ISBN: 978-1-946628-12-1

IN-SILICO DESIGN OF GENETIC MODIFIED MICRO-ORGANISMS (GMO) OF INDUSTRIAL USE, BY USING SYSTEMS BIOLOGY AND (BIO) CHEMICAL ENGINEERING TOOLS

Juniper Publishers
key to the Researchers

By
Gheorghe Maria
Department of Chemical and Biochemical Engineering,
University Politehnica of Bucharest, Romania

In-Silico Design of Genetic Modified Micro-Organisms (GMO) of Industrial Use, by Using Systems Biology and (Bio) Chemical Engineering Tools

Gheorghe Maria*

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

***Corresponding**

Department of Chemical and Biochemical Engineering, University Politehnica of Bucharest, Polizu Str. 1-7, 011061, Bucharest, Romania, Email: gmaria99m@hotmail.com

Published By :

Juniper publishers Inc.

United States

Date : March 21, 2018

Prof. Gheorghe Maria

Department of Chemical & Biochemical Engineering, University Politehnica of Bucharest (UPB), Romania

Dr. Gheorghe Maria is currently professor in Chemical & Biochemical Reaction Engineering with the UPB - University Politehnica of Bucharest (Romania). He received the PhD in 1987 in chemical engineering at UPB (supervisor Prof. Dr. ing. Raul Mihail, in a subject related to statistical estimation of (bio)chemical process kinetic models). On 1982 he joined ICECHIM research center in Bucharest (catalysis, and chemical & biochemical energetics Institute, as a senior researcher), and on 1990 he joined UPB as a lecturer. On 1992 he come in Switzerland for working as Assistant Professor with ETH Zürich (with the late Prof. David W.T. Rippin group of Process System Engineering). On 1997 he return to Romania becoming Associate Professor and then full Professor (1999) with UPB. His research interests concern the fields of (bio) chemical reactor and kinetic modelling, biochemical engineering and bioinformatics, risk analysis of chemical plants, Systems Biology, Process systems engineering, modelling of cell metabolic processes, gene expression and regulatory circuits, and drug release kinetics. Over the past 25 years he participated to various national or intl. Projects (more than 15), making short research stages/visitingship at ETH Zürich (3-months on 1997, SNSF fellow on environmental risk analysis), Univ. des Saarlandes (3-months on 1999, DAAD fellow on modelling complex enzymatic kinetics), TU Erlangen (3-months on 2000, catalytic membrane reactors), Texas A&M University (2002-2003, res. scientist with NIH fellow on modelling gene expression and gene regulatory networks, synthetic biology), TU Braunschweig (2006) and TU Hamburg (3-months on 2009, DFG and DAAD fellows on simulating bacteria resistance to environmental pollutants), Tianjin Inst. Ind. Biotechnology China (2-months on 2010, *in-silico* searching for gene knockout strategies to maximize production of succinate in *E. coli* cell, synthetic biology). He presented more than 31 invited Lectures at various Universities in EU, CAN, USA, China (among them: Princeton Univ. 1994, Texas A&M University 2002, EPFL Lausanne 1997, Queen's Univ. Kingston 1994 Canada, BASF Ludwigshafen 1996). He authored 8 books (with ISBN in RO,USA), 5 teaching books (in Romanian), 6 book chapters, (with ISBN, Engl.), 143 papers in peer reviewed ISI international journals and univ. journals, and 78 in intl. conference proceedings, with more than 1200 citations (h-index 18 and i10 index 41; ResearchGate score = more than 35). Among them, 31papers have been published in Chem. Eng. journals with the highest IF>3. According to the Romanian ranking system, he reported high scores signing NP>108 ISI papers as principal author, with a cummultative impact factor of FIC>155. He is a reviewer for many (bio)chemical engineering journals (25). He co-chaired or was member of the organizing Committees of 16 international conferences (among them: 5th Int. Conference on Computational Bioengineering (ICCB-5), 11-13 September, 2013, Leuven, Belgium; ROMPHYSCHEM-15, 15-th International Conference of Physical Chemistry, 11-13 September, 2013, Bucharest; 13th Edition of Academic Days of Timisoara, June 13-14, 2013, Timişoara (Romania); ELSEDIMA 10th and 12th International Conference (Environmental Legislation, Safety Engineering and Disaster Management), 18-19 September 2014 and 26-28 May 2016 and 17-19 May 2018, respectively," Cluj-Napoca (RO).

- On 1985 he award the Prize of the Romanian Academy for kinetic studies and scalep / plant design for the methanol-to-gasoline process at Brazi petrochemical works (Ploiesti refinery, Romania)". He is member in the scientific/editorial board of Chem. & Biochemical Eng. Q. (Zagreb), Revista de Chimie (Bucharest), The Scientific Bulletin of University POLITEHNICA of Bucharest, Bulletin of Romanian Chemical Engineering Society, ECOTERRA Journal of Environmental Research and Protection (edited by Romanian Society of Environmental Sciences and Engineering, Cluj-Napoca Romania). He was also an expert for various EU and national research programs in the bioengineering area, being also a PhD supervisor at UPB (from 2008, in Chem. & Biochem. Engineering topics), with 4 by 5 PhDs finalized, and 4 PhDs in progress. He participated as researcher or director to a large number of national and international grants. Among them are to be mentioned:

- SNSF Swiss-Romanian Project Project 7IP 050 113 / 1997-1998 on: 'Ecological Design and Operation of Chemical Processes', at ETH Zürich;

- NATO Grant no. 974850-99/1999-2001 on 'Identification, Optimal Monitoring and Risk Limits for a Wastewater Biological Treatment Plants', at Universidade da Porto, Portugal, 1999-2001;
- NIH Project 2002, Department of Chemistry and Biochemistry, Texas A&M University, College Station, USA, on the theme: 'Methodology to construct and simulate molecular-level mechanisms by which living systems grow and divide';
- NIH Project 2002, Texas A&M University, College Station, USA, on 'Kinetics of Programmable Drug Release in Human Plasma';
- DFG Grant SFB-578/2006, Development of biotechnological Processes by Integrating Genetic and Engineering Methods, at TU Braunschweig, Germany;
- National CNCSIS Project nr. 1543/2008-2011 (IDEI) on: 'A nonlinear approach to conceptual design and safe operation of chemical processes' ("O abordare neliniara a problemelor de proiectare conceptuala si de operare in conditii de siguranta a proceselor chimice");
- European Commission Project through European Regional Development Fund and of the Romanian state budget, project POSCCE-O2.1.2-2009-2, ID 691 / 2010-2013, "New mesoporous alumnosilicate matterials for controlled release of biological active substances" (Noi materiale din clasa aluminosilicatilor mezoporosi pentru eliberare controlata de substante biologic active).

Complete list of publications on his Web-page: https://sites.google.com/site/gheorghemariasite/

ORCID ID= J-4840-2012

Supplementary information also on: http://sicr.ro/wp-content/uploads/2017/01/BRChES_2_2016.pdf

Contents

1.Abstract

The work is a systematized and reasonable short review of the main published contributions of Dr. Maria in the field of metabolic processes simulation related to the central carbon metabolism (CCM) and, in particular, modelling the dynamics of the gene expression regulation (GERM) and of the genetic regulation circuits (GRC) in living cells. Application of biochemical engineering, and of nonlinear system control principles and concepts to the modeling of complex cellular processes on deterministic bases are briefly presented with including a rich list of references. These models are essential for understanding and simulating the CCM, useful to *in-silico* design of genetically modified micro-organisms (GMOs) with applications in industrial biosynthesis, medicine, environmental engineering, vaccine production, biosensors, etc.

A special attention is paid to contributions related to dynamics simulation of the gene expression regulatory modules (GERM) and of genetic regulation circuits (GRC) in living cells, by promoting novel concepts of a novel cell modelling framework, that is the so-called "variable-volume-whole-cell" (VVWC) models. The relatively novel concept of "whole-cell" simulation of cell metabolic processes has been reviewed to prove its advantages when building-up dynamic models of modular structures that can reproduce complex metabolic syntheses inside living cells. The advantages of the more realistic VVWC approach are briefly underlined and exemplified when developing kinetic representations of the gene expression regulatory modules (GERM) that control the protein synthesis and homeostasis of metabolic processes. After a brief presentation of the general concepts and particularities of the VVWC modelling, both past and current experience with constructing effective GERM models is reviewed, together with some rules used when linking GERM-s to build-up models for optimized globally efficient genetic regulatory circuits (GRC), by using quantified regulatory indices evaluated vs. simulated dynamic and stationary environmental perturbations.

The topics belongs to the emergent field of 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). Systems Biology 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. The present work makes a short review of the (bio)chemical engineering principles and deterministic modelling rules used by the Systems Biology for modelling cellular metabolic processes. This involves application of the classical (bio)chemical engineering modelling techniques (mass balance, thermodynamic principles), algorithmic rules, nonlinear system control theory, and bioinformatics rules. 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.

Exemplifications of using such modular GRC models for *in-silico* design of GMO of industrial use include the own experience on improving several bioprocesses, that is five case studies of high complexity, referring to: 1) *In-silico* design of a genetic switch in *E. coli* with the role of a biosensor; 2) *In-silico* design of a cloned *E. coli* with a maximized capacity of mercury uptake from wastewaters; 3) *In-silico* design of a genetic modified *E. coli* with a maximized capacity of succinate (SUCC) production; 4) *In-silico* design of a genetic modified *E. coli* with a modified glycolytic oscillator; 5) *In-silico* modulate the bioreactor operating conditions with a modified *E. coli* to maximize the production of tryptophan.

2. Keywords: Systems biology; Cell metabolism; Deterministic modelling; Gene expression modelling; Genetic regulatory circuits (GRC); Pareto design of E.coli to get maximum succinate; Design cloned E.coli to get maximum mercury uptake; Design E.coli with a genetic switch biosensor; Design E.coli with a desired glycolytic oscillator; Tryptophan synthesis; Glycolytic oscillator

3. Motto

"All models are wrong, but some are useful"

Box, George E.P., Draper, Norman, R. (1987). *Empirical model-building and response surfaces*, p. 424, Wiley, ISBN 047 1810339.

4. Introduction - GMO Importance

Over the last decades genetically modified micro-organisms (GMOs) reported increasingly important uses in a broad range of areas, such as: biosynthesis industry (optimized bio-syntheses, or bioprocesses using GMO or modified enzymes produced by cloned GMO; production of vaccines using cloned GMO), environmental engineering (improved biological treatment of wastewaters using adapted GMO), medicine (therapy of diseases), design of new devices based on cell-cell communicators, design of biosensors using GMO or cloned micro-organisms, 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)], the *in-silico* (math model based) design of GMO uses the same modular approach [60] provided examples of biological systems that have evolved in a modular fashion and, in different contexts, perform the same basic functions. Each module, grouping several cell components and reactions, generates an identifiable function (e.g. regulation of a certain reaction, gene expression, etc.). More complex functions, such as regulatory networks, synthesis networks, or metabolic cycles can be built-up using the building blocks rules of the emergent *Synthetic Biology* [61,113].

The emergent *Synthetic Biology* [62] "interpreted as the engineering-driven building of increasingly complex biological entities" [61,63,113], 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 [61,63,113]; 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 the industrial biosynthesis. 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 [64,65]. Recent improvements in the synthetic biotechnology and production of modified enzymes using GMO, exhibiting desired functions, allowed a considerable progress in industrial enzyme technologies and various other applications. 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 [64, 65]. Isolated and immobilized enzymes have been used in large-scale industrial reactors for more than fifty 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) [66]. New efforts are invested in protein engineering (molecular level design), various chemical and physical manipulations coupled with bioactive nanostructure fabrication aiming at optimizing the carrier materials' structure to improve the enzyme stability and 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, their high sensitivity to operating conditions and impurities, too high substrate specificity, and difficult process controllability due to their variable characteristics.

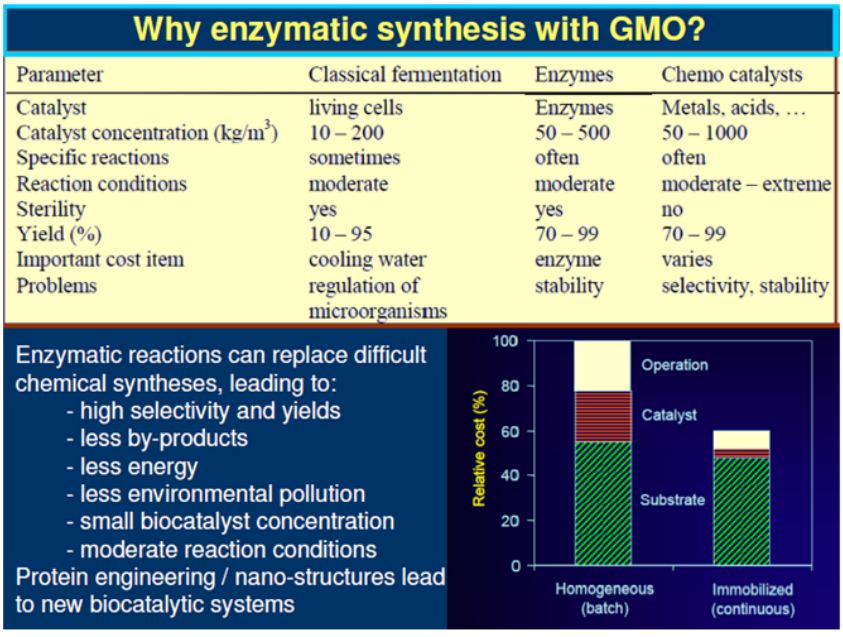


Figure 1: Advantages of (multi-)enzymatic and biological processes vs. chemical synthesis.

Because the isolated/immobilized enzyme is usually the most expensive raw-material of the biochemical process (Figure 1), especially when it is obtained from modified organisms, the right choice and optimization of industrial enzymatic reactors (Figure 2) is a classical but still of high interest subject. As enzymatic processes are usually much slower than the chemical ones (Figure 1), the suitable choice of the constructive solution (continuous mixing vs. plug-flow), enzyme utilization (free-enzyme vs. immobilized enzyme in gel/solid beads), or operation mode (batch/semi-batch vs. continuous reactors) is the main problem to be solved by engineers (Figure 2). Structured information on the enzyme characteristics (stability, activity, sensitivity to operating conditions and products), on immobilization possibilities, and on the process kinetics is essential when deciding on the optimal scaling-up solution. When a process model is available, engineering rules extensively describe the calculation route used to evaluate the reactor efficiency and suggest optimization modes [37].

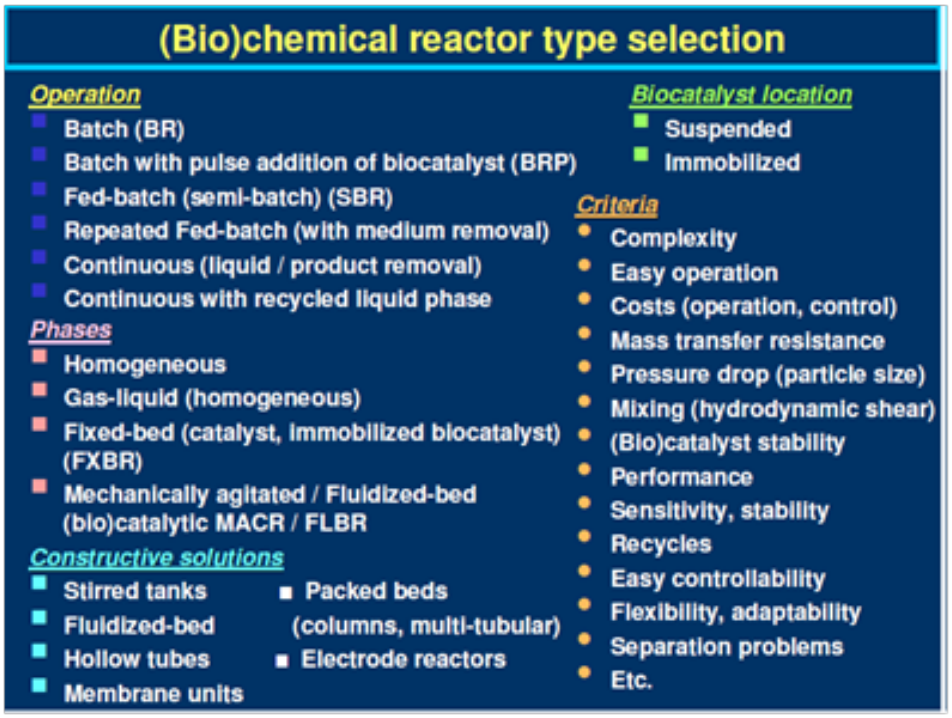


Figure 2: Criteria to choose the enzymatic reactor [37].

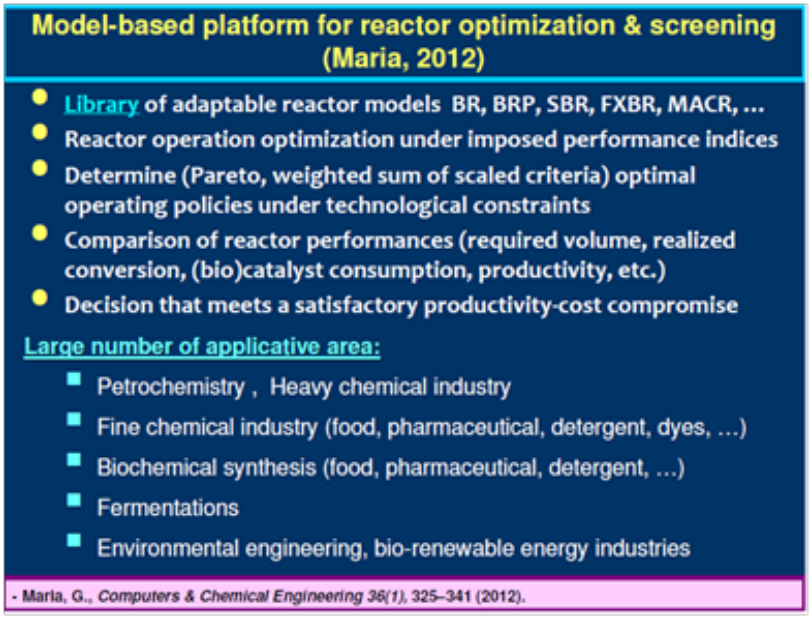


Figure 3: The principle of model-based evaluation of enzymatic reactor performances.

To optimize the design and operation of enzymatic reactors, developed a modular simulation platform to simulate and compare the performance of various enzymatic reactors, with the aim of determining the optimal operation mode for a given enzymatic process of known

kinetic characteristics (Figure 3 & 4). By using ideal enzymatic reactor models [batch reactor with initial addition of enzyme (BR); batch operation with intermittent addition of enzyme (BRP), semi-batch reactor with a constant or variable feed flow rate of enzyme solution (SBR); mechanically agitated continuous reactor (MACR) with immobilized enzyme on suspended porous particles; fixed-bed continuous reactor (FXBR) with immobilized enzyme on porous support packed in columns, (Figure 4), the simulation platform allows considerable savings in design effort, and reactor optimization, by minimizing the enzyme consumption without any loss in productivity or conversion.

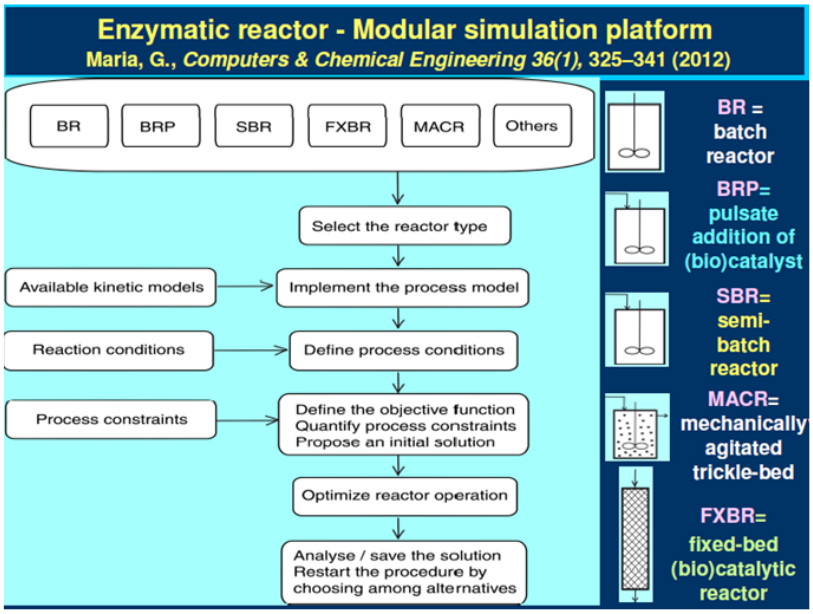


Figure 4: Optimal choice of (multi)enzymatic reactors following the simulation platform of [37].

To save costly experimental effort when design a GMO for a certain application, the use of an *in-silico* technique based on modular math model of the cell metabolism (including the central carbon metabolism CCM, and the GRC interfering with the target metabolites) is the most comprehensive mean for a rational design of GMO, including modified GRC of desired behavior [1,8]. 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). In the case of modelling GRC-s, by chance, the number of interacting GERM-s is limited, one gene interacting with no more than 23-25 [1,7].

Thus, the emergent research area of Systems Biology dealing with *in-silico* design of GMO has caught the contour and developed exponentially over the last 2-3 decades.

Why *in-silico* (model-based) design of GMO?

Classical way followed by biologists to get GMO induces genome modifications in micro-organisms by using a mutagen vector [132], then different strained are isolated and studied to separate and multiply the desirable one. Such an experimental procedure is time consuming (years) and very costly. By contrast, if an adequate cell model is available, the *in-silico* route is faster and allows an easy investigation of a large number of GMO (virtual) alternatives. The limitation of the *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 math model of the target metabolic process. Bioinformatics tools are often used to shorten the computational effort and to better process the huge quantity of information from omics databanks and, in particular, gene expression data [133-137]. The present work makes a short review of the (bio)chemical engineering principles and deterministic modelling rules used by the Systems Biology for modelling cellular metabolic processes aiming at *in-silico* design GMO with desired characteristics useful in medicine, and (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.

Exemplifications includes the own experience on improving several bioprocesses, that is: 1) *In-silico* design of a genetic switch in *E. coli* with the role of a biosensor, [54,56,57] 2) *In-silico* design of a cloned *E. coli* with a maximized capacity of mercury uptake from wastewaters [151-156]; 3) *In-silico* design of a genetic modified E. coli with a maximized capacity of succinate (SUCC) production [157]; 4) *In-silico* design of a genetic modified *E. coli* with a modified glycolytic oscillator [160-162]; 5) *In-silico* modulate the bioreactor operating conditions with suspended *E. coli* to maximize the production of tryptophan [163,164].

5. 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" [9] is one of the modern tools which uses advanced mathematical simulation models for *in-silico* design of GMO 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) must be changed. These metabolic fluxes are mainly determined by the environmental conditions (substrate/nutrients availability), but rather by the characteristics of the enzyme catalyzing these metabolic reactions. As proved by the famous cryptanalyst Alan Turing published on 1952 in his paper "The chemical basis of morphogenesis" [67], there is a (bio)chemical reaction basis of metabolic processes (Figure 21) 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?" [70]. 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) [1,68]. The direct relationship gene – metabolic flux is made via the expressed enzyme (Figure 6), followings subsequent steps: i) to change a certain flux, properties of the concerned protean enzyme should be changed; ii) that means, their encoding gene expression must be changed accordingly; iii) this involves either removal of that gene from the genome ("gene knockout"), its substitution with another gene from another micro-organism, or gene level amplification (by cloning with target plasmids), or eventually, modification of gene/enzymes controlling reactions that interfere with the concerned reaction (Figure 6-10).

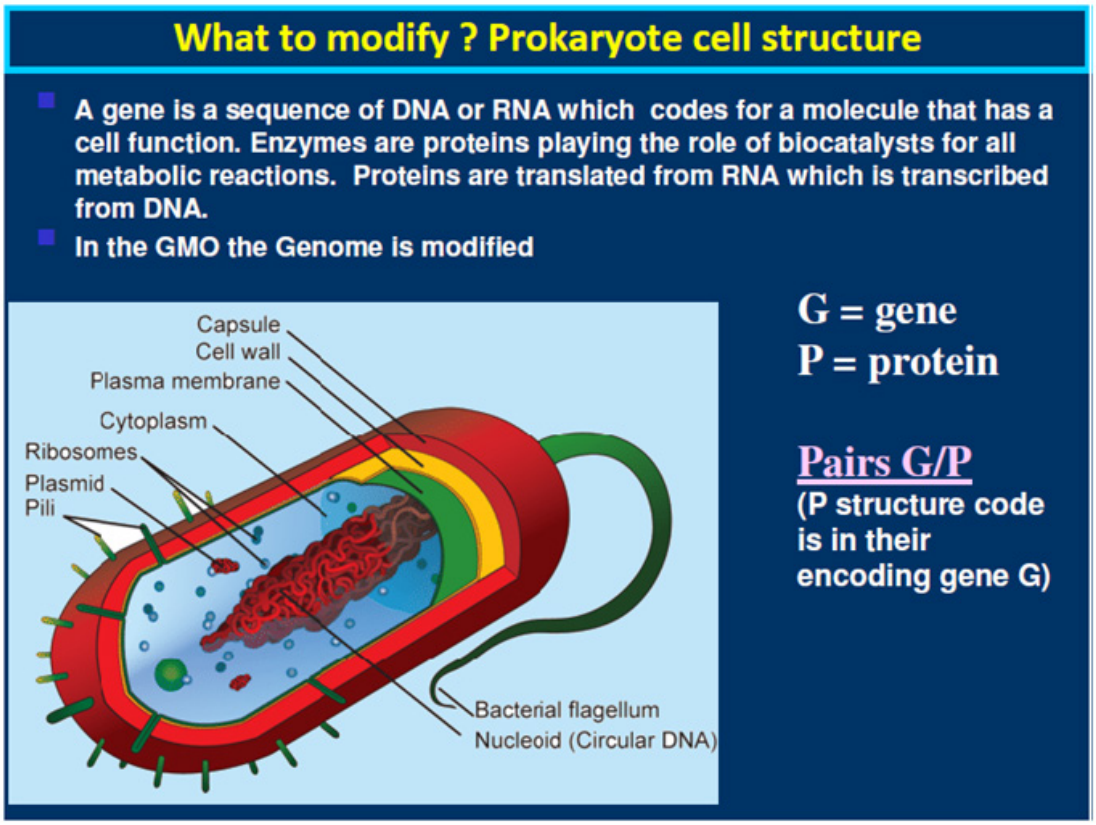


Figure 5: The prokaryote cell structure.

(source= https://en.wikipedia.org/wiki/File:Average_prokaryote_cell-_en.svg)

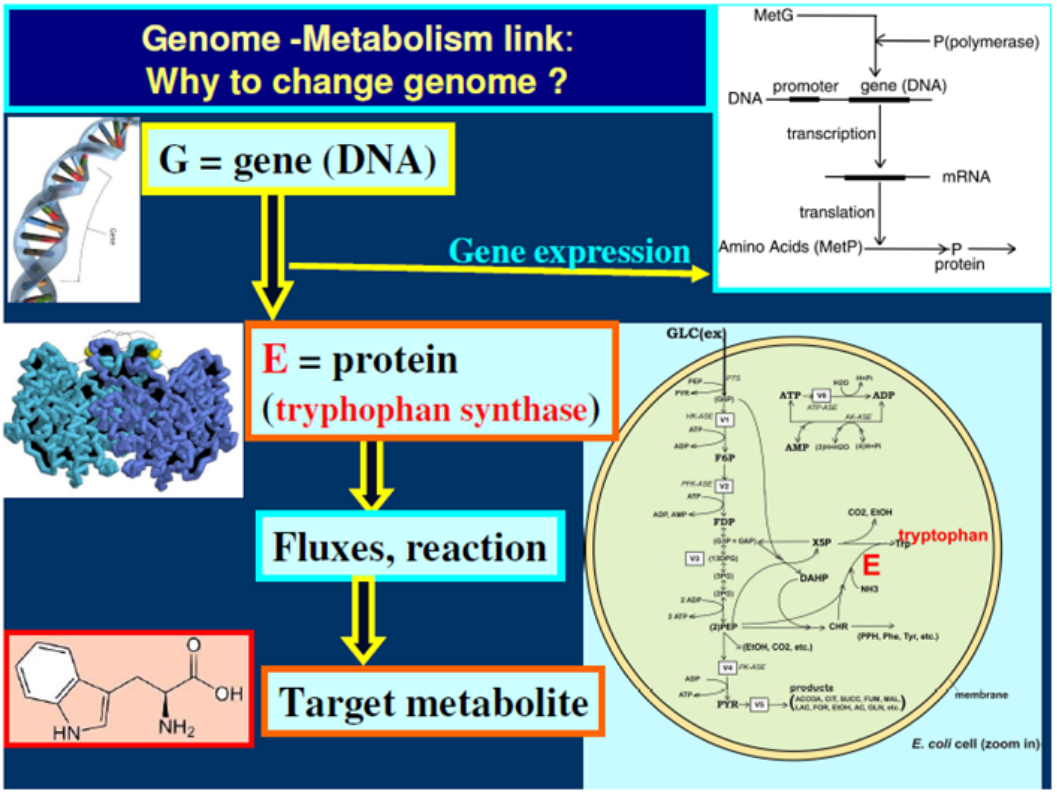


Figure 6: The direct link between genome and the metabolic fluxes of the cell [39,45].

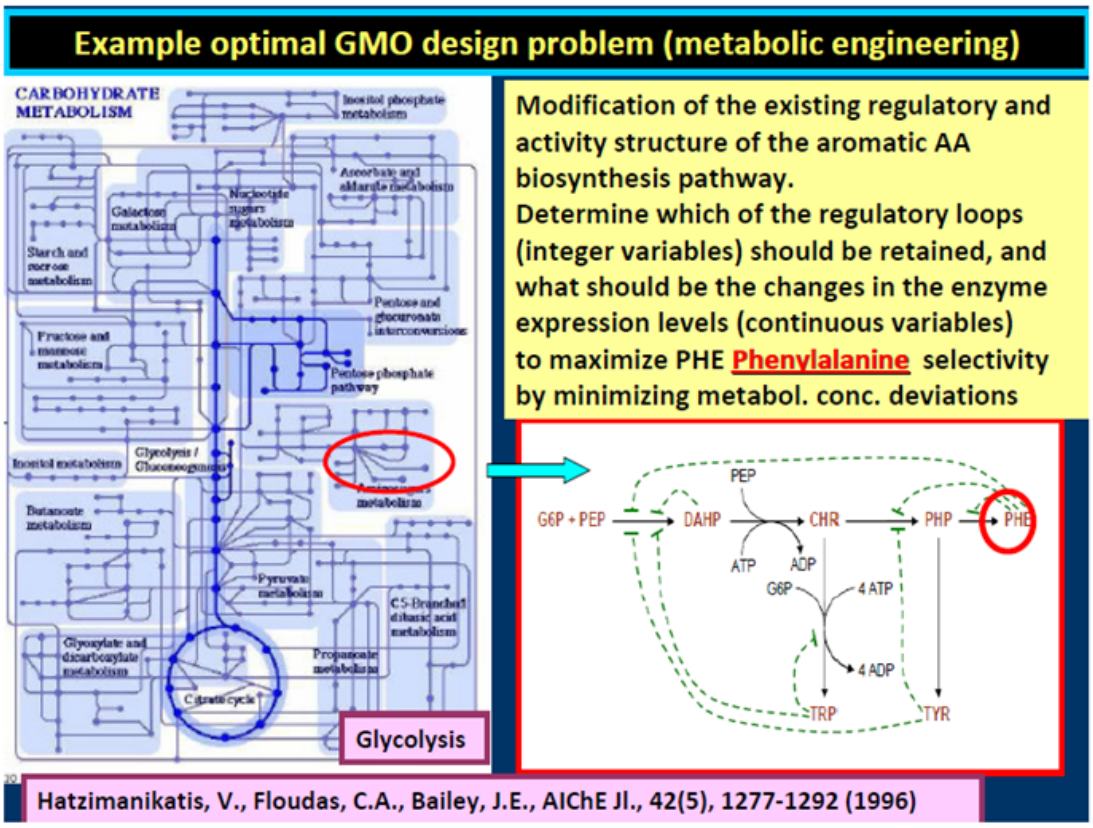


Figure 7: A simple example of a design problem to re-configure the metabolic pathway for Phenyl-alanine synthesis [38].

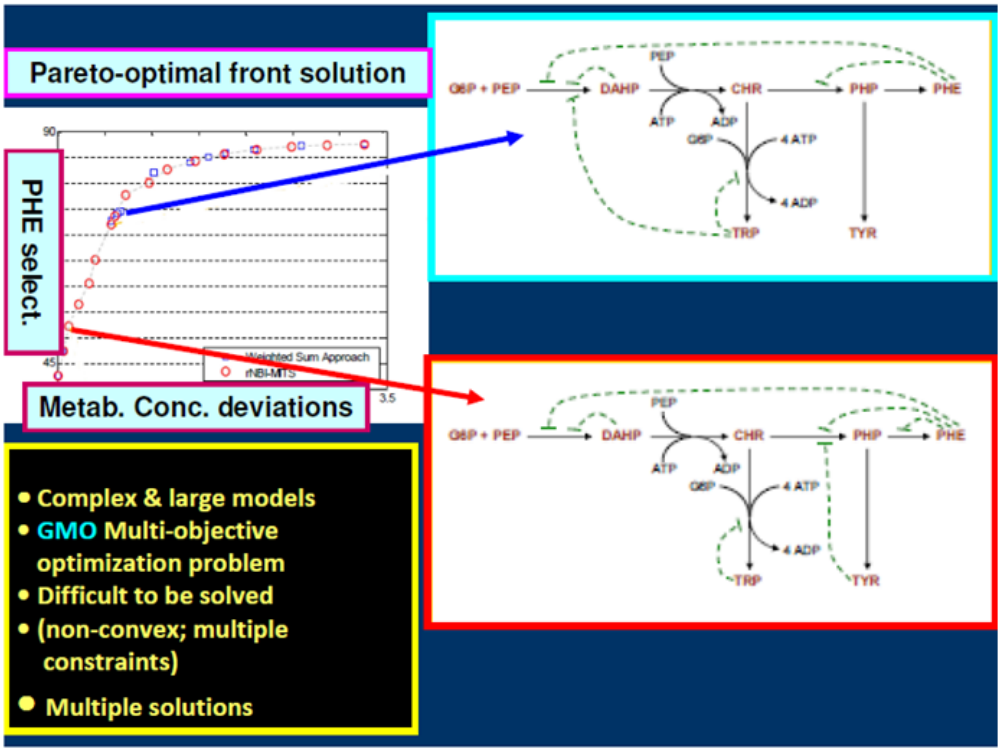


Figure 8: (Continuation). The resulted Pareto-optimal front (of two contrary objectives, i.e. max. selectivity in Phenyl-alanine, and min. deviations in metabolites stationary concentrations) is leading to two alternative metabolic configurations.

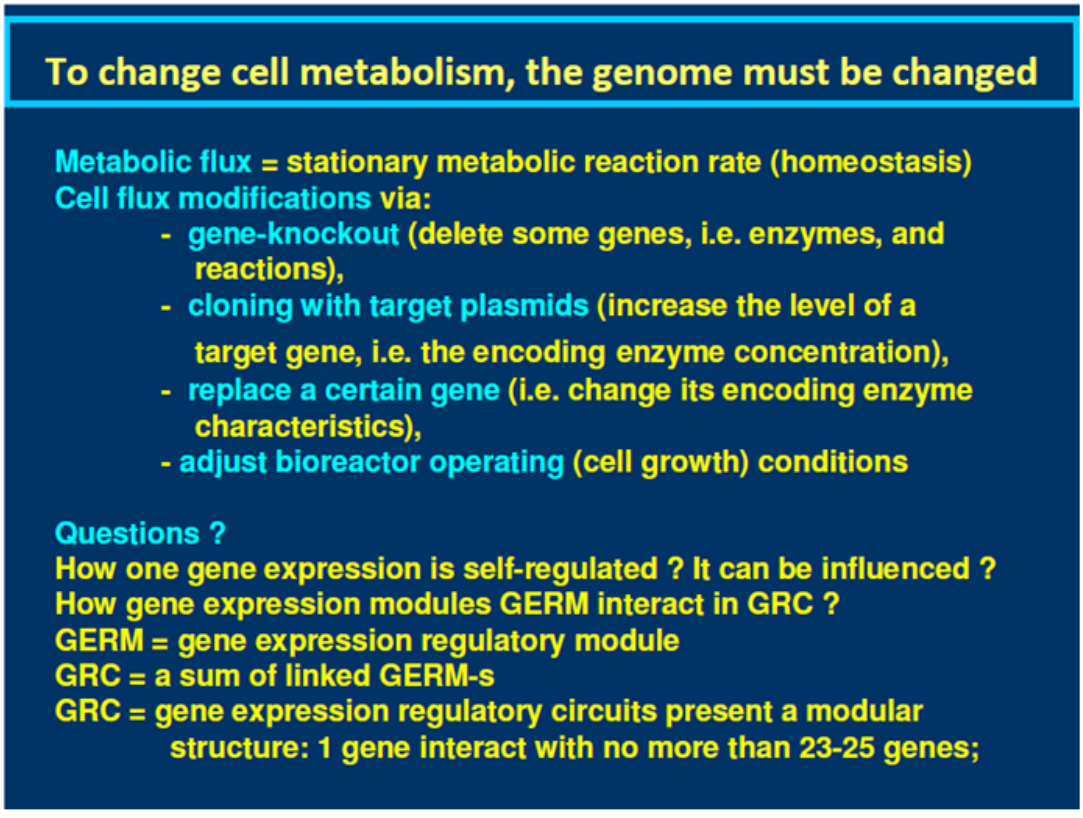


Figure 9: Some usual biological methods to alter the cell genome [39].

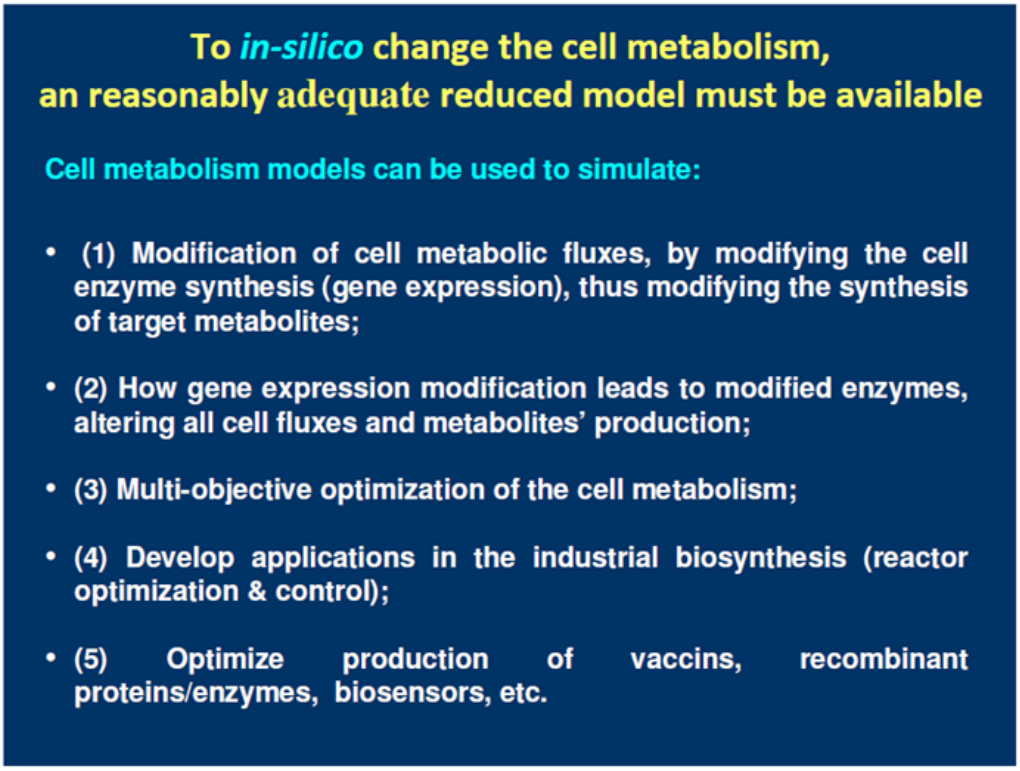


Figure 10: Problem to be solved when applying in-silico design of GMO.

The connection principle between the gene expression, enzymes, and reactions, for one or several interfering enzymes is shown schematically in the Figure 99 [177]. Sometimes, interference among gene expressions is modelled by using simple linear (empirical) models [178-180].

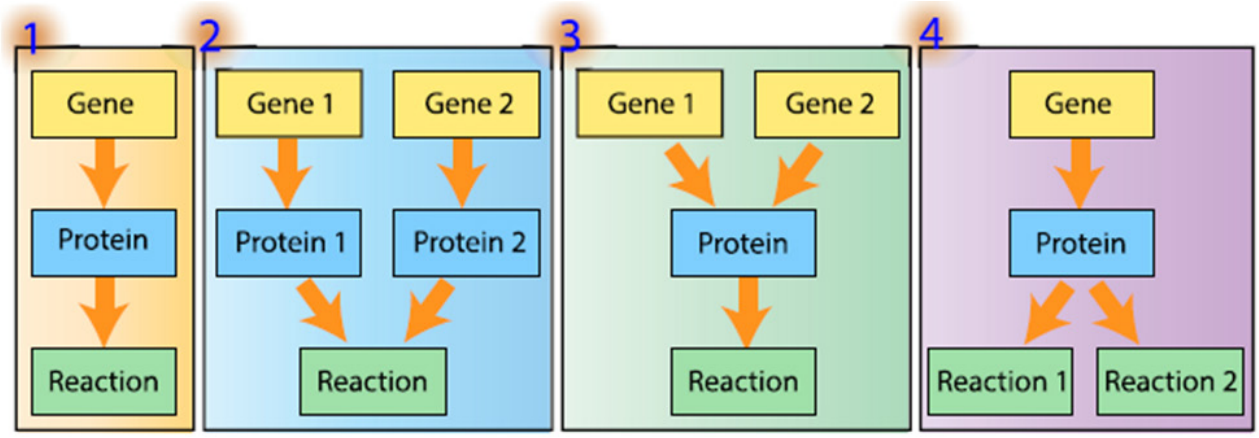


Figure 99: Examples of detailed gene–protein–reaction (GPR) associations. (1) Simple association, in which a single gene encodes a single enzyme. (2) Isozymes, in whichmultiple genes encode distinct proteins carrying out the same function. (3)Multimeric protein complex,wherein multiple genes encoding distinct protein subunits come together to forman active enzyme. (4) Multifunctional protein, in which a single protein can carry out multiple reactions [177].

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) [4], 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 the MCA Web, [5,6].

Due to the large number of genes [O(3-4)], and high complexity of gene interactions (even if one gene interacting with no more than 23-25 other genes [7], the only rational approach of this GMO design problem is to use the math modelling tools of bioinformatics, systems biology, (bio)chemical engineering, and nonlinear systems control theory [1,2,8].

Additionally, due to the modular structure of GRC [1], the modular GRC dynamic models, of an adequate mathematical representation, seem to be the most comprehensive mean for a rational design of the regulatory GRC with desired behavior [8]. The economic advantages offered by novel enzymes used in industrial biosyntheses are very important, that is (Figure 1): 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. All these, make production of novel enzymes by protein engineering (using GMO) to quickly impulse development of the so-called "Computational Systems Biology" [69], based on advanced math models, and extended *omics* databanks.

6.Systems Biology – A Very Short History

The *in-silico* GMO design subject belongs to the Systems Biology area. According to [9], 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 GMO 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 available information and the utilisation scope [2].

At this point, it is to mention that living cells are evolutionary, auto-catalytic, self-adjustable structures able to convert raw materials from environment into additional copies of themselves. Living cells are organized, self-replicating, evolvable, and responsive biological systems to environmental stimuli (Figure 11,12). The structural, functional, and temporal cell organization, including components and reactions, is extremely complex, comprising involving $O(10^{3-4})$ components, $O(10^{3-4})$ 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 [2,3]. Relationships between structure, function and regulation in complex cellular networks are better understood at a low (component) level rather than at the highest-level [1,2,10].

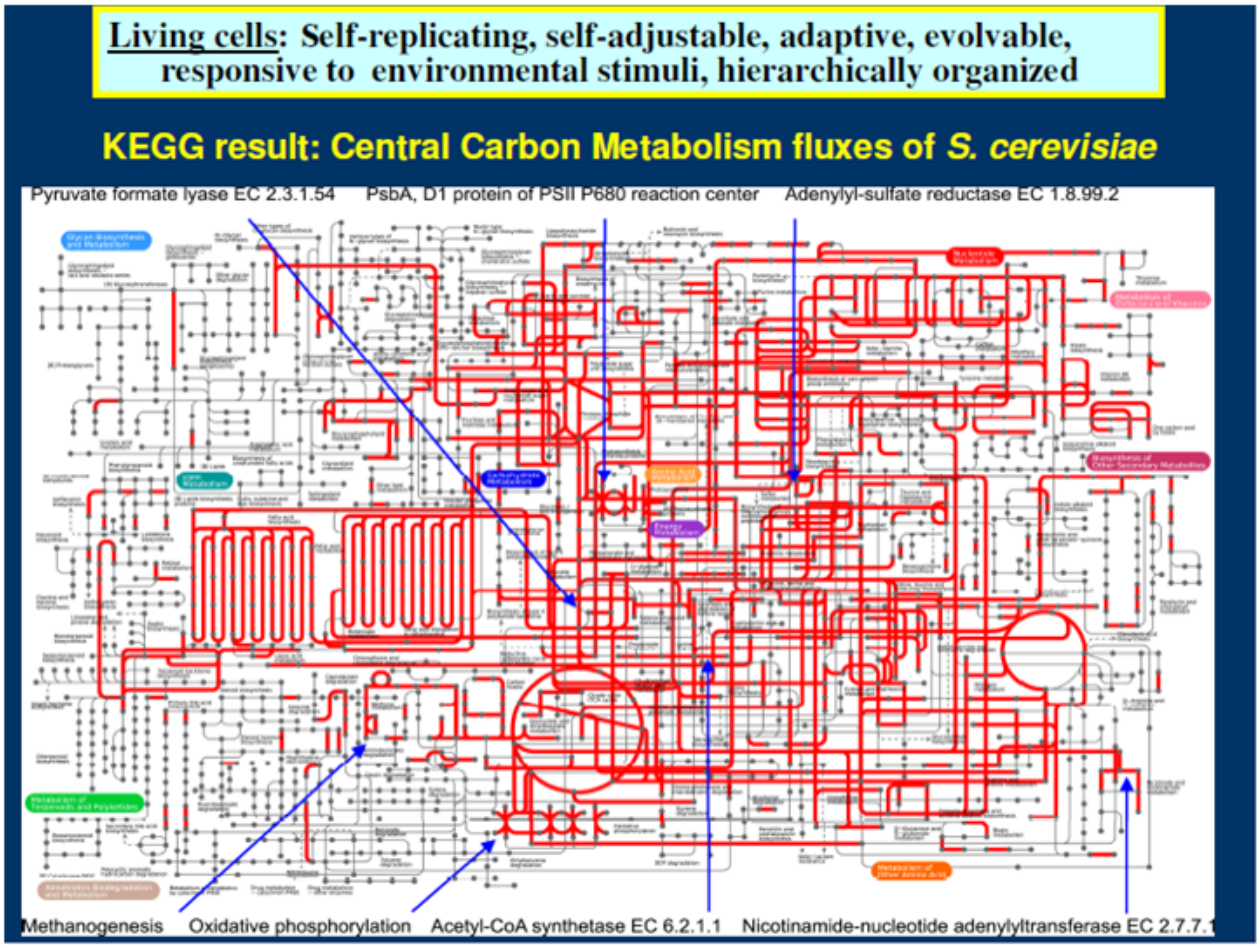


Figure 11: Central carbon metabolism in a cell.

(Source=https://www.researchgate.net/figure/Kyoto-Encyclopedia-of-Genes-and-Genomes-KEGG-pathway-analysis-of-H-pluvialis_262267927)

Living cells: hierarchical organization

Self-replicating apparatus	Time scale separation (slow / fast manifolds)	Self-replication	Regul. net
Replisome, Partitioning apparatus, Z-ring		Nucleoid replication & partitioning, cell division	Cell cycle regulation
Nucleoid		Supercoil and organize genome	Gene expression regulation
Ribosomes, Genome, Energy harnessing apparatus	Intermediate characteristic time	Protein synthesis, Store genetic info, Harness energy	
Cell wall, Nucleic acids, Coenzymes		Metabolic cycles, pathways, Transcription, Translation	Metabolism regulation
Peptidoglycan, Membrane, Protein cplx., Nucleotides	Succession of events	Catalysis, Energy currency	Regulation of enzyme activity
Lipids, Proteins, Nucleosides		Catalysis, Hydrophobic effects	
Saccharides, Fatty acids, Aminoacids	Transient recovering time	Intermediates and building blocks for cell structures and functions	
Simple metabolites			
Raw materials (nutrients)	←Temporal Hierarchy→	Source of energy and material	
←Structural Hierarchy→		←Functional Hierarchy→	

Figure 12: The hierarchical organization of a living cell [2].

In spite of the near astronomic complexity of cell metabolism, attempts to develop math models of metabolic processes are quite old. It is here to mention the early works of [70,71] (Figure 21). But, it is the famous cryptanalyst Alan Turing who clearly proved on 1952 [67] the direct relationships between the (bio)chemical metabolic reactions, the metabolic processes in living cells, and their morphogenesis. When modelling living cells, there are to mention the applied concepts of nonlinear system control theory [6], especially in modelling GRC [6] (Figure 24). They introduced principles to describe systems with interacting components, applicable to biology, cybernetics, and other fields. Bertalanffy proposed that the classical laws of thermodynamics applied to closed systems, to also be applied to the "open systems" such as living cells.

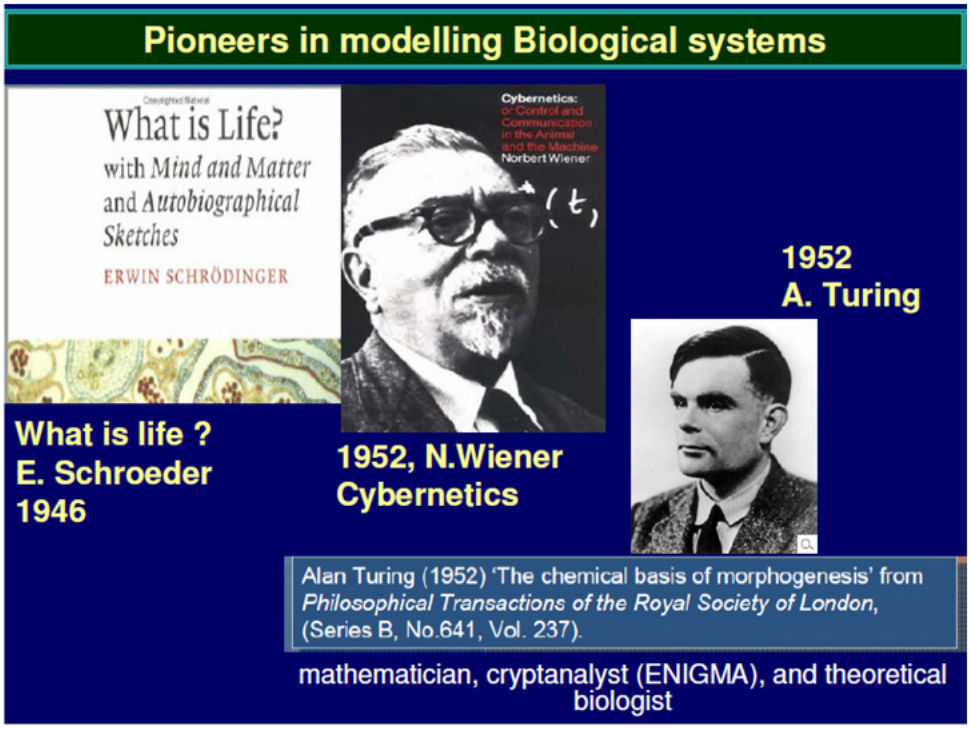


Figure 21: Pioneers in modelling biological systems.

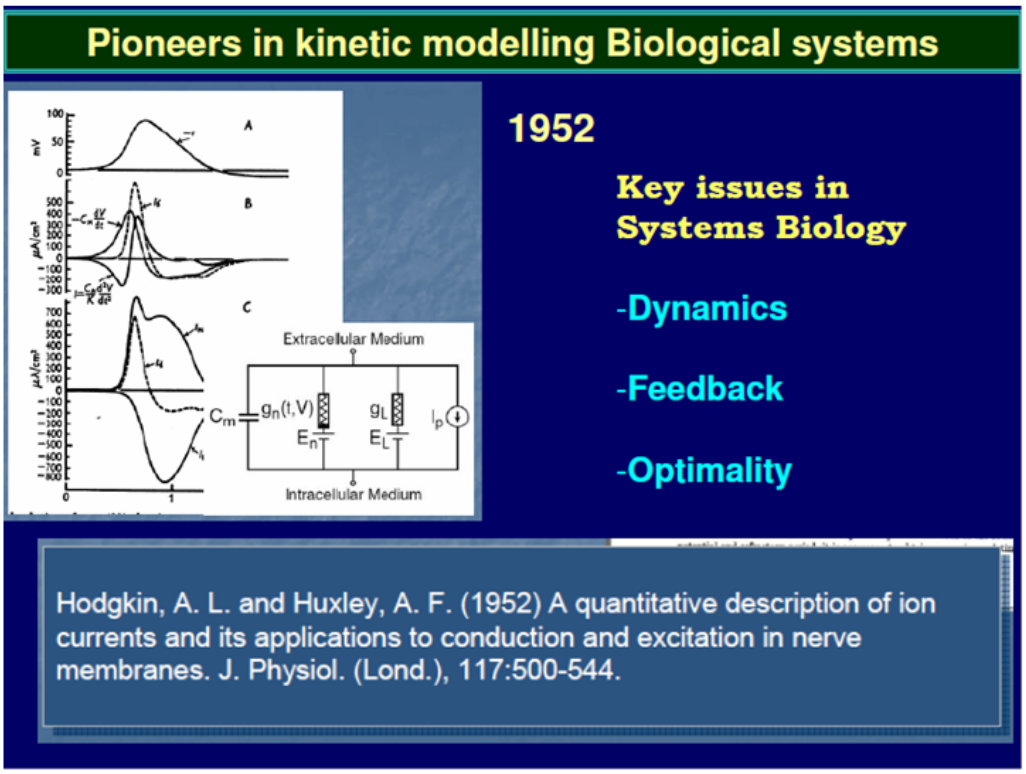


Figure 22: First model attempts are inspired from the electric circuit's theory.

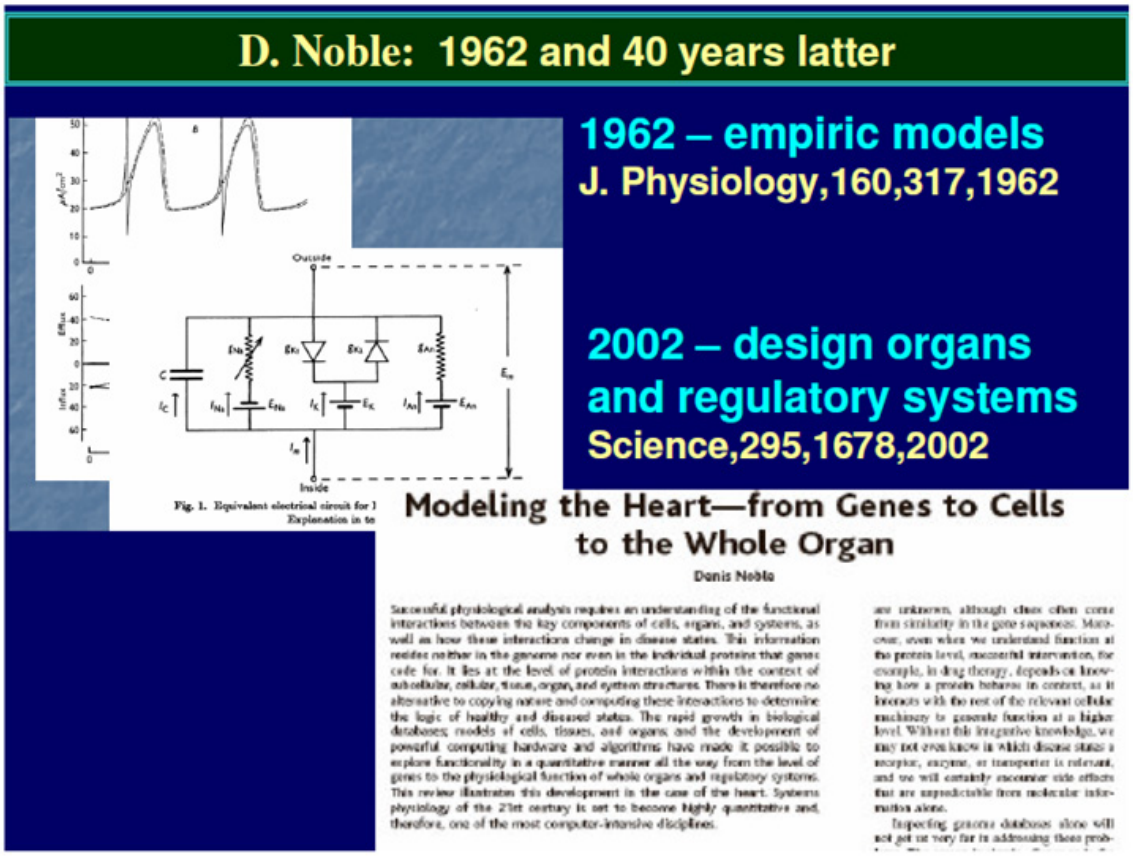


Figure 23: Evolution of Systems Biology from empirical cell models [24] to modelling whole organs [25].

While the first cell modelling attempts have been very empiric, being proposed by electronists [23,24] (Figure 22), 40 years latter Noble published a review paper [25] (Figure 23), pointing-out the tremendous advanced in the Systems Biology for *in-silico* design of GMO, or even tissues, by means of computational systems biology [26,27].

Later, such 'electronic circuits-like' 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 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 below noted: "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" [72].

The first attempts to model the cell metabolism includes modest topological models belonging to the so-called Metabolic Control Analysis [4-6] (Figure 24,27). Such structure-oriented analyses ignore some mechanistic details and the process kinetics, and use the only network topology to quantitatively characterize to what extent the metabolic reactions determine the fluxes and metabolic concentrations [6]. MCA is focus on using various types of sensitivity coefficients (the so-called 'response coefficients'), which are quantitative measures of how much a perturbation [influential variable x(j)] affects the cell-system states y(i) (e.g. r = reaction rates, J = fluxes, C = concentrations) around the "Quasi-Steady-State" (QSS, of index 's'), i.e. $[(\partial y(i)/y(i,s))/(\partial x(j)/x(j,s))]$. The systemic response of fluxes or concentrations to perturbation parameters (i.e. the 'control coefficients'), or of reaction rates to perturbations (i.e. the 'elasticity coefficients') have to fulfil the 'summation theorems', which reflect the network structural properties, and the 'connectivity theorems' related to the properties of single enzymes vs. the system behaviour.

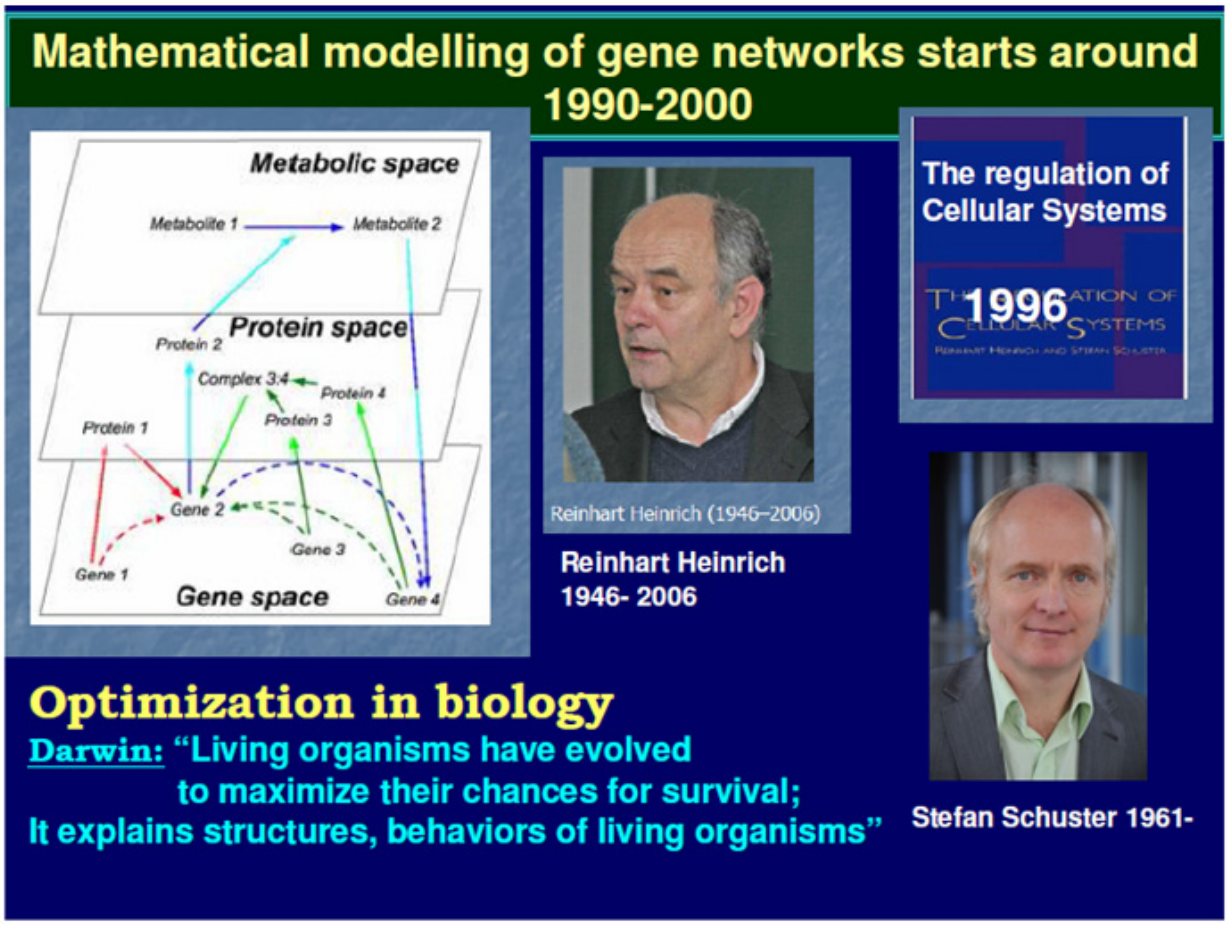


Figure 24: Pioneers in modelling genetic regulatory circuits [6].

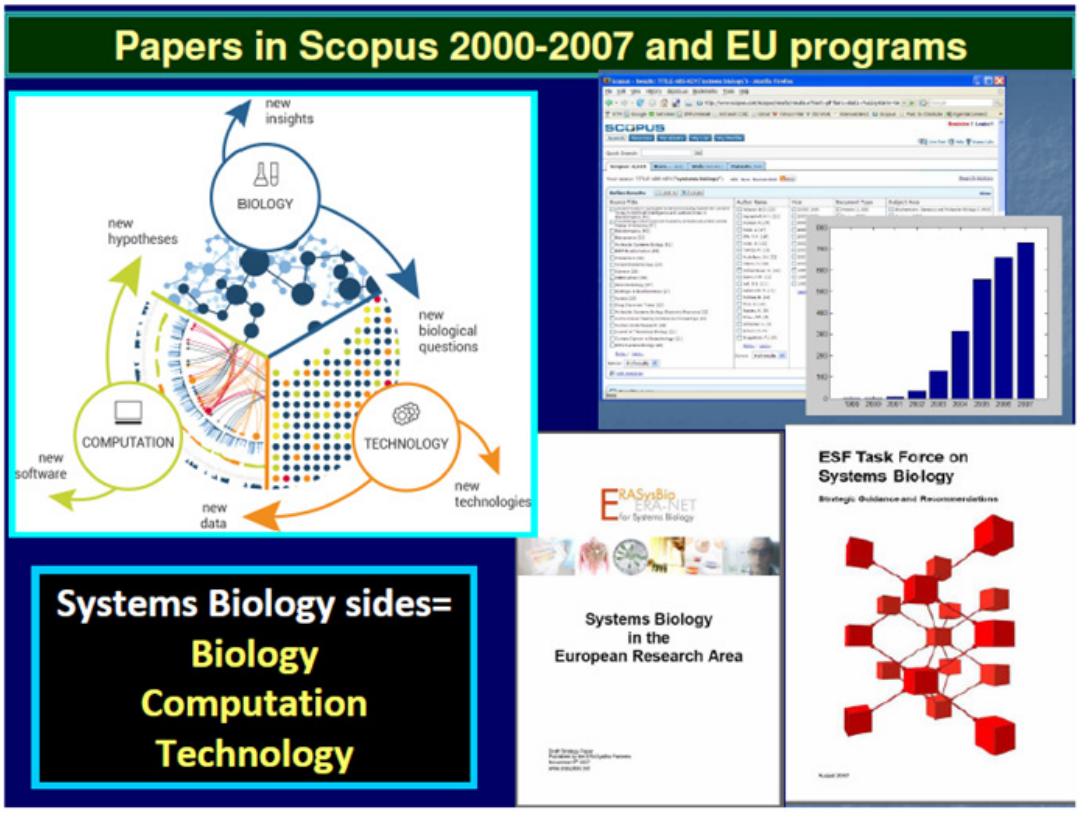


Figure 25: Exponential-like increase of publications (Scopus indexed), and of EU projects on Systems Biology after 2000.

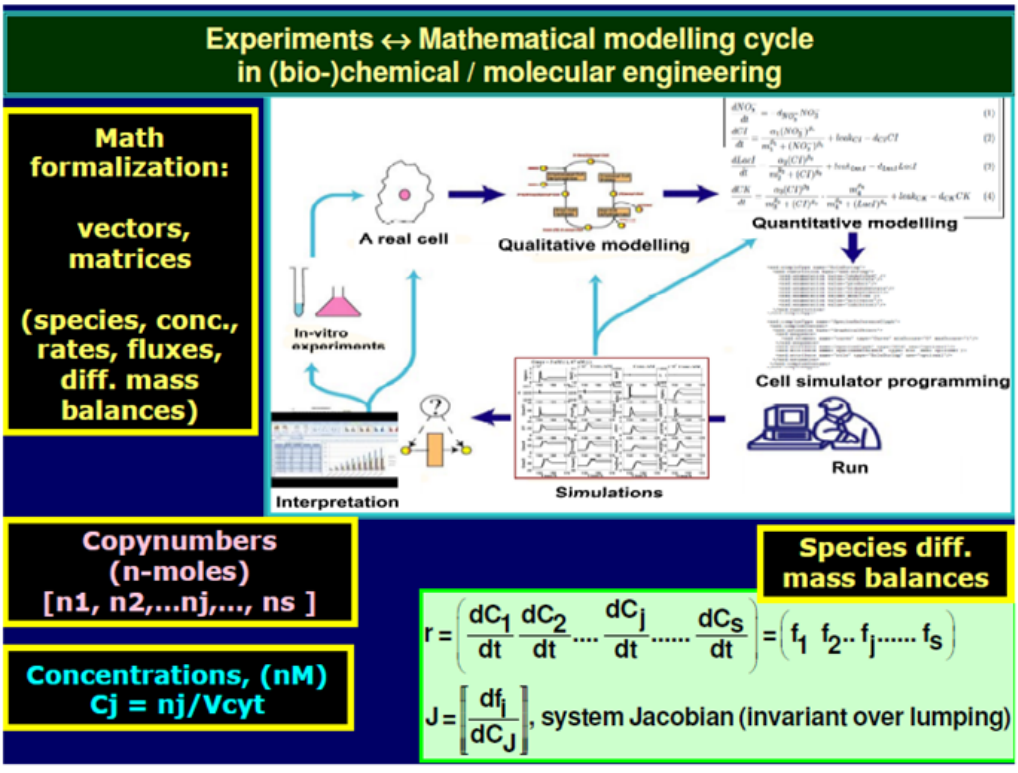


Figure 26: Experiment-modelling cycle to get math models in biochemical/metabolic engineering. Math formalization includes working with vectors and matrices.

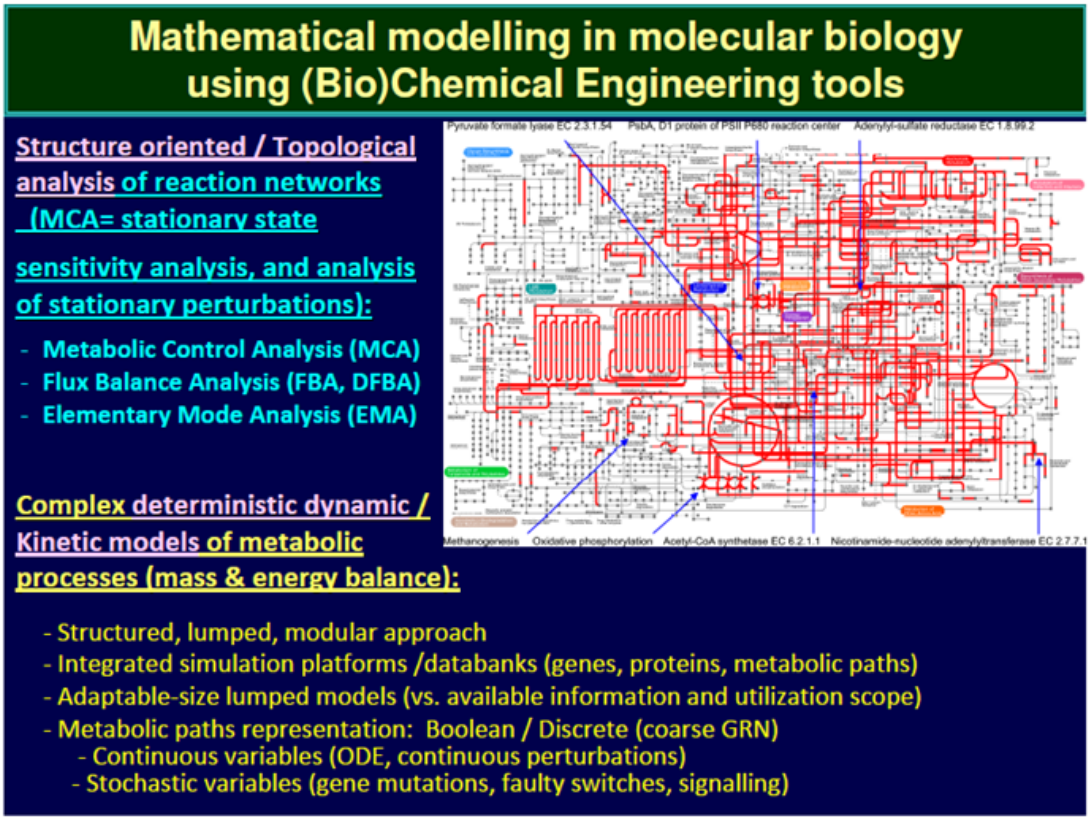


Figure 27: The main types of math models used in Systems Biology.

Originally, MCA has been introduced by Kacser & Burns [4], Heinrich & Rapoport [18], and Burns et al. [19] to quantify the rate limitation in complex enzymatic systems. MCA have been followed by a large number of improvements, mainly dealing with the control analysis of the stationary states, by pointing-out the role of particular reactions and cell components in determining certain metabolic behaviour. Successive extensions of such definitions allow: to study any limit set for non-steady/time-dependent conditions; the flux balance analysis and their optimization (FBA); elementary mode analysis (EMA); dynamic flux balance analysis (DFBA); extreme pathway analysis (ExPA); constrained based modelling of metabolic network (CBM) [17]. MCA was very much applied in determining the metabolic fluxes (i.e. metabolic stationary reaction rates), by applying (non)linear optimization procedures, with applications in metabolic engineering, and GMO design [58,73,74].

Among advantages, MCA methods are able to efficiently characterize the metabolic network robustness and functionality, linked with the cell phenotype and gene regulation. MCA also allows a rapid but rough evaluation of the system response to perturbations (especially of the enzymatic activity), possibilities of control and self-regulation for the whole path or some subunits. Functional subunits are metabolic subsystems, called 'modules', such as amino acid or protein synthesis, protein degradation, mitochondria metabolic path, etc.[20]. Because the living cells are self-evolutive systems, new reactions recruited by cells together with enzyme adaptations can lead to an increase in the cell biological organisation and to optimal performance indices. However, important limitations of MCA, related to any dynamic analysis of the metabolic system, system behaviour linearizations when computing the response coefficients, omission of the influence of the cell variable volume and of the osmotic pressure, any consideration on the regulatory modules vs. stationary and/or dynamic perturbations, all these make application of MCA to be limited. Consequently, when designing GMO, a kinetic analysis of the system with a deterministic adequate model is also necessary.

7.Systems Biology–Modern Concepts

With the accumulation of experimental information, and its storage in omics databanks, the math models and algorithms dedicated to cell process simulation reported a sharply increase after 1990.

Additionally to novel concepts introduced, various definitions of Systems Biology exist in the dedicated literature [28](Figure 28):

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" [9].

II. "System Biology is a comprehensive quantitative analysis of the manner in which all the components of a biological system interact functionally over time" [75].

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).

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. vii)"Modelling is not the final goal, but 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 GRC and modified metabolic fluxes in a GMO with desired characteristics [1,2,8]. The increasingly use of math models and numerical calculus when design GMO, has several reasons (Figure 13-18): i) The tremendous increase of the computers powers; ii) Completion of various genome projects, and starting the post-genomic era; iii) A tremendous increase in (experimental) data in the omics databanks (genomics, proteomics, metabolomics, fluxomics, etc.), and advances in high-throughput experiments; iv) An important experimental effort saving.

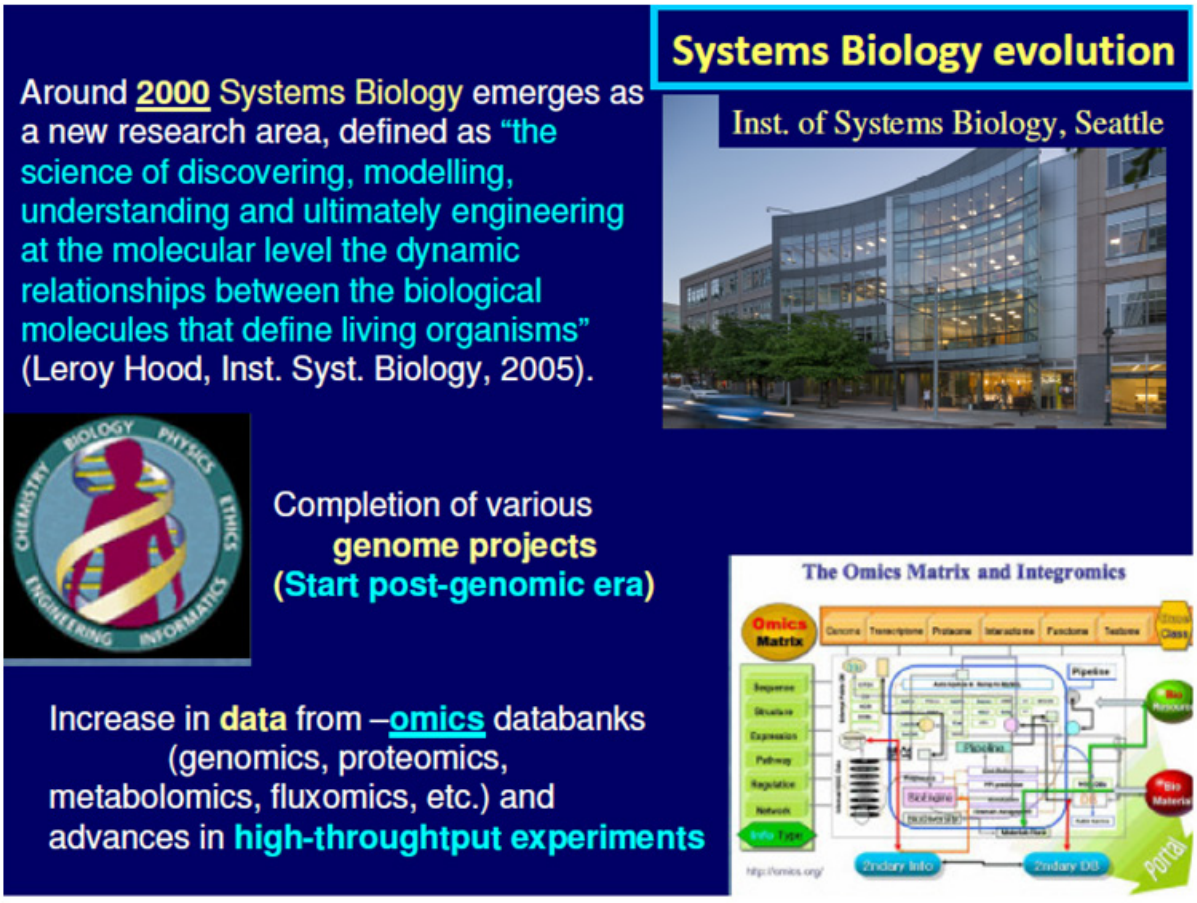


Figure 13: Steps in the evolution of the emergent systems biology.

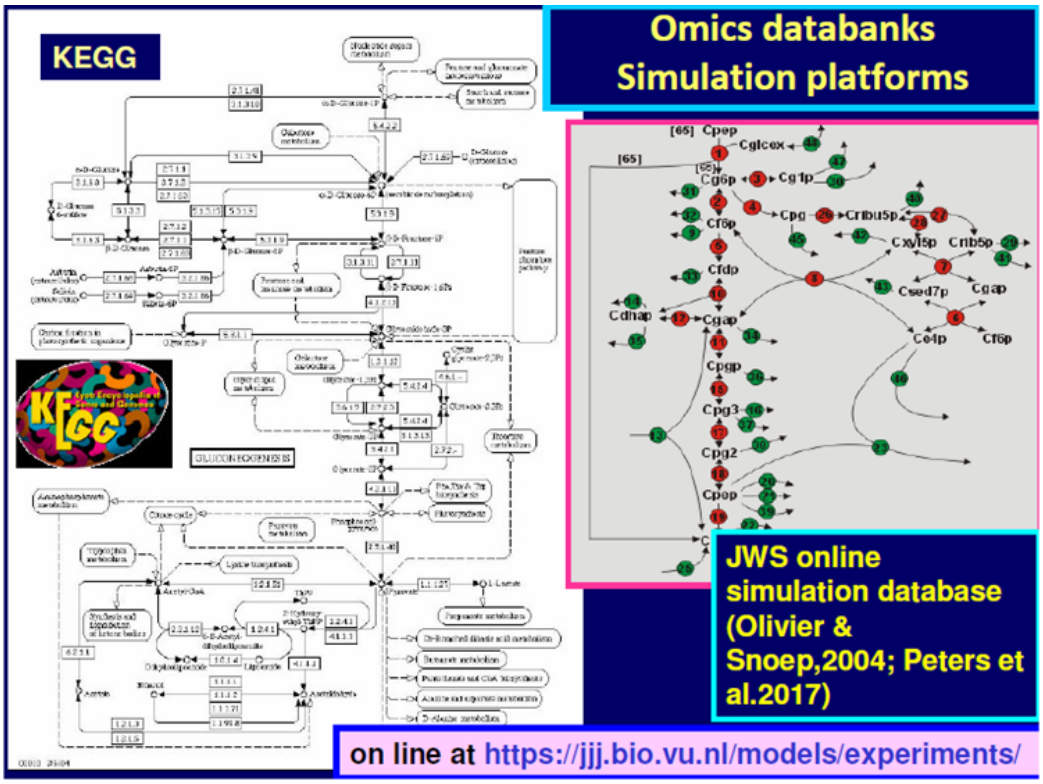


Figure 14: Some developed-omics databanks (e.g. KEGG [15,40], JWS [41,42]).

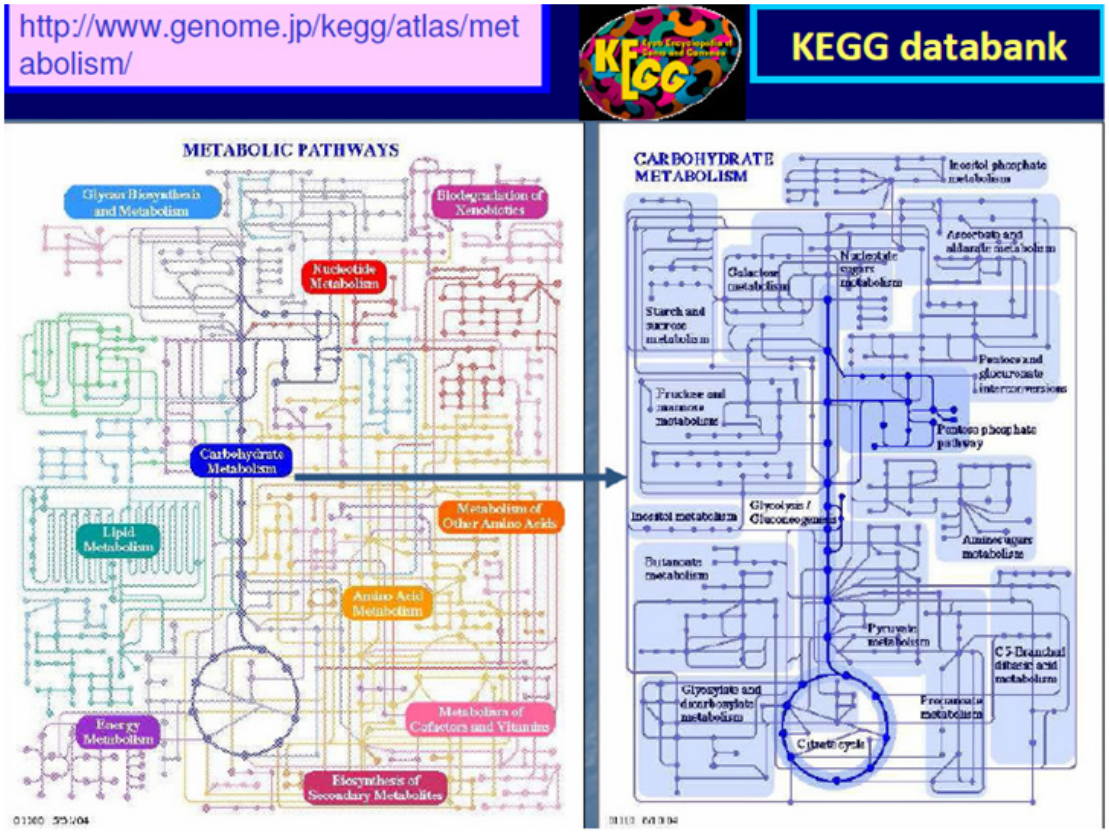


Figure 15: (continuation) KEGG databank [15,40].

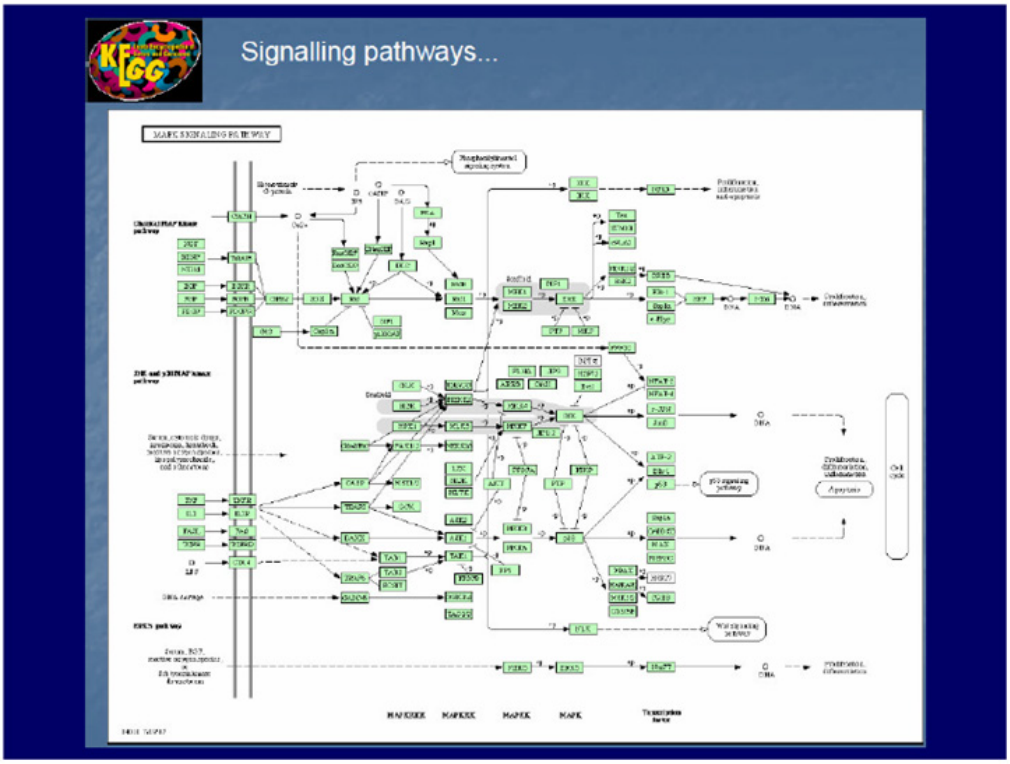


Figure 16: (continuation) KEGG databank [15,40].

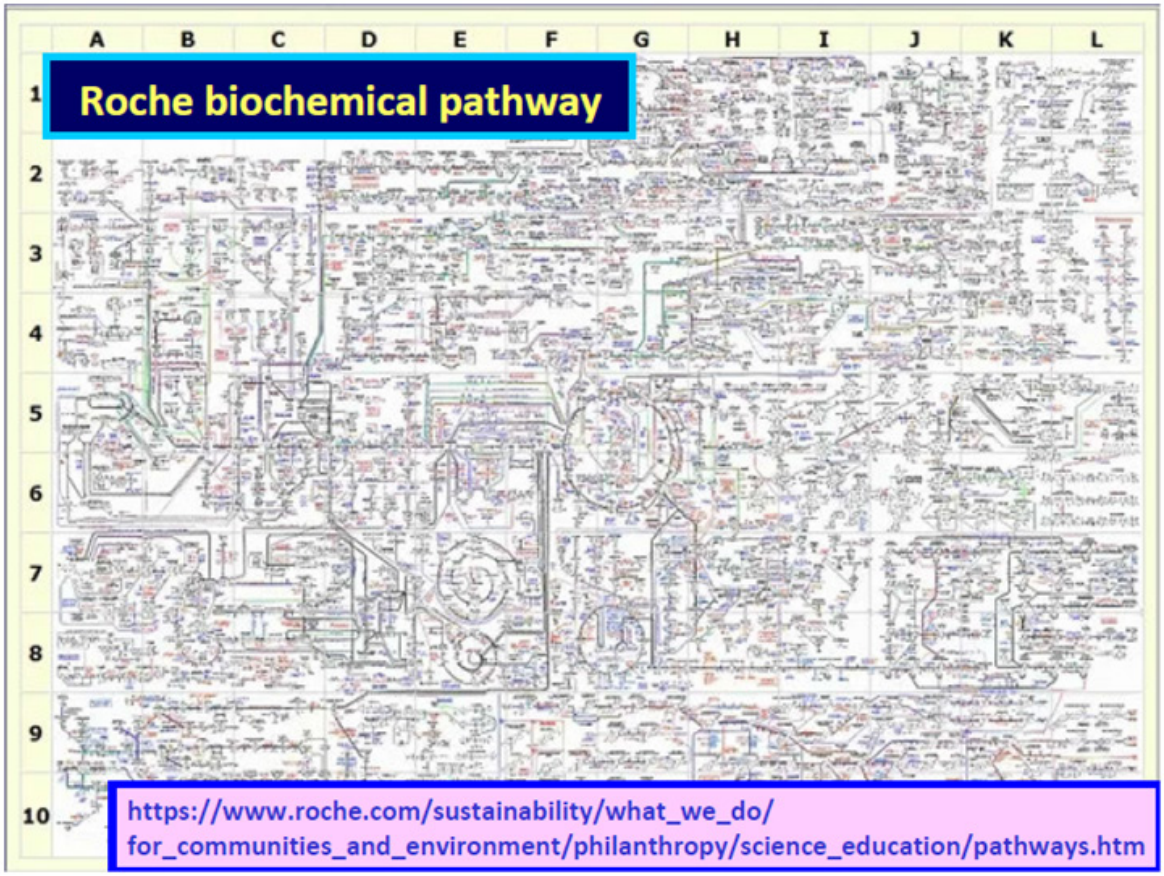


Figure 17: (continuation) Roche biochemical databank [43].

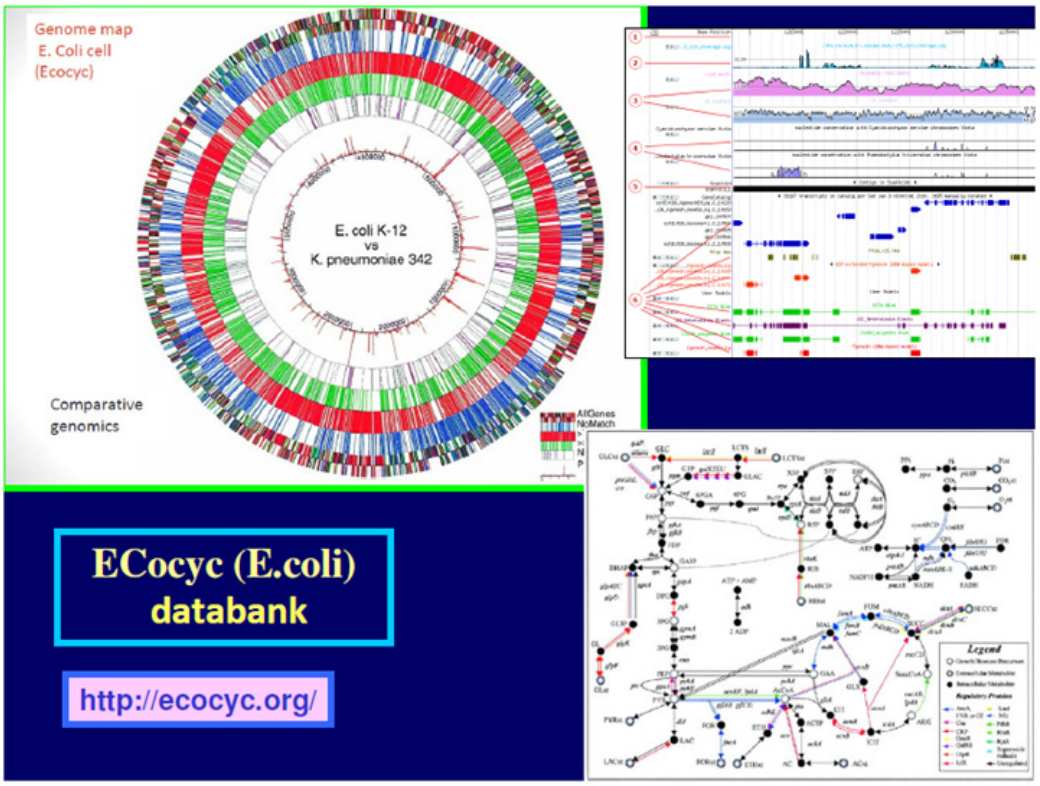


Figure 18: (continuation) EcoCyc biochemical databank [44].

Around 2000 the human genome has been deciphered (Figure 19), which has boosted Systems Biology development. Consequently, in the "post-genomic era" a large number of Systems Biology projects have been developed, and the number of dedicated publications increased near exponentially in Scopus (Figure 25). Among the structured cell simulators able to simulate the whole or parts of the cell metabolism, are to be mentioned [17,46](Figure 36):

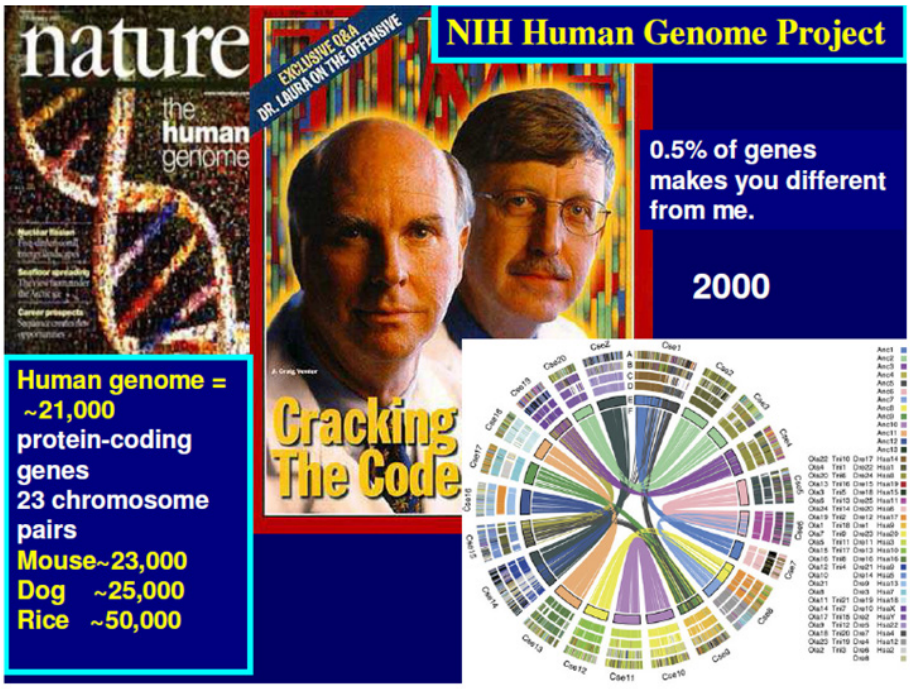


Figure 19: Around 2000 the human genome has been deciphered.

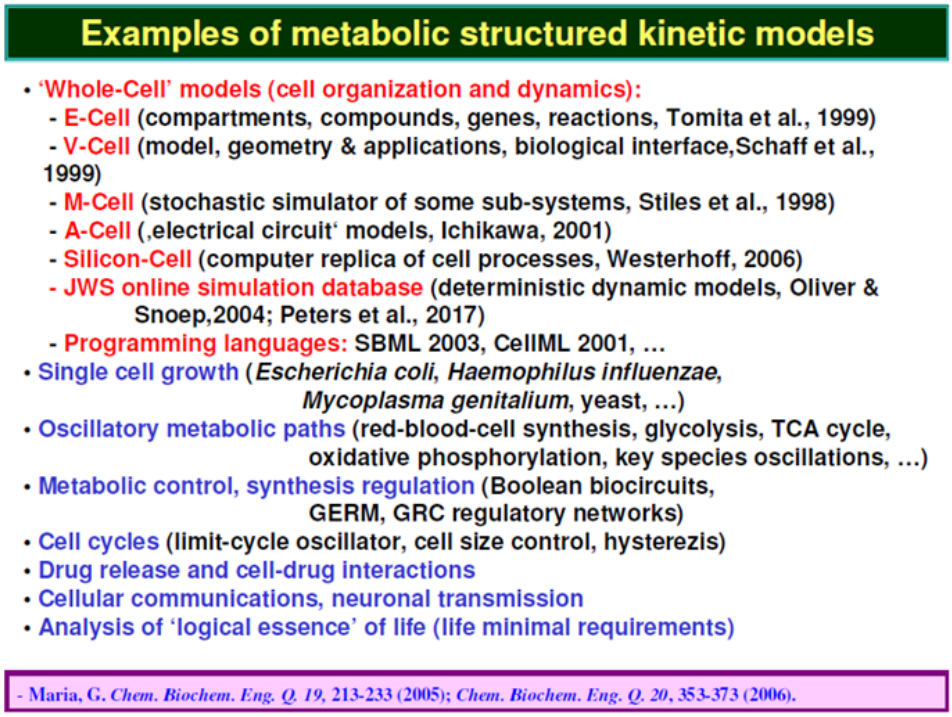


Figure 36: Some examples of structured cell simulators [17,46].

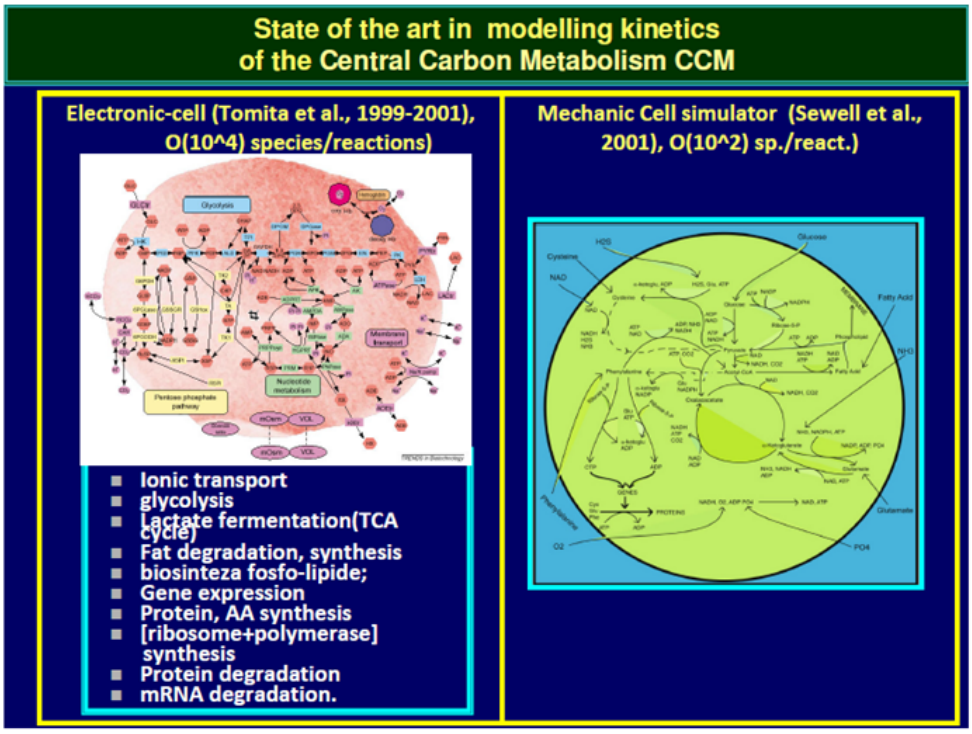


Figure 37: Some simulators of the central carbon metabolism.

i. E-Cell (compartments, compounds, genes, reactions) [76] (Figure 37)

ii. V-Cell (model, geometry & applications, biological interface) [77].

iii. M-Cell (stochastic simulator of some cell sub-systems) [78] (Figure 39)

iv . A-Cell (‚electrical circuit' like models) [47] (Figure 39)

v). Silicon-Cell (computer replica of cell processes to be linked [79] (Figure 38)

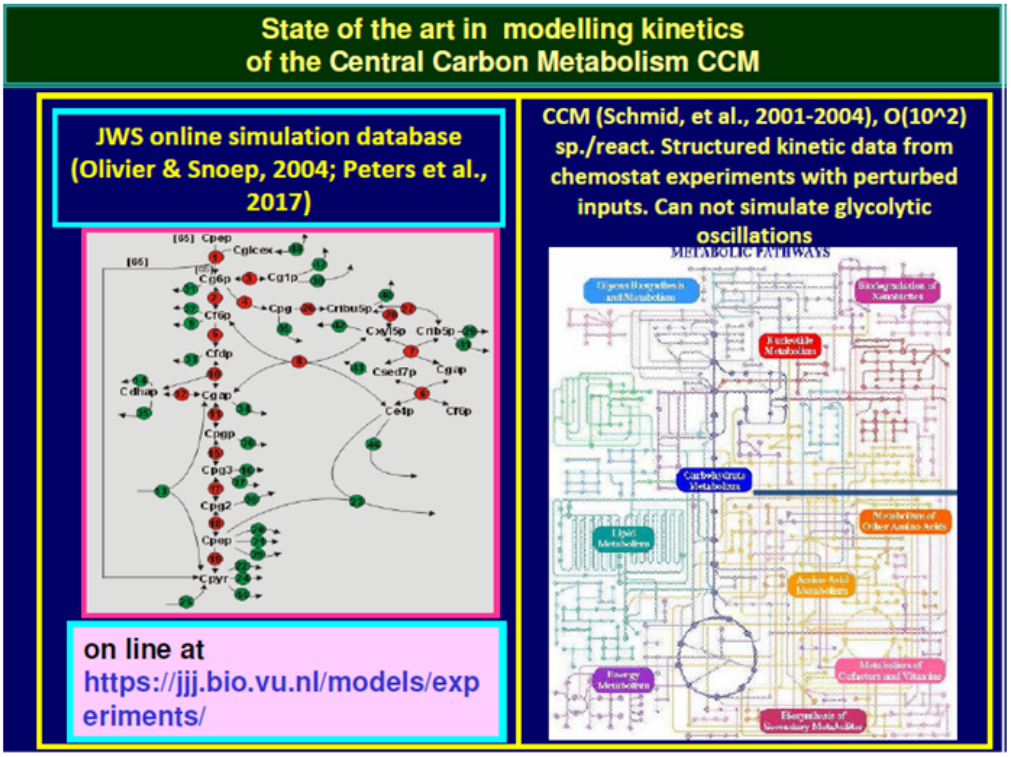


Figure 38: (continued) some simulators of the central carbon metabolism.

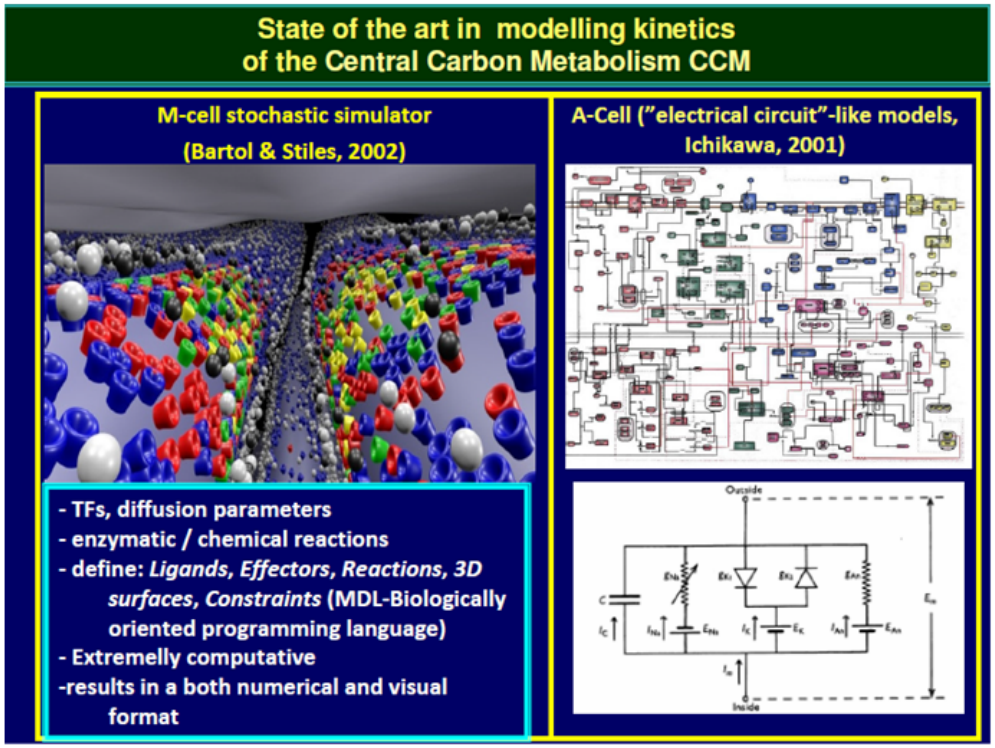


Figure 39: (continued) some cell simulators of stochastic type (left, Bartol & Stiles, (2002) [31]), and of electric-circuit type (right, Ichikawa, 2001 [47]).

vi. Specific programming languages (SBML), or on-line simulation platforms [11], etc. (Figure 38)

vii. Single cell growth (e.g. *Escherichia coli, Haemophilus influenzae, Mycoplasma genitalium*, yeast, etc. (Figure 40)

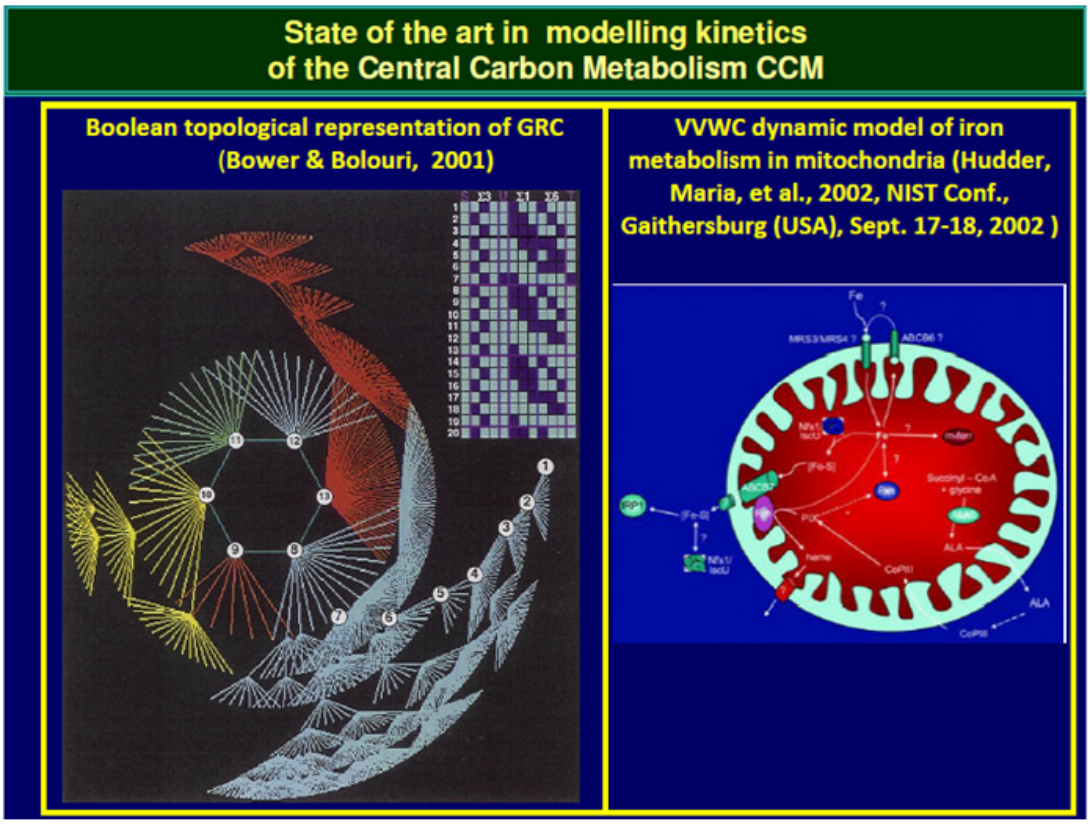


Figure 40: (continued) some cell simulators: one of Boolean type (left, [48]), and one concerning the iron metabolism in mitochondria (right, [49]).

viii. Model metabolic oscillations (red-blood-cell synthesis, glycolysis, TCA cycle, oxidative phosphorylation, key species oscillations, etc.) [160].

ix. Metabolic control of protein synthesis regulation (GERM, GRC) [1,82-86].

x. Modelling the central carbon metabolism (CCM)[59,80,81]; (Figure 82)

xi. FBA based design of GMO [87](Figure 33-35)

xii. Modelling the cell cycle [88,89]

xiii. Modelling the drug release and cell-drug interactions

xv. Modelling cellular communications, neuronal transmission

xvi. Analysis of 'logical essence' of life (life minimal requirements)

At the same time, the exponential-like increase of the experimental biological information lead to development of valuable **omics** databanks, such as (Figure 14-18); [14,15,40,43,44] etc.

However, it is only over the last decades when Systems Biology reported notable successes due to a considerable increase in computing power of the modern computers. It is to mention here, for instance, the cell simulator platforms, and online model repository JWS of Snoep & Olivier [11], or those developed by Rocha et al. [12], or by Tomita et al. [13], together with continuous expansion of the omics databases [14,15], and reported advances in the numerical algorithms used by bioinformatics, (bio)chemical engineering, and nonlinear systems control theory [2].

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 [6,32,33,34,90] 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 25).

As stated by Nandy [29] & Sutherland[36], tremendous applications of Systems Biology have been reported over the next decades in the below areas (Figure 20).

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

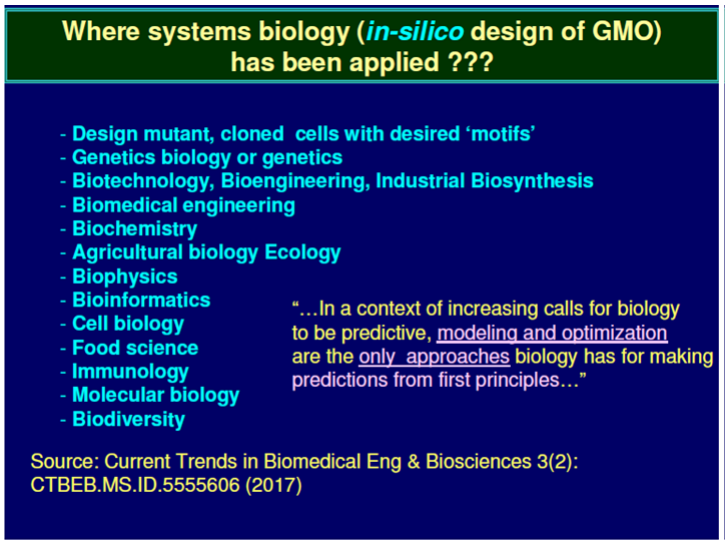


Figure 20: Area where Systems Biology has been applied.

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…" [28]. Due to the computing facilities offered by the algorithmic rules developed by the (bio)chemical engineering and nonlinear systems biology rules (Figure 26-28,31,32), the developed cell math models use a vectorial-matriceal approach (Figure 26), with a continuous model upgrading based on dynamic experimental data recorded in a chemostat (i.e. a continuously operated bioreactor), operated under steady-state, or in a dynamic regime following an input perturbation in the substrate concentration in the solution fed in the bioreactor [59,87].

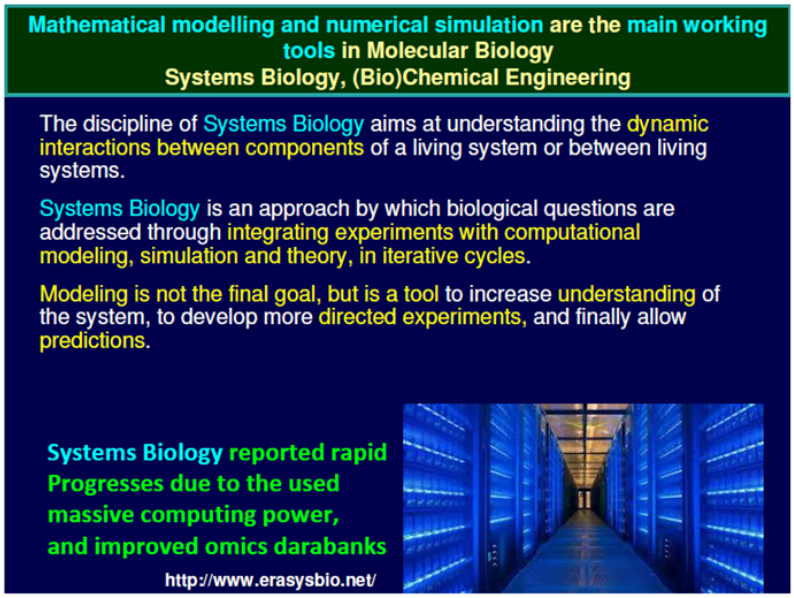


Figure 28: Math models used in Systems Biology requires massive computational efforts and facilities

So that, 40 years after the first reported empirical models of living cells [24], the optimistic researchers advanced very ambitious targets, such as "Modelling the heart – from genes to cells to the whole organ" (Figure 23).

As in all scientific controversies, sceptic opinions also exist, such as (Figure 29-30):

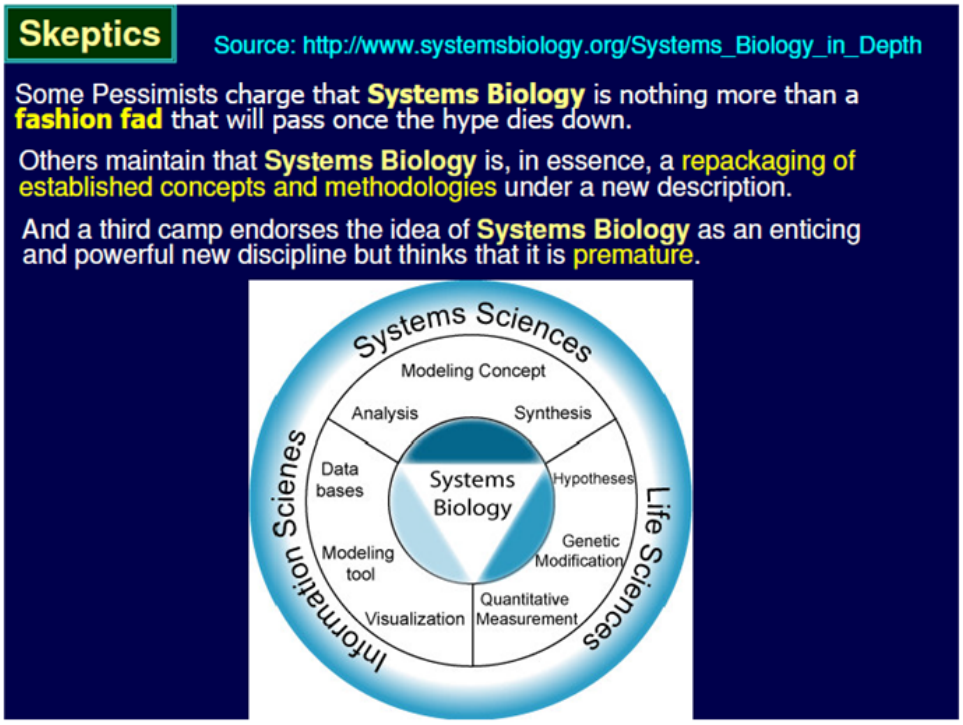


Figure 29: Sceptic opinions on Systems Biology.

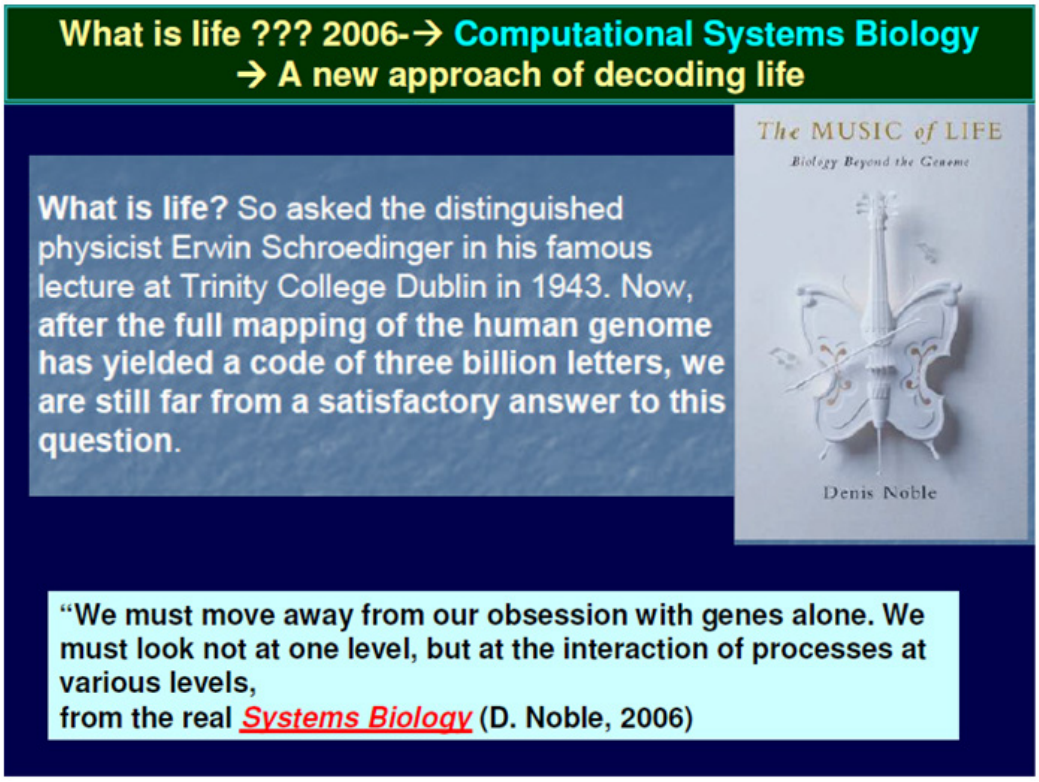


Figure 30: Defining concepts in the computational Systems Biology [45].

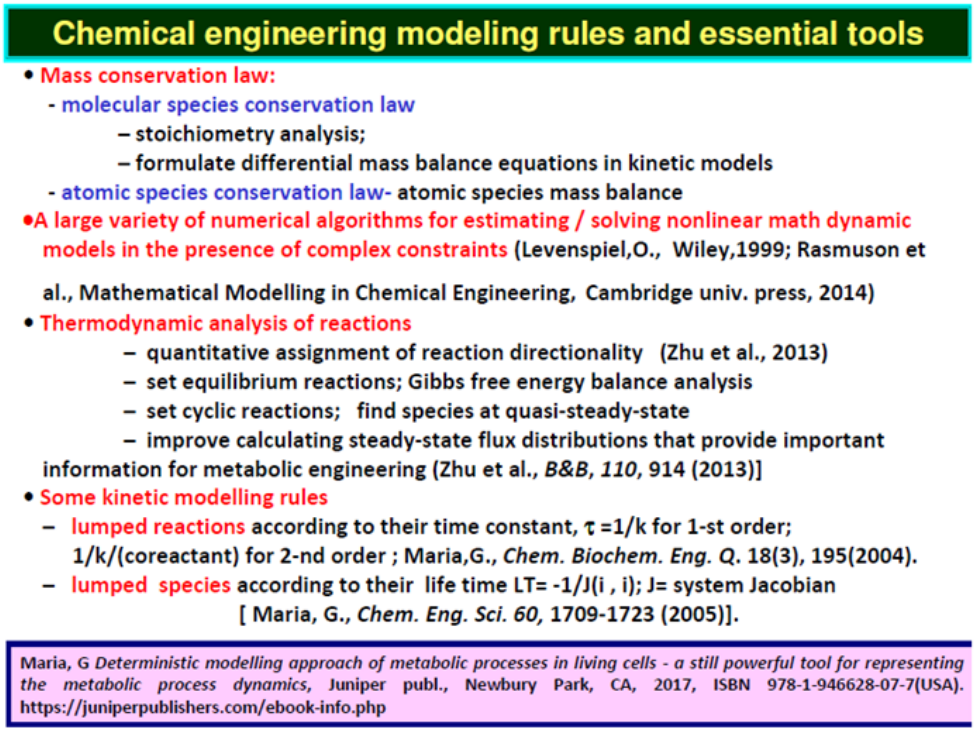


Figure 31: Chemical engineering modelling rules and concepts applied in Systems Biology [1,2].

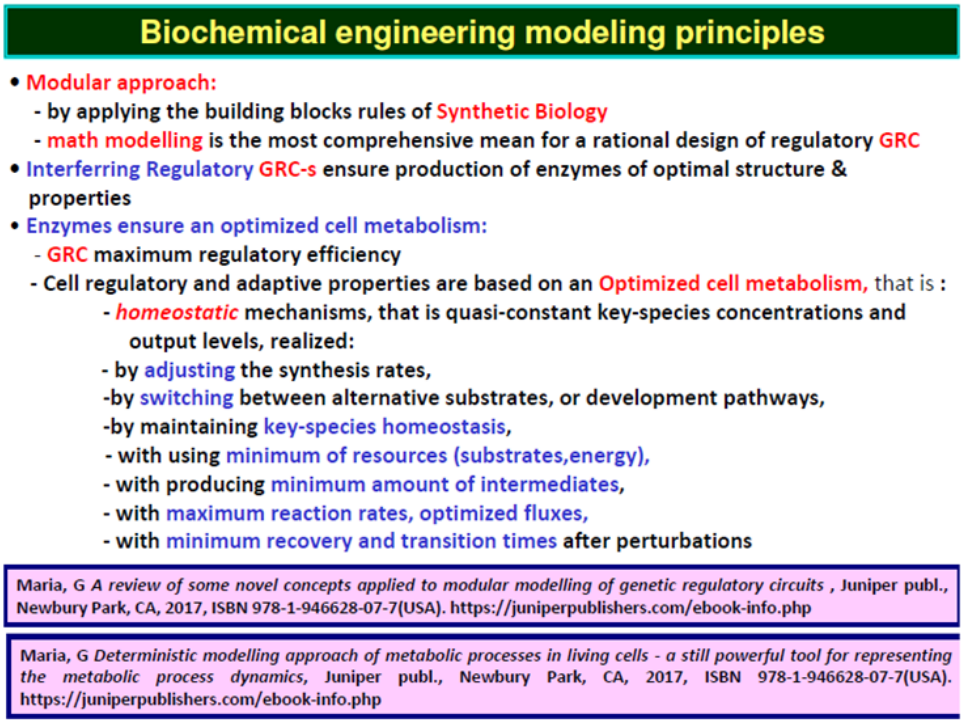


Figure 32: Biochemical engineering modelling rules and some concepts to be applied in Systems Biology.

A. "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 [45]: 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, from the real Systems Biology.; b) The reductionist approach of molecular biology has proved itself immensely powerful. But DNA isn't life."

B. 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.

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.

8.Systems Biology – Math Modelling Tools

Most of GMO used in the industrial synthesis are prokaryote cells (Figure 5). 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 6). This can be done by changing the characteristics or even by removing certain enzymes catalyzing some of the reactions. As 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 9) [39].

i. delete genes (i.e. "gene knock-out", Maria [157, 203-206]),

ii. clone the cell with target plasmids, thus modifying the expression level and the encoded enzyme concentration [151],

iii. replace a target gene with another [68],

iv. *in-silico* (model-based) investigate the bioreactor operating conditions leading to significant changes in the cell target fluxes [163-164] and eventually adjust the cell content dilution rate to match the bioreactor dilution rate [164].

All genome/proteome modifications can be much easily, and with less (experimental) cost investigated if an adequate (extended or reduced) dynamic model of the cell metabolism is available (Figure 10). Such a dynamic structured model, can quickly predict the metabolic fluxes and their dependence on the external, or enzymatic factors. Alternatively, the Flux Balance Analysis [73], can be applied (Figure 33-34).

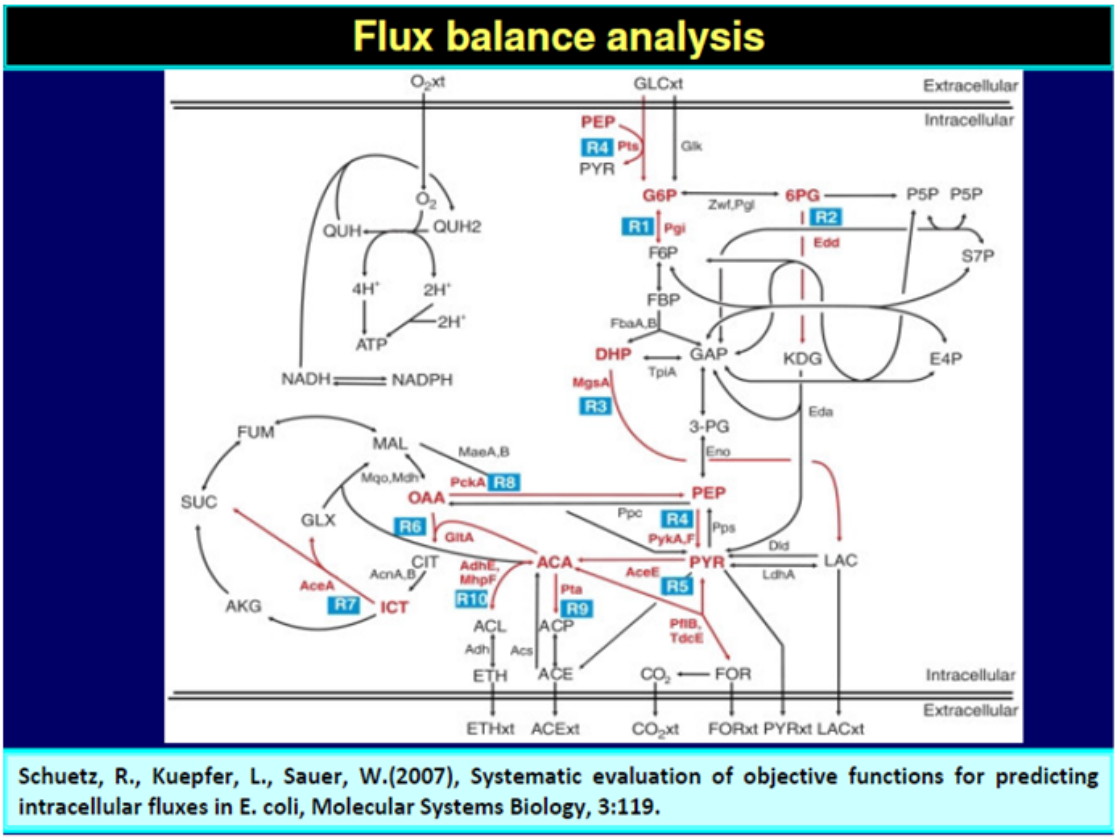


Figure 33: Flux balance analysis-the main objective of metabolic engineering, and an essential preliminary step in design GMO.

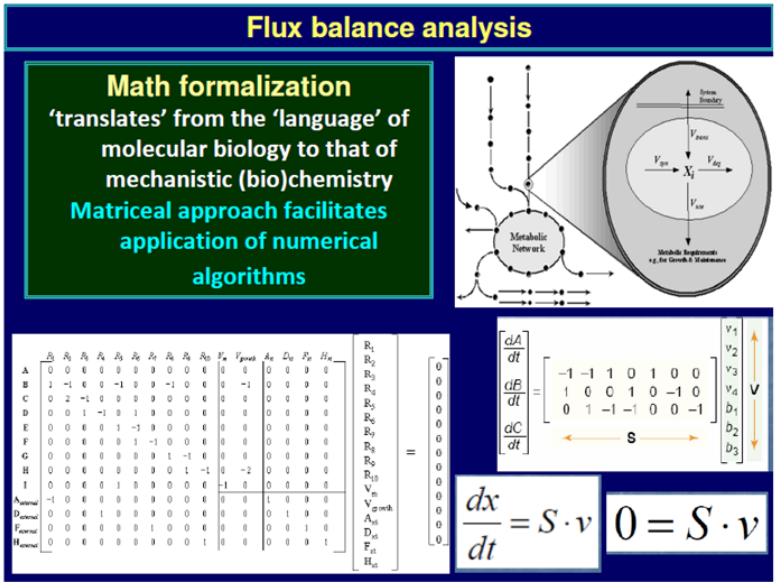


Figure 34: (continuation) Flux balance analysis is working with matrix math models [73].

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 34,35) [28,92]. Here it is to mention the short review of [87] of how to use FBA to get GMO, by also using COBRA toolbox [93] to re-construct reaction pathways by using Kegg Reaction [15,40].

To illustrate in a simple way how a dynamic model of a metabolic pathway can be used to design a GMO, let us consider the problem of [38] (Figure 7): re-configure the metabolic pathway for Phenyl-alanine synthesis in *E. coli*, to maximize its production. 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 retained); 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 8), 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 8), it is to observe the large differences between the two pathways into the cell, fully achievable by genetic engineering.

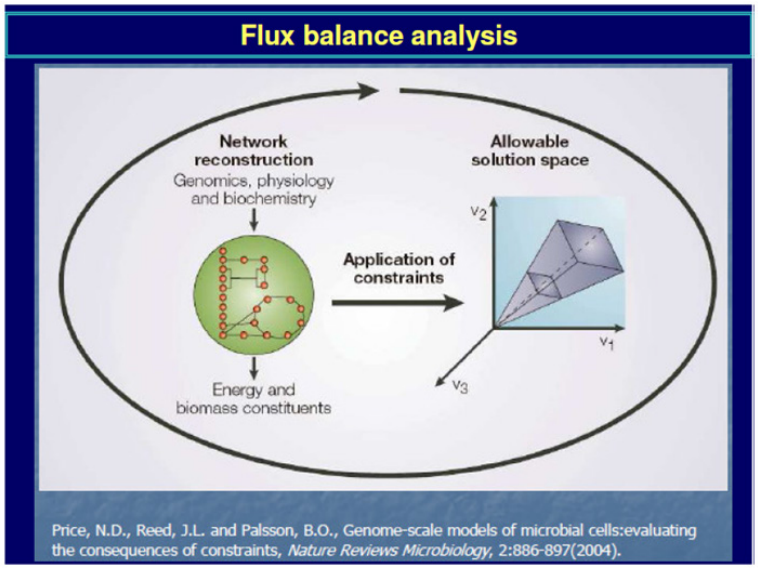


Figure 35: (continuation) solving a Flux balance analysis problem translates in (non)linear programming problem.

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. The first attempts to model the cell metabolism includes modest topological models belonging to the so-called Metabolic Control Analysis [4-6] (Figure 27). Such structure-oriented analyses ignore some mechanistic details and the process kinetics, and use the only network topology to quantitatively characterize to what extent the metabolic reactions determine the fluxes and metabolic concentrations [6]. 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; besides these coefficients have to fulfil the 'summation theorems', which reflect the network structural 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 GMO by optimizing the evolutionary metabolic systems, math models based on the cell bioprocess mechanism are to be used. Such metabolic models, elaborated in a reduced or in an extended form, have been proposed over decades by using the MCA, but also the (bio)chemical engineering, and nonlinear system control concepts and tools (below described), by imposing one or more appropriate performance criteria, such as: maximize production of biomass, maximize synthesis of a target metabolite, maximize reaction rates and steady-state fluxes; minimize metabolic intermediate concentrations; minimize transient times between QSS; optimize the reaction stoichiometry (network topology); maximize thermodynamic efficiency, etc. [6]. Often, a multi-objective optimization is applied to get a desired GMO [157, 203-206]. All these objectives are subjected to various mass balance, thermodynamic, and biological constraints [6]. However, by not accounting for the system dynamics, and grounding the analysis on only the linear system theory, the topological methods (including MCA) presents inherent severe limitations (see for instance some violations of stoichiometric constraints discussed by Atauri et al. [21], or the modified control coefficients of Szedlacsek et al. [22].

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, by preserving the structural hierarchy and component functions. The first dynamic models of some cell processes have been reported by the electronists [23,24]. Later, such 'electronic circuits-like' models have been extensively used to understand intermediate levels of regulation [47], but they failed to reproduce in detail molecular interactions with slow and continuous responses to perturbations and, eventually, they have been abandoned.

However, the electronists underlined the main characteristics of the cell systems, which must be included in any simulation model (Figure 22):

I. the *dynamic* character of species interactions and processes [36];

II. the *feedback* character of processes ensuring their [35];

III. *optimal regulation* [36], with

IV. Consuming *minimum* of resources (nutrients/substrates), and cell energy, but ensuring maximum reaction rates [6].

All these cell metabolic characteristics will be accounted in all the subsequent cell *in-silico* simulators based on extended mathematical models. All these characteristics are in fully agreement to the Darwin theory "Living organisms have evolved to maximize their chances for survival; It explains structures, behaviors of living organisms."(Figure 24). From such very incipient efforts to model-based design GMO, 40 years latter [25] pointed-out the tremendous advanced in the Systems Biology and *in-silico* design of GMO, or even tissues, by means of computational systems biology [26,27].

A review of mathematical model types used to describe metabolic processes is presented by Maria [17], Styczynski & Stephanopoulos [73,94]. Each model type presents advantages but also limitations. Roughly, to model the complex metabolic regulatory mechanisms at a molecular level, two main approaches have been developed over decades: a structure-oriented (topological) analysis, and the dynamic (kinetic) models [10]. Each theory presents strengths and shortcomings in providing an integrated predictive description of the cellular regulatory network.

Structure-oriented analyses or *topological* models ignore some mechanistic details and the process kinetics, and use the only network topology to quantitatively characterize to what extent the metabolic reactions determine the fluxes and metabolic concentrations [6]. The so-called 'metabolic control analysis' (MCA) is focus on using various types of sensitivity coefficients (the so-called 'response coefficients'), which are quantitative measures of how much a perturbation (an influential variable) affects the cell-system states [e.g. reaction rates, metabolic fluxes (stationary reaction rates), species concentrations] around the steady-state (QSS). The systemic response of fluxes or concentrations to perturbation parameters (i.e. the 'control coefficients'), or of reaction rates to perturbations (i.e. the 'elasticity coefficients') have to fulfil the 'summation theorems', which reflect the network structural properties, and the 'connectivity theorems' related to the properties of single enzymes in connection to the system behaviour.

MCA methods are able to efficiently characterize the metabolic network robustness and functionality, linked with the cell phenotype and gene regulation. MCA allows a rapid evaluation of the system response to perturbations (especially of the enzymatic activity), possibilities of control and self-regulation for the whole path or some subunits. Functional subunits are metabolic subsystems, called 'modules', such as amino acid or protein synthesis, protein degradation, mitochondria metabolic path, etc. [20]. Because the living cells are self- evolutionary systems, new reactions recruited by cells together with enzyme adaptations can lead to an increase in the cell biological organisation and to optimal performance indices. When constructing methods to optimize evolutionary metabolic systems, MCA concepts and appropriate performance criteria have been used, leading to: maximize reaction rates and steady-state fluxes; minimize metabolic intermediate concentrations; minimize transient times; optimise the reaction stoichiometry (network topology); maximize thermodynamic efficiency. All these objectives are subjected to various mass balance, thermodynamic, and biological constraints [6] However, by not accounting for the system dynamics, and grounding the analysis on the linear system theory, topological methods presents inherent limitations (see for instance some violations of stoichiometric constraints discussed by Atauri et al. [21], or the use of modified control coefficients [22].

From the mathematical point of view, various structured (mechanism-based) dynamic models have been proposed to simulate the metabolic processes and their regulation, accounting for continuous, discrete, and/or stochastic variables, in a modular construction, 'circuit-like' network, or compartmented simulation platforms [17,48,99]. Briefly, the math models used by Systems Biology are of the following types:

I. *Deterministic continuous* variable models; among other advantages 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* information [6,17,73,48,82, 206].

II. *Boolean* (*discrete*) variable models; such a topological structure is displayed in the (Figure 40) [48,50]. Due to the very large number of states $O(10^3–10^4)$ and $O(10^3)$ of transcriptional factors (TF) involved in the gene expression, such GRC models are organized in clusters, modules, of a multi-layer organization (Figure 40) [48,87];

III. *Stochastic* variable models [100-102].

IV. *Mixed* variable models [48].

A. *Deterministic continuous* variable models; present a large number of advantages, such as: they perfectly can represent the cell response to continuous perturbations; their structure and size can be easily adapted based on the available *omics* information [6,17,48,73,82] etc.. Are also to be underlined the huge advantages coming from the used concepts, rules, and algorithms of (Bio)chemical engineering and nonlinear system control theory, to be below discussed in the chap. 6-7 (Figure 37-38,43-44).

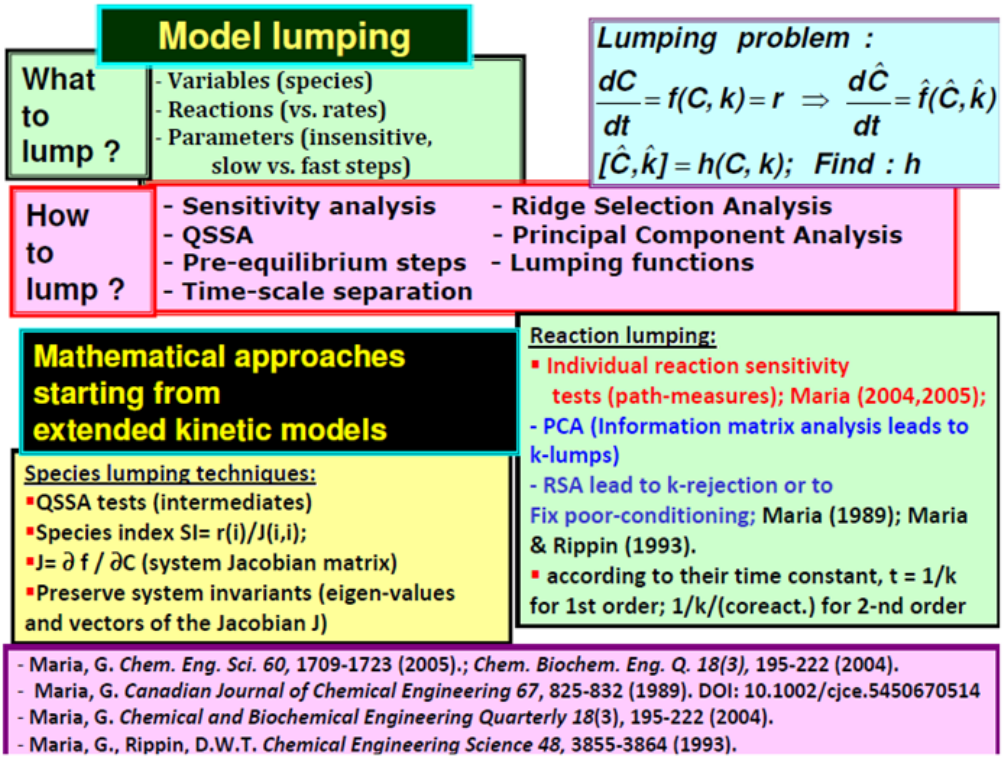


Figure 43: Some guide-lines used in obtaining reduced dynamic models (Maria (2004) [51]).

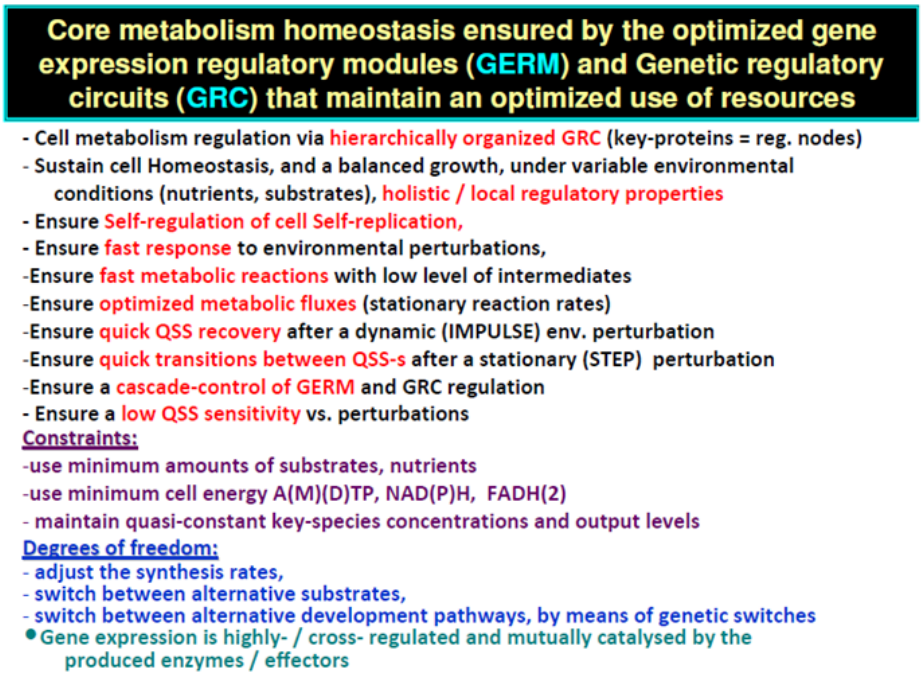


Figure 44: The role of gene expression regulatory modules (GERM) and of the genetic regulatory circuits (GRC) [1].

Classical approach to develop *deterministic* dynamic models is based on a hypothetical reaction mechanism, kinetic equations, and known stoichiometry. This route meets difficulties when the analysis is expanded to large-scale metabolic networks, because the necessary mechanistic details and standard kinetic data to derive the rate constants are difficult to be obtained. However, advances in genomics, transcriptomics, proteomics, and metabolomics, lead to a continuous expansion of bioinformatic databases, while advanced numerical techniques, non-conventional estimation procedures, and massive software platforms reported progresses in formulating such reliable cell models. Valuable structured dynamic models, based on cell biochemical mechanisms, have been developed for simulating various (sub)systems (see chap. 9).

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 [17,96]. Oriented and unified programming languages have been developed (see SBML, JWS, see chap. 4) to include the bio-system organization and complexity in integrated platforms for cellular system simulation (E-Cell, V-Cell, M-Cell, A-Cell, see chap. 4). Such integrated simulation platforms tend to use a large variety of biological databanks including enzymes, proteins and genes characteristics together with metabolic reactions [97,98].

Development of deterministic dynamic models to adequately reproduce such complex synthesis related to the central carbon metabolism [59,80,81,158,160]; but also the genetic regulatory system 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 [1,2,17,51].

By applying various modelling routes, successful *structured* (deterministic) dynamic models have been elaborated to simulate various regulatory mechanisms [17,100,104,106,107]; In fact, as mentioned by [99], 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 [51].

B. In the *Boolean* approach, variables can take only discrete values. Even if less realistic, such an approach is computationally tractable, involving networks of genes that are either "on" or "off" (e.g. a gene is either fully expressed or not expressed at all; Figure 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 process dynamics become available. 'Electronic circuits' structures (see an example in Figure 11 & 39; [47] 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.

C. *Stochastic* models replace the 'average' solution of continuous-variable ODE kinetics (e.g. species concentrations) by a detailed random-based simulator accounting for the exact number of molecules present in the system. Because the small number of molecules for a certain species is more sensitive to stochasticity of a metabolic process than the species present in larger amounts, simulation via continuous models sometimes can lack of enough accuracy for random process representation (as cell signalling, gene mutation, etc.). Monte Carlo simulators are used to predict individual species molecular interactions, while rate equations are replaced by individual reaction probabilities, and the model output is stochastic in nature. Even if the required computational effort is extremely high, stochastic representation is useful to simulate the cell system dynamics by accounting for a large number of species of which spatial location is important [100-102] (Figure 39).

D. Model *lumping*. 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 43).

Reduction in the model structure (via lumping of species, and reactions) is necessary due to [1,2]:

a. The high complexity of cell metabolic processes vs. available data

b. Large number of species, reactions, transport parameters, and interactions

c. low data observability & reproducibility

d. Metabolic process variability

e. Interpretable representation of cell complexity

f. Requirement to get quick simulations of cell behavior under various environmental conditions

g. Computational tractability and easier application of algorithmic rules from (bio)chemical engineering and numerical calculus.

However, a tradeoff between model complexity and adequacy must be maintained [46,51] to use such models for the *in-silico* design GMO, by *in-silico* re-programming the cell metabolism by using a gene knockout strategy [39,157,151,154, 203-206]; or by optimal cell cloning. Application of systematic math-lumping rules to metabolic processes must account for physical significance of lumps, species interactions, and must preserve the *systemic/ holistic* 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* of the ODE model; the ODE model Jacobian being defined as the derivatives of model functions in respect to model states, that is species concentrations the in cell metabolic models) has been proved to be insufficient [108].

This classic approach to develop dynamic reduced models is based on the biochemical reaction engineering rules, that is: propose a hypothetical reaction mechanism, formulate kinetic equations, and 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 are difficult to be obtained. However, advances in genomics, transcriptomics, proteomics, and metabolomics, lead to a continuous expansion of bioinformatic databases, while advanced numerical techniques, non-conventional estimation procedures, and massive software platforms, and lumping algorithms [51], reported progresses in formulating lumped (reduced) but reliable cell models. Valuable structured dynamic models, based on cell biochemical mechanisms, have been developed for simulating various (sub) systems (see examples of chap. 4; [17]).

However, here it is to mention, that the work with *reduced* kinetic models of cell metabolic syntheses and GRC-s, even if computationally very convenient, presents some inherent disadvantages, that is: multiple reduced model structures might exist difficult to be discriminated; 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); a loss in the model prediction capabilities; 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 [46,51,166] (Figure 43):

i. Reduce the list of reactions, by eliminating unimportant side-reactions and/or assuming quasi-equilibrium for some reaction steps; use sensitivity measures of rate constants to detect the redundant part of the model (e.g. ridge selection, principal component analysis, time-scale separation, etc.);

ii. Reduce 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 the target species "i" lifetime LTi = -1/J(i,i), and its production rate r(i), where the Jacobian elements 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 of the cell process, that is:

$$\frac{1}{V(t)}\frac{dn_j}{dt}=\sum_{i=1}^{nr}s_{ij}r_i(\boldsymbol{n}/V,\boldsymbol{k},t)=h_j(\boldsymbol{C},\boldsymbol{k},t)\ ,$$

$$\frac{d(n_j/V)}{dt}=\frac{dC_j}{dt}=\sum_{i=1}^{nr}s_{ij}r_i(\boldsymbol{n}/V,\boldsymbol{k},t)=h_j(\boldsymbol{C},\boldsymbol{k},t)\ \cdots\cdots \quad (1)$$

where notations are the followings: C(j) = (cell-)species j concentration; V = system (cell cytosol) volume; n(j) = species j number of moles; r(j) = j-th reaction rate; 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).

(iii) Decompose the kinetics into fast and slow 'parts' allowing application of the quasi-steady-state-approximation (QSSA) to reduce its dimensionality [51,166].

(iv) When the ODE kinetic model is linear in parameters, then the reduction procedure of Maria [51,166], can be successfully applied by preserving the system Jacobian invariants (eigenvalues, eigenvectors).

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 [46]. 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 metabolic circuits. Such a procedure is very suitable for modelling GRC-s by linking GERM models in such a way to maintain the cell homeostasis, that is to maintain relatively invariant species concentrations despite perturbations [1,166].

The math modelling efforts have intensified a lot after 2000 when the human genome has been deciphered (Figure 19), 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 12). Because many cell regulatory systems are organized as "modules" [114], it is natural to model GRC-s and other metabolic processes by using a *modular* approach [1]. Further analyses including engineered GRC-s can lead to simulate and design of GMO, of desirable characteristics, that is [115]: 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 GRC in GMO one may create potent applications in industrial, environmental, and medical fields (e.g. biosensors, gene therapy). Valuable implementation tools of the design GRC in real cells have been reported over the last years [61,68,113].

Systems Biology using *in-silico* design of GMO is closely related to the *Synthetic Biology*. The emergent field of *Synthetic Biology* [62] "interpreted as the engineering-driven building of increasingly complex biological entities" [61], 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 [61,63]. 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 [1-2]. 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 [61].

To help the efforts of the *Synthetic Biology* to *in-silico* design GMO 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 concepts and tools, together with the control theory of the nonlinear systems to characterize the kinetics and self-regulation of the cell metabolic processes.

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 *in-silico* 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.

9.Deterministic Continuous Variable Math Models in Molecular Biology are Constructed By Using the (Bio)Chemical Engineering Principles and Tools

Metabolic processes at a low(molecular)-level are generally better clarified. Based on that, conventional dynamic models, based on ordinary differential (ODE) species mass balance, 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, 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 [160], and to reflect the species interconnectivity or perturbation effects on cell growth [1,163]. Mixtures of ODE kinetic models with discrete states (i.e. 'continuous logical' models), and of continuous ODE kinetics with stochastic terms can lead to promising mixed models able to simulate both deterministic and non-deterministic cell processes [48]. Representation of metabolic process kinetics is made usually by using rate expressions of extended Michaelis-Menten or Hill type [103,158,160] (Figure 47).

However, the main advantages of deterministic/continuous variable models are coming from the use of experience, concepts, math representation, rules, and algorithms of the biochemical reaction engineering. The reaction rate expressions in the deterministic models are the usual ones of biochemical reaction engineering, that is of Michaelis-Menten, or Hill type (Figure 47) (e.g. see dynamic models of CCM, [59,80,81].

The parameters (rate constants) of deterministic models are estimated using the common (bio)chemical engineering rules [51], by using either dynamic (kinetic) data obtained in a chemostat under transient regime (e.g. pulse-like perturbations in the bioreactor influent [59]), or using steady-state data (metabolites concentrations) obtained at homeostasis (i.e. balanced cell growth [1,2]), by solving the model stationary algebraic set dCj/dt= 0 (where "j" indexes the no. of cell species taken individually or lumped). Parameter estimate must fulfill physico-chemical meaning constraints (related to metabolic reaction stoichiometry; rate constants must be limited by the diffusional processes, and in agreement with the thermodynamic equilibrium steps). Additionally, due to the optimized metabolic cell process, GERM, GRC, and cell models must fulfil some optimality constraints, that is: reaction rates must be maximal, but with rate constants limited by the diffusional processes; the total enzyme (proteine) content of the cell is limited by the isotonicity condition; also the total cell energy (ATP) and reducing agent (NADH) resources are limited; the reaction intermediate level must be minimum; the cell model at homeostasis must be stable, that is will reaching again the steady-state after termination of a perturbation [6], key-species concentration must be constant at homeostasis. Most of the mentioned aspects will be below discussed and exemplified by Maria [1].

Apart from the common estimation criteria used in kinetic model estimation [51], supplementary optimization rules must be applied to determine some rate constants, by imposing optimum regulatory criteria for GERM-s, such as minimum recovering time of the stationary concentrations (homeostasis) after a dynamic ('impulse'-like) perturbation in a key-species [17], by using effective solvers [51], or a quick action of the buffering reactions to get the fastest recovering time of the steady-state (homeostasis) after an external / internal perturbation; smallest sensitivity of the key-species homeostatic levels vs. external perturbations in the nutrient levels; stability highest homeostasis strength, etc. [1,52,17,46,165,152,151,154,158,159]. Estimation rule is based on the fulfilment of the stationary condition imposed to the cell model (dCj/dt = 0), by preserving the system invariants { of the mass balance equations, [51,166], and by imposing optimum regulatory criteria related to GRC efficiency to cope with continuous perturbations, formulated by Maria [1]. In the GRC model cases, also supplementary constraints have to be considered related to the VVWC modelling approach (isotonicity constraints, see below), and constraints related to the GERM, GRC stationary/dynamic regulatory efficiency [1].

Even if complicated and, often over-parameterized, the continuous variable dynamic deterministic ODE models of metabolic pathways and 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/ando-geneous continuous perturbations [17,94]. 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 [6,28], accounting for the regulatory system properties, that is: dynamics, feedback / feedforward, and optimality. And, most important, such ODE kinetic modelling approach allows using the strong tools of the classical (bio)chemical engineering modelling, that is [3] (Figure 31 & 32):

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), [109].

IV. set equilibrium reactions; Gibbs free energy balance analysisset cyclic reactions; find species at quasi-steady-state; improved evaluation of steady-state flux distributions that provide important information for metabolic engineering [110], allowing application of ODE model species and/or reaction lumping rules [51].

The ODE deterministic models have been developed in two alternatives [1,2]:

9.1. Constant volume whole-cell (cvwc) dynamic models

The default Constant Volume Whole-Cell (CVWC) classical continuous variable ODE dynamic models, does not explicitly consider the cell volume exponential increase during the cell growth. When the continuous variable CVWC dynamic models are used to model the cell enzymatic processes, the default-modelling framework is that of a constant volume and, implicitly of a constant osmotic pressure, eventually accounting 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 CVWC formulation results from the species concentration definition of Cj= nj/V, leading to the default kinetic model eq.(1):

The formulation (1) 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 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 [28].

Such a CVWC 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 [88], by distorting very much or even misrepresenting the prediction results, as exemplified by [1,53] (Figure 56) for both stationary or perturbed cell growing conditions.

9.2.Variable Volume Whole-Cell (VVWC) dynamic models

As an alternative, [1-2,17,49,46,51,53-57,151,152,154,158] strongly promoted the holistic *variable-volume whole-cell* (VVWC) modelling framework by explicitly including in the model constraint equations accounting for the cell-volume growth and by preserving the same cell-osmotic pressure. Maria [1,53] also proved the advantages of such a model formulation.

While the past and current CCWC cell dynamic models ensure 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., [84], the VVWC formulation promoted by [1-2] includes thermodynamic isotonicity relationships/ constraints. Thus, it has been proved, step-by-step, 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 [17,46,49,53-57,151,152,154,158].

In the VVWC approach, the continuous ODE model is re-written either in terms of species moles or of species concentrations, as following [1] (Figure 48).

$$\frac{dC_j}{dt}=\frac{1}{V}\frac{dn_j}{dt}-DC_j\ ;\ \frac{1}{V}\frac{dn_j}{dt}=r_j\ ;$$

(j=1,..,no. of species),

Where:

$$D=d(In(V))/dt\ ,\quad \text{..... (2)}$$

where: V=cell volume (in fact cytosol volume); nj= species j number of moles; rj= j-th reaction rate; D= cell-content dilution rate, i.e. cell-volume logarithmic growing rate; species inside the cell are considered individually or lumped; t=time. The mass-balance formulation (2) is that given by Aris [111] for the (bio)chemical reacting systems of variable-volume.

In the VVWC formulation of the cell dynamic cell model an additional constraint must be also considered to preserve the system isotonicity (constancy of the osmotic pressure π) under isothermal conditions. This constraint should be considered together with the ODE model (2), that is the Pfeiffers'law of diluted solutions [112] adopted and promoted by Maria [17], etc., review [1]):

$$V(t)=\frac{RT}{\pi}\sum_{j=1}^{ns} n_j(t) \quad \text{.....} \tag{3}$$

Which, by derivation and division with V leads to (Maria, 2017[1], 2017[55]):

$$D=\frac{1}{V}\frac{dV}{dt}=\left(\frac{RT}{\pi}\right)\sum_{j}^{ns}\left(\frac{1}{V}\frac{dn_j}{dt}\right) \quad \text{.....} \tag{4}$$

In the above relationships, T =absolute temperature, and R =universal gas constant, V =cell (cytosol) volume. As revealed by the Pfeffer's law eqn. (3) in diluted solutions [112], and by the eq. (4), the volume dynamics is directly linked to the molecular species dynamics under isotonic and isothermal conditions. Consequently, the cell dilution D results as a sum of reacting rates of all cell species (individual or lumped). The (RT/ pi) term can be easily deducted in an isotonic cell system, from the fulfilment of the following invariance relationship derived from (3):

$$V(t)=\frac{RT}{\pi}\sum_{j=1}^{ns} n_j(t)\Rightarrow\frac{RT}{\pi}=\frac{V(t)}{\sum_{j=1}^{ns} n_j(t)}=\frac{1}{\sum_{j=1}^{ns} C_j}=\frac{1}{\sum_{j=1}^{ns} C_{jo}}=constant \tag{5}$$

As another observation, from (4) it results that the cell dilution is a complex function D(C,k) being characteristic to each cell and its environmental conditions.

Relationships (4-5) are important constraints imposed to the VVWC cell model (2), eventually leading to different simulation results compared to the CVWC cell kinetic models that neglect the cell volume growth and isotonic effects [1,53]. On the contrary, application of the default classical CVWC ODE kinetic models of eqn. (1) type with neglecting the isotonicity constraints presents a large number of inconveniences, related to ignoring lots of cell properties [1,53],

That is:

I. The influence of the cell ballast in smoothing the homeostasis perturbations;

II. The secondary perturbations transmitted via cell volume following a primary perturbation;

III. The more realistic evaluation of GERM regulatory performance indices (P.I.-s),

IV. The more realistic evaluation of the recovering/transient times after perturbations;

V. Loss of the intrinsic model stability;

VI. Loss of the self-regulatory properties after a dynamic perturbation, etc.

By contrast, the VVWC novel modelling concept/framework promoted by [17], etc., review [1] to derive cell kinetic models, in a holistic approach, ensures cell processes homeostasis, and 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 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. The novel VVWC modelling framework is leading to accurately simulate lot of cell metabolic effects, such as:

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

b. The secondary perturbations transmitted 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 via defined performance indices (P.I.-s).

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

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

g. Allows studying conditions when the system homeostasis intrinsic stability is lost;

h. Allows studying the self-regulatory properties after a dynamic/stationary perturbation, etc.

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

j. allows studying the relationships between the external conditions, species net synthesis reactions, and the cell osmotic pressure,

By using the VVWC modelling framework, dr. Maria (see chap. 9.2) developed structured reduced dynamic models of various complexity to simulate individual GERM-s, but also linked GERM-s in GRC of modulated functions (e.g. toggle-switch, amplitude filters, modified operons, etc.), used to design GMO of practical interest;

The basic equations and hypotheses of a VVWC model are the followings [1,17], (Figure 48):

(i) 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); the cell is separated from the *environment* (containing *nutrients*) by a *membrane*.

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

(iii) The cell is an isothermal system with an uniform content (perfectly-mixed case); species behave ideally, and present uniform concentrations within cell. The cell system is not only homogeneous but also isotonic (constant osmotic pressure), with no inner gradients or species diffusion resistance.

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

(v) To better reproduce the GERM properties interconnected with the rest of the cell, the other cell species are lumped together in the so-called "cell ballast" [1,55].

(vi) The inner osmotic pressure (π cyt) is constant, and all time equal with the environmental pressure, thus ensuring the membrane integrity (πcyt = π env = constant). As a consequence, the isotonic osmolarity under isothermal conditions leads to the equality RT/ π cyt= RT/πenv. Otherwise, the osmosis will eventually lead to an equal osmotic pressure (π cyt = πenv). 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 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 [V=Vo exp(+D•t)] [1,53,55].

(vii) The cell content reports a continuous dilution, that is a species concentration decline due to the continuous increase of the denominator of the expression Cj= nj(t)/V(t). In spite of that, concentrations of key species remain constant because the numerator (copynumbers) increases at the same rate with the denominator. So, the overall concentration of cellular components is time-invariant (homeostasis) [Figure 41].

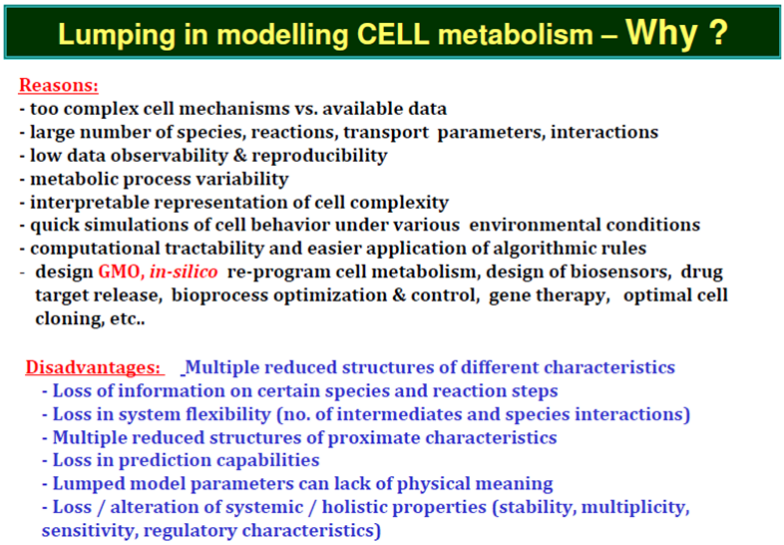


Figure 41: Reasons to use lumped models in description of cell metabolic processes.

(viii) Species concentrations at the cell level are usually expressed in nano-moles, being computed with the relationship [17].

(ix) Cell volume doubles over the cell cycle period (tc), with an average logarithmic growing rate of D = ln(2)/tc [resulted from integrating the definition, D= d(ln(V))/dt, in eq. (4)]. Under stationary growing conditions, that is a constant D over the cell cycle, integration of this relationship indicates an exponential increase of the cell volume, that is [V = Vo exp(+D•t)].

(x) Under stationary growing conditions, species synthesis rates (rj) must equal to first-order dilution rates (DsCjs), leading to time-invariant (index "s") species concentrations Cjs, i.e. the homeostatic conditions (corresponding to a balanced steady-state growth). Under QSS cell growing conditions, the ODE model mass balance eq. (2) is leading to the following nonlinear algebraic mass balance set:

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

$$j = 1,\ldots, n_s \text{ (no. of species); where } D_s = \left(\frac{RT}{\pi}\right)\sum_j^{n_s}\left(\frac{1}{V}\frac{dn_j}{dt}\right)_s \qquad (6)$$

(xi) This QSS mass balance eq.(6) is used to estimate the rate constants k from the known stationary species concentration vector Cs, with also imposing some optimal properties of the cell system [1-2,17,46,49,53,54,56,57,151-154,158].

(xii) It is to observe that, in a continuous variable metabolic kinetic model, species concentrations can present fractional values. When treated deterministically, fractional copy numbers must be loosely interpreted either as time-invariant averages in a population of cells or as a time-dependent average of single cells. For other types of cell kinetic models see the review of [17].

(xiii) A metabolic kinetic model in a VVWC approach should be written in the form (2-5). In such a formulation, all cell species should be considered (individually or lumped), because all species net reaction rates contribute to the cell volume increase (eq. 3). As the cell volume is doubling during the cell cycle, this continuous volume variation cannot be neglected.

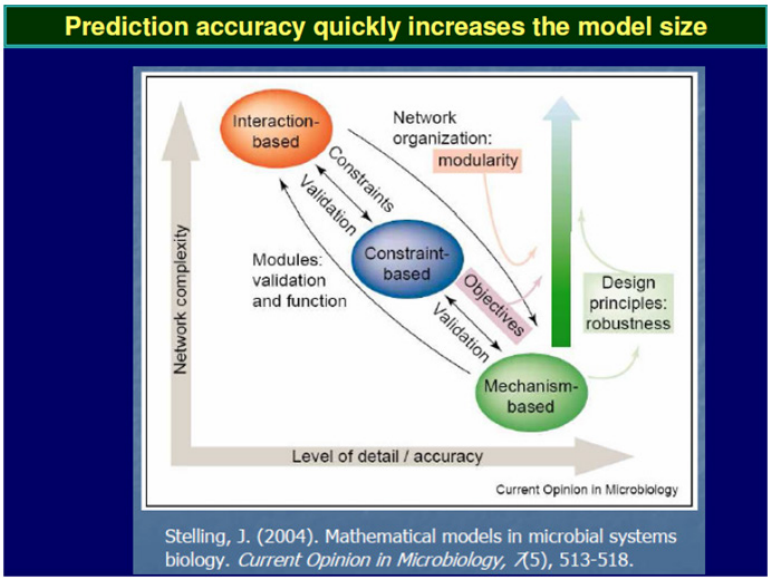


Figure 42: The exponential-like increase of the model size with the prediction accuracy (Stelling (2004) [50]).

However, to not complicate too much the VVWC dynamic model (Figure 42), usually a reduction in the number of cell species and reactions by common lumping rules of (bio)chemical engineering [46,51,166] is usually performed. Such a model reduction strategy of the metabolic kinetic models present a series of disadvantages, such as: a loss in model adequacy; in the simulated system flexibility (due to the reduced number of considered intermediates and species interactions); an increased possibility to get multiple (rival) reduced models of proximate characteristics for the same cell system, difficult to be delimited; a loss 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). In spite of that, lumped deterministic models can successfully simulate a brad range of metabolic processes [2].

A comparison of model prediction quality in the case of a GERM of [G(P)1] type modelled under CVWC or VVWC, clearly indicate that CVWC can lead to biased and distorted conclusions on GERM regulatory performances (under both stationary or as response to dynamic pulse-like perturbations), thus making difficult the modular construction of GRC-s by linking GERM-s (Figure 55-56) [53].

10. Modelling GERM and GRC

When developing deterministic models for the CCM, or for other cell metabolic processes, to be further used for GMO design, an important aspect is to also include math models of individual GERM characterizing the gene expression control of the enzymes production. Also, by linking the interfering GERM modules, complex GRC regulatory chainscan thus be obtained [1,87].

Because the GRC-s are responsible for the control of the cell metabolism, the adequate kinetic modelling of the constitutive GERM-s, 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 GRC-s (where key-proteins play the role of regulatory nodes). Eventually, such models allow simulating the metabolism of modified cells.

Various simple math models have been proposed to represent the elementary metabolic fluxes (Klamt & Stelling [181]), or GRC (Hecker et al. [182]; Wang et al. [183]; Zhdanov [184]; Atkinson et al. [117]; Wall et al. [138]; Hlavacek & Savageau [139]; Kaznessis [115]; Styczynski & Stephanopoulos, [94]; Savageau [90]; Rosenfeld et al. [185]; De Jong [186], [187]; Endy & Brent [188]; Guantes & Poyatos [189]; Hasty et al. [190-192]; Isaacs et al. [193]; Kærn et al. [194]; Klipp [195]; Mochizuki [196]; Samad et al. [197]; Tyson et al. [198]; Widder et al. [199]; Liu et al. [200]; Kaznessis [201]). Eventually, such models allow a multi-criterion design and optimization of a target GRC (Van Someren et al. [202]).

Development of dynamic models to adequately reproduce such complex synthesis related to the CCM [59,80,81], 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 [46,51,166]. However, such structured models are extremely useful for *in-silico* design of novel GRC-s conferring new properties/functions to the mutant cells, in response to external stimuli [8,61,115-122,157,203-206].

A central part of cell metabolic models concerns self-regulation of the metabolic processes via GRC-s. Consequently, one particular application of such dynamic cell models is the study of GRC-s, 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) and a large number of examples have been reported with *in-silico* re-creation of GRC-s 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 GRC-s into organisms, one may create a large variety of mini-functions / tasks (or desired 'motifs') in response to external stimuli.

"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 [73,123,124]. 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 metabolic control analysis or a dynamic sensitivity analysis" [125]. 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 [8,61,115,116,121,122,151,154,157,203-205].

Examples of such GRC modulated functions include:

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

b. 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;

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

d. Specific treatment of external signals by controlled expression such as amplitude filters, noise filters or signal/ stimuli amplifiers;

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

10.1. Modelling individual GERMs

As experimentally proved in the literature [1], the genetic regulatory systems (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 GERM-s. However, it is wellknown that one GERM interacts with no more than other 23-25 GERM-s [7], while most of GERM structures are repeatable. Consequently, in developing the GRC analysis, the modular approach is preferred due to several advantages: a separate analysis of the constitutive GERM-s in conditions that mimic the stationary or perturbed cell growth. The modules are then linked to construct GRC of an optimized regulatory efficiency that ensures key-species homeostasis and cell network holistic properties (Figure

44,46,50). Investigation of GERM and GRC characteristics is focus on the tight control of gene expression, the quick dynamic response, 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 stationary (step-like) perturbation of one of the involved metabolites or nutrients [1](Figure 54,57).

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 45,49. 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 45-down-right, and Figure 49, the row-up), refers to the synthesis of a generic protein P and the simultaneous replication of its encoding gene G. The lumped G(P)1 model includes only one regulatory element (a so-called "effector", that is a fast "buffer" reversible reaction G + P <===> GP(inactive) (Figure 49), 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, 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.

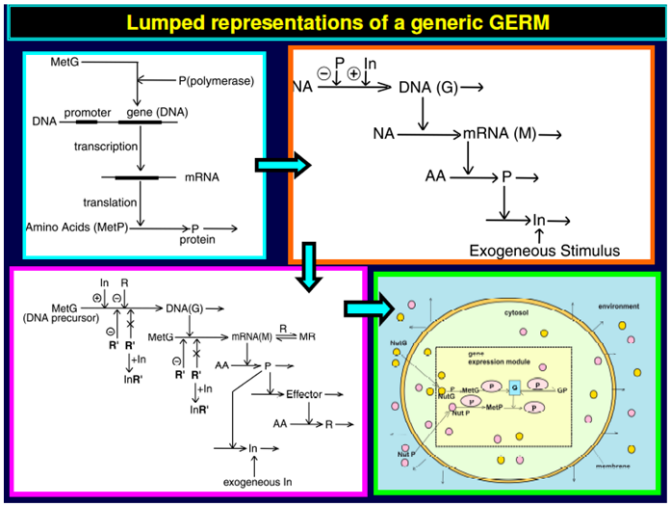


Figure 45: Various simplified representations of a gene expression regulatory modules (GERM) [17]. Down-right. Simplified reaction scheme of a generic gene G expression, by using a regulatory module of G(P)1 type. The model was used to exemplify the synthesis of a generic P protein in the E. coli cell by Maria [17]. 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.

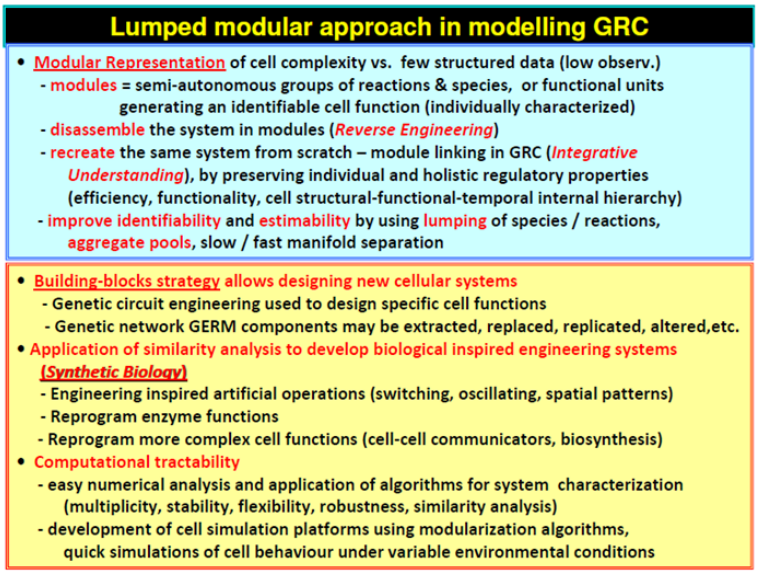


Figure 46: Importance of the lumped modular math modelling for the in-silico design of GMO [1].

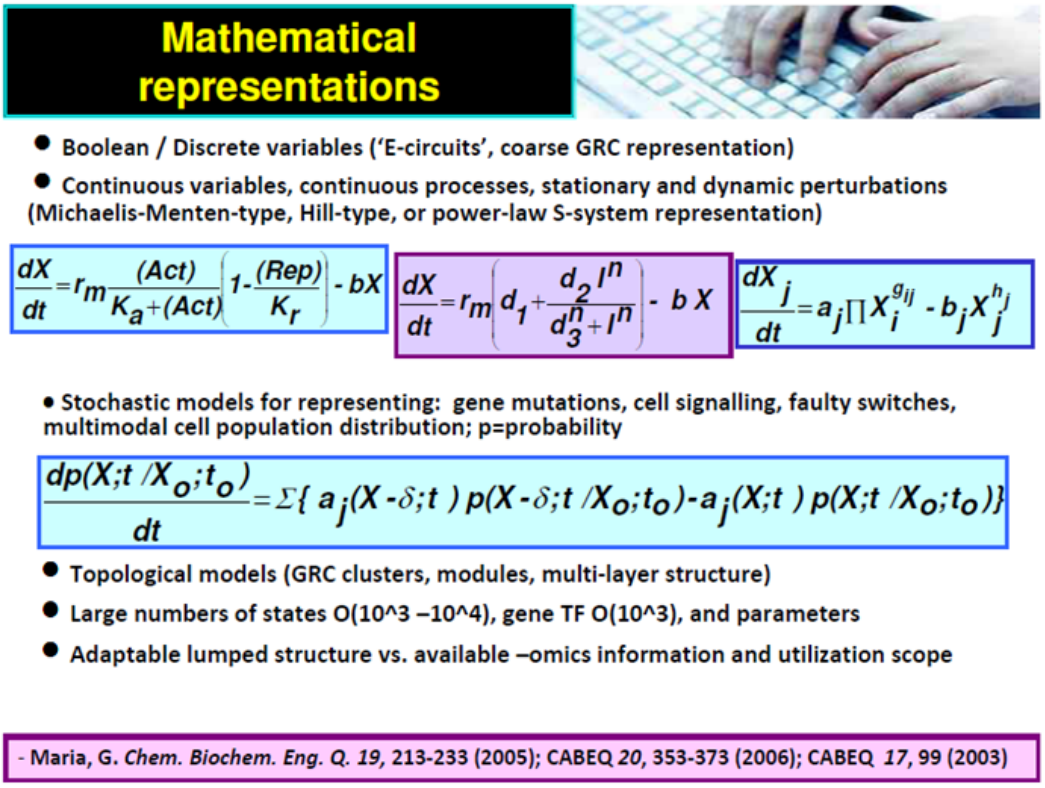


Figure 47. Some math representations of genetic regulatory circuits (GRC) (GERM) [1,17]

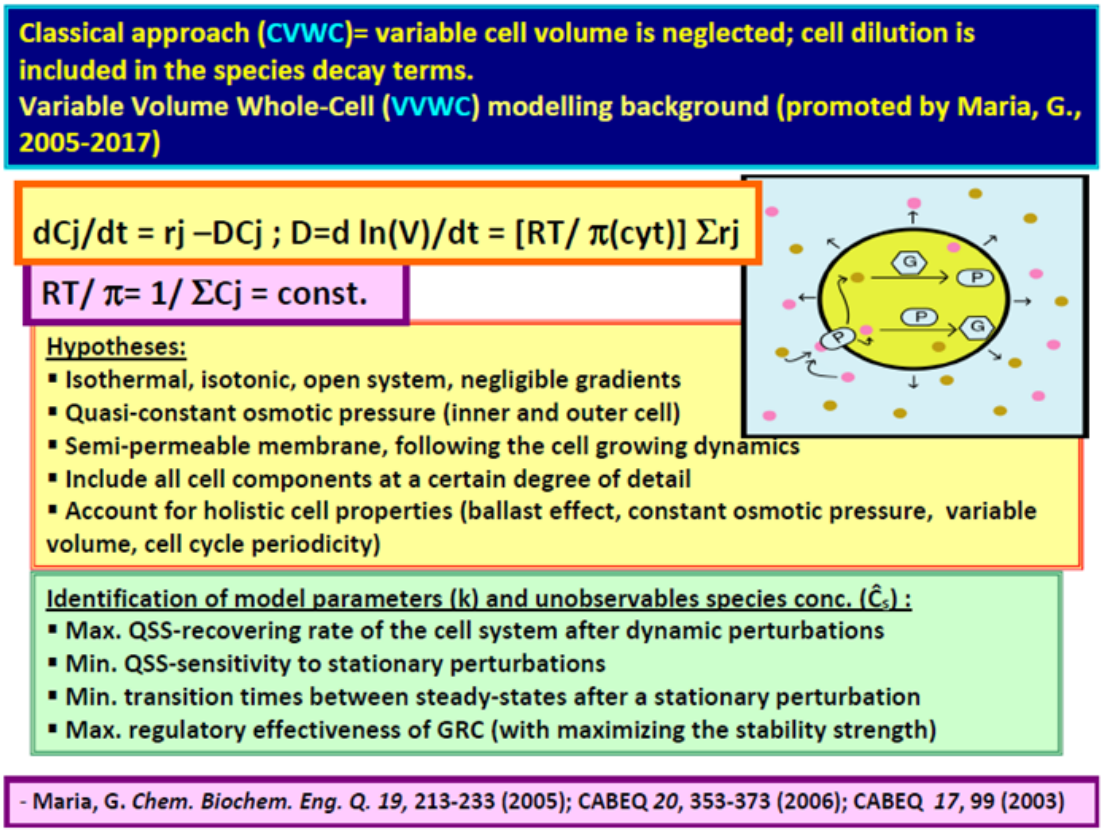


Figure 48: The novel modelling framework of VVWC promoted by Maria [1,17].

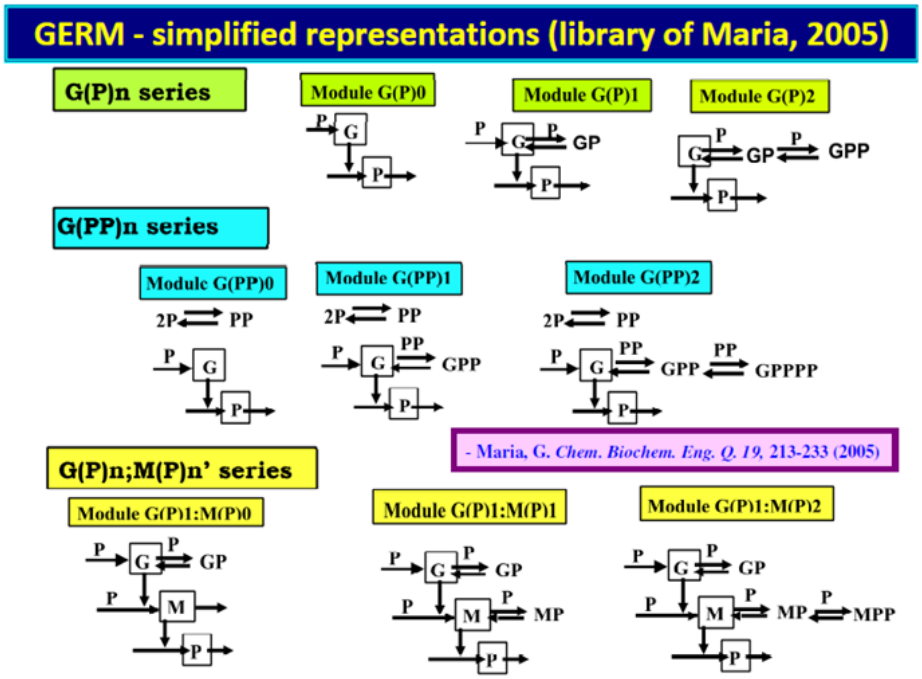


Figure 49: The library of [1,17] with lumped modular models of GERM-s. Simplified representations of a generic gene expression G/P regulatory module (GERM) following [17]. The horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; absence of a substrate or product indicates an assumed concentration invariance of these species; G= gene encoding P; M= mRNA. Up-row: simplified representation of the gene expression model 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 of perturbations in nutrients Nut*, and metabolites Met*, or of internal cell changes, a very rapid buffering reaction G + P<===> GP(inactive) has been added. 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. 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.

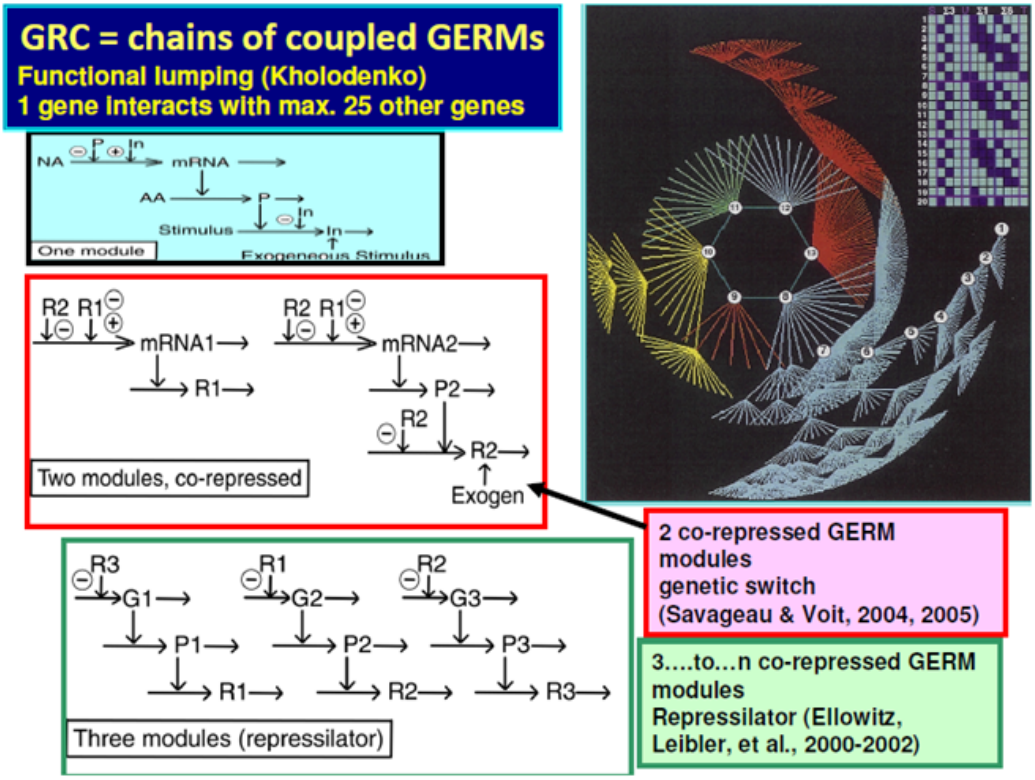


Figure 50: Example of linked GERM-s to form GRC-s(review of [1]).

Notations: G= DNA gene encoding P; M= mRNA; P,PP= allosteric effectors of the transcription/translation.

The module nomenclature used in Figure 49 for such GERM models, proposed by Maria [17,127] is those of : [L1(O1)n1 ; ....; Li(Oi)ni]. 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 or even genes G, P, M, etc.) at which regulatory element acts, and n(i) =0,1,2,... number of 'effector'/TF species O(i) (i.e. 'effectors' P, PP, PPPP, R, RR, RRRR, etc) binding the 'catalyst' L. For instance, a [G(P)2] unit of Figure 49 includes two successive binding steps of G with the product P, that is G + P <===> GP + P <===> GPP, all intermediate species GP, GPP, being inactive catalytically, while the mass conservation law is all time fulfilled, i.e. $\sum_{i=0}^{2}[G(P_i)] = \text{constant.}$ Such a representation accounts for the protein concentration diminishment due to the cell-growth dilution effect, but could also include protein degradation by proteolysis. It is also to observe that such GERM models try to account 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 49 for the 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 G + P <===> GP).

As proved in previous works [1], the performances indices (P.I. see below) of GERM of [G(P)n] type in Figure 49, are as better as the number "n" of buffer reactions increases (Figure 53). Also, [1] proved that when P is acting as a TF, its efficiency is better if it is present in a dimeric form (PP), in GERM-s of [G(PP)n] type in Figure 49 & 53. [1] 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 down-rows of Figure 49), 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. It clearly appears (Figure 54) that, as the number of effectors increases in GERM-s as their P.I. 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-protein stationary level [P]s. These processes occur in addition to those of the ''core'' system (G/P replication over the cell cycle).

GERM regulatory performance indices P.I. are of two types (Maria, 2017[1]) (Figure 51 & 52): stationary and dynamic. Briefly they are presented in the Table 1, together with the associated optimization objective, for a general nonlinear dynamic cell model described by eq. (7). Detailed information is given by Maria [1].

$$dC/dt = h(C,k)\ ;\ C(0)=C_s;\ dA/dt = J_C A\ ;\ A(0)=I;\ J_C=(\partial h(C,k)/\partial C)_s \quad (7)$$

GERM or GRC regulatory performance indices translated in mathematical terms

- Performance indices (QSS= quasi-steady-state):
  - SELECTIVITY: to inner/external stimuli
  - SENSITIVITY: QSS levels vs. endo-/exogeneous stimuli
  - ROBUSTNESS: small sensitivity (States vs. Parameters)
  - EFFICIENCY: large margins of QSS stability vs. small changes, quick recovering of QSS
  - RESPONSIVENESS: small transition time to the new QSS after a step perturbation
  - STABILITY STRENGTH: high QSS recovery rates (RD) after an impulse perturbation;
  - SPECIES CONNECTIVITY during transitions: small St.Dev. of recovering / transition times
- Apply unconventional estimators:
  - Optimise individual module properties
  - Reproduce GRC holistic properties
- Impose physical-biological constraints:
  - Maximum regulatory efficiency
  - Structural, functional, temporal cell hierarchy (systemic properties)
  - Key species homeostasis
  - Cell system flexibility vs. environmental changes
  - Holistic properties of GRC (gene connectivity and sharing functionality, minim basic-levels of effectors / intermediates, system robustness)

Maria, G *A review of some novel concepts applied to modular modelling of genetic regulatory circuits* , Juniper publ., Newbury Park, CA, 2017, ISBN 978-1-946628-07-7(USA). https://juniperpublishers.com/ebook-info.php

- Maria, G. *Chem. Biochem. Eng. Q. 19,* 213-233 (2005); CABEQ *20,* 353-373 (2006); CABEQ *17,* 99 (2003)

Figure 51: Seeking for modelling the GRC regulatory properties (performance indices, [1]).

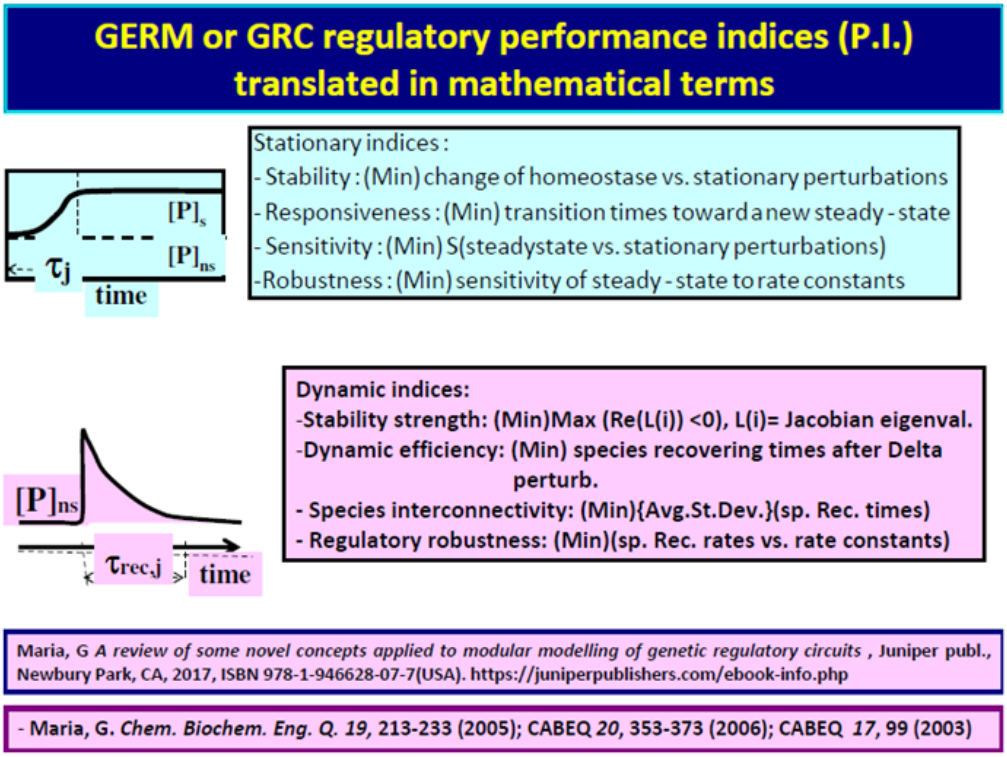


Figure 52: Define some of the GERM regulatory performance indices [1].

By summarizing, the regulatory efficiency P.I.-s proposed to evaluate the perturbation treatment by a GERM or by a chain of GERM-s, are given in the Table 1 (Figure 51).

Table 1: The regulatory efficiency performance indices P.I.-s proposed to evaluate the perturbation treatment efficiency by a GERM following the definitions of (Maria, 2005[17]). Abbreviations: Min = to be minimized; Max = to be maximized. Note: k(syn) and k(decline) refers to the --> P --> overall reaction. Notations used in the Table 1: "n"= nominal value; "s" = stationary value; (*) see eq. (7) and (Maria, 2003[52]) for the monodromy matrix A matrix calculation; Lambda(i) = i-th eigen-value of the JC Jacobian matrix. $J_C=(\partial h(C,k)/\partial C)_s$ defined by equation (7); A Defined by equation (7); A = Monodromy Matrix; tau _(j) = species "j" recovering time; λ_{A_i} = i-th eigen-value of the A matrix; 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 key-protein expressed in the analysed GERM.

Index	Goal	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	Min	$S^i_{NutP_j} = [(\partial C_i / C_{is}) / (\partial C_{Nut_j} / C_{Nut_{js}})]_s$
Stationary Regulation	Min	$S^i_{k_j} = [(\partial C_i / C_{is}) / (\partial k_j / k_j)]_s$
Dynamic Regulation	Min	$R_D = Max(\mathrm{Re}(\lambda_i)); \mathrm{Re}(\lambda_i) < 0$
Dynamic Regulation	Min	$\tau_j; \tau_p$
Regulatory Robustness	Min	$(\partial R_D / \partial_k)$
Species Interconnectivity	Min	$AVG(\tau_j) = average(\tau_j)$

Species Interconnectivity	Min	$STD(\tau_j) = st.dev.(\tau_j)$
QSS Stability	Min	$\mathrm{Re}(\lambda_i) < 0;$ for all i
QSS Stability Strength	Min	$Max(\mathrm{Re}(\lambda_i))$
QSS Stability Strength	Min	$\left\|\lambda_{A_i}\right\| < 1$

A) *Stationary P.I.* are defined as response to a stationary perturbation (Figure 52), that is transition from a QSS to another QSS following a step-like perturbation of one cell component concentration.

Stationary perturbations refer to permanent modifications in the levels of the external nutrients or of the internal metabolites, leading to new stationary component concentrations inside the cell. Referring to a target protein P in a GERM, 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 52). Equivalently, the P-synthesis regulatory module will tend to maintain [P]ns within certain limits, [P]min ≤ [P]ns ≤ [P]max (a relative Rss =±10% maximum deviation has been proposed by Sewell et al. [126], & Yang et al. [127], see Table 1 to get an effective GERM. A measure of the species "i" steady-state concentration (Ci,s) 'resistance' to 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,s; [Nut,j]) respectively, where S(state; perturb.)= ∂(state)/ ∂(perturbation), are the state sensitivities vs. perturbations; (Varma et al., 1999 [128]).

Other stationary P.I. are discussed by Maria [1], that is (Figure 52):

I. *Transition time* necessary to each GERM component to return to their stationary concentration (QSS) after a step-like perturbation in one component;

II. *Responsiveness* to exo/endogeneous signaling species of the analysed GERM or GRC can be represented by the small transient times necessary for a species "j" QSS-level to reach a new QSS (with a certain tolerance) after applying a stationary external stimulus [159]. Consequently, the P.I. measure of the GERM efficiency to move fast to a new QSS is given by the duration of the *transition time* (in the case of P-species in Figure 52) necessary to a certain component to reach the new steady-state concentration. Another regulatory P.I., that is $A_{unsync} = k_{syn} \times k_{decline}$ (Table 1), has been introduced to illustrate the maximum levels of (unsynchronized) stationary perturbations in synthesis or consumption rates of a key-species 'tolerated' by the cell within defined limits (Sauro and Kholodenko, 2004[60]). For instance, in the case of the P-species these rate constants belong to the synthesis and degradation lumped reactions $\xrightarrow{k_{syn}} P \xrightarrow{k_{\deg}}$.

III. *Stationary efficiency*: Another stationary P.I. is related to the small sensitivities S(Ci; Nut,j) of the key-species levels Ci vs. changes in the external nutrient levels Nut(j). These sensitivities are computed from solving a sensitivity nonlinear algebraic set obtained by assuming QSS conditions of the ODE kinetic model eq. (2), or (6), and known nominal species stationary concentrations Cs. Then, differentiation of the steady-state conditions eqn. (6) leads to the evaluation of the state sensitivity vs. nutrient levels, i.e. SCi; Nutj))= (∂(Ci / ∂Nut(j)), [1].

IV. *The steady-state Cs stability strength*: This GERM property is related to the strong capacity of the regulatory system to 'resist' to large external/internal perturbations, thus maintaining the system steady-state Cs and determining very quick recovering paths. As with all other P.I.-s, this GERM property is related to the GERM system characteristics. Basically, as Max(Re(λ (i)))<0 is smaller as this QSS is more stable. Here, the eigenvalues λ (i) of the Jacobian matrix J_C= ∂h(C,k) / ∂C, [1] are evaluated at a checked QSS(Cs).

V. In a more systematic approach, the steady-state Cs *stability strength*. can also be associated to an index against periodic oscillations of key-species synthesis. This index can be evaluated from the linearized form of the system model, by calculating the monodromy matrix A(T) after a checked period T of time (Maria, 2003[52]), by using eq. (7). For a stable Cs, i.e. $|\lambda_{A_i}| < 1$, as $|\lambda_{A_i}|$ are smaller, as the stability of the Cs state is stronger and that **QSS** recovers faster after a small dynamic perturbation. Here, λ_{A_i} denotes the eigenvalues of the *A(T)* matrix, while I = identity matrix. In other words, **QSS** stability strength involves: Min(MAX(Re(λ(i))), with Re(λ(i))<0 for all "i", and (Min)$|\lambda_{A_i}| < 1$.

B) *Dynamic P.I.* are defined are defined as response to a dynamic perturbation (Figure 52), that is the recover of the QSS following an impulse-like perturbation in one cell component concentration. 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 system recovers and returns to its stable nominal state QSS (see Figure 52,57 for a generic P-protein case). The computed recovering time τ (rec,j) necessary to each component "j" to return to their stationary concentration (with a tolerance of 1-5% proposed by Maria 17]) may differ from one species to another depending on how effective are their corresponding regulatory circuits.

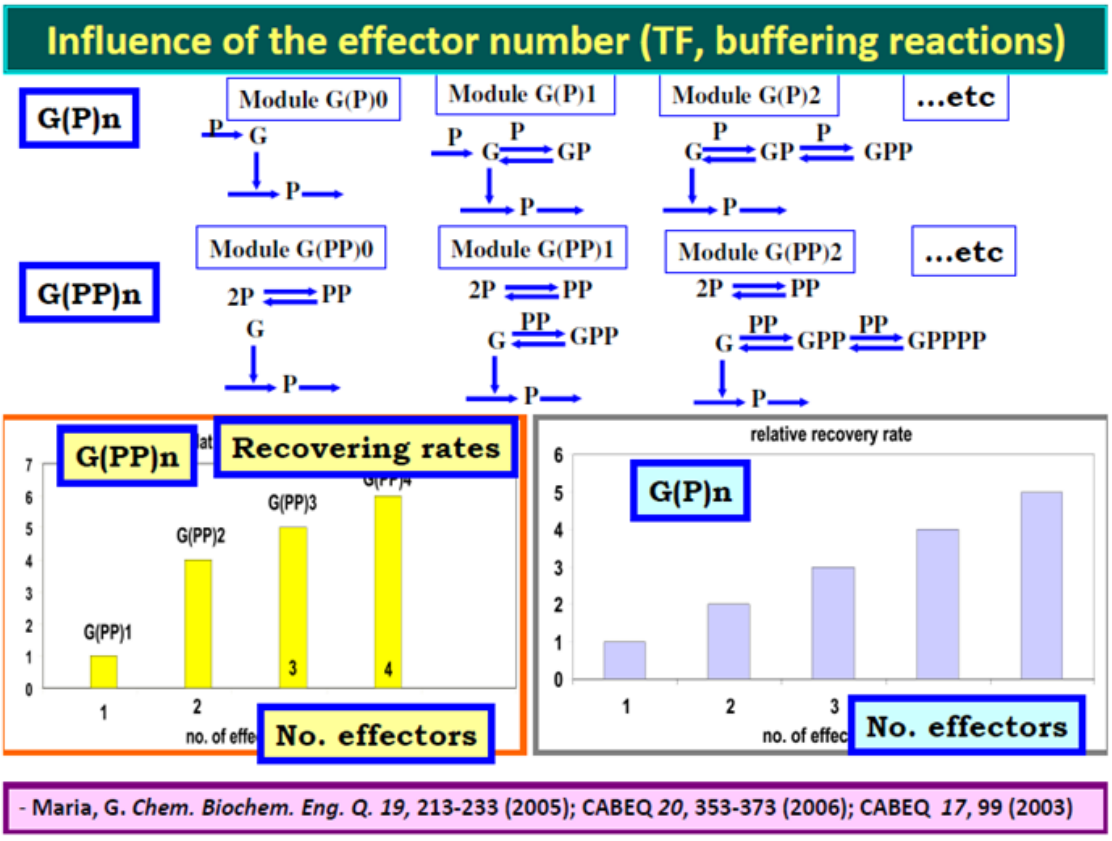


Figure 53: Influence of the number of effectors in a GERM on their regulatory performance indices [1,2].

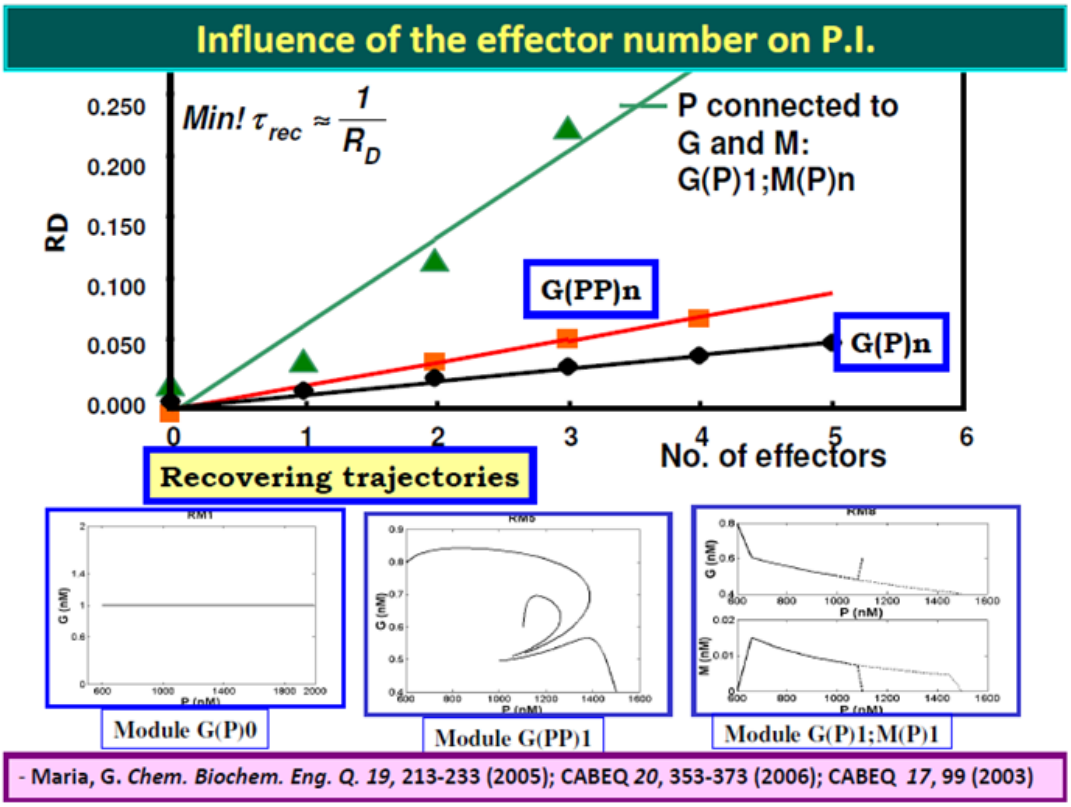


Figure 54: Influence of the GERM number of effectors on their properties [1,2,46,17,52].

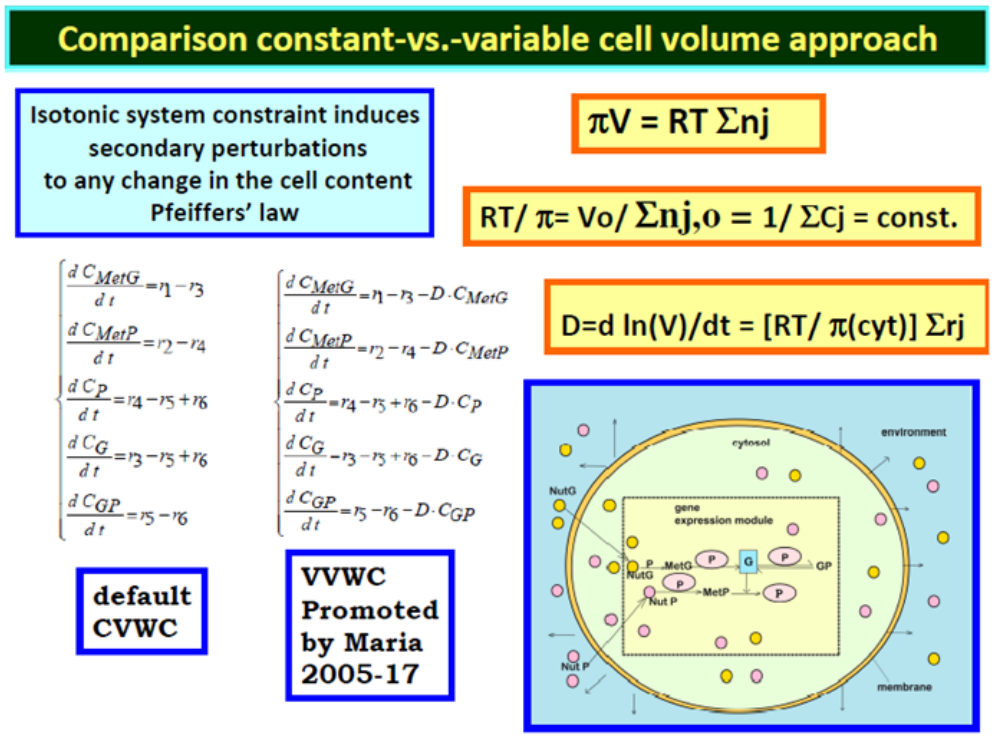


Figure 55: Comparison of the CVWC and VVWC modelling approach in the case of a simple G(P)1 gene expression regulatory module [53].

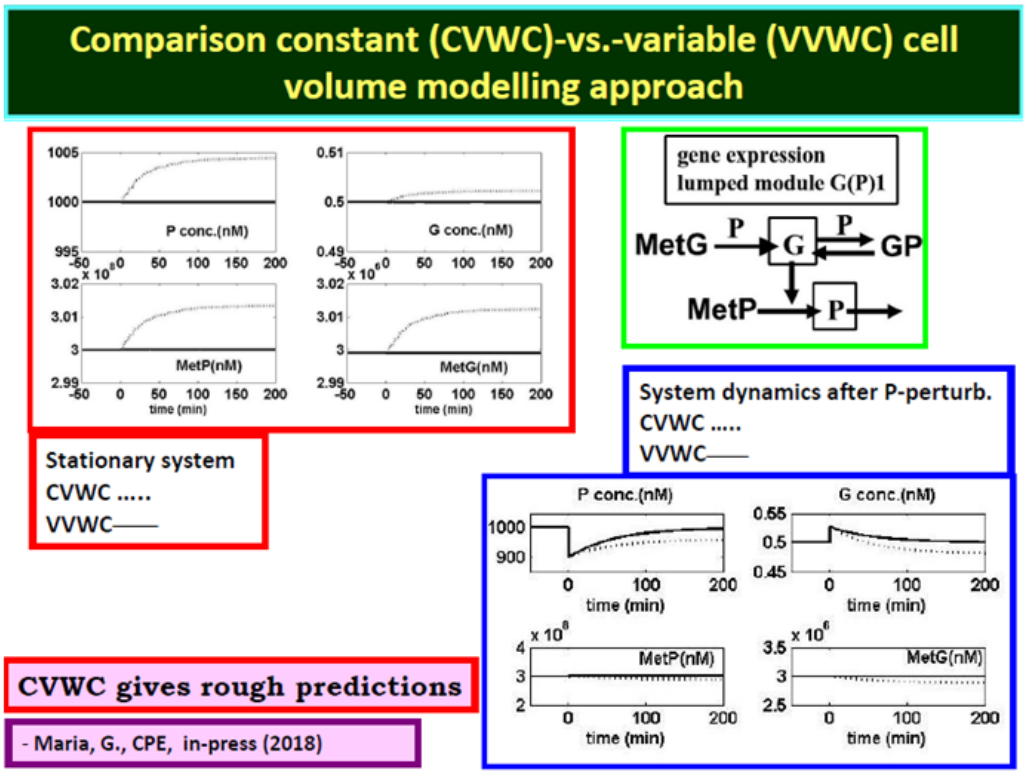


Figure 56: (continued) Comparison of the CVWC and VVWC modelling approach in the case of a simple G(P)1 gene expression regulatory module [53].

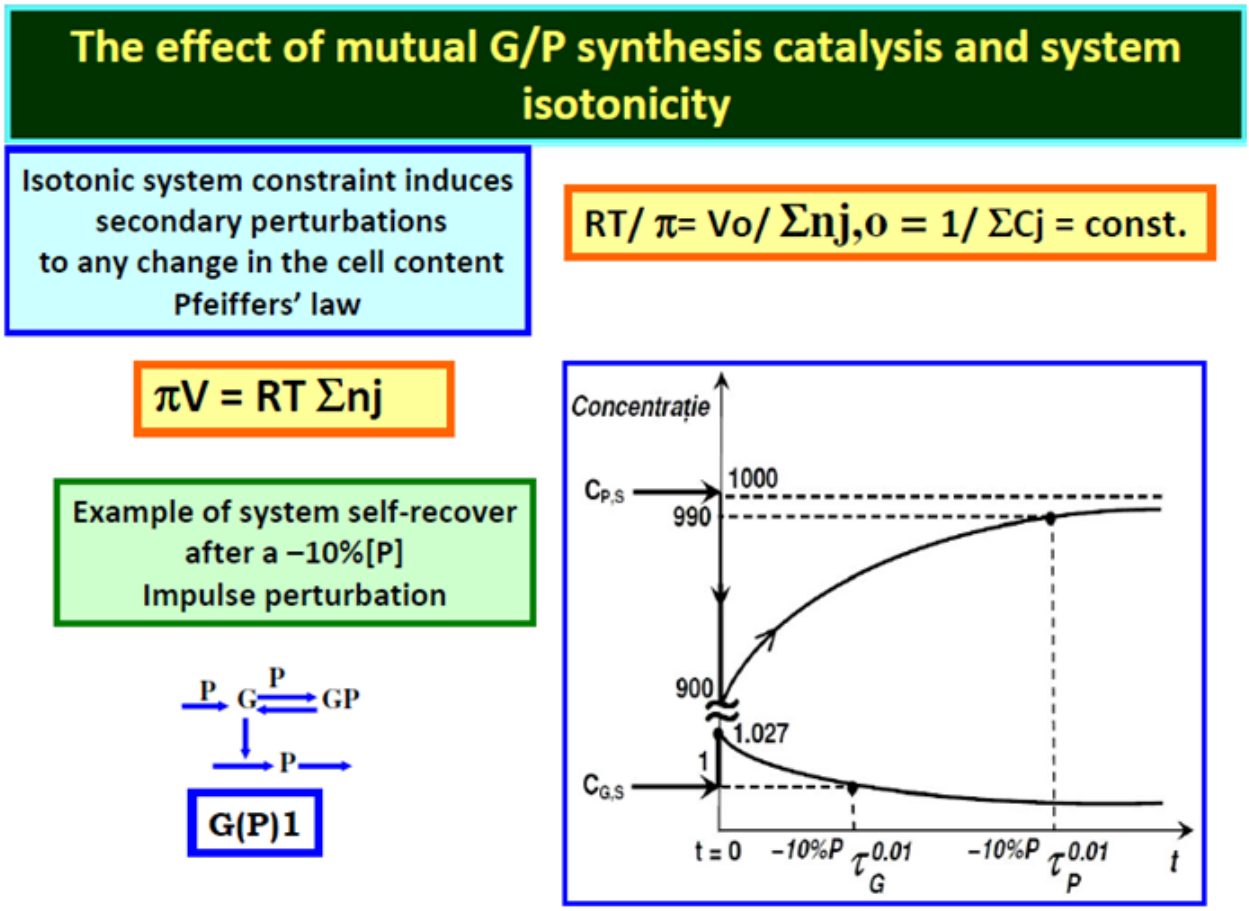


Figure 57: The effect of the mutual G/P catalysis and of the isotonicity in the case of a simple G(P)1 gene expression regulatory module [1,2,17,46].

Recovery rates are properties of all interactions within the system rather than of the individual elements thereof [6]. In terms of the evolution and stability of component QSS concentrations included in a dynamic cell system expressed by an ODE model (1) or (2), these properties can be evaluated from the analysis of the eigenvalues λ (i) (i = no. of species) of the linearized model Jacobian matrix of elements. $J(i,k) = \partial h_i(C,k) / \partial C_k$ [1] If small perturbations of a steady state Cs are considered then, this steady state is asymptotically stable if the real parts of the Jacobian eigenvalues are all negative, Re(λ (i))<0, for all "i" [6,129,130]; Heinrich and. If the 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 (stationary) perturbation.

Here it is to mention the works of [17,52,126] that proved that the optimum concentrations in the "buffering" reactions of GERM-s involving the active and inactive forms of the "catalyst" ensuring the maximum regulation *dynamic efficiency vs*. perturbations (see below) are those of [G]= [GP], [Gi]=[GiPjPj], of [M]= [MP], [Mi]=[MiPjPj], etc. (for the GERM-s of Figure 49).

The main dynamic P.I. are discussed by Maria[1], that is (Figure 52):

a) *Recovering time* necessary to each GERM component to return to their stationary concentration (QSS) after an impulse-like perturbation in one component. This P.I. measure of the GERM efficiency to fast recover the key-species stationary concentrations is given by the time τ (j) necessary to the species "j" to recover its steady-state concentration (with an assumed tolerance of 1%, as proposed by Maria [17]. As an example, in Figure 57&58 is presented how a GERM of [G(P)1] type is working, by recovering the stationary [P]s and [G]s after an impulse perturbation in the [P]s, that is a –10% decline of [P]s at an arbitrary time t=0. As another example, in Figure 59 is presented how the QSS recovering time depends on the GERM structure, and its number of effectors (TF). The regulatory efficiency increases in the order: [G(P)]0 ("1") < [G(P)1] ("2") < [G(P)1;M(P)1] ("3") < [G(PP)2] ("4").

b) *Recovering rate RD* (Table 1) necessary to each GERM component to return to their stationary concentration (QSS) after an impulse-like perturbation in one component.. As an example, in Figure 54 is presented how RD depends on the GERM structure, and its number of

effectors (TF). The recovering rate RD reflects the recovering properties of the regulated key-P synthesis by the GERM system. In a simpler way, the species "j" recovering times $\tau(j) \sim 1/RD$ and trajectories Cj(t) can be obtained by *simulation* , that is by simulating the GERM system dynamics [with using the GERM ODE model eq.(1), or(2,6)] after applying a small impulse perturbation of the species steady-state of +/-10% Cj,s and determining the recovering time until the steady-state Cj,s is reached with a 1% tolerance [17]. Species recovering trajectory and amplitude are both very important (Figure 54,59). As proved by (Maria, 2003[52]), GERM-s display very different recovering trajectories and amplitudes (Figure 54). The most effective are the GERM-s ensuring the smallest amplitude of the recovering pathway, thus not disturbing the other cell metabolic processes. As underlined by Maria [52], the recovering trajectories in the G/P phase plane is more linear for the efficient GERM-s, presenting a lower amplitude, thus not disturbing other cell processes.

c) *Regulatory robustness*. The regulatory robustness of a GERM model is defined by Maria [17] as being the property to realize (Min) $(\partial R_D / \partial \boldsymbol{k})$, where RD denotes the key-species recovering rates, while k= rate constant vector (depending on the micro-organism). This P.I. can be considered a systemic regulatory property, as long as GERM species levels are able to modify the apparent reaction rates. In fact, the cell metabolic network robustness and functionality are linked to the cell phenotype and gene regulation scheme.

d) *Species interconnectivity* in a modular regulatory schema of reactions can be viewed as a degree to which they 'assist' each other and 'cooperate' during the GERM system recovering. Cell species connections appear due to common reactions, or common intermediates participating to chain reactions, or from the common cell volume to which all cell species contribute (under constant osmotic pressure, eq. (2-5), and VWC model hypotheses of the chap. 9.2. [131] reviewed and proposed several quick experimental- computational rules to check a reaction schema via species inter-connectivities. By inducing experimental perturbations to a (bio)chemical system, by means of tracers, or by fluctuating the inputs of the system, one can measure the perturbation propagation through the consecutive/parallel reaction path. Then, various techniques can determine the "distance" among observed species, and rules to include this information in elaborating a reaction schema [17,46,54-57,151,152,154,158]. Proposed a similar approximate measure of species interconnectivity related to the species recovering-times after a dynamic perturbation, that is: AVG(τ(j)) and STD(τ(j)), i.e. the average and the standard deviation of the species individual recovering times τ(j). As AVG and STD are larger, as the cell dynamic regulatory effectiveness is lower, species less interconnected, and components recover more disparately (scattered). The higher the number of effectors and buffering reactions, the better these dynamic regulatory indices of the GERM are [1-2,17,46,52-56,57,151,152,154, 158].

e) *Cell sub-system QSS stability* (and, in particular, of a GERM) refers to the system's capacity to recover the QSS after cessation of a dynamic perturbation. Such a property can be predicted by analysing the QSS of the dynamic GERM ODE model (2) or (6). The stability property can be evaluated from the analysis of the eigenvalues λ (i) (i = no. of species) of the linearized model Jacobian matrix of elements $J_{ik} = \partial h_i (C, k) / \partial C_k$. The QSS is asymptotically stable if the real parts of the Jacobian eigenvalues λ (i) are all negative, that is Re(λ (i))<0 for all "i", [6,129,128,130]. If the 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.

Here it is to mention two important observations:

i. A characteristic of the VVWC models including the Pfeiffers'constraint eq.(3-5) is that they are always stable (intrinsic stability), because, as proved by [88], always Max(Re(λ (i)))= -D (see eq.(2,4)).

ii. By contrast, one fundamental deficiency of the classical CVWC model formulations is the lack of the intrinsic stability of the cell system model, because these models do not include the Pfeiffers'constraint (3-5). Consequently, the GERM regulatory mechanism for recovering the homeostasis (illustrated in Figure 49,57,59), is not longer working and the CVWC model becomes invalid and ineffective.

10.2. Some rules to link GERM when for modelling GRC

When modelling a GRC consisting of a chain of GERM-s, they are two problem to be considered: I) what effective GERM-s have to be chosen to match the individual expression characteristics, and ii) what rules to be applied when linking GERM-s to get the holistic regulatory properties of the GRC in the context of the cell balanced growth. In this chapters some of such linking rules are briefly reviewed following reported results of [1].

When linking GERM-s to construct a GRC reproducing a certain function of the cell, there are two contrary trends: on one hand is using simple GERM structures to reduce the model identification computational effort; on the other hand, it is important to use simple, but effective and flexible GERM-s able to reproduce individual enzyme-synthesis, but also holistic properties of the GRC (of Table 1; complex examples are provided in chap.11). Below, there are reviewed some rules to be applied for adjusting the GERM-s and GRC regulatory properties.

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 49). As proved by Maria [17,46,52] and

by Yang et al. [127], a quasi-linear relationship of P.I. function of no. of regulatory effectors (n) can be derived for every GERM type, of the form $P.I. = a_0 + \Sigma_i a_i n_i$. Here, P.I. denotes the regulatory performance index, such as RD, AVG(τ (j)), STD(τ (j)), stability strength, etc. Also, n(i)= number of effectors (P, PP, O) acting in the "i-th" allosteric regulatory unit Li(Oi)n(i); Notations a(o),a(i) denotes the correlation constants related to the P.I. and module type. Such a dependence can also be observed in the Figure 53&54. In short, [17,46,56,57] proved that (Figure 53,54):

(i) P.I. improves ca. 1.3-2 times (or even more) for every added regulatory unit to the module. Multiple regulatory units lead to an average recovering time AVG(τ (j)) of all GERM species much lower than the cell cycle duration tc, under a constant logarithmic volume growing rate, D= ln(2)/tc. Combinations of regulatory schemes and units (with different effectors) can improve the regulatory P.I.-s.

(ii) 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 [G(P)n;M(P)n1], and of [PP]s in modules [G(PP)n] (Figure 49). Optimal levels of these species can be set accordingly to various optimization criteria, rendering complex regulatory modules to be more flexible in reproducing certain desired cell-synthesis regulatory properties.

II. *Ranging the number of transcription factors TF and buffering reactions:* For selecting 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 (Figure 49,59,61,65) necessary to be included in the model. Evaluation of P.I.-s for a large number of GERM structures (Figure 49) [17,46,49,54-57,151,152,154, 158] indicated that the dynamic regulatory efficiency of [G(P)n] modules is nearly linearly increasing with the number (n) of buffering reactions (correlation $P.I. = a_0 + \Sigma_i a_i n_i$, and Figure 53,54). Moreover, the plots of Figure 54 reveal that this increase is more pronounced in the case of [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)n1] modules that use a control scheme in cascade 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 homeostatic QSS, and species recovering trajectories more linear in the G/P phase plane and of a lower amplitude.

To summarize, when selecting a suitable GERM to be included in a GRC the following issues are to be considered [1-2, 17,46,54,57, 56]:

(i) Modules reporting high stationary-regulation P.I.-s also report high dynamic-regulation P.I.-s.

(ii) 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 54,59), that is lacking of buffering reactions able to modulate the gene G and M catalytic activity) appears to be of lowest regulatory efficiency.

(iii) Multiple copies of effector molecules (i.e. O, R, P in Figure 45,49,50,65), which reversibly and sequentially (allosterically) bind the catalyst (G, M) in negative feedbacks, improve the regulation effectiveness.

(iv) A structured cascade control of the "catalyst" activity, with negative feedback loops at each level as in the [G(P)n;M(P)n1] model series, 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 45,49,50,65 by placing regulatory elements (R,P,PP,…) at the level of mRNA (i.e. species M), and at the level of DNA (i.e. species G) in the [G(P)n;M(P)n1] model is highly effective.

(v) The nearly linear increase of GERM P.I.-s with the number n(i) of effectors (P, PP, O, R, etc.) acting in the i-th allosteric unit [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 1.

(vi) P.I.-s improves ca. 1.3-2 times (or even more) for every added regulatory unit to the same GERM type. Multiple regulatory units lead to much lower average recovering times AVG(τ(j)) than the cell cycle period tc, under constant logarithmic volume growing rate, Ds = ln(2)/ tc.

(vii) Combinations of regulatory schemes and units (with different effectors) might improve the regulatory P.I.-s [46].

(viii) Certain regulatory 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 [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 desired cell-synthesis regulatory properties. Thus, (Maria, 2009[54]) proved existence of an optimal [TF] concentration leading to optimal P.I. of a GERM (Figure 65).

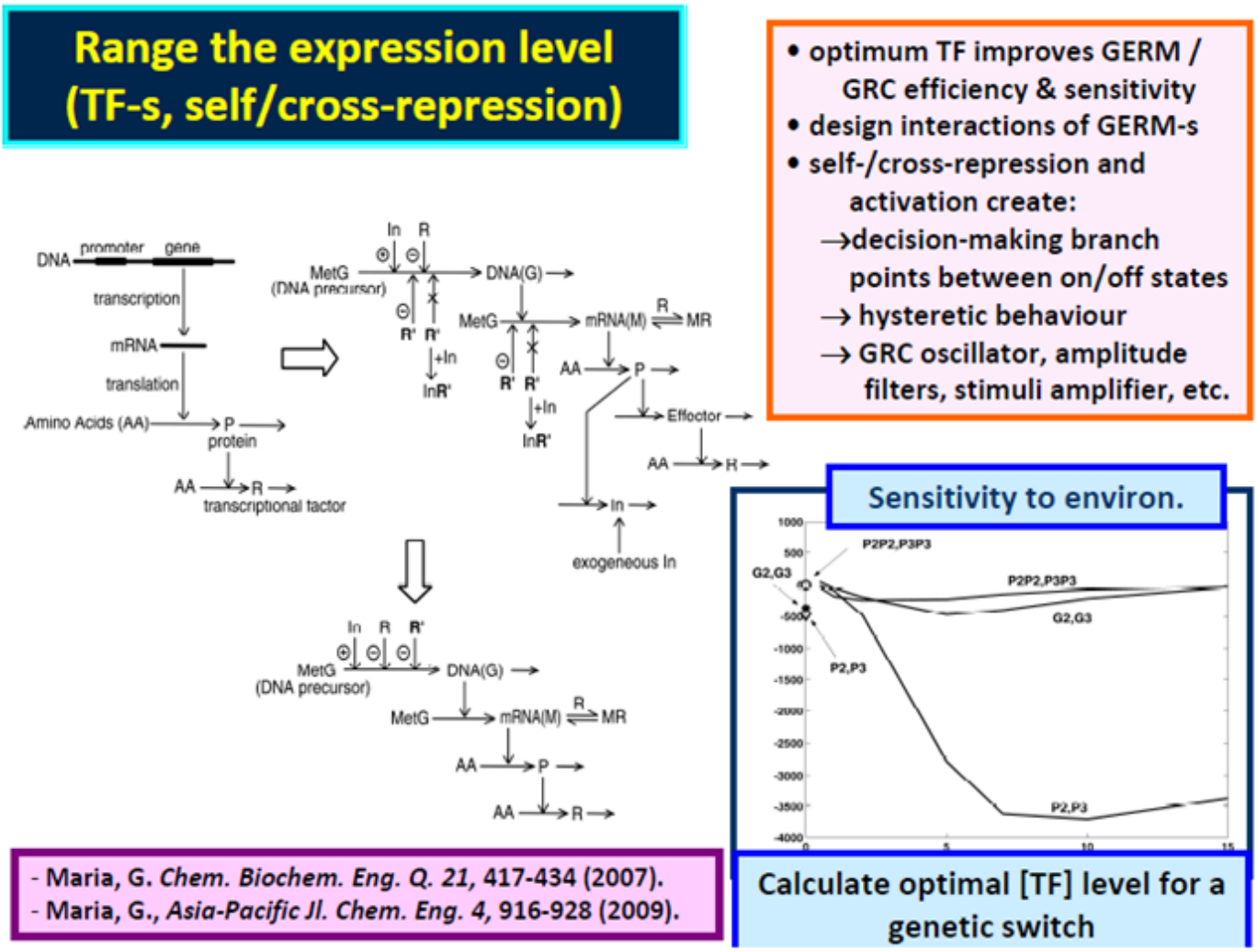
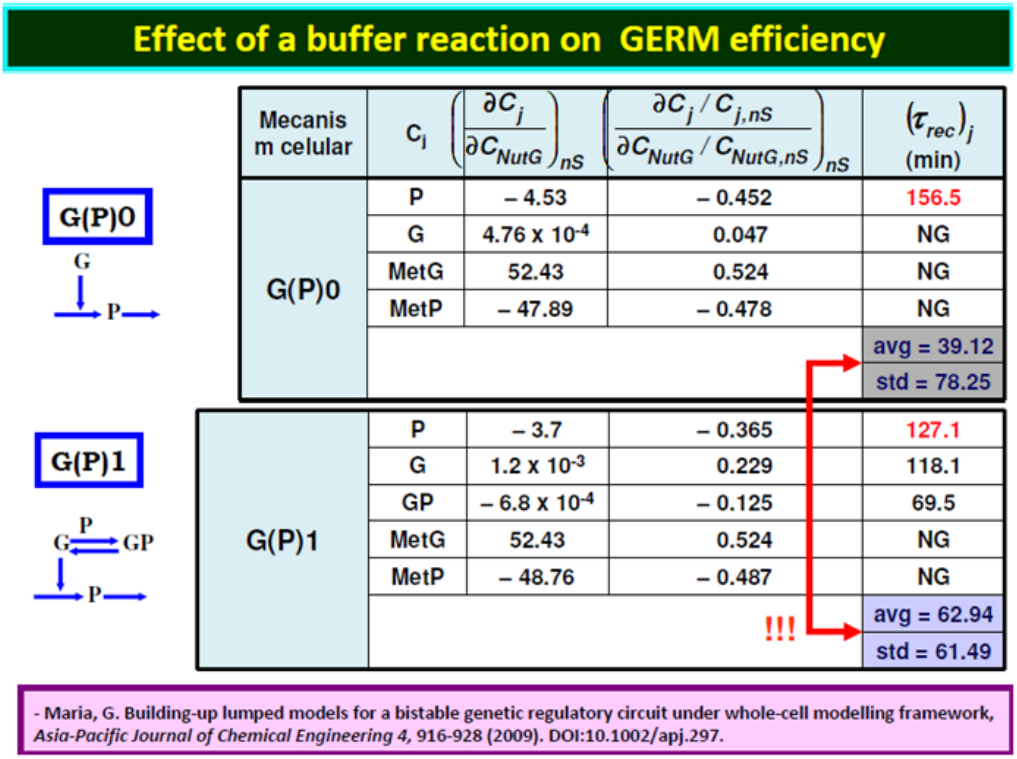


Figure 65: Effect of the TF level on the GERM efficiency [54,56].

III. *The effect of the mutual G/P synthesis catalysis*. One essential aspect of the [G(P)n], [G(PP)n], and [G(P)n;M(P)n1] kinetic models of GERM is the mutual catalysis of G and its encoding protein P synthesis. If one adds the VVWC modelling constraints eqn. (2-6) and the requirement of getting a maximum dynamic responsiveness and efficiency by keeping [G]s = [G(P)]s = [G(PP)]s = ...= [G(P)n]s, as discussed by (Maria, 2003[52], 2005[17], 2017[1]). This direct and indirect link of G and P syntheses ensures a quick recovering of both stationary [G]s and [P]s after any small perturbation. To prove this in a simple way, one considers a synthesis of the G/P pair in a GERM of [G(P)1] type (denoted by "2") or a [G(P)0] type (denoted by "1") (Figure 49,59). After estimating the rate constants from solving the stationary model equations by using the homeostatic concentrations of species in Figure 60 (high ballast cell case), one determines each GERM dynamic efficiency by applying a –10% impulse perturbation in the [P]s = 1000 nM at an arbitrary t=0. The obtained recovering trajectories of P and G obtained by model simulations are plotted in the Figure 57,59. The plots reveal a very good regulatory efficiency of the [G(P)1], both G, and P species presenting relatively short recovering rates, and negligible for the other species. These plots reveal in a simple way the self-regulation of the G/P pair synthesis: after the impulse perturbation leading to the decline of [P]s from 1000 nM to 900 nM, the very fast buffering reaction G + P <===> GP leads to restore the active G, whose concentration quickly increases to [G]= 1.027 nM; 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 58, the dynamic efficiency of the module [G(P)0] is much lower, species recovering their QSS over longer transient times. Also, the species connectivity is better in the [G(P)1] compared to [G(P)0], being reported smaller STD(τ(j)). Consequently, removal of the buffering reaction that automatically adjusts the "catalytic activity" of G, will: decrease the species inter-connectivity (increasing the standard deviation of the recovering times); will increase the species recovering times; will increase the sensitivities of the species steady-state vs. external nutrients (see sensitivity coefficients vs. NutG in the Figure 58). As expected, the P.I.-s of the GERM depend not only on I) the no. of effectors (buffering reactions), but also on ii) the TF type (P, or PP), and even more on the used control scheme (i.e. simple or in cascade).



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	4.76×10^{-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	1.2×10^{-3}	0.229	118.1
	GP	-6.8×10^{-4}	– 0.125	69.5
	MetG	52.43	0.524	NG
	MetP	– 48.76	– 0.487	NG
			!!!	avg = 62.94
				std = 61.49

Figure 58: The effect of a buffer reaction effector (G+P <=> GP) on the G(P)1 gene expression regulatory module [54].

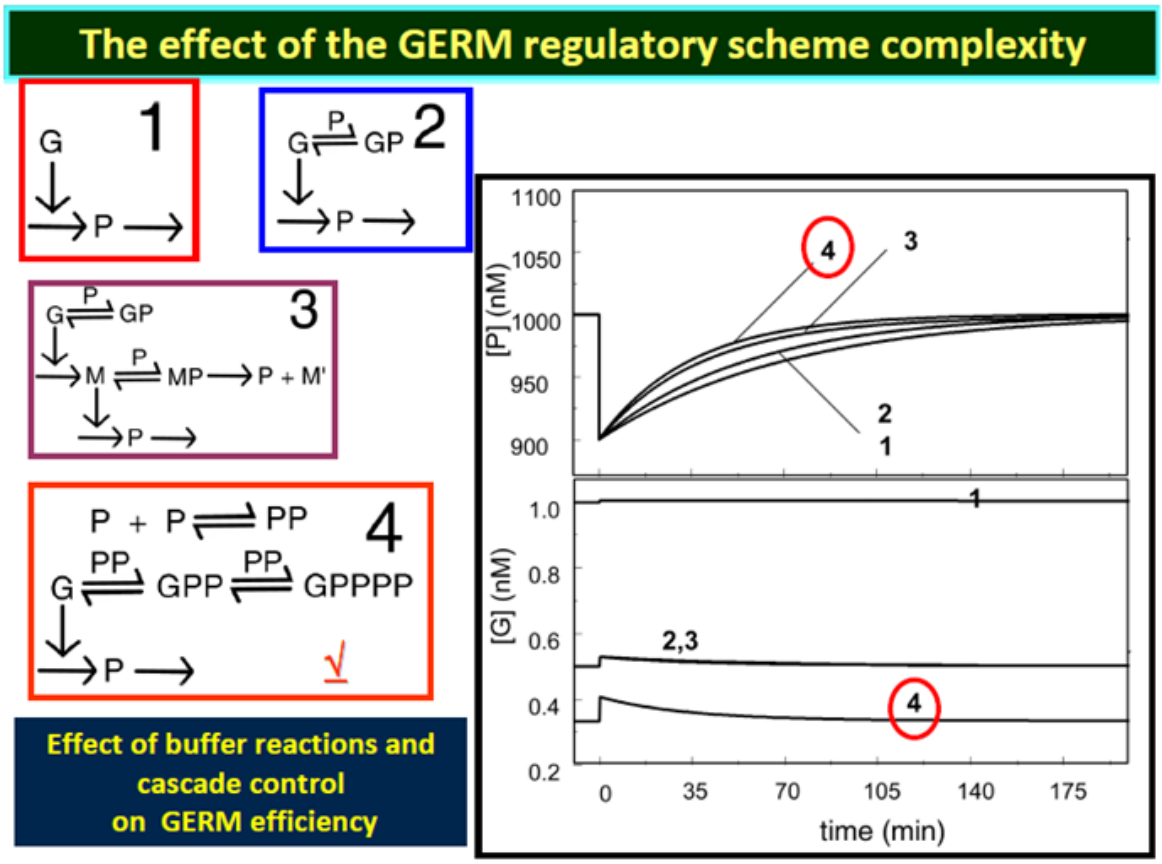


Figure 59: GERM model complexity reflected on its prediction quality [1].

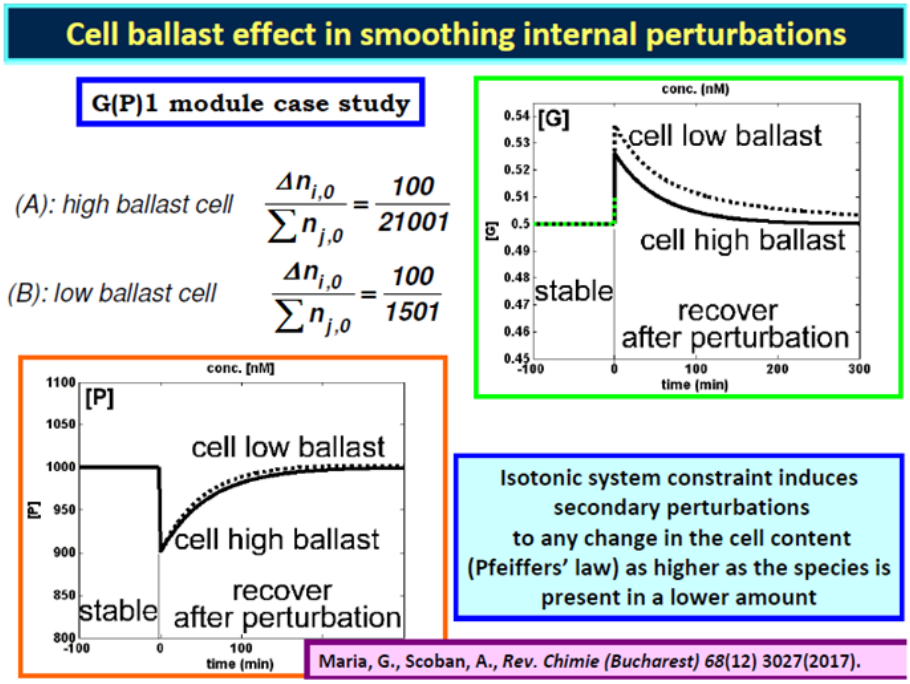


Figure 60: Cell ballast effect on the GERM performance indices [55].

To exemplify these issues, one considers the same a G/P gene expression example with the species homeostatic stationary concentrations given by Maria [55], the high ballast cell case. For comparison, one considers the gene encoding gene G expression by means of GERM-s of various structures given in the Figure 59, that is [G(P)0] without mutual catalysis, [G(P)1] with mutual catalysis and one buffering reaction, or [G(PP)2] with dimeric TF=PP, or even [G(P)1;M(P)1] with mutual catalysis and a cascade control via buffering reactions at the level of G and M. The rate constants have been estimated by solving the stationary form of the GERM model with the stationary concentrations. Additionally, the requirement of getting a maximum dynamic responsiveness and efficiency discussed by Maria [1,17,52], leads to adopt [M]s = [M(P)]s = 0.5 nM and [G]s = [G(PP)]s = [G(PPPP)]s = 1/3 nM. The resulted recovering trajectories of the G and P species after a –10% impulse perturbation in the [P]s = 1000 nM at an arbitrary t=0, are comparatively presented in the Figure 59. It is to remark that the incomplete [G(P)0], reports the worst dynamic efficiency, with very slow recovering tendencies in Figure 59. Better performances are reported by [G(P)1]. Even better regulatory efficiency is reported by the cascade control of separately considered transcription and translation in the [G(P)1;M(P)1] module. The best QSS recovering efficiency is reported by the [G(PP)2] module that uses two buffering reactions and a dimeric PP as TF, quickly synthesized in a small amount (of optimal level [PP]s =0.01nM determined together with the model rate constants to ensure an optimal P.I.).

All the above analyzed GERM-s have been modelled in a VVWC framework (eq. 2-6). For such VVWC kinetic models it is to remark the way by which the variable cell-volume plays an important role to species inter-connectivity (direct or indirect via the cell volume) in the same GERM regulatory module or among linked modules. Even if species connectivity can be expressed in several ways [17,131]; it is directly dependent on the manner by which species in a GERM or in a GRC recover more or less independently after a perturbation. When the species connectivity increases, they recover with a more comparable rate (or equivalently, over the same time), by 'assisting' each other to cope with a perturbation (see the comparison of species recovering times in the Figure 58). By contrary, when the species are more disconnected, they recover in a more disparate way, and the GERM presents weaker P.I.-s. [reporting not only larger species recovering times τ (rec)(j), but also larger state sensitivities to external nutrients].

Thus, the *mutual autocatalysis* G/P appears to interconnect members of the GERM key-components such that they are regulated more as a unit than would otherwise be the case. Interconnectivities (the degree to which a perturbation in one component influences others) may arise from a direct connection between components (e.g. when they are involved in the same chain of reactions), or from an indirect connection (via cell volume changes for an isotonic system). Our analysis indicates that mutual auto-catalysis is a particularly strong type of interaction that unifies the regulatory response, and they serve to "smooth" the effects of perturbations. It also suggests a way to quantitatively evaluate interconnectivities between all cellular components: each component could be perturbed one at a time, and recovery rates or some other measure of regulatory effectiveness could be evaluated for all components. The resulting relationships thus reflecting the holistic properties of the GRC-s.

IV. *The effect of cell system isotonicity*. The effect of the isotonicity constraint eqn. (2-6) of a VVWC cell model can be easily proved [1,55]. By simulating a GERM with a [G(P)1] model type, the effect of applying a –10% impulse perturbation in the key-protein homeostatic level

[P]s = 1000 nM at an arbitrary time t=0, the effect on the key-species (G, P) can be observed in the Figure 57, while species recovering times are given in the Figure 58. By contrast, in a CVWC cell model formulation, when the isotonicity constraint is missing from the model, the key-species do not recover. By contrast, as revealed by the simulations with the [G(P)1], the system isotonicity imposes relatively short recovering rates for the key-species, and negligible for the other GERM species present in a large amount (lumped nutrients and metabolites). As proved by Maria [53,55] the VVWC models, with including the "cell ballast" effect, and the G/P mutual autocatalysis, are more flexible and adaptable to environment constructions, being able to better represent the influence of the environmental changes on the cell homeostasis.

V. *The importance of the adjustable regulatory TF-s in a GERM*. As proved by the example of Figure 59, and those of [17,52], dimeric TF-s, such as PP in [G(PP)n] instead of simple P in [G(P)n], leads to several conclusions:

(i) The dynamic regulatory efficiency increases in the order: [G(P)0] (no buffering reaction) < [G(P)1] (one buffering reaction) < [G(P)1;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 [G(P)n;M(P)n1] and of [PP]s in modules [G(PP)n]. Optimal levels of these species can be set accordingly to various optimization criteria, rendering complex GRC-s to be more flexible in reproducing certain desired cell-synthesis regulatory properties.

(ii) The dynamic regulatory efficiency (defined in Table 1) decreases in the following order [17,52]:

Min(τ (rec)P): [G(PP)2] > [G(P)1;M(P)1] > [G(P)1] > [G(P)0]

Min(STD): G(PP)2 > [G(P)1;M(P)1] > [G(P)1] > [G(P)0],

the stationary regulatory efficiency of [P]s decreases in the order:

Min S([P];[NutG]): [G(PP)2] > [G(P)1;M(P)1] > [G(P)1] > [G(P)0]

VI. *The effect of the cell ballast on the GERM efficiency:* When constructing more or less simplified VVWC cell models, it is important to know what is the minimum level of simplification to not essentially affect the holistic properties of the cell. This paragraph proves why it is essential to include in a VVWC model the so-called "cell ballast", that is the sum of concentrations of all species, which are not accounted in the ODE mass balance of the GRC model. Basically the isotonic constraint 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. As proved by Maria [55], and Figure 60, in such VVWC cell model constructions, the recovery rates are properties of all interactions within the system rather than of the individual elements thereof [6].

However, another important question derived from the isotonicity constraint refers to the degree of importance of the cell content (ballast) for the cell reactions and resistance to 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 from a "poor" cell of low cell content (ballast)? The answer is no. To simply prove that, one considers a GERM of type placed in an *E. coli* cell with two different nominal conditions given in Table 5: a high-ballast cell, and a low-ballast cell. To not complicate these models, lumped gene and protein metabolites have been considered. Being present in a large amount (that is [MetG]= 3E+6 nM, and [MetP]= 3e+8 nM), these components also play the role of cell ballast, their concentrations being set to values much larger than those of the other cell species. Simulation [55], and Figure 60) 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 in the Figure 60.

Selection of appropriate lumped [MetG] and lumped [MetP] will lead to understanding their effect on the cell self-regulatory properties. Low concentrations relative to the total number of other molecules in the cell afforded shorter recovering times τ(rec)(P) for the key-protein P. For instance, in the [G(P)1] module case, with lumped [MetG]=2000 nM, and lumped [MetP]=3000 nM, and (all [Cj]) = 12001 nM, the resulted recovering times of key-species G/P are τ(rec)(P)= 103 min, and τ(rec)(G)=223 min after a –10% impulse perturbation in the [P]s of 1000 nM at an arbitrary t = 0 (Figure 60). Whereas for a cell case with lumped [MetG]=3e+6nM, and lumped [MetP]=3e+8nM, and (all [Cj]) = 6.06e+8nM, the resulted τ(rec)(P)= 133 min, and τ(rec)(G)=93 min after a –10% impulse perturbation in the [P]s of 1000 nM at an arbitrary t=0 (Figure 60).

We refer to this as the *Inertial Effect*. It arises because the invariance relationships described above require that larger rate constants for P and G synthesis be used to counterbalance lower [MetP] and [MetG], and these constants are determinants for key-species recovering rates (τ(rec)(j)) after a perturbation. On the other hand, when metabolite concentrations were low, perturbation of cell volume was greater than when they were high (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 60). 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' shows how all components of the cell are interconnected via volume changes. It represents another holistic property of cells, and it is only evident with only variable-volume VVWC 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" will be insignificant due to the large number of total intracellular species. However, the sum of all perturbations experienced during a cell cycle might be significant.

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 is the degree of detail of the adopted GERM-s to accurately reproduce the GRC regulatory properties. The examples discussed below and by Maria [46] revealed that more important than the number of considered species in the regulatory loops is the selected GERM regulatory scheme, able to render the GRC holistic synchronized response to environmental perturbations.

Consequently, when developing a suitable VVWC kinetic model of a GRC, it is important to adopt a suitable reduced model structure by means of an acceptable trade-off model simplification-vs.-model quality (adequacy).

Adoption of too complex reaction pathways 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 parameters and unknown steady-state concentrations. Beside, cell model constructions with too complex cell modules lead to inoperable large models impossible to be used for cell design purposes. The alternative is to use reduced ODE models with a number of lumped species and enough reactions to fairly reproduce the experimental data, but simple enough to make possible a quick dynamic analysis of the metabolic process and of its regulation properties.

To exemplify how a suitable trade-off between 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 expression of G1/P1 and G2/P2 pairs) such that the resulted GRC to present optimal P.I.-s. To solve this problem, [46] compared two linking alternatives (Figure 61,62):

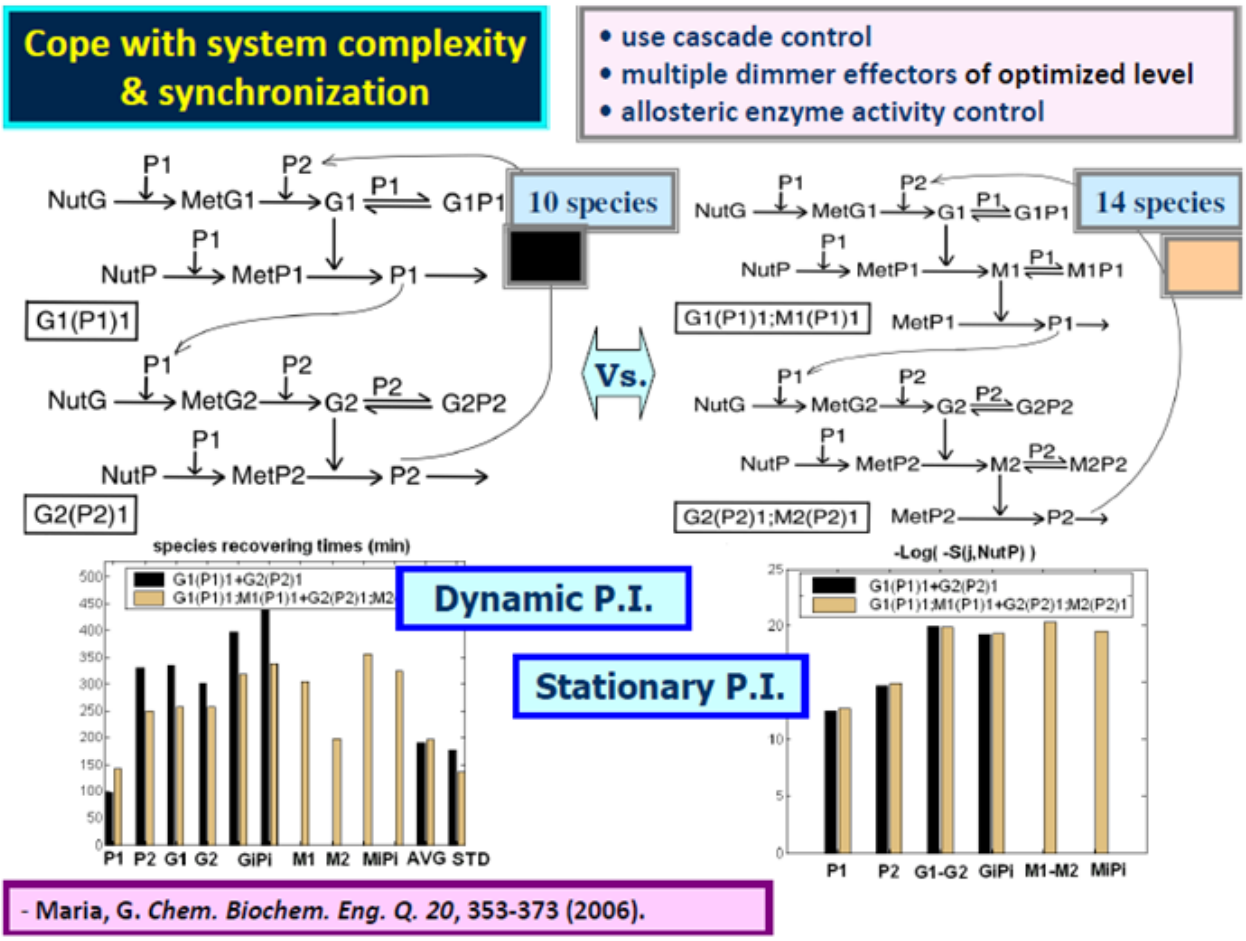


Figure 61: Effect of the GERM model complexity on the GRC stationary performances (P.I.) [46].

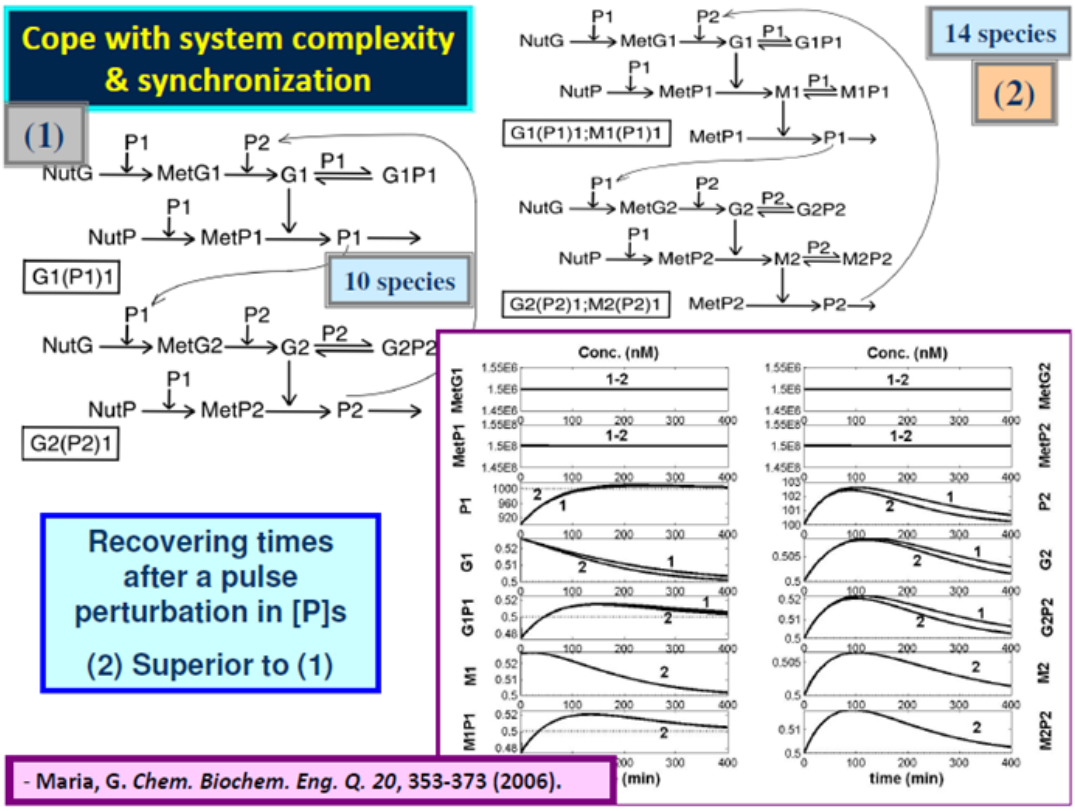


Figure 62: Effect of the GERM model complexity on the GRC dynamic performances (P.I.) [46].

a. Alternative A: [G1(P1)1]+[G2(P2)1] (10 individual and lumped components).

b. 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 as following: the expressed P1 in [G1(P1)1] 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 in the modules [G1(P1)1]+[G2(P2)1].

The GERM linking in the alternative B is as following: : the expressed P1 in [G1(P1)1;M1(P1)1] is the metabolase that converts NutG in MetG2 and NutP in MetP2 in the [G2(P2)1;M2(P2)1]. In turn, the expressed P2 in [G2(P2)1;M2(P2)1] is the polymerase that converts MetG1 in G1 in the module [G1(P1)1;M1(P1)1].

Simulations revealed that alternative B is superior, presenting better stationary P.I. (Figure 61), and dynamic P.I. (Figure 62). The cost is minimum, 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 two more buffering reactions), the GRC variant B presents much better P.I.-s, that is (values not presented here): I) key-species shorter recovering times after an impulse perturbation; ii) lower AVG and STD species connectivity indices; iii) species QSS concentrations lower sensitivity vs. environmental perturbations.

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), the cascade control of the gene expression in modules of [G(P)n;M(P)n1] type (Figure 49) presents superior regulatory properties suitable for designing robust GRC-s, with easily adjustable properties via model parameters, including a better species synchronization when coping with perturbations (i.e. low AVG, STD indices).

VIII. *Cooperative vs. concurrent linking of GERM-s in GRC and species interconnectivity*. When coupling two or more GERM modules into the same cell, the nutrients, and metabolites in the G/P syntheses are roughly the same (Figure 55). The modelling problem is what alternative should be chosen ? (see Figure 63)? 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, there are tested three alternatives of module coupling illustrated in the Figure 63, that is:

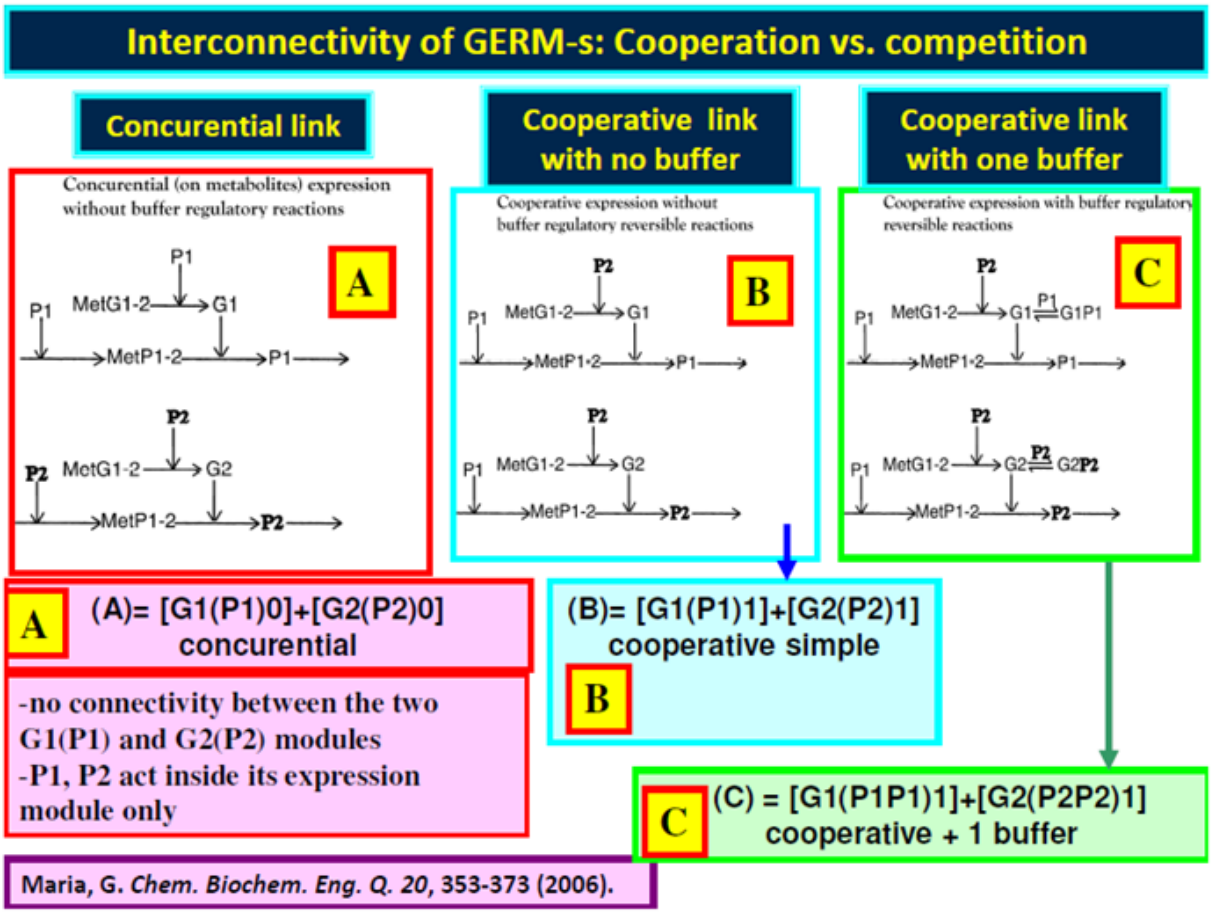


Figure 63: Effect of the GERM inter-connectivity and of individual functions into the cell on the GRC efficiency [46].

a) Alternative A: *Competitive* expression (competition on using the common metabolites) of the type [G1(P1)0]+[G2(P2)0];

b) Alternative B: *Simple cooperative* expression 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.

c) Alternative C: *Complex cooperative* expression (identical to alternative B), but adding buffer reversible regulatory reactions to modulate the G1, G2 catalytic activity in the modules [G1(P1)1]+[G2(P2)1].

Tests were performed by using the VVWC modelling framework, and the nominal high-ballast cell condition of Table 5. Simulations lead to very interesting conclusions Maria Complex [17].

In the Alternative A, one links two modules [G1(P1)0]+[G2(P2)0], both ensuring regulation of the two proteins (P1, P2) synthesis, in an *concurential* disconnected way (Figure 63). For this hypothetic system, synthesis of P1/G1 and P2/G2 from metabolites is realized with any interference between modules {the simulated case study in Figure 64 corresponds to the following steady-state: [P1]s= 1000nM, [P2]s= 100nM, [G1]s= 1nM, [G2]s= 1nM). The only connection between the two GERMs is due to the common cell volume to which both protein syntheses contribute. If one checks the system stability, by applying a ±10% impulse perturbation in [P1]s, it results an *unstable system*, evolving toward the decline and disappearance of one of the proteins (i.e. those presenting the lowest synthesis rate). Consequently, the homeostasis condition is not fulfilled, the cell functions cannot be maintained, and the disconnected protein synthesis results as an unfeasible and less plausible GERM linking alternative [17]. In the Alternative B, the *simple cooperative* linking of [G1(P1)0]+[G2(P2)0] modules in Figure 63 & 64 ensures specific individual functions of each protein, i.e. P1 lumps both the permeases and metabolises functions, while P2 is a polymerase. In the Alternative C, the simple cooperative linking of [G1(P1)0]+[G2(P2)0] system of the Alternative B has been *improved* by adding simple effectors for gene activity control. In the cooperatively linked system , thus resulting the system [G1(P1)1]+[G2(P2)1], of Figure 63-64, where the effectors P1

and P2 act in two buffering reactions, G1+P1 <==> G1P1, and G2+P2 <==> G2P2, respectively, with the stationary states [G1]s= [G1P1]s= 1/2 nM, and [G2]s= [G2P2]s=1/2 nM ensuring maximum dynamic P.I.-s..

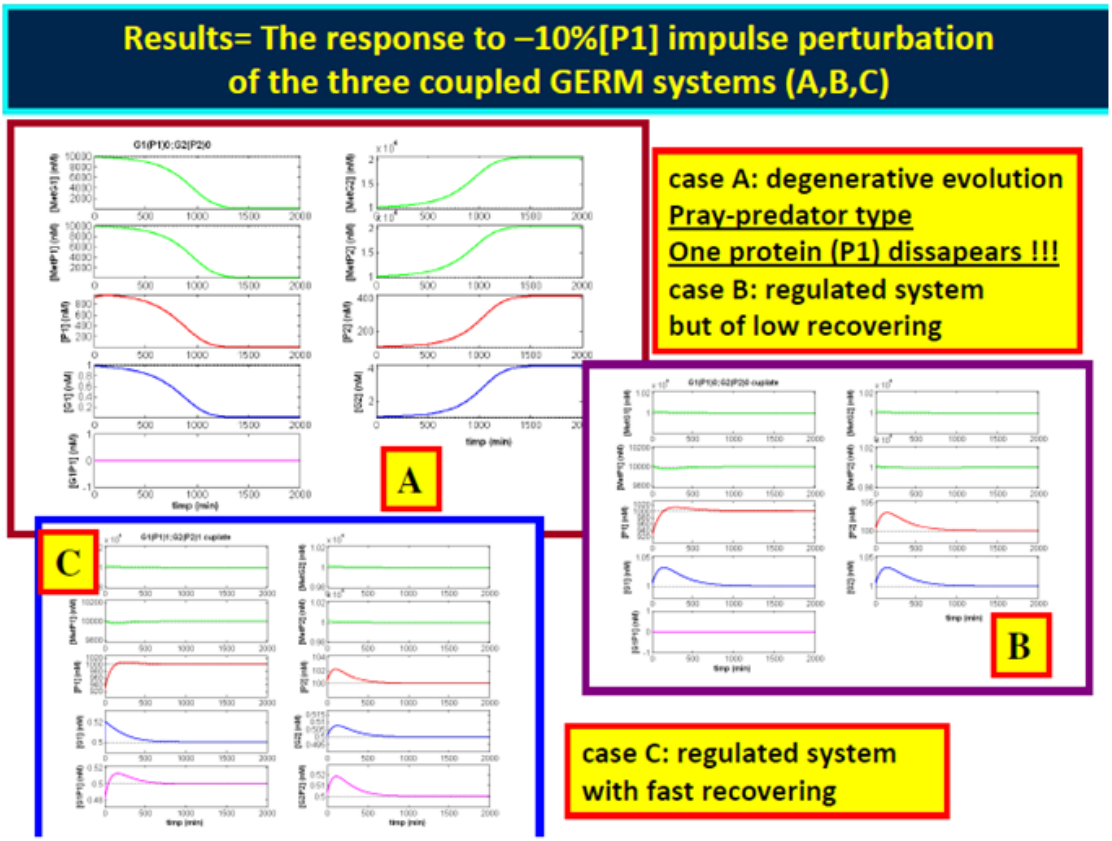


Figure 64: (continued). Effect of the GERM inter-connectivity [46].

The same rule of linking GERM-s can continue in the same way, for instance, also involving [G(PP)n] modules, where the effectors being the dimers PP, acting in "n" buffering reactions of the type, G+PP<==>GPP<==>.....<==>GPn, with the stationary states [G]s= [GP1]s = [GP2]s= ...= [GPn]s= 1/(n+1) nM. The model rate constants should be estimated from the species stationary concentration vector Cs, and by imposing regulatory optimal characteristics discussed by Maria [1] in the chap. 4.3.2). 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). Besides, the dissociation constant of the (L:TFn) complex in the buffering reactions k(diss) >> D has been adopted, e.g. (1e+5 to 1e+7)D, being much higher than other rate constants of the GERM. In subsequent works [1-2,17,46,49,54,53,56,57,151,152,154] also proved that optimization of the GERM P.I.-s with the multi-objective criteria summarized by Maria [17] leads to small values for the intermediate [PP]s (the active parts of dimeric TF).

The stability and the dynamic regulatory characteristics of all three GRC systems have been determined by studying the QSS-recover after a±10%[P1]s impulse perturbation. The results, presented by Maria [17], reveal the following aspects concerning the alternative GRC systems A, B, C:

a. All three systems are stable (that is Max(Re(λ (j)))= -D < 0 (where λ (j) are the eigenvalues of the ODE model Jacobian matrix. Systems B and C recover 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. C alternative presents the best P.I.-s from all three checked alternatives due to the additional regulatory effector of type G+P <==> GP.

b. The system is as better regulated as the effector is more effective (the use of multiple buffering reactions, with dimeric TF, and a cascade control of the expression (not presented here).

c. The use of efficient effectors and multiple regulation units can improve very much the dynamic P.I., in the following ranking: [G(P)n] < [G(PP)n] < [G(P)n;M(P)n1].

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.

In the same way, the regulatory network GRC design procedure can be continued, by accounting for new proteins (and their corresponding GERM-s). For instance, in the simplified representation of [17], a 3-rd GERM for P3 synthesis can be added to the *Alternative C*, 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 protein-metabolites (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], which regulate the synthesis of P1, P2 and P3, in a *cooperative interconnected* way, which preserves the protein functions.

IX. *The optimal value of TF*. It is self-understood that, in a realistic VVWC model, the holistic properties of the cell and of the analysed GRC should be preserved, and modulated via model structure and parameters. One of the cell modelling principles postulates that the concentration of intermediates used in the 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 GRC P.I.-s. Such optimal [TF]s are obtained by solving a multi-objective optimization problem [1,54] . An example was provided by Maria [54] in the case of a genetic switch (GS) in *E. coli* cell, modelled under the VVWC approach. The two considered self- and cross-repressing gene expression modules are of type [G2(P2P2)1(P3P3)1]+[G3(P3P3)1G3(P2P2)1]. These well chosen GERMs 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). Besides, based on adequate math models, [54] proves that it there exists an optimal level of the TF-s (that is [P2P2]s, or [P3P3]s) that are associated to the optimal holistic regulatory properties of the GRC (low sensitivity vs. external nutrients, but high vs. inducers), and that these TF-s are rather dimmers than monomeric molecules. These *in-silico* obtained results have been confirmed by the literature experimental data.

X. *Additional aspects to be considered when linking GERM-s*: Cell GRC-s and, in particular, those involved in some protein synthesis regulation, are poorly understood. The modular approach of studying the regulation path, accounting for its structural and functional organization, seems to be a promising route to be followed. Because a limited number of GERM types exist, individual GERM-s can be separately analysed, as above checked for efficiency in conditions that mimic the stationary and perturbed cell growing conditions. Efficient GERM (of regulatory indices of Table 1) are then linked accordingly to certain rules to mimic the real metabolic process, by ensuring the overall GRC efficiency, system homeostasis, and protein individual functions. Module linking rules are not fully established, but some principles reviewed by Maria [1] and below mentioned should be respected. The hierarchically organised network includes a large number of compounds with strong interactions inside a module and weaker interactions among modules, so that the whole cell system efficiency can be adjusted. By testing several ways to link GERM-s, [17] advanced some rules:

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 intermediate species, and TF levels. To preserve the individual regulatory capacity, the magnitude of linking reactions would have to decline as the number of linked modules increases.

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 path [17,52,117,138,139]. 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 above paragraph "(h)"]. 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 regulatory network nodes. There are many nodes with few connections among proteins and a small, but still significant, number of nodes with many proteic interactions. 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 well-known 'metabolic channelling' (or tunnelling) process, that ensures an efficient intermediate transfer and metabolite consecutive transformation without any release into the cell bulk phase [17].

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 not presenting common reactions, the modules are anyway linked through the cell volume (to which all cell species contribute) and due to some intermediates controlling module interactions in the GRC. The VVWC model is able to account for such cell regulatory characteristics.

v) A natural strategy for building complex and realistic cell models is to analyse independent functional modules or groups of closely interacting cellular components, and then link them. The VVWC approach may facilitate this strategy. Each 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. To allow this, and to reproduce the cell ballast effect, lumped molecular species could be defined into each cell where a GERM is tested, in amounts equal to those of the target cell minus those due to the components of the 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.

vi) The VVWC modelling approach demonstrates 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, 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, while increasing diversity allows individual metabolites to be present at lower concentrations, improving dynamic responses of GERM-s and GRC to perturbations. Another issue, thus far unexamined, is how specific types of interconnectivities affect the regulatory behaviour of cells. This could be probed using experimental methods developed by (Vance et al., 2002[131]) to deduce connectivities in biochemical pathways from the effects of impulse perturbations.

vii) When modelling complex operon structures, simple GERM structures should be adopted to not complicate too much the VVWC model. The default GERM is the [G1(P1)1]. But, according to the experimental data and interactions among genes and proteins, more complicated GERM constructions can be elaborated, as those described in the applications presented in the below chap. 8.

XI. *The effect of cascade control on the GERM efficiency*. Among GERM-s reviewed and tested under VVWC, the most significant are the [G(P)n] of effectiveness nearly linearly increasing with the number (n) of buffering reactions (Figure 53 & 54). Due to their simple structure such GERM-s are most suitable to construct complex GRCs. On the next place, the [G(PP)n] are also favorites, by presenting a more pronounced regulatory efficiency due to the used dimeric TF-s (that is PP). The most effective are the GERM-s with a cascade control of the expression, by means of buffering reactions applied at both gene G, and mRNA (M) catalyst level, that is of type [G(P)n;M(P)n1] (Figure 49). [1,17,46-52,] *in-silico* proved the superiority of the [G(P)n;M(P)n1] gene expression structures.

The conclusions are the followings:

(i) the very rapid buffering reactions, such as G + P <===> GP+ P <===> GPP ……<===> GPn, or M + P <===> MP+ P <===> MPP ……<===> MPn , have been proved to be very effective regulatory elements, by quickly adjusting the active/inactive G/GP/GPP/GPn or M/MP/MPP/MPn ratios thus efficiently coping with the perturbations.

(ii) numerical tests revealed that the P.I.-s of the compared GERM-s increase in the approximate order: [G0] (0 regulatory element) < [G(P)1] (1 regulatory element) < [G(P)1;M(P)1] (2 regulatory elements) < [G(PP)2] (3 regulatory elements),<….[G(P)n;M(P)m] (n+m regulatory elements), etc.

Roughly, the obtained improvement of the P.I. per regulatory element is of ca. 1.3 (under VVWC modelling framework), while the same improvement is of only 2.5 under CVWC modelling framework [17,46,56,54,57]. It clearly appears that the VVWC modelling framework is more realistic, the default CVWC approach tending to over-estimate the P.I.-s of GERM-s.

11. Case Studies of *In-Silico* Design of GMO

Exemplifications of such modular GRC models and *in-silico* design of GMO of industrial use includes several published case studies (Figure 66). Due to the cell metabolism complexity, and existence of both cell-level control parameters together with bioreactor macro-level control variables, optimization of an industrial bioprocess using GMO often translated in a multi-objective optimization problem [140,154,158,164], difficult to be solved. The below case studies exemplify positive experience with *in-silico* design of some GMO for improving certain bioprocesses of practical interest.

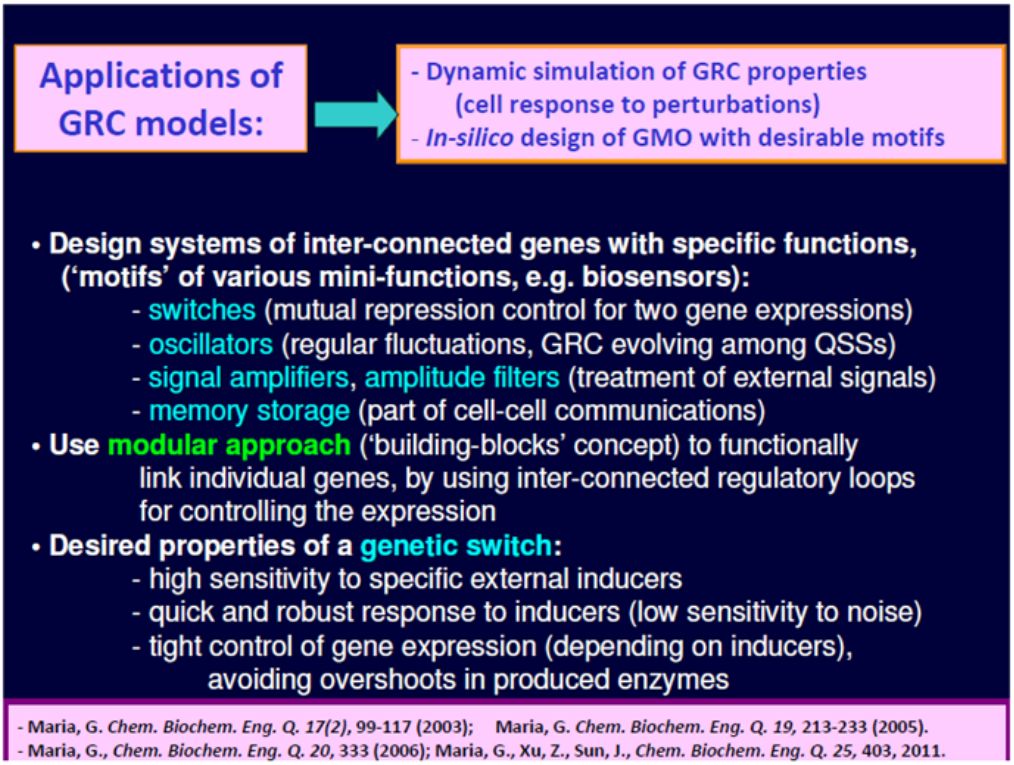


Figure 66: Some applications of GRC models [1].

11.1. *In-silico* design of an *E. coli* cloned bacterium with an improved mercury uptake efficiency, by using a structured dynamic model to simulate the self-control of the GRC responsible for the mer-operon expression [151-156]

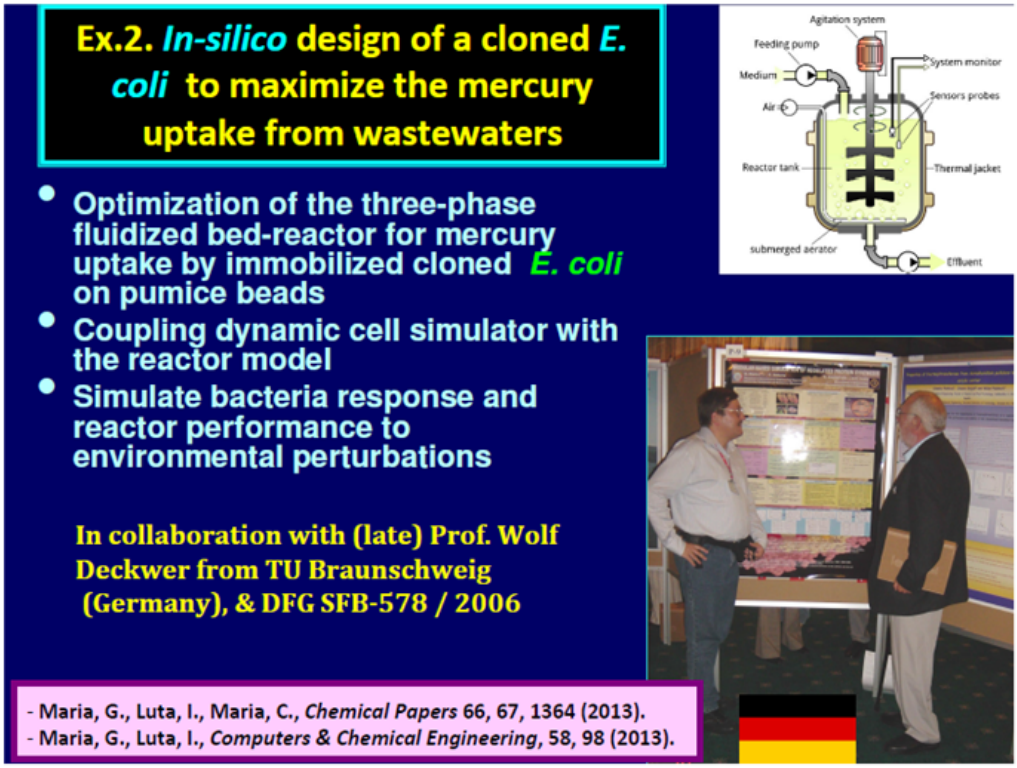


Figure 73: Case study 8.1.*In-silico* design of a cloned E. coli with a maximized capacity of mercury uptake from wastewaters [151-156].

In this industrial application, one uses a complex structured model of mercury ions reduction in *E. coli* to *in-silico* design a GMO (i.e. *E.*

coli cloned with *mer*-plasmids in a degree to be determined) for improving its capacity for mercury uptake from wastewaters (Figure 73). One worthy example of applying VVWC models to adequately represent complex modular GRC-s, is the structured model proposed by Maria [152,154] that reproduces 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 [142-144] experimental data, and the [145] information on *mer*-operon characteristics.

Bacteria resistance to mercury is one of the most studied metallic-ion uptake and release process [145] due to its immediate large-scale application for mercury removal from industrial wastewaters [146,147]. The bacteria response to the presence of toxic mercuric ions in the environment is apparently surprising; instead of building carbon- and energy-intensive disposal "devices" into the cell (like chelate-compounds) to "neutralize" the cytosolic mercury and thus maintaining a tolerable level, a simpler and more efficient defending system is used. The metallic ions are catalytically reduced to the volatile metal, less toxic and easily removable from the cell by simple 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^{2+} in cytosol, and of renewable NAD(P)H reductants able to convert it into neutral metal.

The proposed cell model by Maria [152-154] includes a GRC responsible for the involving the mer-operon expression controls the whole process, by including 4 lumped genes (denoted by GR,GT,GA,GD in Figure 74,77) 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, metabolites, proteins), thus compromising the whole cell metabolism. The GRC model includes four GERM-s of simple but effective [G(PP)1] type as follows:

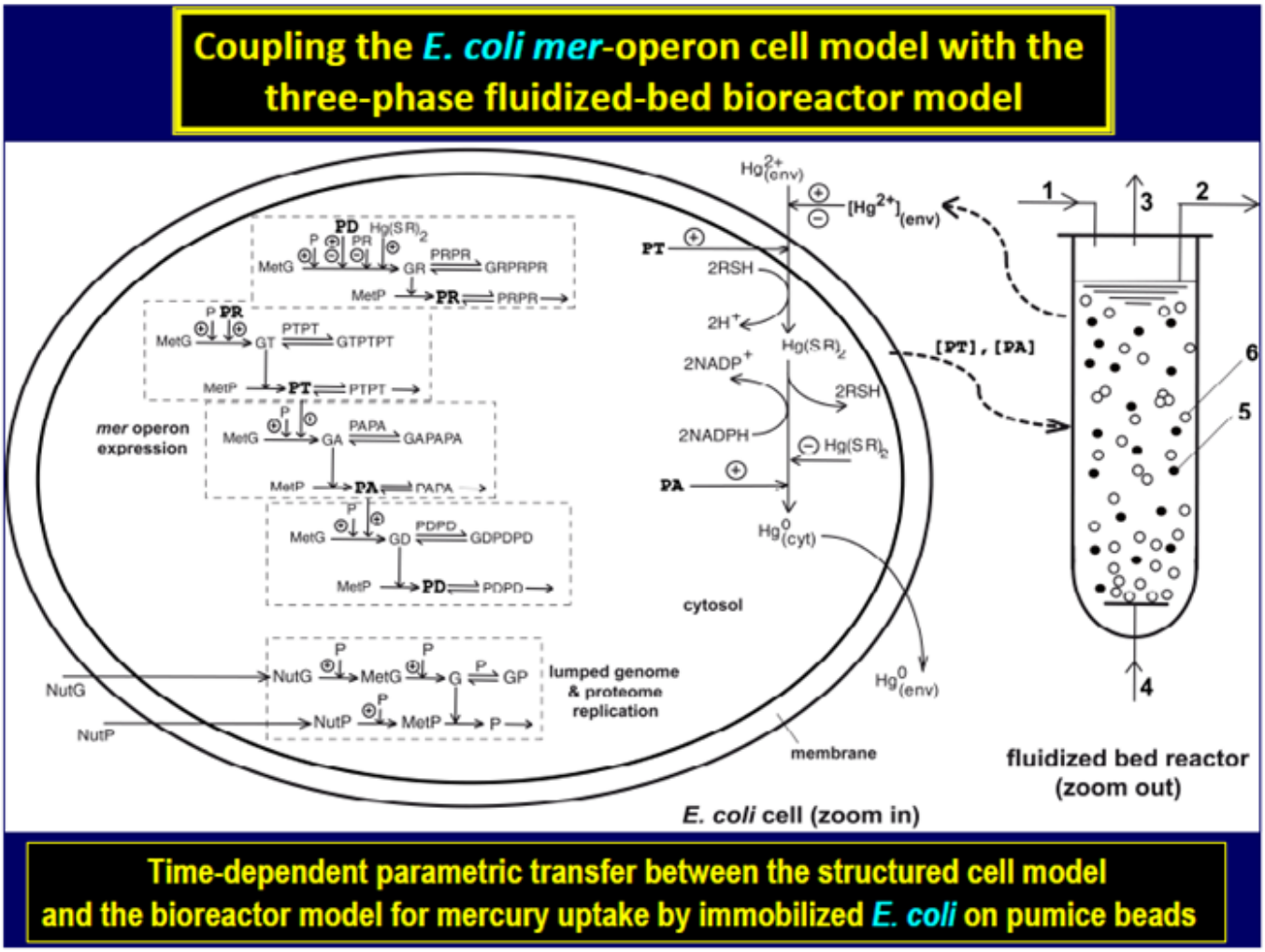


Figure 74: Case study 8.1.Reduced cell structured model, coupled to the three-phase fluidized-bed bioreactor model with suspended immobilized E. coli on alginate beads.

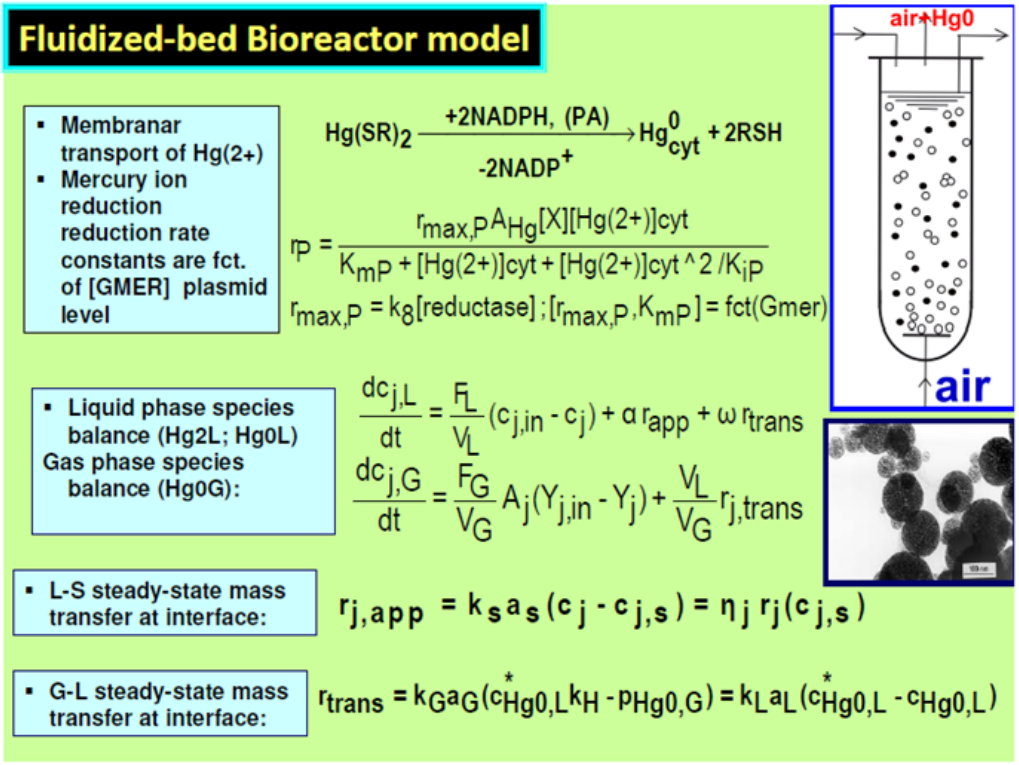


Figure 75: Case study 8.1.The macroscopic model of the three-phase fluidized-bed bioreactor with suspended immobilized E. coli on alginate beads. The Michaelis-Menten rate constants depend on the mer-plasmid level [154].

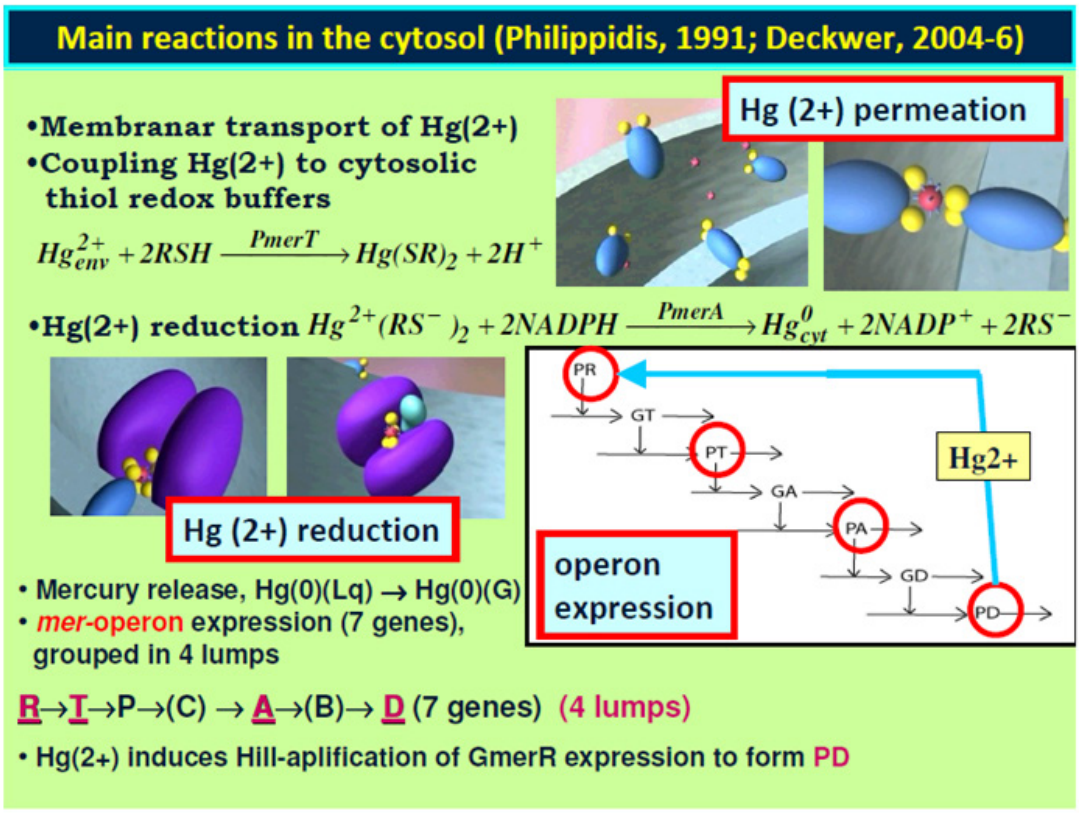


Figure 76: Case study 8.1.The nano-scopic enzymatic process in the immobilized E. coli cells: mer-operon (4 gene lumps) expression, and the main enzymatic reactions (mercury permeation, and its reduction) [154].

i) A GERM to regulate 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 [142] 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 derivative Hg(SR)2 . The GT lumped gene expression to get PT is induced by the regulatory protein PR and 'smoothed' by the 'ballast' effect of proteome lump P.

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 in nM concentrations). This GERM acts as an amplifier of the *mer*-expression leading to a quick (over ca. 30s) cell response and mer-enzyme production. The GR gene expression to get PR is controlled by the protein PD, present in small amounts in the cell.

iii) A GERM to control the expression of PA enzyme responsible for the Hg(SR)2 reduction to metallic mercury (relatively non-toxic for the cell, easily removable through membranar diffusion to be later removed from the bulk liquid phase by the continuously sparged air). The encoding GA gene expression is induced and controlled by the PT protein.

iv) A GERM controlling the protein PD synthesis. This protein has a complex role, by maintaining a certain level of GR expression even when mercury is absent in the cytosol [145].

v) a GERM controlling the replication of the lumped cell proteome (P) and genome (G) (of concentrations 107 nM and 4500nM respectively in immobilized *E. coli* cells; omics data are from Ecocyc [14], Figure 70), thus mimicking the cell 'ballast' effect on the other cell expressions and reactions. The advantage of including the cell content in the VVWC model is the possibility 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. (3-5), and the discussion of [1]}.

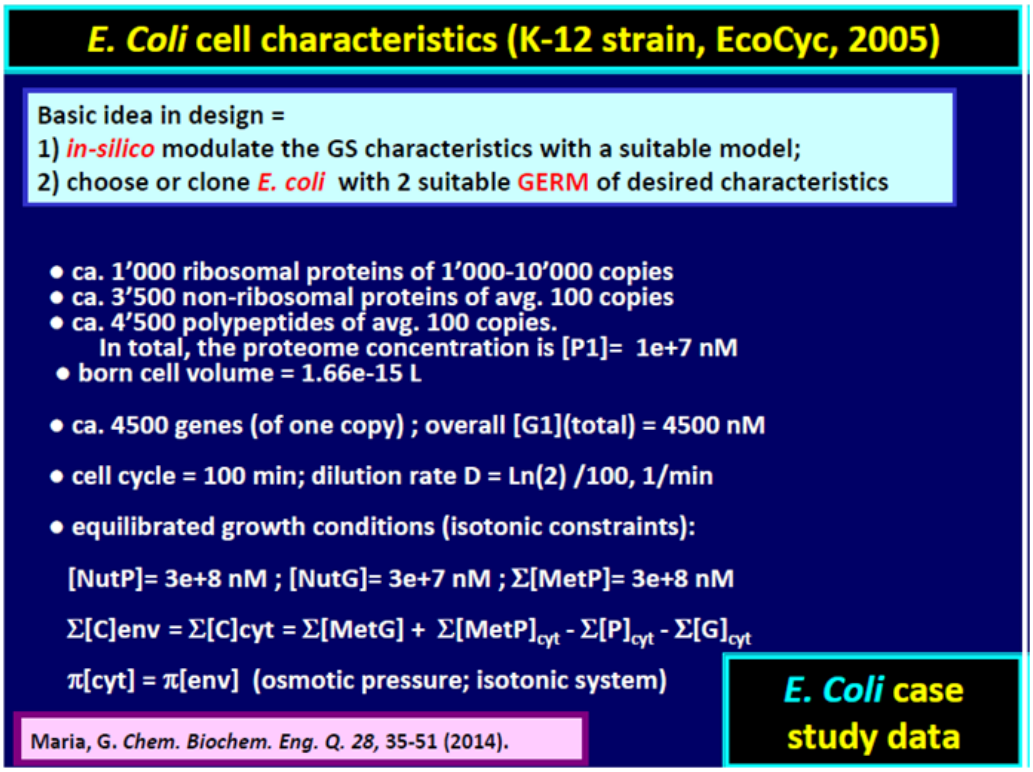


Figure 70: Case study 8.3.(continued). The characteristics of the E. coli (K-12 strain, Ecocyc, 2005) used in the case study.

In total, the *mer*-operon expression GRC model includes 26 individual or lumped cellular species and 33 reactions (Figure 76 & 77). All reactions are considered elementary, excepting some of them for which extended experimental information exists, that is [151,152,154]: i) a Michaelis-Menten rate expression for the mercuric ion permeation through the membrane into the cell; ii) a Michaelis-Menten rate expression for the mercuric ion reduction in cytosol; iii) 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. Dimmerization reactions of TF-s are considered to be much rapid than the enzyme synthesis, while equal concentrations of active (G) and inactive (GPP) forms of the generic gene G are considered

at homeostasis to maximize the GERM efficiency (Figure 76-77). The lumped proteome P, present in a large amount, is included in all gene expression rates, thus leading to more realistic evaluation of the GERM regulatory efficiency indices [56,159]. The model rate constants are estimated from solving the cell stationary mass balances for nominal concentrations of observable species, but also from optimizing the GERM regulatory indices (e.g. adjust the optimum TF level of gene expression to get the minimum recovering times after a 10% dynamic perturbation in the key species, and smallest sensitivities of the homeostatic levels vs. external perturbations). Exceptions are the Michaelis-Menten rate constants for the mercury transport and reduction in cytosol adopted from the [142-144] kinetic data. Thus, the M-M rate constants depends on the amount of Gmer plasmids in the cloned *E. coli* cells (Figure 75 & 76). Simple correlations are used to include this essential aspect in the model.

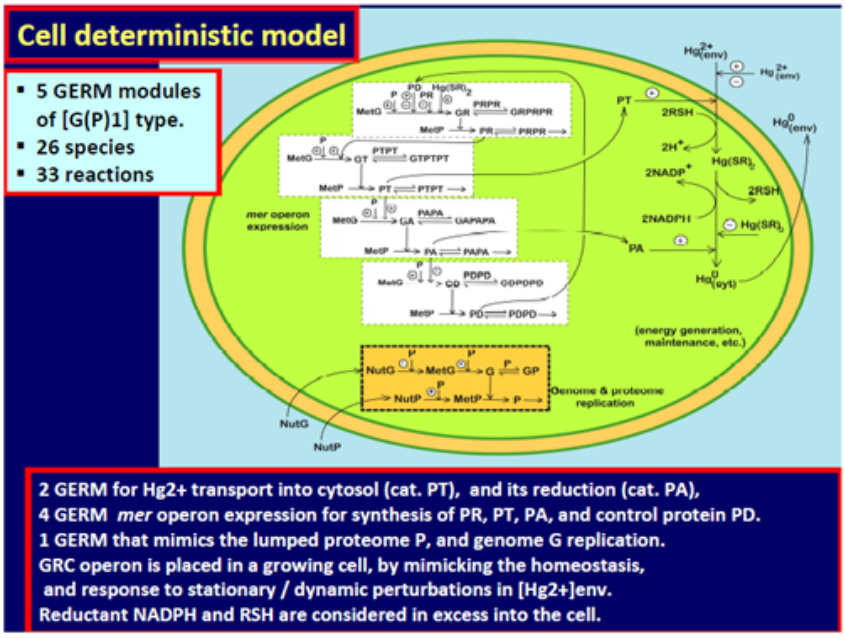


Figure 77: Case study 8.1.The reduced cell model, accounting for the mer-operon expression (4 GERMs), and the main enzymatic reactions [154].

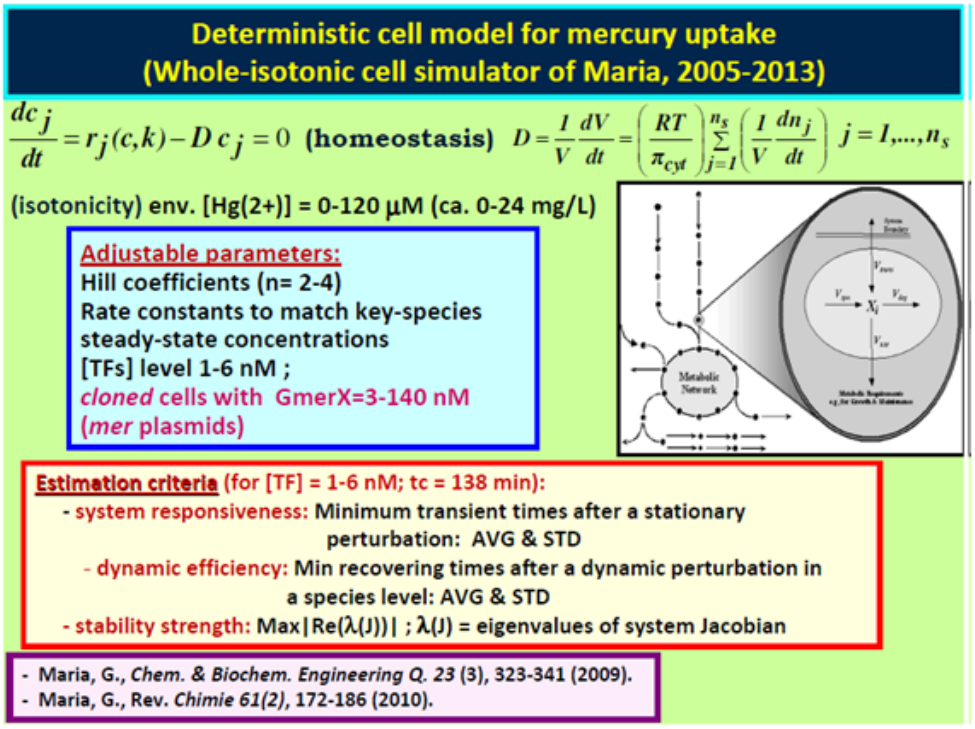


Figure 78: Case study 8.1.Details of the cell model: the differential mass balance equations, adjustable parameters, criteria used to estimate the rate constants [154]

This structured model displays a large number of advantages, being able to:

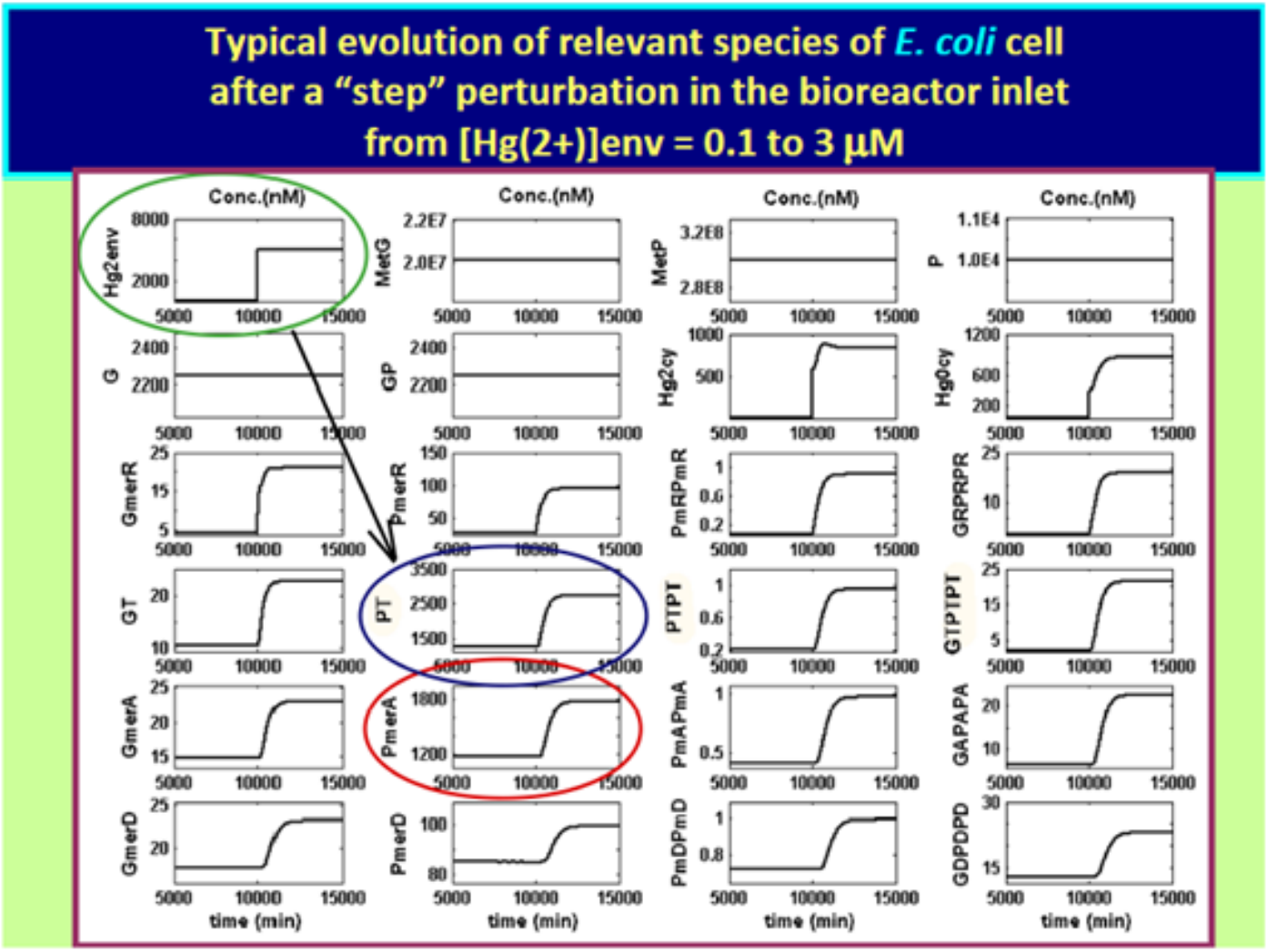


Figure 79: Case study 8.1.Prediction of the cell species concentration dynamics as a response to a "step"-like perturbation of the mercury level in the bioreactor inlet [154].

I. simulate the cell metabolism adaptation when the environmental mercury level changes. Such a reconfiguration of the levels of mer-genes and *mer*-proteins is presented in the Figure 79 as a step response after a step-perturbation in the mercury level from [[Hg(2+)env parenthesis] s = 0.1 µ M to 5mM, in a cloned *E. coli* cell with *mer*-plasmids of [plGmer]= 140nM. 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 environment perturbation is maintained.

II. Because the Hg^{2+} reduction rate constants are dependent on the *mer*-plasmid level, the cell model can predict the maximum level of *mer*-plasmids that can be added to the cell for improving its mercury uptaking rate without exhausting the internal cell resources thus putting in danger the cell survival. Thus, this cell model allows *in-silico* design of modified *E. coli* cloned with a suitable level of *mer*-plasmids to improve its possibility of cleaning wastewaters by improving the mercury uptaking capacity. As an example, in Figure 80, 81 are presented the cell key-component stationary levels, and concentration of mercury in the bioreactor bulk-phase for two GMOs: one cloned with 67 nM, and another one cloned with 140nM. Simulations of [154] revealed that as the mer-plasmids level is increasing, as the mercury uptake capacity in increasing. However, an upper limit exists (around 140nM) over which the cell resources will be exhausted, putting its metabolism in danger.

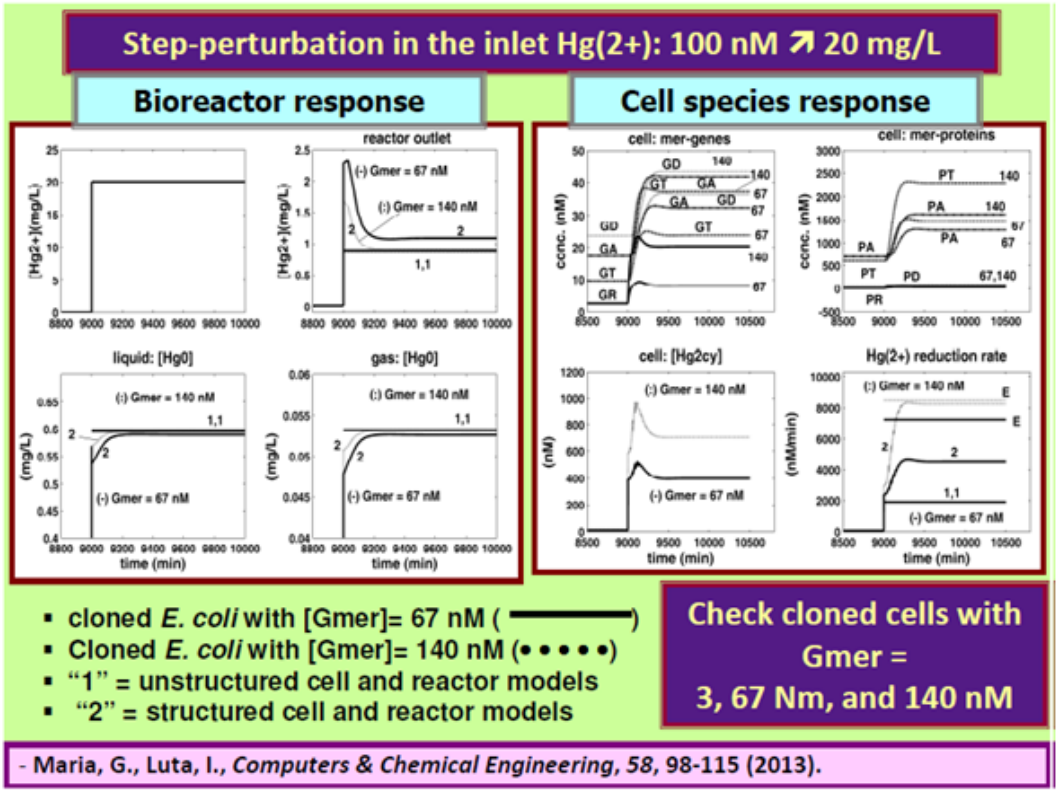


Figure 80: Case study 8.1.Prediction of the bulk-species, and of the cell species response to a "step"-like perturbation of the mercury level in the bioreactor inlet for E. coli cloned with various amounts of mer-plasmids (67 and 140nM) [154].

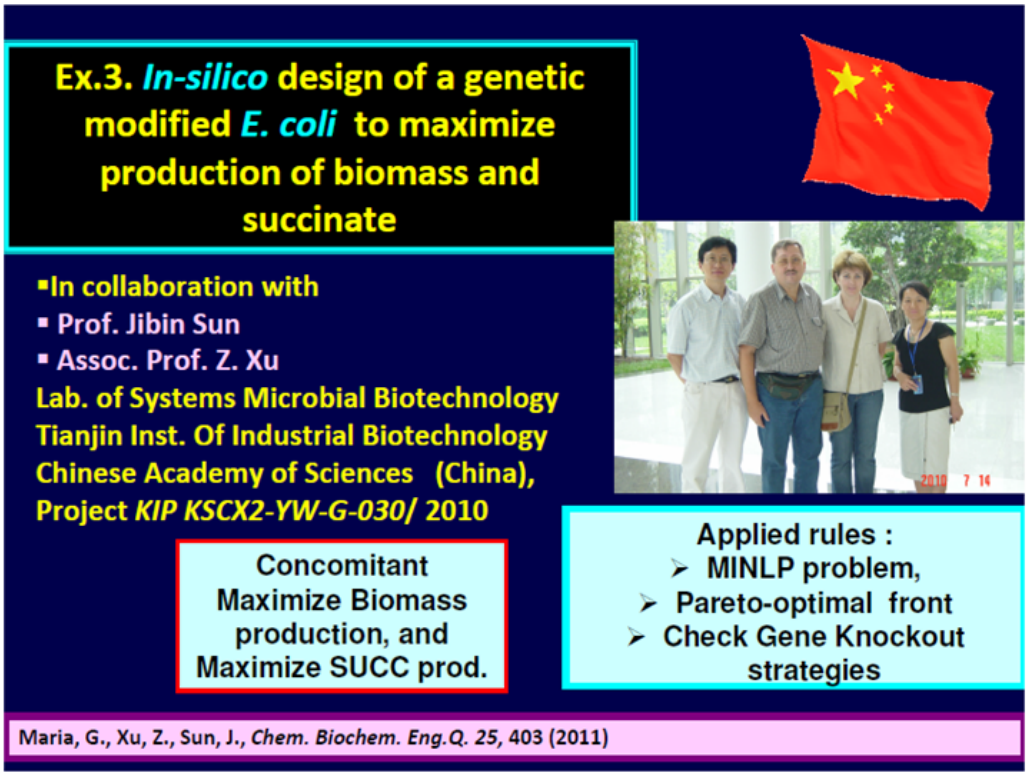


Figure 81: Case study 8.2. In-silico design of a genetic modified E. coli with a maximized capacity of succinate (SUCC) production [157].

III. By coupling the structured cell model with the three-phase continuous bioreactor model (with immobilized *E. coli* cells on alginate beads; see the bioreactor model in the Figure 75), [153,154] have been able to determine the optimal operating policies of the bioreactor in relationship to the culture of cloned cells. Such studies sre reported by Maria [153,155,156].

11.2. In-silico design of an *E. coli* bacterium by using a Pareto-optimal front methodology to get maximum production of biomass and succinate [157]

One important industrial application of GMO is those of maximizing the production of a target metabolite in industrial bioreactors. In the present case study, the problem is to use a structured cell model to *in-silico* design (by using a theoretical gene knockout strategy) an *E. coli* GMO in order to maximize the production of biomass concomitantly with maximization of the succinate (SUCC) production. Succinate is an important biosynthesis product with multiple uses in the pharmaceutical and food industry.

The thus formulated optimization problem is a multi-objective one (Figure 51-52), in the presence of the stoichiometric and flux limitation constraints. To solve it, [157] used a reduced structured CCM kinetic model of *E. coli* cell proposed by Edwards & Palsson [58] accounting for 95 reactions and 72 metabolites (see the pathway schemes in Figure 82-84) to *in-silico* determine what genes (and the corresponding reaction) should be removed (the so-called 'gene knockout' procedure) to realize maximization of both biomass and succinate production.

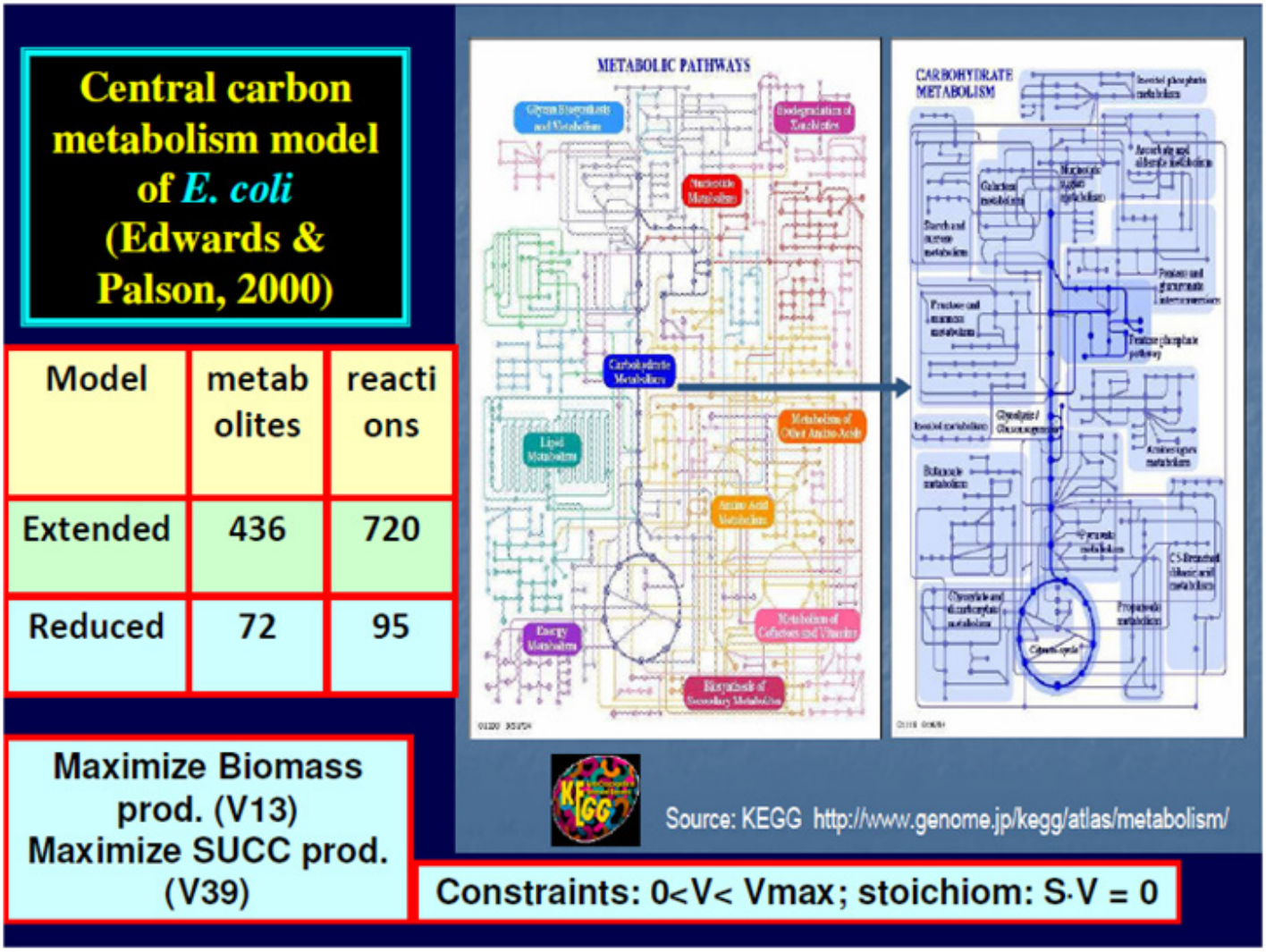


Figure 82: Case study 8.2.Carbohydrate metabolism in E. coli (KEGG [40]). The reduced model of [58] includes 72 metabolites, and 95 reactions.

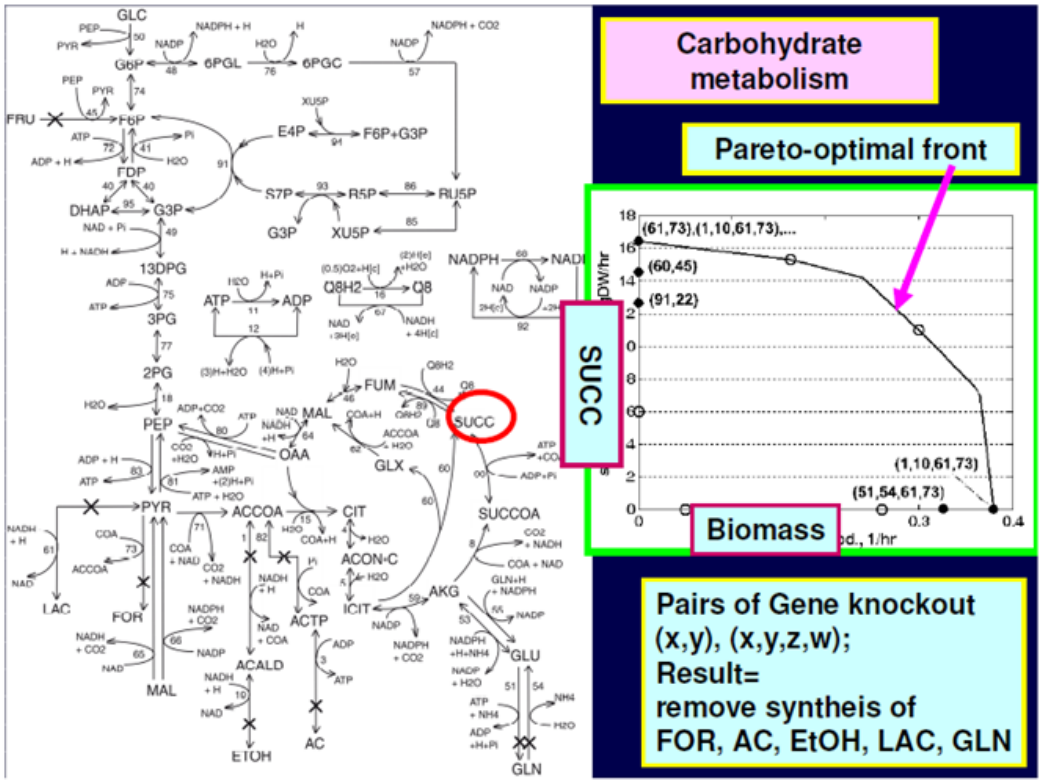


Figure 83: Case study 8.2.Carbohydrate metabolism model in E. coli used by Edwards & Palson [58]. Pareto-front of max. biomass and SUCC given by GMO with knockout genes [157].

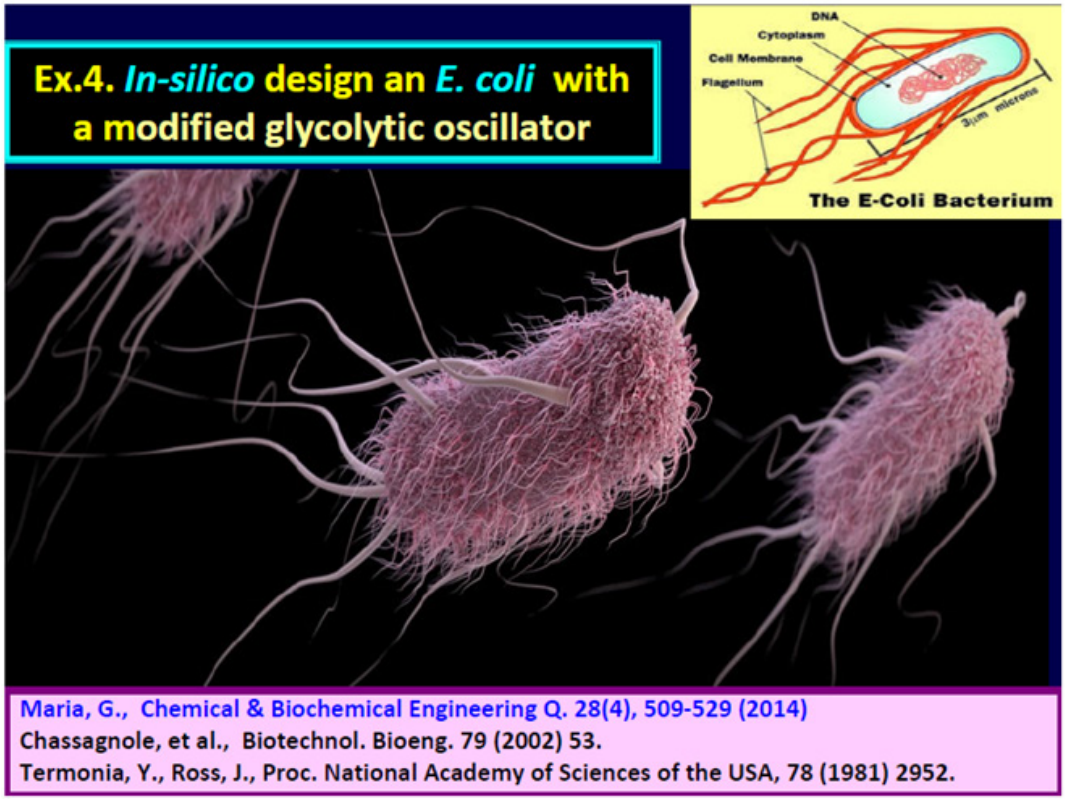


Figure 84: Case study 8.4.In-silico design of a genetic modified E. coli with a modified glycolytic oscillator [160-162].

Beside simulation of cell species dynamics, such a structured cell simulator allows identification of genome modifications leading to the improvement of the formulated optimization objectives. Simultaneous delete of several genes and reaction from the metabolic pathway of Figure 83 is also possible.

The thus formulated MINLP (mixed integer nonlinear programming problem) is difficult to be solved not only because multiple solution may exist, but the searching domain is not convex, thus reducing the possibility to get the global optimum [141].

Maria et al. [157] applied an elegant Pareto-optimal front procedure, and a genetic algorithm to get the Pareto-optimal front (see the curve of Figure 83) including the geometrical locus of all problem solutions realizing the best tradeoff between the two contrary objectives. The used multi-objective Pareto-front procedure was applied by also accounting of the stoichiometric constraints. Problem solutions indicated concomitant removal of 2-4 genes (indicated in parenthesis in Figure 83). Due to the very high complexity of the problem (dozens of hours of computing time), only a few number of alternative solutions have been checked. Results indicated that a better succinate production is obtained when synthesis of formate, acetate, lactate, ethanol, and glutamine are suppressed.

11.3. in-silico design of genetic switches (GS) of biosensor function in an E. coli bacterium, of desired properties, by using an adjustable structured model to characterize the GS (dynamic and stationary) regulatory properties [54,56,57]

A biosensor is an analytical device, used for the detection of an analyte, that combines a biological component with a physico-chemical detector [148,149].

The principle (Figure 67) of the biosensor is the following:

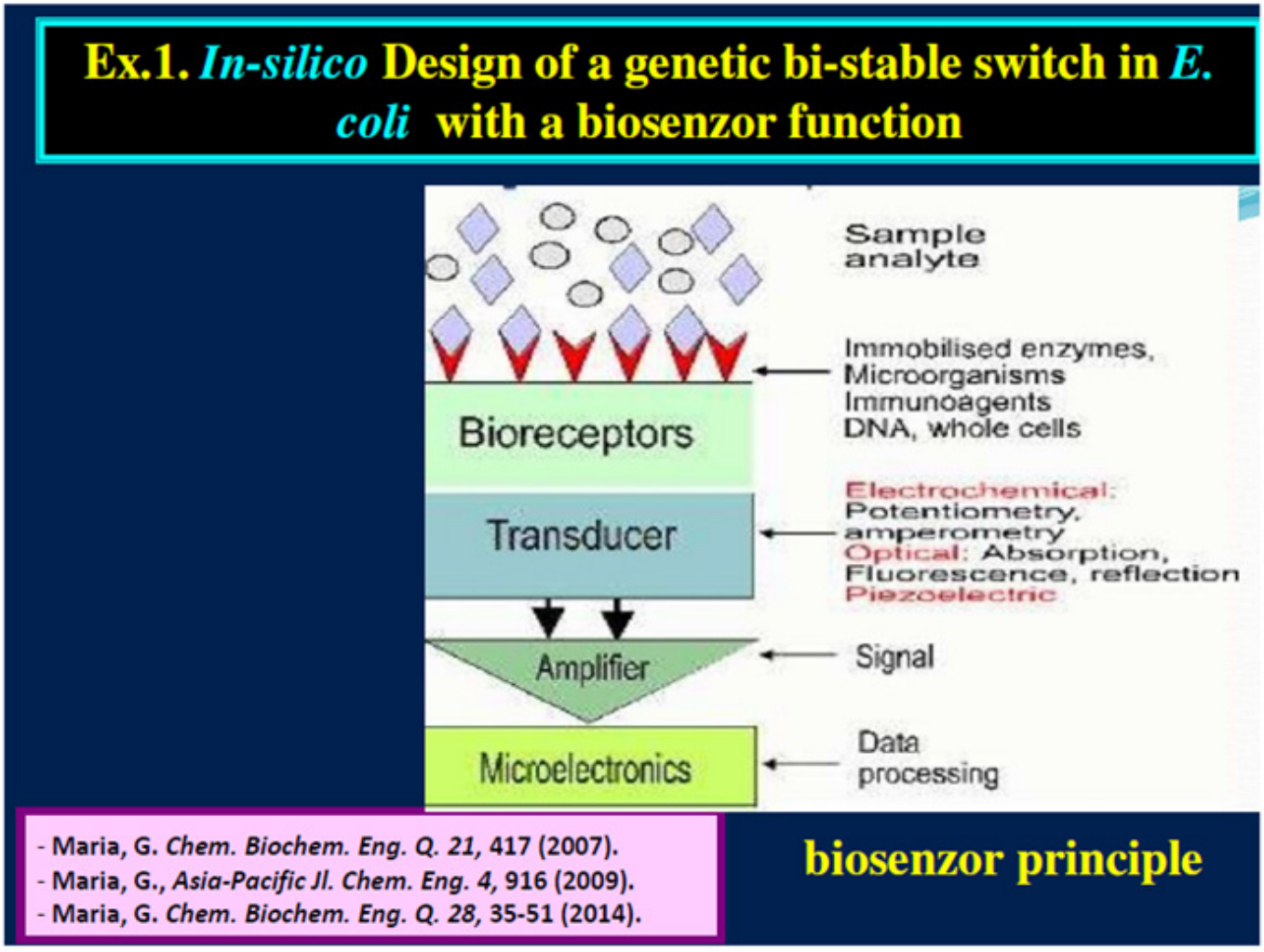


Figure 67: Case study 8.3.In-silico design of a genetic switch in E. coli with the role of a biosensor [54,56,57].

(i) The sensitive biological element (e.g. tissue, microorganisms, cell organelles, cell receptors, enzymes, antibodies, nucleic acids, etc.), is a biologically derived material or biomimetic component that interacts, binds, or recognizes with the analyte (inducer) under study. This bioreceptor is usually located on the cell membrane. The biologically sensitive elements can also be created by biological engineering.

(ii) The transducer or the detector element, which transforms one signal into another one, works in a physico-chemical way: optical, piezoelectric, electrochemical, electro-chemi-luminescence, etc., resulting from the interaction of the analyte with the biological element, to easily measure and quantify. The biosensor reader device with the associated electronics or signal processors that are primarily responsible for the display of the results in a user-friendly way [150]. This sometimes accounts for the most expensive part of the sensor device, however it is possible to generate a user friendly display that includes transducer and sensitive element (holographic sensor). The readers are usually custom-designed and manufactured to suit the different working principles of biosensors.

(iii) A very valuable possibility to construct a cellular biosensor is those of design a genetic-switch (GS) easily to be modulated by gene engineering.

A GS (Figure 68&69,85) consists in two GERM modules, with a mutual repression control. Such a self- and mutual control in two gene expression modules, allows creation of decision-making branch points between on/off states according to the presence of certain inducers. As displayed in the Figure 68, in the presence of the internal inducer I2 (generated by the external inducer NutI2), the GERM G2/P2 is over-expressed, while the GERM G3/P3 is repressed (Figure 69). As a consequence, the concentration of P2 increases very much (being easily to be detected), while P3 concentration decreases very much thus remaining undetectable (Figure 68). Practically, the GS displays two stationary stable states (Figure 98): one with the protein P2 in a large amount and P3 in a small amount, or vice-versa.

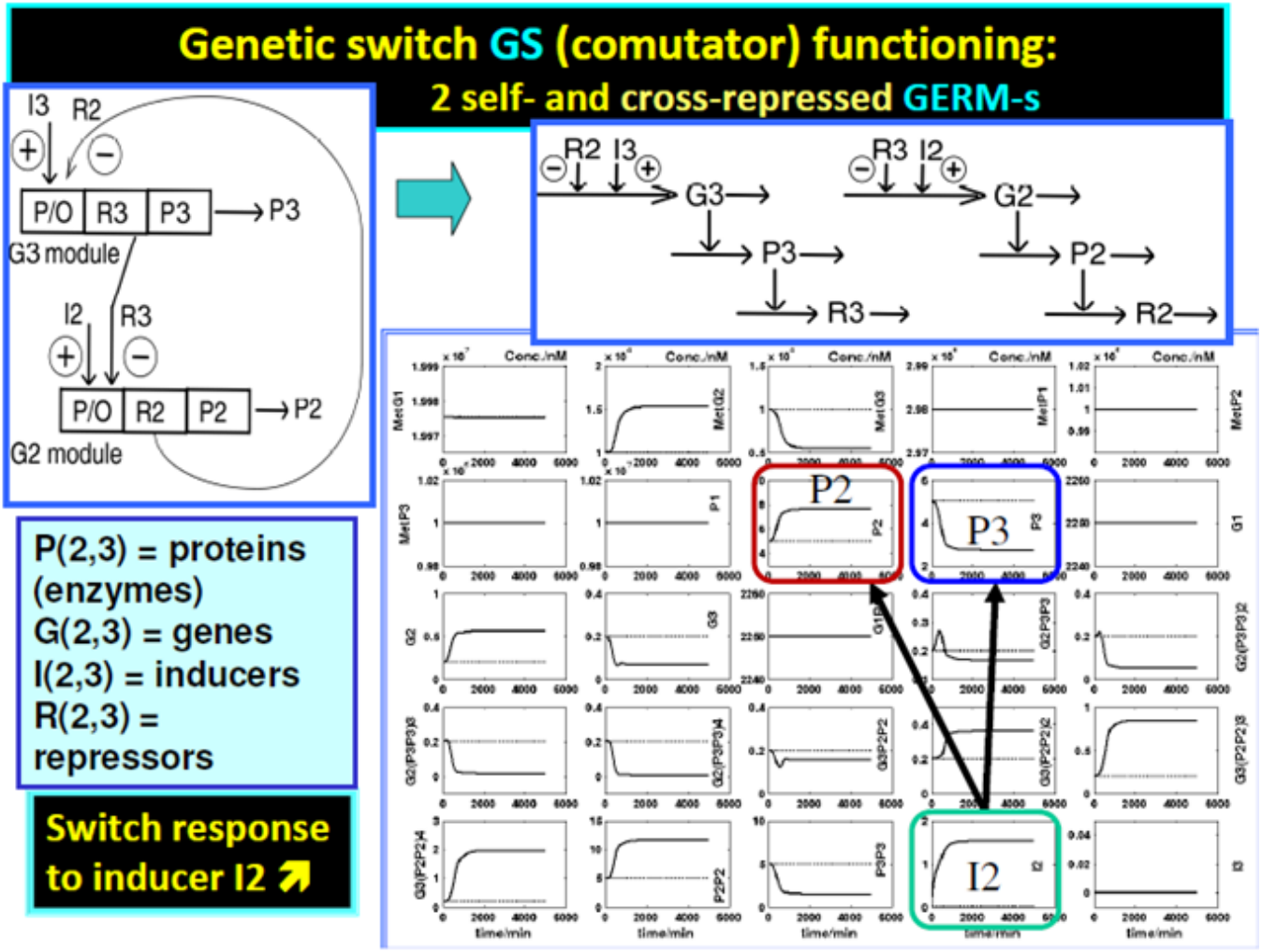


Figure 68: Case study 8.3. (continued). The principle of a genetic switch: two self- and cross-repressed GERM-s, [54, 56-57].

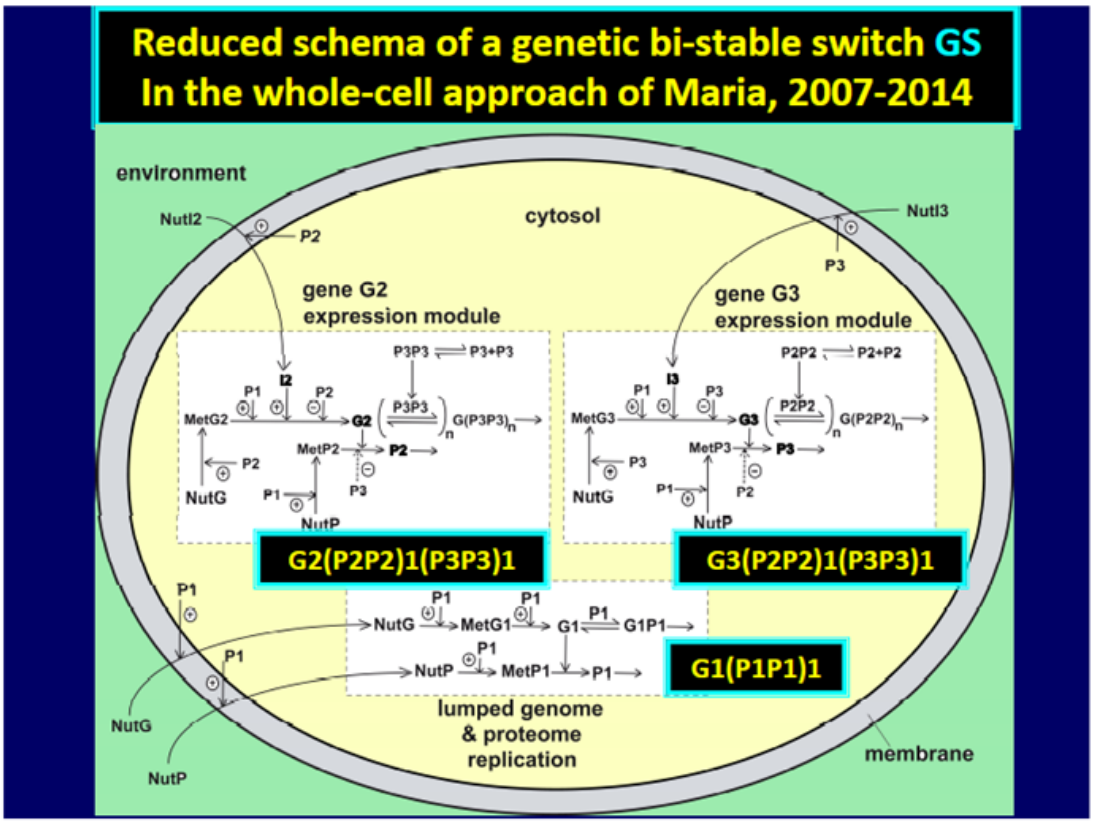


Figure 69: Case study 8.3. (continued). Reduced modular VVWC genetic switch model of [57].

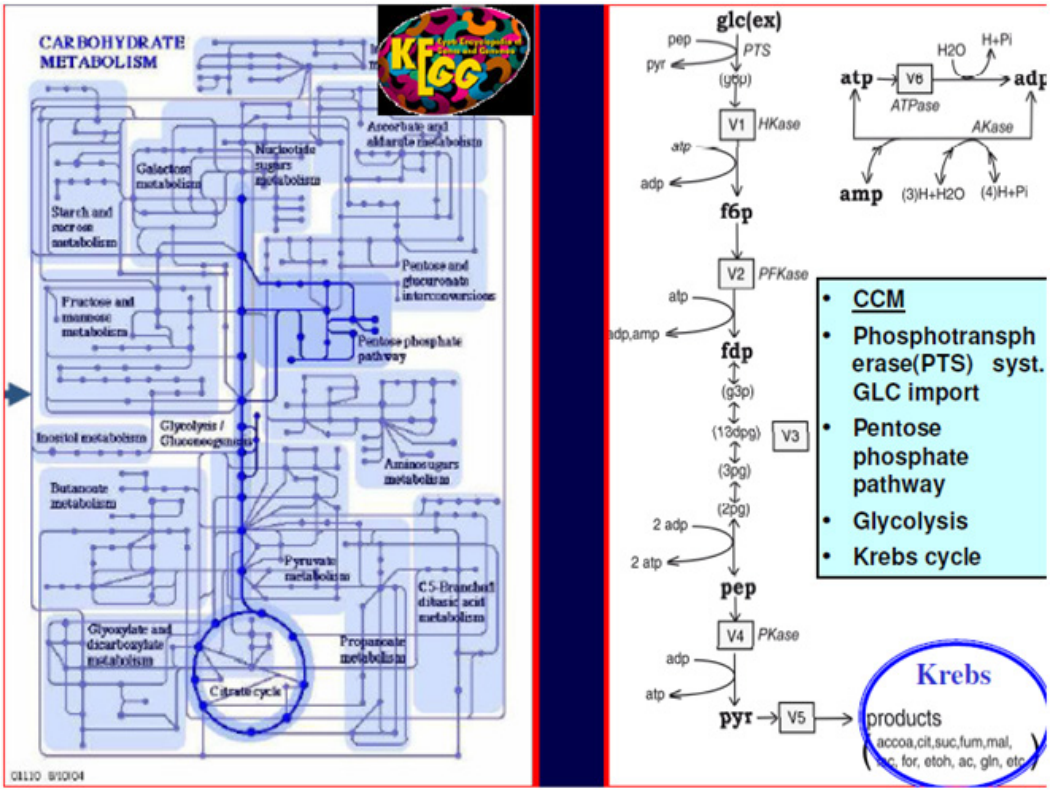


Figure 85: Case study 8.4.The reduced reaction pathway of glycolysis in E. coli used by Maria [160] to simulate the glycolytic oscillations.

In other words (Figure 97), the bistable synthetic gene-regulatory network consists of two promoters and two repressors [167]. The second operon is composed of the promoter 2, which is the repressor for promoter 1 called repressor 1 and a reporter GFP (green fluorescent protein). Therefore the GFP reports for the expression level of repressor 1.

In the absence of inducer there are two stable states possible (Figure 98):

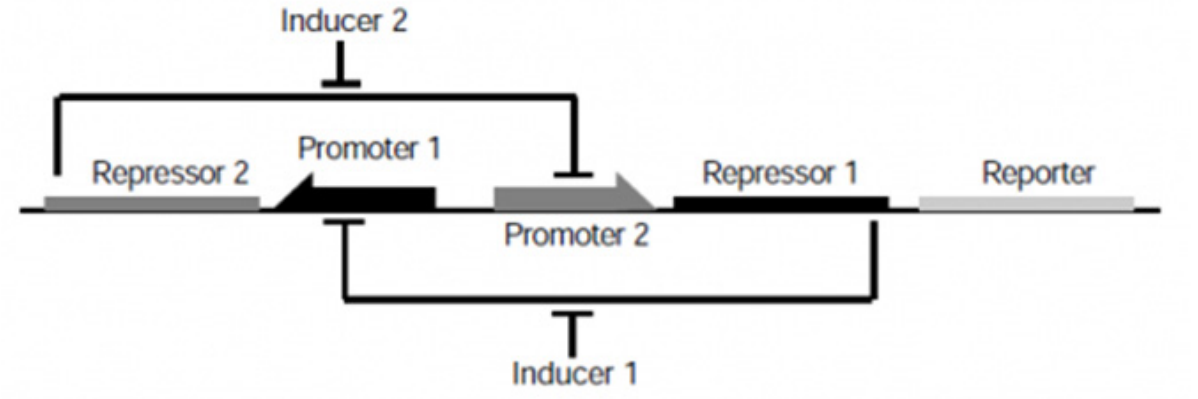


Figure 97: Case study 8.3.The principle of a genetic switch (GS) [148]. Construction of a genetic toggle switch in E. coli" [167].

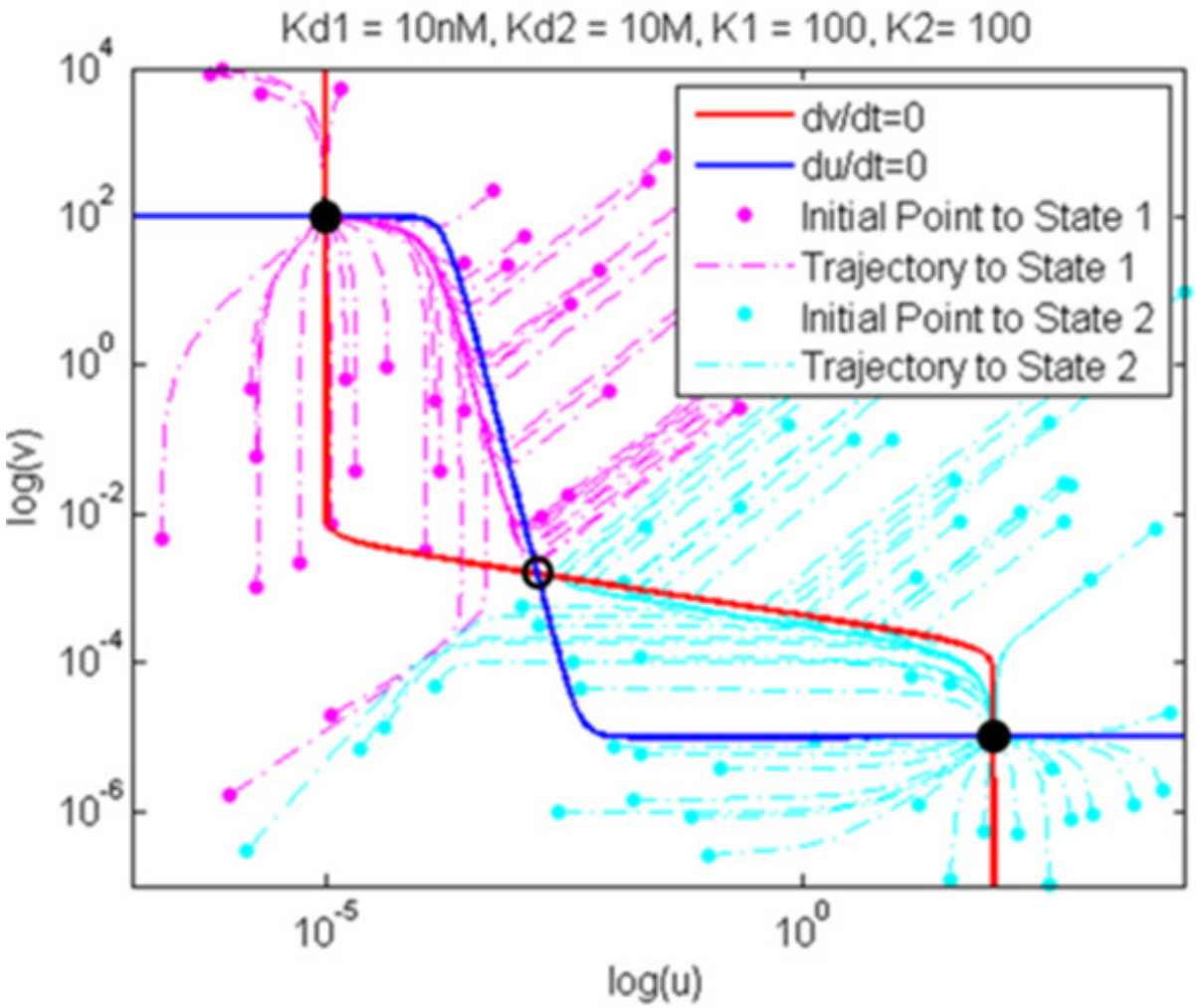


Figure 98: Case study 8.3.The two stationary-states of a genetic switch (GS) (Gardner et al. [167]); (http://2013.igem.org/Team:Duke/Modeling/Kinetic_Model).

a. 'low' state: repress or 2 is transcribed and repressor 1 is repressed.

b. 'high' state: repressor 1 (+GFP) is transcribed and repressor 2 is repressed.

Assuming the switch is in a stable position, addition of inducer for the repressed repressor causes full activation of transcription, until the originally active repressor is repressed. In this manner the inducers allow to switch between the two stable states.

If a good GS model is available, the GS properties can be easily *in-silico* modulated [54,56,57]. Then, there are two possibilities to realized the GMO with including the GS of desirable properties: I) choose from the same cell two GERMs that interact in the desirable way; ii) if not, modify the cell genome by inserting two GERMs taken from another organism, where they behave as a desirable GS.

The design GS presents potential applications in medicine (such as therapy of diseases, that is gene therapy), or in industry, to design new devices based on cell-cell communicators, biosensors, production of vaccines, etc. [158,159,165].

In the present paragraph such a valuable mechanistic GS model [54,56,57] is presented, to be used for further *in-silico* design of GMO with biosensor function.

Genetic switches (GS) are particularly attractive because such a toggle-switch realizes the mutual repression control in two gene expression modules that creates decision-making branch points between on/off states according to the presence of certain external inducers (Figure 68). In fact, GS re-direct the cell metabolism to better adapt to environmental changes. [54,56,57,90,117,138,139]; presented various principles to construct GS models. However, in contrast to the large number of CVWC models of GS from literature [54,56,57], that use apparent kinetic constructions (including Hill or power-law type nonlinear induction/repression models), [5456,57] used mechanistic VVWC models to adequately model GSs of adjustable stationary and dynamic P.I.-s, including the switch efficiency (that is GS certainty, QSS stability, GS sensitivity to inducers, response rate, transition time to another QSS) by means of the no. and type of the included GERM effectors, TF-levels, reaction rate orders of self- and cross-repression (Figure 69,71,72).

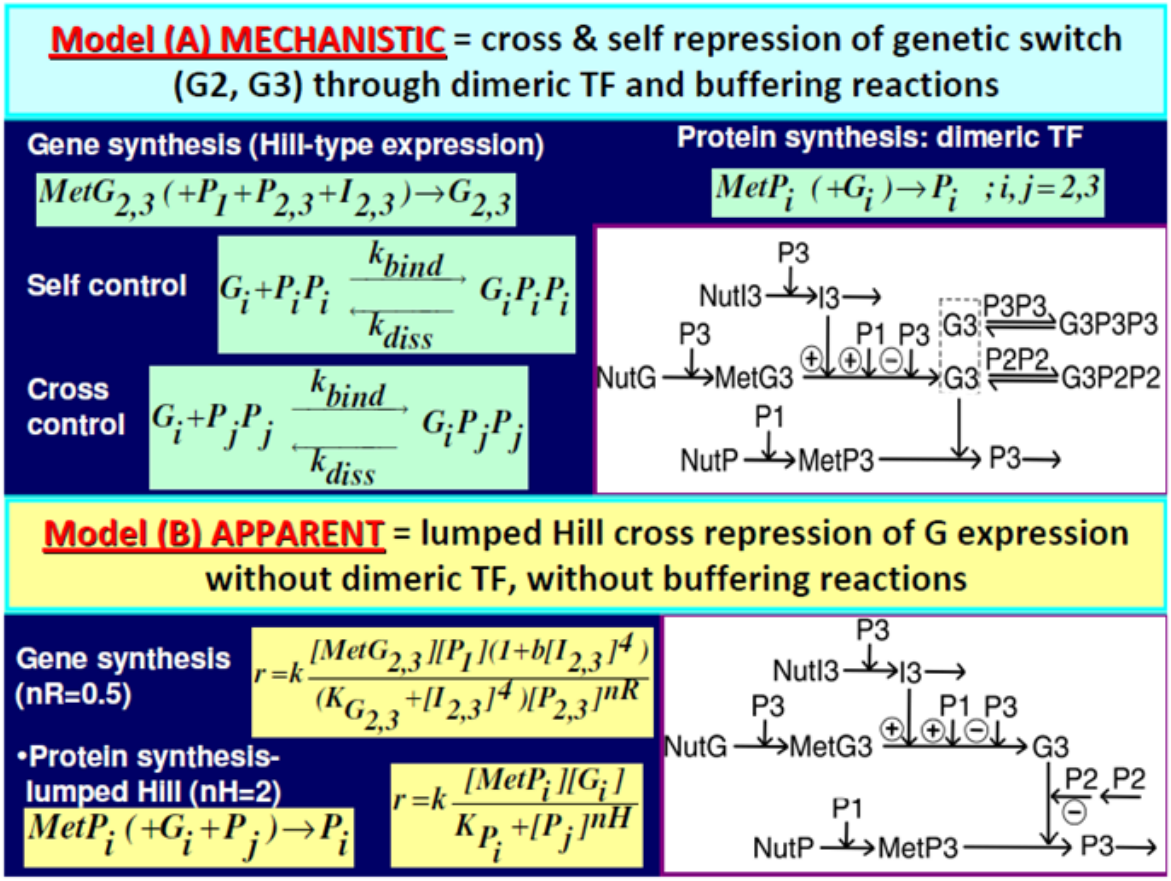


Figure 71: Case study 8.3.(continued). The two models (A, deterministic), and (B, apparent Hill-type) used to represent the genetic switch [57].

Model (A) is superior being easily adaptable

Model	[TF] level (nM)	Stability strength	Stationary efficiency	Switch responsiveness		Dynamic efficiency	
		Max |λ(J)| /1e+5	|d Log[P2,3] / d Log[Nutl2] | / 1e-14	Species transient times (min)		Species recovering times	
				P3	Avg.	P2	Avg.
(B)	-	2.0	47.6	4502	2191	803	224
(A)	1	5.4	8.9	1727	1120	NG	2.6
	7	3.2	362	923	903	NG	2.6
	15	3.1	337	662	865	NG	1.9

Modulate GS characteristics by choosing the right GERM modules and modify TF level
- Good QSS stability (decrease with the TF ↗ and buffer complexity)
- Increased stationary efficiency & responsiveness (best nHill=2; nR= 0.5)
- GS dynamic efficiency (increase with TF, decrease with the complexity of the repression control schema *nHill (nH), n_Self_Rep (nR)*
- Increased GS certainty through:
  - Hill coeficient of Protein synthesis cross-repression (nH↘)
  - Hill coeficient of Gene synthesis self-repression (nR↗) ; [TF] (↘)

Figure 72: Case study 8.3.-(continued). Model (A) superiority in reproducing adjustable switch properties, for both dynamic and stationary regulatory efficiency [57].

It is to be observed that a VVWC model requires that all cell species (individual or lumped) to be considered in the model mass balance (including the lumped genome, proteome, and metabolome), because all these lumps contribute to the volume and dynamics via the isotonicity constraint (eq. 3-6). As proved by Maria [56,57] such an approach offers more realistic P.I-s. of the GS, , in spite of a higher computational effort. The dynamic and stationary P.I-s. are easily adjusted by a suitable selection of the two GERM-s of the GS, and optimization of the GS model parameters such as the no. of buffering reactions, no. of TF per gene operator site, self-repression partial order, TF level and type, Hill-type induction/repression nonlinearity of some lumped steps).

To not complicate the GS model, [G(P)1] lumped GERM-s were adopted for the genome, and proteome replication in the cell. Approximate concentration of lumped genome and proteome can be easily estimated for a large number of micro-organisms by using omics databanks such as [40]. The cell 'ballast' is accounted as a lumped metabolome. The scheme of such GS model includes two [G(P)1] lumped GERM-s. Two GS models have been tested to represent a GS in an easily adjustable way (Figure 71). The mechanistic model (A) includes only a certain number of buffering reactions that quickly adjust the genes activity. By contrast, the lumped (apparent) model (B) includes the cross- and self-repression and fast induction of the two genes G2 and G3 expression, by using two GERMs with lumped apparent rate expressions, that is [P1(P2)(-0.5)(I2)(Hill); G2(P3)(Hill)]+[P1(P3)(-0.5)(I3)(Hill); G3(P2)(Hill)] (eq. 8). The GS scheme is given in the Figure 68-69. The model (B) includes only pseudo-elementary reactions, except those for the switch genes G2 & G3 and proteins P2 & P3 production, for which the apparent Hill-type kinetics of (Voit, 2005[63]) has been adopted, that is:

$$\frac{dc_{G2}}{dt} = k_{G2} c_{MetG2} c_{P1} \left(\frac{1 + B c_{I2}^4}{C_1 + c_{I2}^4} \right) c_{P2}^{-0.5} - D\, c_{G2} \;;\; \frac{dc_{P2}}{dt} = k_{P2} c_{MetP2} c_{G2} \left(\frac{1}{C_2 + c_{P3}^2} \right) - D\, c_{P2} \;;$$

$$\frac{dc_{G3}}{dt} = k_{G3} c_{MetG3} c_{P1} \left(\frac{1 + B c_{I3}^4}{C_3 + c_{I3}^4} \right) c_{P3}^{-0.5} - D\, c_{G3} \;;\; \frac{dc_{P3}}{dt} = k_{P3} c_{MetP3} c_{G3} \left(\frac{1}{C_4 + c_{P2}^2} \right) - D\, c_{P3} \;; \quad (8)$$

where G2 & G3 are the expressed genes, P2 & P3 are the corresponding encoding proteins, while I2 & I3 are the inducers activating the expression of G2 & G3 genes. The GS model (A) of best performances includes (n=4) cross-buffering reactions G2(P3P3)4 and G3(P2P2)4 (-72). The derived GS kinetic model accounts for 19 individual and lumped species and includes 21 rate constants identified from solving the stationary model form with known species stationary concentration, and imposing P.I. optimization [57]. Simulation results of Figure 72 indicated that the model (A) based on the GS reaction mechanism is superior to the lumped model (B), being much easier to be used in *in-silico* analyses. That is because their properties can be easily adjusted by a fewer number of parameters of physical significance, that is: the number "n" of buffering reaction modulating the genes' activity, and the concentration of the dimeric TFs.

Besides, by considering all cell individual or lumped species in the **VVWC** model, the dilution rate is uniform for all species, while the degradative steps have been neglected. The gene G2/G3 expression activation accounts for four molecules (n=4) of inducer, allosterically binding to the promoter site, while a slow self-repression with the product is considered of an (adjustable) -0.5 apparent reaction order. The cross-repression of protein P2/P3 synthesis, of Hill type, accounts for only dimeric repressors (n=2) allosterically binding to the catalytic gene, even if a higher control (with tetramers) has been reported [115,116].

A typical simulation of the individual / lumped GS species response in *E. coli* is plotted in Figure 68 after a "step" perturbation in the environmental stimulus I2 from 0 to 1nM. Predictions are generated by simulation over tenths of cell cycles by using the model with four cross-repressing reactions of {G2,G3} expression, with optimised nR= 3 self-repression exponent and [TF]= 5nM. As expected, species present in large amounts (of order 10^5-10^8nM) display a negligible response to the I2 small perturbation (of 1nM), while the cellular species directly connected to the NutI2/I2 inducer pathway are very strongly affected. Even if the plots of Figure 68 are represented for a large time-scale (thousands of minutes), the transient times for recovering the homeostasis are from the order of minutes up to few cell-cycles (hundreds of minutes), also revealed by Elowitz & Leibler [168] discussion on the transmission of the effect of certain perturbations from cell generation to generation as an expression of cell adaptation to the environment.

As discussed by Maria [54,56,57] such rather mechanistic VVWC models with explicit buffering reactions, by offering better P.I. predictions, seem to be more flexible and versatile, thus being more suitable for *in-silico* design of a GS with desired properties..

11.4. Design of an E.coli bacterium with a modulated glycolytic oscillator [160-162]

In the present paragraph one presents a structured reduced model of the oscillating glycolysis in *E. coli*, very suitable for the *in-silico* design of GMO with modulated glycolytic oscillations. Modulation of the glycolytic oscillations might be important as long as glycolysis is interfering with lot of metabolic processes, and various metabolites synthesis of practical interest (e.g. succinate, citrate, amino-acids, etc.)(Figure 84).

Glycolysis is an essential step of the CCM, being the first stage of the glucose substrate import into the cell, and its usage. In short, glycolysis is the metabolic pathway that converts glucose ($C_6H_{12}O_6$) into pyruvate ($CH_3COCOO-+H^+$)(see its reduced schemes in Figure 87. The free energy released by the subsequent tricarboxylic acid cycle (TCA) metabolic pathway originating from pyruvate is used to form the high-energy molecules of ATP, and NADH that support the glycolysis and a lot of the enzymatic cell syntheses [175].

To *in-silico* design GMO it is indispensable to dispose of a valuable whole-cell simulator of the central carbon metabolism (CCM) (Figure 83,85). The CCM [15,40] includes: the phosphotranspherase (PTS) system for the glucose (Glc) membranar import into the cell; the glycolysis (transformation of Glc in pyruvate Pyr); the pentose phosphate pathway (PPP, which is a metabolic pathway taking place in parallel to the glycolysis; it generates the co-factor NADPH and pentoses as well as ribose-5-phosphate, which is one of the precursors for the synthesis of nucleotides); the tricarboxylic acid (TCA, or Krebs, Figure 85) cycle, which is a series of tightly controlled enzymatic reactions used by all aerobic organisms to release stored energy through the oxidation of acetyl- coenzyme A derived from carbohydrates, fats, and proteins into CO2 and chemical energy in the form of adenosine triphosphate (ATP). In addition, the cycle provides precursors of certain amino acids, as well as the reducing agent NADH, that are used in numerous other biochemical reactions. Its central importance to many biochemical pathways suggests that it was one of the earliest established components of the cellular metabolism and may have originated abiogenically [169].

In this context, one of the most studied module of the CCM is the *glycolysis*. Modelling bacteria glycolysis dynamics is a classical subject but still of high interest, allowing in silico design of GMO with desirable 'motifs' of practical applications in the biosynthesis industry, environmental engineering, and medicine. By using a reduced deterministic kinetic model (Figure 86&87), obtained from reducing the [59] model (Figure 87, left), [160] proposed a reduced glycolysis model able to predict the conditions [161,162] leading to the occurrence of a stable oscillating glycolysis in the *E. coli* cells (experimentally highlighted by Madsen et al, [170] (Figure 88). Even if the reduced model of [160] includes only 9 species and 6 reactions (Figure 86) it has been proved to be able to fairly simulate the stationary or oscillating glycolysis in Escherichia coli bacterium.

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 86: Case study 8.4. A review of available dynamic models in literature for glycolysis, with specification of the accounted number of species, reactions, and parameters [160).

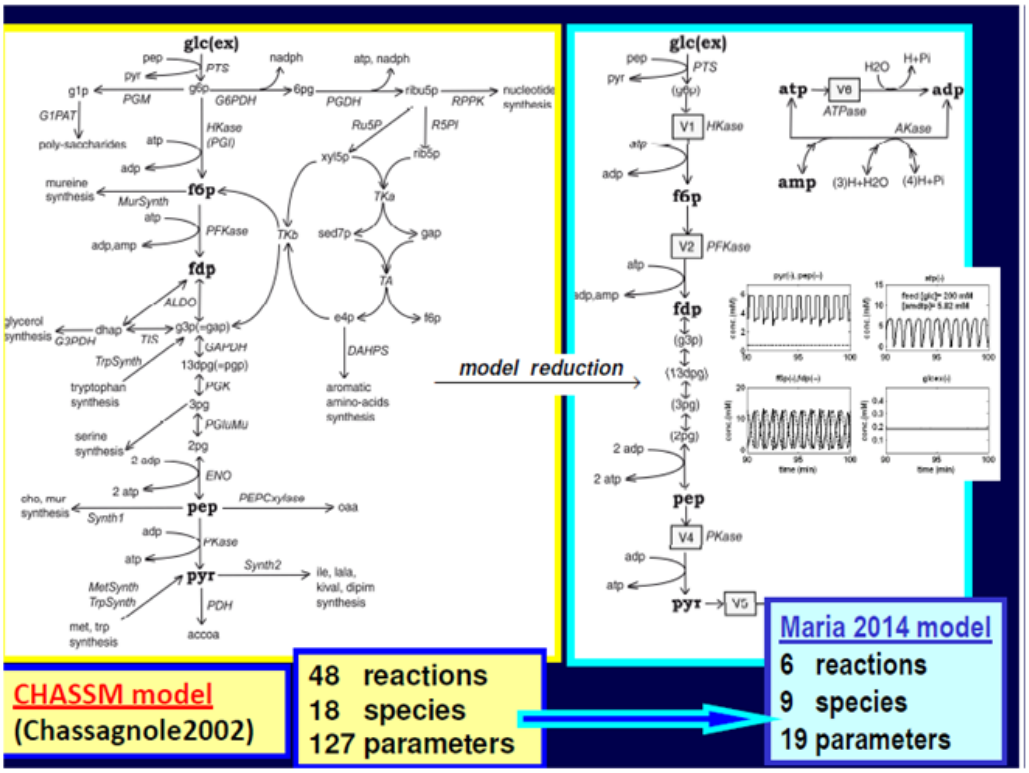


Figure 87: Case study 8.4.A comparison of the glycolysis models of [59], and of (Maria, 2014[160]). The last one allows simulating glycolytic oscillations.

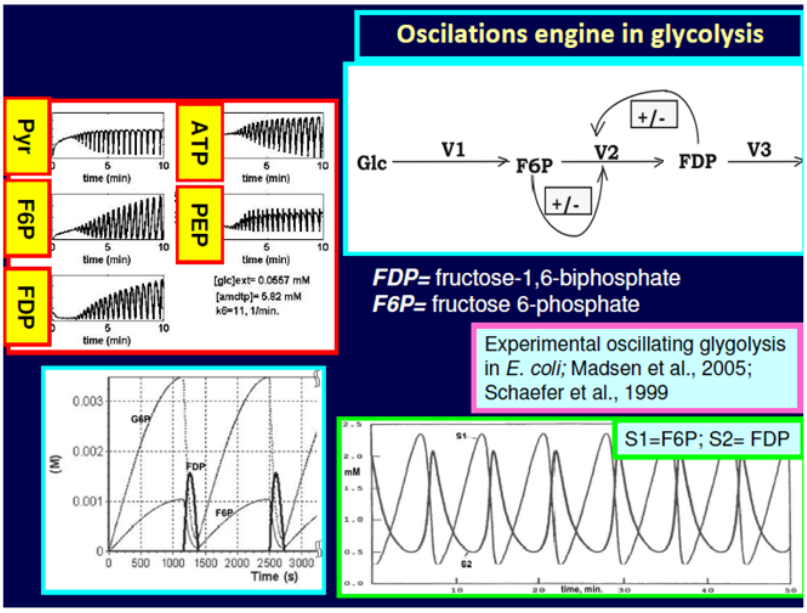


Figure 88: Case study 8.4. Explanation of the regulatory loops simultaneously acting on the reaction V2, that produce glycolytic oscillations [160-162].

Autonomous oscillations of the glycolytic intermediates' concentrations reflect the dynamics of control and regulation of this major catabolic pathway, and the phenomenon has been reported in a broad range of cell types [170]. Understanding glycolytic oscillations might therefore prove crucial for our general understanding of the regulation of metabolism and the interplay among different parts of metabolism as illustrated, for instance, by the hypothesis that glycolytic oscillations play a role in complex pulsatile insulin secretion. The key question in this context is the mechanism(s) of the oscillations but, despite much work over the last 40 years, it remains unsettled.

According to Franck [171], spontaneous occurrence of self-sustained oscillations in chemical systems is due the coupled actions of at least two simultaneous processes (Figure 88&89). Oscillations sourced in a so-called "oscillation node" (that is a chemical species, or a reaction), on which concomitant rapid positive (perturbing) and slow negative (recovering) regulatory loops act. Because the coupling action between the simultaneous processes is mutual, the total coupling effect actually forms closed feedback loops for each kinetic variable involved. There exists a well-established set of essential thermodynamic and kinetics prerequisites for the occurrence of spontaneous oscillations.

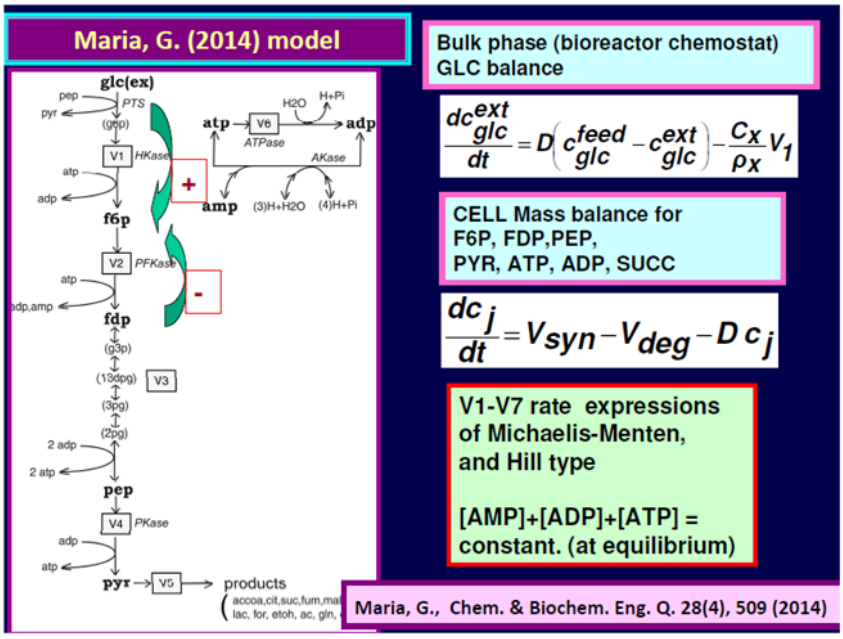


Figure 89: Case study 8.4.The differential mass balance of the [160] kinetic model of glycolysis.

In the glycolytic system case, extensive experiments (Figure 88, right plot, Figure 89) [160,161,162] revealed that self-sustained oscillations are reported in a broad range of cell types. As revealed by [172,173] glycolytic oscillations occurrence is due to the antagonistic action of two processes on regulating the V2 reaction rate that converts F6P in FDP (see reaction scheme in the Figure 88-left, and Figure 89). The glycolytic oscillation occurrence and characteristics (period) are influenced by both external (environmental) and internal (genomic) factors, that is [160-162]:

(i) From one side it is the glucose (Glc) import driving force through the phosphotransferase (PTS)-system (Figure 87,89) regulated and triggered by the external concentration of glucose [Glc]ext and by the PEP and PYR levels;

(ii) However, the Glc import and conversion to PYR requires important amounts of regenerable ATP, and an enough rapid ATP to ADP conversion rate, as well as its quick regeneration;

(iii) On the other hand, limited A(MDT)P cell energy resources exist in the cell, which can slow-down the Glc import if the ATP use/regeneration is not working fast enough.

To *in-silico* modulate the glycolysis oscillations in an *E. coli* cell culture from a chemostat [161,162] used two types of control parameters:

(i) bioreactor operating parameters, such as: the glucose level in the bulk-phase, [Glc]ext (related to the bioreactor operation conditions); the bioreactor cell cculture dilution rate (D) equal to the ratio of the bioreactor fed flow-rate and the volume of the liquid in the bioreactor;

(ii) parameters depending on the cell characteristics, possible to be modulated in a GMO, that is: the Cell content dilution rate [μ=ln(2)/tc], this cell parameter is usually modulated via GMO to become equal to the bioreactor cell culture dilution rate (D)[169]; k6 reaction rate (determined by the ATP-ase characteristics, related to the cell phenotype); k6 and K parameters to be modulated in the glycolysis model are related the AMDTP pathway controlling the ATP to ADP conversion rate, as well as its quick/slow regeneration The characteristics of some key enzymes (e.g. ATP-ase, related to the cell phenotype) involved in the glycolytic oscillations [160-162]; the activity of such enzymes is reflected by the glycolysis model constants (like k6, and K), related to the AMDTP pathway controlling the ATP to ADP conversion rate, as well as its quick/slow regeneration; the activity of the enzymes involved in the AMDTP pathway can be modulated by using common techniques of getting desired GMO (chap. 8).

All other reaction rate constants have been previously estimated from the kinetic data of [59] (Figure 90), and [AMDTP] level are kept unchanged during simulations. The *in-silico* analysis indicated some domains where stationary oscillating glycolysis occurs (Figure 91). The results indicated that glycolytic oscillatins can easily be modulated by using both control variables related to the bioreactor operation, but also GMO displaying certain characteristics of the AMDTP pathway controlling the ATP to ADP conversion rate, as well as its quick/slow regeneration.

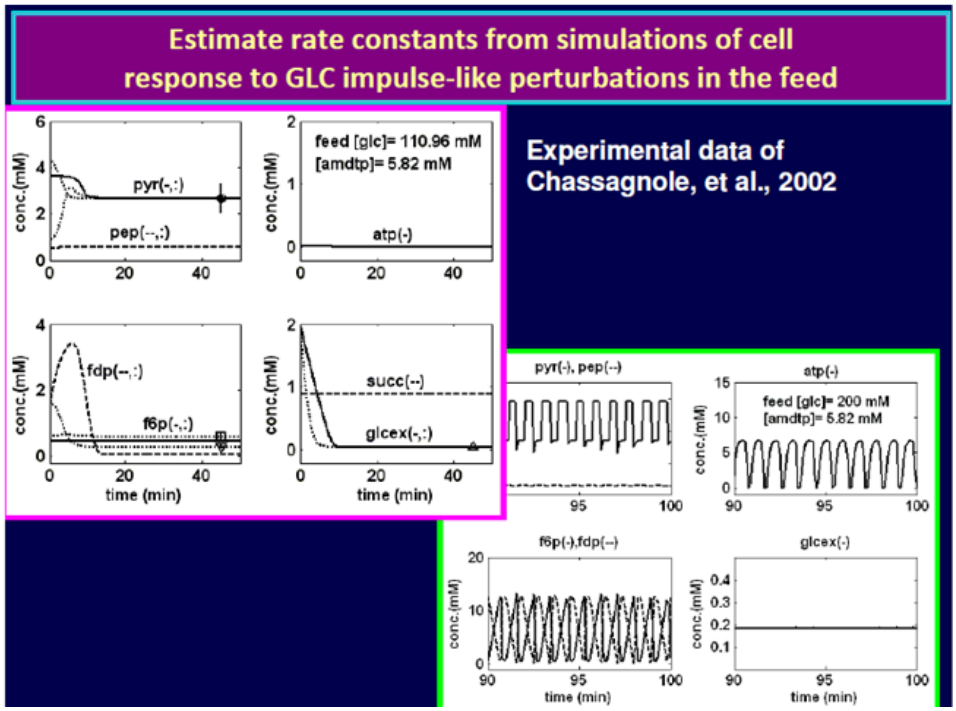


Figure 90: Case study 8.4.The rate constants of the [160] kinetic model of glycolysis are estimated by using the dynamic experimental data of [59]).

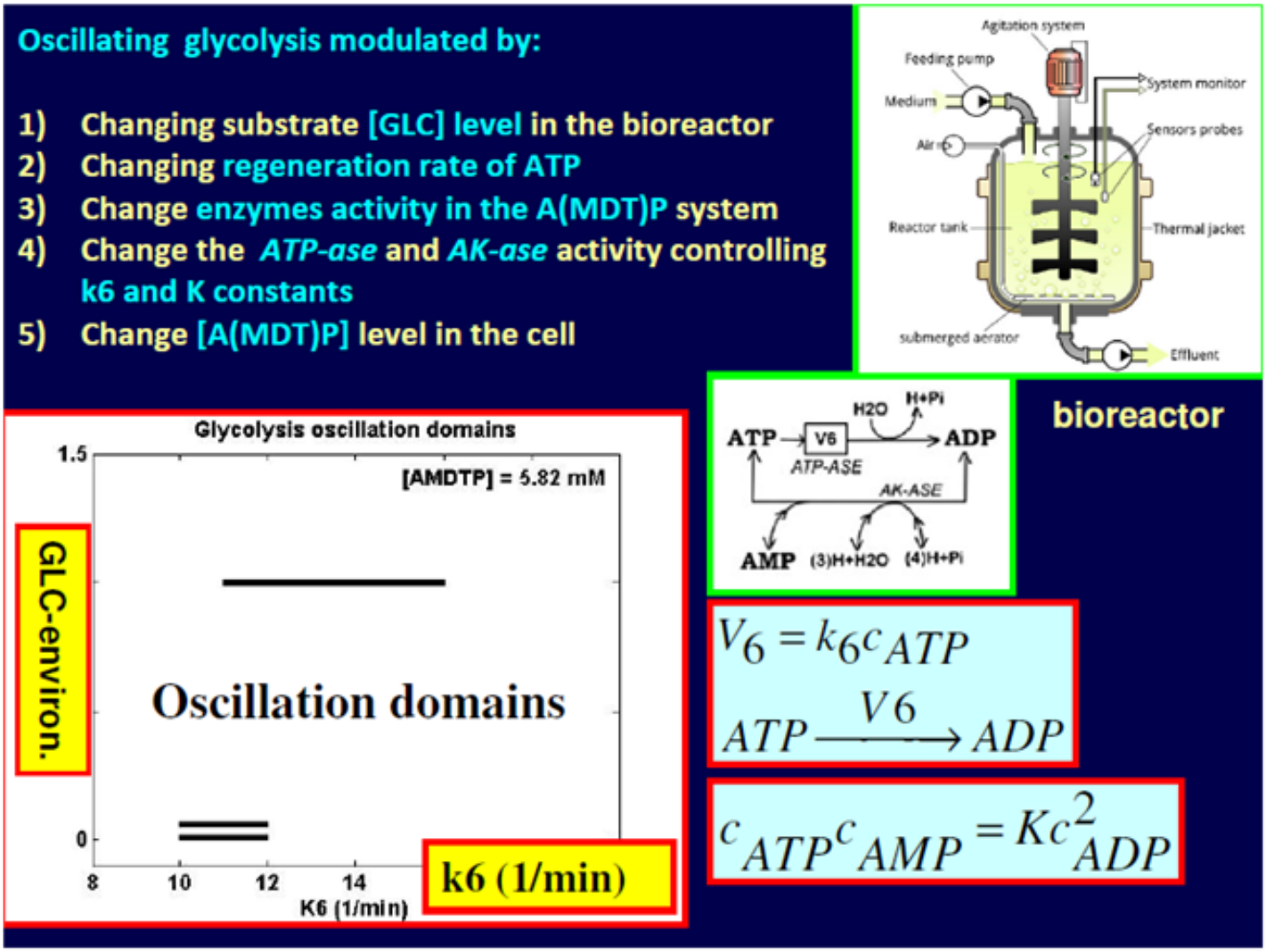


Figure 91: Case study 8.4.Some oscillation domains of the glycolysis in E. coli in-silico simulated by using (Maria, 2014 [160]) model. Suggestions to design GMO for adjusting glycolytic oscillations [161,162].

11.5 *In-silico* Design of a modified E.coli cell culture to optimize the bioreactor used for the tryptophan synthesis [163,164]

In this paragraph is presented an *in-silico* analysis of the operating conditions of a chemostat (bioreactor) for maximizing the production of tryptophan (TRP) in the *E. coli* cells. While some of the control variables depend bioreactor operating conditions, some other parameters depend on the cell characteristics, being possible to be *in-silico* modulated in a GMO.

Autonomous oscillations of glycolytic intermediates' concentrations reflect the dynamics of the control and regulation of this major catabolic pathway with a major role in the cell central carbon metabolism (CCM) in living cells. Consequently, glycolysis model is the 'core' module of any systematic and structured model-based analysis of the cell metabolism. On one hand, tryptophan (TRP) synthesis is another oscillatory metabolic process of high practical interest in the biosynthesis industry, and in medicine. On the other hand, the TRP synthesis is strongly connected to the glycolysis through the PEP (phosphoenolpyruvate) node. By coupling two adequate reduced kinetic models for glycolysis and TRP synthesis in the *E. coli* cells, adopted from literature, this paragraph exemplify how such an in silico analysis can be used to determine the optimal operating conditions of the bioreactor used for tryptophan synthesis, with accounting for the two interfering oscillatory processes (Figure 92).

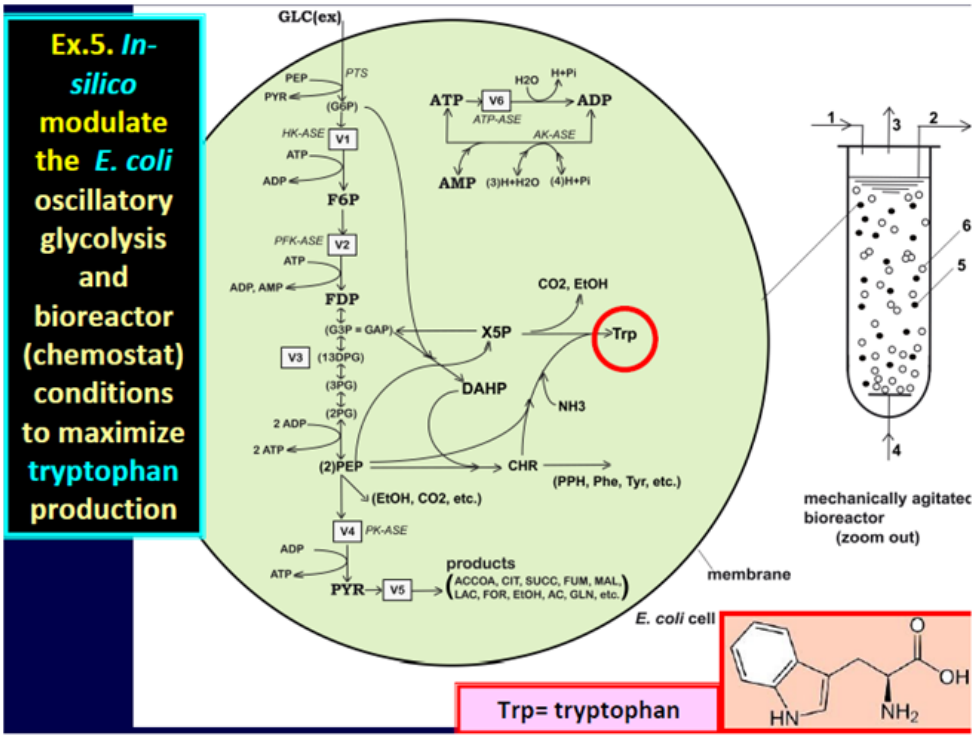


Figure 92: Case study 8.5. In-silico modulate the bioreactor operating conditions with suspended E. coli to maximize the production of tryptophan [163,164].

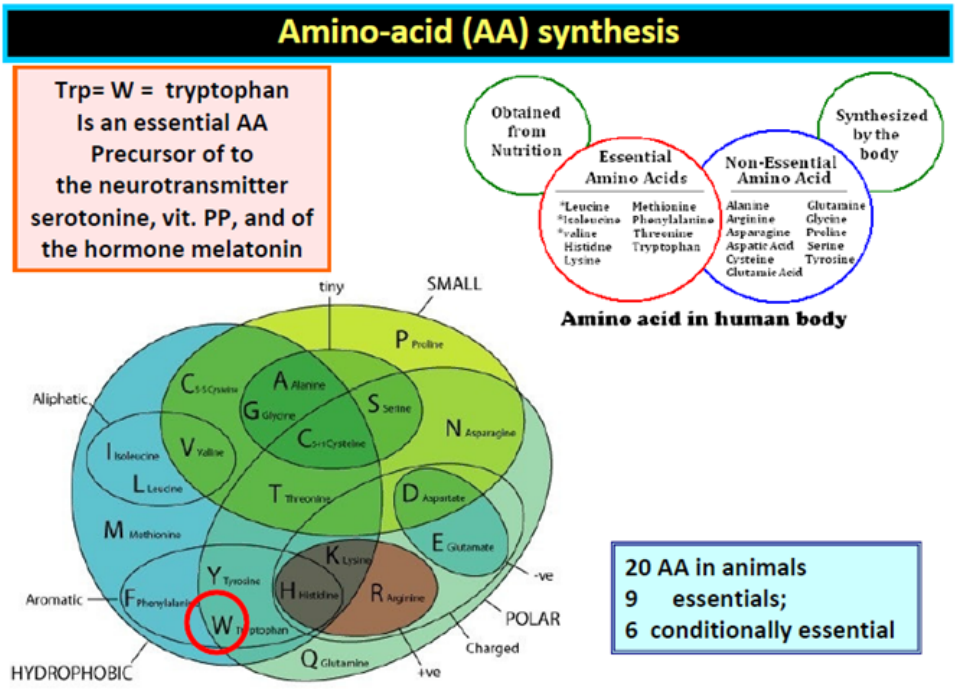


Figure 93: Case study 8.5.The importance of tryptophan (an essential amino-acid AA) in the humans.

Tryptophan is an aromatic non-polar α-amino-acid essential in humans, that is used in the cell biosynthesis of proteins (Figure 93), being also a precursor to the neuro-transmitter serotonin, of the melatonin hormone, and of vitamin PP [174]. Being a metabolite of high practical importance, seeking for its production maximization, intense efforts have been invested to decipher its synthesis regulation mechanism in

various micro-organisms, and for deriving an adequate dynamic model of its oscillatory synthesis. Being strongly connected to the glycolysis through the PEP node, to simulate the dynamics of the TRP-operon expression, a glycolysis model is also necessary. Although such cellular complex structured kinetic models, constructed on a mechanistic basis, are difficult to be derived and coupled, they have been proved to be extremely useful tools for *in-silico* checks of conditions leading to optimization of various biosyntheses [2].

This paragraph is aiming at presenting the results obtained by Maria [163,164] by using a model-based approach, in determining the operating conditions of a mechanically agitated semi-continuous bioreactor (chemostat), with a suspended *E. coli* cells culture, leading to the maximization of the productivity in the tryptophan (TRP) cytosolic (inside cells) synthesis.

For such a purpose, three dynamic models have been coupled, that is:

a) the macroscopic dynamic model of the bioreactor of (Chassagnole et al., 2002[59]) [using state-variables of a (milli-)molar level],

b) The oscillating glycolysis model of [160] to represent this cellular bioprocess by using state-variables of a (nano-, micro-)molar level; the connection between the cellular glycolysis and the bioreactor model is made via glucose (GLC) concentration in the liquid bulk-phase. That is because the first step of the glycolysis consists in the glucose (Glc) membranar import into the cell through the phosphotranspherase (PTS) system (Figure 94);

c) The tryptophan oscillatory synthesis model of (Bhartiya et al. (2006)[176]), modified by (Maria et al., 2018[163]) (Figure 94) for representing this cellular bioprocess by using state-variables of a (nano-, micro-)molar level; the connection between the cellular glycolysis model and the tryptophan synthesis model is made through the PEP node of glycolysis.

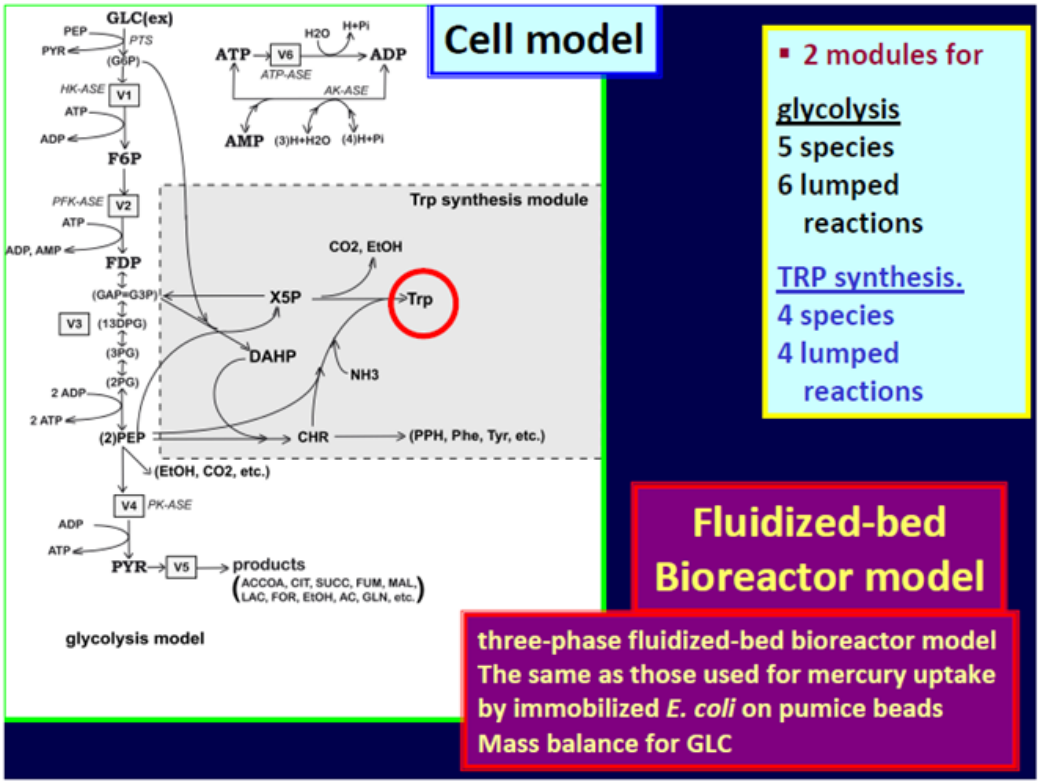


Figure 94: Case study 8.5.The structured dynamic cell model of [163,164] to simulate the oscillating glycolysis and tryptophan synthesis in *E. coli*.

The study of [163,164] includes important elements of novelty. Thus, by coupling two structured reduced (but adequate for the key-species) modular kinetic models for glycolysis and TRP synthesis in the *E. coli* cells (adopted from literature), this paper performs an in silico analysis, for the first time in the literature, of the way by which the two oscillatory processes interfere in the *E. coli* cells through the PEP glycolysis node and on the consequences on the TRP synthesis yield. Moreover, the cellular (nano-,micro-)bioprocess model, directly coupled with the bioreactor macroscopic dynamic model allows determining the bioreactor best operating policy leading to TRP biosynthesis maximization.

To *in-silico* modulate the glycolysis oscillations in an *E. coli* cell culture from a chemostat, and inherently the TRP synthesis yield [163,164] used two types of control parameters:

A. Bioreactor operating parameters, such as:

i) The glucose concentration in the fed solution [Glc]feed;

ii) The initial concentration of glucose [Glc](0);

iii) The bioreactor cell cculture dilution rate (D) equal to the ratio of the bioreactor fed flow-rate and the volume of the liquid in the bioreactor;

iv) The batch time *tf* to get the maximum of TRP productivity; this operating parameter can be easily determined by investigating the bioreactor dynamics over a large time horizon of 100 min. of continuous operation. Thus, *tf* is found as being the batch time when the maximum TRP production is realised for the first time. The TRP production (in micro-M/min) is defined as being the product [bioreactor dilution rate (D) × [max [TRP]*cyt* (t)], where the index "cyt" is referring to the TRP cytosolic concentration.

B. Parameters depending on the cell characteristics, possible to be modulated in a GMO, that is:

a. the cell content dilution rate [μ= ln(2)/cyt]; this parameter is usually is modulated via GMO to equal the bioreactor cell cculture dilution rate (D) [176].

b. characteristics of some key enzymes involved in the glycolytic oscillations (see chap. 11). We are referring here on k6 reaction rate (determined by the ATP-ase characteristics, related to the cell phenotype); k6 and K parameters to be modulated in the glycolysis model are related the AMDTP pathway controlling the ATP to ADP conversion rate, as well as its quick/slow regeneration.

The reported results of [163,164] have revealed the following issues regarding the TRP synthesis yield (Figure 95-96):

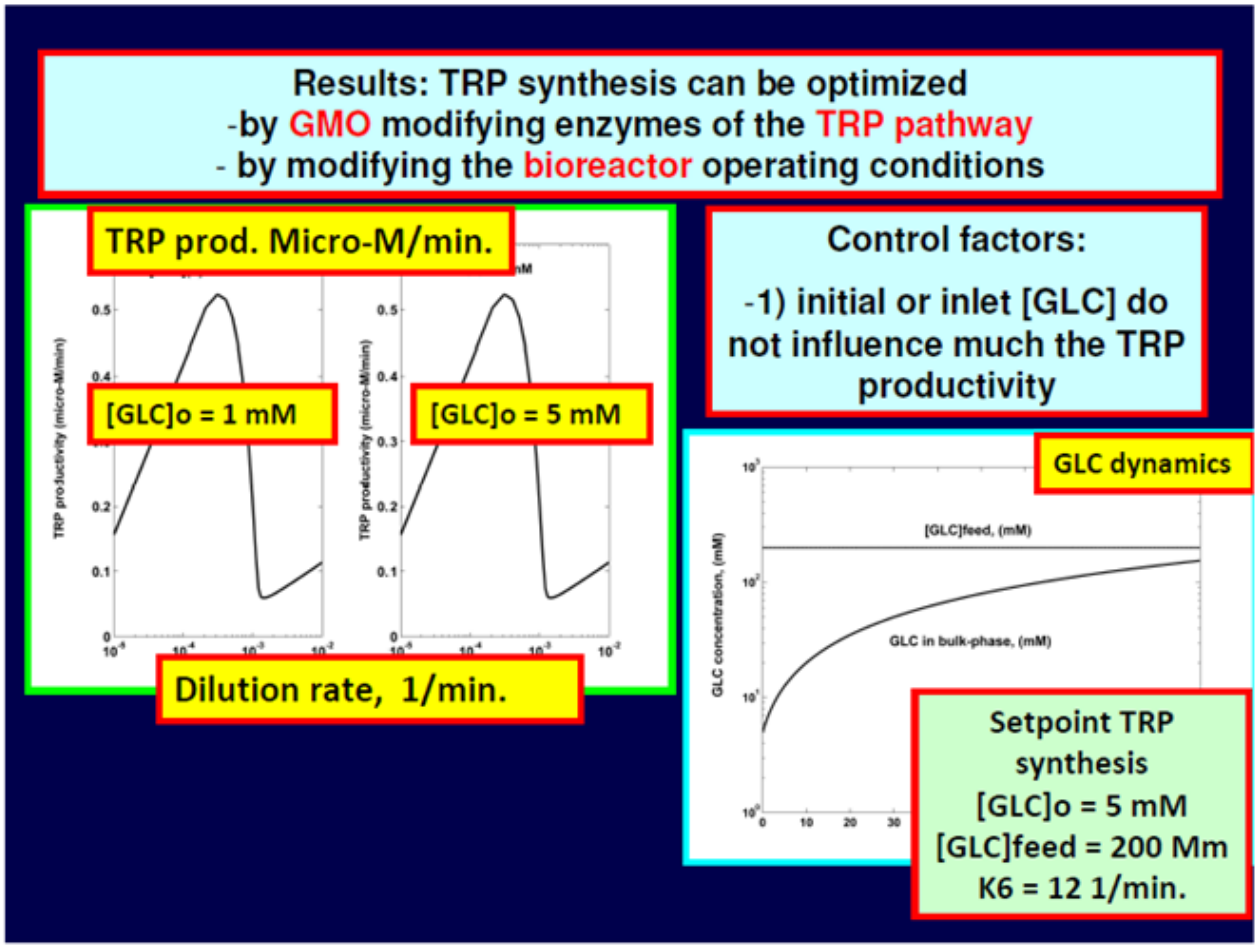


Figure 95: Case study 8.5.In-silico evidence that tryptophan synthesis productivity is not influenced by the initial or inlet concentrations of glucose in the chemostat, [163,164].z

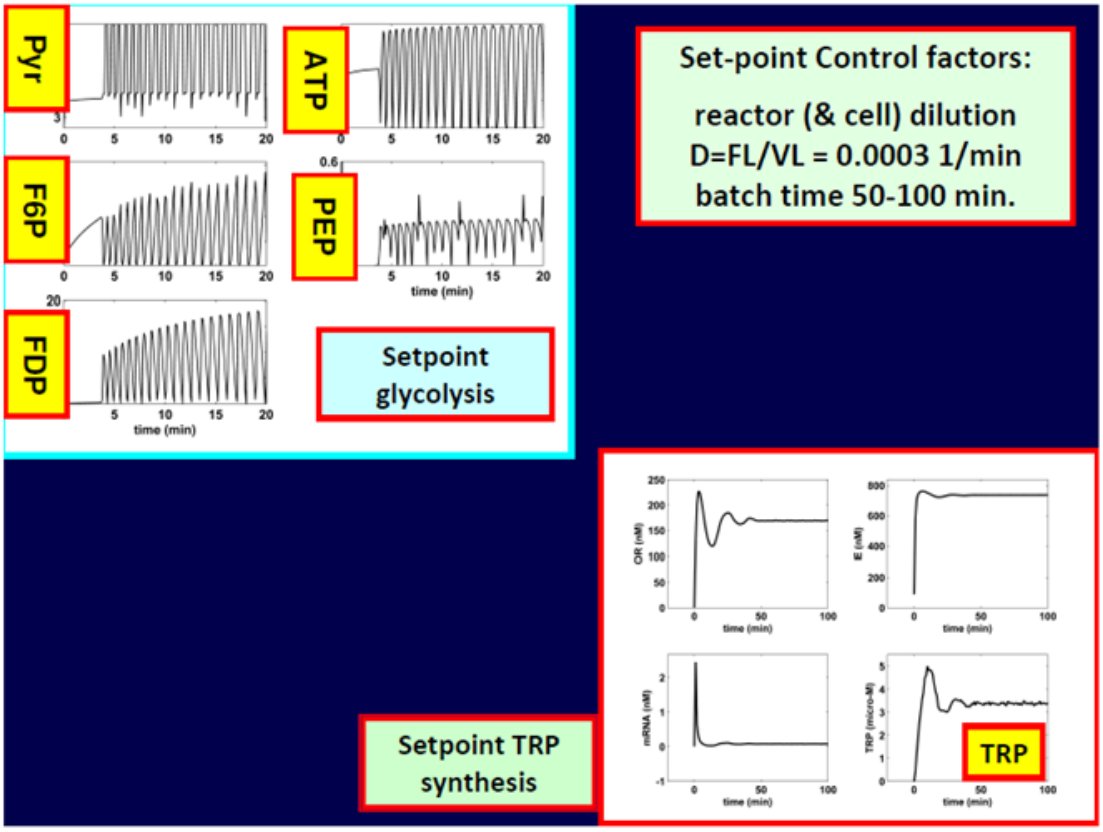


Figure 96: Case study 8.5.In-silico determination of the optimal set-point (bioreactor and cell dilution) leading to maximization of the tryptophan synthesis [163,164].

i. The glucose initial concentration does not influence quantitatively the bioreactor performances;

ii. On the contrary, the TRP production is strongly influenced by the bioreactor dilution rate (D); the maximum TRP production is reached for a value of 0.0003097 (under the Chassagnole et al., 2002[59] bioreactor conditions);

iii. the TRP production is also strongly influenced by the cell content dilution rate [μ= ln(2)/tc]; the maximum TRP production is reached for a value equal to D;

iv. The bioreactor dilution also strongly influences the quasi-steady state regime (QSS) or the stable oscillations regime (OSC) of the two cellular bioprocess es(glycolysis and TRP synthesis).

v. TRP production is not influenced by the glucose concentration in the feeding solution;

In all cases, it is worth noting the firm evolution of the glucose level in the bulk-phase toward its steady-state. Also, it is to remark the strong influence of the dilution on the oscillatory behaviour of the two cell sub-processes. While the glycolytic species present stationary oscillations for D > 0.001(1/min.), these oscillations tend to be amortized for lower D. Concerning the TRP synthesis oscillations, these are very stable, even if of small amplitude, only for D=μ> 0.001(1/min.). For lower values of D=μ, the TRP synthesis tends to steady-state.

12. Conclusions

As a general conclusion, the (bio)chemical engineering principles, modelling rules, and algorithms are fully applicable to modelling cellular metabolic processes, useful to *in-silico* design GMO. 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 cell processes, the adaptable-size and structure of the lumped model depending on available information and model utilisation scope.

The briefly reviewed case studies prove in a relatively simple yet eloquent way how a structured, but enough detailed and adequate dynamic model of metabolic cell sub-processes can support in silico engineering evaluations, if the cell nano-/micro-scale cell model is adequately coupled to the bioreactor macro-scale dynamic model. In such a way it can be realized a direct connection between the macro- and the nano-process variables to finally sustain bioreactor optimization decisions related to both operating conditions and used GMOs.

The structured cell model allows an easy evaluation of cellular metabolic fluxes (i.e. the QSS cell reaction rates). Such an analysis quickly opens the possibility to in silico re-design some cell fluxes, via *in-silico* re-programming the cell metabolism to design GMO with industrial or medical applications (Maria, 2017[1-2]; Visser et al., 2004[125]; Styczynski and Stephanopoulos, 2005[94]; Heinemann and Panke, 2006[61]).

In such a way, the significant large experimental and computational effort to elaborate structured cell models of good quality is fully justified through the benefits of subsequent in silico analyses (as is also the case here).

The cell process modular/structured modelling approach is computationally fully tractable. The deterministic model can be successfully integrated in semi-autonomous modular simulation platforms to study metabolic syntheses regulation properties. To be feasible, the cell lumped models must realize a suitable trade-off between simplicity and model quality vs. physical meaning of (reaction, species) lumps.

GRC representations combining Reverse Engineering and Integrative Understanding (Maria, 2017[1-2]) allow *in-silico* design of GRC inducing specific cell motifs of genetically modified micro-organisms (GMO). Examples includes: Genetic switches of adjustable certainty, sensitivity to exo-/endogeneous stimuli, responsivity, regulatory efficiency;

Even if the cell simulators still present lot of drawbacks and a limited adequacy, they become more and more valuable tools for *in-silico* designing of GMO with desirable characteristics, for obtaining cells with a modified metabolism, by using the gene-knockout procedure, or by cloning cells with target plasmids. Such design GMO present important applications in medicine, such as therapy of diseases (gene therapy), or for obtaining new devices based on cell-cell communicators, biosensors, or even more important in industry (new biotechnological processes, optimization of bioreactors, novel biosyntheses, production of vaccines), etc.

References

1. Maria G (2017) A review of some novel concepts applied to modular modelling of genetic regulatory circuits. Juniper publ., Newbury Park, CA, USA. ISBN: 978-1-946628-03-9.
2. Maria G (2017) Deterministic modeling approach of metabolic processes in living cells-a still powerful tool for representing the metabolic process dynamics. Juniper publ., Newbury Park, CA, USA. ISBN: 978-1-946628-07-7.
3. Maria G (2017) Application of (bio) Chemical engineering principles and lumping analysis in modeling the living systems. Current Trends in Biomedical Engineering & Biosciences 1(4): 1-9.
4. Kacser H, Burns JA (1973) The control of flux. Symp Soc Exp Biol 27: 65-104.
5. MCA Web (2004) The Metabolic Control Analysis Web, The University of Manchester. England.
6. Heinrich R, Schuster S (1996) The regulation of cellular systems. Chapman & Hall, New York.
7. Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, et al. (2004) Programmable cells: Interfacing natural and engineered gene networks. Proc Natl Acad Sci 101(22): 8414-8419.
8. Sotiropoulos V, Kaznessis YN (2007) Synthetic tetracycline-inducible regulatory networks: Computer-aided design of dynamic phenotypes. BMC Syst Biol. pp. 1-7.
9. Leroy H (2017) Statement given by the Centre of Mathematics Applied to the Life Sciences, University of Glasgow, UK.
10. Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED (2002) Metabolic network structure determines key aspects of functionality and regulation. Nature 420(6912): 190-193.
11. Snoep JL, Olivier BG (2003) JWS online cellular systems modelling and microbiology. Microbiology 149(Pt 11): 3045-3047.
12. Rocha I, Maia P, Evangelista P, Vilaça P, Soares S, et al. (2010) OptFlux: an open-source software platform for in silico metabolic engineering. BMC Syst Biol 4: 45.
13. Tomita M, Hashimoto K, Takahashi K, Shimizu T, Matsuzaki Y, et al. (1999) E-Cell: Software environment for whole cell simulation. Bioinformatics 15(1): 72-84.
14. EcoCyc (2005) Encyclopedia of Escherichia coli K-12 genes and metabolism, SRI Intl., The Institute for Genomic Research. University of California at San Diego, USA.
15. Kegg Pathway (2011) Phenylalanine, tyrosine and tryptophan biosynthesis; Kyoto encyclopedia of genes and genomes, Kanehisa Laboratories. Bioinformatics Center of Kyoto University, Japan.
16. Von Bertalanffy (1933) Modern Theories of Development: An Introduction to Theoretical Biology. Oxford University Press. Harper, New York.
17. Maria G (2005) Modular-Based Modelling of Protein Synthesis Regulation. Chemical and Biochemical Engineering Quarterly 19(3): 213-233.
18. Heinrich R, Rapoport TA (1974) A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. Eur J Biochem 42(1): 89-95.
19. Burns JA, Cornish Bowden A, Groen AK, Heinrich R, Kacser H, et al. (1985) Control of metabolic systems. Trends Biochem Sci 10(1): 16.
20. Kholodenko BN, Schuster S, Garci J, Westerhoff HV, Cascante M (1998) Control analysis of metabolic systems involving cvasi ecvilibrium reactions. Biochimica Biophysica Acta 1379(3): 337-352.
21. Atauri P, Curto R, Puigjaner J, Cornish Bowden A, Cascante M (1999) Advantages and disadvantages of aggregating fluxes into synthestic and degradative fluxes when modelling metabolic partways. Eur J Biochem 265(2): 671-679.
22. Szedlacsek ES, Aricescu AR, Havsteen BH (1996) Time-dependent control of metabolic systems by external effectors. J Theor Biol 182(3): 341-350.
23. Hodgkin AL, Huxley AF (1952) A quantitative description of ion currents and its applications to conductions and excitation in nerve membranes. J Physiol 117(4): 500-544.
24. Noble D (1962) A modification of the Hodgkin-Huxley equations applicable to Purkinje fibre action and pace-maker potentials. J Physiol 160: 317-352.
25. Noble D (2002) Modeling the heart-from genes to cells to the whole organ. Science 295(5560): 1678-1682.
26. Kitano H (2002) Computational systems biology. Nature 420(6912): 206-210.
27. Ideker T, Galitski T, Hood L (2001) A new approach to decoding life: Systems Biology. Annu Rev Genomics Hum Genet 2: 343-372.
28. Banga J (2008) Optimization in computational systems biology. 6-th Simon Stevin lecture on optimization in engineering. Leuven-Heverlee, NL. BMC Systems Biology 2: 47.
29. Nandy SK (2017) Bio-XYZ, What is XYZ? Curr Trends Biomedical Eng & Biosci 3(2): 555606.
30. Schaff (2001) Virtual cell (V-Cell) Project, 1st Intl. Symp. on Computational Cell Biology, NRCAM (National Resource for Cell Analysis and Modelling) of the NIH Home-page.
31. Bartol Jr TM, Stiles JR (2002) MCell: A general Monte Carlo simulator of cellular microphysiology. The Salk Institute, Computational Neurobiology Lab. USA.
32. Torres NV, Voit EO (2002) Pathway analysis and optimization in metabolic engineering. Cambridge University Press, Cambridge, USA.
33. Cornish Bowden A (2016) Biochemical evolution-The pursuit of perfection, Garland Science, New York.

34. Brazhnik P, de la Fuente A, Mendes P (2002) Gene networks: how to put the function in genomics. Trends in Biotechnol 20(11): 467-472.

35. Wolkenhauer O, Mesarovic M (2005) Feedback dynamics and cell function: Why Systems Biology. Mol Biosyst 1(1): 14-16.

36. Sutherland W (2005) Optimization-this beguilingly simply idea allows biologists not only to understand current adaptations, but also to predict new designs that may yet evolve. Nature. pp. 435.

37. Maria G (2012) Enzymatic reactor selection and derivation of the optimal operation policy by using a model-based modular simulation platform. Computers & Chemical Engineering 36(1): 325-341.

38. Hatzimanikatis V, Floudas CA, Bailey JE (1996) Analysis and design of metabolic reaction networks via mixed-integer linear optimization. AIChE J 42(5): 1277-1292.

39. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, et al. (2002) Molecular Biology of the Cell, (4th edn). Garland Science, New York.

40. KEGG (2011) Kyoto encyclopedia of genes and genomes, Kanehisa Laboratories. Bioinformatics Center of Kyoto University, Japan.

41. Olivier BG, Snoep JL (2004) JWS Online and the Silicon Cell project; Web-based kinetic modelling using JWS Online. Bioinformatics 20(13): 2143-2144.

42. Peters M, Eicher JJ, Van Niekerk DD, Waltemath D, Snoep JL (2017) The JWS online simulation database. Bioinformatics 33(10): 1589-1590.

43. Michal G (2014) Roche co, Basel CH (Source=https://www.roche.com/dam/jcr:a8123664-2dac-4e4f-a744-0d847446f874/en/biochemical_pathways_factsheet_150514.pdf)

44. EcoCyc (2005) Encyclopedia of Escherichia coli K-12 genes and metabolism, SRI Intl., The Institute for Genomic Research, University of California at San Diego, USA.

45. Noble, D. (2006) The music of life. Oxford University, USA.

46. Maria G (2006) Application of lumping analysis in modelling the living systems-A trade-off between simplicity and model quality. Chemical and Biochemical Engineering Quarterly 20(4): 353-373.

47. Ichikawa K (2001) A-Cell: graphical user interface for the construction of biochemical reaction models. Bioinformatics 17(5): 483-484.

48. Bower JM, Bolouri H (2001) Computational Modeling of Genetic and Biochemical Networks, MIT Press, Cambridge, USA.

49. Hudder B, Yang Q, Bolfing B, Maria G, Morgan JJ, et al. (2002) Computational Modeling of Iron Metabolism in Mitochondria. NIST Conference on 'Systems Biology Approaches to Healt Care: Mitochondrial Proteomics. Gaithersburg, USA, pp. 17-18.

50. Stelling J (2004) Mathematical models in microbial systems biology. Curr opin Microbiol 7(5): 513-518.

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

52. Maria G (2003) Evaluation of Protein Regulatory Kinetics Schemes in Perturbed Cell Growth Environments by Using Sensitivity Methods. Chemical and Biochemical Engineering Quarterly 17(2): 99-117.

53. Maria, G., Gijiu, C.L., Maria, C., Tociu, C., Mihalachi, M. (2018), Importance of considering the isotonic system hypothesis when modelling the self-control of gene expression regulatory modules in living cells, Current Trends in Biomedical Engineering & Biosciences, 12(2), CTBEB.MS.ID.555833, DOI : 10.19080/CTBEB.2018.12.555833.

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

55. Maria G, Scoban AG (2017) Setting some milestones when modelling gene expression regulatory circuits under variable-volume whole-cell modelling framework. 1. Generalities, Revista de Chimie (Bucharest) 68(12): 3027-3037; Part 2. 69(1): 259-266 (2018).

56. Maria G (2007) Modelling bistable genetic regulatory circuits under variable volume framework. Chemical and Biochemical Engineering Quarterly 21(4): 417-434.

57. Maria G (2014) Extended repression mechanisms in modelling bistable genetic switches of adjustable characteristics within a variable cell volume modelling framework. Chemical & Biochemical Engineering Quarterly 28(1): 35-51.

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

59. Chassagnole C, Noisommit Rizz N, Schmid JW, Mauch K, Reuss M (2002) Dynamic modeling of the central carbon metabolism of *Escherichia coli*. Biotechnology and Bioengineering 79(1): 53-73.

60. Sauro HM, Kholodenko BN (2004) Quantitative analysis of signalling networks. Prog Biophys Mol Biol 86(1): 05-43.

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

62. Benner SA, Sismour AM (2005) *Synthetic biology*. Nature Reviews Genetics 6: 533-543.

63. Voit EO (2005) Smooth bistable S-systems. Syst Biol (Stevenage) 152(4): 207-213.

64. Liese A, Seelbach K, Wandrey C (2006) Industrial biotransformations: Wiley-VCH, Weinheim.

65. Wang P (2009) Multi-scale features in recent development of enzymatic biocatalyst systems. Appl Biochem Biotechnol 152(2): 343-352.

66. Moulijn JA, Makkee M, van Diepen A (2001) Chemical process technology. Wiley, New York.

67. Turing A (1952) The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society of London (Series B) 237(641): 37-72.
68. Myers CJ (2009) Engineering genetic circuits. Chapman & Hall/CRC Press. Boca Raton, USA.
69. Kitano H (2002) Computational Systems Biology. Nature 420: 206-210.
70. Schrödinger E (1944) What is life? The physical aspect of the living cell, Cambridge univ. press.
71. Wiener N (1948) Cybernetics or control and communication in the animal and the machine. McGraw-Hill, New York.
72. Mesarovic M (1968) System Theory and Biology. Springer Verlag, Berlin, Germany.
73. Stephanopoulos GN, Aristidou AA, Nielsen J (1998) Metabolic Engineering. Principles and Methodologies. Academic Press, San Diego, CA.
74. Klein S, Heinzle E (2012) Isotope labeling experiments in metabolomics and fluxomics. Wiley Interdiscip Rev Syst Biol Med 4(3): 261-272.
75. Zak DE, Aderem A (2009) Systems biology of innate immunity. Immunol Rev 227(1): 264-282.
76. Tomita M, Hashimoto K, Takahashi K, Shimiz T, Matsuzaki Y, et al. (1999) E-Cell: Software environment for whole cell simulation. Bioinformatics 15(1): 72-84.
77. Schaff (2001) Virtual cell (V-Cell) Project, 1st Intl. Symp on Computational Cell Biology.
78. Bartol Jr TM, Stiles JR (2002) MCell: A general Monte Carlo simulator of cellular microphysiology. The Salk Institute, Computational Neurobiology Lab.
79. Westerhoff HV (2006) Engineering life processes live: the Silicon cell. ESCAPE-16 Conference, Garmisch-Partenkirchen, Germany.
80. Kurata H, Sugimoto Y (2017) Improved kinetic model of Escherichia coli central carbon metabolism in batch and continuous cultures. Journal of Bioscience and Bioengineering, (in-press).
81. Miskovic L, Tokic M, Fengos G, Hatzimanikatis V (2015) Rites of passage: requirements and standards for building kinetic models of metabolic phenotypes. Curr Opin Biotechnol 36: 146-153.
82. Chen T, He HL, Church GM (1999) Modeling gene expression with differential equations. Pac Symp Biocomput. pp. 29-40.
83. Drengstig T, Jolma IW, Ni XY, Thorsen K, Xu XM, et al. (2012) A Basic Set of Homeostatic Controller Motifs. Biophys J 103(9): 1-11.
84. Komasilovs V, Pentjuss A, Elsts A, Stalidzans E (2017) Total enzyme activity constraint and homeostatic constraint impact on the optimization potential of a kinetic model. Biosystems 162: 128-134.
85. Matsuura T, Tanimura N, Hosoda K, Yomo T, Shimizu Y (2017) Reaction dynamics analysis of a reconstituted *Escherichia coli* protein translation system by computational modeling. PNAS 114(8): E1336-E1344.
86. Zhang R, Ehigie JO, Hou X, You X, Yuan C (2017) Steady-State-Preserving Simulation of genetic regulatory systems. Computational and Mathematical Methods in Medicine 2017(2017). pp. 1-16.
87. Blass LK, Weyler C, Heinzle E (2017) Network design and analysis for multi-enzyme biocatalysis. BMC Bioinformatics 18(1): 366.
88. Morgan JJ, Surovtsev IV, Lindahl PA (2004) A framework for whole-cell mathematical modeling. J Theor Biol 231(4): 581-596.
89. Surovstev IV, Morgan JJ, Lindahl PA (2007) Whole-cell modeling framework in which biochemical dynamics impact aspects of cellular geometry. J Theor Biol 244(1): 154-166.
90. Savageau MA (2002) Alternatives designs for a genetic switch: Analysis of switching times using the piecewise power-law representation. Math Biosci 180: 237-253.
91. Banga JR (2008) Optimization in computational systems biology. BMC Syst Biol 28(2): 47.
92. Price ND, Reed JL, Palsson BO (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat Rev Microbiol 2(11): 886-897.
93. Schellenberger J, Que R, Fleming RM, Thiele I, Orth JD, et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: the cobra toolbox v2. 0. Nat Protoc 6(9): 1290-1307.
94. Styczynski MP, Stephanopoulos G (2005) Overview of computational methods for the inference of gene regulatory networks. Comp Chem Eng 29(3): 519-534.
95. Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED (2002) Metabolic network structure determines key aspects of functionality and regulation. Nature 420(6912): 190-193.
96. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, et al. (2003) The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models. Bioinformatics 19(4): 524-531.
97. CRGM-Database (2002) Centro di Ricerca in Genetica Molecolare, Istituto di Istologia e Embriologia Generale. University of Bologna, Italy.
98. NIH Database (2004) National Institute of Health database.
99. Crampin EJ, Schnell S (2004) New approaches to modelling and analysis of biochemical reactions. pathways and networks. Prog Biophys Molecul Biol 86(1): 1-4.
100. McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A 94(3): 814-819.
101. Shimizu TS, Bray D (2002) Computational cell biology-The stochastic approach. In: Shimizu TS, Bray D (Eds.), Computational cell biology-The stochastic approach. MIT Press Math6X9, Cambridge, UK.

102. Gillespie DT, Mangel M (1981) Conditioned averages in chemical kinetics. The Journal of Chemical Physics 75(2): 704-709.

103. Miskovic L, Tokic M, Fengos G, Hatzimanikatis V (2015) Rites of passage: requirements and standards for building kinetic models of metabolic phenotypes. Curr Opin Biotechnol 36: 146-153.

104. McAdams HH, Shapiro L (1995) Circuit simulation of genetic networks. Science 269(5224): 650-656.

105. Tyson JJ, Novak B (2001) Regulation of the eukaryotic cell cycle: Molecular antagonism, hysteresis, and irreversible transitions. J Theor Biol 210(2): 249-263.

106. Heinrich R, Rapoport SM, Rapoport TA (1978) Metabolic regulation and mathematical models. Progress in Biophysics Molecular Biology 32: 1-82.

107. Somogyi R, Sniegoski CA (1996) Modeling the complexity of genetic networks: understanding multigenetic and pleiotropic regulation. Complexity 1(6): 45-63.

108. Tyson JJ, Novak B, Odell GM, Chen K, Thron CD (1996) Chemical kinetic theory: Understanding cell cycle regulation. Trends in Biochemical Sciences 21(3): 89-96.

109. Haraldsdottir HS, Thiele I, Fleming RMT (2012) Quantitative assignment of reaction directionality in a multicompartmental human metabolic reconstruction. Biophys J 102(8): 1703-1711.

110. Zhu Y, Song J, Xu Z, Sun J, Zhang Y, et al. (2013) Development of thermodynamic optimum searching (TOS) to improve the prediction accuracy of flux balance analysis. Biotechnol Bioeng 110(3): 914-923.

111. Aris R (1969) Elementary chemical reactor analysis. In: Aris R (Eds.), Elementary chemical reactor analysis. Prentice-Hall, New Jersey, USA.

112. Wallwork SC, Grant DJW (1977) Physical Chemistry. In: Wallwork SC, Grant DJW (Eds.), Physical Chemistry. (4th edn), Longman, London, UK.

113. Qian Y, McBride C, Del Vecchio D (2017) Programming Cells to Work for Us. Annual Review of Control, Robotics, and Autonomous Systems 1:

114. Kholodenko BN, Kiyatkin A, Bruggeman FJ, Sontag E, Westerhoff HV, Hoek JB (2002) Untangling the wires: A strategy to trace functional interactions in signalling and gene networks. PNAS 99(20): 12841-12846.

115. Kaznessis YN (2006) Multi-scale models for gene network engineering. Chemical Engineering Science 61(3): 940-953.

116. Salis H, Kaznessis Y (2005) Numerical simulation of stochastic gene circuits, Computers &. Chemical Engineering 29(3): 577-588.

117. Atkinson MR, Savageau MA, Myers JT, Ninfa AJ (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli* Cell 113(5): 597-607.

118. Klipp E, Nordlander B, Krüger R, Gennemark P, Hohmann S, et al. (2005) Integrative model of the response of yeast to osmotic shock. Nature Biotechnology 23(8): 975-982.

119. Chen MT, Weiss R (2005) Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signalling elements from Arabidopsis thaliana. Nature Biotechnology 23: 1551-1555.

120. Tian T, Burrage K (2006) Stochastic models for regulatory networks of the genetic toggle switch. PNAS 103(22): 8372-8377.

121. Tomshine J, Kaznessis YN (2006) Optimization of a stochastically simulated gene network model via simulated annealing. Biophys J 91(9): 3196-3205.

122. Zhu R, Ribeiro AS, Salahub D, Kauffman SA (2007) Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models. Journal Theoretical Biology 246(4): 725-745.

123. Bailey JE (1991) Towards a science of metabolic engineering. Science 252(5013): 1668-1675.

124. Nielsen J (1998) Metabolic engineering, techniques for analysis of targets for genetic manipulations. Biotechnol Bioeng 58(2-3): 125-132.

125. Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ, et al. (2004) Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. Metabolic Engineering 6(4): 378-390.

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

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

128. Varma A, Morbidelli M, Wu H (1999) Parametric sensitivity in chemical systems. In: Varma A, Morbidelli M, Wu H (Eds.), Parametric sensitivity in chemical systems. Cambridge University Press, Cambridge, UK.

129. Liao JC, Lightfoot EN (1987) Extending the quasi-steady state concept to analysis of metabolic networks. J Theor Biol 126(3): 253-273.

130. Hofmeyr JHS (2001) Metabolic control analysis in a nutshell, Proc. 2-nd Intl. Conf. on Systems Biology, Madison, WI.

131. Vance W, Arkin A, Ross J (2002) Determination of causal connectivities of species in reaction networks. Proc Natl Acad Sci U S A 99(9): 5816-5821.

132. Puga A, Wallace K (1998) Molecular Biology of the Toxic Response. In: Puga A, Wallace K (Eds.), Molecular Biology of the Toxic Response. CRC Press, Boca Raton, US.

133. Das S, Caragea D, Welch SM, Hsu WH (2010) Handbook of research and computational methodologies in gene regulatory networks. In: Das S, Caragea D, Welch SM, Hsu WH (Eds.), Handbook of research and computational methodologies in gene regulatory networks. Medical information science reference (IGI Global), New York.

134. Luscombe NM, Greenbaum D, Gerstein M (2001) What is bioinformatics? An introduction and overview. Yearbook of Medical Informatics, USA, pp. 83-99.

135. Lopresti D (2010) Introduction to Bioinformatics, course at dept. Computer Science & Engineering, Lehigh university, US.

136. Agbachi CPE (2017) Pathways in Bioinformatics: A Window in Computer Science. International Journal of Computer Trends and Technology (IJCTT) 49(2): 83-90.

137. Xiong J (2006) Essential bioinformatics. In: Xiong J (Ed.), Essential bioinformatics. Cambridge university press, UK.

138. Wall ME, Hlavacek WS, Savageau, MA (2003) Design principles for regulator gene expression in a repressible gene circuit. J Mol Biol 332(4): 861-876.

139. Hlavacek WS, Savageau MA (1997) Completely uncoupled and perfectly coupled gene expression in repressible systems. J Mol Biol 266(3): 538-558.

140. Maria G, Crisan M (2017) Operation of a mechanically agitated semi-continuous multi-enzymatic reactor by using the Pareto-optimal multiple front method. Journal of Process Control 53: 95-105.

141. Burgard AP, Pharkya P, Maranas CD (2003) OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol Bioeng 84(6): 647-657.

142. Philippidis GP, Malmberg LH, Hu WS, Schottel JL (1991) Effect of gene amplification on mercuric ion reduction activity of *Escherichia coli*. Applied & Environmental Microbiology 57(12): 3558-3564.

143. Philippidis GP, Schottel JL, Hu WS (1991b) Mathematical modelling and optimization of complex biocatalysis. A case study of mercuric reduction by *Escherichia coli*. NSF report ECE 8552670, Univ. of Minnesota, pp. 35-49.

144. Philippidis GP, Schottel JL, Hu WS (1991c) A model for mercuric ion reduction in recombinant *Escherichia coli*. Biotechnology Bioengineering 37(1): 47-54.

145. Barkay T, Miller SM, Summers AO (2003) Bacterial mercury resistance from atoms to ecosystems. FEMS Microbiol Rev 27(2-3): 355-384.

146. Wagner-Döbler I, von Canstein H, Li Y, Timmis K, Deckwer WD (2000) Removal of mercury from chemical wastewater by microorganisms in technical scale. Environ Sci Technol 34(21): 4628-4634.

147. Deckwer WD, Becker FU, Ledakowicz S, Wagner-Döbler I (2004) Microbial removal of ionic mercury in a three-phase fluidised bed reactor. Environ Sci Technol 38(6): 1858-1865.

148. Turner A, Wilson G, Kaube I (1987) Biosensors: Fundamentals and applications. In: Turner A, Wilson G, Kaube I (Eds.), Biosensors: Fundamentals and applications. Oxford University Press, UK, p. 770.

149. Bănică FG (2012) Chemical Sensors and Biosensors: Fundamentals and applications. In: Bănică FG (Ed.), Chemical Sensors and Biosensors: Fundamentals and applications. John Wiley & Sons, Chichester, UK. p. 576.

150. Cavalcanti A, Shirinzadeh B, Zhang M, Kretly LC (2008) Nanorobot Hardware Architecture for Medical Defense. Sensors 8 (5): 2932-2958.

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

152. Maria G (2009) A whole-cell model to simulate the mercuric ion reduction by *E coli* under stationary and perturbed conditions, Chemical and Biochemical Engineering Quarterly 23(3): 323-341.

153. Maria G, Luta I, Maria C (2013) Model-based sensitivity analysis of a fluidised-bed bioreactor for mercury uptake by immobilised *Pseudomonas putida* cells, Chemical Papers, 67(11): 1364-1375.

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

155. Scoban AG, Maria G (2016) Model-based optimization of the feeding policy of a fluidized bed bioreactor for mercury uptake by immobilized *Pseudomonas putida* cells. Asia-Pacific Journal of Chemical Engineering 11(5): 721-734.

156. Maria G, Scoban AG (2017) Optimal operating policy of a fluidized bed bioreactor used for mercury uptake from wastewaters by using immobilized *P. putida* cells. Current Trends in Biomedical Engineering & Biosciences 2(4):

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

158. Maria G (2014) Extended repression mechanisms in modelling bistable genetic switches of adjustable characteristics within a variable cell volume modelling framework. Chemical & Biochemical Engineering Quarterly 28(1): 35-51.

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

160. Maria G (2014) In silico derivation of a reduced kinetic model for stationary or oscillating glycolysis in *Escherichia coli* bacterium. Chemical & Biochemical Engineering Quarterly 28(4): 509-529.

161. Maria G, Mihalachi M, Gijiu CL (2018) Model-based identification of some conditions leading to glycolytic oscillations in *E. coli* cells. Chemical and Biochemical Engineering Quarterly accepted.

162. Maria G, Mihalachi M, Gijiu CL (2018) Chemical engineering tools applied to simulate some conditions producing glycolytic oscillations in *E. coli* cells. U P B Sci Bull Series B-Chemie 79(3): 3-19.

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

164. Maria G, Mihalachi M, Gijiu CL (2018) In *silico* optimization of a bioreactor with an *E. coli* culture for tryptophan production by using a structured model coupling the oscillating glycolysis and tryptophan synthesis, under review.

165. Maria G (2007) Modelling bistable genetic regulatory circuits under variable volume framework, Chemical and Biochemical Engineering Quarterly 21: 417-434.

166. Maria G (2005) Relations between Apparent and Intrinsic Kinetics of Programmable Drug Release in Human Plasma. Chemical Engineering Science 60(6): 1709-1723.

167. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. Nature 403: 339-342.

168. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335-338.

169. Wikipedia (2017), The free encyclopedia, TCA cycle, Chemostat.

170. Madsen MF, Dano S, Sorensen PG (2005) On the mechanisms of glycolytic oscillations in yeast, FEBS J 272(11): 2648-2660.

171. Franck UF (1980) Feedback kinetics in physicochemical oscillators. Berichte der Bunsengesellschaft für physikalische Chemie 84(4): 334-341.

172. Termonia Y, Ross J (1981a) Oscillations and control features in glycolysis: Numerical analysis of a comprehensive model. Proc Natl Acad Sci USA. 78(5): 2952-2956.

173. Termonia Y, Ross J (1981b) Oscillations and control features in glycolysis: Analysis of resonance effects. Proc Natl Acad Sci USA 78(6): 3563-3566.

174. Slominski A, Semak I, Pisarchik A, Sweatman T, Szczesniewski A, et al. (2002) Conversion of L-tryptophan to serotonin and melatonin in human melanoma cells. FEBS Lett 511(1-3): 102-126.

175. Lubert S (1995) Glycolysis. In: Lubert S, Fourth, (Eds.), Biochemistry. W.H. Freeman & Company, New York, pp. 483-508.

176. Bhartiya S, Chaudhary N, Venkatesh KV, Doyle III FJ (2006) Multiple feedback loop design in the tryptophan regulatory network of *Escherichia coli* suggests a paradigm for robust regulation of processes in series. J R Soc Interface 3(8): 383-391.

177. Mao L, Nicolae A, Oliveira MAP, He F, Hachi S, Fleming RMT (2015) A constraint-based modelling approach to metabolic dysfunction in Parkinson's disease. Computational and Structural Biotechnology 13: 484-491.

178. Zak DE, Vadigepalli R, Gonye GE, Doyle III FJ, Schwaber JS, et al. (2005) Unconventional systems analysis problems in molecular biology: A case study in gene regulatory network modelling, Comp Chem Eng 29(3): 547-563.

179. de Jong H, Geiselmann J, Hernandez C, Page M (2003) Genetic Network Analyzer: qualitative simulation of genetic regulatory networks. Bioinformatics 19(3): 336-344.

180. Casey R, de Jong H, Gouze JL (2006) Piecewise-linear Models of genetic regulatory networks: Equilibria and their stability. J Math Biol 52: 27-56.

181. Klamt S, Stelling J (2002) Combinatorial complexity of pathway analysis in metabolic networks. Mol Biol Rep 29(1-2): 233-236.

182. Hecker M, Lambeck S, Toepfer S, van Someren E, Guthke R (2009) Gene regulatory network inference: Data integration in dynamic models-A review. BioSystems 96(1): 86-103.

183. Wang X, Li Y, Xua X, Wang Y (2010) Toward a system-level understanding of microRNA pathway via mathematical modeling. BioSystems 100: 31-38.

184. Zhdanov VP (2009) Bistability in gene transcription: Interplay of messenger RNA, protein, and nonprotein coding RNA. BioSystems 95: 75-81

185. Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. J Mol Biol 323(5): 785-793.

186. de Jong H (2002) Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. Journal Computational Biology 9(1): 67-103.

187. De Jong H, Gouze JL, Hernandez C, Page M, Sari T, Geiselmann J (2003) Hybrid modelling and simulation of genetic regulatory networks: a qualitative approach, In: Maler O, Pnueli A (Eds.), Hybrid Systems: Computation and Control. HSCC 2003. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. vol 2623.

188. Endy D, Brent R (2001) Modelling cellular behaviour. Nature 409: 391-395.

189. Guantes R, Poyatos JF (2006) Dynamical principles of two-component genetic oscillators. PLoS Computational Biology 2(3): 188-197.

190. Hasty J, McMillen D, Isaacs F, Collins JJ (2001) Computational studies of gene regulatory networks: in numero molecular biology. Nature Reviews Genetics 2: 268-279.

191. Hasty J, Isaacs F, Dolnik M, McMillen D, Collins JJ (2001) Designer gene networks: Towards fundamental cellular control. Chaos 11(1): 207-220.

192. Hasty J, McMillen D, Collins JJ (2002) Engineered gene circuits. Nature 420: 224-230.

193. Isaacs FJ, Hasty J, Cantor CR, Collins JJ (2003) Prediction and measurement of an autoregulatory genetic module. PNAS 100: 7714-7719.

194. Kaern M, Blake WJ Collins JJ (2003) The engineering of gene regulatory networks. Annu Rev Biomed Eng 5: 179-206.

195. Klipp E (2009) Timing matters. FEBS Letters 583: 4013-4018.

196. Mochizuki A (2005) An analytical study of the number of steady states in gene regulatory networks. Journal of Theoretical Biology 236: 291-310.

197. Samad HE, Khammash M, Petzold L, Gillespie D (2005) Stochastic modelling of gene regulatory networks. Int J Robust Nonlinear Control 15: 691-711.

198. Tyson JJ, Albert R, Goldbeter A, Ruoff P, Sible J (2008) Biological switches and clocks. J R Soc Interface 5: S1-S8.

199. Widder S, Schicho J, Schuster P (2007) Dynamic Patterns of Gene Regulation I: Simple Two Gene Systems. J Theor Bio 246(3): 395-419.

200. Liu XM, Xie HZ, Liu LG, Li ZB (2009) Effect of multiplicative and additive noise on genetic transcriptional regulatory mechanism. Physica A 388: 392-398.

201. Kaznessis YN (2011) SynBioSS-Aided Design of Synthetic Biological Constructs. Methods in Enzymology 498: 137-152.

202. Van Someren EP, Wessels LFA, Backer E, Reinders MJT (2003) Multi-criterion optimization for genetic network modelling. Signal Processing 83(4): 763-775.

203. Xu, Z., Zheng, P., Sun, J., Ma, Y. (2013), ReacKnock: Identifying reaction deletion strategies for microbial strain optimization based on genome-scale metabolic network, PLOS ONE, 8(12), e72150.

204. Xu, Z., Sun, X., Yu, S. (2009), Genome-scale analysis to the impact of gene deletion on the metabolism of E. coli: constraint-based simulation approach, BMC Bioinformatics, 10(Suppl 1): S62

205. Xu, Z., Sun, J., Wu, Q., Zhu, D. (2017), Find_tfSBP: find thermodynamics-feasible and smallest balanced pathways with high yield from large-scale metabolic networks, Nature Scientific Reports, doi:10.1038/s41598-017-17552-2

206. Gilbert, D., Heiner, M., Jayaweera, Y., Rohr, C. (2017), Towards dynamic genome-scale models, Briefings in Bioinformatics, pp. 1-14,•Oct. 2017, doi: 10.1093/bib/bbx096