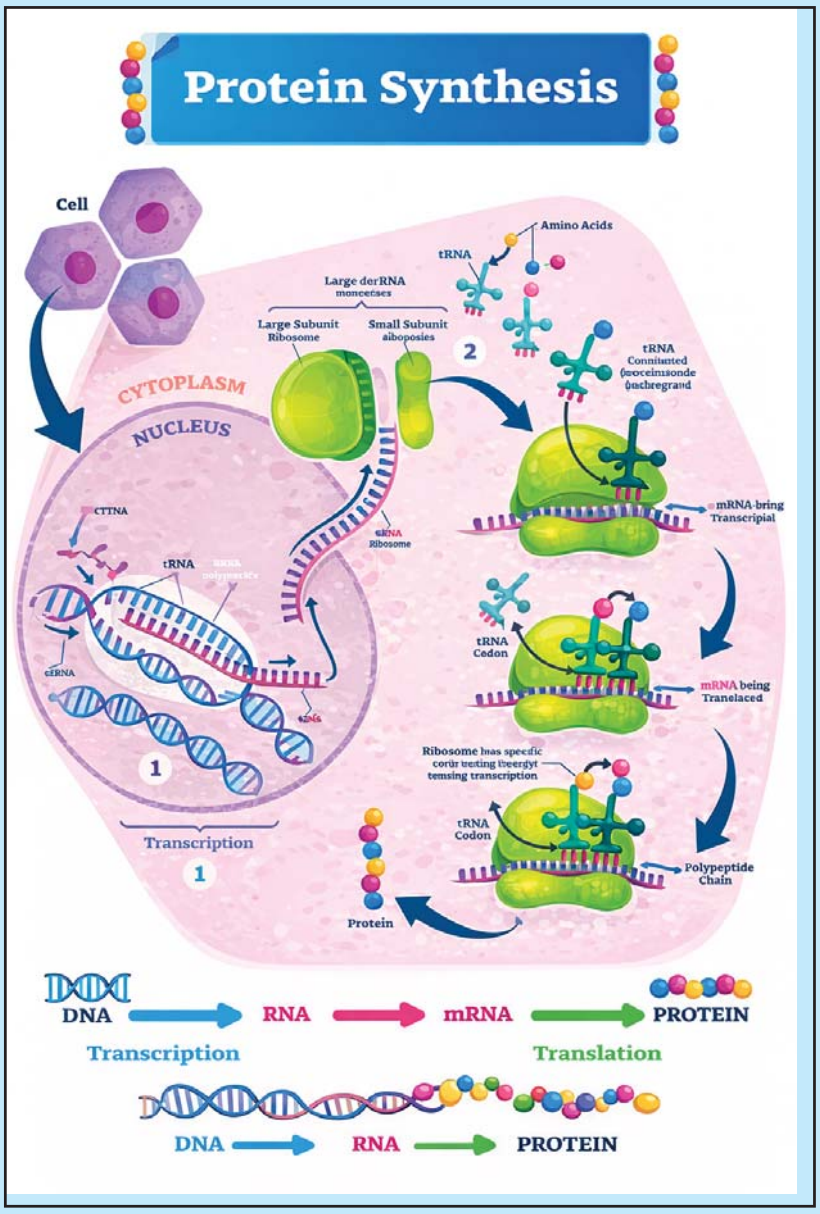
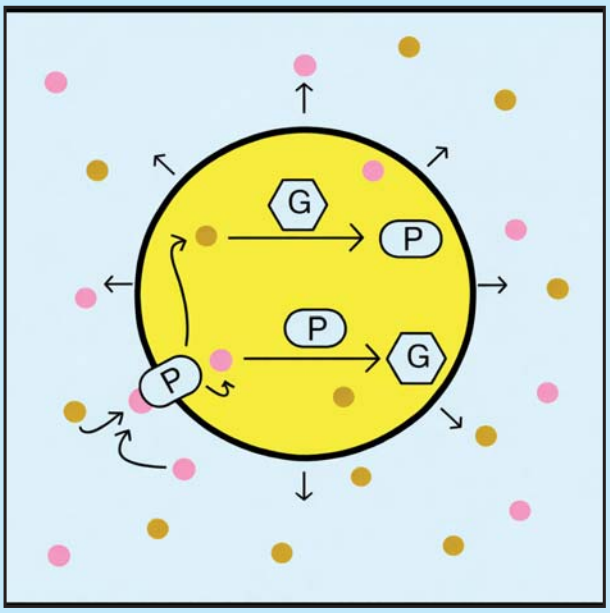
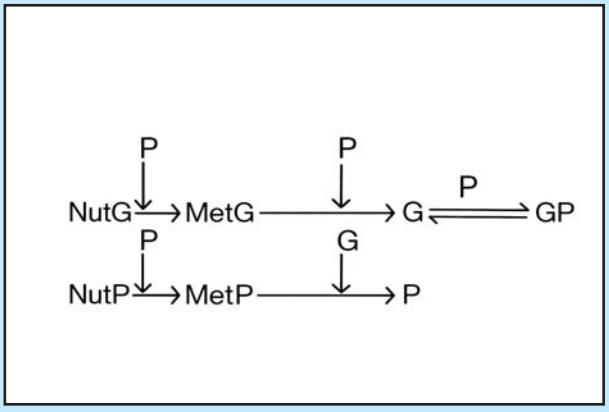
Biology and Medicine

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







Gheorghe MARIA

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Biology and Medicine

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

Gheorghe Maria*

Department of Chemical and Biochemical Engineering, University Politehnica of Bucharest, Romania

Corresponding authors : Gheorghe Maria, Department of Chemical and Biochemical Engineering, University Politehnica of Bucharest, Romania, E-mail: gmaria99m@hotmail.com

Keywords: Modular, Deterministic, Structured, Hybrid (SMDHKM) Dynamic Models; "mechanistic cell" novel modelling concept; Novel kinetic modelling framework "WCVV"; Gene expression regulatory modules (GERM); Genetic Regulatory Circuits (GRC); Central Carbon Metabolism (CCM); Genetically Modified Micro-Organisms (GMO); Bioreactor optimisation; GERM Regulatory Performance Indices;

Printed: United States

Submitted: 17 October, 2025
Accepted: 06 January, 2026
Published: 13 January, 2026

Cover Figures Sources: Maria et al. (2002); Maria,2005,2006,2007. https://st4.depositphotos.com/3900811/22071/v/1600/depositphotos_220715852-stock-illustration-protein-synthesis-vector-illustration-transcription.jpg

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table of Contents — Page Nos

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Abstract

This paper reviews the novel concept of "mechanistic 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" hypotheses of the "WCVV" refer to a novel kinetic modelling framework of the cell metabolic processes - that maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. This WCVV novel modelling framework was introduced by Maria [2002], and by [Maria et al. 2002], while the study of the WCVV properties, features, and advantages was given by Maria over the interval [2005-2023] in papers/booklets published with an almost annual frequency, that is Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c].

This work 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. The more realistic WCVV approach is reviewed when developing modular kinetic representations of the homeostatic gene expression regulatory modules (GERM) that control the protein synthesis, and the homeostasis of metabolic processes. This work reviews the general concepts and hypotheses of the WCVV modelling framework, while the cited literature includes past and current experience with modelling and simulation of the GERM regulatory properties targeted to maintain the intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts.

This work also reviews the linking rules to be used to obtain optimized globally effective kinetic models for the genetic regulatory circuits (GRC) that can reproduce the experimental observations of cell homeostatic stability, and its adaptation to environmental changes (e.g. genetic switches). Based on the novel defined quantitative regulatory indices evaluated *vs.* simulated dynamic and stationary environmental perturbations, this work exemplifies with some GERM-s from E. coli, at a generic level, how this methodology can be extended to other cells.

These contributions, and the novel WCVV modelling framework 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, or of GRCs. The work proves, by using a solved example, how the novel "mechanistic silicon cell", and the WCVV modelling concepts/framework, can be used to build-up complex structured, modular, deterministic, hybrid kinetic models (SMDHKM). Such dynamic hybrid models can be used for various *in-silico* analyses, such as:

I. *in-silico* design of genetically modified micro-organisms (GMO);

II. to characterize the GERM regulatory efficiency, and various local or global (cell level) regulatory properties, such as:

a. The role of the high cell-"ballast" in "smoothing" the effect of internal/external dynamic/stationary perturbations of the cell system homeostasis;

b. The secondary perturbations are transmitted inside the whole cell via the cell volume;

c. 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;

d. Allows comparing the regulatory efficiency of various types of GERM-s;

e. Allows a more realistic evaluation of GERM regulatory performance indices;

f. allows studying the recovering/transient intervals between steady-states (homeostasis) after stationary perturbations;

g. allows studying the species connectivity;

h. allows studying the conditions when the system's homeostasis is lost;

i. allows studying the cell / GERM-s self-regulatory properties after a dynamic/stationary perturbation, etc.,

j. Allows studying the plasmid-level effects in cloned cells.

k. Allows simulation of the cell homeostasis (balanced growth), which maintains the intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts.

III. to build-up modular regulatory chains (GRC-s) of various complexity. An example is provided;

IV. to prove feasibility of the cooperative *vs.* concurrent construction of GRCs that ensures an efficient gene expression, system homeostasis, proteic functions, and a balanced cell growth during the cell cycle;

V. to prove that the classic ("default") kinetic modelling approach (WCCV, "constant-cell-volume, non-isotonic system") is erroneous compared to the novel WCVV, by leading to distorted and wrong simulation results and conclusions [Maria et al., 2018d]. In such a way, the novel WCVV cell kinetic modelling approach replaces the wrong classical (default) WCCV modelling methodology, thus correcting the distorted and false/wrong predictions of WCCV cell kinetic models largely used in the literature;

VI. 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 [example of Maria and Luta, 2013], or from using the literature data [e.g. Yang et al., 2003], or bio-omics databanks [like KEGG, Ecocyc]. The rate constants of the GERM / GRC modular kinetic models were estimated from solving the steady-state (QSS) model equations with using the stationary (homeostatic) concentrations of the key-species [from bio-omics databanks], and by imposing some holistic regulatory properties [Maria and Scoban, 2017,2018].

VII. The use of the novel WCVV math modelling framework to obtain modular, deterministic, structured, hibrid SMDHKM 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 used for *in-silico* engineering evaluations for various purposes, such as:

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

A. To simulate the essential parts of the central carbon metabolism (CCM), such as: glycolysis, TCA cycle, ATP recovery system, under stationary, or oscillating conditions Maria, 2014a, 2021];

B. To simulate the dynamics of various cell systems, and to generate the cell metabolic flux analysis aiming to *in-silico* design of GMOs; examples include: the mercury operon expression Maria and Luta,2013]; the tryptophan operon expression [Maria and Renea, 2021; the succinate synthesis [Maria et al.,2011], etc.;

C. To simulate and/or design various genetic regulatory circuits (GRCs). Examples include: genetic switches (biosensors, [Maria, 2014b]), genetic amplifiers (ex. Mercury-operon quick expression in E. coli, [Maria and Luta,2013; Maria,2023), etc.

D. To *in-silico*, off-line derive optimal operating policies of bioreactors, with a higher precision and level of detail [Maria and Luta,2013; Maria and Renea, 2021].

Foreword

As discussed in section 1.3.3, 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).

I. 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 [Banga, 2008] (Figure 1-23), as followings:"The science of discovering, modelling, understanding and ultimately engineering at the molecular level the dynamic relationships between the biological molecules that define living organisms" [Leroy, 2017].

II. "System Biology is a comprehensive quantitative analysis of the manner in which all the components of a biological system interact functionally over time" [Zak and Aderem, 2009].

III. "Perhaps surprisingly, a concise definition of Systems Biology that most of us can agree upon has yet to emerge" [Ruedi Aebersold, Inst. Systems Biology, Seattle; http://www.systemsbiology.org).

"Systems biology, as defined by Ruedi Aebersold, 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 1-7), Systems biology [Kitano, 2002b], Computational biology [Kitano, 2002], and eventually Synthetic biology [Benner and Sismour, 2005], the *in-silico* analysis (based on mathematical models) of cell processes is experiencing an impetuous development. The used math models are of various types (section 1.3.4): topological, or structured (deterministic with continuous "variables"); Boolean (with discrete variables; stochastic variables; mixed variables.

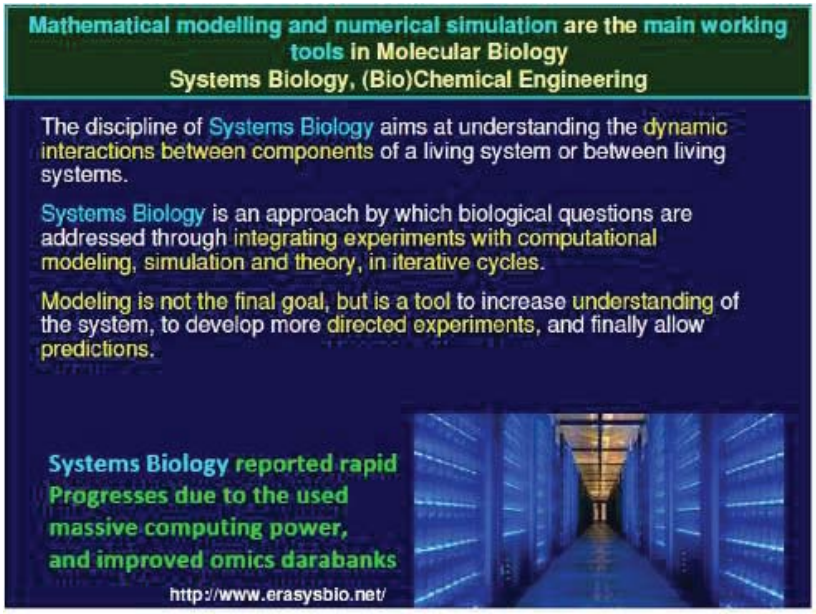


Figure 1-23: Math models used in Systems Biology requires massive computational efforts and facilities [Maria, 2017b, 2018].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

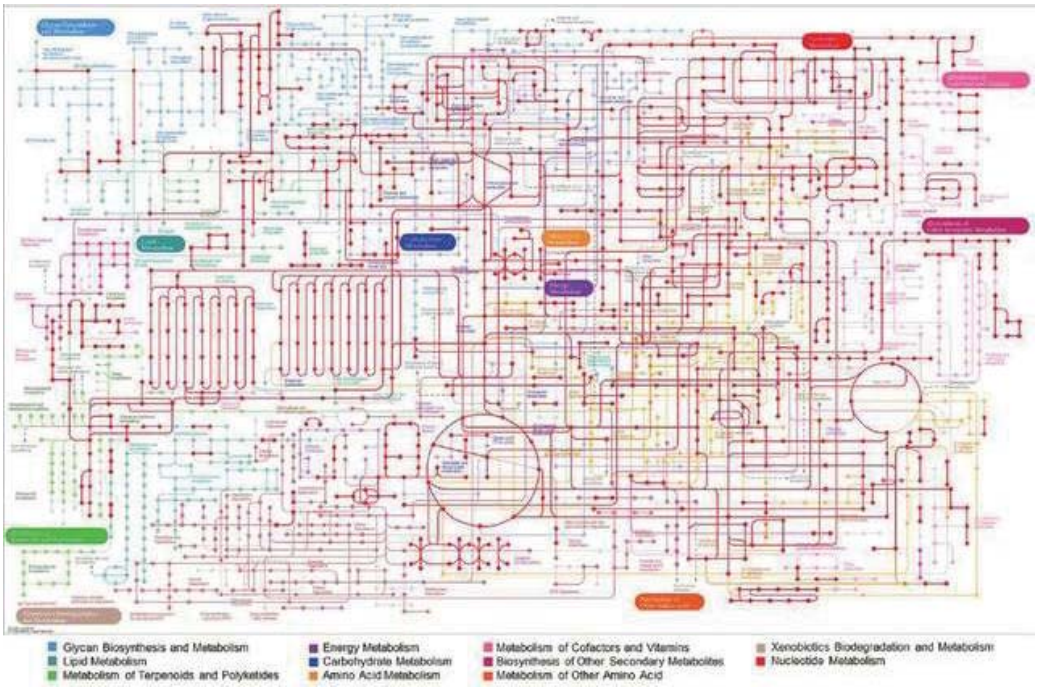


Figure 1-17: Central carbon metabolism in *Haematococcus pluvialis* cell., by using an "electronic circuit-like" model.
This model was included in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of *H. pluvialis* transcripts. Isotigs and contigs of expected values (E-values) below the threshold of 10 −6 were analysed. *H. pluvialis* expressed transcripts for enzymes functioning in most of the major metabolic pathways, including N - or O -glycan, lipid, carbohydrate, energy, amino acid, nucleotide, cofactor and vitamin, and isoprenoid metabolism. (Source= https://www.researchgate.net/figure/Kyoto-Encyclopedia-of-Genes-and-Genomes- KEGG-pathway-analysis-of-H-pluvialis_262267927)

Extended simulation platforms have been reported, able to simulate the cell processes, such as

- E-Cell (compartments, compounds, genes, reactions), able to simulate the whole or parts of the cell metabolism) [Tomita et al., 1999; Tomita, 2001; Kinoshita et al., 2001] (Figure 1-36)
- V-Cell (model, geometry, applications, biological interface) [Schaff et al., 2001].
- M-Cell (stochastic simulator of some cell sub-systems) [Bartol and Stiles, 2002] (Figure 1-37).
- A-Cell ("electrical circuit" like models) [Ichikawa, 2001] (Figure 1-37).
- Silicon-Cell (computer simulation of cell processes) [Westerhoff, 2001, 2006; Westerhoff and Palsson, 2004] (Figure 1-35, and Figure 1-36)

Following important progresses made in the computing rules, numerical algorithms, and computational speed, a large number of insilico studies (based on mathematical models) about cell processes have been reported. As Palsson [2000] noted, such models of the cell metabolic functions / processes are constrained by a lot of physicochemical 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 [2002], and by [Maria et al., 2002], while the study of the WCVV properties, features, and advantages was given by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]. By using the classical modelling approach WCCV (section 2.1), a large number of rather deterministic math 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. [2003] for *Saccharomyces cerevisiae*; Edwards and Palsson [2000]; [Edwards et al., 2001] for *Escherichia coli*; Heinemann et al. [2005] for *Staphylococcus aureus*; RodriguezPrados et al. [2009] for threonine synthesis in *E. coli*.

Also, dedicated software / rules to perform such a genetic network reconstruction, or of the metabolic networks have been reported by [Rocha et al., 2010; Le Phillip et al., 2004; Nemenman et al, 2007].

Consequently, the several review studies have set the milestones [Palsson, 2000; Camacho et al., 2007] for the development of this new research topics (at the interface between several classical fields (biology, control theory of nonlinear systems, numerical computation, (bio)chemical engineering, bioinformatics), that is "the Silicon cell" of Westerhoff [2006], or the "mechanistic silicon whole cell" of Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]. It is worth mentioning that

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

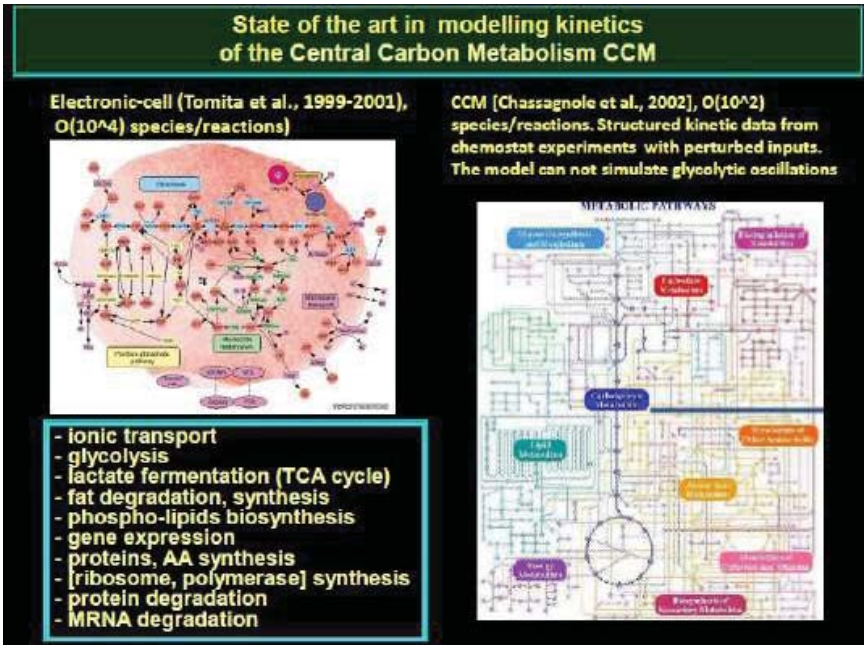


Figure 1-36: Some simulators of the CCM. [Left]. The E-cell of [Tomita et al., 1999]. [Right] The Silicon-cell of [Westerhoff, 2001, 2006; Westerhoff and Palsson, 2004; Chassagnole et al., 2002; Schmid et al.,2004; Visser et al., 2004].

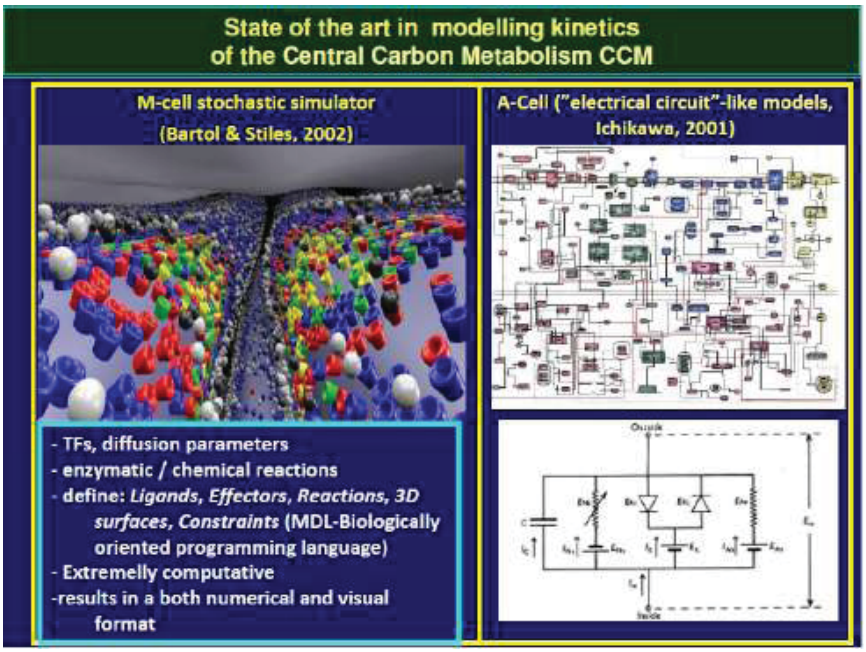


Figure 1-37: Some cell simulators of stochastic type. [Lleft]. The M-cell of [Bartol and Stiles, 2002]. [Right] The "electric-circuit" like A-cell model of [Ichikawa, 2001].

in all *in-silico* alternatives to build-up a cell model, this is an iterative process requiring (Figure 2-48) experimental validation of the essential model parts. According to Westerhoff [2001, 2006], the *in-silico* approach of most of math models follows the chemical and biochemical process engineering concepts, rules, and numerical algorithms (section 1.3.4). Below is reproduced in full his keynote.

"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), 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.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

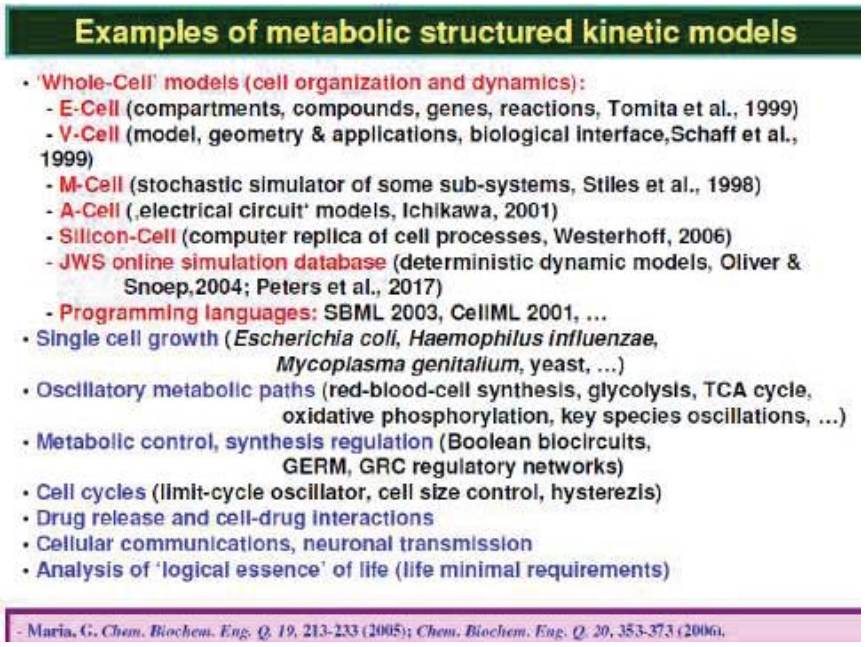


Figure 1-35: Some examples of structured cell simulators. See the review of Maria, 2005, 2006].

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.

I will show that 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. I shall then venture on to *Hierarchical Control Analysis* and Hierarchical Regulation Analysis, both of which allow for adaptation. Perhaps unexpectedly, I shall show how the concepts and principles developed for Life processes also apply to inanimate processes, yielding new principles for process engineering in general.

The presentation will make ample use of a new tool in bioprocess engineering, i.e. the silicon cell (cf. 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*. For the particular case of the engineering of the production of baker's yeast I shall discuss some of the implications of these new approaches." [Westerhoff, 2006].

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], make now possible to record the dynamics of hundreds of cell species, thus generating kinetics trajectories required by the development of cell CCM / GRC dynamic models on a structured basis, and using structured and conventional type of data [Maria, 2023].

As discussed by Kiparissides et al. [2011], 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 [Thilakavathi et al. 2007]. 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 [Covert et al., 2001]. Bailey [1998] 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-organsisms, 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 [Yokobayashi et al., 2003]. Therefore, attempts to wholly model the function of even a single cell are currently non-trivial, if not impossible. The amount of delicate intracellular

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

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 [Sidoli et al., 2004] and on the mechanistic level further complicate this task.

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 [Xie and Wang, 1994], to identify previously ignored growth limiting factors [de Zengotita et al., 2000), to optimize culture growth and productivity [De Tremblay et al., 1992, Dhir et al., 1999, San and Stephanopoulos, 1989], and to apply the process control principles to operate the bioreactors [Frahm et al., 2002].

Pörtner and Schäfer [1996] 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." Kiparissides et al. [2011]

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, especially those related to the central carbon metabolism (CCM) [Maria, 2021; Maria and Renea, 2021], and to complex genetic regulatory circuits GRC-s [Maria, 2010; Maria and Luta, 2013].

In this context, the current work (ebook) is aiming to present a novel WCVV ("whole-cell-variable-volume") kinetic modelling framework, proposed by Maria [2002], and by [Maria et al., 2002], and studied by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; [Maria and Scoban, 2017, 2018; Maria and Luta, 2013], proven to be very effective for accurately model the individual gene expression regulatory modules (GERM-s), but also the complex GRC-s. These WCVV kinetic models can include the "Hierarchical Regulation Analysis", by linking GERM-s in complex GRC-s, under the novel WCVV modelling approach. Thus, the WCVV approach allows modelling the adaptable living cells, by also including genetic switches models able to re-direct the cell metabolism [Maria, 2007, 2009, 2014b, 2018].

1. Generalities

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 [Maria, 2018,2018f]. Where biology enters the arena other than as the usual source of the raw-materials (oil, coal, natural gas) [Maria et al., 2020; Yang, 2007], 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 [Maria, 2017A, 2017B, 2018, 2023]. 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 [Maria, 2017A, 2017B, 2018, 2023]; [Maria, 2023a, 2024, 2024b, 2024c]. Flux balance analysis does not do so either; its pathways are given and assumed to be always expressed." [Westerhoff, 2001; Westerhoff and Palsson, 2004; Zadran and Levine, 2013].

Here it is worth notice that living cells are extraordinarily complex machines, with an astronomical complex organization (Figure 1-2), that exhibit remarkable behaviors including, most predominately, their self-replication. 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 [Palsson, 2000; Tomita, 2001; Kitano, 2002,2002b; Mori, 2004; You, 2004].

"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 [Ideker et al., 2001; Jansen, 2003; Lindon et al., 2004; Rao et al.,2005]." [Surovstev et al.,2007]"

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

<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 rowspan="2">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></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 1-2: The hierarchical organization of living cells [Maria, 2017b].

"Much remains before useful comprehensive mathematical models of cellular behaviour with predictive power can be realized. Obviously, the vast number of unknown kinetic parameters (rate coefficients and 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 physicochemical 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." [Surovstev et al.,2007]

In this work, the novel WCVV ("mechanistic silicon Whole-cell variable volume") modelling framework will then be introduced.

This WCVV novel modelling framework was proposed by Maria [2002], and by [Maria et al., 2002], while the study of the WCVV properties, features, and advantages was given by Maria over the interval [2005-2023], that is Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c].

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). This WCVV novel modelling framework was introduced by Maria [2002], and by [Maria et al., 2002], while the study of the WCVV properties are given by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c].

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 cell", materialized in a novel math (kinetic) modelling framework "WCVV" of the cell metabolic processes, referring to the "wholecell" (WC), variable-volume (VV), of isotonic growing cells. The "mechanistic silicon cell" novel hypotheses of "WCVV" refer to the wholecell, variable-volume, of isotonic growing cells. The paper points-out and proves the advantages of WCVV dynamic models when buildingup modular model structures of simplified form that can reproduce complex protein synthesis regulation inside cells.

Such novel modelling concept of "WCVV - mechanistic silicon cell", proves to be a valuable simulation tool that can overcome many of the simulation difficulties of the dynamics of the evolutionary cellular metabolism mentioned above.

In this work, 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,

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

and the homeostasis of metabolic processes. The paper briefly reviews the general concepts and hypotheses of the WCVV modelling framework, while the cited literature includes past and current experience with GERM regulatory properties, and with the linking rules to be used to obtain optimized globally efficient kinetic models for the genetic regulatory circuits (GRC) that can reproduce the experimental observations. Based on the novel defined quantitative regulatory indices evaluated *vs.* simulated dynamic and stationary environmental perturbations, the paper exemplifies with using simple kinetic models of GERM-s from *E. coli* (by emploing the required -omics data from the Ecocyc, and KEGG databanks). Simple kinetic models at a generic level, were also used to exemplify how this methodology can be extended to reach the following objectives.

The novel "WCVV - mechanistic silicon cell" concept and 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, or of GRCs. As proved in the literature, this WCVV framework is very suitable to model the GRCs [Maria, 2017a, 2017b,2018,2005,2006,2007,2009,2014], but also the modular CCM included in hybrid kinetic models (SMDHKM) [Maria, 2014a, 2021; Maria et al., 2011; Maria and Renea, 2021]. These references proved that the modular modelling approach of CCM is very advantageous, by simplifying the numerical simulation. In fact, as revealed by the (Figure 1-30-32), the main 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 [Maria et al., 2011; Maria, 2014a, Maria, 2021; Maria and Renea, 2021].

Concerning the cell adaptation to the environmental conditions, i.e. the changing of the network structure / behaviour according to the environmental changes, the novel "WCVV - mechanistic silicon cell" math modelling framework, concept, and rules was proved to be very adequate to simulate the metabolism changes (dynamics after perturbations, biomass evolution), caused by the environmental changes, such as:

1) Changes of metabolic pathways determined by the genetic switches (GS) action, induced by the environmental changes [Maria, 2007,2009,2014b, 2018];

2) Oscillatory syntheses induced by the environmental changes [Maria et al., 2018a, 2018b, 2018e; Maria, 2014a, 2020c; Maria et al., 2018c; Atkinson et al., 2003; Guantes and Poyatos, 2006; Elowitz and Leibler, 2000; Gonze, 2010; Bier et al., 1996; Silva and Yunes, 2006; Heinzle et al., 1982; Tyson, 2002; Franck, 1980];

3) GMO design based on the target GRC regulatory adaptive efficiency (see an example for the mercury uptake by *E. coli* given by [Maria,2010; Maria and Luta, 2013; Maria,2023]);

4) 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. The process being better characterized by using the quantitative regulatory performance indices (P.I.-s) of GERM-s, simulated under WCVV modelling framework by [Maria, 2017a, 2018,2023; Maria and Scoban, 2017,2018];

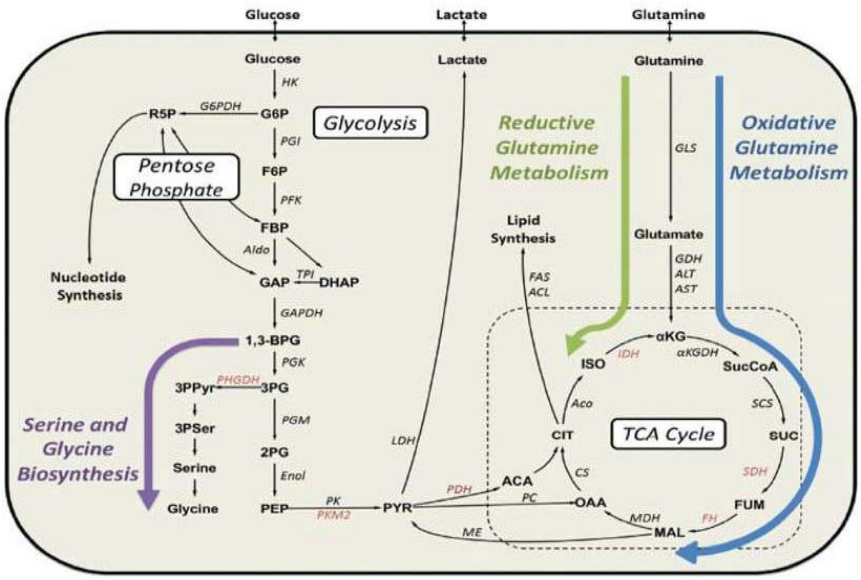


Figure 1-30: Example of the modular organization of the CCM considered in the WCVV hybrid kinetic models (SMDHKM) by [Maria, 2014a; Maria, 2021; Maria et al., 2011; Maria and Renea,2021] [source: https://www.creative-proteomics.com/services/central-carbon-metabolism-analysisservice.htm].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

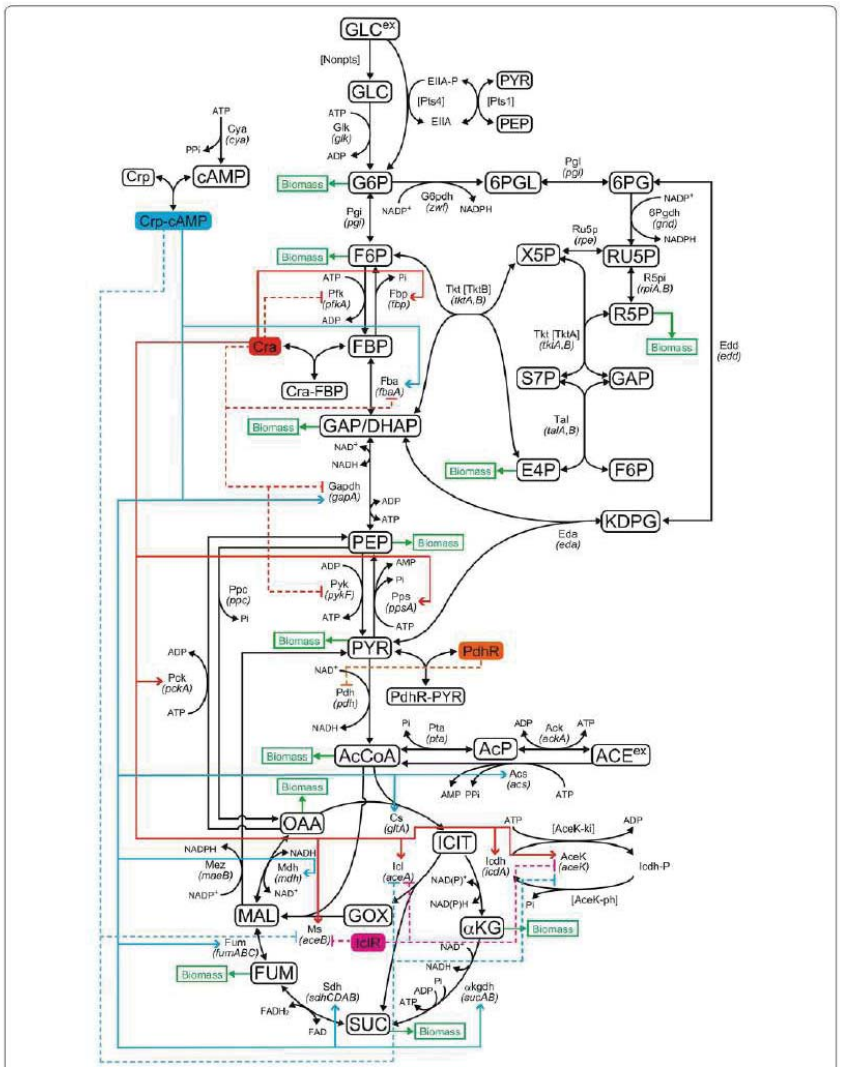


Figure 1-31: Summary diagram of some modules of the CCM (PTS, glycolysis, and TCA cycle) in the "wild" E. coli according to the simplified model of [Jahan et al., 2016]. The ATP regeneration system is shown in Fig. 1-32. The dotted lines indicate the repressor feedback circuits.

5) Better simulation of the (P.I.-s) of GERM-s, defined under WCVV modelling framework by [Maria, 2017A,2017B,2018,2023; Maria, 2023a, 2024, 2024b, 2024c]. As a consequence, the novel "WCVV mechanistic silicon cell" math modelling framework was also proved to be very effective in design effective GMOs with industrial applications [Maria and Luta, 2013; Maria,2021; Maria and Renea,2021].

6) Simulation of a hypothetical "Mechanistic silicon Cell" which is able to maintain the intracellular homeostasis (balanced growth), while growing auto-catalytically on environmental nutrients present in variable amounts [Maria,2002; Maria et al., 2002], and Maria [2017A,2017B,2018,2023, 2023a, 2024, 2024b, 2024c].

1.1. «Mechanistic silicon cell» novel concept and the novel WCVV math modelling framework - Generalities

The "WCVV - mechanistic silicon cell" novel math modelling framework is an outstanding and essential/fundamental contribution of Maria [2002], and [Maria et al., 2002], while further developments/ studies of the WCVV properties were given by Maria over the interval [2005-2023], that is Maria [2002,2003,2005,2006,2007,2009,2014b, 2017A, 2017B, 2018, 2023, 2023 a, 2024, 2024b, 2024c].

The cell kinetic modeling framework "mechanistic silicon whole cell", materialized in the novel "WCVV" modelling framework of cell processes, implicitly is accounting for the isotonic growing cell (of variable-volume). Such WCVV cell models implicitly maintain the intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. In this novel 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).

The main contributions / findings of Maria [2002], and [Maria et al. 2002], derived from the study of the WCVV properties are

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

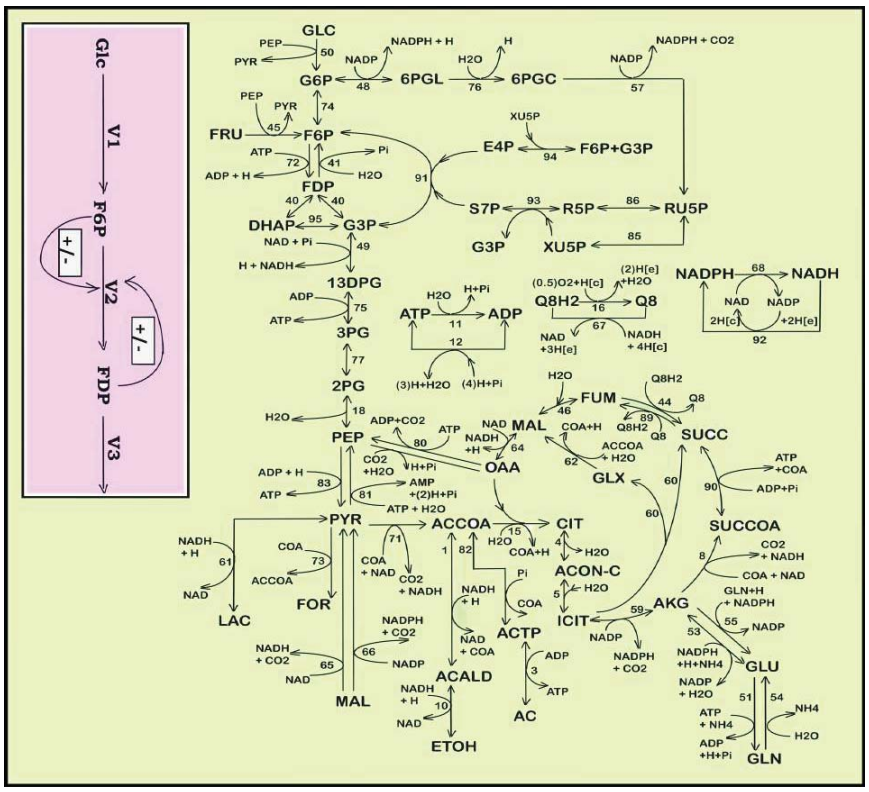


Figure 1-32: Simplified reaction pathway of modular CCM in E. coli considered in the kinetic model of [Maria et al., 2011; Maria et al., 2018A]. This simplified model of Edwards and Palsson [2000] refers to the "wild" strain that includes the glucose (GLC) import PTS system, respectively the sequence of successive reactions GLC → F6P. The fluxes characterizing the membrane transport [Metabolite(environ.) → Metabolite(cyt.)], and the exchange with the environment have been omitted. See [Maria et al., 2011; Maria, 2014A] for details and explanations regarding the numbering of the 95 reactions in which the 72 key metabolites considered in the kinetic model are involved. Notations: [e]= external environment of the cell; [c]= cytosol. Adapted from [Maria et al., 2011] by the courtesy of CABEQ JI. The pink rectangle indicates the "chemical node" that generates glycolytic oscillations [Maria, 2021; Maria et al., 2018E]. The (+) and (-) notations designate the positive and negative feedback/feedback loops. GLC = glucose; F6P = fructose-6-phosphate; FDP = fructose-1,6-diphosphate. See the abbreviation list of [Maria et al., 2011]. V1-V6 = grouped reaction rates are discussed by [Maria, 2021]. Species notations are explained in the abbreviation list of [Maria, 2021].

given by Maria over the interval [2005-2023] (see the above reference list). The main conclusions when modelling metabolic cell processes, and especially the regulatory circuits GERMs/GRCs, when using the WCVV framework can be briefly reviewed as following.

- Fundamental theoretical contributions for the *In-silico* (math model based) dynamic simulation of cell metabolic processes. In particular, dr. Maria introduced and validated the novel "mechanistic silicon cell concept", **materialized in a novel math (kinetic) modelling framework WCVV** of the cell metabolic processes (referring to the "whole-cell, variable-volume", of isotonic growing cells) (see their hypotheses in Table 2-1 of section 2.1).

- The WCVV is particularly suitable to characterize the regulatory properties of selfregulated individual gene expression (GERM), and of genetic regulatory circuits (GRC) (e.g. genetic switches, genetic amplifiers, operon expression, etc.) Dr. Maria also proved that the novel WCVV kinetic modelling framework of cell processes is superior compared to the classical (default) WCCV ("whole-cell, constant-volume" like) kinetic modelling framework, the latter offering distorted, and wrong simulation results. [Maria et al., 2018d; Maria and Scoban,2017, 2018].

- Use of instant cell dilution "Di" in the cell species mass balances of WCVV models instead of the "default" average dilution Dm = Ln(2)/(tc), with tc = cell cycle, or even its omission in the classical WCCV models {see below equations and the Figure 1-3 , the below discussion, and the details of [Maria et al., 2018d; Maria and Scoban,2017,2018], and the books [Maria, 2017a, 2017b, 2018, 2023].

- Applications of *hybrid* structured modular, deterministic kinetic models (SMDHKM) to *in-silico* design of GMOs, and for a more accurate optimization of industrial bioreactors operation, with a higher degree of detail (no. of cell/bioreactor state variables).

- Being very suitable to simulate GERM/GRC, the WCVV was used for the development of a math "LIBRARY" including "template" kinetic models of the main GERM types (Figure 1-4-down). This library is useful to build-up dynamic models of GRC (genetic regulatory circuits), useful for *in-silico* design of GMOs. The use of the novel WCVV math modelling framework to define

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 2-1: 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 [Maria,2005], with the details given by [Maria,2006, 2007, 2014b]; [Maria and Scoban, 2017,2018] ; Maria et al. [2018d, 2018g].

Mass balance and state equations	*Remarks*
$\frac{dC_j}{dt} = \frac{1}{V}\frac{dN_j}{dt} - D_i C_j = g_j(\boldsymbol{C},\boldsymbol{k})$; $\frac{1}{V}\frac{dN_j}{dt} = \sum_{j=1}^{nr} \nu_{ij} r_j(C,k)$; $j = 1,\ldots,n_s$;	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 [Wallwork and Grant, 1977]
$D_i = \frac{1}{V}\frac{dV}{dt} = \left(\frac{RT}{\pi}\right)\sum_{j}^{n_s}\left(\frac{1}{V}\frac{dN_j}{dt}\right)$	D_i = cell content instant dilution rate = cell volume logarithmic instant growing rate
$\frac{RT}{\pi} = \frac{V}{\sum_{j=1}^{n_s} N_j} = \frac{1}{\sum_{j=1}^{n_s} C_j} = \frac{1}{\sum_{j=1}^{n_s} C_{jo}}$ = 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")

Hypotheses:

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)[Surovstev et al.,2007];
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.(2), leading to an average volume growth of $V = V_o e^{D_m t}$ over the 80% of the cell cycle [Surovstev et al.,2007];
g. -- homeostatic stationary growth of $(dC_j/dt)_s = g_j(\boldsymbol{C}_s,\boldsymbol{k}) = 0$;
h. -- perturbations in the cell volume are induced by variations in species copynumbers (number of moles) under the isotonic osmolarity constraint: $V_{perturb}/V = (\sum N_j)_{perturb} / (\sum N_j)$; see the proof of [Maria and Scoban,2017,2018].

Notations:
Cj = cell species "j" concentration;
Di = cell content instant dilution rate = cell volume logarithmic instant growing rate, Eq.(2,10);
Dm = cell content average (apparent) dilution rate = cell volume logarithmic average growing rate Eq.(2);
nr = number of reactions considered in the WCVV kinetic model (individual or lumped);
ns = number of species considered in the WCVV kinetic model (individually, or lumped);
k = rate constants vector;
Nj = the species " j" number of moles (copynumbers);
R= universal gas constant;
rj = the j -th reaction rate;
T= temperature;
t= time;
tc = cell cycle;
V= cell (cytosol) volume (including the membrane);
\pi = osmotic pressure of the cell content;
vij = the stoichiometric coefficient of the species " j" in the reaction "i";
"s" index = quasi-steady-state (cell homeostasis; balanced growth)

novel quantitative performance indices (P.I.-S)(section 2.3) to better characterize the regulation efficiency of individual GERM-s , or of GRC-s related to external perturbations [dynamic ("pulse-like"), or stationary("step-like")]. The use of the GERM library and their P.I.-S to propose a "**buiding-blocks**" strategy and **rules** to connect GERM-s to build-up GRC when designing desirable GMO-s (section 2.3). The use of the GERM -library to study various properties of cell self-regulation effects, impossible to be highlighted by using classical cell WCCV models (section 2.4).

- In other words, this work presents a holistic "closed loop" approach that facilitate the control of the *in vitro* through the

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

in silico development of dynamic models for living cell systems, by deriving *hybrid*, structured, modular, deterministic cell kinetic models (SMDHKM) ,with continuous variables, linking the cell state-variables to the bioreactor ones.

- In a concrete way, the reported **major and essential** contribution of Dr. Maria in the theoretical / math modelling of the metabolism in the living cells (single cell analysis) is the following. He introduced and validated the novel "mechanistic silicon cell concept", materialized in a novel math (kinetic) modelling framework WCVV of cell metabolic processes (referring to the "whole-cell, variable-volume", of isotonic growing cells) (see their hypotheses in the Table 2-1 and in section 2). The WCVV is particularly suitable to characterize the regulatory properties of self-regulated gene (GERM) / operon expression, and of genetic regulatory circuits (GRC), but also the CCM, metabolic processes, in a holistic approach.

- This novel WCVV kinetic modelling approach implicitly ensures the cell natural tendency to reach its homeostasis, and accurately simulate the individual / holistic GRC regulatory properties, by including in a natural way constraints related to the cell system isotonicity, and the variable-volume in relationship to the all species reaction rates [Maria,2017,2017a,2017b,2018; Maria and Scoban,2017,2018] (section 2). Such an isotonicity constraint is required to ensure the cell membrane integrity, but also to preserve the homeostatic properties of the cell system (see their hypotheses in Table 2-1, and section 2), not by imposing "the total enzyme activity" or the "total enzyme concentration" constraints suggested in the literature. [Heinrich and Schuster, 1996]

- By contrast to the novel WCVV, the classical ("default") modelling tools of metabolic cell processes are based on the '*Constant Volume Whole-Cell*' (WCCV) continuous variable ODE dynamic models which, do not explicitly consider the cell volume exponential increase during the cell growth, and ignore the cell content dilution, or consider an average (apparent) dilution "Dm = In (2)/(tc)", where tc = cell cycle. As proved by Maria [2018d], such an approach may lead to **biased, distorted, and wrong** conclusions on the GERM/GRC's regulatory performances, thus making difficult the modular constructions of GRC-s by linking individual GERM-s.

- By contrast to the classical WCCV approach, the holistic '*whole-cell of variable volume*' (WCVV) modelling framework of dr. Maria [Maria,2017,2017a,2017b,2018; Maria,2005,2006,2007,2009,2014] has been proved to be more realistic and robust, by explicitly including in the model relationships the cell-volume growth, with preserving the cell-osmotic pressure (that is the cell membrane integrity). The added isotonicity constraint by Dr. Maria was proved to be essential for predicting more adequate performance regulatory indices of GERM-s or GRC-s. The essential part of the WCVV approach is to explicitly include in the metabolic process kinetic model, the isotonic constraint for variable volume cells, that is [Maria,2017,2017a,2017b,2018; Maria,2005,2006,2007,2009,2014].

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

(Pfeiffers' law of diluted solutions)

which, by derivation and division with *V* leads to [Maria and Scoban, 2017,2018]:

$$D_i=\frac{1}{V}\frac{dV}{dt}=\left(\frac{RT}{\pi}\right)^{ns}\sum_j\left(\frac{1}{V}\frac{dn_j}{dt}\right)$$ (instant cell dilution rate)

Being related to the following cell model (species mass balance):

$$\left(\frac{dC_j}{dt}\right)=\left(\frac{1}{V}\frac{dn_j}{dt}\right)-D_iC_j=h_j(C,k,t)$$

In spite of a continuous cell volume increase, given by the following equation derived by combining the above definitions:

$$V(t)=\left(\frac{RT}{\pi}\right)\sum_j^{ns}n_j(t)\Rightarrow\frac{RT}{\pi}=\frac{V(t)}{\sum_{j=1}^{ns}n_j(t)}=\frac{1}{\sum_{j=1}^{ns}C_j}=\frac{1}{\sum_{j=1}^{ns}C_{jo}}=constant$$

The osmotic pressure remains constant, because of the mass conservation.

At quasi-steady-state conditions (QSS-homeostasis, index "s"), the WCVV model becomes:

$$\left(\frac{dC_j}{dt}\right)_s=\left(\frac{1}{V}\frac{dn_j}{dt}\right)_s-DC_{js}=h_{js}\left(C_s,k,t\right)=0$$

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

with species reaction rates of $\left[\left(\frac{1}{V}\frac{dn_j}{dt}\right)_s = r_{js}(c_s,k,t)\right] J = 1,\ldots$, ns (no. of species); and the stationary cell dilution

$$D = D_{is} = \left(\frac{RT}{\pi}\right)^{ns}\sum_j\left(\frac{1}{V}\frac{dn_j}{dt}\right)_s$$

By contrast, the use of Dm = In (2)/(tc) = average (apparent) cell dilution rate can lead to biased results. The QSS model form is used to estimate their rate constants for known Cs, and also to derive the species concentrations stationary sensitivities $\left(\partial C_i / \partial C_{Nut_j}\right)_s$ to the changes in the environmental nutrients [Nut]j [Maria,2005,2006,2007,2017a,2017b,2018; Maria and Scoban, 2017,2018], to be used to evaluate the cell (GERM-s) regulatory performance indices (P.I.-s).

At quasi-steady-state conditions (QSS-homeostasis, index "s"), the WCVV model becomes:

$$\left(\frac{dC_j}{dt}\right)_s = \left(\frac{1}{V}\frac{dn_j}{dt}\right)_s - DC_{js} = h_{js}(C_s,k,t) = 0$$

with species reaction rates of $\left[\left(\frac{1}{V}\frac{dn_j}{dt}\right)_s = r_{js}(C_s,k,t)\right] J = 1,\ldots,ns$ (no. of species), and the stationary cell dilution

$$D = D_{is} = \left(\frac{RT}{\pi}\right)\sum_j^{ns}\left(\frac{1}{V}\frac{dn_j}{dt}\right)_s .$$

By contrast, the use of Dm = In (2)/(tc) = average (apparent) cell dilution rate can lead to biased results.

Notations: Cj = species "j" concentration; nj = species " j " number of moles; V = cell (cytosol) volume; R = universal gas constant; T = absolute temperature; π = osmotic pressure; t = time; Di = instant cell dilution rate; Dm = Ds = average (apparent) cell dilution rate; "s" index = at stationary QSS (homeostasis) conditions; ns= number of considered cell species (individually, or lumped); Nut = nutrient or another environmental species influencing the cell metabolism; "o" index = initial value.

- The hypotheses of the "mechanistic silicon cell concept" materialized in the novel WCVV kinetic modelling framework in living cells of isotonic variable volume are presented in the Table 2-1 (section 2). The basic relationships of the WCVV theory concern the species mass balances over isotonic constraints (see section 2). Such a constraint allows determining the "instant" cell dilution ("Di") at any moment, to be used during model solving (see section 2, and Figure 1-3). By contrast, the classical (default) WCCV kinetic models use the overall average ("Dm= Ln(2) (cell cycle)") in the species mass balance, or even omit this term.

- The "**mechanistic cell concept**", and WCVV math modelling of cell processes have a huge impact on the cell kinetic models predictions accuracy, and degree of detail, as proved by Maria [2017a,2017b, 2018; Maria et al., 2018d; Maria and Scoban, 2017,2018].

Extensive details about this novel WCVV approach are given by: i) [Maria, 2017a,2017b, 2018]; [Maria et al., 2018d]; [Maria and Scoban, 2017,2018]; ii) the proofs of WCVV superiority *vs.* WCCV given in section 2.3.1. By using this novel WCVV modelling framework, a lot of cell metabolic effects can accurately be simulated (section 2.4), such as: a) the relationships between the external conditions, species net synthesis reactions, osmotic pressure; b) cell content (ballast) influence on smoothing the continuous perturbations in external nutrient concentrations, etc. [Maria,2017a,2017b,2018]; [Maria and Scoban,2017,2018].

- The past and current (**default**) cell dynamic models, are formulated for a constant cell volume system (WCCV), by ensuring some holistic cell properties (such as homeostasis, self-regulation of syntheses, and of gene expression, perturbation treatment, etc.), by imposing lot of constraints, such as "the total enzyme activity" and "total enzyme concentration", etc. [Heinrich and Schuster, 1996]. By contrast, in the "**mechanistic silicon cell concept**", and WCVV math modelling framework, the thermodynamic isotonicity relationships/constraints are included in the cell kinetic model formulation. Thus, dr. Maria proved step-by-step in a MATHEMATICAL way how such constraints ensures cellular intrinsic properties in a natural way that is not derived from artificial hypotheses (as those above mentioned). Such concepts, and rules translated from (bio) chemical engineering principles and nonlinear systems theory are explained, proved, and exemplified in the section 1.2, and by [Maria,2017a,2017b,2018], and by [Maria, 2005,2006,2007,2009]. The WCVV kinetic models of modular GRC-s are successfully used to design GMOs of desirable characteristics. Some examples are mentioned by Maria [2017A,2017B,2018,2023], and by Maria [2023a, 2024, 2024b, 2024c].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

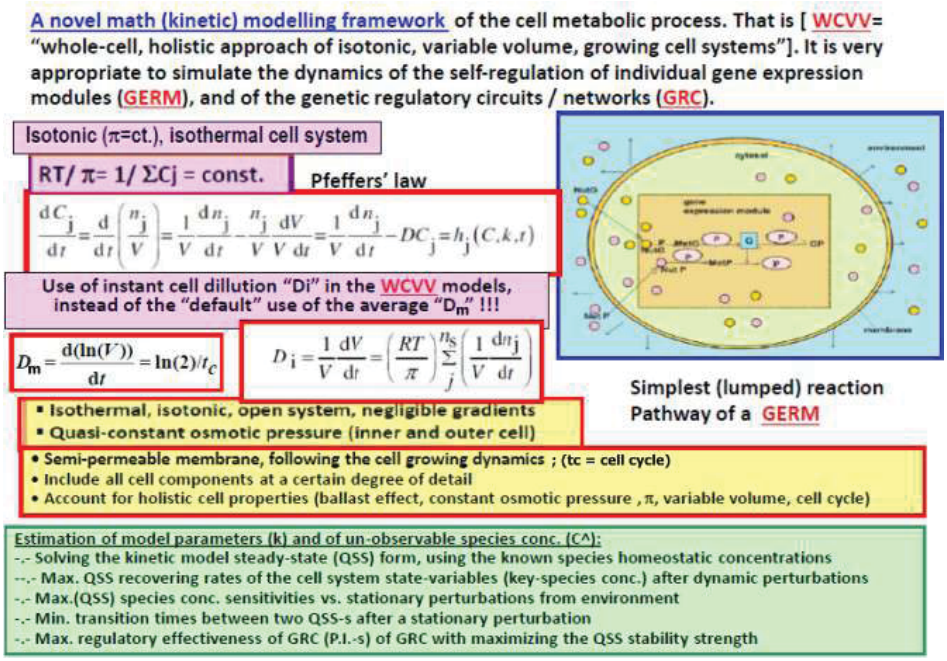


Figure 1-3: Some essential elements of the WCVV modelling framework. Notations:
Cj = species "j" concentration; nj = species "j" number of moles; V= cell (cytosol) volume; R = universal gas constant; T = absolute temperature; π = osmotic pressure; t = time; Di = instant cell dilution rate; Dm = average cell dilution rate (usually Dm = Ds = cell dilution rate under quasi steady-state (homeostatic) conditions).

- Dr Maria proved that the classical "default" WCCV formulation currently used in the literature for the cell processes kinetic models leads to **wrong, distorted and false predictions** concerning the cell (GRC) regulatory performances, i.e. their response to external/internal perturbations [Maria et al., 2017, 2018d]. See the proof in section 2.3.1.

The novel "mechanistic cell concept and WCVV cell modelling framework" proposed by Maria [2002,2003,2005,2006,2007,2009,2014b,2017], and by Maria et al. [2002] present several properties, as quantitatively defined and proved by Maria [2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c], and by Maria et al. [2017]; Maria and Scoban [2017, 2018]. Among them, are to be mentioned the followings:

1) **A novel math (kinetic) modelling concept and framework**. That is, the novel "mechanistic cell concept" (Table 2-1 of section 2), materialized in a novel math (kinetic) modelling framework of cell processes, that is the so-called WCVV ("whole-cell, holistic approach of isotonic, variable volume, growing cell systems"], well described in section 2.2. The WCVV modelling framework allows simulation of the CCM and GERM regulatory properties targeted to maintain the intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. This novel WCVV modelling framework was proposed by Maria [2002,2003,2005,2006,2007,2009, 2014b], and by Maria et al.[2002], with the further developments/ studies of Maria [2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c], The WCVV is particularly very appropriate to simulate the dynamics of the self-regulation of individual gene expression modules (GERM), and of the genetic regulatory circuits / networks (GRC) (e.g. genetic switches, operon expression, genetic amplifiers, etc.), but also the dynamics of metabolic processes belonging to the central carbon metabolism (CCM) (section 2.4). This novel concept implicitely promotes a modular kinetic modelling approach, as proved by Maria [2002,2003,2005,2006,2007,2009,2014b], and by Maria et al.[2002], with the further developments/ studies of Maria [2017A,2017B,2018,2023, 2023a, 2024, 2024b, 2024c].

2) Dr Maria **proved that the classic ("default")** math modelling approach (WCCV, "constant-cell-volume, non-isotonic cell system") **is erroneous** compared to the novel WCVV, by leading to **distorted and wrong simulation results** and conclusions [Maria et al., 2018d]. In such a way, the novel WCVV kinetic modelling approach of cell metabolism replaces the CLASSICAL (default) "constant cell-volume whole-cell" (WCCV) modelling approach, thus correcting the distorted and false/wrong predictions of WCCV cell kinetic models largely used in the literature (chap.2.2) [Maria et al., 2018d].

3) One of the strong aspect of this novel modelling approach is the use of the "**instant cell dilution rate Di**" (section 2) in the species mass balances of the WCVV models instead of the classical/"default" average dilution Dm=Ln(2)/(tc), where tc = cell cycle [Maria et al., 2018d; Maria and Scoban, 2017,2018]. By contrast, the use of the average "Dm", or even the omission of the cell dilution rate in the species mass balance of the classical (default) WCCV kinetic models lead to inaccurate predictions of the cell metabolism dynamics.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

4) **Development of a math "LIBRARY" including "template" kinetic models** of the main GERM types (self-regulated individual gene expressions)(Figure 1-4-down). This library is useful to build-up dynamic models of GRCs (genetic regulatory circuits, useful for *in-silico* design of GMOs) [Maria, 2017A,2017B,2018,2023; 2023a, 2024, 2024b, 2024c].

Dr. Maria used of the novel WCVV kinetic modelling framework and the generic simple GERM kinetic models to quantitatively define the regulatory performance indices (P.I.-S) of GERMs. Based on these (P.I.-s) and on the mentioned GERM "LIBRARY" Maria [2005,2006,2007,2009,2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c] **proposed a "buiding-blocks" strategy and rules to connect GERMs in chains to build-up genetic regulatory circuits (GRC)** when designing GMOs of desirable characteristics. Some of these structures were experimentally validated by [Yang et al.,2003; Sewell et al., 2002; Maria, 2018, 2023].

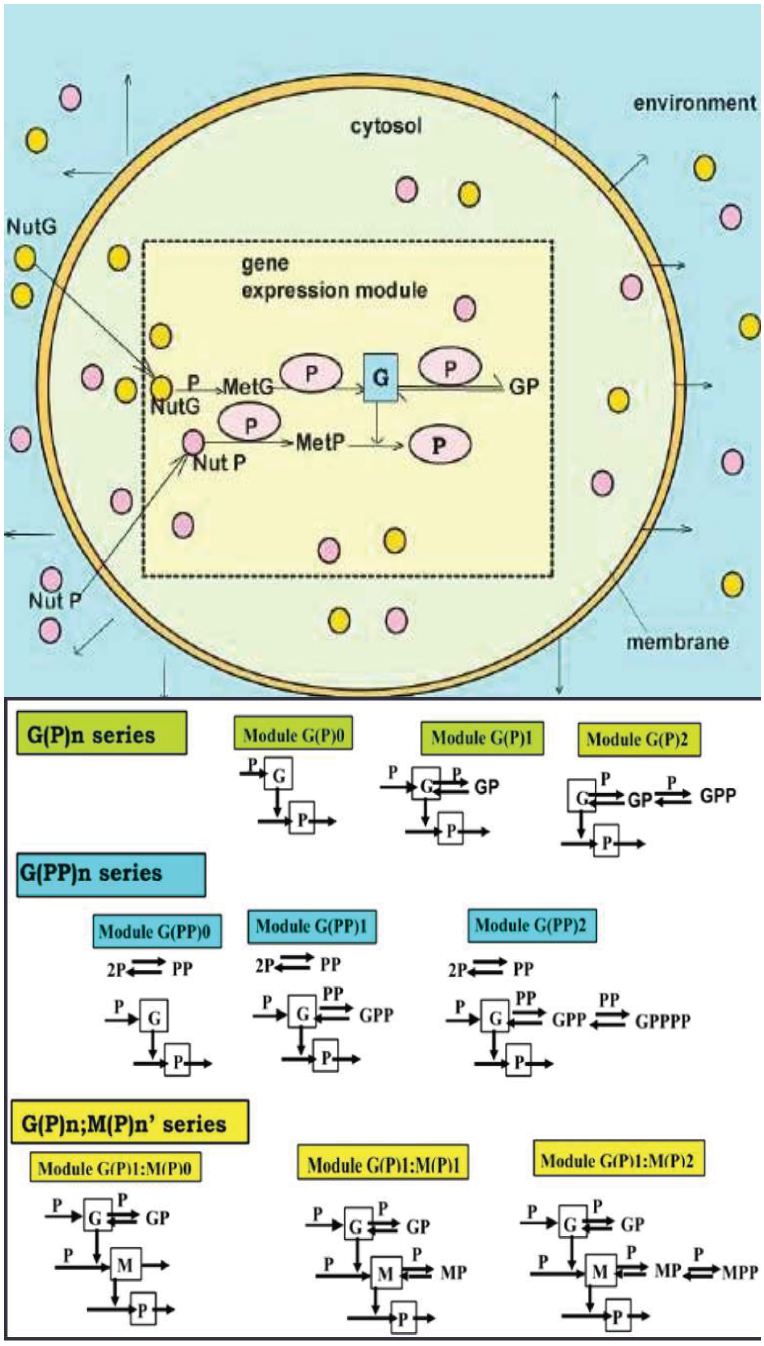


Figure 1-4: [TOP] The simplest reduced schemes of a generic gene expression regulatory module GERM in the holistic approach of Maria [2017a,2017b,2018, 2023, 2023a, 2024, 2024b, 2024c]. [DOWN] The LIBRARY of lumped gene expression regulatory modules GERM–s was developed by Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018, 2023, 2023a, 2024, 2024b, 2024c] in the form of template math (kinetic) models), to be further used to construct genetic regulatory circuits (GRC) in various cells. Notations: G = generic gene (DNA); P = generic protein; M = mRNA; GP, GPP, GPPPP = inactive forms of G; MP, MPP = inactive forms of M.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

5) **Experimental validation** of these GERM / GRC modular lumped pathway structures was done by using the dynamics of state-variables recorded in batch (BR) or fed-batch (FBR) bioreactors [Maria and Luta,2013; Maria et al.,2018a,2018c], or from using the literature data [e.g. Yang et al., 2003, or bio-omics databanks KEGG, Ecocyc]. The rate constants of the GERM / GRC modular kinetic models were estimated from solving the steady-state (QSS) model equations with using the stationary (homeostatic) concentrations of the key-species [from bio-omics databanks], and by imposing some holistic regulatory properties, see [Maria, 2017A, 2017B, 2018, 2023; 2023a, 2024, 2024b, 2024c], and [Maria and Scoban, 2017,2018; Maria et al., 2018d].

6) **Introduction of novel quantitative performance indices** (P.I.) to characterize the regulation efficiency of individual GERMs related to external perturbations (dynamic, or stationary) [Maria, 2017A,2017B,2018,2023; 2023a, 2024, 2024b, 2024c; Maria and Scoban, 2017,2018], and [Maria et al., 2018d; Maria,2003,2005,2006,2007]. The P.I.-s of the developed applicative GERM-s, or of GRC-s in the above cited references were studied by using the novel WCVV modelling framework (section 2.3).

7) **Develop and use** of a GERM -LIBRARY to study and to prove various properties of cell GRC self-regulation effects, such as (section 2.4) [Maria,2003,2005,2006,2007,2009,2014b,2017; Maria and Scoban, 2017,2018; Maria et al., 2017,2018d, 2019; Maria,2020c; Maria and Luta,2013]

- the role of the high cell-"ballast" in "smoothing" the effect of internal/external dynamic/stationary perturbations of the cell system homeostasis;
- the secondary perturbations transmitted inside the whole cell via the cell volume;
- the system isotonicity constraint reveals that every inner primary perturbation in a keyspecies level (following a perturbation from the environment) is followed by a secondary one transmitted to the whole-cell via cell volume;
- allows comparing the regulatory efficiency of various types of GERM-s;
- allows a more realistic evaluation of GERM regulatory performance indices;
- allows studying the recovering/transient intervals between steady-states (homeostasis) after stationary perturbations;
- allows studying the species connectivity;
- allows studying the conditions when the system homeostasis intrinsic stability is lost;
- allows studying the cell / GERM-s self-regulatory properties after a dynamic/stationary perturbation, etc.
- allows studying the plasmid-level effects in cloned cells.

8) **The use of the novel WCVV** math modelling framework to obtain *structured, modular, deterministic, hybrid* kinetic models (SMDHKM), linking the macroscopic state-variables of the bioreactor to the cell (nano-scopic) state-variables belonging to the CCM, thus offering more precise and more detailed (no. of state variables) predictions of the bioreactor/biomass dynamics. Such complex kinetic models are used in engineering rvaluations for various purposes: 1) to simulate the essential parts of the central carbon metabolism (CCM), such as: glycolysis, TCA cycle, ATP recovery system, or various other cell systems, such as: the mercury operon expression; Tryptophan operon expression; succinate synthesis, etc., aiming to *in-silico* design of GMO-s; 2) to simulate and/or design various genetic regulatory circuits (GRC); examples include: genetic switches (GS - biosensors), genetic amplifiers, operons expression. (e.g. GS of [Maria,2007,2009,2014,2018,2023]). 3) to *in-silico*, off-line derive optimal operating policies of bioreactors, with a higher precision and degree of detail. [Maria, 2017A, 2017B, 2018, 2023; 2023a, 2024, 2024b, 2024c].

9) **Applications** of concepts, principles, and numerical rules of the chemical and biochemical engineering (ChBPE, section 1.3), and of the nonlinear systems theory for solving Bioinformatics and Systems Biology problems (that is CCM, GRC model based design of GMOs). It is worth noting that the above essential contributions of dr. Maria in developing and promoting the novel "**mechanistic silicon cell concept**" (see section 2), materialized in developing of a the novel math (kinetic) modelling framework of cell processes, that is the WCVV , proposed by Maria [2002,2003,2005,2006,2007], and by Maria et al.[2002], with the further developments of Maria [2017A,2017B,2018,2023; 2023a, 2024, 2024b, 2024c], [Maria, 2017A,2017B,2018,2023; 2023a, 2024, 2024b, 2024c; Maria and Scoban,2017,2018]. The novel WCVV of Maria [2002,2003,2005,2006,2007], with the further developments of Maria [2017A,2017B,2018,2023; 2023a, 2024, 2024b, 2024c; Maria and Scoban,2017,2018](section 2, see the overview Table 2-1) is based on the application of ChBPE principles, concepts, and numerical rules (section 1.3)(see the overviews Figure 1-5, and Figure 1-6,7), of the math models' lumping rules [Maria,2005,2006,2005b,2019], very well connected to those of the "nonlinear systems theory", and to the "Systems biology", and "Bioinformatics" (Figure 1.-7) (see the books of Maria [2017A,2017B,2018,2023]).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

1.2. GMO importance in the biosynthesis industry

Over the last decades GMOs of superior characteristics reported increasingly important uses in a broad range of areas, such as: a) biosynthesis industry (optimized bio-syntheses, or bioprocesses using GMOs; or b) modified enzymes of superior features (high activity, selectivity, stability) produced by GMOs; c) production of vaccines (monoclonal antibodies - mAbs) in bioreactors by using modified hybridoma cell cultures [Maria,2020a](Figure 1-1), d) environmental engineering (improved biological treatment of wastewaters by using GMOs of an adapted sludge), e) medicine (therapy of diseases), f) design of new devices based on cell-cell communicators, g) design of biosensors using GMOs including design genetic switches [Maria,2018], h) novel and effective biocatalysts including GMO biomass, or enzymes from GMOs, immobilized on stable supports, etc.

Due to the modular organization of the cell [functions, regulatory networks, synthesis networks, gene expression regulatory modules (GERM) and of genetic regulation circuits (GRC), Figure 1-2], the *in-silico* (math model based) design of GMOs uses the

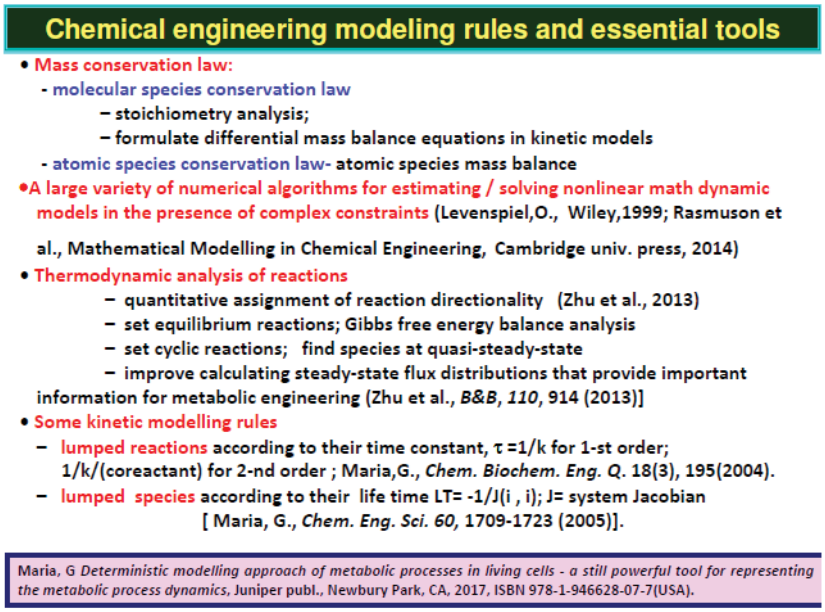


Figure 1-5: The chemical engineering main math modelling rules, and its essential tools [Maria, 2017, 2017a, 2017b, 2018, 2023].

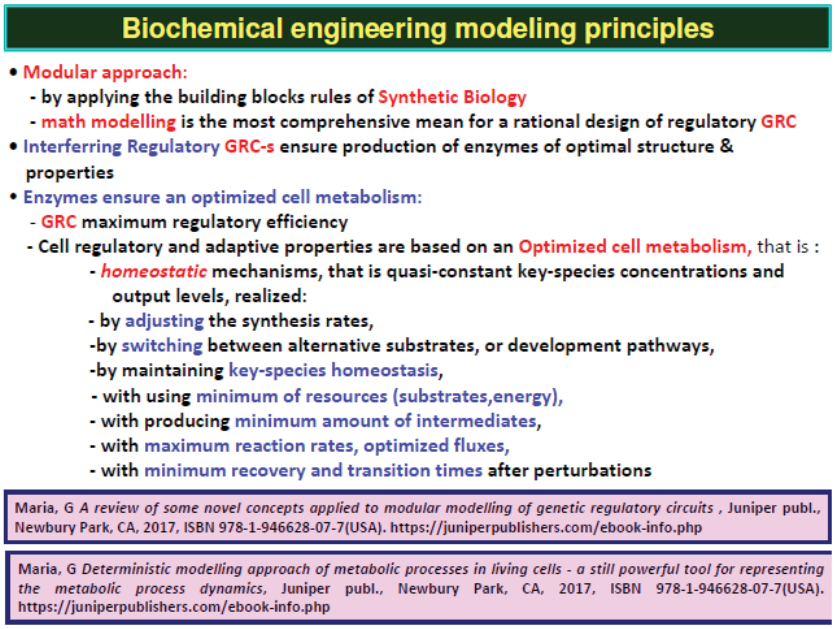


Figure 1-6: The biochemical engineering main math modelling rules, and its essential tools [Maria, 2017, 2017a, 2017b, 2018, 2023].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

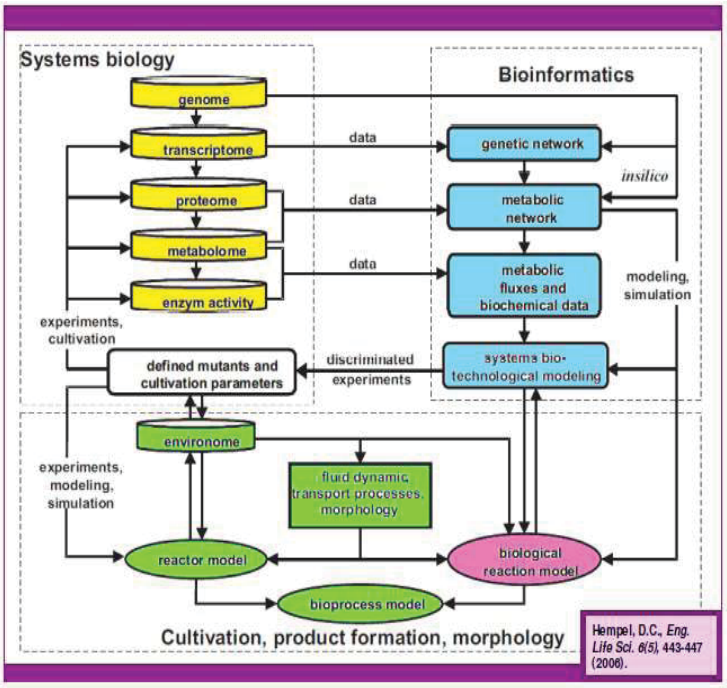


Figure 1-7: The strong connection and equivalence between "Systems biology", and "Bioinformatics" [Hempel, 2006].

same modular approach. Sauro and Kholodenko [2004] 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 builtup using the building blocks rules of the emergent Synthetic Biology are described by Heinemann and Panke [2006]; Qian et al. [2017]; Maria [2018,2023]. Among other advantages, the novel "**mechanistic silicon cell WCVV**" math modelling framework is particularly very suitable to simulate the regulatory circuits, like GERMs, GRCs, and to construct GRC-s from linking GERM-s in chains, as exemplified in this work.

The emergent **Synthetic Biology** [Benner and Sismour, 2005], "interpreted as the engineering-driven building of increasingly complex biological entities" [Heinemann and Panke, 2006; Qian et al.,2017], aims at applying engineering principles of systems design to biology with the idea to produce predictable and robust systems with novel functions in a broad area of applications [Voit, 2005; Heinemann and Panke, 2006; Qian et al.,2017] such as therapy of diseases (gene therapy), design of new biotechnological processes, new devices based on cell-cell communicators, biosensors, etc. By assembling functional parts of an existing cell, such as promoters, ribosome binding sites, coding sequences and terminators, protein domains, or by designing new GRC-s on a modular basis, it is possible to reconstitute an existing cell or to produce novel biological entities with new properties.

One exciting application of GMOs, due to its very high economic impact is those of improvement of industrial biosyntheses. Over the last decades, the general tendency is to replace the classical chemical synthesis (energy-intensive, and producing large quantities of hazardous waste), with enzymatic reactions [Liese et al., 2006]. Recent improvements in the synthetic biotechnology and production of modified enzymes using GMOs, exhibiting desired functions, allowed a considerable progress in industrial enzyme technologies and various other applications [Gavrilescu and Chisti, 2005]. Enzymatic reactions using modified (improved) enzymes, displaying a high selectivity and specificity, are attractive bioengineering routes to obtain a wide range of products in food, pharmaceutical, detergent, and textile industry, biochemical synthesis, or presenting challenging applications in medical tests, bio-sensor production, or emerging bio-renewable energy industries [Wang, 2009; Liese et al., 2006]. Isolated (free, suspended), or immobilized enzymes (on a suitable porous support) have been used in large-scale industrial reactors for more than sixty years, competing in terms of efficiency with the classical chemical synthesis pathways. Biocatalytic processes produce less by-products, consume less energy, and generate less environmental pollution, with using smaller catalyst concentrations and much moderate reaction conditions (Figure 1-8) [Moulijn et al., 2001]. New efforts are invested in the protein engineering (molecules/enzymes design), also implying various chemical and physical manipulations, coupled with bioactive nano-structure fabrication aiming at optimizing the carrier materials' structure to improve the enzyme/biomass stability and its catalytic efficiency. Such efforts are trying to overcome most of difficulties related to the industrial use of biocatalysts concerning the high costs of producing enough stable and long lifetime enzymes/biomass (cell culture), their high sensitivity to operating conditions and impurities, too high substrate specificity, and difficult process controllability due to their variable characteristics. Here it is worth noting that the product cost is also closely related to the production capacity, and market requirements (Figure 1-12).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Because the isolated/immobilized enzyme/biomass is usually the most expensive raw-material of the biochemical/biological process (Figure 1-8), especially when the biocatalyst is obtained from GMO organisms, the right choice and optimization of industrial enzymatic reactors (Figure 1-9) is a classical but a subject of still high interest. As enzymatic/biological processes are usually much slower than the chemical catalytic ones (Figure 1-8), the reactor suitable choice (continuous mixing *vs.* plug-flow), the enzyme utilization (free-enzyme *vs.* immobilized enzyme in gel/solid beads), and the operation mode (batch/semi-batch *vs.* continuous reactors) are the main problems to be solved by the engineers (Figure 1-9). Structured information on the enzyme/biomass culture characteristics (stability, activity, sensitivity to operating conditions and to products), on the immobilization possibilities, and on the process kinetics is essential when deciding on the process optimal scaling-up alternative. When a process model is available, application of the engineering computation rules, extensively described in section 1.3 and by Maria [2012] are usually used to evaluate the reactor efficiency and suggest its operation optimization alternatives.

To optimize the design and operation of enzymatic/biological reactors, Maria [2012] developed a modular simulation platform to simulate and compare the performances of various enzymatic/biological reactors, aiming at determining the optimal constructive and operation mode for a given enzymatic/biological process of known kinetic characteristics, and known kinetic model (Figure 1-10, Figure 1-11). By using ideal simulation models for the main reactor types [Moser,1988; Dutta, 2008; Froment and Bischoff, 1990; Maria,2020a], that is:

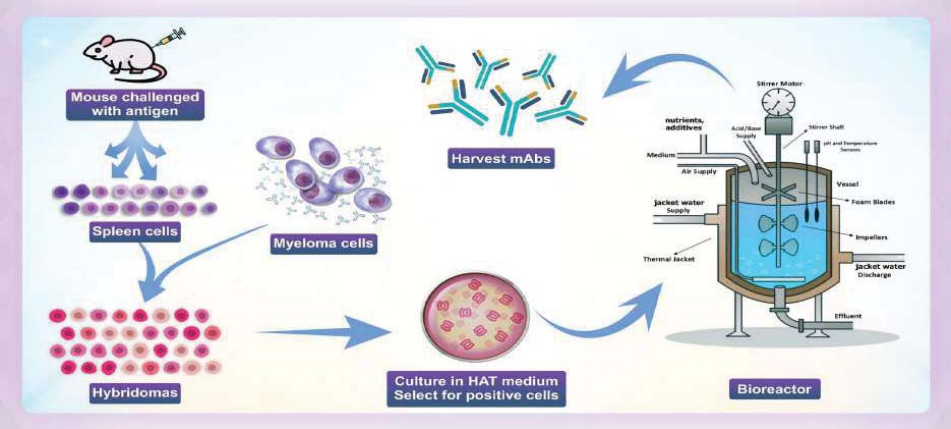


Figure 1-1: A simplified representation of the hybridoma technology used to produce monoclonal antibodies (mAbs)/vaccines [adapted from wikipedia, 2023]. [Right] Oversimplified "scheme of a BR or a FBR used to conduct biological processes. In the BR operating mode, substrate(s), biocatalyst, and additives are only initially loaded in recommended amounts (concentrations). In the FBR operating mode, the substrate(s)/biocatalyst, and additives (nutrients, pH-control substances, etc.) are continuously fed, following a certain (optimal) policy, to be determined"[Maria, 2020a]. The FBR presents a similar construction to a BR, and similar modelling hypotheses. However, unlike BR, the reactants and/or biocatalyst are added during the batch, following a time-step-wise variable (optimal) policy, to be determined offline [Maria,2020a], or on-line [Loeblein et al., 1997].

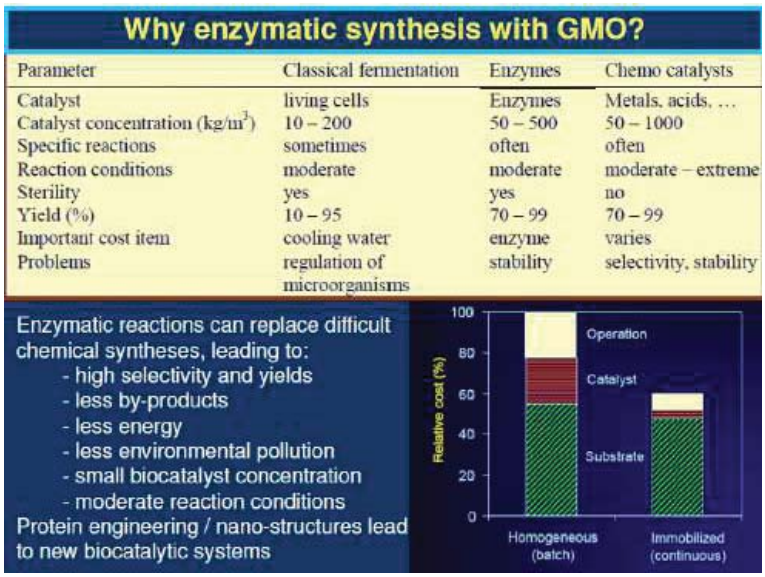


Figure 1-8: Advantages of industrial enzymatic, and biological processes vs. classical chemical catalysis (synthesis). [Bottom Right] Production cost structure in the case of biosyntheses with free enzyme (or free biomass), compared to those with immobilized enzyme (or biomass). Adapted from [Moulijn et al., 2001; Maria, 2018].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

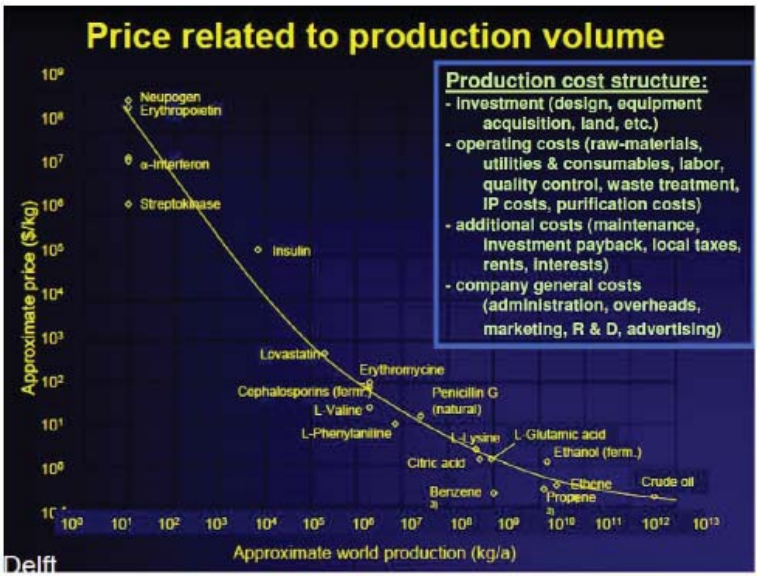


Figure 1-12: Cost-to-world production relationship. [Up-right]. The production cost structure for several industrial products. Adapted from [Moulijn et al., 2001].

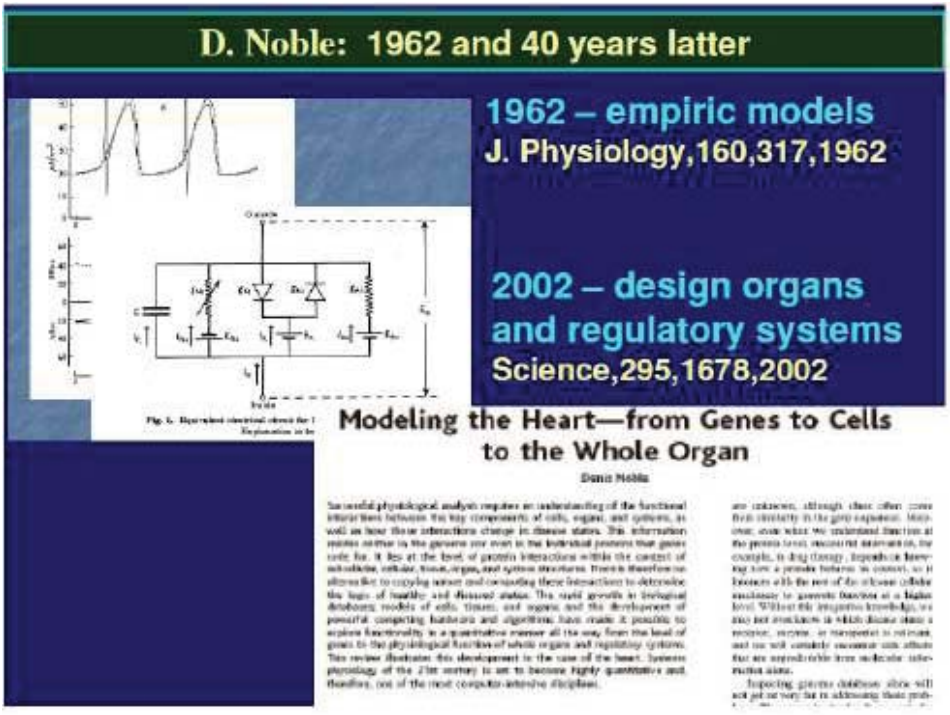


Figure 1-19: Evolution of Systems Biology from the empirical cell models [Noble, 1962] to modelling whole organs [Noble, 2002].

BR - batch reactor with initial addition of the biocatalyst, substrate(s), to be determined by applying an optimization procedure;

BRP - is a BR with an intermittent addition of enzyme or substrate(s), with a frequency, and in quantities to be determined by applying an optimization procedure;

SBR (semi-batch), or FBR ("fed-batch", for biological processes) reactor, with a constant or a variable feed flow rate of substrate(s)/biocatalyst;

MA(S)CR - mechanically agitated (semi)continuous reactor with free (suspended), or immobilized biocatalyst on porous particles;

FXBR - fixed-bed continuous reactor with the biocatalyst immobilized on porous support packed in fixed-bed columns (Figure 1-11).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

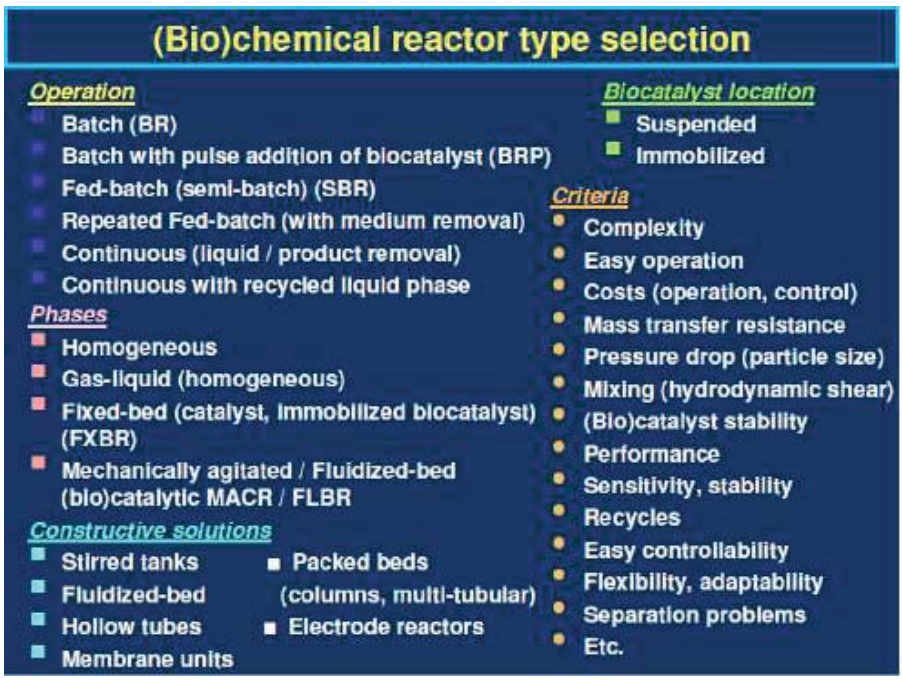


Figure 1-9: Criteria to choose the enzymatic reactor after [Maria, 2012].

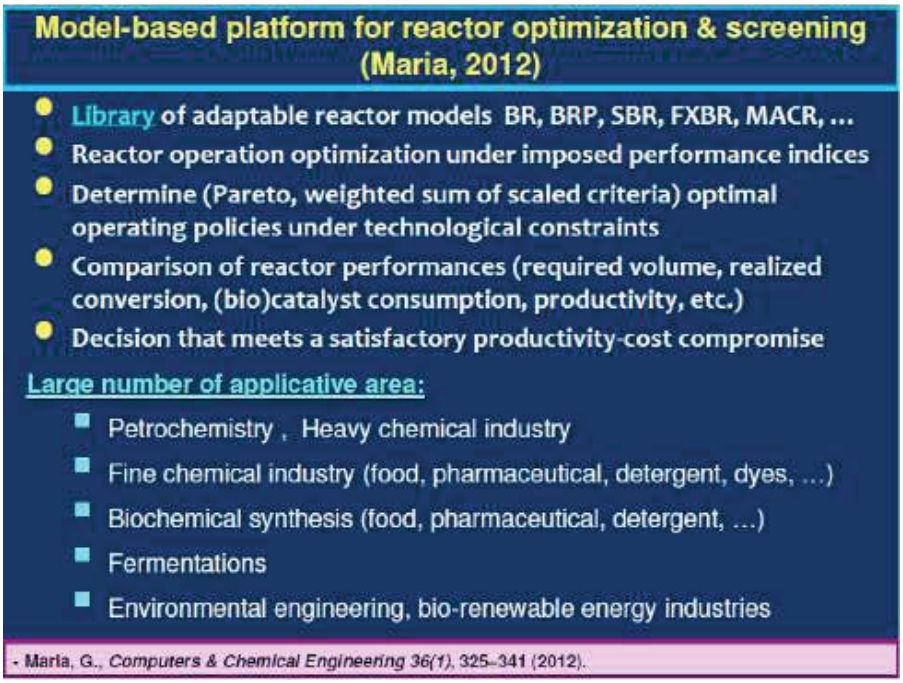


Figure 1-10: The principle of model-based evaluation of enzymatic/biological reactor performances. Adapted after [Maria, 2012].

The simulation platform of Maria [2012] allows considerable savings in the design effort, and reactor operation optimization, by minimizing the biocatalyst consumption without any loss in the target productivity.

When design a GMO for a certain application, a lot of costly experimental effort can be saved if an *in-silico* computational technique is used, by using a *structured, modular, deterministic, hybrid* kinetic model (SMDHKM) of the bioprocess (section 3).

Such a hybrid SMDHKM modular math model can link the cell metabolism (including the essential modules of the central carbon metabolism CCM, and the GRC responsible for the target metabolites' synthesis) to the reactor macro-scopic state-variables [Maria,2023]. As proved by Maria [2017b,2018,2023] such models are the most comprehensive mean for a rational design of GMOs

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

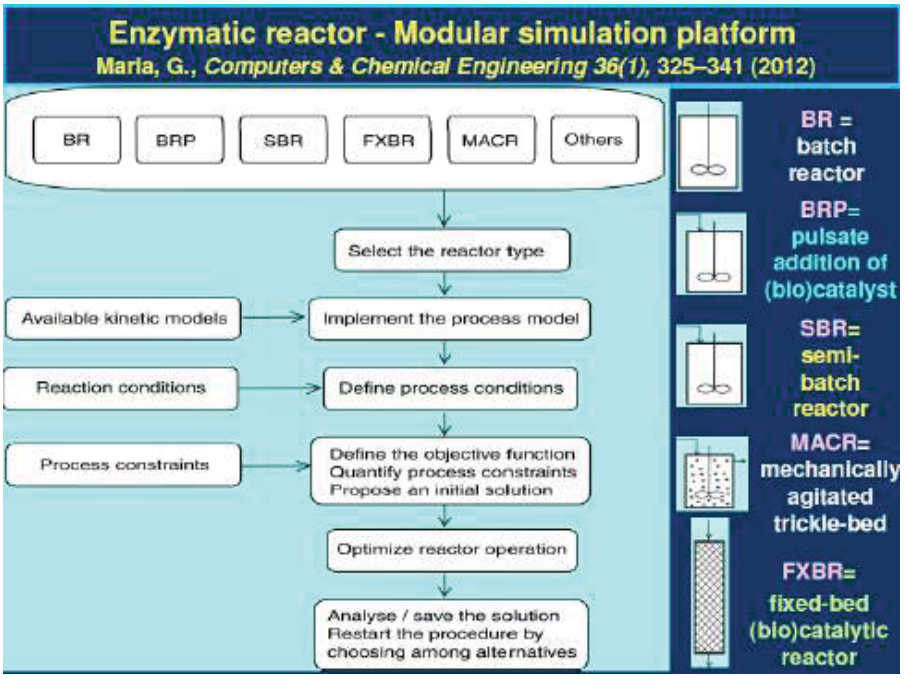


Figure 1-11: Optimal choice of enzymatic/biological reactors according to the simulation platform and rules of Maria [2012]. Some examples are given by Maria [2020a]

[Sotiropoulos and Kaznessis, 2007; Maria, 2018,2023]. 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) and of genetic regulatory circuits (GRC), enzymatic reaction kinetics, energy balance functions for ATP(ADP/AMP) renewable system, electron donor systems of the NADH, NADPH, FADH, FADH2 renewable components, 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.), functions for cell cycle regulation (nucleotide replication and partitioning, cell division). (Figure 1-2) [Maria,2017a, 2017b, 2018]. 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. [Kobayashi et al., 2004; Maria, 2017a,2018].

Thus, the emergent research area of **Systems Biology** dealing with *in-silico* design of GMOs has caught the contour and developed exponentially over the last 2-3 decades (as the number of Scopus citations proved it).

1.3. Why *in-silico* (model-based) design of GMOs ? Systems biology and Synthetic biology

The classical way used by the biologists to obtain GMOs is those to induce genome modifications in micro-organisms by using a mutagen vector [Puga and Wallace, 1998], then different strained are isolated and studied to separate and multiply the desirable one (an example in given in Figure 1-1). Such an experimental procedure is time consuming (years) and very costly. By contrast, if an adequate math (kinetic) model of the essential CCM modules of the cell is available, the *in-silico* route is faster and allows an easy investigation of a large number of GMO (virtual) alternatives [Maria, 2018; Maria et al., 2011; Maria, 2021]. One of the limitations of this *in-silico* route is related to the preliminary effort to acquire enough information (from databanks, and separate experiments), and to the very large computational effort to obtain and to validate an enough adequate SMDHKM math model of the target metabolic process. Such bioinformatics tools are often used to shorten the computational effort and to better manage the huge quantity of information from -omics databanks and, in particular, gene expression data [Das et al.,, 2010; Luscombe et al., 2001; Lopresti, 2010; Agbachi, 2017; Xiong, 2006137].

In short, the WCVV kinetic models of CCM, GERM-s, and GRC-s, eventually integrated and linked to the macroscopic state variables of a bioreactor dynamic model, in a so-called *hybrid* SMDHKM model (structured, modular, deterministic, hybrid kinetic model) can be used for various *in-silico* analyses, such as:

1). *In-silico* design of genetically modified micro-organisms (GMO) (section 3);

2). To characterize the GERM regulatory efficiency, and various local or global (cell level) regulatory properties, such as (details are given in section 2.2.3):

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

A. the role of the high cell-"ballast" in "smoothing" the effect of internal/external dynamic/stationary perturbations of the cell system homeostasis;

B. the secondary perturbations transmitted inside the whole cell via the cell volume;

C. the system isotonicity constraint reveals that every inner primary perturbation in a keyspecies level (following a perturbation from the environment) is followed by a secondary one transmitted to the whole-cell via cell volume;

D. allows comparing the regulatory efficiency of various types of GERM-s;

E. allows a more realistic evaluation of GERM regulatory performance indices;

F. allows studying the recovering/transient intervals between steady-states (homeostasis) after stationary perturbations;

G. allows studying the species connectivity;

H. allows studying the conditions when the system homeostasis intrinsic stability is lost;

I. allows studying the cell / GERM-s self-regulatory properties after a dynamic/stationary perturbation, etc.

J. allows studying the plasmid-level effects in cloned cells.

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

3). To build-up modular regulatory chains (GRC-s) of various complexity. An example is provided by [Maria and Luta,2013];

4). To prove the feasibility of the cooperative *vs.* concurrent construction of GRC-s that ensures an efficient gene expression, system homeostasis, proteic individual functions, and a balanced cell growth during the cell cycle;

5). To prove that the classic ("default") kinetic modelling approach (WCCV, "constant-cell-volume, non-isotonic system", section 2.1) is erroneous compared to the novel WCVV kinetic modelling framework, by leading to distorted and wrong simulation results and conclusions (section 2.3) [Maria et al., 2018d]. In such a way, the novel WCVV cell kinetic modelling approach replaces the wrong classical (default) WCCV modelling methodology, thus correcting the distorted and false/wrong predictions of WCCV cell kinetic models largely used in the literature;

6). Experimental validation of the GERM / GRC modular lumped WCVV kinetic models was done by using the recorded dynamics of state-variables in a batch or a fed-batch bioreactor [example of Maria and Luta,2013], or from using the literature data [e.g. Yang et al., 2003], or by using the bio-omics databanks [like KEGG,2011; Ecocyc,2005]. The rate constants of the GERM / GRC modular kinetic models were estimated from solving the steady-state (QSS) model equations with using the stationary (homeostatic) concentrations of the key-species [from bioomics databanks], and by imposing some holistic regulatory properties [Maria and Scoban, 2017,2018] (see the paragraph "Rate constant estimation in the WCVV kinetic models" in the section 2.2.1).

7). The use of the novel WCVV math modelling framework to obtain modular, deterministic, structured, *hibrid* SMDHKM dynamic models, linking the macroscopic state-variables of the bioreactor to the cell (nanoscopic) state-variables belonging to the CCM, GRCs, thus offering a more precise and a more detailed (as no. of state variables) predictions of the bioreactor/biomass dynamics. Such complex kinetic models are used for *in-silico* engineering evaluations for various purposes, such those of section 3, and the following ones [Maria and Luta,2013; Maria,2021; Maria and Renea, 2021; Maria,2014b]:

A). to simulate the essential parts of the central carbon metabolism (CCM), such as: glycolysis, TCA cycle, ATP recovery system, under stationary, or oscillating conditions Maria, 2014a, 2021];

B). to simulate the dynamics of various cell systems, and to generate the cell metabolic flux analysis aiming to *in-silico* design of GMOs; examples include: the mercury operon expression Maria and Luta,2013]; the tryptophan operon expression [Maria and Renea, 2021; the succinate synthesis [Maria et al.,2011], etc.;

C). to simulate and/or design various genetic regulatory circuits (GRC). Examples include: genetic switches (biosensors, [Maria, 2014b]), genetic amplifiers (ex. Mercury-operon quick expression in E. coli, [Maria and Luta,2013; Maria,2023), etc.

D). to *in-silico*, off-line derive optimal operating policies of bioreactors, with a higher precision and level of detail [Maria and Luta,2013; Maria and Renea, 2021].

The present section makes a short review of the (bio)chemical engineering (ChBPE) principles and deterministic modelling rules used by the Systems Biology for modelling cellular metabolic processes aiming to *in-silico* design GMOs with desired characteristics,

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

useful in the biosynthesis industry. This involves application of the classical modelling techniques (mass balance, thermodynamic principles), algorithmic rules, and nonlinear system control theory. The metabolic pathway representation with continuous and/or stochastic variables remains the most adequate and preferred representation of the cell processes, the adaptable-size and structure of the lumped model depending on available information and the utilisation scope [Maria,2017b,2018].

Exemplifications includes dr. Maria's experience on improving several bioprocesses, that is: 1) *In-silico* design of a genetic switch in *E. coli* with the role of a biosensor, [Maria,2007,2009,2014b]; 2) *In-silico* design of a cloned *E.* coli with a maximized capacity of mercury uptake from wastewaters [Maria, 2009b, 2010; Maria et al.,2013; Maria and Luta, 2013; Scoban and Maria, 2016; Maria et al., 2017d]; 3) *In-silico* design of a genetic modified E. coli with a maximized capacity of succinate (SUCC) production [Maria et al., 2011]; 4) *In-silico* characterize the glycolytic oscillator in E. coli [Maria et al., 2014a, 2018b,2018e]; 5) *In-silico* determine and check a GMO of *E. coli* , together with the FBR bioreactor optimal operating conditions with suspended biomass which can lead to maximization of the tryptophan production [Maria et al., 2018a; Maria and Renea, 2021].

1.3.1. Systems Biology - pioneering works and preliminary considerations

Systems Biology defined as "the science of discovering, modelling, understanding and ultimately engineering at the molecular level the dynamic relationships between the biological molecules that define living organisms" (Leroy Hood, Inst. Systems Biology, Seattle [2017]) is one of the modern tools which uses advanced mathematical simulation models for *in-silico* design of GMOs that possess specific and desired functions and characteristics with various uses (gene therapy of diseases in medicine; industrial biosyntheses; production of vaccines; new devices based on cell-cell communicators, biosensors, etc.).

To confer new characteristics, properties, and desired functions to a living cell, the metabolic fluxes (i.e. the stationary reaction rates of the metabolic enzymatic reactions, under quasi steady state, that is at homeostatic, balaced growth conditions) must be changed. These metabolic fluxes are mainly determined by the environmental conditions (substrate/nutrients availability), but also by the characteristics of the enzymes catalyzing the metabolic reactions. Among the pioneers in modelling biological systems (Figure 1-13), the famous cryptanalyst Alan Turing published on 1952 the paper "The chemical basis of morphogenesis" [Turing, 1952], by proving that there is a (bio)chemical reaction basis of metabolic processes. Studies on this subject were amplified after the famous question formulated by the distinguished physicist Erwin Schroedinger in his famous lecture at Trinity College Dublin on 1943, "What is life?". The emergent field of such efforts is the so-called '*gene circuit engineering*' (GCE) and a large number of examples have been reported with *in-silico* re-creation of genetic regulatory circuits GRC-s conferring new properties/functions to the mutant cells (i.e. desired 'motifs' in response to external stimuli) [Maria,2017a; Myers, 2009]). By taking into account that the genes (generically denoted by G) from the cell genome and the cell proteins (generically denoted by P) form pairs (G/P), as long as the protein synthesis is based on the information owned by its encoding gene, it results that it exists a direct relationship gene - metabolic flux made via the expressed encoded enzyme (Figure 1-14). Any flux change must follow the subsequent steps:

i). to change a certain cellular flux, the properties of the concerned enzyme (protein - P) should be changed: it should be over-expressed, under-expressed, or even it needs to be removed;

ii). the previous step implies that the G-encoding gene expression must be changed accordingly;

iii). the previous steps involve: a) either removal of that gene from the genome ("gene knockout"), or b) its substitution with another gene (plasmids) from another micro-organism, or c) gene copynumbers amplification (by cloning with target

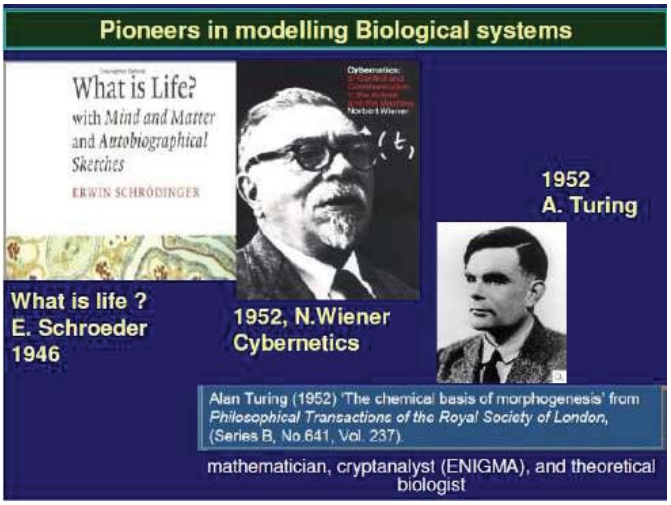


Figure 1-13: Pioneers in modelling biological systems [Maria, 2018].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

plasmids), or d) inhibition of the target gene expression, to decrease the encoded enzyme concentration, and thus to better control the target reaction. The Figure 1-14 exemplifies in brief the control of the tryptophan (TRP) synthesis, through the control of the TRP-synthase expression. The (Figure 1-15), and Figure 1-16) indicate the main techniques to be applied to obtain GMOs, and some of the difficult problems to be solved when applying *in-silico* design of GMOs.

The last way is often chosen because characteristics of an enzyme is in fact a property of the whole cell metabolic system, as proved by the Metabolic Control Analysis (MCA)[Kacser and Burns, 1973], through the so-called "summation theorem" (i.e. the control steps and coefficients are "global" properties, being a "systemic property" of the metabolic systems, dependent on all of its elementary steps; see [MCA Web, 2004; Heinrich and Schuster, 1996].

The change of properties of only one enzyme from the CCM should be made with caution, because, as proved by the Metabolic Control Analysis (MCA), the characteristics of an enzyme is in fact a property of the whole cell metabolic system [Kacser and Burns, 1973], through the so-called "summation theorem". That is, the control steps and coefficients are "global" properties, being a "systemic property" of the metabolic system, being dependent on all of its elementary steps [MCA Web, 2004; Heinrich and Schuster, 1996].

Due to the large number of genes [of the order,0^(3-4)], and the very high complexity of the gene interactions (even if one gene interacts with no more than 23-25 other genes, according to Kobayashi et al. [2004]), the only rational approach of the GMO design problem, and of modelling the kinetics of complex GRC is to use the math modelling tools of bioinformatics, systems biology, (bio) chemical engineering, and nonlinear systems control theory [Maria, 2017, 2017a; Sotiropoulos and Kaznessis, 2007].

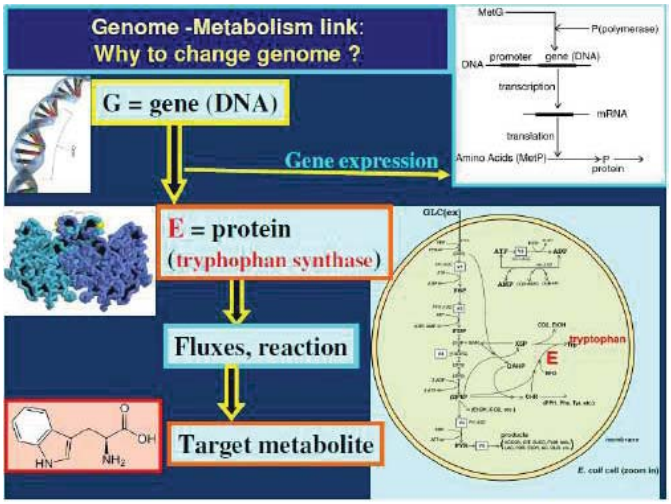


Figure 1-14: The direct link between genome and the metabolic fluxes of the cell [Noble, 2006; Alberts et al., 2002].

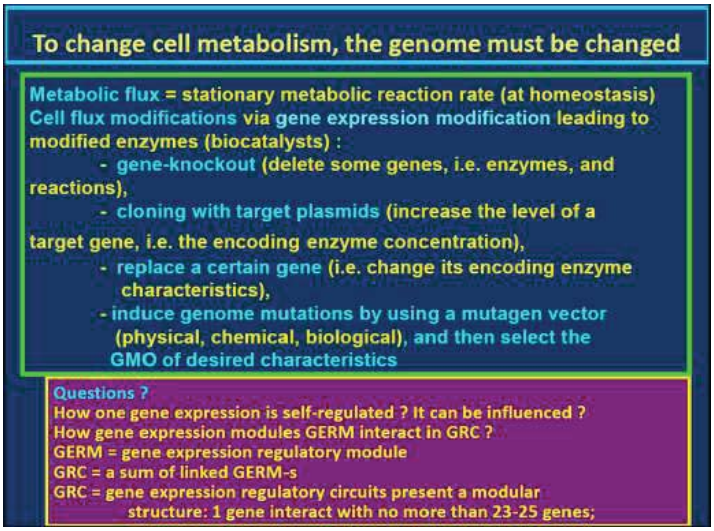


Figure 1-15: Some usual biological methods to alter the cell genome [Alberts et al., 2002].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

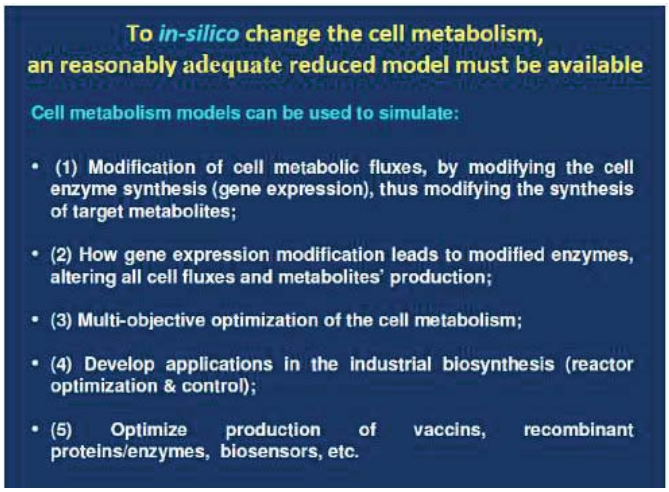


Figure 1-16: Problems to be solved when applying in-silico design of GMOs.

Additionally, due to the modular structure of GRCs [Maria, 2017a,2017b,2018,2005,2006,2007,2009], the modular GRCs dynamic models, of an adequate mathematical representation, seems to be the most comprehensive mean for a rational simulation and design of the regulatory GRC with a desired behavior [Sotiropoulos and Kaznessis, 2007].

The economic advantages offered by novel (modified) enzymes in the industrial biosyntheses are very important, that is (Figure 1-8): high selectivity and yields, less by-products, less consumed energy (processes occurring at low temperature and pressure), less environmental pollution, required small concentration of biocatalyst. Besides, a few numbers of enzymatic reactions can replace a large number of successive difficult chemical steps occurring under severe conditions.

All these, fully justify the quickly increase in the production of novel enzymes by protein engineering (using GMOs), and in the production of novel biocatalyst supports [Buchholz and Hempel, 2006; Hempel, 2006], and also the strong impulse given to the development of the so-called "Computational Systems Biology" [Kitano, 2002], based on advanced math models, and extended *-omics* databanks.

The connection principle between the gene expression, enzymes, and reactions, for one or several interfering enzymes is shown schematically in the Figure 1-58 [Mao et al. 2015]. Sometimes, interference among gene expressions is modeled by using simple linear (empirical) models [Zak et al.,2005; de Jong et al.,2003; Casey et al.,2006].

1.3.2. Systems Biology - a very short history

The *in-silico* GMOs design subject belongs to the **Systems Biology** topics. According to Leroy Hood [Inst. Systems Biology, Seattle, 2017], Systems Biology is defined as "the science of discovering, modelling, understanding and ultimately engineering at the molecular level the dynamic relationships between the biological molecules that define living organisms". Consequently, Systems Biology is a modern tool, which uses advanced mathematical simulation models for *in-silico* design of GMOs that possess specific and desired functions and characteristics. The metabolic pathway representation with continuous and/or stochastic variables remains the most adequate and preferred representation of the cell processes, the adaptable-size and structure of the lumped model depending on the available information and on the utilisation scope [Maria, 2017,2017a,2017b].

At this point, it is to underline that the living cells are evolutionary, auto-catalytic, self-adjustable structures able to convert raw materials from environment into additional copies of themselves (Figure 1-2)[Alberts et al.,2002; Westerhoff and Palsson, 2004]. Living cells are organized, self-replicating, evolvable, and responsive biological systems to environmental stimuli (see Figure 1-2), of a very high complexity (see an example in Figure 1-17).

The structural, functional, and temporal cell organization, including components and reactions (see Figure 1-2) is extremely complex, involving $0(10^{2-4})$ components, $0(10^{2-4})$ transcription factors (TF-s), activators, 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 [Maria, 2017a; Maria, 2017b]. Relationships between structure, function and regulation in complex cellular networks are better understood at a low (component) level rather than at the highest-level [Stelling et al., 2002; Maria, 2017a; Maria, 2017b].

In spite of the near astronomic complexity of cell metabolism, attempts to develop math models of the metabolic processes, at a molecular level, are quite old. The early works belong to [Schrödinger, 1944; Wiener, 1948](Figure 1-13). But, it is the famous cryptanalyst Alan Turing who clearly proved the direct relationships between the (bio)chemical metabolic reactions, the metabolic

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

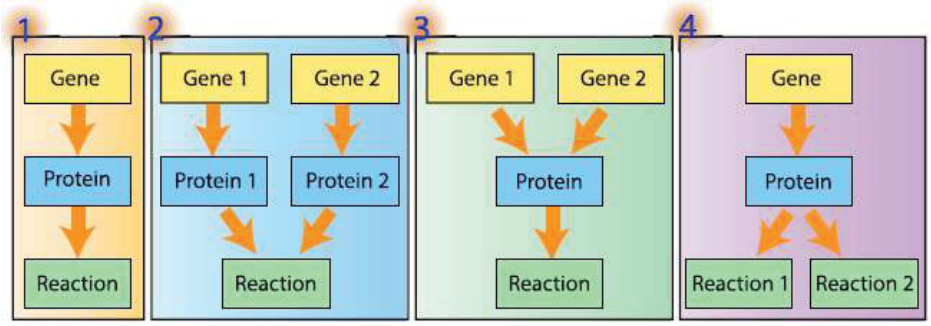


Figure 1-58: Examples of detailed gene–protein–reaction (GPR) associations. (1) Simple association, in which a single gene encodes a single enzyme. (2) Isozymes, in which multiple genes encode distinct proteins carrying out the same function. (3) Multimeric protein complex, wherein multiple genes encoding distinct protein subunits come together to form an active enzyme. (4) Multifunctional protein, in which a single protein can carry out multiple reactions. Adapted from Mao et al., [2015].

processes in living cells, and their morphogenesis [Turing, 1952]. When math modelling living cells, beside application of the ChBPE concepts and computational rules, there are to mention the applied concepts of nonlinear system control theory [Bertalanffy, 1933], especially in modelling GRCs [Heinrich and Schuster, 1996] (Figures 1-18,19). They introduced principles to describe systems with interacting components, applicable to biology, cybernetics, and other fields. Von Bertalanffy [1933] proposed that the classical laws of thermodynamics applied to the closed systems, to also be applied to the "open systems" such as living cells.

While the first cell modelling attempts have been very empiric, being proposed by the electronics engineers [Hodgkin and Huxley, 1952; Noble, 1962](Figure 1-20). Over 40 years, Noble published a review paper [Noble, 2002](Figure 1-21), pointing-out the tremendous advanced in the Systems Biology for *in-silico* design of GMOs, or even tissues, by means of computational systems biology [Kitano, 2002; Ideker et al., 2001].

Later, such 'electronic circuits-like' cell models have been extensively used to understand intermediate levels of gene expression regulation, but they failed to reproduce in detail molecular interactions with slow and continuous responses to perturbations, and their correlation to the metabolic process dynamics. So, eventually, they have been abandoned.

It is around 1968, when *Systems Biology* begins to catch the contour as a border area between (bio)chemical engineering, nonlinear system control theory, bioinformatics, cell / molecular biology, numerical calculus, etc., as pointed-out in the literature:

"The real advance in the application of systems theory to biology will come about only when the biologists start asking questions which are based on the system theoretic concepts rather than using these concepts to represent in still another way the phenomena which are already explained in terms of biophysical or biochemical principles. Then we will [...] have [...] a field of Systems Biology" [Mesarovic, 1968].

The first attempts to model the cell metabolism include modest topological models belonging to the so-called Metabolic Control Analysis (MCA), [Kacser and Burns, 1973; MCA Web, 2004; Heinrich and Schuster, 1996](Figure 1-18, and (Figures 1-22,23). Such structure-oriented analyses ignore the mechanistic details and the process kinetics, and they try in an unsuccessful way to use the network topology to quantitatively characterize to what extent the metabolic reactions determine the cell fluxes and the metabolic species concentrations [Heinrich and Schuster, 1996]. Briefly, the MCA is focus on using various types of sensitivity coefficients (the so-called 'response coefficients'), which are quantitative measures of how much a perturbation [influential variable x(j)] affects the cell-system states y(i) (e.g. r = reaction rates, J = fluxes, C = concentrations), by perturbing its "Quasi-Steady-State" (QSS, of index 's'), i.e. $S(y(i);x(j)) = [(\partial y(i) / y(i,s)) / (\partial x(j) / x(j,s))]$. The systemic response of fluxes or of concentrations *vs.* perturbation parameters (i.e. the 'control coefficients'), or of the reaction rates to perturbations (i.e. the 'elasticity coefficients') have to fulfil the so-called 'summation theorems', which reflect the network structural properties, and the 'connectivity theorems' related to the properties of single enzymes *vs.* the system behaviour.

Originally, MCA has been introduced by Kacser and Burns [1973]; Heinrich and Rapoport [1974]; Burns et al. [1985] to quantify the rate limitation in complex enzymatic systems. The MCA theory has been shortly followed by a large number of improvements, mainly dealing with the control analysis of the stationary states, by pointing-out the role of particular reactions and cell components in determining certain metabolic behaviour. Successive extensions of such definitions allow try to overcome the basic MCA drawbacks. Thus, the novel MCA developments allow: a) to study any limit set for non-steady / time dependent conditions; b) the flux balance analysis (FBA) and their optimization possibilities; c) elementary mode analysis (EMA). The elementary modes define the minimal realizable flow patterns through a biochemical network that can sustain a steady state. Their number, determined from the independent eigen-vectors of the system Jacobian matrix, is related to the system essential enzymes. d) dynamic flux balance analysis (DFBA); e) extreme pathway analysis (ExPA); f) constrained based modelling of metabolic network (CBM). See the review of

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

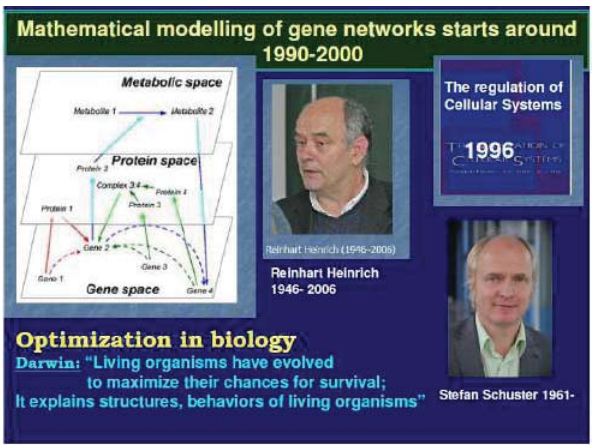


Figure 1-18: Pioneers in modelling genetic regulatory circuits [Heinrich and Schuster, 1996]

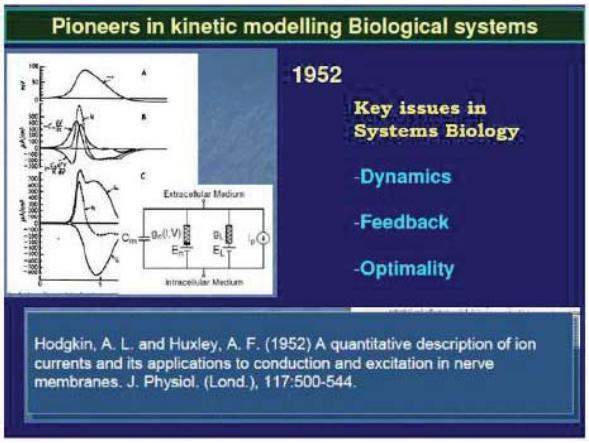


Figure 1-20: First attempts to model metabolic cell processes are inspired from the electric circuit's theory [Hodgkin and Huxley, 1952].

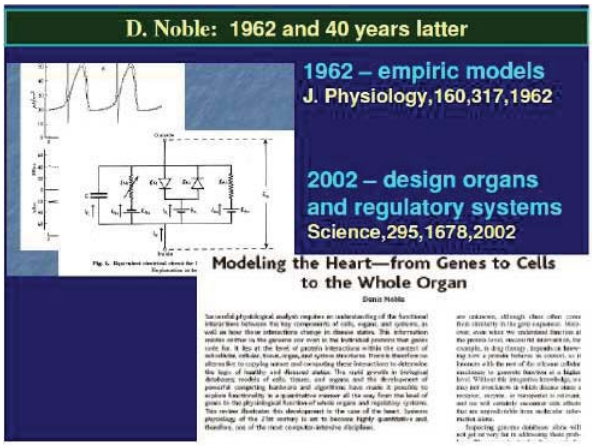


Figure 1-21: Evolution of Systems Biology from empirical cell models [Noble, 1962] to modelling whole animal organs [Noble, 2002].

Maria [2005,2017b]. MCA was very much applied in determining the metabolic fluxes (i.e. the metabolic stationary reaction rates), by applying (non)linear optimization procedures, with applications in metabolic engineering, and GMOs design [Stephanopoulos et al.,1998; Edwards and Palsson, 2000; Klein and Heinzle, 2012].

In spite of their limitations, MCA methods present some advantages when a rapid CCM analysis is required: A) it is able to efficiently characterize the metabolic network robustness and functionality, linked with the cell phenotype and gene regulation. B) MCA also allows a rapid but rough evaluation of the system response (especially the enzymes activity) to perturbations. C) Quickly

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

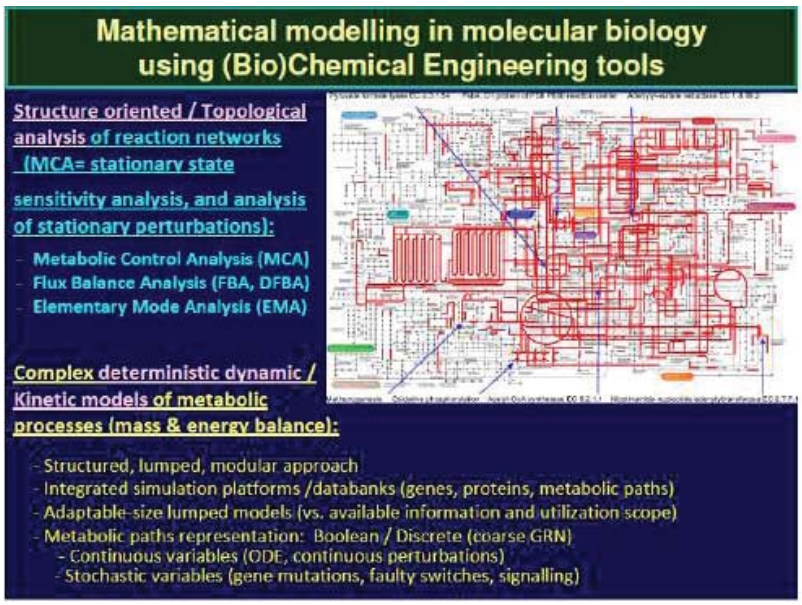


Figure 1-22: The main types of math models used in Systems Biology: topological, deterministic, Boolean, stochastic [Maria,2017a; Maria, 2017b; Maria, 2018].

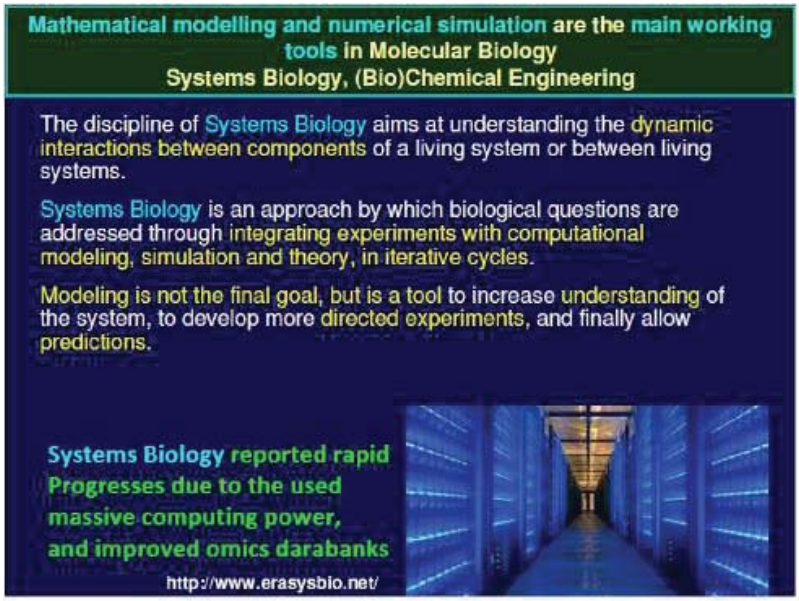


Figure 1-23: Math models used in Systems Biology requires massive computational efforts and facilities [Maria, 2017b, 2018].

determines the possibilities of control and of self-regulation for the whole analysed cellular pathway or of some CCM subunits. *Functional subunits* are metabolic subsystems, called 'modules', such as the amino acid or protein synthesis, protein degradation, mitochondria metabolic pathway, etc. [Kholodenko et al., 1998]. However, important limitations of MCA are related to i) the lack of any dynamic analysis of the metabolic system, ii) the system behaviour linearizations when computing the response coefficients; iii) omission of the influence of the cell variable volume and of the osmotic pressure; iv) neglect the regulatory modules dynamic response *vs.* stationary and/or dynamic perturbations. All these limitations make application of MCA to be quite limited.

Application of the adequate deterministic models when designing GMOs, seems to be more promising [Maria,2017b; Maria,2018,2021].

However, the deterministic models could have some limitations. First, they require extensive experimental and computational efforts to derive enough extended and adequate CCM kinetic models. Secondly, they encounter problems when reproducing the cell evolution. That is because the living cells are self-evolutive systems, caused by the environmental changes. Consequently: a) new reactions can be recruited by cells due to genetic mutations leading to novel produced enzymes; b) external inducers can activate genetic switches able to re-direct the cell metabolism toward more effective pathways [Maria,2007,2009,2014b,2018; Atkinson et

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

al.,2003; Hlavacek and Savageau, 1996,1997; Savageau, 2001,2002; Wall et al.,2003]. Eventually, the environmental changes can lead to an increase in the cell biological organisation and to optimal performance indices.

1.3.3. Systems Biology - modern concepts and tools

With the accumulation of experimental information, and its storage in -omics databanks, the math models and algorithms dedicated to cell process simulation, and "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 [Banga, 2008] (Figure 1-23), 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" [Leroy, 2017].

ii). "System Biology is a comprehensive quantitative analysis of the manner in which all the components of a biological system interact functionally over time" [Zak and Aderem, 2009].

iii). "Perhaps surprisingly, a concise definition of Systems Biology that most of us can agree upon has yet to emerge" [Ruedi Aebersold, Inst. Systems Biology, Seattle; http://www.systemsbiology.org).

"Systems biology, as defined by Ruedi Aebersold, 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."

iv). "The real advance in the application of systems theory to biology will come about only when the biologists start asking questions which are based on the system theoretic concepts rather than using these concepts to represent in still another way the phenomena which are already explained in terms of biophysical or biochemical principles. Then we will [...] have [...] a field of Systems Biology" [Mesarovic (1968)[72]).

v). "The discipline of systems biology aims at understanding the dynamic interaction between components of a living system or between living systems." (http://www.erasysbio.net/);

vi). "Systems biology is an approach by which biological questions are addressed through integrating experiments with computational modelling, simulation and theory, in iterative cycles." (http://www.erasysbio.net/);

vii). "Mathematical modelling is not the final goal, but it is a tool to increase understanding of the system, to develop more directed experiments and finally allow predictions." (http://www.erasysbio.net/)

The use of modular dynamic models, of an adequate mathematical representation, seems to be the most comprehensive mean for a rational design of the regulatory GRCs and of modified metabolic fluxes in a CCM of a GMO with desired characteristics [Sotiropoulos and Kaznessis, 2007; Maria, 2017a, 2017b].

Here it is worth notice that living cells are extraordinarily complex machines, with an astronomical complex organization (Figure 1-2), that exhibit remarkable behaviors including, most predominately, their selfreplication. 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 [Palsson, 2000; Tomita, 2001; Kitano, 2002,2002b; Mori, 2004; You, 2004]. "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 [Ideker et al., 2001; Jansen, 2003; Lindon et al., 2004; Rao et al.,2005]." [Surovstev et al.,2007]"

The increasingly use of math models and numerical calculus when design GMOs, has several reasons (Figure 1-24 up to Figure 1-29), as followings:

i). The tremendous increase of the computing power of the modern computers, allowing working with large and complex math cell models;

ii). Completion of various genome projects, by accumulating a huge amount of information about the cell processes. This allows starting the post-genomic era;

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

iii). A tremendous increase in (experimental) data, stored in structured and integrated *omics* databanks (genomics, proteomics, metabolomics, fluxomics, etc.), and advances in high-throughput experiments (Figure 1-24 up to Figure 1-29),

iv). An important experimental effort saving due to the use of *in-silico* numerical analyses (math model based).

Around 2000 the human genome has been deciphered (Figure 1-33), which has boosted Systems Biology development. Consequently, in the "post-genomic era" (after the "Human genome project, 1990-2003 [wikipedia,2025]) a large number of Systems Biology projects have been developed, and the number of dedicated publications increased near exponentially in Scopus (Figure 1-34). Among the structured cell simulators are to be mentioned some remarkable achievements (see also the reviews of [Maria, 2005, 2006]) (Figure 1-35). In parenthesis are given the main concepts (elements) of the software.

- E-Cell (compartments, compounds, genes, reactions), able to simulate the whole or parts of the cell metabolism) [Tomita et al., 1999] (Figure 1-36)

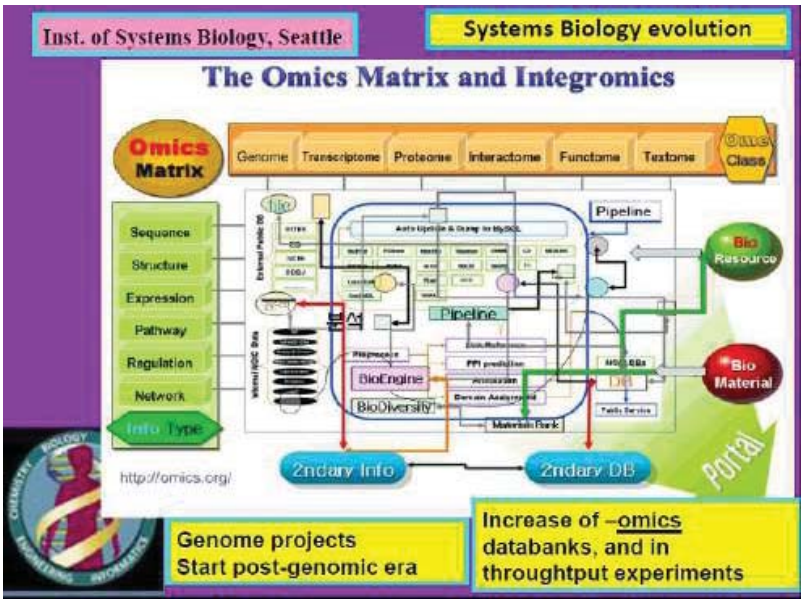


Figure 1-24: Steps in the evolution of the emergent Systems Biology [Maria,2017b,2018], and occurrence of "integromics" between –omics databanks [https://omics.org/File:Omics_matrix_and_integromics_20061103.gif].

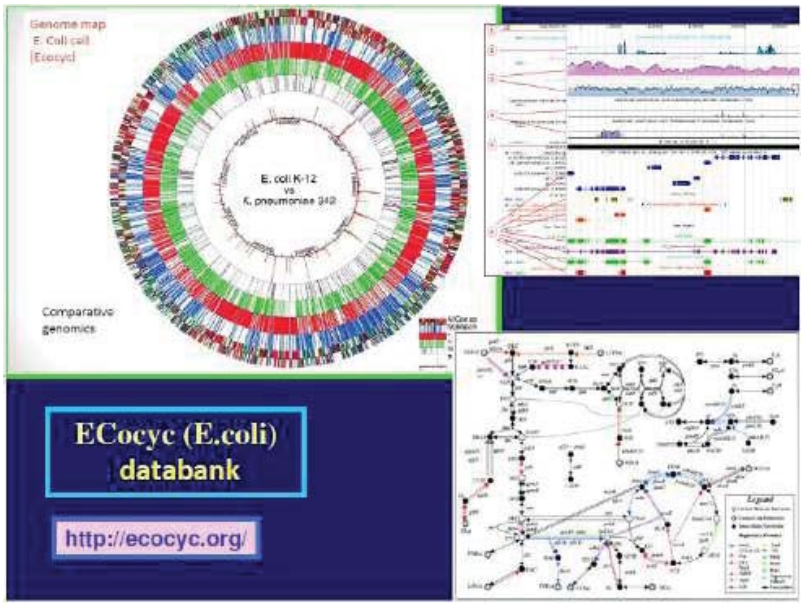


Figure 1-29: EcoCyc biochemical databank [EcoCyc, 2005].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

- V-Cell (model, geometry, applications, biological interface) [Schaff et al., 2001].
- M-Cell (stochastic simulator of some cell sub-systems) [Bartol and Stiles, 2002] (Figure 1-37).
- A-Cell (electrical circuit' like models) [Ichikawa, 2001] (Figure 1-37).
- Silicon-Cell (computer simulation of cell processes) [Westerhoff, 2001, 2006; Westerhoff and Palsson, 2004] (Figure 1-35, and Figure 1-36)
- Specific programming languages (example SBML), or on-line simulation platforms (JWS of Snoep and Olivier [2011]) (Figure 1-25,26, and Figure 1-35).
- Single cell growth, e.g. *Escherichia coli* [Domach et al. 1984; Shuler, 1989; Browning and Shuler, 2001], *Haemophilus influenzae* [Schilling et al., 2000a; Schilling et al., 2000b], *Mycoplasma genitalium* [Tomita et al., 1999; Tomita,2001; Kinoshita et al., 2001], *yeast* [Van Dien and Lidstrom, 2002; Kauffman et al., 2002], etc.

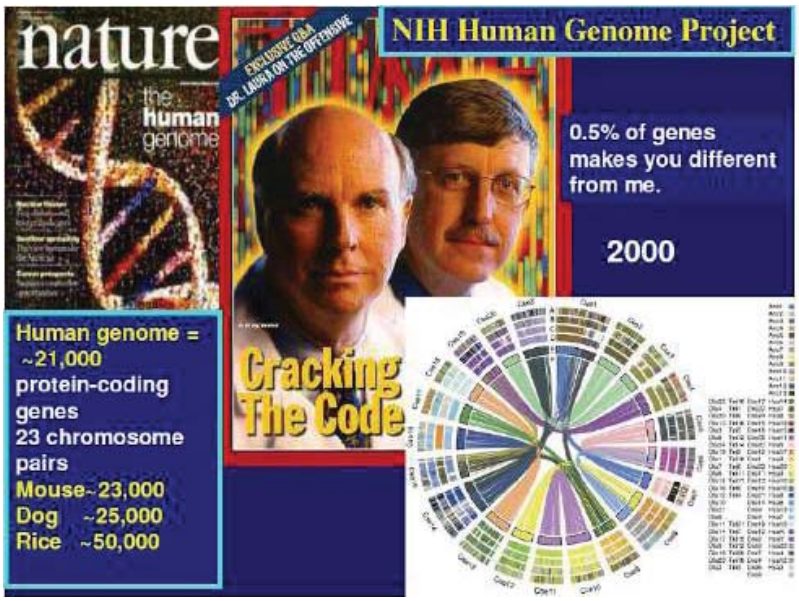


Figure 1-33: Around 2000 the human genome has been deciphered [Carbone and Tharoor, 2000].

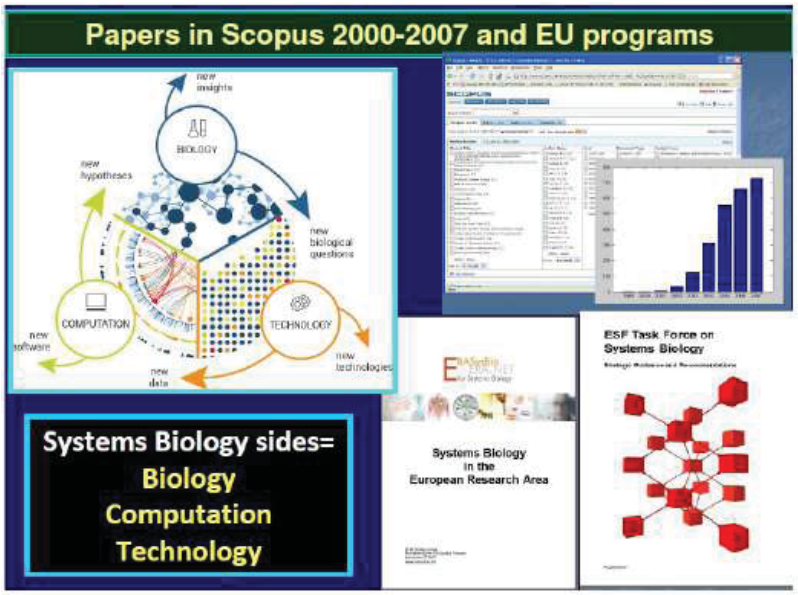


Figure 1-34: Exponential-like increase of publications (Scopus indexed), and of EU projects on Systems Biology after 2000.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

- **Modelling various metabolic oscillations**. Examples include: a) Metabolism of the human red blood cells [Kauffman et al.,2002; Wiback and Palsson, 2002]; b) glycolysis, TCA cycle, oxidative phosphorylation [Heinrich and Schuster, 1996; Riznichenko, 2002; Rensing et al.,2001; Wolf et al.,2000; Demin et al.,2001; Maria, 2014a] (Figure 1-39)

- **Modelling the iron metabolism** regulation under the WCVV framework [Hudder, Maria et al., 2002] (Figure 1-40);

- **Modelling the amino acid synthesis** [Stephanopoulos and Simpson, 1997; Maria,2023b; Maria et al., 2018a; Maria,2021; Maria and Renea, 2021];

- **Metabolic control of protein synthesis regulation** (GERM, GRC modelling), by using the novel WCVV framework [Maria, 2017A,2017B,2018,2023; Maria, 2023a, 2024, 2024b, 2024c], or by using the classical framework [Chen et al., 1999; Drengstig et al., 2012; Komasilovs et al., 2017; Matsuura et al., 2017; Zhang et al., 2017]

- **Modelling the CCM and glycolysis** [review of Maria,2014a; Kurata and Sugimoto, 2017; Chassagnole et al., 2002]; review of Miskovic et al. [2015] (Figure 1-38)

- **Topological FBA** based design of GMOs [Blass et al., 2017] (Figure 1-41-43) [Stephanopoulos et al. 1998; Schuetz, et al., 2007; Price et al. 2004]. The FBA is working with matrix math models, and linear relationships to reflect the cell system reaction network stoichiometry. Under dynamic conditions, dx/dt=S x v, where: S = the stoichiometric matrix; v = the fluxes vector; x = the species vector. However, solving a FBA problem to determine the cell system stationary fluxes (v) translates into a (non) linear programming problem (NLP) under a significant number of constraints [Stephanopoulos et al., 1998; Price et al.,2004]. Several solving alternatives have been reported in the literature.

- **Model regulatory networks** for gene expression by using Boolean 'biocircuit' models for E. coli [Shen-Orr et al. 2002; Mizuno, 1997], for various prokaryotes [McAdams and Shapiro, 1995; McAdams and Arkin, 1997]. Protein synthesis using various models [Maria,2003,2005,2006,2007,2009,2014b, 2017a,2017b,2018; Yang et al., 2003; Maria and Scoban,2017,2018; Sewell et al.,2002].

- **Cell cycles and oscillatory systems** in yeast and eukaryotes, such as: a) limit-cycle oscillator models [Norel and Agur, 1991; Hatzimanikatis et al.,1999; Obeyesekere et al.,1999; Surovstev et al., 2007; Surovtsev et al., 2008]; cell size control, oscillation properties and hysteresis effects [Tyson and Novak, 2001]; key ingredients inducing oscillations [Tyson and Novak, 2001; Hallett, 1989; Maria et al., 2018b; Goldbeter, 1996; Morton-Firth and Bray, 1998; Novak and Tyson, 1997; Rensing et al.,2001; Reijenga et al.,2002].

- **Modelling the drug release** and cell-drug interactions. An example is provided by [Zhang et al. 2003; Maria, 2005b], that is the successive release of drug-ligands groups from a dendrimeric multi-ligand support with disulfide bonds, in a reducing environment that can mediate the thiol-disulfide exchange, that mimics the human plasma environment.

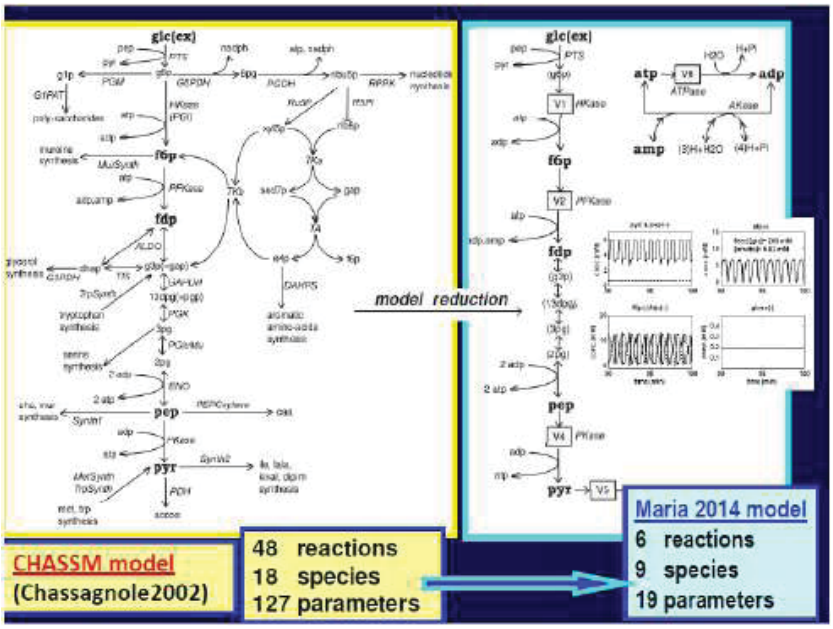


Figure 1-39: A comparison of the glycolysis model of [Chassagnole et al., 2002](left), and those of [Maria, 2014a] (right). Only the last one allows simulating glycolytic oscillations [Maria et al.,2018b].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

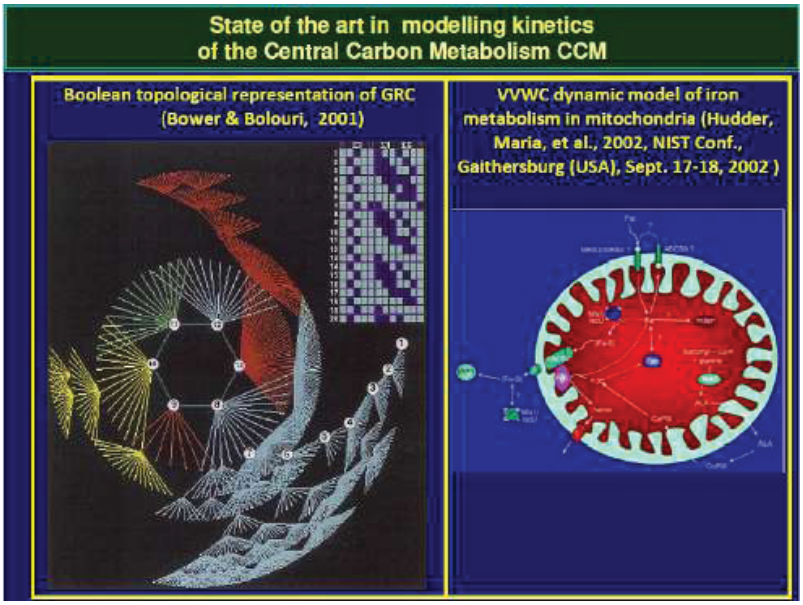


Figure 1-40: [LEFT] A cell simulator of Boolean type of CCM [Bower and Bolouri, 2001]. [RIGHT] Simulation of the iron metabolism in mitochondria under the novel WCVV modelling framework [Hudder, Maria, et al., 2002].

Some glycolysis models from literature

Reference	#Species	#Reactions	#Parameters
Selkov, 1968	5	5	?
TRM, Termonia & Ross, 1981	9	7	19
Maria, 2014	9	6	19
Hatzimanikatis & Bailey, 2006	6	9	?
Bier et al., 1996	7	9	15
Buchholz et al., 2002	3	5	24
CHASSM, Chassagnole et al. 2002	18	48	127
Westermark & Lansner 2003	6	6	?
Degenring et al. 2004	10	22	123
Costa et al., 2008	25	30	116
Kadir et al., 2010	24	30	> 150

Figure 1-38: A review of the kinetic models for glycolysis, with specification of the accounted number of species, reactions, and parameters [Maria, 2014a].

- **Cellular communications**, intracellular signalling, neuronal transmission, networks of nerve cells [Shimizu and Bray, 2002; Stiles et al. 1998; Ichikawa, 2001].
- **Analysis of the "logical essence" of life**, and the life fundamental requirements. [DeBeer and Kourie, 2000].

At the same time, the exponential-like increase of the experimental biological information leads to development of valuable *-omics* databanks, such as (Figure 1-25-29), known as KEGG [KEGG, 2011, KEGG pathway, 2011], JWS [Olivier and Snoep, 2004; Peters et al.,2017], EcoCyc [2005], Roche [Michal, 2014], etc.

However, it is only over the last decades when Systems Biology reported notable successes due to a considerable increase in the computing power of the modern computers. It is to mention here, for instance, the cell simulator platforms, and online model repository JWS of Olivier and Snoep, 2004; Peters et al.,2017], or those developed by Rocha et al. [2010], or by Tomita et al. [1999], together with the continuous expansion of the *-omics* databases (EcoCyc [2005]; KEGG [2011]), and reported advances in the numerical

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

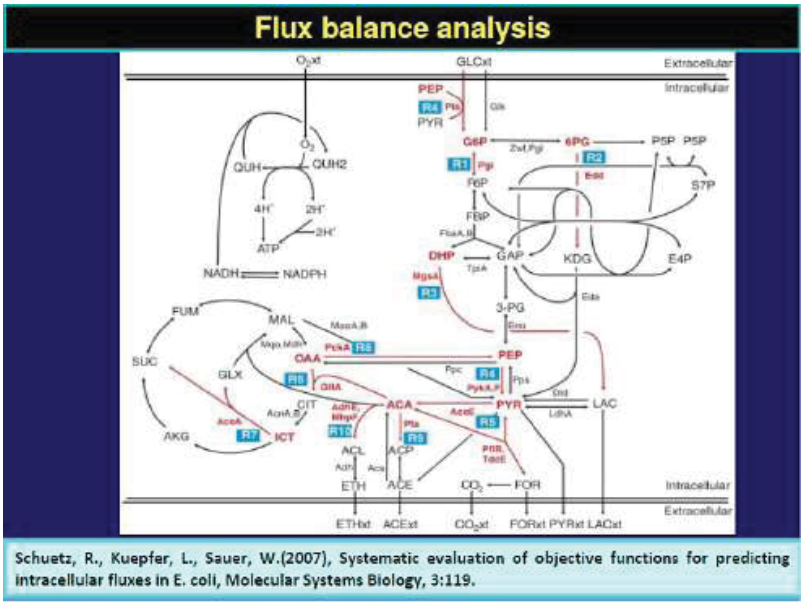


Figure 1-41: Flux balance analysis (FBA) is the main objective of metabolic engineering, and an essential preliminary step in design GMOs. [Schuetz, et al., 2007; Stephanopoulos et al.,1998].

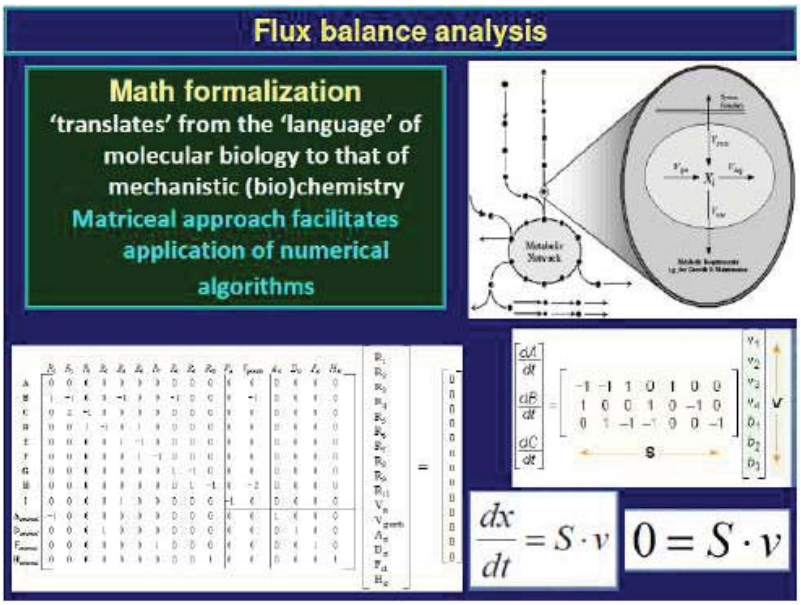


Figure 1-42: Flux balance analysis (FBA) is working with matrix math models [Stephanopoulos et al., 1998]. Notations: S = the stoichiometric matrix; v = the fluxes vector; x = the species vector.

algorithms used by bioinformatics, systems biology, (bio)chemical engineering, and by the nonlinear systems control theory [Maria, 2017a,2017b,2018].

Due to such favourable premises, related to the expansion of *-omics* databanks, and cell metabolism dynamic models, novel works have been reported over the last decades. Among the milestone works in Systems Biology it is to mention the contributions in modelling / design of GRC, GERM, FBA, MCA of [Heinrich and Schuster, 1996; Torres and Voit, 2002; Cornish-Bowden, 2016; Brazhnik et al., 2002; Savageau, 2002, etc.]. The number of published papers in the Systems Biology area increases with two orders of magnitude from 2000 to 2007, and it is still exponentially increasing, most of them being founded by programs of the European Science Foundation (Figure 1-34).

As stated by Nandy [2017], and Sutherland [2005], tremendous applications of Systems Biology have been reported over the next decades in the below areas (Figure 1-44).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

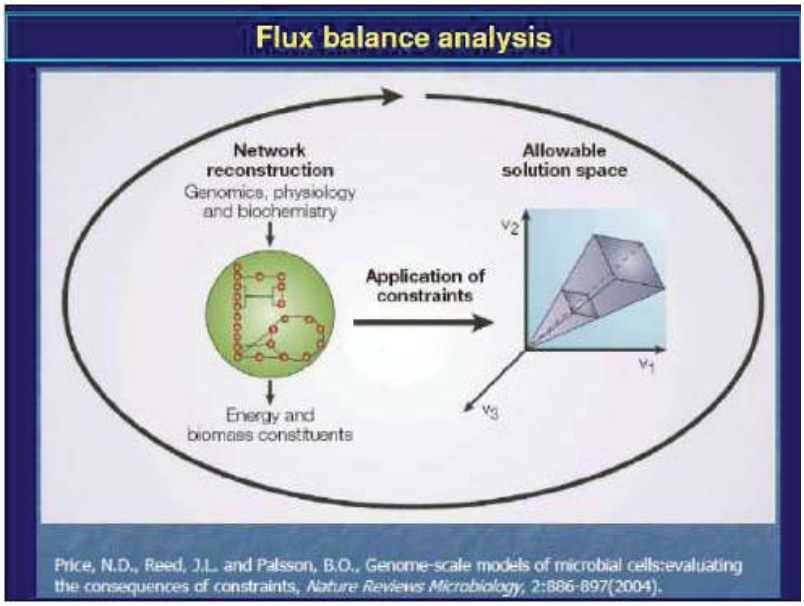


Figure 1-43: Solving a Flux balance analysis (FBA) problem translates into a (non)linear programming problem (NLP) [Stephanopoulos et al., 1998; Price et al.,2004].

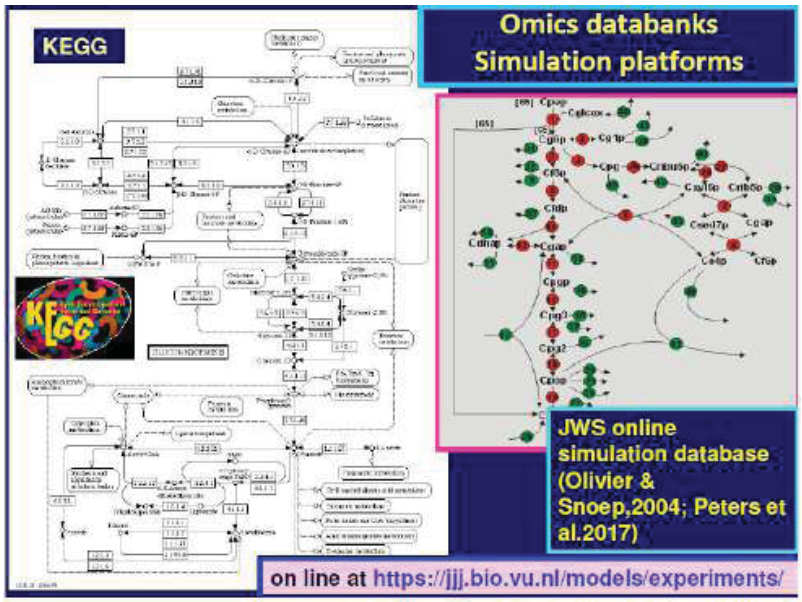


Figure 1-25: [Left]. KEGG –omics databank [KEGG pathway,2011; KEGG,2011]. [Right] the JWS on-line cell simulators [Olivier and Snoep, 2004; Peters et al.,2017].

Designing mutant, cloned cells with desired 'motifs'	Cell biology
Genetics biology or genetics	Food science
Biotechnology, Bioengineering	Immunology
Biomedical engineering	Molecular biology
Biochemistry	Biodiversity
Agricultural biology and Ecology and Biophysics	Bioinformatic

And, "in a context of increasing calls for biology to be predictive, *modelling and optimization are the only approaches* biology has for making predictions from first principles..." [Banga, 2008]. Due to the computing facilities offered by the algorithmic rules developed by the (bio)chemical engineering and nonlinear systems biology rules (Figure 1-45, Figure 1-22, Figure 1-23, Figure 1-5, Figure 1-6), the developed cell math models use a vectorial-matriceal approach (section 2.1)(Figures 1-45,46), with a continuous model upgrading based on dynamic experimental data recorded in a chemostat (i.e. a continuously operated bioreactor), operated

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

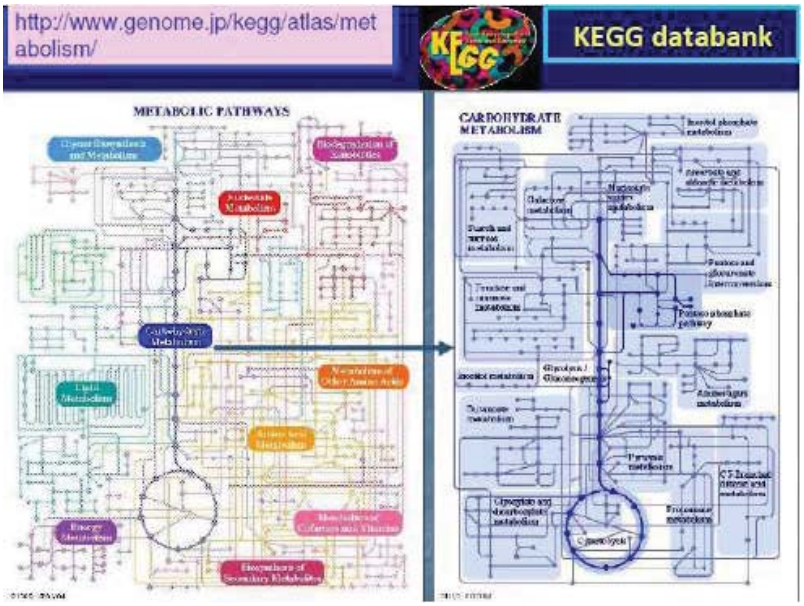


Figure 1-26: KEGG databank - pathways [KEGG pathway,2011; KEGG,2011].

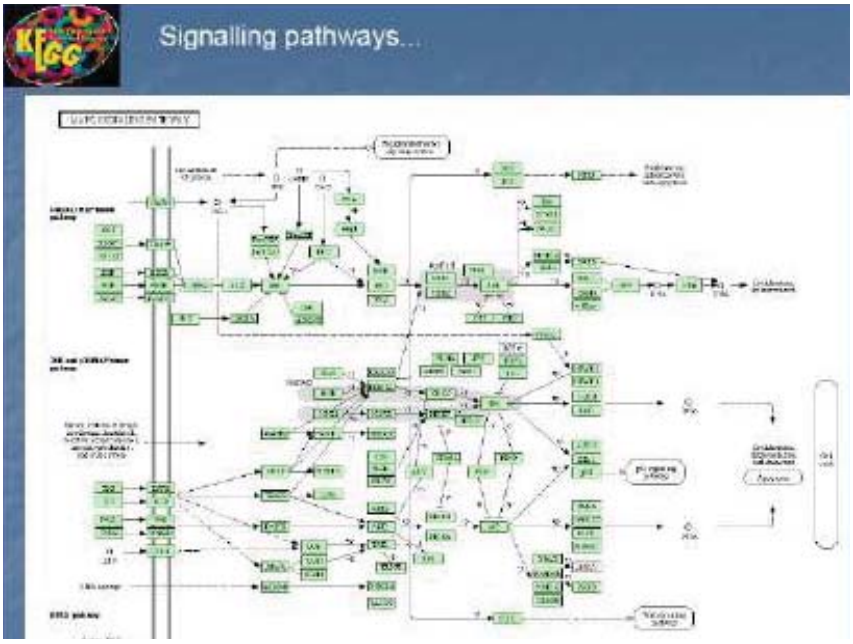


Figure 1-27: KEGG databank - fluxomics [KEGG pathway,2011; KEGG,2011]

under steady-state, or in a dynamic regime following an input perturbation in the substrate concentration in the solution fed in the bioreactor [Chassagnole et al. 2002; Blass et al., 2017](Figure 1-48).

So that, 40 years after the first reported empirical models of living cells [Noble, 1962], the optimistic researchers advanced very ambitious targets, such as "Modelling the heart - from genes to cells to the whole organ" [Noble, 2002] (Figure 1-21).

As in all scientific controversies, sceptic opinions also exist, such as (Figure 1-46,47):

i). "In spite of a full mapping of the human genome which yielded a code of three billion letters, we are still far from a satisfactory answer to the question formulated by the distinguished physicist Erwin Schroedinger in his famous lecture at Trinity College Dublin in 1943:

"What is life?". However, two important observations was made by the world renowned physiologist Noble [2006]: a) "we must move away from our obsession with genes alone. We must look not at one level, but at the interaction of processes at various levels,

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

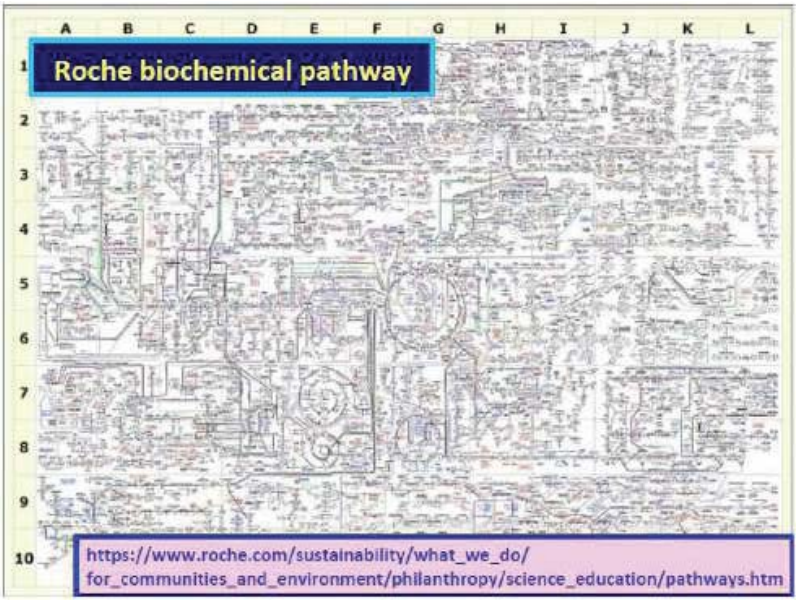


Figure 1-28: Roche biochemical databank [Michal, 2014].

from the real *Systems Biology*; b) The reductionist approach of molecular biology has proved itself immensely powerful. But DNA isn't life."

ii). These Systems Biology tools they really work? (http://www.systemsbiology.org/Systems_Biology_in_Depth). Some pessimists charge that systems biology is nothing more than "a fashion fad" that will pass once the hype dies down. Others maintain that systems biology is, in essence, a "repackaging of established concepts and methodologies" under a new description.

iii). And a third camp endorses the idea of systems biology as an enticing and powerful new discipline but thinks that it's "premature" to be considered.

1.3.4. Systems Biology - Application of chemical and biochemical engineering (ChBPE) concepts/rules when math modelling the metabolic processes

Most of GMOs used in the industrial synthesis are prokaryote cells (Figure 1-49). If one desires to change the cell metabolism, then the metabolic fluxes (i.e. the stationary enzymatic reaction rates occurring inside cell) must be changed, or even deleted (Figure 1-14, Figure 1-15). This can be done by changing the characteristics or even by removing certain enzymes catalyzing some reactions. As the cell enzymes are proteins expressed by their encoding genes, to change enzymes characteristics, it follows that genome has also to be changed, by one of the following alternatives (Figure 1-15) [Alberts et al., 2002]:

i) delete genes (i.e. "gene knock-out", Maria, 2011),

ii) clone the cell with target plasmids, thus modifying the expression level and the encoded enzyme concentration (Maria, 2010),

iii) replace a target gene with another (Myers, 2009),

iv) *in-silico* (model-based) investigate the bioreactor operating conditions leading to significant changes in the cell target fluxes (Maria et al.,2018a,c).

v) Induce genome mutations by using a mutagen vector. A mutagen vector is a physical (x-rays, ionising radiation, etc.), or chemical, or a biological (plasmids, viruses, bacteria) agent that permanently changes the genetic material, usually DNA, in an organism and thus increases the frequency of mutations. [wikipedia, 2025], followed by a process of selecting the cell culture of desired characteristics (an example is given in Figure 1-1).

Besides, it is worth noting that the environmental conditions exert a major influence on the cell fluxes. These inter-dependencies can be *in-silico* better investigated, by using SMDHKM kinetic models [Maria et al.,2018a; Maria et al.,2018e; Maria et al., 2018c].

All **genome / proteome modifications** can be much easily, and with less (experimental) cost investigated if an adequate (reasonably extended) dynamic model of the cell metabolism is available (Figure 1-50). Such a dynamic structured model (preferably of a SMDHKM type [Maria,2023]), can quickly predict the metabolic fluxes and their dependence on the external, or enzymatic factors

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

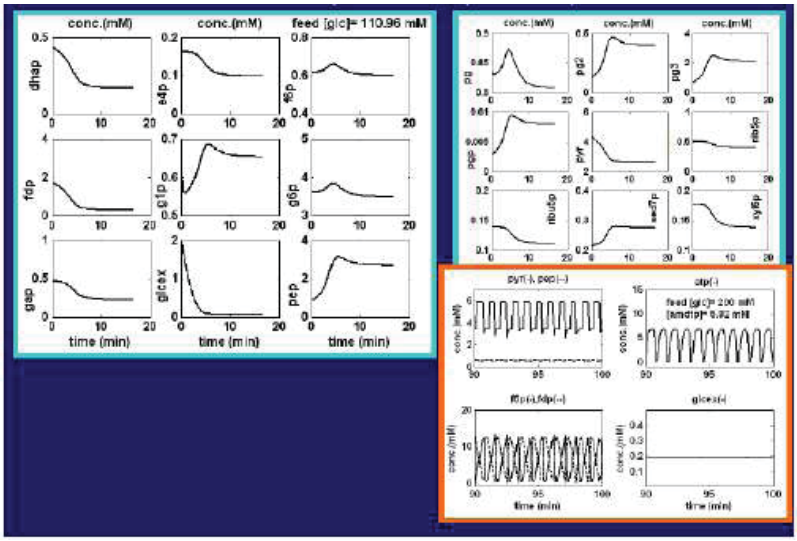


Figure 1-48: Generate structured kinetic data (cell species dynamics) by using chemostat experiments with perturbed inputs (GLC concentration here, denoted as "glcex" here). [Chassagnole et al. 2002; Blass et al., 2017]. Such experiments have been used by Maria [2014a] to build-up a structured kinetic model of the glycolysis, able to simulate the glycolytic oscillations[right-down], under suitable environmental conditions [Maria et al., 2018b].

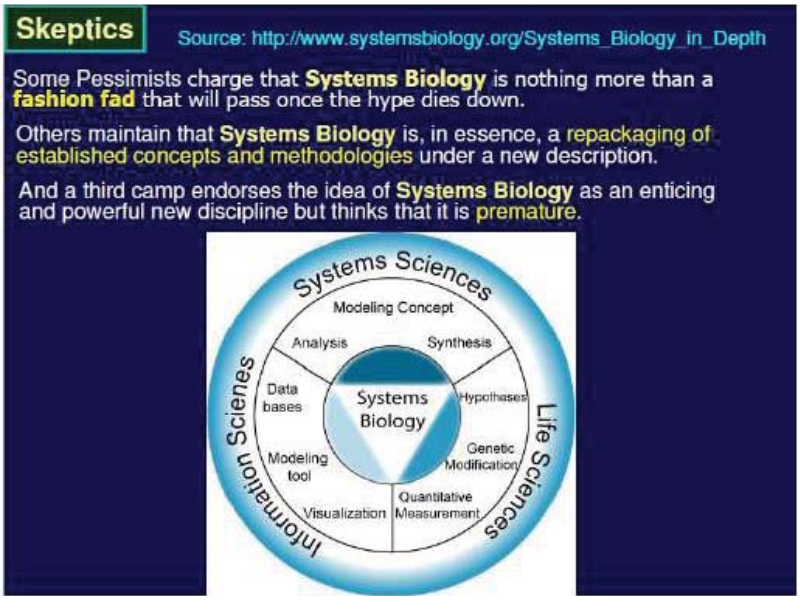


Figure 1-46: Sceptic opinions on Systems Biology [Maria, 2018].

[Maria,2021; Maria et al., 2018b; Maria and Luta,2013]. Alternatively, the Flux Balance Analysis (FBA, [Stephanopoulos et al., 1998]), can be applied (Figure 1-41-43). Cell fluxes can also be estimated by using the measured fluxes of some metabolites in a chemostat (experimental bioreactor), together with a mass balance model of cell metabolites of interest, and applying an optimization method in the presence of system stoiciometric constraints (Figure 1-42, Figure 1-43)[Price et al., 2004; Banga, 2008]. Here it is also to mention the short review of Blass et al. [2017], about how to use FBA to obtain GMOs by also using the COBRA toolbox [Schellenberger et al., 2011] to re-construct reaction pathways by using KEGG REACTION database [KEGG, 2011; KEGG pathway, 2011].

To illustrate in a simple way how a kinetic model of a cell metabolic pathway can be used to design a GMO by using the classic ChBPE rules and algorithms, let us consider the problem of Hatzimanikatis et al. [1996] (Figure 1-51): to solve the NLP equivalent problem to design a GMO by re-configuration of the metabolic pathway for Phenyl-alanine synthesis in *E. coli*, in order to maximize its cellular synthesis.

That implies to modify the structure and activity of the involved enzymes, and modification of the existing regulatory loops. Searching variables of the formulated mixed-integer nonlinear programming (MINLP) optimization problem are the followings: the regulatory loops (that is integer variables, taking " 0 " value when the loop has to be deleted, or the value " 1 " when it has to be

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

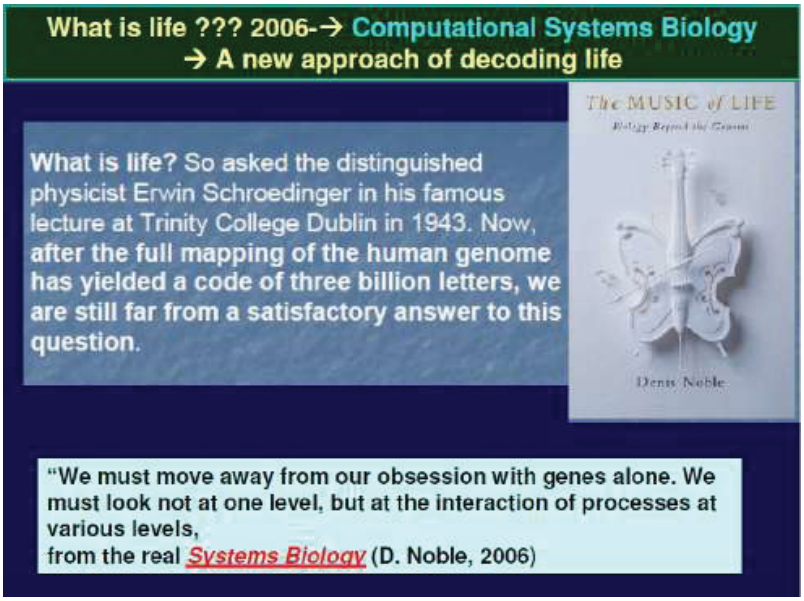


Figure 1-47: Defining concepts in the computational Systems Biology [Noble, 2006].

retained), together with the enzyme expression levels (that is continuous variables), and all these in the presence of the stoichiometric and thermodynamic constraints. To solve this complex optimization problem, two contrary objectives are formulated: maximization of the Phenyl-alanine selectivity, with minimization of cell metabolites' concentration deviations from their homeostatic levels (to avoid an unbalanced cell growth). The elegant solution of the problem is the so-called Pareto-optimal front (Figure 1-52), which is the locus of the best trade-off between the two adverse objectives. By choosing two problem solution alternatives from this Pareto-curve (Figure 1-52), it is to observe the large differences between the two pathways into the cell, fully achievable by using the genetic engineering.

In spite of the near astronomic complexity of the cell metabolism, including a large number of species and enzymatic reactions, various math cell models have been developed over decades, especially those concerning the CCM, see the brief review of Maria [2014a](Figure 1-32, Figure 1-38-40).

The first attempts to model the cell metabolism includes modest topological models belonging to the so-called Metabolic Control Analysis (MCA, [Kacser and Burns, 1973; MCA Web, 2004; Heinrich and Schuster, 1996] (Figure 1-22). Such structure-oriented analyses ignore nonessential mechanistic details and the process kinetics, and use the only network topology to quantitatively characterize to what extent the metabolic reactions determine the fluxes and metabolic concentrations [Heinrich and Schuster, 1996]. The MCA is focus on using various types of so-called "control coefficients", or "elasticity coefficients" to express the response of stationary fluxes or concentrations to perturbation parameters (internal or external species, fluxes). Besides, these coefficients have to fulfil the 'summation theorems', which reflect the network structural ("whole cell") properties, and the 'connectivity theorems' related to the dependence of a single enzyme activity on the whole system behaviour.

When developing *in-silico* methods to obtain GMOs by optimizing the evolutionary metabolic systems, math models based on the cell bioprocess mechanism (that is the 'structured' models) are to be used.

Such metabolic models, elaborated in a reduced or in an extended form, have been proposed over decades by using the topological MCA, but also by using the deterministic approach of the CHBPE, and of the nonlinear system control concepts and tools (below described), by imposing one or more appropriate performance criteria, such as:

a. Maximize the biomass production,

b. Maximize the synthesis of a target metabolite,

c. Maximize target reaction rates (steady-state fluxes);

d. Minimize the metabolic intermediate concentrations;

e. Minimize the transient times between two steady-states (QSS, homeostasis);

f. Optimize the reaction stoichiometry (network topology, by gene mutations);

g. Maximize the thermodynamic efficiency, etc. [Heinrich and Schuster, 1996].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

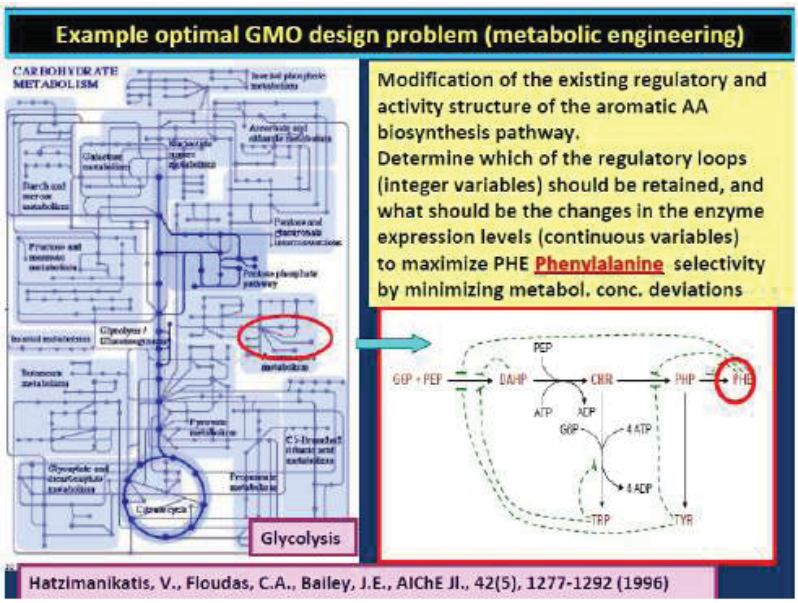


Figure 1-51: A simple example to solve the NLP problem to GMO design by reconfiguration of the metabolic pathway for Phenyl-alanine synthesis [Hatzimanikatis et al. 1996].

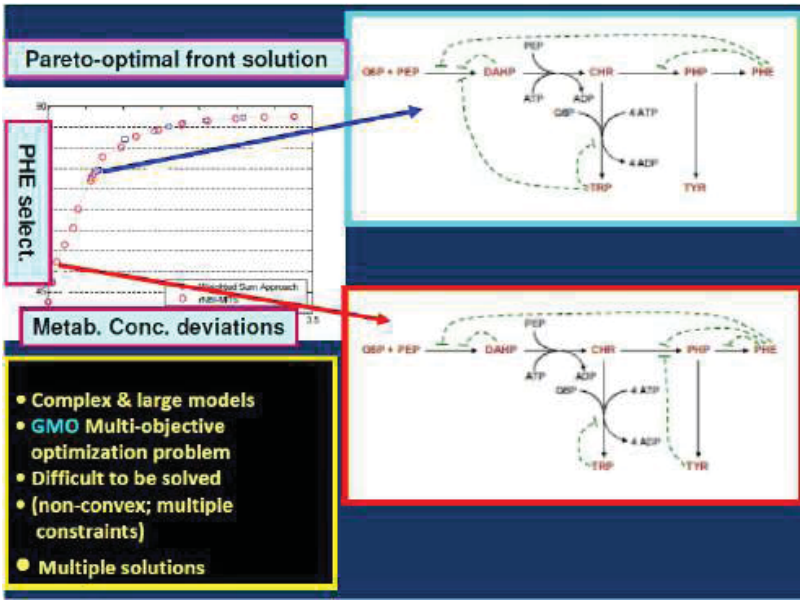


Figure 1-52: The resulted Pareto-optimal front (of two contrary objectives, i.e. max. selectivity in Phenyl-alanine, and min. deviations in metabolites stationary concentrations) for solving the GMO design problem Fig. 1-51. Two alternative solutions have been obtained, defining two metabolic distinct configurations. [Hatzimanikatis et al. 1996].

Often, a multi-objective optimization is applied by combing several of the above objectives, to obtain a desired GMO (see for instance the application of Maria et al. [2011]). All these objectives are subjected to various mass balance, thermodynamic, and biological constraints (some of them formulated in an 'artificial' way) [Heinrich and Schuster, 1996]. However, by not accounting for the system dynamics, and grounding the analysis on only the linear system theory, the topological methods, like those of the MCA presents inherent severe limitations (see for instance some violations of the stoichiometric constraints discussed by Atauri et al. [1999], or the modified control coefficients of Szedlacsek et al. [1996]).

As proved by Maria [section 3], the deterministic structured SMDHKM kinetic models developed under the novel WCVV modelling framework removes most of these shortcomings of the MCA, Booleanapproach, and of the classical (default) WCCV ('whole-cell, constant volume').

Amazing, but the first pioneers in dynamic modelling of biological systems were not the (bio)chemical engineers, which are better trained in 'translating' from the 'language' of molecular biology to that of mechanistic (bio)chemistry,

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

by preserving the structural hierarchy and component functions. The first dynamic models of some cell processes have been reported by the electronics engineers [Hodgkin and Huxley, 1952; Noble, 1962]. Later, such 'electronic circuits-like' models have been extensively used to understand intermediate levels of regulation (for instance, the A-cell of Ichikawa [2001]), but they failed to reproduce in detail molecular interactions with slow and continuous responses to perturbations and, eventually, they have been abandoned. However, the electronics engineers underlined the main characteristics of the cell systems, which must be included in any simulation model, that is (Figure 1-20):

I) the *dynamic character* of species interactions and processes [Sutherland, 2005];

II) the *feedback* character of processes ensuring their optimal regulation [Wolkenhauer and Mesarovic, 2005];

III) *optimal regulation* of the cell syntheses [Sutherland, 2005];

IV) All the above functions are met by *consuming minimum of resources* (nutrients/substrates), and cell energy;

V) All the above functions must ensure *maximum reaction rates* of the cell processes, with *minimum levels of intermediate species* [Heinrich and Schuster, 1996].

All these cell metabolic characteristics will be accounted in all the subsequent cell *in-silico* simulators based on extended mathematical models, that is the SMDHKM kinetic models developed under the novel WCVV modelling framework [Maria, 2017b,2018,2023]. All these characteristics are in fully agreement to the Darwin theory "Living organisms have evolved to maximize their chances for survival; It explains structures, behaviours of living organisms."(Figure 1-18). From such very incipient efforts to model-based design GMOs, 40 years latter it was Noble [2002] who pointed-out the tremendous advanced in the Systems Biology and in the *in-silico* design of GMOs, or even tissues, by means of computational systems biology [Kitano, 2002,2002b; Ideker et al., 2001].

A review of mathematical model types used to describe metabolic processes is presented by [Maria, 2005; Styczynski and Stephanopoulos, 2005; Stephanopoulos et al., 1998]. Each model type presents advantages but also limitations.

Roughly, to model the complex metabolic regulatory mechanisms at a molecular level (i.e. GERMs, and GRCs), two main approaches have been developed over decades: 1) a *structure-oriented (topological) analysis*, and 2) the *dynamic (kinetic), mainly deterministic models* [Stelling et al., 2002]. Each theory presents strengths and shortcomings in providing an integrated predictive description of the cellular regulatory network.

Topological models

The structure-oriented analyses ignore some mechanistic details and the process kinetics, and use the only *network topology*, the so-called 'Metabolic Control Analysis' (MCA) being focus on using various types of sensitivity coefficients (the so-called 'response coefficients'), which are quantitative measures of how much environmental perturbations (influential variable x_j) affects the cell-system states y_j (e.g. r = reaction rates, J = fluxes, C = concentrations) in a neighborhood of the steadystate (QSS, of index 's'), i.e. $\left[\left(\partial y_i / \partial y_{is}\right) / \left(\partial x_j / \partial x_{js}\right)\right]_s$. The systemic response of fluxes (i.e. stationary metabolic reaction rates), or of concentrations to perturbation parameters (i.e. the 'control coefficients'), or of reaction rates to perturbations (i.e. the 'elasticity coefficients') have to fulfil the 'summation theorems', which reflect the network structural properties, and the 'connectivity theorems' related to the properties of single enzymes *vs.* the whole cell system behaviour [Maria, 2005, Heinrich and Schuster, 1996].

The MCA methods present advantages, but also a large number of limitations. They are able to efficiently characterize the metabolic network robustness and functionality, linked with the cell phenotype and gene regulation. Also, MCA allows a rapid evaluation of the system response to perturbations (especially of the enzymatic activity), the possibilities of control and self-regulation for the whole cell pathway or of some CCM subunits. Functional subunits are metabolic subsystems, called 'modules', such as amino acid or protein synthesis, protein degradation, mitochondria metabolic pathways, etc. [Kholodenko et al., 1998]. Because the living cells are self- evolutionary systems, new reactions recruited by cells (by means of adaptive genetic mutations), together with enzyme adaptations (by genetic mutations) can lead to an increase in the cell biological organisation and to optimal regulatory performance indices (section 2.2). When constructing methods to optimize evolutionary metabolic systems, the use of MCA concepts and appropriate regulatory performance criteria, are used to: 1) maximize the metabolic reaction rates, and steady-state fluxes; 2) minimize the metabolic intermediate concentrations; 3) minimize the transient times between steady-states (QSS); 4) optimise the reaction stoichiometry (network topology, genetic modifications); 5) maximize the thermodynamic efficiency. All these objectives are subjected to various mass balance, thermodynamic, and biological constraints [Heinrich and Schuster, 1996] However, by not accounting for the system dynamics, and grounding the analysis on the linear system theory, these topological methods presents a large number of limitations. See for instance some violations of stoichiometric constraints discussed by Atauri et al., [1999], or the need to use corrected MCA coefficients [Szedlacsek et al., 1996].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Structured (mechanism-based) kinetic models

From the mathematical point of view, various structured (mechanism-based) dynamic (kinetic) models have been proposed to simulate the metabolic syntheses and their regulation, accounting for continuous, discrete, and/or stochastic variables, in a modular construction, 'circuit-like' network, or compartmented simulation platforms [Crampin and Schnell, 2004; Maria, 2005,2006,2017b,2018,2023; Bower and Bolouri, 2001].

By applying various modelling routes, successful structured models have been elaborated to simulate various regulatory mechanisms of cell syntheses, that is GERMs and GRCs. [McAdams and Arkin, 1998; Mizuno, 1997; McAdams and Shapiro, 1995; Tyson and Novak, 2001; Somogyi and Sniegoski, 1996; Heinrich et al.,1977; Hyver and Le Guyader, 1990; Wuensche and Lesser, 1992; Bray et al., 1993; Thomas et al.,1995; Tyson et al., 1996] In fact, as mentioned by Crampin and Schnell [2004], a precondition for a reliable modelling is the correct identification of both topological and kinetic properties. As few (kinetic) data are present in a standard form, non-conventional estimation methods have been developed, by accounting for various types of information (even incomplete) and global cell (regulatory) properties. [Maria, 2004; Crampin and Schnell, 2004].

Briefly, the math models used by Systems Biology are of the following four types.

Deterministic continuous variables models

Such models can perfectly represent the cell response to continuous perturbations, and their structure and size can be easily adapted based on the available -omics data information [Maria, 2005, 2006,2017b,2018,2023; Heinrich and Schuster, 1996; Stephanopoulos et al., 1998; Bower and Bolouri, 2001; Chen et al., 1999]. As a special category, the deterministic structured SMDHKM kinetic models developed under the novel WCVV modelling framework removes most of these shortcomings of the MCA, Boolean-approach, and of the classical (default) WCCV ('whole-cell, constant volume'), as proved by Maria [section 2, and section 3].

Deterministic continuous variable models present a large number of advantages, such as:

i). They perfectly can represent the cell response to continuous perturbations [Maria,2002,2005,2006,2007,2009,2017a,2017b,2018]; [Sewell et al.,2002]; [Heinrich and Schuster, 1996];

ii). Their structure and size can be easily adapted based on the available -omics databank information [Maria, 2005, 2006,2017b,2018,2023; Heinrich and Schuster, 1996; Stephanopoulos et al., 1998; Bower and Bolouri, 2001; Chen et al., 1999].

iii). They can satisfactorily simulate GERMs and GRCs [Chen et al., 1999; Maria et al., 2002; Maria, 2002, 2005, 2006,2017b,2018,2023].

iv). The huge advantages coming from the used concepts, rules, and algorithms of the CHBPE, and of the nonlinear system control theory, briefly reviewed in the (Figure 1-5, Figure 1-6, Figure 1-53, Figure 1-54) [Maria,2017].

Classical approach to develop *deterministic dynamic models* is based on a hypothetical reaction mechanism, kinetic equations, and known stoichiometry (see [Aris, 1969; Froment and Bischoff, 1990]). This route meets difficulties when the analysis is expanded to large-scale metabolic networks, because the necessary mechanistic details and standard (structured) kinetic data to derive the rate constants are difficult to be obtained. However, advances in genomics, transcriptomics, proteomics, and metabolomics, lead to a continuous expansion of the bioinformatic omics databases, while advanced numerical techniques, non-conventional estimation procedures, and massive software platforms reported progresses in formulating such reliable cell models. Valuable *structured dynamic models*, based on cell biochemical mechanisms, have been developed for simulating various (sub)systems, [Maria,2005,2006,2007,2009,2014a,2014b,2018,2023, 2021], or [Maria,2021; Maria and Luta,2013; Maria,2010; Maria and Renea,2021].

To model in detail the cell process complexity is a challenging and difficult task. The large number of inner cell species, complex regulatory chains, cell signalling, motility, organelle transport, gene transcription, morphogenesis and cellular differentiation cannot easily be accommodated into existing computer frameworks. Inherently, any model represents a simplification of the real phenomenon, while relevant model parameters are estimated based on the how close the model behaviour is to the real cell behaviour. A large number of software packages have been elaborated allowing the kinetic performance of enzyme pathways to be represented and evaluated quantitatively (reviews of Maria [2005], and of Hucka et al. 2003]). Oriented and unified programming languages have been developed, such as SBML Hucka et al. 2003], JWS [Olivier and Snoep, 2004; Peters et al. 2017], to include the bio-system organization and complexity in integrated platforms for cellular systems simulation (ECell, V-Cell, M-Cell, A-Cell, see section 1.3.3). Such integrated simulation platforms tend to use a large variety of biological databanks including enzymes, and -omics information, together with metabolic reactions [CRGM-database, 2002; NIH-database, 2004].

Development of *deterministic dynamic models* to adequately reproduce such complex synthesis related to the CCM [Maria, 2014a, 2014b, 2021; Kurata and Sugimoto, 2017; Chassagnole et al., 2002; Miskovic et al., 2015] but also the genetic regulatory systems (GRCs) tightly controlling the metabolic syntheses reported significant progresses over the last decades, in spite of the lack of structured experimental kinetic information, being rather based on sparse information from various sources and unconventional identification / lumping algorithms [Maria, 2005,2017a,2017b,2018, 2021, 2014a,2004]).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

By applying various modelling routes, successful *structured, deterministic dynamic* SMDHKM kinetic models, with continuous variables (that is ODE species mass balances) have been elaborated to simulate various regulatory mechanisms [Maria, 2005,2017a,2017b,2018, 2021, 2014a,2004]; [Tyson and Novak, 2001; Heinrich et al., 1977], or using stochastic variables [McAdams and Arkin, 1997; McAdams and Shapiro, 1995; Somogyi and Sniegoski, 1996]. Notation: ODE = ordinary differential equations. In fact, as mentioned by Crampin and Schnell [2004], a precondition for a reliable modelling is the correct identification of both topological and kinetic properties. As few (kinetic) data are present in a standard form, non-conventional estimation methods have been developed, by accounting for various types of information (even incomplete) and global cell (regulatory) properties [Crampin and Schnell, 2004; Maria, 2002,2004,2005,2007,2009,2017a,2017b,2018; Maria and Scoban, 2017,2018; Maria et al.,2002].

In such deterministic models, due to the CCM very high complexity only the essential reactions are retained, the model extension depending on the measurable variables, and on the available information. An important problem to be considered is the distinction between the qualitative and quantitative process knowledge, stability and instability of involved species, the dominant fast and

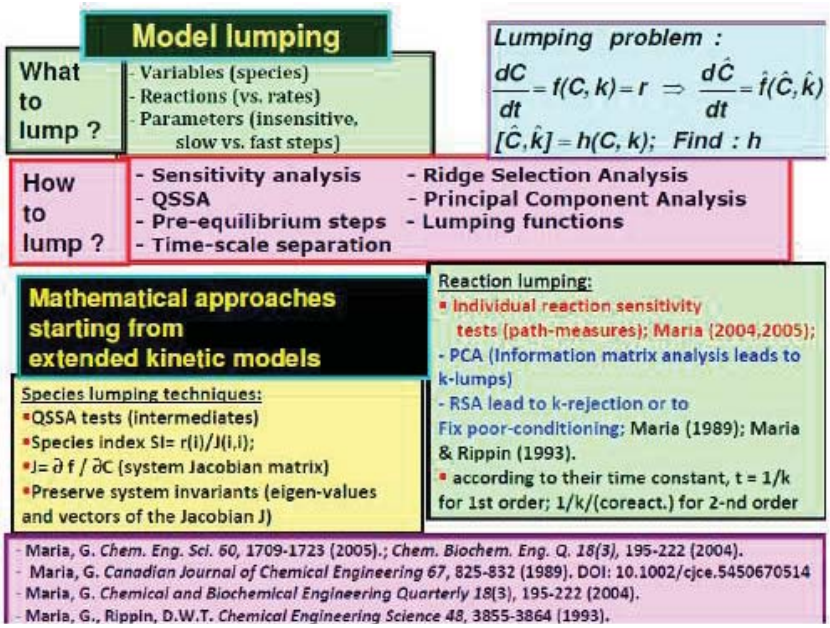


Figure 1-53: Some guidelines used in obtaining reduced dynamic models, according to [Maria, 2004, 2005b, 2019].

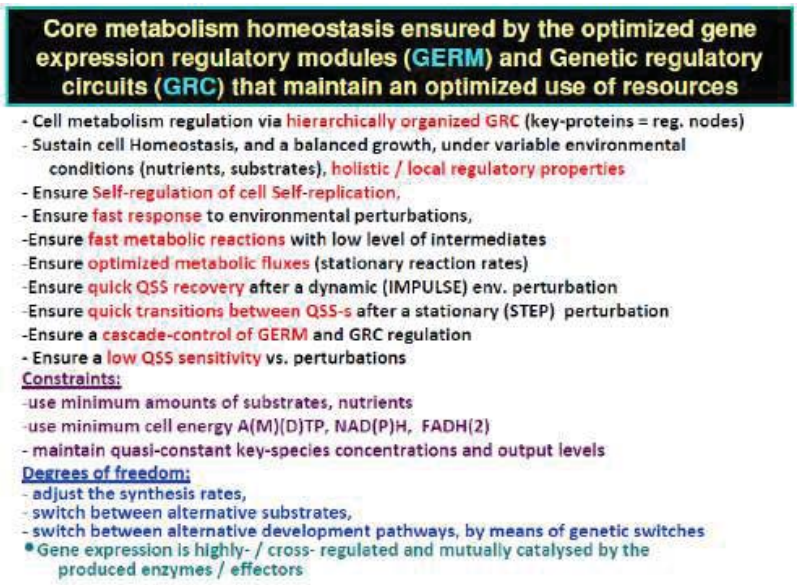


Figure 1-54: The role of the gene expression regulatory modules (GERM-s), and of the genetic regulatory circuits (GRC-s) [Maria,2017a, 2017b,2018], and of Maria [2003,2005,2006,2007,2009,2014].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

slow modes of process dynamics, reaction time constants, macroscopic and microscopic observable elements of the state vector. Such kinetic models can be useful to analyse the regulatory cell-functions, both for stationary and dynamic perturbations, to model cell cycles and oscillatory metabolic paths [Surovstev et al.,2007; Surovtsev et al., 2008, Sewell et al.2002], the effect of external perturbations on the cell metabolism and GRCs [Maria, 2002, 2005,2006,2007,2009,2017a,2017b,2018], and to reflect the species interconnectivity or perturbation effects on cell growth. [Bower and Bolouri, 2001].

Boolean (discrete) variable models

The Boolean or topological approach can not characterize the dynamics of the metabolic processes.

Even if regulation mechanisms are not fully understood, metabolic regulation at a low-level is generally better clarified, and conventional ordinary differential (ODE) kinetic models with continuous variables, with a mechanistic description of reactions tacking place among individual species (proteins, mRNA, intermediates, etc.) have been proved to be a convenient route to analyse continuous metabolic / regulatory processes and continuous perturbations. When systems are too large or poorly understood, coarser and more phenomenological kinetic models may be postulated (e.g. protein complexes, metabolite channelling, etc.). In dynamic ODE models, only essential reactions are retained, the model complexity depending on measurable variables and available information. An important problem to be considered is the distinction between the qualitative and quantitative process knowledge, stability and instability of involved species, the dominant fast and slow modes of process dynamics, reaction time constants, macroscopic and microscopic observable elements of the state vector. Such ODE kinetic models can be useful to analyse the regulatory cell-functions, both for stationary and dynamic perturbations, to model cell cycles and oscillatory metabolic paths, and to reflect the species interconnectivity or perturbation effects on cell growth. [Kobayashi et al. 2004].

Such a topological Boolean model is displayed in the (Figure 1-40) [Stelling, 2004; Bower and Bolouri, 2001; McAdams and Arkin, 1998; McAdams and Shapiro, 1995; Somogyi and Sniegoski, 1996; Akutsu et al., 1998]. Due to the very large number of states $0(10^3 - 10^4)$, and $0(10^3)$of transcriptional factors (TF) involved in the gene expression, such GRC models are organized in clusters, modules, of a multi-layer organization (Figure 1-40) [Bower and Bolouri, 2001; Blass et al., 2017].

In the *Boolean approach*, the variables can take only discrete values. Even if less realistic, such an approach is computationally tractable, involving large networks of genes that are either "on" or "off" (e.g. a gene is either fully expressed or not expressed at all; Figure 1-40) according to simple Boolean relationships, in a finite space. Such a coarse representation is used to obtain a first model for a complex bio-system including a large number of components, until more detailed data on the process dynamics become available. "Electronic circuits" structures. See an example in Figure 1-17, and Figure 1-37 [Ichikawa, 2001]. Such models have been extensively used to understand intermediate levels of regulation, but they cannot reproduce in detail molecular interactions with slow and continuous responses to perturbations. That is because they can not simulate the dynamics of the metabolic processes.

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

Stochastic variable models

The *stochastic models* replace the 'average' solution of continuous variable ODE kinetics (e.g. species concentrations mass balances; ODE = ordinary differential equations) by a detailed random-based simulator accounting for the exact number of molecules present in the system (copy-numbers for a living cell). Because the small number of molecules for a certain species is more sensitive to stochasticity of a metabolic process than the species present in larger amounts, simulation via continuous models sometimes can lack of enough accuracy for random process representation (as cell signalling, gene mutation, etc.). For such random cell processes, the Monte Carlo simulators (stochastic models) are used to predict the individual species molecular interactions, while the rate equations are replaced by individual reaction probabilities, and the model output is stochastic in nature. Even if the required computational effort is extremely high, stochastic representation is useful to simulate the cell system dynamics for species present in a small copynumber, and/or of which spatial location is important [McAdams and Arkin, 1997; Shimizu and Bray, 2002; Gillespie and Mangel, 1981; Gillespie, 1977] (Figure 1-37).

Mixed variable models

Mixtures of ODE dynamic kinetic models (with continuous variables), with kinetic models including discrete state variables (i.e. "continuous logical" models), or mixtures of continuous ODE dynamic kinetic models (of continuous variables), with including stochastic terms can lead to promising mixed models able to simulate combined deterministic and non-deterministic cell processes [Bower and Bolouri, 2001].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

2. Deterministic continuous kinetic models constructed by using the chemical and biochemical reaction engineering (ChBRE) principles and tools

In the deterministic approach, the cell kinetic models correspond to ODE models with continuous variables. Ideally, such models should include the all species of the cell. This is practical impossible to be realized because of the astronomic-like complexity of the cell. The cell system includes a very large number of states $0(10^3 - 10^4)$, and $0(10^3)$ of transcriptional factors (TF) involved in the gene expression (GERMs, and GRCs), the CCM / GRC models includes only the key-species, and the metabolites of interest, and eventually lumped species and/or reactions. The deterministic kinetic models take the advantage of the structural organization of living cells (Figure 1-2). The cell CCM / GRCs are organized in clusters, modules, of a multi-layer organization (Figure 1-2, Figure 1-40, Figure 1-17) [Bower and Bolouri, 2001; Blass et al., 2017].

In other words, the CCM includes a certain number of subunits. *Functional subunits* are metabolic subsystems, called 'modules', such as the amino acid or protein synthesis, protein degradation, mitochondria metabolic pathway, etc. [Kholodenko et al. 1998]. In particular, protein synthesis homeostatic regulation (a chain of GERM modules organized in a GRC) includes a multi-cascade control of the gene expressions with negative feedback loops and allosteric adjustment of the enzymatic activity Kholodenko [2000]; Sewell et al. [2002]. In the case of modelling GRCs, by chance, the number of interacting GERMs is limited, one gene interacting with no more than 23-25 [Kobayashi et al., 2004]. Such a modular construction of the cell models make them computationally very tracktable. It results that, a convenient way to model metabolic processes and regulatory networks is the *modular* approach. Spatial and functional compartments together with functional modules have been defined when developing complex cell simulation platforms: (E-Cell [Tomita et al., 1999], V-Cell [Schaff et al., 2001], M-Cell [Bartol and Stiles, 2002], ACell [Ichikawa, 2001]. Such integrated simulation platforms tend to use a large variety of biological databanks including enzymes, proteins and genes characteristics together with metabolic reactions (CRGM-database [2002]; NIH-database [2004]).

To avoid large deterministic models, difficult to be identified, due to the lack of kinetic information, and difficult to be used, a lumping procedure should be applied (Figure 1-53).

Reduction in the cell model structure (via lumping of species, and reactions) is necessary due to the following reasons [Maria, 2017a, 2017b, 2018];

- the high complexity of cell metabolic processes *vs.* available data
- large number of species, reactions, transport parameters, and species interactions
- low data observability and reproducibility
- metabolic process variability
- interpretable representation of cell complexity
- requirement to obtain quick simulations of the cell behaviour under various environmental conditions
- computational tractability and easier application of algorithmic rules from (bio)chemical engineering and numerical calculus

To exemplify such quick lumping rules, let us consider the classical (default) ODE kinetic model for a couple of cell and reactions inside the cell, that is the WCCV model.

The ODE deterministic models have been developed in two alternatives, [Maria et al., 2018d, 2018g]:

i). The classical (default) Whole-cell Constant Volume (WCCV) kinetic models [Maria, 2017a,2017b,2018,2023; Aris, 1969; Froment and Bischoff, 1990; Levenspiel, 1999].

ii). The novel Whole-cell Variable Volume (WCVV) of the cell metabolic processes (whole-cell, isotonic variable-volume) that maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. This WCVV novel modelling framework was introduced by Maria [2002], and by [Maria et al., 2002], while the study of the WCVV properties are given by Maria over the interval [2005-2023] (see the reference list given in the book abstract).

2.1. Whole-cell Constant Volume (WCCV) classical (default) kinetic models

Definitions:

The default "**Whole-Cell Constant Volume" (WCCV)** classical continuous variable ODE dynamic (kinetic) models of the cell processes, do not explicitly consider the cell volume exponential increase during the cell growth. When the continuous variable WCCV dynamic models are used to model the cell processes, the default-modelling framework is those of the WCCV formulation Eq.(1a).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The Eq.(1b) results from Eq. (1a), by defining the species concentration definition as *Cj = nj/V*. This transformation to the default kinetic model Eq.(1b) 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, becoming double after one cell cycle.

Such WCCV models do not explicitly consider the cell volume exponential increase during the cell growth. When such dynamic models Eq.(1a-c) are used to simulate the cell enzymatic processes, wrong results are expected because the hypotheses of a constant cell volume and, implicitly, of a constant osmotic pressure are not fulfilled. To partially overcome these shortcomings, some WCCV 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.

The formulation Eq.(1a-1c) assumes a homogeneous volume with no inner gradients or species diffusion resistance. 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. Being very overparameterized and strongly nonlinear, parameter estimation of such WCCV models from quasi-steady-state (QSS) data (species 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 [Banga, 2008; Maria,2004].

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 copynumbers: $n = n_1 \ \dots \ n_j \ \dots \ n_{ns}$, and the corresponding concentrations: $C = \left[\frac{n_1}{V} \ \dots \ \frac{n_j}{V} \ \dots \ \frac{n_{ns}}{V}\right] = \left[C_1 \ \dots \ C_j \ \dots \ C_{ns}\right]$.

where "ns" = number of cell species (taken individually or lumped). Then, the kinetic (dynamic) model of a cell process can be formulated as is Eq. (1a) in terms of copynumbers. This formulation (1a) assumes a homogeneous volume with no inner gradients or species diffusion resistance, similar to a batch ideal reactor [Froment and Bischoff, 1990; Dutta, 2008; Moser, 1988]. The used reaction rate expressions for the metabolic reactions are usually those of extended Michaelis-Menten or Hill type. Being very over-parameterized and strongly nonlinear, parameter estimation of such models in the presence of multiple constraints translates into a mixed integer nonlinear progeamming problem (MINLP) difficult to be solved because the searching domain is not convex (Banga, 2008).

Such a WCCV dynamic model might be satisfactory for modelling many cell subsystems, but not for an accurate modelling of cell GRC and holistic cell properties under perturbed conditions, or the division of cells [Morgan et al., 2004; Surovstev et al.,2007], by distorting very much or even misrepresenting the prediction results, as proved by Maria [2017a, 2017b,2018; Maria et al., 2018d; Maria and Luta, 2013] in section 2.3.1. for both stationary, or perturbed cell growing conditions. The basic mass balance equations of a WCCV model, self-understood written for a constant volume (V) of the system are the followings Eq.(1a-c).

$\frac{1}{V(t)}\frac{dN_j}{dt} = \sum_{i=1}^{nr} s_{ij} r_i(\boldsymbol{N}/V, \boldsymbol{k}, t) = h_j(\boldsymbol{C}, \boldsymbol{k}, t)$	**(1a)**
$\frac{d\left(N_j/V\right)}{dt} = \frac{dC_j}{dt} = \sum_{i=1}^{nr} s_{ij} r_i(\boldsymbol{N}/V, \boldsymbol{k}, t) = h_j(\boldsymbol{C}, \boldsymbol{k}, t)$	**(1b)**
$\left(\frac{dC_j}{dt}\right)_s = \sum_{i=1}^{nr} s_{ij} r_i(\boldsymbol{N}/V, \boldsymbol{k}, t) = h_{j,S}(\boldsymbol{C_S}, \boldsymbol{k}, t) = \boldsymbol{0}, \ j = 1,\dots,ns$	**(1c)**

Where notations are the followings: C(j) = (cell-)species j concentration (of vector C); V = system volume (cell cytosol plus membrane); N(j) = (cell-)species j number of moles (copynumbers); r(j) = j-th reaction rate (individual or lumped; s(i,j) = stoichiometric coefficient of the species "j" (individual or lumped) in the reaction "i"; t = time; k = rate constant vector; i =1,...," nr" (no. of reactions). Index "s" denotes the cell steady-state QSS (homeostasis, balanced growth). The model Eq.(1c) characterising the cell steady-state is obtained from Eq.(1b) by setting the derivatives on the left side to zero.

Here it is to observe that Eq.(1b) can be derived from Eq.(1a) only 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 double during a cell cycle (tc), with the following average (apparent) cell dilution rate (Dm):

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

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

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

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

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

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.

Consequently, Eq.(1a) corresponds to that of a constant volume of the system, and, implicitly of a constant osmotic pressure, eventually accounting for the cell-growing rate as a pseudo-'decay' rate of the keyspecies (often lumped with the degrading rate) in a so-called 'diluting' rate. The model Eq. (1c) represents the quasi-steady-state (QSS) of the cell processes (homeostasis, that is an equilibrated growth).

At this point, it is to observe that the metabolic processes at a low(molecular)-level are generally better clarified. Based on that, conventional dynamic models WCCV, based on ordinary differential (ODE) species mass balance Eq.(1a-b,2), with a *mechanistic (deterministic) description* of reactions tacking place among individual species (proteins, mRNA, intermediates, etc.) have been proved to be a convenient route to analyse continuous metabolic / regulatory processes and perturbations. When systems are too large or poorly understood, coarser and more phenomenological kinetic models may be postulated (e.g. protein complexes, metabolite channelling, etc.). In dynamic deterministic models, usually only essential reactions are retained, the model complexity depending on the measurable variables and available information. To reduce the structure of such a model, an important problem to be considered is the distinction between the qualitative and quantitative process knowledge, stability and instability of involved species, the dominant fast and slow modes of process dynamics [represented by the eigenvalues of the ODE model Jacobian matrix Eq.(4)], reaction time constants [Maria et al., 2010], macroscopic and microscopic observable elements of the state vector. Such kinetic models can be useful to analyse the regulatory cell-functions, both for stationary and dynamic perturbations, to model cell cycles and oscillatory metabolic pathways [Maria, 2014a], and to reflect the species interconnectivity or perturbation effects on cell growth [Maria, 2017a,2017b,2018; Maria et al., 2018a].

Representation of metabolic process kinetics is made usually by using rate expressions of extended Michaelis-Menten or Hill type [Miskovic et al., 2015; Maria, 2014a, 2014b]. Some reaction rate expressions used to represent the CCM models, and of GERMs, or of genetic regulatory circuits (GRCs), or of genetic switches (GS) are given in (Figure 1-55) [Maria, 2005, 2017a,2017b,2018].

However, the formulation Eq.(1a) presents several properties, as followings:

1). The rank of the stoichiometric matrix S = {s(i,j)} indicates the minimum number of linearly independent reactions ("modes") which can reproduce the reacting system. Of course, the kinetic model of the cell processes includes a much larger number of reactions due to the complexity of the cell processes, and the requirement to derive an extended representation of them [Froment and Bischoff, 1990; Dutta, 2008; Moser, 1988]. More details are given by [Maria, 2005b].

The stoichiometric matrix **S**, and its rank, can offer more information about the cell system. Thus, conservation relations and reaction invariants can be derived in a simple way. For instance, the independent conservation vectors **g** of molecular species cause some rows of the transpose of the stoichiometric matrix $\mathbf{S}^T$ to be linearly dependent, leading to "ns-rank($\mathbf{S}^T$)" conservation relations [Maria,2005b]:

$$\boldsymbol{g}^T \boldsymbol{s}^T = \boldsymbol{o}^T \Rightarrow \boldsymbol{S} \cdot \boldsymbol{g} = 0 \quad (3)$$

Where the superscript "T" denotes the matrix transpose. The conservation relationships vectors g are, in fact, the nul space of **S**. The conservation relations can be written not only for species (like in Eq.(3)), but also for atoms, electric charge, thermodynamic constraints, or other non-negative constraints [Heinrich and Schuster, 1996].

When proposing a reaction pathway in order to build-up a kinetic model, several additional experimental-numerical techniques can help in pointing out species interactions. For instance, Vance et al. [2002] reviewed rules to check reaction schema via species inter-connectivities, by inducing experimental perturbations to a (bio)chemical system (by means of tracers or fluctuating inputs) and measuring perturbation propagation through consecutive and parallel reactions. "Distances" among observed species can be thus defined and included in proposing the reaction pathway.

2). The Jacobian of the kinetic model WCCV Eq.(1b-c), or WCVV Eq. (11) is defined at QSS conditions by:

$$J_C = (\partial \boldsymbol{h}(\boldsymbol{C},\boldsymbol{k}) / \partial \boldsymbol{C})_S, of\ elements: J_{i,k} = \left(\partial h_i(\boldsymbol{C},\boldsymbol{k}) / \partial C_k\right)_S \quad (4)$$

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

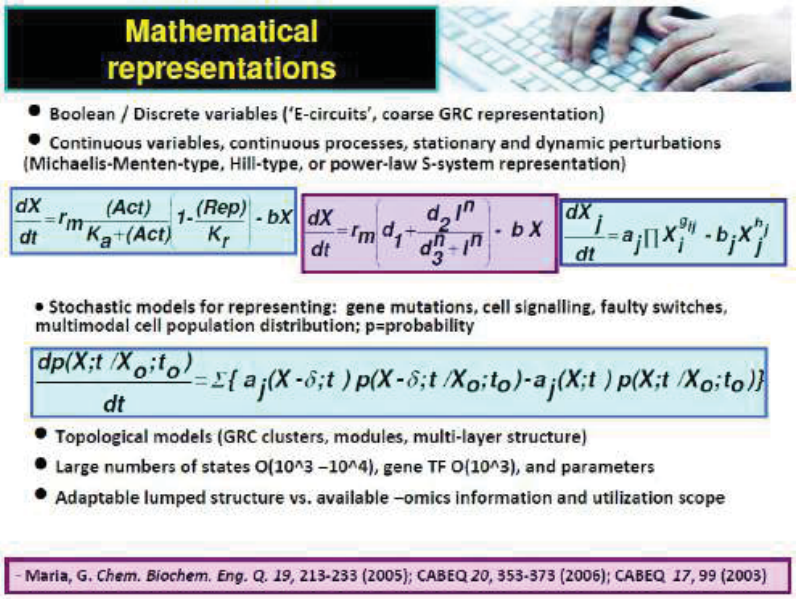


Figure 1-55: Some reaction rate expressions used to represent the CCM models, and those of GERM-s, genetic regulatory circuits (GRC-s), or genetic switches(GS) [Maria, 2005, 2014b, 2017a,2017b,2018].

When lumping an ODE kinetic model, the reactions invariants (that is the eigenvalues of the Jacobian matrix Eq.(4) have to be kept unchanged. Also, the directions and lengths of the eigenvectors have to be kept unchanged [Maria,2005b]. The eigenvalues (λ) and eigenvectors (x) of the Jacobian matrix J are computed with the formula: Jx = λx.

3). The differentiation of the steady-state form Eq.(1c) leads to evaluation of the relative sensitivity coefficients $S_{Nut\ j}^{C_i} = \left(\partial C_i / \partial C_{Nut_j}\right)_s$ of the cell system state variables (Ci) *vs.* the concentration of an external nutrient ($C_{Nut,j}$), explicitly included in the kinetic model, by solving the sensitivity equation [Maria,2017a,2017b; Maria and Scoban, 2017,2018]:

$$\left[\frac{\partial h_i(C,C_{Nut},k)}{\partial C_k}\right]_s \left[\frac{\partial C_k}{\partial C_{Nut_j}}\right]_s + \left[\frac{\partial h_i(C,C_{Nut},k)}{\partial C_{Nut_j}}\right]_s = 0 \qquad (5)$$

In the Eq.(5) linear set of equations, the ODE model Jacobian of Eq.(4), with the elements $J_{i,k} = \left(\partial h_i(\boldsymbol{C},\boldsymbol{k}) / \partial C_k\right)_S$ is numerically evaluated for the cell-system stationary-state Eq.(1c), where "s" index denotes the QSS stationary condition. Such sensitivity coefficients are very important when one evaluates the stability strength of a cell QSS (homeostasis).

Lumping deterministic ODE models

To avoid large deterministic models, difficult to be identified, due to the lack of structured kinetic information, and difficult to be used, a lumping procedure should be applied (Figure 1-53) [Maria, 2004, 2005b, 2019].

When reduce an extended ODE kinetic model, a tradeoff between model complexity and its adequacy must be maintained [Maria, 2004, 2006, 2009]. The lumped ODE kinetic model should be enough complex to include the key-species of interest, for instance for the *in-silico* design of a GMO, by *in-silico* re-programming the cell metabolism [Maria et al., 2011; Alberts et al., 2002], or by optimal cell cloning [Maria, 2010; Maria and Luta, 2013]. Application of systematic mathematical lumping rules [Maria, 2004, 2005b, 2019] to metabolic ODE kinetic models must account for physical significance of lumps, species interactions, and must preserve the *systemic/ holistic* properties of the metabolic pathway [Maria and Scoban, 2017,2018]. The only separation of components and reactions based on the time-constant scale (as in the modal analysis of the Jacobian of the ODE model has been proved to be insufficient [Tyson et al., 1996]).

The classic approach to develop dynamic reduced models is based on the chemical and biochemical reaction engineering (ChBRE) rules, that is: propose a hypothetical reaction mechanism based on literature information (-omics databanks) and own experience, adopt the reaction rate expressions and the reaction stoichiometry, and try to validate them on an experimental basis. This route meets difficulties when the analysis is expanded to large-scale metabolic networks, because the necessary mechanistic details and standard kinetic data to derive the rate constants of the built-up ODE kinetic model are difficult to be obtained. However, advances in the -omics sciences (genomics, transcriptomics, proteomics, and metabolomics) lead to a continuous expansion of bioinformatic databases, while advanced numerical techniques, non-conventional estimation procedures, and massive software platforms, and

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

lumping algorithms [Maria, 2004, 2005b, 2019], reported progresses in formulating lumped (reduced) but reliable cell models. Valuable *structured dynamic* kinetic models, based on cell biochemical mechanisms, have been developed for simulating various (sub) systems. See the examples of Maria [2005b,2006,2007,2009,2014a,2014b,2017a,2017b,2018,2023]).

The rate constants are usually estimated from stationary data Eq. (1c). As revealed by Alon [2007], and by Visser et al. [2004], "Traditionally, kinetic metabolic models are based on mechanistic rate equations, which are derived from in-vitro experiments. However, due to large differences between *in-vivo* and *in-vitro* conditions, it is unlikely that the *in-vitro* obtained parameters are valid *in-vivo*. Thus, the kinetic parameters must be adjusted, using data on *in-vivo* metabolite levels and fluxes obtained in dynamic experiments. Due to the complexity of mechanistic rate equations, which often contain a considerable amount of parameters (rate constants), this requires a large experimental and mathematical effort." Visser et al. [2004]. Such an approach is computational tractable, a large number of ChBRE, and non-linear system control algorithms being available.

Besides, application of lumping rules to metabolic processes must also account for physical significance, species interactions, and for preserving the *systemic* properties of the metabolic pathway. The only separation of components and reactions based on the time-constant scale (as in the modal analysis of the *Jacobian* matrix *J* case, [see Eq.(4) for *J* definition] has been proved to be insufficient.

However, here it is to mention, that the work with *reduced kinetic models* to simulate the cell metabolic syntheses and GRC-s, even if computationally very convenient, presents some inherent disadvantages, that is: i) multiple reduced model structures might exist difficult to be discriminated; ii) a loss of information is reported on certain species, on some reaction steps, and a loss in system flexibility (given by the no. of intermediates and species interactions); iii) a loss in the model prediction capabilities (as precision and detail degree); iv) a lack of physical meaning of some model parameters / constants thus limiting its robustness and portability; alteration of some cell / GRC holistic properties (stability, multiplicity, sensitivity).

Here can be mentioned only a few of the classical chemical engineering rules used for *reducing an extended kinetic model* [Maria, 2004, 2005, 2006; Maria, 2019)(Figure 1-53), as followings:

1). Reduce the list of reactions, by eliminating unimportant sidereactions and/or assuming quasi-equilibrium for some reaction steps. Use sensitivity measures of rate constants to detect the redundant part of the model. Apply for instance the following rules: the ridge selection, the principal component analysis, the time-scale separation, etc. [Maria, 2004, 2019].

2). Reduce the list of species, by eliminating unimportant components and/or lumping some species, by using various measures. For instance, eliminate the species presenting small values for the product of the target species " *i* " lifetime $\mathrm{LTi} = -1/J(i,i)$, and its production rate *r*(*i*), where the Jacobian elements Eq.(4) are $J(i,k) = \partial h(i)(C,k)/\partial C(k)$, where *h*(*i*) are the right-side functions of the ODE kinetic model Eq.(lb) of the cell process.

3). Decompose the kinetics into fast and slow 'parts' allowing application of the quasi-steady-state-approximation (QSSA) of the intermediate species of low and quasi-constant concentration, in order to reduce the model dimensionality [Maria, 2004, 2005b, 2019].

4). When the ODE kinetic model is linear in parameters, then the reduction procedure of Maria [2004, 2005b, 2019]), can be successfully applied by preserving the system Jacobian invariants (eigenvalues and eigenvectors). The elements of the eigenvectors corresponding to the lowest eigenvalue of the model Jacobian can be used to lump model reactions and/or species [Maria, 2005b].

Lumping reactions

Redundant (eliminable) reactions and rate constants can be detected in an extended model by using the sensitivity analysis and the *significance analysis*. One route is to apply the ridge selection completed with parameter *t*-tests and inter-correlation evaluation [Maria and Rippin, 1993; Maria, 1989]. A more detailed parameter sensitivity analysis [Maria, 2005b] can lead to rejecting the insignificant parameters based on individual reaction *j* sensitivity, $S_{g,j} = \sum_u^n \sum_i^r s^*_{ij,u}$, evaluated *vs.* every species *i* and run *u*, i.e. $s^*_{ij,u} = \left|\Delta C_{i,\,predicted}\right| / C_i(t_u, \boldsymbol{k})$, [Turanyi, 1990a-c; Vajda et al., 1985].

An alternative to lump kinetic parameters is based on the principal component analysis (PCA of [Malinowski, 1991]). The algorithm starts with evaluating species relative sensitivities *vs.* kinetic parameters, that is the vectors $S_u = [\partial \ln C_{iu} / \partial \ln k_m]$, and then derivation of the so-called "information" matrix s^Ts, defined in logarithmic terms, $\boldsymbol{S}^T = [\boldsymbol{S}_1 \quad \boldsymbol{S}_2 \quad \dots \quad \boldsymbol{S}_n]$. As revealed by Vajda et al. [1985, 1989], small S^TS eigenvalues (smaller than the smallest noise variance of species concentrations, $\lambda_{min} < \sigma^2$) correspond to eigenvectors X_i presenting proportional components. These eigenvectors are then used to construct approximate "linear dependencies among

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

species sensitivities". The exact linear dependencies correspond to totally singular case of $\lambda_{min}=0$, i.e. a structurally unidentifiable model. Starting from the approximate linear dependencies, Vajda et al. [1985] proposed a way to lump rate constants, of the form $k_i / k_j^{constant}$.

Lumping species

Lumping of species in an ODE kinetic model is performed when there is insufficient information to characterize the dynamics of all compounds, or when by-products and intermediate separate prediction is not crucial for the process analysis. In other terms, the adopted reduced reaction pathway has to contain no redundant species, while the information is condensed in a smaller set including groups of species represented as single variables (lumps).

Species subjected to elimination are identified based on small values for the product of species *i* lifetime (LT_i) and species production rate ($LT_i=-1/J_{ii}$, where $J_{ik}=\partial h_i(\boldsymbol{C},\boldsymbol{k})/\partial C_k$ are elements of the kinetic model Jacobian Eq.(4) [Tomlin et al., 1997; Turanyi et al., 1993]. Model simplifications are realised with the expense of a corresponding bias in the prediction capabilities for species concentrations, that is $e(\boldsymbol{C})=\Delta\boldsymbol{C}_{predicted}$. The same analysis can be carried out in logarithmic terms. Thus, the redundant species index " *i* " are detected based on a small global sensitivity index, $B_i(t)=\sum_{m=1}^{nr} s_{im}^2(t)$, defined for the all reactions, $s_{im}(t)=\partial\ln(r_m)/\partial\ln(C_i)$ [Tomlin et al., 1997].

Linear model lumping

In a more systematic approach [Maria,2005b], the lumped species (superscript " $\wedge$ ") are related to the original ones by a lumping function ***h***, which can be linear or non-linear:

$$\frac{d\boldsymbol{C}}{dt}=\boldsymbol{f}(\boldsymbol{C},\boldsymbol{k})\Rightarrow\frac{d\hat{C}}{dt}=\hat{f}(\hat{C},\hat{k});\hat{C}=\boldsymbol{h}(\boldsymbol{C});\text{size}(\hat{C})<\text{size}(\boldsymbol{C})$$

This technique is particularly attractive to characterize complex kinetic systems due to the links established between species and parameters of the extended and reduced models. Such relations can be used to interpret rate constants of extended models from using the reduced model parameters identified from a few number of observations. Unfortunately, for nonlinear models, there are no general rules to obtain nonlinear lumping of species [Maria, 2005b]

For a linear lumping applied to the original species:

$$\hat{C}_i=m_{i,l}C_l+\ldots+m_{i,n_s}C_{n_s};i=1,\ldots,\hat{n}_s;\hat{C}=\boldsymbol{MC}$$

Wei and Kuo [1969], and Li and Rabitz [1990, 1991a-b] indicate the necessary and sufficient conditions for an exact or approximate lumping, $\hat{f}(\hat{C})=\boldsymbol{M}\boldsymbol{f}(\boldsymbol{C})$.

The *exact* linear lumping matrix *M* can be constructed from the eigenvectors of J^T (i.e. $\boldsymbol{X}^{'}=\left[\boldsymbol{x}_j\right]$ from $\boldsymbol{J}^T\boldsymbol{X}=\lambda\boldsymbol{X}$), because any subspace spanned by a subset of the eigenvectors is a J^T invariant (J^T denotes the transpose of the ODE model Jacobian Eq.(4)). As a consequence, $\left[\text{span}\{0\},\text{span}\{x_l\},\ldots\right]$ give a 1-dimensional lumping matrix *M*; [*span* $\{x_1, x_2\}$, span$\{x_2, x_3\}$,...] give a 2-dimensional lumping matrix *M* with rows formed with the *X*-columns, etc. Then, the lumping matrices of different dimensions can be simply formed by taking the columns of *X*, or any linear combination of them. In the exact lumping, the eigenvalues of the kinetic model Jacobian *J* remain invariant after applying transformations in species concentrations [Maria,2005b]:

$$\lambda\left(\boldsymbol{J}^T(\boldsymbol{C})\right)=\lambda\left(\boldsymbol{J}^T\left(\boldsymbol{M}^{+}\boldsymbol{MC}\right)\right).$$

Lumped models are particularly important when developed reduced kinetic models for complex cell structures, such as CCM, GERM, or various GRC-s (operon expression, genetic switches, genetic amplifiers, etc.)(Figure 1-56) [Maria,2017,2017a,2017b,2018].

Due to the **modular** functional organization of the cell, a worthy route to develop reduced models is to base the analysis on the concepts of '*reverse engineering*' and '*integrative understanding*' of the cell system (reviews of Maria [2006,2017b,2018]). Such a rule allows disassembling the whole system in parts (model functional 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

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

metabolic circuits. Such a procedure is very suitable for modelling GRC-s by linking GERM-s models in such a way to maintain the cell homeostasis, that is to maintain relatively invariant species concentrations despite perturbations. [Maria, 2005, 2017a,2017b,2018].

The math modelling efforts have intensified a lot after 2000 when the human genome has been deciphered (Figure 1-33), being proved that the difficult task to model and design complex biological circuits with a *building blocks strategy* can be accomplished by properly defining the cell basic components, functions, and structural organisation (Figure 1-2). Because many cell regulatory systems are organized as "*modules*" [Kholodenko et al., 2002], it is natural to model GRC-s and other metabolic processes belonging to the CCM by using a *modular approach* [Maria, 2017, 2017A, 2017B, 2018]. Further analyses including engineered GRC-s can lead to simulate and design of GMOs, of desirable characteristics, that is [Kaznessis et al., 2006]: a tight control of gene expression, i.e. low-expression in the absence of inducers and accelerated expression in the presence of specific external signals; a quick dynamic response and high sensitivity to specific inducers; GRC robustness, i.e. a low sensitivity *vs.* undesired inducers (external noise). Through the combination of induced motifs by the modified GRCs in the obtained GMOs one may create potent applications in industrial, environmental, and medical fields (e.g. biosensors, gene therapy). Valuable implementation tools of the design GRCs in real cells have been reported over the last years [Heinemann and Panke, 2006; Qian et al., 2017; Myers, 2009].

Systems Biology using *in silico* design of GMOs is closely related to the *Synthetic Biology*. The emergent field of *Synthetic Biology* [Benner and Sismour, 2005] "interpreted as the engineering-driven building of increasingly complex biological entities" [Heinemann and Panke, 2006], aims at applying engineering principles of systems design to biology with the idea to produce predictable and robust systems with novel functions in a broad area of applications [Heinemann and Panke, 2006; Voit, 2005]. By assembling functional parts of an existing cell, such as promoters, ribosome binding sites, coding sequences and terminators, protein domains, or by designing new GRC-s on a *modular* basis, it is possible to reconstitute an existing cell or to produce novel biological entities with new properties.

Encouraging results have been reported for the design of artificial gene networks for reprogramming signalling pathways, for refactoring of small genomes, or for re-design of metabolic fluxes with using switching genes (review Maria [2017a,2017b,2018]). By assembling functional parts of an existing cell, such as promoters, ribosome binding sites, coding sequences and terminators, protein domains, or by designing new gene regulatory networks on a modular basis, it is possible to reconstitute an existing cell (the so-called "integrative understanding"), or to produce novel biological entities with modified characteristics [Heinemann and Panke, 2006].

To help the efforts of the *Synthetic Biology* to *in-silico* design GMOs of desired characteristics, the emergent border field of the *Systems Biology* has been very quickly developed, based on using mathematical tools and numerical calculus, as well as (bio)chemical engineering (ChBRE) concepts and tools, together with the control theory of the nonlinear systems, and those of the *Bioinformatics* to characterize the kinetics and self-regulation of the cell metabolic processes. The close correspondence between *Systems Biology* and *Bioinformatics* is given in (Figure 1-57) [Hempel, 2006].

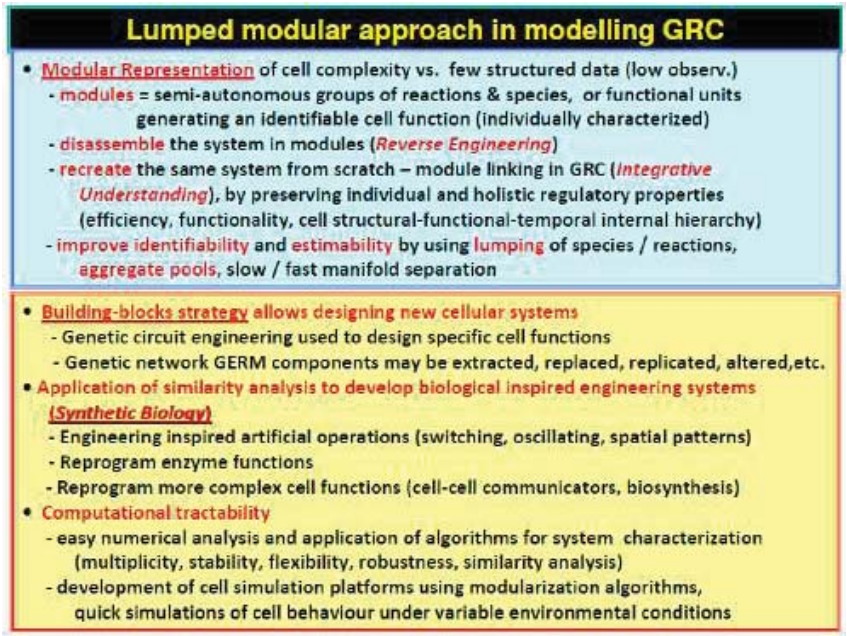


Figure 1-56: Importance of the lumped modular math modelling for the in-silico simulation of GERM–s and of GRC–s [Maria, 2017,2017a,2017b,2018].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

In the *Synthetic Biology*, the genetic components may be considered as "building blocks" because they may be extracted, replicated, altered, and spliced into the new biological organisms. The *Synthetic Biology* is in direct connection with the *Systems Biology* focus on the cell organization, the former being one of the main tools for the insilico design of GMO-s. In such a topics, the metabolism characterization by means of lumped but adequate cell models of the *Systems Biology* plays a central role.

Applications of such cell dynamic simulators, especially of GERMs chains (that is the GRCs) controlling the cell metabolism are immediate: design new micro-organisms of desirable characteristics; insilico re-programming the cell metabolism; design of biosensors; drug target release; industrial bioprocess optimization and control using GMOs; gene therapy; optimal cell cloning, etc. Consequently, the cell metabolism can be changed by modifying/designing GRCs, thus conferring new properties/functions to the mutant cells (i.e. desired 'motifs'), while engineered/synthetic gene circuits can be designed by using the Synthetic Biology tools [Alon, 2007].

Of course, the use of reduced metabolic kinetic models present a series of disadvantages, such as: loss in system flexibility (due to the reduced number of considered intermediates and species interactions); possibility to get multiple (rival) reduced models of proximate characteristics for the same cell system, difficult to be delimited; oss in the model prediction capabilities; lumped model parameters can lack of physical meaning; a loss / alteration of systemic / holistic properties (e.g. cell system stability, multiplicity, sensitivity, regulatory characteristics). Starting from an available extended kinetic model, classical ChBRE rules can be applied to reduce their structure [Maria,2006] aiming at:

(i) Reducing the list of species, by eliminating unimportant components and/or lumping some species (by using various measures, e.g. small values for the product of target species of index " i " lifetime $LT_i = -1/J_{ii}$ and its production rate r_i, where the Jacobian elements are $J_{ik} = \partial f_i(C,k)/\partial C_k$, where f_i are the right-side functions of the **ODE** kinetic model Eq.(1a), Eq.(4);

(ii) Reducing the list of reactions, by eliminating unimportant sidereactions and/or assuming quasi-equilibrium for some reaction steps (or using sensitivity measures of rate constants, such as ridge selection, principal component analysis, time-scale separation, etc. [Maria, 2004, 2019];

(iii) Decomposing the kinetics into fast and slow 'parts' allowing a separate study and application of the quasi-steady-state-approximation (QSSA) to reduce its dimensionality [Maria, 2006].

When the ODE kinetic model is linear in parameters, then the reduction procedure of Maria [2005b] can be applied by preserving the system Jacobian invariants (eigenvalues, eigenvectors).

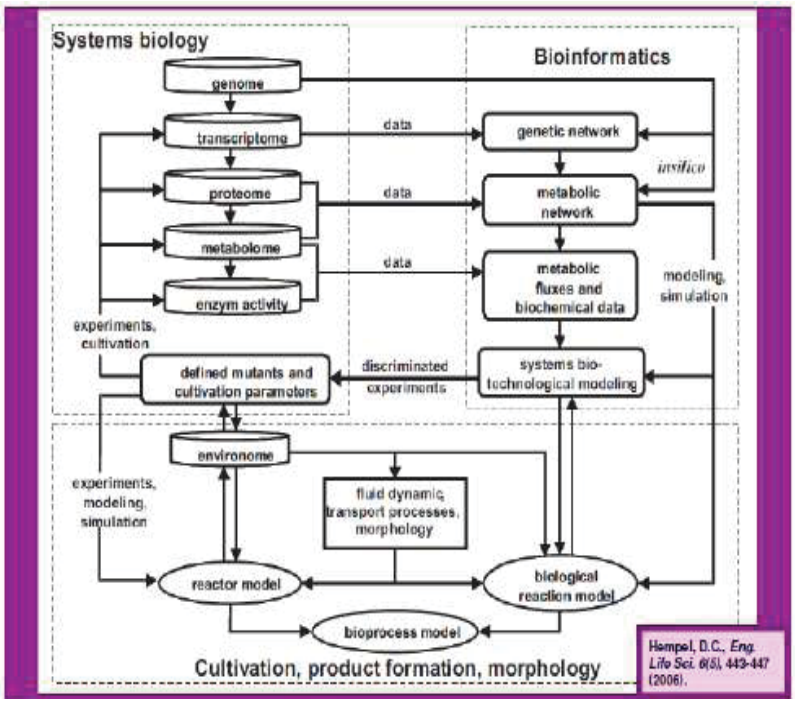


Figure 1-57: The close correspondence between *Systems Biology* and Bioinformatics given by [Hempel, 2006].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

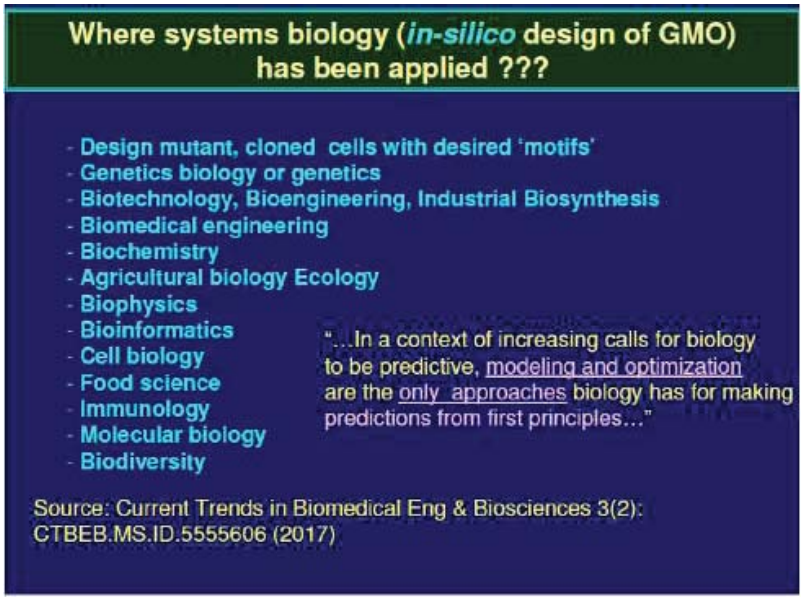


Figure 1-44: Area where Systems Biology has been applied. After [Nandy, 2017].

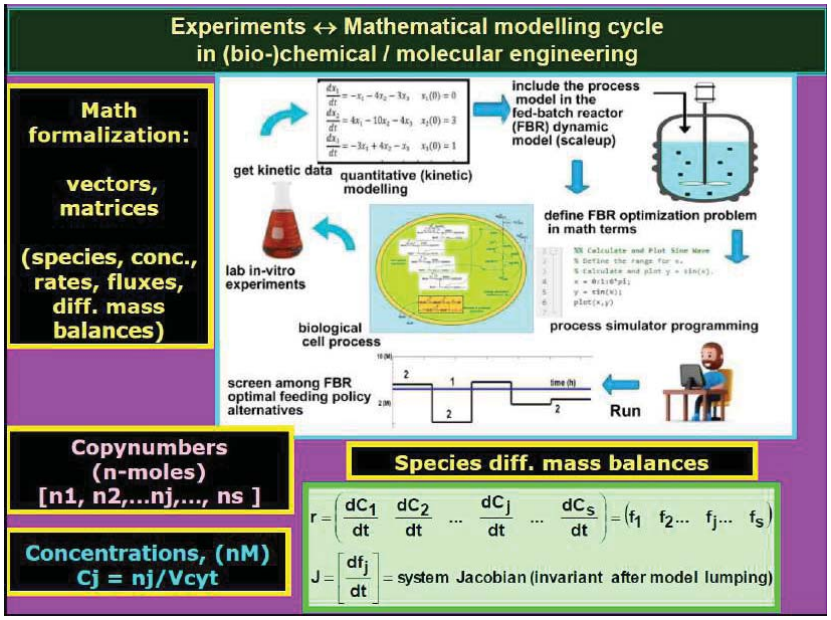


Figure 1-45: Experiment-modelling cycle to get math models in biochemical / metabolic engineering. Math formalization includes working with vectors and matrices. Notations: FBR = fed-batch bioreactor; nj,Cj = cell species "j" number of moles, or concentration, respectively; Vcyt = cell cytosol volume; "s" = no. of species; t = time [Maria,2017b,2018].

Due to the modular functional organization of the cell, a worthy route to develop reduced models is to base the analysis on the concepts of '*reverse engineering*' and '*integrative understanding*' of the cell system [Maria, 2006]. Such a rule allows disassembling the whole system in parts (modules), and then, by performing tests and suitable numerical / sensitivity analysis, to define rules that allow to re-create the whole and its characteristics 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 genetic regulatory circuits (GRCs) by linking gene expression regulatory modules (GERMs) in such a way to maintain the cell homeostasis, that is to maintain relatively invariant species concentrations despite perturbations. [Maria, 2005, 2006, 2007, 2009]; [Maria and Luta,2013].

A potential application of lumped modular GRC models is the socalled 'genetic circuit engineering', by which simulation of gene expression is used to *in-silico* design GMOs that possess specific and desired functions. By inserting new GRCs into micro-organisms,

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

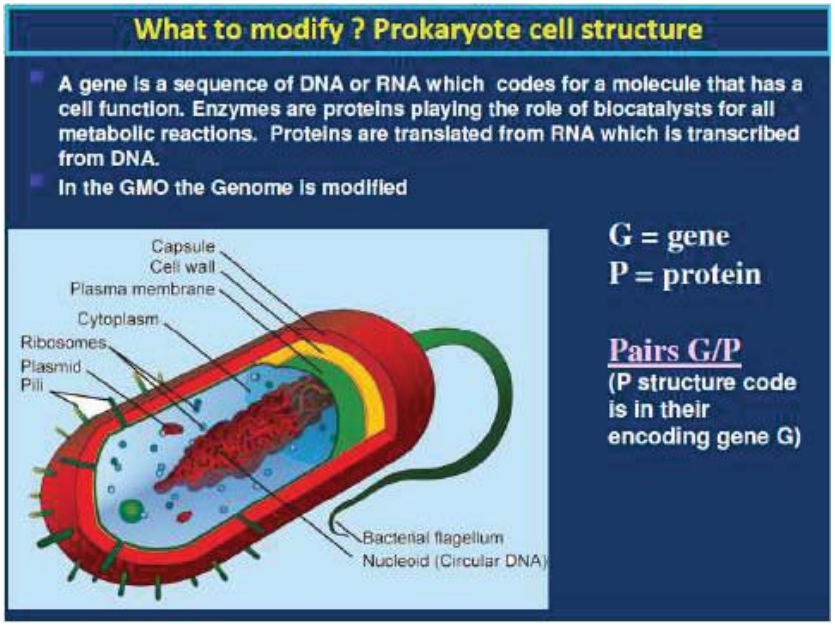


Figure 1-49: The prokaryote cell structure.(source= https://en.wikipedia.org/wiki/File:Averag e_prokaryote_cell-_en.svg).

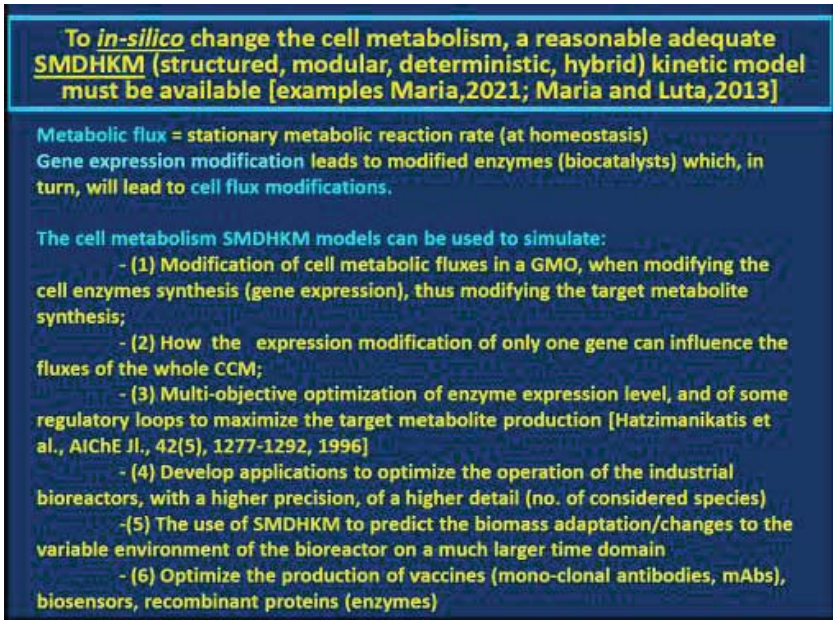


Figure 1-50: Some of the problems to be solved when applying in-silico design of GMOs [Maria,2023].

one may create a large variety of mini-functions / tasks (or desired 'motifs') in response to external stimuli. The induced functions in gene circuits are diverse, such as: switches (decision-making branch points between on/off states according to the presence of inducers), oscillators (cell systems evolving among two or several quasi-steady-states), signal / external stimuli amplifiers, amplitude filters, genetic 'memory' storage. The genetic components may be considered as 'building blocks' because they may be extracted, replicated, altered, and spliced into new biological organisms. Combination of induced motifs in modified cells one may create potent applications in industrial and medical fields, e.g. the production of biosensors used in medicine or environmental engineering applications. Design of modular GRCs must account for some properties (see sections 2.2.1, 2.2.2, 2.2.3): a tight control of gene expression (i.e. low-expression in the absence of inducers and accelerated expression in the presence of specific external signals); a quick dynamic response and high sensitivity to specific inducers [Maria, 2006].

Application of the (ChBRE) concepts, tools, and numerical algorithms to develop deterministic kinetic models of cell processes, with continuously variables, is particularly suitable for several reasons, as followings:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

a). Metabolic processes at a low(molecular)-level are generally better clarified. Based on that, conventional dynamic models, using ordinary differential species mass balances (ODE), with a *mechanistic (deterministic) description* of reactions tacking place among individual species (proteins, mRNA, intermediates, etc.) have been proved to be a convenient route to analyse continuous metabolic / regulatory processes and perturbations.

b). When systems are too large or poorly understood, coarser and more phenomenological ODE kinetic models may be postulated, using lumped reactions/species (e.g. protein complexes, metabolite channelling, etc.). In dynamic deterministic models, usually only essential reactions are retained, the model complexity depending on the measurable variables and available information. To reduce the structure of such a model, an important problem to be considered is the distinction between the qualitative and quantitative process knowledge, stability and instability of involved species, the dominant fast and slow modes of process dynamics, reaction time constants, macroscopic and microscopic observable elements of the state vector.

c). Such deterministic ODE kinetic models can be useful to analyse the regulatory cell-functions, both for stationary and dynamic perturbations, to model cell cycles and oscillatory metabolic paths [Maria, 2014a; Maria et al.,2018a], and to reflect the species interconnectivity or perturbation effects on cell growth [Maria, 2017a, 2017b, 2018]. Similarly to the models used in ChBRE, the representation of metabolic process kinetics is made by using usual rate expressions of enzymatic processes [Segel,1993], such as those of the extended Michaelis-Menten, or Hill type [Miskovic et al., 2015; Maria, 2014b, 2014a] (Figure 1-55). See also the dynamic models of CCM [Miskovic, 2015; Kurata and Sugimoto, 2017; Chassagnole et al., 2002].

The parameters of the deterministic models (that is their rate constants) are estimated by using the common ChBRE rules of Maria [2004], by using either dynamic (kinetic) data obtained in a chemostat under transient regime (e.g. pulse-like perturbations in the bioreactor influent, [Chassagnole et al., 2002], or using steady-state data (metabolites concentrations) obtained at homeostasis (that is under a balanced cell growth, [Maria, 2017a,2017b,2018]), by solving the model stationary algebraic set $dCj/dt\text{-}o$ (where "j" indexes the no. of cell species taken individually or lumped) in Eq.(1c). Parameter estimate must fulfill physico-chemical meaning constraints Eq.(3) (related to metabolic reaction stoichiometry; rate constants must be limited by the diffusional processes, and in agreement with the thermodynamic equilibrium steps [Stephanopoulos et al., 1998]). Additionally, due to the optimized metabolic cell process, the GERM, GRC of the WCVV cell models must fulfil some optimality constraints, that is [Maria and Scoban, 2017, 2018; Maria,2017a,2017b,2018]:

a). reaction rates must be maximal, but with rate constants limited by the diffusional processes;

b). the total enzymes (proteins) content of the cell is limited by the isotonicity condition [Heinrich and Schuster, 1996];

c). the total cell energy (ATP) and reducing agent (NADH) resources are limited;

d). the reaction intermediate level must be minimum;

e). the cell model at homeostasis must be stable, that is will reaching again the steady-state after termination of a perturbation [Heinrich and Schuster, 1996]. This condition can be check by evaluating the eigenvalues (λj) of the ODE model Jacobian Eq.(4); thus the stability condition imposes that the Real part of the Max (λj) to be negative, that is $\mathrm{Re}(\mathrm{Max}(\lambda j)) < 0$;

f). the key-species concentrations must be constant at homeostasis. Most of the mentioned aspects are discussed and exemplified by Maria [2017a,2017b,2018], and by Maria and Scoban [2017,2018].

Apart from the common estimation criteria used by the ChBRE to estimate the kinetic model parameters [Maria, 2004], in the case of GERM / GRC kinetic models of WCVV type, supplementary constraints must be applied to determine some rate constants, by imposing optimum regulatory criteria for GERM/GRC-s, such as:

i). the minimum recovering time of the stationary concentrations (homeostasis) after a dynamic ('impulse'-like) perturbation in a keyspecies [Maria,2005,2017a,2017b,2018], [Maria and Scoban, 2017,2018];

ii). a quick action of the buffering reactions to obtain the fastest recovering time of the steady-state (homeostasis) after an external / internal perturbation;

iii). smallest sensitivity of the key-species homeostatic levels *vs.* external perturbations in the nutrient levels; iv).-- highest homeostasis stability strength, etc. [Maria, 2003, 2005, 2006, 2007, 2009, 2009b, 2010,2014b, 2017a,2017b,2018], Maria and Luta [2013].

Besides, effective solvers must be used for rate constants estimation [Maria, 2004]. In all cases, the estimation rule must be based on the fulfilment of the stationary condition imposed to the ODE cell model ($dCj/dt = 0$), that is Eq.(1c), by preserving the system invariants of the mass balance equations, Eq.(3) [Maria, 2004, 2005], and by imposing optimum regulatory criteria related

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

to GRC efficiency to cope with continuous perturbations, formulated by Maria [2017a, 2017b,2018], [Maria and Scoban, 2017,2018]. In the GRC model cases, also supplementary constraints have to be considered related to the WCVV modelling approach (that is the isotonicity constraints, see section 2.2.1, 2.2.2), and constraints related to the GERM / GRC stationary/dynamic regulatory efficiency [Maria, 2017a; Maria and Scoban,2017,2018].

Even if complicated and, often over-parameterized, the continuous variable dynamic deterministic ODE models of the CCM metabolic pathways, and of GRC-s present a significant number of advantages, being able to reproduce in detail the molecular interactions, the cell slow or fast continuous response to exo/endo-geneous continuous perturbations. [Maria, 2005; Styczynski and Stephanopoulos, 2005]. Besides, the use of ODE kinetic models presents the advantage of being computationally tractable, flexible, easily expandable, and suitable to be characterized using the tools of the nonlinear system theory [Banga, 2008; Heinrich and Schuster, 1996], by accounting for the regulatory system properties, that is: dynamics, feedback / feedforward, and optimality. And, most important, such an ODE kinetic modelling approach allows using the strong tools of the classical ChBRE modelling, that is Maria [2017,2017b,2018], Maria et al. [2022]. (Figure 1-5,6):

i) molecular species conservation law (stoichiometry analysis; species differential mass balance set);

ii) atomic species conservation law (atomic species mass balance);

iii) thermodynamic analysis of reactions (that is quantitative assignment of reaction directionality), [Haraldsdottir et al., 2012];

iv) set equilibrium reactions; Gibbs free energy balance analysis; set cyclic reactions; find species at quasi-steady-state; improved evaluation of steady-state flux distributions that provide important information for metabolic engineering [Zhu et al., 2013], allowing application of ODE model species and/or reaction lumping rules [Maria, 2004,2019].

To overcome some of the above strong limitations of the classic WCCV dynamic models, and to ensure the fulfillment of some holistic cell properties (such as homeostasis, self-regulation of syntheses, and of gene expression self regulation, perturbation treatment, etc.), some extensions chose to impose a lot of constraints (such as "the total enzyme activity" and "total enzyme concentration", etc., [Komasilovs et al., 2017; Heinrich and Schuster, 1996]).

As revealed by this very brief survey, general ChBRE modelling principles are proved to be valuable tools for representing the both stationary and dynamic characteristics of complex cell biochemical processes. Elaboration of reduced models of a satisfactory quality is closely related to the ability of selecting the suitable lumping rules, keyparameters, and influential terms, and to apply multi-objective non-/conventional estimation criteria that realize the best trade-off between model simplicity and its predictive quality [Maria,2017a,2017b,2018]; [Maria and Scoban,2017,2018].

Note. Part of this work was the subject of the following invited plenary lectures:

- Maria, G., 2013, Applications of chemical engineering principles and the lumping analysis in modelling the living cell systems - A trade-off between simplicity and model quality, presented at University BabesBolyai, Cluj (Romania), Department of Chemistry, Nov. 8, 2013. http://www.chem.ubbcluj.ro/~chimie/anunturi.html,

- Maria, G., Applications of chemical engineering principles and the lumping analysis in modelling the living cell systems -A trade-off between simplicity and model quality, Symposium SICHEM-2016, University Politehnica of Bucharest, Romania, 8 Sept. 2016.

2.2. A novel concept of «mechanistic silicon cell» materialized in a novel math (kinetic) modelling framework «WCVV» of the cell metabolic processes (whole-cell, isotonic variable-volume) that maintain intracellular homeostasis while growing auto-catalytically 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 1-2), that exhibit remarkable behaviors including, most predominately, their selfreplication. 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 [Palsson, 2000; Tomita, 2001; Tomita et al., 1999; Kitano, 2002, 2002b; Mori, 2004; You, 2004].

"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" [Ideker et al., 2001; Jahan et al., 2016; Lindon et al., 2004; Rao et al.,2005]. [Surovstev et al.,2007].

Much remains before useful comprehensive mathematical models of cellular behaviour with predictive power can be realized. Obviously, the large number of unknown kinetic parameters of such extended cell kinetic models (rate constants and stationary concentrations of 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 (review of Maria [2017a,2017b,2018,2023]).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

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 models, regardless of the particular set of (bio)chemical reactions and assumed components (individual or lumped species.

The **"WCVV - mechanistic silicon cell"** novel math modelling framework of cell processes, described in this work (section 2.2) is an outstanding and essential/fundamental contribution of Maria [2002], and [Maria et al., 2002], while further developments/ studies of the WCVV properties were given by Maria over the interval [2005-2023], that is Maria [2002,2003,2005,2006,2007,2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c].

As discussed by Maria [2017a, 2017b, 2018, 2023], 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 1-2), involving $O(10^{3-4})$ components, $O(10^{3-4})$ transcription factors (TFs), 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 selfchanging its genome/proteome and metabolism to obtain an optimized and balanced growth with using minimum of resources (nutrients/substrates).

In this section 2.2, one describes the novel cell-modeling framework "mechanistic whole silicon cell", materialized in the novel "WCVV" whole-cell modelling framework of cell processes, implicitly accounting for the isotonic growing cell (of variable-volume). In this novel 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). When appropriate rate constants of the constructed cell kinetic model are obtained from the known stationary (homeostatic) concentrations of the all species included in the kinetic model (taken individually or lumped), the self-replicating behavior of cells can be modelled. Beside, we illustrate the large number of applications, features, and advantages of the novel WCVV modelling framework of the cell kinetic processes. Exemplifications will be made by using simplified kinetic models of GERMs/ GRCs with -omics data of *E. coli* (from Ecocyc, and KEGG), as being the essential modules aiming to regulate the cell syntheses, and to adapt the cell metabolism to the environmental changes.

As briefly presented in (Figure 1-2), cells have a hierarchic organization (structural, functional, and temporal), which is a characteristic of the living matter in general, as followings:

1). the ***structural hierarchy*** includes all cell components from simple molecules (nutrients, saccharides, fatty acids, aminoacids, simple metabolites), then macromolecules or complex molecules (lipids, proteins, nucleotides, peptidoglycans, coenzymes, fragments of proteins, nucleosides, nucleic acids, intermediates), and continuing with wellorganized nano-structures (membranes, ribosomes, genome, operons, energy harnessing apparatus, replisome, partitioning apparatus, Z-ring, etc. [Lodish et al., 2000]). To ensure self-replication of such a complex structure through enzymatic metabolic reactions using nutrients (Nut), metabolites (Met), and substrates (glucose/fructose, N-source, dissolved oxygen, and micro-elements), all the cell components should be associated with specific functions into the cell, as following with:

2). ***functional hierarchy*** according to the species structure; e.g. source of energy, and intermediates. Sauro and Kholodenko [2004] 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, individual gene expression regulation, etc.). More complex functions, as regulatory networks (GRC / GRN), synthesis networks, or metabolic cycles can be built-up from basic ***building blocks***. Such a building blocks structure is very tractable when developing cell reduced dynamic models by defining various metabolic sub-processes, such as: regulatory functions for the enzymatic reactions, energy balance functions for ATP/ADP/AMP renewable system, electron donor systems of the NADH, NADPH, FADH, FADH2 renewable components, 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, harness cell energy), functions for cell cycle regulation (nucleotide replication and partitioning, cell division);

3). *time (temporal) hierarchy*. The wide-separation of time constants of the metabolic reactions in the cell systems is called *time hierarchy*. Thus, the reactions are separated in slow and fast according to their time constant (see the definition given by Maria et al. [2010]). In fact, only fast and slow reactions are of interest, while the very slow processes are neglected or treated as parameters (such as the external nutrient or metabolite evolution). Aggregate pools (that is lumps combining fast reactions) are usually used in building-up cell dynamic models in a way that intermediates are produced in a minimum quantity and consumed only by irreversible reactions. All cell processes obey a certain succession of events, while stationary or dynamic perturbations are treated by maintaining the cell components homeostasis, the recovering or transition times after perturbations being minimal.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

To model such an astronomically complex cell system with a detailed kinetic model is practically impossible even if expandable biomolecular data are continuously added in-omics databanks (e.g. KEGG, EcoCyc, Prodoric, Brenda, CRGM, NIH [Maria,2007,2018]). However, as underlined by Tomita et al. [1999], and Tomita [2001], the "whole-cell (WC) simulation of metabolic processes with mechanistic / deterministic kinetic models of continuous variables, represents "the grand challenge of the 21st century". Such an huge effort is justified by the very large number of immediate applications: design genetically modified microorganisms (GMO) with desirable characteristics to be used in industry (new biotechnological processes, production of vaccines). A large number of GMO applications are in the industrial synthesis, but also in medicine, such as therapy of diseases (gene therapy), or in the novel technologies, such as design of molecular-level devices based on cell-cell communicators, biosensors, etc.

As underlined by Tomita [2001] "Computer models and *in silico* experiments are necessary to understand and predict phenotypes of the cell, especially when they are polygenic phenotypes. After all, most biological and pathological phenomena in which the pharmaceutical industry has a great interest, such as cancer and allergy, are polygenic."

Comparatively to chemical systems, and lacking of enough and reproducible structured experimental data, the cell biochemical processes present the advantage of being approached in a modular way. Almost every CCM subsystem can be separately studied/modeled [Tomita,2001; Maria,2005, 2006, 2014a, 2014b], that is: the phosphotransferase (PTS) system for glucose transport into the cell, the pentose- phosphate pathway (PPP) for nucleotides and amino-acids production, the tricarboxylic acid (TCA) cycle, glycolysis, amino-acids synthesis [Maria,2014a, 2014b, 2021]; regulation of gene expression, cell cycle, signal transduction, and various metabolic pathways [Tomita,2001; Maria,2017a,2017b,2018,2023]. "However, although these models made significant contributions to the development of *in silico* biology, the programs were able to handle only specific subsystems, and it was difficult to combine different subsystem models into one single-cell model." [Tomita,2001]. Encouraging results have been reported for design artificial gene networks (GRCs, like genetic switches), for reprogramming signaling pathways, for refactoring of small genomes, or re-design of metabolic fluxes with using switching genes. Remarcable progresses have been reported by assembling functional parts of an existing cell, or by designing new gene regulatory networks (GRCs) on a modular basis. In such a way, it is possible to reconstitute an existing cell (the so-called 'integrative understanding') or to produce novel biological entities with new properties. The genetic components may be considered as '***building blocks***' because they may be extracted, replicated, altered, and spliced into the new biological organisms [Maria,2005].

In the whole-cell approach, due to the mentioned insuperable detailed modelling difficulties, various cell modelling alternatives have been developed over decades [Maria,2005,2017a,2017b,2018], reviewed in section 1.3.4. The ODE deterministic whole-cells models have been developed in two alternatives, [Maria et al., 2018d, 2018g], that is:

i). The classical (default) Whole-cell Constant Volume (WCCV) kinetic models [Maria, 2017a,2017b,2018,2023; Aris, 1969; Froment and Bischoff, 1990; Levenspiel, 1999]. (see section 2.1).

ii). The novel Whole-cell Variable Volume (WCVV) of the cell metabolic processes (whole-cell, isotonic variable-volume) that maintain intracellular homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. This WCVV novel modelling framework was introduced by Maria [2002], and by Maria et al. [2002] [2002,2003,2005,2006,2007,2009, 2014b,2017,]. The study of the WCVV properties and advantages was given by Maria over the interval [2002-2023], that is Maria [2017A,2017B,2018,2023, 2023a, 2024, 2024b, 2024c] (see the section 2.2.1).

As an alternative, Maria [2002,2005,2006,2007,2009, 2009b,2010,2017a,2017b,2018,2023]; Maria et al. [2002]; Maria and Luta [2013]; Maria [2014a,2014b]; Maria and Scoban [2017,2018]; Maria et al.[2018d] proposed and studied the holistic silicon whole-cell variable volume (WCVV) modelling framework by explicitly including in the model constraint equations accounting for the cell-volume continuous growth over a cell cycle, by preserving the same cell-osmotic pressure (section 2.2).

The ***"WCVV - mechanistic silicon cell"*** novel math modelling framework of cell processes, described in this work is an outstanding and essential/fundamental contribution of Maria [2002], and [Maria et al., 2002], while further developments/ studies of the WCVV properties were given by Maria over the interval [2005-2023], that is Maria [2002, 2003, 2005, 2006, 2007 2009, 2014b, 2017A, 2017B, 2018 2023, 2023a, 2024, 2024b, 2024c].

Maria [2017a,2017b,2018] also proved the advantages of such a model formulation (sections 2.2.3, 2.3. and the below section "Some advantages of the novel WCVV modelling framework" included in the section 2.2.1).

By contrast, the WCVV formulation promoted by (Maria, 2017 reviews), remove these limitations, by including the thermodynamic isotonicity relationships/ constraints. Thus, it has been proved, step-bystep, in a math way (by using simple lumped generic GRC models) how such constraints ensure cellular intrinsic properties in a natural way (that is not derived from artificial hypotheses). Such concepts, and rules translated from (bio)chemical engineering principles and nonlinear systems theory are explained, proved, and exemplified over the mentioned papers of Maria [2005,2006,2007,2009,2010,2014,2014b; Maria and Luta,2013].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

2.2.1. The hypotheses, concepts, and principles of the novel WCVV modelling framework

This section reviews the general concepts and hypotheses of the WCVV modelling framework of cell processes, as proposed by Maria [2002], and by Maria et al. [2002], with briefly reviewing the further developments/ studies of the WCVV properties given by Maria over the interval [2005-2023], that is Maria [2002,2003,2005,2006,2007,2009,2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]. Exemplifications are made with modelling and simulation of individual GERM regulatory properties targeted to maintain the intracellular homeostasis while growing autocatalytically on environmental nutrients present in variable amounts.

Such novel explorations, with using WCVV kinetic models, briefly reviewed in this work, could provide insight into the mechanisms of healthy and diseased cells, as well as a better understanding of how system-level or whole-cell properties emerge from intracellular interactions of molecular components (Lindon et al., 2004).

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 [Aach et al., 2000; Covert et al. 2004). 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 [Maria, 2017a,2017b,2018,2023].

Excellent progress has been made in developing computational approaches for modeling whole-cell processes [Mendes, 1997; Tomita et al., 1999; Schaff et al., 2001; Hucka et al., 2003; Weitzke and Ortoleva, 2003; Takahashi et al., 2004; Sewell et al.,2002; Schuster et al.,2000; Papin et al.,2003; Yang et al., 2003]. However, the constant volume assumptions [see Eq.(1a-1c)]made by these applications 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. [2018d, 2018g], the classic ("default") math modelling approach of cell processes [that is the WCCV, "constant-cell-volume, non-isotonic system", Eq.(1a-c)] *is erroneous* compared to the novel WCVV modelling framework, leading to *distorted and wrong simulation results* and conclusions as proved in section 2.3.1 and by [Maria et al., 2018d; Maria et al.,2002]. In such a way, the novel WCVV kinetic cell modelling approach replaces the CLASSICAL (default) WCCV modelling approach, thus *correcting the distorted* and false/wrong predictions of these cell kinetic models, unfortunately largely used in the literature (sections 2.1, and 2.3.1).

To overcome this major drawback of the classical (default) kinetic models of a WCCV type, when applied to model cell processes, some awkwards trials have been reported. For instance, Perret and Levey [1961] 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. Grainger and Bass [1966], Grainger et al. [1968], Aris [1969], Kacser and Beeby [1984, Brumen et al. [1984], Werner and Heinrich [1985], and Joshi and Palsson [1989a,b; 1990a,b] analyzed similar situations.

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 [Surovstev et al.,2007]. Thereafter, they shift into a division phase in which they ultimately divide into two newborn daughter cells, each equivalent to the original newborn cell in terms of volume and cellular content. These two phases, growth and division, together comprise the cell cycle. The main hypotheses of the WCVV modelling framework are presented in Table 2-1. Details on these hypotheses introduced by Maria [2002], and [Maria et al., 2002], and further developments/ studies of the WCVV properties deriving from these hypotheses were given by Maria over the interval [2005-2023], that is Maria [2002,2003,2005,2006,2007,2009,2014b, 2017A, 2017B, 2018, 2023a, 2024, 2024b, 2024c].

Basically, the WCVV modelling framework presented in the Table 2-1 are given below [Maria,2002,2005]; [Maria et al.,2002].

(1). 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 1-2). The cell is separated from the *environment* (containing *nutrients*) by a *membrane* (Figure 1-4-TOP).

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

(3). The cell is an isothermal system with an uniform content (perfectlymixed 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 [Moser, 1988; Froment and Bischoff, 1990].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

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

(5). To better reproduce the GERM properties (inter-connected with the rest of the cell (Figure 1-4-TOP, Figure 2-7-19), 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 [Maria, 2002,2005,2017a,2017b,2018]; [Maria et al.,2002; Maria and Scoban, 2017,2018; Wallwork and Grant, 1977]:

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, combined with Eq.(2,13), leads to: $$\frac{RT}{\pi} = \frac{V(t)}{\sum_{j=1}^{ns} N_j(t)} = \frac{1}{\sum_{j=1}^{ns} C_j} = \frac{1}{\sum_{j=1}^{ns} C_{j,0}} = constant \quad (6)$$

Where: T = absolute temperature; R = universal gas constant, V = cell (cytosol) volume; $\pi_{cyt} = \pi =$ inner osmotic pressure; t = time; N_j = species j number of moles. Maria and Luta [2013], and Maria [2014b] proposed that lumped genome and proteome replication to also be considered in such cell models. Fulfilment of this Eq.(6) requirement of the WCVV models, implies that all the cell species to be included in the cell model (individually, or lumped). See for instance the lumped genome / proteome used by [Maria and Luta,2013] and by Maria [2014b] in their simulations, in order to "mimic" the cell content regulatory properties. As a very important observation, according to the iso-osmotic constraint Eq.(6) of WCVV kinetic models, the cell model (CCM, GERM, or GRC) should include the all species of the cell, taken individually, or lumped.

(6) 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.

(7) When studying an individual P -synthesis regulatory module (the socalled "gene expression regulatory module" - GERM), the other cell species are lumped together in the so-called "cell ballast" [Maria,2005,2006,2007,2014b].

(8) The inner cell osmotic pressure (π_{cyt}) is constant, and all time is equal with the environmental pressure, thus ensuring the membrane integrity ($\pi_{cyt} = \pi_{env} = \text{constant}$). As a consequence, the isotonic osmolarity under isothermal conditions leads to the equality $RT / \pi_{cyt} = RT / \pi_{env}$, derived from the fulfilment of Eq.(6)(Pfeiffers'law):

$$\pi_{cyt} / RT = \sum_{j}^{all,cell} \frac{N_j}{V_{cell}} = \pi_{env} / RT = \sum_{j}^{all,env} \frac{N_j}{V_{env}};$$

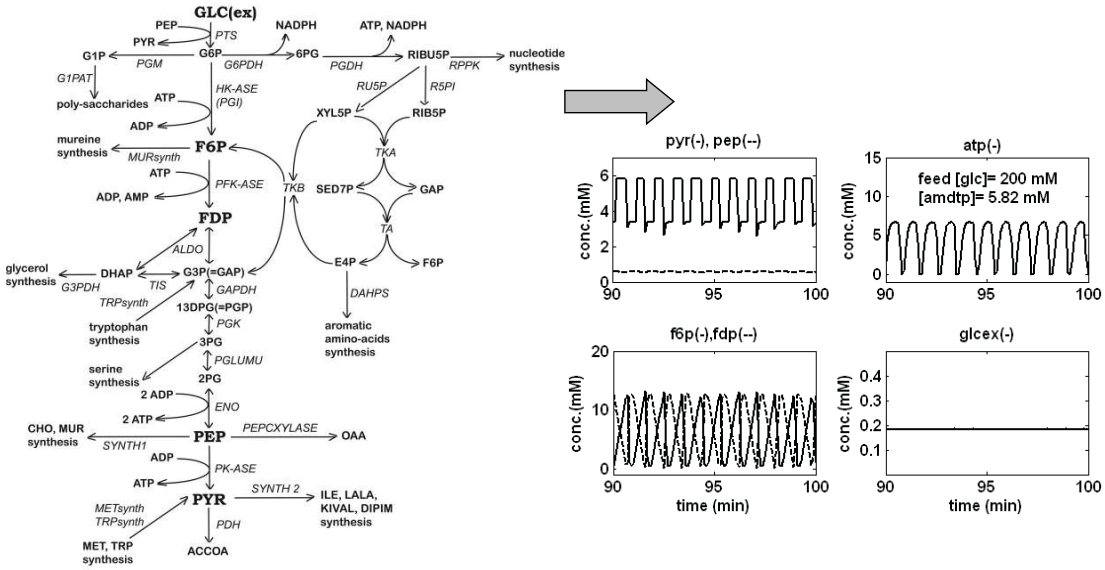


Figure 2-5: The structured reduced model of Maria [2014a] used to simulate the oscillating glycolysis in *Escherichia coli.*

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

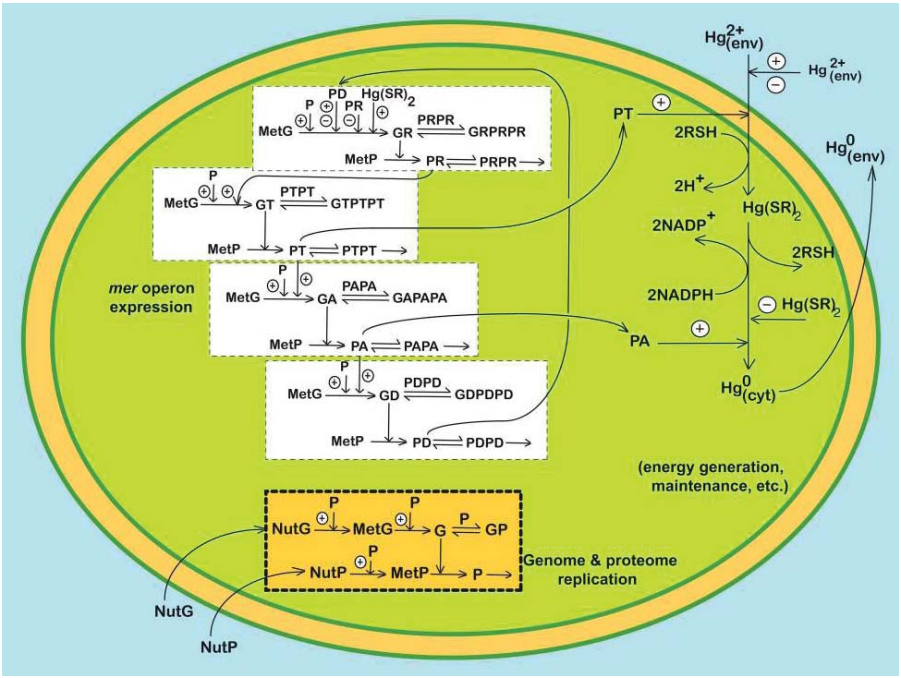


Figure 2-6: Structured cell model of Maria [2010]; [Maria and Luta, 2013] used to simulate the dynamics of mer-operon expression for mercury uptake in *Escherichia coli.*

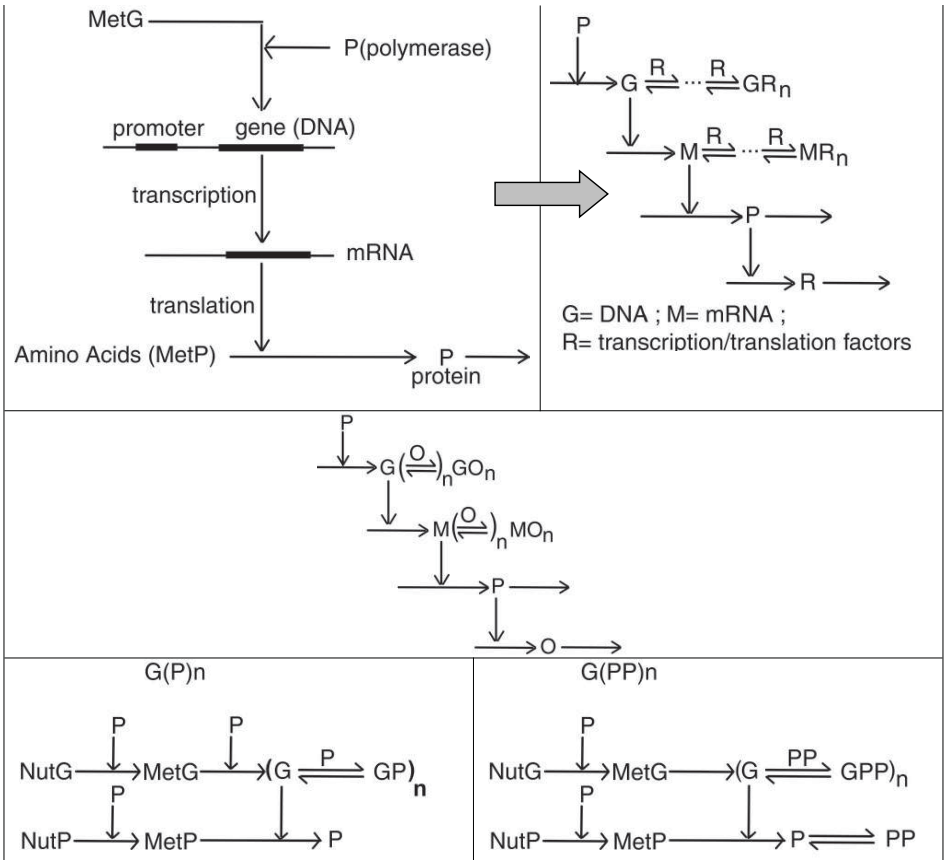


Figure 2-7: Protein P synthesis. The simplified representation of a generic expression regulatory module (GERM) for a generic gene (G) and protein (P). The horizontal arrows indicate reactions. The vertical arrows indicate catalytic actions. The absence of a substrate or product indicates an assumed concentration invariance of these species. Notations: G = gene encoding P; M = mRNA. The top-right structure corresponds to a GERM module of type [G(R)n ; M(R)n]. **Up-row: [Left]** Simplified representation of the two-steps gene expression: transcription and translation. Up-row [Right]. A simplified reaction schema of the gene expression [Maria,2003,2005,2006,2017a,2018]. According to the nomenclature of [Maria,2003,2005,2006,2017a,2018], such a model corresponds to a [G(R)n;M(R)n] GERM model type. [Center-row] . [Maria, 2005]. The simplified protein P synthesis, by using a GERM self-regulated expression module of type [G(O)n; M(O)n] ions; G= DNA gene encoding P; M = mRNA; O = allosteric effectors. **[Down-row]**. [Maria,2005,2006,2009,2017a,2018] two types of GERM simplified representations for protein synthesis, of different types, that is of [G(P)n] (left), and of [G(PP)n] (right).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

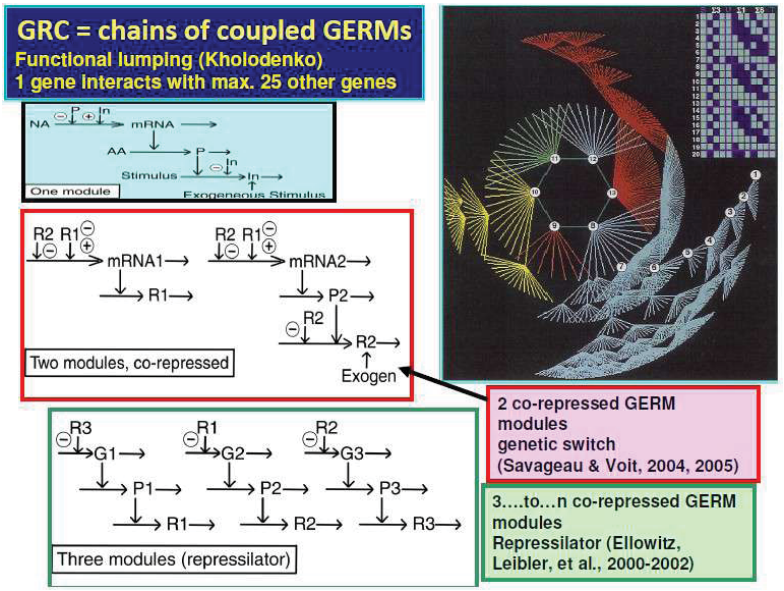


Figure 2-8: Example of linked GERM–s to form GRC–s (review of Maria [2003,2005,2017a,2017b,2018]).

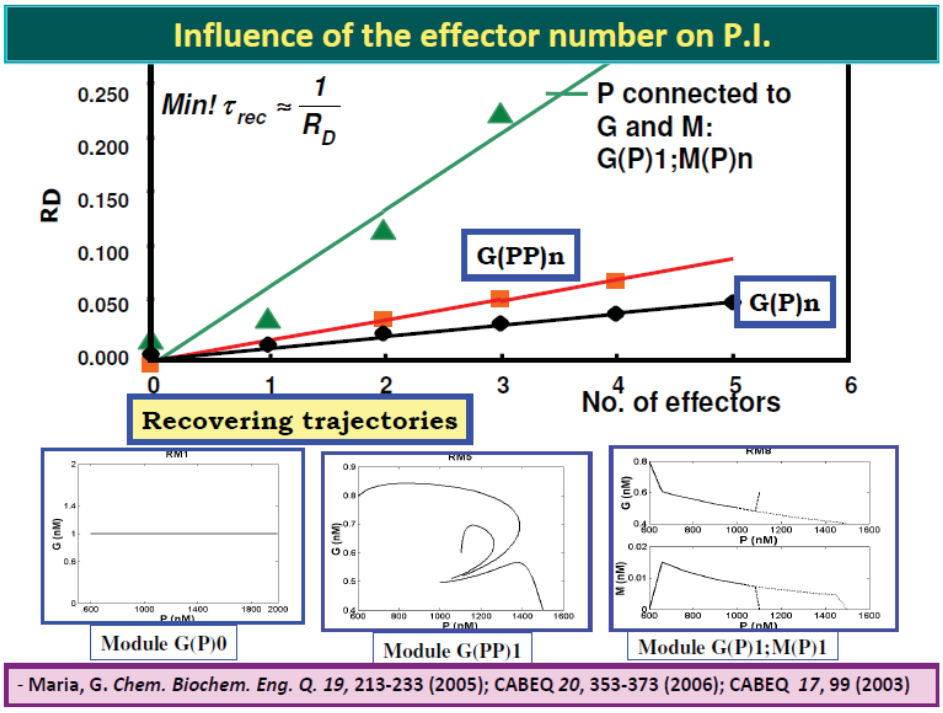


Figure 2-9: [Top]. Influence of the GERM number of effectors denoted by "n" on their properties. Adapted from [Maria, 2003, 2005, 2006, 2017a, 2018]; [Yang et al., 2003]. The approached cases concerns GERM regulatory modules of the following types: [G(PP)n], [G(P)n], [G(P)1;M(P)n] (see Fig. 2-12). RD = recovering rate 22 (Table 2-2) necessary to each GERM component to return to their stationary concentration (QSS) after an impulse-like perturbation in one component of the cell WCVV kinetic model. Here, the impulse-like dynamic perturbation corresponds to a quick decrease of the stationary key-protein [P]s with -10%. The recovering time of the stationary [P]s with a 1% tolerance is denoted by τ (rec). [Down]. The recovery trajectories in the [G-P], or [M-P] phases, after the above mentioned [P]s perturbation for the GERM–s mentioned under the abscissa. Adapted from [Maria, 2003].

Thus, resulting: $\pi_{cyt} = \pi_{cyt}$, and:

$$\left(\sum_{j}^{all} C_j\right)_{cyt} = \left(\sum_{j}^{all} C_j\right)_{env} \quad (7)$$

Otherwise, the osmosis will eventually lead to an equal osmotic pressure $\pi_{cyt} = \pi_{env}$. Even if, in a real cell, such equality is

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

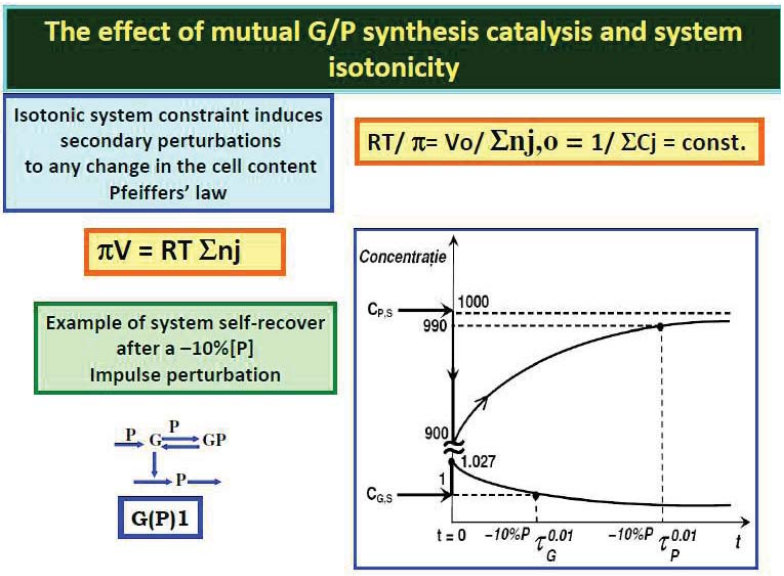


Figure 2-10: The effect of the mutual G/P catalysis and of the isotonicity in the case of a simple gene expression regulatory module (GERM) of [G(P)1] type (see Fig. 1-4-down). For details see [Maria, 2005,2006,2017a,2017b,2018]; [Maria and Scoban, 2017,2018], and section 2.2.4. The impulse-like dynamic perturbation corresponds to a quick decrease of the stationary [P]s with -10%. The recovering time of the stationary [P]s with a 1% tolerance is denoted by τ (P,0.01). The recovering time of the stationary encoding gene [G]s with a 1% tolerance is denoted by τ (P,0.01).

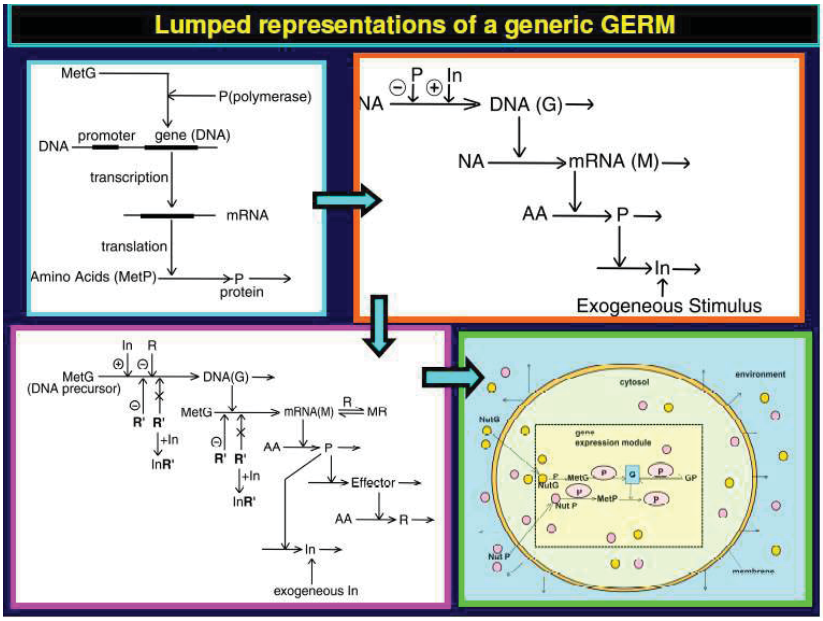


Figure 2-11: Various simplified representations of gene expression regulatory modules (GERM-s) according to Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018]. **[Down-right]**. Simplified reaction scheme of a generic gene G expression, by using a regulatory module of [G(P)1] type. The model was used to exemplify the synthesis of a generic P protein in the E. coli cell by (Maria, 2005). To improve the system homeostasis stability, that is quasi-invariance of key species concentrations (enzymes, proteins, metabolites), despite of perturbations in nutrients Nut*, and metabolites Met*, or of internal cell changes, a very rapid buffering reaction G(active) + P <===> GP(inactive) has been added. Horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; G = active part of the gene encoding protein P; GP = inactive part of the gene encoding protein P; MetG, MetP = lumped DNA and protein precursor metabolites, respectively.

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 are then used to import increasing amounts of nutrients. The net result is an exponential increase of cellular components in time, which translates, through isotonic osmolarity assumption, into an exponential increase in volume with time, Eq.(6) [Maria,2005,2017a,2018]; [Maria and Scoban, 2017,2018]. The overall concentration of cellular components is time-invariant

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

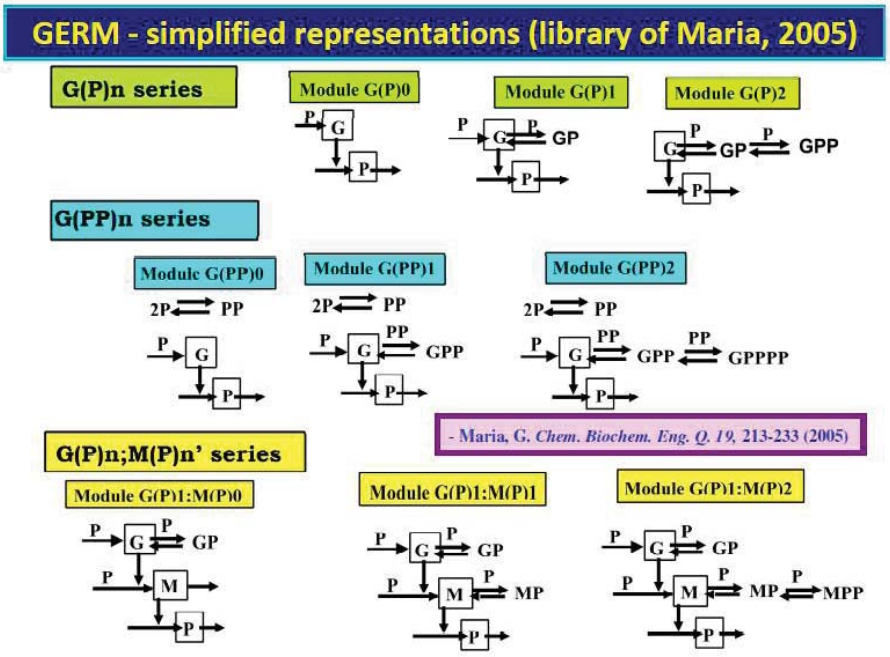


Figure 2-12: The LIBRARY of lumped modular models of GERM–s developed by Maria [2003,2005,2006,2007,2008,2009, 2017a,2017b,2018]. The library includes, in a template form, the kinetic models corresponding to simplified representations of a generic gene expression G/P regulatory module (GERM), assimilated 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 Fig. 2-11, down-right), 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 is the dimmer PP. The other observations from the top-row 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 the transcription / translation. The other observations from the top-row also apply here.

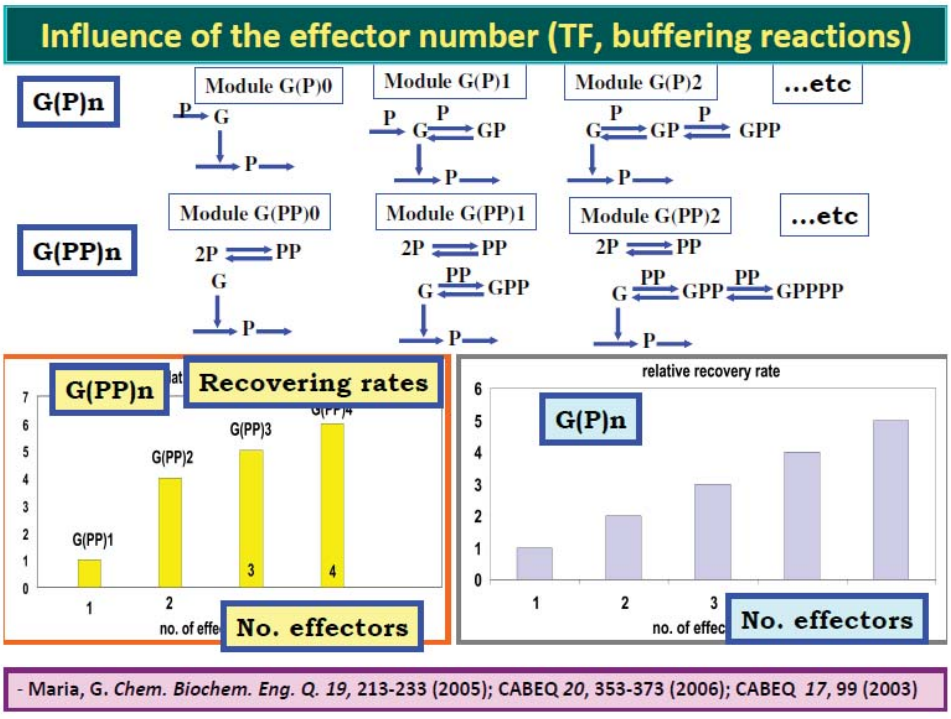


Figure 2-13: Influence of the number of effectors in a GERM on their regulatory performance indices (expressed here as the recovery rate after a dynamic perturbation in the P key-species. Adapted from [Maria, 2017a]. Comparison refers to simple GERM of type [G(P)n], and [G(PP)n]. from Fig. 2-12.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

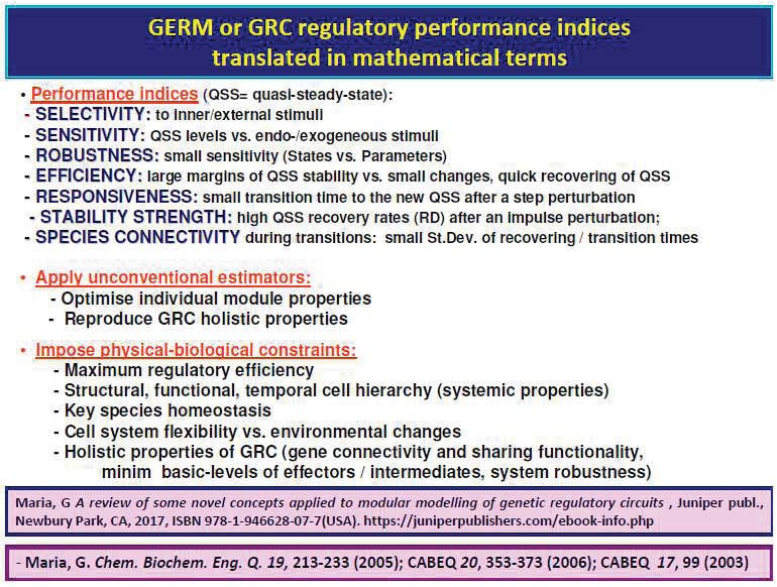


Figure 2-14: Seeking for math modelling of the GERM–s or GRC–s regulatory properties (that is the performance indices P.I.-s of section 2.2.3). See the details of Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018].

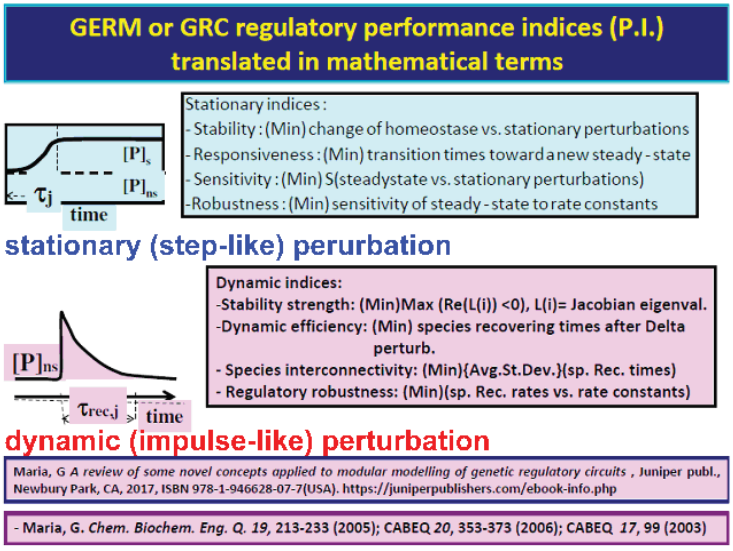


Figure 2-15: Define some of the GERM regulatory performance indices (P.I.-s), in the case of: **[top row]** a stationary ("step"-like) perturbation, **[down row]** a dynamic ("impulse"-like) perturbation). These P.I.-s are extensively discussed in section 2.2.3. See also the details of Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018]; [Maria and Scoban,2017,2018].

(homeostasis), because the rate at which cell-volume increases equals that at which overall number of moles increases, leading to a constant $\left(\sum_{j=1}^{ns} n_j\right)/V$ ratio, Eq.(6). (see Table 2-1).

(9). The species concentrations $C_j(t) = N_j(t)/V(t)$ evaluated at the cell level are usually expressed in nano-moles, being computed with the following relationship [Maria, 2005,2017a,2018]:

$$C_j = \frac{(\text{no.of molecules of species "j", that is copynumbers) per cell}}{N_A\ V_{cyt}} \tag{8}$$

Where: N_A = the Avogadro number; Vcyt = cell volume (cytosol).

For instance, for an *E. coli* cell, with an approximate volume $V_{cyt,o} = 1.66 \times 10^{-15} L$ [Yang et al., 2003], it results for $C_{Gs} = [G]s$ a value of $1/(6.022 \times 10^{23})\ (1.66 \times 10^{-15}) = 1$ nM (that is 10^{-9} mol/L).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

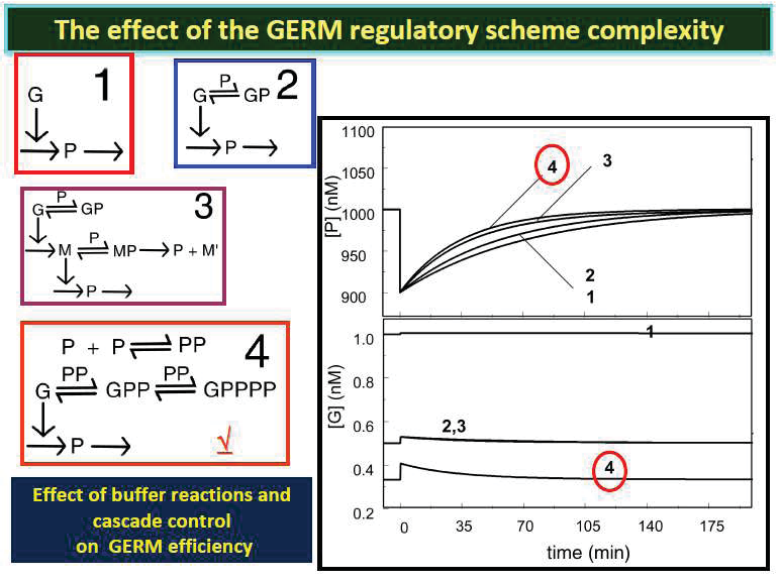
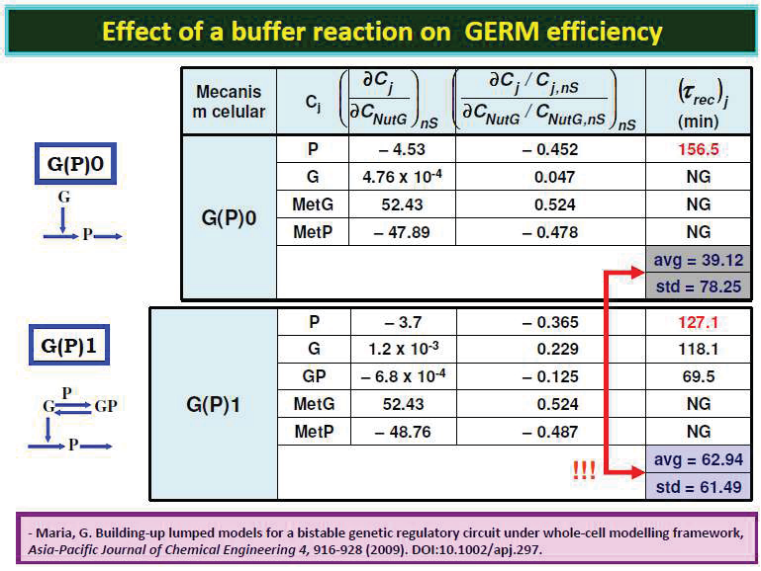


Figure 2-16: GERM model complexity reflected on its dynamic regulatory efficiency to keep the key-species (G,P) homeostasis. Simulations display the P/G recovering trajectories, for every considered GERM model type, after an "impulse"-like dynamic perturbation of -10% [P]s in the key-protein. The GERM models are (see also Fig. 2-12) = "1" = [G(P)0]; "2" = [G(P)1]; "3" = [G(P)1; M(P)1]; "4" = [G(PP)2]. The structure "4" reported the best results, due to the presence of two rapid buffering reactions, while the TF is present in a dimeric (PP) form. See [Maria,2005,2006,2007,2008,2009,2017a,2017b,2018] for details.



Mecanism celular	C_j	$\left(\frac{\partial C_j}{\partial C_{NutG}}\right)_{nS}$	$\left(\frac{\partial C_j / C_{j,nS}}{\partial C_{NutG} / C_{NutG,nS}}\right)_{nS}$	$(\tau_{rec})_j$ (min)
G(P)0	P	– 4.53	– 0.452	156.5
	G	4.76 x 10^{-4}	0.047	NG
	MetG	52.43	0.524	NG
	MetP	– 47.89	– 0.478	NG
				avg = 39.12
				std = 78.25
G(P)1	P	– 3.7	– 0.365	127.1
	G	1.2 x 10^{-3}	0.229	118.1
	GP	– 6.8 x 10^{-4}	– 0.125	69.5
	MetG	52.43	0.524	NG
	MetP	– 48.76	– 0.487	NG
			!!!	avg = 62.94
				std = 61.49

Figure 2-17: The effect of a buffer reaction effector (G+P <=> GP) on the GERM dynamic efficiency; the case of the [G(P)1] type (from Fig. 2-12) [Maria,2005,2009]. Simulations display the P/G recovering trajectories, for every considered GERM model type, after an "impulse"-like dynamic perturbation of -10% [P]s in the keyprotein. Details on the species concentrations at homeostasis (the case of a cell with a high ballast), and on the model rate constants estimation are given by Maria and Scoban [2017,2018].
As resulted, compared to [G(P)0], in the [G(P)1] case, the P recovering rate is higher, while the species are better interconnected (smaller STD).

(10). 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.(2). Under stationary growing conditions, it results an exponential volume growth given by: $V = V_o \exp(+D_m t)$.

(11). 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.(14), of $(dC_j/dt)_s = 0$. The obtained nonlinear algebraic set Eq.(14) under QSS is used to estimate the rate constants K from the known (measurable) stationary concentration vector C_s.

(12). When treated deterministically, the WCVV type kinetic models present solutions with *cell species concentrations* expressed in

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

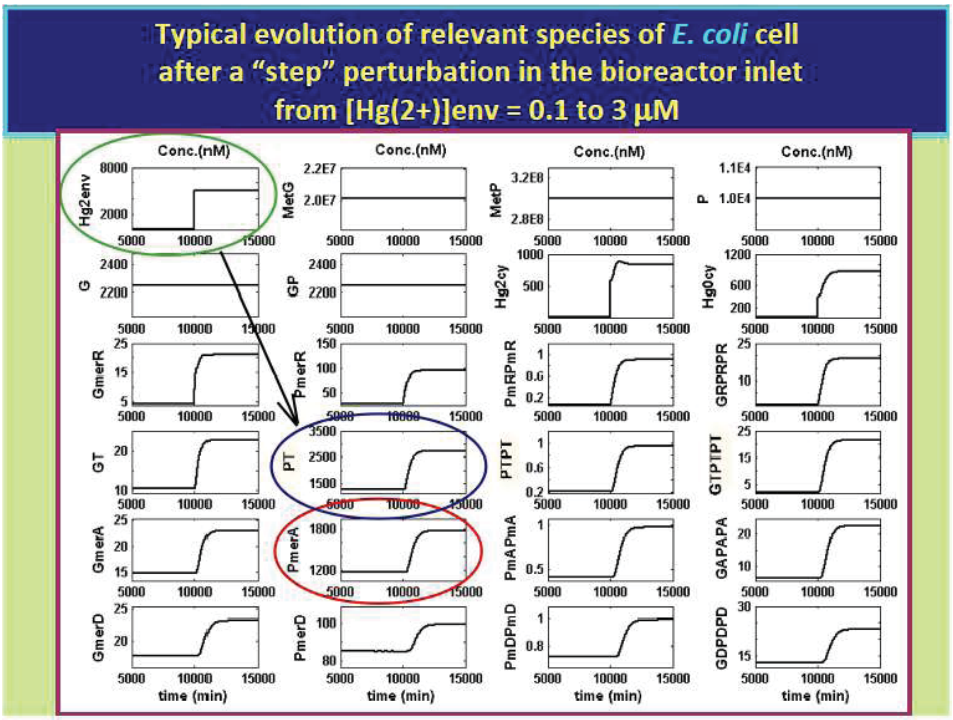


Figure 2-18: Prediction of the cell species concentration dynamics as a response to a "step"-like perturbation of the mercury level in the bioreactor inlet (top-left figure). Such a stationary perturbation will lead to another QSS homeostasis for the all species considered in the mercury-operon expression model of Fig.2-19. Details are given by [Maria and Luta,2013].

real, but *fractional numbers*, which, from a physical point of view, appears to be in contradiction with the reality, because one copynumber can not be divided into pieces. How the WCVV models overcomed this problem? For instance, in a GERM module of [G(P)1] type (Figures 2-24), the fractional copy numbers (e.g. [G]s = 1/2nM) in Table 2-4 . Such fractional concentration of cell species generated in the deterministic

WCVV kinetic models with continues variables, can be interpreted in several ways:

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

2) 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,

3). As a result of the variable cell volume, to which all species contributes, according to the isotonicity constraint, Eq. (11-14), and, in turn the cell volume influence all species concentrations according to the combined relationships Eq.(6; 12), and the Footnote [B] of Table 2-5, that is:

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

Species concentrations in the cell are computed with the formula [Kurata and Sugimoto, 2017]:

$$C_j = \frac{N_j}{N_A \times V_{cyt}} \quad \text{(V2)}$$

where N_A is the Avogadro number.

For instance, for a born E. coli cell, with an approximate volume $V_{cyt,o} = 1.66 \times 10^{-15} L$ [Bar-Joseph et al. 2012], the concentration of one gene G copynumber has a value of:

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

.

Consequently, by combining Eq.(V1) with Eq.(v2), it results:

Species concentrations in the cell are computed with:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

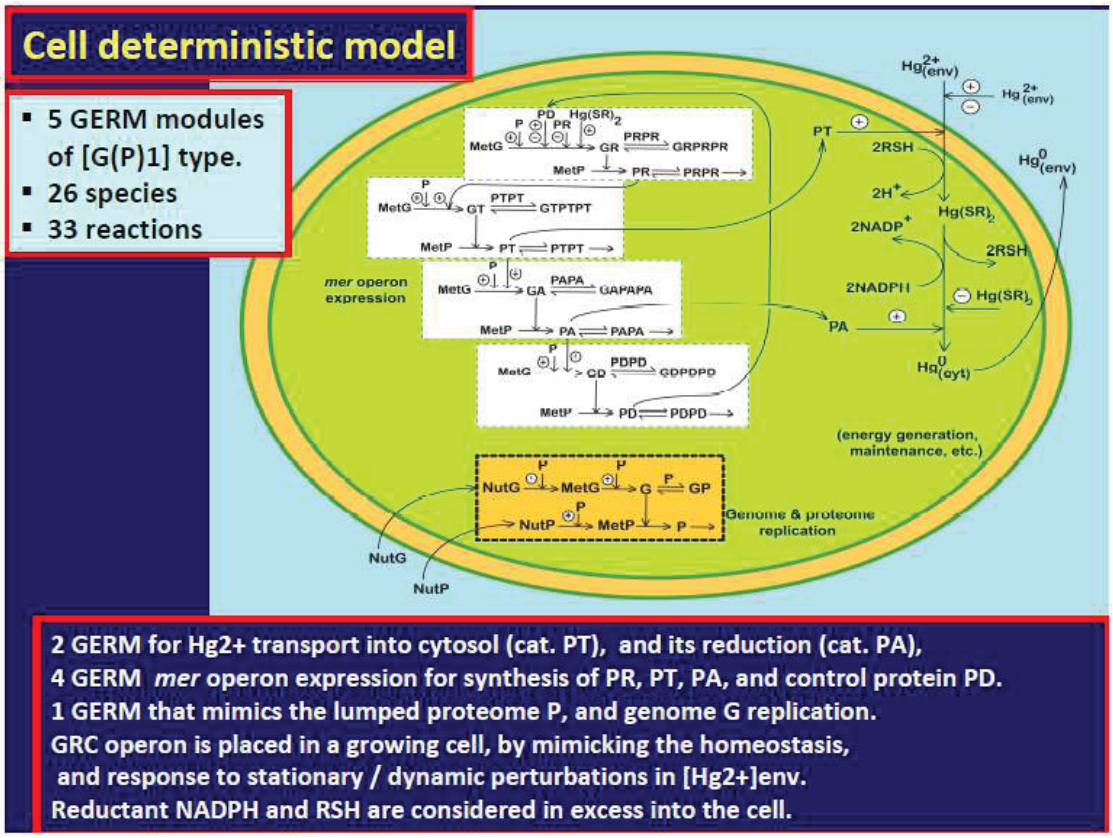


Figure 2-19: The simplified reaction pathway for the WCVV kinetic model of the mercury(mer)-operon expression in E. coli. The model includes 5 GERM-s, and the main enzymatic reactions linked to the mer-operon expression [Maria,2010; Maria and Luta, 2013]. The five GERM−s are of [G(PP)1] type (see Fig. 2-12). More specifically, they are the followings: [GA(PAPA)1] for the reductase PA expression; [GR(PRPR)1] for the protein PR expression. This protein induces the lumped permease PT expression; [GT(PTPT)1] for the lumped permease PT expression;[GD(PDPD)1] for the control protein PD expression; [G(PP)1] is used to mimic the lumped genome and proteome replication. The inclusion in the WCVV model, of this GERM (of lumped genome/proteome) with species (lumped NutG, NutP) of high concentrations is mandatory, in order to fulfil the isotonic law Eq.(6). The GRC is "placed" in the WCVV kinetic model of a growing cell, by mimicking the homeostasis and the cell response to stationary and dynamic perturbations in the environmental [Hg2+]env. The reductant NADPH and RSH are considered in excess into the cell. Notations: P = lumped proteome; G = lumped genome; NutG, NutP = lumped nutrients used for gene and protein synthesis; P• = proteins; G• = genes; RSH = low molecular mass cytosolic thiol redox buffer (such as glutathione). The perpendicular arrows on the reactions indicate the catalytic activation, repressing or inhibition actions. The absence of a substrate, or a product indicates an assumed
concentration invariance of these species. The (+)/(-) symbols indicate positive or negative feedback regulatory loops. This regulatory GRC model was linked with a dynamic model of a fed-batch bioreactor to result a hybrid SMDHKM model. This model, derived under the WCVV novel modelling framework, was used for the insilico
GMO design of *E.coli* to maximize the mercury uptake from wastewaters, and for various engineering analyses [Maria,2010,2008,2009b]; [Maria and Luta,2013; Maria et al.,2013].

$$C_j = \frac{N_j}{N_A \times \left(\frac{RT}{\pi} \sum_{i=1}^{all} N_i \right)} \quad \text{(V3)}$$

where N_A is the Avogadro number. According to the Eq.(V3) and isotonic system hypothesis of the WCVV model, the 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 $\sum_{j=1}^{all} N_i$, and of the cell volume V_{cys}, and so in the C_j of Eq.(V2). For instance, a suddenly increase in the internal lumped metabolites {NutP, NutG, etc.} (Figure 2-24), will lead to a significant increase of their large number of moles, and of the $\sum_{i=1}^{all} N_i$, which, in turn, will lead to a significant increase in V_{cyt}, and to a decrease of C_j in Eq.(V2), particularly for the species " j " present in small amounts (such as G, GP of the GERM kinetic models).

2.2.1.1. WCVV model generic mass balances

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, thet is:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

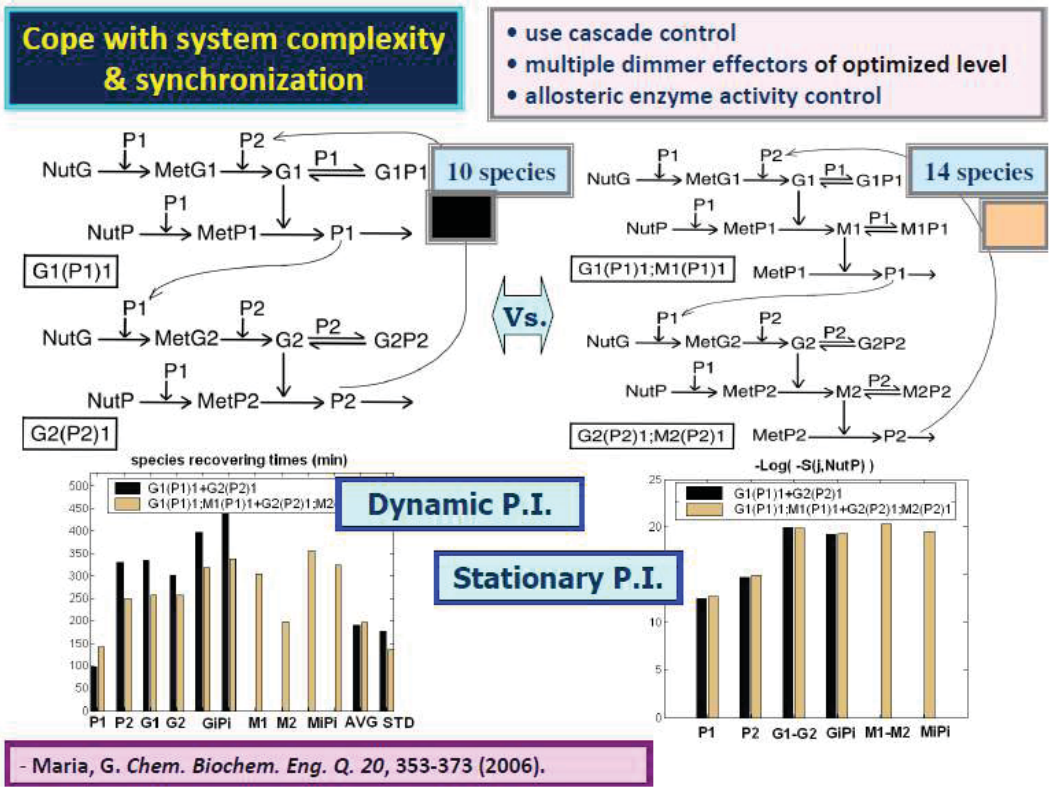


Figure 2-20: Effect of the GERM model complexity on the GRC stationary performances indices (P.I.-s), that is: **[down-left]** species recovering times after a -10%[P1]s dynamic perturbation, and **[down-right]** species stationary le levels Cjs relative sensitivities S(Cj; NutP)s vs. environmental perturbations in the [NutP]s concentration [Maria, 2006]. Figures compare two GRC structures including similar GERM-s, that is **[top-left]** a simpler structure with 10 species and less effectors of [G1(P1)1] linked to [G2(P2)1] type, compared to a more complex structures with 14 species and more effectors of [G2(P1)1 ; M1(P1)1] linked to [G2(P2)1 ; M2(P2)1] . Despite its simplicity, the top-left GRC structure reported better dynamic P.I.-s (down-left), and comparable stationary P.I.-s (down-right).

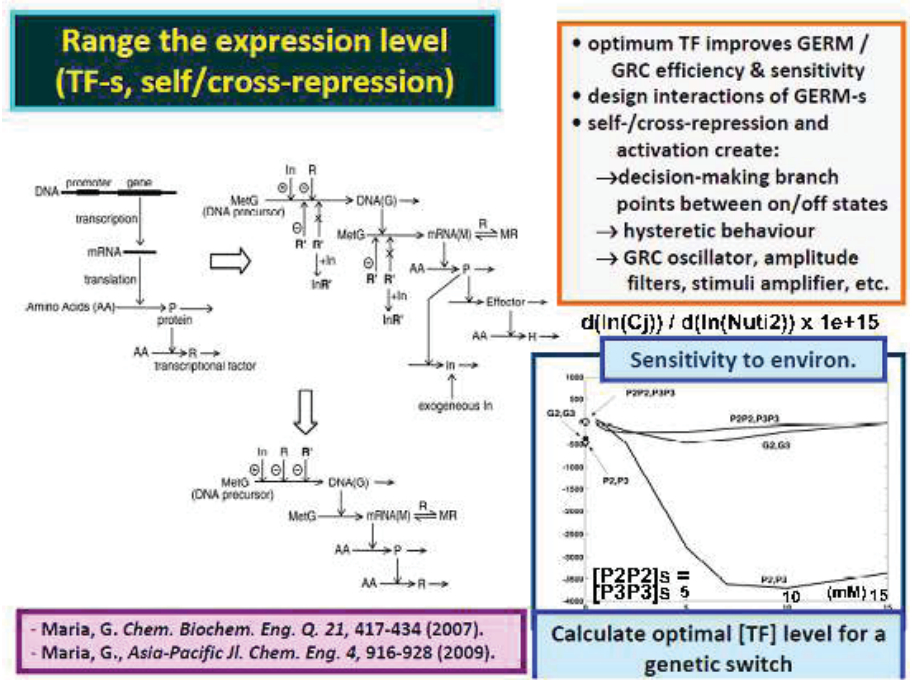
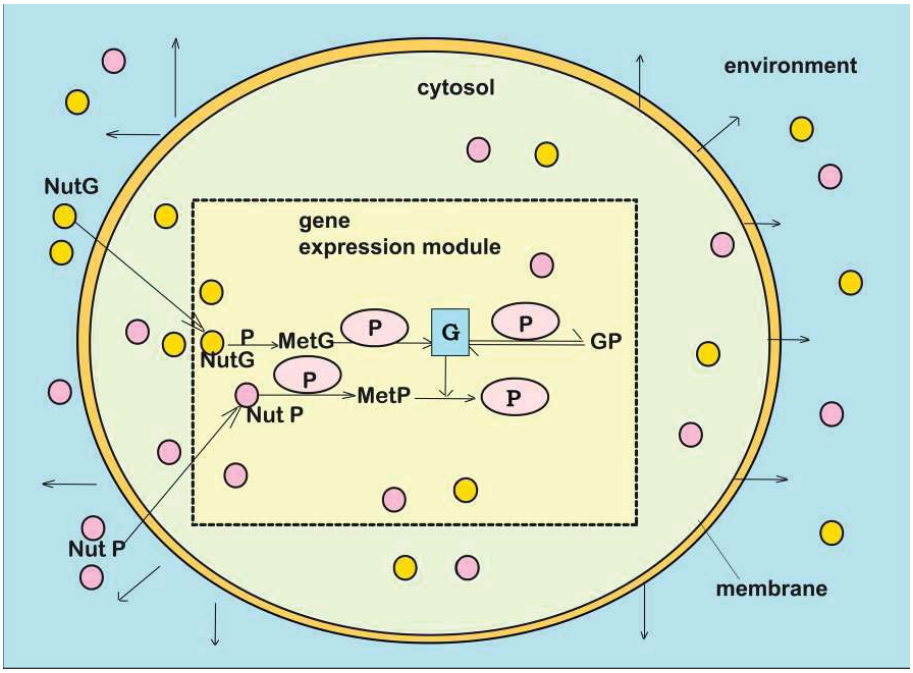


Figure 2-21: Effect of the TF concentration on the GERM efficiency [Maria, 2007,2009,2014b]. The down-plot indicates the key-species stationary sensitivities for various [PiPi] (i=1,2) effector levels predicted by the GRC model of a genetic switch proposed by Maria [2009], of type [G2 (P3P3) n1 (P2P2) m1] + [G3 (P2P2) n2 (P3P3) m2] , with n1,n2 = 1, and m1,m2 = 1 (mixed cross- and selfrepression). The ordinate is the relative (logarithmic) sensitivity of a key-species stationary Cjs vs. the stationary perturbation in the [NutP]s. The chosen optimal [TF] (that is [P2P2]s, and [P3P3]s here) corresponds to the minimum of the P2,P3, and P2P2,P3P3 sensitivity curves vs. the [TF]s. In the example of [Maria, 2009] this optimal value is [P2P2]s = [P3P3]s = 4-10 nM.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124



Or, in a simplistic form:

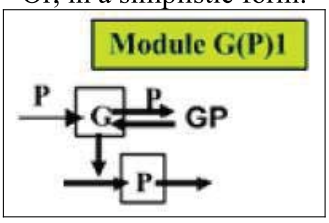


Figure 2-22: The simplified lumped reaction schema for an individual gene expression regulatory module (GERM) of [G(P)1] type of Fig. 2-12. Notations: NutP and NutGare substrates imported from environment, and used in the synthesis of metabolites MetP and MetG used for P and G synthesis; G = a generic gene (DNA); P = encoded protein; M = mRNA; GP = the inactive complex of G with P. The horizontal arrows indicate reactions. The vertical arrows indicate catalytic actions. The absence of a substrate or product indicates an assumed concentration invariance of these species.

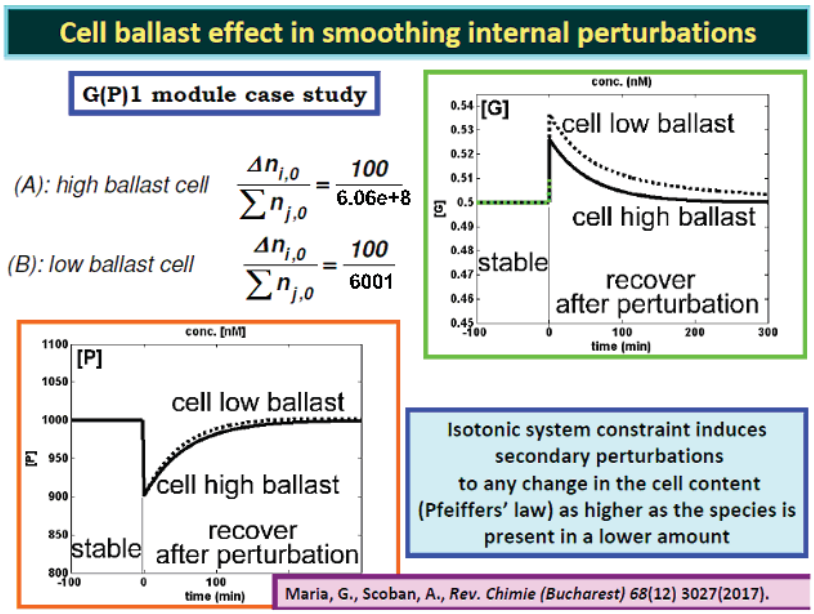


Figure 2-23: Cell ballast effect on the performance indices of a GERM of [G(P)1] type, presented in Fig. 2-22 . The key-species homeostatic concentrations are given in Table 2-4. [Maria and Scoban, 2017,2018].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

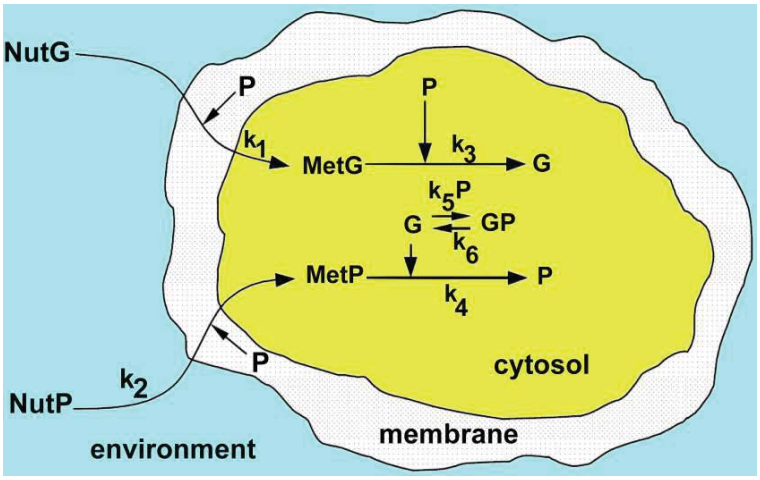


Figure 2-24: The reaction scheme of the [G(P)1] regulatory module GERM of the generic gene G expression used to exemplify the synthesis of a generic P protein in the *E. coli* cell, proposed by Maria [2003,2005,2017a]. To improve the system homeostasis, that is the quasi-invariance of key species concentrations (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 [2005]. See also the case study of section 2.3.1.

$$C_j = \frac{N_j}{N_A \times \left(\frac{RT}{\pi} \sum_{i=1}^{all} N_i \right)} \quad \text{(V3)}$$

where N_A is the Avogadro number. According to the Eq.(V3) and isotonic system hypothesis of the WCVV model, the 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 $\sum_{j=1}^{all} N_i$, and of the cell volume V_{cys}, and so in the C_j of Eq.(V2). For instance, a suddenly increase in the internal lumped metabolites {NutP, NutG, etc.} (Figure 2-24), will lead to a significant increase of their large number of moles, and of the $\sum_{i=1}^{all} N_i$, which, in turn, will lead to a significant increase in V_{cyt}, and to a decrease of C_j in Eq.(V2), particularly for the species " j " present in small amounts (such as G, GP of the GERM kinetic models).

2.2.1.1. WCVV model generic mass balances

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:

$$C_j(t) = \frac{N_j(t)}{V_{cyt}(t)} \quad (9)$$

Consequently, Eq. (1b) should be re-written for a variable volume, in the form:

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

$$\frac{1}{V}\frac{dN_j}{dt} - D_i C_j = h_j(C,k,t);$$

With the instant cell dilution rate of $D_i = \frac{1}{V}\frac{dV}{dt} = \frac{d(\ln(V))}{dt}$ (10)

Or, in a more strict sense, by combining with Eq.(lb), it results the WCVV novel model formulation, as following:

$$\frac{dC_j(t)}{dt} = \frac{1}{V}\frac{dN_j}{dt} - C_j\frac{d(\ln(V))}{dt} = \sum_{i=1}^{nr} s_{ij} r_i - D_i C_j = h_j(\boldsymbol{C},\boldsymbol{k},t) \quad (11)$$

Where:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 2-2: The regulatory efficiency performance quantitative indices (P.I.-s) proposed by Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018] to evaluate the GERM–s regulatory properties in maintaining the species homeostasis following a dynamic (impulse-like), or a stationary (step-like) perturbation in a key-component of the GERM kinetic model of WCVV type.

Index	*Suitable objective*	*Expression*
stationary regulation	Min	$R_{SS} = ([P]_s - [P]_{ns})/ [P]_{ns}$;
stationary regulation	Max	$A_{unsync} = k_{syn} \times k_{decline}$
stationary regulation (vs. environmental perturbations in [Nut(j)]s)	Min	$S^i_{NutP_j} = \left[(\partial C_i / C_{is}) / \left(\partial C_{Nut\,j} / C_{Nut\,js} \right) \right]_s$
stationary regulation (vs. structural perturbations in the rate constant k(j))	Min	$S^i_{k_j} = \left[(\partial C_i / C_{is}) / (\partial k_j / k_j) \right]_s$
dynamic regulation	Min	$R_D = Max\left(Re(\lambda_i)\right);\ Re(\lambda_i) < 0$
dynamic regulation	Min	τ_j; τ_P (P = the expressed key protein)
regulatory robustness	Min	$(\partial R_D / \partial \boldsymbol{k})$
species interconnectivity	Min	$AVG(\tau_j) = \text{average}(\tau_j)$
species interconnectivity	Min	$STD(\tau_j) = \text{st.dev.}(\tau_j)$
QSS stability strength	Min	$Max\left(Re(\lambda_i)\right);\ Re(\lambda_i) < 0$
QSS stability strength*	Min	$\lvert \lambda_A \rvert < 1$

Abbreviations: Min = to be minimized; Max = to be maximized. Note: k(syn) and k(decline) refers to the -➔ P -➔ overall reaction.
Notations: "n"= nominal value; "s" = stationary value; (*) see the definition and the calculation of the monodromy matrix A in section 2.2.3, and [Maria, 2003]; λ(i) = i-th eigenvalue of the Jacobian matrix of Eq.(4); τ(j)= species "j" recovering time after a dynamic (impulse-like) perturbation in a key-species of the GERM kinetic model of WCVV type; Nut= nutrient; Re= real part; AVG= average; STD= standard deviation; Cj= species "j" concentration; RD= dynamic regulatory (recovering) index; QSS = quasi-steady-state; P denotes the expressed key-protein in the analysed GERM kinetic module.

Table 2-3: The main performance indices (P.I.s), that is the quantitative measures of the GERM–s regulatory efficiency, as defined by [Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018]; [Maria and Scoban,2017,2018]. The tested GERM-s kinetic models have been developed, and compared with the literature data, by using the WCVV modelling framework (sections 2.2.1; 2.2.2., and 2.3, and 3). These (P.I.s) are inter-connected to their math formulation of Table 2-2.

Perturbation type (external or internal)	
A).- Stationary P.I.	**B) Dynamic P.I.**
A-i).- Transition time	**B-a).- Recovering time**
A-ii).- Responsiveness to perturbations	**B-b).- Recovering rate**
A-iii).- Stationary efficiency	**B-c).- Regulatory robustness.**
A-iv).- The steady-state Cs stability strength.	**B-d).- Species interconnectivity**
	B-e) Cell sub-system QSS stability

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

.

It is to observe that the instant cell dilution rate "Di" defined in Eq.(10), is different from the average (apparent) cell dilution rate "Dm" defined and evaluated in Eq.(2). 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 (last row of Eq.(6)), that is:

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

,

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 2-4: The nominal (homeostatic QSS) *E. coli* cell conditions of a gene expression module (GERM), modeled with a kinetic model of [G(P)1] type (Fig. 2-22, and Fig. 2-24). Also, the recovering rates after a −10% impulse perturbation in the [P]s of 1000 nM at an arbitrary *t*=0. Cell initial volume of the considered *E. coli* cell, is of Vcyt,0 = $1.66X10^{-15}$ L [c]; tc = 100 min. Adapted from Maria [2005], and Maria and Scoban [2017, 2018].

Species	Low ballast cell (nM)		High ballast cell (nM)	
	QSS conc. (nM)	Recovery time (min)	QSS conc. (nM)	Recovery time (min)
Lumped $C_{Nut}P$	3000	NG	3×10^8	NG
Lumped C_{Nut}	3000	NG	3×10^6	NG
Lump $\sum_j C_{MetG} G_{j,s}$ [a]	~2000	NG	~3×10^6	NG
Lump $\sum_j C_{MetP} P_{j,s}$ [b]	3000	NG	3×10^8	NG
C_{Ps} = [P]s	1000	103	1000	133
C_{Gs} = [G]s	0.5	223	0.5	93
C_{GPs}	0.5	246	0.5	93

Footnotes:

The lump $\sum_j C_{Met} G_{j,s}$ results from the isotonic constraint Eq.(6), and the constraint ($\pi_{cyt} = \pi_{env}$), and $\left(\sum_j^{all} C_j\right)_{cyt} = \left(\sum_j^{all} C_j\right)_{env}$ of Table 2-1, leading to the relationship:

$$C_{NutG} + C_{NutP} = \sum_j C_{MetG_j} + \sum_j C_{MetP_j} + \left(\sum_{\substack{j \neq MetP_j \\ j \neq MetG_j}}^{all\ cell} C_j \right)$$ Consequently, one obtains:

$$\sum_j C_{Met\,G_j} = C_{NutP} + C_{NutG} - \sum_{j \neq MetG}^{all\ cell} \sum_j .$$

The lump $\sum_j C_{MetP\ j,s}$ results analogous to the footnote [A], from the isotonic constraint:

$\sum_j C_{MetP} = C_{NutP} + C_{NutG} - \sum_{j \neq MetP}^{all\ cell}$. As the relationship [A] presents two unknowns {that is $\sum_j C_{Met} G_{j,s}$ and $\sum_j C_{MetP\ j,s}$ }, in the present case, arbitrarily $\sum_j C_{MetP}$ was set to $C_{Nut}P$.

The considered cell cycle is of *tc* = 100 min. The cell-volume average logarithmic growing rate is $Dm = \ln(2)/tc$. The $\mathrm{Max}\left(\mathrm{Re}\left(\lambda_j\right)\right) < 0$ indicates a stable QSS homeostasis of these cell species concentrations, where λ_j are the Jacobian Eq.(4) eigenvalues of the ODE kinetic model of the cell system in a WCVV formulation Eq.(11). The rate constants of the [G(P)1] model Eq.(11) results from solving the stationary model Eq.(14) (that is for *dCj/dt* = 0) with the average $Dm = \ln(2)/tc$, with the cell cycle of *tc* = 100 min., and with the known stationary species concentrations displayed in the above table. The only G + P ⇔ GP buffer reaction was considered with the reverse reaction rate constant of $10^5$1/min, due to the reason presented by Maria [2005, 2006].

Notations: (see also Figure 2-22). $N_{ut}P$ and $N_{ut}G$ are substrates imported from environment, and used in the synthesis of metabolites $M_{et}P$ and $M_{et}G$ used for P and G synthesis; G = a generic gene (DNA); P = a generic protein; M = RNA; GP = the inactive complex of G with P; = species "j" concentration; cyt= cytoplasma; "o"= initial; 's' index refers to the stationary state (QSS); NG = negligible.

Which, by derivation and division with *V* is leading to [Maria,2005,2006,2007]:

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

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

The average (apparent) cell logarithmic dilution rate *Dm* is:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

$$\frac{dV}{V} = D_m \times t; \Rightarrow \int_{V_o}^{2V_o} \frac{dV}{V} = \int_0^{t_c} D_m dt; \Rightarrow D_m = \ln(2)/t_c;$$

$$V = V_o \exp\left(D_m \times t\right) \tag{12}$$

Consequently, the cell dilution rate D_i is linked and depends on the all reaction rates r_j of the cell species N_j (taken individually, or lumped). By contrast, the average (apparent) cell logarithmic dilution rate Dm is equal to In (2)/tc.

As revealed by the Pfeffer's law Eq.(6) in diluted solutions [Wallwork and Grant, 1977], the volume dynamics is linked to the molecular species dynamics under isotonic and isothermal conditions by means of relationship Eq.(6). Consequently, as proved by Eq.(12), 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.(13), derived from the Pfeffer's law in diluted solutions, and accounting for the mass conservation law:

$$V(t) = \frac{RT}{\pi} \sum_{j=1}^{ns} N_j(t), \Rightarrow$$
$$\frac{RT}{\pi} = \frac{V(t)}{\sum_{j=1}^{ns} N_j(t)} = \frac{1}{\sum_{j=1}^{ns} C_j} = \frac{1}{\sum_{j=1}^{ns} C_{j,0}} = constant \tag{13}$$

If this WCVV modelling concept is not applied, the default classical constant volume ODE kinetic modelling of type Eq. (1a-b) has been applied, with a large number of inconveniences, related to ignoring lots of intra-cell effects, such as (section 2.2.3): A). The influence of the cell ballast in smoothing the homeostasis perturbations; B). The secondary perturbations transmitted via the cell volume following a primary perturbation, according to Eq.(13) in isotonic systems; C). The more realistic evaluation of GERM regulatory performances indices (P.I.-s, section 2.2.3), and of the recovering/transient times after perturbations, etc. [Maria, 2005,2006,2007,2017a,2017b,2018].

When elaborating the WCVV kinetic models, or the CCM / GERM / GRC models, the cell-volume growing rate effect is essential for accounting several effects, such as [Maria, 2005,2006,2007,2017a,2017b,2018]: D). The continuous diluting of the cell content, cell-'resistance' to small perturbations in the internal/external components due to the inertial 'big volume' effect; E). The indirect effect of the perturbations in species concentrations on the cell-metabolism through the induced changes in the volume growing rate. F). The volume-growth diluting effect acts as a continuous stationary perturbation for the concentration-levels, and formally can be assimilated with a first order decay rate of all cellular species during a cell-cycle in Eq.(10).

It follows that the cell-volume variation during the most of a cellgrowth cycle (the first 80% [Surovstev et al.,2007; Morgan et al., 2004]) is an essential term to be accounted in every cell-modelling attempt to obtain more satisfactory predictions. Although most of reported models, both deterministic or stochastic, ignore or diminishes such effects and build-up 'constant' cell-volume equations WCCV (section 2.1) written in terms of species concentrations, the new elaborated regulatory/whole cell-models WCVV over the last couple of years accounted the cellvolume growth in an explicit way [Maria, 2005, 2006, 2007, 2009, 2010, 2014b; Maria and Luta, 2013; Morgan et al., 2004; Surovstev et al., 2007; Vilela et al., 2010]. The superiority of WCVV framework models *vs.* the classical WCCV models was discussed and proved in section 2.3.1.

For instance, Sewell et al. [2002] included the volume-diluting effect only for the protein-concentrations through a formally first-order decay rate. Such an approach, even being satisfactory for rapid and simple predictive purposes, suffers of two major disadvantages: (i) the 'decay' rate is not considered for all the species in order to avoid high model complexity, and (ii) the 'decay' rates can report different rate constants for various species, when in reality the same cell diluting rate should be applied for all the considered species in the cell model; (iii) the 'inertial' cell-volume / large copynumber effect to smooth perturbations can not be simply and naturally included in a constant-volume cell model. More details are given in sections 2.2.3., 2.2.4

The same fictive decay-rate approach has been reported by Tomita et al. [1999] in developing an 'E-cell' continuous differential model of a cell, with including a larger number of genes and proteins. As mentioned by Tomita [2001], "the E-cell system also accepts user-defined reactions, making it capable of handling many other phenomena such as diffusion and variable cell volume". Using the EcoCyc and KEGG databases, the authors simulate the dynamics of 127 genes/proteic system for the *M. genitalium* cell. However, this model suffers from several drawbacks such as the lack of autocatalytic effects during a cell-cycle, by considering any replication of the genome, and any cell-division process. Recently, the authors reported some improvements of the E-cell model with including the osmotic pressure balance and volume cell growth without giving any detail [Kinoshita et al.,2001].

Other models, such of Gibson and Bruck [2001], avoid including the cell-volume increase effects when considering only first-order reaction terms into the model equations. And they do not consider the WCVV hypotheses (section 2.2.1). However, the authors signalled that such an approximation "can create a large calculation error".

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

In the WCVV modelling approach, separate equations explicitly link the cell-volume growth by the cell-osmotic pressure, while the continuous ODE model is re-written in terms of species moles under the variable volume of the cell system. In the GERM regulatory models (sections 2.2.2, 2.2.3) the cross-autocatalytic effects are also included when separate protein and gene synthesis catalytic pathways are considered together. Moreover, external cell-factors are better accounted by separately considering the protein and gene 'raw-materials' (see the GERM-s kinetic model formulations in sections 2.2.2., 2.2.3). It is also to observe that, from definition of Di in Eq. (12), it results the cell volume dynamics given in Eq.(2).

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.(2).

For stationary balanced growing conditions (that is at quasi steadystate, QSS), species synthesis rates are equal to first-order dilution rates (DC_j), leading to time-invariant species concentrations (i.e. homeostatic conditions, $\left(dC_j/dt\right)_s = 0$). Thus, from Eq. (11), and combining with Eq.(12), one obtains the WCVV model:

$$\left(\frac{dC_j(t)}{dt}\right)_s = \left(\frac{1}{V}\frac{dN_j}{dt}\right)_s - D_s C_{js} = h_{js}\left(\boldsymbol{C}_s, \boldsymbol{k}, t\right) = 0;$$
$$j = 1,\ldots,n_s;\ D_s = \left(\frac{RT}{\pi}\right)\sum_{j=1}^{ns}\left(\frac{1}{V}\frac{dN_j}{dt}\right)_s \qquad (14)$$

It is to observe that in the WCVV model the instant cell dilution D_i in Eq. (10), is fundamental different to the stationary cell dilution rate D_s of Eq. (14), which, in turn is roughly equal to the average one D_m of Eq.(2).

As proved by Maria et al. [2018d] the use of *Dm* instead of the instant D_i can lead to biased results, and wrong simulation results. This is also valid for the use of classical WCCV formulation instead of the novel WCVV one (sections 2.2.2, 2.2.3, 2.2.4, and 2.3).

2.2.1.2. Some advantages of the novel WCVV modelling framework

When modelling metabolic cell processes with deterministic continuous-variable models by using a certain number of novel concepts related to the WCVV modelling approach, certain advantages have to be underlined, as followings [Maria,2003,2005,2006 ,2007,2008,2009,2009b,2009c,2010,2014a,20 14b]; [Hudder et al.,2002; Maria et al.,2011]. The proofs of these advantages are given in section 2.2.3.

1). The more realistic representation of the GERM / GRC regulatory properties, in a quantitative, easily interpretable way (section 2.2.3). Such an approach was proved to lead to a more effective representation of the cell regulatory processes, being very useful when developing modular kinetic representations of the GRC-s that control the protein synthesis and homeostasis of metabolic processes. Such structured kinetic models were proved to be useful to design GMO-s. [Maria and Luta, 2013; Maria,2018]. As ca. 80% of the cell cycle is the growing phase and, assuming a quasi-constant osmotic pressure and a constant volume growing logarithmic rate, the WCVV cell model can be considered satisfactory to study the cell GRC effectiveness.

2). The novel WCVV modelling framework to buildup kinetic cell models presents the advantage of explicitely inclusion in the model equations of the isotonicity constraints, and of the connection between the variable cell volume and all reactions taking place into the cell.

3). The holistic WCVV approach reveals the role played by the "cell ballast" in smoothing the effect of perturbations coming from the environment. The WCVV modelling framework was succesfully used to derive effective kinetic models describing various genetic regulatory circuits (GRC)(like operon expression, genetic switch, etc.), and individual gene expression modules (GERM) [Maria,2003,2005,2006,2007,2008,2009,2009b,2009c,2010,2014b]; [Hudder et al.,2022; Visser et al.,2004].

4). The WCVV model formulation was proved to be also suitable to accurately model the cell growth and its division [Morgan et al.,2004].

5). The novel WCVV modelling framework allows a realistic analysis of rules to be used for GERM linking when buiding-up complex GRC dynamic models (section 2.2.4), thus offering the possibility i) to simulate the regulatory performances of an individual gene expression, of an operon expression, or ii) of a GRC (switch, amplifier, filter, etc. [Savageau,1987,2002; Wall et al.,2003]), and also iii) to *in-silico* design of GMO or cloned micro-organisms with target plasmides to obtain desirable characteristics for industrial or medical applications. Some examples include: maximization of succinate production in *E. coli* [Maria et al.,2011];, efficient removal of mercury from wastewaters [Maria et al.,2013; Maria and Luta,2013; Maria,2009b,2010]; design of a genetic switch of desirable characteristics [Maria,2007,2009,2014b].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

On the contrary, application of the default classical WCCV ODE kinetic models of Eq. (1a-c) type with neglecting the isotonicity constraints presents a large number of inconveniences, related to ignoring lots of cell properties, discussed in detail by Maria [2017a, 2017b, 2018], that is (see sections 2.2.3, and 2..3):

- the influence of the cell ballast in smoothing the homeostasis perturbations;
- the secondary perturbations transmitted via cell volume following a primary perturbation (see the proof given in the below paragraph);
- the more realistic evaluation of GERM regulatory performance quantitative indices (P.I.-s);
- the more realistic evaluation of the recovering/transient times after perturbations;
- conditions when the intrinsic model stability is lost;
- GERM / GRC self-regulatory properties after a dynamic, or a stationary perturbation, etc.

By contrast, the WCVV novel modelling concept / framework proposed by Maria [2002], and [Maria et al., 2002], with the further developments/ studies of the WCVV properties given by Maria [2002,2003,2005,2006,2007,2009,2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c] was used to derive cell kinetic models, in a holistic approach. Such WCVV cell models, "automatically" maintains the intracellular homeostasis of cell processes, while growing autocatalytically on environmental nutrients present in variable amounts. Also, the WCVV kinetic models of cells allow simulating the individual or the holistic regulatory properties of the GERM / GRCs, by including in a natural way constraints related to the *cell system isotonicity*, and the *variable-volume* in relationship to the species reaction rates, and the *lumped proteome/ genome* replication. Such an *isotonicity constraint* is required to ensure the cell membrane integrity, but also to preserve the homeostatic properties of the cell system, not by imposing "the total enzyme activity" or the "total enzyme concentration" constraints suggested in the literature [Heinrich and Schuster, 1996]. The novel WCVV modelling framework is leading to accurately simulate lot of cell metabolic effects, such as (see the proofs in sections 2.2.3, and 2.3):

A. The role of the high cell-ballast in "smoothing" the continuous perturbations of the cell homeostasis due to external nutrient concentrations;

B. The secondary perturbations transmitted via the cell volume. In other words, the system isotonicity constraint of WCVV models implies 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 the cell volume;

C. The WCVV kinetic models of GERM-s allow comparing the regulatory efficiency in a quantitative way, for various GERM-s types, and for various defined performance indices (P.I.-s).

D. The WCVV kinetic models of GERMs / GRCs allow a more realistic evaluation of their regulatory properties (that is P.I., performance indices, see section 2.2.3);

E. The WCVV kinetic models of cells allow studying the recovering/transient intervals between steady-states (homeostasis) after stationary perturbations in the environment;

F. The WCVV kinetic models allow studying conditions when the cell system homeostasis intrinsic stability is lost;

G. The WCVV kinetic models allow studying the self-regulatory properties of a cell system after a dynamic/stationary perturbation;

I. The WCVV kinetic models allow studying the plasmid-level effects in cloned cells;

J. The WCVV kinetic models allow studying the relationships between the external conditions, species net synthesis reactions, and the cell osmotic pressure.

By using the WCVV novel modelling framework, Maria [2002,2003,2005,2006,2007,2009,2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c] developed structured reduced dynamic models of various complexity to simulate individual GERMs, but also linked GERMs in GRCs of modulated functions (e.g. toggle-switch, amplitude filters, modified operons, etc.), used to design GMOs of practical interest [Maria,2009,2014b,2018,2023].

The novel WCVV modelling framework was used to simulate some cell metabolic processes by using the concept of a modular simulation platform, extensible, and linked to -omic databanks (Ecocyc, KEGG, Brenda, Prodoric, CellML, NIH, etc.) [Maria,2005,2014a,2014b]; [Maria and Luta,2013; Tomita et al.,1999; Tomita, 2001]. By using conventional and non-conventional estimation techniques (see the below paragraph), and usually incomplete experimental data from literature, a couples of notable applications have reproduced the protein synthesis homeostatic regulation, or efficiency of some GRCs [Maria,2003,2005,2006,2007,2008,2009,2009b,2009c,2010,2014b]. Such reduced dynamic deterministic models can fairly simulate

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

some essential cell processes, such as glycolysis in *Escherichia coli* cells (Figure 2-5), under stationary or perturbed growing conditions [Maria,2014a; Maria et al.,2018b,2018e]. Such a modular cell simulator can be used to *in-silico* design genetic modified or cloned micro-organisms with industrial or medical applications [Maria,2003,2005,2006,2007,2008, 2009,2009b,2009c,2010,2014a,2014b]; [Hudder et al.,2002; Visser et al.,2004; Maria et al.,2011].

The mechanistic (deterministic) WCVV approach with contiuous variables was very adequate when developing structured complex kinetic models for various purposes. For instance, the complex WCVV structured kinetic model of Maria [2009b,2010], and Maria and Luta, [2013] with the reaction pathway of (Figure 2-6) is able to adequately simulate the efficiency of the GRC responsible for the mercury induced expression of the mer-operon in gram-negative bacteria (*Pseudomonas sp.,E. coli*) responsible for regulation of the mercury ions uptake and reduction from wastewaters [Maria et al.,2013; Maria and Luta, 2013; Scoban and Maria, 2016; Maria,2009b,2010]. The *hybrid* dynamic model connects the structured cell WCVV model with the dynamic model of the fed-batch bioreactor. By using the literature experimental data the rate constants of this complex model were successfully estimated. This hybrid WCVV model was used: i) to optimize the operation of an industrial fluidized-bed fed-batch bioreactor used for mercury removal from wastewaters, and ii) to *in-silico* design a GMO *E. coli* with an adjustable level of mercury-plasmids to maximize the bioreactor efficiency [Maria et al.,2013; Maria and Luta, 2013; Maria,2009b,2010].

2.2.1.3. Rate constant estimation in the WCVV kinetic models:

In the WCVV differential model of Eq.(11) type, if a large number of species and reactions are considered, a large number of rate constants results. These model parameters (rate constants) are estimated by using several conventional and non-conventional numerical methods. If the stationary cell species concentration vector C_s is known (for the components considered in the kinetic model as individual or lumped species), then, by imposing the quasi-steady-state (QSS, homeostatic) condition, it results the nonlinear set of equations Eq.(15), that is:

$$\left(\frac{dC_j(t)}{dt}\right)_s = \left(\sum_{i=1}^{nr} s_{ij} r_i\right)_s - D_s C_{js} = h_{js}\left(\boldsymbol{C_s}, \boldsymbol{k}, t\right) = 0$$

$$j = 1,...,n_s; D_s = \left(\frac{RT}{\pi}\right)\sum_{j=1}^{ns}\left(\frac{1}{V}\frac{dN_j}{dt}\right)_s \tag{15}$$

The estimated rate constant vector $\hat{k}$ and the estimated unknown species stationary concentrations $\hat{C}_s$ results by solving the nonlinear algebraic set Eq.(15), for every cell subsystem (e.g. the WCVV kinetic model of an individual GERM), by using an effective procedure [Iordache and Maria, 1991; Maria,2004; Smigelschi and Maria, 1986]. Another possibility is to solve Eq.(15) by transforming it into a nonlinear programming problem (NLP), to be solved by using effective optimization routines, like the MMA of Maria [2003b,2004,1998]. Usually, Ds is adopted as being the average cell dilution rate Eq.(2)

As the (RT/ π) term is known from the initial condition Eq.(13), and the number of model parameters is usually higher than the number of observed cell species (even if lumped), supplementary conventional and non-conventional constraints should be applied to the NLP solver, that is:

1). Physical meaning of the solution, $\left[\hat{k}, \hat{c}_s\right] > 0$. Here, the superscript " $\wedge$ " denotes the estimated value.

2). In the case of the WCVV kinetic models of GERMs (section 2.2.2), some constraints should be imposed to ensure the optimal regulatory efficiency of the GERM-s. For instance,

2a). Minimum recovering time of the stationary concentrations (homeostasis) after a dynamic ('impulse'-like) perturbation in a keyspecies [Maria, 2005, 2017a, 2018; Maria and Scoban, 2017, 2018]:

$$\left[\hat{k}, \hat{C}_s\right] = \text{argMin}\left(\tau_P\right) \tag{16}$$

2b). The active $[L]_{active}$ and inactive $[L]_{inactive}$ forms of the regulatory element (that is the "catalyst" here) must be equal. In other words:

$$[L]_{active} / [L]_{total} = 1/2, \text{ (see section 2.2.2)} \tag{17}$$

As proved by Maria [2003, 2005, 2006, 2007, 2009, 2017a], and by Sewell et al. [2002] such a constraint ensures a maximum efficiency of the GERM control of the P-synthesis.

2c). The all forms of the regulatory elements should be constant (the mass conservation), that is:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

$$\sum_{i=0}^{n}\left[L(O_i)\right]=constant\ ;\sum_{i=0}^{n}\left[G(P_i)\right]=constant\ ;\sum_{i=0}^{n}\left[G(PP_i)\right]=constant \quad (18)$$

In Eq.(16), notation τ_p denotes the recovering time of the [P]s stationary concentration after an impulse-like perturbation (where P refers to a generic protein P in a GERM model of section 2.2.2). According to section 2.2.3, the τ_p has been evaluated by applying $a\pm 10\%[\mathrm{P}]_S$ impulse perturbation, and by determining, by using the WCVV simulation model, the recovering time with a tolerance of $1\%[\mathrm{P}]_S$.

In Eq.(17-18) "Li" (e.g. enzymes P , genes G , or M = mRNA) is a GERM component at which the regulatory elements O/R, and the transcription factors (TF) acts (Figure 2-7).

To estimate $\left[\hat{k},\hat{c}_s\right]$, other regulatory global properties can also be used as constraints, together with the constraints Eqns. (13, 14, 16-18) [Maria,2005, Van Someren et al., 2003]. For instance, the reverse reaction rate constants in the rapid buffer reactions of GERM-s, of type L + O ⇔ LO (or, equivalently G + P ⇔ GP, in section 2.2.2) are adopted at values five to seven orders of magnitude higher than the cell average dilution rate D_m (see the proof of Maria [2005,2017a,2006,2018]). The L/ LO denotes here the active / inactive forms of the regulatory element respectively. That is because fast buffering reactions are close to equilibrium and have little effect on the metabolic control coefficients. As a consequence, rate constants of such rapid reactions are much higher than those of the core CCM syntheses and of the cell dilution rate.

2.2.1.4. Proving the importance of the WCVV approach in computing the "secondary perturbations" transmitted via cell volume

When one applies an impulse perturbation to one of the cellspecies (e.g. a generic protein denominated by P), that implies removal (by excretion), or addition (by import) of a certain copynumbers of P from (to) the cell. By keeping constant the cell osmotic pressure π, and the temperature, such a perturbation implies, according to Eqs.(6,13), an immediately contraction (or expansion) of the cell-volume from V to V^*. Consequently, the species concentrations will vary from $C_j=N_j/V$ to $C_j^*=N_j/V^*$, irrespectively if the component " j " suffered any copynumbers variation. This effect due to the cell volume variation is called "*secondary perturbations*" or "indirect" perturbations [Maria, 2005]. The difference in the P copynumbers before and after perturbation, i.e. $\left(N_{P,o}-N_P^*\right)$ (the * denotes the perturbed state), can be easily calculated for a certain imposed final concentration for P , i.e. $c_P^*=N_P^*/V^*$. The final copynumbers N_P^*, for an imposed C_P^* results by re-evaluating the species concentrations, from Eq.(6), with the relationship:

$$N_P^*=V_oC_P^*\frac{RT}{\pi}\frac{\left(\sum_{j=1}^{ns}C_{jo}-C_{P,o}\right)}{\left(1-C_P^*\frac{RT}{\pi}\right)} \quad (A1)$$

To prove the relationship Eq.(A1), one considers that, before and after applying the P perturbation, the sum of copynumbers of all other species remains invariant, that is:

$$\sum_{j=1}^{ns}N_{j,o}-N_{P,o}=\sum_{j=1}^{ns}N_j^*-N_P^* \quad (A2)$$

On the other hand, the volume relationship Eq. (6), written for the initial and final states, and by keeping the same pressure π, is $V_o=\frac{RT}{\pi}\sum_{j=1}^{all}N_{jo}$, and $V^*=\frac{RT}{\pi}\sum_{j=1}^{all}N_j^*$, respectively. By multiplying Eq. (A2) with RT / π constant value, then changing terms from left to right, and including the volume formula, one obtains the net volume variation:

$$\left(V^*-V_o\right)=\frac{RT}{\pi}\left(N_P^*-N_{P,o}\right) \quad (A3)$$

By dividing Eq. (A3) with the product (V^* V), and then multiplying with N_P^*, one obtains:

$$\frac{N_P^*}{V_o}-\frac{N_P^*}{V^*}=\frac{N_P^*}{V^*}\frac{RT}{\pi}\left(\frac{N_P^*}{V_o}-\frac{N_{P,o}}{V_o}\right) \quad (A4)$$

By changing terms, from left to right, the Eq. (A4) can be re-written as followings:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

$$\frac{n_P^*}{V_o}-\frac{n_P^*}{V^*}\frac{RT}{\pi}\frac{n_P^*}{V_o}=\frac{n_P^*}{V^*}-\frac{n_P^*}{V^*}\frac{RT}{\pi}\frac{n_{P,o}}{V_o} \quad \text{(A5)}$$

By re-arranging the Eq.(A5), one results:

$$\frac{N_P^*}{V_o}\left(1-\frac{N_P^*}{V^*}\frac{RT}{\pi}\right)=\frac{N_P^*}{V^*}\frac{RT}{\pi}\left(1/\frac{RT}{\pi}-\frac{N_{P,o}}{V_o}\right) \quad \text{(A6)}$$

By introducing the invariant $\frac{RT}{\pi}=\frac{1}{\sum_{j=1}^{ns}C_{jo}}$ from the right side of Eq.(6), and by substituting with $C_P^*=n_P^*/V^*$, one obtains:

$$\frac{N_P^*}{V_o}\left(1-C_P^*\frac{RT}{\pi}\right)=C_P^*\frac{RT}{\pi}\left(\sum_{J=1}^{ns}C_{j,o}-C_{P,o}\right) \text{(A7)}$$

Relationship Eq.(A7) is identical with Eq.(A1). On the other hand, by using the Eq.(6), the volumes for the initial and final states, and for the same pressure π, are the followings:

$$V_o=\frac{RT}{\pi}\sum_{j=1}^{all}N_{j,o}, and\, V^*=\frac{RT}{\pi}\sum_{j=1}^{all}N_j^*, \text{(A8)}$$

By dividing the two volumes of Eq.(A8), one obtains:

$$V^*/V_o=\left(\sum_{j=1}^{ns}N_j^*/\sum_{j=1}^{ns}N_{j,o}\right) \text{(A9)}$$

Eq.(A9) proves, by using the isotonic WCVV formulation, the cell volume variation when a variation in the species copynumbers occurs. It is also to observe that the sum of species concentrations into the cell (taken individually, or lumped) is a conservative term, as proved by the Eq.(6) combined with Eq.(A9), that is:

$$\sum_{j=1}^{ns}C_j^*=\frac{\sum_{j=1}^{ns}N_j^*}{V^*}=\frac{\sum_{j=1}^{ns}N_j^*}{V_o\left(\sum_{j=1}^{all}N_j^*/\sum_{j=1}^{ns}N_{j,o}\right)}=\frac{\sum_{j=1}^{ns}N_{j,o}}{V_o}=\sum_{j=1}^{ns}C_{j,o}=constant \quad \text{(A10)}$$

2.2.2. Simplified representations of the protein synthesis (gene expression regulatory modules, GERMs), and of the genetic regulatory circuits (GRCs) under a WCVV modelling framework

When developing deterministic kinetic models for the reactions included in the central carbon metabolism (CCM), or for other cell metabolic processes, to be further used for GMO design, an important aspect is to also include math (kinetic) models of individual GERMs characterizing the gene expression control of the enzymes production. Also, by linking the interfering GERMs modules, complex GRC regulatory chains can thus be obtained [Maria, 2017a,2017b,2018; Blass et al., 2017].

Because the GRC-s are responsible for the control of the cell metabolism, the adequate kinetic modelling of the constitutive GERMs, but also the adequate representation of the linked GERM regulatory efficiency in a GRC is an essential step in describing the cell metabolism regulation via the hierarchically organized GRCs (Figure 2-8) (where keyproteins play the role of regulatory nodes). Eventually, such models allow simulating the metabolism of modified cells.

Development of dynamic models to adequately reproduce such complex synthesis related to the central carbon metabolism (CCM) [Kurata and Sugimoto, 2017; Chassagnole et al., 2002; Miskovic et al., 2015], but also to the GRC tightly controlling such metabolic processes reported significant progresses over the last decades in spite of the lack of structured experimental kinetic information, being rather based on sparse information from various sources and unconventional identification / lumping algorithms [Maria, 2004, 2005, 2006,2007,2008,2009,2014b]. However, such structured models are extremely useful for the *in-silico* design of novel GRCs conferring new properties/functions to the mutant cells, in response to external stimuli [Heinemann and Panke, 2006; Salis and Kaznessis, 2005; Kaznessis, 2006; Atkinson et al., 2003; Klipp et al., 2005; Chen and Weiss, 2005; Tian and Burrage, 2006; Sotiropoulos and Kaznessis, 2007; Tomshine and Kaznessis, 2006; Zhu et al., 2007; Maria et al., 2011].

A central part of cell metabolic models concerns self-regulation of the metabolic processes via GRCs. Consequently, one particular application of such dynamic cell models is focus on the study of GRCs properties, in order to predict ways by which biological systems respond to signals, or environmental perturbations. The emergent field of such efforts is the so-called '*gene circuit engineering*' (GCE)

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

and a large number of examples have been reported with *in-silico* re-creation of GRCs conferring new properties/functions to the mutant cells. By using simulation of gene expression, the GCE *in-silico* design GMO that possess specific and desired functions. By inserting new GRCs into organisms, one may create a large variety of mini-functions / tasks (or desired 'motifs') in response to external stimuli [Heinemann and Panke, 2006; Buchholz and Hempel, 2006; Hempel, 2006].

"With the aid of recombinant DNA technology, it has become possible to introduce specific changes in the cellular genome. This enables the directed improvement of certain properties of microorganisms, such as the productivity, which is referred to as *Metabolic Engineering* [Bailey, 1991; Nielsen, 1998; Stephanopoulos et al., 1998]. This is potentially a great improvement compared to earlier random mutagenesis techniques, but requires that the targets for modification are known. The complexity of pathway interaction and allosteric regulation limits the success of intuition-based approaches, which often only take an isolated part of the complete system into account. Mathematical models are required to evaluate the effects of changed enzyme levels or properties on the system as a whole, using the metabolic control analysis (MCA) or a dynamic sensitivity analysis" [Visser et al., 2004]. In this context, GRC dynamic models are powerful tools in developing re-design strategies of modifying genome and gene expression seeking for new properties of the mutant cells in response to external stimuli [Heinemann and Panke, 2006; Salis and Kaznessis, 2005; Kaznessis, 2006; Atkinson et al., 2003; Klipp et al., 2005; Chen and Weiss, 2005; Tian and Burrage, 2006; Sotiropoulos and Kaznessis, 2007; Tomshine and Kaznessis, 2006; Zhu, 2007; Maria et al., 2011; Maria, 2010; Maria and Luta,2013].

Examples of such GRC modulated functions include:

- 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 external inducers;
- Hysteretic GRC behaviour, that is a bio-device able to behave in a history-dependent fashion, in accordance to the presence of a certain inducer in the environment;
- GRC oscillator producing regular fluctuations in network elements and reporter proteins, and making the GRC to evolve among two or several quasi-steady-states;
- Specific treatment of external signals by controlled expression such as amplitude filters, noise filters or signal / stimuli amplifiers;
- GRC signalling circuits and cell-cell communicators, acting as 'programmable' memory units.

Modelling individual GERMs:

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 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 cascade regulation (of the gene transcription and translation) 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. [Maria, 2005,2017a].

To model such a complex metabolic regulatory mechanisms at a molecular level with using ordinary differential equations (ODE) kinetic models, Sewell et al [2002], Savageau [2002], Hlavacek and Savageau [1997]; Maria [2003,2005] proposed simple mechanistic structures by using a modular approach ssee the gene expression regulatory modules (GERM) library of Maria [2005,2017a] in Figure 2-12}, useful in simulating the hierarchical organization of cell regulatory networks (Figure 2-8, right).

Concerning the protein synthesis, this process is presumably regulated by a complex homeostatic mechanism that controls the expression of the encoding genes. On the other hand, cells contain a large number of proteins of well-defined functions [Maria, 2017a,2005], 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 GERM 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. 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 networks/circuits (GRC) by applying certain principles and rules to be further discussed, and then adjusting the network global properties. Accordingly, it is desirable to focus the metabolic regulation and control analysis on the regulatory/control features of functional GERM subunits than to limit the analysis to only kinetic properties of individual enzymes acting over the synthesis pathway.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The modular approach to analyse the gene expression assumes that the reaction mechanism and stoichiometry of various types of individual GERMs are known, while the involved species are completely observable and measurable. Such a hypothesis is rarely fulfilled due to the inherent difficulties in generating reliable experimental (kinetic) data for each individual metabolic subunit. However, incomplete kinetic information can be incorporated by applying suitable lumping algorithms [Maria, 2004], or by exploiting the cell and module global optimal properties during identification steps to build-up GERM kinetic models [Maria and Scoban, 2017,2018]. 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 (see the below section 2.2.4).

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 (see section 2.2.1.3).

When elaborating a protein synthesis regulatory module (i.e. a gene expression regulatory module, GERM), different degrees of simplification of the process complexity can be followed. A GERM is a semi-autonomous regulatory group of reactions and species, linked to efficiently cope with cell perturbations, to adequately ensure the system homeostasis, and an efficient gene expression.

For instance, the gene expression (see schema of Figure 2-7) 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 2-7, up), the protein (P) synthesis rate can be adjusted by the 'catalytic' action of the encoding gene (G). The catalyst activity is in turn allosterically regulated by means of 'effector' molecules (O,P or R) reversibly binding the catalyst G via fast and reversible reactions (the so-called 'buffering' reactions). These simple regulation schemes can be further detailed in order to better reproduce the experimental data, with the expense of a supplementary effort to identify the module kinetic parameters (Figure 2-12). For instance, a two-step cascade control of P -synthesis model also includes the M = mRNA transcript encoding P (Figure 2-7, center). The effector (O), of which synthesis is controlled by the target protein P (Figure 2-7, center), can allostericaly adjust the activity of G and M, i.e. the catalysts for the transcription and translation steps of the gene expression. In such a cascade schema, the rate of the ultimate reaction is amplified, depending on the number of cascade levels and catalysis rates (Figure 2-7, down). More complex regulatory modules have been elaborated [Maria,2007,2017a; Maria and Luta,2013], and used in developing genetic regulatory circuits GRC) 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 (see section 2.2.4). Once elaborated, such a modular structure can be modelled by using a continuous variable ODE kinetic model constructed under a WCVV modelling framework, and then analysed as functional efficiency by means of some quantitative performance indices (P.I.-s) below described (section 2.2.3).

Adequate WCVV kinetic models of GERMs not only can reproduce their regulatory properties, but also can be used to reproduce the regulatory efficiency of linked GERMs in GRCs. Such GRCs are essential parts of a cell model, by describing the cell metabolism regulation via the hierarchically organized GRCs (where key-proteins play the role of regulatory nodes). Also, it is to be underlined that in such GERM models (Figure 2-12), the gene expression is a highly self- / cross-regulated and mutually catalyzed process by means of the produced enzymes / effectors.

As the cell regulatory systems are organized on a module-based, complex feed-back and feed-forward loops are employed for self- or cross-activation / repression of interconnected GERM-s, leading to different interaction alternatives (directly/inversely, perfect/incomplete, coupled/uncoupled connections) of a gene with up to 23-25 other genes, [Maria, 2014], to ensure the key-species homeostasis, holistic and local regulatory properties of the enzymatic reactions. While Maria [2003,2005,2007,2009,2014], Sewell et al [2002], Savageau [2002], Hlavacek and Savageau [1997] used reduced GERM-s structures of up to 10-14 reactions, that ensures a satisfactory trade-off between model simplicity and its predictive quality [Maria,2006], more sophisticated constructions are proposed in the literature [Maria, 2014b, 2017a] (Figure 2-12).

As an example, Kaznessis [2006], and Salis and Kaznessis [2005] designed a bistable genetic switch, by using two gene expression modules GERM-sextracted from the *lac operon* of *E. coli*. The transcriptional regulation is modelled by using a stochastic approach accounting for 40 reactions and 27 species (for the reduced kinetic model) or 70 reactions and 50 species (for the extended kinetic model). Such a regulatory schema (Figure 2-7), including dimeric self-repressors (PP, or OO; [Maria, 2006]) and mutual repression following the presence in excess of one of the activating inducers, can also be illustrated by means of simpler representations of Yang et al. [2003], and Maria [2003,2005,2017a]. The advantage of such a modular approach is the possibility to adapt the model size according to the available information, or to use the same GERM structure to model several gene expressions [Maria and Luta, 2003]. Modular approach can also be useful in simulating the hierarchical organization of the cell regulatory networks.

The modular approach assumes that the reaction mechanism and stoichiometry of the kinetic module are known, while the involved species are completely observable and measurable. Such a hypothesis is rarely fulfilled due to the inherent difficulties in generating reliable experimental (kinetic) data for each individual metabolic subunit. However, incomplete kinetic information can be incorporated by performing a suitable model lumping, or by exploiting the cell and module global properties during identification

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

steps [Maria and Scoban, 2017, 2018]. The regulatory modules can be constructed relatively independent to each other, but the linking procedure has to consider common input/output components, common linking reactions, or even common species. Rate constants can be identified separately for each module, and then extrapolated when simulating the whole regulatory network, by assuming that linking reactions are relatively slow comparatively with the individual module core reactions. In such a manner, linked modules are able to respond to changes in common environment and components such that each module remains fully regulated. (see below section 2.2.4)[Maria, 2005, 2017a]. The advantage of such a modular 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. [Maria, 2009; Maria and Luta, 2013].

To easily study and compare GERM-s regulatory efficiency, Sewell et al. [2002], Yang et al. [2003], and Maria [2003, 2005, 2006, 2007, 2009, 2014, 2017a]; [Maria and Luta,2013] proposed various types of hypothetical GERM-s simplified reaction pathways designed to ensure homeostatic regulation of a generic protein-gene (P/G) pair synthesis (see the GERM "library" of Maria [2003, 2005, 2017a]), with a large number of exemplifications from *E. coli* (some of them are displayed in Figure 2-7, Figure 2-25).

As experimentally proved in the literature (review of Maria [2017a, 2017b, 2018]), the genetic regulatory circuits (GRC), that control the synthesis of all proteins (enzymes) in the cell, present a modular construction, every operon (a cluster of genes under the control of a single promoter) including a variable number of interacting GERMs. Examples of GRCs includes genetic switches, operon expression, genetic amplifiers, or filters, etc. However, it is well-known that one GERM interacts with no more than other 23-25 GERMs [Kobayashi et al., 2004], while most of GERM structures are repeatable. Consequently, in developing the *in-silico* analysis of a GRC dynamics, the WCVV modular approach is preferred due to several advantages: 1) A separate analysis of the constitutive GERM-s in conditions that mimic the stationary or perturbed cell growth. 2) The modules are then linked to reconstruct the GRC of an optimized regulatory efficiency 3) in order to ensures key-species homeostasis and cell network holistic properties (Figure 1-56, Figure 1-54, Figure 2-8). Investigation of GERMs and GRCs characteristics is focus on the tight control of gene expression, the quick dynamic response to perturbations, the high sensitivity to specific inducers, and the GRC robustness (i.e. a low sensitivity *vs.* undesired inducers). Such advanced regulatory structures must ensure the homeostasis (quasi-stationarity) of the regulated key-species, and quick recovery (with a trajectory of minimum amplitude) after a dynamic (impulse-like) or a stationary (step-like) perturbation of one of the involved metabolites or nutrients [Maria, 2003,2005,2006,2007,2008,2009,2014b, 2017a,2017b,2018]) (Figure 2-9, Figure 2-10).

To not complicate the deterministic models, lumped GERM and GRC structures have been adopted in the literature. Some of them are presented in the Figure 2-11, and Figure 2-12. The simplest GERM structure with one regulatory element is those denoted by [G(P)1]. The generic [G(P)1] regulatory module, schematically represented in (Figure 2-24), and in (Figure 2-12, the top-row), refers to the synthesis of a generic protein P and the simultaneous replication of its encoding gene G. The lumped [G(P)1] model (Figure 2-24) includes only one regulatory element (a so-called "effector", that is a fast "buffer" reversible reaction $G + P <===> GP$ (inactive), aiming at controlling the P synthesis rate and its homeostatic (quasi-stationary, QSS) level. In such a generic lumped construction, the protein P and its encoding gene G mutually catalyses the synthesis of each other. The protein P is the "control node" playing multiple roles in such a simplified lumped representation. Thus, (see (Figure 2-11-down-right, and Figures 2-26) P is a permease leading to the import of nutrients NutG, NutP in the cell, but also a metabolase converting the nutrients into precursors MetG and MetP of the G and P respectively. Protein P is also a polymerase catalysing the gene replication. And, finally, the protein P is also a transcriptional factor (TF) by dynamically adjusting the catalytic activity of the G by means of a very rapid "buffer" regulatory reaction $G + P <===> GP$ (inactive). When P is produced in excess, it reversible inactivates more amount of G, which in turn, will slow-down the P synthesis. When P is produced in too low amounts, the regulatory process goes backwards.

Because the GERM structures are repeatable in most of cells, Maria [2003,2005,2006,2007,2009,2017a,2017b,2018, 2023, 2023a, 2024, 2024b, 2024c] developed a math "LIBRARY" including "template" kinetic models of the main GERM types (see Figure 2-12), under the WCVV modelling framework.

The LIBRARY of lumped individual GERMs modules is useful to build-up dynamic models of GRC (genetic regulatory circuits), by linking a large number of GERM-s, useful for *in-silico* design of GMOs. The LIBRARY allows to use the novel WCVV math modelling framework to define novel quantitative performance indices (P.I.-s) in order to characterize and to directly compare the regulation efficiency of the individual GERMs related to the internal/external perturbations [dynamic ("pulse-like"), or stationary("step-like")]. The use of the GERM library and their associated P.I.-S allows an easier application of a "**buiding-blocks**" strategy and **rules** to connect GERMs to build-up GRCs when designing GMOs. Of desirable characteristics. Besides, the GERMs library was proved to be very useful to study various properties of the cell self-regulation effects, impossible to be highlighted by using classical cell WCCV models (see sections 2.2.3, and 2.3).

The module nomenclature used in Figure 2-12 for the GERM models, proposed by Yang et al. [2003], and by Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018] is the following:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

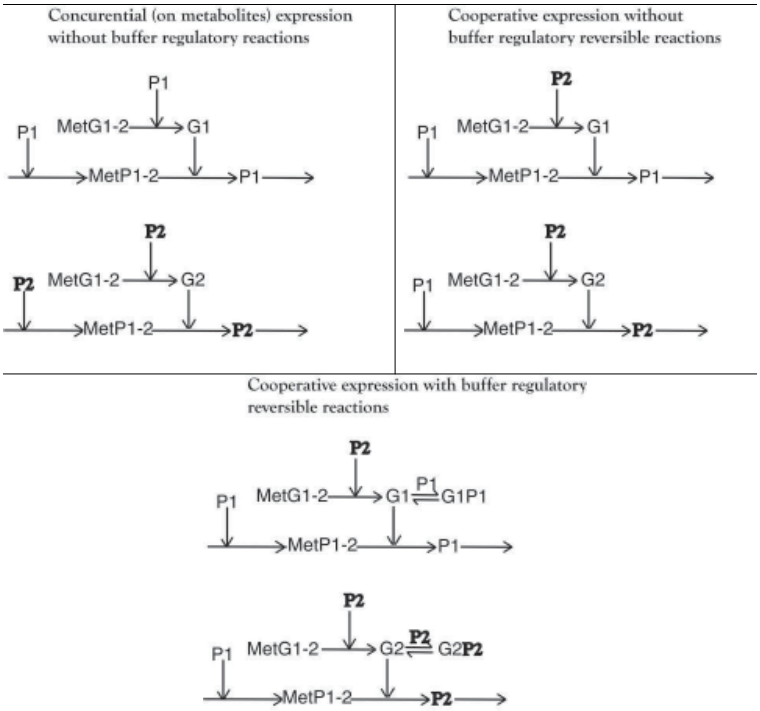


Figure 2-25: Investigate several alternatives to link two GERM-s. **[Up-left]**. **Competitive** (concurrential on common metabolites), linking [G1(P1)0] + [G2(P2)0]; **[Up-right]. Simple cooperative** linking of [G1(P1)0] + [G2(P2)0] modules. P1 is permease and metabolase for both GERM-s; P2 is polymerase for replication of both G1 and G2 genes. **[Down]. Cooperative** linking with buffer reversible regulatory reactions to modulate the G1 and G2 catalytic activity in the modules [G1(P1)1] + [G2(P2)1]. The horizontal arrows indicate reactions; vertical arrows indicate catalytic actions. The absence of a substrate or product indicates an assumed concentration invariance of these species.

$$\left[\mathrm{L_1(O_1)n1;\ldots;L_i(O_i)ni}\right] \tag{19}$$

It includes the assembled regulatory units $L_i(O_i)n_i$. One unit " i " is formed by the component $L(i)$ (e.g. enzymes P or even genes G,M , etc.) at which regulatory element acts. By contrast, $n(i)$ = 0,1,2,... is the number of 'effector'/ TF species $O(i)$ binding the 'catalyst' L. The "effectors" can be P, PP, PPPP, R, RR, RRRR, etc. For instance, a [G(P)2] module of GERM of Figure 2-12 includes two successive binding steps of G with the product P , that is $\mathrm{G+P <==> GP+P <==> GPP}$, all intermediate species GP, GPP, being catalytically inactive, while the mass conservation law is all time fulfilled, i.e. $\sum_{i=0}^{2}\left[G(P_i)\right]=cons\tan t$. Such a representation accounts for the protein concentration diminishment due to the cell-growth dilution effect, but could also include protein degradation by proteolysis. It is also to observe that such GERM models, even simple, try to account 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 the Figure 2-7, 2-8, 2-11, 2-12. For instance, in the simplest [G(P)1] module case, the protein P synthesis is formally catalysed by its encoding gene G . In turn, P protein formally catalyse the G synthesis, but also modulate the G catalyst activity (via the fast buffering reaction $\mathrm{G+P <===> }GP$.

In fact, the repeatable GERM-s structure facilitate the construction of a math "LIBRARY" of the main GERMs types. Maria [2003,2005,2017a] proposed such a "LIBRARY" of the main GERMs . Moreover, he developed simplified (lumped) template kinetic models for these main GERM-s types, ready to be used to simulate individual gene expression [Maria,2017a,2018d], but also to construct GRC-s by linking several GERM-s (see the mercury-operon of 5 linked GERM-s as an example, [Maria and Luta,2013]). The GERM-s kinetic models are constructed to fulfill the chemical and biochemical process engineering (ChBPE) concepts/rules (section 1.3.4). On the other hand, the WCVV "whole-cell" modelling approach, and the kinetic simulation of hypothetical "mechanistic cell" will fulfill the local regulatory properties of the P-synthesis, but also the cell holistic properties that maintain intracellular species homeostasis while growing auto-catalytically on environmental nutrients present in variable amounts. However, unlike the chemical reactions, most of the cell reactions are catalyzed by enzymes (proteins). Such a particularity involves a lot of advantages: a) the reactions are very selective, and of high conversion; b) the reaction rate can be "adjusted" not only from the substrate concentrations, but also by adjusting the biocatalyst (enzyme) concentration. Such an alternative can be done by over-expression of the encoding gene. For instance, by adding a certain number of plasmids in the cell nucleus (see the example of Maria and Ene [2013]). c) A certain reaction can be removed, by removing the biocatalyst synthesis, that is by removing the corresponding encoding gene from the genome ("gene knock-out" rule).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

As proved by Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018], the GERMs performances indices (P.I.-s, see section 2.2.3) are as better as the number of effectors increases. For instance, for the GERM of [**G(P)n**], or of [**G(PP)n**] types in Figure 2-12 , the dynamic regulatory efficiency (expressed here as the recovery rate after a dynamic impulse-like perturbation in the P key species) are as better as the number " n " of buffer reactions increases (Figure 2-13). Of course, also the effector type is important. As proved in section 2.2.3 the [**G(PP)n**] effectors, with a dimeric TF are better than those of [**G(P)n**] type [Maria,2009,2017a,2018].

Also, Maria [2005,2006,2007,2008,2009,2017a,2017b,2018] proved that when P is acting as a TF, its efficiency is better if it is present in a dimeric form (PP), that is in GERMs of [G(PP)n], or [M(PP)n] type in Figure 2-12, and Figure 2-13.

Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018] also proved that the GERM regulatory efficiency is better if TF=PP is acting at both G and M levels of the expression (middle and downrows of Figure 2-12), thus developing a cascade control scheme of the expression, where the transcription and the translation regulatory steps are considered separately in the WCVV kinetic model, that is GERMs of [**G(PP)n; M(PP)n′**] type.

It clearly appears (Figure 2-9) that, as the number of effectors increases in the GERM kinetic model, as its performance regulatory indices (P.I.-S) are better.

Perturbations of the species steady-state (homeostatic) concentrations are caused by environmental processes. In a GERM case, these processes tend to increase or decrease the key-proteins stationary level, generically denoted by [P]s. These processes occur in addition to those of the "core" system (G/P replication over the cell cycle).

The GERM regulatory performance indices P.I.-s are of two types, according to the classification of Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018] (Figure Figure 2-14 , and Figure 2-15), that is, in the response to: 1) stationary (slow, supported, "step"-like) perturbations, and 2) dynamic (quick, unsupported, "impulse"-like) perturbations. Briefly they are presented in the Table 2-2, together with the associated optimization objective, for a general nonlinear dynamic cell WCVV model described, in a compact mode, by the Eq. (20), derived from Eq.(11), and completed with Eq.(4). Detailed information is given by Maria, [2003,2005,2006,2007,2008,2009,2017a,2017b,2018].

$$\begin{aligned} &d\boldsymbol{C}/dt = \boldsymbol{h}(\boldsymbol{C},\boldsymbol{k}); \boldsymbol{C}(t=0) = \boldsymbol{C}_S; d\boldsymbol{A}/dt = \boldsymbol{J}_C \times \boldsymbol{A}; \boldsymbol{A}(t=0) \\ &= \boldsymbol{I}; \boldsymbol{J}_C = (\partial \boldsymbol{h}(\boldsymbol{C},\boldsymbol{k})/\partial \boldsymbol{C})_s \end{aligned} \tag{20}$$

Where C = species concentration vector; "s" index = stationary (QSS) value; k = rate constants vector; J_c = the Jacobian matrix of the WCVV kinetic model Eq.(11), that is Eq.(4). I = the identity matrix.

To better point-out the main GERM-s types of Figure 2-7, Figure 2-25, and Figure2-12, some supplementary observations are necessary.

These simplified GERM-s representations include the essential nutrient lumps (NutP, NutG), metabolites (MetP, MetG), and intermediates involved in the reactions controlling the transcriptional and translation steps of the G expression and P synthesis, and "to mimic" the cell content load ("ballast"), and the environmental influence.

The GERM-s modules nomenclature, proposed by Yang et al. [2003] of type Eq.(19), that is: $[L_1(0_1)nl;...;L_i(0_1)$, includes the assembled regulatory units $^{*}\mathbf{L}_i(\mathbf{O}_i)\mathbf{ni}$ ". One unit " **i** " is formed by the component $\mathbf{L}_i$ [that is enzymes, and TF-s such as P , gene G , or M (mARN), etc.] at which regulatory element acts, and ni = 0,1,2,... number of 'effectors', that is TF species denoted here by O_i (that is P, PP, PPPP, O, OO, OOOO, R, RR, RRRR, etc.) binding the 'catalyst' L (Figure 2-7, Figure 2-25, and Figure 2-21). For instance, a [G(P)5] unit of GERM includes five successive binding steps of G with the product P , that is:

$$G + P \Leftrightarrow GP + P \Leftrightarrow GPP + P \Leftrightarrow GPPP + P \Leftrightarrow GPPPP + P \Leftrightarrow GPPPPPP,$$

all intermediate species GP, GPP, GPP, GPPP, GPPPP, GPPPPP being inactive catalytically, while the mass conservation law is all time fulfilled, i.e. $\sum_{i=0}^{5}\left[G(P_i)\right] = \text{constant}$. Such a representation accounts for the protein concentration diminishment due to the cell-growth dilution effect, but could also include protein degradation by proteolysis. The [G(P)n] units of GERM, even less realistic (Figure 2-7), represent the simplest GERM-s used as control mechanism against which all others are compared. In the simplest [G(P)0] kinetic model of a GERM module (see Figure 2-26), there are only two main synthesis chains. That is, P is a permease that catalyses the import of NutG and NutP from the environment, and a metabolase that converts them into cellular metabolites MetG and MetP. P it is also a polymerase that catalyses the synthesis of G from MetG. Gene G, symbolizing the genome of the cell, functions as a "catalyst" for the synthesis of P from MetP. The result is that G and P syntheses are mutually auto-catalytic. In this [G(P)0] kinetic model of GERM module there are no regulatory elements (no buffering reactions to control the G activity). By comparison, in the [G(P)1] kinetic model of GERM module (Figure 2-24), the negative feedback control

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

of transcription is realized by P itself (as effector), via a rapid buffering reaction, $G + P \Leftrightarrow GP$, leading to the catalytically inactive GP complex. As proved by Maria [2003,2005,2017a], the maximum regulatory efficiency at steady-state (index 's') corresponds to $[G]_S / [G]_{total} = 1/2$, when the maximum regulation sensitivity *vs.* perturbations in $[P]_s$ is reached [Sewell et al., 2002]. Further allosteric control of G activity, leading to inactive species $[GP_n]$, amplifies the regulatory efficiency of such a $[G(P)_n]$ module.

Experimental proofs of GERM-s model structures:

As an experimental evidence of such GERM model structures, prokaryotes commonly bind multiple copies of transcription factors as a means of promoting cooperative effects and thus improving regulatory effectiveness [Yang et al., 2003]. For instance, *dnaA* is an auto-regulated protein and at least five copies can bind to *dnaA* gene in *E. coli*. [Yang et al. 2003; Speck et al. 1999]. Several similar proofs have been reviewed by Yang et al. [2003]. The [G(PP)n] units reflect better the regulatory loops in which multiple copies of effectors (proteins and transcription factors TF-s) bind to promoter sites on the DNA to control the expression of gene G encoding P (see exemplifications from *E. coli* by Yang et al. [2003]. The control is better realized by including the TF as a dimmer PP. This modification, involves a supplementary P dimerization step before the buffering reactions, by keeping a low concentration level of the PP intermediate to reduce substrate consumption, and to keep fast regulatory reactions inside any of GERM. This explains why most of transcription factors bind as oligomers (typically dimers or tetramers) and why they typically bind in multiple copies [Maria, 2005; Yang et al.,2003; Ptashne, 1992].

By contrast to the other series of GERM-s from the "LIBRARY" of Figure 2-12, the $[\mathbf{G(P)n;M(P)n}l]$, or $[\mathbf{G(R)n;M(R)n}l]$, or $[\mathbf{G(O)n;M(O)n}l]$ units (see Figure 2-12, and Figure 2-7 up-right, and center) tries to reproduce more accurately the transcription / translation cascade of reactions during the gene expression, by including an allosteric control at two levels of catalysis: on G (i.e. DNA) and on M (i.e. mRNA). In these GERM modules, M is synthesized from nucleotides under G catalysis (transcription), and then, P is synthesized in a reaction catalyzed by M (translation). Such a supplementary control of mRNA activity is proved to be a more effective means to regulate the protein synthesis. [Maria, 2003, 2005; Yang et al.,2003; Hargrove and Schmidt, 1989].

2.2.3. The regulation efficiency quantitative indices (P.I.-s) of individual GERMs under the novel WCVV modelling framework

By using the theory, concepts, rules, and numerical algorithms of the ChBRE (section 2), and those of the nonlinear systems, Maria [2005,2006,2007,2008,2009,2017a,2017b,2018,2023] defined in a quantitative way, by using simple mathematical (kinetic) models written under the novel WCVV modelling framework, the regulatory performance indices P.I.-s of GERM-s / GRC-s.

According to Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018], the regulatory indices P.I.-s of GERM-s / GRC-s are of two types (Figure 2-14, and Figure 2-15):

A). Those corresponding to a ***stationary*** (continuously supported, "step"-like) perturbation, in the environment or inside the cell (Figure 2-15);

B). Those corresponding to a *dynamic* (very short, "impulse"-like) perturbation, in the environment or inside the cell (Figure 2-15).

Briefly the P.I.-s of GERMs (of Figure 2-12) are presented in the (Table 2-2), together with the associated optimization objective (goal), for a general nonlinear dynamic cell model described by Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018]; [Maria and Scoban, 2017,2018]. See also an intuitive display in (Figure 2-14, and Figure 2-15), and the detailed information given by Maria [2003,2005,2006,2007,2008,2009,2017a,2017b,2018]; [Maria and Scoban, 2017,2018]. According to the mentioned references, the main P.I.s of GERMs are given in the Table 2-3. They significance will be explained in this section, by using simple and generic GERMs.

Most of the species steady-state (homeostatic) concentrations perturbations (dynamic or stationary) are caused by environmental perturbations transmitted inside cell, once a large number of nutrients are transported through the cell membrane [Alberts et al., 2002]. In an individual GERM case, these perturbations (dynamic "impulse"-like, or stationary "step"-like) tend to increase or decrease the key-protein stationary level [P]s. These perturbation processes occur in addition to those of the "core" GERM reaction system (genome/proteome G/P pair replication over the cell cycle).

A) Stationary P.I.

These P.I.-s characterize the GERM-s regulatory efficiency of an individual gene expression (see the model types in the LIBRARY of Figure 2-12) when coping with a *stationary* (continuously supported, "step"like) perturbation, in the environment or inside the cell (Figure 2-15).

The stationary P.I.s are defined as response to a stationary perturbation (Figure 2-15), that is transition from a QSS to another QSS following a step-like perturbation of one cell component concentration.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Stationary perturbations refer to permanent modifications in the homeostatic levels of the external nutrients or of the internal metabolites, leading to new stationary species (components) concentrations inside the cell. Referring to a target protein P in a generic GERM, of type [G(P)1] the regulatory module tends to diminish the deviation [P] s-[P]ns between the 'nominal' QSS (unperturbed set-point, of index 'ns') and the new QSS reached after perturbation (the new setpoint [P]s see Figure 2-15). Equivalently, the P -synthesis regulatory module will tend to maintain [P] ns within certain limits, $[P]\min \le [P]ns \le [P]\max$ (a relative $R_{ss} = \pm 10\%$ maximum deviation has been proposed by Sewell et al. [2002], and by Yang et al. [2003] (see Table 2-2, and Maria [2003]) to obtain an effective GERM. A measure of the species " *i* " steady-state concentration (Ci,s) "resistance" against various stationary perturbations { in the rate constants, kj, or in nutrient concentrations, [Nut,j] } is given by the magnitude of relative sensitivity coefficients at QSS, i.e. S(Ci;Kj) and S(Ci;Sj) [Nut,j]) respectively, where *S* (state; perturbation) = ∂(state) / ∂ (perturbation), denotes the absolute state sensitivities *vs.* perturbations [Varma et al., 1999].

An example of the effect of a stationary perturbation in the bioreactor environment is given by Maria and Luta [2013], and by Maria [2010] in Figure 2-18, for the case of mercury uptake from wastewaters by *E. coli* cells. Thus, Figure 2-18 presents the predictions, obtained with a hybrid SMDHKM model of the cell species concentration dynamics as a response to a "step"-like perturbation of the mercury level in the bioreactor inlet. Such a stationary external perturbation will lead to another QSS homeostasis for the all species considered in the mercuryoperon expression model of Figure 2-19. Very few details about the mercury-operon GRC model are given in the caption of Figure 2-19. More details are given by [Maria, 2010; Maria and Luta, 2013], and in the section 3.

A-i) Transition time

This P.I.-s refers to the transition time necessary to each GERM component to reach the new stationary concentration (quasi-steady-state, QSS), starting from the old one, when a "step-like" continuous and supported perturbation in one (individual or lumped) component of that GERM module. The measure of this P.I. of the GERM efficiency is related to the capacity of a key-species QSS to move fast to a new QSS. For instance, for the key-protein P of a GERM this P.I. translates in the duration of the *transition time* τ(P) (in the case of **P**-species in Figure 2-15) necessary to a certain component to reach the new steady-state concentration.

For a GERM of high efficiency, this transition time after a perturbation has to be minimum, to require less resource consumptions, and to restore the balanced cell growth. By referring to the example of Figure 2-18, and Figure 2-19 of mercury uptake from wastewaters by *E. coli* cells, it is to observe that the species transient times from the old to the new QSS are very short (around 2-3 cell cycles). Such a simulation result indicates that the GRC of the mercury-operon of Figure 2-19 ensures a very effective regulation *vs.* stationary environmental perturbations.

A-ii) Responsiveness to perturbations

This P.I.-s refers to the 'responsiveness' of a certain species QSSlevel of an analysed GERM or GRC *vs.* an exo/endogeneous signaling species. This P.I. can be represented by the small transient times necessary for a species 'j' QSS-level to reach a new QSS (with a certain tolerance, usually 1% - 5%) after applying a stationary external stimulus [Maria, 2009]. Consequently, this P.I. measures the GERM efficiency to move fast to a new QSS if a stationary continuously supported perturbation occurs. Similarly to the "A-i)" index, this efficiency is given by the duration of the *transition time* τ_p (in the case of P-species in Figure 2-15) necessary to a certain component (i.e. key-species P here) to reach the new steady-state concentration. An equivalent regulatory P.I., is expressed by $A_{unsync} = k_{syn} \times k_{decline}$ (Table 2-2). This last index illustrates the maximum levels of (unsynchronized) stationary perturbations in the synthesis (index "syn") or consumption (index "decline") rates of a key-species "tolerated" by the cell within defined limits [Sauro and Kholodenko, 2004]. For instance, in the case of the P-species, these rate constants belong to the following synthesis and degradation / dilution lumped reactions [Maria, 2003, 2017a]:

$$\xrightarrow{k_{syn}} P \xrightarrow{k_{\deg}}$$

A-iii).- Stationary efficiency

This stationary P.I. is related to the small relative sensitivities:

$$S\left(C_{is}; \text{NutP}_{js}\right) = \left(\partial C_{is} / C_{i,ref} \,/\, \partial \text{NutP}_{js} / \text{NutP}_{j,ref}\right)$$

Of the key-species stationary level C(i,s) *vs.* changes in the external nutrient stationary level NutP(j,s). This index is simply denoted by $S^{i}_{Nut_j}$. [here C(i,ref), and NutP(j,ref) denote the reference concentrations of C(i,s) and [NutP(j,s) respectivelly]. These sensitivities are computed from solving a sensitivity nonlinear algebraic set obtained by assuming QSS conditions of the ODE kinetic model of the analysed GERM and known nominal species stationary concentrations *C* s. [see the Eq.(5,11,14) of section 2.1]. Then, differentiation of the model steady-state conditions, leads to the following Eq.(5), which is written down again as Eq.(21):

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

$$\left[\frac{\partial h_i(C,C_{Nut},k)}{\partial C_k}\right]_s\left[\frac{\partial C_k}{\partial C_{Nut\ j}}\right]_s+\left[\frac{\partial h_i(C,C_{Nut},k)}{\partial C_{Nut\ j}}\right]_s=0 \quad (21)$$

By solving the linear algebraic set Eq.(21), with the required derivatives evaluated preferably analytically (by using the Maple ™, or Matlab ™ softwares), or numerically, leads to derive the state absolute sensitivities *vs.* nutrient levels, around the stationary (index "s") conditions, that is [Maria, 2005, 2006, 2007, 2008, 2009, 2017a, 2017b]:

$$s\left(C_i;NutP_j\right)_s=\left(\partial C_i/\partial NutP_j\right)_s \quad (22)$$

A-iv) The steady-state Cs stability strength

This GERM modular self-regulatory systems properties are related to the strong capacity of the GERM to cope (to preserve the homeostasis) in spite of the external/internal small/large perturbations, by maintaining in tight limits the system main-species steady-state **Cs** and, in the case of large perturbations to ensure a very quick recovering path of the key-species. As with the all other P.I.-S, this GERM property is related to the GERM system intrinsic characteristics deriving from the kinetic model construction (that is the reaction pathway, and the considered individual/lumped species). Basically, as Max(Re(λ(i))) < 0 is smaller as the QSS (cvasi-steady-state) concentrations (Cs) of the GERM key-species are more stable. Here, the eigenvalues λ (i) of the model Jacobian matrix, Eq. (4), that is $\boldsymbol{J_C}=(\partial \boldsymbol{h}(\boldsymbol{C},\boldsymbol{k})/\partial \boldsymbol{C})_s$ of the GERM's ODE kinetic model are evaluated at a checked QSS of the species concentration vector (Cs), that is for the kinetic model Eq.(1a,b) or Eq.(11) written at QSS conditions, dCJ/dt=0. For the cell kinetic ODE model written in the WCVV general form of Eq.(11):

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

With:

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

Where: Nj = moles of species "j"; V= cell (cytosol) volume; t = time; sij = stoichiometric coefficients of the species "j" in the reaction "i"; ri = the "i-th" reaction rate; k = rate constant vector, Cj = species "j" concentration (individual or lumped); "s" index = at QSS; R = universal gas constant; T = absolute temperature; π = osmotic pressure. Then, the QSS formulation of the cell kinetic model becomes:

$$\left(\frac{dC_j}{dt}\right)_s=\left(\frac{1}{V}\frac{dN_j}{dt}\right)_s-D_{is}C_{js}=h_{js}\left(\boldsymbol{C}_s,\boldsymbol{k},t\right)=0;\left(\frac{1}{V}\frac{dN_j}{dt}\right)_s=r_{js}$$

Or:

$$\frac{d\boldsymbol{C}}{dt}=\boldsymbol{h}_{\boldsymbol{S}}\left(\boldsymbol{C}_{\boldsymbol{S}},\boldsymbol{k}\right)$$

Where:

$$D_{is}=\left(\frac{RT}{\pi}\right)\sum_j^{n_s}\left(\frac{1}{V}\frac{dN_j}{dt}\right)_s \quad (23)$$

$$\boldsymbol{J}_C=(\partial h(C,\boldsymbol{k})/\partial \boldsymbol{C})_s$$

Notations: Dis = the instant dilution rate, at stationary conditions; J_c = the model Jacobian natrix; ns = number of species considered in the WCVV kinetic model (individual, or lumped).

In a more systematic approach, the steady-state Cs stability strength can be associated to an index against periodic oscillations of keyspecies synthesis. This index can be evaluated from the linearized form of the cell system model Eq.(23), by calculating the monodromy matrix $A(T)$ after a checked period 'T' of time [Maria, 2003]. For a stable homeostasis Cs, that is for $|\lambda_A|<1$, as $|\lambda_A|$ are smaller, as the stability of the Cs state is stronger, and that QSS recovers faster after a small dynamic perturbation. In other words, QSS stability strength involves:

$$\text{Min (Max (Re}(\lambda(\mathrm{i}))\text{)); with Re}(\lambda(\mathrm{i}))<0\text{ for all 'i', and Min}|\lambda_A|<1. \quad (24)$$

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Where: Re(λ(i)) = the real part of λ(i) .

The monodromy matrix $A(t)$ is obtained by integrating the following ODE set concomitantly with the cell process model (the first line):

$$\frac{d\boldsymbol{C}}{dt} = \boldsymbol{h}(\boldsymbol{C},\boldsymbol{k}); \boldsymbol{C}(t=0) = \boldsymbol{C}_s; \frac{d\boldsymbol{A}}{dt} = \boldsymbol{J}_C \boldsymbol{A}; \boldsymbol{A}(t=0) = \boldsymbol{I} \tag{25}$$

Notations: $|\lambda_A|$ = the module of the eigenvalues of the A(T) matrix; I = identity matrix; $J_C = (\partial h(C,k)/\partial C)_s$ = the Jacobian of the GERM process model of Eq.(23).

B) Dynamic P.I.

All the dynamic P.I.-s of a GERM / GRC are defined in relation to the GERM/GRC response to a dynamic perturbation (Figure 2-15), that is the recover of the QSS following an "*impulse*"-like perturbation in the stationary concentration of one of the cell keycomponents.

In other words, the dynamic perturbations refer to instantaneous changes in the concentration of one or more cell components that arise from a process lasting an infinitesimal time ("impulse"-like perturbation). After perturbation, the cell sub-system recovers and returns to its stable nominal state QSS (see Figure 2-15, and Figure 2-10 for a generic **P**-protein case in a GERM of [G(P)1] type). The computed recovering time τ (rej, j) necessary to each component ' j ' to reach-back their stationary concentration Cjs (with a tolerance of 1% - 5% proposed by Maria [2005]) may differ from one species to another depending on how effective are their corresponding regulatory circuit to which it belongs.

Recovery rates are properties of all interactions within the cell system, rather than of the individual elements of the cell [Heinrich and Schuster, 1996]. Anyway, as pointed out by Maria [2005,2006,2007,2008,2009,2017a,2017b,2018], in a WCVV model formulation, all cell components are directly inter-connected via the osmotic pressure and the cell volume to which all species contributes, according to the Pfeffers' law Eq.(6), that is:

$$\left(\frac{RT}{\pi}\right) = 1/\sum_{j=1}^{ns} C_j = \text{constant} \tag{26}$$

Which, by derivation and division with V is leading to evaluation of the cell instant dilution rate [Maria, 2005, 2006, 2007]:

$$D_i = \frac{1}{V}\frac{dV}{dt} = \left(\frac{RT}{\pi}\right)\sum_{j=1}^{ns}\left(\frac{1}{V}\frac{dN_j}{dt}\right) \tag{27}$$

where:

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

.

In terms of the evolution and stability of the species stationary (QSS) concentrations included in a dynamic cell system expressed by an ODE kinetic model (of a GERM, or of a GRC), these properties can be evaluated from the analysis of the eigenvalues λ(i) (i = no. of species considered in the kinetic model) of the Jacobian matrix Eq. (23) of the linearized ODE model, that is $J_C = (\partial h\,(C,k)/\partial C)_s$.

If one considers small perturbations of a steady state **Cs** then, this steady state is asymptotically stable if the real parts of the

Jacobian eigenvalues are all negative, that is $\text{Re}(\lambda(\text{i})) < 0$, for all " **i** " [Heinrich and Schuster, 1996]. If the cell system is stable then, it reaches the same QSS after cessation of a dynamic ("impulse"-like) perturbation, or it reaches another QSS after cessation of a stationary ("step-like") perturbation. The recover trajectory $C(t)$, and the recovery time of kinetic model species (Figure 2-9), can be approximated with the solution of the linearized kinetic model Eq. (25) of the cell process [Maria and Scoban, 2017,2018]:

$$\boldsymbol{C}(t) = \boldsymbol{C_S} + \sum_{i=1}^{ns} d_i \exp(\lambda_i t); \tau_{recover} \approx 1/\text{Min}_i\left|\text{Re}(\lambda_i)\right| \tag{28}$$

Where: C = species concentration vector; t = time; λ(i) (i = 1,...,ns) the eigenvalues of the linearized kinetic model Jacobian matrix J_C = (∂h (C,k) / ∂C) in Eq.(23); bi = constants depending on the cell system characteristics at QSS conditions; ns = number of species considered (individually or lumped) in the cell model.

Here are to be mentioned the works of Maria [2003, 2005, 2017a], and Sewell et al. [2002] proving that the optimum concentrations in the '*buffering*' reactions of GERMs involving the active (G, M) and inactive (GP, GPP, MP, etc.) forms of the '*catalyst*' (that is G, M in Figure 2-12) that ensure the maximum regulation *dynamic efficiency vs.* perturbations are those of equal active and inactive forms of

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

the catalyst, that is: [G]s = [GP]s, [G]s=[GPP]s, [M]s = [MP]s, [M]s=[MPP]s, etc. for the GERM-s types displayed in Figure 2-12) ("s" index denotes the steady-state). The main dynamic P.I.-s discussed by Maria [2005, 2006, 2007, 2008, 2009, 2017a, 2017b, 2018], and by Maria and Scoban [2017, 2018] are the followings (Figure 2-15).

B-a) Recovering time

This dynamic P.I. of a GERM, denoted by τ (rec, j), or simply by τ (j), is the time necessary to each GERM's component to return to their stationary (QSS) concentration (Cs) (with an assumed tolerance of 1%, as proposed by Maria [2005]), after cessation of an "impulse"like perturbation in one of the GERM kinetic model component. This P.I. is a measure of the GERM efficiency in fast recovering the keyspecies stationary concentrations. Of course, this P.I. depends on the GERM model structure and on the WCVV approach.

In general, the theory of nonlinear systems [Stephanopoulos, 2015], postulates that, if a system presents one or more stationary states (QSS-s), after a perturbation the system will evolve toward the QSS which includes that perturbation in their 'attraction' domain, and which present a stronger stability [Maria, 2003]. The system is stable if the real part of the all eigenvalues of the system dynamic model Jacoban matrix are negative, Eq.(4,11).

Example. As an example, in (Figure 2-10, and Figure 2-17) is presented how a simple generic GERM of [G(P)1] type {with only one buffering reaction G + P ↔ GP, adjusting the G-"catalyst" activity} presents a better dynamic efficiency compared to the simplest [G(P)]0 gene-expression module, where any control on the G-activity exists. Thus, the stationary [P]s and | G |s are recovered faster after an impulse perturbation in the [P]s, (that is here, a -10% decline in [P]s at an arbitrary time t = 0) [see the papers of Maria and Scoban, 2017,2018].

The Figure 2-17 exemplifies, in a simple mode, the huge positive effect of a buffer reaction effector (G+P ⇔ GP) on the GERM dynamic efficiency. The chosen case is those of a [G(P)1] type (from Figure 2-12) [Maria, 2005, 2009]. Compared to [G(P)0], in the [G(P)1] case, the P recovering rate is much higher, while the species are better interconnected (smaller STD).

Example. Maria [2017a, 2017b, 2018] proved (Figure 2-16) how this dynamic regulatory efficiency (that is the QSS recovering time after an impulse-like perturbation) of the GERM depends not only on the number of effectors (TFs), but also on the GERM's regulatory scheme structure.

For instance, in the Figure 2-16, are compared the regulatory efficiency of four GERM types of homeostatic self-regulatory mechanism of G expression. Simulations indicate the key-species recovering trajectories and times after an "impulse"-like dynamic perturbation of -10%[P]s in the key-protein.

The considered generic structures of GERMs correspond to the followings regulatory units (see also Figure 2-12):

(1). [G(P)0] with any regulatory effector;

(2). [G(P)1] with one effector, i.e. rapid buffer reaction G + P ⇔ GP;

(3). [G(P)1; M(P)1] with two effectors, and a cascade control at two levels (G and M);

(4). [G(PP)2] with two effectors, that is the rapid buffer reactions $G + PP \Leftrightarrow GPP + PP \Leftrightarrow GPPPP$, while the TF is present in a dimeric form (PP).

See also (Figure 2-12) for the GERM-s model structures.

As a result of model-based simulations, the regulatory efficiency increases in the relative order:

$$[\mathbf{G}(\mathbf{P})]0 < [\mathbf{G}(\mathbf{P})1] < [\mathbf{G}(\mathbf{P})1;\mathbf{M}(\mathbf{P})1] < [\mathbf{G}(\mathbf{PP})2].$$

The index "s" of [P]s denotes the GERM stationary-state QSS (homeostasis). Here, the term "*effectors*" defines the number of transcription factors (TFs), that is P, PP, etc., and the number of the "catalyst" control "buffer" reactions $(G + P \leftrightarrow GP + P \leftrightarrow GPP + \ldots$; $M + P \leftrightarrow MP$, etc.). Of course, as the number of effectors is higher, as the GERM P.I.-s are better (see Figure 2-9).

The structure "4" reported the best results, due to the presence of two rapid buffering reactions, while the TF is present in a dimeric (PP) form. A good regulatory efficiency is given also by the structure " 3 ", due to its cascade control of the expression at two levels (G and M), as it is also in the reality (that is a separate control of the transcription, and translation).

B-b) Recovering rate

This P.I. is in a direct relationship to the P.I. (B-a, "recovering time"). This regulatory index is express by R_D (Table 2-2),

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

which is the rate by which the GERM key-components return to their stationary (QSS) concentration (Cs) after an "impulse"-like perturbation in one key-component. The nonlinear systems theory gives an approximate relationship to a-priori estimate this recovering rate of the GERM ODE model, that is [Zwillinger et al. 1996]:

$$R_D = \left|\mathrm{Max}\left(\mathrm{Re}\left(\lambda_i\right)\right)\right| \text{; where } \mathrm{Re}\left(\lambda_i\right) < 0 \text{, for all } \lambda_I \tag{29}$$

where: Re = real part; λ(i) = the eigenvalues of the linearized kinetic model Jacobian matrix under stationary conditions, Eq.(23), that is $\boldsymbol{J}_C = (\partial \boldsymbol{h}(\boldsymbol{C},\boldsymbol{k}) / \partial \boldsymbol{C})_s$ According to the iso-osmotic constraint Eq.(6) of the WCVV model, the cell models of the central carbon metabolism (CCM), and cell regulatory circuits (GERM, or GRC) should include the all species of the cell, taken individually, or lumped.

Example. In the (Figure 2-9) are presented the recovering rate R_D after an "impulse-like" dynamic perturbation in the expressed P protein for three GERMs structures mentioned under the figures abscissa, that is $[\mathbf{G}(\mathbf{P})0]][\mathbf{G}(\mathbf{P})1]$, and $[\mathbf{G}(\mathbf{P})1;\mathbf{M}(\mathbf{P})1]$. The perturbation corresponds to an "instant-like" decrease of the stationary key-protein [P] s with -10%. The recovering time of the stationary [P] s with a 1% tolerance is denoted by τ_{rec} [Maria, 2005, 2017a]. The top figures demonstrate how R_D depends on the GERM structure, and its number of effectors, that is here, the number of "buffer" reactions controlling the catalyst activity (that is $G + P \Leftrightarrow GP; M + P \Leftrightarrow MP$), but also on the chosen regulatory scheme. Thus, the GERM of type $[\mathbf{G}(\mathbf{P})1;\mathbf{M}(\mathbf{P})1]$ with a cascade control of the expression, done for the both transcription and translation steps, presents the best regulatory efficiency (large R_D, and small τ_{rec}). At the same time, the key-species recovering trajectories present small amplitudes, thus with a minimum disturbance of the other species homeostasis (linked directly to these species, or via the cell volume to which all species contribute, as imposed by Eq.(6).

The species ' j ' recovering times $\tau(\mathbf{j}), R_D$, and the species recover trajectories Cj(t) can easily be obtained by *simulation* with using the analysed cell system WCVV kinetic model (that is the GERM ODE models of Figure 2-12). An example of such a [G(P)1] kinetic model by using the WCVV formulation is given in section 2.3.1 (Figure 2-12, and Figure 2-24).

Species recovering trajectory and amplitude are both very important (Figure 2-9, and Figure 2-16). As proved by [Maria, 2003], GERMs display very different recovering trajectories, amplitudes, and rates according to their reaction pathway structure (Figure 2-9 and Figure 2-16). The most effective are the GERM types that ensure the faster recover of the steady-state (QSS), with a smallest amplitude of the recovering pathway, thus not disturbing the other cell metabolic processes. As underlined by [Maria, 2003], the recovering trajectories in the G/P phase plane is more '*linear*'-like for the efficient GERMs, by presenting a lower amplitude. Index " s " denotes the stationarystate (homeostasis).

B-c) Regulatory robustness

The regulatory robustness of a GERM model is defined by Maria [2005, 2017a] as being the property to realize (Min) $\left(\partial R_D / \partial \boldsymbol{k}\right)$, where R_D denotes the key-species recovering rate (see the above " Bb Recovering rate" section), while ' k ' is rate constant vector of the GERM model (depending on the micro-organism type). In fact, the cell metabolic network robustness and functionality are linked to the cell phenotype and gene regulation scheme (depending on the each individual gene expression). An example of how to choose the GERM types in the GRC model is given by Maria and Luta [2013].

B-d) Species interconnectivity

In a GERM module, of a simplified regulatory schema (see Figure 2-12), including a small number or lumped species and reactions, the interconnectivity of individual or lumped species can be viewed as a degree to which they "assist|" each other, and intrinsically "cooperate" inside the GERM system to realize optimal regulatory performances.

The cell species connections appear due to common reactions, or due to the common intermediates participating to chain (parallel - successive) reactions, or from the common cell volume to which all cell species contribute (under constant osmotic pressure, see the proofs of Maria [2017a, 2017b, 2018], and the WCVV model hypotheses of the (Table 2-1).

In this topics, Vance et al. [2002] reviewed and proposed several quick experimental/computational rules to check a reaction schema via species inter-connectivity's. By inducing experimental / *in-silico* perturbations to a (bio)chemical system, by means of concentration tracers, or by fluctuating the inputs of the system, one can quantitatively determine the perturbation propagation through the consecutive/parallel reaction pathway. Then, various computational techniques can determine the '*distance*' among observed species, and specific rules can be used to include this information in elaborating a novel reaction schema Based on a large number of *in-silico* studies, Maria [2002, 2017a, 2017b, 2018] proposed a *novel measure of species interconnectivity* related to the species recovering-times after a dynamic perturbation, that is: AVG(τ(j)) and STD(τ(j)), i.e. the average and the standard deviation of the species individual recovering times τ(j), respectively (simply denoted by AVG and STD). As AVG and STD are larger, as the cell dynamic regulatory effectiveness is lower, and the GERM / GRC model species are less interconnected, and components recover more disparately (scattered recover times), thus reducing the plausibility of this regulatory structure. The higher the number of effectors

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

and buffering reactions (see GERMs types / units in Figure 2-12), the better dynamic regulatory indices of GERM-s are. [Maria, 2003, 2005, 2006, 2007, 2009, 2010, 2014a, 2014b, 2017a, 2017b, 2018]; [Maria and Luta, 2013; Maria and Scoban, 2017, 2018]. See sections 2.2.3, 2.2.4, and Figure 2-17.

B-e) Cell sub-system QSS stability

The QSS (quasi-steady-state) stability of a cell sub-system and in particular of a GERM , refers to the system's capacity to recover a QSS after cessation of a dynamic "impulse"-like perturbation in one or more key-components.

Such a property can be highlighted by analysing the QSS of the kinetic ODE models of some GERM-s defined in a WCVV modelling framework, see [Maria, 2003, 2005, 2006, 2007, 2009, 2010, 2014a, 2014b, 2017a, 2017b, 2018], and [Maria and Luta, 2013; Maria and Scoban, 2017,2018].

The QSS stability property can be evaluated from the analysis of the eigenvalues $\lambda(i)(i$ = species index) of the linearized model Jacobian matrix Eq.(23), that is $\boldsymbol{J}_C = (\partial \boldsymbol{h}(\boldsymbol{C},\boldsymbol{k}) / \partial \boldsymbol{C})_s$, with the elements $\mathrm{J(i,k)} = \partial \mathrm{h(i)}(\boldsymbol{C},\boldsymbol{k}) / \partial \boldsymbol{C}(\boldsymbol{k})$ evaluated analytically, or numerically [Maria, 2017a, 2017b, 2018]. See also the relationships of section 2.2.3(A-iv). Details are also given by Maria [2005, 2006], and by Maria and Scoban [2017, 2018].

The QSS is asymptotically stable if the real parts of the Jacobian eigenvalues λ(i) are all negative, that is $\mathrm{Re}(\lambda(\mathrm{i})) < 0$ for all components " i " [Heinrich and Schuster, 1996]. If the system is stable then, it reaches the same QSS after cessation of a dynamic "impulse"-like perturbation in one key-component, or it reaches another QSS after cessation of a stationary, continuously, "step"-like perturbation in one key-component of the cell model. Here are to mention two important observations:

1). An essential characteristic of the WCVV models including the Pfeiffers' isotonicity constraint [hypotheses of Table 2-1, and Eq.(23-29)] is that they are always stable (intrinsic stability), because, as proved by [Maria et al., 2002; Morgan et al., 2004], always $\mathrm{Max}(\mathrm{Re}(\lambda(\mathrm{i}))) = -D$, where "D = Di" is the instant cell dilution rate in a WCVV kinetic model formulation. This intrinsic parameter "Di" is evaluated by combining Eq. (6) and Eq.(12), in the following form [Maria, 2005, 2006, 2007, 2017a]:

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

Which, by derivation and division with V is leading to:

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

$$\frac{1}{V}\frac{dN_j}{dt} = r_j, \mathrm{j} = 1,\ldots,\mathrm{ns} . \tag{30}$$

where: Cj = species "j" concentration; Nj = species "j" number of moles; V = cell (cytosol) volume; R = universal gas constant; T = absolute temperature; π = osmotic pressure; t = time; Di = instant cell dilution rate; Dm = average (apparent) cell dilution rate = Ln(2)/(cell cycle), Eq.(12).

2). By contrast to the novel WCVV modelling framework, one fundamental deficiency of the classical "default" (constant-volume like) WCCV kinetic model formulation {Eq.(1a-c) in section 2.1 } is the lack of the intrinsic stability of the cell system kinetic model, because these models do not include neither the Pfeffers' constraint Eq.(6,10-12), nor an equivalent constraint. Consequently, as proved in section 2.3.1, the GERM model formulated in a classic (wrong) WCCV framework, is not able to simulate how the system recover its homeostasis after a dynamic perturbation (as illustrated in Figure 2-17, Figure 2-10, Figure 2-16, from using a novel WCVV model). Besides, when using the classical WCCV wrong modelling framework, its predictions are distorted and inaccurate. Unfortunately, the classical WCCV continue to largely be used in the dedicated literature. As WCCV predictions are proved by Maria [2018d, 2017a, 2017b, 2018] (section 2.3.1) to be invalid and ineffective, the application of the old ("classic/default") WCCV formulation in modelling cell metabolic processes and GERMs/GRCs becomes questionable and not recommended. Such a WCVV isotonic formulation always ensures the stability of the cell system model, and the recover of its stationarity (QSS, homeostasis) after a dynamic or stationary perturbation, as discussed in sections 2.2.3, 2.2.4, and displayed in Figure 2-9, Figure 2-10, Figure 2-16, Figure 2-18, Figure 2-20, Figure 2-23-59, Figure 2-33, Figure 2-37, Figure 2-38, Figure 2-44, Figure 2-58, Figure 2-59 .

Example. The GERM system stability is ensured by the mutual regulation of P and G synthesis. Thus, in a simplified GERM model, by considering the G/P pair synthesis, with P and G as cross catalysts, each one catalysing the other synthesis. As an immediate effect of such a simplified formulation, Figure 2-10 demonstrates the positive effect of such a mutual catalysis on the system stability *vs.* dynamic perturbations.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

2.2.4. Some rules to link GERMs when deriving GRCs models

When modelling a complex GRC consisting of a chain of GERMs (Figure 1-40-left), they are two problem to be considered: I) what effective and suitable GERMs have to be chosen from the GERM's library (Figure 2-12) to match the individual expression characteristics, and ii) what rules to be applied when linking GERM-s to adequately reproduce the holistic regulatory properties of the GRC in the context of the cell balanced growth. In this section, some of such linking rules proposed and tested by Maria [2017a, 2017b, 2018] are briefly reviewed. Such rules are *in-silico* (math model-based) derived by using a large number of simulations. When experimental data were available, the confirmation of the model adequacy was emphasized each time.

When linking GERM-s to construct a complex GRC reproducing a certain function of the cell, there are two contrary goals: (a) on one hand is the use of simple GERM structures to reduce the model identification computational / experimental effort; (b) on the other hand, it is important to use simple, but effective and flexible linking rules of GERM-s able to reproduce individual enzyme-synthesis, but also the individual and holistic regulatory properties (performance indices P.I.-S of Table 2-2) of the GRC. Some math modelling tools in this respect are given by Maria [2017a] (section Systems Biology - Math Modelling Tools). This section reviews, with examples from Maria [2017a, 2018], some adjustable rules to be applied for linking GERM-s to obtain the GRC-s of desirable regulatory properties.

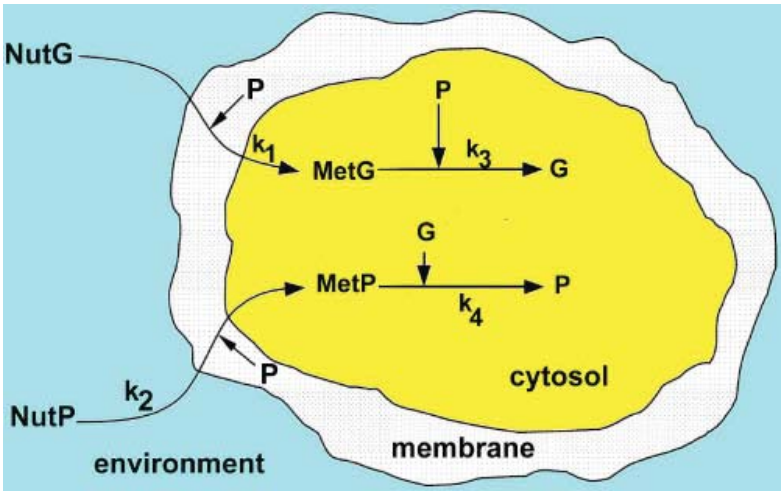


Figure 2-26: The reaction scheme of the simplest regulatory module GERM of the generic gene G expression, denoted by [G(P)0] , used to exemplify the synthesis of a generic P protein in the *E. coli* cell, proposed by Maria [2003,2005,2017a]. Adapted rom Maria [2005].

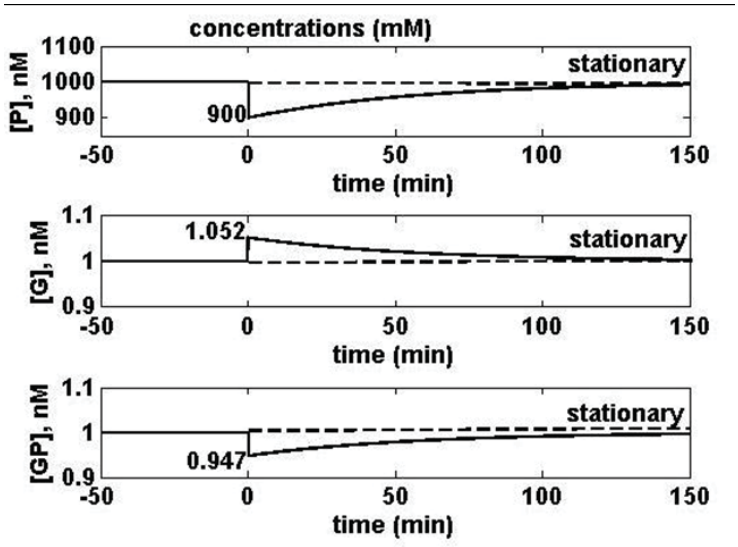


Figure 2-27: Exemplification of the self- and mutual- G/P pair catalysis and regulation after an impulse perturbation in the [P]s = 1000 nM leading to a 10% decline of the steady-state at an arbitrary time t=0 for a generic GERM of [G(P)1] type (Fig.2-24). Simulated results have been generated by using the cell nominal stationary conditions of Table 2-5, but with adopting [G]s =[GP]s = 1 nM.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

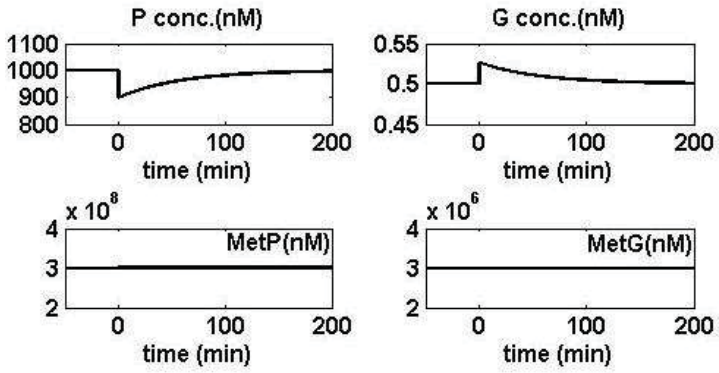


Figure 2-28: Exemplification of the self- and mutual G/P pair catalysis after an impulse perturbation in the [P]s = 1000 nM leading to a 10% decline of the steadystate at an arbitrary time t=0 for a generic GERM from E. coli simulated with a [G(PP)1] model under the WCVV approach. Simulated results have been generated by using the cell nominal stationary conditions of Table 2-5, with adopting [G]s = 0.5 nM.

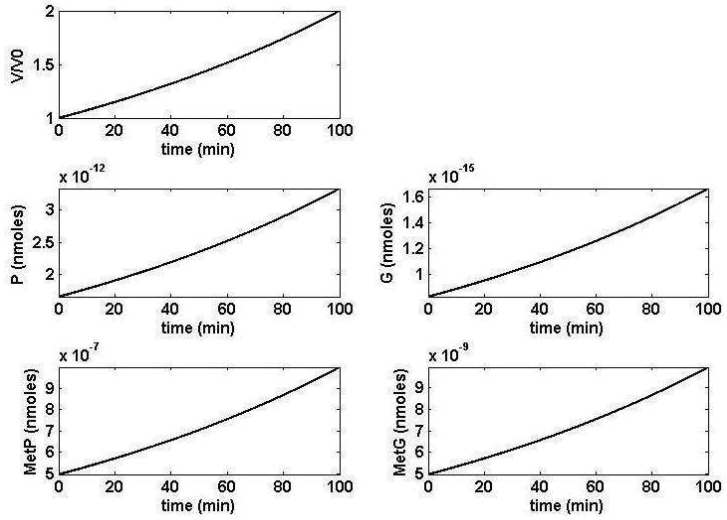


Figure 2-29: Dynamics of the cell volume and of species copynumbers Nj during the cell cycle predicted by the WCVV kinetic model for the G/P expression using a regulatory module of [G(P)1] type (Fig. 2-24). The E. coli cell species homeostatic concentrations are those of Table 2-5.

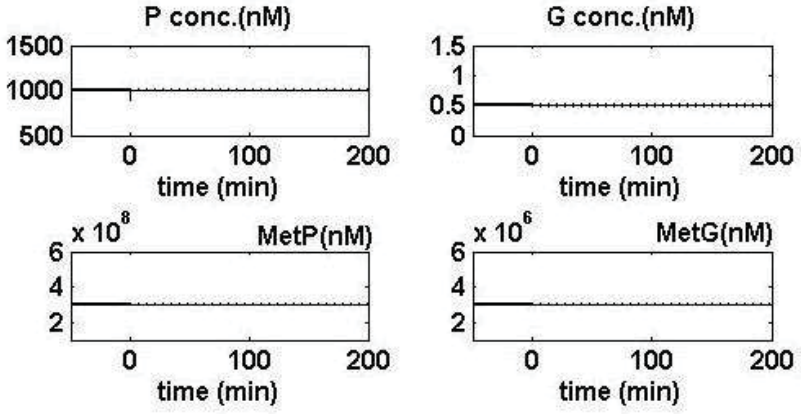


Figure 2-30: Simulation of the G/P expression under stationary (un-perturbed) conditions by using a regulatory module of [G(P)1] type (Fig. 2-24) from the *E. coli* cell in terms of species concentrations The species (individual or lumped) homeostatic concentrations are those of Table 2-5. The dynamics of the key species concentrations (in nM) during the cell cycle is predicted by using two different approaches: i) the classic ("default") constant volume WCCV model (.....), compared to ii) the variable volume WCVV novel modelling framework [----------, with instant dilution Di term and isotonicity constraint included in the model, Eq.(34- 37)]. 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 rate.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

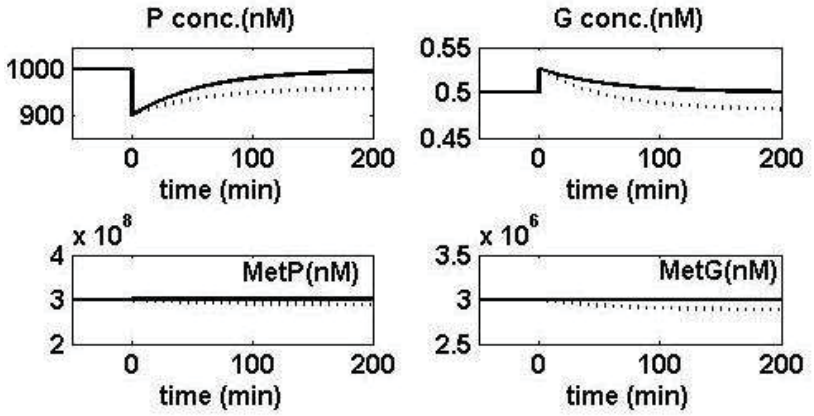


Figure 2-31: Simulation of the G/P expression after a dynamic ("impulse-like") perturbation in the [P]s = 1000 nM leading to a 10% decline of the steady-state at anarbitrary time t=0. The E. coli cell species homeostatic concentrations are those of Table 2-5. Simulations, in terms of species concentrations (nM), have been done by using a regulatory module of [G(P)1] type (Fig. 2-24) from the E. coli cell. The species (individual or lumped) homeostatic concentrations are those of Table 2-5. The dynamics of the key species concentrations during the cell cycle was predicted by using two different approaches: i) the classic ("default") constant volume WCCV model (.....), compared to ii) the variable volume WCVV novel modelling framework [_____ , with instant dilution Di term and isotonicity constraint included in the model, Eq.(34-37)]. Model predictions of the two models are very different, due to the use of the "instant" Di cell dilution rate, and the isotonicity constraint by the WCVV novel modelling framework.

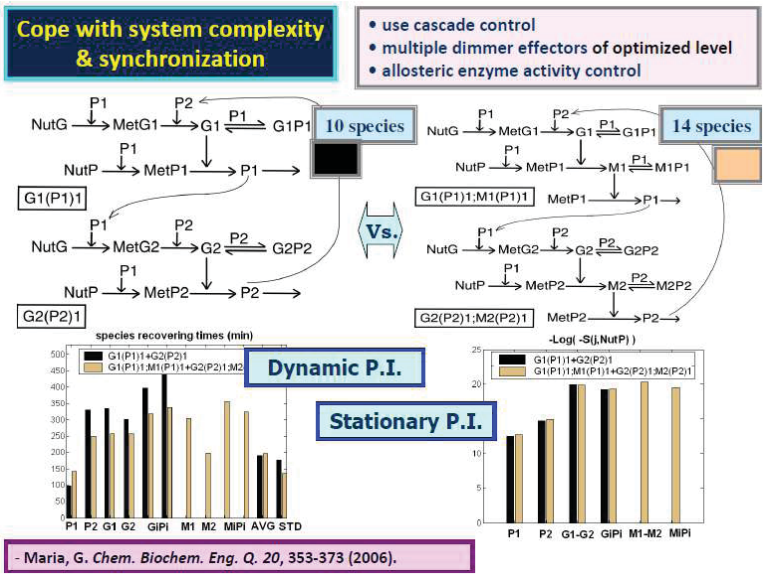


Figure 2-32: Using two GERM-s modules to build-up a GRC. The effect of the GERM model complexity on the GRC stationary regulatory performances. Two linking alternatives are considered: Alternative [A] : [G1(P1)1] + [G2(P2)1] (10 individual and lumped components). Alternative [B] : [G1(P1)1;M1(P1)1] + [G2(P2)1;M2(P2)1] (14 individual and lumped components). Despite its higher complexity (ns = 14 species in alternative [B] vs. ns = 10 species in alternative [A]), the alternative [B] presents better regulatory performances, due to more effective constituent GERM-s. Regulatory performance refer to: i) the GRC response, that is species recovering times τ (rec, j) after a dynamic perturbation impulselike of -10%[P1]s (denoted as "Dynamic P.I."), and ii) the GRC response to a stationary perturbation in the environmental [NutP]s (denoted as "Stationary P.I."), that is the GRC species relative sensitivities vs. a stationary external perturbation [NutP]s, or in math form S(Cj ; τ (rec, j) [NutP]) = ln(Cj)/ ln([NutP]s) (see P.I.-s definitions in Table 2-2) [Maria, 2006,2017A].

2.2.4-I) The effect of the no. of regulatory effectors (n)

By definition, GERM models include an adjustable number of "regulatory effectors", that is: " n " in the $[G(P)n]$, or $[G(PP)n]$ series; " n " and "n1" in the $[G(P)n;M(P)n1]$ series (Figure 2-12). As proved by Maria [2003,2005,2006], and by Yang et al.[2003], a quasi-linear relationship of P.I. function of no. of regulatory effectors (n) can be derived for every GERM type, of the approximate form: $P.I. = a0 + \sum_i a_i \times n_i$. Here, P.I. denotes the regulatory performance index, such as R_D, $AVG(\tau(j))$, $STD(\tau(j))$, stability strength, etc. Also, n(i) number of effectors (P, PP, O) acting in the "i-th" allosteric regulatory unit of type $Li(Oi)n(i)$ (Figure 2-12). Notations a0, a(i) denotes the correlation constants depending on the P.I. and GERM model type. Such a linear dependence can be observed in Figure 2-9, and Figure 2-13. In short, Maria [2005, 2006, 2007, 2014b) proved that (Figure 2-13, Figure 2-9):

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

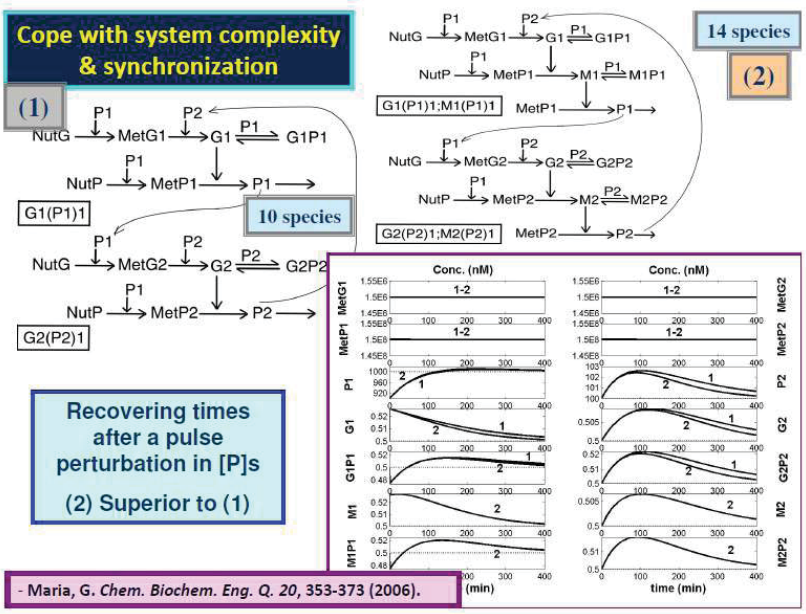


Figure 2-33: Effect of the GERM model complexity on the GRC dynamic performances (that is the key-species recovering times and trajectories after a -10% [P1]s perturbation) [Maria, 2006]. Two linking alternatives are considered: Alternative [A] : [G1(P1)1] + [G2(P2)1] (10 individual and lumped components).Alternative [B] : [G1(P1)1;M1(P1)1] + [G2(P2)1;M2(P2)1] (14 individual and lumped components).

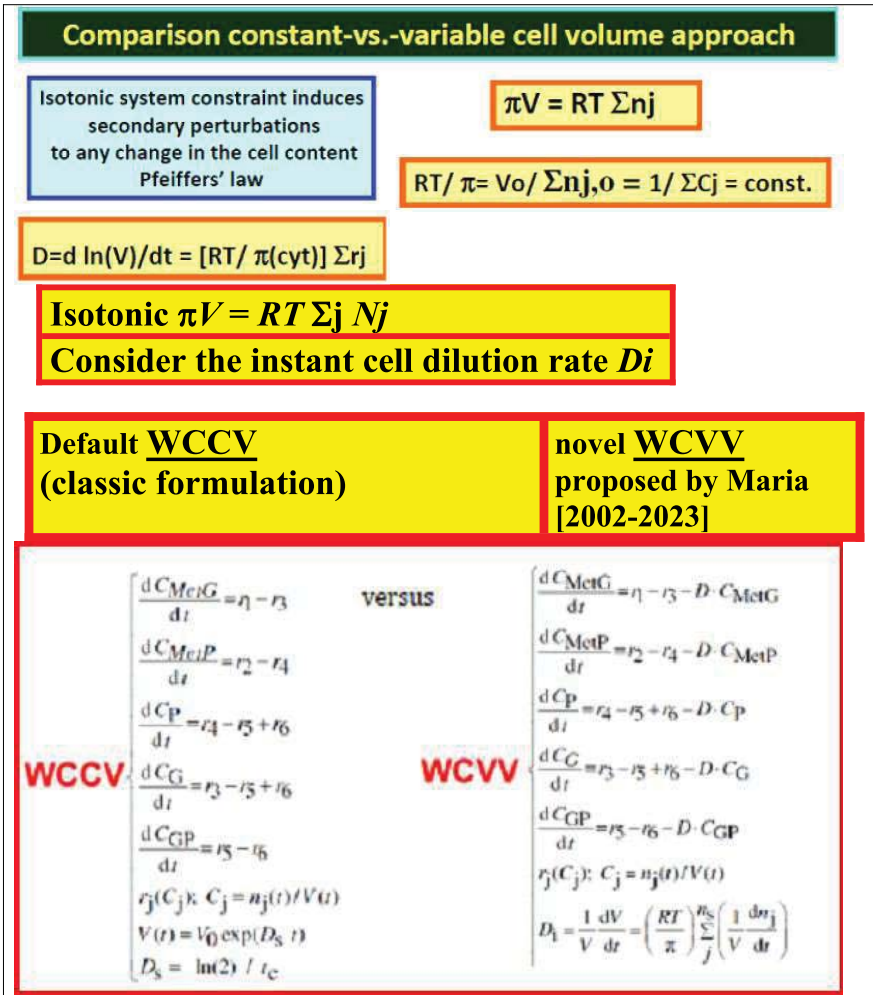


Figure 2-34: Comparison of the WCVV and WCCV modelling approach in the case of a simple [G(P)1] kinetic model to simulate an individual gene expression regulatory module (GERM) [Maria et al.,2018d].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

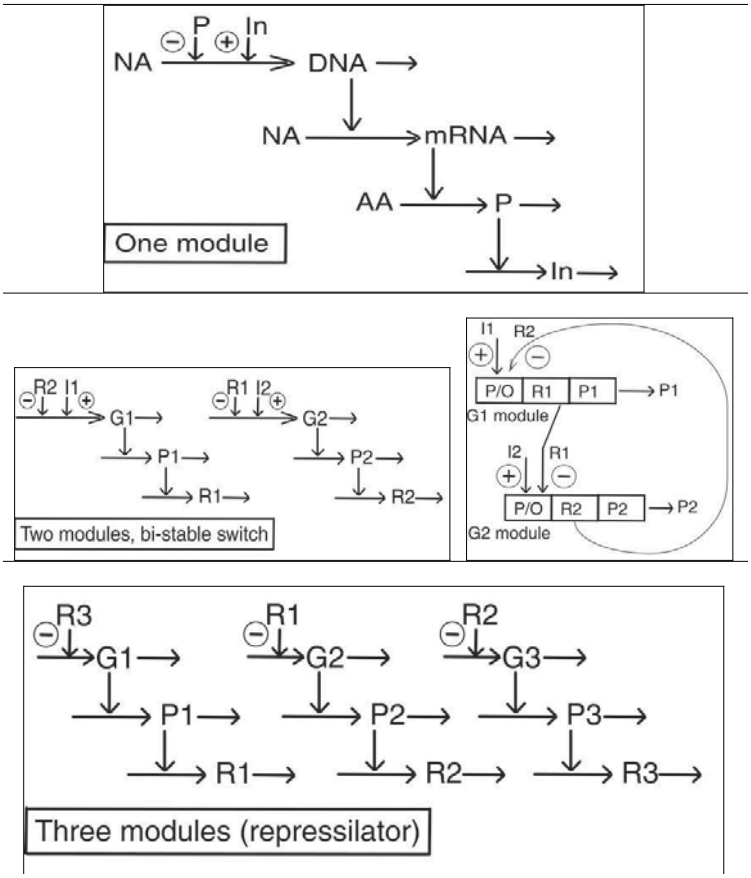


Figure 2-35: Example of lumped representations of simple GRC-s (1-3 regulatory modules) [Maria,2008].
[Top]. One gene expression G/P regulatory module with induction (In) and selfrepression (P) loops [Schwacke and Voit, 2004].
[Middle]. The bistable switch circuit of two gene expression regulatory modules, of type (Fig. 2-12) [G1(R2R2)n(R1)m] + [G2(R1R1)n(R2)m], [Maria, 2006,2007, 2017a,2018].
[Down]. A GRC circular circuit involving three self-/cross-repressed GERM-s modules, i.e. the repressilator from E. coli [Elowitz and Leibler, 2000]. The horizontal arrows indicate reactions. The vertical arrows indicate catalytic actions. The absence of a substrate or product indicates an assumed concentration invariance of these species. Notations: I, In= inducer; AA= aminoacids; P= promoter or the expressed protein, G= gene; O= operators; R= repressor.

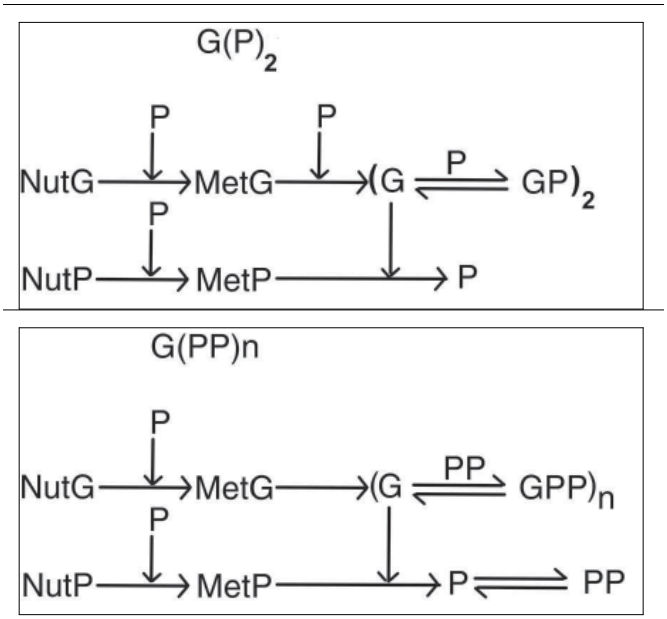


Figure 2-36: Example of two GERM modules [Maria,2008], of high effectiveness (see also the GERM library of Fig. 2-12).
[Top]. A module of [G(P)2] type with two buffering rapid reactions quickly adjusting the 'catalyst' G activity.
[Down]. A module of [G(PP)n] type with "n" buffering rapid reactions quickly adjusting the 'catalyst' G activity. Besides, multiple dimmer effectors (of optimized level) increase the system flexibility.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

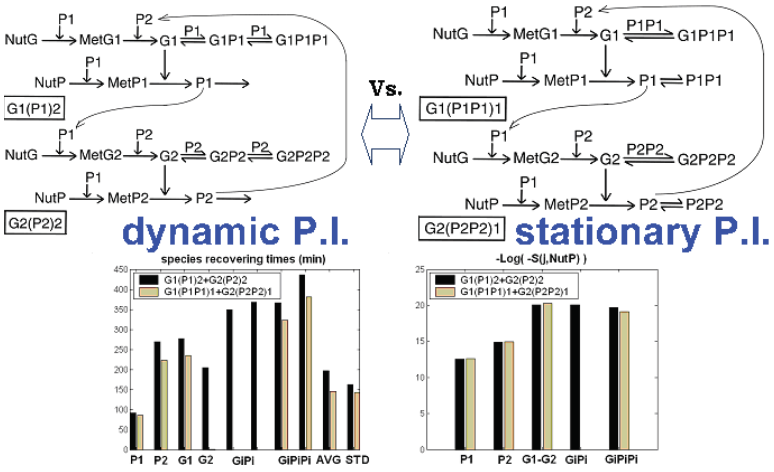


Figure 2-37: Using two GERM-s modules to build-up a GRC. The effect of the GERM model complexity on the GRC stationary regulatory performances. Two linking alternatives are considered: Alternative [A] : [G1(P1)2] + [G2(P2)2] (12 individual and lumped components). Alternative [B] : [G1(P1P1)1] + [G2(P2P2)1] (12 individual and lumped components). Despite the same complexity (ns = 12 species), alternative [B], presents better regulatory performances, due to more effective constituent GERM-s. Regulatory performance refer to: i) the GRC response, that is species recovering times after a dynamic perturbation impulse-like of -10%[P1]s (denoted as "Dynamic P.I."), and ii) the GRC response to a stationary perturbation in the environmental [NutP]s (denoted as "Stationary P.I."), that is the GRC species relative sensitivities vs. a stationary external perturbation [NutP]s, or in math form S(Cj ; [NutP]) = S(Cj ; [NutP]) = $\partial \ln(Cj)/\partial \ln([NutP]s)$ (see P.I.-s definitions in Table 2-2) [Maria, 2006,2017A].

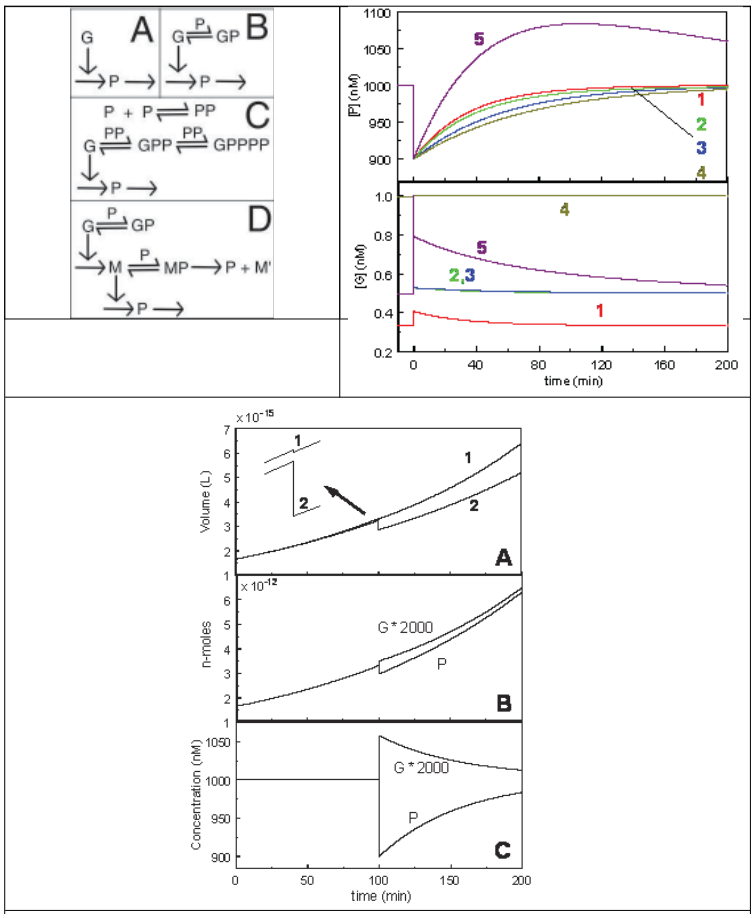


Figure 2-38: Proof the cell "ballast" effect on the species recovery rates after a dynamic "impulse-like " -10%[P]s perturbation. The example of section 2.2.4-VI. ***[TOP-left].*** The compared GERM models chosen in the example of section 2.2.4 , paragraph (VI), that is (see Fig. 2-12): (A) = [G(P)0] ; (B) = [G(P)1] ; (C) = [G(PP)2] ; (D) = [G(P)1 ; M(P)1] . ***[TOP-right].*** Dynamic recovering paths for P and G species after an "impulse-like" -10%[P]s perturbation, predicted by the following GERM models (see Fig. 2-12): "1"= (C) = [G(PP)2] ; "2"= (D) = [G(P)1 ; M(P)1] ; "3"= (B) = [G(P)1] ; "4"= (A) = [G(P)0] ; "5"= [G(P)1] with an instant step increase of [G]s from 0.5 nM to 1 nM at t= 0+. ***[Down].*** Model [G(P)1]: (A) The cell-volume growth for ("1") large ballast into the cell, the case of [MetG]s = [MetP]s = 10,000 nM, and for ("2") low ballast into the cell, the case of [MetG]s = [MetP]s = 250nM. The P and G evolution before and after a 10% perturbation of [P]s in terms of copynumbers (B) and concentrations (C) (large ballast case). The initial values are Vcyt,o = 1.660434503-10^{-15} L [Bar-Joseph et al., 2012]; [G]s = 0.5 nM; [P]s = 1000 nM. At t = 0+ one applies an instant "impulselike" dynamic perturbation of [P]s from 1000 nM to 900 nM.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

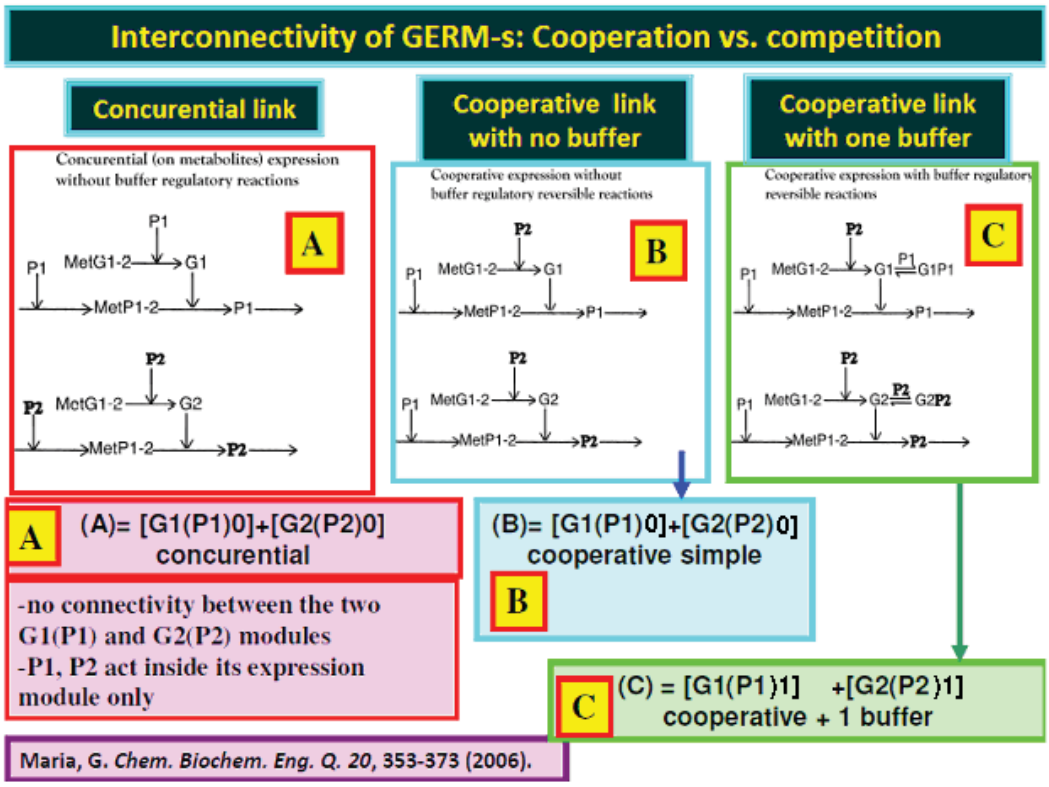


Figure 2-39: Effect of the GERM–s inter-connectivity and of individual functions of cell components into the cell on the GRC efficiency when linking two GERM-s. [Maria, 2005,2006].

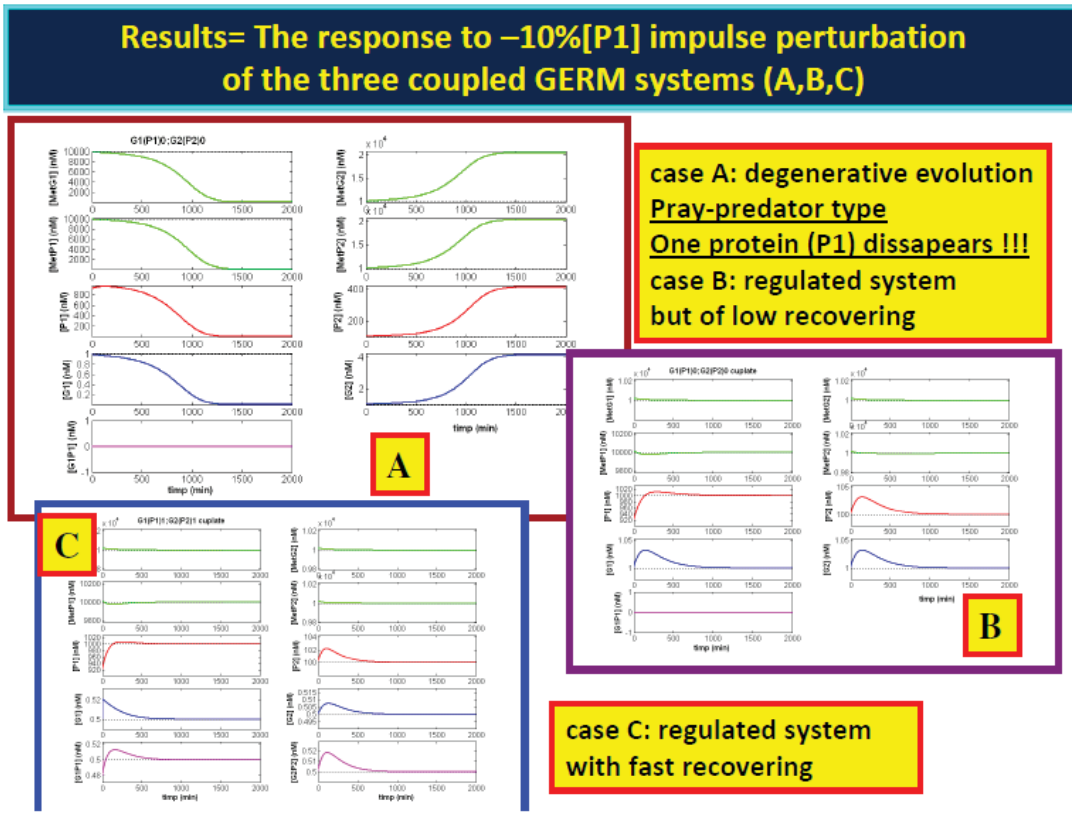


Figure 2-40: The effect of the GERM inter-connectivity on the resulted GRC viability, when linking two GERM-s regulatory units in three alternatives [A-B-C] (see Fig. 2-39) [Maria, 2006]. The plots display the key species recovering trajectories after a dynamic "impulse like" perturbation of -10%[P1]s at an arbitrary time t=0, for the all three inking alternatives of GERM-s.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

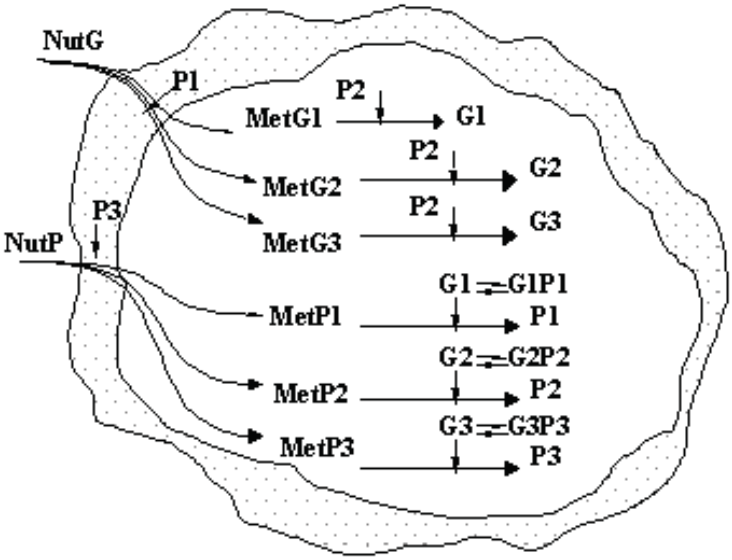


Figure 2-41: Cooperative coupling of three GERM-s , gene expression regulatory modules of types [G1(P1)1] + [G2(P2)1] +[G3(P3)1] for P1, P2, P3 protein synthesis, by giving separate but cooperative functions of proteins into the cell (alternative C of the paragraph) (Cooperative vs. concurrent linking of GERMs in GRC-s and species interconnectivity" of section 2.2.4 VIII).

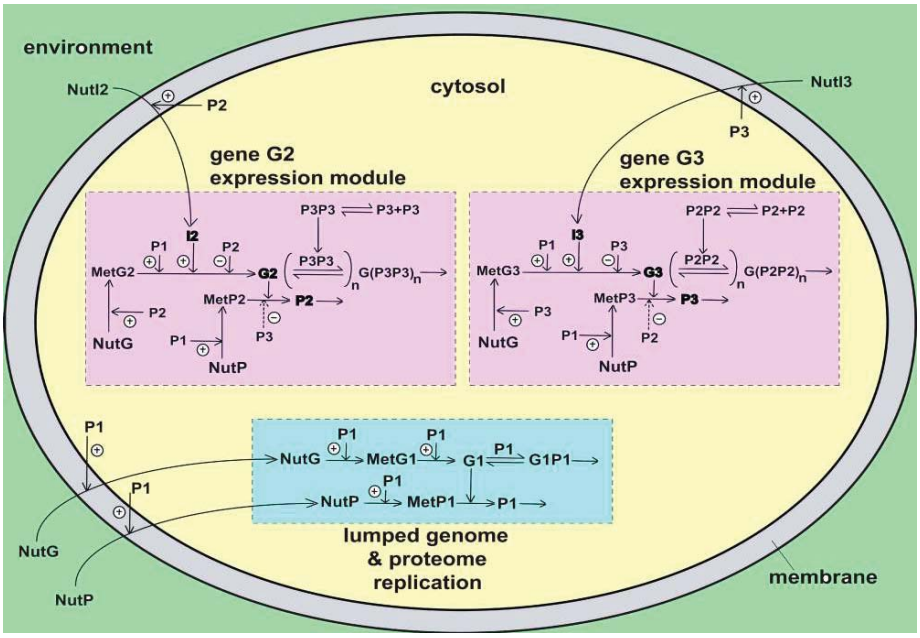


Figure 2-42: The lumped reaction pathway of a lumped genetic switch (see its principle in Fig. 2-43). The reduced modular WCVV model of Maria [2014b] considers twoself- and cross-repressed gene expression modules GERM ensuring replication of the pairs P2/G2 and P3/G3 respectively. Proteins P2/P3 expression is activated by the specific inducers I2 and I3 respectively, and repressed by the R2 and R3 proteins produced by each GERM module. The two GERM modules are of the type: [G(P2P2)1(P3P3)1] + [G3(P3P3)1G3(P2P2)1] . The simple module [G(P)1] simulate the lumped genome (G1) and proteome (P1) regulated replication. It was added to fulfil the cell isotonicity constraints Eq.(10-13). Besides, the P1/G1 high concentrations ensures the "cell ballast" positive effect (section 2.2.4-VI).

- P.I.-s improves ca. 1.3x – 2x (or even more) for every added regulatory unit to the GERM module. Multiple regulatory units lead to an average recovering time AVG($\tau(j)$) of the all GERM-s species much lower than the cell cycle duration *tc*, under an average logarithmic volume growing rate, $Dm = \ln(2) / tc$.
- Combinations of regulatory schemes and units (with different effectors) can improve the regulatory P.I.-s.
- Certain regulatory modules reported an increased flexibility, due to 'adjustable' intermediate species levels. This is the case, for instance, of adjusting [M]s in modules of type [G(P)n;M(P)n1], and of [PP]s, in modules of type

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

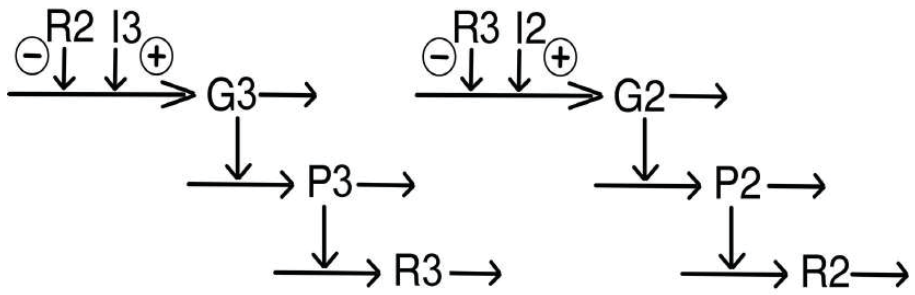


Figure 2-43: The principles of a genetic switch (GS). Two self- and cross-repressed gene expression coupled modules GERM ensure replication of the pairs P2/G2 and P3/G3 respectively. Proteins P2/P3 expression is activated by the specific inducers I2 and I3 respectively, and repressed by the R2 and R3 proteins produced by each GERM module.

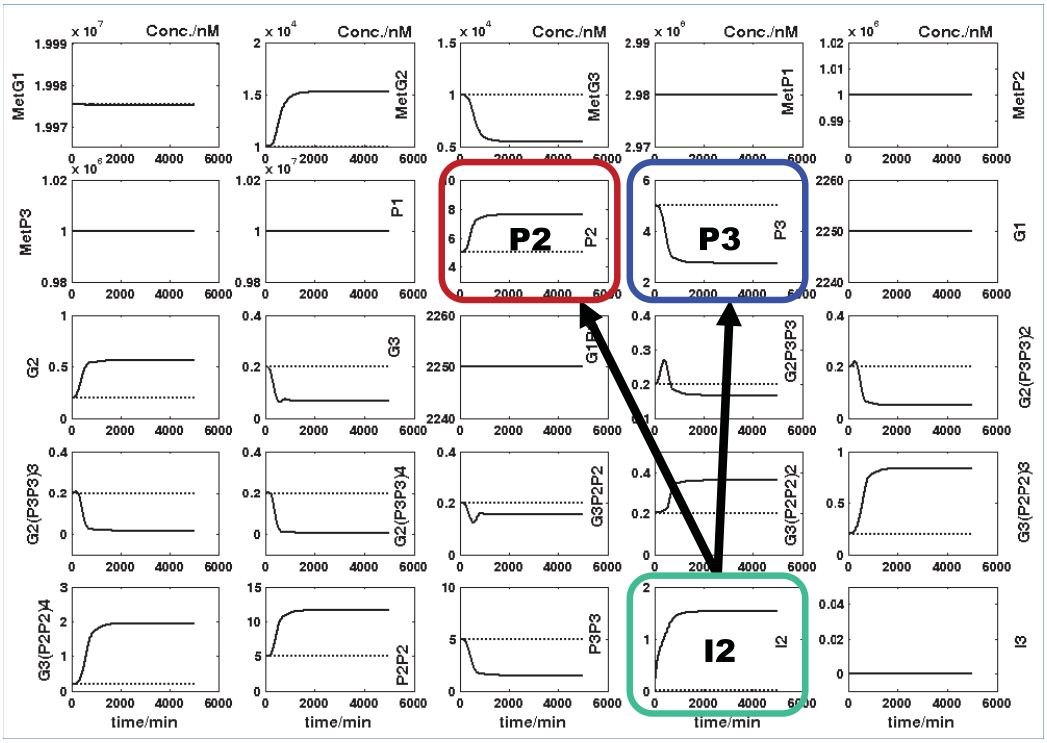


Figure 2-44: The response of the genetic switch of Fig. 2-42, and Fig. 2-43 to a stationary perturbation ("step-like") of the external inducer I2. The result is an overexpression of the P2 protein, leading to a quick and strong under-expression of the protein P3 [Maria, 2007, 2009, 2018].

[G(PP)n] (Figure 2-12). Optimal levels of these intermediate species can be set accordingly to various optimization criteria, rendering complex regulatory modules to be more flexible in reproducing certain desired cell-synthesis regulatory properties.

2.2.4.-II) Adjusting the number and the level of transcription factors TF and buffering reactions

To select the suitable GERM structure that fits the available experimental (kinetic) data, the first problem to be solved is related to the number of buffering reactions of type $G + P <===> GP$, or $M + P <==> MP$, etc. (Figure 2-12, Figure 2-16, Figure 2-20), and also to the optimal stationary concentration of the TF (see Figure 2-21) necessary to be included in the model.

As shown in Figure 2-21, the optimal concentration of [TF]s is given by the optimal value of an P.I. For instance, here corresponds to the minimum of the key-species stationary sensitivities for various [PiPi] (I = 1,2) effector levels predicted by the check GRC model. The chosen optimal TF (that is [P2P2]s, and [P3P3]s here) corresponds to the minimum of the P2,P3, and P2P2, P3P3 sensitivity curves *vs.* the [TF]s. In the example of [Maria, 2009] this optimal value turned out to be [P2P2]s = [P3P3]s = 4-10Nm.. This small TF value satisfies the principle that a cell with an optimally regulated metabolism present intermediates at low levels, and with a short existence (section 1.3.4).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Parameter	Classical fermentation	Enzymes	Chemo catalysts
Catalyst	living cells	Enzymes	Metals, acids, …
Catalyst concentration (kg/m^3)	10 – 200	50 – 500	50 – 1000
Specific reactions	sometimes	often	often
Reaction conditions	moderate	moderate	moderate – extreme
Sterility	yes	yes	no
Yield (%)	10 – 95	70 – 99	70 – 99
Important cost item	cooling water	enzyme	varies
Problems	regulation of microorganisms	stability	selectivity, stability

Enzymatic reactions and bioprocesses can replace difficult chemical syntheses, leading to:
- high selectivity and yields
- less by-products
- less energy
- less environmental pollution
- small biocatalyst concentration
- moderate reaction conditions

Protein engineering / nano-structures / GMO lead to new biocatalytic systems

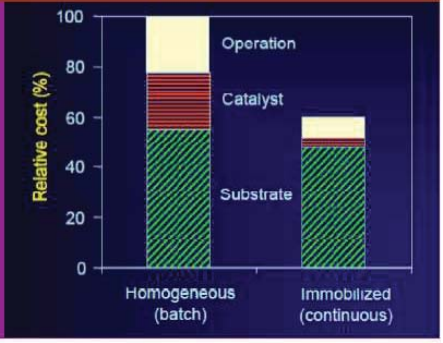


Figure 2-45: Advantages of biosynthesis processes (fermentations in bioreactors, using cell cultures) and enzymatic syntheses, compared to classical chemical catalytic processes. [Bottom-right] Production cost structure in the case of biosyntheses with free enzyme (or biomass) compared to those with immobilized enzyme (or biomass). After [Moulijn, 2001].

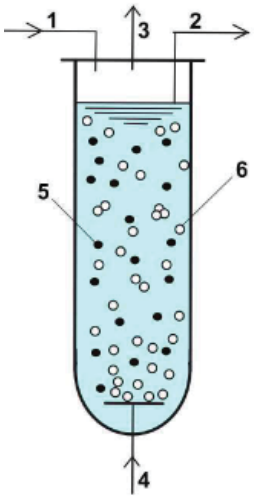


Figure 2-46: Simplified scheme of the three-phase fluidized bed reactor (TPFB) of section 3.2, used by Deckwer et al. [2004] to perform dynamic experiments. Notations : 1 mercury ion fed solution including nutrients (C/saccharides, N/ureea, P/ phosphate salts, mineral sources, pH-buffer additives, anti-bodies, etc.); 2 - liquid outlet; 3 - Hg vapour in air flow; 4 - sterile air input; 5 - immobilized bacteria; 6 - air bubbles. The bioreactor performances depends on the following parameters: i) suspended biomass concentration and efficiency; ii) biomass porous support size; iii) feed flow-rate, inlet [Hg2+]; iv) nutrients, additives, aeration rate, pH, etc. The cell metabolic process efficiency (mercury reduction rate) depends on the PT, PA enzymes concentration (dependent on the cell resources and environmental / cell [Hg^{2+}]).

Evaluation of P.I.-s for a large number of GERM structures (Figure 2-12) as proposed by Maria [2002, 2003, 2005, 2006, 2007, 2008, 2009, 2010, 2014b, 2017a, 2018] indicated that the dynamic regulatory efficiency of the $[\mathbf{G(P)n}]$ module type is nearly linearly increasing with the number (n) of buffering reactions (with a correlated P.I. following the approximate relationship: $P.I. = a0 + \sum_i a_i \times n_i$; see Figure 2-13, Figure 2-9). Moreover, the plots of Figure 2-9 reveal that this increase is more pronounced in the case of $[\mathbf{G(PP)n}]$ model structures that use dimeric TF-s (that is PP instead of simple P), and also for $G(P)n;M(P)nl]$ modular units that use a "cascade" control scheme of the gene expression.

Such a module efficiency ranking concerns not only the dynamic efficiency, but also most of P.I.-s, as discussed in the previous paragraphs, such as the stationary regulatory effectiveness; low sensitivity *vs.* stationary perturbations; stability strength of the

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

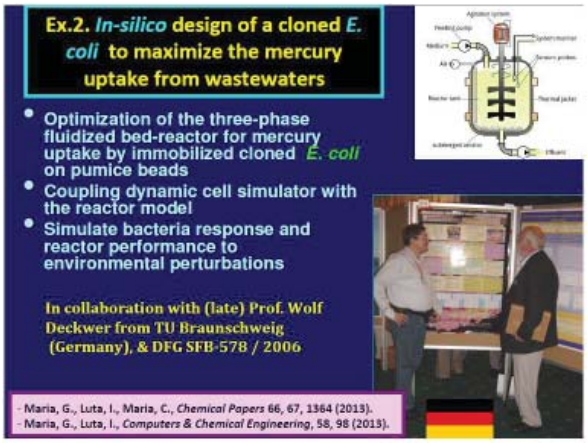


Figure 2-47: Case study of section 3.2 - In-silico design of a cloned *E. coli* with mer-plasmids to maximize its capacity of mercury uptake from wastewaters [Maria, 2009b, 2010,]; [Maria and Luta, 2013; Maria et al., 2013; Scoban and Maria, 2016; Maria and Scoban, 2017b]. [Down-right] Prof. G. Maria and late Prof. W. Deckwer at the 2nd Croatian-German Conference on Enzyme Reaction Engineering, 21-24 Sept. 2005, Dubrovnik (Croatia), sharing opinions about the bioprocess of mercury removal from wastewaters by using cultures of cloned E. coli cells.

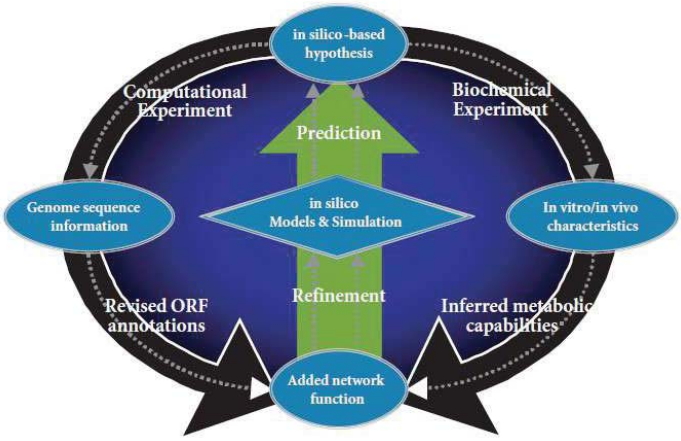


Figure 2-48: Iterative *in silico* model building in biology involves the formulation of experimentally testable hypotheses based on the in silico analysis, collection of experimental data, and subsequent refinement of the models based on these data. After Palsson [2000].

homeostatic QSS. The GERM regulatory units is more effective when the species recovering trajectories after a dynamic small perturbation, display a more linear shape in the G/P phase plane, and of a lower amplitude of the recovery curves (Figure 2-9).

To summarize, when selecting a suitable GERM to be included in a GRC to be built-up, the following issues are to be considered following the experience of Maria [2003,2005,2006,2007,2008,2009,2014b,2017a,2018]:

- The GERM modules reporting stationary-regulation high P.I.-s, also report high dynamic-regulation P.I.-s.
- The catalyst activity control at a single enzyme level (i.e. [G(P)0], [G(PP)0],[G(P)n;M(P)0] structures (Figure 2-9, Figure 2-16), that is lacking of buffering reactions, is not able to modulate the gene G and M catalytic activity, leading to the lowest regulatory efficiency.
- Multiple copies of effector molecules (i.e. O, R, P in Figure 2-11, Figure 212, Figure 2-8, Figure 2-21), which reversibly and sequentially (allosterically) bind the catalyst (G,M) in negative feedbacks, will improve the GERM regulation effectiveness.
- A structured cascade control of the "catalyst" activity, with negative feedback loops at each level as in the [**G(P)n;M(P)n**l] model series (Figure 2-12), improves regulation and amplifies the effect of a change in a stimuli (inducer). The rate of the ultimate reaction is amplified, depending on the number of cascade levels and catalysis rates. As an example in Figure 2-11, Figure 2-12, Figure 2-8, Figure 2-21, by placing regulatory elements (R,P,PP,...) at the level of mRNA (i.e. species denoted by M in the GERM-s of Figure 2-12), and at the level of DNA (i.e. species G in the GERM-s of Figure 2-12), and of multiple buffering reactions in the [**G(P)n;M(P)n**l] models is highly effective.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

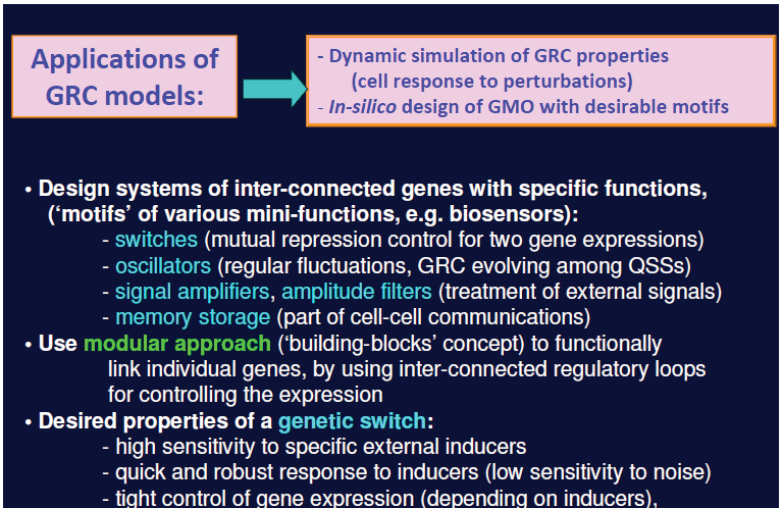


Figure 2-49: Some applications of GRC models, according to Maria [2005, 2017a, 2023].

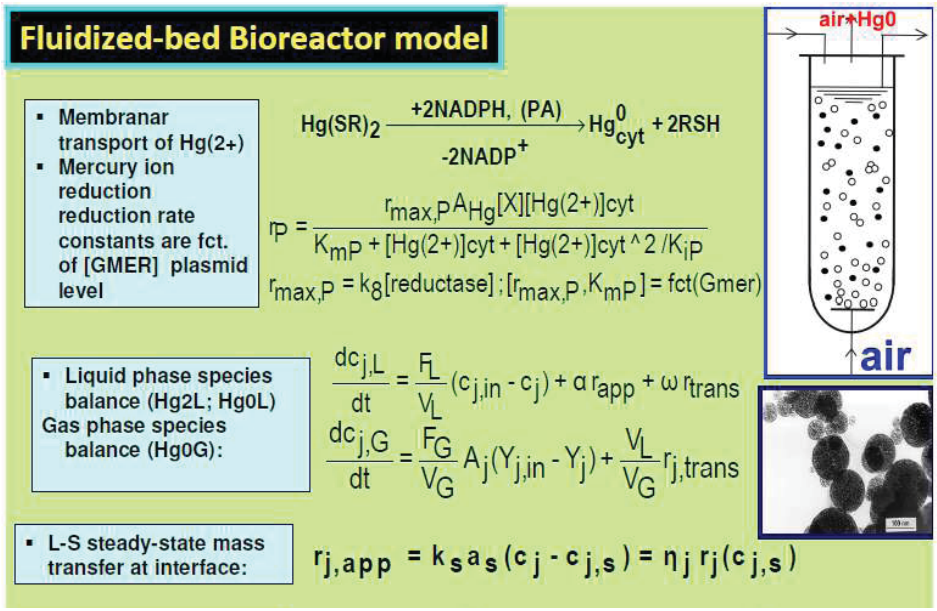


Figure 2-50: The macroscopic model of the three-phase fluidized-bed bioreactor (TPFB) with suspended immobilized E. coli on pumice beads. The Michaelis- Menten rate constants depend on the mer-plasmid level [Maria, 2009b, 2010; Maria and Luta, 2013].

- The nearly linear increase of GERM P.I.-s with the number "n(i)" effectors (P, PP, O, R, etc.) acting in the "i"-th allosteric unit $[\mathbf{L(i)(O(i))n(i)}]$ of buffering reactions applied at various level of control of the gene expression, is valid for both dynamic and stationary P.I.-s of the Table 2-2.

- Roughly, the P.I.-s improve ca. 1.3x-2x (or even more), for every added regulatory unit to the same GERM. Multiple regulatory units lead to much lower average species recovering times AVG(τ(j) than the cell cycle period *tc*, under an average volume growing logarithmic rate of $Dm = \ln(2)/t$.

- Combinations of regulatory schemes and units (with different effectors) improve the regulatory P.I.-s, according to Maria [2005,2006,2007,2017a,2018].

- Certain regulatory GERM modules reported an increased flexibility, due to 'adjustable' intermediate transcription factors TF species levels. This is the case, for instance, of adjusting [M]s in module $[\mathbf{G(P)n;M(P)n1}]$ and of [PP]s in the modules [G(PP) n]. Optimal levels of these species can be set accordingly to various optimization criteria, rendering complex regulatory modules to be more flexible in reproducing certain cell-synthesis regulatory properties. Thus, Maria [2009] proved existence of an optimal [TF] concentration leading to optimal P.I. of a GERM (Figure 2-21).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

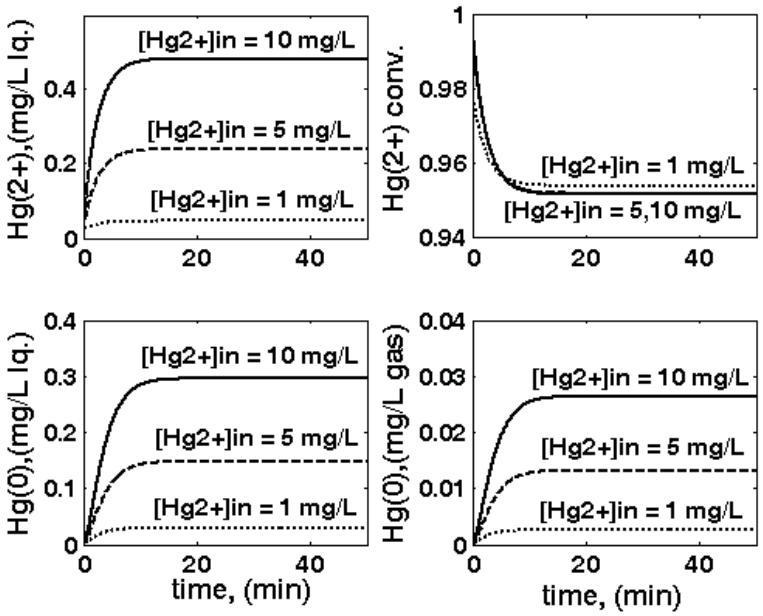


Figure 2-51: Three-phase fluidized bed bioreactor sensitivity related to variations in the inlet $[Hg_1^{2+}]$ in, leading to variations of the following operating variables: outlet $[Hg_1^0]$ (up-left); outlet $[Hg_L^0]$ (down-left); $[Hg^0]$ conversion (up-right); and outlet $[Hg_G^0]$(downright) in the output gas. Nominal conditions: [Gmer] = 3 nM; F_L = 0.02 L/min ; biomass c_X = 1 g/L ; particle size d_p = 1 mm. Notation „in" refers to the bioreactor inlet. All notations are given in the section 3.2.1. Adapted from [Maria and Luta, 2013].

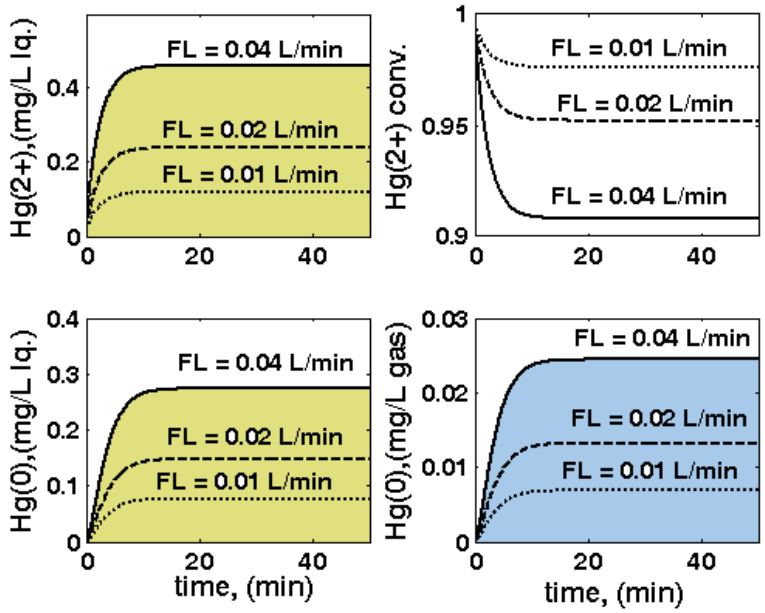


Figure 2-52: Three-phase fluidized bed reactor sensitivity related to inlet liquid flow rate F_L variations, leading to variations of the following operating variables: outlet $[Hg_1^{2+}]$ (up-left); outlet $[Hg_L^0]$(down-left); Hg_1^{2+} conversion (up-right); outlet $[Hg_G^0]$ in the output gas (down-right). Nominal conditions of Fig. 2-51.

2.2.4-III) The effect of the mutual *G/P* synthesis catalysis

One essential aspect of the $[\mathbf{G(P)n}],[\mathbf{G(PP)n}]$, and [G(P)n;M(P)n1] most common GERM kinetic models is the mutual catalysis of *G* and its encoding protein *P* synthesis. If one adds the WCVV modelling constraints Eqn.(11,12,13,14), and the requirement of all times obtaining a maximum dynamic responsiveness and efficiency by keeping in a quasi-equilibrium equal levels for the active and inactive forms of the catalyst $[G]s = [G(P)]s = [G(PP)]s = \ldots = [G(P)n]s$, as discussed by Maria [2003,2005,2006,2007,2008,2009,2014b, 2017a]. This direct and indirect link of G and P syntheses (Figure 2-7, Figure 2-11, Figure 2-21) ensures a quick recovering of both stationary [G]s and [P]s after any small internal/external perturbation.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

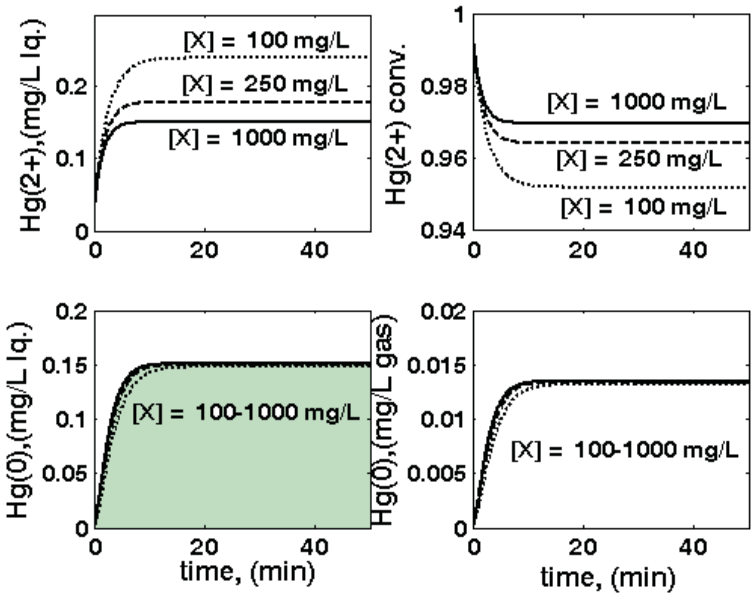


Figure 2-53: Three-phase fluidized bed reactor sensitivity related to biomass concentration c_x variations, leading to variations of the following operating variables: outlet [Hg_1^{2+}](up-left); outlet [Hg_L^0](down-left); Hg_1^{2+} conversion (up-right); outlet [Hg_G^0] in the output gas (down-right). Nominal conditions of Fig. 2-51.

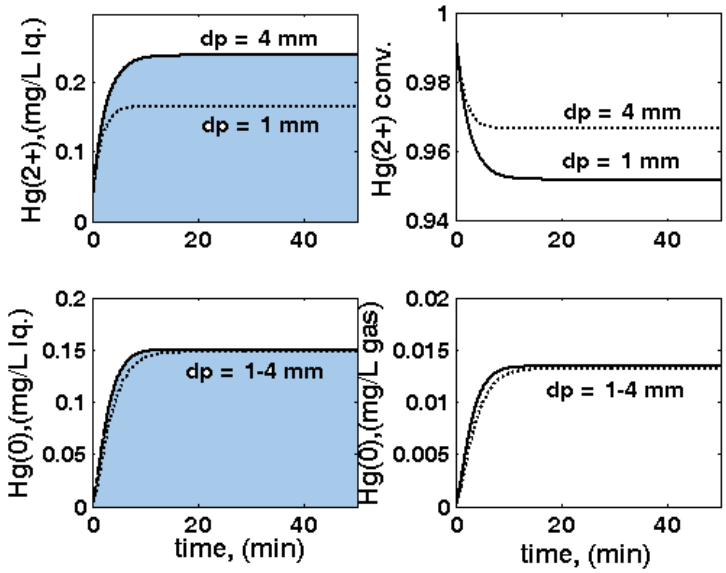


Figure 2-54: Three-phase fluidized bed reactor sensitivity related to particle size d_p variations, leading to variations of the following operating variables: outlet [Hg_1^{2+}] (up-left); outlet [Hg_L^0](down-left); Hg_1^{2+} conversion (up-right); outlet mercury in the outlet-gas [Hg_G^0](down-right). Nominal conditions of Fig. 2-51.

Proof

To prove in a simple way the above statements, one considers a synthesis (replication) of the G/P pair in *E. coli*, with the generic GERM kinetics model of a simple [G(P)1] type (denoted by "2") but with only one effector (that is, the very fast buffering reaction G + P <===> GP) (Figure 2-22), or using a GERM kinetic model of the simplest [G(P)0] type (denoted by "1"), with any effector (Figure 2-12). The rate constants were estimated from solving the stationary model equations with known homeostatic concentrations given in Table 2-4 (the case of a cell with a high ballast), following the data of Maria and Scoban [2017, 2018].

By applying a dynamic perturbation, that is here, a -10% decline in [P]s at an arbitrary time t = 0, the species stationary concentrations recovering times are presented in Figure 2-17. This example reveals that the generic GERM of [G(P)1] type {with only one buffering reaction $G + P \leftrightarrow GP$, adjusting the G-"catalyst" activity, see Figure 2-24} presents a better

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

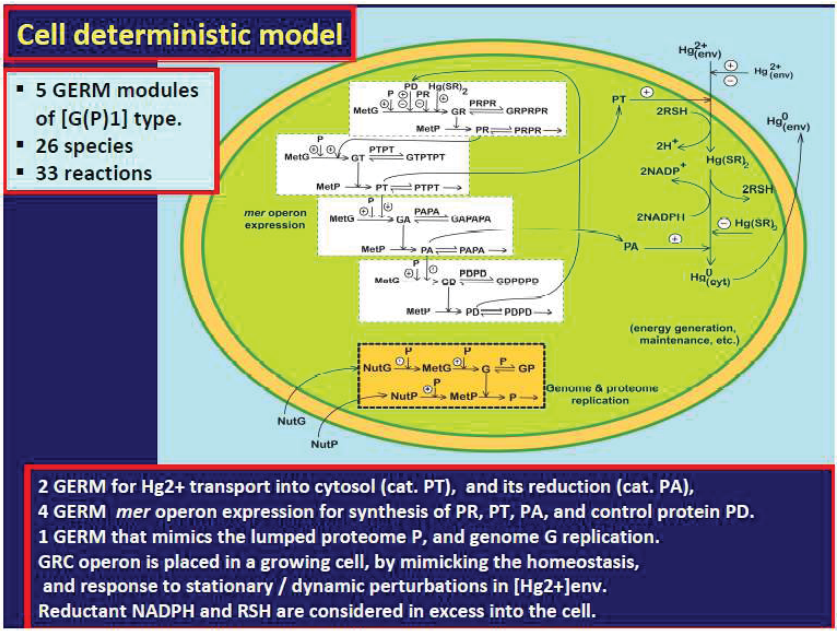


Figure 2-55: The reduced cell model, accounting for the mer-operon expression (5 linked GERM-s), coupled with 2 kinetic modules referring to the main enzymatic reactions. Adapted from [Maria, 2009b, 2010, 2005; Maria and Luta, 2013].

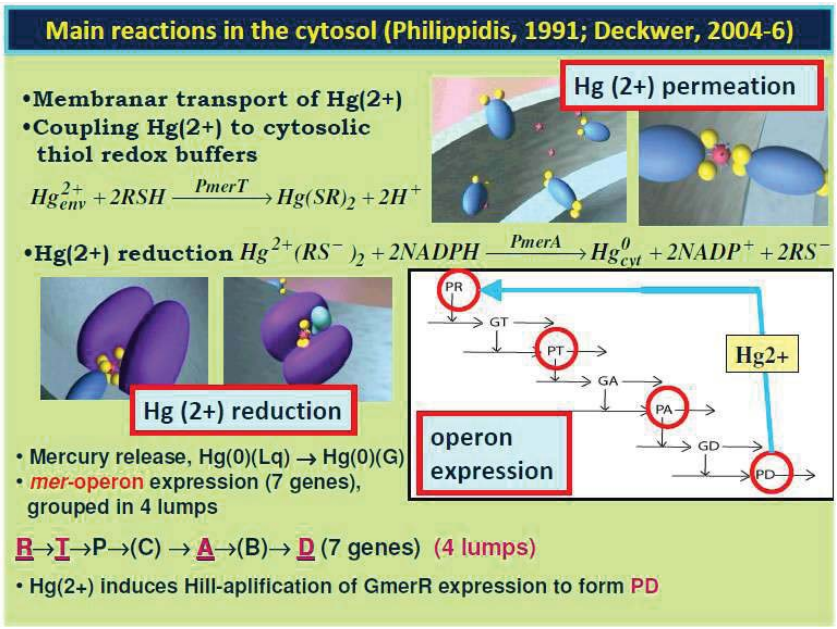


Figure 2-56: The nano-scale cellular enzymatic process in the immobilized E. coli bacteria: mer-operon (4 gene lumps) expression (Fig. 2-55), and the main enzymatic reactions (mercury permeation, and its reduction). Down-right is the sequence of mer-proteins expression. Adapted from [Maria, 2009b, 2010, 2018; Maria and Luta, 2013].

dynamic efficiency compared to the simplest [G(P)]0 gene-expression module, where any control on the G -activity exists. The obtained recovering trajectories of P and G obtained by model simulations (similar plots of Figure 2-16 and Figure 2-10) reveal a very good regulatory efficiency of [G(P)1] modules, both *G*, and *P* species presenting relatively short recovering rates, and negligible for the other species (Figure 2-17). These plots reveal in a simple way the selfregulation of the G/P pair synthesis: after the impulse perturbation leading to the decline of [P]s from 1000 nM to 900 nM. Then, the very fast buffering reaction G + P <===> GP leads to restore the active G, whose concentration quickly increases to [G] = 1.027nM. As a consequence, the synthesis rate of P increases leading to a fast P recovering rate which, in turn, contributes to the recovering of G-lump steady-state. For comparison, as revealed by the results displayed in the Figure 2-17, the dynamic efficiency of the module [G(P)0] is much lower, species recovering their QSS over longer transient times, as revealed by the larger AVG = AVG($\tau(j)$). Also, the species connectivity is better in the [G(P)1]

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

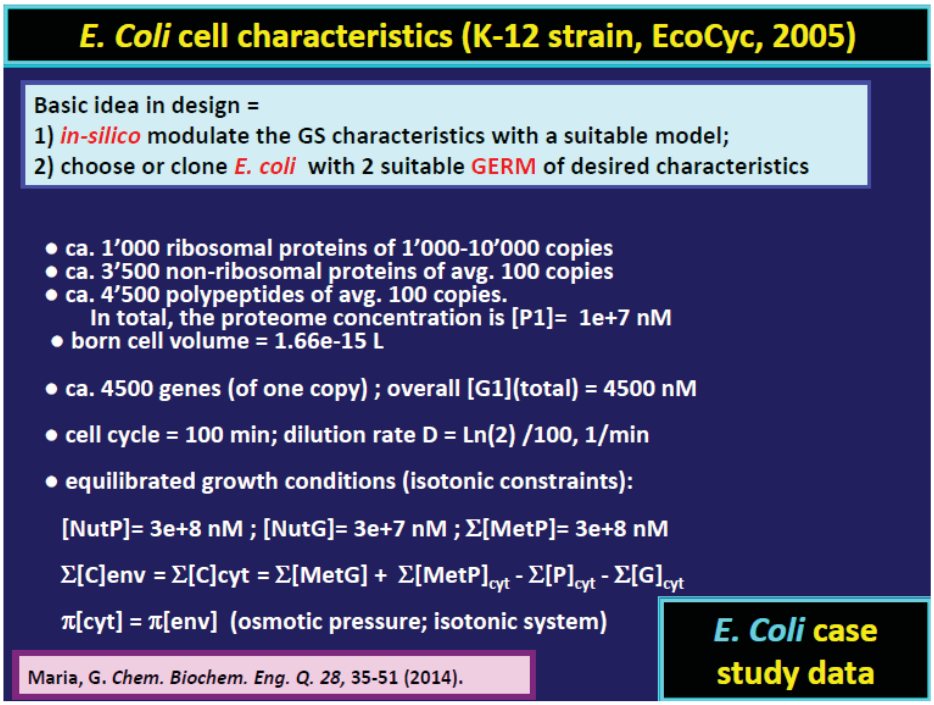


Figure 2-57: The main characteristics of the E. coli (K-12 strain) used in the case study of section 3.3. Data from [EcoCyc, 2005].

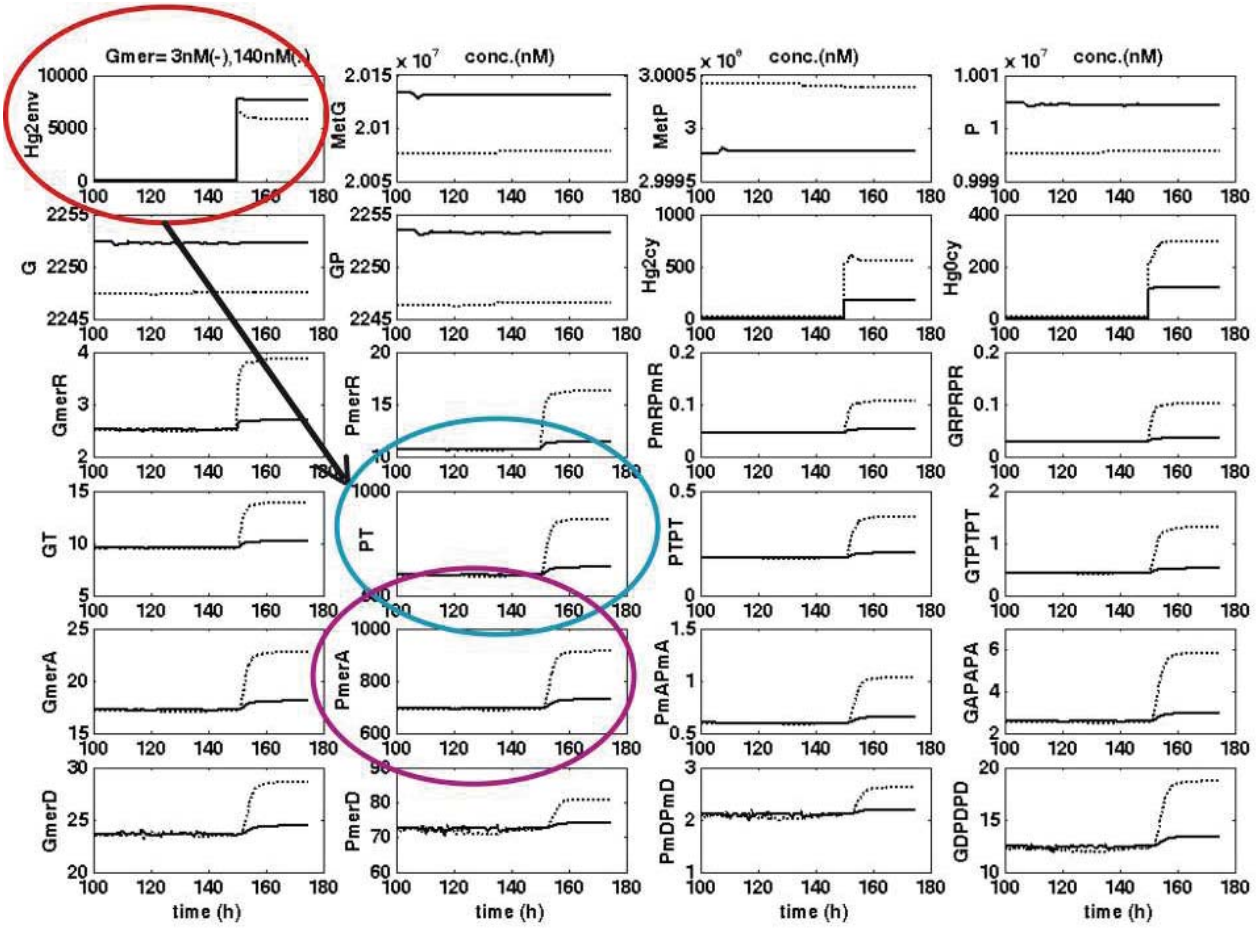


Figure 2-58: Typical evolution of relevant species concentrations predicted by the *E. coli* WCVV cell model of Maria [2009b, 2010] for mercury uptake, after a "step" perturbation in the FBR bioreactor inlet from $[Hg_{env}^{2+}]$s = 0.1 to 10 μM (ca. 2 mg/L), for the case of cell cloned with [Gmer]= 3 nM (___), or with [Gmer]= 140 nM (• • • • •). The arrow indicates the quick and vigorous response of the two key merenzymes: PT = the mercury ions permease, and PA = the mercury ions reductase.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

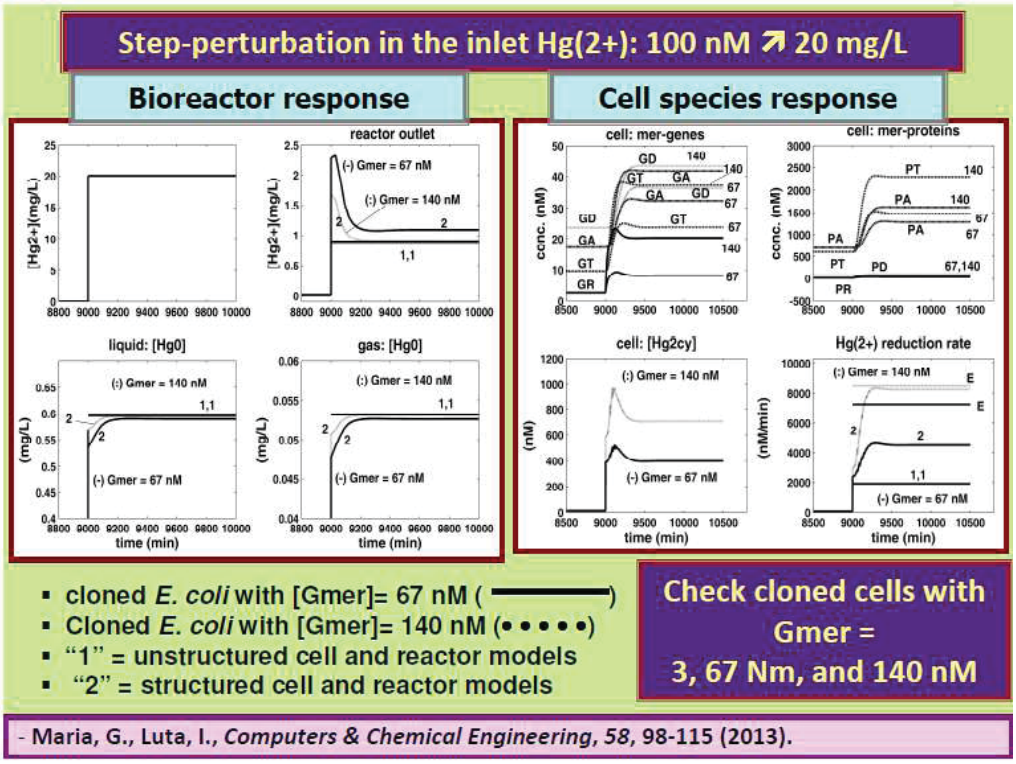


Figure 2-59: [Left]. Dynamics of metallic mercury concentrations in the liquid and gas phases, and in the bioreactor outlet after a "step" perturbation in the reactor inlet from $[Hg_{(l)}^{2+}]$ = 0.1 μM to 100 μM (ca. 20 mg L^{-1}), for immobilized cells on pumice granules under nominal conditions of Table 3-2. Comparison is made for the cases of cells cloned with [Gmer]= 67 nM (___), or with [Gmer]= 140 nM (• • • • •). The curves indexed by "1" denote predictions of the unstructured (apparent) reactor model (Table 3-1, Table 3-3), while curves indexed by "2" denote the predictions of the structured reactor model (including the cell model of Table 3-4). **[Right].** Concentration dynamics of relevant mer-species inside the cell, and the mercury reduction rate following the same "step" perturbation in the reactor inlet from $[Hg_{(l)}^{2+}]$ = 0.1 μM to 100 μM (ca. 20 mg /L), for immobilized *E. coli* cells on pumice granules under nominal conditions of Table 5-2. Notation "E" denotes the experimental curves of Philippidis et al.[1991a-c]. Adapted from [Maria and Luta, 2013].

compared to [G(P)0], being reported smaller STD($\tau(j)$), that is a smaller average STD. Also, the sensitivities of species stationary concentrations *vs.* perturbations in the external species (NutG here) are much smaller in the [G(P)1] case. Consequently, removal of the buffering reaction that automatically adjusts the "catalytic activity" of G, and of any other effector will: a) decrease the species inter-connectivity (increasing the standard deviation of the recovering times); b) will increase the species recovering times; c) will increase the sensitivities of the species steady-states *vs.* the perturbations in the external nutrients concentrations, with a negative effect on the cell metabolism, and GERM efficiency.

In the WCVV kinetic models of GERMs, of [G(P)1] type (Figure 2-24), it is also to mention the way by which the rate constants in the rapid buffering reactions are estimated, that is for the "effector" reaction type {see the nomenclature of Eq.(19)} [Maria,2005]:

$$L + O \underset{k_{diss}}{\overset{k_{bind}}{\rightleftarrows}} LO; K_{LO} = \frac{k_{bind}}{k_{diss}} = \frac{C_{LO,s}\left(1 + \frac{D_S}{k_{diss}}\right)}{C_{L,s}C_{O,s}} \tag{31}$$

As discussed by Kholodenko et al. [1998], fast buffering reactions are close to equilibrium and have little effect on metabolic control coefficients. As a consequence, rate constants of such rapid reactions are much higher than those of the core synthesis and dilution rates. To reduce the vector size of the unknown rate constants of the WCVV kinetic model Eq.(15) (in section 2.2.1.3), large values of $k_{diss} \gg D_s$ can be postulated (5 to 7 orders of magnitude higher) [Maria,2005].

Example

As expected, the P.I.-s of the GERMs depend not only on I) the no. of effectors (buffering reactions, TF), but also on ii) the TF type (P, or PP), and even more iii) on the used control scheme (i.e. simple or in cascade, see alternatives in Figure 2-12).

To exemplify the effect of the G/P *mutual autocatalysis*, one considers the generic G/P gene expression example of [Maria and Scoban, 2017, 2018] with the species homeostatic stationary concentrations given in Table 2-4 (the high ballast cell case). For comparison, one considers the gene *G* expression by means of a GERM-s of various structures (kinetic models) given in the Figure

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

2-16, that is:

"1" [G(P)0] without mutual catalysis, and any effector;

"2" [G(P)1] with mutual catalysis and one rapid buffering reaction ($G + P \Leftrightarrow GP$, to adjust the 'catalyst' G activity);

" 4 " [G(P)2] with dimeric TF = PP,

"3" [G(P)1;M(P)1] with mutual catalysis, and a cascade control via two very rapid buffering reactions at the level of *G* (that is $G + P \Leftrightarrow GP$, to adjust the 'catalyst' G activity), and at the M level (that is $M + P \Leftrightarrow MP$, to adjust the 'catalyst' M activity). The inactive complex MP is also prone to degradation.

In all the cases, the rate constants have been estimated by solving the stationary form of the GERM model Eq.(14), with the known stationary concentrations of Table 2-4 (the high ballast cell case, with $[G]s = 1nM$, $[M]s = 1nM$, for the $[\mathbf{G}(\mathbf{P})0]$, and $[\mathbf{G}(\mathbf{P})1;\mathbf{M}(\mathbf{P})1]$, respectively).

Additional constraints have been imposed to the species involved in the buffering reactions of Tables 2-4. Thus, the requirement of obtaining a maximum dynamic responsiveness and efficiency of GERMs , largely discussed by Maria [2003,2005,2006,2007,2008,2009,2017a,2017b], lead to adopt [M]s = $[M(P)]s = 0.5nm$, and $[G]s = [GP]s$, and also $[G]s = [G(PP)]s = [G(PPPP)]s = 1/3nM$.

To comparatively simulate the GERM efficiency *vs.* a dynamic perturbation, one applied a – 10% impulse perturbation in the [P]s = 1000 nM at an arbitrary t = 0. The resulted recovering trajectories of the G and P species are comparatively presented in the Figure 2-16. It is to remark that the simplest (and incomplete) [G(P)0] model reported the worst dynamic efficiency, with very small recovering rates, and very large recovering times in Figure 2-16. Better performances are reported by [G(P)1]. (see also Table 2-4). However, a better regulatory efficiency was reported by the cascade control of separately considered transcription and translation in the $[\mathbf{G}(\mathbf{P})1;\mathbf{M}(\mathbf{P})1]$ module type. The best dynamic efficiency to recover the stationary state (QSS) was reported by the [G(PP)2] module that uses two buffering reactions and a dimeric PP as TF. The PP intermediate is quickly synthesized in a small amount (of optimal small level [PP]s,= 0.01nM determined together with the model rate constants to ensure the optimal P.I. of a quick [P] s recovering time τ(rec, *P*). As a conclusion, the dynamic P.I. of these GERM-s increase in the relative order:

$$[\mathbf{G}(\mathbf{P})0] < [\mathbf{G}(\mathbf{P})1] < [\mathbf{G}(\mathbf{P})1;\mathbf{M}(\mathbf{P})1] < [\mathbf{G}(\mathbf{PP})2]$$

All the analyzed GERM-s above structures "1-4" have been modelled by using novel WCVV modelling framework Eq.(10-14). For such WCVV kinetic models it is to remark the way by which the variable cell-volume plays an important role to species inter-connectivity (direct via common reactions, or indirect via the cell volume) in the same GERM regulatory module or among linked modules (see section B-d, "Species interconnectivity").

2.2.4-IV). The effect of cell system isotonicity

The effect of the isotonicity constraint Eq.(11-15) in a WCVV cell model can be easily proved [Maria, 2017a; Maria and Scoban,2017,2018]. Thus, by adopting a GERM with a [G(P)1] model type (Figure 2-24), (see section 2.3.1), one can simulate the effect on the cell when applying a -10% impulse perturbation in the keyprotein homeostatic level $[P]s = 1000nM$ at an arbitrary time t = 0. The effect on the key-species (*G*,*P*) can be observed in the Figure 2-10, while species stationary concentrations and recovering times are given in the Figure 2-17, and in Figures 2-27, 2-28. By contrast, in a classic ("default") WCCV cell model formulation Eq.(1a-c), when the isotonicity constraint is missing from the kinetic model, the key-species do not recover after a perturbation, or they recover in an inappropriate way (see section 2.3.1). By contrast, as revealed by the simulations with the [G(P)1], WCVV kinetic model Eq.(11-15), the system isotonicity imposes relatively short recovering rates for the key-species, and negligible for the GERM species present in a large amount (that is the lumped nutrients {NutP, NutG }, and the lumped metabolites MetG, MetP). As proved by Maria [2017a, 2018], the WCVV models, with including the "cell ballast" effect (see the below paragraph no. VI), and the G/P mutual autocatalysis (see the paragraph no. III), are more flexible, adequate, and adaptable models, being able to better represent the influence of the environmental changes on the cell homeostasis.

2.2.4.-V). The importance of the adjustable regulatory effectors in a GERM kinetic WCVV model

As proved by the example of section ["2.2.3. B - a).- Recovering time"], illustrated by Figure 2-16, and by the discussion of Maria [2003, 2005], the use of dimeric TFs, such as PP in [G(PP)n] models (Figure 2-12) instead of simple P in $[\mathbf{G}(\mathbf{P})\mathbf{n}]$ model units, leads to several conclusions, as followings:

i) The dynamic regulatory efficiency increases in the order: $[\mathbf{G}(\mathbf{P})0]$ (no buffering reaction) $< [\mathbf{G}(\mathbf{P})1]$ (one buffering reaction)

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

$<[\mathbf{G(P)}1;\mathbf{M(P)}1]$ (cascade control and also a buffering reaction at the M level) < [G(PP)2] (two buffering reactions, with dimeric TF = PP).

Some GERM modules reported an increased P.I. flexibility, due to 'adjustable' intermediate TF species levels. This is the case, for instance, of adjusting [M]s in the module $[\mathbf{G(P)n};\mathbf{M(P)n}1]$, and of [PP] s in modules $[\mathbf{G(PP)n}]$. Optimal levels of these species can be set accordingly to various optimization criteria, rendering complex GRCs to be more flexible in reproducing certain desired cell-synthesis regulatory properties [Maria and Scoban, 2017, 2018; Maria, 2017a, 2018].

ii) The dynamic regulatory efficiency (defined in Table 2-2) decreases in the following order [Maria, 2003, 2005]:

$\mathrm{Min}(\tau(\mathrm{rec},\mathrm{P})$:

$$[\mathbf{G(PP)}2]>[\mathbf{G(P)}1;\mathbf{M(P)}1]>[\mathbf{G(P)}1]>[\mathbf{G(P)}0]$$

Min(STD):

$$\mathbf{G(PP)}2]>[\mathbf{G(P)}1;\mathbf{M(P)}1]>[\mathbf{G(P)}1]>[\mathbf{G(P)}0]$$

Minimum sensitivity to nutrients changes, that is:

Min S([P]; [NutG]) = (∂[P]/[P] ref / ∂ [NutG] / [NutG] ref) s:

$$[\mathbf{G(PP)}2]>\mathbf{G(P)}1;\mathbf{M(P)}1]>\mathbf{G(P)}1]>\mathbf{G(P)}0]$$

2.2.4. VI). The effect of the cell "ballast"(load) on the GERM efficiency

The cell "ballast" is understood as the total number of moles (copynumbers) of cellular species, that is $\sum_{j=1}^{ns} N_j$ in the WCVV model Eq.(10-14). Of course, this measure plays an essential role in the WCVV model properties, as long it is related to the isotonic constraint Eq.(6).

When constructing more or less simplified WCVV cell models, it is important to know what is the minimum level of simplification to not essentially affect the holistic properties of the cell model. This paragraph proves why it is essential to include in a WCVV model the so-called "**cell ballast**", that is the sum of concentrations of all (individual or lumped) species, which are accounted or not accounted in the ODE mass balance of the GERM / GRC model. Basically the isotonic constraint Eq.(6) imposes that all species (individual or lumped) to be accounted in the cell model, because species concentrations and rates are linked through the common cell volume Eq.(6,10-15). As proved by Maria [2017a], and Figure 2-23, in such WCVV cell model constructions, the recovery rates are properties of all interactions within the system rather than of the individual elements thereof [Heinrich and Schuster, 1996].

Example

Another important question derived from the isotonicity constraint refers to the degree of importance of the cell content 'size' (load, or 'ballast') for the cell reactions and efficiency to cope with the external/internal (dynamic/stationary) perturbations. In other words, the P.I.-s of a GERM are the same in a "rich" cell of high cell content (ballast), compared to those of a "poor" cell of low cell content (ballast)? The answer is no. To simply prove such a remark, one considers a simple generic GERM of $[\mathbf{G(P)}1]$ type from an *E. coli* (Figure 2-24) cell with two hypothetical different nominal conditions given in Table 2-4: a high-ballast cell, and a low-ballast cell. To not complicate these models, lumped gene and protein metabolites have been considered to simulate the cell 'balast'. Being present in a large amount (that is lumped [MetG] = 3E + 6nM, and lumped [MetP] = 3e + 8nM), these components also play the role of cell ballast, their concentrations being set to values much larger than those of other cell species. Simulations of Maria and Scoban [2017, 2018] are given in Figure 2-23). The key-species dynamics allowed obtaining the species trajectories after a – 10% impulse perturbation in the key-protein [P]s of 1000 nM applied at an arbitrary time t = 0. These recovering trajectories are presented comparatively in the Figure 2-23., for the both cases, and using the same GERM WCVV kinetic model.

Selection of appropriate lumped [MetG] and lumped [MetP] will lead to understanding their effect on the cell self-regulatory properties. A low species ballast, relative to the total number of other molecules in the cell lead to longer recovering times τ(rec)(P), and τ(rec)(G), for the keyprotein P , and for the G respectively. For instance, in the [G(P)1] module case, with lumped [MetG]=2000 nM, and lumped [MetP]=3000 nM, and (all [Cj]) = 6001 nM, the resulted recovering times of key-species G/P are long compared to tc 100 min., that is τ(rec)(P)= 103 min, τ(rec) (G)=223 min, and τ(rec)(GP)=246 min, after a –10% impulse perturbation in the [P]s of 1000 nM at an arbitrary t = 0 (Figure 2-23). Whereas for a cell of a high "ballast" case, with lumped [MetG] =3e+6nM, and lumped [MetP]=3e+8nM, and all [Cj]) = 6.06e+8nM (Table 2-4), the resulted recovering times of the same keyspecies are much shorter, that is τ(rec)(P)= 133 min, and τ(rec)(G)=93 min, and τ(rec)(GP)=93 min, after the same "impulse-like" perturbation in the [P]s of 1000

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

nM at an arbitrary t = 0.

We refer to this as the "*Inertial Effect*". It arises because the invariance relationships described above Eq.(6,10-15) require large products rate constants × [precursor metabolite P / G] for P and G synthesis to be used to counterbalance the low [MetP] and [MetG] synthesis rates, and these model rate constants are determinants for keyspecies recovering rates ($\tau(\mathrm{rec})(j)$) after a perturbation. On the other hand, when metabolite concentrations were low (that is the "low ballast" case), perturbation of cell volume was larger than when they were high, that is the cell "high ballast" case (volume increase plots not presented here). The attenuation of perturbation-induced volume changes by large metabolite concentrations is called the "*Ballast Effect*". Ballast diminishes the indirect perturbations, otherwise seen in concentrations of all cellular components. Thus, [G] was perturbed far less, as a result of an impulse perturbation in [P], for the cell containing higher metabolite concentrations than for that containing lower metabolite concentrations (Figure 2-23). Thus, increasing metabolite concentrations attenuates the impact of perturbations on all cellular components but negatively influences recovery times.

In fact, the so-called "*ballast effect*" (that is $\sum_{j=1}^{ns} N_j$, Eq.(6)) shows how all components of the cell are inter-connected via the common volume. It represents another *holistic property* of cells, and it is highlighted and evident with ***only*** using the variable-volume WCVV modelling framework. Its importance is related to the magnitude of perturbations and the total number of species in a cell. For a single perturbation in real cells, the "*Ballast effect*" [that is change of the total copynumbers $\sum_{j=1}^{ns} N_j$, Eq.(6)] will be insignificant due to the large copynumber of the total intracellular species. However, the sum of all perturbations experienced during a cell cycle might be significant.

As another remark, the perturbation in a species present in a low amount (e.g. [P] in Table 2-4) will have an insignificant effect on the species present in large amounts, like the lumped $\sum_j C_{Met} G_{j,s}$, or $\sum_j C_{MetP} P_{j,s}$ in Table 2-4 , their recovering rate after a dynamic perturbation in [P] s being negligible (see Figure. 2-28 as an example). Also, the cell volume Eq.(12), will be insignificantly influenced by such small perturbations. Thus, $\sum_{j=1}^{ns} N_j \approx 3.03 \times 10^8 \mathrm{nM}$ in Table 2-4 for the high ballast cell case, is practically unchanged after a small perturbation of 100 nM in [P]s. The cell volume Eq.(12) remains practically unchanged.

Proof the ballast effect on the cell volume variation for a dynamic perturbation in one cell species

To investigate, in a simple way, how a relatively small perturbation in a key-species concentration (P) from a GERM module of [G(P)1] type affects the main species G/P recovering rates and the cell volume change, one considers two hypothetical (demonstrative) cells with the species stationary (homeostatic) concentrations of Table 2-9. By applying a dynamic "impulse-like" perturbation of -10%[P]s, the species stationary state recovery will depend on the GERM structure. As plotted in Figure 2-38, the P-recovering rate decreases in the following order:

"1" **[G(PP)2]** < "2" **[G(P)1 ; M(P)1]** < "3" **[G(P)1]** < "4" **[G(P)0]**

The GERM structure " 5 " is a hypothetic theoretical structure, that set [G] to 1 nM at the moment of the dynamic perturbation, which is practically impossible to achieve in reality, so it was removed from this discussion.

If one takes the regulatory static and dynamic capabilities of the model $[\mathbf{G}(\mathbf{P})0]$ as basis, the others models predict an improvement with:

32% of the stationary and 19% dynamic regulatory capabilities by [G(P)1];

118% of the stationary and 50% dynamic regulatory capabilities by [G(PP)2];

While the irreversible synthesis reactions of the regulatory schema seems to be responsible on compensating the diluting effect of the cell growth, the fast reversible P1-buffering reactions are responsible for managing the P1 random perturbations which can dynamically occur due to other inner cell-processes. Indeed, if one study the cell-system behaviour after a small, usually ± 10%, perturbation in the P1-level, one can determine by simulation the required time (denoted with $\tau_{rec,j}$) necessary to a species "j" to recover its stationary concentration with a certain tolerance (1%of its nominal P1-level, according to Maria [2005]). As $\tau_{rec,j}$ time is shorter, as the cell-system is better regulated *vs.* internal perturbations. By evaluating $\tau_{rec,j}$ for every GERM model type of Figure 2-12, under the same conditions, a direct comparison of the model dynamic regulatory capabilities is possible (Figure 2-38). The model [G(PP)2] (with three P-buffer reactions) predicted a very quick cell-recovering, faster than those predicted by the other models, like $[\mathbf{G}(\mathbf{P})1]$ (with one P-buffer reaction), and $\mathbf{G}(\mathbf{P})0]$ (with no buffer reaction).

As another important observation in Figure 2-38 [Down, plots of (A)], the cell volume contracts after applying the dynamic perturbation of -10%[P]s, corresponding to the isotonicity constraint Eq.(6). Of course, the cell volume presents a smaller contraction

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 2-5: The stationary homeostatic (QSS) species concentrations "Cj,s" considered for a GERM kinetic model of [G(P)1] type, in the example of section 2.3.1. The reaction pathway, and the reaction rates expressions are given by Eq.(32). These concentrations corresponds to a GERM from E. coli cell K-12 strain [Tian and Burrage, 2006; Myers, 2009; Maria, 2005]. Values are adapted from the example of [Kurata and Sugimoto, 2017]. Notations [footnote F]. The last column displays the species recovering times for this GERM of [G(P)1] type, after a −10% impulse perturbation in the key-protein stationary [P]s produced at an arbitrary moment t=0.

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 $\Sigma_j MetG_{j,s}$	3×10^{6}	Approximated from [footnote C]	NG
Lump $\Sigma_j MetP_{j,s}$	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
cell initial volume, $V_{cyt,o}$ (L)	$1.660434503\times10^{-15}$	Bar-Joseph et al. [2012]	
cell cycle, t_c (min)	100 (min.)	Adopted [Tiruvadi-Krishnan et al.,2022]	
$Ds = Dm$	Ln(2)/ t_c	The average (apparent cell dilution rate); Eq.(12)	
Eigenvalues of the Jacobian of the [G(P)1] kinetic model, Eq. (4,11,34) {∂ h(j)/ $\partial[\Sigma_j MetG_{j,s}$, $\Sigma_j MetP_{j,s}$, $C_{P,s}$, $C_{G,s}$, $C_{GP,s}$]	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) rate is: $Ds = Dm = \mathrm{Ln}(2)/\ t_c$, according to Eq.(12). Lumps NutP and NutG denote substrates used in the synthesis of metabolites MetP and MetG respectively, which, in turn, are further used for the P and G synthesis (see **Fig. 2-24**).

[B] Species copynumbers correspond to the *E. coli* cell K-12 strain [Tian and Burrage, 2006; Maria,2005; Myers, 2009]. Species concentrations in the cell are computed with the formula [Kurata and Sugimoto, 2017]:

$$C_j = \frac{N_j}{N_A \times V_{cyt,o}}$$

where N_A is the Avogadro number. For instance, for a born *E. coli* cell, with an approximate volume $V_{cyt,o} = 1.66\times10^{-15} L$ [Bar-Joseph et al.,2012], the concentration of one gene G copynumber 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.

[C] The lump $\Sigma_j C_{MetG_{j,s}}$ results from the evaluated from the isotonicity constraint $RT/\pi_{cyt} = RT/\pi_{env}$, which translates in $\sum_{all\ j}^{cell} C_{j,s} = \sum_{all\ j}^{env} C_{j,s} = C_{NutG,s} + C_{NutP,s} = \Sigma_j C_{MetG_{j,s}} + \Sigma_j C_{MetP_{j,s}}$.

[C] The lump $\Sigma_j C_{MetG_{j,s}}$ results from the isotonic constraint Eq.(6), that is ($\pi_{cyt} = \pi_{env}$), of **Table 2-1** , leading to the relationship: $C_{NutG} + C_{NutP} = \sum_j C_{MetG_j} + \sum_j C_{MetP_j} + \left(\sum_{\substack{all\ cell \\ j\neq MetP_j \\ j\neq MetG_j}}^{} C_j\right)$ Consequently, one obtains:

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

[D] The $\mathrm{Max}(\mathrm{Re}(\lambda_j)) < 0$ evaluated at the species stationary (**QSS**) values (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 [2002,2005,2006,2007,2008,2209,2017a], Eq.(17) [see section 2.2.1.3), the stationary concentrations of the active ("G") and inactive ("GP") form of the catalyst ("G" here) should be equal.

[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 P; C_j = species "j" concentration; cyt = cytoplasma; "o"= initial; 's' index referring to the steady-state (**QSS**, homeostasis); nM = nano-molar; NG = negligible.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 2-6: The estimated rate constants of the GERM kinetic model (Fig.2-24) of type [G(P)1] from the case study of section 2.3.1, solved under a WCVV modelling 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 2-5. Adopted k6 value much larger than Ds and the other rate constants, see the discussion of Maria [2005].

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
$GP \xrightarrow{k_6} G + P$	$r_6 = k_6 C_{GP}$	$1 \times 10^{+5}$ [*]

[*] The adopted k6 rate constant satisfies the condition k6 ≫Ds=Dm, according to Eq.(31). Here, Dm= the apparent average cell dilution rate Eq.(12), that is: Dm=*Ln*(2)/*tc*. As here, *tc*=100" " min. (Table 2-5), it results *Dm*=6.93×10^{-3} 1/min. This value was used in Eq.(40) to estimate these rate constants.

Table 2-7: Nominal stationary conditions (index 's') for the two considered GERM-s from E. coli cell to be linked to form a complex GRC. These values are explained by Maria [2006] and concern the first example of section 2.2.4-VII, which considers two linking alternatives, as followings: Alternative [A] : [G1(P1)1] + [G2(P2)1] (10 individual and lumped components). Alternative [B] : [G1(P1)1;M1(P1)1] + [G2(P2)1;M2(P2)1] (14 individual and lumped components). Species copynumbers correspond to the E. coli cell K-12 strain [Tian and Burrage, 2006; Maria,2005; Myers, 2009]. Species concentrations in the cell are computed with the formula from the Footnote [B] of Table 2-5.

Symbol	Significance	Value
$V_{cyt,o}$	initial cell (cytoplasma) volume [Bar-Joseph et al.,2012]	1.66×10^{-15} L
t_c	cell cycle duration, adopted from [Tiruvadi-Krishnan et al.,2022]	100 min
D_s	cell-volume average logarithmic growing rate, Eq. (12)	ln(2)/tc, 1/min
$c_{NutG,s}$	nutrient for G-production (concentration in the environment)	3×10^{6} nM
$c_{NutP,s}$	nutrient for P production (concentration in the environment) $\approx \sum_j c_{MetP_{j,s}}$	3×10^{8} nM
$\sum_j c_{MetP_{j,s}}$	overall metabolite concentration for P-production $\left(c_{MetP_{j,s}}\right)_{\text{per module}} = \left(\sum_j c_{MetP_{j,s}}\right)/(\text{no. of modules})$	3×10^{8} nM
$\sum_j c_{MetG_{j,s}}$	overall metabolite concentration for G production (derived from footnote [C] of Table 2-5; $\left(c_{MetG_{j,s}}\right)_{\text{per module}} = \left(\sum_j c_{MetG_{j,s}}\right)/(\text{no. of modules})$	~10^{6} nM
$c_{P1,s}$	protein P1 concentration (in module 1)	1000 nM
$c_{P2,s}$	protein P2 concentration (in module 2)	100 nM
c_{G1s} , c_{G2s}	gene G1 (in module 1) or G2 (in module 2) concentration	0.5 nM
c_{M1s} , c_{M2s}	intermediate M1 (in module 1) or M2 (in module 2) concentration	0.5 nM
$c_{GjPj,s}$	concentration of catalytically inactive forms of G_j (in module *j*)	0.5 nM
$c_{M1P1,s}$, $c_{M2P2,s}$	concentrations of catalytically inactive forms of M1 (in module 1) or M2 (in module 2)	0.5 nM

for the case of a cell with a large content ("ballast"), that is here the hypothetical case of [MetG]s =[MetP]s = 10,000nM (Table 2-9, and Figure 2-38 Down, plots "1"). By contrast, the cell volume presents a significant contraction for the case of a cell with a small content ("ballast"), that is here the hypothetical case of [MetG]s = [MetP]s = 250nM (Table 2-9, and Figure 2-38 Down, plots "2").

Although quick reversible buffer reactions are responsible for dynamic regulatory capabilities of the model, a too large number increases the model complexity, with a certain number of disadvantages: i) makes difficult GERM model identification (due to a larger number of rate constants to be estimated); ii) more important, it will become very difficult to use such effective GERM-s to construct complex GRC-s (see section 2.2.4), due to very high difficulties to estimate the very large number of rate constants from often incomplete / un-structured experimental kinetic data. A roughly 2-3 effectors adopted per GERM module usually leads to satisfactory results (similarly to the allosteric enzymatic reactions [Maria, 2014b; Hammes and Wu, 1974]).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 2-8: Nominal stationary conditions (index 's') for the two considered GERM-s from E. coli cell to be linked to form a complex GRC. The E. coli cell species stationary concentrations are explained by Maria [2008]. These values concern the second example of section 2.2.4-VII, which considers two linking alternatives, as followings: Alternative [A] : [G1(P1)2] + [G2(P2)2] . Alternative [B] : [G1(P1P1)1] + [G2(P2P2)1]. Species copynumbers correspond to the E. coli cell K-12 strain [Tian and Burrage, 2006; Maria,2005; Myers, 2009]. Species concentrations in the cell are computed with the formula from the Footnote [B] of Table 2-5.

Symbol	Significance	Value
$V_{cyt,o}$	initial cell (cytoplasma) volume [Bar-Joseph et al.,2012]	1.66×10^{-15} L
t_c	cell cycle duration, adopted from [Tiruvadi-Krishnan et al.,2022]	100 min
D_s	cell-volume average logarithmic growing rate, Eq. (12)	ln(2)/tc, 1/min
$c_{NutG,s}$	nutrient for G-production (concentration in the environment)	3×10^6 nM
$c_{NutP,s}$	nutrient for P production (concentration in the environment) $\approx \sum_j c_{MetP_{j,s}}$	3×10^8 nM
$\sum_j c_{MetP_{j,s}}$	overall metabolite concentration for P-production; $(c_{MetP_{j,s}})_{\text{per module}} = (\sum_j c_{MetP_{j,s}})/(\text{no. of modules})$	3×10^8 nM
$\sum_j c_{MetG_{j,s}}$	overall metabolite concentration for G production (derived from footnote [C] of Table 2-5; $(c_{MetG_{j,s}})_{\text{per module}} = (\sum_j c_{MetG_{j,s}})/(\text{no. of modules})$	~10^6 nM
$c_{P1,s}$	protein P1 concentration (in module 1)	1000 nM
$c_{P2,s}$	protein P2 concentration (in module 2)	100 nM
$c_{M1P1,s}$, $c_{M2P2,s}$	concentrations of catalytically inactive forms of M1 (in module 1) or M2 (in module 2)	0.5 nM
c_{G1s}, c_{G2s}	gene G1 (in module 1) or G2 (in module 2) concentration	0.5 nM
$c_{GjPj,s}$, $c_{GjPjPj,s}$	concentration of catalytically inactive forms of Gj (in module j)	0.5 nM
$c_{P1P1,s}$, $c_{P2P2,s}$	intermediates of adjustable levels that maximize the P.I.-s of GRC-s (here, the P1,P2 recovering times after a dynamic -10% [P]s perturbation	0.01 nM

Notations: 'cyt' = cytoplasma; G = gene; P = protein; 'o'= initial; M = mRNA; NutP, NutG = nutrients for the MetP and MetG synthesis respectivelly; MetP, MetG = metabolites for the P and G synthesis respectivelly; GP, GPP, MP = catalytically inactive species); PP= dimeric intermediate species of very low concentration used for quickly adjusting the 'catalyst' G activity. They are quickly produced from the expressed P protein.

Calculation of perturbed volume and perturbed species concentrations in a WCVV cell model when removing copynumbers from only one species (P), by keeping all-times the system isotonicity:

To illustrate the cell volume variation when a dynamic perturbation occurs, by removing a certain number of moles (copynumbers) from the cell, one onsiders the following simple numerical proof.

Problem

Given a cell with a certain number of species j = 1,...,ns, of initial concentrations C_{jo}, and the initial volume V_o (before perturbation), How many number of moles of species P have to be removed from the cell, that is ($N_{P,o} - N_P^*$) such that, after recalculating the contracted volume value V^*, and after re-calculating the species new concentrations after contraction c_j^*, the new (perturbed) concentration of P to be an initial imposed value $C_P^* = \frac{N_P^*}{V^*}$? One imposed the system isotonicity constraint, that is all-times the osmotic pressure π= constant, and the sum of concentrations $\sum_{all} C_{jo} = \sum_{all} C_j^* =$ constant (according to Eq.(6), and WCVV model hypotheses of Table 2-1).

Notations of the Problem (continued)

Notations:

V_o = initial (before perturbation) cell volume (cytoplasm);

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 2-9: Nominal stationary conditions (index 's') for considered GERM from E. coli cell of [G(P)1] type, for the example of section 2.2.4 , paragraph (VI). The E. coli cell species stationary concentrations are explained by Maria [2002,2008]. See [Maria, 2005,2006,2007,2008,2009,2017a,2017b,2018] for details. Species copynumbers correspond to the E. coli cell K-12 strain [Tian and Burrage, 2006; Maria,2005; Myers, 2009]. Species concentrations in the cell are computed with the formula from the Footnote [B] of Table 2-5.

Symbol	Significance	Value (nM)	
		Low ballast cell	High ballast cell
[G]s	gene G concentration	0.5	0.5
[P]s	Key protein P concentration (expressed by the encoding gene G)	1000	1000
[GP]s	concentration of catalytically inactive forms of G	0.5	0.5
[NutG]s	nutrient for G-production (concentration in the environment)	100	10^5
[NutP]s	nutrient for P production (concentration in the environment)	100	10^5
[MetG]s	overall metabolite concentration for G production	250	10,000
[MetP]s	overall metabolite concentration for P-production;	250	10,000
$V_{cyt,o}$	initial cell (cytoplasma) volume [Bar-Joseph et al.,2012]	$1.660434503\times10^{-15}$ L	
t_c	cell cycle duration, adopted from [Tiruvadi-Krishnan et al.,2022]	100 min	
D_s	cell-volume average logarithmic growing rate, Eq.(12)	ln(2)/tc, 1/min	

V^* = the cell volume after removing some copynumbers of P ;

C_{Po} = initial concentration of species P;

$\sum_{j\neq P} C_{jo}$ = sum of species concentrations, excepting P ;

$\frac{RT}{\pi} = \frac{1}{\sum_j^{ns} C_{jo}} =$ constant. This is the system invariant (i.e. π = constant), Eq. (13); =ns= number of cell species;

$C_{Po} = \frac{N_p}{V_o} =$ the P initial concentration referred to the initial (stationary) cell volume;

$C_P^T = \frac{N_P^*}{V_o} =$ the P concentration to be set at V_o to lead to a contracted V^* after perturbation, such that to finally obtain the target (desired)

$C_P^* = \frac{N_P^*}{V^*}$. the P concentration referred to the new (perturbed) cell volume

Solution:

The initial number of moles of cell species is:

$N_{PO} = C_{PO}V_O =$ initial no. of moles of species "P";

$\sum_{j\neq P}^{cell} N_{jo} = \sum_{j\neq P}^{cell} C_{jo}V_o =$ initial sum of no. of moles in the cell (that is the cell "ballast"), by excepting P ;

Let us consider as being invariant the number of moles of other species into the cell (other than P), before and after the perturbation by removing some copynumbers of P , that is:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

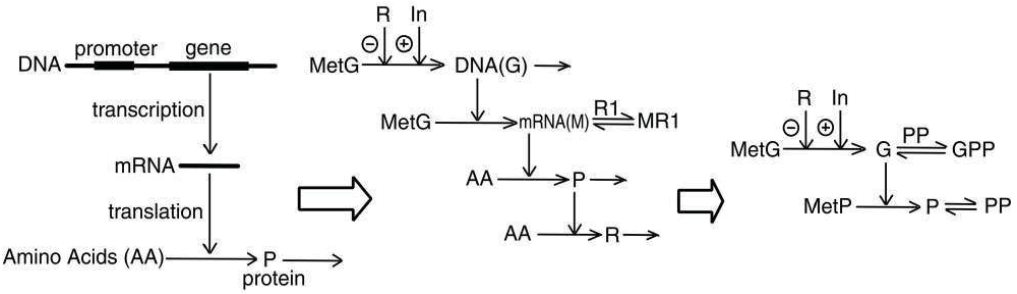


Figure 2-60: Protein P synthesis – simplified representations of a generic GERM regulatory module of [G(PP)1] type (see Fig. 2-12; horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; G = gene encoding P; M = mRNA; R = repressor; In = inducer; MetG, MetP = metabolites used for the synthesis of G,P respectively). According to the studies of Maria [2005, 2017, 2017a, 2018, 2014b]; Maria et al. [2018d, 2019] (sections 2.2.3 – Ba, Bb; 2.2.4-II; 2.2.4-III; 2.2.4-VII) ; Fig. 2-9, Fig. 2-13, Fig. 2-16, and Fig. 2-37), this GERM unit with dimeric TF displays one of the best regulatory efficiency P.I.-s (section 2.2.3).

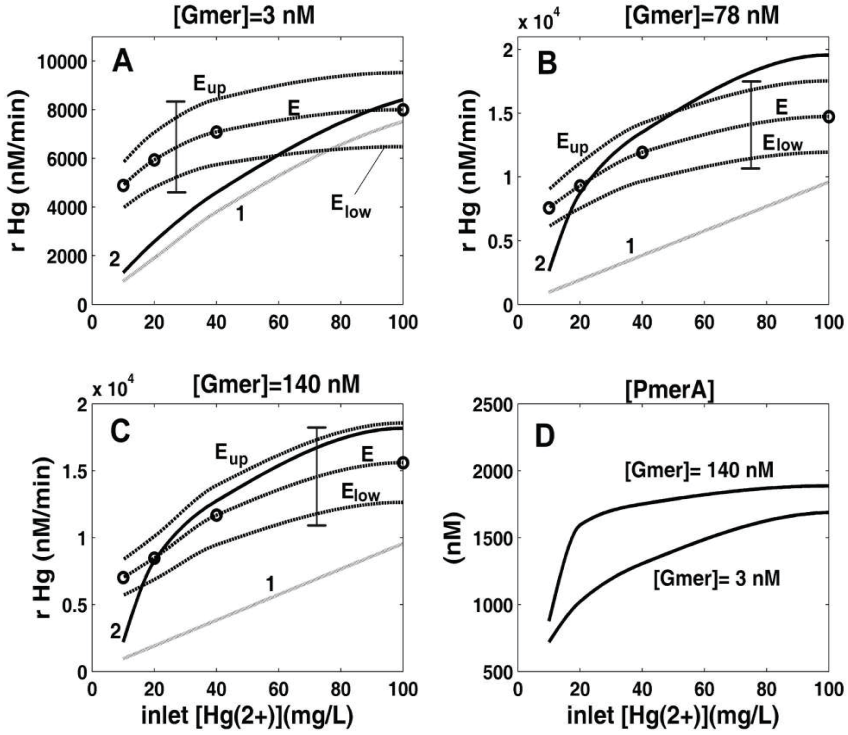


Figure 2-61: The stationary reduction rate (rHg) of mercuric ions by E. coli cells for various inlet [Hg_L^{2+}]. Comparison includes predictions of the structured extended hybrid SMDHKM WCVV dynamic model ("2 ___") vs. the Philippidis et al.[1991ac] experimental curves ("E -----") (diffusion free). The upper / lower ("Eup", "Elow") denote the confidence bounds of the experimental curves, that is in the experimental standard deviation band. Curves denoted by ("1 • • • • •") are the predictions of the unstructured bioreactor model (section 3.2.3; Table 3-1 , and Table 3-3), with including the mass transport terms. Comparison is made for the E. coli cells cloned with *mer*-plasmids in the amount of [Gmer] = 3 nM (A), [Gmer] = 78 nM (B), and [Gmer] = 140 nM (C). Plots (D) display the predicted cytosolic concentration of *mer*-reductase [PmerA] = [PA] for [Gmer] = 3 nM, and [Gmer] = 140 nM. Adapted from [Maria and Luta, 2013].

$$\sum_j^{all\ cell} N_{jo} - N_{Po} = \sum_j^{all\ cell} N_j^* - N_P^* \quad \text{V1}$$

By multiplying Eq. (V1) with RT/ π constant value, one obtains:

$$\frac{RT}{\pi}\left(\sum_j^{all} N_{jo} - N_{Po}\right) = \frac{RT}{\pi}\left(\sum_j^{all} N_j^* - N_P^*\right) \quad \text{V2}$$

One takes into account that the volume calculation relationship, Eq.(6) can be re-written as followings:

$$V_o = \frac{RT}{\pi}\left(\sum_j^{all} N_{jo}\right); V^* = \frac{RT}{\pi}\left(\sum_j^{all} N_j^*\right) \quad \text{V3}$$

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

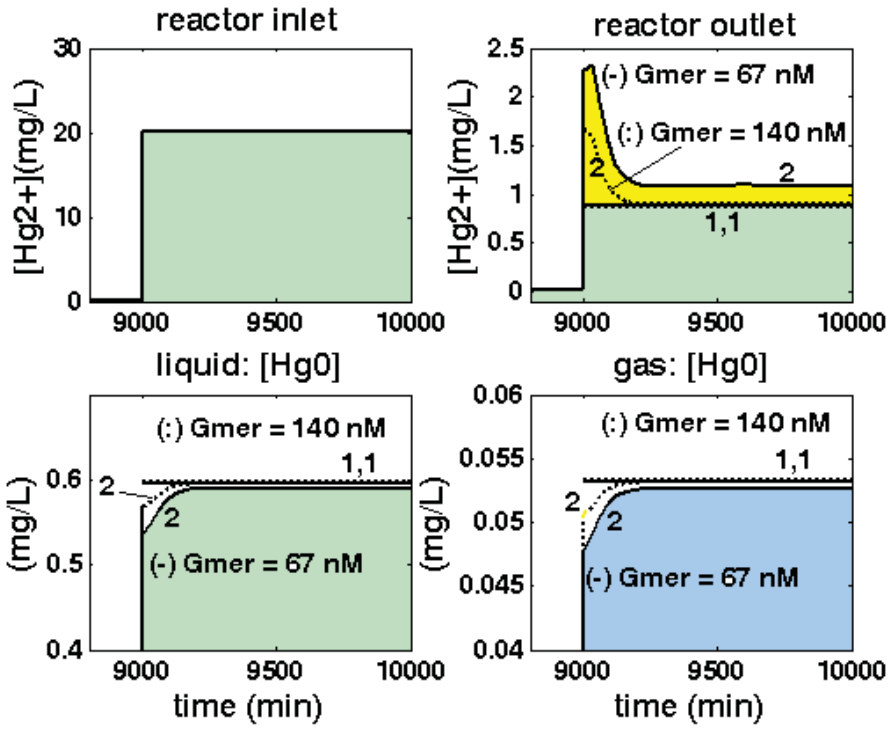


Figure 2-62: Dynamics of mercury concentrations in the liquid and gas phases after a "step" perturbation in the reactor inlet from $[Hg_L^{2+}]$ = 0.1 μM to 100 μM (ca. 20 mg/L), for immobilized cells on pumice granules under nominal conditions of (Table 3-2). Comparison is made for the cases of cells cloned with [Gmer]= 67 nM (___), or with [Gmer]= 140 nM (• • • • •). The curves indexed by "1" denote predictions of the ***unstructured*** (apparent) reactor model (Table 3-1, together with Table 3-3), while curves indexed by "2" denote the predictions of the structured HSMDM reactor model (including the cell model of Table 3-4). Adapted from [Maria and Luta, 2013].

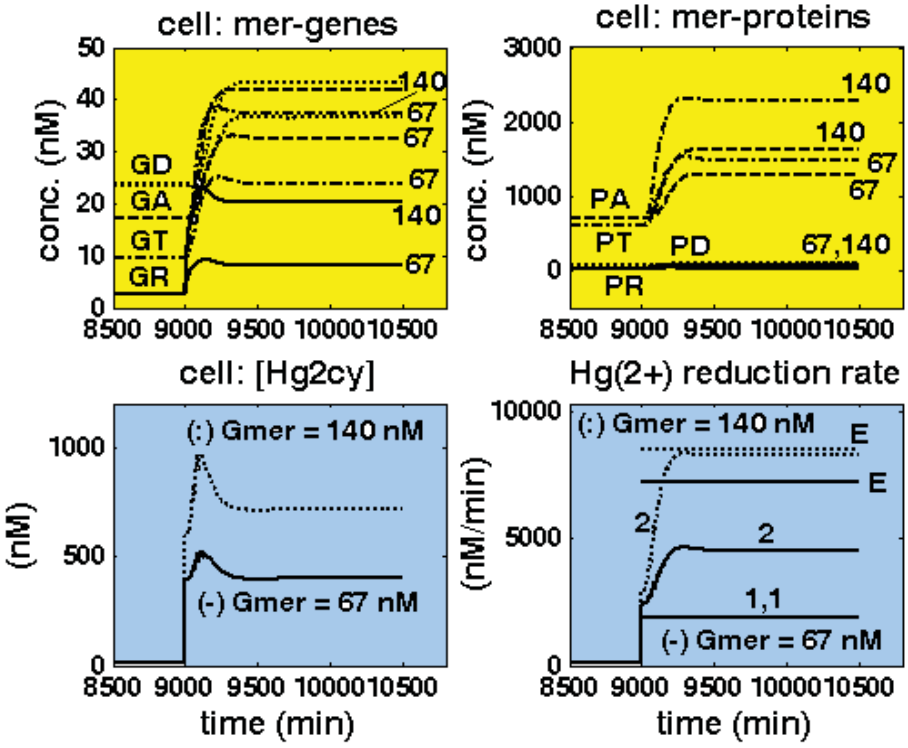


Figure 2-63: Concentration dynamics of relevant cell species, and mercury reduction rate following a "step" perturbation in the reactor inlet from $[Hg_L^{2+}]$ = 0.1 μM to 100 μM (ca. 20 mg/L), for immobilized E. coli cells on pumice granules under nominal conditions of (Table 3-2). Comparison is made for cloned cells with [Gmer] = 67 nM (67, ___), or with [Gmer] = 140 nM (140, • • • • •) mer-plasmids in the nucleus. The curves indexed by "1" denote predictions of the unstructured (apparent) reactor model (Table 3-1, together with Table 3-3), while the curves indexed by "2" denote the predictions of the hybrid SMDHKM structured reactor model (by also including the cell model of Table 3-4). Notation "E" denotes the experimental curves of Philippidis et al. [1991a-c]. Adapted from [Maria and Luta, 2013].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

By writing Eq. (V2), with changing terms from left to right, one obtains:

$$\frac{RT}{\pi}\left(\sum_j^{all} N_j^* - \sum_j^{all} N_{jo}\right) = \frac{RT}{\pi}\left(N_P^* - N_{Po}\right) \quad \text{V4}$$

By applying the volume formula (V3) to (V4), one obtains the volume variation due to the variation in the P content:

$$\left(V^* - V_o\right) = \frac{RT}{\pi}\left(N_P^* - N_{Po}\right) \quad \text{V5}$$

Let us to divide Eq. (V5) by the product (*VV**). One obtains:

$$\frac{1}{V_o} - \frac{1}{V^*} = \frac{V^* - V_o}{V^* V_o} = \frac{1}{V^*}\frac{RT}{\pi}\frac{1}{V_o}\left(N_P^* - N_{Po}\right) \quad \text{V6}$$

By multiplying Eq. (V6) with N_P^*, one obtains:

$$\frac{N_P^*}{V_o} - \frac{N_P^*}{V^*} = \frac{N_P^*}{V^*}\frac{RT}{\pi}\left(\frac{N_P^*}{V_o} - \frac{N_{Po}}{V_o}\right) \quad \text{V7}$$

The number of moles of P to be removed initially from the cell is $\left(N_{Po} - N_P^*\right)$.

Let us denote with $C_P^T = \frac{N_P^*}{V_o}$ the concentration of P to be set at V_0 to lead to a contracted V^* after perturbation. So, we can re-write Eq. (V7) in the new terms, as followings:

$$C_P^T - \frac{N_P^*}{V^*} = \frac{N_P^*}{V^*}\frac{RT}{\pi}\left(\frac{N_P^*}{V_o} - \frac{N_{Po}}{V_o}\right) \quad \text{V8}$$

By changing in Eq. (V8) the terms from right to left, one obtains:

$$C_P^T - \frac{N_P^*}{V^*}\frac{RT}{\pi}\frac{N_P^*}{V_o} = \frac{N_P^*}{V^*} - \frac{N_P^*}{V^*}\frac{RT}{\pi}\frac{N_{Po}}{V_o} \quad \text{V9}$$

In the relationship Eq. (V9), by substituting the definition of $C_P^T = \frac{N_P^*}{V_o}$, and after multiplying the first right term with *RT*/*RT* it can be re-written in the following form:

$$C_P^T - C_P^T\frac{RT}{\pi}\frac{N_P^*}{V^*} = \frac{N_P^*}{V^*}\left(\frac{RT}{\pi} / RT / \pi\right) - \frac{N_P^*}{V^*}\frac{RT}{\pi}\frac{N_{Po}}{V_o} \quad \text{V10}$$

By giving common factor $C^T{}_P$ in left side, and the common factor $\frac{N_P^*}{V^*}\frac{RT}{\pi}$ in the right side, one obtains:

$$C_P^T\left(1 - \frac{N_P^*}{V^*}\frac{RT}{\pi}\right) = \frac{N_P^*}{V^*}\frac{RT}{\pi}\left(1 / \frac{RT}{\pi} - \frac{N_{P,o}}{V_o}\right) \quad \text{V11}$$

By introducing the invariant $\frac{RT}{\pi} = \frac{1}{\sum_j^{all} C_{jo}}$ of Eq.(13) in the right side of Eq.(V11), one obtains:

$$C_P^T\left(1 - \frac{N_P^*}{V^*}\frac{RT}{\pi}\right) = \frac{N_P^*}{V^*}\frac{RT}{\pi}\left(\sum_j^{all} C_{jo} - C_{Po}\right) \quad \text{V12}$$

After substituting $C_P^* = \frac{N_P^*}{V^*}$ in Eq. (V12), one obtains:

$$C_P^T\left(1 - C_P^*\frac{RT}{\pi}\right) = C_P^*\frac{RT}{\pi}\left(\sum_j^{all} C_{jo} - C_{Po}\right) \quad \text{V13}$$

So, the final formula to set the initial perturbed P concentration at V_0 to obtain the target C_P^* after this initial perturbation, is the following:

$$C_P^T = C_P^*\frac{RT}{\pi}\left(\sum_j^{all} C_{jo} - C_{Po}\right) / \left(1 - C_P^*\frac{RT}{\pi}\right) \quad \text{V14}$$

Concerning this short numerical example of how the "cell ballast" is used in the WCVV modelling framework, it is worth introducing some remarks, as followings:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Remark no. 1.

After the cell volume contraction in Eq.(V3), from V_0 to V^*, due to the removed P copynumbers, from the initial $C_{Po}=\frac{N_p}{V_o}$ to the $C_P^T=\frac{N_P^*}{V_o}$, in order to finally obtain the target $C_P^*=\frac{N_P^*}{V^*}$, by preserving the same pressure, the contracted cell volume presents the following value:

$$\frac{V^*}{V_o}=\frac{(RT/\pi)\sum_j^{all}N_j^*}{(RT/\pi)\sum_j^{all}N_{jo}} \quad \text{V15}$$

Remark no. 2.

After removing some species P copynumbers $\left(N_P-N_P^*\right)$ to change the initial $C_{Po}=\frac{N_P}{V_o}$ to the required $C_P^T=\frac{N_P^*}{V_o}$, to obtain an imposed final value $C_P^*=\frac{N_P^*}{V^*}$ after volume contraction, the concentration sum remains unchanged, because also the cell volume is contracted after removing the P copynumbers $\left(N_P-N_P^*\right)$.

Indeed, if one calculates the final concentrations from their definition and substituting the perturbed volume V^* from (V15), one obtains:

$$C_j^*=\frac{N^*}{V^*}=\frac{N^*}{V_o\left(\sum_j^{all}N_j^*/\sum_j^{all}N_{jo}\right)} \quad \text{V16}$$

By summing Eq. (V16) for all species, one obtains:

$$\sum_j^{all}C_j^*=\frac{\sum_j^{all}N_j^*}{V_o\left(\sum_{j=1}^{all}N_j^*/\sum_{j=1}^{all}N_{jo}\right)}=\frac{\sum_j^{all}N_{jo}}{V_o}=\sum_j^{all}C_{jo}=\text{ constant} \quad \text{V17}$$

So, after such a small perturbation, the isotonicity constraint leads to an invariant sum of concentrations, because the small variations in the copynumbers are compensated by small variations in the cell volume, so their ratio is practically unchanged.

Numerical example:

To illustrate the theoretical considerations approached in this paragraph, one used a simple numerical example, of a hypothetical cell with the volume V_0 and only two species:

$C_{Po}=5\text{nM}$

$C_{Go}=5\text{nM}$

In total, according to Eq.(13), it results:

$$\frac{RT}{\pi}=\frac{1}{\sum_j^{all}C_{jo}}=1/10=\text{ constant}.$$

One desires that, after removing some copynumbers of P to finally (after volume contraction and recalculating the concentrations), to have in the perturbed cell a novel concentration of $c_P^*=4\text{nM}$.

Question: How many copynumbers of P have to be removed from the initial cell, and what concentration we have to set for P at V_0 in order to achieve this goal?

Solution:

According to Eq.(V14), one set at V_0 the following P concentration:

$$C_P^T=C_P^*\frac{RT}{\pi}\left(\sum_j^{all}C_{jo}-C_{Po}\right)/\left(1-C_P^*\frac{RT}{\pi}\right)=4(1/10)(10-5)/(1-4/10)=3.33\text{nM}$$

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Proof:

The cell species initial numbers of moles are the followings:

$$N_{Po} = C_{Po}V_O = 5V_O; N_{Go} = C_{Go}V_O = 5V_O$$.

In total, according to Eq.(13),

$$\frac{RT}{\pi} = \frac{1}{\sum_j^{all} C_{jo}} = 1/10 = \text{ constant}$$

According to the $C_P^T = \frac{N_P^*}{V_o}$ definition, the number of moles of P after removing some coynumbers but still referred to V_0 is:

$$N_P^* = C_P^T V_o = 3.33V_o$$.

Consequently, to obtain the $C_P^* = \frac{N_P^*}{V^*}$ the following numbers of moles have to be removed from the cell: $\left(N_{Po} - N_P^*\right) = 5V_O - 3.33V_O = 1.67V_O$

Consequently, according to Eq. (V15), the new volume of the cell is the following:

$$V^* = V_o \frac{(RT/\pi)\sum_j^{all} N_j^*}{(RT/\pi)\sum_j^{all} N_{jo}} = V_o \frac{N_{Go} + N_P^*}{N_{Go} + N_{Po}} = V_o \frac{5V_o + 3.33V_o}{5V_o + 5V_o} =$$
$$V_O\left[5V_O + 3.33V_O\right]/\left[5V_O + 5V_O\right] = 0.833V_O.$$

So, the species new concentrations, after considering the volume contraction, will be:

$$C_P^* = \frac{N_P^*}{V^*} = \left[3.33V_O\right]/\left[0.833V_O\right] = 4;$$
$$C_G^* = \frac{N_G^*}{V^*} = \frac{N_{Go}}{V^*} = \left[5V_O\right]/\left[0.833V_O\right] = 6.$$

So, in total, after the dynamic perturbation in P (removal of $1.67V_o$ copynumbers of P), one obtains: $C_P^* + C_G^* = 10$

As expected, the isotonicity constraint is fulfilled all times in a WCVV model, and the sum of species concentrations remains unchanged, so the osmotic pressure, because of the Eq. (13) relationship:

$$\frac{RT}{\pi} = \frac{1}{\sum_{all} C_{jo}} = \frac{1}{\sum_{all} C_j^*} = \text{constant}$$

2.2.4. VII) The effect of GERM complexity on the resulted GRC efficiency, when linking GERM-s

One important issue to be solved when linking GERM-s to construct a GRC kinetic model, with desirable regulatory properties, is the degree of detail of the adopted GERM-s to accurately reproduce the GRC local/holistic regulatory properties. The example discussed below by Maria [2006, 2008] revealed that more important than the number of considered species in the regulatory loops is the selected GERM regulatory scheme, able to render to the GRC a holistic synchronized response to environmental perturbations.

Consequently, when developing a suitable WCVV kinetic model of a GRC, it is important to adopt a suitable reduced model structure by means of an acceptable trade-off between the model simplification-*vs.*-model quality (adequacy) (see the example of Maria and Luta, 2013]).

Adoption of too complex reaction pathways for the linked GERM-s is not desirable when developing cell simulators, these structures being difficult to be modelled and difficult to be estimated by using ODE kinetic models, due to the very large number of rate constants, and of unknown steady-state concentrations (necessary to estimate the model rate constants). Beside, cell model constructions with too complex GERM-s modules lead to large models impossible to be used for cell design, or for engineering analyses. The alternative is to use reduced ODE models with a relatively small number of lumped species and enough reactions to

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

fairly reproduce the experimental data, but simple enough to make possible a quick dynamic analysis of the metabolic process and of its regulation properties.

Example no. 1.

To exemplify how a suitable trade-off between the GRC model simplicity and its capabilities can be obtained, one considers the problem of adequate and efficient linking of two GERM-s (related to the simultaneous linked expression of G1/P1 and G2/P2 pairs) such that the resulted GRC complex to present optimal P.I.-s. To solve this problem, Maria [2006] compared two linking alternatives (Figure 2-32, and Figure 2-33), that is:

Alternative A: [G1(P1)1] + G2(P2)1] (10 individual and lumped components).

Alternative B: [G1(P1)1;M1(P1)1] + [G2(P2)1;M2(P2)1] (14 individual and lumped components).

The GERM linking in the alternative [A] is the following: the expressed P1 in [G1(P1)1] module is the metabolase that converts NutG in MetG2 and NutP in MetP2 in the [G2(P2)1]. In turn, the expressed P2 in [G2(P2)1] is the polymerase that converts MetG1 in G1, and MetG2 in G2 in the modules [G1(P1)1]; M1(P1)1] + [G2(P2)1;M2(P2)1], respectively.

The GERM linking in the alternative [B] is the following: the expressed P 1 in [G1(P1)1]; M1(P1)1] module is the metabolase that converts NutG in MetG1, MetG2, and NutP in MetP1, MetP2 in the both modules. In turn, the expressed P2 in the [G2(P2)1;M2(P2)1] module is the polymerase that converts MetG1 in G1 in the module[G1(P1)1]; M1(P1)1], and converts MetG2 in G2 in [G2(P2)1;M2(P2)1]. So, clearly, each expressed protein has a well defined function in the cell. The species stationary concentrations for the two GERM modules are given in Table 2-7.

Simulations revealed that the linking alternative [B] is superior compared to the model [A], by presenting better stationary P.I.-s (Figure 2-32), and dynamic P.I.-s (Figure 2-33). The cost of such better P.I.-s is minimal in terms of GRC complexity (that is the required species number), by including 14 species into the model (alternative [B]) compared with only 10 species (alternative [A]). To conclude, in spite of a slightly more complex structure (14 *vs.* 10 individual and lumped components), and of including two more buffering reactions (at the M1, M2 species level), and by adopting a cascade expression control, the GRC alternative [B] presents much better P.I.-s (see details of Maria [2016]), that is: i) key-species shorter recovering times after an impulse perturbation; ii) lower AVG and STD species connectivity indices; iii) lower sensitivity of species QSS concentrations *vs.* environmental perturbations (in [NutP]s here).

Thus, the right choice of the GERM structures in a GRC is an essential modelling step. This example proves how, with the expense of a little increase in the model complexity (4 additional species and 2 buffering reactions), and with including of a cascade control of the gene expression (much close to the real situation) in the modules of [G(P)n;M(P)n1] type (Figure 2-12) the derived GRC-s present superior regulatory properties suitable for designing robust GRC-s, with easily adjustable properties via model parameters. As previously proved, such GRC-s, obtained by applying an optimal building-blocks strategy, present a better species synchronization when coping with perturbations (i.e. low AVG, STD indices, Figure 2-32).

Linking constraints.

In order to exemplify various rules for linking GERM-s modules to build-up a GRC, with optimum regulatory properties (P.I.s, section 2.2.3), the WCVV modelling framework has been used, as being more suitable to quantitatively characterize the GERMs P.I.-s. The WCVV approach (section 2.2) was proposed by Maria [2002], and by [Maria et al., 2002], while the study of the WCVV properties, features, and advantages was given by Maria over the interval [20052023], that is Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c].

In the WCVV approach, the cell volume is explicitly included in the mass-balance equations (sections 2.2.1, and 2.2.2). The kinetic expressions, evaluated at a certain time during the cell growth, are linked through the state-law [that is the osmotic pressure Pfeiffers'law (last row of Eq.(6)), linking the cell volume, osmotic pressure, temperature, and cell content in terms of species number of moles). The main model hypotheses are the followings (Table 2-1): i) negligible inner-cell gradients; ii) open cell system of uniform content; iii) semi-permeable membrane, of negligible volume and resistance to nutrient diffusion, following the cell growing dynamics; iv) constant inner/external osmotic pressure, ensuring the membrane integrity; v) average logarithmic growing rate $\left[Dm = \ln(2)/tc\right.$; tc = cell cycle time, Eq.(12)]; vi) volume growth of the approximate exponential law [$V = V_o \exp\left(D_m \times t\right)$, Eq.(12)]; vii) homeostatic stationary growth, $\left(dC_j/dt\right)_s = 0$ (where C_j = species " j " concentration); viii) according to the isotonic Pfeiffers'law, Eq.(6), the perturbations in the cell volume are induced by variations in species copynumbers under the isotonic osmolarity constraint.

Under such a representation, the individual or lumped reaction rates and the internal/external perturbations in species levels are explicitly linked with the evolution of the cell volume and dillution factor by Eq. (6,12), which in turn will directly influence [

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

by means the so-called 'secondary perturbations', Eq.(9,11,12)] the cell component concentrations. The result is a more accurate representation of the cellgrowth [Maria, 2005, 2006, 2017a], and the cell-division phases [Morgan et al., 2004; Surovstev et al., 2007, 2008], and a more realistic representation of the local and global regulatory properties of the GRC-s [Maria, 2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]. The rates of individual reactions are constrained by the periodicity of the cell-cycle and by the requirement that molar amounts of all components and the volume must double in exactly one cell-cycle [Tyson and Novak, 2001]. To be consistent with the WCVV hypotheses, such a 'whole-cell' modelling framework requires that each of the cell reaction to be included in the model at some level of detail, i.e. as an individual or lumped species and/or reaction, Eq. (12,15). The WCVV model analysis allows the full characterization of the GRC-s properties, by explicitly linking the model parameters to the system properties (stability, responsiveness, etc., Figure 2-14) and its effectiveness (P.I.-s, of Table 2-2). Among module P.I.-s are to be mentioned the most important ones (see Table 2-2, and section 2.2.3) [Maria ,2007,2017a]:

- System *local stability* condition, and *stability strength*:, expressed as: I) *stationary regulation*, i.e. large margins of stability in the state variable space *vs.* stationary perturbations; ii) *dynamic regulation*, i.e. fast $\tau_{rec,j}$ = species " j " recovering time of its stationary (QSS) concentration (*Cj*, *s*) with a tolerance of 5% [Hlavacek and Savageau, 1997] or 1% [Maria,2005], after an impulse-like perturbation;

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

- GRC *high selectivity*, the regulation protein (R1) being sufficiently insensitive to changes in the level of the effector protein (R2) [i.e. small sensitivities $S\left(C_{R1};C_{R2}\right)=\partial\ln\left(C_{R1}\right)/\partial\ln\left(C_{R2}\right)$, see Figure 2-35], or to other species from the GRC;

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

- GRC system *regulatory efficiency* in terms of ensuring an appropriate steady-state stability in response to dynamic perturbations in internal or external species [i.e. small sensitivities of the species stationary concentrations *vs.* perturbations in a certain species, $S\left(C_{j,s};C_{perturb.}\right)$];

- GRC *species connectivity* in terms of synchronized and efficient treatment of a dynamic perturbation for recovering the species stationary concentrations [that is small STD = standard deviation of species recovering times $\tau_{rec,j}$ after a dynamic internal perturbation].

In the next examples, the checked regulatory P.I.-s are the followings (see Table 2-2, and section 2.2.3): a) the species index " j " sensitivity *vs.* a stationary environmental perturbations in the nutrient "NutP" concentration, that is $S\left(C_j;C_{NutP}\right)=\partial\ln\left(C_j\right)/\partial\ln\left(C_{NutP}\right)$. b) The recovering times of the all model species $\tau_{rec,j}$ after a dynamic internal perturbation; AVG = average of $\tau_{rec,j}$ STD = standard deviation of $\tau_{rec,j}$.

Example no. 2.

For exemplification, two GERM-s of very effective types [G(P)1], and IG(P)1; GM(P)1] (Figure 2-36) have been approached (see also the GERM library of Figure 2-12). The species stationary (index "s") growing conditions of an *E. coli* cell are given in the Table 2-8. Two linking alternatives have been considered to built-up an effective GRC.

Alternative [A]: [G1(P1)1] + [G2(P2)1] (10 individual and lumped components).

Alternative [B]: [G1(P1)1;M1(P1)1] + [G2(P2)1;M2(P2)1] (14 individual and lumped components).

For both linking alternatives, the increased GRC complexity is expected to lead to a decline in the whole network P.I.s due to an increased complexity of the system (that is an increased "ns" species number), which, in turn, will lead to an increased difficulty to synchronize the efficient response of all components *vs.* perturbations. Due to such reasons, as the GRC structure is extended, as more effective modules should be included (with cascade control and multiple effectors) to preserve the whole GRC construction desired P.I.-s. For instance, in Figure 2-32 a two-modules GRC is modelled in the simpler Alternative [A] representation (ns = 10 species) comparatively to the Alternative [B](ns = 14 species). In spite of an increased complexity (higher "ns"), the use of a more effective regulatory schem [G(P)1; M(P)1] with a cascade control of the expression leads to better regulatory P.I.-s. Indeed, by comparing the two assembly alternatives [A] *vs.*- [B] in Figure 2-32, and Figure 2-33 , the species recovery times ($\tau_{rec,j}$) are lower for all species (except for P1), while the species interconnectivity index STD τ (rec,j) is better (i.e. a low value), while the species stationary concentrations "resistance" (that is, $S\left(C_j;C_{NutP}\right)=\partial\ln\left(C_j\right)/\partial\ln\left(C_{NutP}\right)$, the species relative sensitivities *vs* external

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

perturbations in [NutP] here) are practically the same for both alternatives.

Example no. 3.

In this example, two GERM-s of effective types [G(P)2], and [G(PP)1] (Figure 2-36) have been approached (see also the GERM library of Figure 2-12). The species stationary (index "s") growing conditions of an *E. coli* cell are given in the Table 2-8. Two linking alternatives have been considered to built-up an effective GRC.

Alternative [A]: [G1(P1)2] + [G2(P2)2] (12 individual and lumped components).

Alternative [B]: [G1(P1P1)1] + [G2(P2P2)1] (12 individual and lumped components).

The two GERM-s link is called "*cooperative link*" (see the below section "2.2.4-VIII)" because each species (especially the proteins) perform distinct functions in GRC, by supporting the reactions in both modules.

The comparison of the two linking alternatives in Figure 2-37 proves that, for the roughly the same GRC complexity, and in spite of two buffering reactions in the [G(P)2] unit, instead of only one in the [G(PP)1] unit, the use of PP dimmers as TF-s in the buffering reaction units [G(PP)n] leads to better dynamic-P.I.s, i.e. lower τ_{p1}, AVG ($\tau_{re\partial j}$) and STD($\tau_{re\partial j}$). The stationary species relative sensitivities $S\left(C_j;C_{NutP}\right)=\partial\ln\left(C_j\right)/\partial\ln\left(C_{NutP}\right)$ are practically the same, while the module complexity is the same). The low-concentrations (ca. 0.1 – 4nM) of the oligomeric effectors (of type PP, PPPP, etc.) are determined not by a quasi-steady-state assumption [Maria, 2019; Maria and Luta, 2013] but from optimising the global properties of the overall modular GRC regulatory chain (that is minimum τ_{p1} here, [Maria and Scoban, 2017, 2018].

Following this positive result, Maria and Luta [2013] used a modular construction of a GRC-operon, by linking individual GERMs but only of [G(PP)1] type.

2.2.4.-VIIC. Preliminary conclusions on linking GERM-s to buildup a GRC

Modelling the dynamics of the gene regulatory circuits (GRC) for the *in-silico* GRC design [Maria and Luta, 2013; Maria, 2007,2009,2014b,2018] is an important step in advancing the understanding on the regulatory cell networks, with important theoretical and practical implications [Bower and Bolouri, 2001] . The modular WCVV approach, with accounting for both local and holistic GRCs regulatory properties, and for the experimental data from bio-omics databanks, proves that this is a very promising and effective computational approach, allowing: similarity analysis of several models (structure *vs.* predictions); model lumping analysis; GRC system characterization (QSS-multiplicity, stability, flexibility, robustness, efficiency); cell system modularisation and development of cell simulation platforms [Maria, 2021; Maria, 2014b,2018,2023; Maria and Luta, 2013].

Even if only generic GRC-s and GERM-s from *E. coli* (Figure 2-12) have been analysed in the present work, the novel WCVV whole-cell modelling framework, with explicitly considering the link between the volume-growth and the reaction rates for all species into the cell, appears to be a more promising approach to evaluate the GRC-s characteristics in a cell, by mimicking the equilibrated or perturbed growth. Such models can avoid over-estimation of some regulatory properties (i.e. responsiveness, efficiency, connectivity), accounting for the role of cellballast in smoothing internal/external perturbations, for direct or indirect perturbations of species levels (transmitted via chain reactions and cellvolume variation).

Lumping rules are proved to be effective tools for modelling the cell regulatory process complexity and dynamics, coping with the cellsystem low observability, identifiability and estimability [Maria, 2004, 2005b, 2019]. Power-law or Hill-type representation of modular GRC-s, including apparent rate constants, can reproduce a wide-range of cell functions and dynamic behaviour [Maria, 2007, 2009, 2014b, 2018, 2023]. However, the model predictability is strongly dependent on the lumping degree, on the key-species selection and ability to realise the suitable trade-off between model simplicity, its predictive power and physical-meaning of terms. A sensitivity analysis applied to model terms can help in relating the GRC holistic properties to the individual regulatory module structure [Maria, 2014b].

Derivation of relations between apparent and extended model structures is of high interest for a process correct interpretation, and to characterise the relative importance of various reaction steps [Maria, 2004, 2005b, 2019]. The apparent steps tend to compensate the loss in system diversity introduced by the lumping rule, and cannot fully describe the real interactions among reaction intermediates. Because the apparent parameters (identified from experimental data) present values smaller or larger than those corresponding to the elementary reaction steps, the physical meaning of reactions / species lumps can also play an important role in choosing the most suitable lumping route from the large number offered by the theoretical analysis. [Maria, 2004, 2005b, 2018, 2019, 2023].

When large cell dynamic models are developed, application of unconventional model reduction strategies are recommended, by combining suitable system modularisation (in functional sub-units) with application of the sensitivity analysis to relate metabolic

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

network holistic properties (hierarchic organization and regulatory efficiency) to the individual module properties. [Maria, 2004, 2005b, 2018, 2019, 2023; Maria and Scoban, 2017, 2018].

In fact, more than these two examples of linking GERM-s to construct a GRC-s can be imagined and tried. But, as general rules, proven in this section some constraints have to be imposed, as followings:

a). Linking reactions between modules must be set slower comparatively to the module core reactions to not disturb the modules individual regulatory properties.

b). Use cooperative and mutual catalysis. Individualized functions must be allocated to each component into the cell. The un-cooperative coupling of GERM-s will lead to the cell to loss its stability and to not survive (see below paragraph "2.2.4-VIII"). That is because the homeostasis condition will be lost, and the whole system will evolve towards disappearance of the proteins presenting the lowest synthesis rate, and the whole cell system will collapse. See the proof of Maria [2005].

c). The intermediate species levels (like [PP] in the [G(PP)n] units) and the allosteric regulation loops (that is the number of the multiple buffering reactions, such as $G + P \Leftrightarrow GP + P \Leftrightarrow GPP$, in the [G(P)n] units) must be adjusted according to the GRC size, and by obtaining optimal P.I.-s of the GRC (see the two examples of this section and [Maria, 2017a; Maria and Scoban, 2018]. Addition of too complex and in a too large number of such GERM-s could growth the GRC complexity. Even if an improvement in the P.I.-s is obtained, the resulted GRC model becomes too difficult to be identified (to estimate its rate constants) from incomplete and often un-structured kinetic data (species concentrations trajectories in time, and their stationary levels), and could raise some questions regarding the time required for all these reactions to occur in a very short time compared to the main reactions (Figure 1-2). Solving several applications, indicated that for the GERM-s modules of type [G(P)n], or [G(P)n; M(P)n]], or [G(PP)n], (Figure 2-12), a reasonable n = 2 - 3 can be adopted, similarly to the allosteric enzymatic reactions [Maria, 2014b; Hammes and Wu, 1974].

d). The results of the previous section "2.2.4-VI" indicated that the novel WCVV isotonic modelling framework must be considered for a more realistic representation of the secondary connectivity effects (cell 'ballast', and 'inertial effect').

e). The results obtained in the second example of this section (2.2.4-VII) indicated that one important rule is to try to use effectors in dimmer form (e.g. PP instead P in the GERM regulatory units of type [G(PP)n]) in order to adjust the gene activity (Figure 2-12, and Figure 2-37). This type of transcriptional control with multiple operators binding repressor dimers in [G(PjPj)n] units is highly effective, the quickly adjustable dimer levels [PjPj]s via fast $Pj + Pj \underset{k_{diss}}{\overset{k_{bind}}{\rightleftarrows}} PjPj$ reactions will confer more flexibility to the gene expression regulatory module in ranging the stability, the dynamic characteristics, and the GRC flexibility *vs.* environmental changes. Experimental evidences of such a mechanism were presented by [Yang et al., 2003].
f). Maria [2017a,2017b,2018] proved that when the expressed protein **P** is acting as a **TF**, in its **GERM**, the module regulatory efficiency of [G(P)n] types is better if the TF is used in a dimeric form (**PP**), that is in GERM-s of [G(PP)n] type in (Figure 2-12).

Besides, Maria [2017a, 2017b, 2018] also proved that the GERM regulatory efficiency is even much better if **TF = PP** is acting at both **G** (DNA) and **M** (mRNA) levels of the expression, that is transcription and translation (Figure 2-7, and the middle and down-rows of Figure 2-12), thus developing a cascade control scheme of the expression where transcription and translation regulatory steps are separately considered, that is GERM-s of [G(PP)n;M(PP)n'] type. See Eq.(19), and the text below Eq.(20) for notations and the nomenclature used to describe the generic gene expression G/P regulatory modules GERMs.

2.2.4. VIII) Cooperative *vs.* concurrent linking of GERM-s in GRC-s and species interconnectivity

When coupling two or more GERM-s modules into the same cell, the nutrients, and metabolites in the G/P syntheses are roughly the same for all genes (denoted by the lumps NutG, MetG), and all proteins (denoted by the lumps NutG, MetG) (Figure 2-24, Figure 2-19), and Figure 2-12). The modelling problem is what alternative should be chosen when linking two (as discussed here) or several GERM-s, once the substrates are similar? (see Figure 2-39)? A competitive scheme (due to the common substrate, i.e. MetG1,2 and MetP1,2), or a cooperative scheme, the two GERM-s assisting each other? For exemplification, one considers the problem of adequate and efficient linking of two GERM-s, related to the expression of G1/P1 and G2/P2 pairs. By using simple [G(P)0], or [G(P)1] modules for the tested GERM-s, there are tested three alternatives of coupling them illustrated in the Figure 2-39, that is:

Alternative A: Competitive expression linking (competition on using the common metabolites) of the two coupled GERM-s modules of type [G1(P1)0] + [G2(P2)0]. The expressed proteins P1 and P2 act as metabolase, and polymerase, only in their own expression module only.

Alternative B: Simple cooperative expression linking [G1(P1)0] + [G2(P2)0] modules for the two GERM-s. Protein P1 is permease and metabolase for both GERM-s. Protein P2 is polymerase for replication of both G1 and G2 genes.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Alternative C: Complex cooperative expression. This case is identical to the alternative [B], but one effector is added to each module. That is, one buffer reversible regulatory reaction G1 + P1 ⇔ G1P1, and G2 + P2 ⇔ G2P2) are added to the [G1(P1)0] and [G2(P2)0] modules, respectively, in order to dynamically quickly adjust the G1, and G2 catalytic activity in the two GERM-s modules, thus becoming of a new type [G1(P1)1] + [G2(P2)1].

Tests were performed by using the novel WCVV modelling framework, and the nominal high-ballast cell condition of Table 2-4. (with [P2]s = 100 nM). Simulations lead to very interesting conclusions pointed-out by Maria [2005], as followings:

i). In the **Alternative A**, one links two modules [G1(P1)0] + [G2(P2)0], of the simplest type (Figure 2-12), without any effector. The way through which they were linked, ensures the regulation of the two proteins (P1, P2) synthesis, even if of a lowest efficiency compared to other GERM structures, but in a *concurential* disconnected way (Figure 2-39). For this hypothetic system, the used linking "alternative A" ensures the synthesis of P1/G1 and P2/G2 pairs from their metabolites {MetP and MetG}, but with any interference between modules, that is the expressed proteins (P1, P2) do not act but only in their own expression module. The simulated case study in Figure 2-40 corresponds to the following steady-state:

$[P1]s = 1000nM, [P2]s = 100nM, [G1]s = 1nM, [G2]s = 1nM$. For the "alternative C" case, $[G]s = [GP]s = 0.5nM$. The only connection between the two GERM-s is due to the common cell volume to which both protein syntheses contribute, Eq.(13). If one checks the system stability, by applying a ± 10% impulse perturbation in [P1]s, it results that this [G1(P1)0] + [G2(P2)0] link is an *unstable system.* That is because the maximum real part of the eigenvalues of the Jacobian matrix of the kinetic model Eq.(4) is positive, that is:

$\mathrm{Max}(\mathrm{Re}(\lambda \mathbf{i})) = 4.2839e-003$, where λi are the eigenvalues of the Jacobian matrix with the elements $J_{i,k} = \left(\partial h_i(\boldsymbol{C}, \boldsymbol{k}) / \partial C_k\right)_S$, referring to the kinetic model Eq. (11). (VIII-1)

Due to such an intrinsic instability, the cell GRC system (alternative A of linking) evolves toward the decline and *disappearance of one of the proteins* (that is, of P1 in Figure 2-40, because it presents the lowest synthesis rate). Consequently, the homeostasis condition is not fulfilled, the cell functions cannot be maintained, and the disconnected protein synthesis "*Alternative A*" results as being an *unfeasible* and *implausible* GERMs linking alternative [Maria, 2005].

ii). In the *Alternative B*, the *simple cooperative* linking of [G1(P1)0] + [G2(P2)0] modules in Figure 2-39, and Figure 2-40, ensures specific individual functions of each protein, i.e. P 1 lumps both the permeases and metabolases functions, while P2 is a polymerase. This GERM linking alternative is a viable one, the system being stable, the key species of both linked [G1(P1)0] + [G2(P2)0] modules recovering after a dynamic impulse-like perturbation of -10%[P1]s at an arbitrary time t=0 Figure 2-40. Such an intrinsic stability is proved by displaying the key species recovering trajectories after a dynamic "impulse like" perturbation of -10%[P1] at an arbitrary time t = 0, for the considered linking alternative "B" of the two GERM-s. However, the resulted GRC regulatory efficiency is low, as long as the recovering times are quite long, that is $\tau_{rec,P1} = 318.1\,\mathrm{min.}, \tau_{rec,P2} =$ 502.2 min ., and $\tau_{rec,G1} = 502.2\,\mathrm{min}$., $\tau_{rec,G2} = 502.2\,\mathrm{min}$.

iii). To fix the GRC regulatory low efficiency of "alternative B", in the *Alternative C*, the *simple cooperative* linking of the two modules [G1(P1)0] + [G2(P2)0] used in the *Alternative B* has been *improved* by adding simple effectors for genes G1,G2 activity control. Thus, the buffering reaction G1+P1<==> G1P1 was added to control the G1 catalytic activity, while the buffering reaction G2+P2<==> G2P2 was added to control the G2 catalytic activity. Thus, the cooperatively linked system, becomes of the form of [G1(P1)1] + [G2(P2)1], of Figure 2-39, and Figure 2-40, where the effectors P1 and P2 act in both GERM-s units, while the stationary states $[G1]s = [G1P1]s = 1/2nM$, and [G2]s $= [G2P2]s = 1/2nM$ ensure the maximum dynamic P.I.-s. As proved by Maria [2005], this small improvement of the two GERM-s regulatory structures reported a significant improvement in the GRC regulatory efficiency. Thus, in the "alternative C" the key species recovery time are much shorter compared to the "alternative B", after applying the same dynamic "impulse like" perturbation of -10%[P1]s at an arbitrary time t = 0, that is $\tau_{rec,P1} = 88.0\,\mathrm{min}$., $\tau_{rec,p2}$ =319.4 min ., and $\tau_{rec,G1}$ = 322.1 min., $\tau_{rec,G2}$ = 288.8 min. [Maria,2005].

The same rule of linking GERM-s can continue in the same way, for instance, also involving [G(PP)n] modules, where the effectors are the dimers PP, acting in " n " buffering reactions of the type, G + PP ⇔ GPP ⇔...⇔ GPn, with the stationary states [G]s = [GP]s = [GPP]s = ...= [GPn]s = 1/(n + 1) nM. The model rate constants should be estimated from the species stationary concentration vector C s, Eq.(14), and by imposing regulatory optimal characteristics discussed by Maria [2017a], and in the present work (see section 2.2.1.3). From the same reasons, stationary levels of active and inactive forms of catalyst should be adopted, [L]s = [TF1]s = [TF2]s = ...= [TFn]s = 1/(n+1) (see Eq.(18,19)). Besides, the dissociation constant of the (L:TFn) complex in the buffering reactions k(diss) >> Dm = ln(2)/tc, has been adopted, e.g. (1e+5 to 1e+7)Dm, being much higher than other rate constants of the GERM [Maria, 2005], (Eq. (31)). In subsequent works, Maria [2002, 2005, 2006, 2007, 2009, 2014b, 2010, 2009b, 2014a, 2017a, 2017b, 2018]; [Maria and Luta, 2013] also proved that optimization of the GERM regulatory performance indices P.I.-s with the multi-objective criteria summarized by Maria [2005] leads to small values (ca. 1-4 nM) for the intermediate [PP]s (the active parts of dimeric **TF**) in the [G(PP)n] regulatory units (Figure 2-12) .

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The stability and the dynamic regulatory characteristics of all three GRC systems above described (that is the alternatives A-B-C) have been determined by studying the stationary (homeostasis, QSS) recover after a ± 10%[P1] s "impulse-like" perturbation at an arbitrary time t = 0 (Figure 2-40). The results, presented in Figure 2-40 [Maria, 2005], reveal the following general aspects concerning the above three alternatives A-B-C to link simple GERM-s to obtain a viable and effective GRC system:

a). The GRC-s generated by the alternatives B and C are stable [that is Max(Re(λ(j)))= -Dm < 0, where λ(j) are the eigenvalues of the ODE model Jacobian matrix). The GRC system generated by the alternative A is unstable, and not-viable, one key-protein (P1) shortly disappearing after any small dynamic perturbation. Such a behavior of alternative **A** is due to the competitive linking (on the same metabolites) of the two GERM regulatory units, and the lack of any cooperative functions of the P1/P2 proteins. Systems B and C are viable GERM linking alternatives. The GRC system recovers after a dynamic perturbation in [P1] s. It results that the *cooperative* module linking is superior to the *competitive* alternative **A**, being only one *viable* solution that ensures the system homeostasis. The **B** and **C** alternatives are superior because they preserve specific functions of each protein inside the cell. Besides, the GRC linking alternative **C** presents the best P.I.-s from all three checked linking alternatives due to the additional regulatory effector of type of a rapid buffer reaction G + P <==> GP.

b). The system is as better regulated as the effector is more effective, that is the use of multiple buffering reactions, with dimeric TF-s (such as PP [Maria and Luta, 2013], or a cascade control of the expression [G(P)n; M(P)m] (not presented here, see Figure 2-12, [Maria,2017a]).

c). The use of efficient effectors and multiple regulation units can improve very much the dynamic P.I.-s (section 2.2.3), in the following ranking: [G(P)n] < [G(PP)n] < [G(P)n;M(P)n1]. Maria, 2003, 2017a, 2018]. (see also Fig 2-13, and Figure 2-15, and Figure 2-38).

d). Dynamic perturbations affect rather species present in small amounts inside the cell, while recovering times for the other species (e.g. large metabolites MetP, MetG) are negligible (see section "2.2.4VI").

By following a similar rule, and taking into account the GERM-s P.I.-s properties underlined in the section 2.2.3, and the above 2.2.4-I-VIII linking rules, the design procedure of the gene regulatory circuits GRC-s can be continued, by accounting for the expression of new proteins (and their corresponding GERM-s). For instance, in the simplified representation of Maria [2005], a 3-rd GERM for P3 synthesis can be added by adopting the above **Alternative C**, linking rule, by allocating specific functions to the P1 , P2, P3 proteins, as follows: P1 and P3 lumps permease and metabolase enzymes, which ensure nutrient import inside the cell, and their transformation in gene-metabolites (MetG1-MetG3) and proteinmetabolites (MetP1-MetP3) respectively. Protein P2 lumps polymerases able to catalyze the genes G1, G2, G3 production. If one considers the simplest effector case, the resulted GRC includes three modules [G1(P1)1] + [G2(P2)1] + [G3(P3)1], (Figure 2-41) which effectively regulate the synthesis of P1, P2 and P3, in a *cooperative interconnected* way, which preserves the all protein functions inside the cell.

2.2.4. IX) The optimal value of TF

It was proved by Maria [2005,2017a,2018] that, in the novel WCVV modelling framework, the holistic properties of the cell and where the GERM or GRC are analysed, should be preserved and, eventually they should be adjusted via the model structure and its parameters [Maria, 2017a; Maria and Scoban, 2017,2018]. One of the cell modelling principles [defined in the section 1.3.4, and numbered with (a-g), or (I-V)] postulates that the concentration of intermediates used in the GERM-s or GRC-s should be maintained at a minimum level to not exhaust the cell resources but, at the same time, at an optimal value to maximize the regulatory performance indiced (P.I.- s) of GERM-s or GRC-s. Such optimal [TF]s are obtained by solving a multi-objective optimization problem, as suggested by Maria [2017a, 2009].

An example was provided by Maria [2009] in the case of a genetic switch (GS) in *E. coli* cell, modelled under the WCVV approach. The two considered self- and cross-repressing gene expression modules are of the following type (see the GERM library of Figure 2-12) [G2(P2P2)1(P3P3)1] + G3(P3P3)1G3(P2P2)1] (Figure 2-42). The principle of a GS is briefly illustrated in Figure 2-43 and Figure 2-44. The two self- and cross-repressed gene expression coupled modules GERM 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.

These well chosen GERM-s type models are very flexible, allowing adjusting the regulatory properties of the **GS** (i.e. switch certainty, good responsivity to inducers, good dynamic and stationary efficiency [Maria, 2009, 2014b]). Besides, based on adequate kinetic WCVV models Maria [2009] proved that it exists an optimal low level of the TF-s (that is [P2P2]s, or [P3P3]s for the present model) that are associated to the optimal holistic regulatory properties of the GRC (low sensitivity *vs.* external nutrients, but high *vs.* inducers). Besides, [Maria, 2009, 2014b] proved that these TF-s are rather dimmers than monomeric molecules, in order to maximize the resulted GS performance indices. An optimal 0.1 - 4 nM (of a very short life) was found to ensure such goals [Maria and Luta, 2013; Maria, 2009, 2014b]. These *in-silico* obtained results have been confirmed by the literature experimental data. Such GS are

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

very suitable to be used as a molecular biosensor [Maria, 2018, 2023], able to detect environmental inducers of micro-/nano- molar concentrations, with application in airports, civil / military industry, etc.

2.2.4, X) Additional aspects to be considered when linking GERM-s

Cell GERM-s, and GRC-s and, in particular, those involved in some protein synthesis regulation, are poorly understood. The deterministic/stochastic modular approach of studying the regulation pathway, accounting for its structural and functional organization, seems to be the most promising routes to be followed. Because a limited number of GERM-s types exist (see the library of Figure 2-12 proposed by Maria [2005, 2006, 2017a, 2017b, 2018]), individual GERM-s can be separately analysed, as above checked for efficiency (section 2.2.3) in conditions that mimic the stationary and perturbed cell growing conditions. Efficient GERM-s (with the regulatory indices P.I.-s defined in Table 2-2), and studied in section 2.2.3) are then linked accordingly to certain rules (this section, paragraphs 2.2.4 I-IX), in order to mimic the real metabolic GRC processes, by ensuring the overall GRC efficiency, cell system homeostasis, and individual functions for all the proteins. Module linking rules are not fully established, but some principles reviewed in this section, (paragraphs 2.2.4 I-IX) [Maria, 2017] must be respected. The hierarchically organised network (Figure 2-8) includes a large number of compounds and GERM-s units with strong interactions inside a module and weaker interactions among modules. Consequently, when developing GRC - WCVV kinetic models, the whole GRC system efficiency must be adjusted by changing the type of the component GERM-s, and some linking parameters. By testing several ways to link GERM-s, Maria [2005, 2017a, 2018] advanced some rules, derived from their separate analysis in this section (paragraphs 2.2.4-I-...-IX), as followings:

i). The linking reactions between GERM-s are set to be relatively slow comparatively with the module core reactions. In such a manner, individual modules remain fully regulated, while the assembly efficiency is adjusted by means of linking reactions, and the intermediate species (TF-s) small concentrations. To preserve the individual regulatory capacity, the magnitude of linking reactions would have to decline as the number of linked modules increases. See the example provided by Maria and Luta [2013], described in section 3.2.

ii). When linking GERM-s, the main questions arise on the connectivity mechanism and on the cooperative *vs.* uncooperative way by which proteins interact over the parallel/consecutive metabolic pathway [Atkinson et al.,2003; Wall et al., 2003; Hlavacek and Savageau, 1997; Maria, 2003, 2005]. In spite of an apparent 'competition' for nutrient consumption, protein synthesis is a closely 'cooperative' process, due to the specific role and function of each protein inside the cell [see the proof in the section 2.2.4.-VIII]. In a *cooperative linking*, common species (or reactions) are used for a cross-control (or cross-catalysis) of the synthesis reactions. Thus, the system stability is strengthened, while species inter-connectivity is increased leading to a better treatment of perturbations.

iii). Protein interactions are very complex, being part of the cell metabolism and distributed over the gene regulatory network nodes. There are many GRC nodes with few connections among proteins and a small, but still significant, number of nodes with many proteic interactions. These highly connected nodes tend to be essential to an organism and to evolve relatively slowly. At a higher level, protein interactions can be organized in 'functional modules', which reflect sets of highly interconnected proteins ensuring certain cell functions. Specific proteins are involved in nutrient permeation (permeases), in metabolite synthesis (metabolases), or in gene production (polymerases). In general, experimental techniques can point-out molecular functions of a large number of proteins, and can identify functional partners over the metabolic pathways. Moreover, protein associations can ensure supplementary cell functions. For instance, enzyme associations (like dimeric or tetrameric TF-s) lead to the wellknown 'metabolic channelling' (or tunnelling) process, that ensures an efficient intermediate transfer and metabolites consecutive transformations without any release into the cell cytosol (similar to a "bulk phase" of a fed-batch bioreactor) [Maria, 2005].

iv). It results that an effective module linking strategy has to ensure the cell-functions of individual proteins and of protein associations over the metabolic synthesis network. As a general observation, even if not all GERM-s include common reactions, the GERM modules are anyway linked through the cell volume [to which all cell species contribute, Eq.(11-12), in an isotonic cell system], and due to some intermediates controlling the module interactions in the GRC-s (see Figure 2-43 as an example). It is only the WCVV modelling framework which is able to account for the complex cell regulatory characteristics, as briefly discussed in this section. 2.2.4.

v). A natural strategy for building-up complex and realistic cell models is to analyse independent functional modules (GERM-s, or belonging to the central carbon metabolism (CCM) or groups of closely interacting cellular components, and then link them [Maria, 2017a, 2018, 2023; Maria and Luta, 2013; Maria, 2021]. The WCVV approach may facilitate this strategy. Each GERM module could be modelled as a separate "entity" growing at the actual rate of the target cell. The volume of the newborn cell and the environment characteristics could match those of the target (see the WCVV hypotheses of Table 2-1). To allow this, and to reproduce the cell ballast effect, lumped molecular species could be defined into each cell where a GERM is analyzed (tested), in amounts equal to those of the target cell minus those due to the components of the analysed GERM module. Thus, each tested cell carrying a certain defined GERM-s would grow at the same observed rate. As a result, linking GERM-s would be a seamless process requiring only that the ballast level to be kept at its experimental level.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

vi). The WCVV modelling approach demonstrates in this section 2.2.4 that each cell component affects, and is affected by, all other cellular components. Indirect interconnectivities arise because all components in a cell contribute to cell volume Eq.(11-13), and cell volume influences component concentrations. Thus, perturbations in one component reverberate throughout the cell. The importance of these indirect relationships will vary with the diversity and complexity of cellular components. Increasing numbers add ballast to the cell, minimizing these indirect relationships (through the cell volume), while the increasing diversity allows individual metabolites to be present at lower concentrations, thus improving the dynamic responses of GERM-s or GRC-s to perturbations. Another issue, thus far unexamined, is how specific types of interconnectivities affect the regulatory behaviour of cells. This could be probed by using the experimental methods developed by Vance et al. [2002] to deduce connectivities in biochemical pathways from the effects of impulse perturbations in a substrate or precursor.

vii). When modelling complex operon structures, simple GERM-s structures should be adopted to not complicate too much the WCVV model. The default GERM is the [G1(P1)1]. (Figure 2-24) But, according to the experimental data and interactions among genes and proteins, more complicated GERM-s constructions can be chosen in building-up such GRC models [Maria, 2017a, 2018, 2014b]. For instance, Maria and Luta [2013], used more effective GERM-s of [G(PP)1] type (Figure 2-19) to develop an adequate WCVV kinetic model for the mercury-GRC-operon in *E. coli*.

2.2.4 XI) The effect of cascade control on the GERM efficiency

Among the GERM-s reviewed (see the library of Figure 2-12) and tested by Maria [2002, 2003, 2005, 2006, 2007, 2008, 2009, 2017a, 2017b, 2018] by using the WCVV modelling framework, the most significant are those of [G(P)n] units, to whom the regulatory effectiveness is increases nearly linearly with the number ("n") of the successive buffering reactions (Figure 2-9, and Figure 2-13). Due to their simple structure such GERM-s are very suitable to construct complex GRC-s. On the next place, the [G(PP)n]] kinetic models are extremely favourite, by presenting a more efficient regulatory efficiency due to the used dimeric TF-s (that is PP) in the adjusting the 'catalyst' G activity through successive buffering reactions of the type: G+PP <===> GPP. The most effective GERM-s models are those with a cascade control of the expression, by means of buffering reactions applied at both gene G (DNA) level, and M (mRNA) level, that is of type [G(P)n ;M(P)n1] (Figure 2-12), and Figure 2-16 , and Figure 2-33. On the other hand, Maria [2003, 2005, 2006, 2017a] *in-silico* proved the superiority of the [G(P)n ; M(P)n1] gene expression structures (see also the above mentioned Figs. 2-16, 2-33. The conclusions are the followings:

i). The very rapid buffering reactions, such as G+P<===>GP+P<===> GPP .....<===> GPn, in the [G(P)n] regulatory units, or those of the form:

ii). M + P <===> MP+ P <===> MPP .....<===> MPn1, in the [G(P)n ; M(P)n1] regulatory units, have been proved to be very effective regulatory elements, by quickly adjusting the active/inactive G / (GP/GPP/GPn), or M / (MP/MPP/MPn1) ratios, thus efficiently coping with the dynamic or stationary perturbations.

iii). The *in-silico* (numerical, math model based) tests revealed that the P.I.-s of the GERM-s, more frequently used to build-up complex GRC-s, increase in the approximate order:

[G(P)0] (0 regulatory "effector") <

[G(P)1] (1 regulatory "effector") <

[G(P)1; M(P)1] (2 regulatory "effectors") <

[G(P)n ; M(P)m] (n + m regulatory "effectors") <

[G(PP)n] (n + 1 regulatory "effectors") <...., etc.

As a conclusion, the gene expression regulatory units of [G(P)n ; M(P)m]] types are one of the most efficient ones (see the library of Figure 2-12), that is because it imitates best the gene expression (that is the transcription and translation steps in Figure 2-7)

Roughly, the improvement of the GERM unit P.I.-s per regulatory "effector" is of ca. 1.3x under the WCVV modelling framework, while the same improvement is of ca. 2 x under the classic WCCV modelling framework, but using more complex model structures and constraints [Maria, 2005, 2006, 2007, 2009, 2014b, 2017a]. It clearly appears that the WCVV modelling framework is more realistic by not imposing but only the isotonicity constraint (Table 2-1), while the default WCCV approach is unable to correctly estimate the P.I.-s of GERM-s (section 2.3.1) [Maria et al., 2018d].

2.2.4- XII) Building-up GERM-s regulatory models under WCVV

As reviewed by Sewell et al. [2002], the regulatory networks, implying a large number of proteins, are poorly understood. They appear to be organised hierarchically, consisting either by a large number of strongly interacting components or of smaller numbers

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

of weakly interacting groups of reactions or modules. The latter alternative seems to be more computationally tractable; each regulatory module can be analysed individually (see P.I.-s of section 2.2.3), and then combined in a functional organized hierarchy following some rules (see the paragraphs I-XI, of this section 2.2.4).

As revealed by this section 2.2.4, the use of simple GERM modules presents some advantages. First, it is possible to easier estimate the model rate constants, and to *in-silico* study the module regulatory properties (P.I.-s of section 2.2.3).

Besides, in a variable cell-volume WCVV approach, for each module it would be possible to estimate the relatively small number of rate constants under the same assumption of average cell-growth rate. Then, a reduced number of linked modules are considered in order to estimate the linkage slow reaction's constants. In such a manner it would be possible to adjust some of the inner 'raw-materials' -levels for each of the protein module in order to match the linkage requirements of optimal individual protein regulation. By applying the same rule with small groups of linked modules it would be possible to step-by-step reconstruct the whole chain of linked optimised GERM modules which can include the cell key-proteins. The advantage of using the WCVV models for every separate GERM unit is the possibility to include the indirect interaction due to the same cell-volume space, Eq.(12-13). The cell volume unique increasing rate will influence all the modules with weights depending on the each module contribution (as number of moles) to the cell entity. Such a modelling of GRC-s in-chain-of-modules will allow considering interactions among modules with no direct linkage reaction via their separate volume contribution. At the same time, it would be easier to simulate the inertial effect of smoothing external perturbations when a large sum of copynumbers is present into the cell (high "cell ballast" case, section 2.2.4-VI).

To simply construct and compare various GERM-s schemes (see their "library" of Maria [2005, 2017a] in Figure 2-12) to regulate the expression of a generic protein "P1" into the cell, under stationary and dynamic perturbation and variable cell-volume conditions (WCVV modelling framework), if any a-priori information is available, a limited number of reactions and species are considered individually, or lumped, that is (Figure 2-24): the key protein "P1", the gene "G1" encoding the protein (i.e. the biocatalyst), external/ internal 'raw-materials' for autocatalytically producing "G1" and "P1" (denoted with "NutP, NutG / MetP, MetG respectively), catalytic inactive intermediates (such as "GP", protein oligomers such as "PP", and by-products like M' in Figure 2-38). These schemes are step-by-step built-up starting from the simplest possible one, and adding reactions according to a certain principle and desired function (see the rules of sections 2.2.4 I-...-XI). The same rules can be continued to develop even more complex / effective regulatory schemes but a trade-off between model complexity/identifiability and their effectiveness should be realised.

For each GERM proposed kinetic schema, an ODE kinetic model is associated under the WCVV framework Eq.(11-15). To be of use, the rate constants of this kinetic model have to be estimated (see section 2.2.1.3). The classical way to estimate the parameters of an ODE kinetic model is based on the product/intermediate analysis, prior information, and (filtered, reconcilied) experimental kinetic curves (e.g. dynamically measured species concentrations versus time) [Maria,2004]. The kinetic parameters are fitted based on a estimation criterion that minimise the residual differences between experimental data and the model predictions in terms of output variables (that is the above key-species in the GERM model case). Due to the noised data and complete/incomplete experiments, sometimes several estimates can be obtained under similar operating conditions. Estimate discrimination is based on its physical significance, and model predictions interpretation. The estimation objective function is linked with the statistical methods because the observed data are always subjected to experimental errors, and several physico-chemical constraints are imposed to the parameters (see review of Maria [2004, 2008b]; [Iordache and Maria, 1991; Maria et al., 2016]).

For the developed regulation models at a cell level, little standard information in the classical sense is available (structured kinetic curves of cell species after a perturbation; see the example of Chassagnole et al. [2002]; Niklas et al. [2011]). The construction of kinetic cell-modules accounts rather disparate qualitative/quantitative information from databanks, being linked to describe the cell functions. By imposing QSSlevels for the key-proteins of the model (known from '-omics' databanks), and by solving the QSS-mass conservation nonlinear set of equations Eq.(11-15) with using common numerical algorithms [Maria, 2004; Iordache and Maria, 1991] most of the model rate constants are obtained. The rest of rate constants can be derived by optimizing a certain regulatory estimation criterion [Maria and Scoban, 2017, 2018], such as (Figure 2-9, Figure 2-13, Figure 2-32, Figure 2-20): fast recover after a dynamic perturbation; smallest amplitude of a recovering path, smallest sensitivity of the stationary concentrations *vs.* external (nutrients) perturbations, etc. In the present study, the criterion of minimum recovering time after an imposed 'impulse' perturbation have been adopted to estimate the model parameters, together with some physical meaning assumptions.

The way in which these separate regulatory modules have been approached in this study can suggest a method of linking GERM regulatory modules when design a complex networks (GRC). If two protein-synthesis regulatory modules are linked, the proteins interact in a common reaction (for instance Figure 2-39 and Figure 2-40, cases BC, section 2.2.4-VIII), or as a catalyst, thus producing a reciprocal perturbation. As long as the linking reactions are slow relative to the core rates of each module, the two proteins (P1,P2) in the linked system would remain regulated. Additional modules could be linked under similar restrictions (Figure 2-41) by forming a network with rapid intra-module reactions and slow inter-module linkages. To assure that individual perturbations would not combine to yield an overall perturbation that exceeds the regulatory capacity of any individual module, the magnitude of each linking interaction would have to decline as the number of linkage reactions increased. The example of linking

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

three modules [G1(P1)1] + [G2(P2)1] + [G3(P3)1], displayed in (Figure 2-41) which effectively regulate the synthesis of P1,P2 and P 3 , in a *cooperative interconnected* way, which preserves the all protein functions inside the cell (section 2.2.4- VIII) eloquently illustrates this linking rule of GERM-s.

2.2.4. XIII). Limitations of the GERM analysis using the WCVV approach

The hypothetical cells analyzed here are several orders of magnitude simpler than even the simplest living cell, and of course the number of data available against which our analysis could be tested. Moreover, the number of possible sets of conditions and cells is infinitely large, rendering a comprehensive study futile. This is why, I restricted the analysis of sections 2.2-2.3 about a single nominal condition that reasonably represented, albeit in a symbolic manner, conditions in real cells. Strictly viewed, any conclusions drawn from the present analysis are applicable only to the hypothetical systems examined and only at those conditions examined. Whether any of this conforms to reality is debatable. On the other hand, these imaginary cells allow extensive analysis of holistic/ emergent properties that may ultimately have some utility and may be required for modelling more realistic and experimentally grounded cells.

2.3. The advantages of using the WCVV modelling framework when modelling the cell GRC-s.

2.3.1. Proving superiority of using the WCVV novel modelling framework compared to the classic ("default") WCCV kinetic models when simulating gene expression regulatory dynamics:

To prove, in a simple way, the superiority of WCVV modelling framework *vs.* the classic ("default") WCCV modelling rules (section 2.1), when simulate the cell key-species dynamics in GERM / GRC , one presents here an example of how wrong predictions can offer a WCCV model when simulating a gene expression regulatory module of simplest [G(P)1] type.

To exemplify how essential conceptual differences are between the novel WCVV modelling framework (section 2.2) and the classic ("default") WCCV models, and how huge are the differences in the simulated results, one considers here the simplest regulated system of one generic gene (G) expression regulatory module used for the synthesis of its encoding protein P . In this respect, one adopts the reaction scheme of Maria [2005] corresponding to the [G(P)1] type module (according to the Figure 2-12). We represented the simplified reaction schema of this lumped pathway, in a simplified manner in (Figure 2-24).

To keep enough generality, the demonstration will concern a generic G/P module, of a simple [G(P)1] type (Figure 2-24), with species concentrations characteristic to the *E. coli* cell [Maria, 2005]. The reaction rate expressions, in a simplified lumped form are the followings:

$$\begin{cases} NutG + P \xrightarrow{k_1} MetG + P\,,\ r_1 = k_1\, C_{NutG}\, C_P \\ NutP + P \xrightarrow{k_2} MetP + P\,,\ r_2 = k_2\, C_{NutP}\, C_P \\ MetG + P \xrightarrow{k_3} G + P\,,\ \ r_3 = k_3\, C_{MetG}\, C_P \\ MetP + G \xrightarrow{k_4} P + G\,,\ \ r_4 = k_4\, C_{MetP}\, C_G \\ P + G \xrightarrow{k_5} GP\,,\ \ r_5 = k_5\, C_P\, C_G \\ GP \xrightarrow{k_6} G + P\,,\ \ r_6 = k_6\, C_{GP} \end{cases} \tag{32}$$

The attached GERM kinetic model written by using the classic ("default") formulation WCCV ("constant volume like", section 2.1), in terms of *Nj* (amount of species *j* in n -moles), and *Cj* (species j concentration, in nM) is the following:

WCCV classic ("default") model [*]

$$\begin{aligned} \frac{1}{V}\frac{dN_{MetG}}{dt} &= r_1 - r_3 \\ \frac{1}{V}\frac{dN_{MetP}}{dt} &= r_2 - r_4 \\ \frac{1}{V}\frac{dN_P}{dt} &= r_4 - r_5 + r_6 \\ \frac{1}{V}\frac{dN_G}{dt} &= r_3 - r_5 + r_6 \\ \frac{1}{V}\frac{dN_{GP}}{dt} &= r_5 - r_6 \end{aligned} \quad \Rightarrow \quad \begin{aligned} \frac{dC_{MetG}}{dt} &= r_1 - r_3 \\ \frac{dC_{MetP}}{dt} &= r_2 - r_4 \\ \frac{dC_P}{dt} &= r_4 - r_5 + r_6 \\ \frac{dC_G}{dt} &= r_3 - r_5 + r_6 \\ \frac{dC_{GP}}{dt} &= r_5 - r_6 \end{aligned}$$

[*] Ds = Dm = Ln(2)/ *tc* = 6.93×10^{-3} 1/min.

(33)

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

In the above classic ("default") kinetic model formulation, despite the cell volume (V) is variable with the time, by doubling after a cell cycle (tc) Eq.(12), in the wrong way the volume variable was considered constant by re-writing "(1/V)(dNj/dt) " as "dCj/dt ". This is a capital (huge) error of these classic WCCV kinetic models which, as it will be further proved, will lead to erroneous simulation results and wrong predictions [Maria et al., 2018d].

By contrast, the kinetic model of this system written in the novel WCVV modelling framework (section 2.2) is the following:

WCVV novel model [*]

$$\begin{cases} \frac{d\,C_{MetG}}{d\,t} = r_1 - r_3 - D_i \cdot C_{MetG} \\ \frac{d\,C_{MetP}}{d\,t} = r_2 - r_4 - D_i \cdot C_{MetP} \\ \frac{d\,C_P}{d\,t} = r_4 - r_5 + r_6 - D_i \cdot C_P \\ \frac{d\,C_G}{d\,t} = r_3 - r_5 + r_6 - D_i \cdot C_G \\ \frac{d\,C_{GP}}{d\,t} = r_5 - r_6 - D_i \cdot C_{GP} \end{cases} \tag{34}$$

[*] With D_i given by Eq.(35-36).

One essential difference of the WCVV model Eq.(34) 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.(10-15), the net difference is done by the introduction of the instant cell dilution rate (Di), Eq.(12) 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. (34) is given by Eq.(12,13), that is:

$$D_i = \frac{RT}{\pi}\sum_{j=1}^{ns}\frac{1}{V}\frac{dN_j}{dt} = \frac{RT}{\pi}\sum_{j=1}^{ns} r_j \tag{35}$$

Which, for the present WCVV model can be re-written as:

$D_i = \frac{RT}{\pi}\left(r_1 + r_2 - r_5 + r_6\right)$, that is, by substituting the rates expressions, it results:

$$\sum_{j=1}^{ns}\left(\frac{1}{V}\frac{dN_j}{dt}\right) = k1\times C_{NutG}\times C_P + k2\times C_{NutP}\times C_P - k5\times C_G\times C_P + k6\times C_{GP} \tag{36}$$

In the above relationships Eq.(33-36), the followings notations have been used: Cj = cell-species j concentration [nM]; V = cell volume (cytosol) [L]; Nj = amount of species j [n-moles]; rj = j-th reaction rate [n-moles/L/min]; Di = cell-content instant dilution rate [1/min]; π = osmotic pressure [units dependent on the R units]; T = temperature [K]; R = universal gas constant; ns = number of species inside the cell (taken individually, or lumped); t = time (min).

The required constant (RT/π) in Eq.(36) is evaluated from Eq.(13) from using the all known initial (stationary) values of the cell species concentrations, that is:

$$\frac{RT}{\pi} = 1/\sum_{j=1}^{ns} C_{j,o} = \text{ constant} \tag{37}$$

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 concentrations for the lumped NutP and NutG (Table 2-4). Solving the models Eq. (33-34) is made with using a stiff integrator ("ODE15s") of Matlab TM package due to the very fast buffering reactions r5-r6 in Eq. (32) compared to the rest of reactions (see the discrepancy in the WCVV model rate constants of the Table 2-6). The number of moles Nj (t) can be calculated at any time with the formula: Nj(t) = Cj(t) V(t), by using the species concentrations Cj(t) derived from solving the ODE model, and the volume obtained from Eq.(12), with the average cell dilution rate Dm = ln(2)/tc, with the cell cycle tc = 100 min (Table 2-5).

The rate constants of the WCVV kinetic model Eq. (34) have been estimated from the steady-state (homeostatic) initial condition of the cell, that is Eq.(14), which corresponds to the following condition: $(dC_j/dt)s = hjs(C_s,k) = 0$. In the present [G(P)1] model case, from Eq.(34) it results:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

$$\begin{cases} 0 = r_{1,s} - r_{3,s} - D_s \cdot C_{MetG,s} \\ 0 = r_{2,s} - r_{4,s} - D_s \cdot C_{MetP,s} \\ 0 = r_{4,s} - r_{5,s} + r_{6,s} - D_s \cdot C_{P,s} \\ 0 = r_{3,s} - r_{5,s} + r_{6,s} - D_s \cdot C_{G,s} \\ 0 = r_{5,s} - r_{6,s} - D_s \cdot C_{GP,s} \end{cases} \quad (38)$$

The analytical solution of Eq.(38), obtained by using the Maple ™ or symbolic Matlab ™ softwares, is the followings:

$$k_1 = \frac{D_s\left(C_{MetG,s} + C_{G,s} + C_{GP,s}\right)}{C_{NutG,s}C_{P,s}} \quad ; k_2 = \frac{D_s\left(C_{MetP,s} + C_{P,s} + C_{GP,s}\right)}{C_{NutP,s}C_{P,s}}$$
$$k_3 = \frac{D_s\left(C_{G,s} + C_{GP,s}\right)}{C_{MetG,s}C_{P,s}} \quad ; k_4 = \frac{D_s\left(C_{P,s} + C_{GP,s}\right)}{C_{MetP,s}C_{G,s}}; \quad (39)$$
$$k_5 = \frac{k_6 C_{G,s} + D_s C_{GP,s}}{C_{P,s}C_{G,s}}$$

By replacing in Eq. (39) the stationary concentrations of Table 2-5 (taken for a GERM from *E. coli* cell K-12 strain, [Maria, 2005]), one obtains the WCVV model rate constants.

The numerical values of the rate constants for the initial cell condition, and an average "Dm" are given in Table 2-6. For the rapid buffer reaction, $G+P \underset{k_6}{\overset{k5}{\rightleftarrows}} GP$, that is 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 Eq.(31), and the discussion of Maria [2006]. Here, Dm denotes the apparent average cell dilution rate Eq.(12), that is: Dm = Ln(2)/ t_c. As in (Table 2-5), *tc* = 100min., it results Dm = 6.93 × 10^{-3} 1/min. The value Ds = Dm was used in Eq.(39) to estimate the rate constants.

It is worth to mention that in the WCCV kinetic model formulation of Eq. (7-left), the model rate constants cannot be estimated on the same way because of singularities of the resulting nonlinear algebraic set. Consequently, the same rate constants of the WCVV model were used instead to make predictions with this classical ("default") WCVV model formulation.

Here it is to observe that in the Table 2-5, the value of the lump $\Sigma_j C_{met} G_{j,s}$, results from the isotonicity constraint (ensuring the membrane integrity) ($RT/\ \pi^{cyt}$) = ($RT/\ \pi^{env}$) which, by using Eq. (12-13), indicates that the sum of cell species concentrations must equal those of the environment, i.e. $\left(\sum_j^{all} C_j\right)_{cyt} = \left(\sum_j^{all} C_j\right)_{env}$. 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.

On the other hand (Figure 2-23, and Figure 2-24), species inside the cell transform the imported nutrients (NutG, NutP) into metabolites (MetG,MetP), which they react to make more cell components, especially to replicate the genome and the proteome. In turn, increased amounts of polymerases (lumped P here) are then used to import increasing amounts of nutrients. The net result is an exponential increase of cellular components in time, which translates, through isotonic osmolarity assumption, into an exponential increase in volume with time Eq. (12-13).

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 2-5, but with adopting $[G]_S$ =$[GP]_S$ = 1 nM, and one applies a dynamic perturbation (impulse like) at an arbitrary moment t=0 (Figure 2-27), by diminishing the stationary $[P]_S$ from 1000 nM to 900 nM (for a GERM of [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. Such very quick changes will, in turn, 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 2-5, 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 2-5), but in cloned cells with plasmids, this level can be higher. It is here worth noting that sometimes the 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) this information on the GERM response to this dynamic perturbation is transmitted from cell generation to generation. [Elowitz and Leibler,2000].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

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 2-5, but with adopting a stationary level of [G]s = [GP]s = 0.5 nM (to obtain a maximum regulatory efficiency; see [Maria,2005,2006]). By applying a dynamic perturbation to [P]s steady-state of a GERM of [G(PP)1] type from *E. coli*, that diminishes the stationary [P]s from 1000 nM to

900 nM, simulation with the identified WCVV model, Eq. (34-39), and Table 2-6, is leading to the species trajectories displayed in Figure 2-28, proving again the good self-control of the P -synthesis produced by the lumped G(P)1 GERM model. The simulation results also underlines the positive effect of a large "cell ballast" (above section "The effect of the cell ballast(load) on the GERM efficiency, section 2.2.4.-VI") in "smoothing" the effect of a dynamic perturbation. Also, it is to remark that the effect of a relatively small perturbation in one species (P) does not affect significantly the species present in large amounts (i.e. MetP and MetG lumps).

It is self-understood that, as the regulatory scheme (in Figure 2-12) 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 [2017a, 2018].

This example demonstrates, in a simple way, how the classic ("default") WCCV modelling is unable / ineffective to reproduce the regulatory properties of GERMs.

In order to better illustrate the discrepancy in the predictive capabilities between the novel WCVV and the classic WCCV kinetic model formulations, and how deceptive can the predictions of the WCCV 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 2-24. The reaction scheme, reaction rate expressions for the [G(P)1] gene expression module are those given in Eq. (32), while the species mass balances are given in Eq. (33) for WCCV, and in Eq.(34-37) for the WCVV formulation.

The rate constant of the WCVV kinetic model were estimated based on Eq.(38), and with using the *E. coli* cell data with the initial homeostatic concentrations of Table 2-5. The same rate constants were used also for the WCCV model because for this last 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 copynumbers (number of moles) plotted in Figure 2-29. It is to observe that, while the cell volume doubles, the species copynumbers double as well. However, there is a very small discrepancy in the predicted cell-volume dynamics because the average *Ds* = *Dm* used by the WCCV model is slightly different from the *Di* used by the WCVV model. However, this difference in the volume-*vs.*-time plot is too small, and the both curves, predicted by the two models (WCCV and WCVV) are practically overlapping.

For the same case study, 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 2-30. 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 WCCV model (-------), compared to ii) the variable volume WCVV novel modelling framework [- , with instant dilution Di term and isotonicity constraint included in the model, Eq.(34-37)]. 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.12) satisfactorily averages the "instant" dilution rate "Di" of Eq. (12), that is Eq.(35-36) here. In Figure 2-30 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 WCCV predictions of the "classic/default" kinetic model, with an average dilution "***Dm= ln(2)/tc"*** in the model (Eq.(33)), are quite close compared to those predicted by the variable volume novel isotonic WCVV model (continuous line ———), employing the cell dilution "***Di***" estimated from the isotonicity constraint, Eq. 35-36).

By making a "zoom" in the Fig. 2-30 , it results Fig. 2-30-zoom. 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. So, the WCCV predictions of the "classic/ default" 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.

However, the situation will change dramatically, and huge differences (discrepancies) 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 WCCV 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 2-31. The differences in model predictions are as larger as the species present a lower level in the cell (that is forP, G, and GP species).

Consequently, while the WCVV kinetic model correctly reproduces the system homeostasis recover after a dynamic small perturbation, species concentrations being kept quasi-constant because both nominator and denominator of the fraction Cj (*t*)= *N*j

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

(*t*) / Vcyt (*t*), [nM], are doubling at the same rate. By contrast, the classic ("default") WCCV model predictions are inaccurate, the predicted species concentration dynamics under perturbed conditions, eventually with using an average formula (Dm, Eq.12) for the cell content dilution rate are wrong, and very far from the reality. So, the WCCV models cannot be used in a satisfactory manner to simulate the regulatory properties of GERM-s or GRC-s [Maria, 2017a, 2018]. One can conclude that the novel WCVV modelling framework introduced by Maria [2002], and by [Maria et al., 2002] {while the study of the WCVV properties, features, and advantages was given by Maria over the interval [2005-2023], that is Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]}, the novel WCVV modelling framework better reflects the GERM-s regulatory properties after dynamic (impulse-like) or stationary (step-like, not discussed here) internal or external perturbations.

As revealed by the simple but eloquent example discussed in this section, there are important issues to be considered when developing modular WCVV models of GERMs, or of GRC-s including a variable number of linked GERMs.

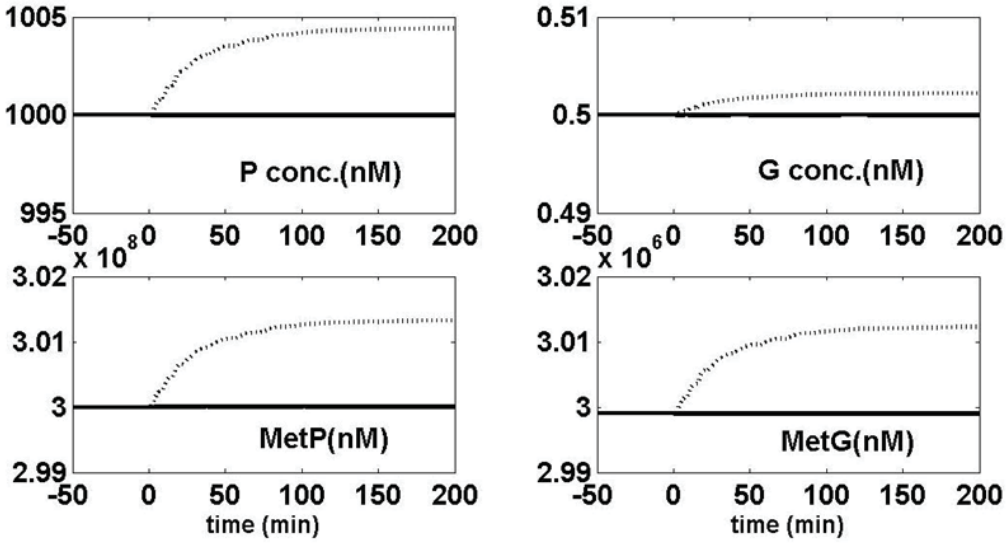


Figure 2-30-zoom: Species concentrations (in nM) dynamics during the cell cycle, under stationary (un-perturbed) conditions, predicted by the "classical" ("default") constant volume WCCV model (dash line ------- , with using an average cell dilution "*Dm= ln(2)/tc*", Eq.(33)), compared to those predicted by the variable volume novel isotonic WCVV model (continuous line ——— , with cell dilution "*Di*" estimated from the isotonicity constraint, Eq. 35-36). The generic G/P expression is using a regulatory module of [G(P)1] type. The *E. coli* cell species homeostatic concentrations are those of Table 2-5.

First, this simple case study of applying WCVV modular kinetic models to simulate GERMs 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 GRCs controlling the cell enzymes syntheses and metabolic fluxes [Maria and Luta, 2013; Maria, 2014b]. 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 [Maria, 2017a, 2017b, 2018].

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 1-2). 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 [Maria, 2006,2017b,2018].

As another observation, by contrast to the novel WCVV modelling framework, one fundamental deficiency of the classical "default" (constant-volume like) WCCV kinetic model formulation {Eq.(1a-c) in section 2.1} is the lack of the intrinsic stability of the cell system kinetic model, because these models do not include neither the Pfeffers' constraint Eq.(6,10-12), nor an equivalent constraint. Consequently, as proved in this section, the GERM model formulated in a classic (wrong) WCCV framework, is not able to simulate how the system recover its homeostasis after a dynamic perturbation (as illustrated in Figure 2-17, Figure 2-10, Figure 2-16, from using a novel WCVV model). Besides, when using the classical WCCV wrong modelling framework, its predictions are distorted and inaccurate. Unfortunately, the classical WCCV continue to largely be used in the dedicated literature. As WCCV predictions are proved by Maria [2018d, 2017a, 2017b, 2018] (section 2.3.1) to be invalid and ineffective, the application of the old ("classic/default") WCCV formulation in modelling cell metabolic processes and GERMs/GRCs becomes questionable and not recommended.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Conclusions. As proved by this simple example, and by Maria [2017a,2017b,2018], the classical (default) WCCV with continuous variable ODE dynamic models are not recommended, despite they continue to be used in simple engineering/chemical calculations. By not explicitly considering the relationship between the cell volume exponential increase during the cell growth, the osmotic pressure, and species reaction rates, these WCCV 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.

2.3.2. The advantages of using the WCVV modelling framework when reproducing the cell GRC holistic properties with examples

By summarising, the study in the section 2.2.4 the characteristics of various GERM-s regulatory kinetic models, several conclusions can be derived, as followings:

1). The *in-silico* (based on a math model) study of the protein-synthesis regulation in a cell, by using numerical simulations via a continuous ODE kinetic model (build-up following the chemical engineering, and the nonlinear systems theory, and a WCVV novel approach), accounting continuous model variables and cell-volume growth, has been proved to be a worthy instrument for a quick cell-cycle evaluation under continuous/random internal-/external variable conditions. These conditions can include 'step-like' perturbations in species concentrations, continuous cell-content dilution due to the cell-volume growth, but also 'impulse-like' dynamic perturbations of the system variables. Based on the known key species nominal concentrations, on the stationary (QSS, homeostatic) mass-law equations, and on optimal regulatory criteria (P.I.s, section 2.2.3), the model kinetic 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 regulatory schemes.

2). The stationary regulatory capability of such a GERM cell-model constructed under the novel WCVV modelling framework seems to be related with the model irreversible synthesis/consumption reactions, which continuously adjust the QSS concentration-levels against the continuous cell-volume dilution effect and step-like (stationary perturbations), or impulse-like (dynamic perturbations) variations in external conditions. At the same time, the presence of as large as possible number of reversible buffering reactions of the 'catalyst' activity can quickly adjust the protein/gene ratios (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 realised. It is also to observe that, during the protein (P) recovering time-interval, the inertial effect created by the large copynumbers into the cell is keeping the cellvolume 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. In a variable-volume WCVV regulatory model such an effect is accurately described, while in a classic ("default") constantvolume WCCV kinetic model, the same effect is 'artificially' averaged by means of fictive reactions, leading to wrong predictions (section 2.3.1).

3). Among the checked GERM-s kinetic models in this work, the bestfound regulatory model [G(PP)n] (Figure 2-12, Figure 2-9, Figure 2-13, Figure 236, Figure 2-37, Figure 2-38, and Figure 2-42) includes a consecutive schema of P key-protein synthesis, and an effective regulatory scheme (tested here for "n=2 " effectors), thus better compensating the cell dilution effect, by adjusting the P/G/GPP/GPPPP ratios. The synthesis of P via an indirect 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. At the same time, the model [G(PP)n] is also superior in dynamic perturbed conditions, due to the presence of an increased number of fast reversible P -buffering reactions.

4). By including into the GERM regulatory model separate 'rawmaterials' species {NutP, MetP} for the synthesis of protein(s), and {NutG, MetG} for the synthesis of their corresponding gene(s), an increase in the model prediction flexibility is thus realised. 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*" to maintain a large sum of species copynumbers into the cell for preventing a too high sensitivity to perturbations (see section 2.2.4.-VI). However, when several proteinic regulatory modules are simulated into the same cell, the presence of separate inner 'raw-materials' for every protein/gene synthesis can offer a more flexible way to increase the regulatory effectiveness for all GERM-s, and to optimize the inner 'raw-material' levels, by maintaining the same large sum of molecules into the cell, thus better reproducing the real homeostasis of living cells.

5). As proven in this work, even if more complex and involving a supplementary identification effort, the advantages of using variablevolume WCVV kinetic models to simulate the GERM-s modules, instead of using the classic ("default") constant-volume WCCV models (sections 2.1, and 2.3.1) for the cell-system simulation are multiple: (i) allows representations of the cell-volume growth under various internal/external-cell perturbations; (ii) allows eliminating 'fictive' reactions accounting the cell-volume growth effect on species [Heinrich and Schuster, 1996] ; (iii) includes all the system species into the overall

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

mass balance (individual or lumped), with all the complexity of direct/indirect interactions (individual or lumped); (iv) 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; (v) allows deriving of more realistic stationary/dynamic quantitative regulatory indices (P.I.-s) for a studied GERM kinetic schema; (vi) allows a more correct ranking of regulatory model alternatives (see their *library* in Figure 2-12); (vii) 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. (vii) Besides, as proved in this work, the GERM-s LIBRARY of Maria [2005, 2017a, 2018] (Figure 2-12) allows the *in-silico* study of the regulatory properties of each GERM, and also comparison of several GERM-s.

The briefly reviewed GERM-s regulatory properties, in this work, proves that it is only WCVV modelling framework (section 2.2) which makes possible the quantitative evaluation of the performance indices (P.I.-s) of various GERM-s types (section 2.2.3). Such an evaluation would be impossible by using the classical ("default") constant-volume WCCV modelling framework (section 2.1). Such regulatory properties, like cell "inertial effect", and "ballast effect", presented in this work (section 2.2.4.-VI), and studied in detail by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017a, 2020c]; Maria and Scoban [2017, 2018]; Maria et al. [2017, 2018d, 2018g]; Maria and Luta [2013], are the followings:

a. The role of the high cell-"ballast" in "smoothing" the perturbations of the cell homeostasis (section 2.2.4-VI);

b. The secondary perturbations transmitted inside the whole cell via the cell volume (section 2.2.4-VI);

c. 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 (section 2.2.4-VI);

d. Allows comparing the regulatory efficiency of various types of GERM-s (section 2.2.3, and sections 2.2.4-I-...-XIII);

e. Allows a more realistic evaluation of GERM regulatory performance indices (sections 2.2.3, 2.2.4-I-...-XIII; section 2.3.1);

f. Allows studying the recovering/transient intervals between steadystates (homeostasis) after dynamic (sections 2.2.4.-I-XIII), or stationary perturbations [Maria and Luta, 2013];

g. Allows studying conditions when the system homeostasis intrinsic stability is lost [Eq.(23-24), section 2.2.3-A-iv];

h. Allows studying the cell / GERM-s self-regulatory properties after a dynamic/stationary perturbation, etc. (section 2.2.3);

i. Allows studying the plasmid-level effects in cloned cells [Maria and Luta, 2013; Maria, 2010].

6) Starting from an effective proteinic regulatory schema GERM (see their library proposed by Maria [2005,2017a,2018] in Figure 2-12), further cell-cycle model developments have to consider several proteinic regulatory modules linked in a regulatory chain, that is a GRC displaying several functions, such as: an operon expression {Figure 2-19, [Maria and Luta, 2013], section 3.2}, or a genetic switch {Fig 2-42, Figure 2-43, and Figure 2-44, [Maria, 2007, 2009, 2014b, 2018] }. The modules can be linked in several ways: i) direct connections through common species(s) and/or common reaction(s) (see the example of Figure 2-41, section 2.2.4VIII); ii) indirect connections by means of the induced cell-volume growth, to which all species contribute; iii) 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 ("second perturbations" in section 2.2.4-VI).

This modular approach to design and to build-up a cell-system GRC by using the rules 2.2.4-I-...-XIII, can allow quick simulation of complex regulatory systems under constant or random perturbed internal/external cell conditions.

To resume the main characteristics, advantages, and applications of the WCVV kinetic models in modelling living cells, the following aspects should be listed:

- P1 In fact, by introducing this **major/essential** and of high impact contribution in *Bioinformatics* and *Systems Biology*, that is the novel "*mechanistic silicon cell concept*", materialized in a novel math (kinetic) modelling framework WCVV of cell metabolic processes (referring to the "whole-cell, variable-volume", of isotonic growing cells) useful in developing cellular metabolic processes numerical simulators, Maria [2005,2017a,2018,2018d] (sections 2.2, and 2.3.1) comes *to correct* the classical "default" WCCV ("constant-volume whole cell") kinetic models where the cell volume growth is often ignored in the species mass balances, or, sometimes, considered through an average dilution constant [Dm=Ln(2)/(cell cycle)] . Here, it is worth mentioning that Maria et al. {2018d, 2017] (section 2.3.1) proved that the classical "default" WCCV formulation currently used in the literature for modelling the cell processes kinetics leads to wrong, distorted and false predictions concerning the cell (GRC) regulatory performances, i.e. their response to external/internal perturbations. In fact, the concepts and rules used in

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

the WCVV math modelling are translated from (bio)chemical engineering principles (section 1.3.4) and from the nonlinear systems theory. See the Table 2-1 for the main hypotheses of the WCVV modelling framework, and of the novel "mechanistic silicon cell concept". The properties, and the advantages of the novel WCVV are extensively explained, proved, and exemplified by Maria [2002], and by [Maria et al., 2002], while the study of the WCVV properties is given by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; Maria and Scoban [2017,2018]; Maria and Luta [2013]. Examples of applying WCVV to build-up kinetic models for various GERM-s, GRC-s, GS-s (genetic switches) are given by Maria [2017a,2017b, 2018, 2023] and in the above mentioned papers.

- P2 The novel "mechanistic silicon cell concept", and the novel math modelling framework WCVV was introduced by Maria [2002], and by [Maria et al., 2002], while the study of the WCVV properties is given by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; Maria and Scoban [2017,2018]; Maria and Luta [2013].

Such a WCVV modelling framework, to develop cell kinetic models in a holistic approach (whole-cell, WC) (see the above references) it is not only more realistic, but it presents a large number of advantages (sections 2.3, and 2.3.2). The WCVV approach is used to construct cell kinetic models that ensure the natural simulation of the cell processes homeostasis, and the individual / holistic GRC regulatory properties, by including in a natural way the constraints related to the cell system isotonicity, and the variable-cell-volume (in an exponential increase) in direct relationship to the species mass balances and the *lumped proteome/ lumped genome* replication [Maria, 2017a, 2017b, 2018]; [Maria and Luta, 2013](see as examples, i) the brown GERM lumped proteome/genome module G(PP)1] in Figure 2-19, for the mercury-operon expression in *E. coli*, and ii) the blue GERM simple module [G(P)1] lumped proteome/genome module, to simulate the lumped genome (G1) and proteome (P1) regulated replication, in Figure 2-42 of a genetic switch of *E. coli*.

In the both models, the simple module [G(PP)1] simulates the lumped genome (G1) and proteome (P1) regulated replication. This GERM unit was added to fulfil the cell isotonicity constraints Eq.(10-13). Besides, the P1/G1 high concentrations ensure the "cell ballast", and "inertial" positive effect on cell processes regulatory efficiency (section 2.2.4-VI). In other words, the use of lumped genome/proteome in the WCVV models is necessary because the isotonicity constraint imposes that all the cell species to be considered in the model (individually, or lumped) because all contribute to the cell volume (see the above Pfeiffers' law, Eq.(6, 10-13)). On the other hand, such an isotonicity constraint is required to ensure the cell membrane integrity, but also to preserve the homeostatic properties of the cell system, not by imposing "artificial" constraints like "the total enzyme activity" or the "*total enzyme concentration*" as suggested in the literature [Heinrich and Schuster, 1996]. The concepts and the basic hypotheses of WCVV dynamic modelling framework in living cells of variable volume are briefly summarized in (Table 2-1). As proved in the cited references of this paragraph "P2", this novel modelling framework WCVV is leading to accurately simulate a lot of cell metabolic intrinsic processes, such as the relationships between the external conditions and the inner cell metabolic processes, like: a). species net synthesis reactions, b). osmotic pressure, c). cell content (ballast) influence on smoothing the continuous perturbations in the external nutrient concentrations, etc. See the examples of the point no. 5 of this section 2.3.2, and [Maria, 2017a, 2018, 2023]. An example is provide by section 3.2.

P3 By contrast to the classical WCCV, by imposing in the WCVV formulation a constraint accounting for the cell-volume growth while preserving a constant osmotic pressure and membrane integrity (see the WCVV model hypotheses in Table 2-1), dr. Maria proved step-by-step the importance of using modular WCVV cell kinetic models when developing GRC dynamic models with including a variable number of GERM-s (see sections 2.2.4, 3.2).

Another major/essential and of high impact contribution of dr. Maria in Bioinformatics and Systems Biology, concerns the construction in a WCVV framework of a MATH LIBRARY including lumped kinetic models, for various types of individual GERMs (Figure 2-12), in a modular template format, ready to be used to construct GRCs (operon expression, genetic switches, genetic amplifiers, etc., see applications of Maria [2018,2023], by linking individual GERMs.

Maria [2005, 2017a, 2018] also derived the **rules to link** such GERM-s in GRC-s (section 2.2.4; with the example of section 3.2)

Maria [2005, 2017a, 2018] also studied the regulatory properties of various GERMs in relation to their model structure. And he developed "regulation efficiency performances indices (P.I.-s) of individual GERM-s" (section 2.2.4).

P3 By contrast to the classical WCCV, by imposing in the WCVV formulation a constraint accounting for the cell-volume growth while preserving a constant osmotic pressure and membrane integrity (see the WCVV model hypotheses in Table 2-1), Maria [2005, 2017a, 2018] proved step-by-step the importance of using modular WCVV cell kinetic models when developing GRC dynamic models with including a variable number of GERM-s. (see sections 2.2.1.2, 2.3, 2.3.2).

P4 By using this modular, structured, and deterministic (reaction mechanism based, of continuous state variables for the cell key-species), dr. Maria developed **hybrid SMDHKM** kinetic models, by linking the nano-scale essential parts of the CCM

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

(glycolysis, ATP recovery system, TCA cycle, GLC import system, etc.), to various operon expression GRC modules, and to the macro-scale bioreactor state-variables (section 3.2). Such hybrid complex dynamic models have been used for *in-silico* design of GMOs [Maria, 2021; Maria and Renea, 2021], or to *in-silico*, off-line derivation of the bioreactor optimal operation/control policy, with a higher accuracy and degree of details (no. of considered state variables) (section 3.2) See the examples provided by Maria [2018, 2021, 2023]; [Maria and Renea, 2021; Maria and Gheorghe, 2024].

P5 The developed novel concepts and rules, introduced together with the novel WCVV modelling framework by dr. Maria [2002, 2005, 2017a, 2018] and described in this work, are, in fact, translated from those of the (bio)chemical engineering and of the nonlinear systems theory (Figure 1-5 and Figure 1-6, sections 1.1, and 2). They are better explained, proved, and exemplified by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; Maria and Scoban [2017,2018]; Maria and Luta [2013].

P6 By contrast to the classical ("default") (constant-volume) WCCV cell dynamic models (section 2.1), Maria [2002], and [Maria et al., 2002], and Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; Maria and Scoban [2017,2018]; Maria and Luta [2013] introduced a novel kinetic modelling WCVV framework (section 2.2), by including the thermodynamic isotonicity relationships/constraints (see its hypotheses in Table 2-1), and, proved step-by-step in a mathematical way how such constraints can reproduce (simulate) the cellular intrinsic regulatory properties in a natural way, by using simple reduced kinetic models (that is not derived from artificially imposed hypotheses/constraints) {see [Maria, 2005, 2017a, 2018], and section 2.2 }. The novel WCVV approach, explicitly linking the volume growth, external conditions, osmotic pressure, cell content "ballast", and the net reaction rates for all cell-components (individual or lumped, Eq. 6, 10-15; Table 2-1) is very suitable for predicting local and holistic regulatory properties of the GERM-s / GRC-s, and of the CCM metabolic network, leading to more realistic predictions of some cell properties (for instance, evaluation of the regulatory efficiency indices of individual GERM-s, or GRC-s; cell content "inertial/smoothing" effect (section 2.2.4-VI) in treating perturbations, etc.). Among the considered GRC regulatory efficiency indices, the most important are reviewed and discussed in the sections 2.2.3, and 2.2.4. The regulatory properties of complex GRC-s and individual GERM-s were defined by using quantitative measures imported from the nonlinear system control theory (section 2.2.3).

P7 By contrast to the classical ("default") WCCV "constant-volume" kinetic modelling approach (section 2.1), the novel WCVV math modelling framework (section 2.2) of cell dynamic processes proposed by Maria [2002], and by [Maria et al., 2002], and studied by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; [Maria and Scoban, 2017, 2018; Maria and Luta, 2013], imposes the isotonicity constraint [that is the Pfeffers' law, Eq.(6,11-15)] to ensure the cell membrane integrity (see the differences between these two approaches when modelling the kinetics of a simple GERM in the section 2.3.1. Such a major constraint requires that all species (individual or lumped) have to be considered in the dynamic cell model [see the proof given by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; [Maria and Scoban, 2017, 2018; Maria and Luta, 2013]. Due to such reasons, Maria included in all their applications with using WCVV kinetic models the lumped genome, proteome, and metabolome. See the followings two examples in this respect: i) the brown module [G(P)1] in Figure 2-6, concerning simulation of the induced mer-operon expression in *E. coli* [Maria, 2010; Maria and Luta, 2013], and ii) the blue module [G(P)1] in Figure 2-42, Figure 2-43, and Figure 2-44, concerning *in-silico* simulation and design of a genetic switch in *E. coli* with the role of a biosensor for detecting molecular inducers of micro-/nano-level concentrations [Maria, 2014b, 2018]. Such a mandatory WCVV modelling approach is necessary because all these big lumps of metabolites, nutrients, genome, and proteome contribute to the cell volume Eq.(6,11-14), and to its dynamics via the isotonicity constraint (adopted Pfeiffers'law of diluted solutions, Eq.(6)). For details see also the review books of Maria [2017a, 2017b, 2018, 2023], and the papers of Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017A, 2017B, 2018, 2023, 2023a, 2024, 2024b, 2024c]; [Maria and Scoban, 2017, 2018; Maria and Luta, 2013].

Repeated *in-silico* analyses of a large number of GERM-s systems approached in a WCVV modelling framework lead to derive a large number of conclusions, all pointing-out the advantage of using the novel WCVV "variable-volume" modelling framework when reproducing the cell self-regulatory properties (impossible to be simulated by using the default / classic WCCV "constant-volume" modelling framework)(see sections 2.2.2, 2.2.3, 2.2.4).

The briefly reviewed GERM-s regulatory properties, in this work (sections 2.2.2, 2.2.3, 2.2.4), proves the same idea: it is only the WCVV modelling framework (section 2.2) which makes possible the quantitative evaluation of the performance indices (P.I.-s) of various GERM-s / GRC-s (section 2.2.3). Such an evaluation would be impossible by using the classical ("default") constant-volume WCCV modelling framework (section 2.1). Such regulatory properties, such as the cell "inertial effect", and "ballast effect", presented in this work (section 2.2.4-VI), and others and studied in detail by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017a, 2020c]; Maria and Scoban [2017,2018]; Maria et al. [2017, 2018d, 2018g]; Maria and Luta [2013], are presented in sections 2.2.2, 2.2.3, 2.2.4), and proved by the example given in section 3.2.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

There are lot of positive outputs when using the novel WCVV modelling framework (sections 2.2.1.2, 2.3. 2.3.2). At the risk of repeating some ideas mentioned in various sections, below are underlined only some of them.

Issue-1).- Application of chemical engineering principles, rules, and algorithms (Figure 1.-5, Figure 1.-6 , and section 1.3.4) as well as those of the nonlinear system control theory, together with model lumping techniques (section 2.1 - lumping rules) to elaborate modular kinetic models of various cell processes, especially those concerning the GRC-s (genetic switches, operons expression, genetic amplifiers of external signals, etc., section 2.2.2). Over the all above mentioned publications, dr. Maria promoted all time the concept of the best trade-off between simplicity and model quality/adequacy.

Issue-2).- Development of new concepts and promotion of a novel dynamic/structured math modelling framework , that is "the mechanistical silicon cell", and its resulted WCVV kinetic modelling framework (the so-called "variable-volume-whole-cell" approach of modular, structured metabolic pathways for isotonic cell systems) to simulate the metabolic cell processes, and especially the GERM / GRC regulatory ones.

The WCVV replaces the classical ("default") constant-volume (WCCV) kinetic modelling (section 2.1), by explicitly including equations linking the cell volume growth, the quasi-constant osmotic pressure (isotonic cell system), environmental conditions, the cell-content (the so-called "cellbalast"), with the net reaction rates of all cell components (accounted individually or lumped)(that is WCVV in section 2.2). Compared to the "default" / classical WCCV kinetic models where the cell volume growth is neglected, or sometimes considered through an average dilution constant (Dm), the novel WCVV math (kinetic) modelling framework of cell processes was proved to offer more realistic predictions of local and global regulation properties of GRC controlling the metabolic syntheses, by employing an "instant" cell dilution rate (Di, see Eq.(11-14), section 2.2). Thus, the WCVV models can simulate, in a more realistic way, the efficient treatment of stationary or dynamic environmental perturbations, species connectivity, cell homeostasis stability and multiplicity, responsivity to perturbations, etc.) {see sections 2.2.3, 2.2.4, 2.3}. See the proof of section 2.3.1. The most studied applications of WCVV kinetic models are related to the GERM-s (protein synthesis), and GRCs. Among the large number of contributions in this area of Maria [2017a, 2017b, 2018, 2023] are to be briefly mentioned the followings:

a. The WCVV approach, explicitly linking the volume growth, external conditions, osmotic pressure, cell content ballast, and the net reaction rates for all cellcomponents. Such a kinetic modelling framework is more suitable for predicting local and holistic properties of the metabolic network, leading to more realistic predictions of some cell properties (e.g. evaluation of the regulatory efficiency (dynamic or stationary, section 2.2.4) of individual GERM-s, or GRC-s; evaluation of cell content inertial/smoothing effect in treating external perturbations, section 2.2.4-VI). The regulation efficiency performances indices (P.I.-s) of individual GERM-s, section 2.2.3), that is: stationary regulation, dynamic regulation, regulatory robustness, species interconnectivity, quasi-steady-state (QSS, homeostasis) stability, QSS stability strength, and others.

b. Promotion of the holistic and modular approach in dynamic modelling of GRC-s; elaboration of various math/kinetic models for GERM-s, organized in a GERM-LIBRARY (Figure 2-12); *in-silico* study of GERM chain properties in a genetic regulatory network (GRC) by *in-silico* "placing" them in a growing cell, and by mimicking (by numerical simulation) the cell behaviour under *in-silico* simulated environmental conditions (stationary or perturbed, section 2.2.4);

c. Development of a methodology for GERM assembling to obtain a GRC-s (e.g. to simulate a defined operon expression [Maria, 2010; Maria and Luta, 2013]). This GERM linking methodology (section 2.2.4) is useful for *in-silico* design of GMO-s with desired motifs for industrial or medical applications;

d. Applications of the modular WCVV variable-volume holistic modelling approach for the *in-silico* design of GRC-s, such as genetic switches acting as biosensors (Figure 2-42, Figure 2-43, Figure 2-44)[Maria, 2007, 2009, 2014b, 2018]. A case study to design a modified *E. coli* cloned cell, with a genetic amplifier circuit to maximize the mercury uptake from wastewaters is reviewed in section 3.2;

e. Study of the criteria associated to gene knockout strategies for *in-silico* design of optimal metabolic fluxes allowing the cell growth with maximizing chemical production targets (a case study for succinate and biomass production concomitant maximization in *E. coli* cells, is given by Maria et al.[2011]; Maria [2018, 2023]).

Issue 3). Building-up hybrid mathematical models, that is structured, modular, deterministic, *hybrid* kinetic models (SMDHKM). Such deterministic kinetic models, with continuous variables, structured and modular, based on the cell metabolic reaction mechanisms, allow linking structured WCVV nano-scale cell-level kinetic models (of GERM / GRC-s, and of the central carbon metabolism CCM) with the macro-scale dynamic models (including macro state-variables) of the industrial bioreactors [Maria, 2020a, 2021; Maria and Renea, 2021].

Issue 4).- Development and promotion of novel concepts for the kinetic structured math modelling of metabolic processes in living cells That is the "mechanistic silicon cell novel concept" materialized in promoting the novel WCVV modelling framework

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

("whole-cell, variable cell-volume of isotonic cell systems"). As proved in this work and cited publications, this novel computational tool is very suitable to develop extended / reduced models for the protein synthesis, that is gene expression (GERM), for the lumped genome/proteome replication, genetic regulatory circuits (GRC), central carbon metabolism (CCM), etc., in a holistic approach (of „whole cell' type). Details about WCVV, are given in the section 2.2, while applications are given by Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017a, 2020c]; Maria and Scoban [2017,2018]; Maria et al.[2017, 2018d, 2018g]; Maria and Luta [2013].

The "mechanistic silicon cell novel concept", and the novel WCVV math modelling concept (section 2.2) replaces the classical ("default") constant-volume (WCCV) kinetic modelling from the current literature (section 2.1). By contrast, Maria [2018d] (section 2.3.1) proved that application of the classical math modelling approach WCCV to simulate cell processes is wrong, leading to distorted/ incorrect predictions (section 2.3.1). The novel WCVV structured kinetic model explicitly includes equations linking the cell volume growth, the quasi-constant osmotic pressure (isotonicity), the environmental conditions, the cell-content (the so-called ,,cell-balast"), and the net reaction rates of all cell components (accounted individually or lumped, section 2.2). Compared to the classical WCCV kinetic models (section 2.1) where the cell volume growth is neglected, or sometimes considered through an average dilution constant (Dm), the novel WCVV was proved to offer more realistic predictions of local and global regulation properties of GERM-s/ GRC-s controlling the metabolic syntheses (e.g. efficient treatment of stationary or dynamic environmental perturbations, species connectivity, cell homeostasis stability and multiplicity, responsivity to perturbations, etc., sections 2.2.3, 2.2.4).

Issue 5). The volume-growth diluting effect explicitly quantified [see the instant "Di" in Eq.(11-15)] in a WCVV kinetic model formulation (section 2.2) acts as a continuous stationary perturbation of all species concentrations, and can formally be assimilated with a first order decay rate of all cellular species during the cell-cycle Maria [2002, 2003, 2005, 2006, 2007, 2009, 2014b, 2017, 2017a, 2020c]; Maria and Scoban [2017,2018]; Maria et al.[2017, 2018d, 2018g]; Maria and Luta [2013].

Issue 6). The use of the WCVV framework to develop and promote structured lumped kinetic models concerning the gene expression individual regulatory lumped modules (GERM). These GERM modules kinetic models, of different complexity have been organized in a LIBRARY (Figure 2-12) including template lumped math (kinetic) models of GERMs, easy to be used to build-up modular genetic regulatory circuits (GRC) (e.g. operon expression, genetic switches, genetic amplifiers, etc., section 2.2.4 [Maria, 2009, 2010, 2014b, 2018, 2023]). These GERM lumped kinetic modular deterministic models can be *in-silico* analyzed concerning their regulatory efficiency and properties, by using quantitative measures (that is regulation performance indices (P.I.-s), section 2.2.3). The quantitative (P.I.-s) have been introduced by Maria [2005, 2017a, 2018] to characterize the GERMs kinetic models. The GERM LIBRARY includes individual GERMs in a template form, ready to be used to construct complex GRCs by using building blocks rules of the *Systems Biology* and *Synthetic Biology* (section 2.2.4, with the examples of Maria [2010, 2013, 2014b]). Rate constants of GERM-s/ GRC-s kinetic models are estimated from the homeostatic cell species concentrations (section 2.2.1.3), , by solving the cell kinetic model rewritten for the quasi-steady-state conditions Eq.(11-15), and from mimicking the GRC response to dynamic or stationary perturbations to optimize the GERM / GRC regulatory properties (section 2.2.3). The GERM, GRC, and CCM kinetic models of WCVV type can be used for the *in-silico* design and test of GMOs of desired characteristics (some examples are given by Maria [2017a, 2018, 2023], and in section 3.2).

Issue 7). The cell system *isotonicity constraint* in a WCVV formulation is leading to accurately simulate a lot of cell metabolic effects, such as (see section 2.2.4): relationships between the external conditions, species net synthesis reactions, osmotic pressure, cell content ("ballast") influence on smoothing the continuous perturbations in external nutrient concentrations, etc. (see section 2.2.4-VI, and the above P7 paragraph).

Issue 8). By contrast to the classical ("default") WCCV cell kinetic modelling approach (section 2.1), the novel proposed WCVV kinetic modelling framework allows a more realistic study of the regulatory properties of complex GRC-s and individual GERM-s by using quantitative measures of their efficiency (P.I.-s in section 2.2.3) derived from the nonlinear system control theory [Maria, 2005, 2017a, 2017b, 2018, 2023]. Based on these P.I.-s quantitative measures, dr. Maria developed *rules of linking* individual GERM-s to construct GRC-s following a "building-blocks" strategy (section 2.2.4), thus allowing development of WCVV kinetic models for GRC-s (operon expression, genetic swiches, etc.), useful for *in-silico* design of GMO-s [Maria, 2017a, 2018, 2023].

Issue 9). The WCVV formulation of various GERM-s in different cells allowed to Maria [2005, 2007, 2014b, 2017a, 2018] to develop a GERMs LIBRARY of reduced kinetic models for various GERM-s types of modules given in (Figure 2-12), ready to be use for the *in-silico* study of the regulatory properties of individual GERM-s in respect to stationary or dynamic perturbations in the cell/ environment, as done in this work (sections 2.2.3, 2.2.4, 2.3). This library of GERM-s lumped kinetic model of all types is ready to be used for the *Synthetic Biology* purposes when designing novel GMO-s of desired characteristics ("motifs"). See some examples in sections 2.2.4, 3.2, provided by Maria [2010, 2014b]; [Maria and Luta, 2013].

Issue 10). *In-silico* prove by using the WCVV models the cooperative *vs.* concurrent linking of GERM-s in a GRC, and species close interconnectivity by preserving specific functions of proteins inside the cell [section 2.2.4.-VIII]. The WCVV kinetic models are also very suitable to study the GERM-s interactions in various GRC-s of different functions (e.g. toggle-switch, amplitude filters, modified operons, etc). [Maria, 2018, 2023]. See the example given in section 3.2.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Issue 11). Maria [2005, 2017a, 2017b, 2018]; Maria and Scoban [2017,2018] exemplified how the kinetic model rate constants can be separately identified for each GERM module of the analysed cell, by using the experimental stationary (homeostatic) concentrations of the species included in the WCVV kinetic model (see section 2.2.1.3), and by imposing some regulatory properties, and then extrapolating them when simulating the whole GRC, by assuming that linking reactions are relatively slow compared to the individual module core reactions. See the developed GERM-s linking rules in section 2.2.4.

In such a manner, linked modules are able to respond to changes in the common environment such that each module remains fully regulated [Maria, 2003, 2005, 2006, 2007, 2014b]; [Maria and Scoban, 2017, 2018; Maria et al., 2018 g]. The advantage of such a modular approach comes from the possibility to reduce the model complexity for the central carbon metabolism (CCM), or large pathways of complex GRC-s, and the size of the identification problem, by assuming, 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 [Styczynski and Stephanopoulos, 2005; Zak et al., 2005].

Issue 12). Such a WCVV holistic modelling framework and isotonicity constraint imposes that all species (individual or lumped) have to be considered in the dynamic cell model Eq.(6, 11-15). Due to such reasons, Maria [2007, 2009, 2014b, 2018, 2023] included in all their WCVV applications with linking GERM-s to build-up complex GRC-s, the lumped genome, lumped proteome, and lumped metabolome, because all these big lumps (of high copynumbers / concentrations) substantially contribute to the cell volume, and volume dynamics Eq.(11-15) via the isotonicity constraint (adopted Pfeiffers'law of diluted solutions, Eq.(6), section 2.2).

Thus, the WCVV approach is used to construct cell kinetic models that ensure the natural simulation of the cell processes homeostasis, and the individual / holistic GRC regulatory properties, by including in a natural way the constraints related to the cell system isotonicity, and the variable-cell-volume (in an exponential increase) in direct relationship to the species mass balances and the lumped proteome/ lumped genome replication [Maria, 2017a, 2017b, 2018]. As examples see: i) the brown GERM lumped proteome/genome module G(PP)1] in Figure 2-19, for the mercury-operon expression in *E. coli*, [Maria and Luta, 2013], and ii) the blue GERM simple module [G(P)1] lumped proteome/genome module, to simulate the lumped genome (G1) and proteome (P1) regulated replication, in Figure 2-42 of a genetic switch of *E. coli*.

In the both models, the simple module [G(PP)1] simulates the lumped genome (G1) and proteome (P1) regulated replication. This GERM unit was added to fulfil the cell isotonicity constraints Eq.(10-13). Besides, the P1/G1 high concentrations ensure the "cell ballast", and "inertial" positive effect on cell processes regulatory efficiency (section 2.2.4-VI). In other words, the use of lumped genome/proteome in the WCVV models is essential and necessary because the isotonicity constraint imposes that all the cell species to be considered in the model (individually, or lumped) because all contribute to the cell volume (see the above Pfeiffers' law, Eq.(6, 10-13)). On the other hand, such an isotonicity constraint is required to ensure the cell membrane integrity, but also to preserve the homeostatic properties of the cell system, not by imposing "artificial" constraints like "the total enzyme activity" or the "total enzyme concentration" as suggested in the literature [Heinrich and Schuster, 1996]. The concepts and the basic hypotheses of WCVV dynamic modelling framework in living cells of variable volume are briefly summarized in (Table 2-1). As proved in the cited references of this paragraph "P2", this novel modelling framework WCVV is leading to accurately simulate a lot of cell metabolic intrinsic processes, such as the relationships between the external conditions and the inner cell metabolic processes, like: a) species net synthesis reactions, b) osmotic pressure, c) cell content (ballast) influence on smoothing the continuous perturbations in the external nutrient concentrations, etc. See the examples of the point no. 5 of this section 2.3.2, and [Maria,2017a, 2018, 2023]. An example is provided by section 3.2.

Issue 13). By solving a certain number of case studies, Maria [2006, 2007, 2009, 2014b, 2017a, 2018] derived a certain number of rules to be followed for the *in-silico* build-up a GRC dynamic math model under the WCVV framework (section 2.2.4). By using simple reduced WCVV kinetic models, various types of lumped kinetic models of GERM modules can individually be analyzed as mechanism, reaction pathway, regulatory characteristics, and regulatory effectiveness (sections 2.2.3, 2.2.4). As a limited number of regulatory module types governs the protein synthesis [Maria, 2005, Sewell et al., 2002; Yang et al., 2003], it is computationally convenient to step-by-step build-up the modular GRCs by applying certain principles and rules {section 2.2.4), and then adjusting the network holistic regulatory properties by means of the reestimated kinetic constants, by also accounting not only on the stationary (known) species concentrations, but also on objectives concerning the GRC regulatory effectiveness [Maria, 2005, 2014b, 2017a; Maria and Scoban, 2017, 2018; Maria and Luta, 2013]. In other words, it is desirable to focus the metabolic regulation and control analysis on the regulatory/control features of the holistic GRC-s, and of the functional GERM-s subunits than to limit the analysis to only kinetic properties of individual enzymes acting over the synthesis pathway.

Issue 14). Maria [2021, 2023], and [Maria and Luta, 2013] introduced *hybrid* structured, modular, deterministic, kinetic models (SMDHKM), derived under the WCVV novel modelling framework for various purposes: i) *in-silico* design of GMOs [Maria and Luta, 2013; Maria, 2021, 2023], and ii) bioreactor optimization with a higher accuracy and degree of detail [Maria and Renea, 2021; Maria, 2020a, 2023]. See an example in the below section 3.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The novel WCVV modelling framework is very suitable to develop **hybrid SMDHKM** structured kinetic models of cell processes, of following essential characteristics [Maria, 2021, 2023; Maria and Luta, 2013; Maria and Renea, 2021; Maria, 2020a] (see also sections 3.1, 3.2): *hybrid kinetic models* able to predict the key-species dynamics at the cell level (excretable or not cell species), in an explicit connection to the predicted macroscopic state-variables at the bioreactor level, (presented in the bioreactor bulk-phase); *structured models*, by including the cell key species involved in the studied bioprocess (belonging to the central carbon metabolism CCM, or to a target GRC); of a *modular construction*, the reaction pathway of modules being interconnected; *deterministic models* (with continuous variables, based on the cell metabolic reaction mechanisms). These cell mathematical **SMDHKM** kinetic models present a detailing degree (that is the number of involved cell key species and reactions, taken individually, or lumped) depending on the each approached case study, but incomparably larger (by orders of magnitude) than the *un-structured (global / apparent)* (Monod / Michaelis-Menten) kinetic models.

Examples of developed *hybrid SMDHKM* structured kinetic models (bilevels), of cell processes and bioreactors, are briefly summarized below. An example is provided in section 3.2.

To realize these ambitious objectives, *Systems Biology* uses a wide range of tools, but mainly complex mathematical simulation models linked to omics databanks [Maria, 2017a, 2017b, 2018]; [Heinemann and Panke, 2006; Myers, 2009; Qian et al., 2017].

Issue 15-ex.1). -- A case study refers to the *mercury*-operon (*mer-operon*) expression in Gram-negative bacteria (such as *E. coli*, *Pseudomonas sp*.) to uptake the mercury ions from wastewaters [Maria, 2010; Maria and Luta, 2013; Maria et al., 2013]. The structured WCVV hybrid SMDHKM kinetic model was used to design GMO-s with cloned *mer*-plasmids ("mer" denotes mercury here) to result a globally efficient mer-operon (i.e. a GRC) leading to optimize the fed-batch bioreactor (FBR) productivity. This extended *hybrid* math dynamic model was proved to be useful for several purposes: i) the process design and ii) bioreactor optimal control, thus allowing predicting the wild/cloned bacteria metabolism adaptation (i.e. the response of mer-operon expression) over hundreds of cell generations under variable operating conditions of the bioreactor as studied by [Maria. 2010; Maria and Luta, 2013; Maria et al., 2013]. This cell structured hybrid dynamic model can *in-silico* predict the maximum level of mer-plasmids that can be added to the GMO cell genome for improving the mercury uptake rate without exhausting the cell resources (that is not putting in danger the cell survival). It is to underline that the parameters of the structured WCVV *hybrid* SMDHKM kinetic model have been estimated based on experimental data. Also the predictions of associated bioreactor optimization, by using this WCVV hybrid SMDHKM kinetic model, have been check on an experimental basis.

Issue 15-ex.2). Another example refers to the *in-silico* design of a GMO *E. coli* including a genetic switch (GS) of adjustable characteristics, amplifying exogeneous stimuli to act as a molecular-level *biosensor*, or involved in signal transduction [Maria, 2007, 2009, 2014b, 2018]. The developed mechanistic modular WCVV kinetic models are proved to be valuable tools in studying and modulating the design desirable GS with fully interpretable parameters, thus avoiding *unstructured* apparent Hill-type lumped models of low quality.

An example was provided by Maria [2009] in the case of a genetic switch (GS) in *E. coli* cell, modelled under the WCVV approach. The two considered self- and cross-repressing gene expression modules in (Figure 2-42) are of the following type (see the GERM library of Figure 2-12) [G2(P2P2)1(P3P3)1] + [G3(P3P3)1G3(P2P2)1]. The principle of a GS is briefly illustrated in Figure 2-43 and Figure 2-44. The two self- and cross-repressed gene expression coupled modules GERM 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. The GERM module of simple [G(P)1] type mimics the lumped genome and proteome replication during the cell cycle. Such a module is absolutely necessary to fulfill the isotonicity constraint of the WCVV model Eq.(6,11-15) requiring that all cell species to be considered in the kinetic model (individually or lumped). Beside, the large concentrations (copynumbers) of the lumped genome/proteome P1/G1 will lead to preserve a high cell load ("ballast"), thus smoothing the small perturbations in the environment (section 2.2.4-VI).

These well chosen GERM-s type models are very flexible, allowing adjusting the regulatory properties of the **GS** (i.e. switch certainty, good responsivity to inducers, good dynamic and stationary efficiency [Maria, 2009, 2014b]). Besides, based on adequate kinetic WCVV models Maria [2009] proved that it there exists an optimal low level of the TF-s (that is [P2P2]s, or [P3P3]s for the GS model) that are associated to the optimal holistic regulatory properties of the GRC (low sensitivity *vs.* external nutrients, but high *vs.* inducers). Besides, Maria [2009, 2014b] proved that these TF-s are rather dimmers than monomeric molecules, in order to maximize the resulted **GS** performance indices. An optimal 0.1-4nM was found to ensure such goals [Maria and Luta, 2013; Maria, 2009, 2014b]. These *in-silico* obtained results have been confirmed by the literature experimental data. Such cell GS-s are very suitable to be used as a molecular biosensor [Maria, 2018, 2023], being able to detect environmental inducers of micro-/nano- molar concentrations, with application in airports, civil / military industry, etc.

Issue 15-ex.3).- Such WCVV *hybrid* SMDHKM structured reduced *modular and bi-level* kinetic models were proved to be effective in developing kinetic models to simulate the dynamics of the essential reaction pathway modules, all belonging to the central carbon metabolism (CCM), in an extended or reduced (lumped) formulation, and in direct relationship to the bioreactor macro-scopic state

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

variables. Examples includes: a) stationary or oscillatory glycolysis, ATP recovering system, PTS system, TCA cycle, and many other modules (iron metabolism in mitochondria, TRP operon expression, SUCC synthesis, etc.) [Maria, 2018, 2023]. As above mentioned, the *hybrid* dynamic models of WCVV- SMDHKM bi-level type in fact are linking the structured cell CCM / GRC models (of a nano-scale representation of the key-species mass balances) to the dynamic model of the bioreactor (of a macro-scale representation of their state variables). [Maria, 2021; Maria and Renea, 2021; Maria and Gheorghe, 2024].

By using the same novel WCVV concept, [Maria, 2014, 2020c; Maria et al., 2018b] proposed a structured reduced dynamic model to simulate the *glycolysis* in prokaryote cells under stationary or *oscillating* dynamic conditions of the bioreactor. The model can be further used as the core of a *hybrid* modular SMDHKM dynamic model used to simulate the central carbon metabolism (CCM) and the regulation of a target metabolite synthesis (that is the tryptophan - TRP), [Maria, 2021; Maria and Renea, 2021, Maria et al., 2018a, 2018c; Maria and Gheorghe, 2024], with applications for the *in silico* reprogramming of the cell metabolism and design of GMO-s of various applications [Maria et al., 2019; Maria, 2018, 2023]. Thus, based on this extended MCSMD hybrid dynamic model for the TRP-synthesis in *E. coli*, a GMO *E. coli* was *in-silico* designed and then experimentally realized in order to improve the tryptophan (TRP) synthesis in a fed-batch (FBR) bioreactor. Thus, compared to a simple batch bioreactor (BR) using a "wild" *E. coli* cell culture, the TRP production was increased with 73% (53%due to the novel GMO *E. coli* strain better performances, and 20% due to the FBR optimization using a *hybrid* modular SMDHKM dynamic model. [Maria, 2020c, 2021, 2023, 2014a; Maria and Renea, 2021; Maria et al., 2018a, 2018c, 2018b; Maria and Gheorghe, 2024]

Issue 15-ex.4).- Hudder, Maria, et al. [2002](see the caption of Figure 1-40) proposed a simplified (lumped) WCVV modular structured kinetic model to simulate the iron metabolism in mitochondria with applications in medicine to simulate the heme synthesis regulation.

3. Importance of the structured, modular, deterministic, hybrid kinetic models (SMDHKM), derived under the WCVV novel modelling framework for the *in-silico* GMOs design and bioreactor optimization

3.1. Hybrid, multi-level SMDHKM dynamic models

According to the previous Issue 14 of section 2.3.2, Maria [2023, 2021]; [Maria and Luta, 2013] introduced the so-called *hybrid structured, modular, deterministic, kinetic models* (SMDHKM), derived under the WCVV novel modelling framework for various purposes: i) the *in-silico* design of GMOs [Maria and Luta, 2013; Maria, 2021, 2023], and ii) bioreactor optimization with a higher accuracy and degree of detail (that is a higher number of considered species in the dynamic model)[Maria and Renea, 2021; Maria, 2021, 2020a, 2023].

Even if very effective, the disadvantages of the *hybrid* SMDHKM kinetic models are related to their high complexity (number of considered species), leading to estimation difficulties of the all rate constants, and to a higher computational time, and additional difficulties to manipulate the model for quick engineering analyses.

The main motivations of developing *hybrid* SMDHKM kinetic models, despite the above mentioned disadvantages of the hybrid SMDHKM kinetic models, are related to the important advantages of using such math models in engineering evaluations, which fully justify their use for the above listed purposes (i-ii). Such strong motivations are related to the huge development over the last decades of the industrial biosynthesis, and its involved problems, as followings:

1). The industrial biosynthesis has become one of the most significant branches of the pharmaceutical and food industry, having a major economic/social impact. Enzymatic and biological production of chemicals is of the order of 10^2 – 10^6 t/year, being extremely profitable, with an average cost of approx. 100 – 1000 $/kg. Biosynthesis successfully replaces complex, energy-consuming chemical syntheses that generate toxic waste. Biosynthesis proceeds with very high selectivity in conditions close to the ambient ones.

2). An essential engineering problem refers to the *in-silico* (mathmodel based) optimization of the biological reactors;

3). The current ("default") approach in bioengineering for *in-silico* solving the design, optimization and control problems of industrial biological reactors is the use of *un-structured (global, apparent)* kinetic models of the Monod type (bioreactors, cell cultures), or MichaelisMenten (enzyme reactors), that ignore the details of the cellular processes, or the interactions between multiple cell enzymatic reactions. These global models provide only an approximate representation of the bioprocess and can not make any correlation between the bioreactor operation and the continuous adaptation of the biomass metabolism. Their predictions are imprecise, with a low degree of detail (number of considered species dynamics), being unsuitable for an effective bioreactor optimization.

The novel WCVV modelling framework is very suitable to develop ***hybrid SMDHKM*** structured kinetic models of cell processes, by considering the following essential characteristics [Maria, 2021, 2023; Maria and Luta, 2013; Maria and Renea, 2021; Maria, 2020a]:

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Process kinetics. A set of ordinary differential equations, characterizing the differential mass balances of cellular species (considered individually or grouped). Such a *hybrid* (bi-level) kinetic model is able to predict the key-species dynamics at the cell level (excretable or not cell species), in an explicit connection to the predicted macroscopic state-variables at the bioreactor level, (presented in the bioreactor bulk-phase). According to the WCVV hypotheses (Table 2-1), and to the isotonicity constraint Eq. (6,11-15), such models should include all the cell species (individually, or grouped).

Deterministic. Kinetic mmodels with continuous variables, based on the mechanism of the cell metabolic reactions, that is on the cell reaction pathway (all reactions taken individually, or lumped). The *hybrid* SMDHKM kinetic models present a detailing degree (that is the number of cell key species and reactions, considered in the model, individually, or lumped) depending on the each approached case study, but incomparably larger (with orders of magnitude) than the *un-structured* (global / apparent) (Monod / Michaelis-Menten) kinetic models.

Modular. The model contains reaction modules that characterize the essential considered sub-processes of the central carbon metabolism (CCM), or of GERM-s, or GRC-s. The modules are considered as being inter-connected (directed, via the common species / reactions, or indirect, via the cell volume to which all species contribute).

Structured. The model includes the cell key species involved in the studied bioprocess (belonging to the central carbon metabolism CCM, or to a target GRC). Ideally, the kinetic SMDHKM model must includes the essential steps of the CCM, (individual, or lumped [Maria, 2021]), plus the target metabolite synthesis module (for instance, the TRP synthesis in Issue 15-ex.3, of section 2.3.2). The cell species are interconnected to the environmental (bioreactor) species directly (via the imported nutrients), or indirectly (via the excreted species / metabolites).

Hybrid. The model should include at least two levels of bioprocess representation: i) species dynamics at the cellular level; ii) species dynamics at the macroscopic level (bioreactor state variables, in the "bulk" liquid phase).

In-silico. A numerical analysis and process simulation by using the SMDHKM kinetic model of the bioprocess / bioreactor.

Examples of CCM modules: PTS system (phosphotransferase system of the glycolysis, that is the first steps of the glucose GLC import into the cell); glycolysis; (TCA, citric acid cycle), or (Krebs cycle); ATP (Adenosine triphosphate) regeneration system; TRP (tryptophan) synthesis and operon expression, etc. [Alberts et al., 2002; Schellenberger et al., 2011; Maria, 2021]

In the case of the biological reactor (with free cell cultures or immobilized on a porous solid support), the trend in the industry is to use complex systems, with genetically modified microorganisms (GMOs), and efficient immobilization systems (which prevent biomass inactivation due to mechanical and chemical stress in the culture medium). These biosyntheses successfully replace energy-consuming and polluting chemical syntheses, by using milder reaction conditions and generating very little waste [Gavrilescu and Chisti, 2005; Hempel, 2006; Buchholz and Hempel, 2006]. Thus, modern biological processes together with enzymatic ones prove to be efficient in the biosynthesis of numerous chemical compounds, competing in terms of efficiency with organic chemical synthesis, proceeding with high selectivity and specificity, producing fewer by-products, reducing energy consumption and generating less environmental pollution (Figure 2-45). This characteristic of industrial biosynthesis is exploited for various economic purposes (industry, medicine, environment, agriculture, fuel production) [Buchholz and Hempel, 2006; Hempel, 2006; Moulijn et al., 2001].

However, industrial bioprocesses still have a limited spread due to the high costs of isolating and stabilizing the biomass on a suitable support, as well as its high sensitivity to operating conditions, the rather low reproducibility of the biological process due to biomass changes from one cell cycle to another, and the difficult controllability of the bioreactor (Figure 2-45). Many of these drawbacks can be overcome by an efficient immobilization of the biomass on suitable supports, by using GMOs with superior catalytic activity or/and by optimizing the working conditions and the operating mode of the selected biological reactor, using advanced mathematical models of the bioprocess and bioreactor.

The present section 3 is dealing with this last aspect of the engineering problem, by using original and modern computational techniques, by following the current trend of using *hybrid*, multi-levels dynamic models of SMDHKM kinetic model type (see the below section 3.2) [Maria, 2023; Maria and Luta, 2013; Maria, 2010]. In this way, by solving some important engineering problems with using adequate SMDHKM kinetic models, the industrial biosyntheses become more efficient valuable alternatives for obtaining a wide range of products in the food and pharmaceutical industries, being also used in various medical applications.

In this context, the present work makes a remarkable contribution to solving engineering problems that arise when scaling up new biological processes, knowing that the engineering part (bioreactor/plant design and their optimal operation, all performed *in-silico*, off-line, through engineering calculations based on a mathematical models) is perhaps the most difficult stage in the development of a new bioprocess, and in the operation of an industrial bioreactor. As has been proven in the dedicated literature, *in-silico* numerical analysis (based on mathematical / kinetic models) of biochemical (enzymatic), or biological processes (the present work case) has proven to be not only an essential tool, but also extremely beneficial for the engineering assessments aimed at determining optimal operating policies for biological reactors [Maria, 2020a; Maria and Renea, 2021; DiBiasio, 1989; Maria et al., 2018C; Scoban and Maria, 2016; Maria et al., 2024d, 2025b; Renea and Maria, 2025].

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The current ("default") approach in the bioengineering practice for insilico solving design, optimization and control problems based on mathematical models of the industrial biological reactors is the use of unstructured Monod-type models (for cell culture bioreactors) or Michaelis-Menten-type models (if only enzymatic reactions are considered) that ignore a detailed representations of cellular processes. The applied engineering rules are similar to those used for chemical processes (section 1.3.4) and are inspired by the theory of nonlinear systems control. However, by only considering the macroscopic process key variables (biomass, substrate and product concentrations), these ***unstructured (apparent, global)*** models do not adequately reflect metabolic changes at the biomass level in the bioreactor, being unsuitable for accurately predicting the cellular response to external perturbations of the culture medium through self-regulated cellular metabolism. These *unstructured* models may be satisfactory for an approximate mathematical modelling of the biological processes, but not for math modelling of cellular metabolic processes, as they cannot make any correlation between the operation of the bioreactor and the continuous adaptation of biomass metabolism to the variable conditions in the bioreactor. Also, the accuracy of the predictions of these global kinetic models is quite low, and often unsatisfactory.

The current trend. The present work, and the *hybrid, multi-levels* kinetic models of SMDHKM type, under the WCVV modelling framework proposed by Maria [2010, 2013, 2018, 2023], follows the current trend for solving such engineering problems with an improved precision, namely the use of the novel *hybrid structured, modular, deterministic* (with continuous variables and based on cellular metabolic reaction mechanisms), kinetic models (SMDHKM), derived under the WCVV novel modelling framework. These mathematical models are structured on several levels of detail of the bioprocess, according to the addressed application, which concomitantly consider the following bioprocess aspects:

a). The dynamics of key cellular species involved in the studied bioprocess at the cellular (nano-scopic) level. These concerns the essential modules of the cell central carbon metabolism (CCM) reactions, the GRC controlling the target metabolite sysnthesis (e.g. TRP of [Maria, 2021; Maria and Renea, 2021], or mercury-operon of [Maria, 2010, 2023; Maria and Luta, 2013]);

b). The dynamics of key cellular species involved in any GRC linked to the target metabolite of interest;

c). The dynamics of some key cellular species involved in the analysed bioprocess, should be connected with the dynamics of macro-scopic state variables of the bioreactor. That is because the cell import nutrients/ substrates (e.g. glucose - GLC), and releases into the environment some metabolites, and metabolic waste.

It is to remark that, if the above cell kinetic model is formulated under the novel WCVV approach, all cell species should be considered individually, or lumped to fulfil the istonicity constraint Eq.(6,11-15), (Table 2-1).

These multi-level *hybrid* SMDHKM models can numerically simulate, with a high precision, both the key steps in the cell CCM, at the cellular scale, responsible for the synthesis of cellular metabolites of industrial interest, as well as the dynamics of the bioreactor state variables, at the macro-scopic level.

In this way, more accurate predictions are obtained for both the dynamics of the biological process at the cellular level and for the dynamics of the state variables of the analyzed industrial bioreactor. The immediate application of *hybrid* SMDHKM models is the more accurate determination of the optimal operating regime of an industrial bioreactor. An example is provided in section 3.2, and others by [Maria, 2021; Maria and Renea, 2021].

Thus, the use of complex *hybrid* SMDHKM models, taking into account all key species of the cellular bioprocess, coupled with advanced numerical multi-objective optimization procedures offers the guarantee of obtaining superior performances for the industrial bioreactor. The use of *hybrid* SMDHKM models taking into account multiple control variables and multiple opposing objectives increases the difficulty of the optimization problem, but, by using the Pareto front technique and by comparing a larger number of optimal operating alternatives, significantly superior were obtained compared to the classical approaches to optimize the bioreactor operation by using unstructured (apparent/global) kinetic models [Maria, 2021; Maria and Renea, 2021].

According to the reviews of Maria [2018, 2023], basically, the *hybrid* SMDHKM models taking into account multiple control variables, are valuable computing tools useful to: 1) Easily evaluate the cell metabolic fluxes (that is the biochemical reaction rates under stationary / homeostatic) conditions; 2) derive the flux balance analysis (FBA) aiming to design GMO-s of desired characteristics ('motifs'); 3). Derive numerical simulations of the bioreactor dynamics with a higher precision, and with a higher degree of details (number of considered species in the model); 4) Allow the use of *hybrid* SMDHKM models for the *in-silico*, off-line optimization of the bioreactor operation; 5) Allows to more accurately predict the biomass response to environmental perturbations, and 5) Allows to predict the biomass adaptation to the variable environmental conditions over hundreds of cell cycles.

Maria [2023, 2018] proved the feasibility and the advantages of using the novel WCVV concept to develop *hybrid* SMDHKM models to couple extended cell-scale CCM-based (central carbon metabolism) structured deterministic kinetic models with bioreactor classical dynamic models (including macro-scale state variables). Thus the work of Maria [2023, 2018] presents a holistic 'closed loop' approach for the development of models of biological systems.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The ever-increasing availability of experimental (qualitative and quantitative) information, at the cell metabolism level, but also on the bioreactors' operation necessitates the advancement of a systematic methodology to organise and utilise these data. The resulted *hybrid* SMDHKM dynamic models were proved to successfully solving more accurately difficult bioengineering problems [Maria, 2023]. In such SMDHKM kinetic models, the cell-scale model part (including nanolevel state variables) is linked to the biological reactor macro-scale state variables for improving the both model prediction quality and its validity range. The three examples approached by [Maria, 2023] include development of SMDHKM kinetic models able to:

i). Case study no.1. Simulate the dynamics and optimize the mercury uptake from wastewaters in a semi-continuous fed-batch (FBR) threephase fluidized bioreactor (TPFB) (below section 3.2);

ii). Case study no.2. Simulate the dynamics of a fed-batch bioreactor (FBR) at both cell- and bulk-phase species levels, aiming to maximize the tryptophan (TRP) production [Maria, 2021; Maria and Renea, 2021];

iii). Case study no.3. Optimize the both production of biomass and succinate (SUCC) in a batch bioreactor (BR) [Maria et al., 2011].

In all the above case studies, *in-silico* designed genetically modified (GMO) *E. coli*, or other bacteria cultures have been used. As proved, there are multiple advantages of using extended SMDHKM-s. Thus, in the case study (*no*.1), a higher prediction detailing degree is reported, that is prediction of the dynamics of [26(cell species) + 3(bulk species)] *vs.* only [3 (bulk species)] by a classical macroscopic FBR- TPFB model, while covering a wider range of input [Hg^{2+}] loads, with using cloned *E. coli* cells with various amounts of mercury-plasmids [Gmer]. Also, the SMDHKM model offers the possibility to predict the bacteria metabolism adaptation to environmental changes over dozens of cell cycles, and the effect of cloning cells to modify their behaviour under stationary or perturbed FBR operating conditions. In the case study (*no*.2) the SMDHKM realizes a higher prediction detailing degree, by predicting the dynamics of [11] (cell species) +4 (bulk species)] *vs.* only [3 (bulk species)] by a classical macro-level FBR model, while covering a wider range of control variables, and various GMO *E. coli* cells strains. Eventually this latter SMDHKM model was used to derive the optimal operating policy of a FBR leading to the TRP production maximization. Besides, the SMDHKM flexibility is high enough to consider a larger number of control variables for the studied bioreactors, that is: (i). biomass concentration, the inlet feed flow-rate, the inlet [Hg^{2+}], and the [Gmer] concentration in the used cloned cells for the case study (no.1) and, (ii). the inlet feed flow-rate, and inlet substrate concentration [GLC] for the case study (no.2). (iii). In the case study (no.3), the SMDHKM generates a Pareto-front of optimal operating alternatives of a BR, by using various *in-silico* design *E. coli* mutants. This approach uses the SMDHKM model and a mixed-integer nonlinear programming (MINLP) rule, coupled with an effective adaptive random search to determine the optimal metabolic fluxes of a GMO in respect to multiple economic objectives associated to the gene knockout strategies. Exemplification is made for the case of designing an *E. coli* GMO that realizes maximization of both biomass and succinate production in a **BR** by using an extended structured central carbon metabolism (CCM) model from literature. Comparatively to the linear procedure **LP** that solves a combinatorial problem in a bi-level optimization approach, **LP** of dimensionality sharply increasing with the number of removed genes, the MINLP alternative using the SMDHKM model includes the nonlinear influence of fluxes, and the number of checked knockout genes to the main goals. Besides, in case study (no.3), the SMDHKM realizes a higher prediction detailing degree, by predicting the metabolic fluxes dynamics of [72(cell species), involved in 95 reactions +1 (bulk species, i.e. the biomass)] *vs.* only [1 (bulk species)] by a classical macroscopic BR unstructured model, while covering a wider range of control variables, and GMO *E. coli* strains.

3.2. Case study. Optimization of a fed-batch bioreactor (FBR) with immobilized *E. coli* cells, cloned with mer-plasmids, used for the mercury uptake from wastewaters by using a hybrid SMDHKM dynamic model under WCVV modelling framework

3.2.1. Symbols used in section 3

^{a}G - G-L specific interfacial area

^{a}L - *L-G specific interfacial area (identical to* a_G*)*

a_{s-} - L-S specific interfacial area

A_{j-} - atomic (molecular) mass of species ,,j"

a,b - rate constants in the Hill-type kinetic expression

C_{j-} - species, j" concentration

D_j - difussivity of species 'j' in a certain phase"

D = Dm - cell content dilution rate (i.e. cell-- content dilution rate (i.e. cenvolume logarithmic growing rate)

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

d_{b-} - bubble average diameter

d_{p-} - particle diameter

d_{r-} - reactor diameter

F - feed flow rate

FL - Liquid feed flow-rate

g - gravitational acceleration

$[\ Hg_L^{2+}\ ]$ - Concentration of the mercury ions in the liquid (bulk) phase of the bioreactor

K_m - Michaelis-Menten constants

k_G - G-L mass transfer coefficient (on gas side)

k_H - Henry constant

k_L - L-G mass transfer coefficient (on liquid side)

k_s - L-S mass transfer coefficient (on liquid side)

k - rate constant

n_H - Hill-coefficient

n_{PD}, N_{PR}- - partial orders of reaction

n_{j-} - number of moles of species ,,j"

n_s - number of species in the cell

N_A - Avogadro number

p - overall pressure

p_j - partial pressure of species j

$\mathrm{Re_L} = \left(\sum_L d_p^4 \rho_L^3\right) / \mu_L^3$ - Reynolds number (liquid)

R_g - universal gas constant

r_j - species j reaction rate

$\mathrm{Sc_L} = \mu_L/(\rho_L D_{s,L})$ - Schmidt number (liquid)

$Sh = (k_S d_p) / D_{S,L}$ - Sherwood number

T - temperature

T - time

t_c - cell-cycle time

u_G - gas superficial velocity

u_L - liquid superficial velocity

V - volume

v_m - maximum reaction rate

X - Biomass in the bioreactor

Y_j - molar ratio of species j to the rest of species in the mixture

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Greeks

α,ω - stoichiometric coefficients

β,Ψ - constants used in evaluation of particle effectiveness in Table 3-3

ε_G - volume fraction of the gas in the bed

ε_L - volume fraction of the liquid in the bed

ε_p - particle porosity

ε_s - volume fraction of particles in the bed

Φ - optimisation objective function

ϕ -Thiele modulus

φ_c - Carman shape factor [Trambouze et al., 1988]

η_j - effectiveness factor of reaction j

μ_L - dynamic viscosity of the liquid

ρ - density

π - osmotic pressure

σ - interfacial tension

Σ_Λ - power dissipated per unit mass of liquid

τ_p - particle tortuosity

Superscript

* - saturation

Index

app - apparent

cell - referring to the *E. coli* cell

cyt - cytoplasma

ef - effective

env - environment

G - referring to gas, or at G-L interface

In - inlet

L - referring to liquid, or at the L-G interface

max - maximum

o - initial

p - particle

ref - reference value

s - referring to particle, or at liquid (L) - solid (S) interface, or referring to the steady-state

trans - referring to the trasport

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Abbreviations

BR - batch reactor

CCM - Central carbon metabolism

FBR - Fed- batch reactor

G-L - gas-liquid

G• - Gene (DNA)

GERM - Individual gene expression regulatory module of reactions

Gmer - *mer* plasmids generating *mer* operons into the cell DNA

GmerX (or GX) - *mer* genes (X = R,T,A,D)

GRC - genetic regulatory circuit

SMDHKM - structured, modular, deterministic, *hybrid* kinetic models

L-S - liquid-solid

Mer - mercury

MetG, MetP - Lumps of metabolites used for the G (gene, DNA) and P (protein) synthesis, respectively

M-M - Michaelis Menten

NADPH - Reduced nicotinamide adenine dinucleotide phosphate

NutG, NutP - Lumps of nutrients used for the MetG and MetP synthesis, respectively

P• - protein

PmerX - *mer* proteins expressed by the mercury (mer) operon

QSS - quasi-steady-state (stationary cell homeostasis)

RSH - compounds including thiol redox groups

S - substrate

SCR - semi-continuous reactor

TF - transcription factor

TPFB - three-phase fluidized bioreactor

WCVV - whole-cell variable volume modelling framework of Maria [2002,2005,2017a]

X - biomass

[.] - Concentration"

3.2.2. Section 3.2 purpose

This section 3.2 exemplifies the use of a complex *hybrid* SMDHKM kinetic model to solve an engineering problem at an industrial pilot scale, that is to insert a complex cell WCVV structured kinetic model of the *mer*-operon GRC expression in *E. coli* (Figure 2-6) into a SMDHKM model, in order to optimize a pilot-scale FBR-TPFB bioreactor (Figure 2-46) used for mercury uptake from wastewaters by immobilized *E. coli* cells cloned with mer-plasmids. The developed SMDHKM dynamic model is linking the cell-scale model part (simulating the dynamics of the nano-scale state variables/species) to the biological reactor macro-scale state variables, aiming at improving the both model prediction quality and its validity range. Eventually, the SMDHKM model was used to *in-silico* design a GMO (i.e. an *E. coli* cloned with mer-plasmids in a degree to be determined) for improving its capacity for mercury uptake from wastewaters (Figure 2-47).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The cell SMDHKM dynamic model of the *E. coli* cloned bacterium is able to simulate the self-control of the GRC responsible for the *mer*-operon expression, and to predict:

(a). the influence of the TPFB bioreactor control variables {such as, the feed flow-rate (FL), the mercury ions concentration $\left[Hg_L^{2+}\right]$, in the feeding liquid, and of the biomass concentration in the bioreactor [X]}; (b).- the influence of various bioreactor running parameters { such as, the size of the solid porous particles (dp) of pumice on which the biomass is immobilized; the concentration [Gmer] of the *mer*-plasmides used in the cloned *E. coli* cells} on the bioreactor performance to uptake the mercury ions from wastewaters, and in eliminating them as mercury vapours by the continuously sparged air into the reactor [Maria, 2009b, 2010; Maria and Luta, 2013; Maria et al., 2013; Scoban and Maria, 2016; Maria and Scoban, 2017d].

This *hybrid* SMDHKM dynamic model is also a worthy example of applying WCVV models (section 2.2), and the GERM-s properties (P.I.-s) described in section 2.2.3 to adequately represent a complex modular GRC-s as it is the mer-operon expression in gram-negative bacteria (such as *E. coli* cells). The structured GRC-WCVV model was proposed by Maria [2009b, 2010] to reproduce the dynamics of the *mer* operon expression in Gram-negative bacteria (*E. coli*, *Pseudomonas sp.*) to uptake the mercury ions from wastewaters under various environmental conditions. The model was constructed and validated by using the Philippidis et al. [1991a, 1991b, 1991c] experimental data, and the Barkay et al. [2003] information on the mer-operon expression characteristics.

Later, based on this information, Maria and Luta [2013] included this cellular GRC kinetic model of the *mer*-operon expression in a *hybrid* SMDHKM dynamic model of the experimental FBR-TPFB bioreactor of Deckwer et al. [2004] in order to simulate its dynamics over a wide range of operating conditions, that is: FL = [0.01-0.04] L/min.; $\left[Hg_L^{2+}\right]$in= [10-40] mg/L; [X] = [250 - 1000] mg/L; dp = [1 - 4] mm; [Gmer] = [3- 140] nM.. Comparison with predictions of a global unstructured model was made by Maria et al. [2013].

As evidenced by this application, the current trend in bioengineering is to use multi-layer (hybrid) kinetic models to extend the detailing degree of the developed bioreactor dynamic models, by also including the dynamics of the concerned key-species of the cell metabolism. Exemplification is made by coupling an *un-structured* dynamic model of a FBR-TPFB, used for mercury uptake from wastewaters by immobilized *E. coli* cells on porous pumice granular support (small particles of 1-4 mm), with a structured cell simulator of the GRC controlling the mercury-operon expression, and the mercuric ion reduction in the bacteria cytosol. After its cytosolic reduction, the metallic mercury leaves the cell by membranar diffusion, and then it is entrained (in the form of vapor) by the air continuously bubbled into the bioreactor (Figure 2-46).

Table 3-1: Un-structured (apparent, reduced) kinetic model of Michaelis-Menten type for mercury ions reduction by E. coli (after [Philippidis et al., 1991a-c]). Notations: substrate $S = Hg_L^{2+} = Hg_{env}^{2+}$; **PT** = lumped permease for the membranar transport of Hg_{env}^{2+} into cytosol Hg_{cyt}^{2+} PA = lumped reductase to reduce cytosolic mercury ions Hg_{cyt}^{2+} to metallic mercury Hg_{cyt}^{0} ; RSH = low molecular-mass thiol redox cytosolic buffers; NADPH = (reduced) nicotinamide adenine dinucleotide phosphate; subscripts: 'env' = environmental; 'cyt' = cytosol. Other notations are given in the symbols list of section 3.2.1. Atfer [Maria and Luta, 2013].

Membranar transport of Hg_{env}^{2+}:

$$Hg_{env}^{2+} \xrightarrow[-2H^+]{+2RSH,\ (PT)} Hg_{cyt}^{2+}$$

Rate expression:

$$r_t = \frac{v_{m,t} c_{Hg_{env}^{2+}}}{K_{mt} + c_{Hg_{env}^{2+}}}\ \text{, (membranar transport)}$$

$$v_{m,t} = r_{max,t} A_{Hg} c_X\ ;$$

$$r_{max,t} = k_7 c_{PT}\ \text{(see Table 3-4)}$$

Parameters:

[Gmer], nM	$r_{max,t}$, nM min^{-1} mgX^{-1}	K_{mt}, nM
3	8.2	3800
67	13.4	4700
78	17.5	3600
124	20.6	5100
140	19.8	6000

Mercury ion reduction:

$$Hg(SR)_2 \xrightarrow[-2NADP^+]{+2NADPH,\ (PA)} Hg_{cyt}^{0} + 2RSH$$

Rate expression:

$$r_P = \frac{v_{m,P} c_{Hg_{cyt}^{2+}}}{K_{mP} + c_{Hg_{cyt}^{2+}} + c^2_{Hg_{cyt}^{2+}} / K_{iP}}\ ,$$

(reduction in cytosol)

$$v_{m,P} = r_{max,P} A_{Hg} c_X\ ;$$

$$r_{max,P} = k_8 c_{PA}\ \text{(see Table 3-4)}$$

Parameters:

[Gmer], nM	$r_{max,P}$, nM min^{-1} mgX^{-1}	K_{mP}, nM	K_{iP}, nM
3	45	12600	96700
67	152	15900	91500
78	168	22800	87400
124	221	15500	75200
140	305	19500	93100

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The obtained results reported a significant improvement in the model prediction quality (ca. 3-12% in state variables, and up to 40% in reduction rate *vs.* experimental information) and in the detailing degree [i.e. simulation of 26 + 3 (cell+bulk) *vs.* only 3 (bulk) variable dynamics]. The major advantages of the hybrid model come from the possibility to predict the bacteria metabolism adaptation to environmental changes over several cell generations, and also the effect of cloning cells with certain plasmids to modify its behaviour under stationary or perturbed conditions.

Basically, this section 3.2 exemplifies the possibility to couple an unstructured TPFB dynamic model with macro-scale state variables [Deckwer et al., 2004] used for mercury uptake by *E. coli* cells immobilized on pumice millimetric size support, with a structured GRC cell model simulating the mercury(*mer*)-operon expression. The advantages of the *hybrid* SMDHKM model developed by Maria [2009b, 2010], and extended by Maria and Luta [2013] are related to: i) the improvement of the prediction accuracy, and of the degree of detail (number of considered species), of the reactor performance / state variable dynamics; ii) the prediction of the bacteria metabolism adaptation to environmental 'step'-like changes in the environmental mercury content $\left[Hg_L^{2+}\right]_{env}$ through the modelled cell GRC related to the *mer*-operon expression, and by mimicking the whole-cell growth under balanced conditions. If promising, such an approach can support the idea of (i) improving the quality of process monitoring (control), that is optimize the bioreactor operation, and (ii) *in-silico* design cloned *E. coli* with an increased content of *mer*-plasmids to maximize the FBR-TPFB reactor performances in uptake the mercury ions from the waste-waters, without exhausting the cell resources. And all these by using complex *hybrid* SMDHKM dynamic model of increased predictive power. The investigation is also supported by the tremendous improvement in the computing power over the last decades, and by the continuous expansion of the available information from cellular *bio-omics* databanks, in spite of steady efforts necessary to elaborate detailed cellular numerical simulators.

Exemplifications of such modular GRC models used for the insilico design of GMO-s of industrial use includes several published case studies by Maria [2018, 2023] (Figure 2-49). Due to the cell metabolism complexity, and existence of control variables at both cell-level, together with the bioreactor macro-level, *in-silico* optimization of an industrial bioprocess by using GMO-s often translated in a multi-objective optimization problem [Maria, 2021; Maria and Renea, 2021; Maria and Luta, 2013; Maria et al., 2011; Maria et al., 2018b; Maria, 2014b], difficult to be solved by using common numerical algorithms. A couple of case studies exemplify the own positive experience with applying *hybrid* SMDHKM dynamic models, also including the cell structured models of CCM, and of various GRC-s, used for optimization of the industrial bioreactors operation, or for design GMO-s for improving certain bioprocesses of practical interest. See examples of [Maria, 2021; Maria and Renea, 2021; Maria, 2018; Maria and Luta, 2013; Maria et al., 2011; Maria et al., 2018c; Maria et al., 2018a].

3.2.3. Mercury ion reduction in bacteria cells - the apparent (unstructured) kinetics

Bacteria resistance to mercury is one of the most studied metallic ion uptake and release process (see the review of Barkay et al. [2003]) due to its immediate large-scale application for mercury removal from industrial wastewaters [Deckwer et al., 2004; Wagner-Döbler et al., 2000; Leonhäuser et al. 2006]. The bacteria response to the presence of toxic mercuric ions in the environment Hg_{env}^{2+} is apparently surprising. Instead of building carbon- and energy-intensive disposal '*devices*' into the cell (like chelate-compounds) to '*neutralize*' the cytosolic mercury Hg_{cyt}^{2+} and thus maintaining a tolerable level, a simpler and more efficient defending system is used. The metallic ions Hg_{cyt}^{2+} imported into cell cytosol are catalytically reduced to the volatile metal Hg_{cyt}^{0}, less toxic and easily removable outside the cell (as $Hg_L^{\ 0}$ in the liquid environment of the TPFB) by simple cell membranar diffusion. Such a process involves less cell resources and is favoured by the large content (millimolar concentrations) of low molecular-mass thiol redox buffers (RSH) able to bond and transport Hg_{cyt}^{2+} in cytosol as Hg(SR)2, and of NAD(P)H reductants able to convert it into neutral metal Hg_{cyt}^{0}. (see the overall reactions of Table 3-1). A genetic regulatory circuit (GRC) responsible for the involved mer-operon expression controls the whole process, by including seven genes of individual expression levels of 5 encoded proteins, of which expression is induced and adjusted according to the level of mercury Hg(SR)2, and other metabolites into cytosol (Figure 2-6). The whole process is tightly cross- and self-regulated to hinder the import of large amounts of mercury into the cell, which eventually might lead to the blockage of cell resources (RSH, NADPH, key-metabolites, and key-proteins), thus compromising the whole cell metabolism.

While the role of each *mer*-genes and *mer*-proteins in the mercury ion reduction process is generally known, not all the regulatory loops of the *mer*-operon expression are perfectly understood, and also the way by which the cell adapts itself to variations of mercuric ion concentrations Hg_{env}^{2+} in the environment. Philippidis et al. [1991a-c] proposed a reduced apparent (*un-structured*) kinetic model, of Michaelis-Menten (M-M) type, to quickly simulate the global (apparent) mercury uptake process by *E. coli*, that is: the membranar transport of environmental Hg_{env}^{2+} into the cell (of r_t rate), and its reduction (^rP rate, see Table 3-1). To highlight the slowest process step, separate experiments have been conducted with cultures of intact cells or '*permeabilized*' cells (with a more permeable cell membrane to metallic ionic species). The results clearly

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 3-2: Nominal operating conditions and characteristics of the fed-batch (FBR) three-phase fluidized bed reactor (TPFB) used for mercuric ions uptake by the immobilized E. coli cells on pumice porous support (Footnote a)[Deckwer et al., 2004; Maria, 2010; Maria and Luta, 2013]. Notations are given in section 3.2.1.

Reactor characteristics [Deckwer et al., 2004]	
Liquid volume	V_L = 1 L
Height / diameter (d_r)	0.2 m / 0.08 m
Volumetric fraction of particles in the fluidized-bed	ε_s = 2 g L^{-1} (0.8 [Deckwer et al., 2004]; up to 3-4 g L^{-1} [Berne and Cordonnier, 1995])
Nominal operating conditions [Deckwer et al., 2004; Maria, 2010]	
Temperature; pressure; pH	37 °C ; 1 atm; pH = 7.2
Liquid inlet/outlet flow rate	F_L = 0.02 L min^{-1}
Liquid inlet concentrations [Hg^{2+}, Hg^0] (polluted water)	$[c_{Hg^{2+}_{env}}, c_{Hg^{o}_{env}}]_{in}$ = [20, 0] mg L^{-1}
Liquid background pollution [Hg^{2+}, Hg^0]	$[c_{Hg^{2+}_{env}}, c_{Hg^{o}_{env}}]$ = [0.02, 0] mg L^{-1}
Gas (air) inlet/outlet flow rate	F_G = 0.2748 moles min^{-1}
Gas inlet concentration of Hg^0	0 mg L^{-1}
Gas (air) superficial velocity	u_G = 0.0233 m s^{-1}
Pumice particle diffusion characteristics [Wagner-Döbler et al., 2000; Whithan and Sparks, 1986]	
Particle specific density (dry)	ρ_s = 2400 kg m^{-3}
Particle average diameter	d_p = 1 mm
Catalyst porosity / BET area	ε_p = 0.8 / S_{BET} = 0.5 m^2 g^{-1}
Apparent G-L mass transfer coefficient (on gas side)(**Footnote b**)	$k_G a_G$ = 0.617 s^{-1}
Apparent G-L mass transfer coefficient (on liquid side)	$k_L a_L$ = 0.022 s^{-1} [Deckwer et al., 2004]
Apparent L-S mass transfer coefficient (on liquid side)	$k_s a_s$ = 0.0114 s^{-1} [Deckwer et al., 2004]
Catalyst tortuosity (**Footnote c**)	τ_p = 1.42
Hg^{2+} effective diffusivity in support (experimental)	D_{ef} = 4.5 · 10^{-10} m^2 s^{-1} [Deckwer et al., 2004]
Hg^{2+} diffusivity in air	D_G = 0.1259 · 10^{-4} m^2 s^{-1} [Lambert, 1996]
Henry constant for Hg^{2+}	He = exp(8.54 - 4150/T)/101.3 atm (mg/L)$^{-1}$ [Deckwer et al., 2004]

(a) fluid physical properties correspond to those of water and air at the operating conditions.
(b) see calculation rule of Table 3-3.
(c) particle tortuosity was evaluated by using [Shen and Chen, 2007] formula $\tau_p^2 = \varepsilon_p / (1 - \sqrt[3]{1 - \varepsilon_p})$

showed that membranar permeation is the rate controlling step, being of one order of magnitude slower than the cytosolic mercuric ion reduction. Identification of rate constants of the two main reactions for cloned *E. coli* cells with an increasingly copynumbers of mer-plasmids (of concentration denoted by [Gmer]), in the range of [Gmer] = 3-140nM, comparatively to [Gmer] = 1- 2nM (for wild-types of *E. coli*) reveals the following aspects (Table 3-1):

(i). The rate constants are strongly dependent on the *mer*-plasmids (genes) level into the cloned *E. coli* cells, the reaction mechanism being more complex than those suggested by the two apparent reaction rates (r_t and *rP*) of Philippidis et al. [1991a-c];

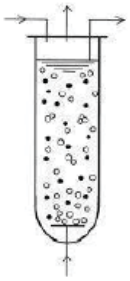
(ii). Such reduced kinetic models can only represent very approximately the overall mercury uptake by the cells, and the steady-state process efficiency, being useful for quick bioprocess scale-up, and rough engineering calculations [Deckwer et al., 2004];

(iii). The *un-structured* models (of M-M type) can not represent the *mer* operon (that is the associated GRC) response to various environmental inducers, or the cell response to stationary or dynamic perturbations in the mercury level ($Hg^{2+}_L = Hg^{2+}_{env}$) in the bioreactor liquid-phase, or the cell response to variations in the mercury concentrations Hg_L^{2+} in the FBR-TPBR feeding. Also, the apparent M-M kinetic model can not explain and simulate/predict the selfregulation of the whole membrabar transport - reduction process, and how the *mer*-gene expression is connected to the cell volume growth and the cell content replication;

(iv). The cellular uptake process may be improved by increasing the mer-permease content into the cell, up to a limit of ca. [Gmer] = 80nM mer-plasmids [Philippidis et al., 1991a-c], higher than the permeation rate remains unchanged, thus preventing exhaustion of the cell metabolic resources, and its death. From the same reason, the cell regulatory system maintains an upper limit for the membranar transport of environmental Hg^{2+}_{env} into the cell (r_t) irrespectively to th Hg^{2+}_{env} concentration in the environment, the cytosolic Hg^{2+}_{cyt} concentration never exceeding a 3,500-5,00nM) level, due to the mentioned reasons.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 3-3: Three-phase fluidized bed (TPFB) reactor model of [Maria, 2009b, 2010, 2012]; [Doran, 1995; Maria and Luta, 2013; Moser, 1988; Trambouze et al., 1988]. Index „s" = concentration at the particle surface (considered identical to those in the bulk-phase). Notations are given in section 3.2.1.

Mercury mass balance equations	**Observations**
- liquid phase: $\frac{dc_j}{dt}=\frac{F_L}{V_L}(c_{j,in}-c_j)+\alpha r_{app}+\omega r_{trans}$; species $j=$ Hg_L^{2+}, Hg_L^0 ; $\alpha=\begin{cases}-1 \text{ for } j=Hg_L^{2+}\\ 1 \text{ for } j=Hg_L^0\end{cases}$; $\omega=\begin{cases}0 \text{ for } j=Hg_L^{2+}\\ -1 \text{ for } j=Hg_L^0\end{cases}$ - gas phase: $\frac{dc_{j,G}}{dt}=\frac{F_G}{V_G}A_j(Y_{j,in}-Y_j)+\frac{V_L}{V_G}r_{j,trans}$; species $j=Hg_L^0$ - apparent $r_{j,app}$ reaction rate is evaluated from equality of L-S mass transfer and surface reaction rate (at quasi-steady-state): $r_{j,app}=k_s a_s(c_j-c_{j,s})=\eta_j r_j(c_{j,s})$ For Michaelis-Menten rate expressions $r=\frac{v_m c_S}{K_m+c_S}$, the solution r_{app} results from solving the equation: $r_{app}^2-r_{app}(\eta v_m+c_{S,s}k_s a_s+K_m k_s a_s)+c_{S,s}k_s a_s \eta v_m=0$, with the constraint: $0\le r_{app}\le r_{max}=\frac{v_m c_{S,s}}{K_m+c_{S,s}}$; (usually $c_{S,s}\approx c_S$): In the present case, $r=r_1$ (Table 3-1; substrate S = $Hg_L^{2+}=Hg_{env}^{2+}$): - apparent r_{trans} transport rate of Hg_L^0 (indexed by 'Hg0') is evaluated from equality of G-L transfer rate at gas interface (at quasi-steady-state): $r_{trans}=k_G a_G(c^*_{Hg0,L}k_H-p_{Hg0,G})=k_L a_L(c^*_{Hg0,L}-c_{Hg0,L})$, that is: $r_{trans}=(c_{Hg0,L}-p_{Hg0,G}/k_H)/[1/(k_G a_G k_H)+1/k_L a_L]$.	 - at $t=0$, $c_j=c_{j,o}$; - hypotheses: dynamic operation, isothermal, iso-pH, perfectly mixed liquid and gas phases, of constant volume; immobilized cell on porous pumice; uniform concentration of solid and biomass in the reactor (by purging).
Model parameters	**Observations**
- biocatalyst effectiveness factors are calculated for every reaction j with the relationships for spherical particles [Doran, 1995] $\eta_j=\frac{\eta_{jo}+\beta\eta_{j1}}{1+\beta}$, $\beta=\frac{K_m}{c_{S,s}}$, $\eta_{jo}=\begin{cases}1 & \text{,for } 0\le\phi_{jo}\le 0.577\\ 1-\left[\frac{1}{2}+\cos\left(\frac{\Psi+4\pi}{3}\right)\right]^3 & \text{for } \phi_{jo}>0.577\end{cases}$ $\eta_{j1}=\frac{1}{3\phi_{j1}^2}(3\phi_{j1}\coth(3\phi_{j1})-1)$.	$\Psi=\arccos\left(\frac{2}{3\phi_{jo}^2}-1\right)$ Thiele moduli are: $\phi_{jo}=\frac{d_p}{6\sqrt{2}}\sqrt{\frac{v_{m,j}}{D_{S,ef}c_{S,s}}}$ $\phi_{j1}=\frac{d_p}{6}\sqrt{\frac{v_{m,j}}{D_{S,ef}K_m}}$ $D_{S,e}=\frac{\varepsilon_p D_{S,L}}{\tau_p}$
- $k_s a_s$ is evaluated by using the relationships [Trambouze et al., 1988] $k_s=Sh\times D_{S,L}/d_p$; $a_s=(6\varepsilon_s)/d_p$	$\Sigma_L=g(u_L+u_G)/\varepsilon_L$ $\varepsilon_L=1-\varepsilon_s$ $u_L=F_L/(\pi d_r^2/4)$
$Sh=(2+0.4\,Re_L^{0.25}\,Sc_L^{0.33})\varphi_C$; $Re_L=\frac{\Sigma_L d_p^4\rho_L^3}{\mu_L^3}$; $Sc_L=\frac{\mu_L}{\rho_L D_{S,L}}$;	$\varphi_C\approx 1$
Experimental $k_s a_s = 0.0114\ s^{-1}$ [Deckwer et al., 2004]	
- $k_L a_L = 0.022\ s^{-1}$ (experimental; [Deckwer et al., 2004]) - $k_L a_L = \text{const.}\times d_r^{0.17}u_G^{0.7}$ [Alvarez et al., 2002]	
- $k_G=6.6D_G/d_b$ (for Hg_G^0)(Shanna's formula; [Trambouze et al., 1988]) - $a_G=6\varepsilon_G/d_b$ (spherical bubbles); - $d_b=26d_t\left(gd_t^2\rho_L/\sigma\right)^{-0.5}\left(gd_t^2\rho_L^2/\mu_L^2\right)^{-0.12}\left(u_G/\sqrt{gd_t}\right)^{-0.12}$ - ε_G is averaged over three relationships Trambouze et al., 1988) $\frac{\varepsilon_G}{(1+\varepsilon_G)^4}=0.14\left(\frac{u_G\mu_L}{\sigma}\right)\left(\frac{\rho_L\sigma^3}{g\mu_L^4}\right)^{7/24}\left(\frac{\rho_L}{\rho_L-\rho_G}\right)^{15/24}\left(\frac{\rho_L}{\rho_G}\right)^{5/72}$, (Mersmann) $\frac{\varepsilon_G}{(1+\varepsilon_G)^4}=0.25\left(\frac{u_G\mu_L}{\sigma}\right)\left(\frac{\rho_L\sigma^3}{g\mu_L^4}\right)^{7/24}$, (Yoshida) $\varepsilon_G=\frac{u_G}{30+2u_G}$, (u_G in cm s^{-1}; Hughmark) - $\varepsilon_G=V_G/(V_L+V_G)$; $k_G a_G\left[\frac{1}{s}\frac{1}{atm}\frac{g-Hg}{L-lq}\right]=k_G a_G\left[\frac{1}{s}\right]\frac{\varepsilon_G}{\varepsilon_L}\frac{A_{Hg}}{R_g T}$.	
- $k_H=\exp(8.54-\frac{4150}{T})/101.3$, atm L mg-1 (for Hg^0; [Deckwer et al., 2004]) - $k_H=p^*_{Hg}/c^*_{Hg0,L}$; $Y_j=y_j/(1-y_j)$; $y_j=p_j/p$; $p_j=(c_{j,G}/A_j)R_g T$, (for gas).	

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

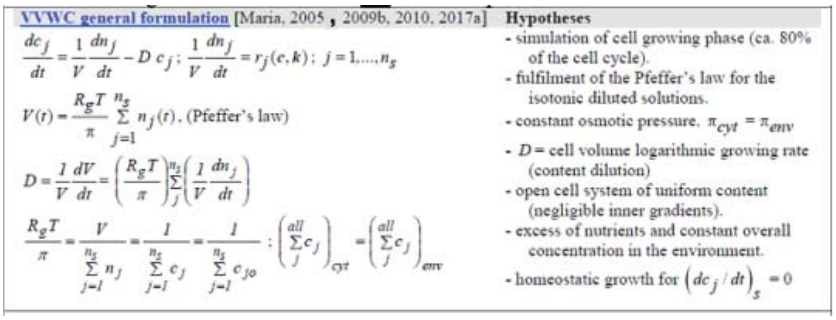
To model and explain these unsolved aspects, an extended structured cell model has been developed by Maria [2009b, 2010], of a WCVV type (see section 2.2) to simulate the mer-operon expression in the "wild", or modified (cloned) *E. coli*. The GRC dynamic model includes seven GERM-s linked by following the rules described in section 2.2.4, and by accounting the few experimental information of Philippidis et al. [1991a-c], and of Barkay et al. [2003]. The derived GRC is able to simulate the *mer*-operon expression, and the process self-control at a molecular level under isothermal and isotonic conditions.

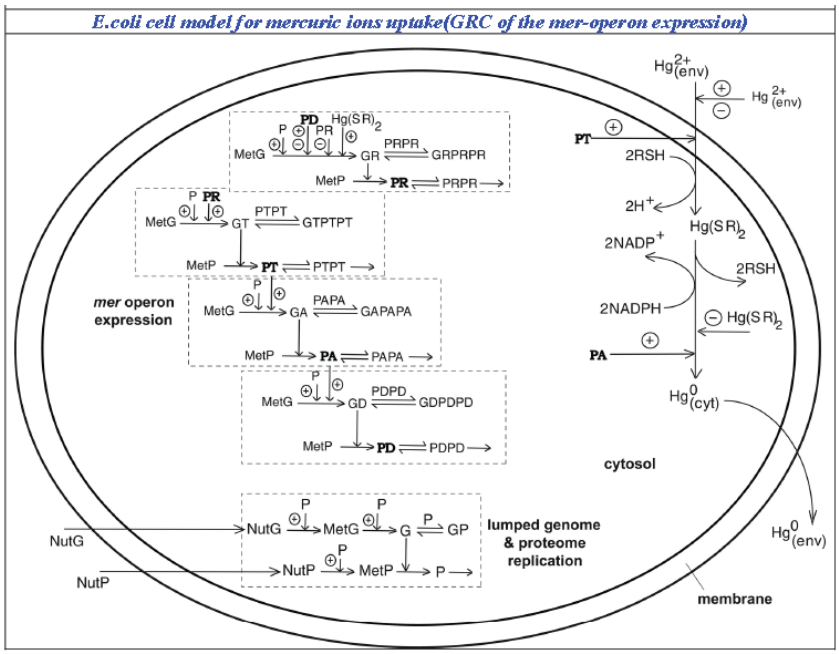
3.2.4. The TPFB bioreactor model and its nominal operating conditions

To exemplify the construction of the *hybrid* SMDHKM dynamic structured model for the analysed mercury removal process, the TPFB bioreactor of Deckwer et al. [2004] was approached. The main characteristics and the nominal operating conditions of this TPFB bioreactor under a semi-continuous (FBR) operation are presented in the Table 3-2. As a first step, in the engineering analysis of this TPFB bioreactor, it is important to emphasize, by using the model-based numerical simulations, its mercury removal performances under various operating conditions. Thus, it resulted that the most influential factors on the mercury uptake efficiency are ranked to be used for the further operating decisions.

The used lab-scale TPFB bioreactor of (Figure 2-46, and Figure 2-47, up-right) includes a resistant *E. coli* cell culture. The bioreactor is completely automated being able to maintain its control parameters of (Table 3-2) at their optimal set-point, by ensuring a constant pH , temperature, a constant feed flow-rate, and inlet mercury concentration [Hg_L^{2+}]in, a constant sparkling air feed flow-rate, and a constant concentration of nutrients used as C/N/P source for the biomass optimal growth. Initially, to study this bioprocess, Deckwer et al. [2004] used *E. coli* bacteria immobilized on alginate beads, but further tests have been extended by using porous pumice granules of 0.9 mm to 4 mm diameter. The pumice carrier checked in the present paper is particularly attractive, the carrier exhibiting a high BET area and porosity, a large pore size (even higher than 10 μm), thus allowing a good diffusion of the substrate (mercuric ions) to

Table 3-4: The structured variable cell-volume whole-cell (WCVV) model proposed for the reduction of cytosolic mercury ions Hg_{cyt}^{2+} to metallic mercury Hg_{cyt}^{0} in cloned E. coli cells. Adapted from [Maria, 2009b, 2010; Maria and Luta, 2013]. Notations are given in section 3.2.1. TF= transcription factor.





ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 3-4a: ●Seven **GERM**-s: two modules for mediated transport of [Hg2+]env into cytosol (catalysed by **PT** lump) and its reduction (catalyzed by **PA**); four **GERM**-s for ***mer***-operon expression including successive synthesis of **PR** (transcriptional activator of other protein synthesis), lumped **PT** permease, **PA** reductase, and the control protein **PD**; one **GERM** for lumped proteome **P** and lumped genome **G** replication control;
● This **GRC** including 7 **GERM**-s for mer-operon expression is *in-silico* "placed" in a simulated growing *E.coli* cell to mimic the cell homeostasis and its response to stationary and dynamic perturbations;
● NADPH and RSH are considered in excess in the *E. coli* cell, of ca. 140 μM NADPH [Philippidis et al., 1991a-c];
● isothermal, iso-pH, homogeneous bulk liquid.

Reaction	Rate expression (Footnote a)	Estimated model parameters (units in min and nM) (Footnote b)	Remarks
Genome and proteome replication			
$NutG \xrightarrow{P} MetG$	$k_1 c_{NutG} c_P$	1.5489×10^{-9}	
$NutP \xrightarrow{P} MetP$	$k_2 c_{NutP} c_P$	2.3875×10^{-9}	
$MetG \xrightarrow{P} G$	$k_3 c_{MetG} c_P$	5.1710×10^{-13}	(Footnote g)
$MetP \xrightarrow{G} P$	$k_4 c_{MetP} c_G$	3.4237×10^{-7}	
$G + P \rightarrow GP$	$k_5 c_G c_P$	$\sim10^{-2}$	
$GP \rightarrow G + P$	$k_6 c_{GP}$	10^{5}	(Footnote c)
Mercury ion import into cytosol			
$Hg^{2+}_{env} \xrightarrow[-2H^+]{+2RSH,\ (PT)} Hg^{2+}_{cyt}$	$\dfrac{k_7 c_{PT} c_{Hg^{2+}_{env}}}{K_{mt} + c_{Hg^{2+}_{env}}}$	$k_7 = 2.9052\times10^{3}$ $[PT]_{s,o} = 2823$ nM $r_{max,t} = k_7 c_{PT}$	$r_{max,t}$, K_{mt} of Table 3-1 [Philippidis et al., 1991a]
Mercury ion reduction			
$Hg(SR)_2 \xrightarrow[-2NADP^+]{+2NADPH,\ (PA)} \rightarrow Hg^0_{cyt} + 2RSH$	$\dfrac{k_8 c_{PA} c_{Hg^{2+}_{cyt}}}{K_{mP} + c_{Hg^{2+}_{cyt}} + c^2_{Hg^{2+}_{cyt}} / K_{iP}}$	$k_8 = 1.2016\times10^{5}$ $[PA]_{s,o} = 1800$ nM $r_{max,P} = k_8 c_{PA}$	$r_{max,P}$, K_{mP}, K_{iP} of Table 3-1 [Philippidis et al., 1991a]
Volatile mercury Hg^0 removal			
$Hg^0_{cyt} \rightarrow Hg^0_{env}$	$k_9 c_{Hg^0_{cyt}}$	1.0105×10^{4}	
Gene GR expression module			
$MetG \xrightarrow{Hg^{2+}_{cyt},\ P,\ PD,\ PR} GR$	$K_{GR} = c^{n_H}_{Hg2cy,ref}$	$k_{10} = 4.5821\times10^{-11}$ $b = 2a^4$ a = variable (Footnote d)	$c_{Hg2cy,ref}$ = avg. experimental value (Philippidis et al., 1991c) n_{PR} of [Voit, 2005];
$k_{10}\, c_P\, c_{MetG}\, c^{n_{PD}}_{PD}\, c^{n_{PR}}_{PR} \times \dfrac{\left(1 + b\,(c_{Hg^{2+}_{cyt}} / c_{Hg2cy,ref})^{n_H}\right)}{(K_{GR} + c^{n_H}_{Hg^{2+}_{cyt}})}$		$c_{Hg2cy,ref} = 1866$ $n_{PD} = 1$; $n_{PR} = -0.5$; $n_H = 2$	n_H = 2-4 similar to genetic switch induction of [Laurent et al., 2005; Gonze, 2010]
$MetP \xrightarrow{GR} PR$	$k_{11} c_{MetP} c_{GR}$	3.3227×10^{-10}	
$GR + PRPR \rightleftarrows GRPRPR$	$k_{12} c_{GR} c_{PRPR}$	2.5000×10^{4}	(Footnote f)
	$k_{13} c_{GRPRPR}$	10^{5}	(Footnote c)
$2PR \rightleftarrows PRPR$	$k_{14} c^2_{PR}$	4.0000×10^{3}	(Footnote e)
	$k_{15} c_{PRPR}$	10^{5}	(Footnote c)

Table 3-4-b. (continued).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Gene GT expression module			
$MetG \xrightarrow{P,\ PR} GT$	$k_{16}{}^{c}MetG\,{}^{c}P\,{}^{c}PR$	1.0725×10^{-16}	(Footnote g)
$MetP \xrightarrow{GT} PT$	$k_{17}{}^{c}MetP\,{}^{c}GT$	4.8253×10^{-9}	
$GT + PTPT \rightleftharpoons GTPTPT$	$k_{18}{}^{c}GT\,{}^{c}PTPT$	2.5000×10^{4}	(Footnote f)
	$k_{19}{}^{c}GTPTPT$	10^{5}	(Footnote c)
$2PT \rightleftharpoons PTPT$	$k_{20}{}^{c}PT^{2}$	5.0210×10^{-2}	(Footnote e)
	$k_{21}{}^{c}PTPT$	10^{5}	(Footnote c)

Gene GA expression module			
$MetG \xrightarrow{P,\ PT} GA$	$k_{22}{}^{c}MetG\,{}^{c}P\,{}^{c}PT$	3.7998×10^{-18}	(Footnote g)
$MetP \xrightarrow{GA} PA$	$k_{23}{}^{c}MetP\,{}^{c}GA$	3.1379×10^{-9}	
$GA + PAPA \rightleftharpoons GAPAPA$	$k_{24}{}^{c}GA\,{}^{c}PAPA$	2.5000×10^{4}	(Footnote f)
	$k_{25}{}^{c}GAPAPA$	10^{5}	(Footnote c)
$2PA \rightleftharpoons PAPA$	$k_{26}{}^{c}PA^{2}$	1.2345×10^{-1}	(Footnote e)
	$k_{27}{}^{c}PAPA$	10^{5}	(Footnote c)

Gene GD expression module			

Table 3-4-c. (continued).

$MetG \xrightarrow{P,\ PA} GA$	$k_{28}{}^{c}MetG\,{}^{c}P\,{}^{c}PA$	5.9584×10^{-18}	(Footnote g)
$MetP \xrightarrow{GD} PD$	$k_{29}{}^{c}MetP\,{}^{c}GD$	3.3227×10^{-10}	
$GD + PDPD \rightleftharpoons GDPDPD$	$k_{30}{}^{c}GD\,{}^{c}PDPD$	2.5000×10^{4}	(Footnote f)
	$k_{31}{}^{c}GDPDPD$	10^{5}	(Footnote c)
$2PD \rightleftharpoons PDPD$	$k_{32}{}^{c}PD^{2}$	40	(Footnote e)
	$k_{33}{}^{c}PDPD$	10^{5}	(Footnote c)

Footnotes:

(a) variable volume formulation of reaction rates $(dn_j/dt)/V$.

(b) identified for Table 3-2 conditions, and [Gmer] = 140 nM.

(c) adopted value of ca. $(10^{6}\text{-}10^{7})\,D_S$ [Klipp et al. 2009; Morgan et al., 2004; Maria, 2005b, 2009b]

(d) the following linear dependence was fitted from data:

$a = a_{\min} + (a_{\max} - a_{\min})(c_{Hg,in} - c_{Hg,in}^{\min}) / (c_{Hg,in}^{\max} - c_{Hg,in}^{\min})$, with $a_{\min} = 2.6$, $a_{\max} = 4.1$, and $a = 3.0$ for [Gmer] > 124 nM. For the outside of the correlation range of $[c_{Hg,in}^{\min}, c_{Hg,in}^{\max}] = [10,100]$ mg L^{-1}, the value $a = 1.5$ should be adopted.

(e) the reaction rate is of ca. 10^{4} 1/nM/min, being similar to the TF (monomer repressor) dimerization with a rate constant of ca. 10^{2} 1/nM/min [Salis and Kaznessis, 2005]

(f) the reaction rate is of ca. 10^{4} 1/nM/min , being similar to the TF binding to the gene operator, with a rate constant of ca. 10^{2} 1/nM/min [Salis and Kaznessis, 2005]

(g) the reaction rate is of ca. $10^{1}\text{-}10^{2}$ 1/nM^{2}/min , being similar to the mRNA (genes) synthesis reactions, of rate constants of ca. 10^{-4} 1/nM/min [Wang et al., 2010]

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 3-5: E. coli cell characteristics and the considered cell key-species stationary concentrations at homeostasis. After [EcoCyc, 2005; Maria, 2009b, 2010; Maria and Luta, 2013].

Monoculture cell:	
Cell cycle time [Trun and Gottesman, 1990]	t_c = 30 min (wild); up to 4-6 h (mutants)
Born cell volume [Kubitschek, 1990]	$V_{cyt,o} = 0.6 \times 10^{-15}$ L
Cell content average dilution rate	$D_s = \ln(2)/t_c = 1.3863$ 1/h
Cell concentration in biomass	$c_{cell} = 2 \cdot 10^{12}$ cells/gX (adjustable; adopted $5 \cdot 10^{12}$ cells/gX by [Philippidis et al., 1991a]
Biomass concentration in the bioreactor	c_X = 1 g/L (ca. 0.6-3 g/L: [Deckwer et al., 2004; Berne and Cordonnier, 1995]
Cell density in the culture medium [Mackay, 2001]	ρ_{cell} = 1 kg/(L cell)

Species concentrations in nM (referring to the cell volume; see the model of **Table 3-4**)(**Footnote a**)	**Reference value** (units in nM)	**Simulated homeostasis for** inlet $[Hg^{2+}]$ = 20 mg/L (units in nM)	**Observations**
Mercury in cytoplasm, $[Hg^{2+}_{cyt}]$	avg. 1866 0-3430	212.7	[Philippidis et al., 1991c]
[NutG] = Lumped nutrients used for the lumped genome **G** synthesis, (ref. to the environmental volume)	$3 \cdot 10^7$	$3 \cdot 10^7$	[Morgan et al., 2004]
[NutP] = Lumped nutrients used for the lumped proteome **P** synthesis, (ref. to the environmental volume)	$3 \cdot 10^8$	$3 \cdot 10^8$	[Morgan et al., 2004]
[MetP] = Lumped metabolites used for the lumped proteome **P** synthesis	$3 \cdot 10^8$	$3 \cdot 10^8$	[Morgan et al., 2004]
[MetG] = Lumped metabolites used for the lumped genome **G** synthesis	$2.01 \cdot 10^7$	$2.01 \cdot 10^7$	**Footnote (b)**
Lumped genome, [G] (active) = [GP](inactive)	4500/2	4500/2	**Footnotes (c,d)**
Lumped proteome, [P]	10^7	10^7	**Footnote (e)**
***Mer*-operon plasmid concentration [Gmer]**	1-2 (wild) 3-140 (cloned)	3 (cloned)	[Philippidis et al., 1991b-c]
Genes expressing the ***mer-*** proteins, [GR], [GT], [GA], [GD]	[Gmer]/3	[4.8,16.4,25.6,31.0]	[Maria, 2010]
Catalytic inactive forms of *mer* genes, [GRPRPR], [GTPTPT], [GAPAPA], [GDPDPD]	[Gmer]/3	[0.2, 2.1, 8.2, 22.0]	[Maria, 2010]; **Footnotes (d,f)**
Proteins initiating and controlling the *mer* operon expression, [PR], [PD]	100	[19.9,84.2]	Fitted from data [Maria, 2010]
Lumped permease PT= PmerT + PmerP + PmerC	500-7000 (avg. 2800)	1021.8	**Footnote (g)**

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Table 3-5-a. (continued).

for the membranar transport of Hg^{2+}_{env}			
[PA] = Reductase of the cytosolic mercury ions Hg^{2+}_{cyt}	500-4000 (avg. 1800)	1022.0	Fitted from data; Footnote (h)
[PXPX] with X= R, T, A, D Proteic dimers of *mer*-proteins with the role of TF-s in the *mer*-gene expression.	4 (1-100)	4	Footnote (i)

Footnotes:

(a) Inner cell concentrations are evaluated with the formula of [Maria, 2005b]: c_j = (copynumbers of species "*j*" per cell) / (N_A V_{cyt}), where $N_A = 6.022 \times 10^{23}$ is the Avogadro number, V_{cyt} = average volume of cell (ca. $1.5 V_{cyt,o}$). Born cell volume $V_{cyt,o} = 0.6 \times \cdot 10^{-15}$ L.

(b) calculated from the state-law constraint for an isotonic and isothermal cell system (subscript 'env' refers to the bulk liquid phase environment):

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

(c) The considered K-12 strain of *E. coli* genome includes ca. 4500 genes [EcoCyc, 2005]

(d) The maximum regulatory effectiveness of expression takes place for equal active and inactive G-forms at steady-state [Maria, 2005b, 2006, 2007], i.e. $[G_j]_s = [G_j TF_j]_s$, where TF denotes the transcription factor adjusting the gene G_j activity. Index "s" denotes the quasi-steady-state (QSS)

(e) *E. coli* genome includes ca. 1000 ribosomal proteins of 1000-10000 copies, ca. 3500 non-ribosomal proteins of avg. 100 copies, and ca. 4500 polypeptides of avg. 100 copies [EcoCyc, 2005]

(f) As reported by Philippidis et al. [1991a], the level of *mer*-genes represents only a fraction of the inserted *mer*-plasmids of [Gmer] concentrations. As an average, this fraction is ca. 1/3 (estimated by Maria [2010])

(g) lumped permease PT concentration is evaluated based on the experimental observations [Barkay et al., 2003], with the approximate formula (in nM): [PT] ≈ 1.5 {[PA] + 15} + [PR]; (see the details of Maria [2009b, 2010])

(h) At cell homeostasis, the PA reductase level (of average 1800 nM) is comparable with those of the permease PT (the Michaelis-Menten constant being 3700 nM for Hg^{2+}_{cyt} binding reaction [Barkay et al., 2003]

(i) Evaluated by optimizing the *mer*-operon GRC holistic properties [Maria, 2009b]. See also [Maria and Scoban, 2017, 2018]

the cells from inside the support, and of the resulted metallic mercury in the bioreactor bulk. The operating conditions are tightly controlled, that is the liquid flow rate, the aeration rate (pO_2), pH, and temperature to ensure an equilibrated bacteria growth (Table 3-2). The supplied oxygen is sufficiently enough to guarantee a good cell metabolism, and a high content of cytosolic NADPH necessary for mercuric ions reduction in the cell cytosol. Beside, the continuously bubbling air plays also the role of a carrier for the resulted volatile metallic mercury, by removing it from the liquid system. Eventually, the mercury vapours from the air leaving the system are condensed and recovered [Deckwer et al., 2004]. A background pollution of ca. 100 nM is considered in the input water (that is ca. 0.02 mg/L, which is smaller than the metabolic regulation threshold of 0.05 mg/L), thus maintaining active the *mer*-operon into the *E. coli* cell. The biomass content (X) of the support is variable (ca. 0.6-3 gX/L, according to [Deckwer et al., 2004; Berne and Cordonnier, 1995], but a quasi-constant level of ca. 1gX/L can be maintained by employing a purge/renewal system for the solid particles. At an industrial-scale, when treating polluted waters, the outlet gas (air) from the bioreactor, containing the volatile metallic mercury, is passed through an adsorption device, or through a desublimator system allowing the recover of metallic mercury [Ledakowicz et al., 1996].

To simulate the (semi-)continuous TPFB bioreactor performance, a dynamic ideal model was considered [Moser, 1988; Trambouze et al., 1988], by assuming homogeneous (perfectly mixed) liquid and gas phases, and an uniform distribution of the solid

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

particles (of uniform characteristics) in the G-L-S fluidized bed. By accounting for only the apparent mercury uptake rate (r_{app}) by the immobilized bacteria (of concentration c_x in Table 3-1) in the spherical solid carrier (of ε_s), the mercury mass balance in the liquid and gas phases is presented in the (Table 3-3). The TPFB reactor dynamic model includes terms referring to the mass balance in the bulk (liquid, L) phase, the gas phase (G), the interphase L-G transport (r_{trans}) of the volatile mercury, and the bioprocess inside the solid particles (r_{app}). The former one, also includes the diffusional resistance of the substrate (mercury ions) / product (dissolved metallic mercury) transport through the pumice support pores.

More specifically, the mercury dfferential mass balance in the TPFB given in (Table 3-3) includes the following terms:

(i). The apparent mercury reduction rate, evaluated at the solid interface;

(ii). The substrate (that is $S = Hg_L^{2+} = Hg_{env}^{2+}$) diffusional transport in the particle, expressed by the effectiveness factor (η) evaluated using the Thiele modulus for an apparent Michaelis-Menten type reaction [Doran, 1995], the effective diffusivity ($D_{s,ef}$) accounting for the molecular diffusion ($D_{S,L}$), the particle porosity (ε_p) and the tortuosity (τ) (other resistances being neglected [Narasimhan et al., 1999];

iii). The apparent rate $r_{j,app}$ was evaluated from solving (on every time-increment, for the considered species "j") the quasi-steady-state equality of mass fluxes at the solid-liquid (S-L) external interface (also including the external diffusion coefficient $k_s a_s$ [Doraiswamy and Sharma, 1984];

iv). The liquid-to-gas transport rate r_{trans} of the metallic mercury (Hg_L^0 to Hg_G^0), evaluated from the quasi-steady-state equality of mass fluxes at bubbles interface, by accounting for the mass transfer coefficients $k_L a_L$ (on the liquid film side; experimental value adopted), and $k_G a_G$ (on the gas film side; value from application of the Sharma's relationship given by Trambouze et al. [1988]).

The resulted hybrid dynamic model of (Table 3-3), with the apparent MM kinetics of (Table 3-1) allows simulating the transient operating conditions of the TPFB bioreactor by using an apparent MichaelisMenten kinetics of Philippidis et al.[1991a-c]. This apparent model of (Figure 2-50) can only give a rough idea on the bioreactor dynamics, but it is unable to describe the biomass adaptation to environmental changes, that is variations in the both inlet feed flow-rate and inlet mercury load in the influent. To offer a prediction to such engineering requirements, a structured kinetic model of the mercury uptake in the *E. coli* bacteria at a cellular level is necessary. The next section describes the WCVV structured cell model proposed by Maria [2009b, 2010] to simulate the dynamics of the mer-operon expression self-regulation in the "wild", or modified *E. coli* under various environmental (bioreactor operating) conditions.

By using the reactor model and the apparent Michaelis-Menten kinetics of Table 3-1, the TPFB dynamics has been simulated under the nominal operating conditions of Table 3-2, but every time varying one operating parameter, that is:

(a). The inlet $\left[Hg_L^{2+}\right]$ in concentration from the background pollution level (0.02mg/L) to successively 1,5 , or 10mg/L levels (Figure 2-51);

(b). The inlet liquid flow rate F_L of 0.01, 0.02, and 0.04 L/min (Figure 252);

(c). The biomass load c_x on the solid support taken as 0.1, 0.25, 1.0 gX/L (referred to the liquid volume) for a constant fraction of the solid in the reactor (Figure 2-53);

(d). The use of particle average diameter d_p of 1 mm, or 4 mm respectively (Figure 2-54).

The results of these TPFB simulations over ca. 50 min. running time reveal that the liquid residence time in the reactor [defined as the ratio between the inlet liquid flow rate, F_L (equal to the liquid exit flow rate), and the quasi-constant liquid volume in the bioreactor], and the biomass content (c_x) are the most influential operating parameters, being directly responsible for the realised mercury uptake conversion. For instance, by doubling the feed flow rate F_L (from 0.02 to 0.04 L/min), the uptake conversion is reduced with aprox. 5%. The particle size is also important, an increase in the average *dp* leading to a higher resistance to the diffusional transport in pores and to a particle effectiveness (η) diminishment (from 0.51 for d_p = 1 mm, to 0.15 for d_p = 4 mm, under nominal conditions).

On the other hand, the biomass average load in the reactor can be adjusted by employing a continuous purge-renewal system of the solid particles. As another remark, the mercury content in the input flow, $\left[Hg_L^{2+}\right]_{in}$ can vary in the range of 1- 40mg/L, and it presents a significant influence on the mercury reduction efficiency and on the mercury vapour content in the output gas (Figure 2-51). As further proved, this last parameter is also of tremendous importance for adaptation of bacteria metabolism when using wild or cloned cells with an increased copynumbers of *mer*-plasmids (denoted here by [Gmer]).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

3.2.5. *E. coli* cell structured dynamic model for mercury uptake, and its use to optimize the TPFB bioreactor

3.2.5.1 Generalities on the proposed extended hybrid SMDHKM model

A first approach to derive a kinetic model for this complex bioprocess / bioreactor is those of Maria [2009b, 2010]. The resulted *hybrid* dynamic model of (Table 3-3), with the apparent M-M kinetics of (Table 3-1) allows simulating the transient operating conditions of the TPFB bioreactor, but of a low quality prediction about the bioreactor dynamics. The main drawback of this TPFB apparent model is coming from its inability to describe the biomass adaptation to environmental changes, that is variations in the both inlet feed flow-rate and mercury load in the influent. And, self-understood this rough model can give no information about the dynamics of the inner cell species related to the mer-operon expression. To offer a prediction to such engineering requirements, a structured kinetic model of the mercury uptake in the *E. coli* bacteria at a cellular level is necessary. The next section describes the WCVV structured cell model proposed by Maria [2009b, 2010] and valorised by Maria and Luta [2013] in a SMDHKM kinetic model able to simulate the dynamics of the *mer*-operon expression self-regulation under various environmental (bioreactor) conditions. This model was constructed and validated by using the experimental data of Philippidis et al. [1991a-c], and the experimental information of Barkay et al.[2003] on the *mer*-operon characteristics. The WCVV cell model was built-up by accounting the GERM-s library of Maria [2005,2017a], and their relative performance indices (P.I.-s) of section 2.2.3, and the linking rules presented in section 2.2.4.

The *E. coli* cell WCVV kinetic model proposed by Maria [2009b, 2010] includes the all GRC-s responsible for the control of the *mer*-operon expression and for the whole process of mercury ions removal. The proposed GRC includes 4 lumped genes (denoted by GR, GT, GA, GD in (Figure 2-19, Figure 2-46, Figure 2-47, Figure 2-55) of individual expression levels induced and adjusted according to the level of mercury and other metabolites into cytosol. The whole process is tightly cross- and self-regulated to hinder the import of large amounts of mercury into the cell, which eventually might lead to the blockage of cell resources (RSH, NADPH, other metabolites, and proteins), thus compromising the whole cell metabolism. This GRC model includes four linked GERM-s of simple but effective [G(PP)1] type (see the discussion of sections 2.2.3, and 2.2.4) as follows (see Figure 2-55 and Figure 2-56):

(i). a GERM to regulate (calibrate) the Hg^{2+} transport across the cellular membrane, mediated by three proteins (PmerP, PmerT, and PmerC) from the periplasmatic space, considered as a lumped permease **PT** in the model. Phillippidis et al. [1991a-c] found this transport step as being energy-dependent and the rate-determining step for the whole mercury uptake process. Once the mercuric ion complex arrives in the cytosol, thiol redox buffers (such as glutathione of millimolar concentrations) form a dithiol compund $Hg(SR)_2$. Almost instantly, the **GT**, that is the lumped gene encoding **PT**, starts its expression induced by the regulatory protein **PR**. Any other perturbation is easily 'smoothed' by the large '**ballast**' effect of the proteome lump **P**. (see section 2.2.4-VI).

(ii). a GERM to control the expression of the PR protein that induces and controls the whole *mer*-operon expression in the presence of cytosolic Hg^{2+} (even if they are present in traces, that is low nM concentrations). This GERM acts as an amplifier of the merexpression leading to a quick (over ca. 30 s) cell response by starting the *mer*-enzymes production [Barkay et al., 2003]. The encoding GR gene expression to obtain **PR** is tightly controlled by the control protein **PD**, present in small amounts into the cell. The **GR** gene expression starts almost instantly in the presence of even small mercury concentrations in the cytosol.

(iii). a GERM to control the expression of PA enzyme responsible for the $Hg(SR)_2$ (that is Hg^{2+}_{cyt}) reduction to metallic mercury (Hg^{0}_{cyt}) into cytosol. The metallic mercury is relatively non-toxic for the cell, being easily removable through membranar diffusion into the bioreactor bulk liquid-phase. From here, metallic mercury (Hg^{O}_{L}) is continuously removed by the sparged air as mercury vapours $\left(Hg^{o}_{G}\right)$, to be later recovered. The encoding **GA** gene expression is induced and controlled by the PT protein.

(iv). a GERM controlling the control protein PD synthesis. This protein has a complex role, by maintaining a certain level of GR expression even when the mercury is absent in the cytosol [Barkay et al., 2003].

(v). a GERM controlling the replication of the lumped cell proteome (**P**) and genome (**G**) (of concentrations 10^7 nM, and 4500 nM , respectively) in the immobilized *E. coli* cells. These data are based on the Ecocyc [2005] databank (Figure 2-57), thus mimicking the cell 'ballast' effect on the cell genes expression, and on the all considered reactions (see section 2.2.4-VI). The need to include the cell content lump (the so-called 'cell ballast', see section 2.2.4-VI) in the WCVV model is due to the isotonicity constraint Eq.(6,11-15), that imposes to include in the model the all cell species, individually, or lumped. This lumped proteome/genome regulated module is legitimated by the possibility offered by such a structured WCVV cell model to 'mimics' the cell response to dynamic/stationary perturbations (section 2.2.3), and to reproduce the "smoothing effect" of perturbations leading to more realistic transient times (comparing to a cell with a 'sparing' content), the synchronized response to certain inducers, and the 'secondary perturbation' effect transmitted via the cell volume to which all cell components contribute {see Eq. (6,11-15), and the discussion of section 2.2.4-VI; Maria [2017a, 2018; Maria and Scoban, 2017, 2018]}.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

In total, the GRC dynamic model describing the mer-operon expression includes only 26 individual or lumped cellular species involved in 33 reactions (Figure 2-55, and Figure 2-56). The structured cell WCVV model is presented in the (Table 3-4). The cell model is coupled with the FBR -TPFB in a SMDHKM dynamic model (Table 3-3) through the $r_{j,app}$ linked to the $r_j\left(c_{j,s}\right)$ occurring inside cell, and by the imported Hg^{2+}_{cyt} and excreted Hg^{0}_{cyt} mass balances. The WCVV cell model includes not only the reactions and the dynamics of the *mer*-GRC, but also the enzymatic reactions directly responsible for the environmental mercury Hg^{2+}_{env} (also denoted by Hg^{2+}_{L}) import into the cell as Hg^{2+}_{cyt}, and for its reduction to cytosolic metallic mercury Hg^{0}_{cyt} in cloned *E. coli* cells.

All reactions in the cell model of (Table 3-4) are considered elementary, excepting some of them for which extended experimental information exists; see the discussion of [Maria, 2009b, 2010; Maria and Luta, 2013]:

(a). a Michaelis-Menten rate expression for the mercuric ion permeation through the membrane into the cell;

(b). a Michaelis-Menten rate expression for the mercuric ion reduction in cytosol;

(c). a Hill type quick induction of the **GR** expression that can rapidly initiate the production of permease **PT** (through the control protein **PR**) when mercuric ions are present in large amounts.

(d). Dimmerization reactions of TF-s are considered to be much rapid than the enzyme synthesis, while equal concentrations of active gene (**G**) and inactive (**GPP**) forms of the generic gene **G** are considered at homeostasis to maximize the GERM efficiency (see section 2.2.4.-III) (Figure 2-55, and Figure 2-56).

(e). The lumped proteome **P**, present in a large amount, is included in all gene expression rates, thus leading to a more realistic evaluation of the GERM regulatory efficiency indices (P.I.-s) (section 2.2.3).

(f). The model rate constants are estimated from solving the cell stationary mass balances for using the known nominal concentrations of observable species, with a rule presented by Eq.(11-15), and by the section 2.2.1.3. , but also from optimizing the GERM-s regulatory indices, as described in the mentioned paragraph. The *E. coli* cell characteristics and the considered cell key-species homeostatic concentrations used in the rate constants estimation are given in (Table 3-5).

Exceptions are the Michaelis-Menten rate constants for the mercury transport and its reduction in cytosol adopted from the Phillipidis et al.[1991a-c] based on experimental kinetic data. Thus, the M-M rate constants depends on the amount of [Gmer] plasmids in the cloned *E. coli* cells (Figure 2-56, and Figure 2-50). Simple correlations are used to include this essential aspect in the model. Some stationary concentrations of species of little importance in the model (e.g. [GRPRPR]) have been estimated together with the model rate constants, by imposing the GRC optimal regulatory (P.I.-s). For instance, by estimating some unknown stationary concentrations of few number of intermediates, and by adjusting the optimum TF level of gene expression, to obtain the minimum recovering times after a 10% dynamic perturbation in the key species, and smallest sensitivities of the homeostatic levels *vs.* external perturbations (see sections 2.2.1.3, and 2.2.4).

As extensively discussed in the section 2.3, such a cell WCVV structured model presents a large number of advantages, being able to:

(1). Simulate the cell metabolism adaptation when the environmental mercury level changes. Such a reconfiguration of the levels of mergenes and *mer*-proteins is presented in the (Figure 2-58) as a step response after a, step'-like perturbation in the mercury level from $[\,Hg^{2+}_{env}\,]s = 0.1\mu M$ to 10 μM (ca. 2 mg/L), in a cloned *E. coli* cell with *mer*-plasmids of [Gmer] = 140 nM, compared to a cell cloned with only [Gmer] = 3 nM. The transient state toward the cell new homeostasis of adapted mer-gene/protein levels stretches over 15-20 cell cycles (of ca. 0.5 h each) as long as the environmental perturbation is maintained. Clearly, an *E. coli* cell with a higher content of merplasmids reacts much strongly to the perturbation, by quickly starting to produce the enzymes responsible for the mercury removal.

(2). Because the Hg^{2+} reduction rate constants are dependent on the *mer* plasmids level into the cell, the WCVV GRC model can predict the maximum level of mer-plasmids that can be added to the cell genome for improving its mercury up-taking rate (and capacity) without exhausting the internal cell resources, thus putting in danger the cell survival. Consequently, it follows that this cell model allows the insilico design of modified *E. coli* cloned with a suitable amount of merplasmids to improve its possibility of cleaning wastewaters by improving its capacity to uptake the environmental mercury. As an example, in the (Figure 2-59) are presented the cell key-component stationary levels, and concentration of mercuric ions in the bioreactor bulk-phase for two analysed GMO-s: one cloned with 67 nM, and another one cloned with 140 nM. Simulations of Maria and Luta [2013] revealed that as the mer-plasmids level is increasing, as the mercury uptake capacity in increasing. However, an upper limit exists (around 140 nM) over which the cell resources will be exhausted, putting its metabolism in danger. The present SMDHKM – WCVV kinetic model put in evidence such a tendency (see Figure 4 of [Maria, 2010]).

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

(3). By coupling the structured cell model of (Table 3-4) with the threephase continuous bioreactor model FBR-TPFB of Table 3-3 (with immobilized *E. coli* cells on pumice beads; see also the bioreactor model in the Figure 2-50), Maria and Luta [2013], and Maria et al. [2013] have been able to determine the optimal operating policies of the bioreactor in relationship to the culture of cloned cells characteristica. Similar studies are reported by Scoban and Maria [2016, 2017d], but using a global (unstructured) kinetic model.

3.2.5.2. Extended SMDHKM - WCVV model structure

To model the mercury uptake by *E. coli* bacteria at a cellular level, Maria [2009b, 2010] proposed a WCVV structured model that reproduces the dynamics of the *mer*-operon expression under various environmental conditions (this section 3.2). The model was constructed and validated by using the Philippidis et al. [1991a-c] experimental data, and the Barkay et al. [2003] experimental information on the *mer* -operon characteristics. The proposed GRC, of modular construction, includes seven GERM-s "placed" in an *E. coli* cell (by also including the lumped genome/proteome replication) of known characteristics (size, cell cycle, copynumbers of all individual or groups of components from EcoCyc databank [2005]). The whole-cell approach allows predicting the dynamic response of the involved mer-operon considered species and of the other lumped cell components to various perturbations, by mimicking the cell growth and its response to external stimuli.

The *mer*-GRC acts by adjusting the expression level of individual *mer*-genes (self-regulation), but also by balancing the expression of related genes through various regulatory loops activated by certain exo-/endogeneous inducers. The basic unit of the construction is the semi-autonomous GERM (sections 2.2.1 and 2.2.2) that control one gene expression [Maria, 2005, 2007; Savageau, 2002]. To not complicate the GRC dynamic model, the GERM module must include the minimum number of individual/lumped species/reactions necessary to adequately represent the experimental information. For instance, the gene expression general schema of (Figure 2-60) can be translated into a modular structure of reactions, more and more reduced depending on the GERM representation importance. Maria [2005], and Yang et al. [2003] proposed a nomenclature of the GERM units that includes negative feedback regulatory loops and a cascade control [see section 2.2.4-III, and Eq. (19)], thus allowing application of the control theory principles to quantitatively evaluate the individual GERM and the whole GRC regulatory performance indices (P.I.-S, see section 2.2.3), such as: *stationary efficiency* to treat 'step'-like perturbations in key species concentrations; *dynamic efficiency* to treat 'impulse'-like perturbations (translated by the species recovering times of the steady-state / homeostasis); *responsiveness* to stimuli (expressed by the transient times from one steady-state to a new one); *species connectivity* (expressed by a synchronised response to perturbations); *system stability region and strength* (see [Maria, 2017a, 2007; Wall et al., 2003] for details).

The GRC lumped reaction pathway for mercury uptake in *E. coli* proposed by Maria [2009b, 2010] is schematically illustrated in (Figure 255). It contains seven linked GERM-s, as follows: (i) a GERM to regulate the Hg^{2+} transport across the cellular membrane, mediated by a lumped permease PT, which is synthesized by the GT gene expression induced by the regulatory protein **PR**. Once the mercuric ions arrive in the cytosol, thiol redox buffers (such as glutathione of millimolar concentrations) form a dithiol derivative $Hg(SR)_2$; (ii) a GERM to control the synthesis of the **PR** protein that induces and controls the whole mer-operon expression in the presence of cytosolic Hg^{2+} (even in nM concentrations). This GERM acts as an amplifier of the *mer* expression leading to a quick (over ca. 30 s [Barkay et al., 2003]) cell response, by starting production of *mer*-enzymes. The gene **GR** expression to obtain **PR** is controlled by the protein **PD**, present in small amounts into the cell; (iii) a GERM to control the synthesis of PA reductase responsible for the $Hg(SR)_2$ reduction to metallic mercury (relatively non-toxic for the cell, easily removable through membranar diffusion). The encoding gene GA expression is induced and controlled by the PT protein; (iv) a GERM controlling the protein PD synthesis, having the role of maintaining a certain GR expression level in the absence of mercury [Barkay et al., 2003]; (*v*) a GERM controlling the replication of the lumped cell proteome (**P**) and genome (**G**) (of large concentrations), thus mimicking the cell 'ballast' and 'smoothing' effect of perturbations on the other gene expressions and reactions of the cell (see section 2.4.-VI); (vi) two reaction modules, characterizing the Hg^{2+} import into the cell (activated / saturated by the environmental Hg^{2+}, mediated by **PT** and RSH), and its cytosolic reduction mediated by **PR** in the presence the renewable co-enzyme NADPH (present in excess in the cell).

The WCVV model equations for the mercury uptake in *E. coli*, together with the general hypotheses of the WCVV model are presented at the top of (Table 3-4). Basically, the cell is considered an open system, of uniform content and negligible inner gradients. To not increase the number of parameters, the structured model includes GERM-s of simple form, but of better regulatory efficiency (see sections 2.2.3, 2.2.4, and Figure, 2-9, Figure 2-13, Figure 2-16, and Figure 2-37) for all mer-genes, by using dimeric TF-s (of Protein:Protein type) to increases the GERM regulatory efficiency, as *in-silico* proved in the above mentioned sections and figures, and by [Maria, 2009, 2017a; Salis and Kaznessis, 2005]. The resulted SMDHKM - WCVV kinetic model includes 26 individual or lumped cellular species and 33 reactions. All reactions are considered elementary, except some for which extended experimental information exists, that is the Michaelis-Menten rate expressions for mercury permeation into the cell, and its reduction in cytosol. A Hill type induction of the **GR** expression is adopted to rapidly amplify the *mer* operon expression when mercuric ions are present in significant amounts inside cell. Dimmerization reactions of **TF**-s are considered to be much rapid than the enzyme synthesis, while equal concentrations of active (G(i)) and inactive (G(i):TFTF) forms of the generic gene

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

G(i) are considered at homeostasis to maximize the GERM efficiency (see sections 2.2.3, 2.2.4, and Fig, 2-9, Figure 2-13, Figure 2-16, and Figure 2-37) [Maria, 2005, 2017a, 2018]. The homeostatic characteristics of *E. coli* cells (belonging to a uniform culture) from the reactor, and the adopted species concentrations are presented in (Table 3-5).

Phillippidis et al. [1991a-c] have proved, on an experimentally basis, that the membranar permeation step is the rate-determining step for the mercury uptake by the bacteria, and they used cloned *E. coli* cells with a higher amount of *mer*-plasmids (Gmer from 3 to 140 nM) to increase the content of *mer*-genes producing mer-enzymes responsible for the mercury uptake process. However, the increased mercury reduction efficiency is obtained up to a certain limit of *mer*-plasmids (with imported mercury ions controlled by the permease PT) to not exhaust the internal cell resources thus putting in danger the cell survival. This limit will be determined here by using the SMDHKM - WCVV kinetic model of Maria [2009b, 2010].

When simulating the TPFB reactor performance (of nominal conditions given in the Table 3-2) by employing a classical unstructured reactor model (see section 3.2.3), the biomass adaptation to variable input loads is not accounted for, that is the maximum rates $v_{m,t}$, $v_{m,p}$, and the Michaelis inhibition constants $K_{m,t}$, $K_{m,p}$,, and K_{ip} of (Table 3-1) are kept constant (eventually depending only on the Gmer plasmid level). In fact, $v_{m,t}$ and $v_{m,p}$ rate constants depend on the cell enzyme average levels of **PA** (reductase) and **PT** (permease) (Figure 2-55),

which vary during the bacteria adaptation in the bioreactor under stationary or transient conditions.

To solve this problem by also offering a more detailed and robust prediction on the cell system dynamics, the TPFB bioreactor and the mer-GRC of *E. coli* cell models were coupled in a *hybrid* structured SMDHKM - WCVV kinetic model. The linked differential mass balance equations (i.e. ODE set) are solved simultaneously, by applying a mutual exchange of input/output parameters $\{\left[Hg^{2+}_{env}\right)],\left[Hg^{0}_{env}\right]$, enzymes **PT** and **PA** concentrations} on every small time increment throughout the solution (integration) of the SMDHKM model, as graphically represented in (Figure 2-55).

The solving rule involves successive integration steps with an adopted time-step-interval equal to the cell cycle (ca. 30 min), enough to obtain the steady-state of the TPFB reactor on every time-interval, as following: ((i)The rule starts with solving the extended hybrid SMDHKM *E. coli* cell model by using the known initial conditions (that is the cell species initial state, or those from the end of the previous integration cycle), and by considering the current concentration of $\left[Hg^{2+}_{env}\right]$ in the bioreactor liquid phase (the nutrients are considered in excess and of constant levels, not included in the SMDHKM model). Thus, the cell species dynamics over one cell cycle is obtained from solving the WCVV cellular model. (ii) Then, the TPFB reactor model is solved over the current time-interval by using the known initial condition (i.e. the reactor state variables from the end of the previous time-interval), and by considering the enzymes PT and PA concentrations resulted from the *E. coli* model solution. [PT] and [PA] are necessary for setting the maximum reaction rates $v_{m,t} = k_7 cPT$ and $v_{m,P} = k_8 cPA$ in the reactor model (Table 3-4-a). The biomass level on the support is taken constant in the simulated case study, but an additional mass balance can be easily added to the reactor model if necessary.

The procedure is repeated of a large number of times, over hundreds of cell cycles. For instance, (Figure 2-58) displays the *E. coli* cell adaptation after a 'step'-like perturbation in the environmental $\left[Hg^{2+}_{env}\right]$s (that is in the bioreactor bulk-phase) from the background level of 0.1 μM to 10μM (ca. 2 mg/L), for the case of cell cloned with [Gmer] = 3 nM (full line), or with [Gmer] = 140 nM (dash line) mer-plasmids. The transient state toward the cell new homeostasis usually lies over 15-20 cell cycles as long as the environmental stationary perturbation is maintained. The simulated species dynamic trajectories in this (Figure 2-58) reveal a vigorous response of the cell *mer*-species (especially of **PT**, and **PA**) to the perturbation in the environmental $\left[Hg^{2+}_{env}\right]$s (i.e. liquid bulkphase).

3.2.5.3. Fitting the extended HSMDM model parameters from experimental data

To overcome the high computational effort necessary to fit the large number of rate constants of this extended WCVV - SMDHKM model, and to increase its confidence and physical meaning, a three-steps procedure was employed based on the available experimental data and additional information from literature, as followings.

(a). **Mass transport parameters** of the TPFB reactor, that is interfacial partial transfer coefficients ($k_S a_S, k_L a_L, k_G a_G$), effective diffusivity (D_{ef}), and particle effectiveness factor (η) have been estimated by using the experimental data of Deckwer et al. [2004], and based on common correlations from the chemical engineering literature (Table 3-3) evaluated for the specified reactor operating conditions of (Table 3-2).

(b). **Cell model parameters are estimated** by using the Philippidis et al. [1991a-c] kinetic data obtained from separate batch experiments with notcloned ("wild") *E. coli* cells. By using the defined cell nominal characteristics of (Table 3-5) (the key

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

notations are the followings: t_c = cell cycle time; $V_{cyt,o}$ born cell volume; ^cX = biomass concentration in the bioreactor; **P** = cell lumped proteome; **G** = cell lumped genome; NutG, NutP = lumped nutrients for the **G** and **P** synthesis, respectively; [Gmer] = mer-plasmid level in the nucleus. The stationary levels of the essential cell mer-proteins (PA, PR, PD), and for the involved TF-s (PRPR, PTPT, PAPA, PDPD) are taken from the literature data, and by maximizing the GERM-s regulatory efficiency P.I.-s (see section 2.2.1.3). The resulted rate constants of (Table 3-4) have been estimated for the most severe experimental conditions of $\left[H_{env}^{2+}\right]s = 120\mu M$, and [Gmer] = 140nM, to broaden the model adequacy domain. The used first guess of Hill-induction rate constants ($n_{PD} = 1, n_{PR} = -0.5, n_H = 2, a = 3$) have been adopted at values recommended in the literature, by similarity with the Hill-induction of gene expression in genetic switches (see the remarks included in Table 3-4), while the Michaelis-Menten rate constants of mercury membranar transport ($r_{\max,t}, K_{mt}$), and its reduction rates' constants ($r_{max}, P, K_{mP}, K_{iP}$) have been kept at the fitted values of Philippidis et al. [1991a-c] (Table 3-1). The reference concentration ^{c}Hg2cy, ref was adopted at the average cytosolic level of mercuric ions detected by Philippidis et al. [1991c]. During model estimation, the GERM regulatory indices have been kept at their optimized levels, which corespond to: (i) Equal concentrations of catalytically active/inactive forms [G(j)]s = [G(j)TFn]s adopted at steady-state to ensure GERM maximum regulatory efficiency *vs.* perturbations (i.e. smallest sensitivities of the homeostatic levels *vs.* external perturbations, see section 2.2.3, and Figure 2.20, Figure 2-32); (ii) adjustable optimum [TF]s level (ca. 4 nM here; Table 3-5) involved in the gene expressions to obtain the minimum recovering times after a dynamic perturbation in the key species (section 2.2.1.3, and [Maria, 2009, 2017a, 2018; Maria and Scoban, 2017, 2018]). Other adjustable parameters, such as the cell concentration in biomass (c_{cell}, are tuned to fit the experimental cellular mercury reduction rate of Philippidis et al. [1991a].

(c). The model validity is extended over a wider experimental range, by covering the input $\left[Hg_L^{2+}\right] = 0 - 100 mg/L$, and plasmid levels of [Gmer] = 3 - 140 nM. All these have been realised by adjusting some keyparameters of the extended WCVV - SMDHKM model, and by applying two concomitant fitting criteria, that is: (1) the cytosolic mercury reduction rate should fit the experimental values of Philippidis et al. [1991a-c], obtained for ‚*permeabilized*' cells (that is using the estimation criterion of "Min $\left\| r_{Hg,\exp} - r_{8,model} \right\|$, see ^{r}8 of Table 3-4); (2) the TPFB reactor output ^{c}Hg, out predicted by the unstructured model (Table 3-1, and Table 3-3, and chap. 3.2.3) with using the Philippidis et al. [1991c] Michaelis-Menten parameters and the mass transfer terms, in order to fit with those of the structured *hybrid* (cell + reactor) SMDHKM dynamic model of (Table 3-3, and Table 3-4). The Hill parameter b = $2a^4$ was adjusted by using the following approximate linear dependence on the inlet mercury load:

$$a = a_{\min} + \left(a_{\max} - a_{\min}\right)\left(c_{Hg,in} - c_{Hg,in}^{\min}\right) / \left(c_{Hg,in}^{\max} - c_{Hg,in}^{\min}\right),$$

(see the fitted values a_{min} and a_{max} in Table 3-4 - footnote d). The above simplified correlation tries to account for the influence of the environmental mercury on the induction characteristics of the *mer*-operon expression, that is a slow induction in the presence of low levels of Hg^{2+} inducer, and a sharp (sigmoidal) mer-operon expression response to high levels of $\left[Hg_{cyt}^{2+}\right]$ inducer. When a certain saturation level is reached, the limited cell resources will impose a '*reduction*' of the metabolic synthesis of mer-proteins, and a decrease of their levels into the cell, irrespectively to the increased amount of mercury in the environment [Philippidis et al., 1991a-c].

The extended *hybrid* SMDHKM model predictions are in a satisfactory agreement with the experimental data. Thus, the predictions of ^{r}*Hg* of (Figure 2-61, A-C) (curves '2') roughly fall within the confidence band (that is the standard deviation limits) [‚Eup', Elow'] of the experimental curves (‚E') of Philippidis et al. [1991a] for three different cloned cell cases (confidence curves being plotted by taking constant the reported maximum relative error of ca. 20%). The unstructured model (section 3.2.3, Table 3-1 , and Table 3-3) predictions (that is curves '1') reported apparent ^{r}Hg of low adequacy, that is lower values due to the rough bioprocess representation by the Michaelis-Menten global model, and due to the inclusion of the mass transport resistance between the three-phases in contact (as also done by the *hybrid* SMDHKM model). The extended *hybrid* SMDHKM model adequacy in terms of standard error 'Std' / average observed value ' Obs' ratio is evaluated for every cloned cell culture, leading to Std/Obs of 22.3%, 16.7%, 20.4%, 19.9%, 18.8% for [Gmer] levels of 3, 67, 78, 124, and 140 nM respectively, that is acceptable values comparable to the maximum experimental error (of 20%). The structured *vs.* unstructured model outputs, in terms of outlet mercury concentration (^{c}*Hg*, out) are also in a fair agreement (curves not display here), that is Std/Obs values of 3.0%, 4.7%, 3.7%, and 12% for [Gmer] levels of 67, 78, 124, and 140 nM respectively. The model poor adequacy for the [Gmer = 3nM data set and low $\left[Hg_L^{2+}\right]$ in the environment (input) might be explained by the use of a less adapted *E. coli* cell than probably those studied by Phillipidis et al. [1991a], reflected by a smaller [PA] initial level of 600 nM (see the PA-curve in Figure 2-61-D). Once a higher level of Gmer - plasmids are introduced into the cell, and a higher Hg_L^{2+} stimulus level is present, then the cytosolic PA level is tripling.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

The extended *hybrid* SMDHKM structured model rate constants (see the above point (b), and Table 3-4) are in a good agreement with the reported data from literature, as pointed-out in the same (Table 3-4). The fitted rate constants multiplied by the reactant lead to reaction rates of the same order of magnitude with those reported in the literature for similar genetic processes, such as the TF (repressor monomer) dimmerization, the TF binding to gene operator, or the mRNA (genes) synthesis reactions (Table 3-4) [Maria, 2009b, 2010]. This observation sustains the physical meaning of model parameters, thus increasing the SMDHKM structured model robustness.

By extending the detailing degree of the bioreactor dynamic model at a cellular level, the resulted *hybrid* SMDHKM structured model preserves the fair adequacy of the unstructured model, but adding the possibility to predict the cell species/fluxes dynamics, and the biomass adaptation to the environmental changes [Maria, 2010]. By offering details on the cell metabolism adaptation, on the 'intrinsic' reduction rate, and possibilities to *in-silico* predict the modified cell response to various stimuli, such hybrid SMDHKM structured models are clearly superior to the unstructured (apparent) bioreactor dynamic models. Consequently, even if complex, extended hybrid SMDHKM structured models worth the supplementary experimental and computational effort to derive them.

3.2.5.4. Prediction of the bioreactor dynamics by using the extended HSMDM model compared to the unstructured model

Simulation of the TPFB bioreactor with the *unstructured* model (including only 3 bulk variables, that is $[Hg^{2+}]$, $[Hg^{0}]$, X) of (Table 3-1 , and Table 3-3) can predict the dynamics of the mercury uptake (3 bulk variables) by the biomass and the resulted metallic mercury transfer in the liquid and gas phase of the TPFB, as displayed in (Figure 2-62) after a 'step'-like perturbation in the Hg^{2+} level of the input flow (Figure 2-62, up-left) from traces of mercury (100 nM) up to 20 mg/L(curves '1'). These predicted trajectories are neither identical, nor close to those predicted by the structured *hybrid* SMDHKM WCVV cell and reactor model (including 26 cellular +3 bulk variables) (curves 2') for two types of cloned cell cases (with [Gmer] = 67 nM and [Gmer] = 140 nM, respectively). Also, against the unstructured model, the hybrid SMDHKM WCVV reactor model correctly predict the $[Hg^{2+}]$ quick increase in the reactor outlet.

The difference between the two process models (structured *vs.* unstructured) are even bigger. Thus, the structured SMDHKM WCVV model is able to '*visualise*' the dynamics of the involved metabolic processes over several cell cycles and the response and the adaptation of the cell target metabolism (*mer*-operon expression here) to the external stimuli. As plotted in (Figure 2-63), the structured bi-level SMDHKM WCVV model can predict the sharp increase in the *mer*-genes (GR, GT, GA, GD) levels and, the encoded mer-proteins (PR, PT, PA, PD), and cytosolic ('intrinsic') reduction rate, as well as the cytosolic content of mercury for both cloned cell cultures following a step-perturbation in the influent Hg^{2+} load (at an arbitrary time t = 9000 min). As the plots indicate, the transient period is of ca. 450 min (over the interval marked [9000, 9450] min on the time-axis), that is over ca. 15 cell cycles. The cell levels of the involved *mer*-genes and of their encoded proteins are continuously changing, while the genetic information is transmitted from one cell generation to the next one. In accordance to its structure, the cell model can also predict the effect of increasing the content of some genes (or of removing them by gene knock-out technique, not approached here). Such modelling aspects are impossible to be represented by an overall / unstructured model. Moreover, the structured SMDHKM WCVV model can predict the cell characteristics and operating conditions required to obtain an imposed efficiency of the bioreactor (not presented here; see an example for another process given by Maria and Renea [2021]).

One disadvantage of the structured *hybrid* SMDHKM WCVV models is related to the large amount of required experimental information on the cell metabolism and on the studied GRC necessary to adequately fit the model parameters. Another disadvantage is related to the required large computational time (up to hours on a common PC using a 'stiff' integrator) to obtain a reasonable prediction of the cell bioprocess evolution. However, this last constraint is not very strong for the process monitoring purposes, as long as most of the industrial biological processes are slow, with a time constant much larger than the required simulation time. Beside, for quick control purposes, a reduced hybrid unstructured model can be used instead, being adjusted by using the experimental data (see the example of Maria [2023]).

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this book.

The author confirm that their paper has no conflict of interest of any kind, and of any nature.

Data availability statement

Experimental datasets and some information used in this study come from the authors own experiments, or are imported from the literature, every-time the source being referred in the text.

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Cover figures sources: Maria et al. (2002); Maria, 2005, 2006, 2007. https://st4.depositphotos.com/3900811/22071/v/1600/depositphotos_220715852-stock-illustration-protein-synthesis-vector-illustration-transcription.jpg

References

1. Aach J, Rindone W, Church GM. Systematic management and analysis of yeast expression data. Genome Res. 2000;10:431–445. Available from: https://doi.org/10.1101/gr.10.4.431
2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular biology of the cell. 4th ed. New York: Garland Science; 2002. Available from: https://www.ncbi.nlm.nih.gov/books/NBK21054/
3. Alon U. An introduction to systems biology: design principles of biological circuits. Boca Raton: Chapman & Hall/CRC; 2007. Available from: https://www.taylorfrancis.com/books/mono/10.1201/9781420011432/introduction-systems-biology-uri-alon
4. Alvarez E, Cancela MA, Navaza JM, Taboas R. Mass transfer coefficients in batch and continuous regime in a bubble column. In: Proceedings of the International Conference on Distillation and Absorption; 2002 Sep; Baden-Baden, Germany. Düsseldorf: GVC–VDI Society of Chemical and Process Engineering; 2002. Available from: https://skoge.folk.ntnu.no/prost/proceedings/distillation02/dokument/6-3.pdf
5. Agbachi CPE. Pathways in bioinformatics: a window in computer science. Int J Comput Trends Technol. 2017;49(2):83–90. Available from: https://www.ijcttjournal.org/archives/ijctt-v49p113
6. Akutsu T, Kuhara S, Maruyama O, Miyano S. A system for identifying genetic networks from gene expression patterns produced by gene disruptions and overexpressions. In: Proceedings of the 9th Workshop on Genome Informatics; 1998 Dec 10; Tokyo, Japan. Available from: https://pubmed.ncbi.nlm.nih.gov/11072331/
7. Aris R. Elementary chemical reactor analysis. New Jersey: Prentice-Hall; 1969. Available from: https://www.sciencedirect.com/book/monograph/9780409902211/elementary-chemical-reactor-analysis
8. Atauri P, Curto R, Puigjaner J, Cornish-Bowden A, Cascante M. Advantages and disadvantages of aggregating fluxes into synthetic and degradative fluxes when modelling metabolic pathways. Eur J Biochem. 1999;265:671–679. Available from: https://doi.org/10.1046/j.1432-1327.1999.00760.x
9. Atkinson MR, Savageau MA, Myers JT, Ninfa AJ. Development of genetic circuitry exhibiting toggle switch or oscillatory behaviour in Escherichia coli. Cell. 2003;113:597–607. Available from: https://doi.org/10.1016/s0092-8674(03)00346-5
10. Bailey J. Complex biology with no parameters. Nat Biotechnol. 2001;19:503–504. Available from: https://doi.org/10.1038/89204
11. Banga J. Optimisation in computational systems biology. In: Proceedings of the 6th Simon Stevin Lecture on Optimisation in Engineering; 2008; Leuven-Heverlee, Netherlands. Also published in: BMC Syst Biol. 2008;2:47. Available from: https://link.springer.com/article/10.1186/1752-0509-2-47
12. Bar-Joseph Z, Gitter A, Simon I. Studying and modelling dynamic biological processes using time-series gene expression data. Nat Rev Genet. 2012;13:552–564. Available from: https://doi.org/10.1038/nrg3244
13. Barkay T, Miller SM, Summers AO. Bacterial mercury resistance from atoms to ecosystems. FEMS Microbiol Rev. 2003;27:355–384. Available from: https://doi.org/10.1016/s0168-6445(03)00046-9
14. Bartol TM Jr, Stiles JR. MCell: a general Monte Carlo simulator of cellular microphysiology. La Jolla (CA): The Salk Institute, Computational Neurobiology Lab; 2002. Available from: http://www.mcell.cnl.salk.edu/
15. Bailey JE. Towards a science of metabolic engineering. Science. 1991;252:1668–1675. Available from: https://doi.org/10.1126/science.2047876
16. Benner SA, Sismour AM. Synthetic biology. Nat Rev Genet. 2005;6:533–543. Available from: https://doi.org/10.1038/nrg1637
17. Berne F, Cordonnier J. Industrial water treatment. Paris: Gulf Publishing Co.; 1995. Available from: https://books.google.co.in/books/about/Industrial_Water_Treatment.html?id=RQlSAAAAMAAJ&redir_esc=y
18. Bier M, Teusink B, Kholodenko BN, Westerhoff HV. Control analysis of glycolytic oscillations. Biophys Chem. 1996;62:15–24. Available from: https://doi.org/10.1016/s0301-4622(96)02195-3
19. Blass LK, Weyler C, Heinzle E. Network design and analysis for multi-enzyme biocatalysis. BMC Bioinformatics. 2017;18:366. Available from: https://doi.org/10.1186/s12859-017-1773-y
20. Bower JM, Bolouri H. Computational modelling of genetic and biochemical networks. Cambridge (MA): MIT Press; 2001. Available from: https://mitpress.mit.edu/9780262024815/computational-modeling-of-genetic-and-biochemical-networks/
21. Bray D, Bourret RB, Simon MI. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. Mol Biol Cell. 1993;4:469–482. Available from: https://doi.org/10.1091/mbc.4.5.469
22. Brazhnik P, de la Fuente A, Mendes P. Gene networks: how to put the function in genomics. Trends Biotechnol. 2002;20:467–472. Available from: https://doi.org/10.1016/s0167-7799(02)02053-x
23. Browning ST, Shuler ML. Towards the development of a minimal cell model by generalisation of a model of E. coli: use of dimensionless rate parameters. Biotechnol Bioeng. 2001;76:187–192. Available from: https://doi.org/10.1002/bit.10007

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

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

25. Buchholz K, Hempel DC. From gene to product (editorial). Eng Life Sci. 2006;6:437.

26. Burns JA, Cornish-Bowden A, Groen AK, Heinrich R, Kacser H, Porteous JW, Rapoport SM, Rapoport TA, Stucki JW, Tager JM, Wanders RJA, Westerhoff HV. Control of metabolic systems. Trends Biochem Sci. 1985;10:16.

27. Camacho D, Vera Licona P, Mendes P, Laubenbacher R. Comparison of reverse-engineering methods using an in silico network. Ann N Y Acad Sci. 2007;1115:73–89. Available from: https://doi.org/10.1196/annals.1407.006

28. Carbone N, Tharoor I. The race is over: cracking the code. The human genome was deciphered by J. Craig Venter and Dr Francis Collins, Time magazine. 2000 Jul 3.

29. Casey R, de Jong H, Gouze JL. Piecewise-linear models of genetic regulatory networks: equilibria and their stability. J Math Biol. 2006;52:27–56. Available from: https://doi.org/10.1007/s00285-005-0338-2

30. Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M. Dynamic modelling of the central carbon metabolism of Escherichia coli. Biotechnol Bioeng. 2002;79:53–73. Available from: https://doi.org/10.1002/bit.10288

31. Chen T, He HL, Church GM. Modelling gene expression with differential equations. In: Proceedings of the Pacific Symposium of Biocomputing; 1999 Jan 4–9; Hawaii. p. 29–40. Available from: https://pubmed.ncbi.nlm.nih.gov/10380183/

32. Chen MT, Weiss R. Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signalling elements from Arabidopsis thaliana. Nat Biotechnol. 2005;23:1551–1555. Available from: https://doi.org/10.1038/nbt1162

33. Cornish-Bowden A. Biochemical evolution: the pursuit of perfection. New York: Garland Science; 2016. Available from: https://www.taylorfrancis.com/books/mono/10.1201/9780429258800/biochemical-evolution-athel-cornish-bowden

34. Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO. Integrating high-throughput and computational data elucidates bacterial networks. Nature. 2004;429:92–96. Available from: https://doi.org/10.1038/nature02456

35. Covert MW, Schilling CH, Famili I, Edwards JS, Goryanin II, Selkov E, Palsson BO. Metabolic modelling of microbial strains in silico. Trends Biochem Sci. 2001;26(3):179–186. Available from: https://doi.org/10.1016/s0968-0004(00)01754-0

36. Crampin EJ, Schnell S. New approaches to modelling and analysis of biochemical reactions, pathways and networks. Prog Biophys Mol Biol. 2004;86:1–4. Available from: https://doi.org/10.1016/j.pbiomolbio.2004.04.001

37. CRGM-Database. Centro di Ricerca in Genetica Molecolare, Istituto di Istologia e Embriologia Generale, University of Bologna; 2002. Available from: http://apollo11.isto.unibo.it/Cells.htm

38. Das S, Caragea D, Welch SM, Hsu WH. Handbook of research and computational methodologies in gene regulatory networks. New York: Medical Information Science Reference (IGI Global); 2010. Available from: https://books.google.co.in/books?id=6X16uAAACAAJ&source=gbs_book_other_versions_r&redir_esc=y

39. DeBeer DJ, Kourie DG. Artificial life: an overview. S Afr J Sci. 2000;96:569–576.

40. Deckwer WD, Becker FU, Ledakowicz S, Wagner-Döbler I. Microbial removal of ionic mercury in a three-phase fluidised bed reactor. Environ Sci Technol. 2004;38:1858–1865. Available from: https://doi.org/10.1021/es0300517

41. de Jong H, Geiselmann J, Hernandez C, Page M. Genetic network analyser: qualitative simulation of genetic regulatory networks. Bioinformatics. 2003;19(3):336–344. Available from: https://doi.org/10.1093/bioinformatics/btf851

42. Demin OV, Goryanin II, Kholodenko BN, Westerhoff HV. Kinetic modelling of energy metabolism and superoxide generation in hepatocyte mitochondria. Mol Biol. 2001;35:1095–1104. Available from: https://pubmed.ncbi.nlm.nih.gov/11771135/

43. Dhir S, Morrow KJ Jr, Rhinehart RR, Wiesner T. Dynamic optimisation of hybridoma growth in a fed-batch bioreactor. Biotechnol Bioeng. 2000;67(2):197–205. Available from: https://pubmed.ncbi.nlm.nih.gov/10592517/

44. DiBiasio D. Introduction to the control of biological reactors. In: Shuler ML, editor. Chemical engineering problems in biotechnology. New York: American Institute of Chemical Engineers; 1989. p. 351–391.

45. Domach MM, Leung SK, Cahn RE, Cocks GG, Shuler ML. Computer model for the glucose-limited growth of a single cell of E. coli B/r-A. Biotechnol Bioeng. 1984;26:203–216. Available from: https://doi.org/10.1002/bit.260260925

46. Doraiswamy LK, Sharma MM. Heterogeneous reactions: analysis, examples, and reactor design. New York: Wiley; 1984. Vol. 1–2. Available from: https://aiche.onlinelibrary.wiley.com/doi/10.1002/aic.690310822

47. Doran PM. Bioprocess engineering principles. Amsterdam: Elsevier; 1995. Available from: https://www.academia.edu/28336164/Bioprocess_Engineering_Principles

48. Drengstig T, Jolma IW, Ni XY, Thorsen K, Xu XM, Ruoff P. A basic set of homeostatic controller motifs. Biophys J. 2012;103:1–11. Available from: https://doi.org/10.1016/j.bpj.2012.09.033

49. Dutta R. Fundamentals of biochemical engineering. Berlin: Springer Verlag; 2008. Available from: https://www.scirp.org/reference/referencespapers?referenceid=927470

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

50. EcoCyc. Encyclopedia of Escherichia coli K-12 genes and metabolism. Menlo Park (CA): SRI International; 2005. Available from: http://ecocyc.org/

51. Edwards JS, Palsson BO. The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. Proc Natl Acad Sci U S A. 2000;97(10):5528–5533. Available from: https://doi.org/10.1073/pnas.97.10.5528

52. Edwards JS, Ibarra RU, Palsson BO. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nat Biotechnol. 2001;19:125–130. Available from: https://doi.org/10.1038/84379

53. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature. 2000;403:335–338. Available from: https://doi.org/10.1038/35002125

54. Forster J, Famili I, Fu P, Palsson BO, Nielsen J. Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. Genome Res. 2003;13:244–253. Available from: https://doi.org/10.1101/gr.234503

55. Frahm B, Lane P, Atzert H, Munack A, Hoffmann M, Hass VC, Pörtner R. Adaptive, model-based control by the open-loop feedback-optimal (OLFO) controller for the effective fed-batch cultivation of hybridoma cells. Biotechnol Prog. 2002;18(5):1095–1103. Available from: https://doi.org/10.1021/bp020035y

56. Franck UF. Feedback kinetics in physicochemical oscillators. Ber Bunsenges Phys Chem. 1980;84:334–341. Available from: https://doi.org/10.1002/bbpc.19800840407

57. Froment GF, Bischoff KB. Chemical reactor analysis and design. New York: Wiley; 1990.

58. Gavrilescu M, Chisti Y. Biotechnology – a sustainable alternative for the chemical industry. Biotechnol Adv. 2005;23:471–499. Available from: https://doi.org/10.1016/j.biotechadv.2005.03.004

59. Gibson MA, Bruck J. A probabilistic model of a prokaryotic gene and its regulation. In: Bower JM, Bolouri H, editors. Computational modelling of genetic and biochemical networks. Cambridge (MA): MIT Press; 2001. Available from: https://www.researchgate.net/publication/228388239_A_probabilistic_model_of_a_prokaryotic_gene_and_its_regulation

60. Gillespie DT, Mangel M. Conditioned averages in chemical kinetics. J Chem Phys. 1981;75:704–709. Available from: https://doi.org/10.1063/1.442111

61. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. J Phys Chem. 1977;81:2340–2361. Available from: https://pubs.acs.org/doi/10.1021/j100540a008

62. Goldbeter A. Biochemical oscillations and cellular rhythms. Cambridge: Cambridge University Press; 1996. Available from: https://doi.org/10.1017/CBO9780511608193

63. Gonze D. Coupling oscillations and switches in genetic networks. BioSystems. 2010;99:60–69. Available from: https://doi.org/10.1016/j.biosystems.2009.08.009

64. Grainger JNR, Bass L. A model of a growing steady state system. J Theor Biol. 1966;10:387–398. Available from: https://doi.org/10.1016/0022-5193(66)90135-4

65. Grainger JNR, Gaffney PE, West TT. A model of a growing steady-state system with a changing surface-volume ratio. J Theor Biol. 1968;21:123–130. Available from: https://doi.org/10.1016/0022-5193(68)90065-9

66. Guantes R, Poyatos JF. Dynamical principles of two-component genetic oscillators. PLoS Comput Biol. 2006;2(3):188–197. Available from: https://doi.org/10.1371/journal.pcbi.0020030

67. Hallett MB. The unpredictability of cellular behaviour: trivial or fundamental importance to cell biology. Perspect Biol Med. 1989;33:110–119.

68. Hammes GG, Wu CW. Kinetics of allosteric enzymes. Annu Rev Biophys Bioeng. 1974;3:133. Available from: https://doi.org/10.1146/annurev.bb.03.060174.000245

69. Haraldsdottir HS, Thiele I, Fleming RMT. Quantitative assignment of reaction directionality in a multicompartmental human metabolic reconstruction. Biophys J. 2012;102(8):1703–1711. Available from: https://doi.org/10.1016/j.bpj.2012.02.032

70. Hargrove JL, Schmidt FH. The role of mRNA and protein stability in gene expression. FASEB J. 1989;3:2360–2370. Available from: https://doi.org/10.1096/fasebj.3.12.2676679

71. Hatzimanikatis V, Floudas CA, Bailey JE. Analysis and design optimisation. AIChE J. 1996;42(5):1277–1292. Available from: https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690420509

72. Hatzimanikatis V, Lee KH, Bailey JE. A mathematical description of regulation of the G1-S transition of the mammalian cell cycle. Biotechnol Bioeng. 1999;65:631–637. Available from: https://doi.org/10.1002/(sici)1097-0290(19991220)65:6<631::aid-bit3>3.0.co;2-7

73. Heinemann M, Kummel A, Ruinatscha R, Panke S. In silico genome-scale reconstruction and validation of the Staphylococcus aureus metabolic network. Biotechnol Bioeng. 2005;92(7):850–864. Available from: https://doi.org/10.1002/bit.20663

74. Heinemann M, Panke S. Synthetic biology: putting engineering into biology. Bioinformatics. 2006;22:2790–2799. Available from: https://doi.org/10.1093/bioinformatics/btl469

75. Heinrich R, Schuster S. The regulation of cellular systems. New York: Chapman & Hall; 1996.

76. Heinrich R, Rapoport TA. A linear steady-state treatment of enzymatic chains. General properties, control, and effector strength. Eur J Biochem. 1974;42:89–95. Available from: https://doi.org/10.1111/j.1432-1033.1974.tb03318.x

77. Heinrich R, Rapoport SM, Rapoport TA. Metabolic regulation and mathematical models. Prog Biophys Mol Biol. 1977;32:1–82. Available from: https://pubmed.ncbi.

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

nlm.nih.gov/343173/

78. Heinzle E, Dunn IJ, Furukawa K, Tanner RD. Modelling of sustained oscillations observed in continuous culture of Saccharomyces cerevisiae. In: Aarne H, editor. Proceedings of the Modelling & Control of Biotechnical Process IFAC Conference; 1982 Aug 17–19; Helsinki, Finland.

79. Hempel DC. Development of biotechnological processes by integrating genetic and engineering methods. Eng Life Sci. 2006;6(5):443–447. Available from: https://doi.org/10.1002/elsc.200620149

80. Hlavacek WS, Savageau MA. Rules for coupled expression of regulator and effector genes in inducible circuits. J Mol Biol. 1996;255:121–139. Available from: https://doi.org/10.1006/jmbi.1996.0011

81. Hlavacek WS, Savageau MA. Completely uncoupled and perfectly coupled gene expression in repressible systems. J Mol Biol. 1997;266:538–558. Available from: https://doi.org/10.1006/jmbi.1996.0811

82. Hodgkin AL, Huxley AF. A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. J Physiol (Lond). 1952;117:500–544. Available from: https://doi.org/10.1113/jphysiol.1952.sp004764

83. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al. The Systems Biology Markup Language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics. 2003;19:524–531. Available from: https://doi.org/10.1093/bioinformatics/btg015

84. Hudder B, Yang Q, Bolfing B, Maria G, Morgan JJ, Lindahl PA. Computational modelling of iron metabolism in mitochondria. In: Proceedings of the NIST Conference on Mitochondrial Proteomics; 2002 Sep 17; Gaithersburg, MD, USA.

85. Human Genome Project. Wikipedia; 1990–2003. Last accessed Aug 5, 2025.

86. Hyver C, Le Guyader H. MPF and cyclin: modelling of the cell cycle minimum oscillator. BioSystems. 1990;24:85–90. Available from: https://doi.org/10.1016/0303-2647(90)90001-h

87. Jansen RC. Studying complex biological systems using multifactorial perturbation. Nat Rev Genet. 2003;4:145–151. Available from: https://doi.org/10.1038/nrg996

88. Ichikawa K. A-Cell: graphical user interface for the construction of biochemical reaction models. Bioinformatics. 2001;17(5):483–484. Available from: https://doi.org/10.1093/bioinformatics/17.5.483

89. Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet. 2001;2:343–372. Available from: https://doi.org/10.1146/annurev.genom.2.1.343

90. Iordache O, Maria G, Corbu S. Modelarea statistică şi estimarea parametrilor proceselor chimice (Statistical modelling and estimation of chemical process models). Bucharest: Romanian Academy Publishing House; 1991. (In Romanian). Available from: https://www.researchgate.net/publication/261767206_Statistical_Modelling_and_Estimation_of_Chemical_Process_Models

91. Jahan M, Maeda K, Matsuoka Y, Sugimoto Y, Kurata H. Development of an accurate kinetic model for the central carbon metabolism of Escherichia coli. Microb Cell Fact. 2016;15:112. Available from: https://doi.org/10.1186/s12934-016-0511-x

92. Joshi A, Palsson BO. Metabolic dynamics in the human red cell. I. A comprehensive kinetic model. J Theor Biol. 1989;141:515–528. Available from: https://doi.org/10.1016/s0022-5193(89)80233-4

93. Joshi A, Palsson BO. Metabolic dynamics in the human red cell. II. Interactions with the environment. J Theor Biol. 1989;141:529–545. Available from: https://doi.org/10.1016/s0022-5193(89)80234-6

94. Joshi A, Palsson BO. Metabolic dynamics in the human red cell. Metabolic reaction rates. J Theor Biol. 1990;142:41–68. Available from: https://doi.org/10.1016/s0022-5193(05)80012-8

95. Joshi A, Palsson BO. Metabolic dynamics in the human red cell. Data prediction and some model computations. J Theor Biol. 1990;142:69–85. Available from: https://doi.org/10.1016/s0022-5193(05)80013-x

96. Kacser H, Burns JA. The control of flux. Symp Soc Exp Biol. 1973;27:65–104. Available from: https://doi.org/10.1042/bst0230341

97. Kacser H, Beeby R. Evolution of catalytic proteins or on the origin of enzyme species by means of natural selection. J Mol Evol. 1984;20:38–51. Available from: https://doi.org/10.1007/bf02101984

98. Kauffman KJ, Pajerowski JD, Jamshidi N, Palsson BO, Edwards JS. Description and analysis of metabolic connectivity and dynamics in the human red blood cell. Biophys J. 2002;83:646–662. Available from: https://doi.org/10.1016/s0006-3495(02)75198-9

99. Kaznessis YN. Multi-scale models for gene network engineering. Chem Eng Sci. 2006;61:940–953. Available from: https://doi.org/10.1016/j.ces.2005.06.033

100. KEGG. Kyoto encyclopedia of genes and genomes. Kyoto: Kanehisa Laboratories, Bioinformatics Centre of Kyoto University; 2011. Available from: http://www.genome.jp/kegg/pathway.html

101. KEGG pathway. Phenylalanine, tyrosine, and tryptophan biosynthesis. Kyoto encyclopedia of genes and genomes. Kyoto: Kanehisa Laboratories, Bioinformatics Centre of Kyoto University; 2011. Available from: http://www.kegg.jp/keggbin/show_pathway?

102. Kholodenko BN, Schuster S, Garcia J, Westerhoff HV, Cascante M. Control analysis of metabolic systems involving quasi-equilibrium reactions. Biochim Biophys Acta. 1998;1379(3):337–352. Available from: https://doi.org/10.1016/s0304-4165(97)00114-1

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

103. Kholodenko BN. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. Eur J Biochem. 2000;267(6):1583–1588. Available from: https://doi.org/10.1046/j.1432-1327.2000.01197.x

104. Kholodenko BN, Kiyatkin A, Bruggeman FJ, Sontag E, Westerhoff HV, Hoek JB. Untangling the wires: a strategy to trace functional interactions in signalling and gene networks. Proc Natl Acad Sci U S A. 2002;99:12841–12846. Available from: https://doi.org/10.1073/pnas.192442699

105. Kinoshita A, Nakayama Y, Tomita M. In silico analysis of human erythrocyte using E-Cell system. In: Proceedings of the 2nd International Conference on Systems Biology; 2001 Nov 4–7; Pasadena, CA: California Institute of Technology.

106. Kiparissides A, Koutinas M, Kontoravdi C, Mantalaris A, Pistikopoulos EN. Closing the loop in biological systems modeling—from the in silico to the in vitro. Automatica. 2011;47:1147–1155. Available from: https://doi.org/10.1016/j.automatica.2011.01.013

107. Kitano H. Computational systems biology. Nature. 2002;420:206–210. Available from: https://doi.org/10.1038/nature01254

108. Kitano H. Systems biology: a brief overview. Science. 2002;295:1662–1664. Available from: https://doi.org/10.1126/science.1069492

109. Klein S, Heinzle E. Isotope labeling experiments in metabolomics and fluxomics. Wiley Interdiscip Rev Syst Biol Med. 2012;4(3):261–272. Available from: https://doi.org/10.1002/wsbm.1167

110. Klipp E, Nordlander B, Krüger R, Gennemark P, Hohmann S. Integrative model of the response of yeast to osmotic shock. Nat Biotechnol. 2005;23:975–982. Available from: https://doi.org/10.1038/nbt1114

111. Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ. Programmable cells: interfacing natural and engineered gene networks. Proc Natl Acad Sci U S A. 2004;101:8414–8419. Available from: https://doi.org/10.1073/pnas.0402940101

112. Komasilovs V, Pentjuss A, Elsts A, Stalidzans E. Total enzyme activity constraint and homeostatic constraint impact on the optimization potential of a kinetic model. BioSystems. 2017;162:128–134. Available from: https://doi.org/10.1016/j.biosystems.2017.09.016

113. Kubitschek HE. Cell volume increase in Escherichia coli after shifts to richer media. J Bacteriol. 1990;172:94–101. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC208405/

114. Kurata H, Sugimoto Y. Improved kinetic model of Escherichia coli central carbon metabolism in batch and continuous cultures. J Biosci Bioeng. 2017;17:S1389–1723. Available from: https://doi.org/10.1016/j.jbiosc.2017.09.005

115. Lambert DP. Ventilation criteria for IDMS facility. DOE contract DE-AC09-89SR18035 (Report WSRC-RP-96-0282). Aiken (SC): Westinghouse Savannah River Company; 1996. Available from: https://inis.iaea.org/records/tntx1-f8j66

116. Ledakowicz S, Becker U, Deckwer WD. Development of mercury biotransformation process in fluidised-bed reactor with immobilized microorganisms. In: Wijffels RH, Buitelaar RM, Bucke C, Tramper J, editors. Immobilized cells: basics and applications. Amsterdam: Elsevier; 1996.

117. Leonhäuser J, Röhricht M, Wagner-Döbler I, Deckwer WD. Reaction engineering aspects of microbial mercury removal. Eng Life Sci. 2006;6:139–148. Available from: https://doi.org/10.1002/elsc.200620904?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle

118. Leroy H. Statement given by the Centre of Mathematics Applied to the Life Sciences, University of Glasgow (UK). 2017. Available from: http://www.gla.ac.uk/research/az/cmals/research/systemsbio logystochasticmodelling/

119. Levenspiel O. Chemical reaction engineering. 3rd ed. New York: Wiley; 1999. Available from: https://www.scribd.com/document/535709235/Edited-Levenspiel

120. Li G, Rabitz H. A general analysis of approximate lumping in chemical kinetics. Chem Eng Sci. 1990;45:977–1002. Available from: https://doi.org/10.1016/0009-2509(90)85020-E

121. Li G, Rabitz H. Determination of constrained lumping schemes for nonisothermal first-order reaction systems. Chem Eng Sci. 1991;46:583–596. Available from: https://doi.org/10.1016/0009-2509(91)80018-T

122. Li G, Rabitz H. New approaches to determination of constrained lumping schemes for a reaction system in the whole composition space. Chem Eng Sci. 1991;46:95–111. Available from: https://ui.adsabs.harvard.edu/abs/1991ChEnS..46...95L/abstract

123. Liese A, Seelbach K, Wandrey C, editors. Industrial biotransformations. Weinheim: Wiley-VCH; 2006. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/3527608184

124. Lindon JC, Holmes E, Nicholson JK. Metabolomics and its role in drug development and disease diagnosis. Expert Rev Mol Diagn. 2004;4:189–199. Available from: https://doi.org/10.1586/14737159.4.2.189

125. Lodish H, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scott MP, Zipursky L, Darnell J. Molecular cell biology. New York: Freeman & Co.; 2000.

126. Loeblein C, Perkins J, Srinivasan B, Bonvin D. Performance analysis of online batch optimisation systems. Comput Chem Eng. 1997;21:S867–S872. Available from: https://doi.org/10.1016/S0098-1354(97)87611-9

127. Lopresti D. Introduction to bioinformatics. Course material, Department of Computer Science and Engineering, Lehigh University; 2010. Available from: https://www.lehigh.edu/~inbios21/PDF/Fall2009/Lopresti11132009.pdf

128. Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? An introduction and overview. Yearb Med Inform. 2001:83–99. Available from: https://pubmed.ncbi.nlm.nih.gov/27701604/

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

129. Mackay D. Multimedia environmental models - The fugacity approach. Boca Raton: Lewis Publ.; 2001. Available from: https://doi.org/10.1201/9781420032543

130. Malinowski ER. Factor analysis in chemistry. New York: Wiley; 1991. Available from: https://www.scirp.org/(S(351jmbnt-vnsjt1aadkozje))/reference/referencespapers?referenceid=2899640

131. Mao L, Nicolae A, Oliveira MAP, He F, Hachi S, Fleming RMT. A constraint-based modelling approach to metabolic dysfunction in Parkinson's disease. Comput Struct Biotechnol J. 2015;13:484–491. Available from: https://doi.org/10.1016/j.csbj.2015.08.002

132. Maria G. An adaptive strategy for solving kinetic model concomitant estimation-reduction problems. Can J Chem Eng. 1989;67:825–832. Available from: https://doi.org/10.1002/cjce.5450670514?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle

133. Maria G, Rippin DWT. A note concerning two techniques for complex kinetic pathway analysis. Chem Eng Sci. 1993;48:3855–3864. Available from: https://doi.org/10.1016/0009-2509(93)80228-I

134. Maria G. Adaptive random search and short-cut techniques for process model identification and monitoring. In: Pekny JF, Blau GE, editors. Proceedings of the FOCAPO98 International Conference on Foundations of Computer Aided Process Operations; 1998 Jul 5–10; Snowbird, UT. Ann Arbor (MI): AIChE Publications, University of Michigan; 1998;351–359.

135. Maria G. Simulation of hypothetical "Mechanical Cells" that maintain intracellular homeostasis while growing autocatalytically on environmental nutrients present in variable amounts. Seminar presented at: Texas A&M University, Department of Chemistry (Prof. P. Lindahl group); 2002 Aug 23; College Station, TX. Personal communication.

136. Maria GC, Morgan JJ, Lindahl PA. Kinetic simulation of hypothetical "Mechanical Cells" that maintain intracellular homeostasis while growing autocatalytically on environmental nutrients present in variable amounts. J Theor Biol. 2002. Manuscript submitted Aug 6, 2002; rejected Sep 2002.

137. Maria G. Evaluation of protein regulatory kinetics schemes in perturbed cell growth environments by using sensitivity methods. Chem Biochem Eng Q. 2003;17(2):99–117. Available from: http://silverstripe.fkit.hr/cabeq/assets/Uploads/Cabeq-2003-02-1.pdf

138. Maria G. ARS combination with an evolutionary algorithm for solving MINLP optimisation problems. In: Hamza MH, editor. Modelling, Identification and Control. Anaheim (CA): IASTED/ACTA Press; 2003. p. 112–118. Available from: https://dblp.org/rec/conf/mic/Maria03.html

139. Maria G. A review of algorithms and trends in kinetic model identification for chemical and biochemical systems. Chem Biochem Eng Q. 2004;18(3):195–222. Available from: http://silverstripe.fkit.hr/cabeq/assets/Uploads/CABEQ-18-3-1.pdf

140. Maria G. Modular-based modelling of protein synthesis regulation. Chem Biochem Eng Q. 2005;19:213–233. Available from: http://silverstripe.fkit.hr/cabeq/past-issues/article/505

141. Maria G. Relations between apparent and intrinsic kinetics of programmable drug release in human plasma. Chem Eng Sci. 2005;60:1709–1723. Available from: https://doi.org/10.1016/j.ces.2004.11.009

142. Maria G. Application of lumping analysis in modelling the living systems—a trade-off between simplicity and model quality. Chem Biochem Eng Q. 2006;20:353–373. Available from: https://www.researchgate.net/publication/27191761_Application_of_Lumping_Analysis_inModelling_the_Living_Systems-A_Trade-off_Between_Simplicity_and_Model_Quality

143. Maria G. Modelling bistable genetic regulatory circuits under variable volume framework. Chem Biochem Eng Q. 2007;21:417–434. Available from: http://silverstripe.fkit.hr/cabeq/assets/Uploads/Cabeq-2007-04-13.pdf

144. Maria G. Reduced modular representations applied to simulate some genetic regulatory circuits. Rev Chim. 2008;59:318–324. Available from: https://doi.org/10.37358/RC.08.3.1756

145. Maria G. Analiza statistică şi corelarea datelor experimentale (bio)chimice. Repartiţii şi estimatori statistici (Statistical data analysis and correlations. Distributions and estimators). Bucharest: Printech Publishing; 2008.

146. Maria G. Building up lumped models for a bistable genetic regulatory circuit under whole-cell modelling framework. Asia Pac J Chem Eng. 2009;4:916–928. Available from: https://doi.org/10.1002/apj.297?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle

147. Maria G. A whole-cell model to simulate the mercuric ion reduction by E. coli under stationary and perturbed conditions. Chem Biochem Eng Q. 2009;23(3):323–341. Available from: http://silverstripe.fkit.hr/cabeq/past-issues/article/295

148. Maria G. A whole-cell approach in modelling genetic circuits in living cells. New Biotechnol. 2009;25:S338. Available from: https://doi.org/10.1016/j.nbt.2009.06.818

149. Maria G, Dan A, Stefan DN. Model-based derivation of the safety operating limits of a semi-batch reactor for the catalytic acetoacetylation of pyrrole using a generalised sensitivity criterion. Chem Biochem Eng Q. 2010;24(3):265–281. Available from: https://scispace.com/pdf/model-based-derivation-of-the-safety-operating-limits-of-a-298l4kz8st.pdf

150. Maria G. A dynamic model to simulate the genetic regulatory circuit controlling the mercury ion uptake by E. coli cells. Rev Chim (Bucharest). 2010;61(2):172–186. Available from: https://www.researchgate.net/publication/259477296_A_Dynamic_Model_to_Simulate_the_Genetic_Regulatory_Circuit_Controlling_the_Mercury_Ion_Uptake_by_E_coli_cells

151. Maria G, Xu Z, Sun J. Multi-objective MINLP optimisation used to identify theoretical gene knockout strategies for E. coli cell. Chem Biochem Eng Q. 2011;25(4):403–424. Available from: https://www.researchgate.net/publication/268370381_Multi-objective_MINLP_Optimization_Used_to_Identify_Theoretical_Gene_Knockout_Strategies_for_E_coli_Cell

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

152. Maria G. Enzymatic reactor selection and derivation of the optimal operation policy by using a model-based modular simulation platform. Comput Chem Eng. 2012;36(1):325–341. Available from: https://doi.org/10.1016/j.compchemeng.2011.06.006

153. Maria G, Luta I, Maria C. Model-based sensitivity analysis of a fluidised-bed bioreactor for mercury uptake by immobilised Pseudomonas putida cells. Chem Pap. 2013;67(11):1364–1375. Available from: https://www.degruyterbrill.com/document/doi/10.2478/s11696-013-0403-z/html?srsltid=AfmBOopwUUtTLbMWrzPDgs1bMV6POrixxSsAzAlFdlyDx8ZdTS8Xvw-0

154. Maria G, Luta I. Structured cell simulator coupled with a fluidised bed bioreactor model to predict the adaptive mercury uptake by E. coli cells. Comput Chem Eng. 2013;58:98–115. Available from: https://doi.org/10.1016/j.compchemeng.2013.06.004

155. Maria G. Insilico derivation of a reduced kinetic model for stationary or oscillating glycolysis in Escherichia coli bacterium. Chem Biochem Eng Q. 2014;28(4):509–529. Available from: https://doi.org/10.15255/CABEQ.2014.2002

156. Maria G. Extended repression mechanisms in modelling bistable genetic switches of adjustable characteristics within a variable cell volume modelling framework. Chem Biochem Eng Q. 2014;28(1):35–51. Available from: https://files01.core.ac.uk/download/pdf/19892014.pdf

157. Maria G, Crisan M, Maria C. Estimarea parametrilor modelelor cinetice ale proceselor (bio)chimice (Parameter estimation of the (bio)chemical process kinetic models). Bucharest: Printech Publishing; 2016. Available from: https://www.researchgate.net/publication/319526546_Estimarea_parametrilor_modelelor_cinetice_ale_proceselor_biochimice_Parameter_estimation_of_the_Biochemical_process_kinetic_models_in_Romanian

158. Maria G. Application of (bio)chemical engineering principles and lumping analysis in modelling the living systems. Curr Trends Biomed Eng Biosci. 2017;1(4):CTBEB.MS.ID.555566. Available from: https://juniperpublishers.com/ctbeb/pdf/CTBEB.MS.ID.555566.pdf

159. Maria G. A review of some novel concepts applied to modular modelling of genetic regulatory circuits. Irvine (CA): Juniper Publishers; 2017. Available from: https://juniperpublishers.com/ebooks/A%20Review%20of%20Some%20Novel%20Concepts%20Applied%20to%20Modular%20Modelling%20of%20Genetic%20Regulatory%20Circuits.pdf

160. Maria G. Deterministic modelling approach of metabolic processes in living cells – a still powerful tool for representing the metabolic process dynamics. Irvine (CA): Juniper Publishers; 2017. Available from: https://juniperpublishers.com/ebooks/Deterministic%20Modelling%20Approach%20of%20Metabolic%20Processes%20in%20Living%20Cells%20-%20A%20Still%20Powerful%20Tool%20for%20Representing%20the%20Metabolic%20Processes%20Dynamics.pdf

161. Maria G, Maria C, Tociu C. 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 Ser B. 2017;79(3):3–19. Available from: https://www.researchgate.net/publication/321027517_Comments_on_two_novel_review_ebooks_in_the_area_of_deterministic_modelling_of_metabolic_processes_and_of_genetic_regulatory_circuits_in_living_cells

162. Maria G, Scoban AG. Optimal operating policy of a fluidised bed bioreactor used for mercury uptake from wastewaters by using immobilised P. putida cells. Curr Trends Biomed Eng Biosci. 2017;2(4):555594. Available from: https://juniperpublishers.com/ctbeb/pdf/CTBEB.MS.ID.555594.pdf

163. Maria G. In-silico design of genetically modified micro-organisms (GMO) of industrial use, by using systems biology and (bio)chemical engineering tools. Irvine (CA): Juniper Publishers; 2018. Available from: https://juniperpublishers.com/ebooks/In-Silico%20Design%20of%20Genetic%20Modified%20Micro-Organisms%20(GMO)%20of%20Industrial%20Use,%20by%20Using%20Systems%20Biology%20and.pdf

164. Maria G, Gijiu CL, Maria C, Tociu C. Interference of the oscillating glycolysis with the oscillating tryptophan synthesis in the E. coli cells. Comput Chem Eng. 2018;108:395–407. Available from: https://doi.org/10.1016/j.compchemeng.2017.10.003

165. Maria G, Mihalachi M, Gijiu CL. Model-based identification of some conditions leading to glycolytic oscillations in E. coli cells. Chem Biochem Eng Q. 2018;32:523–533.

166. Maria G, Mihalachi M, Gijiu CL. Chemical engineering tools applied to simulate some conditions producing glycolytic oscillations in E. coli cells. UPB Sci Bull Ser B Chem. 2018;80:27–38. Available from: https://www.scientificbulletin.upb.ro/rev_docs_arhiva/full6c1_852902.pdf

167. Maria G. 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): Juniper Publishers; 2018. Available from: https://juniperpublishers.com/ebooks/From%20Residual%20Biomass%20and%20Inferior%20Quality%20Coal%20to%20the%20Synthesis%20of%20Methanol%20and%20then%20to%20Hydrocarbons%20and%20Gasoline%20%E2%80%93%20A%20Romanian%20Project%20of%20High%20Success.pdf

168. Maria G, Scoban AG. Setting some milestones when modelling gene expression regulatory circuits under a variable-volume whole-cell modelling framework. 1. Generalities. Rev Chim (Bucharest). 2017;68(12):3027–3037. Available from: https://revistadechimie.ro/Articles.asp?ID=6031

169. Maria G, Scoban AG. Optimal operating policy of a fluidised bed bioreactor used for mercury uptake from wastewaters by using immobilised P. putida cells. Curr Trends Biomed Eng Biosci. 2017;2(4):555594. Available from: https://juniperpublishers.com/ctbeb/pdf/CTBEB.MS.ID.555594.pdf

170. Maria G, Scoban AG. Setting some milestones when modelling gene expression regulatory circuits under a variable-volume whole-cell modelling framework. 2. Case studies. Rev Chim (Bucharest). 2018;69(1):259–266. Available from: https://revistadechimie.ro/Articles.asp?ID=6031

171. Maria G, Mihalachi M, Gijiu CL. In silico optimisation of a bioreactor with an E. coli culture for tryptophan production by using a structured model coupling the oscillating glycolysis and tryptophan synthesis. Chem Eng Res Des. 2018;135:207–221. Available from: https://doi.org/10.1016/j.cherd.2018.05.011

172. Maria G, Gijiu CL, Maria C, Tociu C, Mihalachi M. Importance of considering the isotonic system hypothesis when modelling the self-control of gene expression regulatory modules in living cells. Curr Trends Biomed Eng Biosci. 2018;12(2):CTBEB.MS.ID.555833. Available from: https://ideas.repec.org/a/adp/jctbeb/v12y2018i2p29-48.html

173. Maria G, Maria C, Tociu C. A comparison between two approaches used for deterministic modelling of metabolic processes and of genetic regulatory circuits in

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

living cells. UPB Sci Bull Ser B Chem. 2018;80(1):127–144. Available from: https://www.researchgate.net/publication/323512658_A_comparison_between_two_approaches_used_for_deterministic_modelling_of_metabolic_processes_and_of_genetic_regulatory_circuits_in_living_cells

174. Maria G, Gijiu CL, Crisan M, Maria C, Tociu C. Model-based re-design of some genetic regulatory circuits to get genetically modified micro-organisms (GMO) by using engineering computational tools (a mini-review). Curr Trends Biomed Eng Biosci. 2019;18(3):CTBEB.MS.ID.555988. Available from: https://juniperpublishers.com/ctbeb/CTBEB.MS.ID.555988.php

175. Maria G. Algoritmi numerici de simplificare a modelelor cinetice ale proceselor chimice şi biochimice (Numerical methods to reduce the kinetic models of (bio) chemical processes). Bucharest: Printech Publishing; 2019.

176. Maria G. Model-based optimization of a fed-batch bioreactor for mAb production using a hybridoma cell culture. Molecules. 2020;25(23):5648–5674. Available from: https://doi.org/10.3390/molecules25235648

177. Maria G. 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 Publishing; 2020. Available from: https://www.researchgate.net/publication/344991913_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_numbe

178. Maria G. In-silico determination of some conditions leading to glycolytic oscillations and their interference with some other processes in E. coli cells. Front Chem. 2020;8:526679–526693. Available from: https://doi.org/10.3389/fchem.2020.526679

179. Maria G. A CCM-based modular and hybrid kinetic model to simulate the tryptophan synthesis in a fed-batch bioreactor using modified E. coli cells. Comput Chem Eng. 2021;153:107450–107466. Available from: https://doi.org/10.1016/j.compchemeng.2021.107450

180. Maria G, Renea L. 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. 2021;8(12):210–247. Available from: https://doi.org/10.3390/bioengineering8120210

181. Maria G, Maria C, Renea L. Application of (bio-)chemical engineering concepts and rules in bioinformatics. Review of a CCM-based modular and hybrid kinetic model used to simulate and optimise a bioreactor with genetically modified cells. Curr Trends Biomed Eng Biosci. 2022;20(3):556039. Available from: https://juniperpublishers.com/ctbeb/CTBEB.MS.ID.556039.php

182. Maria G. Hybrid modular kinetic models linking cell-scale structured CCM reaction pathways to bioreactor macro-scale state variables. Applications for solving bioengineering problems. Irvine (CA): Juniper Publishers; 2023.

183. Maria G. Application of (bio)chemical engineering concepts and tools to model GRCs, and some essential CCM pathways in living cells. Part 1. Generalities. Curr Trends Biomed Eng Biosci. 2024;22(1):556080–556104. Available from: https://juniperpublishers.com/ctbeb/pdf/CTBEB.MS.ID.556080.pdf

184. Maria G. In-silico determination of some conditions leading to glycolytic oscillations and their interference with tryptophan synthesis in E. coli cells. In: Sosa FC, editor. Research advances in microbiology and biotechnology. Vol. 5. London: B.P. International; 2023. p. 16–55. Available from: https://doi.org/10.9734/bpi/ramb/v5/5067C

185. Research advances in microbiology and biotechnology. Vol. 5 [Internet]. Available from: https://www.bookpi.org/bookstore/product/research-advances-in-microbiology-and-biotechnology-vol-5/

186. Research advances in microbiology and biotechnology. Vol. 5, article view [Internet]. Available from: https://stm.bookpi.org/RAMB-V5/article/view/10480

187. Maria G. Application of (bio)chemical engineering concepts and tools to model GRCs, and some essential CCM pathways in living cells. Part 2. Mathematical modelling framework. Ann Rev Res. 2024;10(3):555790–555835. Available from: https://doi.org/10.19080/ARR.2024.10.555790

188. Maria G. 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. Arch Biomed Eng Biotechnol. 2024;7(5):ABEB MS.ID.000672. Available from: https://irispublishers.com/abeb/pdf/ABEB.MS.ID.000671.pdf

189. Maria G. 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 GMOs. Ann Syst Biol. 2024;7(1):001–034. Article ID: ASB-7121. Available from: https://juniperpublishers.com/ctbeb/pdf/CTBEB.MS.ID.556080.pdf

190. Maria G, Gijiu CL, Renea L, Gheorghe D. In-silico optimal operating policies of a batch or a fed-batch bioreactor for mAb production using a hybridoma cell culture. Rev Roum Chim. 2024;69(5–6):263–278. Available from: https://juniperpublishers.com/ctbeb/CTBEB.MS.ID.556105.php

191. Maria G, Renea L, Gheorghe D, Muscalu C. In-silico optimisation of a fluidised-bed bioreactor for ethanol production by using several algorithms and operating alternatives. Rev Roum Chim. 2025;70(11–12). In press.

192. Maria G, Gheorghe D. In-silico optimisation of a fed-batch bioreactor for tryptophan production by using a structured hybrid model and several algorithms, including a Pareto optimal front. Algorithms. 2024;17:428–468. Available from: https://doi.org/10.3390/a17100428

193. Matsuura T, Tanimura N, Hosoda K, Yomo T, Shimizu Y. Reaction dynamics analysis of a reconstituted Escherichia coli protein translation system by computational modelling. Proc Natl Acad Sci U S A. 2017. Available from: www.pnas.org/cgi/doi/10.1073/pnas.1615351114

194. McAdams HH, Shapiro L. Circuit simulation of genetic networks. Science. 1995;269:650–656. Available from: https://doi.org/10.1126/science.7624793

195. McAdams HH, Arkin A. Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A. 1997;94:814–819. Available from: https://doi.org/10.1073/pnas.94.3.814

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

196. McAdams HH, Arkin A. Simulation of prokaryotic genetic circuits. Annu Rev Biophys Biomol Struct. 1998;27:199–224. Available from: https://doi.org/10.1146/annurev.biophys.27.1.199

197. MCA Web. The Metabolic Control Analysis Web. Manchester: The University of Manchester; 2004. Available from: https://dbkgroup.org/metabolic-control-analysis/

198. Mesarovic M. System theory and biology. Berlin: Springer-Verlag; 1968. Available from: https://books.google.co.in/books/about/Systems_Theory_and_Biology.html?id=5ra00QEACAAJ&redir_esc=y

199. Michal G. Biochemical pathways in a cell. Basel (CH): Roche; 2014. Available from: https://www.roche.com/about/philanthropy/science-education/biochemical-pathways

200. Miskovic L, Tokic M, Fengos G, Hatzimanikatis V. Rites of passage: requirements and standards for building kinetic models of metabolic phenotypes. Curr Opin Biotechnol. 2015;36:146–153. Available from: https://www.roche.com/about/philanthropy/science-education/biochemical-pathways

201. Mizuno T. Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of Escherichia coli. DNA Res. 1997;4:161–168.

202. Morgan JJ, Surovtsev IV, Lindahl PA. A framework for whole-cell mathematical modelling. J Theor Biol. 2004;231:581–596. Available from: https://doi.org/10.1093/dnares/4.2.161

203. Morgan JJ, Surovtsev IV, Lindahl PA. A framework for whole-cell mathematical modelling. J Theor Biol. 2004;231:581–596. Available from: https://doi.org/10.1016/j.jtbi.2004.07.014

204. Mori H. From the sequence to cell modelling: comprehensive functional genomics in E. coli. J Biochem Mol Biol. 2004;37:83–92. Available from: https://doi.org/10.5483/bmbrep.2004.37.1.083

205. Morton-Firth CJ, Bray D. Predicting temporal fluctuations in an intracellular signalling pathway. J Theor Biol. 1998;192:117–128. Available from: https://doi.org/10.1006/jtbi.1997.0651

206. Moser A. Bioprocess technology: kinetics and reactors. New York: Springer-Verlag; 1988. Available from: https://www.scribd.com/document/358679552/Bioprocess-Technology-Kinetics-and-Reactors-by-Professor-Dr-Anton-Moser

207. Moulijn JA, Makkee M, van Diepen A. Chemical process technology. New York: Wiley; 2001. Available from: https://download.e-bookshelf.de/download/0003/7377/55/L-G-0003737755-0002941059.pdf

208. Myers CJ. Engineering genetic circuits. Boca Raton (FL): Chapman & Hall/CRC Press; 2009. Available from: https://doi.org/10.1201/b16381

209. Nandy SK. BioXYZ, what is XYZ? Curr Trends Biomed Eng Biosci. 2017;3(2):555606. Available from: https://juniperpublishers.com/ctbeb/CTBEB.MS.ID.555606.php

210. Narasimhan B, Mallapragada SK, Peppas NA. Release kinetics, data interpretation. In: Mathiowitz E, editor. Encyclopedia of controlled drug delivery. New York: Wiley; 1999. p. 921–935.

211. Nemenman I, Escola GS, Hlavacek WS, Unkefer PJ, Unkefer CJ, Wall ME. Reconstruction of metabolic networks from high-throughput metabolite profiling data: in silico analysis of red blood cell metabolism. Ann N Y Acad Sci. 2007;1115:102–115. Available from: https://doi.org/10.1196/annals.1407.013

212. Nielsen J. Metabolic engineering: techniques for analysis of targets for genetic manipulations. Biotechnol Bioeng. 1998;58:125–132. Available from: https://doi.org/10.1002/(sici)1097-0290(19980420)58:2/3%3C125::aid-bit3%3E3.0.co;2-n

213. National Institutes of Health. National Institute of Health database [Internet]. 2004. Available from: https://www.nih.gov/

214. Niklas J, Schrader E, Sandig V, Noll T, Heinzle E. Quantitative characterisation of metabolism and metabolic shifts during growth of the new human cell line AGE1.HN using time-resolved metabolic flux analysis. Bioprocess Biosyst Eng. 2011;34:533–545. Available from: https://doi.org/10.1007/s00449-010-0502-y

215. Noble D. A modification of the Hodgkin-Huxley equations applicable to Purkinje fibre action and pace-maker potentials. J Physiol. 1962;160:317–352. Available from: https://doi.org/10.1113/jphysiol.1962.sp006849

216. Noble D. Modelling the heart—from genes to cells to the whole organ. Science. 2002;295:1678. Available from: https://doi.org/10.1126/science.1069881

217. Noble D. The music of life. Oxford: Oxford University Press; 2006. Available from: https://global.oup.com/academic/product/the-music-of-life-9780199295739?cc=in&lang=en&

218. Noble D. The music of life: sourcebook [Internet]. Available from: http://www.musicoflife.website/pdfs/The%20Music%20of%20Lifesourcebook.pdf

219. Norel R, Agur Z. A model for the adjustment of the mitotic clock by cyclin and MPF levels. Science. 1991;251:1076–1078. Available from: https://doi.org/10.1126/science.1825521

220. Novak B, Tyson JJ. Modelling the control of DNA replication in fission yeast. Proc Natl Acad Sci U S A. 1997;94:9147–9152. Available from: https://doi.org/10.1073/pnas.94.17.9147

221. Obeyesekere MN, Zimmerman SO, Tecarro ES, Auchmuty G. A model of cell cycle behaviour dominated by kinetics of a pathway simulated by growth factors. Bull Math Biol. 1999;61:917–934. Available from: https://doi.org/10.1006/bulm.1999.0118

222. Olivier BG, Snoep JL. JWS Online and the Silicon Cell project: web-based kinetic modelling using JWS Online. Bioinformatics. 2004;20:2143–2144. Available from: https://doi.org/10.1093/bioinformatics/bth200

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

223. Palsson BO. The challenges of in silico biology. Nat Biotechnol. 2000;18:1147–1150. Available from: https://doi.org/10.1038/81125

224. Papin JA, Price ND, Wiback SJ, Fell DA, Palsson BO. Metabolic pathways in the post-genome era. Trends Biochem Sci. 2003;28(5):250–258. Available from: https://doi.org/10.1016/s0968-0004(03)00064-1

225. Perret CJ, Levey HC. The theory of uncatalysed linear expanding systems. J Theor Biol. 1961;1:542–550. Available from: https://doi.org/10.1016/S0022-5193(25)00268-1

226. Peters M, Eicher JJ, Van Niekerk DD, Waltemath D, Snoep JL. The JWS online simulation database. Bioinformatics. 2017;33(10):1589–1590. Available from: https://doi.org/10.1093/bioinformatics/btw831

227. Le Phillip P, Bahl A, Ungar LH. Using prior knowledge to improve genetic network reconstruction from microarray data. In Silico Biol. 2004;4:335–353. Available from: https://pubmed.ncbi.nlm.nih.gov/15724284/

228. Philippidis GP, Malmberg LH, Hu WS, Schottel JL. Effect of gene amplification on mercuric ion reduction activity of Escherichia coli. Appl Environ Microbiol. 1991;57:3558–3564. Available from: https://doi.org/10.1128/aem.57.12.3558-3564.1991

229. Philippidis GP, Schottel JL, Hu WS. Mathematical modelling and optimisation of complex biocatalysis: a case study of mercuric reduction by Escherichia coli. In: Expression systems and processes for DNA products. National Science Foundation Report ECE-8552670. Minneapolis (MN): University of Minnesota; 1991.

230. Philippidis GP, Schottel JL, Hu WS. A model for mercuric ion reduction in recombinant Escherichia coli. Biotechnol Bioeng. 1991;37:47–54. Available from: https://doi.org/10.1002/bit.260370108

231. Pörtner R, Schäfer T. Modelling hybridoma cell growth and metabolism—a comparison of selected models and data. J Biotechnol. 1996;49(1–3):119–135. Available from: https://doi.org/10.1016/0168-1656(96)01535-0

232. Price ND, Reed JL, Palsson BO. Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat Rev Microbiol. 2004;2:886–897. Available from: https://doi.org/10.1038/nrmicro1023

233. Ptashne M. A genetic switch: phage lambda and higher organisms. Oxford (UK): Cell Press & Blackwell Scientific Publications; 1992. Available from: https://repository.library.georgetown.edu/handle/10822/545304

234. Puga A, Wallace K. Molecular biology of the toxic response. Boca Raton (FL): CRC Press; 1998. Available from: https://books.google.mv/books?id=TehsK8DX-7YC&printsec=copyright#v=onepage&q&f=false

235. Qian Y, McBride C, Del Vecchio D. Programming cells to work for us. Annu Rev Control Robot Auton Syst. 2017. Available from: https://www.annualreviews.org/content/journals/10.1146/annurev-control-060117-105052

236. Rao BM, Lauffenburger DA, Wittrup KD. Integrating cell-level kinetic modelling into the design of engineered protein therapeutics. Nat Biotechnol. 2005;23:191–194. Available from: https://doi.org/10.1038/nbt1064

237. Reijenga KA, Westerhoff HV, Kholodenko BN, Snoep JL. Control analysis for autonomously oscillating biochemical networks. Biophys J. 2002;82:99–108. Available from: https://doi.org/10.1016/s0006-3495(02)75377-0

238. Renea L, Maria G, Gheorghe D. In-silico derived optimal operation policies of a fed-batch bioreactor to maximise the mAbs production by using Pareto optimal techniques. Rev Roum Chim. 2025;70(5–6):261–278. Available from: https://doi.org/10.33224/rrch.2025.70.5-6.03

239. Rensing L, Meyer-Grahle U, Ruoff P. Biological timing and the clock metaphor: oscillatory and hourglass mechanisms. Chronobiol Int. 2001;18:329–369. Available from: https://doi.org/10.1081/cbi-100103961

240. Riznichenko GY. Mathematical models in biophysics: teaching notes. Moscow: Biological Faculty, Lomonosov Moscow State University; 2002. Available from: www.biophysics.org/btol/img/math-models.pdf

241. Rocha I, Maia P, Evangelista P, Vilaça P, Soares S, Pinto JP, et al. Nielsen J, Patil KR, Ferreira EC, Rocha M. OptFlux: an open-source software platform for in silico metabolic engineering. BMC Syst Biol. 2010;4:45. Available from: https://doi.org/10.1186/1752-0509-4-45

242. Rodriguez-Prados JC, de Atauri P, Maury J, Ortega F, Portais JC, Chassagnole C, et al. Acerenza L, Lindley ND, Cascante M. In Silico strategy to rationally engineer metabolite production: A case study for threonine in Escherichia coli. Biotechnol Bioeng. 2009;103:609–620. Available from: https://doi.org/10.1002/bit.22271

243. Salis H, Kaznessis Y. Numerical simulation of stochastic gene circuits. Comput Chem Eng. 2005;29:577-588. Available from: https://doi.org/10.1016/j.compchemeng.2004.08.017

244. San KY, Stephanopoulos G. Optimisation of fed-batch penicillin fermentation: a case of singular optimal control with state constraints. Biotechnol Bioeng. 1989;34(1):72-8. Available from: https://doi.org/10.1002/bit.260340110

245. Sauro HM, Kholodenko BN. Quantitative analysis of signalling networks. Prog Biophys Mol Biol. 2004;86:5-43. Available from: https://doi.org/10.1016/j.pbiomolbio.2004.03.002

246. Savageau MA, Voit EO. Recasting nonlinear differential equations as S-systems: A canonical nonlinear form. Math Biosci. 1987;87:83-115. Available from: https://doi.org/10.1016/0025-5564(87)90035-6

247. Savageau MA. Design principles for elementary gene circuits: elements, methods, and examples. Chaos. 2001;11:142-159. Available from: https://doi.org/10.1063/1.1349892

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

248. Savageau MA. Alternatives designs for a genetic switch: Analysis of switching times using the piecewise power-law representation. Math Biosci. 2002;180:237-253. Available from: https://doi.org/10.1016/s0025-5564(02)00113-x

249. Schaff JC, et al. Virtual Cell (V-Cell) Project. 1st Intl Symp Comput Cell Biol. 2001. NRCAM (National Resource for Cell Analysis and Modelling) of the NIH. Available from: http://www.nrcam.uchc.edu/

250. Schellenberger J, Que R, Fleming RM, Thiele I, Orth JD, Feist AM, et al. Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, et al. Quantitative prediction of cellular metabolism with constraint-based models: the cobra toolbox v2.0. Nat Protoc. 2011;6(9):1290-307. Available from: https://doi.org/10.1038/nprot.2011.308

251. Schilling CH, Letscher D, Palsson BO. I. Theory for the systematic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective; II. Assessment of the metabolic capabilities of Haemophilus influenzae Rd through a genome-scale pathway analysis. J Theor Biol. 2000;203:229-283. Available from: https://doi.org/10.1006/jtbi.2000.1073

252. Schilling CH, Edwards JS, Letscher D, Palsson BO. Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. Biotechnol Bioeng. 2000;71:286-306. Available from: https://pubmed.ncbi.nlm.nih.gov/11291038/

253. Schrödinger E. What is life? The physical aspect of the living cell. Cambridge: Cambridge University Press; 1944. Available from: https://www.arvindguptatoys.com/arvindgupta/whatislife-schrodinger.pdf

254. Schuetz R, Kuepfer L, Sauer W. Systematic evaluation of objective functions for predicting intracellular fluxes in E. coli. Mol Syst Biol. 2007;3:119. Available from: https://doi.org/10.1038/msb4100162

255. Schuster S, Fell DA, Dandekar T. A general definition of metabolic pathways useful for systematic organisation and analysis of complex metabolic networks. Nat Biotechnol. 2000;18:326-332. Available from: https://doi.org/10.1038/73786

256. Scoban AG, Maria G. Model-based optimisation of the feeding policy of a fluidised bed bioreactor for mercury uptake by immobilised P. putida cells. Asia Pac J Chem Eng. 2016;11(5):721-734. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/apj.2003

257. Segel IH. Enzyme kinetics (behaviour and analysis of rapid equilibrium and steady-state enzyme systems). New York: Wiley; 1993. Available from: https://www.wiley.com/en-us/Enzyme+Kinetics%3A+Behavior+and+Analysis+of+Rapid+Equilibrium+and+Steady-State+Enzyme+Systems-p-9780471303091

258. Sewell C, Morgan J, Lindahl P. Analysis of protein regulatory mechanisms in perturbed environments at steady state. J Theor Biol. 2002;215:151-167. Available from: https://doi.org/10.1006/jtbi.2001.2536

259. Shimizu TS, Bray D. Computational cell biology - The stochastic approach. Cambridge: MIT Press; 2002. Available from: https://www.semanticscholar.org/paper/10-Computational-Cell-Biology-The-Stochastic-Shimizu-Bray/99254dab4778a34ccd2d41cbba02021fc7c8e044

260. Shuler ML. Chemical engineering problems in biotechnology. New York: AIChE; 1989. Available from: https://books.google.com.jm/books?id=FjMeVWP_TAoC&printsec=frontcover&source=gbs_atb#v=onepage&q&f=false

261. Silva AS, Yunes JA. Conservation of glycolytic oscillations in Saccharomyces cerevisiae. Genet Mol Res. 2006;3-5:525-535. Available from: https://pubmed.ncbi.nlm.nih.gov/17117368/

262. Schmid JW, Mauch K, Reuss M, Gilles ED, Kremling A. Metabolic design based on a coupled gene expression–metabolic network model of tryptophan production in Escherichia coli. Metab Eng. 2004;6:364-377. Available from: https://doi.org/10.1016/j.ymben.2004.06.003

263. Schwacke JH, Voit EO. Improved methods for the mathematically controlled comparison of biochemical systems. Theor Biol Med Model. 2004;1:1-18. Available from: https://doi.org/10.1186/1742-4682-1-1

264. Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of E. coli. Nat Genet. 2002;31:64-68. Available from: https://doi.org/10.1038/ng881

265. Shen L, Chen Z. Critical review of the impact of tortuosity on diffusion. Chem Eng Sci. 2007;62:3748-3755. Available from: https://doi.org/10.1016/j.ces.2007.03.041

266. Shimizu TS, Bray D. Computational cell biology - The stochastic approach. Cambridge: MIT Press; 2002. Available from: https://www.esalq.usp.br/lepse/imgs/conteudo_thumb/Computational-Cell-Biology---The-Stochastic-Approach.pdf

267. Sidoli FR, Mantalaris A, Asprey SP. Modelling of mammalian cells and cell culture processes. Cytotechnology. 2004;44(1-2):27-46. Available from: https://doi.org/10.1023/B:CYTO.0000043397.94527.84

268. Smigelschi O, Maria G. A modified half-interval search procedure for solving multivariate nonlinear process models. Hungarian J Ind Chem. 1986;14(4):453-462.

269. Snoep JL, Olivier BG. JWS online cellular systems modelling and microbiology. Microbiology. 2003;149:3045-3047. Available from: https://doi.org/10.1099/mic.0.c0124-0

270. Sotiropoulos V, Kaznessis YN. Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes. BMC Syst Biol. 2007 Jan 9;1:1-7. Available from: https://doi.org/10.1186/1752-0509-1-7

271. Somogyi R, Sniegoski CA. Modelling the complexity of genetic networks: understanding multigenetic and pleiotropic regulation. Complexity. 1996;1:45-63. Available from: https://philpapers.org/rec/SMOMTC

272. Speck C, Weigel C, Messer W. ATP- and ADP-DnaA protein, a molecular switch in gene regulation. EMBO J. 1999;18:6169–6176. Available from: https://doi.

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

org/10.1093/emboj/18.21.6169

273. Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED. Metabolic network structure determines key aspects of functionality and regulation. Nature. 2002;420:190-193. Available from: https://doi.org/10.1038/nature01166

274. Stelling J. Mathematical models in microbial systems biology. Curr Opin Microbiol. 2004;7(5):513-518. Available from: https://doi.org/10.1016/j.mib.2004.08.004

275. Stephanopoulos GN, Aristidou AA, Nielsen J. Metabolic engineering. Principles and methodologies. San Diego: Academic Press; 1998. Available from: https://books.google.co.in/books?id=9mGzkso4NVQC&printsec=frontcover#v=onepage&q&f=false

276. Stephanopoulos G, Simpson TW. Flux amplification in complex metabolic networks. Chem Eng Sci. 1997;52:2607-2627. Available from: https://doi.org/10.1016/S0009-2509(97)00077-8

277. Stephanopoulos G. Chemical process control. India: Pearson; 2015. Available from: https://www.amazon.in/Chemical-Process-Control-George-Stephanopoulos/dp/933254946X

278. Stiles JR, Bartol TM Jr, Salpeter EE, Salpeter MM. Monte Carlo simulation of neurotransmitter release using MCell, a general simulator of cellular physiological processes. In: Bower JM, editor. Computational neuroscience. New York: Plenum Press; 1998. Available from: https://www.semanticscholar.org/paper/Monte-Carlo-simulation-of-neuro-transmitter-release-Stiles-Bartol/45136743875ee08506c479453ed90cc048b459d7

279. Styczynski MP, Stephanopoulos G. Overview of computational methods for the inference of gene regulatory networks. Comput Chem Eng. 2005;29:519-534. Available from: https://doi.org/10.1016/j.compchemeng.2004.08.029

280. Surovstev IV, Morgan JJ, Lindahl PA. Whole-cell modelling framework in which biochemical dynamics impact aspects of cellular geometry. J Theor Biol. 2007;244:154–166. Available from: https://doi.org/10.1016/j.jtbi.2006.07.020

281. Surovtsev IV, Morgan JJ, Lindahl PA. Kinetic modelling of the assembly, dynamic steady state, and contraction of the FtsZ ring in prokaryotic cytokinesis. PLoS Comput Biol. 2008;4:e1000102. Available from: https://doi.org/10.1371/journal.pcbi.1000102

282. Sutherland W. Optimisation—this beguilingly simple idea allows biologists not only to understand current adaptations, but also to predict new designs that may yet evolve. Nature. 2005;435:569. Available from: https://doi.org/10.1038/435569a

283. Szedlacsek ES, Aricescu AR, Havsteen BH. Time-dependent control of metabolic systems by external effectors. J Theor Biol. 1996;182:341-350. Available from: https://doi.org/10.1006/jtbi.1996.0173

284. Thilakavathi M, Basak T, Panda T. Modelling of enzyme production kinetics. Appl Microbiol Biotechnol. 2007;73:991-1007. Available from: https://doi.org/10.1007/s00253-006-0667-0

285. Thomas R, Thieffry D, Kaufman M. Dynamical behaviour of biological regulatory networks. Bull Math Biol. 1995;57:247-295. Available from: https://doi.org/10.1007/BF02460618

286. Tian T, Burrage K. Stochastic models for regulatory networks of the genetic toggle switch. Proc Natl Acad Sci U S A. 2006;103:8372-8377. Available from: https://doi.org/10.1073/pnas.0507818103

287. Tiruvadi-Krishnan S, Maennik J, Kar P, Lin J, Amir A, Maennik J. Coupling between DNA replication, segregation, and the onset of constriction in Escherichia coli. Cell Rep. 2022;38(12):110539. Available from: https://doi.org/10.1016/j.celrep.2022.110539

288. Tomita M, Hashimoto K, Takahashi K, Shimizu T, Matsuzaki Y, Miyoshi F, et al. Saito K, Tanida S, Yugi K, Venter JC. E-Cell: software environment for whole cell simulation. Bioinformatics. 1999;15:72-84. Available from: https://doi.org/10.1093/bioinformatics/15.1.72

289. Tomita M. Whole-cell simulation: a grand challenge of the 21st century. Trends Biotechnol. 2001;19:205-210. Available from: https://doi.org/10.1016/s0167-7799(01)01636-5

290. Tomlin AS, Turanyi T, Pilling MJ. Mathematical tools for the construction, investigation and reduction of combustion mechanisms. In: Pilling MJ, Hancock G, editors. Low temperature combustion and autoignition. Amsterdam: Elsevier; 1997. p. 293-437. Available from: https://doi.org/10.1016/S0069-8040(97)80019-2

291. Tomita M, Hashimoto K, Takahashi K, Shimizu T, Matsuzaki Y, Miyoshi F, et al. Saito K, Tanida S, Yugi K, Venter JC. E-Cell: software environment for whole cell simulation. Bioinformatics. 1999;15:72-84. Available from: https://doi.org/10.1093/bioinformatics/15.1.72

292. Tomita M. Whole-cell simulation: a grand challenge of the 21st century. Trends Biotechnol. 2001;19:205-210. Available from: https://doi.org/10.1016/s0167-7799(01)01636-5

293. Tomshine J, Kaznessis YN. Optimisation of a stochastically simulated gene network model via simulated annealing. Biophys J. 2006;91:3196-3205. Available from: https://doi.org/10.1529/biophysj.106.083485

294. Torres NV, Voit EO. Pathway analysis and optimisation in metabolic engineering. Cambridge: Cambridge University Press; 2002. Available from: https://doi.org/10.1017/CBO9780511546334

295. Trambouze P, Van Landeghem H, Wauquier JP. Chemical reactors: design, engineering, operation. Paris: Edition Technip; 1988. Available from: https://books.google.co.in/books/about/Chemical_Reactors.html?id=e7rpAAAAMAAJ&redir_esc=y

296. Trun NJ, Gottesman S. On the bacterial cell cycle: Escherichia coli mutants with altered ploidy. Genes Dev. 1990;4:2036-2047. Available from: https://doi.

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

org/10.1101/gad.4.12a.2036

297. de Tremblay M, Perrier M, Chavarie C, Archambault J. Fed-batch culture of hybridoma cells: comparison of optimal control approach and closed loop strategies. Bioprocess Eng. 1993;9:13–21. Available from: https://link.springer.com/article/10.1007/BF00389535

298. Turanyi T. Review on sensitivity analysis of complex kinetic systems. Tools and applications. J Math Chem. 1990;5:203-248. Available from: https://link.springer.com/article/10.1007/BF01166355

299. Turanyi T. Reduction of large reaction mechanisms. New J Chem. 1990;14:795-803. Available from: https://www.scirp.org/reference/referencespapers?referenceid=379030

300. Turanyi T. KINAL: A program package for kinetic analysis of reaction mechanisms. Comput Chem. 1990;14:253-254. Available from: http://garfield.chem.elte.hu/Turanyi/pdf/13_Turanyi_CompChem_1990.pdf

301. Turanyi T, Tomlin AS, Pilling MJ. On the error of the quasisteady-state approximation. J Phys Chem. 1993;97:163-172. Available from: https://pubs.acs.org/doi/10.1021/j100103a028

302. Turing A. The chemical basis of morphogenesis. Philos Trans R Soc Lond B Biol Sci. 1952;237:641. Available from: https://www.dna.caltech.edu/courses/cs191/paperscs191/turing.pdf

303. Tyson JJ. Biochemical oscillations. In: Fall CP, Marland ES, Wagner JM, Tyson JJ, editors. Computational Cell Biology. Berlin: Springer Verlag; 2002. chap. 9. Available from: https://doi.org/10.1007/978-0-387-22459-6_9

304. Tyson JJ, Novak B. Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions. J Theor Biol. 2001;210:249-263. Available from: https://doi.org/10.1006/jtbi.2001.2293

305. Tyson JJ, Novak B, Odell GM, Chen K, Thron CD. Chemical kinetic theory: understanding cell cycle regulation. Trends Biochem Sci. 1996;21:89-96. Available from: https://pubmed.ncbi.nlm.nih.gov/8882581/

306. Vajda S, Valko P, Turanyi T. Principal component analysis of kinetic models. Int J Chem Kinet. 1985;17:55-81. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/kin.550170107

307. Vajda S, Rabitz H, Walter E, Lecourtier Y. Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models. Chem Eng Commun. 1989;83:191-219. Available from: https://www.tandfonline.com/doi/abs/10.1080/00986448908940662

308. Van Dien SJ, Lidstrom ME. Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph Methylobacterium extorquens AM1, with application to reconstruction of C3 and C4 metabolism. Biotechnol Bioeng. 2002;78:296-312. Available from: https://discovered.ed.ac.uk/discovery/fulldisplay?docid=cdi_pascalfrancis_primary_14185390&context=PC&vid=44UOE_INST:44UOE_VU2&lang=en&search_scope=UoE&adaptor=Primo%20Central&tab=Everything&query=null,,Biology%20of%20microorganisms&offset=0

309. Van Someren EP, Wessels LFA, Backer E, Reinders MJT. Multi-criterion optimisation for genetic network modelling. Signal Process. 2003;83:763-775. Available from: https://doi.org/10.1016/S0165-1684(02)00473-5

310. Vance W, Arkin A, Ross J. Determination of causal connectivities of species in reaction networks. Proc Natl Acad Sci U S A. 2002;99:5816-5821. Available from: https://doi.org/10.1073/pnas.022049699

311. Varma A, Morbidelli M, Wu H. Parametric sensitivity in chemical systems. Cambridge: Cambridge University Press; 1999. Available from: https://doi.org/10.1017/CBO9780511721779

312. Vilela M, Morgan JJ, Lindahl PA. Mathematical model of a cell size checkpoint. PLoS Comput Biol. 2010;6(12):1-11. Available from: https://doi.org/10.1371/journal.pcbi.1001036

313. Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ. Optimal re-design of primary metabolism in Escherichia coli using linlog kinetics. Metab Eng. 2004;6:378-390. Available from: https://doi.org/10.1016/j.ymben.2004.07.001

314. Voit EO. Smooth bistable S-systems. IEE Proc Syst Biol. 2005;152:207-213. Available from: https://doi.org/10.1049/ip-syb:20050063

315. von Bertalanffy L. Modern theories of development: an introduction to theoretical biology. New York: Oxford University and Harper Press; 1933. Available from: https://openlibrary.org/books/OL6294222M/Modern_theories_of_development

316. Wagner-Döbler I, von Canstein H, Li Y, Timmis K, Deckwer WD. Removal of mercury from chemical wastewater by microorganisms in technical scale. Environ Sci Technol. 2000;34:4628-4634. Available from: https://doi.org/10.1021/es0000652

317. Wall ME, Hlavacek WS, Savageau MA. Design principles for regulator gene expression in a repressible gene circuit. J Mol Biol. 2003;332:861-876. Available from: https://doi.org/10.1016/s0022-2836(03)00948-3

318. Wallwork SC, Grant DJW. Physical chemistry. London: Longman; 1977. Available from: https://www.amazon.in/Physical-Chemistry-Students-Pharmacy-Biology/dp/0582442540

319. Wang P. Multi-scale features in recent development of enzymatic biocatalyst systems. Appl Biochem Biotechnol. 2009;152:343-352. Available from: https://doi.org/10.1007/s12010-008-8243-y

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

320. Wei J, Kuo JCW. A lumping analysis in monomolecular reaction systems. 1. Analysis of the exactly lumpable system. 2. Analysis of approximately lumpable system. Ind Eng Chem Fundam. 1969;8:114-133. Available from: https://pubs.acs.org/doi/10.1021/i160029a020

321. Werner A, Heinrich R. A kinetic model for the interaction of energy-metabolism and osmotic states of human erythrocytes—analysis of the stationary in vivo state and of time-dependent variations under blood preservation conditions. Biomed Biochim Acta. 1985;44:185-212. Available from: https://pubmed.ncbi.nlm.nih.gov/4004830/

322. Westerhoff HV. The Silicon cell, not dead but live! Metab Eng. 2001;3:207-210. Available from: https://doi.org/10.1006/mben.2001.0192

323. Westerhoff HV. Engineering life processes live: the Silicon cell. Oral Lecture ESCAPE 16, Garmisch (DE); 2006.

324. Westerhoff HV, Palsson BO. The evolution of molecular biology into systems biology. Nat Biotechnol. 2004;22(10):1249-1252. Available from: https://doi.org/10.1038/nbt1020

325. Whitham AG, Sparks RSJ. Pumice. Bull Volcanol. 1986;48:209-223. Available from: https://www.scirp.org/reference/referencespapers?referenceid=1894537

326. Wiback SJ, Palsson BO. Extreme pathway analysis of human red blood cell metabolism. Biophys J. 2002;83:808-818. Available from: https://doi.org/10.1016/s0006-3495(02)75210-7

327. Wiener N. Cybernetics or control and communication in the animal and the machine. New York: McGraw-Hill; 1948. Available from: https://uberty.org/wp-content/uploads/2015/07/Norbert_Wiener_Cybernetics.pdf

328. Wikipedia. Mutagen agents. 2025 [cited 2025 Aug 8].

329. Wolf J, Passarge J, Somsen OJG, Snoep JL, Heinrich R, Westerhoff HV. Transduction of intracellular and intercellular dynamics in yeast glycolytic oscillations. Biophys J. 2000;78:1145-1153. Available from: https://doi.org/10.1016/s0006-3495(00)76672-0

330. Wolkenhauer O, Mesarovic M. Feedback dynamics and cell function: why systems biology? Mol Biosyst. 2005;1:14-16. Available from: https://pubs.rsc.org/en/content/articlelanding/2005/mb/b502088n

331. Wuensche A, Lesser MJ. The global dynamics of cellular automata; An atlas of basin of attraction fields of one-dimensional cellular automata. Reading (MA): Santa Fe Institute, Addison-Wesley; 1992. Available from: https://users.sussex.ac.uk/~andywu/downloads/papers/global_dynamics_of_CA.pdf

332. Xie L, Wang DIC. Stoichiometric analysis of animal cell growth and its application in medium design. Biotechnol Bioeng. 1994;43:1164-1174. Available from: https://doi.org/10.1002/bit.260431122

333. Xiong J. Essential bioinformatics. Cambridge: Cambridge University Press; 2006. Available from: https://www.scirp.org/reference/referencespapers?referenceid=451907

334. Yang ST, editor. Bioprocessing for value-added products from renewable resources. Amsterdam: Elsevier; 2007. Available from: https://doi.org/10.1016/B978-0-444-52114-9.X5000-2

335. Yang Q, Lindahl P, Morgan J. Dynamic responses of protein homeostatic regulatory mechanisms to perturbations from steady state. J Theor Biol. 2003;222:407-423. Available from: https://doi.org/10.1016/s0022-5193(03)00052-3

336. Yokobayashi Y, Collins CH, Leadbetter JR, Weiss R, Arnold FH. Evolutionary design of genetic circuits and cell-cell communications. Adv Complex Syst. 2003;6(1):1-9. Available from: https://www.researchgate.net/publication/23750790_Evolutionary_Design_Of_Genetic_Circuits_And_Cell-Cell_Communications

337. You LC. Toward computational systems biology. Cell Biochem Biophys. 2004;40:167-184. Available from: https://doi.org/10.1385/cbb:40:2:167

338. Zhang R, Ehigie JO, Hou X, You X, Yuan C. Steady-state preserving simulation of genetic regulatory systems. Comput Math Methods Med. 2017;2017:2729683. Available from: https://doi.org/10.1155/2017/2729683

339. Zadran S, Levine RD. Perspectives in metabolic engineering: Understanding cellular regulation towards the control of metabolic routes. Appl Biochem Biotechnol. 2013;169:55-65. Available from: https://doi.org/10.1007/s12010-012-9951-x

340. Zak DE, Aderem A. Systems biology of innate immunity. Immunol Rev. 2009;227:264-282. Available from: https://doi.org/10.1111/j.1600-065X.2008.00721.x

341. Zak DE, Vadigepalli R, Gonye GE, Doyle FJ III, Schwaber JS. Unconventional systems analysis problems in molecular biology: A case study in gene regulatory network modelling. Comput Chem Eng. 2005;29(3):547-563. Available from: https://doi.org/10.1016/j.compchemeng.2004.08.016

342. de Zengotita VM, Miller WM, Aunins JG, Zhou W. Phosphate feeding improves high-cell-concentration NS0 myeloma culture performance for monoclonal antibody production. Biotechnol Bioeng. 2000;69(5):566-576. Available from: https://doi.org/10.1002/1097-0290(20000905)69:5%3C566::aid-bit11%3E3.0.co;2-4

343. Zhang W, Tichy SE, Perez LM, Maria GC, Lindahl PA, Simanek EE. Evaluation of multivalent dendrimers based on melamine. Kinetics of dithiothreitol-mediated thiol-disulfide exchange depends on the structure of the dendrimer. J Am Chem Soc. 2003;125:5086-5094. Available from: https://doi.org/10.1021/ja0210906

344. Zhu Y, Song J, Xu Z, Sun J, Zhang Y, Li Y, Ma Y. Development of thermodynamic optimum searching (TOS) to improve the prediction accuracy of flux balance analysis. Biotechnol Bioeng. 2013;110(3):914-923. Available from: https://doi.org/10.1002/bit.24739

345. Zhu R, Ribeiro AS, Salahub D, Kauffman SA. Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models. J Theor Biol. 2007;246:725-745. Available from: https://doi.org/10.1016/j.jtbi.2007.01.021

ISBN:

DOI: https://dx.doi.org/10.17352/ebook10124

346. Zwillinger D, Krantz SG, Rosen KH. Standard mathematical tables and formulae. Boca Raton: CRC Press; 1996. Available from: https://pws.npru.ac.th/sartthong/data/files/CRC_Standard_Mathematical_Tables_and_Formulae__Thirtieth_Edition.pdf

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124

Biology and Medicine

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

Gheorghe Maria*

School of Public Health, University of the Western Cape, Robert Sobukwe Road, Bellville, 7535, South Africa

ISBN:
DOI: https://dx.doi.org/10.17352/ebook10124