

## CELLDESIGNER: A MODELING TOOL FOR BIOCHEMICAL NETWORKS

Akira Funahashi  
Yukiko Matsuoka  
Akiya Jouraku  
Hiroaki Kitano

Norihiro Kikuchi

Mitsui Knowledge Industry Co., Ltd.  
Tokyo 164-8555, JAPAN

Kitano Symbiotic Systems  
Project, ERATO-SORST, JST  
Tokyo 150-0001, JAPAN

### ABSTRACT

Understanding of logic and dynamics of gene-regulatory and biochemical networks is a major challenge of systems biology. To facilitate this research topic, we have developed CellDesigner, a modeling tool of gene-regulatory and biochemical networks. CellDesigner supports users to easily create such networks using solidly defined and comprehensive graphical representation (SBGN: Systems Biology Graphical Notation). CellDesigner is SBML compliant, and is SBW-enabled software so that it can import/export SBML described documents and can integrate with other SBW-enabled simulation/analysis software packages. CellDesigner also supports simulation and parameter search, which is supported by integration with SBML ODE Solver, enabling us to simulate through our sophisticated graphical user interface. We could also browse and modify existing SBML models with references to existing databases. CellDesigner is implemented in Java, thus it runs on various platforms such as Windows, Linux, and MacOS X. CellDesigner is freely available via the Web.

### 1 INTRODUCTION

Systems biology is characterized by synergistic integration of theory, computational modeling, and experiment (Kitano 2002). While software infrastructure is one of the most crucial components of systems biology research, there has been no common infrastructure or standard to enable integration of computational resources. To solve this problem, the Systems Biology Markup Language (SBML, <<http://sbml.org>>) (Hucka et al. 2003) and the Systems Biology Workbench (SBW, <<http://sbw.kgi.edu>>) have been developed (Sauro et al. 2003). SBML is an open, XML-based format for representing biochemical reaction networks, and SBW is a modular, broker-based,

message-passing framework for simplified intercommunication between applications. More than 100 simulation and analysis software packages already support SBML, or are in the process to support them.

Identification of logic and dynamics of gene-regulatory and biochemical networks is a major challenge of systems biology. We believe that the standardized technologies, such as SBML, SBW, and SBGN, play an important role in the software platform of systems biology. As one such approach, we have developed CellDesigner (Funahashi et al. 2003), a process diagram editor for gene-regulatory and biochemical networks.

In the following, we will introduce the main features of CellDesigner.

### 2 FEATURES OF CELLDESIGNER

Broadly classified, the current version (3.2) 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);
- Integration with SBW-enabled simulation/analysis modules;
- Integration with native simulation library (SBML ODE Solver);
- Capability of database connections;
- Extreme portability as a Java application.

The aim in developing CellDesigner is to supply a process diagram editor with the standardized technology (SBML in this case) for every computing platform, so that it could confer benefits to as many users as possible. By using the standardized technology, the model could be easily used with other applications, thereby reducing the



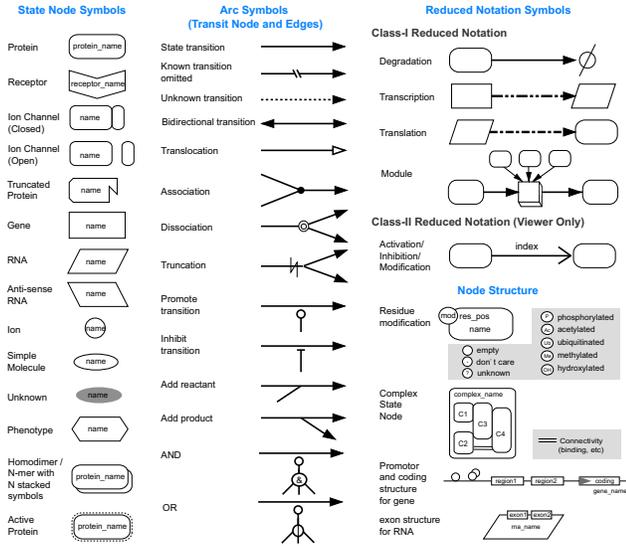


Figure 2: Proposed Set of Symbols for Representing Biological Networks with Process Diagrams

### 2.2 SBML Compliant

CellDesigner is an SBML-compliant application. It supports SBML reading and writing capabilities. 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. SBML is based on XML (eXtensible Markup Language), 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 1, and Level-2 Version 2 will be released soon (as of January 2006). Currently, SBML is supported by over 100 software systems and widely used. CellDesigner uses SBML as its native model description language, and thus once a user create a model by CellDesigner, all information inside the model will be stored in SBML and the model could be used by other software systems without any conversion of the model. 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, and will be incorporated to SBML Level-3. We are currently underway to implement a conversion module to export SBML layout extension from CellDesigner. CellDesigner has an auto layout function so that it could read all SBML Level-1 and Level-2 documents whether the model contains layout information or not. By using this function, users could use

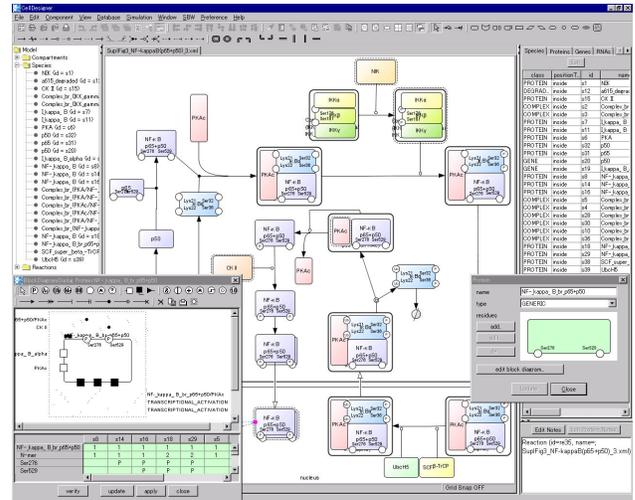


Figure 3: Screenshot of CellDesigner

existing SBML models such as KEGG, BioModels database, etc. We have converted more than 12,000 metabolic pathways of KEGG to SBML (the pathways are available from <http://systems-biology.org/>). Other SBML models are available from the BioModels Database (<http://www.ebi.ac.uk/biomodels/>). We could also use our own SBML models created by CellDesigner on other SBML compliant applications (<http://systems-biology.org/001/>).

### 2.3 SBW Enabled

CellDesigner is an SBW-enabled application. With SBW installed, CellDesigner could integrate all SBW-enabled modules (Figure 4). 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 ODE (ordinary differential equation)-based simulator, stochastic simulator, MATLAB, FORTRAN translator, bifurcation analysis tool, and optimization module. These SBW-enabled modules are freely available from <http://sbw.kgi.edu/>.

### 2.4 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. SBML ODE Solver (Machné et al. 2006) could be invoked directly from CellDesigner, which enables us to run ODE-based simulations. The SBML ODE Solver Library (SOSlib) is a programming library for symbolic and numerical analysis



engine or database integration module. CellDesigner does include a simulation engine provided by the SBML ODE Solver development team, and also it could connect to other SBW-enabled applications so that the user could switch the simulation engine on the fly. Furthermore, we have been converting existing databases to SBML (e.g., KEGG), and one can easily browse all SBML-compliant applications, edit the models, and even simulation 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 (over 100) SBML-compliant applications, and its graphical notation system will make the representation of models in a more efficient and accurate manner.

#### 4 FUTURE WORK

In a future release of CellDesigner, we plan to implement further capabilities. Improvement of auto layout function is a big issue. The bigger (e.g., more than a few hundred nodes) the network diagram becomes, the slower the performance of CellDesigner becomes, which causes our current version not to align nodes and edges quite well. Integration with other modules is also underway, such as other simulation, analysis, and database modules. The current version of CellDesigner has been implemented as a Java application, while we are developing a JWS (Java Web Start) version of CellDesigner so that it could be used as a web-based application as well.

To be widely used 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 (next version) will include layout extension, and we will incorporate the functions in our new release of CellDesigner. BioPAX (<<http://www.biopax.org>>) is another big activity, which tries to connect widely distributed data resources seamlessly. We also plan to connect CellDesigner with the BioPAX data format so that users could use CellDesigner from BioPAX platform and vice versa.

From software development perspectives, providing API, plug-in interface, or open source strategy might be a solution to speed up the development, and enable users to customize the software depending on users needs. While we have been providing a binary program of CellDesigner so far, we are now working to extend our development scheme in such manner.

We wish CellDesigner to be used by anyone who is working in a biology-related field. As described throughout this manuscript, CellDesigner is designed to be user-friendly as much as possible, thus allowing users to draw pathway diagrams as easily as drawing with other drawing tools, such

as Microsoft Visio or Adobe Illustrator. Since our proposed notation itself is rigidly defined, the diagrams could be used for presentation or even for knowledge base. The diagrams could be used as figures in manuscript, or pathway representation of databases. Since definition of the pathway diagram notation is now getting much attention (which has resulted to form an SBGN working group (<<http://sbgn.org>>)), 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.

#### 5 CONCLUSIONS

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, 12,000 downloads has been already accomplished. CellDesigner also aims to support standard graphical notation. Since the standardization process is still underway, 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 3.2-RELEASE, and runs on multiple platforms such as Windows, Linux, and MacOS X, and is freely available from <<http://celldesigner.org/>>.

#### ACKNOWLEDGMENTS

We would like to thank Mineo Morohashi (Human Metabolome Technologies, Inc., Japan), Noriko Hiroi (Kitano Symbiotic Systems Project, JST, Japan), and Naoki Tanimura (Mizuho Information & Research Institute, Inc., Japan) for fruitful discussion. This research is in part 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), Strategic Japanese-Swedish Cooperative Program on “Multidisciplinary BIO” (JST-VINNOVA/SSF), Establishment of a Human Genome Network Platform (MEXT) and through the special coordination funds for promoting science and technology from the Japanese government’s Ministry

of Education, Culture, Sports, Science, and Technology (MEXT).

## REFERENCES

- Cook, D. L., J. F. Farley, and S. J. Tapscott. 2001. A basis for a visual language for describing, archiving and analyzing functional models of complex biological systems. *Genome Biology* 2 (4): RESEARCH0012.
- Funahashi, A., N. Tanimura, M. Morohashi, and H. Kitano. 2003. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *BIOSILICO* 1 (1): 159–162.
- Hucka, M., A. Finney, et al. 2003. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19 (4): 524–531.
- Hucka, M., 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. 2004. Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language (SBML) project. *Systems Biology* 1 (1): 41–53.
- Kitano, H. 2002. Systems biology: A brief overview. *Science* 295 (5560): 1662–1664.
- Kitano, H., A. Funahashi, Y. Matsuoka and K. Oda. 2005. The process diagram for graphical representation of biological networks. *Nature Biotechnology* 23 (8): 961–966.
- Kohn, K. 1999. Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Molecular Biology of the Cell* 10 (8): 2703–2734.
- Kohn, K. 2001. Molecular Interaction Maps as information organizers and simulation guides. *Chaos* 11 (1): 84–97.
- Machné, R., A. Finney, S. Müller, J. Lu, S. Widder and C. Flamm. 2006. The SBML ODE Solver Library: a native API for symbolic and fast numerical analysis of reaction networks. *Bioinformatics* 22 (11): 1406–1407.
- Maimon, R., and S. Browning. 2001. Diagrammatic notation and computational structure of gene networks. In *Proceedings of the Second International Conference on Systems Biology*, 311–317.
- Pirson I., N. Fortemaïson, C. Jacobs, S. Dremier, J. E. Dumont, and C. Maenhaut. 2000. The visual display of regulatory information and networks. *Trends in Cell Biology* 10 (10): 404–408.
- Sauro, H. M., M. Hucka, A. Finney, C. Wellock, H. Bolouri, J. Doyle, and H. Kitano. 2003. Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. *OMICS* 7 (4): 355–372.

## AUTHOR BIOGRAPHIES

**AKIRA FUNAHASHI** is a researcher of Kitano Symbiotic Systems Project, ERATO-SORST, JST. He received the B.E., M.E., and Ph.D. degrees in Computer Science from Keio University, Japan, in 1995, 1997, and 2000 respectively. He was a Research Fellow of the Japan Society of the Promotion of Science (DC1) from April 1997 to March 2000. He was also a Research Associate in the Department of Information Technology, Mie University, Japan, from May 2000 to March 2002. His research interests include the area of systems biology, computational biology, interconnection network and parallel processing. His e-mail address is <funa@symbio.jst.go.jp>, and his web page is <http://www.symbio.jst.go.jp/~funa/>.

**YUKIKO MATSUOKA** is a researcher of Kitano Symbiotic Systems Project, ERATO-SORST, JST. Prior to this assignment, she worked with various software companies, such as Lotus, IBM. She received the B.A. in Liberal Arts from International Christian University, Japan. Her research interests include the area of systems biology, modelling and graphical notations. Her e-mail address is <myukiko@symbio.jst.go.jp>.

**AKIYA JOURAKU** is a Ph.D. candidate in Keio University. He received the B.E. and M.E. degrees from Keio University, Japan, in 1998 and 2000 respectively. His research interests include the area of systems biology, computational biology, interconnection networks, and parallel processing. His e-mail address is <jouraku@am.ics.keio.ac.jp>.

**NORIHIRO KIKUCHI** is a researcher in Mitsui Knowledge Industry. He received the B.S. and M.S. in Applied Biological Science from Tokyo University of Science. His research interests include systems biology, bioinformatics, and glycomics. His e-mail address is <kikuchi-n@mki.co.jp>.

**HIROAKI KITANO** is a director of Sony Computer Science Laboratories, Inc. and President of The Systems Biology Institute, Tokyo Japan. Kitano received Ph.D. in computer science from Kyoto University in 1991. He joined NEC Corporation in 1984, and has been a visiting researcher at Carnegie Mellon University from 1988. Kitano won The Computers and Thought Award in 1991, Prix Ars Electronica Special Award in 2000, and was an invited artist for La Biennale di Venezia 2000, Workspere exhibition at MoMA New York in 2001.