# Evaluation of Protein Regulatory Kinetics Schemes in Perturbed Cell Growth Environments by Using Sensitivity Methods

**G. C. Maria**
Laboratory of Chemical and Biochemical Reaction Engineering,
University Politehnica, Bucharest, Romania
(gmaria99m@hotmail.com)

Original scientific paper
Received: January 30, 2002
Accepted: December 9, 2002

Homeostatic regulatory mechanisms are developed to study the recovering characteristics of the cell components to their steady-states after being perturbed. Such simplified kinetic mechanisms mimic the balanced cell-growth via quasi-invariant cell-component's levels, despite continuous volume growing diluting effects and perturbations in the environment. Recently, *Sewell* et al.[1] proposed nine kinetic schemes to regulate a generic protein P, based on a limited number of reactions. Nominal steady-state (QSS) concentrations of the protein and their encoding gene and maximum regulatory effectiveness constraints allowed ranking the kinetic schemes according to their stationary effectiveness. *Yang* et al.[2] extended this analysis by applying a maximum recovering rate objective function to estimate the rate constants. In the present study, a sensitivity-stability-based analysis of the QSS complete the previous investigations of regulatory models, by accounting the QSS-response surface to stationary perturbations in the P synthesis/dilution rate, characterisation of the QSS quality, sensitivity and local stability, and approaching a multi-criteria ranking of regulatory mechanisms.

*Keywords:*
Protein and gene regulatory kinetic models, cell regulatory module design, sensitivity and stability analysis

## Introduction

Living systems are evolutionary autocatalytic structures capable of converting raw materials from their environment into additional copies of themselves. To explain such a process at a cell level, an impressive number of studies have been reported over the last decades. Although an enormous amount of information on proteins and other molecules inside cells is stored on computer databanks, no consensus exists on how to analyse and model the genomes, the cell metabolism, and inside regulatory functions.

To model the entire cell-system complexity becomes a very difficult computational task because of the large number of inner cell species, successive enzymatic / autocatalytic processes involving a large number of enzymes of various activities, proteinic oligomers, complex intermediates and regulatory chains, cell signalling, motility, organelle transport, gene transcription, morphogenesis and cellular differentiation. Several approaches and a large number of packages have been reported in literature over the last years, several developments being oriented toward area of commercial interest such as biotechnology, food and drug industry, and neurobiology.

One aspect of complex cell-metabolism is to include homeostatic regulatory mechanisms to return most of the cell components to their steady-state (or near steady-state) after being perturbed. Such regulatory kinetic mechanisms ensure a balanced cell-growth via quasi-invariant cell-component's levels despite continuous diluting effects and other perturbations in the environment (nutrient) levels. The maintained metabolic functions achieve the end goal of cell-division and replication. Regulatory mechanisms are mainly dominated by negative feedback elements that control transcriptional processes.[3,4]

The regulatory networks, implying a large number of proteins, are poorly understood. They appear to be organised hierarchically, consisting either of a large number of strongly interacting components or of smaller numbers of weakly interacting groups of reactions or modules. Each regulatory module can be analysed individually, and then combined in a functional organized hierarchy, each one completing a precise cell-function (for proteins, metabolites, membrane processes). Interactions among regulatory kinetic modules happens all-the-time due to complex reactions in which proteins/genes from different modules interact, thus producing mutual perturbations. As the module presents a better regulatory effectiveness, their 'sensitivity' to internal/external perturbations is smaller. If the linkage rates are slow, relative to the core rates of each module,

the proteins in the linked system would remain regulated. To ensure that individual perturbations would not combine to yield an overall perturbation that exceeds the regulatory capacity of any individual module, the magnitude of each linking interaction would have to decline as the number of connections increased. As a higher level, the cell functional hierarchy ensures the cell-self-replication and their global auto-regulation.

Recently, *Sewell* et al.[1] proposed nine ODE (ordinary differential equations) kinetic models to regulate a generic protein P, based on a limited number of elementary reactions for protein synthesis and dilution. The constant volume assumption in the model was corrected by adding a P 'decay' rate to mimic the real cell-volume growth diluting effect. Each proposed scheme is analysed for its capabilities to maintain a nominal P concentration level within a certain small interval of tolerance when continuous perturbations arise. Nominal QSS concentrations have been imposed for the protein P and their encoding gene G to symbolize transcriptional / translational processes for an average protein in *Escherichia coli* cell, allowing to determine average rate constants for the P-synthesis and "consumption" (e.g. dilution) rates. In the absence of systematic experimental kinetic data, supplementary constraints of maximum regulatory effectiveness of the mechanism at QSS, have been used to set some of the rate constants. Such a qualitative-quantitative analysis allowed ranking the regulatory kinetic schemes according to their stationary regulatory effectiveness vs. unsynchronised/synchronised 'step' perturbations in P-synthesis and dilution. The analysis derived useful principles to design modular cell-regulatory schemes, highlighting the role of: cross-autocatalytic P/G synthesis, cell volume growth, reaction invariants from QSS conditions, external factors, response to stationary and dynamic perturbations. Such an approach allows developing whole-cell models by including linked modules, which ensure the main cell-functions of interactivity, regulation, growth, replication, and division.

*Yang* et al.[2] extended the previous analysis of regulatory modules by including new regulatory elements according to previous elaborated principles. In the absence of kinetic data, a maximum recovering rate objective function has been applied to estimate the rate constants for a nominal QSS. The formulated ODE models have been used to rank the mechanisms regulatory effectiveness and to study particular role of reactions and components.

When no standard kinetic data are available, construction of regulatory kinetic models, rate constant estimation, and their analysis have to account rather disparate qualitative/quantitative information from databanks. This is why the rate constants can be also derived by optimising non-conventional estimation criteria such as: faster recovering rate / smallest recovering time after an 'impulse' / dynamic perturbation;[2,5] smallest amplitude of the recovering path; smallest sensitivity of the QSS vs. perturbations; maximum QSS stability; imposed periodicity to the reaction cycle;[6] fulfilment of a nominal QSS and some physical constraints; inflexibility to external condition changes (nutrient's levels); imposed succession of metabolic events into the cell, etc. Each of these criteria can lead to a significant different estimate, which in turn can offer another perspective for analysing the cell-regulatory system. For a real cell simulation, it is probable that a multi-objective criterion, with weights in accordance with the main cell-functions, will offer a more complete process characterisation.

Design of regulatory module, with a specific function into the cell, represents an interesting and promising modelling approach. These simple hypothetical mechanical models can simulate homeostatic regulation functions for one generic component by lumping species and reactions in generic groups. The advantage of this route devolves from the possibility to generalise the simulation conclusions, and to consider some whole-cell properties without requiring currently unavailable details of all the components and reactions of a real cell.

The scope of this paper is to follow this regulatory module design route, by completing the previous kinetic mechanism analysis with a sensitivity-stability-based investigation of cell behaviour around a QSS. The applied procedure allows to estimate the model rate constants, to characterise the QSS response surface under various stationary P-synthesis / dilution perturbations, to characterise the QSS quality and local stability, and to rank the protein regulatory schemes effectiveness.

## Reviews of whole-cell and protein regulatory models

While valuable databanks of individual cells have been developed (such as WormBase, EcoCyc[7]), several ambitious projects have been started with the aim to simulate entire cells or part of cells at a molecular level: **E-Cell** (whole-cell simulation), **V-Cell** (virtual cell modelling), **M-Cell** (a general Monte Carlo simulator of cellular microphysiology), **A-Cell** (construction of biochemical reaction and electrical equivalent circuit models), etc.

The **E-Cell** software allows simulating reaction pathways within a compartment-based cell model, including: nucleotide, phospholipid, and amino acid biosynthesis, energy metabolism, and gene expression.[8] The package works with four predefined

types of subspace: 'environment' corresponding to the extra-cellular space, 'cell' corresponding to the entire cell, 'cytoplasm' corresponding to a region inside the cell, and 'membrane' corresponding to a boundary between any of the others. Simulation of a pre-defined reaction pathway results in dynamic species evolution.[9] The software does not directly account for the cell-volume growth and diffusion effects, genome replication, and cell-division. *Tomita* et al.[10] used the **E-Cell** model linked with the EcoCyc and **KEGG** databases to simulate the dynamics of 127 genes/proteinic system for the *M. genitalium* cell, and some improvements have been recently reported.[11]

The **V-Cell** package,[12] similarly to Cell ML-package, builds-up a cell framework with compartments and membranes, each one including species, reactions and membrane fluxes. The user enters the steady/unsteady mass balance equations into the program, being then solved and analysed for sensitivity in parameters and automatic simplifications by using pseudo-steady-state assumptions. The three package objects (models, geometry and application) with including species diffusion terms allow a spatial dynamic simulation of the concentration field into the cell.

The **M-Cell** package[13] allows simulating high-level complex sub-cellular communications, together with proteins and enzymes involved in exo-/endo-cytosis, synaptic transmission, transport and signal reception. The **M-Cell** simulator, positioned at a biological scale "above molecular dynamics but below whole cell and higher level studies", includes several stochastic algorithms, as Brownian random walk and Monte Carlo, to model the diffusing ligands interactions with individual 3D binding sites. **M-Cell** simulator can thus provide accurate behaviour for the interacting molecules present in the small amounts in a spatially complex environment, and massive parallel computations to reproduce the 3D arrangements of diffusion boundaries and molecules.

The **A-Cell** platform of *Ichikawa*[14] allows constructing a biochemical reaction schema and electrical models for a biological cell and a neuron, by using pre-defined cell-reaction-path, rate constants, and an initial condition.

Several other modelling software and simulators for biological systems have been reported in the literature, including representation of cells, neurons, bio-informatic sequences, bio-polymer sequences, complex molecular structures, gene expression, gene-finding.[14–16] A special attention is paid to the unified SBML language for biology mark-up,[16] which includes representative features of most of the previous mentioned software. The package, based on the XML language, attempts to include the basic route of modelling cellular systems through ODE, DAE (differential-algebraic), and stochastic simulators. A minimal SBML model definition must include most of the following components: compartments (finite well-mixed volumes), species, reactions, model parameters, units, and rules (constraints). Higher-level definitions will add hierarchical sub-models, spatial geometry, and array of components. Given the astronomical complexity and unknown aspects of cell-systems, formulating reliable models with predictive ability remains still a dream. However, advances in genomics and related areas provide hope that this dream might be realized in the near future.

In this computational environment, conventional continuous and mechanistic models can provide a convenient simulation of the cell systems when molecular details are of little importance for the analysis. Such models can be useful for analysing regulatory cell-functions, both for stationary and dynamic perturbations. When species spatial location becomes important, other types of representation can be more successful, for instance the stochastic models. The small number of molecules for a certain species is more sensitive to stochasticity of a reaction process than the species present in larger amounts, and the use of stochastic models can increase the simulation accuracy. In such models, the species concentrations are replaced by individual molecular species, and Monte Carlo methods are used to predict their interactions. Rate equations are replaced by individual reaction probabilities while the model output is stochastic in nature.

One of the key points in elaborating a whole-cell model is the design of regulatory kinetic modules for adjusting cell-component concentrations. A simple regulatory module for a generic protein includes the protein itself, their encoding gene, external/internal 'raw-materials' for protein/gene synthesis, intermediates, oligomers, and by-products. To construct such modules, several alternatives have been considered in the literature: (i) continuous variable ODE models;[17] (ii) discrete (Boolean) variable models;[18,19] iii) mixed continuous-discrete variable models; (iv) stochastic variable models.[20–22] The mechanistic ODE models have been proved to be effective in describing simple cell-systems, accurately predicting the continuous variable evolution as well as continuous perturbations. The Boolean approach, even less realistic, is more computationally tractable for complex systems, involving networks of genes that are either "on" or "off" according to defined Boolean relationships. The mixed models can realise a promising compromise among continuous and discrete variables to model complex systems. The stochastic models can replace the 'average' ODE model solution by a more detailed random-based simulator, but with the expense of a considerable computational effort.

## Homeostatic regulatory kinetic schemes

*Sewell* et al.[1] proposed nine kinetic mechanisms (denoted with RM1-RM9 in Table 1) to homeostatically regulate an average generic protein P in a prokaryote cell such as *E. coli*.[23] In each mechanism, the core cell includes: a) the environmental (nutrients) influence through P synthesis reactions; b) a generic gene G and protein P related to each other through autocatalysis; c) regulatory reactions that facilitate internal homeostasis in a changing cell-volume and environment; and d) perturbation reactions for P-synthesis and P-dilution (due to the cell-growth). Additionally regulatory elements were added allowing the influence of each element to be assessed.

P symbolizes virtually all proteins in a real cell, and has several functions: it is a *permease* that catalyses the import of nutrients into the cell; a *metabolase* that converts those nutrients into intracellular metabolites; a *polymerase* that catalyses the synthesis of G from metabolites. To simplify the regulatory effectiveness analysis, these functions have not been explicitly included in the RM1-RM9 models, while the overall G-level (active and inactive forms) was set constant. A limited number of other species were incorporated: catalytic inactive intermediates (GP, GPP, GPPPP), protein oligomers (PP), mRNA (called M) and their inactive form (MP) involved in a cascade P-synthesis process. The schemes were step-by-step built-up starting from the simplest one and adding reactions according to a certain principle and desired function. The same rule can be continued to develop even more complex/effective regulatory mechanisms[2] but a trade-off between model complexity/identifiability and their effectiveness has to be realised.

The nine protein regulatory mechanisms include consecutive-parallel steps for the enzymatic protein synthesis and consumption, as follows:

– *RM1* ("unregulated mechanism") includes only a G-catalysed synthesis of P and their subsequent first-order decay by dilution; the mechanism has no regulatory elements and is the control scheme against which all others are compared;

– *RM2* ("basic negative feedback") adds to RM1 a transcription control of P/G ratio through their reversible transformation into a catalytically inactive form GP; the equilibrium dissociation constant for this reaction was set to the QSS concentration of protein ($c_{P,s}$) to achieve a maximum regulation sensitivity;

– *RM3* ("multiple binding feedback") adds to RM2 a more accurate description of the G-level control through two successive reversible reactions and inactive forms GP and GPP; prokaryotes commonly bind multiple copies of transcription factors to improve regulatory effectiveness;

– *RM4* ("Boolean negative feedback") considers the RM2 regulatory mechanism with a discontinuous evolution of G over three possible concentration levels $c_G / c_{G_{tot}} = \{0, 0.5, 1\}$ according to the P-level, as an idealised best-case sensitivity;

– *RM5* ("dimerization") considers an intermediate protein P dimerization to PP before the reversible control of G via a reversible PP binding to an inactive form GPP; most, if not all transcription factors, bind promoter sites as oligomers, and GPP mimics this cell property;

– *RM6* ("two-dimer binding") adds to RM5 a more accurate G level control and fast G/P ratio adjustment via two inactive enzymatic forms GPP and GPPPP successively bind on the dimer PP;

– *RM7* ("transcription/translation cascade") introduces a new enzyme (M) responsible for the protein P synthesis; the enzyme M (in fact the mRNA acid) is synthesized from nucleotides under G catalysis, while in a translation step, P is synthesized in a reaction catalysed by M; the control of the G/P ratio through the binding species GP is maintained;

– *RM8* ("transcriptional and translational feedback") improves the RM7 regulatory effectiveness by adding, analogous to the RM2 scheme, a control of the M level through an inactive form MP, to stimulate its degradation;

– *RM9* ("degradation negative feedback") improves the RM2 regulatory scheme by adding an irreversible P-degradation reaction catalysed by the intermediate GP derived from G.

## ODE model formulation

To model the RM1-RM9 homeostatic P-regulatory mechanisms, the constant volume ODE model hypotheses of *Sewell* et al.,[1] have been adopted. In such a continuous formulation, the cell-species concentrations are continuous variables with feasible positive values between zero and large copynumbers. The ODE model reflects in fact the average states of the stochastic processes occurring during the cell metabolism.[5] As a consequence, fractional concentrations can occur during cell simulations with a continuous ODE model. Because 1 nmol $l^{-1}$ concentration corresponds to a copynumber, when treated deterministically, for instance a gene concentration of $c_G$ = 0.5 nmol $l^{-1}$ must be interpreted either as a time-invariant average in a population of cells (e.g. half of all cells contain 1 copynumber of G) or as a time-dependent averages for a single cell

(e.g. that cell contains 1 copynumber of G half of the time).

Irrespectively how the reaction schema was built-up, the ODE model consists of a differential set including the stoichiometric matrix $\nu$ defined over all reactions and all species, with coefficients $\nu_{ij}$ ($j$ = 1,…, $n_s$, number of species; $i$ = 1,…, $n_r$, number of elementary reactions). Thus, a species $j$ reaction rate results as a sum of reactions ($r_i$) in which that species is involved, i.e.:

$$\begin{aligned} \mathrm{d}c_j/\mathrm{d}t &= \sum_{i=1}^{n_r} \nu_{ij} r_i\,; \qquad j = 1,\ldots,n_s, \\ r_i &= k_i \prod_{j=1}^{n_s} c_j^{\beta_{ij}} - k_i' \prod_{j=1}^{n_s} c_j^{\gamma_{ij}}, \end{aligned} \tag{1}$$

($c_j$ = concentration of species $j$; $r_i$ = reaction rate $i$; $k$ = rate constants; $\nu_{ij} < 0$ for reactants and $\nu_{ij} > 0$ for products). If the reactions are considered elementary, the partial orders of reaction $\beta$ and $\gamma$ are identical with the corresponding stoichiometric coefficients. Table 1 displays the RM1-RM9 model rate equations. The constant volume model formulation implies the following hypotheses:

(i) The cell is considered an isotherm system, with a uniform composition field (ideal mixed system). The mass diffusional transport resistance into the cell is neglected.

(ii) The diffusional resistance to the mass transport of the external 'raw-materials' (not accounted as distinct species) through the cell-membrane into the cell is neglected.

(iii) The cell-system is considered an open system. Even if the models does not include terms for delivering by-products in the outer-cell space, further developments can easily incorporate such terms.

(iv) The external conditions are considered constant over a simulated cell-cycle.

(v) The outside-cell space, to which the external 'raw-materials' concentrations are referred, is considered constant comparatively with the growing cell volume.

Table 1 – *Reduced kinetic models for protein P synthesis and homeostatic control. Perturbation reactions are considered of the simplified form:* $*\,*\,* \xrightarrow{k_3} P \xrightarrow{k_4} *\,*\,*$. *The* $c_{G_{tot}}$ *denotes the total concentration of gene (active and inactive forms) catalysing the synthesis of protein P.*

| Kinetic model | Reaction schema | Species rates ($r_j$, $j$ = 1,…,$n_s$) | QSS mass balance equations (index 's') | Observations |
|---|---|---|---|---|
| RM1 | $*\,*\,* + G \xrightarrow{k_1} P + G$<br>$P \xrightarrow{k_2} *\,*\,*$ | $r_P = k_1 c_G - k_2 c_P + k_3 - k_4 c_P$<br>$c_G = c_{G_{tot}}$. | $c_{P,s} = (k_1 c_{G,s} + k_3)/(k_2 + k_4)$;<br>$c_{G,s} = c_{G_{tot},s}$. | species = [*P,G*];<br>$\boldsymbol{k}$ = [$k_1$, $k_2$, $k_3$, $k_4$] |
| RM2 | $*\,*\,* + G \xrightarrow{k_1} P + G$<br>$P \xrightarrow{k_2} *\,*\,*$<br>$G + P \underset{k_6}{\overset{k_5}{\rightleftarrows}} GP$ | $r_P = k_1 c_G - k_2 c_P + k_3 - k_4 c_P$;<br>$-k_5 c_G c_P + k_6 c_{GP}$;<br>$r_G = -k_5 c_G c_P + k_6 c_{GP}$;<br>$r_{GP} = k_5 c_G c_P - k_6 c_{GP}$;<br>$c_{G_{tot}} = c_G + c_{GP}$. | $c_{P,s} = (k_1 c_{G,s} + k_3)/(k_2 + k_4)$;<br>$c_{G,s} = k_6 c_{G_{tot},s}/(k_5 c_{P,s} + k_6)$;<br>$c_{G_{tot},s} = c_{G,s} + c_{GP,s}$. | species = [*P,G,GP*];<br>$\boldsymbol{k}$ = [$k_1$, $k_2$, $k_3$, $k_4$, $k_5$, $k_6$] |
| RM3 | $*\,*\,* + G \xrightarrow{k_1} P + G$<br>$P \xrightarrow{k_2} *\,*\,*$<br>$G + P \underset{k_6}{\overset{k_5}{\rightleftarrows}} GP$<br>$GP + P \underset{k_8}{\overset{k_7}{\rightleftarrows}} GPP$ | $r_P = k_1 c_G - k_2 c_P + k_3 - k_4 c_P$<br>$-k_5 c_G c_P + k_6 c_{GP}$<br>$-k_7 c_{GP} c_P + k_8 c_{GPP}$;<br>$r_G = -k_5 c_G c_P + k_6 c_{GP}$;<br>$r_{GP} = k_5 c_G c_P - k_6 c_{GP}$<br>$-k_7 c_{GP} c_P + k_8 c_{GPP}$;<br>$r_{GPP} = k_7 c_{GP} c_P - k_8 c_{GPP}$;<br>$c_{G_{tot}} = c_G + c_{GP} + c_{GPP}$. | $k_1 c_{G,s} - k_2 c_{P,s} + k_3 - k_4 c_{P,s} = 0$;<br>$c_{G,s} = \dfrac{k_6 k_8 c_{G_{tot},s}}{k_5 k_7 c_{P,s}^2 + k_5 k_8 c_{P,s} + k_6 k_8}$;<br>$c_{GP,s} = \dfrac{k_8 (c_{G_{tot},s} - c_{G,s})}{k_7 c_{P,s} + k_8}$. | species = [*P,G,GP,G,PP*];<br>$\boldsymbol{k}$ = [$k_1$, $k_2$, $k_3$, $k_4$, $k_5$, $k_6$, $k_7$, $k_8$] |
| RM4 | $*\,*\,* + G \xrightarrow{k_1} P + G$<br>$P \xrightarrow{k_2} *\,*\,*$ | $r_P = k_1 c_G - k_2 c_P + k_3 - k_4 c_P$;<br>$c_{G_{tot}} = c_G + c_{GP}$;<br>$\dfrac{c_G}{c_{G_{tot}}} = \begin{cases} 1, & c_P < c_{P,s} \\ 0.5, & c_P = c_{P,s} \\ 0, & c_P > c_{P,s} \end{cases}$. | $k_1 c_{G,s} - k_2 c_{P,s} + k_3 - k_4 c_{P,s} = 0$;<br>$c_{G,s}/c_{G_{tot},s} = 0.5$;<br>$c_{G_{tot},s} = c_{G,s} + c_{GP,s}$. | species = [*P,G,GP*];<br>$\boldsymbol{k}$ = [$k_1$, $k_2$, $k_3$, $k_4$] |

Table 1 – *(continued)*

| Kinetic model | Reaction schema | Species rates ($r_j$, $j = 1,\ldots,n_s$) | QSS mass balance equations (index 's') | Observations |
|---|---|---|---|---|
| RM5 | $*\,*\,* + G \xrightarrow{k_1} P + G$<br>$P \xrightarrow{k_2} *\,*\,*$<br>$2P \underset{k_{10}}{\overset{k_9}{\rightleftharpoons}} PP$<br>$G + PP \underset{k_{12}}{\overset{k_{11}}{\rightleftharpoons}} GPP$ | $r_P = k_1 c_G - k_2 c_P + k_3 - k_4 c_P - 2k_9 c_P^2 + 2k_{10} c_{PP}$;<br>$r_G = -k_{11} c_G c_{PP} + k_{12} c_{GPP}$;<br>$r_{GPP} = k_{11} c_G c_{PP} - k_{12} c_{GPP}$;<br>$r_{PP} = k_9 c_P^2 - k_{10} c_{PP} - k_{11} c_G c_{PP} + k_{12} c_{GPP}$;<br>$c_{G_{tot}} = c_G + c_{GPP}$. | $k_1 c_{G,s} - k_2 c_{P,s} + k_3 - k_4 c_{P,s} = 0$;<br>$c_{G,s} = \dfrac{k_{12} c_{G_{tot},s}}{k_9 k_{11} c_{P,s}^2 / k_{10} + k_{12}}$;<br>$c_{PP,s} = k_9 c_{P,s}^2 / k_{10}$;<br>$c_{GPP,s} = c_{G_{tot},s} - c_{G,s}$. | species = [*P,G,GPP,PP*];<br>$\boldsymbol{k}$ = [$k_1$, $k_2$, $k_3$, $k_4$, $k_9$, $k_{10}$, $k_{11}$, $k_{12}$] |
| RM6 | $*\,*\,* + G \xrightarrow{k_1} P + G$<br>$P \xrightarrow{k_2} *\,*\,*$<br>$2P \underset{k_{10}}{\overset{k_9}{\rightleftharpoons}} PP$<br>$G + PP \underset{k_{12}}{\overset{k_{11}}{\rightleftharpoons}} GPP$<br>$GPP + PP \underset{k_{14}}{\overset{k_{13}}{\rightleftharpoons}} GPPPP$ | $r_P = k_1 c_G - k_2 c_P + k_3 - k_4 c_P - 2k_9 c_P^2 + 2k_{10} c_{PP}$;<br>$r_G = -k_{11} c_G c_{PP} + k_{12} c_{GPP}$;<br>$r_{GPP} = k_{11} c_G c_{PP} - k_{12} c_{GPP} - k_{13} c_{GPP} c_{PP} + k_{14} c_{GPPPP}$;<br>$r_{GPPPP} = k_{13} c_{GPP} c_{PP} - k_{14} c_{GPPPP}$;<br>$r_{PP} = k_9 c_P^2 - k_{10} c_{PP} - k_{11} c_G c_{PP} + k_{12} c_{GPP} - k_{13} c_{GPP} c_{PP} + k_{14} c_{GPPPP}$;<br>$c_{G_{tot}} = c_G + c_{GPP} + c_{GPPPP}$. | $k_1 c_{G,s} - k_2 c_{P,s} + k_3 - k_4 c_{P,s} = 0$;<br>$-k_9 k_{11} c_{G,s} c_{P,s}^2 / k_{10} + k_{12} c_{GPP,s} = 0$;<br>$c_{PP,s} = k_9 c_{P,s}^2 / k_{10}$;<br>$c_{GPP,s} = \dfrac{c_{G_{tot},s} - c_{G,s}}{1 + k_{13} c_{PP,s} / k_{14}}$;<br>$c_{GPPPP,s} = k_{13} c_{GPP,s} c_{PP,s} / k_{14}$. | species = [*P,G,GPP, GPPPP, PP*];<br>$\boldsymbol{k}$ = [$k_1$, $k_2$, $k_3$, $k_4$, $k_9$, $k_{10}$, $k_{11}$, $k_{12}$, $k_{13}$, $k_{14}$] |
| RM7 | $P \xrightarrow{k_2} *\,*\,*$<br>$G + P \underset{k_6}{\overset{k_5}{\rightleftharpoons}} GP$<br>$*\,*\,* + G \xrightarrow{k_{15}} G + M$<br>$M \xrightarrow{k_{16}} *\,*\,*$<br>$*\,*\,* + M \xrightarrow{k_{17}} M + P$ | $r_P = -k_2 c_P + k_3 - k_4 c_P - k_5 c_G c_P + k_6 c_{GP} + k_{17} c_M$;<br>$r_G = -k_5 c_G c_P + k_6 c_{GP}$;<br>$r_{GP} = k_5 c_G c_P - k_6 c_{GP}$;<br>$r_M = k_{15} c_G - k_{16} c_M$;<br>$c_{G_{tot}} = c_G + c_{GP}$. | $-k_2 c_{P,s} + k_3 - k_4 c_{P,s} + k_{15} k_{17} c_{G,s} / k_{16} = 0$;<br>$c_{G,s} = \dfrac{k_6 c_{G_{tot},s}}{k_5 c_{P,s} + k_6}$;<br>$c_{GP,s} = c_{G_{tot},s} - c_{G,s}$;<br>$c_{M,s} = k_{15} c_{G,s} / k_{16}$. | species = [*P,G,GP,M*];<br>$\boldsymbol{k}$ = [$k_2$, $k_3$, $k_4$, $k_5$, $k_6$, $k_{15}$, $k_{16}$, $k_{17}$] |
| RM8 | $P \xrightarrow{k_2} *\,*\,*$<br>$G + P \underset{k_6}{\overset{k_5}{\rightleftharpoons}} GP$<br>$*\,*\,* + G \xrightarrow{k_{15}} G + M$<br>$M \xrightarrow{k_{16}} *\,*\,*$<br>$*\,*\,* + M \xrightarrow{k_{17}} M + P$<br>$M + P \underset{k_{19}}{\overset{k_{18}}{\rightleftharpoons}} MP$<br>$MP \xrightarrow{k_{20}} P + *\,*\,*$ | $r_P = -k_2 c_P + k_3 - k_4 c_P - k_5 c_G c_P + k_6 c_{GP} + k_{17} c_M - k_{18} c_M c_P + k_{19} c_{MP} + k_{20} c_{MP}$;<br>$r_G = -k_5 c_G c_P + k_6 c_{GP}$;<br>$r_{GP} = k_5 c_G c_P - k_6 c_{GP}$;<br>$r_M = k_{15} c_G - k_{16} c_M - k_{18} c_M c_P + k_{19} c_{MP}$;<br>$r_M = k_{18} c_M c_P - k_{19} c_{MP} - k_{20} c_{MP}$;<br>$c_{G_{tot}} = c_G + c_{GP}$. | $-k_2 c_{P,s} + k_3 - k_4 c_{P,s} + k_{17} c_{M,s} = 0$;<br>$c_{G,s} = \dfrac{k_6 c_{G_{tot},s}}{k_5 c_{P,s} + k_6}$;<br>$c_{GP,s} = c_{G_{tot},s} - c_{G,s}$;<br>$k_{15} c_{G,s} - k_{16} c_{M,s} - \dfrac{k_{18} k_{20} c_{M,s} c_{P,s}}{k_{19} + k_{20}} = 0$;<br>$c_{MP,s} = \dfrac{k_{18} c_{M,s} c_{P,s}}{k_{19} + k_{20}}$. | species = [*P,G,GP,M,MP*];<br>$\boldsymbol{k}$ = [$k_2$, $k_3$, $k_4$, $k_5$, $k_6$, $k_{15}$, $k_{16}$, $k_{17}$, $k_{18}$, $k_{19}$, $k_{20}$] |
| RM9 | $*\,*\,* + G \xrightarrow{k_1} P + G$<br>$G + P \underset{k_6}{\overset{k_5}{\rightleftharpoons}} GP$<br>$P + GP \xrightarrow{k_{21}} GP + *\,*\,*$ | $r_P = k_1 c_G + k_3 - k_4 c_P - k_5 c_G c_P + k_6 c_{GP} - k_{21} c_P c_{GP}$;<br>$r_G = -k_5 c_G c_P + k_6 c_{GP}$;<br>$r_{GP} = k_5 c_G c_P - k_6 c_{GP}$;<br>$c_{G_{tot}} = c_G + c_{GP}$. | $k_1 c_{G,s} + k_3 - k_4 c_{P,s} - k_{21} c_{P,s} (c_{G_{tot},s} - c_{G,s}) = 0$;<br>$c_{G,s} = \dfrac{k_6 c_{G_{tot},s}}{k_5 c_{P,s} + k_6}$;<br>$c_{GP,s} = c_{G_{tot},s} - c_{G,s}$; | species = [*P,G,GP*];<br>$\boldsymbol{k}$ = [$k_1$, $k_3$, $k_4$, $k_5$, $k_6$, $k_{21}$] |

(vi) The external and internal 'raw-material' concentrations for P and G synthesis are considered constant, and included in the corresponding kinetic rate constants.

(vii) Some of the reaction by-products are not considered into the simulation.

(viii) The cell volume is considered constant, the system being simulated in terms of concentrations. However, to mimic the cell-volume growth, a diluting rate of P concentration was included in the model. The dilution rate is of first-order in P, of rate constant $k_2$, which corresponds to the cell-volume average logarithmic growing rate[24]. In fact, such a cell-content dilution should be considered, in a more complete model, for every species into the cell-system, and it acts as a continuous perturbation of the system. The osmotic pressure into the cell is also considered constant.

(ix) When separate evaluation of regulatory modules is made, one ignores interactions among modules, even if some of the involved species can be shared in common by several metabolic processes.

Continuous perturbations due to changes in external (nutrient) conditions or due to intracellular processes were included through two P-perturbation reactions, of rate constants $k_3$ and $k_4$ (Table 1). Integrating the dynamical ODE model can simulate the effect of an impulse perturbation of a certain species of QSS concentration level, and determine their recovering path. Due to fast reversible G/P 'buffering' reactions, the ODE model requires a stiff integrator (a Matlab-package environment is usually satisfactory).

If one writes the model (1) in a more general vectorial form:

$$\mathrm{d}c/\mathrm{d}t = h(c,k,t); \quad c(t_0) = c_0, \tag{2}$$

the steady-state mass-balances correspond to the set:

$$h(c_s,k) = 0, \tag{3}$$

(index 's' denotes the steady-state value). By solving the nonlinear algebraic set (3), when nominal QSS concentrations are known, one obtains an estimate of the model kinetic parameters $\hat{k}$. Vice-versa, if the rate constants are known, the set (3) can be used to estimate the QSS concentrations of the system when external conditions of continuous perturbations are imposed (the so-called QSS trajectories).

Predictive, regulatory and model flexibility characteristics can help in discriminating plausible reaction schemes. The current study will consider the following comparison aspects:

– model complexity, i.e. number of reactions and species;

– stationary regulatory effectiveness to maintain a P-concentration level within a certain range under unsynchrone / synchrone perturbations;

– predicted recovering trajectory shape and amplitude after a certain dynamic perturbation;

– predicted QSS's sensitivity to stationary perturbations;

– predicted QSS domain limits for variate stationary perturbations;

– predicted QSS's local stability.

## Stationary (step) and dynamic (impulse) perturbations

Cells are such regulated that their components are maintained at relatively invariant steady-state concentrations despite the presence of perturbations (homeostatic regulation). Protein synthesis involves numerous components. However, in RM1-RM9 mechanisms, except for the gene G encoding P, M, and few related components, all others are ignored and assumed to be present at constant concentrations. Perturbations were divided into *step* (stationary) and *impulse* (dynamic) categories accordingly to the shape of perturbation function of time.

### Steady-state regulatory effectiveness

Stationary perturbations are caused by environmental processes that tend to increase or to decrease P from their *nominal* steady-state level $c^*_{P,s}$ until it reaches another steady-state $c_{P,s}$. The system is homeostatically regulated if the difference between the two perturbed and unperturbed states is below an accepted tolerance $\varepsilon > 0$, i.e.:

$$c_{P,s} = c^*_{P,s}(1 \pm \varepsilon), \tag{4}$$

(where $\varepsilon > 0$ is usually set to 0.05–0.1). By including the min/max $c_{P,s}$ values from (4) into the QSS system mass-balance of Table 1, minimum and maximum admissible values for the perturbation variables can be thus determined for all the models, and then used to rank the model stationary regulatory capability.

This analysis has been done by *Sewell* et al.[1] by introducing in the model the following stationary perturbation reactions for P-species:

$$*\,*\,* \xrightarrow{k_3} P \xrightarrow{k_4} *\,*\,*. \tag{5}$$

Thus, for a nominal condition of a protein from the *E. coli* cell (that is $c^*_{P,s}$ = 1000 nmol l$^{-1}$, $c^*_{G_{tot},s}$ =

1 nmol $l^{-1}$, $c^{*}_{G,s}$ = 1/2 n mol $l^{-1}$), *Sewell* et al.[1] derived a stationary regulatory effectiveness index for P, i.e.: $A_{\text{unsync}} = k_{3,\max} \times k_{4,\max}$. As this index is larger the mechanism regulatory effectiveness is better. Both unsynchronized ($k_3 \neq 0, k_4 = 0$; $k_3 = 0, k_4 \neq 0$) and synchronised ($k_3 \neq 0, k_4 \neq 0$) perturbations have been considered. According to this regulatory index, the models have been ranked in the following decreasing effectiveness order:

$$\text{RM4} > \text{RM6} > \text{RM8m} > \{\text{RM3,RM5,RM9}\} > \{\text{RM2,RM7}\} > \text{RM1}. \quad (6)$$

(in the RM8m model, the $k_{16}$ constant was neglected). The system is better regulated when high rates of P-synthesis/decay are present.

Even if stationary conditions are fulfilled, the continuous cell-volume growth will induce a continuous perturbation by diluting the cell-content. Such an effect is accounted in a 'constant' cell-volume model by adding a 'fictive' first-order decay reaction for P, while the external factors influence is included in the catalysed P-synthesis reaction:

$$*\,*\,* \xrightarrow{k_1 G} P \xrightarrow{k_2} *\,*\,*. \quad (7)$$

In general, when a system variable or parameter (let us denote it with $x_j$) reaches another QSS value, the other system components will be influenced and will reach other QSS levels. The 'resistance' (or the 'sensitivity') of each of the species $i$ to such variations can be numerically expressed as their concentration derivative in respect to the perturbation variable, i.e. the so-called time-dependent sensitivity function, $[\partial c_i / \partial x_j](t)$, or the corresponding sensitivity coefficient at QSS $[\partial c_i / \partial x_j]_s$. The way in which these functions are generated is extensively discussed in the Appendix A. The most interesting for our study are the $[\partial c_P / \partial k_3]_s$ and $[\partial c_P / \partial k_4]_s$ which will be further analysed.

## Dynamic regulatory effectiveness

Homeostatic protein regulatory systems can also be dynamically perturbed. Impulse perturbations are instantaneous changes in the concentration of one or more components of the cell-system, that arise in an event, lasting an infinitesimal length of time. After an impulse, stable systems return to their unperturbed nominal state within a certain recovering time. If one divides the perturbation magnitude by the recovering time, the average recovering rate is thus obtained.[2] A system is better dynamically regulated when the recovering time is minim and the recovering rate maxim.

The recovering time can be evaluated directly from simulating the system recovering dynamics by means of the ODE model until a certain species (i.e. P) recovers to their nominal QSS with an accepted tolerance. The recovering time can be defined for every species, and depends on the initial perturbation magnitude and the reached QSS. The same recovering time is also related to the solution of the linearized dynamic model (3) around the QSS. After applying a perturbation, the ODE set solution consists of a steady-state term and a dynamic term[25], i.e. $c(t) = c_{ns} + \tilde{c}(t)$. The dynamic term is function of the Jacobian's system eigenvalues (Appendix C). The system recovering time is proportional to the inverse of the absolute value of the real part of the Jacobian's eigenvalue whose real part is the least negative. For stable systems, the real parts of all such eigenvalues must be negative (possibly some being zero). This approach was used to optimise and rank regulatory mechanisms from a dynamic perspective.[2]

Dynamic indices are not approached explicitly in the present study because the sensitivity analysis vs. perturbations is mainly focussed on the cell QSS-conditions. However, the recovering trajectory shape and amplitudes after certain impulse perturbations have been determined and used to compare the RM1-RM9 mechanisms. Moreover, the below estimation calculations account for dynamic properties of G-P fast binding reactions to fulfil the QSS-constraints. As proved by *Sewell* et al.[1], the regulatory effectiveness is proportional to the slope of the predicted $\{c_G / c_{G_{tot}}$ vs. $\log(c_P)\}$ curve at QSS. Because G is the catalyst for the P-synthesis, when only one reversible reaction of binding P to G is present (leading to GP inactive species), the curve slope is maximum at $c_{G,s} / c_{G_{tot},s} = 1/2$. Thus, equal possibilities exist for G-species to increase or to diminish in order to adjust the P/G ratio, thus balancing a P-perturbation. The fast P reversible binding reactions of this type are considered to fulfil the equilibrium QSS condition:

$$G + P \underset{k_6}{\overset{k_5}{\rightleftarrows}} GP; \qquad K_{GP} = \frac{c_{GP,s}}{c_{P,s} c_{G,s}}. \quad (8)$$

As the model includes more reversible binding reactions to adjust P/G or M/P ratios, the system is faster regulated vs. perturbations. The maximum effectiveness condition $c_{G,s} / c_{G_{tot},s} = 1/2$ allowed *Sewell* et al.[1] to develop QSS-constraints to be used for estimating the model parameters from regulatory objective functions.

Based on a stationary regulatory analysis, *Sewell* et al.[1] elaborated the main principles to be followed when designing a new regulatory mechanism. Thus, the idealistic RM4 model, with a Boolean variation of G, presents the highest regulatory effectiveness and an infinite $\{c_G / c_{G_{tot}}$ vs.

$\log(c_P)\}$ slope, while the RM1 schema without feedback control, presents the lowest regulatory effectiveness and a zero-slope. The models with two P's, or P-dimer binding G, present the same regulatory effectiveness. Those with negative-feedback-controlled cascade are the most probable, when the ultimate product P inhibits the first-level enzyme G. The cascade regulatory structures can amplify the rates of P synthesis and degradation, thereby improving regulation. The feedback control at each level of each cascade can also improve the regulatory capabilities, as much as the equilibrium points are closer to the point when half active catalyst is bound. As remarked, the real regulatory cycles can be even more complex, by including several transcription/translation sequences with feedback at each level, and leading, to more sensitive reaction schemes.

## Rate constants estimation and sensitivity-to-perturbation analysis

### Sensitivity functions and coefficients

One essential characteristic of a dynamic system is the magnitude, the sense and the route followed by the state variables as a response to perturbations. As revealed in the discussion below to (7), the time-dependent sensitivity functions can be evaluated over a transient path. When recovering after a perturbation, the system asymptotically evolves toward their QSS. The sensitivity functions will also evolve toward the so-called sensitivity coefficients of QSS. The sensitivity functions and coefficients can be numerically evaluated according to the rule described in Appendix A. For the current analysis, of more interest are the relative sensitivities at QSS of $c_{P,s}$ vs. the perturbation parameters $k_j$, $j$ = 3,4 (e.g. the P-synthesis and dilution perturbation rate constants, Table 1):

$$S_{k_j}^{r,c_{is}} = [\partial c_i / c_{i,s}]_s \Big/ [\partial k_j / k_j^*]_{k^*} \,. \qquad (9)$$

Integral averages of the relative sensitivity functions over the recovering path allow defining the average relative sensitivities $\overline{S}_{k_j}^{r,c_i}$. Analogously, the integral averages of the relative sensitivity functions can be evaluated vs. the initial perturbation $\overline{\Phi}_{c_{j,0}}^{r,c_i}$. Because the sensitivity-to-perturbation functions depend on the initial perturbation $c_{j,0}$, the QSS, the model structure and parameters, they can be used under certain conditions to compare the regulatory models.

### Estimation of the model parameters based on a sensitivity criteria

The classical way to estimate the parameters of an ODE kinetic model is based on the product/intermediate analysis, prior information, and (filtered, reconcilied) experimental kinetic curves, i.e. on lumped or individual species concentration evolution over the reaction time. A suitable estimation criterion, minimising the residual differences between data and model predictions in terms of output variables, leads to the estimate. Due to the noised data, several estimates can be obtained under similar operating conditions. The identification problem is formulated as a numerical nonlinear least squares problem and is solved by using several techniques: indirect methods (objective function iterative minimisation with repeatedly model evaluation), or direct methods (based on model transformation and approximate problem solution in one step.[26] The estimation objective function is linked with the statistical methods because the observed data are always subjected to experimental errors, and several physico-chemical constraints are imposed to the parameters.

For the developed regulation models, at a cell level, little standard information in the classical sense is available. The construction of kinetic modules accounts rather disparate qualitative/quantitative information from databanks, being linked to describe the cell functions. This is why, the rate constants are derived, rather, by optimising non-conventional estimation criteria, such as: faster recovering rate or smallest recovering time after an 'impulse' / dynamic perturbation; smallest amplitude of the recovering path; smallest sensitivity of the QSS vs. perturbations; maximum QSS stability; imposed periodicity to the reaction cycle; fulfilment of a nominal QSS and some physical constraints; inflexibility to external condition (nutrient's level) changes; imposed succession of metabolic events into the cell, etc. Each of these criteria can lead to a significant different estimate, which offers another perspective for the analysed cell-regulation system. For a real cell simulation it is probable that a simultaneously multi-objective criterion can be used, with weights in accordance with the main cell-functions.

In the present study, to estimate the model rate constants $\boldsymbol{k}$, one started from the same nominal QSS concentrations of *Sewell* et al.[1], i.e. $c_{P,s}^* = 1000$ nmol l$^{-1}$, $c_{G_{tot},s}^* = 1$ nmol l$^{-1}$. To facilitate the further QSS sensitivity analysis, the estimate has to be related with the sensitivity functions of the system vs. perturbations. Thus, a minimum-sensitivity-to--perturbation criterion at nominal QSS was formu-

lated by summing the P-relative sensitivity coefficients in respect to $k_3$ and $k_4$:

$$\hat{k} = \arg\,\mathrm{Min}\; f_{\mathrm{obj}} = (S_{k_3}^{r,c_{\mathrm{P,s}}} + S_{k_4}^{r,c_{\mathrm{P,s}}}),$$

subjected to the QSS constraints:

$$h(c_{\mathrm{s}}, \hat{k}) = 0; \qquad \hat{k} \geq 0; \qquad g(c_{\mathrm{s}}, \hat{k}) = 0. \quad (10)$$

QSS mass-law constraints in (10) are imposed on the stationary point toward the perturbed system evolves. To ensure also a maximum model regulatory effectiveness at the QSS, supplementary constraints (denoted with ***g***, see Table 2) derived by *Sewell* et al.[1] for every model, have been considered, mainly concerning the P-G binding reaction (see discussion on (8)). Bounding the search region in the parameter space ensured the estimated $\hat{k}$ positiveness. These constraints have been included in the estimation objective function as penalty terms with certain weights $w$:

$$f_{\mathrm{obj}}^{*} = f_{\mathrm{obj}} + \sum_j w_j g_j; \quad \text{(adopted } w_{\mathrm{j}} = 10;\ j = 1\text{–}3). \quad (11)$$

Small $k_3$ = 1e-3 nmol L$^{-1}$ min$^{-1}$ and $k_4$ = 1e-6 min$^{-1}$ values have been assigned to derive the sensitivities in respect to these parameters. The ODE set (Appendix A) was numerically integrated from an impulse-perturbed initial state over a large time interval of 1000 min to reach the QSS with a satisfactory tolerance. As minimisation method, the adaptive random search MMA of *Maria*[27] was used, as being very reliable for solving highly constrained multimodal nonlinear problems ($k_1$ and $k_2$ have been kept at the *Sewell* et al.[1] values). The estimated rate constants are presented for every model in Table 3 together with the obtained equality constraint's fulfilment at the solution. The estimate is close but not identical with those obtained by *Yang* et al. [2] by applying a fast recovering rate estimation criterion. The obtained rate constants will be used to complete the model-to-perturbation sensitivity analysis.

## Model ranking based on the QSS sensitivity characteristics

To compare the QSS sensitivity-to-perturbation coefficients predicted by various models, the same 10 % initial impulse-perturbation in $c_{\mathrm{P,s}}$ and $c_{\mathrm{G,s}}$

Table 2 – *QSS mass-balance and maximum regulatory effectiveness constraints*[1] *imposed to the parameter estimation criterion*

| Model | Constraint $g_1 = 0$ | Constraints $g_2 = 0$; $g_3 = 0$ |
|---|---|---|
| RM1 | $g_1 = c_{\mathrm{G_{tot}}} - \Sigma(c_{\mathrm{G-intermediates}}) = 0$ | – |
| RM2 | Ibidem | $g_2 = \frac{k_5}{k_6} - \frac{1}{c_{\mathrm{P,s}}}$ |
| RM3 | Ibidem | $g_2 = \frac{c_{\mathrm{G,s}}}{c_{\mathrm{G_{tot},s}}} - \frac{1}{1 + \frac{k_5}{k_6} c_{\mathrm{P,s}} + \frac{k_5 k_7}{k_6 k_8} c_{\mathrm{P,s}}^2}$; $g_3 = \frac{c_{\mathrm{G,s}}}{c_{\mathrm{G_{tot},s}}} - \frac{1}{2}$ |
| RM5 | Ibidem | $g_2 = \frac{c_{\mathrm{G,s}}}{c_{\mathrm{G_{tot},s}}} - \frac{1}{1 + \frac{k_9 k_{11}}{k_{10} k_{12}} c_{\mathrm{P,s}}^2}$; $g_3 = \frac{c_{\mathrm{G,s}}}{c_{\mathrm{G_{tot},s}}} - \frac{1}{2}$ |
| RM6 | Ibidem | $g_2 = \frac{c_{\mathrm{G,s}}}{c_{\mathrm{G_{tot},s}}} - \frac{1}{1 + \frac{k_9 k_{11}}{k_{10} k_{12}} c_{\mathrm{P,s}}^2 + \left(\frac{k_9}{k_{10}}\right)^2 \frac{k_{11} k_{13}}{k_{12} k_{14}} c_{\mathrm{P,s}}^4}$; $g_3 = \frac{c_{\mathrm{G,s}}}{c_{\mathrm{G_{tot},s}}} - \frac{1}{2}$ |
| RM7 | Ibidem | $g_2 = \frac{2 k_2 k_{16}}{k_{15} k_{17}} - \frac{1}{c_{\mathrm{P,s}}}$ |
| RM8 | Ibidem | $g_2 = \sqrt{\frac{2 k_2}{k_{\mathrm{c}}}} - \frac{1}{c_{\mathrm{P,s}}}$; $k_{\mathrm{c}} = \frac{k_{15} k_{17} (k_{19} + k_{20})}{k_{18} k_{20}}$ |
| RM9 | Ibidem | $g_2 = \frac{1}{c_{\mathrm{P,s}}} - \frac{k_{21}}{k_1}$; $g_3 = \frac{1}{c_{\mathrm{P,s}}} - \frac{k_5}{k_6}$ |

Table 3 – *Estimated rate constants for RM1-RM9 models by using a sensitivity criterion under QSS constraints of Table 2 ($\Delta g_1$ and $\Delta g_2$ denote the error in fulfilling the constraints by the solution; $\Delta g_3 = 0$; the perturbation constants are null, i.e. $k_3 = k_4 = 0$)*

| $k_j$ | units | RM1 | RM2 | RM3 | RM5 | RM6 | RM7 | RM8 | RM9 |
|---|---|---|---|---|---|---|---|---|---|
| $k_{1(*)}$ | $min^{-1}$ | 20 | 20 | 20 | 20 | 20 | - | - | 20 |
| $k_{2(*)}$ | $min^{-1}$ | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| $k_5$ | L $nmol^{-1}$ $min^{-1}$ | - | 1.9e-5 | 2.89393e-5 | - | - | 6.80000e-3 | 1.89422e-1 | 2.90e-5 |
| $k_6$ | $min^{-1}$ | - | 1.9e-2 | 7.70002e+2 | - | - | 6.80000 | 1.89422e+2 | 2.90e-2 |
| $k_7$ | L $nmol^{-1}$ $min^{-1}$ | - | - | 8.77976e+1 | - | - | - | - | - |
| $k_8$ | $min^{-1}$ | - | - | 3.29985 | - | - | - | - | - |
| $k_9$ | L $nmol^{-1}$ $min^{-1}$ | - | - | - | 1.20036e-9 | 1.80944e-1 | - | - | - |
| $k_{10}$ | $min^{-1}$ | - | - | - | 4.10e-3 | 8.15318e+6 | - | - | - |
| $k_{11}$ | L $nmol^{-1}$ $min^{-1}$ | - | - | - | 9.69000e-2 | 9.99072 | - | - | - |
| $k_{12}$ | $min^{-1}$ | - | - | - | 2.85162e-2 | 2.21721e-1 | - | - | - |
| $k_{13}$ | L $nmol^{-1}$ $min^{-1}$ | - | - | - | - | 4.18779e+1 | - | - | - |
| $k_{14}$ | $min^{-1}$ | - | - | - | - | 1.26076e+7 | - | - | - |
| $k_{15}$ | $min^{-1}$ | - | - | - | - | - | 9.99981e-1 | 7.25565e+1 | - |
| $k_{16}$ | $min^{-1}$ | - | - | - | - | - | 3.74980e-2 | 6.10e-3 | - |
| $k_{17}$ | $min^{-1}$ | - | - | - | - | - | 7.49975e-1 | 1.21278e+3 | - |
| $k_{18}$ | L $nmol^{-1}$ $min^{-1}$ | - | - | - | - | - | - | 5.91193e+1 | - |
| $k_{19}$ | $min^{-1}$ | - | - | - | - | - | - | 8.16933e+1 | - |
| $k_{20}$ | $min^{-1}$ | - | - | - | - | - | - | 6.56863 | - |
| $k_{21}$ | $min^{-1}$ | - | - | - | - | - | - | - | 2.00e-2 |
| $\Delta g_1$ | nmol $L^{-1}$ | 0 | 0 | 5.08e-11 | 2.56e-3 | -7.66e-6 | -8.87e-13 | 5.38e-15 | 0 |
| $\Delta g_2$ | - | - | 0 | 3.82e-18 | 1.28e-3 | -3.83e-6 | 8.69e-16 | -6.50e-19 | 0 |

(*) average rate constants for P-synthesis and dilution rates in *E. coli* adopted by *Sewell* et al.[1]

have been applied to all models (Table 4). The sensitivities in respect to $k_3$ and $k_4$ were derived by considering stationary perturbations of $k_3$ = 1e-1 (nmol $L^{-1}$ $min^{-1}$) and $k_4$ = 1e-4 ($min^{-1}$). The obtained relative sensitivity coefficients for species P of interest are presented in Table 4. Models RM1 and RM3 were not included in this comparison, because of their dynamic evolution to another QSS with different characteristics. As the species P sensitivity coefficients vs. $k_3$ and $k_4$ are smaller, the $c_{P,s}$ is insensitive to the synthesis and decay stationary perturbations. The predicted QSS sensitivities rank the models in the following decreasing order:

$$\{RM2,RM7\} > \{RM5,RM6,RM9\} > RM8. \quad (12)$$

The P average sensitivity coefficients in respect to the perturbed initial states depend on the each checked mechanism, perturbation size, and recovering path. As these coefficients are smaller, the model is more insensitive to initial state. The sum of average sensitivity coefficients of P vs. $c_{P,0}$ and $c_{G,0}$ from Table 4 confirm the previous model ranking:

$$RM9 > \{RM2,RM5,RM7\} > RM6 > RM8. \quad (13)$$

Finally, it is to observe, that due to the stochastic character of perturbations, it can be more realis-

Table 4 – *Species average relative sensitivities* $\overline{\Phi}^{r,c_i}_{c_{j,0}}$ *in respect to the initial perturbed state* $c_{j,0}$ *(determined over 1000 min recovering interval*). Species P relative sensitivities* $S^{r,c_{is}}_{k_j}$ *at QSS in respect to the perturbation parameters* $k_3$ *and* $k_4$.

| Reference perturbation ($c_{i,0}$ or $k$ perturbation) | Model | | | | | |
|---|---|---|---|---|---|---|
| | RM2 | RM5 | RM6 | RM7 | RM8 | RM9 |
| $c_{P,0}$ | 7.11e-2 | 4.24e-2 | 5.59e-2 | 7.00e-2 | 4.76e-2 | 5.48e-2 |
| $c_{G,0}$ | 6.34e+2 | 6.96e+2 | 4.69e+2 | 6.05e+2 | 3.79e+2 | 8.47e+0 |
| $c_{GP,0}$ | 6.01e+2 | - | - | 6.05e+2 | 3.79e+2 | -2.49e+1 |
| $c_{GPP,0}$ | - | 4.56e+2 | 4.67e+2 | - | - | - |
| $c_{GPPPP,0}$ | - | - | 4.67e+2 | - | - | - |
| $c_{M,0}$ | - | - | - | 1.28 | 9.32e-3 | - |
| $c_{MP,0}$ | - | - | - | - | 4.52e-2 | - |
| $c_{PP,0}$ | - | -2.24e+2 | 8.96e-2 | - | - | - |
| $k_3$ | 6.62e-3 | 4.95e-3 | 4.97e-3 | 6.62e-3 | 3.98e-3 | 4.97e-3 |
| $k_4$ | -6.62e-3 | -4.77e-3 | -4.97e-3 | -6.62e-3 | -3.98e-3 | -4.97e-3 |

(*) Initial impulse perturbation $[c_{P,0}, c_{G,0}, c_{GP,0}, c_{GPP,0}, c_{GPPPP,0}]$ was set to [1100,0.6,0.4,0,0](nmol $L^{-1}$) for models RM2, RM7–9, [1100,0.6,0.2,0.2,0] (nmol $L^{-1}$) for model RM3, [1100,0.6,0,0.4,0](nmol $L^{-1}$) for model RM5, and [1100,0.6,0,0.2,0.2](nmol $L^{-1}$) for model RM6; $c_{G_{tot}} = 1$ nmol $L^{-1}$.

tic to consider a random distribution of the stationary perturbations in terms of parameters, but also in initial concentrations, for instance of normal-type with the variance *V(*)*. In this case, the combined random stationary perturbation effect on $c_{P,s}$ variance can be approximated with the error-propagation-formula of Park & Himmelblau:[28]

$$V_P = \sum_{i=1}^{n_s} [S^{c_{Ps}}_{c_{i,0}}]^2 V_{c_{i,0}} + \sum_{i=1}^{n_p} [S^{c_{Ps}}_{k_i}]^2_{k^*} V_{k^*_j} . \quad (14)$$

The previous derived sensitivity functions of the cell-system states (e.g. species concentrations) vs. metabolic-cell perturbations ignore interactions with other cell-regulatory modules. More complex models, including linked modules, can offer the possibility to account for a more complete representation of component sensitivities.

### Recovering trajectory shapes

The effect of an impulse-perturbation can be analysed from various point of views: system recovering rate, recovering time, and recovering path amplitude and shape. In general, if a system presents multiple QSS (not the case here), simulation of recovering path from various initial perturbed states can be used to construct the 'attraction' domain around each QSS. The way and rate of which the cell-system components recover depends on the model regulatory effectiveness, and have been used to compare and rank various mechanisms.[2] To be consistent, such an analysis has to use estimated rate constants from dynamic criteria.

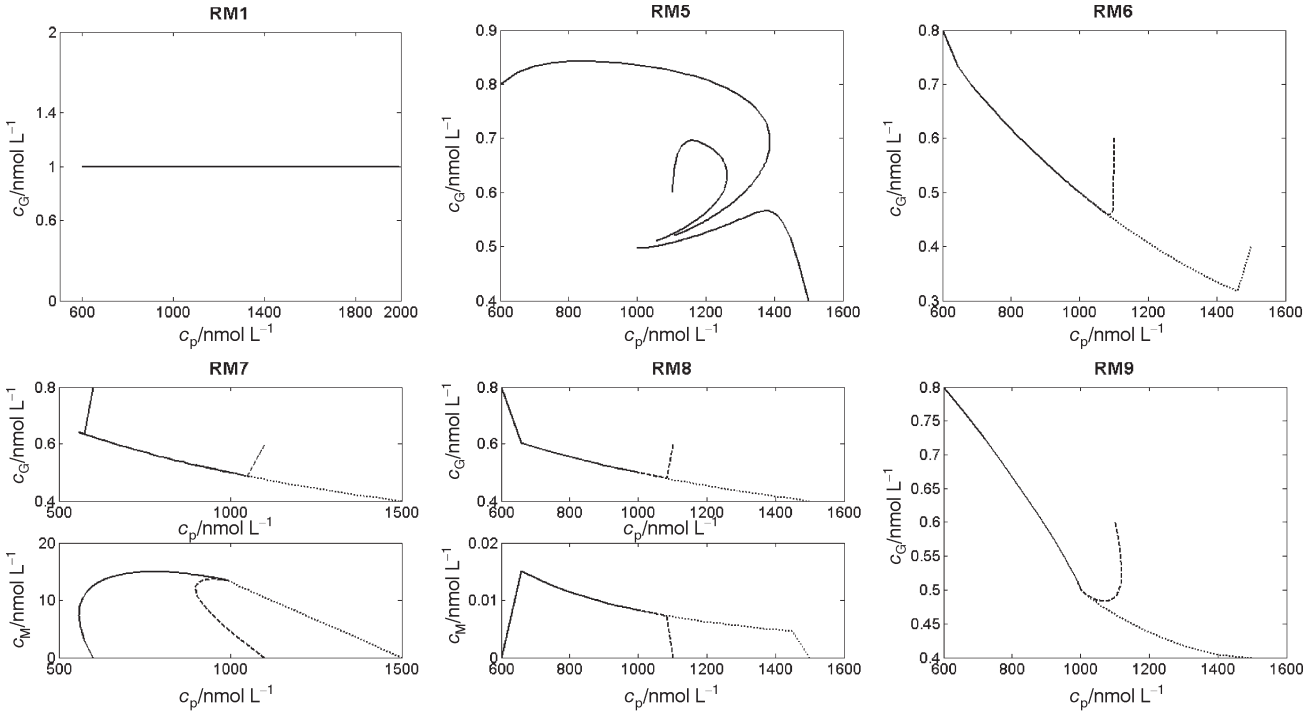
Rapid and straight recovering paths, in terms of species concentrations, can reveal a more effective regulatory model. Figure 1 displays such recovering trajectories for some state-impulse-perturbations. While RM1 and RM4 models predict a Boolean behaviour for P-G species, the other models predict nonlinear trajectories toward QSS. At a first glance, following the recovering shapes, the models can be grouped as follows:

{RM1, RM4}, {RM2, RM9}, {RM3},
{RM5}, {RM6, RM7, RM8}. (15)

Models RM6 and RM8 seem to be more effective in cell regulation, by predicting straighter recovering trajectories, which avoid large P transient states. Such a conclusion is confirmed by the model classification of *Yang* et al.[2] from a more detailed dynamic regulatory analysis.

## QSS admissible region of stability under stationary perturbations

If stationary 'step'-like perturbations occur into the cell, the stationary levels of species concentra-




F i g . 1 – *Recovering paths for P, G, and M species after various 'impulse' perturbations in QSS concentrations, predicted by the models RM1, RM5-RM9. Initial perturbations* $[c_{P,0}, c_{G,0}, c_{GP,0}, c_{GPP,0}, c_{GPPP,0}]$ *of: [1100,0.6,0.4,0,0], [600,0.8,0.2,0,0], [1500,0.4,0.6,0,0] (nmol L*$^{-1}$*) for models RM7–9; [1100,1,0,0,0], [600,1,0,0,0], [1500,1,0,0,0] (nmol L*$^{-1}$*) for models RM1; [1100,0.6,0,0.4,0], [600,0.8,0,0.2,0], [1500,0.4,0,0.6,0] (nmol L*$^{-1}$*) for model RM5; [1100,0.6,0,0.2,0.2], [600,0.8,0,0.1,0.1], [1500,0.4,0,0.3,0.3] (nmol L*$^{-1}$*) for model RM6. (The other species initial concentrations were set to zero).*

tions change also, evolving toward another QSS. Such constant perturbations are generated by the continuous cell-volume growth, but also by changes in external nutrient levels. In the present paragraph, one investigates the variations in the QSS generated by stationary synchronised / unsynchronised perturbations in $k_3$ and $k_4$ parameters, e.g. the P-synthesis and dilution perturbation rate constants. The QSS-trajectories for various $k_3$ and $k_4$ have been traced by means of so-called 'continuation algorithm' (Appendix B). Such a rule highlights the QSS-domain limits tolerated by the cell-system and the QSS-sensitivity to stationary perturbations. The analysis is then extended with characterising the QSS for stability. Finally, a model comparison according to these predicted characteristics is presented.

## Double continuation method to analyse the system QSS

Starting from the model QSS mass balance (3) and applying stationary perturbations $k_3$ and $k_4$ to the system, different QSS-s can be obtained. If one imposes fulfilment of $c_{G_{tot},s} = 1$ nmol L$^{-1}$ constraint, and a perturbation range of $k_3 \in [0,4]$ (nmol L$^{-1}$ min$^{-1}$) and $k_4 \in [0,4e\text{-}3]$ (min$^{-1}$), repeated solutions of mass balance (3) lead to QSS-trajectory plots for all the model species (i.e. the double continuation plots). In Figure 2 are presented the $c_{P,s}$ (nmol L$^{-1}$) trajectories, in logarithmic plots for the models RM6 and RM8. The displayed admissible area corresponds to a variation of $c_{P,s}$ within a ± 5 % $c_{P,s}$ range, as tolerated by the cell-system, without disturbing their metabolic process.

While no multiple QSS-s, bifurcation, or turning point exist, the predicted $c_{P,s}$ value depends on the model structure, parameters, and the stationary perturbation size. The logarithmic plots lead to evaluate the QSS-surface gradients (displayed in Figure 2, the lower panels). The gradients represent the $c_{P,s}$ sensitivities in respect to $k_3$ and $k_4$. It is to remark that, sensitivity vs. $k_4$ is ca. three times higher than those in respect to $k_3$. Such a result is expected because the P-synthesis perturbation reaction (with rate constant $k_3$) is of zero-order, while the P-consumption perturbation reaction (with rate constant $k_4$) is of first-order in P (with concentrations around 1000 nmol L$^{-1}$). The logarithmic QSS-surface gradients correspond in fact to the logarithmic relative sensitivity QSS-coefficients (A2) in Appendix A.

By considering the same perturbation range and nominal conditions of $[c^*_{P,s}, c^*_{G,s}] = [1000,0.5]$ (nmol L$^{-1}$), the QSS-surface predicted by various

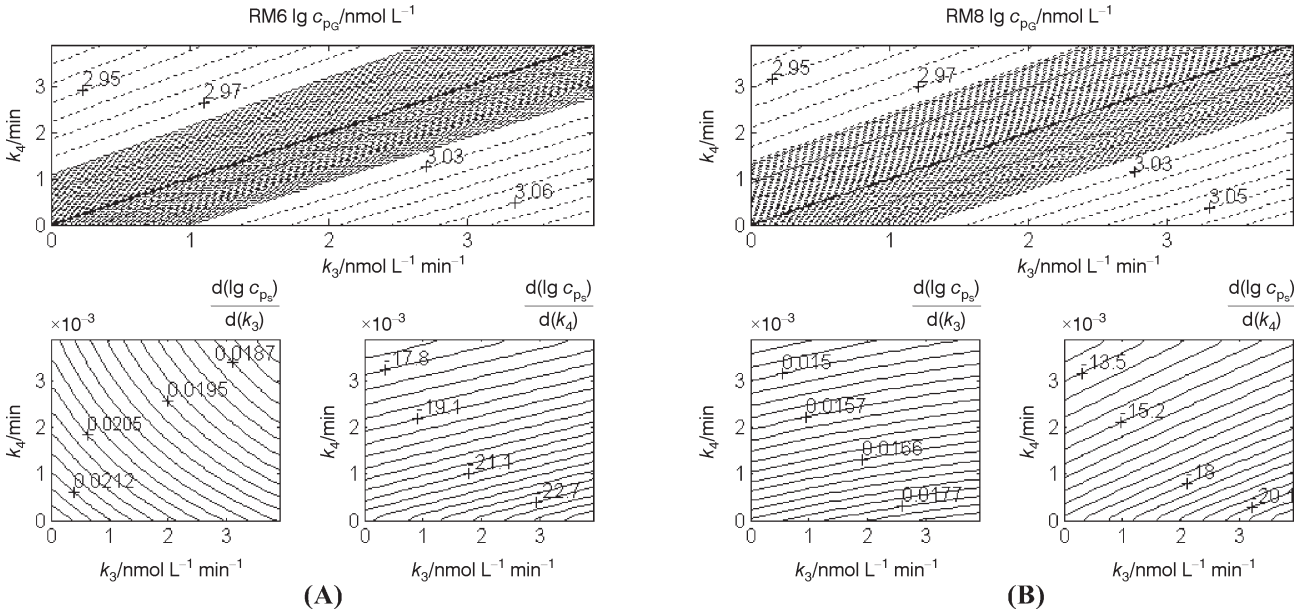



Fig. 2 – *[upper panel] Decimal logarithmic plots for $c_{P,s}$ (in nmol L$^{-1}$) at QSS predicted by models RM6 (A), and RM8 (B). The checked range is $k_3 \in [0,4]$(nmol L$^{-1}$ min$^{-1}$), $k_4 \in [0,4e\text{-}3]$(min$^{-1}$). Within the marked perturbation area, the $c_{P,s}$ (nmol L$^{-1}$) can be maintained within a ± 5 % $c_{P,s}$ admissible range.*
*[lower panel] The response surface gradients in decimal logarithmic plots, e.g. $\partial lg(c_{P,s})/\partial k_3$, and $\partial lg(c_{P,s})/\partial k_4$ (the reference origin of $k_3 = 0$, $k_4 = 0$ corresponds to $lg(c_{P,s}) = 3$).*

models can be directly compared. As the surface gradients are smaller the admissible area is larger and the system is better regulated vs. stationary perturbations. For instance, Figure 2 reveals that QSS-surface gradients for RM8 model are smaller (and the admissible area larger) than those predicted by the RM6 model. The predicted QSS-surface gradients $\partial lg(c_{P,s})/\partial k_3$ and $\partial lg(c_{P,s})/\partial k_4$ on the same perturbation range, rank the models with an increasing 'resistance' to stationary perturbations as follows:

$$\begin{aligned} &\text{RM1} < \{\text{RM2, RM7}\} \\ &< \{\text{RM3, RM5, RM6, RM9}\} < \text{RM8.} \end{aligned} \quad (16)$$

This model classification matches with (6), predicted by *Sewell* et al.[1] when using another stationary regulatory index. The model RM8 predicts the largest admissible QSS-region from all the models.

## Stability checks of QSS

By analysing the protein QSS-levels under various stationary conditions, an important question arises concerning their stability. The analysis is exemplified for the nominal $[c^*_{P,s}, c^*_{G,s}]$ (with $k_3 = k_4 = 0$), but other QSS-s can be similarly approached.

The analysis starts with linearizing the nonlinear QSS-set (3) near the stationary solution $c_s$ (see Appendix C). The evaluated system Jacobian's matrix $\boldsymbol{h}_z$ and their eigenvalues are then analysed. If the real parts are negative for all the eigenvalues (eventually some being zero), it results that the QSS is a stable point. Moreover, one can appreciate the stability strength from how negative these values are. The calculus is repeated for all the models. In Figure 3 (left panels) are displayed, for model RM5 and RM6 cases, the real parts of the Jacobian's eigenvalues corresponding to the P and G mass balance equations, traced over a stabilisation time-interval of 50 min. Model RM6 predicts much smaller $\boldsymbol{h}_z$'s eigenvalues of real parts (1–5 order of magnitude smaller) than RM5 model, while the corresponding stabilisation interval is much shorter. As a direct conclusion, model RM6 predicts stronger stable QSS with a higher recovering potential.

All the investigated models predict non-zero eigenvalues, thus no remnant periodic QSS solution exists. The numerical rule based on the monodromy matrix $\boldsymbol{M}$ eigenvalues (see Appendix C) can check for stability of the periodic solutions. Exception is made by the model RM2 (not displayed here) with a null real part of the $\boldsymbol{h}_z$'s eigenvalue, but not leading to periodic instability due to the less than one $\boldsymbol{M}$'s eigenvalue modulus. For all the models, the predicted $\boldsymbol{M}$ eigenvalue's modulus is less than one. Then, the nominal QSS presents any periodic instability (or Hopf bifurcation points), and the recovering trajectories will converge to the stable point.

Evaluation of $\boldsymbol{M}$ matrix's eigenvalues over a time-interval of 50–100 min, allows to evaluate the stability of a QSS against periodic oscillations. As the eigenvalue's modulus is smaller and reached

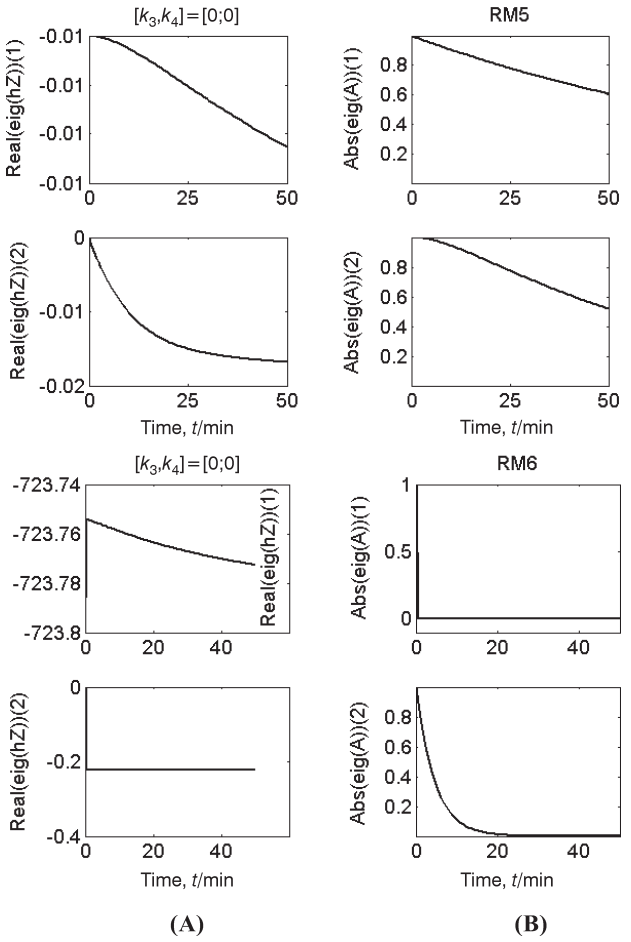



Fig. 3 – *Stability characteristics of QSS predicted by models RM5 (A) and RM6 (B) for* $[c^*_{P,s},c^*_{G,s}] = [1000,0.5]$ *(nmol* $L^{-1}$*), and* $k_3 = k_4 = 0$.

*[A,B-left side] Evolution of the real parts of the* $\boldsymbol{h}_z$ *matrix eigenvalues (corresponding to P and G species).*

*[A,B-right side] Evolution of the eigenvalue modulus of the monodromy matrix* $\boldsymbol{M}$ *(corresponding to P and G species).*

faster, the QSS is more stable. In Figure 3 (right panels) are displayed the eigenvalue's modulus for models RM5 and RM6. The model RM6 predicts a stronger stable QSS point due to a smaller $\boldsymbol{M}$'s eigenvalue modulus, also reached faster.

By comparing the predicted 'degree of stability' for the same nominal QSS, the models can be ranked in the following increasing order:

$$\{RM2,RM9\} < \{RM3,RM5,RM7\} < \{RM6,RM8\}. \quad (17)$$

## Conclusions

When designing a regulatory module for a certain species during a cell-cycle, it is difficult to include all the process complexity in a limited number of reactions. The construction is rather based on disparate qualitative/quantitative information from databanks describing a certain cell function. In the absence of kinetic data, the model parameters are estimated from non-conventional estimation criteria. Each criterion can lead to a significant different estimate, offering a certain perspective on the complexity of the studied cell-system. Simultaneously considering several identification criteria with various weights would lead to more detailed conclusions, with the expense of a considerable computation effort.

The use of simplified kinetic models to mimic proteinic regulatory functions allows a quick evaluation of mechanism effectiveness and highlights contributions of parts of the kinetic schema. When evaluating a regulatory mechanism, several aspects have to be considered: model complexity, rate constant estimability, dynamic and stationary regulatory capability, predicted QSS's stability, QSS sensitivity to perturbations, recovering path after perturbations, QSS multiplicity, QSS-vs.-perturbation surface limits.

The present study completes the *Sewell* et al.[1] and *Yang* et al.[2] stationary and dynamic analyses of variate regulatory mechanisms with including supplementary characterisation aspects. By formulating a minimum-sensitivity-to-perturbation estimation criterion for a nominal QSS, rate constants were obtained and a complex sensitivity-stability analysis was performed. This analysis pointed out differences in regulatory effectiveness, as follows:

(i) model groups presenting similar recovering trajectory shapes after a perturbation:

$$\{RM1,RM4\},\ \{RM2,RM9\},\ \{RM3\},\ \{RM5\},\ \{RM6,RM7,RM8\};$$

(ii) decreasing stationary regulatory effectiveness[1]:

$$RM4 > RM6 > RM8m > \{RM3,RM5,RM9\} > \{RM2,RM7\} > RM1$$

(iii) decreasing QSS sensitivity to perturbation reactions:

$$\{RM2,RM7\} > \{RM5,RM6,RM9\} > RM8$$

(iv) decreasing QSS sensitivity to perturbed initial P/G states:

$$RM9 > \{RM2,RM5,RM7\} > RM6 > RM8$$

(v) [QSS-perturbation] surface gradients / increasing 'resistance' to perturbations:

$$RM1 < \{RM2, RM7\} < \{RM3, RM5, RM6, RM9\} < RM8$$

(vi) QSS decreasing strength of stability:

{RM6, RM8} > {RM3, RM5, RM7} > {RM2, RM9}

As a conclusion, models RM8 and RM6 proved the best low sensitivity-to-perturbations and best stationary regulatory characteristics for all the indices (if we do not include the 'artificial' Boolean model RM4 in this discussion). *Yang* et al.[2] have found the same models very effective for dynamic regulation of the protein levels.

Based on these conclusions, it was possible to highlight the contribution of certain reactions and species to the regulation mechanism and to elaborate general principles for designing new protein regulatory models: (i) use rapid rates for P synthesis and degradation, even if energetically is costly; (ii) use a cascade reaction structure to amplify the synthesis rate of P; (iii) apply feedback relationships (i.e. 'buffer' G/P reaction type) wherever is possible at each level of each cascade; (iv) allow multiple copies of P (or a related transcription factor) to bind sequentially, thus quickly controlling the P/G ratio.

However, a more complete multi-objective analysis, including dynamic recovering effectiveness with a more detailed variable cell-volume model, and possible linked-modules, can confirm such conclusions. A complete analysis of the model regulatory effectiveness lead to adopting the most plausible and simple kinetic scheme to be used when an extended regulation chain is designed. Because the module's inter-connections are expected to be slower than the inner-module reactions, the separated estimated parameters for each module are expected to not vary much when modules are linked. This suggests that such a modular design analysis may be a feasible route to be followed.

*ACKNOWLEDGEMENT*

*The author is grateful to Dr. Paul Lindahl from Texas A&M University (College Station, TX) for his constructive comments in the final stage of the paper.*

## Nomenclature

$A$; $\boldsymbol{A}$ – area; matrices defined in eqn. (A4) or (C3)
$\boldsymbol{B}$ – matrix defined in eqn. (A4)
$\boldsymbol{c}$ – species concentration [$n_s$] vector
$\boldsymbol{f}_{obj}$ – objective function
$\boldsymbol{g}$ – equality constraint functions (Table 2)
$\boldsymbol{h}$ – kinetic model QSS mass balance set
$\boldsymbol{h}_z$ – $\boldsymbol{h}$-matrix derivatives vs. vector $\boldsymbol{C}$ (e.g. the Jacobian matrix)
$\boldsymbol{I}$ – identity matrix
$\boldsymbol{k}$ – (reaction rate) kinetic constants
$\boldsymbol{M}$ – monodromy matrix $\boldsymbol{M}$ (Appendix C)
$n_p$ – number of kinetic parameters
$n_r$ – number of elementary reactions
$n_s$ – number of species
$\boldsymbol{N}$ – augmented mass balance set (Appendix B)
$\boldsymbol{r}$ – reaction rate vector
$s$ – supplementary parameter defined in eqn. (B3)
$S, S', \overline{S}$ – absolute, relative, and average sensitivity functions vs. parameters (Appendix A)
$t$ – time
$T$ – solution period (Appendix C)
$V(*)$ – variance of (*)
$\boldsymbol{w}$ – weights for penalty function during estimation

## Greeks

$\beta$ – partial orders of direct reactions; $\boldsymbol{h}_z$ matrix's eigenvalues
$\varepsilon$ – fractional coefficient defining the protein perturbation magnitude
$\Phi, \Phi^r, \overline{\Phi}$ – absolute, relative, and average sensitivity functions vs. initial conditions (Appendix A)
$\gamma$ – matrix of partial orders of reverse reactions
$\nu$ – stoichiometric coefficient [$n_s \times n_r$] matrix

## Index

o – initial
s – (quasi-)steady-state
tot – total

## Abbreviations

*det(*)* – determinant of the matrix (*)
*rank(*)* – rank of the matrix (*)
*real(*)* – real part of (*)
*trace(*)* – trace of the matrix (*)

## References

1. *Sewell, C., Morgan, J., Lindahl, P.,* J. theor. Biol. **215** (2002) 151.
2. *Yang, Q., Lindahl, P., Morgan, J.,* J. theor. Biol. **215** (2002) (accepted for publishing).
3. *Bower, J. M., Bolouri, H.,* Computational Modeling of Genetic and Biochemical Networks, MIT Press, Cambridge, 2001.
4. *McAdams, H. H., Arkin, A.,* Trends in Genetics **15** (1999) 65.
5. *Axe, D. D., Bailey, J. E.,* Biotech. & Bioeng. **43** (1994) 242.
6. *Tsuchiya, M., Ross, J.,* J. Phys. Chem. **105** (2001) 4052.
7. *Shimizu, T. S., Bray, D.,* Computational Cell Biology – The Stochastic Approach, MIT Press Math6X9, Cambridge, 2002.

8. *Ballet, P., Zemirline, A., Marce, L.,* Course on Cellular Automata, Reaction-Diffusion and Multiagents Systems for Artificial Cell Modeling, Departement d'Informatique, Universite de Bretagne Occidentale (France), 2001. http://www.lami.univ-evry.fr/epigenese/Ateliers/Organisation/PascalBallet-coursHautNiveauBrest1.htm.
9. *Tomita, M.,* Trends in Biotechnology **19** (2001) 205.
10. *Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J. C., Hutchinson III, C. A.,* Bioinformatics **15** (1999) 72.
11. *Kinoshita, A., Nakayama, Y., Tomita, M.,* 2-nd Int. Conf. on Systems Biology, California Institute of Technology, Nov. 4–7, 2001.
12. *NRCAM (The National Resource for Cell Analysis and Modeling),* Virtual Cell Modeling and Simulation Framework, The University of Connecticut Healt Center, Farmington, 2002. http://www.nrcam.uchc.edu/
13. *Bartol Jr., T. M., Stiles, J.R.,* M-Cell: A general Monte Carlo Simulator of Cellular Microphysiology, Computational Neurobiology Lab., The Salk Institute for Biological Studies, La Jolla, California, 2002. http://www.mcell.cnl.salk.edu/
14. *Ichikawa, K.,* **A-Cell**: A Platform for the Construction of Biochemical Reaction and Electrical Equivalent Circuit Models in a Biological Cell and a Neuron, Fuji Xerox Co. Ltd., Kanagawa (Japan), 2000. http://www.fujixerox.co.jp/crc/cng/A-Cell/
15. *CRGM (Centro di Ricerca in Genetica Molecolare),* Annual Report, Istituto di Istologia ed Embriologia Generale, University of Bologna (Italy), 2002. http://apollo11.isto.unibo.it/Cells.htm
16. *Hucka, M. Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., Cuellar, A. A., Dronov, S., Gilles, E. D., Ginkel, M., Gor, V., Goryanin, I. I., Hedley, W. J., Hodgman, T. C., Hofmeyr, J. H., Hunter, P. J., Juty, N. S., Kasberger, J. L., Kremling, A., Kummer, U., Le Novere, N., Loew, L. M., Lucio, D., Mendes, P., Mjolsness, E. D., Nakayama, Y., Nelson, M. R., Nielsen, P. F., Sakurada, T., Schaff, J. C., Shapiro, B. E., Shimizu, T. S., Spence, H. D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J., Wang, J.,*The Systems Biology Markup Language (SBML): A Medium for Representation and Exchange of Biochemical Network Models, Caltech ERATO Kitano Report, System Biology Workbench Development Group, California Institute of Technology, Pasadena, 2002. http://xml.coverpages.org/sbml.html
17. *Chen, T., He, H. L., Church, G. M*., Pacific Symposium of Biocomputing, 1999, p. 29–40.
18. *Somogyi, R., Sniegoski, C.A.,* Complexity **1** (1996) 45.
19. *Akutsu, T., Kuhara, S., Maruyama, O., Miyano, S.,* GIW98 The Ninths Workshop on Genome Informatics, Tokyo (Japan), Dec. 10–11, 1998.
20. *McAdams, H. H., Arkin, A.,* Proc. Natl. Acad. Sci. USA **94** (1997) 814.
21. *Gillespie, D. T.,* Jl. Physical Chemistry **81** (1977) 2340.
22. *Gillespie, D. T., Mangel, M.,* Jl. Chemical Physics **75** (1981) 704.
23. *Neidhardt, F. C., Savageau, M. A.,* Regulation Beyond the Operon, in Neidhardt, F.C.(Ed.), Escherichia coli and Salmonella: Cellular and Molecular Biology, Vol. 1, Ch. 84, ASM Press, Washington D.C., 1996.
24. *Maria, G., Morgan, J., Lindahl, P.,* J. theor. Biol. **215** (2002) (submitted for publishing).
25. *Ingham, J., Dunn, I. J., Heinzle, E., Prenosil, J. E.,* Chemical Engineering Dynamics: Modelling with PC Simulation, VCH, Weinheim, 1994.
26. *Maria, G., Rippin, D.W.T.,* Computers & Chemical Engineering **21** (1997) 1169.
27. *Maria, G.,* FOCAPO98 Int. Conf. on Foundations of Computer Aided Process Operations*,* Snowbird (US), July 5, 1998, p. 351–359.
28. *Park, S. W., Himmelblau, D. M.,* AIChE Jl. **26** (1980) 168.
29. *Tomovic, R.,* Sensitivity analysis of dynamic systems, McGraw-Hill, New York, 1963.
30. *Vajda, S., Rabitz, H., Walter, E., Lecourtier, Y.,* Chem. Eng. Commun*.* **83** (1989) 191.
31. *Froment, G. F., Hosten, L. H.,* Catalytic Kinetics: Modelling, in Anderson, L. and Boudart, M.(Eds.), Catalysis: Science and Technology, Springer-Verlag, Berlin, 1981.
32. *Keller, H. B.,* Numerical Solution of Bifurcation and Nonlinear Eigenvalue Problems, in Rabinowitz, P.H.(Ed.), Applications of bifurcation theory, Academic Press, New York, 1977.
33. *Kubicek, M., Marek, M.,* Computational methods in bifurcation theory and dissipative structures, Springer-Verlag, Berlin, 1983.
34. *Seydel, R., Hlavacek, V.,* Chem. Eng. Sci. **42** (1987) 1281.
35. *Rheinboldt, W. C., Burkardt, J. V.,* ACM TOMS **9** (1983) 215.
36. *Rheinboldt, W. C.,* SIAM Jl. Numer. Anal. **17** (1980) 221.
37. *Juncu, G., Bildea, S.,* Computers & Chemical Engineering **16** (1992) 1059.
38. *DiBiasio, D.,* Introduction to the Control of Biological Reactors, in Shuler, M.L.(Ed.), Chemical engineering problems in biotechnology*,* AIChE Series, Vol.1, New York, 1989.
39. *Holodniok, K., Kubicek, M.,* J. Comput. Physics **55** (1984) 254.
40. *Arnol'd, V. I.,* Geometrical Methods in the Theory of Ordinary Differential Equations, Springer-Verlag, Berlin, 1983.

# Appendix A

## Sensitivity functions

For a kinetic system modelled by means of the differential set (2), the *direct* sensitivity analysis determines the influence of small perturbations of the initial conditions ($c_0$) and parameters ($\boldsymbol{k}$) on the system concentrations ($\boldsymbol{c}$) over a certain time interval, i.e. on a concentration trajectory converging to a QSS[29]. Vice-versa, the *inverse* sensitivity analysis determines the perturbation magnitude of the initial conditions and parameters, which is responsible for a certain perturbation on the system states. Both analyses are based on the sensitivity functions defined in respect to the initial conditions (matrix $\Phi_{n_s \times n_s}$) and parameters (matrix $\boldsymbol{S}_{n_s \times n_p}$) respectively:

$$\Phi = [\Phi^{c_i}_{c_{j,0}}] = [\partial c_i(t) / \partial c_{j,0}];$$
$$\boldsymbol{S} = [S^{c_i}_{k_j}] = [\partial c_i(t) / \partial k_j]_{k^*}. \tag{A1}$$

As proved by *Vajda* et al.[30], it is more convenient to express the absolute sensitivities in relative terms referring to $\boldsymbol{c}_s$ and $\boldsymbol{k}^*$ nominal values:

$$\Phi^r = \Phi; \tag{A2}$$
$$S^{r,c_{is}}_{k_j} = [\partial c_i / c_{i,s}]_s \big/ [\partial k_j / k^*_j]_{k^*} = [\partial \ln c_i / \partial \ln k_j]_s.$$

due to the possibility of a direct comparison of sensitivities and of determining eventually relationships among parameters. An integral average of the relative sensitivities can also be evaluated over a dynamic recovering path after an 'impulse' perturbation:

$$\overline{\Phi} = \frac{1}{(t_f - t_0)} \int_{t_0}^{t_f} \Phi^r(t) \mathrm{d}t;$$
$$\overline{\boldsymbol{S}} = \frac{1}{(t_f - t_0)} \int_{t_0}^{t_f} \boldsymbol{S}^r(t) \mathrm{d}t \tag{A3}$$

These sensitivity functions are numerically evaluated over the integration time by means of so-called sensitivity equations, until the QSS is reached with a certain tolerance at final $t_f$:[31]

$$\begin{cases} \mathrm{d}c/\mathrm{d}t = \boldsymbol{h}(\boldsymbol{c},\boldsymbol{k},\boldsymbol{t}); \quad c(t_0) = c_0 \\ \mathrm{d}S/\mathrm{d}t = AS + B; \; S(t_0) = 0; \; A = [\partial h_i(t)/\partial c_j]; \\ B = [\partial h_i(t)/\partial k_j]; \; \mathrm{d}\Phi/\mathrm{d}t = A\Phi; \; \Phi(t_0) = I \end{cases} \tag{A4}$$

($\boldsymbol{I}$ denotes the identity matrix). At QSS, the sensitivity function's values represent the steady-state sensitivity coefficients.

# Appendix B

## Continuation rule to construct QSS-trajectories

The steady-state mass balance of a regulatory kinetic schema corresponds to a $n_s$-dimensional nonlinear algebraic set (3) (index 's' denotes the steady-state value). By varying some of the $\boldsymbol{k}$-values, one can generate $c_s(k)$ curves in the $[n_s \times n_p]$ space. By inspecting the shape of these curves one can determine the 'boundary' or 'limit' or even 'bifurcation' points controlling the number and type of the QSS-s. This QSS-trajectory generation represents the so-called 'system continuation problem'.[32–34]

Depending on the $\boldsymbol{k}$-values, the set (3) can present no solution, single solution or multiple QSS solutions, leading to partitioning the ($\boldsymbol{c},\boldsymbol{k}$) space in several sub-domains. The domain bounds can point-out 'limits' (because of certain physical constraints), 'turning' points (where solution multiplicity change), or 'bifurcation' points (where $\boldsymbol{c}(\boldsymbol{k})$ solution branches intersect each other or all together). At the so-called turning points ($\boldsymbol{c}^*,\boldsymbol{k}^*$) several conditions hold:

$$det\left[\frac{\partial h(c^*,k^*)}{\partial c}\right] = 0; \quad rank\left[\frac{\partial h(c^*,k^*)}{\partial c}\right] < n_s;$$
$$rank\left[\frac{\partial h(c^*,k^*)}{\partial c} \middle| \frac{\partial h(c^*,k^*)}{\partial k}\right] = n_s, \tag{B1}$$

where:

$$\frac{\partial h(c(k),k)}{\partial c}\frac{\mathrm{d}c}{\mathrm{d}k} = -\frac{\partial h(c(k),k)}{\partial k}. \tag{B2}$$

Supplementary conditions can also be used to check switching branches bifurcation points.[32] In order to numerically determine all the singular (bifurcation) points, when the set functions and derivatives tend to zero, the problem (3) is replaced with solving a normalised one by introducing a new parameter '$s$', such that $c(s)$ and $k(s)$ are dependent on $s$ and the following set holds for a certain proposed function $N$:

$$h(c_s,k) = 0; \qquad N(c_s,k,s) = 0. \tag{B3}$$

The advantage of such a re-parameterisation is that, for a fixed value of $s^*$, the derivative matrix becomes non-singular near the singularities points of the initial matrix, i.e.:

$$rank\left[\frac{\partial h(c(s^*),k(s^*))}{\partial c} \; \frac{\partial h(c(s^*),k(s^*))}{\partial k}; \right. \tag{B4}$$
$$\left. \frac{\partial N(c(s^*),k(s^*))}{\partial c} \; \frac{\partial N(c(s^*),k(s^*))}{\partial k}\right] = n_s + 1,$$

The continuation procedure applied to $\boldsymbol{k}$ can equally be applied to $s$, thus bypassing difficulties encountered near the singularity points when applying the solver algorithms to trace branches $c_s(k)$. From the large number of parameterisation methods reported in the literature, those proposed by *Rheinboldt & Burkardt*[35] were adopted in the present study. This alternative considers one of the system variable $x_m$ as being $s$ and receiving successive increasing values, that is:

$$N(c_s,k,s) := x_m - s; \qquad x := [c,k]; \tag{B5}$$
$$m = 1,\ldots,(n_s + n_p).$$

All of the $[c,k]$ vector components can be chose as being $x_m$ (respectively $s$) parameter. However, *Rheinboldt*[36] suggested using that parameter presenting the larger derivatives $\partial h/\partial x_m$, i.e. the larger sensitivities into the model. The continuation conditions being fulfilled, successive solution of (B3, B5) for various increasingly $x_m$-values lead to tracing the QSS single branch. The QSS solutions for one parameter can be repeated when varying two (or more) parameters, thus leading to double (or multiple) bifurcation branch tracing.[37]

# Appendix C

## Steady-state stability and check for periodic solutions

To investigate the stability of a QSS point $c_s$, the model equations (3) have to be locally linearized via evaluation of the Jacobian matrix $\boldsymbol{h}_z$:

$$dc/dt = h(c,k) \approx h(c_s,k) + h_z(c_s,k)(c - c_s); \qquad \text{(C1)}$$

$$h_z = [\partial h(c,k)/\partial c]_s.$$

By inspecting the matrix $\boldsymbol{h}_z$'s eigenvalues ($\beta_j$, $j = 1,\ldots,n_s$), simple stability conditions can be derived.[38] For real $\beta_j$, a 'nodal' steady-state solution is stable if $\beta_j < 0$ (for all *j*), and unstable if $\beta_j > 0$ (for only some *j*). For complex $\beta_j$, the system presents amortised oscillations toward a stable solution for Re ($\beta_j$) < 0 (for all *j*), remnant oscillations for some Re ($\beta_j$) = 0, and un-amortised divergent oscillations for some Re ($\beta_j$) > 0. In the case of only two-variables, the stability is guaranteed if both following inequalities are satisfied:

$$Trace\,(h_z) < 0; \qquad Det(h_z) > 0. \qquad \text{(C2)}$$

At a bifurcation or limit point in a QSS-trace representation, several types of loss of stability might arise when the real part of an eigenvalue is zero. Of special interest is the case of the Hopf bifurcation points, where periodic solutions emerge.

To check a QSS for Hopf periodic solutions when a perturbation occurs, one method is to determine periodic trajectories starting from a stationary point $c_s$ and integrating over a period-interval $[0,T]$ the set:[39]

$$\begin{bmatrix} \dfrac{dc}{dt} \\ \dfrac{dA}{dt} \end{bmatrix} = \begin{bmatrix} h(c,k) \\ h_z(c,k)A \end{bmatrix} \; ; \; \begin{bmatrix} c(0) - c_s \\ A(0) - I \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad \text{(C3a/b)}$$

Equation (C3b) is integrated on a periodic orbit $\boldsymbol{c(t)}$ defined by (C3a), with the period $T$ and parameters $\boldsymbol{k}$. The $[n_s \times n_s]$ matrix $\boldsymbol{I}$ denotes the identity matrix. The eigenvalues of thus obtained $\boldsymbol{M} = \boldsymbol{A(T)}$, the so-called Floquet multipliers of the monodromy matrix $\boldsymbol{M}$, determine the type of stability of the analysed stationary solution.[40] If they are all inside the unit circle of the complex plane, then the periodic orbit $\boldsymbol{c(t)}$ is stable. The stability is lost when one pair of multipliers crosses the unit circle (Hopf bifurcation point).