Systems Biology Markup Language for R

SBMLR Version 1.26 (10/17/2005)

Tomas Radivoyevitch, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio 44106; radivot@hal.cwru.edu; http://epbi-radivot.cwru.edu/

Background

First order nonlinear ordinary differential equations (ODEs) (Hirsch and Smale, 1974) are commonly used to model biochemical systems. The models have boundary condition nodes which decouple them from the universe, and state variables nodes which respond dynamically to changes in system inputs. Starting in a steady state, model perturbations by changes in metabolite-, drug- or enzyme concentrations create time course responses which may be used to explain or predict data. For example, consider the purine metabolism model (Curto *et al.*, 1997; Curto *et al.*, 1998a; Curto *et al.*, 1998b) shown in Figure 1, where the state variables are boxes, reactions are arrows, and reaction inhibitors and activators are not shown. This model's response to a bolus of phosphoribosyl-pyrophosphate (PRPP) at $t = 0$, simulated by a sudden 10-fold increase in the initial PRPP concentration to a new initial (non-steady) state, is shown in Figure 5.

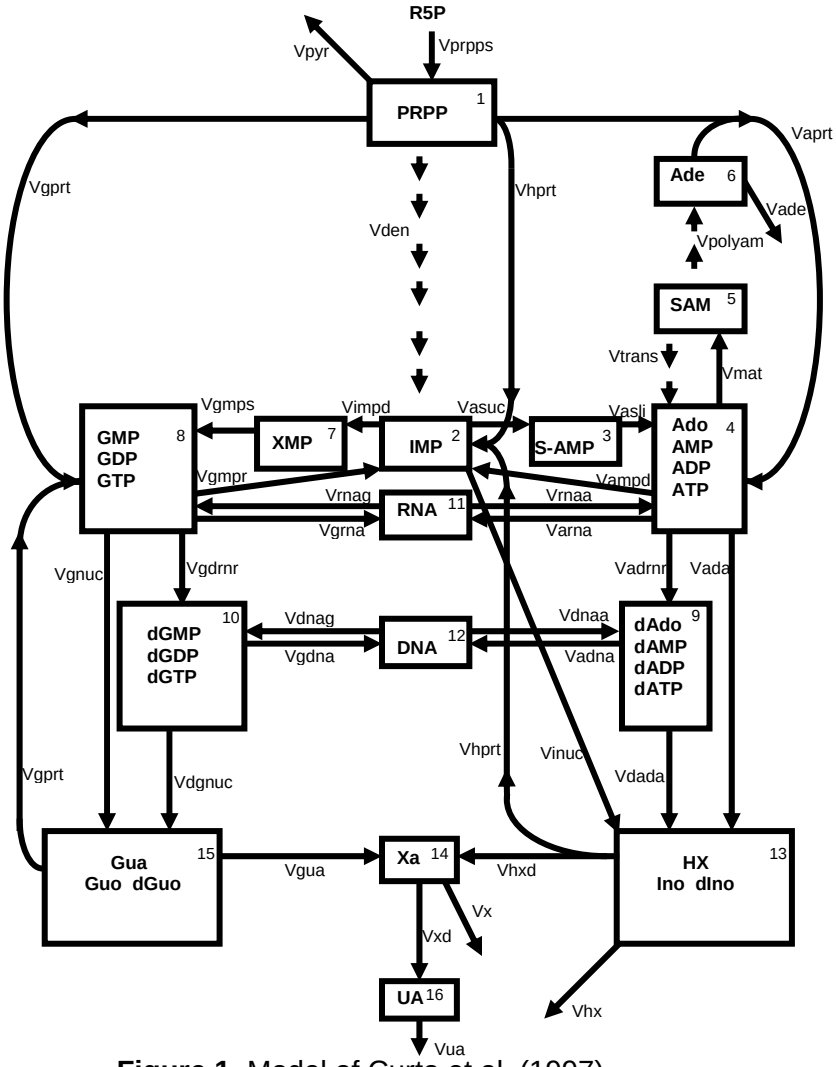

Figure 1. Model of Curto et al. (1997).

The rate at which a state changes equals the sum of the fluxes into its node minus the sum of the fluxes leaving its node; fluxes are determined by current state values, boundary conditions and reaction rate parameters. The coupled nonlinear set of ODEs, the initial conditions, the boundary conditions (a.k.a. forcing functions or independent variables), the parameter values, and any additional state space constraints, together comprise a model which SBML can represent using the list-of-lists file structure shown in Figure 2. Once a particular model is represented in SBML, many different software packages (http://sbml.org/) can then be used to analyze it. Further, it can then be stored and accessed from a repository such as the BioModels Database, see http://www.ebi.ac.uk/biomodels/.

beginning of model definition
list of function definitions (optional)
list of unit definitions (optional)
list of compartments (optional)
list of species (optional)
list of parameters (optional)
list of rules (optional)
list of reactions (optional)
list of events (optional)
end of model definition

Figure 2. SBML file Structure

SBMLR

SBMLR (Radivoyevitch, 2004) includes an SBML-like R model object of class `SBML` and several functions which help users share, edit, test, view and analyze their models, as shown in Fig. 3. Here, `saveSBML` exports SBML model objects to SBML level 2 model files; `readSBML` imports a limited range of SBML level 2 model files back to SBMLR; `saveSBMLR` and `readSBMLR` exchange SBML objects with SBMLR model definition files used for model editing; `simulate` uses `lsoda` of R's `odesolve` package to simulate `SBML` models; `summary` summarizes models, including species and reaction dataframes which contain initial concentrations, initial fluxes, and reaction rate laws (as strings); and == tests the equality of models with respect to these two dataframes.

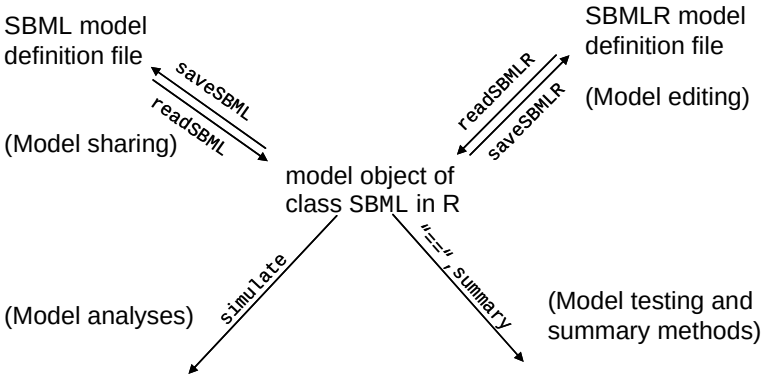

Figure 3. SBMLR objects and functions

SBMLR includes the purine metabolism model of Curto et al. (1997) shown in Figure 1. This model represented as an SBMLR file, is partly shown in Figure 4. An R script which simulates this model's response to a 10-fold phosphoribosylpyrophosphate (PRPP) increase is shown in Figure 5; since PRPP reacts with hypoxanthine (HX) to form inosine monophosphate (IMP), the predicted dynamic responses are reasonable. A more advanced SBMLR use example involving microarray data is described in (Radivoyevitch, 2005). R scripts used in that study are provided in the `BMCcancerFolates` directory of SBMLR (they can also be obtained from http://epbi-radivot.cwru.edu/folates/).

```
model=list(
notes=c("Purine metabolism model.")
compartments=list(list(id= "cell",size=1)),
species=list(
PRPP   =list(id="PRPP", ic=5,compartment="cell", bc=F),
            *** other chemical species ***
 ),
reactions=list(
list(id="gprt", rever=FALSE,
reactants=c("Gua","PRPP"),
modifierss=c("GTP"),
products  =c("GTP"),
parameters=c(agprt =361.69,fgprt1 =1.2,fgprt8 =-1.2,fgprt15 =0.42),
strLaw    = "agprt*PRPP^fgprt1*GTP^fgprt8*Gua^fgprt15"
),
    *** other chemical reactions ***
)
)
```

SBMLR definition of Curto model **Figure 4**

```
library(SBMLR)
library(odesolve)
curto=readSBML("curto.xml")
out1=simulate(curto,seq(-20,0,1))
curto$species$PRPP$ic=50
out2=simulate(curto,0:70)
outs=data.frame(rbind(out1,out2))
attach(outs)
par(mfrow=c(2,1))
plot(time,IMP,type="l",xlab="minutes",ylab="IMP (uM)")
plot(time,HX,type="l",xlab="minutes",ylab="HX (uM)")
par(mfrow=c(1,1))
detach(outs)
```

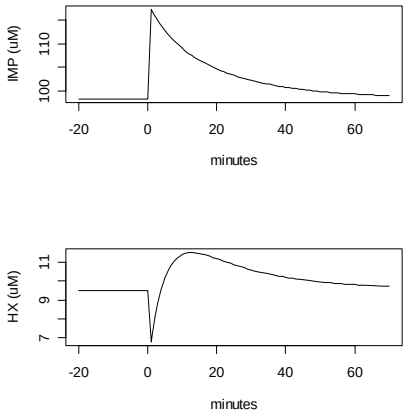

Figure 5. R script used to create the response shown on the right.

Future Work

1. Add events and function definitions.
2. Provide graphic renderings of reaction rate equations in R.
3. Extend SBML to include unigene ID(s) for each reaction. If there is more than one gene involved in the production of a single functional enzyme [i.e. if the enzyme is a complex], a prescription of how gene expression values should be mapped into a modulator of V_{max} for the corresponding reaction rate law should be included in the SBML code.
4. It is anticipated that SBML level 3 will include graphical configuration information. When this happens, the SBMLR structure will be modified to include/carry such information.

1. Curto, R., Voit, E. O. and Cascante, M. Analysis of abnormalities in purine metabolism leading to gout and to neurological dysfunctions in man. *Biochem.J.* **329 (Pt 3)**, 477-487 (1998a).
2. Curto, R., Voit, E. O., Sorribas, A. and Cascante, M. Validation and steady-state analysis of a power-law model of purine metabolism in man. *Biochem.J.* **324 (Pt 3)**, 761-775 (1997).
3. Curto, R., Voit, E. O., Sorribas, A. and Cascante, M. Mathematical models of purine metabolism in man. *Math.Biosci.* **151**, 1-49 (1998b).
4. Hirsch, M. W. and Smale, S. *Differential equations, dynamical systems, and linear algebra* (Academic Press, New York, 1974).
5. Radivoyevitch, T. A two-way interface between limited Systems Biology Markup Language and R. *BMC Bioinformatics* **5**, 190 (2004).
6. Radivoyevitch, T. Folate system correlations in DNA microarray data. *BMC Cancer* **5**, 95 (2005).