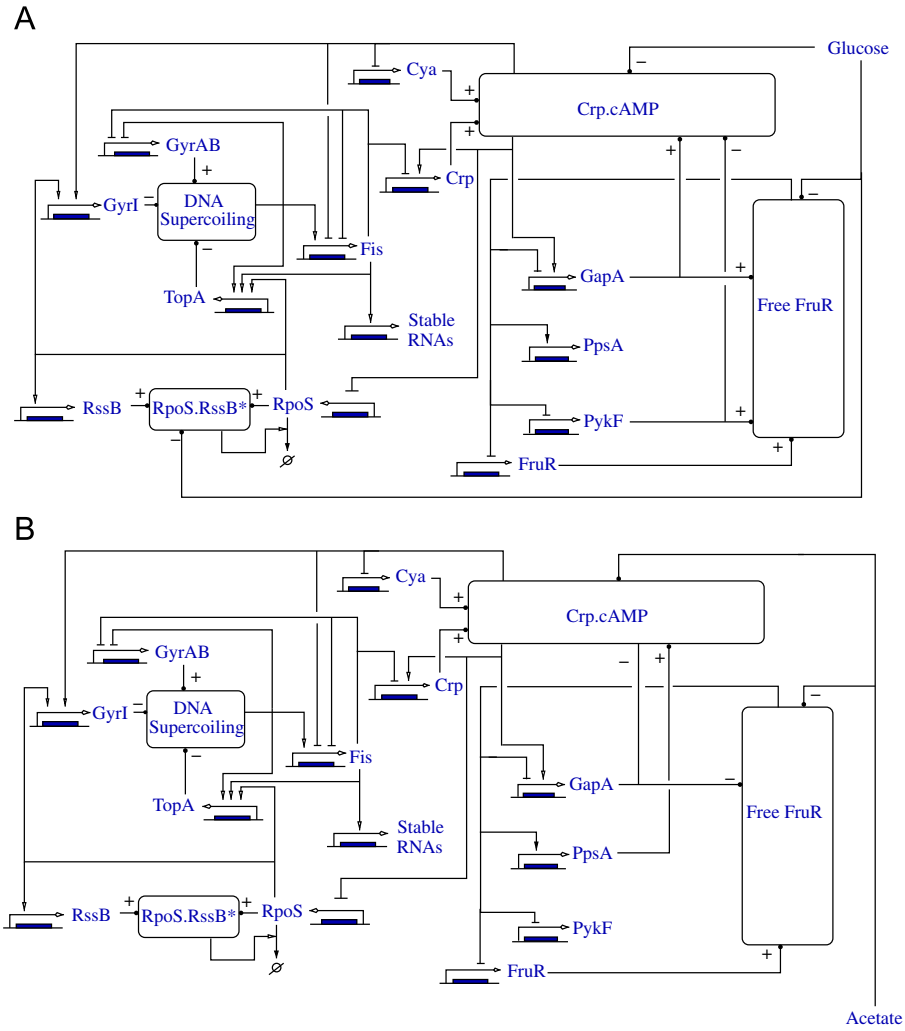


Fig. 4. Gene regulatory networks for the glycolytic (A) and gluconeogenic case (B), obtained from the network of biochemical reactions shown in Fig. 1. The networks are denoted by $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$, respectively. The boxes indicate the coupling species, here Crp · cAMP, free FruR, RpoS · RssB, and DNA supercoiling.

conservative, in the sense that any differences in the qualitative dynamics of the networks with and without metabolic coupling will be exacerbated under a stricter definition of (in)direct interactions.

3.2. *Qualitative modeling of carbon assimilation network*

The networks shown in Figs. 4 and 5 are gene regulatory networks, in the sense that they describe the (direct or indirect) influence of the products of one gene on the expression of another gene (Bolouri, 2008). The interactions can be positive (activation) or negative (inhibition). A variety of methods exist to model and analyze the dynamics of gene regulatory networks (Karlebach and Shamir, 2008). Here we focus on the use of qualitative models, which provide a coarse-grained description of the dynamics of gene expression, in the sense that they do not explicitly specify the biochemical mechanisms. However, they include the logic of gene regulation and allow different expression levels of the genes to be distinguished. They are interesting in their own right, as a way to capture in a simple manner the complex dynamics of a large regulatory network without quantitative data (Morris et al., 2010; de Jong and Ropers, 2006). However, they can also be used as a first step to orient the development of more detailed quantitative ODE models.

More specifically, we use so-called piecewise-linear (PL) differential equations, a formalism originally introduced by Glass and Kauffman (1973) for describing gene regulatory networks (Section 2.3). PL models allow the qualitative dynamics of large regulatory networks to be analyzed using inequality constraints on parameters rather than exact numerical values. The network structures in Figs. 4 and 5 are transformed into qualitative PL models by defining the differential equations for each of the gene products (protein concentrations) as well as parameter constraints that characterize the system dynamics. Intuitively, this amounts to specifying the logic of the regulation of the genes and the relative strength of the different interactions regulating the expression of a gene. The PL models for the $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$ networks extend an existing qualitative model of the network of global regulators in *E. coli* (Ropers et al., 2006), by adding metabolic coupling induced by carbon catabolism. The following principles have guided the modeling of the networks:

- The existence of a single stable attractor in each growth condition (glucose and acetate), and the absence of oscillations in the normal range of operation.
- A parametrization of the models consistent with that of the previous models and with available literature data.

The resulting models include 13 variables and require the specification of 44 or 49 inequality constraints, for the

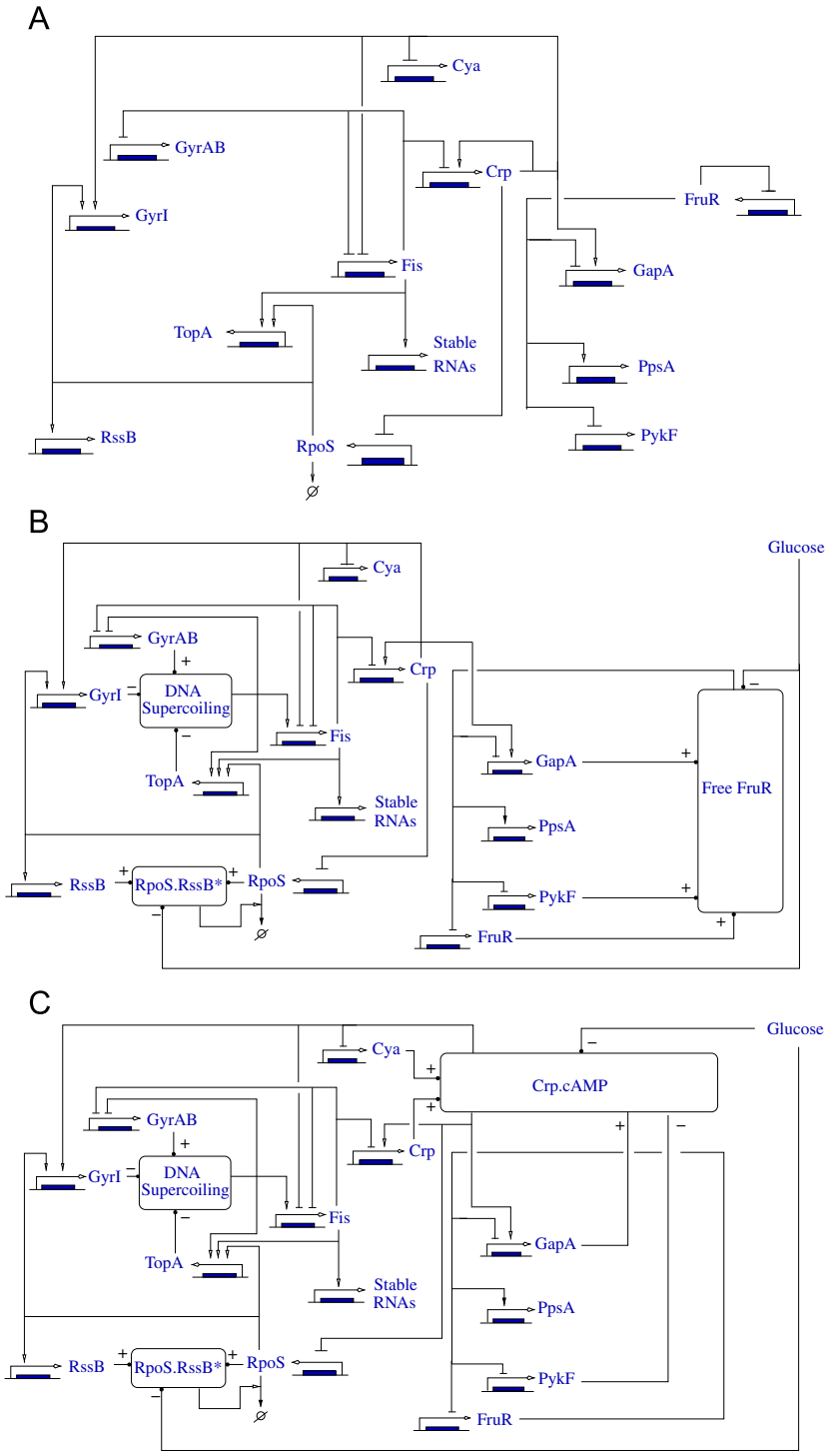


Fig. 5. Gene regulatory networks obtained by omitting some of the indirect interactions in Fig. 4A, during growth on glucose. A: Network including direct transcriptional interactions only (model $\mathcal{M}^0$), B: network obtained by omitting interactions mediated by Crp · cAMP (model $\mathcal{M}^0_{glyco/\text{Crp-cAMP}}$), C: network obtained by omitting interactions mediated by free FruR (model $\mathcal{M}^0_{glyco/\text{free FruR}}$). The corresponding networks during gluconeogenesis are shown in the SI. Note that network $\mathcal{M}^0$ is always the same, regardless of the growth conditions.

gluconeogenic and glycolytic case, respectively (see Section S1 for the fully specified models). The guiding principles listed above were mostly sufficient for determining the inequality constraints between parameters, but in some situations several alternatives remained. In these cases we made an educated guess and chose one specific parametrization. In Section 3.5 we investigate the robustness of the modeling results for variations in parameter constraints.

3.3. Steady states of network during growth on glucose and acetate

The qualitative steady-state concentration levels of the enzymes and regulatory proteins involved in carbon assimilation, both during growth on glucose and acetate, can be predicted from the models $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$, by means of the computer tool GNA (Section 2.3). Both $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$ have a single stable steady state. Given that we focus on qualitative properties of the system,

we are particularly interested in the difference in steady-state concentrations during glycolysis and gluconeogenesis. That is, which genes are higher or lower expressed during growth on acetate as compared to growth on glucose? Table 1 summarizes the results of the comparison of the predicted glycolytic and gluconeogenic steady states. For instance, the table shows that during growth on acetate, instead of glucose, the expression of the gene *ppsA* increases whereas that of *pykF* decreases.

Do these predictions correspond to the available experimental data? Several data sets can be used to verify which genes are up-regulated or down-regulated when the growth conditions change. Oh et al. (2002) monitored transcription of *E. coli* by means of DNA microarrays during steady-state growth on acetate, using growth on glucose as the reference. Liu et al. (2005) also measured the transcription profiles by means of DNA microarrays, for six different carbon sources including glucose and acetate. Table 1 shows the relative expression changes for the genes that are included in the models $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$. As can be seen, the sign of the changes in expression is consistent between the two data sets, bearing in mind the experimental uncertainty. The experiment of Peng and Shimizu (2003) provided direct measurements of the accumulation of the enzymes in central carbon metabolism using two-dimensional electrophoresis, during growth on different carbon sources, including glucose and acetate. The observed differences for the two enzymes included in our model (PykF and PpsA) are consistent with the measured transcription profiles of the corresponding genes.

The predicted changes in expression correspond well to the observations (Table 1). The observed inversion of *pykF* and *ppsA* expression is correctly reproduced, as well as the down-regulation of *gapA* during growth on acetate. A switch in the expression of *pykF* and *ppsA* is reported as a major signature of the glycolysis–gluconeogenesis transition, as it promotes the inversion of metabolic fluxes and the consumption of the new carbon source. The exception to the overall good correspondence between model and experimental data is the predicted higher expression level of Crp in acetate-grown bacteria. The observed decrease in expression is somewhat surprising though, as other measurements obtained under glucose depletion show the contrary (Ishizuka et al., 1993) and are consistent with the model predictions. An alternative explanation is that some regulatory interactions are missing in the model, as the transcriptional control of *crp* is not completely understood.

Is the correspondence with experimental data preserved when the indirect interactions are omitted from the model? In order to answer this question, we first computed the steady states of $\mathcal{M}^0$, the model accounting for purely transcriptional interactions (see Section S2 of the SI). Table 1 shows that $\mathcal{M}^0$ fails to reproduce most of the observed changes in gene expression. Indeed, $\mathcal{M}^0$ has a single steady state for all growth conditions, so that the enzyme concentrations (*e.g.*, PykF and PpsA) remain the same when the carbon source changes. Clearly, metabolic coupling is essential for the adjustment of gene expression when cells are alternatively grown on glucose and acetate.

But which mechanism is mostly responsible for the observed changes in gene expression? Are all indirect interactions equally important for the system response? Further insights into the role of indirect interactions can be gained from the analysis of reference models in which one coupling species (either Crp · cAMP or free FruR) is removed at a time (see Section S2 for the models). The results show that the control of the gene expression response is shared between different mechanisms, in the sense that the absence of a coupling species does not allow all experimental data to be reproduced (Table 1). More precisely, the control exerted via the Crp · cAMP complex affects both the glycolytic enzymes and the global regulator *fis*, thus showing that Crp · cAMP-mediated coupling is at the heart of the regulatory system (Nanchen et al., 2008). Moreover, removal of Crp · cAMP modifies the structure of the phase space itself, in the sense that a second steady state appears for $\mathcal{M}^0_{neo/\text{Crp·cAMP}}$ which causes an ambiguity in the predicted expression levels of the glycolytic enzymes during gluconeogenesis. Metabolic coupling through free FruR has an effect on the expression of glycolytic enzymes, and thus the control of metabolic fluxes, but the expression of global regulators is not affected (Nanchen et al., 2008; Gutierrez-Ríos et al., 2007; Sarkar et al., 2008).

3.4. Dynamic response of gene expression to diauxic shift

When glucose is exhausted the bacteria are able to continue growth on an alternative carbon source like acetate. Such a diauxic shift entails important changes in metabolism, as well as a reorganization of gene expression. The dynamic response of gene expression has two aspects: an initial response on the transcriptional level when the glucose-grown bacteria sense the rapid diminution of substrate levels, and the eventual response on the enzyme level after the bacteria have switched to the utilization of acetate. Both of these responses can be taken into account using the $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$ models. In the former case, we consider the steady state of the glycolytic model and predict the immediate change of the protein synthesis rates during glucose depletion, bearing in mind that the protein synthesis rates are proportional to the mRNA concentrations (Kremling, 2007). In the latter case, we follow the predicted changes in enzyme levels in the gluconeogenic model, starting from an initial state that corresponds to the steady state in glycolysis. We thus make the assumption that on the time-scale of the dynamics of protein concentrations, which is on the order of hours in *E. coli*, the changes in metabolic fluxes are rapid. This assumption is supported by the available experimental data (Hardiman et al., 2007).

The predicted initial response of the bacteria to glucose depletion is shown in Table 2. We compare the synthesis rate of the proteins in the glycolysis steady state and in the successor

Table 1
Predicted and observed differences in gene expression during growth on acetate vs growth on glucose. The measurements concern the differences in mRNA concentrations (Liu et al., 2005; Oh et al., 2002) and protein concentrations (Peng and Shimizu, 2003). The model predictions are the differences in steady-state concentrations (+: higher, −: lower, 0: equal). In some cases the precision of the measurements does not allow to distinguish between qualitative differences, indicated by ?. If the model predictions agree (disagree) with the data, the entry is colored green (red). If the data are ambiguous, the predicted entry is shown in yellow. $\mathcal{M}^0_{neo/\text{Crp·cAMP}}$ has two steady states, leading to ambiguous predictions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

	crp	*fis*	*rpoS*	*fruR*	*gapA*	*ppsA*	*pykF*	Reference vs model
Experimental data	?	-	+	?	-	+	-	Oh et al. (2002)
	-	-	+	+		+	-	Liu et al. (2005)
						+	-	Peng and Shimizu (2003)
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	$\mathcal{M}^0$

Table 2
Predicted and observed initial responses on the transcriptional level when glucose is exhausted. The model predictions are the immediate change in the protein synthesis rate following the exhaustion of glucose, in comparison with the steady-state protein synthesis rate during glycolysis (+: higher, −: lower, 0: equal). The measurements concern the initial change in mRNA concentrations when the bacteria are transferred from a glucose to acetate medium (Kao et al., 2005). See Table 1 for the color codes.

	crp	*fis*	*rpoS*	*fruR*	*gapA*	*ppsA*	*pykF*	Reference/Model
Experimental data	-	-	-	0	-	+	0	Kao et al. (2005)
Model predictions	0	-	-	0	-	+	0	$\mathcal{M}_{glyco}$
	0	0	0	-	-	+	-	$\mathcal{M}^0_{glyco/\text{Crp·cAMP}}$
	0	0	0	0	0	0	0	$\mathcal{M}^0_{glyco/\text{free FruR}}$
	0	0	0	0	0	0	0	$\mathcal{M}^0$

state following the rapid drop in glucose levels (which is modeled in $\mathcal{M}_{glyco}$ by setting the glucose concentration variable below its threshold, see Appendix A). The synthesis rates of most proteins are seen to react immediately to the change in nutrient availability. In fact, the sharp increase of cAMP activates Crp which has a direct effect on many of the targets of this protein, which in our model consist of *fis*, *cya*, *gyrI*, *gapA*, *rpoS*, and *crp* itself.[2] These predictions can be compared with the experimental data of Kao et al. (2005), who used DNA microarrays to measure the transcriptional response of cells grown on glucose to mid-exponential phase and then reinoculated in a medium containing acetate. As can be seen, the predicted signs of the change in mRNA concentrations agree with the signs computed from the data.

The importance of metabolic coupling for the response dynamics can be assessed by comparing the above predictions with those obtained when leaving out some or all of the indirect interactions (Table 2). In the absence of metabolic coupling, the transcription factors Crp and FruR cannot sense the change in concentration of their metabolic effectors (cAMP and FBP), which is known to play a major role in the initiation of the response of the system to glucose exhaustion (Kotte et al., 2010). As a consequence, the model is no longer capable of accounting for the observed changes in gene expression.

The initial response of the bacterial cells occurs minutes after the change of growth conditions. In order to assess the response on the time-scale of protein synthesis and degradation, that is, hours after the change in carbon source, we perform a qualitative simulation with the $\mathcal{M}_{neo}$ model, taking the glycolysis steady state as our initial state. The system is seen to converge to the gluconeogenic steady state, in the sense that all paths in the state transition graph eventually reach this state. The different paths represent possible sequences of qualitative events leading the system to the new steady state (see Fig. 6 for an example). We exploited model-checking techniques to characterize properties common to all paths. Given that the state transition graph is generated from a specified ordering of parameter values, rather than exact numerical values, such properties are robust with respect to small quantitative changes in parameter values and capture fundamental aspects of the underlying regulatory structure of the network.

Following the shift to acetate, the glycolytic enzymes are the first to respond. The level of PpsA, essential for the metabolism of gluconeogenic substrates, starts to increase immediately after the diauxie, whereas PykF and GapA are progressively down-regulated, due to the high levels of free FruR in the new growth conditions. The expression of the genes encoding the global regulators responds later. A sufficiently high level of PpsA is required for the accumulation of Crp · cAMP, and the change in expression of the genes *fis* and *crp*, which are under the control of this regulator. Interestingly, the positive influence of PpsA on the accumulation of Crp · cAMP in the absence of glucose, an indirect interaction inferred by the method described in Section 2.1, is confirmed by the experimental work of Kao et al. (2005). They show that in a *ppsA* mutant the bacteria took much longer to resume growth on acetate after the exhaustion of glucose. The wild-type phenotype could be partially rescued by the addition of cAMP, which suggests that PpsA controls the concentrations of cAMP and Crp · cAMP, as predicted by the model.

3.5. Robustness analysis to parameter changes

As explained in Section 3.2, the qualitative PL models include inequality constraints, some of which are underdetermined by the available experimental data. In particular, the indirect interactions due to metabolic coupling involve parameters that characterize lumped processes occurring on different time-scales, so that a direct measure of their strength is difficult to obtain experimentally. The question can be asked, first, to which extent the results obtained with the models $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$ are robust to changes in parameter constraints. Second, the indirect interactions are not expected to all have the same importance, so it would be interesting to identify those interactions for which a specific ordering between parameters is critical for reproducing the observed behavior of the network.

Within the constraints provided by the criteria introduced in Section 3.2, we analyzed the robustness of our results by randomly changing the ordering of thresholds characterizing the indirect interactions mediated by either Crp · cAMP or free FruR. Each different ordering leads to a different set of inequality constraints, and thus to a different qualitative PL model (see Appendix B for details on the alternative sets of inequality constraints). For each of the thus generated alternative models, the gene expression response during a glucose–acetate diauxie is analyzed in the way described in Sections 3.3 and 3.4, and compared to the prediction of the original models $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$. For each gene in the network, we define a robustness index as the fraction of models for which the predictions remain the same after a change of inequality constraints (see Chaves et al., 2006 for another approach to test the robustness of qualitative models).

The results of this robustness analysis are shown in Table 3, for the predictions of steady-state and dynamic responses. Overall, the model predictions are robust to changes in parameter constraints, thus emphasizing the importance of the network topology for the emergence of the observed response. The major exception are variations in the parameter constraints for free FruR, which may strongly affect the predictions of enzyme steady-state levels. This can be explained by the fact that the space of possible parameter orderings for free FruR is much less constrained by the available biological data than that for Crp · cAMP. Interestingly, from a biological point of view, we find that the control of *gapA* and *ppsA* expression levels via metabolic coupling is essential for the correct system response. In particular, *gapA* autoregulation via FruR should not be so strong as to prevent

[2] Notice that the inhibitory effect of Crp · cAMP complex on the transcription of RpoS remains somewhat controversial (see Mandel and Silhavy, 2005; Venturi, 2003; Zgurskaya et al., 1997 and references therein).

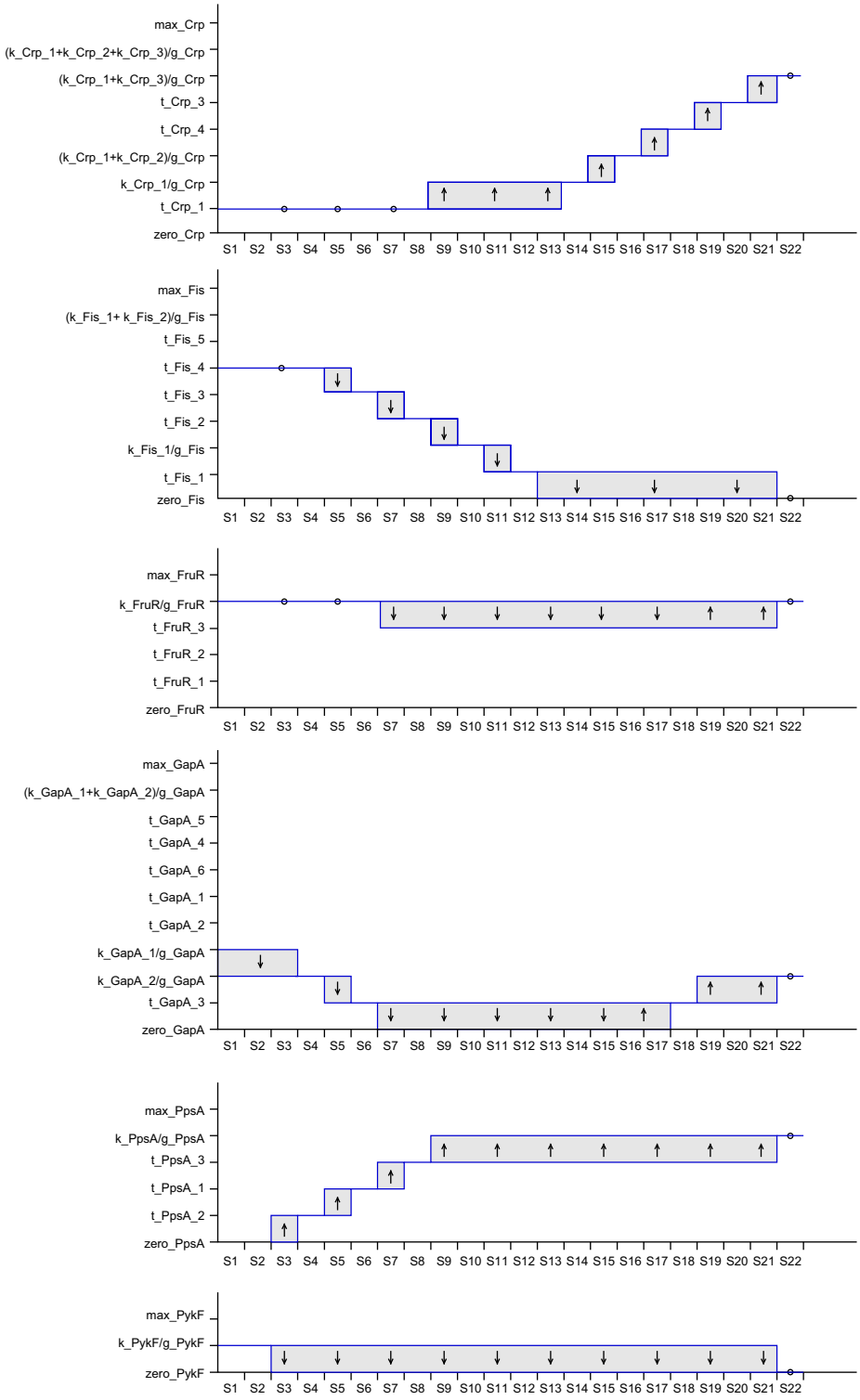


Fig. 6. Example of a qualitative simulation of the glucose–acetate diauxie (Section 2.3). The vertical axis shows the symbolic values of concentration variables, the horizontal axis indicates the qualitative states of the system (note that time is implicit in a qualitative model, given by the ordering of qualitative states). The selected pathway shows the typical dynamics of protein concentrations following the shift to acetate. The glycolytic enzymes (PpsA, GapA, PykF) are the first to respond. Global regulators (Crp, Fis, FruR) respond later, once a sufficiently high level of PpsA is reached. The derivative sign (steady, ○; increasing, ↑; decreasing, ↓) of concentration variables in each qualitative state is explicitly indicated. Transitions between qualitative states correspond to qualitative events, notably threshold crossings of the variables.

the activation of Crp after a glucose–acetate transition. This requires an appropriate balance to be found between the opposite influences of free FruR and Crp · cAMP on *gapA* expression, bearing in mind that both of these coupling species accumulate upon glucose exhaustion.

4. Discussion

The question how interactions between metabolism and gene regulation shape the dynamics of cellular processes has been an active research topic in recent years. One approach is based on

Table 3
Robustness index of *A*: steady-state predictions (Table 1) and *B*: initial transcriptional response (Table 2), when changing the relative strength of indirect interactions in models $\mathcal{M}_{neo}$ (for steady-state results) and $\mathcal{M}_{glyco}$ (for initial transcriptional response). The number of alternative parameter orderings is indicated between parentheses.

Coupling species	*crp*	*fis*	*rpoS*	*fruR*	*gapA*	*ppsA*	*pykF*
A							
Crp · cAMP (20)	0.85	1	1	0.35	1	1	1
Free FruR (15)	0.07	1	1	0.67	0.07	0.07	0.07
Total (35)	0.51	1	1	0.49	0.6	0.6	0.6
B							
Crp · cAMP (20)	0.6	0.6	0.6	1	1	1	1
free FruR (20)	1	1	1	0.8	0.2	0.5	0.5
Total (40)	0.8	0.8	0.8	0.9	0.6	0.75	0.75

fully specified kinetic models of integrated networks of enzymatic reactions and protein synthesis and degradation reactions (*e.g.*, Kotte et al., 2010; Klipp et al., 2005). Such models provide a fine-grained mechanistic picture of the biochemical reaction network. Changes in the environment modify metabolic fluxes, which leads to changes in the activity of transcription factors and other regulatory proteins. The latter affect the expression of genes encoding enzymes, thus bringing about further adjustments of metabolic fluxes. Another type of approach consists of extensions of flux balance analysis (FBA) aiming at the integration of gene regulation with metabolism (*e.g.*, Covert et al., 2004, 2008; Shlomi et al., 2007). These methods sidestep the explicit specification of rate laws. The basic idea is to generate, for a given gene expression state, a metabolic flux distribution that satisfies an objective function, for instance optimal biomass production. Boolean rules then couple the flux distribution to changes in gene expression. These methods enable the prediction of changes in metabolic fluxes over time, under different experimental conditions, and the prediction of the effect of these changes on the activity of transcription factors.

The approach outlined in this paper differs from those outlined above in that we do not explicitly include metabolism in our model nor predict metabolic flux distributions and metabolite concentrations. Rather, we identify *a priori* the indirect interactions mediated by metabolism which couple the expression levels of genes coding for enzymes to those encoding transcription factors. This approach can be seen as a model reduction that uncovers the effective network structure on the time-scale of gene expression (related ideas can be found in Radulescu et al., 2006). The dynamics of the extended network including both direct and indirect interactions can be investigated by means of classical models of gene regulatory networks, such as the PL models used in this study. Another recent example of the integration of metabolism and gene expression in the context of PL models can be found in Oyarzún and Chaves (2011). It provides a detailed analysis of a specific class of networks that can be reduced to a two-variable feedback system on the time-scale of gene expression.

The results of the application of our approach to the *E. coli* network show that indirect interactions play a major role in shaping the dynamics of the system. They completely modify the predicted expression patterns following a glucose-acetate shift, transforming metabolic genes from passive to active players of the adaptive response of the cell. The profiles obtained when including metabolic coupling show good correspondence with the available experimental data, contrary to what is observed when the model is restricted to transcriptional regulation only. In particular, key signatures of growth on a glycolytic substrate (glucose) as compared to growth on a gluconeogenic substrate (acetate) are reproduced, such as the opposite expression levels of the enzymes PpsA and PykF. The indirect interactions arising from metabolic coupling are found to be crucial, because they account for rapid changes in metabolism (fluxes, metabolite concentrations) that control the activity of transcriptional regulators. This allows environmental perturbations, such as the depletion of glucose, to propagate through the network and affect the expression of a large number of genes.

More generally, when are indirect interactions arising from metabolic coupling expected to play a role? Analysis of Eq. (5) provides a suitable framework to address this question. For an enzyme i to have an effect on the synthesis rate v_j^s of a protein, first of all, there should exist a metabolic coupling species k, such that both $\partial g_k(x^s)/\partial x_i^s$ and $\partial v_j^s(x^s, g(x^s))/\partial x_k^f$ are non-zero. That is, the synthesis rate of the protein should be regulated by a transcription factor binding to a signaling metabolite under the control of the enzyme, such as Crp and FruR in the *E. coli* network. Second, the concentration control coefficient $\partial g_k(x^s)/\partial x_i^s$ should be non-negligible. In words, a change in enzyme concentration should affect the activation state of the transcription factor. This was actually the case in Fig. 3, where a variation of the concentration of Cya significantly changed the activation state of Crp. A third condition for metabolic coupling to occur is that the elasticity coefficient $\partial v_j^s(x^s, g(x^s))/\partial x_k^f$ should be such that the change in transcription factor activity falls within the dynamic range of gene regulation. For instance, if the concentration of Crp · cAMP remains far from the threshold $\theta_{Crp\cdot cAMP}$, then the effect of the transcription factor on gene expression would be negligible.

We studied the dynamical consequences of metabolic coupling by means of PL models, because (a) quantitative information about gene regulation in *E. coli* remains scarce, and (b) we are particularly interested in the qualitative effects of metabolic coupling. Despite the fact that PL models and their Boolean relatives are based on a number of modeling approximations (Section 2), they have been shown able to account for the qualitative dynamics of complex biological systems and adequately reproduce experimental data sets (Saez-Rodriguez et al., 2009; Naldi et al., 2010; Li et al., 2006). This has motivated their use in this paper. Notice that the predictions from the *E. coli* model are in good agreement with observed relative changes in gene expression (Tables 1 and 2). We are aware that the use of PL models has its limitations though, and that other formalisms might provide a more precise and quantitative picture of metabolic coupling. Classical kinetic models with fully specified rate laws and parameter values, for instance, would allow the elements of the Jacobian matrix in Eq. (5) to be quantified, and enable the comparison of the relative strengths of indirect interactions in different physiological states. Although qualitative and quantitative modeling may yield different insights, this does not change the main point of this paper that metabolic coupling affects the dynamics of gene regulatory networks, which we believe is rather independent of the particular modeling formalism chosen.

Metabolic coupling is widespread in regulatory networks, as can be illustrated with other examples from the literature. Consider the TOL network in *Pseudomonas putida* mt-2, involved in the degradation of m-xylene in this soil bacterium (Silva-Rocha et al., 2011). The network includes two groups of enzymes catalyzing the transformation of m-xylene to 3-methylbenzoate (3MB) (*upper* pathway) and from 3MB to intermediates of the TCA cycle (*lower* pathway), respectively. Metabolic coupling arises as changes in the concentration of the *upper* enzymes affect the activity of the transcription factor XylS which regulates the expression of the *lower* enzymes. Nitrogen regulation in bacteria and archaea is another typical example of coupling with metabolism (Leigh and Dodsworth, 2007). In these systems, two

metabolites, 2-oxoglutarate and glutamine, respectively precursor and product of the ammonia assimilation pathway, are able to control the activity and the expression of several enzymes, via the intermediary of the protein allosteric effector PII. Indirect interactions can also be mediated by signaling pathways, as illustrated by the ethylene-mediated control of the tomato ripening process (Cara and Giovannoni, 2008). Ethylene belongs to the family of hormones, a group of small signaling molecules involved in the control of plant development and response to environmental factors (Santner and Estelle, 2009). During fruit maturation, the ethylene concentration increases, inducing the expression of a large class of ethylene response factors genes, involved, among other processes, in the up-regulation of cell wall degradation proteins and consequent fruit softening. Two major genes, *rin* and *nor*, control ethylene synthesis in tomato. Through the intermediary of changes in ethylene concentration, the knock-out of one of these genes affects cell wall metabolism and prevents the normal ripening of tomato fruit (Osorio et al., 2011).

In conclusion, we believe that metabolic coupling is an important concept for understanding the dynamics of gene regulatory networks. Among other things, it has implications for reverse engineering gene regulatory networks from genome-wide expression data sets (Hecker et al., 2009). The existence of indirect interactions arising from metabolic coupling puts in doubt the commonly made assumption that gene regulatory networks are sparse, that is, that the expression of a gene is influenced by only a limited number of regulators. Due to metabolic coupling, a gene may actually have a large number of regulators, including metabolic genes whose expression controls metabolic fluxes and the activity state of transcription factors. The resulting picture of the network structure and the dynamics it sustains may thus be much more complex than is sometimes assumed.

Acknowledgments

This work was supported by the European Commission under project EC-MOAN (FP6-2005-NEST-PATH-COM/043235) and the Agence Nationale de la Recherche under project MetaGenoReg (ANR-06-BYOS-0003).

Appendix A. Kinetic model of subnetwork controlling Crp activation

The Crp activation subnetwork is composed of the proteins Cya, Crp, and the complex Crp · cAMP (see Fig. A1). When glucose runs out, Cya is activated through the phosphotransferase system (PTS), and catalyzes the conversion of ATP into cAMP (Saier Jr. et al., 1996). For simplicity, we do not describe the phosphotransferase system in our model, but rather consider a carbon starvation signal that directly stimulates the activity of Cya, and thus the synthesis of cAMP. This signal is identified with the glucose concentration.

In Ropers et al. (2011) we used standard approaches of biochemistry to develop a detailed kinetic model of the carbon activation network

$$\dot{x}_{Cya\ free} = \kappa_y^1 + \kappa_y^2 h^-(x_{Crp\cdot cAMP}, \theta^2_{Crp\cdot cAMP}, m_{Crp\cdot cAMP}) - \gamma_y x_{Cya\ free} + (k_{-1} + k_2 h^-(u_{Glc}, \theta_{Glc}, m_{Glc}))x_{Cya\cdot ATP} - k_1 x_{Cya\ free} u_{ATP}$$

$$\dot{x}_{Crp\ free} = \kappa_c^1 + \kappa_c^2 h^+(x_{Crp\cdot cAMP}, \theta^1_{Crp\cdot cAMP}, m_{Crp\cdot cAMP}) - \gamma_c x_{Crp\ free} + k_{-4} x_{Crp\cdot cAMP} - k_4 x_{Crp\ free} x_{cAMP\ free}$$

$$\dot{x}_{Cya\cdot ATP} = k_1 x_{Cya\ free} u_{ATP} - (k_{-1} + k_2 h^-(u_{Glc}, \theta_{Glc}, m_{Glc}) + \gamma_y) x_{Cya\cdot ATP}$$

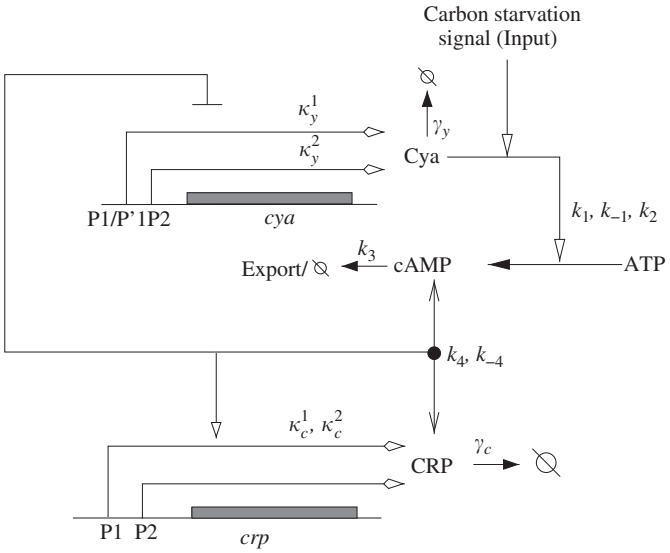


Fig. A1. Crp activation network.

Table A1
Parameter values for the kinetic model of the subnetwork controlling Crp activation (Ropers et al., 2011).

Parameter	Value	Parameter	Value
κ_y^1	$2.6 \times 10^{-13}\ M^{-1}\ s^{-1}$	κ_y^2	$2.27 \times 10^{-13}\ M^{-1}\ s^{-1}$
κ_c^1	$6.89 \times 10^{-9}\ M^{-1}\ s^{-1}$	κ_c^2	$5.75 \times 10^{-8}\ M^{-1}\ s^{-1}$
k_1	$1.32 \times 10^{8}\ M^{-1}\ s^{-1}$	k_{-1}	$0.64\ s^{-1}$
k_2	$135\ s^{-1}$	k_3	$1.007\ s^{-1}$
k_4	$8.28 \times 10^{5}\ M^{-1}\ s^{-1}$	k_{-4}	$0.74\ s^{-1}$
γ_y	$0.0062\ s^{-1}$	γ_c	$0.303\ s^{-1}$
$\theta^1_{Crp\cdot cAMP}$	$8.27 \times 10^{-6}\ M$	$\theta^2_{Crp\cdot cAMP}$	$3.46 \times 10^{-7}\ M$

$$\dot{x}_{Crp\cdot cAMP} = k_4 x_{Crp\ free} x_{cAMP\ free} - (k_{-4} + \gamma_c) x_{Crp\cdot cAMP}$$

$$\dot{x}_{cAMP\ free} = k_2 h^-(u_{Glc}, \theta_{Glc}, m_{Glc}) x_{Cya\cdot ATP} + k_{-4} x_{Crp\cdot cAMP} - k_3 x_{cAMP\ free} - k_4 x_{Crp\ free} x_{cAMP\ free} \quad (A.1)$$

where $x_{Cya\ free}$, $x_{Crp\ free}$, and $x_{cAMP\ free}$ denote the concentrations of free (unbound) Cya, Crp, and cAMP, respectively, and $x_{Cya\cdot ATP}$ and $x_{Crp\cdot cAMP}$ denote the concentrations of the complexes Cya · ATP and Crp · cAMP, respectively. u_{Glc} and u_{ATP} denote the glucose and ATP concentrations. The kinetic constants are associated with the reaction rates in Fig. A1. The above model includes fast and slow reactions, with high and low reactions rates, respectively (see Table A1 for parameter values). Typically, complex formation and enzymatic reactions are fast, while protein synthesis and degradation are slow. Based on this distinction, we define slow variables (total protein concentrations) and fast variables (concentrations of biochemical complexes and metabolites) as linear combinations of the original variables

$$x_{Cya} = x_{Cya\ free} + x_{Cya\cdot ATP}$$

$$x_{Crp} = x_{Crp\ free} + x_{Crp\cdot cAMP}$$

$$x_{cAMP} = x_{cAMP\ free} + x_{Crp\cdot cAMP}$$

The variable transformation ensures that slow variables change through slow reactions only, while fast reactions dominate the

dynamics of fast variables. With the above transformations, the system of Eq. (A.1) can be rewritten as:

Slow system

$$\dot{x}_{Cya} = \kappa_y^1 + \kappa_y^2 h^-(x_{Crp\cdot cAMP}, \theta^2_{Crp\cdot cAMP}, m^2_{Crp\cdot cAMP}) - \gamma_y x_{Cya}$$

$$\dot{x}_{Crp} = \kappa_c^1 + \kappa_c^2 h^+(x_{Crp\cdot cAMP}, \theta^1_{Crp\cdot cAMP}, m^1_{Crp\cdot cAMP}) - \gamma_c x_{Crp} \quad (A.2)$$

Fast system

$$\dot{x}_{Cya\cdot ATP} = k_1(x_{Cya} - x_{Cya\cdot ATP})u_{ATP} - (k_{-1} + k_2 h^-(u_{Glc}, \theta_{Glc}, m_{Glc}) + \gamma_y)x_{Cya\cdot ATP}$$

$$\dot{x}_{Crp\cdot cAMP} = k_4(x_{Crp} - x_{Crp\cdot cAMP})(x_{cAMP} - x_{Crp\cdot cAMP}) - (k_{-4} + \gamma_c)x_{Crp\cdot cAMP}$$

$$\dot{x}_{cAMP} = k_2 h^-(u_{Glc}, \theta_{Glc}, m_{Glc})x_{Cya\cdot ATP} + (k_3 - \gamma_{Crp})x_{Crp\cdot cAMP} - k_3 x_{cAMP} \quad (A.3)$$

We apply the quasi-steady state (QSS) assumption to the reformulated system, by putting the time derivative of fast variables equal to 0 and neglecting the contribution of the slow variables to the dynamics of the fast variables (Heinrich and Schuster, 1996). This results in a system of algebraic equations for the fast variables that can be explicitly solved, leading to the following reduced system at the slow time-scale:

$$\dot{x}_{Cya} = \kappa_y^1 + \kappa_y^2 h^-(x_{Crp\cdot cAMP}, \theta^2_{Crp\cdot cAMP}, m^2_{Crp\cdot cAMP}) - \gamma_y x_{Cya}$$

$$\dot{x}_{Crp} = \kappa_c^1 + \kappa_c^2 h^+(x_{Crp\cdot cAMP}, \theta^1_{Crp\cdot cAMP}, m_{Crp\cdot cAMP}) - \gamma_c x_{Crp}$$

$$x_{Crp\cdot cAMP} = \frac{h^-(u_{Glc}, \theta_{Glc}, m_{Glc})x_{Cya}}{h^-(u_{Glc}, \theta_{Glc}, m_{Glc})x_{Cya} + K_4 K_3} x_{Crp}$$

$$K_3 = \frac{k_3}{k_2}$$

$$K_4 = \frac{k_{-4}}{k_4} \quad (A.4)$$

This model has been used to generate the plot in Fig. 3.

Appendix B. Model variants of $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$ used for robustness tests

In order to assess the robustness of the results obtained with the models $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$, we built a set of model variants by randomly changing the ordering of thresholds characterizing the indirect interactions mediated by either Crp · cAMP or free FruR.

Each model variant satisfies biological and mathematical constraints that reflect current biological knowledge of the system and structural dependencies between variables. We illustrate the general principles on a small example representing indirect interactions mediated by free FruR during gluconeogenesis. The topology of this subnetwork and the corresponding PL model are shown in Fig. B1. The amount of the coupling species free FruR depends on the concentration of proteins FruR and GapA (which control the concentration of FBP). Free FruR regulates the expression of four target genes: *fruR*, *gapA*, *ppsA*, and *pykF*. The corresponding PL model shown in the figure has six distinct thresholds: three for the concentration of FruR ($\theta^1_{FruR}, \theta^2_{FruR}, \theta^3_{FruR}$) and three for GapA ($\theta^1_{GapA}, \theta^2_{GapA}, \theta^3_{GapA}$), where we assume that the control of *pykF* and *ppsA* occurs at the same concentration level. The difficulty consists in choosing a plausible ordering of these thresholds.

A first constraint on the threshold ordering derives from the definition of indirect interactions itself and tightly links the relative order of the θ_{GapA} and the θ_{FruR} thresholds. We briefly explain why. Similar to the example in Fig. 3 of the main text, the concentration of free FruR at steady state can be described as a function of the concentrations of FruR and GapA

$$x_{FruR\ free} = g(x_{FruR}, x_{GapA}, u_{Glc}) \approx \frac{K}{x_{GapA} + K} x_{FruR} \quad (B.1)$$

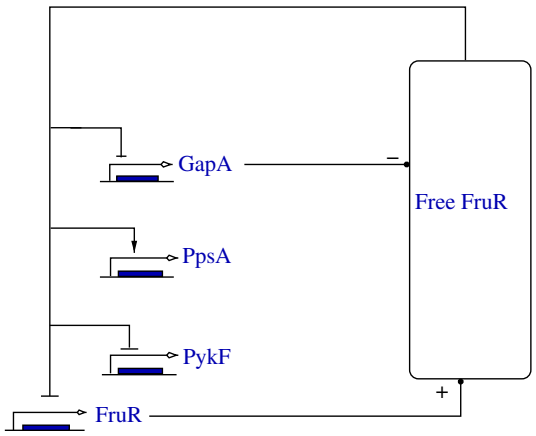
where K is a lumped kinetic parameter. We plot $h^+(x_{FruR\ free}, \theta_{FruR\ free})$ as a function of FruR and GapA in Fig. B2 for two different values of the threshold concentration ($\theta^1_{FruR\ free}$ and $\theta^2_{FruR\ free}$). The bottom panel shows the corresponding cross-sections with $x_{FruR} = 10$ and $x_{GapA} = 10$ planes (arbitrary units).

As expected from the considerations in Section 2.2 of the main text, the surface is sigmoidal in shape, for both variables GapA and FruR, and it can be approximated by a product of Hill functions. In particular, the Hill function will be positive for FruR and negative for GapA, as evident from the slope of the curves in Fig. B2 (see panel *B*), with threshold concentrations $\theta^1_{FruR}, \theta^2_{FruR}$ and $\theta^1_{GapA}, \theta^2_{GapA}$, respectively.

As can be seen, the thresholds for free FruR, GapA and FruR are not independent since, by construction, they are supposed to lie on the same surface. Indeed, a change in the threshold concentration of free FruR causes a change in the corresponding threshold concentrations of GapA and FruR. Moreover, if $\theta^1_{FruR\ free} < \theta^2_{FruR\ free}$, the relative ordering between thresholds for FruR and GapA is imposed by the sigmoidal surface. In particular, the order is preserved for FruR *i.e.* $\theta^1_{FruR} < \theta^2_{FruR}$ whereas it is reversed in the case of GapA, *i.e.* $\theta^1_{GapA} > \theta^2_{GapA}$, as shown in Fig. B2*B*. These dependencies between thresholds of variables controlling a coupling species thus reduce the number of possible orderings that have to be considered as model variants.

Further constraints come from biological knowledge on the system. For instance, the relative ordering of the focal points of GapA is known from the literature: *gapA* is transcribed from several promoters, among which promoter P1 (under the control of free FruR) is the most efficient (Thouvenot et al., 2004). We can therefore assume $\kappa^2_{GapA}/\gamma_{GapA} < \kappa^1_{GapA}/\gamma_{GapA} < (\kappa^1_{GapA} + \kappa^2_{GapA})/\gamma_{GapA}$.

The variants of the models $\mathcal{M}_{glyco}$ and $\mathcal{M}_{neo}$ studied in Section 3.5 of the main text have been generated along the lines sketched above. Despite the above mathematical and biological constraints, the number of possible variants remains too big for manual analysis (480 for our small example). We therefore opted for a random, uniform sampling of the ensemble of model variants.



$$\dot{x}_{FruR} = \kappa_{FruR}\,(1 - s^+(x_{FruR}, \theta^3_{FruR})\,s^-(x_{GapA}, \theta^3_{GapA})) - \gamma_{FruR}\,x_{FruR}$$
$$\dot{x}_{GapA} = \kappa^2_{GapA} + \kappa^1_{GapA}\,(1 - s^+(x_{FruR}, \theta^1_{FruR}) s^-(x_{GapA}, \theta^1_{GapA})) - \gamma_{GapA}\,x_{GapA}$$
$$\dot{x}_{PpsA} = \kappa_{PpsA}\,s^+(x_{FruR}, \theta^2_{FruR})\,s^-(x_{GapA}, \theta^2_{GapA}) - \gamma_{PpsA}\,x_{PpsA}$$
$$\dot{x}_{PykF} = \kappa_{PykF}\,(1 - s^+(x_{FruR}, \theta^2_{FruR})\,s^-(x_{GapA}, \theta^2_{GapA})) - \gamma_{PykF}\,x_{PykF}$$

Fig. B1. Network of indirect interactions mediated by the coupling species free FruR during gluconeogenesis and the corresponding PL model.

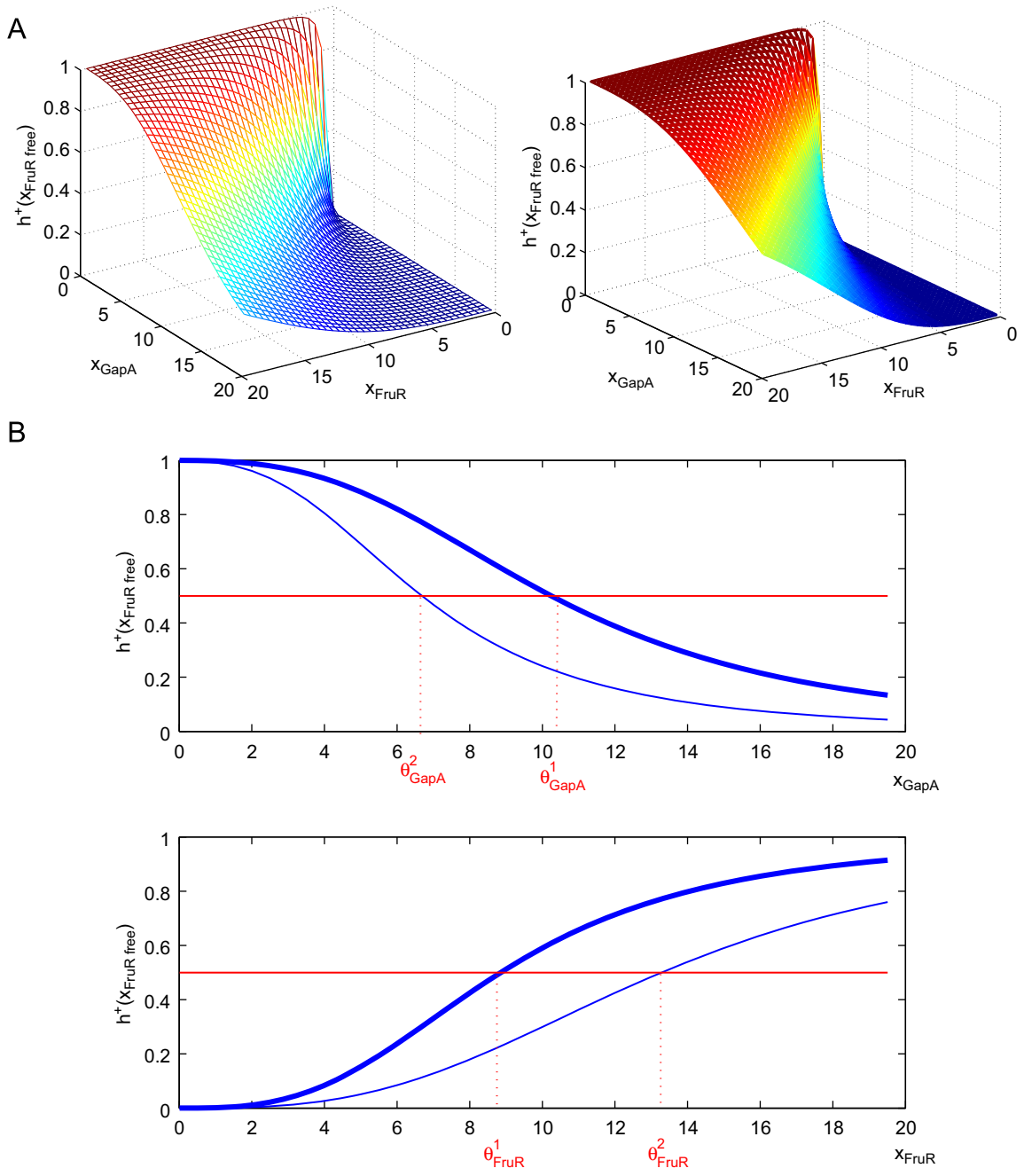


Fig. B2. Plot of $h^+(x_{FruR\ free},\theta_{FruR\ free})$ as a function of x_{FruR} and x_{GapA} (total FruR and GapA concentrations in arbitrary units) for two different values of the threshold concentration $\theta_{FruR\ free}$ (left: $\theta^1_{FruR\ free}$, right: $\theta^2_{FruR\ free}$). B: Sections by the $x_{FruR}=10$ and $x_{GapA}=10$ planes, for the two values of $\theta_{FruR\ free}$. A red line at 0.5 defines the position of the corresponding thresholds concentrations for FruR and GapA (θ^1_{FruR},θ^2_{FruR} and θ^1_{GapA},θ^2_{GapA}). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Appendix C. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jtbi.2011.11.010.

References

Baldazzi, V., Ropers, D., Markowicz, Y., Kahn, D., Geiselmann, J., de Jong, H., 2010. The carbon assimilation network in *Escherichia coli* is densely connected and largely sign-determined by directions of metabolic fluxes. PLoS Comput. Biol. 6, e1000812.

Batt, G., de Jong, H., Page, M., Geiselmann, J., 2008. Symbolic reachability analysis of genetic regulatory networks using discrete abstractions. Automatica 44, 982–989.

Bettenbrock, K., Fischer, S., Kremling, A., Jahreis, K., Sauter, T., Gilles, E., 2005. A quantitative approach to catabolite repression in *Escherichia coli*. J. Biol. Chem. 281, 2578–2584.

Bolouri, H., 2008. Computational Modeling of Gene Regulatory Networks—A Primer. Imperial College Press, London.

Brazhnik, P., de la Fuente, A., Mendes, P., 2002. Gene networks: how to put the function in genomics. Trends Biotechnol. 20, 467–472.

Cara, B., Giovannoni, J., 2008. Molecular biology of ethylene during tomato fruit development and maturation. Plant Sci. 175, 106–113.

Chassagnole, C., Noisommit-Rizzi, N., Schmid, J., Mauch, K., Reuss, M., 2002. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. Biotechnol. Bioeng. 79, 53–73.

Chaves, M., Sontag, E., Albert, R., 2006. Methods for the robustness analysis for Boolean models. IET Syst. Biol. 153, 154–167.

Covert, M., Knight, E., Reed, J., Herrgard, M., Palsson, B., 2004. Integrating high-throughput and computational data elucidates bacterial networks. Nature 429, 92–100.

Covert, M., Xiao, N., Chen, T., Karr, J., 2008. Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. Bioinformatics 24, 2044–2050.

Davidich, M., Bornholdt, S., 2008. The transition from differential equations to Boolean networks: a case study in simplifying a regulatory network model. J. Theor. Biol. 255, 269–277.

de Jong, H., 2002. Modeling and simulation of genetic regulatory systems: a literature review. J. Comput. Biol. 9, 67–103.

de Jong, H., Ropers, D., 2006. Strategies for dealing with incomplete information in the modeling of molecular interaction networks. Brief. Bioinform. 7, 354–363.

de Jong, H., Geiselmann, J., Hernandez, C., Page, M., 2003. Genetic Network Analyzer: qualitative simulation of genetic regulatory networks. Bioinformatics 19, 336–344.

de Jong, H., Gouzé, J.-L., Hernandez, C., Page, M., Sari, T., Geiselmann, J., 2004a. Qualitative simulation of genetic regulatory networks using piecewise-linear models. Bull. Math. Biol. 66, 301–340.

de Jong, H., Geiselmann, J., Batt, G., Hernandez, C., Page, M., 2004b. Qualitative simulation of the initiation of sporulation in *B. subtilis*. Bull. Math. Biol. 66, 261–299.

Fauré, A., Thieffry, D., 2009. Logical modelling of cell cycle control in eukaryotes: a comparative study. Mol. Biosyst. 5, 1569–1581.

Glass, L., Kauffman, S., 1973. The logical analysis of continuous non-linear biochemical control networks. J. Theor. Biol. 39, 103–129.

Gutierrez-Ríos, R., Freyre-Gonzalez, J., Resendis, O., Collado-Vides, J., Saier, M., Gosset, G., 2007. Identification of regulatory network topological units coordinating the genome-wide transcriptional response to glucose in *Escherichia coli*. BMC Microbiol. 7, 53.

Hardiman, T., Lemuth, K., Reuss, M.K.M., Siemann-Herzberg, M., 2007. Topology of the global regulatory network of carbon limitation in *Escherichia coli*. J. Biotechnol. 132, 359–374.

Hecker, M., Lambeck, S., Toepfer, S., vanSomeren, E., Guthke, R., 2009. Gene regulatory network inference: data integration in dynamic models—a review. Biosystems 96, 86–103.

Heinrich, R., Schuster, S., 1996. The Regulation of Cellular Systems. Chapman & Hall, New York.

Ironi, L., Panzeri, L., 2009. A computational framework for qualitative simulation of nonlinear dynamical models of gene-regulatory networks. BMC Bioinf. 10, S14.

Ishizuka, H., Hanamura, A., Kunimura, T., Aiba, H., 1993. A lowered concentration of cAMP receptor protein caused by glucose is an important determinant for catabolite repression in *Escherichia coli*. Mol. Microbiol. 10, 341350.

Kao, K.C., Tran, L.M., Liao, J.C., 2005. A global regulatory role of gluconeogenic genes in *Escherichia coli* revealed by transcriptome network analysis. J. Biol. Chem. 280, 36079–36087.

Kappler, K., Edwards, R., Glass, L., 2003. Dynamics in high-dimensional model gene networks. Signal Process. 83, 789–798.

Karlebach, G., Shamir, R., 2008. Modelling and analysis of gene regulatory networks. Nat. Rev. Mol. Cell Biol. 9, 770–780.

Keseler, I.M., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muñiz-Rascado, L., Bonavides-Martinez, C., Paley, S., Krummenacker, M., Altman, T., Kaipa, P., Spaulding, A., Pacheco, J., Latendresse, M., Fulcher, C., Sarker, M., Shearer, A.G., Mackie, A., Paulsen, I., Gunsalus, R.P., Karp, P.D., 2011. EcoCyc: a comprehensive database of *Escherichia coli* biology. Nucleic Acids Res. 39, D583–D590.

Khalil, H., 2001. Nonlinear Systems, 3rd ed. Prentice Hall, Upper Saddle River, NJ.

Klipp, E., Nordlander, B., Krüger, R., Gennemark, P., Hohmann, S., 2005. Integrative model of the response of yeast to osmotic shock. Nat. Biotechnol. 23, 975–982.

Kotte, O., Zaugg, J., Heinemann, M., 2010. Bacterial adaptation through distributed sensing of metabolic fluxes. Mol. Syst. Biol. 6, 355.

Kremling, A., 2007. Comment on mathematical models which describe transcription and calculate the relationship between mRNA and protein expression ratio. Biotechnol. Bioeng. 96, 815–819.

Leigh, J.A., Dodsworth, J.A., 2007. Nitrogen regulation in bacteria and archaea. Annu. Rev. Microbiol. 61, 349–377.

Li, S., Assmann, S., Albert, R., 2006. Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. PLoS Biol. 4, e312.

Liu, M., Durfee, T., Cabrera, J., Zhao, K., Jin, D., Blattner, F., 2005. Global transcriptional programs reveal a carbon source foraging strategy by *Escherichia coli*. J. Biol. Chem. 280, 15921–15927.

Mandel, M., Silhavy, T., 2005. Starvation for different nutrients in *Escherichia coli* results in differential modulation of RpoS levels and stability. J. Bacteriol. 187, 434–442.

Monteiro, P., Ropers, D., Mateescu, R., Freitas, A., de Jong, H., 2008. Temporal logic patterns for querying dynamic models of cellular interaction networks. Bioinformatics 24, i227–i233.

Morris, M., Saez-Rodriguez, J., Sorger, P., Lauffenburger, D., 2010. Logic-based models for the analysis of cell signaling networks. Biochemistry 49, 3216–3225.

Naldi, A., Carneiro, J., Chaouiya, C., Thieffry, D., 2010. Diversity and plasticity of Th cell types predicted from regulatory network modelling. PLoS Comput. Biol. 6, e1000912.

Nanchen, A., Schicker, A., Revelles, O., Sauer, U., 2008. Cyclic AMP-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in *Escherichia coli*. J. Bacteriol. 190, 2323–2330.

Oh, M.-K., Rohlin, L., Kao, K.C., Liao, J.C., 2002. Global expression profiling of acetate-grown *Escherichia coli*. J. Biol. Chem. 277, 13175–13183.

Osorio, S., Alba, R., Damasceno, C.M.B., Lopez-Casado, G., Lohse, M., Zanor, M.I., Tohge, T., Usadel, B., Rose, J.K.C., Fei, Z., Giovannoni, J.J., Fernie, A.R., 2011. Systems biology of tomato fruit development: combined transcript, protein, and metabolite analysis of tomato transcription factor (nor, rin) and ethylene receptor (Nr) mutants reveals novel regulatory interactions. Plant Physiol. 157, 405–425.

Oyarzún, D., Chaves, M., 2011. Global gene regulation in metabolic networks. In: Proceedings of the 18th IFAC World Congress. Milano, Italy. pp. 14838–14843.

Peng, L., Shimizu, K., 2003. Global metabolic regulation analysis for *Escherichia coli* K12 based on protein expression by 2-dimensional electrophoresis and enzyme activity measurement. Appl. Microbiol. Biotechnol. 61, 163–178.

Plahte, E., Mestl, T., Omholt, S., 1998. A methodological basis for description and analysis of systems with complex switch-like interactions. J. Math. Biol. 36, 321–348.

Polynikis, A., Hogan, S., di Bernardo, M., 2009. Comparing different ODE modelling approaches for gene regulatory networks. J. Theor. Biol. 261, 511–530.

Radulescu, O., Lagarrigue, S., Siegel, A., Veber, P., Borgne, M.L., 2006. Topology and static response of interaction networks in molecular biology. J. R. Soc. Interface 3, 185–196.

Reder, C., 1988. Metabolic control theory: a structural approach. J. Theor. Biol. 135, 175–201.

Ropers, D., de Jong, H., Page, M., Schneider, D., Geiselmann, J., 2006. Qualitative simulation of the carbon starvation response in *Escherichia coli*. Biosystems 84, 124–152.

Ropers, D., Baldazzi, V., de Jong, H., 2011. Model reduction using piecewise-linear approximations preserves dynamic properties of the carbon starvation response in *Escherichia coli*. ACM/IEEE Trans. Comput. Biol. Bioinform. 8, 166–181.

Saez-Rodriguez, J., Alexopoulos, L., Epperlein, J., Samaga, R., Lauffenburger, D., Klamt, S., Sorger, P., 2009. Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. Mol. Syst. Biol. 5, 331.

Saier Jr., M.H., Ramseier, T.M., Reizer, J., 1996. Regulation of carbon utilization. In: Neidhardt, F., Curtiss III, R., Ingraham, J., Lin, E., Low, K., Magasanik, B., Reznikoff, W., Riley, M., Schaechter, M., Umbarger, H. (Eds.), *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. , ASM Press, Washington, DC, pp. 1325–1343.

Santner, A., Estelle, M., 2009. Recent advances and emerging trends in plant hormone signalling. Nature 459, 1071–1078.

Sarkar, D., Siddiquee, K., Araúzo-Bravo, M., Oba, T., Shimizu, K., 2008. Effect of *cra* gene knockout together with *edd* and *iclR* genes knockout on the metabolism in *Escherichia coli*. Arch. Microbiol. 190, 558–571.

Sepulchre, J.-A., Reverchon, S., Nasser, W., 2007. Modeling the onset of virulence in a pectinolytic bacterium. J. Theor. Biol. 44, 239–257.

Shlomi, T., Eisenberg, Y., Sharan, R., Ruppin, E., 2007. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. Mol. Syst. Biol. 3, 101.

Silva-Rocha, R., Tamames, J., Santos, V.D., de Lorenzo, V., 2011. The *logicome* of environmental bacteria: merging catabolic and regulatory events with Boolean formalisms. Environ. Microbiol 13, 2389–2402.

Snoussi, E., 1989. Qualitative dynamics of piecewise-linear differential equations: a discrete mapping approach. Dyn. Stab. Syst. 4, 189–207.

Thouvenot, B., Charpentier, B., Branlant, C., 2004. The strong efficiency of the *Escherichia coli gapA* P1 promoter depends on a complex combination of functional determinants. Biochem. J. 383, 371–382.

Usseglio Viretta, A., Fussenegger, M., 2004. Modeling the quorum sensing regulatory network of human-pathogenic *Pseudomonas aeruginosa*. Biotechnol. Prog. 20, 670–678.

Venturi, V., 2003. Control of *rpoS* transcription in *Escherichia coli* and *Pseudomonas*: why so different? Mol. Microbiol. 49, 1–9.

Wittmann, D., Krumsiek, J., Saez-Rodriguez, J., Lauffenburger, D., Klamt, S., Theis, F., 2009. Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. BMC Syst. Biol. 3, 98.

Zgurskaya, H., Keyhan, M., Matin, A., 1997. The σ^S level in starving *Escherichia coli* cells increases solely as a result of its increased stability despite decreased synthesis. Mol. Microbiol. 24, 643–652.