

Introduction to rule-based modeling with BioNetGen and RuleBender

**Jim Faeder, Cihan Kaya, Sanjana Gupta,
and Kunal Aggarwal**

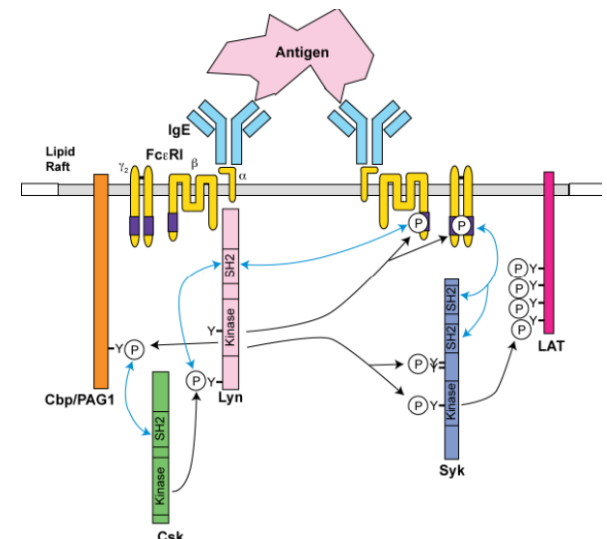
*Department of Computational and Systems Biology
University of Pittsburgh School of Medicine*



Cell Modeling Workshop
Pittsburgh Supercomputing Center
June 26-28, 2017



For additional information and references see
<http://bionetgen.org/index.php/Tutorials>

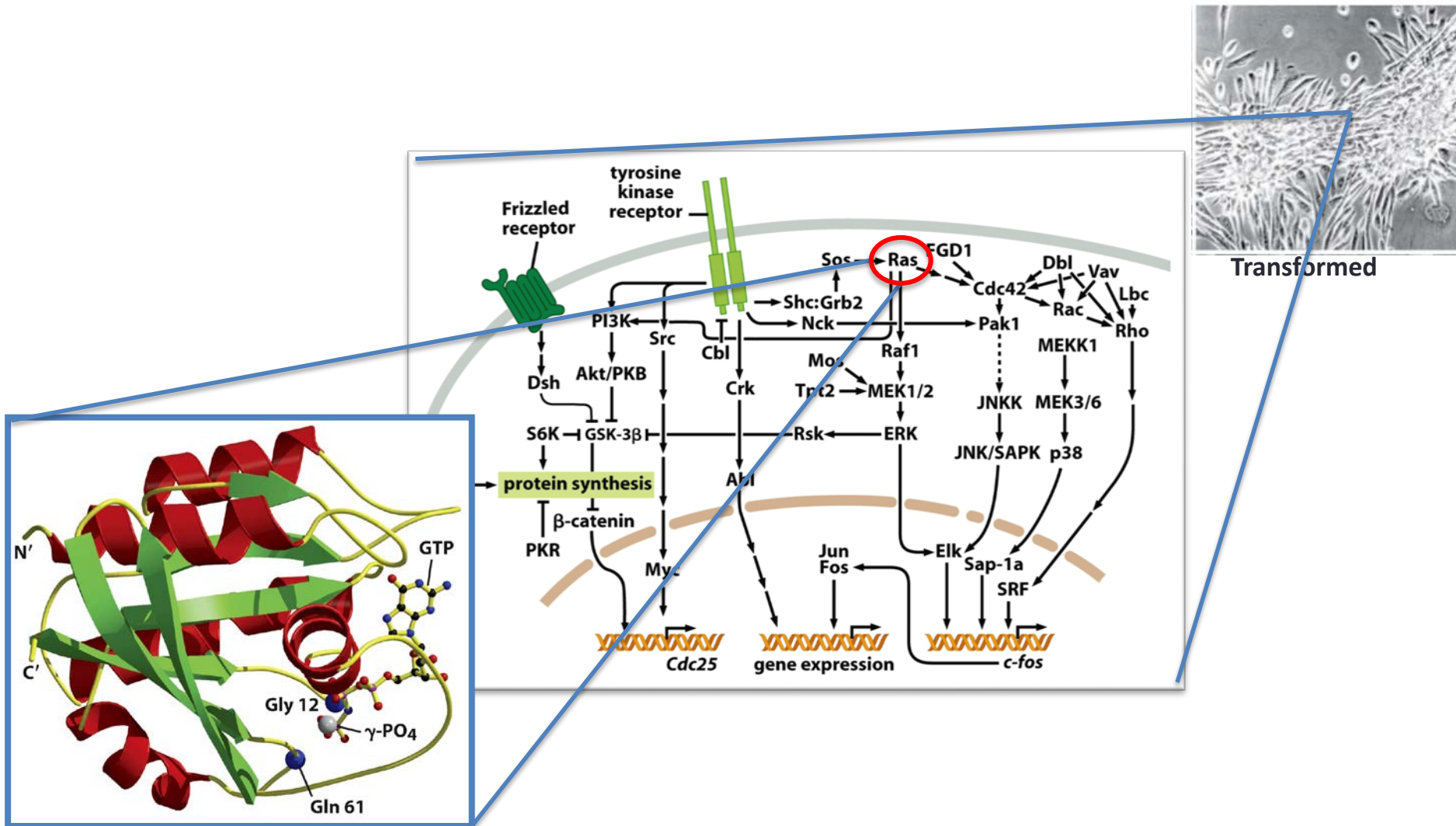


Outline

1. Background and motivation (Jim)
2. Rule-based modeling tools and workflow (Jim)
3. Some example models (Jim)
4. CaMKII example (Cihan)
5. Model visualization (Sanjana)
6. Calibrating models to data (Sanjana)
7. Using Atomizer for model import (Kunal)

The Biology of Cancer (© Garland Science 2007)

Signal flow involves interactions and dynamics at multiple scales



abstraction level

The diagram illustrates the signaling pathways initiated by the tyrosine kinase receptor and the Frizzled receptor. The tyrosine kinase receptor (green) activates pathways involving Src, Cbl, Crk, ABL, and Rsk, leading to gene expression. The Frizzled receptor (green) activates pathways involving Dsh, Akt/PKB, S6K, GSK-3β, and β-catenin, leading to protein synthesis. Both pathways converge on the ERK1/2 cascade, which leads to the activation of transcription factors like Jun, Fos, Elk, and SRF, ultimately resulting in gene expression and c-fos production.

The diagram illustrates the Ras signaling pathway. At the top, EGF binds to EGFR, which activates Ras (GTP-bound). Ras then activates Raf, which activates MEK, which activates ERK. The diagram shows the activation of each component by the previous one, with GTP/GDP states and phosphorylation indicated.

```

graph TD
    EGF[EGF] --> EGFR[EGFR]
    EGFR --> Ras[Ras]
    Ras --> Raf[Raf]
    Raf --> MEK[MEK]
    MEK --> ERK[ERK]
    ERK --> ERK2[ERK2]
    ERK2 --> ERK3[ERK3]
  
```

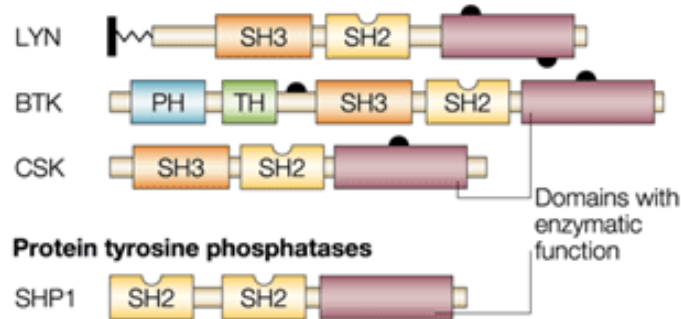
Key components and states shown in the diagram:

- Ras:** GTP-bound (active) and GDP-bound (inactive) states.
- Raf:** Activated by Ras (GTP-bound).
- MEK:** Activated by Raf (GTP-bound).
- ERK:** Activated by MEK (GTP-bound).
- ERK2 and ERK3:** Activated by ERK (GTP-bound).



Challenges of modeling cell regulatory networks

- Proteins are multi-functional

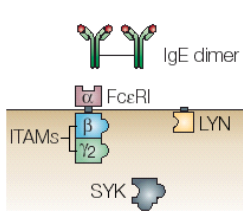


multiple sites of binding

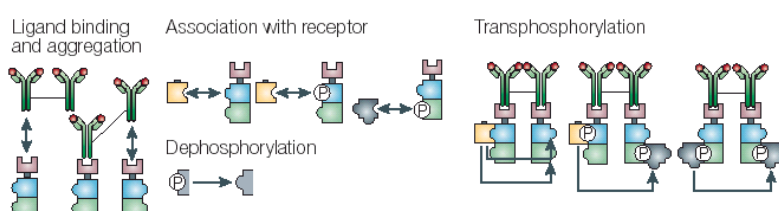
multiple sites of posttranslational modification

- Representing their known interactions requires handling of *combinatorial complexity*

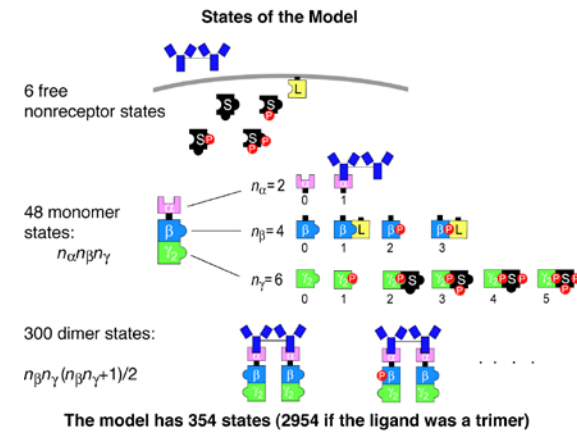
a Components



b Interactions



Small number of rules



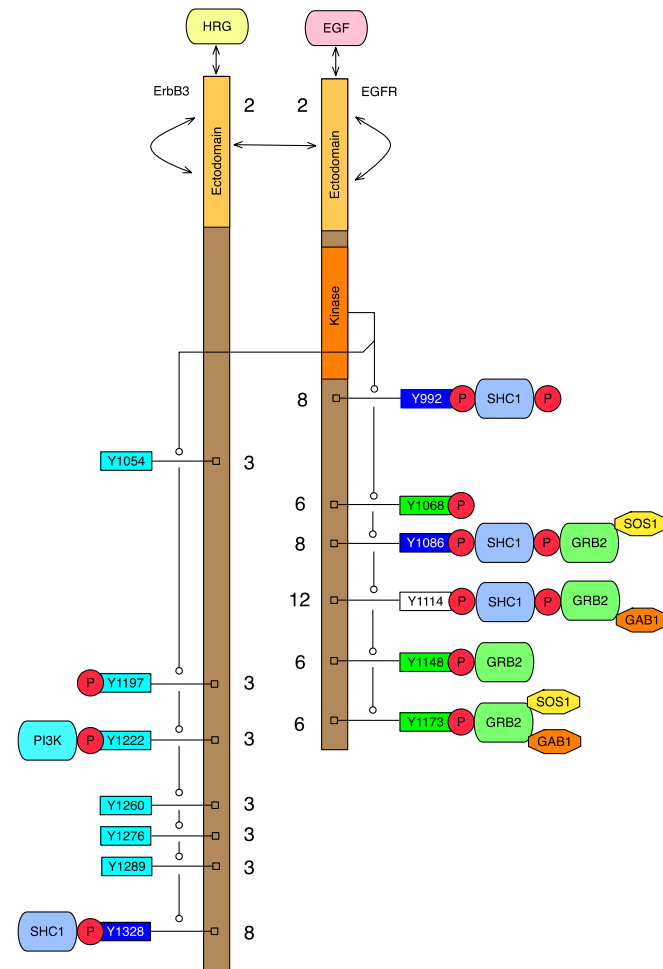
Small number of components and interactions → huge number of possible species and reactions

Combinatorial complexity in a realistic model of EGFR signaling

ErbB3:ErbB1 has $> 3.8 \times 10^9$ states

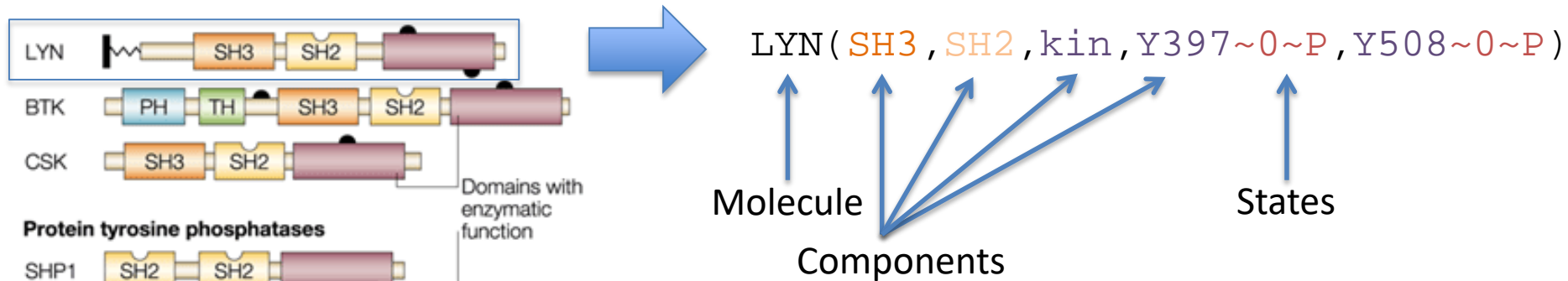
ErbB1:ErbB1 has $> 5.5 \times 10^{10}$ states

Such models can be built from a manageable number of rules – 10's to 100's

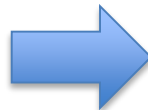
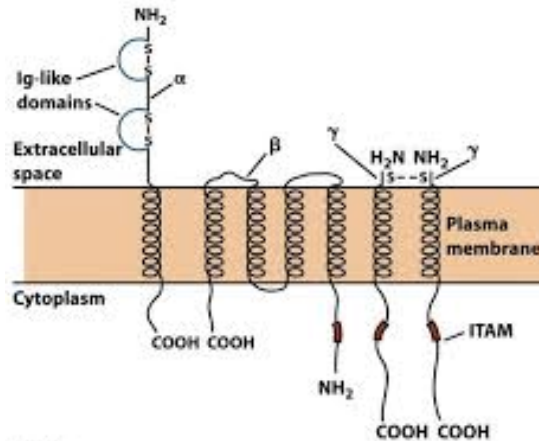


What is Rule-based Modeling (RBM)?

Molecules are modeled as *structured objects*



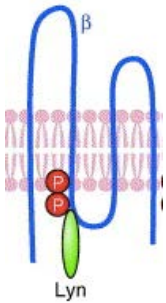
FcεRI: High-affinity IgE receptor



`FcεRI(a_Ig, b_Y218~0~P, g_ITAM~0~P)`

What is Rule-based Modeling (RBM)?

Rules define the interactions of molecules

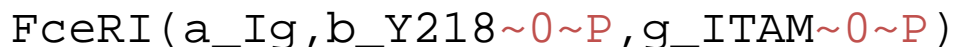
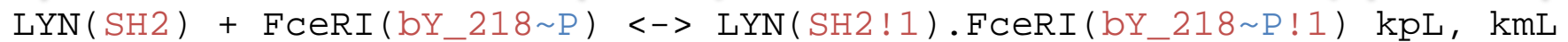


“Lyn SH2 domain binds to phosphorylated Tyr 218 on the β subunit of Fc ϵ RI ”

Reactants

Products

Rate Law



bond

“Don’t write don’t care” – elements not mentioned may be in any state

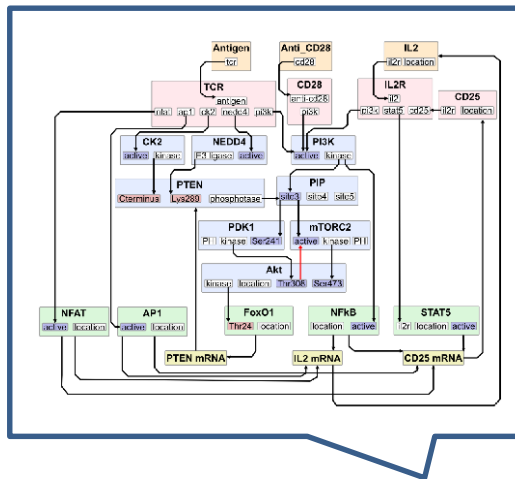
➔ *One rule can generate reactions involving many different species*

Reaction rate determined by **Mass Action kinetics**

$$\text{rate forward} = k_{\text{pL}} * [\text{LYN}(\text{SH2})] * [\text{Fc}\epsilon\text{RI}(\text{bY}_{218}\sim\text{P})]$$

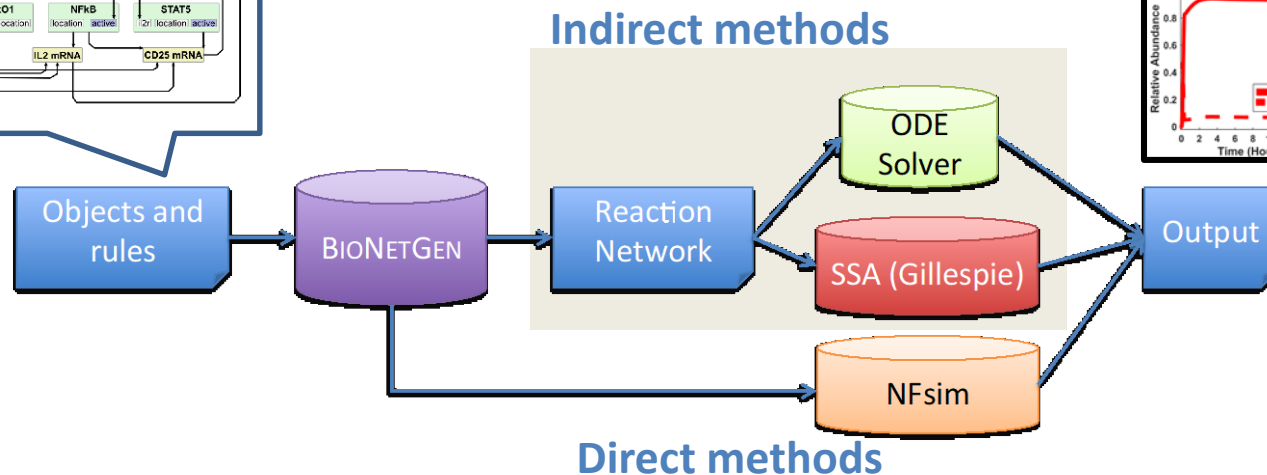
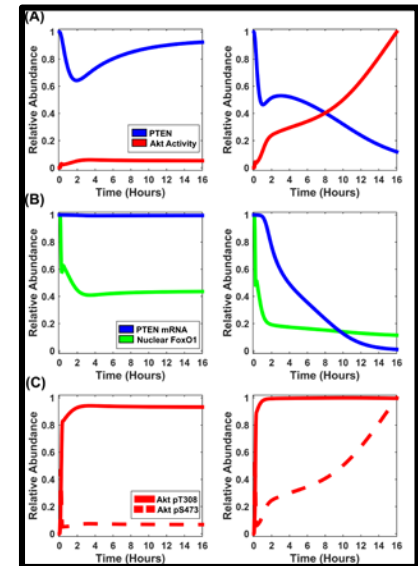
Basic RBM Workflow with BioNetGen

Rule-based models can be simulated using many different approaches



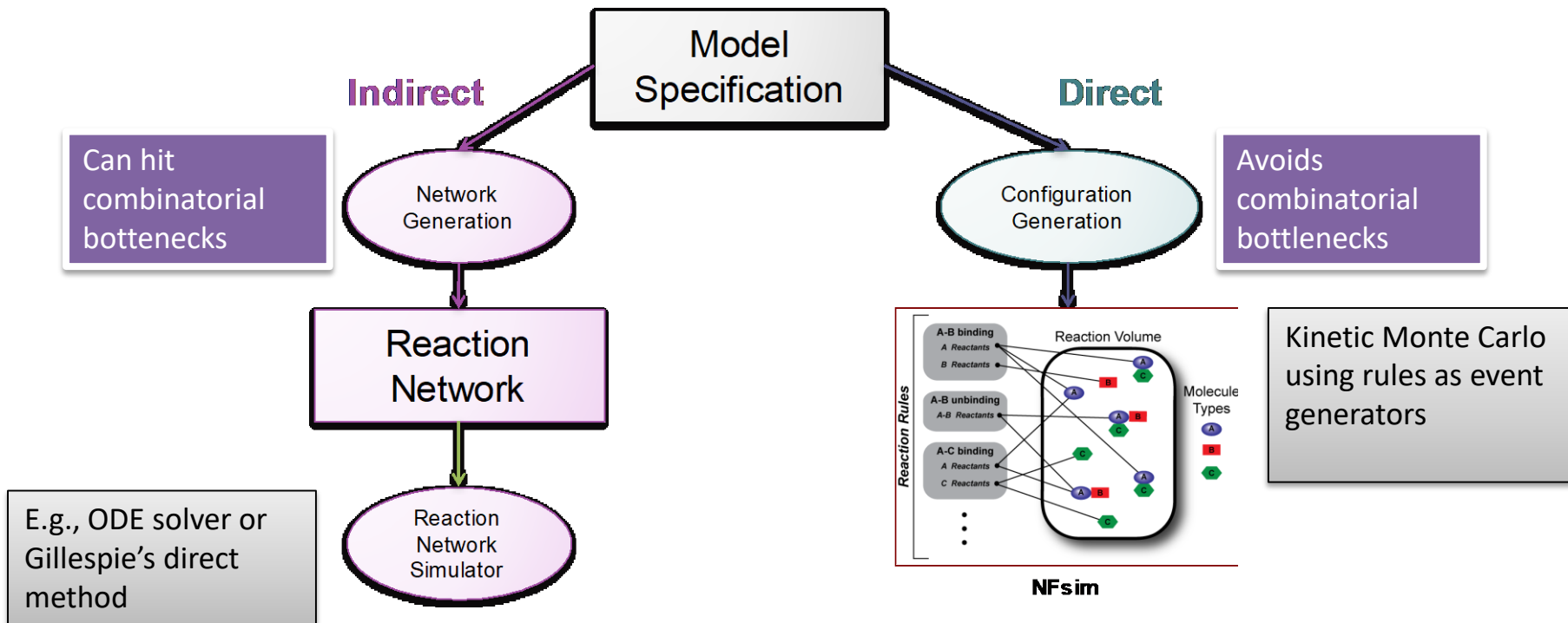
Molecules and interactions

Simulation output



“Network-based” vs. “network-free”

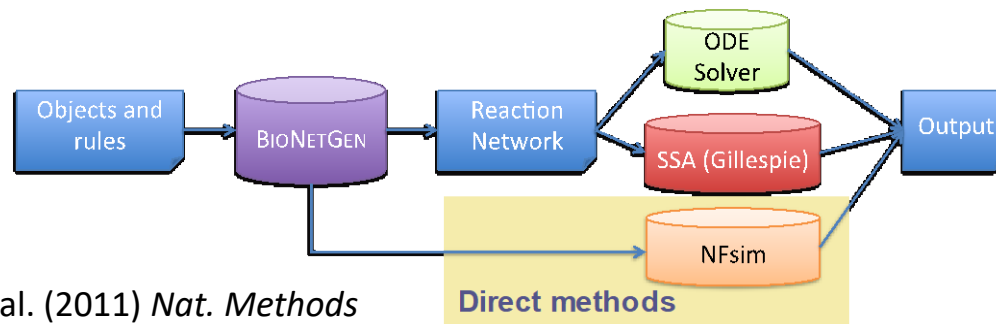
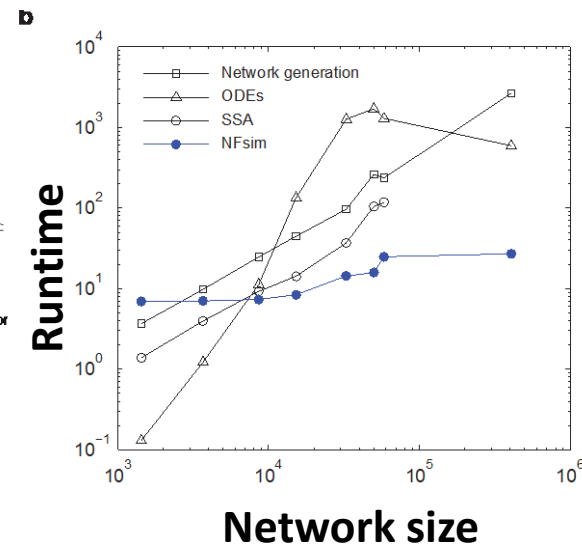
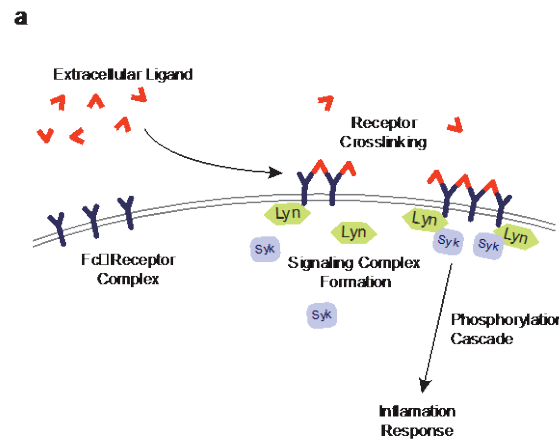
BioNetGen language file



Basic RBM Workflow with BioNetGen - NFsim

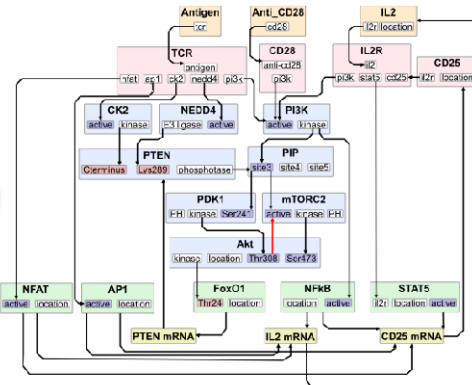
Network-free simulation enables scalable simulation as network size increases

*IgE receptor
signaling
network*

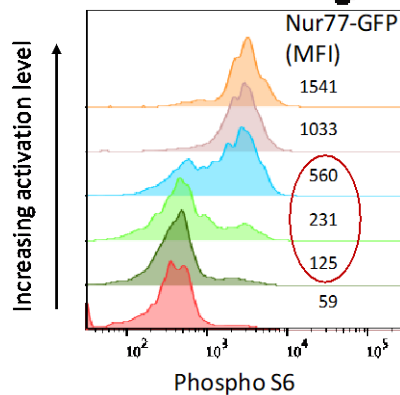


Closed-loop systems biology

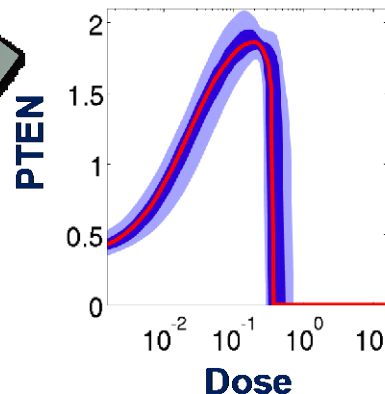
1. Construct model based on **known** and **hypothesized** interactions.



2. Calibrate model parameters to available data.

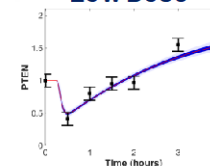


Repeat as needed

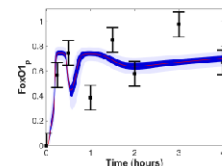
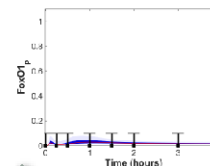
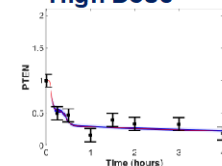


4. Perform experiments to test predictions.

Low Dose

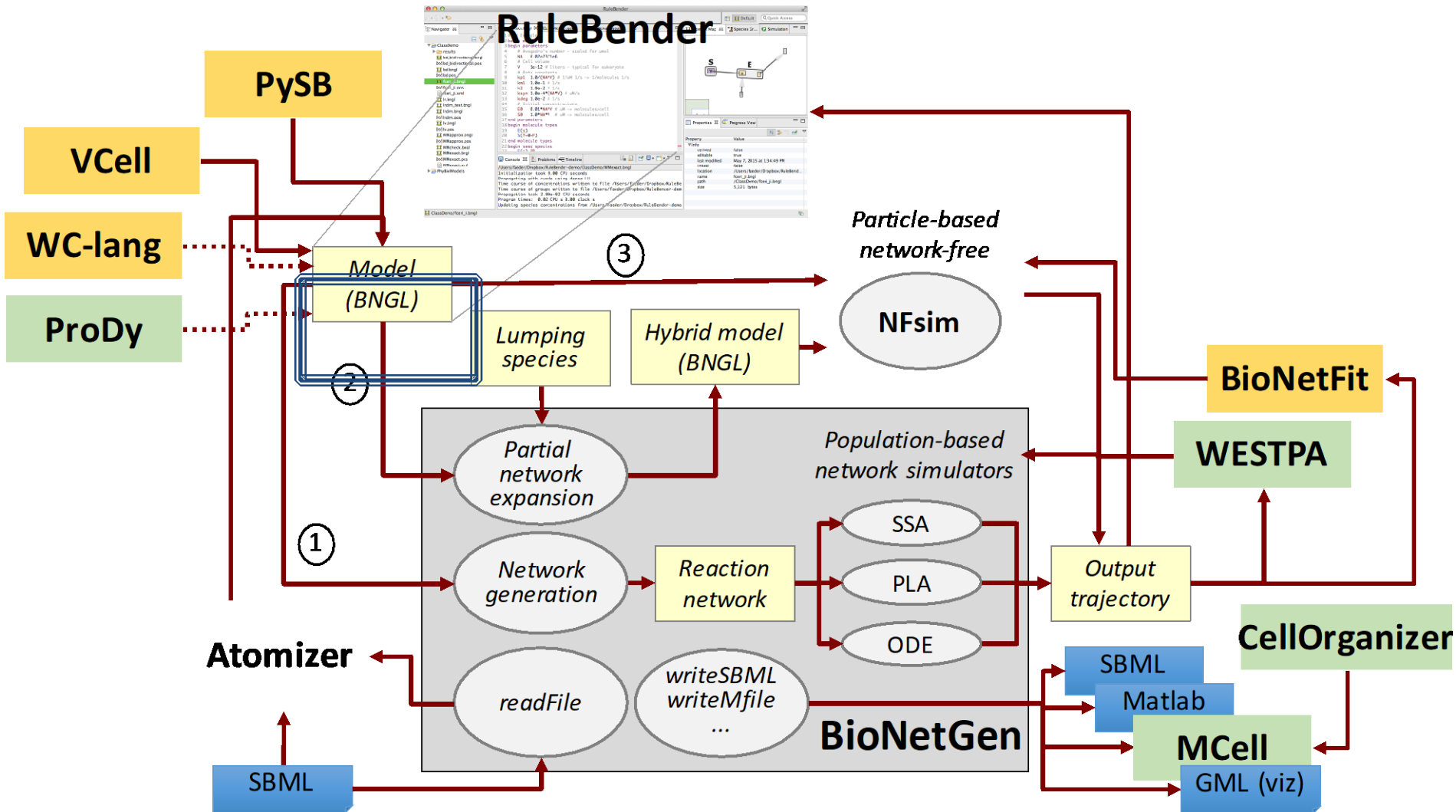


High Dose



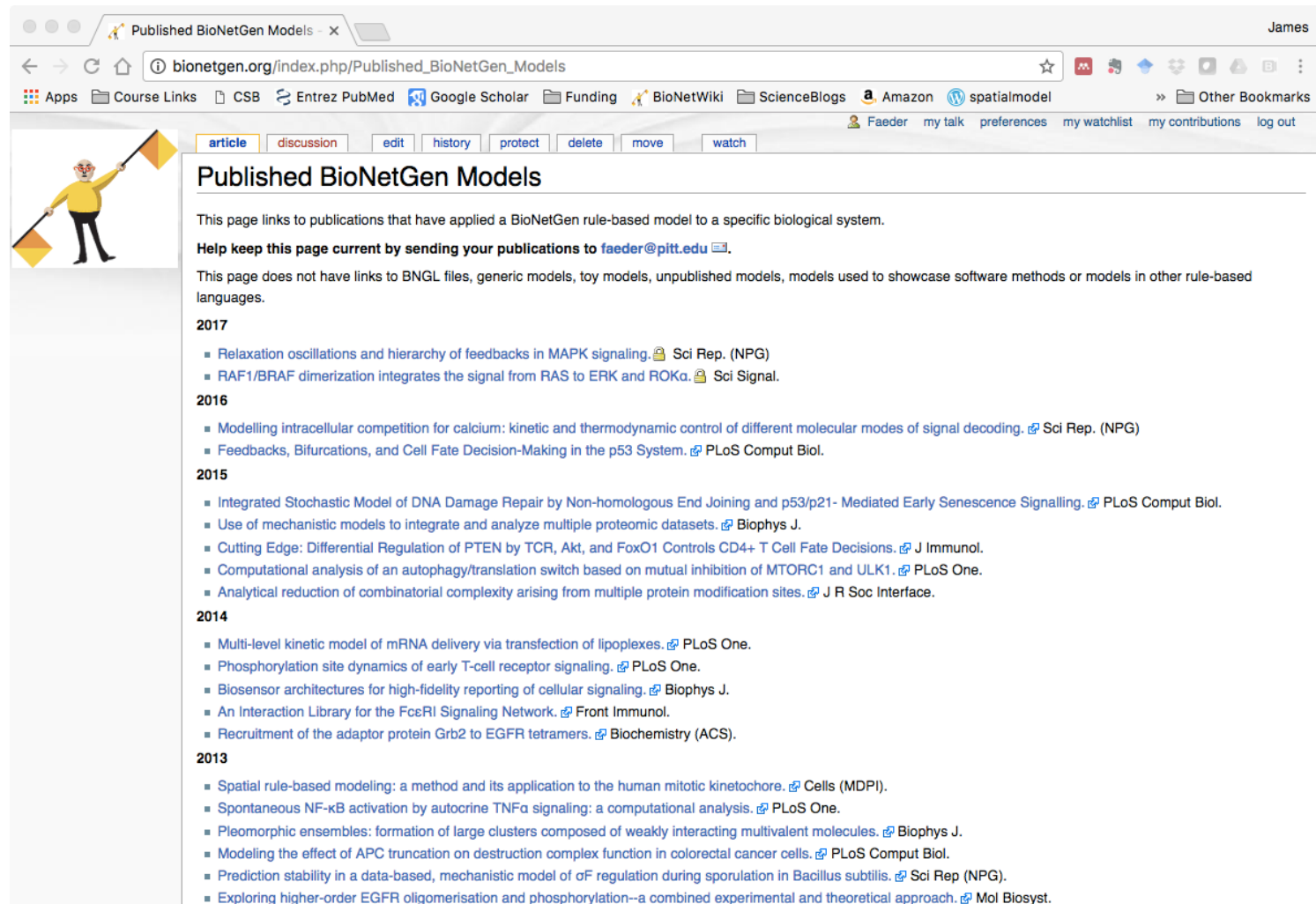
3. Analyze model to find novel behaviors.

BioNetGen and associated modeling tools



Growing number of applications

Models page at bionetgen.org currently lists > 40 papers



Published BioNetGen Models

This page links to publications that have applied a BioNetGen rule-based model to a specific biological system.

Help keep this page current by sending your publications to faeder@pitt.edu.

This page does not have links to BNGL files, generic models, toy models, unpublished models, models used to showcase software methods or models in other rule-based languages.

2017

- [Relaxation oscillations and hierarchy of feedbacks in MAPK signaling.](#) Sci Rep. (NPG)
- [RAF1/BRAF dimerization integrates the signal from RAS to ERK and ROKα.](#) Sci Signal.

2016

- [Modelling intracellular competition for calcium: kinetic and thermodynamic control of different molecular modes of signal decoding.](#) Sci Rep. (NPG)
- [Feedbacks, Bifurcations, and Cell Fate Decision-Making in the p53 System.](#) PLoS Comput Biol.

2015

- [Integrated Stochastic Model of DNA Damage Repair by Non-homologous End Joining and p53/p21- Mediated Early Senescence Signalling.](#) PLoS Comput Biol.
- [Use of mechanistic models to integrate and analyze multiple proteomic datasets.](#) Biophys J.
- [Cutting Edge: Differential Regulation of PTEN by TCR, Akt, and FoxO1 Controls CD4+ T Cell Fate Decisions.](#) J Immunol.
- [Computational analysis of an autophagy/translation switch based on mutual inhibition of MTORC1 and ULK1.](#) PLoS One.
- [Analytical reduction of combinatorial complexity arising from multiple protein modification sites.](#) J R Soc Interface.

2014

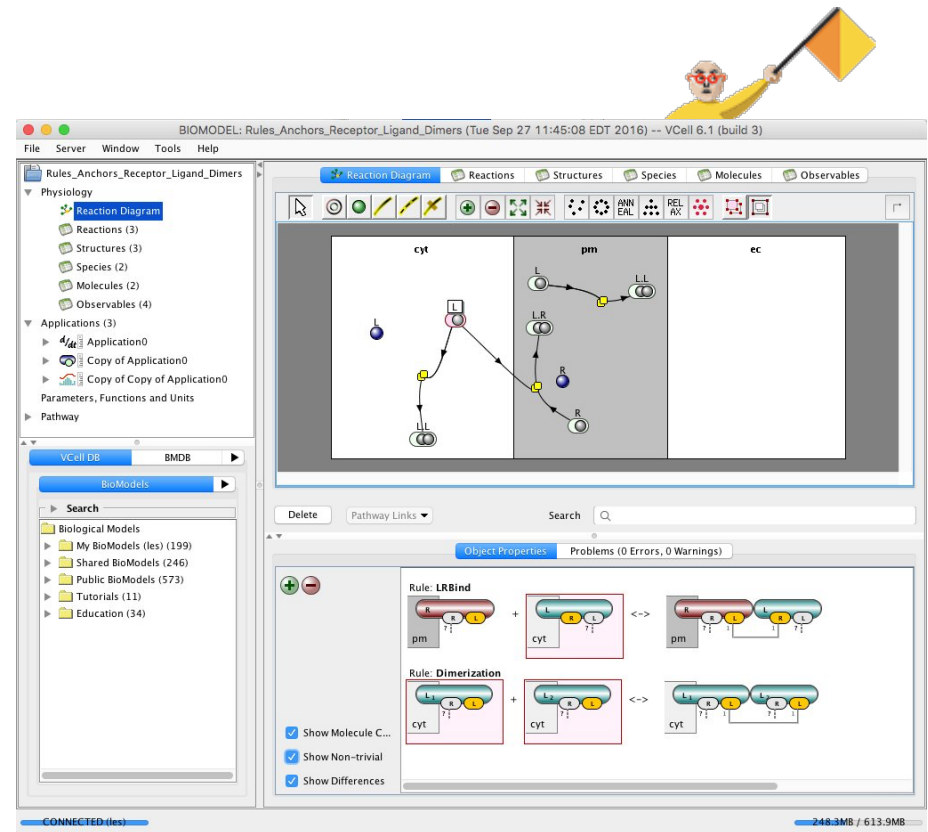
- [Multi-level kinetic model of mRNA delivery via transfection of lipoplexes.](#) PLoS One.
- [Phosphorylation site dynamics of early T-cell receptor signaling.](#) PLoS One.
- [Biosensor architectures for high-fidelity reporting of cellular signaling.](#) Biophys J.
- [An Interaction Library for the FcεRI Signaling Network.](#) Front Immunol.
- [Recruitment of the adaptor protein Grb2 to EGFR tetramers.](#) Biochemistry (ACS).

2013

- [Spatial rule-based modeling: a method and its application to the human mitotic kinetochore.](#) Cells (MDPI).
- [Spontaneous NF-κB activation by autocrine TNFα signaling: a computational analysis.](#) PLoS One.
- [Pleomorphic ensembles: formation of large clusters composed of weakly interacting multivalent molecules.](#) Biophys J.
- [Modeling the effect of APC truncation on destruction complex function in colorectal cancer cells.](#) PLoS Comput Biol.
- [Prediction stability in a data-based, mechanistic model of αF regulation during sporulation in Bacillus subtilis.](#) Sci Rep (NPG).
- [Exploring higher-order EGFR oligomerisation and phosphorylation—a combined experimental and theoretical approach.](#) Mol Biosyst.

Software tools that integrate BioNetGen/NFsim

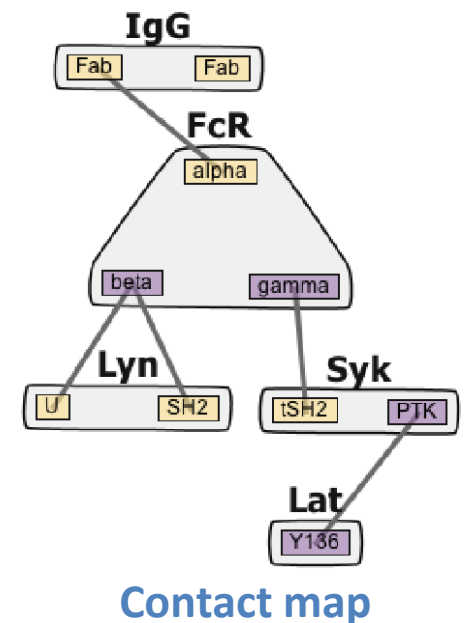
- PySB - INDRA
- BioNetFit
- SBML Multi
- Virtual Cell
- SRsim
- BioUML
- Parts&Pools
- ...



VCell interface highlighting rule-based modeling capability

Why use RBM?

- *Concise and precise* representation of biochemical knowledge
 - rules are simple (less context) when interactions are modular
- *Flexible* with respect to simulation method
 - Deterministic / stochastic
 - Well-mixed / compartmental / spatial
- Structures and rules are *reusable*
 - Rule libraries
- Compact visual representation
 - Contact map and beyond



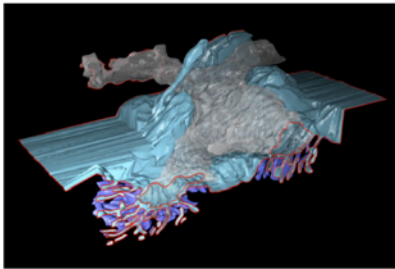
Development of network-free MCell

Beyond a well-mixed cell....

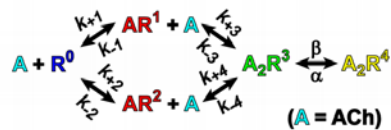
Jose Juan Tapia

MCell

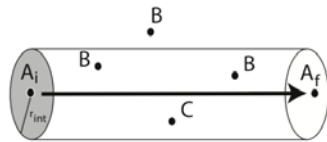
1. Realistic 3D Geometry



3. Stochastic Biochemical State Transition

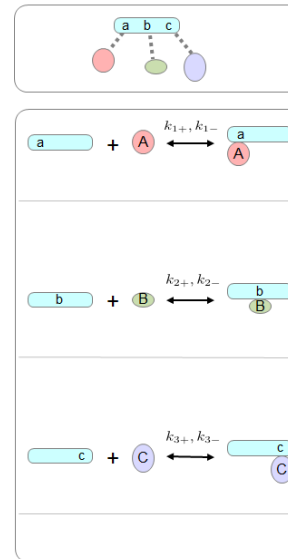


2. Random Walk Diffusion



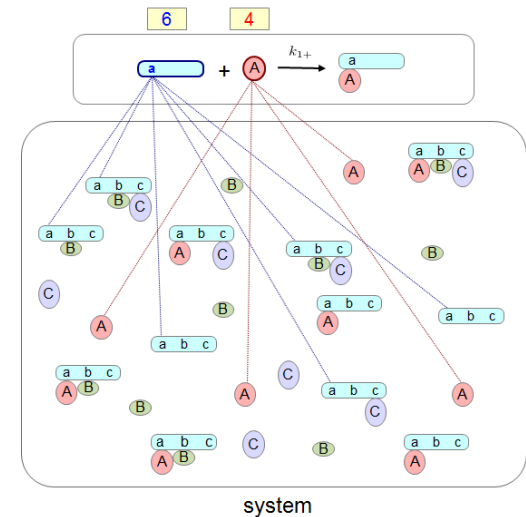
$A + B \rightarrow \text{products}$

contact map



Network-free Simulation

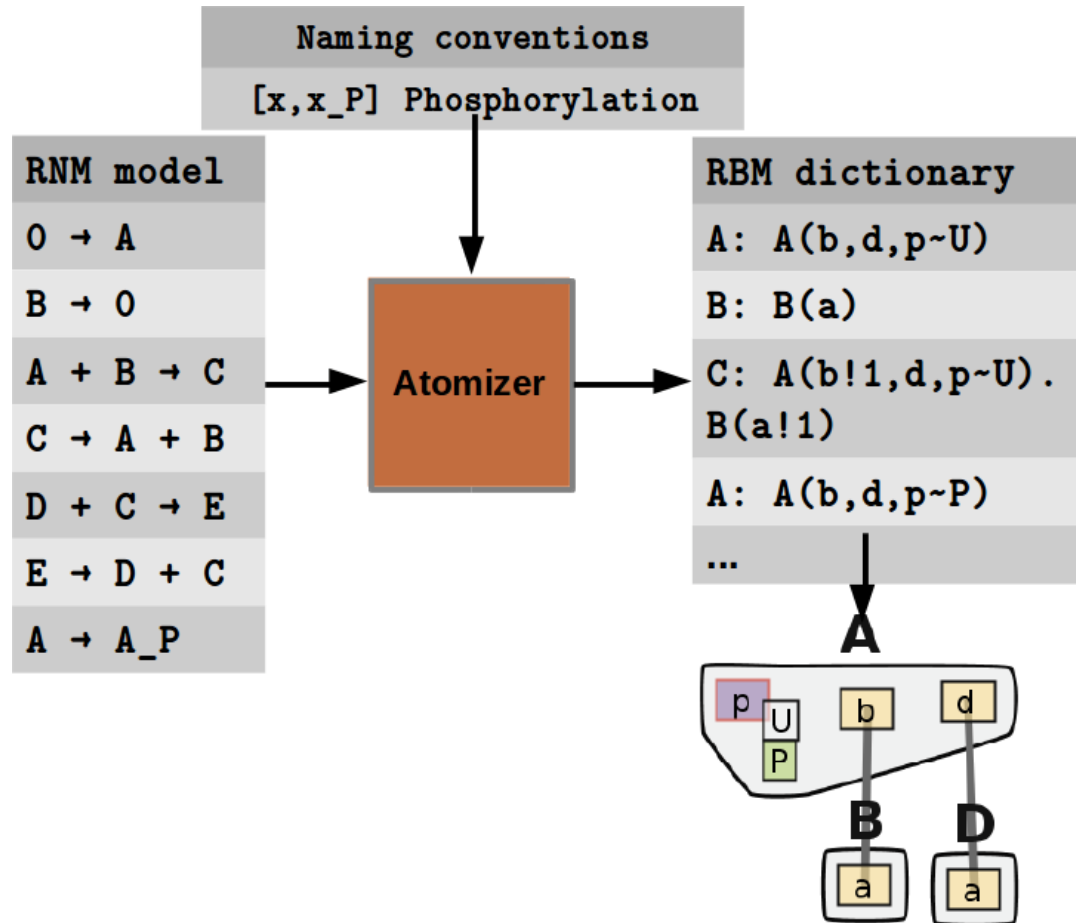
Find reactant sites and compute propensity



Particle-based

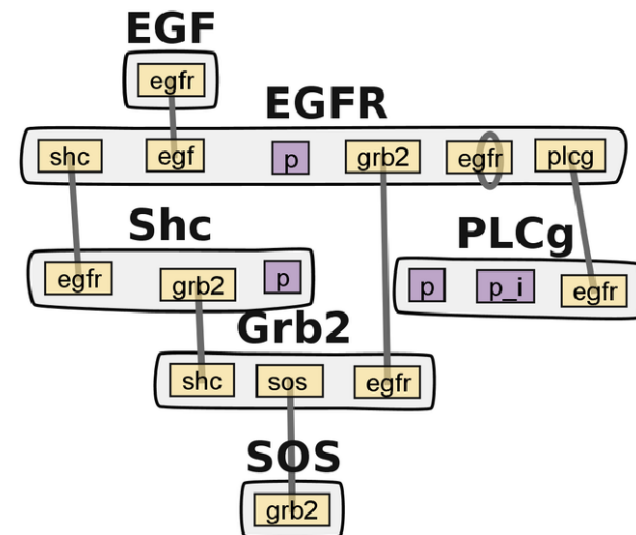
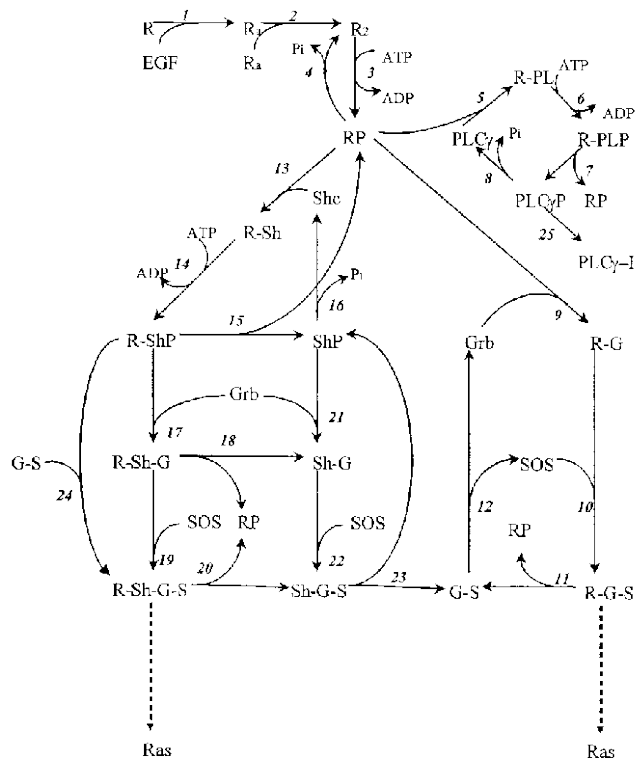
Atomizer can recover implicit molecular structure reaction networks

1. Reaction stoichiometry
2. Naming conventions
3. Annotation



Automating structure recovery

“Atomization”



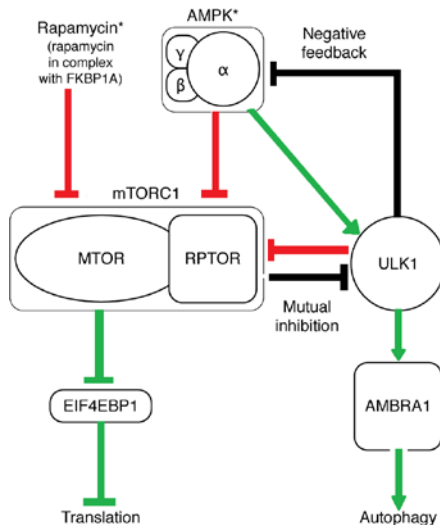
Motivation

- Find molecules and interactions used to build model
- Find out what has been modeled?
- Find the implicit assumptions
- Automate study of relaxing model assumptions

Visualizing rule-based models

John Sekar
Jose Juan Tapia

Autophagy network



Rule-based model

Rules are code!

```

R1: rapa(mtor)+MTOR(HEAT,FRB)<->rapa(mtor!1).MTOR(HEAT,FRB!1) a1,d1
R2: MTOR(HEAT,FRB)+RPTOR(WD40!1)<->MTOR(HEAT!1,FRB).RPTOR(WD40!1) a2,d2
R3a: RPTOR(RNC,ulk1,S792~O,S855_S859~O)+ULK1(straptor) ->RPTOR(RNC,ulk1!1,S792~O,S855_S859~O).ULK1(straptor!1) a3
R3b: RPTOR(ulk1!1).ULK1(straptor!1)->RPTOR(ulk1)+ULK1(straptor) d3
R3c: RPTOR(ulk1!1,S792~P,S855_S859~PP).ULK1(straptor!1)->RPTOR(ulk1,S792~P,S855_S859~PP)+ULK1(straptor) d3max
R4a: RPTOR(RNC,ulk1,S792~O,S855_S859~O)+EIF4EBP1(RCR,S65_T70~O)->RPTOR(RNC!1,ulk1,S792~O,S855_S859~O).EIF4EBP1(RCR!1)
R4b: RPTOR(RNC!1).EIF4EBP1(RCR!1)->RPTOR(RNC)+EIF4EBP1
R5a: AMPK(ulk1,T172~P)+ULK1(stampk,S758~O)->AMPK(ulk1,T172~P)+ULK1(stampk,S758~O)
R5b: AMPK(ulk1!1).ULK1(stampk!1)->AMPK(ulk1)+ULK1(stampk)
R6: MTOR(HEAT!1).RPTOR(WD40!1,ulk1!2).ULK1(straptor!1)->MTOR(HEAT!1).RPTOR(WD40!1,ulk1!2).ULK1(straptor!1)
R7: MTOR(HEAT!1).RPTOR(WD40!1,RNC!2).EIF4EBP1(RNC!2)->MTOR(HEAT!1).RPTOR(WD40!1,RNC!2).EIF4EBP1(RNC!2)
R8a: RPTOR(ulk1!1,S792~O).ULK1(straptor!1,S317~P,S855_S859~O)->RPTOR(ulk1!1,S792~O).ULK1(straptor!1,S317~P,S855_S859~O)
R8b: RPTOR(ulk1!1,S855_S859~P).ULK1(straptor!1,S317~P,S855_S859~O)->RPTOR(ulk1!1,S855_S859~P).ULK1(straptor!1,S317~P,S855_S859~O)
R9: ULK1(straptor,S317~P,S778~P)+AMBRA1(ST~O)->ULK1(straptor,S317~P,S778~P)+AMBRA1(ST~O)
R10: ULK1(straptor,S317~P,S778~P)+AMPK(ST~O)->ULK1(straptor,S317~P,S778~P)+AMPK(ST~O)
R11a: AMPK(ulk1!1,T172~P,ST~O).ULK1(stampk!1,S758~O)->AMPK(ulk1!1,T172~P,ST~O).ULK1(stampk!1,S758~O)
R12: AMPK(T172~P)+RPTOR(S792~O)->AMPK(T172~P)+RPTOR(S792~O)
R13a: ULK1(S758~P)->ULK1(S758~O) u0
R13b: EIF4EBP1(S65_T70~P)->EIF4EBP1(S65_T70~O) u0
R13c: RPTOR(S792~P)->RPTOR(S792~O) u1
R13d: RPTOR(S855_S859~P)->RPTOR(S855_S859~O) u1
R13e: RPTOR(S855_S859~PP)->RPTOR(S855_S859~O) u1
R13f: AMBRA1(ST~P)->AMBRA1(ST~O) u0
R13g: AMPK(ST~P)->AMPK(ST~O) u2
R13h: ULK1(straptor,S317~P)->ULK1(straptor,S317~O) u0
R13i: ULK1(S778~P)->ULK1(S778~O) u0
R13j: RPTOR(S792~P)->RPTOR(S792~O) u0
    
```

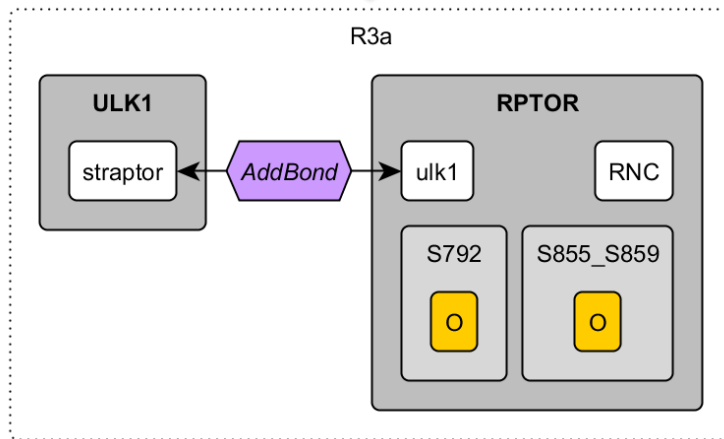
Why rule visualization is essential

1. Debugging
2. Model reuse / sharing
3. Communicating with collaborators
4. Model comparison
5. Model integration

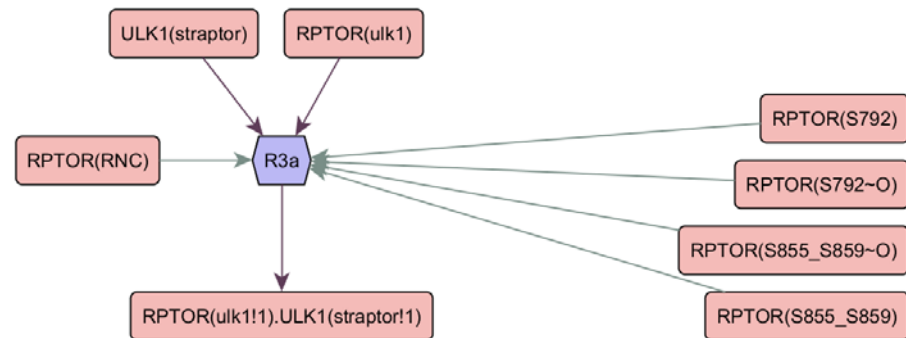
Szymanska et al. (2015), *PLOS One*

Automated rule visualization

R3a: RPTOR(RNC,ulk1,S792~O,S855_S859~O) + ULK1(straptor) -> \ RPTOR(RNC,ulk1!1,S792~O,S855_S859~O).ULK1(straptor!1) a3



Compact rule visualization



Regulatory graph

Model Calibration with Bayesian Parameter Estimation

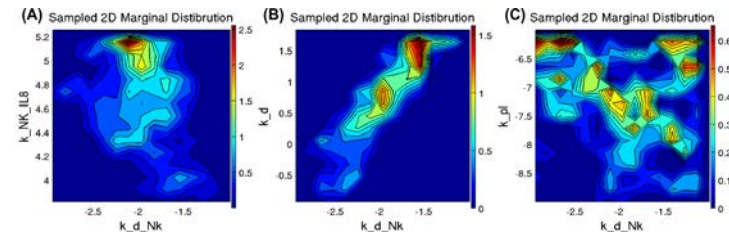
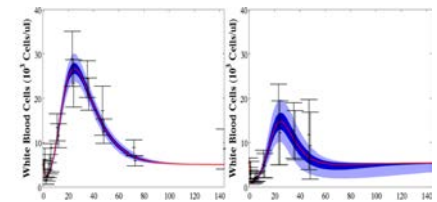
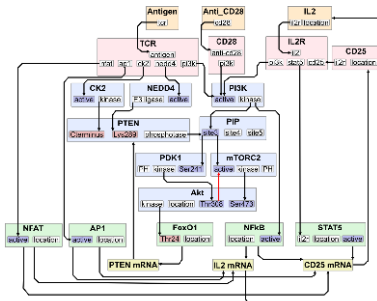
Experimental
Data

“Affinity maturation on
parameters to get best
fitting models.”

ptempest

Best fit parameter
ensembles

BioNetGen
Model



A quick RuleBender demo

- To download RuleBender go to <http://bionetgen.org/Download>
- To find files for this and many other examples, go <http://bionetgen.org/Tutorials>