WinBEST-KIT for Analyzing Multilayer and Multicellular Systems

Tatsuya Sekiguchi
Department of Life Sciences and Informatics,
Faculty of Engineering,
Maebashi Institute of Technology,
Maebashi, Gunma 371-0816, Japan
Email: sekiguchi@maebashi-it.ac.jp

Masahiro Okamoto*
Department of Bioinformatics,
Graduate School of Systems Life Sciences,
Kyushu University,
Maidashi, Fukuoka 812-8582, Japan
Email: okahon@brs.kyushu-u.ac.jp

***Abstract*—Previously, we developed a biochemical reaction simulator called WinBEST-KIT (Biochemical Engineering System analyzing Tool-KIT, which runs under Microsoft Windows) for analyzing complicated metabolic pathways. WinBEST-KIT provides an integrated simulation environment for experimental researchers in metabolic engineering. A particularly notable feature of WinBEST-KIT is that users can easily define and customize reaction symbols in the graphical user interface. Users can use their original kinetic equations, in addition to the pre-installed standard kinetic equations, to represent unknown kinetic mechanisms as reaction steps. However, owing to the increasing size of reaction systems to be analyzed in metabolic pathways, large-scale reaction systems must be divided into several arbitrary compartmental reaction systems and procedures are needed, such as multilayered hierarchical representation, to describe the interactions between the compartmental reaction systems. Accordingly, in this study, we developed a new version of WinBEST-KIT that enables users to construct several arbitrary reaction schemes as layers, to connect the layers, and to analyze the interactions between them. This hierarchical representation is effective for constructing multilayered mathematical models of biochemical systems, such as genome–enzyme–metabolite systems, reaction cascade systems, and multicellular systems.**

I. Introduction

With recent advances in molecular biology, understanding of molecular-level mechanisms in biochemical systems has rapidly improved. A simple collection of static databases on molecular-level mechanisms does not provide insight into the functional properties of a biochemical system, and thus the use of genomic, proteomic, and metabolomic analyses has been growing over the last several years. For example, systems biology is a top-down approach where a system's functional properties are presumed to result from the interactions between system components including genes, proteins, and metabolites. Research strategies in systems biology are 1) system identification, 2) system analysis, 3) system control, and 4) system design. Owing to improvements in computing power, the dynamic (time variant) behavior of large-scale nonlinear networks such as metabolic pathways can now be studied through computer simulations. To understand the dynamic behavior of metabolites and their interactions in metabolic pathways, biochemical reaction simulators are effective tools, which can be used in tasks such as data collection, kinetics modeling, numerical simulation, parameter estimation, and simulation visualization. Accordingly, the development of a biochemical reaction simulator has become an important research topic in systems biology. In recent years, a number of biochemical reaction simulators have been reported, including Gepasi [1], DBSolve [2], E-Cell [3], CellDesigner [4], COPASI [5], and Cell Illustrator [6]. In addition, we developed a biochemical reaction simulator called WinBEST-KIT (Biochemical Engineering System analyzing Tool-KIT, which runs under Microsoft Windows) [7] for analyzing complicated metabolic pathways. Since these simulators can construct and analyze arbitrary reaction schemes without requiring knowledge of information sciences, experimental researchers can effectively understand the target reaction systems.

However, the problem is that the activity of biochemical systems is governed by the interactions of several reaction systems in metabolic pathways. Each reaction system has a specific role and thus can be treated as a "compartment" or a "layer" of a metabolic-level mechanism. (In this study, we use the term "layer.") When viewed in this way, the activity of biochemical systems is governed by the interactions of several layers. For example, in aerobic respiration, ATP is produced by the interactions of the following reaction systems, which can each be treated as a layer of the metabolic mechanism: the glycolytic pathway, the citric acid cycle, and the electron-transport chain. When constructing and analyzing the interactions of several layers, however, the scale of the reaction system to be analyzed in metabolic pathways becomes massive. In most distributed biochemical reaction simulators, including our original version of WinBEST-KIT, users can simultaneously construct and analyze only a single-reaction scheme. Consequently, users must manage a large reaction scheme involving several reaction systems in a single layer. Without the use of multiple layers, users are likely to encounter difficulty in understanding the large-scale reaction systems that constitute metabolic pathways.

To overcome this problem, we drew inspiration from a conventional technology in electrical engineering. In electronic circuit design, the well-known simulator SPICE (Simulation Program with Integrated Circuit Emphasis) can be imple-

* To whom correspondence should be addressed.

mented with a schematic capture front-end that provides a graphical user interface for users to select and connect symbols representing circuit components in order to design and simulate electronic circuits. A practical implementation of an integrated circuit contains numerous circuit components packed onto a single microchip; the integrated circuits function independently and respond specifically to input signals from the electronic circuits to be designed. Thus, we can consider electronic circuits to have a hierarchical tree structure (i.e., a layer structure). The most important feature of SPICE is that symbols representing circuit components of integrated circuits can be treated as connectors to other layers because the internal components of the integrated circuit also consist of electronic circuits. Thus, users can design electronic circuits without considering the internal circuit components of integrated circuits.

In this study, the concept of integrated circuit components in SPICE is applied to our biochemical reaction simulator and a new version of WinBEST-KIT is developed that enables users to construct several arbitrary reaction schemes as layers, to connect the layers, and to analyze the interactions between them.

II. Overview of WinBEST-KIT

Figure 1 shows a screenshot of WinBEST-KIT, which provides an integrated simulation environment for analyzing complicated metabolic pathways. We implemented several of the following features in the original version of WinBEST-KIT:

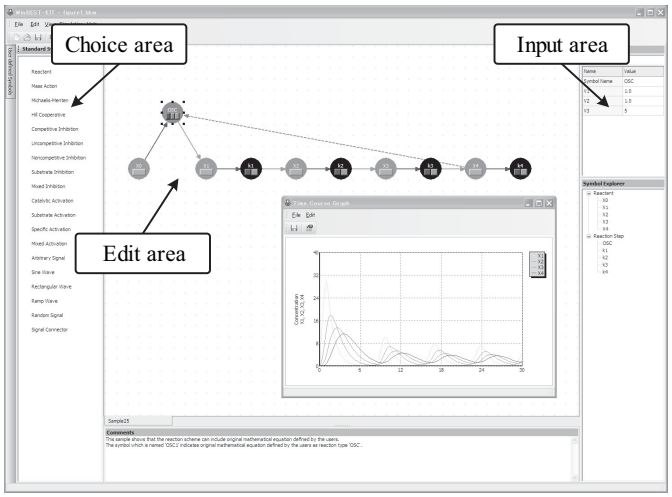


Fig. 1. Screenshot of WinBEST-KIT.

- Users can easily construct and analyze an arbitrary reaction scheme. The process is similar to drawing a picture through the graphical user interface and the kinetic model does not need to be considered. The reaction scheme can be constructed in the Edit area by selecting and connecting suitable "reactant symbols" and "reaction symbols" to represent the reactants and reaction steps, respectively, which are listed in the Choice area. The properties of these symbols can be edited in the Input area.
- The kinetic model consists of simultaneous nonlinear differential equations for the velocity of the reactants (dx/dt, where x and t represent the concentration of the reactant and the time, respectively). The mathematical formalisms in these differential equations are represented by both the mass action law and 11 types of well-known approximated velocity functions of steady-state enzyme kinetics (Michaelis–Menten, Hill cooperative, competitive inhibition, etc.); these kinetic equations are pre-installed and registered as standard reaction symbols in the Choice area.
- Users can define original kinetic equations to represent reaction steps and can customize these equations easily with user-defined reaction symbols in the graphical user interface. The mathematical formalisms in the simultaneous nonlinear differential equations of the kinetic model can also include user-defined equations.
- Complicated simultaneous nonlinear differential equations representing the kinetic model can be automatically derived without the user having to write out and manipulate the equations.
- Numerical calculations can be automatically performed by using an internal interpreter or an external complier.
- Users can use the following powerful analytical methods for system analysis: time-course calculations, parameter scanning, estimation of unknown kinetic parameter values based upon empirical time-course data on the reactants, analysis of reactants' dynamic responses to virtual external perturbations, and real-time simulations.

A particularly notable feature of WinBEST-KIT is that users can easily define and customize reaction symbols in the graphical user interface. Users can supply original kinetic equations, in addition to the pre-installed standard kinetic equations, for representing unknown kinetic mechanisms of reaction steps. For example, Shinto et al. [8], [9] applied WinBEST-KIT to the kinetic modeling and system analysis of acetone-butanol-ethanol (ABE) fermentation in *Clostridium saccharoperbutylacetonicum* N1-4 (ATCC13564). They used their original kinetic equations to represent reaction steps in constructing the target reaction scheme. Their study demonstrates that WinBEST-KIT is practical for experimental researchers because it can construct and analyze complicated reaction schemes involving unknown kinetic mechanisms. The details of customizing user-defined reaction symbols are described elsewhere [7].

III. Layer Structure of Metabolic Pathways

In Section I, we discussed the concept of integrated circuit components in SPICE and the application of this concept to our biochemical reaction simulator. The problem that must be considered here, however, is that the layer structure of metabolic pathways is subtly different from the layer structure of electronic circuits. We discuss the layer structure of metabolic pathways in order to explain the new features of WinBEST-KIT that are notably different from the features of other biochemical reaction simulators.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

Metabolic pathways can be divided into a layer structure for convenience and according to the role of the each pathways; in other words, the division of metabolic pathways is arbitrary and the position of the input and output terminals in each layer can be freely designed. Figure 2 shows a conceptual illustration of the layer structure of metabolic pathways.

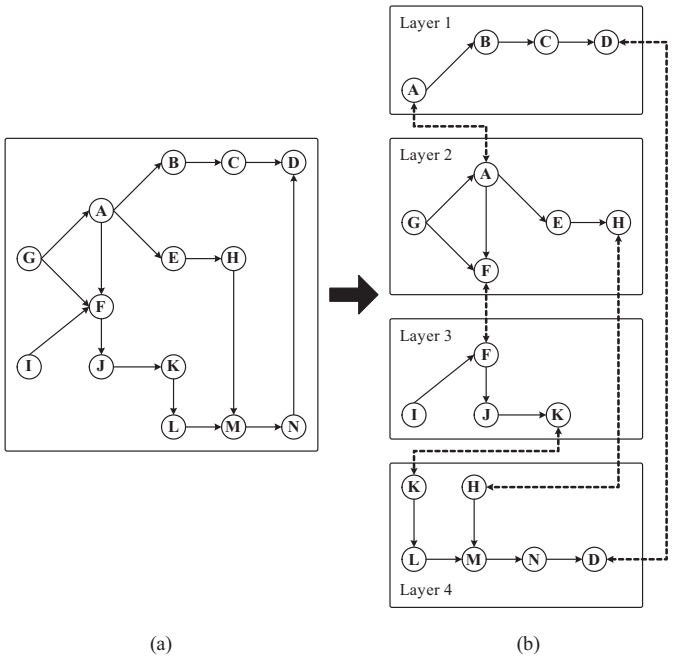


Fig. 2. Conceptual illustration of layer structure of metabolic pathways. Circled letters represent reactants.

As shown in Figure 2(b), the reaction scheme in Figure 2(a) is divided arbitrarily into four layers. The reactants in these layers are represented by circled letters, some of which are connected to other layers. In this case, layer 1 is connected to layers 2 and 4 by reactants A and D, respectively. Similarly, layer 2 is connected to layers 1, 3, and 4; layer 3 is connected to layers 2 and 4; layer 4 is connected to layers 1, 2, and 3. As shown in Figure 2(b), since the layer structure of the metabolic pathways is quite complicated and the related reaction schemes are intricately intertwined, the layers cannot be represented as a hierarchical tree structure. Thus, we can recognize from this figure that the connections between layers in the metabolic pathways do not follow any standard rule. The most important point is that the same reactant is placed in multiple layers to connect them; for example, reactant A in both layers 1 and 2 connects these two layers. As a more practical example, the glycolytic pathway, the citric acid cycle, and the electron-transport chain are typical metabolic pathway layers. Here, pyruvic acid, ATP, ADP, NAD^+, NADH, and so forth, are the reactants connecting the layers. Pyruvic acid connects the glycolytic pathway and the citric acid cycle; ATP, NAD^+, NADP, and NADH connect all three layers.

Next, we discuss a different problem in modeling metabolic pathways. A given reactant, having a certain name and concentration, is shown in multiple layers for users to easily represent and understand the metabolic pathways; for example, ATP, ADP, NAD^+, NADH, and so forth are located in many reaction steps in the metabolic pathways. Even though the reactant has a particular name and concentration, the stoichiometric coefficients of the reactant differ between reactions. For example, both 2ATP and ATP are present in the glycolytic pathway; and $2NAD^+$ and NAD^+ are present in the glycolytic pathway and the citric acid cycle, respectively.

IV. Implementation of Multilayer Structure in WinBEST-KIT

To construct several arbitrary reaction schemes as layers, to connect the layers, and to analyze the interactions between them, we implemented the following four new features in WinBEST-KIT:

i) Users can simultaneously construct several arbitrary reaction schemes as layers; specifically, we implemented multiple Edit areas in WinBEST-KIT. In each Edit area, a reaction scheme can be constructed in a corresponding layer. At the beginning of the simulation, the pointers of all symbols in all layers are assembled. By doing this, all the layers can be analyzed as a single large reaction scheme.

ii) We implemented "shortcut symbols" for the reactant symbols. The shortcut symbols are analogous to the shortcut icons in Microsoft Windows. Users can create any number of shortcut symbols in the Edit area, even if the corresponding target reactant symbol appears in a different layer. Thus, the shortcut symbols in the different layers play the role of connectors between the layers. The reactant symbols and the shortcut symbols have three properties: 1) symbol name, 2) concentration, and 3) stoichiometric coefficient. The symbol name and the concentration can be edited only in the Input area of the reactant symbols. These two properties of the shortcut symbols are synchronized when the two properties of their corresponding target reactant symbol are updated. However, all reactant symbols and all shortcut symbols can be assigned a unique stoichiometric coefficient. The stoichiometric coefficient can be edited in the Input areas of both the reactant symbols and the shortcut symbols.

iii) We implemented a merge function to facilitate efficient construction of reaction schemes. Users can add a saved reaction scheme to the reaction scheme being editing by copying the constructed scheme into the Edit area. In this operation, if the reactant names overlap, the reactant symbols to be merged are automatically converted into shortcut symbols.

iv) We implemented a new, original computer-readable format based on XML called BKML (BEST-KIT Markup Language) for documents of constructed reaction schemes. BKML stores the notations needed for the features of WinBEST-KIT, for example: the original kinetic equations for describing reaction steps, the layers, and the relationships between reactant symbols and their shortcut symbols.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

In the next section, we show typical examples of layer structures and of reaction scheme construction using the new features of WinBEST-KIT.

V. Examples of the New Features of WinBEST-KIT

A. Multilayer Systems

Figure 3 shows a metabolic pathway map of butanol production through synthetic biology techniques [10]. Synthetic butanol production involves two different metabolisms; the left side is *E. coli* metabolism and the right side is *C. acetobutylicum* metabolism. Acetoacetyl-CoA is a metabolite in both *E. coli* and *C. acetobutylicum* metabolisms. Thus, these two metabolisms are connected through acetoacetyl-CoA and butanol is produced via their interactions. Moreover, the same reactants (NAD^+ and NADH) are present in the metabolic pathways of these two different metabolisms. The stoichiometric coefficients of NAD^+ and NADH have a different value in each reaction step. The stoichiometric coefficients of NAD^+ and NADH are balanced in the two metabolisms. Four molecules of NADH are produced in *E. coli* metabolism, and four molecules of NADH are utilized in *C. acetobutylicum* metabolism. Conversely, four molecules of NAD^+ are utilized in *E. coli* metabolism, and four molecules of NAD^+ are produced in *C. acetobutylicum* metabolism.

This complicated metabolism for producing synthetic butanol can be constructed by using the layer structure and shortcut symbols in the new version of WinBEST-KIT. Figure 4 shows a screenshot of WinBEST-KIT being used to construct the reaction scheme presented in Figure 3. The two different metabolisms can be considered as layers connected through the reactant symbol of acetoacetyl-CoA in the *E. coli* metabolism layer and its shortcut symbol in the *C. acetobutylicum* metabolism layer. Since the reactant symbols and the shortcut symbols have independent stoichiometric coefficients, all the reactant symbols and their shortcut symbols have the same name and the same concentration but different stoichiometric coefficients. Thus, all the NAD^+ and NADH molecules can be arranged according to the original metabolic pathway map in Figure 3 by using the shortcut symbols. In Figure 3, at the reaction step of 2glyceraldehyde-3-phosphate to 2-1,3-diphosphoglycerate in the *E. coli* metabolism, two molecules of NADH are produced, and at the reaction step of acetoacetyl-CoA to 3-hydroxybutyryl-CoA in the *C. acetobutylicum* metabolism, one molecule of NADH is utilized. When we construct these two reaction steps, the reactant symbol of NADH and its shortcut symbol are defined, and the stoichiometric coefficients can be set independently as 2 and 1, respectively.

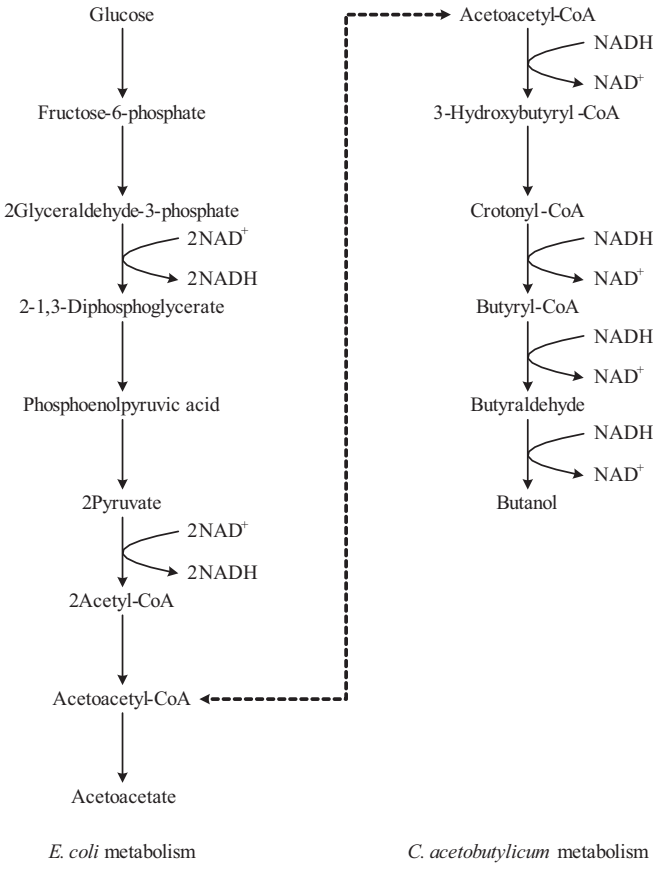
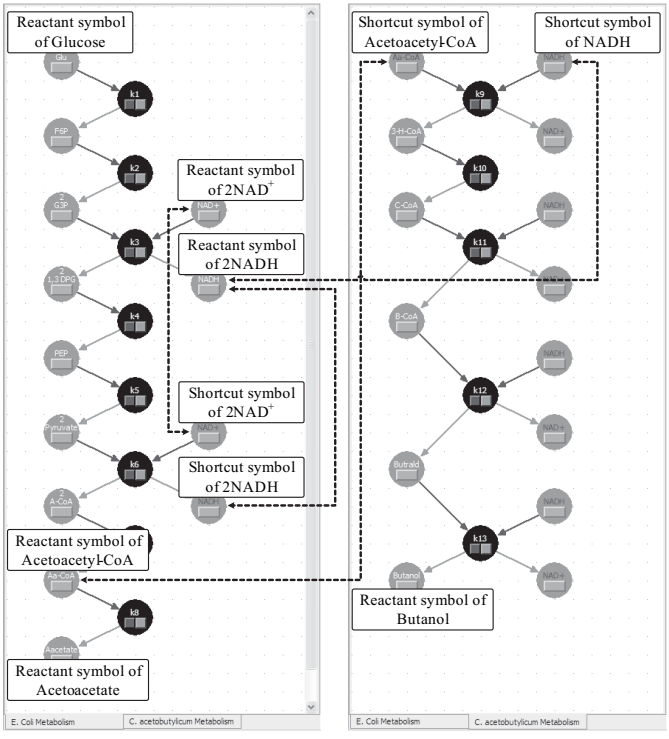


Fig. 3. Metabolic pathway map of butanol production through synthetic biology techniques. In both *E. coli* and *C. acetobutylicum* metabolisms, Acetoacetyl-CoA symbols connected by a bidirectional broken arrow represent the same metabolite.



E. coli metabolism layer *C. acetobutylicum* metabolism layer

Fig. 4. Construction of reaction scheme in WinBEST-KIT for butanol production through synthetic biology techniques. A bidirectional broken arrow indicates a connection between a reactant symbol and its shortcut symbol.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

Figure 5 shows the reactant symbol of NADH and its shortcut symbol in the two reaction steps in WinBEST-KIT. The number above the symbol name is the stoichiometric coefficient of the reactant. The number is not displayed when the value of the stoichiometric coefficient is 1. In Figure 5, the stoichiometric coefficient of the left symbol is 2, and the stoichiometric coefficient of the right symbol is 1.

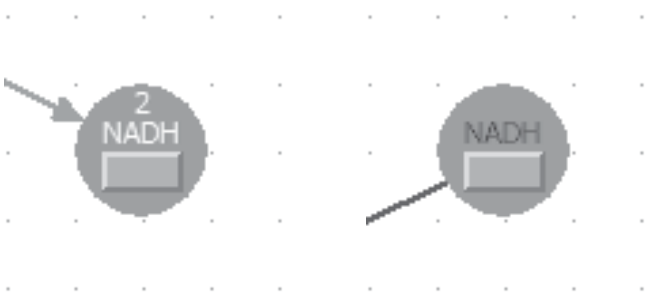


Fig. 5. Reactant symbol of NADH and its shortcut symbol from Figure 4.

B. Multicellular Systems

As previously stated, the layer structure in WinBEST-KIT is implemented for efficiently constructing and analyzing complicated metabolic pathways. In the following, we describe another possible use of the layer structure in WinBEST-KIT.

To obtain sufficient knowledge of biochemical systems with a holistic view of living cells, the dynamics must be understood at both the intracellular and intercellular levels. A system analysis of multicellular systems can be performed by using the layer structure and the shortcut symbols in the new version of WinBEST-KIT. Kaneko and Yomo [11] proposed a mechanism of cell differentiation; this mechanism represents the cell–cell communication involving the metabolic reactions within each cell (intracellular dynamics) and the interactions with other cells through a medium (intercellular dynamics). Their model is quite complicated. In a similar but simpler example of a multicellular system that captures the essence of Kaneko and Yomo's model of cell–cell communication, we use WinBEST-KIT to examine a nutrient consumed competitively according to the growth activity of the cells in the medium.

Figure 6 shows our model of cell–cell communication. Let us assume that the metabolic reactions within each cell are quite simple, as shown. Here, X_{ij} represents the j-th metabolite in the i-th cell, where $0 \leq j \leq 4$ and $0 \leq i \leq 3$. The reaction step of $X_{i0} \rightarrow X_{i1}$ within each cell is regulated by the negative feedback from X_{i4}. The other reaction steps within each cell follow Michaelis–Menten kinetics. Moreover, the nutrient is supplied to X_{i0} only from the medium, and the cell having high growth activity assimilates more of the nutrient. Given that S represents the total concentration of the nutrient in the medium, the velocity of the inflow of the nutrient to the i-th cell can be written as

$$v_i = \frac{X_{i0}}{X_{10} + X_{20} + X_{30}} S \quad (i = 1, 2, 3). \tag{1}$$

However, the reaction step described by Eq. (1) is not one of the pre-installed standard kinetic equations in WinBEST-KIT. In this case, we can use an original kinetic equation to describe the reaction step by defining a reaction symbol in WinBEST-KIT.

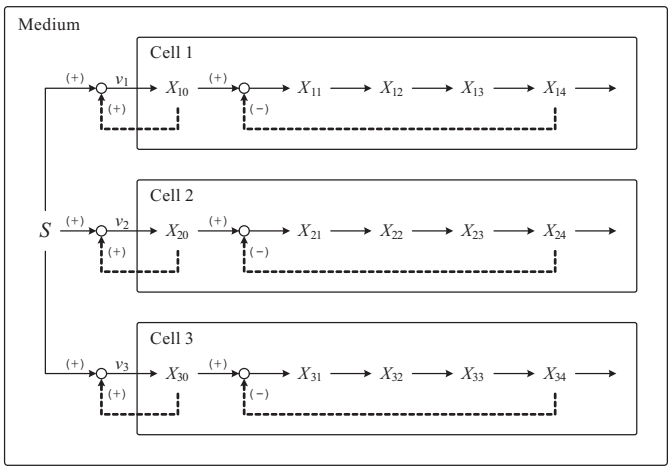


Fig. 6. Simple model of cell–cell communication.

Figure 7 shows a screenshot of WinBEST-KIT being used to construct the reaction scheme presented in Figure 6. One layer can be considered a "medium," and the others, "cell." The layer of cell i $(i = 1, 2, 3)$ is connected to the medium

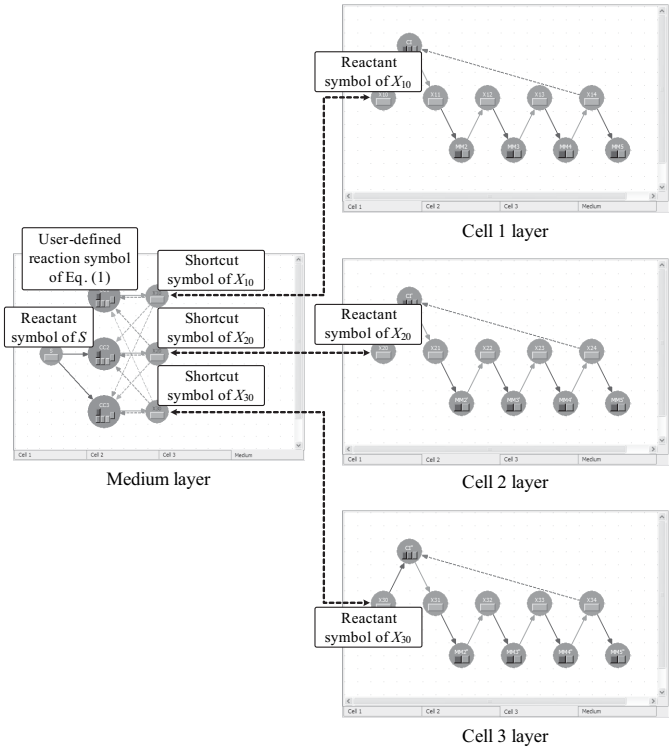


Fig. 7. Construction of reaction scheme for a simple model of cell–cell communication in WinBEST-KIT. A bidirectional broken arrow indicates the connection between a reactant symbol and its shortcut symbol.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

layer through the reactant symbol X_{i0} and its shortcut symbol. Moreover, the shortcut symbol for X_{i0} in the medium layer is connected to the reactant symbol for S by the user-defined reaction symbol for Eq. (1) representing a reaction step. These reaction steps show that the cell having high growth activity assimilates more of the nutrient.

VI. Conclusion

In this study, we developed a new version of WinBEST-KIT that enables users to construct several arbitrary reaction schemes as layers, to connect the layers, and to analyze the interactions between them. To realize the layer structure of metabolic pathways in this new version, we implemented shortcut symbols that have the same name and the same concentration as the corresponding target reactant symbol but that can be assigned a unique stoichiometric coefficient. In our implementation of this concept, users can arrange the reactant symbols at arbitrary positions in the Edit area in reference to the original metabolic pathway map. Moreover, users can create any number of shortcut symbols in the Edit area, even if the corresponding target reactant symbol appears in a different layer. As a result, users can freely divide the target reaction scheme into several arbitrary layers, and can connect the layers by using shortcut symbols. Furthermore, the existing features of WinBEST-KIT can be used to construct and analyze a large-scale reaction scheme that involves several layers. In particular, a notable feature is that users can define original kinetic equations to represent reaction steps with unknown kinetic mechanisms. This feature is useful because reaction systems are likely to include an increasing number of unknown kinetic mechanisms as the size of the reaction systems to be analyzed in metabolic pathways continues to grow.

The new version of WinBEST-KIT runs under Microsoft Windows XP/Vista/7. We are now preparing a website in order to distribute the new version of WinBEST-KIT and its user manual.

A problem to be overcome in WinBEST-KIT is its compatibility with SBML (Systems Biology Markup Language). SBML is one of the best and most widely used computer-readable formats in systems biology application software. Since WinBEST-KIT should work well with other biochemical reaction simulators to aid in the further development of systems biology, we intend in the near future to implement an import/export function for SBML in the next version of WinBEST-KIT.

References

[1] P. Mendes and D. Kell, "Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation," *Bioinformatics*, vol. 14, pp. 869–883, 1998.

[2] I. Goryanin, T. C. Hodgman, and E. Selkov, "Mathematical simulation and analysis of cellular metabolism and regulation," *Bioinformatics*, vol. 15, pp. 749–758, 1999.

[3] M. Tomita, K. Hashimoto, K. Takahashi, T. S. Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. C. Venter, and C. A. Hutchison III, "E-CELL: software environment for whole-cell simulation," *Bioinformatics*, vol. 15, pp. 72–84, 1999.

[4] A. Funahashi, M. Morohashi, H. Kitano, and N. Tanimura, "CellDesigner: a process diagram editor for gene-regulatory and biochemical networks," *BIOSILICO*, vol. 1, pp. 159–162, 2003.

[5] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer, "COPASI – a COmplex PAthway SImulator," *Bioinformatics*, vol. 22, pp. 3067–74, 2006.

[6] M. Nagasaki, A. Saito, E. Jeong, C. Li, K. Kojima, E. Ikeda, and S. Miyano, "Cell Illustrator 4.0: A computational platform for systems biology," *In Silico Biology*, vol. 10, pp. 5–26, 2010.

[7] T. Sekiguchi and M. Okamoto, "WinBEST-KIT: Windows-based Biochemical Reaction Simulator for Metabolic Pathways," *Journal of Bioinformatics and Computational Biology*, vol. 4, pp. 621–638, 2006.

[8] H. Shinto, Y. Tashiro, M. Yamashita, G. Kobayashi, T. Sekiguchi, T. Hanai, Y. Kuriya, M. Okamoto, and K. Sonomoto, "Kinetic Modeling and Sensitivity Analysis of Acetone-Butanol-Ethanol Production," *Journal of Biotechnology*, vol. 131, pp. 45–56, 2007.

[9] H. Shinto, Y. Tashiro, G. Kobayashi, T. Sekiguchi, T. Hanai, Y. Kuriya, M. Okamoto, and K. Sonomoto, "Kinetic study of substrate dependency for higher butanol production in acetone-butanol-ethanol fermentation," *Process Biochemistry*, vol. 43, pp. 1452–1461, 2008.

[10] S. Atsumi, A. F. Cann, M. R. Connor, C. R. Shen, K. M. Smith, M. P. Brynildsen, K. J. Y. Chou, T. Hanai, and J. C. Liao, "Metabolic engineering of Escherichia coli for 1-butanol production," *Metabolic Engineering*, vol. 10, pp. 305–311, 2008.

[11] K. Kaneko and T. Yomo, "Cell division, differentiation, and dynamic clustering," *Physica D*, vol. 75, pp. 89–102, 1994.

[12] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, Goryanin, II, W. J. Hedley, T. C. Hodgman, J. H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novere, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang, "The Systems Biology Markup Language (SBML): A Medium for Representation and Exchange of Biochemical Network Models," *Bioinformatics*, vol. 19, pp. 524–531, 2003.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE