

# Automatic layout tool for large-scale metabolic pathway models based on KEGG Atlas

Nobuhiro Kido<sup>1,2</sup>

dokky@g-language.org

Nobuaki Kono<sup>1,3</sup>

yves@g-language.org

Kazuharu Arakawa<sup>1,3</sup>

gaou@sfc.keio.ac.jp

Masaru Tomita<sup>1,2</sup>

mt@sfc.keio.ac.jp

<sup>1</sup> Institute for Advanced Biosciences, Keio University, Fujisawa 252-8520, Japan

<sup>2</sup> Department of Environment and Information Studies, Keio University, Fujisawa 252-8520, Japan

<sup>3</sup> Graduate School of Media and Governance, Keio University, Fujisawa 252-8520, Japan

## Abstract

In order to gain comprehensive understanding of the systematic dynamical workings of cells, large-scale models encompassing the entire metabolic pathways is critical in the next era of systems biology. However, existing modeling software tools rely on automatic layout to produce graphical representations of pathway models, which often results in overly complex layout, thus making it difficult to manually curate large-scale models. Hence we developed a software tool that lays out large-scale models according to the coordinates of the KEGG Atlas using Systems Biology Graphical Notation (SBGN), so that users can edit the large-scale metabolic model intuitively on CellDesigner.

**Keywords:** metabolic pathway, simulation model, visualization, KEGG Atlas, CellDesigner

## 1 Introduction

A grand challenge of systems biology is the reconstruction of intracellular activities *in silico*, and the understanding of complex systematic properties through computer simulation. With the availability of “omics” information and high-throughput quantitative data, genome-scale models are becoming increasingly realistic. While the modeling process is gradually expanding to handle large-scale pathway models, existing modeling software tools still primarily rely on automatic layout for the visualization of the components, which is not an exception for the popular SBML (Systems Biology Markup Language) [1] compliant graphical modeling software CellDesigner [2].

Here we introduce a web tool that takes models written in standard SBML and adds layout information using CellDesigner extensions with entities laid out based on the coordinates of KEGG (Kyoto Encyclopedia of Genes and Genomes) Atlas [3]. KEGG Atlas is a recently introduced feature in KEGG, which integrates 120 existing local maps into a single global map. By utilizing this manually curated map, our tool allows intuitive modeling of genome-scale models on CellDesigner or other modeling software that supports the graphical notation standard for SBML, SBGN (Systems Biology Graphical Notation) [4]. Genome-scale models laid out using our software tool, originally reconstructed by the GEM System (Genome-based Modeling System) [5], are available at <http://www.g-language.org/gem/>.

## 2 Software Architecture

Although the majority of edges in KEGG Atlas represent single reaction, about one third of the edges represent multiple reactions and thus contain many hidden compounds, mostly due to layout efficiency.

However, this is not sufficient for the visualization in pathway modeling software, since every component should be available for editing and manipulation. Therefore, we have rebuilt the map with hidden compounds and reactions by comprehensively checking through the relations within KEGG Atlas and LIGAND database. Components of the model that are not included in the KEGG Atlas, such as coenzymes or molecules not included in KEGG LIGAND, are automatically placed nearby related components.

Our layout tool outputs models as SBML with CellDesigner SBGN layout annotation, which is compatible with most SBML/SBGN compliant simulation software. CellDesigner can export the model as SBML with layout annotation, plain SBML (without annotation), and in graphic formats such as SVG (Scalable Vector Graphic), PDF, PNG, JPEG and EPS.

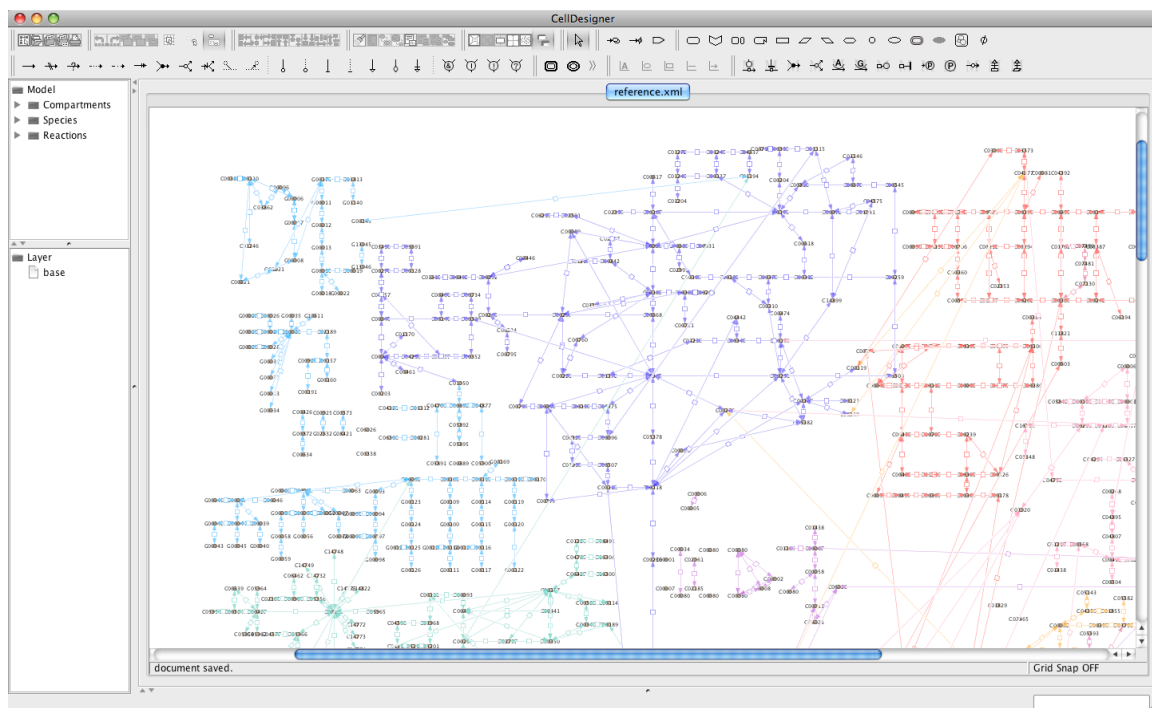


Figure 1: KEGG Atlas reference pathway map drawn on CellDesigner

## References

- [1] Hucka, M., Finney, A., Bornstein, B.J., Keating, S.M., Shapiro, B.E., Matthews, J., Kovitz, B.L., Schilstra, M.J., Funahashi, A., Doyle, J.C. and Kitano, H., Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language (SBML) project. *IEE Systems Biology*. 1:41-53, 2004.
- [2] Funahashi, A., Tanimura, N., Morohashi, M. and Kitano, H.. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *BIOSILICO*. 1:159-162, 2003.
- [3] Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., Goto, S., Kanehisa, M., KEGG Atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Res*. 36:W423-426, 2008.
- [4] Kitano, H., Funahashi, A., Matsuoka, Y., Oda, K., Using process diagrams for graphical representation of biological networks. *Nature Biotechnology*. 23:961-966, 2005.
- [5] Arakawa, K., Yamada, Y., Shinoda, K., Nakayama, Y. and Tomita, M., GEM System: automatic prototyping of cell-wide metabolic pathway models from genomes. *Bioinformatics*. 7:168, 2006.