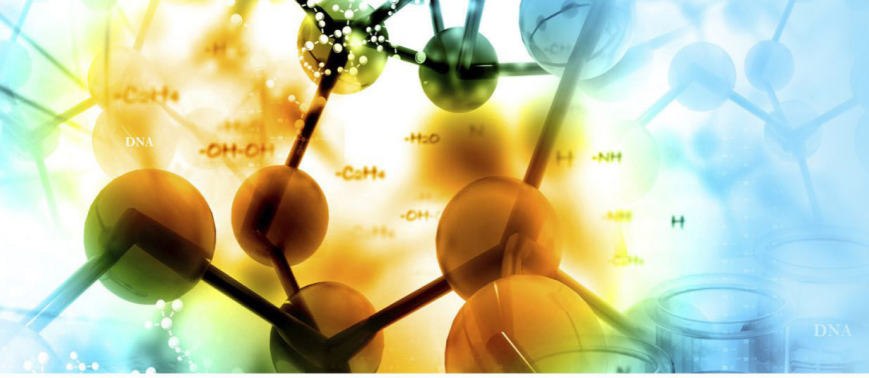


Juniper PUBLISHERS
key to the Researchers

DETERMINISTIC MODELLING APPROACH OF METABOLIC PROCESSES IN LIVING CELLS - A STILL POWERFUL TOOL FOR REPRESENTING THE METABOLIC PROCESSES DYNAMICS

ISBN: 978-1-946628-08-4

Gheorghe Maria

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

Deterministic Modelling Approach of Metabolic Processes in Living Cells - A Still Powerful Tool for Representing the Metabolic Processes Dynamics

Prof. Dr. Gheorghe Maria

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

Romania

Published By :

Juniper publishers Inc.

United States

Date : June 14, 2017

Acknowledgement

Author is grateful to Prof. Dr. Stefan Szedlacsek from the Department of Enzymology of the Institute of Biochemistry of the Romanian Academy Bucharest for the very constructive suggestions and discussions. I am also grateful to the late Prof. Wolf-Dieter Deckwer from GBF-Gesellschaft für Biotechnologische Forschung mbH, Biochemical Engineering Division, Braunschweig, Germany for funding my summer research stay on 2006. Many thanks are also addressed to Prof. Jibin Sun from the Tianjin Institute of Industrial Biotechnology, China for funding my summer research stay on 2010 in Tianjin. Many thanks are also addressed to Prof. An-Ping Zeng from TU Hamburg, Institute of Bioprocess and Biosystems Engineering (Germany) for hosting my summer DAAD fellow on 2009. Finally, I want to address many thanks to my PhD students Diana Ene, Ionela Luţa, Mara Crişan and Andreea Şcoban all being co-authors of some of my works in the bioengineering area.

~ Prof. Dr. Gheorghe Maria

Contents

List of Figures

Abstract

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) is one of the modern tools which uses advanced mathematical simulation models for *in-silico* design of micro-organisms 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 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.

Keywords: Systems biology; Cell metabolism modelling; Deterministic modelling; Gene expression modelling; Genetic regulatory circuits

Abbreviations: GRC: Genetic Regulatory Circuits; TF-s: Transcription Factors; VVWC: Whole-Cell-Variable-Volume; CGE: Gene Circuit Engineering; GERM: Gene Expression Regulatory Modules; CVWC: Constant Volume Whole-Cell; ODE: Ordinary Differential Equations; QSS: Quasi Steady-state; G: Generic Gene; P: Generic Protein; M: mRNA; CCM: Central Carbon Metabolism; GMO: Genetic Modified Organisms; P.I: Performance Indices

Introduction

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. The structural and functional cell organization, including components and reactions, is extremely complex, 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 [1-3]. Relationships between structure, function and regulation in complex cellular networks are better understood at a low (component) level rather than at the highest-level [4].

Cell regulatory and adaptive properties are based on *homeostatic* mechanisms, which maintain quasi-constant key-species concentrations and output levels, by adjusting the synthesis rates, by switching between alternative substrates, or development pathways. Cell regulatory mechanisms include allosteric enzymatic interactions and feedbacks in gene transcription networks, metabolic pathways, signal transduction and other species interactions [5]. In particular, protein synthesis homeostatic regulation includes a multi-cascade control of the gene expression with negative feedback loops and allosteric adjustment of the enzymatic activity [1,6-8].

Cells have a hierarchic organization (structural, functional, and temporal, Figure 1):

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	Source of energy and material	
←Structural Hierarchy→	Hierarchy→	←Functional Hierarchy→	

Figure 1: The hierarchical organization of living cells.

The structural hierarchy

Includes all cell components from simple molecules (nutrients, saccharides, fatty acids, aminoacids, simple metabolites), then macromolecules or complex molecules (lipids, proteins, nucleotides, peptidoglycans, coenzymes, fragments of proteins, nucleosides, nucleic acids, intermediates), and continuing with well-organized nano-structures (membranes, ribosomes, genome, operons, energy harnessing apparatus, replisome, partitioning apparatus, Z-ring, etc. [9]). To ensure self-replication of such a complex structure through enzymatic metabolic reactions using nutrients (Nut), metabolites (Met), and substrates (glucose/fructose, N-source, dissolved oxygen, and micro-elements), all the cell components should be associated with specific functions into the cell, following a

The functional hierarchy

According to the species structure; e.g. sources of energy (ATP, ADP, AMP), reaction intermediates, TF-s. [10] 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, of enzymes' activity, 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 **Synthetic Biology** [11]. This is why, 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 [12]. 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 [13].

The Time Hierarchy

The wide-separation of time constants of the metabolic reactions in the cell systems is called time hierarchy. Thus, the reactions are separated in slow and fast according to their time constant; in fact, only fast and slow reactions are of interest, while the very slow processes are neglected or treated as parameters (such as the external nutrient or metabolite evolution). Aggregate pools (combining fast reactions) are usually used in building-up cell dynamic models in a way that intermediates are produced in a minimum quantity and consumed only by irreversible reactions. All cell processes obey a certain succession of events, while stationary or dynamic perturbations are treated by maintaining the cell components homeostasis (steady-state levels), and by minimizing the recovering or transition times after perturbations.

A central part of such cell models concerns self-regulation of the metabolic processes via GRC-s. Consequently, one particular application of such dynamic deterministic 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 (i.e. desired 'motifs' in response to external stimuli)

[1,14]. Simulation of gene expression and of GRC makes possible *in-silico* design of organisms 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* [15-17]. 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" [18]. 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 [1]. Examples of such GRC modulated functions include:

o 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.

o 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.

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

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

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

The development of dynamic models on a deterministic basis to adequately simulate *in detail* the cell metabolism self-regulation, cell growth, and replication for such an astronomical cell metabolism complexity is practical impossible due to lack of structured information and computational limitations. A review of some trials is presented by Styczynski & Stephanopoulos [19].

In spite of such tremendous modelling difficulties, development of *reduced* dynamic models to adequately reproduce such complex synthesis related to the central carbon metabolism (Figure 2) [18-21], but also to the genetic regulatory system [20] tightly controlling the metabolic processes reported significant progresses over the last decades in spite of the lack of structured experimental kinetic information. In spite of being rather based on sparse information from various sources and unconventional identification /lumping algorithms [1,8], such structured deterministic kinetic models have been proved to be extremely useful for *in-silico* design of novel GRC-s conferring new properties/functions to the mutant cells, that is desired 'motifs' in response to the external stimuli [1].

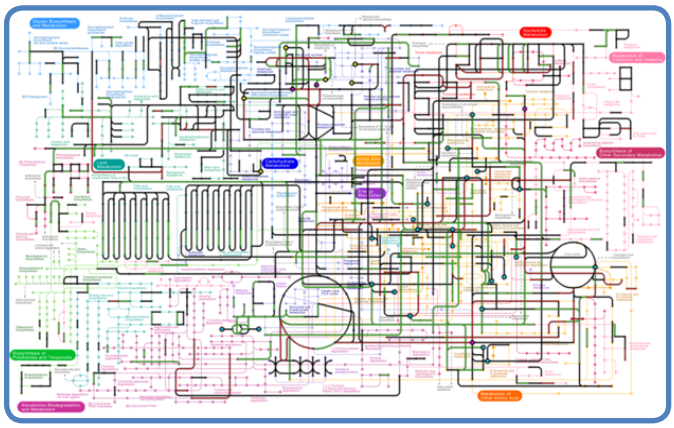


Figure 2: KEGG [34] result: the central carbon metabolic fluxes (metabolic pathway map) of Mycobacterium smegmatis MC2155 (after [2]).

The scope of this paper is to review some novel concepts and (bio) chemical engineering rules applied to modular modelling of gene expression regulatory modules (GERM), GRC-s and other metabolic processes on a deterministic basis by using continuous variable dynamic models under the novel "whole-cell-variable-volume" (VVWC) modelling framework [1].

An Emergent Border Field: Systems Biology

Efforts to understand and to develop mathematical models on a mechanistic (deterministic) basis of the cell metabolic processes started many decades ago, but with modest results. 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?", and after the famous cryptanalyst Alan Turing published on 1952 his paper "The chemical basis of morphogenesis", in Philosophical Transactions of the Royal Society of London (Series B, No. 641, Vol. 237) proving the (bio) chemical reaction basis of metabolic processes. Notable progresses in the structured modelling of GRC and in the study of their regulatory properties have been reported after publication of the book of "General System Theory" by Ludwig von Bertalanffy on 1968.

Amazing, but the first pioneers in dynamic modelling of biological systems were not the (bio)chemical engineers which are better trained to 'translate' 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 on 1952 [22,23]. Later, such 'electronic circuits-like' models have been extensively used to understand intermediate levels of regulation, 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:

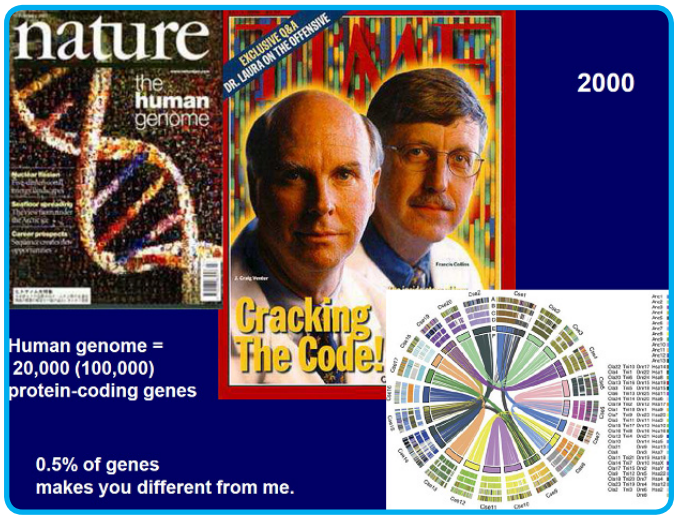
o The dynamic character of species interactions and processes;

o The feedback character of processes ensuring their

o Optimal regulation, with

o Consuming minimum of resources (nutrients/substrates), and cell energy, but ensuring maximum reaction rates.

All these cell metabolic characteristics will be accounted in all the subsequent cell in-silico simulators based on extended mathematical models. All these metabolic process characteristics are also in the agreement of the Darwin theory "Living organisms have evolved to maximize their chances for survival; It explains structures, behaviours of living organisms."



The modelling efforts have intensified a lot after 2000 when the human genome has been deciphered (Figure 3), 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. Because many cell regulatory systems are organized as "modules" [24], it is natural to model GRC-s using a *modular* approach [1]. Further analyses including engineered GRC-s can lead to predict/design desirable cell characteristics, that is [25]: 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 in modified cells 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 [11].

The emergent field of *Synthetic Biology* [26] "interpreted as the engineering-driven building of increasingly complex biological entities" [11], 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 [11,27] 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.

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]. 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 [11].

To help the efforts of the Synthetic Biology to *in-silico* design genetic modified micro-organisms (GMO) with 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 [1], 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 plays a central role, as underlined by the following definition "Systems Biology is 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, president Institute for Systems Biology, Seattle, USA, cited by Banga, 2008). Beside the Institute for Systems Biology in Seattle, a large number of research groups appeared worldwide based on the increased computing power of the new generations of computers [28-32].

Various definitions of Systems Biology exist in the dedicated literature [3]:

o "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).

o System Biology is a comprehensive quantitative analysis of the manner in which all the components of a biological system interact functionally over time (Alan Aderem, Director Inst. Systems Biology, Seattle).

o 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).

o "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" (Mike Mesarovic in System Theory and Biology, 1968).

o "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/);

o "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/);

o "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/)

In the "post-genomic era" a large number of Systems Biology projects have been developed leading to simulate parts of cell metabolism, such as [8]: EcoCyc [33] database; KEGG [34] database; 'Whole-Cell' models (cell organization and dynamics):

a. E-Cell: (compartments, compounds, genes, reactions [35])

b. V-Cell: (model, geometry & applications, biological interface [36])

c. M-Cell: (stochastic simulator of some cell sub-systems [37])

d. A-Cell (‚electrical circuit' models [38])

e. Silicon-Cell: (computer replica of cell processes to be linked [39])

f. Specific programming languages: SBML, JWS [40], etc.

g. Single cell growth (e.g. Escherichia coli, Haemophilus influenzae, Mycoplasma genitalium, yeast, ...)

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

i. Metabolic control of protein synthesis regulation (GERM, GRC)

j. Modelling the cell cycle

k. Modelling the drug release and cell-drug interactions

l. Modelling cellular communications, neuronal transmission

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

Among the milestone works in Systems Biology it is to mention the contributions of some of their pioneers: Heinrich & Schuster [28], Torres & Voit [29], Bowden [30], Brazhnik [31], 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 5).

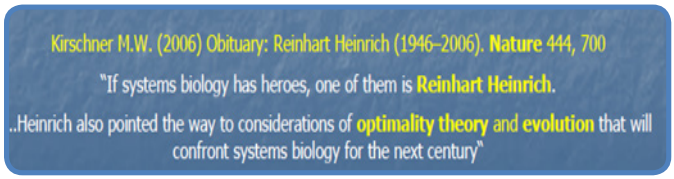


Figure 4: Quoted from the obituary of R Heinrich [3].

Here it is to mention the huge contributions of Reinhart Heinrich (1946-2006) in the field of modelling the regulation of cellular systems [28]. So that, at his obituary on 2006, M.W. Kirschner (Nature 444, 700) said (Figure 4): "If Systems Biology has heroes, one of them is Reinhart Heinrich....."

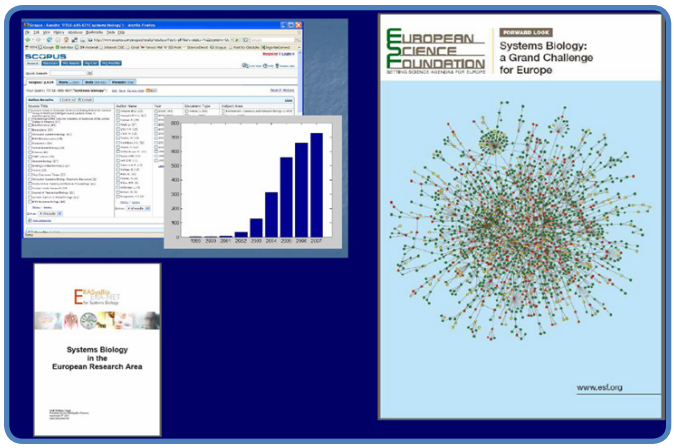


Figure 5: Systems Biology publications and EU programs after 2000.

Tremendous applications of systems biology have been reported over the next decades in the area of [41]:

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

So that, 50 years after the first reported models of living cells [22,23], the optimistic researchers advanced very ambitious targets, such as "Modeling the heart–from genes to cells to the whole organ" [42].

As in all scientific controversies, the are also skeptic opinions, such as [3]

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 were made by the world renowned physiologist Denis Noble in his book" The music of life", 2006:

a. "We must move away from our obsession with genes alone. We must look not at one level, but at the interaction of processes at various levels, from the realm of Systems Biology.

b. The reductionist approach of molecular biology has proved itself immensely powerful. But DNA isn't life."

i. These Systems Biology tools they really work?

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

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

(Bio) Chemical Engineering Deterministic Approach: Rules, Advantages and Limitations

A review of mathematical model types used to describe metabolic processes is presented in [8,19,32]. Each model type presents advantages but also limitations. To model such a complex metabolic regulatory mechanisms at a molecular level, two main approaches have been developed over decades: structure-oriented analysis, and dynamic (kinetic) models [4]. 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 determines the fluxes and metabolic concentrations [28]. 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 vs. the system behaviour.

Originally, MCA has been introduced by Kacser & Burns [8], Heinrich & Rapoport [8] 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 [8]: to study any limit set for non-steady/time-dependent conditions [43,44]; the flux balance analysis and optimization (FBA); elementary mode analysis (EMA); dynamic flux balance analysis (DFBA); extreme pathway analysis (ExPA); constrained based modelling of metabolic network (CBM).

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 [6]. Because the living cells are self-evolutive systems, new reactions recruited by cells together with enzyme adaptations can lead to an increase in the cell biological organisation and to optimal performance indices. When constructing methods to optimize evolutive metabolic systems, MCA concepts and appropriate performance criteria have been used, 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 [28]. 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. [45], or the use of modified control coefficients [46]).

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. "An emergent border field: Systems Biology").

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 [8,47]. Oriented and unified programming languages have been developed (see SBML, JWS, see chap. "An emergent border field: Systems Biology") 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. "An emergent border field: Systems Biology"). Such integrated simulation platforms tend to use a large variety of biological databanks including enzymes, proteins and genes characteristics together with metabolic reactions (CRGM-database [48]; NIH-database [49]).

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 [5,8,50]. Such models can include:

(i) Boolean (discrete) variables; such a topological structure is displayed in the Figure 6 [50]; due to the very large number of states $O(10^3–10^4)$, and $O(10^3)$ of TFs involved in the gene expression, such GRC models are organized in clusters, modules, of a multi-layer organization.

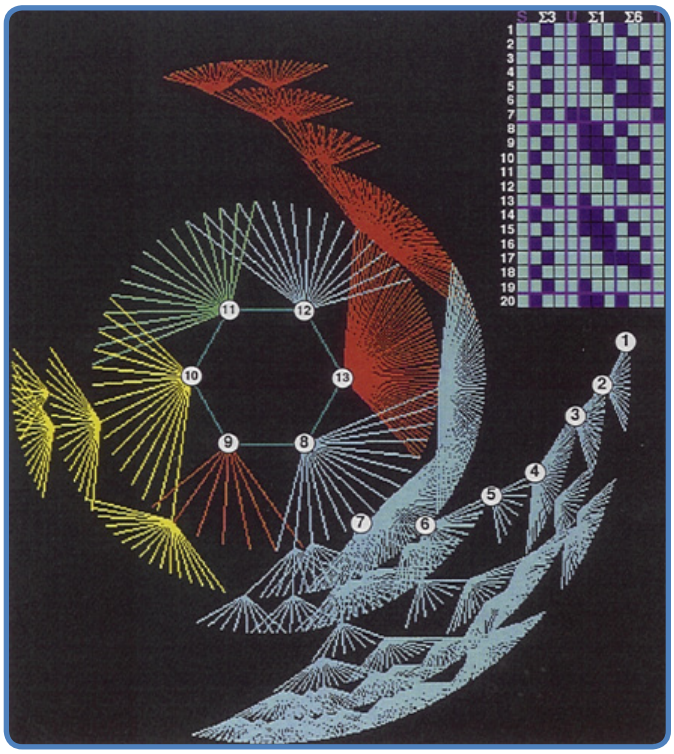


Figure 6: Boolean topological representation of GRC [50].

(ii) 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 [8,28,32,50,51].

(iii) Stochastic variable models [52-54];

(iv) Mixed variable models [50].

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 6) according to simple Boolean relationships, in a finite space. Such a coarse representation is used to obtain a first model for a complex biosystem including a large number of components, until more detailed data on process dynamics become available. 'Electronic circuits' structures (see an example in Figure 7) 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.

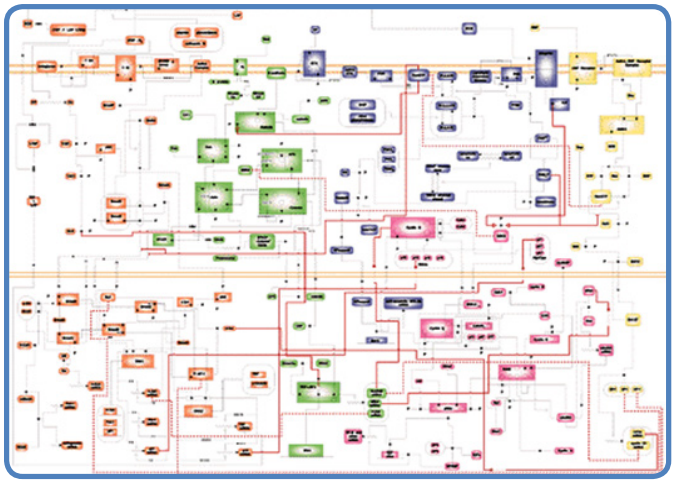


Figure 7: An 'electronic circuit-like' representation of a GRC. A human cancer cell pathway (from http://www.bio-itworld.com/archive/111202/virtual.html)

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 [21] and to reflect the species interconnectivity or perturbation effects on cell growth [1,55]. Mixtures of ordinary differential equation (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 [50]. Representation of metabolic process kinetics is made usually by using rate expressions of extended Michaelis-Menten or Hill type [20,21,55].

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 [52-54].

By applying various modelling routes, successful structured models have been elaborated to simulate various regulatory mechanisms [8,52,56-59]. In fact, as mentioned by Crampin & Schnell [5], 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 [5,61].

Development of deterministic dynamic models to adequately reproduce such complex synthesis related to the central carbon metabolism [20,21] 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 [1,8,61].

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

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

b. Large number of species, reactions, transport parameters, and interaction s

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 trade off between model complexity and adequacy must be maintained [62] to use such models for the *in-silico* design GMO, by *in-silico* re-programming the cell metabolism, or by optimal cell cloning [69,70]. 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 [55,61].

The work with reduced kinetic models of cell 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).

Mathematical Modeling in Molecular Biology Using (Bio) Chemical Engineering Tools

Even if complicated and, often over parameterized, the continuous variable dynamic deterministic ODE models of GRC-s present a significant number of advantages, being able to reproduce in detail the molecular interactions, the cell slow or fast continuous response to exo/endo-geneous continuous perturbations [8,19]. 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 [3,28], accounting for the regulatory system properties, that is: dynamics, feedback/feed forward, and optimality. And, most important, such ODE kinetic modelling approach allows using the strong tools of the classical (bio) chemical engineering modelling, that is [62]:

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) [63];

iv. Set equilibrium reactions; Gibbs free energy balance analysis set cyclic reactions; find species at quasi-steady-state; improved evaluation of steady-state flux distributions that provide important information for metabolic engineering [64], allowing application of ODE model species and/or reaction lumping rules [61].

The ODE deterministic models have been developed in two alternatives:

o The default Constant Volume Whole-Cell (CVWC) classical continuous variable ODE dynamic models, which do 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 frame work eq. (1) 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:

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

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

Where: Cj = (cell-)species j concentration; V = system (cell) volume; nj = species j number of moles; rj = j-th reaction rate; s(i,j) = stoichiometric coefficient of the species "j" (individual or lumped) in the reaction "i"; t = time; j= 1,...,ns = number of cell species (individual or lumped); k = rate constant vector; I= 1,...,nr = number of reactions. The above formulation 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 [3].

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, by distorting very much or even misrepresenting the prediction results, as exemplified by [1].

o As an alternative, Maria [1,8] promoted over the last 15 years 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, while the continuous ODE model was re-written either in terms of species moles or of species concentrations, as following [1]:

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

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

Where:

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

Because:

$$\frac{dC_j}{dt}=\frac{d}{dt}\left(\frac{n_j}{V}\right)=\frac{1}{V}\frac{dn_j}{dt}-C_j\frac{d\left(ln(V)\right)}{dt}=\frac{1}{V}\frac{dn_j}{dt}-DC_j=h_j(C,k,t) \quad \text{.....(3)}$$

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 (2-3) mass-balance formulation is that given by Aris [65] for the (bio) chemical reacting systems of variable-volume.

In the VVWC formulation of the cell dynamic 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 (2-3), that is the Pfeiffers' law of diluted solutions [66] adopted and promoted by Maria [1,8]:

$$V(t)=\frac{RT}{\pi}\sum_{j=1}^{ns}n_j(t) \quad \text{.....(4)}$$

Which, by derivation and division with V leads to [1]:

$$D=\frac{1}{V}\frac{dV}{dt}=\left(\frac{RT}{\pi}\right)\sum_{j}^{ns}\left(\frac{1}{V}\frac{dn_j}{dt}\right) \quad \text{.....(5)}$$

In the above relationships, T=absolute temperature, and R = universal gas constant, V=cell (cytosol) volume. As revealed by the Pfeffer's law eqn. (4) in diluted solutions [66], and by the eq. (5), 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/π) term can be easily deducted in an isotonic cell system, from the fulfilment of the following invariance relationship derived from (4):

$$V(t)=\frac{RT}{\pi}\sum_{j=1}^{ns}n_j(t)\Rightarrow\frac{RT}{\pi}=\frac{V(t)}{\sum_{j=1}^{ns}n_j(t)}=\frac{1}{\sum_{j=1}^{ns}C_j}=\frac{1}{\sum_{j=1}^{ns}C_{jo}}=constant \quad \text{.....(6)}$$

As another observation, from (5) it results that the cell dilution is a complex function D(C,k) being characteristic to each cell and its environmental conditions.

Relationships (5-6) are important constraints imposed to the VVWC cell model (2-3), eventually leading to different simulation results compared to the CVWC cell kinetic models that neglect the cell volume growth and isotonic effects (see an example in [1]).

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 (discussed in detail in [1]), that is:

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

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

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

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

e. Loss of the intrinsic model stability;

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

Mass Balance and State Equations	Remarks
$\frac{dC_j}{dt} = \frac{1}{V}\frac{dn_j}{dt} - D\,C_j = g_j(C,k)$	continuous variable dynamic model representing the cell growing phase (ca. 80% of the cell cycle)
$\frac{1}{V}\frac{dn_j}{dt} = r_j(C,k)\ ;\ j = 1,\ldots,n_s$	
$V(t) = \frac{RT}{\pi}\sum_{j=1}^{n_s} n_j(t)$	Pfeffer's law in diluted solutions
$D = \frac{1}{V}\frac{dV}{dt} = \left(\frac{RT}{\pi}\right)\sum_{j}^{n_s}\left(\frac{1}{V}\frac{dn_j}{dt}\right)$	D = cell content dilution rate = cell volume logarithmic growing rate
$\frac{RT}{\pi} = \frac{V}{\sum_{j=1}^{n_s} n_j} = \frac{1}{\sum_{j=1}^{n_s} C_j} = \frac{1}{\sum_{j=1}^{n_s} C_{jo}}$ constant.	constant osmotic pressure (π) constraint
$\left(\sum_{j}^{all} C_j\right)_{cyt} = \left(\sum_{j}^{all} C_j\right)_{env}$	Derived from the isotonic osmolarity constraint
Hypotheses:	
a. Negligible inner-cell gradients.	
b. Open cell system of uniform content.	
c. Semi-permeable membrane, of negligible volume and resistance to nutrient diffusion, following the cell growing dynamics.	
d. Constant osmotic pressure (the same in cytosol "cyt" and environment "env"), ensuring the membrane integrity ($\pi_{cyt} = \pi_{env}$ = constant).	
e. Nutrient and overall environment species concentration remain unchanged over a cell cycle t_c.	
f. Logarithmic growing rate of average $D_s = \ln(2)/t_c$; volume growth of ; $V = V_o e^{D_s t}$; t_c = duration of the cell cycle.	
g. Homeostatic stationary growth of $\left(dC_j/dt\right)_s = g_j(C_s,k) = 0$.	
h. Perturbations in cell volume are induced by variations in species copynumbers under the isotonic osmolarity constraint: $V_{perturb}/V = \left(\sum n_j\right)_{perturb} / \left(\sum n_j\right)$.	

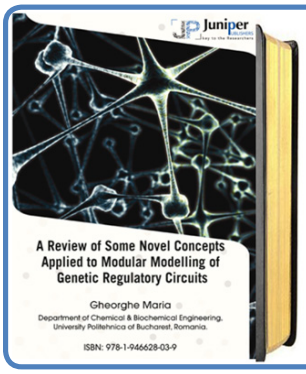
Notations: T = absolute temperature; R = universal gas constant; V= cell (cytosol) volume; π = osmotic pressure; C_j = cell species j concentration; n_j = species j number of moles; r_j = j-th reaction rate; t = time; k =rate constant vector; "s" index indicates the stationary state.

Figure 8: The variable cell-volume whole-cell (VVWC) dynamic modelling framework and its basic hypotheses [1,8].

The basic equations and hypotheses of a VVWC model are presented in Figure 8. Even if all cell regulation mechanisms are not fully understood, metabolic regulation at a low-level is generally better clarified. Based on that, conventional (*deterministic*) dynamic models based on ODE kinetics using continuous variables, approached in this paper, based on a *mechanistic description* of cell reactions taking place among individual species [including proteins, mRNA, DNA, transcription factors TF-s, intermediates, etc.] proved to be a convenient route to analyse continuous metabolic/regulatory processes and perturbations. When systems are too large or poorly understood, coarser and more phenomenological kinetic models may be postulated (e.g. protein complexes, metabolite channelling, etc.). In dynamic models, only essential reactions are retained, species and reactions often being included as lumps, the model complexity depending on measurable variables and available information. Such reduced VVWC kinetic models can be useful to analyse the regulatory cell-functions, and the treatment of both stationary and dynamic perturbations, the cell cycles, oscillatory metabolic paths, and lot of cell biosyntheses related to the central carbon metabolism [1], by reflecting the species interconnectivity or perturbation effects on cell growth.

Modular Modelling of Genetic Regulatory Circuits

One of the very successfully application of VVWC deterministic models with continuous variables is those of simulating the regulatory properties of the individual gene expression regulatory modules (GERM), and of the genetic regulatory circuits (GRC) comprising a certain number of linked GERM-s (no more than 23-25 [13]). A review of the systematic and comprehensive approaches in modelling the dynamics of the GRC-s based on VVWC deterministic models and bio-chemical engineering concepts and principles was presented by Maria in his work (Figure 9) [1].



Published By:
Juniper publishers
1890 W Hillcrest Dr, Newbury Park,
CA 91320,
(United States)
juniperpublishers.com
Date: March 15, 2017

Figure 9: The cover of the ebook of Maria [1], https://juniperpublishers.com/ebook-info.php.

Why the GRC are important to be understood and simulate their properties? That is because the cell core metabolism is ensured by the optimized GERM-s and GRC-s that maintain the optimized protein (enzymes) synthesis and, thus, a balanced cell metabolism and an equilibrated cell growth despite the continuous perturbations in the environment, by also ensuring the cell evolution and competitiveness eventually by gene mutations [3]. It is here to mention only some of the GRC functions:

a. Cell metabolism regulation via hierarchically organized GRC (key-proteins being the regulatory nodes),

b. Sustain cell homeostasis, and a balanced cell growth, under variable environmental conditions (nutrients, substrates),

c. Preserve the holistic and local GRC regulatory properties,

i. Ensure Self-regulation of cell Self-replication,

ii. Ensure fast cell response to environmental perturbations,

iii. Ensure fast metabolic reactions with low level of intermediates,

iv. Ensure optimized metabolic fluxes (stationary reaction rates),

v. Ensure quick recovery of QSS (homeostasis) after a dynamic (impulse-like) environmental perturbation,

vi. Ensure quick transitions between QSS-s after a stationary (step-like) perturbations,

vii. Ensure a cascade-control of GERM and GRC regulation,

viii. Ensure a low QSS sensitivity vs. perturbations

And all these should be accomplished by fulfilling a certain number of constraints, that is:

o By using minimum amounts of substrates, and nutrients,

o By using minimum cell energy A(M)(D)TP, NAD(P)H, FADH(2),

o By maintaining quasi-constant key-species concentrations and output levels,

o By quickly adjusting the synthesis rates,

o By switching between alternative substrates, or development pathways by means of genetic switches.

Applications of GERM chain dynamic simulators in *Synthetic Biology* field are immediate, as long as GRC-s controlling the cell metabolism allows *in-silico* re-programming the cell metabolism by means of modified GRC properties leading to GMO of desirable characteristics. Among essential GRC structures used in this respect are to be mentioned the genetic-switches (decision-making branch points between on/off states according to the presence of inducers), oscillators (cell systems evolving among two or several quasi-steady-states), signal/external stimuli amplifiers, amplitude filters, signal transduction circuits (specific treatment of external signals by controlled gene expression), etc. Modular construction of GRC-s must account for some individual (local) but also for holistic properties of the cell considered by the whole-cell modelling approach [1], such as: 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 [67].

The gene expression is a highly self and mutually regulated process catalysed by the produced enzymes/effectors. Simple generic representations of a gene expression regulatory module (GERM) are given in the Figure 10 (for a generic pair P/G self-catalysed synthesis of the protein P and of its encoding gene G). Such representations include the essential nutrient lumps (NutP, NutG) used to obtain the protein and DNA precursor metabolites (MetP, MetG) respectively, and intermediates (R,R') involved in the reactions controlling the transcriptional and translation steps of the P synthesis. The module nomenclature of such GERM models (Figure 11), proposed by [8,75] is those of [L1(O1)n1,..., Li(Oi)ni], and includes the assembled regulatory units Li(Oi) ni. One unit i is formed by the component Li (e.g. enzymes or even genes G, P, M, etc.) at which regulatory element acts, and ni=0,1,2,... number of 'effector'/(transcriptional factors(TF) species Oi (i.e. 'effectors' P, PP, PPPP, R, RR, RRRR, etc) binding the 'catalyst' L. For instance, a [G(P)2] unit of Figure 11 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. $\sum_{i=0}^{2}[G(P_i)]$ = 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 Figure 10 & 11 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.

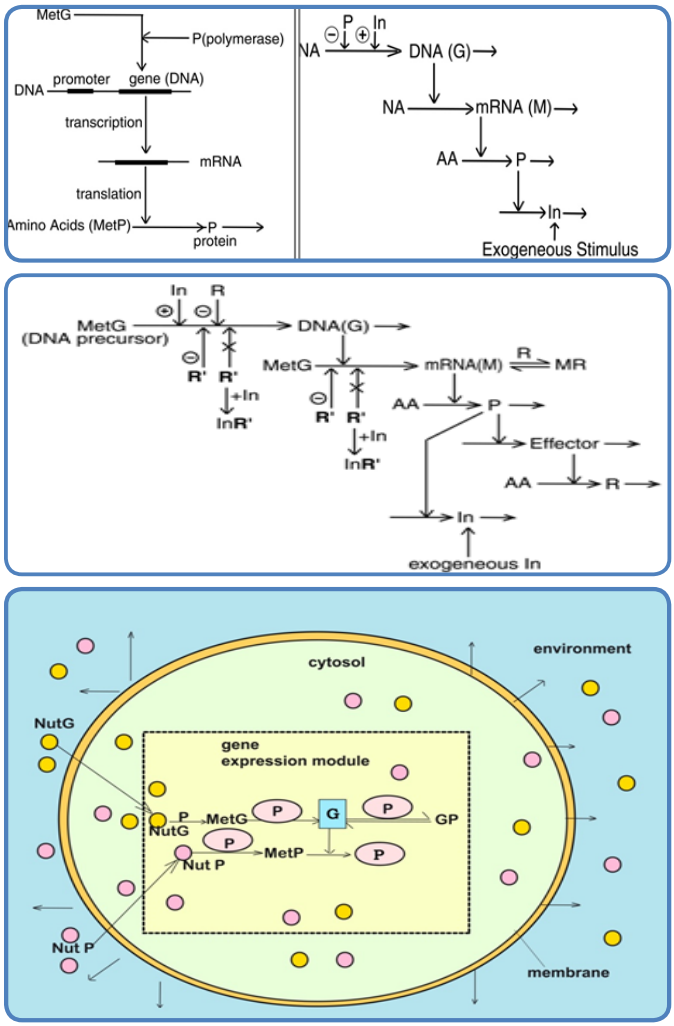


Figure 10: Simplified representations of a regulatory module (GERM) for a generic gene G expression (up-left, down) with perfectly coupled enzyme/regulator P expression. Down-right is a GERM of [G(P)1] type. Such GERM models are further used to construct various GRC models. Notations: In=inducer; AA= aminoacids; horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; G= gene encoding protein P; M = mRNA; R, R' = transcriptional factors (repressors); MetG=DNA precursor metabolites. The enzyme (protein P) interacts with the inducer In for controlling the transcription rate by means of feedback; +/ - positive or negative regulatory loops [1].

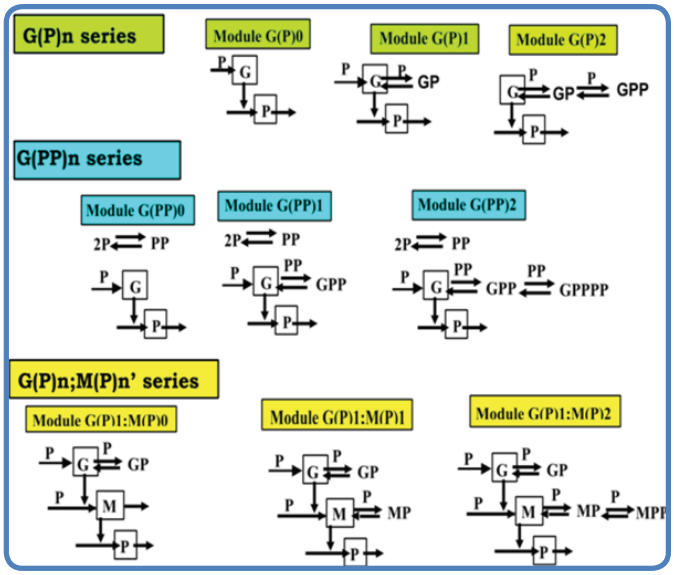


Figure 11: Various types of GERM regulatory modules for protein synthesis [1]. Horizontal arrows indicate reactions; vertical arrows indicate catalytic actions; absence of a substrate or product indicates an assumed concentration invariance of these species.

The GERM model structure can be extended according to experimental information and accounting for individual or lumped species. For instance, at a generic level, in the simplest representation (Figure 10, up), the protein (P) synthesis rate can be adjusted by the 'catalytic' action of the encoding gene (G) (Figure 11, down). The catalyst activity is in turn allosterically regulated by means of 'effector' molecules (P, or PP; Figure 11) reversibly binding the catalyst G (DNA) or M (mRNA) via fast and reversible reactions (the so-called 'buffering' reactions). These simple regulation schemes can be further detailed in order to better reproduce the experimental data, with the expense of a supplementary effort to identify the module kinetic parameters. For instance, a two-step cascade control of P-synthesis model also includes the M=mRNA transcript encoding P (Figure 11, down). The effector (R), of which synthesis is controlled by the target protein P (Figure 10), can allosterically adjust the activity of G and M, i.e. the catalysts for the transcription and translation steps of the gene expression. In such a cascade scheme, the rate of the ultimate reaction is amplified, depending on the number of cascade levels and catalysis rates. More complex regulatory modules have been elaborated [1], and used in developing genetic regulatory circuits (GRC) following a similar route to 'translate' from the 'language' of molecular biology to that of mechanistic chemistry, by preserving the structural hierarchy and component functions. Once elaborated, such a modular structure can be modelled by using a continuous variable ODE kinetic model under a VVWC framework, and then analysed as functional efficiency by means of some quantitative performance indices (P.I.-s) below shortly described.

As the cell regulatory systems are module-based organized, complex feed-back and feed-forward loops are employed for self- or cross-activation/repression of interconnected GERM-s, leading to different interaction alternatives (directly/inversely, perfect/incomplete, coupled/uncoupled connections) of a gene with up to 23-25 other genes." [13], to ensure the key-species homeostasis, holistic and local regulatory properties of the enzymatic reactions. While Maria and others [1] used reduced GERM structures of 10-14 reactions, that ensures a satisfactory trade off between model simplicity and its predictive quality [67], more sophisticated constructions are proposed in the literature [1,20], such as the GRC controlling the lac operon expression in *E. coli*, including 40 reactions and 27 species (reduced model) or 70 reactions and 50 species (extended model) [25]. Eventually, the advantage of such a modular approach is the possibility to adapt the model size according to the available information, or to use the same GERM structure to model several gene expressions. Modular approach can also be useful in simulating the hierarchical organization of the cell regulatory networks.

As discussed by Maria [1,8], fast buffering reactions, like G + P <===> GP, are close to equilibrium and have little effect on metabolic control coefficients. As a consequence, rate constants of such rapid reactions are much higher than those of the core synthesis and dilution rates [1]. By contrast to CVWC modelling, the mechanistic based GERM models [1] in a VVWC framework seem to be more robust, flexible and easily adaptable to different case studies (examples in [1]). The model rate constants and some unknown species concentrations are estimated from solving the nonlinear ODE model equations written for the quasi-stationary conditions (homeostasis) [1], by using the measured species stationary concentrations, with also imposing some optimal properties of the cell system [1]. The GERM regulatory properties (and P.I.) are defined for two types of environmental (in a nutrient) or internal (in a metabolite, protein) perturbations [1]:

o *Dynamic* (impulse-like) perturbations. The GERM properties are: stability strength (minimum changes of the homeostasis vs. dynamic perturbations); dynamic efficiency (species minimum recovering times after an impulse perturbation); species interconnectivity (species minimum average and standard deviation of the recovering times after a dynamic perturbation); robustness (species minimum recovering times vs. rate constants); dynamic efficiency (species minimum recovering times after a dynamic perturbation);

o *Stationary* (step-like) perturbations. The GERM properties are: system stability strength (minimum changes of the homeostasis vs. stationary perturbations); responsiveness (minimum transition times toward a new steady-state); sensitivity (minimum sensitivity coefficients of homeostatic states vs. environmental variables); robustness (minimum sensitivity of the steady-state vs. rate constants);

Each of the mentioned regulatory attributes has also a mathematical formulation [1]. As proved by [1], when elaborating the kinetic model for whatever GERM, VVWC formulation by including the cell-volume growing rate is an essential issue to account for, due to several reasons. Beside, the simplified modular representation of GERM and GRC allow studying their regulatory properties and characterization of their performance indices (P.I.) by using the nonlinear system theory. Thus, by studying

the GERM properties of such simple formulations of Figure 11, several conclusions can be derived [1]:

o The continuous dilution of the cell content, that is concentration decline due to the continuous increase of the denominator of C=n(t)/V(t), where C is the species concentration, n=species number of moles, V=cell volume, t= time; in spite of that, concentrations of key species remain constant because the numerator (copy numbers) increases at the same rate with the denominator;

o The regulatory efficiency of the GERM increases with the number of effectors from its structure, that is: the number of buffering reactions G + P <===> GP, for the G(P)n GERM types (Figure 11), the number of buffering reactions G + PP <===> GPP, for the G(PP)n GERM types (Figure 11), the number of buffering reactions M + P <===> MP, for the [G(P)1;M(P)n] GERM types (Figure 11);

o For an efficient GERM, the species recovering trajectories after a dynamic perturbation toward the steady-state (in the phase diagram) are more direct and straight, thus less disturbing the cell metabolic reactions;

o The VVWC model representation correctly reproduces the system homeostasis, that is the species quasi-constant concentrations because both nominator and denominator of the fraction C=n(t)/V(t) are doubling at the same rate. By contrast, the CVWC model predictions are wrong, the predicted species concentration dynamics C(t) having the same shape and relative growth as with those of the copy numbers n(t) trajectories. 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 (discussed in [1]): the influence of the cell ballast in smoothing the homeostasis perturbations; the secondary perturbations transmitted via cell volume following a primary perturbation; the more realistic evaluation of GERM P.I.-s, and of the recovering/transient times after perturbations; loss of the intrinsic model stability; loss of the self-regulatory properties after a dynamic perturbation, etc.

o The system isotonicity constraint eq. (4-6) is an essential part of such GRC constructions to better reflect their properties. In such a VVWC formulation, all cell species should be considered (individually or lumped), because all species net reaction rates contribute to the cell volume increase (eq. 5). As the cell volume is doubling during the cell cycle, this continuous volume variation cannot be neglected. 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). The system isotonicity is also responsible for the indirect effect of perturbations in concentrations on the cell-metabolism transmitted via induced changes in the volume growing rate (the so-called "secondary perturbation" following the first one);

o The mutual autocatalysis in the GERM constructions appears to *interconnect* 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;

o The 'big volume' or cell big content creates a 'ballast' effect, leading to an increased cell homeostatic (steady-state) concentrations 'resistance' vs. small perturbations in the level of some internal or external components. Thus, the whole-cell content ballast has an essential influence in smoothing the effect of internal/external perturbations on the system homeostasis.

o The GERM complexity when constructing a GRC is also important. It has been proved that cooperative linking of GERM (giving specific role and function of each protein inside the cell) is more efficient, because the system stability is strengthened, while species inter-connectivity is increased leading to a better treatment of perturbations. More important than the number of species considered in the individual GERM-s is the used of a cascade control of the GERM efficiency, that is [G(P)n; M(P)m] structures, dimeric TF-s (e.g. PP) instead of monomeric ones, and an allosteric enzyme activity control.

o Maria [1, 76, 67] proved there exists an optimal level of the TF-s that are associated to the optimal holistic regulatory properties of the GRC (low sensitivity vs. external nutrients, but high vs. inducers).

o 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 [1]) is 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 discussed in [1] should be fulfilled. 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 [76].

Some Examples of Modular Deterministic Dynamic Models

In this chapter some simple examples of using modular deterministic dynamic models are presented.

A whole-cell model to simulate mercuric ion reduction by *E. coli* under stationary or perturbed conditions

One worthy example of applying VVWC models to adequately represent complex modular GRC-s, is the structured model proposed by Maria [68-70] to simulate the dynamics of the *mer*-operon expression in Gram-negative bacteria (like *E. coli, Pseudomonas sp.*) to uptake the mercury ions from wastewaters under various environmental conditions. The model was constructed and validated by using literature experimental data, but also qualitative information [77] on mer-operon characteristics.

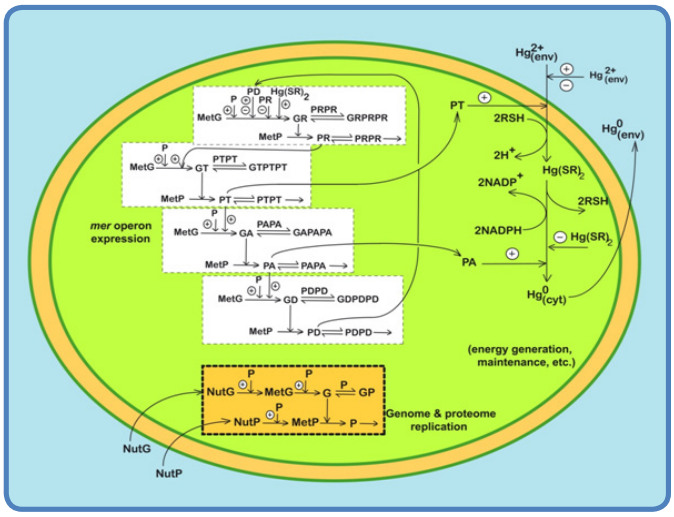


Figure 12: The whole-cell model of Maria [68-70] in the VVWC approach used to simulate the reduction of $Hg^{(2+)}$ ions from environment to volatile Hg(0) in *E. coli bacteri*a. The simplified reaction path includes: Two modules for mediated transport of $Hg^{(2+)}$ into cytosol (catalysed by the enzyme PT) and its reduction (catalysed by the enzyme PA); Five regulatory modules of *mer* operon expression including successive synthesis of the enzyme PR (the transcriptional activator of other protein synthesis), lumped PT permease, PA reductase, and of the control protein PD; One module for the lumped proteome P and genome G replication of [G(P)1] type. The regulatory system is placed in a growing cell, by mimicking the homeostasis and cell response to stationary and dynamic perturbations in the environmental $[Hg^{2+}]$. The reductant NADPH and RSH are considered in excess into the cell. Figure adapted from [68-70].

Notations: P=lumped proteome; G=lumped genome; NutG, NutP = lumped nutrients used for gene and protein synthesis; MetG/MetP=lumped metabolome (DNA or protein precursor); P• = proteins; G• = genes; RSH= low molecular mass cytosolic thiol redox buffer (such as glutathione); perpendicular arrows on the reaction path indicate the catalytic activation, repressing or inhibition actions; absence of a substrate or product indicates an assumed concentration invariance of these species; ± positive or negative feedback regulatory loops.

Bacteria resistance to mercury is one of the most studied metallic-ion uptake and release process (see the review of [77] due to its immediate large-scale application for mercury removal from industrial wastewaters [78]. 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. A genetic regulatory circuit responsible for the mer-operon expression controls the whole process, by including 4 lumped genes (denoted by GR,GT,GA,GD in Figure 12) 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 too 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 (Figure 12).

One additional GERM module is added to the whole-cell model to simulate the lumped proteome P and genome G replication of [G(P)1] type. The regulatory system of 7 GERM-s is placed in a growing cell, by mimicking the *E. coli* cell homeostasis and its response to stationary and dynamic perturbations in the environmental $Hg^{(2+)}$. The reductant NADPH and RSH are considered in excess into the cell.

In-silico design of a genetic modified *E. coli* cell to concomitantly maximize the production of biomass and succinate

Beside simulation of cell GRC-s, one of the objectives of the structured deterministic modular cell simulators is to identify genome modifications leading to the improvement of some of the cell characteristics.

For instance, Maria [71] used a reduced *E. coli* cell whole-cell reduced dynamic model from literature including 95 reactions and 72 metabolites (Figure 13), obtained by the reduction of an extended deterministic model of 720 reactions and 436 metabolites to *in-silico* determine what genes should be removed (the so-called 'gene knockout' procedure) to realize the maximization of both biomass and succinate production. The optimization problem has been solved by using a mixed integer nonlinear programming method (MINLP) [71]. Being an optimization problem with two contrary objectives, an elegant option is to obtain the set of Pareto optimal solutions, also called Pareto front for the case of at least two adverse objectives. A Pareto solution is one where any improvement in one objective can only take place at the cost of another objective. The Pareto-front procedure was applied by also accounting for the stoichiometric constraints. The problem presents multiple solutions indicating concomitant removal of 2-4 genes (indicated in parentheses in the down-right plot of Figure 13, together with the Pareto-front).

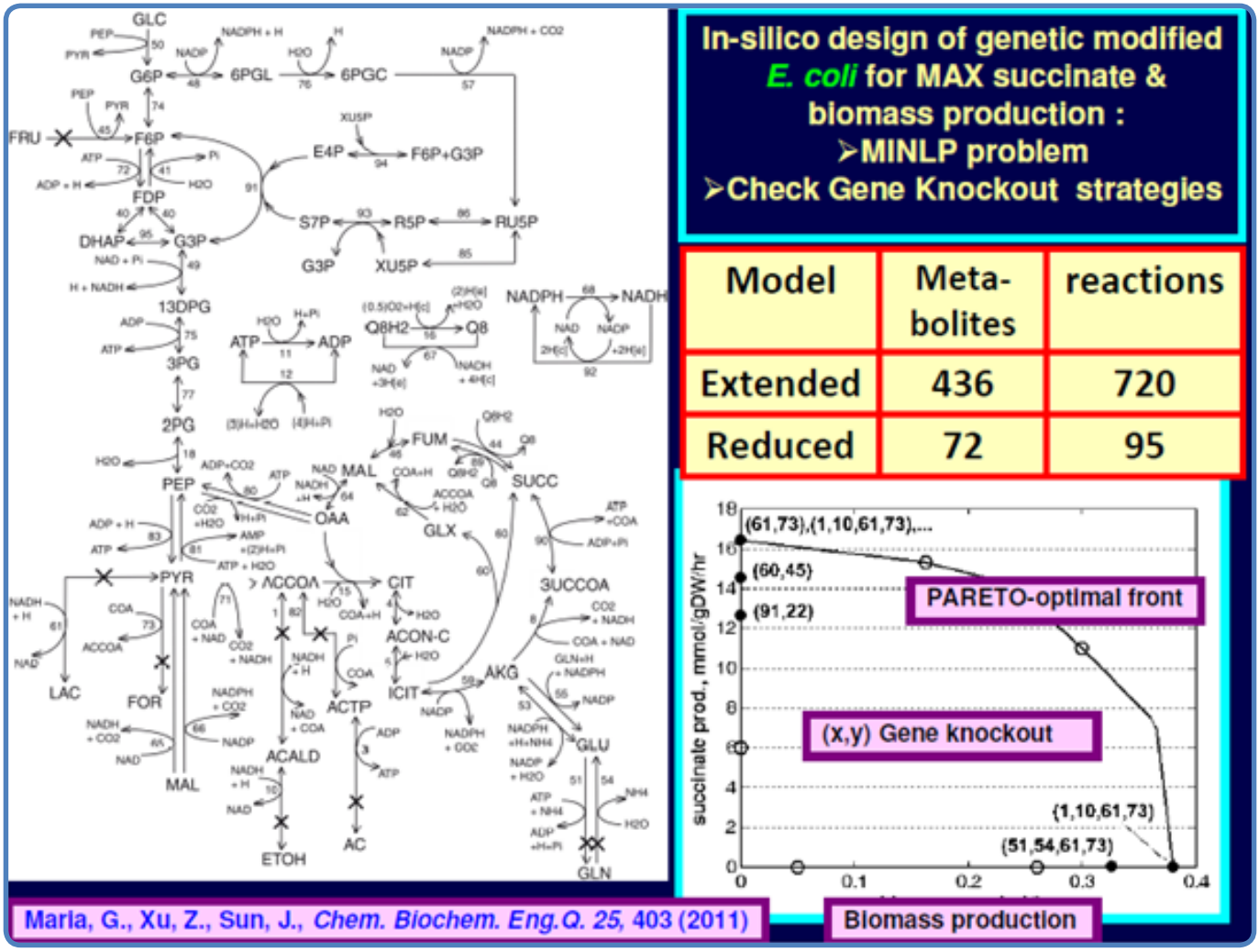


Figure 13: *In-silico* design of a genetic modified *E. coli* cell to concomitantly maximize the production of biomass and succinate (the Pareto-optimal front method, below-right). In the parentheses are the deleted gene numbers from the genome (see the corresponding reaction in the left scheme). The used structured reduced model is those of Edwards and Palsson [72]. See the computing details in [71].

In-silico study of glycolytic oscillations occurrence in *E. coli* cell by using a reduced modular deterministic model

To in-silico design GMO it is indispensable to dispose of a valuable whole-cell simulator of the central carbon metabolism. The central carbon metabolism (CCM, see KEGG [34]) 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) 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 [79-82]. 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 [83].

One of the most studied modules of the CCM is the glycolysis. In this context, 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 (denoted by mTRM in Figure 14, right), obtained from reducing the Chassagnole et al. [73] model (Figure 14, left), Maria [21] has simulated the conditions leading to the occurrence of a stable oscillating glycolysis in the E. coli cells (experimentally highlighted by Madsen et al. [79]).

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 [79]. 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.

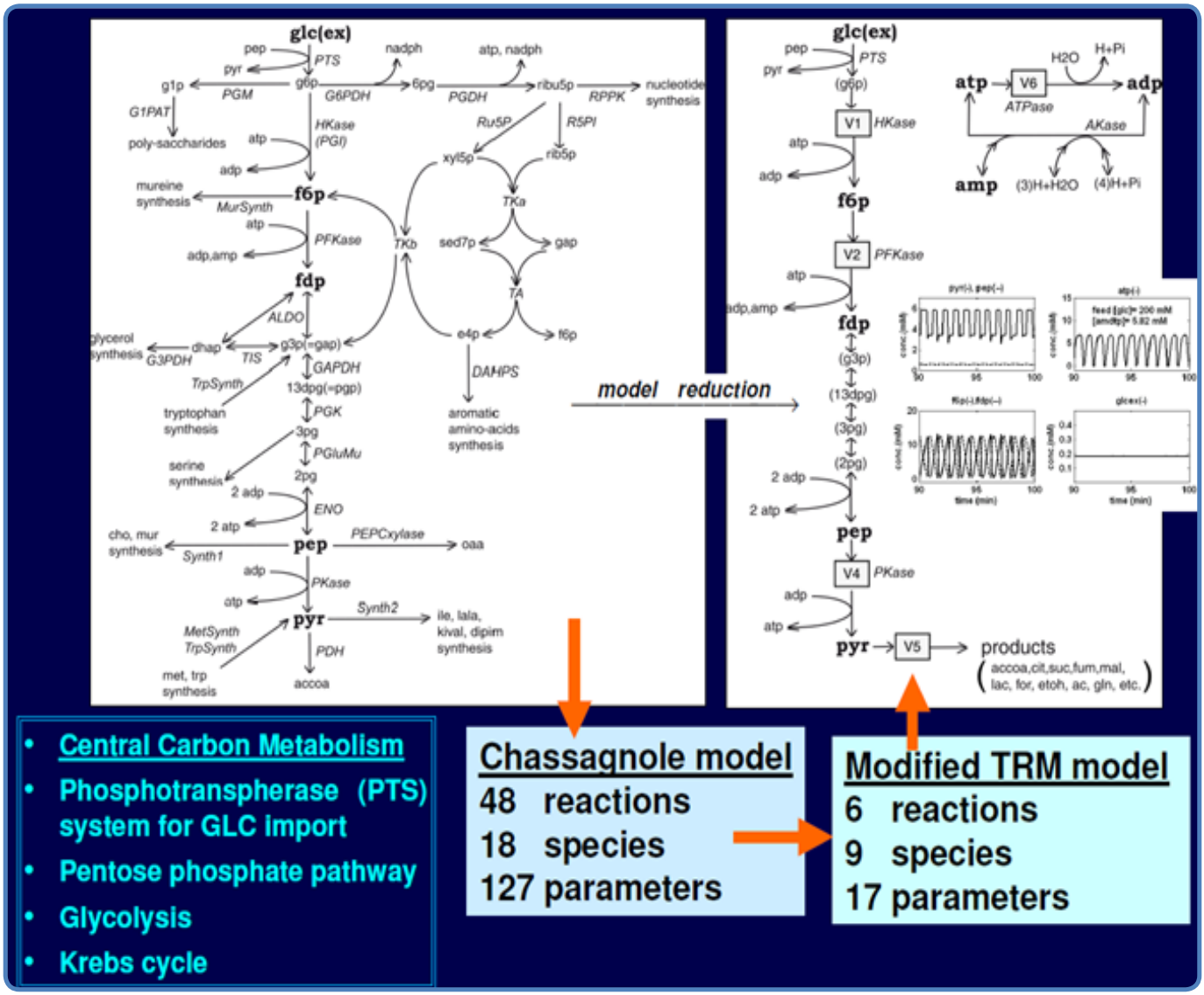


Figure 14: *In-silico* study of glycolytic oscillations occurrence in *E. coli* cell by using the mTRM model of Maria [21] (right) obtained by reducing the Chassagnole et al. [73].

According to Franck [74], spontaneous occurrence of self-sustained oscillations in chemical systems is due the coupled actions of at least two simultaneous processes (Figure 16). 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 [74].

In the glycolytic system case, extensive experiments (Figure 15, right plot, [79]) have revealed that self-sustained oscillations are reported in a broad range of cell types [79]. As revealed by Termonia & Ross [80] glycolytic oscillations occurrence is due to the antagonistic action of two processes on regulating the V2 reaction rate that converts F6P in FDP (Figure 15, left). The glycolytic oscillation occurrence and characteristics (period) are influenced by both external (environmental) and internal (genomic) factors, that is [81,82]:

i. From one side it is the glucose (Glc) import driving force through the phosphotransferase (PTS)-system (Figure 15) 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 [82].

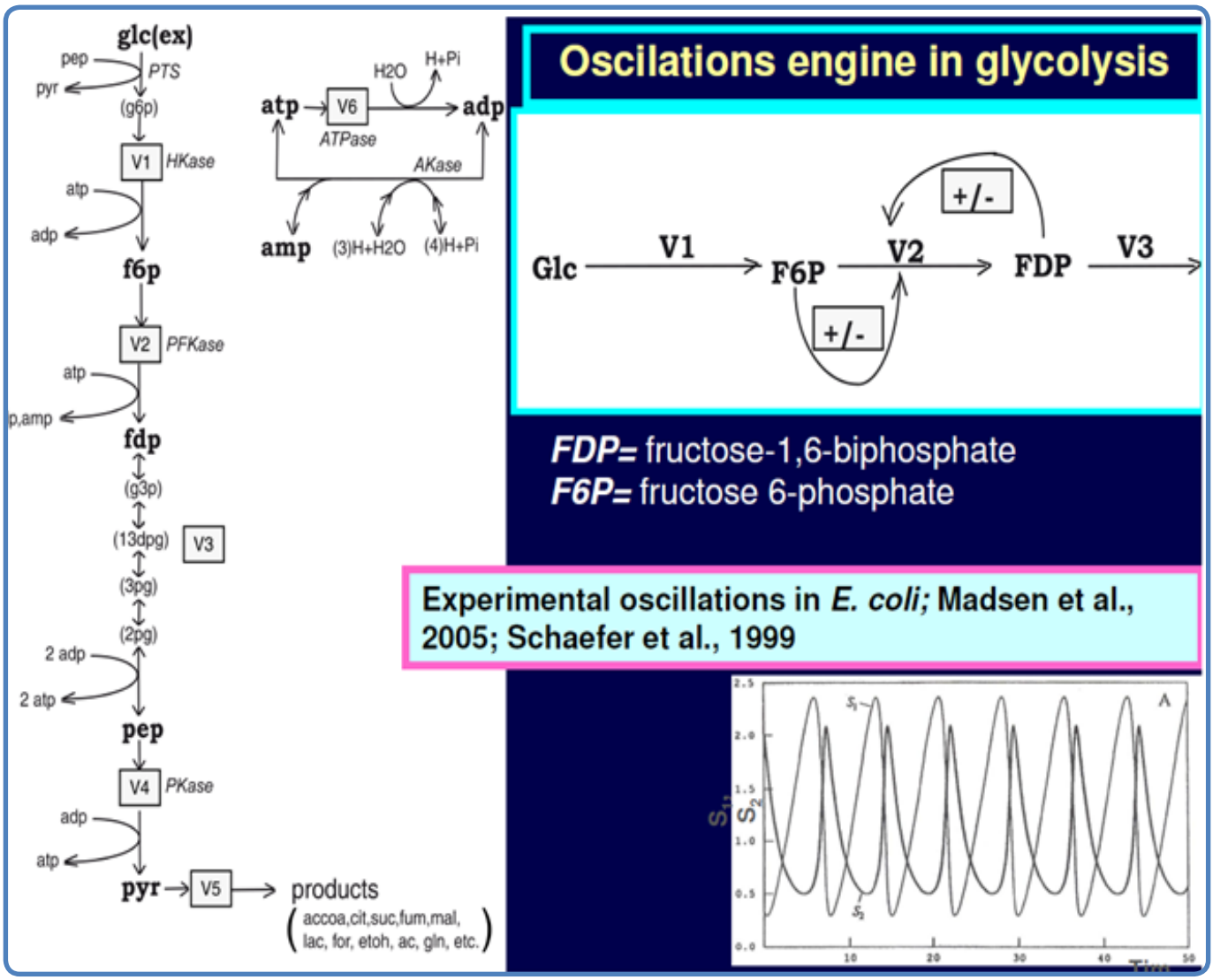


Figure 15: Chemical node inducing glycolytic oscillations [21]. +/ - denotes the feedback positive or negative regulatory loops. Glc=glucose; F6P= fructose-6-phosphate; FDP = fructose-1,6-biphosphate; V1-V3 = reaction rates belonging to the glycolysis reduced model (left) of Maria [21].

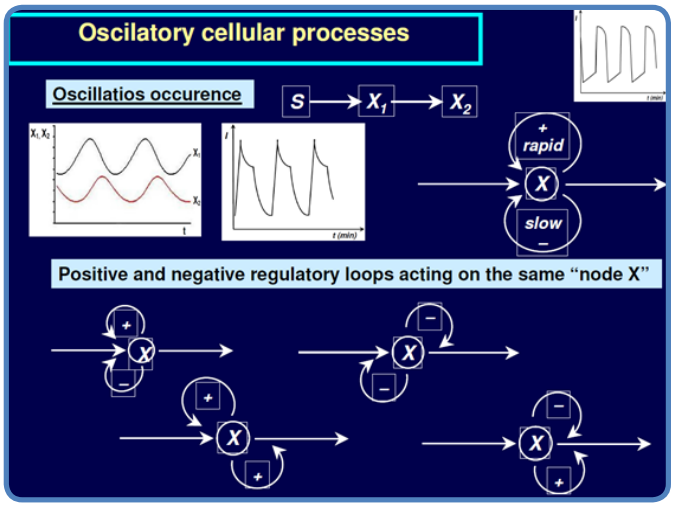


Figure 16: Oscillation occurrence in chemical systems Franck [74]. ± denotes the feedback positive or negative regulatory loops. X is a generic species denoting the engine node.

Conclusions

As a general conclusion, the bio-chemical engineering principles and modelling rules are fully applicable to modelling cellular metabolic processes. 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 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* [1,8] allows *in-silico* design of GRC inducing specific cell motifs of genetically modified micro-organisms (GMO). Examples includes [1]: Genetic switches of adjustable certainty, sensitivity to exo-/endogeneous stimuli, responsivity, regulatory efficiency; Metabolism behavior of wild or cloned cells with plasmids, with potential applications in medicine, such as therapy of diseases (gene therapy), new devices based on cell-cell communicators, biosensors, production of vaccines, etc.

Even if the cell simulators still present lot of drawbacks and present a limited adequacy, they become more and more valuable tools in designing GMO with desirable characteristics, or for obtaining micro-organisms cloned with desirable plasmids with important applications in industry (new biotechnological processes, optimization of bioreactors, production of vaccines), or in medicine. As mentioned by G.E.P Box (Professor of statistics at the University of Wisconsin, and a pioneering the areas of quality control, time series analysis, design of experiments, and Bayesian inference): "All models are wrong, but some are useful."

References

1. Maria G (2017) A review of some novel concepts applied to modular modelling of genetic regulatory circuits, Series: Current Trends in Biomedical Engineering & Biosciences. Juniper publishing, USA.
2. Baloni P, Padiadpu J, Singh A, Gupta KR, Chandra N (2014) Identifying feasible metabolic routes in Mycobacterium smegmatis and possible alterations under diverse nutrient conditions. BMC Microbiol 14: 276.
3. Banga J (2008) Optimization in computational systems biology BMC Syst Biol 2: 47.
4. 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.
5. Crampin EJ, Schnell S (2004) New approaches to modelling and analysis of biochemical reactions, pathways and networks. Prog Biophys Mol Biol 86(1): 1-4.
6. Kholodenko BN, Schuster S, Garcia J, Westerhoff HV, Cascante M (1998) Control analysis of metabolic systems involving quasi-equilibrium reactions. Biochimica Biophysica Acta 1379: 337-352.
7. 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.
8. Maria G (2005) Modular-based modelling of protein synthesis regulation. Chemical and Biochemical Engineering Quarterly 19: 213-233.
9. Lodish H, Berk A, Matsudaira P, Kaise CA, Krieger M, et al. (2000) Molecular cell biology. Freeman & Co, New York, USA.
10. Sauro HM, Kholodenko BN (2004) Quantitative analysis of signalling networks. Prog Biophys Mol Biol 86(1): 5-43.
11. Heinemann M, Panke S (2006) Synthetic Biology - putting engineering into biology. Bioinformatics 22: 2790-2799.
12. Sotiropoulos V, Kaznessis YN (2007) Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes. BMC Syst Biol 9: 1-7.
13. 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 USA 101: 8414-8419.
14. Myers CJ (2009) Engineering genetic circuits, Chapman & Hall/CRC Press, Boca Raton, USA.
15. Bailey JE (1991) Towards a science of metabolic engineering. Science 252(501): 1668-1675.
16. Nielsen J (1998) Metabolic engineering, techniques for analysis of targets for genetic manipulations. Biotechnol Bioeng 58(2-3): 125-132.
17. Stephanopoulos GN, Aristidou AA, Nielsen J (1998) Metabolic Engineering. Principles and Methodologies. Academic Press, San Diego, CA.
18. Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ (2004) Optimal re-design of primarymetabolism in *Escherichia coli* using linlog kinetics. Metabolic Engineering 6(4): 378-390.
19. Styczynski MP Stephanopoulos G (2005) Overview of computational methods for the inference of gene regulatory networks. Comp Chem Eng 29: 519-534.
20. Maria G (2014) Extended repression mechanisms in modelling bistable genetic switches of adjustable characteristics within a variable cell volume modelling framework. Chemical and Biochemical Engineering Quarterly 28(1): 83-99.
21. Maria G (2014) Insilico derivation of a reduced kinetic model for stationary or oscillating glycolysis in *Escherichia coli* bacterium. Chemical and Biochemical Engineering Quarterly 28(4): 509-529.
22. Hodgkin AL, Huxley AF (1952) A quantitative description of ion currents and its applications to conductions and excitation in nerve membranes. J Physiol (Lond) 117: 500-544.
23. Noble D (1952) A modification of the Hodgkin-Huxley equations applicable to Purkinje fibre action and pace-maker potentials. J Physiol 160: 317-352.
24. Kholodenko BN, Kiyatkin A, Bruggeman FJ, Sontag E, Westerhoff HV, et al. (2002) Untangling the wires: A strategy to trace functional interactions in signalling and gene networks. Proc Natl Acad Sci USA 99: 12841-12846.
25. Kaznessis YN (2006) Multi-scale models for gene network engineering. Chemical Engineering Science 61:940-953.
26. Benner SA, Sismour AM (2005) Synthetic biology. Nature Reviews Genetics 6: 533-543.
27. Voit EO (2005) Smooth bistable S-systems. IEE Proc. Syst Biol 152: 207-213.
28. Heinrich R, Schuster S (1996) The regulation of cellular systems. Chapman & Hall, London, UK.
29. Torres NV, Voit EO (2002) Pathway analysis and optimization in metabolic engineering. Cambridge University Press, Cambridge (MS), USA.
30. Cornish-Bowden A (2016) Biochemical evolution- The pursuit of perfection. Garland Science, New York, USA.
31. Brazhnik P, de la Fuente A, Mendes P (2002) Gene networks: how to put the function in genomics. Trends in Biotechnology 20(11): 467-472.
32. Stephanopoulos GN, Aristidou AA, Nielsen J (1998) Metabolic Engineering. Principles and Methodologies. Academic Press, San Diego, USA.
33. EcoCyc (2005) Encyclopedia of *Escherichia coli* K-12 genes and metabolism, SRI Intl., The Institute for Genomic Research, Univ. of California at San Diego.
34. KEGG (2011) Kyoto encyclopedia of genes and genomes, Kanehisa Laboratories, Bioinformatics Center of Kyoto University, USA.
35. Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, et al. (1999) E-Cell: Software environment for whole cell simulation. Bioinformatics 15(1): 72-84.
36. 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.

37. Bartol TM, Stiles JR (2002) MCell: A general Monte Carlo simulator of cellular microphysiology, The Salk Institute, Computational Neurobiology Lab.

38. Ichikawa K (2000) A-Cell: a platform for the construction of biochemical reaction and electrical equivalent circuit models in a biological cell and a neuron, Juki Xerox Co. Ltd., Kanagawa (JP).

39. Westerhoff HV (2006) Engineering life processes live: the Silicon cell. ESCAPE-16 Conference, Garmisch-Partenkirchen, Germany.

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

41. Nand SK (2017) BioXYZ, what is XYZ? Curr Trends Biomedical Eng & Biosci 3(2): 555606.

42. Noble D (2002) Modeling the heart-from genes to cells to the whole organ. Science 295(5560): 678.

43. Markus M, Hess B (1990) Control of metabolic oscillations: unpredictability, critical slowing down, optimal stability and hysteresis, in: Cornish-Bowden A, Cárdenas M (Eds.), Control of metabolic processes, Plenum Press, New York, USA. pp. 303-313.

44. Acerenza L, Sauro HM, Kacser H (1989) Control analysis of time-dependent metabolic systems. J Theor Biol 137(4): 423-444.

45. 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: 671-679.

46. Szedlacsek ES, Aricescu AR, Havsteen BH (1996) Time-dependent control of metabolic systems by external effectors. J theor Biol 182(3): 341-350.

47. 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.

48. CRGM-Database (2002) Centro di Ricerca in Genetica Molecolare, Istituto di Istologia e Embriologia Generale, University of Bologna.

49. NIH Database (2004) National Institute of Health database.

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

51. Chen T, He HL, Church GM (1999) Modeling gene expression with differential equations. Pacific Symposium on Biocomputing 4: 29-40.

52. McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A 94(3): 814-819.

53. Shimizu TS, Bray D (2002) Computational cell biology - The stochastic approach, MIT Press Math6X9, Cambridge, USA.

54. Gillespie DT, Mangel M (1981) Conditioned averages in chemical kinetics. Jl Chemical Physics 75: 704-709.

55. Maria G, Szedlacsek S, Laslo AC Zaharia E (2005) Modular-based simulation of regulated protein synthesis, Conference on Enzyme Reaction Engineering, Dubrovnik (Croatia), Sept. 21.

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

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

58. Heinrich R, Rapoport SM, Rapoport TA (1977) Metabolic regulation and mathematical models. Prog Biophys Mol Biol 32(1): 1-82.

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

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

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

62. Maria G (2017) Application of (bio) chemical engineering principles and lumping analysis in modelling the living systems, Current Trends in Biomedical Engineering & Biosciences 1(4).

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

64. 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. Biotechnology and Bioengineering 110(3): 914-923.

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

66. Wallwork SC, Grant DJW (1977) Physical Chemistry. Longman, London, UK.

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

68. 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.

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

70. 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.

71. Maria G, Xu, Z, Sun J (2011) Investigating alternatives to in-silico find optimal fluxes and theoretical gene knockout strategies for *E. coli* cell. Chemical & Biochemical Engineering Quarterly 25(4): 403-424.

72. Edwards JS, Palsson BO (2000) The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. Proc Natl Acad Sci USA 97(10): 5528.

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

74. Franck UF (1980) Feedback kinetics in physicochemical oscillators. Ber Bunsenges Phys Chem 84: 334-341.

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

76. 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.

77. Barkay T, Miller SM, Summers AO (2003) Bacterial mercury resistance from atoms to ecosystems. FEMS Microbiology Reviews 27: 355-384.

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

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

80. Termonia Y, Ross J (1998) Oscillations and control features in glycolysis: Analysis of resonance effects, Proceedings of the National Academy of Sciences of the USA,

81. Westermark PO, Lansner A (2003) A model of phosphofructokinase and glycolytic oscillations in the pancreatic b-cell. Biophysical Journal 85(1): 126-139.

82. Bier M, Teusink B, Kholodenko BN, Westerhoff HV (1996) Control analysis of glycolytic oscillations. Biophysical Chemistry 62: 15-24.

83. Wikipedia (edited on 2 May 2017) The free encyclopedia, TCA cycle.