Modelling Bistable Genetic Regulatory Circuits Under Variable Volume Framework

G. Maria
Laboratory of Chemical & Biochemical Reaction Engineering,
University Politehnica of Bucharest, P.O. 35-107 Bucharest, Romania
Email: gmaria99m@hotmail.com

Paper dedicated to the memory of
Prof. Wolf-Dieter Deckwer (TU Braunschweig)

Adequate modelling of Genetic Regulatory Circuits (GRC) allows a deeper understanding of the regulatory and control mechanism of gene expression in living cells, but also *in-silico* design of synthetic cell-systems exhibiting desired mini-functions (i.e. motifs, such as bistable-switches, oscillators, amplitude filters, etc.), with various practical applications in medical, industrial, or environmental fields. Modular lumped dynamic models have been reported as being valuable tools to adequately reproducing a wide-range of cell nonlinear dynamics, such as saturation, inhibition, sigmoidals, multiple steady-states, stable oscillations. In the present work, the analysis of a bistable-switch formed by two gene-expression modules is performed in a variable-volume and isotonic modelling framework, by mimicking the *E. coli* cell growth. A combination of lumped models, of Hill-type activation – repression, with including quick buffering reversible reactions using dimeric intermediates, is proved to offer a more flexible representation of the bistable genetic-switch than the classical lumped power-law approach. Intermediate species, of adjustable levels, allow a fine-tuning of GRC properties, in terms of stability strength, responsiveness and selectivity to external stimuli, regulatory efficiency and species connectivity in the gene-expression modules, under stationary and dynamic perturbations.

Key words:
Bistable switch, genetic regulatory circuits, dynamic models, variable volume

Introduction

Cell metabolism is characterized by a high complexity, including a large number of species, intermediates, elementary reactions, complex interactions and regulatory networks, all organised in a structural, functional, and temporal hierarchy. Detailed modelling of evolvable and adaptable cell metabolic processes with conventional approach is a difficult task, due to the known system low observability, indentifiability, and reproducibility. However, the increased availability of -omics data (genomics, proteomics, metabolomics, fluxomics) lead to a considerably progress in elaborating topological and dynamic models for some metabolic processes or for even the whole-cell evolution, and valuable structural kinetic models with accounting for individual or lumped components have been reported:[1,2] single cell growth models, oscillatory metabolic pathways, genetic regulatory networks/circuits, cell cycles, cell communications, cell-drug interactions, intracellular signalling, neuronal transmission, analysis of 'logical essence' of life, life minimal requirements, etc. Extensive simulation platforms have been developed based on cell databanks, such as: EcoCyc[3] databank on *E. coli* cell; KEGG[4] databank on 36 eukaryotes, 415 bacteria and 31 archaea; BRENDA[5] databank on enzymes; Prodoric[6] databank on gene regulation in prokaryotes, etc. Modelling studies not only allow a deep understanding of the cell process mechanisms, but also lead to immediate practical applications in industrial and medicine fields: new strains with improved performances in production of various bioproducts in food, agriculture, drug-industry or medicine, improved biotreatment of pollutants in environmental engineering.

The emergent *Synthetic Biology,*[7] interpreted as the engineering-driven building of complex biological entities, aims at applying engineering principles of systems design to biology with the idea to produce predictable and robust systems with novel functions in a broad area of applications,[8,9] such as therapy of diseases (gene therapy), design of new biotechnological processes, new devices based on cell-cell communicators, biosensors, etc. Encouraging results have been reported for the design of arti-

ficial gene networks,[10] for reprogramming signalling pathways,[11] for refactoring of small genomes,[12] or re-design of metabolic fluxes with using switching genes.[13,14,15] By assembling functional parts of an existing cell, such as promoters, ribosome binding sites, coding sequences and terminators, protein domains, or by designing new gene regulatory networks on a modular basis, it is possible to reconstitute an existing cell (the so-called 'integrative understanding') or to produce novel biological entities with new properties.[9] The genetic components may be considered as 'building blocks' because they may be extracted, replicated, altered, and spliced into the new biological organisms.

One particular application of such dynamic models is the study of genetic regulatory networks (GRN), in order to predict the way in which biological systems are self-regulated and respond to signals. The emergent field of such efforts is the so-called '*gene circuit engineering'* and a large number of examples have been reported with *in-silico* creation of novel GRC conferring new properties/functions to the mutant cells (i.e. desired 'motifs' in response to external stimuli), such as:[9,16–24]

– toggle-switch, i.e. mutual repression control in two gene expression modules, and creation of decision-making branch points between on/off states according to the presence of certain inducers;

– hysteretic GRC behaviour, that is a bio-device able to behave in a history-dependent fashion, in accordance to the presence of a certain inducer in the environment;

– GRC oscillator producing regular fluctuations in network elements and reporter proteins, and making the GRC to evolve among two or several quasi-steady-states;

– specific treatment of external signals by controlled expression such as amplitude filters, noise filters or signal / stimuli amplifiers;

– GRC signalling circuits and cell-cell communicators, acting as 'programmable' memory units.

The difficult task to design complex biological circuits with a building blocks strategy can be accomplished by properly defining the basic components, functions, structural organisation. Because many cell regulatory systems are organized as modules (up to 23-25),[25] a modular design strategy including engineered GRC leads to desirable characteristics:[17] a tight control of gene expression (i.e. low-expression in the absence of inducers and accelerated expression in the presence of specific external signals); a quick dynamic response and high sensitivity to specific inducers; gene circuit robustness, i.e. a low sensitivity vs. undesired inducers (external noise). Combination of induced motifs in modified cells one may create potent applications in industrial, environmental, and medical fields (e.g. biosensors, gene therapy). Valuable implementation tools of the design GRC in real cells have been reported over the last years.[9]

The modular GRC dynamic models, of an adequate mathematical representation, seem to be the most comprehensive mean for a rational design of the regulatory GRC with desired behaviour.[22] However, the lack of detailed information on reactions, rates and intermediates make the extensive representation of the large-scale GRC difficult for both deterministic and stochastic approach.[21] When continuous variable dynamic models are used, the default framework is that of a constant volume/osmotic pressure system, accounting for the cell-growing rate as a 'decay' rate of key-species (often lumped with the degrading rate) in a so-called 'diluting' rate. Such a representation might be satisfactory for many applications, but not for accurate modelling of cell regulatory / metabolic processes under perturbed conditions, or for division of cells, distorting the prediction quality. The variable-volume modelling framework, with explicitly linking the volume growth, external conditions, osmotic pressure, cell content ballast and net reaction rates for all cell-components, is proved as being more promising in predicting local and holistic properties of the metabolic network,[1,2,26] while the classical one tends to over-estimate some of the regulatory dynamic properties.[27,28]

The paper objective is to investigate the properties of a bistable-switch formed by two gene-expression modules, when placed in a cell of known characteristics, and when a variable-volume whole-cell modelling framework is used instead of the analysis of individual modules take out of cell. Exemplification and comparison is made for Hill-type cross-/self-repression models and for a *E. coli* cell framework. Some terms of a lumped GRC model have been detailed in order to obtain a more flexible representation of the switch with desired properties in terms of: responsiveness, local stability, selectivity, robustness, regulatory efficiency, species connectivity. The explicit constraint of isotonic osmolarity under variable cell-volume, and the whole-cell approach with including the inertial and ballast effects under perturbed conditions, can offer a more detailed analysis of the GRC properties.

Modelling genetic regulatory circuits (GRC)

From the mathematical point of view, a large variety of lumped dynamic models have been developed to represent GRC in various cells.[29] Most of these models consider rather the transcriptional control than the translation as governing the genetic ex-

pression regulation, use gene-modular approach constructions, are based on continuous/discrete variables in a deterministic/stochastic representation (or a mixture of them), account for rather lumped elementary reactions and species (aggregate pools), and try to separate the fast and slow parts of the kinetics.[1,2]

Due to the metabolic cell process high complexity and low observability, identifiability, and reproducibility, a worthy route to develop models is to base the analysis on the concepts of '*reverse engineering*' and '*integrative understanding'* of the cell system.[29,30] Such a rule seeks for disassembling the whole system in parts (functional modules), which can be individually studied and characterized. Then, following an appropriate linking algorithm, and performing predictive tests and property / sensitivity analysis, the whole GRN is recreated for reproducing the real system. Such an approach allows reducing the model complexity by relating the cell response to stimuli to the response of only few regulatory loops, instead of the response of thousands of regulatory circuits in gene expression and metabolic pathway. The lumping approach disadvantages concern the decrease in the model predictability, loss of information on some intermediate species and elementary reactions, loss in model properties (holistic optimality, flexibility, multiplicity, sensitivity, regulatory characteristics), and a lack of physical meaning for some model terms and parameters. Derivation of a satisfactory model is closely related to the ability of selecting the suitable lumping rules, key-parameters, and influential terms that better realize a trade-off between model simplicity and its predictive quality.[2]

Semi-autonomous lumped modules are elaborated for representing various regulatory units used in gene expression / protein synthesis, and then linked to efficiently cope with cell perturbations, and to ensure an equilibrated growth during the cell cycle, with an optimised resource consumption (substrate, metabolic energy). This approach allows reducing the analysis complexity by investigating individual modules, and then to relate them to the holistic cell properties. Moreover, the rule is computationally tractable, allowing application of numerical algorithms from the automated system theory. By simulating the transcriptional mechanism and gene interactions, the architecture of the cell regulatory network can be related to the physiological characteristics of the organism.[29,31]

Binary Boolean variable models assume that genes, or other states, exist in only discrete states.[29] In this representation, the genes are either ON or OFF, that is fully expressed or not expressed. Even less realistic, simple Boolean logical functions link the genes in a coarse representation of the GRC (electronic circuit like). This intuitively modelling framework is easily expandable to larger scale, implying no stability issues to be dealt with. As disadvantages, the Boolean models cannot capture some GRN behaviour, such as negative feedback loops (increasing the network stability, as the oscillation theory proves),[32] processes randomness, or macromolecular transport and heterogeneity. The rich dynamical structure of Boolean models, predicting a large number of attractor states and cycles, is however rarely observed in the nature, while some predicted steady-states by continuous models may have no analogy in the Boolean representation.

The continuous variable models, including ordinary differential balance equations (ODE), are suitable to represent metabolic kinetics under various continuous perturbations, binding thermodynamics, cell-division cycle, oscillatory processes, or molecular diffusion. Due to the large size of the identification problem and the time-course data type, co-regulated genes are clustered together, and an unstructured or structured model is derived. Such an extended model can include $O(10^3–10^4)$ number of states (usually species concentrations, mRNA-levels), $O(10^3)$ gene transcription factors TF, while the number of parameters corresponds in principles to the number of states multiplied with the number of TF-s. However, because the biological network is known to be sparsely interconnected, most of such model parameters will be zero, the identification problem becoming a problem of simultaneous structure and parameter identification.

The difficulty to precise the very large number of parameters in complex GRN leads to account for lumped representations, such as gene clustering and path structure reduction based on various system constraints (stability, sensitivity, multiplicity).[2,29,30] For instance, simple rate expressions of power-law type (the so-called S-systems) are obtained by recasting the elementary steps and intermediates in a lumped representation including apparent rate constants and reaction orders:[33,34]

$$\frac{dc_j}{dt} = \alpha_j \sum_{i=1}^{n_S} c_i^{g_{ij}} - \beta_j \sum_{i=1}^{n_S} c_i^{h_{ij}} \qquad (1)$$

(where: c_j = species (lump, or 'pool') concentration; α, β, g, h = vectors of model parameters). Even if the resulted fractional orders of reactions produce a biased representation of the real process, such models are very versatile, simple, and allow application of dynamic system theory, being able to simulate cell system non-linearity (saturation, sigmoidals), multi-stability, bifurcations, oscillatory behaviour, hysteresis, responsivity to perturbations.[31,33,35] Various criteria to define the modular system functional effectiveness have been defined in terms of stability, responsiveness, selectivity, robustness, effi-

ciency,[36] while multi-objective criteria allow identification and optimization of GRN in terms of gene connectivity, stability, redundancy, robustness/low sensitivity vs. external noise, high regulatory performance (P.I.), response rate and overshoot.[1,37,38] Alternative lumped modular GRN are discriminated based on the system constraints, experimental observation, physical meaning of lumped components and reactions. However, a too advanced lumping can lead to diminish some network properties (local stability strength, efficient responsiveness, flexibility), several measures detecting the loss in system adequacy and predictability power (such as Akaike's information criterion).[2,29] As many *in vivo* cell-systems operate close to nominal steady-states (equilibrate growth with key-species at homeostasis), the power-law representation has been proved to be a valid modelling alternative, even for systems far from their nominal steady-state.

One favourite lumped representation of the non-linear enzymatic kinetics is the Michaelis-Menten hyperbolic type model. For the protein (P) synthesis, this model can represent the TF binding reaction to DNA, with explicitly including the activator (Act) / repressor (Rep) influence on the transcription/translation processes and operon activity, in one simple rate expression:[39–41]

$$\frac{dc_P}{dt} = \frac{r_{P,\max} c_S}{K + c_S} \frac{c_{Act}}{K_A + c_{Act}} \left(1 - \frac{c_{Rep}}{K_R}\right) - \beta_P c_P , \quad (2)$$

(where dilution and degradation are included in the decay constant β). One problem with this formalism is the slow response to changes in the substrate (S), Act, or inducer (In) concentrations. This is why a Hill-type model is often preferred due to the possibility to adjust the steepness of the response via increasing the Hill-coefficient n>1 (Eq. 3).[39,41] The general Hill-type expression can be used to include the quick activation of the P transcription due to an inducer *In* (c^n_{Act} being replaced by c^n_{In}), or the quick deactivation due to a repressor *Rep*:

$$\frac{dc_P}{dt} = \frac{r_{P,\max} c_S^n}{\overline{K} + c_S^n} \frac{c_{In}^{n'}}{\overline{K}_A + c_{In}^{n'}} \frac{1}{\overline{K}_R + c_{Rep}^{n''}} - \beta_P c_P , \quad (3)$$

Here n' or n'' represents the number of substrate binding sites per molecule of *In* or *Rep*.[42] The Hill-type model has been successfully applied to represent the gene expression in a wide range of GRC, highlighting some conclusions:[1,27,33,39,41,43,44] negative auto-regulation speeds the response time to perturbations; cooperativity in the repression of TF-s decrease the rise-time required by a gene product to reach half of its steady-state level; for rapidly degradable proteins (with lifetime much smaller than the cell-cycle time) a strongly negatively auto-regulated circuit produces a rise-time five-times less than those produced by a single transcription unit; explicit treatment of mRNA in gene-expression module representation (i.e. a cascade control) increases the regulatory unit effectiveness. On the other hand, the synthesis of active proteins in GRC presents a certain delay vs. transcription initiation, due to the successive steps of elongation, termination, ribosome binding, peptide elongation, protein folding, (dimeric) complexes formation, and their diffusion to the DNA-binding site. An optimal level of the repression strength can be predicted with using dynamic models, leading to rapid rise-times, but also avoiding undesirable overshoots in protein level (that increase production costs, and lead to a longer time to be removed).

Limitations of deterministic dynamic models are coming from the average behaviour description of the large populations of cells, and from the impossibility to predict fluctuations in different cells and multimodal population distributions.[21] For such reasons, stochastic GRC models have been developed by accounting for the individual molecule-to-molecule interactions; the rate equations are replaced by individual reaction probability, while the model output also becomes stochastic.[17,22–24] Hybrid models (stochastic-discrete or stochastic-continuous variables) have been reported as being feasible to represent inducible GRN.[22,24] Various random processes can also be better represented by stochastic models, such as gene mutation, interactions of species present in small amounts (as the TF-s and mRNA in the gene transcription or translation) for which spatial location is important, random faulty in GRC (e.g. faulty switches), random evolutive processes, cell signalling, etc. As an example, Salis & Kaznessis[16,17] designed a bistable switch of two gene-modules from the *lac operon* of *E. coli*, in which the transcriptional regulation is modelled by using a stochastic approach accounting for 40 elementary reactions and 27 species (reduced version) or 70 reactions and 50 species (extended version). Such a regulatory schema (Fig. 1) includes repressors in dimeric form (RR), thus increasing the efficiency of mutual-repression following the presence in excess of one of the activating inducers.

As reviewed by Heinemann & Panke,[9] the model-based design of signalling and GRC requires progresses in the following areas: I) an improved quantitative understanding of the regulatory and signalling processes and cellular organization; ii) development of effective rules and standards to characterize modules; iii) improved multiscale simulation algorithms (including modular simulation platforms). When linking regulatory modules, various system properties have to be ensured in terms

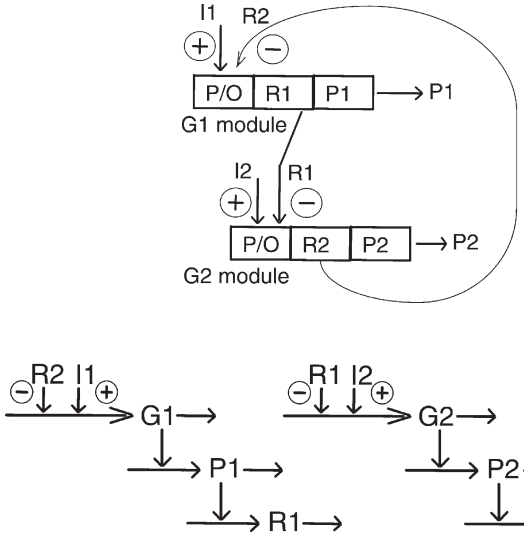
of metabolic efficiency (minimum energy and substrate consumption), individual or associative functions, hierarchic organization, system homeostasis, equilibrated cell growth, minimum intermediate levels, etc. Several linking rules have been advanced, such as:[1,45] linking reactions between modules must be set slower comparatively to the module core reactions; use cooperative – mutual catalysis; individualized functions must be allocated to each component into the cell; intermediate species levels and allosteric regulation loops must be adjusted accordingly to the GRN size; variable cell-volume and isotonic modelling environment must be considered for a more realistic representation of the secondary connectivity effects transmitted via cell-volume (cell 'ballast' and 'inertia'). When modelling large GRN, lumping rules are based on a modular representation combined with a sensitivity analysis to relate the GRN local and holistic properties to the structure, function and efficiency of individual modules to cope with perturbations.[1,46] Topological measures can help in decomposing the regulatory network due to the GRN multi-layer hierarchical structure inferring global regulators with network motifs.[47,48]

Bistable switch of two gene expression modules

Power law model with Hill-type switch

As response to an external stimulus, the expression of some genes into the cell changes, the regulator gene affecting the expression of a set of effector genes by means of a signal responsive protein, which in turn will change the production of enzymes, proteins, and other effector species.[49] In an inducible GRC, an increase in the inducer level leads to an increase in the effector gene expression, while in a repressible GRC leads to the decrease in the expression. In a bistable GRC, the two involved genes can be regulator or effectors genes, the result being a two-module system displaying one or two stable steady-states under different environmental conditions. The system switches from one steady-state to the other by increasing stimulation or inhibition or by changing the regulatory mechanism.[21] Typical examples from *Escherichia coli* cell are the Trp-repressor of the tryptophan (*trp*) operon,[49] or the lactose (*lac*) operon activation.[9,18,50]

In the schematic representation of Fig. 1, the toggle switch consists of two repressors (R1, R2) that control the expression of each other, modifying the gene G1, G2 activity in expressing the encoding proteins P1, P2. The addition of an inducer will deactivate one of the repressing proteins, leading to the increase in the corresponding gene expression,



F i g . 1 – *The bistable switch circuit of two genes designed by Salis & Kaznessis[16,17] (Up), and the lumped schema of the two modules used by the model of Wall et al.[49] (Down) (Notations: P = promoter or protein, G = gene; O = operators; I = inducer, R = repressor; ⊕ / Θ = pozitive / negative regulatory loops).*

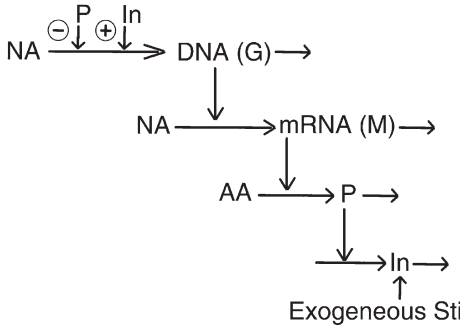
and vice-versa for the addition of the second inducer. The resulting "genetic-switch" is either "On" when [P2]>[P1], or "off" otherwise. The certainty of the switch is high when the difference in concentrations of the two co-expressed proteins P1 and P2 is high. The short transient times, high sensitivity to specific inducers, robust dynamic response, and tight control of gene expression are the common target goals when designing an optimised genetic switch circuit.[17] Following the definitions of Wall et al.,[49] various alternatives can be accounted for in such constructions: directly coupled circuits (when changes in the regulator and effector genes are coordinated), inversely coupled (when changes are opposite), or uncoupled circuits (when regulator gene expression remains constant while effector gene expression changes).

For instance, in the case of Lac-operon system of *E. coli*, the lactose (*lac*) is the inducer that disrupts the ability of the Lac-repressor (a symmetric tetramer protein) to bind and block the three operator sites of the gene expressing one of the switch proteins.[16,17] Such a Lac-system can be easily adapted for constructing genetic switches or oscillators.[17,44] Other examples of such cross-/self- repressing circuits are the directly coupled circuit of tryptophan (*trp*)-system in *E. coli,* or the histidine (*hut*)-system in *Salmonella typhimurium*, the uncoupled circuit of aspargine (*asn*)-system in *E. coli*, the purine (*pur*)- and guanine (*gua*)-operator system in *E. coli*, the oscillator (repressilator) including three gene-modules repressed by *lacI*, *tetR*, and *cI* in *E. coli*, etc.[23,36,44,49,51–53]

One of the first representation of an inducible bistable system in a bacterial cell is the Jacob-Monod trigger system.[39,50,54] The JM kinetic model accounts for only two lumped species (P1,P2), playing the role of cross-repressors and expressed proteins, while genes are not included explicitly.[54] The Hill-type rate equations describing the cross-repression of P1 and P2 can create, for Hill-coefficients higher than 1, a bistable switch where the two stable states are defined by the levels [P1-high, P2-low] and [P1-low, P2-high] respectively. Otherwise, the system has a single stable point. The competition of species (enzymes) suggests that the system can be thrown over the separatix either by adding sufficient amount of P1 that was the minimal in the initial state, or parametrically (through a genetic program) changing the system characteristics so that the initial state becomes unstable. By writing the model in the general form, $d\boldsymbol{c}/dt = \boldsymbol{g}(\boldsymbol{c},\boldsymbol{k},t)$, and by using the quasi-linearization around the steady-state, $d\boldsymbol{c}/dt \approx \boldsymbol{J}^T\boldsymbol{c}$ (where $\boldsymbol{J} = d\boldsymbol{g}/d\boldsymbol{c}$ is the Jacobian matrix, $\boldsymbol{c}$ = species concentrations; t = time), one can check: the Routh-Hurwitz conditions of local stability;[55] the necessary conditions of oscillations;[56] the Friedrichs' sufficient conditions for bifurcation $[tr(\boldsymbol{J}) = 0,\ tr(\partial\boldsymbol{J}/\partial\boldsymbol{k}) \neq 0]$ leading to bi-stability.[57] Even if intuitive, the JM model is excessively lumped, lacking on any information on the inducers' influence, on detailed activation-repressing mechanism of gene expression, on the intermediate/metabolite levels, or the separate influence of protein degradation and dilution rate. Consequently, its predictions in terms of regulatory properties are usually unrealistic.

For a better switch simulation/design, various deterministic dynamic models describing the transcription cross-regulation of the two-interconnected genes have been developed. One of the alternative is to use simple lumped power-law models of type (1), based on reaction schema including 3 – 4 lumped species per gene-module (Figs. 1 – 2).[33–36,49,58–61] As proved, the power-law representation can successfully replace the extended kinetic models including a large number of elementary steps (difficult to be observed and identified), by means of a reduced number of lumped species, lumped reactions and rate parameters. In such a manner, the large variety of nonlinear behaviours and functions can be recast in a canonical form called S-system, resulting a simpler dynamic model based on aggregated reactions forming and degrading the X_i state variables ($i = 1,\ldots n_S$ species lumped concentrations).

The versatility of power-law models allows designing coupled or uncoupled genetic circuits, such as bistable-switches. In a reduced representation, by not-explicitly including the DNA synthesis,



F i g . 2 – *Gene module type, with perfectly coupled enzyme/regulator P expression, used to represent the genetic switches by Voit[33], and Savageau[34] (In = inducer; AA = aminoacids). The enzyme (protein P) interacts with the inducer In for controlling the transcription rate by means of feedback pozitive* ⊕ *or negative* Θ *regulatory loops.*

Savageau[34] derived a gene-module dynamic model including three state-variables **X** = [M,P,In] (M = mRNA, P = enzyme/regulator protein, In = inducer, see Fig. 2). The switch effectiveness and the appropriate steady-state condition have been tuned by choosing suitable parameters values and lower/upper threshold concentrations for the inducer. The model analysis allows the full characterization of the bistable-switch, by explicitly linking the model parameters to the system properties and its effectiveness (P.I.).[1,28,36,49,60,61] Among module P.I.-s (from which some are used in the present study) are to be mentioned the followings:

– system local *stability* condition, and stability strength: I) stationary regulation, i.e. large margins of stability in the state variable space vs. stationary perturbations; ii) dynamic regulation, i.e. fast $\tau_{rec,j}$ = species j recovering time of the steady-state (QSS) with a tolerance of 5 %[36] or 1 %,[1] after an impulse-like perturbation;

– high *responsiveness* to (exo/endogeneous) signalling species of repression or de-repression, that is small rise-times (transition times τ_j) and tolerable overshoots in the level of enzymes repressing or de-repressing the gene expression;

– GRC *selectivity*, the regulator protein being sufficiently insensitive to changes in the level of effector protein [i.e. small sensitivities $S(c_{R1};c_{R2}) = \partial\ln(c_{R1})/\partial\ln(c_{R2})$, see Fig. 1], or to other species from the GRC;

– GRC *robustness*, that is small sensitivities of the system performances vs. its parameters [i.e. small sensitivities $S(c_{j,s};\boldsymbol{k}) = \partial\ln(c_{j,s})/\partial\ln(\boldsymbol{k})$, or $S(\tau_{rec,j};\boldsymbol{k}) = \partial\ln(\tau_{rec,j})/\partial\ln(\boldsymbol{k})$];

– system *regulatory efficiency* in terms of ensuring an appropriate steady-state stability in response to dynamic perturbations in internal or external species [i.e. small QSS-sensitivities $S(c_{j,s};c_{perturb.})$];

– species *connectivity* in terms of synchronized and efficient treatment of a dynamic perturbation for recovering the steady-state [i.e. small STD = standard deviation of $\tau_{rec,j}$].

The variation intervals for some model parameters can be experimentally checked, being related to the cell strain type and cycle period (see for instance the model identification for the (*trp*)-system case in *E. coli*)[36,49] This genetic-switch representation is quite flexible, the positive or negative regulation in a coupled or uncoupled circuit being adjusted by means of only one parameter. The other model parameters result from an adequate fitting, by imposing certain constraints and holistic properties to the design system (e.g. working as an effective biosensor by signalling the presence of certain type of molecules in the environment).

The repressible gene-expression module of Savageau[34] has been completed by Voit[33] (Fig. 2), with separately considering the DNA and mRNA species, and by including a Hill-type sigmoidal switch for DNA synthesis, thus allowing saturation and cooperativity in individual power-law terms:

$$\begin{aligned}
\frac{dX_1}{dt} &= \alpha_1\left(\frac{1+BX_4^n}{C+X_4^n}\right)X_3^{g_{13}} - \beta_1 X_1^{h_{11}} \\
\frac{dX_2}{dt} &= \alpha_2 X_1^{g_{21}} - \beta_2 X_2^{h_{22}} \\
\frac{dX_3}{dt} &= \alpha_3 X_2^{g_{32}} - \beta_3 X_3^{h_{33}} \\
\frac{dX_4}{dt} &= \alpha_4 X_3^{g_{43}} - \beta_4 X_4^{h_{44}}
\end{aligned} \tag{4}$$

[where: $X_1 = c_G$ = gene G concentration (DNA); $X_2 = c_M$ = concentration of mRNA; $X_3 = c_P$ = concentration of repressor & co-expressed protein; $X_4 = c_{in}$ = concentration of inducer; *n,* α*,* β*,* ***g, h*** = model parameters). The gene induction term allows tuning the baseline level of expression α_1 when c_{In} is close to zero, a maximum level $1+B$ when c_{In} is large, a switch domain determined by C, and the allosteric steepness of the response for $n>1$.

For the studied case of (*trp*)-system in *E. coli,* with a cell cycle period of $t_c = 40$ min, after tuning the model parameters Hlavacek & Savageau[36] predict the transient time of ca. t_c for a gene-module switch de-repression with coupled negative feedback control, and of ca. $2t_c$ for uncoupled switch control. By considering a similar bistable-switch example, the model (4) parameters recommended by Voit[33] are: $\alpha_1 = 4,\ B = 2,\ n = 4,\ C = X_{4,ref}^4 = 256, g_{13} = -0.5,\ \beta_1 = 0.5,\ h_{11} = 0.5,\ \alpha_2 = 1,\ g_{21} = 0.5,\ \beta_2 = 5,\ h_{22} = 0.5,\ \alpha_3 = 2,\ g_{32} = 1,\ \beta_3 = 3,\ h_{33} = 1,\ \alpha_4 = 3,\ g_{43} = 1,\ \beta_4 = 12,\ h_{44} = 0.75$ (by taking the dimensions in minutes and nM). The reported recovering times to one of the two stable QSS, $\boldsymbol{X}_S = [13.598, 7.3964, 4.9309, 1.3218, 4.0943]$ or $\boldsymbol{X}_S = [27.251, 29.705, 19.804, 8.4381, 11.6154]$ from the unstable QSS, is of ca. 200-900 min, that is much larger than t_c. As observed by Elowitz & Leibler[44] in the oscillatory GRC, the transition time in the inductive/co-repressed GRC systems can be much larger than the cell-division cycle, the state of the oscillator or transient switch being transmitted from generation to generation. As another observation, the stationary dilution rate $D_S = \ln(2)/t_c$ (due to the cell-volume growth) in the models of Savageau[34] and Voit[33] is usually lumped together with the species degradation rate, leading to non-uniform parameters $\beta_1 - \beta_4$.

The power-law models are computationally tractable, being very convenient to analyse the GRC characteristics, and to analytically derive conditions of stability and performances. However, they also present some weaknesses. By using lumped model terms (which shorten the list of species and reactions), a loss of information on elementary steps and on some intermediates will lead to a simulated GRC-system of lower flexibility/adaptability and prediction capabilities, and to a bias in GRC behaviour vs. perturbations. The lack of intermediate species of adjustable levels, which can better optimise the system, will lead to a difficult identification step of model parameters (sometimes lacking of physical meaning), due to the requirement of matching the mentioned system local and holistic properties (stability, multiplicity, sensitivity, effectiveness)[2] with a fewer number of variables. For instance, by imposing a stable QSS of $\boldsymbol{X}_S = [4, 4, 4, 40]$ to the model (4), identification of rate constants lead to the same values as identified by Voit[33], excepting for $\beta_1 = 2,\ \beta_2 = 1,\ \beta_3 = 2,\ \beta_4 = 0.75446$. However, this stationary point presents a very small region of stability lacking of the switch properties. As a conclusion, identification of parameters needs to account for a larger number of constraints, and requires more effective numerical search algorithms, which can become problematic for solving large GRC-systems.[33]

Moreover, by ignoring the explicit cell content evolution and its influence on the GRC, over-estimated conclusions might be derived in terms of design GRC-switch performances. As proved by Morgan et al.[36] and Maria,[1,2] in a whole-cell representation the cell large ballast tends to stabilize the system and to smooth the perturbations, thus decreasing the GRC responsiveness comparatively to an isolate system behaviour. By considering gene-expression modules together with the cell growth and its content replication, and by explicitly including

lumped interactions with the genome and proteome, other important effects can be modelled, such as the cell content inertial effect in treating perturbations, the effect of the indirect or secondary perturbations transmitted via cell-volume variation under isotonic osmolarity conditions.[1]

Modelling gene-expression switch modules under variable volume framework

Variable cell-volume dynamic models

In order to simulate the gene expression regulation in connection with the whole-cell evolution, a continuous variable differential model has been adopted, accounting for the variable cell-volume and osmotic pressure but neglecting the inner-cell gradients.[1,2,26,62] The cell-model of Maria[1] is based on the hypothesis that, the quasi-constant osmotic pressure in the cell will connect the volume variation to the molecular species dynamics by means of the Pfeffer's law in diluted isotonic solutions:[63]

$$\frac{dc_j}{dt} = \frac{1}{V}\frac{dn_j}{dt} - Dc_j = g_j(\boldsymbol{c},\boldsymbol{k}):$$

(at steady-state: $g_j(\boldsymbol{c}_S,\boldsymbol{k}) = 0$);

$$\pi V(t) = RT\sum_{j=1}^{n_S} n_j(t), \text{ (Pfeffer's law)} \Rightarrow$$
$$D = \frac{1}{V}\frac{dV}{dt} = \left(\frac{RT}{\pi}\right)\sum_{j=1}^{n_S}\left(\frac{1}{V}\frac{dn_j}{dt}\right), \quad (5)$$

(where: V = cell volume; n_j = species j number of moles; D = cell-content dilution rate, i.e. the cell-volume logarithmic growing rate; π = osmotic pressure; T = temperature; R = universal gas constant; n_S = number of species inside the cell). Such an assumption, made for the growing phase of the cell (ca. 80 % of the cell-cycle), implies that perturbation in one species will influence the volume growth, which in turn will perturb the other cell component concentrations (the so-called 'secondary' perturbation).[1] So, apart from their direct and indirect connectivity, the species are also inter-connected via the common cell-volume to which all species contribute. This is an important feature of the variable-volume / whole-cell model vs. the classical approach.

The model hypotheses and the basic equations are presented in Table 1. The main assumptions

Table 1 – *Variable cell-volume dynamic modelling framework, and its basic hypotheses*[1]

Mass balance and state equations	Remarks
$\frac{dc_j}{dt} = \frac{1}{V}\frac{dn_j}{dt} - Dc_j = g_j(\boldsymbol{c},\boldsymbol{k});$ $\frac{1}{V}\frac{dn_j}{dt} = r_j(\boldsymbol{c},\boldsymbol{k}); \quad j = 1,\ldots,n_S;$	continuous variable dynamic model representing the cell growing phase (ca. 80% of the cell cycle)
$V(t) = \frac{RT}{\pi}\sum_{j=1}^{n_S} n_j(t);$	Pfeffer's law in diluted solutions[63]
$D = \frac{1}{V}\frac{dV}{dt} = \left(\frac{RT}{\pi}\right)\sum_{j}^{n_S}\left(\frac{1}{V}\frac{dn_j}{dt}\right)$	D = cell content dilution rate = cell volume logarithmic growing rate
$\frac{RT}{\pi} = \frac{V}{\sum_{j=1}^{n_S} n_j} = \frac{1}{\sum_{j=1}^{n_S} c_j} = \frac{1}{\sum_{j=1}^{n_S} c_{j0}} = \text{constant}.$	constant osmotic pressure constraint
$\left(\sum_{j}^{all} c_j\right)_{cyt} = \left(\sum_{j}^{all} c_j\right)_{env}$	isotonic osmolarity constraint

Hypotheses:

– negligible inner-cell gradients;
– open cell system of uniform content;
– semi-permeable membrane, of negligible volume and resistance to nutrient diffusion, following the cell growing dynamics;
– constant osmotic pressure, ensuring the membrane integrity ($\pi_{cyt} = \pi_{env}$ = constant);
– nutrient and overall environment concentration remain unchanged over a cell cycle;
– logarithmic growing rate of average $D_s = \ln(2)/t_c$; volume growth of $V = V_0 e^{D_s t}$; t_c = cell cycle time;
– homeostatic stationary growth of $(dc_j/dt)_s = g_j(\boldsymbol{c}_s,\boldsymbol{k}) = 0$;
– perturbations in cell volume are induced by variations in species copynumbers under the isotonic osmolarity constraint: $V_{perturb}/V = (\Sigma n_j)_{perturb}/(\Sigma n_j)$.

concern the space division, osmotic pressure invariance, and the state-law linking volume, pressure, temperature, and cell content. In such a manner, the dynamics of molecular biological processes occurring in the growing phase and dividing phase is directly linked with the cell-volume evolution, the rates of individual reactions within cell being constrained by the periodicity of the cell-cycle and by the requirement that molar amounts of all components and the volume must double in exactly one cell-cycle. As an important observation, the whole-cell modelling under variable volume framework requires that each and every process be included at some level of detail (i.e. individual or lumped species) in order to fulfil the constraint (5).

This modelling approach is not only more complex vs. the classical constant-volume framework, but introduction of the Pfeffer's law restriction (5) leads to a better simulation of the species interconnectivity via direct, indirect, and secondary perturbations transmitted via the common volume. By explicitly accounting for the main cell characteristics and linking the other species/lumps evolution, these models leads to usually different results vs. the classical approach, i.e. significantly different P.I.-s for the individual genetic regulatory modules. For instance, by keeping the same module pathway schema, the recovering rates for minor species (present in small amounts) after a dynamic perturbation can be several times slower in a whole-cell/variable-volume formulation, while the sensitivities to stationary perturbations can be for some species of orders of magnitude smaller than those predicted by the constant-volume formulation (see the results of Maria[1,2] compared to the analysis of Yang et al.[27] and Maria[28]).

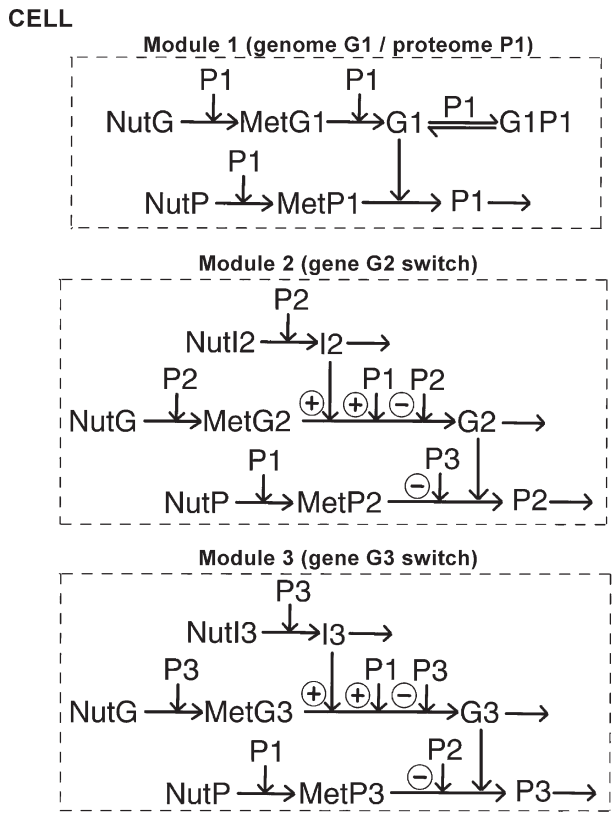
The allosteric control of enzyme activity has been accounted for, by explicitly including fast reversible 'buffering' reactions, standardized in the form of the so-called regulatory units $L_i(R_i)n_i$. One unit i is formed by the component L_i [e.g. enzymes or even genes (G), proteins (P), mRNA (M), etc.] at which regulatory element acts, and n_i = 0,1,2,… number of 'effector' species Ri (e.g. P, PP, PPPP, etc.) binding the 'catalyst' L_i. For instance, a *G(P)4* unit includes four successive binding steps of G with the protein P (of type G + P ⇔ GP), all intermediate species GP, GPP, GPPP, GPPPP being inactive catalytically, while the mass conservation law is all time fulfilled, i.e. $\sum_{i=0}^{4}[G(P_i)]=$ constant. Such a representation accounts for the protein concentration diminishment due to the cell-growth dilution effect, but could also include the protein degradation by proteolysis.

Using this framework, cell regulatory modules can be modelled based on the generated identifiable function and regulatory properties. For instance, sets of semi-autonomous modules can be used to build-up regulatory chains of protein synthesis sustaining the cell homeostasis during the equilibrated growth.[1] The model parameters, i.e. the rate constants and the unobservable stationary concentrations of some intermediates, have been identified based on the QSS-condition of stability (accounting for homeostatic quasi-constancy of concentrations for some key-species during the cell growth), and on extremization of an optimality criterion of the whole network, such as:[1,27,28,37,38,45] i) maximum recovering rates after a dynamic perturbation; ii) minimum levels for intermediate concentrations; iii) maximum system flexibility vs. external changes (large margins of stability due to multiple levels of intermediates and homeostasis); iv) smallest QSS-sensitivity vs. stationary perturbations, and robustness vs. noise; v) maximum regulatory loop sensitivity vs. perturbations; vi) imposed properties of the internal oscillators; vii) minimum number of inter-connected genes (sparsely connection gene criterion), etc. An important number of constraints must also be added: physical meaning of parameters (positive and bounded rate constants); fulfilment of cell structural, functional and temporal internal hierarchy; genetic network redundancy; isotonic osmolarity.

By relating the module structure to the local and global regulatory network properties, several conclusions can be derived in terms of regulatory effectiveness for various effector types and number, regulatory units properties, module effectiveness, linking rules for weakly-interconnected modules, variable volume and pressure influence on regulatory properties, etc.[1,2,26] Based on these conclusions, some rules can be proposed for building-up modular simulation platforms for GRC with desired identifiable functions and properties.[1,29]

Hill-type induction and cross-repressive switch (model 3M-A)

To exemplify in a simple way how can be modelled the properties of a GRC under variable cell-volume and isotonic osmolarity, one places a bistable switch of two gene-expression regulatory modules in a *E. coli* cell (K-12 strain, Ecocyc[3]), of initial volume $V_{cyt,o} = 1.66 \times 10^{-15}$ L and dilution rate of $D_s = \ln(2)/100$ min^{-1} ($t_c = 100$ min). The cell lumped representation, denoted as model 3M-A, includes only three modules (Fig. 3), i.e. for synthesis of the lumped pair G1/P1, and for the



F i g . 3 – *Model 3M-A: placing two gene switch modules (for G2 and G3 expression) in a variable volume cell mimicking the E. coli growth. The G1 module expression mimics the whole cell content replication. P1 lump plays the role of a permease for NutG, NutP import, of a metabolase for MetG1, MetP1, MetP2, MetP3 synthesis, and of a polymerase for G1, G2, G3 production. P2 plays the role of a permease/metabolase for exogeneous NutI2 (stimulus) import and I2 (inducer) production, and of a self- or cross-repressor for G2 and G3 synthesis respectively. P3 plays the role of a permease/metabolase for exogeneous NutI3 (stimulus) import and I3 (inducer) production, and of a self- or cross-repressor for G3 and G2 synthesis respectively. Inducers I2, I3 activate the G2, G3 synthesis. (Notations: G1 = lumped genome; P1 = lumped proteome; MetG1, MetP1 = lumped metabolome; NutP, NutG = lumped external nutrients; P2, P3 = two individual proteins; G2, G3 = two individual genes; MetG2, MetG3, MetP2, MetP3 = individual metabolites; I2, I3 = inducers; NutI2, NutI3 = external stimuli; ⊕ / Θ = positive / negative regulatory loops).*

synthesis of individual pairs G2/P2 and G3/P3. The designed genetic switch, even kept as generic, is able to describe the evolution of the main involved species/lumps, being similar to several existing or designed switches from *E. coli,* such as the *lacI* and *λ cI* inter-connected genes encoding the transcriptional regulatory proteins LacR and *λ* CI cross-repressing the corresponding gene expression promoters.[25]

The *Module 1*, denoted as G1(P1)1 following the Yang et al.[27] standardization, mimics the genome and proteome replication and, due to their large concentrations, the cell content 'balast' and 'inertial' effect when coping with perturbations. To not complicate the model, the lumped proteome P1 synthesis is controlled by means of the lumped genome G1 activity using the simplest regulatory schema with one rapid buffering reaction G1 + P1 ⇔ G1P1, close to its equilibrium, with a dissociation constants much larger than those of the core synthesis (e.g. $k_{diss} \approx 10^7 D_s$ in the present example). The [P1] and [G1] concentrations at steady-state, displayed in Table 2, have been computed by accounting for the born-cell volume [see Table 2 footnotes (c-d)] and by summing over all genes and proteins, protein fragments and complexes. At QSS, equal concentrations of catalytically active/inactive forms $[G1]_s = [G1P1]_s$ ensures maximum regulation sensitivity vs. perturbations.[1] The nominal environmental conditions consider the lumped nutrient concentrations of $[NutG]_s = 3\times10^7$ nM, $[NutP]_s = 3\times10^8$ nM, and the traces of signalling species at insignificant levels of $[NutI2]_s = [NutI3]_s \approx 10^{-4}$ nM. Concentrations of inner metabolites MetGj, MetPj (j = 1–3) are set to high values, i.e. $\sum_j^{all} c_{MetPj,cyt} = 3\times10^8\, nM$, while $\sum_j^{all} c_{MetGj,cyt}$ results from fulfilling the inner/outer equal osmotic pressure condition $\pi_{cyt} = \pi_{env}$ that is $\sum_j^{all} c_{j,cyt} = \sum_j^{all} c_{j,env}$. Such high levels of P1, G1, MetGj, MetPj will be less influenced by small size perturbations due to the cell 'core' and 'inertial' effects.

The bistable switch is formed by two connected gene-expression modules G2/P2 and G3/P3, denoted as *Module 2* and *Module 3* in Fig. 3. The induction/repression schema is of type $[P1(P2)_{-0.5}(I2)_{Hill};G2(P3)_{Hill}] + [P1(P3)_{-0.5}(I3)_{Hill};G3(P2)_{Hill}]$, being modelled using the Voit[33] kinetic expressions (4). As most of the switches are designed to respond to nonendogeneous inducers, only external stimuli have been considered. In such a grouped representation, P1 lump plays the role of a permease for NutG, NutP import, of a metabolase for MetG1, MetP1, MetP2, MetP3 synthesis, and of a polymerase for G1, G2, G3 production. P2 & P3 proteins play the role of permeases and metabolases for exogeneous NutI2 & NutI3 (stimuli) import and for I2 & I3 (inducers) production. P2 & P3 are also cross- or self- repressors for the genes G2 and G3 synthesis respectively, activated by the inducers I2 & I3.

The model includes only pseudo-elementary reactions, except those for the switch genes G2 & G3 and proteins P2 & P3 production, for which the apparent Hill-type kinetics of Voit[33] has been adopted according to the model (4):

T a b l e 2 – *Cell lumped content under stationary growth, and the species sensitivities vs. the quasi-stationary environmental conditions for the considered E. coli cell. Predicted quasi-steady-states (QSS) by various cell models under perturbed environmental conditions (NutI2, NutI3 denote the external stimuli which activate the bistable G2/G3 gene expression switch).*

Species/ Model	$c_{j,s}$ (nM) (nominal steady-state)	$\left[\frac{\partial \ln(c_j)}{\partial \ln(c_{NutI2})}\right]_s \times 10^{13}$			$c_{j,s}$ (nM) (NutI2 ext. stimulus present; G3-repressed)			$c_{j,s}$ (nM) (NutI3 ext. stimulus present; G2-repressed)		
	Model 3M-C	*3M-A*	*3M-B*	*3M-C*	*Model 3M-A*	*Model 3M-B*	*Model 3M-C*	*Model 3M-A*	*Model 3M-B*	*Model 3M-C*
Environment										
NutG (f)	3×10^{7}	–	–	–	3×10^{7}	3×10^{7}	3×10^{7}	3×10^{7}	3×10^{7}	3×10^{7}
NutP (g)	3×10^{8}	–	–	–	3×10^{8}	3×10^{8}	3×10^{8}	3×10^{8}	3×10^{8}	3×10^{8}
NutI2	10^{-4}	–	–	–	10	10	10	10^{-4}	10^{-4}	10^{-4}
NutI3	10^{-4}	–	–	–	10^{-4}	10^{-4}	10^{-4}	10	10	10
Inner cell										
MetG1 (a)	1.9975×10^{7}	–3.02	–3.02	–2.67	1.9974×10^{7}	1.9975×10^{7}	1.9974×10^{7}	1.9974×10^{7}	1.9975×10^{7}	1.9974×10^{7}
MetG2	10^{4}	–2.95	+0.81	–3.24	4.0340×10^{4}	3.0347×10^{4}	2.9325×10^{4}	7.544×10^{3}	1.710×10^{3}	1.7075×10^{4}
MetG3	10^{4}	–2.96	+0.75	–3.24	7.544×10^{3}	1.710×10^{3}	1.7075×10^{4}	4.0340×10^{4}	3.0347×10^{4}	2.9325×10^{4}
MetP1 (b)	2.98×10^{8}	–2.97	–2.97	–2.63	2.9798×10^{8}	2.9799×10^{8}	2.9798×10^{8}	2.9798×10^{8}	2.9799×10^{8}	2.9798×10^{8}
MetP2	10^{6}	–3.03	–3.03	–2.68	9.9991×10^{5}	9.9993×10^{5}	9.9991×10^{5}	9.9992×10^{5}	9.9997×10^{5}	9.9992×10^{5}
MetP3	10^{6}	–3.03	–3.03	–2.68	9.9992×10^{5}	9.9997×10^{5}	9.9992×10^{5}	9.9991×10^{5}	9.9993×10^{5}	9.9991×10^{5}
P1 (c)	10^{7}	–4.82	–4.82	–4.07	9.9987×10^{6}	9.9994×10^{6}	9.9987×10^{6}	9.9987×10^{6}	9.9994×10^{6}	9.9987×10^{6}
P2	1	–4.75	–0.98	–4.71	4.03	3.03	2.93	0.75	0.17	1.70
P3	1	–4.75	–1.03	–4.71	0.75	0.17	1.70	4.03	3.03	2.93
G1 (d)	4500/2	–3.64	–3.64	–3.01	2249	2250	2249	2249	2250	2249
G2 (d)	1/3	–3.60	–0.69	–1.89	4.01	3.38	0.88	0.86	0.04	0.33
G3 (d)	1/3	–3.61	–0.77	–1.89	0.86	0.04	0.33	4.01	3.38	0.88
G1P1	4500/2	–8.47	–8.47	–7.08	2249	2250	2249	2249	2250	2249
G2P2P2	1/3	–	–	–4.24	–	–	1.89	–	–	0.24
G2P3P3	1/3	–	–2.75	–4.24	–	9.89×10^{-2}	0.64	–	3.72×10^{-1}	0.72
G3P3P3	1/3	–	–	–4.24	–	–	0.24	–	–	1.89
G3P2P2	1/3	–	–2.73	–4.24	–	3.72×10^{-1}	0.72	–	9.89×10^{-2}	0.64
P2P2 (e)	2	–	–1.95	–2.35	–	18.41	4.30	–	5.84×10^{-2}	1.45
P3P3 (e)	2	–	–2.07	–2.35	–	5.84×10^{-2}	1.45	–	18.41	4.30
I2	10^{-4}	1.0×10^{13}	1.0×10^{13}	8.84×10^{12}	40.34	30.34	29.32	7.5×10^{-5}	1.7×10^{-5}	1.7×10^{-4}
I3	10^{-4}	–2.96	+0.75	–3.24	7.5×10^{-5}	1.7×10^{-5}	1.7×10^{-4}	40.34	30.34	29.32
$\lambda(J)$ (stability)	All negatives (stable QSS)	–	–	–	All negatives (stable QSS)	All negatives (stable QSS)	All negatives (stable QSS)	All negatives (stable QSS)	All negatives (stable QSS)	All negatives (stable QSS)

(a) calculated from $\sum_j^{all} c_{j,cyt} = \sum_j^{all} c_{j,env}$, i.e.

$$\sum_j^{all} c_{MetGj,cyt} = \sum_j^{all} c_{j,env} - \sum_j^{all} c_{MetPj,cyt} - \sum_j^{all} c_{Gj,cyt} - \sum_j^{all} c_{Pj,cyt};$$

(b) $\sum_j^{all} c_{MetPj,cyt} = 3\times10^{8}\, nM$;

(c) $\sum_j^{all} c_{Pj,cyt} = 10^{7}\, nM$, evaluated with the formula[1]:

$c_j = \frac{no.\ of\ copies\ j}{(6.022\times10^{23})\times V_{cyt,o}}$; *E.coli* cell was considered as including ca. 1000 ribosomal proteins of 1000-10000 copies, ca. 3500 non-ribosomal proteins of avg. 100 copies, and ca. 4500 polypeptides of avg. 100 copies (K-12 strain EcoCyc[3,64]);

(d) the *E. coli* genome includes ca. 4500 genes (of one copy; K-12 strain, EcoCyc[3]; to allow a maximum regulatory effectiveness of the buffering reactions adjusting the gene (G) catalytic activity, concentrations of active and inactive G-forms have been taken equal at QSS,[1,28,45] i.e. $[G1]_s=[G1P1]_s=4500/2$; $[G2]_s=[G2P2P2]_s=[G2P3P3]_s=1/3$; $[G3]_s=[G3P3P3]_s=[G3P2P2]_s=1/3$;

(e) $[P2P2]_s > 1nM$, $[P3P3]_s > 1nM$ are optimised in order to confer some cell characteristics: QSS-stability and module high regulatory efficiency[1];

(f) Maria[1]

(g) Morgan et al.[26]

Notations:

$\lambda(J)$= eigenvalues of the dynamic model Jacobian; 'o'= initial (born cell); 's'= stationary; $V_{cyt,o} = 1.66 \times 10^{-15}$ L = initial cell (cytoplasma) volume; D_s = ln(2)/100 min^{-1} = cell-volume stationary logarithmic growing rate; *Model 3M-A* accounts for three gene lumped modules with the Hill-type induction rate and implicit Hill-type repressor of gene expression; *Model 3M-B* is the model 3M-A with explicitly including rapid buffering reactions for cross-repression by means of species G2P3P3 and G3P2P2; *Model 3M-C* is the model 3M-A with explicitly including rapid buffering reactions for self- and cross-repression by means of species (G2P2P2, G2P3P3) and respectively (G3P3P3, G3P2P2).

$$\frac{dc_{G2}}{dt} = k_{G2}c_{MetG2}c_{P1}\left(\frac{1+Bc_{I2}^4}{C_1+c_{I2}^4}\right)c_{P2}^{-0.5} - Dc_{G2};$$

$$\frac{dc_{P2}}{dt} = k_{P2}c_{MetP2}c_{G2}\left(\frac{1}{C_2+c_{P3}^2}\right) - Dc_{P2};$$

$$\frac{dc_{G3}}{dt} = k_{G3}c_{MetG3}c_{P1}\left(\frac{1+Bc_{I3}^4}{C_3+c_{I3}^4}\right)c_{P3}^{-0.5} - Dc_{G3}; \quad (6)$$

$$\frac{dc_{P3}}{dt} = k_{P3}c_{MetP3}c_{G3}\left(\frac{1}{C_4+c_{P2}^2}\right) - Dc_{P3}.$$

The dilution rate is uniform for all species in the model, while the degradative steps have been neglected. The gene G2/G3 expression activation accounts for four molecules (n = 4) of inducer, allosterically binding to the promoter site, while a slow self-repression with the product is considered of a –0.5 apparent reaction order. The cross-repression of protein P2/P3 synthesis, of Hill type, accounts for only dimeric repressors (n = 2) allosterically binding to the catalytic gene, even if a higher control (with tetramers) have been reported.[16,17]

The derived 3M-A kinetic model accounts for 19 individual and lumped species and includes 21 rate constants identified from solving the stationary model equations (5,6) in the form $\boldsymbol{g}(\boldsymbol{c},\boldsymbol{k},t)_s = 0$ with substituted observed nominal concentrations of Table 2 (2nd column). Similarly to Voit[33], Hill-constants are set to $B = 2$, $C_I = c_{I2,ref}^4 = 1^4$ (nM^4), $C_3 = c_{I3,ref}^4 = 1^4$ (nM^4), $C_2 = c_{P3,ref}^2 = 10^2$ (nM^2), $C_4 = c_{P2,ref}^2 = 10^2$ (nM^2).

By simulating the cell model at stationary conditions it is possible to predict the species QSS-concentrations under lack of stimuli ($[\text{NutI2}]_s$ = $[\text{NutI3}]_s \approx 10^{-4}$ nM), but also when one of the external stimuli reach a certain concentration, e.g. $[\text{NutI2}]_s$ = 10 nM, or $[\text{NutI3}]_s$ = 10 nM (see the results of Table 2). In all the cases, the obtained QSS is stable, all system Jacobian eigenvalues $\lambda(\boldsymbol{J})$ being real and negative. As an observation, the use of G1(P1)1 regulation schema for the lumped genome/proteome instead of a unregulated G1(P1)0 schema is conferring a better stability strength to the QSS, that is a $\text{Max}|\lambda(\boldsymbol{J})|$ value of 2×10^5, comparatively to 1.4×10^{-2}. In order to analyse the model 3M-A properties, several performance indices (P.I.) described in Section 3.1 have been evaluated, as followings.

After application of an external stimulus (a stationary perturbation in $[\text{NutI2}]_s$ or $[\text{NutI3}]_s$), one gene is over-expressed while the other is repressed, and the species concentrations reach a new QSS after a certain transient time τ_j (with a 1 % tolerance), with a rate depending on the bistable switch schema effectiveness. The average transient time $AVG(\tau_j)$ can be used to approximate the overall *responsiveness* of the modular GRC, while the standard deviation $STD(\tau_j)$ can be used as a global measure of the species *connectivity* (synchronisation) during the transition [see τ_j, $AVG(\tau_j)$, $STD(\tau_j)$ in Fig. 4]. As expected, the obtained transition times are longer for species present in small amounts (e.g. P2, P3, G2, G3,…) and negligible for major species belonging to the cell ballast (P1, G1, MetP1-MetP3, MetG1). The average $AVG(\tau_j)$ = 2191 min [and $STD(\tau_j)$ = 2074 min] is much larger than those of ca. 900 min estimated by Voit[33] in analysing the individual switch module taken out of *E. coli* cell. As a conclusion, by explicitly accounting for the variable-volume and whole cell content in a real cell, the lumped power-law model predictions appear to be of lower quality, requiring completion with supplementary regulatory details of the mechanism.

The species QSS-level sensitivity vs. stationary perturbations from the environment $S(c_j;c_{Nut_i}) = (\partial c_j/\partial c_{Nut_i})_s$ can be calculated from solving the differentiated set of stationary mass balance equations (Table 1, first line, $\boldsymbol{g}(\boldsymbol{c},\boldsymbol{k},t)_s = 0$), that is:

$$\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{c}}\right]_s\left[\frac{\partial \boldsymbol{c}}{\partial \boldsymbol{c}_{Nut}}\right]_s + \left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{c}_{Nut}}\right]_s = 0. \quad (7)$$

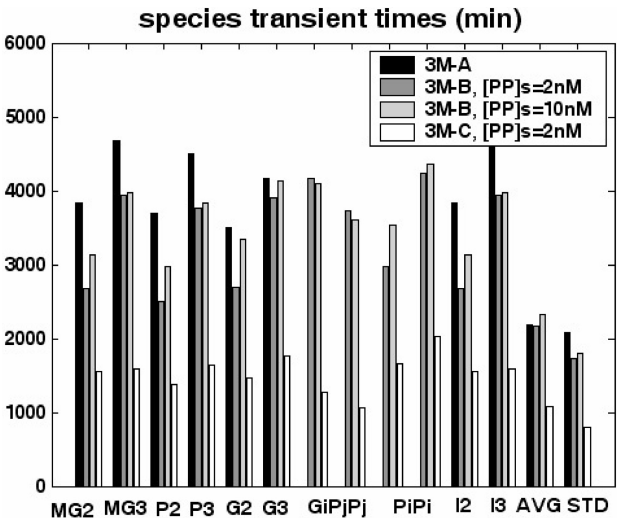


Fig. 4 – *Comparative species transient times τ_j, predicted by various gene-switch models (3M-A, 3M-B, 3M-C), from the cell nominal steady-state (column 2, Table 2; $[\text{NutI2}]_s$ = $[\text{NutI3}]_s = 10^{-4}$ nM) to another steady-state, under the influence of the stationary external stimulus of $[\text{NutI2}]_s$ = 10 nM (columns 6-8, Table 2). AVG= average of the transient times; STD= standard deviation of the transient times. Notation GiPjPj denotes the species G2P3P3 and G3P2P2. Notation PiPi denotes the species P2P2 and P3P3. The transient times of species not-plotted are negligible (less than 1 min, that is for MetG1=MG1, MetP1, MetP2, MetP3, P1, G1, G1P1).*

(where $\boldsymbol{J} = d\boldsymbol{g}/d\boldsymbol{c}$ is the system Jacobian matrix). By referring these sensitivities to a stationary state, the relative sensitivities are obtained in the form $\widetilde{S}(c_j; c_{Nut_i}) = (c_{Nut_i,s}/c_{j,s})(\partial c_j/\partial c_{Nut_i})_s = (\partial \ln(c_j)/\partial \ln(c_{Nut_i}))_s$. The results, presented in Table 2 (3rd column), reveal comparable relative sensitivities of species vs. NutI2 (or NutI3) changes, excepting high values for I2 and I3 species, directly affected by the presence of external stimuli. It is also to observe that protein P1-P3 levels are more sensitive to signalling NutI2/NutI3 changes than the genes G1-G3 levels.

To test the protein synthesis *regulation efficiency*, one applies a dynamic, impulse-like, instantaneous perturbation in $[P2]_s$ (i.e. ±10 % $c_{P2,s}$ in Fig. 5), and one determines the species recovering times $\tau_{rec,j}$ to their QSS-levels (with a 1 % tolerance), the average $AVG(\tau_{rec,j})$ and the standard deviation $STD(\tau_{rec,j})$. Small dynamic perturbations are frequent during the cell growth due to inherent interactions among species participating to complex (parallel) metabolic processes. The results, presented in Fig. 5, indicate QSS-recovering times much shorter than the transient times (from one QSS to another QSS), but still longer than the cell-division cycle of t_c = 100 min [i.e. $AVG(\tau_{rec,j})$ = 216 min, $STD(\tau_{rec,j})$ = 363 min]. Such a result indicates the model 3M-A as predicting a satisfactory responsive bistable GRC-switch, but of a low regulatory efficiency of protein synthesis vs. inner dynamic perturbations. In fact, the power-law model weakness consists in the too advanced lumping of terms (species and elementary reactions), lacking of intermediates of quickly adjustable levels in buffering reactions, thus leading to an alteration in GRC-system properties (such as cell flexibility and QSS-stability characteristics).

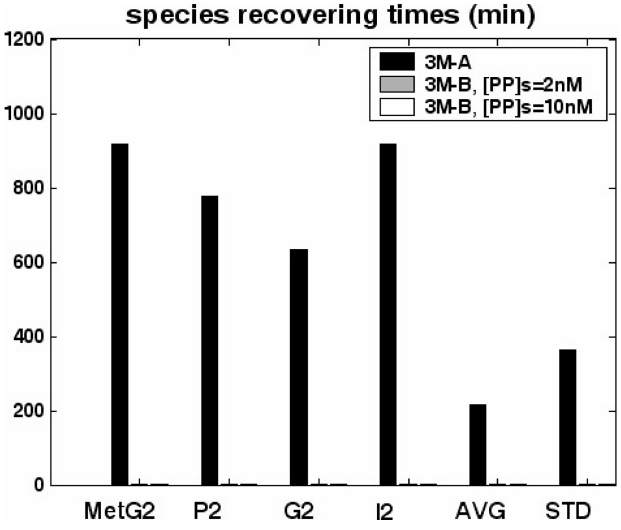


Fig. 5 – *Comparative species recovering times to QSS ($\tau_{rec,j}$, with 1% tolerance), predicted by various gene-switch models, after a ±10 % $c_{P2,s}$ impulse perturbation, in the presence of the stationary external stimulus $[NutI2]_s$ = 10 nM ($[NutI3]_s = 10^{-4}$ nM). The predicted $\tau_{rec,j}$ by the model 3M-B (with $[P2P2]_s = [P3P3]_s$ = 2nM or 10nM), and by the model 3M-C are smaller than 1 min. AVG = average of the recovering times; STD = standard deviation of the recovering times. The recovering times of species not-plotted are negligible (less than 1 min, that is for MetG1, MetG3, MetP1-MetP3, P1, P3, G1, G3, G1P1, GiPjPj, GiPiPi, PiPi, I3).*

Mixed Hill-type switch with elementary buffering reactions (models 3M-B and 3M-C)

In order to improve the model 3M-A by better accounting for the adjustable genetic switch properties (regulatory efficiency, flexibility) the previous model 3M-A has been completed with some elementary reactions detailing the cross-repression of proteins P2 and P3 synthesis. Thus, the Hill-type kinetics will be replaced by rapid buffering reactions adjusting the G2 and G3 catalytic activity by reversible binding to the dimmers of PP-type, in a Gi(PjPj)1 regulatory unit[1,27] (see also Fig. 6):

(models 3M-B and 3M-C)

$$\text{G2(P3P3)1: } G2 + P3P3 \underset{k_{diss}}{\overset{k_{bind}}{\rightleftharpoons}} G2P3P3$$

$$\text{G3(P2P2)1: } G3 + P2P2 \underset{k_{diss}}{\overset{k_{bind}}{\rightleftharpoons}} G3P2P2$$

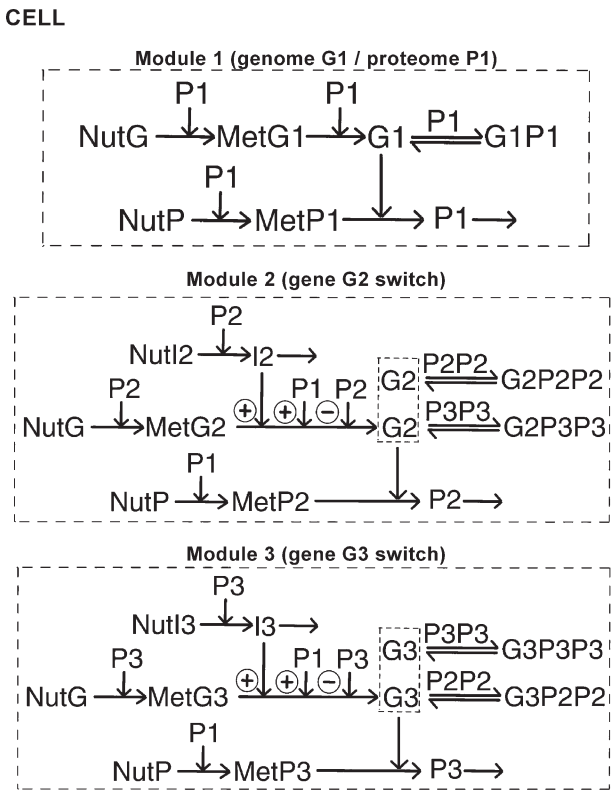
$$K_{GiPjPj} = \frac{k_{bind}}{k_{diss}} = \frac{c_{GiPjPj,s}\left(1+\dfrac{D_s}{k_{diss}}\right)}{c_{Gi,s}c_{PjPj,s}};\quad k_{diss} >> D_s. \quad (8)$$

Such a gene activity control by means of dimer effectors are proved to be highly efficient in dynamic regulation, especially when an allosteric cross-inhibition with PP is used.[1] As exemplified by Salis & Kaznessis[16] for a bistable-switch, "the ability for the repressor to bind to two operators at once allows the repressor to form a stable complex, increasing the transcriptional control." Moreover, the optimised intermediates [P2P2]s and [P3P3]s levels can confer a certain regulator flexibility, increasing the QSS-stability strength and recovering rates.

To extend the comparison terms, one completes the model 3M-B with considering a better regulatory loop for the self-repressing control of the protein synthesis, by means of a highly effective regulatory unit, that is Gi(PjPj)1 and Gi(PiPi)1 (see also Fig. 6):

(model 3M-C only)

$$\begin{aligned}&\text{G2(P2P2)1: } G2 + P2P2 \underset{k_{diss}}{\overset{k_{bind}}{\rightleftharpoons}} G2P2P2\\ &\text{G3(P3P3)1: } G3 + P3P3 \underset{k_{diss}}{\overset{k_{bind}}{\rightleftharpoons}} G3P3P3.\end{aligned} \quad (9)$$



F i g . 6 – *Models 3M-B and 3M-C: the lumped gene expression system placed in a variable volume cell mimicking the E. coli growth. The individual G2 and G3 gene expression is regulated by buffering reactions of type G2(P3P3)$_n$(P2P2)$_m$ and G3(P2P2)$_n$(P3P3)$_m$ (m,n = 0,1,2,…; see also Notations of Fig. 3). For Model 3M-B, n=1,m=0 (cross repression), and for Model 3M-C, n=1,m=1 (mixed cross- and auto-repression).*

As remarked by Salis & Kaznessis,[16] such a cross-repression doubled by a self-repression might have a positive effect on the system responsiveness by decreasing the concentration of both repressor and co-expressed proteins at steady-state (lumped together in P2 for G2-module, or in P3 for G3-module).

The kinetic model 3M-B accounts for 23 individual and lumped species and includes 27 rate constants, while the model 3M-C for 25 individual and lumped species and includes 31 rate constants. The kinetic constants are identified by solving the stationary model equations (5,8,9) with the substituted observed nominal concentrations of Table 2 (2nd column). The PiPi intermediate stationary-concentrations have been varied to study their effect on the module P.I.-s. The Hill-type kinetic term for the gene synthesis activation remains unchanged as for the model 3M-A [first term of dc_{Gi}/dt in (6)], with the same rate constants: $B = 2$, $C_1 = c_{12,ref}^4 = 1^4$ (nM^4), $C_3 = c_{13,ref}^4 = 1^4$ (nM^4). The P2 and P3 synthesis regulation uses the buffer reactions (8-9) in negative feedback loops, with $k_{diss} \approx 10^7 D_s$, while $[P2P2]_s$ and $[P3P3]_s$ are adjusted in order to confer stability strength to the nominal steady-state and a high system responsiveness and connectivity when stationary environmental stimuli are applied.

By simulating the cell dynamics when one of the external stimuli reach a certain concentration, e.g. $[NutI2]_s$ = 10 nM, or $[NutI3]_s$ = 10 nM, it is possible to predict the species new QSS-concentrations (see QSS in Table 2 for models 3M-B and 3M-C). In all the cases, the obtained stationary QSS is stable, all system Jacobian eigenvalues $\lambda(\boldsymbol{J})$ being real and negative for $[P2P2]_s > 1$ and $[P3P3]_s > 1$. In the other hand, Max$|\lambda(\boldsymbol{J})|$ value, measuring the QSS-strength, is of 9×10^5 for model 3M-B and $[P2P2]_s = [P3P3]_s = 2$ nM, of 4.1×10^6 for model 3M-B and $[P2P2]_s = [P3P3]_s = 10$nM, and of 3.8×10^5 for model 3M-C and $[P2P2]_s = [P3P3]_s = 2$nM. In conclusion, model 3M-B seems to predict an increased QSS-strength, while the increase in $[P2P2]_s$ and $[P3P3]_s$ levels exacerbate this trend. As observed by Salis & Kaznessis,[16] the use of both cross- and self-repression in model 3M-C slowly decrease the over-expressed protein level (e.g. P2), but will attenuate the repression of the other protein (e.g. P3, see QSS in Table 2). As another remark, the adjustable levels of $[P2P2]_s$ and $[P3P3]_s$ can indirectly control the overshoot in the level of enzymes repressing or de-repressing the gene expression.

In order to evaluate the bistable-switch responsiveness to external stationary perturbations, one change instantaneously the level of one of the stimulus $[NutI2]_s$ or $[NutI3]_s$ from nearly 0 to 10nM. As a result, one of the switch-genes (G2 or G3) is over-expressed, while the other is repressed, and the species concentrations evolve toward a new QSS. The species transient times τ_j (with a 1% tolerance), their $AVG(\tau_j)$ and $STD(\tau_j)$ are presented comparatively in Fig. 4 for all the models (3M-A, 3M-B, 3M-C). It is to remark that, by explicitly including the rapid buffering reactions which adjust the G2/G3 catalytic activity via Gi(PjPj)1 regulatory units, a more accurate representation of transient times is obtained. Thus, τ_j predicted by model 3M-B for target proteins P2 and P3 are smaller in comparison to the model 3M-A prediction, while the model 3M-C predicts the best responsiveness. The realized AVG and STD for 3M-C are close to the experimental observations of Voit[33] even if accounting for the whole cell content, ballast effect and the variable-volume equations. As a conclusion, it appears that by only using lumped power-law terms in constructing whole-cell dynamic models with modular chains, bias results can be obtained. Adjustments in local/holistic properties can be performed by means of adjustable intermediate levels included in some elementary reactions.

It is also to remark that the use of self-repression together with cross-repression in the model 3M-C considerably reduces the transient times, comparatively to only cross-repression in models 3M-A and 3M-B, being in concordance with the observations of Salis & Kaznessis[16]. At the same time, the transcriptional control with multiple operators binding repressor dimers Gi(PjPj)n is highly effective, increasing the switch-sensitivity to the concentration of repressor in solution. Moreover, the quickly adjustable dimer levels, via fast $Pj + Pj \underset{k_{diss}}{\overset{k_{bind}}{\rightleftharpoons}} PjPj$ reactions, will confer more flexibility to the switch in ranging the stability and dynamic characteristics. As revealed by Fig. 4, a higher level of $[PjPj]_s$, for instance 10nM comparatively to 2nM in model 3M-B, will slowly decrease the switch-responsiveness, i.e. will increase the AVG of species transient times in spite of a better QSS-stability strength. Because a too high level of $[PjPj]_s$ is not energetically favourable, an optimal value is expected according to a certain efficiency criterion.

The species level sensitivity vs. stationary perturbations in the environment (NutI2 or NutI3), comparatively presented in Table 2 (3-5th columns) for all models, reveal comparable values, excepting those of MetG2/MetG3 species for model 3M-B due to the insufficiently balanced repression of G2/G3 activity.

To prove the effectiveness of keeping some intermediates and explicit elementary reactions in lumped power-law/Hill-type models when the whole-cell properties are analysed, one applies a dynamic impulse-like ± 10 % $c_{P2,s}$ perturbation, and one determines the species recovering times $\tau_{rec,j}$ to their QSS-levels (with a 1% tolerance). As presented in Fig. 5, the predicted QSS-recovering times are negligible (being smaller than 1 min) for both models 3M-B and 3M-C, the regulatory efficiency being in concordance to the observations[36] and of better quality than the basic Hill-type model 3M-A.

Conclusions

Modelling synthetic gene circuits for *in-silico* GRC-design is an important step in advancing the understanding on the regulatory cell network, with important theoretical and practical implications. The modular approach, with accounting for both local and holistic GRC properties and observations from bio-molecular databanks, makes this computational approach effective allowing: similarity analysis of models (structure vs. predictions); lumping analysis; system characterization (QSS-multiplicity, stability, flexibility, robustness, efficiency); system modularisation and development of cell simulation platforms.

Even if a generic bistable switch from *E. coli* has been analysed, the variable-volume and whole-cell modelling framework, with explicitly considering the link between the volume-growth and the reaction rates for all species into the cell, appears to be a more promising approach to evaluate the GRC characteristics in a cell, by mimicking the equilibrated or perturbed growth. Such models can avoid over-estimation of some regulatory properties (i.e. responsiveness, efficiency, connectivity),[1,2,27,28] accounting for the role of cell-ballast in smoothing internal/external perturbations, for direct or indirect perturbations of species levels (transmitted via chain reactions and cell-volume variation).

Lumping rules are proved to be effective tools for modelling the cell regulatory process complexity and dynamics, coping with the cell-system low observability, identifiability and estimability. Power-law or Hill-type representation of modular GRC, including apparent rate constants, can reproduce a wide-range of cell functions and dynamic behaviour. However, the model predictability is strongly dependent on the lumping degree, on the key-species selection and ability to realise the suitable trade-off between model simplicity, its predictive power and physical-meaning of terms. A sensitivity analysis applied to model terms can help in relating the GRC holistic properties to the individual regulatory module structure.

Under more complex modelling framework, including the variable cell-volume and whole-cell approach, the requirement to realise a fine tuning of the GRC regulatory properties leads to develop mixed models that include both lumped but also elementary steps and key-intermediates of adjustable levels, able to range the system stability strength, its responsiveness and selectivity to external stimuli, and the regulatory efficiency in the gene-expression modules. Exemplification is made by simulating the dynamics of a bistable genetic-switch placed in a *E. coli* cell, by considering the cell content influence on the GRC characteristics, and a Hill-type activation completed by explicit rapid buffering reactions with adjustable intermediate levels controlling the catalyst activity. Such a modelling approach presents better possibilities to: range the QSS-stability strength (margins of stability); optimise the regulatory efficiency for key-species synthesis (increasing recovering rates); improve the sensitivity and quick responsiveness of gene-expression to external stimuli (reducing the transient times); tune interactions between cross- and self-repression of gene-expression (increasing the switch responsiveness).

Further model developments can consider the separate role of transcriptional and translational control of the gene-expression, by including [G(PP)n;mRNA(PP)n'] regulatory units, as studied by Maria[1,2], instead of simply [G(PP)n] units. The role of randomness in controlling the gene expression, probable faulty switches in a small number of cells from the colony, and other random processes can be better treated with a stochastic modelling approach.[16,17,21]

Nomenclature

A, B – rate constants
c_j – species (lump, or 'pool') concentration
D – cell content dilution rate (i.e. cell-volume logarithmic growing rate)
$\boldsymbol{g}$, $\boldsymbol{h}$ – vectors of model parameters
$\boldsymbol{g}$ – kinetic model function vector
$\boldsymbol{J} = d\boldsymbol{g}/d\boldsymbol{c}$ – kinetic model Jacobian matrix
$\boldsymbol{k}$ – kinetic constant vector
K, K_A, K_R – equilibrium or kinetic constants
n – Hill-coefficient
n_j – species j number of moles, or number of effector species binding the 'catalyst' L
n_s – no. of species
r_j – species j reaction rate
R – universal gas constant
$S(y;x) = \partial \ln(y)/\partial \ln(x)$ – relative sensitivity of y vs. x
t – time
t_c – cell-cycle time
T – temperature
V – cell volume
$\boldsymbol{X}$ – state vector

Greeks

α, β – vectors of model parameters
$\lambda(\boldsymbol{J})$ – eigenvalues of the dynamic model Jacobian
π – osmotic pressure
$\tau_{rec,j}$ – species j recovering time of the steady-state
τ_j – species j transition time from one steady-state to another

Index

bind – binding
cyt – cytoplasma
diss – dissociation
env – environment
max – maximum value
o – initial
ref – reference
s – steady-state
syn – synthesis

Superscript

T – transpose

Abbreviations

AA – aminoacids
Act – activator
AVG(τ_j) – average of τ_j
G – gene
GRC – genetic regulatory circuits
GRN – genetic regulatory networks
In, I – inducer
L – species at which regulatory element acts
M – mRNA
Met – metabolite
nM – nmol L^{-1}, nano-molar (i.e. 10^{-9} mol L^{-1} concentration)
Nut – nutrient
O, O(.) – operator, or order of magnitude
ODE – ordinary differential equations
P – protein
P.I. – regulatory performance index
QSS – quasi steady-state
R, Rep – repressor
S – substrate
STD(τ_j) – standard deviation of τ_j
tr(.) – trace
TF – transcription factor
[.] – concentration

References

1. *Maria, G.,* Modular-based modelling of protein synthesis regulation, Chemical and Biochemical Engineering Quarterly **19** (2005) 213-233.
2. *Maria, G.*, Application of lumping analysis in modelling the living systems. A trade-off between simplicity and model quality, Chemical and Biochemical Engineering Quarterly **20** (2006) 353-373.
3. EcoCyc, Encyclopedia of Escherichia coli K-12 genes and metabolism, SRI Intl., The Institute for Genomic Research, Univ. of California at San Diego, 2005. http://ecocyc.org/
4. KEGG, Kyoto encyclopedia of genes and genomes, Kyoto University and Univ. of Tokyo, 2007. /www.genome.jp/keg/
5. BRENDA, The comprehensive enzyme information system, Univ. of Köln, 2007. /www.brenda.uni-koeln.de/
6. Prodoric, Prokaryotic database of gene regulation, TU Braunschweig, 2007. /www.prodoric.tu-bs.de/
7. *Benner, S. A., Sismour, A. M.*, Synthetic biology, Nat. Rev. Genet. **6** (2005) 533-543.
8. *Voigt, C. A., Keasling, J. D.*, Programming cellular function, Nat. Chem. Biol. **1** (2005) 304-307.

9. *Heinemann, M., Panke, S.*, Synthetic Biology – putting engineering into biology, Bioinformatics **22** (2006) 2790-2799.
10. *Sprinzak, D., Elowitz, M.B.*, Reconstruction of genetic circuits, Nature **438** (2005) 443-448.
11. *Dueber, J. E., Yeh, B. J., Chak, K., Lim, W. A.*, Reprogramming control of an allosteric signalling switch through modular recombination, Science **301** (2003) 1904-1908.
12. *Chan, L. Y., Kosuri, S., Endy, D.*, Refactoring bacteriophage T7, Mol. Syst. Biol. **1**:2005.0018 (2005).
13. *Schügerl, K., Bellgardt, K. H.*, Bioreaction engineering, Springer Verlag, Berlin, 2000.
14. *Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D., Keasling, J. D.*, Engineering a mevalonate pathway in E. coli for production of terpenoids, Nature Biotechnol. **21** (2003) 796-802.
15. *Selvarasu, S., Lee, D. Y., Karimi, I. A.*, Identifying synergistically switching pathways for multi-product strain improvement using multiobjective flux balance analysis, Proc. ESCAPE-17 Symp., Bucharest, 2007.
16. *Salis, H., Kaznessis, Y.*, Numerical simulation of stochastic gene circuits, Comp. Chem. Eng. **29** (2005) 577-588.
17. *Kaznessis, Y. N.*, Multi-scale models for gene network engineering, Chemical Engineering Science **61** (2006) 940-953.
18. *Atkinson, M. R., Savageau, M. A., Myers, J. T., Ninfa, A. J.*, Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli, Cell **113** (2003) 597-607.
19. *Klipp, E., Nordlander, B., Krüger, R., Gennemark, P., Hohmann, S.*, Integrative model of the response of yeast to osmotic shock, Nature Biotechnology **23** (2005) 975-982.
20. *Chen, M. T., Weiss, R.*, Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signalling elements from Arabidopsis thaliana, Nat. Biotechnol. **23** (2005) 1551-1555.
21. *Tian, T., Burrage, K.,* Stochastic models for regulatory networks of the genetic toggle switch, Proc. Natl. Acad. Sci. USA. **103** (2006) 8372-8377.
22. *Sotiropoulos, V., Kaznessis, Y. N.*, Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes, BMC Syst Biol., Jan 9 (2007) 1-7 (doi:10.1186/1752-0509-1-7).
23. *Tomshine, J., Kaznessis, Y. N.*, Optimization of a stochastically simulated gene network model via simulated annealing, Biophys J. **91** (2006) 3196-3205.
24. *Zhu, R., Ribeiro, A. S., Salahub, D., Kauffman, S. A.*, Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models, J Theor Biol. **246** (2007) 725-45.
25. *Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R., Collins, J. J.*, Programmable cells: Interfacing natural and engineered gene networks, PNAS **101** (2004) 8414–8419
26. *Morgan, J. J., Surovtsev, I. V., Lindahl, P. A.*, A framework for whole-cell mathematical modelling, Jl. theor. Biology **231** (2004) 581-596.
27. *Yang, Q., Lindahl, P., Morgan, J.*, Dynamic responses of protein homeostatic regulatory mechanisms to perturbations from steady state, J. theor. Biol. **222** (2003) 407-423.
28. *Maria, G.*, Evaluation of protein regulatory kinetics schemes in perturbed cell growth environments by using sensitivity methods, Chemical and Biochemical Engineering Quarterly **17** (2003) 99-117.
29. *Styczynski, M. P., Stephanopoulos, G.*, Overview of computational methods for the inference of gene regulatory networks, Comp. Chem. Eng. **29** (2005) 519-534.
30. *Zak, D. E., Vadigepalli, R., Gonye, G. E., Doyle III, F. J., Schwaber, J. S., Ogunnaike, B. A.*, Unconventional systems analysis problems in molecular biology: A case study in gene regulatory network modelling, Comp. Chem. Eng. **29** (2005) 547-563.
31. *Angeli, D., Ferrell, J. E. Jr., Sontag, E. D.*, Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems, Proceedings of the National Academy of Sciences of the USA **101** (2004) 1822-1827.
32. *Franck, U. F.*, Feedback kinetics in physicochemical oscillators, Ber. Bunsenges. Phys. Chem. **84** (1980) 334-341.
33. *Voit, E. O.*, Smooth bistable S-systems, IEE Proc. Syst. Biol. **152** (2005) 207-213.
34. *Savageau, M. A.*, Alternatives designs for a genetic switch: Analysis of switching times using the piecewise power-law representation, Math. Biosciences **180** (2002) 237-253.
35. *Savageau, M. A., Voit, E. O.*, Recasting nonlinear differential equations as S-systems: A canonical nonlinear form, Math. Biosciences **87** (1987) 83-115.
36. *Hlavacek, W. S., Savageau, M. A.*, Completely uncoupled and perfectly coupled gene expression in repressible systems, J. Mol. Biol. **266** (1997) 538-558.
37. *Van Someren, E. P., Wessels, L. F. A., Backer, E., Reinders, M. J. T.*, Multi-criterion optimization for genetic network modelling, Signal Processing **83** (2003) 763-775.
38. *Schwacke, J. H., Voit, E. O.*, Improved methods for the mathematically controlled comparison of biochemical systems, Theoretical Biology and Medical Modelling **1** (2004) 1-18.
39. *Eden, P.*, Course on bioinformatics and computational biology, Faculty of Medicine, Dept. of Theoretical Physics, Lund University (Sweden), 2003. webmaster@thep.lu.se.
40. *Ofiteru, D. I.*, Optimal control strategies for bioprocesses, PhD Thessis, University Politehnica of Bucharest, 2006.
41. *Alon, U.*, An introduction to systems biology. Design principles of biological circuits, Chapman & Hall / CRC, Boca Raton, 2007.
42. *Segel, I. H.*, Enzyme kinetics: Behaviour and analysis of rapid equilibrium and steady-state enzyme systems, Wiley, New York, 1975.
43. *Rosenfeld, N., Elowitz, M. B., Alon, U.*, Negative autoregulation speeds the response times of transcription networks, J. Mol. Biol. **323** (2002) 785-793.
44. *Elowitz, M. B., Leibler, S.*, A synthetic oscillatory network of transcriptional regulators, Nature **403** (2000) 335-338.
45. *Sewell, C., Morgan, J., Lindahl, P.*, Analysis of protein regulatory mechanisms in perturbed environments at steady state, J. theor. Biol. **215** (2002) 151-167.
46. *Kholodenko, B. N., Kiyatkin, A., Bruggeman, F. J., Sontag, E., Westerhoff, H. V., Hoek, J. B.*, Untangling the wires: A strategy to trace functional interactions in signalling and gene networks, Proceedings of the National Academy of Sciences of the USA **99** (2002) 12841-12846.
47. *Silva, M. R.*, Bioinformatics tools for the visualization and structural analysis of metabolic networks, PhD Thesis, TU Braunschweig (Germany), 2006.
48. *Ma, H. W., Zhao, X. M., Yuan, Y. J., Zeng, A. P.*, Decomposition of metabolic network into functional modules based on the global connectivity structure of reaction graph, Bioinformatics **20** (2004) 1870-1876.
49. *Wall, M. E., Hlavacek, W. S., Savageau, M. A.*, Design principles for regulator gene expression in a repressible gene circuit, J. Mol. Biol. **332** (2003) 861-876.
50. *Jacob, F., Monod, J.*, Genetic regulatory mechanisms in the synthesis of proteins, J. Mol. Biol. **3** (1961) 318-356.

51. *Xu, H., Moraitis, M., Reedstrom, R. J., Matthews, K. S.*, Kinetic and thermodynamic studies of purine repressor binding to corepressor and operator DNA, Jl. Biological Chemistry **273** (1998) 8958-8964.
52. *Meng, L. M., Kilstrup, M., Nygaard, P.*, Autoregulation of PurR repressor synthesis and involvement of purr in the regulation of purB, purC, purl, purMN and guaBA expression in Escherichia coli, Eur. J. Biochem. **187** (1990) 373-379.
53. *Shen-Orr, S. S., Milo, R., Mangan, S., Alon, U.*, Network motifs in the transcriptional regulation network of Escherichia coli, Nature Genetics **31** (2002) 64-68.
54. *Riznichenko, G. Y.*, Mathematical models in biophysics, Biological Faculty, Lomonosov Moscow State University, 2002; www.biophysics. org/btol/img/math-models.pdf
55. *Murray, J. D.*, Mathematical biology. I An introduction, Springer Verlag, Berlin, 2001.
56. *Heinrich, R., Schuster, S.*, The regulation of cellular systems, Chapman & Hall, New York, 1996.
57. *Agrawal, P., Lee, C., Lim, H. C., Ramkrishna, D.*, Theoretical investigations of dynamic behavior of isothermal continuous stirred tank biological reactors, Chemical Engineering Science **37** (1982) 453-462.
58. *Savageau, M. A.*, Concepts relating the behaviour of biochemical systems to their underlying molecular properties, Archives of Biochemistry and Biophysics **145** (1971), 612-621.
59. *Savageau, M. A., Sorribas, A.*, Constraints among molecular and systemic properties: Implications for physiological genetics, Journal of Theoretical Biology **141** (1989) 93-115.
60. *Hlavacek, W. S., Savageau, M. A.*, Subunit structure of regulator proteins influences the design of gene circuitry: Analysis of perfectly coupled and completely uncoupled circuits, Journal of Molecular Biology **248** (1995), 739-755.
61. *Hlavacek, W. S., Savageau, M. A.*, Rules for coupled expression of regulator and effector genes in inducible circuits, Journal of Molecular Biology **255** (1996), 121-139.
62. *Grainger, J. N. R., Gaffney, P. E., West, T. T.*, A model of a growing steady-state system with a changing surface-volume ratio, J. theor. Biol. **21** (1968) 123-130.
63. *Wallwork, S. C., Grant, D. J. W.*, Physical Chemistry, Longman, London, 1977.
64. *Allen, G. C. Jr., Kornberg, A.*, The priB gene encoding the primosomal replication n-protein of E. coli, Jl. Biological Chemistry **266** (1991) 11610-11613.