SBML2TikZ: supporting the SBML render extension in LaTeX

Si Yuan Shen*, Frank Bergmann and Herbert M. Sauro

Department of Bioengineering, University of Washington, Seattle, WA 98195, USA

Associate Editor: Trey Ideker

ABSTRACT

Motivation: The SBML Render Extension enables coloring and shape information of biochemical models to be stored in the Systems Biology Markup Language (SBML). Rendering of this stored graphical information in a portable and well supported system such as TeX would be useful for researchers preparing documentation and presentations. In addition, since the Render Extension is not yet supported by many applications, it is helpful for such rendering functionality be extended to the more popular CellDesigner annotation as well.

Results: SBML2TikZ supports automatic generation of graphics for biochemical models in the popular TeX typesetting system. The library generates a script of TeX macro commands for the vector graphics languages PGF/TikZ that can be compiled into scalable vector graphics described in a model.

Availability: Source code, documentation and compiled binaries for the SBML2TikZ library can be found at http://www.sbml2tikz.org. In addition, a web application is available at http://www.sys-bio.org/layout

Contact: hsauro@u.washington.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on May 25, 2010; revised on August 17, 2010; accepted on August 31, 2010

1 INTRODUCTION

The Systems Biology Markup Language (SBML) is an extensible format for storing biochemical models (Hucka *et al*., 2003). SBML promotes model exchange and longevity, and has become widely supported by numerous software tools (Alves *et al*., 2006).

In order to augment model data stored in SBML with graphical information, a render extension specification was proposed to enable the storing of coloring and shape information (Gauges *et al*., 2009). The SBML Render Extension is currently supported by a collection of libraries written by Frank Bergmann which can generate images in PNG, BMP and PDF format (http://sbmllayout.sf.net). In addition to exactly rendering SBML Render Extension information, these libraries are also capable of assigning Systems Biology Graphical Notation Styles (SBGN) to existing model layout and rendering information (Le Novère *et al*., 2009). SBGN is an increasingly popular standard notation for representing models; such support is similarly seen in tools such as Arcadia and CellDesigner that automatically render SBGN (Funahashi *et al*., 2006; Villéger *et al*., 2010). In addition to the PNG, BMP and PDF formats supported by Bergmann's libraries, the SVG format is supported as well

via a stylesheet written by Ralph Gauges (http://otto.bioquant.uni-heidelberg.de/sbml/).

SBML2TikZ extends existing support for the SBML Render Extension to TeX by generating a script of macros for PGF/TikZ. PGF/TikZ are a pair of languages for creating vector graphics through TeX macros (http://tug.ctan.org/tex-archive/graphics/pgf/base/doc/generic/pgf/pgfmanual.pdf). These macros are translated by the PGF/TikZ interpreter into images in the PDF or PS format. The benefit of having a TeX intermediate is the ease it allows graphics to be integrated with text within a TeX document. Instead of relying on imported images, a graph can be embedded within a document as a TeX script that is converted to a resolution independent image during document compilation. Once the image is compiled, it can be used in other documents as well.

The ubiquity of TeX support and extensive documentation on PGF/TikZ syntax (http://www.texample.net/tikz/) makes it possible for users to quickly edit rendering information post-file generation. This adds a layer of flexibility to graphics generation since users can now modify their graphics outside of whichever modeling environments were used to create the SBML file. At the same time, less experienced TeX users are insulated from the PGF/TikZ commands since SBML2TikZ performs all essential rendering functions such as scaling, page size selection and accurate graph generation automatically.

2 APPROACH

We extended Bergmann's libraries' classes in SBML2TikZ using extension methods in C#. This allowed us to take advantage of existing class architecture and SBML parsing capabilities while focusing attention on generating PGF/TikZ commands.

SBML2TikZ can be run from either a graphical user interface or from a console. It runs with a .NET implementation such as the .NET Framework or Mono.

In order to compile the PGF/TikZ commands generated by SBML2TikZ it is recommended that users install the PGF and geometry packages. Installation of PGF is particularly convenient with the MikTeX typesetting system (http://miktex.org) or with the Debian Universal Operating System (http://www.debian.org). The PGF/TikZ manual provides additional information on PGF installation (http://tug.ctan.org/tex-archive/graphics/pgf/base/doc/generic/pgf/pgfmanual.pdf).

3 METHODS

Two dimensional shapes are represented exactly in PGF/TikZ as rectangles, ellipses and connected segments of straight lines and Bezier curves. Colors are also implemented exactly, using their alpha, red, green and blue values; gradients of colors are implemented approximately using PGF's shading definition macros. Two dimensional transformations are applied exactly to

---
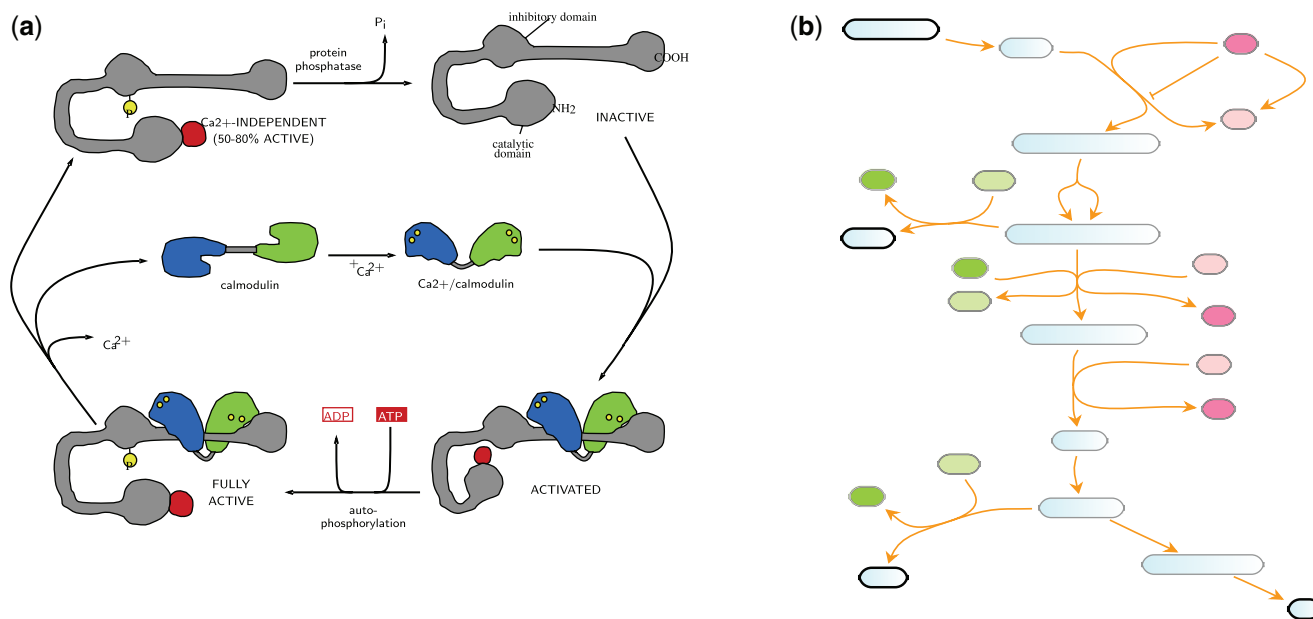
*To whom correspondence should be addressed.

Fig. 1. Graphics generated with SBML2TikZ: (**a**) Calcium/Calmodulin protein kinase model graph. The graphical information was obtained from an SBML file written by Ralph Gauges and consists of a large collection of Bezier curves joined together to form the amorphous shapes rendered above. (**b**) Glycolysis reaction pathway with species names removed to improve visual clarity. The graph makes use of rounded rectangles filled with gradients of colors and arrowhead line endings defined as filled polygons. This graph was obtained from a glycolysis model file that comes with the Systems Biology Workbench (Bergmann and Sauro, 2006; Sauro *et al.*, 2003).

all non-text elements. Figure 1 illustrates two graphs rendered using these methods.

Due to the various implementation methods for different fonts in LaTeX as well as the differing collections of fonts that come with different TeX distributions, SBML2TikZ only supports the common PostScript fonts that are guaranteed to be installed with every TeX distribution (http://ctan.tug.org/tex-archive/macros/latex/required/psnfss). Users can enable additional fonts by editing a configuration file for SBML2TikZ's graphical user interface.

To encourage user-modification of the output TeX file, sections of PGF/TikZ commands are automatically prefaced with comments describing the parts of the graph the commands render. Colors and gradients are defined at the start of the TeX file, so users can make document-wide changes to coloring information simply by editing the definitions at the beginning.

4 CONCLUSION

SBML2TikZ enables rapid and automatic generation of TeX encoded graphics from SBML files with layout and rendering information. A variety of examples and tutorials, as well as downloads for the application and its source code (available under the new BSD license) can be found at http://sbml2tikz.org

ACKNOWLEDGEMENTS

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of Interest: none declared.

REFERENCES

Alves,R. *et al*. (2006) Tools for kinetic modeling of biochemical networks. *Nat. Biotechnol.*, **24**, 667–672.

Bergmann,F.T. and Sauro,H. (2006) SBW - a modular framework for systems biology. In *Proceedings of the 38th Conference on Winter Simulation, Series WSC '06*, Winter Simulation Conference, pp. 1637–1645.

Funahashi,A. *et al*. (2006) CellDesigner: a modeling tool for biochemical networks. In *Proceedings for the 38th Conference on Winter Simulation, Series WSC '06*, Winter Simulation Conference, pp. 1707–1712.

Gauges,R. *et al*. (2009) Complementing layout information with render information in SBML files. University of Heidelberg. Available at http://otto.bioquant.uni-heidelberg.de/sbml/ (last accessed date September 12, 2010).

Hucka,M. *et al*. (2003) The systems biology markup language (SBML): a medium for presentation and exchange of biochemical network models. *Bioinformatics*, **19**, 524–531.

Novère,L.N. *et al*. (2009) The systems biology graphical notation. *Nat. Biotechnol.*, **19**, 735–741.

Sauro,H. *et al*. (2003) Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. *OMICS*, **7**, 355–372.

Villéger,A. *et al*. (2010) Arcadia: a visualization tool for metabolic pathways. *Bioinformatics*, **26**, 1470–1471.