INVITED PAPER

CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks

This tool uses developing standards for graphical representation of biological systems, and for intercommunications between biological objects and interactions, to allow researchers to easily create network models.

By Akira Funahashi, Yukiko Matsuoka, Akiya Jouraku, Mineo Morohashi, Norihiro Kikuchi, and Hiroaki Kitano

ABSTRACT | Understanding of the logic and dynamics of gene-regulatory and biochemical networks is a major challenge of systems biology. To facilitate this research topic, we have developed a modeling/simulating tool called CellDesigner. CellDesigner primarily has capabilities to visualize, model, and simulate gene-regulatory and biochemical networks. Two major characteristics embedded in CellDesigner boost its usability to create/import/export models: 1) solidly defined and comprehensive graphical representation (systems biology graphical notation) of network models and 2) systems biology markup language (SBML) as a model-describing basis, which function as intertool media to import/export SBML-based models. In addition, since its initial release in 2004, we have extended various capabilities of CellDesigner. For example, we integrated other Systems Biology Workbench enabled simulation/analysis software packages. CellDesigner also supports simulation and parameter search, supported by integration with SBML ODE Solver, enabling users to simulate through our sophisticated graphical user interface. Users can also browse and modify existing models by referring to existing databases directly through CellDesigner. Those extended functions empower CellDesigner as not only a modeling/simulating tool but also an integrated analysis suite. CellDesigner is implemented in Java and thus supports various platforms (i.e., Windows, Linux, and MacOS X). CellDesigner is freely available via our Web site.

KEYWORDS | Biochemical simulation; kinetic modeling; SBGN; SBML; systems biology

I. INTRODUCTION

Systems biology is characterized by synergistic integration of theory, computational modeling, and experiments [1]. Identification of the logic and dynamics of gene-regulatory and biochemical networks is a major challenge of systems biology. From the view of computational modeling, a model is used to understand the dynamics of biological phenomena. The model consists of molecules and reactions that represents gene regulatory and biochemical network (such as transcription, translation, protein–protein interaction, enzymatic reaction, etc.), and contains a mathematical equation for each reaction. So that the model contains mathematical equations inside, it would be possible to simulate the dynamics of the model and compare the simulation results with their experiments; even more, it would be possible to tune the parameters in the model to fit with the experimental results. This workflow is important to understand unknown function or structure of biological phenomena, so development of software infrastructure to support this

Manuscript received November 29, 2007; revised January 27, 2008. This work was supported by the ERATO-SORST program (Japan Science and Technology Agency), the International Standard Development area of the International Joint Research Grant (NEDO, Japanese Ministry of Economy, Trade, and Industry), the Strategic Japanese-Swedish Cooperative Program on "Multidisciplinary BIO" (JST-VINNOVA/SSF), the Establishment of a Human Genome Network Platform (MEXT) and through the Japanese Ministry of Education, Culture, Sports, Science, and Technology.
A. Funahashi and **A. Jouraku** are with the Department of Biosciences and Informatics, Keio University, Yokohama 223-8522, Japan (e-mail: funa@bio.keio.ac.jp; jouraku@bio.keio.ac.jp).
Y. Matsuoka is with Kitano Symbiotic Systems Project, ERATO-SORST, Tokyo 150-0001, Japan (e-mail: myukiko@symbio.jst.go.jp).
M. Morohashi is with PRTM, Tokyo 163-0430, Japan (e-mail: moro@mineo.com).
N. Kikuchi is with Mitsui Knowledge Industry Co., Ltd., Tokyo 164-8555, Japan (e-mail: kikuchi-norihiro@mki.co.jp).
H. Kitano is with The Systems Biology Institute, Shibuya, Tokyo 150-0001, Japan (e-mail: kitano@sbi.jp).
Digital Object Identifier: 10.1109/JPROC.2008.925458

workflow is essential for systems biology research. While the software infrastructure is one of the most crucial components in systems biology research, there has been almost no common infrastructure or standard to enable integration of computational resources. For example, researchers built their model with their specific application or inside their simulator as a source code so that it was difficult to port their model to be used on other applications. Since there was no gold-standard software for systems biology research, at that time, researchers had to use multiple applications to solve their problem. They had to switch their software to run simulations, analyze the model, and fit parameters with their experimental results. To solve this problem, the Systems Biology Markup Language[1] (SBML) [2], [3] and the Systems Biology Workbench[2] (SBW) have been developed [4]. SBML is an open, Extensible Markup Language (XML)-based format for representing biochemical reaction networks, which enables researchers to share their model between different software applications, while SBW is a modular, broker-based message-passing framework for simplified intercommunication between applications. Rapid acceptance of this standard is proved by the fact that more than 110 simulation and analysis software packages already support SBML or are in the process of supporting the standard.

We believe that the standardized technologies, such as SBML, SBW, and Systems Biology Graphical Notation (SBGN—a graphical notation for network diagrams of biological models), play a critical role as the software platform to tackle this challenge. As an approach, we have developed CellDesigner [5], a process diagram editor for gene-regulatory and biochemical networks. CellDesigner currently supports model creation, simulation, and database integration—those features are significant for users willing to create their model from scratch.

II. FEATURES OF CELLDESIGNER

The current version (3.5.2, as of June 2008) of CellDesigner has the following features:

- representation of biochemical semantics;
- detailed description of state transition of proteins;
- SBML compliant (SBML Level-1 and Level-2 Version-1);
- integration with SBW-enabled simulation/analysis modules;
- integration with native simulation library (SBML ODE Solver [6]);
- database connectivity;
- platform independent.

The aim of developing CellDesigner is to supply a process diagram editor utilizing standardized technology (SBML and SBGN in this case) for every computing platform, so that it could confer benefits to as many users as possible. By using the standardized technology, any model can be easily ported to other applications, thereby reducing the cost to create a specific model from scratch. The main standardized features that CellDesigner supports are summarized as "graphical notation," "model description," and "application integration environment." The standard for graphical notation plays an important role for efficient and accurate dissemination of knowledge [7], and these standards for model description enhance the portability of models among various software tools and aid human readability. Similarly, the standard for application integration environment will help software developers to provide the ability for their applications to communicate with other tools.

A. Symbols/Expressions and SBGN

CellDesigner supports graphical notation and listing of symbols based on a proposal by our group [7]. While we have proposed our original notation system, graphical notation has now been developed as an international community based activities called SBGN.[3] So far, several graphical notation systems already have been proposed [8]–[12]. The goal of SBGN is to design a graphical notation system expressing sufficient information in more visible and more unambiguous way, as we proposed [7]. We expect that these features will become part of the standardized technology in systems biology field. The key components of SBGN are:

- to allow representation of diverse biological objects and interactions;
- to be semantically and visually unambiguous;
- to be able to incorporate other notations;
- to allow software tools to convert a graphically represented model into mathematical formulas for analysis and simulation;
- to have software support to draw diagrams;
- to make the notation scheme of SBGN freely available.

To accomplish the above requirements for the notation, we first decided to define a notation by using a process diagram [7]. The notation graphically represents state transitions of the molecules involved. In the process diagram representation, each node represents the state of the molecules and complex, and each arrow represents state transitions among the states of a molecule. In the conventional entity-relationship diagrams, an arrow generally represents activation of the molecule. However, this confuses the semantics of the diagram, as well as limits possible molecular processes that can be represented [7]. A process diagram represents a more intuitive way for model definition than entity-relationship diagrams. One of the reasons is that the

[1]http://www.sbml.org.
[2]http://sys-bio.org.
[3]http://www.sbgn.org.

process diagram could be explicitly represented as a temporal sequence of events whereas an entity-relationship cannot. For example, in a process of mitosis-promoting factor (MPF) activation in cell cycle, Wee1 phosphorylates residues of Cdc2 (Cell Division Cycle 2), is one of the components of MPF (Fig. 1). However, MPF is not yet activated by this phosphorylation. If we use an arrow for activation, we cannot properly represent the case. In the process diagram, on the other hand, whether a molecule is active or not is represented as a state of the node instead of an arrow symbol for activation. Promoting and inhibition of catalysis are represented as a modifier of state transition using a circle-headed line and a bar-headed line, respectively.

Another benefit of the process diagram is that the state transition representation of molecules will fit with a semantic of biochemical simulation model. Usually, biochemical reaction represents a state transition of molecule, not just for the binding process but also for the activation/inhibition of proteins and enzymatic reactions. While creating a biochemical computational model, it is important to add a reaction considering the transition of the target molecule with the reaction. This is of course an obvious procedure if users build their model only with mathematical equations, but with the graphical notation it is hard to represent a sequence of state transition of molecules, such as entity-relationship notation. While a process diagram is the preferred solution for representing temporal sequences, either a process diagram or entity-relationship approach could be used, depending upon the purpose of the diagram. Both notations could actually maintain compatible information internally but differ in visualization. We propose, as a basis of SBGN, a set of notations that enhances the formality and richness of the information represented. The symbols used to represent molecules and interactions are shown in Fig. 2.

The goal of SBGN is to define a comprehensive system of notation for visually describing biological networks and processes, thereby contributing to the eventual formation of a standard notation. For such a graphical notation to be practical and to be accepted by the community, it is essential that software tools and data resources be made available. Even if the proposed notation system satisfies the requirements of biologists, lack of software supports will drastically decrease its advantages. CellDesigner currently supports the majority of the process diagram notation proposed and will fully implement the all features in the near future (Fig. 3).

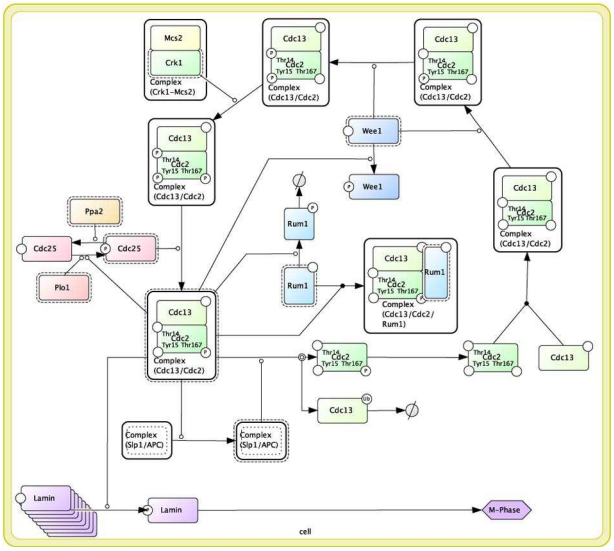

Fig. 1. ***Process diagram representation of MPF cycle. Abbreviations of protein names are as follows. Crk: cyclin-dependent kinase-related kinase; Mcs: mitotic catastrophe suppressor; Cdc: cell division cycle; Wee: small-sized mutant in Scottish; Ppa2: type 2 serine/threonine protein phosphatase a; Plo: polo kinase; Rum: strange shaped mutant; Slp: sleepy homolog; Apc: anaphase-promoting complex subunit; M-phase: mitotic phase.***

B. SBML Compliant

CellDesigner supports both reading and writing capabilities of SBML. SBML is a tool-neutral computer-readable format for representing models of biochemical reaction networks, applicable to metabolic networks, cell signaling pathways, gene regulatory networks, and other modeling problems in systems biology [2], [3]. SBML is based on XML, a simple, flexible text format for exchanging a wide variety of data. The initial version of the specification was released on March 2001 as SBML Level-1. The most recent released version of SBML is Level-2 Version 3 (as of June 2008). Currently, SBML is supported by more than 110 software systems and is now widely accepted and used. CellDesigner uses SBML as its native model description language; therefore once a model is created using CellDesigner, all the information inside the model will be stored in SBML, resulting in high model portability. For example, genes and proteins are stored as a list of ⟨species⟩ under ⟨listOfSpecies⟩ tag, and reactions are stored as a list of ⟨reactions⟩ under ⟨listOfReactions⟩ tag. Kinetic laws, which are required for ordinary differential equation (ODE)-based simulation, are stored under ⟨kineticLaw⟩ tags, which are also compatible with the Mathematical Markup Language (MathML) standard (Fig. 4). As mentioned, CellDesigner draws a pathway with its specialized graphical notation. Since such layout information has not been supported by SBML, CellDesigner stores its layout information under an ⟨annotation⟩ tag, which does not conflict with the current SBML specification. There is a working group of layout extension for SBML that will be incorporated in SBML Level-3. We are currently under way to implement a conversion module to export SBML layout extension from CellDesigner. If the SBML model has no CellDesigner compatible layout information, an autolayout function can be run to lay out SBML Level-1 and Level-2 models. By using this

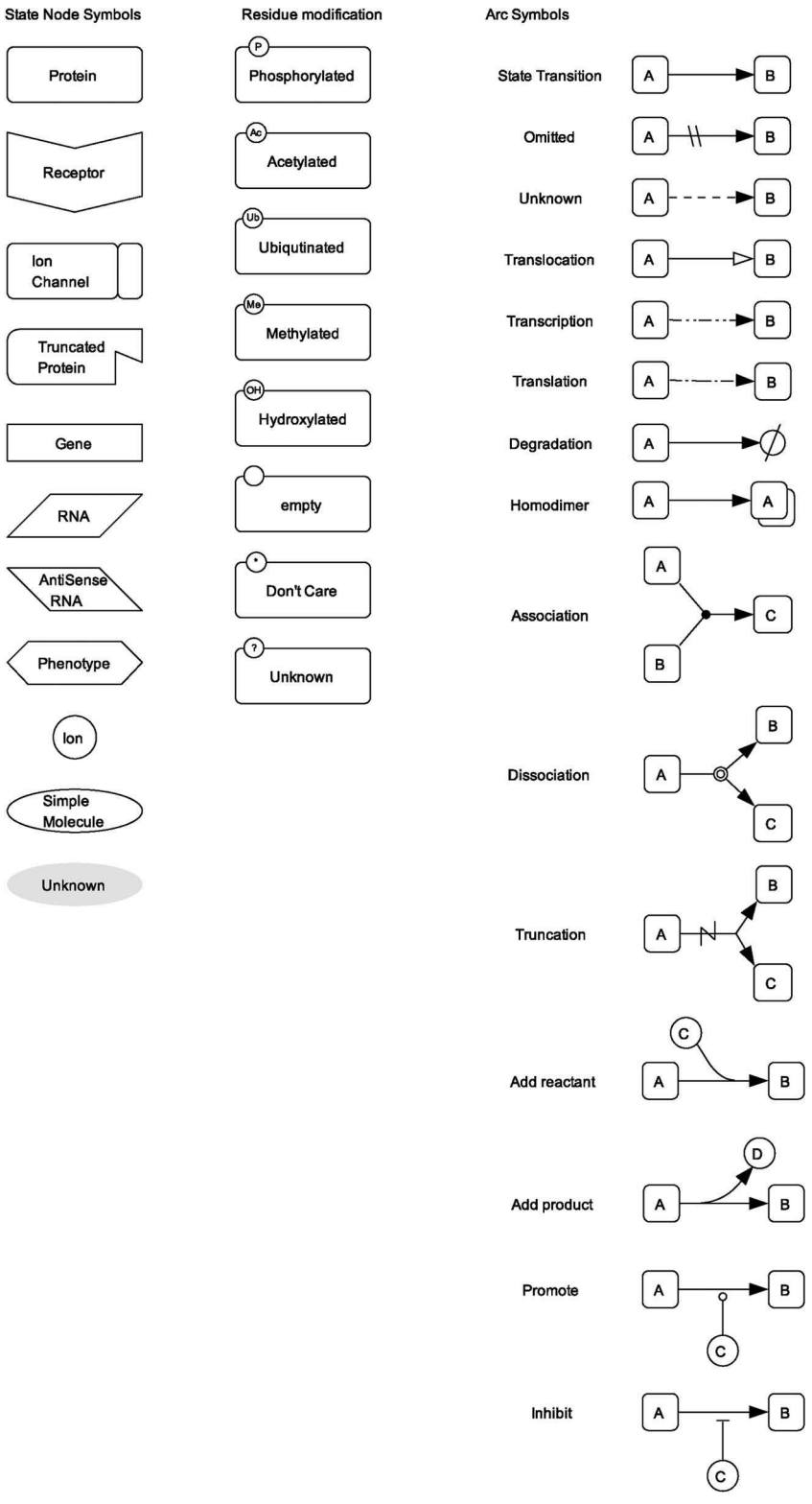

Fig. 2. ***Proposed set of symbols for representing biological networks with process diagrams.***

function, users can quickly lay out existing SBML models such as the Kyoto Encyclopedia of Genes and Genomes (KEGG), a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals, [13] converted models, and models from the BioModels [14] database. We have converted more than 12 000 metabolic

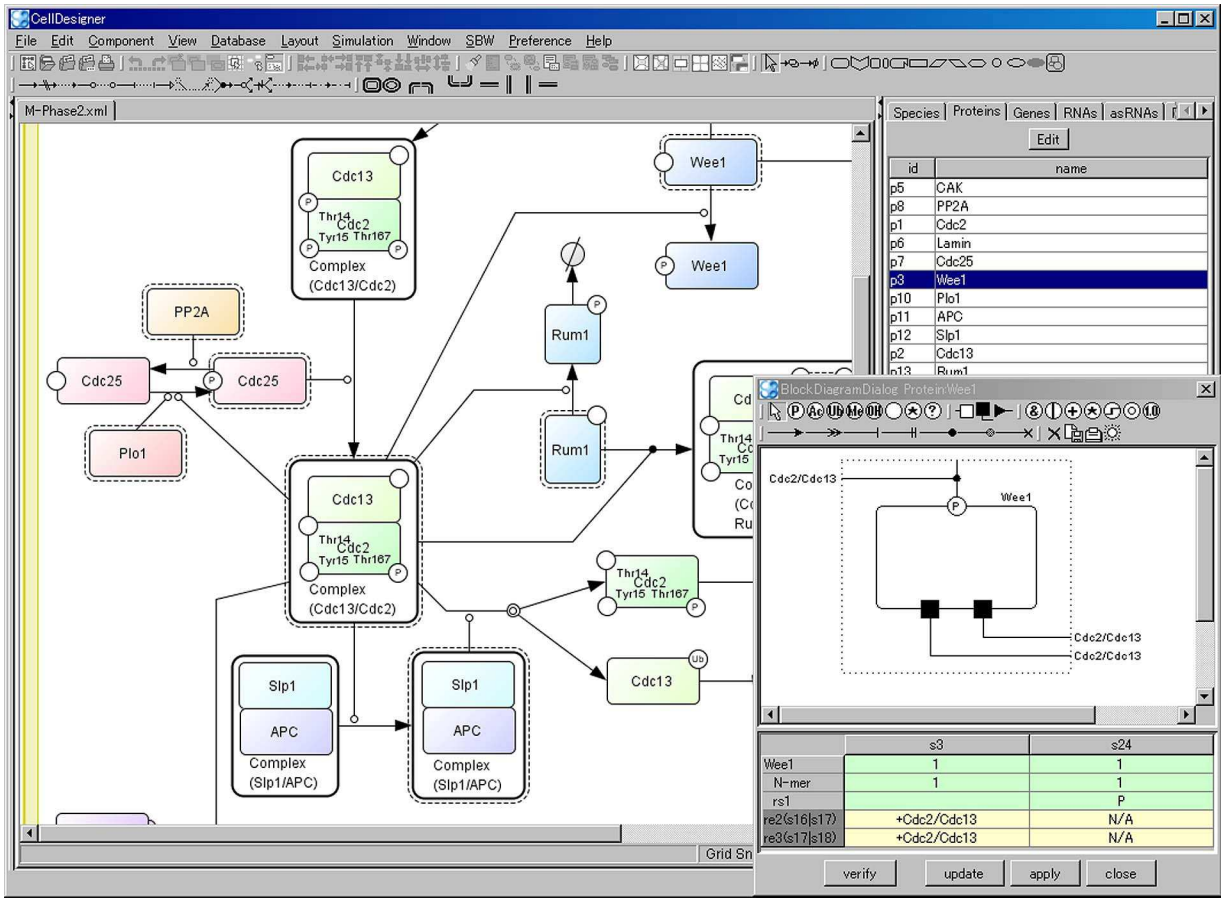

Fig. 3. ***Screenshot of CellDesigner.***

pathways of KEGG to SBML.[4] Other SBML models are available from the BioModels database.[5] Users can also use our own SBML models created by CellDesigner on other SBML compliant applications.[6]

C. SBW Enabled

CellDesigner is an SBW-enabled [4] application. In other words, CellDesigner could integrate all SBW-enabled modules (Fig. 5). For example, users could browse or modify a model converted from an existing database with CellDesigner and launch a simulator from CellDesigner (by selecting Simulation Service or Jarnac Simulation Service from the SBW menu) to run simulations in real time. There are many other SBW-enabled modules, such as the ODE-based simulator, stochastic simulator, MATLAB, FORTRAN translator, bifurcation analysis tool, and optimization module.[7]

D. Simulation Capability

One of our aims is to use CellDesigner as a simulation platform, and thus integration capability with native simulation library has been implemented. An SBML ODE solver [6] could be invoked directly from CellDesigner, which enables users to run ODE-based simulations. The SBML ODE Solver Library (SOSlib) is a programming library for symbolic and numerical analysis of chemical reaction network models encoded in SBML.

It is written in ISO C and is distributed under the open-source GNU Lesser General Public License. The SBML ODE Solver can read SBML models by using libSBML[8] and then construct a set of ODEs and their Jacobian matrix, and so forth. The SBML ODE Solver uses SUNDIALS CVODES [15] for numerical integration and sensitivity analysis. CVODES is a solver for stiff and nonstiff ODE systems (initial value problem) with sensitivity analysis capabilities (both forward and adjoint modes). The methods used in CVODES are variable-order variable-step multistep methods. For nonstiff problems, CVODES includes the Adams–Moulton formulas. For stiff problems,

[4]The pathways are available from http://www.systems-biology.org.
[5]See http://www.ebi.ac.uk/biomodels/.
[6]See http://www.systems-biology.org/001/.
[7]These SBW-enabled modules are freely available from http://sys-bio.org.
[8]http://www.sbml.org/software/libsbml/.

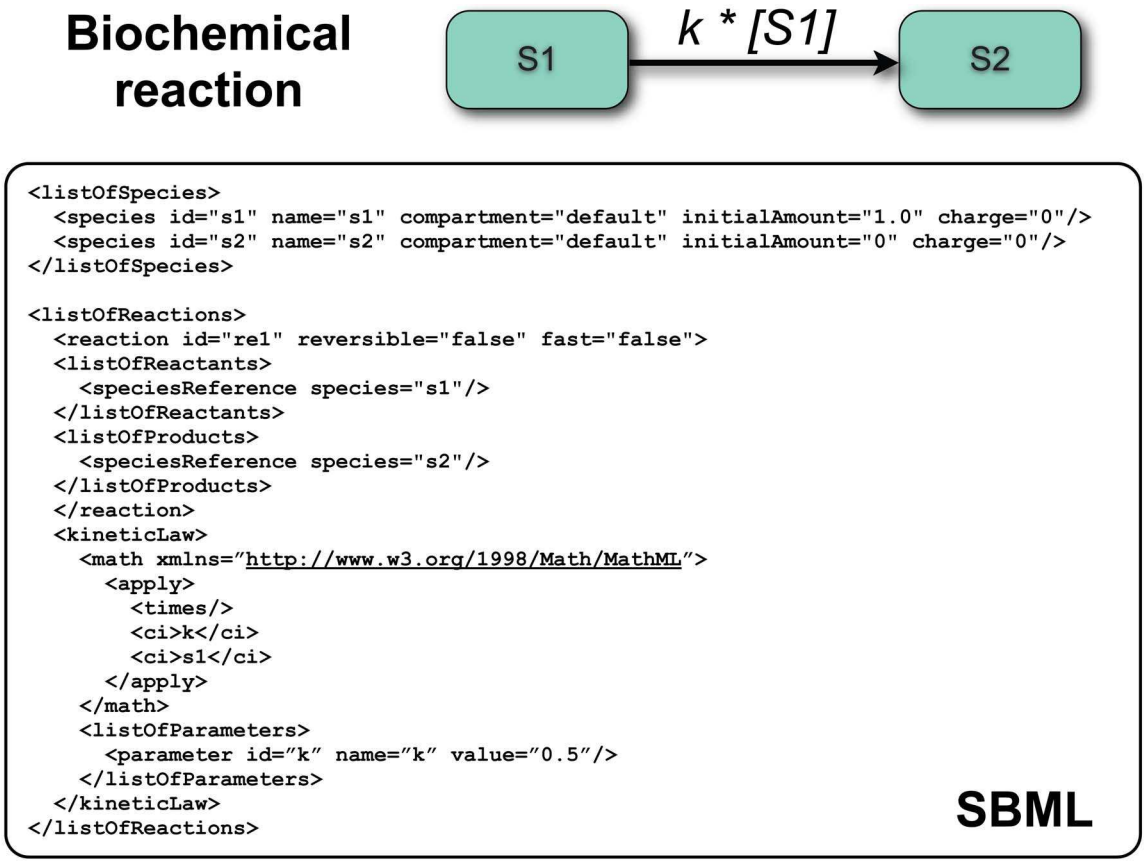

Fig. 4. *SBML representation of biochemical reaction with kinetic law.*

CVODE includes the backward differentiation formulas (BDFs) in so-called fixed-leading coefficient form. Both integration methods (Adams–Moulton and BDF) and the corresponding nonlinear iteration methods, as well as all linear solver and preconditioner modules, are available for the integration of the original ODEs, the sensitivity systems, or the adjoint system.

The performance of the simulation engine is a critical issue for a simulation platform, so we have wrapped the C application programming interface (API) of the SBML ODE Solver from Java by using Java Native Interface.[9] This resulted in small overhead of simulation execution time compared with the native library, and thus the broad support of multiple OSs. The simulation engine itself is executed by the native library, and the results are shown in a graphical user interface window written in Java (Fig. 6). Fig. 6 shows a simulation result of a mitogen-activated protein kinase (MAPK) cascades model proposed by Kholodenko [16]. Each line represents the oscillatory behavior of MAPK concentration. The dynamics of the model will change depending on the set of parameters in the model. In CellDesigner, users can change the value of the selected parameter through a control panel. Users are often required to execute multiple times of simulation with different parameter set within a specified parameter range to find an exact parameter set to reproduce the results obtained from experiments. Through the interface, it is possible to execute multiple simulations as a "batch" function with different parameter set. It is also possible to execute this batch function with two different parameters, with different parameter range for each parameter. An intuitive interface such as sliderbars is also implemented. Users can change the parameter just by dragging the sliderbar, and the simulation result plot will be generated immediately. This allows users to easily understand the behavior of a model. Furthermore, the control panel allows users to change the concentration of molecules or parameter values at specified time during the simulation. This feature is useful because some biological experiments are not just observing a stable state of a biological system but also observing a response of the system (in other words, how the biological system reaches to stable state) from external stimuli. This feature enables users to simulate the dynamics of a model with under such condition. The simulation results can be exported to CSV so that it will be used for analytical work, and exporting to JPEG, PNG format, and various bitmap formats is also supported.

[9]http://java.sun.com/j2se/1.5.0/docs/guide/jni/.

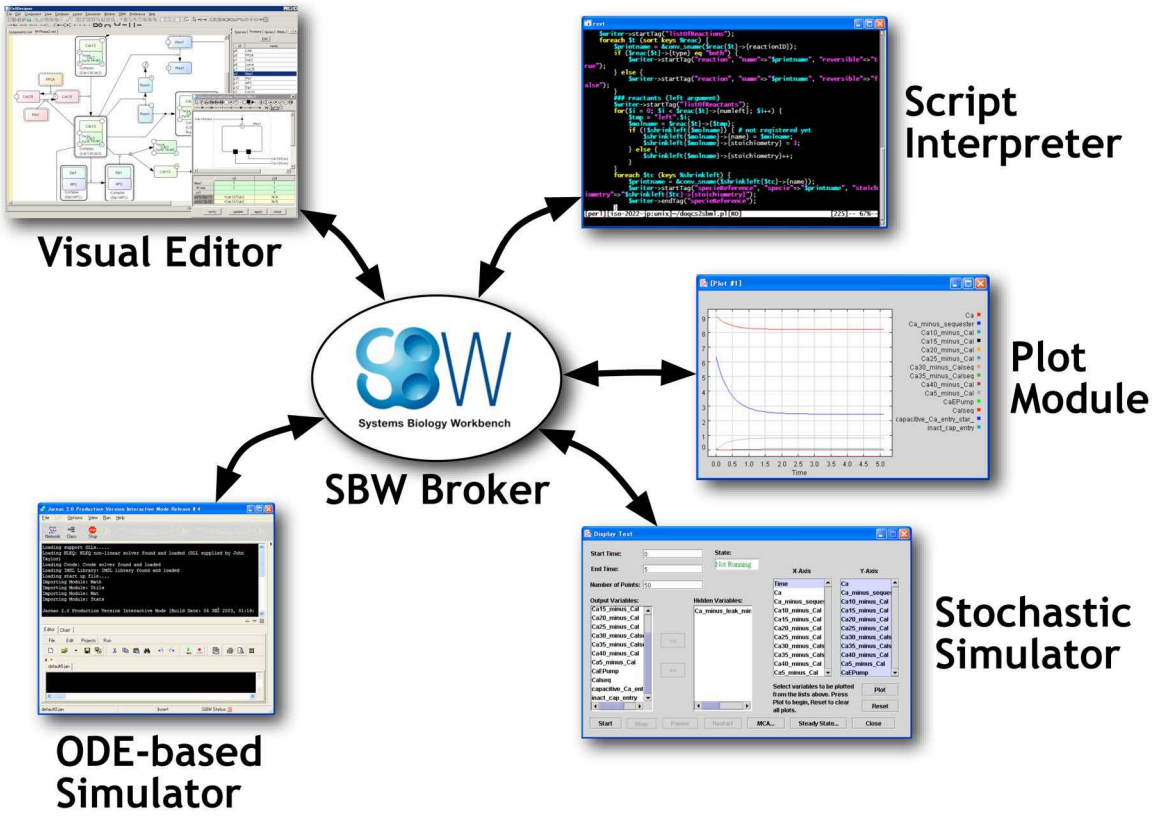

Fig. 5. *Illustration of the relationship between SBW broker and SBW modules.*

E. Database Connection Capability

To efficiently conduct network analysis, connection with databases is significant because users may want to further examine network characteristics. We have added this capability, enabling direct connection with the following databases:

- BioModels: database of annotated computational models [14][10];
- SGD: Saccharomyces Genome Database [17][11];
- DBGET: database retrieval system for a diverse range of molecular biology databases [18][12];
- iHOP: Information Hyperlinked Over Proteins [19][13];
- PubMed[14];
- Entrez Gene [20].[15]

Once a node or an edge on the process diagram is selected, users can query databases via a popup menu, from which the database could be chosen to query according to the internal information of the selected object. For example, the PubMed ID search utilizes notes written in the components. The BioModels database connection allows importing SBML-based models, which are curated computational models prepared for simulations. This enables users to efficiently open and simulate the BioModels inventory.

F. General Workflow of CellDesigner

CellDesigner consists of four areas such as Draw, List, Notes, and Tree Area. Draw is the main part of CellDesigner, which is used to draw and edit a model on a canvas. A diagram drawn on the canvas will be treated as an SBML model in CellDesigner. The model consists of species (chemical or biological molecules) as nodes and reactions as edges, and also, the model may have compartments that represent an area of reaction space (e.g., cell, nucleus, etc.) and mathematical equations/rules. The List Area displays a list of components of the model and is also an editable area. Users can modify a name value of each component in this area. The Notes Area displays the "notes" elements of the selected component (nodes, edges, and compartments). In CellDesigner, each component can store external information (e.g., database accession number, URL, etc.) in notes, and such information is used to call databases from CellDesigner (as shown in Section II-E). The Notes Area is also an editable area so that users can add external information to the selected component.

[10]http://www.biomodels.net/.
[11]http://www.yeastgenome.org.
[12]http://www.genome.ad.jp/dbget/.
[13]http://www.ihop-net.org/.
[14]http://www.pubmed.gov.
[15]http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

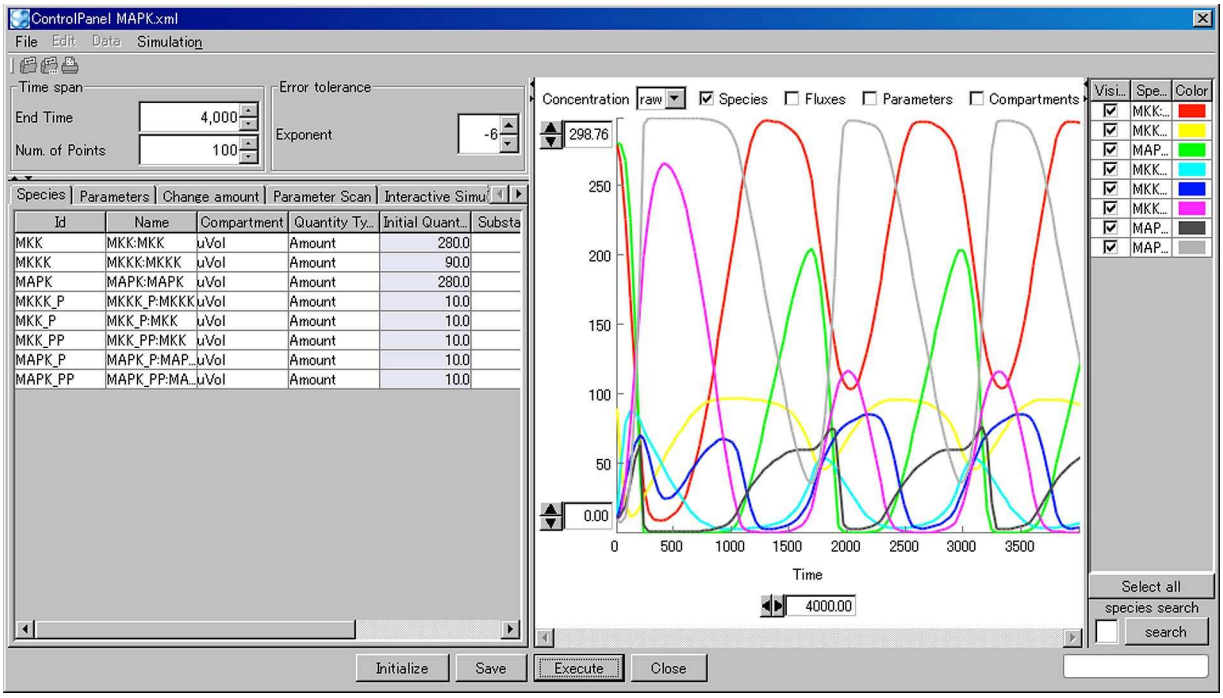

Fig. 6. ***Screenshot of a simulation result obtained by integration with SBML ODE solver.***

Building models with CellDesigner is quite straightforward. To create a model, select "New" from the "File" menu and then input the name and size of a model—a new canvas will appear in the Draw Area. The name specified in this procedure will also be the default file name of this model. Users can then place a species on the canvas as a node such as protein, gene, RNA, ion, simple molecule, and so forth from the notation defined by us [7]. When adding a species on the canvas, CellDesigner will ask a name of the species. Users can move the species by dragging and dropping, and change the size of each species by dragging the corner of species. It is possible to define the default size and color of each species from "Components Color & Shape" from the "Preference" menu.

To draw a reaction between species, at first a type of reaction should be selected from the user interface buttons. Reactant species (start node) and product species (end node) should be selected after then. To add more reactants to an existing reaction, select the "Add reactant" button and then choose a species for the reactant and the reaction. CellDesigner can also represent common types of reactions, such as catalysis, inhibition, activation, and so forth. The procedure for representing such reactions is just the same as adding reactants (or products) to an existing reaction; select a species (which will be a modifier of targeted reaction) followed by the target reaction. Users can also easily modify symbols of proteins with modification residues, and hence, can describe detailed state transitions between species of an identical protein by adding different modifications.

CellDesigner is not just a drawing tool for biochemical networks; it is also possible to add mathematics into the model so that the model will be simulated by one of available simulators. We assume that the users will use ODE to represent a rate equation in the model. By choosing "Edit Reaction" from a right-click menu on a selected reaction, a new dialog will open that enables users to add kinetic equation, parameters, and properties to the reaction. Free-text format is used to input the kinetic equation from a text field, but it will be stored as a MathML object inside CellDesigner. Assigning kinetic equations, required parameters to reactions, and initial values (concentration or amount) to all the species in the model, the model will be a computational model.

To run a simulation with the model, select the "Simulation" menu, which in turn calls the SBML ODE Solver for solving a set of ODEs included in the model. A Control Panel then appears, enabling users to specify the details of parameters, to change amount of specific species, to conduct parameter search, and to run a simulation interactively.

The created model will be stored as an SBML document, which contains all the necessary information referring to species, reactions, modifiers, layout information (geometry), state transitions of proteins, modification residues, and so on. Although the above information (especially on the graphical part) is not fully supported by the current version of SBML, it is possible to store such information inside SBML as external information (annotation) so that the created SBML model can be used on other SBML-compliant

applications. This feature is easily used through SBW menu if users have installed SBW and SBW-enabled applications. CellDesigner will pass the model to other third-party applications via SBW so that the model will be simulated/analyzed with other specific applications.

III. UNIQUE CELLDESIGNER ASPECTS

Currently, many other applications support pathway design features. Here is a list of applications that contains pathway design features for computational models:

- BioTapestry: freely available (open source), platform independent, SBML supported [21];
- BioUML: freely available (open source), platform independent, SBML supported, built-in simulator, plug-in architecture [22];
- Cellware: freely available, platform independent, SBML supported, built-in simulator, grid enabled [23];
- E-Cell: freely available (open source), SBML supported, built-in simulator, extendable architecture [24];
- Edinburgh Pathway Editor (EPE): freely available, platform independent, SBML supported, SBGN support ongoing, plug-in architecture [25];
- JDesigner: freely available (open source), SBML and SBW supported[16];
- Narrator: freely available (Java applet), platform independent, partially supports SBML, built-in simulator [26];
- PathwayLab: commercial, SBML supported, built-in simulator[17];
- ProMoT: freely available, SBML supported, simulation environment (DIANA) is also available, modular models [27];
- SimBiology: commercial (requires MATLAB), platform independent, SBML supported, built-in simulator[18];
- SmartCell: freely available, platform independent, SBML supported, built-in simulator, diffusion-reaction model [28];
- TERANODE Design Suite: commercial, SBML supported, built-in simulator, hierarchical pathways.[19]

The advantages of CellDesigner over other tools are as follows:

- based on standard technology (i.e., SBML compliant and SBW enabled);
- supports clearly expressive and unambiguous graphical notation systems (e.g., clear representation of eventual standard formation);
- platform independent (i.e., Windows, Mac OS X, Linux).

[16]http://www.jdesigner.org/.
[17]http://www.innetics.com/.
[18]http://www.mathworks.com/products/simbiology/.
[19]http://www.teranode.com/products/tds/index.php.

As described above, the aim of the development of CellDesigner is to supply a process diagram editor with standardized technology for every computing platform, so that it will benefit as many biological researchers as possible. Some of the existing applications are SBML-compliant and some run on multiple computer platforms. These tools are powerful in some aspects. However, they are not intended to support the features as CellDesigner. Some of them have the facility to create pathways and some also include a simulation engine or database integration module. CellDesigner does include a simulation engine provided by the SBML ODE Solver development team, and it is able to cooperate with other SBW-enabled applications so that the user could switch the simulation engines on the fly. Furthermore, we have converted some existing databases to SBML (e.g., KEGG) so that one can easily browse them with other SBML-compliant applications, edit the models, and even simulate via CellDesigner. The overriding advantage of CellDesigner is that it uses open and standard technologies. The models created by CellDesigner could be used on many other (more than 110) SBML-compliant applications, and its graphical notation system will make the representation of models in a more efficient and accurate manner. Survey results of standards on systems biology [29] show that about 80% of the survey respondents consider that the creation of standards is necessary or desirable because the standards will improve a collaboration, communication between software tools, and reduce the duplication of work. Not just the standard model description language and the integration framework for software tools, graphical representation of biochemical networks is also listed as the need for the standardization.

IV. FUTURE WORK

In future releases of CellDesigner, we plan to implement further capabilities. Integration with other modules is under way, such as other simulation, analysis, and database modules. The current version of CellDesigner has been implemented as a Java application, but we are developing a Java Web Start version of CellDesigner so that it could be used as a Web-based application as well. To be widely used by users from biologists to theorists, we believe that it is essential to meet the standard. We are thus actively working as SBML and SBGN working group members, which aims to establish *de facto* standards in systems biology field; the former one seems to have already become *de facto* as model description language. SBML Level-3 (the next version) will include layout extension, and we will incorporate the functions in our new release of CellDesigner.

BioPAX[20] [30] is another big activity that tries to connect widely distributed data resources seamlessly. We also plan to connect CellDesigner with the BioPAX data

[20]http://www.biopax.org.

format so that users can use CellDesigner from BioPAX platform and vice versa. From software development perspectives, providing API, plug-in interface, or open-source strategy is a solution to speed up the development and enable users to customize the software depending on users' needs. While we have been providing CellDesigner as a binary program so far, we have been working to extend our development scheme in such a manner. Currently, an alpha release version of CellDesigner (4.0 alpha) supports plug-in development framework so that users can call CellDesigner's API from their plug-in using Java. Plug-in API enables one to obtain and modify information of the model, which includes the graphical (layout) information and simulation parameters and all of the SBML elements. Some other enhancement is also under development. We are now implementing a new integration scheme with SABIO-RK [31], which has the potential to expand connectivity and semi-automate visualization and model building. SABIO-RK contains information about biochemical reactions, related kinetic equations, and parameters. Also information about the experimental conditions under which these parameters were measured is stored. By using the Web service [32] API provided by the SABIO-RK team [33], the integration will enable CellDesigner to directly connect to the database, send search queries by ID or the name of its component, and then import the query results into CellDesigner.

We wish CellDesigner to be used by anyone who is working in a biology-related field. As described throughout this paper, CellDesigner is designed to be user-friendly as much as possible, allowing users to draw pathway diagrams as easily as drawing with other drawing tools. Since our proposed notation itself (and along SBGN definition in future releases) is rigidly defined, the diagrams could be used for a presentation or even for a knowledge base. The diagrams could be used as figures in a manuscript or a pathway representation of databases. Since the definition of the pathway diagram notation is now getting much attention (which has resulted in the formation of an SBGN working group,[21] we hope the notation will be much refined as a *de facto* standard representation, which will be reflected in the representation manner of CellDesigner as well.

Our concept for developing CellDesigner is "easy to create a model, to run a simulation, and to use analysis tools." This will be achieved by extending the development of corresponding native libraries or SBW-enabled modules. Improvement of the graphical-user interface is also required, including the mathematical equation editor, so that the user could easily write equations by selecting and dragging a species.

V. CONCLUSION

We have introduced CellDesigner, a process diagram editor for gene-regulatory and biochemical networks based on standardized technologies and with wide transportability to other SBML-compliant applications and SBW-enabled modules. Since the first release of CellDesigner, 21 000 downloads have already occurred. CellDesigner also aims to support the standard graphical notation. Since the standardization process is still under way, our technologies are still changing and evolving. As we are in partnership with the SBML, SBW, and SBGN working groups, we will go through with these standardization projects and hence improve the quality of CellDesigner. The current version of CellDesigner is 3.5.2, which runs on multiple platforms such as Windows, Linux, and Mac OS X.[22] ■

Acknowledgment

The authors thank R. Machné and Dr. C. Flamm (University of Vienna) for providing them a library version of the SBML ODE Solver; F. Bergmann and Dr. H. Sauro (University of Washington) for providing them a new simulation driver module for SBW-2.x; and Dr. D. Murray (Keio University) and Dr. N. Hiroi (University of Vienna) for fruitful discussions. They also thank members of the SBML and SBGN community for fruitful discussions on standardization.

[21]http://www.sbgn.org.

[22]CellDesigner is freely available from http://www.celldesigner.org.

REFERENCES

[1] H. Kitano, "Systems biology: A brief overview," *Science*, vol. 295, no. 5560, pp. 1662–1664, 2002.

[2] M. Hucka *et al.*, "The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models," *Bioinformatics*, vol. 19, no. 4, pp. 524–531, 2003.

[3] M. Hucka, A. Finney, B. J. Bornstein, S. M. Keating, B. E. Shapiro, J. Matthews, B. L. Kovitz, M. J. Schilstra, A. Funahashi, J. C. Doyle, and H. Kitano, "Evolving a Lingua Franca and associated software infrastructure for computational systems biology: The Systems Biology Markup Language (SBML) Project," *Inst. Elect. Eng. Syst. Biol.*, vol. 1, no. 1, pp. 41–53, 2004.

[4] H. M. Sauro, M. Hucka, A. Finney, C. Wellock, H. Bolouri, J. C. Doyle, and H. Kitano, "Next generation simulation tools: The Systems Biology Workbench and BioSPICE integration," *OMICS*, vol. 7, no. 4, pp. 355–372, 2003.

[5] A. Funahashi, N. Tanimura, M. Morohashi, and H. Kitano, "CellDesigner: A process diagram editor for gene-regulatory and biochemical networks," *BIOSILICO*, vol. 1, no. 1, pp. 159–162, 2003.

[6] R. Machné, A. Finney, S. Müller, J. Lu, S. Widder, and C. Flamm, "The SBML ODE solver library: A native API for symbolic and fast numerical analysis of reaction networks," *Bioinformatics*, vol. 22, no. 11, pp. 1406–1407, 2006.

[7] H. Kitano, A. Funahashi, Y. Matsuoka, and K. Oda, "Using process diagrams for the graphical representation of biological networks," *Nature Biotechnol.*, vol. 23, no. 8, pp. 961–966, 2005.

[8] K. Kohn, "Molecular interaction map of the mammalian cell cycle control and DNA repair systems," *Mol. Biol. Cell*, vol. 10, no. 8, pp. 2703–2734, 1999.

[9] K. Kohn, "Molecular interaction maps as information organizers and simulation guides," *Chaos*, vol. 11, no. 1, pp. 84–97, 2001.

[10] I. Pirson, N. Fortemaison, C. Jacobs, S. Dremier, J. E. Dumont, and C. Maenhaut, "The visual display of regulatory information and networks," *Trends Cell Biol.*, vol. 10, no. 10, pp. 404–408, 2000.

[11] D. L. Cook, J. F. Farley, and S. J. Tapscott, "A basis for a visual language for describing, archiving and analyzing functional models of complex biological systems," *Genome Biol.*, vol. 2, no. 4, 2001, RESEARCH0012.

[12] R. Maimon and S. Browing, "Diagramatic notation and computational structure of gene networks," in *Proc. 2nd Int. Conf. Syst. Biol.*, 2001, pp. 311–317.

[13] M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa, "From genomics to chemical genomics: New developments in KEGG," *Nucl. Acids Res.*, vol. 34, pp. D354–D357, 2006.

[14] N. L. Novère, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M. Schilstra, B. Shapiro, J. L. Snoep, and M. Hucka, "BioModels database: A free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems," *Nucl. Acids Res.*, vol. 34, pp. D689–D691, 2006.

[15] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward, "SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers," *ACM Trans. Math. Software*, vol. 31, no. 3, pp. 363–396, 2005.

[16] B. N. Kholodenko, "Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades," *Eur. J. Biochem.*, vol. 267, no. 6, pp. 1583–1588, 2000.

[17] J. M. Cherry, C. Adler, C. Ball, S. A. Chervitz, S. S. Dwight, E. T. Hester, Y. Jia, G. Juvik, T. Roe, M. Schroeder, S. Weng, and D. Botstein, "SGD: Saccharomyces genome database," *Nucl. Acids Res.*, vol. 26, no. 1, pp. 73–79, 1998.

[18] W. Fujibuchi, S. Goto, H. Migimatsu, I. Uchiyama, A. Ogiwara, Y. Akiyama, and M. Kanehisa, "DBGET/LinkDB: An integrated database retrieval system," in *Proc. Pacific Symp. Biocomput.*, 1998, pp. 683–694.

[19] R. Hoffmann and A. Valencia, "A gene network for navigating the literature," *Nature Genetics*, vol. 36, no. 7, p. 664, 2004.

[20] D. Maglott, J. Ostell, K. D. Pruitt, and T. Tatusova, "Entrez gene: Gene-centered information at NCBI," *Nucl. Acids Res.*, vol. 35, pp. D26–D31, 2007.

[21] W. J. R. Longabaugh, E. H. Davidson, and H. Bolouri, "Computational representation of developmental genetic regulatory networks," *Dev. Biol.*, vol. 283, pp. 1–16, 2005.

[22] F. A. Kolpakov, "BIOUML—Framework for visual modeling and simulation biological systems," in *Proc. Int. Conf. Bioinf. Genome Regulat. Structure (BGRS'2002)*, 2002.

[23] P. Dhar, T. C. Meng, S. Somani, L. Ye, A. Sairam, M. Chitre, Z. Hao, and K. Sakharkar, "Cellware—A multi-algorithmic software for computational systems biology," *Bioinformatics*, vol. 20, no. 8, pp. 1319–1321, 2004.

[24] K. Takahashi, N. Ishikawa, Y. Sadamoto, H. Sasamoto, S. Ohta, A. Shiozawa, F. Miyoshi, Y. Naito, Y. Nakayama, and M. Tomita, "E-Cell 2: Multi-platform E-Cell simulation system," *Bioinformatics*, vol. 19, no. 13, pp. 1727–1729, 2003.

[25] A. Sorokin, K. Paliy, A. Selkov, O. Demin, S. Dronov, P. Ghazal, and I. Goryanin, "The Pathway editor: A tool for managing complex biological networks," *IBM J. Res. Develop.*, vol. 6, pp. 561–575, 2006.

[26] J. J. Mandel, H. Fuß, N. M. Palfreyman1, and W. Dubitzky, "Modeling biochemical transformation processes and information processing with Narrator," *BMC Bioinf.*, vol. 8, p. 103, 2007.

[27] M. Ginkel, A. Kremling, T. Nutsch, R. Rehner, and E. D. Gilles, "Modular modeling of cellular systems with ProMoT/Diva," *Bioinformatics*, vol. 19, no. 9, pp. 1169–1176, 2003.

[28] M. Ander *et al.*, "SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: Analysis of simple networks," *Syst. Biol.*, vol. 1, pp. 129–138, 2004.

[29] E. Klipp, W. Liebermeister, A. Helbig, A. Kowald, and J. Schaber, "Systems biology standards-the community speaks," *Nature Biotechnol.*, vol. 25, pp. 390–391, 2007.

[30] J. S. Luciano and R. D. Stevens, "e-Science and biological pathway semantics," *BMC Bioinf.*, vol. 8, no. 3, p. S3, 2007.

[31] U. Wittig, M. Golebiewski, R. Kania, O. Krebs, S. Mir, A. Weidemann, S. Anstein, J. Saric, and I. Rojas, "SABIO-RK: Integration and curation of reaction kinetics data," in *Proc. 3rd Int. Workshop Data Integr. Life Sci. 2006 (DILS'06)*, vol. 4075, *Lecture Notes in Bioinformatics*, Hinxton, U.K., 2006, pp. 94–103.

[32] W3C Working Group Note, *Web Services Architecture*. [Online]. Available: http://www.w3.org/TR/ws-arch/

[33] *SABIO-RK Web Service Reference Manual*. [Online]. Available: http://www.sabio.villa-bosch.de/SABIORK/webservicedoc.jsp

ABOUT THE AUTHORS

Akira Funahashi received the B.E. degree in electrical engineering and the M.E. and Ph.D. degrees in computer science from Keio University, Japan, in 1995, 1997, and 2000, respectively.

His research interests include the areas of systems biology, computational biology, interconnection networks, and parallel processing. He was a Research Fellow with the Japan Society of the Promotion of Science (DC1) from 1997 to 2000 and a Research Associate with the Department of Information Technology, Mie University, Japan, from 2000 to 2002. He then joined the Kitano Symbiotic Systems Project, JST, and The Systems Biology Institute as a Researcher before joining Keio University in 2007. He is now an Assistant Professor in the Department of Biosciences and Informatics, Keio University.

Yukiko Matsuoka received the B.A. degree in liberal arts from International Christian University, Japan.

Her research interests include the area of systems biology, modeling, and graphical notations. She is a Researcher with the Kitano Symbiotic Systems Project, ERATO-SORST, JST. Previously, she was with various software companies, such as Lotus.

Akiya Jouraku received the B.E. degree in electrical engineering and the M.E. and Ph.D. degrees in computer science from Keio University, Japan, in 1998, 2000, and 2007, respectively.

His research interests include the areas of systems biology, computational biology, interconnection networks, and parallel processing. He currently is with Keio University as a Postdoctoral Researcher.

Mineo Morohashi received the B.E. degree in electrical engineering, the M.E. degree in computer science, and the Ph.D. degree in computational biology from Keio University, Tokyo, in 1996, 1998, and 2007, respectively.

His areas of interest include computational biology, metabolomics, and computer science. Previously, he was with Texas Instruments Japan as a Design Engineer to design DSP, ERATO Kitano Symbiotic Systems Project, Japan, as a Researcher to work on developing CellDesigner and establishing a methodology on computational analysis in various computational models. He then joined Human Metabolome Technologies, Japan, as a Manager of the Bioinformatics group, where he led the group to develop software platform to analyze metabolome data. He currently is with PRTM, Japan, a management consulting company, as a Consultant.

Norihiko Kikuchi received the B.S. and M.S. degrees in applied biological science from Tokyo University of Science, Tokyo, Japan, in 1997 and 1999, respectively, and the Ph.D. degree in medical science from University of Tsukuba, Japan, in 2006.

His research interests include systems biology, bioinformatics, and glycomics. He currently is with Mitsui Knowledge Industry as a Researcher.

Hiroaki Kitano received the Ph.D. degree in computer science from Kyoto University, Kyoto, Japan, in 1991.

He joined NEC Corporation in 1984 and has been a Visiting Researcher at Carnegie-Mellon University, Pittsburgh, PA, since 1988. He is now a Director of Sony Computer Science Laboratories, Inc., and President of The Systems Biology Institute, Tokyo, Japan.

Dr. Kitano received the Computers and Thought Award in 1991 and Prix Ars Electronica Special Award in 2000. He was an invited artist for La Biennale di Venezia 2000 and Worksphere exhibition at the Museum of Modern Art, New York, in 2001.