ISSN: 2687-8100

Archives in Biomedical Engineering & Biotechnology

DOI: 10.33552/ABEB.2026.06.000691

Iris Publishers

Research Article

A Novel Kinetic Modelling Framework of Gene Expression Self-Regulation, Able to Maintain Intracellular Homeostasis While Growing Auto-catalytically on Environmental Nutrients Present in Variable Amounts

Gheorghe Maria[1,2*]

[1]*Department of Chemical & Biochemical Engineering, University Politehnica of Bucharest, Romania*

[2]*Romanian Academy, Calea Victoriei 125, Bucharest, Romania*

***Corresponding author:** Gheorghe Maria, Department of Chemical & Biochemical Engineering, University Politehnica of Bucharest, Romania

Received Date: May 15, 2026

Published Date: June 30, 2026

Abstract

Systems Biology is one of the modern tools, which uses advanced mathematical simulation models for in-silico design of micro-organisms that possess desired characteristics. Due to the near astronomic complexity of cell processes, simulation of the cell metabolism dynamics by means of deterministic models are based on a reduced reaction mechanism, employing both individual and lumped species and reactions, of known kinetics and stoichiometry [1-3]. A central part of such models concerns the adequate simulation of the protein synthesis homeostatic self-regulation present in any gene expression regulatory module (GERM) that produces enzymes controlling the whole cell metabolism with negative feedback loops and rapid adjustments of the enzymatic activity. The deterministic approach uses ODE kinetic models of cell processes with continuous variables (cell individual or lumped species).

As proved in this paper, the classical formulations by using the default "whole-cell quasi-constant volume" kinetic models (CVKM) approach, do not explicitly consider the cell volume exponential increase during the cell growth leading to biased and distorted/wrong conclusions on GERM regulatory performances. When the continuous variable CVKM dynamic models are used to model the cell processes, the default-modelling framework is those defining the species differential mass balances (ODE set) in terms of species concentrations but ignoring the cell volume increase during ca. 80% of the cell cycle. The classical formulation is valid if and only if the cell system would present a constant volume. But such a hypothesis is not valid for the cell system, of which the cell volume is growing continuously, doubling after one cell cycle. When such dynamic models are used to simulate the cell enzymatic processes, wrong results are expected because the hypothesis of a constant cell volume and, implicitly, of a constant osmotic pressure. To partially overcome these shortcomings, some CVKM models eventually account for the cell-growing rate as a pseudo-'decay' rate of the key-species (often lumped with the degrading rate) in a so-called 'diluting' rate, or they use the average cell dilution rate Dm = ln(2)/(cell cycle) instead of the instant correct cell dilution "Di" as proved here.

To overcome such serious drawbacks of the classical CVKM cell kinetic models, Maria [1-7] introduced the novel concept of "mechanistic whole silicon cell", materialized in a novel math (kinetic) modelling framework "WCVV" of the cell metabolic processes, referring to the "whole silicon cell" (WC), of variable-volume (VV), and characterizing assumed isotonic growing cells. The WCVV kinetic model hypotheses refer to a novel kinetic modelling framework of the cell metabolic processes that is able to maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. As proved by Maria [7] this WCVV novel modelling framework can simulate a lot of GERM/GRC regulatory properties, related to the cell response to the dynamic ("impulse"-like), or stationary ("step"-like) perturbations in the environment, otherwise impossible to be reproduced by the classic CVKM without imposing "the total enzyme activity" or the "total enzyme concentration" constraints suggested in the literature [8].

This paper exemplifies the overwhelming importance of using a holistic variable-volume whole silicon cell (WCVV) modelling framework with explicitly including constraints accounting for the cell-volume growth while preserving a constant osmotic pressure and membrane integrity. In this novel WCVV approach, the cell volume change arises naturally from the chemical dynamics associated to an assumed mechanism of the cell biochemical reactions (related to the central carbon metabolism –CCM, and genetic regulatory circuits - GRCs). In fact, the novel WCVV modelling framework lead to a paradigm shift when simulating the dynamics of cell metabolic processes from the CCM, and especially the self-regulatory properties of GERMs, or of GRCs [7]. The provided example proves the significant improvements offered by the novel WCVV modelling framework in both precision and simulation potential.

Keywords: Modular, deterministic, structured kinetic models of cell processes; 'mechanistic whole silicon cell' novel modelling concept; novel kinetic modelling framework '**WCVV**'; gene expression regulatory modules (**GERM**); homeostatic regulation; variable cell-volume, isotonic approach

Introduction

As reviewed by Maria [9-12], with the accumulation of experimental information, and its storage in bio-*omics* databanks, the mathematical models and algorithms dedicated to cell process simulation, in the area of "Systems Biology" reported a sharply exponentially-like increase after 1990 (see the Scopus citations in this area). Beside its quick development, and the novel concepts introduced in this topics, the definitions of *Systems Biology* existing in the literature have evolved accordingly, trying to better specify its importance in the *In-silico* study of the cell metabolism [13], as followings:

i).--"The science of discovering, modelling, understanding and ultimately engineering at the molecular level the dynamic relationships between the biological molecules that define living organisms" [14].

ii).-- "System Biology is a comprehensive quantitative analysis of the manner in which all the components of a biological system interact functionally over time" [15].

iii).--"Perhaps surprisingly, a concise definition of Systems Biology that most of us can agree upon has yet to emerge" [16], that is: "Systems biology is a holistic, quantitative approach to understanding biological systems by analyzing the interactions of all their components over time. It emphasizes the interconnectedness of biological components and aims to understand how these interactions lead to emergent properties of the system, which cannot be predicted from studying individual parts alone."

Following the tremendous progresses made in the Bioinformatics (Figure 1), Systems biology [17], Computational biology [18], and eventually Synthetic biology [19], the *in-silico* analysis (based on mathematical models) of cell processes is experiencing an impetuous development. The used math models are of various types [1-3,7]: topological, or structured. In turn, the structured dynamic models of living cells are of the following types: deterministic (with continuous variables); Boolean (with discrete variables); stochastic (with statistically distributed variables); or with mixed variables.

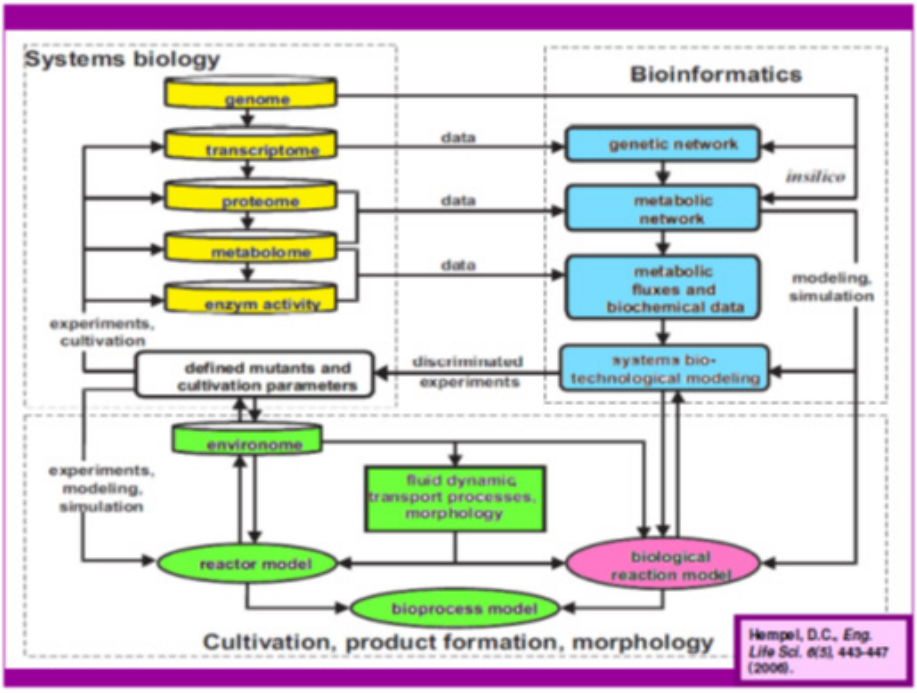

Figure 1: The strong connection and equivalence between "Systems biology", and "Bioinformatics". Adapted from [20].

The concepts/rules/algorithms of the Chemical and biochemical process engineering (ChBPE), and those of the nonlinear systems theory has been proved to be quite successful for the optimization of production in inanimate systems [3,21]. Where biology enters the arena other than as the usual source of the raw-materials (oil, coal, natural gas) [22,23], the success rate of process engineering tends to drop. "Biological systems appear to be more resistant to control by the engineer; they are more 'stubborn'. In this presentation we shall attempt to relate this phenomenon to a feature that emerges from an entirely different field, i.e. 'functional genomics'. In the latter field, one regularly observes the phenomena of redundancy and robustness. Where the former used to be explained on the basis of parallel pathways, both may be more related to the processes of adaptation that are so characteristic of live and certainly Life systems [1-3,24]. Although metabolic engineering did interface with genomics, it rarely incorporated the aspect of adaptation, i.e. of the changing of the structure of the network with conditions [1-3,9-12,24]. Flux balance analysis does not do so either; its pathways are given and assumed to be always expressed." [25-27].

Here it is worth notice that living cells are extraordinarily complex machines, with an astronomical complex organization (Figure 2), that exhibit remarkable behaviors including, most predominately, their self-replication where GERM-s and GRC-s play the essential role. Although much progress has been made in understanding the mechanisms underlying this behavior, further efforts on both experimental and theoretical fronts will be required to deepen our understanding of it. Towards this end are efforts to simulate global cellular behavior starting from the molecular level of detail [17,18,28,29,30,31].

<table>
<tr><td>Self-replicating apparatus</td><td rowspan="3">Time scale separation (slow/fast manifolds)</td><td>Self-replication</td><td>Regulation network (GRC)</td></tr>
<tr><td>Replisome, Partitioning apparatus, Z-ring</td><td>Nucleotid replication and partitioning, cell division</td><td>Cell cycle regulation</td></tr>
<tr><td>Nucleotid</td><td>Supercoil and organize genome</td><td rowspan="2">Gene expression regulation (GERM)</td></tr>
<tr><td>Ribosomes, Genome, Energy harnessing apparatus</td><td rowspan="2">Intermediate characteristic time</td><td>Protein synthesis, Store genetic info, Harness energy</td></tr>
<tr><td>Cell wall, Nucleic acids, Coenzymes</td><td>Metabolic cycles, pathways, Transcription, Translation</td><td>Metabolism regulation (GRC)</td></tr>
<tr><td>Peptidoglycan, Membrane, Protein cplx., Nucleosides</td><td rowspan="2">Succession of events</td><td>Catalysis, Energy currency</td><td rowspan="2">Regulation of enzyme activity</td></tr>
<tr><td>Lipids, proteins, nucleosides</td><td>Catalysis, Hydrophobic effects</td></tr>
<tr><td>Saccharides, Fatty acids, Aminoacids</td><td>Transient recovering time</td><td colspan="2" rowspan="2">Intermediates and building blocks for cell structures and fubctions</td></tr>
<tr><td>Simple metabolites</td><td></td></tr>
<tr><td>Raw materials (nutrients)</td><td rowspan="2">← Temporal Hierarchy →</td><td colspan="2">Source of energy and material</td></tr>
<tr><td>← Structural Hierarchy →</td><td colspan="2">← Functional Hierarchy →</td></tr>
</table>

Figure 2: The hierarchical complex organization of living cells. Adapted after [2].

"Ideally, such whole-cell models would integrate massive amounts of biological information into compact, unambiguous and testable hypotheses which could be used to explore the entire panoply of cellular relationships in both healthy and diseased states [32-36]. "Much remains before useful comprehensive mathematical models of cellular behaviour with predictive power can be realized. Obviously, the vast number of unknown parameters of such kinetic models (rate coefficients and cell species concentrations), and ambiguities in chemical mechanisms occurring within a cell hinder progress, but there is also no consensus as to what modelling framework will ultimately prove most effective. By framework, we mean a group of general assumptions and procedures that can transform a physico-chemical model of a cell into a set of mathematical expressions. Such a framework would be applicable to all such models, regardless of the particular set of chemical reactions and components assumed" [36].

In this work, the novel WCVV ("mechanistic Whole-cell variable volume") modelling framework will then be proposed. This WCVV novel modelling framework was proposed by Maria [1,4-7],

while the study of the WCVV properties, features, and advantages were given by Maria [1-3,6,7,9-12,24,37,38,40-43]. In this novel WCVV approach, the cell volume change arises naturally from the chemical dynamics associated to an assumed mechanism of the cell biochemical reactions (related to the CCM and GRCs). As we will prove here, when appropriate rate constants of the constructed cell kinetic model are obtained from the known stationary (homeostatic) concentrations of the all species of the model (taken individually or lumped), the self-replicating dynamic behavior of cells can be modeled. Beside, we illustrate the large number of applications and features, of such a novel WCVV cell kinetics modelling framework. Exemplifications will be made by using simplified kinetic models with –omics data of *E.coli* (from Ecocyc, and KEGG).

In short, the work reviews the novel concept of "mechanistic silicon whole cell", materialized in a novel math (kinetic) modelling framework "WCVV" of the cell metabolic processes, referring to the "whole-cell" (WC), variable-volume (VV), of isotonic growing cells. The "mechanistic cell" novel hypotheses of "WCVV" refer to the whole-cell, variable-volume, of isotonic growing cells. The paper points-out and proves the advantages of WCVV dynamic models when building-up modular model structures of simplified form that can reproduce complex protein synthesis regulation inside cells. These WCVV kinetic models of cells are able to maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts.

The paper proves, in a simple way how the novel modelling concept of "WCVV - mechanistic silicon whole cell" is a extremelly valuable simulation tool that can overcome most of the simulation difficulties of the dynamics of the evolutionary cellular metabolism mentioned above. Thus, the more realistic WCVV approach is reviewed, by exemplifying and by pointing-out its advantages when developing modular kinetic representations of the homeostatic gene expression regulatory modules (GERM) that control the protein synthesis, and the homeostasis of metabolic processes. The paper briefly reviews the general concepts and hypotheses of the WCVV modelling framework. Simple kinetic models at a generic level, were used by Maria [7] to exemplify how this methodology can be extended to reach the following objectives.

A) The novel WCVV math modelling framework and rules lead to a paradigm shift when simulating the dynamics of cell metabolic processes from the central carbon metabolism (CCM), and especially the self-regulatory properties of GERMs. As proved in the literature, the WCVV framework is very suitable to model the GRCs [1-3,6,38,39,41,42,44], but also the modular CCM included in hybrid kinetic models [44-47]. These references proved that the modular modelling approach of CCM is very advantageous, by simplifying the numerical simulation. In fact, as revealed by the Maria [44-46], some of the essential modules of the CCM are: the glycolysis; the phosphotransferase (PTS)-system, or an equivalent one for GLC-uptake; the pentose-phosphate pathway (PPP); tricarboxylic acid cycle (TCA), the ATP recovery system [44-47].

B) The novel WCVV math modelling framework was proved to be very adequate to simulate the cell adaptation to the environmental conditions, i.e. the changing of the metabolism behavior (dynamics after perturbations, biomass evolution), caused by the environmental changes, such as:

i) Changes of metabolic pathways determined by the genetic switches (GS) action, induced by the environmental changes [3,39,41,42];

ii) Oscillatory syntheses induced by the environmental changes [44,48- 50,52,54-63]

C) GMO design based on the target GRC regulatory adaptive efficiency (see an example for the mercury uptake by *E.coli* given by [7,24,64,65]

D) Better simulation of the cell regulatory GRC response to stationary perturbations, leading to the transition from one steady-state (stable balanced homeostasis) to another cell homeostasis.

E) Better characterization of the self-regulatory efficiency of GERM-s, by introducing for the first time in the literature the quantitative regulatory performance indices (P.I.-s) of GERM-s, *in-silico* tested under WCVV modelling framework by Maria [1-3,7,24,66,67].

F) Based on the (P.I.-s) and of a library of GERM-s structures, defined under WCVV modelling framework by Maria [1-3,7,9-12], the novel WCVV math modelling framework was also proved to be very effective in design effective GMOs with industrial applications [45,46,65]. Such a GERM-s library is proving its value for the *in-silico* (math model based) design of GMOs by using the modular approach. Such an approach is based on the observation of Sauro and Kholodenko [68] about the modular structure of living cells. They provided 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. regulation of a certain reaction, gene expression, etc.). More complex functions, such as regulatory networks, synthesis networks, or metabolic cycles can be built-up using the *building blocks* rules of the emergent *Synthetic Biology* are described by Heinemann and Panke [69]; Qian et al. [70]; Maria [3,24]. Moreover, such a library proves its efficiency, because the GERM structures are repeatable (similar) in most micro-organisms. As another observation, by chance such a "building blocks" cell structure is computationally very tractable when developing cell reduced dynamic models, by defining and characterizing various metabolic sub-processes, such as: regulatory functions of the gene expression regulatory modules (GERM-s), and of genetic regulatory circuits (GRC-s), kinetics of complex enzymatic reactions, energy balance functions for ATP(ADP/AMP) renewable systems, electron donor renewable systems of the NADH, NADPH, FADH, FADH2 compounds, hydrophobic effects, or functions related to the metabolism regulation (regulatory components/reactions of the metabolic cycles,

gene transcription and translation); genome replication/gene expression regulation (protein synthesis, storage of the genetic information, etc.)(Figure 2) [1-3]. When modelling GRC-s under the WCVV modelling framework, by chance, the number of interacting GERM-s is limited, one gene interacting with no more than 23-25 [1-3,71]. Consequently, such a *modular approach* allows disassembling the whole cell system in parts (modules) and then, by performing tests and applying suitable numerical procedures, to define rules that allow recreating the whole and its characteristics by reproducing the real system. Such an approach, combined with derivation of lumped modules, allows reducing the model complexity by relating the cell response to certain perturbations to the response of few inner regulatory loops instead of the response of thousands of gene expression and metabolic circuits. Such a procedure is very suitable for modelling GRC-s by linking GERM models in such a way to maintain the cell homeostasis that is to maintain relatively invariant species concentrations despite perturbations. See Maria [1-3,7] for an extended discussion about this subject.

G) Based on these (P.I.–s) and on the mentioned GERM–s "*library*" Maria [1,7] proposed a "buiding-blocks" strategy and rules to connect GERMs in chains to build-up complex GRC-s (Figure 3). See the examples provided by Maria and Luta [65]; Maria [6,38,41,42,64]. Some of these structures were experimentally validated by [3,7,72,73].

H) As proved by Maria [7], the kinetic models of GERM-s/GRC-s derived by using the WCVV modelling framework are able to simulate the maintaining of intracellular homeostasis (balanced growth), while growing auto-catalytically on environmental nutrients present in variable amounts.

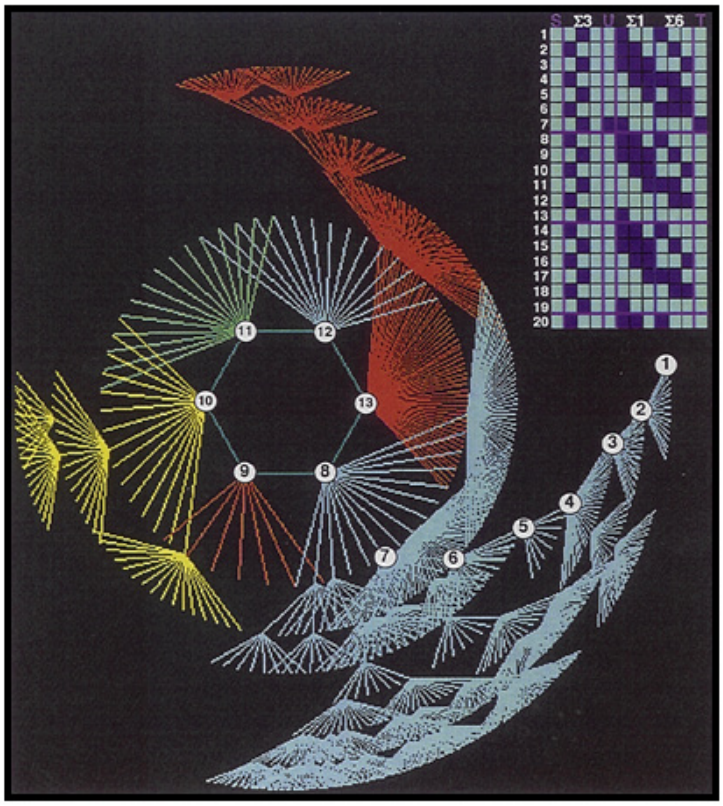

Figure 3: A cell GRC simulator of Boolean type. Adapted from [74].

2. The Silicon Cell and Gene Expression Regulation Analysis

According to the *silicon cell* project (www.siliconcell.net), "the silicon cell is a collection of computer replica of processes in living organisms, which should be linked up to produce models of larger networks." They can be also used to play with and engineer biological processes on line, in silico. An example is provided by Westerhoff [75]. Following important progresses made in the computing rules, numerical algorithms, and computational speed, a large number of *in-silico* studies (based on mathematical models) about cell processes have been reported. As Palsson [28] noted, such models of the cell metabolic functions/processes are constrained by a lot of physico-chemical factors. One of them concerns the osmotic pressure. This iso-tonicity (constant osmotic pressure) is the basis of the novel WCVV modelling framework

proposed by Maria [1,4-7], while the study of the WCVV properties, features, and advantages were given by Maria [1-3,6,7,9-12,24,37-42,76]. By using the classical (default) modelling approach CVKM , a large number of rather deterministic kinetic models have been developed to simulate the central carbon metabolism (CCM), that is the so-called "Genome-Scale Reconstruction", such those of Forster, et al. [77] for *Saccharomyces cerevisiae*; or those of [78,79] for *Escherichia coli*; or those of Heinemann et al. [80] for *Staphylococcus aureus*; or those of Rodriguez-Prados et al. [81] for threonine synthesis in *E.coli*.

Also, dedicated software / rules to perform such a genetic network reconstruction, or reconstruction of the metabolic networks have been reported by [82-84]. Consequently, several review studies have set the milestones [28,85] for the development of this new research topics (at the interface between several classical fields, that is: biology, control theory of nonlinear systems, numerical computation, (bio)chemical engineering, bioinformatics. This topics is called "the Silicon cell" by Westerhoff [75], or the "mechanistic silicon whole cell" by Maria [1,4-7,9-12]. It is worth mentioning that in all *in-silico* alternatives to build-up a cell model, this is an iterative process requiring (Figure 4) experimental validation of the essential model parts. According to Westerhoff [25,75], the *in-silico* approach of most of math models follows the chemical and biochemical process engineering concepts, rules, and numerical algorithms briefly mentioned by Maria [1-3,7,24].

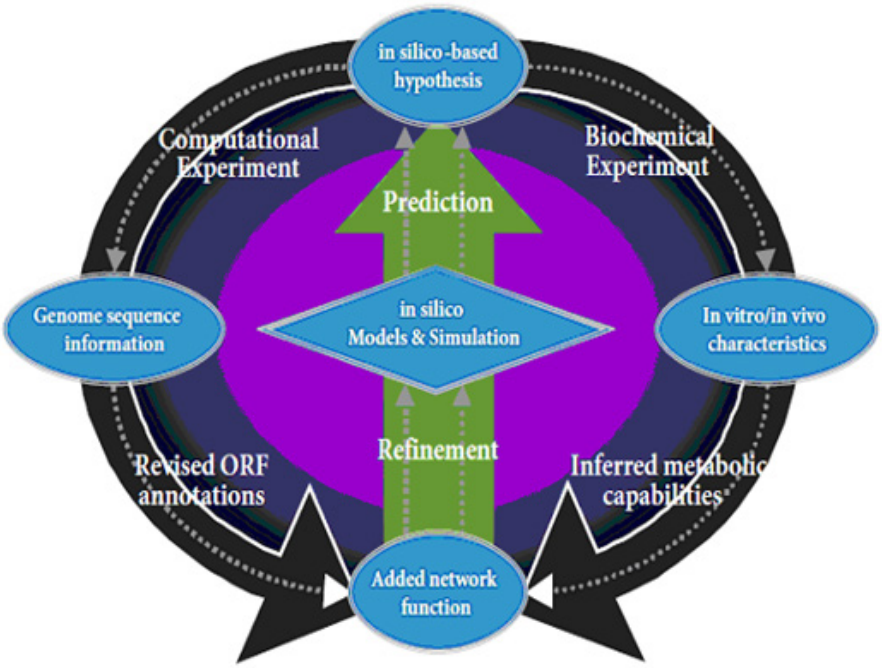

Figure 4: Iterative *in silico* model building in biology involves the formulation of experimentally testable hypotheses based on the *in silico* analysis, collection of experimental data, and subsequent refinement of the mathematical models based on these data. Adapted from Palsson [28].

As commented by Westerhoff [75], "Process Engineering has been quite successful for the optimization of production in inanimate systems. Where biology enters the arena other than as the usual source of the material (oil, sewage, etc. [22]), the success rate of process engineering tends to drop. Biological systems appear to be "more resistant" to the control by the engineer; they are more "stubborn." This phenomenon is related to a feature that emerges from an entirely different field, i.e. *functional genomics*. In the latter field, one regularly observes the phenomena of redundancy and robustness. Where the former used to be explained on the basis of parallel pathways, both may be more related to the processes of adaptation that are so characteristic of live and certainly Life systems. Although *metabolic engineering* did interface with genomics, it rarely incorporated the aspect of adaptation, i.e. of the changing of the structure of the network with conditions. Flux balance analysis does not do so either; its pathways are given and assumed to be always expressed.

Although metabolic control analysis deals with the engineering of Life processes in much the same way, it already explains part of the resistance of biological systems to process engineering. Cell adaptation to the environment is rather given by the 'Hierarchical Control Analysis' and by '*Hierarchical Regulation Analysis*', both being part of the hierarchical complex organization of living cells." (Figure 2) As proved by Maria [7], the novel WCVV modelling framework is very suitable to simulate the GERM-s/GRC-s responsible to cell response to the environmental perturbations, and the cell

metabolism adaptation to environmental changes (for instance, by using genetic switches to re-direct the entire metabolism toward a more efficient homeostasis. In fact "the concepts and principles developed for Life processes also apply to inanimate processes, yielding new principles for process engineering in general." [75].

According to Kiparissides, et al. [2011], "the mathematical models of biological systems evolve toward a holistic 'closed loop' approach that will facilitate the control of the *in vitro* [biological process] through the *in silico*." (that is a math model-based full control of the bioreactor). Recent advances in developing the so-called "micro-array data technique" [Le Phillip et al., 2004], it is now possible to record the dynamics of hundreds of cell species, thus generating kinetic trajectories required by the development of cell CCM/GRC_dynamic models on a structured basis, and using structured and conventional type of data [Maria, 2017b,2018,2023].

As discussed by Kiparissides, et al. [86], math models of biological systems developed over the last decades incorporate various degrees of cell structure and mathematical complexity. "Models of single cells, cell populations and microbial cultures have been central in the understanding and improvement of biological systems, as well as in the optimization and control of bioprocesses [87]. The large-scale generation of biological data obtained with the development of a variety of high-throughput experimental technologies demand for mathematical model building to become a centre of importance in biology [88]. Bailey [89,90] argued the development of mathematically and computationally orientated research has failed to catch up with the recent developments in biology. Furthermore, he concluded that the little attention that mathematical modelling of biological systems receives from experimentalists could be partly attributed to the lack of effective communication of the benefits of formulating and using a mathematical model.

Even relatively simple micro-organisms, which have been extensively studied, are hosts to a complex network of interconnected processes occurring on diverse time scales within a confined volume. The multi-level nature of the regulatory network of cells and the interactions occurring at the intracellular level further augment this complexity [91]. Therefore, attempts to wholly model the function of even a single cell are currently non-trivial, if not impossible. The amount of delicate intracellular measurements required to validate such a model is exhaustive both in terms of labour as well as cost. Uncertainties introduced on the parameter identifiability level [92] and on the mechanistic level further complicate this task [86].

"Borrowing research principles from the Process Systems Engineering paradigm, mathematical modelling of biological systems can provide a systematic means to quantitatively study the characteristics of the multilevel interactions that occur in cell bio-processing. The literature around mathematical modelling of biological systems, be they prokaryotic or eukaryotic, is arguably too vast to summarize within the limited space of this work. Indicatively, mathematical models have been successfully used to design optimal media [93], to identify previously ignored growth limiting factors [94], to optimize culture growth and productivity [95-97], and to apply the process control principles to operate the bioreactors." [86,98].

Pörtner and Schäfer [99] compared a selection of models that existed in the literature at that time and carried out an analytic error and range of validity analysis. They found significant variations in the values of maximum growth rate, yields and nutrient Monod constants used by researchers. They came to the conclusion that the models' predictions involved significant errors, particularly due to the lack of understanding of cellular metabolism and the limited data ranges within which the model was valid. They further pointed out that the majority of studies presented either utilized literature data to validate the models or generate their own experimental data without any form of systematic design of experiments. In order to maximize the gains from the ever increasing influx of biological information, the approach to modelling needs to shift towards a systematic framework from conception to optimization. Herein we attempt to formalize a structure upon which experimental and mathematical biology can interact seamlessly. The use of model-based techniques can facilitate the reduction of unnecessary experimentation hence reducing operating labour and cost by indicating the most informative experiments and providing strategies to optimize and automate the process at hand." [86].

Despite the above difficulties, accumulation of cell data (stored in bio-omics databanks), and recent advances in the computational methodology, make it possible to achieve substantial progresses in the detailed structured modelling of cellular processes [1-3,7,24], especially those related to the central carbon metabolism (CCM) [45,46], and to complex genetic regulatory circuits GRC-s [64,65]. In this context, the current work is aiming to shortly review the novel WCVV ("whole silicon cell of variable-volume") kinetic modelling framework, proposed by Maria [1,4-7], while the study of the WCVV properties, features, and advantages were given by Maria [1-3,6,7,9-12,24,37,38,40-43,64-67]. The novel WCVV was proved to be very effective and appropriate for accurately model the individual GERM modules, but also the complex GRC-s. These WCVV kinetic models can include the "*Hierarchical Regulation Analysis*", by linking GERM-s in complex GRC-s, following the rules studied by Maria [1-3,7,24]. Thus, the WCVV approach allows modelling the adaptable living cells, by also including genetic switches models able to re-direct the cell metabolism [3,39,41,42].

Such WCVV kinetic models can be used for various *in-silico* analyses, such as [7]:

A) *in-silico* design of genetically modified micro-organisms (GMO)

B) To characterize the GERM regulatory efficiency, and various local or global (cell level) regulatory properties, such as

i) The role of the high cell-"ballast" (the sum of cell species copynumbers) in "smoothing" the effect of internal/external dynamic or stationary perturbations of the cell system homeostasis

ii) The system isotonicity constraint reveals that every inner

primary perturbation in a key-species level (following a perturbation from the environment) is followed by a *secondary one* transmitted to the whole-cell via cell volume

iii) Allows comparing the regulatory efficiency of various types of GERM-s

iv) Allows a more realistic evaluation of GERM regulatory performance quantitative indices (P.I.-s), defined by the first time in the literature by Maria [1-3,7]

v) Allows studying the recovering/transient intervals between steady-states (homeostasis) after stationary ('step'-like) perturbations

vi) Allows studying the species connectivity

vii) Allows studying the conditions when the system homeostasis 'intrinsic stability' is lost

viii) Allows studying the GERM-s/GRC-s self-regulatory properties after a dynamic ('impulse'-like), or a stationary ('step'-like) perturbation

ix) Allows studying the plasmid-level effects in cloned cells

x) Allows simulation of the cell balanced growth, that maintain the intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts

xi) Allows to build-up modular regulatory chains (GRC-s) of various complexity

xii) Prove feasibility of the cooperative vs. concurrent construction of GRC-s that ensures an efficient gene expression, system homeostasis, proteic functions, and a balanced cell growth during the cell cycle

xiii) Prove that the classic ("default") kinetic modelling approach (CVKM, "whole cell of quasi-constant volume", non-isotonic cell systems) is erroneous compared to the novel WCVV, by leading to distorted and wrong simulation results and conclusions [7]. In such a way, the novel WCVV cell kinetic modelling approach replaces the wrong classical (default) CVKM modelling methodology, thus correcting the distorted and false/wrong predictions of CVKM cell kinetic models unfortunately, still largely used in the literature

xiv) An experimental validation of the GERM / GRC modular lumped WCVV kinetic models was done by using the state-variables recorded dynamics in batch or fed-batch bioreactors [65], or from using the literature data, e.g. those of Yang et al. [72], or bio-omics databanks [like KEGG, Ecocyc]. The rate constants of the GERM / GRC modular kinetic models were estimated from solving the quasi-steady-state (QSS) model equations with using the stationary (homeostatic) concentrations of the key-species [determined experimentally, or from bio-omics databanks], and by imposing some holistic regulatory properties [7,66,67].

xv) The use of the novel WCVV mathematical modelling framework to obtain modular, deterministic, structured, *hybrid* dynamic models, linking the macroscopic state-variables of the bioreactor to the cell (nano-scopic) state-variables belonging to the central carbon metabolism (CCM), thus offering a more precise and a more detailed (no. of state variables) predictions of the bioreactor/biomass dynamics. Such complex kinetic models are very effective for *in-silico* engineering evaluations [3,7,24].

3. Some simple GERM models

To perform *in-silico* analyses of the cell gene expressions self-regulation properties, shortly reviewed in section 2, an adequate kinetic model of the target GERM module is necessary. In a deterministic approach the model should include the all reactions and species relate to that GERM module, taken individually, or lumped. As proved by Maria [1-3,7], adequate WCVV kinetic models for GERM modules have multiple purposes, such as:

i) Build-up/design GRC-s of desired characteristics by using a 'building-blocks strategy, and defined rules

ii) To quantitatively characterize the self-regulation properties of various GERM-s types (the so-called 'performance indices – P.I.s), and to compare between them. Formally, a GRC is a chain of GERM-s (Figure 3) linked following certain rules studied by Maria [1-3,7] by using the WCVV modelling approach, impossible to be derived by using the classic CVWK modelling framework; iii) to establish the role of each GERM in a GRC construction [42,64,65]. And, eventually, a good GRC model can help for *in-silico* design of GMO-s.

One of the very promising applications of WCVV kinetic models with continuous variables is the study of individual gene expression regulatory modules GERM-s, and of the genetic regulatory circuits (GRC-s) (e.g. applications of Maria [6,7,39,40,-42,64,65]), in order to predict the way in which biological systems are self-regulated and respond to signals. The emergent field of such efforts is the so-called '*gene circuit engineering*' and a large number of examples (reviews of Maria [3,6,24]) have been reported with *in-silico* creation of novel GRC-s conferring new properties/functions to the mutant cells (i.e. desired 'motifs' in response to external stimuli), such as (Figure 5).

a) Toggle-switch, i.e. mutual repression control in two gene expression modules, and creation of decision making branch points between on/off states according to the presence of certain inducers (Maria [39,41,42]; hysteretic GRC behavior, that is a bio-device able to behave in a history-dependent fashion, in accordance to the presence of a certain inducer in the environment; GRC oscillator producing regular fluctuations in network elements and reporter proteins [44,48-51,57,58], and making the GRC to evolve among two or several quasi-steady-states

b) Specific treatment of external signals by controlled expression such as amplitude filters, noise filters, or signal/stimuli amplifiers [41,42,64,65].

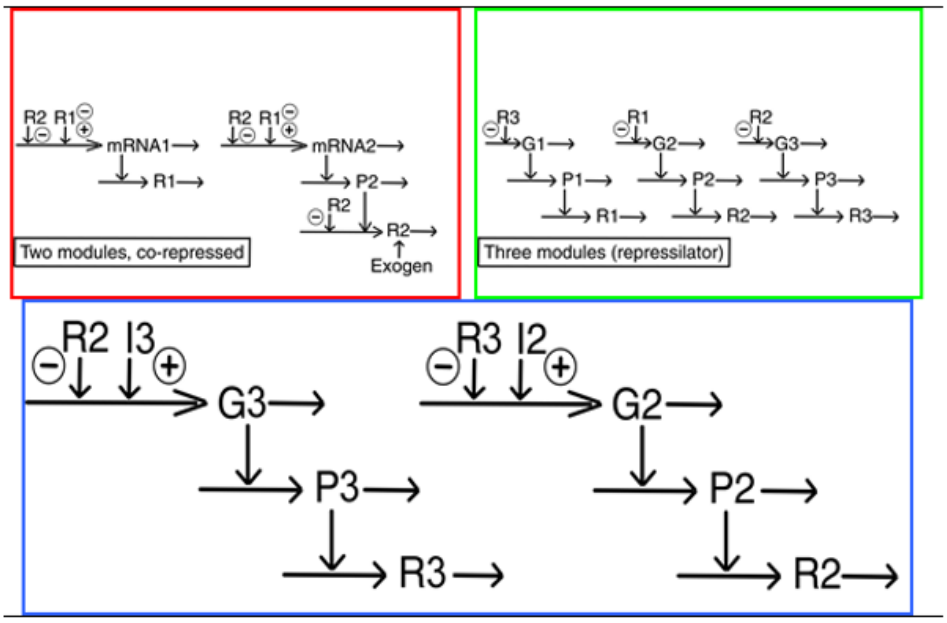

Figure 5: (top-left) The simplified scheme of a genetic switch GS [two mutually self- and cross-repressed GERM-s), after Wall et al. [103]. (top-right) The 'repressilator' of *lac*-operon expression in *E. coli* (three mutually repressed GERM-s; each expressing a protein that represses the next gene in the loop). Adapted from Ellowitz and Leibler [57]. (down). The principles of a genetic switch (GS). Two self- and cross-repressed gene expression coupled modules GERM–s that ensure replication of the pairs P2/G2 and P3/G3 respectively. Proteins P2/P3 expression is activated by the specific inducers I2 and I3 respectively, and repressed by the R2 and R3 proteins produced by each GERM module. Horizontal arrows indicate reactions; vertical arrows indicate catalytic actions. Adapted after Maria [38]. Absence of a substrate or product indicates an assumed concentration invariance of these species.

This subject is of tremendous importance as long as GERM-s and GRC-s are the essential parts used to re-design the cell metabolism.

Here it is worth to mention the modular construction of GRC-s. Kholodenko, et al. [100] provided examples of biological systems that have evolved in a *modular* fashion and, in different contexts, perform the same basic functions. In fact, a GRC presents a modular construction including chains of GERM-s. Each GERM groups the cell components and reactions that are linked to generate an identifiable function (e.g. regulation of a certain reaction, gene expression, etc.). More complex functions, such as regulatory networks, synthesis networks, or metabolic cycles can be built-up using the *building blocks* rules (reviewed by Heinemann and Panke [69], and by Maria [1,7]). The modular organization of cell regulatory systems is computational very tractable. Moreover, it is well-known that one gene expression interacts with no more than other 23-25 GERM-s [71], while most of GERM structures are repeatable. Consequently, in developing the GRC analysis, the modular approach under a WCVV modelling framework is preferred due to various advantages:

i) A separate analysis of the constitutive GERM-s in conditions that mimic the stationary and /or perturbed cell growth;

ii) Investigation of module links used to construct the whole GRC of an optimized regulatory efficiency that ensures key-species homeostasis and network holistic properties;

iii) Investigation of GRC characteristics such as the tight control of gene expression, the quick dynamic response, and the high sensitivity to specific inducers, and

iv) The GRC robustness (i.e. a low sensitivity vs. undesired inducers). The WCVV advanced modelling framework of regulatory structures are able to maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. In other words, such a WCVV kinetic model of GERM-s/GRC-s must ensure the homeostasis (quasi-stationarity) of the regulated key-species, and quick recovery (with a trajectory of minimum amplitude) of the same steady-state (homeostasis) after a dynamic ('impulse'-like) perturbation, or of a novel steady-state (homeostasis) after a stationary ('step'-like) perturbation of one or more involved metabolites or nutrients.

The key element to such cell GRC dynamic simulators is the adopted kinetic model of the constituents GERM-s. A significant number of GERM-s reduced (lumped) models exists in the literature, based on a hypothetical mechanism [6,37,39,41,42,55,73,101-104]. Because the GERM structures are modular and repeatable in most of micro-organisms, according to the above discussion, Maria [1-3,6] built-up a "LIBRARY" of GERM-s (Figure 6). A reduced GERM structure usually includes no more than 10-14 reactions, thus ensuring a satisfactory trade-off between model simplicity and its predictive quality [7]. When constructing a GRC for a certain

cell metabolic pathway, there are two problems, which must be addressed properly [7].

A. How to choose the suitable GERM–s structures of the GRC chain, by screening among existing alternatives of the GERM-s "LIBRARY", for selecting simple structures which are suitable to confer the regulatory properties to the built-up GRC. Some examples are provided by [6,39,41,42,65,73,101,102]. Even if more sophisticated constructions are proposed in the literature [42,64], such a GERM selection must be based on their regulatory properties (that is quantitative performance indices P.I.-s defined and studied by Maria [1,3,7]) matching with the experimental data.

B. What rules should be applied to link such GERM-s to reproduce the cell system holistic behaviour and its self-regulation of all GERM–s and GRC-s. Following an extended *in-silico* analysis of the GERM–s included in the attached "LIBRARY", Maria [1,7] derived the main rule when constructing a modular GRC.

C. There are three modelling and control levels of the regulatory circuits

- A single gene expression module GERM (local, or individual level)
- The GRC including several GERM-s retaining in the regulation nodes protein complexes resulted from proteins (enzymes) interactions that promote a catalytically efficient sequence of reactions over the so-called 'channelling intermediate metabolites' [105,106]
- The whole-cell replication regulation to be included in the novel silicon whole-cell modelling concept WCVV, so that the whole cell model to be able to maintain intracellular homeostasis (i.e. quasi-constancy of the key-species concentrations) during its auto-catalytic growth on environmental nutrients present in variable amounts.

D. The following concepts derived from the experimental evidence should be included in the WCVV models of GERM-s and GRC-s [1-3,7]

All regulatory performance indices P.I.-s should be optimal at both GERM and GRC levels. That is because all cell metabolic processes occur with optimal performances to ensure the cell replication over an exact cell cycle. So, the metabolic reactions must occur with maximum reaction rates, with using minimum of resources (substrates, energy), and producing minimum amount of reaction intermediates; reactions and key-species homeostasis should be less influenced by the environmental perturbation even if involving simple GRC-s with a preferable cascade control of the gene expression that minimize the transition or recovering times of a quasi-steady-state QSS [see the proof of Maria [1,2,7,38]. Besides, the computed species sensitivities vs. environmental nutrient levels "$S(C_j, Nut_i) = \partial \ln C_j / \partial Nut_i$ should be minimum in a GRC structure. See Maria [1-3.7].

In all WCVV cell models the isotonicity hypothesis requires to consider the lumped genome, proteome, metabolome, due to their high concentrations (the so-called "cell ballast"), who exercises a big influence on the cell metabolism, on the volume increase, and on the treatment of perturbations. The genome/proteome replication was considered by Maria [3,7,39,41,42] in a simple way, by means of a lumped GERM. of simple structure. Simple examples can be provided, that is

i) The mercury-operon expression in *E.coli* [64,65]

ii) A genetic switch in *E.coli* [39,41,42]

iii) Consideration of the lumped proteome and genome replication in all WCVV cell models is mandatory in order to fulfil the isotonicity constraint

iv) In fact, such a constraint requires to consider all cell species (individually or lumped), because all species net reaction rates contribute to the cell volume increase, and to the instant cell dilution rate "Di" [section 5]. Some examples are presented by Maria [6,38,39,41,42,65]

v) According to [3,7,107], the modular WCVV dynamic models of GRC-s, of an adequate mathematical representation (deterministic or stochastic), seem to be the most comprehensive mean for a rational design of the regulatory GRC–s with desired behaviour. However, the lack of detailed information on reactions, rate's expression and intermediates make the extensive representation of the large-scale GRC-s difficult for both deterministic and stochastic approaches [108]

In spite of the gene expression process complexity, to be easy to use, most of the GERM kinetic models proposed in the literature present a simplified structure, by including only lumped species and reactions, but able to confer optimized regulatory P.I.-s to the GERM regulatory efficiency (see [1,3,7,38] for an extended discussion). Besides, as discussed in the section 5, the GERM WCVV models must fulfill some constraints, that is

i) The metabolic reaction stoichiometry

ii) Reaction rates must be maximal, but with rate constants limited by the diffusional processes and in agreement with the thermodynamic equilibrium steps

iii) The total enzyme (proteine) content of the cell is limited by the isotonicity condition

iv) The total cell energy (ATP) and reducing agent (NADH) resources are limited

v) The reaction intermediate level must be minimum

vi) The cell model at homeostasis must be stable, reaching again the steady-state after termination of a dynamic ('impulse'-like), or a stationary ('step'-like) perturbation (see [1-3,7,8,38] for details)

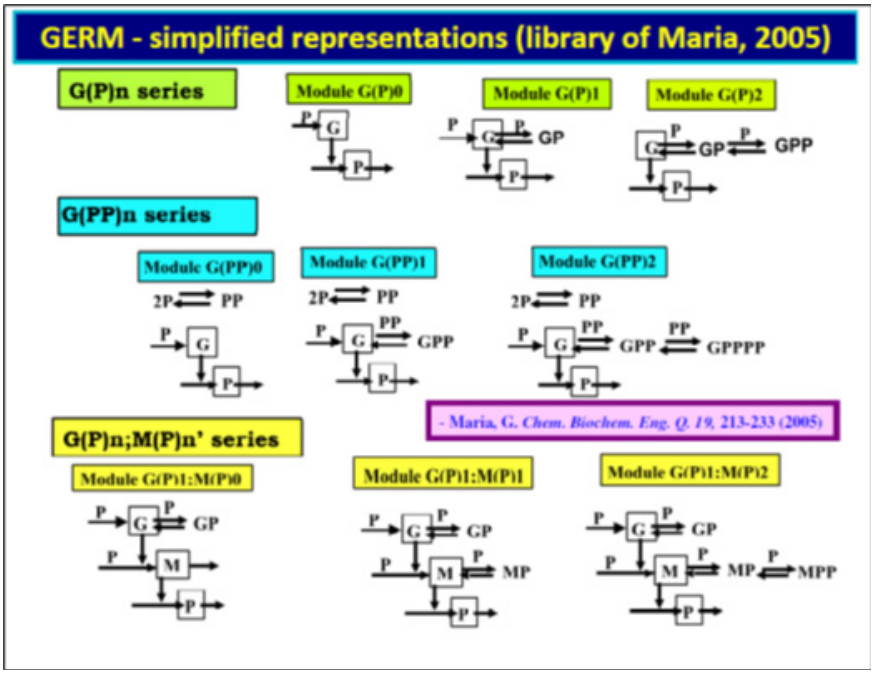

Figure 6: The LIBRARY of lumped modular models of GERM–s developed by Maria [1-3,6,37-41]. The library includes, in a template form, the kinetic models corresponding to simplified representations of a generic gene expression G/P regulatory module (GERM), assimilated with formal mutual catalytic actions of G and P. The horizontal arrows indicate reactions, while the vertical arrows indicate catalytic actions. The absence of a substrate or product indicates an assumed concentration invariance of these species. Notations: G = gene encoding P; M = mRNA.

[Up-row]. Simplified representation of the gene expression models corresponding to [G(P)n] regulatory module types. The transcriptional factor is the protein P itself, the self-regulation over the transcription and translation steps being lumped together. To improve the system homeostasis stability and self-regulation, despite perturbations in nutrients Nut*, and metabolites Met* (see Figure 7), or of internal cell changes, a very rapid buffering reaction G + P <===> GP(inactive) has been added, once time, or several times.

[Middle-row]. Simplified representation of the gene expression model corresponding to a [G(PP)n] regulatory module types. The transcriptional factor (TF) is the dimmer PP. The top-row observations also apply here.

[Down –row]. Simplified representation of the gene expression model corresponding to [G(P)1; M(PP)n] regulatory module types. The models account for the cascade control of the expression via the separate transcription and translation steps.

Notations: G= DNA gene encoding P; M= mRNA; P,PP= allosteric effectors of transcription/translation. The top-row observations also apply here. Adapted from Maria [38].

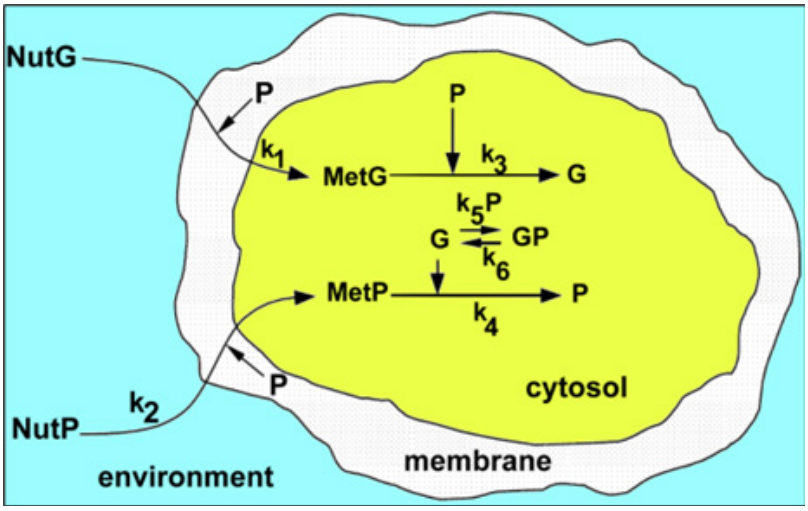

Figure 7: The reaction scheme of the [G(P)1] regulatory module GERM of the generic gene G expression used to exemplify the synthesis of a generic P protein in the *E. coli* cell, proposed by Maria [1,6,37]. To improve the system homeostasis, that is the quasi-invariance of key species concentrations (namely enzymes, proteins, metabolites), despite of external perturbations in nutrients {NutP. NutG} here, and metabolites {MetP, MetG}, or despite the cell internal changes in the species stationary concentrations), a very rapid "buffering" reaction G + P ⇔ (GP) inactive has been added. Adapted from Maria [6].

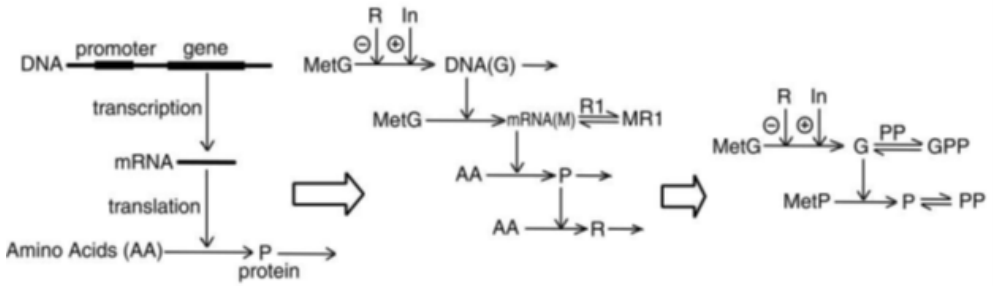

Figure 8: Protein P synthesis – simplified representations of a generic GERM regulatory module of [G(PP)1] type (see Figure 6). Horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; G = gene (DNA) encoding P; M = mRNA; R = repressor; In = inducer; MetG, MetP = metabolites used for the synthesis of G,P respectively); AA = amino-acids. According to the studies of Maria [1-3,6,7,41,42], this GERM structure with dimeric ttranscription factors (TF = PP) displays one of the best regulatory efficiency P.I.s (defined and tested by Maria [7]).

Protein synthesis by gene expression is a highly regulated process to ensure a balanced and flexible cell growth under indefinitely variate environmental conditions. How this very complex process occurs is partially understood, but a multi-cascade control of the transcription and translation steps (see some simplified GERM–s representations in Figure 6, Figure 7 & Figure 8, and of GRC-s in (Figure 5) with negative feedback loops seems to be the key element. Enzymes catalyzing the synthesis are allosterically regulated by means of positive or negative effector molecules (transcriptional factors TF), while cooperative binding and structured regulation amplify the effect of a change in an exo/endo-geneous inducer. Gene expression is also highly regulated to flexibly respond to the environmental stress. 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 [6].

How works such a simple GERM self-regulatory module? For instance, in the GERM of Figure 7, that reproduce the expression of a generic pair G/P, protein P plays multiple roles: is a permease controlling the nutrients {NutG, NutP} into the cell, but also metabolases catalyzing their transformation to inner metabolites {MetG, MetP} respectively (rate constants k1,k2), and is also a polymerase catalyzing the gene G synthesis (rate constant k3). In turn, the gene G catalyze (rate constant k4) the protein P synthesis. Protein P is also a transcription factor (TF), by controlling the gene G catalytic activity through a "buffering reaction", that is a very fast reversible reaction G + P <==> GP(inactive). When P is formed in too high amounts, it will inactivate G by means of higher amounts of GP. The result is a decrease of P synthesis rate, and a decrease of P concentration. In turn, lower amounts of P will direct the "buffer" reaction towards GP decomposition, which will increase G concentration and the regulatory loop is closed. This module is denoted by [G(P)1] to point out the existence of a single "buffer" reaction involving P as TF, that is only one regulatory effector.

As proved by Maria [1,7] the GERM-s efficiency depends on the GERM structure (Figure 6) and the number and the type of effectors. For instance, a more effective GERM is those of Figure 8, where the TF is a dimmer PP, present in small concentrations (1-5 nM, [65]). An even more effective GERM is those from the series [G(P)n; M(P)n1] of Figure 6 with a cascade control of the expression, the translation and transcription steps being considered separately, each step being controlled by a "buffer" reaction, that is G + P $\Leftrightarrow$ GP(inactive), and M + P $\Leftrightarrow$ MP(inactive), respectively.

The protein synthesis is 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. 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. The novel WCVV modelling framework hypotheses account for the perturbations effect inside cell, in a direct mode, or as a "secondary" perturbation transmitted via the cell volume to which all species contributes in an isotonic cell system.

The *modular approach* to analyze the gene expression assumes that the reaction mechanism and stoichiometry of various types of kinetic lumped modules are known, while the involved species (even if lumped) 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 procedure [109,110], or by exploiting the cell and module global optimal properties during identification steps [111,39]. 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 [6,7,38]. 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. In such a manner, linked modules are able to respond to changes in common environment and components such that each module remains fully regulated [1-2,7]. "The advantage of such a modular WCVV approach is the possibility to reduce the system model complexity and the size of the identification problem, by understanding, for instance, the gene expression response to a perturbation as the response of a few genetic regulatory loops instead of the response of thousands of genetic circuits in the metabolic pathway" [41,42,7].

Various types of kinetic modules can be analyzed individually as mechanism, reaction pathway, 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 (GRC) by applying certain principles and rules tested and discussed in detail by Maria [7], and then adjusting the network global regulatory properties. Accordingly, it is desirable to focus the metabolic regulation and control analysis on the regulatory/control features of functional GERM subunits than to limit the analysis to only kinetic properties of individual enzymes acting over the synthesis pathway.

The difficulty to precise the very large number of parameters in complex GRC-s leads to account for lumped representations, such as gene clustering and path structure reduction based on various system properties (stability, sensitivity, multiplicity). Consequently, to model individual GERM-s and GRC-s, at a molecular level, several authors [6,37,73,101,102] proposed simple mechanistic structures by using a modular approach, useful in simulating the hierarchical organization of cell regulatory networks (Figure 2). Some simple GERM models are given in (Figures 5-8). While Sewell, et al. [73], and Maria [1-3,6,7,42] used ordinary differential (ODE) kinetic models with continuous variables, based on classical common rate expressions (of Michaelis-Menten, or Hill type), Savageau [101], Hlavacek and Savageau [102] and [112,113] used simple rate expressions of power-law type (the so-called S-systems) obtained by recasting the elementary steps and intermediates in a lumped representation including apparent rate constants and reaction orders. Such apparent power-law models with fractional orders of reactions can generate a biased representation of the real process and suffer of a series of inconveniences related to the predicted distorted GERM properties (multi-stationarity, sub-optimal P.I.-s). By contrast, the mechanistic based GERM models [6,39,41,42] seem to be more robust, flexible and easily adaptable [1-3,7].

On the other hand, a too advanced lumping can lead to diminish some network properties (local stability strength, efficient responsiveness, flexibility; see the P.I.-s measures defined by Maria [3,7]). One disadvantage of using continuous variable formulations is the possibility of translating fractional concentrations to fraction of copy numbers. For instance, for a born *E. coli* cell volume of $[Vcyt,o] = 1.66 \cdot 10^{-15}$ L, one gene G copy number translates to 1 nM concentration. Such a concentration can be easily derived based on the WCVV model hypotheses of section 5.

Thus, when treated deterministically, the WCVV type kinetic models present solutions with *cell species concentrations* expressed in real, but *fractional numbers*, which, from a physical point of view, appears to be in contradiction with the reality, because one copy number can not be divided into pieces. How the WCVV models over comed this problem? For instance, in a GERM module of [G(P)1] type (Figure 7), the fractional copy numbers, e.g. [G]s = ½ nM, are assumed at homeostasis. Such fractional concentrations of cell species generated in the deterministic WCVV kinetic model approach with continues variables, can be interpreted in several ways:

a) Loosely either as time-invariant averages in a population of cells (e.g. that half of all cells contain 1 G-copy number), or,

b) As a time-dependent averages of single cells (e.g. that the cell contains 1 copy of G half time of the analysed cell cycle), or

c) As a result of the variable cell volume, to which all species contributes, according to the isotonicity constraint (Eq. (7,9), **Table 1** of section 5), in turn the cell volume influence all species concentrations according to the combined relationships (Eq.(10,11) of section 5), it results the isotonicity constraint (Pfeiffers' law of diluted solutions, [1-3,7]), of the form:

Table 1: The basic hypotheses of the "mechanistic silicon cell concept" materialized in the novel WCVV dynamic modelling framework in living cells of isotonic variable volume. Adapted from [6], with the details given by Maria [6,7,38,39,41,42,51,53,66,67].

Mass balance and state equations	Remarks
$\frac{dC_j}{dt} = \frac{1}{V}\frac{dN_j}{dt} - D_i C_j = h_j(\boldsymbol{C},\boldsymbol{k});$ $\frac{1}{V}\frac{dN_j}{dt} = \sum_{j=1}^{mr} v_{ij} r_j(C,k);\ j = 1,\ldots,n_s;$	continuous variable dynamic model representing the cell growing phase (ca. 80% of the cell cycle)
$V(t) = \frac{RT}{\pi}\sum_{j=1}^{n_s} N_j(t);$;	Pfeffer's law in diluted solutions [142]

$D_i = \frac{1}{V}\frac{dV}{dt} = \left(\frac{RT}{\pi}\right)\sum_j^{n_S}\left(\frac{1}{V}\frac{dN_j}{dt}\right)$	D_i = cell content instant dilution rate = cell volume logarithmic instant growing rate. Derived from the Pfeffer's law.
$\frac{RT}{\pi} = \frac{V}{\sum_{j=1}^{n_S} N_j} = \frac{1}{\sum_{j=1}^{n_S} C_j} = \frac{1}{\sum_{j=1}^{n_S} C_{jo}} =$ = constant.	constant osmotic pressure constraint
$\left(\sum_j^{all} C_j\right)_{cyt} = \left(\sum_j^{all} C_j\right)_{env}$	isotonic osmolarity constraint (see hypothesis "d")

Footnote. The WCVV model hypotheses are the followings:

a.-- negligible inner-cell gradients

b. -- open cell system of uniform content

c. -- semi-permeable membrane, of negligible volume and resistance to nutrient diffusion, following the cell growing dynamics

d. -- constant osmotic pressure π, (the same in cytosol "cyt" and environment "env"), thus ensuring the membrane integrity ($\pi_{cyt} = \pi_{env}$ = constant)[36]

e. -- nutrients and the overall environment concentration remain unchanged over a cell cycle (t_c)

f. -- logarithmic growing rate of average (apparent) $D_S = D_m = ln(2)/t_C$,Eq.(6), leading to an average volume growth of $V = Vo\ \exp(Dm \times t)$, over the 80% of the cell cycle [36]

g. -- homeostatic stationary growth of $\left(dC_j/dt\right)_s = h_j(\boldsymbol{C}_s, \boldsymbol{k}) = 0$

h. -- perturbations in the cell volume are induced by variations in species copy numbers (number of moles) under the isotonic osmolarity constraint: $V_{perturb}/V = \left(\sum N_j\right)_{perturb} / \left(\sum N_j\right)$; see the proof of Maria [7,66,67]

$$V(t) = \frac{RT}{\pi_{cyt}}\sum_{j=1}^{ns} N_j(t) \qquad (1)$$

The cell isotonicity constraint (section 5) imposes that all cell species to be included in the model, individually, or lumped.

Species concentrations in the cell are computed with the formula of Kurata and Sugimoto [114]

$$C_j = \frac{N_j}{N_A \times V_{cyt}} \qquad (2)$$

For instance, for a born *E. coli* cell, with an approximate volume [Vcyt,o] = 1.66 · 10^{-15} L, [115], the concentration of one gene G copy number at homeostasis has a value of

[G]s = $1/\left(6.022\times10^{23}\right)\left(1.66\times10^{-15}\right) \approx$ 10^{-9} mol/L = 1 nM.

Consequently, by combining Eq.(1) with Eq.(2), it results the following formula to evaluate the species concentrations inside the cell

$$C_j = \frac{N_j}{N_A \times \left(\frac{RT}{\pi}\sum_{i=1}^{all} N_i\right)} \qquad (3)$$

According to the Eq.(3) and isotonic system hypothesis of the WCVV model given in section 5, species concentrations could present fractional concentrations due to the various (long stationary, or short dynamic) perturbations into the cell, which are manifested by a variation in the number of moles

$\sum_{i=1}^{all} N_i$,

and of the cell volume V_{cyt}, and so in the C_j of Eq.(2). For instance, a suddenly increase in the species present in large amounts (for instance the lumped nutrients {NutP, NutG} or metabolites {MetP, MetG} in Figure 7) will lead to a significant increase of the total number of moles

$$\sum_{i=1}^{all} N_i ,$$

which, in turn, will lead to a significant increase in the V_{cyt}, and to a decrease of C_j in Eq.(2), particularly for the species "j" present in small amounts (such as G, GP of the GERM kinetic models).

When elaborating a protein synthesis GERM model, different degrees of simplification of the process complexity can be followed. For instance, the gene expression (see the scheme of Figures 7&8) 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 (Figure 9), 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 (P, or R) reversibly binding the catalyst via fast and reversible reactions (the so-called 'buffering' reactions). This simplest regulation scheme can be further detailed (Figure 6) 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 (Figure 9).

The effector (R), 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 respectively, 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 [1-3,7] 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 mentioned alternatives (deterministic, stochastic, Boolean), and then analyzed as functional efficiency by means of the performance indices (P.I.-s) defined by Maria [1-3,7].

A few number of GERM-s from Figure 6 have been studied to compare their regulatory efficiency by several authors [72,73]. Unfortunately, their studies have been done by using the classical CVKM approach with exemplifications from *E. coli* cells, leading to distorted and erroneous results, as proved in section 6. On the contrary, Maria [1-3,7] studied similar structures properties but using the novel WCVV approach with superior results and conclusions.

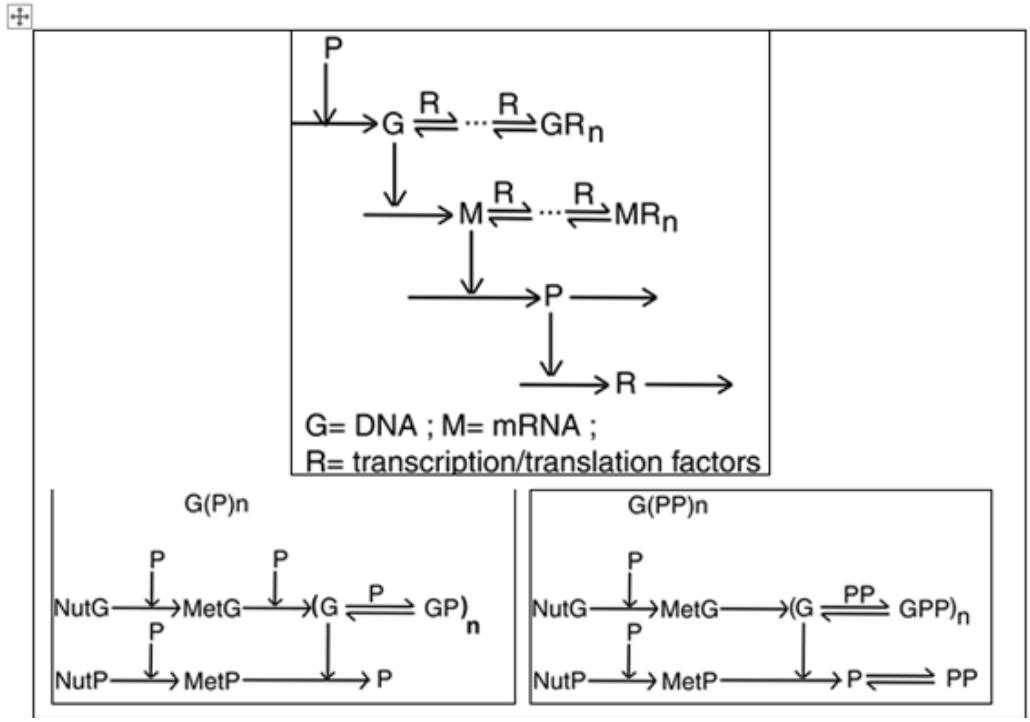

Figure 9: Protein P synthesis. Simplified representations of a generic gene expression G/P regulatory module (GERM). The horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; absence of a substrate or product indicates an assumed concentration invariance of these species; G= gene (DNA) encoding P; M= mRNA.

Up-row: simplified representation of the gene expression self-regulation over the transcription and translation steps. The model corresponds to a [G(R)n; M(R)n] regulatory module type of Figure 6. R= allosteric effectors of the transcription / translation.

Down-row: two types of GERM simplified representations of the self-regulated protein P synthesis: [G(P)n] (left) and [G(PP)n] (right). Adapted from Maria [6,38] by courtesy of CABEQ JI.

Simplified representations of Figure 7 and Figure 9 include the essential nutrient lumps (NutP, NutG) used to obtain the protein and DNA precursor metabolites (MetP, MetG) respectively, and intermediates involved in the reactions controlling the transcriptional and translation steps of the P synthesis. The GERM-s *module nomenclature*, is those proposed by [6,72]. The abbreviations of (Figure 6) are of the generic type $[L_1(O_1)n_1;...; L_i(O_i)n_i]$. Such a notation includes the assembled regulatory units "$L_i(O_i)n_i$". One regulatory unit "*i*" is formed by the component L_i (e.g. enzymes P, or even genes G, M, etc.) at which regulatory element acts, and n_i = 0,1,2,... number of 'effectors'/TF species (O_i) (e.g. P, PP, PPPP, R, RR, RRRR, etc.) binding the 'catalyst' L. For instance, a [G(P)2] unit of Figure 6 includes two successive binding steps of G with the product P, that is G + P ⇔ GP + P ⇔ GPP, all intermediate species GP, GPP, being inactive catalytically, of equal concentrations at homeostasis ([G] = [GP] = [GPP] = 1/3 nM) while the mass conservation law is all time fulfilled, i.e. $\Sigma_{i=0}^{2}[G(P_i]$ = constant. Such a representation accounts for the protein concentration diminishment due to the cell-volume growth dilution effect, but could also include protein degradation by proteolysis. It is also to observe that such GERM–s lumped models try to account for the essential properties of the gene expression, which is a highly self-/cross- regulated and mutually catalyzed process by means of the produced enzymes / effectors. As depicted in Figure 6,7 for the [G(P)1] module case, the protein P synthesis is formally catalysed by its encoding gene G. In turn, P protein formally catalyses the G synthesis, but also modulate the G catalyst activity via the fast buffering reaction G + P <===> GP. The GERM model structure can be extended according to experimental information and accounted individual or lumped species. The right choice of a GERM in a GRC depends on its regulatory properties and according to the linking rules proposed, both being reviewed by Maria [7].

The effective GERM-s which belong to the series [G(P)n; M(P)n1] of Figure 6 with a cascade control of the expression tries to reproduce more accurately, even if in a lumped way, 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, P is synthesized in a reaction catalyzed by M (translation). Each step is controlled by a "buffer" reaction, that is G + P ⇔ GP(inactive) ⇔ ... ⇔ G(P)n, and M + P ⇔ MP(inactive) ⇔ .. ⇔ M(P)n1, respectively. Such a supplementary control of M (i.e. mRNA) activity is proved to be a more effective mean of regulating protein synthesis [6,37,72,116].

It is also to mention the way by which the rate constants in the rapid buffering reactions are estimated, that is in a WCVV model for the "effector"/"buffering reaction" of the following form [6]:

$$L + O \underset{k_{diss}}{\overset{k_{bind}}{\rightleftarrows}} LO\,;K_{LO} = \frac{k_{bind}}{k_{diss}} = \frac{C_{LO,s}(1+\frac{D_m}{k_{diss}})}{C_{L,s}C_{O,s}} \qquad (4)$$

When estimating the WCVV kinetic model parameters from the model QSS condition Eq.(10,13), several optimal regulatory properties should be imposed, as discussed by Maria [7], for instance:

a) Physical meaning of the solution, $[\hat{k}, \hat{C}_s] > 0$. Here, the superscript " ^ " denotes the estimated value.

b) In the case of the WCVV kinetic models of GERM-s (sections 5,6), some constraints should be imposed to ensure the optimal regulatory efficiency of the GERM-s. For instance:

i) Minimum recovering time of the stationary concentrations (homeostasis) after a dynamic ('impulse'-like) perturbation in a key-species Maria [1-3,6,66,67]

$$[\hat{k}, \hat{C}_s] = \arg\min(\tau_P)$$

Here, notation τ_P denotes the recovering time of the [P]s stationary concentration after an impulse-like perturbation of a ±10%[P]s (where P refers to the key-protein P in a GERM model described in this section 3. This recovering time was determined by using the WCVV simulation model, with a tolerance of 1-5%[P]s [6]..

ii) The active $[L]_{active}$ and inactive $[L]_{inactive}$ forms of the regulatory element (that is the "catalyst" here) must be equal. In other words, according to the above GERM nomenclature:

$[L]_{active}/[L]_{total}$ = 1/ 2.

Here, "L" (e.g. enzymes P, genes G, or M = mRNA) is a GERM component at which the regulatory elements and the transcription factors (TF) act (Figure 6-9).

As proved by Maria [1,6,37,38,39,41,42], and by Sewell, et al. [73] such a constraint ensures a maximum efficiency of the GERM self-control of the P-synthesis.

iii) The all forms of the regulatory elements should be constant (the mass conservation), that is

$$\sum_{i=0}^{n}[L(O_i)] = const\,ant \quad ; \quad \sum_{i=0}^{n}[G(P_i)] = const\,ant \quad ;$$

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

As discussed by Kholodenko et al. (1998), fast buffering reactions are close to equilibrium and have little effect on metabolic control coefficients of GERM-s. As a consequence, rate constants of such rapid reactions are much higher than those of the core synthesis and dilution rates. To reduce the size of the unknown vector during rate constant estimation of a GERM model, in Eq. (4) large values of $k_{diss} > D_m$ can be postulated (5 to 7 orders of magnitude higher), where "*Dm*" is the average logarithmic growing

rate of the cell volume $D_m = \ln(2)/tc$ (where *"tc"* = cell cycle) [6]. Numerical rules used to lump species and/or reactions are discussed by Maria [38,7].

4. The Classic (default) Kinetic CVKM Modelling Framework of GERM-s/GRC-s

The term "whole-cell" (ideally) assumes that all the cell species (from cytosol and membrane) are included in the model (individually, or lumped), with the following copy numbers:

$$N = \left[N_1 \quad \ldots \quad N_j \quad \ldots \quad N_{ns} \right],$$

And the corresponding concentrations:

$$C = \left[C_1 \quad \ldots \quad C_j \quad \ldots \quad C_{ns} \right] = \left[\frac{N_1}{V} \quad \ldots \quad \frac{N_j}{V} \quad \ldots \quad \frac{N_{ns}}{V} \right]$$

where "ns" denotes the number of cell species (taken individually or lumped). Then, the kinetic model of a cell process can be formulated in the "classical" (default) approach as is Eq. (5a) in terms of numbers of moles (copy numbers).

Such default "*Whole-Cell Constant Volume" (CVKM)* classical continuous variable ODE kinetic models of the cell processes, do not explicitly consider the cell volume exponential increase during the cell growth. When the continuous variable CVKM dynamic models are used to model the cell processes, the default-modelling framework is those of the CVKM formulation Eq.(5b) which results from Eq.(5a), by using the species concentration definition as $C_j = N_j/V$. (where V denotes the cell cytosol including the membrane). This transformation to the default kinetic model Eq.(5b) is valid *if and only if* the cell system presents a constant volume. Such a hypothesis is not valid for a cell system, of which the cell volume (cytoplasm) is growing continuously, doubling after one cell cycle.

$$\frac{1}{V(t)} \frac{dN_j}{dt} = \sum_{i=1}^{nr} s_{ij} r_i \left(\mathbf{N}/V, \boldsymbol{k}, t \right) = h_j \left(\mathbf{C}, \boldsymbol{k}, t \right) \quad \textbf{(5a)}$$

$$\frac{d\left(N_j/V\right)}{dt} = \frac{dC_j}{dt} = \sum_{i=1}^{nr} s_{ij} r_i \left(\boldsymbol{N}/V, \boldsymbol{k}, t \right) = h_j \left(\boldsymbol{C}, \boldsymbol{k}, t \right)$$

(5b)

$$\left(\frac{dC_j}{dt} \right)_S = \sum_{i=1}^{nr} s_{ij} r_i \left(\mathbf{N}/V, \boldsymbol{k}, t \right) = h_{j,s} \left(\mathbf{C}_S, \boldsymbol{k}, t \right) = \boldsymbol{0}, \; j = 1, \ldots, ns \quad \textbf{(5c)}$$

The rate constants (vector "***k***") are estimated from the stationary conditions, by solving the nonlinear model (5c) with assumed known stationary concentrations [7]. The model Eq.(5c) characterizing the cell steady-state is obtained from Eq.(5b) by setting the derivatives on the left side to zero.

Such CVKM models, unfortunately largely present in the dedicated literature, do not explicitly consider the cell volume exponential increase during the cell growth. When such dynamic models Eq.(5a-c) are used to simulate the cell enzymatic processes, wrong results are expected because the hypothesis of a constant cell volume is wrong. To partially overcome these shortcomings, some CVKM models eventually account for the cell-growing rate as a pseudo-'decay' rate of key-species (often lumped with the degrading rate) in a so-called 'diluting' rate [1-3,6,7].

The formulation Eq.(5a-5c) assumes a homogeneous volume of the cell system with no inner gradients or species diffusion resistance, similar to a batch ideal reactor [117-119]. The used reaction rate expressions for the metabolic reactions are usually those of ChBRE, that is of the extended Michaelis-Menten or of Hill type [1-3,42,44,120]. Being very over-parameterized and strongly nonlinear, parameter estimation of such CVKM models from quasi-steady-state (QSS) Eq.(5c), and homeostatic data (assumed known species stationary concentrations) in the presence of multiple constraints translates into a mixed integer nonlinear programming problem (MINLP) difficult to be solved because the search domain is not convex, and the model is also highly nonlinear [13,109]. Lumping algorithms of species/reactions in the kinetic models are given by Maria [7,109,110], but they are difficult to be applied for the cell models due to their extended structure.

Such a classic CVKM dynamic model Eq.(5a-c) might be satisfactory for modelling many cell subsystems, but not for an accurate modelling of cell GRC–s and holistic cell properties under perturbed conditions, or the division of cells [36,121], by distorting very much or even misrepresenting the prediction results, as proved by Maria [1-3,51,65], and in the below section 6, for both stationary, or perturbed cell growing conditions. Here it is to observe that Eq.(5b) can be derived from Eq.(5a) *if and only if* the cell system volume (V) (cytoplasm plus membrane) is constant. In the chemical systems, such a constraint is usually fulfilled, but not in a biochemical one, as long as the cell volume doubles during a cell cycle (tc), with the following average (apparent) cell dilution rate (Dm)

$\frac{dV}{V} = D_m t$, which, after integration over a cell cycle is leading to:

$$\frac{1}{V} \frac{dV}{dt} = D_i;$$

By averaging over cell cycle, it results: $\int_{V_0}^{2V_0} \frac{dV}{V} = D_m \int_0^{t_c} dt$;

$$D_m = \frac{\ln(2)}{t_c}; V\left(t = t_c\right) = 2 \cdot V_o \Rightarrow V = V_o \exp\left(+\mathrm{D_m} t\right), \; for 0 \le t \le t_c \tag{6}$$

Where: index "o" denotes the initial volume of the newborn cell; D_i = cell content instant dilution rate = cell volume logarithmic instant growing rate; D_m = cell content average (apparent) dilution rate = cell volume logarithmic average growing rate.

5. The "mechanistic silicon cell" and the novel kinetic modelling framework, WCVV, of GERM-s/GRC-s able to maintain intracellular homeostasis while growing autocatalytically on environmental nutrients present in variable amounts

Here it is worth notice that living cells are extraordinarily complex machines, with an astronomical complex organization (Figure 2), that exhibit remarkable behaviors including, most predominately, their self-replication. As discussed by Maria [1-3,7,24], living cells are self-replicating complex biological structures, able to convert environmental nutrients to replicate the cell content in exactly one cell cycle. Cells present such a highly sophisticated structure (Figure 2), involving O($10^{3\text{-}4}$) components, O($10^{3\text{-}4}$) transcription factors (TF-s), activators, and inhibitors, and at least one order of magnitude higher number of (bio)chemical reactions, all ensuring a fast adaptation of the cell to the changing environment. Cell is highly responsive to the environmental stimuli and highly evolvable by self-changing its genome/proteome and metabolism to obtain an optimized and balanced growth with using minimum of resources (nutrients/substrates).

Although much progress has been made in understanding the mechanism(s) underlying this behavior, further efforts on both experimental and theoretical fronts will be required to deepen our understanding of it. Towards this end are efforts to simulate global cellular behavior starting from the molecular level of detail [17,18,28-31,122]. "Ideally, such whole-cell models would integrate massive amounts of biological information into compact, unambiguous and testable hypotheses which could be used to explore the entire panoply of cellular relationships in both healthy and diseased states" [32,34-36,123]. Much remains before useful comprehensive mathematical models of cellular behaviour with predictive power can be realized. This is due to the large number of unknown kinetic parameters of such extended cell kinetic models (rate constants, and unknown stationary concentrations of some individual species), and ambiguities in the biochemical mechanism occurring within a cell hinder fast progress in this direction, even if recent progresses have been made (reviews of Maria [1-3,7,24]).

Beside, there is also no consensus as to what modelling framework will ultimately prove most effective. By framework, we mean a group of general assumptions and procedures that can transform a physico-chemical model of a cell into a set of mathematical expressions (differential mass balance equations ODE, accompanied by algebraic equations/relationships). Such a framework would be applicable to all such cell models (CCM, individual GERM-s, or GRC-s), regardless of the particular set of (bio)chemical reactions and assumed components (individual or lumped species. Given the enormous complexity and unknown aspects of living systems, formulating predictive whole-cell molecular-level models will require complete and massive genomic, transcriptomic, proteomic, and metabolomic data sets [124,125]. Unfortunately, such data sets are not generally available, precluding significant advances in whole-cell modeling efforts. However, it is not simply the lack of data that inhibits such advances, but also the modeling framework [1-3,7,24].

Excellent progress has been made in developing computational approaches for modeling whole-cell processes [29,72,73,122,126-130]. However, the constant volume assumptions [see Eq.(5a-5c)] made by all these modelling trials may not be appropriate for modeling all of the processes occurring in whole cells, processes which ultimately convert environmental nutrients into another copy of the original (or newborn) cell as a product of cell growth and division. As proved by Maria, et al. [51,53], and in the below section 6, the classic ("default") math modelling approach of cell processes [that is the CVKM, "constant-cell-volume, non-isotonic system", Eq.(5a-c)] is erroneous compared to the novel WCVV modelling framework, leading to distorted and wrong simulation results and conclusions as proved in section 6, and by Maria et al. [7,51,53]. In such a way, the novel WCVV kinetic silicon cell modelling approach replaces the *classical (default) CVKM* modelling approach, thus *correcting the distorted* and false/wrong predictions of these cell kinetic models, unfortunately largely used in the literature.

To overcome this major drawback of the classical (default) kinetic models of a CVKM type, when applied to model cell processes, some awkwards trials have been reported. For instance, Perret and Levey [131] developed a model of cell volume expansion, in which a cell was assumed to house an un-catalysed metabolic pathway that converted nutrients into metabolites and to expand in proportion to the amount of cellular components. Such a cell could exist in an expanding steady-state in which volume and amounts of metabolites increased exponentially while concentrations remained invariant. Similar situations have been analyzed by Grainger et al. [132,133], Aris [134], Kacser and Beeby [135], Brumen et al. [136], Werner and Heinrich [137], and by Joshi and Palsson [138-141]. Although these approaches are more realistic than those which assume constant-volume, their volumes expand indefinitely in contrast to real cells which exist in a growth phase only until a critical size is achieved (at ca. 80% of the cell cycle [36]).

Thereafter, they shift into a division phase in which they ultimately divide into two newborn daughter cells, each equivalent to the original cell in terms of volume and cellular content. These two phases, growth and division, together comprise the cell cycle. The basic hypotheses of the WCVV modelling framework are presented in Table 1 being introduced by Maria [1-3,5-7]. Briefly, they can be formulated as follows:

a) The cell system consists in a sum of hierarchically organized components, e.g. *metabolites, genes DNA, proteins, RNA, intermediates,* etc. (interrelated through transcription, translation and DNA replication and other processes) (Figure 2). The cell is separated from the *environment (*containing *nutrients)* by a *membrane* (Figure 7).

b) The membrane, of negligible volume, presents a negligible resistance to nutrient diffusion; the membrane dynamics being neglected, and assumed to follow the cell growing dynamics.

c) The cell is an isothermal system with an uniform content (perfectly-mixed case). Species behave ideally, and present uniform concentrations within cell. The cell system is not only homogeneous but also isotonic (constant osmotic pressure), with no inner gradients or species diffusion resistance. Formally, it is similar to a batch/fed-batch biological reactor [117-119].

d) The cell is an open system interacting with the environment through a semi-permeable membrane.

e) To better reproduce the GERM properties (inter-connected with the rest of the cell (Figure 7,8), the other cell species are lumped together in the so-called "cell ballast"**.**

Thus, the adopted Pfeiffers'law of diluted solutions for the cell content is fulfilled, that is [1-3,5-7,66,67, 142]:

$$\text{Pfeiffers'law: } \pi_{cyt} V = \sum_{j=1}^{ns} N_j RT .$$

It results:

$$V(t) = \frac{RT}{\pi_{cyt}} \sum_{j=1}^{ns} N_j(t),$$

Which, is valid initially, but also at any time, that is:

$$\frac{RT}{\pi} = \frac{V(t)}{\sum_{j=1}^{ns} N_j(t)} = \frac{1}{\sum_{j=1}^{ns} C_j} = \frac{1}{\sum_{j=1}^{ns} C_{j,0}} = constant \quad (7)$$

Such a constraint imposes that all cell species to be included in the kinetic model, individually or lumped, especially the lumps of high concentrations (lumped genome/proteome, lumped nutrients/metabolites). In this context, Maria and Luta [65], and Maria [42] proposed that lumped genome and proteome replication to be considered in a WCVV cell model in the form of the simplest GERM model, of [G(P)1] type (Figure 6,7), in order to "mimic" the cell content regulatory properties.

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

g) When studying an individual P-synthesis regulatory module (i.e. gene expression regulatory module – GERM), the other cell species are lumped together in the so-called "cell ballast" [6,38,39,41,42]. For instance {MetP,MetG} in Figure 7.

h) The inner cell osmotic pressure (π_{cyt}) is constant, and all time equal with the environmental pressure, thus ensuring the membrane integrity (π_{cyt} = π_{env} = constant). As a consequence, the isotonic osmolarity under isothermal conditions leads to the equality $RT / \pi_{cyt} = RT / \pi_{env}$, derived from the fulfilment of the Pfeiffers'law.

Otherwise, the osmosis will eventually lead to an equal osmotic pressure $\pi_{cyt} = \pi_{env}$. Even if, in a real cell, such equality is approximately fulfilled, due to the environmental 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, increased amounts of polymerases, and permeases are then used to import increasing amounts of nutrients. The net result is an exponential increase of cellular components in time, at the same rate with the cell volume increase which translates, through isotonic osmolarity assumption, into an exponential increase in volume with time, Eq.(7) [1-3,6,7,66,67]. The overall concentration of cellular components is time-invariant (homeostasis), because the rate at which cell-volume increases equals that at which overall number of moles increases, leading to a constant $\left(\sum_{j=1}^{n_S} N_j \right) / V$ ratio in Eq. (7), and Table 1.

i)The species concentrations $C_j(t) = N_j(t)/V(t)$ evaluated at the cell level are usually expressed in nano-molar units (nM), being computed with the Eq. (2).

j)Cell volume doubles over the cell cycle period (t_c), with an average logarithmic growing rate of $D_m = ln(2)/t_c$, given by Eq. (6).

k) For stationary growing conditions, species synthesis rates are equal to first-order dilution rates ($D_s C_{js}$), leading to time-invariant species concentrations, that is the quasi-steady-state (QSS) homeostatic conditions in Eq.(10), of $\left(dC_j/dt\right)_s = 0$. The obtained nonlinear algebraic set Eq. (13) is used to estimate the rate constants *k* from the known (measurable) stationary concentration vector $\boldsymbol{C_s}$., with also imposing some optimal regulatory properties of the GERM/ GRC analyzed system [see the discussion below Eq.(4)].

l) When treated deterministically, the WCVV type kinetic models present solutions with *cell species concentrations* expressed in real, but *fractional numbers*, which, from a physical point of view, appears to be in contradiction with the reality, because one copy number can not be divided into pieces. How the WCVV models over comed this problem? For instance, in a GERM module of [G(P)1] type (Figure 7), the fractional copy numbers (e.g. [G]s = ½ nM) in section 6. Such fractional concentration of cell species generated in the deterministic WCVV kinetic models with continues variables, can be interpreted in several ways, as discussed and exemplified in section 3, by using Eq.(1-3).

The core relationship of the WCVV modelling framework relates to the dependence of the species inner concentration on the both variable number of moles (copynumbers), and cell volume, that is:

$$C_j(t)=\frac{N_j(t)}{V_{cyt}(t)} \quad \textbf{(8)}$$

Consequently, Eq. (5b) should be re-written for a variable volume, in the following form:

$$\frac{dC_j(t)}{dt}=\frac{d}{dt}\left(\frac{N_j(t)}{V(t)}\right)=\frac{V\dfrac{dN_j}{dt}-N_j\dfrac{dV}{dt}}{V^2}=\frac{1}{V}\frac{dN_j}{dt}-C_j\frac{d\left(\ln(V)\right)}{dt}=$$

$$\frac{1}{\mathrm{V}}\frac{dN_j}{dt}-D_iC_j=h_j(C,k,t); \qquad (9)$$

With the instant cell dilution rate of $D_i=\frac{1}{V}\frac{dV}{dt}=\frac{d(\ln(V))}{dt}$

In a more strict sense, by combining Eq.(9) with Eq.(5b), it results the WCVV novel model formulation, as follows:

$$\frac{dC_j(t)}{dt}=\frac{1}{V}\frac{dN_j}{dt}-C_j\frac{d\left(\ln(V)\right)}{dt}=\sum_{i=1}^{nr}s_{ij}r_i-\mathrm{D_i}C_j=h_j(C,k,t)$$

Where: $\frac{1}{V}\frac{dN_j}{dt}=r_j,\ \mathrm{j}\ =\ 1,\ldots,\mathrm{ns}.$ (10)

It is to observe that the instant cell dilution rate "*Di*" defined in Eq.(9), is structurally different from the average (apparent) cell dilution rate "*Dm*" defined and evaluated by Eq.(6). The instant dilution is dependent on the all reaction rates of the cell, being a holistic property. It results from derivation of the Pfeiffers'law Eq.(1), that is:

Which, by derivation and division with *V* is leading to [6,7]:

$$D_i=\frac{1}{V}\frac{dV}{dt}=\left(\frac{RT}{\pi}\right)\sum_{j=1}^{ns}\left(\frac{1}{V}\frac{dN_j}{dt}\right) \quad \text{where: (11)}$$

$$\frac{1}{V}\frac{dN_j}{dt}=r_j,\ \mathrm{j}\ =\ 1,\ldots,\mathrm{ns}.$$

For comparison, the average (apparent) cell logarithmic dilution rate *Dm* is given by Eq.(6) in the form of: $D_m=\ln(2)/t_c$

By contrast to the average D_m, the cell instant dilution rate D_i is linked and depends on the all reaction rates r_j of the cell species *Nj* (taken individually, or lumped).

As revealed by the Pfeffer's law Eq.(1) in diluted solutions [142], the volume dynamics is linked to the molecular species dynamics under isotonic and isothermal conditions by means of relationship Eq.(11). Consequently, as proved by Eq.(11), the instant dilution rate D_i results as a sum of the reacting rates of all cell species (individual or lumped), The RT/π term can be easily deducted in an isotonic cell system, from the fulfilment of the following invariance relationship Eq.(12), derived from the Pfeffer's law in diluted solutions, by accounting for the mass conservation law:

$$V(t)=\frac{RT}{\pi}\sum_{j=1}^{ns}N_j(t)$$

$$\frac{\mathrm{RT}}{\pi}=\frac{V(t)}{\sum_{j=1}^{ns}N_j(t)}=\frac{1}{\sum_{j=1}^{ns}C_j}=\frac{1}{\sum_{j=1}^{ns}C_{j,0}}=\ \text{constant} \quad (12)$$

The cell volume doubles over the cell cycle period (t_c), with an average logarithmic growing rate of $D_m=\ln(2)/t_c$ given in Eq.(6).

For stationary balanced growing conditions (that is at quasi steady-state, QSS), species synthesis rates are equal to first-order dilution rates (DiC_j), leading to time-invariant species concentrations (i.e. homeostatic conditions, $\left(dC_j/dt\right)_s=0$). Thus,

from Eq. (10), and combining with Eq.(12), one obtains the WCVV model:

$$\left(\frac{dC_j(t)}{dt}\right)_s = \left(\frac{1}{V}\frac{dN_j}{dt}\right)_s - D_s C_{js} = h_{js}(C_s, K, t) = 0;$$

$$j = 1,\ldots,n_s; \quad D_s = \left(\frac{RT}{\pi}\right)\sum_{j=1}^{ns}\left(\frac{1}{V}\frac{dN_j}{dt}\right)_s \qquad (13)$$

It is to observe that in the WCVV model the instant cell dilution D_i in Eq. (10), is different from the stationary cell dilution rate D_s of Eq. (13), which, in turn is roughly equal to the average one D_m of Eq.(6). As proved by Maria et al. [7,51,53] the use of D_m instead of the instant D_i can lead to biased results, and wrong simulation results. This is also valid for the use of classical CVKM formulation instead of the novel WCVV one (section 6).

6. Proving Superiority of using the WCVV Novel Modelling Framework Compared to the Classic ("Default") CVKM Kinetic Models when Simulating Gene Expression Regulatory Dynamics

To prove, in a simple way, the superiority of WCVV modelling framework (section 5) vs. the classic ("default") CVKM modelling rules (section 4), when simulate the cell key-species dynamics especially for GERM / GRC systems, one presents here an example of how wrong predictions can offer a classic CVKM model when simulating a gene expression regulatory module of simplest [G(P)1] type. To point-out the essential conceptual differences between the novel WCVV modelling framework and the „default" CVKM models, and how huge are the differences in the simulated results, one considers the case of the simplest regulated system of a generic gene (G) expression regulatory module GERM from the *E. coli* cell, used for the synthesis of its encoding protein P, and one mimicks the cell homeostasis and GERM response to a dynamic ('impulse'-like) perturbation. In this respect, one adopts the reaction scheme of Maria [6] corresponding to a regulatory module of [G(P)1] type, according to the simplified reaction pathway of Figure 7, that is:

$$\begin{cases} NutG + P \xrightarrow{k_1} MetG + P, & r_1 = k_1 C_{NutG} C_P \\ NutP + P \xrightarrow{k_2} MetP + P, & r_2 = k_2 C_{NutP} C_P \\ MetG + P \xrightarrow{k_3} G + P, & r_3 = k_3 C_{MetG} C_P \\ MetP + G \xrightarrow{k_4} P + G, & r_4 = k_4 C_{MetP} C_G \\ P + G \xrightarrow{k_5} GP, & r_5 = k_5 C_P C_G \\ GP \xrightarrow{k_6} G + P, & r_6 = k_6 C_{GP} \end{cases} \qquad (14)$$

To keep enough generality, the proof will concern a generic G/P synthesis regulation module, with species concentrations characteristic to the *E. coli* cell [6]. The reaction rate expressions of the lumped model are given in Eq.(14). The kinetic model of this GERM derived by using the classic ("default") formulation CVKM ("constant volume like", of section 4), in terms of *Nj* (species "*j*" number of moles), and *Cj* (species "j" concentrations, in nM) is given by Eq.(15).

CVKM classic ("default") kinetic model(*)

$$\begin{cases} \frac{1}{V}\frac{dN_{MetG}}{dt} = r_1 - r_3 \\ \frac{1}{V}\frac{dN_{MetP}}{dt} = r_2 - r_4 \\ \frac{1}{V}\frac{dN_P}{dt} = r_4 - r_5 + r_6 \\ \frac{1}{V}\frac{dN_G}{dt} = r_3 - r_5 + r_6 \\ \frac{1}{V}\frac{dN_{GP}}{dt} = r_5 - r_6 \end{cases} \Rightarrow \begin{cases} \frac{dC_{MetG}}{dt} = r_1 - r_3 \\ \frac{dC_{MetP}}{dt} = r_2 - r_4 \\ \frac{dC_P}{dt} = r_4 - r_5 + r_6 \\ \frac{dC_G}{dt} = r_3 - r_5 + r_6 \\ \frac{dC_{GP}}{dt} = r_5 - r_6 \end{cases} \qquad (15)$$

[*] with the cell average dilution of Eq.(6): Ds = Dm = Ln(2)/ *tc* = 6.93 × 10^{-3} 1/min, for a cell with *tc* = 100 min.

In the classic ("default") CVKM kinetic model formulation Eq.(15), despite the cell volume (*V*) is variable with the time, by doubling after a cell cycle (*tc*) Eq.(6), in the wrong way the volume variable was considered constant by re-writing the term

$$\frac{1}{V}\frac{dN_j}{dt}$$

as .

$$\frac{dC_j}{dt}$$

This is a capital (huge) *error* of these classic CVKM kinetic models which, as it will be further proved, will lead to erroneous simulation results and wrong predictions [7,51,53].

By contrast, the kinetic model of this system written in the novel WCVV modelling framework (section 5) is given in Eq.(16).

WCVV novel modelling framework (*)

$$\begin{cases} \frac{dC_{MetG}}{dt} = r_1 - r_3 - D_i \cdot C_{MetG} \\ \frac{dC_{MetP}}{dt} = r_2 - r_4 - D_i \cdot C_{MetP} \\ \frac{dC_P}{dt} = r_4 - r_5 + r_6 - D_i \cdot C_P \\ \frac{dC_G}{dt} = r_3 - r_5 + r_6 - D_i \cdot C_G \\ \frac{dC_{GP}}{dt} = r_5 - r_6 - D_i \cdot C_{GP} \end{cases} \qquad (16)$$

[*] The instant cell dilution rate " D_i " is given by Eq.(17-18).

One essential difference of the novel WCVV model Eq.(16) is due to the isotonicity constraint of the cell system, Eq.(6). Due

to such a requirement to keep a variable cell volume during the whole cell cycle, Eq.(10-11-12-13), the net difference is done by the introduction of the instant cell dilution rate (*Di*), Eq.(11) which reflects the major influence of the volume continuous increase on the species reaction rates, and vice-versa. The (*Di*) value necessary to solve the model Eq.(16) is given by Eq.(11,12) which, for this case study translates in Eq.(17).

$$D_i = \frac{RT}{\pi}\sum_{j=1}^{ns}\frac{1}{V}\frac{dN_j}{dt} = \frac{RT}{\pi}\sum_{j=1}^{ns} r_j$$

Is written for this case study as follows:

$$D_i = \frac{RT}{\pi}\left(r_1 + r_2 - r_5 + r_6\right), \text{ that is,} \quad (17)$$

by substituting the rates expressions, it results:

$$r_1 + r_2 - r_5 + r_6 =$$
$$k1\times C_{NutG}\times C_P + k2\times C_{NutP}\times C_P - k5\times C_G\times C_P + k6\times C_{GP}$$

In the above Eq.(17), the followings units are used: *Cj* [nM]; *V* [L]; *Nj* [n-moles]; *rj* [n-moles/L/min]; *Di* [1/min]; π [atm]; *T* [K]; *R* = [L.atm/n-moles.K]; *t* [min]. The required constant RT/π is evaluated from Eq.(12) from using the all known initial (stationary) values of the cell species concentrations, that is:

$$\frac{RT}{\pi} = 1\Big/\sum_{j=1}^{ns} C_{j,o} = \text{constant} \quad (18)$$

At this point, it is to observe that, in a WCVV model formulation, all cell species should be considered (individually or lumped), because all species net reaction rates contribute to the cell volume increase Eq. (6). In the present case, the rest of the cell content was mimicked by adopting large stationary concentrations for the lumped nutrients {NutP,NutG} and metabolites {MetP,MetG}. Species concentrations at homeostasis are given in (**Table 2**). Solving the models Eq. (15-16) is made by using the stiff integrator ("ODE15s") of Matlab package due to the very fast buffering reactions *r5-r6* in Eq. (14) compared to the rest of reactions (see the discrepancy in the WCVV model rate constants of **Table 3**). The number of moles *N*j (*t*) can be calculated at any time with the formula: *N*j(*t*)= *C*j(*t*) *V*(*t*) , by using the species concentrations *C*j(*t*) derived from solving the ODE model, and the volume obtained from Eq.(11), with using instant dilution "*Di*", or for a rough evaluation, with using the average cell dilution rate Dm = ln(2)/tc, with the cell cycle tc=100 min (**Table 2**).

The rate constants of the WCVV kinetic model Eq. (16) have been estimated from the steady-state (homeostatic) cell species concentrations of (Table 2). Such a QSS condition imposed to Eq.(10), and Eq.(16), requires setting the derivatives with respect to time to zero, that is: $\left(\frac{dC_j}{dt}\right)_s = h_{j,s}(C_s,k) = 0$, with the constraints discussed below Eq.(4) of section 3. In the present [G(P)1] model case, from Eq.(16) it results the nonlinear set Eq.(19):

$$\begin{cases} 0 = r_{1,s} - r_{3,s} - D_s\cdot C_{MetG,s} \\ 0 = r_{2,s} - r_{4,s} - D_s\cdot C_{MetP,s} \\ 0 = r_{4,s} - r_{5,s} + r_{6,s} - D_s\cdot C_{P,s} \\ 0 = r_{3,s} - r_{5,s} + r_{6,s} - D_s\cdot C_{G,s} \\ 0 = r_{5,s} - r_{6,s} - D_s\cdot C_{GP,s} \end{cases} \quad (19)$$

Table 2: The stationary homeostatic (QSS) species concentrations "*Cj,s*" considered for a GERM kinetic model of [G(P)1] type, in the example of section 6. Concentrations correspond to a GERM from *E. coli* cell K-12 strain [6,108,143]. Values are adapted from the example of Kurata and Sugimoto [114]. Notations are given in [footnote F]. The last column displays the GERM species recovering times of their homeostasis after a –10% impulse perturbation in the key-protein stationary [P]s produced at an arbitrary moment *t*=0. NG = negligible.

Species [footnote A]	Stationary conc. (nM)[footnote B]	Remarks	Recovery time (min.)
Lump $C_{NutP,s}$	3×10^8	Adopted	NG
Lump $C_{NutG,s}$	3×10^6	Adopted	NG
Lump $\sum_j MetG_{j,s}$	$\sim3\times10^6$	Approximated [footnote C]	NG
Lump $\sum_j MetP_{j,s}$	3×10^8	Adopted	NG
$C_{P,s}$ = [P]s	1000	Adopted	133
$C_{G,s}$ = [G]s	0.5	Adopted [footnote E]	93
$C_{GP,s}$ = [GP]s	0.5	Adopted [footnote E]	93
initial volume, $V_{cyt,o}$ (L)	$1.660434503\times10^{-15}$	[115]	
cell cycle, t_c	100 (min.)	Adopted [144]	
Ds = *Dm*	Ln(2)/ t_c	The average (apparent cell dilution rate); Eq.(6)	

Eigenvalues of the Jacobian of **[G(P)1]** kinetic model, Eq.(10,16) {∂ h(j)/ $\partial[\sum_j MetG_{j,s}, \sum_j MetP_{j,s}, C_{P,s},$ $C_{G,s}, C_{\emptyset,s}]\}$	Re(λ(1)) = -2.0005e+005 Re(λ(2)) = -1.7331e-002 Re(λ(3)) = -6.9315e-003 Re(λ(4)) = -6.9315e-003 Re(λ(5)) = -6.9315e-003	Ordered in the decreasing order [footnote D]

Footnotes:

[A] The cell volume average logarithmic growing rate (that is, the apparent cell dilution rate) is: $Ds = Dm = \mathrm{Ln}(2)/t_c$, according to Eq.(6,13). Lumps NutP and NutG denote substrates used in the synthesis of lumped metabolites MetP and MetG respectively, which, in turn, are further used for the P and G synthesis (see Fig. 7).

[B] Species copynumbers correspond to the *E. coli* cell K-12 strain [6,108,143]. Species concentrations in the cell are computed with the Eq.(2) relationship. For one gene [G]s = 1 nM in *E.coli* (see below Eq.(2)).

[C] The lump $\Sigma_j C_{MetG_{j,s}}$ results from the isotonicity constraint of **Table 1**, that is $RT/\pi\mathrm{cyt} = RT/\pi\mathrm{env}$, which translates in $\sum_{all\ j}^{cell} C_{j,s} = \sum_{all\ j}^{env} C_{j,s}$. That is:

$$C_{NutG,s} + C_{NutP,s} = \Sigma_j C_{MetG_{j,s}} + \Sigma_j C_{MetP_{j,s}} + \sum_{j \neq MetG, MetP}^{all\ cell} C_{j,s}.$$

Consequently, one obtains:

$$\Sigma_j C_{MetG_j} = C_{NutP,s} + C_{NutG,s} - \sum_{j \neq MetG_j}^{all\ cell} C_{j,s}.$$

[D] The $\mathrm{Max}(\mathrm{Re}(\lambda_j)) < 0$ evaluated at steady-state (**QSS**) (homeostasis) indicates a stable **QSS** cell homeostasis, where λ_j are the Jacobian eigenvalues of the **WCVV** kinetic model of the approached **GERM** model **[G(P)1]**.

[E] As proved by Maria [1,3,5-7], the stationary concentrations of the active ("G") and inactive ("GP") form of the catalyst ("G" here) should be equal at homeostasis. See the discussion below Eq.(4).

[F] Notations: G = a generic gene (DNA) from *E. coli* cell; P = the protein encoded by G; M = mRNA; GP = the inactive complex of G with the transcription factor (TF = P here); cyt = cytoplasma; "o"= initial; 's' index referring to the steady-state (**QSS**, homeostasis); nM = nano-molar; NG = negligible.

Table 3: The estimated rate constants of the GERM kinetic model of [G(P)1] type, as displayed in (Figure 7). In the case study of section 6, the model was written in the WCVV framework. The units of the 1-st order reactions are 1/min, while those of the 2-nd order are 1/(nM×min). Species stationary concentrations are those of Table 2. Adopted *k6* value much larger than *Ds* and the other rate constants, see the discussion of Maria [6], and those below Eq.(4).

[*] The adopted k6 rate constant satisfies the condition k6 >> *Ds* = *Dm*, according to Eq.(4). Here, Dm = the apparent average cell dilution rate Eq.(6), that is: Dm = Ln(2)/ *tc* For a *tc* = 100 min. (**Table 2**), it results *Dm* = 6.93 × 10^{-3} 1/min. This value was used in Eq.(19) (*Ds* = *Dm*) to estimate the WCVV model rate constants.

Reaction	Rate expression	Estimated rate constant (units in nM, and min.)
$NutG + P \xrightarrow{k_1} MetG + P$	$r_1 = k_1 C_{NutG} C_P$	6.929161 × 10^{-6}
$NutP + P \xrightarrow{k_2} MetP + P$	$r_2 = k_1 C_{NutP} C_P$	6.931494 × 10^{-6}
$MetG + P \xrightarrow{k_3} G + P$	$r_3 = k_3 C_{MetG} C_P$	2.311261 × 10^{-12}
$MetP + G \xrightarrow{k_4} P + G$	$r_4 = k_4 C_{MetP} C_G$	4.623291 × 10^{-8}
$P + G \xrightarrow{k_5} GP$	$r_5 = k_5 C_P C_G$	1.0000069
$GP \xrightarrow{k_6} G + P$	$r_6 = k_6 C_{GP}$	1 × 10^{+5} [*]

The analytical solution of Eq.(19), obtained by using the Maple or the symbolic Matlab softwares, is the followings:

$$k_1 = \frac{D_S\left(C_{MetG,s} + C_{G,s} + C_{GP,s}\right)}{C_{NutG,s}C_{P,s}}; k_2 = \frac{D_S\left(C_{MetP,s} + C_{P,s} + C_{GP,s}\right)}{C_{NutP,s}C_{P,s}}$$
$$k_3 = \frac{D_S\left(C_{G,s} + C_{GP,s}\right)}{C_{MetG,s}C_{P,s}}; k_4 = \frac{D_S\left(C_{P,s} + C_{GP,s}\right)}{C_{MetP,s}C_{G,s}};$$
$$k_5 = \frac{k_6 C_{G,s} + D_S C_{GP,s}}{C_{P,s}C_{G,s}}. \quad (20)$$

By replacing in Eq. (20) the stationary concentrations of Table 2 (taken for a GERM of *E. coli* cell K-12 strain, [6]), one obtains the WCVV model rate constants.

The numerical values of the rate constants, for the initial cell condition, and an average "*Ds* = *Dm*", are given in Table 3. For the rapid buffer reaction, ,

$$G + P \underset{k_6}{\overset{k_5}{\rightleftarrows}} GP$$

the reverse reaction rate constant "k6" (not estimable) was adopted at a value much larger than "Dm" (that is k6 >> Ds = Dm). To justify such a choice, see discussion below Eq.(4), and those of Maria [1,5,38]. For *tc* = 100 min. of Table 2, it results Dm = Ln(2)/*tc* = 6.93×10^{-3} 1/min (Eq.(6)). By adopting Ds = Dm in Eq.(20), the rate constants of Table 3 result. It is worth to mention that in the CVKM kinetic model classic formulation of Eq. (15), the rate constants cannot be estimated from the QSS form of the model, that is by setting the derivatives with respect to time to zero, that is: (d*Cj*/d*t*)s = *h*js(*Cs*,*k*) = 0. That is because of singularities of the resulting nonlinear algebraic set. Consequently, the same rate constants of the WCVV model (Table 3) were used in this section to make predictions with this classical ("default") CVKM model formulation.

Here it is to observe that in the Table 2, the value of the lump $\sum_j C_{MetG_{j,s}}$ results from the isotonicity constraint (ensuring the membrane integrity) $\pi_{cyt} = \pi_{env}$ from the hypotheses of WCVV models of Table 1. In other words, the sum of cell species concentrations must equal those of the environment (footnote [C] of Table 2). Otherwise, the osmosis will eventually lead to an equal osmotic pressure $\pi_{cyt} = \pi_{env}$. Even if, in a real cell, such 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 quasi-unchanged during the cell cycle.

To exemplify how the self-control of the protein synthesis works due to the rapid *buffer* reaction G + P ⇔ GP, one starts from the cell nominal stationary conditions (QSS) of Table 2, but with adopting [G]s =[GP]s = 1 nM, and one applies a dynamic perturbation ('impulse' like) at an arbitrary moment t=0 (Figure 10), by diminishing the stationary [P]s from 1000 nM to 900 nM (for a GERM of simplest [G(P)1] type). As a result, the regulatory buffer system leads to a quickly [G]s increase from 1 nM to 1.052 nM (due to less P in the buffer reaction, on the expense of GP which displays an accordingly decrease of [GP]s from 1 nM to 0.947 nM. In turn, such very quick changes will speed-up the P synthesis enough to recover the initial [P]s in ca. 127min. (with an acceptable tolerance of 1%). The same regulatory mechanism also applies to the [G]s, controlled by the same buffer reaction, the recovering time being of an acceptable 118 min. If one repeats the simulation, but with an initial [G]s =[GP]s = 0.5 nM then, as presented in the Table 2, the recovering time of [P]s is 133 min., and 93 min. for [G]s, respectively. It is to be noted that in "wild" *E. coli* cells [G]s is around 1 nM (see the proof in the footnote [B] of Table 2), but in cloned cells with plasmids, this level can be higher.

It is here worth noting that sometimes the predicted species recovering times after a perturbation could be longer than the cell cycle (100 min. here). In this case, we consider two alternatives to deal with this situation:

i) use another, more effective type of GERM, or

ii) if this has been experimentally demonstrated then, the recovering information on the GERM response to this dynamic perturbation is transmitted from cell generation to generation, as pointed-out by Elowitz and Leibler [57].

To complete this discussion about the advantages of using GERM models formulated under the WCVV approach, one simulates the dynamics of cell components present in large amounts (MetP, MetG), with using the same initial QSS of Table 2, but with adopting a stationary level of [G]s = [GP]s = 0.5 nM (to obtain a maximum regulatory efficiency; see [6,38]). By applying a dynamic perturbation to [P]s steady-state of a GERM of [G(PP)1] type from *E. coli*, (Figure 7), that diminishes the stationary [P]s from 1000 nM to 900 nM, simulation with the identified WCVV model, Eq. (16-18), and Table 3, is leading to the species trajectories displayed in Figure 11, proving again the good self-control of the P-synthesis produced by the lumped [G(P)1] model of GERM . The simulation results also underlines the positive effect of a large "cell ballast" (the sum of cell species concentrations [7]) on the GERM efficiency in "smoothing" the effect of a dynamic perturbation. Also, it is to remark the small effect of this small perturbation in [P]s on the homeostasis of species present in large amounts (i.e. MetP and MetG lumps).

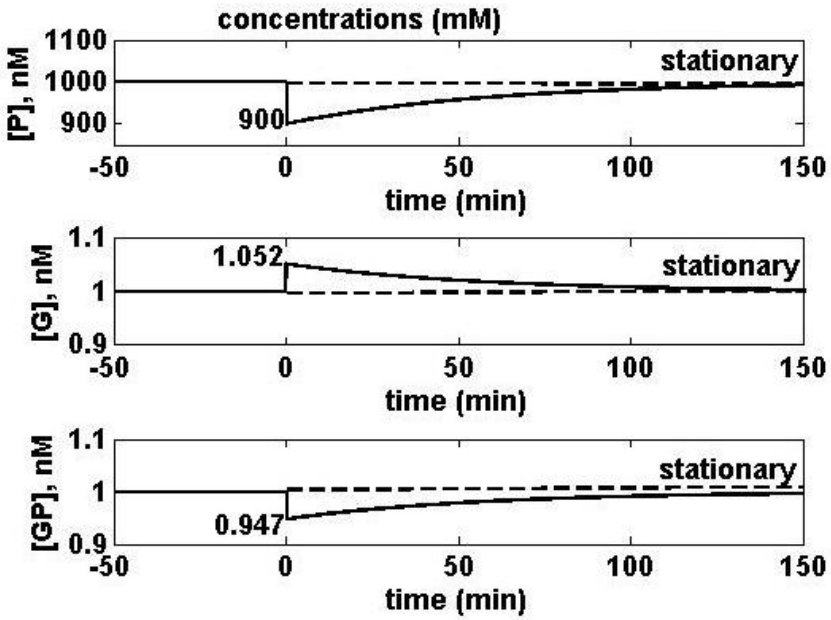

Figure 10: Exemplification of the self- and mutual- G/P pair catalysis and regulation after an impulse perturbation in the [P]s = 1000 nM leading to a -10% decline of the steady-state at an arbitrary time t=0 for a generic GERM of [G(P)1] type (Figure 7). Simulated results have been generated by using the WCVV kinetic model Eq.(16) with the rate constants of Table 3, for the cell nominal stationary conditions of Table 2, but with adopting $[G]_s$ =$[GP]_s$ = 1 nM.

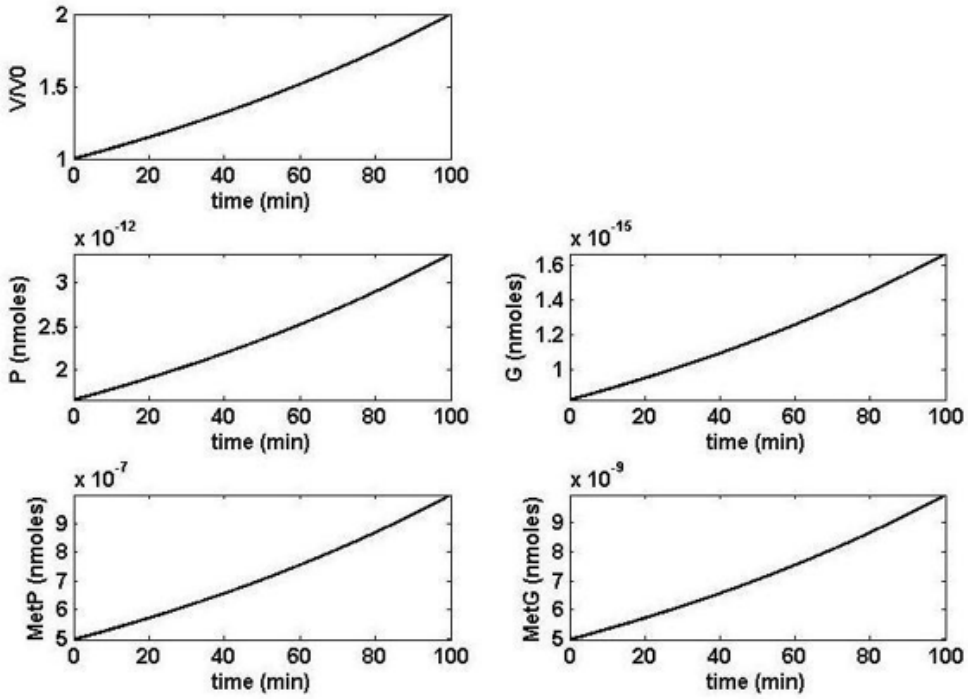

Figure 12: Dynamics of the cell volume and of species copy numbers Nj during the cell cycle predicted by the WCVV kinetic model for the G/P expression using a regulatory module of [G(P)1] type (Figure 7). The *E. coli* cell species homeostatic concentrations are those of Table 2.

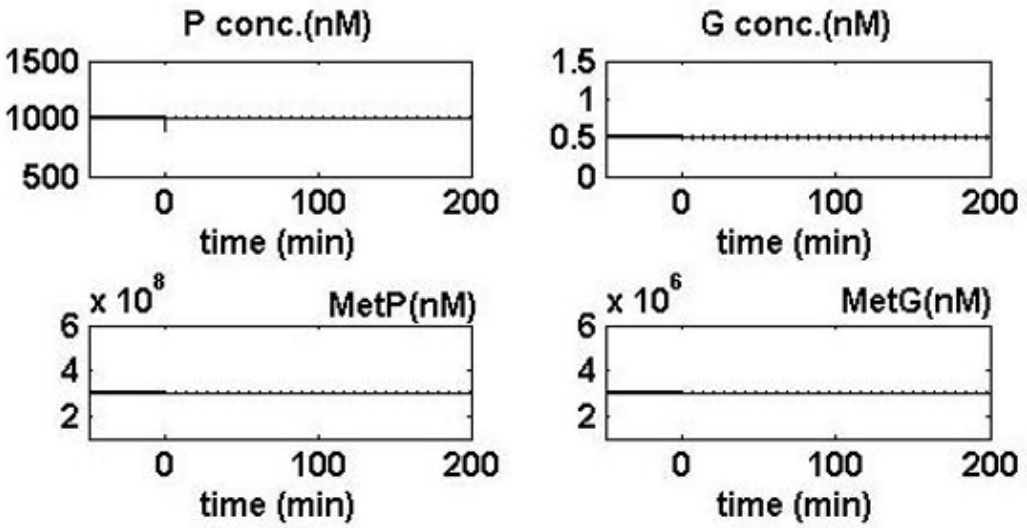

Figure 13: Simulation of the G/P expression under stationary (un-perturbed) conditions by using a regulatory module of [G(P)1] type (Figure 7) from the *E. coli* cell in terms of species concentrations The species (individual or lumped) homeostatic concentrations are those of Table 2. The dynamics of the key species concentrations (in nM) during the cell cycle is predicted by using two different approaches:

i) The classic ("default") constant volume CVKM model (------- , dash line with average "Dm" cell dilution Eq.(15)), compared to

ii) The variable volume WCVV novel modelling framework [_____ , continuous line with instant dilution "Di" Eq.(17) and isotonicity constraint included in the model, Eq.(16-18)]. Model predictions of the two models are practically overlaid, due to the small difference (in this case) between the instant (Di) and the average (Dm) cell dilution rates.

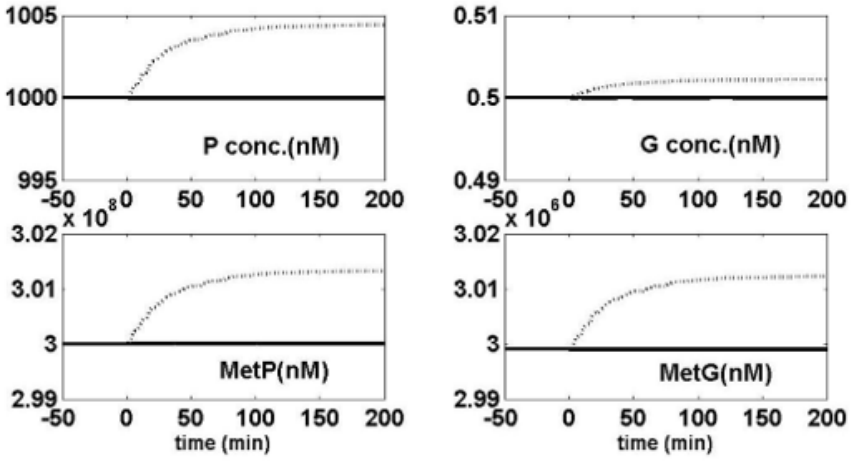

Figure 14: Species concentrations (in nM) dynamics during two cell cycles, under stationary (*un-perturbed*) conditions, predicted by i) the "classical" ("default") constant volume CVKM model (------- ,dash line with average "Dm" cell dilution Eq.(15)), compared to ii) the variable volume WCVV novel modelling framework [_____ , continuous line with instant dilution "Di" Eq.(17) and isotonicity constraint included in the model, Eq.(16-18)]. The generic G/P expression is simulated using a regulatory module of [G(P)1] type. The *E. coli* cell species homeostatic concentrations are those of Table 2.

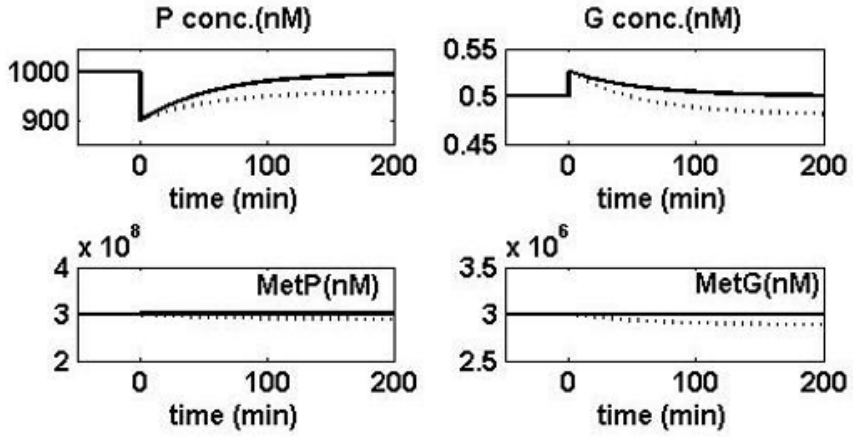

Figure 15: Simulation of the G/P expression after a dynamic ("impulse-like") perturbation in the [P]s = 1000 nM leading to a -10% decline of the steady-state at an arbitrary time t=0. The *E. coli* cell species homeostatic concentrations are those of Table 2. Simulations, in terms of species concentrations (nM), have been done by using a regulatory module of [G(P)1] type (Figure 7) from the *E. coli* cell. The species (individual or lumped) stationary (homeostatic) concentrations are those of Table 2. The dynamics of the key species concentrations during the cell cycle was predicted by using two different approaches: i) the classic ("default") constant volume CVKM model (------- ,dash line with average "Dm" cell dilution Eq.(15)), compared to ii) the variable volume WCVV novel modelling framework [_____ , continuous line with instant dilution "Di" Eq.(17) and isotonicity constraint included in the model, Eq.(16-18)]. Model predictions of the two models are very different, due to the use of the "instant" Di cell dilution rate, and of the isotonicity constraint by the WCVV novel modelling framework.

It is self-understood that, as the regulatory scheme (in Figure 6) is more effective, as the GERM regulatory efficiency is better. For an extensive discussion on regulatory properties of various GERM-s the reader is referred to the review of Maria [1-3,7]. This example demonstrates, in a simple way, how the classic ("default") CVKM modelling is unable/ineffective to reproduce even simple regulatory properties of GERM-s. In order to better illustrate the discrepancy in the predictive capabilities between the novel WCVV and the classic CVKM kinetic model formulations, and how deceptive can the predictions of the CVKM kinetic models be, one considers the same example of a generic GERM in the *E. coli* cell of [G(P)1] type illustrated in Figure 7. The reaction scheme, reaction rate expressions for the [G(P)1] gene expression module are those given in Eq.(16-18), and Table 3, for the WCVV formulation, and in Eq.(15) and Table 3, for the CVKM formulation.

The rate constants of the WCVV kinetic model were estimated based on Eq.(16,19), and with using the *E. coli* cell data with the initial homeostatic concentrations of Table 2. The same rate constants were used also for the CVKM model because for this model it is impossible to estimate the rate constants from the stationary conditions, due to the present singularities. Then, simulations of one cell cycle by using the WCVV model lead to obtaining the dynamics of species copy numbers (number of moles) plotted in Figure 12. It is to observe that, while the cell volume doubles, the species copy numbers double as well. However, there is a very small discrepancy in the predicted cell-volume dynamics because the average *Ds=Dm* Eq.(15) used by the CVKM model is slightly different from the instant *Di* Eq.(17) used by the WCVV model. However, this difference in the volume-vs.-time plot is small, and the both curves, predicted by the two models (WCCV and WCVV) are practically overlapping.

If the species dynamics is plotted in terms of concentrations (referred to the cell cytosol volume), the predicted trajectories for the *un-perturbed (stationary)* cell growth case, are given in Figure 13. The dynamics of the key species concentrations (in nM) during the cell cycle is predicted by using two different approaches: i) the classic ("default") constant volume CVKM model (------- , with average "Dm" cell dilution Eq.(15)), compared to ii) the variable volume WCVV novel modelling framework [—— , with instant dilution Di Eq.(17) and isotonicity constraint included in the model, Eq.(16-18)]. Model predictions of the two models are practically overlaid, the difference being negligible (in relative terms). Such a result can be explained by the fact that, over a large time-domain (cell cycle) the average cell dilution rate ("Dm", Eq.6) satisfactorily averages the "instant" dilution rate "Di" of Eq. (11,16-18). In Figure 13 both models correctly indicate how the key-species stationary values are preserved under stationary cell growth conditions, which corresponds to a stable system.

Apparently, the CVKM predictions of the "classic/default" kinetic model, with an average dilution "*Dm= ln(2)/tc*" in the model (Eq.(15)), are quite close compared to those predicted by the variable volume novel isotonic WCVV model (continuous line ——), employing the cell instant dilution "*Di*" Eq.(17) estimated from the isotonicity constraint included in the model, Eq.(16-18). By making a "zoom" in the Figure 13, it results Figure 14. Here, it clearly appears the differences in species predictions between the two models. The differences are small in relative terms, for species present in small amounts. By contrast, for the species present

in larger amount into the cell, such differences are important. As a conclusion, the predictions of the "classic/default" CVKM kinetic models, with employing an average dilution "*Dm= ln(2)/ tc*" are *inaccurate*, biased, and can not be used to perform precise evaluations of the GERM-s/GRC-s regulatory properties.

Discrepancies between the two kinetic model formulations increase dramatically, and huge differences in the predictions of the two models will appear *under perturbed* growing conditions. Thus, the species recovering trajectories and recovering times, after a dynamic perturbation, predicted by the CVKM model, and by the WCVV model after an impulse-like perturbation by diminishing the stationary [P]s from 1000 nM to 900 nM, are quite different as revealed by the plots of Figure 15. The differences in model predictions are as larger as the species present a lower level in the cell (that is for P, G, and GP species).

Consequently, while the WCVV kinetic models correctly reproduces the self-regulatory property of the cell system, its homeostasis being quickly recovered after a dynamic small perturbation, species concentrations being kept quasi-constant because both nominator and denominator of the fraction $Cj\ (t) = Nj\ (t) / Vcyt\ (t)$, [nM], are doubling at the same rate. By contrast, the classic ("default") CVKM model predictions *are inaccurate*, the predicted species concentration dynamics under perturbed conditions, eventually with using an average formula "Dm" Eq.(6,15) for the cell content dilution rate are *wrong*, and very far from the reality. So, the CVKM models can not be used in a satisfactory manner to simulate the regulatory properties of GERM–s or GRC–s [1-3,7]. One can conclude that the novel WCVV modelling framework introduced by Maria [1-7] better reflects the GERM–s regulatory properties after dynamic (impulse-like) or stationary (step-like, not discussed here) internal or external perturbations.

7 Conclusions

The simple case study approached in this paper to exemplify the application advantage of the novel WCVV modular kinetic modelling framework to simulate GERM-s dynamics, and regulatory properties/efficiency, proved that the chemical and biochemical engineering principles and numerical algorithms, together with those of the nonlinear systems control theory are fully applicable to modelling complex metabolic cell processes, including sophisticated GRC-s controlling the cell enzymes syntheses and metabolic fluxes. The deterministic ODE kinetic models with continuous variables are fully feasible alternatives to well describe the cell response to stationary or dynamic continuous perturbations from the environment [1-3,7,65].

In fact, such WCVV cell process models 'translate' from the 'language' of molecular biology to that of mechanistic chemistry and mathematics/computing languages, trying to preserve the structural, functional, and timing hierarchy of the cell components and functions (Figure 2). To avoid extended ODE cell kinetic models, difficult to identify, and to be used, lumped deterministic model structures have been proved to efficiently represent the metabolic cell processes. The lumping degree should be chosen according to the available kinetic data, and utilisation scope, to ensure a satisfactory trade-off between model simplicity and its predictive quality [1-3,6,7].

As another observation, by contrast to the novel WCVV modelling framework, one fundamental deficiency of the classical "default" (constant-volume like) CVKM kinetic model formulation is the lack of the intrinsic stability of the cell system kinetic model {that is "Max(Re(li)) = -Di", where "li" are the eigenvalues of the kinetic model Jacobian matrix; see the proof of Maria [1-3]}, because these models do not include neither the Pfeffers' isotonicity constraint**,** nor an equivalent constraint. Consequently, as proved here, the GERM model formulated in a classic CVKM framework is wrong, being not able to simulate how the system recovers its homeostasis after a dynamic perturbation**.** Besides, predictions of CVKM models are distorted and inaccurate even when using an average cell dilution. Unfortunately, the classical CVKM continue to largely be used in the dedicated literature with wrong / inaccurate predictions.

By not explicitly considering the relationships between the cell volume exponential increase during the cell growth, the osmotic pressure, and species reaction rates, the classic CVKM kinetic models lead to biased and distorted conclusions on GERM–s regulatory performances (i.e. the response to perturbations), thus making difficult the modular construction of GRCs by linking effective GERMs. Maria [7] reviewed a large number of advantages offered by the use of the novel WCVV modelling framework. Among these, the following are worth mentioning

i) The *in-silico* study of the protein-synthesis regulation in a cell, by using numerical simulations via a continuous ODE kinetic model and a WCVV novel modelling approach, accounting for continuous variables and cell-volume growth, has been proved to be a worthy instrument for its quick evaluation under continuous/random internal-/external variable conditions. These conditions can include stationary ('step-like') perturbations in species concentrations, continuous cell-content dilution due to the cell-volume growth, but also 'impulse-like' dynamic perturbations of the cell system components. Based on the known stationary (homeostatic) concentrations of the key species considered in the QSS mass-law equations of the WCVV model, and on optimal regulatory performance quantitative indices (P.I.-s), the kinetic model rate constants can be estimated. Even if incomplete experimental information is available, the same rule has been proven to be effective in ranking and discriminating among various GERM-s / GRC-s regulatory schemes.

ii) The stationary regulatory capability of GERM models constructed under the novel WCVV modelling framework seems to be related to the *regulatory effectors* (i.e. very rapid reversible "buffering" reactions) which continuously adjust the 'catalyst' activity in the G/P syntheses, thus keeping the QSS (homeostatic) concentration-levels against the continuous cell-volume dilution effect and 'step-like' (stationary perturbations), or 'impulse-like' (dynamic perturbations) variations in the environment. At the same time, the presence of as large as possible number of *regulatory effectors* can quickly adjust the protein/gene ratio (during their expression) to cope with the quick recovering after an 'impulse' external/ internal perturbation in species concentration levels. When designing a cell-regulatory model, both these aspects have to

be considered, and a trade off between model effectiveness and model complexity has to be realized. It is also to observe that, during the protein (P) recovering time-interval, the inertial effect created by the large copy numbers into the cell is keeping the cell-volume growing with a near the same average rate irrespectively to small perturbations, thus fulfilling the goal to become of critical size after a certain time to trigger the cell division process [36]. In a variable-volume WCVV regulatory model such an effect is accurately described, while in a classic ("default") constant-volume CVKM kinetic model, the same effect is 'artificially' averaged by means of fictive reactions and hypotheses [8], leading to wrong predictions.

iii) Among the checked GERM-s kinetic models of (Figure 6), the best-found regulatory model [G(PP)n] includes an effective regulatory consecutive "buffering" schema for the P key-protein synthesis, thus better compensating the cell dilution effect, by adjusting the P/G/GPP/GPPPP ratios (for n = 4 case). The synthesis of P via a cascade control, involving an intermediate species M with an increased number of concurrent reactions (that is the model [G(P)n; M(P)n1]) was proved to be more sensitive to inner/external conditions, and then less effective for stationary regulation. By contrast, the model [G(PP)n] was proved to be superior in dynamic perturbed conditions, due to the presence of dimeric transcription factors (PP) of adjustable small concentration, involved in fast reversible P-synthesis "buffering" reactions.

iv) By including in the GERM regulatory WCVV models separate 'raw-materials' species {NutP, MetP} for the synthesis of protein(s) (Figure 7), and {NutG, MetG} for the synthesis of the corresponding gene(s), an increase in the model prediction flexibility is thus realized. This allows simulating the cell-behaviour under various external conditions (that is for various {NutP, NutG} levels), but also various internal 'raw-materials' conditions (that is for various {MetP, MetG} levels). When only a proteinic regulatory module is simulated (expression of a G/P pair), it will be difficult to set the optimal inner 'raw-material' levels to obtain an optimal regulatory index, due to the required system isotonicity constraint, and due to the "*volume-inertial constraint*" for preventing a too high sensitivity to perturbations [38].

v) As proven in this work, and by Maria [7,51,53], even if more complex and involving a supplementary identification effort, the advantages of using variable-volume WCVV kinetic models to simulate the GERM-s modules, instead of using the classic ("default") constant-volume CVKM models for the cell-system simulation are multiple:

a. allows representations of the cell-volume growth under various internal/external-cell perturbations; allows eliminating 'fictive' reactions accounting the cell-volume growth effect on species [8]

b. includes all the system species into the overall mass balance (individual or lumped), with all the complexity of direct/indirect interactions (individual or lumped)

c. allows to include the volume inertial effect due to the presence of a large number of molecules into the cell, thus smoothing the small environmental perturbations

d. allows deriving of more realistic stationary/dynamic quantitative regulatory indices (P.I.-s) for a studied GERM kinetic schema

e. allows a more correct ranking of regulatory model alternatives (see their *library* in Figure 6)

f. allows a more detailed and accurate linkage of the regulatory modules to construct GRC-s, in order to simulate several protein regulatory processes in a more flexible way and including the direct but also the indirect interacting effects among modules via the common cell-volume increase to which all species participate

g. Besides, as proved by Maria [7] the GERM-s LIBRARY of Maria [1-3,6] (Figure 6) allows the *in-silico* comparative study of the regulatory properties of all GERM-s

The extensive discussion of the GERM-s regulatory properties by Maria [38] proved that it is only WCVV modelling framework which makes possible the quantitative evaluation of the performance indices (P.I.-s) of various GERM–s types. Such an evaluation would be impossible by using the classical ("default") constant-volume CVKM modelling framework (as they are for example the cell "inertial effect", or the cell "ballast effect", or the "second perturbation", in a perturbed environment).

vi) Starting from an effective proteinic regulatory schema GERM (from the library of Figure 6), effective modular GRC–s can be constructed by linking proteinic GERM-s in a regulatory chain, by employing a "building blocks" strategy and certain linking rules derived by Maria [1,7]. The obtained GRC can display several functions, such as: an operon expression [65], or a genetic switch [3,39,41,42]. The GERM modules can be linked in several ways (see examples of Maria [7]), that is

a) direct connections through common species(s) and/or common reaction(s)

b) indirect connections by means of the induced cell-volume growth, to which all species contribute

c) indirect connections by means of the cell volume, which is influenced by the species levels perturbation which, in turn will influence the all species from the cell through the common cell volume (the so-called "second perturbations")

Acknowledgment

None.

Conflict of Interest

No conflict of Interest.

Notations

Cj	Species "j" concentration, of vector ***C***
Di	cell content instant dilution rate = cell volume logarithmic instant growing rate, Eq.(10,11)
Dm	The average logarithmic growing rate of the cell volume, Eq.(6)
h(j)	Kinetic model j-th function, corresponding the the "j" species mass balance, Eq.(5a-c)
K, k	Rate constants, of vector ***k***
Nj	Species "j" number of moles, of vector ***N***
N_A	The Avogadro number.
$nr = n_r$	Number of reactions considered in the kinetic model
$ns = n_s$	Number of cell species considered in the kinetic model
r_i	*i*-th reaction rate (individual or lumped)
R	Universal gas constant;
$s(i,j) = \nu_{ij}$	The stoichiometric coefficients of species "j" in the reaction "i"
T	Absolute temperature
t	Time
tc	Cell cycle
V	Cell (cytosol) volume
π	Osmotic pressure of the cell content
Index	
0,o	Initial
cyt	Cytosol
env	Environment
s	At QSS, that is at cell homeostasis (balanced growth)
Abbreviations	
CCM	Central carbon metabolism
ChBPE	Chemical and biochemical process engineering
CVKM	"Whole-cell quasi-constant volume" kinetic models
GERM	Gene expression regulatory module
GMO	genetically modified micro-organisms
GRC	Genetic regulatory circuits
GS	Genetic switches
MetG	Lumped metabolites used for the synthesis of G genes (DNA)
MetP	Lumped metabolites used for the synthesis of P proteins
MINLP	Mixed integers NLP optimization problems
NLP	Nonlinear programming optimization problems (with continuous variables)
NutG	Lumped nutrients used for the synthesis of lumped MetG metabolites
NutP	Lumped nutrients used for the synthesis of lumped MetP metabolites
ODE	Ordinary differential set of equations
P.I.-s	Quantitative regulatory performance indices of GERM-s,
QSS	Quasi-steady-state
TF	Transcriptional factors
WCVV	The novel "mechanistic silicon whole cell" kinetic modelling framework of variable volume for isotonic cells, able to maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts

References

1. Maria, G (2017A) A review of some novel concepts applied to modular modelling of genetic regulatory circuits. Irvine (CA, USA): Juniper publ.
2. Maria, G (2017B) Deterministic modelling approach of metabolic processes in living cells - a still powerful tool for representing the metabolic process dynamics. Irvine (CA, USA): Juniper publ.
3. Maria, G. (2018). In-silico design of Genetic Modified Micro-organisms (GMO) of industrial use, by using Systems Biology and (Bio)Chemical Engineering tools. Irvine (CA, USA): Juniper publ.
4. Maria, G (2002) Seminar "Simulation of hypothetical 'Mechanical Cells' that maintain intracellular homeostasis while growing autocatalytically on environmental nutrients present in variable amounts". Personal communication, Texas A&M Univ., Dept. of chemistry (prof. P. Lindahl group), Aug. 23.
5. Maria G.C, Morgan J.J, Paul A. Lindahl P.A (2002) Kinetic simulation of hypothetical "Mechanical Cells" that maintain intracellular homeostasis while growing autocatalytically on environmental nutrients present in variable amounts. Jl. Theor. Biology, submitted Aug. 6, 2002, rejected Sept. 2002.
6. Maria, G (2005) Modular-Based Modelling of Protein Synthesis Regulation, Chemical and Biochemical Engineering Quarterly, 19: 213-233.
7. Maria, G (2026) The mechanistic silicon whole cell - A novel kinetic modelling framework of the cell metabolic processes, able to maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. Los Angeles (USA): Peertechz publ.
8. Heinrich R, Schuster S (1996) The regulation of cellular systems. New York: Chapman & Hall.
9. Maria, G (2024a) Application of (bio)chemical engineering concepts and tools to model GRCs, and some essential CCM pathways in living cells. Part 1. Generalities. Curr Trends in Biomedical Eng & Biosci., 22: 556080-556104.
10. Maria, G (2024b) Application of (bio)chemical engineering concepts and tools to model GRCs, and some essential CCM pathways in living cells. Part 2. Mathematical modelling framework, Annals of Reviews & Research. 10: 555790-555835.
11. Maria, G (2024c) Application of (bio)chemical engineering concepts and tools to model GRCs, and some essential CCM pathways in living cells. Part 3. Applications in the bioengineering area, Archives in Biomedical Engineering & Biotechnology. 7: ABEB MS.ID.000672.
12. Maria, G (2024d) Application of (bio)chemical engineering concepts and tools to model GRCs, and some essential CCM pathways in living cells. Part 4. Applications in design some GMO-s. Annals of Systems Biology. 7: ID: ASB-7-121.
13. Banga, J (2008) Optimization in computational systems biology, Proc. 6-th Simon Stevin lecture on optimization in engineering, Leuven-Heverlee (NL). In: BMC Systems Biology. 2: 47.
14. Leroy H (2017) Statement given by the Centre of Mathematics Applied to the Life Sciences, University of Glasgow (UK).
15. Zak D.E, Aderem A (2009) Systems biology of innate immunity. Immunol Rev 227(1): 264-282.
16. Aebersold R (2026) Inst. Systems Biology, Seattle.
17. Kitano H (2002b) Systems biology: a brief overview. Science 295(5560): 1662-1664.
18. Kitano H (2002) Computational Systems Biology. Nature 420(6912): 206-210.
19. Benner S.A, Sismour A.M (2005) Synthetic biology. Nature Reviews Genetics 6(7): 533-543.
20. Hempel D.C (2006) Development of biotechnological processes by integrating genetic and engineering methods. Eng. Life Sci 6(5): 443-447.
21. Maria G (2018f) From residual biomass and inferior quality coal to the synthesis of methanol and then to hydrocarbons and gasoline – a Romanian project of high success. Irvine (CA, USA): Juniper publ.
22. Maria G, Gijiu C.L, Dinculescu D, Titica M, Juncu G (2020) A review of unconventional technologies for capitalization of cheap natural resources (natural gas, lower coal), greenhouse gases (CO2) and renewable biomass for the production via methanol of a large number of high value-added chemicals and fuel by using technologies based on modern tools and concepts of chemical and biochemical engineering. Bucharest: Printech Publ (in Romanian).
23. Yang S.T (2007) Bioprocessing for value-added products from renewable resources. Amsterdam: Elsevier.
24. Maria G (2023) Hybrid modular kinetic models linking cell-scale structured CCM reaction pathways to bioreactor macro-scale state variables. Applications for solving bioengineering problems. Irvine (CA, USA): Juniper publ.
25. Westerhoff H.V (2001) The Silicon cell, not dead but live! Metabolic Engineering 3(3): 207-210.
26. Westerhoff H.V, Palsson B.O (2004) The evolution of molecular biology into systems biology. Nature Biotechnology 22(10): 1249-1252.
27. Zadran S, Levine R.D (2013) Perspectives in Metabolic Engineering: Understanding cellular regulation towards the control of metabolic routes. Appl Biochem Biotechnol. 169: 55-65.
28. Palsson B.O (2000) The challenges of in silico biology. Nat Biotechnol 18(11):1147-1150.
29. Tomita M (2001) Whole-cell simulation: a grand challenge of the 21-st century. Trends in Biotechnology 19(6): 205-210.
30. Mori H (2004) From the sequence to cell modeling: comprehensive functional genomics in E. coli. J Biochem Mol Biol 37(1): 83-92.
31. You L.C (2004) Toward computational systems biology. Cell Biochem Biophys 40(2): 167-184.
32. Ideker T, Galitski T, Hood L (2001) A new approach to decoding life: Systems Biology. Annu Rev Genomics Hum Genet 2: 343-372.
33. Jansen R.C (2003) Studying complex biological systems using multifactorial perturbation. Nat Rev Genet 4(2): 145-151.
34. Lindon J.C, Holmes E, Nicholson J.K (2004) Metabolomics and its role in drug development and disease diagnosis. Exp Rev Mol Diagn 4(2): 189-199.
35. Rao B.M, Lauffenburger D.A, Wittrup K.D (2005) Integrating cell level kinetic modeling into the design of engineered protein therapeutics. Nat Biotech 23(2): 191-194.
36. Surovstev I.V, Morgan J.J, Lindahl P.A (2007) Whole-cell modeling framework in which biochemical dynamics impact aspects of cellular geometry. Journal of Theoretical Biology 244(1): 154-166.
37. Maria G (2003) Evaluation of protein regulatory kinetics schemes in perturbed cell growth environments by using sensitivity methods. Chemical and Biochemical Engineering Quarterly 17: 99-117.
38. Maria G (2006) Application of lumping analysis in modelling the living systems -A trade-off between simplicity and model quality. Chemical and Biochemical Engineering Quarterly 20: 353-373.
39. Maria G (2007) Modelling bistable genetic regulatory circuits under variable volume framework. Chemical and Biochemical Engineering Quarterly 21: 417-434.
40. Maria G (2008) Reduced modular representations applied to simulate some genetic regulatory circuits. Revista de Chimie 59(3): 318-324.

41. Maria G (2009) Building-up lumped models for a bistable genetic regulatory circuit under whole-cell modelling framework, Asia-Pacific Journal of Chemical Engineering 4: 916-928.

42. Maria G (2014b) Extended repression mechanisms in modelling bistable genetic switches of adjustable characteristics within a variable cell volume modelling framework, Chemical & Biochemical Engineering Quarterly 28: 35-51.

43. Maria G, Maria C, Tociu C (2017) Comments on two novel review ebooks in the area of deterministic modelling of metabolic processes and of genetic regulatory circuits in living cells. UPB Bull. Sci., series B 79: 3-19.

44. Maria G (2014a) Insilico derivation of a reduced kinetic model for stationary or oscillating glycolysis in Escherichia coli bacterium, Chemical & Biochemical Engineering Quarterly 28: 509-529.

45. Maria G (2021) A CCM-based modular and hybrid kinetic model to simulate the tryptophan synthesis in a fed-batch bioreactor using modified E. coli cells. Computers & Chemical Engineering 153: 107450-107466.

46. Maria G, Renea L (2021) Tryptophan production maximization in a fed-batch bioreactor with modified E. coli cells, by optimizing its operating policy based on an extended structured cell kinetic model. Bioengineering-Basel 8(12): 210-247.

47. Maria G, Xu Z, Sun J (2011) Multi-objective MINLP optimization used to identify theoretical gene knockout strategies for E. coli cell, Chemical & Biochemical Engineering Quarterly 25(4): 403-424.

48. Maria G, Gijiu C.L, Maria C, Tociu C (2018a) Interference of the oscillating glycolysis with the oscillating tryptophan synthesis in the E. coli cells. Computers & Chemical Engineering 108: 395-407.

49. Maria G, Mihalachi M, Gijiu C.L (2018b) Model-based identification of some conditions leading to glycolytic oscillations in E. coli cells. Chemical and Biochemical Engineering Quarterly 32(4): 523-533.

50. Maria G, Mihalachi M, Gijiu C.L (2018c) In silico optimization of a bioreactor with an E. coli culture for tryptophan production by using a structured model coupling the oscillating glycolysis and tryptophan synthesis. Chemical Eng. Res. and Design. 135: 207-221.

51. Maria G, Gijiu C.L, Maria C, Tociu C, Mihalachi M (2018d) Importance of considering the isotonic system hypothesis when modelling the self-control of gene expression regulatory modules in living cells, Current Trends in Biomedical Engineering & Biosciences, 12: CTBEB. MS.ID.555833.

52. Maria G, Mihalachi M, Gijiu C.L (2018e) Chemical engineering tools applied to simulate some conditions producing glycolytic oscillations in E. coli cells. U.P.B. Sci. Bull. Series B – Chemie 80: 27-38.

53. Maria G, Maria C, Tociu C (2018g) A comparison between two approaches used for deterministic modelling of metabolic processes and of genetic regulatory circuits in living cells. U.P.B. Sci Bull series B – Chemie 80(1): 127-144.

54. Maria G (2020C) In-silico determination of some conditions leading to glycolytic oscillations and their interference with some other processes in E. coli cells. Frontiers in Chemistry 8: 526679-526693.

55. Atkinson M.R, Savageau M.A, Myers J.T, Ninfa A.J (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. Cell 113(5): 597-607.

56. Guantes R, Poyatos J.F (2006) Dynamical principles of two-component genetic oscillators. PLoS Computational Biology 2(3): e30.

57. Elowitz M.B, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403(6767): 335-338.

58. Gonze D (2010) Coupling oscillations and switches in genetic networks. BioSystems 99(1): 60-69.

59. Bier M, Teusink B, Kholodenko B.N, Westerhoff H.V (1996) Control analysis of glycolytic oscillations, Biophysical Chemistry 62(3): 15-24.

60. Silva A.S, Yunes J.A (2006) Conservation of glycolytic oscillations in Saccharomyces cerevisiae and human pancreatic beta-cells: a study of metabolic robustness. Genetics and Molecular Research 5(3): 525-535.

61. Heinzle E, Dunn I.J, Furukawa K, Tanner R.D (1982) Modelling of sustained oscillations observed in continuous culture of Saccharomyces Cerevisiae, In: Aarne, H. (Ed.), Proc. Modelling & Control of Biotechnical Process IFAC Conference, Helsinki (Finland), Aug. 17-19, 1982.

62. Tyson J.J (2002) Biochemical Oscillations. In: Fall, C.P., Marland, E.S., Wagner, J.M., Tyson, J.J. (Eds.), Computational Cell Biology. Berlin: Springer verlag, chap. 9.

63. Franck U.F (1980) Feedback kinetics in physicochemical oscillators. Ber Bunsenges Phys Chem 84(4): 334-341.

64. Maria G (2010) A dynamic model to simulate the genetic regulatory circuit controlling the mercury ion uptake by E. coli cells. Revista de Chimie 61(2): 172-186.

65. Maria G, Luta I (2013) Structured cell simulator coupled with a fluidized bed bioreactor model to predict the adaptive mercury uptake by E. coli cells. Computers & Chemical Engineering. 58: 98-115.

66. Maria G, Scoban A.G (2017) Setting some milestones when modelling gene expression regulatory circuits under variable-volume whole-cell modelling framework. 1. Generalities. Revista de Chimie 68(12): 3027-3037.

67. Maria G, Scoban A.G. (2018) Setting some milestones when modelling gene expression regulatory circuits under variable-volume whole-cell modelling framework. 2. Revista de Chimie 69(1): 259-266.

68. Sauro H.M, Kholodenko B.N (2004) Quantitative analysis of signalling networks. Prog Biophys Mol Biol 86(1): 5-43.

69. Heinemann M, Panke S (2006) Synthetic Biology - putting engineering into biology. Bioinformatics 22(22): 2790-2799.

70. Qian Y, McBride C, Del Vecchio D (2017) Programming cells to work for us. Annual Review of Control, Robotics, and Autonomous Systems 1: 411-440.

71. Kobayashi H, Kaern M, Araki M, Chung K, Gardner T.S (2004) Programmable cells: Interfacing natural and engineered gene networks. Proc. Natl. Acad. Sci. USA 101(22): 8414-8419.

72. Yang Q, Lindahl P, Morgan J (2003) Dynamic responses of protein homeostatic regulatory mechanisms to perturbations from steady state. J theor Biol 222(4): 407-423.

73. Sewell C, Morgan J, Lindahl P (2002) Analysis of protein regulatory mechanisms in perturbed environments at steady state. J theor Biol 215(2): 151-167.

74. Bower J.M, Bolouri H (2001) Computational Modeling of Genetic and Biochemical Networks, Cambridge: MIT Press.

75. Westerhoff H.V (2006) Engineering life processes live: the Silicon cell, Oral Lecture ESCAPE 16 conference, Garmisch (Germany).

76. Maria G (2017) Application of (bio) chemical engineering principles and lumping analysis in modelling the living systems. Current Trends in Biomedical Engineering & Biosciences 1(4): CTBEB.MS.ID.555566.

77. Forster J, Famili I, Fu P, Palsson B.O, Nielsen J (2003) Genome-Scale reconstruction of the Saccharomyces cerevisiae metabolic network. Genome Res 13(2): 244-253.

78. Edwards J.S, Palsson B.O (2000) The Escherichia coli MG1655 in silico metabolic genotype - Its definition, characteristics, and capabilities. Proc Natl Acad Sci U S A 97(10): 5528-5533.

79. Edwards J.S, Ibarra R.U, Palsson B.O (2001) In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nature Biotechnology 19(2): 125-130.

80. Heinemann M, Kummel A, Ruinatscha R, Panke S (2005) In Silico Genome-Scale Reconstruction and Validation of the Staphylococcus

aureus Metabolic Network. Biotechnology and Bioengineering 92(7): 850-864.

81. Rodriguez-Prados J.C, de Atauri P, Maury J, Ortega F, Portais J.C (2009) In Silico strategy to rationally engineer metabolite production: A case study for threonine in Escherichia coli. Biotechnol Bioeng 103(3): 609-620.

82. Rocha I, Maia P, Evangelista P, Vilaca P, Soares S, (2010) OptFlux: an open-source software platform for in silico metabolic engineering. BMC Syst Biol 4: 45.

83. Le Phillip P, Bahl A, Ungar L.H (2004) Using prior knowledge to improve genetic network reconstruction from microarray data. Silico Biol 4(3): 335-353.

84. Nemenman I, Escola G.S, Hlavacek W.S, Unkefer P.J, Unkefer C.J (2007) Reconstruction of metabolic networks from high-throughput metabolite profiling data - In Silico analysis of red blood cell metabolism. Ann NY Acad Sci 1115: 102-115.

85. Camacho D, Vera Licona P, Mendes P, Laubenbacher R (2007) Comparison of reverse-engineering methods using an in-silico network. Ann NY Acad Sci 1115: 73-89.

86. Kiparissides A, Koutinas M, Kontoravdi C, Mantalaris A, Pistikopoulos E.N (2011) 'Closing the loop' in biological systems modeling—From the in silico to the in vitro. Automatica 47(6): 1147-1155.

87. Thilakavathi M, Basak T, Panda T (2007) Modeling of enzyme production kinetics. Applied microbiology and biotechnology 73(5): 991-1007.

88. Covert M.W, Schilling C.H, Famili I, Edwards J.S, Goryanin I.I (2001) Metabolic modeling of microbial strains in silico. Trends Biochem Sci 26(3): 179-186.

89. Bailey J (2001) Complex biology with no parameters. Nat Biotechnol 19(6): 503-504.

90. Bailey J.E (1991) Towards a science of metabolic engineering. Science 252(5013): 1668-1675.

91. Yokobayashi Y, Collins C.H, Leadbetter J.R, Weiss R, Arnold F.H (2003) Evolutionary design of genetic circuits and cell-cell communications. Advances in Complex Systems 6(1): 1-9.

92. Sidoli FR, Mantalaris A, Asprey SP (2004) Modelling of mammalian cells and cell culture processes. Cytotechnology 44(2): 27-46.

93. Xie L, Wang D.I.C (1994) Stoichiometric analysis of animal cell growth and its application in medium design. Biotech Bioeng 43(11): 1164-1174.

94. deZengotita VM, Miller WM, Aunins JG, Zhou W (2000) Phosphate feeding improves high-cell-concentration NS0 myeloma culture performance for monoclonal antibody production. Biotechnol Bioeng 69(5): 566-576.

95. de Tremblay M, Perrier M, Chavarie C, et al. (1993) Fed-batch culture of hybridoma cells: comparison of optimal control approach and closed loop strategies. Bioprocess Engineering 9: 13-21.

96. Dhir S, Morrow Jr. K.J, Rhinehart R.R, Wiesner T (2000) Dynamic optimization of hybridoma growth in a fed-batch bioreactor. Biotechnology and Bioengineering 67(2): 197-205.

97. San KY, Stephanopoulos G (1989) Optimization of fed-batch penicillin fermentation: a case of singular optimal control with state constraints. Biotechnol Bioeng 34(1): 72-78.

98. Frahm B, Lane P, Atzert H, Munack A, Hoffmann M (2002) Adaptive, Model-Based Control by the Open-Loop-Feedback-Optimal (OLFO) Controller for the Effective Fed-Batch Cultivation of Hybridoma Cells. Biotechnology progress 18(5): 1095-1103.

99. Pörtner R, Schäfer T (1996) Modelling hybridoma cell growth and metabolism — a comparison of selected models and data. Journal of Biotechnology 49(3): 119-135.

100. Kholodenko B.N, Kiyatkin A, Bruggeman F.J, Sontag E, Westerhoff H.V (2002) Untangling the wires: A strategy to trace functional interactions in signalling and gene networks. Proc Natl Acad Sci USA 99(20): 12841-12846.

101. Savageau M.A (2002) Alternatives designs for a genetic switch: Analysis of switching times using the piecewise power-law representation. Math. Biosciences 180: 237-253.

102. Hlavacek W.S, Savageau M.A (1997) Completely uncoupled and perfectly coupled gene expression in repressible systems. J. Mol. Biol 266(3): 538-558.

103. Wall M.E, Hlavacek W.S, Savageau M.A (2003) Design principles for regulator gene expression in a repressible gene circuit. J Mol Biol 332(4): 861-876.

104. Salvador A, Savageau M.A (2003) Quantitative evolutionary design of glucose 6-phosphate dehydrogenase expression in human erythrocytes. Proc Natl Acad Sci USA 100(24): 14463-14468.

105. Sorribas A, Savageau M.A (1989) A comparison of variant theories of intact biochemical systems. II. Flux-oriented and metabolic control theories. Math Biosci 94(2): 195-238.

106. Gabaldon T, Huynen M.A (2004) Prediction of protein function and pathways in the genome era. Cell Mol Life Sci 61(8): 930-944.

107. Sotiropoulos V, Kaznessis Y.N (2007) Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes. BMC Syst Biol 1-7.

108. Tian T, Burrage K (2006) Stochastic models for regulatory networks of the genetic toggle switch. Proc Natl Acad Sci USA 103(22): 8372-8377.

109. Maria G (2004) A review of algorithms and trends in kinetic model identification for chemical and biochemical systems. Chemical and Biochemical Engineering Quarterly 18(3): 195-222.

110. Maria G (2019) Numerical methods to reduce the kinetic models of (bio)chemical processes. Bucharest: Printech Publ. (in Romanian).

111. Lodish H, Berk A, Matsudaira P, Kaiser C.A, Krieger M (2000) Molecular cell biology. New York: Freeman & Co.

112. Savageau M.A, Voit E.O (1987) Recasting nonlinear differential equations as S-systems: A canonical nonlinear form. Math. Biosciences 87(1): 83-115.

113. Voit E.O (2005) Smooth bistable S-systems. Syst Biol 152(4): 207-213.

114. Kurata H, Sugimoto Y (2017) Improved kinetic model of Escherichia coli central carbon metabolism in batch and continuous cultures. Journal of Bioscience and Bioengineering 17: S1389-1723.

115. Bar-Joseph Z, Gitter A, Simon I (2012) Studying and modelling dynamic biological processes using time-series gene expression data. Nature Reviews Genetics 13(8): 552-564.

116. Hargrove J.L, Schmidt F.H (1989) The role of mRNA and protein stability in gene expression. FASEB J 3(12): 2360-2370.

117. Froment G.F, Bischoff K.B (1990) Chemical reactor analysis and design, New York: Wiley.

118. Dutta R (2008) Fundamentals of biochemical engineering. Berlin: Springer verlag.

119. Moser A (1988) Bioprocess technology: kinetics and reactors. New York: Springer-Verlag.

120. Miskovic L, Tokic M, Fengos G, Hatzimanikatis V (2015) Rites of passage: requirements and standards for building kinetic models of metabolic phenotypes. Current Opinion in Biotechnology 36: 146-153.

121. Morgan J.J, Surovtsev I.V, Lindahl P.A (2004) A framework for whole-cell mathematical modelling. J theor Biology 231(4): 581-596.

122. Tomita M, Hashimoto K, Takahashi K, Shimizu T, Matsuzaki, Y (1999) E-Cell: Software environment for whole cell simulation. Bioinformatics. 15(1): 72-84.

123. Jahan M, Maeda K, Matsuoka Y, Sugimoto Y, Kurata H (2016) Development of an accurate kinetic model for the central carbon metabolism of Escherichia coli. Microb Cell Fact. 15(1):112.

124. Aach J, Rindone W, Church G.M (2000) Systematic Management and Analysis of Yeast Expression Data. Genome Res 10(4): 431-445.

125. Covert M.W, Knight E.M, Reed J.L, Herrgard M.J, Palsson B.O (2004) Integrating high-throughput and computational data elucidates bacterial networks. Nature 429(6987): 92-96.

126. Mendes P (1999) The Metabolic Control Analysis Web, Course notes, (In: Snoep, J., Mendes, P., Westerhoff, H., 1999, Teaching MCA and kinetic modelling: toward a portable teaching module. The Biochemist. Feb. 1999).

127. Schaff et al. (2001) Virtual cell (V-Cell) Project. 1st Intl. Symp. on Computational Cell Biology, NIH-NRCAM.

128. Hucka M, Finney A, Sauro H.M, Bolouri H, et al. (2003) The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models. Bioinformatics. 19(4): 524-531.

129. Schuster S, Fell D.A, Dandekar T (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nature Biotechnology 18(3): 326-332.

130. Papin J.A, Price N.D, Wiback S.J, Fell D.A, Palsson B.O (2003) Metabolic pathways in the post-genome era. Trends in Biochemical Sciences 28(5): 250-258

131. Perret C.J, Levey H.C (1961) The theory of uncatalysed linear expanding systems. J Theor Biol 1: 542-550.

132. Grainger J.N.R, Bass, L (1966) A model of a growing steady state system. J. Theor. Biol. 10(3): 387-398.

133. Grainger J.N.R, Gaffney P.E, West T.T (1968) A model of a growing steady-state system with a changing surface-volume Ratio. J Theor Biol 21(1): 123-130.

134. Aris R (1969) Elementary chemical reactor analysis, New Jersey: Prentice-Hall.

135. Kacser H, Beeby R (1984) Evolution of catalytic proteins or on the origin of enzyme species by means of natural-selection. J Mol Evol 20(1): 38-51.

136. Brumen M, Svetina S, Heinrich R (1984) Relations among variations in red-cell properties—a comparison of the cation impermeable and metabolic osmotic models, Biomed. Biochim. Acta 43: S13-S14.

137. Werner A, Heinrich R (1985) A kinetic model for the interaction of energy-metabolism and osmotic states of human erythrocytes—analysis of the stationary invivo state and of time-dependent variations under blood preservation conditions. Biomed Biochim Acta 44(2): 185-212.

138. Joshi A, Palsson B.O (1989a) Metabolic dynamics in the human red cell. 1. A comprehensive kinetic model. J. Theor. Biol. 141(4): 515-528.

139. Joshi A, Palsson B.O (1989b) Metabolic dynamics in the human red cell. 2. Interactions with the environment. J Theor Biol 141(4): 529-545.

140. Joshi A, Palsson B.O (1990a) Metabolic dynamics in the human red cell. 3. Metabolic Reaction Rates. J Theor Biol 142(1): 41-68.

141. Joshi A, Palsson B.O (1990b) Metabolic dynamics in the human red cell. 4. Data Prediction and some model computations. J Theor Biol 142(1): 69-85.

142. Wallwork S.C, Grant D.J.W (1977) Physical Chemistry. London: Longman.

143. Myers C.J (2009) Engineering genetic circuits, Boca Raton (USA): Chapman & Hall/CRC Press.

144. Tiruvadi-Krishnan S, Maennik j, Kar P, Lin J, Amir A (2022) Coupling between DNA replication, segregation, and the onset of constriction in Escherichia coli. Cell Rep 38(12): 110539.