

# The Systems Biology Markup Language (SBML) Level 2

Michael Hucka<sup>1,2</sup>, Andrew Finney<sup>1,2,3</sup>, Benjamin J. Bornstein<sup>1,2,4</sup>  
Bruce E. Shapiro<sup>1,2,4</sup>, John C. Doyle<sup>1,2</sup>, Hiroaki Kitano<sup>1,2,5</sup>

<sup>1</sup> California Institute of Technology, CA, USA

<sup>2</sup> JST/ERATO Kitano Symbiotic Systems Project, Tokyo, Japan

<sup>3</sup> University of Hertfordshire, Hatfield, UK

<sup>4</sup> NASA Jet Propulsion Laboratory, Pasadena, CA, USA

<sup>5</sup> The Systems Biology Institute, Tokyo, Japan

Over the years, a variety of biochemical network modeling packages have been developed and used by researchers interested in understanding cellular networks. No single software package currently answers all the needs of the biology community, nor is one likely to do so in the near future because the range of tools needed is vast and new techniques are emerging rapidly. It seems unavoidable that, for the foreseeable future, systems biology researchers will continue using multiple software packages to carry out their work. This diversity of resources results in several problems, including difficulties in moving models between packages and a lack of straightforward mechanisms for researchers to publish models in electronic form.

In response to these needs, the Caltech ERATO Kitano group has developed the Systems Biology Markup Language (SBML) in collaboration with a community of software developers and computational modelers. SBML is a machine-readable model definition language based upon XML, which in turn is a simple and portable text-based medium that has been gaining widespread acceptance in computational biology and bioinformatics. Many analysis, modeling and database packages now support SBML, including the following: *BASIS* (University of Newcastle), *CellDesigner* (ERATO Kitano), *Cellerator* (JPL), *Cytoscape* (ISB), *Dizzy* (ISB), *Gepasi* (Virginia Tech.), *Jarnac* (Keck Graduate Institute), *JDesigner* (Keck Graduate Institute), *JigCell* (Virginia Tech.), *BioSketchPad* (BBN), *NetBuilder* (University of Hertfordshire), *SBW* (Caltech & University of Hertfordshire), *SigPath* (Mount Sinai School of Medicine), *StochSim* (Cambridge University), and *Virtual Cell* (UCHC). In addition, SBML is the de facto model interchange standard used by the DARPA BioSPICE and IECA consortia.

SBML's current vocabulary consists of structures for specifying fundamental aspects of a system of biochemical reactions, in particular the chemical species involved in the reactions, the reaction kinetics, the volumetric compartments in which the reactions take place, definitions of mathematical functions, parameters in the system, units on quantities, discrete events and delays, and additional mathematical constraints. The SBML community has recently settled the feature set for SBML Level 2 (the latest version of SBML), and a final specification will be issued later this year. Work has already begun on SBML Level 3.

In this presentation, I will provide an overview of SBML with an emphasis on the main changes introduced in Level 2. These changes include: replacing SBML Level 1's text-string based format for mathematical expressions with MathML, introducing support for metadata using the same metadata specification as CellML, introducing support for named function definitions, and introducing new constructs for discrete events and time delays. I will also briefly discuss some of the features currently being developed for SBML Level 3, and two open-source software libraries (LIBSBML and MathSBML) for working with SBML.

More information about SBML is available online at <http://www.sbml.org>.