2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Integration of large-scale metabolic, signaling, and gene regulatory networks with application to infection responses

Guilhem Richard, Hyeygjeon Chang, Igor Cizelj, Calin Belta, A. Agung Julius, and Salomon Amar

Abstract— **Mathematical models of biochemical networks, such as metabolic, signaling, and gene networks, have been studied extensively and have been shown to provide accurate descriptions of various cell processes. Nevertheless, their usage is restricted by the fact that they are usually studied in isolation, without feedback from the environment in which they evolve. Integrating these models in a global framework is a promising direction in order to increase both their accuracy and predictive capacity. In this paper, we describe the integration of large-scale metabolic and signaling networks with a regulatory gene network. We focus on the response to infection in mouse macrophage cells. Our computational framework allows to virtually simulate any type of infection and to follow its effect on the cell. The model comprises 3,507 chemical species involved in 4,630 reactions evolving at the fast time scale of metabolic and signaling processes. These interact with 20 genes evolving at the slow time scale of gene expression and regulation. We develop a simulator for this model and use it to study infections with *Porphyromonas gingivalis*.**

I. INTRODUCTION

Signal transduction pathways, such as the Toll-Like Receptor (TLR) signaling pathways, are an essential component of the immune system, which they activate by detecting pathogens and cytokines [1]. The TLRs are highly conserved proteins and recognize specific molecules of bacterial and viral origin. Once triggered, the TLRs initiate several signaling cascades leading to the activation of transcription factors (TFs) that regulate the expression of specific target genes. Depending on the nature of the stimulus (*i.e.* bacterial/viral infection, stress signal), different TFs are activated, which then regulate the necessary response of the cell.

Many studies have emphasized the significance of the TLRs within the immune system. Nevertheless, recent works have shown that they could also provoke deleterious effects in some infections [2], [3]. These infections can significantly increase the concentration of inflammatory cytokines, which produce harmful effects to the host. These studies have suggested that a better response to infection can be obtained in mice deficient in one of the TLRs. Targeting these pathways can lead to a decrease in cytokine concentration and an increase in survival rate. These results demonstrate that targeting the TLRs pathways may prove to be beneficial in fighting infections, and that accurate models of these pathways are necessary in order to find drug targets and to follow the evolution of the response to infection.

Mathematical models of biochemical networks have been used extensively to study the metabolism of bacteria and mammals [4], signaling pathways involved in immunity [5], [6], and gene regulatory interactions [7]. Within the systems and control community, most of the previous works in this area have focused on studying such networks in isolation [8], [9], [10], [11], [12]. However, recent studies show that the integration of mathematical models of different types of biochemical networks, together with environmental conditions, increases the accuracy of the models, as well as their predictive capacity [13], [14].

In this paper, we integrate large-scale metabolic and signaling networks with a regulatory gene network, with the goal of creating a model capable of simulating the response to infection in mouse macrophage cells. To perform this integration, we first merge the iMM1415 [4] mouse genome-scale metabolic model with the *ihs*TLR v1.0 model of the TLR signaling pathways [6] to create a mathematical model with "fast" dynamics. We combine these networks with a regulatory gene network linked to the TLR pathways that we have recently identified [15]. This network was obtained by applying our reconstruction algorithm [16] to gene expression data from [17] to construct a mathematical model for the gene network in the form of a discrete-time piecewise affine system with "slow" dynamics. The overall integration exploits the difference in time-scales, and is based on a stoichiometric steady-state model for the combined metabolic and signaling networks. We use this computational framework to study infections with *P. gingivalis* (*P.g*), a common pathogen involved in periodontal disease.

To the best of our knowledge, this is the first attempt to construct a large scale mathematical model for the response to infection that incorporates metabolism, signaling, and gene regulation. The model is not limited to *P.g*-infection and it can be used to simulate virtually any type of infection by triggering different combinations of TLRs. Our framework allows to follow simultaneously the effects of an infection on the signaling pathways of the immune system, metabolism, and gene expression. It can also be used to assess the effects of gene deletions in the response to infection.

The remainder of this paper is organized as follows. In Sec. II, we provide a short overview of stoichiometric quasi-steady state analysis. The mathematical models that we use for the metabolic, signaling, and gene regulatory networks are described in Sec. III and their integration is given in Sec. IV. Simulation results are included in Sec. V.

This work was partially supported by grant R01 DE15989 from the National Institutes of Health.

G. Richard, H. Chang, I. Cizelj, C. Belta, and S. Amar are with Boston University, Boston, MA 02215, USA {grichard, hjchang, icizelj, cbelta, samar}@bu.edu A.A. Julius is with the Dept. of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA agung@ecse.rpi.edu

II. STOICHIOMETRIC MODELING AND QUASI-STEADY STATE ANALYSIS

In this section, we provide a short description of quasi-steady state analysis of stoichiometric models. We represent a mass-balanced network involving m species and r reactions with a stoichiometric matrix $S \in \mathbb{R}^{m\times r}$. Each entry S_{ij} specifies the stoichiometric coefficient for species i in reaction j. Let $\mathcal{M} = \{1,\dots,m\}$ and $\mathcal{R} = \{1,\dots,r\}$ denote the set of all species and reactions in the network, respectively. Let $x(t) = (x_1(t),\dots,x_s(t))$ and $v(t) = (v_1(t),\dots,v_r(t))$ denote the vectors of all species concentrations and all reaction fluxes, respectively, at time t. The rates of change of species concentrations are governed by:

$$\dot{x} = Sv, \alpha \le v \le \beta, \tag{1}$$

where $\alpha, \beta \in \mathbb{R}^r$ are the lower and upper bounds of v. Models of type (1) have been widely used for genome-scale metabolic networks. For such systems, the set of reactions is given by all known reactions of small molecule biochemistry occurring inside the cell, including transport and core metabolism. The set of species include external (extracellular) metabolites such as oxygen, hydrogen, or glucose, which are typically essential for cell survival and require transport reactions to traverse cell membranes.

Flux Balance Analysis (FBA) is a method used extensively in the study of metabolic networks [18]. FBA assumes that (1) is in a quasi-steady state since metabolic reactions are fast when compared to gene regulation dynamics [6], [19]. Under this assumption, (1) becomes:

$$Sv = 0, \alpha \le v \le \beta. \tag{2}$$

Let

$$K = \{v \in \mathbb{R}^r | Sv = 0, \alpha \le v \le \beta\} \tag{3}$$

denote the set of all feasible fluxes, which can be easily shown to be a polyhedral cone. A reaction $o \in \mathcal{R}$ is called *producible* if and only if

$$\exists v \in K \text{ s.t. } v_o > 0. \tag{4}$$

The essence of FBA is the definition of a "biomass" reaction (the flux through this reaction is usually denoted by μ) that models the growth of the cell. In other words, a biomass reaction takes as substrates all the small molecule species that are necessary for the survival and growth of the cell. The stoichimetric coefficients are their relative proportions. A flux distribution through all the reactions in the network can then be easily found by maximizing μ under the constraint that the fluxes are feasible. Computationally, this leads to a Linear Program (LP). The maximal growth assumption is reasonable in bacteria, where it can be assumed that cells tend to grow as fast as possible [13], [19]. Biomass functions, and objective functions in general, are hard to formulate in the case of multicellular organisms where cellular objectives depend primarily on the cell type.

Dynamic Flux Balance Analysis (dFBA) is a method which captures the use of a media by a population of cells [19]. In brief, dFBA uses FBA over multiple iterations to determine the consumption of external metabolites and the growth of a cell population in time. Let $E \subset \mathcal{M}$ denote the set of all external metabolites. We assume that there is exactly one transport reaction for each external metabolite. Therefore, E also denotes the set of transport reactions. Let $B(t)$ denote the cell concentration at time t. The equations governing the cell concentration and extracellular metabolite concentrations are given by:

$$\begin{aligned} \dot{B}(t) &= \mu B(t), \\ \dot{x}_e(t) &= -v_e B(t), e \in E. \end{aligned} \tag{5}$$

Assuming that there are T time steps of equal length h in the simulation, the algorithm proceeds as follows:

a) Initialization: $B(0)$, $x_e(0)$,

$$\beta_e(0) = \frac{x_e(0)}{B(0)h}, e \in E; \tag{6}$$

b) Iteration: At each time step $n \in \{0,\dots,T-1\}$, FBA is performed to maximize μ. The solution will give specific values for all the fluxes, including those of the transport reactions. The general discrete-time form of (5) during the time interval $[nh,(n+1)h)$ is given by

$$\begin{aligned} B((n+1)h) &= B(nh)e^{\mu h} \\ x_e((n+1)h) &= x_e(nh) + \frac{v_e B(nh)}{\mu}(1-e^{\mu h}), e \in E, \end{aligned} \tag{7}$$

with $B(nh)$ and $x_e(nh)$ known from the previous iteration and v_e determined from the solution to the FBA. $B((n+1)h)$, $x_e((n+1)h)$, and $\beta_e((n+1)h) = \frac{x_e((n+1)h)}{B((n+1)h)h}$, $e \in E$ are computed and the procedure is reiterated.

The time step h is typically several minutes [14], which is in accordance with the growth dynamics. Since metabolism is much faster, the quasi-steady state assumption is justified.

III. MODELS AND INTERACTIONS

In this section we describe the different models that we use for metabolism, signaling, and gene regulation. The starting point for our approach is the observation that gene regulations are slow in comparison to metabolic and signaling reactions. Indeed, metabolic and signaling reactions take less than 1 sec, while transcriptional regulation and receptor internalization take on the order of 10^2 sec [5]. This motivates the separation of the overall dynamics into fast and slow dynamics (see Fig. 1). We first present the metabolic and signaling networks and their integration into a single model with fast dynamics. Then, we briefly review the derivation of a mathematical model for the gene regulatory network by applying our recently developed identification method on the gene expression dataset from [17].

A. The fast dynamics: signaling and metabolism

Our metabolic model is based on the recently published mouse iMM1415 metabolic model [4]. This model, which consists of 2,775 metabolites involved in 3,727 reactions, is a comprehensive description of the model metabolism, *i.e.* it contains all the metabolic reactions that occur in mouse cells, regardless of the cell type. The signaling part of our model

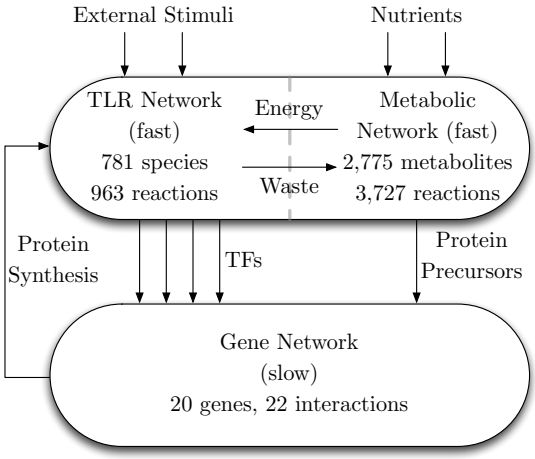

Fig. 1. Schematic representation of the overall model. The fast dynamics comprise reactions from the signaling pathways (TLR network) and from the metabolic network. Metabolism produces energy that is used by the TLR network. The waste produced by the phosphorylation reactions in the signaling network is recycled by metabolism. In addition to producing energy, metabolism forms all the precursors necessary for protein synthesis. Gene expression is controlled by transcription factors (TFs) whose activation depends on the outcome of the signaling pathways. In response, the gene network synthesizes proteins used in both signaling and metabolic reactions.

is based on the *ihs*TLR v1.0 model of the TLR signaling pathways [6]. This describes the reactions involved in the transmission of a stimulus from the TLRs to the nucleus. The system is triggered by 14 inputs and leads to the activation of six different outputs: AP-1, CREB, IRF3, IRF7, Reactive Oxygen Species (ROS), and NF-κB. With the exception of ROS, the outputs are transcription factors (TFs) that regulate the expression of several other genes. These compounds have been shown to play a major role in the immune system [20]. The *ihs*TLR v1.0 model comprises 781 species involved in 963 reactions [6].

Since the metabolic and signaling networks have similar time scales [5], we merged the two networks into a single combined metabolic/signaling (*fast*) network, which consists of three types of reactions (i) reactions specific to metabolism, (ii) reactions specific to signaling pathways, and (iii) reactions shared by metabolic and signaling pathways. Merging the two networks required some careful analysis in order to be able to construct a coherent stoichiometric model of the form given in Eqn. (2). Specifically, in the case of a shared reaction, we only keep the metabolic reaction in order to be able to use the flux bounds defined in the metabolic model. If there are several signaling reactions which jointly correspond to a metabolic reaction, we keep the signaling reactions to preserve more signaling pathways, and delete the metabolic reaction. We also delete several import reactions specific to the signaling network.

As a result of the merging procedure, we obtained a fast network described by a quasi-steady state stoichiometric model of the form (2), where S has 3,507 rows and 4,630 columns. As it will become clear in Sec. IV, we use this model as part of a modified dFBA procedure (see Sec. II), which captures gene regulatory events in addition to external metabolite availability and cell growth. In fact, since macrophages do not grow, we have $\mu = 0$, which implies $B(t) = B(0)$. As a result, Eqn. (7) reduces to:

$$x_e((n+1)h) = x_e(nh) + v_e Bh, e \in E. \tag{8}$$

The objective function that we maximize in the FBA problem that we solve at each step is a combination of three components: (1) a cell maintenance component, which uses 41 metabolites, such as amino acids and lipids [4], (2) a signaling component, which maximizes a linear combination of the signaling reactions producing active outputs, and (3) a protein synthesis component, which captures the transcription and translation processes, and which is modulated according to the activity of the gene network.

B. The slow dynamics: gene regulation and expression

We first briefly review the gene network reconstruction algorithm from [16] and its application to construct a discrete-time piecewise linear mathematical model for a gene network that interacts with the TLR signaling pathways [15].

1) Gene network reconstruction algorithm: Assume we are given experimental data for the genes in a set $\mathcal{G}$ as time-series expression data at $N+1$ time points in the form $x_{g,n}$, $g \in \mathcal{G}, n = 0, \ldots, N$. Our goal is to construct a mathematical model for the gene network dynamics that is compatible with the gene expression data. We focus on a particular class of discrete time systems of the following form

$$\begin{aligned} x_g((n+1)H) = x_g(nH) \quad &+H\Bigg[\sum\nolimits_{k\in\mathcal{G}_g^R} f_{g,k}(x_k(nH)) \\ &-\lambda_g x_g(nH)\Bigg], \end{aligned} \tag{9}$$

where $x_g(nH)$ denotes the concentration of protein expressed from gene $g \in \mathcal{G}$ at time nH, $\lambda_g \geq 0$ is its decay rate, $\mathcal{G}_g^R$ is the set of regulators for gene g, $f_{g,k}(\cdot)$ is a function describing the regulation of gene g by gene k, and H is the time step. We assume that each function $f_{g,k}(\cdot)$, $g \in \mathcal{G}$ and $k \in \mathcal{G}_g^R$, is continuous, non-negative, and monotone. The time step H is chosen in accordance with the slow time scales of gene regulation. As it will become clear in Sec IV, H will be an integer multiple of the time step h from Eqn. (8).

The set of regulator genes $\mathcal{G}_g^R$ for all $g \in \mathcal{G}$ are identified by checking the feasibility of a Linear Programming (LP) problem. From an implementation viewpoint, it is more efficient to reformulate the LP as a Quadratic Programming (QP) problem based on the definition of slack variables $\epsilon := (\epsilon f, \epsilon\Delta)$ (see [16] for detailed LP and QP formulations).

It is important to note that, as a byproduct of the QP problem, we get numerical values for the decay rates λ_g and the regulation functions $f_{g,k}$ at the time points corresponding to the experimental data. Given that the gene expression data is over the same time points, the latter can be easily converted to a finite number of values for each function $f_{g,k}(x_k)$. By simple linear interpolation of these values, we can construct a piecewise linear model of the form given in Eqn. (9).

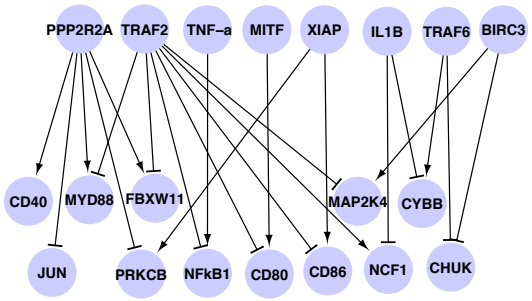

Fig. 2. Reconstructed genetic regulatory network [15]. The 8 genes in the upper region and the 12 genes in the lower region correspond to the genes in $\mathcal{G}_I$ and $\mathcal{G}_O$, respectively.

2) Application of the gene network reconstruction algorithm: We describe in [15] the application of the reconstruction algorithm to reconstruct a gene network that interacts with the TLR signaling pathways. Briefly, the first step for the construction of the gene network shown at the bottom of Fig. 1 is the selection of the set of genes $\mathcal{G}$. This is, in fact, an iterative process. By using the KEGG database (http://www.genome.jp/kegg/) and the *ihs*TLR v1.0 model, we selected a set of 9 genes that are directly regulated by the TFs that are activated as part of the TLR signaling network. This set of genes form the "input layer" $\mathcal{G}_I$ of our gene network. Similarly, we identified a set of 31 "output layer" genes $\mathcal{G}_O$ that code for proteins involved in the TLR signaling pathways. Our goal is to connect each gene in the input layer to at least one gene in the output layer through regulation interactions, possibly using some other intermediate genes. As noted in [15], we considered regulation of the output genes by (sets of) input genes.

Fig. 2 shows the result of the reconstruction of the gene regulatory network [15]. We can see that each gene in $\mathcal{G}_I$ is connected to at least one gene in $\mathcal{G}_O$. A total of 12 output genes are linked to the input layer. The discrete-time piecewise affine model is omitted due to space constraints.

IV. INTEGRATED FRAMEWORK

A. Network Interconnections

The integrated overall model combines the fast dynamics of the metabolic and signaling pathways with the slow dynamics of the gene regulatory network. The interactions among these networks occur at four levels (see Fig. 1).

1) Metabolism to Signaling: Metabolism creates energy in the form of molecules of ATP, which is an essential metabolite in signaling pathways. ATP is most often used during phosphorylation events. These reactions produce ADP as a by-product, which is recycled by metabolism to create new molecules of ATP. The combined metabolic/signaling network integrates theses connections.

2) Metabolism to Gene regulation: In addition to creating energy, metabolism uses nutrients to produce all the necessary building blocks of the cell. These components are essential for protein synthesis. The demand of these precursors fluctuates according to the expression of the genes in the regulatory network. The objective function used during FBA analysis takes this demand into consideration.

3) Signaling to Gene regulation: The overall goal of the signaling pathways is to transmit external stimuli to the nucleus. This is done through the activation of TFs, which regulate the expression of target genes. If TF_o enhances the expression of gene $g \in \mathcal{G}_I$, then:

$$\dot{x}_g(t) = \begin{cases} \rho + \rho_0 - \lambda_g x_g(t) & \text{if } \delta TF_o > 0 \\ \rho_0 - \lambda_g x_g(t) & \text{if } \delta TF_o \leq 0, \end{cases} \quad (10)$$

where ρ and ρ_0 are the production and basal production rates of g, respectively, and λ_g is its decay rate. These parameters can be easily determined from the gene expression data from [17]. δTF_o is positive if the rate of activation of TF_o increases through time, negative if it decreases. The direction of the inequalities is reverted if TF_o represses g. In our implementation, we use a discretized version of (10).

4) Gene regulation to Signaling and Metabolism: The regulatory constraints from the gene network are incorporated in the fast dynamics through an *incidence matrix*. This matrix was first defined in [14] as a Boolean matrix of size $r \times T$, with r the number of reactions in the fast network and T the number of time steps in the simulation. Each entry $I(j,n)$ specifies if reaction $j \in \mathcal{R}$ is enabled or not at the n-th time step. At time step n, the value in $I(j,n)$ is used to constraint the lower and upper bounds of v_j:

$$\alpha_j I(j,n) \leq v_j \leq \beta_j I(j,n). \quad (11)$$

$I(j,n) = 0$ is equivalent to removing reaction j. I is updated at every iteration according to the state of the gene network. In this paper, we extend the framework from [14] by considering an incidence matrix with entries varying in a continuous range from 0 to 1. This extension is necessary to account for the continuous changes in the expressions of the genes from the regulatory network. Let $J \subseteq \mathcal{R}$ denote the set of reactions *enabled* by a gene. A gene enables a reaction j if it codes for a protein, typically an enzyme, used in j. In our model, a reaction is only enabled by a single gene. Therefore, J also denotes the set of genes enabling reactions. The expression of gene j directly influences the activity of reaction j. To capture this constraint, we consider the normalized gene expression of gene j:

$$I(j,n+1) = \frac{x_j(n) - x_j^{min}}{x_j^{max} - x_j^{min}}, j \in J, \quad (12)$$

where x_j^{min} and x_j^{max} are the minimum and maximum expression values for gene j, respectively, derived from the expression data in [17]. Using (11), a low gene expression for gene j does not completely shut down reaction j, but rather diminishes its maximum capacity. For all reactions $j \in \mathcal{R} \setminus J$, $I(j,n) = 1$ at all time steps.

B. Overall simulation algorithm

The framework for the integration follows the procedure from [14]. Our implementation differs from this approach by considering gene expression as a continuous variable. A schematic representation of the framework is shown in Fig. 3.

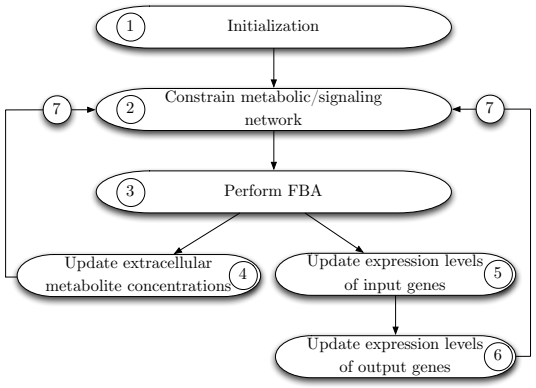

Fig. 3. Integrated framework: 1) The simulation starts by initializing the biomass B, the extracellular metabolite concentrations x_e, and the gene expression levels of the regulatory network x_g. 2) The biomass and the extracellular metabolite concentrations are used to compute the upper bounds β_e for the transport reactions. The gene expression levels determine the incidence matrix I at the current time step. I is used to constrain the lower and upper bounds of all reactions in the metabolic/signaling network. 3) FBA is performed to obtain flux values for all the signaling and metabolic reactions; in particular, it gives flux values for the transport reactions and for the outputs of the signaling pathways. 4) The fluxes of the transport reactions give the new external metabolite concentrations. 5) The outputs of the signaling pathways determine the expression levels of the input genes (*i.e.* the genes in $\mathcal{G}_I$). 6) The output genes (*i.e.* the genes in $\mathcal{G}_O$) are updated according to the state of the input genes. The update is done with some delay τ. 7) The new extracellular metabolite concentrations and expression levels are used to constrain the metabolic/signaling network for the next iteration.

a) Initialization: B, $x_e(0)$, $\beta_e(0) = \frac{x_e(0)}{Bh}$, $e \in E$, x_g^{min}, $x_g^{max}, g \in \mathcal{G}$. $x_g(0)$ are either specified or generated randomly. In the latter case, an expression value is arbitrarily chosen between x_g^{min} and x_g^{max}. $I(j,0) = \frac{x_j(0)-x_j^{min}}{x_j^{max}-x_j^{min}}, j \in J$. FBA is performed at time 0 on the unstimulated model (*i.e.* without infection) to determine the "baseline" flux value $v_{TF_o}(0)$ for each output o of the signaling pathways.

b) Iteration: At each time step $n \in \{0, \ldots, T-1\}$:

1) Constrain the fast dynamics: The regulatory constraints defined in I are included in the fast dynamics using (11). An infection is modeled by constraining the input reactions of the signaling pathways: the flux of an input i is constrained to $1 \leq v_i \leq \beta$ for a predefined duration. If no infection is present, we apply the constraint $v_i = 0$.

2) FBA: FBA is performed to maximize the objective function described in section III-A.

3) Update the concentrations of external metabolites: $x_e((n+1)h), e \in E$ is computed with (8). $\beta_e((n+1)h) = \frac{x_e((n+1)h)}{Bh}$, $e \in E$. The biomass remains unchanged since macrophages do not grow.

4) Determine the activity of the TFs: $\delta TF_o = v_{TFo}((n+1)h) - v_{TFo}(0)$. δTF_o being positive, zero, or negative corresponds to an increased, stable, or decreased rate of activation of TF_o, respectively.

5) Update gene expression levels: The time scale H of the gene network is much larger than the time step h used during dFBA [14]. Nevertheless, the gene network must capture the events of the signaling network, *i.e.* the activations of the TFs, which happen at the small time step. Gene expression is thus updated at every time step h. Each gene $g \in \mathcal{G}_I$ is updated using the discrete-time version of (10), with h as the constant integration time step. The expression level of the genes $g \in \mathcal{G}_O$ is updated using (9). We introduce a delay during the update procedure to account for the gene slow dynamics (*i.e.* H). The state of each gene is updated with some delay τ, where τ is the number of time steps required for a regulator to change the expression level of a target gene. The time scale H in (9) is then expressed as $H = \tau h$.

6) Update I: I is updated according to (12) for all reactions $j \in J$, and the procedure is reiterated.

V. SIMULATIONS OF INFECTIONS WITH *Porphyromonas gingivalis*

We first simulated the system without infection to check whether or not the system could reach a steady state. We ran 10 simulations with $T = 1,000$, $h = 6$ minutes, and $\tau = 30$ ($H = \tau h = 3$ hours). Gene expression was initialized randomly. Most of the genes reached a steady state value after 150 iterations. Interestingly, in all the runs, the systems reached the same steady state after the 1,000 time steps. The values corresponding to this steady state were used as initial conditions for all further simulations. As a sanity check, we started a simulation at the steady state. Gene expression levels stayed unchanged throughout the simulation.

P.g is known to engage the TLR1/TLR2, TLR2, TLR4, and TLR2/TLR6 dimers [21], [22]. To simulate an infection, we forced the fluxes through the reactions activating these receptors to be greater than 1. During the simulations, *P.g* infections were able to increase the activation rates of AP-1, CREB, and NF-κB. As a consequence, these TFs enhance the expression of TNF-α. Our observations partially match previous studies which have reported increased activation of AP-1, NF-κB, and production of TNF-α [23]. Nevertheless, instead of observing an increase of ROS production, as it is usually the case in *P.g* infections [24], the simulated infection provoked a decrease of ROS. This discrepancy may arise from the objective function considered during the FBA optimization. This function maximizes the production of all outputs, and it appears that it becomes more "efficient" to increase the production of NF-κB at the expense of ROS. This is a clear limitation of the FBA model.

Fig. 4 shows the time course of TNF-α expression in our simulated data and in the experimental data from [17]. Both simulations and experiments produced similar results. The expression level of TNF-α increased shortly after the initiation of the infection. Both expression levels reached a comparable plateau. The infection remained active in our simulation for 10 hours. Once the infection was removed, the expression level of TNF-α quickly returned to its original steady state value. This step took 1.5 hours in our simulation, compared to $\sim$20 hours in the biological experiments. This difference suggests that the decay rate in our simulations might have been too large.

We next assessed the effect of the deletion of MyD88 on the response to *P.g* infections. MyD88 is known to be a critical gene involved in the response to infection. It has been

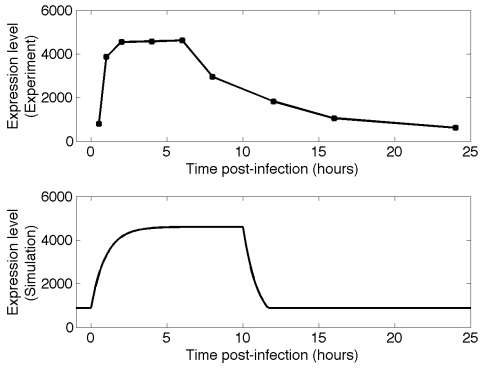

Fig. 4. Expression levels of TNF-α. The top plot shows the expression of the gene as present in the experimental data in [17]. The bottom plot presents the expression as obtained in our simulation. The same trend is observed in both cases.

reported that its deletion disrupts major cell responses [23]. In our simulation, the knockout of MyD88 disrupted several pathways from the TLR network: the deletion rendered the reactions activating the TLR1/TLR2, TLR4, and TLR2/TLR6 inputs non-producible. Among the four inputs triggered by *P.g*, only TLR2 remained producible after the deletion of MyD88. In addition, the network could no longer sustain the activation of AP-1.

VI. CONCLUSION

We described a computational framework for the integration and simulation of the TLR signaling network with the mouse genome scale metabolic network and a gene network that regulates the expression of some genes involved in the TLR pathways. The mathematical model for the gene regulatory network was obtained through a new identification method applied to a recently published gene expression data set for mouse immune cells. Our tool can simulate any type of infection. We present simulation results for infection with *P.g*, which are similar to biological observations.

REFERENCES

[1] T Kaisho and S Akira. Toll-like receptor function and signaling. *J Allergy Clin Immunol*, 117(5):979–87; quiz 988, May 2006.

[2] EA Kurt-Jones et al. Herpes simplex virus 1 interaction with toll-like receptor 2 contributes to lethal encephalitis. *Proc Natl Acad Sci USA*, 101(5):1315–20, Feb 2004.

[3] BB Gowen et al. Tlr3 deletion limits mortality and disease severity due to phlebovirus infection. *J Immunol*, 177(9):6301–7, Nov 2006.

[4] MI Sigurdsson et al. A detailed genome-wide reconstruction of mouse metabolism based on human recon 1. *BMC Syst Biol*, 4(1):140, Oct 2010.

[5] JA Papin and B O Palsson. The jak-stat signaling network in the human b-cell: an extreme signaling pathway analysis. *Biophys J*, 87(1):37–46, Jul 2004.

[6] F Li et al. Identification of potential pathway mediation targets in toll-like receptor signaling. *PLoS Comput Biol*, 5(2):e1000292, Feb 2009.

[7] I Shmulevich et al. Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2):261–74, Feb 2002.

[8] TM Yi et al. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc Natl Acad Sci USA*, 97(9):4649–53, Apr 2000.

[9] E Sontag et al. Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data. *Bioinformatics*, 20(12):1877–86, Aug 2004.

[10] D Del Vecchio and ED Sontag. Engineering principles in biomolecular systems: From retroactivity to modularity. *European Journal of Control (Special Issue)*, 2009.

[11] N Motee et al. Performance limitations in autocatalytic networks in biology. *49th IEEE Conference on Decision and Control (CDC), 2010*, pages 4715 – 4720, 2010.

[12] M Imielinski and C Belta. Deep epistasis in human metabolism. *Chaos*, 20(2):026104, Jun 2010.

[13] A Goelzer et al. Reconstruction and analysis of the genetic and metabolic regulatory networks of the central metabolism of bacillus subtilis. *BMC Syst Biol*, 2:20, Jan 2008.

[14] JM Lee et al. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput Biol*, 4(5):e1000086, May 2008.

[15] H Chang et al. An application of monotone functions decomposition to the reconstruction of gene regulatory networks. *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2011.

[16] A Julius and C Belta. Genetic regulatory network identification using monotone functions decomposition. *18th IFAC World Congress, 2011*.

[17] I Amit et al. Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science*, 326(5950):257–63, Oct 2009.

[18] JS Edwards et al. In silico predictions of escherichia coli metabolic capabilities are consistent with experimental data. *Nat Biotechnol*, 19(2):125–30, Feb 2001.

[19] A Varma and BO Palsson. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type escherichia coli w3110. *Appl Environ Microbiol*, 60(10):3724–31, Oct 1994.

[20] R Medzhitov and T Horng. Transcriptional control of the inflammatory response. *Nat Rev Immunol*, 9(10):692–703, Oct 2009.

[21] K Tabeta et al. Toll-like receptors confer responsiveness to lipopolysaccharide from porphyromonas gingivalis in human gingival fibroblasts. *Infect Immun*, 68(6):3731–5, Jun 2000.

[22] Ge Hajishengallis et al. Differential interactions of fimbriae and lipopolysaccharide from porphyromonas gingivalis with the toll-like receptor 2-centred pattern recognition apparatus. *Cell Microbiol*, 8(10):1557–70, Oct 2006.

[23] Q Zhou and S Amar. Identification of signaling pathways in macrophage exposed to porphyromonas gingivalis or to its purified cell wall components. *J Immunol*, 179(11):7777–90, Dec 2007.

[24] DV Sculley and SC Langley-Evans. Salivary antioxidants and periodontal disease status. *Proc Nutr Soc*, 61(1):137–43, Feb 2002.