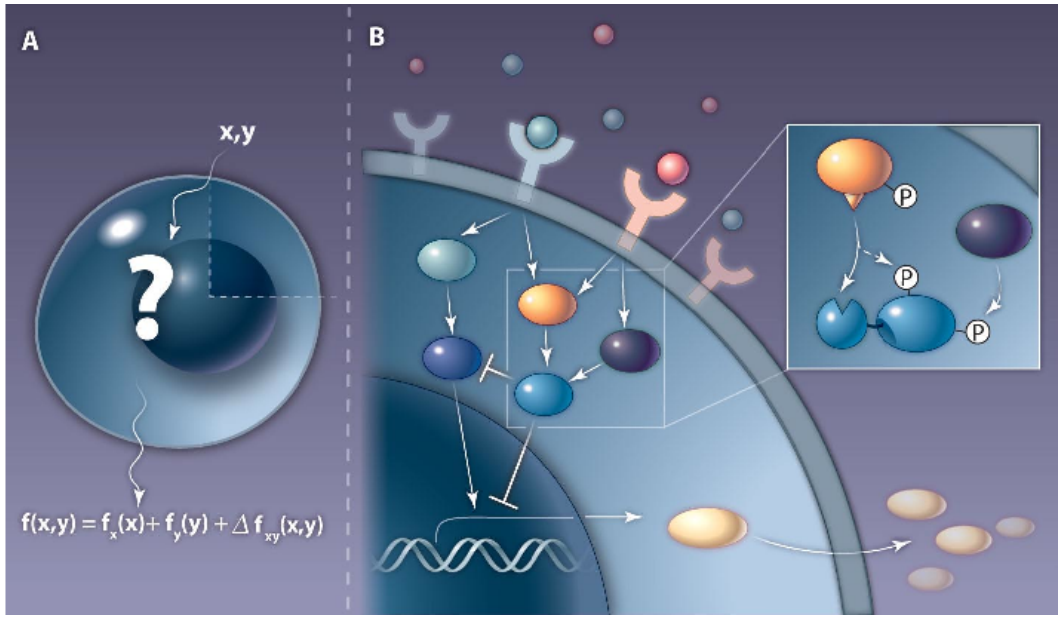
The atomizer: RNM2RBM

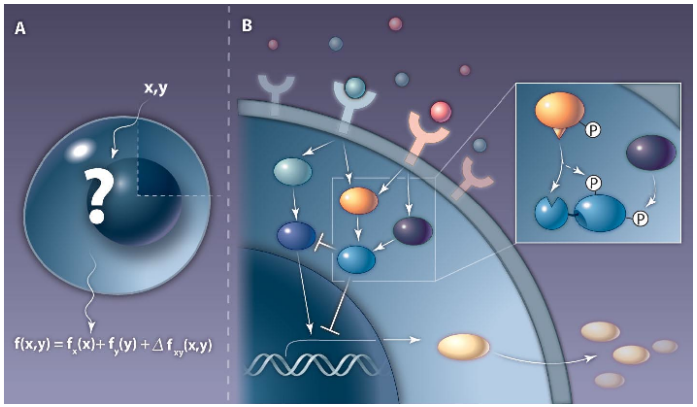
Jose Juan Tapia
John Sekar
James R. Faeder
University of Pittsburgh

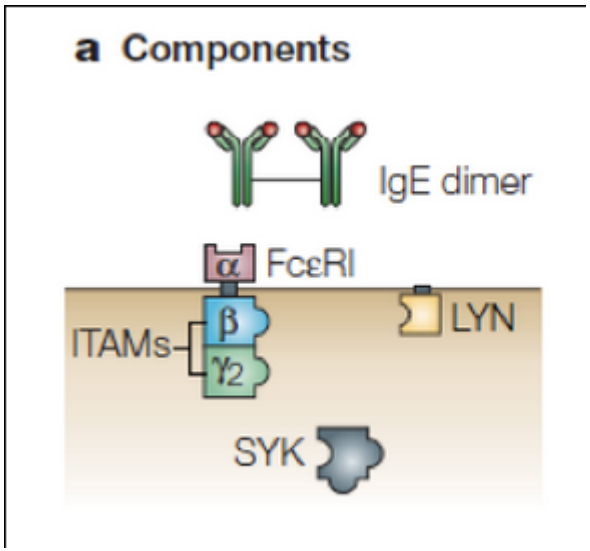
Rule-based models are …



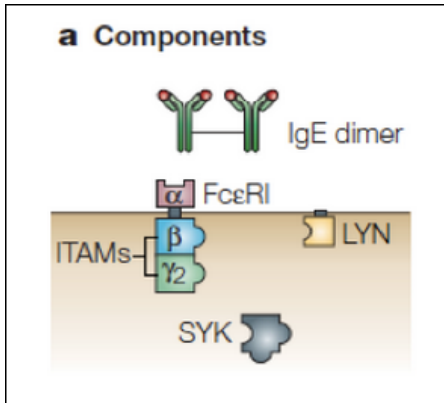
W. S. Hlavacek and J. R. Faeder, Sci. Signal. 2, pe46 (2009).

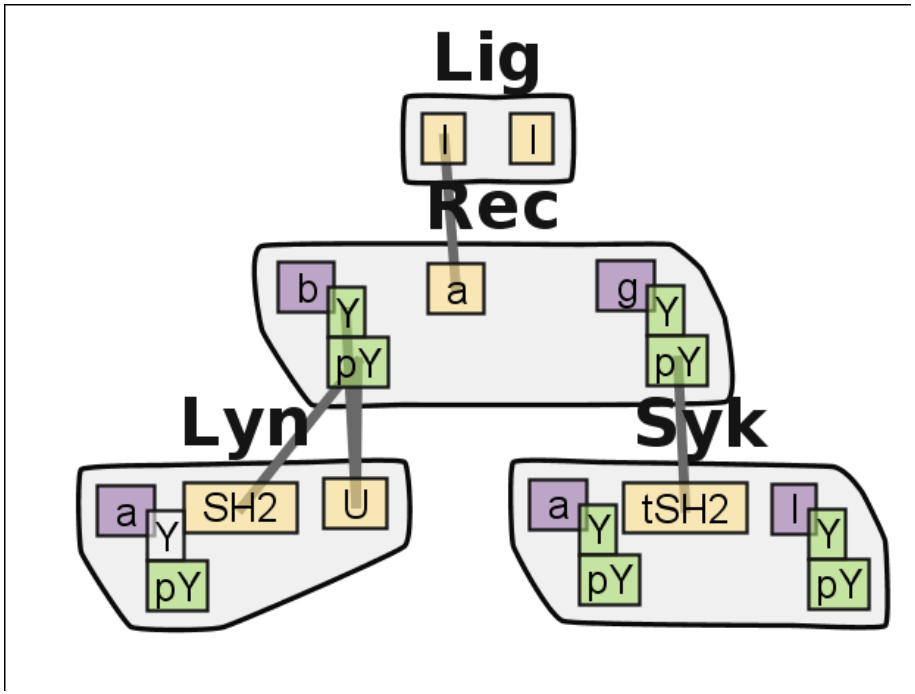
From a biological system to a mental model





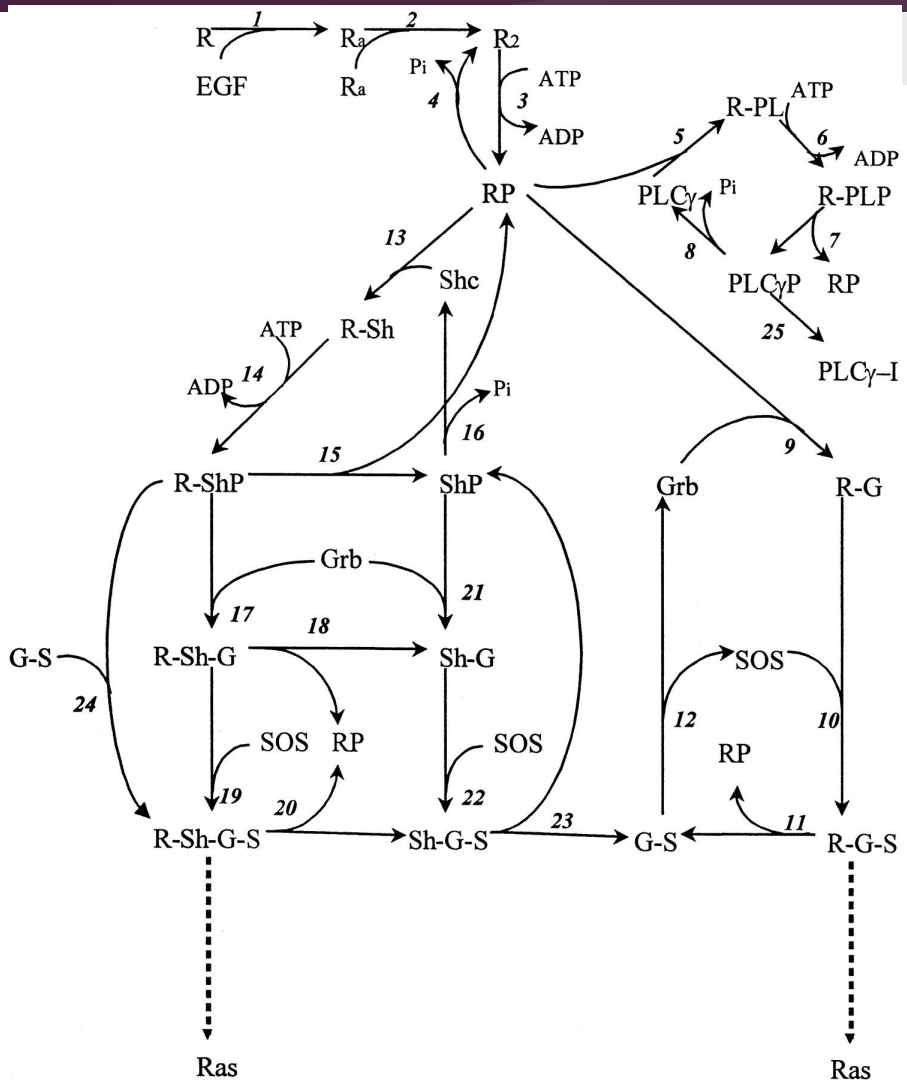
Mental model vs RBM





Contact map

Reaction Network model



Kholodenko EGFR model 1999

Syntax comparison

RNM

```
A + B -> A_B
```

RBM

```
A(b) + B(a) -> A(b!1).B(a!1)
```

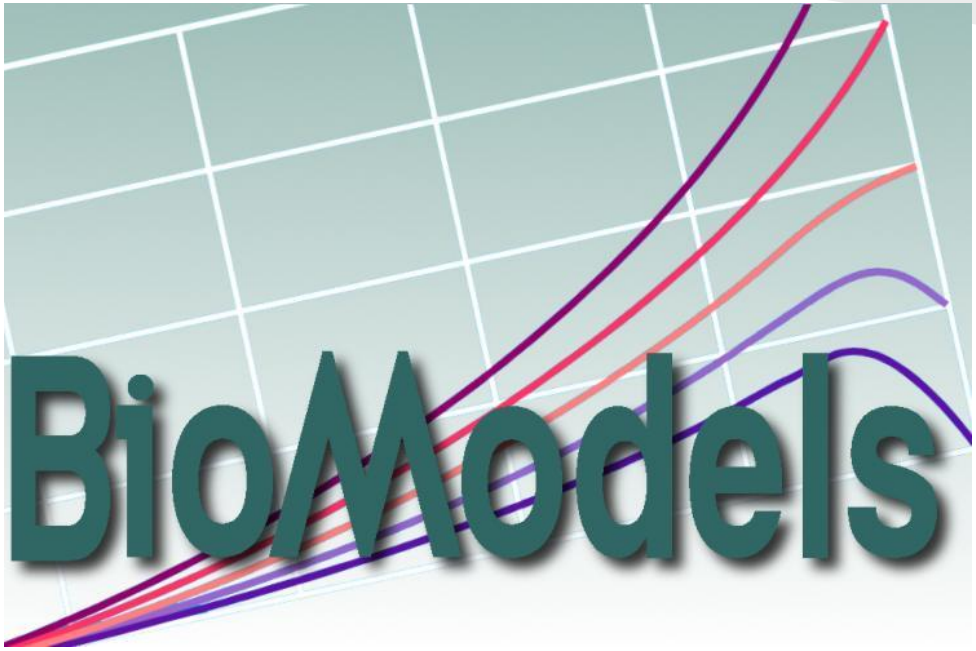
Rule-based modeling

- Rich syntax that allows the modeler to encode structural and contextual information
- Very scalable
- Syntax may be overkill for smaller models or phenomenological models.

Reaction Network Modeling

- Well understood theory of differential equations based chemical kinetics.
- Streamlined representation suitable for medium-sized models.
- Large body of models encoded using RNM
- Biological structural and contextual information is lost.

Large collection of RNM's



Our goals are

Find a way to recover structural and contextual information that is no longer explicit in RNM models.

Obtain an RBM representation of the same model with this information.

Understand and build upon years of RNM modeling knowledge from an RBM point of view in a semi-automated way.

Presenting…

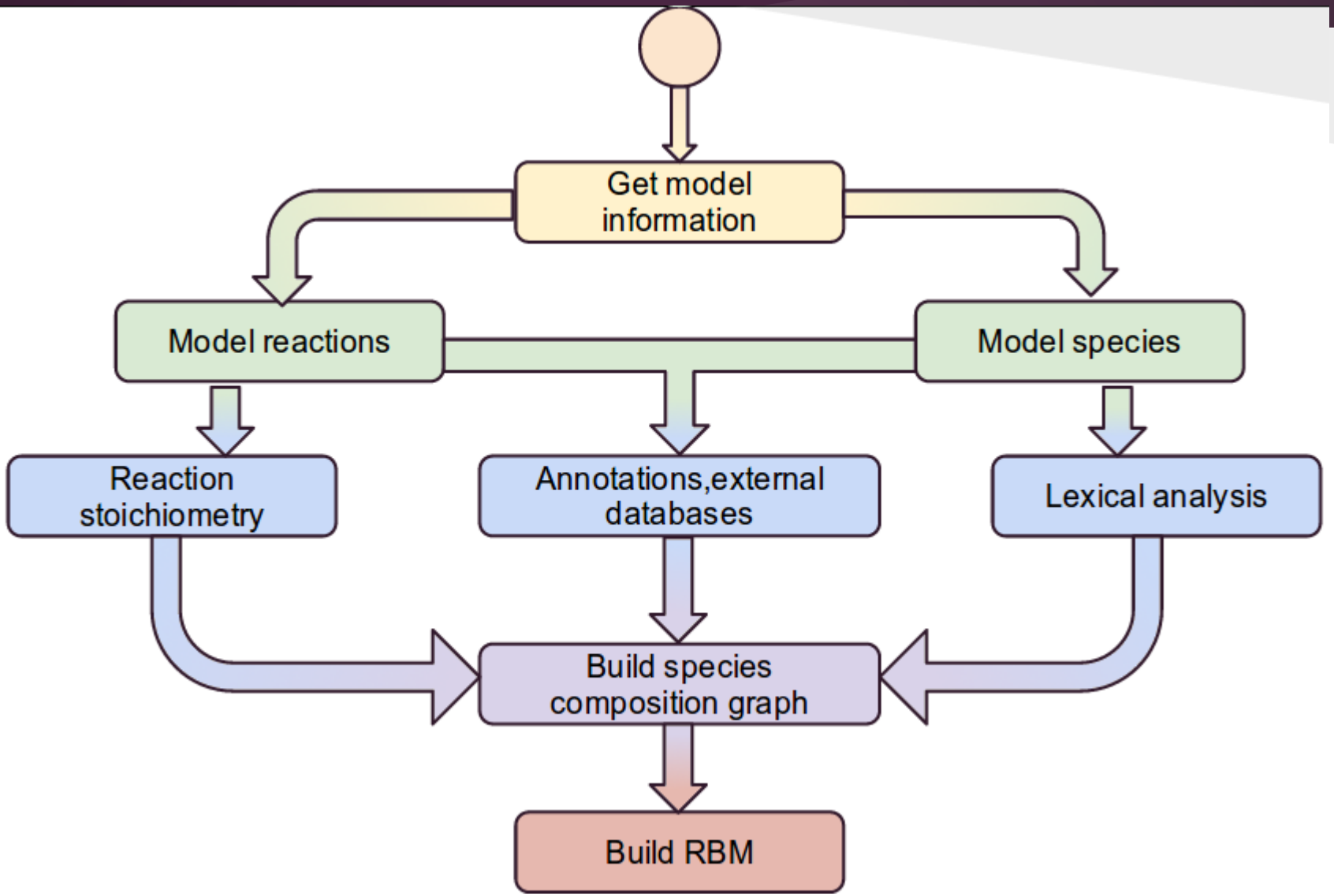
The

Atomizer

RBM

RNM

How does it work?



Toy example

```
A + B <-> A_B
B -> B_P
A + B_P -> A_P_B
```

Atomizing in a nutshell

- Reaction stoichiometry information

`A + B <-> A_B` *Complexation*

- Lexical analysis

`B -> B_P`

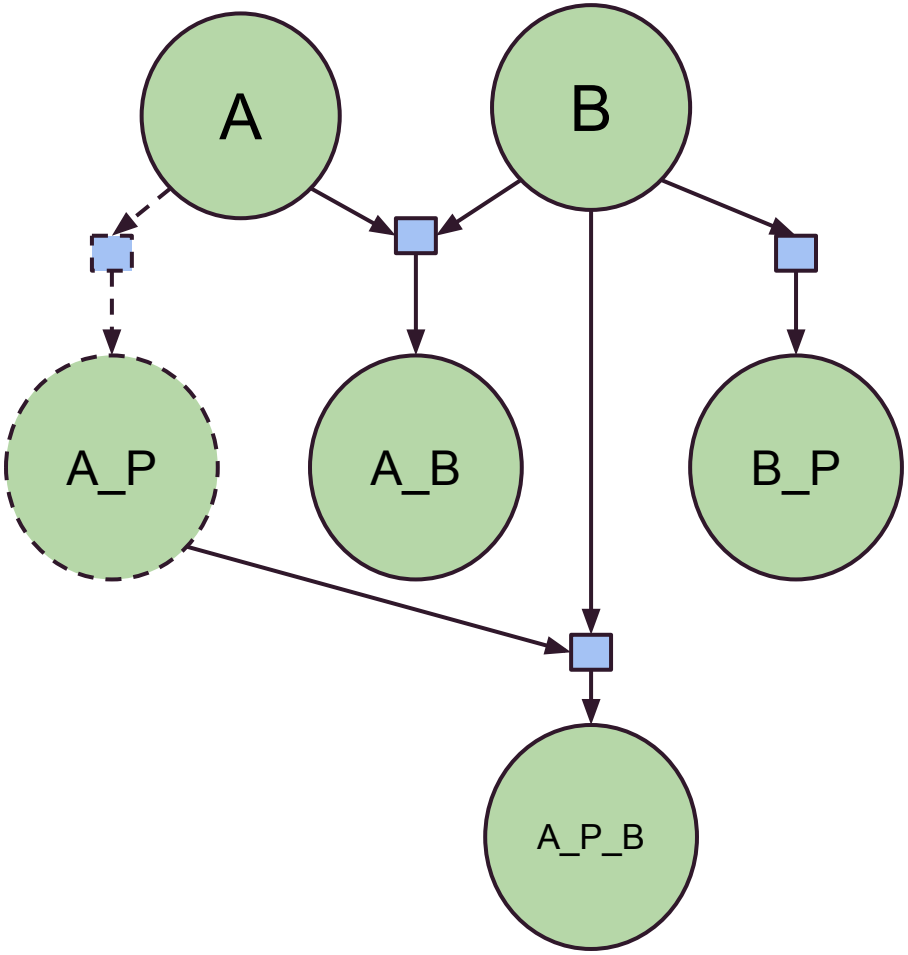
Phosphorylation

`A + B_P -> A_P_B`

Complexation and phosphorylation transfer

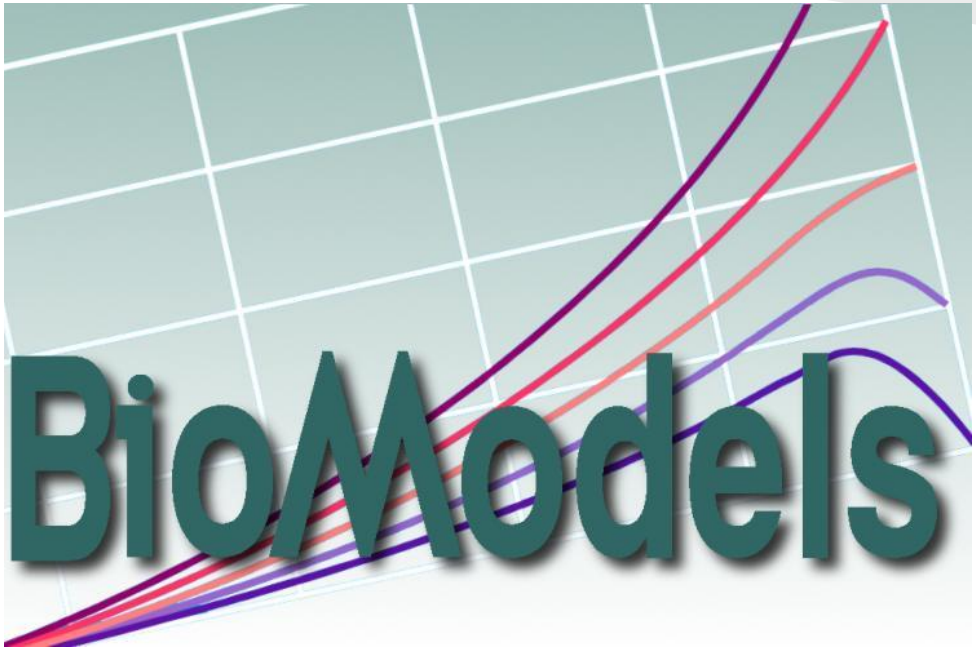
- Annotation information,interaction databases

Compositional Graph

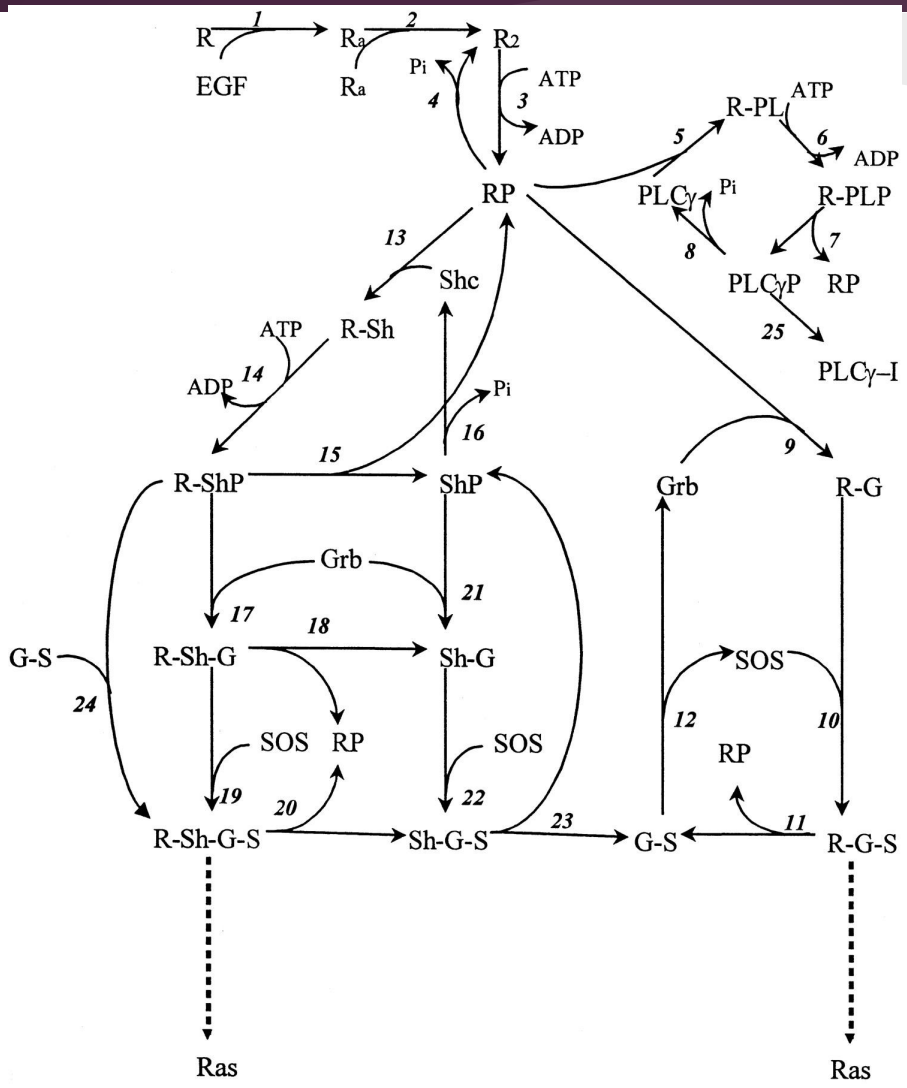



```
A(b,phospho~U~P)
B(a,phospho~U~P)
```

Atomize BioModels

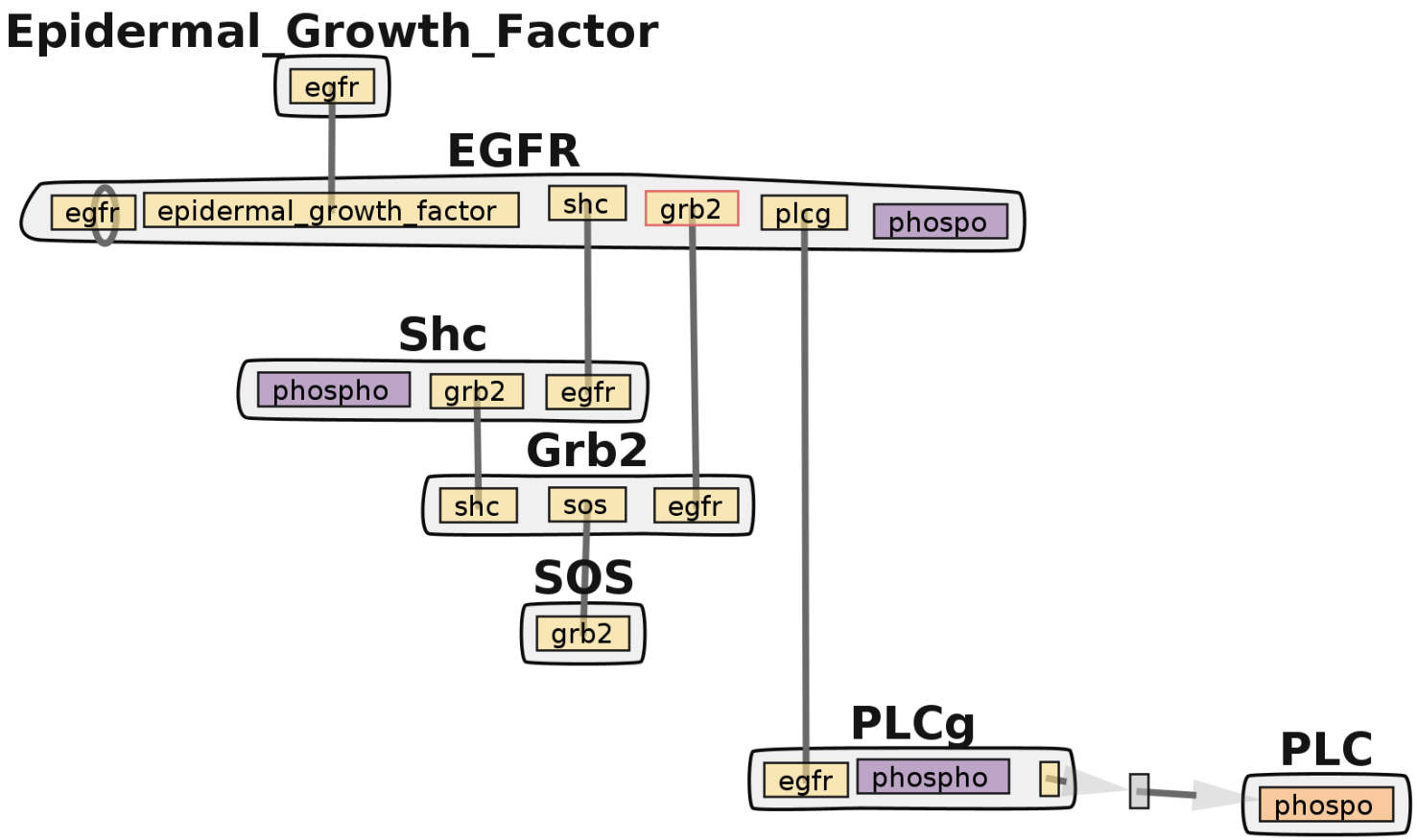


Epidermal Growth Factor

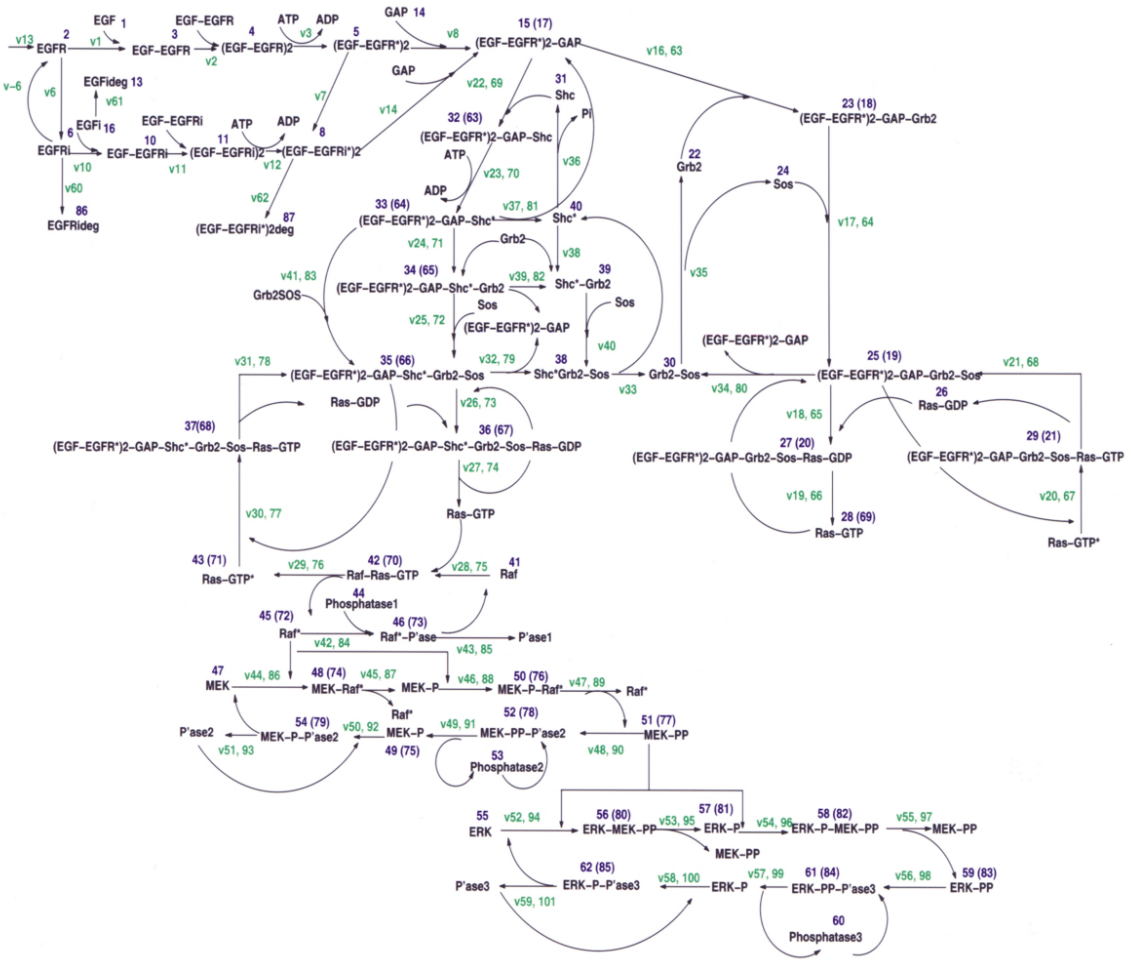


Kholodenko 1999

BioModels 48 atomized

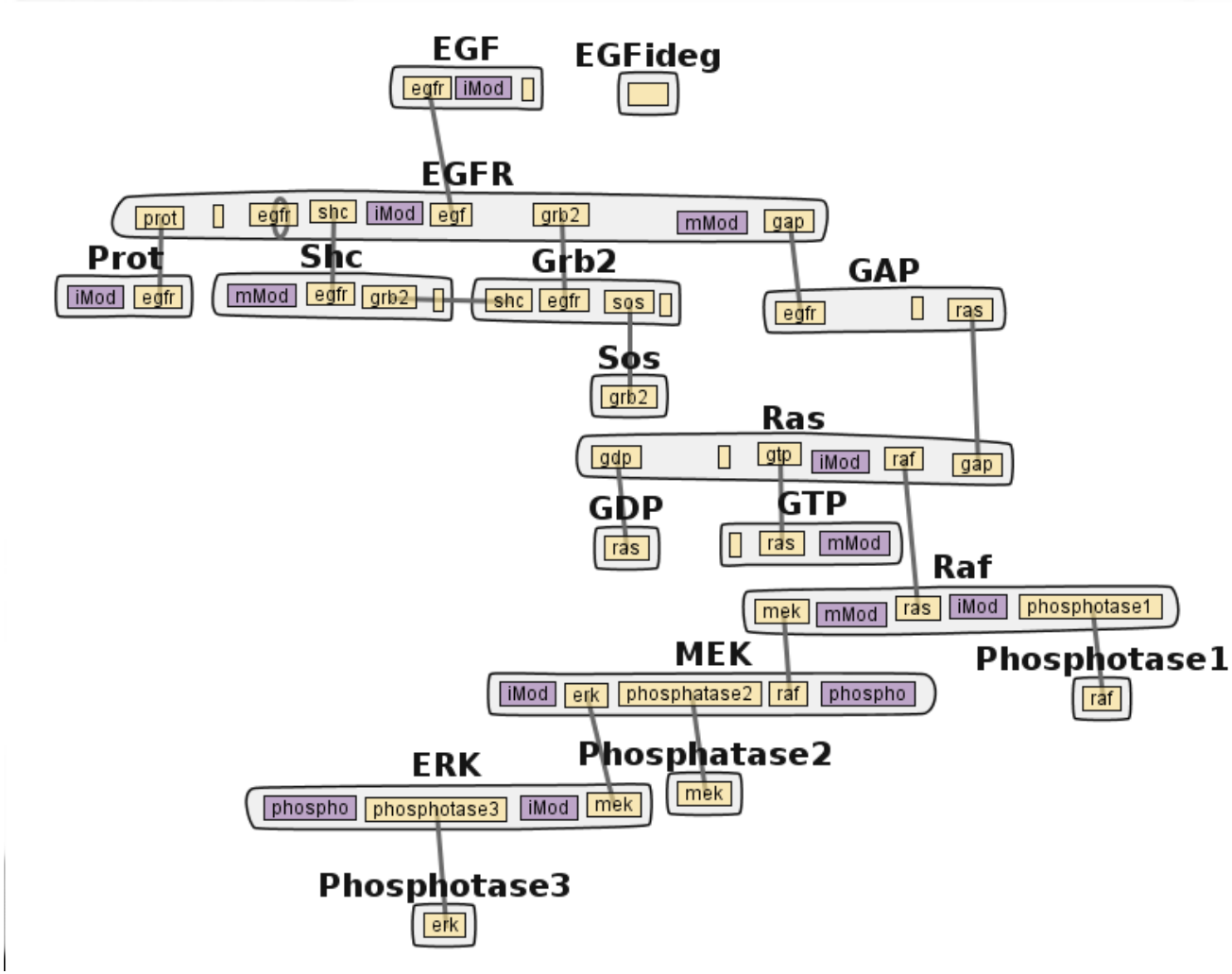


Bigger example

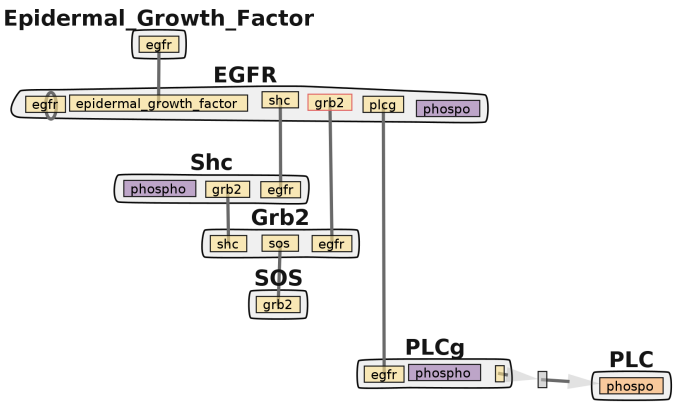


Schoeberl et. al.

BioModels 19 atomized



In depth:Dependency analysis



Molecule EGFR

- **Requires** component **_Pmod** to be in state **_P** for component **grb2** to bind
- **Requires** component **_Pmod** to be in state **_P** for component **plcg** to bind
- **Requires** component **_Pmod** to be in state **_P** for component **shc** to bind
- **Requires** component **egfr** to be bound for component **_Pmod** to acquire state **_P**
- **Requires** component **epidermal_growth_factor** to be bound for component **egfr** to bind
- Components **grb2,plcg,shc** **cannot be bound** at the same time

Molecule Grb2

- Components shc,egfr **cannot be bound** at the same time

Molecule PLCg

- **Requires** component genericMod to be **off** for component egfr to bind
- **Requires** component egfr to be off for component genericMod to be in state PLCgP_I

How do I compare these models?

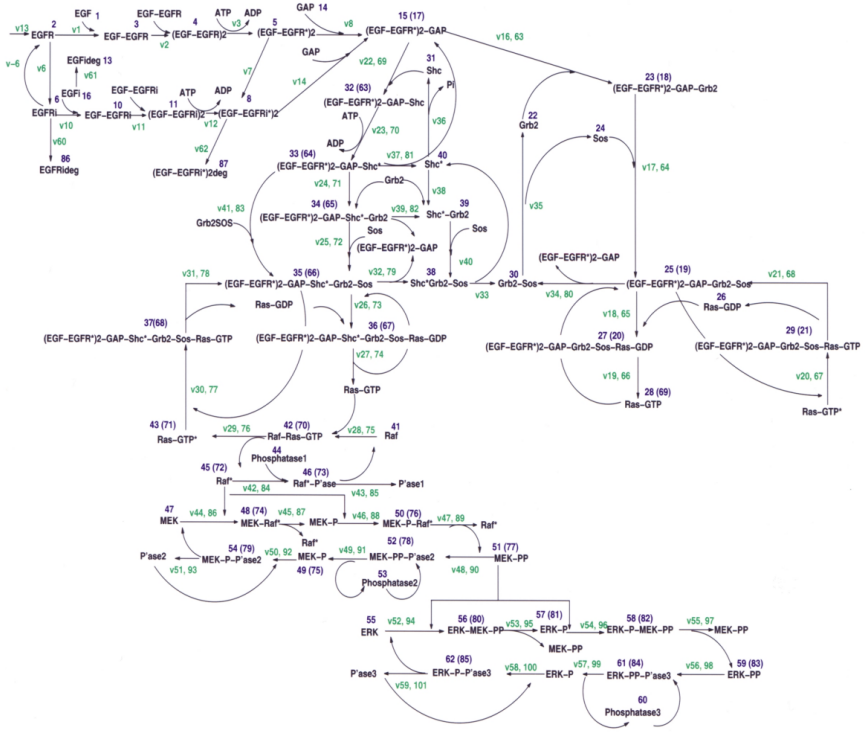


Image taken from Schoeberl et. al.

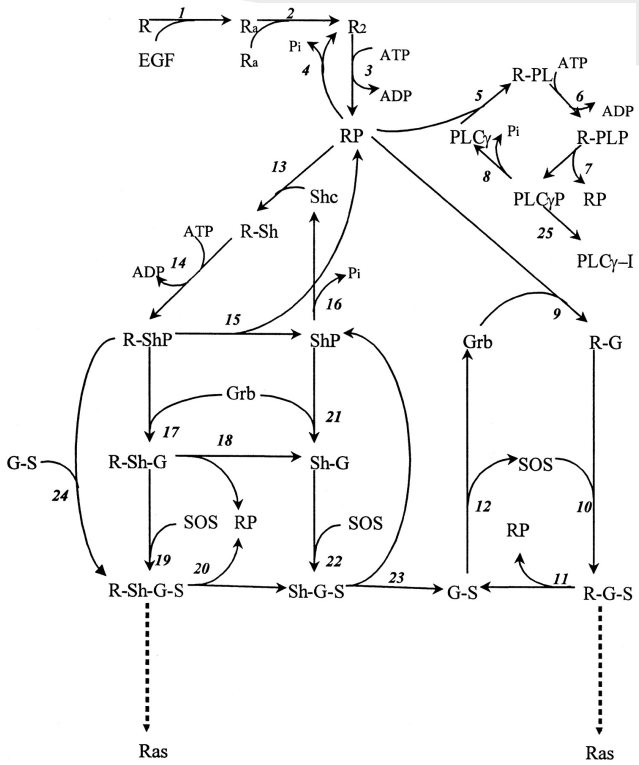
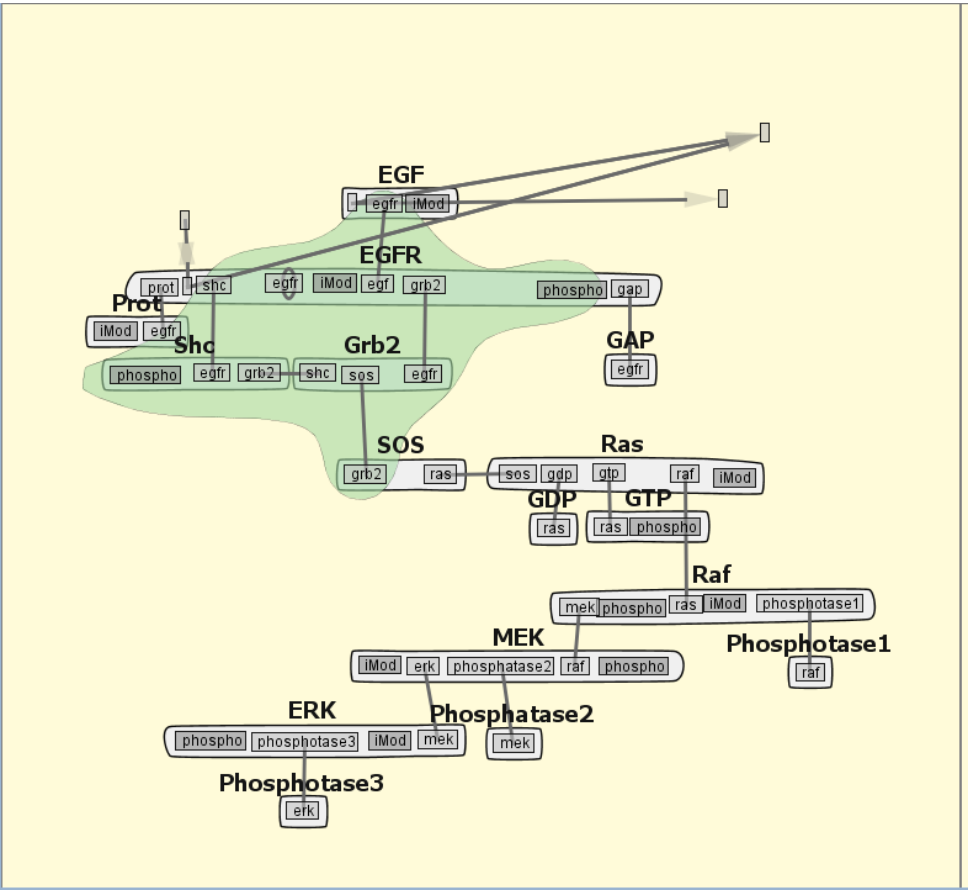
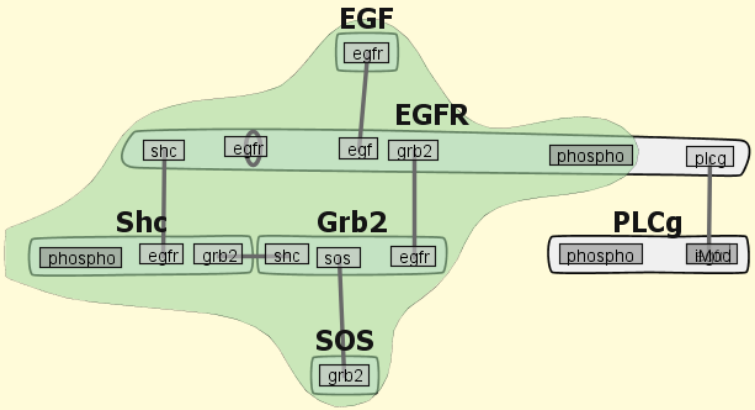


Image taken from Kholodenko et. al.

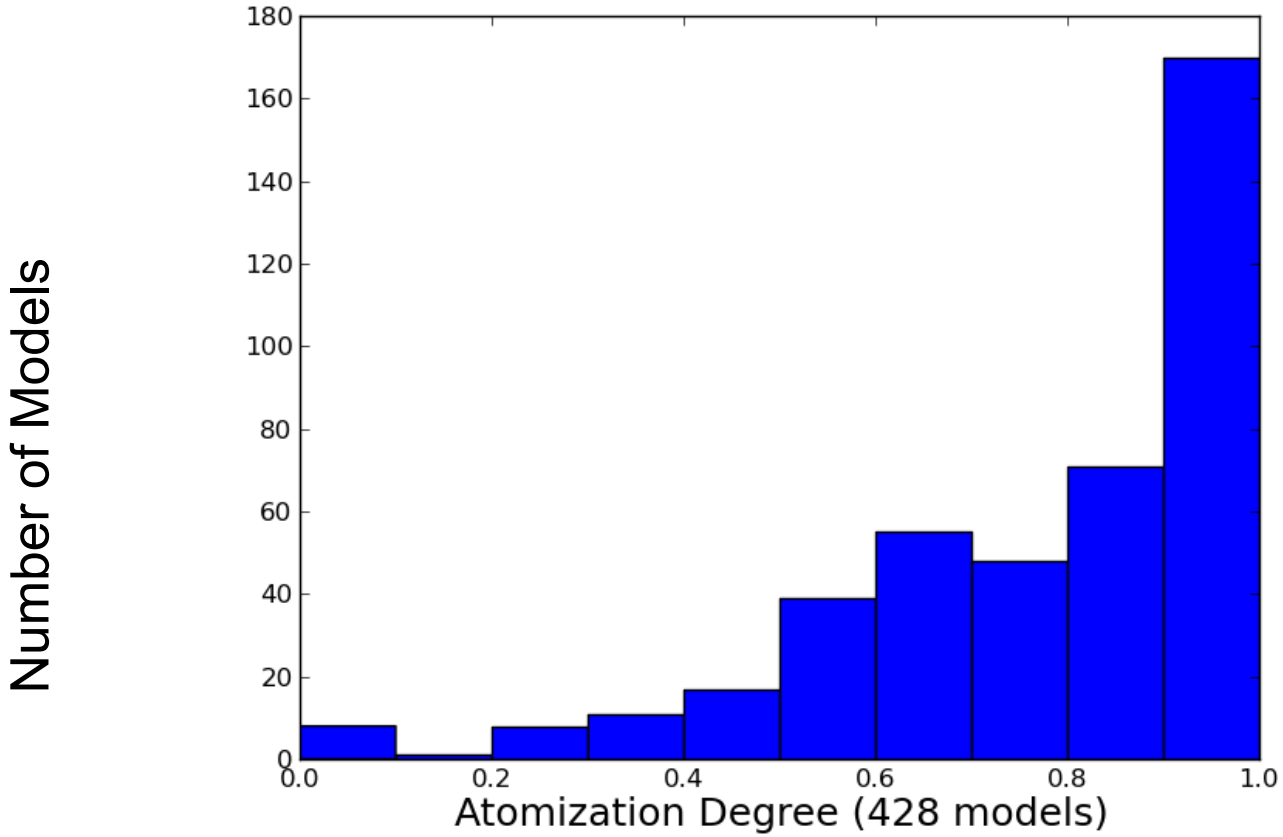
Atomized model alignment





MOSBIE: A Tool of Comparison and Analysis of Rule-Based Biochemical Models
John E. Wenskovitch Jr., Leonard A. Harris, Jose-Juan Tapia, James R. Faeder and G. Elisabeta

Atomization degree = structured / total SBML species

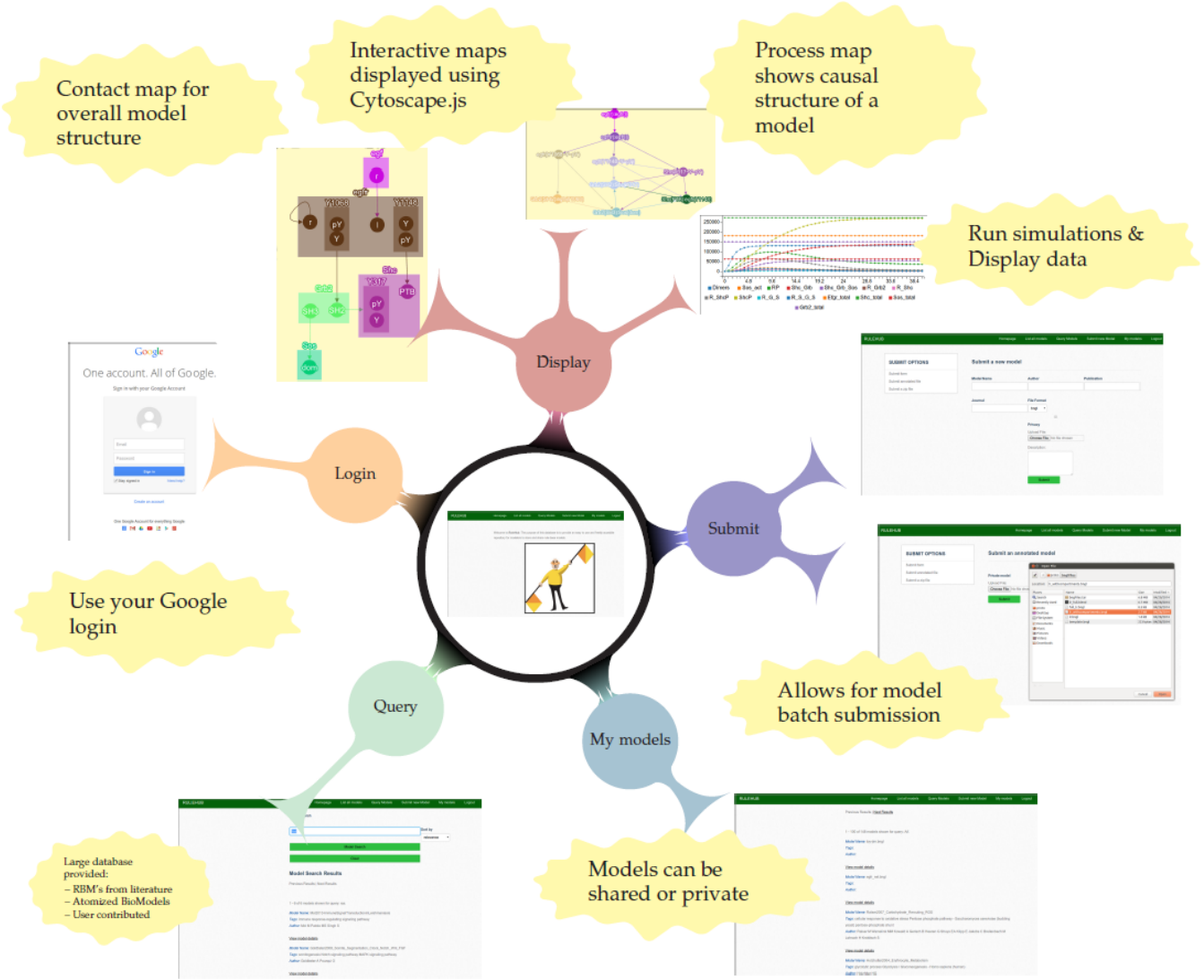


SBML Limitations

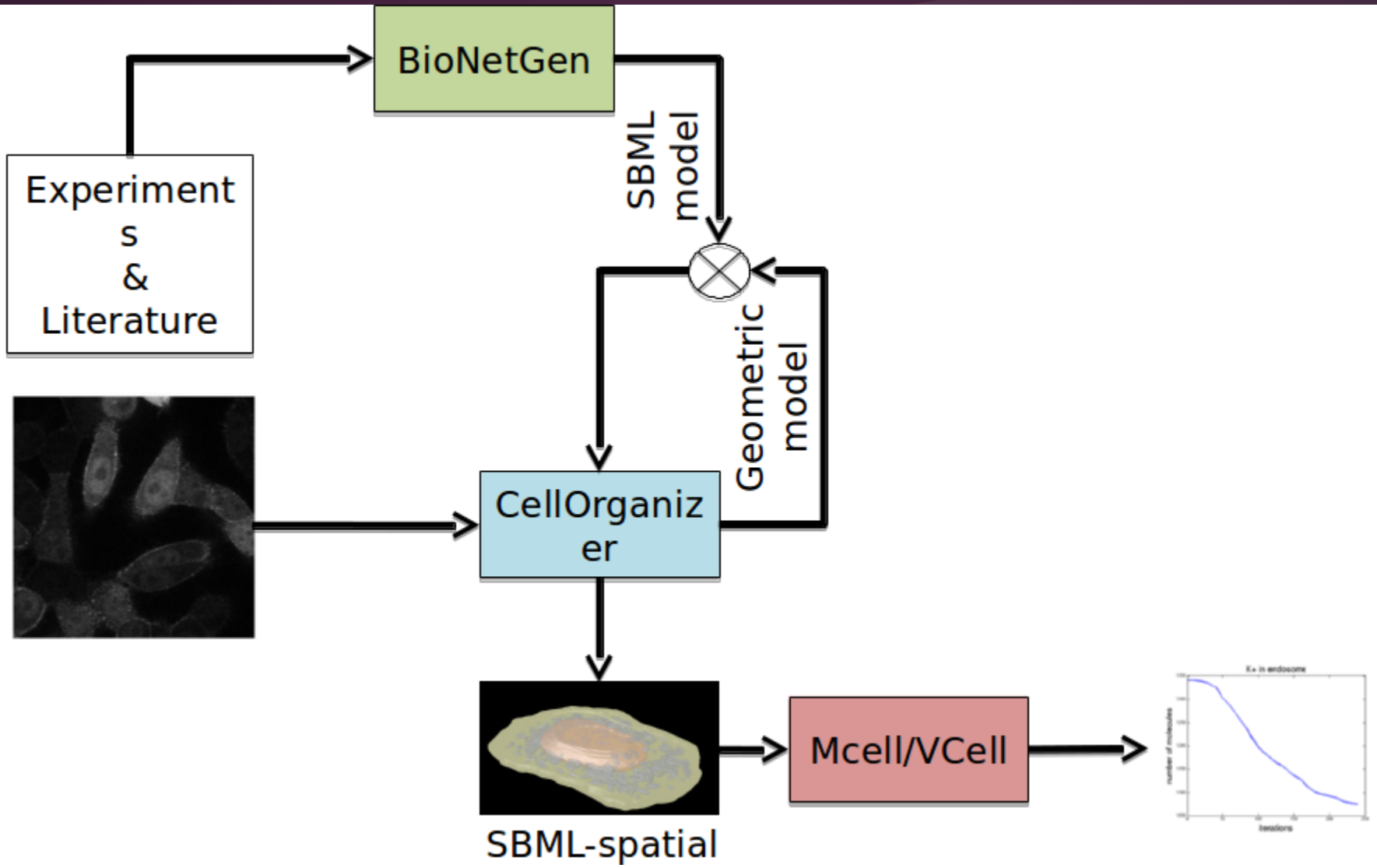
Atomization works best with ODE-based models.

Molecular structure, cooperativity analysis and such can still be obtained for models with non ODE based dynamics (events, rate rules, delays, etc) however their simulation can not always be obtained.

RuleHub


Contact map for overall model structure
Interactive maps displayed using Cytoscape.js
Process map shows causal structure of a model
Run simulations & Display data
Display
Login
Use your Google login
Submit
Allows for model batch submission
Query
My models
Large database provided:
– RBM's from literature
– Atomized BioModels
– User contributed
Models can be shared or private

Pipeline


BioNetGen
Experiments & Literature
SBML model
Geometric model
CellOrganizer
SBML-spatial
Mcell/VCell

Now you can atomize too!

http://ratomizer.appspot.com/translate

```
BNG2.pl <sbmlfile.xml>
```

```
readFile(<name.xml>)
```

Thank you

Faeder Lab:

- Dr. James Faeder
- Dr. Justin Hogg
- Dr. Leonard Harris
- John Sekar

MCell Team

- Jacob Czech
- Markus Dittrich
- Devin Sullivan

NIH grant P41 GM103712 and NSF Expeditions in Computing Grant (award 0926181)

Q&A

Extra slides

Naming convention analysis

Patterns	Associated process	Example
$[`+P', `+p', `+P\_', `+\_P']$	Phosporylation	x → xP
$[`+PP', `+\_PP']$	Double phosporylation	xP → xPP
$[`+i']$	Internalization	x− > xi
$[`-\_n', `+\_c']$	Compartment transfer from nucleus to cytoplasm	x_n, x_c
$[`+\_ubiq']$	Ubiquitination	x− > x_ubiq
$['K', 'KK']$	Kinase, Kinase kinase	MAP, MAPK, MAPKK
$[`+H']$	Adding a hydrogen-related modification	NAD, NADH
$[`+R']$	Receptor	EGF, EGFR
$[`+c']$	[Cyclic version, cytoplasm, casp3 substrate]	x, cx
$[`+2', `+3', `+4']$	[Dimer, Trimer, Tetramer]-[Protein family]	x → x2

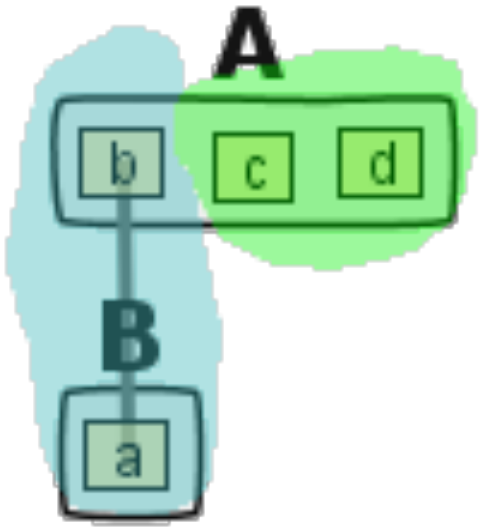
Naming convention analysis

+ P	27.4669509595
+ p	21.5778251599
- T+ D	9.0618336887
+ 2	7.4669509595
+ a	6.908315565
- D+ T	6.7356076759
- P+ M	5.5991471215
- n+ c	4.8614072495

Metric is the product of the number of times an annotation appears across the database multiplied by the percentage of models it appears in

Reaction center and context

A(b,c,d) + B(a) -> A(b!1,c,d).B(a!1)



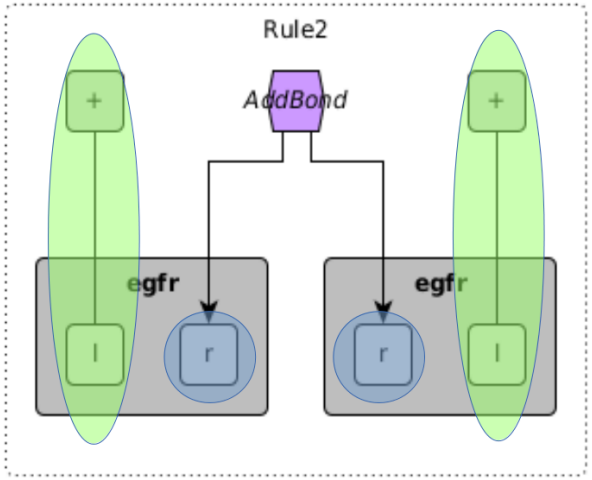
SBML Molecules contain minimal context information

```
A + B <-> C
```

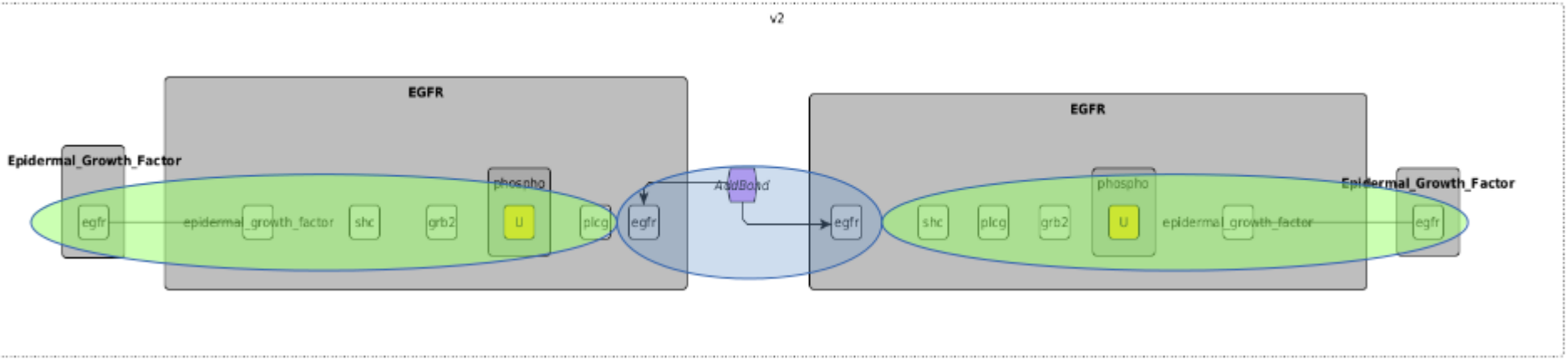
What is A? (Single molecule, umbrella name for a compound series of A molecules, etc).

Impossible to know without extensive annotation information. So we have to take them at face value.

Kholodenko's reaction context

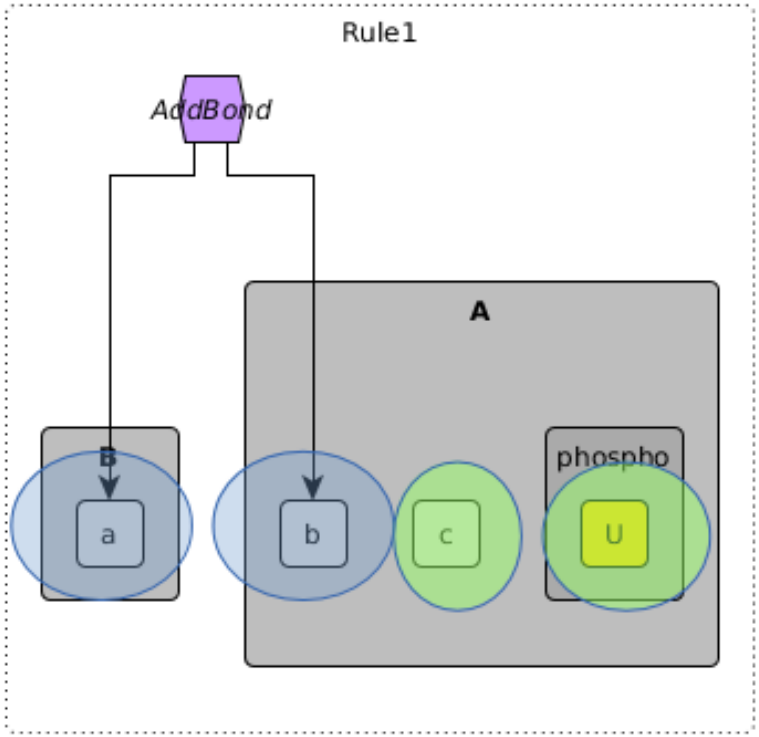


Manually constructed RBM version of Kholenko's model.

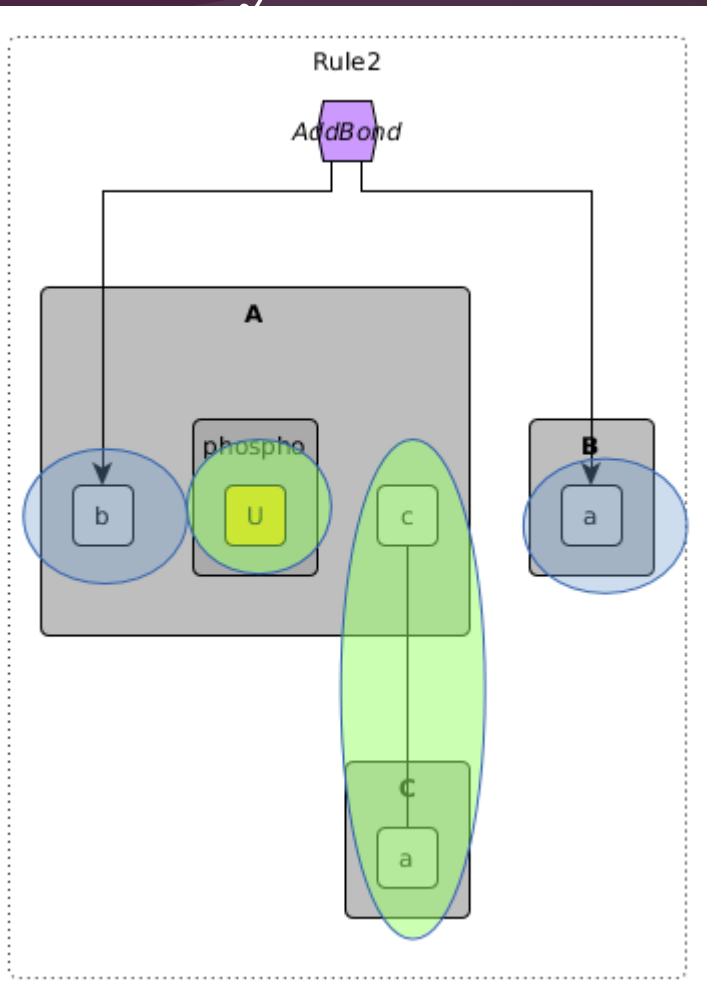


Automatically translated version of Kholodenko's model

Redundancy is the redundant way to redundantly go redundantly

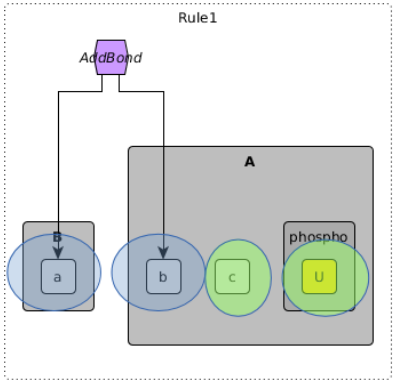


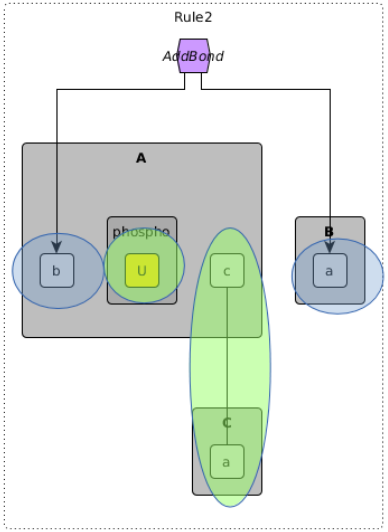
For ‘A’ and ‘B’ to bind all other states must be unoccupied.

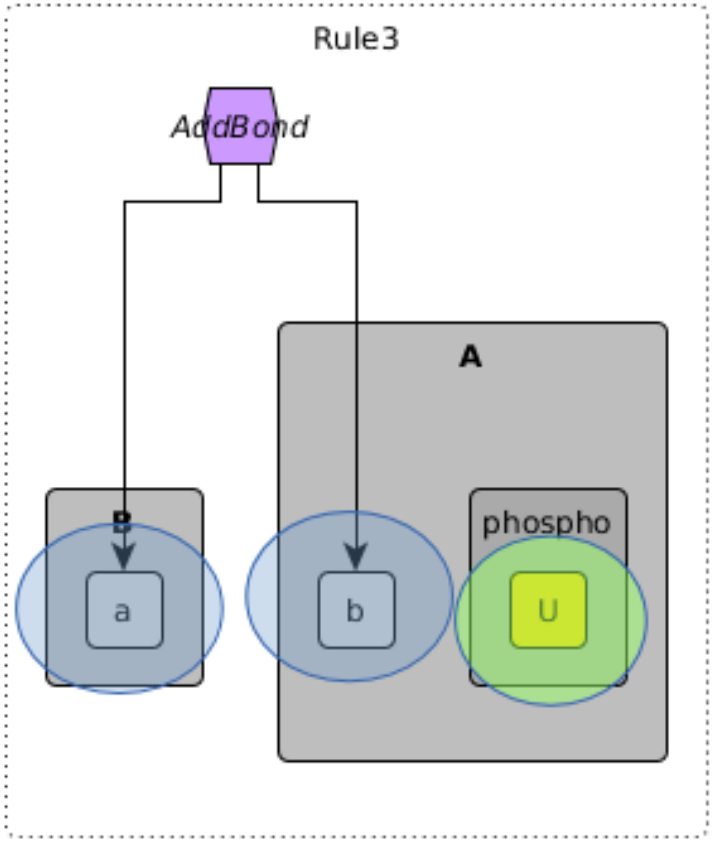


For ‘A’ and ‘B’ to bind ‘A’ must be bound to ‘C’ already

Context factorization







For A and B to bind it does not matter if C is part of the complex or not