Journal of Biotechnology 261 (2017) 215–220

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

COPASI and its applications in biotechnology☆

Frank T. Bergmann[a,*], Stefan Hoops[b], Brian Klahn[b], Ursula Kummer[a,**], Pedro Mendes[c], Jürgen Pahle[d], Sven Sahle[a]

[a] BioQuant/COS, Heidelberg University, Heidelberg, Germany
[b] Biocomplexity Institute of Virginia Tech, Blacksburg, VA, USA
[c] Center for Quantitative Medicine, UConn Health, Farmington, CT, USA
[d] BIOMS/BioQuant, Heidelberg University, Heidelberg, Germany

ARTICLE INFO

Keywords:
COPASI
Biotechnology
Systems biology

ABSTRACT

COPASI is software used for the creation, modification, simulation and computational analysis of kinetic models in various fields. It is open-source, available for all major platforms and provides a user-friendly graphical user interface, but is also controllable via the command line and scripting languages. These are likely reasons for its wide acceptance. We begin this review with a short introduction describing the general approaches and techniques used in computational modeling in the biosciences. Next we introduce the COPASI package, and its capabilities, before looking at typical applications of COPASI in biotechnology.

1. Introduction

Computational modeling has been used in the context of biology for decades. Early studies on enzyme kinetics have already employed small computational models and their simulation (e.g. Chance, 1949). Also, in the context of population ecology and epidemiology, computational modeling has been involved for many years to describe and predict population sizes of species or to analyze the spread of diseases. Cell biological questions, e.g. the investigation of the mechanisms behind calcium oscillations (for a review see e.g. Schuster et al., 2002) or the understanding of the regulation of glycolysis (Sel'Kov, 1968) are another context in which models have played a role since the 1960's. However, despite the successes and the knowledge about the usefulness of the respective techniques, usage only started to virtually explode with the turn of the century, when systems biology developed as its own discipline (Hübner et al., 2011). This is probably due to the advances both on the experimental side (e.g. live-cell imaging, high-throughput data), as well as on the computational side (widespread availability of computers, faster machines, more and tailored algorithms). Another likely contributor to this virtual explosion in computational modeling, has been the emergence of user-friendly software, which makes model set-up and analysis much more accessible, and practical, for non-mathematicians.

Different computational methods are employed on different scales of abstraction and detail. In the areas of biochemistry and cell biology the two largest classes (Hübner et al., 2011) of models are: (a) stoichiometric models including constraint based genome scale models, and (b) kinetic models. The former are based on static reaction schemes and used to calculate potential flux distributions mostly in metabolic networks. The latter allows for a more detailed picture and for more fine-grained analysis, but also demands more knowledge, namely about the kinetics of individual processes.

For these two classes of computational models the German Network for Bioinformatics Infrastructure, de.NBI, hosts two widely used resources. Stoichiometric models, including genome scale constraint based models, can be set-up and analyzed using CellNetAnalyzer (Klamt et al., 2007), whereas COPASI (Hoops et al., 2006; Mendes et al., 2009) specializes in setting up and analyzing kinetic models while also providing some basic stoichiometric analyses.

Computational models are widely employed in biotechnology. The following sections will go into some detail about this. Generally speaking, often the goal is to manipulate systems to either maximize the yield, flux, or titer of desirable end-products, or to minimize the amounts or fluxes of undesirable side-products.

Below, we will first give an overview of the current features of the COPASI software. Then we describe its application in the context of biotechnology by presenting concrete examples where COPASI was applied to systems of biotechnological interest with varying levels of complexity.

☆ Authors are in alphabetical order.

* Corresponding author.
** Co-corresponding author.
E-mail address: frank.bergmann@bioquant.uni-heidelberg.de (F.T. Bergmann).

http://dx.doi.org/10.1016/j.jbiotec.2017.06.1200
Received 14 February 2017; Received in revised form 19 June 2017; Accepted 22 June 2017
Available online 24 June 2017

2. The COPASI software

COPASI (Hoops et al., 2006; Mendes et al., 2009, 2009b) is a modeling and simulation environment, jointly founded by the groups of Pedro Mendes (now at UConn Health) and Ursula Kummer (now at Heidelberg University). Currently the groups of Stefan Hoops, Jürgen Pahle, and Sven Sahle are also part of the consortium. All 3 have been former group members and are now PIs on their own, and are actively participating in development of the software. The software is platform-independent, developed as open source, and distributed available under the terms of the Artistic License 2.0,[1] which makes it free to use for any purpose. Source and binary distributions are available from the COPASI Website (http://copasi.org) for Windows, Linux, and Mac OS X. This makes it very easy for researchers to install COPASI and get started quickly. Those who prefer it can also download the source code and compile their own specific versions of the software.

2.1. Features

Model editing. Models are entered in (bio)chemical terms. (Bio) chemical species, located in compartments, can be transformed through reactions. The rate of each reaction is given by a kinetic rate law, that can be either selected from an included function database (filtered to apply to the selected reaction), or freely defined. Additionally, discrete events can be added to the model, that change model elements based on arbitrary expressions. From this formalism, COPASI derives the mathematical representation of the model automatically. Thus users do not have to specify the differential equations themselves. However, expert users can add additional differential equations to the model, including for compartment volumes or any other variable (i.e. not just chemical species).

Model standards. COPASI stores the model and information what to do with it in its own specialized XML format. However, wherever possible, applicable community standards are supported. Models can be imported and exported in the Systems Biology Markup Language (SBML, Hucka et al., 2003) format, where all major versions are supported. This enables COPASI to consume models from a variety of online model repositories (such as the BioModels Database Le Novère et al., 2006), as well as from many other software tools that support the SBML standard.

Where possible, the analysis tasks are exported to the Simulation Experiment Description Markup Language (SED-ML, Waltemath et al., 2011). Specifically, time course simulations and parameter scans can be exported to SED-ML, as SED-ML does not specify further simulation types yet.

Currently the COPASI team is extending COPASI to additionally support the COMBINE Archive (Bergmann et al., 2014) which will make it easier to exchange COPASI models together with experimental data.

Simulation algorithms. At the core, COPASI supports two main formalisms for simulating the dynamics of the defined model (time course simulations). One is the traditional chemical kinetics approach of using ordinary differential equations (ODE), where the software uses the LSODAR integrator (Petzold, 1983). The second is the stochastic formalism (Pahle, 2009), where individual reaction events are drawn from probability distributions, using either of the following algorithms: Gillespie's Direct Method (Gillespie, 1976), Gibson–Bruck (Gibson and Bruck, 2000), τ-Leap, or adaptive SSA/τ-leap (Cao et al., 2006). To aid users in adapting deterministic rate equations into their stochastic equivalent, COPASI features an option that splits up reversible reactions into their forward and backward-directional forms. Additionally there is an option to apply stochastic corrections to rate equations.

Since stochastic simulations can be rather time consuming, a third class of simulation algorithms available consist of hybrid solvers. The solvers “Hybrid (Runge–Kutta)” and “Hybrid (LSODA)” employ a dynamic partitioning strategy that treats all reactions involving species that have a particle number below a specified threshold with the stochastic approach, while the remaining reactions are treated with the differential equation approach. A new solver was recently introduced, “Hybrid (RK-45)”, which allows modelers to define this partitioning themselves, and thus control which reactions are to be simulated stochastically or deterministically.

In the future, solvers for Stochastic Differential Equations (SDEs) and Delay Differential Equations (DDEs) will also be added.

Analysis tasks. The model evaluations through time course simulations or steady state computations at the core of COPASI support a number of additional analysis tasks.

- *Parameter scan:* The parameter scan task can be used to run any of the other tasks repeatedly. Frequently one needs to examine how the model behaves when some parameter(s) vary, and this task allows programming this parameter variation in an easy way. Parameter values can be either “scanned” in a given interval (i.e. their values changed through regular intervals), or sampled from a random distribution; these methods can be intermixed such that some parameters can be scanned and others sampled. This can be applied to time course simulations or steady state simulations, but also to *any* other task in this list. This allows for complex operations, such as multi-dimensional parameter scans, repeating stochastic time course simulations to determine the time-dependent probability distributions for each model variable, creating bifurcation plots (running the cross section task for different values of a parameter), evaluating the fitness landscape around an optimization solution, or even determining profile likelihood plots (Schaber, 2012).
- *Optimization:* In the optimization task users can enter an arbitrary expression of model quantities (or even results from the execution of other tasks) that should be maximized or minimized. Additionally the model elements that should be modified, where appropriate with optional constraints, are specified. COPASI implements a wide selection of local (Hooke & Jeeves: Hooke and Jeeves, 1961, Levenberg–Marquardt: Levenberg, 1944, Nelder–Mead: Nelder and Mead, 1965, Praxis: Brent, 1973, Steepest Descent, Truncated Newton: Nash, 1984) and global optimization methods (Differential Evolution: Storn and Price, 1997, Evolutionary Strategy: Runarsson and Yao, 2000, Evolutionary Programming: Fogel et al., 1992, Genetic Algorithm: Michalewicz, 1994, Particle Swarm: Kennedy and Eberhart, 1995, Scatter Search: Egea et al., 2009, Random Search, Simulated Annealing: Corana et al., 1987), that can be used for the optimization.
- *Parameter estimation:* Parameter estimation is the activity of changing the model's parameter values in order to make the behavior of the model as close as possible to a set of experimental measurements. Parameter estimation is a special form of optimization, in which the distance between the model and the experimental data is minimized. This is achieved by changing user-selected model parameters through specific optimization algorithms (the same ones that are available for general optimization described above). COPASI supports both experimental data in the form of time courses and steady state measurements; several data sets of either type can be mixed together. Users can specify as many data sets as required. A new feature allows validation data sets to be specified as well. These experimental data sets are not used during the fitting process, but only during a validation phase to measure how robust the model predictions are. The middle window in Fig. 1 shows the definition of parameters to be modified in a parameter estimation run.
- *Metabolic control analysis:* Metabolic control analysis (MCA) is a special type of sensitivity analysis that calculates how much a perturbation in the rate of a reaction affects the steady state concentrations or fluxes (Fell, 1996). These are systems level coefficients and are called “control coefficients”. MCA establishes a link

[1] https://opensource.org/licenses/Artistic-2.0.