

Biochemical Modelling Tools and Applications to Metabolic Engineering

[Biyokimyasal Modelleme Gereçleri ve Metabolik Mühendislik Uygulamaları]

Işıl Aksan Kurnaz

Yeditepe Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Genetik ve Biyomühendislik Bölümü

Yazışma Adresi

[Correspondence Address]

Yeditepe Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Genetik ve Biyomühendislik Bölümü, 34755 Kayışdağı, İstanbul

Tel: (0216) 578 04 10

Faks: (0216) 578 04 00

iakurnaz@yeditepe.edu.tr

Kayıt tarihi 01.10.2004; kabul tarihi 25.02.2005
[Received 01.10.2004; accepted 25.02.2005]

ÖZET

Metabolik mühendislik, biyolojik sistemlerin sanayi uygulamalarına yönelik iyileştirilmesi amacıyla metabolik yolların manipülasyonudur ve günümüzde bu amaç için kullanılan en yaygın yöntem rekombinant DNA teknolojisidir. Ancak canlı bir sistemi etkili biçimde değiştirebilmek için ilk adım o sistemin çalışma prensiplerini anlamak olmalıdır. Özellikle genom projelerinin ve büyük-ölçekli genomik ve proteomik araştırma teknolojilerinin gelişmesiyle beraber anlaşılmıştır ki laboratuvar koşullarında her bir molekülün işlevini anlamak teknik açıdan oldukça zorlayıcı olacaktır. Canlı sistemlerin in silico araştırmaları için kullanılan hesaplamalı modelleme araçları, metabolik yolları kontrol eden anahtar kontrolcü sistemleri ve genetik devreleri tanımlamakta ve organizmanın genomundaki değişikliklerin olası sonuçlarını tahmin etmekte cazip bir alternatif oluşturmaktadır. Bu derlemede, bazı popüler modelleme araçları tanıtılacak, ve özellikle Gepasi programı avantajları ve dezavantajları ile tartışılacaktır. Aynı zamanda bu program kullanılarak yapılmış bazı modeller ve bu tarz benzetim ortamlarının metabolik mühendisliği uygulamalarındaki potansiyelinden bahsedilecektir.

Anahtar Kelimeler: metabolik mühendislik, hesaplamalı modelleme, Gepasi, biyokimyasal benzetim, genetik devreler

ABSTRACT

Metabolic engineering is the manipulation of metabolic pathways in order to improve biological systems for industrial applications, and recombinant DNA technology is the most commonly used tool today for this purpose. However, in order to effectively alter a living system, the first key step should be to understand the inner workings of that system. Especially with the advent of genome projects and new technology for large-scale genomic and proteomic studies, it becomes clear that understanding the function of every component under laboratory conditions will prove technically challenging. Computational modeling tools for the in silico study of living systems offer a nice alternative for identification of key regulatory pathways and genetic circuits that control metabolic pathways, in order to predict potential outcomes of alterations to the organism's genome. In this review, it is aimed to introduce some of the popular modeling tools, and concentrate on one particular program, Gepasi, with its advantages and disadvantages. I will also discuss some of the models generated using this model, and discuss the potential of such simulation environments in metabolic engineering.

Key Words: metabolic engineering, computational modeling, Gepasi, biochemical simulation, genetic circuitry

INTRODUCTION

One of the most important goals of post-genomic biology is the understanding of the fundamental working principles of living systems. With the increased use of the computational tools and the exponentially growing number of completely sequenced genomes, we are ever more in need of identifying each cellular component and their structure-function relationship as well as the interactions thereof (the interactome). At the end of the day, these interactions are what makes cells respond to external signals and adapt to their environment.

The general norm at one time was the study of intermediary and secondary metabolism, and it was widely believed that if you start out with analyzing and understanding single components *in vitro*, it would be possible to reconstitute living systems from their individual components. We now know that the behavior of these components may vary significantly under different conditions. Although we know quite a bit about the individual kinetics enzymes and metabolites in certain experimental systems, still a lot is missing, and we certainly know very little about the control mechanisms and the integration of various pathways. Therefore it becomes extremely difficult to predict the cellular responses to changes in the environment, or during development. In addition, we have to consider is the genetic control circuitry and signaling pathways that either directly or indirectly regulate the metabolic pathways before we set out to manipulate and exploit these pathways in the context of metabolic engineering.

Considering that even a small bacterium such as *Escherichia coli* has around 4500 genes in its genome, it should be clear that putting all these genes and gene products, along with metabolites and metabolic pathways, within the context of a "minimal genome", such as that of the *Mycoplasma genitalium*, becomes a daunting task (1). Consider engineering of a 6000-gene yeast or a 35,000-gene human red blood cell, and the task becomes even more challenging (2).

At this stage, what the researchers and companies choose to do is to conduct *in silico* experiments, in other words perform the experiment in computer simulations and predict potential outcomes of any modifications of the organism. In addition to saving time and money, simulations can also help reveal novel interactions or relations in the intracellular pathways. To understand intracellular pathways using computational approach, most scientists study parts of specific biochemical pathways, such as the cell cycle, replication, glycolysis, MAPK signaling, etc. The ultimate goal is to construct an integrated model that represents the entire cellular content, where metabolism, genetic circuitry and cellular signaling components are all represented, allowing for the discussion of key regulatory units prone to modification (4). This review intends to give a brief overview of the biochemistry and genetics of the cell, and introduce some simulation platforms

and their potential uses in metabolic engineering applications.

GENETIC EVENTS AND HOW THEY LINK TO CONTROL OF METABOLISM

Genetic networks are represented by ensembles of interacting genes to one another and to transcription factors, which activate or repress other genes in the network. The products of these genes may in turn act on other target genes, which ultimately act on structural genes. Thus the genetic network can be considered as a switchboard of genes turning each other on and off. This switchboard enables the cell to respond to changing conditions, which the cell senses through a complex network of what are known as "signal transduction pathways" (3, 5-6).

The cell-to-cell and environment-to-cell communication, and the following intracellular transfer of information (biological activation/ inhibition) occurs through these signalling pathway. In each signal transduction system, an activation/ inhibition signal from a biologically active molecule (hormone, neurotransmitter, ligand, extracellular matrix attachment) is mediated via the coupling of usually an enzyme receptor to downstream intracellular components, eventually leading to the regulation of transcription factors. Signaling pathways and transcription factors together also regulate the accessibility of the DNA through chromatin modifications, the dense packaging of eukaryotic genomes (7).

Technological developments such as cDNA microarrays, real-time PCR analysis and large-scale proteomic studies have increased the rate of massively parallel biological data acquisition, opening the door to a more complex understanding of molecular biology. In the light of these data, it will be possible to both determine the roles of individual genes, and to study cells as a complex network of biochemical factors. Along with metabolic models, it should be possible to construct an integrated model representing the complex cellular system, to identify key regulatory patterns and modify them for the purposes of metabolic engineering by improving certain metabolic pathways. In the future, engineered control of cellular function should be possible through the design of synthetic genetic networks (6).

QUANTITATIVE MODELS OF CELLULAR EVENTS

Computers have been used for modeling since 1940s, but their impact on molecular biology has never been more pronounced than in this post-genomic era. Nowadays computational models are used to check or validate a proposed model based on *in vitro* measurements, to make predictions based on this model, to be later confirmed by laboratory experiments, for example the dynamic simulation of the metabolic network in the human red blood cell (2, 4). The applications are also varied:

In silico models can be used in the rational design of improved metabolic pathways with industrial significance, large-scale screening of novel drug targets for the pharmaceutical industry, design of optimized metabolic pathways producing a particular recombinant drug in vitro using metabolic engineering tools, identification of possible side-effects of novel drugs in model systems, design of customized genetic circuits for the biotechnology industry, identification of minimal genomes and many others (8-10).

Computational whole-cell models can be created using a variety of simulation platforms, most of which can be obtained free of charge for academic use, while others have to be purchased in order to enjoy the full capabilities of the system. I will be introducing some of these programs below, and give examples on the program of preference, which are by no means the only tools available for biochemical modeling. These programs almost inevitably rely on the mathematical depictions of biochemical systems that constitute a particular system.

Most biochemical reactions follow the Michaelis-Menten kinetic scheme, which is based on differential equations defining the change in each component of the system. Most quantitative models try to solve these equations using a variety of different methods, using enzyme concentration, metabolite concentration, reaction rate and the so-called Michaelis-Menten constants as reaction parameters (3). Enzymes themselves can be controlled in three major ways: their amount, their activity, and their compartmentalization. Since enzymes are encoded by the genome of the organism, genetic control mechanisms should also be considered when constructing a computer model of the pathway under study (4).

1. MATLAB simulation tool (<http://www.mathworks.com>)

Matlab (Matrix laboratory) is an interactive software system for numerical computations using matrices and vectors: solving systems of linear equations, computing eigenvalues and eigenvectors, factoring matrices, and so on. It also has a variety of graphical capabilities, and can be extended through programs written in its own programming language. Programs also exist which extend MatLab's capabilities, such as the solution of initial value problems for ordinary differential equations.

Programs using MatLab can be used to simulate and analyze biochemical kinetics, such as metabolic drug-drug interactions (9), using several built-in modules such as the ode (ordinary differential equations) solver. MatLab consists of a command window, and metabolic pathways can be simulated as several functions that define and solve the differential equations defining a metabolic system, after which the results can be depicted as graphs (see <http://www.mathworks.com> for a demo). You are advised to download a MatLab tutorial from the web or refer to a manual in order to get acquainted with the pro-

gramming involved.

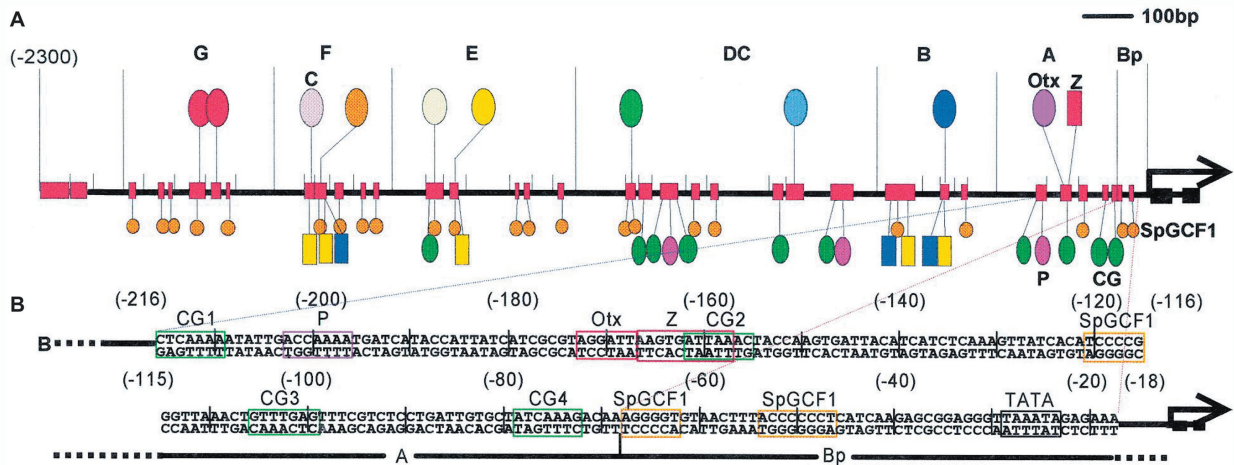
A genetic toggle switch in *Escherichia coli* was designed in MatLab environment (11) and experimentally verified, demonstrating the useability of such self-contained, programmable, genetic circuits can be used in biotechnology, biocomputing and gene therapy. Another example of how MatLab can be used to understand biological systems is the simulation of genetic regulatory network of the *endo16* promoter (12, Fig.1). By transforming the mechanism of regulation on this promoter into mathematical terms (Fig.2), it was shown that one of the regulatory modules was important in spatial regulation of gene expression, the other 6 modules involved in the regulation of the network itself (12).

2. Jarnac (<http://64.17.162.114/software/jarnac.htm>)

Jarnac is a language for simulating cellular models that aid in the analysis of metabolic pathways, signal transduction cascades and gene networks, by performing arithmetic operations, including matrix and vector arithmetic, numerical differentiation and integration and as a function or data grapher, available from the Systems Biology web site (see <http://www.sys-bio.org>). Metabolic support is similar to that found in SCAMP, a previous version of Jarnac, but more extensive and much more flexible. In principle there is no limit to the number of reactions or metabolites a model can contain but there will be practical limits imposed by the computing hardware. The built-in computational support features dynamic simulation, steady state analysis, stability analysis, matrix arithmetic, metabolic control analysis, metabolic structural analysis, among many others.

3. Virtual Cell (<http://www.nrcam.uchc.edu/>)

The National Resource for Cell Analysis and Modeling (NRCAM), developer of the Virtual Cell Modeling and Simulation Framework, is a national resource center supported by the National Center for Research Resources (NCR), at the National Institutes of Health (NIH). The Virtual Cell is a simulation tool designed for researchers from fields as wide as experimental cell biologists and biochemists to computational scientists and theoretical biologists. The Virtual Cell allows the user to build complex models with a web-based Java interface, where compartmental topology and geometry, molecular characteristics, and relevant interaction parameters can be specified to be used in further simulations. The Virtual Cell automatically converts the biological description into a corresponding mathematical system of ordinary and/or partial differential equations. The Virtual Cell will then generate the appropriate code in order to solve these equations. Simulation results are then displayed in various formats. The Virtual Cell can be used to simulate a number of different pathways, including Ca^{++} signaling, IP_3 signaling, G-protein cascade, and many others (13, 14).



C Module A functions:

- Vegetal plate expression in early development:
- Synergism with modules B and G enhancing endoderm expression in later development:
- Repression in ectoderm (modules E and F) and skeletogenic mesenchyme (module DC):
- Modules E, F and DC with LiCl treatment:

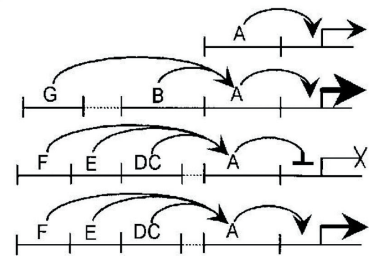


Figure 1. The regulatory modules in the endo16 gene promoter (courtesy of Yuh et al and AAAS). (A) A cartoon diagram showing different modules with various transcription factor binding sites. (B) DNA sequence of the regulatory module A. (C) A summary of the mechanism of action of the A module during various stages of development.

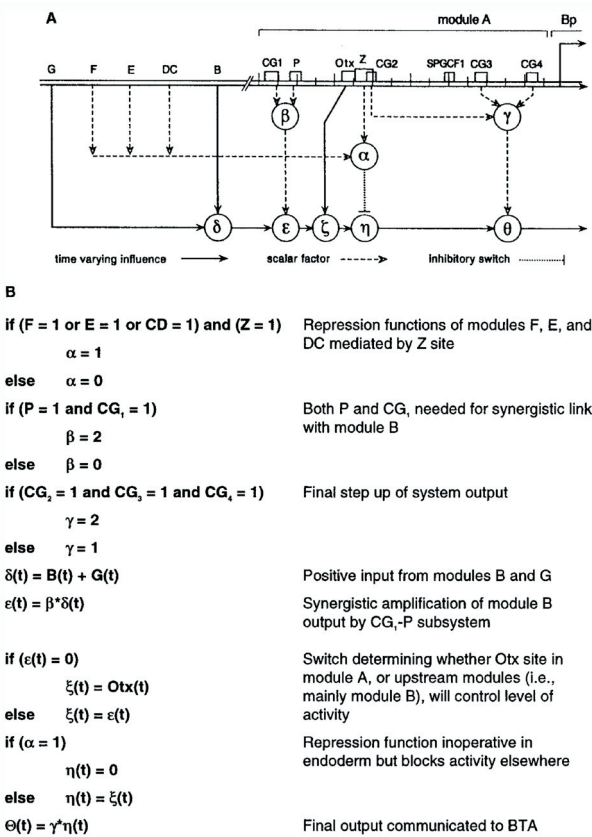


Figure 2. The mathematical formulation of the endo16 regulatory mechanism (courtesy of Yuh et al and AAAS). (A) A schematic of the modular interactions on the endo16 promoter. (B) Mathematical depiction of the regulatory interactions as used in MatLab simulations.

4. E-Cell (<http://www.e-cell.org/>)

E-CELL is a computer software environment for modeling and simulation of the cell, a generic object-oriented environment for simulating molecular processes in user-definable models. Graphical interfaces are provided to allow observation and interaction throughout the simulation process. Written in C++, E-CELL provides a unified, object-oriented framework for modeling and simulation of the complex interactions among the gene products of completed genomes, and can be used to study diverse cellular processes such as gene expression, signaling and metabolism, to construct a cell model for conducting experiments in silico. Using E-CELL, a hypothetical, minimal cell model was generated based on the gene set of *Mycoplasma genitalium*, the self-replicating microorganism (1), and a minimal set of genes required for independent growth was identified computationally (8). E-Cell version 3.1.102 was recently released (see web site, <http://www.e-cell.org>). E-cell uses a main window for the entry of reactions and entities, including the respective compartments, and displays the simulation outcome in Gnuplot (see Gepasi below).

5. Gepasi (<http://www.gepasi.org>)

Gepasi is a Microsoft Windows-based program intended for the simulation of the kinetics of systems of chemi-

cal and biochemical reactions. Written in C++, it simulates the kinetics of biochemical systems, optimize any function of the model, and perform metabolic control analysis (15-18). Gepasi is a user-friendly program that can be obtained free-of-charge, and assists the user in converting the biochemical reactions into mathematical representations. The user supplies the program with information about the stoichiometric structure of the pathway, kinetics of each reaction, volumes of the compartments and initial concentration of all chemical species. The program then builds the differential equations that govern the behaviour of the system and solves them. The data can be plotted in 2D and 3D using the graphical representation package Gnuplot that is distributed with the program itself (see official Gepasi webpage for downloads).

Being initially intended for education purposes and classroom simulations, Gepasi is extremely easy to use even for the computer-unfriendly, so long as they are well-equipped with the required biochemical background. It can theoretically hold unlimited number of reactions, the only constraint being the technical configuration of your computer hardware. The program has so far been used to simulate a number of different cellular processes, including MAPK signal transduction, melanogenesis, and chromatin modifications of immediate-early gene promoters, inhibition mechanisms of HIV proteinase, or simply nonlinear regression analysis (7, 10, 17, 19-21). The advised pattern of study for any biochemical

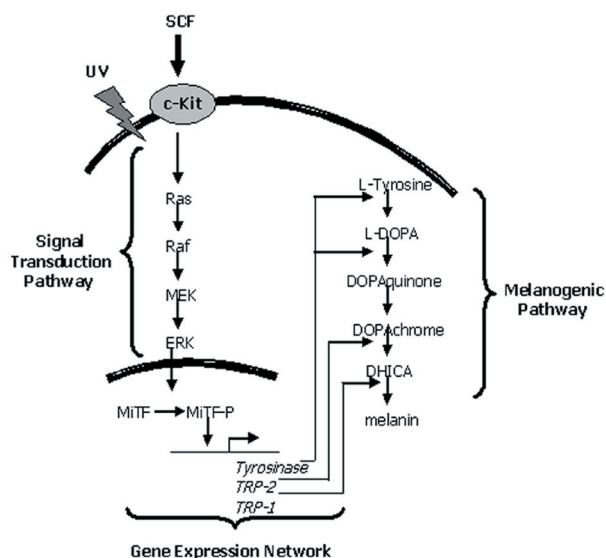


Figure 3. A cartoon diagram showing a typical interactions and reactions map that constitutes the basic framework for a computational model (courtesy of Emir and Kurnaz, 2003). This schematic shows a signal transduction pathway that feeds into a genetic circuit, whereby pigment-specific genes are transcribed. Enzyme products of these genes, in turn, regulate the pigment-synthesis pathway (melanogenesis). (For detailed explanation of the model, please refer to the relevant paper).

modeling is first to construct a map of all the reactions, metabolites and interactions on a “map” (Fig.3). Then, all the reaction steps are entered in the Reactions window of Gepasi one by one (Fig.4 A and B). Once the reactions steps are entered, the most crucial part is the definition of the kinetic type for each reaction, including all the necessary reaction parameters (Fig.4C). Once the initial concentrations of the metabolites are also defined (Fig.4D), the simulation program can be started, scanning a time period predefined by the user.

The scans are depicted as a Gnuplot graph (Fig.5), and all the values in the time course scan are saved as a simulation results file that can be imported into MS Excel or Sigma Plot application software for analysis. Typical SigmaPlot results for a chromatin modification model is shown in Fig.6.

The program is being constantly improved with new additions, such as MEG, the Model Extender for Gepasi, that allows several different Gepasi units to “talk” to each other and exchange metabolites, thereby simulating models of liquid cultures, 2-dimensional tissues and 2-dimensional tissues immersed in a liquid medium (18, Fig.7). Unfortunately, current MEG only allows for multiples of the same Gepasi unit to exchange information, in other words it is not yet possible to simulate cell-cell communication among different cell types. This disadvantage will hopefully be overcome in COPASI, an improved program based on Gepasi (http://www.vbi.vt.edu/research/projects/resproj_mendes_copasi.htm).

In addition, The ERATO Systems Biology Workbench Development Team has developed an XML (Extensible Markup Language) Schema for SBML (Systems Biology Markup Language), “a description language for simulations in systems biology (<http://www.sbw-sbml.org>) for the representation of biochemical networks via exchange of models between different simulation/analysis tools (22-24). SBML has merged features from simulation systems including, but not restricted to, E-Cell, Gepasi, Jarnac, and Virtual Cell, in collaboration with authors of these packages. This will enable researchers to integrate single models generated by a variety of modeling tools such as MAPK signaling, cell cycle and nucleotide metabolism, thereby enabling a fully-integrated model of complex cellular systems in the future.

CONCLUSION

Metabolic Engineering can be defined as the targeted and purposeful modification of metabolic pathways in an organism for the improved use of cellular pathways for chemical transformation, energy transduction, macromolecular synthesis or breakdown, potentially benefiting the society by producing biological substitutes

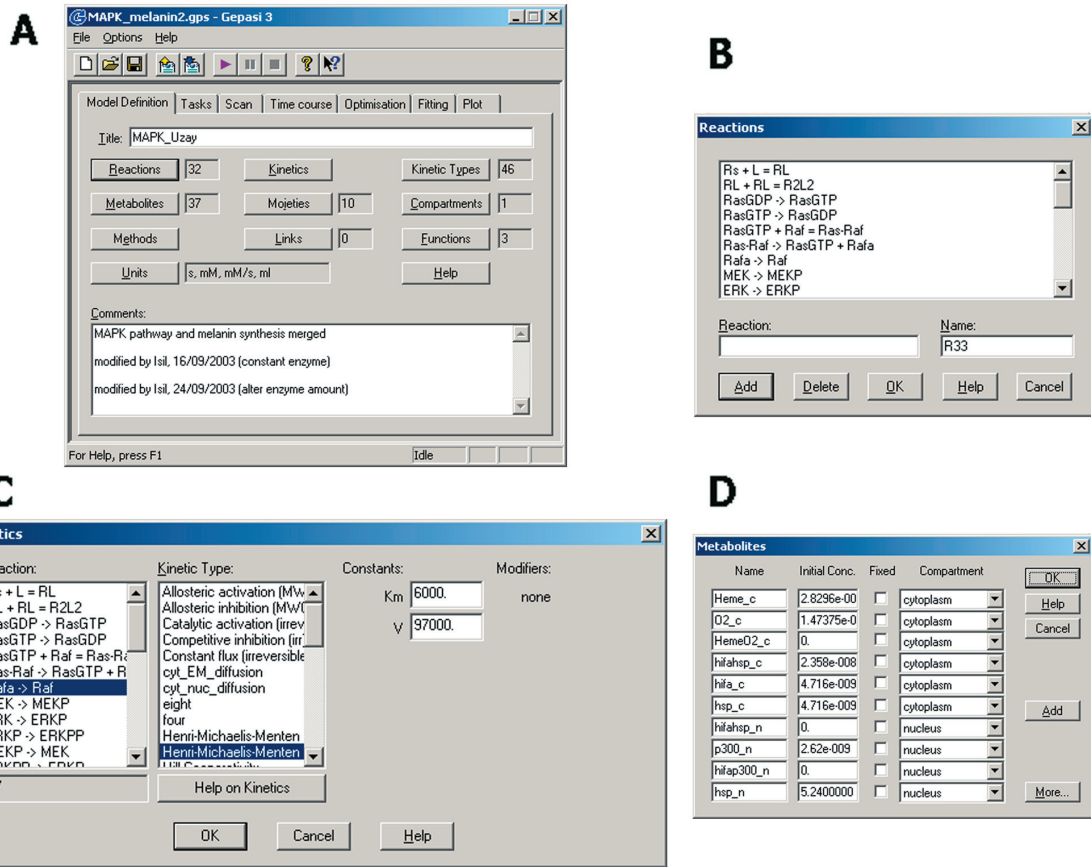


Figure 4. Gepasi biochemical simulation environment user interface (15-18). (A) A typical Gepasi model page where different components of the model can be entered (courtesy of Emir and Kurnaz, MAPK and melanogenesis model). (B) The reactions page is used to enter each biochemical reaction or interaction one by one (courtesy of Emir and Kurnaz, MAPK and melanogenesis model). (C) Each reaction is then defined in terms of its kinetic type, and kinetic parameters (courtesy of Yucel and Kurnaz, unpublished VEGF and angiogenesis model). (D) The next step is the definition of each metabolite concentration from the Metabolites page (courtesy of Yucel and Kurnaz, unpublished model).

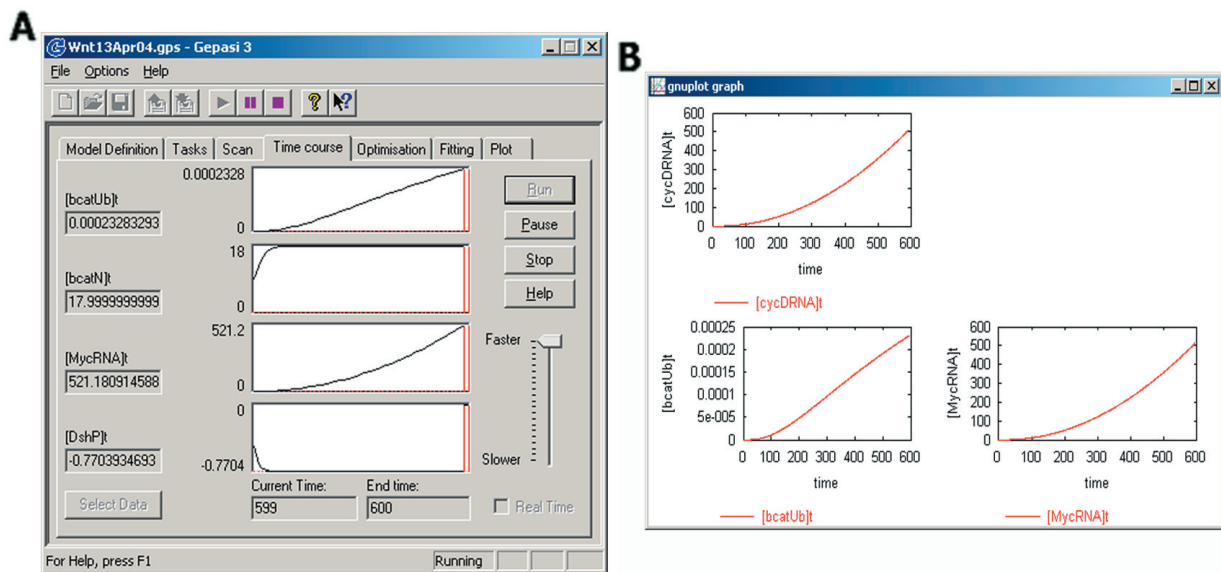


Figure 5. Typical plotting possibilities of Gepasi simulation software. (A) The time course plot generated by the Gepasi software itself (courtesy of Aksan Kurnaz, Wnt pathway time course scans, unpublished VEGF and angiogenesis model). (B) Gnuplot graphs of user-selected metabolites generated by Gepasi software add-in module (courtesy of Aksan Kurnaz, Wnt pathway figures, unpublished model).

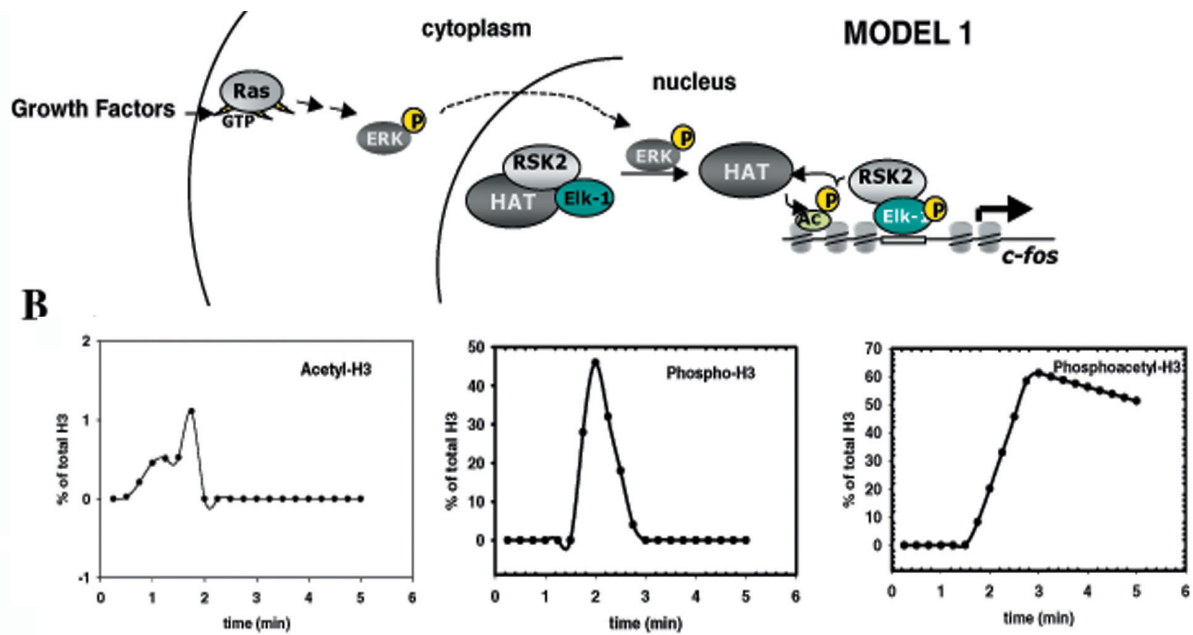


Figure 6. A typical figure generated for a Gepasi model (courtesy of Aksan Kurnaz 2004). (A) An interactions map for the model of chromatin modifications simulated in Gepasi environment. In this scheme, growth factor stimulation of a receptor tyrosine kinase leads to the activation of the Ras/MAPK pathway, thr downstream effectors of which are the RSK2 kinase and the Elk-1 transcription factor. In Model 1 shown here, RSK2, Elk-1 and histone acetyltransferase HAT are assumed to be ina trimeric complex. Upon phosphorylation and activation of RSK2 and Elk-1 by ERK MAPK, HAT dissociates from the complex and is activated, acetylating histone proteins on the local chromatin of the *c-fos* immediate-early gene. (7,20) (B) A typical SigmaPlot graph generated to demonstrate the changes of several metabolites in time, in this case changes in acetyl-histone H3, phospho-histone H3 and acetyl-phospho-H3. Acetyl-H3, phospho-H3 and acetyl-phospho H3 amounts are shown against time as percentage of total H3 protein in the chromatin region. t=0 shows the period of growth factor stimulation (7).

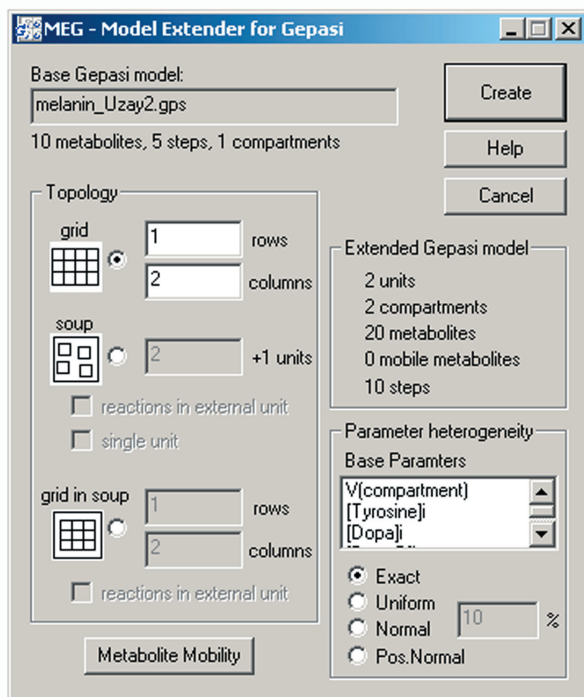


Figure 7. A screenshot of the MEG (Model Extender for Gepasi) module of Gepasi (courtesy of Yucel and Kurnaz, unpublished VEGF and angiogenesis model) (18). It is possible to allow for cross-talk between several repeating units of a Gepasi model, mimicking a population of cells, either in the grid (for example, epithelial tissue), or the soup (for example, cells in the bloodstream) model.

for toxic chemicals, by increasing agricultural production, improvement of industrial fermentation processes, the production of completely new compounds, or by an understanding of the molecular mechanism underlying medical conditions in order to develop new cures.

The discovery, characterization, and mapping of molecular interactions, regulatory patterns, cellular signaling pathways and genetic regulatory circuits offer a critical link between gene discovery, target validation, screening and optimization of novel drug targets. Computational biology appears to be a useful new tool for these purposes.

References

1. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, et al. (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science*, 270:397-403.
2. Jamshidi N, Edwards JS, Fahland T, Chrucho GM, Palsson BO (2001). Dynamic simulation of the human red blood cell metabolic network. *Bioinformatics*, 17(3): 286 – 287.
3. Mathews CK, van Holde KE and Ahern KG. (1999). *Biochemistry*. 3rd Ed. Benjamin Cummings Pub.
4. Voit EO (2000). *Computational Analysis of Biochemical Systems : A Practical Guide for Biochemists and Molecular Biologists*. Cambridge University Press
5. Herrgard MJ, Palsson BO. (2004). Flagellar biosynthesis in silico: building quantitative models of regulatory networks. *Cell*, 117(6):689-90.
6. Hasty J, Isaacs F, Dolnik M, McMillen D, Collins JJ. (2001). Designer gene networks: Towards fundamental cellular control. *Chaos*, 11(1):207-220.
7. Aksan I (2004) A kinetic analysis of RSK2 and Elk-1 interaction on the serum response element and implications for cellular engineering. *Biotech. Bioeng.* 88(7): 890 – 900.
8. Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, Miyoshi F, Saito K, Tanida S, Yugi K, Venter JC and Hutchison III CA (1999). E-CELL: software environment for whole-cell simulation. *Bioinformatics*, 15(1): 72-84.
9. Yang J, Kjellsson M, Rostami-Hodjegan A and Tucker GT. (2003). The effects of dose staggering on metabolic drug-drug interactions. *Eur J Pharm Sci.*, 20(2):223-32.
10. Emir UE and Aksan Kurnaz I (2003). An integrated model for melanocyte-specific gene expression and melanogenesis. *Signal Transduction: Receptors, mediators and genes*, 3(5-6): 209 – 217.
11. Gardner ST, Cantor CR and Collins JJ (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403: 339 – 342.
12. Yuh C-H, Bolouri H, Davidson EH (1998). Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* 279: 1896 – 1902.
13. Loew LM and Schaff JC (2001). The Virtual Cell: a software environment for computational cell biology. *Trends in Biotech.* 19(10): 401-406.
14. Xu C, Watras J, Loew LM. (2003). Kinetic analysis of receptor-activated phosphoinositide turnover. *J Cell Biol.* 161(4): 779 – 791.
15. Mendes P. (1993) GEPASI: A software package for modelling the dynamics, steady states and control of biochemical and other systems. *Comput. Applic. Biosci.* 9, 563-571.
16. Mendes P. (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends Biochem. Sci.* 22, 361-363.
17. Mendes P and Kell DB. (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics* 14, 869-883.
18. Mendes P and Kell DB (2001). MEG (Model Extender for Gepasi): a program for the modeling of complex, heterogenous, cellular systems. *Bioinformatics*, 17(3): 288-289.
19. Brightman FA and Fell DA. (2000). Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells. *FEBS Lett.* 482(3):169-74.
20. Aksan I and Kurnaz ML. (2003) A computer-based model for the regulation of MAPK activation. *J. Rec. Signal Transduc.* 23(2-3): 197 – 209.
21. Martins AM, Mendes P, Cordeiro C and Freire AP. (2001). In situ kinetic analysis of glyoxalase I and glyoxalase II in *Saccharomyces cerevisiae*. *Eur. J. Biochem.* 268: 3930 – 3936.
22. Shapiro BE, Hucka M, Finney A, Doyle J. (2004) MathSBML: a package for manipulating SBML-based biological models. *Bioinformatics*, e-pub ahead of print
23. Finney A, Hucka M. (2003) Systems biology markup language: Level 2 and beyond. *Biochem Soc Trans.* 31(Pt 6):1472-3.
24. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novere N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J; SBML Forum. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524-31.