

SynBioSS: The Synthetic Biology Modeling Suite

Anthony D. Hill^{1,2}, Jonathan R. Tomshine^{1,2}, Emma M. B. Weeding^{1,2}, Vassilios Sotiropoulos^{1,2} and Yiannis N. Kaznessis^{1,2,*}

¹Department of Chemical Engineering and Materials Science, University of Minnesota, 421 Washington Ave. SE, Minneapolis, MN 55455-0132, USA.

²Digital Technology Center, University of Minnesota, 117 Pleasant Ave. SE, Minneapolis, MN 55455, USA.

Associate Editor: Prof. Martin Bishop

ABSTRACT

Summary: SynBioSS is a suite of software for the modeling and simulation of synthetic genetic constructs. SynBioSS utilizes the registry of standard biological parts, a database of kinetic parameters, and both graphical and command-line interfaces to multiscale simulation algorithms.

Availability: SynBioSS is available under the GNU General Public License at <http://synbioass.sourceforge.net>.

Contact: yiannis@cems.umn.edu

Supplementary Information: Tutorials and documentation are available at the SynBioSS web site: <http://synbioass.sourceforge.net>.

1 INTRODUCTION

SynBioSS stands for Synthetic Biology Software Suite. As with any engineering problem, modeling the behavior of a putative synthetic biological system is a key step in the design process. To date, many of these simulations of synthetic systems have used coarse-grained models and deterministic kinetic equations (Kærn, et al., 2003; McAdams and Arkin, 1998). There have been a few notable exceptions, where detailed stochastic models spanning multiple time scales have been used (Samad, et al., 2005; Sotiropoulos and Kaznessis, 2007; Taylor, et al., 2008; Tuttle, et al., 2005). Major bottlenecks to using such models are the relative sophistication of accelerated stochastic simulation algorithms and the computational power required for a multiscale simulation of realistic complexity (Kaznessis, 2007). Moreover, while many of the promoters, operators, repressors, etc. used in synthetic biology have been well studied, the necessary quantitative information is scattered throughout the literature, making model construction challenging. SynBioSS addresses the challenges of data collection, curation, and kinetic simulation in a user-friendly fashion, facilitating the construction of detailed and accurate models along with their simulation with efficient multiscale algorithms.

2 SYNBIOSS DESIGNER

SynBioSS Designer is an application for the automatic generation of sets of biomolecular reactions. This software allows a user to

input the molecular parts involved in gene expression and regulation (e.g. promoters, transcription factors, ribosomes, etc.) The software then generates complete networks of reactions that represent transcription, translation, regulation, induction and degradation of those parts.

The BioBricks Foundation provides a database of genetic parts (<http://www.partsregistry.org>) — standardized DNA sequences that may be composed into arbitrarily complex synthetic networks through standardized ligation and cloning procedures (Canton, et al., 2008). The expanding use of these BioBricks provides a particularly interesting avenue for application of *de novo* gene network modeling. To facilitate the creation of detailed kinetic models of synthetic gene networks composed of BioBricks, we have adapted SynBioSS Designer to automatically generate a kinetic model from a construct composed entirely of BioBricks (A pictorial of the process is presented in Figure 1). A Systems Biology Markup Language (SBML) (Hucka, et al., 2003) or NetCDF (Rew and Davis, 1990) file is generated for simulations.

As described in the Supplementary text in detail, SynBioSS Designer implements simple rules of how biological molecules interact in gene regulatory networks, from transcription (e.g. RNAP holoenzyme binding a promoter), to translation (e.g. translational elongation), regulation, induction, etc.

3 SYNBIOSS WIKI

The inaccessibility of requisite kinetic data complicates the generation of detailed mechanistic models. We address this barrier by creating a web accessible database curated by users in a “Wiki” format (Leuf and Cunningham, 2001). SynBioSS Wiki is a significant extension of the open-source Mediawiki software (<http://www.mediawiki.org>). Database structures and special pages have been added to Mediawiki to support the storage, retrieval, viewing and editing of species and reaction data (see Supplementary Material for details). The Wiki (Fig. 1c) stores reaction kinetic data in a formatted and searchable scheme with references to the relevant literature. This framework allows for the input of reactions whose rates are described either by elementary first and second order rate equations or any arbitrarily complex rate equation defined using MathML (e.g. Hill type reactions). Reactions can be searched via participating molecules which may be proteins, DNA

*To whom correspondence should be addressed.

sequences, small molecules, etc. Once located, reactions of interest (along with their associated kinetic data) can be collected. The completed model can be exported in SBML format (Fig. 1d) for additional editing or simulation in SynBioSS simulation packages.

It is also through the SynBioSS Wiki databases that SynBioSS Designer can access and proliferate kinetic information related to the simulation of BioBricks, thus extending the utility of the database for the synthetic biology community. To jump-start the process, we have entered the known biomolecular interactions in the expression and regulation of well-studied operons, such as the lactose and the tetracycline operons (Sotiropoulos and Kaznessis, 2007; Tomshine and Kaznessis, 2006; Tuttle, et al., 2005).

4 SIMULATOR IMPLEMENTATION

As synthetic biology operates in femtoliter volumes, even modest concentrations equate to fewer than one million molecules per cell. Importantly, genomic DNA exists with precisely one copy per cell and the number of plasmids is often countable. Such systems are far from the thermodynamic limit and consequently cannot be accurately modeled using deterministic integrators (McAdams and Arkin, 1999; McQuarrie, et al., 1964). While the exact Stochastic Simulation Algorithm (Gillespie, 1976) can be easily overwhelmed by the multiple time scales present in protein-DNA interactions, many accelerated methods exist (Kaznessis, 2007). Both the graphical desktop and command-line supercomputer simulators in SynBioSS are based on Hy3S (Salis, et al., 2006), a hybrid jump-continuous Markov process integrator. This algorithm maintains accuracy while significantly accelerating time-consuming high-concentration interactions such as protein-protein and protein-inducer interactions. Hy3S is written in FORTRAN90 utilizing NetCDF for data storage and MPI for message passing on supercomputing clusters.

4.1 Desktop Simulator

The desktop version of the SynBioSS simulator is implemented in Python (<http://www.python.org>) using GTK+ (<http://www.gtk.org>) to provide a graphical interface. These choices enable cross-platform deployment of the same code.

Figure 16 in the supplementary material shows the main user interface in MS Windows, with the main user interaction elements. First, users can load a model. SynBioSS DS uses libSBML (Bornstein, et al., 2008) to read models specified in SBML and can also read models created directly for Hy3S as NetCDF files. Models can then be edited or, alternatively, created within the model interface and saved. The user can subsequently specify simulation parameters. Typically, this is merely the amount of simulation time and the number of stochastic trajectories to be sampled (stochastic simulation requires multiple samples to construct population distributions). The simulator includes default simulation parameters for stable simulation of *E. coli*. Simulations can either be run locally or exported to a NetCDF file appropriate for the supercomputer simulator. If conducted locally, the results can be exported as either an ASCII comma-separated-value (CSV) file for import into any spreadsheet program, or as a NetCDF file appropriate for MatLab.

4.2 Supercomputer Simulator

The supercomputer simulator uses MPI to coordinate parallel execution of multiple trajectories on multiple processors. Such parallel execution is an example of a trivially parallelizable problem, where perfect scaling is achieved and the number of processors utilized is limited only by the number of trajectories simulated. Parallel execution may be necessary for algorithms to determine the optimum set of parts for a synthetic network to match a targeted biological phenotype (Tomshine and Kaznessis, 2006). Once completed, the simulation results are stored in a NetCDF file and can be loaded into MatLab or converted to an ASCII CSV file.

5 SUMMARY

SynBioSS is a complete software suite for each step in the construction of a synthetic network model: creating a kinetic model, locating kinetic information, and simulating that model accurately and efficiently. The resulting probability distributions of dynamic biological phenotypes can provide insight useful for engineering synthetic biological networks.

ACKNOWLEDGEMENTS

Funding: Funding was provided by the US National Science Foundation Grant CBET-0644792 (CAREER award to YNK).

Conflict of Interest: none declared.

REFERENCES

- Bornstein, B.J., et al. (2008) LibSBML: an API Library for SBML, *Bioinformatics*, **24**, 880-881.
- Canton, B., et al. (2008) Refinement and standardization of synthetic biological parts and devices, *Nature Biotechnology*, **26**, 787-793.
- Gillespie, D.T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *Journal of Computational Physics*, **22**, 403-434.
- Hucka, M., et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models, *Bioinformatics*, **19**, 524-531.
- Kærn, M., et al. (2003) The engineering of gene regulatory networks, *Annual Review of Biomedical Engineering*, **5**, 179-206.
- Kaznessis, Y.N. (2007) Models for synthetic biology, *BMC Systems Biology*, **1**, 47.
- Leuf, B. and Cunningham, W. (2001) *The Wiki way: quick collaboration on the Web*. Addison-Wesley Longman Publishing Co., Inc.
- McAdams, H.H. and Arkin, A. (1998) Simulation of prokaryotic genetic circuits, *Annual Review of Biophysics and Biomolecular Structure*, **27**, 199-224.
- McAdams, H.H. and Arkin, A. (1999) It's a noisy business! Genetic regulation at the nanomolar scale, *Trends in Genetics*, **15**, 65-69.
- McQuarrie, D.A., et al. (1964) Kinetics of small systems. II, *The Journal of Chemical Physics*, **40**, 2914-2921.
- Rew, R. and Davis, G. (1990) Data management: NetCDF: an interface for scientific data access, *IEEE Computer Graphics and Applications*, **10**, 76-82.
- Salis, H., et al. (2006) Multiscale Hy3S: hybrid stochastic simulation for supercomputers, *BMC Bioinformatics*, **7**, 93.
- Samad, H.E., et al. (2005) Stochastic modeling of gene regulatory networks, *International Journal of Robust and Nonlinear Control*, **15**, 691-711.
- Sotiropoulos, V. and Kaznessis, Y.N. (2007) Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes, *BMC Systems Biology*, **1**, 7.
- Taylor, S.R., et al. (2008) Sensitivity measures for oscillating systems: application to mammalian circadian gene network, *Automatic Control, IEEE Transactions on*, **53**, 177-188.
- Tomshine, J.R. and Kaznessis, Y.N. (2006) Optimization of a stochastically simulated gene network model via simulated annealing, *Biophysical Journal*, **91**, 3196-3205.
- Tuttle, L.M., et al. (2005) Model-driven designs of an oscillating gene network, *Biophysical Journal*, **89**, 3873-3883.

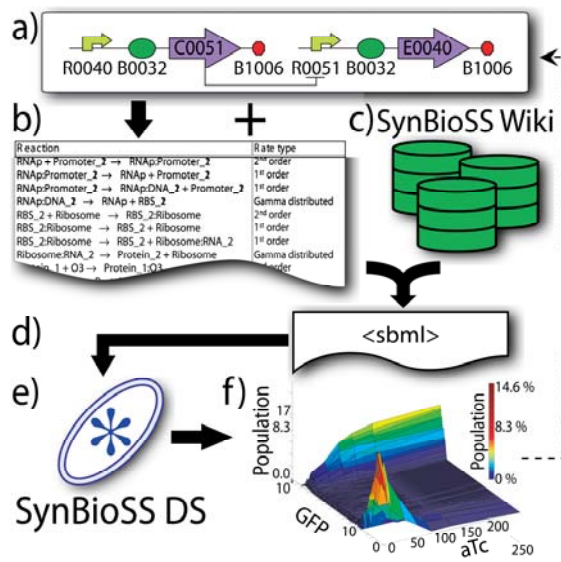


Figure 1. Overview of the SynBioSS workflow. The user starts at a) a synthetic gene construct (BioBrick nomenclature). Using the SynBioSS Designer creates b) a kinetic model of the reaction network. Adding c) the SynBioSS Wiki, d) a kinetic model in SBML is generated. This model is then simulated with e) the SynBioSS simulator, resulting in f) a population distribution of each species over time. The process can then be iterated until the desired phenotype is achieved.