Modular-Based Modelling of Protein Synthesis Regulation

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

Original scientific paper
Received: February 18, 2005
Accepted: May 25, 2005

A variable-volume modelling framework has been used to build-up modular structures that can reproduce complex protein syntheses inside cells. Methodology is based on a modular kinetic representation of the homeostatic regulatory network that control the metabolic processes, and on a globally efficient module-linking rule that optimise the chain performance indices. The paper exemplifies, at a generic level, how this methodology can be applied to: i) characterize the module efficiency, species connectivity, system stability, based on proposed regulatory indices vs. dynamic and stationary environmental perturbations; ii) build-up modular regulatory chains of various complexity; iii) prove feasibility of the cooperative vs. uncooperative construction that ensures gene expression, system homeostasis, proteic functions, and an equilibrated cell growth during the cell cycle. The more realistic variable-cell-volume dynamic modelling allows an accurate evaluation of individual effector/unit efficiency and perturbation propagation inside a cell, pointing out the influence of module type and enzyme activity allosteric control on regulatory indices.

Key words:
Modular modelling, homeostatic regulation, protein synthesis

Introduction

Living cells are evolutionary, auto-catalytic, self-adjustable structures able to convert raw materials from environment into additional copies of themselves. During the cell growth, protein functions are ensured by internal regulatory networks that adjust the metabolic synthesis to maintain homeostasis, i.e. quasi-invariance of key species concentrations (enzymes, proteins, metabolites), despite of external perturbations (in nutrients and metabolites) or internal cell changes. Due to the highly complex and partly unknown aspects of the metabolic process, the detailed modelling of the whole-cell remains still an unsettled issue, even if recent trials have been reported based on massive databanks and advances in bioinformatics, genomics, transciptomics, proteomics, and metabolomics (e.g. cell simulation platforms).[1,2]

During cell content replication, it was pointed-out that proteins (enzymes) present interactions leading to multi-enzyme complexes that promote a catalytically efficient sequence of reactions over the so-called 'channelling intermediate metabolites'.[3,4] Proteic complexes also act as regulation nodes that provide a balanced response to perturbations, avoiding dysfunctional responses in branched pathways. The regulatory information is transmitted via short-circuits that bypass the cause-effect sequence of intervening reactions. To understand and simulate such a complex mutual-assistance related to the individual proteic functions, various representations of the homeostatic regulatory network have been proposed. The modular approach seems to be preferred, being based on coupled semi-autonomous regulatory groups, linked to efficiently cope with cell perturbations and to ensure an equilibrated cell growth during the cell cycle.[4,5] Because of a limited number of types, individual regulatory modules can be separately analysed and checked for efficiency in conditions that mimic the stationary and perturbed cell growth. Then, they are linked accordingly to certain rules in a hierarchical structure that ensures the overall network efficiency, system homeostasis, and protein functions.

The scope of this paper is to exemplify, at a generic level, the modelling principle of assembling regulatory modules for gene expression, in order to build-up metabolic regulatory networks of protein synthesis. Methodology is based on the modular kinetic representation of the homeostatic regulation during metabolic synthesis. Various types of regulatory modules can be thus proposed and characterized by means of proposed "efficiency indices" (P.I.), which evaluate the species connectivity, system stability, and recovery effectiveness of a steady state after stationary and dynamic perturbations, respectively. Modules are then linked to ensure a globally efficient regulatory network that ensures gene expression, system homeostasis, proteic functions, and an equilibrated cell growth during the

cell cycle. The paper exemplifies how such a methodology should be used to step-by-step build-up a regulatory chain, and how the feasibility and effectiveness of the (cooperatively-linked) modular construction can be checked in every modelling step by means of P.I.-s.

The modelling framework uses continuous-variable ordinary differential (ODE) kinetic models to check the synthesis regulatory efficiency, by explicitly accounting for cell-volume variation during the cell cycle, protein interactions and regulatory loops under stationary and dynamic environmental perturbations. The more realistic modelling approach allows an accurate evaluation of the individual effector/unit relative importance, perturbation propagation inside a cell, and the influence of module type and enzyme activity allosteric control on its regulatory P.I.-s. Once the methodology elaborated, such a cell modelling platform can be detailed and used to simulate complex protein / species interactions, cooperative regulation of the metabolic synthesis, and other processes during the cell growth.

Modelling metabolic regulation – A short review

Protein synthesis by gene expression is a highly regulated process which ensure a balanced and flexible cell growth under an indefinitely variate environmental conditions. How this very complex process occurs is partially understood, but a multi-cascade control with negative feedback loops seems to be the key element.[6–9] Enzymes catalysing the synthesis are allosterically regulated by means of positive or negative effector molecules, while cooperative binding and structured cascade regulation amplify the effect of a change in a signal. Gene expression is also highly regulated to flexibly respond to the environmental stress, in a scheme looking like a 'genetic circuit'.[10] While the effect of changes in system parameters on the system states (variables) is approached by the metabolic control analysis, the metabolic regulator features are determined by its ability to efficiently vary species flows and concentrations under changing environmental conditions so, that a stationary state of the key metabolite concentrations can be maintained inside the cell.

To model such a complex metabolic regulatory mechanisms at a molecular level, two main approaches have been developed over decades: structure-oriented analysis, and dynamic (kinetic) models.[11] Each theory presents strengths and shortcomings in providing an integrated predictive description of the cellular regulatory network.

Structure-oriented analyses ignore some mechanistic details and the process kinetics, and use the only network topology to quantitatively characterize to what extent the metabolic reactions determine the fluxes and metabolic concentrations.[12] The so-called 'metabolic control analysis' (MCA) is focus on using various types of sensitivity coefficients (the so-called 'response coefficients'), which are quantitative measures of how much a perturbation (influential variable x_j) affects the cell-system states y_i (e.g. r = reaction rates, J = fluxes, c = concentrations) around the steady-state (QSS, of index 's'), i.e. $\left[(\partial y_i / y_{is})/(\partial x_j / x_{js})\right]_s$. The systemic response of fluxes or concentrations to perturbation parameters (i.e. the 'control coefficients'), or of reaction rates to perturbations (i.e. the 'elasticity coefficients') have to fulfil the 'summation theorems', which reflect the network structural properties, and the 'connectivity theorems' related to the properties of single enzymes vs. the system behaviour.

Originally, MCA has been introduced by *Kacser & Burns*,[13] *Heinrich & Rapoport*,[14] and *Burns* et al.[15] to quantify the rate limitation in complex enzymatic systems. MCA have been followed by a large number of improvements, mainly dealing with the control analysis of the stationary states, by pointing-out the role of particular reactions and cell components in determining certain metabolic behaviour. Successive extensions of such definitions allow: to study any limit set for non-steady/time-dependent conditions;[16,17] the flux balance analysis and optimization (FBA);[11,12,18–22] elementary mode analysis (EMA);[11,18,19] dynamic flux balance analysis (DFBA);[23] extreme pathway analysis (ExPA);[18,19] constrained based modelling of metabolic network (CBM).[24]

MCA methods are able to efficiently characterize the metabolic network robustness and functionality, linked with the cell phenotype and gene regulation. MCA allows a rapid evaluation of the system response to perturbations (especially of the enzymatic activity), possibilities of control and self-regulation for the whole path or some subunits. Functional subunits are metabolic subsystems, called 'modules', such as amino acid or protein synthesis, protein degradation, mitochondria metabolic path, etc.[25] Because the living cells are self-evolutive systems, new reactions recruited by cells together with enzyme adaptations, can lead to an increase in the cell biological organisation and to optimal performance indices. When constructing methods to optimize evolutive metabolic systems, MCA concepts and appropriate performance criteria have been used to: maximize reaction rates and steady-state fluxes; minimize metabolic intermediate concentrations; minimize transient times; optimise the reac-

tion stoichiometry (network topology); maximize thermodynamic efficiency. All these objectives are subjected to various mass balance, thermodynamic, and biological constraints.[12] However, by not accounting for the system dynamics, and grounding the analysis on the linear system theory, topological methods present inherent limitations (see for instance some violations of stoichiometric constraints discussed by *Atauri* et al.,[26] or the modified control coefficients of *Szedlacsek* et al.[27,28]).

Classical approach to develop dynamic models is based on an hypothetical reaction mechanism, kinetic equations, and known stoichiometry. This route meets difficulties when the analysis is expanded to large-scale metabolic networks, because the necessary mechanistic details and standard kinetic data to derive the rate constants are difficult to be obtained. However, advances in genomics, transcriptomics, proteomics, and metabolomics, lead to a continuous expansion of bioinformatic databases, while advanced numerical techniques, non-conventional estimation procedures, and massive software platforms reported progresses in formulating reliable cell models. Valuable *structured dynamic* models, based on cell biochemical mechanisms, have been developed for simulating various (sub)systems, such as:

– entire cell: 'whole-cell' models, as E-Cell, V-Cell, M-Cell, A-Cell, etc.;[1]

– single cell growth: *Escherichia coli*;[29–31] *Haemophilus influenzae*;[18,19] *Mycoplasma genitalium*;[32–34] yeast;[35,36]

– various (oscillatory) metabolic paths: metabolism of human red blood cells;[36,37] glycolysis, citric acid cycle, oxidative phosphorylation[12,38–40]; iron metabolism;[41] amino acid synthesis;[20]

– regulatory networks for: gene expression via Boolean 'biocircuit' models (*Escherichia coli*;[42,43] prokaryotes;[44,45] protein synthesis[1,5,9]);

– cell cycles and oscillatory systems in yeast and eukaryotes: limit-cycle oscillator models;[46–48] cell size control, oscillation properties and hysteresis effects;[49] key ingredients inducing oscillations;[49–55] whole-cell framework for cell cycle simulation;[56]

– cellular communications, intracellular signalling, neuronal transmission, networks of nerve cells;[57–59]

– analysis of "logical essence" of life, and the life fundamental requirements.[60]

To model in detail the cell process complexity is a challenging and difficult task. The large number of inner cell species, complex regulatory chains, cell signalling, motility, organelle transport, gene transcription, morphogenesis and cellular differentiation cannot easily be accommodated into existing computer frameworks. Inherently, any model represents a simplification of the real phenomenon, while relevant model parameters are estimated based on the how close the model behaviour is to the real cell behaviour. A large number of software packages have been elaborated allowing the kinetic performance of enzyme pathways to be represented and evaluated quantitatively.[61] Oriented and unified programming languages have been developed (CellML;[62] SBML[61]) to include the bio-system organization and complexity in integrated platforms for cellular system simulation (E-Cell;[32,33] V-Cell;[63,64] M-Cell;[58] A-Cell[59]). Such integrated simulation platforms tend to use a large variety of biological databanks including enzymes, proteins and genes characteristics together with metabolic reactions (CRGM-database;[65] NIH-database[66]).

From the mathematical point of view, various structured (mechanism-based) dynamic models have been proposed to simulate the metabolic regulation, accounting for continuous, discrete, and/or stochastic variables, in a modular construction, 'circuit-like' network, or compartmented simulation platforms.[1,2,67] Such models can include:

– Boolean (discrete) variables;[10,44,68–70]

– continuous variable models;[9,12,68,71]

– stochastic variable models;[45,57,72,73]

– mixed variable models.[68]

In the Boolean approach, variables can take only discrete values. Even if less realistic, such an approach is computationally tractable, involving networks of genes that are either "on" or "off" (e.g. a gene is either fully expressed or not expressed at all) according to simple Boolean relationships, in a finite space. Such a coarse representation is used to obtain a first model for a complex biosystem including a large number of components, until more detailed data on process dynamics become available. 'Electronic circuits' structures have been extensively used to understand intermediate levels of regulation, but they cannot reproduce in detail molecular interactions with slow and continuous responses to perturbations.

Even if regulation mechanisms are not fully understood, metabolic regulation at a low-level is generally better clarified. Based on that, conventional dynamic models (ODE kinetics), with a mechanistic description of reactions taking place among individual species (proteins, mRNA, intermediates, etc.) have been proved to be a convenient route to analyse continuous metabolic/regulatory processes and perturbations. When systems are too large or poorly understood, coarser and more phenomenological kinetic models may be postulated (e.g. protein complexes, metabolite channelling, etc.). In dynamic models, only essential reactions are retained, the model complexity depending

on measurable variables and available information. An important problem to be considered is the distinction between the qualitative and quantitative process knowledge, stability and instability of involved species, the dominant fast and slow modes of process dynamics, reaction time constants, macroscopic and microscopic observable elements of the state vector. Such kinetic models can be useful to analyse the regulatory cell-functions, both for stationary and dynamic perturbations, to model cell cycles and oscillatory metabolic paths,[56] and to reflect the species interconnectivity or perturbation effects on cell growth.[74] Mixtures of ODE kinetics with discrete states (i.e. 'continuous logical' models), and of continuous ODE kinetics with stochastic terms, can lead to promising mixed models able to simulate, both, deterministic and non-deterministic cell processes.[68]

Stochastic models replace the 'average' solution of continuous-variable ODE kinetics (e.g. species concentrations) by a detailed random-based simulator accounting for the exact number of molecules present in the system. Because the small number of molecules for a certain species is more sensitive to stochasticity of a metabolic process than the species present in larger amounts, simulation via continuous models lacks of accuracy for random process representation (as cell signalling, gene mutation, etc.). Monte Carlo simulators are used to predict individual species molecular interactions, while rate equations are replaced by individual reaction probabilities, and the model output is stochastic in nature. Even if the required computational effort is very high, stochastic representation is useful to simulate the cell system dynamics by accounting for a large number of species of which spatial location is important.

By applying various modelling routes, successful structured models have been elaborated to simulate various regulatory mechanisms.[10,43,44,49,69,75–79,92] In fact, as mentioned by *Crampin & Schnell*,[67] a precondition for a reliable modelling is the correct identification of both topological and kinetic properties. As few (kinetic) data are present in a standard form, non-conventional estimation methods have been developed by accounting for various types of information (even incomplete) and global cell (regulatory) properties.[2,67]

Modular structures for protein synthesis regulation

Lumped representation

Living cells are organized, self-replicating, evolvable, and responsive to environmental biologycal systems. The structural and functional cell organization, including components and reactions, is very complex. Relationships between structure, function and regulation in complex cellular networks, are better understood at a low (component) level rather than at the highest-level.[11] Cell regulatory and adaptive properties are based on *homeostatic* mechanisms, which maintain quasi-constant key-species concentrations and output levels, by adjusting the synthesis rates, by switching between alternative substrates, or development pathways. Cell regulatory mechanisms include allosteric enzymatic interactions and feedback in gene transcription networks, metabolic pathways, signal transduction, and other species interactions.[67] In particular, protein synthesis homeostatic regulation includes a multi-cascade control with negative feedback loops and allosteric adjustment of the enzymatic activity.[8,9]

A convenient way to model metabolic processes and regulatory networks is the *modular* approach. Spatial and functional compartments together with functional modules have been defined when developing complex cell simulation platforms.[9,12,32,61,67,93] This increasing trend is based on the observations of *Kholodenko* et al.,[90] that metabolic networks can be decomposed in functional sub-units called *modules*. Grouping enzymes and other species into "modules", according to existence of functional units (i.e. particular pathways but also spatial structures), leads to a "modular" approach applied, both, for a structure-oriented analysis and for deriving dynamic models.[12] *Sauro & Kholodenko*[91] provide examples of biological systems that have evolved in a modular fashion and, in different contexts, perform the same basic functions. Each module, grouping several cell components and reactions, generates an identifiable function (e.g. synthesis regulation for a certain component, regulation of a certain reaction, etc.). More complex functions, as regulatory networks, synthesis networks, or metabolic cycles can be built-up from basic building blocks. The modular approach is also computationally tractable, allowing elaboration of complex simulation platforms accounting for separate cell sub-units, metabolic functional sub-units (synthesis, degradation, etc.), or certain component metabolism.[1,2] Modular approach can also be useful in simulating the hierarchical organization of cell regulatory networks.

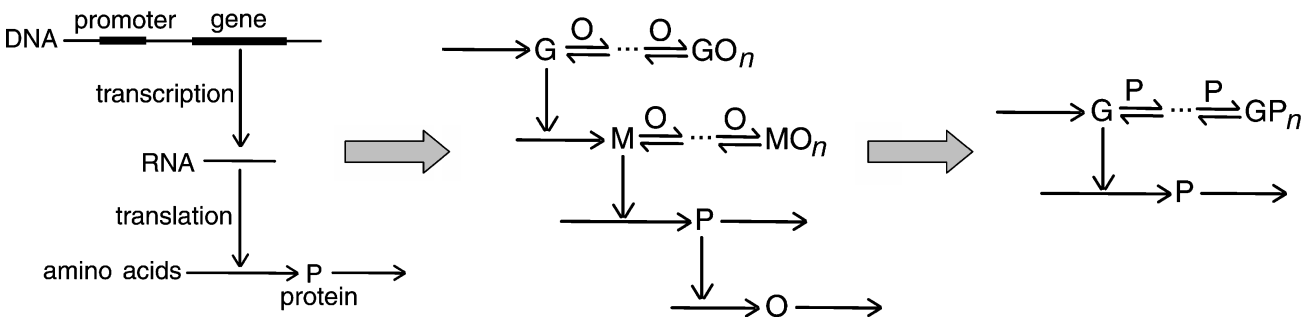
Concerning the protein synthesis, this process is presumably regulated by a complex homeostatic mechanism that controls the expression of the encoding genes. On the other hand, cells contain a large number of proteins of well-defined functions, but strongly interrelated to ensure an efficient metabolism and cell growth under certain environmental conditions. Proteins interact during the synthesis and, as a consequence, the homeostatic systems

perturb and are perturbed by each other.[9] To understand and simulate such a complex regulatory process, the modular approach is preferred, being based on coupled semi-autonomous regulatory groups (of reactions and species), linked to efficiently cope with cell perturbations, to ensure system homeostasis, and an equilibrated cell growth.[9] Various types of kinetic modules can be analysed individually as mechanism, reaction path, regulatory characteristics, and effectiveness. As a limited number of regulatory module types govern the protein synthesis, it is computationally convenient to step-by-step build-up the modular regulatory network by applying certain principles and rules, and then adjusting the network global properties. Accordingly, it is desirable to focus the metabolic regulation and control analysis on the regulatory/control features of functional subunits than to limit the analysis to only kinetic properties of individual enzymes acting over the synthesis path.

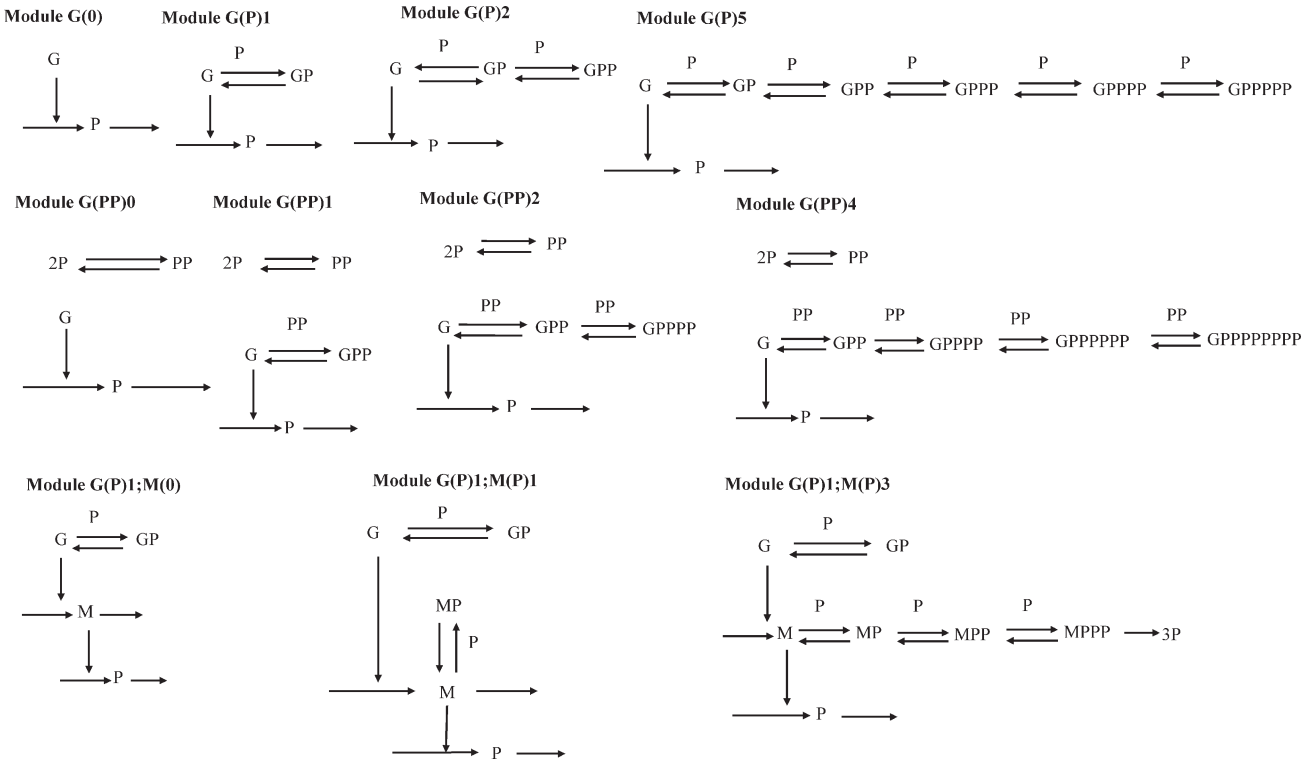
The modular approach assumes that the reaction mechanism and stoichiometry of the kinetic module are known, while the involved species are completely observable and measurable. Such a hypothesis is rarely fulfilled due to the inherent difficulties in generating reliable experimental (kinetic) data for each individual metabolic subunit. However, incomplete kinetic information can be incorporated by performing a suitable model lumping,[2] or by exploiting the cell and module global properties during identification steps.[1,5] The regulatory modules can be constructed relatively independent to each other, but the linking procedure has to consider common input/output components, common linking reactions, or even common species. Rate constants can be identified separately for each module, and then extrapolated when simulating the whole regulatory network, by assuming that linking reactions are relatively slow comparatively with the individual module core reactions.[9] In such a manner, linked modules are able to respond to changes in common environment and components such, that each module remains fully regulated.

When elaborating a protein synthesis regulatory module, different degrees of simplification of the process complexity can be followed. For instance, the gene expression (see schema of Fig. 1) can be translated into a modular structure of reactions, more or less extended, accounting for individual or lumped species. At a generic level, in the simplest representation (Fig. 1, right), the protein (P) synthesis rate can be adjusted by the 'catalytic' action of the encoding gene (G). The catalyst activity is in turn allosterically regulated by means of 'effector' molecules (O or P), reversibly binding the catalyst via fast and reversible reactions (the so-called 'buffering' reactions). This simplest regulation schema can be further detailed in order to better reproduce the real process, with the expense of a supplementary effort to identify the module kinetic parameters. For instance, a two-step cascade control of P-synthesis includes the M = mRNA transcript encoding P (Fig. 1). The effector (O), of which synthesis is controlled by the target protein P, can allosterically adjust the activity of G and M, i.e. the catalysts for the transcription and translation steps of the gene expression. In such a cascade schema, the rate of the ultimate reaction is amplified, depending on the number of cascade levels and catalysis rates. More complex regulatory modules can be elaborated following a similar route to 'translate' from the 'language' of molecular biology to that of mechanistic chemistry, by preserving the structural hierarchy and component functions. Once elaborated, such a modular structure can be modelled by using one of the previously described alternatives, and then analysed as functional efficiency by means of some defined performance indices.

Recently, *Sewell* et al.[9] and *Yang* et al.[5] studied various types of hypothetical modules designed to ensure homeostatic regulation of a generic protein-gene (P/G) pair synthesis, with exemplifications from *E. coli* (some of them are displayed in Fig. 2). Simplified representations include the essential nutrient lumps (NutP, NutG), metabolites



F i g . 1 – *Protein P synthesis – simplified representations of the regulatory module (horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; absence of a substrate or product indicate an assumed concentration invariance of these species; G= gene encoding P; M= mRNA; O= allosteric effectors).*

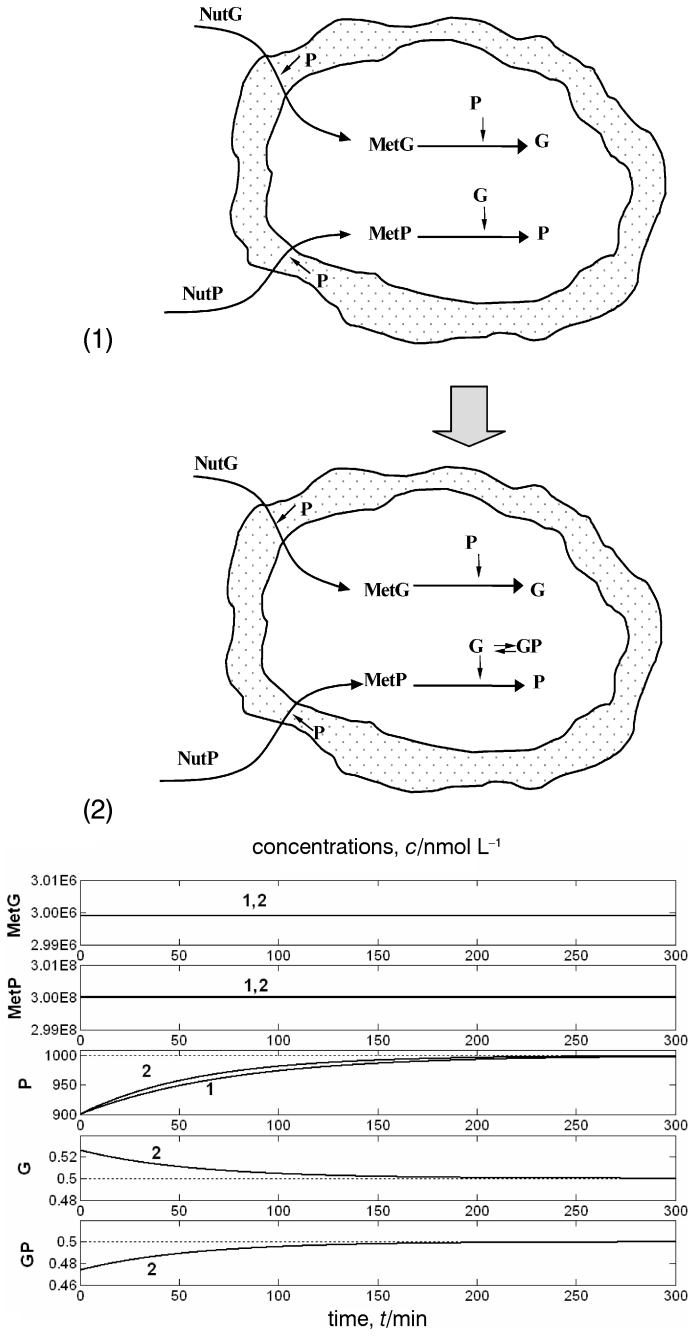


F i g . 2 – *Various types of regulatory modules for protein synthesis (from Yang et al.[5]; horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; absence of a substrate or product indicate an assumed concentration invariance of these species).*

(MetP, MetG), and intermediates involved in the reactions, controlling the transcriptional and translation steps of the P synthesis. The module nomenclature, proposed by *Yang* et al.,[5] of type $[L_1(O_1)n_1; \ldots; L_i(O_i)n_i]$ includes the assembled regulatory units $L_i(O_i)n_i$. One unit *i* is formed by the component L_i (e.g. enzymes or even genes G, P, M, etc.) at which regulatory element acts, and n_i = 0,1,2,… number of species O_i (i.e. 'effectors' P, PP, PPPP, etc.) binding the 'catalyst' L. For instance, a *G(P)5* unit includes five successive binding steps of G with the product P, all intermediate species GP, GPP, GPP, GPPP, GPPPP, GPPPPP being inactive catalytically, while the mass conservation law is all time fulfilled, i.e. $\sum_{i=0}^{5}[G(P_i)] = \text{constant}$. Such a representation accounts for the protein concentration diminishment due to the cell-growth dilution effect, but could also include protein degradation by proteolysis.

The G(P)*n* type of units, even less realistic, represent the simplest regulatory module used as control mechanism against which all others are compared. In a G(P)0 module (see also Fig. 3), there are only two main synthesis chains. P is a permease that catalyses the import of NutG and NutP from the environment, and a metabolase that converts them into cellular metabolites MetG and MetP. P is also a polymerase that catalyses the synthesis of G from MetG. Gene G, symbolizing the genome of the cell, functions as catalyst for the synthesis of P from MetP. The result is that G and P syntheses are mutually autocatalytic. In *G(P)0* there are no regulatory elements. In *G(P)1*, the negative feedback control of transcription is realised by P itself (as effector), via a rapid buffering reaction, G + P <==> GP, leading to the catalytically inactive GP. As proved, the maximum regulatory efficiency at steady-state (index 's') corresponds to $[G]_s/[G]_{total}$ = 1/2, when the maximum regulation sensitivity vs. perturbations in $[P]_s$ is reached.[9] Further allosteric control of G activity, leading to inactive species $[GP_n]$, amplifies the regulatory efficiency of the module. As an example, prokaryotes commonly bind multiple copies of transcription factors as a means of promoting cooperative effects and thus improving regulatory effectiveness.[5] For instance, DnaA is an autoregulated protein and at least five copies can bind to *dnaA* gene in *E. coli*.[5,80]

The G(PP)*n* units reflect better the regulatory loops in which multiple copies of effectors (proteins and transcription factors) bind to promoter sites on the DNA that control expression of gene G encoding P (see exemplifications from *E. coli* by *Yang* et al.[5]). The control is better realised by including a supplementary P dimerization step before



F i g . 3 – *Constructing the G(P)1 regulatory module (2) by adding a buffering reaction to the G(P)0 module (1). Cell components recover after a –10 % impulse perturbation in* $[P]_s$ *= 1000 nmol* L^{-1} *at t = 0 (down).*

the buffering reactions. This explains why most of transcription factors bind as oligomers (typically dimers or tetramers) and why they typically bind in multiple copies.[5] For instance, in *E. coli* the monomeric $\pi^{35.0}$ protein of plasmid R6K forms dimers that bind to the operator of the *pir* gene that encodes the protein, and represses its own synthesis.[5,89] The λ repressor transcription factor is a dimer, three copies of which bind the operator region of the gene that it regulates.[5,81]

Module [G(P)*n*; M(P)*n′*] tries to reproduce more accurately the transcription/translation cascade of reactions during the gene expression, by including an allosteric control at two levels of catalysis: on G (i.e. DNA) and on M (i.e. mRNA). M is synthesized from nucleotides under G catalysis, and then, in a translation step, P is synthesized in a reaction catalysed by M. Such a supplementary control of mRNA activity is proved to be a more effective means of regulating protein synthesis.[5,82]

Module efficiency indices

To evaluate and compare the regulatory efficiency of various kinetic modules, when maintaining cell homeostasis, two categories of indices have been used, defined under stationary ('step' like) or dynamic ('impulse' like) continuous perturbations of species stationary concentrations. Random perturbations, due to interactions of P-synthesis module with other metabolic processes, or due to environmental changes, lead to a module response that tends to maintain the key-component functions. Module efficiency depends on the regulatory structure, species inter-connectivity, QSS-characteristics, cell size, and perturbation magnitude.

Stationary perturbations refers to permanent modifications of nutrient/metabolite levels, leading to new stationary concentrations inside the cell. Referring to the target protein P, the regulatory module tends to diminish the deviation $[P]_s - [P]_{ns}$ between the 'nominal' QSS (unperturbed set-point, of index 'n') and the new QSS reached after perturbation. Equivalently, the P-synthesis regulatory module will tend to maintain $[P]_{ns}$ within certain limits, $[P]_{min} \leq [P]_{ns} \leq [P]_{max}$ (a relative $R_{SS} = \pm$ 10 % maximum deviation has been proposed by *Sewell* et al.[9]). A measure of steady-state (c_{is}) 'resistance' to various perturbations (in rate coefficients, k_j, or in nutrient concentrations, c_{NUT_j}) is given by the magnitude of relative sensitivity coefficients at QSS, i.e. $S^{c_i}_{k_j}$ and $S^{c_i}_{Nut_j}$, respectively [where $S^{state}_{perturb}$ = ∂(State)/∂(Perturbation) are the state sensitivities vs. perturbations[1]]. A regulatory index, $A_{unsync} = k_{syn} \times k_{decline}$, has been introduced to illustrate the maximum levels of (unsynchronised) stationary perturbations in synthesis or consumption rates of a key-species tolerated by the cell within defined limits.[9]

Dynamic perturbations are instantaneous changes in the concentration of one or more components that arise from a process lasting an infinitesimal time (impulse-perturbation). After perturbation, the system recovers and returns to their stable nominal QSS. The recovering time and rate can be approximated from the solution of the liniarized system model:[12]

$$d\boldsymbol{c} / dt = \boldsymbol{h}(\boldsymbol{c}, \boldsymbol{k}); \quad \boldsymbol{c}(0) = \boldsymbol{c}_s;$$

$$\boldsymbol{c}(t) = \boldsymbol{c}_s + \sum_{i=1}^{n_s} d_i b_i \exp(\lambda_i t), \qquad (1)$$

(where: $\boldsymbol{c}$ = concentration vector; λ_i = eigenvalues of the system Jacobian matrix at QSS, $\boldsymbol{J}_c = (\partial \boldsymbol{h}(\boldsymbol{c}, \boldsymbol{k})/\partial \boldsymbol{c})_s$; b_i, d_i = constants, depending on the system characteristics at stationary conditions; t = time). If the real parts of eigenvalues are all negative, then the stationary state $\boldsymbol{c}_s$ is stable. The recovering rate index R_D of *Yang* et al.[5] reflects the recovering properties of the regulated P-synthesis system. The species j recovering time $\tau_j \sim 1/R_D$ is necessary to the system to return to the stationary c_{js} concentration, with a certain tolerance and for a defined perturbation magnitude (see below the proposed 1 % tolerance for a ±10 % c_{js} impulse perturbation).

Steady-state $\boldsymbol{c}_s$ stability strength is related to the system characteristics. As $\text{Max}(\text{Re}(\lambda_i)) < 0$ is smaller, the QSS is more stable. When analysing the predicted QSS and regulatory characteristics of a P-synthesis module, the stability strength can also be related to an index against periodic oscillations. This can be evaluated from the liniarized form of the system model, by calculating the monodormy matrix ***A(T)*** after a checked period T of time:[1]

$$d\boldsymbol{c} / dt = \boldsymbol{h}(\boldsymbol{c}, \boldsymbol{k}); \quad \boldsymbol{c}(0) = \boldsymbol{c}_s;$$

$$d\boldsymbol{A} / dt = \boldsymbol{J}_c \boldsymbol{A}; \quad \boldsymbol{A}(0) = \boldsymbol{I}. \qquad (2)$$

For a stable QSS, i.e. $|\lambda_{A_i}| < 1$, as $|\lambda_{A_i}|$ are smaller, the system stability strength is higher [λ_{A_i} are the eigenvalues of the ***A(T)*** matrix].

Species interconnectivity in a modular regulatory schema of reactions can be viewed as a degree of which they assist each other during the system recovering. Cell species connections appear due to common reactions, or common intermediates participating to chain reactions, or from the common cell volume to which all species contribute (under constant osmotic pressure, see below). *Vance* et al.[83] reviewed and proposed several quick experimental – computational rules to check a reaction schema via species inter-connectivities. By inducing experimental perturbations to a (bio)chemical system, by means of tracers, or by fluctuating the inputs of the system, one can measure the perturbation propagation through the consecutive/parallel reaction path. Then, various techniques can determine the "distance" among observed species, and rules to include this information in elaborating a reaction schema. In the present study, one proposes a similar approximate measure of species interconnectivity related to the species recovering-times after a dynamic perturbation, that is: AVG(τ_j) and STD(τ_j), i.e. the average and standard deviation of τ_j. As AVG and STD are larger, as the cell dynamic regulatory effectiveness is lower, species less interconnected, and components recover more disparately (scattered). As the number of effectors and buffering reactions is higher, these dynamic regulatory indices of the module are smaller.

By summarizing, the regulatory efficiency indices, proposed to evaluate a module or a modular chain behaviour vs. perturbations, are the following (in parenthesis are mentioned the suitable Min/Max objectives):

– stationary regulation: (Min)$R_{ss} = ([P]_s - [P]_{ns})/[P]_{ns}$; (Max) $A_{unsync} = k_{syn} \times k_{decline}$; (Min) $S^i_{NutP_j} = [(\partial c_i/c_{is})/(\partial c_{Nut_j}/c_{Nut_{js}}]_s$; (Min) $S^i_{k_j} = [(\partial c_i/c_{is})/(\partial k_j/k_j)]_s$;

– dynamic regulation: (Min) $R_D = \text{Max}(\text{Re}(\lambda_i))$, $\text{Re}(\lambda_i) < 0$; (Min) τ_j;

– regulatory robustness: (Min) $(\partial R_D/\partial k)$

– species interconnectivity: (Min) AVG(τ_j) = average (τ_j); (Min) STD(τ_j)= st.dev.(τ_j);

– QSS stability strength: (Min) $\text{Max}(\text{Re}(\lambda_i))$, $\text{Re}(\lambda_i) < 0$; (Min) $|\lambda_{A_i}| < 1$.

To evaluate the module regulatory indices and to simulate the system dynamic evolution after a perturbation, rate constants (***k*** vector) and stationary states (species concentrations $\boldsymbol{c}_s$) for a certain cell system are necessary to be known. In the absence of standard kinetic data,[2] rate constants have been estimated from system invariants and from some cell properties. Assuming that most of stationary $\boldsymbol{c}_s$ are known under an equilibrated cell growth (from experiments or databanks), the time-invariant condition imposed to eq. (1) lead to a (non)linear algebraic set to be used for estimating ***k***-s. If the size of ***k***-vector together with unobservable $\boldsymbol{c}_s$ is higher than the observable $\boldsymbol{c}_s$ vector size, the unknowns have been estimated by imposing some properties to the system (under some physical constraints), such as: maximum recovering rate after a perturbation (R_D of *Yang* et al.[5]), smallest sensitivity of steady-states vs. perturbations,[1] stability strength of QSS,[1] oscillatory properties, system flexibility, etc.[1,2] This is the case, for instance, of estimating the $[PP]_s$ level in modules *G(PP)n*, or of $[M]_s$ level in modules *G(P)n;M(P)n'*, by minimizing a performance index, such as AVG(τ_j) or τ_j for a key-species (see below).

The modular representation of protein synthesis allows an easier and quantitative analysis of individual regulatory modules, by comparing various mechanisms and effectors (as number and nature) from the efficiency point of view. Several conclusions can be derived:[5,74]

– Performance indices (P.I.) depend on the type of the regulatory module used by the protein syn-

thesis. The increasing regulatory effectiveness follows the following order: *G(P)n* < *G(PP)n* < *G(P)n;M(P)n'*. Modules reporting high stationary-regulation P.I.-s also report high dynamic-regulation P.I.-s.

– The catalyst activity control at a single enzyme level appears to be of lowest efficiency.

– Multiple copies of effector molecules, which reversibly and sequentially (allosterically) bind the catalyst (G, M) in negative feedbacks, improve the regulation effectiveness.

– A structured cascade control of several enzyme activities, with negative feedback loops at each level, improves regulation and amplifies the effect of a change in a signal. The rate of the ultimate reaction is amplified, depending on the number of cascade levels and catalysis rates. As an example, placing regulatory elements at the level of mRNA is highly effective (i.e. species M in the module *G(P)n; M(P)n'*).

– For the same module, each added regulatory element linearly improves the regulatory index, by an approximate relationship:[5]

$$\text{P.I.} = a_0 + \sum_i a_i n_i , \tag{3}$$

(where P.I. = regulatory performance index, such as R_D, AVG(τ_j), STD(τ_j); n_i = number of effectors (P, PP, O) acting in the *i-th* allosteric unit $L_i(O_i)n_i$; a_0, a_i = constants related to the P.I. and module type).

– P.I. improves ca. 1.3–2 times (or even more) for every added regulatory unit to the module. Multiple regulatory units lead to average recovering times AVG(τ_j), much lower than the cell cycle period t_c, under constant logarithmic volume growing rate, $D = \ln(2)/t_c$.

– Combinations of regulatory schemes and units (with different effectors) improve the regulatory P.I.-s.

– Certain regulatory modules reported an increased flexibility, due to 'adjustable' intermediate species levels. This is the case, for instance, of adjusting $[M]_s$ in module *G(P)n; M(P)n'* and of $[PP]_s$ in module *G(PP)n*. Optimal levels of these species can be set accordingly to various optimization criteria, rendering complex regulatory modules to be more flexible in reproducing certain desired cell-synthesis regulatory properties.

Modelling cooperative linking of regulatory modules

Model hypotheses

To study the protein synthesis regulation, a continuous-variable ODE model of a classical formulation has been adopted for the cell system, accounting for variable-volume and osmotic pressure, but neglecting the inner-cell gradients:[84,85]

$$\frac{\mathrm{d}c_j}{\mathrm{d}t} = \frac{1}{V}\frac{\mathrm{d}n_j}{\mathrm{d}t} - Dc_j ; \quad \frac{1}{V}\frac{\mathrm{d}n_j}{\mathrm{d}t} = r_j ; \quad j = 1,\ldots,n_s ;$$

$$V(t) = \frac{RT}{\pi}\sum_{j=1}^{n_s} n_j(t); \tag{4}$$

$$D = \frac{1}{V}\frac{\mathrm{d}V}{\mathrm{d}t} = \left(\frac{RT}{\pi}\right)\sum_{j}^{n_s}\left(\frac{1}{V}\frac{\mathrm{d}n_j}{\mathrm{d}t}\right).$$

(where: c_j = cell-species j concentration; V = cell volume; n_j = amount of species j; r_j = j-th reaction rate; D = cell-content dilution rate, i.e. cell-volume logarithmic growing rate; π = osmotic pressure; T = temperature; R = universal gas constant; n_s = number of species inside the cell; t = time). Volume dynamics have been linked to molecular species dynamics and to (constant) osmotic pressure by means of the Pfeffer's law in diluted solutions.[86] Such a variable-volume formulation is suitable to accurately model the cell growth,[56] the response to perturbations, cell-'ballast' and 'inertial' effects vs. continuous changes in cell and environment. As ca. 80 % of the cycle period is the growing phase,[56] and assuming a quasi-constant osmotic pressure and a constant volume growing logarithmic rate, model (4) can be considered satisfactory to study the regulatory network effectiveness. The constant osmotic pressure assumption leads to fulfilment of the following invariance relationship:

$$\frac{RT}{\pi} = \frac{V}{\sum_{j=1}^{n_s} n_j} = \frac{1}{\sum_{j=1}^{n_s} c_j} = \frac{1}{\sum_{j=1}^{n_s} c_{j0}} = \text{constant}. \tag{5}$$

By summarizing, the used continuous-variable model (4) has been based on the following hypotheses:

(i) The cell is an isotherm open system with an uniform content (well-mixed reactor case); species behave ideally, and present uniform concentrations within cell.

(ii) The open environment interacts with the cell through a semi-permeable membrane.

(iii) The membrane, of negligible volume, presents a negligible resistance to nutrient diffusion; membrane dynamics is neglected in the model, being assumed to follow the cell growing dynamics.

(iv) When studying an individual P-synthesis regulatory module, other cell species are lumped together.

(v) The inner osmotic pressure is constant, and all the time equal with the environmental pressure, thus ensuring the membrane integrity ($\pi_{cyt} = \pi_{env} =$ constant). As a consequence, the isotonic osmolarity under isothermal conditions leads to the equality $RT/\pi_{cyt} = RT/\pi_{env}$, which indicates that the sum of cell species concentrations must equal those of the environment, i.e. $\left(\sum_j^{all} c_j\right)_{cyt} = \left(\sum_j^{all} c_j\right)_{env}$. Even if in a real cell such an equality is approximately fulfilled due to perturbations and transport gradients, and in spite of migrating nutrients from environment into the cell, the overall environment concentration is considered to remain unchanged. On the other hand, species inside the cell transform the nutrients into metabolites and react to make more cell components. In turn, increasing amounts of polymerases are used to import increasing amounts of nutrients. The net result is an exponential increase of cellular components in time, which translates, through isotonic osmolarity assumption, into an exponential increase in volume with time.[56] The overall concentration of cellular components is time-invariant, because the rate at which cell-volume increases equals (in an normal cell) that at which overall amount increases, leading to a constant $V/\sum_{j=1}^{n_s} n_j$ ratio.

(vi) Cell volume doubles over the cell cycle period (t_c), with an average logarithmic growing rate of $D = \ln(2)/t_c$. Under stationary growing conditions, it results from (4) an exponential volume increase, i.e. $V = V_0 e^{D_s t}$.

(vii) For stationary growing conditions, species synthesis rates are equal to first-order dilution rates (Dc_j), leading to time-invariant species concentrations [i.e. homeostatic conditions, $(dc_j/dt)_s = 0$].

A variable-volume cell model formulation presents an important number of advantages:

– the estimated rate constants are more realistic comparatively with those derived from constant-volume model formulations (usually used in modelling biochemical systems);

– some simplifications, such as dilution terms defined for only key species[5,9] are removed, and all species are treated on the same basis;

– species inter-connectivity (i.e. the degree to which a perturbation in one component influences others) is better characterized by including direct interrelations (via common reactions and intermediates) but also indirect relationships via the common cell-volume to which all species contribute (see eq. 4);

– possible perturbations in the volume size and osmotic pressure can be also considered;

– perturbations applied to components of large concentrations lead to an important cell volume perturbation, which in turn lead to large perturbations of other cell component concentrations (i.e. the so-called 'secondary' or 'indirect' perturbations); vice-versa, perturbations in species of low levels will have a low effect on the cell volume, and then a small secondary effect on other components, because: $V_{perturb}/V_0 = \left(\sum_j n_j\right)_{perturb} \Big/ \left(\sum_j n_j\right)_0$.

– cells of large content (large 'ballast') diminishes the indirect perturbations (the so-called 'inertial' effect, or perturbation smoothing); the ballast effect shows how all cell components are interconnected via volume changes;

– the derived performance indices for regulatory modules under variable volume conditions are more realistic comparatively with those derived from the constant-volume models.

It is also to be observed that, in a variable-volume model all species (individual or lumped) have to be included in the model, because all contribute to the volume dynamics. In such a manner, the number of rate constants increases leading to a corresponding increase in the identification effort. When comparing two regulatory modules under variable-volume conditions, the comparison terms must be kept constant (i.e. the same cell 'ballast' and environmental conditions).

When a cell reaches a critical size and a certain level of the surface-area-to-volume ratio, the division phase begins, lasting the last 20 % of cell cycle. Over this phase (not analysed here), specialized proteins constrict the cell about its equator, thus leading to cell division. The duplicated content is thus partitioned, more or less evenly, between daughters cells. To model such a phase, supplementary terms must explicitly account for membrane dynamics.[56]

Stationary cell growth conditions lead to the nonlinear set of equations:

$$\left(\frac{dc_j}{dt}\right)_s = \left(\frac{1}{V}\frac{dn_j}{dt}\right)_s - D_s c_{js} = h_{js}(c_s, k) = 0;\ j = 1,\ldots,n_s;$$

$$D_s = \left(\frac{RT}{\pi}\right)\sum_j^{n_s}\left(\frac{1}{V}\frac{dn_j}{dt}\right)_s. \qquad (6)$$

As the (RT/π) term is known from the initial condition (5), the time-invariant set (6) can be used to estimate $\boldsymbol{k}$-s (and some unobservable $\tilde{c}_{js}$). Estimation rule is based on fulfilment of QSS station-

ary condition (6), system invariants (mass balance equations[2]), and on imposing optimum regulatory criteria, such as:

$$[\hat{\boldsymbol{k}}, \hat{\tilde{c}}_s] = \arg Min(\tau_P),$$

subjected to:

$$\boldsymbol{h}_{js}(c_s, \boldsymbol{k}) = 0;\ j = 1,\ldots,n_s;\ \text{(eq. 6)}$$

$$[\hat{\boldsymbol{k}}, \hat{\tilde{c}}_s] > 0;$$

$$\sum_{i=0}^{n}[G(P_i)] = \text{constant};\quad \sum_{i=0}^{n}[G(PP_i)] = \text{constant};$$

$$\sum_{i=0}^{n}[L(O_i)] = \text{constant, etc.}$$

$$[L]_{active}/[L_{total}] = \tfrac{1}{2}\text{, etc.}$$

$$\left(\sum_{j}^{all} c_j\right)_{cell} = \left(\sum_{j}^{all} c_j\right)_{env} = \text{constant}, \qquad (7)$$

(where τ_P has been evaluated by applying a 10 % $[P]_s$ impulse perturbation and by determining the recovering time with a tolerance of 1 % $[P]_s$). To estimate $[\hat{\boldsymbol{k}}, \hat{\tilde{c}}_s]$, other regulatory global properties can also be used together with the constraints of (7). It is also to be mentioned here the special case of estimating rate constants in rapid buffering reactions, of type:

$$L + O \underset{k_{diss}}{\overset{k_{bind}}{\rightleftharpoons}} LO$$

$$K_{LO} = \frac{k_{bind}}{k_{diss}} = \frac{c_{LO,s}}{c_{L,s}c_{O,s}};$$

(constant volume equilibrium)

$$K_{LO} = \frac{\boldsymbol{k}_{bind}}{\boldsymbol{k}_{diss}} = \frac{c_{LO,s}\left(1 + \dfrac{D_s}{k_{diss}}\right)}{c_{L,s}c_{O,s}};$$

(variable volume equilibrium, eq. 6). (8)

As discussed by *Kholodenko* et al.,[25] fast buffering reactions are close to equilibrium and have little effect on metabolic control coefficients. As a consequence, rate constans of such rapid reactions are much higher than those of the core synthesis and dilution rates. To reduce the size of the vector of unknowns in (7), large values of $\boldsymbol{k}_{diss} \gg D_s$ can be postulated, while K_{LO} becomes closer to its constant-volume value in (8).

Assuming a known nominal set of species stationary concentrations $\boldsymbol{c}_s$, differentiation of the steady-state conditions (6) lead to evaluate the state sensitivity vs. nutrient levels [i.e. $S^{j}_{Nut_j} = (\partial c_i / c_{Nut_j})_s$] from:

$$\boldsymbol{h}_{js}(c_s, c_{Nut,s}, \boldsymbol{k}) = 0;\quad j = 1,\ldots,n_s, \Rightarrow \qquad (9)$$

$$\left[\frac{\partial h_i(\boldsymbol{c}, \boldsymbol{c}_{Nut}, \boldsymbol{k})}{\partial c_j}\right]_s \left[\frac{\partial c_j}{\partial \boldsymbol{c}_{Nut}}\right]_s + \left[\frac{\partial h_i(\boldsymbol{c}, \boldsymbol{c}_{Nut}, \boldsymbol{k}}{\partial \boldsymbol{c}_{Nut}}\right]_s = 0.$$

In the previous relationship, the Jacobian $J_c = \left[\partial \boldsymbol{h}_i / \partial c_j\right]_s$ is evaluated numerically at the cell-system stationary-state.

Module linking rules

Cell regulatory networks, and in particular protein synthesis regulation, are poorly understood. The modular approach of studying the regulation path, accounting for its structural and functional organization, seems to be a promising route to be followed. Because a limited number of module types exists, individual regulatory modules can be separately analysed and checked for efficiency in conditions that mimic the stationary and perturbed cell growing conditions. A module is as efficient as the stationary performance indices are more favourable and sensitivity to perturbations is lower. Then, they are linked accordingly to certain rules to mimic the real metabolic process, by ensuring the overall network efficiency, system homeostasis, and protein functions. Module linking rules are not fully established. It seems that modular network is hierarchically organised, and includes a large number of compounds with strong interactions inside a module and weaker interactions among modules, so that the whole cell system efficiency can be adjusted.

For instance, *Sewell* et al.[9] offered some module linking suggestions, accounting for automated system principles. Thus, module interactions can be defined kinetically, implying common reactions and species and common input-output components. Different module types can be identified separately based on stationary compound concentrations, and by optimising a regulatory efficiency index. When the network is constructed, linking reactions between modules are set to be relatively slow comparatively with the module core reactions. In such a manner, individual modules remain fully regulated, while the assembly efficiency is adjusted by means of linking reaction and intermediate levels. To preserve the individual regulatory capacity, the magnitude of linking reactions would have to decline as the number of linked modules increases.[9]

When linking regulatory modules, the main questions arise on the connectivity mechanism and on the cooperative vs. uncooperative way of which

proteins interact over the parallel/consecutive metabolic path. In spite of an apparent 'competition' for nutrient consumption, protein synthesis is a closely cooperative process, due to the specific role and function of each protein inside the cell.[3,12] Protein interactions are very complex, being part of the cell metabolism and distributed over regulatory network nodes. There are many nodes with few connections among proteins and a small, but still significant, number of nodes with many proteic interactions.[4] These highly connected nodes tend to be essential to an organism and to evolve relatively slowly. At a higher level, protein interactions can be organised in 'functional modules', which reflect sets of highly interconnected proteins ensuring certain cell functions. Specific proteins are involved in nutrient permeation (permeases), in metabolite synthesis (metabolases), or in gene production (polymerases). In general, experimental techniques can point-out molecular functions of a large number of proteins, and can identify functional partners over the metabolic pathways. Moreover, protein associations can ensure supplementary cell functions. For instance, enzyme associations lead to the well-known 'metabolic channelling' (or tunnelling) process, that ensures an efficient intermediate transfer and metabolite consecutive transformation without any release into the cell bulk phase.[87]

It results that an effective module linking strategy has to ensure the cell-functions of individual proteins and of protein associations over the metabolic synthesis network. As a general observation, even not presenting common reactions, the modules are anyway linked through the cell volume (to which all cell species contribute) and due to some signalling intermediates controlling the regulatory network construction. The variable cell-volume model is able to account for such cell regulatory characteristics. From the theoretical point of view, several linking alternatives can be investigated, such as cooperative vs. uncooperative linking. In a cooperative linking, common species (or reactions) are used for a cross-control (or cross-catalysis) of the synthesis reactions. Thus, the system stability is strengthened, while species inter-connectivity is increased leading to a better treatment of perturbations.

Step-by-step building-up of a regulatory modular structure – An example

The present study aims to exemplify, at a generic level and by using the variable cell-volume model, various alternatives to link regulatory modules, and to simulate protein interrelations during metabolic synthesis. It also investigates the feasibility of a module linking strategy to elaborate a regulatory network, and the way in which a step-by-step increase in the network complexity is reflected by the regulatory efficiency of the whole path.

To exemplify such an analysis in a simple way, one considers a hypothetical cell, similar to *E. coli*, in an equilibrated growth at a constant temperature, with a cell cycle period of t_c = 100 min, and a quasi-constant logarithmic growing rate of $D_s = \ln(2)/t_c$. Concentrations of lumped nutrients in the environment have been considered constant, i.e. $c_{NutG,s} = 3\times10^6$ nmol L^{-1}, $c_{NutP,s} = 3\times10^8$ nmol L^{-1} (similarly to *Morgan* et al.[56]). As only a few number of individual species are accounted in the model, the cell 'ballast' is mimicked by adopting high levels for metabolite concentrations, i.e. $\sum_j c_{MetP_{j,s}} = 3\times10^8$ nmol L^{-1}; $\sum_j c_{MetG_{j,s}} = c_{NutG,s} + c_{NutP,s} - \sum_{j\neq MetG_j}^{cell} c_{js} \approx 10^6$ n mol L^{-1}. For the genes, proteins, and other intermediates, the stationary concentrations are displayed in the footnote of Table 1. Species in the environment have been lumped, and only two groups (NutP and NutG) have been considered. Not to complicate the model, only protein concentration diminishment due to the cell-growth dilution effect has been considered, the protein degradation being neglected. Once the network construction methodology elaborated, more detailed modular representations can include mechanism extensions in the analysis.

The network construction starts with considering the simplest system, i.e. a cell which includes only one module for P-protein synthesis, in an unregulated alternative *G(P)0.* In such a simplified representation, a simple reaction, catalysed by P, is used to kinetically describe the membrane permeation and nutrient transformation into metabolites (see Fig. 3). The considered stationary concentrations for the P/G pair are c_{Ps} = 1000 nmol L^{-1} and c_{Gs} = 1 nmol L^{-1}. The value of 1 nmol L^{-1} for gene concentration in a newborn cell results from applying the formula:[56]

$$\text{concentration} = \frac{\text{no. of copies / cell}}{N_A \times V_{cyt,0}} \qquad (10)$$

(where N_A is the Avogadro number). As for an *E. coli* cell, with an approximate $V_{cyt,0} = 1.66 \times 10^{-15}$ L (see *Yang* et al.[5]), it results for c_{Gs} a value of $1/(6.022 \times 10^{23})(1.66 \times 10^{-15}) = 1$ nmol L^{-1} (i.e. 10^{-9} mol L^{-1}). The other species of the cell are included in the cell 'ballast' by adopting high levels for metabolites (~10^8 n mol L^{-1}). Then, the module is improved by adding a regulatory element. For instance, by including a simple effector and a buffering reaction, one obtains the module *G(P)1* (Fig. 2 and Fig. 3, $c_{Gs} = c_{GPs} = ½$ nmol L^{-1}). Rate constants are then estimated from the QSS condition (6).

Table 1 – *Stationary (S^j_{NutP}) and dynamic (τ_j) regulatory effectiveness of (un)coupled regulatory modules [$S^j_{NutP} = (\partial c_j/\partial c_{NutP)s}$ = species j level sensitivity to NutP; τ_j = species j recovering time to QSS, with a 1% tolerance, after a ± 10 % $c_{P1,s}$ impulse perturbation].*

Species	G(P)0		G(P)1		G1(P1)0+G2(P2)0		G1(P1)1+G2(P2)1		G1(P1P1)1+G2(P2P2)1		G1(P1)1+G2(P2)1+G3(P3)1	
	τ_j (min)	S^j_{NutP}	τ_j (min)	S^j_{NutP}	τ_j (min)	S^j_{NutP}	τ_j (min)	S^j_{NutP}	τ_j (min)	S^j_{NutP}	τ_j (min)	S^j_{NutP}
Met_{G1-G3}	NG	-9.9×10^{-3}	NG	-9.9×10^{-3}	NG	-4.9×10^{-3}	NG	-4.9×10^{-3}	NG	-4.9×10^{-3}	384.1	-3.3×10^{-3}
Met_{P1-P3}	NG	9.9×10^{-3}	NG	9.9×10^{-3}	NG	4.9×10^{-3}	NG	4.9×10^{-3}	NG	4.9×10^{-3}	NG	3.3×10^{-3}
P1	166.2	-4.9×10^{-6}	131.7	-3.9×10^{-6}	318.1	-4.9×10^{-6}	88.0	-3.9×10^{-6}	76.6	-3.3×10^{-6}	347.9	-3.9×10^{-6}
P2	-	-	-		502.2	-4.9×10^{-7}	319.4	-3.9×10^{-7}	212.7	-3.3×10^{-7}	NG	-3.9×10^{-7}
P3	-	-	-		-	-	-		-		NG	-3.9×10^{-6}
G1	NG	-6.6×10^{-9}	92.0	-2.3×10^{-9}	502.2	-6.6×10^{-9}	322.1	-2.3×10^{-9}	224.7	-1.6×10^{-9}	355.5	-2.3×10^{-9}
G2	-	-	-	-	502.2	-6.6×10^{-9}	288.8	-2.3×10^{-9}	NG	-1.6×10^{-9}	456.8	-2.3×10^{-9}
G3	-	-	-	-	-	-	-	-	-	-	456.5	-2.3×10^{-9}
G1P1	-	-	92.0	-4.3×10^{-9}	-	-	394.1	-4.3×10^{-9}	-	-	584.9	-4.3×10^{-9}
G2P2	-	-	-	-	-	-	425.5	-4.3×10^{-9}	-	-	581.6	-4.3×10^{-9}
G3P3	-	-	-	-	-	-	-	-	-	-	581.9	-4.3×10^{-9}
G1P1P1	-	-	-	-	-	-	-	-	320.8	-4.9×10^{-9}	-	-
G2P2P2	-	-	-	-	-	-	-	-	370.1	-4.9×10^{-9}	-	-
P1P1	-	-	-	-	-	-	-	-	173.4	-6.6×10^{-11}	-	-
P2P2	-	-	-	-	-	-	-	-	318.1	-6.6×10^{-11}	-	-
AVG[a]	41.5	-	55.9	-	304.1	-	229.7	-	169.6	-	340.8	-
STD[b]	71.9	-	57.6	-	224.6	-	162.4	-	136.0	-	224.2	-
$\max(\mathrm{Re}(\lambda_j))$	-D (stable)	-	-D (stable)	-	-D (stable)	-	-D (stable)	-	-D (stable)	-	-D (stable)	-

a) average of τ_j / min; b) standard deviation of τ_j / min; λ_j = QSS-system Jacobian eigenvalues; $\max(\mathrm{Re}(\lambda_j)) < 0$ indicates a stable QSS; NG = negligible; stationary concentrations are: $[c_{NutG,s}, c_{NutP,s}, \Sigma_j c_{MetG_j,s}, \Sigma_j c_{MetP_j,s}, c_{P1,s}, c_{P2,s}, c_{P3,s}, c_{G1,G2,G3,s}, c_{G1P1,s}, c_{G2P2,s}, c_{G3P3,s}, c_{G1P1P1,s}, c_{G2P2P2,s}, c_{P1P1,s}, c_{P2P2,s}] = [3\times10^6, 3\times10^8, \sim10^6, 3\times10^8, 10^3, 10^2, 10^3, 1(\text{or } ½), 0(\text{or } ½), 0(\text{or } ½), 0(\text{or } ½), 0(\text{or } ½), 0(\text{or } ½), 10^{-2}, 10^{-2}](\text{nmol L}^{-1})$; $\Sigma_j c_{MetG_j} = c_{NutG} + c_{NutP} - \sum_{j\neq MetG_j}^{cell} c_j$; life cycle t_c = 100 min; cell-volume logarithmic growing rate $D = \ln 2/t_c$; 's' index refers to the QSS.

Based on that, dynamic regulatory effectiveness [τ_j; AVG(τ_j); STD(τ_j)], stationary regulatory effectiveness (S^j_{NutP}), and system stability are evaluated [by integrating eq. (4-5) after an impulse perturbation, or by solving the set (9) for $j = 1,\ldots, n_s$]. The results, presented in Table 1 and Fig. 3, lead to several conclusions.

i) Both systems are stable [$\max(\mathrm{Re}(\lambda_j)) = -D < 0$]. Each system recovers after a dynamic perturbation in c_{Ps} (Fig. 3, down).

ii) Systems are unregulated or very poorly regulated, the AVG(τ_j) being large (41 min and 56 min), some species presenting recovering times τ_j close or larger than the cell cycle period t_c = 100 min.

iii) The dynamic regulatory index for the key-species P improves from τ_P = 166.2 min for G(P)0 to τ_P = 131.7 min for G(P)1, that is with a factor of 1.26. Consequently, it is expected that an allosteric gene G activity control of type G(P)*n* to

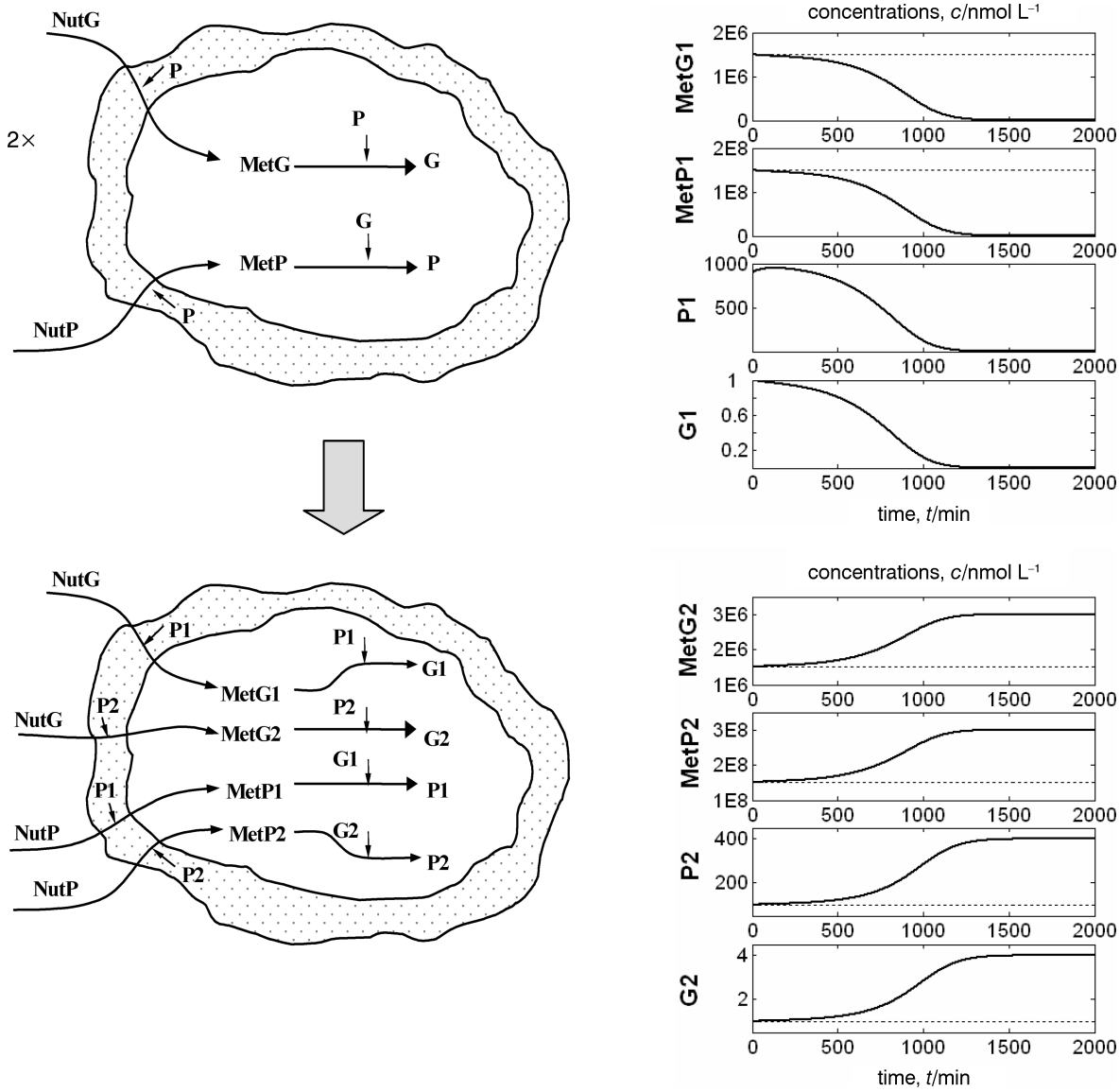


Fig. 4 – *Un-cooperative coupling of two modules [G1(P1)0] + [G2(P2)0] for P1 and P2 protein synthesis (left). Degenerative evolution of cell components after a – 10 % impulse perturbation in* $[P1]_s$ *= 1000 nmol* L^{-1} *at* $t = 0$ *(right).*

roughly improve ($n \times 1.26$) times the P.I. (see eq. 3), leading to τ_P values smaller than t_c. It is also expected that another effector and a more effective regulatory unit to lead to even much better P.I. and smaller τ_P values.

iv) The dynamic perturbations affect rather species present in small amounts inside the cell, while recovering times for major species (e.g. metabolites MetP, MetG) are negligible (Table 1, Fig. 3).

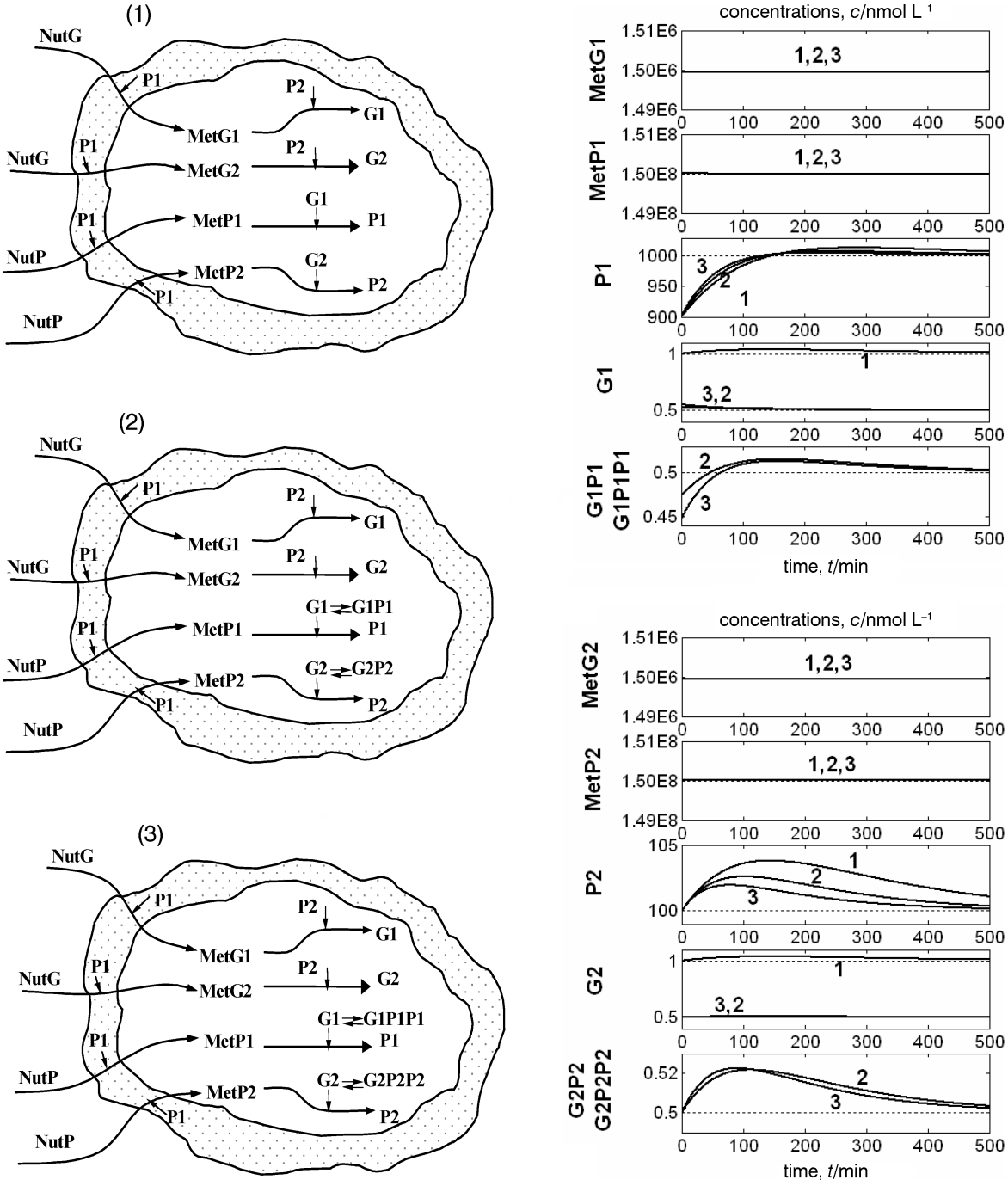
v) By increasing the complexity from G(P)0 to G(P)1 (with one effector and a buffering reaction), the whole module recovers slower after an impulse perturbation (i.e. the resulted AVG is higher). At the same time, species interconnectivity increases (i.e. STD becomes lower), leading to a better regulatory index for the target P protein synthesis.

vi) The stationary regulatory indices (sensitivities vs. nutrients) follow the same trend. For instance, the sensitivity of P vs. NutP is decreasing from $|S^{P}_{NutP}| = 4.9\times10^{-6}$ to $|S^{P}_{NutP}| = 3.9 \times 10^{-6}$, i.e. with the same 1.26 improvement factor.

In the second stage of the analysis, one links two modules *G1(P1)0 + G2(P2)0,* ensuring regulation of two protein synthesis (P1 and P2), in an uncooperative disconnected way (Fig. 4). For this hypothetic system, synthesis of P1/G1 and P2/G2 from metabolites is realised with any inference between modules (c_{P1s} = 1000 nmol L^{-1}, c_{P2s} = 100 nmol L^{-1}, c_{G1s} = 1 nmol L^{-1}, c_{G2s} = 1 nmol L^{-1}). The only connection is due to the common cell volume to which both protein syntheses contribute. If one checks this system for stability, by applying a ±10 % impulse perturbation in $c_{P1,s}$, it results an un-

stable system, evolving toward the decline and disappearance of one of the proteins (i.e. those presenting the lowest synthesis rate, Fig. 4). Consequently, the homeostasis condition is not fulfilled, the cell functions cannot be maintained, and the disconnected protein synthesis results as a unfeasible and less plausible linking alternative.

The cooperative connection of [*G1(P1)0 + G2(P2)0*] modules in Fig. 5 [sub-case (1)] ensures specific functions of proteins, i.e. P1 lumps, both, the permeases and metabolases, while P2 is a polymerase. For comparison, the same two-module system has been improved by adding simple effectors for gene activity control. In the cooperatively linked system [*G1(P1)1 + G2(P2)1*], [see Fig. 5, sub-case (2), and Fig. 2], the effectors P1 and P2 act in two buffering reactions, G1+P1 ⇔ G1P1, and G2+P2 ⇔ G2P2, with the stationary states $c_{G1,s}$ = $c_{G1P1,s}$ = 1/2 nmol L^{-1}, and $c_{G2,s}$ = $c_{G2P2,s}$ = 1/2 nmol L^{-1}. In the coupled system [*G1(P1P1)1 +*



F i g . 5 – *Cooperative coupling of two regulatory modules (up): [G1(P1)0] + [G2(P2)0] (case 1); [G1(P1)1] + [G2(P2)1] (case 2); [G1(P1P1)1] + [G2(P2P2)1] (case 3). Cell components recover after a – 10% impulse perturbation in* $[P1]_s$ *= 1000 nmol L*$^{-1}$ *at t = 0 (down).*

G2(P2P2)1], [see Fig. 5, sub-case (3), and Fig. 2], the effectors are the dimers P1P1 and P2P2, acting in two buffering reactions, G1 + P1P1 ⇔ G1P1P1, and G2 + P2P2 ⇔ G2P2P2, with the stationary states $c_{G1,s} = c_{G1P1P1,s} = 1/2$ nmol L^{-1}, and $c_{G2,s} = c_{G2P2P2,s} = 1/2$ nmol L^{-1}. Rate constants have been estimated from the stationary concentrations ($c_{j,s}$ in Table 1, adopted from the *E. coli* cell) and by imposing regulatory optimal characteristics given by the criterion (7). The $k_{diss} >> D_s$ in eq. (8) has been adopted as being ca. 10^7 D_s, while system optimization with criterion (7) leads to small values for $c_{P1P1,s}$ and $c_{P2P2,s}$ (i.e. the active parts of dimers). Such a result is in agreement with the reported values of *Sewell* et al.[9] and *Yang* et al.[5] for a one-module, constant-volume system. The system stability and dynamic regulatory characteristics have been determined by studying the QSS-recover after a ± 10 % $c_{P1,s}$ impulse perturbation (Fig. 5, right). The results, presented in Table 1, reveal the following aspects concerning the systems [*G1(P1)0 + G2(P2)0*], [*G1(P1)1 + G2(P2)1*], and [*G1(P1P1)1 + G2(P2P2)1*].

i) All three systems (curves 1-3 of Fig. 5) are stable [$\max(\mathrm{Re}(\lambda_j)) = -D < 0$]. Each system recovers after a dynamic perturbation in $c_{P1,s}$ (Fig. 5, right). It results that the cooperative module linking, by preserving specific functions of each protein inside the cell, is a viable solution ensuring system homeostasis.

ii) The system is as better regulated as the effector is more effective, the AVG(τ_j) being reduced from 304 min, to 229 min, and then to 169 min for the previously mentioned systems. These values are larger than the cell cycle period $t_c = 100$ min, but an allosteric regulation can drastically improve this index.

iii) The dynamic regulatory index for the key-species P is significantly improved from $\tau_P =$ 318 min [case *G1(P1)0 + G2(P2)0*] to $\tau_P = 88$ min [for *G1(P1)1 + G2(P2)1*], or to $\tau_P = 77$ min [case *G1(P1P1)1 + G2(P2P2)1*], that is with a factor of 3.6-4.1. It is expected that, by using an allosteric gene activity control of type *G(P)n* or *G(PP)n*, this index to be still improved *n* times (see eq. 3), leading to τ_P values much smaller than t_c. Thus, the use of efficient effectors and multiple regulation units can improve very much the dynamic index, in the following ranking: G(P)*n* < G(PP)*n* < M(P)*n*.

iv) Dynamic perturbations affect rather species present in small amounts inside the cell, while recovering times for other species (e.g. metabolites MetP, MetG) are negligible (Table 1, Fig. 5).

v) By adding to [*G1(P1)0 + G2(P2)0*] regulation effectors of type *G(P)1*, or even more effective of type *G(PP)1,* the whole system recovers faster after an impulse perturbation (i.e. the resulted AVG is lower), but also the species interconnectivity increases (i.e. STD becomes lower), with a positive effect on the target P-synthesis regulation.

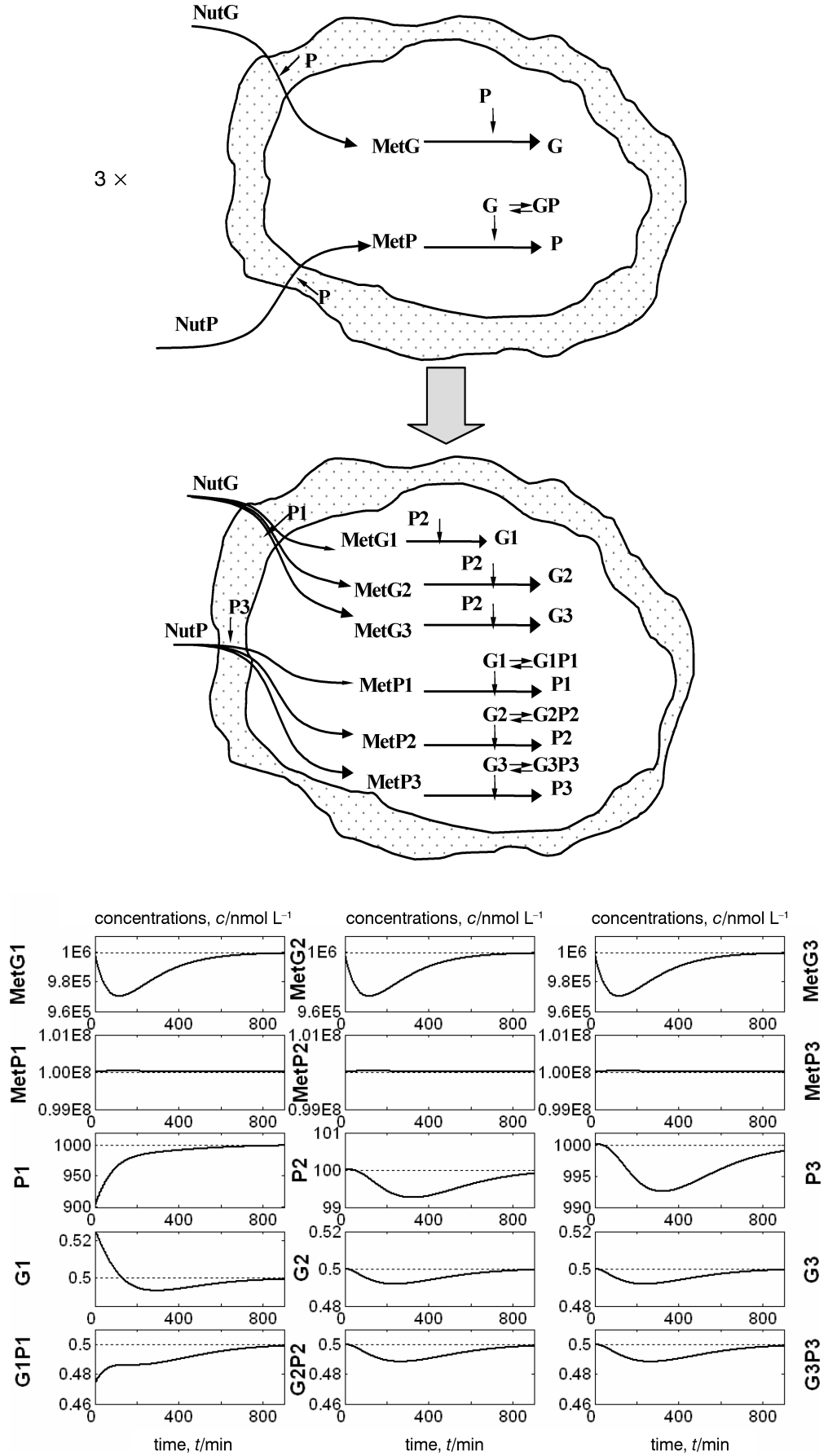
vi) The stationary regulatory indices (sensitivities of states vs. nutrients) follow the same trend. For instance, the sensitivity of P vs. NutP is decreasing from $|S^{P}_{NutP}| = 6.6\times10^{-9}$, to 2.3×10^{-9}, and 1.6×10^{-9}, and that is with an improvement factor of 2.8 times for *G(P)1* vs. *G(P)0*, and of 4.1 times for *G(PP)1* vs. *G(P)0*.

The regulatory network design procedure can be continued, by accounting for a new protein (and synthesis module). For instance, in the simplified representation of Fig. 6, P1 and P3 lumps permease and metabolase enzymes, which ensure nutrient import inside the cell, and transformation in gene-metabolites (MetG1-MetG3) and protein-metabolites (MetP1-MetP3) respectively. Proteic unit P2 lumps polymerases able to catalyse the gene production. If one considers the simplest effector case, the resulted cell includes three modules *G1(P1)1 + G2(P2)1 +G3(P3)1,* which regulate the synthesis of P1, P2 and P3, in a cooperative interconnected way of preserving protein functions. To exemplify, one considers the stationary concentrations of Table 1 (footnote), with $c_{P1,s} = 1000$ nmol L^{-1}, $c_{P2,s} = 100$ nmol L^{-1}, $c_{P3,s} = 1000$ nmol L^{-1}.

The regulatory effectors are the proteins P1-P3 themselves, while the buffering reactions are of type G(P)1, that is G1+P1 ⇔ G1P1, G2+P2 ⇔ G2P2, G3+P3 ⇔ G3P3, with stationary states $c_{G1,s} = c_{G1P1,s} = 1/2$ nmol L^{-1}, $c_{G2,s} = c_{G2P2,s} = 1/2$ nmol L^{-1}, and $c_{G3,s} = c_{G3P3,s} = 1/2$ nmol L^{-1}. Rate constants have been estimated from the stationary concentrations and by imposing optimal characteristics to the buffering reactions [i.e. $\boldsymbol{k}_{diss} >> D_s$ in eq. (8)]. The system stability and dynamic regulatory characteristics have been determined by studying the QSS-recover after a ± 10 % $c_{P1,s}$ impulse perturbation (Fig. 6, down), while the stationary regulatory characteristics from solving the set eq. (9). The results, presented in Table 1, reveal the following aspects.

i) The system is stable [$\max(\mathrm{Re}(\lambda_j)) = -D < 0$], and recovers after a dynamic perturbation (Fig. 6, down). It results that the cooperative module linking, by preserving the specific functions of each protein inside the cell, remains the only viable solution ensuring homeostasis.

ii) The dynamic regulatory index AVG(τ_j), τ_{P1}, and species connectivity STD-index become weaker comparatively to the module *G1(P1)1+G2(P2)1* case. One can supposes that, the decline in the dynamic regulatory indices is the 'price' to be paid by the increase in the cell complexity, leading to an in-



F i g . 6 – *Cooperative coupling of three regulatory modules [G1(P1)1] + [G2(P2)1 + [G3(P3)1] for P1, P2, P3 protein synthesis (up). Cell components recover after a – 10 % impulse perturbation in* $[P1]_s$ *= 1000 nmol* L^{-1} *at t = 0 (down).*

crease difficulty in realising a non-scattered recovering for a larger number of species. However, it is expected that, by using an allosteric gene activity control of type *G(P)n* or *G(PP)n*, coupled with a cascade control involving *M(P)n* units, these dynamic indices to be drastically improved (see eq. 3), thus τ_j becoming much smaller than t_c.

iii) Dynamic perturbations affect rather species present in small amounts inside the cell, while recovering times for other species are negligible (Table 1, Fig. 6).

iv) The stationary regulatory indices (sensitivities of states vs. nutrients) remain unchanged, as for the two-modules or one G(P)1 module. Then, the increase in the network complexity by adding modules of the same type seems not to affect the stationary effectiveness of modules but only the dynamic one.

v) As another positive effect of a larger network, the system robustness and stability increases with the cell content, due to the 'ballast' and 'inertial' effects, propagated through the cell-volume to which all-inner species contributes.

The same rule to develop the regulatory network can be continued, step-by-step accounting for new proteins and their specific functions (for instance, including individual metabolases for distinct metabolite synthesis). By extending this rule, other metabolic syntheses can be modelled in a similar way.

Conclusions

The kinetic modular approach, designed to reproduce the regulation mechanism of each protein synthesis, can be used to build-up metabolic regulation networks and to model cooperative interplays of proteins. In the proposed variable-volume modelling framework, modules are individually analysed and linked one-by-one, in a cooperative manner that ensures specific functions of each protein inside the cell. As more proteins are accounted, the system becomes more complex and more robust to environmental changes. The alternative, that is an uncooperative way of linking, in which modules are only connected by means of the common cell volume, is proved as not being a viable alternative: the stationary growth becomes impossible, the homeostasis condition not fulfilled, and the whole system evolves towards disappearance of some proteins (presenting the lowest synthesis rate).

In spite of an apparent 'competition' for nutrient consumption, protein synthesis is a closely cooperative process, in which syntheses are mutually assisted due to the specific protein functions inside the cell.[3,88] Protein complex interactions are part of the cell metabolism and distributed over regulatory network nodes.

When linking regulatory modules, species connectivity, system stability, and regulatory performance indices vs. stationary and dynamic perturbations are evaluated and analysed. Thus, after each added module, the network regulatory characteristics are used in adjusting the model parameters to fulfil the global properties of the whole system. For the same module type, it appears that an increase in the network complexity seems to affect rather the dynamic regulatory indices than the stationary ones. Dynamic indices can be adjusted by considering an allosteric control of the enzyme activity, by means of a suitable effector, and in a structured cascade control. The improvement factor is ca. 1.3-2 (or even higher) for each buffering regulatory reaction added in the allosteric schema. Multiple copies of effector molecules, allosterically binding the catalyst in negative feedbacks, improve the regulation effectiveness, species inter-connectivity, and network flexibility; species recover less disparately after a perturbation and in a shorter time. Performance indices also depend on the type of the regulatory module used for protein synthesis, improving in the order *G(P)n < G(PP)n < G(P)n; M(P)n'*. Thus, the highly interconnected and cooperative linking way becomes favourable for the whole construction, ensuring a fast and economic way to cope with perturbations over the cell growth.

Dynamic perturbations affect rather species (intermediates) present in small amounts inside the cell than those existing in large amounts. However, the levels of intermediate species are responsible for optimising the whole network construction efficiency. Global optimization criteria are used to estimate such network parameters of the model. The cell system complexity and size are responsible for the 'ballast' and 'inertial' effect in smoothing dynamic perturbations from environment.

The variable-volume modelling framework appears to be more realistic, allowing an accurate evaluation of the network regulatory indices, but also the study of complex species inter-connectivities (direct and indirect), effector's relative importance, and of dynamic primary and secondary perturbations (propagated by the volume dynamics).

Even if the module-linking rule has been exemplified by accounting only for a generic system (similar to *E. coli*), particularisations to various metabolic syntheses can easily be realised.[1,2,5]

Functional modules, including several regulatory-effectors and a functional-base linking strategy, can lead to simulate complex metabolic paths, opening the way of developing multi-compartmented simulation platforms of the cell.

Nomenclature

$\boldsymbol{A}$ – monodromy matrix (defined in eq. 2)
A_{unsync} – stationary regulatory index
a_i – constants defined in eq. (3)
b_i – constants in eq. (1)
c_j – species j concentration
D – cell-content dilution rate (i.e. cell-volume logarithmic growing rate)
d_i – constants in eq. (1)
$\boldsymbol{h}$ – model function vector
$\boldsymbol{I}$ – identity matrix
$\boldsymbol{J_c} = \partial \boldsymbol{h}/\partial \boldsymbol{c}$ – system Jacobian matrix
K – equilibrium constant
k_j – rate constant of reaction j
L_i – component i at which regulatory element acts
n_j – amount of species j (number of moles)
N_A – the Avogadro number
n, n_i – number of species O_i (effectors)
n_s – number of species inside the cell
O_i – regulatory element i (effector)
R – universal gas constant
R_D – dynamic regulatory (recovering) index
R_{ss} – relative stationary regulatory index
r_j – species j reaction rate
$S^{\text{state}}_{\text{perturb}}$ – steady-state sensitivity coefficient vs. perturbations
T – period, or absolute temperature
t – time
t_c – cell cycle period
V – cell system volume
x_j – independent variable j
y_i – dependent variable i

Greeks

λ_i – i-th eigenvalue
π – osmotic pressure
τ_j – species j recovering time

Index

cyt – cytoplasma
diss – dissociation
env – environment
n – nominal value
0 – initial
perturb – perturbed
s – (quasi-)steady-state
syn – synthesis

Superscripts

~ – unobservable variable
^ – estimated value

Abbreviations

arg – argument
AVG – average
G – gene
M – mRNA
Max – maximum
Met – metabolite
Min – minimum
Nut – nutrient
P – protein
P.I. – regulatory performance index
QSS – quasi-steady-state
Re – real part
STD – standard deviation
[.] – concentration

References

1. *Maria, G.*, Chemical and Biochemical Engineering Quarterly **17** (2003) 99.
2. *Maria, G.*, Chemical and Biochemical Engineering Quarterly **18** (2004) 195.
3. *Sorribas, A., Savageau, M. A.*, Math Biosci **94** (1989) 161.
4. *Gabaldon, T., Huynen, M. A.*, Cell Mol Life Sci **61** (2004) 930.
5. *Yang, Q., Lindahl, P., Morgan, J.*, J. theor. Biol. **222** (2003) 407.
6. *Stadtman, E. R., Chock, P. B.*, Proc. Natl. Acad. Sci. USA **74** (1977) 2761.
7. *Ferrell, J. E. Jr.*, Trends Biochem. Sci. **22** (1997) 288.
8. *Kholodenko, B. N.*, Eur. J. Biochem. **267** (2000) 1583.
9. *Sewell, C., Morgan, J., Lindahl, P.*, J. theor. Biol. **215** (2002) 151.
10. *McAdams, H. H., Arkin, A.*, Ann. Rev. Biophys. Biomol. Struct. **27** (1998) 199.
11. *Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E. D.*, Nature **420** (2002) 190.
12. *Heinrich, R., Schuster, S.*, The regulation of cellular systems, Chapman & Hall, New York, 1996.
13. *Kacser, H., Burns, J. A.*, The control of flux, Proc. Symp. Soc. Exper. Biol. Rate Control Biol. Proc., Cambridge Univ. Press, Cambridge, 1973, pp. 65-104.
14. *Heinrich, R., Rapoport, T. A.*, Eur. J. Biochem. **42** (1974) 89.
15. *Burns, J. A., Cornish-Bowden, A., Groen, A. K., Heinrich, R., Kacser, H., Porteous, J. W., Rapoport, S. M., Rapoport T. A., Stucki, J. W., Tager, J. M., Wanders, R. J. A., Westerhoff, H. V.*, Trends Biochem. Sci. **10** (1985) 16.
16. *Markus, M., Hess, B.,* Control of metabolic oscillations: unpredictability, critical slowing down, optimal stability and hysteresis, in: Cornish-Bowden, A., Cárdenas, M. (Eds.), Control of metabolic processes, Plenum Press, New York, 1990, pp. 303-313.

17. *Acerenza, L., Sauro, H. M., Kacser, H.,* Control analysis of time-dependent metabolic systems, J. theor. Biol. **137** (1989) 423.
18. *Schilling, C. H., Letscher, D., Palsson, B. O.*, J. Theor. Biol. **203** (2000a) 229.
19. *Schilling, C. H., Edwards, J. S., Letscher, D., Palsson, B. O.*, Biotechnology and Bioengineering **71** (2000b) 286.
20. *Stephanopoulos, G., Simpson, T. W.*, Chem. Eng. Sci. **52** (1997) 2607.
21. *Pfeiffer, T., Sanchez-Valdenbro, I., Nuno, J. C., Montero, F., Schuster, S.*, Bioinfomatics **15** (1999) 251.
22. *Segre, D., Vitkup, D., Church, G. M.*, Proc Natl Acad Sci USA **99** (2002) 15112.
23. *Mahadevan, R., Edwards, J. S., Doyle, F. J.*, Biophysical Journal **83** (2002) 1331.
24. *Edwards, J. S., Ramakrishna, R., Palsson, B. O.*, Biotechnology and Bioengineering **77** (2002) 27.
25. *Kholodenko, B. N., Schuster, S., Garcia, J., Westerhoff, H. V., Cascante, M.*, Biochimica Biophysica Acta **1379** (1998) 337.
26. *Atauri, P., Curto, R., Puigjaner, J., Cornish-Bowden, A., Cascante, M.*, Eur. J. Biochem. **265** (1999) 671.
27. *Szedlacsek, E. S., Cardenas, M. L., Cornish-Bowden, A.*, Eur. J. Biochem. **204** (1992) 807.
28. *Szedlacsek, E. S., Aricescu, A. R., Havsteen, B. H.*, J. theor. Biol. **182** (1996) 341.
29. *Domach, M. M., Leung, S. K., Cahn, R. E., Cocks, G. G., Shuler, M. L.*, Biotechnology and Bioengineering **26** (1984) 203.
30. *Shuler, M. L.*, Chemical Engineering Problems in Biotechnology, AIChE, New York, 1989.
31. *Browning, S. T., Shuler, M. L.*, Biotechnology and Bioengineering **76** (2001) 187.
32. *Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J. C.*, Bioinformatics **15** (1999) 72.
33. *Tomita, M.*, Trends in Biotechnology **19** (2001) 205.
34. *Kinoshita, A., Nakayama, Y., Tomita, M.*, In silico analysis of human erythrocyte using E-Cell system, 2nd Int. Conf. on Systems Biology, California Institute of Technology, Nov. 4, 2001.
35. *Van Dien, S. J., Lidstrom, M. E.*, Biotechnology and Bioengineering **78** (2002) 296.
36. *Kauffman, K. J., Pajerowski, J. D., Jamshidi, N., Palsson, B. O., Edwards, J. S.*, Biophysical Journal **83** (2002) 646.
37. *Wiback, S. J., Palsson, B. O.*, Biophysical Jl. **83** (2002) 808.
38. *Riznichenko, G. Y.*, Mathematical models in biophysics: Teaching notes, Biological Faculty, Lomonosov Moscow State University, 2002; www.biophysics.org/btol/img/math-models.pdf.
39. *Wolf, J., Passarge, J., Somsen, O. J. G., Snoep, J. L., Heinrich, R., Westerhoff, H. V.*, Biophysical J. **78** (2000) 1145.
40. *Demin, O. V., Goryanin, I. I., Kholodenko, B. N., Westerfhoff, H. V.,* Molecular Biology **35** (2001) 1095.
41. *Hudder, B., Yang, Q., Bolfing, B., Maria, G., Morgan, J. J., Lindahl, P. A.*, Computational modeling of iron metabolism in mitochondria, NIST Conf. Mitochondrial Proteomics, Gaithersburg (USA), Sept. 17, 2002.
42. *Shen-Orr, S. S., Milo, R., Mangan, S., Alon, U.*, Nature Genetics **31** (2002) 64.
43. *Mizuno, T.*, DNA Res. **4** (1997) 161.
44. *McAdams, H. H., Shapiro, L.*, Science **269** (1995) 650.
45. *McAdams, H. H., Arkin, A.,* Proc. Natl. Acad. Sci. USA **94** (1997) 814.
46. *Norel, R., Agur, Z.*, Science **251** (1991) 1076.
47. *Hatzimanikatis, V., Lee, K. H., Bailey, J. E.*, Biotechnology and Bioengineering **65** (1999) 631.
48. *Obeyesekere, M. N., Zimmerman, S. O., Tecarro, E. S., Auchmuty, G.*, Bull. Math. Biol. **61** (1999) 917.
49. *Tyson, J. J., Novak, B.*, J. theor. Biol. **210** (2001) 249.
50. *Hallett, M. B.*, Perspect. Biol. Med. **33** (1989) 110.
51. *Goldbeter, A.*, Biochemical oscillations and cellular rhythms, Cambridge University Press, Cambridge, 1996.
52. *Morton-Firth, C. J., Bray, D.*, J. theor. Biol. **192** (1998) 117.
53. *Novak, B., Tyson, J. J.*, Proc. Natl. Acad. Sci. USA **94** (1997) 9147.
54. *Rensing, L., Meyer-Grahle, U., Ruoff, P.*, Chronobiology International **18** (2001) 329.
55. *Reijenga, K. A., Westerhoff, H. V., Kholodenko, B. N., Snoep, J. L.*, Biophysical Journal **82** (2002) 99.
56. *Morgan, J. J., Surovtsev, I. V., Lindahl, P. A.*, J. theor. Biology **231** (2004) 581.
57. *Shimizu, T. S., Bray, D.*, Computational cell biology – The stochastic approach, MIT Press Math6X9, Cambridge, 2002.
58. *Stiles, J. R., Bartol, T. M. Jr., Salpeter, E. E., Salpeter, M. M.*, Monte Carlo simulation of neurotransmitter release using MCell, a general simulator of cellular physiological processes, in: Bower, J. M. (Ed.), Computational neuroscience, Plenum Press, New York, 1998.
59. *Ichikawa, K.*, Bioinformatics Application Note **17** (2001) 483.
60. *DeBeer, D. J., Kourie, D. G.*, South African Jl. of Science **96** (2000) 569.
61. *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., Minch, E., 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.,* Bioinformatics **19** (2003) 524.
62. *Hedley, W. J., Nelson, M. R., Bullivant, D. P., Nielson, P. F.,* Philosophical Transactions of the Royal Society of London A **359** (2001) 1073.
63. *Schaff, J. C., Loew, L. M.,* The Virtual Cell, Pacific Symposium on Biocomputing **4** (1999) 228.
64. *Schaff, J. C., Slepchenko, B. M., Choi, Y. S., Wagner, J. M., Resasco, D., Loew, L. M.,* Chaos **11** (2001) 115.
65. *CRGM*-Database, Centro di Ricerca in Genetica Molecolare, Istituto di Istologia e Embriologia Generale, University of Bologna, 2002. http://apollo11.isto.unibo.it/Cells.htm
66. *NIH Database*, National Institute of Health database, 2004. http:// molbio.info.nih.gov/ molbio/ db.html.
67. *Crampin, E. J., Schnell, S.*, Progress in Biophysics & Molecular Biology **86** (2004) 1.
68. *Bower, J. M., Bolouri, H.,* Computational Modeling of Genetic and Biochemical Networks, MIT Press, Cambridge, 2001.
69. *Somogyi, R., Sniegoski, C. A.*, Complexity **1** (1996) 45.

70. *Akutsu, T., Kuhara, S., Maruyama, O., Miyano, S.*, A system for identifying genetic networks from gene expression patterns produced by gene disruptions and overexpressions, 9th Workshop on Genome Informatics, Tokyo, Dec. 10, 1998.
71. *Chen, T., He, H. L., Church, G. M.*, Pacific Symposium on Biocomputing **4** (1999) 29.
72. *Gillespie, D. T.*, Jl. Physical Chemistry **81** (1977) 2340.
73. *Gillespie, D. T., Mangel, M.*, Jl. Chemical Physics **75** (1981) 704.
74. *Maria, G., Szedlacsek, S., Laslo, A. C., Zaharia, E.*, Modular-based simulation of regulated protein synthesis, Conference on Enzyme Reaction Engineering, Dubrovnik (Croatia), Sept. 21, 2005.
75. *Heinrich, R., Rapoport, S. M., Rapoport, T. A.*, Prog. Biophys. Mol. Biol **32** (1977) 1.
76. *Hyver, C., Le Guyader, H.*, Biosystems **24** (1990) 85.
77. *Wuensche, A., Lesser, M. J.*, The global dynamics of cellular automata; An atlas of basin of attraction fields of one-dimensional cellular automata, Santa Fe Institute, Addison-Wesley, Reading (MA), 1992. http://www.cogs.susx.ac.uk/users/andywu/gdca.html
78. *Bray, D., Bourret, R. B., Simon, M. I.*, Mol. Biol. Cell **4** (1993) 469.
79. *Thomas, R., Thieffry, D., Kaufman, M.*, Bull. Math. Biol. **57** (1995) 247.
80. *Speck, C., Weigel, C., Messer, W.,* EMBO J. **18** (1999) 6169.
81. *Ptashne, M.*, A Genetic Switch: Phage Lambda and Higher Organisms, Cell Press & Blackwell Scientific Publs., Cambridge (Mass.), 1992, pp. 82.
82. *Hargrove, J. L., Schmidt, F. H.*, FASEB Jl. **3** (1989) 2360.
83. *Vance, W., Arkin, A., Ross, J.,* Proceedings of the National Academy of Science USA **99** (2002) 5816.
84. *Aris, R.*, Elementary chemical reactor analysis, Prentice-Hall, New Jersey, 1969.
85. *Grainger, J. N. R., Gaffney, P. E., West, T. T.*, J. theor. Biol. **21** (1968) 123.
86. *Wallwork, S. C., Grant, D. J. W.*, Physical Chemistry, Longman, London, 1977.
87. *Milani, M., Pesce, A., Bolognesi, M., Bocedi, A., Ascenzi, P.*, Biochemistry and Molecular Biology Education **31** (2003) 228.
88. *Welch, G. R., Keleti, T.*, Trends Biochem. Sci. **12** (1987) 216.
89. *Chen, D., Feng, J., Krüger, R., Urh, M., Inman, R. B., Filutowicz, M.*, J. Mol. Biol. **282** (1998) 775.
90. *Kholodenko B. N., Schuster S., Rohwer J. M., Cascante M., Westerhoff H. V.*, Febs Letters. **368** (1995) 1.
91. *Sauro, H. M., Kholodenko B. N.*, Prog. Biophys. Mol. Biol. **86** (2004) 5.
92. *Tyson, J. J., Novak, B., Odell, G. M., Chen., K., Thron, C. D.*, Trends Biochem. Sci. **221** (1996) 89.
93. *Ballet, P., Zemirline, A., Marce, L.,* Course on 'Cellular automata, reaction-diffusion and multiagents systems for artificial cell modeling', Universite de Bretagne Occidentale, Laboratoire de Methodes Informatiques, Evry (France), 2001; http://www.lami.univ-evry.fr/